(19)中华人民共和国国家知识产权局

(12)发明专利

(10)授权公告号 CN 103589692 B
(45)授权公告日 2018.07.24

(21)申请号 201310326869.0
(22)申请日 2002.11.12
(65)同一申请的已公布的文献号
申请公布号 CN 103589692 A
(43)申请公布日 2014.02.19
(30)优先权数据
60/350,607 2001.11.13 US
60/341,117 2001.12.17 US
60/377,066 2002.05.01 US
60/386,675 2002.06.05 US
(62)分案原申请数据
02827059.2 2002.11.12
(73)专利权人 宾夕法尼亚大学托管会
地址 美国宾夕法尼亚州
(72)发明人 G·高 J·M·威尔森
M·阿尔维拉
(74)专利代理机构 上海专利商标事务所有限公司 31100
代理人 余颖
(51)Int.Cl.
C12N 7/01(2006.01)
C07K 14/075(2006.01)
A61K 39/235(2006.01)
C12N 15/34(2006.01)
C12N 1/21(2006.01)
C12N 1/19(2006.01)
C12N 5/10(2006.01)
C12N 1/15(2006.01)
C12R 1/93(2006.01)
审查员 于娜

权利要求书1页 说明书47页
序列表147页 附图105页

(54)发明名称

检测和/或鉴定腺伴随病毒(AAV)序列以及分离所鉴定的新型序列的方法

(57)摘要

本发明提供了从组织或细胞来源样品的DNA中检测和分类AAV序列的方法。本发明还提供了用该方法鉴定的AAV序列以及用这些序列构建的载体。

CN 103589692 B

1.一种重组腺伴随病毒，包含至少一种功能性AAVrh.1衣壳，所述衣壳由vp1蛋白、vp2蛋白和vp3蛋白构成，其中，所述vp1衣壳蛋白由SEQ ID NO:100中第1至736位氨基酸构成，所述重组腺伴随病毒还包含3’和5’AAV反向末端重复序列和与调节序列操作性相连的转基因编码序列，所述调节序列指导所述转基因编码序列在宿主细胞中的表达，“AAV”表示腺伴随病毒。

2.如权利要求1所述的重组腺伴随病毒，其中，所述vp2蛋白由SEQ ID NO:100中第138至736位氨基酸构成。

3.如权利要求1所述的重组腺伴随病毒，其中，所述vp3蛋白由SEQ ID NO:100中第203至736位氨基酸构成。

4.如权利要求1-3中任一项所述的重组腺伴随病毒，其中，所述反向末端重复序列来自相对于AAV rh.1而言的异源腺伴随病毒。

5.一种组合物，包含权利要求1至4中任一项所述重组腺伴随病毒和生理相容性载体。

6.一种分离的腺伴随病毒衣壳蛋白，所述腺伴随病毒衣壳蛋白选自：由SEQ ID NO:100中第1至736位氨基酸构成的vp1衣壳蛋白；由SEQ ID NO:100中第138至736位氨基酸构成的vp2衣壳蛋白；和由SEQ ID NO:100中第203至736位氨基酸构成的vp3衣壳蛋白。

7.一种分离或合成的编码权利要求6所述蛋白的核酸分子。

8.权利要求7所述的核酸分子，所述核酸分子还包含编码功能性AAV rep蛋白的腺伴随病毒序列。

9.一种产生重组腺伴随病毒的方法，包括培养含有以下物质的宿主细胞的步骤，所述宿主细胞含有(a)权利要求7所述的核酸分子，(b)功能性rep基因和(c)AAV反向末端重复序列和转基因，以及(d)E1a基因产物、E1b基因产物、E2a基因产物和/或E4ORF6基因产物。

10.一种宿主细胞，转染了权利要求7或8所述的核酸分子，所述转染体外进行。

11.一种组合物，包含生理相容性载体和至少一种重组腺伴随病毒，所述重组腺伴随病毒包含功能性AAVrh.1衣壳，所述衣壳由vp1蛋白、vp2蛋白和vp3蛋白构成，其中，所述vp3衣壳蛋具有SEQ ID NO:100中第203至736位氨基酸构成的氨基酸序列，所述重组腺伴随病毒还包含3’和5’AAV反向末端重复序列和与调节序列操作性相连的转基因编码序列，所述调节序列指导所述转基因编码序列在宿主细胞中的表达。

12.一种分离或合成的核酸分子，它编码(a)氨基酸序列如SEQ ID NO:100中第1至736位氨基酸所示的AAVrh.1vp1衣壳蛋白，(b)功能性rep基因，(c)腺伴随病毒反向末端重复序列和转基因和(d)将腺伴随病毒反向末端重复序列和转基因包装到AAV衣壳蛋白内所需足够的辅助功能。

检测和/或鉴定腺伴随病毒(AAV)序列以及分离所鉴定的新型序列的方法

[0001] 本申请是PCT国际申请中国国家阶段02827059.2号发明专利申请的分案申请。原PCT国际申请的申请号为PCT/US02/33629,国际申请日为2002年11月12日。

技术领域

[0002] 本发明提供了从组织或细胞来源样品的DNA中检测和分类AAV序列的方法。本发明还提供了用该方法鉴定的AAV序列以及用这些序列构建的载体。

背景技术

[0003] 腺伴随病毒(AVV)是细小病毒家族的一个成员,是一个小的无包膜的二十面体,其基因组为4.7kb-6kb的单链线性DNA分子。AVV开始时被归为依赖病毒,因为这个病毒是作为纯化的腺病毒株中的污染物而被发现的。AVV的生命周期包括潜伏期和感染期,在潜伏期内,AVV基因组在感染后以位点特异性的方式整合到宿主染色体内,在感染期内,伴随着腺病毒或单纯疱疹病毒的感染,整合的病毒基因组随后恢复、复制并包装成具有感染性的病毒颗粒。AVV因其具有非致病性、广泛的宿主范围包括非分裂细胞以及位点特异性整合等特征而成为一个极具吸引力的基因转移工具。

[0004] 最近的研究表明,AVV载体是一个优选的基因治疗工具。到目前为止,已经有6个不同血清型的AVV被从人或非人灵长类动物(NHP)体内分离出来并被鉴定。其中,人血清2型首先被开发为基因转移工具,现已被广泛用于在不同靶组织和动物模型中进行基因转移试验。利用AVV2来源的载体治疗某些人类疾病的临床试验也在进行之中,其中包括囊性纤维化和血友病B。我们所想要的是用于基因转移的AAV来源的载体。

发明内容

[0005] 本发明的一个方面涉及根据在缺少辅助病毒共感染的情况下AVV的潜伏和整合特性,利用生物信息学分析、PCR扩增以及克隆技术从各种人源及非人灵长类动物(NHP)组织的细胞DNA中检测和鉴定AVV序列的新方法。

[0006] 本发明的另外一个方面涉及利用本发明的上述方法所检测到的AVV序列的分离方法。本发明还包括利用这些新的AVV血清型制备载体的方法,仅仅根据衣壳的基因序列和rep/cap基因连接处的结构就可以用这些载体进行血清学和基因转移的研究。

[0007] 本发明的另外一个方面涉及利用本发明所描述的试剂,如通用引物对/探针和定量实时PCR进行血清学、流行病学、生物分布及传播模式的研究。

[0008] 本发明的另外一个方面涉及利用RACE和其他分子技术从不同来源的细胞DNA中分离新AVV血清型的完整且有感染性的基因组的方法。

[0009] 本发明还涉及利用不同来源的辅助腺病毒从人或NHP细胞系中恢复新的AVV血清型基因组的方法。

[0010] 本发明的另一个方面涉及新的AVV血清型、含该血清型的载体及利用该血清型的

方法。

[0011] 从下面对发明的详细描述中可以很容易地理解本发明的这些方面及其他方面。

附图说明

[0012] 图1A到1AAAR描述的是至少编码AAV cap蛋白的核酸序列的比对。这些图列出了包含ITR、rep区和衣壳区的如下新型AAV血清型的全长序列：新型AAV血清型7[SEQ ID NO:1]，以及以前报道的AAV1[SEQ IN NO:6]，AAV2[SEQ ID NO:7]和AAV3[SEQ ID NO:8]。新型AAV血清型AAV8[SEQ ID NO:4]和AAV9[SEQ ID NO:5]是该共同提交的申请的主题。这个比对中所涉及的本发明的其他新型克隆包括42-2[SEQ ID NO:9]，42-8[SEQ ID NO:27]，42-15[SEQ ID NO:28]，42-5b[SEQ ID NO:29]，42-1b[SEQ ID NO:30]；42-13[SEQ ID NO:31]，42-3a[SEQ ID NO:32]，42-4[SEQ ID NO:33]，42-5a[SEQ ID NO:34]，42-10[SEQ ID NO:35]，42-3b[SEQ ID NO:36]，42-11[SEQ ID NO:37]，42-6b[SEQ ID NO:38]，43-1[SEQ ID NO:39]，43-5[SEQ ID NO:40]，43-12[SEQ ID NO:41]，43-20[SEQ ID NO:42]，43-21[SEQ ID NO:43]，43-23[SEQ ID NO:44]，43-25[SEQ ID NO:45]，44.1[SEQ ID NO:47]，44.5[SEQ ID NO:47]，223.10[SEQ ID NO:48]，223.2[SEQ ID NO:49]，223.4[SEQ ID NO:50]，223.5[SEQ ID NO:51]，223.6[SEQ ID NO:52]，223.7[SEQ ID NO:53]，A3.4[SEQ ID NO:54]，A3.5[SEQ ID NO:55]，A3.7[SEQ ID NO:56]，A3.3[SEQ ID NO:57]，42.12[SEQ ID NO:58]，44.2[SEQ ID NO:59]。AAV10[SEQ ID NO:117]，AAV11[SEQ ID NO:118]和AAV12[SEQ ID NO:119]的特征区域序列也列在图中。AAV基因组结构中的关键标记也在图中标识出来。断裂的空间用黑点表示。AAV1[SEQ ID NO:6]的3'ITR用与以前报道的序列同样的结构表示。TRS代表末端进化位点。要注意的是AAV7是所报道的唯一利用GTG作为VP3起始密码子的AAV。

[0013] 图2A到2F显示的是以前所报道的AAV血清型1[SEQ ID NO:64]，AAV2[SEQ ID NO:70]，AAV3[SEQ ID NO:71]，AAV4[SEQ ID NO:63]，AAV5[SEQ ID NO:114]，和AAV6[SEQ ID NO:65]以及本发明的新型AAV序列的vp1衣壳蛋白的氨基酸序列的比对，本发明的新型AAV序列包括:C1[SEQ ID NO:60]，C2[SEQ ID NO:61]，C5[SEQ ID NO:62]，A3-3[SEQ ID NO:66]，A3-7[SEQ ID NO:67]，A3-4[SEQ ID NO:68]，A3-5[SEQ ID NO:69]，3.3b[SEQ ID NO:62]，223.4[SEQ ID NO:73]，223-5[SEQ ID NO:74]，223-10[SEQ ID NO:75]，223-2[SEQ ID NO:76]，223-7[SEQ ID NO:77]，223-6[SEQ ID NO:78]，44-1[SEQ ID NO:79]，44-5[SEQ ID NO:80]，44-2[SEQ ID NO:81]，42-15[SEQ ID NO:84]，42-8[SEQ ID NO:85]，42-13[SEQ ID NO:86]，42-3A[SEQ ID NO:87]，42-4[SEQ ID NO:88]，42-5A[SEQ ID NO:89]，42-1B[SEQ ID NO:90]，42-5B[SEQ ID NO:91]，43-1[SEQ ID NO:92]，43-12[SEQ ID NO:93]，43-5[SEQ ID NO:94]，43-21[SEQ ID NO:96]，43-25[SEQ ID NO:97]，43-20[SEQ ID NO:99]，24.1[SEQ ID NO:101]，42.2[SEQ ID NO:102]，7.2[SEQ ID NO:103]，27.3[SEQ ID NO:104]，16.3[SEQ ID NO:105]，42.10[SEQ ID NO:106]，42-3B[SEQ ID NO:107]，42-11[SEQ ID NO:108]，F1[SEQ ID NO:109]，F5[SEQ ID NO:110]，F3[SEQ ID NO:111]，42-6B[SEQ ID NO:112]，42-12[SEQ ID NO:113]。新血清型AAV8[SEQ ID NO:95]和AAV9[SEQ ID NO:100]是此共同提交的专利申请的主题。

[0014] 图3A到3C显示的是AAV7rep蛋白[SEQ ID NO:3]的氨基酸序列。

具体实施方式

[0015] 在本发明中，发明者发现了一种方法可以利用腺伴随病毒(AVV)在缺少辅助病毒共感染的情况下具有穿过细胞核整合到细胞DNA并达到潜伏感染的能力。该方法利用聚合酶链式反应(PCR)为基础的技术可以从人和非人灵长类动物来源的或其他来源的组织DNA内检测、鉴定和/或分离AVV序列。另外，该方法还适于检测、鉴定和/或分离如下所描述的其他整合的病毒及非病毒序列。

[0016] 本发明还涉及利用本发明的方法所鉴定的核酸序列。其中一种腺伴随病毒是一个新的血清型，本文中被称为血清型7(AVV7)。本文所涉及的其他新型腺伴随病毒血清型包括AAV10、AVV11和AVV12。其他利用本发明的方法所鉴定的新型AVV血清型也包括在本说明书中。参见附图和所列的序列，本文已纳入作为参考。

[0017] 本发明还涉及这些AAV序列的片段。其中特别有意义的AAV片段是cap蛋白，包括vp1、vp2、vp3和超变区，rep蛋白，其中包括rep78、rep68、rep52和rep40，以及编码这些蛋白的序列。这些片段都可被用于各种载体系统和宿主细胞中。这种片段可以单独使用，也可以和其他AAV序列或片段、或者其他AAV或非AAV病毒序列的元件联合使用。在一个特别优选的实施方式中，载体含有本发明的AAV cap和/或rep序列。

[0018] 如本文所描述，核酸序列是用任何一种公用的或商业的多序列比对程序如“Clustal W”来比对的，这些程序可从互联网的服务器上下载。另外，还使用了Vector NTI程序。还可以使用本领域所熟知的许多算法来测算核苷酸序列的同一性，其中包括那些上面所描述的程序中的序列。另外，多聚核苷酸序列还可用Fasta进行比较，这是GCG6.1版本中的一个程序。Fasta能将所查询的序列与所检索到的序列之间最佳重叠区域进行比对并计算出其序列同一性的百分率。例如，两个核酸序列之间同一性的百分率可以利用Fasta软件，参考GCG6.1版本中所提供的默认参数(字号为6，评分矩阵采用NOPAM因子)来确定，本文已纳入作为参考。氨基酸序列也可用相似的程序来处理，如“Clustal X”程序。一般来说，这些程序的设置都是默认的，但是如果有必要本领域的技术人员也可以改变这些设置。另外，本领域的技术人员也可以使用其他的算法或计算机程序，只要这些程序能如参考算法和程序一样测算序列的同一性或比对核酸序列。

[0019] 术语“基本同源”或“基本一致”用于核酸或核酸片段时是指在一个核酸序列中适当插入或删除一些核苷酸后与另一个核酸序列比较，其核苷酸序列同一性至少占到所比较序列的95％到99％。当称两个序列同源时必须比较其全长序列、其开放阅读框架或至少含15个核苷酸的适宜核酸片段。适宜核酸片段的例子在本文中有描述。

[0020] 术语“基本同源”或“基本一致”用于氨基酸或其片段时是指在一个氨基酸片段中适当插入或删除一些氨基酸后与另一个氨基酸片段比较，其氨基酸序列同一性至少占到所比较序列的95％到99％。当称两个序列同源时必须比较其全长序列、其蛋白质如cap蛋白、rep蛋白，或至少含8个氨基酸的适宜片段，较好的是含15个氨基酸的片段。适宜片段的例子在本文中有描述。

[0021] 术语“高度保守的”是指至少80％的序列同一性，优选的为至少90％的序列同一性，更优选的为超过97％的序列同一性。本领域的技术人员可以很容易地通过本领域所熟知的算法和计算机程序来确定序列的同一性。

[0022] 术语“序列同一性百分率”或“一致的”用于核酸序列时是指当将两个序列一起比对时其核苷酸残基是一样的。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，或至少约含500-5000个核苷酸的片段。但是，也可以比较小片段之间的同一性，如约含9个核苷酸的片段，通常含至少约20-24个核苷酸、至少约28-32个核苷酸、至少约36个或更多个核苷酸。同样，“序列同一性百分率”也可以用于确定氨基酸序列，确定全长蛋白质或其片段的同一性。适宜的片段至少约含8个氨基酸，也可以最多达700个氨基酸。适宜片段的例子在本文中也有描述。

[0023] AAV序列及其片段可用于制备rAAV，也可作为反义转移载体、基因治疗载体或疫苗载体。本发明还涉及核酸分子、基因转移载体以及含本发明的AAV序列的宿主细胞。

[0024] 如本文所描述，含本发明所述的AAV衣壳蛋白的本发明的载体特别适用于利用中和抗体消除AAV血清型来源载体以及其他病毒载体的效应。本发明的rAAV载体特别适合rAAV再注射和重复的基因治疗。

[0025] 本发明的这些和其他实施方式以及优点下面有更详细的描述。如本说明书和权利要求书中所描述的，术语“含有”和“包括”以及其他类似的词语是指还可以包含其他的部分、元件、整体、步骤等。相反的，术语“由…组成”及其类似的词语是指不包含其他的部分、元件、整体、步骤等。

[0026] I.本发明的方法

[0027] A.通过分子克隆技术检测序列

[0028] 本发明的一个方面涉及检测和/或鉴定样品中靶核酸序列的方法。该方法特别适于检测整合到细胞染色体上的病毒序列，如腺伴随病毒(AAV)和逆转录病毒等。本说明书以AAV为参照，以它为例子进行说明。但是，根据本说明书，本领域的技术人员也可以很容易地利用逆转录病毒(如猫白血病病毒((FeLV)，HTLVI和HTLVII)和慢病毒(如人免疫缺陷病毒(HIV)、猿免疫缺陷病毒(SIV)、猫免疫缺陷病毒(FIV)、马感染性贫血病毒、以及泡沫病毒)等来实施本发明。另外，本发明的方法还可用于检测整合到宿主细胞染色体内或未整合到宿主细胞染色体内的其他病毒或非病毒序列。

[0029] 本文所描述的样品可以是任何含核酸的样品，如组织、组织培养物、细胞、细胞培养物、以及生物液体，其中包括而不限于尿液和血液。这些核酸序列可以是质粒来源的DNA或RNA、任何来源的天然DNA或RNA，其中包括细菌、酵母、病毒、以及较高等的生物如植物或动物。DNA或RNA的提取可以使用本领域技术人员所熟知的任何方法，如Sambrook编写的“分子克隆:实验室手册”(Molecular Cloning:A Laboratory Manual)(纽约:冷泉港实验室)所描述的方法。样品的来源以及提取核酸的方法并不是对本发明所作出的限制。另外，本发明的方法可以直接用于含DNA的样品或从该样品中得到(或提取)的核酸。

[0030] 本发明的方法包括利用聚合酶链式反应(PCR)扩增含DNA的样品，所用的第一组引物是双链核酸序列第一区特异性的，因此可以得到扩增序列。

[0031] 如本文所描述，每个区域都是基于至少两个血清型(即AAV)或病毒株(即慢病毒)的核酸序列的比对而确定的，其中所述每个区所含有的序列5’末端高度保守，中间部分优选可变的但不是必须的，3’末端也是高度保守的，这些序列保守或可变是相对于至少两个比对的AAV血清型的序列而言的。5’和/或3’高度保守的核苷酸至少约有9个，较好的是至少有18个碱基对。对于可变区来说，不需要有保守的序列，这些序列可以是相对保守的，即比

对的血清型或病毒株之间具有小于90％、80％或70％的同一性。

[0032] 每个区可以跨越约100bp到10kb碱基对的长度，但是特别优选的是其中一个被称为特征区域的，即该区具有足够的独特性，能用来鉴定靶组织来源的扩增序列。例如，在一个实施方式中，第一区是一个约250bp的片段，与任何已知的AAV序列都能区别开，该序列是AAV来源的经正性鉴定的扩增区。而且，该区内的各种序列都具有足够的独特性，可用于鉴定扩增序列所来源的血清型。一旦被扩增(及被检测)后，序列可用常规的限制性酶切技术鉴定，通过与AAV1、AAV2、AAV3、AAV4、AAV5或AAV6或AAV7、AAV10、AAV11、AAV12或其他任何经本发明鉴定的新血清型内该区的限制性酶切图谱比较就可以鉴定该序列。上述血清型都是本发明所鉴定和涉及的。

[0033] 根据本文所描述的方法，本领域的技术人员可以很容易地鉴定其他整合病毒上的这种区，并且易于检测和鉴定这些序列。因此，一对理想的通用引物应被设计在高度保守的末端，这组引物可用于扩增样品中所选择的区。本发明的这一方面可用于设计检测靶序列(如AAV)是否存在以及鉴定AAV血清型的诊断试剂盒，所用的判断标准可以包括本文所描述的或利用本文所描述的技术分离出的血清型的限制性酶切图谱。例如，PCR产物的快速鉴定或分子血清型的确定可以通过酶切PCR产物并与限制性酶切图谱比较来完成。

[0034] 因此，在一个实施方式中，AAV的“特征区域”可能跨越AAV1[SEQ ID NO:6]的约2800到3200区域，以及AAV2、AAV3、AAV4、AAV5和AAV6上的相应区域。优选的区域约为250bp，位于AAV1[SEQ ID NO:6]的2886到3143区，以及AAV2[SEQ ID NO:7]，AAV3[SEQ ID NO:8]以及其他AAV血清型上的相应区域。见图1。为了快速检测样品中的AAV，所用的引物为该特征区域特异性的。但是，本发明并不仅限于那些与本文所鉴定的AAV特征区域完全相配的序列，本领域的技术人员可以对这个区加以改动，使之成为比特征区域短点或长点的片段。

[0035] PCR引物是用本领域技术人员所熟知的方法合成的。每对PCR引物都由5’引物和3’引物组成，见Sambrook等，本文已引用。术语“引物”指一个寡核苷酸片段，在适当的条件下，与核酸链互补的引物延伸产物在该点开始合成。但是，也可以使用双链引物，只要在制备延伸产物前将其处理使之分开就可以了。引物含有约15到25个或更多个核苷酸，优选至少18个核苷酸。但是在某些情况下较短的如7到15个核苷酸也可以使用。

[0036] 引物要选择与所要扩增的特定序列的不同链足够互补的序列，并且还能与其各自的链杂交。因此，引物序列并不需要是所要扩增区域的精确序列。例如，可以将一段非互补的核苷酸片段连接到引物的5’末端，但是剩余的引物部分要与该链完全互补。另外，非互补碱基或较长的序列可以分散在引物内，只要引物序列与所要扩增的序列足够互补，能与之杂交，该链可以作为模板来合成引物的延伸产物就可以了。

[0037] 根据本发明，特征区域的PCR引物是基于两个或多个比对序列(即两个或多个AAV血清型)的高度保守区而设计的。引物的序列在5’末端或中间可以加以改动，并不一定要和两个或多个比对的AAV血清型完全互补。但是，引物的3’末端至少要有5个以上的核苷酸是与两个或多个比对的AAV血清型完全互补的，优选9个以上的碱基对完全互补，更优选的是18个碱基对完全互补。这样，引物的3’末端由至少5个与比对序列100％一致的核苷酸组成。但是，也可以在引物的3’末端加上1、2、或多个变性的核苷酸。

[0038] 例如，AAV特征区域的引物对是基于AAV衣壳内的如下特异区而设计的。5’引物是基于AAV2[SEQ ID NO:7]的nt2867-2891，5’-GGTAATTCCTCCGGAAATTGGCATT3’，见图1。3’引

物是基于AAV2[SEQ ID NO:7]的nt2867-2891,5'-GACTCATCAACAACAACTGG GGATTC-3'。但是,本领域的技术人员可以基于AAV1,AAV3,AAV4,AAV5,AAV6的相应区域,或基于AAV7,AAV10,AAV11,AAV12或本发明的其他AAV所提供的信息来设计引物对。另外,还可以利用本领域技术人员所熟知的方法设计其他的引物对来扩增这个特征区域。

[0039] B.靶序列的分离

[0040] 如本文所描述,本发明的第一引物对用于特异性地扩增靶序列即AAV血清型的特征区域,以便于检测该靶序列。如果还需要检测其他的序列,即,如果需要鉴定一个新的血清型,那么特征区域需要延伸,这样,本发明还需要一个或多个附加的引物对。这些引物对适宜设计成包含第一引物对的5'引物或3'引物和该引物对独特的第二引物,这样,该引物对可以扩增出特征区域的5'区或3'区,该序列与特征区域的5'末端或3'末端互补。例如,第一引物对由5'引物P1和3'引物P2组成以扩增特征区域。为了延伸特征区域的3'末端,第二引物对应由引物P1和3'引物P4组成,该引物对可以扩增特征区域及其下游的序列。为了延伸特征区域的5'末端,第三引物对应由5'引物P5和引物P2组成,该引物对可以扩增特征区域及其上游的序列。如果需要,这些延伸步骤可以重复(或同时进行)。因此,从这些扩增步骤中得到的产物与常规步骤所得到的产物接合就得到了所需要长度的独立序列。

[0041] 第二和第三引物对与特征区域的引物对一起用于扩增比对序列上的高度保守区。本文所用的术语"第二"或"第三"引物对只是为了便于描述,并不表示这些引物添加到反应混合物中或用于扩增的顺序。第二引物对所扩增的区域是经过选择的,这样一旦扩增出来后其5'末端就可以与特征区域的3'末端互补杂交。同样,第三引物对所扩增的区域也是经过选择的,这样一旦扩增出来后其3'末端就可以与特征区域的5'末端互补杂交。还可以设计出其他的引物对,这些引物对所扩增的区域可以与第二或第三引物对以及下面的引物对扩增产物的5'末端或3'末端互补杂交。

[0042] 例如,如果以AAV为靶序列,第一引物对(P1和P2)用于从样品中扩增特征区域。在一个优选实施方式中,特征区域位于AAV衣壳之内。第二引物对(P1和P4)用于将特征区域的3'末端延伸到AAV序列3'ITR之前的区域,即以特征区域为锚扩增出含AAV衣壳整个3'末端的产物。在一个实施方式中,P4引物位于AAV2[SEQ ID NO:7]的nt4435-4462,以及其他AAV血清型的相应序列上。结果得到一个约1.6kb的产物,该产物含有0.25kb的特征区域。第三引物对(P3和P2)用于将特征区域的5'末端延伸到AAV序列rep基因的3'末端,即以特征区域为锚扩增出含AAV衣壳整个5'末端的产物。在一个实施方式中,P3引物位于AAV2[SEQ ID NO:7]的nt1384-1409,以及其他AAV血清型的相应序列上。结果得到一个约1.7kb的产物,该产物含有0.25kb的特征区域。另外,第四引物对用于将含AAV衣壳整个5'末端的延伸产物进一步延长到包含rep序列。在一个实施方式中,P5引物位于AAV2[SEQ ID NO:7]的nt108-133,以及其他AAV血清型的相应序列上,与P2引物合用。

[0043] 完成了所希望数量的延伸步骤之后,利用特征区域为锚或标记将各种延伸产物连接起来就得到了一个完整的序列。在本文所描述的一个实施例中,至少含有一个完整的AAV cap基因的AAV序列被扩增出来,当然也可以扩增出更大的序列,这要根据所进行的扩增步骤的多少。

[0044] 将延伸产物组装成一个完整的AAV序列所用的方法是本领域技术人员所熟知的。例如,用DraIII消化扩增产物,可以在特征区域的DraIII位点上将其裂解,得到一个限制性

酶切片段，将此片段与含完整AAV cap基因分产物从新连接起来。但是也可以利用其他合适的技术将延伸产物组装成一个完整的序列，见Sambrook等，本文已引用。

[0045] 除了上面所描述的多步骤延伸方法，本发明的另外一个实施方式是直接扩增出一个3.1kb的片段，该片段含有全长的cap序列。为了从NHP组织或血液DNA中直接扩增3.1kb的全长cap片段，需要鉴定AAV基因组中其他两个高度保守区以便于用PCR方法扩增这个大片段。保守区内的一个引物位于rep基因的中间(AV1ns:5'GCTGCGTCAACTGGACCAATGAGAAC3'，SEQ ID NO:6的nt)，另一个引物位于cap基因下游的另一个保守区内(AV2cas:5'CGCAGAGACCAAAGTTCAACTGAAACGA3'，SEQ ID NO:7)，而这结合可以扩增出含全长cap基因的AAV序列。一般来说，扩增完成以后，产物经克隆和测序其准确率都可以达到99.9以上。利用该方法，发明者至少分离出50个衣壳克隆并随之进行了鉴定。其中37个克隆来源于猕猴(rh.1-rh.37)，6个克隆来源于短尾猴(cy.1-cy.6)，2个克隆来源于狒狒(bb.1和bb.2)，5个克隆来源于黑猩猩(ch.1-ch.5)。这些克隆在说明书的其他地方也作了鉴定，连同其来源的物种以及存在新型序列的动物组织。

[0046] C.分离新型AAV的其他方法

[0047] 本发明的另外一个方面涉及从细胞中分离新型AAV的其他方法。该方法包括用具有AAV辅助病毒功能的载体感染细胞；分离含AAV的感染克隆；分离AAV的测序；以及将分离的AAV序列与已知的AAV血清型进行比较，如果分离的AAV与已知的AAV血清型序列有区别就说明存在新型AAV。

[0048] 在一个实施方式中，具有辅助功能的载体提供了腺病毒的基本功能，包括如E1a，E1b，E2a，E4ORF6。在一个实施方式中，辅助功能是由腺病毒提供。腺病毒可以是野生型的，可以是人源的也可以是非人源的，最好是非人灵长类动物(NHP)来源的。各种腺病毒的DNA序列可以从Genbank上找到，包括Ad5型[Genbank登录号M73260]。腺病毒序列可从已知的任何类型的腺病毒中获得，如血清型2，3，4，7，12和40，还包括任何本发明已经鉴定的人源类型[见Horwitz，上面已引用]。同样，已知可以感染非人动物(如黑猩猩)的腺病毒也可用于构建本发明的载体。见美国专利No.6，083，716。除了野生型腺病毒之外，具有必须辅助功能的重组病毒或非病毒载体(如质粒、游离体等)也可使用。这些重组病毒是本领域所熟知的，可以通过已发表的方法制备。见美国专利No.5，871，982和6，251，677，其中描述了一种杂合的Ad/AAV病毒。对腺病毒类型的选择并不预示着对下列发明的限制。从美国模式培养物保藏所(American Type Culture Collection，Manassas，Virginia)可以得到各种腺病毒株，另外也可以从各种商业机构或研究机构获得。而且，许多这种病毒株的序列都可以从各种数据库中查询到，如PubMed和GenBank。

[0049] 另外，感染性AAV可以用基因组步行技术(Siebert等，1995，Nucleic Acid Research，23:1087-1088，Friezner-Degen等，1986，J.Biol.Chem.261:6972-6985，BD Biosciences Clontech，Palo Alto，CA)分离。基因组步行技术特别适于鉴定和分离与本发明方法所鉴定的新型序列相邻的序列。例如，这个技术可用于分离新型AAV血清型的反向末端重复序列(ITR)，其根据就是用本发明的方法鉴定的新型AAV的衣壳和/或rep序列。该技术也可用于分离与本发明所鉴定和分离的其他AAV或非AAV序列相邻的序列。见实施例3和4。

[0050] 本发明的方法可用于各种流行病学研究，生物分布研究、通过AAV载体或其他整合

病毒来源的载体进行基因治疗的检测。因此，这些方法可用于制备预包装的试剂盒便于临床医生、研究人员和流行病学家使用。

[0051] II.诊断试剂盒

[0052] 本发明的另外一个方面涉及检测样品中是否已知的或未知的腺伴随病毒(AAV)。这种试剂盒包括AAV核酸序列特征区域特异性的第一对5'和3'PCR引物。另外，该试剂盒还可以包含3.1kb片段特异性的第一对5'和3'PCR引物，该片段含有本文所鉴定的全长AAV衣壳核酸序列(即AV1ns和AV2cas引物)。试剂盒还可以任选添加本文所描述的两对或多对5'和3'引物，以及PCR探针。这些引物和探针按照本发明可用于扩增每种AAV血清型的特征区域，如用定量PCR方法。

[0053] 本发明还涉及用于鉴定用本发明的方法检测到的AAV血清型，和/或将新的血清型与已知的血清型区别开来。这种试剂盒还包含一种或多种限制性内切酶，AAV血清型的标准以用于"标记的限制性酶切分析"，和/或其他确定所检测到的AAV血清型的试剂。

[0054] 另外，本发明的试剂盒还包含说明书、阴性和/或阳性对照、容器、样品的稀释液和缓冲液、标记比较的对照表、一次性手套、防污染说明、加样棒或容器、样品制备试管、以及其他任何需要的试剂，如介质、洗涤试剂和浓缩试剂。这些试剂可从本文所描述的试剂中选择，液可以从常规的试剂中选择。在一个优选实施方式中，洗涤试剂是等渗盐溶液，可用于缓冲生理pH，如磷酸缓冲盐(PBS)；洗脱试剂是含0.4M NaCl的PBS，以及浓缩试剂和设备。例如，本发明的技术人员都知道可以使用试剂如聚乙烯甘油(PEG)或NH_4SO_4，设备如过滤装置。例如，含有100K膜的过滤装置可以浓缩rAAV。

[0055] 本发明的试剂盒可用于实现本文所描述的方法，也可用于生物分布、流行病学、新型AAV病毒在人和NHP传播模式的研究。

[0056] 因此，本发明的方法和试剂盒可用于靶病毒序列的检测、鉴定和分离。该方法和试剂盒特别适用于检测、鉴定和分离AAV序列，其中包括新的AAV血清型。

[0057] 在一个值得注意的实施例中，发明者利用本发明的方法促进了对克隆AAV序列的分析，结果显示不同动物来源的克隆片段之间的原病毒序列具有异质性，所有的序列都与已知的6个AAV血清型不同，其变异区主要集中在衣壳蛋白的超变区。令人吃惊的是AAV序列的分散性在从一个短尾猴的单一组织如淋巴结中分离出来的克隆中特别明显。这种异质性的最好解释是动物个体内AAV序列存在明显的进化，部分可能是由于有限数目的共感染亲本病毒之间发生同源重组。这些研究说明天然AAV感染过程中广泛传播病毒的序列进化可能导致了一群准种的形成，这种准种在衣壳超变区的比对上是互不相同的。这是第一个DNA病毒发生快速分子进化的例子，这种方式以前被认为只发生于RNA病毒。

[0058] 本文描述了几种用本发明的方法鉴定的新型AAV血清型及其特征。

[0059] III.新型AAV血清型

[0060] A.核酸序列

[0061] 本文描述了用本发明的方法所鉴定的新型AAV血清型的核酸序列。见SEQ ID NO:1,9-59,以及117-120，本文已纳入作为参考。也可参见图1和表中所列的序列。

[0062] 新血清型AAV7的全长序列，包括AAV5'ITR，衣壳，rep，和AAV3'ITR显示在SEQ ID NO:1中。

[0063] 对于本发明的其他新的AAV血清型，本文描述了按照本发明的方法分离的约3.1kb

的片段。该片段含有编码全长衣壳蛋白的序列和编码rep蛋白的全部或部分序列。这些序列包括下面所鉴定的克隆。

[0064] 对于其他的新AAV血清型来说，本文描述了编码衣壳蛋白的特征区域。例如，本发明的AAV10核酸序列包含了图1所显示的序列[见SEQ ID NO:117，255bp长]。本发明的AAV11核酸序列包含了图1所显示的序列[见SEQ ID NO:118，258bp长]。本发明的AAV12核酸序列包含了图1所显示的序列[见SEQ ID NO:119，255bp长]。利用本文所描述的方法，AAV10、AAV11和AAV12序列可以很容易地被鉴定，并用于各种目的，如AAV7和本文所描述的其他新血清型所用的各种目的。

[0065] 图1显示的是本发明的非人灵长类(NHP)AAV核酸序列，以前文献所报道的AAV血清型AAV1[SEQ ID NO:6]、AAV2[SEQ ID NO:7]和AAV3[SEQ ID NO:8]与其一起比对。这些新型NHP序列包括下面表1中所列的序列，该序列是通过克隆数目来鉴定的。

[0066] 表1

[0067]

AAV Cap 序列	克隆标号	来源		
		物种	组织	序列号(DNA)
Rh.1	克隆 9(AAV9)	猕猴	心脏	5
Rh.2	克隆 43.1	猕猴	MLN	39
Rh.3	克隆 43.5	猕猴	MLN	40
Rh.4	克隆 43.12	猕猴	MLN	41
Rh.5	克隆 43.20	猕猴	MLN	42
Rh.6	克隆 43.21	猕猴	MLN	43
Rh.7	克隆 43.23	猕猴	MLN	44
Rh.8	克隆 43.25	猕猴	MLN	45
Rh.9	克隆 44.1	猕猴	肝脏	46
Rh.10	克隆 44.2	猕猴	肝脏	59
Rh.11	克隆 44.5	猕猴	肝脏	47
Rh.12	克隆 42.1B	猕猴	MLN	30
Rh.13	克隆 42.2	猕猴	MLN	9
Rh.14	克隆 42.3A	猕猴	MLN	32

[0068]

Rh.15	克隆 42.3B	猕猴	MLN	36
Rh.16	克隆 42.4	猕猴	MLN	33
Rh.17	克隆 42.5A	猕猴	MLN	34
Rh.18	克隆 42.5B	猕猴	MLN	29
Rh.19	克隆 42.6B	猕猴	MLN	38
Rh.20	克隆 42.8	猕猴	MLN	27
Rh.21	克隆 42.10	猕猴	MLN	35
Rh.22	克隆 42.11	猕猴	MLN	37
Rh.23	克隆 42.12	猕猴	MLN	58
Rh.24	克隆 42.13	猕猴	MLN	31
Rh.25	克隆 42.15	猕猴	MLN	28
Rh.26	克隆 223.2	猕猴	肝脏	49
Rh.27	克隆 223.4	猕猴	肝脏	50
Rh.28	克隆 223.5	猕猴	肝脏	51
Rh.29	克隆 223.6	猕猴	肝脏	52
Rh.30	克隆 223.7	猕猴	肝脏	53
Rh.31	克隆 223.10	猕猴	肝脏	48
Rh.32	克隆 C1	猕猴	脾、十二指肠、肾脏和肝脏	19
Rh.33	克隆 C3	猕猴		20
Rh.34	克隆 C5	猕猴		21
Rh.35	克隆 F1	猕猴	肝脏	22
Rh.36	克隆 F3	猕猴		23
Rh.37	克隆 F5	猕猴		24
Cy.1	克隆 1.3	短尾猴	血	14
Cy.2	克隆 13.3B	短尾猴	血	15
Cy.3	克隆 24.1	短尾猴	血	16
Cy.4	克隆 27.3	短尾猴	血	17
Cy.5	克隆 7.2	短尾猴	血	18
Cy.6	克隆 16.3	短尾猴	血	10

[0069]

bb.1	克隆 29.3	猕猴	血	11
bb.2	克隆 29.5	猕猴	血	13
Ch.1	克隆 A3.3	黑猩猩	血	57
Ch.2	克隆 A3.4	黑猩猩	血	54
Ch.3	克隆 A3.5	黑猩猩	血	55
Ch.4	克隆 A3.7	黑猩猩	血	56

[0070] 将两个黑猩猩腺病毒的衣壳片段插入到AAV rep载体中得到了一个新型NHP克隆。这个新型克隆A3.1也被称为Ch.5[SEQ ID NO:20]。另外，本发明还包括两个人AAV序列，分别被称为H6[SEQ ID NO:25]和H2[SEQ ID NO:26]。

[0071] 本发明的AAV核酸序列还包括与图1中所列序列、序列[SEQ ID NO:1,9-59,117-120]互补的链以及图1所列序列、序列[SEQ ID NO:1,9-59,117-120]及其互补链相应的RNA和cDNA。本发明的核酸序列还包括图1所列序列、序列[SEQ ID NO:1,9-59,117-120]及其互补链的天然突变株以及经基因工程修饰的序列。这些修饰包括本领域所熟知的标记、甲基化、以及用变性核苷酸替代一个或多个天然核苷酸。

[0072] 本发明的核酸序列还包括与图1所列序列、序列[SEQ ID NO:1,9-59,117-120]具有大于85％的同一性或同源性的序列，较好的是至少约90％，更好的是至少约95％，最好的是至少约98％-99％的同一性或同源性。这些术语已在本文定义。

[0073] 本发明还包括用本文所描述的方法鉴定的新型AAV序列的片段。适宜的片段至少要有15个核苷酸的长度，并且包含功能片段，即该片段具有生物学意义。在一个实施方式中，这些片段是图1所列新型序列、序列[SEQ ID NO:1,9-59,117-120]、其互补链及其互补的cDNA和RNA的片段。

[0074] 适宜的片段是那些位于AAV1、AAV2或AAV7上的片段。但是，利用本文所描述的比对方法(用Clustal W程序比对，默认设置)，或与其他新血清型比对的相似方法，本领域的技术人员可以很容易地确定目的片段的起始密码子和终止密码子。

[0075] 适宜的片段包括编码AAV衣壳三个可变蛋白(vp)的序列，它们具有不同的剪接突变：vp1[即AAV7的nt825to3049,SEQ ID NO:1]；vp2[即AAV7的nt1234-3049,SEQ ID NO:1]；以及vp3[即AAV7的nt1434-3049,SEQ ID NO:1]。值得注意的是AAV7具有不同寻常的起始密码子。除了几个看家基因以外，以前很少报道DNA病毒具有这种起始密码子。其他AAV血清型vp1、vp2和vp3的起始密码子相信也是如此，该特点使其在细胞内产生vp1、vp2和vp3蛋白的比例分别为10%:10%:80%，从而能有效地装配成病毒颗粒。但是，研究表明即使有这个稀有的GTG起始密码子AAV7病毒颗粒也能有效装配。这样，发明者设想是否能改变其他AAV血清型vp3的起始密码子使其含有这个稀有GTG起始密码子，从而提高包装效率，改变病毒颗粒的结构和/或改变其他AAV血清型表面抗原表位(即中和抗体表位)的位置。起始密码子可以通过常规的方法改变，如直接位点突变法。因此，本发明包括经改造的任何所选择的血清型AAV病毒颗粒，其中含有起始密码子变为GTG的vp3和/或再加上vp1和/或vp2。

[0076] AAV的其他适宜片段包括含AAV衣壳蛋白起始密码子的片段[即AAV7的nt468-3090,SEQ ID NO:1;AAV7的nt725to3090,SEQ ID NO:1,以及其他AAV血清型的相应区域]。

AAV7和用本文所描述的方法鉴定的其他新型AAV血清型的其他片段包括那些编码rep蛋白的序列，其中包括rep78[即图1中AAV7的起始密码子334]、rep68[图1中AAV7的起始密码子nt334]、rep52[图1中AAV7的起始密码子1006]以及rep40[图1中AAV7的起始密码子1006]。其他感兴趣的片段还有AAV的5'末端反向重复序列ITR[图1中AAV7的nt1-107]；AAV3'ITR[图1中AAV7的nt4704-4721]，P19序列，AAV P40序列，rep结合位点以及末端分辨位点(TRS)。其他适宜的片段对于本领域的技术人员来说是显而易见的。本发明其他新血清型上的相应区域通过参考图1或利用常规的比对方法与本文所描述的序列对照后就可以很容易地确定。

[0077] 除了包括图中及序列表中所列的核酸序列以外，本发明还包括用于表达本发明AAV血清型的氨基酸序列、蛋白质和多肽的核酸分子和序列。因此，本发明包括编码下列新型AAV氨基酸序列的核酸序列：C1[SEQ ID NO:60]，C2[SEQ ID NO:61]，C5[SEQ ID NO:62]，A3-3[SEQ ID NO:66]，A3-7[SEQ ID NO:67]，A3-4[SEQ ID NO:68]，A3-5[SEQ ID NO:69]，3.3b[SEQ ID NO:62]，223.4[SEQ ID NO:73]，223-5[SEQ ID NO:74]，223-10[SEQ ID NO:75]，223-2[SEQ ID NO:76]，223-7[SEQ ID NO:77]，223-6[SEQ ID NO:78]，44-1[SEQ ID NO:79]，44-5[SEQ ID NO:80]，44-2[SEQ ID NO:81]，42-15[SEQ ID NO:84]，42-8[SEQ ID NO:85]，42-13[SEQ ID NO:86]，42-3A[SEQ ID NO:87]，42-4[SEQ ID NO:88]，42-5A[SEQ ID NO:89]，42-1B[SEQ ID NO:90]，42-5B[SEQ ID NO:91]，43-1[SEQ ID NO:92]，43-12[SEQ ID NO:93]，43-5[SEQ ID NO:94]，43-21[SEQ ID NO:96]，43-25[SEQ ID NO:97]，43-20[SEQ ID NO:99]，24.1[SEQ ID NO:101]，42.2[SEQ ID NO:102]，7.2[SEQ ID NO:103]，27.3[SEQ ID NO:104]，16.3[SEQ ID NO:105]，42.10[SEQ ID NO:106]，42-3B[SEQ ID NO:107]，42-11[SEQ ID NO:108]，F1[SEQ ID NO:109]，F5[SEQ ID NO:110]，F3[SEQ ID NO:111]，42-6B[SEQ ID NO:112]，和/或42-12[SEQ ID NO:113]，以及利用这些序列和/或其独特片段准备的人工AAV血清型。

[0078] 如本文所描述，人工AAV血清型包括而不限于具有非天然衣壳蛋白的AAV。这种人工衣壳可以用任何适宜的方法制备，利用本发明的新型AAV序列(即vp1衣壳蛋白的片段)加上异源序列，这些异源序列可以来自其他AAV血清型(已知的或新的)、同一AAV血清型的不连续部分、非AAV病毒、或非病毒序列。人工AAV血清型可能而不限于具有嵌合的AAV衣壳、重组AAV衣壳或“人工”AAV衣壳。

[0079] B.AAV氨基酸序列、蛋白质和多肽。

[0080] 本发明涉及本文所鉴定的新型AAV血清型核酸序列所编码的蛋白质及其片段，其中包括AAV7[AAV7的nt825-3049，SEQ ID NO:1]和本文所描述的其他新血清型。因此，本发明的新血清型的衣壳蛋白包括：H6[SEQ ID NO:25]，H2[SEQ ID NO:26]，42-2[SEQ ID NO:9]，42-8[SEQ ID NO:27]，42-15[SEQ ID NO:28]，42-5b[SEQ ID NO:29]，42-1b[SEQ ID NO:30]；42-13[SEQ ID NO:31]，42-3a[SEQ ID NO:32]，42-4[SEQ ID NO:33]，42-5a[SEQ ID NO:34]，42-10[SEQ ID NO:35]，42-3b[SEQ ID NO:36]，42-11[SEQ ID NO:37]，42-6b[SEQ ID NO:38]，43-1[SEQ ID NO:39]，43-5[SEQ ID NO:40]，43-12[SEQ ID NO:41]，43-20[SEQ ID NO:42]，43-21[SEQ ID NO:43]，43-23[SEQ ID NO:44]，43-25[SEQ ID NO:45]，44.1[SEQ ID NO:47]，44.5[SEQ ID NO:47]，223.10[SEQ ID NO:48]，223.2[SEQ ID NO:49]，223.4[SEQ ID NO:50]，223.5[SEQ ID NO:51]，223.6[SEQ ID NO:52]，223.7[SEQ ID

NO:53],A3.4[SEQ ID NO:54],A3.5[SEQ ID NO:55],A3.7[SEQ ID NO:56],A3.3[SEQ ID NO:57],42.12[SEQ ID NO:58],和44.2[SEQ ID NO:59],可以利用常规的技术从上述列表的克隆中的开放阅读框架来制备。

[0081] 本发明还包括利用本发明的新型AAV血清型的序列制备的AAV血清型,制备方法包括本领域技术人员所熟知的合成技术、重组技术以及其他技术。本发明并不限于本发明的新型AAV核酸序列所表达的新型AAV氨基酸序列、多肽和蛋白质,也包括用本领域所熟知的其他方法制备的氨基酸序列、多肽和蛋白质,如化学合成技术、其他合成技术或其他方法。例如,下列序列中的任何一个:C1[SEQ ID NO:60],C2[SEQ ID NO:61],C5[SEQ ID NO:62],A3-3[SEQ ID NO:66],A3-7[SEQ ID NO:67],A3-4[SEQ ID NO:68],A3-5[SEQ ID NO:69],3.3b[SEQ ID NO:62],223.4[SEQ ID NO:73],223-5[SEQ ID NO:74],223-10[SEQ ID NO:75],223-2[SEQ ID NO:76],223-7[SEQ ID NO:77],223-6[SEQ ID NO:78],44-1[SEQ ID NO:79],44-5[SEQ ID NO:80],44-2[SEQ ID NO:81],42-15[SEQ ID NO:84],42-8[SEQ ID NO:85],42-13[SEQ ID NO:86],42-3A[SEQ ID NO:87],42-4[SEQ ID NO:88],42-5A[SEQ ID NO:89],42-1B[SEQ ID NO:90],42-5B[SEQ ID NO:91],43-1[SEQ ID NO:92],43-12[SEQ ID NO:93],43-5[SEQ ID NO:94],43-21[SEQ ID NO:96],43-25[SEQ ID NO:97],43-20[SEQ ID NO:99],24.1[SEQ ID NO:101],42.2[SEQ ID NO:102],7.2[SEQ ID NO:103],27.3[SEQ ID NO:104],16.3[SEQ ID NO:105],42.10[SEQ ID NO:106],42-3B[SEQ ID NO:107],42-11[SEQ ID NO:108],F1[SEQ ID NO:109],F5[SEQ ID NO:110],F3[SEQ ID NO:111],42-6B[SEQ ID NO:112],和/或42-12[SEQ ID NO:113],利用各种技术可以很容易地制备。

[0082] 适宜的制备技术是本领域技术人员所熟知的。见Sambrook等,分子克隆:实验室手册,冷泉港出版社(冷泉港,NY)。另外,多肽也可以用大家所熟知的固相多肽合成方法合成(Merrifield,J.Am.Chem.Soc.,85:2149(1962);Stewart和Young,固相肽合成(Solid Phase Peptide Synthesis)(Freeman,旧金山,1969),27-62页)。这些方法和其他适宜的制备方法是本领域技术人员所熟知的,并不是对本发明的限制。

[0083] 特别优选的蛋白质包括由上述方法鉴定的核苷酸序列编码的AAV衣壳蛋白。本发明的许多衣壳蛋白的序列列在图2和/或序列表的SEQ ID NO:2和60-115中,本文已纳入作为参考。AAV衣壳由三种蛋白质组成,vp1、vp2和vp3,它们是选择性的剪接突变株。这些图中所列的全长序列是vp1的序列。基于AAV7衣壳[SEQ ID NO:2]的编号,vp2的序列跨越AAV7的氨基酸138-737,vp3的序列跨越AAV7的氨基酸203-737。根据这一信息,本领域的技术人员可以很容易地确定vp2和vp3在本发明的其他新型血清型上的位置。

[0084] 衣壳蛋白的其他优选蛋白和片段包括位于两个超变区(HPV)之间的恒定区和可变区以及HPV区本身的序列。利用一种分析AAV2上序列分散区域的算法检测出了12个超变区(HVR),其中5个与以前所描述的可变区有重叠或是其一部分[Chiorini等,J.Virol,73:1309-19(1999);Rutledge等,J.Virol.,72:309-319]。利用这种算法和/或本文所描述的比对方法,可以确定新型AAV血清型的HVR。例如,对于AAV2的vp1[SEQ ID NO:70]来说,HVR的位置如下:HVR1,aa146-152;HVR2,aa182-186;HVR3,aa262-264;HVR4,aa381-383;HVR5,aa450-474;HVR6,aa490-495;HVR7,aa500-504;HVR8,aa514-522;HVR9,aa534-555;HVR10,aa581-594;HVR11,aa658-667;以及HVR12,aa705-719。根据由传统方法推算出的比对方法

和本文所描述的新型序列[见图2]，可以很容易地确定HVR在本发明的新型AAV血清型上的位置。例如，根据图2，可以很容易地确定HVR在AAV7上的位置，HVR1位于aa146-152；HVR2位于aa182-187；HVR3位于aa263-266，HVR4位于aa383-385，HVR5位于aa451-475；HVR6位于aa491-496；HVR7位于aa501-505；HVR8位于aa513-521；HVR9位于aa533-554；HVR10位于aa583-596；HVR11位于aa660-669；HVR12位于aa707-721。根据本文所提供的信息，其他新型AAV血清型的HVR也很容易确定。

[0085] 另外，在衣壳内也鉴定出了具有同一性的氨基酸表达盒。这些表达盒很有意义，因为可以利用它们构建人工血清型，比如用一种血清型的HVR1表达盒替代另一种血清型的HVR1表达盒。这些表达盒的同一性在图1的HVR2中有显示。见图2。如果用Clustal X比对方法，则该方法有一个位于序列下的标尺，起始于第一个残基，并标识为1。标尺上的线用于标记高度保守的区域。使用了三种字符(*、:、.)，“*”代表单一的完全保守的位置；“:”代表完全保守的一个“强”组，“.”代表完全保守的一个“弱”组。在Gonnet Pam250矩阵中有全部的正分组。强组的分数大于0.5，弱组的分数小于0.5。

[0086] 另外，AAV衣壳其他适宜的片段包括AAV2[SEQ ID NO:70]上的aa24-42，aa25-28；aa81-85；aa133-165；aa134-165；aa137-143；aa154-156；aa194-208；aa261-274；aa262-274；aa171-173；aa413-417；aa449-478；aa494-525；aa534-571；aa581-601；aa660-671；aa709-723。其他的优选片段包括AAV7上的aa1-184，aa199-259；aa274-446；aa603-659；aa670-706；aa724-736；aa185-198；aa260-273；aa447-477；aa495-602；aa660-669；以及aa707-723。根据AAV7[SEQ ID NO:2]的编号，其他优选的区域还可以旋子含有如下片段的一组：aa185-198；aa260-273；aa447-477；aa495-602；aa660-669；以及aa707-723。利用默认设置的Clustal X程序或默认设置的其他商业的或公共的比对软件，根据本文所比对的序列，可以很容易地确定本发明的新型AAV衣壳中的相应片段。

[0087] 其他优选的蛋白质是AAV rep蛋白[AAV7SEQ ID NO:3的aa1-623]及其功能片段，包括SEQ ID NO:3的aa1-171，aa172-372，aa373-444，aa445-623。优选的这种片段是具有至少8个氨基酸长度的。见图3。利用本文所描述的技术和本领域所熟知的技术，可以确定本发明的其他新型AAV蛋白上的类似区域。另外，也可以使用其他所需要长度的片段。这些片段可以通过重组方法或其他适宜的方法如化学合成方法制备。

[0088] 本发明的序列、蛋白或片段可以通过任何适宜的方法制备，包括重组方法、化学合成方法或其他合成方法。这些制备方法都是本领域技术人员所熟知的，并不代表对本发明的限制。

[0089] IV.具有新型AAV衣壳的rAAV的制备

[0090] 本发明包括通过本发明的方法鉴定的新的野生型的AAV血清型，这些野生型AAV血清型的序列不含有天然病毒所具有的DNA和/或细胞组分。本发明的另一个方面涉及利用本发明的新型AAV序列，包括其片段制备用于将异源基因或其他核酸序列导入到靶细胞内的分子。

[0091] 本发明的分子含有本发明的新型AAV血清型序列，其中包括任何能被导入到宿主细胞内的基因元件(载体)，如裸DNA、质粒、噬菌体、转座子、粘粒、游离体、非病毒转移载体内的蛋白(即脂质载体)、病毒等，这些分子可以转移其所携带的序列。载体可通过任何适宜的方法转导，包括转染、电穿孔、脂质体转移、膜融合、高电压的DNA包被颗粒、病毒感染和原

生质融合。用于实施本发明的任何实施方式的方法都是核酸操作技术领域的技术人员所熟知的，包括基因工程技术、重组工程技术及合成技术，Sambrook等，分子克隆：实验室手册，冷泉港出版社，冷泉港，NY。

[0092] 在一个实施方式中，本发明的载体含有编码本发明新型AAV衣壳的序列(即衣壳的AAV7衣壳，AAV44-2(rh.10)，AAV10衣壳，AAV11衣壳，AAV12衣壳)，或一种或多种这些AAV衣壳的片段。另外，载体还可以含有衣壳蛋白及其片段。

[0093] 本发明的载体还可以含有编码AAV rep蛋白的序列。这种rep序列可以来自于具有rep序列的同一AAV血清型。另外，本发明的载体所含有的rep序列与cap序列还可以来自于不同的AAV血清型。在一个实施方式中，rep和cap序列是不同表达来源的(即不同的载体，或一个宿主细胞和一个载体)。在另外一个实施方式中，这些rep序列和cap序列的表达来源相同。在这种实施方式中，rep序列可以和不同AAV血清型的cap序列在一个框架内熔合形成一个嵌合的AAV载体。另外，本发明的载体还可以含有一个小基因，这个小基因包含一个筛选基因，AAV5'ITR和AAV3'ITR位于其两侧。

[0094] 因此，在一个实施方式中，本文所描述的载体含有编码完整AAV衣壳的核酸序列，该衣壳可以来自单一的AAV血清型(即AAV7或其他的新型AAV)。另外，这些载体还可以含有编码人工衣壳的序列，这种衣壳是由AAV7(或其他新型AAV)衣壳的一个或多个片段与异源AAV或非AAV衣壳蛋白(或片段)熔合而成的。这些人工的衣壳蛋白选自AAV7(或其他新型AAV)衣壳的不连续部分或其他AAV血清型的衣壳。例如，我们可能希望修饰一个或多个AAV vp1上的编码区，即一个或多个超变区上的编码区(HPV1-12)，或vp2，和/或vp3。在另外一个实施方式中，我们可能希望将vp3的起始密码子改变成GTG。这些修饰可能能够增加表达量，产出量，和/或改善蛋白在选择性表达系统中的纯化，或者其他目的(如改变包涵体或改变中和抗体表位)。

[0095] 本文所描述的载体即质粒可用于各种目的，但是特别适用于制备含衣壳的rAAV，该衣壳具有AAV序列或其片段。这些载体，包括rAAV、其元件、构建体及其应用，在本文中都有详细描述。

[0096] 本发明的一个方面涉及制备具有AAV血清型7(或其他新型AAV)衣壳或其片段的重组腺伴随病毒(AAV)。这种方法包括培养宿主细胞，该细胞含有编码如本文所定义的腺伴随病毒(AAV)血清型7(或其他新型AAV)衣壳蛋白或其片段的核酸序列；具有功能的rep基因；至少含有AAV末端反向重复序列(ITR)和转基因的小基因；以及足够的辅助功能以便于将小基因包装导AAV7(或其他新型AAV)的衣壳蛋白内。

[0097] 需要在宿主细胞内培养并将AAV小基因包装到AAV衣壳内的组分可用转导的方式提供给宿主细胞。另外，所需组分的任何一种或多种(即小基因，rep序列，cap序列，和/或辅助功能)也可以通过一种稳定的宿主细胞提供，该细胞经基因改造后含有一种或多种所需要的组分，改造方法是本领域技术人员所熟知的。最好的这种稳定宿主细胞所含有的所需组分应该位于可诱导的启动子控制之下。但是，所需组分也可以位于其自身组成启动子控制之下。适宜的可诱导启动子和自身组成启动子在本文的转基因适宜调节元件的讨论部分有描述。另外，选择性稳定的宿主细胞还可以含有自身组成启动子控制下的可选组分以及一个或多个可诱导启动子控制下的其他可选组分。例如，稳定的宿主细胞可以来源于293细胞(该细胞含有E1辅助功能，位于自身组成启动子控制之下)，但是它也含有可诱导启动子

控制之下的rep和/或cap蛋白。其他稳定的宿主细胞本领域的技术人员也可以制备。

[0098] 制备本发明rAAV所需要的小基因、rep序列、cap序列以及辅助功能可以任何基因元件的形式导入到包装细胞内，这些元件可以将所携带的序列导入到细胞内。这些选择性的基因元件可以通过任何适宜的方法导入，包括本文所描述的方法。用于实施本发明的任何实施方式的方法都是核酸操作技术领域的技术人员所熟知的，包括基因工程技术、重组工程技术及合成技术，见Sambrook等，分子克隆:实验室手册，冷泉港出版社，冷泉港，NY。同样，制备rAAV病毒颗粒的方法也是本领域所熟知的，选择何种适宜的方法不代表对本发明的限制。见K.Fisher等，J.Virol.，70:520-532(1993)和美国专利5,478,745号。

[0099] A.小基因

[0100] 小基因至少由转基因及其调节序列、AAV5'和3'末端反向重复序列(ITR)组成。正是这个小基因被包装到衣壳蛋白中并被导入到选择性的宿主细胞内。

[0101] 1.转基因

[0102] 转基因是一种核酸序列，与其两侧的载体序列不同源，编码目的多肽、蛋白或其他产物。核酸编码序列通过人工方式连接上调节元件，这些调节元件可以使转基因在宿主细胞内转录、翻译和/或表达。

[0103] 转基因的组成根据其载体用途的不同而不同。例如，一种类型的转基因序列包含一个报告基因序列，其表达后可以产生一个可检测信号。这种报告基因包括而不限于编码下列蛋白的DNA序列:β-内酰胺酶、β-半乳糖苷酶(LacZ)、碱性磷酸酶、胸腺激酶、绿色荧光蛋白(GFP)、氯霉素乙酰转移酶(CAT)、荧光素酶、膜结合蛋白如CD2、CD4、CD8、流感血凝素蛋白以及本领域所熟知的直接存在的或通过常规方法制备的高亲和力抗体可以结合的其他蛋白，含有膜结合蛋白与血凝素或Myc来源的抗原标签区适当熔合而形成的融合蛋白。

[0104] 这些编码序列与控制其表达的调节序列结合后就可以提供能用常规方法检测的信号，包括酶催化分析方法、放射性同位素测定方法、比色法、荧光或其他化学发光法、荧光活化细胞分类分析方法、免疫分析方法如酶联免疫分析方法(ELISA)、放射免疫分析方法(RIA)和免疫组化分析方法。例如，当标记序列是LacZ基因时，带有这个信号的载体可以通过分析β半乳糖苷酶的活性来检测。当转基因是绿色荧光蛋白或荧光素酶时，携带该信号的载体发出的颜色或光信号可在显微镜下观察到。

[0105] 但是，转基因也可以是非标记序列，编码用于生物学或药学的产物，如蛋白、多肽、RNA、酶或催化RNA。优选的RNA分子包括tRNA、dsRNA、核糖体RNA、催化RNAs，和反义RNAs。一个有用的RNA序列的例子是能灭活被处理动物体内靶核酸序列的表达的RNA。

[0106] 转基因可用于矫正或改善基因缺陷，其中包括正常基因表达水平低于正常值的基因缺陷或功能基因产物不表达的基因缺陷。一种优选的转基因序列编码治疗性蛋白或多肽，该蛋白或多肽在宿主细胞内表达。本发明还涉及使用多个基因以矫正或改善由多亚单位蛋白引起的基因缺陷。在某些情况下，不同的转基因用于编码蛋白的不同亚单位，或编码不同的多肽或蛋白。比较理想的是编码蛋白亚单位的DNA较大，如免疫球蛋白、血小板来源的生长因子或营养障碍蛋白。为了使细胞产生多亚单位的蛋白质，需要用能同时表达不同亚单位的重组病毒转染细胞。另外，蛋白的不同亚单位也可以用同一转基因编码。在这种情况下，这一个转基因包含了编码不同亚单位的DNA，编码每个亚单位的DNA用内部核糖体结合位点(IRES)分隔开。我们也希望编码每个亚单位的DNA是较小的，如编码亚单位的DNA和

IRES的总大小在5kb以内。除了IRES外，DNA也可以用编码2A多肽的序列分隔开，这个多肽在翻译后可以自身裂解。见M.L.Donnelly等，J.Gen.Virol.，78(第一部分)：13-21(1997.1)；Furler，S等，Gene Ther.，8(11)：864-873(2001.6)；Klump H等，Gene Ther.，8(10)：811-817(2001.5)。这个2A多肽比IRES小很多，因此很适于片段大小成为限制因素的时候。但是，选择性转基因可以编码任何具有生物活性的产物或其他产物，如用于研究的产物。

[0107] 本领域的技术人员可以很容易地选择适宜的转基因。转基因的选择不意味着对本发明的限制。

[0108] 2.调节元件

[0109] 除了上述小基因的主要元件以外，载体还可以含有一些必须的常规控制元件，这些元件人工连接到转基因上，控制其在转染质粒载体或感染本发明的病毒的细胞内的转录、翻译和/或表达。如本文所述，人工连接的序列包括与转基因相邻的表达控制序列和反式作用的表达控制序列或在远端调节转基因的表达控制序列。

[0110] 表达控制序列包含适当的转录起始点、终止点、启动子和增强子序列；有效的RNA处理信号如剪接位点和多聚腺苷酸(polyA)信号；稳定胞浆mRNA的序列；提高翻译效率的序列(即Kozak共有序列)；提高蛋白稳定性的序列；如果需要还可以含有增加编码产物分泌的序列。有许多表达控制序列，包括本领域所熟知的天然的、自身组成的、可诱导的和/或组织特异性的启动子都可以使用。

[0111] 自身组成启动子的例子包括而不限于逆转录病毒劳氏肉瘤病毒(RSV)LTR启动子(还可以加上增强子)、巨细胞病毒(CMV)启动子(还可以加上增强子)[见Boshart等，Cell，41:521-530(1985)]、SV40启动子、二氢叶酸还原酶启动子、β-肌动蛋白启动子、磷酸甘油激酶(PGK)启动子、以及EF1α启动子[Invitrogen]。

[0112] 可诱导启动子能够调节基因的表达，并且其本身也可以被外源添加的化合物、环境中的因子所调节，这些因素包括温度、或特异性的生理状态如急性期、细胞的分化阶段或只有在复制的细胞内。可诱导启动子和可诱导系统可从许多公司获得，其中有Invitrogen、Clontech和Ariad。其他许多系统也有报道，本领域的技术人员很容易选择。能被外源添加启动子调节的可诱导启动子包括锌诱导的绵羊金属硫蛋白(MT)启动子、地塞米松(Dex)诱导的鼠乳腺致癌因子(MMTV)启动子、T7多聚酶启动子系统[WO98/10088]、昆虫蜕皮激素启动子[No等，Proc.Natl.Acad.Sci.USA，93:3346-3351(1996)]、四环素可抑制系统[Gossen等，Proc.Natl.Acad.Sci.USA，89:5547-5551(1992)]、四环素可诱导系统[Gossen等，Science，268:1766-1769(1995)，也参见Harvey等，Curr.Opin.Chem.Biol.，2:512-518(1998)]、RU486可诱导系统[Wang等，Nat.Biotech.，15:239-243(1997)和Wang等，Gene Ther.，4:432-441(1997)]以及雷帕霉素可诱导系统[Magari等，J.Clin.Invest.，100:2865-2872(1997)]。其他可用于本发明的可诱导启动子是那些由特异性的生理状态所调节的启动子，如稳定、急性期、细胞的特定分化状态或复制状态。

[0113] 在另外一个实施方式中所使用的是转基因本身的天然启动子。如果只是想要使转基因达到天然的表达水平，那么天然启动子是优选的。如果转基因的表达必须受暂时的调节、发育的调节、或以组织特异的方式调节、或者受某种特异性的转录刺激信号的调节，那么也可以用天然的启动子。在另外一个实施方式中，也可以用其他的天然表达控制元件如增强子、多聚腺苷酸位点或Kozak共有序列来模拟天然表达。

[0114] 转基因的另外一个实施方式是将转基因与组织特异性启动子相连接。例如，如果希望在骨骼肌内表达，应该使用在肌肉组织内有活性的启动子。这些启动子包括编码下列蛋白质的基因的启动子：骨骼肌β-肌动蛋白、肌球蛋白轻链2A、营养不良素、肌肉肌酸激酶、以及比天然启动子活性高的合成的肌肉启动子(见Li等，Nat.Biotech.，17:241-245(1999))。已知的组织特异性启动子是肝组织特异性的(白蛋白，Miyatake等，J.Virol.，71:5124-32(1997)；乙型肝炎核启动子，Sandig等，Gene Ther.，3:1002-9(1996)；α-甲胎蛋白(AFP)，Arbuthnot等，Hum.Gene Ther.，7:1503-14(1996))，骨钙素(Stein等，Mol.Biol.Rep.，24:185-96(1997))；骨唾液蛋白(Chen等，J.Bone Miner.Res.，11:654-64(1996))，淋巴细胞特异性的(CD2，Hansal等，J.Immunol.，161:1063-8(1998)；免疫球蛋白重链；T细胞受体α链)，神经元特异性的如神经元特异性的烯醇化酶(NSE)启动子(Andersen等，Cell.Mol.Neurobiol.，13:503-15(1993))，神经丝轻链基因(Piccioli等，Proc.Natl.Acad.Sci.USA，88:5611-5(1991))，以及神经元特异性的vgf基因(Piccioli等，Neuron，15:373-84(1995))。

[0115] 另外，携带治疗性转基因的质粒也可以包含选择标记基因或报告基因，如编码遗传霉素、潮霉素或嘌呤霉素(purimycin)耐药的序列。这种选择报告基因或标记基因(优选位于基因组的外面以便于用本发明的方法恢复)可用于显示细菌细胞内是否存在质粒，如氨苄青霉素耐药基因。质粒的其他元件还可以包括复制起点。这些元件和其他的启动子以及载体元件的选择是常规的，有许多可以供选择[见Sambrook等，本文已引用作为参考]。

[0116] 转基因、启动子/增强子、以及5'和3'ITR一起被称为“小基因”以便于本文参考。根据本发明的提示，通过常规的技术可以设计出这种小基因。

[0117] 3.将小基因导入包装宿主细胞内

[0118] 小基因可以插入到任何适宜的载体内，如质粒内，然后被导入到宿主细胞内。用于本发明的质粒可以经基因工程改造以使其能在原核细胞、或哺乳动物细胞、或者在两者内复制并整合。这些质粒(或携带5'AAV ITR-异源分子-3'ITR)含有能使小基因在真核细胞和/或原核细胞内复制的序列以及这些系统的筛选标记。筛选标记或报告基因可以包括编码遗传霉素、潮霉素或嘌呤霉素耐药的序列。质粒还可以含有可用于显示细菌细胞内是否存在质粒的某种选择报告基因或标记基因，如氨苄青霉素耐药基因。质粒的其他元件包括复制起点和扩增子，如Epstein Barr病毒核抗原的扩增子系统。这种扩增子系统或其他类似的扩增子元件使载体在细胞内以游离体形式高拷贝复制。携带小基因的分子优选转染到细胞内，其在细胞内可以是瞬时存在的。另外，小基因(携带5'AAV ITR-源分子-3'ITR)可以稳定整合到宿主细胞的基因组内，或者是整合到染色体上，或者以游离体的形式存在。在某些实施方式中，小基因以多拷贝的形式存在，形成头-头相连、头-尾相连或尾-尾相连的串联体。适宜的转染技术是大家所熟知的，都可用于将小基因转染到宿主细胞内。

[0119] 当以转染的方式转导含小基因的载体时，Epstein Barr载体一般的量为约5μg到100μg DNA，优选约10到50μg DNA，细胞数约为1×10^4到1×10^{13}，优选约10^5个细胞。但是，载体DNA与宿主细胞的相对量可以调整，这主要考虑所选择的载体、转导方法、所选择的宿主细胞等因素。

[0120] B.rep和cap序列

[0121] 除了小基因以外，宿主细胞还含有驱动新型AAV衣壳蛋白(如AAV7或其他新型AAV

衣壳，或含有一个或多个这种衣壳的片段的人工衣壳蛋白）宿主细胞内表达的序列以及与小基因内AAV ITR的血清型来源相同的rep序列。AAV的cap序列和rep序列可以各自独立地从一个AAV来源中获得，然后用上面所描述的本领域技术人员所熟知的任何方法导入到宿主细胞内。另外，当假设一类本发明的新型AAV衣壳时，编码每一rep蛋白的序列可以来源于同一AAV血清型，也可以来源于不同的血清型（如AAV1，AAV2，AAV3，AAV4，AAV5，AAV6或本文所鉴定的一个新血清型）。例如，rep78/68序列可以来自AAV2，而rep52/40序列可以来自AAV1。

[0122] 在一个实施方式中，宿主细胞稳定地表达适宜启动子控制下的衣壳蛋白，如上面所描述的。在这个实施方式中，衣壳蛋白最好位于可诱导启动子的控制之下。在另一个实施方式中，衣壳蛋白以反式方式提供给宿主细胞。当以反式方式导入到宿主细胞内时，衣壳蛋白可以通过质粒导入，该质粒含有调节衣壳蛋白在宿主细胞内表达的必须序列。当以反式方式导入到细胞内时，携带衣壳蛋白的质粒最好也携带rAAV包装所必须的其他序列，如rep序列。

[0123] 在另外一个实施方式中，宿主细胞稳定地包含适宜启动子控制下的rep序列，如上面所描述的。在这个实施方式中，rep蛋白最好位于可诱导启动子的控制之下。在另一个实施方式中，rep蛋白以反式方式提供给宿主细胞。当以反式方式导入到宿主细胞内时，rep蛋白可以通过质粒导入，该质粒含有调节rep蛋白在宿主细胞内表达所必须的序列。当以反式方式导入到细胞内时，携带衣壳蛋白的质粒最好也携带rAAV包装所必须的其他序列，如rep序列和cap序列。

[0124] 因此，在一个实施方式中，rep和cap序列位于同一条核酸分子上被转染到宿主细胞内，并且以游离体的形式稳定地存在。在另外一个实施方式中，rep和cap序列稳定地整合到宿主细胞的基因组内。还有一个实施方式，rep和cap序列在宿主细胞内的表达是瞬时的。例如，一种经这种方式转染的核酸分子含有一个5’到3’的启动子，选择地添加一个插入片段，该片段位于启动子和rep基因、AAV rep基因和AAV cap基因的起始位点之间。

[0125] 另外，含有rep和/或cap序列的载体还可以含有其他需要被导入到宿主细胞内的序列。例如，载体可以包括带有小基因的rAAV构建体。载体可以含有一个或多个编码辅助功能的基因，如腺病毒蛋白E1，E2a，和E4ORF6，以及VAI RNA基因。

[0126] 用于这个构建体中的启动子可以是本领域技术人员所熟知的或上面所描述的任何组成启动子、可诱导启动子或天然启动子。在一个实施方式中使用的是AAV P5启动子。选择那种AAV来提供这些序列不是对本发明的限制。

[0127] 在另外一个优选实施方式中，rep的启动子是一种可诱导启动子，如上面所描述的那样与转基因调节元件相连接。控制rep表达的一个优选启动子是T7启动子。含有cap基因和T7启动子控制的rep基因的载体转染或转化到细胞内可以表达T7多聚酶。见WO98/10088，出版日1998年3月12日。

[0128] 插入片段在载体的设计中是一个可选择的元件。插入片段是一个插在启动子和rep基因ATG起始位点之间的DNA序列。可以按照需要任意设计，就是说它可以是随机序列的核苷酸，也可以是编码基因产物如标记基因的序列。插入片段可以含有带起始/终止位点和polyA位点的基因。插入片段可以是来源于原核细胞或真核细胞的非编码DNA序列、重复的非编码序列、无转录控制的编码序列或有转录控制的编码序列。插入片段来源的两个例子

是λ噬菌体梯级序列和酵母梯级序列，这两个插入片段可以买到，如从Gibco或Invitrogen等公司购买。插入片段可以是任何长度的，只要它能够足以降低rep78和rep68基因的表达，使rep52、rep40和cap基因达到正常表达水平。因此，插入片段的长度范围可以从约10bp到约10.0kb，优选约100bp到约8.0kb。为了降低重组的可能性，插入片段最好小于2kb，但是，本发明并不对此作限制。

[0129] 尽管rep和cap的分子可以瞬时存在于宿主细胞内(如通过转染)，但是最好是rep蛋白和cap蛋白中的一个或两个以及控制其表达的启动子能在宿主细胞内稳定表达，即作为游离体存在于宿主细胞内或整合到宿主细胞染色体内。构建本发明的该实施方式的方法是常规的基因工程技术或重组工程技术，如上面参考文献所描述的那些技术。虽然本说明书举例说明了一些特异的构建体，但是本领域的技术人员根据本文所描述的内容，选择插入片段、P5启动子以及其他元件，至少包括一个翻译起始信号和一个终止信号以及任选添加多聚腺苷酸位点，就可以选择和设计其他适宜的构建体。

[0130] 在本发明的另外一个实施方式中，宿主细胞可以稳定地提供rep或cap蛋白。

[0131] C.辅助功能

[0132] 包装宿主细胞还需要辅助功能才能包装本发明的rAAV。这些功能还可以由疱疹病毒提供。必须的辅助功能最好是人或非人灵长类动物腺病毒来源的，如上面所描述的那些和/或其他来源，包括美国模式培养物保藏所(ATCC)，Manassas，VA(US)。在当前的一个优选实施方式中，宿主细胞提供和/或含有E1a、E1b、E2a和/或E4ORF6基因产物。宿主细胞可以含有其他的腺病毒基因，如VAI RNA，但这些基因不是必须的。在一个优选实施方式中，宿主细胞没有其他的腺病毒基因或基因功能。

[0133] 表达E1a基因产物的腺病毒DNA是指任何编码E1a或任何功能性E1a蛋白的腺病毒序列。表达E2a基因产物的腺病毒DNA和表达E4ORF6基因产物的腺病毒DNA的定义也是一样的。另外还包括腺病毒基因或其功能片段的任何等位基因或其他修饰序列。这种修饰可以通过常规的基因工程技术或突变技术故意引入以便于以某些方式增强腺病毒的功能，也可是是天然发生的等位基因突变。这种修饰的DNA及其修饰方法是本领域技术人员所熟知的。

[0134] 腺病毒E1a，E1b，E2a和/或E4ORF6基因产物以及其他任何希望的辅助功能可以通过使其在细胞内表达的任何方式提供。编码这些产物的每个序列可以是在不同的载体上，也可以是在同一个载体上的一个或多个基因。载体可以是本领域所熟知的或本文所描述的任何载体，包括质粒、粘粒和病毒。载体导入到宿主细胞可以通过本领域所熟知的或本文所描述的任何方法实现，如转染、感染、电穿孔、脂质体转导、膜融合技术、高压DNA包被颗粒、病毒感染和原生质融合等。一个或多个腺病毒基因可以稳定地整合到宿主细胞基因组内，以游离体的形式稳定表达或瞬时表达。基因产物可能全部是瞬时表达的，也可能全部是在游离体内表达的或稳定整合的，还可能是一部分基因产物稳定表达而另一部分基因产物瞬时表达。而且，每个腺病毒基因的启动子都是独立选择的，可以是组成启动子、可诱导启动子或天然腺病毒启动子。启动子可以被生物体或细胞的特殊生理状态调节(如分化状态或复制和静止状态)，或者被外源添加的因子所调节。

[0135] D.宿主细胞和包装细胞系

[0136] 宿主细胞本身可以来源于任何生物体，包括原核细胞(如细菌)、真核细胞，如昆虫细胞、酵母细胞和哺乳动物细胞。特别优选的宿主细胞选自任何哺乳动物，包括而不限于

A549,WEHI,3T3,10T1/2,BHK,MDCK,COS1,COS7,BSC1,BSC40,BMT10,VERO,WI38,HeLa,293细胞(可以表达功能性的腺病毒E1),Saos,C2C12,L细胞,HT1080,HepG2以及来源于哺乳动物包括人、猴、小鼠、大鼠、兔和仓鼠的原代成纤维细胞、肝细胞和成肌细胞。提供宿主细胞的哺乳动物物种不是对本发明的限制,哺乳动物细胞的类型也不是对本发明的限制,如成纤维细胞、肝细胞、肿瘤细胞等。宿主细胞最好不含除E1、E2a和/或E4OR F6之外的任何腺病毒基因,也不含在rAAV产生过程中能与污染病毒发生同源重组的任何其他病毒的基因;宿主细胞能感染或转染DNA并能使转染的DNA表达。在一个优选实施方式中,宿主细胞是稳定转染rep和cap的细胞。

[0137] 一种用于本发明的细胞是稳定转化编码rep和cap序列的宿主细胞,该细胞转染了E1、E2a和/或E4OR F6以及携带上述小基因的构建体。稳定表达rep和/或cap的细胞系,如B-50(PCT/US98/19463)或那些美国专利No.5,658,785所描述的细胞也同样适用。另外一个优选的宿主细胞含有最少的腺病毒DNA,这种DNA足以表达E4ORF6。利用本发明的新型AAV rep和cap序列还可以构建出其他的细胞系。

[0138] 本发明的宿主细胞的制备方法包括选择性DNA序列的装配。这种装配可以利用常规技术。包括cDNA和基因组克隆技术,这是大家所熟知的,在Sambrook等中也有描述,利用了腺病毒和AAV基因组的重叠寡核苷酸序列,结合使用聚合酶链式反应、合成方法以及能提供所需核苷酸序列的任何其他适宜方法。

[0139] 分子(质粒或病毒)导入宿主细胞也可以用熟练技术人员所熟知的以及整个说明书中所讨论的技术进行。在一个优选实施方式中使用的是标准转染技术,如CaPO4转染或电穿孔方法,和/或通过腺病毒/AAV载体杂合体感染细胞系如人胚胎肾细胞系HEK293(一种含功能性腺病毒E1基因的人肾细胞系,该基因编码顺式作用的E1蛋白)。

[0140] 这些本领域技术人员制备的新型AAV来源的载体有益于将基因导入到所选择的细胞内并用于病人的基因治疗,因为在人群中还没有发现AAV7的中和抗体。而且,早期的研究表明在短尾猴和黑猩猩中也没有中和抗体,猕猴中AAV7的交叉反应性低于15%,这个血清型是从该物种中分离出来的。本领域的技术人员利用其熟知的技术可以很容易地制备含本文所描述的AAV7衣壳蛋白的其他rAAV病毒载体。用本发明的其他新型AAV制备的载体也具有同样的优点。

[0141] 因此,本领域的技术人员很容易理解本发明的AAV7序列适于制备这些载体及其他病毒载体,用于体外和体内的基因转移。同样,本领域的技术人员也容易选择本发明新型AAV基因组的其他片段用于各种rAAV及非rAAV载体系统。这些载体包括慢病毒、逆转录病毒、痘病毒、痘苗病毒等。选择何种病毒系统不是对本发明的限制。

[0142] 因此,本发明还涉及利用本发明的新型AAV的核酸和氨基酸序列制备的载体。这种载体有各种用途,包括转移治疗分子和用于制备疫苗。转移治疗分子优选含本发明新型AAV衣壳的重组AAV。这些载体或其他含本发明的新型AAV序列的载体也可用于疫苗治疗方案,如和细胞因子共转染或转染免疫原自身。

[0143] V.重组病毒及其应用

[0144] 利用本文所描述的方法,本领域的技术人员可以制备含本发明的新型血清型衣壳或带有用本发明的方法鉴定的AAV血清型的一个或多个新型片段的新型衣壳的rAAV。在一个实施方式中,所用的全长衣壳来自一个血清型即AAV7[SEQ ID NO:2]。在另一个实施方式

中，全长的衣壳是通过将本发明的一种新型血清型的一个或多个片段与其他AAV血清型的序列融合在一个框架内而制备的。例如，一种rAAV可以含有本发明的一种AAV血清型的一个或多个新型超变区。另外，本发明的独特AAV血清型也可用于含其他病毒或非病毒序列的载体中。

[0145] 本领域的技术人员很容易理解包含本发明的某种特定血清型的实施方式特别适用于某种特定的目的。例如，本发明AAV7衣壳来源的载体特别适用于肌肉中，而本发明的rh.10(44-2)衣壳来源的载体特别适用于肺。这种载体的使用没有如此的限制，本领域的技术人员也可将这些载体用于其他类型的细胞、组织或器官中。而且，本发明的其他衣壳来源的载体也可用于转染这些及其他细胞、组织或器官。

[0146] A.转基因的传递

[0147] 本发明的另外一个方面涉及向宿主传递转基因的方法，包括用本发明的AAV序列制备的载体转染或感染所选择的宿主细胞。基因转移方法是本领域技术人员所熟知的，并不代表对本发明的限制。

[0148] 在一个优选实施方式中，本发明提供了一种用AAV介导的转移方法将转基因导入到宿主中的方法。该方法包括用重组病毒载体转染或感染所选择的宿主细胞，该载体含有控制序列控制下的选择转基因，控制序列可以调节转基因和AAV衣壳蛋白的表达。

[0149] 宿主来源的样品还可以选择地分析其是否存在所选AAV血清型的抗体。各种检测中和抗体方法是本领域技术人员所熟知的。分析方法的选择不意味着对本发明的限制。见Fisher等，Nature Med.，3(3)：306-312(1997.3)和W.C.Manning等，Human Gene Therapy，9:477-485(1998.3.1)。分析结果可用于确定含一种特定血清型衣壳蛋白的AAV载体是否适于基因转移，即是否不存在该血清型衣壳特异性的抗体。

[0150] 在本方法的一个方面，具有所选AAV衣壳蛋白的载体的转移可以在通过含不同AAV血清型衣壳蛋白的载体转移一个基因之前或之后进行。因此，通过rAAV载体转移基因的方法可以向一个宿主细胞中重复转移基因。后转移的rAAV载体可以携带与第一个rAAV载体相同的基因，但其所含有的衣壳蛋白和第一个载体的衣壳蛋白可以具有不同的血清型来源。例如，如果第一个载体具有AAV7衣壳蛋白[SEQ ID NO:2]，后转移的载体所含有的衣壳蛋白可以选自其他血清型，如AAV1，AAV2，AAV3A，AAV3B，AAV4，AAV6，AAV10，AAV11，和AAV12，以及本文所鉴定的任何其他新型AAV衣壳，包括而不限于A3.1，H2，H6，C1，C2，C5，A3-3，A3-7，A3-4，A3-5，3.3b，223.4，223-5，223-10，223-2，223-7，223-6，44-1，44-5，44-2，42-15，42-8，42-13，42-3A，42-4，42-5A，42-1B，42-5B，43-1，43-12，43-5，43-21，43-25，43-20，24.1，42.2，7.2，27.3，16.3，42.10，42-3B，42-11，F1，F5，F3，42-6B，和/或42-12。

[0151] 上述重组载体可用文献报道的方法导入到宿主细胞内。RAAV最好凝胶、悬浮于生理相容性的介质中，然后注射给患病的人或非人哺乳动物。本领域的技术人员可以根据转移的病毒来选择适宜的介质。例如，适宜的介质包括盐溶液，可以用各种缓冲液(如磷酸盐缓冲液)混和制成制剂。其他的介质包括无菌盐溶液、乳糖溶液、蔗糖溶液、磷酸钙溶液、葡萄糖、琼脂、果胶、花生油、芝麻油、以及水。介质的选择不意味着对本发明的限制。

[0152] 除了rAAV和介质外，本发明的组合物还可以含有其他常规的药用组分，如防腐剂或化学稳定剂。适宜的防腐剂包括三氯叔丁醇、山梨酸钾、山梨酸、二氧化硫、没食子酸丙酯、对羟基苯甲酸酯、乙基香草醛、甘油、苯酚以及对氯酚。适宜的化学稳定剂包括凝胶和白

蛋白。

[0153] 病毒载体应以足够的量转染细胞才能得到足够高的转染效率和表达水平，这样才能获得无意外副作用的治疗效果，或达到可预期的生理效应，医学领域的技术人员可以确定这些效果。常规的和药用的给药途径包括而不限于直接注射到所选器官(如静脉灌注到肝脏)、口服、吸入(包括鼻内吸入和气管内吸入)、滴眼、静脉注射、肌肉注射、皮下注射、真皮内注射以及其他系统给药途径。如果需要也可以联合使用几种给药途径。

[0154] 病毒载体的剂量主要根据所患疾病的种类、病人的年龄、体重和健康状况等因素确定，不同的病人剂量也不同。例如，病毒载体治疗人的有效剂量一般约在1ml到100ml溶液之间，所含有的病毒数约为1×10^{9}-1×10^{16}。优选的人体剂量约为10^{13}-1×10^{16}AAV基因组。根据治疗效果及副作用的平衡可以调整剂量，这种剂量根据重组病毒所要达到的治疗目的不同而不同。可以监测转基因的表达水平来确定给药的频率，优选含有小基因的AAV载体。也可以用与治疗方案相同的剂量方案进行免疫。

[0155] 本发明的含AAV的载体转染进行治疗或免疫的例子在下面有描述。这些载体可用于各种治疗方案或疫苗免疫方案，如本文所描述。另外，这些载体还可以和所需治疗和/或疫苗方案中的一种或多种其他载体或活性组分一起给药。

[0156] B.治疗性转基因

[0157] 转基因编码的治疗产物包括激素和生长因子及分化因子，其中包括而不限于胰岛素、胰高血糖素、生长激素(GH)、甲状旁腺激素(PTH)、生长激素释放因子(GFR)、促性腺激素(FSH)、黄体化激素(LH)、人绒毛膜促性腺激素(hCG)、血管内皮生长因子(VEGF)、促进血管生成素、血管抑制素、粒细胞集落刺激因子(GCSF)、红细胞生成素(EPO)、结缔组织生长因子(CTGF)、碱性成纤维细胞生长因子(bFGF)、酸性成纤维细胞生长因子(aFGF)、表皮生长因子(EGF)、转化生长因子α(TGFα)、血小板源性生长因子(PDGF)、胰岛素样生长因子I和I(IGF-I和IGF-II)、转化生长因子β超家族的任一成员、包括TGFβ、苯丙酸诺龙、抑制素、骨形态发生蛋白(BMP)BMPs1-15的任意一个、生长因子的调蛋白/神经调节蛋白/ARIA/neu分化因子(NDF)家族的任意一个、神经生长因子(NGF)、脑源性神经生长因子(BDNF)、神经营养因子NT-3和NT-4/5、睫状神经营养因子(CNTF)、神经胶质细胞源性神经营养因子(GDNF)、neurturin、集聚蛋白、脑信号蛋白/瓦解蛋白家族的任意一员、导蛋白-1和导蛋白-2、肝细胞生长因子(HGF)、肝配蛋白、头蛋白、sonic hedgehog和酪氨酸水解酶。

[0158] 其他有用的转基因产物包括能调节免疫系统的蛋白质，其中包括而不限于细胞因子和淋巴因子，如血小板生成素(TPO)、白细胞介素(IL)IL-1到IL-25(包括IL-2，IL-4，IL-12，和IL-18)、单核细胞趋化蛋白、白血病抑制因子、粒细胞巨噬细胞集落刺激因子、Fas配基、肿瘤坏死因子α和β、干扰素α、β和γ、干细胞因子、flk-2/flt3配基。免疫系统产生的基因产物也可用于本发明。其中包括而不限于免疫球蛋白IgG、IgM、IgA、IgD和IgE、嵌合的免疫球蛋白、人源化的抗体、单链抗体、T细胞抗体、I类和II类MHC分子、以及基因工程改造的免疫球蛋白和MHC分子。有用的基因产物还包括补体调节蛋白，如补体调节蛋白、膜辅因子蛋白(MCP)、加速衰变因子(DAF)、CR1、CF2和CD59。

[0159] 其他有用的基因产物还包括激素、生长因子、细胞因子、淋巴因子、调节蛋白及免疫系统蛋白的任一受体。本发明包括胆固醇调节素的受体，包括低密度脂蛋白(LDL)受体、高密度脂蛋白(HDL)受体、极低密度脂蛋白(VLDL)受体、以及清道夫受体。本发明包括的基

因产物还有类固醇激素受体超家族的成员，包括糖皮质激素受体和雌激素受体、维生素D受体和其他核受体。另外，有用的基因产物还包括转录因子，如jun、fos、max、mad、血清反应因子(SRF)、AP-1、AP2、myb、MyoD和肌细胞生成蛋白、ETS-盒包含蛋白、TFE3、E2F、ATF1、ATF2、ATF3、ATF4、ZF5、NFAT、CREB、HNF-4、C/EBP、SP1、CCAAT-盒结合蛋白、干扰素调节因子(IRF-1)，肾母细胞瘤蛋白、ETS-结合蛋白、STAT、GATA-盒结合蛋白、即GATA-3、以及翼状螺旋蛋白的forkhead家族。

[0160] 其他有用的基因产物包括氨基甲酰合成酶I、鸟氨酸氨甲酰基转移酶、精氨酸琥珀酰合成酶、精氨酸琥珀酰裂解酶、精氨酸酶、延胡索酸乙酰乙酸水解酶、苯丙氨酸水解酶、α-1抗胰蛋白酶、葡萄糖-6-磷酸酶、胆色素原脱氨基酶、因子VIII、因子IX、胱硫醚β合成酶、支链酮酸脱羧酶、白蛋白、异戊酰辅酶A脱氢酶、丙酰辅酶A羧化酶、甲基丙二酸单酰辅酶A变位酶、戊二酰辅酶A脱氢酶、胰岛素、β-葡萄糖苷酶、丙酮酸羧化酶、肝磷酸化酶、磷酸化酶激酶、甘氨酸脱羧酶、H-蛋白、T-、蛋白、囊性纤维化跨膜调节因子(CFTR)序列、以及营养不良素cDNA序列。其他有用的基因产物还包括酶如用于替代治疗的酶，可用于各种酶活性缺乏所导致的疾病。例如，含甘露糖-6-磷酸盐的酶可用于治疗溶酶体贮积病(如编码β-葡萄糖醛酸酶(GUSB))。

[0161] 其他有用的基因产物包括非天然的多肽，如具有非天然氨基酸序列的嵌合多肽或杂合多肽，其中有氨基酸的插入、删除或替代。例如，人工单链免疫球蛋白可用于治疗某些免疫缺陷病人。其他类型的非天然基因序列包括反义分子和催化核酸，如核酶，可用于降低靶基因的过量表达。

[0162] 降低和/或改变基因的表达特别适于治疗过度增殖的疾病，这些疾病的特征是细胞的过度增殖，如肿瘤和鳞癣。靶多肽包括那些过度表达或高水平表达的多肽，这些多肽在过度增殖的细胞内比正常细胞内表达水平高许多。靶抗原包括编码癌基因的多肽，如myb，myc，fyn，和易位基因bcr/abl，ras，src，P53，neu，trk以及EGRF。除了作为靶抗原的癌基因产物以外，用于抗肿瘤基因治疗和保护性治疗措施的靶多肽还有B细胞淋巴瘤产生的各种抗体的可变区以及T细胞淋巴瘤的T细胞受体的可变区，在某些实施方式中，这些可变区也可作为靶抗原用于治疗自身免疫疾病。其他一些肿瘤相关多肽也可以作为靶多肽，如在肿瘤细胞内高水平表达的多肽，其中包括可被单克隆抗体17-A识别的多肽和叶酸结合多肽。

[0163] 其他适宜的治疗多肽和蛋白还包括那些可用于治疗自身免疫病患者和患有由抗靶位的广泛的保护性免疫反应而引起的疾病患者的多肽和蛋白，这种保护性免疫反应与细胞受体的自身免疫及产生自身免疫抗体的细胞有关。T细胞介导的自身免疫疾病包括类风湿关节炎(RA)、多发性硬化症(MS)、Sjögren综合征、结节病、胰岛素依赖型糖尿病(IDDM)、自身免疫性甲状腺炎、反应性关节炎、强直性脊柱炎、硬皮病、多肌炎、皮肌炎、鳞癣、脉管炎、Wegener肉芽肿病、节段性肠炎、以及溃疡性结肠炎、每种疾病的特征都是T细胞受体与内源性的抗原结合从而激活与自身免疫疾病有关的炎症级联反应。

[0164] C.免疫原性转基因

[0165] 另外，本发明的载体还含有本发明的AAV序列和编码肽、多肽和蛋白质的转基因，这些肽、多肽和蛋白质能够诱导对靶免疫原的免疫反应。例如，免疫原可以选择各种病毒家族。其中包括细小核糖核酸病毒家族，如鼻病毒类，占感冒致病因素的50％；肠道病毒类，包括脊髓灰质炎病毒、柯萨奇病毒、埃可病毒、以及人肠道病毒如甲型肝炎病毒；以及

apthovirus属，它是引起非人动物口蹄疫的主要病毒。细小核糖核酸病毒家族成员内的靶抗原包括VP1、VP2、VP3、VP4和VPG。另外一个病毒家族是calcivirus家族，包括Norwalk族病毒，是流行性肠胃炎的主要致病病毒。诱导人和非人动物体内免疫反应的靶抗原所针对的病毒家族还有披膜病毒家族，包括甲病毒类，如辛德毕斯病毒、RossRiver病毒、委内瑞拉马脑炎病毒、东方马脑炎病毒、西方马脑炎病毒及风疹病毒属的风疹病毒。黄病毒家族包括登革热病毒、黄热病病毒、日本脑炎病毒、圣-路易斯脑炎病毒和蜱脑炎病毒。其他的靶抗原可以从乙型肝炎病毒和冠状病毒家族中制备，其中包括多种非人源病毒，如感染性支气管炎病毒(家禽)、猪传播性胃肠道病毒(猪)、猪凝血性脑脊髓炎病毒(猪)、猫感染性腹膜炎病毒(猫)、犬冠状病毒(狗)以及人呼吸道冠状病毒，该病毒可以引起普通感冒和/和非甲、非乙非丙型肝炎。冠状病毒家族来源的靶抗原包括E1(也称M或基质蛋白)、E2(也称S或Spike蛋白)、E3(也称HE或hemagglutin-elterose)糖蛋白(不是所有的冠状病毒都存在)或N(核壳体)。其他的抗原可能是弹状病毒家族的靶位，其中包括水泡性病毒属(如水泡口炎病毒)和普通的狂犬病病毒(如狂犬病)。在弹状病毒家族内适宜的抗原可能来自G蛋白或N蛋白。线状病毒科包括出血热病毒如Marburg和Ebola病毒，也是适宜的抗原来源。副黏液病毒家族包括1型副流感病毒、3型副流感病毒、牛3型副流感病毒、rubulavirus(流行性腮腺炎病毒、2型副流感病毒、4型副流感病毒)、新城疫病毒(鸡)、牛疫病毒、麻疹病毒，包括麻疹和犬瘟热、肺病毒，包括呼吸道合胞体病毒。流感病毒属于正粘病毒类，也是合适的抗原来源(如HA蛋白、N1蛋白)。布尼亚病毒家族包括布尼亚病毒属(加利福尼亚脑炎病毒，La Crosse)、白蛉病毒属(立夫特山谷热)、汉他病毒属(puremala是一种hemahagin fever病毒)、内罗病毒属(内罗毕羊病)以及各种未分类的bungaviruses。沙粒病毒家族的LCM和拉萨热病毒也是抗原的来源之一。呼肠孤病毒家族包括呼肠孤病毒属、轮状病毒(引起儿童的急性胃肠炎)、环状病毒、cultivirus(科罗拉多蜱热，Lebombo(人)，马脑炎，blue tongue)。

[0166] 逆转录病毒家族包括oncorivirinal亚科，其中有引起人或兽疾病的猫白血病病毒、HTLVI和HTLVII、慢病毒亚科(其中包括人免疫缺陷病毒(HIV)、猿免疫缺陷病毒(SIV)、猫免疫缺陷病毒(FIV)、马感染性贫血病毒以及泡沫病毒亚科)。HIV和SIV中有许多适宜的抗原，本文已描述可供选择。适宜的HIV和SIV抗原包括而布限于gag，pol，Vif，Vpx，VPR，Env，Tat和Rev蛋白，及其各种片段。另外，本文也描述了对这些抗原的各种修饰。为此目的的适宜抗原是本领域技术人员所熟知的。例如，可以选择编码gag，pol，Vif，和Vpr，Env，Tat和Rev以及其他蛋白的一个序列。见美国专利5,972,596所描述的经修饰的gag蛋白。也可参见D.H.Barouch等，J.Virol.，75(5)：2462-2467(2001.3)，和R.R.Amara等，Science，292：69-74(2001.4.6)所描述的HIV和SIV蛋白。这些蛋白及其亚单位可以单独转导，也可以通过不同载体或单一载体联合转导。

[0167] 乳头多瘤空泡病毒家族包括多瘤病毒亚科(BKU和JCU病毒)和乳头状瘤病毒亚科(与肿瘤或乳头状瘤的恶性变有关)。腺病毒家族包括病毒(EX，AD7，ARD，O.B.)，主要引起呼吸道疾病和/或肠炎。细小病毒家族包括猫细小病毒(猫肠炎)、猫panleucopeniavirus、犬细小病毒和猪细小病毒。疱疹病毒家族包括甲型疱疹病毒亚科，其中有单纯病毒属(HSVI，HSVII)、水痘病毒属(伪狂犬病、水痘、带状疱疹)和β疱疹病毒亚属，其中包括巨细胞病毒属(HCMV、鼠巨细胞病毒)，以及丙型疱疹病毒亚科，其中包括淋巴滤泡病毒亚科、EBV(伯基特淋巴瘤)、感染性鼻气管炎、Marek病病毒以及rhadinovirus。痘病毒家族包括

chordopoxvirinae亚属，其中有正痘病毒属(天花和牛痘)、副痘病毒、禽痘病毒、羊痘病毒、兔痘病毒、猪痘病毒以及昆虫痘病毒亚科。肝DNA病毒家族包括乙型肝炎病毒。一种未分类的可作为抗原来源的病毒是δ肝炎病毒。其他的病毒包括禽感染性粘液囊病病毒和猪呼吸及生殖综合征病毒。甲病毒家族包括马的动脉炎病毒和各种脑炎病毒。

[0168] 本发明还包括免疫原，这些免疫原用于免疫人或非人动物以抵抗病原如细菌、真菌、寄生的微生物或寄生虫，这些病原可感染人或非人脊椎动物，或者来自与癌症或肿瘤细胞。细菌病原包括致病性革兰氏阳性球菌如肺炎球菌、葡萄球菌和链球菌；致病性的肠革兰氏阴性球菌如肠内菌科(enterobacteriaceae)；脑膜炎球菌、淋球菌。致病性的肠革兰氏阴性杆菌如假单胞菌；acinetobacteria；艾肯氏菌；类鼻疽；沙门氏菌；志贺氏菌；嗜血杆菌；莫拉杆菌；杜克雷嗜血杆菌(H.ducreyi)(引起软下疳)；布鲁氏菌；Franisella tularensis(兔热病)；耶尔森菌(巴斯德菌)；链杆菌属的念球棘球菌和螺旋菌；革兰氏阳性杆菌包括单核细胞增多性李斯特杆菌；猪红斑丹毒丝菌；白喉棒状杆菌(白喉)；霍乱；炭疽杆菌(炭疽)；腹股沟肉芽肿(腹股沟肉芽肿)；以及杆菌状巴尔通体。由致病性厌氧菌引起的疾病包括破伤风；肉毒中毒；其他梭菌引起的疾病；结核杆菌；麻风；以及其他分枝杆菌。致病性螺旋菌引起的疾病包括梅毒；treponematoses；印度痘；南美锥虫病和地方性梅毒；以及钩端螺旋体病。其他由高等致病性细菌和致病性真菌引起的感染包括放线菌病；诺卡放线菌病；隐球菌病；酿母菌病；荚膜组织胞浆菌病；球孢子菌病；念珠菌病；曲霉病；毛霉菌病；孢子丝菌病；paracoccidiodomycosis，petriellidiosis，torulopsosis，足分枝菌病和着色真菌病；以及脚气。立克次体感染包括斑疹伤寒、落基山斑疹热、Q热和立克次体热。支原体和衣原体感染包括支原体肺炎；性病淋巴肉芽肿；鹦鹉热；以及围产期的衣原体感染。致病性真核生物包括致病性的寄生虫和蠕虫及其引起的感染，如阿米巴病；疟疾；利什曼病；锥形虫病；弓形体病；卡氏肺孢子虫病；Trichans；刚地弓形虫(Toxoplasma gondii)；巴贝西虫病；梨形鞭毛虫病；毛线虫病；丝虫病；血吸虫病；线虫病；吸虫病或吸虫；以及绦虫感染。

[0169] 这些生物中的许多及其产生的毒素已被疾病控制中心[(CDC)，健康和人类服务部(Department of Heath and Human Services)，USA]所鉴定，其中有些有可能用作生物攻击的试剂。例如，其中包括炭疽杆菌(炭疽)、肉毒梭菌及其毒素(肉毒中毒)、鼠疫杆菌(鼠疫)、重型天花(天花)、土拉热弗朗西丝菌(土热病)、以及病毒性出血热，所有这些现在都被归为A类因子；伯氏考克斯氏体(Q热)、布鲁氏菌属(布鲁氏菌病)、鼻疽假单胞菌(鼻疽)、蓖麻及其毒素(蓖麻蛋白)、产气荚膜梭菌及其毒素(ε毒素)、葡萄球菌属及其毒素(肠毒素B)，所有这些现在都被归为B类因子；Nipan病毒和汉他病毒现在被归为C类因子。另外，其他如此分类或未如此分类的其他微生物将来也可能被鉴定并被用于此目的。很容易理解病毒载体和本文所描述的其他构建体可用于转移这些生物体、病毒、其毒素或其他副产物的抗原，预防和/或治疗由这些生物因子所引起的感染或其他副作用。

[0170] 通过本发明的载体转运抗T细胞可变区的免疫原可以激发包括CTL在内的免疫反应以清除这些T细胞。在类风湿关节炎(RA)中发挥作用的T细胞受体(TCRs)上的几个特异性可变区已被鉴定，这些TCRs包含V-3，V-14，V-17和Vα-17。因此，转导编码至少一个这种多肽的核酸序列可以激发针对RA中T细胞的免疫反应。在多发性硬化症(MS)中发挥作用的T细胞受体(TCRs)上的几个特异性可变区已被鉴定，这些TCRs包含V-7和Vα-10。因此，转导编码至少一个这种多肽的核酸序列可以激发针对MS中T细胞的免疫反应。在硬皮病中发挥作用的T

细胞受体(TCRs)上的几个特异性可变区已被鉴定,这些TCRs包含V-6,V-8,V-14,Va-16,Va-3C,Va-7,Va-14,Va-15,Va-16,Va-28和Va-12。因此,转导编码至少一个这种多肽的核酸序列可以激发针对硬皮病中T细胞的免疫反应。

[0171] 另外,本发明含AAV序列的载体还可用基础-加强方案(prime-boost regimen)转移。本领域已报道了许多这样的方案,很容易选择。见WO00/11140,出版日2000年3月2日,本文已纳入作为参考。

[0172] 这种基础-加强方案一般包括先给予DNA(如质粒)载体以初步激发免疫系统,然后通过给予传统的抗原如蛋白质或携带编码该抗原的序列的重组病毒来加强这种免疫反应。在一个实施方式中,本发明提供了一种基础-加强免疫方案激发针对所选择抗原的免疫反应,该方案先转移携带所述抗原的一个质粒DNA载体,然后进行加强,就是在转移含本发明AAV序列的载体。

[0173] 在一个实施方式中,基础-加强方案包括基础和/或加强载体内多蛋白的表达。见R.R.Amara,Science,292:69-74(6April2001),其中描述了表达蛋白亚单位的多蛋白方案用于激发抗HIV和SIV的免疫反应。例如,一个DNA基础载体可以用一个转录子转移Gag,Pol,Vif,VPX和Vpr和Env,Tat,和Rev。另外,SIV的Gag,Pol和HIV-1的Env也被转移。

[0174] 但是,基础-加强方案并不限于HIV的免疫或转移这些抗原。例如,基础免疫是转移本发明的第一黑猩猩载体,然后用第二黑猩猩载体或含蛋白形式的抗原的组合物进行加强免疫。在一个实施例中,基础-加强免疫方案可以提供对抗原所来源的病毒、细菌或其他微生物的保护性免疫反应。在另一个优选实施方式中,基础-加强免疫方案可以达到治疗效果,利用常规的分析方法可以检测到这种效果。

[0175] 基础疫苗可以剂量依赖的方式注射到身体的各种部位,主要根据所要激发的免疫反应的靶抗原。本发明不限制注射的量或位点以及药用载体。基础免疫步骤的治疗方案可以是单一剂量,也可以是每小时、每天、每周、每月或每年的剂量。例如,哺乳动物可以接受1-2次注射,注射剂量为约含10μg到50μg质粒的载体。优选的基础免疫量或基础DNA疫苗组合物的剂量约为1μg到10,000μg DNA疫苗。剂量也可以为每公斤体重约1μg到1000μg DNA。注射的量或位点的选择要根据免疫动物的同一性和身体状况。

[0176] 本文描述了适于将抗原转移到哺乳动物体内所需DNA疫苗的剂量单位。用于注射的DNA疫苗应悬到或溶解到药用介质或生理介质中,如等渗盐、等渗盐溶液或其他制剂,这些都是药物注射人员所熟知的。本领域的技术人员很清楚哪些是合适的介质,这主要是根据注射的途径来选择。本发明的组合物可以通过上述的途径注射,用生物降解的生物相容性聚合物制成的缓释制剂的形式,或用微胶粒、凝胶和脂质体制成的在位注射制剂的形式注射。

[0177] 另外,本发明的初级免疫步骤还包括注射初级DNA疫苗组合物、适宜量的佐剂,如本文所定义的。

[0178] 对于哺乳动物来说,加强免疫最好在注射初级DNA疫苗后约2到27周进行。加强免疫组合物的折射是用有效量的加强疫苗组合物,该组合物含有或能够转移基础DNA疫苗所给予的同一抗原。加强免疫组合物由来源于同一病毒或不同病毒的重组病毒载体组成。另外,“加强免疫组合物”可以含有与基础DNA疫苗所编码抗原相同的抗原,只是该抗原是以蛋白或肽的形式存在,该组合物可以诱导宿主的免疫反应。在另外一个实施方式中,加强免疫

组合物所包含的一个组合物含有一种DNA序列，该序列编码调节序列控制下的抗原，调节序列可以控制其在哺乳动物细胞内的表达，及载体如已知的细菌或病毒载体。加强免疫组合物的一个基本要求是疫苗组合物中的抗原要与DNA疫苗所编码的抗原相同或是交叉反应抗原。

[0179] 本发明的载体也适用于含非AAV载体或蛋白、肽或其他生物治疗或免疫化合物的免疫方案中。这些方案特别适于通过基因转移进行治疗和免疫，包括诱导保护性免疫。这些应用是本领域技术人员所熟知的。

[0180] 本发明的载体还是有效的基因治疗工具，它可以在体内或体外将所选择的转基因导入到所选择的宿主细胞中，即使生物体内含有一种或多种AAV血清型的中和抗体。在一个实施方式中，载体(即rAAV)和细胞可以在体外混合；用常规方法培养感染的细胞；然后再将转导的细胞注入病人体内。而且，本发明的载体还可用于在体外制备所需要的基因产物。对于体外制备来说，先用含编码目的产物的分子的rAAV转染宿主细胞，在合适其表达的条件下培养细胞，然后就可以从细胞培养物中得到所需要的产物(如蛋白质)。然后将表达产物纯化分离。适宜的转染、细胞培养、纯化和分离技术都是本领域技术人员所熟知的。

[0181] 下面的实施例解释了本发明的几个方面和实施方式。

[0182] 实施例

[0183] 实施例1：新型AAV序列的PCR扩增、克隆和特征鉴定

[0184] 基于已知AAV高度保守区的寡核苷酸设计引物用PCR方法筛选非人灵长类动物组织内的AAV序列。跨越AAV1[SEQ ID NO:6]2886-3143bp的一段AAV序列作为PCR扩增子，其中衣壳蛋白(Cap)上的超变区与其他所有已知的AAV血清型都不同，本文命名为“特征区域”，其两侧是保守序列。经过后来的分析发现这个特征区域位于超变区3的保守残基之间。

[0185] 通过对一组非人灵长类动物外周血样品的初步筛查发现能检测到AAV的灵长类亚类是猕猴、短尾猴、黑猩猩和狒狒。但是，在检测的某些其他亚类的动物中没有发现AAV，其中包括日本猿、猪尾猴和松鼠猴。通过对宾夕发尼亚大学所饲养的猕猴的组织进行了载体分布的集中分析，并对动物尸体进行解刨发现了AAV序列在组织内的整体分布状况。

[0186] A.AAV特征区域的扩增

[0187] AAV1-6以及从鹅和鸭体内分离的AAV的DNA序列用“Clustal W”以默认设置彼此比对。图1显示的是AAV1-6的比对序列及AAV7的信息。AAV序列的相似性作了比较。

[0188] 在一组研究中，一个跨越AAV1[SEQ ID NO:6]2886bp-3143bp的257bp的区域以及AAV2-6基因组上相应区域[见图1]被发明者鉴定。该区域位于AAV衣壳基因上，其5’末端和3’末端都是高度保守的，中间区域是相对可变的。另外，该区域含有一个DraIII限制性酶切位点(CACCACGTC，SEQ ID NO:15)。发明者发现该区域可以作为已知的所有类型AAV DNA的特异特征区域。换言之，经过PCR反应、已知血清型特异性的内切酶消化和凝胶电泳，这个区域可用于将扩增的DNA定义为血清型1，2，3，4，5，6或其他血清型。

[0189] 以AAV1、2和5DNA为对照设计、确证引物和优化PCR条件。引物的设计基于AAV2的序列：5’引物，1S:bp AAV2(SEQ ID NO:7)的2867-2891和3’primer，18as，AAV2(SEQ ID NO:7)的bp3095-3121。

[0190] 不同猕猴的不同组织如血液、脑、肝、肺、睾丸等来源的细胞DNA用上述方法分析，结果显示从这些猴子不同组织的DNA中都可以得到很强的PCR扩增产物。PCR产物的限制性

酶切分析表明这些产物是从AAV序列中扩增出来的，与已发表的所有AAV序列都不同。

[0191] 将各种猴组织DNA中扩增出的PCR产物(约255bp长)克隆和测序。对这些新型AAV序列进行生物信息学分析发现它们是新型AAV序列的衣壳基因，并且彼此不同。图1显示的AAV10-12的新型AAV特征区域[SEQ ID NO:117，118和119]。利用Clustal W(1.81)程序进行多序列比对分析。AAV1-7和AAV10-12基因组的特征区域之间的序列同一性百分率见下。

[0192] 表1.所分析的序列

[0193]

序列号	AAV血清型	片段大小(bp)
1	AAV1	258
2	AAV2	255
3	AAV3	255
4	AAV4	246
5	AAV5	258
6	AAV6	258
7	AAV7	258
10	AAV10	255
11	AAV11	258
12	AAV12	255

[0194] 表3.配对比对(同一性的百分率)

[0195]

	AAV2	AAV3	AAV4	AAV5	AAV6	AAV7	AAV10	AAV11	AAV12
AAV1	90	90	81	76	97	91	93	94	93
AAV2		93	79	78	90	90	93	93	92
AAV3			80	76	90	92	92	92	92
AAV4				76	81	84	82	81	79
AAV5					75	78	79	79	76
AAV6						91	92	94	94
AAV7							94	92	92
AAV10								95	93
AAV11									94

[0196] 一共分离和测序了300多个含新型AAV血清型的克隆，这些血清型都有所选择的257bp的区域。通过对这300多个克隆的生物信息学分析发现这个257bp的区域可以作为一个优越的标记序列用于新型AAV血清型的分离和鉴定

[0197] B.利用特征区域进行PCR扩增

[0198] 257bp的特征区域可以作为PCR锚将PCR扩增子延伸到基因组的5'末端覆盖住rep基因和cap基因(1398bp-3143bp，SEQ ID NO:6)的连接区，还可以延伸到基因组的3'末端得到完整的cap基因序列(2866bp-4600bp，SEQ ID NO:6)。利用标准方法进行PCR扩增，95°C变性0.5-1min，60-65°C退火0.5-1min，72°C延伸1min/kb，28-42个循环。

[0199] 利用"A"中所描述的比对序列，在两个序列中又发现了其他两个相对保守的区域，

其中一个序列位于rep基因的3'末端，其5'末端指向257bp序列，另一个序列位于257bp片段的下游但是位于AAV3'ITR之前。为了扩增出新型AAV血清型的完整衣壳序列和部分rep序列，设计出两对新引物并优化了PCR条件。为了提高特异性，5'末端扩增引物：5'引物为AV1Ns，GCTGCGTCAACTGGACCAATGAGAAC[AAV1的nt1398-1423，SEQ ID NO:6]，3'引物为上述的18as。3'扩增引物：5'引物为上述的1s，3'引物为AV2Las，TCGTTTCAGTTGAACTTTGGTCTCTGCG[AAV2的nt4435-4462，SEQ ID NO:7]。

[0200] 对于这些PCR扩增来说，257bp区作为PCR锚和登陆标记制备重叠片段以构建完整的衣壳基因，如本文所描述的那样扩增出5'延伸片段和3'延伸片段后通过特征区域内DraIII位点融合。为了较特异地扩增出完整的AAV7cap基因，按照说明书分别将三个扩增产物(a)特征区域序列；(b)5'延伸序列；(c)3'延伸序列克隆到pCR4-Topo[Invitrogen]质粒中。然后用DraIII消化质粒并重组形成完整的cap基因。

[0201] 在本组实验中，分离和分析了AAV7和AAV8约80％的衣壳序列。另外一个新血清型，AAV9在猴子#2也被发现。

[0202] 利用上述PCR条件，分离和克隆了AAV rep基因的剩余部分，所用的引物可以扩增出AAV基因组的108bp到1461bp的区域(基于AAV2的编号SEQ ID NO:7计算出的)。该克隆经测序后用于构建无ITR的完整AAV7基因组。

[0203] C.直接扩增3.1kb的Cap片段

[0204] 为了从NHP组织和血液DNA中直接扩增3.1kb的全长Cap片段，在AAV基因组中鉴定出两段其他的高度保守区用于大片段的PCR扩增。选择了一个位于rep基因中间保守区内引物(AV1ns:5'GCTGCGTCAACTGGACCAATGAGAAC3'，SEQ ID NO:6的nt1398-1423)和位于Cap基因下游另一保守区内的3'引物(AV2cas:5'CGCAGAGACCAAAGTT CAACTGAAACGA3'，SEQ ID NO:7)扩增全长的cap片段。按照厂家的说明书(Invitrogen)将PCR产物克隆到Topo中，用Qiagengenomics(Qiagengenomics，Seattle，WA)进行序列分析，结果显示精确率大于99.9%。总共分离和鉴定了50个衣壳序列克隆，其中37个克隆来源于猕猴(rh.1-rh.37)，6个克隆来源于短尾猴(cy.1-cy.6)，2个克隆来源于狒狒(bb.1和bb.2)，5个克隆来源于黑猩猩(ch.1-ch.5)。

[0205] 为了排除新型AAV家族内序列多样性是由于PCR错误引起的可能性，如通过DNA模板不同部分与同源序列之间重叠延伸而导致的PCR介导的基因剪接，或者细菌内重组而导致的结果，在同样的条件下又进行了一系列VP1扩增实验，所用的模板为细胞总DNA。首先，将完整AAV7和AAV8等比例混和后进行系列稀释，系列稀释的混合物作为模板PCR扩增3.1kb的VP1片段，所用的引物和扩增条件与DNA扩增相同，观察是否出现任何杂合的PCR产物。将混合物转化到细菌中，分离出转化子，观察杂合克隆是否可能来自于细菌内的重组过程。在不同的试验中，我们用Msp I，Ava I和HaeI限制性酶切AAV7和AAV8，所有的酶都可以在基因组的不同位置切多次，将不同组合的消化产物混和并用其作为模板PCR扩增VP1片段，所用的条件相同，观察是否会因为部分AAV序列的重叠延伸而产生任何的PCR产物。在另一个试验中，将凝胶纯化的与特征区域重叠的1.5kb AAV7VP15'片段和1.7kb AAV8VP13'片段混和，进行系列稀释后用于PCR扩增，体系中加入或不加200ng来源于猴细胞系的细胞DNA，该DNA经TaqMan分析确定不含有AAV序列。在基因组DNA Cap扩增的条件下所有这些试验都没有发现PCR介导的重叠序列产生(数据未列出)。为了进一步确证，设计了3组引物，分别位于

不同的HVR上，并且是猕猴F953的克隆42特异性的，用不同的引物组合从F953的肠系膜淋巴结(MLN)中扩增较短的片段，克隆42就是从F953中分离的。在全长Cap克隆中鉴定出的所有序列突变株都存在于这些短片段中(数据未列出)。

[0206] 实施例2:在自然感染过程中腺伴随病毒在灵长类动物体内经历了实质进化

[0207] 将分离的AAV进行序列分析发现整个基因组的分散性主要集中在衣壳蛋白的可变区。流行病学调查结果发现所有已知的血清型都是灵长类动物部位特异性的，尽管临床分离株只限于从婴儿的肛拭子和咽拭子中分离AAV2和AAV3，从人的湿疣中分离AAV5。还没有发现有临床并发症与AAV感染有关。

[0208] 为了更好地了解AAV的生物学特性，用非人灵长类动物为模型以观察自然感染的并发症。基于已知AAV高度保守区设计寡核苷酸引物(见实施例1)，用本发明的PCR方法从非人灵长类动物的组织中筛选AAV序列。跨越AAV1[SEQ ID NO:6]2886-3143bp的一段AAV序列作为PCR扩增子，其中保守序列位于超变区的两侧，这个超变区与其他所有已知的AAV血清型都不同，本文命名为“特征区域”。

[0209] 通过对一组非人灵长类动物外周血样品的初步筛查发现能检测到AAV的灵长类亚类是猕猴、短尾猴、黑猩猩和狒狒。通过对宾夕发尼亚大学所饲养的猕猴的组织进行了载体分布的集中分析，并对动物尸体进行解刨发现了AAV序列在组织内的整体分布状况。

[0210] 扩增的特征区域序列亚克隆到质粒中，每个转化子都进行测序。结果显示不同动物来源的克隆的核苷酸序列都有实质的变化。特征区域序列的变化在一个动物的不同克隆内也很明显。从两个具有不同标记序列的动物中取下组织进行进一步分析，利用高度保守区的寡核苷酸作为引物通过PCR方法扩增出延伸的片段。利用这个方法从新构建了两种组织来源的病毒基因组。这些原病毒与已知的其他AAV在cap基因的序列分散性上有明显不同

[0211] 下一步的试验用于确证非人类人猿动物组织内的AAV序列代表了感染病毒的原病毒基因组，并且这些原病毒基因组可以被恢复形成病毒颗粒。动物98E056肝组织的基因组DNA可以检测到AAV8特征区域序列，用一种AAV序列内不含其位点的内切酶消化该基因组DNA，并将其和一种质粒一起转染到293细胞内，该质粒含有E1缺失的人腺病毒血清型5基因组，作为辅助功能的来源。将得到的裂解产物在293细胞上传代一次，回收裂解物，利用具有多价反应性的Cap蛋白的多克隆抗体检测是否存在AAV Cap蛋白以及利用PCR方法扩增AAV8来源的AAV原病毒以检测DNA序列的丰度。通过Western印迹方法检测Cap蛋白可以证明限制性内切酶消化的心脏DNA和腺病毒辅助质粒共转染可以产生高滴度的AAV8病毒，每个293细胞中含有10^4AAV8载体基因组。大规模地制备细胞裂解物并通过氯化铯离心沉淀技术从中纯化。纯化的产物中含有26nm二十面体(icosohedral)结构，看起来与AAV血清型2一致。单独用腺病毒辅助载体转染不产生AAV蛋白或基因组，这就排除了恢复的AAV作为污染源的可能性。

[0212] 为了进一步确定动物种内及种间AAV标记序列的变异情况，挑选组织进行延伸PCR以扩增整个Cap蛋白的开放阅读框架。

[0213] 将所得到的片段克隆到细菌质粒中，分离单个转化子并测序。所取的组织包括3个猕猴的肠系膜淋巴结(Tulane/V223-6克隆;Tulane/T612-7克隆;Tulane/F953-14克隆)、2个猕猴的肝脏(Tulane/V251-3克隆;Penn/00E033-3克隆)、1个猕猴的脾脏(Penn/97E043-3克隆)、1个猕猴的心脏(IHGT/98E046-1克隆)、1个黑猩猩(New Iberia/X133-5克隆)、6个短

尾猴(Charles River/A1378,A3099,A3388,A3442,A2821,A3242-6克隆)和1个狒狒(SFRB/8644-2克隆)的血液。15个动物来源的50个克隆中,30个被认为是非多余的,因为发现其中至少有7个氨基酸是彼此不同的。非重复的VP1克隆以其被分离的顺序作为编号,前缀表示其来源的非人灵长类动物的种类。利用SplitsTree程序[Huson,D.H.SplitsTree:analyzing and visualizing evolutionary data.Bioinformatics14,68-73(1998)]加上分裂分解方法分析了这30个非重度的克隆与以前所描述的8个AAV血清型之间的结构关系。分析结果显示树状网络而不是分枝树状网络中的一组序列之间存在相似性。其优点在于能检测会聚而形成的分组及显示其系统发生关系,即使这种关系可能被平行发生的事件所逆转。因此还需要大量的系统发生研究以说明AAV的进化,本发明只是对不同克隆进行了分组以显示其序列相似性。

[0214] 为了确定新型VP1序列是感染病毒基因组来源的,从具有高丰度病毒基因组的组织中提取细胞DNA,用在AAV内无酶切位点的限制性内切酶消化并转染293细胞,然后用腺病毒感染293细胞。这样就可以从两种不同动物来源的组织DNA中恢复和扩增AAV基因组(数据未列出)。

[0215] 进一步分析新型AAV的VP1序列以确定氨基酸变异的性质和位置。将彼此氨基酸序列差异大于1%的所有30个VP1克隆比对在一起评价每个残基上变异的分数。根据一种分析序列分散区域的算法得到了12个超变区(HVR),其中5个与以前所描述的可变区重叠或是其一部分[Kotin,上面已引用;Rutledge,上面已引用]。三叠近端峰(three-fold-proximal peaks)包含了变异性的大部分(HVR5-10).令人感兴趣的是2倍轴和8倍轴上的袢也是变异较多的区域。衣壳蛋白N端有HVR1和2,X射线衍射并不能发现这两个区域,说明VP1蛋白的N端暴露在病毒颗粒表面。

[0216] 用实时PCR从21只猕猴的组织中定量检测AAV序列,所用的引物和探针是针对已知AAV Rep(1对)和Cap(2对)的高度保守区的。每个数据点代表从一个动物的一种组织中扩增的数据。结果发现AAV序列具有广泛的分布,虽然不同动物之间所含有的AAV的量不同。动物的来源及其背景和所受的处理并不影响AAV序列在猕猴内的分布。AAV表达水平最高的是在肠系膜淋巴结中,13个阳性动物中平均每个二倍体基因中含0.01个拷贝。肝和脾也含有高丰度的病毒DNA。有些组织中含有非常高拷贝的AAV,如猕猴98E056的心脏组织、猕猴97E043的脾组织以及猕猴RQ4407的肝组织,它们中每个二倍体基因组分别含有1.5,3和20拷贝的AAV序列。外周血单核细胞中病毒DNA的水平相对较低,说明组织中数据不是由于组织中残存的血液引起的(数据未列出)。应当注意的是这种方法并不能检测出非灵长类动物体内的所有AAV,因为检测出的序列需要与寡核苷酸和实时PCR探针具有高度的同源性。通过DNA杂交技术进一步分析具有高丰度AAV DNA的动物组织以确定DNA的状态,通过原位杂交来分析其在细胞内的分布情况。

[0217] 不同动物间及同一动物不同组织间AAV原病毒片段存在序列变异使人联想到许多RNA病毒在流行过程中,甚至在感染一个个体时会发生进化。在许多情况下都可以发现野生型病毒被一群所代替,这些准种在选择压力存在的情况下通过快速复制和突变而发生了进化。其中一个例子是HIV感染,这个病毒会因免疫压力或药物压力而发生进化。RNA病毒突变的高发生率有几种机制,其中包括逆转录酶的低保真度和缺乏纠错能力,以及非同源重组和同源重组。

[0218] 在本研究中可以通过多个原病毒片段克隆的系统测序来说明AAV准种形成的原因。实际上，从不同动物或同一动物内分离的任何延伸克隆中都找不到完全一致的序列。这种序列进化的机制可能是亲本病毒之间高频率的同源重组。结果导致Cap蛋白的超变区频繁交换，从而形成了一组嵌合体，这种嵌合体具有不同的小趋向性和不同的血清型特异性(即获得了逃避免疫反应的能力，特别是中和抗体的攻击)。同源重组发生的机制还不清楚。一种可能是不同单链AAV基因组的+链和-链在复制时杂交，如以前所报道的在有许多AAV重组子感染时。还不清楚是否存在其他机制导致了AAV感染时发生序列进化。AAV复制时发生的突变总频率相对较低，试验数据也不支持出现高频率的复制错误。但是，已有报道证实在裂解性感染过程中AAV基因组的重排可以导致干扰缺损颗粒的形成。不论导致序列变异的机制如何，无一例外的是准种的vp1结构还是保持完整的，没有发生移码突变和无义突变，这说明由于种群动力学的原因竞争选择使病毒具有了最佳的结构及形态组合。

[0219] 这些研究可用于生物学及医学的几个领域。快速病毒进化的概念以前被认为只是RNA病毒的特性，现在发现DNA病毒也可能发生，这一点已被血清学分析所证明。对于细小病毒来说，发展一种新方法来描述病毒分离株是十分重要的，这种描述应该能够涵盖其结构和生物特征的复杂性，如HIV，就被分类为具有相同结构和功能的一个通用家族，被称为Clades。现在正在寻找其他的策略通过血清型特异性对分离株进行分类，并提出描述血清学组群内变异的标准。

[0220] 实施例3:利用含AAV2rep基因和新型AAV cap基因的嵌合质粒构建携带AAV2ITR的重组AAV基因组载体用于不同动物模型的血清学和基因转移研究

[0221] 将AAV2rep与新型AAV血清型的cap序列融合构建嵌合的包装载体。这些嵌合包装载体最初被用于携带AAV2ITR的假型(pseudotyping)重组AAV基因组，所用的方法是利用Ad5辅助质粒三重转染293细胞。在分离出这些新型血清型的完整的有感染性的病毒之前，这些假型载体被用于评价以转导为基础的血清学研究方法的可行性及新型AAV血清型在不同动物模型包括NHP和啮齿类动物中的基因转移效率。

[0222] A.pAAV2GFP

[0223] AAV2质粒含有AAV2 ITR和组成启动子控制下的绿色荧光蛋白。这个质粒包含如下元件:AAV2 ITR、CMV启动子和 GFP编码序列。

[0224] B.反式质粒的克隆

[0225] 为了构建嵌合的反式质粒来制备重组的假型AAV7载体，p5E18质粒(Xiao等，1999，J.Virol 73:3994-4003)用Xho I部分消化以使质粒线性化，该质粒只在3169bp处有一个Xho I位点。然后将Xho I酶切末端补平然后从新连接起来。这种经改造的p5E18质粒用Xba I和Xho I完全酶切以去除AAV2 cap基因序列，并用一个含AAV7 cap 基因的2267 bp Spe I/Xho I片段代替，该片段是从pCRAAV7 6-5+15-4质粒上切下的。

[0226] 所得到的质粒含有AAV2 P5启动子控制下的 AAV2 rep序列Rep78/68和AAV2 P19启动子控制下的AAV2 rep序列Rep52/40。AAV7衣壳序列位于AAV2 P40启动子控制之下，该序列位于Rep序列内。这个质粒还含有插入片段5’ rep ORF。

[0227] C.假型 rAAV的制备

[0228] rAAV颗粒(含AAV7衣壳的AAV2载体)是用无腺病毒的方法制备的。主要步骤为，将顺式质粒(含AAV2 ITR 的pAAV2.1 lacZ质粒)、反式质粒pCRAAV7 6-5+15-4(含AAV2 rep和

AAV7 cap)以及辅助质粒同时转染293细胞,其比例为1:1:2,转染方法为磷酸钙沉淀法。

[0229] 为了构建pAd辅助质粒,从Microbix(Canada)购买了质粒pBG10 。从pBHG10中切下一个含L2和L3的RsrII片段得到第一辅助质粒pAdΔF13。从pBHG10中切下一个含PmeI/Sgf1缺失的Asp700/SalI片段克隆到Bluescript中获得质粒AdΔ F1。从pAdΔF1 中删除 MLP,L2,L2和L3。然后分别删除一个2.3 kb NruI片段和一个0.5kb RsrII/NruI片段得到辅助质粒pAdΔF5和pAdΔF6。被命名为pΔF6的辅助质粒具有E2a 和 E4 ORF6所必须的辅助功能,这是E1表达辅助细胞所不具有的,但是这个质粒不含有腺病毒衣壳蛋白和功能性的E1区。

[0230] 通常将50 μg DNA(顺式质粒:反式质粒:辅助质粒)转染到一个150 mm 组织培养皿中。转染后72小时收获293细胞,超声破碎并用0.5% 去氧胆酸纳处理(37,10min.)。细胞裂解物经两次CsCl离心沉淀。收集含rAAV载体的峰值部分,置于容器中用PBS透析。

[0231] 实施例4:制备携带完整新型AAV血清型的感染性克隆在人和NHP来源的细胞系中研究其基本的病毒学特性,并在NHP和其他动物模型中评价新型AAV血清型的致病性。

[0232] 为达到此目的,用基因组步行系统获取5'和3'末端序列(ITR)并完成含完整新型AAV血清型基因组的克隆的构建。

[0233] 简言之,利用商业化的Universal Genome Walker Kit[Clontech],将猴组织和细胞系来源的基因组DNA用Dra I,EcoR V,Pvu II和Stu I限制性内切酶消化并连接到Genome Walker Adaptor上得到4个个体的Genome Walker Libraries(GWLs),这些基因组DNA中已被证实含有AAV7序列。已GWLs的DNA为模板,PCR扩增出AAV7及其相邻的基因组序列,所用的引物为衔接子引物1(AP1,试剂盒提供)和AAV7特异性引物1,然后再用衔接子引物2(AP2)和另一个AAV7特异性引物2进行嵌套PCR,这两个引物都在第一组引物内。将嵌套PCR所得到的主要产物克隆并进行序列分析。

[0234] 在这个试验中,涵盖257bp片段或AAV基因组其他标记片段的引物被用于PCR扩增从人和NHP细胞系中提取的细胞DNA以鉴定和鉴别是否存在AAV序列。利用不同物种和不同株的辅助腺病毒可以从阳性细胞系中恢复这种潜伏的AAV基因组。

[0235] 为了从NHP来源的细胞系内分类具有感染性的AAV克隆,从ATCC获得所需的细胞系,然后用PCR方法筛选以鉴定257bp的扩增子,即本发明的特征区域。将257bp的PCR产物克隆并通过测序分析其血清型。这些含AAV7序列的细胞系用从ATCC购买的猿猴腺病毒SV-15、人Ad5感染,或者用含人Ad基因的质粒转染,这个Ad基因编码AAV辅助功能。感染或转染48小时猴,收获细胞,按照Xiao等,1999,J.Virol,73:3994-4003的描述制备Hirt DNA用于AAV7基因组的克隆。

[0236] 实施例5:AAV载体的构建

[0237] 利用与实施例3中AAV1/7所采取的同样的假型策略构建包装有AAV1,AAV5和AAV8衣壳蛋白的AAV2载体。简要叙述如下,通过顺式质粒、腺病毒辅助质粒和嵌合包装载体三重转染293细胞包装携带AAV2ITR的重组AAV基因组,其中AAV2rep基因与新型AAV血清型的cap基因融合在一起。为了构建嵌合的包装载体,去除质粒p5E18上3169bp处的Xho I位点,修饰的质粒再用Xba I和Xho I完全酶切以去除AAV2cap基因,并用一个含AAV8cap基因的2267bp Spe I/Xho I片段代替[Xiao,W.等,(1999)J Virol73,3994-4003]。用同样的策略构建AAV2/1和AAV2/5的嵌合包装质粒。除了AAV2/2外,所有的重组载体都用标准CsCl2沉降法纯化,AAV2/2是用肝素色谱柱通过一步纯化的。

[0238] AAV载体在基因组中的拷贝(GC)滴度是用AV40polyA区特异性的探针和引物通过TaqMan分析来确定的，如以前文献所报道的[Gao，G.等，(2000)Hum Gene Ther11，2079-91]。

[0239] 每种血清型都构建载体以便于体内和体外实验。将8个不同转基因表达盒插入到载体中，然后制备每种血清型的重组病毒颗粒。根据其基因组拷贝数，恢复的病毒计数列在下面的表4中。每种血清型的载体产量都很高，但血清型之间的差异不同。表中所列的数据是多批50个平板(150mm)转染的基因组拷贝数的平均值加减标准差$\times 10^{13}$。

[0240] 表4.重组病毒的制备

[0241]

	AAV2/1	AAV2/2	AAV2/5	AAV2/7	AAV2/8
CMV LacZ	7.30 ± 4.33 (n=9)	4.49 ± 2.89 (n=6)	5.19 ± 5.19 (n=8)	3.42 (n=1)	0.87 (n=1)
CMV EGFP	6.43 ± 2.42 (n=2)	3.39 ± 2.42 (n=2)	5.55 ± 6.49 (n=4)	2.98 ± 2.66 (n=2)	3.74 ± 3.88 (n=2)
TBG LacZ	4.18 (n=1)	0.23 (n=1)	0.704 ± 0.43 (n=2)	2.16 (n=1)	0.532 (n=1)
Alb AlAT	4.67 ± 0.75 (n=2)	4.77 (n=1)	4.09 (n=1)	5.04 (n=1)	2.02 (n=1)
CB AlAT	0.567 (n=1)	0.438 (n=1)	2.82 (n=1)	2.78 (n=1)	0.816 ± 0.679 (n=2)
TBG rhCG	8.51 ± 6.65 (n=6)	3.47 ± 2.09 (n=5)	5.26 ± 3.85 (n=4)	6.52 ± 3.08 (n=4)	1.83 ± 0.98 (n=5)
TBG cFIX	1.24 ± 1.29 (n=3)	0.63 ± 0.394 (n=6)	3.74 ± 2.48 (n=7)	4.05 (n=1)	15.8 ± 15.0 (n=5)

[0242] 实施例6：假型载体的血清学分析

[0243] 给C57BL/6小鼠肌肉注射AAVCBA1AT载体的不同血清型(5×10^{11}GC)，34天后收集样品。为了检测血清中是否存在每种AAV血清型的中和抗体和交叉中和抗体，用以转导为基础的中和抗体分析方法分析血清[Gao，G.P.等，(1996)J Virol70，8934-43]。为了提高特异性，通过评价血清抑制不同血清型的报告病毒(AAVCMVEGFP)转导84-31细胞的能力来确定中和抗体的存在。每种血清型的报告病毒AAVCMVEGFP[感染多重性(multiplicity of infection)(MOI)能使90％的指示细胞感染]与接受不同AAV血清型的动物或对照鼠来源的热灭活血清共孵育。37℃孵育1小时后，将病毒加到含84-31细胞的96孔板中，根据不同的病毒血清型，孵育48或72小时。用FluoroImagin(Molecular Dynamics)测定GF P的表达并用Image Quant Software定量。中和抗体的滴度定义为能将转导抑制到50％以下的最高血清稀释度。

[0244] GFP表达载体的应用简化了中和抗体的分析方法，该方法是基于对允许细胞系(即稳定表达Ad5的E4的293细胞)转导的抑制。AAV血清型的抗血清是通过给动物肌肉注射重组病毒而制备的。1:20和1:80稀释度的抗血清对AAV转导的中和作用进行了分析(见下面的表5)。AAV1，AAV2，AAV5和AAV8的抗血清可以抑制该抗血清来源的血清型的转导(AAV5和AAV8的抑制作用比AAV1和AAV2弱)，但是不能抑制其他血清型的转导(即没有迹象表明存在交叉中和作用，说明AAV8是真正的独特血清型)。

[0245] 表5.新型AAV血清型的血清学分析

[0246]

		84-31细胞被AAVCMVEGFP病毒感染的百分比									
		AAV2/1		AAV2/2		AAV2/5		AAV2/7		AAV2/8	
		血清稀释度		血清稀释度		血清稀释度		血清稀释度		血清稀释度	
血清	免疫载体	1/20	1/80	1/20	1/80	1/20	1/80	1/20	1/80	1/20	1/80
组1	AAV2/1	0	0	100	100	100	100	100	100	100	100
组2	AAV2/2	100	100	0	0	100	100	100	100	100	100
组3	AAV2/5	100	100	100	100	16.5	16.5	100	100	100	100
组4	AAV2/7	100	100	100	100	100	100	61.5	100	100	100
组5	AAV2/8	100	100	100	100	100	100	100	100	26.3	60

[0247] 从52个正常人的血清中筛选能中和所选血清型的样本。结果没有找到能中和AAV2/7和AAV2/8血清样本，而20％和10％的血清分别可以中和AAV2/2和AAV2/1载体。一种混合IgG代表了所收集的60000人的血清样本，该混合物不能中和AAV2/7和AAV2/8，但是AAV2/2和AAV2/1载体分别被滴度为1/1280和1/640样品所中和。

[0248] 实施例7:AAV载体不同血清型的体内评价

[0249] 本试验中，7种重组AAV基因组，AAV2CBhA1AT，AAV2AlbhA1AT，AAV2CMVrhCG，AAV2TBGrhCG，AAV2TBGcFIX，AAV2CMVLacZ和AAV2TBGLacZ包装有不同血清型的衣壳蛋白。在所有7中载体中，小基因表达盒的两侧都是AAV2ITR。人α-抗胰蛋白酶(A1AT)[Xiao，W.等，(1999)J Virol73，3994-4003]、猕猴绒毛膜促性腺激素(CG)β-亚单位[Zoltick，P.W.& Wilson，J.M.(2000)Mol Ther2，657-9]、犬因子IX[Wang，L.等，(1997)Proc Natl Acad Sci U S A94，11563-6]和细菌β-半乳糖苷酶(即Lac Z)基因的cDNA作为报告基因。对于直接的肝基因转移来说，用鼠白蛋白基因启动子(Alb)[Xiao，W.(1999)，上面已引用]或人甲状腺激素结合球蛋白基因启动子(TBG)[Wang(1997)，上面已引用]驱动报告基因在肝脏内的组织特异性表达。对直接的肌肉基因转移来说，用巨细胞病毒早期启动子(CMV)或鸡β-肌动蛋白启动子加上CMV增强子(CB)来控制报告基因的表达。

[0250] 如果通过直接的肌肉注射进行基因转移，就是将载体注射到4-6周龄NCR裸鼠或C57BL/6小鼠(Taconic，Germantown，NY)的右侧胫前腓肌内。如果是直接的肝基因转移，就是将载体通过门静脉注射到到7-9周龄NCR裸鼠或C57BL/6小鼠(Taconic，Germantown，NY)的肝脏内。载体注射后在不同的时间点收集血清样本。在不同的时间点采集注射LacZ载体

的小鼠肌肉和肝组织制作冷冻切片并进行Xgal组织化学染色。在重复给药的试验中，C56BL/6小鼠首先被肌肉注射AAV2/1,2/2,2/5,2/7和2/8CBA1AT载体，检测A1AT基因的表达7周。然后门静脉内注射AAV2/8TBGcFIX并检测cFIX基因的表达。

