(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 107295802 B
(45)授权公告日 2021.06.29

(21)申请号 201580063926.8
(22)申请日 2015.09.23
(65)同一申请的已公布的文献号
申请公布号 CN 107295802 A
(43)申请公布日 2017.10.24
(30)优先权数据
62/054,899 2014.09.24 US
62/063,587 2014.10.14 US
62/209,862 2015.08.25 US
(85)PCT国际申请进入国家阶段日
2017.05.24
(86)PCT国际申请的申请数据
PCT/US2015/051785 2015.09.23
(87)PCT国际申请的公布数据
WO2016/049230 EN 2016.03.31
(73)专利权人 希望之城
地址 美国加利福尼亚州
(72)发明人 S·查特瑞 L·史密斯 K·王
(74)专利代理机构 北京市金杜律师事务所 11256
代理人 陈文平 侯宝光
(51)Int.Cl.
C12N 7/01(2006.01)
C12N 15/864(2006.01)
C12N 15/90(2006.01)
(56)对比文件
WO 2012051343 A1,2012.04.19
WO 2012051343 A1,2012.04.19
US 2011294218 A1,2011.12.01
Iram F Khan et al..AAV-mediated gene targeting methods for human cells.《Nature Protocols》.2011,
Daniel G Miller et al..Gene targeting in vivo by adeno-associated virus vectors.《NATURE BIOTECHNOLOGY》.2006,
王海鹰等.AAV载体的最新研究进展.《现代生物医学进展》.2008,
审查员 管冰

权利要求书6页 说明书63页
序列表83页 附图70页

(54)发明名称

用于高效基因组编辑的腺相关病毒载体变异体和其方法

(57)摘要

本发明提供用于精确编辑细胞基因组的腺相关病毒(AAV)分化体F载体或AAV载体变异体(相对于AAV9)以及其方法和试剂盒。使用本文所提供的AAV分化体F载体或AAV载体变异体的靶向基因组编辑所发生的出现频率的效率展示为先前所报告的1,000到100,000倍。本发明还提供在个体中治疗疾病或病症的方法,其通过编辑所述个体的细胞的基因组来进行,所述编辑经由用如本文所描述的AAV分化体F载体或AAV载体变异体转导所述细胞和将经过转导的细胞进一步移植到所述个体中来进行,以治疗所述个体的所述疾病或病症。本文还提供在个体中通过体内基因组编辑来治疗疾病或病症的方法,所述体内基因组编辑通过向所述个体直接投与如本文所描述的AAV分化体F载体或AAV载体变异体来进行。

CN 107295802 B

1.一种复制缺陷型腺相关病毒(AAV)在制备用于在不共转导或共投予外源性核酸酶或编码外源性核酸酶的核苷酸的情况下编辑细胞中基因组的目标基因座的药物中的用途,其中所述AVV包含:

a)AAV衣壳,其包含:

(i)VP3衣壳蛋白,其中所述VP3衣壳蛋白的氨基酸序列由SEQ ID NO:2的氨基酸203-736组成;VP2衣壳蛋白,其中所述VP2衣壳蛋白的氨基酸序列由SEQ ID NO:2的氨基酸138-736组成;和VP1衣壳蛋白,其中所述VP1衣壳蛋白的氨基酸序列由SEQ ID NO:2的氨基酸1-736组成;

(ii)VP3衣壳蛋白,其中所述VP3衣壳蛋白的氨基酸序列由SEQ ID NO:3的氨基酸203-736组成;VP2衣壳蛋白,其中所述VP2衣壳蛋白的氨基酸序列由SEQ ID NO:3的氨基酸138-736组成;和VP1衣壳蛋白,其中所述VP1衣壳蛋白的氨基酸序列由SEQ ID NO:3的氨基酸1-736组成;

(iii)VP3衣壳蛋白,其中所述VP3衣壳蛋白的氨基酸序列由SEQ ID NO:4的氨基酸203-736组成;VP2衣壳蛋白,其中所述VP2衣壳蛋白的氨基酸序列由SEQ ID NO:4的氨基酸138-736组成;和VP1衣壳蛋白,其中所述VP1衣壳蛋白的氨基酸序列由SEQ ID NO:4的氨基酸1-736组成;

(iv)VP3衣壳蛋白,其中所述VP3衣壳蛋白的氨基酸序列由SEQ ID NO:5的氨基酸203-736组成;VP2衣壳蛋白,其中所述VP2衣壳蛋白的氨基酸序列由SEQ ID NO:5的氨基酸138-736组成;和VP1衣壳蛋白,其中所述VP1衣壳蛋白的氨基酸序列由SEQ ID NO:5的氨基酸1-736组成;

(v)VP3衣壳蛋白,其中所述VP3衣壳蛋白的氨基酸序列由SEQ ID NO:6的氨基酸203-736组成;VP2衣壳蛋白,其中所述VP2衣壳蛋白的氨基酸序列由SEQ ID NO:6的氨基酸138-736组成;和VP1衣壳蛋白,其中所述VP1衣壳蛋白的氨基酸序列由SEQ ID NO:6的氨基酸1-736组成;

(vi)VP3衣壳蛋白,其中所述VP3衣壳蛋白的氨基酸序列由SEQ ID NO:7的氨基酸203-736组成;VP2衣壳蛋白,其中所述VP2衣壳蛋白的氨基酸序列由SEQ ID NO:7的氨基酸138-736组成;和VP1衣壳蛋白,其中所述VP1衣壳蛋白的氨基酸序列由SEQ ID NO:7的氨基酸1-736组成;

(vii)VP3衣壳蛋白,其中所述VP3衣壳蛋白的氨基酸序列由SEQ ID NO:8的氨基酸203-736组成;VP2衣壳蛋白,其中所述VP2衣壳蛋白的氨基酸序列由SEQ ID NO:8的氨基酸138-736组成;和VP1衣壳蛋白,其中所述VP1衣壳蛋白的氨基酸序列由SEQ ID NO:8的氨基酸1-736组成;

(viii)VP3衣壳蛋白,其中所述VP3衣壳蛋白的氨基酸序列由SEQ ID NO:9的氨基酸203-736组成;VP2衣壳蛋白,其中所述VP2衣壳蛋白的氨基酸序列由SEQ ID NO:9的氨基酸138-736组成;和VP1衣壳蛋白,其中所述VP1衣壳蛋白的氨基酸序列由SEQ ID NO:9的氨基酸1-736组成;

(ix)VP3衣壳蛋白,其中所述VP3衣壳蛋白的氨基酸序列由SEQ ID NO:10的氨基酸203-736组成;VP2衣壳蛋白,其中所述VP2衣壳蛋白的氨基酸序列由SEQ ID NO:10的氨基酸138-736组成;和VP1衣壳蛋白,其中所述VP1衣壳蛋白的氨基酸序列由SEQ ID NO:10的氨基酸1-

736组成；

(x)VP3衣壳蛋白，其中所述VP3衣壳蛋白的氨基酸序列由SEQ ID NO:11的氨基酸203-736组成；VP2衣壳蛋白，其中所述VP2衣壳蛋白的氨基酸序列由SEQ ID NO:11的氨基酸138-736组成；和VP1衣壳蛋白，其中所述VP1衣壳蛋白的氨基酸序列由SEQ ID NO:11的氨基酸1-736组成；

(xi)VP3衣壳蛋白，其中所述VP3衣壳蛋白的氨基酸序列由SEQ ID NO:12的氨基酸203-736组成；VP2衣壳蛋白，其中所述VP2衣壳蛋白的氨基酸序列由SEQ ID NO:12的氨基酸138-736组成；和VP1衣壳蛋白，其中所述VP1衣壳蛋白的氨基酸序列由SEQ ID NO:12的氨基酸1-736组成；

(xii)VP3衣壳蛋白，其中所述VP3衣壳蛋白的氨基酸序列由SEQ ID NO:13的氨基酸203-736组成；VP2衣壳蛋白，其中所述VP2衣壳蛋白的氨基酸序列由SEQ ID NO:13的氨基酸138-736组成；和VP1衣壳蛋白，其中所述VP1衣壳蛋白的氨基酸序列由SEQ ID NO:13的氨基酸1-736组成；

(xiii)VP3衣壳蛋白，其中所述VP3衣壳蛋白的氨基酸序列由SEQ ID NO:14的氨基酸203-736组成；VP2衣壳蛋白，其中所述VP2衣壳蛋白的氨基酸序列由SEQ ID NO:14的氨基酸138-736组成；和VP1衣壳蛋白，其中所述VP1衣壳蛋白的氨基酸序列由SEQ ID NO:14的氨基酸1-736组成；

(xiv)VP3衣壳蛋白，其中所述VP3衣壳蛋白的氨基酸序列由SEQ ID NO:15的氨基酸203-736组成；VP2衣壳蛋白，其中所述VP2衣壳蛋白的氨基酸序列由SEQ ID NO:15的氨基酸138-736组成；和VP1衣壳蛋白，其中所述VP1衣壳蛋白的氨基酸序列由SEQ ID NO:15的氨基酸1-736组成；

(xv)VP3衣壳蛋白，其中所述VP3衣壳蛋白的氨基酸序列由SEQ ID NO:16的氨基酸203-736组成；VP2衣壳蛋白，其中所述VP2衣壳蛋白的氨基酸序列由SEQ ID NO:16的氨基酸138-736组成；和VP1衣壳蛋白，其中所述VP1衣壳蛋白的氨基酸序列由SEQ ID NO:16的氨基酸1-736组成；或

