(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 102181480 B
(45)授权公告日 2016.01.27

(21)申请号 201110081897.1

(22)申请日 2002.11.12

(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 崔佳佳

(51)Int.Cl.
C12N 15/86(2006.01)
C12N 15/63(2006.01)
C12N 5/10(2006.01)
C07K 14/005(2006.01)
A61K 48/00(2006.01)

(56)对比文件
WO 9720943 A1,1997.06.12,
CN 1233291 A,1999.10.27,
伍志坚等.一种高效的重组腺伴随病毒载体生产系统.《中国科学C辑》.2001,(第5期),
putledge e a et al..Infectious clones and vectors derived from adeno-associatedvirus(AVV)serotypes other than AVV type 2.《 Joural of Virology72 》.1998, 第72卷

审查员 张丽玮

权利要求书1页 说明书48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。

1.一种重组腺伴随病毒(AAV),所述病毒具有 vp1 蛋白、vp2 蛋白和 vp3 蛋白构成的 AAV8 衣壳,其中,所述 vp1 蛋白由 SEQ ID NO:95 所示全长序列构成,所述 AAV8 衣壳包封的核酸具有 AAV 5’反向末端重复序列、与调节序列操作性连接的异源基因和 AAV 3’反向末端重复序列,所述调节序列指导所述异源基因在宿主细胞内表达。

2.如权利要求 1 所述的重组 AAV,其特征在于,所述 AAV vp2 衣壳蛋白由 SEQ ID NO:95 中第 138-738 位所示序列构成。

3.如权利要求 1 所述的重组 AAV,其特征在于,所述 AAV vp3 衣壳蛋白由 SEQ ID NO:95 中第 203-738 位所示序列构成。

4.如权利要求 1 所述的重组 AAV,其特征在于,所述异源基因选自:因子 VIII 序列、α-1 抗胰蛋白酶序列和 HIV 序列。

5.如权利要求 1 所述的重组 AAV,其特征在于,所述调节序列包括组织特异性启动子。

6.如权利要求 5 所述的重组 AAV,其特征在于,所述组织特异性启动子是肝脏特异性启动子。

7.一种制备含 AAV8 衣壳的重组腺伴随病毒的方法,所述方法包括培养宿主细胞的步骤,该宿主细胞包含:(a) 编码 AAV8 衣壳的分子,所述 AAV8 衣壳由 vp1、vp2 和 vp3 蛋白组成,其中 vp1 蛋白由 SEQ ID NO:95 的氨基酸 1-738 所示全长序列构成;(b) 功能性 rep 基因;(c) 包含 AAV 反向末端重复序列(ITR) 和转基因的小基因;以及 (d) 能将该小基因包装到 AAV 衣壳蛋白中的足够的辅助功能。

8.一种用于制备重组腺伴随病毒 AAV8 的包装宿主细胞,所述宿主细胞包含:编码氨基酸序列如 SEQ ID NO:95 所示 AAV8 衣壳蛋白的分子,具有 AAV 反向末端重复序列和与调节序列操作性连接的异源基因的小基因,所述调节序列指导所述异源基因在宿主细胞内表达;和功能性 rep 基因。

9.一种在包含如权利要求 1-6 中任一项所述重组腺伴随病毒的培养基中的宿主细胞。

10.一种包含如权利要求 1-6 中任一项所述重组腺伴随病毒和生理相容性载体的组合物。

11.如权利要求 1-6 中任一项所述重组腺伴随病毒在制备用于将转基因递送到细胞的药物中的应用。

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] 本申请是国际申请号为PCT/US02/33629,国际申请日为2002年11月12日的PCT国际申请进入中国国家阶段后的申请,申请号为02827059.2,发明名称为“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”的发明专利申请的分案申请。