[0251] ELISA分析用于定量检测上述hA1AT,rhCG和cFIX蛋白的血清水平[Gao,G.P.等,(1996)J Virol70,8934-43;Zoltick,P.W.&Wilson,J.M.(2000)Mol Ther2,657-9;Wang,L.等,Proc Natl Acad Sci U S A94,11563-6]。试验结束时处死动物，采集肌肉和肝组织，从中提取DNA，利用与测定制备载体的滴度一样的一组引物和探针通过TaqMan方法定量分析靶组织基因组内存在的本发明的载体拷贝数[Zhang,Y.等,(2001)Mol Ther3,697-707]。

[0252] 新型血清型来源的载体在直接的肌肉和肝基因转移的鼠模型中进行评价，并与已知的血清型AAV1、AAV2和AAV5来源的载体进行比较。表达分泌蛋白的载体(α-抗胰蛋白酶(A1AT)和绒毛膜促性腺激素(CG))用作指标通过血清的ELISA分析来定量测定不同血清型之间的相对转导效率。靶器官内转导载体在细胞内的分布用LacZ表达载体和X-gal组织化学方法进行评价。

[0253] AAV载体在骨骼肌中的作用通过将载体直接注射到小鼠的胫前腓肌内进行分析。载体含有相同的AAV2基因组和CMV的早期即刻基因，该基因的表达受AMC增强的β-肌动蛋白启动子控制。以前的研究表明当AAV载体表达人A1AT蛋白时，具有完全免疫能力的C57BL/6小鼠体内能激发针对该蛋白的有限的体液免疫反应[Xiao,W.等,(1999)J Virol73,3994-4003]。

[0254] 在各种载体中，AAV2/1载体产生的A1AT水平最高，AAV2/2载体最低，AAV2/7和AAV2/8载体处于二者之间。nu/nu NCR小鼠注射载体28天后CG的表达水平达到峰值，其中AAV2/7最高，AAV2/2最低，AAV2/8和AAV2/1位于二者之间。注射AAV2/1和AAV2/7lacZ载体后可以在注射部位的所有肌纤维中产生基因表达，而注射AAV2/2和AAV2/8载体后能观察到的lacZ阳性肌纤维要低的多。这些数据表明AAV2/7载体在骨骼肌中的转导效率和AAV2/1是一样的，并且是以前所报道的血清型中骨骼肌转导效率最高的[Xiao,W.(1999)，上文引用；Chao,H.等,(2001)Mol Ther4,217-22;Chao,H.等,(2000)Mol Ther2,619-23]。

[0255] 用同样的鼠模型来评价直接的肝基因转移。同样剂量的基因组拷贝来源的载体注射到小鼠的门静脉。，然后分析转基因的表达。每种载体都含有AAV2基因组，用上述的肝特异性启动子(即白蛋白或甲状腺激素结合球蛋白的启动子)来驱动转基因的表达。更特殊的是直接的肌肉和肝基因转移分别用CMVCG和TBGCG小基因表达盒。RhCG的表达水平用相对单位(RUs×10^3)来表示。数据来自于载体注射后28天的血清样品(每组4只动物)。如表3所示，衣壳蛋白对A1AT载体在nu/nu和C57BL/6小鼠体内的转导效率以及CG载体在C57BL/6小鼠体内的转导效率的影响是一致的(见表6)。

[0256] 表6.猕猴绒毛膜促性腺激素(rhCG)β-亚单位的表达

[0257]

载体	肌肉	肝
AAV2/1	4.5 ± 2.1	1.6 ± 1.0
AAV2	0.5 ± 0.1	0.7 ± 0.3
AAV2/5	ND*	4.8 ± 0.8
AAV2/7	14.2 ± 2.4	8.2 ± 4.3
AAV2/8	4.0 ± 0.7	76.0 ± 22.8

[0258] *本试验中未检测

[0259] 在所有的例子中，AAV2/8载体产生的转基因表达水平最高，为16-110，高于AAV2/2载体；AAV2/5和AAV2/7载体介于二者之间，但AAV2/7又高于AAV2/5。注射相应的lacZ载体的动物的肝组织切片经X-gal染色后发现转导细胞的数目与转基因的表达总水平之间存在相关关系。利用实时PCR技术分析接受A1AT载体的C57BL/6小鼠肝组织内提取的DNA中载体DNA的丰度。

[0260] 注射后56天肝组织内载体DNA的量与转基因的表达水平成相关关系(见表7)。这个试验中，设计了一组载体基因组SV40polyA区特异性的探针和引物用于TaqMan PCR。所表示的数值是三只动物的平均值加减标准差。注射后56天杀死动物，采集肝组织提取其DNA。这些试验表明AAV8是直接的肝基因转移最有效的载体，因为其转导的肝细胞数目最多。

[0261] 表7—注射$1x10^{11}$基因组拷贝的载体后用实时PCR分析nu/nu小鼠肝组织内AAV载体的丰度

[0262]

AAV 载体/剂量	每个细胞内的基因组拷贝数
AAV2/1A1bA1AT	0.6 ± 0.36
AAV2A1bA1AT	0.003 ± 0.001
AAV2/5A1bA1AT	0.83 ± 0.64
AAV2/7A1bA1AT	2.2 ± 1.7
AAV2/8A1bA1AT	18 ± 11

[0263] 上面的血清学数据表明用其他血清型免疫后体内的AAV2/8载体不能被中和。C57BL/6小鼠肌肉注射不同血清型的A1AT载体56天后门静脉注射表达犬因子IX(10^{11}基因组拷贝)的AAV2/8载体。因子IX表达的最高水平出现在AAV2/8载体(17±2μg/ml，n=4)注射空白动物后14天，与注射AAV2/1(31±23μg/ml，n=4)、AAV2/2(16μg/ml，n=2)和AAV2/7(12μg/ml，n=2)的动物无显著性差异。这结果与AAV2/8因子IX载体注射的AAV2/8免疫动物中所观察到的结果相反，其中没有检测到因子IX(<0.1μg/ml，n=4)。

[0264] 利用cap基因保守区特异性的寡核苷酸确实可以从猕猴中扩增出代表独特AAV的序列。至少两个不同克隆的猕猴的多种组织内可以检测到一致的cap特征区域序列。从一个来源中分离全长的rep和cap开放阅读框架并测序。只需要评价新型AAV的cap开放阅读框架作为载体的能力，因为含AAV7或AAV8衣壳的载体是用AAV2的ITR和rep制备的。这也可以简

化不同载体的比较，因为不同载体血清型的实际载体基因组是一致的。实际上，利用这种方法制备的重组载体的产生在不同的血清型之间无差异。

[0265] AAV7和AAV8载体具有独特的免疫学特征(即它们不能被抗其他血清型的抗体所中和)。而且，人血清不能中和AAV7和AAV8载体的转导，因此它们比现在正在研究的人源[Chirmule,N.等,(1999)Gene Ther6,1574-83]。

[0266] 每种新型载体的趋向性都适于体内应用。AAV2/7载体转导骨骼肌的效率与AAV2/1相似，后者是到目前为止所检测到的灵长类动物AAV在骨骼肌中转导效率最高的[Xiao,W.,上面已引用;Chou(2001)，上面已引用,以及Chou(2000)，上面已引用]。但值得注意的是，在肝的基因转移中，AA2/8的效率要比其他血清型高，到目前为止，还没有找到能稳定转染肝细胞的令人满意的载体。与其他载体相比，AAV2/8的基因转移效率可以提高10到100倍。AAV2/8的高转导效率的机制还不清楚，尽管有人推测可能是由于其摄取所依赖的受体与其他的载体不同，该受体在肝细胞的基侧表面具有较高的活性。这种高效率的载体对于直接的肝基因转移相当有用，因为在治疗尿素循环障碍和家族性高胆固醇血症时转导细胞的数目是很关键的。

[0267] 因此，本发明提供了一种新型方法通过PCR检索基因组序列来分离新型AAV。扩增出的序列易于克隆到载体中并在动物体内检测。由于人体内不存在对AAV7的免疫并且该载体的趋向性适于转染肌肉组织，因此AAV7很适于作为载体用于人体的基因治疗以及其他的体内应用。同样，由于缺乏对本发明AAV血清型的免疫及其趋向性，使这些AAV可用于转移治疗分子或其他有用的分子。

[0268] 实施例9—组织趋向性研究

[0269] 在设计高通量功能性筛选方案构建新型AAV载体时，选择非组织特异性的及高活性的启动子，CB启动子(CMV增强的鸡β肌动蛋白启动子)来驱动一个易于检测的并且可以定量检测的报告基因-人α抗胰蛋白酶基因。每个新型AAV克隆只需要构建一个载体来进行基因转移的研究，检测特定AAV载体对三种不同靶组织，肝、肺和肌肉的趋向性。下表概括了4种新型AAV载体在组织趋向性研究(AAVCBA1AT)中所得到的数据。发现其中新型AAV衣壳克隆44.2在所有3种组织中都是极好的基因转移工具，特别是在肺组织中。表8报告了试验第14天的数据(μg A1AT/mL血清)。

[0270] 表8

[0271]

载体	靶组织		
	肺	肝	肌肉
AAV2/1	ND	ND	45±11
AAV2/5	0.6±0.2	ND	ND
AAV2/8	ND	84±30	ND
AAV2/rh.2(43.1)	14±7	25±7.4	35±14
AAV2/rh.10(44.2)	23±6	53±19	46±11
AAV2/rh.13(42.2)	3.5±2	2±0.8	3.5±1.7
AAV2/rh.21(42.10)	3.1±2	2±1.4	4.3±2

[0272] 另外还进行了另外的试验来进一步确证AAV44.2在肺组织中的超级趋向性。首先，将携带肺组织特异性表达的CC10hA1AT小基因的AAV载体与新型AAV的衣壳假型，然后注射给免疫缺陷动物(NCR裸鼠)，如下表所示，通过气管灌注等量的载体(每个动物50μl原液，不稀释)。在表9中，每只NCR裸鼠50μl原液，检测下限为0.033μg/ml，检测时间为第28天。

[0273] 表9

[0274]

载体	50μl 载体中的总 GC	50μl 载体中 A1AT/ml 的 μg 数	$1x10^{11}$ 载体中 A1AT/ml 的 μg 数	与 rh.10(克隆 44.2) 相比较的相对基因转移
2/1	$3x10^{12}$	2.6±0.5	0.09±0.02	2.2
2/2	$5.5x10^{11}$	<0.03	<0.005	<0.1
2/5	$3.6x10^{12}$	0.65±0.16	0.02±0.004	0.5
2/7	$4.2x10^{12}$	1±0.53	0.02±0.01	0.5
2/8	$7.5x10^{11}$	0.9±0.7	0.12±0.09	2.9
2/ch.5(A.3.1)	$9x10^{12}$	1±0.7	0.01±0.008	0.24
2/rh.8(43.25)	$4.6x10^{12}$	26±21	0.56±0.46	13.7
2/rh.10(44.2)	$2.8x10^{12}$	115±38	4.1±1.4	100
2/rh.13(42.2)	$6x10^{12}$	7.3±0.8	0.12±0.01	2.9
2/rh.21(42.10)	$2.4x10^{12}$	9±0.9	0.38±0.04	9.3
2/rh.22(42.11)	$2.6x10^{12}$	6±0.4	0.23±0.02	5.6
2/rh.24(42.13)	$1.1x10^{11}$	0.4±0.3	0.4±0.3	1

[0275] 也将载体注射给具有完全免疫能力的动物(C57BL/6)，所用的基因拷贝数相同($1x10^{11}$GC)，如表10所示。(每只动物$1x10^{11}$GC，C57BL/6，14天，检测下限为0.033μg/ml)。

[0276] 表10

[0277]

AAV 载体	$1x10^{11}$ 载体中 A1AT/ml 的 μg 数	与 rh.10(克隆 44.2)相比较的相对基因转移
2/1	0.076±0.031	2.6
2/2	0.1±0.09	3.4
2/5	0.0840.033	2.9
2/7	0.33±0.01	11
2/8	1.92±1.3	2.9
2/ch.5(A.3.1)	0.048±0.004	1.6
2/rh.8(43.25)	1.7±0.7	58
2/rh.10(44.2)	2.93±1.7	100
2/rh.13(42.2)	0.45±0.15	15
2/rh.21(42.10)	0.86±0.32	29
2/rh.22(42.11)	0.38±0.18	13
2/rh.24(42.13)	0.3±0.19	10

[0278] 两个试验的数据都确证了克隆44.2在直接的肺基因转移中的超级趋向性。

[0279] 令人感兴趣的是,克隆44.2在直接的肝和肌肉基因转移中的表现也是十分突出的,接近最佳的肝转导载体AAV8和最佳的肌肉转导载体AAV1所达到的水平,说明这个新型AAV具有极其引人的生物意义。

[0280] 为了研究这些新型血清型的血清学特性,制备了假型AAVGFP载体免疫兔,并且在存在或缺少不同衣壳抗血清的情况下体外转导84-31细胞。数据概述如下。

[0281] 表11a.8431细胞内的交叉-NAB分析及Adv共感染

[0282] 8431细胞感染下列载体(与Adv共感染):

[0283]

血清来源兔免疫所用载体:	10^9 GC	10^9 GC	10^9 GC	10^{10} GC
	rh.13	***rh.21***	***rh.22***	***rh.24***
	AAV2/42.2	AAV2/42.10	AAV2/42.11	AAV2/42.13
AAV2/1	1/20	1/20	1/20	No NAB
AAV2/2	1/640	1/1280	1/5120	No NAB
AAV2/5	No NAB	1/40	1/160	No NAB

[0284]

AAV2/7	1/81920	1/81920	1/40960	1/640
AAV2/8	1/640	1/640	1/320	1/5120
Ch. 5 AAV2/A3	1/20	1/160	1/640	1/640
rh. 8 AAV2/43.25	1/20	1/20	1/20	1/320
rh. 10 AAV2/44.2	No NAB	No NAB	No NAB	1/5120
rh. 13 AAV2/42.2	1/5120	1/5120	1/5120	No NAB
rh. 21 AAV2/42.10	1/5120	1/10240	1/5120	1/20
rh. 22 AAV2/42.11	1/20480	1/20480	1/40960	No NAB
rh. 24 AAV2/42.13	No NAB	1/20	1/20	1/5120

[0285] 表11b.8431细胞内的交叉-NAB分析及Adv共感染

[0286] 8431细胞感染下列载体(与Adv共感染)：

[0287]

	10^{8} GC	10^{10} GC	10^{10}GC	10^{9} GC	10^{9} GC
血清来源兔免疫所用	*rh. 12*	*ch. 5*	*rh. 8*	*rh. 10*	*rh. 20*
载体:	AAV2/42.1B	AAV2/A3	AAV2/43.25	AAV2/44.2	AAV2/42.8.2
AAV2/1	No NAB	1/20480	No NAB	1/80	ND
AAV2/2	1/20	No NAB	No NAB	No NAB	ND
AAV2/5	No NAB	1/320	No NAB	No NAB	ND
AAV2/7	1/2560	1/640	1/160	1/81920	ND
AAV2/8	1/10240	1/2560	1/2560	1/81920	ND
ch. 5 AAV2/A3	1/1280	1/10240	ND	1/5120	1/320
rh. 8 AAV2/43.25	1/1280	ND	1/20400	1/5120	1/2560
rh. 10 AAV2/44.2	1/5120	ND	ND	1/5120	1/5120
rh. 13 AAV2/42.2	1/20	ND	ND	No NAB	1/320
rh. 21 AAV2/42.10	1/20	ND	ND	1/40	1/80
rh. 22 AAV2/42.11	No NAB	ND	ND	ND	No NAB
rh. 24 AAV2/42.13	1/5120	ND	ND	ND	1/2560

[0288] 表12

[0289]

兔血清的滴度			加强免疫后的滴度
载体		第 21 天的滴度	
ch. 5	AAV2/A3	1/10,240	1/40,960
rh. 8	AAV2/43. 25	1/20,400	1/163,840
rh. 10	AAV2/44. 2	1/10,240	1/527,680
rh. 13	AAV2/42. 2	1/5,120	1/20,960
rh. 21	AAV2/42. 10	1/20,400	1/81,920
rh. 22	AAV2/42. 11	1/40,960	ND
rh. 24	AAV2/42. 13	1/5,120	ND

[0290] 表13a.用GFP载体感染8431细胞(Adv共感染)

[0291]

	10^9 GC/孔	10^9 GC/孔	10^9 GC/孔	10^9 GC/孔	10^9 GC/孔	10^9 GC/孔
						ch. 5
	AAV2/1	AAV2/2	AAV2/5	AAV2/7	AAV2/8	AAV2/A3
# GFU/组	128	>200	95	56	13	1
	83	>200	65	54	11	1

[0292] 表13b.用GFP载体感染8431细胞(共感染Adv)

[0293]

	10^9 GC/孔	10^9 GC/孔	10^9 GC/孔	10^9 GC/孔	10^9 GC/孔	10^9 GC/孔	10^9 GC/孔
	rh. 8	*rh. 10*	*rh. 13*	*rh. 21*	*rh. 22*	*rh. 24*	*rh. 12*
	AAV2/43. 25	AAV2/44. 2	AAV2/42. 2	AAV2/42. 10	AAV2/42. 11	AAV2/42. 13	AAV2/42. 1B
# GFU/组	3	13	54	62	10	3	18
	2	12	71	60	14	2	20
			48	47	16	3	12

[0294] 实施例10—家族性高胆固醇血症的鼠模型

[0295] 下面的试验用于证明本发明的AAV2/7载体能转移LDL受体并可以表达有效量的LDL受体以降低家族性高胆固醇血症动物模型体内的血浆胆固醇水平和甘油三脂的水平。

[0296] A.载体的构建

[0297] 带有AAV7或AAV8衣壳蛋白的AAV载体利用假型策略构建[Hildinger M等,

J.Virol2001;75:6199-6203]。携带AAV2末端反向重复序列(ITR)的重组AAV基因组是通过用顺式质粒、腺病毒辅助质粒和一个嵌合的包装载体三重转染293细胞而包装成的，嵌合的包装载体是由新型AAV血清型的衣壳和AAV2的rep基因融合而成。嵌合的包装质粒是通过以前所描述的方法构建的[Hildinger等，上面已引用]。重组载体用标准的CsCl2沉淀方法纯化。为了确定病毒的产生量，利用载体SV40polyA区特异性的探针和引物进行TaqMan(Applied Biosystems)分析[Gao GP等，Hum Gene Ther.20000ct10;11(15):2079-91]。得到的载体表达人甲状腺激素结合球蛋白基因启动子(TBG)控制下的转基因。

[0298] B.动物

[0299] C57Bl/6来源的LDL受体缺陷小鼠是从Jackson实验室(Bar Harbor，ME，USA)购买的，作为生殖克隆来饲养。不限制小鼠饮水，载体注射前3周开始喂养高脂Western Diet(胆固醇百分含量高)。通过眶后取血在0天和7天采集血液样品，测定血脂含量。小鼠随机分为7组。如前所述通过门静脉注射载体[Chen SJ等，Mol Therapy2000;2(3)，256-261]。简言之，小鼠用氯胺酮和赛拉嗪麻醉，用一个30g的注射器吸取适当剂量的用100ul PBS稀释的载体直接注射到门静脉中。按压注射部位以止血。缝合皮肤伤口，在接下来的时间内仔细检测小鼠。从直接肝基因转移后14天开始每周采血以测定血脂含量。在注射后6周和12周时每组处死2只动物，测定粥样动脉硬化的大小和受体的表达情况。其余的小鼠在20周时处死并测定粥样动脉硬化的大小和受体的表达情况。

[0300] 表14

[0301]

	载体	剂量	n
组1	AAV2/7-TBG-hLDLr	$1x10^{12}$gc	12
组2	AAV2/7-TBG-hLDLr	$3x10^{11}$gc	12
组3	AAV2/7-TBG-hLDLr	$1x10^{11}$gc	12
组4	AAV2/8-TBG-hLDLr	$1x10^{12}$gc	12
组5	AAV2/8-TBG-hLDLr	$3x10^{11}$gc	12
组6	AAV2/8-TBG-hLDLr	$1x10^{11}$gc	12
组7	AAV2/7-TBG-LacZ	$1x10^{11}$gc	16

[0302] C.血清脂蛋白和肝功能分析

[0303] 禁食6小时后从眶后血管丛采集血液样品。通过离心从血浆中制备血清。用临床生化自动分析仪(ACE，Schiapparelli Biosystems，Alpha Wassermann)测定血清中脂蛋白和肝转氨酶的血浆含量。

[0304] D.转基因表达的检测

[0305] 通过免疫荧光染色和Western印迹分析LDL受体的表达情况。用裂解缓冲液(20mM Tris，pH7.4，130mM NaCl，1%Triton X100，蛋白酶抑制剂(完全的，无EDTA，Roche，Mannheim，Germany))将冷冻的肝组织匀浆用于Western印迹分析。利用Micro BCA Protein Assay Reagent Kit(Pierce，Rockford，IL)测定蛋白浓度。40μg蛋白溶于4-15%Tris-HCl Ready Gels(Biorad，Hercules，CA)中，然后转移到硝酸纤维素膜(Invitrogen，)上。为了制备抗-hLDL受体的抗体，给兔静脉注射AdhLDLr prep($1x10^{13}$GC)。4周后采集兔血清用于Western印迹。用1:100稀释的血清作为一抗，以连接HRP的抗兔IgG，然后进行ECL化学发光

检测(ECL Western印迹检测试剂盒,Amersham,Arlington Heights,IL)。

[0306] E.免疫组织化学

[0307] 用免疫组织化学染色方法分析冷冻肝组织切片中LDL受体的表达。10um冷冻切片或者用丙酮固定5分钟,或者不固定。与10%的山羊血清共孵育1小时进行。然后切片与一抗在室温下共孵育1小时,按照厂家的说明稀释兔抗人LDL多克隆抗体(Biomedical Technologies Inc.,Stoughton,MA),然后加到切片上。用PBS洗涤切片,然后与1:100稀释的荧光标记羊抗兔IgG(Sigma,St Louis,MO)共孵育。最后在荧光显微镜Nikon Microphot-FXA观察切片。在所有的例子中,每一步孵育结束后都要用PBS彻底洗涤。阴性对照预孵育用PBS,用同型相配的非免疫对照抗体来代替一抗。每个试验在同一天设置上面提到的三种类型的对照。

[0308] F.基因转移效率

[0309] 在预定的时间点处死动物后取其肝组织,组织在液氮中迅速冷冻,然后储存在-80℃备用。用QIAamp DNA Mini Kit(QIAGEN GmbH,Germany)按照说明书从肝组织中提取DNA。利用上述的AV40poly(A)尾巴特异性的引物和探针通过Taqman分析来测定肝组织中AAV载体的基因组拷贝数。

[0310] G.动脉粥样硬化斑块的测量

[0311] 为了测量鼠动脉内粥样斑块的大小,浆小鼠麻醉(10%氯胺酮和赛拉嗪,静脉注射),打开胸腔,从左心室向动脉系统中灌注冰冻的磷酸缓冲盐。然后小心地将动脉切下,从主动脉弓到股动脉延腹侧中线切开,然后用甲醛固定。富含脂质的动脉粥样硬化斑块用Sudan IV(Sigma,Germany)染色,动脉用大头针钉在黑色石蜡表面,用Sony DXC-960MD彩色摄像机照相。粥样斑块的面积及其在整个动脉表面所占的比例用Phase3Imaging Systems(Media Cybernetics)分析。

[0312] H.I^{125}LDL的清除率

[0313] 每组检测两只动物。每隔30秒通过尾静脉将I^{125}-标记的LDL(一般由Dan Rader,U Penn提供)缓慢注入到小鼠体内(每只动物注射用100μl无菌PBS稀释的1,000,000计数的[I^{125}]-LDL)。分别在注射后3min,30min,1.5hr,3hr,6hr从眶后动脉丛采集血液样品。从全血中分离血浆,取10μl血浆在γ计数仪上计数。最终的分数代谢率从脂蛋白清除数据中计算。

[0314] I.肝脂质蓄积的评价

[0315] 将肝切片用油红染色以测定脂质蓄积程度。冷冻肝切片先用蒸馏水漂洗,然后在无水丙二醇中孵育2分钟。将切片放入到油红溶液(0.5%,溶于丙二醇中)染色16小时,然后用Mayer苏木精溶液复染30秒,固定在加热的甘油胶溶液中。

[0316] 为了确定肝组织中胆固醇和甘油三脂的含量,将肝组织切片匀浆,然后在氯仿/甲醇(2:1)中孵育过夜。加入0.05%H_2SO_4后离心10分钟,收集每个样品的下层,分为两管,在液氮中冻干。为了测定胆固醇含量,将第一管中冻干的脂质溶于用氯仿溶解的1%Triton X-100中。溶解后用液氮冻干。将脂质溶于双氧水中后在37℃孵育30分钟,然后用Total Cholesterol Kit(Wako Diagnostics)测定总胆固醇的浓度。将第二管冻干的脂质溶于乙醇KOH中,在60℃孵育30分钟。然后加入1M $MgCl_2$,置于冰上孵育10分钟,然后14,000rpm离心30分钟。测定上清液中的甘油三脂含量(Wako Diagnostics)。

[0317] 与对照组相比，假型in AAV2/8或AAV2/7的所有载体都能降低总胆固醇、LDL和甘油三脂的含量。这些被检测的载体能以剂量依赖的方式矫正高胆固醇血症的表型、AAV2/8和AAV2/7小鼠内的粥样斑块面积在第一个试验中也可以观察到缩小(2个月)，并且这种效果可以持续到整个试验(6个月)。

[0318] 实施例—功能因子IX的表达和血友病的矫正

[0319] A.基因敲除小鼠

[0320] 对血友病B小鼠体内功能性犬因子IX(FIX)的表达情况作了检测。构建携带AAV1，AAV2，AAV5，AAV7或AAV8衣壳序列的载体，该载体的结构为AAV25'ITR-肝组织特异性启动子[LSP]-犬FIX-北美土拨鼠肝炎病毒转录后调节元件(WPRE)-AAV23'ITR。如Wang等，2000，Molecular Therapy2:154-158)所描述利用适当的衣壳构建这种载体。

[0321] 基因敲除小鼠的制备方法见Wang等，1997.Proc.Natl.Acad.Sci.USA94:11563-11566的描述。这个模型最接近人血友病B的表型。

[0322] 不同血清型(AAV1，AAV2，AAV5，AAV7和AAV8)的载体一次通过门静脉注射到成年血友病C57Bl/6小鼠体内，其中5种不同血清型的注射剂量为$1x10^{11}$GC/鼠，AAV8载体组注射较低的剂量，$1x10^{10}$GC/鼠。对照组注射$1x10^{11}$GC的AAV2/8TBG LacZ3。每组动物含5-10只雄性和雌性小鼠。载体注射后每两周采集一次血样。

[0323] 1.ELISA

[0324] 鼠血浆种的犬FIX浓度通过犬因子IX特异性的ELISA分析方法来检测，如Axelrod等，1990，Proc.Natl.Acad.Sci.USA，87:5173-5177所描述，本发明稍加改动。山羊抗犬因子IX抗体(Enzyme Research Laboratories)作为一抗，兔抗犬因子IX抗体(Enzyme Research Laboratories)作为二抗。注射后两周开始测定所有试验载体组的cFIX血浆水平。结果发现其血浆水平增高，在整个试验过程中，即到12周时，能维持在治疗水平。治疗水平被认为是正常水平的5%，即约250ng/mL。

[0325] 最高的表达水平出现在AAV2/8(10^{11})和AAV2/7载体组，达到超生理的cFIX浓度(比正常水平高10倍)。AAV2/8(10^{11})的表达水平大约比AAV2/2和AAV2/8(10^{10})组高10倍。最低的表达水平出现在AAV2/5组，虽然也在治疗水平范围内。

[0326] 2.体外活化部分凝血激酶时间(aPTT)的检测

[0327] FIX基因敲除小鼠血浆中功能性因子IX的活性通过体外活化部分凝血激酶时间(aPTT)分析方法来确定-从眶后动脉丛采集小鼠的血液样品加入到1/10体积的柠檬酸盐缓冲液中。aPTT试验方法如Wang等，1997，Proc.Natl.Acad.Sci.USA94:11563-11566所描述。

[0328] 通过aPTT分析发现，在注射后2周时所有载体注射鼠血浆样品的凝血时间都在正常范围内(约60秒)，正常的凝血时间和稍短于正常凝血时间能持续整个试验周期(12周)。

[0329] 持续最短的凝血时间出现在接受AAV2/8(10^{11})和AAV2/7的小鼠中。到12周时，AAV2/2也能诱导出与AAV2/8和AAV2/7组相同的凝血时间。但是，这种较短的凝血时间知道12周时也没在AAV2/2组观察到，而在2周时开始就可以在AAV2/8和AAV2/7组观察到较短的凝血时间(25-40秒之间)。

[0330] 从某些治疗小鼠中采集肝组织进行免疫组化染色。注射AAV2/8.cFIX载体的小鼠肝组织中约有70-80%的肝细胞是犬FIX阳性的。

[0331] B.血友病B犬

[0332] 通过模型研究发现，犬F.IX基因催化区的点突变可以使该基因表达的蛋白不稳定，导致血友病B[Evans等，1989，Proc.Natl.Acad.Sci.USA，86:10095-10099]。在University of North Carolina，Chapel Hill有一群这种犬已经饲养了20年以上。这些犬的自我平衡参数已经有了详细的描述，其中包括缺乏血浆F.IX抗原、全血凝血时间超过60分钟(正常犬为6-8分钟)以及活化部分凝血激酶时间延长至50-80秒(正常犬为13-28秒)。这些犬会反复出现自发性出血。一般来说，出现自发性出血时一次静脉注射10ml/kg正常犬血浆就可以制止；偶尔需要反复输注才能止血。

[0333] 按照下面的方案给4只犬通过门静脉注射AAV.cFIX。第一只犬一次注射剂量为$3.7x10^{11}$基因组拷贝(GC)/kg的AAV2/2.cFIX。第二只犬先注射AAV2/2.cFIX($2.8x10^{11}$GC/kg)，然后在第1180天第二次注射AAV2/7.cFIX($2.3x10^{13}$GC/kg)。第三只犬一次注射剂量为$4.6x10^{12}$GC/kg的AAV2/2.cFIX。第四只犬先注射AAV2/2.cFIX($2.8x10^{12}$GC/kg)，在第995天再注射AAV2/7.cFIX($5x10^{12}$GC/kg)。

[0334] .血友病犬全身麻醉后剪掉腹部的毛，手术切开腹部，一次将载体注射到门静脉内。为了防止手术期间出血，动物预先注射正常的犬血浆。将犬镇静后，插管诱导全身麻醉，腹部剃毛备皮。切开腹部后将脾移到手术视野内，找到脾静脉的位置，将一条缝合线松弛地记在近端，在静脉的远端切开一个小的切口，将针头快速插入静脉中，然后将缝合线松开，将一根5F套管栓到门静脉分叉处附近的静脉上。止血完毕后导管气囊内充气，在5分钟内将约5.0ml用PBS稀释的载体注入到门静脉内。然后将气囊放气，去除导管，完成静脉止血。然后将脾复位，烧灼出血的血管，缝合伤口。动物完全耐受手术后取下插管。血样的分析如前所述[Wang等，2000，Molecular Therapy2:154-158]。

[0335] 结果表明如预期的那样AAV2/7可以矫正和部分矫正血友病的症状。

[0336] 本说明书引用的所有文献都已纳入本文作为参考。在参考特别优选实施方式对本发明进行描述的同时，只要不背离本发明的精神对本发明作出修改也是值得赞赏的。这种修改应该包括在本发明权利要求的范围内。

序列表

<110> 宾夕法尼亚大学托管会(The Trustees of The University of Pennsylvania)

<120> 检测和/或鉴定腺伴随病毒(AAV)序列以及分离所鉴定的新型序列的方法

<130> UPN-02735PCT

<150> US 60/350,607
<151> 2001-11-13

<150> US 60/341,117
<151> 2001-12-17

<150> US 60/377,066
<151> 2002-05-01

<150> US 60/386,675
<151> 2002-06-05

<160> 120

<170> PatentIn version 3.1

<210> 1
<211> 4721
<212> DNA
<213> 腺伴随病毒血清型 7

<400> 1

[0001]

ttggccactc cctctatgcg cgctcgctcg ctcggtgggg cctgcggacc aaaggtccgc 60
agacggcaga gctctgctct gccggcccca ccgagcgagc gagcgcgcat agagggagtg 120
gccaactcca tcactagggg tacgcgcgaag cgcctcccac gctgccgcgt cagcgctgac 180
gtaaatcacg tcatagggga gtggtcctgt attagctgtc acgtgagtgc tttgcgacca 240
tttgcgacca ccacgtggcc atttgaggta tatatggccg agtgagcgag caggatctcc 300
atttgaccgc cgaaatttga acgagcagca gccatgccgg gtttctacga gatcgtgatc 360
aaggtgccga gcgacctgga cgagcacctg ccgggcattt ctgactcgtt tgtgaactgg 420
gtggccgaga aggaatgggg gctgcccccg gattctgaca tggatctgaa tctgatcgag 480
caggcacccc tgaccgtggc cgagaagctg cagcgcgact tcctggtcca atggcgccgc 540
gtgagtaagg ccccggaggc cctgttcttt gttcagttcg agaagggcga gagctacttc 600
cacatccacg ttctggtgga gaccacgggg gtcaagtcca tggtgctagg ccgcttcctg 660
agtcagatte gggagaagct ggtccagacc atctaccgcg gggtcgagcc cacgctgccc 720
aactggttcg cggtgaccaa gacgcgtaat ggcgccggcg gggggaacaa ggtggtggac 780
gagtgctaca tccccaacta catcttcctc aagacccagc ccgagctgca gtgggcgtgg 840
actaacatgg aggagtatat aagcgcgtgt ttgnacctgg ccgaacgcaa acggctcgtg 900
gcgcagcacc tgacccacgt cagccagacg caggagcaga acaaggagaa tctgaacccc 960
aattctgacg cgccccgtga caggtcaaaa accttccgcc gctacatgga gctggtcggg 1020
tggctggtgg accggggcat cacctccgag aagcagtgga tccaggagga ccaggcctcg 1080
tacatctcct tcaacgccgc ctccaactcg cggtcccaga tcaaggccgc gctggacaat 1140
gccggcaaga tcatggcgct gaccaaatcc gcgccccgct accttggtgg gcccttcgctg 1200
ccccgcggaca ttaaaaccaa ccgcatctac cgcatcctgg agctgaacgg gtacgatcct 1260
gcctacgccg gctccgtctt tctcggctgg gcccagaaaa agttcgggaa gcgcaacacc 1320
atctggctgt ttgggccccgc caccaccggc aagaccaaca ttgcggaagc catcgcccac 1380
gccgtgccct tctacggctg cgtcaactgg accaatgaga actttccctt caacgattgc 1440
gtcgacaaga tggtgatctg gtgggaggag ggcaagatga cggccaaggt cgtggagtcc 1500
gccaaggcca ttctcggcgg cagcaaggtg cgcgtggacc aaaagtgcaa gtcgtccgcc 1560
cagatcgacc ccaccccccgt gatcgtcacc tccaacacca acatgtgcgc cgtgattgac 1620
gggaacagca ccaccttcga gcaccagcag ccgttgcagg accggatgtt caaatttgaa 1680
ctcacccgcc gtctggagca cgactttggc aaggtgacga agcaggaagt caaagagttc 1740
ttccgctggg ccagtgatca cgtgaccgag gtggcgcatg agttctacgt cagaaagggc 1800
ggagccagca aaagaccccc ccccgatgac gcggatatca gcgagcccaa gcgggcctgc 1860
ccctcagtcg cggatccatc gacgtcagac gcggaaggag ctccggtgga ctttgccgac 1920

[0002]

```
aggtaccaaa acaaatgttc tcgtcacgcg ggcatgattc agatgctgtt tccctgcaaa  1980
acgtgcgaga gaatgaatca gaatttcaac atttgcttca cacacggggt cagagactgt  2040
ttagagtgtt tccccggcgt gtcagaatct caaccggtcg tcagaaaaaa gacgtatcgg  2100
aaactctgcg cgattcatca tctgctgggg cgggcgcccg agattgcttg ctcggcctgc  2160
gacctggtca acgtggacct ggacgactgc gtttctgagc aataaatgac ttaaaccagg  2220
tatggctgcc gatggttatc ttccagattg gctcgaggac aacctctctg agggcattcg  2280
cgagtggtgg gacctgaaac ctggagcccc gaaacccaaa gccaaccagc aaaagcagga  2340
caacggccgg ggtctggtgc ttcctggcta caagtacctc ggacccttca acggactcga  2400
caagggggag cccgtcaacg cggcggacgc agcggccctc gagcacgaca aggcctacga  2460
ccagcagctc aaagcgggtg acaatccgta cctgcggtat aaccacgccg acgccgagtt  2520
tcaggagcgt ctgcaagaag atacgtcatt tgggggcaac ctcgggcgag cagtcttcca  2580
ggccaagaag cgggttctcg aacctctcgg tctggttgag gaaggcgcta agacggctcc  2640
tgcaaagaag agaccggtag agccgtcacc tcagcgttcc cccgactcct ccacgggcat  2700
cggcaagaaa ggccagcagc ccgccagaaa gagactcaat ttcggtcaga ctggcgactc  2760
agagtcagtc cccgaccctc aacctctcgg agaacctcca gcagcgccct ctagtgtggg  2820
atctggtaca gtggctgcag gcggtggcgc accaatggca gacaataacg aaggtgccga  2880
cggagtgggt aatgcctcag gaaattggca ttgcgattcc acatggctgg gcgacagagt  2940
cattaccacc agcacccgaa cctgggccct gcccacctac aacaaccacc tctacaagca  3000
aatctccagt gaaactgcag gtagtaccaa cgacaacacc tacttcggct acagcaccct  3060
ctgggggtat tttgacttta acagatttca ctgccactte tcaccacgtg actggcagcg  3120
actcatcaac aacaactggg gattccggcc caagaagctg cggttcaagc tcttcaacat  3180
ccaggtcaag gaggtcacga cgaatgacgg cgttacgacc atcgctaata accttaccag  3240
cacgattcag gtattctcgg actcggaata ccagctgccg tacgtcctcg gctctgcgca  3300
ccagggctgc ctgcctccgt tcccggcgga cgtcttcatg attcctcagt acggctacct  3360
gactctcaac aatggcagtc agtctgtggg acgttcctcc ttctactgcc tggagtactt  3420
ccctctcag atgctgagaa cgggcaacaa ctttgagttc agctacaget tcgaggacgt  3480
gcctttccac agcagctacg cacacagcca gagcctggac cggctgatga atccccttat  3540
cgaccagtac ttgtactacc tggccagaac acagagtaac ccaggaggca cagctggcaa  3600
tcgggaactg cagttttacc agggcgggcc ttcaactatg gccgaacaag ccaagaattg  3660
gttacctgga ccttgcttcc ggcaacaaag agtctccaaa acgctggatc aaaacaacaa  3720
cagcaacttt gcttggactg gtgccaccaa atatcacctg aacggcagaa actcgttggt  3780
taatcccggc gtcgccatgg caacticaca ggacgacgag gaccgctttt tcccatccag  3840
cggagtcctg attttggaaa aaactggagc aactaacaca actacattgg aaaatgtgtt  3900
aatgacaaat gaagaagaaa ttcgtcctac taatcctgta gccacggaag aatacgggat  3960
agtcagcagc aacttacaag cggctaatac tgcagcccag acacaagttg tcaacaacca  4020
gggagcctta cctggcatgg tctggcagaa cgggacgtg tacctgcagg gtcccatctg  4080
ggccaagatt cctcacacgg atggcaactt tcacccgtct cctttgatgg gcggctttgg  4140
acttaaacat ccgcctcctc agatcctgat caagaacact cccgttcccg ctaatcctcc  4200
ggaggtgttt actcctgccc agttttcttc gttcatcaca cagtacagca ccggacaagt  4260
cagcgtggaa atcgagtggg agctgcagaa ggaaaacagc aagcgctgga acccggagat  4320
tcagtacacc tccaactttg aaaagcagac tggtgtggac tttgccgttg acagccaggg  4380
tgtttactct gagcctcgcc ctattggcac tcgttacctc acccgtaatc tgtaattgca  4440
tgttaatcaa taaaccggtt gattcgtttc agttgaactt tggtctcctg tgcttcttat  4500
cttatcggtt tccatagcaa ctggttacac attaactgct tgggtgcgct tcacgataag  4560
aacactgacg tcaccgcggt acccctagtg atggagttgg ccactccctc tatgcgcgct  4620
cgctcgctcg gtgggggcctg cggaccaaag gtccgcagac ggcagagctc tgctctgccg  4680
gcccccacga gcgagcgagc gcgcatagag ggagtggcca a                       4721
```

<210> 2
<211> 737
<212> PRT
<213> 腺伴随病毒血清型7的衣壳蛋白

<400> 2

```
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 Asp Leu Lys Pro Gly Ala Pro Lys Pro
            20                  25                  30
Lys Ala Asn Gln Gln Lys Gln Asp Asn Gly Arg Gly Leu Val Leu Pro
        35                  40                  45
Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro
    50                  55                  60
```

[0003]

```
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 Arg Tyr Asn His Ala
                85                  90                  95
Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly
            100                 105                 110
Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
        115                 120                 125
Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Ala Lys Lys Arg
    130                 135                 140
Pro Val Glu Pro Ser Pro Gln Arg Ser Pro Asp Ser Ser Thr Gly Ile
145                 150                 155                 160
Gly Lys Lys Gly Gln Gln Pro Ala Arg Lys Arg Leu Asn Phe Gly Gln
                165                 170                 175
Thr Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Leu Gly Glu Pro
            180                 185                 190
Pro Ala Ala Pro Ser Ser Val Gly Ser Gly Thr Val Ala Ala Gly Gly
        195                 200                 205
Gly Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn
    210                 215                 220
Ala Ser Gly Asn Trp His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val
225                 230                 235                 240
Ile Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His
                245                 250                 255
Leu Tyr Lys Gln Ile Ser Ser Glu Thr Ala Gly Ser Thr Asn Asp Asn
            260                 265                 270
Thr 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 Lys Leu Arg Phe Lys Leu Phe Asn Ile
305                 310                 315                 320
Gln Val Lys Glu Val Thr Thr Asn Asp Gly Val Thr Thr Ile Ala Asn
                325                 330                 335
Asn Leu Thr Ser Thr Ile Gln Val Phe Ser Asp Ser Glu Tyr Gln Leu
            340                 345                 350
Pro Tyr Val Leu Gly Ser Ala His Gln 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 Asn
    370                 375                 380
Gly Ser Gln Ser 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 Glu Phe Ser Tyr Ser
                405                 410                 415
Phe Glu Asp 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 Ala
        435                 440                 445
Arg Thr Gln Ser Asn Pro Gly Gly Thr Ala Gly Asn Arg Glu Leu Gln
    450                 455                 460
Phe Tyr Gln Gly Gly Pro Ser Thr Met Ala Glu Gln Ala Lys Asn Trp
465                 470                 475                 480
Leu Pro Gly Pro Cys Phe Arg Gln Gln Arg Val Ser Lys Thr Leu Asp
                485                 490                 495
Gln Asn Asn Asn Ser Asn Phe Ala Trp Thr Gly Ala Thr Lys Tyr His
            500                 505                 510
Leu Asn Gly Arg Asn Ser Leu Val Asn Pro Gly Val Ala Met Ala Thr
        515                 520                 525
His Lys Asp Asp Glu Asp Arg Phe Phe Pro Ser Ser Gly Val Leu Ile
    530                 535                 540
Phe Gly Lys Thr Gly Ala Thr Asn Lys Thr Thr Leu Glu Asn Val Leu
545                 550                 555                 560
```

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

<210> 3
<211> 623
<212> PRT
<213> 豚伴随病毒血清型 7 的 rep 蛋白

[0004] <400> 3
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
Val Gln 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 Leu 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 Val Gln Thr Ile Tyr Arg Gly Val 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 Glu Tyr Ile
165 170 175
Ser Ala Cys Leu Asn Leu Ala 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 Leu 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 Arg 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 Ala Leu Thr Lys Ser Ala Pro Asp Tyr Leu Val Gly Pro Ser
275 280 285
Leu Pro Ala Asp Ile Lys Thr Asn Arg Ile Tyr Arg Ile Leu Glu Leu
290 295 300
Asn Gly Tyr Asp Pro Ala Tyr Ala Gly Ser Val Phe Leu Gly Trp Ala
305 310 315 320
Gln 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 Ala 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 Glu His Asp Phe Gly Lys Val Thr Lys Gln
450 455 460
Glu Val Lys Glu Phe Phe Arg Trp Ala Ser Asp His Val Thr Glu Val
465 470 475 480
Ala His Glu Phe Tyr Val Arg Lys Gly Gly Ala Ser Lys Arg Pro Ala
485 490 495
[0005] Pro Asp Asp Ala Asp Ile Ser Glu Pro Lys Arg Ala Cys Pro Ser Val
500 505 510
Ala Asp Pro Ser Thr Ser Asp Ala Glu Gly Ala Pro Val Asp Phe Ala
515 520 525
Asp Arg Tyr Gln Asn Lys Cys Ser Arg His Ala Gly Met Ile Gln Met
530 535 540
Leu Phe Pro Cys Lys Thr Cys Glu Arg Met Asn Gln Asn Phe Asn Ile
545 550 555 560
Cys Phe Thr His Gly Val Arg Asp Cys Leu Glu Cys Phe Pro Gly Val
565 570 575
Ser Glu Ser Gln Pro Val Val Arg Lys Lys Thr Tyr Arg Lys Leu Cys
580 585 590
Ala Ile His His Leu Leu Gly Arg Ala Pro Glu Ile Ala Cys Ser Ala
595 600 605
Cys Asp Leu Val Asn Val Asp Leu Asp Asp Cys Val Ser Glu Gln
610 615 620

<210> 4
<211> 4393
<212> DNA
<213> 腺伴随病毒血清型 8

<400> 4
cagagaggga gtggccaact ccatcactag gggtagcgcg aagcgcctcc cacgctgccg 60
cgtcagcgct gacgtaaatt acgtcatagg ggagtggtcc tgtattagct gtcacgtgag 120
tgcttttgcg gcattttgcg acaccacgtg gccatttgag gtatatatgg ccgagtgagc 180
gagcaggatc tccattttga ccgcgaaatt tgaacgagca gcagccatgc cgggcttcta 240
cgagatcgtg atcaaggtgc cgagcgacct ggacgagcac ctgccgggca tttctgactc 300
gtttgtgaac tgggtggccg agaaggaatg ggagctgccc ccggattctg acatggatcg 360
gaatctgatc gaggcaggcc cctgaccgt ggccgagaag ctgcagcgcg acttcctggt 420
ccaatggcgc cgcgtgagta aggccccgga ggccctcttc tttgtttcagt tcgagaaggg 480

[0006]

cgagagctac tttcacctgc acgttctggt cgagaccacg ggggtcaagt ccatggtgct 540
aggccgcttc ctgagtcaga ttcgggaaaa gcttggtcca gaccatctac ccgcgggctc 600
gagccccacc ttgcccaact ggttcgcggt gaccaaagac gcggtaatgg cgccggcggg 660
ggggaacaag gtggtggacg agtgctacat ccccaactac ctcctgccca agactcagcc 720
cgagctgcag tgggcgtgga ctaacatgga ggagtatata agcgcgtgct tgaacctggc 780
cgagcgcaaa cggctcgtgg cgcagcacct gaccccacgt agccagacgc aggagcagaa 840
caaggagaat ctgaaccccg attctgacgc gccccgtgat caggtcaaaa acctccgcgc 900
ctatatggag ctggtcgggt ggctggtgga ccgggggcat cacctccgaga agcagtggat 960
ccaggaggac caggcctcgt acatctcctt caacgccgcc tccaactcgc ggtccccagat 1020
caaggccgcg ctggacaatg ccggcaagat catggcgctg accaaatccg cgccccgacta 1080
cctggtgggg ccctcgctgc ccgcggacat tacccagaac cgcatctacc gcatcctcgc 1140
tctcaacggc tacgaccctg cctacgccgg ctccgtcttt ctcggcatgg ctcagaaaaa 1200
gttcgggaaa cgcaacacca tctggctgtt tggacccgcc accaccggca agaccaacat 1260
tgcggcagcc atcgcccacg ccgtgccctt ctacggctgc gtcaactgga ccaatgagaa 1320
ctttcccttc aatgattgcg tcgacaagat ggtgatctgg tgggaggagg gcaagatgac 1380
ggccaaggtc gtggagtccg ccaaggccat tctcggcggc agcaaggtgc gcgtggacca 1440
aaagtgccaag tcgtccgccc agatcgaccc caccccccgtg atcgtcacct ccaacaccaa 1500
catgtgcgcc gtgattgacg ggaacagcac caccttcgag caccagcagc ctctccagga 1560
ccggatgttt aagttcgaac tcacccgccg tctggagcac gactttggca aggtgacaaa 1620
gcaggaagtc aaagagttct tccgctgggc cagtgatcac gtgaccgagg tggcgcatga 1680
gttttacgtc agaaagggcg gagccagcaa aagacccgcc ccccgatgacg cggataaaag 1740
cgagccccag cgggcctgcc cctcagtcgc ggatccatcg acgtcagacg cggaaggagc 1800
tccggtggac tttgccgacc ggtaccaaaa caaatgttct cgtcacgcgg gcatgcttca 1860
gatgctgttt ccctgcaaaa cgtgcgagag aatgaatcag aatttcaaca tttgcttcac 1920
accacggggtc agagactgct cagagtgttt ccccggcgtg tcagaatctc aaccggtcgt 1980
cagaaagagg acgtatcgga aactctgtgc gattcatcat ctgctggggc gggctcccga 2040
gattgcttgc tcggcctgcg atctggtcaa cgtggacctg gatgactgtg tttctgagca 2100
ataaatgact taaaccaggt atggctgccg atggttatct tccagattgg ctcgaggaca 2160
acctctctga gggcattcgc gagtggtggg cgctgaaacc tggagccccg aagcccaaag 2220
ccaaccagca aaagcaggac gacggccggg gtctggtgcc tcctggctac aagtacctcg 2280
gacccttcaa cggactcgac aagggggagc ccgtcaacgc ggcggacgca gcggccctcg 2340
agcacgacaa ggcctacgac cagcagctgc aggcgggtga caatccgtac ctgcggtata 2400
accacgccga cgccgagttt caggagcgtc tgcaagaaga tacgtctttt gggggcaacc 2460
tcgggcgagc agtcttccag gccaagaagc gggttctcga acctctcggt ctggttgagg 2520
aaggcgctaa gacggctcct ggaaagaaga gaccggtaga gccatcaccc cagcgttctc 2580
cagactcctc tacgggcatc ggcaagaaag gccaacagcc cgccagaaaa agactcaatt 2640
ttggtcagac tggcgactca gagtcagttc cagacctca acctctcgga gaacctccag 2700
cagcgcctc tggtgtggga cctaatacaa tggctgcagg cggtggcgca ccaatggcag 2760
acaataacga aggcgccgac ggagtgggta gttcctcggg aaattggcat tgcgattcca 2820
catggctggg cgacagagtc atcaccacca gcacccgaac ctgggccctg cccacctaca 2880
acaaccacct ctacaagcaa atctccaacg ggacatcggg aggagccacc aacgacaaca 2940
cctacttcgg ctacagcacc cctggggggt attttgactt taacagattc cactgcccact 3000
tttcaccacg tgactggcag cgactcatca acaacaactg gggatccgg cccaagagac 3060
tcagcttcaa gctcttcaac atccaggtca aggaggtcac gcagaatgaa ggcaccaaga 3120
ccatcgccaa taacctcacc agcaccatcc aggtgtttac ggactcggag taccagctgc 3180
cgtacgttct cggctctgcc caccagggct gcctgcctcc gttcccggcg gacgtgttca 3240
tgattcccca gtacggctac ctaacactca acaacggtag tcaggccgtg ggacgctcct 3300
ccttctactg cctggaatac tttccttcgc agatgctgag aacggcaac aacttccagt 3360
ttacttacac cttcgaggac gtgcctttcc acagcagcta cgcccacagc cagagcttgg 3420
accggctgat gaatcctctg attgaccagt acctgtacta cttgtctcgg actcaaacaa 3480
caggaggcac ggcaaatacg cagactctgg gcttcagcca aggtgggcct aatacaatgg 3540
ccaatcaggc aaagaactgg ctgccaggac cctgttaccg ccaacaacgc gtctcaacga 3600
caaccgggca aaacaacaat agcaactttg cctggactgc tgggaccaaa taccatctga 3660
atggaagaaa ttcattggct aatcctggca tcgctatggc aacacacaaa gacgacgagg 3720
agcgtttttt tccagtaac gggatcctga ttttttggcaa acaaaatgct gccagagaca 3780
atgcggatta cagcgatgtc atgctcaccaa gcgaggaaga aatcaaaacc aactaccctg 3840
tggctacaga ggaatacggt atcgtggcag ataacttgca gcagcaaaac acggctcctc 3900
aaattggaac tgtcaacagc cagggggcct taccoggtat ggtctggcag aaccgggacg 3960
tgtacctgca gggtcccatc tgggccaaga ttcctcacac ggacggcaac ttccaccogt 4020
ctccgctgat gggcggcttt ggcctgaaac atcctccgcc tcagatcctg atcaagaaca 4080
cgcctgtacc tgcggatcct ccgaccacct tcaaccagtc aaagctgaac tctttcatca 4140
cgcaatacag caccggacag gtcagcgtgg aaattgaatg ggagctgcag aaggaaaaca 4200

gcaagcgctg gaaccccgag atccagtaca cctccaacta ctacaaatct acaagtgtgg 4260
actttgctgt taatacagaa ggcgtgtact ctgaaccccg ccccattggc accegttacc 4320
tcaccegtaa tctgtaattg cctgttaatc aataaaccgg ttgattcgtt tcagttgaac 4380
tttggtctct gcg 4393

<210> 5
<211> 4385
<212> DNA
<213> 腺伴随病毒血清型 9

<400> 5
cagagagggа gtggccaact ccatcactag gggtaatcgc gaagcgcctc ccacgctgcc 60
gcgtcagcgc tgacgtagat tacgtcatag gggagtggtc ctgtattagc tgtcacgtga 120
gtgcttttgc gacattttgc gacaccacat ggccatttga ggtatatatg gccgagtgag 180
cgagcaggat ctccattttg accgcgaaat ttgaacgagc agcagccatg ccgggcttct 240
acgagattgt gatcaaggtg ccgagcgacc tggacgagca cctgccgggc atttctgact 300
cttttgtgaa ctgggtggcc gagaaggaat gggagctgcc cccggattct gacatggatc 360
ggaatctgat cgagcaggca ccccctgacc tggccgagaa gctgcagcgc gacttcctgg 420
tccaatggcg ccgcgtgagt aaggccccgg aggccctctt ctttgttcag ttcgagaagg 480
gcgagagcta ctttcacctg cacgttctgg tcgagaccac gggggtcaag tccatggtgc 540
taggccgctt cctgagtcag attcgggaga agctggtcca gaccatctac cgcgggatcg 600
agcgcaccct gaccaactgg ttcgcggtga ccaagacgcg taatggcgcc ggcgggggcga 660
acaaggtggt ggacgagtgc tacatccccа actacctcct gcccaagact cagccccgag 720
tgcagtgggc gtggactaac atggaggagt atataagcgc gtgcttgaac ctggccgagc 780
gcaaacggct cgtggcgcag cacctgaccc acgtcagcca gacgcaggag cagaacaagg 840
agaatctgaa ccccaattct gacgcgcccg tgatcaggtc aaaaacctcc gcgcgctaca 900
tggagctggt cgggtggctg gtggaccggg gcatcaccte cgagaagcag tggatccagg 960
aggaccaggc ctcgtacatc tccttcaacg ccgcctccaa ctcgcggtcc cagatcaagg 1020
ccgcgctgga caatgccggc aagatcatgg cgctgaccaa atccgcgccc gactacctgg 1080
taggcccttc atttcggtg gacattacgc agaaccgcat ctaccgcatc ctgcagctca 1140
acggctacga ccctgtctac gccggctccg tctttctcgg ctgggcacaa aagaagttcg 1200
ggaaacgcaa caccatctgg ctgtttgggc cggccaccac gggaaagacc aacatcgcag 1260
aagccattgc ccacgccgtg ccctttctacg gctgcgtcaa ctggaccaat gagaacttte 1320
ccttcaacga ttgcgtcgac aagatggtga tctggtggga ggagggcaag atgacggcca 1380
aggtcgtgga gtccgccaag gccattctcg acggcagcaa ggtgcgcgtg gaccaaaagt 1440
gcaagtcgtc cgccccagatc gacccccactc ccgtgatcgt cacctccaac accaacatgt 1500
gcgccgtgat tgacgggaac agcaccacct tcgagcacca gcagcctctc caggaccgga 1560
tgtttaagtt cgaactcacc cgccgtctgg agcacgactt tggcaaggtg acaaagcagg 1620
aagtcaaaga gttcttccgc tgggccagtg atcacgtgac cgaggtggcg catgagtttt 1680
acgtcagaaa gggcggagcc agcaaaagac ccgccccccga tgacgcggat aaaagcgagc 1740
ccaagcgggc ctgccctca gtcgcggatc catcgacgtc agacgcggaa ggagctccgg 1800
tggactttgc cgacaggtac caaaacaaat gttctcgtca cgcgggcatg cttcagatgc 1860
tgcttcctg caaaacgtgc gagagaatga atcagaattt caacattgc ttcacacacg 1920
gggtcagaga ctgctcagag tgtttccccg gcgtgtcaga atctcaaccg gtcgtcagaa 1980
agaggacgta tcggaaactc tgtgcgattc atcatctgct ggggcgggct cccgagattg 2040
cttgctcggc ctgcgatctg gtcaacgtgg acctggatga ctgtgtttct gagcaataaa 2100
tgacttaaac caggtatggc tgccgatggt tattttccag attggctcga ggacaacctc 2160
tctgagggca ttcgcgagtg gtgggcgctg aaacctggag ccccgaagcc caaagccaac 2220
cagcaaaagc aggacgacgg ccggggtctg gtgcttcctg gctacaagta cctcggaccc 2280
ttcaacggac tcgacaaggg ggagccccgtc aacgcggcgg acgcagcggc cctcgagcac 2340
ggcaaggcct acgaccagca gctgcaggcg ggtgacaatc cgtacctgcg gtataaccac 2400
gccgacgccg agtttcagga gcgtctgcaa gaagatacgt cttttggggg caacctcggg 2460
cgagcagtct tccaggccaa gaagcgggtt ctcgaacctc tcggtctggt tgaggaaggc 2520
gctaagacgg ctcctggaaa gaagagaccg gtagagccat cacccagcg ttctccagac 2580
tctctacgg gcatcggcaa gaaaggccaa cagccgcca gaaaaagact caattttggt 2640
cagactgggg actcagagtc agttccagac cctcaacctc tcggagaacc tccagcagcg 2700
ccctctggtg tgggacctaa tacaatgggt gcaggcggtg gcgcaccaat ggcagacaat 2760
aacgaaggcg ccgacggagt gggtaattcc tcgggaaatt ggcattgcga ttccacatgg 2820
ctgggggaca gagtcatcac caccagcacc cgaacctggg cattgcccac ctacaacaac 2880
cacctctaca agcaaatctc caatggaaca tcgggaggaa gcaccaacga caacacctac 2940
tttggctaca gcaccccctg ggggtatttt gacttcaaca gattccactg ccacttctca 3000
ccacgtgact ggcagcgact catcaacaac aactggggat tccggccaaa gagactcaac 3060

[0007]

ttcaagctgt tcaacatcca ggtcaaggag gttacgacga acgaaggcac caagaccatc 3120
gccaataacc ttaccagcac cgtccaggtc tttacggact cggagtacca gctaccgtac 3180
gtcctaggct ctgcccacca aggatgcctg ccaccgtttc ctgcagacgt cttcatggtt 3240
cctcagtacg gctacctgac gctcaacaat ggaagtcaag cgttaggacg ttcttctttc 3300
tactgtctgg aatacttccc ttctcagatg ctgagaaccg gcaacaactt tcagttcagc 3360
tacactttcg aggacgtgcc tttccacagc agctacgcac acagccagag tctagatcga 3420
ctgatgaacc cctcatcga ccagtacctа tactacctgg tcagaacaca gacaactgga 3480
actgggggaa ctcaaacttt ggcatcagc caagcaggcc ctagctcaat ggccaatcag 3540
gctagaaact gggtacccgg gccttgctac cgtcagcagc gcgtctccac aaccaccaac 3600
caaaataaca acagcaactt tgcgtggacg ggagctgcta aattcaagct gaacgggaga 3660
gactcgctaa tgaatcctgg cgtggctatg gcatcgcaca aagacgacga ggaccgcttc 3720
tttccatcaa gtggcgtttt catatttggc aagcaaggag ccgggaacga tggagtcgac 3780
tacagccagg tgctgattac agatgaggaa gaaattaaag ccaccaaccc tgtagccaca 3840
gaggaatacg gagcagtggc catcaacaac caggccgcta acacgcaggc gcaaaactgga 3900
cttgtgcata accagggagt tattcctggt atggtctggc agaaccggga cgtgtacctg 3960
cagggcccta tttgggctaa aataccтcac acagatggca actttcaccc gtctcctctg 4020
atgggtggat ttggactgaa acacccacct ccacagattc taattaaaaa tacaccagtg 4080
ccggcagatc ctcctcttac cttcaatcaa gccaagctga actctttcat cacgcagtac 4140
agcacgggac aagtcagcgt ggaaatcgag tgggagctgc agaaagaaaa cagcaagcgc 4200
tggaatccag agatccagta tacttcaaac tactacaaat ctacaaatgt ggactttgct 4260
gtcaatacca aaggtgttta ctctgagcct cgccccattg gtactcgtta cctcacccgt 4320
aatttgtaat tgcctgttaa tcaataaacc ggttaatttg tttcagttga actttggtct 4380
ctgcg 4385

<210> 6
<211> 4718
<212> DNA
<213> 腺伴随病毒血清型 1

<400> 6

[0008]

ttgcccactc cctctctgcg cgctcgctcg ctcggtgggg cctgcggacc aaaggtccgc 60
agacggccaga gctctgctct gccggcccca ccgagcgagc gagcgcgcag agagggagtg 120
ggcaactcca tcactagggg taatcgcgaa gcgcctccca cgctgccgcg tcagcgctga 180
cgtaaattac gtcataggggg agtggtcctg tattagctgt cacgtgagtg cttttgcgac 240
attttgcgac accacgtggc catttagggt atatatggcc gagtgagcga gcaggatctc 300
cattttgacc gcgaaatttg aacgagcagc agccatgccg ggcttctacg agatcgtgat 360
caaggtgccg agcgacctgg acgagcacct gccgggcatt tctgactcgt ttgtgagctg 420
ggtggccgag aaggaatggg agctgccccc ggattctgac atggatctga atctgattga 480
gcaggcaccc ctgaccgtgg ccgagaagct gcagcgcgac ttcctggtcc aatggcgccg 540
cgtgagtaag gccccggagg ccctcttctt tgttcagttc gagaagggcg agtcctactt 600
ccacctccat attctggtgg agaccacggg ggtcaaatcc atggtgctgg gcgcttcct 660
gagtcagatt agggacaagc tggtgcagac catctaccgc gggatcgagc cgaccctgcc 720
caactggttc gcggtgacca agacgcgtaa tggcgccgga gggggaaca aggtggtgga 780
cgagtgctac atccccaact acctcctgcc caagactcag cccgagctgc agtgggcgtg 840
gactaacatg gaggagtata tagcgccctg tttgaacctg gccgagcgca aacggctcgt 900
ggcgcagcac ctgacccacg tcagccagac ccaggagcag aacaaggaga atctgaaccc 960
caatttctgac gcgcctgtca tccggtcaaa aacctccgcg cgctacatgg agctggtcgg 1020
gtggctggtg gaccggggca tcacctccga gaagcagtgg atccaggagg accaggcctc 1080
gtacatctcc ttcaacgccg cttccaactc gcggtcccag atcaaggccg ctctggacaa 1140
tgccggcaag atcatggcgc tgaccaaatc cgcgccccgac tacctggtag gccccgctcc 1200
gccgcggac attaaaacca accgcatcta ccgcatcctg gagctgaacg gctacgaacc 1260
tgcctacgcc ggctccgtct ttctcggctg ggcccagaaa aggttcggga agcgcaacac 1320
catctggctg tttgggccgg ccaccacggg caagaccaac atcgcggaag ccatcgccca 1380
cgccgtgccc ttctacggct gcgtcaactg gaccaatgag aactttcctt tcaatgattg 1440
cgtcgacaag atggtgatct ggtgggagga gggcaagatg acggccaagg tcgtggagtc 1500
ggccaaggcc attctcggcg gcagcaaggt gcgcgtggac caaaagtgca agtcgtccgc 1560
ccagatcgac cccaccccca tgatcgtcac ctccaacacc aacatgtgcg ccgtgattga 1620
cgggaacagc accaccttcg agcaccagca gccgttgcag gaccggatgt tcaaatttga 1680
actcaccccc cgtctggagc atgacttigg caaggtgaca aagcaggaag tcaaagagtt 1740
cttccgctgg gcgcaggatc acgtgaccga ggtggcgcat gagttctacg tcagaaaggg 1800
tggagccaac aaaagacccg ccccccgatga cgcggataaa agcgagccca agcgggcctg 1860
ccccтcagtc gcggatccat cgacgtcaga cgcggaagga gctccggtgg actttgccga 1920

[0009]

```
caggtaccaa aacaaaatgtt ctcgtcacgc gggcatgctt cagatgctgt ttcctgcaa   1980
gacatgcgag agaatgaatc agaatttcaa catttgcttc acgcacggga cgagagactg   2040
ttcagagtgc ttccccggcg tgtcagaatc tcaaccggtc gtcagaaaga ggacgtatcg   2100
gaaactctgt gccattcatc atctgctggg gcgggctccc gagattgctt gctcggcctg   2160
cgatctggtc aacgtggacc tggatgactg tgtttctgag caataaatga cttaaaccag   2220
gtatggctgc cgatggttat cttccagatt ggctcgagga caacctctct gagggcattc   2280
gcgagtggtg ggacttgaaa cctggagccc cgaagcccaa agccaaccag caaaagcagg   2340
acgacggccg gggtctggtg cttcctggct acaagtacct cggacccttc aacggactcg   2400
acaaggggga gcccgtcaac gcggcggacg cagcggcgct cgagcacgac aaggcctacg   2460
accagcagct caaagcgggt gacaatccgt acctgcggta taaccacgcc gacgccgagt   2520
ttcaggagcg tctgcaagaa gatacgtctt ttgggggcaa cctcgggcga gcagtcttcc   2580
aggccaagaa gcgggttctc gtaccctctg gtctggttga ggaaggcgct aagacggctc   2640
ctggaaagaa acgtccggta gagcagtcgc cacaagagcc agactcctcc tcgggcatcg   2700
gcaagacagg ccagcagccc gctaaaaaga gactcaattt tggtcagact ggcgactcag   2760
agtcagtccc cgatccacaa cctctcggag aacctccagc aacccccgct gctgtgggac   2820
ctactacaat ggcttcaggc ggtggcgcac caatggcaga caataacgaa ggcgccgacg   2880
gagtgggtaa tgcctcagga aattggcatt gcgattccac atggctgggc gacagagtca   2940
tcaccaccag cacccgcacc tgggccttgc ccacctacaa taaccacctc tacaagcaaa   3000
tctccagtgc ttcaacgggg gccagcaacg acaaccacta cttcggctac agcaccccct   3060
gggggtattt tgatttcaac agattccact gccacttttc accacgtgac tggcagcgac   3120
tcatcaacaa caattgggga ttccggccca agagactcaa cttcaaactc ttcaacatcc   3180
aagtcaagga ggtcacgacg aatgatggcg tcacaaccat cgctaataac cttaccagca   3240
cggttcaagt cttctcggac tcggagtacc agttccgtta cgtcctcggc tctgcgcacc   3300
agggctgcct ccctccgttc ccggcggacg tgttcatgat tccgcaatac ggctacctga   3360
cgctcaacaa tggcagccaa gccgtgggac gttcatcctt ttactgcctg gaatatttcc   3420
cttctcagat gctgagaacg ggcaacaact ttaccttcag ctacaccttt gaggaagtgc   3480
ctttccacag cagctacgcg cacagccaga gcctggaccg gctgatgaat cctctcatcg   3540
accaatacct gtattacctg aacagaactc aaaatcagtc cggaagtgcc caaaacaagg   3600
actgctgtt tagccgtggg tctccagctg gcatgtctgt tcagcccaaa aactggctac   3660
ctggaccctg ttatcggcag cagcgcgttt ctaaaacaaa aacagacaac aacaacagca   3720
attttacctg gactggtgct tcaaaatata acctcaatgg gcgtgaatcc atcatcaacc   3780
ctggcactgc tatggcctca cacaaagacg acgaagacaa gttctttccc atgagcggtg   3840
tcatgatttt tggaaaagag agcgccggag cttcaaacac tgcattggac aatgtcatga   3900
ttacagacga agaggaaatt aaagccacta accctgtggc caccgaaaga tttgggaccg   3960
tggcagtcaa tttccagagc agcagcacag accctgcgac cggagatgtg catgctatgg   4020
gagcattacc tggcatggtg tggcaagata gagacgtgta cctgcagggt ccccatttggg   4080
ccaaaattcc tcacacagat ggacacttte accegtetee tettatgggc ggctttggac   4140
tcaagaaccc gcctcctcag atcctcatca aaaacacgcc tgtcctgcg aatcctccgg   4200
cggagtttte agctacaaag tttgctteat tcatcaccca atactccaca ggacaagtga   4260
gtgtggaaat tgaatgggag ctgcagaaag aaaacagcaa gcgctggaat ccegaagtgc   4320
agtacacatc caattatgca aaatctgcca acgttgattt tactgtggac aacaatggac   4380
tttatactga gcctcgcccc attggcaccc gttaccttac ccgtcccctg taattacgtg   4440
ttaatcaata aacggttga ttcgtttcag ttgaacttg gtctcctgtc cttcttatct   4500
tatcggttac catggttata gcttacacat taactgcttg gttgcgcttc gcgataaaag   4560
acttacgtca tcgggttacc cctagtgatg gagttgccca ctccctctct gcgcgctcgc   4620
tcgctcggtg gggcctgcgg accaaaggtc cgcagacggc agagctctgc tctgccggcc   4680
ccacgagcg agcgagcgcg cagagaggga gtgggcaa                            4718

<210> 7
<211> 4675
<212> DNA
<213> 腺伴随病毒血清型 2

<400> 7
ttggccactc cctctctgcg cgctcgctcg ctcactgagg ccgggcgacc aaaggtcgcc     60
cgacgcccgg gctttgcccg ggcggcctca gtgagcgagc gagcgcgcag agagggagtg    120
ggccaactcca tcactagggg ttcctggagg ggtggagtcg tgacgtgaat tacgtcatag    180
ggttagggag gtcctgtatt agaggtcacg tgagtgtttt gcgacatttt gcgacaccat    240
gtggtcacgc tgggtattta agcccgagtg agcacgcagg gtctccattt tgaagcggga    300
gatttgaacg cgcagccgcc atgccggggt tttacgagat tgtgattaag gtccccagcg    360
accttgacgg gcatctgccc ggcatttctg acagctttgt gaactgggtg gccgagaagg    420
aatgggagtt gccgccagat tctgacatgg atctgaatct gattgagcag gcacccctga    480
```

[0010]

ccgtggccga gaagctgcag cgcgactttc tgacggaatg gcgccgtgtg agtaaggccc 540
cggaggccct tttctttgtg caatttgaga agggagagag ctacttccac atgcacgtgc 600
tcgtggaaac caccggggtg aaatccatgg ttttgggacg tttcctgagt cagattcgcg 660
aaaaactgat tcagagaatt taccgcggga tcgagccgac tttgccaaac tggttcgcgg 720
tcacaaagac cagaaatggc gccggaggcg ggaacaaggt ggtggatgag tgctacatcc 780
ccaattactt gctcccccaaa accccagcctg agctccngtg ggcgtggact aatatggaac 840
agtatttaag cgcctgtttg aatctcacgg agcgtaaacg gttggtggcg cagcatctga 900
cgcacgtgtc gcagacgcag gagcagaaca aagagaatca gaatcccaat tctgatgcgc 960
cggtgatcag atcaaaaact tcagccaggt acatggagct ggtcgggtgg ctcgtggaca 1020
aggggattac ctcggagaag cagtggatcc aggaggacca ggcctcatac atctccttca 1080
atgcggcctc caactcgcgg tcccaaatca aggctgcctt ggacaatgcg ggaaagatta 1140
tgagcctgac taaaaccgcc ccgactacc tggtgggcca gcagccccgtg gaggacattt 1200
ccagcaatcg gatttataaa attttggaac taaacgggta cgatccccaa tatgcggctt 1260
ccgtctttct gggatgggcc acgaaaaagt tcggcaagag gaacaccatc tggctgtttg 1320
ggcctgcaac taccgggaag accaacatcg cggaggccat agcccacact gtgccctttct 1380
acgggtgcgt aaactggacc aatgagaact ttccctcaa cgactgtgtc gacaagatgg 1440
tgatctggtg ggaggagggg aagatgaccg ccaaggtcgt ggagtcggcc aaagccattc 1500
tcggaggaag caaggtgcgc gtggaccaga aatgcaagtc ctcggcccag atagacccga 1560
ctcccgtgat cgtcacctcc aacaccaaca tgtgcgccgt gattgacggg aactcaacga 1620
ccttcgaaca ccagcagccg ttgcaagacc ggatgttcaa atttgaactc acccgccgtc 1680
tggatcatga ctttgggaag gtcaccaagc aggaagtcaa agacttttic cggtgggcaa 1740
aggatcacgt ggttgaggtg gagcatgaat tctacgtcaa aaagggtgga gccaagaaaa 1800
gacccgcccc cagtgacgca gatataagtg agcccaaacg ggtgcgcgag tcngttgcgc 1860
agccategac gtcagacgcg gaagcttcga tcaactacgc agacaggtac caaaacaaat 1920
gttctcgtca cgtgggcatg aatctgatgc tgtttcccig cagacaatgc gagagaatga 1980
atcagaattc aaatatctgc ttcactcacg gacagaaaga ctgtttagag tgctttcccg 2040
tgtcagaatc tcaacccgtt tctgtcgtca aaaaggcgta tcagaaactg tgctacattc 2100
atcatatcat gggaaaggtg ccagacgctt gcactgcctg cgatctggtc aatgtggatt 2160
tggatgacig catctttgaa caataaatga tttaaatcag gtatggcigc cgatggttat 2220
cttccagatt ggctcgagga cactctcict gaaggaataa gacagtggtg gaagctcaaa 2280
cctggcccac caccaccaaa gcccgcagag cggcataagg acgacagcag gggtcttgtg 2340
cttcctgggt acaagtacct cggacccttc aacggactcg acaagggaga gccggtcaac 2400
gaggcagacg ccgcggccct cgagcacgta caaagcctac gaccggcagc tcgacagcgg 2460
agacaacccg tacctcaagt acaaccacgc cgacgcggag tttcaggagc gccttaaaga 2520
agatacgtct tttgggggca acctcggacg agcagtcttc caggcgaaaa agagggttct 2580
tgaacctctg ggcctggttg aggaacctgt taagacggct ccgggaaaaa agaggccggt 2640
agagcactct cctgtggagc cagactcctc ctcgggaacc ggaaaggcgg gccagcagcc 2700
tgcaagaaaa agattgaatt ttggtcagac tggagacgca gactcagtac ctgaccccca 2760
gcctctcgga cagccaccag cagccccctc tggtctggga actaatacga tggctacagg 2820
cagtggcgca ccaatggcag acaataacga gggcgccgac ggagtgggta attcctccgg 2880
aaattggcat tgcgattcca catggatggg cgacagagtc atcaccacca gcaccogaac 2940
ctgggccctg cccacctaca acaaccacct ctacaaacaa atttccagcc aatcaggagc 3000
ctcgaacgac aatcactact ttggctacag cacccottgg gggtatttig acttcaacag 3060
attccactgc cacttttcac cacgtgactg gcaaagactc atcaacaaca actggggatt 3120
ccgacccaag agactcaact tcaagctctt tnacattcaa gtcaaagagg tcacgcagaa 3180
tgacggtacg acgacgattg ccaataacct taccagcacg gttcaggtgt ttactgactc 3240
ggagtaccag ctcccgtacg tcctcggctc ggcgcatcaa ggatgcctcc cgccgttccc 3300
agcagacgtc ttcatggtgc cacagtatgg atacctcacc ctgaacaacg ggagtcaggc 3360
agtaggacgc tcttcatttt actgcctgga gtacttcct tctcagatgc tgcgtaccgg 3420
aaacaacttt accttcagct acactttttga ggacgttcct ttccacagca gctacgctca 3480
cagccagagt ctggaccgtc tcatgaatcc tctcatcgac cagtacctgt attacttgag 3540
cagaacaaac actccaagtg gaaccaccac gcagtcaagg cttcagtttt ctcaggccgg 3600
agcgagtgac attcgggacc agtctaggaa ctggcttcct ggaccctgtt accgtcagca 3660
gcgagtatca aagacatctg cggataacaa caacagtgaa tactcgtgga ctggagctac 3720
caagtaccac ctcaatggca gagactctct ggtgaatccg gccatggcaa gccacaagga 3780
cgatgaagaa aagttttttc ctcagagcgg ggttctcatc tttgggaagc aaggctcaga 3840
gaaaacaaat gtgaacattg aaaaggtcat gattacagac gaagaggaaa tcggaacaac 3900
caatcccgtg gctacggagc agtatggtic tgtatctacc aacctccaga gaggcaacag 3960
acaagcagct acgcagatg tcaacacaca aggcgttctt ccaggcatgg tctggcagga 4020
cagagatgtg taccttcagg ggcccatctg ggcaaagatt ccacacacgg acggacattt 4080
tcacccctct ccctcatgg gtggattcgg acttaaacac cctcctccac agattctcat 4140
caagaacacc ccggtacctg cgaatcttic gaccacctic agtgcggcaa agtttgcttc 4200

cttcatcaca cagtactcca cgggacacgg tcagcgtgga gatcgagtgg gagctgcaga 4260
aggaaaacag caaacgctgg aatcccgaaa ttcagtacac ttccaactac aacaagtctg 4320
ttaatcgtgg acttaccgtg gatactaatg gcgtgtattc agagcctcgc cccattggca 4380
ccagatacct gactcgtaat ctgtaattgc ttgttaatca ataaaccgtt taattcgttt 4440
cagttgaact ttggtctctg cgtatttctt tcttatctag tttccatggc tacgtagata 4500
agtagcatgg cgggttaatc attaactaca aggaacccct agtgatggag ttggccactc 4560
cctctctgcg cgctcgctcg ctcactgagg ccgggcgacc aaaggtcgcc cgacgcccgg 4620
gctttgcccg ggcggcctca gtgagcgagc gagcgcgcag agagggagtg gccaa 4675

<210> 8
<211> 4726
<212> DNA
<213> 腺伴随病毒血清型 3

<400> 8
ttggccactc cctctatgcg cactgctcgc ctcggtgggg cctggcgacc aaaggtcgcc 60
agacggacgt gctttgcacg tccggccccc ccgagcgagc gagtgcgcat agagggagtg 120
gccaactcca tcactagagg tatggcagtg acgtaacgcg aagcgcgcga agcgagacca 180
cgcctaccag ctgcgtcagc agtcaggtga ccctttgcgc acagtttgcg acaccacgtg 240
gccgctgagg gtatatattc tcgagtgagc gaaccaggag ctccattttg accgcgaaat 300
ttgaacgagc agcagccatg ccgggggttct acgagattgt cctgaaggtc ccgagtgacc 360
tggacgagcg cctgccgggc atttctaact cgtttgttaa ctgggtggcc gagaaggaat 420
gggacgtgcc gccggattct gacatggatc cgaatctgat tgagcaggca ccccctgccg 480
tggccgaaaa gcttcagcgc gagttcctgg tggagtggcg ccgcgtgagt aaggccccgg 540
aggccctctt ttttgtccag ttcgaaaagg gggagaccta cttccacctg cacgtgctga 600
ttgagaccat cggggtcaaa tccatggtgg tcggccgcta cgtgagccag attaaagaga 660
agctggtgac ccgcatctac cgcggggtcg agccgcagct tccgaactgg ttcgcggtga 720
ccaaaacgcg aaatggcgcc ggcggcggga acaaggtggt ggacgactgc tacatccccg 780
actacctgct ccccaagacc cagcccgagc tccagtgggc gtggactaac atggaccagt 840
attaagcgc ctgttlgaat ctcgcggagc gtaaacggct ggtggcgcag catcgacgc 900
acgtgtcgca gacgcaggag cagaacaaag agaatcagaa ccccaattct gacgcgccgg 960
tcatcaggtc aaaaacctca gccaggtaca tggagctggt cgggtggctg gtggaccgcg 1020
ggatcacgtc agaaaagcaa tggattcagg aggaccaggc ctcgtacatc tccttcaacg 1080
ccgcctccaa ctcgcggtcc cagatcaagg ccgcgctgga caatgcctcc aagatcatga 1140
gcctgacaaa gacggctccg gactacctgg tgggcagcaa cccgccggag gacattacca 1200
aaaatcggat ctaccaaatc ctggagctga acgggtacga tccgcagtac gcggcctccg 1260
tcttcctggg ctgggcgcaa aagaagtccg ggaagaggaa caccatctgg ctctttgggc 1320
cggccacgac gggtaaaacc aacatcgcgg aagccatcgc ccacgccgtg ccctttctacg 1380
gctgcgtaaa ctggaccaat gagaactttc cttcaacga ttgcgtcgac aagatggtga 1440
tctggtggga ggagggcaag atgacggcca agtcgtggga gagcgccaag gccattctgg 1500
gcggaagcaa ggtgcgcgtg gaccaaaagt gcaagtcatc ggcccagatc gaacccactc 1560
ccgtgatcgt cacctccaac accaacatgt gcgccgtgat tgacgggaac agcaccacct 1620
tcgagcatca gcagccgctg caggaccgga tgtttgaatt tgaacttacc cgccgtttgg 1680
accatgactt tgggaaggtc accaaacagg aagtaaagga ctttttccgg tgggcttccg 1740
atcacgtgac tgacgtggct catgagttct acgtcagaaa gggtggagct aagaaacgcc 1800
cgccctccaa tgacgcggat gtaagcgagc caaaacggga gtgcacgtca ctlgcgcagc 1860
cgacaacgtc agacgcggaa gcaccggcgg actacgcgga caggtaccaa aacaaatgtt 1920
ctcgtcacgt gggcatgaat ctgatgcttt ttccctgtaa aacatgcgag agaatgaatc 1980
aaatttccaa tgtctgtttt acgcatggtc aaagagactg tggggaatgc ttccctggaa 2040
tgtcagaatc tcaacccgtt tctgtcgtca aaaagaagac ttatcagaaa ctgtgtccaa 2100
ttcatcatat cctgggaagg gcacccgaga ttgcctgttc ggcctgcgat ttggccaatg 2160
tggacttgga tgactgtgtt tctgagcaat aaatgactta aaccaggtat ggctgctgac 2220
ggttatcttc cagattggct cgaggacaac ctttctgaag gcattcgtga gtggtgggct 2280
ctgaaacctg gagtccctca acccaaagcg aaccaacaac accaggacaa ccgtcggggt 2340
cttgtgcttc cgggttacaa atacctcgga cccggtaacg gactcgacaa aggagagccg 2400
gtcaacgagg cggacgcggc agcctcgaa cacgacaaag cttacgacca gcagctcaag 2460
gcgggtgaca accccgtacct cangtacaac cacgccgacg ccgagtttca ggagcgtctt 2520
caagaagata cgtctttttgg gggcaacctt ggcagagcng tcttccaggc caaaaagagg 2580
atccttgagc ctcttggtct ggttgaggaa gcagctaaaa cggctcctgg aaagaagggg 2640
gctgtagatc agtctcctca ggaaccggac tcatcatctg gtgttggcaa ategggcaaa 2700
cagcctgcca gaaaaagact aaatttcggt cagactggag actcagagtc agtccccagac 2760
cctcaacctc tcggagaacc accagcagcc cccacaagtt tgggatctaa tacaatggct 2820

[0011]

[0012]

```
tcaggcggtg gcgcaccaat ggcagacaat aacgagggtg ccgatggagt gggtaattcc   2880
tcaggaaatt ggcattgcga ttccaatgg ctgggcgaca gagtcatcac caccagcacc   2940
agaacctggg ccctgcccac ttacaacaac catctctaca agcaaatctc cagccaatca   3000
ggagcttcaa acgacaacca ctactttggc tacagcaccc cttgggggta ttttgacttt   3060
aacagattcc actgccnctt ctcaccacgt gactggcagc gactcattaa caacaactgg   3120
ggattcggc ccaagaaact cagcttcaag ctcttcnaca tccaagttag aggggtcacg   3180
cagaacgatg gcacgacgac tattgccaat aaccttacca gcacggttca agtgtttacg   3240
gactcggagt atcagctccc gtacgtgctc gggtcggcgc accaaggctg tctcccgccg   3300
tttccagcgg acgtcttcat ggtccctcag tatggatacc tcaccctgaa caacggaagt   3360
caagcggtgg gacgctcatc cttttactgc ctggagtact tccttcgca gatgctaagg   3420
actggaaata actttcaatt cagctatacc ttcgaggatg taccttttca cagcagctac   3480
gctcacagcc agagtttgga tcgcttgatg aatcctctta ttgatcagta tctgtactac   3540
ctgaacagaa cgcaaggaac aacctctgga acaaccaacc aatcacggct gcttttttagc   3600
caggctgggc ctcagtctat gtctttgcag gccagaaatt ggctacctgg gccctgctac   3660
cggcaacaga gactttcaaa gactgctaac gacaacaaca acagtaactt tcattggaca   3720
gcggccagca aatatcatct caatggccgc gactcgctgg tgaatccagg accagctatg   3780
gccagtcaca aggacgatga agaaaaattt ttccctatgc acggcaatct aatatttggc   3840
aaagaaggga caacggcaag taacgcagaa ttagataatg taatgattac ggatgaagaa   3900
gagattcgta ccaccaatcc tgtggcaaca gagcagtatg gaactgtggc aaataacttg   3960
cagagctcaa atacagctcc cacgactgga actgtcaatc atcaggggcc cttacctggc   4020
atggtgtggc aagategtga cgtgtacctt caaggacctа tctgggcaaa gattcctcac   4080
acggatggac actttcatcc ttctcctctg atgggaggct ttggactgaa acatccgcct   4140
cctcaaatca tgatcaaaaa tactccggta ccggcaaatc ctccgacgac tttcagccca   4200
gccaagtttg cttcatttat cactcagtac tccactggac aggtcagcgt ggaaattgag   4260
tgggagctac agaaagaaaa cagcaaacgt tggaatccag agattcagta cacttccaac   4320
tacaacaagt ctgttaatgt ggactttact gtagacacta atggtgttta tagtgaacct   4380
cgcccctattg gaacccggta tctcacacga aacttgtgaa tcctggttaa tcaataaacc   4440
gtttaattcg tttcagttga actttggctc ttgtgcactt ctttatcttt atcttgtttc   4500
catggctact gcgtagataa gcagcggcct gcggcgcttg cgcttcgcgg tttacaactg   4560
ctggttaata tttaactctc gccatacctc tagtgatgga gttggccact ccctctatgc   4620
gcactcgctc gctcggtggg gcctggcgac caaaggtcgc cagacggacg tgctttgcac   4680
gtccggcccc accgagcgag cgagtgcgca tagagggagt ggccaa                  4726
```

<210> 9
<211> 3098
<212> DNA
<213> 新 AAV 血清型, 克隆 42.2

<400> 9

```
gaattcgccc tttctacggc tgcgtcaact ggaccaatga gaactttccc ttcaacgatt     60
gcgtcgacaa gatggtgatc tggtgggagg agggcaagat gacggccaag gtcgtggagt    120
ccgccaaggc catctcggcc ggcagcaagg tgcgcgtgga ccaaaagtgc aagtcttccg    180
cccagatcga tcccaccccc gtgatcgtca cttccaacac caacatgtgc gctgtgattg    240
acgggaacag caccaccttc gagcaccagc agccgttaca agaccggatg ttcaaatttg    300
aactcacccg ccgtctggag cacgacttig gcaaggtgac aaagcaggaa gtcaaagagt    360
tcttccgctg ggcgcaggat cacgtgaccg aggtggcgca tgagttctac gtcagaaagg    420
gtggagccaa caagagaccc gcccccgatg acgcggatаа aagcgagccc aagcgggcct    480
gcccctcagt cgcggatcca tcgacgtcag acgcggaagg agctccggtg gactttgccg    540
acaggtacca aaacaaatgt tctcgtcacg cgggcatgct tcagatgctg tttccctgca    600
agacatgcga gagaatgaat cagaatttca acatttgctt cacgcacggg accagagact    660
gttcagaatg tttccccggc gtgtcagaat ctcaaccggt cgtcagaaag aggacgtatc    720
ggaaactctg tgccattcat catctgctgg ggcgggctcc cgagattgct tgctcggcct    780
gcgatctggt caacgtggac ctggatgacc gtgtttctga gcaataaatg acttaaacca    840
ggtatggctg ccgatggtta ttccagat tggctcgagg acaacctctc tgagggcatt    900
cgcgagtggt gggacttgaa acctggagcc ccgaaaccca aagccaacca gcaaaagcag    960
gacgacggcc gggtctggt gattcctggc tacaagtacc tcggacccтt caacggactc   1020
gacaagggag agccggtcaa cgaggcagac gccgcggccc tcgagcacga caaggcctac   1080
gacaagcagc tcgagcaggg ggacaacccg tacctcaagt acaaccacgc cgacgccgag   1140
ttccaggagc gtctccaaga agatacgtct tttggggcca acctcgggcg agcagtcttc   1200
caggccaaga agcgggttct cgaacctctc ggtctggttg aggaaggcgc taagacggct   1260
cctggaaaga agagacccat agaatccccc gactcctcca cgggcatcgg caagaaaggc   1320
cagcagcccg ctaaaaagaa gctcaacttt gggcagactg gcgactcaga gtcagtgccc   1380
```

gaccccccaac ctctcggaga acctcccgcc gcgccctcag gtctgggatc tggtacaatg 1440
gctgcaggcg gtggcgcacc aatggcagac aataacgaag gcgccgacgg agtgggtaat 1500
gcctccggaa attggcattg cgattccaca tggctgggcg acagagtcat caccaccagc 1560
acccgcacct gggccctgcc cacctacaac aaccacctct acaagcagat atcaagtcag 1620
agcggggcta ccaacgacaa ccacttcttc ggctacagca ccccctgggg ctattttgac 1680
ttcaacagat tccactgcca cttctcacca cgtgactggc agcgactcat caacaacaac 1740
tggggattcc ggccccagaaa gctgcggttc aagttgttca acatccaggt caaggaggtc 1800
acgacgaacg acggcgttac gaccatcgct aataacctta ccagcacgat tcaggtcttc 1860
tcggactcgg agtaccaact gccgtacgtc ctcggctctg cgcaccaggg ctgcctccct 1920
ccgttcctg cggacgtgtt catgattcct cagtacggat atctgactct aaacaacggc 1980
agtcagtctg tgggacgttc ctccttctac tgcctggagt actttcttc tcagatgctg 2040
agaacgggca atcaacttga attcagctac accttgagg aagtgccttt ccacagcagc 2100
tatgcgcaca gccagagcct ggaccggctg atgaatcccc tcatcgacca gtacctgtac 2160
tacctggccc ggacccagag cactacgggg tccacaaggg agctgcagtt ccatcaggct 2220
gggcccaaca ccatggccga gcaatcaaag aactggctgc ccggaccctg ttatcggcag 2280
cagagactgt caaaaaacat agacagcaac aacaacagta actttgcctg gaccgggcc 2340
actaaatacc atctgaatgg tagaaattca ttaaccaacc cgggcgtagc catggccacc 2400
aacaaggacg acgaggacca gttctttccc atcaacggag tgctggtttt tggcgaaacg 2460
ggggctgcca acaagacaac gctggaaaac gtgctaatga ccagcgagga ggagatcaaa 2520
accaccaatc ccgtggctac agaagaatac ggtgtggtct ccagcaacct gcaatcgtct 2580
acggccggac cccagacaca gactgtcaac agccagggcg ctctgcccgg catggtctgg 2640
cagaaccggg acgtgtacct gcagggtccc atctgggcca aaattcctca cacggacggc 2700
aactttcacc cgtctcccct gatgggcgga tttggactca aacacccgcc tcctcaaatt 2760
ctcatcaaaa acacccccggt acctgctaat cctccagagg tgtttactcc tgccaagttt 2820
gcctcattta tcacgcagta cagcaccggc caggtcagcg tggagatcga gtgggaactg 2880
cagaaagaaa acagcaaacg ctggaatcca gagattcagt acacctcaaa ttatgccaag 2940
tctaataatg tggaatttgc tgtcaacaac gaaggggttt atactgagcc tcgccccatt 3000
ggcacccgtt acctcacccg taacctgtaa ttgcctgtta atcaataaac cggttaattc 3060
gtttcagttg aactttggtc tctgcgaagg gcgaatte 3098

[0013]

<210> 10
<211> 3098
<212> DNA
<213> 新AAV血清型,克隆16.3

<400> 10
gaattcgccc ttcgcagaga ccaaagttca actgaaacga atcaaccggt ttattgatta 60
acaagtaatt acaggttacg ggtgaggtaa cgggtgccaa tggggcgagg ctcagtataa 120
acccccttcgt tgtttgacagc aaattccaca ttattagact tggcataatt tgaggtgtac 180
tgaatctctg gattccagcg tttgctgttt tctttctgca gttcccactc gatctccacg 240
ctgacctggc cggtgctgta ctgcgtgata aatgaggcaa actaggcagg agtaaacacc 300
cctggaggat tagcaggtac cggggtgttt ttgatgagaa tttgaggagg cgggtgtttg 360
agtccaaatc cgcccatcag gggagacggg tgaaagttgc cgtccgtgtg aggaattttg 420
gcccagatgg gacccctgcag gtacacgtcc cggttctgcc agaccatgcc gggcagagcc 480
ccctggctgt tgacagtctg tgtctggggt ccggccgtag acgattgcag gttgctggag 540
accacaccgt attcttctgt agccacggga ttggtggttt tgatctccta ctcgctggtc 600
attagcacgt tttccagcgt tgtcttgttg gcagcccccg ttttgccaaa aaccagcact 660
ccgttgatgg gaaagaactg gcccttcgtcg tccttgttgg tggccatggc tacgccccgg 720
ttggttaatg aatttctacc attcagatgg tatttagtgg ccccggtcca ggcaaagtta 780
ctgttgttgt tgctgtctat gttttttgac agtctctgct gccgataaca gggtccgggc 840
agccagttct ttgattgctc ggccatggtg ttgggcccag cctgatggaa ctgcagctcc 900
cttgtggacc ccgtagtgct ctgggtccgg gccaggtagt acaggtactg gtcgatgagg 960
ggattcatca gccggtccag gctctggctg tgcgcatagc tgctgtggaa aggcacttcc 1020
tcaaaggtgt agctgaattc aaagttattg cccgttctca gcatctgaga aggaaagtac 1080
tccaggcagt agaaggagga acgtcccata gactgactgc cgttgtttag agtcagatat 1140
ccgtactgag gaatcatgaa cacgtccgca gggaacggag ggaggcagcc ctggtgcgca 1200
gagccgagga cgtacggcag ttggtactcc gagtccgaga agacctgaat cgtgctggta 1260
aggttattag cgatggtcgt aacgccgtcg ttcgtcgtga cctccttgac ctggatgttg 1320
aacaacttga accgcagctt tctgggccgg aatccccagt tgttgttgat gagtcgctgc 1380
cagtcacgtg gtgagaagtg gcagtggaat ctgttgaagt caaaatagcc ccagggggtg 1440
ctgtagccga agaagtggtt gtcgttggta gccccgctct gacttgatat ctgcttgtag 1500
aggtggttgt tgtaggtggg cagggccccag gtgcgggtgc tggtggtgat gactctgtcg 1560

cccagccatg tggaatcgca atgccaattt ccggaggcat tacccactcc gtcggcgcct 1620
tcgttattgt ctgccattgg tgcgccaccg cctgcagcca ttgtaccaga tcccagacct 1680
gagggcgcgg cgggaggttc tccgagaggt tgggggtcgg gcactgactc tgagtcgcca 1740
gtctgcccaa agttgagctt ctttttagcg ggctgctggc ctttcttgcc gatgcccgtg 1800
gaggagtcgg gggattctat gggtctcttc tttccaggag ccgtcttagc gccttcctca 1860
accagaccga gaggttcgag aacccgcttc ttggcctgga agactgctcg cccgaggttg 1920
ccccccaaaag acgtatcttc ttgaagacgc tcctgaaact cagcgtcggc gtggttgtac 1980
ttgaggtacg ggttgtcccc ctgctcgagc tgcttgtcgt aggccttgtc gtgctcgagg 2040
gccgcggcgt ctgcctcgtt gaccggctct cccttgtcga gtccgttgaa gggtccgagg 2100
tacttgtagc caggaagcac cagacccccgg ccgtcgtcct gcttttgctg gttggctttg 2160
ggtttcgggg ctccaggttt caagtcccac cactcgcgaa tgccctcaga gaggttgtcc 2220
tcgagccaat ctggaagata accatcggca gccataccig gtttaagtca tttatigctc 2280
agaaacacag tcatccaggt ccacgttgac cagatcgcag gccgagcaag caatctcggg 2340
agcccgcccc agcagatgat gaatggcaca gagtttccga tacgtcctct ttctgacgac 2400
cggttgagat tctgacacgc cggggaaaca ttctgaacag tctctggtcc cgtgcgtgaa 2460
gcaaatgttg aaattctgat tcattctctc gcatgtcttg cagggaaaca gcatctgaag 2520
catgcccgcg tgacgagaac atttgttttg gtacctgtcg gcaaagtcca ccggagctcc 2580
ttccgcgtct gacgtcgatg gatccgcgac tgaggggcag gcccgcttgg gctcgctttt 2640
atccgcgtca tcgggggcgg gcctcttgtt ggctccaccc tttctgacgt agaactcatg 2700
cgccacctcg gtcacgtgat cctgcgccca gcggaagaac tctttgactt cctgctttgt 2760
caccttgcca aagtcctgct ccagacggcg ggtgagttca aatttgaaca tccggtcttg 2820
taanggctgc tggtgctcga aggtggtgct gttcccgtca atcacggcgc acatgttggt 2880
gttggaagtg acgatcacgg gggtgggatc gatctgggcg gacgacttgc actttggtcc 2940
cacgcgcacc ttgctgccgc cgagaatggc cttggcggac tccacgacct tggccgtcat 3000
cttgccctcc tcccaccaga tcaccatctt gtcgacgcaa tcgttgaagg gaaagttctc 3060
attggtccag ttgacgcagc cgtagaaagg gcgaattc 3098

<210> 11
<211> 3121
<212> DNA
<213> 新 AAV 血清型, 克隆 29.3

[0014]

<400> 11
gaattcgccc ttcgcagaga ccaaagttca actgaaacga atcaaccggt ttattgatta 60
acaagcaatt acagattacg ggtgaggtaa cgggtgccga tggggcgagg ctcagaataa 120
gtgccatctg tgttaacagc aaagtccaca tttgtagatt tgtagtagtt ggaagtgtat 180
tgaatctctg ggttccagcg tttgctgttt tcttttctgca gctcccattc aatttccacg 240
ctgacctgtc cggtgctgta ctgcgtgatg aacgacgcca gcttagcttg actgaaggta 300
gttggaggat ccgcgggaac aggtgtattc ttaatcagga tctgaggagg cgggtgtttc 360
agtccaaagc ccccccatcag cggcgaggga tgaaagtttc cgtccgtgtg aggaatcttg 420
gcccagatag gaccctgcag gtacacgtcc cggttctgcc agaccatgcc aggtaaggct 480
ccttgactgt tgacggcccc tacaatagga gcggcgtttt gctgttgcag gttatcggcc 540
accacgccgt actgttctgt ggcccactggg ttggtggttt taatttcttc ctcactggtt 600
agcataacgc tgctatagtc cacgttgcct tttccagctc cctgtttccc aaacattaag 660
actccgctgg acggaaaaaa tcgctcttcg tcgtccttgt gggttgccat agcgacaccg 720
ggatttacca gagagtctct gccattcaga tgatacttgg tggcacccgt ccaggcaaag 780
ttgctgttgt tatttlgcga cagtgtcgtg gagacgcgtt gctgccggta gcagggcccg 840
ggtagccagt ttttggcctg agccgacatg ttattaggcc cggcctgaga aaatagcaac 900
tgctgagttc ctgcggtacc tcccgtggac tgagtccgag acaggtagta caggtactgg 960
tcgatgaggg ggttcatcag ccggtccagg ctttggctgt gcgcgtagct gctgtgaaaa 1020
ggcacgtcct caaactggta gctgaactca aagttgttgc ccgttctcag catttgagaa 1080
ggaaagtact ccaggcagta gaaggaggaa cggcccacgg cctgactgcc attgttcaga 1140
gtcaggtacc cgtactgagg aatcatgaag acgtccgccg ggaacggagg caggtagccc 1200
tggcgcgcag agccgaggac gtacgggagc tggtattccg agtccgtaaa gacctgaatc 1260
gtgctggtaa ggttattggc gatggtcttg gtgccttcat tctgcgtgac ctccttgacc 1320
tggatgttga agagcttgaa gttgagtctc ttgggccgga atcccccagtt gttgttgatg 1380
agtcgctgcc agtcacgtgg tgagaagtgg cagtggaatc tgtaaaagtc aaaatacccc 1440
cagggggtgc tgtagccgaa gtaggtgttg tcgttggtgc ttcctcccga agtcccgttg 1500
gagatttgct tgtagaggtg gttgttgtag gtggggaggg cccaggttcg ggtgctggtg 1560
gtgatgactc tgtcgcccag ccatgtggaa tcgcaatgcc aatttcctga ggaactaccc 1620
actccgtcgg cgccttcgtt attgtctgcc attggagcgc caccgcctgc agccattgta 1680
ccagatccca gaccagaggg gcctgcgggg ggttctccga ttggttgagg gtcgggcact 1740

gactctgagt cgccagtctg cccaaagttg agtctctttt tcgcgggctg ctggcctttc 1800
ttgccgatgc ccgtagtgga gtctggagaa cgctggggtg atggctctac cggtctcttc 1860
tttccaggag ccgtcttagc gccttcctca accagaccga gaggttcgag aacccgcttc 1920
ttggcctgga agactgctcg tccgaggttg ccccccaaaag acgtatcttc ttgcagacgc 1980
tcctgaaact cggcgtcggc gtggttatac cgcaggtacg gattgtcacc cgctttgagc 2040
tgctggtcgt aggccttgtc gtgctcgagg gccgctgcgt ccgccgcgtt gacgggctcc 2100
cccttgtcga gtccgttgaa gggtccgagg tacttgtagc caggaagcac cagacccccgg 2160
ccgtcgtcct gcttttgctg gttggctttg ggcttcgggg ctccaggttt cagcgcccac 2220
cactcgcgaa tgccctcaga gaggttgtcc tcgagccaat ctggaagata accatcggca 2280
gccatacctg atctaaatca tttattgttc aaagatgcag tcatccaaat ccacattgac 2340
cagatcgcag gcagtgcaag cgtctggcac ctttccccatg atatgatgaa tgtagcacag 2400
tttctgatac gccttttiga cgacagaaac gggttgagat tctgaccacgg gaaagcactc 2460
taaacagtct ttctgtccgt gagtgaagca gatatttgaa ttctgattca ttctctcgca 2520
ttgtctgcag ggaaacagca tcagatttat gcccacgtga cgagaaacatt tgttttggta 2580
ccigtccgcg tagttgatcg aagcttccgc gtctgacgtc gatggctgcg caactgactc 2640
gcgcacccgt ttgggctcac ttatatctgc gtcactgggg gcgggtcttt tcttggctcc 2700
accctttttg acgtagaatt catgctccac ctcaaccacg tgatcctttg cccaccggaa 2760
aaagtctttg acttcctgct tggtgacctt cccaaagtca tgatccagac ggcgggtgag 2820
ttcaaatttg aacatccggt cttgcaacgg ctgctggtgt tcgaaggtcg ttgagttccc 2880
gtcaatcacg gcgcacatgt tggtgttgga ggtgacgatc acgggagtcg ggtctatctg 2940
ggccgaggac ttgcatttct ggtccacgcg caccttgctt cctccgagaa tggctttggc 3000
cgactccacg accttggcgg tcatcttccc ctcctcccac cagatcacca tcttgtcgac 3060
acagtcgttg aagggaaagt tctcattggt ccagttgacg cagccgtaga agggcgaatt 3120
c 3121

<210> 12
<211> 3121
<212> DNA
<213> 新AAV血清型,克隆29.4

[0015]

<400> 12
gaattcgccc ttctacggct gcgtcaactg gaccaatgag aactttccct tcaacgactg 60
tatcgacaag atggtgatct ggtgggagga ggggaagatg accgccaagg tcgtggagtc 120
ggccaaagcc attctcggag gaagcaaggt gcgcgtggac cagaaatgca agtcctcggc 180
ccagatagac ccgactcccg tgatcgtcac ctccaacacc aacatgtgcg ccgtgattga 240
cgggaactca acgaccttcg aacaccagca gccgttgcaa gacccggatgt tcaaatttga 300
actcacccgc cgtctggatc atgactttgg gaaggtcacc aagcaggaag tcaaagactt 360
tttccggtgg gcaaaggatc acgtggttga ggtggagcac gaattctacg tcaaaaaggg 420
tggagccaag aaaagacccg ccccccagtga cgcagatata agtgagccca aacgggtgcg 480
cgagtcagtt gcgcagccat cgacgtcaga cgcggaagct tcgatcaact acgcagacag 540
gtaccaaaac aaatgttctc gtcacgcggg catgaatctg atgctgtttc cctgcagaca 600
atgcgagagg atgaatcaga attcaaatat ctgcttcact cacggacaga aagactgttt 660
agagtgcttt cccgtgtcag aatctcaacc cgtttctgtc gtcaaaaagg cgtatcagaa 720
actgtgctac attcatcata tcatgggaaa ggtgccagac gcttgcactg cctgcgatct 780
ggtcgatgtg gatttggatg actgcatctt tgaacaataa atgatttaaa tcaggtatgg 840
ctgccgatgg ttatcttcca gatiggctcg aggacaacct ctctgagggc attcgcgagt 900
ggtgggcgct gaaacctgga gccccgaagc ccaaagccaa ccagcaaaag caggacggcg 960
gccggggtct ggtgcttcct ggctacaagt acctcggacc cttcaacgga ctcgacaagg 1020
gggagccccgt caacgcggcg gacgcagcgg ccctcgagca cgacaaggcc tacgaccagc 1080
agctcaaagc gggtgacaat ccgtacctgc ggtataacca cgccgacgcc gagtttcagg 1140
agcgtctgca agaagatacg tcttttgggg gcaacctcgg gcgagcagtc ttccaggcca 1200
agaagcgggt tctcgaacct ctcggtctgg ttgaggaagg cgctaagacg gctcctggaa 1260
agaagagacc ggtagagcca tcaccccagc gttctccaga ctcctctacg ggcatcggca 1320
agaaaggcca gcagcccgcg aaaaagagac tcaactttgg gcagactggc gactcagagt 1380
cagtgcccga ccctcaacca atcggagaac ccccgcagg ccccctctggt ctgggatctg 1440
gtacaatggc tgcaggcggt ggcgctccaa tggcagacaa taacgaaggc gccgacggag 1500
tgggtagttc ctcaggaaat tggcattgcg attccacatg gctgggcgac tgagtcatca 1560
ccaccagcac ccgaacctgg gccctcccca cctacaacaa ccacctctac aagcaaatct 1620
ccaacgggac ttcgggagga agccaccaacg acaacaccta cttcggctac agcaccccct 1680
ggggtattt tgactttaac agattccact gccacttctc accacgtgac tggcagcgac 1740
tcatcaacaa caactgggga ttccggccca agagactcaa cttcaagctc ttcaacatcc 1800
aggtcaagga ggtcacgcag aatgaaggca ccaagaccat cgccaataac cttaccagca 1860

cgatcaggt ctttacggac tcggaatacc agctcccgta cgtcctcggc tctgcgcacc 1920
agggctgcct gcctccgttc ccggcggacg tcttcatgat tcctcagtac gggtacctga 1980
ctctgaacaa tggcagtcag gccgtgggcc gttcctcctt ctactgcctg gagtactttc 2040
cttctcaaat gctgagaacg ggcaacaact ttgagttcag ctaccagttt gaggacgtgc 2100
ctttcacag cagctacgcg cacagccaaa gcctggaccg gctgatgaac ccctcatcg 2160
accagtacct gtactacctg tctcggactc agtccacggg aggtaccgca ggaactcagc 2220
agttgctatt ttctcaggcc gggcctaata acatgtcggc tcaggccaaa aactggctac 2280
ccgggccctg ctaccggcag taacgcgtct ccacgacact gtcgcaaaat aacaacagca 2340
actttgtctg gaccggtgcc accaagtatc atctgaatgg cagagactct ctggtagatc 2400
ccggtgtcgc tatggcaacc cacaaggacg acgaagagcg atttttccg tccagcggag 2460
tcataatgtt tgggaaacag ggagctggaa aagacaacgt ggactatagc agcgtcatgc 2520
taaccagtga ggaagaaatt aaaaccacca acccagtggc cacagaacag tacggcgtgg 2580
tggccgataa cctgccaacag caaaacgccg ctcctattgt aggggccgtc aacagtcaag 2640
gagccttacc tggcatggtc tggcagaacc gggacgtgta cctgcagggt cctaccctgg 2700
ccaagattcc tcacacggac ggaaacttte atccctcgcc gctgatggga ggctttggac 2760
tgaaacaccc gcctcctcag atcctgatta agaatacacc tgttccgcg gatcctccaa 2820
ctaccttcag tcaagctaag ctggcgtcgt tcatcacgca gtacagcacc ggacaggtca 2880
gcgtggaaat tgaatgggag ctgcaggaag aaaacagcaa acgctggaac ccagagattc 2940
aatacacttc caactactac aaatctacaa atgtggactt tgctgttaac acagatggca 3000
cttattctga gcctcgcccc atcggcaccc gttacctcac ccgtaatctg taattgcttg 3060
ttaatcaata aaccggttga ttcgtttcag ttgaactttg gtctctgcga agggcgaatt 3120
c 3121

<210> 13
<211> 3121
<212> DNA
<213> 新 AAV 血清型, 克隆 29.5

[0016]

<400> 13
gaattcgccc ttcgcgagac caaagttcaa ctgaaacgaa tcaaccggtt tattgattaa 60
caagcaatta cagattacgg gtgaggtaac gggtgccgat ggggcgaggc tcagaataag 120
tgccatctgt gttaacagca aagtccacat ttgtagattt gtagtagttg gaagtgtatt 180
gaatctctgg gttccagcgt ttgctgtttt ctttctgcag ctcccattca atttccacgc 240
tgacctgtcc ggtgctgtac tgcgtgatga acgacgccag cttagcttga ctgaaggtag 300
ttggaggatc cgcgggaaca ggtgtattct taatcaggat ctgaggaggc gggtgtttca 360
gtccaaagcc tcccatcagc ggcgagggat gaaagtttcc gtccgtgtga ggaatcttgg 420
cccagatagg accctgcagg tacacgtccc ggttctgcca gaccatgcca ggtaaggctc 480
cttgactgtt gacggccccc acaataggag cggcgttttg ctgttgcagg ttatcggcca 540
ccacgccgta ctgttctgtg gccactgggt tggtggtttt aatttcttcc tcactggtta 600
gcataacgct gctatagtcc acgttgtctt ttccagctcc ctgtttccca aacattaaga 660
ctccgctgga cggaaaaaat cgctcttcgt cgtccttgtg ggttgccata gcgacaccgg 720
gatttaccag agagtctctg ccattcagat gatacttggt ggcaccggtc caggcaaagt 780
tgctgttgtc attttgcgac agtgtcgtgg agacgcgttg ctgccggtag cagggcccgg 840
gtagccagtt tttggcctga gccgacatgt tattaggccc ggcctgagaa aatagcaact 900
gctgagttcc tgcggtacct cccgtggact gagtccgaga caggtagtac aggtactggt 960
cgatgagggg gttcatcagc cggtccaggc tttggctgtg cgcgtagctg ctgtgaaaag 1020
gcacgtcctc aaactggtag ctgaactcaa agttgttgcc cgttctcagc attgagaag 1080
gaaagtactc caggcagtag aaggaggaac ggcccacggc ctgactgcca ttgttcagag 1140
tcaggtaccc gtactgagga atcatgaaga cgtccgccgg gaacggaggc aggcagccct 1200
ggtgcgcaga gccgaggacg tacgggagct ggtattccga gtccgtaaag acctgaatcg 1260
tgctggtaag gttattggcg atggtcttgg tgccttcatt ctgcgtgacc tccttgacct 1320
ggatgttgaa gagcttgaag ttgaggctct tgggccggaa tcccccagttg ttgttgatga 1380
gtcgctgcca gtcacgtggt gagaagtggc agtggaatct gttaaagtca aaatacccgc 1440
agggggtgct gtagccgaag taggtgttgt cgttggtgct tcctccgaa gtcccgttgg 1500
agatttgctt gtagaggtgg ttgttgtagg tggggagggc ccaggttcgg gtgctggtgg 1560
tgatgactcc gtcgcccagc catgtggaat cgcaatgcca atttcctgag gaactaccca 1620
ctccgtcggc gcttcgtta ttgtctgcca ttggagcgcc accgcctgca gccattgtac 1680
cagatcccag accagagggg cctgcggggg gttctccgat tggttgaggg tcgggcactg 1740
actctgagtc gccagtctgc ccaaagttga gtctcttttt cgcgggctgc tggcctttct 1800
tgccgatgcc cgtagaggag tctggagaac gctggggtga tggctctacc ggtctcttct 1860
ttccaggagc cgttcttagcg ccttcctcaa ccagaccgag aggttcgaga acccgcttct 1920
tggcctggaa gactgctcgc ccgaggttgc ccccaaaaga cgtatcttct tgcagacgct 1980

cctgaaactc ggcgtcggcg tggttatacc gcaggtacgg attgtcaccc gctttgagct 2040
gctggtcgta ggccttgtcg tgctcgaggg ccgctgcgtc cgccgcgttg acgggctccc 2100
ccttgtcgag tccgttgaag ggtccgaggt acttgtagcc aggaagcacc agacccccggc 2160
cgtcgtcctg cttttgctgg ttggctttgg gcttcggggc tccaggtttc agcgcccacc 2220
actgcgcaat gccctcngag aggttgtcct cgagccaatc tggaagataa ccatcggcag 2280
ccatacctga tttaaatcat ttattgttca aagatgcagt catccaaatc cacattgacc 2340
agatcgcagg cagtgcaagc gtctggcacc tttcccatga tatgatgaat gtagcacagt 2400
ttctgatacg ccttttttgac gacagaaacg ggttgagatt ctgacacggg aaagcactct 2460
aaacagtctt tctgtccgtg agtgaagcag atatttgaat tctgattcat tctctcgcat 2520
tgtctgcagg gaaacagcat cagattcatg cccacgtgac gagaacattt gttttggtac 2580
ctgtctgcgt agttgatcga agcttccgcg tctgacgtcg atggctgcgc aactgactcg 2640
cgcacccgtt tgggctcacl tatatctgcg tcactggggg cgggtctttt cttggctcca 2700
ccctttttga cgtagaattc atgctccacc tcaaccacgt gatcctttgc ccaccggaaa 2760
aagtctttga cttcctgctt ggtgaccttc ccaaagtcat gatccagacg gcgggtgagt 2820
tcaaatttga acatccggtc ttgcaacggc tgctggtgtt cgaaggtcgt tgagttcccg 2880
tcaatcacgg cgcacatgtt ggtgttggag gtgacgatca cgggagtcgg gtctatctgg 2940
gccgaggact tgcatttctg gtccacgcgc accttgcttc ctccgagaat ggctttggcc 3000
gactccacga ccttggcggt catcttcccc tcctcccacc agatcaccat cttgtcgaca 3060
cagtcgttga agggaaagtt ctcattggtc cagttgacgc agccgtagaa agggcgaatt 3120
c 3121

<210> 14
<211> 3131
<212> DNA
<213> 新 AAV 血清型, 克隆 1-3

<400> 14

[0017]

gcggccgcga attcgccctt ggctgcgtca actggaccaa tgagaacttt ccctteaatg 60
attgcgtcga caagatggtg atctggtggg aggagggcaa gatgacggcc aaggtcgtgg 120
agtccgccaa ggccattctc ggcggcagca aggtgcgcgt ggaccaaaag tgcaagtcgt 180
ccgcccagat cgaccccacc ccccgtgatcg tcacctccaa caccaacatg tgcgccgtga 240
ttgacgggaa cagcaccacc ttcgagcacc agcagcctct ccaggaccgg atgtttaagt 300
tcgaactcac ccgccgtctg gagcacgact ttggcaaggt gacaaagcag gaagtcaaag 360
agttctttcg ctgggccagt gatcacgtga ccgaggtggc gcatgagttt tacgtcagaa 420
agggcggagc cagcaaaaga ccccgccccg atgacgcgga taaaagcgag ccccaagcggg 480
cctgccccte agtcgcggat ccatcgacgt cagacgcgga aggagctccg gtggacttig 540
ccgacaggta ccaaaaacaaa tgttctcgtc acgcgggcat gcttcagatg ctgtttccct 600
gcaaaacgtg cgagagaatg aategggaatt tcaacatttg cttcacacac ggggtcagag 660
actgctcaga gtgtttcccc ggcgtgtcag aatctcaacc ggtcgtcaga aagaggacgt 720
atcggaaact ccgtgcgatt catcatctgc tggggcgggc tccgagatt gcttgctcgg 780
cctgcgatct ggtcaacgtg gacctggatg actgtgtttc tgagcaataa atgacttaaa 840
ccaggtatgg ctgccgatgg ttatcttcca gattggctcg aggacaacct ctctgagggc 900
attcgcgagt ggtgggcgct gaaacctgga gccccgaagc ccaaagccaa ccagcaaaag 960
caggacgacg gccggggtct ggtgcttcct ggctacaagt acctcggacc cttcaacgga 1020
ctcgacaagg gggagcccgt caacgcggcg gacgcagcgg ccctcgagca cgacaaggct 1080
tacgaccagc agctgcaggc gggtgacaat ccgtacctgc ggtataacca cgccgacgcc 1140
gagtttcagg agcgtctgca agaagatacg tctttgggag gcaacctcgg gcgagcagtc 1200
ttccaggcca agaagcgggt tctcgaacct ctcggtctgg ttgaggaagg cgctaagacg 1260
gctcctggaa agaagagacc ggtagagcca tcacccccagc gttctccaga ctcctctacg 1320
ggcatcggca agaaaggcca acagcccgcc agaaaaagac tcaattttgg tcagactggc 1380
gactcagagt cagttccaga cctcaacct ctcggagaac ctccagcagc gcctctggt 1440
gtgggaccta atacaatggc tgcaggcggt ggcgcaccaa tggcagacaa taacgaaggc 1500
gccgacggag tgggtagttc ctcgggaaat tggcattgcg attccacatg gctgggcgac 1560
agagtcatca ccaccagcac ccgaacctgg gccctgccca cctacaacaa ccacctctac 1620
aagcaaatct ccaacgggac atcgggagga gccaccaacg acaacaccta cttcggctac 1680
agcaccccct ggggqtattt tgactttaac agattccact gccacttic accacgtgac 1740
tggcagcgac tcatcaacaa caactgggga ttccgaccca agagactcag cttcaagctc 1800
ttcaacatcc aggtcaagga ggtcacgcag aatgaaggca ccaagaccat cgccaataac 1860
ctcaccagca ccatccaggt gtttacggac tcggagtacc agctgccgta cgttctcggc 1920
tctgtccacc agggctgcct gcctccgttc ccggcggacg tgttcatgat tccccagtac 1980
ggctacctaa cactcaacaa cggtagtcag gccgtgggac gctcttcctt ctactgcctg 2040
gaatacttic cttcgcagat gctgagaacc ggcaacaact tccagtttac ttcaccttc 2100

gaggacgtgc ctttccacag cagctacgcc cacagctaga gcttggaccg gctgatgaat 2160
cctctgattg accagtacct gtactacttg tctcggactc aaacaacagg aggcacggca 2220
aatacgcaga ctctgggctt cagccaaggt gggcctaata caatggccaa tcaggcaaag 2280
aactggctgc caggaccctg ttaccgccaa caacgcgtct caacgacaac cgggcaaaac 2340
aacaatagca actttgcctg gactgctggg accaaatacc atctgaatgg aagaaattca 2400
ttggctaatc ctggcatcgc tatgggaaca cacaaagacg acgaggagcg ttttttttcc 2460
agtaacggga tcctgatttt tggcaaacaa aatgctgcca gagacaatgc ggattacagc 2520
gatgtcatgc tcaccagcga ggaagaaatc aaaaccacta accctgtggc tacagaggaa 2580
tacggtatcg tggcagataa cttgcagcag caaaacacgg ctcctcaaat tggaactgtc 2640
aacagccagg gggccttacc cggtatggtc tggcagaacc gggacgtgta cctgcagggt 2700
cccatctggg ccaagattcc tcacacggac ggcaacttcc accccgtctc cgtgatgggc 2760
ggctttggcc tgaaacatcc tccgcctcag atcctgatca agaacacgcc tgtacctgcg 2820
gatcctccga ccaccttcaa ccagtcaaag ctgaactctt tcatcacgca atacagcacc 2880
ggacaggtca gcgtggaaat tgaatgggag ctgcagaagg aaaacagcaa gcgctggaac 2940
cccgagatcc agtacacctc caactactac aaatctataa gtgtggactt tgctgttaat 3000
acagaaggcg tgtactctga accccgcccc attggcaccc gttacctcac ccgtaatctg 3060
taattgcctg ttaatcaata aaccggttga ttcgtttcag ttgaactttg gtctctgcga 3120
agggcgaatt c 3131

<210> 15
<211> 3127
<212> DNA
<213> 新 AAV 血清型,克隆 13-3b

[0018]

<400> 15
gcggccgcga attcgccctt cgcagagacc aaagttcaac tgaaacgaat caaccggttt 60
attgattaac atgcaattac agattacggg tgaggtaacg agtgccaata gggcgaggct 120
cagagtaaac accctggctg tcaacggcaa agtccacacc agtctgtttt tcaaagttgg 180
aggtgtactg aatctccggg tcccagcgct tgctgtttc ctttgcagc tcccactcga 240
tttccacgct gacttgtccg gtgctgtact gtgtgatgaa cgaagcaaac ttggcaggag 300
taaacacctc cggaggatta gcgggaacgg gagtgttctt gatcaggatc tgaggaggcg 360
gatgtttaag tccaaagccg ccatcaaag gagacgggtg aaagttgcca tccgtgtgag 420
gaatcttggc ccagatggga ccctgcaggt acacgtcccg gttctgccag accatgccag 480
gtaaggctcc ctggttgttg acaacttgtg tctgggctgc agtattagcc gcttgtaagt 540
tgctgctgac tatcccgtat tcttccgtgg ctacaggatt agtaggacga atttcttctt 600
catttgtcat taacaccattt tccaatgtag ttttgttagt tgctccagtt tttccaaaaa 660
tcaggactcc gctggatggg aaaaagcggt cctcgtcgtc cttgtgagtt gccatggcga 720
cgccgggatt aaccaacgag tttctgccgt tcaggtgata tttggtggca ccagtccaag 780
caaagttgct gttgttgttt tgatccagcg ttttggagac cctttgttgc cggcagcagg 840
gtccaggtaa ccaattcttg gcttgttcgg ccatagttga aggccccgcc tggtaaaact 900
gcagttcccg attgccagct gtgcctcctg ggtcactctg tgttctggcc aggtagtaca 960
agtactggtc gatgagggga ttcatcagcc ggtccaggct ctggctgtgt gcgtagctgc 1020
tgtggaaagg cacgtcctcg aagctgtagc tgaactcaaa gttgttgccc gttctcagca 1080
tctgagaggg gaagtactcc aggcagtaga aggaggaacg tcccacagac tgactgccat 1140
tgttgaggct caggtagccg tactgaggaa tcatgaagac gtccgccggg aacggaggca 1200
ggcagccctg gtgcgcagag ccgaggacgt acggcagctg gtattccgag tccgagaata 1260
cctgaatcgt gctggtaagg ttattagcga tggtcgtaac gccgtcattc gtcgtgacct 1320
ccttgacctg gatgttgaag agcttgaacc gcagcttctt gggccggaat ccccagttgt 1380
tgttgatgag tcgctgccag tcacgtggtg agaagtggca gtggaatctg ttaaagtcaa 1440
aataccccca gggggtgctg tagccgaagt aggtgttgtc gttggtacta cctgcagttt 1500
cactggagat ttgctcgtag aggtggttgt tgtaggtggg cagggcccag gttcgggtgc 1560
tggtggtaat gactctgtcg cccagccatg tggaatcgca atgccaattt cctgaggcat 1620
taccctctcc gtcggcacct tcgttattgt ctgccattgg tgcgccaccg cctgcagcca 1680
ctgtaccaga tcccacacta gagggcgctg ctggaggttc tccgagaggt tgagggtcgg 1740
ggactgactc tgagtcgcca gtctgaccga aattgagtct ctttctggcg ggctgctggc 1800
ccttcttgcc gatgcccgtg gaggagtcgg gggaacgctg aggtgacggc tctaccggtc 1860
tcttctttgc aggagccgtc ttagcgcctt cctcaaccag accgagaggt tcgagaaccc 1920
gcttcttggc ctggaagact gctcgcccga ggttgccccc aaatgacgta tcttcttgca 1980
gacgctcctg aaactcggcg tcggcgtggt tataccgcag gtacgggttg tcaccccgcat 2040
tgagctgctg gtcgtaggcc ttgtcgtgct cgagggcccgc tgcgtccgcc gcgttgacgg 2100
gctcccccti gtcgagtccg ttgaagggtc cgaggtactt gtagccagga agcaccagac 2160
cccggccgtt gtcctgcttt tgctggttgg ctttgggttt cggggctcca ggtttcaggt 2220

cccaccactc gcgaatgccc tcagagaggt tgtcctcgag ccaatctgga agataaccat 2280
cggcagccat acctgattta aatcatttat tgttcaaaga tgcagtcatc caaatccaca 2340
ttgaccagat cgcaggcagt gcaagcgtct ggcacctttc ccatgatatg atgaatgtag 2400
cacagtttct gatacgcctt tttgacgaca gaaacgggtt tagattctga cacgggaaag 2460
cactctaaac agtctttctg tccgtgagtg aagcagatat ttgaattctg attcattctc 2520
tgcaattgtc tgcagggaaa cagcatcaga ttcatgccca cgtgacgaga acatttgttt 2580
tggtacctgt ctgcgtagtt gatcgaagct tccgcgtctg acgtcgatgg ctgcgcaact 2640
gactcgcgca cccgtttggg ctcacttata tctgcgtcac tgggggcggg tcttttcttg 2700
gctccaccct ttttgacgta gaattcatgc tccacctcaa ccacgtaatc ctttgcccac 2760
cggaaaaagt ctttgacttc ctgcttggtg accttcccaa agtcatgatc cagacggcgg 2820
gtgagttcaa atttgaacat ccggtcttgc aacggctgct ggtgttcgaa ggtcgttgag 2880
ttcccgtcga tcacggcgca catgttggtg ttggagatga cgatcgcggg agtcgggtct 2940
atctgggccg aggacttgca tttctggtcc acgcgcacct tgcttcctcc gagaatggct 3000
ttggccgact ccacgacctt ggcggtcatc ttcccctcct cccaccagat caccatcttg 3060
tcgacacagt cgttgaaggg aaagttctca ttggtccagt tgacgcagcc gtagaaaggg 3120
cgaattc 3127

<210> 16
<211> 3106
<212> DNA
<213> 新 AAV 血清型, 克隆 24-1

[0019]

<400> 16
gcggccgcga attcgccctt cgcagagacc aaagttcaac tgaaacgaat caaccggttt 60
attgattaac aagtaattac aggttacggg tgaggtaacg ggtgccaatg gggcgaggct 120
cagtataaac ccttcgttg ttgacagcaa attccacatt attagacttg gcataatttg 180
aggtgtactg aatctctgga ttccagcgtt tgctgtttc tttctgcagt tcccactcga 240
tctccacgct gacctggccg gtgctgtact gcgtgataaa tgaggcaaac ttggcaggag 300
taaacaccic tggaggatta gcagglaccg gggtgtttt gatgagaatt tgaggaggcg 360
ggtgtttgag tccaaaatcg ccgatcaggg gagacgggtg aaagttgccg tccgtgtgag 420
gaatttggc ccagatggga ccctgcaggc acacgtcccg gttctgccag accatgccgg 480
gcagagcccc ctggctgttg acagtctgtg tctggggtcc ggccgtagac gattgcaggt 540
tgctggagac cacaccgtat tcttctgtag ccacgggatt ggtggttttg atctcctcct 600
cgctggtcat tagcacgttt tccagcgttg tcttgttggc agcccccgtt ttgccaaaaa 660
ccagcactcc gttgatggga aagaactggt cctcgtcgtc cttgttggtg gccatggcta 720
cgcccgggtt ggttaatgaa tttctaccat tcagatggta tttagtggcc ccggtccagg 780
caaagttact gttgttgttg ctgtctatgt ttttttgacag tctctgctgc cgataacagg 840
gtccgggcag ccagttcttt gattgctcgg ccatggtgtt gggcccagcc tgatggaact 900
gcagctcccc tgtggacccc gtagtgctct gggtccgggc caggtagtac aggtactggt 960
cgatgagggg attcatcagc cggtctaggc tctggctgtg cacatagctg ctgtggaaag 1020
gcacttcctc aaaggtgtag ctgaattcaa agttattgcc cgttctcagc atctgagaag 1080
gaaagtactc caggcagtag aaggaggaac gtcccacaga ctgactgccg ttgtttagag 1140
tcagatatcc gtactgagga atcatgaaca cgtccgcagg gaacggaggg aggcagcctc 1200
ggtgcgcaga gccgaggacg tacggcagtt ggtactccga gtccgagaag acctgaatcg 1260
tgctggtaag gttattagcg atggtcgtaa cgccgtcgtt cgtcgtgacc tccttgacct 1320
ggatgttgaa caacttgaac cgcagcttic tgggccggaa tcccccagtig ttgttgatga 1380
gtcgctgcca gtcacgtggt gagaagtggc agtggaatct gttgaagtca aaatagcccc 1440
agggggtgct gtagctgaag aagtggttgt cgttggtagc cccgctctga cttgatatct 1500
gcttgtagag gtggttgttg taggtgggca gggcccaggt gcgggtgctg gtggtgatga 1560
ctctgtcgcc cagccatgtg gaatcgcaat gccaatttcc ggaggcatta cccactccgt 1620
cggcgccttc gttattgtct gccattggtg cgccaccgcc tgcagccatt gtaccagatc 1680
ccagacctga gggcgcggcg ggaggttctc cgagaggttg ggggtcgggc actgactctg 1740
agtcgccagt ctgcccaaag ttgagcttct ttttagcggg ctgctggcct ttcttgccga 1800
tgcccgtgga ggagtcgggg gattctatgg gtctcttctt tccaggagcc gtcttagcga 1860
cttcctcaac cagaccgaga ggttcgagaa cccgcttctt ggcctggaag actgctcgcc 1920
cgaggttgtc cccaaaagac gtatcttctt gaagacgctc ctgaaactcg gcgtcggcgt 1980
ggttgtactt gaggtacggg ttgtccccct gctcgagctg cttgtcgtag gccttgtcgt 2040
gctcgagggc cgcggcgtct gcctcgttga ccggctctcc cttgtcgagt ccgttgaagg 2100
gtctgaggta cttgtagcca ggaagcacca gaccccggcc gtcgtcctgc ttttgctggt 2160
taggtttggg tttcggggct ccaggtttca agtcccacca ctcgcgaatg ccctcagaga 2220
ggttgtcctc gagccaatct ggaagataac catcggcagc catacctggt ttaagtcatt 2280
tattgctcag aaacacagtc atccaggtcc acgttgacca gatcgcaggc cgagcaagca 2340

atctcgggag cccgccccag cagatgatga atggcacaga gtttccgata cgtcctcttt 2400
ctgacgaccg gttgagattc tgacacgccg gggaaacatt ctgaacagtc tctggtcccg 2460
tgcgtgaagc aaatgttgaa attctgattc actctctcgc atgtcttgca gggaaacagc 2520
atctgaagra tgcccgcgtg acgagaacat ttgttttggt accigtcggc aaagtccacc 2580
ggagctcctt ccgcgtctga cgtcgatgga ttcgcgactg aggggcaggc ccgcttgggc 2640
tcgcttttat ccgcgtcatc gggggcgggt ctcttgttgg ccccaccctt tctgacgtag 2700
aacccatgcg ccacctcggt cacgtgatcc tgcgcccagc ggaagaacct tttgacttcc 2760
tgctttgtca ccttgccaaa gttatgctcc agacggcggg tgggttcaaa tttgaacatc 2820
cggtcctgca acggctgctg gtgctcgaag gtggcgctgt tcccgtcaat cacggcgcac 2880
atgttggtgt tggaggtgac ggtcacgggg gtggggtcga tctgggcgga cgacttgcac 2940
ttttggtcca cgcgcacctt gctgccgccg agaatggcct tggcggactc cacgaccttg 3000
gccgtcatct tgccctcctc ccaccagatc accatcttgt cggcgccaatc gttgaaggga 3060
aagttctcat tggtccagtt gacgcagccg tagaaagggc gaattc 3106

<210> 17
<211> 3102
<212> DNA
<213> 新AAV血清型,克隆27 3

<400> 17
gcggccgcga attcgccctt cgcagagacc aaagttcaac tgaaacgaat caaccggttt 60
attgattaac aagtaattac aggttacggg tgaggtaacg ggtgccaatg gggcgaggct 120
cagtataaac cccttcgttg ttgacagcaa attccacatt attagacttg gcataatttg 180
aggtgtactg aatctctgga ttccagcgtt tgctgttttc tttctgcagt tcccactcga 240
tctccacgct gacctggccg gtgctgtact gcgtgataaa tgaggcaaac ttggcaggag 300
taaacacctc tggaggatta gcaggtaccg gggtgttttt gatgagaatt tgaggaggcg 360
ggtgtttgag tccaaatccg cccatcaggg gagacgggtg aaagttgccg tccgtgtgag 420
gaatttcggc ccagatggga cctgcaggt acacgtcccg gttctgccag accatgccgg 480
gcagagcccc ctggctgttg acagtctgtg tccgggctcc ggccgtagac gattgcaggt 540
tgctggagac cacaccgtat tcttctgtag ccacgggatt ggtggtttig atctcctcct 600
cgctggtcat tagcacgttt tccagcgttg tcttgttggc agccccccgtt ttgccaaaaa 660
ccagcactcc gttgatggga aggaactggt cctcgtcgtc cttgttggtg gccatggcta 720
cgcccgggtt ggttaatgaa tttctaccat tcagatggta tttagtggcc ccggtccagg 780
caaagttact gttgttgttg ctgtctatgt tttttgacag tctctgctgc cgataacagg 840
gtccgggcag ccagttcttt gattgctcgg ccacggtgtt gggcccagcc tgatggaact 900
gcagctcccc tgtggacccc gtagtgctct gggtccgggc caggtagtac aggtactggt 960
cgatgagggg attcatcagc cggtccaggc tctggctgtg cgcatagctg ctgtggaaag 1020
gcacttcctc aaaggtgtag ctgaattcaa agttattgcc cgttctcagc atctgagaag 1080
gaaagtactc caggcagcag aaggaggaac gtcccacaga ctgactgccg ttgtttagag 1140
tcagatatcc gtactgagga atcatgaaca cgtccgcagg gaacggaggg aggcagcctc 1200
ggtgcgcaga gccgaggacg tacggcagtt ggtactccga gtccgagaag acctgaatcg 1260
tgctggtaag gttattagcg atggtcgtaa cgccgtcgtt cgtcgtgacc tccttgacct 1320
ggatgttgaa caacttgaac cgcagctttc tgggccggaa tcccccagttg ttgttgatga 1380
gtcgctgcca gtcacgtggt gagaagtggc agtggaatct gttgaagtca aaatagcccc 1440
aggggtgct gtagccgaag aagtggttgt cgttggtagc cccgctctga cttgatatct 1500
gcttgtagag gtggttgttg taggtgggca gggcccaggt gcgggtgctg gtggtgatga 1560
ctctgtcgcc cagccatgtg gaatcgccaat gccaatttcc ggaggcattа cccactccgt 1620
cggcgccttc gttattgtct gccattggtg cgccacccgcc tgcagccatt gtaccagatc 1680
ccagacctga gggcgcggcg ggaggttctc cgagaggttg ggggtcgggc actgactctg 1740
agtcgccagt ctgcccaaag ttgagcttct ttttagcggg ctgctggcct ttcttgccga 1800
tgcccgtgga ggagtcgggg gattctatgg gtctcttctt tccggaagcc gtcttagcgc 1860
cttcctcaac cagaccgaga ggttcgagaa cccgcttctt ggcctggaag actgctcgcc 1920
cgaggttgcc cccaaaagac gtatcttctt gaagacgctc ctgaaactcg gcgtcggcgt 1980
ggttgtactt gaggtacggg ttgtcccccct gctcgagctg cttgtcgtag gccttgtcgt 2040
gctcgagggc cgcggcgtct gcctcgttga ccggctctcc cttgtcgagt ccgttgaagg 2100
gtccgaggta cttgtagccа ggaagcacca gacccccggcc gtcgtcctgc ttttgctggt 2160
tggctttggg tttcgggggct ccaggtttca agtcccacca ctcgcgaatg cctcagaga 2220
ggttgtcctc gagccaatct ggaagataac catcggcagc catacctggt ttaagtcatt 2280
tattgctcag aaacacagtc atccaggtcc acgttgacca gatcgcaggc cgagcaagca 2340
atctcgggag cccgccccag cagatgatga atggcacaga gtttccgata cgtcctcttt 2400
ctgacgaccg gttgagattc tgacacgccg gggaaacatt ctgaacagtc tctggtcccg 2460
tgcgtgaagc aaatgttgaa attctgattc attctctcgc atgtcttgca gggaaacagc 2520

[0020]

atctgaagca tgcccgcgtg acgagaacat ttgttttggt acctgtcggc aaagtccacc 2580
ggagctcctt ccgcgtctga cgtcgatgga tccgcgactg aggggcaagc ccgcttgggc 2640
tcgcttttat ccgcgtcatc gggggcgggt ctcttgttgg ctccaccctt tctgacgtag 2700
aactcatgcg ccacctcggt cacgtgatcc tgcgcccagc ggaagaactc tttgacttcc 2760
tgctttgtca ccttgccaaa gtcatgctcc agacgggcgg tgagttcaaa tttgaacatc 2820
cggtcttgta acggctgctg gtgctcgaag gtggtgctgt tcccgtcaat cacggcgcac 2880
atgttggtgt tggaagtgac gatcacgggg gtgggatcga tctgggcgga cgacttgcac 2940
ttttggtcca cgcgcacctt gctgccgccg agaatggcct tggcggactc cacgaccttg 3000
gccgtcatct tgccctcctc ccaccagatc accatcttgt cgacgcaatc gttgaaggga 3060
aagttctcat tggtccagtt gacgcagccg aagggcgaat tc 3102

<210> 18
<211> 3106
<212> DNA
<213> 新AAV血清型,克隆7-2

[0021]

<400> 18
gcggccgcga attcgccctt cgcagagacc aaagttcaac tgaaacgaat cagccggttt 60
attgattaac aagtaattac aggttacggg tgaggtaacg ggtgccaatg gggcgaggct 120
cagtataaac cccttcgttg ttgacagcaa attccacatt attagacttg gcataatttg 180
aggtgtactg aatctctgga ttccagcgtt tgctgttttc tttctgcagt tcccactcga 240
tcttccacgct gacctggccg gtgctgtact gcgtgataaa tgaggcaaac ttggcaggag 300
taaacacctc tggaggatta gcaggtaccg gggtgttttt gatgagaatt tgaggaggcg 360
ggtgtttgag tccaaatccg ccccatcaggg gagacgggtg aaagttgccg tccgtgtgag 420
gaattttggc ccagatggga ccctgcaggt acacgtcccg gttctgccag accatgccgg 480
gcagagcccc ctggctgttg acagtctgtg tctggggtcc ggccgtagac gattgcaggt 540
tgctggagac cacaccgtat tcttctgtag ccacgggatt ggtggttttg atctcctcct 600
cgctggtcat tagcacgttt tccagcgttg tcttgttggc agccccccgtt ttgccaaaaa 660
ccagcactcc gttgatggga aagaactggt cctcgtcgtc cttgttggtg gccatggcta 720
cgcccgggtt ggttaatgaa tttctaccat tcagatggta tttagtggcc ccggtccagg 780
caaagttact gttgttgttg ctgtctatgt ttttttgacag tctctgctgc cgataacagg 840
gtccgggcag ccagttcttt gattgctcgg ccatggtgtt gggcccagcc tgatggaact 900
gcagctccct tgtggacccc gtagtgctct gggtccgggc caggtagtac aggtactggt 960
cgatgagggg attcatcagc cggtccaggc tctggctgtg cgcatagctg ctgtggaaag 1020
gcacttcctc aaaggtgtag ctgaattcaa agttategcc cgttctcagc atctgagaag 1080
gaaagtactc caggcagtag aaggaggaac gtcccacaga ctgactgccg ttgtttagag 1140
tcagatatcc gtactgagga atcatgaaca cgtccgcagg gaacggaggg aggcagccct 1200
ggtgcgcaga gccgaggacg tacggcagtt ggtactccga gtccgagaag acctgaatcg 1260
tgctggtaag gttattagcg atggtcgtaa cgccgtcgtt cgtcgtgacc tccttgacct 1320
ggatgttgaa caacttgaac cgcagctttc tgggccggaa tcccccagttg ttgttgatga 1380
gtcgctgcca gtcacgtggt gagaagtggc agtggaatct gttgaagtca aaatagcccc 1440
agggggtgct gtagccgaag aagtggttgt cgttggtagc cccgctctga cttgatatct 1500
gcttgtagag gtggttgttg taggtgggca gggccccaggt gcgggtgctg gtggtgatga 1560
ctctgtcgcc cagccatgtg gaatcgcaat gccaatttcc ggaggcatta cccactccgt 1620
cggcgccttc gttattgtct gccattggtg cgccaccgcc tgcagccatt gtaccagatc 1680
ccagacctga gggcgcggcg ggaggttctc cgagaggttg ggggtcgggc actgactctg 1740
agtcgccagt ctgcccaaag ttgagcttct ttttagcggg cggctggccg ttcttgccga 1800
tgcccgtgga ggagtcgggg gattctatgg gtctcttctt tccaggagcc gtcttagcgc 1860
cttcctcaac cagaccgaga ggttcgagaa cccgcttctt ggcctggaag actgctcgcc 1920
cgaggttgcc cccaaaagac gtatcttctt gaagacgctc ctgaaactcg gcgtcggcgt 1980
ggttgtactt gaggtacggg ttgtcccccct gctcgagctg cttgtcgtag gccttgtcgt 2040
gctcgagggc cgcggcgtct gcctcgttga ccggctctcc cttgtcgagt ccgttgaagg 2100
gtccgaggta cctgtagcca ggaagcacca gaccccggcc gtcgtcctgc ttttgctggt 2160
tggctttggg tttcgggggct ccaggtttca agtcccacca ctcgcgaatg ccctcagaga 2220
ggttgccctc gagccaatct ggaagataac catcggcagc catacctggt ttaagtcatt 2280
tattgctcag aaacacagtc atccaggtcc acgttggcca gatcgcaggc cgagcaagca 2340
atctcgggag ccccccccag cagatgatga atggccacaga gtttccgata cgtcctcttt 2400
ctgacgaccg gttgagattc tgacacgccg gggaaacatt ctgaacagtc tctggtcccg 2460
tgcgtgaagc aaatgttgaa attctgattc attctctcgc atgtcttgca ggggaacagc 2520
atctgaagca tgcccgcgtg acgagaacat ttgttttggt acctgtcggc aaagtccacc 2580
ggagctcctt ccgcgtctga cgtcgatgga tccgcgactg aggggcaggc ccgcttgggc 2640
tcgcttttat ccgcgtcatc gggggcgggt ctcttgttgg ctccaccctt tctgacgtag 2700

aactcatacg ccacctcggt cacgtgatcc tgcgcccagc ggaagaactc tttgacttcc 2760
tgctttgtca ccttgccaaa gtcatgctcc agacggcggg tgagttcaaa tttgaacatc 2820
cggtcttgta acggctgctg gtgctcgaag gtggtgctgt tcccgtcaat cacggcgcac 2880
atgttggtgt tggaagtgac gatcacgggg gtgggatcga tctgggcgga cgacttgcac 2940
ttttggtcca cgcgcacctt gctgccgcrg agaatggcct tggcggactc cacgaccttg 3000
gccgtcatcc tgcctcctc ccaccagatc accatcttgt cgacgcaatc gttgaaggga 3060
aagttctcat tggtccagtt gacgcagccg tagaaagggc gaattc 3106