(xvi)VP3衣壳蛋白，其中所述VP3衣壳蛋白的氨基酸序列由SEQ ID NO:17的氨基酸203-736组成；VP2衣壳蛋白，其中所述VP2衣壳蛋白的氨基酸序列由SEQ ID NO:17的氨基酸138-736组成；和VP1衣壳蛋白，其中所述VP1衣壳蛋白的氨基酸序列由SEQ ID NO:17的氨基酸1-736组成；和

b)校正基因组，其包含(i)编辑元件，所述编辑元件包含用于整合到细胞染色体的目标基因座中的核苷酸间键或核苷酸序列，(ii)位于所述编辑元件5'端的5'同源臂核苷酸序列，所述核苷酸序列与所述染色体中相对于所述目标基因座的5'区具有同源性，以及(iii)位于所述编辑元件3'端的3'同源臂核苷酸序列，所述核苷酸序列与所述染色体中相对于所述目标基因座的3'区具有同源性。

2.根据权利要求1所述的用途，其中所述细胞在哺乳动物个体中，并且向所述个体施用有效量的AAV以在所述个体中转导所述细胞。

3.根据权利要求1所述的用途，其中所述校正基因组不存在可操作地连接于所述编辑元件核苷酸序列的启动子。

4.根据权利要求1所述的用途，其中所述校正基因组进一步包含可操作地连接于所述

编辑元件的外源性启动子。

5.根据权利要求1-4任一项所述的用途，其中每个5'和3'同源臂核苷酸序列独立地具有50到2000个核苷酸的核苷酸长度。

6.根据权利要求1-4任一项所述的用途，其中所述编辑元件包含基因或其片段、基因的编码序列、外显子、内含子、5'非翻译区(UTR)、3'UTR、启动子、剪接供体、剪接受体、对应于非编码RNA的序列、绝缘子(insulator)、或其组合。

7.根据权利要求1-4任一项所述的用途，其中所述编辑元件是包含相对于所述染色体的目标基因座插入、缺失或取代的核苷酸序列。

8.根据权利要求1-4任一项的用途，其中所述目标基因座相对于相对应的野生型染色体中的基因座，包含一个或多个突变，任选地其中所述目标基因座包含无效(amorphic)突变、新效(neomorphic)突变、反效(antimorphic)突变、常染色体显性突变、常染色体隐性突变、或其组合，以及任选地其中所述目标基因座选自染色体内的启动子、强化子、信号序列、内含子、外显子、剪接供体位点、剪接受体位点、内部核糖体进入位点、反向外显子、绝缘子、基因或其片段、染色体反转和染色体易位。

9.根据权利要求1-4任一项所述的用途，其中所述校正基因组包含位于所述5'同源臂核苷酸序列5'端的5'反向末端重复(5'ITR)核苷酸序列和位于所述3'同源臂核苷酸序列3'端的3'反向末端重复(3'ITR)核苷酸序列，并且任选地其中：

(a)所述5'ITR核苷酸序列与SEQ ID NO:36具有至少95％序列一致性，并且所述3'ITR核苷酸序列与SEQ ID NO:37具有至少95％序列一致性；或

(b)所述5'ITR核苷酸序列与SEQ ID NO:38具有至少95％序列一致性，并且所述3'ITR核苷酸序列与SEQ ID NO:39具有至少95％序列一致性。

10.根据权利要求1-4任一项所述的用途，其中所述细胞是哺乳动物细胞，任选地其中所述细胞是成肌细胞、内皮细胞、肝脏细胞、成纤维细胞、乳房细胞、淋巴细胞或视网膜细胞。

11.根据权利要求1-4任一项所述的用途，其中所述细胞是干细胞，任选地其中所述干细胞是造血干细胞、脐血干细胞、骨髓干细胞、胎儿肝脏干细胞或周边血液干细胞，并且任选地其中所述干细胞是$CD34^+$干细胞。

12.一种复制缺陷型腺相关病毒(AAV)，其包含：

a)AAV衣壳，其包含：

(i)VP3衣壳蛋白，其中所述VP3衣壳蛋白的氨基酸序列由SEQ ID NO:2的氨基酸203-736组成；VP2衣壳蛋白，其中所述VP2衣壳蛋白的氨基酸序列由SEQ ID NO:2的氨基酸138-736组成；和VP1衣壳蛋白，其中所述VP1衣壳蛋白的氨基酸序列由SEQ ID NO:2的氨基酸1-736组成；

(ii)VP3衣壳蛋白，其中所述VP3衣壳蛋白的氨基酸序列由SEQ ID NO:3的氨基酸203-736组成；VP2衣壳蛋白，其中所述VP2衣壳蛋白的氨基酸序列由SEQ ID NO:3的氨基酸138-736组成；和VP1衣壳蛋白，其中所述VP1衣壳蛋白的氨基酸序列由SEQ ID NO:3的氨基酸1-736组成；

(iii)VP3衣壳蛋白，其中所述VP3衣壳蛋白的氨基酸序列由SEQ ID NO:4的氨基酸203-736组成；VP2衣壳蛋白，其中所述VP2衣壳蛋白的氨基酸序列由SEQ ID NO:4的氨基酸138-

736组成;和VP1衣壳蛋白,其中所述VP1衣壳蛋白的氨基酸序列由SEQ ID NO:4的氨基酸1-736组成;

(iv)VP3衣壳蛋白,其中所述VP3衣壳蛋白的氨基酸序列由SEQ ID NO:5的氨基酸203-736组成;VP2衣壳蛋白,其中所述VP2衣壳蛋白的氨基酸序列由SEQ ID NO:5的氨基酸138-736组成;和VP1衣壳蛋白,其中所述VP1衣壳蛋白的氨基酸序列由SEQ ID NO:5的氨基酸1-736组成;

(v)VP3衣壳蛋白,其中所述VP3衣壳蛋白的氨基酸序列由SEQ ID NO:6的氨基酸203-736组成;VP2衣壳蛋白,其中所述VP2衣壳蛋白的氨基酸序列由SEQ ID NO:6的氨基酸138-736组成;和VP1衣壳蛋白,其中所述VP1衣壳蛋白的氨基酸序列由SEQ ID NO:6的氨基酸1-736组成;

(vi)VP3衣壳蛋白,其中所述VP3衣壳蛋白的氨基酸序列由SEQ ID NO:7的氨基酸203-736组成;VP2衣壳蛋白,其中所述VP2衣壳蛋白的氨基酸序列由SEQ ID NO:7的氨基酸138-736组成;和VP1衣壳蛋白,其中所述VP1衣壳蛋白的氨基酸序列由SEQ ID NO:7的氨基酸1-736组成;

(vii)VP3衣壳蛋白,其中所述VP3衣壳蛋白的氨基酸序列由SEQ ID NO:8的氨基酸203-736组成;VP2衣壳蛋白,其中所述VP2衣壳蛋白的氨基酸序列由SEQ ID NO:8的氨基酸138-736组成;和VP1衣壳蛋白,其中所述VP1衣壳蛋白的氨基酸序列由SEQ ID NO:8的氨基酸1-736组成;

(viii)VP3衣壳蛋白,其中所述VP3衣壳蛋白的氨基酸序列由SEQ ID NO:9的氨基酸203-736组成;VP2衣壳蛋白,其中所述VP2衣壳蛋白的氨基酸序列由SEQ ID NO:9的氨基酸138-736组成;和VP1衣壳蛋白,其中所述VP1衣壳蛋白的氨基酸序列由SEQ ID NO:9的氨基酸1-736组成;

(ix)VP3衣壳蛋白,其中所述VP3衣壳蛋白的氨基酸序列由SEQ ID NO:10的氨基酸203-736组成;VP2衣壳蛋白,其中所述VP2衣壳蛋白的氨基酸序列由SEQ ID NO:10的氨基酸138-736组成;和VP1衣壳蛋白,其中所述VP1衣壳蛋白的氨基酸序列由SEQ ID NO:10的氨基酸1-736组成;

(x)VP3衣壳蛋白,其中所述VP3衣壳蛋白的氨基酸序列由SEQ ID NO:11的氨基酸203-736组成;VP2衣壳蛋白,其中所述VP2衣壳蛋白的氨基酸序列由SEQ ID NO:11的氨基酸138-736组成;和VP1衣壳蛋白,其中所述VP1衣壳蛋白的氨基酸序列由SEQ ID NO:11的氨基酸1-736组成;

(xi)VP3衣壳蛋白,其中所述VP3衣壳蛋白的氨基酸序列由SEQ ID NO:12的氨基酸203-736组成;VP2衣壳蛋白,其中所述VP2衣壳蛋白的氨基酸序列由SEQ ID NO:12的氨基酸138-736组成;和VP1衣壳蛋白,其中所述VP1衣壳蛋白的氨基酸序列由SEQ ID NO:12的氨基酸1-736组成;

(xii)VP3衣壳蛋白,其中所述VP3衣壳蛋白的氨基酸序列由SEQ ID NO:13的氨基酸203-736组成;VP2衣壳蛋白,其中所述VP2衣壳蛋白的氨基酸序列由SEQ ID NO:13的氨基酸138-736组成;和VP1衣壳蛋白,其中所述VP1衣壳蛋白的氨基酸序列由SEQ ID NO:13的氨基酸1-736组成;

(xiii)VP3衣壳蛋白,其中所述VP3衣壳蛋白的氨基酸序列由SEQ ID NO:14的氨基酸

203-736组成;VP2衣壳蛋白,其中所述VP2衣壳蛋白的氨基酸序列由SEQ ID NO:14的氨基酸138-736组成;和VP1衣壳蛋白,其中所述VP1衣壳蛋白的氨基酸序列由SEQ ID NO:14的氨基酸1-736组成;