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 IDNO：1]，以及以前报道的AAV1[SEQ IN NO：6]，AAV2[SEQ ID NO：7]和AAV3[SEQ IDNO：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 IDNO：30]；42-13[SEQ ID NO：31]，42-3a[SEQ ID NO：32]，42-4[SEQ ID NO：33]，42-5a[SEQID 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 IDNO：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 IDNO：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 IDNO：55]，A3.7[SEQ ID NO：56]，A3.3[SEQ ID NO：57]，42.12[SEQ ID NO：58]，44.2[SEQID 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 IDNO：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 IDNO：62]，A3-3[SEQ ID NO：66]，A3-7[SEQ ID NO：67]，A3-4[SEQ ID NO：68]，A3-5[SEQID 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[SEQID NO：78]，44-1[SEQ ID NO：79]，44-5[SEQ ID NO：80]，44-2[SEQ ID NO：81]，42-15[SEQID 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 IDNO：91]，43-1[SEQ ID NO：92]，43-12[SEQ ID NO：93]，43-5[SEQ ID NO：94]，43-21[SEQID 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 IDNO：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AAV7 rep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VectorNTI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GCG 6.1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GCG 6.1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:ALaboratory 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 AAV 1[SEQ ID NO:6] 的约 2800 到 3200 区域，以及 AAV 2、AAV3、AAV4、AAV5 和 AAV6 上的相应区域。优选的区域约为 250bp，位于 AAV 1[SEQ ID NO:6] 的 2886 到 3143 区，以及 AAV 2[SEQ IDNO: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 nt 2867-2891,5’-GGTAATTCCTCCGGAAATTGGCATT3’,见图 1。3’引物是基于 AAV2[SEQ ID NO:7] 的 nt 2867-2891,5’-GACTCATCAACAACAACTGGGGATTC-3’。但是,本领域的技术人员可以基于 AAV 1,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 nt 4435-4462,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 nt 1384-1409,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 nt 108-133,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’GCTGCGTCAACTGGACCAATGAGAAC 3’，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’CGCAGAGACCAAAGTTCAACTGAAACGA 3’，SEQ IDNO：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 IDNO：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 AAV 5' ITR，衣壳，rep，和 AAV 3' 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 AAV 1[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

[0068]

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
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

[0069]

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
bb.1	克隆 29.3	狒狒	血	11
bb.2	克隆 29.5	狒狒	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nt 825to 3049，SEQ ID NO:1]；vp2[即AAV7的nt 1234-3049，SEQ ID NO:1]；以及vp3[即AAV7的nt 1434-3049，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nt 725 to 3090,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rep 78[即图1中AAV7的起始密码子334]、rep 68[图1中AAV7的起始密码子nt 334]、rep 52[图1中AAV7的起始密码子1006]以及rep 40[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nt 1-107];AAV 3'ITR[图1中AAV7的nt 4704-4721],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[SEQID NO:62],A3-3[SEQ ID NO:66],A3-7[SEQ ID NO:67],A3-4[SEQ ID NO:68],A3-5[SEQID 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[SEQID NO:78],44-1[SEQ ID NO:79],44-5[SEQ ID NO:80],44-2[SEQ ID NO:81],42-15[SEQID 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 IDNO:91],43-1[SEQ ID NO:92],43-12[SEQ ID NO:93],43-5[SEQ ID NO:94],43-21[SEQID 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 IDNO: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nt 825-3049, 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 IDNO: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 IDNO:36],42-11[SEQ ID NO:37],42-6b[SEQ ID NO:38],43-1[SEQ ID NO:39],43-5[SEQID 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 IDNO: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 IDNO:54],A3.5[SEQ ID NO:55],A3.7[SEQ ID NO:56],A3.3[SEQ ID NO:57],42.12[SEQID 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 IDNO:61],C5[SEQ ID NO:62],A3-3[SEQ ID NO:66],A3-7[SEQ ID NO:67],A3-4[SEQID 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[SEQID 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 IDNO:86],42-3A[SEQ ID NO:87],42-4[SEQ ID NO:88],42-5A[SEQ ID NO:89],42-1B[SEQID 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 IDNO:99],24.1[SEQ ID NO:101],42.2[SEQ ID NO:102],7.2[SEQ ID NO:103],27.3[SEQID 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,aa 146-152 ;HVR2,aa 182-186 ;HVR3,aa 262-264 ;HVR4,aa 381-383 ;HVR5,aa 450-474 ;HVR6,aa 490-495 ;HVR7,aa500-504 ;HVR8,aa 514-522 ;HVR9,aa 534-555 ;HVR10,aa 581-594 ;HVR11,aa 658-667 ;以及 HVR12,aa 705-719。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 aa 146-152 ;HVR2 位于 aa182-187 ;HVR3 位于 aa 263-266,HVR4 位于 aa 383-385,HVR5 位于 aa 451-475 ;HVR6 位于 aa 491-496 ;HVR7 位于 aa 501-505 ;HVR8 位于 aa 513-521 ;HVR9 位于 aa 533-554 ;HVR10 位于 aa 583-596 ;HVR11 位于 aa 660-669 ;HVR12 位于 aa 707-721。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 aa 24-42,aa 25-28 ;aa 81-85 ;aa133-165 ;aa 134-165 ;aa 137-143 ;aa 154-156 ;aa 194-208 ;aa 261-274 ;aa 262-274 ;aa 171-173 ;aa 413-417 ;aa 449-478 ;aa 494-525 ;aa 534-571 ;aa 581-601 ;aa 660-671 ;aa 709-723。其他的优选片段包括 AAV7 上的 aa 1-184,aa 199-259 ;aa 274-446 ;aa 603-659 ;aa 670-706 ;aa 724-736 ;aa 185-198 ;aa 260-273 ;aa 447-477 ;aa 495-602 ;aa 660-669 ;以及 aa 707-723。根据 AAV7[SEQID 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 :aa 185-198 ;aa 260-273 ;aa 447-477 ;aa 495-602 ;aa 660-669 ;以及 aa 707-723。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 [AAV7 SEQ ID NO :3 的 aa 1-623] 及其功能片段,包括 SEQ ID NO :3 的 aa 1-171,aa 172-372,aa 373-444,aa 445-623。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，AAV 44-2(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，AAV 5’ TTR 和 AAV 3’ 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白 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、AAV 5’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[WO 98/10088]、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)] 以及雷怕霉素可诱导系统 [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AAVcap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 WO 98/10088，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 E4 ORF6 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，COS 1，COS 7，BSC 1，BSC 40，BMT 10，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 E4 OR 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 E4 OR 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 E4 ORF6。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 $CaPO_4$ 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)BMPs 1-15 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,Vα-16,Vα-3C,Vα-7,Vα-14,Vα-15,Vα-16,Vα-28和Vα-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WO 00/11140,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(6April 2001),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 AAV 1[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 bp 3095-3121。