<210> 19
<211> 3105
<212> DNA
<213> 新AAV血清型,克隆C1

<400> 19
gaattcgccc ttgctgcgtc aactggacca atgagaactt tcccttcaac gattgcgtcg 60
acaagatggt gatctggtgg gaggagggca agatgacccc caaggtcgtg gagtccgcca 120
aggccattct gggcggaagc aaggtgcgcg tggaccaaaa gtgcaagtca tcggcccaga 180
tcgaccccac gccgtgatc gtcacctcca acaccaacat gtgcgccgtg atcgacggga 240
acagcaccac cttcgagcac cagcagccgc tgcaggaccg catgttcaag ttcgagctca 300
cccgccgtct ggagcacgac tttggcaagg tgaccaagca ggaagtcaaa gagttcttcc 360
gctgggctca ggatcacgtg actgaggtgg cgcatgagtt ctacgtcaga aagggcggag 420
ccaccaaaag acccgccccc agtgacgcgg atataagcga gcccaagcgg gcctgcccct 480
cagttgcgga gccatcgacg tcagacgcgg aagcaccggt ggactttgcg gacaggtacc 540
aaaacaaatg ttctcgtcac gcgggcatgc ttcagatgct gtttcctgc aagacatgcg 600
agagaatgaa tcagaatttc aacgtctgct tcacgcacgg ggtcagagac tgctcagagt 660
gcttccccgg cgcgtcagaa tctcaacccg tcgtcagaaa aaagacgtat cagaaactgt 720
gcgcgattca tcatctgctg gggcgggcac ccgagattgc gtgttcggcc cgcgatctcg 780
tcaacgtgga cttggatgac tgtgtttctg agcaataaat gacttaaacc aggtatggct 840
gctgacggtt atcttccaga ttggctcgag gacaacctct ctgagggcat tcgcgagtgg 900
tgggacctga aacctggagc cccaagccc aaggccaacc agcagaagca ggacgacggc 960
cggggtctgg tgcttcctgg ctacaagtac ctcggaccct tcaacggact cgacaagggg 1020
gagccccgtca acgcggcgga cgcagcggcc ctcgagcacg acaaggccta cgaccagcag 1080
ctcaaagcgg gtgacaatcc gtacctgcgg tataaccacg ccgacgccga gtttcaggag 1140
cgtctgcaag aagatacgtc ttttgggggc aacctcgggc gagcagtctt ccaggccaag 1200
aagagggtac tcgaacctct gggcctggtt gaagaaggtg ctaagacggc tcctggaaag 1260
aagagaccgt tagagtccac acaagagccc gactcctcct caggaatcgg caaaaaaggc 1320
aaacaaccag ccaaaaagag actcaacttt gaagaggaca ctggagccgg agacggaccc 1380
cctgaaggat cagataccag cgccatgtct tcagacattg aaatgcgtgc agcacgggac 1440
ggaaatgctg tcgatgcggg acaaggttcc gatggagtgg gtaatgcctc gggtgattgg 1500
cattgcgatt ccacctggtc tgagggcaag gtcacaacaa cctcgaccag aacctgggtc 1560
ttgccacct acaacaacca cttgtacctg cggctcggaa caacatcaaa cagcaacacc 1620
tacaacggat tctccacccc ctggggatac tttgacttta acagatttca ctgtcacttc 1680
tcaccacgtg actggcaaag acicatcaac aacaactggg gactacgacc aaaagccatg 1740
cgcgttaaaa tcttcaatat ccaagttaag gaggtcacaa cgtcgaacgg cgagactacg 1800
gtcgctaata accttaccag cacggttcag atattgcgg actcgtcgta tgagctcccg 1860
tacgtgatgg acgctggaca agagggaagt ctgtctcctt tcccaatga cgtcttcatg 1920
gtgccttcaat atggctactg tggcattgtg actggcgaaa atcagaacca gacggacaga 1980
aatgctttct actgcctgga gtattttcct tcacaaatgc tgagaactgg caataacttt 2040
gaaatggctt acaactttgg gaaggtgccg ttccactcaa tgtatgctta cagccagagc 2100
ccggacagac tgatgaatcc cctcctggac cagtacctgt ggcacttaca gtcgaccacc 2160
tctggagaga ctctgaatca aggcaatgca gcaaccacat ttggaaaaat caggagtgga 2220
gactttgcct tttacagaaa gaactggctg cctgggcctt gtgttaaaca gcagagactc 2280
tcaaaaactg ccagtcaaaa ttacaagatt cctgccagcg ggggcaacgc tctgttaaag 2340
tatgacaccc actatacctt aaacaaccgc tggagcaaca tagcgcctgg acctccaatg 2400
gcaacagctg gaccttcaga tggggacttc agcaacgccc agctcatctt ccctggacca 2460
tcagtcaccg gaaacacaac aacctcagca aacaatctgt tgtttacatc agaagaagaa 2520
attgctgccc ccaaccccag agcaacggac atgttttggt cagattgctg caataatcag 2580
aatgctacaa ctgctcccat aacggcaac gtgactgcta tgggagtgct tcctggcatg 2640
gtgtggcaaa acagagacat ttactaccaa gggccaattt gggccaagat cccacacgcg 2700
gacggacatt ttcatcctic accgctaatt ggcggtttig gactgaaaca tccgcctccc 2760
cagatattta tcaaaaacac ccccgtacct gccaatcctg cgacaacctt cactgcagcc 2820
agagtggact ctttcatcac acaatacagc accggccagg tcgctgttca gattgaatgg 2880

[0022]

gaaatcgaaa aggaacgctc caaacgctgg aatcctgaag tgcagtttac ttcaaactat 2940
gggaaccagt cttctatgtt gtgggctccc gatacaactg ggaagtatac agagccgcgg 3000
gttattggct ctcgttattt gactaatcat ttgtaactgc ctagttaatc aataaaccgt 3060
gtgattcgtt tcagttgaac tttggtctct gcgaagggcg aattc 3105

<210> 20
<211> 3105
<212> DNA
<213> 新 AAV 血清型, 克隆 C3

<400> 20
gaattcgccc ttgctgcgtc aactggacca atgagaactt tccattcaac gattgcgtcg 60
acaagatggt gatctggtgg gaggagggca agatgaccgc caaggtcgtg gagtccgcca 120
aggccattct gggcggaagc aaggtgcgcg tggaccaaaa gtgcaagtca tcggcccaga 180
tcgaccccac gccccgtgat gtcacctcca acaccaacat gtgcgccgtg atcgacggga 240
acagcaccac cttcgagcac cagcagccgc tgcaggaccg catgttcaag ttcgagctca 300
cccgccgtct ggagcacgac tttggcaagg tgaccaagca ggaagtcaaa gagttcttcc 360
gctgggctca ggatcacgtg actgaggtgg cgcatgagtt ctacgtcaga aagggcggag 420
ccaccaaaag acccgccccc agtgacgcgg atataagcga gcccaagcgg gcctgcccct 480
cagttgcgga gccatcgacg tcagacgcgg aagcaccggt ggactttgcg gacaggtacc 540
aaaacaaatg ttctcgtcac gcgggcatgc ttcagatgct gtttccctgc aagacatgcg 600
agagaatgaa tcagaattte aacgtctgct tcacgcacgg ggtcagagac tgctcagagt 660
gcttcccggg cgcgtcagaa tctcaacccg tcgtcagaaa aaagacgtat cagaaactgt 720
gcgcgattca tcatctgctg gggcgggcac ccgagattgc gtgttcggcc tgcgatctcg 780
tcaacgtgga cttggatgac tgtgtttctg agcaataaat gacttaaacc aggtatggct 840
gctgacggtt atcttccaga ttggctcgag gacaacctct ctgagggcat tcgcgagtgg 900
tgggacctga aacctggagc ccccaagctc aaggccaacc agcagaagca ggacgacggc 960
cggggtctgg tgcttcctgg ctacaagtac ctcggaccct tccacggact cgacaagggg 1020
gagcccgtca acgcggcgga cgcagcggcc ctcgagcacg acaaggccta cgaccagcag 1080
ctcaaagcgg gtgacaatcc gtacctgcgg tataaccacg ccgacgccga gtttcaggag 1140
cgtctgcaag aagatacgtc ttttgggggc aacctcgggc gagcagtctt ccaggccaag 1200
aagagggtac tcgaaccact gggcctggtt gaagaaggtg ctaagacggc tcctggaaag 1260
aagagaccgt tagagtcacc acaagagccc gactcctcct caggaatcgg caaaaaaggc 1320
aaacaaccag ccaaaaagag actcaacttt gaagaggaca ctggagccgg agacggaccc 1380
cctgaaggat cagataccag cgccatgtct tcagacattg aaatgcgtgc agccaccgggc 1440
ggaaatgctg tcgatgcggg acaaggttcc gatggagtgg gtaatgcctc gggtgattgg 1500
cattgcgatt ccacctggtc tgagggcaag gtcacaacaa cctcgaccag aacctgggtc 1560
ttgcccacct acaacaacca cttgtacctg cggctcggaa caacatcaaa cagcaacacc 1620
tacaacggat tctccaccccc ctggggatac tttgacttta acagattcca ctgtcacttc 1680
tcaccacgtg actggcaaag actcatcaac aacaactggg gactacgacc aaaagccatg 1740
cgcgttaaaa tcttcaatat ccaagttaag gaggtcacaa cgtcgaacgg cgagactacg 1800
gtcgctaata accttaccag cacggttcag atatttgcgg actcgtcgta tgagctcccg 1860
tacgtgatgg acgctggaca agagggaagt ctgcctcctt tccccaatga cgtcttcatg 1920
gtgccttaat atggctactg tggcattgtg actggcgaaa atcagaacca gacggacaga 1980
aatgcttict actgcctgga gtatttttct tcacaaatgc tgagaactgg caataacttt 2040
gaaatggctt acaacttiga gaaggtgccg tttcactcaa tgtatgctca cagccagagc 2100
ctggacagac tgatgaatcc cctcctggac caglacctgt ggcacttaca gtcgaccacc 2160
tctggagaga ctctgaatca aggcaatgca gcaaccacat ttggaaaaat caggagtgga 2220
gactttgcct tttacagaaa gaactggctg cctgggcctt gtgttaaaca gcagagattc 2280
tcaaaaactg ccagtcaaaa ttacaagatt cctgccagcg ggggcaacgc tctgttaaag 2340
tatgacaccc actatacctt aaacaaccgc tggagcaaca tagcgcctgg acctccaatg 2400
gcaacagctg gaccttcaga tggggacttc agcaacgccc agctcatctt cctggacca 2460
tcagtcaccg gaaacacaac aacctcagca aacaatctgt tgtttacatc agaaggagaa 2520
attgctgcca ccaacccaag agacacggac atgtttggtc agattgctga caataatcag 2580
aatgctacaa ctgctcccat aacggcaac gtgactgcta tgggagtgct tcctggcatg 2640
gtgtggcaaa acagagacat ttactaccaa gggccaattt gggccaagat ccccacacgcg 2700
gacggacatt ttcatcctic acccctaatt ggcggttttg gactgaaaca tccgcctccc 2760
cagatattta tcaaaaacac ccccgtacct gccaatcctg cgaccaacct cactgcagcc 2820
agagtggact ctttcatcac acaatacagc accggccagg tcgctgttca gattgaatgg 2880
gaaatcgaaa aggaacgctc caaacgccgg aatcctgaag tgcagtttac ttcaaactat 2940
gggaaccagt cttctatgtt gtgggctccc gatacaactg ggaagtatac agagccgcgg 3000
gttattggct ctcgttattt gactaatcat ttgtaactgc ctagttaatc aataaaccgt 3060

[0023]

gtgattcgtt tcagttgaac tttggtctct gcgaagggcg aattc 3105

<210> 21
<211> 3105
<212> DNA
<213> 新 AAV 血清型, 克隆 C5

<400> 21

[0024]

gaattcgccc ttcgcagaga ccaaagttca actgaaacga atcacacggt ttattgatta 60
actaggcagt tacaaatgat tagtcaaata acgagagcca ataacccgcg gctctgtata 120
cttcccagtt gtatcgggag cccacaacat agaagactgg ttcccacagt ttgaagtaaa 180
ctgcacttca ggattccagc gtttggagcg ttcctttttg atttcccatt caatctgaac 240
agcgacctgg ccggtgctgt attgtgtgat gaaagagtcc actctggctg cagtgaaggt 300
tgtcgcagga taggcaggta cgggggtgtt tttgataaat atctggggag gcggatgttt 360
cagtccaaaa ccgccaatta gcggtgaagg atgaaaatgt ccgtccgcgt gtgggatctt 420
ggcccaaatt ggcccttggt agtaaatgtc tctgttttgc cacaccatgc caggaagcac 480
tcccatagca gtcacgttgc cggttatggg agcagttgta gcattctgat tattgtcagc 540
aatctgacca aacatgtccg tgtctcttgg gttggtggca gcaatttctt cttctgatgt 600
aaacaacaga ttgtttgctg aggttgttgt gtttccggtg actgatggtc cagggaagat 660
gagctgggcg ttgctgaagt ccccatctga aggtccagct gttgccattg gaggtccagg 720
cgctatgttg ctccagcggt tgtttaaggt atagtgggtg tcatacttta acagagcgtt 780
gccccgctg gcaggaatct tgtaattttg actggcagtt tttgagaatc tctgctgttt 840
aacacaaggc ccaggcagcc agttctttct gtaaaaggca aagtctccac tcctgatttt 900
tccaaatgtg gttgctgcat tgccttgatt cagagtctct ccagaggtgg tcgactgtaa 960
gtgccacagg tactggtcca ggaggggatt catcagtccg tccaggctct ggctgtgagc 1020
atacattgag tggaacggca ccttctcaaa gttgtaagcc gtttcaaagt tattgccagt 1080
tctcagcatt tgtgaaggaa aatactccag gcagtagaaa gcatttctgt ccgtctggtt 1140
ctgattttcg ccagtcacaa tgccacagta gccatattga ggcaccatga agacgtcatt 1200
ggggaaagga ggcagacttc cctcttgtcc agcgtccatc acgtacggga gctcatacga 1260
cgagtccgca aatatctgaa ccgtgctggt aaggttatta gcgaccgtag tctcgccgtt 1320
cgacgttgtg acctccttaa cttggatatt gaagatttta acgcgcatgg cttttggtcg 1380
tagtcccag ttgttgttga tgagtctttg ccagtcacgt ggtgagaagt gacagtggaa 1440
tctgttaaag tcaaagtatc cccaggggct ggagaatccg ttgtaggtgt tgctgtttga 1500
tgttgttccg agccgcaggt acaagtggtt gttgtaggtg ggcaagaccc aggttctggt 1560
cgaggttgtt gtgaccttgc cctcagacca ggtggaatcg caatgccaat caccgaggcc 1620
attacccact ccatcggaac cttgtcccgc atcgacagca tttccgcccg gtgctgcacg 1680
catttcaatg tctgaagaca tggcgctggt atctgatcct tcaggggctc cgtctccggc 1740
tccagtgtcc tcttcaaagt tgagtctctt tttggctggt tgtttgcctt ttttgccgat 1800
tcctgaggag gagtcgggct cttgtggtga ctctaacggt ctcttctttc caggagccgt 1860
cttagcacct tcttcaacca ggcccagagg ttcgagtacc ctcttcttgg cctggaagac 1920
tgctcgcccg aggttgcccc caaaagacgt atcttcttgc agacgctcct gaaactcggc 1980
gtcggcgtgg ttataccgca ggtacggatt gtcacccgct ttgagctgct ggtcgtaggc 2040
cttgtcgtgc tcgagggccg ctgcgtccgc cgcgttgacg ggctccccct tgtcgagtcc 2100
gttgaagggt ccgaggtact cgtagccagg aagcaccaga ccccggccgt cgtcctgctt 2160
ctgctggttg gccttgggct tgggggctcc aggtttcagg tcccaccact cgcgaatgcc 2220
ctcagagagg ttgtcctcga gccaatctgg aagataaccg tcagcagcca tacctggttt 2280
aagtcattta ttgctcagaa acacagtcat ccaagtccac gttgacgaga tcgcaggccg 2340
aacacgcaat ctcgggtgcc cgccccagca gatgatgaat cgcgcacagt ttctgatacg 2400
tctttttct gacgacgggt tgagattctg acgcgccggg gaagcactct gagcagtctc 2460
tgaccccgtg cgtgaagcag acgttgaaat tctgatttat tctctcgcat gtcttgcagg 2520
gaaacagcat ctgaagcatg cccgcgtgac gagaacattt gttttggtac ctgtccgcaa 2580
ggtccaccgg tgcttccgcg tctgacgtcg atggctccgc aactgagggg caggccgct 2640
tgggctcgct tatatccgcg tcactggggg cgggtctttt ggtggctccg cctttctga 2700
cgtagaactc atgcgccacc tcagtcacgt gatcctgagc ccagcggaag aactctttga 2760
cttcctgctt ggtcaccttg ccaaagtcgt gctccagacg gcgggtgagc tcgaacttga 2820
acatgcggtc ctgcagcggc tgctggtgct cgaaggtggt gctgttcccg tcgatcacgg 2880
cgcacatgtt ggtgttggag gtgacgatca cgggcgtggg gtcgatctgg gccgatgact 2940
tgcacttttg gtcccacgcg accttgcttc cgcccagaat ggccttggcg gactccacga 3000
ccttggcggt catcttgccc tcctcccacc agatcaccat cttgtcgacg caatcgttga 3060
agggaaagtt ctcattggtc cagttgacgc agcaagggcg aattc 3105

<210> 22
<211> 3094
<212> DNA
<213> 新 AAV 血清型,克隆 F1

<400> 22

gaattcgccc ttgctgcgtc aactggacca agagaacttt cccttcaacg attgcgtcga 60
caagatggtg atctggtggg aggagggcaa gatgacggcc aaggtcgtgg agtccgccaa 120
agccattctg ggcggaagca aggtgcgcgt cgaccaaaag tgcaagtcct cggcccagat 180
cgatcccacc ccgtgatcg tcacctccaa caccaacatg tgcgccgtga tcgacgggaa 240
cagcaccacc ttcgagcacc agcagccgtt gcaggaccgg atgttcaaat ttgaactcac 300

ccgccgtctg gaacacgact ttggcaaggt gaccaagcag gaagtcaaag agttcttccg 360
ctgggctagt gatcacgtga ctgagggtgc gcatgagttc tacgtcagaa agggcggagc 420
cagcaaaaga ccccgcccccg atgacgcgga tataagcgag cccaagcggg cctgtccctc 480
agtcacggac ccatcgacgt cagacgcgga aggagctccg gtggactttg ccgacaggta 540
ccaaaacaaa tgttctcgtc acgcgggcat gcttcagatg ctgtttccct gcaaaacgtg 600
cgagagaatg aatcagaatt tcaacatttg cttcacgcac ggggtcagag actgtttaga 660
atgtttcccc ggcgtgtcag aatctcaacc ggtcgtcaga aaaaagacgt atcggaagct 720
gtgtgcgatt catctctgc tggggcgggc acccgagatt gcttgctcgg cctgcgacct 780
ggtcaacgtg gacctggacg actgtgtttc tgagcaataa atgacttaaa ccgggtatgg 840
ctgccgatgg ttatcttcca gattggctcg aggacaacct ctctgagggc attcgcgagt 900
ggtgggacct gaaacctgga gccccgaaac ccaaagccaa ccagcaaaag caggacgacg 960
gccggggtct ggtgcttcct ggctacaagt acctcggacc cttcaacgga ctcgacaagg 1020
gggagccccgt caacgcggcg gacgcagcgg ccctcgagca cgacaaggcc tacgaccagc 1080
agctcaaagc gggtgacaat ccgtacctgc ggtataacca cgccgacgcc gagtttcagg 1140
agcgtctgca agaagatacg tcatttgggg gcaacctcgg gcgagcagtc ttccaggccg 1200
agaagcgggt tctcgaacct ctcggtctgg ttgaggaagg cgctaagacg gctcctggaa 1260
agaagagacc catagactct ccagactcct ccacgggcat cggcaaaaaa ggccagcagc 1320
ccgctaaaaa gaagctcaat tttggtcaga ctggcgactc agagtcagtc ccccgaccct 1380
aacctctigg agaacctcca gcagcgccct ctagtgtggg atctggtaca atggctgcag 1440
gcggtggcgc accaatggca gacaataacg aaggtgccga cggagtgggt aatgcctcag 1500
gaaattggca ttgcgattcc acatggctgg gcgacagagt catcaccacc agcaccagaa 1560
cctgggccct ccccacctac aacaaccacc tctacaagca aatctccagc agcagctcag 1620
gagccaccaa tgacaaccac tacttcggct acagcacccc ctggggtat tttgacttta 1680
acagattcca ctgccacttc tcaccacgtg actggcagcg actcatcaac aacaactggg 1740
gattccggcc caagaagctg cggttcaagc tcttcaacat ccaggtcaag gaggtcacaa 1800
cgaatgacgg cgtcacgacc atcgctaata accttaccag cacggttcag gtcttctcgg 1860
actcggaata ccagctgccg tacgtcctcg gctctgcgca ccagggctgc ctgcctccgt 1920
tcccggcgga cgtcttcatg attcctcagt acggctacct gactctgaac aacggcagcc 1980
aatcggtggg ccgttcctcc ttctactgcc tggaatattt ccctctcaa atgctgagaa 2040
cgggcaacaa ctttgagttc agttacagct tcgaggacgt gcctttccac agcagctacg 2100
cgcacagcca gagcctagac cggctgatga accctctcat cgaccagtac ctgtactacc 2160
tggcccggac ccagagcacc acgggttcca ccagggaact gcaatttcat caagctgggc 2220
ccaatactat ggccgagcag tcaaagaact ggctgcctgg accctgctat aggcaacagg 2280
gactgtcaaa gaacttggac tttaacaaca acagcaattt tgcctggact gctgccacta 2340
aatatcatct gaatggcaga aactctttga ccaatcctgg cattcccatg gcaaccaaca 2400
aggatgatga ggaccagttc tttcccatca acggggtact ggttttggc aagacgggag 2460
ctgccaacaa aactacgctg gaaaacgttc tgatgaccag cgaggaggag atcaagacca 2520
ctaaccctgt ggctacagaa gaatacggtg tggtctccag caacctgcag cgtctacag 2580
ccgggcctca atcacagact atcaacagcc agggagcact gcctggcatg gtctggcaga 2640
accgggacgt gtatctgcag ggtcccatct gggccaaaat tcctcacacg gatggcaact 2700
ttcaccgtc tctctgatg ggcggttttg gactcaaaca ccccgcctcca cagatcctga 2760
tcaaaaacac acctgtacct gctaatcctc cggaggtgtt tactcctgcc aagtttgcct 2820
ccttcatcac gcagtacagc accggacaag tcagcgtgga aatcgagtgg gagctgcaga 2880
aagaaaacag caagcgctgg aacccagaaa ttcagtatac ttccaattat gccaagtcta 2940
ataatgttga atttgctgtg aacctgatg gtgtttatac tgagcctcgc cccattggca 3000
ctcgttacct ccccgtaat ctgtaattgc ttgttaatca ataaaccggt tgattcgttt 3060
cagttgaact ttggtctctg cgaagggcga attc 3094

<210> 23
<211> 3095

[0025]

<212> DNA
<213> 新 AAV 血清型, 克隆 F3

<400> 23

gaattcgccc ttcgcagaga ccaaagttca actgaaacga atcaaccggt ttattgatta 60
acaagcaatt acagattacg ggtgaggtaa cgagtgccaa tggggcgagg ctcagtataa 120
acaccatcag ggttcacagc aaattcaaca ttattagact tggcataatt ggaagtatac 180
tgaatttctg ggttccagcg cttgctgttt tctttctgca gctcccactc gatttccacg 240
ctgacttgtc cggtgctgta ctgcgtgatg aaggaggcaa acttggcagg agtaaacacc 300
tccggaggat tagcaggtac aggtgtgttt ttgatcagga tctgtggagg cgggtgtttg 360
agtccaaaac cgcccatcag aggagacggg tgaaagttgc catccgtgtg aggaattttg 420
gcccagatgg gaccctgcag atacacgtcc cggttctgcc agaccatgcc aggcagtgct 480
ccctggctgt tgatagtctg tgattgaggc ccggctgtag acgactgcag gttgctggag 540
accacaccgt attcttctgt agccacaggg ttagtggtct tgatctcctc ctcgctggtc 600
atcagaacgt tttccagcgt agttttgttg gcagctcccg tcttgcccaaa aaccagtacc 660
ccgttgatgg gaaagaactg gtcctcatca tcctttgttgg ttgccatggg aatgccagga 720
ttggtcaaag agtttctgcc attcagatga tatttagtgg cagcagtcca ggcaaaattg 780
ctgttgttgt taaagtccaa gttctttgac agtctctgtt gcctatagca gggtccaggc 840
agccagttct ttgactgctc ggccatagta ttgggcccag cttgatgaaa ttgcagttcc 900
ctggtggaac ccgtggtgct ctgggtccgg gccagglagt acaggtactg gtcgatgaga 960
gggttcatca gccggtctag gctctggctg tgcgcgtagc tgctgtggaa aggcacgtcc 1020
tcgaagctgt aactgaactc aaagttgttg cccgttctca gcatttgaga ggggaaatat 1080
tccaggcagt agcaggagga acggcccacc gattggctgc cgttgtccag agtcaggtag 1140
ccgtactgag gaatcatgaa gacgtccgcc gggaacggag gcaggcagcc ctggtgcgca 1200
gagccgagga cgtacggcag ctggtattcc gagtccgaga agacctgaac cgtgctggta 1260
aggttattag cgatggtcgt gacgccgtca ttcgttgtga cctccttgac ctggatgttg 1320
aggagcttga accgcagctt cttgggccgg aatccccagt tgttgttgat gagtcgctgc 1380
cagtcacgtg gtgagaagtg gcagtggaat ctgttaaagt caaaataccc ccagggggtg 1440
ctgtagccga agtagtggtt gtcattggtg gctcctgagc tgctgctgga gatttgcttg 1500
tagaggtggt tgttgtaggt ggggagggcc caggttctgg tgctggtggt gatgactctg 1560
[0026] tcgcccagcc atgtggaatc gcaatgccaa tttcctgagg cattaccccac tccgtcggca 1620
ccttcgttat tgtctgccat tggtgcgcca ccgcctgcag ccattgtacc agatccccaca 1680
ctagagggcg ctgctggagg tictccaaga ggttgagggt cggggactga ctctgagtcg 1740
ccagtctgac caaaattgag cttctttita gcgggctgct ggccttttit gccgatgccc 1800
gtggaggagt ctggagagcc tatgggtctc ttctttccag gagccgtctt agcgccttcc 1860
tcaaccagac cgagaggttc gagaacccgc ttcttggcct ggaagactgc tcgcccgagg 1920
ttgccccccaa atgacgtatc ttcttgcaga cgtccctgaa actcggcgtc ggcgtggtta 1980
taccgcaggt acggattgtc acccgctttg agctgctggt cgtaggcctt gtcgtgctcg 2040
agggccgctg cgtccgcccgc gttgacgggc tcccccttgt cgagtccgtt gaagggtccg 2100
aggtacttgt agccaggaag caccagaccc cggccgtcgt cctgcttttg ctggttggct 2160
ttggctttcg gggctccagg tttcaggtcc caccactcgc gaatgccctc agagaggttg 2220
tcctcgagcc aatctggaag ataaccatcg gcagccatac ctggtttaag tcatttattg 2280
ctcagaaaca cagtcgtcca ggtccacgtt gaccaggtcg caggccgagc aagcaatctc 2340
gggtgcccgc cccagcagat gatgaatcgc acacagcttc cgatacgtct tttttctgac 2400
gaccggttga gattctgaca cgccggggaa acattctaaa cagtctctga ccccgtgcgt 2460
gaagcaaatg ttgaaattct gattcattct ctcgcacgtt ttgcagggaa acagcacctg 2520
aagcatgccc gcgtgacgag aacatttgtt ttggtacctg tcggcaaagt ccaccggagc 2580
tccttccgcg tctgacgtcg atgggtccgt gactgaggga cgggcccgct tgggctcgct 2640
tatatccgcg tcatcggggg cgggtctttt gctggctccg cccttctga cgtagaactc 2700
atgcgtcacc tcagtcacgt gatcactagc ccagcggaag aactctttga cttcctgctt 2760
tgtcaccttg ccaaagtcgt gttccagacg gcgggtgagt tcaaatttga acatccggtc 2820
ctgcaacggt tgctggtgct cgaaggtggt gctgttcccg tcgatcacgg cgcacatgtt 2880
ggtgttggag gtgacgatca cgggggtggg atcgatctgg gcggacgact tgcacttttg 2940
gtccacgcgc accttgctgc cgccgagaat ggccttggcg gactccacga ccttggccgt 3000
catcttgccc tcctcccacc agatcaccat cttgtcgacg caatcgttga agggaaagtt 3060
ctcattggtc cagttgacgc agcaagggcg aattc 3095

<210> 24
<211> 3095
<212> DNA
<213> 新 AAV 血清型, 克隆 F5

<400> 24

gaattcgccc ttcgcagaga ccaaagttca actgaaacga atcaaccggt ttattgatta 60
acaagcaatt acagattacg ggtgaggtaa cgagtgccaa tggggcgagg ctcagtataa 120
acaccatcag ggttcacagc aaattcaaca ttattagact tggcataatt ggaagtatac 180
tgaatttctg ggttccagcg cttgctgttt tctttctgca gctcccactc gatttccacg 240
ctgacttgtc cggtgctgta ctgcgtgatg aaggaggcaa acttggcagg agtaaacacc 300
tccggaggat tagcaggtac aggtgtgttt ttgatcagga tctgtggagg cgggtgttcg 360
agtccaaaac cgcccatcag aggagacggg tgaaagttgc catccgtgtg aggaattttg 420
gcccagatgg gaccctgcag atacacgtcc cggttctgcc agaccatgcc aggcagtgct 480
ccctggctgt tgatagtctg tgattgaggc ccggctgtag acgactgcag gttgctggag 540
accacaccgt attcttcigt agcccacaggg ttagtggtct tgatctcctc ctcgctggtc 600
atcagaacgt tttccagcgt agttttgttg gcagctcccg tcttgccaaa aaccagtacc 660
ccgttgatgg gaaagaactg gtcctcatca tcctigttgg ttgccatggg aatgccagga 720
ttggtcaaag agttictgcc attcagatga tatttagtgg cagcagtcca ggcaaaattg 780
ctgttgttgt taaagtccaa gttctttgac agtctctgtt gcctatagca gggtccaggc 840
agccagttct ttgactgctc ggccatagta ttgggcccag cttgatgaaa ttgcagttcc 900
ctggtggaac ccgtggtgct ctgggtccgg gccaggtagt acaggtactg gtcgatgaga 960
gggttcatca gccggtctag gctctggctg tgcgcgtagc tgctgtggaa aggcacgtcc 1020
tcgaagctgt aactgaactc aaagttgttg cccgttctca gcatttgaga ggggaaatat 1080
tccaggcagt agaaggagga acggcccacc gattggctgc cgttgttcag agtcaggtag 1140
ccgtactgag gaatcatgaa gacgtccgcc gggaacggag gcaggcagcc ctggtgcgca 1200
gagccgagga cgtacggcag ctggtattcc gagtccgaga agacctgaac cgtgctggta 1260
aggttattag cgatggtcgt gacgccgtca ttcgttgtga cctccttgac ctggatgttg 1320
aagagcttga acagcagctt cttgggccgg aatccccagt tgttgttgat gagtcgctgc 1380
cagtcacgtg gtgagaagtg gcagtggaat ctgttaaagt caaaataccc ccagggggtg 1440
ctgtagccga agtagtggtt gtcattggtg gctcctgagc tgctgctgga gatttgcttg 1500
tagaggtggt tgttgtaggt ggggagggcc caggttctgg tgctggtggt gatgactctg 1560
tgccccagcc atgtggaatc gcaatgccaa tttcctgagg cattacccac tccgtcggca 1620
ccttcgttat tgtctgccgt tggtgcgcca ccgcctgcag ccattgtacc agatccccaca 1680
ctagagggcg ctgctggagg ttctccaaga ggttgagggt cggggactga ctctgagtcg 1740
ccagtctgac caaaattgag cttcllllla gcgggctgct ggccllllll gccgatgccc 1800
gtggaggagt ctggagagtc tatgggtctc ttctttccag gagccgtctt agcgccttcc 1860
tcaaccagac cgagaggttc gagaacccgc ttcttggcct ggaagactgc tcgcccgagg 1920
ttgccccccaa atgacgtatc ttcttgcagg cgctcctgaa actcggcgtc ggcgtggtta 1980
taccgcaggt acggattgtc acccgctttg agctgctggt cgtaggcctt gtcgtgctcg 2040
agggccgctg cgtccgccgc gttgacgggc tcccccttgt cgagtccgtt gaagggtccg 2100
aggtacttgt agccaggaag caccagaccc cggccgtcgt cctgcttttg ctggttggct 2160
ttgggtttcg gggctccagg tttcaggtcc caccactcgc gaatgccctc agagaggttg 2220
tcctcgagcc aatctggaag ataaccatcg gcagccatac ctggtttaag ccatttattg 2280
ctcagaaaca cagtcgtcca ggtccacgtt gaccaggtcg caggccgagc aggcaatctc 2340
gggtgcccgc cccagcagat gatgaatcgc acacagcttc cgatacgtct ttttctgtac 2400
gacccggttgn gattctgaca cgccgggggaa acattctaaa cagtctctga cccccgtgcgt 2460
gaagcaaatg ttgaaattct gattcattct ctcgcacgtt ttgcagggaa acagcatctg 2520
aagcatgccc gcgtggcgag aacatttgtt ttggtacctg tcggcaaagt ccacccggagc 2580
tccttccgcg tctgacgtcg atgggtccgt gactgaggga caggcccgct tgggctcgct 2640
tatatccgcg tcatcgggggg cgggtctttt gctggctccg cccttctga cgtagaactc 2700
atgcgtcacc tcagtcacgt gatcactagc ccagcggaag aactctttga cttcctgctt 2760
tgtcacctig ccaaagtcgt gttccagacg gcgggtgagt tcaaatttga acatccggtc 2820
ctgcaacggc tgctggtgct cgaaggtggt gctgttcccg tcgatcacgg cgcgcatgtt 2880
ggtgttggag gtgacgatca cggggtggg atcgatctgg gcggacgact tgcacttttg 2940
gtccacgcgc accttgctgc cgccgagaat ggccttggcg gactccacga cttggccgt 3000
catcttgccc tcctcccacc agatcaccat cttgtcgacg caatcgttga agggaaagtt 3060
ctcattggtc cagttgacgc agcaagggcg aattc 3095

[0027]

<210> 25
<211> 3142
<212> DNA
<213> 新 AAV 血清型, 克隆 H6

<400> 25

aaaacgacgg gccagtgatt gtaatacgac tcactatagg gcgaaattga attagcggc 60

[0028]

cgcgaattcg cctttcgcag agaccaaagt tcaactgaaa cgaattaaac ggtttattga 120
ttaacaagca attacagatt acgagtcagg tatctggtgc caatggggcg aggctctgaa 180
tacacaccat tagtgtccac agtaaagtcc acattaacag acttgttgta gttggaagtg 240
tactgaattt cgggattcca gcgtttgctg ttctccttct gcagctccca ctcgatctcc 300
acgctgacct gtcccgtgga atactgtgtg atgaaagaag caaacttggc agaactgaag 360
tttgtgggag gattggctgg aacgggagtg tttttgatca tgatctgagg aggcgggtgt 420
ttgagtccaa aacctcccat cagtggagaa ggatgaaagt gtccatcggt gtgaggaatc 480
ttggcccaaa tgggtccctg caggtacacg tctcgatcct gccacaccat accaggtaac 540
gctccttggt gattgacagt tccagtagtt ggaccagtgt ttgagttttg caaattattt 600
gacacagtcc cgtactgctc cgtagccacg ggattggtgg ccctgatttc ttcttcatct 660
gtaatcatga cattttccaa atccgcgtcg ttggcatttg ttccttgttt accaaatatc 720
agggttccat gcatgggggaa aaacttttct tcgtcatcct tgtgactggc catagctggt 780
cctggattaa ccaacgagtc ccggccattt agatgatact ttgtagctgc agtccaggga 840
aagttgctgt tgttgttgtc gtttgcctgt tttgacagac gctgctgtct gtagcaaggt 900
ccaggcagcc agtttttagc ttgaagagac atgttggttg gtccagcttg gctaaacagt 960
agccgagact gctgaagagt tccactattt gtttgtgtct tgttcagata atacaggtac 1020
tggtcgatca gaggattcat cagccgatcc agactctggc tgtgagcgta gctgctgtgg 1080
aaaggcacgt cttcaaaagt gtagctgaac tgaaagttgt ttccagtacg cagcatctga 1140
gaaggaaagt actccaggca gtaaaaggaa gagcgtccta ccgcctgact ccccgttgtt 1200
agggtgaggt atccatactg tgggaccatg aagacgtccg ctggaaacgg cgggaggcat 1260
ccttgatgcg ccgagcccag gacgtacggg agctggtact ccgagtcagt aaacacctga 1320
accgtgctgg taaggttatt ggcaatcgtc gtcgtaccgt cattctgcgt gacctctttg 1380
acttgaatat taaagagctt gaagttgagt ctttgggcc ggaatccccg gttgttgttg 1440
acgagtcttt gccagtcacg tggtgaaaag tggcagtgga atctgttgaa gtcaaaatac 1500
ccccagggcg tgctgtagcc aaagtagtgg ttgtcgttgc tggctcctga ttggctggag 1560
atttgcttgt agaggtggtt gttgtatgtg ggcagggccc aggttcgggt gctggtggtg 1620
atgactctgt cgcccagcca ttgggaatcg caatgccaat ttcctgagga attaccccact 1680
ccatcggcac cctcgttatt gtctgccatt ggtgcgccac tgcctgtagc cattgtagta 1740
gatcccagac cagagggggc tgctggtggc tgtccgagag gctgggggtc aggtacggag 1800
tctgcgtctc cagtctgacc aaaatttaat ctttttcttg caggctgctg gcccgctttt 1860
ccggttcccg aggaggagtc tggctccaca ggagagtgct ctaccggcct ctttttttcc 1920
ggagccgtct taacaggctc ctcaaccagg cccagaggtt caagaaccct ctttttcgcc 1980
tggaagactg ctcgtccgag gttgccccca aaagacgtat cttctttaag gcgctcctga 2040
aactctgcgt cggcgtggtt gtacttgagg tacgggttgt ctccgctgtc gagctgccgg 2100
tcgtaggcct tgtcgtgctc gagggccgcg gcgtctgcct cgttgaccgg ctcccccttg 2160
tcgagtccgt tgaagggtcc gaggtacttg tacccaggaa gcacaagacc cctgctgtcg 2220
tccttatgcc gctctgcggg ctttggtggt ggtgggccag gtttgagctt ccaccactgt 2280
cttattcctt cagagagagt gtcctcgagc caatctggaa gataaccatc ggcagccata 2340
cctgatttaa atcatttatt gttcagagat gcagtcatcc aaatccacat tgaccagatc 2400
gcaggcagtg caagcgtctg gcacctttcc catgatatga tgaatgtagc acagtttctg 2460
atacgccttt ttgacgacag aaacgggttg agattctgac acgggaaagc actctaaaca 2520
gtctttctgt ccgtgagtga agcagatatt tgaattctga ttcattctct cgcattgtct 2580
gcagggaaac agcatcagat tcatgcccac gtgacgagaa catttgtttt ggtacctgtc 2640
cgcgtagttg atcgaagctt ccgcgtctga cgtcgatggc tgcgcaactg actcgcgcgc 2700
ccgtttgggc tcacttatat ctgcgtcact gggggcgggt ctttttttag ctccaccctt 2760
tttgacgtag aattcatgct ccacctcaac cacgtgatcc tttgcccacc ggaaaaagtc 2820
tttcacttcc tgcttggtga ccttttccaaa gtcatgatcc agacggcggg taagttcaaa 2880
tttgaacatc cggtcttgca acggctgctg gtgctcgaag gtcgttgagt tcccgtcaat 2940
cacggcgcac atgttggtgt tggaggtgac gatcacggga gtcgggtcta tctgggccga 3000
ggacttgcat ttctggtcca cacgcacctt gcttcctcca agaatggctt tggccgactc 3060
cacgaccttg gcggtcatct tccctcctc ccaccagatc accatcttgt cgacgcaatg 3120
gtaaaaggaa agttctcatt gg 3142

<210> 26
<211> 3075
<212> DNA
<213> 新 AAV 血清型, 克隆 H2

<400> 26
tgagaacttt cctttcaacg attgcgtcgg acaagatggt gatctggtgg gaggagggga 60
agatgaccgc caaggtcgtg gagtcggcca aagccattct tggaggaagc aaggtgcgtg 120
tggaccagaa atgcaagtcc tcggcccaga tagaccccgac tcccgtgatc gtcacctcca 180

[0029]

acaccaacat gtgcgccgtg attgacggga actcaacgac cttcgagcac cagcagccgt 240
tgcaagaccg gatgttcaaa tttgaacttа cccgccgtct ggatcatgac tttggaaagg 300
tcaccaagca ggaagtgaaa gacttttтcc ggtgggcaaa ggatcacgtg gttgaggtgg 360
agcatgaatt ctacgtcaaa aagggtggag ctaagaaaag acccgccccc agtgacgcag 420
atataagtga gcccaaacgg gcgcgcgagt cagttgcgca gccatcaacg tcagacgcgg 480
aagcttcgat caactacgcg gacaggtacc aaaaacnaat gttctcgtca cgtgggcatg 540
aatctgatgc tgtttccctg cagacaatgc gagagaatga atcagaattc aaatatctgc 600
ttcactcacg gacagaaaga ctgtttagag tgctttcccg tgtcagaatc tcaacccgtt 660
tctgtcgtca aaaaggcgta tcagaaactg tgctacattc atcatatcat gggaaaggtg 720
ccagacgctt gcactgcctg cgatctggtc aatgtggatt tggatgactg catctctgaa 780
caataaatga tttaaaatcag gtatggctgc cgatggttat cctccagatt ggctcgagga 840
cacictctci gaagggalaa gacagtggtg gaagctcaaa cctggcccac caccaccaaa 900
gcccgcagag cggcataagg acgacagcag gggtcttgtg cttcctgggt acaagtacct 960
cggacccttc aacggactcg acaaggggga gccggtcaac gaggcagacg ccgcggccct 1020
cgagcacgac aaggcctacg accggcagct cgacagcgga gacaacccgt acctcaagta 1080
caaccacgcc gacgcagagt ttcaggagcg ccttaaagaa gatacgtctt ttgggggcaa 1140
cctcggacga gcagtcttcc aggcgaaaaa gagggttctt gaacctctgg gcctggttga 1200
ggaacctgtt aagacggctc cgggaaaaaa gaggccggta gagcactctc ctgtggagcc 1260
agactcctcc tcgggaaccg gaaaagcggg ccagcggcct gcaagaaaaa gattaaattt 1320
tggtcagact ggagacgcag actccgtacc tgacccccag cctctcggac agccaccagc 1380
agccccctct ggtctgggat ctactacaat ggctacaggc agtggcgcac caatggcaga 1440
caataacgag ggtgccgatg gagtgggtaa ttcctcagga aattggcatt gcgattccca 1500
atggctgggc gacagagtca tcaccaccag caccccgaacc tgggccctgc ccacatacaa 1560
caaccacctc tacaagcaaa tctccagcca atcaggagcc agcaacgaca accactactt 1620
tggctacagc acccccctggg ggtattttga cttcaacaga ttccactgcc acttttcacc 1680
acgtgactgg caaagactca tcaacaacaa ctggggattc cggcccaaaa gactcaactt 1740
caagctcttt aatattcaag tcaaagaggt cacgcagaat gacggtacga cgacgattgc 1800
caataacctt accagcacgg ttcaggtgtt tactgactcg gagtaccagc tcccgtacgt 1860
cctgggctcg gcgcatcaag gatgcctccc gccgttccca gcggacgtct tcatggtccc 1920
acaglatgga tacctcaccc tgaacaacgg gagtcaggcg gtaggacgct cttcttttа 1980
ctgcctggag tactttcctt ctcagatgct gcgtactgga aacaactttc agttcagcta 2040
cacttttgaa gacgtgcctt tccacagcag ctacgctcac agccagagtc tggatcggct 2100
gatgaatcct ctgatcgacc agtacctgta ttatctgaac aagacacaaa caaatagtgg 2160
aactcttcag cagtctcggc tactgtttag ccaagctgga ccaaccaaca tgtctcttca 2220
agctaaaaac tggctgcctg gaccttgcta cagacagcag cgtctgtcaa aacaggcaaa 2280
cgacaacaac aacagcaact ttcctggac tgcagctaca aagtatcatc taaatggccg 2340
ggactcgttg gttaatccag gaccagctat ggccagtcac aaggatgacg aagaaaagtt 2400
tttccccatg catggaaccc tgatatttgg taaacaagga acaaatgcca acgacgcgga 2460
tttggaaaat gtcatgatta cagatgaaga agaaatcagg gccaccaatc ccgtggctac 2520
ggagcagtac gggactgtgt caaataattt gcaaaactca aacactggtc caactactgg 2580
aactgtccat cgccaaggag cgttacctgg tatggtgtgg caggatcgag acgtgtacct 2640
gcagggaccc atttgggcca agattcctca caccgatgga cacttтcatc cttctccact 2700
gatgggaggt tttggactca aacacccgcc tcctcagatc atgatcaaaa acactcccgt 2760
tccagccaat cctcccacaa actcagttc tgccaagttt gcttctttca tcacacagta 2820
ttccacggga caggtcagcg tggagatcga gtgggagctg cagaaggaga acagcaaacg 2880
ctggaatccc gaaattcagt acacttccaa ctacaacaag tctgttaatg tggactttac 2940
tgtggacact aatggtgtgt attcagagcc tcgccccatt ggcaccagat acctgactcg 3000
taatctgtaa ttgcttgtta atcaataaac cgtttaattc gtttcagttg aactttggtc 3060
tctgcgaagg gcgaa 3075

<210> 27
<211> 3128
<212> DNA
<213> 42.8

<400> 27
gaattcgccc tttctacggc tgcgtcaact ggaccaatga gaactttccc ttcaacgatt 60
gcgtcgacaa gatggtgatc tggtgggagg agggcaagat gacggccaag gtcgtggagt 120
ccgccaaggc cattctcggc ggcagcaagg tgcgcgtgga ccaaaagtgc aagtctccg 180
cccagatcga tcccacccca gtgatcgtca cttccaacac caacatgtgc gccgtgattg 240
acgggaacag caccaccttc gagcaccagc agccgttaca agaccggatg ttcaaatttg 300
aactcacccg ccgtctggag cacgactttg gcaaggtgac aaagccaggaa gtcaaagagt 360

[0030]

tcttccgctg ggcgcaggat cacgtgaccg aggtggcgca tgagttctac gtcagaaagg 420
gtggagccaa caagagaccc gccccgatg acgcggataa aagcgagccc aagcgggcct 480
gccccagt cgcggatcca tcgacgtcag acgcggaagg agctccggtg gactttgccg 540
acaggtacca aaacaaatgt tctcgtcacg cgggcatgct tcagatgctg tttccctgca 600
agacatgcga gagaatgaat cagaatttca acatttgctt cacgcacggg acccagagact 660
gttcngaatg tttccccggc gtgtcagaat ctcaaccggt cgtcagaaag aggacgtatc 720
ggaaaactctg tgccattcat catctgctag ggcgggctcc cgagattgct tgctcggcct 780
gcgatctggt caacgtggac ctggatgact gtgtttctga gcaataaatg acttaaacca 840
ggtatggctg ccgatggtta tcttccagat tggctcgagg acaacctctc tgagggcatt 900
cgcgagtggt gggacttgaa acctggagcc ccgaaaccca aagccaacca gcaaaagcag 960
gacgacggcc ggggtctggt gcttcctggc tacaagtacc tcggaccctt caacggactc 1020
gacaaggggg agcccgtcaa cgcggcggac gcagcggccc tcgagcacga caaggcctac 1080
gaccagcagc tcaaagcggg tgacaatccg tacctgcggt ataaccacgc cgacgccgag 1140
tttcaggagc gtctgcaaga agatacgtct tttgggggca acctcgggcg agcagtcttc 1200
caggccaaga agcgggttct cgaacctctc ggtctggttg aggaaggcgc taagacggct 1260
cctggaaaga agagaccggt agagccatca ccccagcgtt ctccagactc ctctacgggc 1320
atcggcaaga caggccagca gccccgcgaaa aagagactca actttgggca gactggcgac 1380
tcagagtcag tgccccgaccc tcaaccaatc ggagaacccc ccgcaggccc ctctggtctg 1440
ggatctggta caatggctgc aggcggtggc gctccaatgg cagacaataa cgaaggcgcc 1500
gacggagtgg gtagttcctc aggaaattgg cattgcgatt ccacatggct gggcgacaga 1560
gtcatcacca ccagcacccg aacctgggcc ctccccacct acaacaacca cctctacaag 1620
caaatctcca acgggacatc gggaggaagc accaacgaca acaccttactt cggctacagc 1680
acccccctggg ggtattttga ctttaacaga ttccactgcc acttctcacc acgtgactgg 1740
cagcgactca tcaacaacaa ctggggattc cggccccaaga gactcaactt caagctcttc 1800
aacatccagg tcaaggaggt cacgcagaat gaaggcacca agaccatcgc caataacctt 1860
accagcacga ttcaggtctt tacggactcg gaataccagc tcccgtacgt cctcggctct 1920
gcgcaccagg gctgcctgcc tccgttcccg gcggacgtct tcatgattcc tcagtacggg 1980
tacctgactc tgaacaacgg cagtcaggcc gtgggccgtt cctccttcta ctgcctggag 2040
tactttcctt ctcaaatgct gagaacgggc aacaacttig agttcagcta ccagtttgag 2100
gacgtgcctt ttcacagcag ctacgcgcac agccaaagcc tggaccggct gatgaaccct 2160
ctcatcgacc agtacctgta ctacctgtct cggactcagt ccacgggagg taccgcagga 2220
actcagcagt tgctatttfc tcaggccggg cctaataaca tgtcggctca ggccaaaaac 2280
tggctacccg ggccctgcta ccggcagcaa cgcgtctcca cgacactgtc gcaaaataac 2340
aacagcaact ttgcttggac cggtgccacc aagtatcatc tgaatggcag agactctctg 2400
gtaaatcccg gtgtcgctat ggcaacgcac aaggacgacg aagagcgatt ttttccatcc 2460
agcggagtct tgatgtttgg gaaacagggа gctggaaaag acaacgtgga ctatagcagc 2520
gttatgctaa ccagtgagga agaaatcaaa accaccaacc cagtggccac agaacagtac 2580
ggcgtggtgg ccgataacct gcaacagcaa aacgccgctc ctattgtagg ggccgtcaac 2640
agtcaaggag ccttacctgg catggtctgg cagaaccggg acgtgtacct gcagggtcct 2700
atctgggcca agattcctca cacggacggc aacttfcatc cttcgccgct gatgggaggc 2760
tttggactga aacacccgcc tcctcagatc ctgattaaga atacaccigt tcccgcggat 2820
cctccaacta ccttcagtca agccaagctg gcgtcgttca tcacgcagta cagcacccgga 2880
caggtcagcg tggaaattga atgggagctg cagaaagaga acagcaagcg ctggaacccа 2940
gagatticagt atacttccaa ctactacaaa tctacaaatg tggactttgc tgtcaatact 3000
gagggtactt attcagagcc tcgcccccatt ggcacccgtt acctcacccg taacctgtaa 3060
ttgcctgtta atcaataaac cggctaattc gtttcagttg aactttggtc tctgcgaagg 3120
gcgaattc 3128

<210> 28
<211> 3128
<212> DNA
<213> 新AAV血清型,克隆42.15

<400> 28
gaattcgccc tttctacggc tgcgtcaact ggaccaatga gaactttccc ttcaacgatt 60
gcgtcgacaa gatggtgatc tggtgggagg agggcaagat gacggccaag gtcgtggagt 120
ccgccaaggc catctcggcc ggcagcaagg tgcgcgtgga ccaaaagtgc aagtcgtccg 180
cccagatcga ccccacccca gtgatcgtca cctccaacac caacatgtgc gccgtgattg 240
acgggaacag caccaccttc gagcaccagc agccgttgca ggaccggatg ttcaaatttg 300
aactcacccg ccgtctggag catgactttg gcaaggtgac aaagcaggaa gtcaaagagt 360
tcttccgctg ggcgcaggat cacgtgaccg aggtggcgca tgagttctac gtcagaaagg 420
gtggagccaa caagagaccc gccccgatg acgcggataa aagcgagccc aagcgggcct 480

[0031]

gccectcagt cgcggatcca tcgacgtcag acgcggaagg agctccggtg gactttgccg 540
acaggtacca aaacaaatgt tctcgtcacg cgggcatgct tcagatgctg tttcectgca 600
agacatgcga gagaatgaat cagaatttca acatttgctt cacgcgcggg accagagact 660
gttcagaatg tttcccgggc gtgtcagaat ctcaaccggt cgtcagaaag aggacgtatc 720
ggaaactctg tgccattcat catctgctgg ggcgggctcc cgagattgct tgctcggcct 780
gcgatctggt caacgtggac ctggatgact gtgtttctga gcaataaatg aattaaacca 840
ggtatggctg ccgatggtta ttttccagat tggctcgagg acaacctctc tgagggcatt 900
cgcgagtggt gggacttgaa acctggagcc ccgaaaccca aagccaacca gcaaaagcag 960
gacgacggcc ggggtctggt gattcctggc tacaagtacc tcggaccctt caacggactc 1020
gacaaggggg agcccgtcaa cgcggcggac gcagcggccc tcgagcacga caaggcctac 1080
gaccagcagc tcaaagcggg tgacaatccg tacctgcggt ataaccacgc cgacgccgag 1140
tttcaggagc gtctgcaaga agatacgtct tttgggggca acctcgggcg agcagtcttc 1200
caggccaaga agcgggttct cgaacctctc ggtctggttg aggaaggcgc taagacggct 1260
cctggaaaga agagaccggt agagccatca ccccagcgtt ctccagactc ctctacgggc 1320
atcggcaaga caggccagca gcccgcgaaa aagagactca actttgggca gactggcgac 1380
tcagagtcag tgcccgaccc tcaaccaatc ggagaacccc ccgcaggccc ctctggtctg 1440
ggatctggta caatggctgc aggcggtggc gctccaatgg cagacaataa cgaaggcgcc 1500
gacggagtgg gtagttcctc aggaaattgg cattgcgatt ccacatggct gggcgacaga 1560
gtcatcacca ccagcacccg aacctgggcc ctccccacct acaacaacca cctctacaag 1620
caaatctcca acgggacatc gggaggaagc accaacgacg acacctactt cggctacagc 1680
accccctggg ggtattttga ctttaacaga ttccactgcc acttctcacc acgtgactgg 1740
cagcgactca tcaacaacaa ctggggattc cggccccaaga gactcaactt caagctcttc 1800
aacatccagg tcaaggaggt cacgcagaat gaaggcacca agaccatcgc caataacctt 1860
accagcacga ttcaggtctt tacggactcg gaataccagc tccctacgt cctcggctct 1920
gcgcaccagg gctgccccgcc tccgttcccg gcggacgtct tcatgattcc tcagtacggg 1980
tacctgactc tgaacaacgg cagtcaggcc gtgggccgtt cctccttcta ctgcctggag 2040
tactttcctt ctcaaatgcg gagaacgggc aacaacttta agttcagcta ccagtttgag 2100
gacgtgcctt ttcacagcag ctacgcgcat agccaaagcc tggaccggct gatgaaccct 2160
ctcatcgacc agtacctgta ctacctgtct cggactcagt ccacgggagg taccgcagga 2220
actcagcagt tgctatttct tcaggccggg cctaatggca tgtcggctca ggccaaaaac 2280
tggctacccg ggccctgcta ccggcagcaa cgcgtctcca cgacactgtc gcaaaataac 2340
aacagcaact ttgcttggac cggtgccacc aagtatcatc tgaatggcag agactctctg 2400
gtaaatcccg gtgtcgctat ggcaacgcac aaggacgacg aagagcgatt ttttccatcc 2460
agcggagtct tgatgtttgg gaaacagggg gctggaaaag acaacgtgga ctatagcagc 2520
gttatgctaa ccagtgagga agaaatcaaa accaccaacc cagtggccac agaacagtac 2580
ggcgtggtgg ccgataacct gcaacagcaa aacgccgctc ctattgtagg ggccgtcaac 2640
agtcaaggag ccttacctgg catggtctgg cagaaccggg acgtgtacct gcagggtcct 2700
atctgggcca agattcctca cacggacggc aactttcatc cttcgccgct gatgggaggc 2760
tttggactga aacaccccgcc tcctcagatc ctgattaaga atacaacctgt tcccgcggat 2820
cctccaacta ccttcagtca agccaagctg gcgtcgttca tcacgcagta cagcaccgga 2880
caggtcagcg tggaaattga atgggagctg cagaaagaga acagcaagcg ctggaaccca 2940
gagatccagt atacttccaa ctactacaaa tctacaaatg tggactttgc tgtcaatact 3000
gagggtactt attcagagcc tcgcccccatt ggcaccegtt acctcacccg taacctgtaa 3060
ttgcctgtta atcaataaac cggttaattc gtttcagttg aactttggtc tctgcgaagg 3120
gcgaattc 3128

<210> 29
<211> 3197
<212> DNA
<213> 新AAV血清型.克隆42.5b

<400> 29
gaattcgccc tttctacggc tgcgtcaact ggaccaatga gaactttccc ttcaacgatt 60
gcgtcgacaa gatggtgatc tggtgggagg agggcaagat gacggccaag gtcgtggagt 120
ccgccaaggc cattctcggc ggcagcaagg tcgcgtggga ccaaaagtgc aagtcgtccg 180
cccagatcga ccccacccccc gtgatcgtca ccttccaacac caacatgtgc gcgtgattg 240
acgggaacag caccaccttc gagcaccagc agccgttaca agaccggatg ttcaaatttg 300
aactcacccg ccgtctggag cacgacttig gcaaggtgac aaagcaggaa gtcaaagagt 360
tcttccgctg ggcgcaggat cacgtgaccg aggtggcgca tgagttctac gtcagaaagg 420
gtggagccaa caagagaccc gccccgatg acgcggataa aagcgagccc aagcgggcct 480
gcccctcagt cgcggatcca tcgacgtcag acgcggaagg agctccggtg gactttgccg 540
acaggtacca aaacaaatgt tctcgtcacg cgggcatgct tcagatgctg tttcectgca 600

[0032]

agacatgcga gagaatgaat cagaatttca acatttgctt cacgcacggg accagagact 660
gttcagaatg tttccccggc gtgtcagaat ctcaaccggt cgtcagaaag aggacgtatc 720
ggaaactctg tgccattcat catctgctgg ggcgggctcc cgagattgct tgctcggcct 780
gcgatctggt caacgtggac ctggatgact gtgtttctga gcaataaatg acttaaacca 840
ggtatggctg ccgatggtta tcttccagat tggctcgagg acaacctcta tgagggcatt 900
cgcgngtggt gggacttgaa acctggagcc ccgaaaccca aagccaacca gcaaaagcag 960
gacgacggcc ggggtctggt gattcctggc tacaagtacc tcggaccctt caacggactc 1020
gacaagggag agccggtcaa cgaggcagac gccgcggccc tcgagcacga caaggcctac 1080
gacaagcagc tcgagcaggg ggacaacccg tacctcaagt acaaccacgc cgacgccgag 1140
tttcaggagc gtcttcaaga agatacgtct tttgggggca acctcgggcg agcagtcttc 1200
caggccaaga agcgggttct cgaacctctc ggtctggttg aggaaggcgc taagacggct 1260
cctggaaaga agaggccggt agagccatca ccccagcgtt ctccagactc ctctacgggc 1320
atcggcaaga caggccagca gccgcgcaaa aagagactca actttgggca gactggcgac 1380
tcagagtcag tgccccgacc tcaaccaatc ggagaacccc ccgcaggccc ctctggtctg 1440
ggatctggta caatggctgc aggcggtggc gctccaatgg cagacaataa cgaaggcgcc 1500
gacggagtgg gtagttcctc aggaaattgg cattgcgatt ccacatggct gggcgacaga 1560
gtcatcacca ccagcacccg aacctgggcc ctccccacct acaacaacca cctctacaag 1620
caaatctcca acgggacatc gggaggaagc accaacgaca acacctactt cggctacagc 1680
acccctgggg ggtattttga ctttaacaga ttccactgcc acttctcacc acgtgactgg 1740
cagcgactca tcaacaacaa ctggggattc cggccccaaga gactcaactt caagctcttc 1800
aacatccagg tcaaggaggt cacgcagaat gaaggcacca agaccatcgc caataacctt 1860
accagcacga ttcaggtctt tacggacteg gaataccagc tcccgtacgt cctcggctct 1920
gcgcaccagg gctgcctgcc tcgttcccg gcggacgtct tcatgattcc tcagtacggg 1980
tacctgactc tgaacaacgg cagtcaggcc gtgggccgtt cctccttcta ctgcctggag 2040
tactttcctt ctcaaatgct gagaacgggc aacaactttg agttcagcta ccagtttgag 2100
gacgtgcctt ttcacagcag ctacgcgcac agccaaagcc tggaccggct gatgaaccct 2160
ctcatcgacc agtacctgta ctacctgtct cggactcagt ccacgggagg taccgcagga 2220
actcagcagt tgctatttte tcaggccggg cctaataaca tgtcggctca ggccaaaaac 2280
tggctaccgg ggccctgcta ccggcagcaa cgcgtctcca cgacactgtc gcaaaataac 2340
aacagcaact ttgcttggac cggtgccacc aagtatcatc tgaatggcag agactctctg 2400
gtaaatcccg gtgtcgctat ggcaacgcac aaggacgacg aagagcgatt ttttccatcc 2460
agcggagtct tgatgtttgg gaaacagggа gctggaaaag acaacgtgga ctatagcagc 2520
gttatgctaa ccagtgagga agaaatcaaa accaccaacc cagtggccac agaacagtac 2580
ggcgtggtgg ccgataacct gcaacagcaa aacgccgctc ctattgtagg ggccgtcaac 2640
agtcaaggag ccttacctgg catggtctgg cagaaccggg acgtgtacct gcagggtcct 2700
atctgggcca agattcctca cacggacggc aactttcatc cttcgccgct gatgggaggc 2760
tttggactga aacaccccgc tcctcagatc ctgattaaga atacacctgt tcccgcggat 2820
cctccaacta ccttcagtca agccaagctg gcgtcgttca tcacgcagta cagcacggga 2880
caggtcagcg tggaaattga atgggagctg cagaaagaga acagcaagcg ctggaacccа 2940
gagattcagt atacttccaa ctactacaaa tctacaaatg tggactttgc tgtcaatact 3000
gagggtactt attcagagcc tcgccccatt ggcaccegtt acctcacccg taacctgtaa 3060
ttgcctgtta atcaataaac cggttaattc gtttcagttg aactttggtc tctgcgaagg 3120
gcgaattcgt ttaaacctgc aggactagtc cctttagtga gggttaattc tgagcttggc 3180
gtaatcatgg gtcatag 3197

<210> 30
<211> 2501
<212> DNA
<213> 新 AAV 血清型,克隆 42.1b

<400> 30
gaattcgccc ttggctgcgt caactggacc aatgagaact ttcccttcaa cgattgcgtc 60
gacaagatgg tgatctggtg ggaggagggc aagatgacgg ccaaggtcgt ggagtccgcc 120
aaggccattc atcatctgct ggggcgggct cccgagattg cttgctcggc ctgcgatctg 180
gtcaacgtgg acctggatga ctgtgtttct gagcaataaa tgacttaaac caggtatggc 240
tgccgatggt tatcttccag attggctcga ggacaacctc tctgagggca ttcgcgagtg 300
gtgggactig agacctggag ccccgaaacc caaagccaac cagcaaaagc aggacgacgg 360
ccggggtctg gtgcttcctg gctacaagta cctcggaccc ttcaacggac tcgacaaggg 420
agagccggtc aacgaggcag acgccgcggc cctcgagcac gacaaggcct acgacaagca 480
gctcgagcag ggggacaacc cgtacctcaa gtacaaccac gccgacgccg agtttcagga 540
gcgtcttcaa gaagatacgt cttttggggg caacctcggg cgagcagtct tccaggccaa 600
gaagcgggtt ctcgaacctc tcggtctggt tgaggaaggc gctaagacgg ctcctggaaa 660

```
gaagagaccc atagaatccc ccgactcctc cacgggcatc ggcaagaaag gccagcagcc   720
cgctaaaaag agactcaact ttgggcagac tggcgactca gagtcagtgc ccgaccctca   780
accaatcgga gaaccccccg caggcccctc tggtctggga tctggcacaa tggctgcagg   840
cggtggcgct ccaatggcag acaataacga aggcgccgac ggagtgggta gttcctcagg   900
aaattggcat tgcgattcca catggctggg cgacagagtc atcaccaccc gcacccgaac   960
ctgggccctc ccccactaca acaaccacct ctacaagcaa atctccaacg ggacatcggg  1020
aggaagcacc aacgacaaca cctacttcgg ctacagcacc ccctgggggt attttgactt  1080
taacagattc cactgcccct tctcacccacg tgactggcag cgactcatca acaacaactg  1140
gggattccgg cccaagagac tcaacttcaa gctcttcaac atccaggtca aggaggtcac  1200
gcagaatgaa ggcaccaaga ccatcgccaa taaccttacc agcacgattc aggtctttac  1260
ggactcggaa taccagctcc cgtacgtcct cggctctgcg caccagggct gcctgcctcc  1320
gttcccggcg gacgtcttca tgattcctca gtacgggtac ctgactctga acaacggcag  1380
tcaggccgtg ggccgttcct ccttctactg cctggagtac tttccttctc aaatgctgag  1440
aacgggcaac aactttgagt tcagctacca gtttgaggac gtgccttttc acagcagcta  1500
tgcgcacagc caaagcctgg accggctgat gaaccccctc atcgaccagt acctgtacta  1560
cctgtctcgg actcagtcca cgggaggtac cgcaggaact cagcagttgc tattttctca  1620
ggccgggcct aataacatgt cggctcaggc caaaaactgg ctaccccggg cctgctaccg  1680
gcagcaacgc gtctccacga cagtgtcgca aaataacaac agcaactttg cttggaccgg  1740
tgccaccaag tatcatctga atggcagaga ctctctggta aatccccggg tcgctatggc  1800
aacgcacaag ggcgacgaag agcgatttit tccatccagc ggagtcttga tgtttgggaa  1860
acagggagct ggaaaagaca acgtagacta tagcagcgtt atgctaacca gtgaggaaga  1920
aatcaaaacc accaaccccg tggccacaga acagtacggc gtggtggccg ataacctgca  1980
acagcaaaac gccgctccta ttgtagggac cgtcaacagt caaggagcct tacctggcat  2040
ggtctggcag aaccgggacg tgtacctgca gggtcctatc tgggccaaga ttcctcacac  2100
ggacggcaac tttcatcctt cgccgctgat gggaggcttt ggactgaaac accccctcc  2160
tcagatcctg attaagaata cacctgttcc cgcggatcct ccaactacct tcagtcaagc  2220
caagctggcg tcgttcatca cgcagtacag cacccggacag gtcagcgtgg aaattgaatg  2280
ggagctgcag aaagagaaca gcaagcgctg gaacccagag attcagtata cttccaacta  2340
ctacaaatct acaaatgtgg acttigctgt caatactgag ggtactiatt cagagcctcg  2400
ccccattggc acccgttacc tcacccgtaa cctgtaattg ccigttaatc aataaaccgg  2460
ttgattcgtt tcagttgaac tttggtctca agggcgaatt c                      2501
```

[0033]

<210> 31
<211> 3113
<212> DNA
<213> 新 AAV 血清型, 克隆 42.13

<400> 31

```
gaattcgccc tttctacggc tgcgtcaact ggaccaatga gaactticcc ttcaacgatt     60
gcgtcgacaa gatggtgatc tggtgggagg agggcaagat gacggccaag gtcgtggagt    120
ccgccaaggc cattctcggc ggcagcaagg tgcgcgtgga ccaaaagtgc aagtcgtccg    180
cccagatcga tcccacccca gtgatcgtca cttccaacac caacatgtgc gccgtgattg    240
acgggaacag caccaccttc gagcaccagc agccgttaca agaccggatg ttcaaatttg    300
aactcacccg ccgtctggag catgactttg gcaaggtgac aaagcaggaa gtcaaagagt    360
tcttccgctg ggcgcaggat cacgtgaccg aggtggcgca tgagttctac gtcagaaagg    420
gtggagccaa caagagaccc gcccccgatg acgcggataa aagcgagccc aagcgggcct    480
gcccctcagt cgcggatcca tcgacgtcag acgcggaagg agctccggtg gacttgccg     540
acaggtacca aaacaaatgt tctcgtcacg cgggcatgct tcagatgctg tttccctgca    600
agacatgcga gagaatgaat cagaatttca acatttgctt cacgcacggg accagagact    660
gttcagaatg tttccccggc gtgtcagaat ctcaaccggt cgtcagaaag aggacgtatc    720
ggaaactctg tgccattcat catctgctgg ggcgggctcc cgagattgct tgctcggcct    780
gcgatctggt caacgtggac ctggatgact gtgtttctga gcaataaatg acttaaacca    840
ggtatggctg ccgatggtta tcttccagat tggctcgagg acaacctctc tgagggcatt    900
cgcgagtggt gggacttgaa acctggagcc ccgaaaccca aagccaacca gcaaaagcag    960
gacgacggcc ggggtctggt gcttcctggc tacaagtacc tcggacccit caacggactc   1020
gacaaggggg agcccgtcaa cgcggcggac gcagcggccc tcgagcacga caaggcctac   1080
gaccagcagc tcaaagcggg tgacaatccg tacctgcggt ataaccacgc cgacgcggag   1140
tttcaggagc gtcttcaaga agatacgtct tttgggggca acctcgggcg agcagtcttc   1200
caggccaaga agcgggttct cgaacctctc ggtctggttg aggaaggcgc taagacggct   1260
cctggaaaga agagacccgt agaatccccc gactcctcca cgggcatcgg caagaaaggc   1320
cagcagcccg ctaaaaagaa gctcaacttt gggcagactg gcgactcaga gtcagtgccc   1380
gaccctcaac caatcggaga acccccgcca ggcccctctg gtctgggatc tggtacaatg   1440
```

gctgcaggcg gtggcgctcc aatggcagac aataacgaag gcgccgacgg agtgggtagt 1500
tcctcaggaa attggcattg cgattccaca tggctgggcg acagagtcat caccaccagc 1560
acccgaacct gggccctccc cacctacaac aaccacctct acaagcaaat ctccaacggg 1620
acatcgggag gaagcaccaa cgacaacacc tacttcggct acagcacccc ctggggtat 1680
tttgacttta acagatccca ctgccacttc tcaccacgtg actggcagcg actcatcaac 1740
aacaactggg gattccggcc caagagactc aacttcaagc tcttcaacat ccaggtcaag 1800
gaggtcacgc agaatgaagg caccaagacc ategccaata accttaccag cacgatccag 1860
gtctttacgg actcggaata ccagctcccg tacgtcctcg gctctgcgca ccagggctgc 1920
ctgcctccgt tcccggcgga cgtcttcatg attcctcagt acgggtacct gactctgaac 1980
aacggcagtc aggccgtggg ccgttcctcc ttctactgcc tggagtactt tccttctcaa 2040
atgctgagaa cgggcaacaa ctttgagttc agctaccagt ttgaggacgt gcctttccac 2100
agcagctatg cgcacagcca aagcctggac cggctgatga acccctcat cgaccagtac 2160
ctgtactacc tgtctcggac tcagtccacg ggaggtaccg caggaactca gcagttgcta 2220
ttttctcagg ccgggcctaa tagcatgtcg gctcaggcca aaaactggct accgggccc 2280
tgctaccggc agcaacgcgt ctccacgaca gtgtcgcaaa ataacaacag caactttgct 2340
tggaccggtg ccaccaagta tcatctgaat ggcagagact ctctggtaaa tcccggtgtc 2400
gctatggcaa cgcacaaggg cgacgaagag cgatttttc catccagcgg agtcttgatg 2460
tttgggaaac agggagctgg aaaagacaac gtggactata gcagcgttat gctaaccagt 2520
gaggaagaaa tcaaaaccac caacccagtg gccacagaac agtacggcgt ggtggccgat 2580
aacctgcaac agcaaaacgc cgctcctatt gtaggggccg tcaacagtca aggagcctta 2640
cctggcatgg tctggcagaa ccgggacgtg tacctgcagg gtcctatctg ggccaagatt 2700
cctcacacgg acggcaactt tcatcctteg ccgctgatgg gaggctttgg actgaaacac 2760
cgccctcctc agatcctgat taagaataca cctgttcccg cggatcctac aactaccttc 2820
agtcaagcca agctggcgtc gttcatcacg cagtacagca ccggacaggt cagcgtggaa 2880
attgaatggg agctgcagaa agagaacagc aagcgctgga acccagagat tcagtatact 2940
tccaactact acaaatctac aaatgtggac tttgctgtca atactgaggg tacttattca 3000
gagcctcgcc ccattggcac ccgttacctc accccgtagcc tgtaattgcc tgttaatcaa 3060
taaaccggtt gattcgtttc agttgaactt tggtctctgc gaagggcgaa ttc 3113

[0034]

<210> 32
<211> 3113
<212> DNA
<213> 新AAV血清型,克隆42.3a

<400> 32
gaattcgccc tttctacggc tgcgtcaact ggaccaatga gaactttccc ttcaacgatt 60
gcgtcgacaa gatggtgatc tggtgggagg agggcaagat gacggccaag gtcgtggagt 120
ccgccaaggc cattctcggc ggcagcaagg tgcgcgtgga ccaaaagtgc aagtcgtccg 180
cccagatcga tcccacccct gtgatcgtca cttccaacac caacatgtgc gccgtgattg 240
acgggaacag caccaccttc gagcaccagc agccgttaca agaccggatg ttcaaatttg 300
aactcacccg ccgtctggag catgacttig gcaaggtgac aaagcaggaa gtcaaagagt 360
tcttccgctg ggcgcaggat cacgtgaccg aggtggcgca tgagttctac gtcagaaagg 420
gtggagccaa caagagaccc gcccccgatg acgcggataa aagcgagccc aagcgggcct 480
gcccctcagt cgcggatcca tcgacgtcag acgcggaagg agctccggtg gactttgccg 540
acaggtacca aaacaaatgt tctcgtcacg cgggcatgct tcagatgctg cttcctgca 600
agacatgcga gagaatgaat cagaattica gcatttgctt cacgcacggg accagagact 660
gttcagaatg tttccccggc gtgtcagaat ctcaaccggt cgtcagaaag aggacgtatc 720
ggaaactctg tgccattcat catctgctgg ggcgggctcc cgagattgct tgctcggcct 780
gcgatctggt caacgtggac ctggatgact gtgtttctga gcaataaatg acttaaacca 840
ggtatggctg ccgatggtca ttttccagat tggctcgagg acaacctctc tgagggcatt 900
cgcgagtggt gggacttgaa acctggagcc ccgaaacccа aagccaacca gcaaaagcag 960
gacgacggcc ggggtctggt gcttcctggc tacaagtacc tcggacccтt caacggactc 1020
gacaaggggg agccccgtcaa cgcggcggac gcagcggccc tcgagcacga caaggcctac 1080
gaccagcagc tcaaagcggg tgacaatccg tacctgcggt ataaccacgc cgacgccgag 1140
tttcaggagc gtcttcaaga agatacgtct tttgggggca acctcgggcg agcagtcttc 1200
caggccaaga agcgggttct cgaacctctc ggtctggttg aggaaggcgc taagacggct 1260
cctggaaaga agagaccccat agnatccccc gactcctcca cgggcatcgg caagaaaggc 1320
cagcagcccg ctaaaaagan gctcaacttt gggcagactg gcgactcaga gtcagtgccc 1380
gacccicaac caateggaga acccccgca ggcccctctg gtctgggatc tggtacaatg 1440
gctgcaggcg gtggcgctcc aatggcagac aataacgaag gcgccgacgg agtgggtagt 1500
tcctcaggaa attggcattg cgattccaca tagctgggcg acagagtcat caccaccagc 1560
acccgaacct gggccctccc cacctacaac aaccacctct acaagcaaat ctccaacggg 1620

acatcgggag gaagcaccaa cgacaacacc tacttcggct acagcacccc ctgggggtat 1680
tttgacttta acagatccca ctgccacttc tcaccacgtg actggcagcg actcatcaac 1740
aacagctggg gattccggcc caagagactc aacttcaagc tcttcaacat ccaggtcaag 1800
gaggtcacgc agaatgaagg caccaagacc atcgccaata accttaccag cacgattcag 1860
gtctttacgg actcggaata ccagctccrg tacgtcctcg gctctgrgca ccagggctgc 1920
ctgccteegt tcccggcgga cgtcttcatg attcctcagt acgggtacct gactctgaac 1980
aacggcagtc aggccgtggg ccgttcctcc ttctactgcc tggagtactt tccttctcaa 2040
atgctgagaa cgggcaacaa ctttgagttc agctaccagt ttgaggacgt gcctttcac 2100
agcagctacg cgcacagcca aagcctggac cggctgatga accccctcat cgaccagtac 2160
ctgtactacc tgtctcggac tcagtccacg ggaggtaccg caggaactca gcagttgcta 2220
ttttctcagg ccgggcctaa taacatgtcg gctcaggcca aaaactggct accccgggcc 2280
tgctaccggc agcaacgcgt ctccacgaca ctgtcgcaaa ataacaacag caacttlgct 2340
tggaccggtg ccaccaagta tcatctgaat ggcagagact ctctggtaaa tccccgtgtc 2400
gctatggcaa cgcacaagga cgacgaagag cgattttttc catccagcgg agtcttgatg 2460
tttgggaaac agggagctgg aaaagacaac gtggactata gcagcgttat gctaaccagt 2520
gaggaagaaa tcaaaaccac caaccccagtg gccacagaac agtacggcgt ggtggccgat 2580
aacctgcaac agcaaaacgc cgctcctatt gtaggggccg tcaacagtca aggagcctta 2640
cctggcatgg tctggcagaa ccgggacgtg tacctgcagg gtcctatctg ggccaagatt 2700
cctcacacgg acggcaactt tcatcctteg ccgctgatgg gaggctttgg actgaaacac 2760
ccgcctcctc agatcctgat taagaatacg cctgttcccg cggatcctcc aactaccttc 2820
agtcaagccn agctggcgtc gttcatcacg cagtacagca ccggacaggt cagcgtggaa 2880
attgaatggg agctgcagaa agnagaacagc angcgctgga acccagagat tcagtatact 2940
tccaactact acaaatctac aaatgtggac tttgctgtca atactgaggg tacttattca 3000
gagcctcgcc ccattggcac ccgttacctc acccgtaacc tgtaattgcc tgttaatcaa 3060
taaaccggtt aattcgtttc agttgaactt tggtctctgc gaagggcgaa ttc 3113

<210> 33
<211> 2504
<212> DNA
<213> 新AAV血清型,克隆42.4

[0035]

<400> 33
gaattcgccc tttctacggc tgcgtcaact ggaccaatga gaacttlccc ttcaacgatt 60
gcgtcgacaa gatggtgatc tggtgggagg agggcaagat gacggccaag gtcgtggagt 120
ccgccaaggc cattcatcat ctgctggggc gggctcccga gattgcttgc tcggcctgcg 180
atctggtcaa cgtggacctg gatgactgtg tttctgagca ataaatgact taaaccaggt 240
atggctgccg atggttatct tccagattgg ctcgaggaca acctctctga gggcattcgc 300
gagtggtggg acttgaaacc tggagccccg aaacccaaag ccaaccagca aaagcaggac 360
gacggccggg gtctggtgct tcctggctac aagtacctcg gacccttcaa cggactcgac 420
aagggagagc cggtcaacga ggcagacgcc gcggccctcg agcacgacaa ggcctacgac 480
aagcagctcg agcaggggga caaccccgtac ctcaagtacn accacgccga cgccgagttt 540
cgggagcgtc ttcaagaaga tacgtctttt gggggcaacc tcgggcgagc agtcttccag 600
gccaagaagc gggttctcga acctctcggt ctggttgagg aaggcgctaa gacggctcct 660
ggaaagaaga gacccataga atccccgac tcctccacgg gcatcggcaa gaaaggccag 720
cagcccgcta aaaagnagct caacttlggg cagactggcg actcagagtc agtgcccgac 780
cctcaaccaa tcggagaacc ccccgcaggc ccctctggtc tgggatctgg tacaatggct 840
gcaggcggtg gcgctccaat ggcagacaat aacgaaggcg ccgacggagt gggtaatgcc 900
tccggaaatt ggcattgcga ttccacatgg ctgggcgaca gagtcatcac caccagcacc 960
cgcacctggg ccctgcccac ctacaacaac cacctctaca agcagatatc aagtcagagc 1020
ggggctacca acgacaacca cttcttcggc tacagcaccc cctggggcta ttttgacttc 1080
aacagattcc actgccactt ctcatcacgt gactggcagc gactcatcaa caacaactgg 1140
ggattccggc ccaagagact caacttcaag ctcttcaaca tccaggtcaa ggaggtcacg 1200
cagaatgaag gcaccaagac categccaat aaccttacca gcacgattca ggtctttacg 1260
gactcggaat accggctccc gtacgtcctc ggctctgcgc accagggctg cctgcctccg 1320
ttcccggcgg acgtcttcat gattcctcag tacgggtacc tgactctgaa caacggcagt 1380
caggccgtgg gccgttcctc cttctactgc ctggagtact ttcttctca aatgctgaga 1440
acgggcaaca actttgagtt cagctaccag tttgaggacg tgcctttlca cagcagctac 1500
gcgcacagcc aaagcctgga ccggctgatg aacccctca tcgaccagta cctgtactac 1560
ctgtctcgga ctcagtccac gggaggtacc gcaggaactc agcagttgct attttctcag 1620
gccgggccta ataacatgtc ggctcaggcc aaaaactggc taccccgggc ctgctaccgg 1680
cagcaacgcg tctccacgac actgtcgcaa aataacaaca gcaactttgc ttggaccggt 1740
gccaccaagt atcatctgaa tggcagagac tctctggtaa atccccggtgt cgctatggca 1800

acgcacaagg acgacgaaga gcgatttttt ccatccagcg gagtcttgat gtttgggaaa 1860
cagggagctg gaaaagacaa cgtggactat agcagcgtta tgctaaccag tgaggaagaa 1920
atcaaaacca ccaacccagt ggccacagaa cagtacggcg tggtggccga taacctgcaa 1980
cagcaaaacg ccgctcctat tgtagggccc gtcaacagtc aaggagcctt acctggcatg 2040
gtctggcaga accgggacgt gtacctgcag ggtcctatct gggccaagat tcctcacacg 2100
gacggcaact ttcatccttc gccgctgatg ggaggctttg gactgaaaca ccccgcctcct 2160
cagatcctga ttaagaatac acctgttccc gcggatcctc caactacctt cagtcaagcc 2220
aagccggcgt cgttcatcac gcagtacagc accggacagg tcagcgtgga aattgaatgg 2280
gagctgcaga aagagaacag caagcgctgg aacccagaga ttcagtatac ttccaactac 2340
tacaaatcta caaatgtgga ctttgctgtc aatactgagg gtacttattc agagcctcgc 2400
cccattggca ccccgttacct cacccgtaac ctgtaattgc ctgttaatca ataaacccgt 2460
taatttgttt cagttgaact ttggtctctg cgaagggcga attc 2504

<210> 34
<211> 3106
<212> DNA
<213> 新 AAV 血清型，克隆 42.5a

<400> 34

[0036]

gaattcgccc ttctacggct gcgtcaactg gaccaatgag aactttcctt tcaacgattg 60
cgtcgacaag atggtgatct ggtgggagga gggcaagatg acggccaagg tcgtggagtc 120
cgccaaggcc attctcggcg gcagcaaggt gcgcgtggac caaaagtgca agtcgtccgc 180
ccagatcgac cccacccccg tgatcgtcac ctccaacacc aacatgtgcg ccgtgattga 240
cgggaacagc accaccttcg agcaccagca gccgttgcag gaccggatgt tcaaatttga 300
actcacccgc cgtctggagc atgactttgg caaggcgaca aagcaggaag tcaaagagtt 360
cttccgctgg gcgcaggatc acgtgaccga ggtggcgcat gagttctacg tcagaaaggg 420
tggagccaac aagagacccg ccccccgatga cgcggataaa agcgagccca agcgggcccg 480
ccctcagtc gcggatccat cgacgtcaga cgccggaagga gctccggtgg actttgccga 540
caggtaccaa aacaaatgtt ctcgtcacgc gggcatgctt cagatgctgt ttcccigcaa 600
aacatgcgag agaatgaatc agaatttcaa cattgcttc acgcacggga ccagagactg 660
ttcagaatgt ttccccggcg tgtcagaatc tcaaccggtc gtcagaaaga ggacgtatcg 720
gaaactctgt gccatttatc atctgctggg gcgggctccc gagattgctt gctcggcctg 780
cgatctggtc aacgtggacc tggatgactg tgtttctgag caataaatga cttaaaccag 840
gtatggctgc cgatggttat cttccagatt ggctcgagga caacctctct gagggcattc 900
gcgagtggtg ggacttgaaa cctggagccc cgaaacccaa agccaaccag caaaagcagg 960
acgacggccg gggtctggtg cttcctggct acaagtacct cggacccttc aacggactcg 1020
acaagggaga gccggtcaac gaggcagacg ccgcggccct cgagcacgac aaggcctacg 1080
acaagcagct cgagcagggg gacaacccgt acctcaagta caaccacgcc gacgccgagt 1140
ttcaggagcg tcttcaagaa gatacgtctt ttgggggcaa cctcgggcga gcagtcttcc 1200
gggccaagaa gcgggttctc gaacctctcg gtctggttga ggaaggcgct aagacggctc 1260
ctggaaagaa gagacccata gaatcccccg actcctccac gggcatcggc aagaaaggcc 1320
agcagccccgc taaaaagaag ctcaacttig ggcagactgg cgactcagag tcagtgcccg 1380
acccccaacc tctcggagaa cctcccgccg cgcccctcagg tctgggatct ggtacaatgg 1440
ctgcaggcgg tggcgcacca atggcagaca ataacgaagg cgccgacgga gtgggtaatg 1500
cctccggaaa tiggcattgc gattccacat ggctgggcga cagagtcatc accaccagca 1560
cccgcaccig ggccctgccc acctacaaca accacctcta caagcagata tcaagtcaga 1620
acgcgggctac caacgacaac cacttcttcg gctacagcac ccccigggcc tattigact 1680
tcaacagatt ccactgccac ttctcaccac gtgactggca gcgactcatc aacaacaacc 1740
ggggattccg gcccagaaag ctgcggttca agttgttcaa catccaggtc aaggaggtca 1800
cgacgaacga cggcgttacg accatcgcta ataaccttac cagcacgatt caggtcttct 1860
cggactcgga gtaccaactg ccgtacgtcc tcggctctgc gcaccagggc tgcctcccic 1920
cgttccctgc ggacgtgttc atgattcctc agtacggata tctgactcta aacaacggca 1980
gtcagtctgt gggacgttcc tccttctact gcctggagta ctttcttct cagatgctga 2040
gaacgggcaa taactttgaa ttcagctacc agtttgagga cgtgccctit cacagcagct 2100
acgcgcacag ccaaagcctg gaccggctga tgaaccccct catcgaccag tacctgtact 2160
acctgtctcg gactcagtcc acgggaggta ccgcaggaac tcagcagttg ctatttttct 2220
aggcgcgggcc taataacatg tcggctcagg ccaaaaactg gctacccggg cctgctacc 2280
ggcagcaacg cgtctccacg acactgtcgc aaaatancaa cagcaactit gcttggaccg 2340
gtgccaccaa gtatcatctg aatggcagag actctctggt aaatccccgt gtcgctatgg 2400
caacgcacaa ggacgacgaa gagcgatttt ttccatccag cggagtcttg atgtttggga 2460
aacagggagc tggaaaagac aacgtggact atagcagcgt tatgctaacc agtgaggaag 2520
aaatcaaaac caccaaccca gtggccacag aacagtacgg cgtggtggcc gataacctgc 2580

aacagcaaaa cgccgctcct attgtagggg ccgtcaacag tcaaggagcc ttacctggca 2640
tggcctggca gaaccgggac gtgtacctgc agggtcctat ctgggccaag attcctcaca 2700
cggacggcaa ctttcatcct tcgccgctga tgggaggctt tggactgaaa caccccgcct 2760
ctcagatcct gattaagaat acacctgttc ccgcggatcc tccaactacc ttcagtcaag 2820
ccaagctggc gtcgttcatc acgcagtaca gcaccggaca ggtcagcgtg gaaattgaat 2880
gggagctgca gaaagagaac agcaagcgct ggaacccaga gattcagtat acttccaact 2940
actacaaatc tacaaatgtg gactttgctg tcaatactga gggtacttat tcagagcctc 3000
gccccattgg cacccgttac ctcaccccta acctgtaatt gcctgttaat caataaaccg 3060
gttaattcgt ttcagttgaa ctttggtctc tgcgaagggc gaattc 3106

<210> 35
<211> 2489
<212> DNA
<213> 新 AAV 血清型, 克隆 42.10

<400> 35
gaattcgccc tttctacggc tgcgtcaact ggaccaatga gaactttccc ttcaacgatt 60
gcgtcgacaa gatggtgatc tggtgggagg agggcaagat gacggccaag gtcgtgaagt 120
ccgccaaggc cattcatcat ctgctggggc gggctccgga gattgcttgc tcggcctgcg 180
atctggtcaa cgtggacctg gatgactgtg tttctgagca ataaatgact taaaccaggt 240
atggctgccg atggttatct tccagattgg ctcgaggaca acctctctga gggcattcgc 300
gagtggtggg acttgaaacc tggagccccg aaacccaaag ccaaccagca aaagcaggac 360
gacggccggg gtctggtgct tcctggctac aagtacctcg gacccttcaa cggactcgac 420
aagggagagc cggtcaacga ggcagacgcc gcggccctcg agcacgacaa ggcctacgac 480
agcagctcg agcagggggga caacccgtac ctcaagtaca accacgccga cgccgagttt 540
caggagcgtc ttcaagaaga tacgtctttt gggggcaacc tcgggcgagc agtcttccag 600
gccaagaagc gggttctcga acctctcggt ctggttgagg aaggcgctaa gacggctcct 660
ggaaagaaga gaccccatag atccccccgc tcctccacgg gcatcggcag gaaaggccag 720
cagcccgcta aaaagaagct caactttggg cagactggcg actcagagtc agtgcccgac 780
cctcaaccaa tcggagaacc ccccgcaggc ccctctggtc tgggatctgg tacaatggct 840
gcaggcggtg gcgctccaat ggcagacaat aacgaaggcg ccgacggagt gggtaatgcc 900
tccggaaatt ggcattgcga ttccacatgg ctgggcgaca gagtcatcac caccagcacc 960
cgcacctggg ccctgcccac ctacaacaac cacctctaca agcagatatc aagtcagagc 1020
ggggctacca acgacaacca cttcttcggc tacagcaccc cctggggcta ttttgacttc 1080
aacagattcc actgccactt ctcaccacgt gactggcagc gactcatcaa caacaactgg 1140
ggattcrggc ccagaaagct gcggttcaag ttgttcaaca tccaggtcaa ggaggtcacg 1200
acgaacgacg gcgttacgac catcgccaat aaccttacca gcacgattca ggtcttctcg 1260
gactcggagt accaactgcc gtacgtcctc ggctctgcgc accagggctg cctccctccg 1320
ttccctgcgg acgtgttcat gattcctcag tacggatatc tgactctaaa caacggcagt 1380
cagtctgtgg gacgttcctc cttctactgc ctggagtact ttccttctca gatgctgaga 1440
acgggcaata actttgaatt cagctacacc tttgaggaag tgcctttcca cagcagctat 1500
gcgcacagcc agagcctgga ccggctgatg aatccgctca tcgaccagta cctgtactac 1560
ctggcccgga cccagagcac tacggggtcc acaagggagc tgcagttcca tcaggctggg 1620
cccaacacca tggccgagca atcaaagaac tggctgcccg gaccctgtta tcggcagcag 1680
agactgtcaa aaaacataga cagcaacaac aacagtaact ttgcctggac cggggccact 1740
aaataccatc tgaatggtag aaattcatta accaacccgg gcgtagccat ggccaccaac 1800
aaggacgacg aggaccagtt ctttcccatc aacggagtgc tggttttttgg caaaacgggg 1860
gctgccaaca agacaacgct ggaaaacgtg ctaatgttca gcgaggagga gatcaaaacc 1920
accaatcccg tggctacaga agaatacggt gtggtctcca gcaacctgca atcgtctacg 1980
gccggacccc agacacagac tgtcaacagc caggggggctc tgcccggcat ggtctggcag 2040
aaccgggacg tgtacctgca gggtcccatc tgggccaaaa ttcctcacac ggacggcaac 2100
tttcacccgt ctcccctgat gggcggattt ggactcaaac accccctcc tcaaattctc 2160
atcaaaaaca ccccggtacc tgctaatcct ccagaggtgt ttactcctgc caagttgacc 2220
tcatttatca cgcagtacag caccggccag gtcagcgtgg agatcgagtg ggaactgcag 2280
aaagaaaaca gcaaacgctg gaatccagag attcagtaca cctcaaatta tgccaagtct 2340
aataatgtgg aattlgcigt caacaacgaa ggggtttata clgagcctcg ccccattggc 2400
accgttacc tcaccctaa cctgtaattg cctgttaatc aataaaccgg ttaattcgtt 2460
tcagttgaac tttggtcaag ggcgaattc 2489

<210> 36
<211> 2495
<212> DNA

[0037]

<213> 新 AAV 血清型, 克隆 42.3b

<400> 36

gaattcgccc tttctacggc tgcgtcaact agaccaatga gaactttccc ttcaacgatt 60
gcgtcgacaa gatggtgatc tggtgggagg agggcaagat gacggccaag gtcgtggagt 120
ccgccaaggc cattcatcat ctgctggggc gggctcccga gattgcttgc tcggcctgcg 180
atctggtcaa cgtggacctg gatgactgtg tttctgagca ataaatgact taaaccaggt 240
atggctgccg atggttatct tccagattgg ctcgaggaca acctctctga gggcattcgc 300
gagtggtggg acttgaaacc tggagccccg aaacccaaag ccaaccagca aaagcaggac 360
gacggccggg gtctggtgct tcctggctac aagtacctcg gacccttcaa cggactcgac 420
aagggagagc cggtcaacga ggcagacgcc gcggccctcg agcacgacaa ggcctacgac 480
aagcagctcg agcagggggg caacccgtac ctcaagtaca accacgccga cgccgagttt 540
caggagcgtc ttcaagaaga tacgtctttt gggggcaacc tcgggcgagc agtcttccag 600
gccaagaagc gggttctcga acctctcggt ctggttgagg aaggcgctaa gacggctcct 660
ggaaagaaga gaccccatgg atcccccgac tcctccacgg gcatcggcaa gaaaggccag 720
cagcccgcta aaaagaagct caactttggg cagactggcg actcagagtc agtgccccac 780
cctcaaccaa tcggagaacc cccgcaggc cctctggtc tgggatctgg tacaatggct 840
gcaggcggtg gcgctccaat ggcagacaat aacgaaggcg ccgacggagt gggtaatgcc 900
tccggaaatt ggcattgcga ttccacatgg ctgggcgaca gagtcatcac caccagcacc 960
cgcacctggg ccctgcccac ctacaacaac cacctctaca agcagatatc aagtcagagc 1020
ggggctacca acgacaacca cttcttcggc tacagcaccc cctggggcta ttttgacttc 1080
aacagattcc actgccactt ctccaccacgt gactggcagc gactcatcaa caacaactgg 1140
ggattccggc ccagaaagct gcggttcaag ttgttcaaca tccaggtcaa ggaggtcacg 1200
acgaacgacg gcgttacgac catcgctaat aaccttacca gcacgattca ggtcttctcg 1260
gactcggagt accagctgcc gtacgtcctc ggctctgcgc accagggctg cctccctccg 1320
ttccctgcgg acgtgttcat gattcctcag tacggatatc tgactctaaa caacggcagt 1380
cagtctgtgg gacgttcctc cttctactgc ctggagtact ttccttctca gatgctgaga 1440
acgggcaata actttgaatt cagctacacc tttgaggaag tgccttttca cagcagctat 1500
gcgcacagcc agagcctgga ccggctgatg aatccgctca tcgaccagta cctgtactac 1560
ctggcccgga cccagagcac tacgggqtcc acaagggagc tgcagttcca tcaggctggg 1620
cccaacacca tggcgcgagca atcaaagaac tggctgcccg gaccctgtta tcggcagcag 1680
agactgtcaa aaaacataga cagcaacaac accagtaact ttgcctggac cgggqccact 1740
aaataccatc tgaatggtag aaattcatta accaacccgg gcgtagccat ggccaccaac 1800
aaggacgacg aggaccagtt ctttccgatc aacggagtgc tggttttgg caaaacgggg 1860
gctgccaaca agacaacgct ggaaaacgtg ctaatgacca gcgaggagga gatcaaaacc 1920
accaatccgg tggctacaga acagtacggt gtggtctcca gcaacctgca ategtctacg 1980
gccggacccc agacacagac tgtcaacagc caggggqctc tgcccggcat ggtctggcag 2040
aaccgggacg tgtacctgca gggtccgatc tgggccaaaa ttcctcacac ggacggcaac 2100
tttcaccgt ctcccctgat gggcggattt ggactcaaac accgcctcc tcaaattctc 2160
atcaaaaaca ccccggtacc tgctaatcct ccagaggtgt ttactcctgc caagtttgcc 2220
tcatttatca cgcagtacag caccggccag gtcagcgtgg agatcgagtg ggaactgcag 2280
aaagaaaaca gcaaacgctg gaatccagag atccagtaca cctcaaatta tgccaagtct 2340
aataatgtgg aatttgctgt caacaacgaa ggggtttata ctgagcctcg ccccattggc 2400
accegttacc tcaccegtaa cctgtaattg cctgttaatc aataaaccgg ttaattcgtt 2460
tcagttgaac tttggtctct gcgaagggcg aattc 2495

<210> 37
<211> 3098
<212> DNA
<213> 新 AAV 血清型, 克隆 42.11

<400> 37

gaattcgccc tttctacggc tgcgtcaact ggaccaatga gaactttccc ttcaacgatt 60
gcgtcgacaa gatggtgatc tggtgggagg agggcaagat gacggccaag gtcgtggagt 120
ccgccaaggc cattctcggc ggcagcaagg tgcgcgtgga ccaaaagtgc aagtcttccg 180
cccagatcga tcccacccct gtgatcgtca cttccaacac caacatgtgc gccgtgattg 240
acgggaacag caccaccttc gagcaccagc agccgttaca agaccggatg ttcaaatttg 300
aactcacccg ccgtctggag cacgactttg gcaaggtgac aaagcaggaa gtcaaagagt 360
tcttccgctg ggcgcaggat cacgtgaccg aggtggcgca tgagttctac gtcagaaagg 420
gtggagccaa caagagaccc gcccccgatg acgcggataa aagcgagccc aagcgggcct 480
gcccctcagt cgcggatcca tcgacgtcag acgcggaagg agctccggtg gactttgccg 540
acaggtacca aaacaaatgt tctcgtcacg cgggcatgct tcagatgctg tttccctgca 600

[0039]

agacatgcga gagaatgaat cagaatttca acatttgctt cacgcacggg acccggagact 660
gttcagaatg tttccccggc gtgtcagaat ctcaaccggt cgtcagaaag aggacgtatc 720
ggaaactctg tgccattcat catctgctgg ggcgggctcc cgagattgct tgctcggcct 780
gcgatctggt caacgtggac ctggatgact gtgtttctga gcaataaatg acttaaacca 840
ggtatggctg ccgatggtta tcttccagat tggctcgagg acaacctctc tgagggcatt 900
cgcgngtggt gggacttgaa acctggagcc ccgaaaccca aagccaacca gcaaaagcag 960
gacgacggcc ggggtctggt gcttcctggc tacaagtacc tcggacccct caacggactc 1020
gacaagggag agccggtcaa cgcggcggac gcagcggccc tcgagcacga caaggcctac 1080
gaccagcagc tcaaagcggg tgacaatccg tacctgcggt ataaccacgc cgacgccgag 1140
tttcaggagc gtcttcaaga agatacgtct tttgggggca acctcgggcg agcagtcttc 1200
caggccaaga agcgggttct cgaacctctc ggtctggttg aggaaggcgc taagacggct 1260
cctggaaaga agagacccat agaatccccc gactcctcca cgggcatcgg caagaaaggc 1320
cagcagcccg ctaaaaagaa gctcaacttt gggcagactg gcgactcaga gtcagtgccc 1380
gaccctcaac caatcggaga accccccgca ggcccctctg gtctgggatc tggtacaatg 1440
gctgcaggcg gtggcgctcc aatggcagac aataacgaag gcgccgacgg agtgggtaat 1500
gcctccggaa attggcattg cgattccaca tggctgggcg acagagtcat caccaccagc 1560
accccgcacct gggccctgcc cacctacaac aaccacctct acaagcagat atcaagtcag 1620
agcggggcta ccaacgacaa ccacttcttc ggctacagca ccccctgggg ctattttgac 1680
ttcaacagat tccactgcca cttctcacca cgtgactggc agcgacteat caacaacaac 1740
tggggattcc ggcccagaaa gctgcggttc aagttgttca acatccaggt caaggaggtc 1800
acgacgaacg acggcgttac gaccatcgct aatagcctta ccagcacgat tcaggtcttc 1860
tcggactcgg agtaccaact gccgtacgtc ctcggctctg cgcaccaggg ctgcctccct 1920
ccgttccctg cggacgtgtt catgattcct cagtacggat atctgactct aaacaacggc 1980
agtcagtctg tgggacgttc ctccttctac tgcctggagt actttcttc tcagatgctg 2040
agaacgggca ataactttga attcagctac acctttgagg aagtgccttt ccacagcagc 2100
tatgcgcaca gccagagcct ggaccggctg atgaatcccc tcatcgacca gtacctgtac 2160
tacctggccc ggacccagag cactacgggg tccacaaggg agctgcagtt ccatcaggct 2220
gggccccaaca ccatggccga gcaatcaaag aactggctgc ccggacccctg ttatcggcgg 2280
cagagactgt caaaagacat agacagcaac aacaacagta actttgcctg gaccggggcc 2340
actaaatacc atctgaatgg tagaaatttca ttaaccaacc cgggcgtagc catggccacc 2400
aacaaggacg acgaggacca gttctttccc atcaacggag tgctggtttt tggcaaaacg 2460
ggggctgcca acaagacaac gctggaaaac gtgctaatga ccagcgagga ggagatcaaa 2520
accaccaatc ccgtggctac agaagaatac ggtgtggtct ccagcaacct gcaatcgtct 2580
acggccggac cccagacaca gactgtcaac agccagggcg ctctgcccgg catggtctgg 2640
cagaaccggg acgtgtacct gcagggtccc atctgggcca aaattcctca cacggacggc 2700
aactttcacc cgtctcccct gatgggcgga tttggactca aacaccccgc tcctcaaatt 2760
ctcatcaaaa acaccccggt acctgctaat cctccagagg tgtttactcc tgccaagttt 2820
gcctcattta tcacgcagta cagcaccggc caggtcagcg tggagatcga gtgggaactg 2880
cagaaagaga acagcaaacg ctggaatcca gagatttcagt acacctcaaa ttatgccaag 2940
tctaataatg tggaatttgc tgtcaacaac gaagggtttt atactgagcc tcgccccatt 3000
ggcaccegtt acctcacccg taacctgtaa ttacttgtta atcaataaac cggttgattc 3060
gttcagttg aactttggtc tctgcgaagg gcgaattc 3098

<210> 38
<211> 3276
<212> DNA
<213> 新 AAV 血清型，克隆 42.6a

<400> 38
gaattcgccc ttcgcagaga ccaaagttca actgaaacga attaaccggt ttattgatta 60
acaggcaatt acaggttacg ggtgaggtaa cgggtgccaa tgggggcgagg ctcagtataa 120
acccccttcgt tgttgacagc aaattccaca ttattagact tggcataatt tgaggtgtac 180
tgaatctctg gattccagcg tttgctgttt tctttctgca gttcccactc gatctccacg 240
ctgacctggc cggtgctgta ctgcgtgata aatgaggcaa acttggcagg agtaaacacc 300
tctggaggat tagcaggtac cggggtgttt ttgatgagaa tttgaggagg cgggtgtttg 360
agtccaaatc cgtcccatcag gggagacggg tgaaagttgc cgtccgtgtg aggaattttg 420
gcccagatgg gacccigcag gtacacgtcc cggttctgcc agaccatgcc gggcagagcc 480
cctggctgt tgacagtctg tgtctggggt ccggccgtag acgattgcag gttgctggag 540
accacaccgt attcttctgt agcccacggga ttggtggttt tgatctcctc ctcgctggtc 600
attagcacgt tttccagcgt tgtcttgttg gcagcccccg ttttgccaaa aaccagcact 660
ccgttgatgg gaaagaactg gtcctcgtcg tccttgttgg tggccatggc tacgcccggg 720
ttggttaatg aatttctacc attcagatgg tatttagtgg ccccggtcca ggcaaagtta 780

[0040]

ctgttgttgt tgctgtctat gtttttttgac agtctctgct gccgataaca gggtccgggc 840
agccagttct ttgattgctc ggccatggtg ttgggcccag cctgatggaa ctgcagctcc 900
cttgtggacc ccgtagtgct ctgggtccgg gccaggtagt acaggtactg gtcgatgagg 960
ggattcatca gccggtccag gctctggcta tgcgcatagc tgctgtggaa aggcacttcc 1020
tcaaaggtgt agctgaattc aaagttattg cccgttctca gcatctgaga aggaaagtac 1080
tccaggcagt agaaggagga acgtcccaca gactgactgc cgttgtttag agtcagatat 1140
ccgtactgag gaatcatgaa cacgtccgca gggaacggag ggaggcagcc ctggtgcgca 1200
gagccgagga cgtacggcag ttggtactcc gagtccgaga agacctgaat cgtgctggta 1260
aggttattag cgatggtcgt aacgccgtcg tccgtcgtga cctccttgac ctggatgttg 1320
aacaacttga accgcagctt tctgggccgg aatccccagt tgttgttgat gagtcgctgc 1380
cagtcacgtg gtgagaagtg gcagtggaat ctgttaaagt caaaataccc ccaggggtg 1440
ctgtagccga agtaggtgtt gtcgttggtg ctttcttccg atgtcccgtt ggagatttgc 1500
ttgtagaggt ggttgttgta ggtggggagg gcccaggttc gggtgctggt ggtgatgact 1560
ctgtcgccca gccatgtgga atcgcaatgc caatttcctg aggaactacc cactccgtcg 1620
gcgccttcgt tattgtctgc cattggagcg ccaccgcctg cagccattgt accagatccc 1680
agaccagagg ggcctgcggg gggttcttcg attggttgag ggtcgggcac tgactctgag 1740
tcgccagtct gcccaaagtt gagtctcttt ttcgcgggct gctggcctgt cttgccgatg 1800
cccgtagagg agtctggaga acgctggggt gatggctcta ccggtctctt ctttccagga 1860
gccgttctag cgccttcctc aaccagaccg agaggttcga gaacccgctt cttggcctgg 1920
aagactgctc gcccgaggtt gccccccaaaa gacgtatctt cttgaagacg ctcctgaaac 1980
tcggcgtcgg cgtggttgta cttgaggtac gggttgtccc cctgctcgag ctgcttgtcg 2040
taggccttgt cgtgctcgag ggcgcgggcg tctgccttgt tgaccggctc tccttgtcg 2100
agtccgttga agggtccgag gtacttgtag ccaggaagca ccagacccccg gccgtcgtcc 2160
tgcttttgct ggttggcttt gggtttcggg gctccaggtt tcaagtccca ccactcgcga 2220
atgccctcag agaggttgtc ctcgagccaa tctggaagat aaccatcggc agccatacct 2280
ggtttaagtc atttattgct cagaaacaca gtcatccagg tccacgttga ccagatcgca 2340
ggccgagcaa gcaatctcgg gagccccgccc cagcagatga tgaatggcac agagtttccg 2400
atacgtcctc tttctgacga ccggttgaga ttctgacacg ccggggaaac attctgaaca 2460
gtctctggtc ccgtgcgtga agcaaatgtt gaaattctga ttcattctct cgcatgtctt 2520
gcagggaaac agcatctgaa gcatgcccgc gtgacgagaa cacttgtttt ggtacctgtc 2580
ggcaaagtcc accggagctc cttccgcgtc tgacgtcgat ggatgcaaaa tgtcgcaaaa 2640
gcactcacgt gacagctaat acaggaccac tcccctatga cgtgatttac gtcagcgcta 2700
tacccgcgtg acgagaacat ttgttttggt acctgtcggc aaagtccacc ggagctcctt 2760
ccgcgtctga cgtcgatgga tccgcgactg aggggcaggc ccgcttgggc tcgcttttat 2820
ccgcgtcatc ggggggcggt ctcttgttgg ctccacccct tctgacgtag aactcatgcg 2880
ccacctcggt cacgtgatcc tgcgcccagc ggaagaactc tttgacttcc tgctttgtca 2940
ccttgccaaa gtcatgctcc agacggcggg tgagttcaaa tttgaacatc cggtcctgca 3000
acggctgctg gtgctcgaag gtggtgctgt tcccgtcaat cacggcgcac atgttggtgt 3060
tggaagtgac gatcacgggg gtgggatcga tctgggcgga agacttgcac ttttggtcca 3120
cgcgcacctt gctgccgccg agaatggcct tggcggactc cacgaccttg gccgtcatct 3180
tgccctcctc ccaccagatc accatcttgt cgacgcaatc gttgaaggga aagttctcat 3240
tggtccagtt gacgcagccg tagaaagggc gaattc 3276

<210> 39
<211> 3084
<212> DNA
<213> 43.1

<400> 39
gaattcgccc tttctacggc tgcatcaact ggaccaatga gaactttccc ttcaacgatt 60
gcgtcgacaa gatggtgatc tggtgggagg agggcaagat gacggccaag gtcgtggagt 120
ccgccaaggc cattctggcc ggcagcaagg tgcgcgtgga ccaaaagtgc aagtcgtccg 180
cccagatcga ccccacccccc gtgatcgtca cctccaacac caacatgtgc gccgtgattg 240
acgggaacag caccaccttc gagcaccagc agccgttgca ggaccggatg ttcaagttcg 300
aactcacccg ccgtctggag cacgactttg gcaaggtgac caagcaggaa gtcaaagagt 360
tcttccgctg ggcgcaggat cacgtgaccg aggtggcgca tgagttctac gtcagaaagg 420
gcggagccag caaaagaccc gccccgatg acgcggatat aagcgagccc aagcgggcct 480
gccccgagt cgcggatcca tcgacgtcag acgcggaagg agctccggtg gacttgccg 540
acaggtacca aaacaaatgt tctcgtcacg cgggcatgct tcagatgctg tttccctgca 600
aaacgtgcga gaaaatgaat cagaatttca acatttgctt cacgcacggg gtcagagact 660
gctcagaatg tttccccggt gcatcagaat ctcaaccggt cgtcagaaaa aaaacgtatc 720
agaaaactgtg tgccattcat catctgctgg ggcgggcacc cgagattgct tgctcggcct 780

[0041]

gcgatctggt caacgtggac ctggacgact gtgtttctga gcaataaatg acttaaacca 840
ggtatggctg ccgatggtta tcttccagat tggcttgagg acaacctctc tgagggcatt 900
cgcgagtggt gggacctgaa acctggagcc ccgaaaccca aagccaacca gcaaaagcag 960
gacgacggcc ggggtctggt gcttcctggc tacaagtacc tcggacccct caacggactc 1020
gacaaggggg agcccgtcaa cgcggcggac gcagcggccc tcgagcacga caaggcctac 1080
gaccagcagc tcaaagcggg tgacaatccg tacctgcggt ataaccacgc cgacgccgag 1140
ttccaggagc gtctgcaaga agatacgtct tttgggggca acctcgggcg agcagtcttc 1200
caggccaaga agcgggttct cgaacctctc ggtctggttg aggaaggcgc taagacggct 1260
cctggaaaga agagaccggt agagccatca cctcagcgtt ccccgactc ctccacgggc 1320
atcggcaaga aaggccacca gccgcgaga aagagactga actttgggca gactggcgac 1380
tcggagtcag tccccgaccc tcaaccaatc ggagaaccac cagcaggccc ctctggtctg 1440
ggatctggta caatggcigc aggcggtggc gctccaatgg cagacaataa cgaaggcgcc 1500
gacggagtgg gtagttcctc aggaaattgg cattgcgatt ccacatggct gggcgacaga 1560
gtcatcacca ccagcacccg aacctgggcc ctgcccacct acaacaacca tctctacaag 1620
caaatctcca acgggacatc gggaggaagc actaacgaca acacctactt tggctacagc 1680
acccccctgg ggtattttga cttcaacaga ttccactgcc acttctcacc acgtgactgg 1740
cagcgactca tcaacaataa ctggggattc cggcccaaga gactcaactt caagctcttc 1800
aacatccagg tcaaggaggt cacgcagaat gaaggcacca agaccatcgc caataacctt 1860
accagcacga ttcaggtgtt tacggactcg gaataccagc tcccgtacgt cccccggctct 1920
gcgcaccagg gctgcctccc tccgttcccg gcggacgtct tcatgattcc tcagtacggg 1980
tatctgaccc taaacaatgg cagtcaggct gtgggccgtt cctccttcta ctgcctggaa 2040
tacttccctt ctcaaatgct gaggacgggc aacaacttig aattcagcta caccttcgag 2100
gacgtgcctt tccacagcag ctacgcgcac agccagagcc tggaccggct gatgaaccct 2160
ctcatcgacc agtacctgta ttacttatcc agaactcagt ccacaggagg aactcaaggt 2220
actcagcaat tgttatttic tcaagccggg cccgcaaaca tgtcggctca ggccaagaac 2280
tggctacctg gaccgtgtta ccgtcagcaa cgagtttcca cgacactgtc gcaaaacaac 2340
aacagcaatt ttgcttggac cggtgccacc aagtatcacc tgaatggcag agactccctg 2400
gttaatcccg gcgttgccat ggctacccac aaggacgacg aggagcgctt cttccgtca 2460
agcggagtic taatgttigg caagcagggg gcigaaaaag acaatgtgga ctacagcagc 2520
gtgatgctca ccagcgaaga agaaattaaa actactaacc cagtggctac agagcagtat 2580
ggtgtggtgg cagacaacct gcagcagacc aacggagctc ccattgtggg aactgtcaac 2640
agccaggggg ccttaccigg tatggtctgg caaaaccggg acgtgtacct gcagggcccc 2700
atctgggcca aaattcctca cacggacggc aactttcatc cttcgccgct gatgggaggc 2760
tttggactga aacaccccgcc tcctcagatc ctggtgaaaa acactcctgt tcctgcggat 2820
cctccgacca ccttcagcca ggccaagctg gcttctttta tcacgcagta cagcaccgga 2880
caggtcagcg tggaaatcga atgggagctg cagaaagaaa acagcaagcg ctggaacccg 2940
gagattcagt atacttccaa ctactacaaa tctacaaatg tggactttgc tgtcaatact 3000
gagggtactt attcagagcc tcgccccatt ggcactcgtt atctcaccog taatctgtaa 3060
ttgcttgtta atcaataaac cggt 3084

<210> 40
<211> 2370
<212> DNA
<213> 新 AAV 血清型, 克隆 43.5

<400> 40
gaattcgccc tttctacggc tgcgtcaact ggaccaatga gaactticcc ttcaacgatt 60
gcgtcgacaa gatggtgatc tggtgggagg agggcaagat gacggccaag gtcgtggagt 120
ccgccaaggc cattctcggc ggcagcaagg tgcgcgtgga ccaaaagtgc aagtcgtccg 180
cccagatcga ccccacccct gtgatcgtca cctccaacac caacatgtgc gccgtgattg 240
acgggaacag caccaccttc gagcaccagc agccgttgca ggaccggatg ttcaagttcg 300
aactcacccg ccgtctggag cacgactttg gcaaggtgac caagcaggaa gtcaaagagt 360
tcttccgctg ggcgcaggat cacgtgaccg aggtggcgca tgagttctac gtcagaaagg 420
gcggagccag caaaagaccc gcccccgatg acgcggatat aagcgagccc aagcgggcct 480
gcccctcagt cgcggatcca tcgacgtcag acgcggaagg agctccggtg gactttgccg 540
acaggtacca aaacaaatgt tctcgtcacg cgggcatgct tcagacgctg tttcccigca 600
aaacgtgcga gagaatgaat cagaatttca acattigctt cacgcacggg gtcagagact 660
gctcagaatg tttcccggct gcatcagaat ctcaaccggt cgtcagaaaa aaaacgtatc 720
agaaaactgtg tgccattcat catctgctgg ggcgggcacc cgagattgct tgctcggcct 780
gcgatctggt caacgtggac ctggacgact gtgtttctga gcaataaatg acttaaacca 840
ggtatggctg ccgatggtta tcttccagat tggcttgagg acaacctctc tgagggcatt 900
cgcgagtggt gggacctgaa acctggagcc ccgaaaccca aagccaacca gcaaaagcag 960

gacgacggcc ggggtctggt gcttcctggc tacaagtacc tcggacccct caacggactc 1020
gacaaggggg agcccgtcaa cgcggcggac gcagcggccc tcgagcacga caaggcctac 1080
gaccagcagc tcaaagcggg tgacaatccg tacctgcggt ataaccacgc cgacgccgag 1140
tttcaggagc gtctgcaaga agatacgtct tttgggggca acctcgggcg agcagtcttc 1200
caggccaaga agcgggttct cgaacctctc ggtctggttg aggaaggcgc taagacggct 1260
cctggaaaga agagaccggt agagccatca cctcagcgtt ccccgactc ctccacgggc 1320
atcggcaaga aaggccacca gccgcgagca aagagactga actttgggca gactggcgac 1380
tcggagtcag tccccgaccc tcaaccaatc ggagaaccac cagcaggccc ctctggtctg 1440
ggatctggta caatggctgc aggcggtggc gctccaatgg cagacaataa cgaaggcgcc 1500
gacggagtgg gtagttcctc aggaaattgg cattgcgatt ccacatggct gggcgacaga 1560
gtcatcacca ccagcacccg aacctgggcc ctgcccacct acaacaacca tctctacaag 1620
caaaictcca acgggacatc gggaggaagc actaacgaca acacctactt tggctacagc 1680
acccctgggg ggtatttiga cttcaacaga ttccactgcc acttctcacc acgtgactgg 1740
cagcgactca tcaacaataa ctggggattc cggccccaaga gactcaactt caagctcttc 1800
aacatccagg tcaaggaggt cacgcagaat gaaggcacca agaccatcgc caataacctt 1860
accagcacga ttcaggtgtt tacggactcg gaataccagc tcccgtacgt cctcggctct 1920
gcgcaccagg gctgcctccc tccgttcccg gcggacgtct tcatgattcc tcagtacggg 1980
tatctgaccc taaacaatgg cagtcaggct gtgggccgtt cctccttcta ctgcctggaa 2040
tacttccctt ctcaaatgct gaggacgggc aacaactttg aattcagcta caccttcgag 2100
gacgtgcctt tccacagcag ctacgcgcac agccagagcc tggaccggct gatgaaccct 2160
ctcatcgacc agtacctgta ttacttatcc agaactcagt ccacaggagg aactcaaggt 2220
actcagcaat tgttatttic tcaagccggg cccgcaaaca tgtyggctca ggccaagaac 2280
tggctacctg gaccgtgtta ccgtcagcaa cgagtttcca cgacactgtc gcaaaacaac 2340
aacagcaatt ttgctggacc ggtgccacca 2370

<210> 41
<211> 3123
<212> DNA
<213> 43.12

[0042] <400> 41
gaattcgccc ttggctgcgt caactggacc aatgagaact ttcccttcaa cgattgcgtc 60
gacaagatgg tgatctggtg ggaggagggc aagatgacgg ccaaggtcgt ggagtccgcc 120
aaggccattc tcggcggcag caaggtgcgc gtggaccaaa agtgcaagtc gtccgcccag 180
atcgacccca cccccgtgat cgtcacctcc aacaccaaca tgtgcgccgt gattgacggg 240
aacagcacca ccttcgagca ccagcagccg ttgcaggacc ggatgttcaa gttcgaactc 300
acccgccgtc tggagcacga ctttggcaag gtgaccaagc aggaagtcaa agagttcttc 360
cgctgggcgc aggatcacgt gaccgaggtg gcgcatgagt tctacgtcag aaagggcgga 420
gccagcaaaa gacccgcccc cgatgacgcg gatataagcg agcccaagcg ggcctgcccc 480
tcagtcgcgg atccatcgac gtcagacgcg gaaggagctc cggtggactt tgccgacagg 540
taccaaaaca aatgttctcg tcacgcgggc atgctccaga tgctgtttcc ctgcaaaacg 600
tgcgagagaa tgaatcagaa tttcaacatt tgcttcacgc acggggtcag agactgctca 660
gaatgtttcc ccggtgcatc agaatctcaa ccggtcgtca gaaaaaaaac gtatcagaaa 720
ctgtgtgcca ttcatcatct gctgggcggg gcacccgaga ttgcttgctc ggcctgcgat 780
ctggtcaacg tggacctgga cgactgtgtt tctgagcaat aaatgactta aaccaggtat 840
ggctgccgat ggttatcttc cagattggct tgaggacaac ctctctgagg gcattgcgca 900
gtggtgggac ctgaaacctg gagccccgaa accaaagcc aaccagcaaa agcaggacga 960
cggccgggat ctggtgcttc ctggctacaa gtacctcgga cccttcaacg gactcgacaa 1020
gggggagccc gtcaacgcgg cggacgcagc ggccctcgag cacgacaagg cctacgacca 1080
gcagctcaaa gcgggtgaca atccgtacct gcggtataac cacgccgacg ccgagtttca 1140
ggagcgtctg caagaagata cgtcttttgg gggcaacctc gggcgagcag tcttcaggc 1200
caagaagcgg gttctcgaac ctctcggtct ggttgaggaa ggcgctaaga cggctcctgg 1260
aaagaagaga ccggtagagc catcaccica gcgttcccc gactctica cgggcatcgg 1320
caagaaaggc caccagcccg cgagaaagag actgaacttt gggcagactg gcgactcgga 1380
gtcagtcccc gacctcaac caatcggaga accaccagca ggccccctctg gtctgggatc 1440
tggtacaatg gctgcaggcg gtggcgctcc aatggcagac aataacgaag gcgccgacgg 1500
agtgggtagt tcctcaggaa attggcattg cgattccaca tggctgggcg acagagtcat 1560
caccaccagc acccgaacct gggccctgcc cacctacaac aaccatctct acaagcaaat 1620
ctccaacggg acatcgggag gaagcactaa cgacaacacc tactttggct acagcacccc 1680
ctggggtat tttgacttca acagattcca ctgccacttc tcaccacgtg actggcagcg 1740
actcatcaac aataactggg gattccggcc caagagactc aacttcaagc tcttcaacat 1800
ccaggtcaag gaggtcacgc agaatgaagg caccaagacc atcgccaata accttaccag 1860

cacgattcag gtgtttacgg actcggaata ccagctcccg tacgtcctcg gctctgcgca 1920
ccagggctgc ctccctccgt tcccggcgga cgtcttcatg attcctcagt acgggtatct 1980
gaccctaaac aatggcagtc aggctgtggg ccgttcctcc ttctactgcc tggaatactt 2040
cccttctcaa atgctgagga cgggcaacaa ctttgaattc agctacacct tcgaggacgt 2100
gccttttcac agcagctacg cgcacagcca gagcctggac cggctgatga accctctcat 2160
cgaccagtac ctgtattact tatccagaac tcagtcccca ggaggaactc aaggtactca 2220
gcaattgtta ttttctcaag ccgggcccgc aaacatgtcg gctcaggcca agaactggct 2280
acctggaccg tgttacgtct agcaacgagt ttccacgaca ctgtcgcaaa acaacaacag 2340
caattttgct tggaccggtg ccaccaagta tcacctgaat ggcagagact ccctggttaa 2400
tcccggcgtt gccatggcta cccacaagga cgacgaggag cgcttcttcc cgtcaagcgg 2460
agttctaatg tttggcaagc agggggctgg aaaagacaat gtggactaca gcagcgtgat 2520
gctcaccagc gaagaagaaa ttaaaactac taacccagtg gctacagagc agtatggtgt 2580
ggtggcagac aacctgcagc agaccaacgg agctcccatt gtgggaactg tcaacagcca 2640
gaggggctta cctggtatgg tctggcaaaa ccgggacgtg tacctgcagg gccccatctg 2700
ggccaaaatt cctcacacgg acggcaactt tcatcctccg ccgctgatgg gaggctttgg 2760
actgaaacac ccgcctcctc agatcctggt gaaaaacact cctgttcctg cggatcctcc 2820
gaccacctte agccaggcca agctggcttc ttttatcacg cagtacagca ccggacaggt 2880
cagcgtggaa atcgaatggg agctgcagaa agaaaacagc aagcgctgga acccagagat 2940
tcagtatact tccaactact acaaaatctac aaatgtggac tttgctgtca atactgaggg 3000
tacttattca gagcctcgcc ccattggcac tcgttatctc accegtaatc tgtaattgct 3060
tgttaatcan taaaccggtt aattcgtttc agttgaactt tggtctctgc gaagggcgaa 3120
ttc 3123

<210> 42
<211> 3122
<212> DNA
<213> 43.20

[0043]

<400> 42
gaattcgccc tttctacggc tgcgtcaact ggaccaatga gaacttccc ttcaacgatt 60
gcgtcgacaa gatggtgatc tggtgggagg agggcaagat gacggccaag gtcgtggagt 120
ccgccaaggc cattctcggc ggcagcaagg tgcgtgtgga ccaaaagtgc aagtcttccg 180
cccagatcga tcccaccccc gtgatcgtca cctccaacac caacatgtgc gccgtgattg 240
acgggaacag cgccaccttc gagcaccagc agccgttgca ggaccggatg ttcaaatttg 300
aactcacccg ccgtctggag catgacttig gcaaggtgac gaagcaggaa gtcaaagagt 360
tcttccgctg ggcgcaggat cacgtgaccg aggtggcgca tgagttccac gtcagaaagg 420
gtggagccaa caagagaccc gccccgatg acgcggatat aagcgagccc aagcgggcct 480
gcccctcagt cgcggatcca tcgacgtcag acgcggaagg agctccggtg gactttgccg 540
acaggtacca aaacaaatgt tctcgtcacg cgggcatgct tcagatgctg tttcctgca 600
agacatgcga gagaatgaat cagaatttca acatttgctt cacgcacggg accagagact 660
gttcagaatg tttccccggc gtgtcagaat ctcaaccggt cgtcagaaag aggacgtatc 720
ggaaaactctg tgcgattcat catctgctgg ggcgggctcc cgagattgct tgctcggcct 780
gcgatctggt caacgtggac ctggatgact gtgtttctga gcaataaatg acttaaacca 840
ggtatggctg ccgatggtta tcttccagat tggctcgagg acaacctctc tgagggcatt 900
cgcgagtggt gggacttgaa acctggagcc ccgaaaccca aagccaacca gcaaaagcag 960
gacgacggcc ggggtctggt gcttcctggc tacaagtacc tcggaccctt caacggactc 1020
gacaaggggg agccccgtcaa cgcggcggac gcagcggccc tcgagcacga caaagcctac 1080
gaccagcagc tcaaagcggg tgacaatccg tacctgcggt ataatcacgc cgacgccgag 1140
tttcaggagc gtctgcaaga agatacgtct tttgggggca acctcgggcg agcagtcttc 1200
caggccaaga agcgggttct cgaacctctc ggtctggttg aggaaggcgc taagacggct 1260
cctggaaaga agagactggt agagcagtcg ccacaagagc cagactcctc ctcgggcatc 1320
ggcaagacag gccagcagcc cgctaaaaag agactcaatt ttggtcagac tggcgactca 1380
gagtcagtcc ccgacccaca acctctcgga gaacctccag cagccccctc aggtctggga 1440
cctaatacaa tggcttcagg cggtggcgct ccaatggcag acaataacga aggcgccgac 1500
ggagtgggta attcctcggg aaattggcat tgcgattcca catggctggg ggacagagtc 1560
atcaccacca gcacccgaac ctgggccctg ccccactacs acaaccacct ctacaagcaa 1620
atctccaacg gcacctcggg aggaagcacc aacgacaaca cctattttgg ctacagcacc 1680
cctgggggt attttgactt caacagattc cactgtcact tttcaccacg tgactggcaa 1740
cgactcatca acaacaattg gggattcggg ccccaaaagac tcaacttcaa gctgttcaac 1800
atccaggtca aggaagtcac gacgaacgaa ggcaccaaga ccatcgccaa taatctcacc 1860
agcaccgtgc aggtctttac ggactcggag taccagttac cgtacgtgct aggatccgct 1920
caccagggat gtctgcctcc gttcccggcg gacgtcttca cggttcctca gtacggctat 1980

ttaactttaa acaatggaag ccaagccctg ggacgttcct ccttctactg tctggagtat 2040
ttcccatcgc agatgctgag aacccggcaac aactttcagt tcagctacac cttcgaggac 2100
gtgcctttcc acagcagcta cgcgcacagc cagagcctgg acaggctgat gaatccgctc 2160
atcgaccagt acctgtacta cctggtcaga acgcaaacga ctggaactgg agggacgcag 2220
actctggcat tcagccaagc gggtcctagc tcaatggcca accaggctag aaattgggtg 2280
cccggacctt gctacggcga gcagcgcgtc tccacgacaa ccaaccagaa caacaacagc 2340
aactttgcct ggacgggagc tgccaagttt aagctgaacg gccgagactc tctaatgaat 2400
ccgggcgtgg caatggcttc ccacaaggat gacgacgacc gcttcttccc ttcgagcggg 2460
gtcctgattt ttggcaagca aggagccggg aacgatggag tggattacag ccaagtgctg 2520
attacagatg aggaagaaat caaggctacc aaccccgtgg ccacagaaga atatggagca 2580
gtggccatca acaaccaggc cgccaatacg caggcgcaga ccggactcgt gcacaaccag 2640
ggggtgattc ccggcatggt ggtgcagaat agagacgtgt accttgcagg gccaatctgg 2700
gccaaaattc ctcacacgga cggcaacttt cacccgtctc ccctgatggg cggctttgga 2760
ctgaagcacc cgcctcctca aattctcatc aagaacacac cggttccagc ggacccgccg 2820
cttaccttca accaggccaa gctgaactct ttcatcacgc agtacagcac cggacaggtc 2880
agcgtggaaa tcgagtggga gctgcagaaa gaaaacagca aacgctggaa tccagagatt 2940
caatacactt ccaactacta caaatctaca aatgtggact ttgctgtcaa cacggaagga 3000
gtttatagcg agcctcgccc cattggcacc cgttacctca cccgcaacct gtaattacat 3060
gttaatcaat aaaccggtta attcgtttca gttgaacttt ggtctctgcg aagggcgaat 3120
tc 3122

<210> 43
<211> 3117
<212> DNA
<213> 43.21

<400> 43

[0044]

gaattcgccc ttggctgcgt caactggacc aatgagaact ttcccttcaa cgattgcgtc 60
gacaagatgg tgatcttgtg ggaggagggc aagatgacgg ccaaggtcgt ggagtccgcc 120
aaggccatic tcggcggcag caaggtgcgt gtggaccaaa agtgcaagtc ttccgcccag 180
atcgatccca ccccccgtgat cgtcacctcc aacaccaaca tgtgcgccgt gattgacggg 240
aacagcacca ccttcgagca ccagcagccg ttgcaggacc ggatgttcaa atttgaactc 300
acccgccgtc tggagcatga ctttggcaag gtgacgaagc aggaagtcaa agagttcttc 360
cgctgggcgc aggatcacgt gaccgaggtg gcgcatgagt tccacgtcag aaagggtgga 420
gccaacaaga gacccgcccc cgatgacgcg gatataagcg agccccaagcg ggcctgcccc 480
tcagtcgcgg atccatcgac gtcagacgcg gaaggagctc cggtggactt tgccgacagg 540
taccaaaaca aatgttctcg tcacgcgggc atgcttcaga tgctgtttcc ctgcaagaca 600
tgcgagagaa tgaatcagaa tttcaacatt tgcttcacgc acgggaccag agactgttca 660
gaatgtttcc ccggcgtgtc agsatctcaa ccggtcgtca gaaagaggac gtatcggaaa 720
ctctgtgcga ttcatcatct gctggggcgg gctcccgaga ttgcttgctc ggcctgcgat 780
ctggtcaacg tggacctgga tgactgtgtt tctgagcaat aaatgactta aaccaggtat 840
ggctgccgat ggttatcttc cagattggct cgaggacaac ctctctgagg gcattcgcga 900
gtggtgggac ttgaaaactg gagccccgaa accccaaagcc aaccagcaaa agcaggacga 960
cggccggggt ctggtgcttc ctggctacaa gtacctcgga cccttcaacg gactcgacaa 1020
gggggagccc gtcaacgcgg cggacgcagc ggccctcgag cacgacaaag cctacgacca 1080
gcagctcaaa gcgggtgacn atccgtacct gcggtatnat cacgccgacg ccgagtttca 1140
ggagcgtctg caagaagata cgtctttngg gggcaacctc gggcgagcag tcttccaggc 1200
caagaagcgg gttctcgaac ctctcggtct ggttgaggaa ggcgctaaga cggctcctgg 1260
aaagaagaga ccggtagagc agtcgccaca agagccagac tcctcctcgg gcatcggcaa 1320
gacaggccag cagcccgcta aaaagagact caattttggt cagactggcg actcagagtc 1380
agtccccgac ccacaacctc tcggagaacc tccagcagcc ccctcaggtc tgggacctaa 1440
tacaatggct tcaggcggtg gcgctccaat ggcagacaat aacgaaggcg ccgacggagt 1500
gggtaattcc tcgggaaatt ggcattgcga ttccacatgg ctgggggaca gagtcatcac 1560
caccagcacc cgaacctggg ccctgcccac ctacaacaac cacctctaca agcaaatctc 1620
caacggcacc tcgggaggaa gcaccaacga caacacctat tttggctaca gcaccccctg 1680
ggggtatttt gacttcaaca gattccactg tcacttttca ccacgtgact ggcaacgact 1740
catcaacaac aattggggat tcoggcccaa aagactcaac ttcaagctgt tcaacatcca 1800
ggtcaaggaa gtcacgacga acgaaggcac caagaccatc gccaataatc tcaccagcac 1860
cgtgcgggtc tttacggact cggagtacca gttaccgtac gtgctaggat ccgctcacca 1920
gggatgtctg cctccgttcc cggcggacgt cttcatggtt cctcagtacg gctatttaac 1980
tttaaacaat ggaagccaag ccctgggacg ttcctcttc tactgtctgg agtatttccc 2040
atcgcagatg ctgagaaccg gcaacaactt tcagttcagc tacaccttcg aggacgtgcc 2100

tttccacagc agctacgcgc acagccagag cctggacagg ctgatgaatc ccctcatcga 2160
ccagtacctg tactacctgg tcagaacgca aacgactgga actggaggga cgcagactct 2220
ggcattcagc caagcgggtc ctagctcaat ggccaaccag gctagaaatt gggtgcccgg 2280
accttgctac cggcagcagc gcgtctccac gacaaccaac cagagcaaca acagcaactt 2340
tgcctggacg ggagctgcca agtttaagct gaacggccga gactctctaa tgaatccggg 2400
cgtggcaatg gcttcccaca aggatgacga cgaccgcttc ttcccttcga gcggggtcct 2460
gatttttggc aagcaaggag ccgggaacga tggagtggat tacagccaag tgctgattac 2520
agatgaggaa gaaatcaagg ctaccaaccc cgtggccaca gaagaatatg gagcagtggc 2580
catcaacaac caggccgcca atacgcaggc gcagaccgga ctcgtgcaca accaggggct 2640
gattcccggc atggtgtggc agaatagaga cgtgtacctg cagggtccca tctgggccaa 2700
aattcctcac acggacggca actttcaccc gtctcccctg atgggcggct ttggactgaa 2760
gcaccccgcc cctcaaatta tcatcaagaa cacaccggtt ccagcggacc cgccgcttac 2820
cttcaaccag gccaagctga actctttcat cacgcagtac agcaccggac aggtcagcgt 2880
ggaaatcgag tgggagctgc agaagaaaaa cagcaaacgc tggaatccag agattcaata 2940
cacttccaac tactacaaat ctacaaatgt ggactttgct gtcaacacgg aaggagttta 3000
tagcgagcct cgccccattg gcacccgtta cctcacccgc aacctgtaat tacatgttaa 3060
tcaataaacc ggttaattcg tttcagttga actttggtct ctgcgaaggg cgaattc 3117

<210> 44
<211> 3121
<212> DNA
<213> 43.23

<400> 44
gaattcgccc ttctacggct gcgtcaactg gaccaatgag aacttccct tcaacgattg 60
cgtcgacaag atggtgatct ggtgggagga gggcaagatg acggccaagg tcgtggagtc 120
cgccaaggcc attctcggcg gcagcaaggt gcgtgtggac caaaagtgca agtcttccgc 180
ccagatcgat cccaccccg tgatcgtcac ctccaacacc aacatgtgcg ccgtgattga 240
cgggaacagc accaccttcg agcaccagcc gctgcaggac cgcggcatgt tcaaattiga 300
actcaccccc cgtctggagc atgacttigg caaggtgacg aagcaggaag tcaaagagtt 360
cttccgctgg gcgcaggatc acgtgaccga ggtggcgcat gagttccacg tcagaaaggg 420
tggcgccaac aagagacccg ccccgatga cgcggatata agcgagccca agcgggcctg 480
ccctcagtc gcggatccat cgacgtcaga cgcggaagga gctccggtgg actttgccga 540
caggtaccaa aacaaatgtt ctcgtcacgc gggcatgctt cagatgctgt ttcctgcaa 600
gacatgcgag agaatgaatc agaatttcaa catttgcttc acgcacggga ccagagactg 660
ttcagaatgt ttccccggcg tgtcagaatc tcaaccggtc gtcagaaaga ggacgtatcg 720
gaaactctgt gcgatteatc atctgctggg gcgggctccc gagattgctt gctcggcctg 780
cgatctggtc aacgtggacc tggatgactg tgtttctgag caataaatga cttaaaccag 840
gtatggctgc cgatggttat cttccagatt ggctcgagga caacctctct gagggcattc 900
gcgagtggtg ggacttgaaa cctggagccc cgaaacccaa agccaaccag caaaagcagg 960
acgacggccg gggtctggtg cttcctggct acaagtacct cggaccctic aacggactcg 1020
acaagggggc gcccgtcaac gcggcggacg cagcggccct cgagcacgac aaagcctacg 1080
accagcagct caaagcgggt gacaatccgt acctgcggta taatcacgcc gacgccgagt 1140
ttcaggagcg tctgcaagaa gatacgtcct ttgggggcaa cctcgggcga gcagtcttcc 1200
aggccaagaa gcggglctic gaacctctcg gtctggttga ggaaggcgct aagacggctc 1260
ctggaaagaa gagaccggta gagcagtcgc cacaagagcc agactcctcc tcgggcatcg 1320
gcaagacagg ccagcagccc gctaaaaaga gactcaatti tggtcagact ggcgactcag 1380
agtcagtccc cgaccccacaa cctctcggag aacctccagc agcccccica ggtctgggac 1440
ctaatacaat ggcttcaggc ggtggcgctc caatggcaga caataacgaa ggcgccgacg 1500
gagtgggtaa ttcctcggga aattggcatt gcgattccac atggctgggg gacagagtca 1560
tcaccaccag cacccgaacc tgggcccigc ccacctacaa caaccacctc tacaagcaaa 1620
tctccaacgg cacctcggga ggaagcacca acgacaacac ctattttggc tacagcaccc 1680
cctgggggta ttttgacttc aacagattcc actgtcactt ttcaccacgt gactggcaac 1740
gactcatcaa caacaattgg ggattccggc ccaaaagact caacttcaag ctgttcaaca 1800
tccaggtcaa ggaagtcacg acgaacgaag gcaccaagac catcgccaat aatctcacca 1860
gcaccgtgca ggtctttacg gacttggagt accagttacc gtacgtgcta ggatccgctc 1920
accagggatg tctgcctccg ttcccggcgg acgtcttcat ggticctcag tacggctatt 1980
taactttaaa caatggaagc caagcctigg gacgttcctc cttctactgt ctggagtatt 2040
tcccatcgca gatgccgaga accggcaaca acttccagtt cagctacacc ttcgaggacg 2100
tgcctttcca cagcagctac gcgcacagcc agagcctgga caggctgatg aatcccctca 2160
tcgaccagta cctgtactac ctggtcagaa cgcaaacgac tggaactgga gggacgcaga 2220
ctctggcatt cagccaagcg ggtcctagct caatggccaa ccaggctaga aattgggtgc 2280