(xiv)VP3衣壳蛋白,其中所述VP3衣壳蛋白的氨基酸序列由SEQ ID NO:15的氨基酸203-736组成;VP2衣壳蛋白,其中所述VP2衣壳蛋白的氨基酸序列由SEQ ID NO:15的氨基酸138-736组成;和VP1衣壳蛋白,其中所述VP1衣壳蛋白的氨基酸序列由SEQ ID NO:15的氨基酸1-736组成;

(xv)VP3衣壳蛋白,其中所述VP3衣壳蛋白的氨基酸序列由SEQ ID NO:16的氨基酸203-736组成;VP2衣壳蛋白,其中所述VP2衣壳蛋白的氨基酸序列由SEQ ID NO:16的氨基酸138-736组成;和VP1衣壳蛋白,其中所述VP1衣壳蛋白的氨基酸序列由SEQ ID NO:16的氨基酸1-736组成;或

(xvi)VP3衣壳蛋白,其中所述VP3衣壳蛋白的氨基酸序列由SEQ ID NO:17的氨基酸203-736组成;VP2衣壳蛋白,其中所述VP2衣壳蛋白的氨基酸序列由SEQ ID NO:17的氨基酸138-736组成;和VP1衣壳蛋白,其中所述VP1衣壳蛋白的氨基酸序列由SEQ ID NO:17的氨基酸1-736组成;和

b)校正基因组,其包含(i)编辑元件,所述编辑元件包含用于整合到细胞染色体的目标基因座中的核苷酸间键或核苷酸序列,(ii)位于所述编辑元件5'端的5'同源臂核苷酸序列,所述核苷酸序列与所述染色体中相对于所述目标基因座的5'区具有同源性,以及(iii)位于所述编辑元件3'端的3'同源臂核苷酸序列,所述核苷酸序列与所述染色体中相对于所述目标基因座的3'区具有同源性,

其中所述AAV在不共转导或共投予外源性核酸酶或编码外源性核酸酶的核苷酸的情况下将所述编辑元件整合到所述细胞中所述染色体的所述目标基因座中。

13.根据权利要求12所述的AAV,其中所述校正基因组不存在可操作地连接于所述编辑元件核苷酸序列的启动子。

14.根据权利要求12所述的AAV,其中所述校正基因组进一步包含可操作地连接于所述编辑元件的外源性启动子。

15.根据权利要求12-14任一项所述的AAV,其中每个5'和3'同源臂核苷酸序列独立地具有50到2000个核苷酸的核苷酸长度。

16.根据权利要求12-14任一项所述的AAV,其中所述编辑元件包含基因或其片段、基因的编码序列、外显子、内含子、5'非翻译区(UTR)、3'UTR、启动子、剪接供体、剪接受体、对应于非编码RNA的序列、绝缘子(insulator)、或其组合。

17.根据权利要求12-14任一项所述的AAV,其中所述编辑元件是包含相对于所述染色体的目标基因座插入、缺失或取代的核苷酸序列。

18.根据权利要求12-14任一项所述的AAV,其中所述目标基因座相对于相对应的野生型染色体中的基因座,包含一个或多个突变,任选地其中所述目标基因座包含无效(amorphic)突变、新效(neomorphic)突变、反效(antimorphic)突变、常染色体显性突变、常染色体隐性突变、或其组合,以及任选地其中所述目标基因座选自染色体内的启动子、强化子、信号序列、内含子、外显子、剪接供体位点、剪接受体位点、内部核糖体进入位点、反向外显子、绝缘子、基因或其片段、染色体反转和染色体易位。

19.根据权利要求12-14任一项所述的AAV,其中所述校正基因组包含位于所述5'同源臂核苷酸序列5'端的5'反向末端重复(5'ITR)核苷酸序列和位于所述3'同源臂核苷酸序列3'端的3'反向末端重复(3'ITR)核苷酸序列,并且任选地其中:

(a)所述5'ITR核苷酸序列与SEQ ID NO:36具有至少95%序列一致性,并且所述3'ITR核苷酸序列与SEQ ID NO:37具有至少95%序列一致性;或

(b)所述5'ITR核苷酸序列与SEQ ID NO:38具有至少95%序列一致性,并且所述3'ITR核苷酸序列与SEQ ID NO:39具有至少95%序列一致性。

20.根据权利要求12-14任一项所述的AAV,其中所述细胞是哺乳动物细胞,任选地其中所述细胞是成肌细胞、内皮细胞、肝脏细胞、成纤维细胞、乳房细胞、淋巴细胞或视网膜细胞。

21.根据权利要求12-14任一项所述的AAV,其中所述细胞是干细胞,任选地其中所述干细胞是造血干细胞、脐血干细胞、骨髓干细胞、胎儿肝脏干细胞或周边血液干细胞,并且任选地其中所述干细胞是$CD34^+$干细胞。

22.一种用于重组制备腺相关病毒(AAV)的包装系统,其中所述包装系统包含

编码一个或多个AAV Rep蛋白质的Rep核苷酸序列;

编码一个或多个根据权利要求12-14中任一项所述的衣壳蛋白的核苷酸序列;以及

根据权利要求12到14中任一项所述的校正基因组;

其中所述包装系统在细胞中可操作地用于将所述校正基因组包封于所述衣壳中以形成所述腺相关病毒。

23.一种用于重组制备AAV的方法,其中所述方法包含在可操作地用于将校正基因组包封于衣壳中的条件下将根据权利要求22所述的包装系统引入细胞以形成AAV。

用于高效基因组编辑的腺相关病毒载体变异体和其方法

[0001] 相关申请的交叉参考

[0002] 本申请要求2015年8月25日提交的美国临时申请第62/209,862号、2014年10月14日提交的美国临时申请第62/063,587号和2014年9月24日提交的美国临时申请第62/054,899号的优先权日的权益。这些所提及的申请中的每一个的内容都以全文引入的方式并入本文中。

[0003] 政府利益的陈述

[0004] 本发明在政府支持下在由美国国家卫生研究院(National Institutes of Health)授予的授权号HL087285下进行。政府在本发明中享有一定权利。

背景技术

[0005] 腺相关病毒(AAV)基因组由约4.9千碱基长的正义或负义的单链脱氧核糖核酸(ssDNA)构建。基因组在DNA链的两端包含反向末端重复(ITR),以及两个开放阅读框(ORF):rep和cap。Rep由编码AAV生命周期所需rep蛋白质的四个重叠基因组成,并且cap含有衣壳蛋白的重叠核苷酸序列:VP1、VP2和VP3,其在一起相互作用以形成具有二十面体对称性的衣壳。

[0006] 来源于复制缺陷型人类细小病毒AAV2的重组腺相关病毒(rAAV)载体验证为安全并且有效的基因转移媒剂,其尚未决定性地鉴别为具有病原性或致癌性[3-4,6,18-19,26,31]。rAAV转导非分裂原代细胞,免疫原性较低,并且直接支持体内转基因表达[6,10,20]。野生型AAV的感染与致癌性转型相关联,并且AAV反向末端重复可以实际上赋予致癌保护[2,28,52-55]。对各组人类组织的最近调查发现,骨髓和肝脏是天然存在的AAV分离株在人类中的两个最常见位点,表明AAV对骨髓细胞的感染并非罕见。

[0007] 已经长期考虑使用病毒载体用于基因疗法。由于造血系统用于长期校正和易于离体操控的潜能,其是基因疗法最早目标中的一个。然而,尽管进行了大量努力,但实际治疗性成功仍难以实现[5]。这是由于公认大部分病毒载体不可能高效地转导静止非分裂造血干细胞(HSC)[23]以及由插入型瘤形成产生安全性问题[15,22]。然而,已经成功地展现通过rAAV向鼠类和人类HSC两者的稳定基因转移[8,11-12,24,27,29-30,37]。

[0008] 另外也难以有效地在用于治疗神经病状,尤其中枢神经系统疾病或病症的基因疗法中使用病毒载体,这是因为难以穿过血-脑屏障,即通过限制溶质通过的内皮细胞之间的紧密接合点而产生循环血液与大脑细胞外流体的细胞和代谢隔离。

[0009] CD34是细胞表面醣蛋白和细胞-细胞粘附因子。CD34蛋白质在早期造血和血管组织中表达,并且将表达CD34的细胞命名为$CD34^+$。已经在免疫缺陷型NOD-SCID小鼠中展现rAAV在人类$CD34^+$HSC中的染色体整合[8,12,16,29]和能够支持原代和二代多谱系植入的原始、多能、自我更新的人类HSC的高效转导[29]。展示能够支持连续植入的原始HSC的转导可归因于rAAV高效地转导驻留在G0中的原始静止CD34+CD38-细胞的倾向[24]。尽管存在成功进行rAAV介导的基因转移体外进入人类HSC中和体内进入类和非人类灵长类动物HSC中的若干报告,但关于rAAV用于HSC转导的效用的争论仍持续。这些不符主要由通过表达谱评

定转导的短期体外研究产生，并且可归因与已经鉴别的对由rAAV2进行转基因表达的限制，包括病毒脱壳[35]、细胞内运输(intracellular trafficking)[33]、核转运(nuclear transport)和第二链合成[36]。

[0010] 虽然AAV2仍是用于AAV类载体的研究最多的原型病毒[1,13,18,21]，但大量新AAV血清型的鉴别显著地增加潜在基因转移载体的谱系[14]。AAV1、3和4以腺病毒储备液的污染物形式分离，并且AAV5从人类湿疣(condylomatous wart)中分离。AAV6以AAV1与AAV2之间的实验室重组形式产生。最近，在人类和非人类灵长类动物组织中鉴别出天然存在的AAV的超过100种不同分离株。这导致使用来源于一些这些分离株的衣壳用于假型化，用新颖包膜替换AAV2的包膜蛋白，从而然后使用AAV2rep和新颖衣壳基因来包装rAAV2基因组。使用作为病毒壳层一部分的蛋白质的新颖衣壳在与AAV2相关联的转基因表达中引起对许多限制的规避[32,35-36]。