[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℃变性 0.5-1min，60-65℃退火 0.5-1min，72℃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 AAV 3’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 nt 4435-4462，SEQID 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'GCTGCGTCAACTGGACCAATGAGAAC 3'，SEQ ID NO :6 的 nt 1398-1423) 和位于 Cap 基因下游另一保守区内的 3' 引物 (AV2cas :5' CGCAGAGACCAAAGTTCAACTGAAACGA 3'，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 Hae I 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 AAV7 VP15' 片段和 1.7kb AAV8 VP13' 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 10^4 AAV8 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.Bioinformatics 14, 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 AAV2 rep 基因和新型 AAV cap 基因的嵌合质粒构建携带 AAV2 ITR 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 AAV2 rep 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 AAV2 ITR 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 pCRAAV76-5+15-4 质粒上切下的。

[0226] 所得到的质粒含有 AAV2P5 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 AdΔF1。从 pAdΔF1 中删除 MLP，L2，L2 和 L3。然后分别删除一个 2.3kb 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50μg DNA(顺式质粒：反式质粒：辅助质粒)转染到一个150mm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%去氧胆酸钠处理(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AAV2 ITR的重组AAV基因组，其中AAV2 rep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AAV2 cap基因，并用一个含AAV8 cap基因的2267bp Spe I/Xho I片段代替[Xiao，W.等，(1999)J Virol 73，3994-4003]。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$CsCl_2$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AV40 poiyA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 Ther 11，2079-91]。

[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 A1AT	4.67 ± 0.75 (n=2)	4.77 (n=1)	4.09 (n=1)	5.04 (n=1)	2.02 (n=1)
CB A1AT	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 Virol 70,8934-43]。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 GFP 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 AAV2 ITR。人 α－抗胰蛋白酶 (A1AT)[Xiao，W. 等，(1999)J Virol 73,3994-4003]、猕猴绒毛膜促性腺激素 (CG) β－亚单位 [Zoltick，P.W.& Wilson，J.M.(2000)Mol Ther 2,657-9]、犬因子 IX[Wang，L. 等，(1997)Proc Natl Acad Sci USA 94,11563-6] 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 Virol 70,8934-43；Zoltick，P.W.& Wilson，J.M.(2000)Mol Ther 2,657-9；Wang，L. 等，Proc Natl Acad Sci USA 94,11563-6]。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 Ther 3，697-707]。

[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 Virol 73,3994-4003]。

[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AAV2/7 lacZ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AAV 2/8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 Ther 2,619-23]。

[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($RUs \times 10^3$)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 1×10^{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
AAV2/1AlbA1AT	0.6 ± 0.36
AAV2AlbA1AT	0.003 ± 0.001
AAV2/5AlbA1AT	0.83 ± 0.64
AAV2/7AlbA1AT	2.2 ± 1.7
AAV2/8AlbA1AT	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(1011 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 Ther 6,1574-83]。

[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 AAV 44.2 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 50 μl 原液,不稀释)。在表 9 中,每只 NCR 裸鼠 50 μl 原液,检测下限为 0.033 μg/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(1×10^{11}GC),如表 10 所示。(每只动物 1×10^{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*

[0284]

	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
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]

血清来源兔免疫所用载体:	10^9 GC *rh. 12* AAV2/42. 1B	10^{10} GC *ch. 5* AAV2/A3	10^{10}GC *rh. 8* AAV2/43. 25	10^9 GC *rh. 10* AAV2/44. 2	10^9 GC *rh. 20* 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

[0288]

rh. 24 AAV2/42. 13	1/5120	ND	ND	ND	1/2560

[0289] 表 12

[0290]