[0045]

ccggaccttg ctaccggcag cagcgcgtct ccacgacaac caaccagaac aacaacagca 2340
actttgcctg gacgggagct gccaagttta agctgaacgg ccgagactct ctaatgaatc 2400
cgggcgtggc aatggcttcc cacaaggatg acgacgaccg cttcttccct tcgagcgggg 2460
tcctgatttt tggcaagcaa ggagccggga acgatggagt ggattacagc caagtgctga 2520
ttacagatga ggaagaaatc aaggctacca accccgtggc cacagaagaa tatggagcag 2580
tggccatcaa caaccaggcc gccaatacgc aggcgcngac cggactcgtg cacaaccagg 2640
gggtgattcc cggcatggtg tggcagaata gagacgtgta cctgcagggt ccccatctggg 2700
ccaaaattcc tcacacggac ggcaactttc accegtctcc cctgatgggc ggctttggac 2760
tgaagcaccc gcctcctcaa attctcatca agaacacacc ggttccagcg gacccgccgc 2820
ttaccttcaa ccaggccaag ctgaactctt tcatcacgca gtacagcacc ggacaggtca 2880
gcgtggaaat cgagtgggag ctgcagaaag aaaacagcaa acgctggaat ccagagattc 2940
aatacacttc caactaciac aaaictacaa atgtggacti tgctgtcaac acggaaggag 3000
tttatagcga gcctcgcccc attggcaccc gttaccicac ccgcaacctg taattacatg 3060
ttaatcaata aacoggttaa ttcgtttcag ttgaactttg gtctctgcga agggcgaatt 3120
c 3121

<210> 45
<211> 3122
<212> DNA
<213> 43.25

<400> 45

[0046]

gaattcgccc tttctacggc tgcgtcaact ggaccaatga gaactttccc ttcaacgatt 60
gcgtcgacaa gatggtgatc tggtgggagg agggcaagat gacggccaag gtcgtggagt 120
cogccaaggc cattctcggc ggcagcaagg tgcgtgtgga ccaaaagtgc aagtcttccg 180
cccagatcga tcccaccccc gtgatcgtca cctccaacac caacatgtgc gccgtgattg 240
acgggaacag caccaccttc gagcaccagc agccgttgca ggaccggatg ttcaaatttg 300
aactcacccg ccgtctggag catgacttig gcaaggtgac gaagcaggaa gtcaaagggt 360
tcttccgctg ggcgcaggat cacgtgaccg aggtggcgca tgagttccac gtgcgagccc 420
aagcgggcct gcccctcagt cgcggatcca tcgacgtcag accagaaagg gtggagccaa 480
caagagaccc gcccccgatg acgcggatat aagcggaagg agctccggtg gactttgccg 540
acaggtacca aaacaaatgt tctcgtcacg cgggcatgct tcagatgctg tttcctgca 600
agacatgcga gagaatgaat cagaatttca acatttgctt cacgcacggg accagagact 660
gttcagaatg tttccccggc gtgtcagaat ctcaaccggt cgtcagaaag aggacgtatc 720
ggaaaactctg tgcgattcat catctgctgg ggcgggctcc cgagattgct tgctcggcct 780
gcgatctggt caacgtggac ctggatgact gtgtttctga gcaataaatg acttaaacca 840
ggtatggctg ccgatggtta ttttccagat tggctcgagg acaacctctc tgagggcatt 900
cgcgagtggt gggacttgaa acctggagcc ccgaaaccca aagccaacca gcaaaagcag 960
gacgacggcc ggggtctggt gcttcctggc tacaagtacc tcggacccit caacggactc 1020
gacaaggggg agcccgtcaa cgcggcggac gcagcggccc tcgagcacga caaagcctac 1080
gaccagcagc tcaaagcggg tgacaatccg tacctgcggt ataatcacgc cgacgccgag 1140
ttccaggagc gtctgcaaga agatacgtct tttgggggca acctcgggcg agcagtcttc 1200
caggccaaga agcgggttct cgaacctctc ggtctggttg aggaaggcgc taagacggct 1260
cctggaaaga agagaccggt agagcagtcg ccacaagagc cagactcctc ctcgggcatc 1320
ggcaagacag gccagcagcc cgctaaaaag agactcaatt ttggtcagac tggcgactca 1380
gagtcagtcc ccgacccaca accictcgga gaacctccag cagccccctc aggtctggga 1440
cctaatacaa tggcttcagg cggtggcgct ccaatggcag acaataacga aggcgccgac 1500
ggagtgggta attcctcggg aaattggcat tgcgattcca catggctggg ggacagagtc 1560
atcaccacca gcacccgaac ctgggccctg cccacctaca acaaccacct ctacaagcaa 1620
atctccaacg gcacctcggg aggaagcacc aacgacaaca tctattttgg ctacagcacc 1680
ccctgggggt attttgactt caacagattc cactgtcact tttccaccacg tgactggcaa 1740
cgactcatca acaacaattg gggattccgg cccaaaagac tcaacttcaa gctgttcaac 1800
atccaggtca aggaagtcac gacgaacgaa ggcaccaaga ccatcgccaa taatctcacc 1860
agcaccgtgc aggtctttac ggactcggag taccagttac cgtacgtgct aggatccgct 1920
caccagggat gtctgcctcc gttcccggcg gacgtcttca tggttcctca gtacggctat 1980
ttaactttaa acaatggaag ccaagcctig ggacgttcct ccttctactg tctggagtat 2040
ttcccatcgc agatgctgag aacccggcaac aacttticagt tcagctacac cttcgaggac 2100
gtgcctttcc acagcagcta cgcgcacagc cagagcctgg accggctgat gaatcccctc 2160
atcgaccagt acctgtacta cctggtcaga acgcaaacga ctggaactgg agggacgcag 2220
actctggcat tcagccaagc gggtcctagc tcaatggcca accaggctag aaattgggtg 2280
cccggacctt gctaccggca gcagcgcgtc tccacgacaa ccaaccagaa caacaacagc 2340
aactttgcct ggacgggagc tgccaagttt aagctgaacg gccgagactc tctaatgaat 2400

ccgggcgtgg caatggcttc ccacaaggat gacgacgacc gcttcttccc ttcgagcggg 2460
gtcctgattt ttggcaagca aggagccggg aacgatggag tggattacag ccaagtgctg 2520
attacagatg aggaagaaat caaggctacc aaccccgtgg ccacagaaga atatggagca 2580
gtggccatca acaaccaggc cgccaatacg caggcgcaga ccggactcgt gcacaaccag 2640
ggggtgattc ccggcatggt gtggcagaat agagacgtgt acctgcaggg tcccatctgg 2700
gccaaaattc ctcacacgga cggcaacttt cacccgtctc ccctgatggg cggctttgga 2760
ctgaagcacc cgcctcctca aattctcatc aagaacacac cggttccagc ggacccgccg 2820
cttaccttca accaggccaa gctgaactct ttcatcacgc agtacagcac cggacaggtc 2880
agcgtggaaa tcgagtggga gctgcagaaa gaaaacagca aacgctggaa tccagagatt 2940
caatacactt ccaactacta caaatctaca aatgtggact ttgctgtcaa cacggagggg 3000
gtttatagcg agcctcgccc cattggcacc cgttacctca cccgcaacct gtaattacat 3060
gttaatcaat aaaccggtta attcgtttca gttgaacttt ggtctctgcg aagggcgaat 3120
tc 3122

<210> 46
<211> 3128
<212> DNA
<213> 44.1

<400> 46

[0047]

gaattcgccc tttctacggc tgcgtcaact ggaccaatga gaacttttcc ttcaacgatt 60
gcgtcgacaa gatgttgatc tggtgggagg agggcaagat gacggccaag gtcgtggagt 120
cgccaaggc cattctggcc ggcagcaaag tgcgcgtgga ccaaaagtgc aagccgtccg 180
cccagatcga ccccacccca gtgatcgtca cctccaacac caacatgtgc gccgtgattg 240
acgggaacag caccaccttc gagcaccagc agccgttgcg ggaccggatg ttcaagtttg 300
aactcacccg ccgtctggag cacgactttg gcaaggtgac aaagcaggaa gtcagagagt 360
tcttccgctg ggcgcaggat cacgtgaccg aggtggcgca cgagttctac gtcagaaagg 420
gtggagccaa caagagaccc gccccccgatg acgcggataa aagcgagccc aagcgggcct 480
gccccgagt cgcggatcca tcgacgtcag acgcggaagg agctccggtg gacttgccg 540
acaggtacca aaacaaaatgt tctcgtcacg cgggcatgct tcagatgctg tttccctgca 600
aaacatgcga gagaatgaat cagaatttca acatttgctt cacgcacggg accagagact 660
gttcagaatg tttccccggc gtgtcagaat ctcaaccggt cgtcagaaaa aagacgtatc 720
gaaaactctg tgcgattcat catctgctgg ggcgggcacc cgagattgct tgctcggcct 780
gcgatctggt caacgtggac ctagatgact gtgtttctga gcaataaatg acttaaacca 840
ggtatggctg ccgatggtta ttttccagat tggctcgagg acaacctctc tgagggcatt 900
cgcgagtggt gggacttgaa acctggagcc ccgaaaccca aagccaacca gcaaaagcag 960
gacgacggcc ggggtctggt gcttcctggc tacaagtacc tcggaccctt caacggactc 1020
gacaaggggg agccccgtcaa cgcggcggac gcagcggccc tcgagcacga caaggcctac 1080
gaccagcgac tcaaagcggg tgacaatccg tacctgcggt ataaccacgc cgacgccgag 1140
tttcaggagc gtctgcaaga agatacgtct tttgggggca acctcgggcg agcagtcttc 1200
caggccaaga agcgggttct cgaacctctc ggtctggttg aggaaggcgc taagacggct 1260
cctggaaagn agagaccggt agagccatca ccccagcgtt ctccagactc ctctacgggc 1320
atcggcaaga aaggccagca gcccgcgaaa aagagactca actttgggca gactggcgac 1380
tcagagtcag tgccccgacc tcaaccaatc ggagaacccc ccgcaggccc ctctggtctg 1440
ggatctggta caatggctgc aggcggtggc gctccaatgg cagacaataa cgaaggcgcc 1500
gacggagtgg gtagttcctc aggaaattgg cattgcgatt ccacatggct gggcgacaga 1560
gtcatcacca ccagcacccg aacctgggcc ctccccacct acaacaacca cctctacaag 1620
caaatctcca acgggactic gggaggaagc accaacgaca acacctactt cggctacagc 1680
accccctggg ggtattttga ctttaacaga ttccactgcc acttctcacc acgtgactgg 1740
cagcgactca tcaacaacaa ctggggattc cggccccaaga gactcaactt caagctcttc 1800
aacatccagg tcaaggaggt cacgcagaat gaaggcacca agaccatcgc caataacctt 1860
accagcacga ttcaggtctt tacggactcg gaataccagc tccgtacgt cctcggctct 1920
gcgcaccagg gctgcctgcc tccgttcccg gcggacgtct tcatgattcc tcagtacggg 1980
tacctgactc tgaacaatgg cagtcaggcc gtgggccgtt cctccttcta ctgcctggag 2040
tactttcctt ctcaaatgct gagaacgggc aacaactttg agttcagcta ccagtttgag 2100
gacgtgcctt ttcacagcag ctacgcgcac agccaaagcc tggaccggct gatgaaccct 2160
ctcatcgacc agtacctgta ctacctgtct cggactcagt ccacgggagg taccgcagga 2220
actcagcagt tgctatttct tcaggccggg cctaataaca tgtcggctca ggccaaaaac 2280
tggctacccg ggccctgcta ccggcagcaa cgcgtctcca cgacactgtc gcaaaataac 2340
aacagcaact gtaaatcccg gtgtcgctat ggcaacccac aaggacgacg aagagcgatt 2400
ttgcctggac cggtgccacc aagtatcatc tgaatggcag agactctctg ttttccgtcc 2460
agcggagtct taatgtttgg gaaacagggg gctggaaaag acaacgtgga ctatagcagc 2520

gttatgctaa ccagtgagga agaaattaaa accaccaacc cagtggccac ggaacagtac 2580
ggcgtggtgg ccgataacct gcaacagcaa aacgccgctc ctattgtagg ggccgtcaac 2640
agtcaaggag ccttacctgg catggtctgg cagaaccggg acgtgtacct gcagggtcct 2700
atctgggcca agattcctca cacggacgga aactttcatc cctcgccgct gatgggaggc 2760
tttggactga aacacccgcc tcctcagatc ctgattaaga atacacctgt tccccgcggat 2820
cctccaacta ccttcagtca agctaagctg gcgtcgttca tcacgcagta cagcaccgga 2880
caggtcagcg tggaaattga atgggagctg cagaaagaaa acagcaaaag ctggaacccc 2940
gagattcaat acacttccaa ctactacaaa tctacaaatg tggacttcgc tgttaacaca 3000
gatggcactt attctgagcc tcgccccatt ggcacccgtt acctcacccg taatctgtaa 3060
ttgctcgtta atcaataaac cggttgattc gtttcagttg aactttggtc tctgcgaagg 3120
gcgaattc 3128

<210> 47
<211> 3128
<212> DNA
<213> 44.5

<400> 47
gaattcgccc tttctacggc tgcgtcaact ggaccaatga gaactttccc ttcaacgatt 60
gcgtcgacaa gatggtgatc tggtgggagg agggcaagat gacggccaag gtcgtggagt 120
ccgccaaggc cattctcggc ggcagcaaag tgcgcgtgga ccaaaagtgc aagtcgtccg 180
cccagatcga ccccacccca gtgatcgtca cctccaacac caacatgtgc gccgtgattg 240
acgggaacag caccaccttc gagcaccagc agccgttgca ggaccggatg ttcaagtttg 300
aactcacccg ccgtctggag cacgacttig gcaaggtgac aaagcaggaa gtcagagagt 360
tcttccgctg ggcgcaggat cacgtgaccg aggtggcgca cgagttctac gtcagaaagg 420
gtggagccaa caagagaccc gccccgatg acgcggataa aagcgagccc aagcgggcct 480
gccccgagt cgcggatcca tcgacgtcag acgcggaagg agctccggtg gactttgccg 540
acaggtacca aaacaaatgt tctcgtcacg cgggcatgct tcagatgctg tttccctgca 600
aaacatgcga gagaatgaat cagaatttca acatttgctt cacgcacggg accagagact 660
gttcagaatg tttccccggc gtgtcagaat ctcaaccggt tgtcagaaaa aagacgtatc 720
ggaaactctg tgcgattcat catctgctgg ggcgggcacc cgagattgct tgctcggcct 780
gcgatctggt caacgtggac ctagatgact gtgtttctga gcaataaatg acttaaacca 840
ggtatggctg ccgatggtta ttttccagat tggctcgagg acaacctctc tgagggcatt 900
cgcgagtggt gggacttgaa acctggagcc ccgaaaccca aagccaacca gcaaaagcag 960
gacgacggcc ggggtctggt gcttcctggc tacaagtacc tcggaccctt caacggactc 1020
gacaaggggg agccccgtcaa cgcggcggac gcagcggccc tcgagcacga caaggcctac 1080
gaccagcagc tcaaagcggg tgacaatccg tacctgcggt ataaccacgc cgacgccgag 1140
tttcaggagc gtctgcaaga agatacgtct tttgggggca acctcgggcg agcagtcttc 1200
caggccaaga agcgggttct cgaacctctc ggtctggttg aggaggcgcc taagacggct 1260
cctggaaagn agagaccggt agagccatca ccccagcgtt ctccagactc ctctacgggc 1320
atcggcaaga aaggccagca gcccgcgaaa aagagactca actttgggca gactggcgac 1380
tcagagtcag tgccccgacc tcaaccaatc ggagaacccc ccgcaggccc ctctggtctg 1440
ggatctggta caatggctgc aggcggtggc gctccaatgg cagacaataa cgaaggcgcc 1500
gacggagtgg gtagttcctc aggaaattgg cattgcgatt ccacatggct gggcgacaga 1560
gtcatcacca ccagcacccg aacctgggcc ctccccacct acaacaacca cctctacaag 1620
caaatctcca acgggacttc gggaggaagc accaacgaca acacctactt cggctacagc 1680
accccctggg ggtatttiga ctttaacaga ttccactgcc acttctcacc acgtgactgg 1740
cagcgactca tcaacaacaa ctgggattc cggcccaaga gacccaactt caagctcttc 1800
aacatccagg tcaaggaggt cacgcagaat gaaggcacca agaccatcgc caataacctt 1860
accagcacga ttcaggtctt tacggactcg gaataccagc tcccgtacgt cctcggctct 1920
gcgcaccagg gctgcctgcc tccgttcccg gcggacgtct tcatgattcc tcagtacggg 1980
tacctgactc tgaacaatgg cagtcaggcc gtgggccgtt cctccttcta ctgcctggag 2040
tactttcctt ctcaaatgct gagaacgggc aacaactttg agttcagcta ccagtttgag 2100
gacgtgcctt ttcacagcag ctacgcgcac agccaaagcc tggaccggct gatgaacccc 2160
ctcatcgacc agtacctgta ctacctgtct cggactcagt ccacgggagg taccgcagga 2220
actcagcagt tgctatttic tcaggccggg cctaataaca tgtcggctca ggccaaaaac 2280
tggctacccg ggccctgcta ccggcagcaa cgcgtctcca cgacactgtc gcaaaataac 2340
aacagcaact ttgcctggac cggtgccacc aagtatcatc tgaatggcag agactctctg 2400
gtaaatcccg gtgtcgctat ggcaacccac aaggacgacg aagagcgatt ttttccgtcc 2460
agcggagtct taatgtttgg gaaacaggga gctggaaaag acaacgtgga ctatagcagc 2520
gttatgctaa ccagtgagga agaaattaaa accaccaacc cagtggccac agaacagtac 2580
ggcgtggtgg ccgataacct gcaacagcaa aacgccgctc ctattgtagg ggccgtcaac 2640

[0048]

agtcaaggag ccttacctgg catggtctgg cagaaccggg acgtgtacct gcagggtcct 2700
atctgggcca agattcctca cacggacgga aactttcatc cctcgccgct gatgggaggc 2760
tttggactga aacacccgcc tcctcagatc ctgattaaga atacacctgt tcccgcggat 2820
cctccaacta ccttcagtca agctaagctg gcgtcgttca tcacgcagta cagcaccgga 2880
caggtcagcg tggaaattga atgggagctg cagaaagaaa acagcaaacg ctggaaccca 2940
gagattcaat acacttccaa ctactacaaa tctacaaatg tggactttgc tgttaacaca 3000
gatggcactt attctgagcc tcgccccatt ggcacccgtt acctcacccg taatctgtaa 3060
ttgcttgtta atcaataaac cggttgattc gttttcagtt gaactttggt ctgcgaagg 3120
gcgaattc 3128

<210> 48
<211> 1933
<212> DNA
<213> 新AAV血清型,克隆223.10

<220>
<221> misc_feature
<222> (1302)..(1302)
<223> can be a, c, g or t
<400> 48
caaggcctac gaccagcagc tcaaagcggg tgacaatccg tacctgcggt ataaccacgc 60
cgacgccgag tttcaggagc gtcttcaaga agatacgtct tttgggggca acctcgggcg 120
agcagtcttc caggccaaaa agcgggttct cgaacctctt ggtctggttg agacgccagc 180
taagacggca cctggaaaga agcgaccggt agactcgcca gactccacct cgggcatcgg 240
caagaaaggc cagcagccag cgaaaaagag actcaacttt gggcagactg gcgactcaga 300
gtcagtcccc gaccctcaac caatcggaga accaccagca ggccccctctg gtctgggatc 360
tggtacaatg gctgcaggcg gtggcgcacc aatggctgac aataacgagg gcgccgacgg 420
agtgggtaat gcctcaggaa attggcattg cgattccaca tggctgggcg acagagtcat 480
caccaccagc acccgaacct gggccctgcc cacctacaac aaccacctct acaagcaaat 540
ctccagtcag tcagcaggga gcaccaacga taacgtctat ttcggctaca gcaccccctg 600
gagggtatttt gacttcaaca gattccattg ccacttctca ccacgtgact ggcagcgact 660
tatcaacaac aactggggat tccgcgcccaa gaagctcaac ttcaagctct tcaacatcca 720
ggtcaaggag gtcacgacga atgacggtgt cacaaccatc gctaataacc ttaccagcac 780
ggttcaggtc ttttcggact cggaatatca actgccgtac gtcctcggct ccgcgcacca 840
ggctgcctg cctccgttcc cggcagacgt gttcatgatt ccgcagtacg gatacctgac 900
tctgaacaat ggcagccaat cggtaggccg ttcctcttc tactgcctgg agtactttcc 960
ttctcagatg ctgagaacgg gcaacaactt caccttcagc tacaccttcg aggacgtgcc 1020
tttccacagc agctacgcgc acagccagag tctggaccgg ctgatgaatc cctcatcga 1080
ccagtacctg tactacttgg ccagaacaca gagcaacgca ggaggtactg ctggcaatcg 1140
ggaactgcag ttttatcagg gcggacctac caccatggcc gaacaagcaa agaactggct 1200
gcccggacct tgcttccggc aacagagagt atccaagacg ctggatcaaa ataacaacag 1260
caactttgcc tggactggtg ccacaaaata ccatttaaat gnaagaaatt cattggttaa 1320
tcccggtgtc gccatggcaa cccacaagga cgacgaggaa cgcttcttcc cttcgagcgg 1380
agttctaatt tttggcaaaa ctggagcagc taataaaact acattagaaa acgtgctcat 1440
gacaaatgaa gaagaaattc gtcctaccaa cccggtagct acggaggaat acgggattgt 1500
aagcagcaac ttgcaggcgg ctagcacccc agcccagaca caagttgtta acaaccaggg 1560
agccttacct ggcatggtct ggcagaaccg ggacgtgtac ctgcaaggtc ccattgggcc 1620
caagatttcc cacacggacg gcaacttcca cccgtctcct ctaatgggtg gctttggact 1680
gaaacacccg cctcccaaga tcctgatcaa aaacacaccg gtacctgcta atcctccaga 1740
agtgtttact cctgccaagt ttgcttcctt catcacgcag tacagcaccg ggcaagtcag 1800
cgttgagatc gagtgggagc tgcagaaaga gaacagcaag cgctggaacc cagagattca 1860
gtacacctcc aactttgaca aacagactgg agtggacttt gctgttgaca gccagggtgt 1920
ttactctgag cct 1933

<210> 49
<211> 1933
<212> DNA
<213> 新AAV血清型,克隆223.2

<400> 49
caaggcctac gaccagcagc tcaaagcggg tgacaatccg tacctgcggt ataaccacgc 60
cgacgccgag tttcaggagt gtcttcaaga agatacgtct tttgggggca acctcgggcg 120

[0049]

```
agcagtcttc caggccaaaa agcgggttct cgaacctctt ggtctggttg agacgccagc   180
taagacggca cctggaaaga agcgaccggt agactcgcca gactccacct cgggcatcgg   240
caagaaaggc cagcagcccg cgaaaaagag actcaacttt gggcagactg gcgactcaga   300
gtcagtcccc gacctcaac caatcggaga accaccagca ggccccctctg gtctgggatc   360
tggtacaatg gttgcaggcg gtggcgcacc aatggctgac aataacgagg gcgccgacgg   420
agtgggtaat gcctcaggaa attggcattg cgattccaca tggctgggcg acagagtcat   480
caccaccagc accgaacct gggccctgcc cacctacaac aaccacctct acaagcaaat   540
ctccagtcag tcagcaggga gcaccaacga taacgtctat ttcggctaca gcaccccctg   600
ggggtatttt gacttcaaca gattccattg ccacttctca ccacgtgact ggcagcgact   660
tatcaacaac aactggggat tccggccccaa gaagctcaac ttcaagctct tcaacatcca   720
ggtcaaggag gtcacgacga atgacggtgt cacaaccatc gctaataacc ttaccagcac   780
ggttcaggtc ttttcggact cggaatatca actgccgtac gtcctcggct ccgcgcacca   840
gggctgcctg cctccgttcc cggcagacgt gttcatgatt ccgcagtacg gatacctgac   900
tctgaacaat ggcagccaat cggtaggccg ttcctccttc tactgcctgg agtactttcc   960
ttctcagatg ctgagaacgg gcaacaactt caccttttagc tacaccttcg aggacgtgcc  1020
tttccacagc agctacgcgc acagccagag tctggaccgg ctgatgaatc cctcatcgga  1080
ccagtacctg tactacttgg ccagaacaca gagcaacgca ggaggtactg ctggcaatcg  1140
ggaactgcag ttttatcagg gcggacctac caccatggcc gaacaagcaa agaactggct  1200
gcccggacct tgcttccggc aacagagagt atccaagacg ctggatcaaa atcacaacag  1260
caactttgcc tggactggtg ccacaaaata ccatttaaat ggaagaaatt cattggttaa  1320
tcccggtgtc gccatggcaa ccacaagga cgacgaggaa cgcttcttcc cttcgagcgg  1380
agttctaatt tttggcaaaa ctggagcagc taataaaact acattagaaa acgtgctcat  1440
gacaaatgaa gaagaaatte gtcctaccaa ccoggtagct accgaggaat acgggattgt  1500
aagcagcaac ttgcaggcgg ctagcaccgc agcccagaca caagttgtta acaaccaggg  1560
agccttacct ggcatggtct ggcagaaccg ggacgtgtac ctgcaaggtc ccatttgggc  1620
caagattcct cacacggacg gcaactttca cccgtctcct ctaatgggtg gctttggact  1680
gaaacacccg cctcccaga tcctgatcaa aaacacgccg gtacctgcta atcctccaga  1740
agtgtttact cctgccaagt ttgcttcctt catcacgcag tacagcaccg ggcaagtcag  1800
cgttgagatc gagtgggagc tgcagaaaga gaacagcaag cgctggaacc cagagattca  1860
gtacacctcc aacttlgaca aacagacigg agtggacttl gctgttgaca gccagggtgt  1920
ttactctgag cct                                                     1933
```

[0050]

<210> 50
<211> 1933
<212> DNA
<213> 新 AAV 血清型, 克隆 223.4

<400> 50

```
caaggcctac gaccagcagc tcaaagcggg tgacaatccg tacctgcggt ataaccacgc     60
cgacgccgag tttcaggagc gtcttcaaga agataccgtct tttgggggca acctcgggcg    120
agcagtcttc caggccaaaa agcgggttct cgaacctctt ggtctggttg agacgccagc    180
taagacggca cctggaaaga agcgaccggt agactcgcca gactccacct cgggcatcgg    240
caagaaaggc cagcagcccg cgaaaaagag actcaacttt gggcagactg gcgactcaga    300
gccagtcccc gacctcaac caatcggaga accaccagca ggccccctctg gtctgggatc    360
tggtacaatg gctgcaggcg gtggcgcacc aatggctgac aataacgagg gcgccgacgg    420
agtgggtaat gcctcaggaa attggcattg cgattccaca cggctgggcg acagagtcat    480
caccaccagc accgaacct gggccctgcc cacctacaac aaccacctct acaagcaaat    540
ctccagtcag tcagcaggga gcaccaacga taacgtctat ttcggctaca gcaccccctg    600
ggggtatttt gacttcaaca gattccattg ccacttctca ccacgtgact ggcagcgact    660
tatcaacaac aactggggat tccggccccaa gaagctcaac ttcaagctct tcaacatcca    720
ggtcaaggag gtcacgacga atgacggcgt cacaaccatc gctaataacc ttaccagcac    780
ggttcaggtc ttttcggact cggaatatca actgccgtac gtcctcggct ccgcgcacca    840
gggctgcctg cctccgttcc cggcagacgt gttcatgatt ccgcagtacg gatacctgac    900
tctgaacaat ggcagccaat cggtaggccg ttcctccttc tactgcctgg agtactttcc    960
ttctcagatg ctgagaacgg gcaacaactt caccttttagc tacaccttcg aggacgtgcc   1020
tttccacagc agctacgcgc acagccagag tctgggccgg ctgatgaatc cctcatcgga   1080
ccagtacctg tactacttgg ccagaacaca gagcaacgca ggaggtactg ctggcaatcg   1140
ggaactgcag ttttatcagg gcggacctac caccatggcc gaacaagcaa agaactggct   1200
gcccggacct tgcttccggc aacagagagt atccaagacg ctggatcaaa atcacaacag   1260
caactttgcc tggactggtg ccacaaaata ccatttaaat ggaagaaatt cattggttaa   1320
tcccggtgtc gccatggcaa ccacaagga cgacgaggaa cgcttcttcc cttcgagcgg   1380
agttctaatt tttggcaaaa ctggagcagc taataaaact acattagaaa acgtgctcat   1440
```

gacaaatgaa gaagaaattc gtcctaccaa cccggtagct accgaggaat acgggattgt 1500
aagcagcaac ttgcaggcgg ctagcaccgc agcccagaca caagttgtta acaaccaggg 1560
agccttacct ggcatggtct ggcagaaccg ggacgtgtac ctgcaaggtc ccatttgggc 1620
caagattcct cacacggacg gcaactttca cccgtctcct ctaatgggtg gctttggact 1680
gaaacacccg cctcccccaga tcctgatcaa aaacacaccg gtacctgcta atcctccaga 1740
agtgtttact ccigccaagt ttgcttcctt catcacgcag tacagcaccg ggcaagtcag 1800
cgttgagatc gaatgggagc tgcagaaaga gaacagcaag cgctggaacc cagagattca 1860
gtacacctcc aactttgaca aacagactgg agtggacttt gctgttgaca gccagggtgt 1920
ttactctgag cct 1933

<210> 51
<211> 1933
<212> DNA
<213> 新AAV血清型,克隆223.5

<400> 51
caaggcctac gaccagcagc tcaaagcggg tgacaatccg tacctgcggt ataaccacgc 60
cgacgccgag tttcaggagc gtcttcaaga agatacgtct tttgggggca acctcgggcg 120
agcagtcttc caggccaaaa agcgggttct cgaacctctt ggtctggttg agacgccagc 180
taagacggca cctggaaaga agcgaccggt agactcgcca gactccacct cgggcatcgg 240
caagaaaggc cagcagcccg cgaaaaagag actcaacttt gggcagactg gcgactcaga 300
gtcagtcccc gaccctcaac caatcggaga accaccagca ggcccctctg gtctgggatc 360
tggtacaatg gctgcaggcg gtggcgcacc aatggctgac aatacgaggg gcgccgacgg 420
agtgggtaat gcctcaggaa attggcattg cgattccaca cggctgggcg acagagtcat 480
caccaccagc acccgaacct gggccctgcc cacctacaac aaccacctct acaagcaaat 540
ctccagtcag tcagcaggga gcaccaacga taacgtctat ttcggctaca gcacccctg 600
ggggtatttt gacttcaaca gattccattg ccacttctca ccacgtgact ggcagcgact 660
tatcaacaac aactggggat tccggcccaa gaagctcaac ttcaagctct tcaacatcca 720
ggtcaaggag gtcacgacga atgacggcgt cacaaccatc gctaataacc ttaccagcac 780
ggttcaggtc tttcaggact cggaatatca actgccgtac gtcctcggct ccgcgcacca 840
gagctgcctg ccicccgttcc cggcagacgt gttcatgatt ccgcagtacg gatacctgac 900
tctgaacaat ggcagccaat cggtaggccg ttcctcattc tactgcctgg agtactttcc 960
ttctcagatg ctgagaacgg gcaacaactt caccttiagc tacaccttcg aggacgtgcc 1020
tttccacagc agctacgcgc acagccagag tctgggccgg ctgatgaatc ccctcatcga 1080
ccagtacctg tactacttgg ccagaacaca gagcaacgca ggaggtactg ctggcaatcg 1140
ggaactgcag ttttatcagg gcggacctac caccatggcc gaacaagcaa agaactggct 1200
gcccggacct tgcttccggc aacagagagt atccaagacg ctggatcaaa ataacaacag 1260
caactttgcc tggactggtg ccacaaaata ccatttaaat ggaagaaatt cattggttaa 1320
tcccggtgtc gccatggcaa cccacaagga cgacgaggaa cgcttcttcc cttcgagcgg 1380
agttctaatt tttggcaaaa ctggagcagc taataaaact acattagaaa acgtgctcat 1440
gacaaatgaa gaagaaattc gtcctaccaa cccggtagct accgaggaat acgggattgt 1500
aagcagcaac ttgcaggcgg ctagcaccgc agcccagaca caagttgtta acaaccaggg 1560
agccttacct ggcatggtct ggcagaaccg ggacgtgtac ctgcaaggtc ccatttgggc 1620
caagattcct cacacggacg gcaactttca cccgtctcct ctaatgggtg gctttggact 1680
gaaacacccg cctcccccaga tcctgatcaa aaacacaccg gtacctgcta atcctccaga 1740
agtgtttact ccigccaagt ttgcttcctt catcacgcag tacagcaccg ggcaagtcag 1800
cgttgagatc gaatgggagc tgcagaaaga gaacagcaag cgctggaacc cagagattca 1860
gtacacctcc aactttgaca aacagactgg agtggacttt gctgttgaca gccagggtgt 1920
ttactctgag cct 1933

<210> 52
<211> 1933
<212> DNA
<213> 新AAV血清型,克隆223.6

<400> 52
caaggcctac gaccagcagc tcaaagcggg tgacaatccg tacctgcggt ataaccacgc 60
cgacgccgag tttcaggagc gtcttcaaga agatacgtct tttgggggca acctcgggcg 120
agcagtcttc caggccaaaa agcgggttct cgaacctctt ggtctggttg agacgccagc 180
taagacggca cctggaaaga agcgaccggt agactcgcca gactccacct cgggcatcgg 240
caagaaaggc cagcagcccg cgaaaaagag actcaacttt gggcagactg gcgactcaga 300
gtcagtcccc gaccctcaac caatcggaga accaccagca ggcccctctg gtctgggatc 360

[0051]

tggtacaatg gctgcaggcg gtggcgcacc aatggctgac aatagcgagg gcgccgacgg 420
agtgggtaat gcctcaggaa attggcattg cgattccaca tggctgggcg acagagtcat 480
caccaccagc acccgaacct gggccctgcc cacctacaac aaccacctct acaagcaaat 540
ctccagtcag tcagcaggga gcaccaacga taacgtctat ttcggctaca gcacccctg 600
ggggtatttt gacttcaaca gattccattg ccacttctca ccacgtgact ggcagcgact 660
tatcaacaac aactggggat tccggcccaa gaagctcaac ttcaagctct tcaacatcca 720
ggtcaaggag gtcacgacga atgacggtgt cacaaccatc gctaataacc ttaccagcac 780
ggttcaggtc ttttcggact cggaatatca actgccgtac gtcctcggct ccgcgcacca 840
gggctgcctg cctccgttcc cggcagacgt gttcatgatt ccgcagtacg gatacctgac 900
tctgaacaat ggcagccaat cggtaggccg ttcctccttc tactgcctgg agtactttcc 960
ttctcagatg ctgagaacgg gcaacaactt caccttttagc tacaccttcg aggacgtgcc 1020
tttccacagc agctacgcgc acagccagag tctggaccgg ctgatgaatc cctcatcgca 1080
ccagtacctg tactacttgg ccagaacaca gagcaacgca ggaggtactg ctggcaatcg 1140
ggaactgcag ttttatcagg gcggacctac caccatggcc gaacaagcaa agaactggct 1200
gcccggacct tgcttccggc aacagagagt atccaagacg ctggatcaaa ataacaacag 1260
caactttgcc tggactggtg ccacaaaata ccatttaaat ggaagaaatt cattggttaa 1320
tcccggtgtc gccatggcaa cccacaagga cgacgaggaa cgcttcttcc cttcgagcgg 1380
agttctaatt tttggcaaaa ctggagcagc taataaaact acattagaaa acgtgctcat 1440
gacaaatgaa gaagaaattc gtcctaccaa cccggtagct accgaggaat acgggattgt 1500
aagcagcaac ttgcaggcgg ctagcaccgc agcccagaca caagttgtta acaaccaggg 1560
agccttacct ggcatggtct ggcagaaccg ggacgtgtac ctgcaaggtc ccatttgggc 1620
caagattcct cacacggacg gcaactttca cccgtctcct ctaatgggtg gctttggact 1680
gaaacacccg cctcccagaa tcctgatcaa aaacacaccg gtacctgcta atcctccaga 1740
agtgtttact cctgccaagc ttgcttcctt catcacgcag tacagcaccg ggcaagtcag 1800
cgttgagatc gagtgggagc tgcagaaaga gaacagcaag cgctggaacc cagagattca 1860
gtacacctcc aactttgaca aacagactgg agtggacttt gctgttgaca gccagggtgt 1920
ttactctgag cct 1933

<210> 53
<211> 1933
<212> DNA
[0052] <213> 新 AAV 血清型, 克隆 223.7

<400> 53
caaggcctac gaccagcagc tcaaagcggg tgacaatccg tacctgcggt ataaccacgc 60
cgacgccgag tttcaggagc gtcttcaaga agatacgtct tttgggggca acctcgggcg 120
agcagtcttc caggccaaaa agcgggttct cgaacctctt ggtctggttg agacgccagc 180
tsagacggca cctggaaaga agcgaccggt agactcgcca gactccacct cgggcatcgg 240
caagaaaggc cagcagccgc cgaaaaagag actcaacttt gggcagactg gcgactcaga 300
gtcagtcccc gacctccaac caatcggaga accaccagca ggcccctctg gtctgggatc 360
tggtacaatg gctgcaggcg gtggcgcacc aatggctgac aataacgagg gcgccgacgg 420
agtgggtaat gcctcaggaa attggcattg cgattccaca tggctgggcg acagagtcat 480
caccaccagc acccgaacct gggccctgcc cacctacaac aaccacctct acaagcaaat 540
ctccagtcag tcagcaggga gcaccaacga taacgtctat ttcggctaca gcacccctg 600
ggggtatttt gacttcaaca gattccattg ccacttctca ccacgtgact ggcagcgact 660
tatcaacaac aactggggat tccggcccaa gaagctcaac ttcaagctct tcaacatcca 720
ggtcaaggag gtcacgacga atgacggcgt cacaaccatc gctaataacc ttaccagcac 780
ggttcaggtc ttttcggacc cggaatatca actgccgtac gtcctcggct ccgcgcacca 840
gggctgcctg cctccgttcc cggcagacgt gttcatgatt ccgcagtacg gatacctgac 900
tctgaacaat ggcagccaat cggtaggccg ttcctccttc tactgcctgg agtactttcc 960
ttctcagatg ctgagaacgg gcaacaactt caccttttagc tacaccttcg aggacgtgcc 1020
tttccacagc agctacgcgc acagccagag tctggaccgg ctgatgaatc cctcatcgca 1080
ccagtacctg tactacttgg ccagaacaca gagcaacgca ggaggtactg ctggcaatcg 1140
ggaactgcag ttttatcagg gcggacctac caccatggcc gaacaagcaa agaactggct 1200
gcccggacct tgcttccggc aacagagagt atccaagacg ctggatcaaa ataacaacag 1260
caactttgcc tggactggtg ccacaaaata ccatttaaat ggaagaaatt cattggttaa 1320
tcccggtgtc gccatggcaa cccacaagga cgacgaggaa cgcttcttcc cttcgagcgg 1380
agttctaatt tttggcaaaa ctggagcagc taataaaact acattagaaa acgtgctcat 1440
gacaaatgaa gaagaaattc gtcctaccaa cccggtagct accgaggaat acgggattgt 1500
aagcagcaac ttgcaggcgg ctagcaccgc agcccagaca caagttgtta acaaccaggg 1560
agccttacct ggcatggtct ggcagaaccg ggacgtgtac ctgcaaggtc ccatttgggc 1620
caagattcct cacacggacg gcaactttca cccgtctcct ctaatgggtg gctttggact 1680

gaaacacccg cctccccaga tcctgatcaa aaacacaccg gtacctgcta atcctccaga 1740
agtgtttact cctgccaaga ttgcttcctt catcacgcag tacagcaccg ggcaagtcag 1800
cgttgagatc gagtgggagc tgcagaaaga gaacagcaag cgctggaacc cagagattca 1860
gtacacctcc aactttgaca aacagactgg agtggacttt gctgttgaca gccagggtgt 1920
ttactctgag cct 1933

<210> 54
<211> 3123
<212> DNA
<213> 新 AAV 血清型，克隆 A3.4

<400> 54
gaattcgccc tttctacggc tgcgtcaact ggaccaatga aaactttccc ttcaacgatt 60
gcgtcgacaa gatggtgatc tggtgggagg agggaaagat gaccgccaag gtcgtggaat 120
ctgccaaagc cattctgggt ggaagcaagg ttcgtgtgga ccagaaatgc aagtcttcgg 180
cccagatcga cccgactccg gtgattgtca cctctaacac caacatgtgc gccgtgattg 240
acggaaacte gaccaccttc gagcaccagc agccgttgca agaccggatg ttcaaatttg 300
aacttacccg ccgtttggat catgacttig ggaaggtcac caagcaggaa gtcaaagact 360
ttttcggtig ggctcaagat cacgtgactg aggtggagca tgagttctac gtcaaaaagg 420
gtggagccaa gaaaaggccc gcccccgatg atgtatatat aaatgagccc aagcgggcgc 480
gcgagtcagt tgcgcagcca tcgacgtcag acgcggaagc ttcgataaac tacgcgggca 540
ggtaccaaaa caaatgttct cgtcacgtgg gcatgaatct gatgctgttt ccctgtcgac 600
aatgcgaaag aatgaatcag aattcaaata tctgcttcac acacgggcaa aangactgtt 660
tggaatgctt tcccgtgtca gaatctcaac ccgtttctgt cgtcagnaaa acgtatcaga 720
aactttgtta catteatcat atcatgggaa aagaaccaga cgcctgcact gcctgcgacc 780
tggtaaatgt ggacttggat gactgtattt ctgagcaata aatgacttaa atcaggtatg 840
gctgctgacg gttatcttcc agattggctc gaggacactc tctctgaagg aatcagacag 900
tggtggaagc tcaaacctgg cccaccaccg ccgaaaccta accaacaaca ccgggacgac 960
agtaggggtc ttgtgcttcc tgggtacaag tacctcggac ccttcaacgg actcgacaaa 1020
ggagagccgg tcaacgaggc agacgccgcg gcctcgagca acgacaaagc ctacgaccac 1080
cagctcaagc aaggggacaa cccgtacctc aaatacaacc acgcggacgc tgaatttcag 1140
gagcgtcttc aagaagatac gtctttcggg ggcaacctcg ggcgagcagt cttccaggcc 1200
aaaaagaggg tactcgagcc tcttggtctg gttgaggaag ctgttaagac ggctcctgga 1260
aaaaagagac ctatagagca gtctcctgca gaaccggact cttcttcggg catcggcgaa 1320
tcaggccagc agcccgctaa gaaaagactc aattttggtc agactggcga cacagagtca 1380
gtcccagacc ctcaaccaat cggagaaccc ccgcagccc cctctggtgt gggatctaat 1440
acaatggctt caggcggtgg ggccaccaatg gcagacgata acgaaggcgc cgacggagtg 1500
ggtaattcct cgggaaattg gcattgcgat tccacatgga tgggcgacag agttatcacc 1560
accagcacaa gaacctgggc cctcccccc tacaataatc acctctacaa gcaaatctcc 1620
agcgaatcgg gagccaccaa cgacaaccac tacttcggct acagcacccc ctgggggtat 1680
tttgacttta acagatttca ctgtcacttc tcaccacgtg actggcagcg actcatcaac 1740
aacaactggg gatttagacc caagaaactc aatttcaagc tcttcaacat ccaagtcaag 1800
gaggtcacgc agaatgatgg aacccacgacc ategccaata accttaccag cacggtgcag 1860
gtcttcacag actctgagta ccagctgccc tacgtcctcg gttcggctca ccagggctgc 1920
cttccgccgt tcccagcaga cgtcttcatg attcctcagt acggctactt gactctgaac 1980
aatggcagcc aagcggtagg acgttcttca ttctactgtc tagagtattt tccctctcag 2040
atgctgagga cgggaaacaa cttcaccttc agctacactt ttgaagacgt gccttccac 2100
agcagctacg cgcacagcca gagtctggat cggctgatga atcctctcat tgaccagtac 2160
ctgtattacc tgagcaaaac tcagggtaca agtggaacaa cgcagcaatc gagactgcag 2220
ttcagccaag ctgggcctag ctccatggct cagcaggcca aaaactggct accgggaccc 2280
agctaccgac agcagcgaat gtctaagacg gctaatgaca acaacaacag tgaatttgct 2340
tggactgcag ccaccaaata ttacctgaat ggaagaaatt ctctggtcaa tcccgggccc 2400
ccaatggcca gtcacaagga cgatgaggaa aagtatttcc ccatgcacgg aaatctcatc 2460
tttggaaaac aaggcacagg aactaccaat gtggacattg aatcagtgct tattacagac 2520
gaagaagaaa tcagaacaac taatcctgtg gctacagaac aatacggaca ggttgccacc 2580
aaccatcaga gtcaggacac cacagcttcc tatggaagtg tggacagcca gggaatctta 2640
cctggaatgg tgtggcagga ccgcgatgtc tatcttcaag gtcccatttg ggccaaaact 2700
cctcacacgg acggcaactt tcatcttctt ccgctcatgg gaggcttigg actgaaacac 2760
cctcctcccc agatcctgat caaaaacaca cctgtgccag cgaatcccgc gaccactttc 2820
actcctggaa agtttgcttc gttcattacc cagtatttcca ccggacaggt cagcgtggaa 2880
atagagtggg agctgcagaa agaaaacagc aaacgctgga acccagaaat tcagtacacc 2940
tccaactaca acaagtcggt gaatgtggag tttaccgtgg acgcaaacgg tgtttattct 3000

[0053]

gaaccccgcc ctattggcac tcgttacctt accccggaact tgtaatttcc tgttaatgaa 3060
taaaccgatt tatgcgtttc agttgaactt tggtctctgc gaagggcgaa ttcgcggccg 3120
cta 3123

<210> 55
<211> 3113
<212> DNA
<213> 新AAV血清型,克隆A3.5

<400> 55
gaattcgccc tttctacggc tgcgtcaact ggaccaatga aaactttccc ttcaacgatt 60
gcgtcgacaa gatggtgatc tggtgggagg agggaaagat gaccgccaag gtcgtggaat 120
ctgccaaagc catcttgggt ggaagcaagg ttcgtgtgga ccagaaatgc aagtcttcgg 180
cccagatcga cccgactccg gtgattgtca cctctaacac caacatgtgc gccgtgattg 240
acggaaacte gaccaccttc gagcaccagc agccgttgca agaccggatg ttcaaatttg 300
aacttacccg ccgtttggat catgactttg ggaaggtcac caagcaggaa gtcaaagact 360
ttttccggtg ggctcaagat cacgtgactg aggtggagca tgagttctac gtcaaaaagg 420
gtggagccaa gaaaaggccc gcccccgatg atgtatatat aaatgagccc aagcgggcgc 480
gcgagtcagt tgcgcagcca tcgacgtcag acgcggaagc ttcgataaac tacgcggaca 540
ggtaccaaaa caaatgttct cgtcacgtgg gcatgaatct gatgctgttt ccctgtcgac 600
aatgcgaaag aatgaatcag aattcaaata tctgcttcac acacgggcaa aaagactgtt 660
tggaatgctt tcccgtgtca gaatctcaac ccgttcctgt cgtcagaaaa acgtatcaga 720
aactttgtta catttcatcat atcatgggaa aagtaccaga cgcctgcact gcctgcgacc 780
tggtaaatgt ggacttggat gactgtattt ctgagcaata aatgacttaa atcaggtatg 840
gctgctgacg gttatcttcc agattggctc gaggacactc tctctgaagg aatcagacag 900
tggtggaagc tcaaacctgg cccaccaccg ccgaaaccta accaacaaca ccgggacgac 960
agtaggggtc ttgtgcttcc tgggtacaag tacctcggac ccttcaacgg actcgacaaa 1020
ggagagccgg tcaacgaggc agacgccgcg gccctcgagc acgacaaagc ctacgaccac 1080
cagctcaagc aaggggacaa cccgtacctc aaatacaacc acgcggacgc tgaatttcag 1140
gagcgtctic aagaagatac gtcttcgggg gccaacctcg ggcgagcagt cttccaggcc 1200
aaaaagaggg tactcgagcc tcttggtctg gttgaggaag ctgttaagac ggctcctgga 1260
aaaaagagac ctatagagca gtctcctgca gaaccggact cttcctcggg catcggcaaa 1320
tcaggccagc agcccgctaa gaaaagactc aattttggtc agactggcga cacagagtca 1380
gtcccagacc ctcaaccaat cggagaaccc ccgcagccc cctctggtgt gggatctaat 1440
acaatggctt caggcggtgg ggcaccaatg gcagacaata acgaaggcgc cgacggagtg 1500
ggtaattcct cgggaaattg gcattgcgat tccacatgga tgggcgacag agttatcacc 1560
accagcacaa gaacctgggc cctcccccac tacaataatc acctctacaa gcaaatctcc 1620
agcgaatcgg gagccaccaa cgacaaccac tacttcggct acagcacccc ctgggggtat 1680
tttgacttta acagatccca ctgtcacttc tcaccacgtg actggcagcg actcatcaat 1740
aacaactggg gatttagacc caagaagctc aatttcaagc tcttcaacat ccaagtcaag 1800
gaggtcacgc agaatgatgg aacccacgacc atcgccaata acttaccag caccgtgcag 1860
gtcttcacag actctgagta ccagctgccc tacgtcctcg gttcggctca ccagggctgc 1920
cttccgccgt tcccagcaga cgtcttcatg attcctcagt acggctactt gactctgaac 1980
aatggcagcc aagcggtagg acgttcttca ttctactgtc tagagtattt tcctctcag 2040
atgctgagga cgggaaacaa cttcaccttc agctacactt ttgaagacgt gccttccac 2100
agcagctacg cgcacagccg gagtctggat cggctgatga atcctctcat tgaccagtac 2160
ctgtattacc tgagcaaaac tcagggtaca agtggaacaa cgcagcaatc gagactgcag 2220
ttcaaccaag ctgggcctag ctccatggct cagcaggcca aaaactggct accgggacca 2280
agctaccgac agcagcgaat gtctaagacg gctaatgaca acaacaacag tgaatttgct 2340
tggactgcag ccaccaaata ttaccgaat ggaagaaatt ctctggtcaa tcccgggccc 2400
ccaatggcca gtcacaagga cgatgaggaa aagtatttcc ccatgcacgg aaatctcatc 2460
tttggaaaac aaggcacagg aactaccaat gtggacattg aatcagtgct tattacagac 2520
gaagaagaaa tcagaacgac taatcctgtg gctacagaac aatacggaca ggttgccacc 2580
aaccgtcaga gtcagaacac cacagcttcc tatggaagtg tggacagcca gggaatctta 2640
cctggaatgg tgtggcagga ccgcgatgtc tatcttcaag gtcccatttg ggccaaaact 2700
cctcacacgg acggacactt tcatccttct ccgctcatgg gaggctttgg actgaaacac 2760
cctcctcccc agatcctgat caaaaacaca cctgtgccag cgaatcccgc gaccacttic 2820
actcctggaa agtttgcttc gttcattacc cagtattcca ccggacaggt cagcgtggaa 2880
atagagtggg agctgcagaa agaaaacagc aaacgctgga acccggaaat tcagtacacc 2940
tccaactaca acaagtcggt gaatgtggag tttaccgtgg acgcaaacgg tgtttattct 3000
gaaccccgcc ctattggcac tcgttacctt accccggaact tgtaatttcc tgttaatgaa 3060
taaaccgatt tatgcgtttc agttgaactt tggtctctgc gaagggcgaa ttc 3113

[0054]

<210> 56
<211> 3122
<212> DNA
<213> 新 AAV 血清型, 克隆 A3.7

<400> 56

[0055]

agcggccgcg aattcgccct ttctacggct gcgtcaactg gaccaatgaa aactttcctc 60
tcaacgattg cgtcgacaag atggtgatct ggtgggagga gggaaagatg accgccaagg 120
tcgtggaatc tgccaaagcc attctgggtg gaagcaaggt tcgtgtggac cagaaatgca 180
ggtcttcggc ccagatcgac ccgactccgg tgattgtcac ctctaacacc aacatgtgcg 240
ccgtgattga cggaaactcg accaccttcg agcaccagca gccgttgcaa gaccggatgt 300
tcaaatttga acttacccgc cgtttggatc atgacttttgg gaaggtcacc aagcaggaag 360
tcaaagactt tttccggtgg gctcaagatc acgtgactga ggtggagcat gagttctacg 420
tcaaaaaggg tggagccaag aaaaggcccg ccccccgatga tgtatatata aatgagccca 480
agcgggcgcg cgagtcagtt gcgcagccat cgacgtcaga cgcggaagct tcgataaact 540
acgcggacag gtaccaaaac aaatgttctc gtcacgtggg catgaatctg atgctgtttc 600
cctgtcgaca atgcgaaaga atgaatcaga attcaaatat ctgcttcaca cacgggcaaa 660
aagactgttt ggaatgcttt cccgtgtcag aatctcaacc cgtttctgtc gtcagaaaaa 720
cgtatcagaa actttgttac attcatcata tcatgggaaa agtaccagac gcctgcactg 780
cctgcgacct ggtaaatgtg gacttggatg actgtatttc tgagcaataa atgacttaaa 840
tcaggtatgg ctgctgacgg ttatcttcca gattggctcg aggacactct ctctgaagga 900
atcagacagt ggtggaagct caaacctggc ccaccaccgc cgaaacctaa ccaacaacac 960
cgggacgaca gtaggggtct tgtgcttcct gggtacaagt acctcggacc cttcaacgga 1020
ctcgacaaag gagagccggt caacgaggca gacgccgcgg ccctcgagca cgacaaagcc 1080
tacgaccacc agctcaagca aggggacaac ccgtacctca aatacaacca cgcggacgct 1140
gaatttcagg agcgtcttca agaagatacg tctttcgggg gcaacctcgg gcgagcagtc 1200
ttccaggcca aaaagagggt actcgagcct cttggtctgg ttgaggaagc tgttaagacg 1260
gctcctggaa aaaagagacc tatagagcag tctcctgcag aaccggactc ttcctcgggc 1320
atcggcaaat caggccagca gcccgctaag aaaagactca attttggtca gactggcgac 1380
acagagtcag tccccagacc tcaaccaatc ggagaacccc ccgcagcccc ctctggtgtg 1440
ggatctaata caatggcttc aggcggtggg gcaccaatgg cagacaataa cgaaggcgcc 1500
gacggagtgg gtaattcctc gggaaattgg cattgcgatt ccacatggat gggcgacaga 1560
gttatcacca ccagcacaag aacctgggcc ctccccacct acaataatcg cctctacaag 1620
caaatctcca gcgaatcggg agccaccaac gacaacccact acttcggcta cagcaccccc 1680
tgggggtatt ttgactttaa cagattccac tgtcacttct caccacgtga ctggcagcga 1740
ctcatcaaca acaactgggg atttagaccc aagaaactca atttcaagct cttcaacatc 1800
caagtcaagg aggtcacgca gaatgatgga accacgacca tcgccaataa ccttaccagc 1860
acggtgcagg tctttcacaga ctctgagtac cagctgccct acgtcctcgg ttcggctcac 1920
cagggctgcc ttccgccgtt cccagcagac gtcttcatga ttcctcagta cggctacttg 1980
actctgaaca atggcagcca agcggtagga cgttcttcat tctactgtct agagtatttt 2040
ccctctcaga tgctgaggac gggaaacaac ttcaccttca gctacacttt tgaagacgtg 2100
cctttccaca gcagctacgc gcacagccag agtctggatc ggctgatgaa tcctctcatt 2160
gaccagtacc tgtattacct gagcaaaact cagggtacaa gtggaacaac gcagcaatcg 2220
agactgcagt tcagccaagc tgggcctagc tccatggctc agcaggccaa aaactggcta 2280
ccgggaccca gctaccgaca gcagcgaatg tctaagacgg ctaatgacaa caacaacagt 2340
gaatttgctt ggactgcagc caccaaatat tacctgaatg gaagaaattc tctggtcaat 2400
cccgggcccc caatggccag tcacaaggac gatgaggaaa agtatttccc catgcacgga 2460
aatctcatct ttggaaaaca aggcacagga actaccaatg tggacattga atcagtgctt 2520
attacagacg aagaagaaat cagaacaact aatccigtgg ctacagaaca atacggacag 2580
gttgccacca accatcagag tcagaacacc acagcttcct atggaagtgt ggacagccag 2640
ggaatcttac ctggaatggt gtggcaggac cgcgatgtct atcttcaagg tcccatttgg 2700
gccaaaactc ctcacacgga cggacacttt catccttctc cgctcatggg aggctttgga 2760
ctgaaacacc ctcctcccca gatcctgatc aaaaacacac ctgtgccagc gaatcccgcg 2820
accactttca ctcctggaaa gtttgcttcg ttcattaccc agtattccac cggacaggtc 2880
agcgtggaaa tagagtggga gctgcagaaa gaaaacagca aacgctggaa ccccagaaatt 2940
cagtacacct ccaactacaa caagtcggtg aatgtggagt ttaccgtgga cgcaaacggt 3000
gtttattctg aacccccgcc tattggcact cgttacctta ccccggaactt gtaatttcct 3060
gttaatgaat aaaccgattt atgcgtttca gttgaacttt ggtctctgcg aagggcgaat 3120
tc 3122

<210> 57
<211> 3123
<212> DNA
<213> 新 AAV 血清型，克隆 A3.3

<400> 57

[0056]

gaattcgccc tttctacggc tgcgtcaact ggaccaatga aaactttccc ttcaacgatt 60
gcgtcgacaa gatggtgatc tggtgggagg agggaaagat gaccgccaag gtcgtggaat 120
ctgccaaagc cattctgggt ggaggcaagg ttcgtgtgga ccagaaatgc aagtcttcgg 180
cccagatcga cccgactccg gtgattgtca cctctaacac caacatgtgc gccgtgattg 240
acggaaactc gaccaccttc gagcaccagc agccgttgca agacccggat gttcaaattg 300
aacttacccg ccgtttggat catgactttg ggaaggtcac caagcaggaa gtcaaagact 360
ttttccggtg ggctcaagat cacgtgactg aggtggagca tgagttctac gtcaaaaagg 420
gtggagccaa gaaaaggccc gcccccgatg atgtatatat aaatgagccc aagcgggcgc 480
gcgagtcagt tgcgcagcca tcgacgtcag acgcggaagc ttcgataaac tacgcggaca 540
ggtaccaaaa caaatgttct cgtcacgtgg gcatgaatct gatgctgttt ccctgtcgac 600
aatgcgaaag aatgaatcag aattcaaata tctgcttcac acacgggcaa aaagactgtt 660
tggaatgctt tcccgtgtca gaatctcaac ccgtttctgt cgtcagaaaa acgtatcaga 720
aactttgtta cattcatcat atcatgggaa aagtaccaga cgcctgcact gcctgcgacc 780
tggtaaatgt ggacttggat gactgtattt ctgagcaata aatgacttaa atcaggtatg 840
gctgctgacg gttatcttcc agattggctc gaggacactc tctctgaagg aatcagacag 900
tggtggaagc tcaaacctgg ccaccaccgc ccgaaaccta accaacaaca ccgggacgac 960
agtaggggtc ttgtgcttcc tgggtacaag tacctcggac ccttcaacgg actcgacaaa 1020
ggagagccgg tcaacgaggc agacgccgcg gccctcgagc acgacaaagc ctacgaccac 1080
cagctcaagc aaggggacaa cccgtacctc aaatacaacc acgcggacgc tgaatttcag 1140
gagcgtcttc aagaagatac gtctttcggg ggcaacctcg ggcgagcagt cttccaggcc 1200
aaaaagaggg tactcgagcc tcttggtctg gttgaggaag ctgttaagac ggctcctgga 1260
aaaaagagac ctatagagca gtctcctgca gaaccggact cttcttcggg catcggcaaa 1320
tcaggccagc agcccgctaa gaaaagactc aattttggtc agactggcga cacagagtca 1380
gtcccaggcc ctcaaccaat cggagaaccc cccgcagccc cctctggtgt gggatctaat 1440
acaatggctt caggcggtgg ggcaccaatg gcagacaata acgaaggcgc cgacggagtg 1500
ggtaattcct cgggaaattg gcattgcgat tccacatgga tgggcgacag agttatcacc 1560
accagcacaa gaacctgggc cctcccctcc tacaataatc acctctacaa gcaaaatctcc 1620
agcgaatcgg gagccaccaa cgacaaccac tacttcggct acagcacccc ctgggggtat 1680
tttgacttta acagatccca ctgtcacttc tcaccacgtg actggcagcg actcatcaac 1740
aacaactggg gatttagacc caagaaactc aatttcaagc tcttcaacat ccaagtcaag 1800
gaggtcacgc agaatgatgg aaccacgacc atcgccaata accttaccag cgcggtgcag 1860
gtcttcacag actctgagta ccagctgccc tacgtcctcg gttcggctca ccagggctgc 1920
cttccgccgt tcccagcaga cgtcttcatg attcctcagt acggctactt gactctgaac 1980
aatggcagcc aagcggtagg acgttcttca ttctactgtc tagagtattt tcctctcag 2040
atgctgagga cgggaaacaa cttcaccttc agctacactt ttgaagacgt gccttttcac 2100
agcagctacg cgcacagcca gagtctggat cggctgatga atcctctcat tgaccagtac 2160
ctgtattacc tgagcaaaac tcagggtaca agtggaacaa cgcagcaatc gagactgcag 2220
ttcagccaag ctgggcctag ctccatggct cagcaggcca aaaactggct accgggaccc 2280
agctaccgac agcagcgaat gtctaagacg gctaatgaca acaacaacag tgaattgct 2340
tggactgcag ccaccaaata ttacctgaat ggaagaaatt ctctggtcaa tcccgggccc 2400
ccagtggcca gtcacaagga cgatgaggaa aagtatttcc ccatgcacgg aaatctcatc 2460
tttggaaaac aaggcacagg aactaccaat gtggacattg aatcagtgct tattacagac 2520
gaagaagaaa tcagaacaac taatcctgtg gctacagaac aatcggaca ggttgccacc 2580
aaccatcaga gtcagaacac cacagcttcc tatggaagtg tggacagcca gggaatctta 2640
cctggaatgg tgtggcagga ccgcgatgtc tatcttcaag gtcccattg ggccaaaact 2700
cctcacacgg acggacactt tcatcctct ccgctcatgg gaggctttgg actgaaacac 2760
cctcctcccc agatcctgat caaaaacaca cctgtgccag cgaatccccc gaccactttc 2820
actcctggaa agtttgcttc gttcattacc cagtattcca cctgacaggt cagcgtggaa 2880
atagagtggg agctgcagaa agaaaacagc aaacgctgga acccagaaat tcagtacacc 2940
tccaactaca acaagtcggt gaatgtggag tttaccgtgg acgcaaacgg tgtttattct 3000
gaacccccgc ctattggcac tcgttacctt acccggaacc tgtaatttcc tgttaatgaa 3060
taagccgatt tatgcgtttc agttgaactt tggtctctgc gaagggcgaa ttcgtttaaa 3120
cct 3123

<210> 58
<211> 2969

<212> DNA
<213> 42.12

<400> 58
gaattcgccc tttctacggc tgcgtcaact ggaccaatga gaactttccc ttcaacgatt 60
gcgtcgacaa gatggtgatc tggtgggagg agggcaagat gacggcccaag gtcgtggagt 120
ccgccaaggc catctcggc ggcagcaagg tgcgcgtgga ccaaaagtgc aagtcgtccg 180
cccagatcga ccccaccccc gtgatcgtca ctctccaacac caacatgtgc gccgtgattg 240
acgggaacag caccaccttc gagcaccagc agccgttaca agaccggatg ttcaaatttg 300
aactcacccg ccgtctggag cacgactttg gcaaggtgac aaagcaggaa gtcaaagagt 360
tctttcgctg ggcgcaggat cacgtgaccg aggtggcgca tgagttctac gtcagaaagg 420
gtggagccaa caagagaccc gcccccgatg acgcggataa aagcgagccc aagcgggcct 480
gcccctcagt cgcggatcca tcgacgtcag acgcggaagg agctccggtg gactttgccg 540
acaggtacca aaacaaatgt tctcgtcacg cgggcatgct tcagatgctg tttcctgtca 600
agacatgcga gagaatgaat cagaattica acattigctt cacgcacggg accagagact 660
gttcagaatg tttccccggc gtgtcagaat ctcaaccggt cgtcagaaag aggacgtatc 720
ggaaactctg tgccattcat catctgctgg ggcgggctcc cgagattgct tgctcggcct 780
gcgatctggt caacgtggac ctggatgact gtgtttctga gcaataaatg acttaaacca 840
ggtatggctg ccgatggtta tcttccagat tggctcgagg acaacctctc tgagggcatc 900
cgcgagtggt gggacttgaa acctggagcc ccgaaaccca aagccaacca gcaaaagcag 960
gacgacggcc ggggtctggt gcttcctggc tacaagtacc tcggaccctt caacggactc 1020
gacaagggag agccggtcaa cgaggcagac gccgcggccc tcgagcacga caaggcctac 1080
gacaagcagc tcgagcaggg ggacaacccg tacctcaagt acaaccacgc cgacgccgag 1140
tttcaggagc gtcttcaaga agatacgtct tttgggggca acctcgggcg agcagtcttc 1200
caggccaaga agcgggttct cgaacctctc ggtctggttg aggaaggcgc taagacggct 1260
cctggaaaga agagaccggt agagccatca ccccagcgtt ctccagactc ctctacgggc 1320
atcggcaaga caggccagca gccgcgcgaaa aagagactca actttgggca gactggcgac 1380
tcagagtcag tgccccgaccc tcaaccaatc ggagaacccc ccgcaggccc ctctggtctg 1440
ggatctggta caatggctgc aggcggtggc gctccaatgg cagacaataa cgaaggcgcc 1500
gacggagtgg gtagttcctc aggaaattgg cattgcgatt ccacatggct gggcgacaga 1560
gtcatcacca ccagcacccg aacctgggcc ctcccaacct acaacaacca cctctacaag 1620
caaatctcca acgggacatc gggaggaagc accaacgaca acacctactt cggctacagc 1680
accccctggg ggtattttga ctttaacaga ttccactgcc acttctcacc acgtgactgg 1740
cagcgactca tcaacaacaa ctgcgggattc cggccccaaga gactcaactt caagctcttc 1800
aacatccagg tcaaggaggt cacgcagaat gaaggcacca agaccatcgc caataacctt 1860
accagcacga ttcaggtctt tacggactcg gaataccagc tccctacgt cctcggctct 1920
gcgcaccagg gctgcctgcc tccgttcccg gcggacgtct tcatgattcc tcagtacggg 1980
tacctgactc tgaacaacgg cagtcaggcc gtgggccgtt cctcttcta ctgcctggag 2040
tactttcctt ctcaaatgct gagaacgggc aacaacttig agttcagcta ccagtttgag 2100
gacgtgcctt ttcacagcag ctacgcgcac agccaaagcc tggaccggct gacgaacccc 2160
atcatcgacc agtacctgta ctacctggcc cggacccaga gcactacggg gtccacaagg 2220
gggctgcagt tccatcaggc tgggcccaac accatggccg agcaatcaaa gaactggctg 2280
cccggaccct gttatcggca gcagagactg tcaaaaaaca tagacagcaa caacaacagt 2340
aactttgcct ggaccggggc cactaaatac catctgaatg gtagaaattc attaaccaac 2400
ccgggcgtag ccatggccac caacaaggac gacgaggacc agttctttcc catcaacgga 2460
gtgctggttt ttggcaaaac gggggctgcc aacaagacaa cgctggaaaa cgtgctaatg 2520
accagcgagg aggagatcaa aaccaccaat cccgtggcta cagaagaata cggtgtggtc 2580
tccagcaacc tgcaatcgtc tacggccgga ccccagacac agactgtcaa cagccagggg 2640
gctctgcccg gcatggtctg gcagaaccgg gacgtgtacc tgcagggtcc catctgggcc 2700
aaaattcctc acacggacgg caacttccac ccgtctcccc tgatgggcgg atttggactc 2760
aaacaccccc ctcctcaaat tctcatcaag tatacttcca actactacaa atctacaaat 2820
gtggactttg ctgtcaatac tgagggtact tattcagagc ctcgccccat tggcacccgt 2880
tacctcaccc gtaacctgta attgcctgtt aatcaataaa ccggttaatt cgtttcagtt 2940
gaactttggt ctctgcgaag ggcgaattc 2969

<210> 59
<211> 3129
<212> DNA
<213> 44.2

<400> 59
gaattcgccc tttctacggc tgcgtcaact ggaccaatga gaactttccc ttcaacgatt 60

[0057]

[0058]

```
gcgtcgacaa gatggtgatc tggtgggagg agggcaagat gacggccaag gtcgtggagt    120
ccgccaaggc catctcggcc ggcagcaaag tgcgcgtgga ccaaaagtgc aagtcgtccg    180
cccagatcga ccccacccca gtgatcgtca cctccaacac caacatgtgc gccgtgattg    240
acgggaacag caccacctic gagcaccagc agccgttgca ggaccggatg ttcaagtttg    300
aactcacccg ccgtctggag cacgactttg gcaaggtgac aaagcaggaa gtcagagagt    360
tcttccgctg ggcgcaggat cacgtgaccg aggtggcgca cgagttctac gtcatgaagg    420
gtggagccaa caagagaccc gccccgatg acgcggataa aagcgagccc aagcgggcct    480
gccccicagt cgcggatcca tcgacgtcag acgcggaagg agctccggtg gactttgccg    540
acaggtacca aaacaaatgt tctcgtcacg cgggcatgct tcagatgctg tttccctgca    600
aaacatgcga gagaatgaat cagaatttca acatttgctt cacgcacggg accagagact    660
gttcagaatg tttccccggc gtgtcagaat ctcaaccggt cgtcagaaaa gagacgtatc    720
ggaaactctg tgcgattcat catctgctgg gggcgggcac ccgagattgc ttgctcggcc    780
tgcgatctgg tcaacgtgga cctagatgac tgtgtttctg agcaataaat gacttaaacc    840
aggtatggct gccgatggtt atcttccaga ttggctcgag gacaacctct ctgagggcat    900
tcgcgagtgg tgggacttga aacctggagc cccgaaaccc aaagccaacc agcaaaagca    960
ggacgacggc cggggtctgg tgcttcctgg ctacaagtac ctcggacccc tcaacggact   1020
cgacaagggg gagccccgtca acgcggcgga cgcagcggcc ctcgagcacg acaaggccta   1080
cgaccagcag ctcaaagcgg gtgacaatcc gtacctgcgg tataaccacg ccgacgccga   1140
gtttcaggag cgtctgcaag aagatacgtc ttttgggggc aacctcgggc gagcagtctt   1200
ccaggccaag aagcgggttc tcgaacctct cggtctggtt gaggaaggcg ctaagacggc   1260
tcctggaaag aagagaccgg tagagccatc accccagcgt tctccagact cctctacggg   1320
catcggcaag aaaggccagc agccccgcgaa aaagagactc aactttgggc agactggcga   1380
ctcagagtca gtgccngacc ctcaaccaat cggagaaccc ccgcaggcc cctctggtct   1440
gggatctggt acaatggctg caggcggtgg cgctccaatg gcagacaata acgaaggcgc   1500
cgacggagtg ggtagttcct caggaaattg gcattgcgat tccacatggc tgggcgacag   1560
agtcatcacc accagcaccc gaacctgggc cctccccacc tacaacaacc acctctacaa   1620
gcaaatctcc aacgggactt cgggaggaag caccaacgac aacacctact tcggctacag   1680
caccccctgg gggtatttig actttaacag attccactgc cacttctcac cacgtgactg   1740
gcagcgacic atcaacaaca actggggatt ccggcccaag agactcaact tcaagctctt   1800
caacatccag gtcaaggagg tcacgcagaa tgaaggcacc aagaccatcg ccaataacct   1860
taccagcacg attcaggtct ttacggactc ggaataccag ctcccgtacg tcctcggctc   1920
tgcgcaccag ggctgcctgc ctccgttccc ggcggacgtc ttcatgattc ctcagtacgg   1980
gtacctgact ctgaacaatg gcagtcaggc cgtgggccgt tcctccttct actgcctgga   2040
gtactttcct tctcaaatgc tgagaacggg caacaacttt gagttcagct accagtttga   2100
ggacgtgcct tttcacagca gctacgcgca cagccaaagc ctggaccggc tgatgaaccc   2160
cctcatcgac cagtacctgt actacctgtc tcggactcag tccacgggag gtaccgcagg   2220
aactcagcag ttgctatttt ctcaggccgg gcctaataac atgtcggctc aggccaaaaa   2280
ctggctaccc gggccctgct accggcagca acgcgtctcc acgacactgt cgcaaaataa   2340
caacagcaac tttgcctgga ccggtgccac caagtatcat ctgaatggca gagactctct   2400
ggtaaatccc ggtgtgctgt tggcaaccca caaggacgac gaagagcgat ttttccgtc    2460
cagcggagtc ttaatgtttg ggaaacaggg agctggaaaa gacaacgtgg actatagcag   2520
cgttatgcta accagtgagg aagaaattaa aaccaccaac ccagtggcca cagaacagta   2580
cggcgtggtg gccgataacc tgcaacagca aaacgccgct cctattgtag gggccgtcaa   2640
cagtcaagga gccttaccig gcatggtctg gcagaaccgg gacgtgtacc tgcagggtcc   2700
tatctgggcc aagattcctc acacggacgg aaactttcat ccctcgccgc tgatgggagg   2760
ctttggactg aaacaccccc ctcctcagat cctgattaag aatacacctg ttcccgcgga   2820
tccttccaact accttcagtc aagctaagct ggcgtcgttc atcacgcagt acagcaccgg   2880
acaggtcagc gtggaaattg aatgggagct gcagaaagaa aacagcaaac gctggaaccc   2940
agagattcaa tacacttcca actactacaa atctacaaat gtggactttg ctgttaacac   3000
agatggcact tattctgagc ctcgccccat cggcacccgt tacctcaccc gtaatctgta   3060
attgcttgtt aatcaataaa ccggttgatt cgtttcagtt gaactttggt ctctgcgaag   3120
ggcgaattc                                                          3129
```

<210> 60
<211> 733
<212> PRT
<213> AAV 血清型的衣壳蛋白, 克隆 CIVP1

<400> 60

```
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 Asp Leu Lys Pro Gly Ala Pro Lys Pro
```

[0059]

```
            20                  25                  30
Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro
        35                  40                  45
Gly Tyr Lys Tyr Leu Gly Pro Phe 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 Arg Tyr Asn His Ala
                85                  90                  95
Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly
            100                 105                 110
Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
        115                 120                 125
Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg
    130                 135                 140
Pro Leu Glu Ser Pro Gln Glu Pro Asp Ser Ser Ser Gly Ile Gly Lys
145                 150                 155                 160
Lys Gly Lys Gln Pro Ala Lys Lys Arg Leu Asn Phe Glu Glu Asp Thr
                165                 170                 175
Gly Ala Gly Asp Gly Pro Pro Glu Gly Ser Asp Thr Ser Ala Met Ser
            180                 185                 190
Ser Asp Ile Glu Met Arg Ala Ala Pro Gly Gly Asn Ala Val Asp Ala
        195                 200                 205
Gly Gln Gly Ser Asp Gly Val Gly Asn Ala Ser Gly Asp Trp His Cys
    210                 215                 220
Asp Ser Thr Trp Ser Glu Gly Lys Val Thr Thr Thr Ser Thr Arg Thr
225                 230                 235                 240
Trp Val Leu Pro Thr Tyr Asn Asn His Leu Tyr Leu Arg Leu Gly Thr
                245                 250                 255
Thr Ser Asn Ser Asn Thr Tyr Asn Gly Phe Ser Thr Pro Trp Gly Tyr
            260                 265                 270
Phe Asp Phe Asn Arg Phe His Cys His Phe Ser Pro Arg Asp Trp Gln
        275                 280                 285
Arg Leu Ile Asn Asn Asn Trp Gly Leu Arg Pro Lys Ala Met Arg Val
    290                 295                 300
Lys Ile Phe Asn Ile Gln Val Lys Glu Val Thr Thr Ser Asn Gly Glu
305                 310                 315                 320
Thr Thr Val Ala Asn Asn Leu Thr Ser Thr Val Gln Ile Phe Ala Asp
                325                 330                 335
Ser Ser Tyr Glu Leu Pro Tyr Val Met Asp Ala Gly Gln Glu Gly Ser
            340                 345                 350
Leu Ser Pro Phe Pro Asn Asp Val Phe Met Val Pro Gln Tyr Gly Tyr
        355                 360                 365
Cys Gly Ile Val Thr Gly Glu Asn Gln Asn Gln Thr Asp Arg Asn Ala
    370                 375                 380
Phe Tyr Cys Leu Glu Tyr Phe Pro Ser Gln Met Leu Arg Thr Gly Asn
385                 390                 395                 400
Asn Phe Glu Met Ala Tyr Asn Phe Gly Lys Val Pro Phe His Ser Met
                405                 410                 415
Tyr Ala Tyr Ser Gln Ser Pro Asp Arg Leu Met Asn Pro Leu Leu Asp
            420                 425                 430
Gln Tyr Leu Trp His Leu Gln Ser Thr Thr Ser Gly Glu Thr Leu Asn
        435                 440                 445
Gln Gly Asn Ala Ala Thr Thr Phe Gly Lys Ile Arg Ser Gly Asp Phe
    450                 455                 460
Ala Phe Tyr Arg Lys Asn Trp Leu Pro Gly Pro Cys Val Lys Gln Gln
465                 470                 475                 480
Arg Leu Ser Lys Thr Ala Ser Gln Asn Tyr Lys Ile Pro Ala Ser Gly
                485                 490                 495
Gly Asn Ala Leu Leu Lys Tyr Asp Thr His Tyr Thr Leu Asn Asn Arg
            500                 505                 510
Trp Ser Asn Ile Ala Pro Gly Pro Pro Met Ala Thr Ala Gly Pro Ser
```

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

[0060]

<210> 61
<211> 733
<212> PRT
<213> AAV 血清型的衣壳蛋白,克隆 C2VP1

```
<400> 61
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 Asp Leu Lys Pro Gly Ala Pro Lys Leu
            20                  25                  30
Lys Ala Asn Gln Gln Lys Glu Asp Asp Gly Arg Gly Leu Val Leu Pro
        35                  40                  45
Gly Tyr Lys Tyr Leu Gly Pro Phe His 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 Arg Tyr Asn His Ala
                85                  90                  95
Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly
            100                 105                 110
Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
        115                 120                 125
Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg
    130                 135                 140
Pro Leu Glu Ser Pro Gln Glu Pro Asp Ser Ser Ser Gly Ile Gly Lys
145                 150                 155                 160
Lys Gly Lys Gln Pro Ala Lys Lys Arg Leu Asn Phe Glu Glu Asp Thr
                165                 170                 175
Gly Ala Gly Asp Gly Pro Pro Glu Gly Ser Asp Thr Ser Ala Met Ser
            180                 185                 190
Ser Asp Ile Glu Met Arg Ala Ala Pro Gly Gly Asn Ala Val Asp Ala
        195                 200                 205
Gly Gln Gly Ser Asp Gly Val Gly Asn Ala Ser Gly Asp Trp His Cys
    210                 215                 220
```

Asp Ser Thr Trp Ser Glu Gly Lys Val Thr Thr Thr Ser Thr Arg Thr
225 230 235 240
Trp Val Leu Pro Thr Tyr Asn Asn His Leu Tyr Leu Arg Leu Gly Thr
245 250 255
Thr Ser Asn Ser Asn Thr Tyr Asn Gly Phe Ser Thr Pro Trp Gly Tyr
260 265 270
Phe Asp Phe Asn Arg Phe His Cys His Phe Ser Pro Arg Asp Trp Gln
275 280 285
Arg Leu Ile Asn Asn Asn Trp Gly Leu Arg Pro Lys Ala Met Arg Val
290 295 300
Lys Ile Phe Asn Ile Gln Val Lys Glu Val Thr Thr Ser Asn Gly Glu
305 310 315 320
Thr Thr Val Ala Asn Asn Leu Thr Ser Thr Val Gln Ile Phe Ala Asp
325 330 335
Ser Ser Tyr Glu Leu Pro Tyr Val Met Asp Ala Gly Gln Glu Gly Ser
340 345 350
Leu Pro Pro Phe Pro Asn Asp Val Phe Met Val Pro Gln Tyr Gly Tyr
355 360 365
Cys Gly Ile Val Thr Gly Glu Asn Gln Asn Gln Thr Asp Arg Asn Ala
370 375 380
Phe Tyr Cys Leu Glu Tyr Phe Pro Ser Gln Met Leu Arg Thr Gly Asn
385 390 395 400
Asn Phe Glu Met Ala Tyr Asn Phe Glu Lys Val Pro Phe His Ser Met
405 410 415
Tyr Ala His Ser Gln Ser Leu Asp Arg Leu Met Asn Pro Leu Leu Asp
420 425 430
Gln Tyr Leu Trp His Leu Gln Ser Thr Thr Ser Gly Glu Thr Leu Asn
435 440 445
Gln Gly Asn Ala Ala Thr Thr Phe Gly Lys Ile Arg Ser Gly Asp Phe
450 455 460
[0061] Ala Phe Tyr Arg Lys Asn Trp Leu Pro Gly Pro Cys Val Lys Gln Gln
465 470 475 480
Arg Phe Ser Lys Thr Ala Ser Gln Asn Tyr Lys Ile Pro Ala Ser Gly
485 490 495
Gly Asn Ala Leu Leu Lys Tyr Asp Thr His Tyr Thr Leu Asn Asn Arg
500 505 510
Trp Ser Asn Ile Ala Pro Gly Pro Pro Met Ala Thr Ala Gly Pro Ser
515 520 525
Asp Gly Asp Phe Ser Asn Ala Gln Leu Ile Phe Pro Gly Pro Ser Val
530 535 540
Thr Gly Asn Thr Thr Thr Ser Ala Asn Asn Leu Leu Phe Thr Ser Glu
545 550 555 560
Gly Glu Ile Ala Ala Thr Asn Pro Arg Asp Thr Asp Met Phe Gly Gln
565 570 575
Ile Ala Asp Asn Asn Gln Asn Ala Thr Thr Ala Pro Ile Thr Gly Asn
580 585 590
Val Thr Ala Met Gly Val Leu Pro Gly Met Val Trp Gln Asn Arg Asp
595 600 605
Ile Tyr Tyr Gln Gly Pro Ile Trp Ala Lys Ile Pro His Ala Asp Gly
610 615 620
His Phe His Pro Ser Pro Leu Ile Gly Gly Phe Gly Leu Lys His Pro
625 630 635 640
Pro Pro Gln Ile Phe Ile Lys Asn Thr Pro Val Pro Ala Asn Pro Ala
645 650 655
Thr Thr Phe Thr Ala Ala Arg Val Asp Ser Phe Ile Thr Gln Tyr Ser
660 665 670
Thr Gly Gln Val Ala Val Gln Ile Glu Trp Glu Ile Glu Lys Glu Arg
675 680 685
Ser Lys Arg Arg Asn Pro Glu Val Gln Phe Thr Ser Asn Tyr Gly Asn
690 695 700
Gln Ser Ser Met Leu Trp Ala Pro Asp Thr Thr Gly Lys Tyr Thr Glu
705 710 715 720

Pro Arg Val Ile Gly Ser Arg Tyr Leu Thr Asn His Leu
725 730

<210> 62
<211> 733
<212> PRT
<213> AAV 血清型的衣壳蛋白, 克隆 C5VP1@2

<400> 62
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 Asp Leu Lys Pro Gly Ala Pro Lys Pro
20 25 30
Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro
35 40 45
Gly Tyr Glu Tyr Leu Gly Pro Phe 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 Arg Tyr Asn His Ala
85 90 95
Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly
100 105 110
Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
115 120 125
Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg
130 135 140
Pro Leu Glu Ser Pro Gln Glu Pro Asp Ser Ser Ser Gly Ile Gly Lys
145 150 155 160
Lys Gly Lys Gln Pro Ala Lys Lys Arg Leu Asn Phe Glu Glu Asp Thr
165 170 175
[0062] Gly Ala Gly Asp Gly Pro Pro Glu Gly Ser Asp Thr Ser Ala Met Ser
180 185 190
Ser Asp Ile Glu Met Arg Ala Ala Pro Gly Gly Asn Ala Val Asp Ala
195 200 205
Gly Gln Gly Ser Asp Gly Val Gly Asn Ala Ser Gly Asp Trp His Cys
210 215 220
Asp Ser Thr Trp Ser Glu Gly Lys Val Thr Thr Thr Ser Thr Arg Thr
225 230 235 240
Trp Val Leu Pro Thr Tyr Asn Asn His Leu Tyr Leu Arg Leu Gly Thr
245 250 255
Thr Ser Asn Ser Asn Thr Tyr Asn Gly Phe Ser Thr Pro Trp Gly Tyr
260 265 270
Phe Asp Phe Asn Arg Phe His Cys His Phe Ser Pro Arg Asp Trp Gln
275 280 285
Arg Leu Ile Asn Asn Asn Trp Gly Leu Arg Pro Lys Ala Met Arg Val
290 295 300
Lys Ile Phe Asn Ile Gln Val Lys Glu Val Thr Thr Ser Asn Gly Glu
305 310 315 320
Thr Thr Val Ala Asn Asn Leu Thr Ser Thr Val Gln Ile Phe Ala Asp
325 330 335
Ser Ser Tyr Glu Leu Pro Tyr Val Met Asp Ala Gly Gln Glu Gly Ser
340 345 350
Leu Pro Pro Phe Pro Asn Asp Val Phe Met Val Pro Gln Tyr Gly Tyr
355 360 365
Cys Gly Ile Val Thr Gly Glu Asn Gln Asn Gln Thr Asp Arg Asn Ala
370 375 380
Phe Tyr Cys Leu Glu Tyr Phe Pro Ser Gln Met Leu Arg Thr Gly Asn
385 390 395 400
Asn Phe Glu Thr Ala Tyr Asn Phe Glu Lys Val Pro Phe His Ser Met
405 410 415
Tyr Ala His Ser Gln Ser Leu Asp Gly Leu Met Asn Pro Leu Leu Asp

420 425 430
Gln Tyr Leu Trp His Leu Gln Ser Thr Thr Ser Gly Glu Thr Leu Asn
435 440 445
Gln Gly Asn Ala Ala Thr Thr Phe Gly Lys Ile Arg Ser Gly Asp Phe
450 455 460
Ala Phe Tyr Arg Lys Asn Trp Leu Pro Gly Pro Cys Val Lys Gln Gln
465 470 475 480
Arg Phe Ser Lys Thr Ala Ser Gln Asn Tyr Lys Ile Pro Ala Ser Gly
485 490 495
Gly Asn Ala Leu Leu Lys Tyr Asp Thr His Tyr Thr Leu Asn Asn Arg
500 505 510
Trp Ser Asn Ile Ala Pro Gly Pro Pro Met Ala Thr Ala Gly Pro Ser
515 520 525
Asp Gly Asp Phe Ser Asn Ala Gln Leu Ile Phe Pro Gly Pro Ser Val
530 535 540
Thr Gly Asn Thr Thr Thr Ser Ala Asn Asn Leu Leu Phe Thr Ser Glu
545 550 555 560
Glu Glu Ile Ala Ala Thr Asn Pro Arg Asp Thr Asp Met Phe Gly Gln
565 570 575
Ile Ala Asp Asn Asn Gln Asn Ala Thr Thr Ala Pro Ile Thr Gly Asn
580 585 590
Val Thr Ala Met Gly Val Leu Pro Gly Met Val Trp Gln Asn Arg Asp
595 600 605
Ile Tyr Tyr Gln Gly Pro Ile Trp Ala Lys Ile Pro His Ala Asp Gly
610 615 620
His Phe His Pro Ser Pro Leu Ile Gly Gly Phe Gly Leu Lys His Pro
625 630 635 640
Pro Pro Gln Ile Phe Ile Lys Asn Thr Pro Val Pro Ala Tyr Pro Ala
645 650 655
Thr Thr Phe Thr Ala Ala Arg Val Asp Ser Phe Ile Thr Gln Tyr Ser
660 665 670
[0063] Thr Gly Gln Val Ala Val Gln Ile Glu Trp Glu Ile Glu Lys Glu Arg
675 680 685
Ser Lys Arg Trp Asn Pro Glu Val Gln Phe Thr Ser Asn Cys Gly Asn
690 695 700
Gln Ser Ser Met Leu Trp Ala Pro Asp Thr Thr Gly Lys Tyr Thr Glu
705 710 715 720
Pro Arg Val Ile Gly Ser Arg Tyr Leu Thr Asn His Leu
725 730

<210> 63
<211> 734
<212> PRT
<213> AAV 血清型的衣壳蛋白, 克隆 AAV4VP1

<400> 63
Met Thr Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser Glu
1 5 10 15
Gly Val Arg Glu Trp Trp Ala Leu Gln Pro Gly Ala Pro Lys Pro Lys
20 25 30
Ala Asn Gln Gln His Glu Asp Asn Ala Arg Gly Leu Val Leu Pro Gly
35 40 45
Tyr Lys Tyr Leu Gly Pro Gly Asn Gly Leu Asp Lys Gly Glu Pro Val
50 55 60
Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp Gln
65 70 75 80
Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala Asp
85 90 95
Ala Glu Phe Gln Gln Arg Leu Gln Gly Asp Thr Ser Phe Gly Gly Asn
100 105 110
Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro Leu
115 120 125

[0064]

```
Gly Leu Val Glu Gln Ala Gly Glu Thr Ala Pro Gly Lys Lys Arg Pro
    130                 135                 140
Leu Ile Glu Ser Pro Gln Gln Pro Asp Ser Ser Thr Gly Ile Gly Lys
145                 150                 155                 160
Lys Gly Lys Gln Pro Ala Lys Lys Lys Leu Val Phe Glu Asp Glu Thr
                165                 170                 175
Gly Ala Gly Asp Gly Pro Pro Glu Gly Ser Thr Ser Gly Ala Met Ser
            180                 185                 190
Asp Asp Ser Glu Met Arg Ala Ala Ala Gly Gly Ala Ala Val Glu Gly
        195                 200                 205
Gly Gln Gly Ala Asp Gly Val Gly Asn Ala Ser Gly Asp Trp His Cys
    210                 215                 220
Asp Ser Thr Trp Ser Glu Gly His Val Thr Thr Thr Ser Thr Arg Thr
225                 230                 235                 240
Trp Val Leu Pro Thr Tyr Asn Asn His Leu Tyr Lys Arg Leu Gly Glu
                245                 250                 255
Ser Leu Gln Ser Asn Thr Tyr Asn Gly Phe Ser Thr Pro Trp Gly Tyr
            260                 265                 270
Phe Asp Phe Asn Arg Phe His Cys His Phe Ser Pro Arg Asp Trp Gln
        275                 280                 285
Arg Leu Ile Asn Asn Asn Trp Gly Met Arg Pro Lys Ala Met Arg Val
    290                 295                 300
Lys Ile Phe Asn Ile Gln Val Lys Glu Val Thr Thr Ser Asn Gly Glu
305                 310                 315                 320
Thr Thr Val Ala Asn Asn Leu Thr Ser Thr Val Gln Ile Phe Ala Asp
                325                 330                 335
Ser Ser Tyr Glu Leu Pro Tyr Val Met Asp Ala Gly Gln Glu Gly Ser
            340                 345                 350
Leu Pro Pro Phe Pro Asn Asp Val Phe Met Val Pro Gln Tyr Gly Tyr
        355                 360                 365
Cys Gly Leu Val Thr Gly Asn Thr Ser Gln Gln Gln Thr Asp Arg Asn
    370                 375                 380
Ala Phe Tyr Cys Leu Glu Tyr Phe Pro Ser Gln Met Leu Arg Thr Gly
385                 390                 395                 400
Asn Asn Phe Glu Ile Thr Tyr Ser Phe Glu Lys Val Pro Phe His Ser
                405                 410                 415
Met Tyr Ala His Ser Gln Ser Leu Asp Arg Leu Met Asn Pro Leu Ile
            420                 425                 430
Asp Gln Tyr Leu Trp Gly Leu Gln Ser Thr Thr Thr Gly Thr Thr Leu
        435                 440                 445
Asn Ala Gly Thr Ala Thr Thr Asn Phe Thr Lys Leu Arg Pro Thr Asn
    450                 455                 460
Phe Ser Asn Phe Lys Lys Asn Trp Leu Pro Gly Pro Ser Ile Lys Gln
465                 470                 475                 480
Gln Gly Phe Ser Lys Thr Ala Asn Gln Asn Tyr Lys Ile Pro Ala Thr
                485                 490                 495
Gly Ser Asp Ser Leu Ile Lys Tyr Glu Thr His Ser Thr Leu Asp Gly
            500                 505                 510
Arg Trp Ser Ala Leu Thr Pro Gly Pro Pro Met Ala Thr Ala Gly Pro
        515                 520                 525
Ala Asp Ser Lys Phe Ser Asn Ser Gln Leu Ile Phe Ala Gly Pro Lys
    530                 535                 540
Gln Asn Gly Asn Thr Ala Thr Val Pro Gly Thr Leu Ile Phe Thr Ser
545                 550                 555                 560
Glu Glu Glu Leu Ala Ala Thr Asn Ala Thr Asp Thr Asp Met Trp Gly
                565                 570                 575
Asn Leu Pro Gly Gly Asp Gln Ser Asn Ser Asn Leu Pro Thr Val Asp
            580                 585                 590
Arg Leu Thr Ala Leu Gly Ala Val Pro Gly Met Val Trp Gln Asn Arg
        595                 600                 605
Asp Ile Tyr Tyr Gln Gly Pro Ile Trp Ala Lys Ile Pro His Thr Asp
    610                 615                 620
```

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

<210> 64
<211> 736
<212> PRT
<213> AAV 血清型的衣壳蛋白, 克隆 AAV1

<400> 64
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 Asp Leu Lys Pro Gly Ala Pro Lys Pro
20 25 30
Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro
35 40 45
Gly Tyr Lys Tyr Leu Gly Pro Phe 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 Arg Tyr Asn His Ala
85 90 95
Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly
100 105 110
Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
115 120 125
Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg
130 135 140
Pro Val Glu Gln Ser Pro Gln Glu Pro Asp Ser Ser Ser Gly Ile Gly
145 150 155 160
Lys Thr Gly Gln Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr
165 170 175
Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Leu Gly Glu Pro Pro
180 185 190
Ala Thr Pro Ala Ala Val Gly Pro Thr Thr Met Ala Ser Gly Gly Gly
195 200 205
Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ala
210 215 220
Ser Gly Asn Trp His Cys Asp Ser Thr 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 Ser Ala Ser Thr Gly Ala Ser Asn Asp Asn His
260 265 270
Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe
275 280 285
His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn
290 295 300
Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile Gln
305 310 315 320
Val Lys Glu Val Thr Thr Asn Asp Gly Val Thr Thr Ile Ala Asn Asn

[0065]

325 330 335
Leu Thr Ser Thr Val Gln Val Phe Ser Asp Ser Glu Tyr Gln Leu Pro
340 345 350
Tyr Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala
355 360 365
Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly
370 375 380
Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro
385 390 395 400
Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Thr Phe Ser Tyr Thr Phe
405 410 415
Glu Glu Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp
420 425 430
Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Asn Arg
435 440 445
Thr Gln Asn Gln Ser Gly Ser Ala Gln Asn Lys Asp Leu Leu Phe Ser
450 455 460
Arg Gly Ser Pro Ala Gly Met Ser Val Gln Pro Lys Asn Trp Leu Pro
465 470 475 480
Gly Pro Cys Tyr Arg Gln Gln Arg Val Ser Lys Thr Lys Thr Asp Asn
485 490 495
Asn Asn Ser Asn Phe Thr Trp Thr Gly Ala Ser Lys Tyr Asn Leu Asn
500 505 510
Gly Arg Glu Ser Ile Ile Asn Pro Gly Thr Ala Met Ala Ser His Lys
515 520 525
Asp Asp Glu Asp Lys Phe Phe Pro Met Ser Gly Val Met Ile Phe Gly
530 535 540
Lys Glu Ser Ala Gly Ala Ser Asn Thr Ala Leu Asp Asn Val Met Ile
545 550 555 560
Thr Asp Glu Glu Glu Ile Lys Ala Thr Asn Pro Val Ala Thr Glu Arg
565 570 575
[0066] Phe Gly Thr Val Ala Val Asn Phe Gln Ser Ser Ser Thr Asp Pro Ala
580 585 590
Thr Gly Asp Val His Ala Met Gly Ala 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 His Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Leu
625 630 635 640
Lys Asn Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala
645 650 655
Asn Pro Pro Ala Glu Phe Ser Ala Thr Lys Phe Ala 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 Val Gln Tyr Thr Ser Asn
690 695 700
Tyr Ala Lys Ser Ala Asn Val Asp Phe Thr Val Asp Asn Asn Gly Leu
705 710 715 720
Tyr Thr Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Pro Leu
725 730 735

<210> 65
<211> 736
<212> PRT
<213> AAV 血清型的衣壳蛋白, 克隆 AAV6VP1

<400> 65
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 Asp Leu Lys Pro Gly Ala Pro Lys Pro
20 25 30

Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro
35 40 45
Gly Tyr Lys Tyr Leu Gly Pro Phe 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 Arg Tyr Asn His Ala
85 90 95
Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly
100 105 110
Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
115 120 125
Phe Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg
130 135 140
Pro Val Glu Gln Ser Pro Gln Glu Pro Asp Ser Ser Ser Gly Ile Gly
145 150 155 160
Lys Thr Gly Gln Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr
165 170 175
Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Leu Gly Glu Pro Pro
180 185 190
Ala Thr Pro Ala Ala Val Gly Pro Thr Thr Met Ala Ser Gly Gly Gly
195 200 205
Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ala
210 215 220
Ser Gly Asn Trp His Cys Asp Ser Thr 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 Ser Ala Ser Thr Gly Ala Ser Asn Asp Asn His
260 265 270
[0067] Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe
275 280 285
His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn
290 295 300
Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile Gln
305 310 315 320
Val Lys Glu Val Thr Thr Asn Asp Gly Val Thr Thr Ile Ala Asn Asn
325 330 335
Leu Thr Ser Thr Val Gln Val Phe Ser Asp Ser Glu Tyr Gln Leu Pro
340 345 350
Tyr Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala
355 360 365
Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly
370 375 380
Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro
385 390 395 400
Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Thr Phe Ser Tyr Thr Phe
405 410 415
Glu Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp
420 425 430
Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Asn Arg
435 440 445
Thr Gln Asn Gln Ser Gly Ser Ala Gln Asn Lys Asp Leu Leu Phe Ser
450 455 460
Arg Gly Ser Pro Ala Gly Met Ser Val Gln Pro Lys Asn Trp Leu Pro
465 470 475 480
Gly Pro Cys Tyr Arg Gln Gln Arg Val Ser Lys Thr Lys Thr Asp Asn
485 490 495
Asn Asn Ser Asn Phe Thr Trp Thr Gly Ala Ser Lys Tyr Asn Leu Asn
500 505 510
Gly Arg Glu Ser Ile Ile Asn Pro Gly Thr Ala Met Ala Ser His Lys
515 520 525

Asp Asp Lys Asp Lys Phe Phe Pro Met Ser Gly Val Met Ile Phe Gly
530 535 540
Lys Glu Ser Ala Gly Ala Ser Asn Thr Ala Leu Asp Asn Val Met Ile
545 550 555 560
Thr Asp Glu Glu Glu Ile Lys Ala Thr Asn Pro Val Ala Thr Glu Arg
565 570 575
Phe Gly Thr Val Ala Val Asn Leu Gln Ser Ser Ser Thr Asp Pro Ala
580 585 590
Thr Gly Asp Val His Val Met Gly Ala 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 His Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Leu
625 630 635 640
Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala
645 650 655
Asn Pro Pro Ala Glu Phe Ser Ala Thr Lys Phe Ala 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 Val Gln Tyr Thr Ser Asn
690 695 700
Tyr Ala Lys Ser Ala Asn Val Asp Phe Thr Val Asp Asn Asn Gly Leu
705 710 715 720
Tyr Thr Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Pro Leu
725 730 735

<210> 66
<211> 735
<212> PRT
<213> AAV血清型的衣壳蛋白,克隆A3.3

[0068]

<400> 66
Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Thr Leu Ser
1 5 10 15
Glu Gly Ile Arg Gln Trp Trp Lys Leu Lys Pro Gly Pro Pro Pro Pro
20 25 30
Lys Pro Asn Gln Gln His Arg Asp Asp Ser Arg Gly Leu Val Leu Pro
35 40 45
Gly Tyr Lys Tyr Leu Gly Pro Phe 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
His Gln Leu Lys Gln Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala
85 90 95
Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly
100 105 110
Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
115 120 125
Leu Gly Leu Val Glu Glu Ala Val Lys Thr Ala Pro Gly Lys Lys Arg
130 135 140
Pro Ile Glu Gln Ser Pro Ala Glu Pro Asp Ser Ser Ser Gly Ile Gly
145 150 155 160
Lys Ser Gly Gln Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr
165 170 175
Gly Asp Thr Glu Ser Val Pro Gly Pro Gln Pro Ile Gly Glu Pro Pro
180 185 190
Ala Ala Pro Ser Gly Val Gly Ser Asn Thr Met Ala Ser Gly Gly Gly
195 200 205
Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ser
210 215 220
Ser Gly Asn Trp His Cys Asp Ser Thr Trp Met 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 Ser Glu Ser Gly Ala Thr Asn Asp Asn His Tyr
260 265 270
Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His
275 280 285
Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Trp
290 295 300
Gly Phe Arg Pro Lys Lys Leu Asn Phe Lys Leu Phe Asn Ile Gln Val
305 310 315 320
Lys Glu Val Thr Gln Asn Asp Gly Thr Thr Thr Ile Ala Asn Asn Leu
325 330 335
Thr Ser Ala Val Gln Val Phe Thr Asp Ser Glu Tyr Gln Leu Pro Tyr
340 345 350
Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala Asp
355 360 365
Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly Ser
370 375 380
Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser
385 390 395 400
Gln Met Leu Arg Thr Gly Asn Asn Phe Thr Phe Ser Tyr Thr Phe Glu
405 410 415
Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp Arg
420 425 430
Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser Lys Thr
435 440 445
Gln Gly Thr Ser Gly Thr Thr Gln Gln Ser Arg Leu Gln Phe Ser Gln
450 455 460
Ala Gly Pro Ser Ser Met Ala Gln Gln Ala Lys Asn Trp Leu Pro Gly
465 470 475 480

[0069]

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

725 730 735

<210> 67
<211> 735
<212> PRT
<213> AAV 血清型的衣壳蛋白，克隆 A3.7

<400> 67

[0070]

```
Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Thr Leu Ser
1               5                  10                  15
Glu Gly Ile Arg Gln Trp Trp Lys Leu Lys Pro Gly Pro Pro Pro Pro
            20                  25                  30
Lys Pro Asn Gln Gln His Arg Asp Asp Ser Arg Gly Leu Val Leu Pro
        35                  40                  45
Gly Tyr Lys Tyr Leu Gly Pro Phe 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
His Gln Leu Lys Gln Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala
                85                  90                  95
Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly
            100                 105                 110
Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
        115                 120                 125
Leu Gly Leu Val Glu Glu Ala Val Lys Thr Ala Pro Gly Lys Lys Arg
    130                 135                 140
Pro Ile Glu Gln Ser Pro Ala Glu Pro Asp Ser Ser Ser Gly Ile Gly
145                 150                 155                 160
Lys Ser Gly Gln 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 Asn Thr Met Ala Ser Gly Gly Gly
        195                 200                 205
Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ser
    210                 215                 220
Ser Gly Asn Trp His Cys Asp Ser Thr Trp Met Gly Asp Arg Val Ile
225                 230                 235                 240
Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn Arg Leu
                245                 250                 255
Tyr Lys Gln Ile Ser Ser Glu Ser Gly Ala Thr Asn Asp Asn His Tyr
            260                 265                 270
Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His
        275                 280                 285
Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Trp
    290                 295                 300
Gly Phe Arg Pro Lys Lys Leu Asn Phe Lys Leu Phe Asn Ile Gln Val
305                 310                 315                 320
Lys Glu Val Thr Gln Asn Asp Gly Thr Thr Thr Ile Ala Asn Asn Leu
                325                 330                 335
Thr Ser Thr Val Gln Val Phe Thr Asp Ser Glu Tyr Gln Leu Pro Tyr
            340                 345                 350
Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala Asp
        355                 360                 365
Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly Ser
    370                 375                 380
Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser
385                 390                 395                 400
Gln Met Leu Arg Thr Gly Asn Asn Phe Thr Phe Ser Tyr Thr Phe Glu
                405                 410                 415
Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp Arg
            420                 425                 430
```

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

[0071]

<210> 68
<211> 735
<212> PRT
<213> AAV血清型的衣壳蛋白,克隆A3.4

<400> 68
Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Thr Leu Ser
1 5 10 15
Glu Gly Ile Arg Gln Trp Trp Lys Leu Lys Pro Gly Pro Pro Pro Pro
20 25 30
Lys Pro Asn Gln Gln His Arg Asp Asp Ser Arg Gly Leu Val Leu Pro
35 40 45
Gly Tyr Lys Tyr Leu Gly Pro Phe 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
His Gln Leu Lys Gln Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala
85 90 95
Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly
100 105 110
Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
115 120 125
Leu Gly Leu Val Glu Glu Ala Val Lys Thr Ala Pro Gly Lys Lys Arg

[0072]

```
    130                 135                 140
Pro Ile Glu Gln Ser Pro Ala Glu Pro Asp Ser Ser Ser Gly Ile Gly
145                 150                 155                 160
Glu Ser Gly Gln 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 Asn Thr Met Ala Ser Gly Gly Gly
        195                 200                 205
Ala Pro Met Ala Asp Asp Asn Glu Gly Ala Asp Gly Val Gly Asn Ser
    210                 215                 220
Ser Gly Asn Trp His Cys Asp Ser Thr Trp Met 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 Ser Glu Ser Gly Ala Thr Asn Asp Asn His Tyr
            260                 265                 270
Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His
        275                 280                 285
Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Trp
    290                 295                 300
Gly Phe Arg Pro Lys Lys Leu Asn Phe Lys Leu Phe Asn Ile Gln Val
305                 310                 315                 320
Lys Glu Val Thr Gln Asn Asp Gly Thr Thr Thr Ile Ala Asn Asn Leu
                325                 330                 335
Thr Ser Thr Val Gln Val Phe Thr Asp Ser Glu Tyr Gln Leu Pro Tyr
            340                 345                 350
Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala Asp
        355                 360                 365
Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly Ser
    370                 375                 380
Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser
385                 390                 395                 400
Gln Met Leu Arg Thr Gly Asn Asn Phe Thr Phe Ser Tyr Thr Phe Glu
                405                 410                 415
Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp Arg
            420                 425                 430
Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser Lys Thr
        435                 440                 445
Gln Gly Thr Ser Gly Thr Thr Gln Gln Ser Arg Leu Gln Phe Ser Gln
    450                 455                 460
Ala Gly Pro Ser Ser Met Ala Gln Gln Ala Lys Asn Trp Leu Pro Gly
465                 470                 475                 480
Pro Ser Tyr Arg Gln Gln Arg Met Ser Lys Thr Ala Asn Asp Asn Asn
                485                 490                 495
Asn Ser Glu Phe Ala Trp Thr Ala Ala Thr Lys Tyr Tyr Leu Asn Gly
            500                 505                 510
Arg Asn Ser Leu Val Asn Pro Gly Pro Pro Met Ala Ser His Lys Asp
        515                 520                 525
Asp Glu Glu Lys Tyr Phe Pro Met His Gly Asn Leu Ile Phe Gly Lys
    530                 535                 540
Gln Gly Thr Gly Thr Thr Asn Val Asp Ile Glu Ser Val Leu Ile Thr
545                 550                 555                 560
Asp Glu Glu Glu Ile Arg Thr Thr Asn Pro Val Ala Thr Glu Gln Tyr
                565                 570                 575
Gly Gln Val Ala Thr Asn His Gln Ser Gln Asp Thr Thr Ala Ser Tyr
            580                 585                 590
Gly Ser Val Asp Ser Gln Gly Ile Leu Pro Gly Met Val Trp Gln Asp
        595                 600                 605
Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Thr Pro His Thr
    610                 615                 620
Asp Gly His Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Leu Lys
```

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

<210> 69
<211> 735
<212> PRT
<213> AAV 血清型的衣壳蛋白,克隆 A3.5

<400> 69
Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Thr Leu Ser
1 5 10 15
Glu Gly Ile Arg Gln Trp Trp Lys Leu Lys Pro Gly Pro Pro Pro Pro
20 25 30
Lys Pro Asn Gln Gln His Arg Asp Asp Ser Arg Gly Leu Val Leu Pro
35 40 45
Gly Tyr Lys Tyr Leu Gly Pro Phe 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
His Gln Leu Lys Gln Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala
[0073] 85 90 95
Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly
100 105 110
Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
115 120 125
Leu Gly Leu Val Glu Glu Ala Val Lys Thr Ala Pro Gly Lys Lys Arg
130 135 140
Pro Ile Glu Gln Ser Pro Ala Glu Pro Asp Ser Ser Ser Gly Ile Gly
145 150 155 160
Lys Ser Gly Gln 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 Asn Thr Met Ala Ser Gly Gly Gly
195 200 205
Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ser
210 215 220
Ser Gly Asn Trp His Cys Asp Ser Thr Trp Met 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 Ser Glu Ser Gly Ala Thr Asn Asp Asn His Tyr
260 265 270
Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His
275 280 285
Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Trp
290 295 300
Gly Phe Arg Pro Lys Lys Leu Asn Phe Lys Leu Phe Asn Ile Gln Val
305 310 315 320
Lys Glu Val Thr Gln Asn Asp Gly Thr Thr Thr Ile Ala Asn Asn Leu
325 330 335

[0074]

```
Thr Ser Thr Val Gln Val Phe Thr Asp Ser Glu Tyr Gln Leu Pro Tyr
            340                 345                 350
Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala Asp
        355                 360                 365
Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly Ser
    370                 375                 380
Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser
385                 390                 395                 400
Gln Met Leu Arg Thr Gly Asn Asn Phe Thr Phe Ser Tyr Thr Phe Glu
                405                 410                 415
Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp Arg
            420                 425                 430
Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser Lys Thr
        435                 440                 445
Gln Gly Thr Ser Gly Thr Thr Gln Gln Ser Arg Leu Gln Phe Asn Gln
    450                 455                 460
Ala Gly Pro Ser Ser Met Ala Gln Gln Ala Lys Asn Trp Leu Pro Gly
465                 470                 475                 480
Pro Ser Tyr Arg Gln Gln Arg Met Ser Lys Thr Ala Asn Asp Asn Asn
                485                 490                 495
Asn Ser Glu Phe Ala Trp Thr Ala Ala Thr Lys Tyr Tyr Pro Asn Gly
            500                 505                 510
Arg Asn Ser Leu Val Asn Pro Gly Pro Pro Met Ala Ser His Lys Asp
        515                 520                 525
Asp Glu Glu Lys Tyr Phe Pro Met His Gly Asn Leu Ile Phe Gly Lys
    530                 535                 540
Gln Gly Thr Gly Thr Thr Asn Val Asp Ile Glu Ser Val Leu Ile Thr
545                 550                 555                 560
Asp Glu Glu Glu Ile Arg Thr Thr Asn Pro Val Ala Thr Glu Gln Tyr
                565                 570                 575
Gly Gln Val Ala Thr Asn Arg Gln Ser Gln Asn Thr Thr Ala Ser Tyr
            580                 585                 590
Gly Ser Val Asp Ser Gln Gly Ile Leu Pro Gly Met Val Trp Gln Asp
        595                 600                 605
Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Thr Pro His Thr
    610                 615                 620
Asp Gly His Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Leu Lys
625                 630                 635                 640
His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala Asn
                645                 650                 655
Pro Ala Thr Thr Phe Thr Pro Gly Lys Phe Ala Ser Phe Ile Thr Gln
            660                 665                 670
Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln Lys
        675                 680                 685
Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn Tyr
    690                 695                 700
Asn Lys Ser Val Asn Val Glu Phe Thr Val Asp Ala Asn Gly Val Tyr
705                 710                 715                 720
Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu
                725                 730                 735
```

<210> 70
<211> 735
<212> PRT
<213> AAV 血清型的衣壳蛋白, 克隆 AAV2

<400> 70

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

```
        35                  40                  45
Gly Tyr Lys Tyr Leu Gly Pro Phe 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
Arg Gln Leu Asp Ser 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 Val Leu Glu Pro
        115                 120                 125
Leu Gly Leu Val Glu Glu Pro Val Lys Thr Ala Pro Gly Lys Lys Arg
    130                 135                 140
Pro Val Glu His Ser Pro Val Glu Pro Asp Ser Ser Ser Gly Thr Gly
145                 150                 155                 160
Lys Ala Gly Gln Gln Pro Ala Arg Lys Arg Leu Asn Phe Gly Gln Thr
                165                 170                 175
Gly Asp Ala Asp Ser Val Pro Asp Pro Gln Pro Leu Gly Gln Pro Pro
            180                 185                 190

Ala Ala Pro Ser Gly Leu Gly Thr Asn Thr Met Ala Thr Gly Ser Gly
        195                 200                 205
Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ser
    210                 215                 220
Ser Gly Asn Trp His Cys Asp Ser Thr Trp Met 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 Ser Gln Ser Gly Ala Ser Asn Asp Asn His Tyr
            260                 265                 270
Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His
        275                 280                 285
Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Trp
    290                 295                 300
Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile Gln Val
305                 310                 315                 320
Lys Glu Val Thr Gln Asn Asp Gly Thr Thr Thr Ile Ala Asn Asn Leu
                325                 330                 335
Thr Ser Thr Val Gln Val Phe Thr Asp Ser Glu Tyr Gln Leu Pro Tyr
            340                 345                 350
Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala Asp
        355                 360                 365
Val Phe Met Val Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly Ser
    370                 375                 380
Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser
385                 390                 395                 400
Gln Met Leu Arg Thr Gly Asn Asn Phe Thr Phe Ser Tyr Thr Phe Glu
                405                 410                 415
Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp Arg
            420                 425                 430
Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser Arg Thr
        435                 440                 445
Asn Thr Pro Ser Gly Thr Thr Thr Gln Ser Arg Leu Gln Phe Ser Gln
    450                 455                 460
Ala Gly Ala Ser Asp Ile Arg Asp Gln Ser Arg Asn Trp Leu Pro Gly
465                 470                 475                 480
Pro Cys Tyr Arg Gln Gln Arg Val Ser Lys Thr Ser Ala Asp Asn Asn
                485                 490                 495
Asn Ser Glu Tyr Ser Trp Thr Gly Ala Thr Lys Tyr His Leu Asn Gly
            500                 505                 510
Arg Asp Ser Leu Val Asn Pro Gly Pro Ala Met Ala Ser His Lys Asp
        515                 520                 525
```

[0075]

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

<210> 71
<211> 736
<212> PRT
<213> AAV血清型的衣壳蛋白,克隆AAV3

[0076]

```
<400> 71
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 Val Pro Gln Pro
            20                  25                  30
Lys Ala Asn Gln Gln His Gln Asp Asn Arg 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 Gln Glu Asp Thr Ser Phe Gly Gly
            100                 105                 110
Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Ile Leu Glu Pro
        115                 120                 125
Leu Gly Leu Val Glu Glu Ala Ala Lys Thr Ala Pro Gly Lys Lys Gly
    130                 135                 140
Ala Val Asp Gln Ser Pro Gln Glu Pro Asp Ser Ser Ser Gly Val Gly
145                 150                 155                 160
Lys Ser Gly Lys Gln Pro Ala Arg Lys Arg Leu Asn Phe Gly Gln Thr
                165                 170                 175
Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Leu Gly Glu Pro Pro
            180                 185                 190
Ala Ala Pro Thr Ser Leu Gly Ser Asn Thr Met Ala Ser Gly Gly Gly
        195                 200                 205
Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn 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 Ser Gln Ser Gly Ala Ser Asn Asp Asn His Tyr
260 265 270
Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His
275 280 285
Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Trp
290 295 300
Gly Phe Arg Pro Lys Lys Leu Ser Phe Lys Leu Phe Asn Ile Gln Val
305 310 315 320
Arg Gly Val Thr Gln Asn Asp Gly Thr Thr Thr Ile Ala Asn Asn Leu
325 330 335
Thr Ser Thr Val Gln Val Phe Thr Asp Ser Glu Tyr Gln Leu Pro Tyr
340 345 350
Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala Asp
355 360 365
Val Phe Met Val Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly Ser
370 375 380
Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser
385 390 395 400
Gln Met Leu Arg Thr Gly Asn Asn Phe Gln Phe Ser Tyr Thr Phe Glu
405 410 415
Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp Arg
420 425 430
Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Asn Arg Thr
435 440 445
Gln Gly Thr Thr Ser Gly Thr Thr Asn Gln Ser Arg Leu Leu Phe Ser
450 455 460
Gln Ala Gly Pro Glu Ser Met Ser Leu Gln Ala Arg Asn Trp Leu Pro
465 470 475 480

[0077]

Gly Pro Cys Tyr Arg Gln Gln Arg Leu Ser Lys Thr Ala Asn Asp Asn
485 490 495
Asn Asn Ser Asn Phe Pro Trp Thr Ala Ala Ser Lys Tyr His Leu Asn
500 505 510
Gly Arg Asp Ser Leu Val Asn Pro Gly Pro Ala Met Ala Ser His Lys
515 520 525
Asp Asp Glu Glu Lys Phe Phe Pro Met His Gly Asn Leu Ile Phe Gly
530 535 540
Lys Glu Gly Thr Thr Ala Ser Asn Ala Glu Leu Asp Asn Val Met Ile
545 550 555 560
Thr Asp Glu Glu Glu Ile Arg Thr Thr Asn Pro Val Ala Thr Glu Gln
565 570 575
Tyr Gly Thr Val Ala Asn Asn Leu Gln Ser Ser Asn Thr Ala Pro Thr
580 585 590
Thr Gly Thr Val Asn His Gln Gly Ala 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 His Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Leu
625 630 635 640
Lys His Pro Pro Pro Gln Ile Met Ile Lys Asn Thr Pro Val Pro Ala
645 650 655
Asn Pro Pro Thr Thr Phe Ser Pro Ala Lys Phe Ala 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 Asn Lys Ser Val Asn Val Asp Phe Thr Val Asp Thr Asn 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> 72
<211> 737
<212> PRT
<213> AAV 血清型的衣壳蛋白，克隆 3.3bVP1

<400> 72

[0078]

```
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 Asp Leu Lys Pro Gly Ala Pro Lys Pro
            20                  25                  30
Lys Ala Asn Gln Gln Lys Gln Asp Asn Gly Arg Gly Leu Val Leu Pro
        35                  40                  45
Gly Tyr Lys Tyr Leu Gly Pro Phe 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 Asn Ala Gly Asp Asn Pro Tyr Leu Arg Tyr Asn His Ala
                85                  90                  95
Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly
            100                 105                 110
Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
        115                 120                 125
Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Ala Lys Lys Arg
    130                 135                 140
Pro Val Glu Pro Ser Pro Gln Arg Ser Pro Asp Ser Ser Thr Gly Ile
145                 150                 155                 160
Gly Lys Lys Gly Gln Gln Pro Ala Arg Lys Arg Leu Asn Phe Gly Gln
                165                 170                 175
Thr Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Leu Gly Glu Pro
            180                 185                 190
Pro Ala Ala Pro Ser Ser Val Gly Ser Gly Thr Val Ala Ala Gly Gly
        195                 200                 205
Gly Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn
    210                 215                 220
Ala Ser Gly Asn Trp His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val
225                 230                 235                 240
Ile Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His
                245                 250                 255
Leu Tyr Glu Gln Ile Ser Ser Glu Thr Ala Gly Ser Thr Asn Asp Asn
            260                 265                 270
Thr 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 Lys Leu Arg Phe Lys Leu Phe Asn Ile
305                 310                 315                 320
Gln Val Lys Glu Val Thr Thr Asn Asp Gly Val Thr Thr Ile Ala Asn
                325                 330                 335
Asn Leu Thr Ser Thr Ile Gln Val Phe Ser Asp Ser Glu Tyr Gln Leu
            340                 345                 350
Pro Tyr Val Leu Gly Ser Ala His Gln 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 Asn
    370                 375                 380
Gly Ser Gln Ser 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 Glu Phe Ser Tyr Ser
                405                 410                 415
Phe Glu Asp 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 Ala
435 440 445
Arg Thr Gln Ser Asp Pro Gly Gly Thr Ala Gly Asn Arg Glu Leu Gln
450 455 460
Phe Tyr Gln Gly Gly Pro Ser Thr Met Ala Glu Gln Ala Lys Asn Trp
465 470 475 480
Leu Pro Gly Pro Cys Phe Arg Gln Gln Arg Val Ser Lys Thr Leu Asp
485 490 495
Gln Asn Asn Asn Ser Asn Phe Ala Trp Thr Gly Ala Thr Lys Tyr His
500 505 510
Leu Asn Gly Arg Asn Ser Leu Val Asn Pro Gly Val Ala Met Ala Thr
515 520 525
His Lys Asp Asp Glu Asp Arg Phe Phe Pro Ser Ser Gly Val Leu Ile
530 535 540
Phe Gly Lys Thr Gly Ala Thr Asn Lys Thr Thr Leu Glu Asn Val Leu
545 550 555 560
Met Thr Asn Glu Glu Glu Ile Arg Pro Thr Asn Pro Val Ala Thr Glu
565 570 575
Glu Tyr Gly Ile Val Ser Ser Asn Leu Gln Ala Ala Asn Thr Ala Ala
580 585 590
Gln Thr Gln Val Val Asn Asn Gln Gly Ala Leu Pro Gly Met Val Trp
595 600 605
Gln Asn Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro
610 615 620
His Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly
625 630 635 640
Leu Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro
645 650 655
Ala Asn Pro Pro Glu Val Phe Thr Pro Ala Lys Phe Ala Ser Phe Ile
660 665 670
Thr Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu
675 680 685
Gln Lys Glu Asn Ser Lys Arg Trp Asp Pro Glu Ile Gln Tyr Thr Ser
690 695 700
Asn Phe Glu Lys Gln Thr Gly Val Asp Phe Ala Val Asp Ser Gln Gly
705 710 715 720
Val Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn
725 730 735
Leu

[0079]

<210> 73
<211> 644
<212> PRT
<213> AAV血清型的衣壳蛋白,克隆223-4

<400> 73
Lys Ala Tyr Asp Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Arg
1 5 10 15
Tyr Asn His Ala Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr
20 25 30
Ser Phe Gly Gly Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg
35 40 45
Val Leu Glu Pro Leu Gly Leu Val Glu Thr Pro Ala Lys Thr Ala Pro
50 55 60
Gly Lys Lys Arg Pro Val Asp Ser Pro Asp Ser Thr Ser Gly Ile Gly
65 70 75 80
Lys Lys Gly Gln Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr
85 90 95
Gly Asp Ser Glu Pro Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro
100 105 110
Ala Gly Pro Ser Gly Leu Gly Ser Gly Thr Met Ala Ala Gly Gly Gly

[0080]

```
          115                 120                 125
Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ala
      130                 135                 140
Ser Gly Asn Trp His Cys Asp Ser Thr Arg Leu Gly Asp Arg Val Ile
145                 150                 155                 160
Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu
                165                 170                 175
Tyr Lys Gln Ile Ser Ser Gln Ser Ala Gly Ser Thr Asn Asp Asn Val
            180                 185                 190
Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe
          195                 200                 205
His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn
      210                 215                 220
Trp Gly Phe Arg Pro Lys Lys Leu Asn Phe Lys Leu Phe Asn Ile Gln
225                 230                 235                 240
Val Lys Glu Val Thr Thr Asn Asp Gly Val Thr Thr Ile Ala Asn Asn
                245                 250                 255
Leu Thr Ser Thr Val Gln Val Phe Ser Asp Ser Glu Tyr Gln Leu Pro
            260                 265                 270
Tyr Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala
          275                 280                 285
Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly
      290                 295                 300
Ser Gln Ser Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro
305                 310                 315                 320
Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Thr Phe Ser Tyr Thr Phe
                325                 330                 335
Glu Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Gly
            340                 345                 350
Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ala Arg
          355                 360                 365
Thr Gln Ser Asn Ala Gly Gly Thr Ala Gly Asn Arg Glu Leu Gln Phe
      370                 375                 380
Tyr Gln Gly Gly Pro Thr Thr Met Ala Glu Gln Ala Lys Asn Trp Leu
385                 390                 395                 400
Pro Gly Pro Cys Phe Arg Gln Gln Arg Val Ser Lys Thr Leu Asp Gln
                405                 410                 415
Asn Asn Asn Ser Asn Phe Ala Trp Thr Gly Ala Thr Lys Tyr His Leu
            420                 425                 430
Asn Gly Arg Asn Ser Leu Val Asn Pro Gly Val Ala Met Ala Thr His
          435                 440                 445
Lys Asp Asp Glu Glu Arg Phe Phe Pro Ser Ser Gly Val Leu Ile Phe
      450                 455                 460
Gly Lys Thr Gly Ala Ala Asn Lys Thr Thr Leu Glu Asn Val Leu Met
465                 470                 475                 480
Thr Asn Glu Glu Glu Ile Arg Pro Thr Asn Pro Val Ala Thr Glu Glu
                485                 490                 495
Tyr Gly Ile Val Ser Ser Asn Leu Gln Ala Ala Ser Thr Ala Ala Gln
            500                 505                 510
Thr Gln Val Val Asn Asn Gln Gly Ala Leu Pro Gly Met Val Trp Gln
          515                 520                 525
Asn Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His
      530                 535                 540
Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Leu
545                 550                 555                 560
Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala
                565                 570                 575
Asn Pro Pro Glu Val Phe Thr Pro Ala Lys Phe Ala Ser Phe Ile Thr
            580                 585                 590
Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln
          595                 600                 605
Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn
```

610 615 620
Phe Asp Lys Gln Thr Gly Val Asp Phe Ala Val Asp Ser Gln Gly Val
625 630 635 640
Tyr Ser Glu Pro

<210> 74
<211> 644
<212> PRT
<213> AAV血清型的衣壳蛋白,克隆223.5

<400> 74
Lys Ala Tyr Asp Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Arg
1 5 10 15
Tyr Asn His Ala Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr
20 25 30
Ser Phe Gly Gly Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg
35 40 45
Val Leu Glu Pro Leu Gly Leu Val Glu Thr Pro Ala Lys Thr Ala Pro
50 55 60
Gly Lys Lys Arg Pro Val Asp Ser Pro Asp Ser Thr Ser Gly Ile Gly
65 70 75 80
Lys Lys Gly Gln Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr
85 90 95
Gly Asp Ser Glu Pro Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro
100 105 110
Ala Gly Pro Ser Gly Leu Gly Ser Gly Thr Met Ala Ala Gly Gly Gly
115 120 125
Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ala
130 135 140
[0081] Ser Gly Asn Trp His Cys Asp Ser Thr Arg Leu Gly Asp Arg Val Ile
145 150 155 160
Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu
165 170 175
Tyr Lys Gln Ile Ser Ser Gln Ser Ala Gly Ser Thr Asn Asp Asn Val
180 185 190
Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe
195 200 205
His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn
210 215 220
Trp Gly Phe Arg Pro Lys Lys Leu Asn Phe Lys Leu Phe Asn Ile Gln
225 230 235 240
Val Lys Glu Val Thr Thr Asn Asp Gly Val Thr Thr Ile Ala Asn Asn
245 250 255
Leu Thr Ser Thr Val Gln Val Phe Ser Asp Ser Glu Tyr Gln Leu Pro
260 265 270
Tyr Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala
275 280 285
Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly
290 295 300
Ser Gln Ser Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro
305 310 315 320
Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Thr Phe Ser Tyr Thr Phe
325 330 335
Glu Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Gly
340 345 350
Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ala Arg
355 360 365
Thr Gln Ser Asn Ala Gly Gly Thr Ala Gly Asn Arg Glu Leu Gln Phe
370 375 380
Tyr Gln Gly Gly Pro Thr Thr Met Ala Glu Gln Ala Lys Asn Trp Leu
385 390 395 400

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

[0082]

```
<210>  75
<211>  644
<212>  PRT
<213>  AAV 血清型的衣壳蛋白, 克隆 223.10

<220>
<221>  MISC_FEATURE
<222>  (434)..(434)
<223>  can be any amino acid
<400>  75
Lys Ala Tyr Asp Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Arg
1               5                   10                  15
Tyr Asn His Ala Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr
            20                  25                  30
Ser Phe Gly Gly Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg
        35                  40                  45
Val Leu Glu Pro Leu Gly Leu Val Glu Thr Pro Ala Lys Thr Ala Pro
    50                  55                  60
Gly Lys Lys Arg Pro Val Asp Ser Pro Asp Ser Thr Ser Gly Ile Gly
65                  70                  75                  80
Lys Lys Gly Gln Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr
                85                  90                  95
Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro
            100                 105                 110
Ala Gly Pro Ser Gly Leu Gly Ser Gly Thr Met Ala Ala Gly Gly Gly
        115                 120                 125
Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ala
    130                 135                 140
Ser Gly Asn Trp His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val Ile
```

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

[0083]

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

<210> 76
<211> 644
<212> PRT
<213> AAV 血清型的衣壳蛋白, 克隆 223.2

<400> 76

[0084]

Lys Ala Tyr Asp Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Arg
1 5 10 15
Tyr Asn His Ala Asp Ala Glu Phe Gln Glu Cys Leu Gln Glu Asp Thr
20 25 30
Ser Phe Gly Gly Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg
35 40 45
Val Leu Glu Pro Leu Gly Leu Val Glu Thr Pro Ala Lys Thr Ala Pro
50 55 60
Gly Lys Lys Arg Pro Val Asp Ser Pro Asp Ser Thr Ser Gly Ile Gly
65 70 75 80
Lys Lys Gly Gln Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr
85 90 95
Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro
100 105 110
Ala Gly Pro Ser Gly Leu Gly Ser Gly Thr Met Val Ala Gly Gly Gly
115 120 125
Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ala
130 135 140
Ser Gly Asn Trp His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val Ile
145 150 155 160
Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu
165 170 175
Tyr Lys Gln Ile Ser Ser Gln Ser Ala Gly Ser Thr Asn Asp Asn Val
180 185 190
Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe
195 200 205
His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn
210 215 220
Trp Gly Phe Arg Pro Lys Lys Leu Asn Phe Lys Leu Phe Asn Ile Gln
225 230 235 240
Val Lys Glu Val Thr Thr Asn Asp Gly Val Thr Thr Ile Ala Asn Asn
245 250 255
Leu Thr Ser Thr Val Gln Val Phe Ser Asp Ser Glu Tyr Gln Leu Pro
260 265 270
Tyr Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala
275 280 285
Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly
290 295 300
Ser Gln Ser Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro
305 310 315 320
Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Thr Phe Ser Tyr Thr Phe
325 330 335
Glu Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp
340 345 350
Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ala Arg
355 360 365
Thr Gln Ser Asn Ala Gly Gly Thr Ala Gly Asn Arg Glu Leu Gln Phe
370 375 380
Tyr Gln Gly Gly Pro Thr Thr Met Ala Glu Gln Ala Lys Asn Trp Leu
385 390 395 400
Pro Gly Pro Cys Phe Arg Gln Gln Arg Val Ser Lys Thr Leu Asp Gln
405 410 415
Asn Asn Asn Ser Asn Phe Ala Trp Thr Gly Ala Thr Lys Tyr His Leu
420 425 430

```
Asn Gly Arg Asn Ser Leu Val Asn Pro Gly Val Ala Met Ala Thr His
        435                 440                 445
Lys Asp Asp Glu Glu Arg Phe Ser Pro Ser Ser Gly Val Leu Ile Phe
    450                 455                 460
Gly Lys Thr Gly Ala Ala Asn Lys Thr Thr Leu Glu Asn Val Leu Met
465                 470                 475                 480
Thr Asn Glu Glu Glu Ile Arg Pro Thr Asn Pro Val Ala Thr Glu Glu
                485                 490                 495
Tyr Gly Ile Val Ser Ser Asn Leu Gln Ala Ala Ser Thr Ala Ala Gln
            500                 505                 510
Thr Gln Val Val Asn Asn Gln Gly Ala Leu Pro Gly Met Val Trp Gln
        515                 520                 525
Asn Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His
    530                 535                 540
Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Leu
545                 550                 555                 560
Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala
                565                 570                 575
Asn Pro Pro Glu Val Phe Thr Pro Ala Lys Phe Ala Ser Phe Ile Thr
            580                 585                 590
Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln
        595                 600                 605
Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn
    610                 615                 620
Phe Asp Lys Gln Thr Gly Val Asp Phe Ala Val Asp Ser Gln Gly Val
625                 630                 635                 640
Tyr Ser Glu Pro
```

[0085]

<210> 77
<211> 644
<212> PRT
<213> AAV血清型的衣壳蛋白,克隆223.7

<400> 77

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

```
    210                 215                 220
Trp Gly Phe Arg Pro Lys Lys Leu Asn Phe Lys Leu Phe Asn Ile Gln
225                 230                 235                 240
Val Lys Glu Val Thr Thr Asn Asp Gly Val Thr Thr Ile Ala Asn Asn
                245                 250                 255
Leu Thr Ser Thr Val Gln Val Phe Ser Asp Pro Glu Tyr Gln Leu Pro
            260                 265                 270
Tyr Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala
        275                 280                 285
Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly
    290                 295                 300
Ser Gln Ser Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro
305                 310                 315                 320
Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Thr Phe Ser Tyr Thr Phe
                325                 330                 335
Glu Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp
            340                 345                 350
Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ala Arg
        355                 360                 365
Thr Gln Ser Asn Ala Gly Gly Thr Ala Gly Asn Arg Glu Leu Gln Phe
    370                 375                 380
Tyr Gln Gly Gly Pro Thr Thr Met Ala Glu Gln Ala Lys Asn Trp Leu
385                 390                 395                 400
Pro Gly Pro Cys Phe Arg Gln Gln Arg Val Ser Lys Thr Leu Asp Gln
                405                 410                 415
Asn Asn Asn Ser Asn Phe Ala Trp Thr Gly Ala Thr Lys Tyr His Leu
            420                 425                 430
Asn Gly Arg Asn Ser Leu Val Asn Pro Gly Val Ala Met Ala Thr His
        435                 440                 445
Lys Asp Asp Glu Glu Arg Phe Phe Pro Ser Ser Gly Val Leu Ile Phe
    450                 455                 460
Gly Lys Thr Gly Ala Ala Asn Lys Thr Thr Leu Glu Asn Val Leu Met
465                 470                 475                 480
Thr Asn Glu Glu Glu Ile Arg Pro Thr Asn Pro Val Ala Thr Glu Glu
                485                 490                 495
Tyr Gly Ile Val Ser Ser Asn Leu Gln Ala Ala Ser Thr Ala Ala Gln
            500                 505                 510
Thr Gln Val Val Asn Asn Gln Gly Ala Leu Pro Gly Met Val Trp Gln
        515                 520                 525
Asn Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His
    530                 535                 540
Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Leu
545                 550                 555                 560
Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala
                565                 570                 575
Asn Pro Pro Glu Val Phe Thr Pro Ala Lys Ile Ala Ser Phe Ile Thr
            580                 585                 590
Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln
        595                 600                 605
Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn
    610                 615                 620
Phe Asp Lys Gln Thr Gly Val Asp Phe Ala Val Asp Ser Gln Gly Val
625                 630                 635                 640
Tyr Ser Glu Pro
```

[0086]

<210> 78
<211> 644
<212> PRT
<213> AAV 血清型的衣壳蛋白, 克隆 223.6

<400> 78

[0087]

```
Lys Ala Tyr Asp Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Arg
1               5                  10                  15
Tyr Asn His Ala Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr
            20                  25                  30
Ser Phe Gly Gly Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg
        35                  40                  45
Val Leu Glu Pro Leu Gly Leu Val Glu Thr Pro Ala Lys Thr Ala Pro
    50                  55                  60
Gly Lys Lys Arg Pro Val Asp Ser Pro Asp Ser Thr Ser Gly Ile Gly
65                  70                  75                  80
Lys Lys Gly Gln Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr
                85                  90                  95
Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro
            100                 105                 110
Ala Gly Pro Ser Gly Leu Gly Ser Gly Thr Met Ala Ala Gly Gly Gly
        115                 120                 125
Ala Pro Met Ala Asp Asn Ser Glu Gly Ala Asp Gly Val Gly Asn Ala
    130                 135                 140
Ser Gly Asn Trp His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val Ile
145                 150                 155                 160
Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu
                165                 170                 175
Tyr Lys Gln Ile Ser Ser Gln Ser Ala Gly Ser Thr Asn Asp Asn Val
            180                 185                 190
Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe
        195                 200                 205
His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn
    210                 215                 220
Trp Gly Phe Arg Pro Lys Lys Leu Asn Phe Lys Leu Phe Asn Ile Gln
225                 230                 235                 240
Val Lys Glu Val Thr Thr Asn Asp Gly Val Thr Thr Ile Ala Asn Asn
                245                 250                 255
Leu Thr Ser Thr Val Gln Val Phe Ser Asp Ser Glu Tyr Gln Leu Pro
            260                 265                 270
Tyr Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala
        275                 280                 285
Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly
    290                 295                 300
Ser Gln Ser Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro
305                 310                 315                 320
Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Thr Phe Ser Tyr Thr Phe
                325                 330                 335
Glu Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp
            340                 345                 350
Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ala Arg
        355                 360                 365
Thr Gln Ser Asn Ala Gly Gly Thr Ala Gly Asn Arg Glu Leu Gln Phe
    370                 375                 380
Tyr Gln Gly Gly Pro Thr Thr Met Ala Glu Gln Ala Lys Asn Trp Leu
385                 390                 395                 400
Pro Gly Pro Cys Phe Arg Gln Gln Arg Val Ser Lys Thr Leu Asp Gln
                405                 410                 415
Asn Asn Asn Ser Asn Phe Ala Trp Thr Gly Ala Thr Lys Tyr His Leu
            420                 425                 430
Asn Gly Arg Asn Ser Leu Val Asn Pro Gly Val Ala Met Ala Thr His
        435                 440                 445
Lys Asp Asp Glu Glu Arg Phe Phe Pro Ser Ser Gly Val Leu Ile Phe
    450                 455                 460
Gly Lys Thr Gly Ala Ala Asn Lys Thr Thr Leu Glu Asn Val Leu Met
465                 470                 475                 480
Thr Asn Glu Glu Glu Ile Arg Pro Thr Asn Pro Val Ala Thr Glu Glu
                485                 490                 495
```

Tyr Gly Ile Val Ser Ser Asn Leu Gln Ala Ala Ser Thr Ala Ala Gln
500 505 510
Thr Gln Val Val Asn Asn Gln Gly Ala Leu Pro Gly Met Val Trp Gln
515 520 525
Asn Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His
530 535 540
Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Leu
545 550 555 560
Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala
565 570 575
Asn Pro Pro Glu Val Phe Thr Pro Ala Lys Leu Ala Ser Phe Ile Thr
580 585 590
Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln
595 600 605
Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn
610 615 620
Phe Asp Lys Gln Thr Gly Val Asp Phe Ala Val Asp Ser Gln Gly Val
625 630 635 640
Tyr Ser Glu Pro

<210> 79
<211> 738
<212> PRT
<213> AAV 血清型的衣壳蛋白, 克隆 44.1

<400> 79
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 Asp Leu Lys Pro Gly Ala Pro Lys Pro
20 25 30
Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro
35 40 45
Gly Tyr Lys Tyr Leu Gly Pro Phe 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 Arg Tyr Asn His Ala
85 90 95
Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly
100 105 110
Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
115 120 125
Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg
130 135 140
Pro Val Glu Pro Ser Pro Gln Arg Ser Pro Asp Ser Ser Thr Gly Ile
145 150 155 160
Gly Lys Lys Gly Gln Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln
165 170 175
Thr Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro
180 185 190
Pro Ala Gly Pro Ser Gly Leu Gly Ser Gly Thr Met Ala Ala Gly Gly
195 200 205

Gly Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser
210 215 220
Ser Ser Gly Asn Trp His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val
225 230 235 240
Ile Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His
245 250 255
Leu Tyr Lys Gln Ile Ser Asn Gly Thr Ser Gly Gly Ser Thr Asn Asp
260 265 270

[0088]

Asn Thr Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn
275 280 285
Arg Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn
290 295 300
Asn Asn Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn
305 310 315 320
Ile Gln Val Lys Glu Val Thr Gln Asn Glu Gly Thr Lys Thr Ile Ala
325 330 335
Asn Asn Leu Thr Ser Thr Ile Gln Val Phe Thr Asp Ser Glu Tyr Gln
340 345 350
Leu Pro Tyr Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe
355 360 365
Pro Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn
370 375 380
Asn Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr
385 390 395 400
Phe Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Glu Phe Ser Tyr
405 410 415
Gln Phe Glu Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser
420 425 430
Leu Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu
435 440 445
Ser Arg Thr Gln Ser Thr Gly Gly Thr Ala Gly Thr Gln Gln Leu Leu
450 455 460
Phe Ser Gln Ala Gly Pro Asn Asn Met Ser Ala Gln Ala Lys Asn Trp
465 470 475 480
Leu Pro Gly Pro Cys Tyr Arg Gln Gln Arg Val Ser Thr Thr Leu Ser
485 490 495
Gln Asn Asn Asn Ser Asn Phe Ala Trp Thr Gly Ala Thr Lys Tyr His
[0089] 500 505 510
Leu Asn Gly Arg Asp Ser Leu Val Asn Pro Gly Val Ala Met Ala Thr
515 520 525
His Lys Asp Asp Glu Glu Arg Phe Phe Pro Ser Ser Gly Val Leu Met
530 535 540
Phe Gly Lys Gln Gly Ala Gly Lys Asp Asn Val Asp Tyr Ser Ser Val
545 550 555 560
Met Leu Thr Ser Glu Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr
565 570 575
Glu Gln Tyr Gly Val Val Ala Asp Asn Leu Gln Gln Gln Asn Ala Ala
580 585 590
Pro Ile Val Gly Ala Val Asn Ser Gln Gly Ala Leu Pro Gly Met Val
595 600 605
Trp Gln Asn Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile
610 615 620
Pro His Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe
625 630 635 640
Gly Leu Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val
645 650 655
Pro Ala Asp Pro Pro Thr Thr Phe Ser Gln Ala Lys Leu Ala Ser Phe
660 665 670
Ile Thr Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu
675 680 685
Leu Gln Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr
690 695 700
Ser Asn Tyr Tyr Lys Ser Thr Asn Val Asp Phe Ala Val Asn Thr Asp
705 710 715 720
Gly Thr Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg
725 730 735
Asn Leu