[0011] 在努力规避这些限制的情况下，最近研究已经展示，新颖衣壳序列引起蛋白酶体介导的衣壳降解减少，核运输和滞留增加。其中许多利用新颖受体的新颖衣壳扩大rAAV的向性，允许高效转导先前难治性组织，并且提供规避高度普遍预先存在的针对AAV2的血清免疫性的手段。值得注意地，一些新颖衣壳似乎改变rAAV的细胞内处理。举例来说，与原生AAV2衣壳相比，在AAV8的情形下，脱壳和转基因表达加速。最近，展示不论反向末端重复(ITR)的血清型来源如何，转基因表达都基于衣壳蛋白。

[0012] 在细胞激素预致敏的周边血液干细胞中可鉴别出天然存在的AAV。这些AAV的衣壳序列是独特的。这些衣壳能够对重组AAV2基因组进行假型化。美国专利公开第20130096182A1号描述衣壳AAVF1-17以及其用于细胞转导和基因转移的用途。在基因疗法领域中关于永久和可逆基因转移与表达的出于治疗性目的的任何改良都将是所属领域中的显著改进。此外，尽管进行了数十年的研究，但向干细胞中的安全并且高效的基因递送仍是所属领域中的明显挑战。因此，安全地对干细胞进行基因修饰的能力将表示显著进步。

[0013] 此外，通过在基因组中特定位点处进行基因靶向或校正而进行基因组编辑但不在基因组中留下足迹对于遗传性和获得性疾病的精确校正有吸引力。目前的技术通过使用外源性核酸内切酶来实现这种编辑，所述外源性核酸酶如锌指核酸酶、TAL核酸内切酶或卡斯蛋白酶9/CRISPR系统。然而，这些“传统”方法与核酸内切酶裂解的毒性和脱靶效果相关联。因此，在不需要外源性核酸内切酶裂解的情况下以高频率安全并且高效地对干细胞进行基因修饰的能力将表示显著进步。

[0014] 此外，目前对人类HSC进行基因转导的方法涉及离体转导经过纯化的供体干细胞，接着移植到通常“经过适应的”接受者中。细胞采集程序是创伤性的，并且涉及骨髓采集或对供体进行多日粒细胞菌落刺激因子(G-CSF)预致敏接着进行血浆分离置换(apheresis)。离体转导程序可能影响干细胞的造血潜能。此外，在移植之前必须测试经过体外转导的细胞的无菌性、毒性等。在移植到接受者中之前，干细胞常常必须经历使用化学疗法或辐射的适应以确保植入。所述过程通常需要患者住院至少若干天并且有时更长。在整体上，这是极大地限制干细胞基因疗法效用的艰难、昂贵并且高风险的程序。需要为目前的干细胞转导方法提供更好的替代方案而不需要纯化和离体转导的程序。

发明内容

[0015] 本文提供用于经由同源重组来编辑细胞基因组的腺相关病毒(AAV)载体(例如,分化体(Clade)F载体,如包含包封于衣壳中的校正基因组的复制缺陷型腺相关病毒(AAV),所述衣壳是AAV分化体F衣壳),以及其使用方法和试剂盒。

[0016] 在一些方面,本发明提供一种复制缺陷型腺相关病毒(AAV),其包含包封于衣壳中的校正基因组,所述衣壳是AAV分化体F衣壳;并且所述校正基因组包含(a)编辑元件,所述编辑元件选自用于整合到哺乳动物染色体的目标基因座中的核苷酸间键或核苷酸序列,(b)位于所述编辑元件5'端的5'同源臂核苷酸序列,所述核苷酸序列与所述哺乳动物染色体中相对于所述目标基因座的5'区具有同源性,以及(c)位于所述编辑元件3'端的3'同源臂核苷酸序列,所述核苷酸序列与所述哺乳动物染色体中相对于所述目标基因座的3'区具有同源性。在一些方面,本发明还提供复制缺陷型腺相关病毒(AAV),其包含包封于衣壳中的校正基因组,所述衣壳是AAV分化体F衣壳;并且所述校正基因组包含用于整合到哺乳动物染色体的目标基因座中的编辑元件核苷酸序列,所述校正基因组基本不存在可操作地连接于所述编辑元件核苷酸序列的启动子。在其它方面,本发明提供一种复制缺陷型腺相关病毒(AAV),其包含包封于衣壳中的校正基因组,所述衣壳是AAV分化体F衣壳;所述校正基因组包含编辑元件,所述编辑元件选自用于整合到细胞中哺乳动物染色体的目标基因座中的核苷酸间键或核苷酸序列;并且所述AAV用于将所述编辑元件整合到所述细胞中所述哺乳动物染色体的所述目标基因座中的染色体整合效率为至少约1%。本发明的其它方面涉及一种包含复制缺陷型腺相关病毒(AAV)的基因编辑载体,所述AAV包含包封于AAV衣壳中的校正基因组,所述校正基因组包含编辑元件,所述编辑元件选自用于整合到哺乳动物细胞染色体的目标基因座中的核苷酸间键或核苷酸序列;位于所述编辑元件5'端的5'同源臂核苷酸序列,所述核苷酸序列与所述染色体中相对于所述目标基因座的5'区具有同源性;位于所述编辑元件3'端的3'同源臂核苷酸序列,所述核苷酸序列与所述染色体中相对于所述目标基因座的3'区具有同源性;并且其中所述AAV用于在无外源性核酸酶存在下将所述编辑元件整合到所述哺乳动物细胞染色体的所述目标基因座中的染色体整合效率为至少约10%。

[0017] 在本文所提供的AAV中的任一个的一些实施例中,AAV在无外源性核酸酶存在下用于将编辑元件整合到细胞中哺乳动物染色体的目标基因座中的染色体整合效率为至少约1%。

[0018] 在本文所提供的AAV中的任一个的一些实施例中,校正基因组包含位于5'同源臂核苷酸序列5'端的5'反向末端重复(5'ITR)核苷酸序列和位于3'同源臂核苷酸序列3'端的3'反向末端重复(3'ITR)核苷酸序列。在一些实施例中,5'ITR核苷酸序列和3'ITR核苷酸序列分别与AAV2病毒5'ITR和AAV2病毒3'ITR基本上一致。在一些实施例中,5'ITR核苷酸序列和3'ITR核苷酸序列基本上是彼此的镜像。在一些实施例中,5'ITR核苷酸序列与SEQ ID NO:36具有至少95%序列一致性,并且3'ITR核苷酸序列与SEQ ID NO:37具有至少95%序列一致性。在一些实施例中,5'ITR核苷酸序列和3'ITR核苷酸序列分别与AAV5病毒5'ITR和AAV5病毒3'ITR基本上一致。在一些实施例中,5'ITR核苷酸序列和3'ITR核苷酸序列基本上是彼此的镜像。在一些实施例中,5'ITR核苷酸序列与SEQ ID NO:38具有至少95%序列一致性,并且3'ITR核苷酸序列与SEQ ID NO:39具有至少95%序列一致性。

[0019] 在本文所提供的AAV中的任一个的一些实施例中，校正基因组基本不存在可操作地连接于编辑元件核苷酸序列的启动子。在本文所提供的AAV中的任一个的一些实施例中，校正基因组进一步包含可操作地连接于编辑元件的外源性启动子。

[0020] 在本文所提供的AAV中的任一个的一些实施例中，复制缺陷型AAV基因组包含基本不存在AAV rep基因和AAV cap基因。

[0021] 在本文所提供的AAV中的任一个的一些实施例中，5'和3'同源臂核苷酸序列中的每一个的核苷酸长度都独立地介于约500到1000个核苷酸之间或介于约600到1000个核苷酸之间。在一些实施例中，5'和3'同源臂核苷酸序列具有基本上相等的核苷酸长度。在一些实施例中，5'和3'同源臂核苷酸序列具有不对称的核苷酸长度。在一些实施例中，5'同源臂核苷酸序列与哺乳动物染色体中相对于目标基因座的5'区具有至少约95％核苷酸序列一致性。在一些实施例中，3'同源臂核苷酸序列与哺乳动物染色体中相对于目标基因座的3'区具有至少约95％核苷酸序列一致性。在一些实施例中，5'同源臂核苷酸序列与哺乳动物染色体中相对于目标基因座的5'区具有100％一致性，并且3'同源臂核苷酸序列与哺乳动物染色体中相对于目标基因座的3'区具有100％一致性。

[0022] 在本文所提供的AAV中的任一个的一些实施例中，编辑元件由一个核苷酸组成。在一些实施例中，目标基因座是由一个核苷酸组成的核苷酸序列，并且所述目标基因座表示哺乳动物染色体的点突变。

[0023] 在一些实施例中，目标基因座可以包含哺乳动物染色体的内含子。在一些实施例中，目标基因座可以包含哺乳动物染色体的外显子。在一些实施例中，目标基因座可以包含哺乳动物染色体的非编码区。在一些实施例中，目标基因座可以包含哺乳动物染色体的调节区。在一些实施例中，目标基因座可以是与如本文所描述的疾病病况相关联的基因座。

[0024] 在本文所提供的AAV中的任一个的一些实施例中，编辑元件包含至少1、2、10、100、200、500、1000、1500、2000、3000、4000或5000个核苷酸。在一些实施例中，编辑元件包含1到5500、1到5000、1到4500、1到4000、1到3000、1到2000、1到1000、1到500、1到200或1到100个核苷酸，或2到5500、2到5000、2到4500、2到4000、2到3000、2到2000、2到1000、2到500、2到200或2到100个核苷酸，或10到5500、10到5000、10到4500、10到4000、10到3000、10到2000、10到1000、10到500、10到200或10到100个核苷酸。