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

[0291] 表 13a. 用 GFP 载体感染 8431 细胞 (Adv 共感染)

[0292]

	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
	128	>200	95	56	13	1
# GFU/组	83	>200	65	54	11	1

[0293] 表 13b. 用 GFP 载体感染 8431 细胞（共感染 Adv）

[0294]

	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
	3	13	54	62	10	3	18
# GFU/组	2	12	71	60	14	2	20
			48	47	16	3	12

[0295] 实施例 10- 家族性高胆固醇血症的鼠模型

[0296] 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。

[0297] A. 载体的构建

[0298] 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. Virol 2001 ;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 $CsCl_2$ 沉淀方法纯化。为了确定病毒的产生量，利用载体 SV40 polyA 区特异性的探针和引物进行 TaqMan(Applied Biosystems) 分析 [Gao GP 等，Hum Gene Ther. 2000 Oct 10 ;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。

[0299] B. 动物

[0300] 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 Therapy 2000 ;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。

[0301] 表14

[0302]

	载体	剂量	n
组1	AAV2/7-TBG-hLDLr	1×10^{12}gc	12
组2	AAV2/7-TBG-hLDLr	3×10^{11}gc	12
组3	AAV2/7-TBG-hLDLr	1×10^{11}gc	12
组4	AAV2/8-TBG-hLDLr	1×10^{12}gc	12
组5	AAV2/8-TBG-hLDLr	3×10^{11}gc	12
组6	AAV2/8-TBG-hLDLr	1×10^{11}gc	12
组7	AAV2/7-TBG-LacZ	1×10^{11}gc	16

[0303]

[0304] C. 血清脂蛋白和肝功能分析

[0305] 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。

[0306] D. 转基因表达的检测

[0307] 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 X 100,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AdhLDLrprep(1×10^{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)。

[0308] E. 免疫组织化学

[0309] 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。

[0310] F.基因转移效率

[0311] 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。

[0312] G.动脉粥样硬化斑块的测量

[0313] 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-960 MD 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 Phase 3 Imaging Systems(Media Cybernetics) 分析。

[0314] H. I^{125}LDL 的清除率

[0315] 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。

[0316] I.肝脂质蓄积的评价

[0317] 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。

[0318] 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)。

[0319] 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)。

[0320] 实施例 - 功能因子 IX 的表达和血友病的矫正

[0321] A.基因敲除小鼠

[0322] 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 AAV2 5’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)-AAV2 3’ITR。如 Wang 等，2000，Molecular Therapy 2:154-158) 所描述利用适当的衣壳构建这种载体。

[0323] 基因敲除小鼠的制备方法见 Wang 等，1997. Proc. Natl. Acad. Sci. USA 94：11563-11566 的描述。这个模型最接近人血友病 B 的表型。

[0324] 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 1×10^{11}GC/ 鼠，AAV8 载体组注射较低的剂量，1×10^{10}GC/ 鼠。对照组注射 1×10^{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。

[0325] 1. ELISA

[0326] 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。

[0327] 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。

[0328] 2. 体外活化部分凝血激酶时间（aPTT）的检测

[0329] 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. USA 94：11563-11566 所描述。

[0330] 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）。

[0331] 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）。

[0332] 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。

[0333] B. 血友病 B 犬

[0334] 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。

[0335] 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 3.7×10^{11} 基因组拷贝 (GC)/kg 的 AAV2/2.cFIX。第二只犬先注射 AAV2/2.cFIX(2.8×10^{11}GC/kg)，然后在第 1180 天第二次注射 AAV2/7.cFIX(2.3×10^{13}GC/kg)。第三只犬一次注射剂量为 4.6×10^{12}GC/kg 的 AAV2/2.cFIX。第四只犬先注射 AAV2/2.cFIX(2.8×10^{12}GC/kg)，在第 995 天再注射 AAV2/7.cFIX(5×10^{12}GC/kg)。

[0336] .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 Therapy 2 :154-158]。