<210> 80
<211> 738
<212> PRT
<213> AAV重组型的衣壳蛋白,克隆44.5

<400> 80

[0090]

```
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 Asp Leu Lys Pro Gly Ala Pro Lys Pro
            20                  25                  30
Lys Ala Asn Gln Glu Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro
        35                  40                  45
Gly Tyr Lys Tyr Leu Gly Pro Phe 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 Arg Tyr Asn His Ala
                85                  90                  95
Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly
            100                 105                 110
Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
        115                 120                 125
Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg
    130                 135                 140
Pro Val Glu Pro Ser Pro Gln Arg Ser Pro Asp Ser Ser Thr Gly Ile
145                 150                 155                 160
Gly Lys Lys Gly Gln Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln
                165                 170                 175
Thr Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro
            180                 185                 190
Pro Ala Gly Pro Ser Gly Leu Gly Ser Gly Thr Met Ala Ala Gly Gly
        195                 200                 205
Gly Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser
    210                 215                 220
Ser Ser Gly Asn Trp His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val
225                 230                 235                 240
Ile Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His
                245                 250                 255
Leu Tyr Lys Gln Ile Ser Asn Gly Thr Ser Gly Gly Ser Thr Asn Asp
            260                 265                 270
Asn Thr Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn
        275                 280                 285
Arg Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn
    290                 295                 300
Asn Asn Trp Gly Phe Arg Pro Lys Arg Pro Asn Phe Lys Leu Phe Asn
305                 310                 315                 320
Ile Gln Val Lys Glu Val Thr Gln Asn Glu Gly Thr Lys Thr Ile Ala
                325                 330                 335
Asn Asn Leu Thr Ser Thr Ile Gln Val Phe Thr Asp Ser Glu Tyr Gln
            340                 345                 350
Leu Pro Tyr Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe
        355                 360                 365
Pro Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn
    370                 375                 380
Asn Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr
385                 390                 395                 400
Phe Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Glu Phe Ser Tyr
                405                 410                 415
Gln Phe Glu Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser
            420                 425                 430
Leu Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu
        435                 440                 445
```

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

[0091]

<210> 81
<211> 738
<212> PRT
<213> AAV血清型的衣壳蛋白,克隆44.2

<400> 81
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 Asp Leu Lys Pro Gly Ala Pro Lys Pro
20 25 30
Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro
35 40 45
Gly Tyr Lys Tyr Leu Gly Pro Phe 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 Arg Tyr Asn His Ala
85 90 95
Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly
100 105 110
Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
115 120 125
Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg

130 135 140
Pro Val Glu Pro Ser Pro Gln Arg Ser Pro Asp Ser Ser Thr Gly Ile
145 150 155 160
Gly Lys Lys Gly Gln Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln
165 170 175
Thr Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro
180 185 190
Pro Ala Gly Pro Ser Gly Leu Gly Ser Gly Thr Met Ala Ala Gly Gly
195 200 205
Gly Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser
210 215 220
Ser Ser Gly Asn Trp His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val
225 230 235 240
Ile Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His
245 250 255
Leu Tyr Lys Gln Ile Ser Asn Gly Thr Ser Gly Gly Ser Thr Asn Asp
260 265 270
Asn Thr Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn
275 280 285
Arg Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn
290 295 300
Asn Asn Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn
305 310 315 320
Ile Gln Val Lys Glu Val Thr Gln Asn Glu Gly Thr Lys Thr Ile Ala
325 330 335
Asn Asn Leu Thr Ser Thr Ile Gln Val Phe Thr Asp Ser Glu Tyr Gln
340 345 350
Leu Pro Tyr Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe
355 360 365
Pro Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn
370 375 380
[0092] Asn Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr
385 390 395 400
Phe Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Glu Phe Ser Tyr
405 410 415
Gln Phe Glu Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser
420 425 430
Leu Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu
435 440 445
Ser Arg Thr Gln Ser Thr Gly Gly Thr Ala Gly Thr Gln Gln Leu Leu
450 455 460
Phe Ser Gln Ala Gly Pro Asn Asn Met Ser Ala Gln Ala Lys Asn Trp
465 470 475 480
Leu Pro Gly Pro Cys Tyr Arg Gln Gln Arg Val Ser Thr Thr Leu Ser
485 490 495
Gln Asn Asn Asn Ser Asn Phe Ala Trp Thr Gly Ala Thr Lys Tyr His
500 505 510
Leu Asn Gly Arg Asp Ser Leu Val Asn Pro Gly Val Ala Met Ala Thr
515 520 525
His Lys Asp Asp Glu Glu Arg Phe Phe Pro Ser Ser Gly Val Leu Met
530 535 540
Phe Gly Lys Gln Gly Ala Gly Lys Asp Asn Val Asp Tyr Ser Ser Val
545 550 555 560
Met Leu Thr Ser Glu Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr
565 570 575
Glu Gln Tyr Gly Val Val Ala Asp Asn Leu Gln Gln Gln Asn Ala Ala
580 585 590
Pro Ile Val Gly Ala Val Asn Ser Gln Gly Ala Leu Pro Gly Met Val
595 600 605
Trp Gln Asn Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile
610 615 620
Pro His Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe

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

<210> 82
<211> 738
<212> PRT
<213> AAV 血清型的衣壳蛋白,克隆 29.3VP1

<400> 82
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 Lys Pro
20 25 30
Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro
35 40 45
Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro
50 55 60
[0093] 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 Arg Tyr Asn His Ala
85 90 95
Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly
100 105 110
Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
115 120 125
Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg
130 135 140
Pro Val Glu Pro Ser Pro Gln Arg Ser Pro Asp Ser Thr Thr Gly Ile
145 150 155 160
Gly Lys Lys Gly Gln Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln
165 170 175
Thr Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro
180 185 190
Pro Ala Gly Pro Ser Gly Leu Gly Ser Gly Thr Met Ala Ala Gly Gly
195 200 205
Gly Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser
210 215 220
Ser Ser Gly Asn Trp His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val
225 230 235 240
Ile Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His
245 250 255
Leu Tyr Lys Gln Ile Ser Asn Gly Thr Ser Gly Gly Ser Thr Asn Asp
260 265 270
Asn Thr Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn
275 280 285
Arg Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn
290 295 300
Asn Asn Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn
305 310 315 320

Ile Gln Val Lys Glu Val Thr Gln Asn Glu Gly Thr Lys Thr Ile Ala
325 330 335
Asn Asn Leu Thr Ser Thr Ile Gln Val Phe Thr Asp Ser Glu Tyr Gln
340 345 350
Leu Pro Tyr Val Leu Gly Ser Ala Arg Gln Gly Cys Leu Pro Pro Phe
355 360 365
Pro Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn
370 375 380
Asn Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr
385 390 395 400
Phe Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Glu Phe Ser Tyr
405 410 415
Gln Phe Glu Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser
420 425 430
Leu Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu
435 440 445
Ser Arg Thr Gln Ser Thr Gly Gly Thr Ala Gly Thr Gln Gln Leu Leu
450 455 460
Phe Ser Gln Ala Gly Pro Asn Asn Met Ser Ala Gln Ala Lys Asn Trp
465 470 475 480
Leu Pro Gly Pro Cys Tyr Arg Gln Gln Arg Val Ser Thr Thr Leu Ser
485 490 495
Gln Asn Asn Asn Ser Asn Phe Ala Trp Thr Gly Ala Thr Lys Tyr His
500 505 510
Leu Asn Gly Arg Asp Ser Leu Val Asn Pro Gly Val Ala Met Ala Thr
515 520 525
His Lys Asp Asp Glu Glu Arg Phe Phe Pro Ser Ser Gly Val Leu Met
530 535 540
Phe Gly Lys Gln Gly Ala Gly Lys Gly Asn Val Asp Tyr Ser Ser Val
545 550 555 560
[0094] Met Leu Thr Ser Glu Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr
565 570 575
Glu Gln Tyr Gly Val Val Ala Asp Asn Leu Gln Gln Gln Asn Ala Ala
580 585 590
Pro Ile Val Gly Ala Val Asn Ser Gln Gly Ala Leu Pro Gly Met Val
595 600 605
Trp Gln Asn Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile
610 615 620
Pro His Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe
625 630 635 640
Gly Leu Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val
645 650 655
Pro Ala Asp Pro Pro Thr Thr Phe Ser Gln Ala Lys Leu Ala Ser Phe
660 665 670
Ile Thr Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu
675 680 685
Leu Gln Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr
690 695 700
Ser Asn Tyr Tyr Lys Ser Thr Asn Val Asp Phe Ala Val Asn Thr Asp
705 710 715 720
Gly Thr Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg
725 730 735
Asn Leu

<210> 83
<211> 738
<212> PRT
<213> AAV 血清型的衣壳蛋白, 克隆 29.6VP1

<400> 83
Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser

[0095]

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

Pro Ala Gly Pro Ser Gly Leu Gly Ser Gly Thr Met Ala Ala Gly Gly
        195                 200                 205
Gly Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser
    210                 215                 220
Ser Ser Gly Asn Trp His Cys Asp Ser Thr Trp Leu Gly Asp Gly Val
225                 230                 235                 240
Ile Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His
                245                 250                 255
Leu Tyr Lys Gln Ile Ser Asn Gly Thr Ser Gly Gly Ser Thr Asn Asp
            260                 265                 270
Asn Thr Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn
        275                 280                 285
Arg Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn
    290                 295                 300
Asn Asn Trp Gly Phe Arg Pro Lys Ser Leu Asn Phe Lys Leu Phe Asn
305                 310                 315                 320
Ile Gln Val Lys Glu Val Thr Gln Asn Glu Gly Thr Lys Thr Ile Ala
                325                 330                 335
Asn Asn Leu Thr Ser Thr Ile Gln Val Phe Thr Asp Ser Glu Tyr Gln
            340                 345                 350
Leu Pro Tyr Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe
        355                 360                 365
Pro Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn
    370                 375                 380
Asn Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr
385                 390                 395                 400
Phe Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Glu Phe Ser Tyr
                405                 410                 415
Gln Phe Glu Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser
            420                 425                 430
Leu Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu
        435                 440                 445
Ser Arg Thr Gln Ser Thr Gly Gly Thr Ala Gly Thr Gln Gln Leu Leu
    450                 455                 460
Phe Ser Gln Ala Gly Pro Asn Asn Met Ser Ala Gln Ala Lys Asn Trp
465                 470                 475                 480
Leu Pro Gly Pro Cys Tyr Arg Gln Gln Arg Val Ser Thr Thr Leu Ser
                485                 490                 495
```

```
Gln Asn Asp Asn Ser Asn Phe Ala Trp Thr Gly Ala Thr Lys Tyr His
            500                 505                 510
Leu Asn Gly Arg Asp Ser Leu Val Asn Pro Gly Val Ala Met Ala Thr
        515                 520                 525
His Lys Asp Asp Glu Glu Arg Phe Phe Pro Ser Ser Gly Val Leu Met
    530                 535                 540
Phe Gly Lys Gln Gly Ala Gly Lys Asp Asn Val Asp Tyr Ser Ser Val
545                 550                 555                 560
Met Leu Thr Ser Glu Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr
                565                 570                 575
Glu Gln Tyr Gly Val Val Ala Asp Asn Leu Gln Gln Gln Asn Ala Ala
            580                 585                 590
Pro Ile Val Gly Ala Val Asn Ser Gln Gly Ala Leu Pro Gly Met Val
        595                 600                 605
Trp Gln Asn Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile
    610                 615                 620
Pro His Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe
625                 630                 635                 640
Gly Leu Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val
                645                 650                 655
Pro Ala Asp Pro Pro Thr Thr Phe Ser Gln Ala Lys Leu Ala Ser Phe
            660                 665                 670
Ile Thr Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu
        675                 680                 685
Leu Gln Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr
    690                 695                 700
Ser Asn Tyr Tyr Lys Ser Thr Asn Val Asp Phe Ala Val Asn Thr Asp
705                 710                 715                 720
Gly Thr Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg
                725                 730                 735
Asn Leu
```

[0096]

<210> 84
<211> 738
<212> PRT
<213> AAV 血清型的衣壳蛋白, 克隆 42.15

<400> 84

```
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 Asp Leu Lys Pro Gly Ala Pro Lys Pro
            20                  25                  30
Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro
        35                  40                  45
Gly Tyr Lys Tyr Leu Gly Pro Phe 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 Arg Tyr Asn His Ala
                85                  90                  95
Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly
            100                 105                 110
Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
        115                 120                 125
Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg
    130                 135                 140
Pro Val Glu Pro Ser Pro Gln Arg Ser Pro Asp Ser Ser Thr Gly Ile
145                 150                 155                 160
Gly Lys Thr Gly Gln Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln
                165                 170                 175
Thr Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro
```

[0097]

```
                180                 185                 190
Pro Ala Gly Pro Ser Gly Leu Gly Ser Gly Thr Met Ala Ala Gly Gly
        195                 200                 205
Gly Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser
    210                 215                 220
Ser Ser Gly Asn Trp His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val
225                 230                 235                 240
Ile Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His
                245                 250                 255
Leu Tyr Lys Gln Ile Ser Asn Gly Thr Ser Gly Gly Ser Thr Asn Asp
            260                 265                 270
Asn Thr Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn
        275                 280                 285
Arg Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn
    290                 295                 300
Asn Asn Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn
305                 310                 315                 320
Ile Gln Val Lys Glu Val Thr Gln Asn Glu Gly Thr Lys Thr Ile Ala
                325                 330                 335
Asn Asn Leu Thr Ser Thr Ile Gln Val Phe Thr Asp Ser Glu Tyr Gln
            340                 345                 350
Leu Pro Tyr Val Leu Gly Ser Ala His Gln Gly Cys Pro Pro Pro Phe
        355                 360                 365
Pro Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn
    370                 375                 380
Asn Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr
385                 390                 395                 400
Phe Pro Ser Gln Met Arg Arg Thr Gly Asn Asn Phe Glu Phe Ser Tyr
                405                 410                 415
Gln Phe Glu Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser
            420                 425                 430
Leu Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu
        435                 440                 445
Ser Arg Thr Gln Ser Thr Gly Gly Thr Ala Gly Thr Gln Gln Leu Leu
    450                 455                 460
Phe Ser Gln Ala Gly Pro Asn Asn Met Ser Ala Gln Ala Lys Asn Trp
465                 470                 475                 480
Leu Pro Gly Pro Cys Tyr Arg Gln Gln Arg Val Ser Thr Thr Leu Ser
                485                 490                 495
Gln Asn Asn Asn Ser Asn Phe Ala Trp Thr Gly Ala Thr Lys Tyr His
            500                 505                 510
Leu Asn Gly Arg Asp Ser Leu Val Asn Pro Gly Val Ala Met Ala Thr
        515                 520                 525
His Lys Asp Asp Glu Glu Arg Phe Phe Pro Ser Ser Gly Val Leu Met
    530                 535                 540
Phe Gly Lys Gln Gly Ala Gly Lys Asp Asn Val Asp Tyr Ser Ser Val
545                 550                 555                 560
Met Leu Thr Ser Glu Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr
                565                 570                 575
Glu Gln Tyr Gly Val Val Ala Asp Asn Leu Gln Gln Gln Asn Ala Ala
            580                 585                 590
Pro Ile Val Gly Ala Val Asn Ser Gln Gly Ala Leu Pro Gly Met Val
        595                 600                 605
Trp Gln Asn Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile
    610                 615                 620
Pro His Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe
625                 630                 635                 640
Gly Leu Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val
                645                 650                 655
Pro Ala Asp Pro Pro Thr Thr Phe Ser Gln Ala Lys Leu Ala Ser Phe
            660                 665                 670
Ile Thr Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu
```

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

<210> 85
<211> 738
<212> PRT
<213> AAV 血清型的衣壳蛋白,克隆 42.8

<400> 85
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 Asp Leu Lys Pro Gly Ala Pro Lys Pro
20 25 30
Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro
35 40 45
Gly Tyr Lys Tyr Leu Gly Pro Phe 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 Arg Tyr Asn His Ala
85 90 95
Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly
100 105 110
Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
115 120 125
Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg
130 135 140
Pro Val Glu Pro Ser Pro Gln Arg Ser Pro Asp Ser Ser Thr Gly Ile
145 150 155 160
Gly Lys Thr Gly Gln Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln
165 170 175
Thr Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro
180 185 190
Pro Ala Gly Pro Ser Gly Leu Gly Ser Gly Thr Met Ala Ala Gly Gly
195 200 205
Gly Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser
210 215 220
Ser Ser Gly Asn Trp His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val
225 230 235 240
Ile Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His
245 250 255
Leu Tyr Lys Gln Ile Ser Asn Gly Thr Ser Gly Gly Ser Thr Asn Asp
260 265 270
Asn Thr Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn
275 280 285
Arg Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn
290 295 300
Asn Asn Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn
305 310 315 320
Ile Gln Val Lys Glu Val Thr Gln Asn Glu Gly Thr Lys Thr Ile Ala
325 330 335
Asn Asn Leu Thr Ser Thr Ile Gln Val Phe Thr Asp Ser Glu Tyr Gln
340 345 350
Leu Pro Tyr Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe
355 360 365

[0098]

Pro Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn
370 375 380
Asn Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr
385 390 395 400
Phe Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Glu Phe Ser Tyr
405 410 415
Gln Phe Glu Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser
420 425 430
Leu Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu
435 440 445
Ser Arg Thr Gln Ser Thr Gly Gly Thr Ala Gly Thr Gln Gln Leu Leu
450 455 460
Phe Ser Gln Ala Gly Pro Asn Asn Met Ser Ala Gln Ala Lys Asn Trp
465 470 475 480
Leu Pro Gly Pro Cys Tyr Arg Gln Gln Arg Val Ser Thr Thr Leu Ser
485 490 495
Gln Asn Asn Asn Ser Asn Phe Ala Trp Thr Gly Ala Thr Lys Tyr His
500 505 510
Leu Asn Gly Arg Asp Ser Leu Val Asn Pro Gly Val Ala Met Ala Thr
515 520 525
His Lys Asp Asp Glu Glu Arg Phe Phe Pro Ser Ser Gly Val Leu Met
530 535 540
Phe Gly Lys Gln Gly Ala Gly Lys Asp Asn Val Asp Tyr Ser Ser Val
545 550 555 560
Met Leu Thr Ser Glu Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr
565 570 575
Glu Gln Tyr Gly Val Val Ala Asp Asn Leu Gln Gln Gln Asn Ala Ala
580 585 590
Pro Ile Val Gly Ala Val Asn Ser Gln Gly Ala Leu Pro Gly Met Val
595 600 605
Trp Gln Asn Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile
610 615 620
Pro His Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe
625 630 635 640
Gly Leu Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val
645 650 655
Pro Ala Asp Pro Pro Thr Thr Phe Ser Gln Ala Lys Leu Ala Ser Phe
660 665 670
Ile Thr Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu
675 680 685
Leu Gln Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr
690 695 700
Ser Asn Tyr Tyr Lys Ser Thr Asn Val Asp Phe Ala Val Asn Thr Glu
705 710 715 720
Gly Thr Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg
725 730 735
Asn Leu

[0099]

<210> 86
<211> 733
<212> PRT
<213> AAV 血清型的氨基酸, 克隆 42.13

<400> 86
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 Asp Leu Lys Pro Gly Ala Pro Lys Pro
20 25 30
Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro
35 40 45
Gly Tyr Lys Tyr Leu Gly Pro Phe 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 Arg Tyr Asn His Ala
                85                  90                  95
Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly
            100                 105                 110
Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
        115                 120                 125
Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg
    130                 135                 140
Pro Ile Glu Ser Pro Asp Ser Ser Thr Gly Ile Gly Lys Lys Gly Gln
145                 150                 155                 160
Gln Pro Ala Lys Lys Lys Leu Asn Phe Gly Gln Thr Gly Asp Ser Glu
                165                 170                 175
Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro Ala Gly Pro Ser
            180                 185                 190
Gly Leu Gly Ser Gly Thr Met Ala Ala Gly Gly Gly Ala Pro Met Ala
        195                 200                 205
Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser Ser Ser Gly Asn Trp
    210                 215                 220
His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val Ile Thr Thr Ser Thr
225                 230                 235                 240
Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu Tyr Lys Gln Ile
                245                 250                 255
Ser Asn Gly Thr Ser Gly Gly Ser Thr Asn Asp Asn Thr Tyr Phe Gly
            260                 265                 270
Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His Cys His
        275                 280                 285
Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Trp Gly Phe
    290                 295                 300
Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile Gln Val Lys Glu
305                 310                 315                 320
Val Thr Gln Asn Glu Gly Thr Lys Thr Ile Ala Asn Asn Leu Thr Ser
                325                 330                 335
Thr Ile Gln Val Phe Thr Asp Ser Glu Tyr Gln Leu Pro Tyr Val Leu
            340                 345                 350
Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala Asp Val Phe
        355                 360                 365
Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly Ser Gln Ala
    370                 375                 380
Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser Gln Met
385                 390                 395                 400
Leu Arg Thr Gly Asn Asn Phe Glu Phe Ser Tyr Gln Phe Glu Asp Val
                405                 410                 415
Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp Arg Leu Met
            420                 425                 430
Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser Arg Thr Gln Ser
        435                 440                 445
Thr Gly Gly Thr Ala Gly Thr Gln Gln Leu Leu Phe Ser Gln Ala Gly
    450                 455                 460
Pro Asn Asn Met Ser Ala Gln Ala Lys Asn Trp Leu Pro Gly Pro Cys
465                 470                 475                 480
Tyr Arg Gln Gln Arg Val Ser Thr Thr Val Ser Gln Asn Asn Asn Ser
                485                 490                 495
Asn Phe Ala Trp Thr Gly Ala Thr Lys Tyr His Leu Asn Gly Arg Asp
            500                 505                 510
Ser Leu Val Asn Pro Gly Val Ala Met Ala Thr His Lys Gly Asp Glu
        515                 520                 525
Glu Arg Phe Phe Pro Ser Ser Gly Val Leu Met Phe Gly Lys Gln Gly
    530                 535                 540
Ala Gly Lys Asp Asn Val Asp Tyr Ser Ser Val Met Leu Thr Ser Glu
```

[0100]

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

<210> 87
<211> 733
<212> PRT
<213> AAV 重组型的衣壳蛋白,克隆 42.3A

<400> 87
[0101]
Met Ala Ala Asp Gly His Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser
1 5 10 15
Glu Gly Ile Arg Glu Trp Trp Asp Leu Lys Pro Gly Ala Pro Lys Pro
20 25 30
Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro
35 40 45
Gly Tyr Lys Tyr Leu Gly Pro Phe 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 Arg Tyr Asn His Ala
85 90 95
Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly
100 105 110
Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
115 120 125
Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg
130 135 140
Pro Ile Glu Ser Pro Asp Ser Ser Thr Gly Ile Gly Lys Lys Gly Gln
145 150 155 160
Gln Pro Ala Lys Lys Lys Leu Asn Phe Gly Gln Thr Gly Asp Ser Glu
165 170 175
Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro Ala Gly Pro Ser
180 185 190
Gly Leu Gly Ser Gly Thr Met Ala Ala Gly Gly Gly Ala Pro Met Ala
195 200 205
Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser Ser Ser Gly Asn Trp
210 215 220
His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val Ile Thr Thr Ser Thr
225 230 235 240
Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu Tyr Lys Gln Ile
245 250 255

[0102]

```
Ser Asn Gly Thr Ser Gly Gly Ser Thr Asn Asp Asn Thr Tyr Phe Gly
            260                 265                 270
Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His Cys His
        275                 280                 285
Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Ser Trp Gly Phe
    290                 295                 300
Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile Gln Val Lys Glu
305                 310                 315                 320
Val Thr Gln Asn Glu Gly Thr Lys Thr Ile Ala Asn Asn Leu Thr Ser
                325                 330                 335
Thr Ile Gln Val Phe Thr Asp Ser Glu Tyr Gln Leu Pro Tyr Val Leu
            340                 345                 350
Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala Asp Val Phe
        355                 360                 365
Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly Ser Gln Ala
    370                 375                 380
Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser Gln Met
385                 390                 395                 400
Leu Arg Thr Gly Asn Asn Phe Glu Phe Ser Tyr Gln Phe Glu Asp Val
                405                 410                 415
Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp Arg Leu Met
            420                 425                 430
Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser Arg Thr Gln Ser
        435                 440                 445
Thr Gly Gly Thr Ala Gly Thr Gln Gln Leu Leu Phe Ser Gln Ala Gly
    450                 455                 460
Pro Asn Asn Met Ser Ala Gln Ala Lys Asn Trp Leu Pro Gly Pro Cys
465                 470                 475                 480
Tyr Arg Gln Gln Arg Val Ser Thr Thr Leu Ser Gln Asn Asn Asn Ser
                485                 490                 495
Asn Phe Ala Trp Thr Gly Ala Thr Lys Tyr His Leu Asn Gly Arg Asp
            500                 505                 510
Ser Leu Val Asn Pro Gly Val Ala Met Ala Thr His Lys Asp Asp Glu
        515                 520                 525
Glu Arg Phe Phe Pro Ser Ser Gly Val Leu Met Phe Gly Lys Gln Gly
    530                 535                 540
Ala Gly Lys Asp Asn Val Asp Tyr Ser Ser Val Met Leu Thr Ser Glu
545                 550                 555                 560
Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr Glu Gln Tyr Gly Val
                565                 570                 575
Val Ala Asp Asn Leu Gln Gln Gln Asn Ala Ala Pro Ile Val Gly Ala
            580                 585                 590
Val Asn Ser Gln Gly Ala Leu Pro Gly Met Val Trp Gln Asn Arg Asp
        595                 600                 605
Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His Thr Asp Gly
    610                 615                 620
Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Leu Lys His Pro
625                 630                 635                 640
Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala Asp Pro Pro
                645                 650                 655
Thr Thr Phe Ser Gln Ala Lys Leu Ala Ser Phe Ile Thr Gln Tyr Ser
            660                 665                 670
Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln Lys Glu Asn
        675                 680                 685
Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn Tyr Tyr Lys
    690                 695                 700
Ser Thr Asn Val Asp Phe Ala Val Asn Thr Glu Gly Thr Tyr Ser Glu
705                 710                 715                 720
Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu
                725                 730
```

<210> 88
<211> 731
<212> PRT
<213> AAV 血清型的衣壳蛋白, 克隆 42.8

<400> 88

[0103]

```
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 Asp Leu Lys Pro Gly Ala Pro Lys Pro
            20                  25                  30
Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro
        35                  40                  45
Gly Tyr Lys Tyr Leu Gly Pro Phe 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
Lys Gln Leu Glu Gln Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala
                85                  90                  95
Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly
            100                 105                 110
Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
        115                 120                 125
Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg
    130                 135                 140
Pro Ile Glu Ser Pro Asp Ser Ser Thr Gly Ile Gly Lys Lys Gly Gln
145                 150                 155                 160
Gln Pro Ala Lys Lys Lys Leu Asn Phe Gly Gln Thr Gly Asp Ser Glu
                165                 170                 175
Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro Ala Gly Pro Ser
            180                 185                 190
Gly Leu Gly Ser Gly Thr Met Ala Ala Gly Gly Gly Ala Pro Met Ala
        195                 200                 205
Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ala Ser Gly Asn Trp
    210                 215                 220
His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val Ile Thr Thr Ser Thr
225                 230                 235                 240
Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu Tyr Lys Gln Ile
                245                 250                 255
Ser Ser Gln Ser Gly Ala Thr Asn Asp Asn His Phe Phe Gly Tyr Ser
            260                 265                 270
Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His Cys His Phe Ser
        275                 280                 285
Ser Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Trp Gly Phe Arg Pro
    290                 295                 300
Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile Gln Val Lys Glu Val Thr
305                 310                 315                 320
Gln Asn Glu Gly Thr Lys Thr Ile Ala Asn Asn Leu Thr Ser Thr Ile
                325                 330                 335
Gln Val Phe Thr Asp Ser Glu Tyr Arg Leu Pro Tyr Val Leu Gly Ser
            340                 345                 350
Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala Asp Val Phe Met Ile
        355                 360                 365
Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly Ser Gln Ala Val Gly
    370                 375                 380
Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser Gln Met Leu Arg
385                 390                 395                 400
Thr Gly Asn Asn Phe Glu Phe Ser Tyr Gln Phe Glu Asp Val Pro Phe
                405                 410                 415
His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp Arg Leu Met Asn Pro
            420                 425                 430
Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser Arg Thr Gln Ser Thr Gly
        435                 440                 445
```

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

[0104]

<210> 89
<211> 731
<212> PRT
<213> AAV血清型的衣壳蛋白,克隆42.5A

<400> 89
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 Asp Leu Lys Pro Gly Ala Pro Lys Pro
20 25 30
Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro
35 40 45
Gly Tyr Lys Tyr Leu Gly Pro Phe 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
Lys Gln Leu Glu Gln Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala
85 90 95
Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly
100 105 110
Asn Leu Gly Arg Ala Val Phe Arg Ala Lys Lys Arg Val Leu Glu Pro
115 120 125
Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg
130 135 140
Pro Ile Glu Ser Pro Asp Ser Ser Thr Gly Ile Gly Lys Lys Gly Gln

145 150 155 160
Gln Pro Ala Lys Lys Lys Leu Asn Phe Gly Gln Thr Gly Asp Ser Glu
165 170 175
Ser Val Pro Asp Pro Gln Pro Leu Gly Glu Pro Pro Ala Ala Pro Ser
180 185 190
Gly Leu Gly Ser Gly Thr Met Ala Ala Gly Gly Gly Ala Pro Met Ala
195 200 205
Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ala Ser Gly Asn Trp
210 215 220
His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val Ile Thr Thr Ser Thr
225 230 235 240
Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu Tyr Lys Gln Ile
245 250 255
Ser Ser Gln Ser Gly Ala Thr Asn Asp Asn His Phe Phe Gly Tyr Ser
260 265 270
Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His Cys His Phe Ser
275 280 285
Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Arg Gly Phe Arg Pro
290 295 300
Arg Lys Leu Arg Phe Lys Leu Phe Asn Ile Gln Val Lys Glu Val Thr
305 310 315 320
Thr Asn Asp Gly Val Thr Thr Ile Ala Asn Asn Leu Thr Ser Thr Ile
325 330 335
Gln Val Phe Ser Asp Ser Glu Tyr Gln Leu Pro Tyr Val Leu Gly Ser
340 345 350
Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala Asp Val Phe Met Ile
355 360 365
Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly Ser Gln Ser Val Gly
370 375 380
Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser Gln Met Leu Arg
385 390 395 400
[0105] Thr Gly Asn Asn Phe Glu Phe Ser Tyr Gln Phe Glu Asp Val Pro Phe
405 410 415
His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp Arg Leu Met Asn Pro
420 425 430
Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser Arg Thr Gln Ser Thr Gly
435 440 445
Gly Thr Ala Gly Thr Gln Gln Leu Leu Phe Ser Gln Ala Gly Pro Asn
450 455 460
Asn Met Ser Ala Gln Ala Lys Asn Trp Leu Pro Gly Pro Cys Tyr Arg
465 470 475 480
Gln Gln Arg Val Ser Thr Thr Leu Ser Gln Asn Asn Asn Ser Asn Phe
485 490 495
Ala Trp Thr Gly Ala Thr Lys Tyr His Leu Asn Gly Arg Asp Ser Leu
500 505 510
Val Asn Pro Gly Val Ala Met Ala Thr His Lys Asp Asp Glu Glu Arg
515 520 525
Phe Phe Pro Ser Ser Gly Val Leu Met Phe Gly Lys Gln Gly Ala Gly
530 535 540
Lys Asp Asn Val Asp Tyr Ser Ser Val Met Leu Thr Ser Glu Glu Glu
545 550 555 560
Ile Lys Thr Thr Asn Pro Val Ala Thr Glu Gln Tyr Gly Val Val Ala
565 570 575
Asp Asn Leu Gln Gln Gln Asn Ala Ala Pro Ile Val Gly Ala Val Asn
580 585 590
Ser Gln Gly Ala Leu Pro Gly Met Ala Trp Gln Asn Arg Asp Val Tyr
595 600 605
Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His Thr Asp Gly Asn Phe
610 615 620
His Pro Ser Pro Leu Met Gly Gly Phe Gly Leu Lys His Pro Pro Pro
625 630 635 640
Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala Asp Pro Pro Thr Thr

645 650 655
Phe Ser Gln Ala Lys Leu Ala Ser Phe Ile Thr Gln Tyr Ser Thr Gly
660 665 670
Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln Lys Glu Asn Ser Lys
675 680 685
Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn Tyr Tyr Lys Ser Thr
690 695 700
Asn Val Asp Phe Ala Val Asn Thr Glu Gly Thr Tyr Ser Glu Pro Arg
705 710 715 720
Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu
725 730

<210> 90
<211> 733
<212> PRT
<213> AAV 血清型的衣壳蛋白, 克隆 42.1B

<400> 90
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 Asp Leu Arg Pro Gly Ala Pro Lys Pro
20 25 30
Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro
35 40 45
Gly Tyr Lys Tyr Leu Gly Pro Phe 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
Lys Gln Leu Glu Gln Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala
85 90 95
[0106] Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly
100 105 110
Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
115 120 125
Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg
130 135 140
Pro Ile Glu Ser Pro Asp Ser Ser Thr Gly Ile Gly Lys Lys Gly Gln
145 150 155 160
Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr Gly Asp Ser Glu
165 170 175
Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro Ala Gly Pro Ser
180 185 190
Gly Leu Gly Ser Gly Thr Met Ala Ala Gly Gly Gly Ala Pro Met Ala
195 200 205
Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser Ser Ser Gly Asn Trp
210 215 220
His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val Ile Thr Thr Ser Thr
225 230 235 240
Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu Tyr Lys Gln Ile
245 250 255
Ser Asn Gly Thr Ser Gly Gly Ser Thr Asn Asp Asn Thr Tyr Phe Gly
260 265 270
Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His Cys His
275 280 285
Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Trp Gly Phe
290 295 300
Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile Gln Val Lys Glu
305 310 315 320
Val Thr Gln Asn Glu Gly Thr Lys Thr Ile Ala Asn Asn Leu Thr Ser
325 330 335
Thr Ile Gln Val Phe Thr Asp Ser Glu Tyr Gln Leu Pro Tyr Val Leu
340 345 350

Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala Asp Val Phe
355 360 365
Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly Ser Gln Ala
370 375 380
Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser Gln Met
385 390 395 400
Leu Arg Thr Gly Asn Asn Phe Glu Phe Ser Tyr Gln Phe Glu Asp Val
405 410 415
Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp Arg Leu Met
420 425 430
Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser Arg Thr Gln Ser
435 440 445
Thr Gly Gly Thr Ala Gly Thr Gln Gln Leu Leu Phe Ser Gln Ala Gly
450 455 460
Pro Asn Asn Met Ser Ala Gln Ala Lys Asn Trp Leu Pro Gly Pro Cys
465 470 475 480
Tyr Arg Gln Gln Arg Val Ser Thr Thr Val Ser Gln Asn Asn Asn Ser
485 490 495
Asn Phe Ala Trp Thr Gly Ala Thr Lys Tyr His Leu Asn Gly Arg Asp
500 505 510
Ser Leu Val Asn Pro Gly Val Ala Met Ala Thr His Lys Gly Asp Glu
515 520 525
Glu Arg Phe Phe Pro Ser Ser Gly Val Leu Met Phe Gly Lys Gln Gly
530 535 540
Ala Gly Lys Asp Asn Val Asp Tyr Ser Ser Val Met Leu Thr Ser Glu
545 550 555 560
Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr Glu Gln Tyr Gly Val
565 570 575
Val Ala Asp Asn Leu Gln Gln Gln Asn Ala Ala Pro Ile Val Gly Ala
580 585 590
Val Asn Ser Gln Gly Ala Leu Pro Gly Met Val Trp Gln Asn Arg Asp
595 600 605
Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His Thr Asp Gly
610 615 620
Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Leu Lys His Pro
625 630 635 640
Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala Asp Pro Pro
645 650 655
Thr Thr Phe Ser Gln Ala Lys Leu Ala Ser Phe Ile Thr Gln Tyr Ser
660 665 670
Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln Lys Glu Asn
675 680 685
Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn Tyr Tyr Lys
690 695 700
Ser Thr Asn Val Asp Phe Ala Val Asn Thr Glu Gly Thr Tyr Ser Glu
705 710 715 720
Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu
725 730

[0107]

<210> 91
<211> 738
<212> PRT
<213> AAV 血清型的衣壳蛋白, 克隆 42.5B

<400> 91
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 Asp Leu Lys Pro Gly Ala Pro Lys Pro
20 25 30
Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro
35 40 45
Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro

[0108]

```
    50                  55                  60
Val Asn Glu Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp
65                  70                  75                  80
Lys Gln Leu Glu Gln Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala
                85                  90                  95
Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly
            100                 105                 110
Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
        115                 120                 125
Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg
    130                 135                 140
Pro Val Glu Pro Ser Pro Gln Arg Ser Pro Asp Ser Ser Thr Gly Ile
145                 150                 155                 160
Gly Lys Thr Gly Gln Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln
                165                 170                 175
Thr Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro
            180                 185                 190
Pro Ala Gly Pro Ser Gly Leu Gly Ser Gly Thr Met Ala Ala Gly Gly
        195                 200                 205
Gly Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser
    210                 215                 220
Ser Ser Gly Asn Trp His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val
225                 230                 235                 240
Ile Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His
                245                 250                 255
Leu Tyr Lys Gln Ile Ser Asn Gly Thr Ser Gly Gly Ser Thr Asn Asp
            260                 265                 270
Asn Thr Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn
        275                 280                 285
Arg Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn
    290                 295                 300
Asn Asn Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn
305                 310                 315                 320
Ile Gln Val Lys Glu Val Thr Gln Asn Glu Gly Thr Lys Thr Ile Ala
                325                 330                 335
Asn Asn Leu Thr Ser Thr Ile Gln Val Phe Thr Asp Ser Glu Tyr Gln
            340                 345                 350
Leu Pro Tyr Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe
        355                 360                 365
Pro Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn
    370                 375                 380
Asn Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr
385                 390                 395                 400
Phe Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Glu Phe Ser Tyr
                405                 410                 415
Gln Phe Glu Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser
            420                 425                 430
Leu Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu
        435                 440                 445
Ser Arg Thr Gln Ser Thr Gly Gly Thr Ala Gly Thr Gln Gln Leu Leu
    450                 455                 460
Phe Ser Gln Ala Gly Pro Asn Asn Met Ser Ala Gln Ala Lys Asn Trp
465                 470                 475                 480
Leu Pro Gly Pro Cys Tyr Arg Gln Gln Arg Val Ser Thr Thr Leu Ser
                485                 490                 495
Gln Asn Asn Asn Ser Asn Phe Ala Trp Thr Gly Ala Thr Lys Tyr His
            500                 505                 510
Leu Asn Gly Arg Asp Ser Leu Val Asn Pro Gly Val Ala Met Ala Thr
        515                 520                 525
His Lys Asp Asp Glu Glu Arg Phe Phe Pro Ser Ser Gly Val Leu Met
    530                 535                 540
Phe Gly Lys Gln Gly Ala Gly Lys Asp Asn Val Asp Tyr Ser Ser Val
```

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

<210> 92
<211> 738
<212> PRT
<213> AAV 血清型的衣壳蛋白, 克隆 43.1

[0109]

```
<400> 92
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 Asp Leu Lys Pro Gly Ala Pro Lys Pro
            20                  25                  30
Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro
        35                  40                  45
Gly Tyr Lys Tyr Leu Gly Pro Phe 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 Arg Tyr Asn His Ala
                85                  90                  95
Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly
            100                 105                 110
Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
        115                 120                 125
Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg
    130                 135                 140
Pro Val Glu Pro Ser Pro Gln Arg Ser Pro Asp Ser Ser Thr Gly Ile
145                 150                 155                 160
Gly Lys Lys Gly His Gln Pro Ala Arg Lys Arg Leu Asn Phe Gly Gln
                165                 170                 175
Thr Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro
            180                 185                 190
Pro Ala Gly Pro Ser Gly Leu Gly Ser Gly Thr Met Ala Ala Gly Gly
        195                 200                 205
Gly Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser
    210                 215                 220
Ser Ser Gly Asn Trp His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val
225                 230                 235                 240
```

Ile Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His
245 250 255
Leu Tyr Lys Gln Ile Ser Asn Gly Thr Ser Gly Gly Ser Thr Asn Asp
260 265 270
Asn Thr Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn
275 280 285
Arg Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn
290 295 300
Asn Asn Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn
305 310 315 320
Ile Gln Val Lys Glu Val Thr Gln Asn Glu Gly Thr Lys Thr Ile Ala
325 330 335
Asn Asn Leu Thr Ser Thr Ile Gln Val Phe Thr Asp Ser Glu Tyr Gln
340 345 350
Leu Pro Tyr Val Pro Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe
355 360 365
Pro Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn
370 375 380
Asn Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr
385 390 395 400
Phe Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Glu Phe Ser Tyr
405 410 415
Thr Phe Glu Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser
420 425 430
Leu Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu
435 440 445
Ser Arg Thr Gln Ser Thr Gly Gly Thr Gln Gly Thr Gln Gln Leu Leu
450 455 460
Phe Ser Gln Ala Gly Pro Ala Asn Met Ser Ala Gln Ala Lys Asn Trp
465 470 475 480
[0110] Leu Pro Gly Pro Cys Tyr Arg Gln Gln Arg Val Ser Thr Thr Leu Ser
485 490 495
Gln Asn Asn Asn Ser Asn Phe Ala Trp Thr Gly Ala Thr Lys Tyr His
500 505 510
Leu Asn Gly Arg Asp Ser Leu Val Asn Pro Gly Val Ala Met Ala Thr
515 520 525
His Lys Asp Asp Glu Gln Arg Phe Phe Pro Ser Ser Gly Val Leu Met
530 535 540
Phe Gly Lys Gln Gly Ala Gly Lys Asp Asn Val Asp Tyr Ser Ser Val
545 550 555 560
Met Leu Thr Ser Glu Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr
565 570 575
Glu Gln Tyr Gly Val Val Ala Asp Asn Leu Gln Gln Thr Asn Gly Ala
580 585 590
Pro Ile Val Gly Thr Val Asn Ser Gln Gly Ala Leu Pro Gly Met Val
595 600 605
Trp Gln Asn Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile
610 615 620
Pro His Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe
625 630 635 640
Gly Leu Lys His Pro Pro Pro Gln Ile Leu Val Lys Asn Thr Pro Val
645 650 655
Pro Ala Asp Pro Pro Thr Thr Phe Ser Gln Ala Lys Leu Ala Ser Phe
660 665 670
Ile Thr Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu
675 680 685
Leu Gln Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr
690 695 700
Ser Asn Tyr Tyr Lys Ser Thr Asn Val Asp Phe Ala Val Asn Thr Glu
705 710 715 720
Gly Thr Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg
725 730 735

Asn Leu

<210> 93
<211> 738
<212> PRT
<213> AAV 血清型的衣壳蛋白, 克隆 43.12

<400> 93
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 Asp Leu Lys Pro Gly Ala Pro Lys Pro
20 25 30
Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro
35 40 45
Gly Tyr Lys Tyr Leu Gly Pro Phe 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 Arg Tyr Asn His Ala
85 90 95
Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly
100 105 110
Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
115 120 125
Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg
130 135 140
Pro Val Glu Pro Ser Pro Gln Arg Ser Pro Asp Ser Ser Thr Gly Ile
145 150 155 160
[0111] Gly Lys Lys Gly His Gln Pro Ala Arg Lys Arg Leu Asn Phe Gly Gln
165 170 175
Thr Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro
180 185 190
Pro Ala Gly Pro Ser Gly Leu Gly Ser Gly Thr Met Ala Ala Gly Gly
195 200 205
Gly Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser
210 215 220
Ser Ser Gly Asn Trp His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val
225 230 235 240
Ile Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His
245 250 255
Leu Tyr Lys Gln Ile Ser Asn Gly Thr Ser Gly Gly Ser Thr Asn Asp
260 265 270
Asn Thr Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn
275 280 285
Arg Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn
290 295 300
Asn Asn Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn
305 310 315 320
Ile Gln Val Lys Glu Val Thr Gln Asn Glu Gly Thr Lys Thr Ile Ala
325 330 335
Asn Asn Leu Thr Ser Thr Ile Gln Val Phe Thr Asp Ser Glu Tyr Gln
340 345 350
Leu Pro Tyr Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe
355 360 365
Pro Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn
370 375 380
Asn Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr
385 390 395 400
Phe Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Glu Phe Ser Tyr

405 410 415
Thr Phe Glu Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser
420 425 430
Leu Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu
435 440 445
Ser Arg Thr Gln Ser Thr Gly Gly Thr Gln Gly Thr Gln Gln Leu Leu
450 455 460
Phe Ser Gln Ala Gly Pro Ala Asn Met Ser Ala Gln Ala Lys Asn Trp
465 470 475 480
Leu Pro Gly Pro Cys Tyr Arg Gln Gln Arg Val Ser Thr Thr Leu Ser
485 490 495
Gln Asn Asn Asn Ser Asn Phe Ala Trp Thr Gly Ala Thr Lys Tyr His
500 505 510
Leu Asn Gly Arg Asp Ser Leu Val Asn Pro Gly Val Ala Met Ala Thr
515 520 525
His Lys Asp Asp Glu Glu Arg Phe Phe Pro Ser Ser Gly Val Leu Met
530 535 540
Phe Gly Lys Gln Gly Ala Gly Lys Asp Asn Val Asp Tyr Ser Ser Val
545 550 555 560
Met Leu Thr Ser Glu Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr
565 570 575
Glu Gln Tyr Gly Val Val Ala Asp Asn Leu Gln Gln Thr Asn Gly Ala
580 585 590
Pro Ile Val Gly Thr Val Asn Ser Gln Gly Ala Leu Pro Gly Met Val
595 600 605
Trp Gln Asn Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile
610 615 620
Pro His Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe
625 630 635 640
Gly Leu Lys His Pro Pro Pro Gln Ile Leu Val Lys Asn Thr Pro Val
645 650 655
[0112] Pro Ala Asp Pro Pro Thr Thr Phe Ser Gln Ala Lys Leu Ala Ser Phe
660 665 670
Ile Thr Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu
675 680 685
Leu Gln Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr
690 695 700
Ser Asn Tyr Tyr Lys Ser Thr Asn Val Asp Phe Ala Val Asn Thr Glu
705 710 715 720
Gly Thr Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg
725 730 735
Asn Leu

<210> 94
<211> 738
<212> PRT
<213> AAV 血清型的衣壳蛋白, 克隆 43.5

<400> 94
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 Asp Leu Lys Pro Gly Ala Pro Lys Pro
20 25 30
Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro
35 40 45
Gly Tyr Lys Tyr Leu Gly Pro Phe 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 Arg Tyr Asn His Ala
85 90 95

Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly
100 105 110
Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
115 120 125
Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg
130 135 140
Pro Val Glu Pro Ser Pro Gln Arg Ser Pro Asp Ser Ser Thr Gly Ile
145 150 155 160
Gly Lys Lys Gly His Gln Pro Ala Arg Lys Arg Leu Asn Phe Gly Gln
165 170 175
Thr Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro
180 185 190
Pro Ala Gly Pro Ser Gly Leu Gly Ser Gly Thr Met Ala Ala Gly Gly
195 200 205
Gly Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser
210 215 220
Ser Ser Gly Asn Trp His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val
225 230 235 240
Ile Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His
245 250 255
Leu Tyr Lys Gln Ile Ser Asn Gly Thr Ser Gly Gly Ser Thr Asn Asp
260 265 270
Asn Thr Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn
275 280 285
Arg Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn
290 295 300
Asn Asn Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn
305 310 315 320
Ile Gln Val Lys Glu Val Thr Gln Asn Glu Gly Thr Lys Thr Ile Ala
325 330 335
[0113] Asn Asn Leu Thr Ser Thr Ile Gln Val Phe Thr Asp Ser Glu Tyr Gln
340 345 350
Leu Pro Tyr Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe
355 360 365
Pro Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn
370 375 380
Asn Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr
385 390 395 400
Phe Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Glu Phe Ser Tyr
405 410 415
Thr Phe Glu Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser
420 425 430
Leu Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu
435 440 445
Ser Arg Thr Gln Ser Thr Gly Gly Thr Gln Gly Thr Gln Gln Leu Leu
450 455 460
Phe Ser Gln Ala Gly Pro Ala Asn Met Ser Ala Gln Ala Lys Asn Trp
465 470 475 480
Leu Pro Gly Pro Cys Tyr Arg Gln Gln Arg Val Ser Thr Thr Leu Ser
485 490 495
Gln Asn Asn Asn Ser Asn Phe Ala Trp Thr Gly Ala Thr Lys Tyr His
500 505 510
Leu Asn Gly Arg Asp Ser Leu Val Asn Pro Gly Val Ala Met Ala Thr
515 520 525
His Lys Asp Asp Glu Glu Arg Phe Phe Pro Ser Ser Gly Val Leu Met
530 535 540
Phe Gly Lys Gln Gly Ala Gly Lys Asp Asn Val Asp Tyr Ser Ser Val
545 550 555 560
Met Leu Thr Ser Glu Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr
565 570 575
Glu Gln Tyr Gly Val Val Ala Asp Asn Leu Gln Gln Thr Asn Gly Ala
580 585 590

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

<210> 95
<211> 738
<212> PRT
<213> AAV 血清型的衣壳蛋白,克隆 AAV8

<400> 95
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 Lys Pro
20 25 30
[0114] Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro
35 40 45
Gly Tyr Lys Tyr Leu Gly Pro Phe 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 Glu Ala Gly Asp Asn Pro Tyr Leu Arg Tyr Asn His Ala
85 90 95
Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly
100 105 110
Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
115 120 125
Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg
130 135 140
Pro Val Glu Pro Ser Pro Gln Arg Ser Pro Asp Ser Ser Thr Gly Ile
145 150 155 160
Gly Lys Lys Gly Gln Gln Pro Ala Arg Lys Arg Leu Asn Phe Gly Gln
165 170 175
Thr Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Leu Gly Glu Pro
180 185 190
Pro Ala Ala Pro Ser Gly Val Gly Pro Asn Thr Met Ala Ala Gly Gly
195 200 205
Gly Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser
210 215 220
Ser Ser Gly Asn Trp His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val
225 230 235 240

Ile Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His
245 250 255
Leu Tyr Lys Gln Ile Ser Asn Gly Thr Ser Gly Gly Ala Thr Asn Asp
260 265 270

Asn Thr Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn
275 280 285
Arg Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn
290 295 300
Asn Asn Trp Gly Phe Arg Pro Lys Arg Leu Ser Phe Lys Leu Phe Asn
305 310 315 320
Ile Gln Val Lys Glu Val Thr Gln Asn Glu Gly Thr Lys Thr Ile Ala
325 330 335
Asn Asn Leu Thr Ser Thr Ile Gln Val Phe Thr Asp Ser Glu Tyr Gln
340 345 350
Leu Pro Tyr Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe
355 360 365
Pro Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn
370 375 380
Asn Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr
385 390 395 400
Phe Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Gln Phe Thr Tyr
405 410 415
Thr Phe Glu Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser
420 425 430
Leu Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu
435 440 445
Ser Arg Thr Gln Thr Thr Gly Gly Thr Ala Asn Thr Gln Thr Leu Gly
450 455 460
Phe Ser Gln Gly Gly Pro Asn Thr Met Ala Asn Gln Ala Lys Asn Trp
465 470 475 480
Leu Pro Gly Pro Cys Tyr Arg Gln Gln Arg Val Ser Thr Thr Thr Gly
485 490 495
Gln Asn Asn Asn Ser Asn Phe Ala Trp Thr Ala Gly Thr Lys Tyr His
[0115] 500 505 510
Leu Asn Gly Arg Asn Ser Leu Ala Asn Pro Gly Ile Ala Met Ala Thr
515 520 525
His Lys Asp Asp Glu Glu Arg Phe Phe Pro Ser Asn Gly Ile Leu Ile
530 535 540
Phe Gly Lys Gln Asn Ala Ala Arg Asp Asn Ala Asp Tyr Ser Asp Val
545 550 555 560
Met Leu Thr Ser Glu Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr
565 570 575
Glu Glu Tyr Gly Ile Val Ala Asp Asn Leu Gln Gln Gln Asn Thr Ala
580 585 590
Pro Gln Ile Gly Thr Val Asn Ser Gln Gly Ala Leu Pro Gly Met Val
595 600 605
Trp Gln Asn Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile
610 615 620
Pro His Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe
625 630 635 640
Gly Leu Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val
645 650 655
Pro Ala Asp Pro Pro Thr Thr Phe Asn Gln Ser Lys Leu Asn Ser Phe
660 665 670
Ile Thr Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu
675 680 685
Leu Gln Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr
690 695 700
Ser Asn Tyr Tyr Lys Ser Thr Ser Val Asp Phe Ala Val Asn Thr Glu
705 710 715 720
Gly Val Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg
725 730 735
Asn Leu

<210> 96
<211> 736
<212> PRT
<213> AAV 血清型的衣壳蛋白, 克隆 43.21

<400> 96

[0116]

```
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 Asp Leu Lys Pro Gly Ala Pro Lys Pro
            20                  25                  30
Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro
        35                  40                  45
Gly Tyr Lys Tyr Leu Gly Pro Phe 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 Arg Tyr Asn His Ala
                85                  90                  95
Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly
            100                 105                 110
Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
        115                 120                 125
Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg
    130                 135                 140
Pro Val Glu Gln Ser Pro Gln Glu Pro Asp Ser Ser Ser Gly Ile Gly
145                 150                 155                 160
Lys Thr Gly Gln Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr
                165                 170                 175
Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Leu Gly Glu Pro Pro
            180                 185                 190
Ala Ala Pro Ser Gly Leu Gly Pro Asn Thr Met Ala Ser Gly Gly Gly
        195                 200                 205
Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ser
    210                 215                 220

Ser Gly Asn Trp His Cys Asp Ser Thr 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 Gly Thr Ser Gly Gly Ser Thr Asn Asp Asn
            260                 265                 270
Thr 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 Thr Asn Glu Gly Thr Lys Thr Ile Ala Asn
                325                 330                 335
Asn Leu Thr Ser Thr Val Arg Val Phe Thr Asp Ser Glu Tyr Gln Leu
            340                 345                 350
Pro Tyr Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro
        355                 360                 365
Ala Asp Val Phe Met Val Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn
    370                 375                 380
Gly Ser Gln Ala Leu 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 Thr
                405                 410                 415
Phe Glu Asp 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 Val
```

435 440 445
Arg Thr Gln Thr Thr Gly Thr Gly Gly Thr Gln Thr Leu Ala Phe Ser
450 455 460
Gln Ala Gly Pro Ser Ser Met Ala Asn Gln Ala Arg Asn Trp Val Pro
465 470 475 480
Gly Pro Cys Tyr Arg Gln Gln Arg Val Ser Thr Thr Thr Asn Gln Ser
485 490 495
Asn Asn Ser Asn Phe Ala Trp Thr Gly Ala Ala Lys Phe Lys Leu Asn
500 505 510
Gly Arg Asp Ser Leu Met Asn Pro Gly Val Ala Met Ala Ser His Lys
515 520 525
Asp Asp Asp Asp Arg Phe Phe Pro Ser Ser Gly Val Leu Ile Phe Gly
530 535 540
Lys Gln Gly Ala Gly Asn Asp Gly Val Asp Tyr Ser Gln Val Leu Ile
545 550 555 560
Thr Asp Glu Glu Glu Ile Lys Ala Thr Asn Pro Val Ala Thr Glu Glu
565 570 575
Tyr Gly Ala Val Ala Ile Asn Asn Gln Ala Ala Asn Thr Gln Ala Gln
580 585 590
Thr Gly Leu Val His Asn Gln Gly Val Ile Pro Gly Met Val Trp Gln
595 600 605
Asn 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 Leu
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 Leu Thr Phe Asn Gln Ala 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

[0117]

Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn
690 695 700
Tyr Tyr Lys Ser Thr Asn Val Asp 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> 97
<211> 736
<212> PRT
<213> AAV 血清型的衣壳蛋白, 克隆 43.25

<400> 97
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 Asp Leu Lys Pro Gly Ala Pro Lys Pro
20 25 30
Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro
35 40 45
Gly Tyr Lys Tyr Leu Gly Pro Phe 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 Arg Tyr Asn His Ala
85 90 95
Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly
100 105 110
Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
115 120 125
Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg
130 135 140

[0118]

```
Pro Val Glu Gln Ser Pro Gln Glu Pro Asp Ser Ser Ser Gly Ile Gly
145                 150                 155                 160
Lys Thr Gly Gln Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr
                165                 170                 175
Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Leu Gly Glu Pro Pro
            180                 185                 190
Ala Ala Pro Ser Gly Leu Gly Pro Asn Thr Met Ala Ser Gly Gly Gly
        195                 200                 205
Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ser
    210                 215                 220
Ser Gly Asn Trp His Cys Asp Ser Thr 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 Gly Thr Ser Gly Gly Ser Thr Asn Asp Asn
            260                 265                 270
Thr 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 Thr Asn Glu Gly Thr Lys Thr Ile Ala Asn
                325                 330                 335
Asn Leu Thr Ser Thr Val Gln Val Phe Thr Asp Ser Glu Tyr Gln Leu
            340                 345                 350
Pro Tyr Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro
        355                 360                 365
Ala Asp Val Phe Met Val Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn
    370                 375                 380
Gly Ser Gln Ala Leu 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 Thr
                405                 410                 415
Phe Glu Asp 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 Val
        435                 440                 445
Arg Thr Gln Thr Thr Gly Thr Gly Gly Thr Gln Thr Leu Ala Phe Ser
    450                 455                 460
Gln Ala Gly Pro Ser Ser Met Ala Asn Gln Ala Arg Asn Trp Val Pro
465                 470                 475                 480
Gly Pro Cys Tyr Arg Gln Gln Arg Val Ser Thr Thr Thr Asn Gln Asn
                485                 490                 495
Asn Asn Ser Asn Phe Ala Trp Thr Gly Ala Ala Lys Phe Lys Leu Asn
            500                 505                 510
Gly Arg Asp Ser Leu Met Asn Pro Gly Val Ala Met Ala Ser His Lys
        515                 520                 525
Asp Asp Asp Asp Arg Phe Phe Pro Ser Ser Gly Val Leu Ile Phe Gly
    530                 535                 540
Lys Gln Gly Ala Gly Asn Asp Gly Val Asp Tyr Ser Gln Val Leu Ile
545                 550                 555                 560
Thr Asp Glu Glu Glu Ile Lys Ala Thr Asn Pro Val Ala Thr Glu Glu
                565                 570                 575
Tyr Gly Ala Val Ala Ile Asn Asn Gln Ala Ala Asn Thr Gln Ala Gln
            580                 585                 590
Thr Gly Leu Val His Asn Gln Gly Val Ile Pro Gly Met Val Trp Gln
        595                 600                 605
Asn 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 Leu
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 Leu Thr Phe Asn Gln Ala 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 Thr Asn Val Asp 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> 98
<211> 736
<212> PRT
<213> AAV 血清型的衣壳蛋白, 克隆 43.23

<400> 98
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 Asp Leu Lys Pro Gly Ala Pro Lys Pro
20 25 30
Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro
35 40 45
Gly Tyr Lys Tyr Leu Gly Pro Phe 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 Arg Tyr Asn His Ala
85 90 95
[0119] Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly
100 105 110
Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
115 120 125
Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg
130 135 140
Pro Val Glu Gln Ser Pro Gln Glu Pro Asp Ser Ser Ser Gly Ile Gly
145 150 155 160
Lys Thr Gly Gln Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr
165 170 175
Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Leu Gly Glu Pro Pro
180 185 190
Ala Ala Pro Ser Gly Leu Gly Pro Asn Thr Met Ala Ser Gly Gly Gly
195 200 205
Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ser
210 215 220
Ser Gly Asn Trp His Cys Asp Ser Thr 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 Gly Thr Ser Gly Gly Ser Thr Asn Asp Asn
260 265 270
Thr 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 Thr Asn Glu Gly Thr Lys Thr Ile Ala Asn
325 330 335
Asn Leu Thr Ser Thr Val Gln Val Phe Thr Asp Leu Glu Tyr Gln Leu

340 345 350
Pro Tyr Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro
355 360 365
Ala Asp Val Phe Met Val Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn
370 375 380
Gly Ser Gln Ala Leu Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe
385 390 395 400
Pro Ser Gln Met Pro Arg Thr Gly Asn Asn Phe Gln Phe Ser Tyr Thr
405 410 415
Phe Glu Asp 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 Val
435 440 445
Arg Thr Gln Thr Thr Gly Thr Gly Gly Thr Gln Thr Leu Ala Phe Ser
450 455 460
Gln Ala Gly Pro Ser Ser Met Ala Asn Gln Ala Arg Asn Trp Val Pro
465 470 475 480
Gly Pro Cys Tyr Arg Gln Gln Arg Val Ser Thr Thr Thr Asn Gln Asn
485 490 495
Asn Asn Ser Asn Phe Ala Trp Thr Gly Ala Ala Lys Phe Lys Leu Asn
500 505 510
Gly Arg Asp Ser Leu Met Asn Pro Gly Val Ala Met Ala Ser His Lys
515 520 525
Asp Asp Asp Asp Arg Phe Phe Pro Ser Ser Gly Val Leu Ile Phe Gly
530 535 540
Lys Gln Gly Ala Gly Asn Asp Gly Val Asp Tyr Ser Gln Val Leu Ile
545 550 555 560
Thr Asp Glu Glu Glu Ile Lys Ala Thr Asn Pro Val Ala Thr Glu Glu
565 570 575
Tyr Gly Ala Val Ala Ile Asn Asn Gln Ala Ala Asn Thr Gln Ala Gln
580 585 590
[0120] Thr Gly Leu Val His Asn Gln Gly Val Ile Pro Gly Met Val Trp Gln
595 600 605
Asn 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 Leu
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 Leu Thr Phe Asn Gln Ala 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 Thr Asn Val Asp 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> 99
<211> 736
<212> PRT
<213> AAV血清型的衣壳蛋白,克隆43.20

<400> 99
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 Asp Leu Lys Pro Gly Ala Pro Lys Pro
20 25 30
Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro

[0121]

```
          35                  40                  45
Gly Tyr Lys Tyr Leu Gly Pro Phe 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 Arg Tyr Asn His Ala
                85                  90                  95
Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly
            100                 105                 110
Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
        115                 120                 125
Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg
    130                 135                 140
Leu Val Glu Gln Ser Pro Gln Glu Pro Asp Ser Ser Ser Gly Ile Gly
145                 150                 155                 160
Lys Thr Gly Gln Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr
                165                 170                 175
Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Leu Gly Glu Pro Pro
            180                 185                 190
Ala Ala Pro Ser Gly Leu Gly Pro Asn Thr Met Ala Ser Gly Gly Gly
        195                 200                 205
Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ser
    210                 215                 220
Ser Gly Asn Trp His Cys Asp Ser Thr 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 Gly Thr Ser Gly Gly Ser Thr Asn Asp Asn
            260                 265                 270
Thr 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 Thr Asn Glu Gly Thr Lys Thr Ile Ala Asn
                325                 330                 335
Asn Leu Thr Ser Thr Val Gln Val Phe Thr Asp Ser Glu Tyr Gln Leu
            340                 345                 350
Pro Tyr Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro
        355                 360                 365
Ala Asp Val Phe Thr Val Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn
    370                 375                 380
Gly Ser Gln Ala Leu 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 Thr
                405                 410                 415
Phe Glu Asp 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 Val
        435                 440                 445
Arg Thr Gln Thr Thr Gly Thr Gly Gly Thr Gln Thr Leu Ala Phe Ser
    450                 455                 460
Gln Ala Gly Pro Ser Ser Met Ala Asn Gln Ala Arg Asn Trp Val Pro
465                 470                 475                 480
Gly Pro Cys Tyr Arg Gln Gln Arg Val Ser Thr Thr Thr Asn Gln Asn
                485                 490                 495
Asn Asn Ser Asn Phe Ala Trp Thr Gly Ala Ala Lys Phe Lys Leu Asn
            500                 505                 510
Gly Arg Asp Ser Leu Met Asn Pro Gly Val Ala Met Ala Ser His Lys
        515                 520                 525
Asp Asp Asp Asp Arg Phe Phe Pro Ser Ser Gly Val Leu Ile Phe Gly
```

```
    530                 535                 540
Lys Gln Gly Ala Gly Asn Asp Gly Val Asp Tyr Ser Gln Val Leu Ile
545                 550                 555                 560
Thr Asp Glu Glu Glu Ile Lys Ala Thr Asn Pro Val Ala Thr Glu Glu
                565                 570                 575
Tyr Gly Ala Val Ala Ile Asn Asn Gln Ala Ala Asn Thr Gln Ala Gln
            580                 585                 590
Thr Gly Leu Val His Asn Gln Gly Val Ile Pro Gly Met Val Trp Gln
        595                 600                 605
Asn 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 Leu
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 Leu Thr Phe Asn Gln Ala 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 Thr Asn Val Asp 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> 100
<211> 736
<212> PRT
<213> AAV 血清型的衣壳蛋白, 克隆 AAV9

[0122]

```
<400> 100
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 Asp Leu Lys Pro Gly Ala Pro Lys Pro
            20                  25                  30
Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro
        35                  40                  45
Gly Tyr Lys Tyr Leu Gly Pro Phe 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 Arg Tyr Asn His Ala
                85                  90                  95
Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly
            100                 105                 110
Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
        115                 120                 125
Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg
    130                 135                 140
Pro Val Glu Gln Ser Pro Gln Glu Pro Asp Ser Ser Ser Gly Ile Gly
145                 150                 155                 160
Lys Ser Gly Gln Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr
                165                 170                 175
Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Leu Gly Glu Pro Pro
            180                 185                 190
Glu Ala Pro Ser Gly Leu Gly Pro Asn Thr Met Ala Ser Gly Gly Gly
        195                 200                 205
Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ser
    210                 215                 220
Ser Gly Asn Trp His Cys Asp Ser Thr 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 Gly Thr Ser Gly Gly Ser Thr Asn Asp Asn
260 265 270
Thr 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 Thr Asn Glu Gly Thr Lys Thr Ile Ala Asn
325 330 335
Asn Leu Thr Ser Thr Val Gln Val Phe Thr Asp Ser Glu Tyr Gln Leu
340 345 350
Pro Tyr Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro
355 360 365
Ala Asp Val Phe Met Val Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn
370 375 380
Gly Ser Gln Ala Leu 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 Thr
405 410 415
Phe Glu Asp 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 Val
435 440 445
Arg Thr Gln Thr Thr Gly Thr Gly Gly Thr Gln Thr Leu Ala Phe Ser
450 455 460
Gln Ala Gly Pro Ser Ser Met Ala Asn Gln Ala Arg Asn Trp Val Pro
465 470 475 480
[0123] Gly Pro Cys Tyr Arg Gln Gln Arg Val Ser Thr Thr Thr Asn Gln Asn
485 490 495
Asn Asn Ser Asn Phe Ala Trp Thr Gly Ala Ala Lys Phe Lys Leu Asn
500 505 510
Gly Arg Asp Ser Leu Met Asn Pro Gly Val Ala Met Ala Ser His Lys
515 520 525
Asp Asp Glu Asp Arg Phe Phe Pro Ser Ser Gly Val Leu Ile Phe Gly
530 535 540
Lys Gln Gly Ala Gly Asn Asp Gly Val Asp Tyr Ser Gln Val Leu Ile
545 550 555 560
Thr Asp Glu Glu Glu Ile Lys Ala Thr Asn Pro Val Ala Thr Glu Glu
565 570 575
Tyr Gly Ala Val Ala Ile Asn Asn Gln Ala Ala Asn Thr Gln Ala Gln
580 585 590
Thr Gly Leu Val His Asn Gln Gly Val Ile Pro Gly Met Val Trp Gln
595 600 605
Asn 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 Leu
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 Leu Thr Phe Asn Gln Ala 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 Thr Asn Val Asp 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> 101
<211> 728
<212> PRT
<213> AAV 血清型的衣壳蛋白, 克隆 24.1

<400> 101
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 Asp Leu Lys Pro Gly Ala Pro Lys Pro
20 25 30
Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro
35 40 45
Gly Tyr Lys Tyr Leu Arg Pro Phe 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
Lys Gln Leu Glu Gln Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala
85 90 95
Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly
100 105 110
Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
115 120 125
Leu Gly Leu Val Glu Glu Val Ala Lys Thr Ala Pro Gly Lys Lys Arg
130 135 140
Pro Ile Glu Ser Pro Asp Ser Ser Thr Gly Ile Gly Lys Lys Gly Gln
145 150 155 160
Gln Pro Ala Lys Lys Lys Leu Asn Phe Gly Gln Thr Gly Asp Ser Glu
165 170 175
Ser Val Pro Asp Pro Gln Pro Leu Gly Glu Pro Pro Ala Ala Pro Ser
180 185 190
[0124] Gly Leu Gly Ser Gly Thr Met Ala Ala Gly Gly Gly Ala Pro Met Ala
195 200 205
Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ala Ser Gly Asn Trp
210 215 220
His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val Ile Thr Thr Ser Thr
225 230 235 240
Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu Tyr Lys Gln Ile
245 250 255
Ser Ser Gln Ser Gly Ala Thr Asn Asp Asn His Phe Phe Ser Tyr Ser
260 265 270
Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His Cys His Phe Ser
275 280 285
Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Trp Gly Phe Arg Pro
290 295 300
Arg Lys Leu Arg Phe Lys Leu Phe Asn Ile Gln Val Lys Glu Val Thr
305 310 315 320
Thr Asn Asp Gly Val Thr Thr Ile Ala Asn Asn Leu Thr Ser Thr Ile
325 330 335
Gln Val Phe Ser Asp Ser Glu Tyr Gln Leu Pro Tyr Val Leu Gly Ser
340 345 350
Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala Asp Val Phe Met Ile
355 360 365
Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly Ser Gln Ser Val Gly
370 375 380
Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser Gln Met Leu Arg
385 390 395 400
Thr Gly Asn Asn Phe Glu Phe Ser Tyr Thr Phe Glu Glu Val Pro Phe
405 410 415
His Ser Ser Tyr Val His Ser Gln Ser Leu Asp Arg Leu Met Asn Pro
420 425 430
Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ala Arg Thr Gln Ser Thr Thr

435 440 445
Gly Ser Thr Arg Glu Leu Gln Phe His Gln Ala Gly Pro Asn Thr Met
450 455 460
Ala Glu Gln Ser Lys Asn Trp Leu Pro Gly Pro Cys Tyr Arg Gln Gln
465 470 475 480
Arg Leu Ser Lys Asn Ile Asp Ser Asn Asn Asn Ser Asn Phe Ala Trp
485 490 495
Thr Gly Ala Thr Lys Tyr His Leu Asn Gly Arg Asn Ser Leu Thr Asn
500 505 510
Pro Gly Val Ala Met Ala Thr Asn Lys Asp Asp Glu Asp Gln Phe Phe
515 520 525
Pro Ile Asn Gly Val Leu Val Phe Gly Lys Thr Gly Ala Ala Asn Lys
530 535 540
Thr Thr Leu Glu Asn Val Leu Met Thr Ser Glu Glu Glu Ile Lys Thr
545 550 555 560
Thr Asn Pro Val Ala Thr Glu Glu Tyr Gly Val Val Ser Ser Asn Leu
565 570 575
Gln Ser Ser Thr Ala Gly Pro Gln Thr Gln Thr Val Asn Ser Gln Gly
580 585 590
Ala Leu Pro Gly Met Val Trp Gln Asn Arg Asp Val Cys Leu Gln Gly
595 600 605
Pro Ile Trp Ala Lys Ile Pro His Thr Asp Gly Asn Phe His Pro Ser
610 615 620
Pro Leu Met Gly Gly Phe Gly Leu Lys His Pro Pro Pro Gln Ile Leu
625 630 635 640
Ile Lys Asn Thr Pro Val Pro Ala Asn Pro Pro Glu Val Phe Thr Pro
645 650 655
Ala Lys Phe Ala Ser Phe Ile Thr Gln Tyr Ser Thr Gly Gln Val Ser
660 665 670
Val Glu Ile Glu Trp Glu Leu Gln Lys Glu Asn Ser Lys Arg Trp Asn
675 680 685
[0125] Pro Glu Ile Gln Tyr Thr Ser Asn Tyr Ala Lys Ser Asn Asn Val Glu
690 695 700
Phe Ala Val Asn Asn Glu Gly Val Tyr Thr Glu Pro Arg Pro Ile Gly
705 710 715 720
Thr Arg Tyr Leu Thr Arg Asn Leu
725

<210> 102
<211> 728
<212> PRT
<213> AAV 血清型的衣壳蛋白, 克隆 42.2REAL

<400> 102
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 Asp Leu Lys Pro Gly Ala Pro Lys Pro
20 25 30
Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro
35 40 45
Gly Tyr Lys Tyr Leu Gly Pro Phe 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
Lys Gln Leu Glu Gln Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala
85 90 95
Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly
100 105 110
Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
115 120 125
Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg

130 135 140
Pro Ile Glu Ser Pro Asp Ser Ser Thr Gly Ile Gly Lys Lys Gly Gln
145 150 155 160
Gln Pro Ala Lys Lys Lys Leu Asn Phe Gly Gln Thr Gly Asp Ser Glu
165 170 175
Ser Val Pro Asp Pro Gln Pro Leu Gly Glu Pro Pro Ala Ala Pro Ser
180 185 190
Gly Leu Gly Ser Gly Thr Met Ala Ala Gly Gly Gly Ala Pro Met Ala
195 200 205
Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ala Ser Gly Asn Trp
210 215 220
His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val Ile Thr Thr Ser Thr
225 230 235 240
Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu Tyr Lys Gln Ile
245 250 255
Ser Ser Gln Ser Gly Ala Thr Asn Asp Asn His Phe Phe Gly Tyr Ser
260 265 270
Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His Cys His Phe Ser
275 280 285
Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Trp Gly Phe Arg Pro
290 295 300
Arg Lys Leu Arg Phe Lys Leu Phe Asn Ile Gln Val Lys Glu Val Thr
305 310 315 320
Thr Asn Asp Gly Val Thr Thr Ile Ala Asn Asn Leu Thr Ser Thr Ile
325 330 335
Gln Val Phe Ser Asp Ser Glu Tyr Gln Leu Pro Tyr Val Leu Gly Ser
340 345 350
Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala Asp Val Phe Met Ile
355 360 365
Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly Ser Gln Ser Val Gly
370 375 380
[0126] Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser Gln Met Leu Arg
385 390 395 400
Thr Gly Asn Asn Phe Glu Phe Ser Tyr Thr Phe Glu Glu Val Pro Phe
405 410 415
His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp Arg Leu Met Asn Pro
420 425 430
Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ala Arg Thr Gln Ser Thr Thr
435 440 445
Gly Ser Thr Arg Glu Leu Gln Phe His Gln Ala Gly Pro Asn Thr Met
450 455 460
Ala Glu Gln Ser Lys Asn Trp Leu Pro Gly Pro Cys Tyr Arg Gln Gln
465 470 475 480
Arg Leu Ser Lys Asn Ile Asp Ser Asn Asn Asn Ser Asn Phe Ala Trp
485 490 495
Thr Gly Ala Thr Lys Tyr His Leu Asn Gly Arg Asn Ser Leu Thr Asn
500 505 510
Pro Gly Val Ala Met Ala Thr Asn Lys Asp Asp Glu Asp Gln Phe Phe
515 520 525
Pro Ile Asn Gly Val Leu Val Phe Gly Glu Thr Gly Ala Ala Asn Lys
530 535 540
Thr Thr Leu Glu Asn Val Leu Met Thr Ser Glu Glu Glu Ile Lys Thr
545 550 555 560
Thr Asn Pro Val Ala Thr Glu Glu Tyr Gly Val Val Ser Ser Asn Leu
565 570 575
Gln Ser Ser Thr Ala Gly Pro Gln Thr Gln Thr Val Asn Ser Gln Gly
580 585 590
Ala Leu Pro Gly Met Val Trp Gln Asn Arg Asp Val Tyr Leu Gln Gly
595 600 605
Pro Ile Trp Ala Lys Ile Pro His Thr Asp Gly Asn Phe His Pro Ser
610 615 620
Pro Leu Met Gly Gly Phe Gly Leu Lys His Pro Pro Pro Gln Ile Leu

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

<210> 103
<211> 728
<212> PRT
<213> AAV 血清型的衣壳蛋白,克隆 7.2VP1

<400> 103
Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Gly Asn Leu Ser
1 5 10 15
Glu Gly Ile Arg Glu Trp Trp Asp Leu Lys Pro Gly Ala Pro Lys Pro
20 25 30
Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro
35 40 45
Gly Tyr Arg Tyr Leu Gly Pro Phe 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
Lys Gln Leu Glu Gln Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala
85 90 95
Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly
100 105 110
Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
115 120 125
Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg
130 135 140
Pro Ile Glu Ser Pro Asp Ser Ser Thr Gly Ile Gly Lys Asn Gly Gln
145 150 155 160
Pro Pro Ala Lys Lys Lys Leu Asn Phe Gly Gln Thr Gly Asp Ser Glu
165 170 175
Ser Val Pro Asp Pro Gln Pro Leu Gly Glu Pro Pro Ala Ala Pro Ser
180 185 190
Gly Leu Gly Ser Gly Thr Met Ala Ala Gly Gly Gly Ala Pro Met Ala
195 200 205
Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ala Ser Gly Asn Trp
210 215 220
His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val Ile Thr Thr Ser Thr
225 230 235 240
Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu Tyr Lys Gln Ile
245 250 255
Ser Ser Gln Ser Gly Ala Thr Asn Asp Asn His Phe Phe Gly Tyr Ser
260 265 270
Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His Cys His Phe Ser
275 280 285
Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Trp Gly Phe Arg Pro
290 295 300
Arg Lys Leu Arg Phe Lys Leu Phe Asn Ile Gln Val Lys Glu Val Thr
305 310 315 320
Thr Asn Asp Gly Val Thr Thr Ile Ala Asn Asn Leu Thr Ser Thr Ile
325 330 335

[0127]

Gln Val Phe Ser Asp Ser Glu Tyr Gln Leu Pro Tyr Val Leu Gly Ser
340 345 350
Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala Asp Val Phe Met Ile
355 360 365
Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly Ser Gln Ser Val Gly
370 375 380
Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser Gln Met Leu Arg
385 390 395 400
Thr Gly Asp Asn Phe Glu Phe Ser Tyr Thr Phe Glu Glu Val Pro Phe
405 410 415
His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp Arg Leu Met Asn Pro
420 425 430
Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ala Arg Thr Gln Ser Thr Thr
435 440 445
Gly Ser Thr Arg Glu Leu Gln Phe His Gln Ala Gly Pro Asn Thr Met
450 455 460
Ala Glu Gln Ser Lys Asn Trp Leu Pro Gly Pro Cys Tyr Arg Gln Gln
465 470 475 480
Arg Leu Ser Lys Asn Ile Asp Ser Asn Asn Asn Ser Asn Phe Ala Trp
485 490 495
Thr Gly Ala Thr Lys Tyr His Leu Asn Gly Arg Asn Ser Leu Thr Asn
500 505 510
Pro Gly Val Ala Met Ala Thr Asn Lys Asp Asp Glu Asp Gln Phe Phe
515 520 525
Pro Ile Asn Gly Val Leu Val Phe Gly Lys Thr Gly Ala Ala Asn Lys
530 535 540
Thr Thr Leu Glu Asn Val Leu Met Thr Ser Glu Glu Glu Ile Lys Thr
545 550 555 560
Thr Asn Pro Val Ala Thr Glu Glu Tyr Gly Val Val Ser Ser Asn Leu
565 570 575
Gln Ser Ser Thr Ala Gly Pro Gln Thr Gln Thr Val Asn Ser Gln Gly
580 585 590
Ala Leu Pro Gly Met Val Trp Gln Asn Arg Asp Val Tyr Leu Gln Gly
595 600 605
Pro Ile Trp Ala Lys Ile Pro His Thr Asp Gly Asn Phe His Pro Ser
610 615 620
Pro Leu Met Gly Gly Phe Gly Leu Lys His Pro Pro Pro Gln Ile Leu
625 630 635 640
Ile Lys Asn Thr Pro Val Pro Ala Asn Pro Pro Glu Val Phe Thr Pro
645 650 655
Ala Lys Phe Ala Ser Phe Ile Thr Gln Tyr Ser Thr Gly Gln Val Ser
660 665 670
Val Glu Ile Glu Trp Glu Leu Gln Lys Glu Asn Ser Lys Arg Trp Asn
675 680 685
Pro Glu Ile Gln Tyr Thr Ser Asn Tyr Ala Lys Ser Asn Asn Val Glu
690 695 700
Phe Ala Val Asn Asn Glu Gly Val Tyr Thr Glu Pro Arg Pro Ile Gly
705 710 715 720
Thr Arg Tyr Leu Thr Arg Asn Leu
725

[0128]

<210> 104
<211> 728
<212> PRT
<213> AAV 血清型的衣壳蛋白, 克隆 27.3VP1

<400> 104
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 Asp Leu Lys Pro Gly Ala Pro Lys Pro
20 25 30
Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro

[0129]

```
          35                  40                  45
Gly Tyr Lys Tyr Leu Gly Pro Phe 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
Lys Gln Leu Glu Gln Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala
                85                  90                  95
Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly
            100                 105                 110
Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
        115                 120                 125
Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Ser Gly Lys Lys Arg
    130                 135                 140
Pro Ile Glu Ser Pro Asp Ser Ser Thr Gly Ile Gly Lys Lys Gly Gln
145                 150                 155                 160
Gln Pro Ala Lys Lys Lys Leu Asn Phe Gly Gln Thr Gly Asp Ser Glu
                165                 170                 175
Ser Val Pro Asp Pro Gln Pro Leu Gly Glu Pro Pro Ala Ala Pro Ser
            180                 185                 190
Gly Leu Gly Ser Gly Thr Met Ala Ala Gly Gly Gly Ala Pro Met Ala
        195                 200                 205
Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ala Ser Gly Asn Trp
    210                 215                 220
His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val Ile Thr Thr Ser Thr
225                 230                 235                 240
Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu Tyr Lys Gln Ile
                245                 250                 255
Ser Ser Gln Ser Gly Ala Thr Asn Asp Asn His Phe Phe Gly Tyr Ser
            260                 265                 270
Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His Cys His Phe Ser
        275                 280                 285
Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Trp Gly Phe Arg Pro
    290                 295                 300
Arg Lys Leu Arg Phe Lys Leu Phe Asn Ile Gln Val Lys Glu Val Thr
305                 310                 315                 320
Thr Asn Asp Gly Val Thr Thr Ile Ala Asn Asn Leu Thr Ser Thr Ile
                325                 330                 335
Gln Val Phe Ser Asp Ser Glu Tyr Gln Leu Pro Tyr Val Leu Gly Ser
            340                 345                 350
Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala Asp Val Phe Met Ile
        355                 360                 365
Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly Ser Gln Ser Val Gly
    370                 375                 380
Arg Ser Ser Phe Cys Cys Leu Glu Tyr Phe Pro Ser Gln Met Leu Arg
385                 390                 395                 400
Thr Gly Asn Asn Phe Glu Phe Ser Tyr Thr Phe Glu Glu Val Pro Phe
                405                 410                 415
His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp Arg Leu Met Asn Pro
            420                 425                 430
Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ala Arg Thr Gln Ser Thr Thr
        435                 440                 445
Gly Ser Thr Arg Glu Leu Gln Phe His Gln Ala Gly Pro Asn Thr Val
    450                 455                 460
Ala Glu Gln Ser Lys Asn Trp Leu Pro Gly Pro Cys Tyr Arg Gln Gln
465                 470                 475                 480
Arg Leu Ser Lys Asn Ile Asp Ser Asn Asn Asn Ser Asn Phe Ala Trp
                485                 490                 495
Thr Gly Ala Thr Lys Tyr His Leu Asn Gly Arg Asn Ser Leu Thr Asn
            500                 505                 510
Pro Gly Val Ala Met Ala Thr Asn Lys Asp Asp Glu Asp Gln Phe Leu
        515                 520                 525
Pro Ile Asn Gly Val Leu Val Phe Gly Lys Thr Gly Ala Ala Asn Lys
```

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

<210> 105
<211> 728
<212> PRT
<213> AAV 血清型的衣壳蛋白, 克隆 16.3VP1

[0130]

```
<400> 105
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 Asp Leu Lys Pro Gly Ala Pro Lys Pro
            20                  25                  30
Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro
        35                  40                  45
Gly Tyr Lys Tyr Leu Gly Pro Phe 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
Lys Gln Leu Glu Gln Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala
                85                  90                  95
Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly
            100                 105                 110
Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
        115                 120                 125
Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg
    130                 135                 140
Pro Ile Glu Ser Pro Asp Ser Ser Thr Gly Ile Gly Lys Lys Gly Gln
145                 150                 155                 160
Gln Pro Ala Lys Lys Lys Leu Asn Phe Gly Gln Thr Gly Asp Ser Glu
                165                 170                 175
Ser Val Pro Asp Pro Gln Pro Leu Gly Glu Pro Pro Ala Ala Pro Ser
            180                 185                 190
Gly Leu Gly Ser Gly Thr Met Ala Ala Gly Gly Gly Ala Pro Met Ala
        195                 200                 205
Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ala Ser Gly Asn Trp
    210                 215                 220
His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val Ile Thr Thr Ser Thr
225                 230                 235                 240
```

Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu Tyr Lys Gln Ile
245 250 255
Ser Ser Gln Ser Gly Ala Thr Asn Asp Asn His Phe Phe Gly Tyr Ser
260 265 270
Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His Cys His Phe Ser
275 280 285
Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Trp Gly Phe Arg Pro
290 295 300
Arg Lys Leu Arg Phe Lys Leu Phe Asn Ile Gln Val Lys Glu Val Thr
305 310 315 320
Thr Asn Asp Gly Val Thr Thr Ile Ala Asn Asn Leu Thr Ser Thr Ile
325 330 335
Gln Val Phe Ser Asp Ser Glu Tyr Gln Leu Pro Tyr Val Leu Gly Ser
340 345 350
Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala Asp Val Phe Met Ile
355 360 365
Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly Ser Gln Ser Met Gly
370 375 380
Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser Gln Met Leu Arg
385 390 395 400
Thr Gly Asn Asn Phe Glu Phe Ser Tyr Thr Phe Glu Glu Val Pro Phe
405 410 415
His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp Arg Leu Met Asn Pro
420 425 430
Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ala Arg Thr Gln Ser Thr Thr
435 440 445
Gly Ser Thr Arg Glu Leu Gln Phe His Gln Ala Gly Pro Asn Thr Met
450 455 460
Ala Glu Gln Ser Lys Asn Trp Leu Pro Gly Pro Cys Tyr Arg Gln Gln
465 470 475 480
[0131] Arg Leu Ser Lys Asn Ile Asp Ser Asn Asn Asn Ser Asn Phe Ala Trp
485 490 495
Thr Gly Ala Thr Lys Tyr His Leu Asn Gly Arg Asn Ser Leu Thr Asn
500 505 510
Pro Gly Val Ala Met Ala Thr Asn Lys Asp Asp Glu Gly Gln Phe Phe
515 520 525
Pro Ile Asn Gly Val Leu Val Phe Gly Lys Thr Gly Ala Ala Asn Lys
530 535 540
Thr Thr Leu Glu Asn Val Leu Met Thr Ser Glu Glu Glu Ile Lys Thr
545 550 555 560
Thr Asn Pro Val Ala Thr Glu Glu Tyr Gly Val Val Ser Ser Asn Leu
565 570 575
Gln Ser Ser Thr Ala Gly Pro Gln Thr Gln Thr Val Asn Ser Gln Gly
580 585 590
Ala Leu Pro Gly Met Val Trp Gln Asn Arg Asp Val Tyr Leu Gln Gly
595 600 605
Pro Ile Trp Ala Lys Ile Pro His Thr Asp Gly Asn Phe His Pro Ser
610 615 620
Pro Leu Met Gly Gly Phe Gly Leu Lys His Pro Pro Pro Gln Ile Leu
625 630 635 640
Ile Lys Asn Thr Pro Val Pro Ala Asn Pro Pro Glu Val Phe Thr Pro
645 650 655
Ala Leu Phe Ala Ser Phe Ile Thr Gln Tyr Ser Thr Gly Gln Val Ser
660 665 670
Val Glu Ile Glu Trp Glu Leu Gln Lys Glu Asn Ser Lys Arg Trp Asn
675 680 685
Pro Glu Ile Gln Tyr Thr Ser Asn Tyr Ala Lys Ser Asn Asn Val Glu
690 695 700
Phe Ala Val Asn Asn Glu Gly Val Tyr Thr Glu Pro Arg Pro Ile Gly
705 710 715 720
Thr Arg Tyr Leu Thr Arg Asn Leu
725

<210> 106
<211> 728
<212> PRT
<213> AAV 血清型的衣壳蛋白, 克隆 42.10

<400> 106
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 Asp Leu Lys Pro Gly Ala Pro Lys Pro
20 25 30
Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro
35 40 45
Gly Tyr Lys Tyr Leu Gly Pro Phe 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
Lys Gln Leu Glu Gln Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala
85 90 95
Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly
100 105 110
Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
115 120 125
Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg
130 135 140
Pro Ile Glu Ser Pro Asp Ser Ser Thr Gly Ile Gly Arg Lys Gly Gln
145 150 155 160
Gln Pro Ala Lys Lys Lys Leu Asn Phe Gly Gln Thr Gly Asp Ser Glu
165 170 175
Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro Ala Gly Pro Ser
180 185 190
[0132] Gly Leu Gly Ser Gly Thr Met Ala Ala Gly Gly Gly Ala Pro Met Ala
195 200 205
Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ala Ser Gly Asn Trp
210 215 220
His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val Ile Thr Thr Ser Thr
225 230 235 240
Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu Tyr Lys Gln Ile
245 250 255
Ser Ser Gln Ser Gly Ala Thr Asn Asp Asn His Phe Phe Gly Tyr Ser
260 265 270
Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His Cys His Phe Ser
275 280 285
Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Trp Gly Phe Arg Pro
290 295 300
Arg Lys Leu Arg Phe Lys Leu Phe Asn Ile Gln Val Lys Glu Val Thr
305 310 315 320
Thr Asn Asp Gly Val Thr Thr Ile Ala Asn Asn Leu Thr Ser Thr Ile
325 330 335
Gln Val Phe Ser Asp Ser Glu Tyr Gln Leu Pro Tyr Val Leu Gly Ser
340 345 350
Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala Asp Val Phe Met Ile
355 360 365
Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly Ser Gln Ser Val Gly
370 375 380
Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser Gln Met Leu Arg
385 390 395 400
Thr Gly Asn Asn Phe Glu Phe Ser Tyr Thr Phe Glu Glu Val Pro Phe
405 410 415
His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp Arg Leu Met Asn Pro
420 425 430
Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ala Arg Thr Gln Ser Thr Thr

435 440 445
Gly Ser Thr Arg Glu Leu Gln Phe His Gln Ala Gly Pro Asn Thr Met
450 455 460
Ala Glu Gln Ser Lys Asn Trp Leu Pro Gly Pro Cys Tyr Arg Gln Gln
465 470 475 480
Arg Leu Ser Lys Asn Ile Asp Ser Asn Asn Asn Ser Asn Phe Ala Trp
485 490 495
Thr Gly Ala Thr Lys Tyr His Leu Asn Gly Arg Asn Ser Leu Thr Asn
500 505 510
Pro Gly Val Ala Met Ala Thr Asn Lys Asp Asp Glu Asp Gln Phe Phe
515 520 525
Pro Ile Asn Gly Val Leu Val Phe Gly Lys Thr Gly Ala Ala Asn Lys
530 535 540
Thr Thr Leu Glu Asn Val Leu Met Thr Ser Glu Glu Glu Ile Lys Thr
545 550 555 560
Thr Asn Pro Val Ala Thr Glu Glu Tyr Gly Val Val Ser Ser Asn Leu
565 570 575
Gln Ser Ser Thr Ala Gly Pro Gln Thr Gln Thr Val Asn Ser Gln Gly
580 585 590
Ala Leu Pro Gly Met Val Trp Gln Asn Arg Asp Val Tyr Leu Gln Gly
595 600 605
Pro Ile Trp Ala Lys Ile Pro His Thr Asp Gly Asn Phe His Pro Ser
610 615 620
Pro Leu Met Gly Gly Phe Gly Leu Lys His Pro Pro Pro Gln Ile Leu
625 630 635 640
Ile Lys Asn Thr Pro Val Pro Ala Asn Pro Pro Glu Val Phe Thr Pro
645 650 655
Ala Lys Phe Ala Ser Phe Ile Thr Gln Tyr Ser Thr Gly Gln Val Ser
660 665 670
Val Glu Ile Glu Trp Glu Leu Gln Lys Glu Asn Ser Lys Arg Trp Asn
675 680 685
[0133] Pro Glu Ile Gln Tyr Thr Ser Asn Tyr Ala Lys Ser Asn Asn Val Glu
690 695 700
Phe Ala Val Asn Asn Glu Gly Val Tyr Thr Glu Pro Arg Pro Ile Gly
705 710 715 720
Thr Arg Tyr Leu Thr Arg Asn Leu
725

<210> 107
<211> 728
<212> PRT
<213> AAV 血清型的衣壳蛋白, 克隆 42.3B

<400> 107
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 Asp Leu Lys Pro Gly Ala Pro Lys Pro
20 25 30
Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro
35 40 45
Gly Tyr Lys Tyr Leu Gly Pro Phe 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
Lys Gln Leu Glu Gln Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala
85 90 95
Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly
100 105 110
Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
115 120 125
Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg
130 135 140

[0134]

```
Pro Ile Glu Ser Pro Asp Ser Ser Thr Gly Ile Gly Lys Lys Gly Gln
145                 150                 155                 160
Gln Pro Ala Lys Lys Lys Leu Asn Phe Gly Gln Thr Gly Asp Ser Glu
                165                 170                 175
Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro Ala Gly Pro Ser
            180                 185                 190
Gly Leu Gly Ser Gly Thr Met Ala Ala Gly Gly Gly Ala Pro Met Ala
        195                 200                 205
Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ala Ser Gly Asn Trp
    210                 215                 220
His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val Ile Thr Thr Ser Thr
225                 230                 235                 240
Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu Tyr Lys Gln Ile
                245                 250                 255
Ser Ser Gln Ser Gly Ala Thr Asn Asp Asn His Phe Phe Gly Tyr Ser
            260                 265                 270
Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His Cys His Phe Ser
        275                 280                 285
Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Trp Gly Phe Arg Pro
    290                 295                 300
Arg Lys Leu Arg Phe Lys Leu Phe Asn Ile Gln Val Lys Glu Val Thr
305                 310                 315                 320
Thr Asn Asp Gly Val Thr Thr Ile Ala Asn Asn Leu Thr Ser Thr Ile
                325                 330                 335
Gln Val Phe Ser Asp Ser Glu Tyr Gln Leu Pro Tyr Val Leu Gly Ser
            340                 345                 350
Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala Asp Val Phe Met Ile
        355                 360                 365
Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly Ser Gln Ser Val Gly
    370                 375                 380
Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser Gln Met Leu Arg
385                 390                 395                 400
Thr Gly Asn Asn Phe Glu Phe Ser Tyr Thr Phe Glu Glu Val Pro Phe
                405                 410                 415
His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp Arg Leu Met Asn Pro
            420                 425                 430
Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ala Arg Thr Gln Ser Thr Thr
        435                 440                 445
Gly Ser Thr Arg Glu Leu Gln Phe His Gln Ala Gly Pro Asn Thr Met
    450                 455                 460
Ala Glu Gln Ser Lys Asn Trp Leu Pro Gly Pro Cys Tyr Arg Gln Gln
465                 470                 475                 480
Arg Leu Ser Lys Asn Ile Asp Ser Asn Asn Thr Ser Asn Phe Ala Trp
                485                 490                 495
Thr Gly Ala Thr Lys Tyr His Leu Asn Gly Arg Asn Ser Leu Thr Asn
            500                 505                 510
Pro Gly Val Ala Met Ala Thr Asn Lys Asp Asp Glu Asp Gln Phe Phe
        515                 520                 525
Pro Ile Asn Gly Val Leu Val Phe Gly Lys Thr Gly Ala Ala Asn Lys
    530                 535                 540
Thr Thr Leu Glu Asn Val Leu Met Thr Ser Glu Glu Glu Ile Lys Thr
545                 550                 555                 560
Thr Asn Pro Val Ala Thr Glu Gln Tyr Gly Val Val Ser Ser Asn Leu
                565                 570                 575
Gln Ser Ser Thr Ala Gly Pro Gln Thr Gln Thr Val Asn Ser Gln Gly
            580                 585                 590
Ala Leu Pro Gly Met Val Trp Gln Asn Arg Asp Val Tyr Leu Gln Gly
        595                 600                 605
Pro Ile Trp Ala Lys Ile Pro His Thr Asp Gly Asn Phe His Pro Ser
    610                 615                 620
Pro Leu Met Gly Gly Phe Gly Leu Lys His Pro Pro Pro Gln Ile Leu
625                 630                 635                 640
```

Ile Lys Asn Thr Pro Val Pro Ala Asn Pro Pro Glu Val Phe Thr Pro
645 650 655
Ala Lys Phe Ala Ser Phe Ile Thr Gln Tyr Ser Thr Gly Gln Val Ser
660 665 670
Val Glu Ile Glu Trp Glu Leu Gln Lys Glu Asn Ser Lys Arg Trp Asn
675 680 685
Pro Glu Ile Gln Tyr Thr Ser Asn Tyr Ala Lys Ser Asn Asn Val Glu
690 695 700
Phe Ala Val Asn Asn Glu Gly Val Tyr Thr Glu Pro Arg Pro Ile Gly
705 710 715 720
Thr Arg Tyr Leu Thr Arg Asn Leu
725

<210> 108
<211> 728
<212> PRT
<213> AAV 血清型的衣壳蛋白, 克隆 42.11

<400> 108
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 Asp Leu Lys Pro Gly Ala Pro Lys Pro
20 25 30
Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro
35 40 45
Gly Tyr Lys Tyr Leu Gly Pro Phe 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 Arg Tyr Asn His Ala
85 90 95
[0135] Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly
100 105 110
Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
115 120 125
Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg
130 135 140
Pro Ile Glu Ser Pro Asp Ser Ser Thr Gly Ile Gly Lys Lys Gly Gln
145 150 155 160
Gln Pro Ala Lys Lys Lys Leu Asn Phe Gly Gln Thr Gly Asp Ser Glu
165 170 175
Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro Ala Gly Pro Ser
180 185 190
Gly Leu Gly Ser Gly Thr Met Ala Ala Gly Gly Gly Ala Pro Met Ala
195 200 205
Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ala Ser Gly Asn Trp
210 215 220
His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val Ile Thr Thr Ser Thr
225 230 235 240
Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu Tyr Lys Gln Ile
245 250 255
Ser Ser Gln Ser Gly Ala Thr Asn Asp Asn His Phe Phe Gly Tyr Ser
260 265 270
Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His Cys His Phe Ser
275 280 285
Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Trp Gly Phe Arg Pro
290 295 300
Arg Lys Leu Arg Phe Lys Leu Phe Asn Ile Gln Val Lys Glu Val Thr
305 310 315 320
Thr Asn Asp Gly Val Thr Thr Ile Ala Asn Asn Leu Thr Ser Thr Ile
325 330 335
Gln Val Phe Ser Asp Ser Glu Tyr Gln Leu Pro Tyr Val Leu Gly Ser

[0136]

```
            340                 345                 350
Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala Asp Val Phe Met Ile
        355                 360                 365
Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly Ser Gln Ser Val Gly
    370                 375                 380
Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser Gln Met Leu Arg
385                 390                 395                 400
Thr Gly Asn Asn Phe Glu Phe Ser Tyr Thr Phe Glu Glu Val Pro Phe
                405                 410                 415
His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp Arg Leu Met Asn Pro
            420                 425                 430
Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ala Arg Thr Gln Ser Thr Thr
        435                 440                 445
Gly Ser Thr Arg Glu Leu Gln Phe His Gln Ala Gly Pro Asn Thr Met
    450                 455                 460
Ala Glu Gln Ser Lys Asn Trp Leu Pro Gly Pro Cys Tyr Arg Arg Gln
465                 470                 475                 480
Arg Leu Ser Lys Asp Ile Asp Ser Asn Asn Asn Ser Asn Phe Ala Trp
                485                 490                 495
Thr Gly Ala Thr Lys Tyr His Leu Asn Gly Arg Asn Ser Leu Thr Asn
            500                 505                 510
Pro Gly Val Ala Met Ala Thr Asn Lys Asp Asp Glu Asp Gln Phe Phe
        515                 520                 525
Pro Ile Asn Gly Val Leu Val Phe Gly Lys Thr Gly Ala Ala Asn Lys
    530                 535                 540
Thr Thr Leu Glu Asn Val Leu Met Thr Ser Glu Glu Glu Ile Lys Thr
545                 550                 555                 560
Thr Asn Pro Val Ala Thr Glu Glu Tyr Gly Val Val Ser Ser Asn Leu
                565                 570                 575
Gln Ser Ser Thr Ala Gly Pro Gln Thr Gln Thr Val Asn Ser Gln Gly
            580                 585                 590
Ala Leu Pro Gly Met Val Trp Gln Asn Arg Asp Val Tyr Leu Gln Gly
        595                 600                 605
Pro Ile Trp Ala Lys Ile Pro His Thr Asp Gly Asn Phe His Pro Ser
    610                 615                 620
Pro Leu Met Gly Gly Phe Gly Leu Lys His Pro Pro Pro Gln Ile Leu
625                 630                 635                 640
Ile Lys Asn Thr Pro Val Pro Ala Asn Pro Pro Glu Val Phe Thr Pro
                645                 650                 655
Ala Lys Phe Ala Ser Phe Ile Thr Gln Tyr Ser Thr Gly Gln Val Ser
            660                 665                 670
Val Glu Ile Glu Trp Glu Leu Gln Lys Glu Asn Ser Lys Arg Trp Asn
        675                 680                 685
Pro Glu Ile Gln Tyr Thr Ser Asn Tyr Ala Lys Ser Asn Asn Val Glu
    690                 695                 700
Phe Ala Val Asn Asn Glu Gly Val Tyr Thr Glu Pro Arg Pro Ile Gly
705                 710                 715                 720
Thr Arg Tyr Leu Thr Arg Asn Leu
                725
```

<210> 109
<211> 729
<212> PRT
<213> AAV 血清型的衣壳蛋白, 克隆 F1VP1

```
<400> 109
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 Asp Leu Lys Pro Gly Ala Pro Lys Pro
            20                  25                  30
Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro
        35                  40                  45
```

[0137]

```
Gly Tyr Lys Tyr Leu Gly Pro Phe 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 Arg Tyr Asn His Ala
                85                  90                  95
Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly
            100                 105                 110
Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
        115                 120                 125
Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg
    130                 135                 140
Pro Ile Asp Ser Pro Asp Ser Ser Thr Gly Ile Gly Lys Lys Gly Gln
145                 150                 155                 160
Gln Pro Ala Lys Lys Lys Leu Asn Phe Gly Gln Thr Gly Asp Ser Glu
                165                 170                 175
Ser Val Pro Asp Pro Gln Pro Leu Gly Glu Pro Pro Ala Ala Pro Ser
            180                 185                 190
Ser Val Gly Ser Gly Thr Met Ala Ala Gly Gly Gly Ala Pro Met Ala
        195                 200                 205
Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ala Ser Gly Asn Trp
    210                 215                 220
His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val Ile Thr Thr Ser Thr
225                 230                 235                 240
Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu Tyr Lys Gln Ile
                245                 250                 255
Ser Ser Ser Ser Ser Gly Ala Thr Asn Asp Asn His Tyr Phe Gly Tyr
            260                 265                 270
Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His Cys His Phe
        275                 280                 285
Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Trp Gly Phe Arg
    290                 295                 300
Pro Lys Lys Leu Arg Phe Lys Leu Phe Asn Ile Gln Val Lys Glu Val
305                 310                 315                 320
Thr Thr Asn Asp Gly Val Thr Thr Ile Ala Asn Asn Leu Thr Ser Thr
                325                 330                 335
Val Gln Val Phe Ser Asp Ser Glu Tyr Gln Leu Pro Tyr Val Leu Gly
            340                 345                 350
Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala Asp Val Phe Met
        355                 360                 365
Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly Ser Gln Ser Val
    370                 375                 380
Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser Gln Met Leu
385                 390                 395                 400
Arg Thr Gly Asn Asn Phe Glu Phe Ser Tyr Ser Phe Glu Asp Val Pro
                405                 410                 415
Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp Arg Leu Met Asn
            420                 425                 430
Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ala Arg Thr Gln Ser Thr
        435                 440                 445
Thr Gly Ser Thr Arg Glu Leu Gln Phe His Gln Ala Gly Pro Asn Thr
    450                 455                 460
Met Ala Glu Gln Ser Lys Asn Trp Leu Pro Gly Pro Cys Tyr Arg Gln
465                 470                 475                 480
Gln Gly Leu Ser Lys Asn Leu Asp Phe Asn Asn Asn Ser Asn Phe Ala
                485                 490                 495
Trp Thr Ala Ala Thr Lys Tyr His Leu Asn Gly Arg Asn Ser Leu Thr
            500                 505                 510
Asn Pro Gly Ile Pro Met Ala Thr Asn Lys Asp Asp Glu Asp Gln Phe
        515                 520                 525
Phe Pro Ile Asn Gly Val Leu Val Phe Gly Lys Thr Gly Ala Ala Asn
    530                 535                 540
```

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

<210> 110
<211> 729
<212> PRT
<213> AAV 血清型的衣壳蛋白, 克隆 F5VP1@3

[0138]
<400> 110
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 Asp Leu Lys Pro Gly Ala Pro Lys Pro
20 25 30
Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro
35 40 45
Gly Tyr Lys Tyr Leu Gly Pro Phe 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 Arg Tyr Asn His Ala
85 90 95
Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly
100 105 110
Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
115 120 125
Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg
130 135 140
Pro Ile Asp Ser Pro Asp Ser Ser Thr Gly Ile Gly Lys Lys Gly Gln
145 150 155 160
Gln Pro Ala Lys Lys Lys Leu Asn Phe Gly Gln Thr Gly Asp Ser Glu
165 170 175
Ser Val Pro Asp Pro Gln Pro Leu Gly Glu Pro Pro Ala Ala Pro Ser
180 185 190
Ser Val Gly Ser Gly Thr Met Ala Ala Gly Gly Gly Ala Pro Thr Ala
195 200 205
Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ala Ser Gly Asn Trp
210 215 220
His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val Ile Thr Thr Ser Thr
225 230 235 240
Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu Tyr Lys Gln Ile

245 250 255
Ser Ser Ser Ser Ser Gly Ala Thr Asn Asp Asn His Tyr Phe Gly Tyr
260 265 270
Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His Cys His Phe
275 280 285
Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Trp Gly Phe Arg
290 295 300
Pro Lys Lys Leu Arg Phe Lys Leu Phe Asn Ile Gln Val Lys Glu Val
305 310 315 320
Thr Thr Asn Asp Gly Val Thr Thr Ile Ala Asn Asn Leu Thr Ser Thr
325 330 335
Val Gln Val Phe Ser Asp Ser Glu Tyr Gln Leu Pro Tyr Val Leu Gly
340 345 350
Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala Asp Val Phe Met
355 360 365
Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly Ser Gln Ser Val
370 375 380
Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser Gln Met Leu
385 390 395 400
Arg Thr Gly Asn Asn Phe Glu Phe Ser Tyr Ser Phe Glu Asp Val Pro
405 410 415
Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp Arg Leu Met Asn
420 425 430
Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ala Arg Thr Gln Ser Thr
435 440 445
Thr Gly Ser Thr Arg Glu Leu Gln Phe His Gln Ala Gly Pro Asn Thr
450 455 460
Met Ala Glu Gln Ser Lys Asn Trp Leu Pro Gly Pro Cys Tyr Arg Gln
465 470 475 480
Gln Arg Leu Ser Lys Asn Leu Asp Phe Asn Asn Asn Ser Asn Phe Ala
[0139] 485 490 495
Trp Thr Ala Ala Thr Lys Tyr His Leu Asn Gly Arg Asn Ser Leu Thr
500 505 510
Asn Pro Gly Ile Pro Met Ala Thr Asn Lys Asp Asp Glu Asp Gln Phe
515 520 525
Phe Pro Ile Asn Gly Val Leu Val Phe Gly Lys Thr Gly Ala Ala Asn
530 535 540
Lys Thr Thr Leu Glu Asn Val Leu Met Thr Ser Glu Glu Glu Ile Lys
545 550 555 560
Thr Thr Asn Pro Val Ala Thr Glu Glu Tyr Gly Val Val Ser Ser Asn
565 570 575
Leu Gln Ser Ser Thr Ala Gly Pro Gln Ser Gln Thr Ile Asn Ser Gln
580 585 590
Gly Ala Leu Pro Gly Met Val Trp Gln Asn Arg Asp Val Tyr Leu Gln
595 600 605
Gly Pro Ile Trp Ala Lys Ile Pro His Thr Asp Gly Asn Phe His Pro
610 615 620
Ser Pro Leu Met Gly Gly Phe Gly Leu Glu His Pro Pro Pro Gln Ile
625 630 635 640
Leu Ile Lys Asn Thr Pro Val Pro Ala Asn Pro Pro Glu Val Phe Thr
645 650 655
Pro Ala Lys Phe Ala Ser Phe Ile Thr Gln Tyr Ser Thr Gly Gln Val
660 665 670
Ser Val Glu Ile Glu Trp Glu Leu Gln Lys Glu Asn Ser Lys Arg Trp
675 680 685
Asn Pro Glu Ile Gln Tyr Thr Ser Asn Tyr Ala Lys Ser Asn Asn Val
690 695 700
Glu Phe Ala Val Asn Pro Asp Gly Val Tyr Thr Glu Pro Arg Pro Ile
705 710 715 720
Gly Thr Arg Tyr Leu Thr Arg Asn Leu
725

<210> 111
<211> 729
<212> PRT
<213> AAV血清型的衣壳蛋白,克隆F3VP1

<400> 111

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 Asp Leu Lys Pro Gly Ala Pro Lys Pro
20 25 30
Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro
35 40 45
Gly Tyr Lys Tyr Leu Gly Pro Phe 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 Arg Tyr Asn His Ala
85 90 95

Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly
100 105 110
Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
115 120 125
Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg
130 135 140
Pro Ile Gly Ser Pro Asp Ser Ser Thr Gly Ile Gly Lys Lys Gly Gln
145 150 155 160
Gln Pro Ala Lys Lys Lys Leu Asn Phe Gly Gln Thr Gly Asp Ser Glu
165 170 175
Ser Val Pro Asp Pro Gln Pro Leu Gly Glu Pro Pro Ala Ala Pro Ser
180 185 190
Ser Val Gly Ser Gly Thr Met Ala Ala Gly Gly Gly Ala Pro Met Ala
195 200 205
Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ala Ser Gly Asn Trp
210 215 220
His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val Ile Thr Thr Ser Thr
225 230 235 240
Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu Tyr Lys Gln Ile
245 250 255
Ser Ser Ser Ser Ser Gly Ala Thr Asn Asp Asn His Tyr Phe Gly Tyr
260 265 270
Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His Cys His Phe
275 280 285
Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Trp Gly Phe Arg
290 295 300
Pro Lys Lys Leu Arg Phe Lys Leu Leu Asn Ile Gln Val Lys Glu Val
305 310 315 320
Thr Thr Asn Asp Gly Val Thr Thr Ile Ala Asn Asn Leu Thr Ser Thr
325 330 335
Val Gln Val Phe Ser Asp Ser Glu Tyr Gln Leu Pro Tyr Val Leu Gly
340 345 350
Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala Asp Val Phe Met
355 360 365
Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asp Asn Gly Ser Gln Ser Val
370 375 380
Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser Gln Met Leu
385 390 395 400
Arg Thr Gly Asn Asn Phe Glu Phe Ser Tyr Ser Phe Glu Asp Val Pro
405 410 415
Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp Arg Leu Met Asn
420 425 430
Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ala Arg Thr Gln Ser Thr

[0140]

435 440 445
Thr Gly Ser Thr Arg Glu Leu Gln Phe His Gln Ala Gly Pro Asn Thr
450 455 460
Met Ala Glu Gln Ser Lys Asn Trp Leu Pro Gly Pro Cys Tyr Arg Gln
465 470 475 480
Gln Arg Leu Ser Lys Asn Leu Asp Phe Asn Asn Asn Ser Asn Phe Ala
485 490 495
Trp Thr Ala Ala Thr Lys Tyr His Leu Asn Gly Arg Asn Ser Leu Thr
500 505 510
Asn Pro Gly Ile Pro Met Ala Thr Asn Lys Asp Asp Glu Asp Glu Phe
515 520 525
Phe Pro Ile Asn Gly Val Leu Val Phe Gly Lys Thr Gly Ala Ala Asn
530 535 540
Lys Thr Thr Leu Glu Asn Val Leu Met Thr Ser Glu Glu Glu Ile Lys
545 550 555 560
Thr Thr Asn Pro Val Ala Thr Glu Glu Tyr Gly Val Val Ser Ser Asn
565 570 575
Leu Gln Ser Ser Thr Ala Gly Pro Gln Ser Gln Thr Ile Asn Ser Gln
580 585 590
Gly Ala Leu Pro Gly Met Val Trp Gln Asn Arg Asp Val Tyr Leu Gln
595 600 605
Gly Pro Ile Trp Ala Lys Ile Pro His Thr Asp Gly Asn Phe His Pro
610 615 620
Ser Pro Leu Met Gly Gly Phe Gly Leu Lys His Pro Pro Pro Gln Ile
625 630 635 640
Leu Ile Lys Asn Thr Pro Val Pro Ala Asn Pro Pro Glu Val Phe Thr
645 650 655
Pro Ala Lys Phe Ala Ser Phe Ile Thr Gln Tyr Ser Thr Gly Gln Val
660 665 670
Ser Val Glu Ile Glu Trp Glu Leu Gln Lys Glu Asn Ser Lys Arg Trp
675 680 685
[0141] Asn Pro Glu Ile Gln Tyr Thr Ser Asn Tyr Ala Lys Ser Asn Asn Val
690 695 700
Glu Phe Ala Val Asn Pro Asp Gly Val Tyr Thr Glu Pro Arg Pro Ile
705 710 715 720
Gly Thr Arg Tyr Leu Thr Arg Asn Leu
725

<210> 112
<211> 735
<212> PRT
<213> AAV 血清型的衣壳蛋白, 克隆 42.6B

<400> 112
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 Asp Leu Lys Pro Gly Ala Pro Lys Pro
20 25 30
Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro
35 40 45
Gly Tyr Lys Tyr Leu Gly Pro Phe 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
Lys Gln Leu Glu Gln Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala
85 90 95
Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly
100 105 110
Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
115 120 125
Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg
130 135 140

[0142]

```
Pro Val Glu Pro Ser Pro Gln Arg Ser Pro Asp Ser Ser Thr Gly Ile
145                 150                 155                 160
Gly Lys Thr Gly Gln Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln
                165                 170                 175
Thr Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro
            180                 185                 190
Pro Ala Gly Pro Ser Gly Leu Gly Ser Gly Thr Met Ala Ala Gly Gly
        195                 200                 205
Gly Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser
    210                 215                 220
Ser Ser Gly Asn Trp His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val
225                 230                 235                 240
Ile Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His
                245                 250                 255
Leu Tyr Lys Gln Ile Ser Asn Gly Thr Ser Gly Gly Ser Thr Asn Asp
            260                 265                 270
Asn Thr Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn
        275                 280                 285
Arg Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn
    290                 295                 300
Asn Asn Trp Gly Phe Arg Pro Arg Lys Leu Arg Phe Lys Leu Phe Asn
305                 310                 315                 320
Ile Gln Val Lys Glu Val Thr Thr Asp Asp Gly Val Thr Thr Ile Ala
                325                 330                 335
Asn Asn Leu Thr Ser Thr Ile Gln Val Phe Ser Asp Ser Glu Tyr Gln
            340                 345                 350
Leu Pro Tyr Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe
        355                 360                 365
Pro Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn
    370                 375                 380
Asn Gly Ser Gln Ser Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr
385                 390                 395                 400
Phe Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Glu Phe Ser Tyr
                405                 410                 415
Thr Phe Glu Glu Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser
            420                 425                 430
Leu Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu
        435                 440                 445
Ala Arg Thr Gln Ser Thr Thr Gly Ser Thr Arg Glu Leu Gln Phe His
    450                 455                 460
Gln Ala Gly Pro Asn Thr Met Ala Glu Gln Ser Lys Asn Trp Leu Pro
465                 470                 475                 480
Gly Pro Cys Tyr Arg Gln Gln Arg Leu Ser Lys Asn Ile Asp Ser Asn
                485                 490                 495
Asn Asn Ser Asn Phe Ala Trp Thr Gly Ala Thr Lys Tyr His Leu Asn
            500                 505                 510
Gly Arg Asn Ser Leu Thr Asn Pro Gly Val Ala Met Ala Thr Asn Lys
        515                 520                 525
Asp Asp Glu Asp Gln Phe Phe Pro Ile Asn Gly Val Leu Val Phe Gly
    530                 535                 540
Lys Thr Gly Ala Ala Asn Lys Thr Thr Leu Glu Asn Val Leu Met Thr
545                 550                 555                 560
Ser Glu Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr Glu Glu Tyr
                565                 570                 575
Gly Val Val Ser Ser Asn Leu Gln Ser Ser Thr Ala Gly Pro Gln Thr
            580                 585                 590
Gln Thr Val Asn Ser Gln Gly Ala Leu Pro Gly Met Val Trp Gln Asn
        595                 600                 605
Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His Thr
    610                 615                 620
Asp Gly Asn Phe His Pro Ser Pro Leu Met Asp Gly Phe Gly Leu Lys
625                 630                 635                 640
```

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

<210> 113
<211> 685
<212> PRT
<213> AAV 血清型的衣壳蛋白, 克隆 42.12

<400> 113
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 Asp Leu Lys Pro Gly Ala Pro Lys Pro
20 25 30
Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro
35 40 45
Gly Tyr Lys Tyr Leu Gly Pro Phe 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
Lys Gln Leu Glu Gln Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala
85 90 95
[0143] Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly
100 105 110
Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
115 120 125
Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg
130 135 140
Pro Val Glu Pro Ser Pro Gln Arg Ser Pro Asp Ser Ser Thr Gly Ile
145 150 155 160
Gly Lys Thr Gly Gln Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln
165 170 175
Thr Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro
180 185 190
Pro Ala Gly Pro Ser Gly Leu Gly Ser Gly Thr Met Ala Ala Gly Gly
195 200 205
Gly Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser
210 215 220
Ser Ser Gly Asn Trp His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val
225 230 235 240
Ile Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His
245 250 255
Leu Tyr Lys Gln Ile Ser Asn Gly Thr Ser Gly Gly Ser Thr Asn Asp
260 265 270
Asn Thr Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn
275 280 285
Arg Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn
290 295 300
Asn Asn Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn
305 310 315 320
Ile Gln Val Lys Glu Val Thr Gln Asn Glu Gly Thr Lys Thr Ile Ala
325 330 335
Asn Asn Leu Thr Ser Thr Ile Gln Val Phe Thr Asp Ser Glu Tyr Gln

[0144]

```
            340                 345                 350
Leu Pro Tyr Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe
        355                 360                 365
Pro Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn
    370                 375                 380
Asn Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr
385                 390                 395                 400
Phe Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Glu Phe Ser Tyr
                405                 410                 415
Gln Phe Glu Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser
            420                 425                 430
Leu Asp Arg Leu Thr Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu
        435                 440                 445
Ala Arg Thr Gln Ser Thr Thr Gly Ser Thr Arg Gly Leu Gln Phe His
    450                 455                 460
Gln Ala Gly Pro Asn Thr Met Ala Glu Gln Ser Lys Asn Trp Leu Pro
465                 470                 475                 480
Gly Pro Cys Tyr Arg Gln Gln Arg Leu Ser Lys Asn Ile Asp Ser Asn
                485                 490                 495
Asn Asn Ser Asn Phe Ala Trp Thr Gly Ala Thr Lys Tyr His Leu Asn
            500                 505                 510
Gly Arg Asn Ser Leu Thr Asn Pro Gly Val Ala Met Ala Thr Asn Lys
        515                 520                 525
Asp Asp Glu Asp Gln Phe Phe Pro Ile Asn Gly Val Leu Val Phe Gly
    530                 535                 540
Lys Thr Gly Ala Ala Asn Lys Thr Thr Leu Glu Asn Val Leu Met Thr
545                 550                 555                 560
Ser Glu Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr Glu Glu Tyr
                565                 570                 575
Gly Val Val Ser Ser Asn Leu Gln Ser Ser Thr Ala Gly Pro Gln Thr
            580                 585                 590
Gln Thr Val Asn Ser Gln Gly Ala Leu Pro Gly Met Val Trp Gln Asn
        595                 600                 605
Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His Thr
    610                 615                 620
Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Leu Lys
625                 630                 635                 640
His Pro Pro Pro Gln Ile Leu Ile Lys Tyr Thr Ser Asn Tyr Tyr Lys
                645                 650                 655
Ser Thr Asn Val Asp Phe Ala Val Asn Thr Glu Gly Thr Tyr Ser Glu
            660                 665                 670
Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu
        675                 680                 685

<210>  114
<211>  724
<212>  PRT
<213>  AAV血清型的衣壳蛋白,克隆AAV5CAP

<400>  114
Met Ser Phe Val Asp His Pro Pro Asp Trp Leu Glu Glu Val Gly Glu
1               5                   10                  15
Gly Leu Arg Glu Phe Leu Gly Leu Glu Ala Gly Pro Pro Lys Pro Lys
            20                  25                  30
Pro Asn Gln Gln His Gln Asp Gln Ala Arg Gly Leu Val Leu Pro Gly
        35                  40                  45
Tyr Asn Tyr Leu Gly Pro Gly Asn Gly Leu Asp Arg Gly Glu Pro Val
    50                  55                  60
Asn Arg Ala Asp Glu Val Ala Arg Glu His Asp Ile Ser Tyr Asn Glu
65                  70                  75                  80
Gln Leu Glu Ala Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala Asp
                85                  90                  95
```

[0145]

```
Ala Glu Phe Gln Glu Lys Leu Ala Asp Asp Thr Ser Phe Gly Gly Asn
            100                 105                 110
Leu Gly Lys Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro Phe
        115                 120                 125
Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Thr Gly Lys Arg Ile
    130                 135                 140
Asp Asp His Phe Pro Lys Arg Lys Lys Ala Arg Thr Glu Glu Asp Ser
145                 150                 155                 160
Lys Pro Ser Thr Ser Ser Asp Ala Glu Ala Gly Pro Ser Gly Ser Gln
                165                 170                 175
Gln Leu Gln Ile Pro Ala Gln Pro Ala Ser Ser Leu Gly Ala Asp Thr
            180                 185                 190
Met Ser Ala Gly Gly Gly Gly Pro Leu Gly Asp Asn Asn Gln Gly Ala
        195                 200                 205
Asp Gly Val Gly Asn Ala Ser Gly Asp Trp His Cys Asp Ser Thr Trp
    210                 215                 220
Met Gly Asp Arg Val Val Thr Lys Ser Thr Arg Thr Trp Val Leu Pro
225                 230                 235                 240
Ser Tyr Asn Asn His Gln Tyr Arg Glu Ile Lys Ser Gly Ser Val Asp
                245                 250                 255
Gly Ser Asn Ala Asn Ala Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr
            260                 265                 270
Phe Asp Phe Asn Arg Phe His Ser His Trp Ser Pro Arg Asp Trp Gln
        275                 280                 285
Arg Leu Ile Asn Asn Tyr Trp Gly Phe Arg Pro Arg Ser Leu Arg Val
    290                 295                 300
Lys Ile Phe Asn Ile Gln Val Lys Glu Val Thr Val Gln Asp Ser Thr
305                 310                 315                 320
Thr Thr Ile Ala Asn Asn Leu Thr Ser Thr Val Gln Val Phe Thr Asp
                325                 330                 335
Asp Asp Tyr Gln Leu Pro Tyr Val Val Gly Asn Gly Thr Glu Gly Cys
            340                 345                 350
Leu Pro Ala Phe Pro Pro Gln Val Phe Thr Leu Pro Gln Tyr Gly Tyr
        355                 360                 365
Ala Thr Leu Asn Arg Asp Asn Thr Glu Asn Pro Thr Glu Arg Ser Ser
    370                 375                 380
Phe Phe Cys Leu Glu Tyr Phe Pro Ser Lys Met Leu Arg Thr Gly Asn
385                 390                 395                 400
Asn Phe Glu Phe Thr Tyr Asn Phe Glu Glu Val Pro Phe His Ser Ser
                405                 410                 415
Phe Ala Pro Ser Gln Asn Leu Phe Lys Leu Ala Asn Pro Leu Val Asp
            420                 425                 430
Gln Tyr Leu Tyr Arg Phe Val Ser Thr Asn Asn Thr Gly Gly Val Gln
        435                 440                 445
Phe Asn Lys Asn Leu Ala Gly Arg Tyr Ala Asn Thr Tyr Lys Asn Trp
    450                 455                 460
Phe Pro Gly Pro Met Gly Arg Thr Gln Gly Trp Asn Leu Gly Ser Gly
465                 470                 475                 480
Val Asn Arg Ala Ser Val Ser Ala Phe Ala Thr Thr Asn Arg Met Glu
                485                 490                 495
Leu Glu Gly Ala Ser Tyr Gln Val Pro Pro Gln Pro Asn Gly Met Thr
            500                 505                 510
Asn Asn Leu Gln Gly Ser Asn Thr Tyr Ala Leu Glu Asn Thr Met Ile
        515                 520                 525
Phe Asn Ser Gln Pro Ala Asn Pro Gly Thr Thr Ala Thr Tyr Leu Glu
    530                 535                 540
Gly Asn Met Leu Ile Thr Ser Glu Ser Glu Thr Gln Pro Val Asn Arg
545                 550                 555                 560
Val Ala Tyr Asn Val Gly Gly Gln Met Ala Thr Asn Asn Gln Ser Ser
                565                 570                 575
Thr Thr Ala Pro Ala Thr Gly Thr Tyr Asn Leu Gln Glu Ile Val Pro
            580                 585                 590
```

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

<210> 115
<211> 9
<212> DNA
<213> DraIII 限制性酶切位点

<400> 115
caccacgtc 9

<210> 116
<211> 28
<212> DNA
<213> AV2cas

[0146]

<400> 116
cgcagagacc aaagttcaac tgaaacga 28

<210> 117
<211> 255
<212> DNA
<213> 腺伴随病毒血清型 10

<400> 117
ggtaattcct ccggaaaatg gcattgcgat tccacatggc tgggcgacag agtcatcacc 60
accagcaccc gaacctgggt cctgccraac tacaacaacc acatctacaa gcaaatctcc 120
agcgagacag gagccaccaa cgacaaccac tacttcggct acagcacccc ctggggtat 180
tttgacttta acagattcca ctgccacttt tcaccacgtg actggcagcg actcatcaac 240
aacaactggg gattc 255

<210> 118
<211> 258
<212> DNA
<213> 腺伴随病毒血清型 11

<400> 118
ggtaattcct ccggaaaatg gcattgcgat tccacatggc tgggcgacag agtcatcacc 60
accagcaccc gaacctgggc cctgccaacc tacaacaacc acctctacaa acaaatctcc 120
agcgcttcaa cgggggccag caacgacaac cactactttg gctacagcac ccctgggggg 180

tattttgact ttaacagatt ccactgccac ttctcaccac gtgactggca gcgactcatc 240
aacaacaact ggggattc 258

<210> 119
<211> 255

<212> DNA
<213> 腺伴随病毒血清型 12

<400> 119
ggtaattcct ccggaaattg gcattgcgat tccacatggc tgggcgaccg agtcattacc 60
accagcaccc ggacttgggc cctgcccacc tacaacaacc acctctacaa gcaaatctcc 120
agccaatcgg gtgccaccaa cgacaaccac tacttcggct acagcacccc ttgggggtat 180
tttgatttca acagattcca ctgccatttc tcaccacgtg actggcagcg actcatcaac 240
aacaactggg gattc 255

<210> 120
<211> 2205
<212> DNA
<213> 腺伴随病毒血清型, 克隆 A3.1vp1

<400> 120
atggctgccg atggttatct tccagattgg ctcgaggaca tctctctga aggaatcaga 60
cagtggtgga agctcaaacc tggcccacca ccgccgaaac ctaaccaaca acaccgggac 120
gacagtaggg gtcttgtgct tcctgggtac aagtacctcg gacccttcaa cggactcgac 180
aaaggagagc cggtcaacga ggcagacgcc gcggccctcg agcacgacaa agcctacgac 240
caccagctca agcaagggga caacccgtac ctcaaataca accacgcgga cgctgaattt 300
caggagcgtc ttcaagaaga tacgtctttc gggggcaacc tcgggcgagc agtcttccag 360
gccaaaaaga gggtactcga gcctcttggt ctggttgagg aagctgttaa gacggctcct 420
ggaaaaaaga gacctataga gcagtctcct gcagaaccgg actcttcctc gggcatcggc 480
aaatcaggcc agcagcccgc taagaaaaga ctcaattttg gtcagactgg cgacacagag 540
tcagtcccag acctcaacc aatcggagaa cccccgcag ccccctctgg tgtgggatct 600
aatacaatgg cttcaggcgg tggggcacca atggcagaca ataacgaagg cgccgacgga 660
gtgggtaatt cctcgggaaa ttggcattgc gattccacat ggatgggcga cagagttatc 720
accaccagca caagaacctg ggccctcccc acctacaata atcacctcta caagcaaatc 780
tccagcgaat cgggagccac caacgacaac cactacttcg gctacagcac ccctggggg 840
tattttgact ttaacagatt ccactgtcac ttctcaccac gtgactggca gcgactcatc 900
aacaacaact ggggatttag acccaagaaa ctcaatttca agctcttcaa catccaagtc 960
aaggaggtca cgcagaatga tggaaccacg accatcgcca ataaccttac cagcacggtg 1020

caggtcttca cagactctga gtaccagctg ccctacgtcc tcggttcggc tcaccagggc 1080
tgccttccgc cgttcccagc agacgtcttc atgattcctc agtacggcta cttgactctg 1140
aacaatggca gccaagcggt aggacgttct tcattctact gtctagagta ttttcctct 1200
cagatgctga ggacgggaaa caacttcacc ttcagctaca cttttgaaga cgtgcctttc 1260
cacagcagct acgcgcacag ccagagtctg gatcggctga tgaatcctct cattgaccag 1320
tacctgtatt acctgagcaa aactcagggt acaagtggaa caacgcagca atcgagactg 1380
cagttcagcc aagctgggcc tagctccatg gctcagcagg ccaaaaactg gctaccggga 1440
cccagctacc gacagcagcg aatgtctaag acggctaatg acaacaacaa cagtgaattt 1500
gcttggactg cagccaccaa atattacctg aatggaagaa attctctggt caatcccggg 1560
ccccaatgg ccagtcacaa ggacgatgag gaaaagtatt tccccatgca cggaaatctc 1620
atctttggaa aacaaggcac aggaactacc aatgtggaca ttgaatcagt gcttattaca 1680
gacgaagaag aaatcagaac aactaatcct gtggctacag aacaatacgg acaggttgcc 1740
accaaccatc agagtcagaa caccacagct tcctatggaa gtgtggacag ccagggaatc 1800
ttacctggaa tggtgtggca ggaccgcgat gtctatcttc aaggtcccat ttgggccaaa 1860
actcctcaca cggacggaca ctttcatcct tctccgctca tgggaggctt tggactgaaa 1920
cacctcctc cccagatcct gatcaaaaac acacctgtgc cagcgaatcc cgcgaccact 1980
ttcactcctg gaaagtttgc ttcgttcatt acccagtatt ccaccggaca ggtcagcgtg 2040
gaaatagagt gggagctgca gaaagaaaac agcaaacgct ggaacccaga aattcagtac 2100
acctccaact acaacaagtc ggtgaatgtg gagtttaccg tggacgcaaa cggtgtttat 2160
tctgaacccc gcctattgg cactcgttac cttaccccgga acttg 2205

[0147]

```
          1                                                  50
  42_2    .......... .......... .......... .......... ..........
  42_8    .......... .......... .......... .......... ..........
 42_15    .......... .......... .......... .......... ..........
 42_5b    .......... .......... .......... .......... ..........
 42_1b    .......... .......... .......... .......... ..........
 42_13    .......... .......... .......... .......... ..........
 42_3a    .......... .......... .......... .......... ..........
  42_4    .......... .......... .......... .......... ..........
 42_5a    .......... .......... .......... .......... ..........
 42_10    .......... .......... .......... .......... ..........
 42_3b    .......... .......... .......... .......... ..........
 42_11    .......... .......... .......... .......... ..........
 42_6b    .......... .......... .......... .......... ..........
  43_1    .......... .......... .......... .......... ..........
  43_5    .......... .......... .......... .......... ..........
 43_12    .......... .......... .......... .......... ..........
 43_20    .......... .......... .......... .......... ..........
 43_21    .......... .......... .......... .......... ..........
 43_23    .......... .......... .......... .......... ..........
 43_25    .......... .......... .......... .......... ..........
  44_1    .......... .......... .......... .......... ..........
  44_5    .......... .......... .......... .......... ..........
223_10    .......... .......... .......... .......... ..........
 223_2    .......... .......... .......... .......... ..........
 223_4    .......... .......... .......... .......... ..........
 223_5    .......... .......... .......... .......... ..........
 223_6    .......... .......... .......... .......... ..........
 223_7    .......... .......... .......... .......... ..........
  A3_4    .......... .......... .......... .......... ..........
  A3_5    .......... .......... .......... .......... ..........
  A3_7    .......... .......... .......... .......... ..........
  A3_3    .......... .......... .......... .......... ..........
 42_12    .......... .......... .......... .......... ..........
  AAV1    TTGCCCACTC CCTCTCTGCG CGCTCGCTCG CTCGGTGGGG CCTGCGGACC
  AAV2    TTGGCCACTC CCTCTCTGCG CGCTCGCTCG CTCACTGAGG CCGGGCGACC
  AAV3    TTGGCCACTC CCTCTATGCG CACTCGCTCG CTCGGTGGGG CCTGGCGACC
  AAV8    .......... .......... .......... .......... ..........
  AAV9    .......... .......... .......... .......... ..........
  AAV7    TTGGCCACTC CCTCTATGCG CGCTCGCTCG CTCGGTGGGG CCTGCGGACC
  44_2    .......... .......... .......... .......... ..........
```

图1A

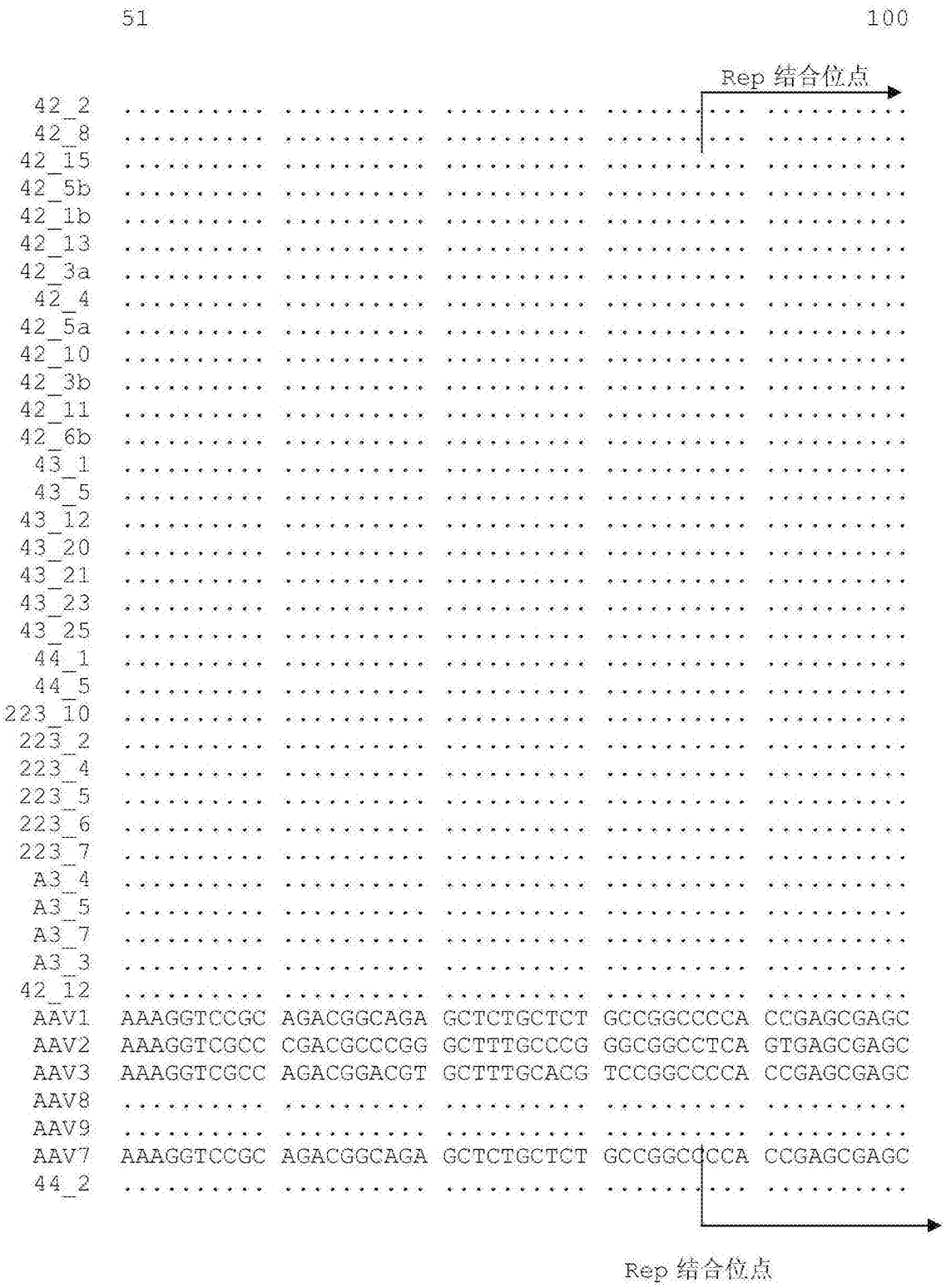

51
100
Rep 结合位点
42_2 ....
42_8 ....
42_15 ....
42_5b ....
42_1b ....
42_13 ....
42_3a ....
42_4 ....
42_5a ....
42_10 ....
42_3b ....
42_11 ....
42_6b ....
43_1 ....
43_5 ....
43_12 ....
43_20 ....
43_21 ....
43_23 ....
43_25 ....
44_1 ....
44_5 ....
223_10 ....
223_2 ....
223_4 ....
223_5 ....
223_6 ....
223_7 ....
A3_4 ....
A3_5 ....
A3_7 ....
A3_3 ....
42_12 ....
AAV1 AAAGGTCCGC AGACGGCAGA GCTCTGCTCT GCCGGCCCCA CCGAGCGAGC
AAV2 AAAGGTCGCC CGACGCCCGG GCTTTGCCCG GGCGGCCTCA GTGAGCGAGC
AAV3 AAAGGTCGCC AGACGGACGT GCTTTGCACG TCCGGCCCCA CCGAGCGAGC
AAV8 ....
AAV9 ....
AAV7 AAAGGTCCGC AGACGGCAGA GCTCTGCTCT GCCGGCCCCA CCGAGCGAGC
44_2 ....
Rep 结合位点

图1B

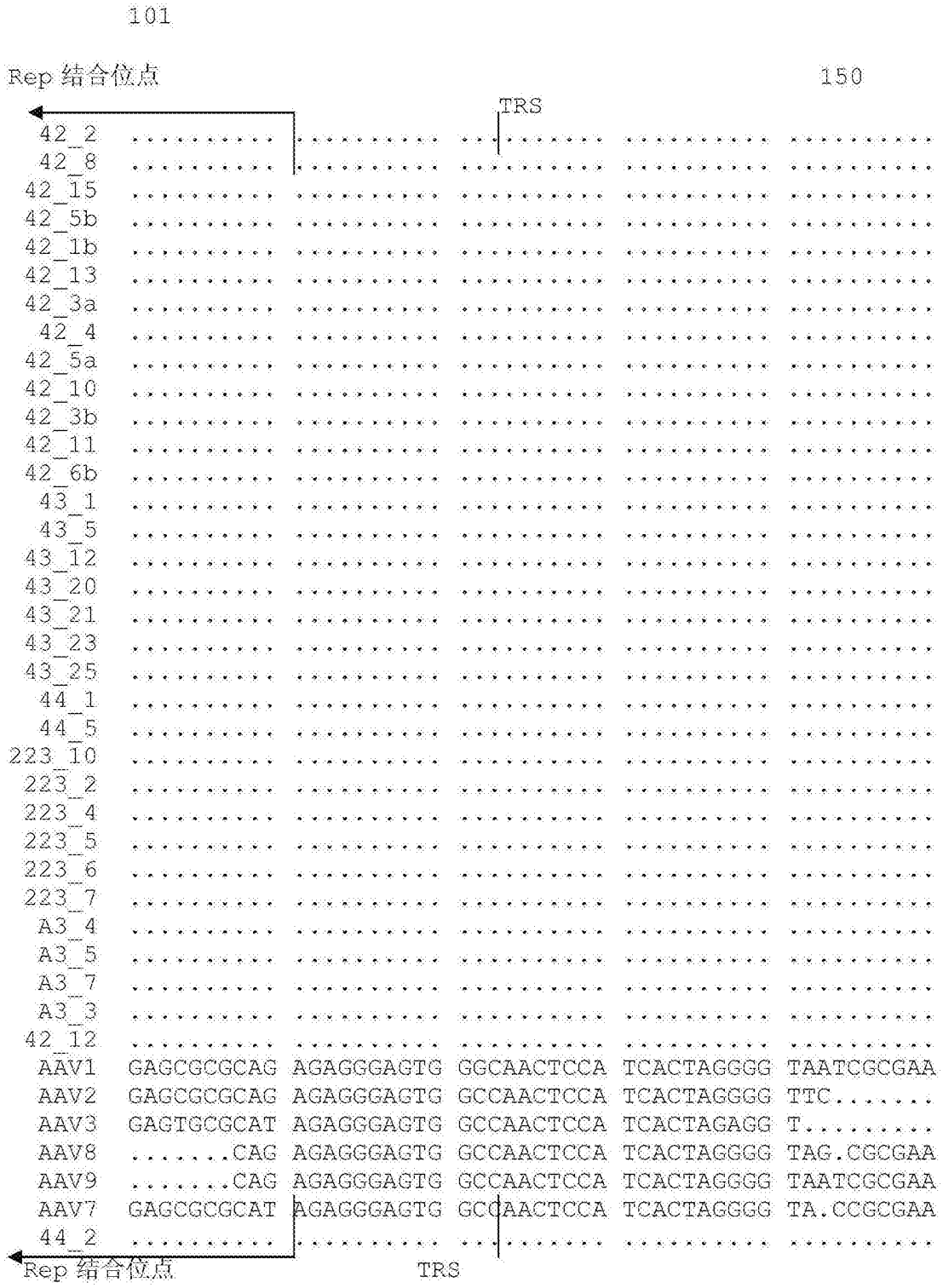

101
Rep 结合位点
150
TRS
42_2 ....
42_8 ....
42_15 ....
42_5b ....
42_1b ....
42_13 ....
42_3a ....
42_4 ....
42_5a ....
42_10 ....
42_3b ....
42_11 ....
42_6b ....
43_1 ....
43_5 ....
43_12 ....
43_20 ....
43_21 ....
43_23 ....
43_25 ....
44_1 ....
44_5 ....
223_10 ....
223_2 ....
223_4 ....
223_5 ....
223_6 ....
223_7 ....
A3_4 ....
A3_5 ....
A3_7 ....
A3_3 ....
42_12 ....
AAV1 GAGCGCGCAG AGAGGGAGTG GGCAACTCCA TCACTAGGGG TAATCGCGAA
AAV2 GAGCGCGCAG AGAGGGAGTG GCCAACTCCA TCACTAGGGG TTC.....
AAV3 GAGTGCGCAT AGAGGGAGTG GCCAACTCCA TCACTAGAGG T.....
AAV8 .....CAG AGAGGGAGTG GCCAACTCCA TCACTAGGGG TAG.CGCGAA
AAV9 .....CAG AGAGGGAGTG GCCAACTCCA TCACTAGGGG TAATCGCGAA
AAV7 GAGCGCGCAT AGAGGGAGTG GCCAACTCCA TCACTAGGGG TA.CCGCGAA
44_2 ....
Rep 结合位点
TRS

图1C