[0025] 在本文所提供的AAV中的任一个的一些实施例中，编辑元件包含外显子、内含子、5'非翻译区(UTR)、3'UTR、启动子、剪接供体、剪接受体、编码或非编码RNA的序列、绝缘子(insulator)、基因或其组合。在本文所提供的AAV中的任一个的一些实施例中，编辑元件是目标基因座内或跨目标基因座的基因的编码序列的片段。

[0026] 在本文所提供的AAV中的任一个的一些实施例中，目标基因座是包含n个核苷酸的核苷酸序列，其中n是大于或等于1的整数；编辑元件包含m个核苷酸，其中m是等于n的整数；并且所述编辑元件表示对哺乳动物染色体目标基因座的取代。在本文所提供的AAV中的任一个的一些实施例中，目标基因座是包含n个核苷酸的核苷酸序列，其中n是大于或等于1的整数；编辑元件包含m个核苷酸，其中m是大于n的整数；并且所述编辑元件表示对哺乳动物染色体目标基因座的取代性添加。在本文所提供的AAV中的任一个的一些实施例中，目标基因座是包含n个核苷酸的核苷酸序列，其中n是大于或等于2的整数；编辑元件包含m个核苷酸，其中m是小于n的整数；并且所述编辑元件表示对哺乳动物染色体目标基因座的取代性

缺失。在本文所提供的AAV中的任一个的一些实施例中，目标基因座是核苷酸间键；编辑元件包含m个核苷酸，其中m是大于或等于1的整数；并且所述编辑元件表示对哺乳动物染色体目标基因座的添加。

[0027] 在本文所提供的AAV中的任一个的一些实施例中，编辑元件是核苷酸间键。在一些实施例中，目标基因座是包含一个或多个核苷酸的核苷酸序列，并且编辑元件包含对哺乳动物染色体目标基因座的缺失。

[0028] 在本文所提供的AAV中的任一个的一些实施例中，哺乳动物染色体的目标基因座是相对于相对应的野生型哺乳动物染色体，包含一个或多个突变型核苷酸的突变型目标基因座。在一些实施例中，突变型目标基因座包含点突变、错义突变、无义突变、一个或多个核苷酸的插入、一个或多个核苷酸的缺失，或其组合。在一些实施例中，突变型目标基因座包含无效(amorphic)突变、新效(neomorphic)突变或反效(antimorphic)突变。在一些实施例中，突变型目标基因座包含常染色体显性突变、常染色体隐性突变、异型接合子突变、同型接合子突变，或其组合。在一些实施例中，突变型目标基因座选自哺乳动物染色体内的启动子、强化子、信号序列、内含子、外显子、剪接供体位点、剪接受体位点、内部核糖体进入位点、反向外显子、绝缘子、基因、染色体反转和染色体易位。

[0029] 在本文所提供的AAV中的任一个的一些实施例中，AAV分化体F衣壳包含选自分化体F VP1、分化体F VP2和分化体F VP3的至少一种蛋白质。在一些实施例中，AAV分化体F衣壳包含选自分化体F VP1、分化体F VP2和分化体F VP3的至少两种蛋白质。在一些实施例中，AAV分化体F衣壳包含分化体F VP1、分化体F VP2和分化体F VP3蛋白质。在一些实施例中，AAV分化体F衣壳包含分别与SEQ ID NO:1的氨基酸1到736、氨基酸138到736或氨基酸203到736具有至少90％氨基酸序列一致性的VP1、VP2或VP3蛋白质，所述氨基酸分别对应于AAV9衣壳蛋白VP1、VP2和VP3的氨基酸序列。在一些实施例中，AAV分化体F衣壳包含分别与SEQ ID NO:1的氨基酸1到736和氨基酸138到736具有至少90％氨基酸序列一致性的VP1和VP2蛋白质，所述氨基酸分别对应于AAV9衣壳蛋白VP1和VP2的氨基酸序列；分别与SEQ ID NO:1的氨基酸1到736和氨基酸203到736具有至少90％氨基酸序列一致性的VP1和VP3蛋白质，所述氨基酸分别对应于AAV9衣壳蛋白VP1和VP3的氨基酸序列；或分别与SEQ ID NO:1的氨基酸138到736和氨基酸203到736具有至少90％氨基酸序列一致性的VP2和VP3蛋白质，所述氨基酸分别对应于AAV9衣壳蛋白VP2和VP3的氨基酸序列。在一些实施例中，AAV分化体F衣壳包含分别与SEQ ID NO:1的氨基酸1到736、氨基酸138到736或氨基酸203到736具有至少90％氨基酸序列一致性的VP1、VP2和VP3蛋白质，所述氨基酸分别对应于AAV9衣壳蛋白VP1、VP2和VP3的氨基酸序列。在一些实施例中，AAV分化体F衣壳包含分别与SEQ ID NO:2、3、5、6、11、7、8、9、10、4、12、14、15、16、17或13中任一个的氨基酸1到736、氨基酸138到736或氨基酸203到736具有至少90％氨基酸序列一致性的VP1、VP2或VP3蛋白质，所述氨基酸分别对应于AAVF1到AAVF9和AAVF11到AAVF17衣壳蛋白VP1、VP2和VP3的氨基酸序列。在一些实施例中，AAV分化体F衣壳包含分别与SEQ ID NO:2、3、5、6、11、7、8、9、10、4、12、14、15、16、17或13中任一个的氨基酸1到736和氨基酸138到736具有至少90％氨基酸序列一致性的VP1和VP2蛋白质，所述氨基酸分别对应于AAVF1到AAVF9和AAVF11到AAVF17衣壳蛋白VP1和VP2的氨基酸序列；分别与SEQ ID NO:2、3、5、6、11、7、8、9、10、4、12、14、15、16、17或13中任一个的氨基酸1到736和氨基酸203到736具有至少90％氨基酸序列一致性的VP1和VP3蛋白质，所述

氨基酸分别对应于AAVF1到AAVF9和AAVF11到AAVF17衣壳蛋白VP1和VP3的氨基酸序列;或分别与SEQ ID NO:2、3、5、6、11、7、8、9、10、4、12、14、15、16、17或13中任一个的氨基酸138到736和氨基酸203到736具有至少90％氨基酸序列一致性的VP2和VP3蛋白质,所述氨基酸分别对应于AAVF1到AAVF9和AAVF11到AAVF17衣壳蛋白VP2和VP3的氨基酸序列。在一些实施例中,AAV分化体F衣壳包含分别与SEQ ID NO:2、3、5、6、11、7、8、9、10、4、12、14、15、16、17或13中任一个的氨基酸1到736、氨基酸138到736和氨基酸203到736具有至少90％氨基酸序列一致性的VP1、VP2和VP3蛋白质,所述氨基酸分别对应于AAVF1到AAVF9和AAVF11到AAVF17衣壳蛋白VP1、VP2和VP3的氨基酸序列。在一些实施例中,AAV分化体F衣壳包含分别由与SEQ ID NO:18包含至少90％核苷酸序列一致性的核苷酸序列编码的VP1、VP2或VP3蛋白质,所述核苷酸序列分别对应于编码AAV9衣壳蛋白VP1、VP2和VP3的核苷酸序列。在一些实施例中,AAV分化体F衣壳包含由与SEQ ID NO:18包含至少90％核苷酸序列一致性的核苷酸序列编码的VP1和VP2蛋白质;由与SEQ ID NO:18包含至少90％核苷酸序列一致性的核苷酸序列编码的VP1和VP3蛋白质;或由与SEQ ID NO:18包含至少90％核苷酸序列一致性的核苷酸序列编码的VP2和VP3蛋白质。在一些实施例中,AAV分化体F衣壳包含由与SEQ ID NO:18包含至少90％核苷酸序列一致性的核苷酸序列编码的VP1、VP2和VP3蛋白质,所述核苷酸序列对应于编码AAV9衣壳蛋白VP1、VP2和VP3的核苷酸序列。在一些实施例中,AAV分化体F衣壳包含由分别与SEQ ID NO:20、21、22、23、25、24、27、28、29、26、30、31、32、33、34或35中的任一个包含至少90％核苷酸序列一致性的核苷酸序列编码的VP1、VP2或VP3蛋白质,所述核苷酸序列分别对应于编码AAVF1到AAVF9和AAVF11到AAVF17衣壳蛋白VP1、VP2和VP3的核苷酸序列。在一些实施例中,AAV分化体F衣壳包含由与SEQ ID NO:20-35中的任一个包含至少90％核苷酸序列一致性的核苷酸序列编码的VP1和VP2蛋白质;由与SEQ ID NO:20-35中的任一个包含至少90％核苷酸序列一致性的核苷酸序列编码的VP1和VP3蛋白质;或由与SEQ ID NO:20-35中的任一个包含至少90％核苷酸序列一致性的核苷酸序列编码的VP2和VP3蛋白质。在一些实施例中,AAV分化体F衣壳包含由与SEQ ID NO:20、21、22、23、25、24、27、28、29、26、30、31、32、33、34或35中的任一个包含至少90％核苷酸序列一致性的核苷酸序列编码的VP1、VP2和VP3蛋白质,所述核苷酸序列分别对应于编码AAVF1到AAVF9和AAVF11到AAVF17衣壳蛋白VP1、VP2和VP3的核苷酸序列。在一些实施例中,AAV分化体F衣壳包含AAV9VP1、VP2或VP3衣壳蛋白,其分别对应于如SEQ ID NO:1中所阐述的氨基酸1到736、氨基酸138到736和氨基酸203到736。在一些实施例中,AAV分化体F衣壳包含AAV9VP1和VP2衣壳蛋白,所述衣壳蛋白分别对应于如SEQ ID NO:1中所阐述的氨基酸1到736和氨基酸138到736;AAV9VP1和VP3衣壳蛋白,所述衣壳蛋白分别对应于如SEQ ID NO:1中所阐述的氨基酸1到736和氨基酸203到736;或AAV9VP2和VP3衣壳蛋白,所述衣壳蛋白分别对应于如SEQ ID NO:1中所阐述的氨基酸138到736和氨基酸203到736。在一些实施例中,AAV分化体F衣壳包含AAV9衣壳蛋白VP1、VP2和VP3,其分别对应于如SEQ ID NO:1中所阐述的氨基酸1到736、氨基酸138到736和氨基酸203到736。在一些实施例中,AAV分化体F衣壳包含选自AAVF1到AAVF9和AAVF11到AAVF17中任一个的VP1衣壳蛋白的VP1衣壳蛋白,所述衣壳蛋白分别对应于如SEQ ID NO:2、3、5、6、11、7、8、9、10、4、12、14、15、16、17或13中所阐述的氨基酸1到736。在一些实施例中,AAV分化体F衣壳包含独立地选自AAVF1到AAVF9和AAVF11到AAVF17中任一个的VP1和VP2衣壳蛋白的VP1和VP2衣壳蛋白,所述衣壳蛋白分别对应于如SEQ ID NO:2、3、5、6、11、7、8、9、