[0337] 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。

[0338] 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。

	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 1AA0

	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

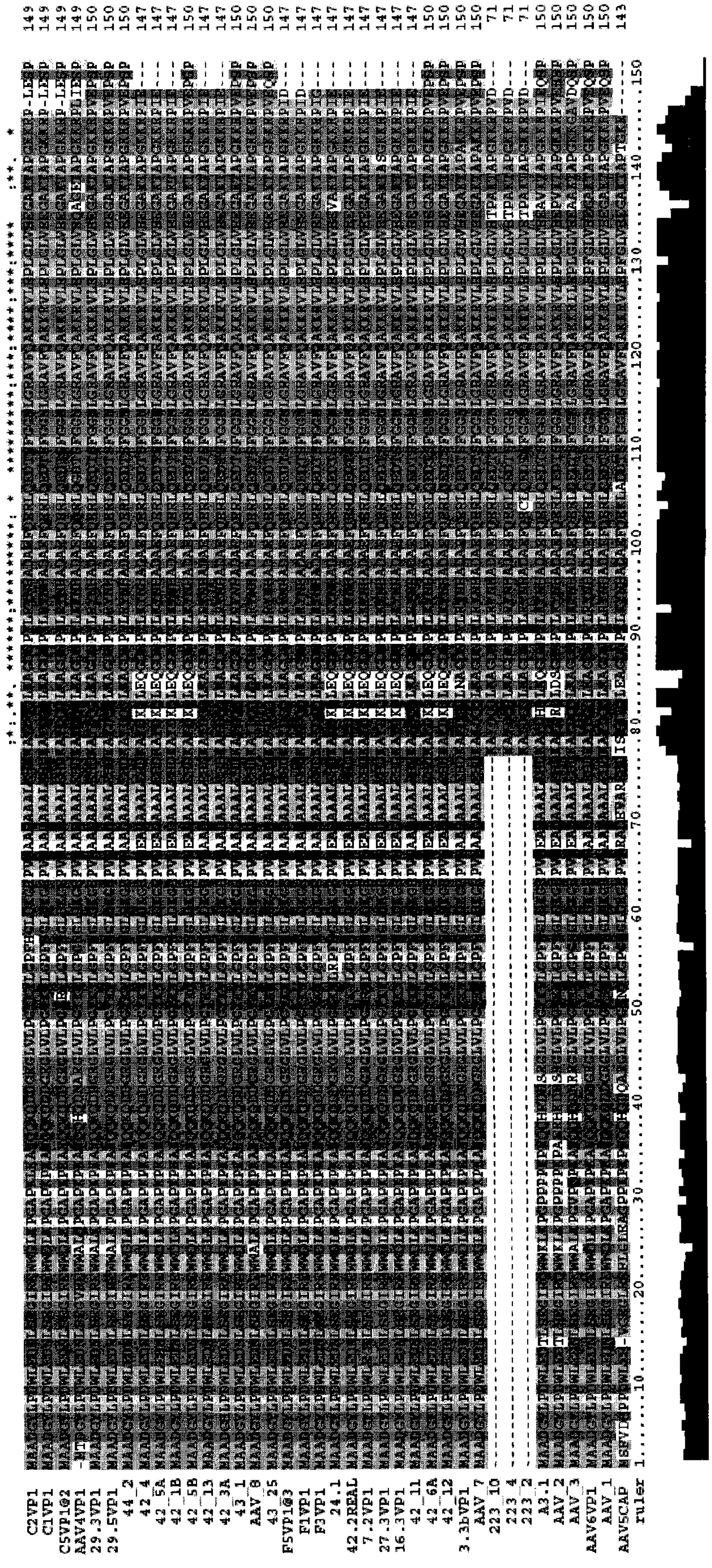

图 2A

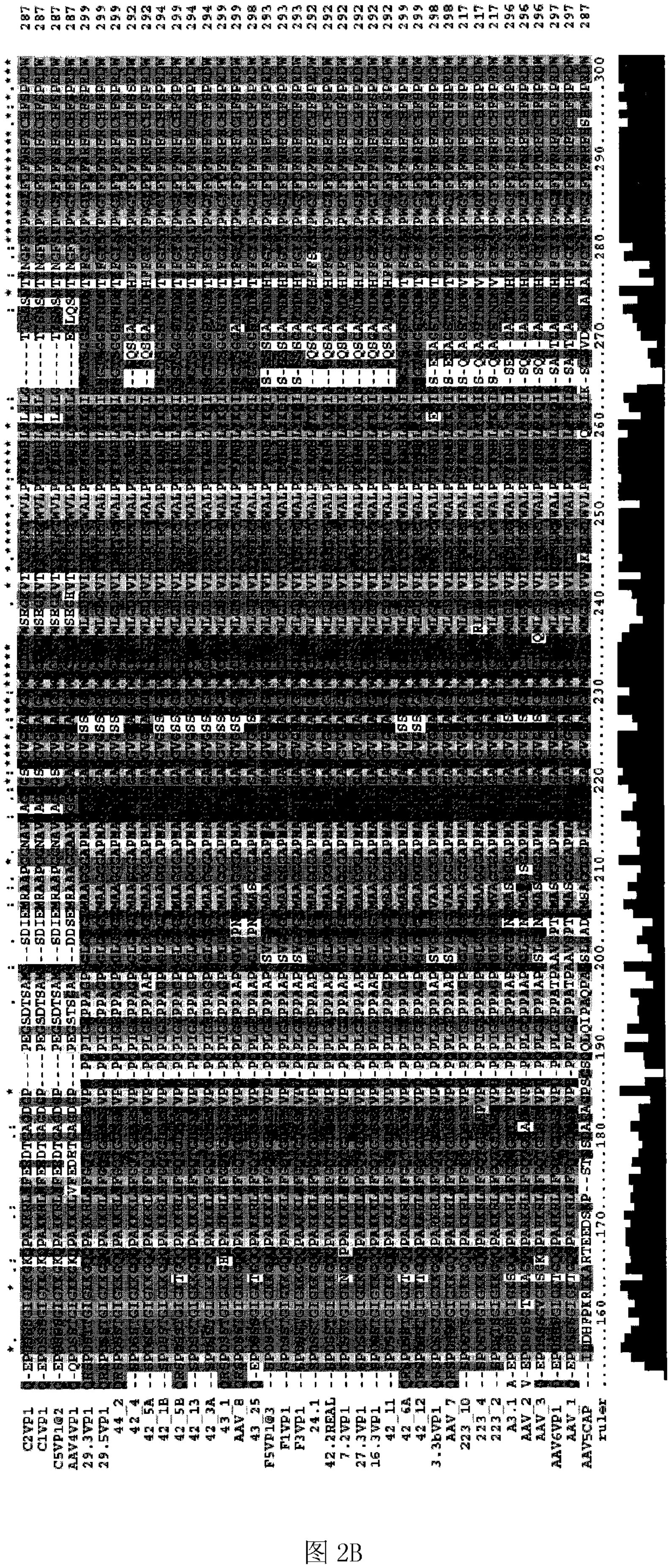