```
          151                                                        200
   42_2   .......... .......... .......... .......... ..........
   42_8   .......... .......... .......... .......... ..........
  42_15   .......... .......... .......... .......... ..........
  42_5b   .......... .......... .......... .......... ..........
  42_1b   .......... .......... .......... .......... ..........
  42_13   .......... .......... .......... .......... ..........
  42_3a   .......... .......... .......... .......... ..........
   42_4   .......... .......... .......... .......... ..........
  42_5a   .......... .......... .......... .......... ..........
  42_10   .......... .......... .......... .......... ..........
  42_3b   .......... .......... .......... .......... ..........
  42_11   .......... .......... .......... .......... ..........
  42_6b   .......... .......... .......... .......... ..........
   43_1   .......... .......... .......... .......... ..........
   43_5   .......... .......... .......... .......... ..........
  43_12   .......... .......... .......... .......... ..........
  43_20   .......... .......... .......... .......... ..........
  43_21   .......... .......... .......... .......... ..........
  43_23   .......... .......... .......... .......... ..........
  43_25   .......... .......... .......... .......... ..........
   44_1   .......... .......... .......... .......... ..........
   44_5   .......... .......... .......... .......... ..........
 223_10   .......... .......... .......... .......... ..........
  223_2   .......... .......... .......... .......... ..........
  223_4   .......... .......... .......... .......... ..........
  223_5   .......... .......... .......... .......... ..........
  223_6   .......... .......... .......... .......... ..........
  223_7   .......... .......... .......... .......... ..........
   A3_4   .......... .......... .......... .......... ..........
   A3_5   .......... .......... .......... .......... ..........
   A3_7   .......... .......... .......... .......... ..........
   A3_3   .......... .......... .......... .......... ..........
  42_12   .......... .......... .......... .......... ..........
   AAV1   GCGCCTCCCA CGCTGCCGCG TCAGCGCTGA CGTAAATTAC GTCATAGGGG
   AAV2   .......CTG GAGGGGTGGA GTCGTGACGT GAATTACGTC ATAGGGTTAG
   AAV3   .......ATG GCAGTGACGT AACGCGAAGC GCGCGAAGCG AGACCACGCC
   AAV8   GCGCCTCCCA CGCTGCCGCG TCAGCGCTGA CGTAAATTAC GTCATAGGGG
   AAV9   GCGCCTCCCA CGCTGCCGCG TCAGCGCTGA CGTAGATTAC GTCATAGGGG
   AAV7   GCGCCTCCCA CGCTGCCGCG TCAGCGCTGA CGTAAATCAC GTCATAGGGG
   44_2   .......... .......... .......... .......... ..........
```

图1D

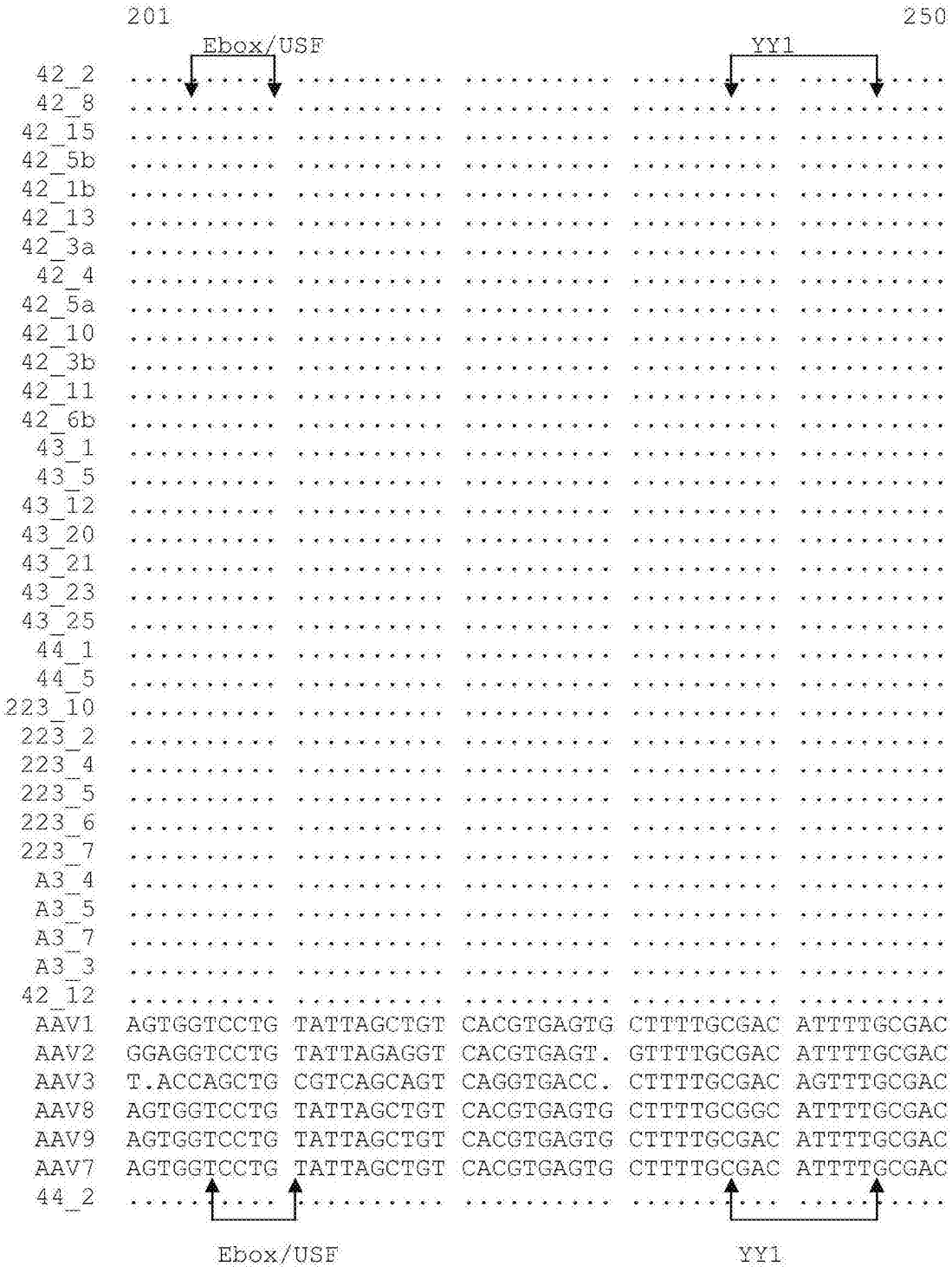

201
250
Ebox/USF
YY1
42_2
42_8
42_15
42_5b
42_1b
42_13
42_3a
42_4
42_5a
42_10
42_3b
42_11
42_6b
43_1
43_5
43_12
43_20
43_21
43_23
43_25
44_1
44_5
223_10
223_2
223_4
223_5
223_6
223_7
A3_4
A3_5
A3_7
A3_3
42_12
AAV1 AGTGGTCCTG TATTAGCTGT CACGTGAGTG CTTTTGCGAC ATTTTGCGAC
AAV2 GGAGGTCCTG TATTAGAGGT CACGTGAGT. GTTTTGCGAC ATTTTGCGAC
AAV3 T.ACCAGCTG CGTCAGCAGT CAGGTGACC. CTTTTGCGAC AGTTTGCGAC
AAV8 AGTGGTCCTG TATTAGCTGT CACGTGAGTG CTTTTGCGGC ATTTTGCGAC
AAV9 AGTGGTCCTG TATTAGCTGT CACGTGAGTG CTTTTGCGAC ATTTTGCGAC
AAV7 AGTGGTCCTG TATTAGCTGT CACGTGAGTG CTTTTGCGAC ATTTTGCGAC
44_2
Ebox/USF
YY1

图1E

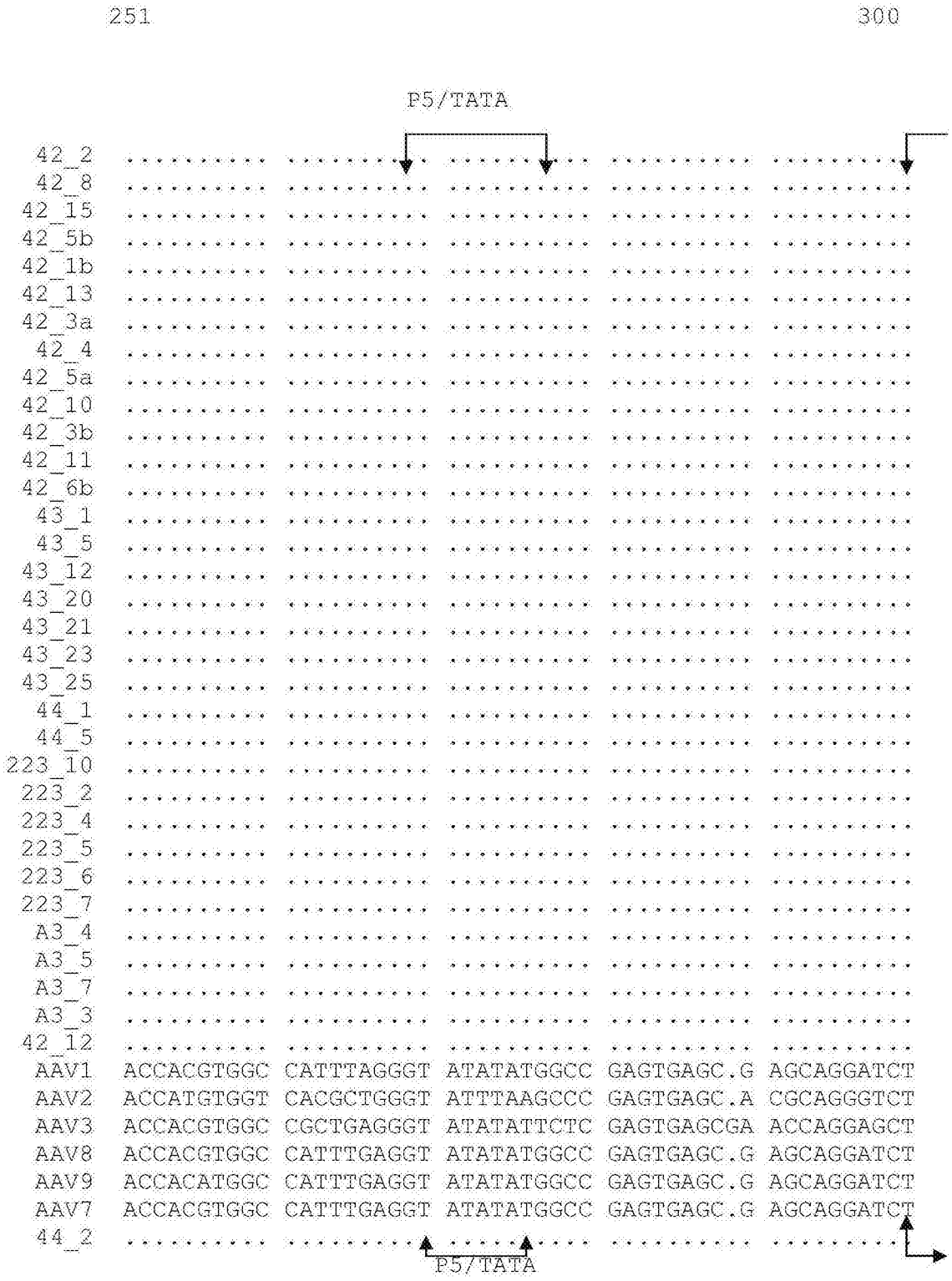

251
300
P5/TATA
42_2
42_8
42_15
42_5b
42_1b
42_13
42_3a
42_4
42_5a
42_10
42_3b
42_11
42_6b
43_1
43_5
43_12
43_20
43_21
43_23
43_25
44_1
44_5
223_10
223_2
223_4
223_5
223_6
223_7
A3_4
A3_5
A3_7
A3_3
42_12
AAV1 ACCACGTGGC CATTTAGGGT ATATATGGCC GAGTGAGC.G AGCAGGATCT
AAV2 ACCATGTGGT CACGCTGGGT ATTTAAGCCC GAGTGAGC.A CGCAGGGTCT
AAV3 ACCACGTGGC CGCTGAGGGT ATATATTCTC GAGTGAGCGA ACCAGGAGCT
AAV8 ACCACGTGGC CATTTGAGGT ATATATGGCC GAGTGAGC.G AGCAGGATCT
AAV9 ACCACATGGC CATTTGAGGT ATATATGGCC GAGTGAGC.G AGCAGGATCT
AAV7 ACCACGTGGC CATTTGAGGT ATATATGGCC GAGTGAGC.G AGCAGGATCT
44_2
P5/TATA

图1F

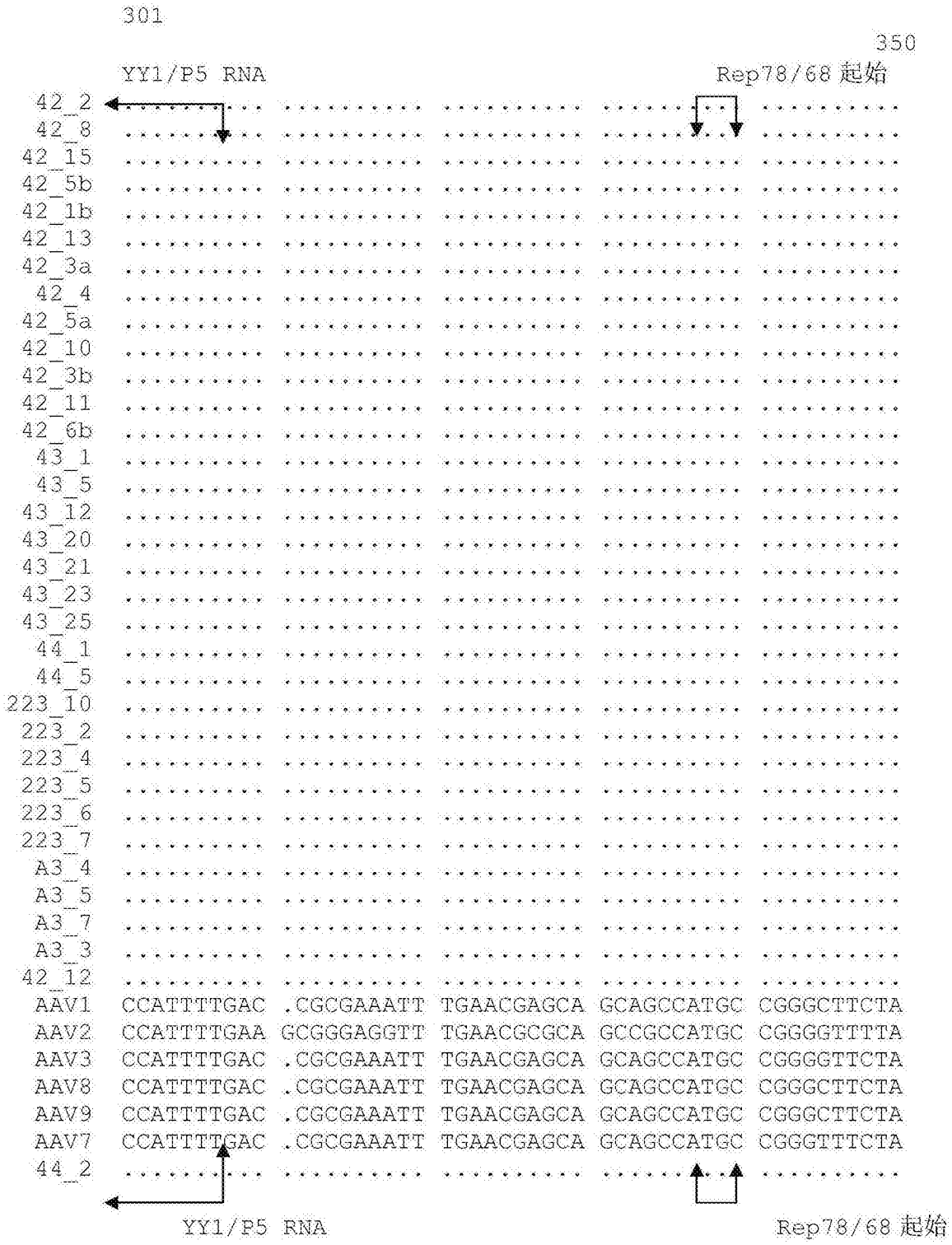

301
350
YY1/P5 RNA
Rep78/68 起始
42_2
42_8
42_15
42_5b
42_1b
42_13
42_3a
42_4
42_5a
42_10
42_3b
42_11
42_6b
43_1
43_5
43_12
43_20
43_21
43_23
43_25
44_1
44_5
223_10
223_2
223_4
223_5
223_6
223_7
A3_4
A3_5
A3_7
A3_3
42_12
AAV1 CCATTTTGAC .CGCGAAATT TGAACGAGCA GCAGCCATGC CGGGCTTCTA
AAV2 CCATTTTGAA GCGGGAGGTT TGAACGCGCA GCCGCCATGC CGGGGTTTTA
AAV3 CCATTTTGAC .CGCGAAATT TGAACGAGCA GCAGCCATGC CGGGGTTCTA
AAV8 CCATTTTGAC .CGCGAAATT TGAACGAGCA GCAGCCATGC CGGGCTTCTA
AAV9 CCATTTTGAC .CGCGAAATT TGAACGAGCA GCAGCCATGC CGGGCTTCTA
AAV7 CCATTTTGAC .CGCGAAATT TGAACGAGCA GCAGCCATGC CGGGTTTCTA
44_2
YY1/P5 RNA
Rep78/68 起始

图1G

```
          351                                                        400
  42_2    .......... .......... .......... .......... ..........
  42_8    .......... .......... .......... .......... ..........
 42_15    .......... .......... .......... .......... ..........
 42_5b    .......... .......... .......... .......... ..........
 42_1b    .......... .......... .......... .......... ..........
 42_13    .......... .......... .......... .......... ..........
 42_3a    .......... .......... .......... .......... ..........
  42_4    .......... .......... .......... .......... ..........
 42_5a    .......... .......... .......... .......... ..........
 42_10    .......... .......... .......... .......... ..........
 42_3b    .......... .......... .......... .......... ..........
 42_11    .......... .......... .......... .......... ..........
 42_6b    .......... .......... .......... .......... ..........
  43_1    .......... .......... .......... .......... ..........
  43_5    .......... .......... .......... .......... ..........
 43_12    .......... .......... .......... .......... ..........
 43_20    .......... .......... .......... .......... ..........
 43_21    .......... .......... .......... .......... ..........
 43_23    .......... .......... .......... .......... ..........
 43_25    .......... .......... .......... .......... ..........
  44_1    .......... .......... .......... .......... ..........
  44_5    .......... .......... .......... .......... ..........
223_10    .......... .......... .......... .......... ..........
 223_2    .......... .......... .......... .......... ..........
 223_4    .......... .......... .......... .......... ..........
 223_5    .......... .......... .......... .......... ..........
 223_6    .......... .......... .......... .......... ..........
 223_7    .......... .......... .......... .......... ..........
  A3_4    .......... .......... .......... .......... ..........
  A3_5    .......... .......... .......... .......... ..........
  A3_7    .......... .......... .......... .......... ..........
  A3_3    .......... .......... .......... .......... ..........
 42_12    .......... .......... .......... .......... ..........
  AAV1    CGAGATCGTG ATCAAGGTGC CGAGCGACCT GGACGAGCAC CTGCCGGGCA
  AAV2    CGAGATTGTG ATTAAGGTCC CCAGCGACCT TGACGGGCAT CTGCCCGGCA
  AAV3    CGAGATTGTC CTGAAGGTCC CGAGTGACCT GGACGAGCGC CTGCCGGGCA
  AAV8    CGAGATCGTG ATCAAGGTGC CGAGCGACCT GGACGAGCAC CTGCCGGGCA
  AAV9    CGAGATTGTG ATCAAGGTGC CGAGCGACCT GGACGAGCAC CTGCCGGGCA
  AAV7    CGAGATCGTG ATCAAGGTGC CGAGCGACCT GGACGAGCAC CTGCCGGGCA
  44_2    .......... .......... .......... .......... ..........
```

图1H

```
          401                                                        450
  42_2    .......... .......... .......... .......... ..........
  42_8    .......... .......... .......... .......... ..........
 42_15    .......... .......... .......... .......... ..........
 42_5b    .......... .......... .......... .......... ..........
 42_1b    .......... .......... .......... .......... ..........
 42_13    .......... .......... .......... .......... ..........
 42_3a    .......... .......... .......... .......... ..........
  42_4    .......... .......... .......... .......... ..........
 42_5a    .......... .......... .......... .......... ..........
 42_10    .......... .......... .......... .......... ..........
 42_3b    .......... .......... .......... .......... ..........
 42_11    .......... .......... .......... .......... ..........
 42_6b    .......... .......... .......... .......... ..........
  43_1    .......... .......... .......... .......... ..........
  43_5    .......... .......... .......... .......... ..........
 43_12    .......... .......... .......... .......... ..........
 43_20    .......... .......... .......... .......... ..........
 43_21    .......... .......... .......... .......... ..........
 43_23    .......... .......... .......... .......... ..........
 43_25    .......... .......... .......... .......... ..........
  44_1    .......... .......... .......... .......... ..........
  44_5    .......... .......... .......... .......... ..........
223_10    .......... .......... .......... .......... ..........
 223_2    .......... .......... .......... .......... ..........
 223_4    .......... .......... .......... .......... ..........
 223_5    .......... .......... .......... .......... ..........
 223_6    .......... .......... .......... .......... ..........
 223_7    .......... .......... .......... .......... ..........
  A3_4    .......... .......... .......... .......... ..........
  A3_5    .......... .......... .......... .......... ..........
  A3_7    .......... .......... .......... .......... ..........
  A3_3    .......... .......... .......... .......... ..........
 42_12    .......... .......... .......... .......... ..........
  AAV1    TTTCTGACTC GTTTGTGAGC TGGGTGGCCG AGAAGGAATG GGAGCTGCCC
  AAV2    TTTCTGACAG CTTTGTGAAC TGGGTGGCCG AGAAGGAATG GGAGTTGCCG
  AAV3    TTTCTAACTC GTTTGTTAAC TGGGTGGCCG AGAAGGAATG GGACGTGCCG
  AAV8    TTTCTGACTC GTTTGTGAAC TGGGTGGCCG AGAAGGAATG GGAGCTGCCC
  AAV9    TTTCTGACTC TTTTGTGAAC TGGGTGGCCG AGAAGGAATG GGAGCTGCCC
  AAV7    TTTCTGACTC GTTTGTGAAC TGGGTGGCCG AGAAGGAATG GGAGCTGCCC
  44_2    .......... .......... .......... .......... ..........
```

图1I

```
          451                                                        500
  42_2    .......... .......... .......... .......... ..........
  42_8    .......... .......... .......... .......... ..........
 42_15    .......... .......... .......... .......... ..........
 42_5b    .......... .......... .......... .......... ..........
 42_1b    .......... .......... .......... .......... ..........
 42_13    .......... .......... .......... .......... ..........
 42_3a    .......... .......... .......... .......... ..........
  42_4    .......... .......... .......... .......... ..........
 42_5a    .......... .......... .......... .......... ..........
 42_10    .......... .......... .......... .......... ..........
 42_3b    .......... .......... .......... .......... ..........
 42_11    .......... .......... .......... .......... ..........
 42_6b    .......... .......... .......... .......... ..........
  43_1    .......... .......... .......... .......... ..........
  43_5    .......... .......... .......... .......... ..........
 43_12    .......... .......... .......... .......... ..........
 43_20    .......... .......... .......... .......... ..........
 43_21    .......... .......... .......... .......... ..........
 43_23    .......... .......... .......... .......... ..........
 43_25    .......... .......... .......... .......... ..........
  44_1    .......... .......... .......... .......... ..........
  44_5    .......... .......... .......... .......... ..........
223_10    .......... .......... .......... .......... ..........
 223_2    .......... .......... .......... .......... ..........
 223_4    .......... .......... .......... .......... ..........
 223_5    .......... .......... .......... .......... ..........
 223_6    .......... .......... .......... .......... ..........
 223_7    .......... .......... .......... .......... ..........
  A3_4    .......... .......... .......... .......... ..........
  A3_5    .......... .......... .......... .......... ..........
  A3_7    .......... .......... .......... .......... ..........
  A3_3    .......... .......... .......... .......... ..........
 42_12    .......... .......... .......... .......... ..........
  AAV1    CCGGATTCTG ACATGGATCT GAATCTGATT GAGCAGGCAC CCCTGACCGT
  AAV2    CCAGATTCTG ACATGGATCT GAATCTGATT GAGCAGGCAC CCCTGACCGT
  AAV3    CCGGATTCTG ACATGGATCC GAATCTGATT GAGCAGGCAC CCCTGACCGT
  AAV8    CCGGATTCTG ACATGGATCG GAATCTGATC GAGCAGGCAC CCCTGACCGT
  AAV9    CCGGATTCTG ACATGGATCG GAATCTGATC GAGCAGGCAC CCCTGACCGT
  AAV7    CCGGATTCTG ACATGGATCT GAATCTGATC GAGCAGGCAC CCCTGACCGT
  44_2    .......... .......... .......... .......... ..........
```

图1J

```
          501                                                        550
  42_2    .......... .......... .......... .......... ..........
  42_8    .......... .......... .......... .......... ..........
 42_15    .......... .......... .......... .......... ..........
 42_5b    .......... .......... .......... .......... ..........
 42_1b    .......... .......... .......... .......... ..........
 42_13    .......... .......... .......... .......... ..........
 42_3a    .......... .......... .......... .......... ..........
  42_4    .......... .......... .......... .......... ..........
 42_5a    .......... .......... .......... .......... ..........
 42_10    .......... .......... .......... .......... ..........
 42_3b    .......... .......... .......... .......... ..........
 42_11    .......... .......... .......... .......... ..........
 42_6b    .......... .......... .......... .......... ..........
  43_1    .......... .......... .......... .......... ..........
  43_5    .......... .......... .......... .......... ..........
 43_12    .......... .......... .......... .......... ..........
 43_20    .......... .......... .......... .......... ..........
 43_21    .......... .......... .......... .......... ..........
 43_23    .......... .......... .......... .......... ..........
 43_25    .......... .......... .......... .......... ..........
  44_1    .......... .......... .......... .......... ..........
  44_5    .......... .......... .......... .......... ..........
223_10    .......... .......... .......... .......... ..........
 223_2    .......... .......... .......... .......... ..........
 223_4    .......... .......... .......... .......... ..........
 223_5    .......... .......... .......... .......... ..........
 223_6    .......... .......... .......... .......... ..........
 223_7    .......... .......... .......... .......... ..........
  A3_4    .......... .......... .......... .......... ..........
  A3_5    .......... .......... .......... .......... ..........
  A3_7    .......... .......... .......... .......... ..........
  A3_3    .......... .......... .......... .......... ..........
 42_12    .......... .......... .......... .......... ..........
  AAV1    GGCCGAGAAG CTGCAGCGCG ACTTCCTGGT CCAATGGCGC CGCGTGAGTA
  AAV2    GGCCGAGAAG CTGCAGCGCG ACTTTCTGAC GGAATGGCGC CGTGTGAGTA
  AAV3    GGCCGAAAAG CTTCAGCGCG AGTTCCTGGT GGAGTGGCGC CGCGTGAGTA
  AAV8    GGCCGAGAAG CTGCAGCGCG ACTTCCTGGT CCAATGGCGC CGCGTGAGTA
  AAV9    GGCCGAGAAG CTGTAGCGCG ACTTCCTGGT CCAATGGCGC CGCGTGAGTA
  AAV7    GGCCGAGAAG CTGCAGCGCG ACTTCCTGGT CCAATGGCGC CGCGTGAGTA
  44_2    .......... .......... .......... .......... ..........
```

图1K

```
          551                                                        600
  42_2    .......... .......... .......... .......... ..........
  42_8    .......... .......... .......... .......... ..........
 42_15    .......... .......... .......... .......... ..........
 42_5b    .......... .......... .......... .......... ..........
 42_1b    .......... .......... .......... .......... ..........
 42_13    .......... .......... .......... .......... ..........
 42_3a    .......... .......... .......... .......... ..........
  42_4    .......... .......... .......... .......... ..........
 42_5a    .......... .......... .......... .......... ..........
 42_10    .......... .......... .......... .......... ..........
 42_3b    .......... .......... .......... .......... ..........
 42_11    .......... .......... .......... .......... ..........
 42_6b    .......... .......... .......... .......... ..........
  43_1    .......... .......... .......... .......... ..........
  43_5    .......... .......... .......... .......... ..........
 43_12    .......... .......... .......... .......... ..........
 43_20    .......... .......... .......... .......... ..........
 43_21    .......... .......... .......... .......... ..........
 43_23    .......... .......... .......... .......... ..........
 43_25    .......... .......... .......... .......... ..........
  44_1    .......... .......... .......... .......... ..........
  44_5    .......... .......... .......... .......... ..........
223_10    .......... .......... .......... .......... ..........
 223_2    .......... .......... .......... .......... ..........
 223_4    .......... .......... .......... .......... ..........
 223_5    .......... .......... .......... .......... ..........
 223_6    .......... .......... .......... .......... ..........
 223_7    .......... .......... .......... .......... ..........
  A3_4    .......... .......... .......... .......... ..........
  A3_5    .......... .......... .......... .......... ..........
  A3_7    .......... .......... .......... .......... ..........
  A3_3    .......... .......... .......... .......... ..........
 42_12    .......... .......... .......... .......... ..........
  AAV1    AGGCCCCGGA GGCCCTCTTC TTTGTTCAGT TCGAGAAGGG CGAGTCCTAC
  AAV2    AGGCCCCGGA GGCCCTTTTC TTTGTGCAAT TTGAGAAGGG AGAGAGCTAC
  AAV3    AGGCCCCGGA GGCCCTCTTT TTTGTCCAGT TCGAAAAGGG GGAGACCTAC
  AAV8    AGGCCCCGGA GGCCCTCTTC TTTGTTCAGT TCGAGAAGGG CGAGAGCTAC
  AAV9    AGGCCCCGGA GGCCCTCTTC TTTGTTCAGT TCGAGAAGGG CGAGAGCTAC
  AAV7    AGGCCCCGGA GGCCCTGTTC TTTGTTCAGT TCGAGAAGGG CGAGAGCTAC
  44_2    .......... .......... .......... .......... ..........
```

图1L

```
          601                                                        650
  42_2    .......... .......... .......... .......... ..........
  42_8    .......... .......... .......... .......... ..........
 42_15    .......... .......... .......... .......... ..........
 42_5b    .......... .......... .......... .......... ..........
 42_1b    .......... .......... .......... .......... ..........
 42_13    .......... .......... .......... .......... ..........
 42_3a    .......... .......... .......... .......... ..........
  42_4    .......... .......... .......... .......... ..........
 42_5a    .......... .......... .......... .......... ..........
 42_10    .......... .......... .......... .......... ..........
 42_3b    .......... .......... .......... .......... ..........
 42_11    .......... .......... .......... .......... ..........
 42_6b    .......... .......... .......... .......... ..........
  43_1    .......... .......... .......... .......... ..........
  43_5    .......... .......... .......... .......... ..........
 43_12    .......... .......... .......... .......... ..........
 43_20    .......... .......... .......... .......... ..........
 43_21    .......... .......... .......... .......... ..........
 43_23    .......... .......... .......... .......... ..........
 43_25    .......... .......... .......... .......... ..........
  44_1    .......... .......... .......... .......... ..........
  44_5    .......... .......... .......... .......... ..........
223_10    .......... .......... .......... .......... ..........
 223_2    .......... .......... .......... .......... ..........
 223_4    .......... .......... .......... .......... ..........
 223_5    .......... .......... .......... .......... ..........
 223_6    .......... .......... .......... .......... ..........
 223_7    .......... .......... .......... .......... ..........
  A3_4    .......... .......... .......... .......... ..........
  A3_5    .......... .......... .......... .......... ..........
  A3_7    .......... .......... .......... .......... ..........
  A3_3    .......... .......... .......... .......... ..........
 42_12    .......... .......... .......... .......... ..........
  AAV1    TTCCACCTCC ATATTCTGGT GGAGACCACG GGGGTCAAAT CCATGGTGCT
  AAV2    TTCCACATGC ACGTGCTCGT GGAAACCACC GGGGTGAAAT CCATGGTTTT
  AAV3    TTCCACCTGC ACGTGCTGAT TGAGACCATC GGGGTCAAAT CCATGGTGGT
  AAV8    TTTCACCTGC ACGTTCTGGT CGAGACCACG GGGGTCAAGT CCATGGTGCT
  AAV9    TTTCACCTGC ACGTTCTGGT CGAGACCACG GGGGTCAAGT CCATGGTGCT
  AAV7    TTCCACCTTC ACGTTCTGGT GGAGACCACG GGGGTCAAGT CCATGGTGCT
  44_2    .......... .......... .......... .......... ..........
```

图1M

```
          651                                                        700
  42_2    .......... .......... .......... .......... ..........
  42_8    .......... .......... .......... .......... ..........
 42_15    .......... .......... .......... .......... ..........
 42_5b    .......... .......... .......... .......... ..........
 42_1b    .......... .......... .......... .......... ..........
 42_13    .......... .......... .......... .......... ..........
 42_3a    .......... .......... .......... .......... ..........
  42_4    .......... .......... .......... .......... ..........
 42_5a    .......... .......... .......... .......... ..........
 42_10    .......... .......... .......... .......... ..........
 42_3b    .......... .......... .......... .......... ..........
 42_11    .......... .......... .......... .......... ..........
 42_6b    .......... .......... .......... .......... ..........
  43_1    .......... .......... .......... .......... ..........
  43_5    .......... .......... .......... .......... ..........
 43_12    .......... .......... .......... .......... ..........
 43_20    .......... .......... .......... .......... ..........
 43_21    .......... .......... .......... .......... ..........
 43_23    .......... .......... .......... .......... ..........
 43_25    .......... .......... .......... .......... ..........
  44_1    .......... .......... .......... .......... ..........
  44_5    .......... .......... .......... .......... ..........
223_10    .......... .......... .......... .......... ..........
 223_2    .......... .......... .......... .......... ..........
 223_4    .......... .......... .......... .......... ..........
 223_5    .......... .......... .......... .......... ..........
 223_6    .......... .......... .......... .......... ..........
 223_7    .......... .......... .......... .......... ..........
  A3_4    .......... .......... .......... .......... ..........
  A3_5    .......... .......... .......... .......... ..........
  A3_7    .......... .......... .......... .......... ..........
  A3_3    .......... .......... .......... .......... ..........
 42_12    .......... .......... .......... .......... ..........
  AAV1    GGGCCGCTTC CTGAGTCAGA TTAGGGACAA GCT.GGTGCA GACCATCTAC
  AAV2    GGGACGTTTC CTGAGTCAGA TTCGCGAAAA ACT..GATTC AGAGAATTTA
  AAV3    CGGCCGCTAC GTGAGCCAGA TTAAAGAGAA GCT..GGTGA CCCGCATCTA
  AAV8    AGGCCGCTTC CTGAGTCAGA TTCGGGAAAA GCTTGGTCCA GACCATCTAC
  AAV9    AGGCCGCTTC CTGAGTCAGA TTCGGGAGAA GCT.GGTCCA GACCATCTAC
  AAV7    AGGCCGCTTC CTGAGTCAGA TTCGGGAGAA GCT....G.. GTCCAGACCA
  44_2    .......... .......... .......... .......... ..........
```

图1N

```
          701                                                       750
  42_2    .......... .......... .......... .......... ..........
  42_8    .......... .......... .......... .......... ..........
 42_15    .......... .......... .......... .......... ..........
 42_5b    .......... .......... .......... .......... ..........
 42_1b    .......... .......... .......... .......... ..........
 42_13    .......... .......... .......... .......... ..........
 42_3a    .......... .......... .......... .......... ..........
  42_4    .......... .......... .......... .......... ..........
 42_5a    .......... .......... .......... .......... ..........
 42_10    .......... .......... .......... .......... ..........
 42_3b    .......... .......... .......... .......... ..........
 42_11    .......... .......... .......... .......... ..........
 42_6b    .......... .......... .......... .......... ..........
  43_1    .......... .......... .......... .......... ..........
  43_5    .......... .......... .......... .......... ..........
 43_12    .......... .......... .......... .......... ..........
 43_20    .......... .......... .......... .......... ..........
 43_21    .......... .......... .......... .......... ..........
 43_23    .......... .......... .......... .......... ..........
 43_25    .......... .......... .......... .......... ..........
  44_1    .......... .......... .......... .......... ..........
  44_5    .......... .......... .......... .......... ..........
223_10    .......... .......... .......... .......... ..........
 223_2    .......... .......... .......... .......... ..........
 223_4    .......... .......... .......... .......... ..........
 223_5    .......... .......... .......... .......... ..........
 223_6    .......... .......... .......... .......... ..........
 223_7    .......... .......... .......... .......... ..........
  A3_4    .......... .......... .......... .......... ..........
  A3_5    .......... .......... .......... .......... ..........
  A3_7    .......... .......... .......... .......... ..........
  A3_3    .......... .......... .......... .......... ..........
 42_12    .......... .......... .......... .......... ..........
  AAV1    C.GCGGGATC GAGCCG.ACC CTGCCCAACT GGTTCGCGGT GACCAA.GAC
  AAV2    CCGCGGGATC GAGCCG.ACT TTGCCAAACT GGTTCGCGGT CACAAA...G
  AAV3    CCGCGGGGTC GAGCCG.CAG CTTCCGAACT GGTTCGCGGT GACCAA...A
  AAV8    CCGCGGGGTC GAGCCCCACC TTGCCCAACT GGTTCGCGGT GACCAAAGAC
  AAV9    C.GCGGGATC GAGCCG.ACC CTGCCCAACT GGTTCGCGGT GACCAA.GAC
  AAV7    TCTACCGCGG GGTCGAGCCC ACGCTGCCCA ACTGGTTCGC GGTGACCAAG
  44_2    .......... .......... .......... .......... ..........
```

图10

```
          751                                                       800
  42_2    .......... .......... .......... .......... ..........
  42_8    .......... .......... .......... .......... ..........
 42_15    .......... .......... .......... .......... ..........
 42_5b    .......... .......... .......... .......... ..........
 42_1b    .......... .......... .......... .......... ..........
 42_13    .......... .......... .......... .......... ..........
 42_3a    .......... .......... .......... .......... ..........
  42_4    .......... .......... .......... .......... ..........
 42_5a    .......... .......... .......... .......... ..........
 42_10    .......... .......... .......... .......... ..........
 42_3b    .......... .......... .......... .......... ..........
 42_11    .......... .......... .......... .......... ..........
 42_6b    .......... .......... .......... .......... ..........
  43_1    .......... .......... .......... .......... ..........
  43_5    .......... .......... .......... .......... ..........
 43_12    .......... .......... .......... .......... ..........
 43_20    .......... .......... .......... .......... ..........
 43_21    .......... .......... .......... .......... ..........
 43_23    .......... .......... .......... .......... ..........
 43_25    .......... .......... .......... .......... ..........
  44_1    .......... .......... .......... .......... ..........
  44_5    .......... .......... .......... .......... ..........
223_10    .......... .......... .......... .......... ..........
 223_2    .......... .......... .......... .......... ..........
 223_4    .......... .......... .......... .......... ..........
 223_5    .......... .......... .......... .......... ..........
 223_6    .......... .......... .......... .......... ..........
 223_7    .......... .......... .......... .......... ..........
  A3_4    .......... .......... .......... .......... ..........
  A3_5    .......... .......... .......... .......... ..........
  A3_7    .......... .......... .......... .......... ..........
  A3_3    .......... .......... .......... .......... ..........
 42_12    .......... .......... .......... .......... ..........
  AAV1    GCG.TAATGG CGCCGGAGGG GGG.AACAAG GTGGTGGACG AGTGCTACAT
  AAV2    ACCAGAAATG GCGCCGGAGG CGGGAACAAG GTGGTGGATG AGTGCTACAT
  AAV3    ACGCGAAATG GCGCCGGGGG CGGGAACAAG GTGGTGGACG ACTGCTACAT
  AAV8    GCGGTAATGG CGCCGGCGGG GGGGAACAAG GTGGTGGACG AGTGCTACAT
  AAV9    GCG.TAATGG CGCCGGCGGG GGG.AACAAG GTGGTGGACG AGTGCTACAT
  AAV7    ACGCGTAATG GCGCCGGCGG GGGGAACAAG GTGGTGGACG AGTGCTACAT
  44_2    .......... .......... .......... .......... ..........
```

图1P

```
          801                                                       850
  42_2    .......... .......... .......... .......... ..........
  42_8    .......... .......... .......... .......... ..........
 42_15    .......... .......... .......... .......... ..........
 42_5b    .......... .......... .......... .......... ..........
 42_1b    .......... .......... .......... .......... ..........
 42_13    .......... .......... .......... .......... ..........
 42_3a    .......... .......... .......... .......... ..........
  42_4    .......... .......... .......... .......... ..........
 42_5a    .......... .......... .......... .......... ..........
 42_10    .......... .......... .......... .......... ..........
 42_3b    .......... .......... .......... .......... ..........
 42_11    .......... .......... .......... .......... ..........
 42_6b    .......... .......... .......... .......... ..........
  43_1    .......... .......... .......... .......... ..........
  43_5    .......... .......... .......... .......... ..........
 43_12    .......... .......... .......... .......... ..........
 43_20    .......... .......... .......... .......... ..........
 43_21    .......... .......... .......... .......... ..........
 43_23    .......... .......... .......... .......... ..........
 43_25    .......... .......... .......... .......... ..........
  44_1    .......... .......... .......... .......... ..........
  44_5    .......... .......... .......... .......... ..........
223_10    .......... .......... .......... .......... ..........
 223_2    .......... .......... .......... .......... ..........
 223_4    .......... .......... .......... .......... ..........
 223_5    .......... .......... .......... .......... ..........
 223_6    .......... .......... .......... .......... ..........
 223_7    .......... .......... .......... .......... ..........
  A3_4    .......... .......... .......... .......... ..........
  A3_5    .......... .......... .......... .......... ..........
  A3_7    .......... .......... .......... .......... ..........
  A3_3    .......... .......... .......... .......... ..........
 42_12    .......... .......... .......... .......... ..........
  AAV1    CCCCAACTAC CTCCTGCCCA AGACTCAGCC CGAGCTGCAG TGGGCGTGGA
  AAV2    CCCCAATTAC TTGCTCCCCA AAACCCAGCC TGAGCTCCAG TGGGCGTGGA
  AAV3    CCCCAACTAC CTGCTCCCCA AGACCCAGCC CGAGCTCCAG TGGGCGTGGA
  AAV8    CCCCAACTAC CTCCTGCCCA AGACTCAGCC CGAGCTGCAG TGGGCGTGGA
  AAV9    CCCCAACTAC CTCCTGCCCA AGACTCAGCC CGAGCTGCAG TGGGCGTGGA
  AAV7    CCCCAACTAC CTCCTGCCCA AGACCCAGCC CGAGCTGCAG TGGGCGTGGA
  44_2    .......... .......... .......... .......... ..........
```

图1Q

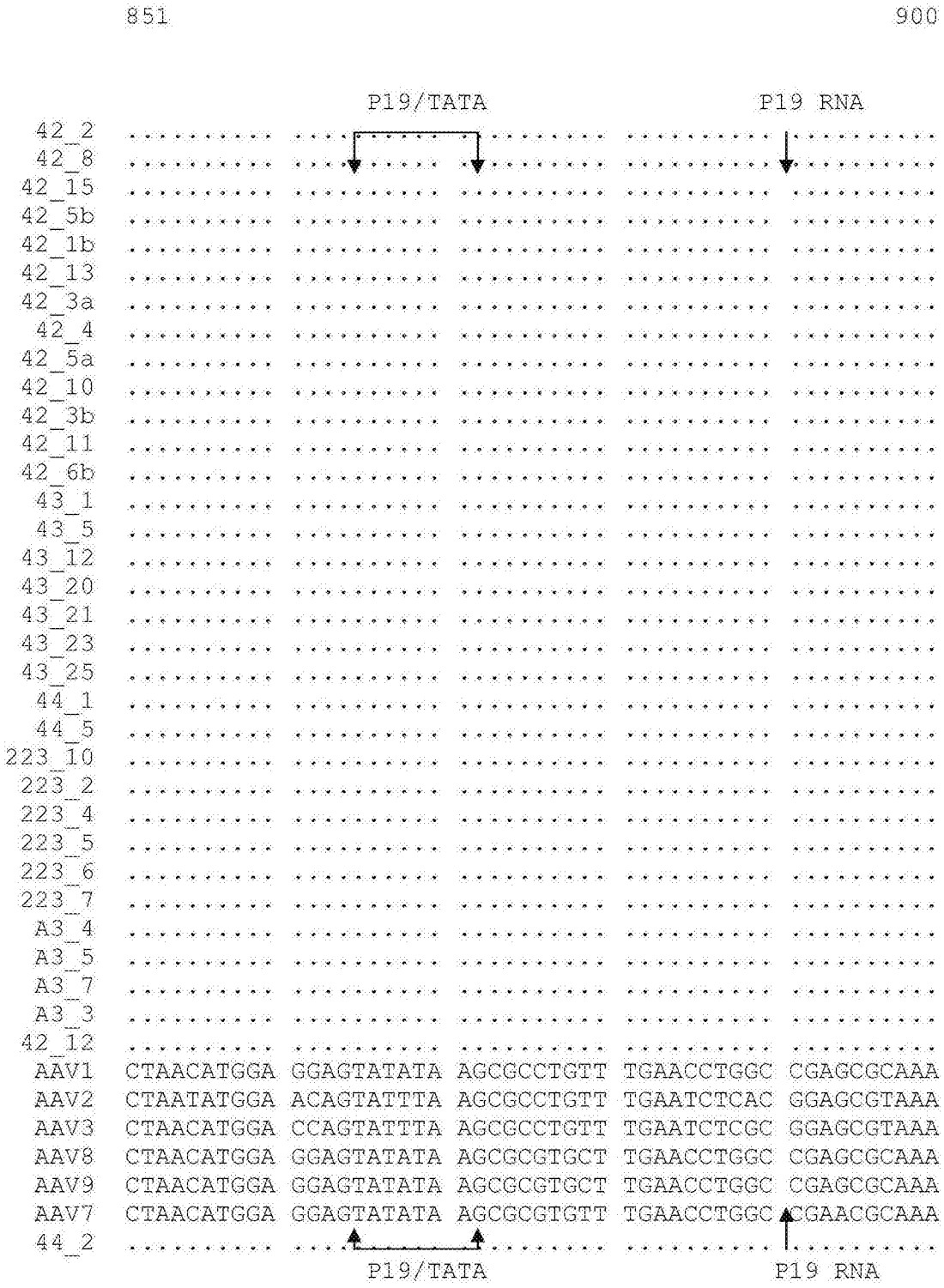

851
900
P19/TATA
P19 RNA
42_2
42_8
42_15
42_5b
42_1b
42_13
42_3a
42_4
42_5a
42_10
42_3b
42_11
42_6b
43_1
43_5
43_12
43_20
43_21
43_23
43_25
44_1
44_5
223_10
223_2
223_4
223_5
223_6
223_7
A3_4
A3_5
A3_7
A3_3
42_12
AAV1 CTAACATGGA GGAGTATATA AGCGCCTGTT TGAACCTGGC CGAGCGCAAA
AAV2 CTAATATGGA ACAGTATTTA AGCGCCTGTT TGAATCTCAC GGAGCGTAAA
AAV3 CTAACATGGA CCAGTATTTA AGCGCCTGTT TGAATCTCGC GGAGCGTAAA
AAV8 CTAACATGGA GGAGTATATA AGCGCGTGCT TGAACCTGGC CGAGCGCAAA
AAV9 CTAACATGGA GGAGTATATA AGCGCGTGCT TGAACCTGGC CGAGCGCAAA
AAV7 CTAACATGGA GGAGTATATA AGCGCGTGTT TGAACCTGGC CGAACGCAAA
44_2
P19/TATA
P19 RNA

图1R

```
          901                                                        950
  42_2    .......... .......... .......... .......... ..........
  42_8    .......... .......... .......... .......... ..........
 42_15    .......... .......... .......... .......... ..........
 42_5b    .......... .......... .......... .......... ..........
 42_1b    .......... .......... .......... .......... ..........
 42_13    .......... .......... .......... .......... ..........
 42_3a    .......... .......... .......... .......... ..........
  42_4    .......... .......... .......... .......... ..........
 42_5a    .......... .......... .......... .......... ..........
 42_10    .......... .......... .......... .......... ..........
 42_3b    .......... .......... .......... .......... ..........
 42_11    .......... .......... .......... .......... ..........
 42_6b    .......... .......... .......... .......... ..........
  43_1    .......... .......... .......... .......... ..........
  43_5    .......... .......... .......... .......... ..........
 43_12    .......... .......... .......... .......... ..........
 43_20    .......... .......... .......... .......... ..........
 43_21    .......... .......... .......... .......... ..........
 43_23    .......... .......... .......... .......... ..........
 43_25    .......... .......... .......... .......... ..........
  44_1    .......... .......... .......... .......... ..........
  44_5    .......... .......... .......... .......... ..........
223_10    .......... .......... .......... .......... ..........
 223_2    .......... .......... .......... .......... ..........
 223_4    .......... .......... .......... .......... ..........
 223_5    .......... .......... .......... .......... ..........
 223_6    .......... .......... .......... .......... ..........
 223_7    .......... .......... .......... .......... ..........
  A3_4    .......... .......... .......... .......... ..........
  A3_5    .......... .......... .......... .......... ..........
  A3_7    .......... .......... .......... .......... ..........
  A3_3    .......... .......... .......... .......... ..........
 42_12    .......... .......... .......... .......... ..........
  AAV1    CGGCTCGTGG CGCAGCACCT GACCCACGTC AGCCAGACCC AGGAGCAGAA
  AAV2    CGGTTGGTGG CGCAGCATCT GACGCACGTG TCGCAGACGC AGGAGCAGAA
  AAV3    CGGCTGGTGG CGCAGCATCT GACGCACGTG TCGCAGACGC AGGAGCAGAA
  AAV8    CGGCTCGTGG CGCAGCACCT GACCCACGTC AGCCAGACGC AGGAGCAGAA
  AAV9    CGGCTCGTGG CGCAGCACCT GACCCACGTC AGCCAGACGC AGGAGCAGAA
  AAV7    CGGCTCGTGG CGCAGCACCT GACCCACGTC AGCCAGACGC AGGAGCAGAA
  44_2    .......... .......... .......... .......... ..........
```

图1S

```
          951                                                        1000
  42_2    .......... .......... .......... .......... ..........
  42_8    .......... .......... .......... .......... ..........
 42_15    .......... .......... .......... .......... ..........
 42_5b    .......... .......... .......... .......... ..........
 42_1b    .......... .......... .......... .......... ..........
 42_13    .......... .......... .......... .......... ..........
 42_3a    .......... .......... .......... .......... ..........
  42_4    .......... .......... .......... .......... ..........
 42_5a    .......... .......... .......... .......... ..........
 42_10    .......... .......... .......... .......... ..........
 42_3b    .......... .......... .......... .......... ..........
 42_11    .......... .......... .......... .......... ..........
 42_6b    .......... .......... .......... .......... ..........
  43_1    .......... .......... .......... .......... ..........
  43_5    .......... .......... .......... .......... ..........
 43_12    .......... .......... .......... .......... ..........
 43_20    .......... .......... .......... .......... ..........
 43_21    .......... .......... .......... .......... ..........
 43_23    .......... .......... .......... .......... ..........
 43_25    .......... .......... .......... .......... ..........
  44_1    .......... .......... .......... .......... ..........
  44_5    .......... .......... .......... .......... ..........
223_10    .......... .......... .......... .......... ..........
 223_2    .......... .......... .......... .......... ..........
 223_4    .......... .......... .......... .......... ..........
 223_5    .......... .......... .......... .......... ..........
 223_6    .......... .......... .......... .......... ..........
 223_7    .......... .......... .......... .......... ..........
  A3_4    .......... .......... .......... .......... ..........
  A3_5    .......... .......... .......... .......... ..........
  A3_7    .......... .......... .......... .......... ..........
  A3_3    .......... .......... .......... .......... ..........
 42_12    .......... .......... .......... .......... ..........
  AAV1    CAAGGAGAAT CTGAACCCCA ATTCTGACGC GCCTGTCATC CGGTCAAAAA
  AAV2    CAAAGAGAAT CAGAATCCCA ATTCTGATGC GCCGGTGATC AGATCAAAAA
  AAV3    CAAAGAGAAT CAGAACCCCA ATTCTGACGC GCCGGTCATC AGGTCAAAAA
  AAV8    CAAGGAGAAT CTGAACCCCA ATTCTGACGC GCCCGTGATC AGGTCAAAAA
  AAV9    CAAGGAGAAT CTGAACCCCA ATTCTGACGC GCCCGTGATC AGGTCAAAAA
  AAV7    CAAGGAGAAT CTGAACCCCA ATTCTGACGC GCCCGTGATC AGGTCAAAAA
  44_2    .......... .......... .......... .......... ..........
```

图1T

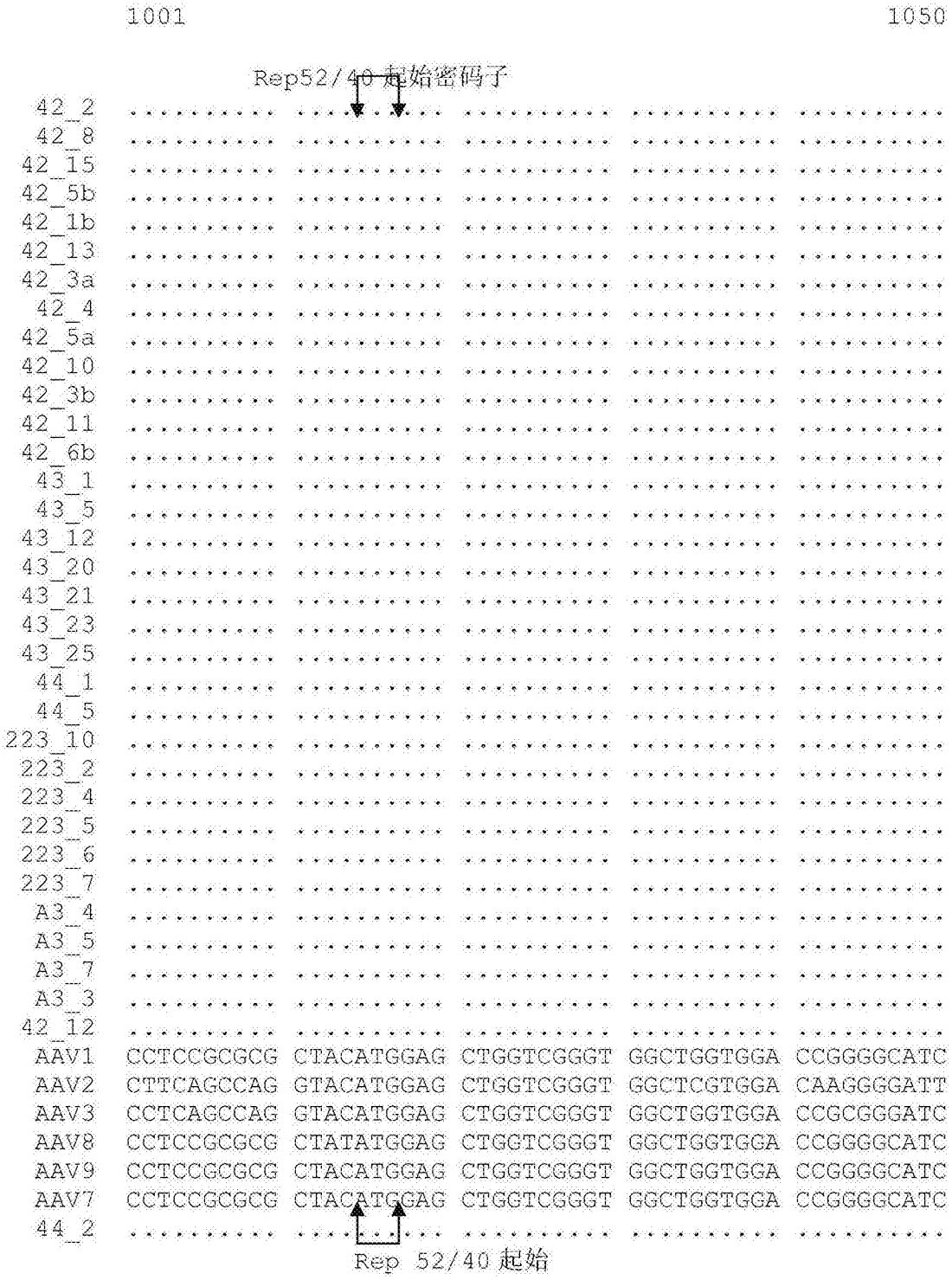

1001
1050
Rep52/40 起始密码子
42_2 ....
42_8 ....
42_15 ....
42_5b ....
42_1b ....
42_13 ....
42_3a ....
42_4 ....
42_5a ....
42_10 ....
42_3b ....
42_11 ....
42_6b ....
43_1 ....
43_5 ....
43_12 ....
43_20 ....
43_21 ....
43_23 ....
43_25 ....
44_1 ....
44_5 ....
223_10 ....
223_2 ....
223_4 ....
223_5 ....
223_6 ....
223_7 ....
A3_4 ....
A3_5 ....
A3_7 ....
A3_3 ....
42_12 ....
AAV1 CCTCCGCGCG CTACATGGAG CTGGTCGGGT GGCTGGTGGA CCGGGGCATC
AAV2 CTTCAGCCAG GTACATGGAG CTGGTCGGGT GGCTCGTGGA CAAGGGGATT
AAV3 CCTCAGCCAG GTACATGGAG CTGGTCGGGT GGCTGGTGGA CCGCGGGATC
AAV8 CCTCCGCGCG CTATATGGAG CTGGTCGGGT GGCTGGTGGA CCGGGGCATC
AAV9 CCTCCGCGCG CTACATGGAG CTGGTCGGGT GGCTGGTGGA CCGGGGCATC
AAV7 CCTCCGCGCG CTACATGGAG CTGGTCGGGT GGCTGGTGGA CCGGGGCATC
44_2 ....
Rep 52/40 起始

图1U

```
          1051                                                      1100
  42_2    .......... .......... .......... .......... ..........
  42_8    .......... .......... .......... .......... ..........
 42_15    .......... .......... .......... .......... ..........
 42_5b    .......... .......... .......... .......... ..........
 42_1b    .......... .......... .......... .......... ..........
 42_13    .......... .......... .......... .......... ..........
 42_3a    .......... .......... .......... .......... ..........
  42_4    .......... .......... .......... .......... ..........
 42_5a    .......... .......... .......... .......... ..........
 42_10    .......... .......... .......... .......... ..........
 42_3b    .......... .......... .......... .......... ..........
 42_11    .......... .......... .......... .......... ..........
 42_6b    .......... .......... .......... .......... ..........
  43_1    .......... .......... .......... .......... ..........
  43_5    .......... .......... .......... .......... ..........
 43_12    .......... .......... .......... .......... ..........
 43_20    .......... .......... .......... .......... ..........
 43_21    .......... .......... .......... .......... ..........
 43_23    .......... .......... .......... .......... ..........
 43_25    .......... .......... .......... .......... ..........
  44_1    .......... .......... .......... .......... ..........
  44_5    .......... .......... .......... .......... ..........
223_10    .......... .......... .......... .......... ..........
 223_2    .......... .......... .......... .......... ..........
 223_4    .......... .......... .......... .......... ..........
 223_5    .......... .......... .......... .......... ..........
 223_6    .......... .......... .......... .......... ..........
 223_7    .......... .......... .......... .......... ..........
  A3_4    .......... .......... .......... .......... ..........
  A3_5    .......... .......... .......... .......... ..........
  A3_7    .......... .......... .......... .......... ..........
  A3_3    .......... .......... .......... .......... ..........
 42_12    .......... .......... .......... .......... ..........
  AAV1    ACCTCCGAGA AGCAGTGGAT CCAGGAGGAC CAGGCCTCGT ACATCTCCTT
  AAV2    ACCTCGGAGA AGCAGTGGAT CCAGGAGGAC CAGGCCTCAT ACATCTCCTT
  AAV3    ACGTCAGAAA AGCAATGGAT TCAGGAGGAC CAGGCCTCGT ACATCTCCTT
  AAV8    ACCTCCGAGA AGCAGTGGAT CCAGGAGGAC CAGGCCTCGT ACATCTCCTT
  AAV9    ACCTCCGAGA AGCAGTGGAT CCAGGAGGAC CAGGCCTCGT ACATCTCCTT
  AAV7    ACCTCCGAGA AGCAGTGGAT CCAGGAGGAC CAGGCCTCGT ACATCTCCTT
  44_2    .......... .......... .......... .......... ..........
```

图1V

```
          1101                                                      1150
  42_2    .......... .......... .......... .......... ..........
  42_8    .......... .......... .......... .......... ..........
 42_15    .......... .......... .......... .......... ..........
 42_5b    .......... .......... .......... .......... ..........
 42_1b    .......... .......... .......... .......... ..........
 42_13    .......... .......... .......... .......... ..........
 42_3a    .......... .......... .......... .......... ..........
  42_4    .......... .......... .......... .......... ..........
 42_5a    .......... .......... .......... .......... ..........
 42_10    .......... .......... .......... .......... ..........
 42_3b    .......... .......... .......... .......... ..........
 42_11    .......... .......... .......... .......... ..........
 42_6b    .......... .......... .......... .......... ..........
  43_1    .......... .......... .......... .......... ..........
  43_5    .......... .......... .......... .......... ..........
 43_12    .......... .......... .......... .......... ..........
 43_20    .......... .......... .......... .......... ..........
 43_21    .......... .......... .......... .......... ..........
 43_23    .......... .......... .......... .......... ..........
 43_25    .......... .......... .......... .......... ..........
  44_1    .......... .......... .......... .......... ..........
  44_5    .......... .......... .......... .......... ..........
223_10    .......... .......... .......... .......... ..........
 223_2    .......... .......... .......... .......... ..........
 223_4    .......... .......... .......... .......... ..........
 223_5    .......... .......... .......... .......... ..........
 223_6    .......... .......... .......... .......... ..........
 223_7    .......... .......... .......... .......... ..........
  A3_4    .......... .......... .......... .......... ..........
  A3_5    .......... .......... .......... .......... ..........
  A3_7    .......... .......... .......... .......... ..........
  A3_3    .......... .......... .......... .......... ..........
 42_12    .......... .......... .......... .......... ..........
  AAV1    CAACGCCGCT TCCAACTCGC GGTCCCAGAT CAAGGCCGCT CTGGACAATG
  AAV2    CAATGCGGCC TCCAACTCGC GGTCCCAAAT CAAGGCTGCC TTGGACAATG
  AAV3    CAACGCCGCC TCCAACTCGC GGTCCCAGAT CAAGGCCGCG CTGGACAATG
  AAV8    CAACGCCGCC TCCAACTCGC GGTCCCAGAT CAAGGCCGCG CTGGACAATG
  AAV9    CAACGCCGCC TCCAACTCGC GGTCCCAGAT CAAGGCCGCG CTGGACAATG
  AAV7    CAACGCCGCC TCCAACTCGC GGTCCCAGAT CAAGGCCGCG CTGGACAATG
  44_2    .......... .......... .......... .......... ..........
```

图1W

```
          1151                                                      1200
  42_2    .......... .......... .......... .......... ..........
  42_8    .......... .......... .......... .......... ..........
 42_15    .......... .......... .......... .......... ..........
 42_5b    .......... .......... .......... .......... ..........
 42_1b    .......... .......... .......... .......... ..........
 42_13    .......... .......... .......... .......... ..........
 42_3a    .......... .......... .......... .......... ..........
  42_4    .......... .......... .......... .......... ..........
 42_5a    .......... .......... .......... .......... ..........
 42_10    .......... .......... .......... .......... ..........
 42_3b    .......... .......... .......... .......... ..........
 42_11    .......... .......... .......... .......... ..........
 42_6b    .......... .......... .......... .......... ..........
  43_1    .......... .......... .......... .......... ..........
  43_5    .......... .......... .......... .......... ..........
 43_12    .......... .......... .......... .......... ..........
 43_20    .......... .......... .......... .......... ..........
 43_21    .......... .......... .......... .......... ..........
 43_23    .......... .......... .......... .......... ..........
 43_25    .......... .......... .......... .......... ..........
  44_1    .......... .......... .......... .......... ..........
  44_5    .......... .......... .......... .......... ..........
223_10    .......... .......... .......... .......... ..........
 223_2    .......... .......... .......... .......... ..........
 223_4    .......... .......... .......... .......... ..........
 223_5    .......... .......... .......... .......... ..........
 223_6    .......... .......... .......... .......... ..........
 223_7    .......... .......... .......... .......... ..........
  A3_4    .......... .......... .......... .......... ..........
  A3_5    .......... .......... .......... .......... ..........
  A3_7    .......... .......... .......... .......... ..........
  A3_3    .......... .......... .......... .......... ..........
 42_12    .......... .......... .......... .......... ..........
  AAV1    CCGGCAAGAT CATGGCGCTG ACCAAATCCG CGCCCGACTA CCTGGTAGGC
  AAV2    CGGGAAAGAT TATGAGCCTG ACTAAAACCG CCCCCGACTA CCTGGTGGGC
  AAV3    CCTCCAAGAT CATGAGCCTG ACAAAGACGG CTCCGGACTA CCTGGTGGGC
  AAV8    CCGGCAAGAT CATGGCGCTG ACCAAATCCG CGCCCGACTA CCTGGTGGGG
  AAV9    CCGGCAAGAT CATGGCGCTG ACCAAATCCG CGCCCGACTA CCTGGTAGGC
  AAV7    CCGGCAAGAT CATGGCGCTG ACCAAATCCG CGCCCGACTA CCTGGTGGGG
  44_2    .......... .......... .......... .......... ..........
```

图1X

```
          1201                                                      1250
  42_2    .......... .......... .......... .......... ..........
  42_8    .......... .......... .......... .......... ..........
 42_15    .......... .......... .......... .......... ..........
 42_5b    .......... .......... .......... .......... ..........
 42_1b    .......... .......... .......... .......... ..........
 42_13    .......... .......... .......... .......... ..........
 42_3a    .......... .......... .......... .......... ..........
  42_4    .......... .......... .......... .......... ..........
 42_5a    .......... .......... .......... .......... ..........
 42_10    .......... .......... .......... .......... ..........
 42_3b    .......... .......... .......... .......... ..........
 42_11    .......... .......... .......... .......... ..........
 42_6b    .......... .......... .........GA ATTCGCCCTT TCTACGGCTG
  43_1    .......... .......... .......... .......... ..........
  43_5    .......... .......... .......... .......... ..........
 43_12    .......... .......... .......... .......... ..........
 43_20    .......... .......... .......... .......... ..........
 43_21    .......... .......... .......... .......... ..........
 43_23    .......... .......... .......... .......... ..........
 43_25    .......... .......... .......... .......... ..........
  44_1    .......... .......... .......... .......... ..........
  44_5    .......... .......... .......... .......... ..........
223_10    .......... .......... .......... .......... ..........
 223_2    .......... .......... .......... .......... ..........
 223_4    .......... .......... .......... .......... ..........
 223_5    .......... .......... .......... .......... ..........
 223_6    .......... .......... .......... .......... ..........
 223_7    .......... .......... .......... .......... ..........
  A3_4    .......... .......... .......... .......... ..........
  A3_5    .......... .......... .......... .......... ..........
  A3_7    .......... .......... .......... .......... ..........
  A3_3    .......... .......... .......... .......... ..........
 42_12    .......... .......... .......... .......... ..........
  AAV1    CCCGCTCCGC CCGCGGACAT TAAAACCAAC CGCATCTACC GCATCCTGGA
  AAV2    CAGCAGCCCG TGGAGGACAT TTCCAGCAAT CGGATTTATA AAATTTTGGA
  AAV3    AGCAACCCGC CGGAGGACAT TACCAAAAAT CGGATCTACC AAATCCTGGA
  AAV8    CCCTCGCTGC CCGCGGACAT TACCCAGAAC CGCATCTACC GCATCCTCGC
  AAV9    CCTTCACTTC CGGTGGACAT TACGCAGAAC CGCATCTACC GCATCCTGCA
  AAV7    CCCTCGCTGC CCGCGGACAT TAAAACCAAC CGCATCTACC GCATCCTGGA
  44_2    .......... .......... .......... .......... ..........
```

图1Y

```
          1251                                                      1300
  42_2    .......... .......... .......... .......... ..........
  42_8    .......... .......... .......... .......... ..........
 42_15    .......... .......... .......... .......... ..........
 42_5b    .......... .......... .......... .......... ..........
 42_1b    .......... .......... .......... .......... ..........
 42_13    .......... .......... .......... .......... ..........
 42_3a    .......... .......... .......... .......... ..........
  42_4    .......... .......... .......... .......... ..........
 42_5a    .......... .......... .......... .......... ..........
 42_10    .......... .......... .......... .......... ..........
 42_3b    .......... .......... .......... .......... ..........
 42_11    .......... .......... .......... .......... ..........
 42_6b    CGTCAACTGG ACCAATGAGA ACTTTCCCTT CAACGATTGC GTCGACAAGA
  43_1    .......... .......... .......... .......... ..........
  43_5    .......... .......... .......... .......... ..........
 43_12    .......... .......... .......... .......... ..........
 43_20    .......... .......... .......... .......... ..........
 43_21    .......... .......... .......... .......... ..........
 43_23    .......... .......... .......... .......... ..........
 43_25    .......... .......... .......... .......... ..........
  44_1    .......... .......... .......... .......... ..........
  44_5    .......... .......... .......... .......... ..........
223_10    .......... .......... .......... .......... ..........
 223_2    .......... .......... .......... .......... ..........
 223_4    .......... .......... .......... .......... ..........
 223_5    .......... .......... .......... .......... ..........
 223_6    .......... .......... .......... .......... ..........
 223_7    .......... .......... .......... .......... ..........
  A3_4    .......... .......... .......... .......... ..........
  A3_5    .......... .......... .......... .......... ..........
  A3_7    .......... .......... .......... .......... ..........
  A3_3    .......... .......... .......... .......... ..........
 42_12    .......... .......... .......... .......... ..........
  AAV1    GCTGAACGGC TACGAACCTG CCTACGCCGG CTCCGTCTTT CTCGGCTGGG
  AAV2    ACTAAACGGG TACGATCCCC AATATGCGGC TTCCGTCTTT CTGGGATGGG
  AAV3    GCTGAACGGG TACGATCCGC AGTACGCGGC CTCCGTCTTC CTGGGCTGGG
  AAV8    TCTCAACGGC TACGACCCTG CCTACGCCGG CTCCGTCTTT CTCGGCTGGG
  AAV9    GCTCAACGGC TACGACCCTG CCTACGCCGG CTCCGTCTTT CTCGGCTGGG
  AAV7    GCTGAACGGG TACGATCCTG CCTACGCCGG CTCCGTCTTT CTCGGCTGGG
  44_2    .......... .......... .......... .......... ..........
```

图1Z

```
          1301                                                      1350
  42_2    .......... .......... .......... .......... ..........
  42_8    .......... .......... .......... .......... ..........
 42_15    .......... .......... .......... .......... ..........
 42_5b    .......... .......... .......... .......... ..........
 42_1b    .......... .......... .......... .......... ..........
 42_13    .......... .......... .......... .......... ..........
 42_3a    .......... .......... .......... .......... ..........
  42_4    .......... .......... .......... .......... ..........
 42_5a    .......... .......... .......... .......... ..........
 42_10    .......... .......... .......... .......... ..........
 42_3b    .......... .......... .......... .......... ..........
 42_11    .......... .......... .......... .......... ..........
 42_6b    TGGTGATCTG GTGGGAGGAG GGCAAGATGA CGGCCAAGGT CGTGGAGTCC
  43_1    .......... .......... .......... .......... ..........
  43_5    .......... .......... .......... .......... ..........
 43_12    .......... .......... .......... .......... ..........
 43_20    .......... .......... .......... .......... ..........
 43_21    .......... .......... .......... .......... ..........
 43_23    .......... .......... .......... .......... ..........
 43_25    .......... .......... .......... .......... ..........
  44_1    .......... .......... .......... .......... ..........
  44_5    .......... .......... .......... .......... ..........
223_10    .......... .......... .......... .......... ..........
 223_2    .......... .......... .......... .......... ..........
 223_4    .......... .......... .......... .......... ..........
 223_5    .......... .......... .......... .......... ..........
 223_6    .......... .......... .......... .......... ..........
 223_7    .......... .......... .......... .......... ..........
  A3_4    .......... .......... .......... .......... ..........
  A3_5    .......... .......... .......... .......... ..........
  A3_7    .......... .......... .......... .......... ..........
  A3_3    .......... .......... .......... .......... ..........
 42_12    .......... .......... .......... .......... ..........
  AAV1    CCCAGAAAAG GTTCGGGAAG CGCAACACCA TCTGGCTGTT TGGGCCGGCC
  AAV2    CCACGAAAAA GTTCGGCAAG AGGAACACCA TCTGGCTGTT TGGGCCTGCA
  AAV3    CGCAAAAGAA GTTCGGGAAG AGGAACACCA TCTGGCTCTT TGGGCCGGCC
  AAV8    CTCAGAAAAA GTTCGGGAAA CGCAACACCA TCTGGCTGTT TGGACCCGCC
  AAV9    CACAAAAGAA GTTCGGGAAA CGCAACACCA TCTGGCTGTT TGGGCCGGCC
  AAV7    CCCAGAAAAA GTTCGGGAAG CGCAACACCA TCTGGCTGTT TGGGCCCGCC
  44_2    .......... .......... .......... .......... ..........
```

图1AA

```
          1351                                                    1400

  42_2    .......... .......... .......... .......GAA TTCGCCCTTT
  42_8    .......... .......... .......... .......GAA TTCGCCCTTT
 42_15    .......... .......... .......... .......GAA TTCGCCCTTT
 42_5b    .......... .......... .......... .......GAA TTCGCCCTTT
 42_1b    .......... .......... .......... .......... ..........
 42_13    .......... .......... .......... .......GAA TTCGCCCTTT
 42_3a    .......... .......... .......... .......GAA TTCGCCCTTT
  42_4    .......... .......... .......... .......... ..........
 42_5a    .......... .......... .......... ........GA ATTCGCCCTT
 42_10    .......... .......... .......... .......... ..........
 42_3b    .......... .......... .......... .......... ..........
 42_11    .......... .......... .......... .......GAA TTCGCCCTTT
 42_6b    GCCAAGGCCA TTCTCGGCGG CAGCAAGGTG CGCGTGGACC AAAAGTGCAA
  43_1    .......... .......... .......... .......GAA TTCGCCCTTT
  43_5    .......... .......... .......... .......GAA TTCGCCCTTT
 43_12    .......... .......... .......... .......GAA TTCGCCCTT.
 43_20    .......... .......... .......... .......GAA TTCGCCCTTT
 43_21    .......... .......... .......... .......GAA TTCGCCCTT.
 43_23    .......... .......... .......... .......GAA TTCGCCCTT.
 43_25    .......... .......... .......... .......GAA TTCGCCCTTT
  44_1    .......... .......... .......... .......GAA TTCGCCCTTT
  44_5    .......... .......... .......... .......GAA TTCGCCCTTT
223_10    .......... .......... .......... .......... ..........
 223_2    .......... .......... .......... .......... ..........
 223_4    .......... .......... .......... .......... ..........
 223_5    .......... .......... .......... .......... ..........
 223_6    .......... .......... .......... .......... ..........
 223_7    .......... .......... .......... .......... ..........
  A3_4    .......... .......... .......... ........GA ATTCGCCCTT
  A3_5    .......... .......... .......... ........GA ATTCGCCCTT
  A3_7    .......... .......... .........A GCGGCCGCGA ATTCGCCCTT
  A3_3    .......... .......... .......... ........GA ATTCGCCCTT
 42_12    .......... .......... .......... .......GAA TTCGCCCTTT
  AAV1    ACCACGGGCA AGACCAACAT CGCGGAAGCC ATCGCCCACG CCGTGCCCTT
  AAV2    ACTACCGGGA AGACCAACAT CGCGGAGGCC ATAGCCCACA CTGTGCCCTT
  AAV3    ACGACGGGTA AAACCAACAT CGCGGAAGCC ATCGCCCACG CCGTGCCCTT
  AAV8    ACCACCGGCA AGACCAACAT TGCGGAAGCC ATCGCCCACG CCGTGCCCTT
  AAV9    ACCACGGGAA AGACCAACAT CGCAGAAGCC ATTGCCCACG CCGTGCCCTT
  AAV7    ACCACCGGCA AGACCAACAT TGCGGAAGCC ATCGCCCACG CCGTGCCCTT
  44_2    .......... .......... .......... ........GA ATTCGCCCTT
```

图1AB

```
          1401                                                      1450
  42_2    .CTACGGCTG CGTCAACTGG ACCAATGAGA ACTTTCCCTT CAACGATTGC
  42_8    .CTACGGCTG CGTCAACTGG ACCAATGAGA ACTTTCCCTT CAACGATTGC
 42_15    .CTACGGCTG CGTCAACTGG ACCAATGAGA ACTTTCCCTT CAACGATTGC
 42_5b    .CTACGGCTG CGTCAACTGG ACCAATGAGA ACTTTCCCTT CAACGATTGC
 42_1b    .......... .......... .......... .......... ..........
 42_13    .CTACGGCTG CGTCAACTGG ACCAATGAGA ACTTTCCCTT CAACGATTGC
 42_3a    .CTACGGCTG CGTCAACTGG ACCAATGAGA ACTTTCCCTT CAACGATTGC
  42_4    .......... .......... .......... .......... ..........
 42_5a    .CTACGGCTG CGTCAACTGG ACCAATGAGA ACTTTCCCTT CAACGATTGC
 42_10    .......... .......... .......... .......... ..........
 42_3b    .......... .......... .......... .......... ..........
 42_11    .CTACGGCTG CGTCAACTGG ACCAATGAGA ACTTTCCCTT CAACGATTGC
 42_6b    .GTCTTCCGC CCAGATCGAT CCCACCCCCG TGATCGTCAC TTCCAACACC
  43_1    .CTACGGCTG CATCAACTGG ACCAATGAGA ACTTTCCCTT CAACGATTGC
  43_5    .CTACGGCTG CGTCAACTGG ACCAATGAGA ACTTTCCCTT CAACGATTGC
 43_12    .....GGCTG CGTCAACTGG ACCAATGAGA ACTTTCCCTT CAACGATTGC
 43_20    .CTACGGCTG CGTCAACTGG ACCAATGAGA ACTTTCCCTT CAACGATTGC
 43_21    .....GGCTG CGTCAACTGG ACCAATGAGA ACTTTCCCTT CAACGATTGC
 43_23    .CTACGGCTG CGTCAACTGG ACCAATGAGA ACTTTCCCTT CAACGATTGC
 43_25    .CTACGGCTG CGTCAACTGG ACCAATGAGA ACTTTCCCTT CAACGATTGC
  44_1    .CTACGGCTG CGTCAACTGG ACCAATGAGA ACTTTCCCTT CAACGATTGC
  44_5    .CTACGGCTG CGTCAACTGG ACCAATGAGA ACTTTCCCTT CAACGATTGC
223_10    .......... .......... .......... .......... ..........
 223_2    .......... .......... .......... .......... ..........
 223_4    .......... .......... .......... .......... ..........
 223_5    .......... .......... .......... .......... ..........
 223_6    .......... .......... .......... .......... ..........
 223_7    .......... .......... .......... .......... ..........
  A3_4    TCTACGGCTG CGTCAACTGG ACCAATGAAA ACTTTCCCTT CAACGATTGC
  A3_5    TCTACGGCTG CGTCAACTGG ACCAATGAAA ACTTTCCCTT CAACGATTGC
  A3_7    TCTACGGCTG CGTCAACTGG ACCAATGAAA ACTTTCCCTT CAACGATTGC
  A3_3    TCTACGGCTG CGTCAACTGG ACCAATGAAA ACTTTCCCTT CAACGATTGC
 42_12    .CTACGGCTG CGTCAACTGG ACCAATGAGA ACTTTCCCTT CAACGATTGC
  AAV1    .CTACGGCTG CGTCAACTGG ACCAATGAGA ACTTTCCCTT CAATGATTGC
  AAV2    .CTACGGGTG CGTAAACTGG ACCAATGAGA ACTTTCCCTT CAACGACTGT
  AAV3    .CTACGGCTG CGTAAACTGG ACCAATGAGA ACTTTCCCTT CAACGATTGC
  AAV8    .CTACGGCTG CGTCAACTGG ACCAATGAGA ACTTTCCCTT CAATGATTGC
  AAV9    .CTACGGCTG CGTCAACTGG ACCAATGAGA ACTTTCCCTT CAACGATTGC
  AAV7    .CTACGGCTG CGTCAACTGG ACCAATGAGA ACTTTCCCTT CAACGATTGC
  44_2    TCTACGGCTG CGTCAACTGG ACCAATGAGA ACTTTCCCTT CAACGATTGC
```

图1AC

```
          1451                                                      1500
  42_2    GTCGACAAGA TGGTGATCTG GTGGGAGGAG GGCAAGATGA CGGCCAAGGT
  42_8    GTCGACAAGA TGGTGATCTG GTGGGAGGAG GGCAAGATGA CGGCCAAGGT
 42_15    GTCGACAAGA TGGTGATCTG GTGGGAGGAG GGCAAGATGA CGGCCAAGGT
 42_5b    GTCGACAAGA TGGTGATCTG GTGGGAGGAG GGCAAGATGA CGGCCAAGGT
 42_1b    .......... .......... .......... .......... ..........
 42_13    GTCGACAAGA TGGTGATCTG GTGGGAGGAG GGCAAGATGA CGGCCAAGGT
 42_3a    GTCGACAAGA TGGTGATCTG GTGGGAGGAG GGCAAGATGA CGGCCAAGGT
  42_4    .......... .......... .......... .......... ..........
 42_5a    GTCGACAAGA TGGTGATCTG GTGGGAGGAG GGCAAGATGA CGGCCAAGGT
 42_10    .......... .......... .......... .......... ..........
 42_3b    .......... .......... .......... .......... ..........
 42_11    GTCGACAAGA TGGTGATCTG GTGGGAGGAG GGCAAGATGA CGGCCAAGGT
 42_6b    AACATGTGCG CCGTGATTGA CGGGAACAGC ACCACCTTCG AGCACCAGCA
  43_1    GTCGACAAGA TGGTGATCTG GTGGGAGGAG GGCAAGATGA CGGCCAAGGT
  43_5    GTCGACAAGA TGGTGATCTG GTGGGAGGAG GGCAAGATGA CGGCCAAGGT
 43_12    GTCGACAAGA TGGTGATCTG GTGGGAGGAG GGCAAGATGA CGGCCAAGGT
 43_20    GTCGACAAGA TGGTGATCTG GTGGGAGGAG GGCAAGATGA CGGCCAAGGT
 43_21    GTCGACAAGA TGGTGATCTG GTGGGAGGAG GGCAAGATGA CGGCCAAGGT
 43_23    GTCGACAAGA TGGTGATCTG GTGGGAGGAG GGCAAGATGA CGGCCAAGGT
 43_25    GTCGACAAGA TGGTGATCTG GTGGGAGGAG GGCAAGATGA CGGCCAAGGT
  44_1    GTCGACAAGA TGTTGATCTG GTGGGAGGAG GGCAAGATGA CGGCCAAGGT
  44_5    GTCGACAAGA TGGTGATCTG GTGGGAGGAG GGCAAGATGA CGGCCAAGGT
223_10    .......... .......... .......... .......... ..........
 223_2    .......... .......... .......... .......... ..........
 223_4    .......... .......... .......... .......... ..........
 223_5    .......... .......... .......... .......... ..........
 223_6    .......... .......... .......... .......... ..........
 223_7    .......... .......... .......... .......... ..........
  A3_4    GTCGACAAGA TGGTGATCTG GTGGGAGGAG GGAAAGATGA CCGCCAAGGT
  A3_5    GTCGACAAGA TGGTGATCTG GTGGGAGGAG GGAAAGATGA CCGCCAAGGT
  A3_7    GTCGACAAGA TGGTGATCTG GTGGGAGGAG GGAAAGATGA CCGCCAAGGT
  A3_3    GTCGACAAGA TGGTGATCTG GTGGGAGGAG GGAAAGATGA CCGCCAAGGT
 42_12    GTCGACAAGA TGGTGATCTG GTGGGAGGAG GGCAAGATGA CGGCCAAGGT
  AAV1    GTCGACAAGA TGGTGATCTG GTGGGAGGAG GGCAAGATGA CGGCCAAGGT
  AAV2    GTCGACAAGA TGGTGATCTG GTGGGAGGAG GGGAAGATGA CCGCCAAGGT
  AAV3    GTCGACAAGA TGGTGATCTG GTGGGAGGAG GGCAAGATGA CGGCCAAGGT
  AAV8    GTCGACAAGA TGGTGATCTG GTGGGAGGAG GGCAAGATGA CGGCCAAGGT
  AAV9    GTCGACAAGA TGGTGATCTG GTGGGAGGAG GGCAAGATGA CGGCCAAGGT
  AAV7    GTCGACAAGA TGGTGATCTG GTGGGAGGAG GGCAAGATGA CGGCCAAGGT
  44_2    GTCGACAAGA TGGTGATCTG GTGGGAGGAG GGCAAGATGA CGGCCAAGGT
```

图1AD

```
          1501                                                        1550
  42_2    CGTGGAGTCC GCCAAGGCCA TTCTCGGCGG CAGCAAGGTG CGCGTGGACC
  42_8    CGTGGAGTCC GCCAAGGCCA TTCTCGGCGG CAGCAAGGTG CGCGTGGACC
 42_15    CGTGGAGTCC GCCAAGGCCA TTCTCGGCGG CAGCAAGGTG CGCGTGGACC
 42_5b    CGTGGAGTCC GCCAAGGCCA TTCTCGGCGG CAGCAAGGTG CGCGTGGACC
 42_1b    .......... .......... .......... .......... ..........
 42_13    CGTGGAGTCC GCCAAGGCCA TTCTCGGCGG CAGCAAGGTG CGCGTGGACC
 42_3a    CGTGGAGTCC GCCAAGGCCA TTCTCGGCGG CAGCAAGGTG CGCGTGGACC
  42_4    .......... .......... .......... .......... ..........
 42_5a    CGTGGAGTCC GCCAAGGCCA TTCTCGGCGG CAGCAAGGTG CGCGTGGACC
 42_10    .......... .......... .......... .......... ..........
 42_3b    .......... .......... .......... .......... ..........
 42_11    CGTGGAGTCC GCCAAGGCCA TTCTCGGCGG CAGCAAGGTG CGCGTGGACC
 42_6b    GCCGTTGCAG GACCGGATGT TCAAATTTGA ACTCACCCGC CGTCTGGAGC
  43_1    CGTGGAGTCC GCCAAGGCCA TTCTCGGCGG CAGCAAGGTG CGCGTGGACC
  43_5    CGTGGAGTCC GCCAAGGCCA TTCTCGGCGG CAGCAAGGTG CGCGTGGACC
 43_12    CGTGGAGTCC GCCAAGGCCA TTCTCGGCGG CAGCAAGGTG CGCGTGGACC
 43_20    CGTGGAGTCC GCCAAGGCCA TTCTCGGCGG CAGCAAGGTG CGTGTGGACC
 43_21    CGTGGAGTCC GCCAAGGCCA TTCTCGGCGG CAGCAAGGTG CGTGTGGACC
 43_23    CGTGGAGTCC GCCAAGGCCA TTCTCGGCGG CAGCAAGGTG CGTGTGGACC
 43_25    CGTGGAGTCC GCCAAGGCCA TTCTCGGCGG CAGCAAGGTG CGTGTGGACC
  44_1    CGTGGAGTCC GCCAAGGCCA TTCTCGGCGG CAGCAAAGTG CGCGTGGACC
  44_5    CGTGGAGTCC GCCAAGGCCA TTCTCGGCGG CAGCAAAGTG CGCGTGGACC
223_10    .......... .......... .......... .......... ..........
 223_2    .......... .......... .......... .......... ..........
 223_4    .......... .......... .......... .......... ..........
 223_5    .......... .......... .......... .......... ..........
 223_6    .......... .......... .......... .......... ..........
 223_7    .......... .......... .......... .......... ..........
  A3_4    CGTGGAATCT GCCAAAGCCA TTCTGGGTGG AAGCAAGGTT CGTGTGGACC
  A3_5    CGTGGAATCT GCCAAAGCCA TTCTGGGTGG AAGCAAGGTT CGTGTGGACC
  A3_7    CGTGGAATCT GCCAAAGCCA TTCTGGGTGG AAGCAAGGTT CGTGTGGACC
  A3_3    CGTGGAATCT GCCAAAGCCA TTCTGGGTGG AGGCAAGGTT CGTGTGGACC
 42_12    CGTGGAGTCC GCCAAGGCCA TTCTCGGCGG CAGCAAGGTG CGCGTGGACC
  AAV1    CGTGGAGTCC GCCAAGGCCA TTCTCGGCGG CAGCAAGGTG CGCGTGGACC
  AAV2    CGTGGAGTCG GCCAAAGCCA TTCTCGGAGG AAGCAAGGTG CGCGTGGACC
  AAV3    CGTGGAGAGC GCCAAGGCCA TTCTGGGCGG AAGCAAGGTG CGCGTGGACC
  AAV8    CGTGGAGTCC GCCAAGGCCA TTCTCGGCGG CAGCAAGGTG CGCGTGGACC
  AAV9    CGTGGAGTCC GCCAAGGCCA TTCTCGGCGG CAGCAAGGTG CGCGTGGACC
  AAV7    CGTGGAGTCC GCCAAGGCCA TTCTCGGCGG CAGCAAGGTG CGCGTGGACC
  44_2    CGTGGAGTCC GCCAAGGCCA TTCTCGGCGG CAGCAAAGTG CGCGTGGACC
```

图1AE

```
          1551                                                      1600
  42_2    AAAAGTGCAA GTCTTCCGCC CAGATCGATC CCACCCCCGT GATCGTCACT
  42_8    AAAAGTGCAA GTCTTCCGCC CAGATCGATC CCACCCCCGT GATCGTCACT
 42_15    AAAAGTGCAA GTCGTCCGCC CAGATCGACC CCACCCCCGT GATCGTCACC
 42_5b    AAAAGTGCAA GTCGTCCGCC CAGATCGACC CCACCCCCGT GATCGTCACC
 42_1b    .......... .......... .......... .......... ..........
 42_13    AAAAGTGCAA GTCGTCCGCC CAGATCGATC CCACCCCCGT GATCGTCACT
 42_3a    AAAAGTGCAA GTCGTCCGCC CAGATCGATC CCACCCCCGT GATCGTCACT
  42_4    .......... .......... .......... .......... ..........
 42_5a    AAAAGTGCAA GTCGTCCGCC CAGATCGACC CCACCCCCGT GATCGTCACC
 42_10    .......... .......... .......... .......... ..........
 42_3b    .......... .......... .......... .......... ..........
 42_11    AAAAGTGCAA GTCTTCCGCC CAGATCGATC CCACCCCCGT GATCGTCACT
 42_6b    ATGACTTTGG CAAGGTGACA AAGCAGGAAG TCAAAGAGTT CTTCCGCTGG
  43_1    AAAAGTGCAA GTCGTCCGCC CAGATCGACC CCACCCCCGT GATCGTCACC
  43_5    AAAAGTGCAA GTCGTCCGCC CAGATCGACC CCACCCCCGT GATCGTCACC
 43_12    AAAAGTGCAA GTCGTCCGCC CAGATCGACC CCACCCCCGT GATCGTCACC
 43_20    AAAAGTGCAA GTCTTCCGCC CAGATCGATC CCACCCCCGT GATCGTCACC
 43_21    AAAAGTGCAA GTCTTCCGCC CAGATCGATC CCACCCCCGT GATCGTCACC
 43_23    AAAAGTGCAA GTCTTCCGCC CAGATCGATC CCACCCCCGT GATCGTCACC
 43_25    AAAAGTGCAA GTCTTCCGCC CAGATCGATC CCACCCCCGT GATCGTCACC
  44_1    AAAAGTGCAA GCCGTCCGCC CAGATCGACC CCACCCCCGT GATCGTCACC
  44_5    AAAAGTGCAA GTCGTCCGCC CAGATCGACC CCACCCCCGT GATCGTCACC
223_10    .......... .......... .......... .......... ..........
 223_2    .......... .......... .......... .......... ..........
 223_4    .......... .......... .......... .......... ..........
 223_5    .......... .......... .......... .......... ..........
 223_6    .......... .......... .......... .......... ..........
 223_7    .......... .......... .......... .......... ..........
  A3_4    AGAAATGCAA GTCTTCGGCC CAGATCGACC CGACTCCGGT GATTGTCACC
  A3_5    AGAAATGCAA GTCTTCGGCC CAGATCGACC CGACTCCGGT GATTGTCACC
  A3_7    AGAAATGCAG GTCTTCGGCC CAGATCGACC CGACTCCGGT GATTGTCACC
  A3_3    AGAAATGCAA GTCTTCGGCC CAGATCGACC CGACTCCGGT GATTGTCACC
 42_12    AAAAGTGCAA GTCGTCCGCC CAGATCGACC CCACCCCCGT GATCGTCACC
  AAV1    AAAAGTGCAA GTCGTCCGCC CAGATCGACC CCACCCCCGT GATCGTCACC
  AAV2    AGAAATGCAA GTCCTCGGCC CAGATAGACC CGACTCCCGT GATCGTCACC
  AAV3    AAAAGTGCAA GTCATCGGCC CAGATCGAAC CCACTCCCGT GATCGTCACC
  AAV8    AAAAGTGCAA GTCGTCCGCC CAGATCGACC CCACCCCCGT GATCGTCACC
  AAV9    AAAAGTGCAA GTCGTCCGCC CAGATCGACC CCACTCCCGT GATCGTCACC
  AAV7    AAAAGTGCAA GTCGTCCGCC CAGATCGACC CCACCCCCGT GATCGTCACC
  44_2    AAAAGTGCAA GTCGTCCGCC CAGATCGACC CCACCCCCGT GATCGTCACC
```

图1AF

```
          1601                                                      1650
  42_2    TCCAACACCA ACATGTGCGC TGTGATTGAC GGGAACAGCA CCACCTTCGA
  42_8    TCCAACACCA ACATGTGCGC CGTGATTGAC GGGAACAGCA CCACCTTCGA
 42_15    TCCAACACCA ACATGTGCGC CGTGATTGAC GGGAACAGCA CCACCTTCGA
 42_5b    TCCAACACCA ACATGTGCGC CGTGATTGAC GGGAACAGCA CCACCTTCGA
 42_1b    .......... .......... .......... .......... ..........
 42_13    TCCAACACCA ACATGTGCGC CGTGATTGAC GGGAACAGCA CCACCTTCGA
 42_3a    TCCAACACCA ACATGTGCGC CGTGATTGAC GGGAACAGCA CCACCTTCGA
  42_4    .......... .......... .......... .......... ..........
 42_5a    TCCAACACCA ACATGTGCGC CGTGATTGAC GGGAACAGCA CCACCTTCGA
 42_10    .......... .......... .......... .......... ..........
 42_3b    .......... .......... .......... .......... ..........
 42_11    TCCAACACCA ACATGTGCGC CGTGATTGAC GGGAACAGCA CCACCTTCGA
 42_6b    GCGCAGGATC ACGTGACCGA GGTGGCGCAT GAGTTCTACG TCAGAAAGGG
  43_1    TCCAACACCA ACATGTGCGC CGTGATTGAC GGGAACAGCA CCACCTTCGA
  43_5    TCCAACACCA ACATGTGCGC CGTGATTGAC GGGAACAGCA CCACCTTCGA
 43_12    TCCAACACCA ACATGTGCGC CGTGATTGAC GGGAACAGCA CCACCTTCGA
 43_20    TCCAACACCA ACATGTGCGC CGTGATTGAC GGGAACAGCG CCACCTTCGA
 43_21    TCCAACACCA ACATGTGCGC CGTGATTGAC GGGAACAGCA CCACCTTCGA
 43_23    TCCAACACCA ACATGTGCGC CGTGATTGAC GGGAACAGCA CCACCTTCGA
 43_25    TCCAACACCA ACATGTGCGC CGTGATTGAC GGGAACAGCA CCACCTTCGA
  44_1    TCCAACACCA ACATGTGCGC CGTGATTGAC GGGAACAGCA CCACCTTCGA
  44_5    TCCAACACCA ACATGTGCGC CGTGATTGAC GGGAACAGCA CCACCTTCGA
223_10    .......... .......... .......... .......... ..........
 223_2    .......... .......... .......... .......... ..........
 223_4    .......... .......... .......... .......... ..........
 223_5    .......... .......... .......... .......... ..........
 223_6    .......... .......... .......... .......... ..........
 223_7    .......... .......... .......... .......... ..........
  A3_4    TCTAACACCA ACATGTGCGC CGTGATTGAC GGAAACTCGA CCACCTTCGA
  A3_5    TCTAACACCA ACATGTGCGC CGTGATTGAC GGAAACTCGA CCACCTTCGA
  A3_7    TCTAACACCA ACATGTGCGC CGTGATTGAC GGAAACTCGA CCACCTTCGA
  A3_3    TCTAACACCA ACATGTGCGC CGTGATTGAC GGAAACTCGA CCACCTTCGA
 42_12    TCCAACACCA ACATGTGCGC CGTGATTGAC GGGAACAGCA CCACCTTCGA
  AAV1    TCCAACACCA ACATGTGCGC CGTGATTGAC GGGAACAGCA CCACCTTCGA
  AAV2    TCCAACACCA ACATGTGCGC CGTGATTGAC GGGAACTCAA CGACCTTCGA
  AAV3    TCCAACACCA ACATGTGCGC CGTGATTGAC GGGAACAGCA CCACCTTCGA
  AAV8    TCCAACACCA ACATGTGCGC CGTGATTGAC GGGAACAGCA CCACCTTCGA
  AAV9    TCCAACACCA ACATGTGCGC CGTGATTGAC GGGAACAGCA CCACCTTCGA
  AAV7    TCCAACACCA ACATGTGCGC CGTGATTGAC GGGAACAGCA CCACCTTCGA
  44_2    TCCAACACCA ACATGTGCGC CGTGATTGAC GGGAACAGCA CCACCTTCGA
```

图1AG

```
          1651                                                      1700
  42_2    GCACCAGCAG CCGTTACAAG ACCGGATGTT CAAATTTGAA CTCACCCGCC
  42_8    GCACCAGCAG CCGTTACAAG ACCGGATGTT CAAATTTGAA CTCACCCGCC
 42_15    GCACCAGCAG CCGTTGCAGG ACCGGATGTT CAAATTTGAA CTCACCCGCC
 42_5b    GCACCAGCAG CCGTTACAAG ACCGGATGTT CAAATTTGAA CTCACCCGCC
 42_1b    .......... .......... .......... .......... ..........
 42_13    GCACCAGCAG CCGTTACAAG ACCGGATGTT CAAATTTGAA CTCACCCGCC
 42_3a    GCACCAGCAG CCGTTACAAG ACCGGATGTT CAAATTTGAA CTCACCCGCC
  42_4    .......... .......... .......... .......... ..........
 42_5a    GCACCAGCAG CCGTTGCAGG ACCGGATGTT CAAATTTGAA CTCACCCGCC
 42_10    .......... .......... .......... .......... ..........
 42_3b    .......... .......... .......... .......... ..........
 42_11    GCACCAGCAG CCGTTACAAG ACCGGATGTT CAAATTTGAA CTCACCCGCC
 42_6b    TGGAGCCAAC AAGAGACCCG CCCCCGATGA CGCGGATAAA AGCGAGCCCA
  43_1    GCACCAGCAG CCGTTGCAGG ACCGGATGTT CAAGTTCGAA CTCACCCGCC
  43_5    GCACCAGCAG CCGTTGCAGG ACCGGATGTT CAAGTTCGAA CTCACCCGCC
 43_12    GCACCAGCAG CCGTTGCAGG ACCGGATGTT CAAGTTCGAA CTCACCCGCC
 43_20    GCACCAGCAG CCGTTGCAGG ACCGGATGTT CAAATTTGAA CTCACCCGCC
 43_21    GCACCAGCAG CCGTTGCAGG ACCGGATGTT CAAATTTGAA CTCACCCGCC
 43_23    GCACCAGCAG CCGTTGCAGG ACCGGATGTT CAAATTTGAA CTCACCCGCC
 43_25    GCACCAGCAG CCGTTGCAGG ACCGGATGTT CAAATTTGAA CTCACCCGCC
  44_1    GCACCAGCAG CCGTTGCGGG ACCGGATGTT CAAGTTTGAA CTCACCCGCC
  44_5    GCACCAGCAG CCGTTGCAGG ACCGGATGTT CAAGTTTGAA CTCACCCGCC
223_10    .......... .......... .......... .......... ..........
 223_2    .......... .......... .......... .......... ..........
 223_4    .......... .......... .......... .......... ..........
 223_5    .......... .......... .......... .......... ..........
 223_6    .......... .......... .......... .......... ..........
 223_7    .......... .......... .......... .......... ..........
  A3_4    GCACCAGCAG CCGTTGCAAG ACCGGATGTT CAAATTTGAA CTTACCCGCC
  A3_5    GCACCAGCAG CCGTTGCAAG ACCGGATGTT CAAATTTGAA CTTACCCGCC
  A3_7    GCACCAGCAG CCGTTGCAAG ACCGGATGTT CAAATTTGAA CTTACCCGCC
  A3_3    GCACCAGCAG CCGTTGCAAG ACCGGATGTT CAAATTTGAA CTTACCCGCC
 42_12    GCACCAGCAG CCGTTACAAG ACCGGATGTT CAAATTTGAA CTCACCCGCC
  AAV1    GCACCAGCAG CCGTTGCAGG ACCGGATGTT CAAATTTGAA CTCACCCGCC
  AAV2    ACACCAGCAG CCGTTGCAAG ACCGGATGTT CAAATTTGAA CTCACCCGCC
  AAV3    GCATCAGCAG CCGCTGCAGG ACCGGATGTT TGAATTTGAA CTTACCCGCC
  AAV8    GCACCAGCAG CCTCTCCAGG ACCGGATGTT TAAGTTCGAA CTCACCCGCC
  AAV9    GCACCAGCAG CCTCTCCAGG ACCGGATGTT TAAGTTCGAA CTCACCCGCC
  AAV7    GCACCAGCAG CCGTTGCAGG ACCGGATGTT CAAATTTGAA CTCACCCGCC
  44_2    GCACCAGCAG CCGTTGCAGG ACCGGATGTT CAAGTTTGAA CTCACCCGCC
```

图1AH

```
          1701                                                      1750
  42_2    GTCTGGAGCA CGACTTTGGC AAGGTGACAA AGCAGGAAGT CAAAGAGTTC
  42_8    GTCTGGAGCA CGACTTTGGC AAGGTGACAA AGCAGGAAGT CAAAGAGTTC
 42_15    GTCTGGAGCA TGACTTTGGC AAGGTGACAA AGCAGGAAGT CAAAGAGTTC
 42_5b    GTCTGGAGCA CGACTTTGGC AAGGTGACAA AGCAGGAAGT CAAAGAGTTC
 42_1b    .......... .......... .......... .......... ..........
 42_13    GTCTGGAGCA TGACTTTGGC AAGGTGACAA AGCAGGAAGT CAAAGAGTTC
 42_3a    GTCTGGAGCA TGACTTTGGC AAGGTGACAA AGCAGGAAGT CAAAGAGTTC
  42_4    .......... .......... .......... .......... ..........
 42_5a    GTCTGGAGCA TGACTTTGGC AAGGCGACAA AGCAGGAAGT CAAAGAGTTC
 42_10    .......... .......... .......... .......... ..........
 42_3b    .......... .......... .......... .......... ..........
 42_11    GTCTGGAGCA CGACTTTGGC AAGGTGACAA AGCAGGAAGT CAAAGAGTTC
 42_6b    AGCGGGCCTG CCCCTCAGTC GCGGATCCAT CGACGTCAGA CGCGGAAGGA
  43_1    GTCTGGAGCA CGACTTTGGC AAGGTGACCA AGCAGGAAGT CAAAGAGTTC
  43_5    GTCTGGAGCA CGACTTTGGC AAGGTGACCA AGCAGGAAGT CAAAGAGTTC
 43_12    GTCTGGAGCA CGACTTTGGC AAGGTGACCA AGCAGGAAGT CAAAGAGTTC
 43_20    GTCTGGAGCA TGACTTTGGC AAGGTGACGA AGCAGGAAGT CAAAGAGTTC
 43_21    GTCTGGAGCA TGACTTTGGC AAGGTGACGA AGCAGGAAGT CAAAGAGTTC
 43_23    GTCTGGAGCA TGACTTTGGC AAGGTGACGA AGCAGGAAGT CAAAGAGTTC
 43_25    GTCTGGAGCA TGACTTTGGC AAGGTGACGA AGCAGGAAGT CAAAGGGTTC
  44_1    GTCTGGAGCA CGACTTTGGC AAGGTGACAA AGCAGGAAGT CAGAGAGTTC
  44_5    GTCTGGAGCA CGACTTTGGC AAGGTGACAA AGCAGGAAGT CAGAGAGTTC
223_10    .......... .......... .......... .......... ..........
 223_2    .......... .......... .......... .......... ..........
 223_4    .......... .......... .......... .......... ..........
 223_5    .......... .......... .......... .......... ..........
 223_6    .......... .......... .......... .......... ..........
 223_7    .......... .......... .......... .......... ..........
  A3_4    GTTTGGATCA TGACTTTGGG AAGGTCACCA AGCAGGAAGT CAAAGACTTT
  A3_5    GTTTGGATCA TGACTTTGGG AAGGTCACCA AGCAGGAAGT CAAAGACTTT
  A3_7    GTTTGGATCA TGACTTTGGG AAGGTCACCA AGCAGGAAGT CAAAGACTTT
  A3_3    GTTTGGATCA TGACTTTGGG AAGGTCACCA AGCAGGAAGT CAAAGACTTT
 42_12    GTCTGGAGCA CGACTTTGGC AAGGTGACAA AGCAGGAAGT CAAAGAGTTC
  AAV1    GTCTGGAGCA TGACTTTGGC AAGGTGACAA AGCAGGAAGT CAAAGAGTTC
  AAV2    GTCTGGATCA TGACTTTGGG AAGGTCACCA AGCAGGAAGT CAAAGACTTT
  AAV3    GTTTGGACCA TGACTTTGGG AAGGTCACCA AACAGGAAGT AAAGGACTTT
  AAV8    GTCTGGAGCA CGACTTTGGC AAGGTGACAA AGCAGGAAGT CAAAGAGTTC
  AAV9    GTCTGGAGCA CGACTTTGGC AAGGTGACAA AGCAGGAAGT CAAAGAGTTC
  AAV7    GTCTGGAGCA CGACTTTGGC AAGGTGACGA AGCAGGAAGT CAAAGAGTTC
  44_2    GTCTGGAGCA CGACTTTGGC AAGGTGACAA AGCAGGAAGT CAGAGAGTTC
```

图1AI

```
          1751                                                      1800
  42_2    TTCCGCTGGG CGCAGGATCA CGTGACCGAG GTGGCGCATG AGTTCTACGT
  42_8    TTCCGCTGGG CGCAGGATCA CGTGACCGAG GTGGCGCATG AGTTCTACGT
 42_15    TTCCGCTGGG CGCAGGATCA CGTGACCGAG GTGGCGCATG AGTTCTACGT
 42_5b    TTCCGCTGGG CGCAGGATCA CGTGACCGAG GTGGCGCATG AGTTCTACGT
 42_1b    .......... .......... .......... .......... ..........
 42_13    TTCCGCTGGG CGCAGGATCA CGTGACCGAG GTGGCGCATG AGTTCTACGT
 42_3a    TTCCGCTGGG CGCAGGATCA CGTGACCGAG GTGGCGCATG AGTTCTACGT
  42_4    .......... .......... .......... .......... ..........
 42_5a    TTCCGCTGGG CGCAGGATCA CGTGACCGAG GTGGCGCATG AGTTCTACGT
 42_10    .......... .......... .......... .......... ..........
 42_3b    .......... .......... .......... .......... ..........
 42_11    TTCCGCTGGG CGCAGGATCA CGTGACCGAG GTGGCGCATG AGTTCTACGT
 42_6b    GCTCCGGTGG ACTTTGCCGA CAGGTACCAA AACAAATGTT CTCGTCACGC
  43_1    TTCCGCTGGG CGCAGGATCA CGTGACCGAG GTGGCGCATG AGTTCTACGT
  43_5    TTCCGCTGGG CGCAGGATCA CGTGACCGAG GTGGCGCATG AGTTCTACGT
 43_12    TTCCGCTGGG CGCAGGATCA CGTGACCGAG GTGGCGCATG AGTTCTACGT
 43_20    TTCCGCTGGG CGCAGGATCA CGTGACCGAG GTGGCGCATG AGTTCCACGT
 43_21    TTCCGCTGGG CGCAGGATCA CGTGACCGAG GTGGCGCATG AGTTCCACGT
 43_23    TTCCGCTGGG CGCAGGATCA CGTGACCGAG GTGGCGCATG AGTTCCACGT
 43_25    TTCCGCTGGG CGCAGGATCA CGTGACCGAG GTGGCGCATG AGTTCCACGT
  44_1    TTCCGCTGGG CGCAGGATCA CGTGACCGAG GTGGCGCACG AGTTCTACGT
  44_5    TTCCGCTGGG CGCAGGATCA CGTGACCGAG GTGGCGCACG AGTTCTACGT
223_10    .......... .......... .......... .......... ..........
 223_2    .......... .......... .......... .......... ..........
 223_4    .......... .......... .......... .......... ..........
 223_5    .......... .......... .......... .......... ..........
 223_6    .......... .......... .......... .......... ..........
 223_7    .......... .......... .......... .......... ..........
  A3_4    TTCCGGTGGG CTCAAGATCA CGTGACTGAG GTGGAGCATG AGTTCTACGT
  A3_5    TTCCGGTGGG CTCAAGATCA CGTGACTGAG GTGGAGCATG AGTTCTACGT
  A3_7    TTCCGGTGGG CTCAAGATCA CGTGACTGAG GTGGAGCATG AGTTCTACGT
  A3_3    TTCCGGTGGG CTCAAGATCA CGTGACTGAG GTGGAGCATG AGTTCTACGT
 42_12    TTCCGCTGGG CGCAGGATCA CGTGACCGAG GTGGCGCATG AGTTCTACGT
  AAV1    TTCCGCTGGG CGCAGGATCA CGTGACCGAG GTGGCGCATG AGTTCTACGT
  AAV2    TTCCGGTGGG CAAAGGATCA CGTGGTTGAG GTGGAGCATG AATTCTACGT
  AAV3    TTCCGGTGGG CTTCCGATCA CGTGACTGAC GTGGCTCATG AGTTCTACGT
  AAV8    TTCCGCTGGG CCAGTGATCA CGTGACCGAG GTGGCGCATG AGTTTTACGT
  AAV9    TTCCGCTGGG CCAGTGATCA CGTGACCGAG GTGGCGCATG AGTTTTACGT
  AAV7    TTCCGCTGGG CCAGTGATCA CGTGACCGAG GTGGCGCATG AGTTCTACGT
  44_2    TTCCGCTGGG CGCAGGATCA CGTGACCGAG GTGGCGCACG AGTTCTACGT
```

图1AJ

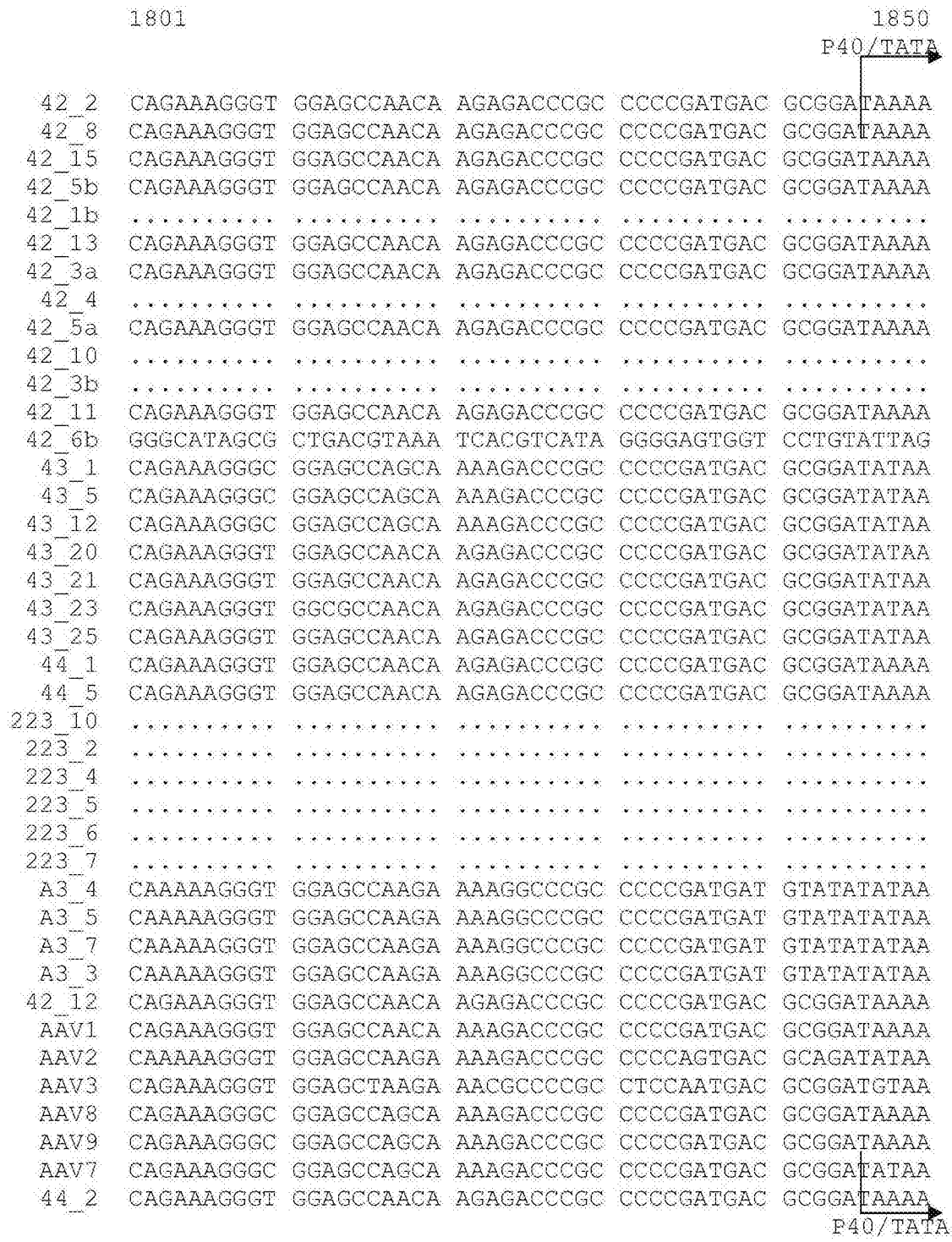

1801
1850
P40/TATA
42_2 CAGAAAGGGT GGAGCCAACA AGAGACCCGC CCCCGATGAC GCGGATAAAA
42_8 CAGAAAGGGT GGAGCCAACA AGAGACCCGC CCCCGATGAC GCGGATAAAA
42_15 CAGAAAGGGT GGAGCCAACA AGAGACCCGC CCCCGATGAC GCGGATAAAA
42_5b CAGAAAGGGT GGAGCCAACA AGAGACCCGC CCCCGATGAC GCGGATAAAA
42_1b .....
42_13 CAGAAAGGGT GGAGCCAACA AGAGACCCGC CCCCGATGAC GCGGATAAAA
42_3a CAGAAAGGGT GGAGCCAACA AGAGACCCGC CCCCGATGAC GCGGATAAAA
42_4 .....
42_5a CAGAAAGGGT GGAGCCAACA AGAGACCCGC CCCCGATGAC GCGGATAAAA
42_10 .....
42_3b .....
42_11 CAGAAAGGGT GGAGCCAACA AGAGACCCGC CCCCGATGAC GCGGATAAAA
42_6b GGGCATAGCG CTGACGTAAA TCACGTCATA GGGGAGTGGT CCTGTATTAG
43_1 CAGAAAGGGC GGAGCCAGCA AAAGACCCGC CCCCGATGAC GCGGATATAA
43_5 CAGAAAGGGC GGAGCCAGCA AAAGACCCGC CCCCGATGAC GCGGATATAA
43_12 CAGAAAGGGC GGAGCCAGCA AAAGACCCGC CCCCGATGAC GCGGATATAA
43_20 CAGAAAGGGT GGAGCCAACA AGAGACCCGC CCCCGATGAC GCGGATATAA
43_21 CAGAAAGGGT GGAGCCAACA AGAGACCCGC CCCCGATGAC GCGGATATAA
43_23 CAGAAAGGGT GGCGCCAACA AGAGACCCGC CCCCGATGAC GCGGATATAA
43_25 CAGAAAGGGT GGAGCCAACA AGAGACCCGC CCCCGATGAC GCGGATATAA
44_1 CAGAAAGGGT GGAGCCAACA AGAGACCCGC CCCCGATGAC GCGGATAAAA
44_5 CAGAAAGGGT GGAGCCAACA AGAGACCCGC CCCCGATGAC GCGGATAAAA
223_10 .....
223_2 .....
223_4 .....
223_5 .....
223_6 .....
223_7 .....
A3_4 CAAAAAGGGT GGAGCCAAGA AAAGGCCCGC CCCCGATGAT GTATATATAA
A3_5 CAAAAAGGGT GGAGCCAAGA AAAGGCCCGC CCCCGATGAT GTATATATAA
A3_7 CAAAAAGGGT GGAGCCAAGA AAAGGCCCGC CCCCGATGAT GTATATATAA
A3_3 CAAAAAGGGT GGAGCCAAGA AAAGGCCCGC CCCCGATGAT GTATATATAA
42_12 CAGAAAGGGT GGAGCCAACA AGAGACCCGC CCCCGATGAC GCGGATAAAA
AAV1 CAGAAAGGGT GGAGCCAACA AAAGACCCGC CCCCGATGAC GCGGATAAAA
AAV2 CAAAAAGGGT GGAGCCAAGA AAAGACCCGC CCCCAGTGAC GCAGATATAA
AAV3 CAGAAAGGGT GGAGCTAAGA AACGCCCCGC CTCCAATGAC GCGGATGTAA
AAV8 CAGAAAGGGC GGAGCCAGCA AAAGACCCGC CCCCGATGAC GCGGATAAAA
AAV9 CAGAAAGGGC GGAGCCAGCA AAAGACCCGC CCCCGATGAC GCGGATAAAA
AAV7 CAGAAAGGGC GGAGCCAGCA AAAGACCCGC CCCCGATGAC GCGGATATAA
44_2 CAGAAAGGGT GGAGCCAACA AGAGACCCGC CCCCGATGAC GCGGATAAAA
P40/TATA

图1AK

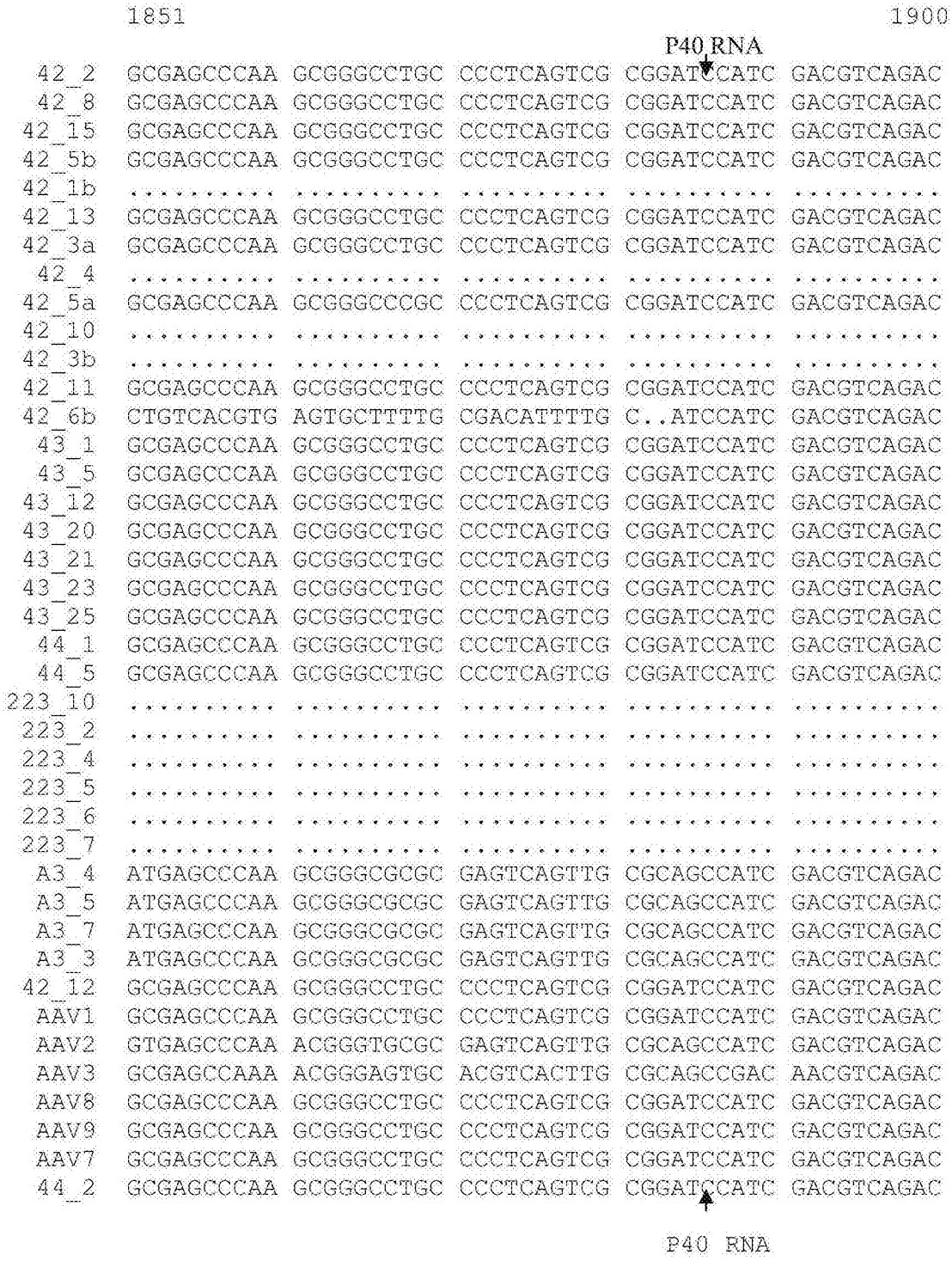

1851
1900
P40 RNA
42_2 GCGAGCCCAA GCGGGCCTGC CCCTCAGTCG CGGATCCATC GACGTCAGAC
42_8 GCGAGCCCAA GCGGGCCTGC CCCTCAGTCG CGGATCCATC GACGTCAGAC
42_15 GCGAGCCCAA GCGGGCCTGC CCCTCAGTCG CGGATCCATC GACGTCAGAC
42_5b GCGAGCCCAA GCGGGCCTGC CCCTCAGTCG CGGATCCATC GACGTCAGAC
42_1b ....
42_13 GCGAGCCCAA GCGGGCCTGC CCCTCAGTCG CGGATCCATC GACGTCAGAC
42_3a GCGAGCCCAA GCGGGCCTGC CCCTCAGTCG CGGATCCATC GACGTCAGAC
42_4 ....
42_5a GCGAGCCCAA GCGGGCCCGC CCCTCAGTCG CGGATCCATC GACGTCAGAC
42_10 ....
42_3b ....
42_11 GCGAGCCCAA GCGGGCCTGC CCCTCAGTCG CGGATCCATC GACGTCAGAC
42_6b CTGTCACGTG AGTGCTTTTG CGACATTTTG C..ATCCATC GACGTCAGAC
43_1 GCGAGCCCAA GCGGGCCTGC CCCTCAGTCG CGGATCCATC GACGTCAGAC
43_5 GCGAGCCCAA GCGGGCCTGC CCCTCAGTCG CGGATCCATC GACGTCAGAC
43_12 GCGAGCCCAA GCGGGCCTGC CCCTCAGTCG CGGATCCATC GACGTCAGAC
43_20 GCGAGCCCAA GCGGGCCTGC CCCTCAGTCG CGGATCCATC GACGTCAGAC
43_21 GCGAGCCCAA GCGGGCCTGC CCCTCAGTCG CGGATCCATC GACGTCAGAC
43_23 GCGAGCCCAA GCGGGCCTGC CCCTCAGTCG CGGATCCATC GACGTCAGAC
43_25 GCGAGCCCAA GCGGGCCTGC CCCTCAGTCG CGGATCCATC GACGTCAGAC
44_1 GCGAGCCCAA GCGGGCCTGC CCCTCAGTCG CGGATCCATC GACGTCAGAC
44_5 GCGAGCCCAA GCGGGCCTGC CCCTCAGTCG CGGATCCATC GACGTCAGAC
223_10 ....
223_2 ....
223_4 ....
223_5 ....
223_6 ....
223_7 ....
A3_4 ATGAGCCCAA GCGGGCGCGC GAGTCAGTTG CGCAGCCATC GACGTCAGAC
A3_5 ATGAGCCCAA GCGGGCGCGC GAGTCAGTTG CGCAGCCATC GACGTCAGAC
A3_7 ATGAGCCCAA GCGGGCGCGC GAGTCAGTTG CGCAGCCATC GACGTCAGAC
A3_3 ATGAGCCCAA GCGGGCGCGC GAGTCAGTTG CGCAGCCATC GACGTCAGAC
42_12 GCGAGCCCAA GCGGGCCTGC CCCTCAGTCG CGGATCCATC GACGTCAGAC
AAV1 GCGAGCCCAA GCGGGCCTGC CCCTCAGTCG CGGATCCATC GACGTCAGAC
AAV2 GTGAGCCCAA ACGGGTGCGC GAGTCAGTTG CGCAGCCATC GACGTCAGAC
AAV3 GCGAGCCAAA ACGGGAGTGC ACGTCACTTG CGCAGCCGAC AACGTCAGAC
AAV8 GCGAGCCCAA GCGGGCCTGC CCCTCAGTCG CGGATCCATC GACGTCAGAC
AAV9 GCGAGCCCAA GCGGGCCTGC CCCTCAGTCG CGGATCCATC GACGTCAGAC
AAV7 GCGAGCCCAA GCGGGCCTGC CCCTCAGTCG CGGATCCATC GACGTCAGAC
44_2 GCGAGCCCAA GCGGGCCTGC CCCTCAGTCG CGGATCCATC GACGTCAGAC
P40 RNA

图1AL

```
          1901                                                      1950
  42_2    GCGGAAGGAG CTCCGGTGGA CTTTGCCGAC AGGTACCAAA ACAAATGTTC
  42_8    GCGGAAGGAG CTCCGGTGGA CTTTGCCGAC AGGTACCAAA ACAAATGTTC
 42_15    GCGGAAGGAG CTCCGGTGGA CTTTGCCGAC AGGTACCAAA ACAAATGTTC
 42_5b    GCGGAAGGAG CTCCGGTGGA CTTTGCCGAC AGGTACCAAA ACAAATGTTC
 42_1b    .......... .......... .......... .......... ..........
 42_13    GCGGAAGGAG CTCCGGTGGA CTTTGCCGAC AGGTACCAAA ACAAATGTTC
 42_3a    GCGGAAGGAG CTCCGGTGGA CTTTGCCGAC AGGTACCAAA ACAAATGTTC
  42_4    .......... .......... .......... .......... ..........
 42_5a    GCGGAAGGAG CTCCGGTGGA CTTTGCCGAC AGGTACCAAA ACAAATGTTC
 42_10    .......... .......... .......... .......... ..........
 42_3b    .......... .......... .......... .......... ..........
 42_11    GCGGAAGGAG CTCCGGTGGA CTTTGCCGAC AGGTACCAAA ACAAATGTTC
 42_6b    GCGGAAGGAG CTCCGGTGGA CTTTGCCGAC AGGTACCAAA ACAAGTGTTC
  43_1    GCGGAAGGAG CTCCGGTGGA CTTTGCCGAC AGGTACCAAA ACAAATGTTC
  43_5    GCGGAAGGAG CTCCGGTGGA CTTTGCCGAC AGGTACCAAA ACAAATGTTC
 43_12    GCGGAAGGAG CTCCGGTGGA CTTTGCCGAC AGGTACCAAA ACAAATGTTC
 43_20    GCGGAAGGAG CTCCGGTGGA CTTTGCCGAC AGGTACCAAA ACAAATGTTC
 43_21    GCGGAAGGAG CTCCGGTGGA CTTTGCCGAC AGGTACCAAA ACAAATGTTC
 43_23    GCGGAAGGAG CTCCGGTGGA CTTTGCCGAC AGGTACCAAA ACAAATGTTC
 43_25    GCGGAAGGAG CTCCGGTGGA CTTTGCCGAC AGGTACCAAA ACAAATGTTC
  44_1    GCGGAAGGAG CTCCGGTGGA CTTTGCCGAC AGGTACCAAA ACAAATGTTC
  44_5    GCGGAAGGAG CTCCGGTGGA CTTTGCCGAC AGGTACCAAA ACAAATGTTC
223_10    .......... .......... .......... .......... ..........
 223_2    .......... .......... .......... .......... ..........
 223_4    .......... .......... .......... .......... ..........
 223_5    .......... .......... .......... .......... ..........
 223_6    .......... .......... .......... .......... ..........
 223_7    .......... .......... .......... .......... ..........
  A3_4    GCGGA...AG CTTCGATAAA CTACGCGGGC AGGTACCAAA ACAAATGTTC
  A3_5    GCGGA...AG CTTCGATAAA CTACGCGGAC AGGTACCAAA ACAAATGTTC
  A3_7    GCGGA...AG CTTCGATAAA CTACGCGGAC AGGTACCAAA ACAAATGTTC
  A3_3    GCGGA...AG CTTCGATAAA CTACGCGGAC AGGTACCAAA ACAAATGTTC
 42_12    GCGGAAGGAG CTCCGGTGGA CTTTGCCGAC AGGTACCAAA ACAAATGTTC
  AAV1    GCGGAAGGAG CTCCGGTGGA CTTTGCCGAC AGGTACCAAA ACAAATGTTC
  AAV2    GCGGA...AG CTTCGATCAA CTACGCAGAC AGGTACCAAA ACAAATGTTC
  AAV3    GCGGA...AG CACCGGCGGA CTACGCGGAC AGGTACCAAA ACAAATGTTC
  AAV8    GCGGAAGGAG CTCCGGTGGA CTTTGCCGAC AGGTACCAAA ACAAATGTTC
  AAV9    GCGGAAGGAG CTCCGGTGGA CTTTGCCGAC AGGTACCAAA ACAAATGTTC
  AAV7    GCGGAAGGAG CTCCGGTGGA CTTTGCCGAC AGGTACCAAA ACAAATGTTC
  44_2    GCGGAAGGAG CTCCGGTGGA CTTTGCCGAC AGGTACCAAA ACAAATGTTC
```

图1AM

```
          1951                                                        2000
  42_2    TCGTCACGCG GGCATGCTTC AGATGCTGTT TCCCTG.CAA GACATGCGAG
  42_8    TCGTCACGCG GGCATGCTTC AGATGCTGTT TCCCTG.CAA GACATGCGAG
 42_15    TCGTCACGCG GGCATGCTTC AGATGCTGTT TCCCTG.CAA GACATGCGAG
 42_5b    TCGTCACGCG GGCATGCTTC AGATGCTGTT TCCCTG.CAA GACATGCGAG
 42_1b    .......... .......... ....GAATTC GCCCTT.... .GGCTGCGTC
 42_13    TCGTCACGCG GGCATGCTTC AGATGCTGTT TCCCTG.CAA GACATGCGAG
 42_3a    TCGTCACGCG GGCATGCTTC AGATGCTGCT TCCCTG.CAA GACATGCGAG
  42_4    .......... .......... ....GAATTC GCCCTTTCTA CGGCTGCGTC
 42_5a    TCGTCACGCG GGCATGCTTC AGATGCTGTT TCCCTG.CAA AACATGCGAG
 42_10    .......... .......... ....GAATTC GCCCTTTCTA CGGCTGCGTC
 42_3b    .......... .......... ....GAATTC GCCCTTTCTA CGGCTGCGTC
 42_11    TCGTCACGCG GGCATGCTTC AGATGCTGTT TCCCTG.CAA GACATGCGAG
 42_6b    TCGTCACGCG GGCATGCTTC AGATGCTGTT TCCCTG.CAA GACATGCGAG
  43_1    TCGTCACGCG GGCATGCTTC AGATGCTGTT TCCCTG.CAA AACGTGCGAG
  43_5    TCGTCACGCG GGCATGCTTC AGACGCTGTT TCCCTG.CAA AACGTGCGAG
 43_12    TCGTCACGCG GGCATGCTCC AGATGCTGTT TCCCTG.CAA AACGTGCGAG
 43_20    TCGTCACGCG GGCATGCTTC AGATGCTGTT TCCCTG.CAA GACATGCGAG
 43_21    TCGTCACGCG GGCATGCTTC AGATGCTGTT TCCCTG.CAA GACATGCGAG
 43_23    TCGTCACGCG GGCATGCTTC AGATGCTGTT TCCCTG.CAA GACATGCGAG
 43_25    TCGTCACGCG GGCATGCTTC AGATGCTGTT TCCCTG.CAA GACATGCGAG
  44_1    TCGTCACGCG GGCATGCTTC AGATGCTGTT TCCCTG.CAA AACATGCGAG
  44_5    TCGTCACGCG GGCATGCTTC AGATGCTGTT TCCCTG.CAA AACATGCGAG
223_10    .......... .......... .......... .......... ..........
 223_2    .......... .......... .......... .......... ..........
 223_4    .......... .......... .......... .......... ..........
 223_5    .......... .......... .......... .......... ..........
 223_6    .......... .......... .......... .......... ..........
 223_7    .......... .......... .......... .......... ..........
  A3_4    TCGTCACGTG GGCATGAATC TGATGCTGTT TCCCTG.TCG ACAATGCGAA
  A3_5    TCGTCACGTG GGCATGAATC TGATGCTGTT TCCCTG.TCG ACAATGCGAA
  A3_7    TCGTCACGTG GGCATGAATC TGATGCTGTT TCCCTG.TCG ACAATGCGAA
  A3_3    TCGTCACGTG GGCATGAATC TGATGCTGTT TCCCTG.TCG ACAATGCGAA
 42_12    TCGTCACGCG GGCATGCTTC AGATGCTGTT TCCCTG.CAA GACATGCGAG
  AAV1    TCGTCACGCG GGCATGCTTC AGATGCTGTT TCCCTG.CAA GACATGCGAG
  AAV2    TCGTCACGTG GGCATGAATC TGATGCTGTT TCCCTG.CAG ACAATGCGAG
  AAV3    TCGTCACGTG GGCATGAATC TGATGCTTTT TCCCTG.TAA AACATGCGAG
  AAV8    TCGTCACGCG GGCATGCTTC AGATGCTGTT TCCCTG.CAA AACGTGCGAG
  AAV9    TCGTCACGCG GGCATGCTTC AGATGCTGCT TCCCTG.CAA AACGTGCGAG
  AAV7    TCGTCACGCG GGCATGATTC AGATGCTGTT TCCCTG.CAA AACGTGCGAG
  44_2    TCGTCACGCG GGCATGCTTC AGATGCTGTT TCCCTG.CAA AACATGCGAG
```

图1AN

```
          2001                                                      2050
  42_2    AGAATGAATC AGAATTTCAA CATTTGCTTC ACGCACGGGA CCAGAGACTG
  42_8    AGAATGAATC AGAATTTCAA CATTTGCTTC ACGCACGGGA CCAGAGACTG
 42_15    AGAATGAATC AGAATTTCAA CATTTGCTTC ACGCGCGGGA CCAGAGACTG
 42_5b    AGAATGAATC AGAATTTCAA CATTTGCTTC ACGCACGGGA CCAGAGACTG
 42_1b    A.ACTGGACC A..ATGAGAA CTTTCCCTTC A........A CGATTGCGTC
 42_13    AGAATGAATC AGAATTTCAA CATTTGCTTC ACGCACGGGA CCAGAGACTG
 42_3a    AGAATGAATC AGAATTTCAG CATTTGCTTC ACGCACGGGA CCAGAGACTG
  42_4    A.ACTGGACC A..ATGAGAA CTTTCCCTTC A........A CGATTGCGTC
 42_5a    AGAATGAATC AGAATTTCAA CATTTGCTTC ACGCACGGGA CCAGAGACTG
 42_10    A.ACTGGACC A..ATGAGAA CTTTCCCTTC A........A CGATTGCGTC
 42_3b    A.ACTAGACC A..ATGAGAA CTTTCCCTTC A........A CGATTGCGTC
 42_11    AGAATGAATC AGAATTTCAA CATTTGCTTC ACGCACGGGA CCGGAGACTG
 42_6b    AGAATGAATC AGAATTTCAA CATTTGCTTC ACGCACGGGA CCAGAGACTG
  43_1    AAAATGAATC AGAATTTCAA CATTTGCTTC ACGCACGGGG TCAGAGACTG
  43_5    AGAATGAATC AGAATTTCAA CATTTGCTTC ACGCACGGGG TCAGAGACTG
 43_12    AGAATGAATC AGAATTTCAA CATTTGCTTC ACGCACGGGG TCAGAGACTG
 43_20    AGAATGAATC AGAATTTCAA CATTTGCTTC ACGCACGGGA CCAGAGACTG
 43_21    AGAATGAATC AGAATTTCAA CATTTGCTTC ACGCACGGGA CCAGAGACTG
 43_23    AGAATGAATC AGAATTTCAA CATTTGCTTC ACGCACGGGA CCAGAGACTG
 43_25    AGAATGAATC AGAATTTCAA CATTTGCTTC ACGCACGGGA CCAGAGACTG
  44_1    AGAATGAATC AGAATTTCAA CATTTGCTTC ACGCACGGGA CCAGAGACTG
  44_5    AGAATGAATC AGAATTTCAA CATTTGCTTC ACGCACGGGA CCAGAGACTG
223_10    .......... .......... .......... .......... ..........
 223_2    .......... .......... .......... .......... ..........
 223_4    .......... .......... .......... .......... ..........
 223_5    .......... .......... .......... .......... ..........
 223_6    .......... .......... .......... .......... ..........
 223_7    .......... .......... .......... .......... ..........
  A3_4    AGAATGAATC AGAATTCAAA TATCTGCTTC ACACACGGGC AAAAAGACTG
  A3_5    AGAATGAATC AGAATTCAAA TATCTGCTTC ACACACGGGC AAAAAGACTG
  A3_7    AGAATGAATC AGAATTCAAA TATCTGCTTC ACACACGGGC AAAAAGACTG
  A3_3    AGAATGAATC AGAATTCAAA TATCTGCTTC ACACACGGGC AAAAAGACTG
 42_12    AGAATGAATC AGAATTTCAA CATTTGCTTC ACGCACGGGA CCAGAGACTG
  AAV1    AGAATGAATC AGAATTTCAA CATTTGCTTC ACGCACGGGA CGAGAGACTG
  AAV2    AGAATGAATC AGAATTCAAA TATCTGCTTC ACTCACGGAC AGAAAGACTG
  AAV3    AGAATGAATC AAATTTCCAA TGTCTGTTTT ACGCATGGTC AAAGAGACTG
  AAV8    AGAATGAATC AGAATTTCAA CATTTGCTTC ACACACGGGG TCAGAGACTG
  AAV9    AGAATGAATC AGAATTTCAA CATTTGCTTC ACACACGGGG TCAGAGACTG
  AAV7    AGAATGAATC AGAATTTCAA CATTTGCTTC ACACACGGGG TCAGAGACTG
  44_2    AGAATGAATC AGAATTTCAA CATTTGCTTC ACGCACGGGA CCAGAGACTG
```

图1A0

```
          2051                                                    2100
  42_2    TTCAGAATGT TTCCCCGGCG TGTCAGAATC TCAACC.... ..GGTCGTCA

  42_8    TTCAGAATGT TTCCCCGGCG TGTCAGAATC TCAACC.... ..GGTCGTCA
 42_15    TTCAGAATGT TTCCCGGGCG TGTCAGAATC TCAACC.... ..GGTCGTCA
 42_5b    TTCAGAATGT TTCCCCGGCG TGTCAGAATC TCAACC.... ..GGTCGTCA
 42_1b    GACAAGATGG TGATCTGGTG GG..AGGAGG GCAAGA.... ..TGACGGCC
 42_13    TTCAGAATGT TTCCCCGGCG TGTCAGAATC TCAACC.... ..GGTCGTCA
 42_3a    TTCAGAATGT TTCCCCGGCG TGTCAGAATC TCAACC.... ..GGTCGTCA
  42_4    GACAAGATGG TGATCTGGTG GG..AGGAGG GCAAGA.... ..TGACGGCC
 42_5a    TTCAGAATGT TTCCCCGGCG TGTCAGAATC TCAACC.... ..GGTCGTCA
 42_10    GACAAGATGG TGATCTGGTG GG..AGGAGG GCAAGA.... ..TGACGGCC
 42_3b    GACAAGATGG TGATCTGGTG GG..AGGAGG GCAAGA.... ..TGACGGCC
 42_11    TTCAGAATGT TTCCCCGGCG TGTCAGAATC TCAACC.... ..GGTCGTCA
 42_6b    TTCAGAATGT TTCCCCGGCG TGTCAGAATC TCAACC.... ..GGTCGTCA
  43_1    CTCAGAATGT TTCCCCGGTG CATCAGAATC TCAACC.... ..GGTCGTCA
  43_5    CTCAGAATGT TTCCCCGGTG CATCAGAATC TCAACC.... ..GGTCGTCA
 43_12    CTCAGAATGT TTCCCCGGTG CATCAGAATC TCAACC.... ..GGTCGTCA
 43_20    TTCAGAATGT TTCCCCGGCG TGTCAGAATC TCAACC.... ..GGTCGTCA
 43_21    TTCAGAATGT TTCCCCGGCG TGTCAGAATC TCAACC.... ..GGTCGTCA
 43_23    TTCAGAATGT TTCCCCGGCG TGTCAGAATC TCAACC.... ..GGTCGTCA
 43_25    TTCAGAATGT TTCCCCGGCG TGTCAGAATC TCAACC.... ..GGTCGTCA
  44_1    TTCAGAATGT TTCCCCGGCG TGTCAGAATC TCAACC.... ..GGTCGTCA
  44_5    TTCAGAATGT TTCCCCGGCG TGTCAGAATC TCAACC.... ..GGTTGTCA
223_10    .......... .......... .......... .......... ..........
 223_2    .......... .......... .......... .......... ..........
 223_4    .......... .......... .......... .......... ..........
 223_5    .......... .......... .......... .......... ..........
 223_6    .......... .......... .......... .......... ..........
 223_7    .......... .......... .......... .......... ..........
  A3_4    TTTGGAATGC TTTCCCG... TGTCAGAATC TCAACCCGTT TCTGTCGTCA
  A3_5    TTTGGAATGC TTTCCCG... TGTCAGAATC TCAACCCGTT CCTGTCGTCA
  A3_7    TTTGGAATGC TTTCCCG... TGTCAGAATC TCAACCCGTT TCTGTCGTCA
  A3_3    TTTGGAATGC TTTCCCG... TGTCAGAATC TCAACCCGTT TCTGTCGTCA
 42_12    TTCAGAATGT TTCCCCGGCG TGTCAGAATC TCAACC.... ..GGTCGTCA
  AAV1    TTCAGAGTGC TTCCCCGGCG TGTCAGAATC TCAACC.... ..GGTCGTCA
  AAV2    TTTAGAGTGC TTTCCCG... TGTCAGAATC TCAACCCGTT TCTGTCGTCA
  AAV3    TGGGGAATGC TTCCCTGGAA TGTCAGAATC TCAACCCGTT TCTGTCGTCA
  AAV8    CTCAGAGTGT TTCCCCGGCG TGTCAGAATC TCAACC.... ..GGTCGTCA
  AAV9    CTCAGAGTGT TTCCCCGGCG TGTCAGAATC TCAACC.... ..GGTCGTCA
  AAV7    TTTAGAGTGT TTCCCCGGCG TGTCAGAATC TCAACC.... ..GGTCGTCA
  44_2    TTCAGAATGT TTCCCCGGCG TGTCAGAATC TCAACC.... ..GGTCGTCA
```

图1AP