10、4、12、14、15、16、17或13中所阐述的氨基酸1到736和氨基酸138到736。在一些实施例中，AAV分化体F衣壳包含独立地选自AAVF1到AAVF9和AAVF11到AAVF17中任一个的VP1、VP2和VP3衣壳蛋白的VP1、VP2和VP3衣壳蛋白，所述衣壳蛋白分别对应于如SEQ ID NO:2、3、5、6、11、7、8、9、10、4、12、14、15、16、17或13中所阐述的氨基酸1到736、氨基酸138到736和氨基酸203到736。在一些实施例中，AAV分化体F衣壳包含AAVF1到AAVF9和AAVF11到AAVF17中任一个的VP1、VP2和VP3衣壳蛋白中的每一个，所述衣壳蛋白分别对应于如SEQ ID NO:2、3、5、6、11、7、8、9、10、4、12、14、15、16、17或13中所阐述的氨基酸1到736、氨基酸138到736和氨基酸203到736。

[0030] 在所提供的AAV中的任一个的一些实施例中，分化体F衣壳包含与选自AAVF5(SEQ ID NO:11)、AAVF7(SEQ ID NO:8)、AAVF12(SEQ ID NO:12)、AAVF15(SEQ ID NO:16)、AAVF17(SEQ ID NO:13)、AAVF9(SEQ ID NO:10)、AAVF16(SEQ ID NO:17)、其变异体、片段、突变体以及任何组合的群组的多肽序列具有至少95％序列一致性百分比的多肽序列。如本文所用，AAVF1、AAVF2、AAVF3、AAVF4、AAVF5、AAVF6、AAVF7、AAVF8、AAVF9、AAVF10、AAVF11、AAVF12、AAVF13、AAVF14、AAVF15、AAVF16和AAVF17也分别被称作AAVHSC1、AAVHSC2、AAVHSC3、AAVHSC4、AAVHSC5、AAVHSC6、AAVHSC7、AAVHSC8、AAVHSC9、AAVHSC10、AAVHSC11、AAVHSC12、AAVHSC13、AAVHSC14、AAVHSC15、AAVHSC16和AAVHSC17。换言之，对AAVF1、AAVF2、AAVF3、AAVF4、AAVF5、AAVF6、AAVF7、AAVF8、AAVF9、AAVF10、AAVF11、AAVF12、AAVF13、AAVF14、AAVF15、AAVF16或AAVF17的任何叙述分别等效于AAVHSC1、AAVHSC2、AAVHSC3、AAVHSC4、AAVHSC5、AAVHSC6、AAVHSC7、AAVHSC8、AAVHSC9、AAVHSC10、AAVHSC11、AAVHSC12、AAVHSC13、AAVHSC14、AAVHSC15、AAVHSC16或AAVHSC17，并且可以由其替换。

[0031] 在本文所提供的AAV中的任一个的一些实施例中，哺乳动物染色体选自人类染色体1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、X以及Y。在本文所提供的AAV中的任一个的一些实施例中，哺乳动物染色体选自小鼠染色体1、2、3、4、5、6、7、8、9、10,11、12、13、14、15、16、17、18、19、X以及Y。在一些实施例中，哺乳动物染色体不是人类染色体19。

[0032] 在本文所提供的AAV中的任一个的一些实施例中，哺乳动物染色体是体细胞染色体。在一些实施例中，体细胞来自选自由以下组成的群组的组织：结缔组织(包括血液)、肌肉组织、神经组织和上皮组织。在一些实施例中，体细胞来自选自由以下组成的群组的器官：肺、心脏、肝脏、肾脏、肌肉、大脑、眼睛、乳房、骨骼和软骨。在一些实施例中，体细胞是CD34+细胞。

[0033] 在本文所提供的AAV中的任一个的一些实施例中，所述细胞是干细胞。在一些实施例中，干细胞是造血干细胞、脐血干细胞、骨髓干细胞、胎儿肝脏干细胞或周边血液干细胞。在一些实施例中，细胞选自由以下组成的群组：CD34+造血干细胞系(HSC)、K562CD34+白血病细胞系、HepG2人类肝脏细胞系、周边血液干细胞、脐血干细胞、CD34+周边血液干细胞、WI-38人类二倍体成纤维细胞细胞系、MCF7人类乳癌细胞系、Y79人类成视网膜细胞瘤细胞系、SCID-X1LBL人类EBV永生化B细胞系、原代人类肝细胞、原代肝窦状隙内皮细胞以及原代骨骼肌成肌细胞。

[0034] 在本文所提供的AAV中的任一个的一些实施例中，AAV用于将编辑元件整合到细胞中哺乳动物染色体的目标基因座中的染色体整合效率为至少约5％。在一些实施例中，AAV

用于将编辑元件整合到细胞中哺乳动物染色体的目标基因座中的染色体整合效率为至少约10％。

[0035] 本发明的其它方面涉及包含如本文所描述的AAV的组合物，其中所述组合物呈药学上可接受的调配物形式。在一些实施例中，调配物被构成为用于向哺乳动物投与。在一些实施例中，调配物被构成为用于经由静脉内注射、皮下注射、肌肉内注射、自体细胞转移或同种异体细胞转移向哺乳动物投与。在一些实施例中，药学上可接受的调配物包含赋形剂。在一些实施例中，赋形剂选自载剂、佐剂和媒剂，或其组合。

[0036] 本发明的又其它方面涉及用于重组制备腺相关病毒(AAV)的包装系统，其中所述包装系统包含编码AAV分化体F衣壳的一个或多个AAV Rep蛋白质的Rep核苷酸序列；编码其一个或多个AAV Cap蛋白质的Cap核苷酸序列；以及如本文所描述的校正基因组；其中所述包装系统在细胞中可操作地用于将校正基因组包封于衣壳中以形成腺相关病毒。在一些实施例中，包装系统包含有包含Rep核苷酸序列和Cap核苷酸序列的第一载体以及包含校正基因组的第二载体。在一些实施例中，AAV分化体F衣壳包含选自分化体F VP1、分化体F VP2和分化体F VP3的至少一种蛋白质。在一些实施例中，AAV分化体F衣壳包含选自分化体F VP1、分化体F VP2和分化体F VP3的至少两种蛋白质。在一些实施例中，AAV分化体F衣壳包含分化体F VP1、分化体F VP2和分化体F VP3蛋白质。在一些实施例中，AAV分化体F衣壳是如本文所描述的任何AAV分化体F衣壳。在一些实施例中，Rep核苷酸序列编码AAV2Rep蛋白质。在一些实施例中，所编码的AAV2Rep蛋白质是Rep 78/68或Rep 68/52中的至少一个。在一些实施例中，编码AAV2Rep蛋白质的核苷酸序列包含与SEQ ID NO:40的AAV2Rep核苷酸序列具有最小序列一致性百分比的核苷酸序列，其中所述最小序列一致性百分比是跨编码AAV2Rep蛋白质的核苷酸序列的长度至少70％。在所提供的包装系统中的任一个的一些实施例中，包装系统进一步包含第三载体，其中所述第三载体是辅助病毒载体。在一些实施例中，辅助病毒载体是独立的第三载体。在一些实施例中，辅助病毒载体与第一载体整合。在一些实施例中，辅助病毒载体与第二载体整合。在一些实施例中，第三载体包含编码辅助病毒蛋白质的基因。在一些实施例中，辅助病毒选自由以下组成的群组：腺病毒、疱疹病毒(包括单纯性疱疹病毒(HSV))、牛痘病毒和巨细胞病毒(CMV)。在一些实施例中，辅助病毒是腺病毒。在一些实施例中，腺病毒基因组包含一个或多个选自由以下组成的群组的腺病毒RNA基因：E1、E2、E4和VA。在一些实施例中，辅助病毒是HSV。在一些实施例中，HSV基因组包含选自由以下组成的群组的HSV基因中的一个或多个：UL5/8/52、ICPO、ICP4、ICP22以及UL30/UL42。在一些实施例中，第一载体和第三载体含于第一转染质体内。在一些实施例中，第二载体和第三载体的核苷酸含于第二转染质体内。在一些实施例中，第一载体和第三载体的核苷酸克隆到重组辅助病毒中。在一些实施例中，第二载体和第三载体的核苷酸克隆到重组辅助病毒中。在一些实施例中，AAV衣壳是选自由以下组成的群组的分化体F AAV的衣壳：AAV9、AAVF1、AAVF2、AAVF3、AAVF4、AAVF5、AAVF6、AAVF7、AAVF8、AAVF9、AAVF11、AAVF12、AAVF13、AAVF14、AAVF15、AAVF16、AAVF17、AAVHU31以及AAVHU32。在一些实施例中，本文所描述的包装系统中的任一个包含于试剂盒内。