图 2B

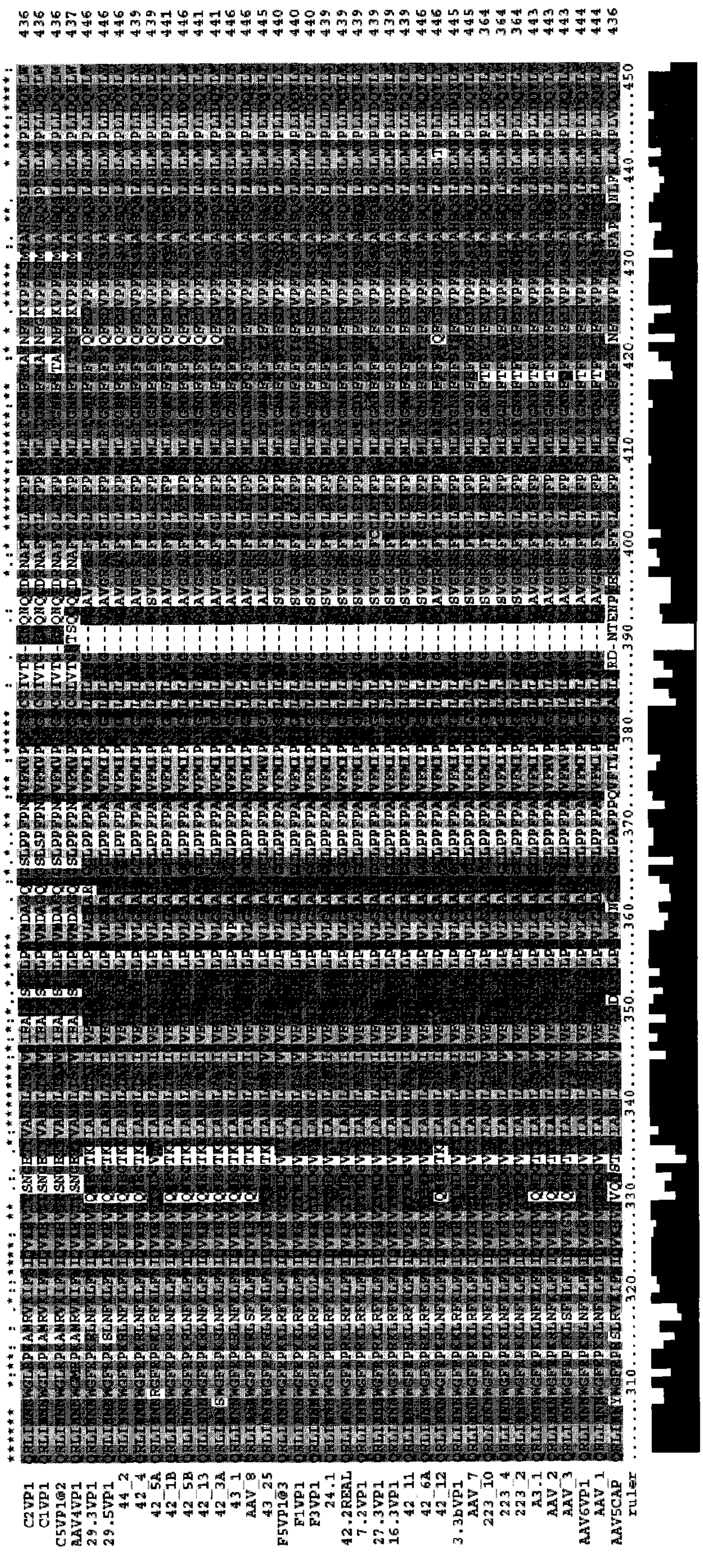
C2VP1
C1VP1
C5VP102
AAV4VP1
29.3VP1
29.5VP1
44_2
42_4
42_5A
42_1B
42_5B
42_13
42_3A
43_1
AAV_8
43_25
F5VP103
F1VP1
F3VP1
24.1
42.2REAL
7.2VP1
27.3VP1
16.3VP1
42_11
42_6A
42_12
3.3bVP1
AAV_7
223_10
223_4
223_2
A3.1
AAV_2
AAV_3
AAV6VP1
AAV_1
AAV5CAP
ruler