```
          2101                                                      2150
  42_2    GAAAGAGGAC GTATCGGAAA CTCTGTGCCA TTCATCATCT GCTGGGG.CG

  42_8    GAAAGAGGAC GTATCGGAAA CTCTGTGCCA TTCATCATCT GCTAGGG.CG
 42_15    GAAAGAGGAC GTATCGGAAA CTCTGTGCCA TTCATCATCT GCTGGGG.CG
 42_5b    GAAAGAGGAC GTATCGGAAA CTCTGTGCCA TTCATCATCT GCTGGGG.CG
 42_1b    .AAGGTCGTG GAGTCCGCCA AG....GCCA TTCATCATCT GCTGGGG.CG
 42_13    GAAAGAGGAC GTATCGGAAA CTCTGTGCCA TTCATCATCT GCTGGGG.CG
 42_3a    GAAAGAGGAC GTATCGGAAA CTCTGTGCCA TTCATCATCT GCTGGGG.CG
  42_4    .AAGGTCGTG GAGTCCGCCA AG....GCCA TTCATCATCT GCTGGGG.CG
 42_5a    GAAAGAGGAC GTATCGGAAA CTCTGTGCCA TTCATCATCT GCTGGGG.CG
 42_10    AA....GGTC GTGAAGTCCG CCAAG.GCCA TTCATCATCT GCTGGGG.CG
 42_3b    AA....GGTC GTGGAGTCCG CCAAG.GCCA TTCATCATCT GCTGGGG.CG
 42_11    GAAAGAGGAC GTATCGGAAA CTCTGTGCCA TTCATCATCT GCTGGGG.CG
 42_6b    GAAAGAGGAC GTATCGGAAA CTCTGTGCCA TTCATCATCT GCTGGGG.CG
  43_1    GAAAAAAAAC GTATCAGAAA CTGTGTGCCA TTCATCATCT GCTGGGG.CG
  43_5    GAAAAAAAAC GTATCAGAAA CTGTGTGCCA TTCATCATCT GCTGGGG.CG
 43_12    GAAAAAAAAC GTATCAGAAA CTGTGTGCCA TTCATCATCT GCTGGGG.CG
 43_20    GAAAGAGGAC GTATCGGAAA CTCTGTGCGA TTCATCATCT GCTGGGG.CG
 43_21    GAAAGAGGAC GTATCGGAAA CTCTGTGCGA TTCATCATCT GCTGGGG.CG
 43_23    GAAAGAGGAC GTATCGGAAA CTCTGTGCGA TTCATCATCT GCTGGGG.CG
 43_25    GAAAGAGGAC GTATCGGAAA CTCTGTGCGA TTCATCATCT GCTGGGG.CG
  44_1    GAAAAAAGAC GTATCGGAAA CTCTGTGCGA TTCATCATCT GCTGGGG.CG
  44_5    GAAAAAAGAC GTATCGGAAA CTCTGTGCGA TTCATCATCT GCTGGGG.CG
223_10    .......... .......... .......... .......... ..........
 223_2    .......... .......... .......... .......... ..........
 223_4    .......... .......... .......... .......... ..........
 223_5    .......... .......... .......... .......... ..........
 223_6    .......... .......... .......... .......... ..........
 223_7    .......... .......... .......... .......... ..........
  A3_4    GAAAAACG.. .TATCAGAAA CTTTGTTACA TTCATCATAT CATGGGA.AA
  A3_5    GAAAAACG.. .TATCAGAAA CTTTGTTACA TTCATCATAT CATGGGA.AA
  A3_7    GAAAAACG.. .TATCAGAAA CTTTGTTACA TTCATCATAT CATGGGA.AA
  A3_3    GAAAAACG.. .TATCAGAAA CTTTGTTACA TTCATCATAT CATGGGA.AA
 42_12    GAAAGAGGAC GTATCGGAAA CTCTGTGCCA TTCATCATCT GCTGGGG.CG
  AAV1    GAAAGAGGAC GTATCGGAAA CTCTGTGCCA TTCATCATCT GCTGGGG.CG
  AAV2    AAAAGGCG.. .TATCAGAAA CTGTGCTACA TTCATCATAT CATGGGA.AA
  AAV3    AAAAGAAGAC TTATCAGAAA CTGTGTCCAA TTCATCATAT CCTGGGA.AG
  AAV8    GAAAGAGGAC GTATCGGAAA CTCTGTGCGA TTCATCATCT GCTGGGG.CG
  AAV9    GAAAGAGGAC GTATCGGAAA CTCTGTGCGA TTCATCATCT GCTGGGG.CG
  AAV7    GAAAAAAGAC GTATCGGAAA CTCTGCGCGA TTCATCATCT GCTGGGG.CG
  44_2    GAAAAAAGAC GTATCGGAAA CTCTGTGCGA TTCATCATCT GCTGGGGGCG
```

图1AQ

```
          2151                                                      2200
  42_2    GGCTCCCGAG ATTGCTTGCT CGGCCTGCGA TCTGGTCAAC GTGGACCTGG
  42_8    GGCTCCCGAG ATTGCTTGCT CGGCCTGCGA TCTGGTCAAC GTGGACCTGG
 42_15    GGCTCCCGAG ATTGCTTGCT CGGCCTGCGA TCTGGTCAAC GTGGACCTGG
 42_5b    GGCTCCCGAG ATTGCTTGCT CGGCCTGCGA TCTGGTCAAC GTGGACCTGG
 42_1b    GGCTCCCGAG ATTGCTTGCT CGGCCTGCGA TCTGGTCAAC GTGGACCTGG
 42_13    GGCTCCCGAG ATTGCTTGCT CGGCCTGCGA TCTGGTCAAC GTGGACCTGG
 42_3a    GGCTCCCGAG ATTGCTTGCT CGGCCTGCGA TCTGGTCAAC GTGGACCTGG
  42_4    GGCTCCCGAG ATTGCTTGCT CGGCCTGCGA TCTGGTCAAC GTGGACCTGG
 42_5a    GGCTCCCGAG ATTGCTTGCT CGGCCTGCGA TCTGGTCAAC GTGGACCTGG
 42_10    GGCTCCCGAG ATTGCTTGCT CGGCCTGCGA TCTGGTCAAC GTGGACCTGG
 42_3b    GGCTCCCGAG ATTGCTTGCT CGGCCTGCGA TCTGGTCAAC GTGGACCTGG
 42_11    GGCTCCCGAG ATTGCTTGCT CGGCCTGCGA TCTGGTCAAC GTGGACCTGG
 42_6b    GGCTCCCGAG ATTGCTTGCT CGGCCTGCGA TCTGGTCAAC GTGGACCTGG
  43_1    GGCACCCGAG ATTGCTTGCT CGGCCTGCGA TCTGGTCAAC GTGGACCTGG
  43_5    GGCACCCGAG ATTGCTTGCT CGGCCTGCGA TCTGGTCAAC GTGGACCTGG
 43_12    GGCACCCGAG ATTGCTTGCT CGGCCTGCGA TCTGGTCAAC GTGGACCTGG
 43_20    GGCTCCCGAG ATTGCTTGCT CGGCCTGCGA TCTGGTCAAC GTGGACCTGG
 43_21    GGCTCCCGAG ATTGCTTGCT CGGCCTGCGA TCTGGTCAAC GTGGACCTGG
 43_23    GGCTCCCGAG ATTGCTTGCT CGGCCTGCGA TCTGGTCAAC GTGGACCTGG
 43_25    GGCTCCCGAG ATTGCTTGCT CGGCCTGCGA TCTGGTCAAC GTGGACCTGG
  44_1    GGCACCCGAG ATTGCTTGCT CGGCCTGCGA TCTGGTCAAC GTGGACCTAG
  44_5    GGCACCCGAG ATTGCTTGCT CGGCCTGCGA TCTGGTCAAC GTGGACCTAG
223_10    .......... .......... .......... .......... ..........
 223_2    .......... .......... .......... .......... ..........
 223_4    .......... .......... .......... .......... ..........
 223_5    .......... .......... .......... .......... ..........
 223_6    .......... .......... .......... .......... ..........
 223_7    .......... .......... .......... .......... ..........
  A3_4    AGAACCAGAC ...GCCTGCA CTGCCTGCGA CCTGGTAAAT GTGGACTTGG
  A3_5    AGTACCAGAC ...GCCTGCA CTGCCTGCGA CCTGGTAAAT GTGGACTTGG
  A3_7    AGTACCAGAC ...GCCTGCA CTGCCTGCGA CCTGGTAAAT GTGGACTTGG
  A3_3    AGTACCAGAC ...GCCTGCA CTGCCTGCGA CCTGGTAAAT GTGGACTTGG
 42_12    GGCTCCCGAG ATTGCTTGCT CGGCCTGCGA TCTGGTCAAC GTGGACCTGG
  AAV1    GGCTCCCGAG ATTGCTTGCT CGGCCTGCGA TCTGGTCAAC GTGGACCTGG
  AAV2    GGTGCCAGAC ...GCTTGCA CTGCCTGCGA TCTGGTCAAT GTGGATTTGG
  AAV3    GGCACCCGAG ATTGCCTGTT CGGCCTGCGA TTTGGCCAAT GTGGACTTGG
  AAV8    GGCTCCCGAG ATTGCTTGCT CGGCCTGCGA TCTGGTCAAC GTGGACCTGG
  AAV9    GGCTCCCGAG ATTGCTTGCT CGGCCTGCGA TCTGGTCAAC GTGGACCTGG
  AAV7    GGCGCCCGAG ATTGCTTGCT CGGCCTGCGA CCTGGTCAAC GTGGACCTGG
  44_2    GGCACCCGAG ATTGCTTGCT CGGCCTGCGA TCTGGTCAAC GTGGACCTAG
```

图1AR

```
          2201                                                                2250

                                          Rep 78 终止                vp1 起始
  42_2    ATGACCGTGT TTCTGAGCAA TAAATGACTT AAACCAGGTA TGGCTGCCGA
  42_8    ATGACTGTGT TTCTGAGCAA TAAATGACTT AAACCAGGTA TGGCTGCCGA
 42_15    ATGACTGTGT TTCTGAGCAA TAAATGACTT AAACCAGGTA TGGCTGCCGA
 42_5b    ATGACTGTGT TTCTGAGCAA TAAATGACTT AAACCAGGTA TGGCTGCCGA
 42_1b    ATGACTGTGT TTCTGAGCAA TAAATGACTT AAACCAGGTA TGGCTGCCGA
 42_13    ATGACTGTGT TTCTGAGCAA TAAATGACTT AAACCAGGTA TGGCTGCCGA
 42_3a    ATGACTGTGT TTCTGAGCAA TAAATGACTT AAACCAGGTA TGGCTGCCGA
  42_4    ATGACTGTGT TTCTGAGCAA TAAATGACTT AAACCAGGTA TGGCTGCCGA
 42_5a    ATGACTGTGT TTCTGAGCAA TAAATGACTT AAACCAGGTA TGGCTGCCGA
 42_10    ATGACTGTGT TTCTGAGCAA TAAATGACTT AAACCAGGTA TGGCTGCCGA
 42_3b    ATGACTGTGT TTCTGAGCAA TAAATGACTT AAACCAGGTA TGGCTGCCGA
 42_11    ATGACTGTGT TTCTGAGCAA TAAATGACTT AAACCAGGTA TGGCTGCCGA
 42_6b    ATGACTGTGT TTCTGAGCAA TAAATGACTT AAACCAGGTA TGGCTGCCGA
  43_1    ACGACTGTGT TTCTGAGCAA TAAATGACTT AAACCAGGTA TGGCTGCCGA
  43_5    ACGACTGTGT TTCTGAGCAA TAAATGACTT AAACCAGGTA TGGCTGCCGA
 43_12    ACGACTGTGT TTCTGAGCAA TAAATGACTT AAACCAGGTA TGGCTGCCGA
 43_20    ATGACTGTGT TTCTGAGCAA TAAATGACTT AAACCAGGTA TGGCTGCCGA
 43_21    ATGACTGTGT TTCTGAGCAA TAAATGACTT AAACCAGGTA TGGCTGCCGA
 43_23    ATGACTGTGT TTCTGAGCAA TAAATGACTT AAACCAGGTA TGGCTGCCGA
 43_25    ATGACTGTGT TTCTGAGCAA TAAATGACTT AAACCAGGTA TGGCTGCCGA
  44_1    ATGACTGTGT TTCTGAGCAA TAAATGACTT AAACCAGGTA TGGCTGCCGA
  44_5    ATGACTGTGT TTCTGAGCAA TAAATGACTT AAACCAGGTA TGGCTGCCGA
223_10    .......... .......... .......... .......... ..........
 223_2    .......... .......... .......... .......... ..........
 223_4    .......... .......... .......... .......... ..........
 223_5    .......... .......... .......... .......... ..........
 223_6    .......... .......... .......... .......... ..........
 223_7    .......... .......... .......... .......... ..........
  A3_4    ATGACTGTAT TTCTGAGCAA TAAATGACTT AAATCAGGTA TGGCTGCTGA
  A3_5    ATGACTGTAT TTCTGAGCAA TAAATGACTT AAATCAGGTA TGGCTGCTGA
  A3_7    ATGACTGTAT TTCTGAGCAA TAAATGACTT AAATCAGGTA TGGCTGCTGA
  A3_3    ATGACTGTAT TTCTGAGCAA TAAATGACTT AAATCAGGTA TGGCTGCTGA
 42_12    ATGACTGTGT TTCTGAGCAA TAAATGACTT AAACCAGGTA TGGCTGCCGA
  AAV1    ATGACTGTGT TTCTGAGCAA TAAATGACTT AAACCAGGTA TGGCTGCCGA
  AAV2    ATGACTGCAT CTTTGAACAA TAAATGATTT AAATCAGGTA TGGCTGCCGA
  AAV3    ATGACTGTGT TTCTGAGCAA TAAATGACTT AAACCAGGTA TGGCTGCTGA
  AAV8    ATGACTGTGT TTCTGAGCAA TAAATGACTT AAACCAGGTA TGGCTGCCGA
  AAV9    ATGACTGTGT TTCTGAGCAA TAAATGACTT AAACCAGGTA TGGCTGCCGA
  AAV7    ACGACTGCGT TTCTGAGCAA TAAATGACTT AAACCAGGTA TGGCTGCCGA
  44_2    ATGACTGTGT TTCTGAGCAA TAAATGACTT AAACCAGGTA TGGCTGCCGA
                                Rep78 终止                vp1 起始
```

图1AS

```
          2251                                                    2300
                                                        Rep68 终止
  42_2    TGGTTATCTT CCAGATTGGC TCGAGGACAA CCTCTCTGAG GGCATTCGCG
  42_8    TGGTTATCTT CCAGATTGGC TCGAGGACAA CCTCTCTGAG GGCATTCGCG
 42_15    TGGTTATCTT CCAGATTGGC TCGAGGACAA CCTCTCTGAG GGCATTCGCG
 42_5b    TGGTTATCTT CCAGATTGGC TCGAGGACAA CCTCTCTGAG GGCATTCGCG
 42_1b    TGGTTATCTT CCAGATTGGC TCGAGGACAA CCTCTCTGAG GGCATTCGCG
 42_13    TGGTTATCTT CCAGATTGGC TCGAGGACAA CCTCTCTGAG GGCATTCGCG
 42_3a    TGGTCATCTT CCAGATTGGC TCGAGGACAA CCTCTCTGAG GGCATTCGCG
  42_4    TGGTTATCTT CCAGATTGGC TCGAGGACAA CCTCTCTGAG GGCATTCGCG
 42_5a    TGGTTATCTT CCAGATTGGC TCGAGGACAA CCTCTCTGAG GGCATTCGCG
 42_10    TGGTTATCTT CCAGATTGGC TCGAGGACAA CCTCTCTGAG GGCATTCGCG
 42_3b    TGGTTATCTT CCAGATTGGC TCGAGGACAA CCTCTCTGAG GGCATTCGCG
 42_11    TGGTTATCTT CCAGATTGGC TCGAGGACAA CCTCTCTGAG GGCATTCGCG
 42_6b    TGGTTATCTT CCAGATTGGC TCGAGGACAA CCTCTCTGAG GGCATTCGCG
  43_1    TGGTTATCTT CCAGATTGGC TTGAGGACAA CCTCTCTGAG GGCATTCGCG
  43_5    TGGTTATCTT CCAGATTGGC TTGAGGACAA CCTCTCTGAG GGCATTCGCG
 43_12    TGGTTATCTT CCAGATTGGC TTGAGGACAA CCTCTCTGAG GGCATTCGCG
 43_20    TGGTTATCTT CCAGATTGGC TCGAGGACAA CCTCTCTGAG GGCATTCGCG
 43_21    TGGTTATCTT CCAGATTGGC TCGAGGACAA CCTCTCTGAG GGCATTCGCG
 43_23    TGGTTATCTT CCAGATTGGC TCGAGGACAA CCTCTCTGAG GGCATTCGCG
 43_25    TGGTTATCTT CCAGATTGGC TCGAGGACAA CCTCTCTGAG GGCATTCGCG
  44_1    TGGTTATCTT CCAGATTGGC TCGAGGACAA CCTCTCTGAG GGCATTCGCG
  44_5    TGGTTATCTT CCAGATTGGC TCGAGGACAA CCTCTCTGAG GGCATTCGCG
223_10    .......... .......... .......... .......... ..........
 223_2    .......... .......... .......... .......... ..........
 223_4    .......... .......... .......... .......... ..........
 223_5    .......... .......... .......... .......... ..........
 223_6    .......... .......... .......... .......... ..........
 223_7    .......... .......... .......... .......... ..........
  A3_4    CGGTTATCTT CCAGATTGGC TCGAGGACAC TCTCTCTGAA GGAATCAGAC
  A3_5    CGGTTATCTT CCAGATTGGC TCGAGGACAC TCTCTCTGAA GGAATCAGAC
  A3_7    CGGTTATCTT CCAGATTGGC TCGAGGACAC TCTCTCTGAA GGAATCAGAC
  A3_3    CGGTTATCTT CCAGATTGGC TCGAGGACAC TCTCTCTGAA GGAATCAGAC
 42_12    TGGTTATCTT CCAGATTGGC TCGAGGACAA CCTCTCTGAG GGCATCCGCG
  AAV1    TGGTTATCTT CCAGATTGGC TCGAGGACAA CCTCTCTGAG GGCATTCGCG
  AAV2    TGGTTATCTT CCAGATTGGC TCGAGGACAC TCTCTCTGAA GGAATAAGAC
  AAV3    CGGTTATCTT CCAGATTGGC TCGAGGACAA CCTTTCTGAA GGCATTCGTG
  AAV8    TGGTTATCTT CCAGATTGGC TCGAGGACAA CCTCTCTGAG GGCATTCGCG
  AAV9    TGGTTATCTT CCAGATTGGC TCGAGGACAA CCTCTCTGAG GGCATTCGCG
  AAV7    TGGTTATCTT CCAGATTGGC TCGAGGACAA CCTCTCTGAG GGCATTCGCG
  44_2    TGGTTATCTT CCAGATTGGC TCGAGGACAA CCTCTCTGAG GGCATTCGCG
                                                        Rep 68 终止
```

图1AT

```
          2301                                                      2350
  42_2    AGTGGTGGGA CTTGAAACCT GGAGCCCCGA AACCCAAAGC CAACCAGCAA

  42_8    AGTGGTGGGA CTTGAAACCT GGAGCCCCGA AACCCAAAGC CAACCAGCAA
 42_15    AGTGGTGGGA CTTGAAACCT GGAGCCCCGA AACCCAAAGC CAACCAGCAA
 42_5b    AGTGGTGGGA CTTGAAACCT GGAGCCCCGA AACCCAAAGC CAACCAGCAA
 42_1b    AGTGGTGGGA CTTGAGACCT GGAGCCCCGA AACCCAAAGC CAACCAGCAA
 42_13    AGTGGTGGGA CTTGAAACCT GGAGCCCCGA AACCCAAAGC CAACCAGCAA
 42_3a    AGTGGTGGGA CTTGAAACCT GGAGCCCCGA AACCCAAAGC CAACCAGCAA
  42_4    AGTGGTGGGA CTTGAAACCT GGAGCCCCGA AACCCAAAGC CAACCAGCAA
 42_5a    AGTGGTGGGA CTTGAAACCT GGAGCCCCGA AACCCAAAGC CAACCAGCAA
 42_10    AGTGGTGGGA CTTGAAACCT GGAGCCCCGA AACCCAAAGC CAACCAGCAA
 42_3b    AGTGGTGGGA CTTGAAACCT GGAGCCCCGA AACCCAAAGC CAACCAGCAA
 42_11    AGTGGTGGGA CTTGAAACCT GGAGCCCCGA AACCCAAAGC CAACCAGCAA
 42_6b    AGTGGTGGGA CTTGAAACCT GGAGCCCCGA AACCCAAAGC CAACCAGCAA
  43_1    AGTGGTGGGA CCTGAAACCT GGAGCCCCGA AACCCAAAGC CAACCAGCAA
  43_5    AGTGGTGGGA CCTGAAACCT GGAGCCCCGA AACCCAAAGC CAACCAGCAA
 43_12    AGTGGTGGGA CCTGAAACCT GGAGCCCCGA AACCCAAAGC CAACCAGCAA
 43_20    AGTGGTGGGA CTTGAAACCT GGAGCCCCGA AACCCAAAGC CAACCAGCAA
 43_21    AGTGGTGGGA CTTGAAACCT GGAGCCCCGA AACCCAAAGC CAACCAGCAA
 43_23    AGTGGTGGGA CTTGAAACCT GGAGCCCCGA AACCCAAAGC CAACCAGCAA
 43_25    AGTGGTGGGA CTTGAAACCT GGAGCCCCGA AACCCAAAGC CAACCAGCAA
  44_1    AGTGGTGGGA CTTGAAACCT GGAGCCCCGA AACCCAAAGC CAACCAGCAA
  44_5    AGTGGTGGGA CTTGAAACCT GGAGCCCCGA AACCCAAAGC CAACCAGCAA
223_10    .......... .......... .......... .......... ..........
 223_2    .......... .......... .......... .......... ..........
 223_4    .......... .......... .......... .......... ..........
 223_5    .......... .......... .......... .......... ..........
 223_6    .......... .......... .......... .......... ..........
 223_7    .......... .......... .......... .......... ..........
  A3_4    AGTGGTGGAA GCTCAAACCT GGCCCACCAC CGCCGAAACC TAACCAACAA
  A3_5    AGTGGTGGAA GCTCAAACCT GGCCCACCAC CGCCGAAACC TAACCAACAA
  A3_7    AGTGGTGGAA GCTCAAACCT GGCCCACCAC CGCCGAAACC TAACCAACAA
  A3_3    AGTGGTGGAA GCTCAAACCT GGCCCACCAC CGCCGAAACC TAACCAACAA
 42_12    AGTGGTGGGA CTTGAAACCT GGAGCCCCGA AACCCAAAGC CAACCAGCAA
  AAV1    AGTGGTGGGA CTTGAAACCT GGAGCCCCGA AGCCCAAAGC CAACCAGCAA
  AAV2    AGTGGTGGAA GCTCAAACCT GGCCCACCAC CACCAAAGCC CGCAGAGCGG
  AAV3    AGTGGTGGGC TCTGAAACCT GGAGTCCCTC AACCCAAAGC GAACCAACAA
  AAV8    AGTGGTGGGC GCTGAAACCT GGAGCCCCGA AGCCCAAAGC CAACCAGCAA
  AAV9    AGTGGTGGGC GCTGAAACCT GGAGCCCCGA AGCCCAAAGC CAACCAGCAA
  AAV7    AGTGGTGGGA CCTGAAACCT GGAGCCCCGA AACCCAAAGC CAACCAGCAA
  44_2    AGTGGTGGGA CTTGAAACCT GGAGCCCCGA AACCCAAAGC CAACCAGCAA
```

图1AU

```
          2351                                                      2400
  42_2    AAGCAGGACG ACGGCCGGGG TCTGGTGCTT CCTGGCTACA AGTACCTCGG
  42_8    AAGCAGGACG ACGGCCGGGG TCTGGTGCTT CCTGGCTACA AGTACCTCGG
 42_15    AAGCAGGACG ACGGCCGGGG TCTGGTGCTT CCTGGCTACA AGTACCTCGG
 42_5b    AAGCAGGACG ACGGCCGGGG TCTGGTGCTT CCTGGCTACA AGTACCTCGG
 42_1b    AAGCAGGACG ACGGCCGGGG TCTGGTGCTT CCTGGCTACA AGTACCTCGG
 42_13    AAGCAGGACG ACGGCCGGGG TCTGGTGCTT CCTGGCTACA AGTACCTCGG
 42_3a    AAGCAGGACG ACGGCCGGGG TCTGGTGCTT CCTGGCTACA AGTACCTCGG
  42_4    AAGCAGGACG ACGGCCGGGG TCTGGTGCTT CCTGGCTACA AGTACCTCGG
 42_5a    AAGCAGGACG ACGGCCGGGG TCTGGTGCTT CCTGGCTACA AGTACCTCGG
 42_10    AAGCAGGACG ACGGCCGGGG TCTGGTGCTT CCTGGCTACA AGTACCTCGG
 42_3b    AAGCAGGACG ACGGCCGGGG TCTGGTGCTT CCTGGCTACA AGTACCTCGG
 42_11    AAGCAGGACG ACGGCCGGGG TCTGGTGCTT CCTGGCTACA AGTACCTCGG
 42_6b    AAGCAGGACG ACGGCCGGGG TCTGGTGCTT CCTGGCTACA AGTACCTCGG
  43_1    AAGCAGGACG ACGGCCGGGG TCTGGTGCTT CCTGGCTACA AGTACCTCGG
  43_5    AAGCAGGACG ACGGCCGGGG TCTGGTGCTT CCTGGCTACA AGTACCTCGG
 43_12    AAGCAGGACG ACGGCCGGGG TCTGGTGCTT CCTGGCTACA AGTACCTCGG
 43_20    AAGCAGGACG ACGGCCGGGG TCTGGTGCTT CCTGGCTACA AGTACCTCGG
 43_21    AAGCAGGACG ACGGCCGGGG TCTGGTGCTT CCTGGCTACA AGTACCTCGG
 43_23    AAGCAGGACG ACGGCCGGGG TCTGGTGCTT CCTGGCTACA AGTACCTCGG
 43_25    AAGCAGGACG ACGGCCGGGG TCTGGTGCTT CCTGGCTACA AGTACCTCGG
  44_1    AAGCAGGACG ACGGCCGGGG TCTGGTGCTT CCTGGCTACA AGTACCTCGG
  44_5    AAGCAGGACG ACGGCCGGGG TCTGGTGCTT CCTGGCTACA AGTACCTCGG
223_10    .......... .......... .......... .......... ..........
 223_2    .......... .......... .......... .......... ..........
 223_4    .......... .......... .......... .......... ..........
 223_5    .......... .......... .......... .......... ..........
 223_6    .......... .......... .......... .......... ..........
 223_7    .......... .......... .......... .......... ..........
  A3_4    CACCGGGACG ACAGTAGGGG TCTTGTGCTT CCTGGGTACA AGTACCTCGG
  A3_5    CACCGGGACG ACAGTAGGGG TCTTGTGCTT CCTGGGTACA AGTACCTCGG
  A3_7    CACCGGGACG ACAGTAGGGG TCTTGTGCTT CCTGGGTACA AGTACCTCGG
  A3_3    CACCGGGACG ACAGTAGGGG TCTTGTGCTT CCTGGGTACA AGTACCTCGG
 42_12    AAGCAGGACG ACGGCCGGGG TCTGGTGCTT CCTGGCTACA AGTACCTCGG
  AAV1    AAGCAGGACG ACGGCCGGGG TCTGGTGCTT CCTGGCTACA AGTACCTCGG
  AAV2    CATAAGGACG ACAGCAGGGG TCTTGTGCTT CCTGGGTACA AGTACCTCGG
  AAV3    CACCAGGACA ACCGTCGGGG TCTTGTGCTT CCGGGTTACA AATACCTCGG
  AAV8    AAGCAGGACG ACGGCCGGGG TCTGGTGCTT CCTGGCTACA AGTACCTCGG
  AAV9    AAGCAGGACG ACGGCCGGGG TCTGGTGCTT CCTGGCTACA AGTACCTCGG
  AAV7    AAGCAGGACA ACGGCCGGGG TCTGGTGCTT CCTGGCTACA AGTACCTCGG
  44_2    AAGCAGGACG ACGGCCGGGG TCTGGTGCTT CCTGGCTACA AGTACCTCGG
```

图1AV

```
          2401                                                      2450
  42_2    ACCCTTCAAC GGACTCGACA AGGGAGAGCC GGTCAACGAG GCAGACGCCG
  42_8    ACCCTTCAAC GGACTCGACA AGGGGGAGCC CGTCAACGCG GCGGACGCAG
 42_15    ACCCTTCAAC GGACTCGACA AGGGGGAGCC CGTCAACGCG GCGGACGCAG
 42_5b    ACCCTTCAAC GGACTCGACA AGGGAGAGCC GGTCAACGAG GCAGACGCCG
 42_1b    ACCCTTCAAC GGACTCGACA AGGGAGAGCC GGTCAACGAG GCAGACGCCG
 42_13    ACCCTTCAAC GGACTCGACA AGGGGGAGCC CGTCAACGCG GCGGACGCAG
 42_3a    ACCCTTCAAC GGACTCGACA AGGGGGAGCC CGTCAACGCG GCGGACGCAG
  42_4    ACCCTTCAAC GGACTCGACA AGGGAGAGCC GGTCAACGAG GCAGACGCCG
 42_5a    ACCCTTCAAC GGACTCGACA AGGGAGAGCC GGTCAACGAG GCAGACGCCG
 42_10    ACCCTTCAAC GGACTCGACA AGGGAGAGCC GGTCAACGAG GCAGACGCCG
 42_3b    ACCCTTCAAC GGACTCGACA AGGGAGAGCC GGTCAACGAG GCAGACGCCG
 42_11    ACCCTTCAAC GGACTCGACA AGGGAGAGCC GGTCAACGCG GCGGACGCAG
 42_6b    ACCCTTCAAC GGACTCGACA AGGGAGAGCC GGTCAACGAG GCAGACGCCG
  43_1    ACCCTTCAAC GGACTCGACA AGGGGGAGCC CGTCAACGCG GCGGACGCAG
  43_5    ACCCTTCAAC GGACTCGACA AGGGGGAGCC CGTCAACGCG GCGGACGCAG
 43_12    ACCCTTCAAC GGACTCGACA AGGGGGAGCC CGTCAACGCG GCGGACGCAG
 43_20    ACCCTTCAAC GGACTCGACA AGGGGGAGCC CGTCAACGCG GCGGACGCAG
 43_21    ACCCTTCAAC GGACTCGACA AGGGGGAGCC CGTCAACGCG GCGGACGCAG
 43_23    ACCCTTCAAC GGACTCGACA AGGGGGAGCC CGTCAACGCG GCGGACGCAG
 43_25    ACCCTTCAAC GGACTCGACA AGGGGGAGCC CGTCAACGCG GCGGACGCAG
  44_1    ACCCTTCAAC GGACTCGACA AGGGGGAGCC CGTCAACGCG GCGGACGCAG
  44_5    ACCCTTCAAC GGACTCGACA AGGGGGAGCC CGTCAACGCG GCGGACGCAG
223_10    .......... .......... .......... .......... ..........
 223_2    .......... .......... .......... .......... ..........
 223_4    .......... .......... .......... .......... ..........
 223_5    .......... .......... .......... .......... ..........
 223_6    .......... .......... .......... .......... ..........
 223_7    .......... .......... .......... .......... ..........
  A3_4    ACCCTTCAAC GGACTCGACA AAGGAGAGCC GGTCAACGAG GCAGACGCCG
  A3_5    ACCCTTCAAC GGACTCGACA AAGGAGAGCC GGTCAACGAG GCAGACGCCG
  A3_7    ACCCTTCAAC GGACTCGACA AAGGAGAGCC GGTCAACGAG GCAGACGCCG
  A3_3    ACCCTTCAAC GGACTCGACA AAGGAGAGCC GGTCAACGAG GCAGACGCCG
 42_12    ACCCTTCAAC GGACTCGACA AGGGAGAGCC GGTCAACGAG GCAGACGCCG
  AAV1    ACCCTTCAAC GGACTCGACA AGGGGGAGCC CGTCAACGCG GCGGACGCAG
  AAV2    ACCCTTCAAC GGACTCGACA AGGGAGAGCC GGTCAACGAG GCAGACGCCG
  AAV3    ACCCGGTAAC GGACTCGACA AAGGAGAGCC GGTCAACGAG GCGGACGCGG
  AAV8    ACCCTTCAAC GGACTCGACA AGGGGGAGCC CGTCAACGCG GCGGACGCAG
  AAV9    ACCCTTCAAC GGACTCGACA AGGGGGAGCC CGTCAACGCG GCGGACGCAG
  AAV7    ACCCTTCAAC GGACTCGACA AGGGGGAGCC CGTCAACGCG GCGGACGCAG
  44_2    ACCCTTCAAC GGACTCGACA AGGGGGAGCC CGTCAACGCG GCGGACGCAG
```

图1AW

```
          2451                                                       2500
  42_2    CGGCCCTCGA GCACG.ACAA GGCCTACGAC AAGCAGCTCG AGCAGGGGGA
  42_8    CGGCCCTCGA GCACG.ACAA GGCCTACGAC CAGCAGCTCA AAGCGGGTGA
 42_15    CGGCCCTCGA GCACG.ACAA GGCCTACGAC CAGCAGCTCA AAGCGGGTGA
 42_5b    CGGCCCTCGA GCACG.ACAA GGCCTACGAC AAGCAGCTCG AGCAGGGGGA
 42_1b    CGGCCCTCGA GCACG.ACAA GGCCTACGAC AAGCAGCTCG AGCAGGGGGA
 42_13    CGGCCCTCGA GCACG.ACAA GGCCTACGAC CAGCAGCTCA AAGCGGGTGA
 42_3a    CGGCCCTCGA GCACG.ACAA GGCCTACGAC CAGCAGCTCA AAGCGGGTGA
  42_4    CGGCCCTCGA GCACG.ACAA GGCCTACGAC AAGCAGCTCG AGCAGGGGGA
 42_5a    CGGCCCTCGA GCACG.ACAA GGCCTACGAC AAGCAGCTCG AGCAGGGGGA
 42_10    CGGCCCTCGA GCACG.ACAA GGCCTACGAC AAGCAGCTCG AGCAGGGGGA
 42_3b    CGGCCCTCGA GCACG.ACAA GGCCTACGAC AAGCAGCTCG AGCAGGGGGA
 42_11    CGGCCCTCGA GCACG.ACAA GGCCTACGAC CAGCAGCTCA AAGCGGGTGA
 42_6b    CGGCCCTCGA GCACG.ACAA GGCCTACGAC AAGCAGCTCG AGCAGGGGGA
  43_1    CGGCCCTCGA GCACG.ACAA GGCCTACGAC CAGCAGCTCA AAGCGGGTGA
  43_5    CGGCCCTCGA GCACG.ACAA GGCCTACGAC CAGCAGCTCA AAGCGGGTGA
 43_12    CGGCCCTCGA GCACG.ACAA GGCCTACGAC CAGCAGCTCA AAGCGGGTGA
 43_20    CGGCCCTCGA GCACG.ACAA AGCCTACGAC CAGCAGCTCA AAGCGGGTGA
 43_21    CGGCCCTCGA GCACG.ACAA AGCCTACGAC CAGCAGCTCA AAGCGGGTGA
 43_23    CGGCCCTCGA GCACG.ACAA AGCCTACGAC CAGCAGCTCA AAGCGGGTGA
 43_25    CGGCCCTCGA GCACG.ACAA AGCCTACGAC CAGCAGCTCA AAGCGGGTGA
  44_1    CGGCCCTCGA GCACG.ACAA GGCCTACGAC CAGCAGCTCA AAGCGGGTGA
  44_5    CGGCCCTCGA GCACG.ACAA GGCCTACGAC CAGCAGCTCA AAGCGGGTGA
223_10    .......... .......CAA GGCCTACGAC CAGCAGCTCA AAGCGGGTGA
223_2     .......... .......CAA GGCCTACGAC CAGCAGCTCA AAGCGGGTGA
223_4     .......... .......CAA GGCCTACGAC CAGCAGCTCA AAGCGGGTGA
223_5     .......... .......CAA GGCCTACGAC CAGCAGCTCA AAGCGGGTGA
223_6     .......... .......CAA GGCCTACGAC CAGCAGCTCA AAGCGGGTGA
223_7     .......... .......CAA GGCCTACGAC CAGCAGCTCA AAGCGGGTGA
  A3_4    CGGCCCTCGA GCACG.ACAA AGCCTACGAC CACCAGCTCA AGCAAGGGGA
  A3_5    CGGCCCTCGA GCACG.ACAA AGCCTACGAC CACCAGCTCA AGCAAGGGGA
  A3_7    CGGCCCTCGA GCACG.ACAA AGCCTACGAC CACCAGCTCA AGCAAGGGGA
  A3_3    CGGCCCTCGA GCACG.ACAA AGCCTACGAC CACCAGCTCA AGCAAGGGGA
 42_12    CGGCCCTCGA GCACG.ACAA GGCCTACGAC AAGCAGCTCG AGCAGGGGGA
  AAV1    CGGCCCTCGA GCACG.ACAA GGCCTACGAC CAGCAGCTCA AAGCGGGTGA
  AAV2    CGGCCCTCGA GCACGTACAA AGCCTACGAC CGGCAGCTCG ACAGCGGAGA
  AAV3    CAGCCCTCGA ACACG.ACAA AGCTTACGAC CAGCAGCTCA AGGCCGGTGA
  AAV8    CGGCCCTCGA GCACG.ACAA GGCCTACGAC CAGCAGCTGC AGGCGGGTGA
  AAV9    CGGCCCTCGA GCACG.GCAA GGCCTACGAC CAGCAGCTGC AGGCGGGTGA
  AAV7    CGGCCCTCGA GCACG.ACAA GGCCTACGAC CAGCAGCTCA AAGCGGGTGA
  44_2    CGGCCCTCGA GCACG.ACAA GGCCTACGAC CAGCAGCTCA AAGCGGGTGA
```

图1AX

```
          2501                                                      2550
  42_2    CAACCCGTAC CTCAAGTACA ACCACGCCGA CGCCGAGTTT CAGGAGCGTC
  42_8    CAATCCGTAC CTGCGGTATA ACCACGCCGA CGCCGAGTTT CAGGAGCGTC
 42_15    CAATCCGTAC CTGCGGTATA ACCACGCCGA CGCCGAGTTT CAGGAGCGTC
 42_5b    CAACCCGTAC CTCAAGTACA ACCACGCCGA CGCCGAGTTT CAGGAGCGTC
 42_1b    CAACCCGTAC CTCAAGTACA ACCACGCCGA CGCCGAGTTT CAGGAGCGTC
 42_13    CAATCCGTAC CTGCGGTATA ACCACGCCGA CGCCGAGTTT CAGGAGCGTC
 42_3a    CAATCCGTAC CTGCGGTATA ACCACGCCGA CGCCGAGTTT CAGGAGCGTC
  42_4    CAACCCGTAC CTCAAGTACA ACCACGCCGA CGCCGAGTTT CAGGAGCGTC
 42_5a    CAACCCGTAC CTCAAGTACA ACCACGCCGA CGCCGAGTTT CAGGAGCGTC
 42_10    CAACCCGTAC CTCAAGTACA ACCACGCCGA CGCCGAGTTT CAGGAGCGTC
 42_3b    CAACCCGTAC CTCAAGTACA ACCACGCCGA CGCCGAGTTT CAGGAGCGTC
 42_11    CAATCCGTAC CTGCGGTATA ACCACGCCGA CGCCGAGTTT CAGGAGCGTC
 42_6b    CAACCCGTAC CTCAAGTACA ACCACGCCGA CGCCGAGTTT CAGGAGCGTC
  43_1    CAATCCGTAC CTGCGGTATA ACCACGCCGA CGCCGAGTTT CAGGAGCGTC
  43_5    CAATCCGTAC CTGCGGTATA ACCACGCCGA CGCCGAGTTT CAGGAGCGTC
 43_12    CAATCCGTAC CTGCGGTATA ACCACGCCGA CGCCGAGTTT CAGGAGCGTC
 43_20    CAATCCGTAC CTGCGGTATA ATCACGCCGA CGCCGAGTTT CAGGAGCGTC
 43_21    CAATCCGTAC CTGCGGTATA ATCACGCCGA CGCCGAGTTT CAGGAGCGTC
 43_23    CAATCCGTAC CTGCGGTATA ATCACGCCGA CGCCGAGTTT CAGGAGCGTC
 43_25    CAATCCGTAC CTGCGGTATA ATCACGCCGA CGCCGAGTTT CAGGAGCGTC
  44_1    CAATCCGTAC CTGCGGTATA ACCACGCCGA CGCCGAGTTT CAGGAGCGTC
  44_5    CAATCCGTAC CTGCGGTATA ACCACGCCGA CGCCGAGTTT CAGGAGCGTC
223_10    CAATCCGTAC CTGCGGTATA ACCACGCCGA CGCCGAGTTT CAGGAGCGTC
 223_2    CAATCCGTAC CTGCGGTATA ACCACGCCGA CGCCGAGTTT CAGGAGTGTC
 223_4    CAATCCGTAC CTGCGGTATA ACCACGCCGA CGCCGAGTTT CAGGAGCGTC
 223_5    CAATCCGTAC CTGCGGTATA ACCACGCCGA CGCCGAGTTT CAGGAGCGTC
 223_6    CAATCCGTAC CTGCGGTATA ACCACGCCGA CGCCGAGTTT CAGGAGCGTC
 223_7    CAATCCGTAC CTGCGGTATA ACCACGCCGA CGCCGAGTTT CAGGAGCGTC
  A3_4    CAACCCGTAC CTCAAATACA ACCACGCGGA CGCTGAATTT CAGGAGCGTC
  A3_5    CAACCCGTAC CTCAAATACA ACCACGCGGA CGCTGAATTT CAGGAGCGTC
  A3_7    CAACCCGTAC CTCAAATACA ACCACGCGGA CGCTGAATTT CAGGAGCGTC
  A3_3    CAACCCGTAC CTCAAATACA ACCACGCGGA CGCTGAATTT CAGGAGCGTC
 42_12    CAACCCGTAC CTCAAGTACA ACCACGCCGA CGCCGAGTTT CAGGAGCGTC
  AAV1    CAATCCGTAC CTGCGGTATA ACCACGCCGA CGCCGAGTTT CAGGAGCGTC
  AAV2    CAACCCGTAC CTCAAGTACA ACCACGCCGA CGCGGAGTTT CAGGAGCGCC
  AAV3    CAACCCGTAC CTCAAGTACA ACCACGCCGA CGCCGAGTTT CAGGAGCGTC
  AAV8    CAATCCGTAC CTGCGGTATA ACCACGCCGA CGCCGAGTTT CAGGAGCGTC
  AAV9    CAATCCGTAC CTGCGGTATA ACCACGCCGA CGCCGAGTTT CAGGAGCGTC
  AAV7    CAATCCGTAC CTGCGGTATA ACCACGCCGA CGCCGAGTTT CAGGAGCGTC
  44_2    CAATCCGTAC CTGCGGTATA ACCACGCCGA CGCCGAGTTT CAGGAGCGTC
```

图1AY

```
          2551                                                      2600
  42_2    TTCAAGAAGA TACGTCTTTT GGGGGCAACC TCGGGCGAGC AGTCTTCCAG
  42_8    TGCAAGAAGA TACGTCTTTT GGGGGCAACC TCGGGCGAGC AGTCTTCCAG
 42_15    TGCAAGAAGA TACGTCTTTT GGGGGCAACC TCGGGCGAGC AGTCTTCCAG
 42_5b    TTCAAGAAGA TACGTCTTTT GGGGGCAACC TCGGGCGAGC AGTCTTCCAG
 42_1b    TTCAAGAAGA TACGTCTTTT GGGGGCAACC TCGGGCGAGC AGTCTTCCAG
 42_13    TTCAAGAAGA TACGTCTTTT GGGGGCAACC TCGGGCGAGC AGTCTTCCAG
 42_3a    TTCAAGAAGA TACGTCTTTT GGGGGCAACC TCGGGCGAGC AGTCTTCCAG
  42_4    TTCAAGAAGA TACGTCTTTT GGGGGCAACC TCGGGCGAGC AGTCTTCCAG
 42_5a    TTCAAGAAGA TACGTCTTTT GGGGGCAACC TCGGGCGAGC AGTCTTCCGG
 42_10    TTCAAGAAGA TACGTCTTTT GGGGGCAACC TCGGGCGAGC AGTCTTCCAG
 42_3b    TTCAAGAAGA TACGTCTTTT GGGGGCAACC TCGGGCGAGC AGTCTTCCAG
 42_11    TTCAAGAAGA TACGTCTTTT GGGGGCAACC TCGGGCGAGC AGTCTTCCAG
 42_6b    TTCAAGAAGA TACGTCTTTT GGGGGCAACC TCGGGCGAGC AGTCTTCCAG
  43_1    TGCAAGAAGA TACGTCTTTT GGGGGCAACC TCGGGCGAGC AGTCTTCCAG
  43_5    TGCAAGAAGA TACGTCTTTT GGGGGCAACC TCGGGCGAGC AGTCTTCCAG
 43_12    TGCAAGAAGA TACGTCTTTT GGGGGCAACC TCGGGCGAGC AGTCTTCCAG
 43_20    TGCAAGAAGA TACGTCTTTT GGGGGCAACC TCGGGCGAGC AGTCTTCCAG
 43_21    TGCAAGAAGA TACGTCTTTT GGGGGCAACC TCGGGCGAGC AGTCTTCCAG
 43_23    TGCAAGAAGA TACGTCCTTT GGGGGCAACC TCGGGCGAGC AGTCTTCCAG
 43_25    TGCAAGAAGA TACGTCTTTT GGGGGCAACC TCGGGCGAGC AGTCTTCCAG
  44_1    TGCAAGAAGA TACGTCTTTT GGGGGCAACC TCGGGCGAGC AGTCTTCCAG
  44_5    TGCAAGAAGA TACGTCTTTT GGGGGCAACC TCGGGCGAGC AGTCTTCCAG
223_10    TTCAAGAAGA TACGTCTTTT GGGGGCAACC TCGGGCGAGC AGTCTTCCAG
 223_2    TTCAAGAAGA TACGTCTTTT GGGGGCAACC TCGGGCGAGC AGTCTTCCAG
 223_4    TTCAAGAAGA TACGTCTTTT GGGGGCAACC TCGGGCGAGC AGTCTTCCAG
 223_5    TTCAAGAAGA TACGTCTTTT GGGGGCAACC TCGGGCGAGC AGTCTTCCAG
 223_6    TTCAAGAAGA TACGTCTTTT GGGGGCAACC TCGGGCGAGC AGTCTTCCAG
 223_7    TTCAAGAAGA TACGTCTTTT GGGGGCAACC TCGGGCGAGC AGTCTTCCAG
  A3_4    TTCAAGAAGA TACGTCTTTC GGGGGCAACC TCGGGCGAGC AGTCTTCCAG
  A3_5    TTCAAGAAGA TACGTCTTTC GGGGGCAACC TCGGGCGAGC AGTCTTCCAG
  A3_7    TTCAAGAAGA TACGTCTTTC GGGGGCAACC TCGGGCGAGC AGTCTTCCAG
  A3_3    TTCAAGAAGA TACGTCTTTC GGGGGCAACC TCGGGCGAGC AGTCTTCCAG
 42_12    TTCAAGAAGA TACGTCTTTT GGGGGCAACC TCGGGCGAGC AGTCTTCCAG
   AAV1   TGCAAGAAGA TACGTCTTTT GGGGGCAACC TCGGGCGAGC AGTCTTCCAG
   AAV2   TTAAAGAAGA TACGTCTTTT GGGGGCAACC TCGGACGAGC AGTCTTCCAG
   AAV3   TTCAAGAAGA TACGTCTTTT GGGGGCAACC TTGGCAGAGC AGTCTTCCAG
   AAV8   TGCAAGAAGA TACGTCTTTT GGGGGCAACC TCGGGCGAGC AGTCTTCCAG
   AAV9   TGCAAGAAGA TACGTCTTTT GGGGGCAACC TCGGGCGAGC AGTCTTCCAG
   AAV7   TGCAAGAAGA TACGTCATTT GGGGGCAACC TCGGGCGAGC AGTCTTCCAG
  44_2    TGCAAGAAGA TACGTCTTTT GGGGGCAACC TCGGGCGAGC AGTCTTCCAG
```

图1AZ

```
          2601                                                      2650
  42_2    GCCAAGAAGC GGGTTCTCGA ACCTCTCGGT CTGGTTGAGG AAGGCGCTAA
  42_8    GCCAAGAAGC GGGTTCTCGA ACCTCTCGGT CTGGTTGAGG AAGGCGCTAA
 42_15    GCCAAGAAGC GGGTTCTCGA ACCTCTCGGT CTGGTTGAGG AAGGCGCTAA
 42_5b    GCCAAGAAGC GGGTTCTCGA ACCTCTCGGT CTGGTTGAGG AAGGCGCTAA
 42_1b    GCCAAGAAGC GGGTTCTCGA ACCTCTCGGT CTGGTTGAGG AAGGCGCTAA
 42_13    GCCAAGAAGC GGGTTCTCGA ACCTCTCGGT CTGGTTGAGG AAGGCGCTAA
 42_3a    GCCAAGAAGC GGGTTCTCGA ACCTCTCGGT CTGGTTGAGG AAGGCGCTAA
  42_4    GCCAAGAAGC GGGTTCTCGA ACCTCTCGGT CTGGTTGAGG AAGGCGCTAA
 42_5a    GCCAAGAAGC GGGTTCTCGA ACCTCTCGGT CTGGTTGAGG AAGGCGCTAA
 42_10    GCCAAGAAGC GGGTTCTCGA ACCTCTCGGT CTGGTTGAGG AAGGCGCTAA
 42_3b    GCCAAGAAGC GGGTTCTCGA ACCTCTCGGT CTGGTTGAGG AAGGCGCTAA
 42_11    GCCAAGAAGC GGGTTCTCGA ACCTCTCGGT CTGGTTGAGG AAGGCGCTAA
 42_6b    GCCAAGAAGC GGGTTCTCGA ACCTCTCGGT CTGGTTGAGG AAGGCGCTAA
  43_1    GCCAAGAAGC GGGTTCTCGA ACCTCTCGGT CTGGTTGAGG AAGGCGCTAA
  43_5    GCCAAGAAGC GGGTTCTCGA ACCTCTCGGT CTGGTTGAGG AAGGCGCTAA
 43_12    GCCAAGAAGC GGGTTCTCGA ACCTCTCGGT CTGGTTGAGG AAGGCGCTAA
 43_20    GCCAAGAAGC GGGTTCTCGA ACCTCTCGGT CTGGTTGAGG AAGGCGCTAA
 43_21    GCCAAGAAGC GGGTTCTCGA ACCTCTCGGT CTGGTTGAGG AAGGCGCTAA
 43_23    GCCAAGAAGC GGGTTCTCGA ACCTCTCGGT CTGGTTGAGG AAGGCGCTAA
 43_25    GCCAAGAAGC GGGTTCTCGA ACCTCTCGGT CTGGTTGAGG AAGGCGCTAA
  44_1    GCCAAGAAGC GGGTTCTCGA ACCTCTCGGT CTGGTTGAGG AAGGCGCTAA
  44_5    GCCAAGAAGC GGGTTCTCGA ACCTCTCGGT CTGGTTGAGG AAGGCGCTAA
223_10    GCCAAAAAGC GGGTTCTCGA ACCTCTTGGT CTGGTTGAGA CGCCAGCTAA
 223_2    GCCAAAAAGC GGGTTCTCGA ACCTCTTGGT CTGGTTGAGA CGCCAGCTAA
 223_4    GCCAAAAAGC GGGTTCTCGA ACCTCTTGGT CTGGTTGAGA CGCCAGCTAA
 223_5    GCCAAAAAGC GGGTTCTCGA ACCTCTTGGT CTGGTTGAGA CGCCAGCTAA
 223_6    GCCAAAAAGC GGGTTCTCGA ACCTCTTGGT CTGGTTGAGA CGCCAGCTAA
 223_7    GCCAAAAAGC GGGTTCTCGA ACCTCTTGGT CTGGTTGAGA CGCCAGCTAA
  A3_4    GCCAAAAAGA GGGTACTCGA GCCTCTTGGT CTGGTTGAGG AAGCTGTTAA
  A3_5    GCCAAAAAGA GGGTACTCGA GCCTCTTGGT CTGGTTGAGG AAGCTGTTAA
  A3_7    GCCAAAAAGA GGGTACTCGA GCCTCTTGGT CTGGTTGAGG AAGCTGTTAA
  A3_3    GCCAAAAAGA GGGTACTCGA GCCTCTTGGT CTGGTTGAGG AAGCTGTTAA
 42_12    GCCAAGAAGC GGGTTCTCGA ACCTCTCGGT CTGGTTGAGG AAGGCGCTAA
   AAV1   GCCAAGAAGC GGGTTCTCGA ACCTCTCGGT CTGGTTGAGG AAGGCGCTAA
   AAV2   GCGAAAAAGA GGGTTCTTGA ACCTCTGGGC CTGGTTGAGG AACCTGTTAA
   AAV3   GCCAAAAAGA GGATCCTTGA GCCTCTTGGT CTGGTTGAGG AAGCAGCTAA
   AAV8   GCCAAGAAGC GGGTTCTCGA ACCTCTCGGT CTGGTTGAGG AAGGCGCTAA
   AAV9   GCCAAGAAGC GGGTTCTCGA ACCTCTCGGT CTGGTTGAGG AAGGCGCTAA
   AAV7   GCCAAGAAGC GGGTTCTCGA ACCTCTCGGT CTGGTTGAGG AAGGCGCTAA
  44_2    GCCAAGAAGC GGGTTCTCGA ACCTCTCGGT CTGGTTGAGG AAGGCGCTAA
```

图1AAA

```
          2651                                                2700

          vp2 起始
  42_2    GACGGCTCCT GGAAAGAAGA GACCCATAGA ...ATCCCCC ..........
  42_8    GACGGCTCCT GGAAAGAAGA GACCGGTAGA GCCATCACCC CAGCGTTCTC
 42_15    GACGGCTCCT GGAAAGAAGA GACCGGTAGA GCCATCACCC CAGCGTTCTC
 42_5b    GACGGCTCCT GGAAAGAAGA GACCGGTAGA GCCATCACCC CAGCGTTCTC
 42_1b    GACGGCTCCT GGAAAGAAGA GACCCATAGA ..,ATCCCCC ..........
 42_13    GACGGCTCCT GGAAAGAAGA GACCCATAGA ...ATCCCCC ..........
 42_3a    GACGGCTCCT GGAAAGAAGA GACCCATAGA ..,ATCCCCC ..........
  42_4    GACGGCTCCT GGAAAGAAGA GACCCATAGA ...ATCCCCC ..........
 42_5a    GACGGCTCCT GGAAAGAAGA GACCCATAGA ...ATCCCCC ..........
 42_10    GACGGCTCCT GGAAAGAAGA GACCCATAGA ...ATCCCCC ..........
 42_3b    GACGGCTCCT GGAAAGAAGA GACCCATAGA ...ATCCCCC ..........
 42_11    GACGGCTCCT GGAAAGAAGA GACCCATAGA ...ATCCCCC ..........
 42_6b    GACGGCTCCT GGAAAGAAGA GACCGGTAGA GCCATCACCC CAGCGTTCTC
  43_1    GACGGCTCCT GGAAAGAAGA GACCGGTAGA GCCATCACCT CAGCGTTCCC
  43_5    GACGGCTCCT GGAAAGAAGA GACCGGTAGA GCCATCACCT CAGCGTTCCC
 43_12    GACGGCTCCT GGAAAGAAGA GACCGGTAGA GCCATCACCT CAGCGTTCCC
 43_20    GACGGCTCCT GGAAAGAAGA GACTGGTAGA GCAGTCGCCA CAAGAG...C
 43_21    GACGGCTCCT GGAAAGAAGA GACCGGTAGA GCAGTCGCCA CAAGAG...C
 43_23    GACGGCTCCT GGAAAGAAGA GACCGGTAGA GCAGTCGCCA CAAGAG...C
 43_25    GACGGCTCCT GGAAAGAAGA GACCGGTAGA GCAGTCGCCA CAAGAG...C
  44_1    GACGGCTCCT GGAAAGAAGA GACCGGTAGA GCCATCACCC CAGCGTTCTC
  44_5    GACGGCTCCT GGAAAGAAGA GACCGGTAGA GCCATCACCC CAGCGTTCTC
223_10    GACGGCACCT GGAAAGAAGC GACCGGTAGA CTCGCCA... ..........
 223_2    GACGGCACCT GGAAAGAAGC GACCGGTAGA CTCGCCA... ..........
 223_4    GACGGCACCT GGAAAGAAGC GACCGGTAGA CTCGCCA... ..........
 223_5    GACGGCACCT GGAAAGAAGC GACCGGTAGA CTCGCCA... ..........
 223_6    GACGGCACCT GGAAAGAAGC GACCGGTAGA CTCGCCA... ..........
 223_7    GACGGCACCT GGAAAGAAGC GACCGGTAGA CTCGCCA... ..........
  A3_4    GACGGCTCCT GGAAAAAAGA GACCTATAGA GCAGTCTCCT GCAGAA...C
  A3_5    GACGGCTCCT GGAAAAAAGA GACCTATAGA GCAGTCTCCT GCAGAA...C
  A3_7    GACGGCTCCT GGAAAAAAGA GACCTATAGA GCAGTCTCCT GCAGAA...C
  A3_3    GACGGCTCCT GGAAAAAAGA GACCTATAGA GCAGTCTCCT GCAGAA...C
 42_12    GACGGCTCCT GGAAAGAAGA GACCGGTAGA GCCATCACCC CAGCGTTCTC
  AAV1    GACGGCTCCT GGAAAGAAAC GTCCGGTAGA GCAGTCGCCA CAAGAG...C
  AAV2    GACGGCTCCG GGAAAAAAGA GGCCGGTAGA GCACTCTCCT GTGGAG...C
  AAV3    AACGGCTCCT GGAAAGAAGG GGGCTGTAGA TCAGTCTCCT CAGGAA...C
  AAV8    GACGGCTCCT GGAAAGAAGA GACCGGTAGA GCCATCACCC CAGCGTTCTC
  AAV9    GACGGCTCCT GGAAAGAAGA GACCGGTAGA GCCATCACCC CAGCGTTCTC
  AAV7    GACGGCTCCT GCAAAGAAGA GACCGGTAGA GCCGTCACCT CAGCGTTCCC
  44_2    GACGGCTCCT GGAAAGAAGA GACCGGTAGA GCCATCACCC CAGCGTTCTC
          vp2 起始
```

图1AAB

```
          2701                                                      2750
  42_2    ..GACTCCTC CACGGGCATC GGCAAGAAAG GCCAGCAGCC CGCTAAAAAG
  42_8    CAGACTCCTC TACGGGCATC GGCAAGACAG GCCAGCAGCC CGCGAAAAAG
 42_15    CAGACTCCTC TACGGGCATC GGCAAGACAG GCCAGCAGCC CGCGAAAAAG
 42_5b    CAGACTCCTC TACGGGCATC GGCAAGACAG GCCAGCAGCC CGCGAAAAAG
 42_1b    ..GACTCCTC CACGGGCATC GGCAAGAAAG GCCAGCAGCC CGCTAAAAAG
 42_13    ..GACTCCTC CACGGGCATC GGCAAGAAAG GCCAGCAGCC CGCTAAAAAG
 42_3a    ..GACTCCTC CACGGGCATC GGCAAGAAAG GCCAGCAGCC CGCTAAAAAG
  42_4    ..GACTCCTC CACGGGCATC GGCAAGAAAG GCCAGCAGCC CGCTAAAAAG
 42_5a    ..GACTCCTC CACGGGCATC GGCAAGAAAG GCCAGCAGCC CGCTAAAAAG
 42_10    ..GACTCCTC CACGGGCATC GGCAGGAAAG GCCAGCAGCC CGCTAAAAAG
 42_3b    ..GACTCCTC CACGGGCATC GGCAAGAAAG GCCAGCAGCC CGCTAAAAAG
 42_11    ..GACTCCTC CACGGGCATC GGCAAGAAAG GCCAGCAGCC CGCTAAAAAG
 42_6b    CAGACTCCTC TACGGGCATC GGCAAGACAG GCCAGCAGCC CGCGAAAAAG
  43_1    CCGACTCCTC CACGGGCATC GGCAAGAAAG GCCACCAGCC CGCGAGAAAG
  43_5    CCGACTCCTC CACGGGCATC GGCAAGAAAG GCCACCAGCC CGCGAGAAAG
 43_12    CCGACTCCTC CACGGGCATC GGCAAGAAAG GCCACCAGCC CGCGAGAAAG
 43_20    CAGACTCCTC CTCGGGCATC GGCAAGACAG GCCAGCAGCC CGCTAAAAAG
 43_21    CAGACTCCTC CTCGGGCATC GGCAAGACAG GCCAGCAGCC CGCTAAAAAG
 43_23    CAGACTCCTC CTCGGGCATC GGCAAGACAG GCCAGCAGCC CGCTAAAAAG
 43_25    CAGACTCCTC CTCGGGCATC GGCAAGACAG GCCAGCAGCC CGCTAAAAAG
  44_1    CAGACTCCTC TACGGGCATC GGCAAGAAAG GCCAGCAGCC CGCGAAAAAG
  44_5    CAGACTCCTC TACGGGCATC GGCAAGAAAG GCCAGCAGCC CGCGAAAAAG
223_10    ..GACTCCAC CTCGGGCATC GGCAAGAAAG GCCAGCAGCC CGCGAAAAAG
 223_2    ..GACTCCAC CTCGGGCATC GGCAAGAAAG GCCAGCAGCC CGCGAAAAAG
 223_4    ..GACTCCAC CTCGGGCATC GGCAAGAAAG GCCAGCAGCC CGCGAAAAAG
 223_5    ..GACTCCAC CTCGGGCATC GGCAAGAAAG GCCAGCAGCC CGCGAAAAAG
 223_6    ..GACTCCAC CTCGGGCATC GGCAAGAAAG GCCAGCAGCC CGCGAAAAAG
 223_7    ..GACTCCAC CTCGGGCATC GGCAAGAAAG GCCAGCAGCC CGCGAAAAAG
  A3_4    CGGACTCTTC CTCGGGCATC GGCGAATCAG GCCAGCAGCC CGCTAAGAAA
  A3_5    CGGACTCTTC CTCGGGCATC GGCAAATCAG GCCAGCAGCC CGCTAAGAAA
  A3_7    CGGACTCTTC CTCGGGCATC GGCAAATCAG GCCAGCAGCC CGCTAAGAAA
  A3_3    CGGACTCTTC CTCGGGCATC GGCAAATCAG GCCAGCAGCC CGCTAAGAAA
 42_12    CAGACTCCTC TACGGGCATC GGCAAGACAG GCCAGCAGCC CGCGAAAAAG
  AAV1    CAGACTCCTC CTCGGGCATC GGCAAGACAG GCCAGCAGCC CGCTAAAAAG
  AAV2    CAGACTCCTC CTCGGGAACC GGAAAGGCGG GCCAGCAGCC TGCAAGAAAA
  AAV3    CGGACTCATC ATCTGGTGTT GGCAAATCGG GCAAACAGCC TGCCAGAAAA
  AAV8    CAGACTCCTC TACGGGCATC GGCAAGAAAG GCCAACAGCC CGCCAGAAAA
  AAV9    CAGACTCCTC TACGGGCATC GGCAAGAAAG GCCAACAGCC CGCCAGAAAA
  AAV7    CCGACTCCTC CACGGGCATC GGCAAGAAAG GCCAGCAGCC CGCCAGAAAG
  44_2    CAGACTCCTC TACGGGCATC GGCAAGAAAG GCCAGCAGCC CGCGAAAAAG
```

图1AAC

```
          2751                                                      2800
  42_2    AAGCTCAACT TTGGGCAGAC TGGCGACTCA GAGTCAGTGC CCGACCCCCA
  42_8    AGACTCAACT TTGGGCAGAC TGGCGACTCA GAGTCAGTGC CCGACCCTCA
 42_15    AGACTCAACT TTGGGCAGAC TGGCGACTCA GAGTCAGTGC CCGACCCTCA
 42_5b    AGACTCAACT TTGGGCAGAC TGGCGACTCA GAGTCAGTGC CCGACCCTCA
 42_1b    AGACTCAACT TTGGGCAGAC TGGCGACTCA GAGTCAGTGC CCGACCCTCA
 42_13    AAGCTCAACT TTGGGCAGAC TGGCGACTCA GAGTCAGTGC CCGACCCTCA
 42_3a    AAGCTCAACT TTGGGCAGAC TGGCGACTCA GAGTCAGTGC CCGACCCTCA
  42_4    AAGCTCAACT TTGGGCAGAC TGGCGACTCA GAGTCAGTGC CCGACCCTCA
 42_5a    AAGCTCAACT TTGGGCAGAC TGGCGACTCA GAGTCAGTGC CCGACCCCCA
 42_10    AAGCTCAACT TTGGGCAGAC TGGCGACTCA GAGTCAGTGC CCGACCCTCA
 42_3b    AAGCTCAACT TTGGGCAGAC TGGCGACTCA GAGTCAGTGC CCGACCCTCA
 42_11    AAGCTCAACT TTGGGCAGAC TGGCGACTCA GAGTCAGTGC CCGACCCTCA
 42_6b    AGACTCAACT TTGGGCAGAC TGGCGACTCA GAGTCAGTGC CCGACCCTCA
  43_1    AGACTGAACT TTGGGCAGAC TGGCGACTCG GAGTCAGTCC CCGACCCTCA
  43_5    AGACTGAACT TTGGGCAGAC TGGCGACTCG GAGTCAGTCC CCGACCCTCA
 43_12    AGACTGAACT TTGGGCAGAC TGGCGACTCG GAGTCAGTCC CCGACCCTCA
 43_20    AGACTCAATT TTGGTCAGAC TGGCGACTCA GAGTCAGTCC CCGACCCACA
 43_21    AGACTCAATT TTGGTCAGAC TGGCGACTCA GAGTCAGTCC CCGACCCACA
 43_23    AGACTCAATT TTGGTCAGAC TGGCGACTCA GAGTCAGTCC CCGACCCACA
 43_25    AGACTCAATT TTGGTCAGAC TGGCGACTCA GAGTCAGTCC CCGACCCACA
  44_1    AGACTCAACT TTGGGCAGAC TGGCGACTCA GAGTCAGTGC CCGACCCTCA
  44_5    AGACTCAACT TTGGGCAGAC TGGCGACTCA GAGTCAGTGC CCGACCCTCA
223_10    AGACTCAACT TTGGGCAGAC TGGCGACTCA GAGTCAGTCC CCGACCCTCA
 223_2    AGACTCAACT TTGGGCAGAC TGGCGACTCA GAGTCAGTCC CCGACCCTCA
 223_4    AGACTCAACT TTGGGCAGAC TGGCGACTCA GAGCCAGTCC CCGACCCTCA
 223_5    AGACTCAACT TTGGGCAGAC TGGCGACTCA GAGCCAGTCC CCGACCCTCA
 223_6    AGACTCAACT TTGGGCAGAC TGGCGACTCA GAGTCAGTCC CCGACCCTCA
 223_7    AGACTCAACT TTGGGCAGAC TGGCGACTCA GAGTCAGTCC CCGACCCTCA
  A3_4    AGACTCAATT TTGGTCAGAC TGGCGACACA GAGTCAGTCC CAGACCCTCA
  A3_5    AGACTCAATT TTGGTCAGAC TGGCGACACA GAGTCAGTCC CAGACCCTCA
  A3_7    AGACTCAATT TTGGTCAGAC TGGCGACACA GAGTCAGTCC CAGACCCTCA
  A3_3    AGACTCAATT TTGGTCAGAC TGGCGACACA GAGTCAGTCC CAGGCCCTCA
 42_12    AGACTCAACT TTGGGCAGAC TGGCGACTCA GAGTCAGTGC CCGACCCTCA
  AAV1    AGACTCAATT TTGGTCAGAC TGGCGACTCA GAGTCAGTCC CCGATCCACA
  AAV2    AGATTGAATT TTGGTCAGAC TGGAGACGCA GACTCAGTAC CTGACCCCCA
  AAV3    AGACTAAATT TCGGTCAGAC TGGAGACTCA GAGTCAGTCC CAGACCCTCA
  AAV8    AGACTCAATT TTGGTCAGAC TGGCGACTCA GAGTCAGTTC CAGACCCTCA
  AAV9    AGACTCAATT TTGGTCAGAC TGGCGACTCA GAGTCAGTTC CAGACCCTCA
  AAV7    AGACTCAATT TCGGTCAGAC TGGCGACTCA GAGTCAGTCC CCGACCCTCA
  44_2    AGACTCAACT TTGGGCAGAC TGGCGACTCA GAGTCAGTGC CCGACCCTCA
```

图1AAD

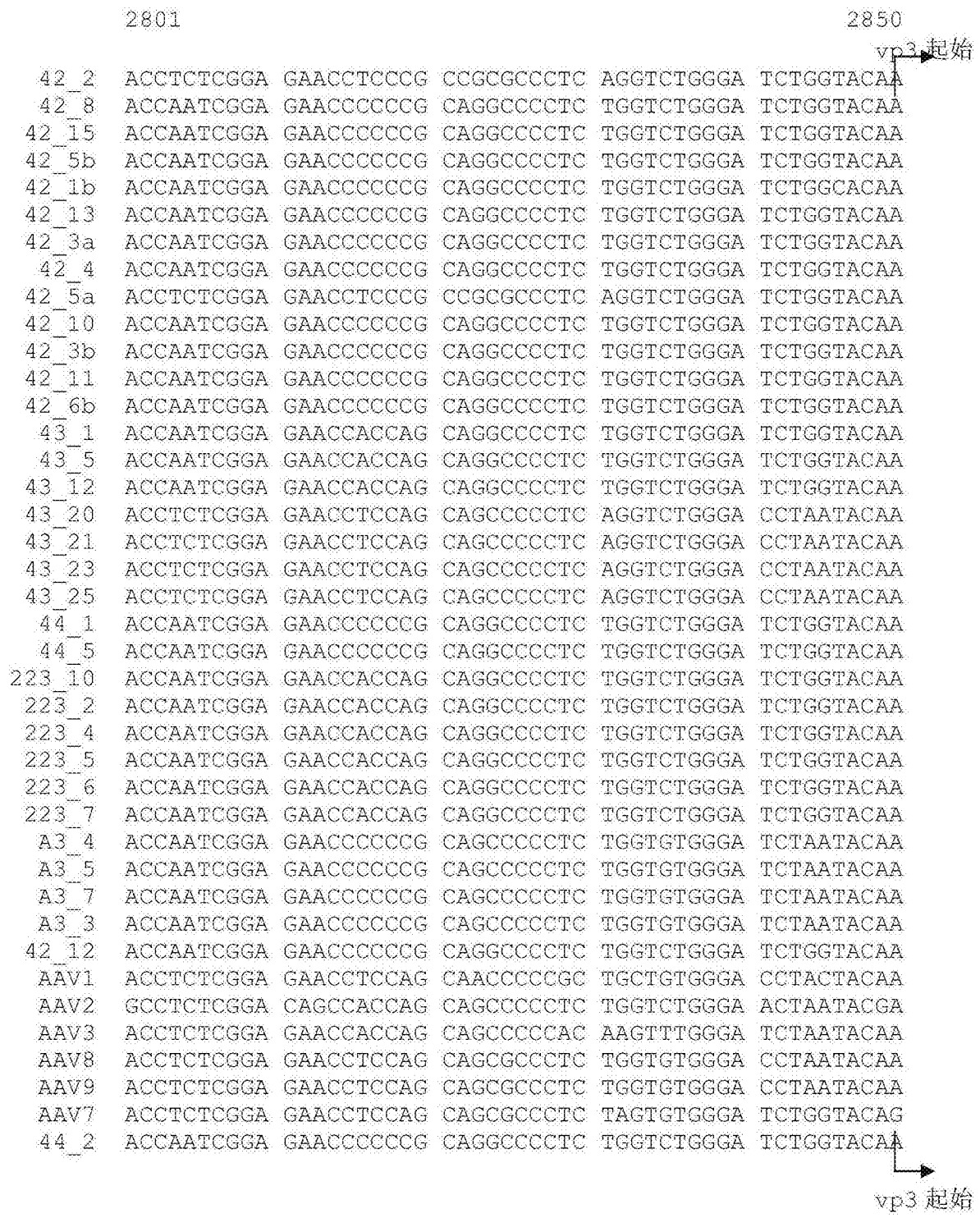

2801
2850
vp3 起始
42_2 ACCTCTCGGA GAACCTCCCG CCGCGCCCTC AGGTCTGGGA TCTGGTACAA
42_8 ACCAATCGGA GAACCCCCCG CAGGCCCCTC TGGTCTGGGA TCTGGTACAA
42_15 ACCAATCGGA GAACCCCCCG CAGGCCCCTC TGGTCTGGGA TCTGGTACAA
42_5b ACCAATCGGA GAACCCCCCG CAGGCCCCTC TGGTCTGGGA TCTGGTACAA
42_1b ACCAATCGGA GAACCCCCCG CAGGCCCCTC TGGTCTGGGA TCTGGCACAA
42_13 ACCAATCGGA GAACCCCCCG CAGGCCCCTC TGGTCTGGGA TCTGGTACAA
42_3a ACCAATCGGA GAACCCCCCG CAGGCCCCTC TGGTCTGGGA TCTGGTACAA
42_4 ACCAATCGGA GAACCCCCCG CAGGCCCCTC TGGTCTGGGA TCTGGTACAA
42_5a ACCTCTCGGA GAACCTCCCG CCGCGCCCTC AGGTCTGGGA TCTGGTACAA
42_10 ACCAATCGGA GAACCCCCCG CAGGCCCCTC TGGTCTGGGA TCTGGTACAA
42_3b ACCAATCGGA GAACCCCCCG CAGGCCCCTC TGGTCTGGGA TCTGGTACAA
42_11 ACCAATCGGA GAACCCCCCG CAGGCCCCTC TGGTCTGGGA TCTGGTACAA
42_6b ACCAATCGGA GAACCCCCCG CAGGCCCCTC TGGTCTGGGA TCTGGTACAA
43_1 ACCAATCGGA GAACCACCAG CAGGCCCCTC TGGTCTGGGA TCTGGTACAA
43_5 ACCAATCGGA GAACCACCAG CAGGCCCCTC TGGTCTGGGA TCTGGTACAA
43_12 ACCAATCGGA GAACCACCAG CAGGCCCCTC TGGTCTGGGA TCTGGTACAA
43_20 ACCTCTCGGA GAACCTCCAG CAGCCCCCTC AGGTCTGGGA CCTAATACAA
43_21 ACCTCTCGGA GAACCTCCAG CAGCCCCCTC AGGTCTGGGA CCTAATACAA
43_23 ACCTCTCGGA GAACCTCCAG CAGCCCCCTC AGGTCTGGGA CCTAATACAA
43_25 ACCTCTCGGA GAACCTCCAG CAGCCCCCTC AGGTCTGGGA CCTAATACAA
44_1 ACCAATCGGA GAACCCCCCG CAGGCCCCTC TGGTCTGGGA TCTGGTACAA
44_5 ACCAATCGGA GAACCCCCCG CAGGCCCCTC TGGTCTGGGA TCTGGTACAA
223_10 ACCAATCGGA GAACCACCAG CAGGCCCCTC TGGTCTGGGA TCTGGTACAA
223_2 ACCAATCGGA GAACCACCAG CAGGCCCCTC TGGTCTGGGA TCTGGTACAA
223_4 ACCAATCGGA GAACCACCAG CAGGCCCCTC TGGTCTGGGA TCTGGTACAA
223_5 ACCAATCGGA GAACCACCAG CAGGCCCCTC TGGTCTGGGA TCTGGTACAA
223_6 ACCAATCGGA GAACCACCAG CAGGCCCCTC TGGTCTGGGA TCTGGTACAA
223_7 ACCAATCGGA GAACCACCAG CAGGCCCCTC TGGTCTGGGA TCTGGTACAA
A3_4 ACCAATCGGA GAACCCCCCG CAGCCCCCTC TGGTGTGGGA TCTAATACAA
A3_5 ACCAATCGGA GAACCCCCCG CAGCCCCCTC TGGTGTGGGA TCTAATACAA
A3_7 ACCAATCGGA GAACCCCCCG CAGCCCCCTC TGGTGTGGGA TCTAATACAA
A3_3 ACCAATCGGA GAACCCCCCG CAGCCCCCTC TGGTGTGGGA TCTAATACAA
42_12 ACCAATCGGA GAACCCCCCG CAGGCCCCTC TGGTCTGGGA TCTGGTACAA
AAV1 ACCTCTCGGA GAACCTCCAG CAACCCCCGC TGCTGTGGGA CCTACTACAA
AAV2 GCCTCTCGGA CAGCCACCAG CAGCCCCCTC TGGTCTGGGA ACTAATACGA
AAV3 ACCTCTCGGA GAACCACCAG CAGCCCCCAC AAGTTTGGGA TCTAATACAA
AAV8 ACCTCTCGGA GAACCTCCAG CAGCGCCCTC TGGTGTGGGA CCTAATACAA
AAV9 ACCTCTCGGA GAACCTCCAG CAGCGCCCTC TGGTGTGGGA CCTAATACAA
AAV7 ACCTCTCGGA GAACCTCCAG CAGCGCCCTC TAGTGTGGGA TCTGGTACAG
44_2 ACCAATCGGA GAACCCCCCG CAGGCCCCTC TGGTCTGGGA TCTGGTACAA
vp3 起始

图1AAE

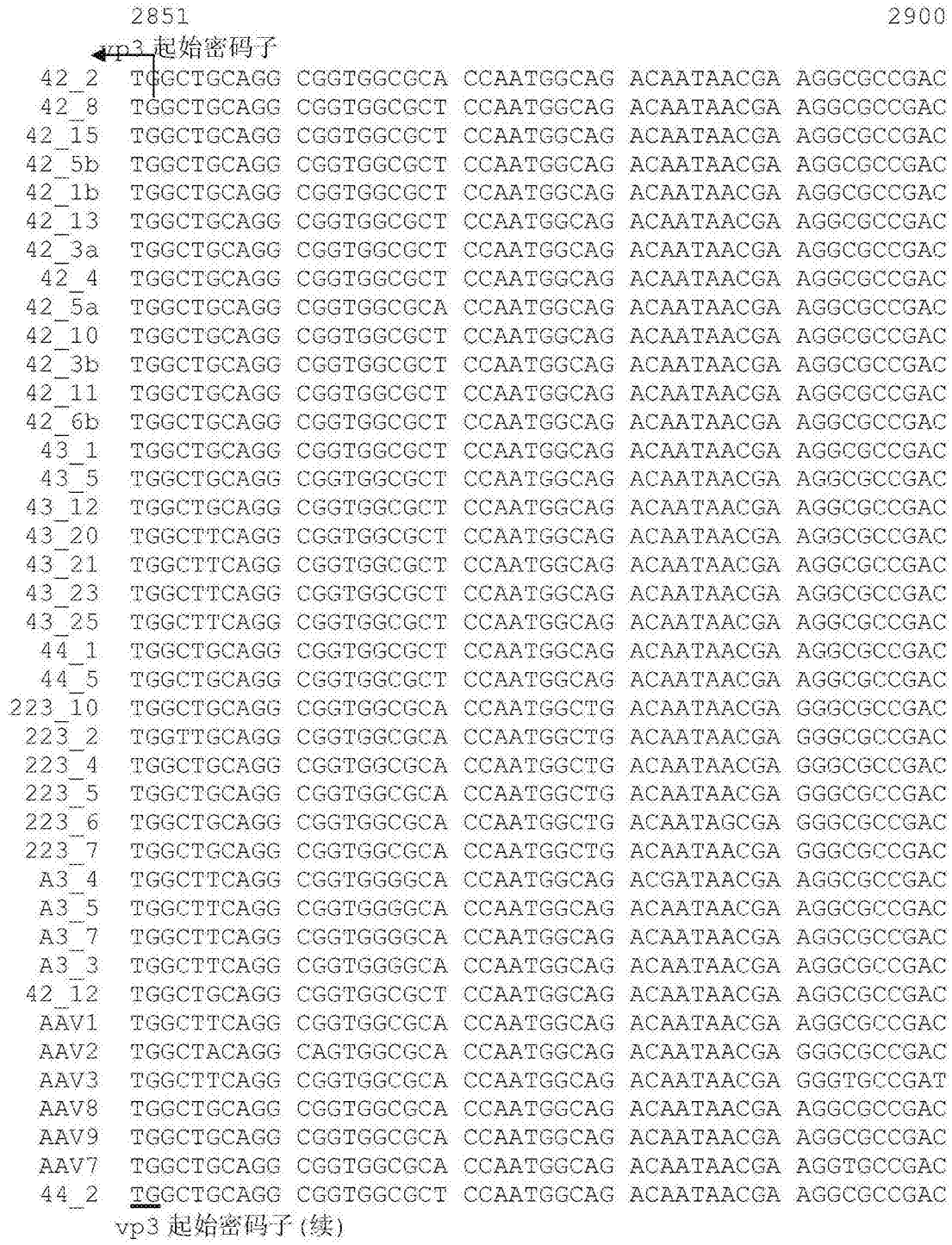

2851
2900
vp3 起始密码子
42_2 TGGCTGCAGG CGGTGGCGCA CCAATGGCAG ACAATAACGA AGGCGCCGAC
42_8 TGGCTGCAGG CGGTGGCGCT CCAATGGCAG ACAATAACGA AGGCGCCGAC
42_15 TGGCTGCAGG CGGTGGCGCT CCAATGGCAG ACAATAACGA AGGCGCCGAC
42_5b TGGCTGCAGG CGGTGGCGCT CCAATGGCAG ACAATAACGA AGGCGCCGAC
42_1b TGGCTGCAGG CGGTGGCGCT CCAATGGCAG ACAATAACGA AGGCGCCGAC
42_13 TGGCTGCAGG CGGTGGCGCT CCAATGGCAG ACAATAACGA AGGCGCCGAC
42_3a TGGCTGCAGG CGGTGGCGCT CCAATGGCAG ACAATAACGA AGGCGCCGAC
42_4 TGGCTGCAGG CGGTGGCGCT CCAATGGCAG ACAATAACGA AGGCGCCGAC
42_5a TGGCTGCAGG CGGTGGCGCA CCAATGGCAG ACAATAACGA AGGCGCCGAC
42_10 TGGCTGCAGG CGGTGGCGCT CCAATGGCAG ACAATAACGA AGGCGCCGAC
42_3b TGGCTGCAGG CGGTGGCGCT CCAATGGCAG ACAATAACGA AGGCGCCGAC
42_11 TGGCTGCAGG CGGTGGCGCT CCAATGGCAG ACAATAACGA AGGCGCCGAC
42_6b TGGCTGCAGG CGGTGGCGCT CCAATGGCAG ACAATAACGA AGGCGCCGAC
43_1 TGGCTGCAGG CGGTGGCGCT CCAATGGCAG ACAATAACGA AGGCGCCGAC
43_5 TGGCTGCAGG CGGTGGCGCT CCAATGGCAG ACAATAACGA AGGCGCCGAC
43_12 TGGCTGCAGG CGGTGGCGCT CCAATGGCAG ACAATAACGA AGGCGCCGAC
43_20 TGGCTTCAGG CGGTGGCGCT CCAATGGCAG ACAATAACGA AGGCGCCGAC
43_21 TGGCTTCAGG CGGTGGCGCT CCAATGGCAG ACAATAACGA AGGCGCCGAC
43_23 TGGCTTCAGG CGGTGGCGCT CCAATGGCAG ACAATAACGA AGGCGCCGAC
43_25 TGGCTTCAGG CGGTGGCGCT CCAATGGCAG ACAATAACGA AGGCGCCGAC
44_1 TGGCTGCAGG CGGTGGCGCT CCAATGGCAG ACAATAACGA AGGCGCCGAC
44_5 TGGCTGCAGG CGGTGGCGCT CCAATGGCAG ACAATAACGA AGGCGCCGAC
223_10 TGGCTGCAGG CGGTGGCGCA CCAATGGCTG ACAATAACGA GGGCGCCGAC
223_2 TGGTTGCAGG CGGTGGCGCA CCAATGGCTG ACAATAACGA GGGCGCCGAC
223_4 TGGCTGCAGG CGGTGGCGCA CCAATGGCTG ACAATAACGA GGGCGCCGAC
223_5 TGGCTGCAGG CGGTGGCGCA CCAATGGCTG ACAATAACGA GGGCGCCGAC
223_6 TGGCTGCAGG CGGTGGCGCA CCAATGGCTG ACAATAGCGA GGGCGCCGAC
223_7 TGGCTGCAGG CGGTGGCGCA CCAATGGCTG ACAATAACGA GGGCGCCGAC
A3_4 TGGCTTCAGG CGGTGGGGCA CCAATGGCAG ACGATAACGA AGGCGCCGAC
A3_5 TGGCTTCAGG CGGTGGGGCA CCAATGGCAG ACAATAACGA AGGCGCCGAC
A3_7 TGGCTTCAGG CGGTGGGGCA CCAATGGCAG ACAATAACGA AGGCGCCGAC
A3_3 TGGCTTCAGG CGGTGGGGCA CCAATGGCAG ACAATAACGA AGGCGCCGAC
42_12 TGGCTGCAGG CGGTGGCGCT CCAATGGCAG ACAATAACGA AGGCGCCGAC
AAV1 TGGCTTCAGG CGGTGGCGCA CCAATGGCAG ACAATAACGA AGGCGCCGAC
AAV2 TGGCTACAGG CAGTGGCGCA CCAATGGCAG ACAATAACGA GGGCGCCGAC
AAV3 TGGCTTCAGG CGGTGGCGCA CCAATGGCAG ACAATAACGA GGGTGCCGAT
AAV8 TGGCTGCAGG CGGTGGCGCA CCAATGGCAG ACAATAACGA AGGCGCCGAC
AAV9 TGGCTGCAGG CGGTGGCGCA CCAATGGCAG ACAATAACGA AGGCGCCGAC
AAV7 TGGCTGCAGG CGGTGGCGCA CCAATGGCAG ACAATAACGA AGGTGCCGAC
44_2 TGGCTGCAGG CGGTGGCGCT CCAATGGCAG ACAATAACGA AGGCGCCGAC
vp3 起始密码子(续)

图1AAF

```
          2901                                                    2950
  42_2    GGAGTGGGTA ATGCCTCCGG AAATTGGCAT TGCGATTCCA CATGGCTGGG
  42_8    GGAGTGGGTA GTTCCTCAGG AAATTGGCAT TGCGATTCCA CATGGCTGGG
 42_15    GGAGTGGGTA GTTCCTCAGG AAATTGGCAT TGCGATTCCA CATGGCTGGG
 42_5b    GGAGTGGGTA GTTCCTCAGG AAATTGGCAT TGCGATTCCA CATGGCTGGG
 42_1b    GGAGTGGGTA GTTCCTCAGG AAATTGGCAT TGCGATTCCA CATGGCTGGG
 42_13    GGAGTGGGTA GTTCCTCAGG AAATTGGCAT TGCGATTCCA CATGGCTGGG
 42_3a    GGAGTGGGTA GTTCCTCAGG AAATTGGCAT TGCGATTCCA CATAGCTGGG
  42_4    GGAGTGGGTA ATGCCTCCGG AAATTGGCAT TGCGATTCCA CATGGCTGGG
 42_5a    GGAGTGGGTA ATGCCTCCGG AAATTGGCAT TGCGATTCCA CATGGCTGGG
 42_10    GGAGTGGGTA ATGCCTCCGG AAATTGGCAT TGCGATTCCA CATGGCTGGG
 42_3b    GGAGTGGGTA ATGCCTCCGG AAATTGGCAT TGCGATTCCA CATGGCTGGG
 42_11    GGAGTGGGTA ATGCCTCCGG AAATTGGCAT TGCGATTCCA CATGGCTGGG
 42_6b    GGAGTGGGTA GTTCCTCAGG AAATTGGCAT TGCGATTCCA CATGGCTGGG
  43_1    GGAGTGGGTA GTTCCTCAGG AAATTGGCAT TGCGATTCCA CATGGCTGGG
  43_5    GGAGTGGGTA GTTCCTCAGG AAATTGGCAT TGCGATTCCA CATGGCTGGG
 43_12    GGAGTGGGTA GTTCCTCAGG AAATTGGCAT TGCGATTCCA CATGGCTGGG
 43_20    GGAGTGGGTA ATTCCTCGGG AAATTGGCAT TGCGATTCCA CATGGCTGGG
 43_21    GGAGTGGGTA ATTCCTCGGG AAATTGGCAT TGCGATTCCA CATGGCTGGG
 43_23    GGAGTGGGTA ATTCCTCGGG AAATTGGCAT TGCGATTCCA CATGGCTGGG
 43_25    GGAGTGGGTA ATTCCTCGGG AAATTGGCAT TGCGATTCCA CATGGCTGGG
  44_1    GGAGTGGGTA GTTCCTCAGG AAATTGGCAT TGCGATTCCA CATGGCTGGG
  44_5    GGAGTGGGTA GTTCCTCAGG AAATTGGCAT TGCGATTCCA CATGGCTGGG
223_10    GGAGTGGGTA ATGCCTCAGG AAATTGGCAT TGCGATTCCA CATGGCTGGG
 223_2    GGAGTGGGTA ATGCCTCAGG AAATTGGCAT TGCGATTCCA CATGGCTGGG
 223_4    GGAGTGGGTA ATGCCTCAGG AAATTGGCAT TGCGATTCCA CACGGCTGGG
 223_5    GGAGTGGGTA ATGCCTCAGG AAATTGGCAT TGCGATTCCA CACGGCTGGG
 223_6    GGAGTGGGTA ATGCCTCAGG AAATTGGCAT TGCGATTCCA CATGGCTGGG
 223_7    GGAGTGGGTA ATGCCTCAGG AAATTGGCAT TGCGATTCCA CATGGCTGGG
  A3_4    GGAGTGGGTA ATTCCTCGGG AAATTGGCAT TGCGATTCCA CATGGATGGG
  A3_5    GGAGTGGGTA ATTCCTCGGG AAATTGGCAT TGCGATTCCA CATGGATGGG
  A3_7    GGAGTGGGTA ATTCCTCGGG AAATTGGCAT TGCGATTCCA CATGGATGGG
  A3_3    GGAGTGGGTA ATTCCTCGGG AAATTGGCAT TGCGATTCCA CATGGATGGG
 42_12    GGAGTGGGTA GTTCCTCAGG AAATTGGCAT TGCGATTCCA CATGGCTGGG
  AAV1    GGAGTGGGTA ATGCCTCAGG AAATTGGCAT TGCGATTCCA CATGGCTGGG
  AAV2    GGAGTGGGTA ATTCCTCCGG AAATTGGCAT TGCGATTCCA CATGGATGGG
  AAV3    GGAGTGGGTA ATTCCTCAGG AAATTGGCAT TGCGATTCCC AATGGCTGGG
  AAV8    GGAGTGGGTA GTTCCTCGGG AAATTGGCAT TGCGATTCCA CATGGCTGGG
  AAV9    GGAGTGGGTA ATTCCTCGGG AAATTGGCAT TGCGATTCCA CATGGCTGGG
  AAV7    GGAGTGGGTA ATGCCTCAGG AAATTGGCAT TGCGATTCCA CATGGCTGGG
 AAV10          GGTA ATTCCTCCGG AAATTGGCAT TGCGATTCCA CATGGCTGGG
 AAV11          GGTA ATTCCTCCGG AAATTGGCAT TGCGATTCCA CATGGCTGGG
 AAV12          GGTA ATTCCTCCGG AAATTGGCAT TGCGATTCCA CATGGCTGGG
  44_2    GGAGTGGGTA GTTCCTCAGG AAATTGGCAT TGCGATTCCA CATGGCTGGG
```

图1AAG

```
          2951                                                        3000
  42_2    CGACAGAGTC ATCACCACCA GCACCCGCAC CTGGGCCCTG CCCACCTACA
  42_8    CGACAGAGTC ATCACCACCA GCACCCGAAC CTGGGCCCTC CCCACCTACA
 42_15    CGACAGAGTC ATCACCACCA GCACCCGAAC CTGGGCCCTC CCCACCTACA
 42_5b    CGACAGAGTC ATCACCACCA GCACCCGAAC CTGGGCCCTC CCCACCTACA
 42_1b    CGACAGAGTC ATCACCACCA GCACCCGAAC CTGGGCCCTC CCCACCTACA
 42_13    CGACAGAGTC ATCACCACCA GCACCCGAAC CTGGGCCCTC CCCACCTACA
 42_3a    CGACAGAGTC ATCACCACCA GCACCCGAAC CTGGGCCCTC CCCACCTACA
  42_4    CGACAGAGTC ATCACCACCA GCACCCGCAC CTGGGCCCTG CCCACCTACA
 42_5a    CGACAGAGTC ATCACCACCA GCACCCGCAC CTGGGCCCTG CCCACCTACA
 42_10    CGACAGAGTC ATCACCACCA GCACCCGCAC CTGGGCCCTG CCCACCTACA
 42_3b    CGACAGAGTC ATCACCACCA GCACCCGCAC CTGGGCCCTG CCCACCTACA
 42_11    CGACAGAGTC ATCACCACCA GCACCCGCAC CTGGGCCCTG CCCACCTACA
 42_6b    CGACAGAGTC ATCACCACCA GCACCCGAAC CTGGGCCCTC CCCACCTACA
  43_1    CGACAGAGTC ATCACCACCA GCACCCGAAC CTGGGCCCTG CCCACCTACA
  43_5    CGACAGAGTC ATCACCACCA GCACCCGAAC CTGGGCCCTG CCCACCTACA
 43_12    CGACAGAGTC ATCACCACCA GCACCCGAAC CTGGGCCCTG CCCACCTACA
 43_20    GGACAGAGTC ATCACCACCA GCACCCGAAC CTGGGCCCTG CCCACCTACA
 43_21    GGACAGAGTC ATCACCACCA GCACCCGAAC CTGGGCCCTG CCCACCTACA
 43_23    GGACAGAGTC ATCACCACCA GCACCCGAAC CTGGGCCCTG CCCACCTACA
 43_25    GGACAGAGTC ATCACCACCA GCACCCGAAC CTGGGCCCTG CCCACCTACA
  44_1    CGACAGAGTC ATCACCACCA GCACCCGAAC CTGGGCCCTC CCCACCTACA
  44_5    CGACAGAGTC ATCACCACCA GCACCCGAAC CTGGGCCCTC CCCACCTACA
223_10    CGACAGAGTC ATCACCACCA GCACCCGAAC CTGGGCCCTG CCCACCTACA
 223_2    CGACAGAGTC ATCACCACCA GCACCCGAAC CTGGGCCCTG CCCACCTACA
 223_4    CGACAGAGTC ATCACCACCA GCACCCGAAC CTGGGCCCTG CCCACCTACA
 223_5    CGACAGAGTC ATCACCACCA GCACCCGAAC CTGGGCCCTG CCCACCTACA
 223_6    CGACAGAGTC ATCACCACCA GCACCCGAAC CTGGGCCCTG CCCACCTACA
 223_7    CGACAGAGTC ATCACCACCA GCACCCGAAC CTGGGCCCTG CCCACCTACA
  A3_4    CGACAGAGTT ATCACCACCA GCACAAGAAC CTGGGCCCTC CCCACCTACA
  A3_5    CGACAGAGTT ATCACCACCA GCACAAGAAC CTGGGCCCTC CCCACCTACA
  A3_7    CGACAGAGTT ATCACCACCA GCACAAGAAC CTGGGCCCTC CCCACCTACA
  A3_3    CGACAGAGTT ATCACCACCA GCACAAGAAC CTGGGCCCTC CCCACCTACA
 42_12    CGACAGAGTC ATCACCACCA GCACCCGAAC CTGGGCCCTC CCCACCTACA
  AAV1    CGACAGAGTC ATCACCACCA GCACCCGCAC CTGGGCCTTG CCCACCTACA
  AAV2    CGACAGAGTC ATCACCACCA GCACCCGAAC CTGGGCCCTG CCCACCTACA
  AAV3    CGACAGAGTC ATCACCACCA GCACCAGAAC CTGGGCCCTG CCCACTTACA
  AAV8    CGACAGAGTC ATCACCACCA GCACCCGAAC CTGGGCCCTG CCCACCTACA
  AAV9    GGACAGAGTC ATCACCACCA GCACCCGAAC CTGGGCATTG CCCACCTACA
  AAV7    CGACAGAGTC ATTACCACCA GCACCCGAAC CTGGGCCCTG CCCACCTACA
 AAV10    CGACAGAGTC ATCACCACCA GCACCCGAAC CTGGGTCCTG CCCACCTACA
 AAV11    CGACAGAGTC ATCACCACCA GCACCCGAAC CTGGGCCCTG CCAACCTACA
 AAV12    CGACCGAGTC ATTACCACCA GCACCCGGAC TTGGGCCCTG CCCACCTACA
  44_2    CGACAGAGTC ATCACCACCA GCACCCGAAC CTGGGCCCTC CCCACCTACA
```

图1AAH

```
          3001                                                    3050
  42_2    ACAACCACCT CTACAAGCAG ATATCAA..G TCAGAGCGGG GCT....ACC
  42_8    ACAACCACCT CTACAAGCAA ATCTCCAACG GGACATCGGG AGGAAGCACC
 42_15    ACAACCACCT CTACAAGCAA ATCTCCAACG GGACATCGGG AGGAAGCACC
 42_5b    ACAACCACCT CTACAAGCAA ATCTCCAACG GGACATCGGG AGGAAGCACC
 42_1b    ACAACCACCT CTACAAGCAA ATCTCCAACG GGACATCGGG AGGAAGCACC
 42_13    ACAACCACCT CTACAAGCAA ATCTCCAACG GGACATCGGG AGGAAGCACC
 42_3a    ACAACCACCT CTACAAGCAA ATCTCCAACG GGACATCGGG AGGAAGCACC
  42_4    ACAACCACCT CTACAAGCAG ATATCAA... .GTCAGAGCG GGGC..TACC
 42_5a    ACAACCACCT CTACAAGCAG ATATCAA... .GTCAGAGCG GGGC..TACC
 42_10    ACAACCACCT CTACAAGCAG ATATCAA..G TCAGAGCGGG GCTA....CC
 42_3b    ACAACCACCT CTACAAGCAG ATATCAA..G TCAGAGCGGG GCTA....CC
 42_11    ACAACCACCT CTACAAGCAG ATATCAA..G TCAGAGCGGG GCTA....CC
 42_6b    ACAACCACCT CTACAAGCAA ATCTCCAACG GGACATCGGG AGGAAGCACC
  43_1    ACAACCATCT CTACAAGCAA ATCTCCAACG GGACATCGGG AGGAAGCACT
  43_5    ACAACCATCT CTACAAGCAA ATCTCCAACG GGACATCGGG AGGAAGCACT
 43_12    ACAACCATCT CTACAAGCAA ATCTCCAACG GGACATCGGG AGGAAGCACT
 43_20    ACAACCACCT CTACAAGCAA ATCTCCAACG GCACCTCGGG AGGAAGCACC
 43_21    ACAACCACCT CTACAAGCAA ATCTCCAACG GCACCTCGGG AGGAAGCACC
 43_23    ACAACCACCT CTACAAGCAA ATCTCCAACG GCACCTCGGG AGGAAGCACC
 43_25    ACAACCACCT CTACAAGCAA ATCTCCAACG GCACCTCGGG AGGAAGCACC
  44_1    ACAACCACCT CTACAAGCAA ATCTCCAACG GGACTTCGGG AGGAAGCACC
  44_5    ACAACCACCT CTACAAGCAA ATCTCCAACG GGACTTCGGG AGGAAGCACC
223_10    ACAACCACCT CTACAAGCAA ATCTCCAGTC AGTCAGCAGG GAG...CACC
 223_2    ACAACCACCT CTACAAGCAA ATCTCCAGTC AGTCAGCAGG GAG...CACC
 223_4    ACAACCACCT CTACAAGCAA ATCTCCAGTC AGTCAGCAGG GAG...CACC
 223_5    ACAACCACCT CTACAAGCAA ATCTCCAGTC AGTCAGCAGG GAG...CACC
 223_6    ACAACCACCT CTACAAGCAA ATCTCCAGTC AGTCAGCAGG GAG...CACC
 223_7    ACAACCACCT CTACAAGCAA ATCTCCAGTC AGTCAGCAGG GAG...CACC
  A3_4    ATAATCACCT CTACAAGCAA ATCTCCA... GCGAATCGGG AGC...CACC
  A3_5    ATAATCACCT CTACAAGCAA ATCTCCA... GCGAATCGGG AGC...CACC
  A3_7    ATAATCGCCT CTACAAGCAA ATCTCCA... GCGAATCGGG AGC...CACC
  A3_3    ATAATCACCT CTACAAGCAA ATCTCCA... GCGAATCGGG AGC...CACC
 42_12    ACAACCACCT CTACAAGCAA ATCTCCAACG GGACATCGGG AGGAAGCACC
  AAV1    ATAACCACCT CTACAAGCAA ATCTCCAGTG CTTCAACGGG .GG..CCAGC
  AAV2    ACAACCACCT CTACAAACAA ATTTCCA... GCCAATCAGG AGC...CTCG
  AAV3    ACAACCATCT CTACAAGCAA ATCTCCA... GCCAATCAGG AGC...TTCA
  AAV8    ACAACCACCT CTACAAGCAA ATCTCCAACG GGACATCGGG AGGAGCCACC
  AAV9    ACAACCACCT CTACAAGCAA ATCTCCAATG GAACATCGGG AGGAAGCACC
  AAV7    ACAACCACCT CTACAAGCAA ATCTCCAGTG AAACTGCAGG TAG...TACC
 AAV10    ACAACCACAT CTACAAGCAA ATCTCCAGCG AGACAGGAGC CACCAACGAC
 AAV11    ACAACCACCT CTACAAACAA ATCTCCAGCG CTTCAACGGG GGCCAGCAAC
 AAV12    ACAACCACCT CTACAAGCAA ATCTCCAGCC AATCGGGTGC CACCAACGAC
  44_2    ACAACCACCT CTACAAGCAA ATCTCCAACG GGACTTCGGG AGGAAGCACC
```

图1AAI

```
          3051                                                      3100
  42_2    AACGACAACC ACTTCTTCGG CTACAGCACC CCCTGGGGCT ATTTTGACTT
  42_8    AACGACAACA CCTACTTCGG CTACAGCACC CCCTGGGGGT ATTTTGACTT
 42_15    AACGACAACA CCTACTTCGG CTACAGCACC CCCTGGGGGT ATTTTGACTT
 42_5b    AACGACAACA CCTACTTCGG CTACAGCACC CCCTGGGGGT ATTTTGACTT
 42_1b    AACGACAACA CCTACTTCGG CTACAGCACC CCCTGGGGGT ATTTTGACTT
 42_13    AACGACAACA CCTACTTCGG CTACAGCACC CCCTGGGGGT ATTTTGACTT
 42_3a    AACGACAACA CCTACTTCGG CTACAGCACC CCCTGGGGGT ATTTTGACTT
  42_4    AACGACAACC ACTTCTTCGG CTACAGCACC CCCTGGGGCT ATTTTGACTT
 42_5a    AACGACAACC ACTTCTTCGG CTACAGCACC CCCTGGGGCT ATTTTGACTT
 42_10    AACGACAACC ACTTCTTCGG CTACAGCACC CCCTGGGGCT ATTTTGACTT
 42_3b    AACGACAACC ACTTCTTCGG CTACAGCACC CCCTGGGGCT ATTTTGACTT
 42_11    AACGACAACC ACTTCTTCGG CTACAGCACC CCCTGGGGCT ATTTTGACTT
 42_6b    AACGACAACA CCTACTTCGG CTACAGCACC CCCTGGGGGT ATTTTGACTT
  43_1    AACGACAACA CCTACTTTGG CTACAGCACC CCCTGGGGGT ATTTTGACTT
  43_5    AACGACAACA CCTACTTTGG CTACAGCACC CCCTGGGGGT ATTTTGACTT
 43_12    AACGACAACA CCTACTTTGG CTACAGCACC CCCTGGGGGT ATTTTGACTT
 43_20    AACGACAACA CCTATTTTGG CTACAGCACC CCCTGGGGGT ATTTTGACTT
 43_21    AACGACAACA CCTATTTTGG CTACAGCACC CCCTGGGGGT ATTTTGACTT
 43_23    AACGACAACA CCTATTTTGG CTACAGCACC CCCTGGGGGT ATTTTGACTT
 43_25    AACGACAACA CCTATTTTGG CTACAGCACC CCCTGGGGGT ATTTTGACTT
  44_1    AACGACAACA CCTACTTCGG CTACAGCACC CCCTGGGGGT ATTTTGACTT
  44_5    AACGACAACA CCTACTTCGG CTACAGCACC CCCTGGGGGT ATTTTGACTT
223_10    AACGATAACG TCTATTTCGG CTACAGCACC CCCTGGGGGT ATTTTGACTT
 223_2    AACGATAACG TCTATTTCGG CTACAGCACC CCCTGGGGGT ATTTTGACTT
 223_4    AACGATAACG TCTATTTCGG CTACAGCACC CCCTGGGGGT ATTTTGACTT
 223_5    AACGATAACG TCTATTTCGG CTACAGCACC CCCTGGGGGT ATTTTGACTT
 223_6    AACGATAACG TCTATTTCGG CTACAGCACC CCCTGGGGGT ATTTTGACTT
 223_7    AACGATAACG TCTATTTCGG CTACAGCACC CCCTGGGGGT ATTTTGACTT
  A3_4    AACGACAACC ACTACTTCGG CTACAGCACC CCCTGGGGGT ATTTTGACTT
  A3_5    AACGACAACC ACTACTTCGG CTACAGCACC CCCTGGGGGT ATTTTGACTT
  A3_7    AACGACAACC ACTACTTCGG CTACAGCACC CCCTGGGGGT ATTTTGACTT
  A3_3    AACGACAACC ACTACTTCGG CTACAGCACC CCCTGGGGGT ATTTTGACTT
 42_12    AACGACAACA CCTACTTCGG CTACAGCACC CCCTGGGGGT ATTTTGACTT
  AAV1    AACGACAACC ACTACTTCGG CTACAGCACC CCCTGGGGGT ATTTTGATTT
  AAV2    AACGACAATC ACTACTTTGG CTACAGCACC CCTTGGGGGT ATTTTGACTT
  AAV3    AACGACAACC ACTACTTTGG CTACAGCACC CCTTGGGGGT ATTTTGACTT
  AAV8    AACGACAACA CCTACTTCGG CTACAGCACC CCCTGGGGGT ATTTTGACTT
  AAV9    AACGACAACA CCTACTTTGG CTACAGCACC CCCTGGGGGT ATTTTGACTT
  AAV7    AACGACAACA CCTACTTCGG CTACAGCACC CCCTGGGGGT ATTTTGACTT
 AAV10    AACCACTACT TCGGCTACAG C......ACC CCCTGGGGGT ATTTTGACTT
 AAV11    ...GACAACC ACTACTTTGG CTACAGCACC CCCTGGGGGT ATTTTGACTT
 AAV12    AACCACTACT TCGGCTA... ...CAGCACC CCTTGGGGGT ATTTTGATTT
 44_2     AACGACAACA CCTACTTCGG CTACAGCACC CCCTGGGGGT ATTTTGACTT
```

图1AAJ

```
          3101                                                        3150
  42_2    TAACAGATTC CACTGCCACT TCTCACCACG TGACTGGCAG CGACTCATCA
  42_8    TAACAGATTC CACTGCCACT TCTCACCACG TGACTGGCAG CGACTCATCA
 42_15    TAACAGATTC CACTGCCACT TCTCACCACG TGACTGGCAG CGACTCATCA
 42_5b    TAACAGATTC CACTGCCACT TCTCACCACG TGACTGGCAG CGACTCATCA
 42_1b    TAACAGATTC CACTGCCACT TCTCACCACG TGACTGGCAG CGACTCATCA
 42_13    TAACAGATTC CACTGCCACT TCTCACCACG TGACTGGCAG CGACTCATCA
 42_3a    TAACAGATTC CACTGCCACT TCTCACCACG TGACTGGCAG CGACTCATCA
  42_4    CAACAGATTC CACTGCCACT TCTCATCACG TGACTGGCAG CGACTCATCA
 42_5a    CAACAGATTC CACTGCCACT TCTCACCACG TGACTGGCAG CGACTCATCA
 42_10    CAACAGATTC CACTGCCACT TCTCACCACG TGACTGGCAG CGACTCATCA
 42_3b    CAACAGATTC CACTGCCACT TCTCACCACG TGACTGGCAG CGACTCATCA
 42_11    CAACAGATTC CACTGCCACT TCTCACCACG TGACTGGCAG CGACTCATCA
 42_6b    TAACAGATTC CACTGCCACT TCTCACCACG TGACTGGCAG CGACTCATCA
  43_1    CAACAGATTC CACTGCCACT TCTCACCACG TGACTGGCAG CGACTCATCA
  43_5    CAACAGATTC CACTGCCACT TCTCACCACG TGACTGGCAG CGACTCATCA
 43_12    CAACAGATTC CACTGCCACT TCTCACCACG TGACTGGCAG CGACTCATCA
 43_20    CAACAGATTC CACTGTCACT TTTCACCACG TGACTGGCAA CGACTCATCA
 43_21    CAACAGATTC CACTGTCACT TTTCACCACG TGACTGGCAA CGACTCATCA
 43_23    CAACAGATTC CACTGTCACT TTTCACCACG TGACTGGCAA CGACTCATCA
 43_25    CAACAGATTC CACTGTCACT TTTCACCACG TGACTGGCAA CGACTCATCA
  44_1    TAACAGATTC CACTGCCACT TCTCACCACG TGACTGGCAG CGACTCATCA
  44_5    TAACAGATTC CACTGCCACT TCTCACCACG TGACTGGCAG CGACTCATCA
223_10    CAACAGATTC CATTGCCACT TCTCACCACG TGACTGGCAG CGACTTATCA
 223_2    CAACAGATTC CATTGCCACT TCTCACCACG TGACTGGCAG CGACTTATCA
 223_4    CAACAGATTC CATTGCCACT TCTCACCACG TGACTGGCAG CGACTTATCA
 223_5    CAACAGATTC CATTGCCACT TCTCACCACG TGACTGGCAG CGACTTATCA
 223_6    CAACAGATTC CATTGCCACT TCTCACCACG TGACTGGCAG CGACTTATCA
 223_7    CAACAGATTC CATTGCCACT TCTCACCACG TGACTGGCAG CGACTTATCA
  A3_4    TAACAGATTC CACTGTCACT TCTCACCACG TGACTGGCAG CGACTCATCA
  A3_5    TAACAGATTC CACTGTCACT TCTCACCACG TGACTGGCAG CGACTCATCA
  A3_7    TAACAGATTC CACTGTCACT TCTCACCACG TGACTGGCAG CGACTCATCA
  A3_3    TAACAGATTC CACTGTCACT TCTCACCACG TGACTGGCAG CGACTCATCA
 42_12    TAACAGATTC CACTGCCACT TCTCACCACG TGACTGGCAG CGACTCATCA
  AAV1    CAACAGATTC CACTGCCACT TTTCACCACG TGACTGGCAG CGACTCATCA
  AAV2    CAACAGATTC CACTGCCACT TTTCACCACG TGACTGGCAA AGACTCATCA
  AAV3    TAACAGATTC CACTGCCACT TCTCACCACG TGACTGGCAG CGACTCATTA
  AAV8    TAACAGATTC CACTGCCACT TTTCACCACG TGACTGGCAG CGACTCATCA
  AAV9    CAACAGATTC CACTGCCACT TCTCACCACG TGACTGGCAG CGACTCATCA
  AAV7    TAACAGATTC CACTGCCACT TCTCACCACG TGACTGGCAG CGACTCATCA
 AAV10    TAACAGATTC CACTGCCACT TTTCACCACG TGACTGGCAG CGACTCATCA
 AAV11    TAACAGATTC CACTGCCACT TCTCACCACG TGACTGGCAG CGACTCATCA
 AAV12    CAACAGATTC CACTGCCATT TCTCACCACG TGACTGGCAG CGACTCATCA
  44_2    TAACAGATTC CACTGCCACT TCTCACCACG TGACTGGCAG CGACTCATCA
```

图1AAK

```
          3151                                                        3200
  42_2    ACAACAACTG GGGATTCCGG CCCAGAAAGC TGCGGTTCAA GTTGTTCAAC
  42_8    ACAACAACTG GGGATTCCGG CCCAAGAGAC TCAACTTCAA GCTCTTCAAC
 42_15    ACAACAACTG GGGATTCCGG CCCAAGAGAC TCAACTTCAA GCTCTTCAAC
 42_5b    ACAACAACTG GGGATTCCGG CCCAAGAGAC TCAACTTCAA GCTCTTCAAC
 42_1b    ACAACAACTG GGGATTCCGG CCCAAGAGAC TCAACTTCAA GCTCTTCAAC
 42_13    ACAACAACTG GGGATTCCGG CCCAAGAGAC TCAACTTCAA GCTCTTCAAC
 42_3a    ACAACAGCTG GGGATTCCGG CCCAAGAGAC TCAACTTCAA GCTCTTCAAC
  42_4    ACAACAACTG GGGATTCCGG CCCAAGAGAC TCAACTTCAA GCTCTTCAAC
 42_5a    ACAACAACCG GGGATTCCGG CCCAGAAAGC TGCGGTTCAA GTTGTTCAAC
 42_10    ACAACAACTG GGGATTCCGG CCCAGAAAGC TGCGGTTCAA GTTGTTCAAC
 42_3b    ACAACAACTG GGGATTCCGG CCCAGAAAGC TGCGGTTCAA GTTGTTCAAC
 42_11    ACAACAACTG GGGATTCCGG CCCAGAAAGC TGCGGTTCAA GTTGTTCAAC
 42_6b    ACAACAACTG GGGATTCCGG CCCAGAAAGC TGCGGTTCAA GTTGTTCAAC
  43_1    ACAATAACTG GGGATTCCGG CCCAAGAGAC TCAACTTCAA GCTCTTCAAC
  43_5    ACAATAACTG GGGATTCCGG CCCAAGAGAC TCAACTTCAA GCTCTTCAAC
 43_12    ACAATAACTG GGGATTCCGG CCCAAGAGAC TCAACTTCAA GCTCTTCAAC
 43_20    ACAACAATTG GGGATTCCGG CCCAAAAGAC TCAACTTCAA GCTGTTCAAC
 43_21    ACAACAATTG GGGATTCCGG CCCAAAAGAC TCAACTTCAA GCTGTTCAAC
 43_23    ACAACAATTG GGGATTCCGG CCCAAAAGAC TCAACTTCAA GCTGTTCAAC
 43_25    ACAACAATTG GGGATTCCGG CCCAAAAGAC TCAACTTCAA GCTGTTCAAC
  44_1    ACAACAACTG GGGATTCCGG CCCAAGAGAC TCAACTTCAA GCTCTTCAAC
  44_5    ACAACAACTG GGGATTCCGG CCCAAGAGAC CCAACTTCAA GCTCTTCAAC
223_10    ACAACAACTG GGGATTCCGG CCCAAGAAGC TCAACTTCAA GCTCTTCAAC
 223_2    ACAACAACTG GGGATTCCGG CCCAAGAAGC TCAACTTCAA GCTCTTCAAC
 223_4    ACAACAACTG GGGATTCCGG CCCAAGAAGC TCAACTTCAA GCTCTTCAAC
 223_5    ACAACAACTG GGGATTCCGG CCCAAGAAGC TCAACTTCAA GCTCTTCAAC
 223_6    ACAACAACTG GGGATTCCGG CCCAAGAAGC TCAACTTCAA GCTCTTCAAC
 223_7    ACAACAACTG GGGATTCCGG CCCAAGAAGC TCAACTTCAA GCTCTTCAAC
  A3_4    ACAACAACTG GGGATTTAGA CCCAAGAAAC TCAATTTCAA GCTCTTCAAC
  A3_5    ATAACAACTG GGGATTTAGA CCCAAGAAAC TCAATTTCAA GCTCTTCAAC
  A3_7    ACAACAACTG GGGATTTAGA CCCAAGAAAC TCAATTTCAA GCTCTTCAAC
  A3_3    ACAACAACTG GGGATTTAGA CCCAAGAAAC TCAATTTCAA GCTCTTCAAC
 42_12    ACAACAACTG GGGATTCCGG CCCAAGAGAC TCAACTTCAA GCTCTTCAAC
  AAV1    ACAACAATTG GGGATTCCGG CCCAAGAGAC TCAACTTCAA ACTCTTCAAC
  AAV2    ACAACAACTG GGGATTCCGA CCCAAGAGAC TCAACTTCAA GCTCTTTAAC
  AAV3    ACAACAACTG GGGATTCCGG CCCAAGAAAC TCAGCTTCAA GCTCTTCAAC
  AAV8    ACAACAACTG GGGATTCCGG CCCAAGAGAC TCAGCTTCAA GCTCTTCAAC
  AAV9    ACAACAACTG GGGATTCCGG CCAAAGAGAC TCAACTTCAA GCTGTTCAAC
  AAV7    ACAACAACTG GGGATTCCGG CCCAAGAAGC TGCGGTTCAA GCTCTTCAAC
 AAV10    ACAACAACTG GGGATTC
 AAV11    ACAACAACTG GGGATTC
 AAV12    ACAACAACTG GGGATTC
  44_2    ACAACAACTG GGGATTCCGG CCCAAGAGAC TCAACTTCAA GCTCTTCAAC
```

图1AAL

```
          3201                                                  3250
  42_2    ATCCAGGTCA AGGAGGTCAC GACGAACGAC GGCGTTACGA CCATCGCTAA
  42_8    ATCCAGGTCA AGGAGGTCAC GCAGAATGAA GGCACCAAGA CCATCGCCAA
 42_15    ATCCAGGTCA AGGAGGTCAC GCAGAATGAA GGCACCAAGA CCATCGCCAA
 42_5b    ATCCAGGTCA AGGAGGTCAC GCAGAATGAA GGCACCAAGA CCATCGCCAA
 42_1b    ATCCAGGTCA AGGAGGTCAC GCAGAATGAA GGCACCAAGA CCATCGCCAA
 42_13    ATCCAGGTCA AGGAGGTCAC GCAGAATGAA GGCACCAAGA CCATCGCCAA
 42_3a    ATCCAGGTCA AGGAGGTCAC GCAGAATGAA GGCACCAAGA CCATCGCCAA
  42_4    ATCCAGGTCA AGGAGGTCAC GCAGAATGAA GGCACCAAGA CCATCGCCAA
 42_5a    ATCCAGGTCA AGGAGGTCAC GACGAACGAC GGCGTTACGA CCATCGCTAA
 42_10    ATCCAGGTCA AGGAGGTCAC GACGAACGAC GGCGTTACGA CCATCGCCAA
 42_3b    ATCCAGGTCA AGGAGGTCAC GACGAACGAC GGCGTTACGA CCATCGCTAA
 42_11    ATCCAGGTCA AGGAGGTCAC GACGAACGAC GGCGTTACGA CCATCGCTAA
 42_6b    ATCCAGGTCA AGGAGGTCAC GACGGACGAC GGCGTTACGA CCATCGCTAA
  43_1    ATCCAGGTCA AGGAGGTCAC GCAGAATGAA GGCACCAAGA CCATCGCCAA
  43_5    ATCCAGGTCA AGGAGGTCAC GCAGAATGAA GGCACCAAGA CCATCGCCAA
 43_12    ATCCAGGTCA AGGAGGTCAC GCAGAATGAA GGCACCAAGA CCATCGCCAA
 43_20    ATCCAGGTCA AGGAAGTCAC GACGAACGAA GGCACCAAGA CCATCGCCAA
 43_21    ATCCAGGTCA AGGAAGTCAC GACGAACGAA GGCACCAAGA CCATCGCCAA
 43_23    ATCCAGGTCA AGGAAGTCAC GACGAACGAA GGCACCAAGA CCATCGCCAA
 43_25    ATCCAGGTCA AGGAAGTCAC GACGAACGAA GGCACCAAGA CCATCGCCAA
  44_1    ATCCAGGTCA AGGAGGTCAC GCAGAATGAA GGCACCAAGA CCATCGCCAA
  44_5    ATCCAGGTCA AGGAGGTCAC GCAGAATGAA GGCACCAAGA CCATCGCCAA
223_10    ATCCAGGTCA AGGAGGTCAC GACGAATGAC GGTGTCACAA CCATCGCTAA
 223_2    ATCCAGGTCA AGGAGGTCAC GACGAATGAC GGTGTCACAA CCATCGCTAA
 223_4    ATCCAGGTCA AGGAGGTCAC GACGAATGAC GGCGTCACAA CCATCGCTAA
 223_5    ATCCAGGTCA AGGAGGTCAC GACGAATGAC GGCGTCACAA CCATCGCTAA
 223_6    ATCCAGGTCA AGGAGGTCAC GACGAATGAC GGTGTCACAA CCATCGCTAA
 223_7    ATCCAGGTCA AGGAGGTCAC GACGAATGAC GGCGTCACAA CCATCGCTAA
  A3_4    ATCCAAGTCA AGGAGGTCAC GCAGAATGAT GGAACCACGA CCATCGCCAA
  A3_5    ATCCAAGTCA AGGAGGTCAC GCAGAATGAT GGAACCACGA CCATCGCCAA
  A3_7    ATCCAAGTCA AGGAGGTCAC GCAGAATGAT GGAACCACGA CCATCGCCAA
  A3_3    ATCCAAGTCA AGGAGGTCAC GCAGAATGAT GGAACCACGA CCATCGCCAA
 42_12    ATCCAGGTCA AGGAGGTCAC GCAGAATGAA GGCACCAAGA CCATCGCCAA
   AAV1   ATCCAAGTCA AGGAGGTCAC GACGAATGAT GGCGTCACAA CCATCGCTAA
   AAV2   ATTCAAGTCA AAGAGGTCAC GCAGAATGAC GGTACGACGA CGATTGCCAA
   AAV3   ATCCAAGTTA GAGGGGTCAC GCAGAACGAT GGCACGACGA CTATTGCCAA
   AAV8   ATCCAGGTCA AGGAGGTCAC GCAGAATGAA GGCACCAAGA CCATCGCCAA
   AAV9   ATCCAGGTCA AGGAGGTTAC GACGAACGAA GGCACCAAGA CCATCGCCAA
   AAV7   ATCCAGGTCA AGGAGGTCAC GACGAATGAC GGCGTTACGA CCATCGCTAA
  44_2    ATCCAGGTCA AGGAGGTCAC GCAGAATGAA GGCACCAAGA CCATCGCCAA
```

图1AAM

```
          3251                                                    3300
  42_2    TAACCTTACC AGCACGATTC AGGTCTTCTC GGACTCGGAG TACCAACTGC
  42_8    TAACCTTACC AGCACGATTC AGGTCTTTAC GGACTCGGAA TACCAGCTCC
 42_15    TAACCTTACC AGCACGATTC AGGTCTTTAC GGACTCGGAA TACCAGCTCC
 42_5b    TAACCTTACC AGCACGATTC AGGTCTTTAC GGACTCGGAA TACCAGCTCC
 42_1b    TAACCTTACC AGCACGATTC AGGTCTTTAC GGACTCGGAA TACCAGCTCC
 42_13    TAACCTTACC AGCACGATTC AGGTCTTTAC GGACTCGGAA TACCAGCTCC
 42_3a    TAACCTTACC AGCACGATTC AGGTCTTTAC GGACTCGGAA TACCAGCTCC
  42_4    TAACCTTACC AGCACGATTC AGGTCTTTAC GGACTCGGAA TACCGGCTCC
 42_5a    TAACCTTACC AGCACGATTC AGGTCTTCTC GGACTCGGAG TACCAACTGC
 42_10    TAACCTTACC AGCACGATTC AGGTCTTCTC GGACTCGGAG TACCAACTGC
 42_3b    TAACCTTACC AGCACGATTC AGGTCTTCTC GGACTCGGAG TACCAACTGC
 42_11    TAACCTTACC AGCACGATTC AGGTCTTCTC GGACTCGGAG TACCAACTGC
 42_6b    TAACCTTACC AGCACGATTC AGGTCTTCTC GGACTCGGAG TACCAACTGC
  43_1    TAACCTTACC AGCACGATTC AGGTGTTTAC GGACTCGGAA TACCAGCTCC
  43_5    TAACCTTACC AGCACGATTC AGGTGTTTAC GGACTCGGAA TACCAGCTCC
 43_12    TAACCTTACC AGCACGATTC AGGTGTTTAC GGACTCGGAA TACCAGCTCC
 43_20    TAATCTCACC AGCACCGTGC AGGTCTTTAC GGACTCGGAG TACCAGTTAC
 43_21    TAATCTCACC AGCACCGTGC GGGTCTTTAC GGACTCGGAG TACCAGTTAC
 43_23    TAATCTCACC AGCACCGTGC AGGTCTTTAC GGACTTGGAG TACCAGTTAC
 43_25    TAATCTCACC AGCACCGTGC AGGTCTTTAC GGACTCGGAG TACCAGTTAC
  44_1    TAACCTTACC AGCACGATTC AGGTCTTTAC GGACTCGGAA TACCAGCTCC
  44_5    TAACCTTACC AGCACGATTC AGGTCTTTAC GGACTCGGAA TACCAGCTCC
223_10    TAACCTTACC AGCACGGTTC AGGTCTTTTC GGACTCGGAA TATCAACTGC
 223_2    TAACCTTACC AGCACGGTTC AGGTCTTTTC GGACTCGGAA TATCAACTGC
 223_4    TAACCTTACC AGCACGGTTC AGGTCTTTTC GGACTCGGAA TATCAACTGC
 223_5    TAACCTTACC AGCACGGTTC AGGTCTTTTC GGACTCGGAA TATCAACTGC
 223_6    TAACCTTACC AGCACGGTTC AGGTCTTTTC GGACTCGGAA TATCAACTGC
 223_7    TAACCTTACC AGCACGGTTC AGGTCTTTTC GGACCCGGAA TATCAACTGC
  A3_4    TAACCTTACC AGCACGGTGC AGGTCTTCAC AGACTCTGAG TACCAGCTGC
  A3_5    TAACCTTACC AGCACGGTGC AGGTCTTCAC AGACTCTGAG TACCAGCTGC
  A3_7    TAACCTTACC AGCACGGTGC AGGTCTTCAC AGACTCTGAG TACCAGCTGC
  A3_3    TAACCTTACC AGCGCGGTGC AGGTCTTCAC AGACTCTGAG TACCAGCTGC
 42_12    TAACCTTACC AGCACGATTC AGGTCTTTAC GGACTCGGAA TACCAGCTCC
   AAV1   TAACCTTACC AGCACGGTTC AAGTCTTCTC GGACTCGGAG TACCAGCTTC
   AAV2   TAACCTTACC AGCACGGTTC AGGTGTTTAC TGACTCGGAG TACCAGCTCC
   AAV3   TAACCTTACC AGCACGGTTC AAGTGTTTAC GGACTCGGAG TATCAGCTCC
   AAV8   TAACCTCACC AGCACCATCC AGGTGTTTAC GGACTCGGAG TACCAGCTGC
   AAV9   TAACCTTACC AGCACCGTCC AGGTCTTTAC GGACTCGGAG TACCAGCTAC
   AAV7   TAACCTTACC AGCACGATTC AGGTATTCTC GGACTCGGAA TACCAGCTGC
  44_2    TAACCTTACC AGCACGATTC AGGTCTTTAC GGACTCGGAA TACCAGCTCC
```

图1AAN

```
          3301                                                    3350
  42_2    CGTACGTCCT CGGCTCTGCG CACCAGGGCT GCCTCCCTCC GTTCCCTGCG
  42_8    CGTACGTCCT CGGCTCTGCG CACCAGGGCT GCCTGCCTCC GTTCCCGGCG
 42_15    CGTACGTCCT CGGCTCTGCG CACCAGGGCT GCCCGCCTCC GTTCCCGGCG
 42_5b    CGTACGTCCT CGGCTCTGCG CACCAGGGCT GCCTGCCTCC GTTCCCGGCG
 42_1b    CGTACGTCCT CGGCTCTGCG CACCAGGGCT GCCTGCCTCC GTTCCCGGCG
 42_13    CGTACGTCCT CGGCTCTGCG CACCAGGGCT GCCTGCCTCC GTTCCCGGCG
 42_3a    CGTACGTCCT CGGCTCTGCG CACCAGGGCT GCCTGCCTCC GTTCCCGGCG
  42_4    CGTACGTCCT CGGCTCTGCG CACCAGGGCT GCCTGCCTCC GTTCCCGGCG
 42_5a    CGTACGTCCT CGGCTCTGCG CACCAGGGCT GCCTCCCTCC GTTCCCTGCG
 42_10    CGTACGTCCT CGGCTCTGCG CACCAGGGCT GCCTCCCTCC GTTCCCTGCG
 42_3b    CGTACGTCCT CGGCTCTGCG CACCAGGGCT GCCTCCCTCC GTTCCCTGCG
 42_1     CGTACGTCCT CGGCTCTGCG CACCAGGGCT GCCTCCCTCC GTTCCCTGCG
 42_6b    CGTACGTCCT CGGCTCTGCG CACCAGGGCT GCCTCCCTCC GTTCCCTGCG
  43_1    CGTACGTCCC CGGCTCTGCG CACCAGGGCT GCCTCCCTCC GTTCCCGGCG
  43_5    CGTACGTCCT CGGCTCTGCG CACCAGGGCT GCCTCCCTCC GTTCCCGGCG
 43_12    CGTACGTCCT CGGCTCTGCG CACCAGGGCT GCCTCCCTCC GTTCCCGGCG
 43_20    CGTACGTGCT AGGATCCGCT CACCAGGGAT GTCTGCCTCC GTTCCCGGCG
 43_21    CGTACGTGCT AGGATCCGCT CACCAGGGAT GTCTGCCTCC GTTCCCGGCG
 43_23    CGTACGTGCT AGGATCCGCT CACCAGGGAT GTCTGCCTCC GTTCCCGGCG
 43_25    CGTACGTGCT AGGATCCGCT CACCAGGGAT GTCTGCCTCC GTTCCCGGCG
  44_1    CGTACGTCCT CGGCTCTGCG CACCAGGGCT GCCTGCCTCC GTTCCCGGCG
  44_5    CGTACGTCCT CGGCTCTGCG CACCAGGGCT GCCTGCCTCC GTTCCCGGCG
223_10    CGTACGTCCT CGGCTCCGCG CACCAGGGCT GCCTGCCTCC GTTCCCGGCA
 223_2    CGTACGTCCT CGGCTCCGCG CACCAGGGCT GCCTGCCTCC GTTCCCGGCA
 223_4    CGTACGTCCT CGGCTCCGCG CACCAGGGCT GCCTGCCTCC GTTCCCGGCA
 223_5    CGTACGTCCT CGGCTCCGCG CACCAGGGCT GCCTGCCTCC GTTCCCGGCA
 223_6    CGTACGTCCT CGGCTCCGCG CACCAGGGCT GCCTGCCTCC GTTCCCGGCA
 223_7    CGTACGTCCT CGGCTCCGCG CACCAGGGCT GCCTGCCTCC GTTCCCGGCA
  A3_4    CCTACGTCCT CGGTTCGGCT CACCAGGGCT GCCTTCCGCC GTTCCCAGCA
  A3_5    CCTACGTCCT CGGTTCGGCT CACCAGGGCT GCCTTCCGCC GTTCCCAGCA
  A3_7    CCTACGTCCT CGGTTCGGCT CACCAGGGCT GCCTTCCGCC GTTCCCAGCA
  A3_3    CCTACGTCCT CGGTTCGGCT CACCAGGGCT GCCTTCCGCC GTTCCCAGCA
 42_12    CGTACGTCCT CGGCTCTGCG CACCAGGGCT GCCTGCCTCC GTTCCCGGCG
  AAV1    CGTACGTCCT CGGCTCTGCG CACCAGGGCT GCCTCCCTCC GTTCCCGGCG
  AAV2    CGTACGTCCT CGGCTCGGCG CATCAAGGAT GCCTCCCGCC GTTCCCAGCA
  AAV3    CGTACGTGCT CGGGTCGGCG CACCAAGGCT GTCTCCCGCC GTTTCCAGCG
  AAV8    CGTACGTTCT CGGCTCTGCC CACCAGGGCT GCCTGCCTCC GTTCCCGGCG
  AAV9    CGTACGTCCT AGGCTCTGCC CACCAAGGAT GCCTGCCACC GTTTCCTGCA
  AAV7    CGTACGTCCT CGGCTCTGCG CACCAGGGCT GCCTGCCTCC GTTCCCGGCG
  44_2    CGTACGTCCT CGGCTCTGCG CACCAGGGCT GCCTGCCTCC GTTCCCGGCG
```

图1AAO

```
          3351                                                        3400
  42_2    GACGTGTTCA TGATTCCTCA GTACGGATAT CTGACTCTAA ACAACGGCAG
  42_8    GACGTCTTCA TGATTCCTCA GTACGGGTAC CTGACTCTGA ACAACGGCAG
 42_15    GACGTCTTCA TGATTCCTCA GTACGGGTAC CTGACTCTGA ACAACGGCAG
 42_5b    GACGTCTTCA TGATTCCTCA GTACGGGTAC CTGACTCTGA ACAACGGCAG
 42_1b    GACGTCTTCA TGATTCCTCA GTACGGGTAC CTGACTCTGA ACAACGGCAG
 42_13    GACGTCTTCA TGATTCCTCA GTACGGGTAC CTGACTCTGA ACAACGGCAG
 42_3a    GACGTCTTCA TGATTCCTCA GTACGGGTAC CTGACTCTGA ACAACGGCAG
  42_4    GACGTCTTCA TGATTCCTCA GTACGGGTAC CTGACTCTGA ACAACGGCAG
 42_5a    GACGTGTTCA TGATTCCTCA GTACGGATAT CTGACTCTAA ACAACGGCAG
 42_10    GACGTGTTCA TGATTCCTCA GTACGGATAT CTGACTCTAA ACAACGGCAG
 42_3b    GACGTGTTCA TGATTCCTCA GTACGGATAT CTGACTCTAA ACAACGGCAG
  42_1    GACGTGTTCA TGATTCCTCA GTACGGATAT CTGACTCTAA ACAACGGCAG
 42_6b    GACGTGTTCA TGATTCCTCA GTACGGATAT CTGACTCTAA ACAACGGCAG
  43_1    GACGTCTTCA TGATTCCTCA GTACGGGTAT CTGACCCTAA ACAATGGCAG
  43_5    GACGTCTTCA TGATTCCTCA GTACGGGTAT CTGACCCTAA ACAATGGCAG
 43_12    GACGTCTTCA TGATTCCTCA GTACGGGTAT CTGACCCTAA ACAATGGCAG
 43_20    GACGTCTTCA CGGTTCCTCA GTACGGCTAT TTAACTTTAA ACAATGGAAG
 43_21    GACGTCTTCA TGGTTCCTCA GTACGGCTAT TTAACTTTAA ACAATGGAAG
 43_23    GACGTCTTCA TGGTTCCTCA GTACGGCTAT TTAACTTTAA ACAATGGAAG
 43_25    GACGTCTTCA TGGTTCCTCA GTACGGCTAT TTAACTTTAA ACAATGGAAG
  44_1    GACGTCTTCA TGATTCCTCA GTACGGGTAC CTGACTCTGA ACAATGGCAG
  44_5    GACGTCTTCA TGATTCCTCA GTACGGGTAC CTGACTCTGA ACAATGGCAG
223_10    GACGTGTTCA TGATTCCGCA GTACGGATAC CTGACTCTGA ACAATGGCAG
 223_2    GACGTGTTCA TGATTCCGCA GTACGGATAC CTGACTCTGA ACAATGGCAG
 223_4    GACGTGTTCA TGATTCCGCA GTACGGATAC CTGACTCTGA ACAATGGCAG
 223_5    GACGTGTTCA TGATTCCGCA GTACGGATAC CTGACTCTGA ACAATGGCAG
 223_6    GACGTGTTCA TGATTCCGCA GTACGGATAC CTGACTCTGA ACAATGGCAG
 223_7    GACGTGTTCA TGATTCCGCA GTACGGATAC CTGACTCTGA ACAATGGCAG
  A3_4    GACGTCTTCA TGATTCCTCA GTACGGCTAC TTGACTCTGA ACAATGGCAG
  A3_5    GACGTCTTCA TGATTCCTCA GTACGGCTAC TTGACTCTGA ACAATGGCAG
  A3_7    GACGTCTTCA TGATTCCTCA GTACGGCTAC TTGACTCTGA ACAATGGCAG
  A3_3    GACGTCTTCA TGATTCCTCA GTACGGCTAC TTGACTCTGA ACAATGGCAG
 42_12    GACGTCTTCA TGATTCCTCA GTACGGGTAC CTGACTCTGA ACAACGGCAG
   AAV1   GACGTGTTCA TGATTCCGCA ATACGGCTAC CTGACGCTCA ACAATGGCAG
   AAV2   GACGTCTTCA TGGTGCCACA GTATGGATAC CTCACCCTGA ACAACGGGAG
   AAV3   GACGTCTTCA TGGTCCCTCA GTATGGATAC CTCACCCTGA ACAACGGAAG
   AAV8   GACGTGTTCA TGATTCCCCA GTACGGCTAC CTAACACTCA ACAACGGTAG
   AAV9   GACGTCTTCA TGGTTCCTCA GTACGGCTAC CTGACGCTCA ACAATGGAAG
   AAV7   GACGTCTTCA TGATTCCTCA GTACGGCTAC CTGACTCTCA ACAATGGCAG
  44_2    GACGTCTTCA TGATTCCTCA GTACGGGTAC CTGACTCTGA ACAATGGCAG
```

图1AAP

```
          3401                                                      3450
  42_2    TCAGTCTGTG GGACGTTCCT CCTTCTACTG CCTGGAGTAC TTTCCTTCTC
  42_8    TCAGGCCGTG GGCCGTTCCT CCTTCTACTG CCTGGAGTAC TTTCCTTCTC
 42_15    TCAGGCCGTG GGCCGTTCCT CCTTCTACTG CCTGGAGTAC TTTCCTTCTC
 42_5b    TCAGGCCGTG GGCCGTTCCT CCTTCTACTG CCTGGAGTAC TTTCCTTCTC
 42_1b    TCAGGCCGTG GGCCGTTCCT CCTTCTACTG CCTGGAGTAC TTTCCTTCTC
 42_13    TCAGGCCGTG GGCCGTTCCT CCTTCTACTG CCTGGAGTAC TTTCCTTCTC
 42_3a    TCAGGCCGTG GGCCGTTCCT CCTTCTACTG CCTGGAGTAC TTTCCTTCTC
  42_4    TCAGGCCGTG GGCCGTTCCT CCTTCTACTG CCTGGAGTAC TTTCCTTCTC
 42_5a    TCAGTCTGTG GGACGTTCCT CCTTCTACTG CCTGGAGTAC TTTCCTTCTC
 42_10    TCAGTCTGTG GGACGTTCCT CCTTCTACTG CCTGGAGTAC TTTCCTTCTC
 42_3b    TCAGTCTGTG GGACGTTCCT CCTTCTACTG CCTGGAGTAC TTTCCTTCTC
 42_11    TCAGTCTGTG GGACGTTCCT CCTTCTACTG CCTGGAGTAC TTTCCTTCTC
 42_6b    TCAGTCTGTG GGACGTTCCT CCTTCTACTG CCTGGAGTAC TTTCCTTCTC
  43_1    TCAGGCTGTG GGCCGTTCCT CCTTCTACTG CCTGGAATAC TTCCCTTCTC
  43_5    TCAGGCTGTG GGCCGTTCCT CCTTCTACTG CCTGGAATAC TTCCCTTCTC
 43_12    TCAGGCTGTG GGCCGTTCCT CCTTCTACTG CCTGGAATAC TTCCCTTCTC
 43_20    CCAAGCCCTG GGACGTTCCT CCTTCTACTG TCTGGAGTAT TTCCCATCGC
 43_21    CCAAGCCCTG GGACGTTCCT CCTTCTACTG TCTGGAGTAT TTCCCATCGC
 43_23    CCAAGCCCTG GGACGTTCCT CCTTCTACTG TCTGGAGTAT TTCCCATCGC
 43_25    CCAAGCCCTG GGACGTTCCT CCTTCTACTG TCTGGAGTAT TTCCCATCGC
  44_1    TCAGGCCGTG GGCCGTTCCT CCTTCTACTG CCTGGAGTAC TTTCCTTCTC
  44_5    TCAGGCCGTG GGCCGTTCCT CCTTCTACTG CCTGGAGTAC TTTCCTTCTC
223_10    CCAATCGGTA GGCCGTTCCT CCTTCTACTG CCTGGAGTAC TTTCCTTCTC
 223_2    CCAATCGGTA GGCCGTTCCT CCTTCTACTG CCTGGAGTAC TTTCCTTCTC
 223_4    CCAATCGGTA GGCCGTTCCT CCTTCTACTG CCTGGAGTAC TTTCCTTCTC
 223_5    CCAATCGGTA GGCCGTTCCT CCTTCTACTG CCTGGAGTAC TTTCCTTCTC
 223_6    CCAATCGGTA GGCCGTTCCT CCTTCTACTG CCTGGAGTAC TTTCCTTCTC
 223_7    CCAATCGGTA GGCCGTTCCT CCTTCTACTG CCTGGAGTAC TTTCCTTCTC
  A3_4    CCAAGCGGTA GGACGTTCTT CATTCTACTG TCTAGAGTAT TTTCCCTCTC
  A3_5    CCAAGCGGTA GGACGTTCTT CATTCTACTG TCTAGAGTAT TTTCCCTCTC
  A3_7    CCAAGCGGTA GGACGTTCTT CATTCTACTG TCTAGAGTAT TTTCCCTCTC
  A3_3    CCAAGCGGTA GGACGTTCTT CATTCTACTG TCTAGAGTAT TTTCCCTCTC
 42_12    TCAGGCCGTG GGCCGTTCCT CCTTCTACTG CCTGGAGTAC TTTCCTTCTC
  AAV1    CCAAGCCGTG GGACGTTCAT CCTTTTACTG CCTGGAATAT TTCCCTTCTC
  AAV2    TCAGGCAGTA GGACGCTCTT CATTTTACTG CCTGGAGTAC TTTCCTTCTC
  AAV3    TCAAGCGGTG GGACGCTCAT CCTTTTACTG CCTGGAGTAC TTCCCTTCGC
  AAV8    TCAGGCCGTG GGACGCTCCT CCTTCTACTG CCTGGAATAC TTTCCTTCGC
  AAV9    TCAAGCGTTA GGACGTTCTT CTTTCTACTG TCTGGAATAC TTCCCTTCTC
  AAV7    TCAGTCTGTG GGACGTTCCT CCTTCTACTG CCTGGAGTAC TTCCCCTCTC
  44_2    TCAGGCCGTG GGCCGTTCCT CCTTCTACTG CCTGGAGTAC TTTCCTTCTC
```

图1AAQ