[0037] 本发明的其它方面涉及一种用于重组制备腺相关病毒(AAV)的方法，其中所述方法包含用如本文所描述的包装系统在可操作地用于将校正基因组包封于衣壳中的条件下转染或转导细胞以形成AAV。

[0038] 在其它方面，本发明提供一种用于编辑哺乳动物基因组的目标基因座的方法，其中所述方法包含用如本文所描述的腺相关病毒(AAV)转导包含所述哺乳动物基因组的细胞。在一些实施例中，细胞是哺乳动物干细胞。在一些实施例中，哺乳动物细胞来自选自由以下组成的群组的组织：结缔组织(包括血液)、肌肉组织、神经组织和上皮组织。在一些实施例中，哺乳动物细胞来自选自由以下组成的群组的器官：肺、心脏、肝脏、肾脏、肌肉、大脑、眼睛、乳房、骨骼和软骨。在一些实施例中，哺乳动物细胞是干细胞。在一些实施例中，干细胞是造血干细胞、脐血干细胞或周边血液干细胞。在一些实施例中，哺乳动物细胞是成肌细胞、内皮细胞、肝脏细胞、成纤维细胞、乳房细胞、淋巴细胞或视网膜细胞。本发明的其它方面涉及可通过本文所描述的任何方法获得的细胞。

[0039] 本发明的另一方面涉及一种用于编辑哺乳动物基因组的目标基因座的方法，其中所述方法包含：(a)从哺乳动物获得哺乳动物细胞；(b)离体培养所述哺乳动物细胞以形成离体培养物；(c)在所述离体培养物中用如本文所描述的腺相关病毒(AAV)转导所述哺乳动物细胞以形成经过转导的哺乳动物细胞；以及(d)向所述哺乳动物投与所述经过转导的哺乳动物细胞。

[0040] 在其它方面，本发明提供一种用于编辑哺乳动物基因组的目标基因座的方法，其中所述方法包含：(a)从第一哺乳动物获得哺乳动物细胞；(b)离体培养所述哺乳动物细胞以形成离体培养物；(c)在所述离体培养物中用如本文所描述的腺相关病毒(AAV)转导所述哺乳动物细胞以形成经过转导的哺乳动物细胞；以及(d)向第二哺乳动物投与所述经过转导的哺乳动物细胞。在一些实施例中，第一哺乳动物和第二哺乳动物是相同物种。在一些实施例中，哺乳动物细胞来自选自由以下组成的群组的组织：结缔组织(包括血液)、肌肉组织、神经组织和上皮组织。在一些实施例中，哺乳动物细胞来自选自由以下组成的群组的器官：肺、心脏、肝脏、肾脏、肌肉、大脑、眼睛、乳房、骨骼和软骨。在一些实施例中，哺乳动物细胞是干细胞。在一些实施例中，干细胞是造血干细胞、脐血干细胞或周边血液干细胞。在一些实施例中，哺乳动物细胞是CD34+细胞。在一些实施例中，哺乳动物细胞是成肌细胞、内皮细胞、肝脏细胞、成纤维细胞、乳房细胞、淋巴细胞或视网膜细胞。

[0041] 本发明的另一方面提供一种用于编辑哺乳动物基因组的目标基因座的方法，其中所述方法包含以有效用AAV体内转导哺乳动物细胞的量向哺乳动物投与如本文所描述的AAV或如本文所描述的组合物。

[0042] 在所提供的方法中的任一个的一些实施例中，AAV在不共转导或共投与外源性核酸酶或编码外源性核酸酶的核苷酸序列的情况下进行转导或投与。

[0043] 在所提供的方法中的任一个的一些实施例中，所述AAV用于将编辑元件整合到哺乳动物染色体的目标基因座中的染色体整合效率为至少约1％。在一些实施例中，AAV用于将编辑元件整合到哺乳动物染色体的目标基因座中的染色体整合效率是至少约2％、3％、4％或5％。在一些实施例中，校正基因组的编辑元件以哺乳动物细胞至少10％、20％、40％或50％的染色体整合效率整合到哺乳动物染色体的目标基因座中。在一些实施例中，校正基因组的编辑元件以介于哺乳动物细胞10％到70％、20％到70％、40％到70％或50％到70％范围内的染色体整合效率整合到哺乳动物染色体的目标基因座中。

[0044] 在所提供的方法中的任一个的一些实施例中，AAV的染色体整合效率的另一个特征为，对于包含整合到哺乳动物染色体的目标基因座中的编辑元件的等位基因，细胞群体

中的等位基因出现频率为至少约10％。在一些实施例中，AAV的染色体整合效率的另一个特征为，对于包含整合到哺乳动物染色体的目标基因座中的编辑元件的等位基因，细胞群体中的等位基因出现频率为至少约50％。在一些实施例中，AAV的染色体整合效率的另一个特征为，对于包含整合到哺乳动物染色体的目标基因座中的编辑元件的等位基因，细胞群体中的等位基因出现频率为至少约75％。在一些实施例中，细胞群体中的等位基因出现频率是体外细胞群体中的等位基因出现频率。

[0045] 本发明的其它方面涉及一种用于产生转基因非人类动物的方法，所述方法包含向非人类动物投与如本文所描述的AAV或如本文所描述的组合物；或用如本文所描述的AAV或如本文所描述的组合物转导非人类动物细胞，并且在足以由宿主非人类动物产生转基因非人类动物的条件下将所述细胞植入到所述宿主非人类动物中（例如，通过使所植入的细胞形成或变为胚胎的一部分，然后其在宿主中发育成转基因非人类动物）。在一些实施例中，使转基因非人类动物与另一种非人类动物杂交以产生其它转基因非人类动物。在一些实施例中，非人类动物细胞来源于非人类动物的接合子或胚胎。在一些实施例中，非人类动物是小鼠、大鼠、兔、猪、牛、绵羊、山羊、鸡、猫、犬、雪貂或灵长类动物。

[0046] 本发明的其它方面涉及一种转基因非人类动物，其可通过本文所描述的方法，如上文所描述的方法获得。在一些实施例中，转基因非人类动物是小鼠、大鼠、兔、猪、牛、绵羊、山羊、鸡、猫、犬、雪貂或灵长类动物。

[0047] 本发明的又其它方面涉及来源于如本文所描述的转基因非人类动物的组织。在一些实施例中，组织选自由以下组成的群组：结缔组织（包括血液）、肌肉组织、神经组织、内皮组织和上皮组织。在一些实施例中，组织来自选自由以下组成的群组的器官：肺、心脏、肝脏、肾脏、肌肉、大脑、眼睛、乳房、骨骼和软骨。

[0048] 本发明的其它方面涉及一种来源于如本文所描述的转基因非人类动物的细胞。在一些实施例中，细胞是原代细胞。在一些实施例中，细胞是CD34+细胞、成肌细胞、内皮细胞、肝脏细胞、成纤维细胞、乳房细胞、淋巴细胞或视网膜细胞。在一些实施例中，细胞是诱导型多能干（iPS）细胞。在一些实施例中，细胞来自选自由以下组成的群组的组织：结缔组织（包括血液）、肌肉组织、神经组织、内皮组织和上皮组织。在一些实施例中，细胞来自选自由以下组成的群组的器官：肺、心脏、肝脏、肾脏、肌肉、大脑、眼睛、乳房、骨骼和软骨。在一些实施例中，细胞是干细胞。在一些实施例中，干细胞是造血干细胞、脐血干细胞或周边血液干细胞。

[0049] 根据某些实施例，提供用于编辑细胞基因组的腺相关病毒（AAV）分化体F载体（例如，包含包封于分化体F衣壳中的校正基因组的复制缺陷型AAV）或AAV载体变异体（例如，包含相对于AAV9衣壳的衣壳变异体的复制缺陷型AAV）。在某些实施例中，AAV分化体F载体或AAV载体变异体可以包含一个或多个分化体F衣壳或一个或多个衣壳变异体（相对于AAV9衣壳）、包含一个或多个待整合到基因组的目标基因座（在本文中也称为目标位点）中的治疗性核苷酸序列的编辑元件（在本文中也称为靶向盒）、侧接所述编辑元件（靶向盒）并且与位于所述目标基因座（目标位点）上游的区具有同源性的5'同源臂多核苷酸序列以及侧接所述编辑元件（靶向盒）并且与位于所述目标基因座（目标位点）下游的区具有同源性的3'同源臂多核苷酸序列。编辑元件（靶向盒）可以含于如本文所描述的校正基因组内，所述校正基因组包含如本文所描述的反向末端重复（ITR）。在某些实施例中，一个或多个分化体F衣