图 2C

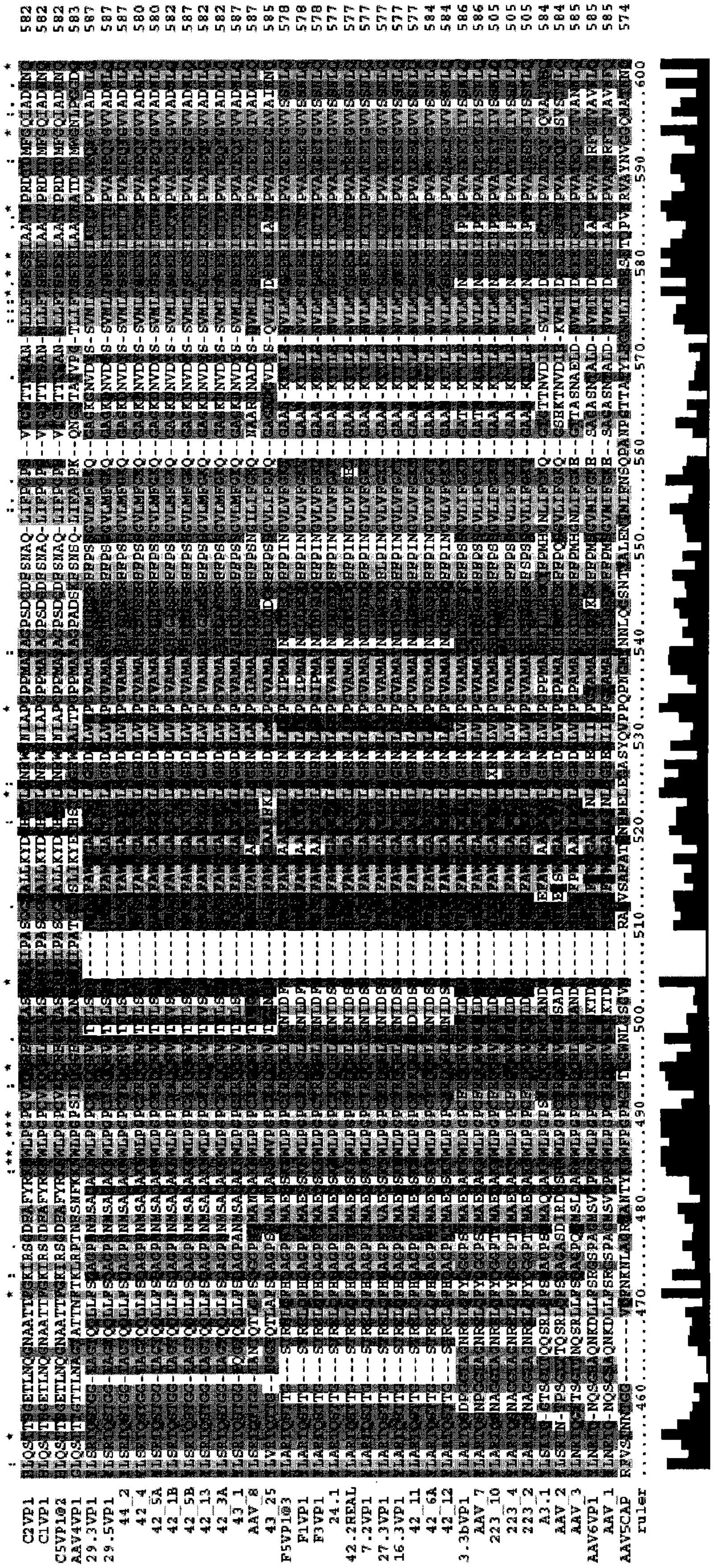

图 2D

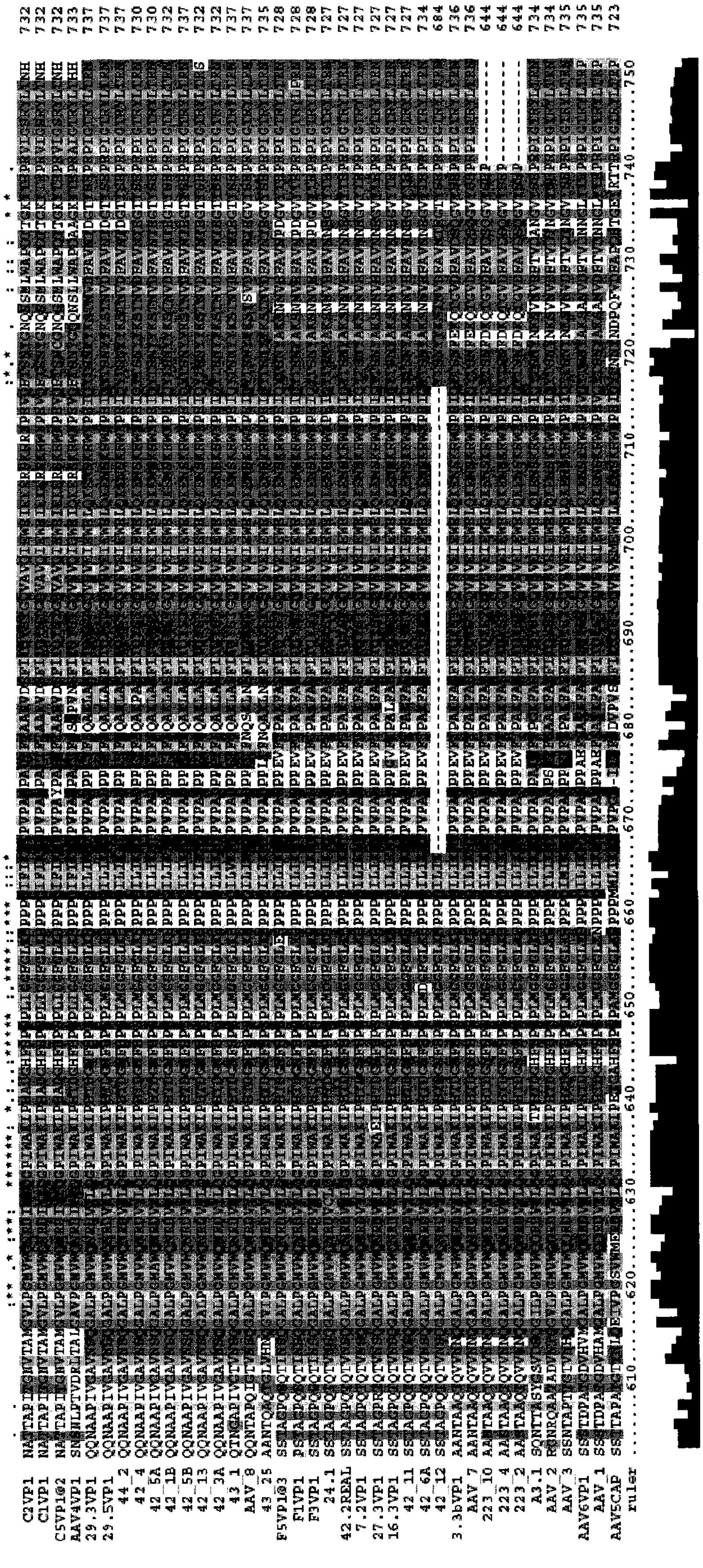

图 2E

C2VP1	L	733
C1VP1	L	733
C5VP1@2	L	733
AAV4VP1	L	734
29.3VP1	L	738
29.5VP1	L	738
44_2	L	738
42_4	L	731
42_5A	L	731
42_1B	L	733
42_5B	L	738
42_13	L	733
42_3A	L	733
43_1	L	738
AAV_8	L	738
43_25	L	736
F5VP1@3	L	729
F1VP1	L	729
F3VP1	L	729
24.1	L	728
42.2REAL	L	728
7.2VP1	L	728
27.3VP1	L	728
16.3VP1	L	728
42_11	L	728
42_6A	L	735
42_12	L	685
3.3bVP1	L	737
AAV_7	L	737
223_10	-	644
223_4	-	644
223_2	-	644
A3.1	L	735
AAV_2	L	735
AAV_3	L	736
AAV6VP1	L	736
AAV_1	L	736
AAV5CAP	L	724
ruler	.	

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