```
          3451                                                       3500
  42_2    AGATGCTGAG AACGGGCAAT AACTTTGAAT TCAGCTACAC CTTTGAGGAA
  42_8    AAATGCTGAG AACGGGCAAC AACTTTGAGT TCAGCTACCA GTTTGAGGAC
 42_15    AAATGCGGAG AACGGGCAAC AACTTTGAGT TCAGCTACCA GTTTGAGGAC
 42_5b    AAATGCTGAG AACGGGCAAC AACTTTGAGT TCAGCTACCA GTTTGAGGAC
 42_1b    AAATGCTGAG AACGGGCAAC AACTTTGAGT TCAGCTACCA GTTTGAGGAC
 42_13    AAATGCTGAG AACGGGCAAC AACTTTGAGT TCAGCTACCA GTTTGAGGAC
 42_3a    AAATGCTGAG AACGGGCAAC AACTTTGAGT TCAGCTACCA GTTTGAGGAC
  42_4    AAATGCTGAG AACGGGCAAC AACTTTGAGT TCAGCTACCA GTTTGAGGAC
 42_5a    AGATGCTGAG AACGGGCAAT AACTTTGAAT TCAGCTACCA GTTTGAGGAC
 42_10    AGATGCTGAG AACGGGCAAT AACTTTGAAT TCAGCTACAC CTTTGAGGAA
 42_3b    AGATGCTGAG AACGGGCAAT AACTTTGAAT TCAGCTACAC CTTTGAGGAA
 42_11    AGATGCTGAG AACGGGCAAT AACTTTGAAT TCAGCTACAC CTTTGAGGAA
 42_6b    AGATGCTGAG AACGGGCAAT AACTTTGAAT TCAGCTACAC CTTTGAGGAA
  43_1    AAATGCTGAG GACGGGCAAC AACTTTGAAT TCAGCTACAC CTTCGAGGAC
  43_5    AAATGCTGAG GACGGGCAAC AACTTTGAAT TCAGCTACAC CTTCGAGGAC
 43_12    AAATGCTGAG GACGGGCAAC AACTTTGAAT TCAGCTACAC CTTCGAGGAC
 43_20    AGATGCTGAG AACCGGCAAC AACTTTCAGT TCAGCTACAC CTTCGAGGAC
 43_21    AGATGCTGAG AACCGGCAAC AACTTTCAGT TCAGCTACAC CTTCGAGGAC
 43_23    AGATGCCGAG AACCGGCAAC AACTTTCAGT TCAGCTACAC CTTCGAGGAC
 43_25    AGATGCTGAG AACCGGCAAC AACTTTCAGT TCAGCTACAC CTTCGAGGAC
  44_1    AAATGCTGAG AACGGGCAAC AACTTTGAGT TCAGCTACCA GTTTGAGGAC
  44_5    AAATGCTGAG AACGGGCAAC AACTTTGAGT TCAGCTACCA GTTTGAGGAC
223_10    AGATGCTGAG AACGGGCAAC AACTTCACCT TTAGCTACAC CTTCGAGGAC
 223_2    AGATGCTGAG AACGGGCAAC AACTTCACCT TTAGCTACAC CTTCGAGGAC
 223_4    AGATGCTGAG AACGGGCAAC AACTTCACCT TTAGCTACAC CTTCGAGGAC
 223_5    AGATGCTGAG AACGGGCAAC AACTTCACCT TTAGCTACAC CTTCGAGGAC
 223_6    AGATGCTGAG AACGGGCAAC AACTTCACCT TTAGCTACAC CTTCGAGGAC
 223_7    AGATGCTGAG AACGGGCAAC AACTTCACCT TTAGCTACAC CTTCGAGGAC
  A3_4    AGATGCTGAG GACGGGAAAC AACTTCACCT TCAGCTACAC TTTTGAAGAC
  A3_5    AGATGCTGAG GACGGGAAAC AACTTCACCT TCAGCTACAC TTTTGAAGAC
  A3_7    AGATGCTGAG GACGGGAAAC AACTTCACCT TCAGCTACAC TTTTGAAGAC
  A3_3    AGATGCTGAG GACGGGAAAC AACTTCACCT TCAGCTACAC TTTTGAAGAC
 42_12    AAATGCTGAG AACGGGCAAC AACTTTGAGT TCAGCTACCA GTTTGAGGAC
  AAV1    AGATGCTGAG AACGGGCAAC AACTTTACCT TCAGCTACAC CTTTGAGGAA
  AAV2    AGATGCTGCG TACCGGAAAC AACTTTACCT TCAGCTACAC TTTTGAGGAC
  AAV3    AGATGCTAAG GACTGGAAAT AACTTCCAAT TCAGCTATAC CTTCGAGGAT
  AAV8    AGATGCTGAG AACCGGCAAC AACTTCCAGT TTACTTACAC CTTCGAGGAC
  AAV9    AGATGCTGAG AACCGGCAAC AACTTTCAGT TCAGCTACAC TTTCGAGGAC
  AAV7    AGATGCTGAG AACGGGCAAC AACTTTGAGT TCAGCTACAG CTTCGAGGAC
  44_2    AAATGCTGAG AACGGGCAAC AACTTTGAGT TCAGCTACCA GTTTGAGGAC
```

图1AAR

```
          3501                                                      3550
  42_2    GTGCCTTTCC ACAGCAGCTA TGCGCACAGC CAGAGCCTGG ACCGGCTGAT
  42_8    GTGCCTTTTC ACAGCAGCTA CGCGCACAGC CAAAGCCTGG ACCGGCTGAT
 42_15    GTGCCTTTTC ACAGCAGCTA CGCGCATAGC CAAAGCCTGG ACCGGCTGAT
 42_5b    GTGCCTTTTC ACAGCAGCTA CGCGCACAGC CAAAGCCTGG ACCGGCTGAT
 42_1b    GTGCCTTTTC ACAGCAGCTA TGCGCACAGC CAAAGCCTGG ACCGGCTGAT
 42_13    GTGCCTTTTC ACAGCAGCTA TGCGCACAGC CAAAGCCTGG ACCGGCTGAT
 42_3a    GTGCCTTTTC ACAGCAGCTA CGCGCACAGC CAAAGCCTGG ACCGGCTGAT
  42_4    GTGCCTTTTC ACAGCAGCTA CGCGCACAGC CAAAGCCTGG ACCGGCTGAT
 42_5a    GTGCCCTTTC ACAGCAGCTA CGCGCACAGC CAAAGCCTGG ACCGGCTGAT
 42_10    GTGCCTTTCC ACAGCAGCTA TGCGCACAGC CAGAGCCTGG ACCGGCTGAT
 42_3b    GTGCCTTTCC ACAGCAGCTA TGCGCACAGC CAGAGCCTGG ACCGGCTGAT
 42_11    GTGCCTTTCC ACAGCAGCTA TGCGCACAGC CAGAGCCTGG ACCGGCTGAT
 42_6b    GTGCCTTTCC ACAGCAGCTA TGCGCATAGC CAGAGCCTGG ACCGGCTGAT
  43_1    GTGCCTTTCC ACAGCAGCTA CGCGCACAGC CAGAGCCTGG ACCGGCTGAT
  43_5    GTGCCTTTCC ACAGCAGCTA CGCGCACAGC CAGAGCCTGG ACCGGCTGAT
 43_12    GTGCCTTTCC ACAGCAGCTA CGCGCACAGC CAGAGCCTGG ACCGGCTGAT
 43_20    GTGCCTTTCC ACAGCAGCTA CGCGCACAGC CAGAGCCTGG ACAGGCTGAT
 43_21    GTGCCTTTCC ACAGCAGCTA CGCGCACAGC CAGAGCCTGG ACAGGCTGAT
 43_23    GTGCCTTTCC ACAGCAGCTA CGCGCACAGC CAGAGCCTGG ACAGGCTGAT
 43_25    GTGCCTTTCC ACAGCAGCTA CGCGCACAGC CAGAGCCTGG ACAGGCTGAT
  44_1    GTGCCTTTTC ACAGCAGCTA CGCGCACAGC CAAAGCCTGG ACCGGCTGAT
  44_5    GTGCCTTTTC ACAGCAGCTA CGCGCACAGC CAAAGCCTGG ACCGGCTGAT
223_10    GTGCCTTTCC ACAGCAGCTA CGCGCACAGC CAGAGTCTGG ACCGGCTGAT
 223_2    GTGCCTTTCC ACAGCAGCTA CGCGCACAGC CAGAGTCTGG ACCGGCTGAT
 223_4    GTGCCTTTCC ACAGCAGCTA CGCGCACAGC CAGAGTCTGG GCCGGCTGAT
 223_5    GTGCCTTTCC ACAGCAGCTA CGCGCACAGC CAGAGTCTGG GCCGGCTGAT
 223_6    GTGCCTTTCC ACAGCAGCTA CGCGCACAGC CAGAGTCTGG ACCGGCTGAT
 223_7    GTGCCTTTCC ACAGCAGCTA CGCGCACAGC CAGAGTCTGG ACCGGCTGAT
  A3_4    GTGCCTTTCC ACAGCAGCTA CGCGCACAGC CAGAGTCTGG ATCGGCTGAT
  A3_5    GTGCCTTTCC ACAGCAGCTA CGCGCACAGC CAGAGTCTGG ATCGGCTGAT
  A3_7    GTGCCTTTCC ACAGCAGCTA CGCGCACAGC CAGAGTCTGG ATCGGCTGAT
  A3_3    GTGCCTTTCC ACAGCAGCTA CGCGCACAGC CAGAGTCTGG ATCGGCTGAT
 42_12    GTGCCTTTTC ACAGCAGCTA CGCGCACAGC CAAAGCCTGG ACCGGCTGAC
  AAV1    GTGCCTTTCC ACAGCAGCTA CGCGCACAGC CAGAGCCTGG ACCGGCTGAT
  AAV2    GTTCCTTTCC ACAGCAGCTA CGCTCACAGC CAGAGTCTGG ACCGTCTCAT
  AAV3    GTACCTTTTC ACAGCAGCTA CGCTCACAGC CAGAGTTTGG ATCGCTTGAT
  AAV8    GTGCCTTTCC ACAGCAGCTA CGCCCACAGC CAGAGCTTGG ACCGGCTGAT
  AAV9    GTGCCTTTCC ACAGCAGCTA CGCACACAGC CAGAGTCTAG ATCGACTGAT
  AAV7    GTGCCTTTCC ACAGCAGCTA CGCACACAGC CAGAGCCTGG ACCGGCTGAT
  44_2    GTGCCTTTTC ACAGCAGCTA CGCGCACAGC CAAAGCCTGG ACCGGCTGAT
```

图1AAS

```
          3551                                                        3600
  42_2    GAATCCCCTC ATCGACCAGT ACCTGTACTA CCTGGCCCGG ACCCAGAGCA
  42_8    GAACCCCCTC ATCGACCAGT ACCTGTACTA CCTGTCTCGG ACTCAGTCCA
 42_15    GAACCCCCTC ATCGACCAGT ACCTGTACTA CCTGTCTCGG ACTCAGTCCA
 42_5b    GAACCCCCTC ATCGACCAGT ACCTGTACTA CCTGTCTCGG ACTCAGTCCA
 42_1b    GAACCCCCTC ATCGACCAGT ACCTGTACTA CCTGTCTCGG ACTCAGTCCA
 42_13    GAACCCCCTC ATCGACCAGT ACCTGTACTA CCTGTCTCGG ACTCAGTCCA
 42_3a    GAACCCCCTC ATCGACCAGT ACCTGTACTA CCTGTCTCGG ACTCAGTCCA
  42_4    GAACCCCCTC ATCGACCAGT ACCTGTACTA CCTGTCTCGG ACTCAGTCCA
 42_5a    GAACCCCCTC ATCGACCAGT ACCTGTACTA CCTGTCTCGG ACTCAGTCCA
 42_10    GAATCCCCTC ATCGACCAGT ACCTGTACTA CCTGGCCCGG ACCCAGAGCA
 42_3b    GAATCCCCTC ATCGACCAGT ACCTGTACTA CCTGGCCCGG ACCCAGAGCA
 42_11    GAATCCCCTC ATCGACCAGT ACCTGTACTA CCTGGCCCGG ACCCAGAGCA
 42_6b    GAATCCCCTC ATCGACCAGT ACCTGTACTA CCTGGCCCGG ACCCAGAGCA
  43_1    GAACCCTCTC ATCGACCAGT ACCTGTATTA CTTATCCAGA ACTCAGTCCA
  43_5    GAACCCTCTC ATCGACCAGT ACCTGTATTA CTTATCCAGA ACTCAGTCCA
 43_12    GAACCCTCTC ATCGACCAGT ACCTGTATTA CTTATCCAGA ACTCAGTCCA
 43_20    GAATCCCCTC ATCGACCAGT ACCTGTACTA CCTGGTCAGA ACGCAAACGA
 43_21    GAATCCCCTC ATCGACCAGT ACCTGTACTA CCTGGTCAGA ACGCAAACGA
 43_23    GAATCCCCTC ATCGACCAGT ACCTGTACTA CCTGGTCAGA ACGCAAACGA
 43_25    GAATCCCCTC ATCGACCAGT ACCTGTACTA CCTGGTCAGA ACGCAAACGA
  44_1    GAACCCCCTC ATCGACCAGT ACCTGTACTA CCTGTCTCGG ACTCAGTCCA
  44_5    GAACCCCCTC ATCGACCAGT ACCTGTACTA CCTGTCTCGG ACTCAGTCCA
223_10    GAATCCCCTC ATCGACCAGT ACCTGTACTA CTTGGCCAGA ACACAGAGCA
 223_2    GAATCCCCTC ATCGACCAGT ACCTGTACTA CTTGGCCAGA ACACAGAGCA
 223_4    GAATCCCCTC ATCGACCAGT ACCTGTACTA CTTGGCCAGA ACACAGAGCA
 223_5    GAATCCCCTC ATCGACCAGT ACCTGTACTA CTTGGCCAGA ACACAGAGCA
 223_6    GAATCCCCTC ATCGACCAGT ACCTGTACTA CTTGGCCAGA ACACAGAGCA
 223_7    GAATCCCCTC ATCGACCAGT ACCTGTACTA CTTGGCCAGA ACACAGAGCA
  A3_4    GAATCCTCTC ATTGACCAGT ACCTGTATTA CCTGAGCAAA ACTCAGGGTA
  A3_5    GAATCCTCTC ATTGACCAGT ACCTGTATTA CCTGAGCAAA ACTCAGGGTA
  A3_7    GAATCCTCTC ATTGACCAGT ACCTGTATTA CCTGAGCAAA ACTCAGGGTA
  A3_3    GAATCCTCTC ATTGACCAGT ACCTGTATTA CCTGAGCAAA ACTCAGGGTA
 42_12    GAACCCCCTC ATCGACCAGT ACCTGTACTA CCTGGCCCGG ACCCAGAGCA
   AAV1   GAATCCTCTC ATCGACCAAT ACCTGTATTA CCTGAACAGA ACTCAAA.AT
   AAV2   GAATCCTCTC ATCGACCAGT ACCTGTATTA CTTGAGCAGA ACAAACACTC
   AAV3   GAATCCTCTT ATTGATCAGT ATCTGTACTA CCTGAACAGA ACGCAAGGAA
   AAV8   GAATCCTCTG ATTGACCAGT ACCTGTACTA CTTGTCTCGG ACTCAAACAA
   AAV9   GAACCCCCTC ATCGACCAGT ACCTATACTA CCTGGTCAGA ACACAGACAA
   AAV7   GAATCCCCTC ATCGACCAGT ACTTGTACTA CCTGGCCAGA ACACAGAGTA
  44_2    GAACCCCCTC ATCGACCAGT ACCTGTACTA CCTGTCTCGG ACTCAGTCCA
```

图1AAT

```
          3601                                                    3650
  42_2    CTACGG..GG TCCACAAGGG AGCTGCA.GT TCCA...... TCAGGCTGGG
  42_8    CGGGA...GG TACCGCAGGA ACTCAGCAGT TGCTATTTTC TCAGGCCGGG
 42_15    CGGGA...GG TACCGCAGGA ACTCAGCAGT TGCTATTTTC TCAGGCCGGG
 42_5b    CGGGA...GG TACCGCAGGA ACTCAGCAGT TGCTATTTTC TCAGGCCGGG
 42_1b    CGGGA...GG TACCGCAGGA ACTCAGCAGT TGCTATTTTC TCAGGCCGGG
 42_13    CGGGA...GG TACCGCAGGA ACTCAGCAGT TGCTATTTTC TCAGGCCGGG
 42_3a    CGGGA...GG TACCGCAGGA ACTCAGCAGT TGCTATTTTC TCAGGCCGGG
  42_4    CGGGA...GG TACCGCAGGA ACTCAGCAGT TGCTATTTTC TCAGGCCGGG
 42_5a    CGGGA...GG TACCGCAGGA ACTCAGCAGT TGCTATTTTC TCAGGCCGGG
 42_10    CTACG...GG GTCCACAAGG GAGCTGCAGT TCCA...... TCAGGCTGGG
 42_3b    CTACG...GG GTCCACAAGG GAGCTGCAGT TCCA...... TCAGGCTGGG
 42_11    CTACG...GG GTCCACAAGG GAGCTGCAGT TCCA...... TCAGGCTGGG
 42_6b    CTACG...GG GTCCACAAGG GAGCTGCAGT TCCA...... TCAGGCTGGG
  43_1    CAGGA...GG AACTCAAGGT ACTCAGCAAT TGTTATTTTC TCAAGCCGGG
  43_5    CAGGA...GG AACTCAAGGT ACTCAGCAAT TGTTATTTTC TCAAGCCGGG
 43_12    CAGGA...GG AACTCAAGGT ACTCAGCAAT TGTTATTTTC TCAAGCCGGG
 43_20    CT......GG AACTGGAGGG ACGCAGACTC TGGCATTCAG CCAAGCGGGT
 43_21    CT......GG AACTGGAGGG ACGCAGACTC TGGCATTCAG CCAAGCGGGT
 43_23    CT......GG AACTGGAGGG ACGCAGACTC TGGCATTCAG CCAAGCGGGT
 43_25    CT......GG AACTGGAGGG ACGCAGACTC TGGCATTCAG CCAAGCGGGT
  44_1    CGGGA...GG TACCGCAGGA ACTCAGCAGT TGCTATTTTC TCAGGCCGGG
  44_5    CGGGA...GG TACCGCAGGA ACTCAGCAGT TGCTATTTTC TCAGGCCGGG
223_10    ACGCAGGAGG TACTGCTGGC AATCGGGAAC TGCAGTTTTA TCAGGGCGGA
 223_2    ACGCAGGAGG TACTGCTGGC AATCGGGAAC TGCAGTTTTA TCAGGGCGGA
 223_4    ACGCAGGAGG TACTGCTGGC AATCGGGAAC TGCAGTTTTA TCAGGGCGGA
 223_5    ACGCAGGAGG TACTGCTGGC AATCGGGAAC TGCAGTTTTA TCAGGGCGGA
 223_6    ACGCAGGAGG TACTGCTGGC AATCGGGAAC TGCAGTTTTA TCAGGGCGGA
 223_7    ACGCAGGAGG TACTGCTGGC AATCGGGAAC TGCAGTTTTA TCAGGGCGGA
  A3_4    CAAG...TGG AACAACGCAG CAATCGAGAC TGCAGTTCAG CCAAGCTGGG
  A3_5    CAAG...TGG AACAACGCAG CAATCGAGAC TGCAGTTCAA CCAAGCTGGG
  A3_7    CAAG...TGG AACAACGCAG CAATCGAGAC TGCAGTTCAG CCAAGCTGGG
  A3_3    CAAG...TGG AACAACGCAG CAATCGAGAC TGCAGTTCAG CCAAGCTGGG
 42_12    CTACG...GG GTCCACAAGG GGGCTGCAGT TCCA...... TCAGGCTGGG
   AAV1   CAGTCC..GG AAGTGCCCAA AACAAGGACT TGCTGTTTAG CCGTGGGTCT
   AAV2   CAAG...TGG AACCACCACG CAGTCAAGGC TTCAGTTTTC TCAGGCCGGA
   AAV3   CAACCTCTGG AACAACCAAC CAATCACGGC TGCTTTTTAG CCAGGCTGGG
   AAV8   CAGGAG..GC .ACGGCAAAT ACGCAGACTC TGGGCTTCAG CCAAGGTGGG
   AAV9   CTGGA..... .ACTGGGGGA ACTCAAACTT TGGCATTCAG CCAAGCAGGC
   AAV7   ACCCAGGAGG CACAGCTGGC AATCGGGAAC TGCAGTTTTA CCAGGGCGGG
  44_2    CGGGA...GG TACCGCAGGA ACTCAGCAGT TGCTATTTTC TCAGGCCGGG
```

图1AAU

```
          3651                                                    3700
  42_2    CCCAACACCA TGGCCGAGCA ATCAAAGAAC TGGCTGCCCG GACCCTGTTA
  42_8    CCTAATAACA TGTCGGCTCA GGCCAAAAAC TGGCTACCCG GGCCCTGCTA
 42_15    CCTAATAACA TGTCGGCTCA GGCCAAAAAC TGGCTACCCG GGCCCTGCTA
 42_5b    CCTAATAACA TGTCGGCTCA GGCCAAAAAC TGGCTACCCG GGCCCTGCTA
 42_1b    CCTAATAACA TGTCGGCTCA GGCCAAAAAC TGGCTACCCG GGCCCTGCTA
 42_13    CCTAATAACA TGTCGGCTCA GGCCAAAAAC TGGCTACCCG GGCCCTGCTA
 42_3a    CCTAATAACA TGTCGGCTCA GGCCAAAAAC TGGCTACCCG GGCCCTGCTA
  42_4    CCTAATAACA TGTCGGCTCA GGCCAAAAAC TGGCTACCCG GGCCCTGCTA
 42_5a    CCTAATAACA TGTCGGCTCA GGCCAAAAAC TGGCTACCCG GGCCCTGCTA
 42_10    CCCAACACCA TGGCCGAGCA ATCAAAGAAC TGGCTGCCCG GACCCTGTTA
 42_3b    CCCAACACCA TGGCCGAGCA ATCAAAGAAC TGGCTGCCCG GACCCTGTTA
 42_11    CCCAACACCA TGGCCGAGCA ATCAAAGAAC TGGCTGCCCG GACCCTGTTA
 42_6b    CCCAACACCA TGGCCGAGCA ATCAAAGAAC TGGCTGCCCG GACCCTGTTA
  43_1    CCCGCAAACA TGTCGGCTCA GGCCAAGAAC TGGCTACCTG GACCGTGTTA
  43_5    CCCGCAAACA TGTCGGCTCA GGCCAAGAAC TGGCTACCTG GACCGTGTTA
 43_12    CCCGCAAACA TGTCGGCTCA GGCCAAGAAC TGGCTACCTG GACCGTGTTA
 43_20    CCTAGCTCAA TGGCCAACCA GGCTAGAAAT TGGGTGCCCG GACCTTGCTA
 43_21    CCTAGCTCAA TGGCCAACCA GGCTAGAAAT TGGGTGCCCG GACCTTGCTA
 43_23    CCTAGCTCAA TGGCCAACCA GGCTAGAAAT TGGGTGCCCG GACCTTGCTA
 43_25    CCTAGCTCAA TGGCCAACCA GGCTAGAAAT TGGGTGCCCG GACCTTGCTA
  44_1    CCTAATAACA TGTCGGCTCA GGCCAAAAAC TGGCTACCCG GGCCCTGCTA
  44_5    CCTAATAACA TGTCGGCTCA GGCCAAAAAC TGGCTACCCG GGCCCTGCTA
223_10    CCTACCACCA TGGCCGAACA AGCAAAGAAC TGGCTGCCCG GACCTTGCTT
 223_2    CCTACCACCA TGGCCGAACA AGCAAAGAAC TGGCTGCCCG GACCTTGCTT
 223_4    CCTACCACCA TGGCCGAACA AGCAAAGAAC TGGCTGCCCG GACCTTGCTT
 223_5    CCTACCACCA TGGCCGAACA AGCAAAGAAC TGGCTGCCCG GACCTTGCTT
 223_6    CCTACCACCA TGGCCGAACA AGCAAAGAAC TGGCTGCCCG GACCTTGCTT
 223_7    CCTACCACCA TGGCCGAACA AGCAAAGAAC TGGCTGCCCG GACCTTGCTT
  A3_4    CCTAGCTCCA TGGCTCAGCA GGCCAAAAAC TGGCTACCGG GACCCAGCTA
  A3_5    CCTAGCTCCA TGGCTCAGCA GGCCAAAAAC TGGCTACCGG GACCCAGCTA
  A3_7    CCTAGCTCCA TGGCTCAGCA GGCCAAAAAC TGGCTACCGG GACCCAGCTA
  A3_3    CCTAGCTCCA TGGCTCAGCA GGCCAAAAAC TGGCTACCGG GACCCAGCTA
 42_12    CCCAACACCA TGGCCGAGCA ATCAAAGAAC TGGCTGCCCG GACCCTGTTA
  AAV1    CCAGCTGGCA TGTCTGTTCA GCCCAAAAAC TGGCTACCTG GACCCTGTTA
  AAV2    GCGAGTGACA TTCGGGACCA GTCTAGGAAC TGGCTTCCTG GACCCTGTTA
  AAV3    CCTCAGTCTA TGTCTTTGCA GGCCAGAAAT TGGCTACCTG GGCCCTGCTA
  AAV8    CCTAATACAA TGGCCAATCA GGCAAAGAAC TGGCTGCCAG GACCCTGTTA
  AAV9    CCTAGCTCAA TGGCCAATCA GGCTAGAAAC TGGGTACCCG GGCCTTGCTA
  AAV7    CCTTCAACTA TGGCCGAACA AGCCAAGAAT TGGTTACCTG GACCTTGCTT
  44_2    CCTAATAACA TGTCGGCTCA GGCCAAAAAC TGGCTACCCG GGCCCTGCTA
```

图1AAV

	3701				3750
42_2	TCGGCAGCAG	AGACTGTCAA	AAAACATAGA	CAGCAACAAC	AACAGTAACT
42_8	CCGGCAGCAA	CGCGTCTCCA	CGACACTGTC	GCAAAATAAC	AACAGCAACT
42_15	CCGGCAGCAA	CGCGTCTCCA	CGACACTGTC	GCAAAATAAC	AACAGCAACT
42_5b	CCGGCAGCAA	CGCGTCTCCA	CGACACTGTC	GCAAAATAAC	AACAGCAACT
42_1b	CCGGCAGCAA	CGCGTCTCCA	CGACAGTGTC	GCAAAATAAC	AACAGCAACT
42_13	CCGGCAGCAA	CGCGTCTCCA	CGACAGTGTC	GCAAAATAAC	AACAGCAACT
42_3a	CCGGCAGCAA	CGCGTCTCCA	CGACACTGTC	GCAAAATAAC	AACAGCAACT
42_4	CCGGCAGCAA	CGCGTCTCCA	CGACACTGTC	GCAAAATAAC	AACAGCAACT
42_5a	CCGGCAGCAA	CGCGTCTCCA	CGACACTGTC	GCAAAATAAC	AACAGCAACT
42_10	TCGGCAGCAG	AGACTGTCAA	AAAACATAGA	CAGCAACAAC	AACAGTAACT
42_3b	TCGGCAGCAG	AGACTGTCAA	AAAACATAGA	CAGCAACAAC	ACCAGTAACT
42_11	TCGGCGGCAG	AGACTGTCAA	AAGACATAGA	CAGCAACAAC	AACAGTAACT
42_6b	TCGGCAGCAG	AGACTGTCAA	AAAACATAGA	CAGCAACAAC	AACAGTAACT
43_1	CCGTCAGCAA	CGAGTTTCCA	CGACACTGTC	GCAAAACAAC	AACAGCAATT
43_5	CCGTCAGCAA	CGAGTTTCCA	CGACACTGTC	GCAAAACAAC	AACAGCAATT
43_12	CCGTCAGCAA	CGAGTTTCCA	CGACACTGTC	GCAAAACAAC	AACAGCAATT
43_20	CCGGCAGCAG	CGCGTCTCCA	CGACAACCAA	CCAGAACAAC	AACAGCAACT
43_21	CCGGCAGCAG	CGCGTCTCCA	CGACAACCAA	CCAGAGCAAC	AACAGCAACT
43_23	CCGGCAGCAG	CGCGTCTCCA	CGACAACCAA	CCAGAACAAC	AACAGCAACT
43_25	CCGGCAGCAG	CGCGTCTCCA	CGACAACCAA	CCAGAACAAC	AACAGCAACT
44_1	CCGGCAGCAA	CGCGTCTCCA	CGACACTGTC	GCAAAATAAC	AACAGCAACT
44_5	CCGGCAGCAA	CGCGTCTCCA	CGACACTGTC	GCAAAATAAC	AACAGCAACT
223_10	CCGGCAACAG	AGAGTATCCA	AGACGCTGGA	TCAAAATAAC	AACAGCAACT
223_2	CCGGCAACAG	AGAGTATCCA	AGACGCTGGA	TCAAAATAAC	AACAGCAACT
223_4	CCGGCAACAG	AGAGTATCCA	AGACGCTGGA	TCAAAATAAC	AACAGCAACT
223_5	CCGGCAACAG	AGAGTATCCA	AGACGCTGGA	TCAAAATAAC	AACAGCAACT
223_6	CCGGCAACAG	AGAGTATCCA	AGACGCTGGA	TCAAAATAAC	AACAGCAACT
223_7	CCGGCAACAG	AGAGTATCCA	AGACGCTGGA	TCAAAATAAC	AACAGCAACT
A3_4	CCGACAGCAG	CGAATGTCTA	AGACGGCTAA	TGACAACAAC	AACAGTGAAT
A3_5	CCGACAGCAG	CGAATGTCTA	AGACGGCTAA	TGACAACAAC	AACAGTGAAT
A3_7	CCGACAGCAG	CGAATGTCTA	AGACGGCTAA	TGACAACAAC	AACAGTGAAT
A3_3	CCGACAGCAG	CGAATGTCTA	AGACGGCTAA	TGACAACAAC	AACAGTGAAT
42_12	TCGGCAGCAG	AGACTGTCAA	AAAACATAGA	CAGCAACAAC	AACAGTAACT
AAV1	TCGGCAGCAG	CGCGTTTCTA	AAACAAAAAC	AGACAACAAC	AACAGCAATT
AAV2	CCGCCAGCAG	CGAGTATCAA	AGACATCTGC	GGATAACAAC	AACAGTGAAT
AAV3	CCGGCAACAG	AGACTTTCAA	AGACTGCTAA	CGACAACAAC	AACAGTAACT
AAV8	CCGCCAACAA	CGCGTCTCAA	CGACAACCGG	GCAAAACAAC	AATAGCAACT
AAV9	CCGTCAGCAG	CGCGTCTCCA	CAACCACCAA	CCAAAATAAC	AACAGCAACT
AAV7	CCGGCAACAA	AGAGTCTCCA	AAACGCTGGA	TCAAAACAAC	AACAGCAACT
44_2	CCGGCAGCAA	CGCGTCTCCA	CGACACTGTC	GCAAAATAAC	AACAGCAACT

图1AAW

```
          3751                                                      3800
  42_2    TTGCCTGGAC CGGGGCCACT AAATACCATC TGAATGGTAG AAATTCATTA
  42_8    TTGCTTGGAC CGGTGCCACC AAGTATCATC TGAATGGCAG AGACTCTCTG
 42_15    TTGCTTGGAC CGGTGCCACC AAGTATCATC TGAATGGCAG AGACTCTCTG
 42_5b    TTGCTTGGAC CGGTGCCACC AAGTATCATC TGAATGGCAG AGACTCTCTG
 42_1b    TTGCTTGGAC CGGTGCCACC AAGTATCATC TGAATGGCAG AGACTCTCTG
 42_13    TTGCTTGGAC CGGTGCCACC AAGTATCATC TGAATGGCAG AGACTCTCTG
 42_3a    TTGCTTGGAC CGGTGCCACC AAGTATCATC TGAATGGCAG AGACTCTCTG
  42_4    TTGCTTGGAC CGGTGCCACC AAGTATCATC TGAATGGCAG AGACTCTCTG
 42_5a    TTGCTTGGAC CGGTGCCACC AAGTATCATC TGAATGGCAG AGACTCTCTG
 42_10    TTGCCTGGAC CGGGGCCACT AAATACCATC TGAATGGTAG AAATTCATTA
 42_3b    TTGCCTGGAC CGGGGCCACT AAATACCATC TGAATGGTAG AAATTCATTA
 42_11    TTGCCTGGAC CGGGGCCACT AAATACCATC TGAATGGTAG AAATTCATTA
 42_6b    TTGCCTGGAC CGGGGCCACT AAATACCATC TGAATGGTAG AAATTCATTA
  43_1    TTGCTTGGAC CGGTGCCACC AAGTATCACC TGAATGGCAG AGACTCCCTG
  43_5    TTGCTTGGAC CGGTGCCACC AAGTATCACC TGAATGGCAG AGACTCCCTG
 43_12    TTGCTTGGAC CGGTGCCACC AAGTATCACC TGAATGGCAG AGACTCCCTG
 43_20    TTGCCTGGAC GGGAGCTGCC AAGTTTAAGC TGAACGGCCG AGACTCTCTA
 43_21    TTGCCTGGAC GGGAGCTGCC AAGTTTAAGC TGAACGGCCG AGACTCTCTA
 43_23    TTGCCTGGAC GGGAGCTGCC AAGTTTAAGC TGAACGGCCG AGACTCTCTA
 43_25    TTGCCTGGAC GGGAGCTGCC AAGTTTAAGC TGAACGGCCG AGACTCTCTA
  44_1    TTGCCTGGAC CGGTGCCACC AAGTATCATC TGAATGGCAG AGACTCTCTG
  44_5    TTGCCTGGAC CGGTGCCACC AAGTATCATC TGAATGGCAG AGACTCTCTG
223_10    TTGCCTGGAC TGGTGCCACA AAATACCATT TAAATGNAAG AAATTCATTG
 223_2    TTGCCTGGAC TGGTGCCACA AAATACCATT TAAATGGAAG AAATTCATTG
 223_4    TTGCCTGGAC TGGTGCCACA AAATACCATT TAAATGGAAG AAATTCATTG
 223_5    TTGCCTGGAC TGGTGCCACA AAATACCATT TAAATGGAAG AAATTCATTG
 223_6    TTGCCTGGAC TGGTGCCACA AAATACCATT TAAATGGAAG AAATTCATTG
 223_7    TTGCCTGGAC TGGTGCCACA AAATACCATT TAAATGGAAG AAATTCATTG
  A3_4    TTGCTTGGAC TGCAGCCACC AAATATTACC TGAATGGAAG AAATTCTCTG
  A3_5    TTGCTTGGAC TGCAGCCACC AAATATTACC CGAATGGAAG AAATTCTCTG
  A3_7    TTGCTTGGAC TGCAGCCACC AAATATTACC TGAATGGAAG AAATTCTCTG
  A3_3    TTGCTTGGAC TGCAGCCACC AAATATTACC TGAATGGAAG AAATTCTCTG
 42_12    TTGCCTGGAC CGGGGCCACT AAATACCATC TGAATGGTAG AAATTCATTA
  AAV1    TTACCTGGAC TGGTGCTTCA AAATATAACC TCAATGGGCG TGAATCCATC
  AAV2    ACTCGTGGAC TGGAGCTACC AAGTACCACC TCAATGGCAG AGACTCTCTG
  AAV3    TTCCTTGGAC AGCGGCCAGC AAATATCATC TCAATGGCCG CGACTCGCTG
  AAV8    TTGCCTGGAC TGCTGGGACC AAATACCATC TGAATGGAAG AAATTCATTG
  AAV9    TTGCGTGGAC GGGAGCTGCT AAATTCAAGC TGAACGGGAG AGACTCGCTA
  AAV7    TTGCTTGGAC TGGTGCCACC AAATATCACC TGAACGGCAG AAACTCGTTG
  44_2    TTGCCTGGAC CGGTGCCACC AAGTATCATC TGAATGGCAG AGACTCTCTG
```

图1AAX

```
          3801                                                      3850
  42_2    ACCAACCCGG GCGTAGCCAT GGCCACCAAC AAGGACGACG AGGACCAGTT
  42_8    GTAAATCCCG GTGTCGCTAT GGCAACGCAC AAGGACGACG AAGAGCGATT
 42_15    GTAAATCCCG GTGTCGCTAT GGCAACGCAC AAGGACGACG AAGAGCGATT
 42_5b    GTAAATCCCG GTGTCGCTAT GGCAACGCAC AAGGACGACG AAGAGCGATT
 42_1b    GTAAATCCCG GTGTCGCTAT GGCAACGCAC AAGGGCGACG AAGAGCGATT
 42_13    GTAAATCCCG GTGTCGCTAT GGCAACGCAC AAGGGCGACG AAGAGCGATT
 42_3a    GTAAATCCCG GTGTCGCTAT GGCAACGCAC AAGGACGACG AAGAGCGATT
  42_4    GTAAATCCCG GTGTCGCTAT GGCAACGCAC AAGGACGACG AAGAGCGATT
 42_5a    GTAAATCCCG GTGTCGCTAT GGCAACGCAC AAGGACGACG AAGAGCGATT
 42_10    ACCAACCCGG GCGTAGCCAT GGCCACCAAC AAGGACGACG AGGACCAGTT
 42_3b    ACCAACCCGG GCGTAGCCAT GGCCACCAAC AAGGACGACG AGGACCAGTT
 42_11    ACCAACCCGG GCGTAGCCAT GGCCACCAAC AAGGACGACG AGGACCAGTT
 42_6b    ACCAACCCGG GCGTAGCCAT GGCCACCAAC AAGGACGACG AGGACCAGTT
  43_1    GTTAATCCCG GCGTTGCCAT GGCTACCCAC AAGGACGACG AGGAGCGCTT
  43_5    GTTAATCCCG GCGTTGCCAT GGCTACCCAC AAGGACGACG AGGAGCGCTT
 43_12    GTTAATCCCG GCGTTGCCAT GGCTACCCAC AAGGACGACG AGGAGCGCTT
 43_20    ATGAATCCGG GCGTGGCAAT GGCTTCCCAC AAGGATGACG ACGACCGCTT
 43_21    ATGAATCCGG GCGTGGCAAT GGCTTCCCAC AAGGATGACG ACGACCGCTT
 43_23    ATGAATCCGG GCGTGGCAAT GGCTTCCCAC AAGGATGACG ACGACCGCTT
 43_25    ATGAATCCGG GCGTGGCAAT GGCTTCCCAC AAGGATGACG ACGACCGCTT
  44_1    GTAAATCCCG GTGTCGCTAT GGCAACCCAC AAGGACGACG AAGAGCGATT
  44_5    GTAAATCCCG GTGTCGCTAT GGCAACCCAC AAGGACGACG AAGAGCGATT
223_10    GTTAATCCCG GTGTCGCCAT GGCAACCCAC AAGGACGACG AGGAACGCTT
 223_2    GTTAATCCCG GTGTCGCCAT GGCAACCCAC AAGGACGACG AGGAACGCTT
 223_4    GTTAATCCCG GTGTCGCCAT GGCAACCCAC AAGGACGACG AGGAACGCTT
 223_5    GTTAATCCCG GTGTCGCCAT GGCAACCCAC AAGGACGACG AGGAACGCTT
 223_6    GTTAATCCCG GTGTCGCCAT GGCAACCCAC AAGGACGACG AGGAACGCTT
 223_7    GTTAATCCCG GTGTCGCCAT GGCAACCCAC AAGGACGACG AGGAACGCTT
  A3_4    GTCAATCCCG GGCCCCCAAT GGCCAGTCAC AAGGACGATG AGGAAAAGTA
  A3_5    GTCAATCCCG GGCCCCCAAT GGCCAGTCAC AAGGACGATG AGGAAAAGTA
  A3_7    GTCAATCCCG GGCCCCCAAT GGCCAGTCAC AAGGACGATG AGGAAAAGTA
  A3_3    GTCAATCCCG GGCCCCCAGT GGCCAGTCAC AAGGACGATG AGGAAAAGTA
 42_12    ACCAACCCGG GCGTAGCCAT GGCCACCAAC AAGGACGACG AGGACCAGTT
  AAV1    ATCAACCCTG GCACTGCTAT GGCCTCACAC AAAGACGACG AAGACAAGTT
  AAV2    GTGAATCC.. GGCC....AT GGCAAGCCAC AAGGACGATG AAGAAAAGTT
  AAV3    GTGAATCCAG GACCAGCTAT GGCCAGTCAC AAGGACGATG AAGAAAAATT
  AAV8    GCTAATCCTG GCATCGCTAT GGCAACACAC AAAGACGACG AGGAGCGTTT
  AAV9    ATGAATCCTG GCGTGGCTAT GGCATCGCAC AAAGACGACG AGGACCGCTT
  AAV7    GTTAATCCCG GCGTCGCCAT GGCAACTCAC AAGGACGACG AGGACCGCTT
  44_2    GTAAATCCCG GTGTCGCTAT GGCAACCCAC AAGGACGACG AAGAGCGATT
```

图1AAY

```
          3851                                                    3900
  42_2    CTTTCCCATC AACGGAGTGC TGGTTTTTGG CGAAACGGGG GCTGCCAACA
  42_8    TTTTCCATCC AGCGGAGTCT TGATGTTTGG GAAACAGGGA GCTGGAAA..
 42_15    TTTTCCATCC AGCGGAGTCT TGATGTTTGG GAAACAGGGA GCTGGAAA..
 42_5b    TTTTCCATCC AGCGGAGTCT TGATGTTTGG GAAACAGGGA GCTGGAAA..
 42_1b    TTTTCCATCC AGCGGAGTCT TGATGTTTGG GAAACAGGGA GCTGGAAA..
 42_13    TTTTCCATCC AGCGGAGTCT TGATGTTTGG GAAACAGGGA GCTGGAAA..
 42_3a    TTTTCCATCC AGCGGAGTCT TGATGTTTGG GAAACAGGGA GCTGGAAA..
  42_4    TTTTCCATCC AGCGGAGTCT TGATGTTTGG GAAACAGGGA GCTGGAAA..
 42_5a    TTTTCCATCC AGCGGAGTCT TGATGTTTGG GAAACAGGGA GCTGGAAA..
 42_10    CTTTCCCATC AACGGAGTGC TGGTTTTTGG CAAAACGGGG GCTGCCAACA
 42_3b    CTTTCCCATC AACGGAGTGC TGGTTTTTGG CAAAACGGGG GCTGCCAACA
 42_11    CTTTCCCATC AACGGAGTGC TGGTTTTTGG CAAAACGGGG GCTGCCAACA
 42_6b    CTTTCCCATC AACGGAGTGC TGGTTTTTGG CAAAACGGGG GCTGCCAACA
  43_1    CTTCCCGTCA AGCGGAGTTC TAATGTTTGG CAAGCAGGGG GCTGGAAA..
  43_5    CTTCCCGTCA AGCGGAGTTC TAATGTTTGG CAAGCAGGGG GCTGGAAA..
 43_12    CTTCCCGTCA AGCGGAGTTC TAATGTTTGG CAAGCAGGGG GCTGGAAA..
 43_20    CTTCCCTTCG AGCGGGGTCC TGATTTTTGG CAAGCAAGGA GCCGGGAA..
 43_21    CTTCCCTTCG AGCGGGGTCC TGATTTTTGG CAAGCAAGGA GCCGGGAA..
 43_23    CTTCCCTTCG AGCGGGGTCC TGATTTTTGG CAAGCAAGGA GCCGGGAA..
 43_25    CTTCCCTTCG AGCGGGGTCC TGATTTTTGG CAAGCAAGGA GCCGGGAA..
  44_1    TTTTCCGTCC AGCGGAGTCT TAATGTTTGG GAAACAGGGA GCTGGAAA..
  44_5    TTTTCCGTCC AGCGGAGTCT TAATGTTTGG GAAACAGGGA GCTGGAAA..
223_10    CTTCCCTTCG AGCGGAGTTC TAATTTTTGG CAAAACTGGA GCAGCTAATA
 223_2    CTCCCCTTCG AGCGGAGTTC TAATTTTTGG CAAAACTGGA GCAGCTAATA
 223_4    CTTCCCTTCG AGCGGAGTTC TAATTTTTGG CAAAACTGGA GCAGCTAATA
 223_5    CTTCCCTTCG AGCGGAGTTC TAATTTTTGG CAAAACTGGA GCAGCTAATA
 223_6    CTTCCCTTCG AGCGGAGTTC TAATTTTTGG CAAAACTGGA GCAGCTAATA
 223_7    CTTCCCTTCG AGCGGAGTTC TAATTTTTGG CAAAACTGGA GCAGCTAATA
  A3_4    TTTCCCCATG CACGGAAATC TCATCTTTGG AAAACAAGGC ACAGGAAC..
  A3_5    TTTCCCCATG CACGGAAATC TCATCTTTGG AAAACAAGGC ACAGGAAC..
  A3_7    TTTCCCCATG CACGGAAATC TCATCTTTGG AAAACAAGGC ACAGGAAC..
  A3_3    TTTCCCCATG CACGGAAATC TCATCTTTGG AAAACAAGGC ACAGGAAC..
 42_12    CTTTCCCATC AACGGAGTGC TGGTTTTTGG CAAAACGGGG GCTGCCAACA
  AAV1    CTTTCCCATG AGCGGTGTCA TGATTTTTGG AAAAGAGAGC GCCGGAGC..
  AAV2    TTTTCCTCAG AGCGGGGTTC TCATCTTTGG GAAGCAAGGC TCAGAGAA..
  AAV3    TTTCCCTATG CACGGCAATC TAATATTTGG CAAAGAAGGG ACAACGGC..
  AAV8    TTTTCCCAGT AACGGGATCC TGATTTTTGG CAAACAAAAT GCTGCCAG..
  AAV9    CTTTCCATCA AGTGGCGTTC TCATATTTGG CAAGCAAGGA GCCGGGAA..
  AAV7    TTTCCCATCC AGCGGAGTCC TGATTTTTGG AAAAACTGGA GCAACTAACA
  44_2    TTTTCCGTCC AGCGGAGTCT TAATGTTTGG GAAACAGGGA GCTGGAAA..
```

图1AAZ

```
         3901                                                    3950
  42_2   AGACAACGCT GGAA...... AACGTGCTAA TGACCAGCGA GGAGGAGATC
  42_8   AGACAACG.T GGACTATAGC AGCGTTATGC TAACCAGTGA GGAAGAAATC
 42_15   AGACAACG.T GGACTATAGC AGCGTTATGC TAACCAGTGA GGAAGAAATC
 42_5b   AGACAACG.T GGACTATAGC AGCGTTATGC TAACCAGTGA GGAAGAAATC
 42_1b   AGACAACG.T AGACTATAGC AGCGTTATGC TAACCAGTGA GGAAGAAATC
 42_13   AGACAACG.T GGACTATAGC AGCGTTATGC TAACCAGTGA GGAAGAAATC
 42_3a   AGACAACG.T GGACTATAGC AGCGTTATGC TAACCAGTGA GGAAGAAATC
  42_4   AGACAACG.T GGACTATAGC AGCGTTATGC TAACCAGTGA GGAAGAAATC
 42_5a   AGACAACG.T GGACTATAGC AGCGTTATGC TAACCAGTGA GGAAGAAATC
 42_10   AGACAACGCT GGAA...... AACGTGCTAA TGACCAGCGA GGAGGAGATC
 42_3b   AGACAACGCT GGAA...... AACGTGCTAA TGACCAGCGA GGAGGAGATC
 42_11   AGACAACGCT GGAA...... AACGTGCTAA TGACCAGCGA GGAGGAGATC
 42_6b   AGACAACGCT GGAA...... AACGTGCTAA TGACCAGCGA GGAGGAGATC
  43_1   AGACAATG.T GGACTACAGC AGCGTGATGC TCACCAGCGA AGAAGAAATT
  43_5   AGACAATG.T GGACTACAGC AGCGTGATGC TCACCAGCGA AGAAGAAATT
 43_12   AGACAATG.T GGACTACAGC AGCGTGATGC TCACCAGCGA AGAAGAAATT
 43_20   CGATGGAG.T GGATTACAGC CAAGTGCTGA TTACAGATGA GGAAGAAATC
 43_21   CGATGGAG.T GGATTACAGC CAAGTGCTGA TTACAGATGA GGAAGAAATC
 43_23   CGATGGAG.T GGATTACAGC CAAGTGCTGA TTACAGATGA GGAAGAAATC
 43_25   CGATGGAG.T GGATTACAGC CAAGTGCTGA TTACAGATGA GGAAGAAATC
  44_1   AGACAACG.T GGACTATAGC AGCGTTATGC TAACCAGTGA GGAAGAAATT
  44_5   AGACAACG.T GGACTATAGC AGCGTTATGC TAACCAGTGA GGAAGAAATT
223_10   AAACTACATT AGAA...... AACGTGCTCA TGACAAATGA AGAAGAAATT
 223_2   AAACTACATT AGAA...... AACGTGCTCA TGACAAATGA AGAAGAAATT
 223_4   AAACTACATT AGAA...... AACGTGCTCA TGACAAATGA AGAAGAAATT
 223_5   AAACTACATT AGAA...... AACGTGCTCA TGACAAATGA AGAAGAAATT
 223_6   AAACTACATT AGAA...... AACGTGCTCA TGACAAATGA AGAAGAAATT
 223_7   AAACTACATT AGAA...... AACGTGCTCA TGACAAATGA AGAAGAAATT
  A3_4   TACCAATG.T GGACATTGAA TCAGTGCTTA TTACAGACGA AGAAGAAATC
  A3_5   TACCAATG.T GGACATTGAA TCAGTGCTTA TTACAGACGA AGAAGAAATC
  A3_7   TACCAATG.T GGACATTGAA TCAGTGCTTA TTACAGACGA AGAAGAAATC
  A3_3   TACCAATG.T GGACATTGAA TCAGTGCTTA TTACAGACGA AGAAGAAATC
 42_12   AGACAACGCT GGAA...... AACGTGCTAA TGACCAGCGA GGAGGAGATC
  AAV1   TTCAAACA.C TGCATTGGAC AATGTCATGA TTACAGACGA AGAGGAAATT
  AAV2   AACAAATG.T GAACATTGAA AAGGTCATGA TTACAGACGA AGAGGAAATC
  AAV3   AAGTAACG.C AGAATTAGAT AATGTAATGA TTACGGATGA AGAAGAGATT
  AAV8   AGACAATG.C GGATTACAGC GATGTCATGC TCACCAGCGA GGAAGAAATC
  AAV9   CGATGGAG.T CGACTACAGC CAGGTGCTGA TTACAGATGA GGAAGAAATT
  AAV7   AAACTACATT GGAA...... AATGTGTTAA TGACAAATGA AGAAGAAATT
  44_2   AGACAACG.T GGACTATAGC AGCGTTATGC TAACCAGTGA GGAAGAAATT
```

图1AAAA

```
          3951                                                    4000
  42_2    AAAACCACCA ATCCCGTGGC TACAGAAGAA TACGGTGTGG TCTCCAGCAA
  42_8    AAAACCACCA ACCCAGTGGC CACAGAACAG TACGGCGTGG TGGCCGATAA
 42_15    AAAACCACCA ACCCAGTGGC CACAGAACAG TACGGCGTGG TGGCCGATAA
 42_5b    AAAACCACCA ACCCAGTGGC CACAGAACAG TACGGCGTGG TGGCCGATAA
 42_1b    AAAACCACCA ACCCAGTGGC CACAGAACAG TACGGCGTGG TGGCCGATAA
 42_13    AAAACCACCA ACCCAGTGGC CACAGAACAG TACGGCGTGG TGGCCGATAA
 42_3a    AAAACCACCA ACCCAGTGGC CACAGAACAG TACGGCGTGG TGGCCGATAA
  42_4    AAAACCACCA ACCCAGTGGC CACAGAACAG TACGGCGTGG TGGCCGATAA
 42_5a    AAAACCACCA ACCCAGTGGC CACAGAACAG TACGGCGTGG TGGCCGATAA
 42_10    AAAACCACCA ATCCCGTGGC TACAGAAGAA TACGGTGTGG TCTCCAGCAA
 42_3b    AAAACCACCA ATCCCGTGGC TACAGAACAG TACGGTGTGG TCTCCAGCAA
 42_11    AAAACCACCA ATCCCGTGGC TACAGAAGAA TACGGTGTGG TCTCCAGCAA
 42_6b    AAAACCACCA ATCCCGTGGC TACAGAAGAA TACGGTGTGG TCTCCAGCAA
  43_1    AAAACTACTA ACCCAGTGGC TACAGAGCAG TATGGTGTGG TGGCAGACAA
  43_5    AAAACTACTA ACCCAGTGGC TACAGAGCAG TATGGTGTGG TGGCAGACAA
 43_12    AAAACTACTA ACCCAGTGGC TACAGAGCAG TATGGTGTGG TGGCAGACAA
 43_20    AAGGCTACCA ACCCCGTGGC CACAGAAGAA TATGGAGCAG TGGCCATCAA
 43_21    AAGGCTACCA ACCCCGTGGC CACAGAAGAA TATGGAGCAG TGGCCATCAA
 43_23    AAGGCTACCA ACCCCGTGGC CACAGAAGAA TATGGAGCAG TGGCCATCAA
 43_25    AAGGCTACCA ACCCCGTGGC CACAGAAGAA TATGGAGCAG TGGCCATCAA
  44_1    AAAACCACCA ACCCAGTGGC CACGGAACAG TACGGCGTGG TGGCCGATAA
  44_5    AAAACCACCA ACCCAGTGGC CACAGAACAG TACGGCGTGG TGGCCGATAA
223_10    CGTCCTACCA ACCCGGTAGC TACCGAGGAA TACGGGATTG TAAGCAGCAA
 223_2    CGTCCTACCA ACCCGGTAGC TACCGAGGAA TACGGGATTG TAAGCAGCAA
 223_4    CGTCCTACCA ACCCGGTAGC TACCGAGGAA TACGGGATTG TAAGCAGCAA
 223_5    CGTCCTACCA ACCCGGTAGC TACCGAGGAA TACGGGATTG TAAGCAGCAA
 223_6    CGTCCTACCA ACCCGGTAGC TACCGAGGAA TACGGGATTG TAAGCAGCAA
 223_7    CGTCCTACCA ACCCGGTAGC TACCGAGGAA TACGGGATTG TAAGCAGCAA
  A3_4    AGAACAACTA ATCCTGTGGC TACAGAACAA TACGGACAGG TTGCCACCAA
  A3_5    AGAACGACTA ATCCTGTGGC TACAGAACAA TACGGACAGG TTGCCACCAA
  A3_7    AGAACAACTA ATCCTGTGGC TACAGAACAA TACGGACAGG TTGCCACCAA
  A3_3    AGAACAACTA ATCCTGTGGC TACAGAACAA TACGGACAGG TTGCCACCAA
 42_12    AAAACCACCA ATCCCGTGGC TACAGAAGAA TACGGTGTGG TCTCCAGCAA
   AAV1    AAAGCCACTA ACCCTGTGGC CACCGAAAGA TTTGGGACCG TGGCAGTCAA
   AAV2    GGAACAACCA ATCCCGTGGC TACGGAGCAG TATGGTTCTG TATCTACCAA
   AAV3    CGTACCACCA ATCCTGTGGC AACAGAGCAG TATGGAACTG TGGCAAATAA
   AAV8    AAAACCACTA ACCCTGTGGC TACAGAGGAA TACGGTATCG TGGCAGATAA
   AAV9    AAAGCCACCA ACCCTGTAGC CACAGAGGAA TACGGAGCAG TGGCCATCAA
   AAV7    CGTCCTACTA ATCCTGTAGC CACGGAAGAA TACGGGATAG TCAGCAGCAA
  44_2     AAAACCACCA ACCCAGTGGC CACAGAACAG TACGGCGTGG TGGCCGATAA
```

图1AAAB

```
          4001                                                    4050
  42_2    CCTGCAATCG TCTACGGCCG GACCCCAGAC ACAGACTGTC AACAGCCAGG
  42_8    CCTGCAACAG CAAAACGCCG CTCCTATTGT AGGGGCCGTC AACAGTCAAG
 42_15    CCTGCAACAG CAAAACGCCG CTCCTATTGT AGGGGCCGTC AACAGTCAAG
 42_5b    CCTGCAACAG CAAAACGCCG CTCCTATTGT AGGGGCCGTC AACAGTCAAG
 42_1b    CCTGCAACAG CAAAACGCCG CTCCTATTGT AGGGGCCGTC AACAGTCAAG
 42_13    CCTGCAACAG CAAAACGCCG CTCCTATTGT AGGGGCCGTC AACAGTCAAG
 42_3a    CCTGCAACAG CAAAACGCCG CTCCTATTGT AGGGGCCGTC AACAGTCAAG
  42_4    CCTGCAACAG CAAAACGCCG CTCCTATTGT AGGGGCCGTC AACAGTCAAG
 42_5a    CCTGCAACAG CAAAACGCCG CTCCTATTGT AGGGGCCGTC AACAGTCAAG
 42_10    CCTGCAATCG TCTACGGCCG GACCCCAGAC ACAGACTGTC AACAGCCAGG
 42_3b    CCTGCAATCG TCTACGGCCG GACCCCAGAC ACAGACTGTC AACAGCCAGG
 42_11    CCTGCAATCG TCTACGGCCG GACCCCAGAC ACAGACTGTC AACAGCCAGG
 42_6b    CCTGCAATCG TCTACGGCCG GACCCCAGAC ACAGACTGTC AACAGCCAGG
  43_1    CCTGCAGCAG ACCAACGGAG CTCCCATTGT GGGAACTGTC AACAGCCAGG
  43_5    CCTGCAGCAG ACCAACGGAG CTCCCATTGT GGGAACTGTC AACAGCCAGG
 43_12    CCTGCAGCAG ACCAACGGAG CTCCCATTGT GGGAACTGTC AACAGCCAGG
 43_20    CAACCAGGCC GCCAATACGC AGGCGCAGAC CGGACTCGTG CACAACCAGG
 43_21    CAACCAGGCC GCCAATACGC AGGCGCAGAC CGGACTCGTG CACAACCAGG
 43_23    CAACCAGGCC GCCAATACGC AGGCGCAGAC CGGACTCGTG CACAACCAGG
 43_25    CAACCAGGCC GCCAATACGC AGGCGCAGAC CGGACTCGTG CACAACCAGG
  44_1    CCTGCAACAG CAAAACGCCG CTCCTATTGT AGGGGCCGTC AACAGTCAAG
  44_5    CCTGCAACAG CAAAACGCCG CTCCTATTGT AGGGGCCGTC AACAGTCAAG
223_10    CTTGCAGGCG GCTAGCACCG CAGCCCAGAC ACAAGTTGTT AACAACCAGG
 223_2    CTTGCAGGCG GCTAGCACCG CAGCCCAGAC ACAAGTTGTT AACAACCAGG
 223_4    CTTGCAGGCG GCTAGCACCG CAGCCCAGAC ACAAGTTGTT AACAACCAGG
 223_5    CTTGCAGGCG GCTAGCACCG CAGCCCAGAC ACAAGTTGTT AACAACCAGG
 223_6    CTTGCAGGCG GCTAGCACCG CAGCCCAGAC ACAAGTTGTT AACAACCAGG
 223_7    CTTGCAGGCG GCTAGCACCG CAGCCCAGAC ACAAGTTGTT AACAACCAGG
  A3_4    CCATCAGAGT CAGGACACCA CAGCTTCCTA TGGAAGTGTG GACAGCCAGG
  A3_5    CCGTCAGAGT CAGAACACCA CAGCTTCCTA TGGAAGTGTG GACAGCCAGG
  A3_7    CCATCAGAGT CAGAACACCA CAGCTTCCTA TGGAAGTGTG GACAGCCAGG
  A3_3    CCATCAGAGT CAGAACACCA CAGCTTCCTA TGGAAGTGTG GACAGCCAGG
 42_12    CCTGCAATCG TCTACGGCCG GACCCCAGAC ACAGACTGTC AACAGCCAGG
  AAV1    TTTCCAGAGC AGCAGCACAG ACCCTGCGAC CGGAGATGTG CATGCTATGG
  AAV2    CCTCCAGAGA GGCAACAGAC AAGCAGCTAC CGCAGATGTC AACACACAAG
  AAV3    CTTGCAGAGC TCAAATACAG CTCCCACGAC TGGAACTGTC AATCATCAGG
  AAV8    CTTGCAGCAG CAAAACACGG CTCCTCAAAT TGGAACTGTC AACAGCCAGG
  AAV9    CAACCAGGCC GCTAACACGC AGGCGCAAAC TGGACTTGTG CATAACCAGG
  AAV7    CTTACAAGCG GCTAATACTG CAGCCCAGAC ACAAGTTGTC AACAACCAGG
  44_2    CCTGCAACAG CAAAACGCCG CTCCTATTGT AGGGGCCGTC AACAGTCAAG
```

图1AAAC

```
          4051                                                        4100
  42_2    GGGCTCTGCC CGGCATGGTC TGGCAGAACC GGGACGTGTA CCTGCAGGGT
  42_8    GAGCCTTACC TGGCATGGTC TGGCAGAACC GGGACGTGTA CCTGCAGGGT
 42_15    GAGCCTTACC TGGCATGGTC TGGCAGAACC GGGACGTGTA CCTGCAGGGT
 42_5b    GAGCCTTACC TGGCATGGTC TGGCAGAACC GGGACGTGTA CCTGCAGGGT
 42_1b    GAGCCTTACC TGGCATGGTC TGGCAGAACC GGGACGTGTA CCTGCAGGGT
 42_13    GAGCCTTACC TGGCATGGTC TGGCAGAACC GGGACGTGTA CCTGCAGGGT
 42_3a    GAGCCTTACC TGGCATGGTC TGGCAGAACC GGGACGTGTA CCTGCAGGGT
  42_4    GAGCCTTACC TGGCATGGTC TGGCAGAACC GGGACGTGTA CCTGCAGGGT
 42_5a    GAGCCTTACC TGGCATGGCC TGGCAGAACC GGGACGTGTA CCTGCAGGGT
 42_10    GGGCTCTGCC CGGCATGGTC TGGCAGAACC GGGACGTGTA CCTGCAGGGT
 42_3b    GGGCTCTGCC CGGCATGGTC TGGCAGAACC GGGACGTGTA CCTGCAGGGT
 42_11    GGGCTCTGCC CGGCATGGTC TGGCAGAACC GGGACGTGTA CCTGCAGGGT
 42_6b    GGGCTCTGCC CGGCATGGTC TGGCAGAACC GGGACGTGTA CCTGCAGGGT
  43_1    GGGCCTTACC TGGTATGGTC TGGCAAAACC GGGACGTGTA CCTGCAGGGC
  43_5    GGGCCTTACC TGGTATGGTC TGGCAAAACC GGGACGTGTA CCTGCAGGGC
 43_12    GGGCCTTACC TGGTATGGTC TGGCAAAACC GGGACGTGTA CCTGCAGGGC
 43_20    GGGTGATTCC CGGCATGGTG TGGCAGAATA GAGACGTGTA CCTGCAGGGT
 43_21    GGGTGATTCC CGGCATGGTG TGGCAGAATA GAGACGTGTA CCTGCAGGGT
 43_23    GGGTGATTCC CGGCATGGTG TGGCAGAATA GAGACGTGTA CCTGCAGGGT
 43_25    GGGTGATTCC CGGCATGGTG TGGCAGAATA GAGACGTGTA CCTGCAGGGT
  44_1    GAGCCTTACC TGGCATGGTC TGGCAGAACC GGGACGTGTA CCTGCAGGGT
  44_5    GAGCCTTACC TGGCATGGTC TGGCAGAACC GGGACGTGTA CCTGCAGGGT
223_10    GAGCCTTACC TGGCATGGTC TGGCAGAACC GGGACGTGTA CCTGCAAGGT
 223_2    GAGCCTTACC TGGCATGGTC TGGCAGAACC GGGACGTGTA CCTGCAAGGT
 223_4    GAGCCTTACC TGGCATGGTC TGGCAGAACC GGGACGTGTA CCTGCAAGGT
 223_5    GAGCCTTACC TGGCATGGTC TGGCAGAACC GGGACGTGTA CCTGCAAGGT
 223_6    GAGCCTTACC TGGCATGGTC TGGCAGAACC GGGACGTGTA CCTGCAAGGT
 223_7    GAGCCTTACC TGGCATGGTC TGGCAGAACC GGGACGTGTA CCTGCAAGGT
  A3_4    GAATCTTACC TGGAATGGTG TGGCAGGACC GCGATGTCTA TCTTCAAGGT
  A3_5    GAATCTTACC TGGAATGGTG TGGCAGGACC GCGATGTCTA TCTTCAAGGT
  A3_7    GAATCTTACC TGGAATGGTG TGGCAGGACC GCGATGTCTA TCTTCAAGGT
  A3_3    GAATCTTACC TGGAATGGTG TGGCAGGACC GCGATGTCTA TCTTCAAGGT
 42_12    GGGCTCTGCC CGGCATGGTC TGGCAGAACC GGGACGTGTA CCTGCAGGGT
   AAV1   GAGCATTACC TGGCATGGTG TGGCAAGATA GAGACGTGTA CCTGCAGGGT
   AAV2   GCGTTCTTCC AGGCATGGTC TGGCAGGACA GAGATGTGTA CCTTCAGGGG
   AAV3   GGGCCTTACC TGGCATGGTG TGGCAAGATC GTGACGTGTA CCTTCAAGGA
   AAV8   GGGCCTTACC CGGTATGGTC TGGCAGAACC GGGACGTGTA CCTGCAGGGT
   AAV9   GAGTTATTCC TGGTATGGTC TGGCAGAACC GGGACGTGTA CCTGCAGGGC
   AAV7   GAGCCTTACC TGGCATGGTC TGGCAGAACC GGGACGTGTA CCTGCAGGGT
  44_2    GAGCCTTACC TGGCATGGTC TGGCAGAACC GGGACGTGTA CCTGCAGGGT
```

图1AAAD

```
          4101                                                    4150
  42_2    CCC.ATCTGG GCCAAAATTC CTCACACGGA CGGCAACTTT CACCCGTCTC
  42_8    CCT.ATCTGG GCCAAGATTC CTCACACGGA CGGCAACTTT CATCCTTCGC
 42_15    CCT.ATCTGG GCCAAGATTC CTCACACGGA CGGCAACTTT CATCCTTCGC
 42_5b    CCT.ATCTGG GCCAAGATTC CTCACACGGA CGGCAACTTT CATCCTTCGC
 42_1b    CCT.ATCTGG GCCAAGATTC CTCACACGGA CGGCAACTTT CATCCTTCGC
 42_13    CCT.ATCTGG GCCAAGATTC CTCACACGGA CGGCAACTTT CATCCTTCGC
 42_3a    CCT.ATCTGG GCCAAGATTC CTCACACGGA CGGCAACTTT CATCCTTCGC
  42_4    CCT.ATCTGG GCCAAGATTC CTCACACGGA CGGCAACTTT CATCCTTCGC
 42_5a    CCT.ATCTGG GCCAAGATTC CTCACACGGA CGGCAACTTT CATCCTTCGC
 42_10    CCC.ATCTGG GCCAAAATTC CTCACACGGA CGGCAACTTT CACCCGTCTC
 42_3b    CCC.ATCTGG GCCAAAATTC CTCACACGGA CGGCAACTTT CACCCGTCTC
 42_11    CCC.ATCTGG GCCAAAATTC CTCACACGGA CGGCAACTTT CACCCGTCTC
 42_6b    CCC.ATCTGG GCCAAAATTC CTCACACGGA CGGCAACTTT CACCCGTCTC
  43_1    CCC.ATCTGG GCCAAAATTC CTCACACGGA CGGCAACTTT CATCCTTCGC
  43_5    CCC.ATCTGG GCCAAAATTC CTCACACGGA CGGCAACTTT CATCCTTCGC
 43_12    CCC.ATCTGG GCCAAAATTC CTCACACGGA CGGCAACTTT CATCCTTCGC
 43_20    CCC.ATCTGG GCCAAAATTC CTCACACGGA CGGCAACTTT CACCCGTCTC
 43_21    CCC.ATCTGG GCCAAAATTC CTCACACGGA CGGCAACTTT CACCCGTCTC
 43_23    CCC.ATCTGG GCCAAAATTC CTCACACGGA CGGCAACTTT CACCCGTCTC
 43_25    CCC.ATCTGG GCCAAAATTC CTCACACGGA CGGCAACTTT CACCCGTCTC
  44_1    CCT.ATCTGG GCCAAGATTC CTCACACGGA CGGAAACTTT CATCCCTCGC
  44_5    CCT.ATCTGG GCCAAGATTC CTCACACGGA CGGAAACTTT CATCCCTCGC
223_10    CCC.ATTTGG GCCAAGATTC CTCACACGGA CGGCAACTTT CACCCGTCTC
 223_2    CCC.ATTTGG GCCAAGATTC CTCACACGGA CGGCAACTTT CACCCGTCTC
 223_4    CCC.ATTTGG GCCAAGATTC CTCACACGGA CGGCAACTTT CACCCGTCTC
 223_5    CCC.ATTTGG GCCAAGATTC CTCACACGGA CGGCAACTTT CACCCGTCTC
 223_6    CCC.ATTTGG GCCAAGATTC CTCACACGGA CGGCAACTTT CACCCGTCTC
 223_7    CCC.ATTTGG GCCAAGATTC CTCACACGGA CGGCAACTTT CACCCGTCTC
  A3_4    CCC.ATTTGG GCCAAAACTC CTCACACGGA CGGACACTTT CATCCTTCTC
  A3_5    CCC.ATTTGG GCCAAAACTC CTCACACGGA CGGACACTTT CATCCTTCTC
  A3_7    CCC.ATTTGG GCCAAAACTC CTCACACGGA CGGACACTTT CATCCTTCTC
  A3_3    CCC.ATTTGG GCCAAAACTC CTCACACGGA CGGACACTTT CATCCTTCTC
 42_12    CCC.ATCTGG GCCAAAATTC CTCACACGGA CGGCAACTTT CACCCGTCTC
   AAV1   CCC.ATTTGG GCCAAAATTC CTCACACAGA TGGACACTTT CACCCGTCTC
   AAV2   CCC.ATCTGG GCAAAGATTC CACACACGGA CGGACATTTT CACCCCTCTC
   AAV3   CCT.ATCTGG GCAAAGATTC CTCACACGGA TGGACACTTT CATCCTTCTC
   AAV8   CCC.ATCTGG GCCAAGATTC CTCACACGGA CGGCAACTTC CACCCGTCTC
   AAV9   CCCTATTTGG GCTAAAATAC CTCACACAGA TGGCAACTTT CACCCGTCTC
   AAV7   CCC.ATCTGG GCCAAGATTC CTCACACGGA TGGCAACTTT CACCCGTCTC
  44_2    CCT.ATCTGG GCCAAGATTC CTCACACGGA CGGAAACTTT CATCCCTCGC
```

图1AAAE

```
          4151                                                    4200
  42_2    CCCTGATGGG CGGATTTGGA CTCAAACACC CGCCTCCTCA AATTCTCATC
  42_8    CGCTGATGGG AGGCTTTGGA CTGAAACACC CGCCTCCTCA GATCCTGATT
 42_15    CGCTGATGGG AGGCTTTGGA CTGAAACACC CGCCTCCTCA GATCCTGATT
 42_5b    CGCTGATGGG AGGCTTTGGA CTGAAACACC CGCCTCCTCA GATCCTGATT
 42_1b    CGCTGATGGG AGGCTTTGGA CTGAAACACC CGCCTCCTCA GATCCTGATT
 42_13    CGCTGATGGG AGGCTTTGGA CTGAAACACC CGCCTCCTCA GATCCTGATT
 42_3a    CGCTGATGGG AGGCTTTGGA CTGAAACACC CGCCTCCTCA GATCCTGATT
  42_4    CGCTGATGGG AGGCTTTGGA CTGAAACACC CGCCTCCTCA GATCCTGATT
 42_5a    CGCTGATGGG AGGCTTTGGA CTGAAACACC CGCCTCCTCA GATCCTGATT
 42_10    CCCTGATGGG CGGATTTGGA CTCAAACACC CGCCTCCTCA AATTCTCATC
 42_3b    CCCTGATGGG CGGATTTGGA CTCAAACACC CGCCTCCTCA AATTCTCATC
 42_11    CCCTGATGGG CGGATTTGGA CTCAAACACC CGCCTCCTCA AATTCTCATC
 42_6b    CCCTGATGGA CGGATTTGGA CTCAAACACC CGCCTCCTCA AATTCTCATC
  43_1    CGCTGATGGG AGGCTTTGGA CTGAAACACC CGCCTCCTCA GATCCTGGTG
  43_5    CGCTGATGGG AGGCTTTGGA CTGAAACACC CGCCTCCTCA GATCCTGGTG
 43_12    CGCTGATGGG AGGCTTTGGA CTGAAACACC CGCCTCCTCA GATCCTGGTG
 43_20    CCCTGATGGG CGGCTTTGGA CTGAAGCACC CGCCTCCTCA AATTCTCATC
 43_21    CCCTGATGGG CGGCTTTGGA CTGAAGCACC CGCCTCCTCA AATTCTCATC
 43_23    CCCTGATGGG CGGCTTTGGA CTGAAGCACC CGCCTCCTCA AATTCTCATC
 43_25    CCCTGATGGG CGGCTTTGGA CTGAAGCACC CGCCTCCTCA AATTCTCATC
  44_1    CGCTGATGGG AGGCTTTGGA CTGAAACACC CGCCTCCTCA GATCCTGATT
  44_5    CGCTGATGGG AGGCTTTGGA CTGAAACACC CGCCTCCTCA GATCCTGATT
223_10    CTCTAATGGG TGGCTTTGGA CTGAAACACC CGCCTCCCCA GATCCTGATC
 223_2    CTCTAATGGG TGGCTTTGGA CTGAAACACC CGCCTCCCCA GATCCTGATC
 223_4    CTCTAATGGG TGGCTTTGGA CTGAAACACC CGCCTCCCCA GATCCTGATC
 223_5    CTCTAATGGG TGGCTTTGGA CTGAAACACC CGCCTCCCCA GATCCTGATC
 223_6    CTCTAATGGG TGGCTTTGGA CTGAAACACC CGCCTCCCCA GATCCTGATC
 223_7    CTCTAATGGG TGGCTTTGGA CTGAAACACC CGCCTCCCCA GATCCTGATC
  A3_4    CGCTCATGGG AGGCTTTGGA CTGAAACACC CTCCTCCCCA GATCCTGATC
  A3_5    CGCTCATGGG AGGCTTTGGA CTGAAACACC CTCCTCCCCA GATCCTGATC
  A3_7    CGCTCATGGG AGGCTTTGGA CTGAAACACC CTCCTCCCCA GATCCTGATC
  A3_3    CGCTCATGGG AGGCTTTGGA CTGAAACACC CTCCTCCCCA GATCCTGATC
 42_12    CCCTGATGGG CGGATTTGGA CTCAAACACC CGCCTCCTCA AATTCTCATC
  AAV1    CTCTTATGGG CGGCTTTGGA CTCAAGAACC CGCCTCCTCA GATCCTCATC
  AAV2    CCCTCATGGG TGGATTCGGA CTTAAACACC CTCCTCCACA GATTCTCATC
  AAV3    CTCTGATGGG AGGCTTTGGA CTGAAACATC CGCCTCCTCA AATCATGATC
  AAV8    CGCTGATGGG CGGCTTTGGC CTGAAACATC CTCCGCCTCA GATCCTGATC
  AAV9    CTCTGATGGG TGGATTTGGA CTGAAACACC CACCTCCACA GATTCTAATT
  AAV7    CTTTGATGGG CGGCTTTGGA CTTAAACATC CGCCTCCTCA GATCCTGATC
  44_2    CGCTGATGGG AGGCTTTGGA CTGAAACACC CGCCTCCTCA GATCCTGATT
```

图1AAAF

```
          4201                                                      4250
  42_2    AAAAACACCC CGGTACCTGC TAATCCTCCA GAGGTGTTTA CTCCTGCCAA
  42_8    AAGAATACAC CTGTTCCCGC GGATCCTCCA ACTACCTTCA GTCAAGCCAA
 42_15    AAGAATACAC CTGTTCCCGC GGATCCTCCA ACTACCTTCA GTCAAGCCAA
 42_5b    AAGAATACAC CTGTTCCCGC GGATCCTCCA ACTACCTTCA GTCAAGCCAA
 42_1b    AAGAATACAC CTGTTCCCGC GGATCCTCCA ACTACCTTCA GTCAAGCCAA
 42_13    AAGAATACAC CTGTTCCCGC GGATCCTCCA ACTACCTTCA GTCAAGCCAA
 42_3a    AAGAATACAC CTGTTCCCGC GGATCCTCCA ACTACCTTCA GTCAAGCCAA
  42_4    AAGAATACAC CTGTTCCCGC GGATCCTCCA ACTACCTTCA GTCAAGCCAA
 42_5a    AAGAATACAC CTGTTCCCGC GGATCCTCCA ACTACCTTCA GTCAAGCCAA
 42_10    AAAAACACCC CGGTACCTGC TAATCCTCCA GAGGTGTTTA CTCCTGCCAA
 42_3b    AAAAACACCC CGGTACCTGC TAATCCTCCA GAGGTGTTTA CTCCTGCCAA
 42_11    AAAAACACCC CGGTACCTGC TAATCCTCCA GAGGTGTTTA CTCCTGCCAA
 42_6b    AAAAACACCC CGGTACCTGC TAATCCTCCA GAGGTGTTTA CTCCTGCCAA
  43_1    AAAAACACTC CTGTTCCTGC GGATCCTCCG ACCACCTTCA GCCAGGCCAA
  43_5    AAAAACACTC CTGTTCCTGC GGATCCTCCG ACCACCTTCA GCCAGGCCAA
 43_12    AAAAACACTC CTGTTCCTGC GGATCCTCCG ACCACCTTCA GCCAGGCCAA
 43_20    AAGAACACAC CGGTTCCAGC GGACCCGCCG CTTACCTTCA ACCAGGCCAA
 43_21    AAGAACACAC CGGTTCCAGC GGACCCGCCG CTTACCTTCA ACCAGGCCAA
 43_23    AAGAACACAC CGGTTCCAGC GGACCCGCCG CTTACCTTCA ACCAGGCCAA
 43_25    AAGAACACAC CGGTTCCAGC GGACCCGCCG CTTACCTTCA ACCAGGCCAA
  44_1    AAGAATACAC CTGTTCCCGC GGATCCTCCA ACTACCTTCA GTCAAGCTAA
  44_5    AAGAATACAC CTGTTCCCGC GGATCCTCCA ACTACCTTCA GTCAAGCTAA
223_10    AAAAACACAC CGGTACCTGC TAATCCTCCA GAAGTGTTTA CTCCTGCCAA
 223_2    AAAAACACGC CGGTACCTGC TAATCCTCCA GAAGTGTTTA CTCCTGCCAA
 223_4    AAAAACACAC CGGTACCTGC TAATCCTCCA GAAGTGTTTA CTCCTGCCAA
 223_5    AAAAACACAC CGGTACCTGC TAATCCTCCA GAAGTGTTTA CTCCTGCCAA
 223_6    AAAAACACAC CGGTACCTGC TAATCCTCCA GAAGTGTTTA CTCCTGCCAA
 223_7    AAAAACACAC CGGTACCTGC TAATCCTCCA GAAGTGTTTA CTCCTGCCAA
  A3_4    AAAAACACAC CTGTGCCAGC GAATCCCGCG ACCACTTTCA CTCCTGGAAA
  A3_5    AAAAACACAC CTGTGCCAGC GAATCCCGCG ACCACTTTCA CTCCTGGAAA
  A3_7    AAAAACACAC CTGTGCCAGC GAATCCCGCG ACCACTTTCA CTCCTGGAAA
  A3_3    AAAAACACAC CTGTGCCAGC GAATCCCGCG ACCACTTTCA CTCCTGGAAA
 42_12    A...A..... .......... .......... .......... ..........
  AAV1    AAAAACACGC CTGTTCCTGC GAATCCTCCG GCGGAGTTTT CAGCTACAAA
  AAV2    AAGAACACCC CGGTACCTGC GAATCCTTCG ACCACCTTCA GTGCGGCAAA
  AAV3    AAAAATACTC CGGTACCGGC AAATCCTCCG ACGACTTTCA GCCCGGCCAA
  AAV8    AAGAACACGC CTGTACCTGC GGATCCTCCG ACCACCTTCA ACCAGTCAAA
  AAV9    AAAAATACAC CAGTGCCGGC AGATCCTCCT CTTACCTTCA ATCAAGCCAA
  AAV7    AAGAACACTC CCGTTCCCGC TAATCCTCCG GAGGTGTTTA CTCCTGCCAA
  44_2    AAGAATACAC CTGTTCCCGC GGATCCTCCA ACTACCTTCA GTCAAGCTAA
```

图1AAAG

```
          4251                                                      4300
  42_2    GTTTGCCTCA TTTATCACGC AGTACAGCAC CGGCCA.GGT CAGCGTGGAG
  42_8    GCTGGCGTCG TTCATCACGC AGTACAGCAC CGGACA.GGT CAGCGTGGAA
 42_15    GCTGGCGTCG TTCATCACGC AGTACAGCAC CGGACA.GGT CAGCGTGGAA
 42_5b    GCTGGCGTCG TTCATCACGC AGTACAGCAC CGGACA.GGT CAGCGTGGAA
 42_1b    GCTGGCGTCG TTCATCACGC AGTACAGCAC CGGACA.GGT CAGCGTGGAA
 42_13    GCTGGCGTCG TTCATCACGC AGTACAGCAC CGGACA.GGT CAGCGTGGAA
 42_3a    GCTGGCGTCG TTCATCACGC AGTACAGCAC CGGACA.GGT CAGCGTGGAA
  42_4    GCCGGCGTCG TTCATCACGC AGTACAGCAC CGGACA.GGT CAGCGTGGAA
 42_5a    GCTGGCGTCG TTCATCACGC AGTACAGCAC CGGACA.GGT CAGCGTGGAA
 42_10    GTTTGCCTCA TTTATCACGC AGTACAGCAC CGGCCA.GGT CAGCGTGGAG
 42_3b    GTTTGCCTCA TTTATCACGC AGTACAGCAC CGGCCA.GGT CAGCGTGGAG
 42_11    GTTTGCCTCA TTTATCACGC AGTACAGCAC CGGCCA.GGT CAGCGTGGAG
 42_6b    GTTTGCCTCA TTTATCACGC AGTACAGCAC CGGCCA.GGT CAGCGTGGAG
  43_1    GCTGGCTTCT TTTATCACGC AGTACAGCAC CGGACA.GGT CAGCGTGGAA
  43_5    GCTGGCTTCT TTTATCACGC AGTACAGCAC CGGACA.GGT CAGCGTGGAA
 43_12    GCTGGCTTCT TTTATCACGC AGTACAGCAC CGGACA.GGT CAGCGTGGAA
 43_20    GCTGAACTCT TTCATCACGC AGTACAGCAC CGGACA.GGT CAGCGTGGAA
 43_21    GCTGAACTCT TTCATCACGC AGTACAGCAC CGGACA.GGT CAGCGTGGAA
 43_23    GCTGAACTCT TTCATCACGC AGTACAGCAC CGGACA.GGT CAGCGTGGAA
 43_25    GCTGAACTCT TTCATCACGC AGTACAGCAC CGGACA.GGT CAGCGTGGAA
  44_1    GCTGGCGTCG TTCATCACGC AGTACAGCAC CGGACA.GGT CAGCGTGGAA
  44_5    GCTGGCGTCG TTCATCACGC AGTACAGCAC CGGACA.GGT CAGCGTGGAA
223_10    GTTTGCTTCC TTCATCACGC AGTACAGCAC CGGGCA.AGT CAGCGTTGAG
 223_2    GTTTGCTTCC TTCATCACGC AGTACAGCAC CGGGCA.AGT CAGCGTTGAG
 223_4    GTTTGCTTCC TTCATCACGC AGTACAGCAC CGGGCA.AGT CAGCGTTGAG
 223_5    GTTTGCTTCC TTCATCACGC AGTACAGCAC CGGGCA.AGT CAGCGTTGAG
 223_6    GCTTGCTTCC TTCATCACGC AGTACAGCAC CGGGCA.AGT CAGCGTTGAG
 223_7    GATTGCTTCC TTCATCACGC AGTACAGCAC CGGGCA.AGT CAGCGTTGAG
  A3_4    GTTTGCTTCG TTCATTACCC AGTATTCCAC CGGACA.GGT CAGCGTGGAA
  A3_5    GTTTGCTTCG TTCATTACCC AGTATTCCAC CGGACA.GGT CAGCGTGGAA
  A3_7    GTTTGCTTCG TTCATTACCC AGTATTCCAC CGGACA.GGT CAGCGTGGAA
  A3_3    GTTTGCTTCG TTCATTACCC AGTATTCCAC CGGACA.GGT CAGCGTGGAA
 42_12    .......... .......... .......... .......... ..........
  AAV1    GTTTGCTTCA TTCATCACCC AATACTCCAC AGGACA.AGT GAGTGTGGAA
  AAV2    GTTTGCTTCC TTCATCACAC AGTACTCCAC GGGACACGGT CAGCGTGGAG
  AAV3    GTTTGCTTCA TTTATCACTC AGTACTCCAC TGGACA.GGT CAGCGTGGAA
  AAV8    GCTGAACTCT TTCATCACGC AATACAGCAC CGGACA.GGT CAGCGTGGAA
  AAV9    GCTGAACTCT TTCATCACGC AGTACAGCAC GGGACA.AGT CAGCGTGGAA
  AAV7    GTTTGCTTCG TTCATCACAC AGTACAGCAC CGGACA.AGT CAGCGTGGAA
  44_2    GCTGGCGTCG TTCATCACGC AGTACAGCAC CGGACA.GGT CAGCGTGGAA
```

图1AAAH

```
          4301                                                      4350
  42_2    ATCGAGTGGG AACTGCAGAA AGAAAACAGC AAACGCTGGA ATCCAGAGAT
  42_8    ATTGAATGGG AGCTGCAGAA AGAGAACAGC AAGCGCTGGA ACCCAGAGAT
 42_15    ATTGAATGGG AGCTGCAGAA AGAGAACAGC AAGCGCTGGA ACCCAGAGAT
 42_5b    ATTGAATGGG AGCTGCAGAA AGAGAACAGC AAGCGCTGGA ACCCAGAGAT
 42_1b    ATTGAATGGG AGCTGCAGAA AGAGAACAGC AAGCGCTGGA ACCCAGAGAT
 42_13    ATTGAATGGG AGCTGCAGAA AGAGAACAGC AAGCGCTGGA ACCCAGAGAT
 42_3a    ATTGAATGGG AGCTGCAGAA AGAGAACAGC AAGCGCTGGA ACCCAGAGAT
  42_4    ATTGAATGGG AGCTGCAGAA AGAGAACAGC AAGCGCTGGA ACCCAGAGAT
 42_5a    ATTGAATGGG AGCTGCAGAA AGAGAACAGC AAGCGCTGGA ACCCAGAGAT
 42_10    ATCGAGTGGG AACTGCAGAA AGAAAACAGC AAACGCTGGA ATCCAGAGAT
 42_3b    ATCGAGTGGG AACTGCAGAA AGAAAACAGC AAACGCTGGA ATCCAGAGAT
 42_11    ATCGAGTGGG AACTGCAGAA AGAGAACAGC AAACGCTGGA ATCCAGAGAT
 42_6b    ATCGAGTGGG AACTGCAGAA AGAAAACAGC AAACGCTGGA ATCCAGAGAT
  43_1    ATCGAATGGG AGCTGCAGAA AGAAAACAGC AAGCGCTGGA ACCCAGAGAT
  43_5    ATCGAATGGG AGCTGCAGAA AGAAAACAGC AAGCGCTGGA ACCCAGAGAT
 43_12    ATCGAATGGG AGCTGCAGAA AGAAAACAGC AAGCGCTGGA ACCCAGAGAT
 43_20    ATCGAGTGGG AGCTGCAGAA AGAAAACAGC AAACGCTGGA ATCCAGAGAT
 43_21    ATCGAGTGGG AGCTGCAGAA AGAAAACAGC AAACGCTGGA ATCCAGAGAT
 43_23    ATCGAGTGGG AGCTGCAGAA AGAAAACAGC AAACGCTGGA ATCCAGAGAT
 43_25    ATCGAGTGGG AGCTGCAGAA AGAAAACAGC AAACGCTGGA ATCCAGAGAT
  44_1    ATTGAATGGG AGCTGCAGAA AGAAAACAGC AAACGCTGGA ACCCAGAGAT
  44_5    ATTGAATGGG AGCTGCAGAA AGAAAACAGC AAACGCTGGA ACCCAGAGAT
223_10    ATCGAGTGGG AGCTGCAGAA AGAGAACAGC AAGCGCTGGA ACCCAGAGAT
 223_2    ATCGAGTGGG AGCTGCAGAA AGAGAACAGC AAGCGCTGGA ACCCAGAGAT
 223_4    ATCGAATGGG AGCTGCAGAA AGAGAACAGC AAGCGCTGGA ACCCAGAGAT
 223_5    ATCGAATGGG AGCTGCAGAA AGAGAACAGC AAGCGCTGGA ACCCAGAGAT
 223_6    ATCGAGTGGG AGCTGCAGAA AGAGAACAGC AAGCGCTGGA ACCCAGAGAT
 223_7    ATCGAGTGGG AGCTGCAGAA AGAGAACAGC AAGCGCTGGA ACCCAGAGAT
  A3_4    ATAGAGTGGG AGCTGCAGAA AGAAAACAGC AAACGCTGGA ACCCAGAAAT
  A3_5    ATAGAGTGGG AGCTGCAGAA AGAAAACAGC AAACGCTGGA ACCCGGAAAT
  A3_7    ATAGAGTGGG AGCTGCAGAA AGAAAACAGC AAACGCTGGA ACCCAGAAAT
  A3_3    ATAGAGTGGG AGCTGCAGAA AGAAAACAGC AAACGCTGGA ACCCAGAAAT
 42_12    .......... .......... .......... .......... ..........
  AAV1    ATTGAATGGG AGCTGCAGAA AGAAAACAGC AAGCGCTGGA ATCCCGAAGT
  AAV2    ATCGAGTGGG AGCTGCAGAA GGAAAACAGC AAACGCTGGA ATCCCGAAAT
  AAV3    ATTGAGTGGG AGCTACAGAA AGAAAACAGC AAACGTTGGA ATCCAGAGAT
  AAV8    ATTGAATGGG AGCTGCAGAA GGAAAACAGC AAGCGCTGGA ACCCCGAGAT
  AAV9    ATCGAGTGGG AGCTGCAGAA AGAAAACAGC AAGCGCTGGA ATCCAGAGAT
  AAV7    ATCGAGTGGG AGCTGCAGAA GGAAAACAGC AAGCGCTGGA ACCCGGAGAT
  44_2    ATTGAATGGG AGCTGCAGAA AGAAAACAGC AAACGCTGGA ACCCAGAGAT
```

图1AAAI

```
          4351                                                    4400
  42_2    TCAGTACACC TCAAATTATG CCAAGTCTAA TAAT.GTGGA ATTTGCTGTC
  42_8    TCAGTATACT TCCAACTACT ACAAATCTAC AAAT.GTGGA CTTTGCTGTC
 42_15    TCAGTATACT TCCAACTACT ACAAATCTAC AAAT.GTGGA CTTTGCTGTC
 42_5b    TCAGTATACT TCCAACTACT ACAAATCTAC AAAT.GTGGA CTTTGCTGTC
 42_1b    TCAGTATACT TCCAACTACT ACAAATCTAC AAAT.GTGGA CTTTGCTGTC
 42_13    TCAGTATACT TCCAACTACT ACAAATCTAC AAAT.GTGGA CTTTGCTGTC
 42_3a    TCAGTATACT TCCAACTACT ACAAATCTAC AAAT.GTGGA CTTTGCTGTC
  42_4    TCAGTATACT TCCAACTACT ACAAATCTAC AAAT.GTGGA CTTTGCTGTC
 42_5a    TCAGTATACT TCCAACTACT ACAAATCTAC AAAT.GTGGA CTTTGCTGTC
 42_10    TCAGTACACC TCAAATTATG CCAAGTCTAA TAAT.GTGGA ATTTGCTGTC
 42_3b    TCAGTACACC TCAAATTATG CCAAGTCTAA TAAT.GTGGA ATTTGCTGTC
 42_11    TCAGTACACC TCAAATTATG CCAAGTCTAA TAAT.GTGGA ATTTGCTGTC
 42_6b    TCAGTACACC TCAAATTATG CCAAGTCTAA TAAT.GTGGA ATTTGCTGTC
  43_1    TCAGTATACT TCCAACTACT ACAAATCTAC AAAT.GTGGA CTTTGCTGTC
  43_5    TCAGTATACT TCCAACTACT ACAAATCTAC AAAT.GTGGA CTTTGCTGTC
 43_12    TCAGTATACT TCCAACTACT ACAAATCTAC AAAT.GTGGA CTTTGCTGTC
 43_20    TCAATACACT TCCAACTACT ACAAATCTAC AAAT.GTGGA CTTTGCTGTC
 43_21    TCAATACACT TCCAACTACT ACAAATCTAC AAAT.GTGGA CTTTGCTGTC
 43_23    TCAATACACT TCCAACTACT ACAAATCTAC AAAT.GTGGA CTTTGCTGTC
 43_25    TCAATACACT TCCAACTACT ACAAATCTAC AAAT.GTGGA CTTTGCTGTC
  44_1    TCAATACACT TCCAACTACT ACAAATCTAC AAAT.GTGGA CTTCGCTGTT
  44_5    TCAATACACT TCCAACTACT ACAAATCTAC AAAT.GTGGA CTTTGCTGTT
223_10    TCAGTACACC TCCAACTTTG ACAAACAGAC TGGA.GTGGA CTTTGCTGTT
 223_2    TCAGTACACC TCCAACTTTG ACAAACAGAC TGGA.GTGGA CTTTGCTGTT
 223_4    TCAGTACACC TCCAACTTTG ACAAACAGAC TGGA.GTGGA CTTTGCTGTT
 223_5    TCAGTACACC TCCAACTTTG ACAAACAGAC TGGA.GTGGA CTTTGCTGTT
 223_6    TCAGTACACC TCCAACTTTG ACAAACAGAC TGGA.GTGGA CTTTGCTGTT
 223_7    TCAGTACACC TCCAACTTTG ACAAACAGAC TGGA.GTGGA CTTTGCTGTT
  A3_4    TCAGTACACC TCCAACTACA ACAAGTCGGT GAAT.GTGGA GTTTACCGTG
  A3_5    TCAGTACACC TCCAACTACA ACAAGTCGGT GAAT.GTGGA GTTTACCGTG
  A3_7    TCAGTACACC TCCAACTACA ACAAGTCGGT GAAT.GTGGA GTTTACCGTG
  A3_3    TCAGTACACC TCCAACTACA ACAAGTCGGT GAAT.GTGGA GTTTACCGTG
 42_12    ...GTATACT TCCAACTACT ACAAATCTAC AAAT.GTGGA CTTTGCTGTC
  AAV1    GCAGTACACA TCCAATTATG CAAAATCTGC CAAC.GTTGA TTTTACTGTG
  AAV2    TCAGTACACT TCCAACTACA ACAAGTCTGT TAATCGTGGA CTT.ACCGTG
  AAV3    TCAGTACACT TCCAACTACA ACAAGTCTGT TAAT.GTGGA CTTTACTGTA
  AAV8    CCAGTACACC TCCAACTACT ACAAATCTAC AAGT.GTGGA CTTTGCTGTT
  AAV9    CCAGTATACT TCAAACTACT ACAAATCTAC AAAT.GTGGA CTTTGCTGTC
  AAV7    TCAGTACACC TCCAACTTTG AAAAGCAGAC TGGT.GTGGA CTTTGCCGTT
  44_2    TCAATACACT TCCAACTACT ACAAATCTAC AAAT.GTGGA CTTTGCTGTT
```

图1AAAJ

```
          4401                                                    4450
  42_2    AACAACGAAG GGGTTTATAC TGAGCCTCGC CCCATTGGCA CCCGTTACCT
  42_8    AATACTGAGG GTACTTATTC AGAGCCTCGC CCCATTGGCA CCCGTTACCT
 42_15    AATACTGAGG GTACTTATTC AGAGCCTCGC CCCATTGGCA CCCGTTACCT
 42_5b    AATACTGAGG GTACTTATTC AGAGCCTCGC CCCATTGGCA CCCGTTACCT
 42_1b    AATACTGAGG GTACTTATTC AGAGCCTCGC CCCATTGGCA CCCGTTACCT
 42_13    AATACTGAGG GTACTTATTC AGAGCCTCGC CCCATTGGCA CCCGTTACCT
 42_3a    AATACTGAGG GTACTTATTC AGAGCCTCGC CCCATTGGCA CCCGTTACCT
  42_4    AATACTGAGG GTACTTATTC AGAGCCTCGC CCCATTGGCA CCCGTTACCT
 42_5a    AATACTGAGG GTACTTATTC AGAGCCTCGC CCCATTGGCA CCCGTTACCT
 42_10    AACAACGAAG GGGTTTATAC TGAGCCTCGC CCCATTGGCA CCCGTTACCT
 42_3b    AACAACGAAG GGGTTTATAC TGAGCCTCGC CCCATTGGCA CCCGTTACCT
 42_11    AACAACGAAG GGGTTTATAC TGAGCCTCGC CCCATTGGCA CCCGTTACCT
 42_6b    AACAACGAAG GGGTTTATAC TGAGCCTCGC CCCATTGGCA CCCGTTACCT
  43_1    AATACTGAGG GTACTTATTC AGAGCCTCGC CCCATTGGCA CTCGTTATCT
  43_5    AATACCGAGG GTACTTATTC AGAGCCTCGC CCCATTGGCA CTCGTTATCT
 43_12    AATACTGAGG GTACTTATTC AGAGCCTCGC CCCATTGGCA CTCGTTATCT
 43_20    AACACGGAAG GAGTTTATAG CGAGCCTCGC CCCATTGGCA CCCGTTACCT
 43_21    AACACGGAAG GAGTTTATAG CGAGCCTCGC CCCATTGGCA CCCGTTACCT
 43_23    AACACGGAAG GAGTTTATAG CGAGCCTCGC CCCATTGGCA CCCGTTACCT
 43_25    AACACGGAGG GGGTTTATAG CGAGCCTCGC CCCATTGGCA CCCGTTACCT
  44_1    AACACAGATG GCACTTATTC TGAGCCTCGC CCCATTGGCA CCCGTTACCT
  44_5    AACACAGATG GCACTTATTC TGAGCCTCGC CCCATTGGCA CCCGTTACCT
223_10    GACAGCCAGG GTGTTTACTC TGAGCCT... .......... ..........
 223_2    GACAGCCAGG GTGTTTACTC TGAGCCT... .......... ..........
 223_4    GACAGCCAGG GTGTTTACTC TGAGCCT... .......... ..........
 223_5    GACAGCCAGG GTGTTTACTC TGAGCCT... .......... ..........
 223_6    GACAGCCAGG GTGTTTACTC TGAGCCT... .......... ..........
 223_7    GACAGCCAGG GTGTTTACTC TGAGCCT... .......... ..........
  A3_4    GACGCAAACG GTGTTTATTC TGAACCCCGC CCTATTGGCA CTCGTTACCT
  A3_5    GACGCAAACG GTGTTTATTC TGAACCCCGC CCTATTGGCA CTCGTTACCT
  A3_7    GACGCAAACG GTGTTTATTC TGAACCCCGC CCTATTGGCA CTCGTTACCT
  A3_3    GACGCAAACG GTGTTTATTC TGAACCCCGC CCTATTGGCA CTCGTTACCT
 42_12    AATACTGAGG GTACTTATTC AGAGCCTCGC CCCATTGGCA CCCGTTACCT
  AAV1    GACAACAATG GACTTTATAC TGAGCCTCGC CCCATTGGCA CCCGTTACCT
  AAV2    GATACTAATG GCGTGTATTC AGAGCCTCGC CCCATTGGCA CCAGATACCT
  AAV3    GACACTAATG GTGTTTATAG TGAACCTCGC CCTATTGGAA CCCGGTATCT
  AAV8    AATACAGAAG GCGTGTACTC TGAACCCCGC CCCATTGGCA CCCGTTACCT
  AAV9    AATACCGAAG GTGTTTACTC TGAGCCTCGC CCCATTGGTA CTCGTTACCT
  AAV7    GACAGCCAGG GTGTTTACTC TGAGCCTCGC CCTATTGGCA CTCGTTACCT
  44_2    AACACAGATG GCACTTATTC TGAGCCTCGC CCCATCGGCA CCCGTTACCT
```

图1AAAK

```
          4451                                                                4500
                                VP1-3 终止          Poly A 信号
  42_2    CACCCGTAAC CTGTAATTGC CTGTTAATCA ATAAACCGGT TAATTCGTTT
  42_8    CACCCGTAAC CTGTAATTGC CTGTTAATCA ATAAACCGGC TAATTCGTTT
 42_15    CACCCGTAAC CTGTAATTGC CTGTTAATCA ATAAACCGGT TAATTCGTTT
 42_5b    CACCCGTAAC CTGTAATTGC CTGTTAATCA ATAAACCGGT TAATTCGTTT
 42_1b    CACCCGTAAC CTGTAATTGC CTGTTAATCA ATAAACCGGT TGATTCGTTT
 42_13    CACCCGTAGC CTGTAATTGC CTGTTAATCA ATAAACCGGT TGATTCGTTT
 42_3a    CACCCGTAAC CTGTAATTGC CTGTTAATCA ATAAACCGGT TAATTCGTTT
  42_4    CACCCGTAAC CTGTAATTGC CTGTTAATCA ATAAACCGGT TAATTCGTTT
 42_5a    CACCCGTAAC CTGTAATTGC CTGTTAATCA ATAAACCGGT TAATTCGTTT
 42_10    CACCCGTAAC CTGTAATTGC CTGTTAATCA ATAAACCGGT TAATTCGTTT
 42_3b    CACCCGTAAC CTGTAATTGC CTGTTAATCA ATAAACCGGT TAATTCGTTT
 42_11    CACCCGTAAC CTGTAATTAC TTGTTAATCA ATAAACCGGT TGATTCGTTT
 42_6b    CACCCGTAAC CTGTAATTGC CTGTTAATCA ATAAACCGGT TAATTCGTTT
  43_1    CACCCGTAAT CTGTAATTGC TTGTTAATCA ATAAACCGGT ..........
  43_5    CACCCGTAAT CTGTAATTGC TTGTTAATCA ATAAACCGGT TAATTCGTTT
 43_12    CACCCGTAAT CTGTAATTGC TTGTTAATCA ATAAACCGGT TAATTCGTTT
 43_20    CACCCGCAAC CTGTAATTAC ATGTTAATCA ATAAACCGGT TAATTCGTTT
 43_21    CACCCGCAAC CTGTAATTAC ATGTTAATCA ATAAACCGGT TAATTCGTTT
 43_23    CACCCGCAAC CTGTAATTAC ATGTTAATCA ATAAACCGGT TAATTCGTTT
 43_25    CACCCGCAAC CTGTAATTAC ATGTTAATCA ATAAACCGGT TAATTCGTTT
  44_1    CACCCGTAAT CTGTAATTGC TCGTTAATCA ATAAACCGGT TGATTCGTTT
  44_5    CACCCGTAAT CTGTAATTGC TTGTTAATCA ATAAACCGGT TGATTCGTTT
223_10    .......... .......... .......... .......... ..........
 223_2    .......... .......... .......... .......... ..........
 223_4    .......... .......... .......... .......... ..........
 223_5    .......... .......... .......... .......... ..........
 223_6    .......... .......... .......... .......... ..........
 223_7    .......... .......... .......... .......... ..........
  A3_4    TACCCGGAAC TTGTAATTTC CTGTTAATGA ATAAACCGAT TTATGCGTTT
  A3_5    TACCCGGAAC TTGTAATTTC CTGTTAATGA ATAAACCGAT TTATGCGTTT
  A3_7    TACCCGGAAC TTGTAATTTC CTGTTAATGA ATAAACCGAT TTATGCGTTT
  A3_3    TACCCGGAAC TTGTAATTTC CTGTTAATGA ATAAGCCGAT TTATGCGTTT
 42_12    CACCCGTAAC CTGTAATTGC CTGTTAATCA ATAAACCGGT TAATTCGTTT
  AAV1    TACCCGTCCC CTGTAATTAC GTGTTAATCA ATAAACCGGT TGATTCGTTT
  AAV2    GACTCGTAAT CTGTAATTGC TTGTTAATCA ATAAACCGTT TAATTCGTTT
  AAV3    CACACGAAAC TTGTGAATCC TGGTTAATCA ATAAACCGTT TAATTCGTTT
  AAV8    CACCCGTAAT CTGTAATTGC CTGTTAATCA ATAAACCGGT TGATTCGTTT
  AAV9    CACCCGTAAT TTGTAATTGC CTGTTAATCA ATAAACCGGT TAATTCGTTT
  AAV7    CACCCGTAAT CTGTAATTGC ATGTTAATCA ATAAACCGGT TGATTCGTTT
  44_2    CACCCGTAAT CTGTAATTGC TTGTTAATCA ATAAACCGGT TGATTCGTTT
                                vp1-3 终止          PolyA 信号
```

图1AAAL

```
          4501                                                    4550

  42_2    CAGTTGAACT TTGGTCTC.T GCGAAGGGCG AATTC..... ..........
  42_8    CAGTTGAACT TTGGTCTC.T GCGAAGGGCG AATTC..... ..........
 42_15    CAGTTGAACT TTGGTCTC.T GCGAAGGGCG AATTC..... ..........
 42_5b    CAGTTGAACT TTGGTCTC.T GCGAAGGGCG AATTCGTTTA AACCTGCAGG
 42_1b    CAGTTGAACT TTGGTCTC.. ...AAGGGCG AATTC..... ..........
 42_13    CAGTTGAACT TTGGTCTC.T GCGAAGGGCG AATTC..... ..........
 42_3a    CAGTTGAACT TTGGTCTC.T GCGAAGGGCG AATTC..... ..........
  42_4    CAGTTGAACT TTGGTCTC.T GCGAAGGGCG AATTC..... ..........
 42_5a    CAGTTGAACT TTGGTCTC.T GCGAAGGGCG AATTC..... ..........
 42_10    CAGTTGAACT TTGGTC.... ...AAGGGCG AATTC..... ..........
 42_3b    CAGTTGAACT TTGGTCTC.T GCGAAGGGCG AATTC..... ..........
 42_11    CAGTTGAACT TTGGTCTC.T GCGAAGGGCG AATTC..... ..........
 42_6b    CAGTTGAACT TTGGTCTC.T GCGAAGGGCG AATTC..... ..........
  43_1    .......... .......... .......... .......... ..........
  43_5    CAGTTGAACT TTGGTCTC.T GCGAAGGGCG AATTCGTTTA AACCTGCAGG
 43_12    CAGTTGAACT TTGGTCTC.T GCGAAGGGCG AATTC..... ..........
 43_20    CAGTTGAACT TTGGTCTC.T GCGAAGGGCG AATTC..... ..........
 43_21    CAGTTGAACT TTGGTCTC.T GCGAAGGGCG AATTC..... ..........
 43_23    CAGTTGAACT TTGGTCTC.T GCGAAGGGCG AATTC..... ..........
 43_25    CAGTTGAACT TTGGTCTC.T GCGAAGGGCG AATTC..... ..........
  44_1    CAGTTGAACT TTGGTCTC.T GCGAAGGGCG AATTC..... ..........
  44_5    CAGTTGAACT TTGGTCTC.T GCGAAGGGCG AATTC..... ..........
223_10    .......... .......... .......... .......... ..........
 223_2    .......... .......... .......... .......... ..........
 223_4    .......... .......... .......... .......... ..........
 223_5    .......... .......... .......... .......... ..........
 223_6    .......... .......... .......... .......... ..........
 223_7    .......... .......... .......... .......... ..........
  A3_4    CAGTTGAACT TTGGTCTC.T GCGAAGGGCG AATTCGC.GG CCGCTA....
  A3_5    CAGTTGAACT TTGGTCTC.T GCGAAGGGCG AATTC..... ..........
  A3_7    CAGTTGAACT TTGGTCTC.T GCGAAGGGCG AATTC..... ..........
  A3_3    CAGTTGAACT TTGGTCTC.T GCGAAGGGCG AATTCGT.TT AAACCT....
 42_12    CAGTTGAACT TTGGTCTC.T GCGAAGGGCG AATTC..... ..........
  AAV1    CAGTTGAACT TTGGTCTCCT GTCCTTCTTA TCTTATCGGT TACCATGGTT
  AAV2    CAGTTGAACT TTGGTCTC.T GCGTATTTCT ..TTCTT.AT CTAGTTTCCA
  AAV3    CAGTTGAACT TTGGCTCT.T GTGCACTTCT TTATCTTTAT CTTGTTTCCA
  AAV8    CAGTTGAACT TTGGTCTC.T GCG....... .......... ..........
  AAV9    CAGTTGAACT TTGGTCTC.T GCG....... .......... ..........
  AAV7    CAGTTGAACT TTGGTCTCCT GTGCTTCTTA TCTTATCGGT TTCCATAGCA
  44_2    CAGTTGAACT TTGGTCTC.T GCGAAGGGCG AATTC..... ..........
```

图1AAAM

```
          4551                                                        4600
  42_2    .......... .......... .......... .......... ..........
  42_8    .......... .......... .......... .......... ..........
 42_15    .......... .......... .......... .......... ..........
 42_5b    ACTAGTCCCT TTAGTGAGGG TTAATTCTGA G......... ..........
 42_1b    .......... .......... .......... .......... ..........
 42_13    .......... .......... .......... .......... ..........
 42_3a    .......... .......... .......... .......... ..........
  42_4    .......... .......... .......... .......... ..........
 42_5a    .......... .......... .......... .......... ..........
 42_10    .......... .......... .......... .......... ..........
 42_3b    .......... .......... .......... .......... ..........
 42_11    .......... .......... .......... .......... ..........
 42_6b    .......... .......... .......... .......... ..........
  43_1    .......... .......... .......... .......... ..........
  43_5    AC........ .......... .......... .......... ..........
 43_12    .......... .......... .......... .......... ..........
 43_20    .......... .......... .......... .......... ..........
 43_21    .......... .......... .......... .......... ..........
 43_23    .......... .......... .......... .......... ..........
 43_25    .......... .......... .......... .......... ..........
  44_1    .......... .......... .......... .......... ..........
  44_5    .......... .......... .......... .......... ..........
223_10    .......... .......... .......... .......... ..........
 223_2    .......... .......... .......... .......... ..........
 223_4    .......... .......... .......... .......... ..........
 223_5    .......... .......... .......... .......... ..........
 223_6    .......... .......... .......... .......... ..........
 223_7    .......... .......... .......... .......... ..........
  A3_4    .......... .......... .......... .......... ..........
  A3_5    .......... .......... .......... .......... ..........
  A3_7    .......... .......... .......... .......... ..........
  A3_3    .......... .......... .......... .......... ..........
 42_12    .......... .......... .......... .......... ..........
  AAV1    ATAGCTTACA CATTAACTGC TTGGTTGCGC T......... ..........
  AAV2    TGGCTAC... GTAGATAAGT AGC....... .......... ..........
  AAV3    TGGCTACTGC GTAGATAAGC AGCGGCCTGC GGCGCTTGCG CTTCGCGGTT
  AAV8    .......... .......... .......... .......... ..........
  AAV9    .......... .......... .......... .......... ..........
  AAV7    ACTGGTTACA CATTAACTGC TTGGGTGCGC TTCACGATAA GAACACTGAC
  44_2    .......... .......... .......... .......... ..........
```

图1AAAN

```
          4601                                                      4650
  42_2    .......... .......... .......... .......... ..........
  42_8    .......... .......... .......... .......... ..........
 42_15    .......... .......... .......... .......... ..........
 42_5b    ....CTTGGC GTAATCATGG GTCATAG... .......... ..........
 42_1b    .......... .......... .......... .......... ..........
 42_13    .......... .......... .......... .......... ..........
 42_3a    .......... .......... .......... .......... ..........
  42_4    .......... .......... .......... .......... ..........
 42_5a    .......... .......... .......... .......... ..........
 42_10    .......... .......... .......... .......... ..........
 42_3b    .......... .......... .......... .......... ..........
 42_11    .......... .......... .......... .......... ..........
 42_6b    .......... .......... .......... .......... ..........
  43_1    .......... .......... .......... .......... ..........
  43_5    .......... .......... .......... .......... ..........
 43_12    .......... .......... .......... .......... ..........
 43_20    .......... .......... .......... .......... ..........
 43_21    .......... .......... .......... .......... ..........
 43_23    .......... .......... .......... .......... ..........
 43_25    .......... .......... .......... .......... ..........
  44_1    .......... .......... .......... .......... ..........
  44_5    .......... .......... .......... .......... ..........
223_10    .......... .......... .......... .......... ..........
 223_2    .......... .......... .......... .......... ..........
 223_4    .......... .......... .......... .......... ..........
 223_5    .......... .......... .......... .......... ..........
 223_6    .......... .......... .......... .......... ..........
 223_7    .......... .......... .......... .......... ..........
  A3_4    .......... .......... .......... .......... ..........
  A3_5    .......... .......... .......... .......... ..........
  A3_7    .......... .......... .......... .......... ..........
  A3_3    .......... .......... .......... .......... ..........
 42_12    .......... .......... .......... .......... ..........
  AAV1    ....TCGCGA TAAAAGACTT ACGTCATCGG GTTACCCCTA GTGATGGAGT
  AAV2    ....ATGGCG GGTTAATCAT TAACTACAAG GA.ACCCCTA GTGATGGAGT
  AAV3    TACAACTGCT GGTTAATATT TAACTCTCGC CATACCTCTA GTGATGGAGT
  AAV8    .......... .......... .......... .......... ..........
  AAV9    .......... .......... .......... .......... ..........
  AAV7    .......... .......... ..GTCACCGC GGTACCCCTA GTGATGGAGT
  44_2    .......... .......... .......... .......... ..........
```

图1AAAO

```
          4651                                                        4700
  42_2    .......... .......... .......... .......... ..........
  42_8    .......... .......... .......... .......... ..........
 42_15    .......... .......... .......... .......... ..........
 42_5b    .......... .......... .......... .......... ..........
 42_1b    .......... .......... .......... .......... ..........
 42_13    .......... .......... .......... .......... ..........
 42_3a    .......... .......... .......... .......... ..........
  42_4    .......... .......... .......... .......... ..........
 42_5a    .......... .......... .......... .......... ..........
 42_10    .......... .......... .......... .......... ..........
 42_3b    .......... .......... .......... .......... ..........
 42_11    .......... .......... .......... .......... ..........
 42_6b    .......... .......... .......... .......... ..........
  43_1    .......... .......... .......... .......... ..........
  43_5    .......... .......... .......... .......... ..........
 43_12    .......... .......... .......... .......... ..........
 43_20    .......... .......... .......... .......... ..........
 43_21    .......... .......... .......... .......... ..........
 43_23    .......... .......... .......... .......... ..........
 43_25    .......... .......... .......... .......... ..........
  44_1    .......... .......... .......... .......... ..........
  44_5    .......... .......... .......... .......... ..........
223_10    .......... .......... .......... .......... ..........
 223_2    .......... .......... .......... .......... ..........
 223_4    .......... .......... .......... .......... ..........
 223_5    .......... .......... .......... .......... ..........
 223_6    .......... .......... .......... .......... ..........
 223_7    .......... .......... .......... .......... ..........
  A3_4    .......... .......... .......... .......... ..........
  A3_5    .......... .......... .......... .......... ..........
  A3_7    .......... .......... .......... .......... ..........
  A3_3    .......... .......... .......... .......... ..........
 42_12    .......... .......... .......... .......... ..........
  AAV1    TGCCCACTCC CTCTCTGCGC GCTCGCTCGC TCGGTGGGGC CTGCGGACCA
  AAV2    TGGCCACTCC CTCTCTGCGC GCTCGCTCGC TCACTGAGGC CGGGCGACCA
  AAV3    TGGCCACTCC CTCTATGCGC ACTCGCTCGC TCGGTGGGGC CTGGCGACCA
  AAV8    .......... .......... .......... .......... ..........
  AAV9    .......... .......... .......... .......... ..........
  AAV7    TGGCCACTCC CTCTATGCGC GCTCGCTCGC TCGGTGGGGC CTGCGGACCA
  44_2    .......... .......... .......... .......... ..........
```

图1AAAP

```
          4701                                                        4750
  42_2    .......... .......... .......... .......... ..........
  42_8    .......... .......... .......... .......... ..........
 42_15    .......... .......... .......... .......... ..........
 42_5b    .......... .......... .......... .......... ..........
 42_1b    .......... .......... .......... .......... ..........
 42_13    .......... .......... .......... .......... ..........
 42_3a    .......... .......... .......... .......... ..........
  42_4    .......... .......... .......... .......... ..........
 42_5a    .......... .......... .......... .......... ..........
 42_10    .......... .......... .......... .......... ..........
 42_3b    .......... .......... .......... .......... ..........
 42_11    .......... .......... .......... .......... ..........
 42_6b    .......... .......... .......... .......... ..........
  43_1    .......... .......... .......... .......... ..........
  43_5    .......... .......... .......... .......... ..........
 43_12    .......... .......... .......... .......... ..........
 43_20    .......... .......... .......... .......... ..........
 43_21    .......... .......... .......... .......... ..........
 43_23    .......... .......... .......... .......... ..........
 43_25    .......... .......... .......... .......... ..........
  44_1    .......... .......... .......... .......... ..........
  44_5    .......... .......... .......... .......... ..........
223_10    .......... .......... .......... .......... ..........
 223_2    .......... .......... .......... .......... ..........
 223_4    .......... .......... .......... .......... ..........
 223_5    .......... .......... .......... .......... ..........
 223_6    .......... .......... .......... .......... ..........
 223_7    .......... .......... .......... .......... ..........
  A3_4    .......... .......... .......... .......... ..........
  A3_5    .......... .......... .......... .......... ..........
  A3_7    .......... .......... .......... .......... ..........
  A3_3    .......... .......... .......... .......... ..........
 42_12    .......... .......... .......... .......... ..........
  AAV1    AAGGTCCGCA GACGGCAGAG CTCTGCTCTG CCGGCCCCAC CGAGCGAGCG
  AAV2    AAGGTCGCCC GACGCCCGGG CTTTGCCCGG GCGGCCTCAG TGAGCGAGCG
  AAV3    AAGGTCGCCA GACGGACGTG CTTTGCACGT CCGGCCCCAC CGAGCGAGCG
  AAV8    .......... .......... .......... .......... ..........
  AAV9    .......... .......... .......... .......... ..........
  AAV7    AAGGTCCGCA GACGGCAGAG CTCTGCTCTG CCGGCCCCAC CGAGCGAGCG
  44_2    .......... .......... .......... .......... ..........
```

图1AAAQ

```
          4751                                4774
 42_2     .......... .......... ....
 42_8     .......... .......... ....
42_15     .......... .......... ....
42_5b     .......... .......... ....
42_1b     .......... .......... ....
42_13     .......... .......... ....
42_3a     .......... .......... ....
 42_4     .......... .......... ....
42_5a     .......... .......... ....
42_10     .......... .......... ....
42_3b     .......... .......... ....
42_11     .......... .......... ....
42_6b     .......... .......... ....
 43_1     .......... .......... ....
 43_5     .......... .......... ....
43_12     .......... .......... ....
43_20     .......... .......... ....
43_21     .......... .......... ....
43_23     .......... .......... ....
43_25     .......... .......... ....
 44_1     .......... .......... ....
 44_5     .......... .......... ....
223_10    .......... .......... ....
223_2     .......... .......... ....
223_4     .......... .......... ....
223_5     .......... .......... ....
223_6     .......... .......... ....
223_7     .......... .......... ....
 A3_4     .......... .......... ....
 A3_5     .......... .......... ....
 A3_7     .......... .......... ....
 A3_3     .......... .......... ....
42_12     .......... .......... ....
 AAV1     AGCGCGCAGA GAGGGAGTGG GCAA
 AAV2     AGCGCGCAGA GAGGGAGTGG CCAA
 AAV3     AGTGCGCATA GAGGGAGTGG CCAA
 AAV8     .......... .......... ....
 AAV9     .......... .......... ....
 AAV7     AGCGCGCATA GAGGGAGTGG CCAA
 44_2     .......... .......... ....
```

图1AAAR

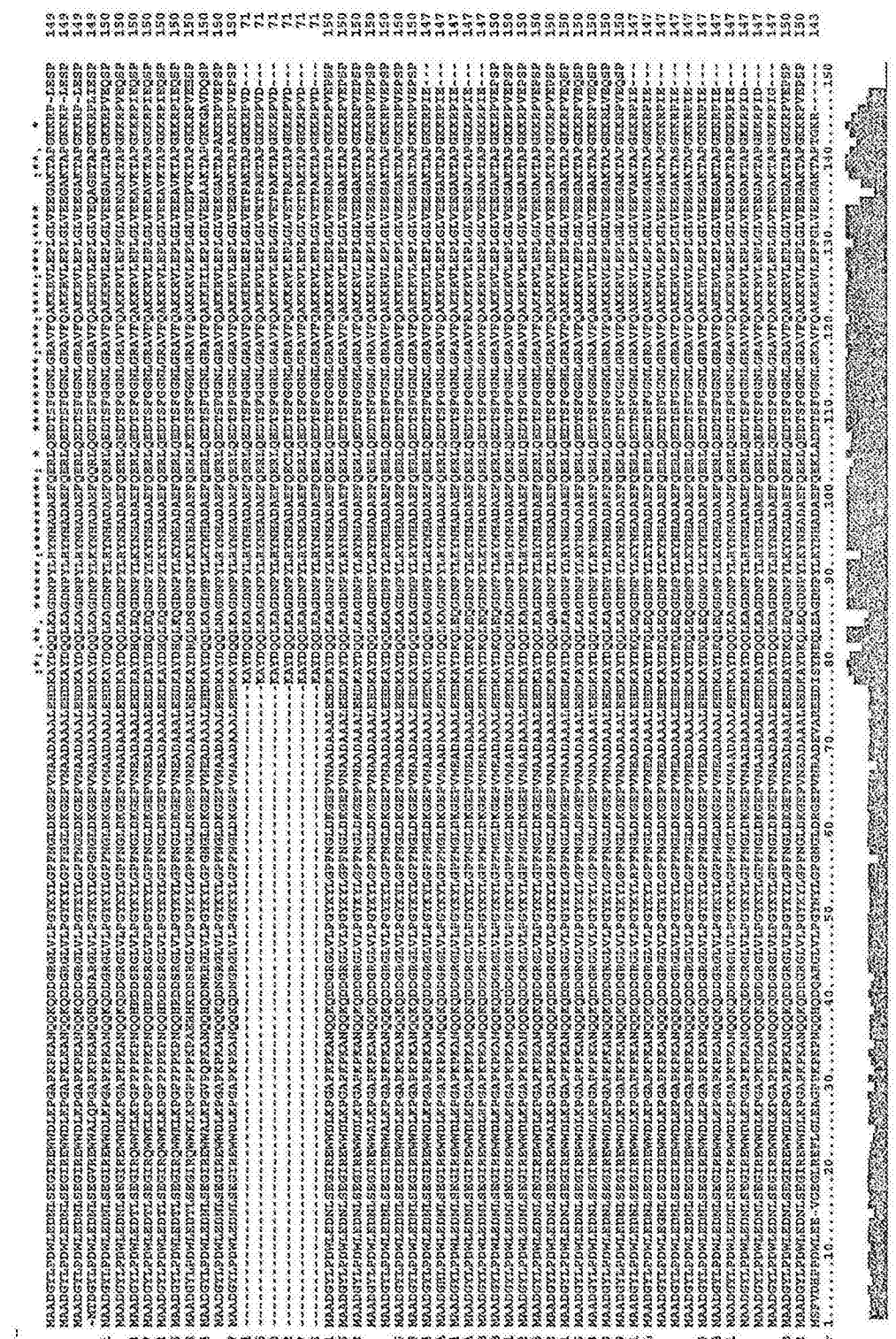


图2A

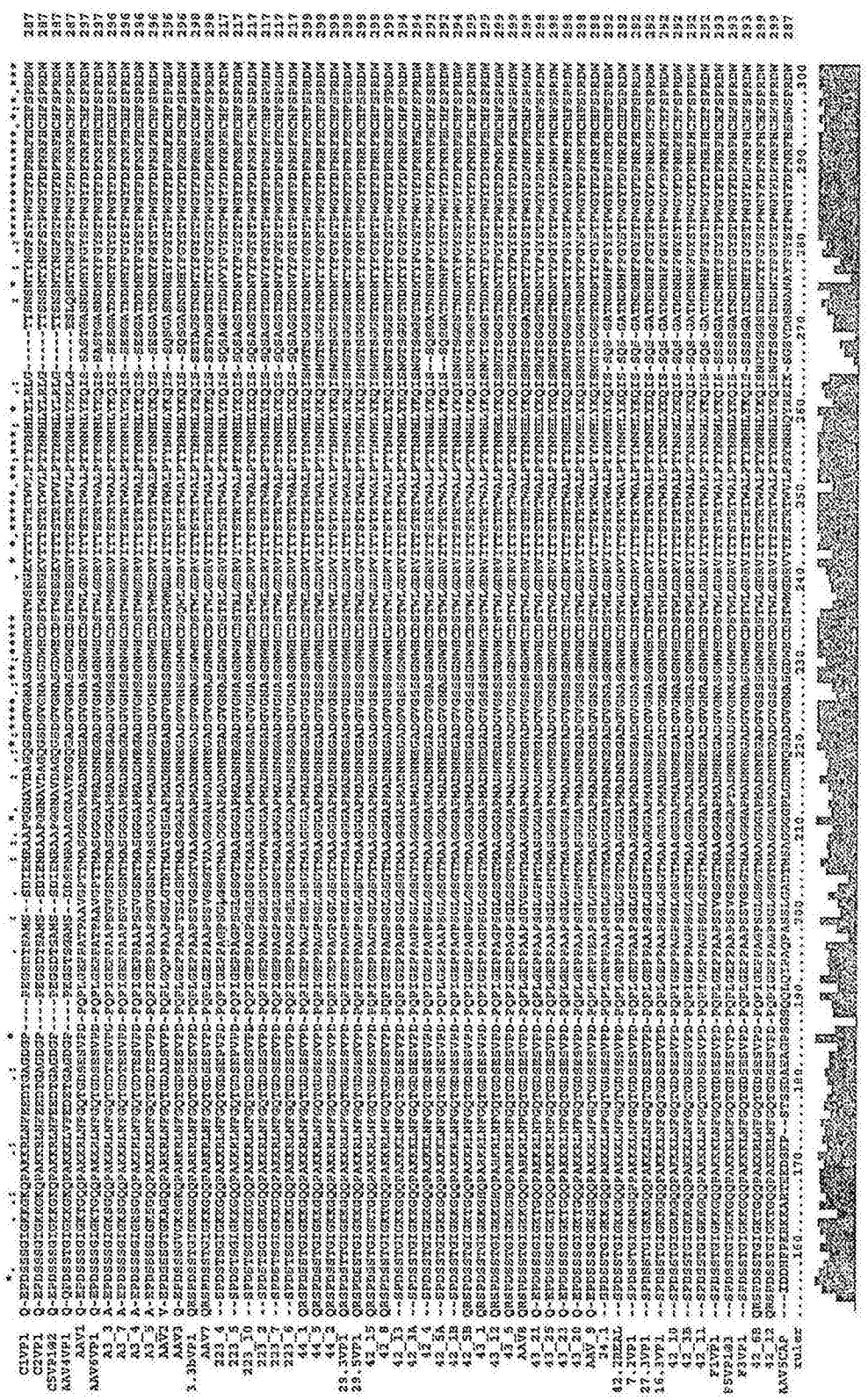


图2B

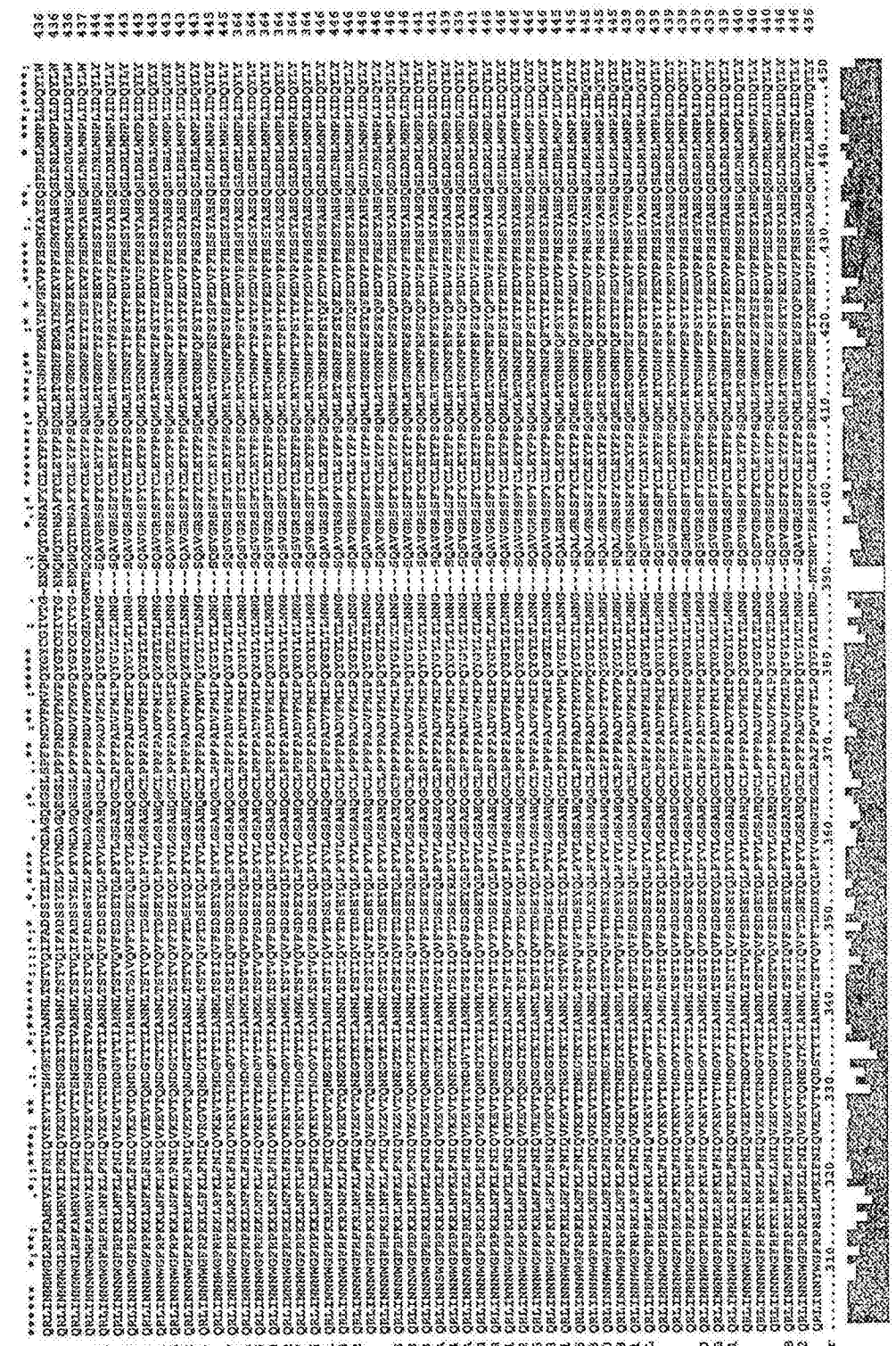


图2C

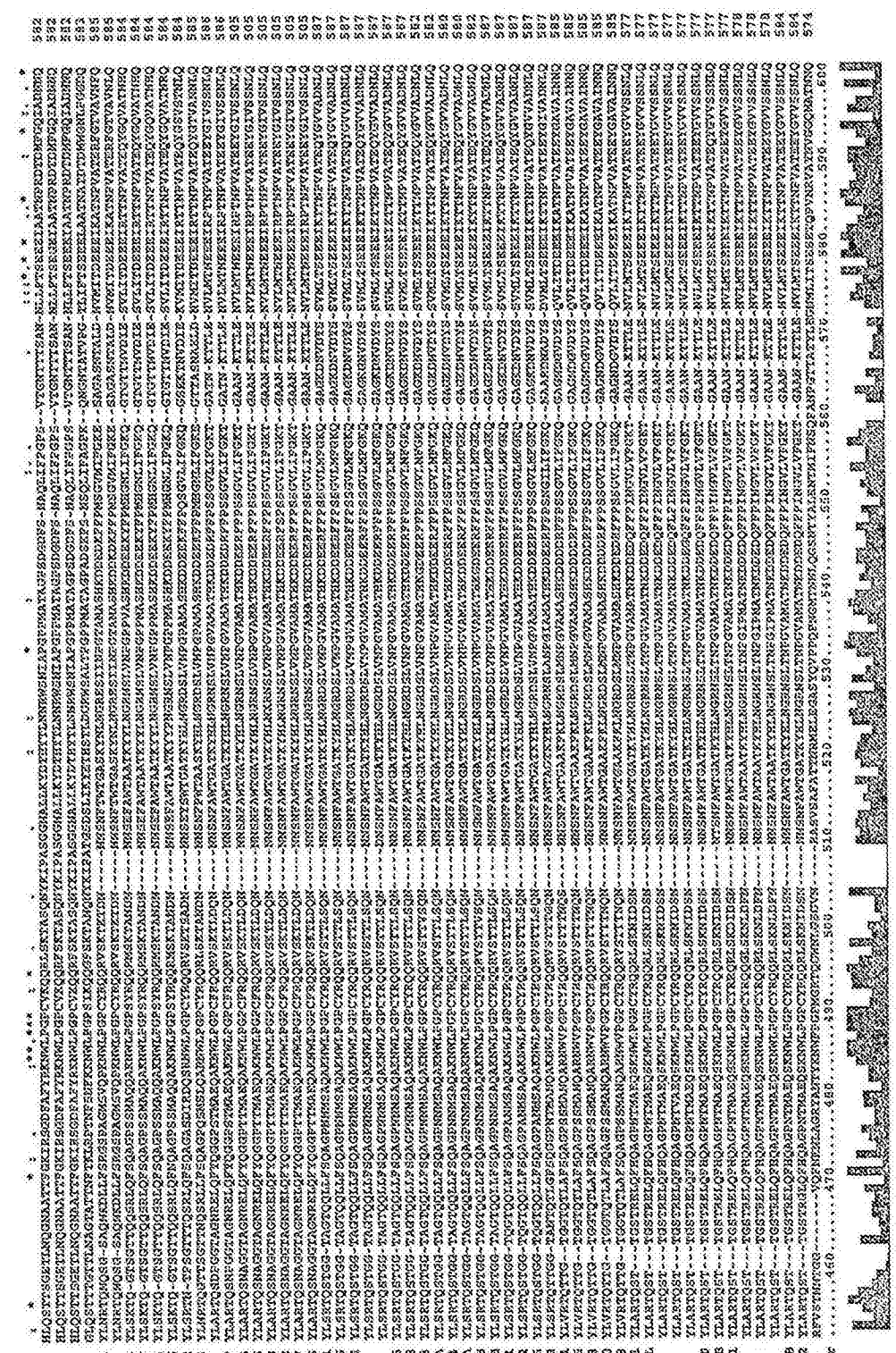


图2D

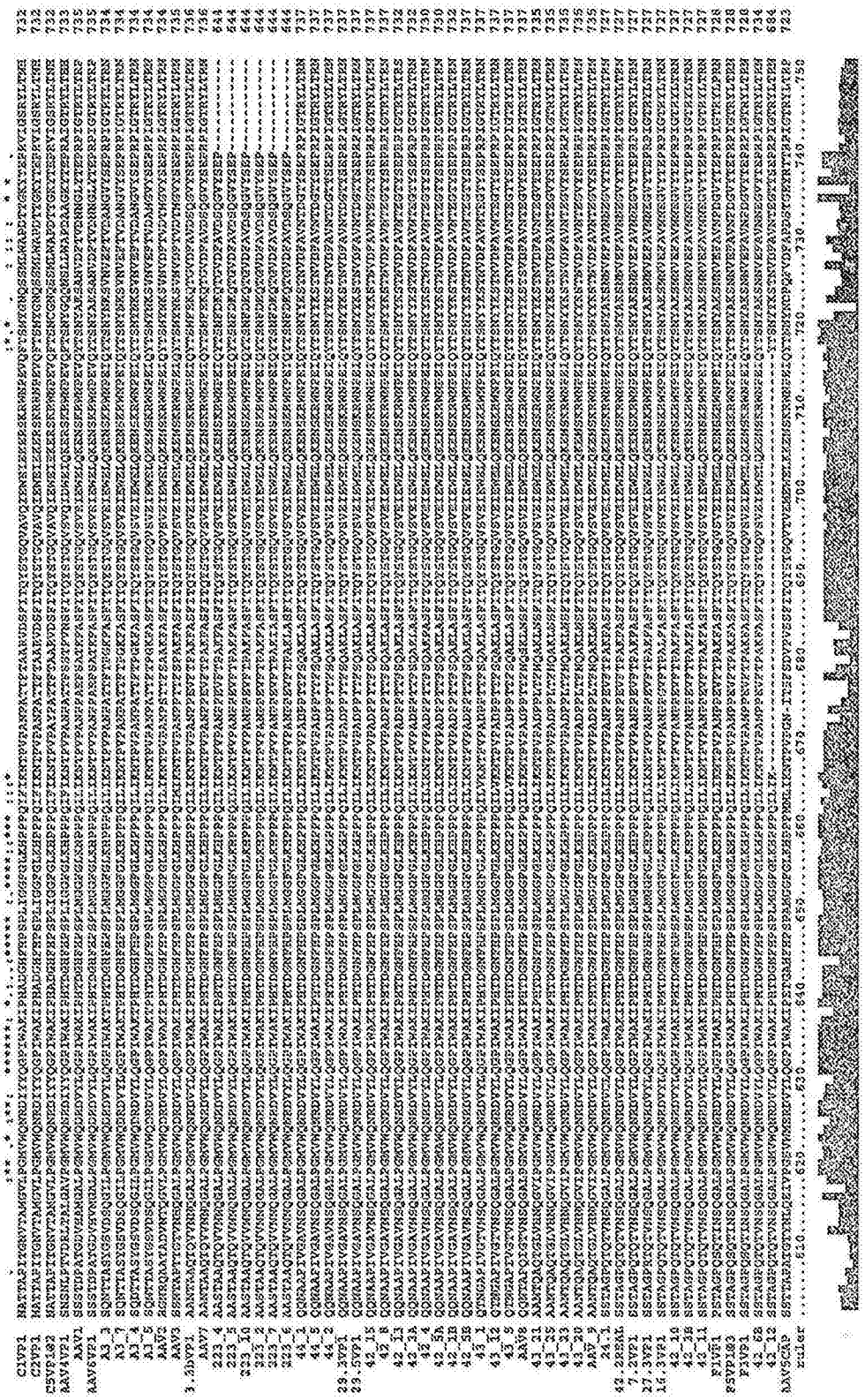


图2E

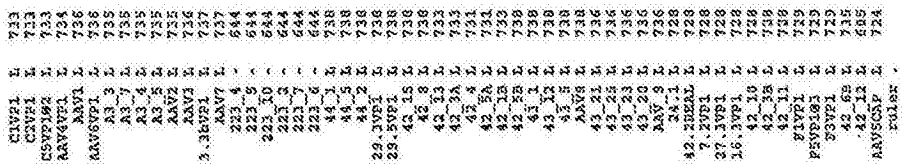


图2F

```
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

Val Gln 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 Leu 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 Val Gln Thr Ile Tyr Arg Gly Val 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 Glu Tyr Ile
                165                 170                 175

Ser Ala Cys Leu Asn Leu Ala 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 Leu Asn
        195                 200                 205

Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr
    210                 215                 220
```

图3A

```
Met Glu Leu Val Gly Trp Leu Val Asp Arg 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 Ala Leu Thr Lys Ser Ala Pro Asp Tyr Leu Val Gly Pro Ser
        275                 280                 285

Leu Pro Ala Asp Ile Lys Thr Asn Arg Ile Tyr Arg Ile Leu Glu Leu
    290                 295                 300

Asn Gly Tyr Asp Pro Ala Tyr Ala Gly Ser Val Phe Leu Gly Trp Ala
305                 310                 315                 320

Gln 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 Ala 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
```

图3B

```
Glu Leu Thr Arg Arg Leu Glu His Asp Phe Gly Lys Val Thr Lys Gln
    450                 455                 460

Glu Val Lys Glu Phe Phe Arg Trp Ala Ser Asp His Val Thr Glu Val
465                 470                 475                 480

Ala His Glu Phe Tyr Val Arg Lys Gly Gly Ala Ser Lys Arg Pro Ala
                485                 490                 495

Pro Asp Asp Ala Asp Ile Ser Glu Pro Lys Arg Ala Cys Pro Ser Val
            500                 505                 510

Ala Asp Pro Ser Thr Ser Asp Ala Glu Gly Ala Pro Val Asp Phe Ala
        515                 520                 525

Asp Arg Tyr Gln Asn Lys Cys Ser Arg His Ala Gly Met Ile Gln Met
    530                 535                 540

Leu Phe Pro Cys Lys Thr Cys Glu Arg Met Asn Gln Asn Phe Asn Ile
545                 550                 555                 560

Cys Phe Thr His Gly Val Arg Asp Cys Leu Glu Cys Phe Pro Gly Val
                565                 570                 575

Ser Glu Ser Gln Pro Val Val Arg Lys Lys Thr Tyr Arg Lys Leu Cys
            580                 585                 590

Ala Ile His His Leu Leu Gly Arg Ala Pro Glu Ile Ala Cys Ser Ala
        595                 600                 605

Cys Asp Leu Val Asn Val Asp Leu Asp Asp Cys Val Ser Glu Gln
    610                 615                 620
```

图3C