壳或衣壳变异体可以是本文所描述的分化体F衣壳或衣壳变异体中的任一个。在某些实施例中，一个或多个分化体F衣壳或衣壳变异体可以包含选自以下群组的多肽序列：AAVF7(SEQ ID NO:8)、AAVF12(SEQ ID NO:12)、AAVF15(SEQ ID NO:16)、AAVF17(SEQ ID NO:13)、其变异体、片段、突变体以及任何组合。在某些实施例中，一个或多个分化体F衣壳或一个或多个衣壳变异体可以包含选自以下群组的多肽序列：AAVF5(SEQ ID NO:11)、AAVF7(SEQ ID NO:8)、AAVF12(SEQ ID NO:12)、AAVF15(SEQ ID NO:16)、AAVF17(SEQ ID NO:13)、AAVF9(SEQ ID NO:10)、AAVF16(SEQ ID NO:17)、其变异体、片段、突变体以及任何组合。在某些实施例中，一个或多个分化体F衣壳或衣壳变异体可以包含与选自以下群组的多肽序列具有至少95％序列一致性百分比的多肽序列：AAVF7(SEQ ID NO:8)、AAVF12(SEQ ID NO:12)、AAVF15(SEQ ID NO:16)、AAVF17(SEQ ID NO:13)、其变异体、片段、突变体以及任何组合。在某些实施例中，一个或多个分化体F衣壳或衣壳变异体可以包含与选自以下群组的多肽序列具有至少95％序列一致性百分比的多肽序列：AAVF5(SEQ ID NO:11)、AAVF7(SEQ ID NO:8)、AAVF12(SEQ ID NO:12)、AAVF15(SEQ ID NO:16)、AAVF17(SEQ ID NO:13)、AAVF9(SEQ ID NO:10)、AAVF16(SEQ ID NO:17)、其变异体、片段、突变体以及任何组合。在某些实施例中，目标基因座(目标位点)可以是安全港(safe harbor)位点。在某些实施例中，安全港位点可以是染色体19上的AAVS1基因座。在某些实施例中，目标基因座(目标位点)可以是与如本文所描述的疾病病况相关联的基因座。在某些实施例中，细胞可以是干细胞。在某些实施例中，干细胞可以是造血干细胞、多能干细胞、胚胎干细胞或间叶细胞干细胞。

[0050] 根据某些实施例，提供编辑细胞基因组的方法。在某些实施例中，编辑细胞基因组的方法可以包含用一种或多种如本文所描述的AAV分化体F载体(例如，包含包封于分化体F衣壳中的校正基因组的复制缺陷型AAV)或AAV载体变异体(例如，包含相对于AAV9衣壳的衣壳变异体的复制缺陷型AAV)转导所述细胞。在某些实施例中，转导可以在没有其它外源性核酸酶的情况下进行。在某些实施例中，AAV分化体F载体或AAV载体变异体可以包含一个或多个分化体F衣壳或衣壳变异体(相对于AAV9衣壳)、包含一个或多个待整合到基因组的目标基因座(目标位点)中的治疗性核苷酸序列的编辑元件(靶向盒)、侧接所述编辑元件(靶向盒)并且与位于所述目标基因座(目标位点)上游的区具有同源性的5'同源臂多核苷酸序列以及侧接所述编辑元件(靶向盒)并且与位于所述目标基因座(目标位点)下游的区具有同源性的3'同源臂多核苷酸序列。编辑元件(靶向盒)可以含于包含如本文所描述的ITR的如本文所描述的校正基因组内。在某些实施例中，一个或多个分化体F衣壳或衣壳变异体可以包含选自以下群组的多肽序列：AAVF7(SEQ ID NO:8)、AAVF12(SEQ ID NO:12)、AAVF15(SEQ ID NO:16)、AAVF17(SEQ ID NO:13)、其变异体、片段、突变体以及任何组合。在某些实施例中，一个或多个分化体F衣壳或衣壳变异体可以包含选自以下群组的多肽序列：AAVF5(SEQ ID NO:11)、AAVF7(SEQ ID NO:8)、AAVF12(SEQ ID NO:12)、AAVF15(SEQ ID NO:16)、AAVF17(SEQ ID NO:13)、AAVF9(SEQ ID NO:10)、AAVF16(SEQ ID NO:17)、其变异体、片段、突变体以及任何组合。在某些实施例中，一个或多个分化体F衣壳或衣壳变异体可以包含与选自以下群组的多肽序列具有至少95％序列一致性百分比的多肽序列：AAVF7(SEQ ID NO:8)、AAVF12(SEQ ID NO:12)、AAVF15(SEQ ID NO:16)、AAVF17(SEQ ID NO:13)、其变异体、片段、突变体以及任何组合。在某些实施例中，一个或多个分化体F衣壳或衣壳变异体可以包含与选自以下群组的多肽序列具有至少95％序列一致性百分比的多肽序列：AAVF5(SEQ ID

NO:11)、AAVF7(SEQ ID NO:8)、AAVF12(SEQ ID NO:12)、AAVF15(SEQ ID NO:16)、AAVF17(SEQ ID NO:13)、AAVF9(SEQ ID NO:10)、AAVF16(SEQ ID NO:17)、其变异体、片段、突变体以及任何组合。在某些实施例中,AAV分化体F载体或AAV载体变异体不含有用于一个或多个治疗性核苷酸序列的启动子。在某些实施例中,目标基因座(目标位点)可以是安全港位点。在某些实施例中,安全港位点可以是染色体19上的AAVS1基因座。在某些实施例中,目标基因座(目标位点)可以是与如本文所描述的疾病病况相关联的基因座。在某些实施例中,细胞可以是干细胞。在某些实施例中,干细胞可以是造血干细胞、多能干细胞、胚胎干细胞或间叶细胞干细胞。

[0051] 根据某些实施例,提供在个体中通过编辑所述个体的细胞基因组来治疗疾病或病症的方法。在某些实施例中,在个体中通过编辑所述个体的细胞基因组来治疗疾病或病症的方法包括以下步骤:用如本文所描述的AAV分化体F载体或AAV载体变异体转导所述个体的所述细胞和将经过转导的细胞移植到所述个体中,其中所述经过转导的细胞治疗所述疾病或病症。在某些实施例中,细胞的转导可以在没有其它外源性核酸酶的情况下进行。在某些实施例中,AAV分化体F载体或AAV载体变异体可以包含一个或多个分化体F衣壳或衣壳变异体(相对于AAV9衣壳)、包含一个或多个待整合到基因组的目标基因座(目标位点)中的治疗性核苷酸序列的编辑元件(靶向盒)、侧接所述编辑元件(靶向盒)并且与位于所述目标基因座(目标位点)上游的区具有同源性的5'同源臂多核苷酸序列以及侧接所述编辑元件(靶向盒)并且与位于所述目标基因座(目标位点)下游的区具有同源性的3'同源臂多核苷酸序列。编辑元件(靶向盒)可以含于包含如本文所描述的ITR的如本文所描述的校正基因组内。在某些实施例中,分化体F衣壳或衣壳变异体可以包含选自以下群组的多肽序列:AAVF7(SEQ ID NO:8)、AAVF12(SEQ ID NO:12)、AAVF15(SEQ ID NO:16)、AAVF17(SEQ ID NO:13)、其变异体、片段、突变体以及任何组合。在某些实施例中,一个或多个分化体F衣壳或衣壳变异体可以包含选自以下群组的多肽序列:AAVF5(SEQ ID NO:11)、AAVF7(SEQ ID NO:8)、AAVF12(SEQ ID NO:12)、AAVF15(SEQ ID NO:16)、AAVF17(SEQ ID NO:13)、AAVF9(SEQ ID NO:10)、AAVF16(SEQ ID NO:17)、其变异体、片段、突变体以及任何组合。在某些实施例中,一个或多个分化体F衣壳或衣壳变异体可以包含与选自以下群组的多肽序列具有至少95%序列一致性百分比的多肽序列:AAVF7(SEQ ID NO:8)、AAVF12(SEQ ID NO:12)、AAVF15(SEQ ID NO:16)、AAVF17(SEQ ID NO:13)、其变异体、片段、突变体以及任何组合。在某些实施例中,一个或多个分化体F衣壳或衣壳变异体可以包含与选自以下群组的多肽序列具有至少95%序列一致性百分比的多肽序列:AAVF5(SEQ ID NO:11)、AAVF7(SEQ ID NO:8)、AAVF12(SEQ ID NO:12)、AAVF15(SEQ ID NO:16)、AAVF17(SEQ ID NO:13)、AAVF9(SEQ ID NO:10)、AAVF16(SEQ ID NO:17)、其变异体、片段、突变体以及任何组合。在某些实施例中,AAV分化体F载体或AAV载体变异体不含有用于一个或多个治疗性核苷酸序列的启动子。在某些实施例中,目标基因座(目标位点)可以是安全港位点。在某些实施例中,安全港位点可以是染色体19上的AAVS1基因座。在某些实施例中,目标基因座(目标位点)可以是与如本文所描述的疾病病况相关联的基因座。在某些实施例中,细胞可以是干细胞。在某些实施例中,干细胞可以是造血干细胞、多能干细胞、胚胎干细胞或间叶细胞干细胞。在某些实施例中,疾病或病症可以由细胞基因组中的一个或多个突变造成。在某些实施例中,疾病或病症可以选自遗传性代谢病、溶酶体贮积病(lysosomal storage disease)、粘多糖病

(mucopolysaccharidodosis)、免疫缺陷疾病以及血红蛋白病和感染。

[0052] 本文还公开在个体中通过体内基因组编辑来治疗疾病或病症的方法，所述体内基因组编辑通过向所述个体直接投与如本文所描述的AAV分化体F载体或AAV载体变异体来进行。在某些实施例中，公开在个体中通过对所述个体的细胞进行体内基因组编辑来治疗疾病或病症的方法，其通过向所述个体直接投与AAV分化体F载体或AAV载体变异体来进行。在某些实施例中，AAV分化体F载体或AAV载体变异体可以包含一个或多个分化体F衣壳或衣壳变异体(相对于AAV9衣壳)、包含一个或多个待整合到基因组的目标基因座(目标位点)中的治疗性核苷酸序列的编辑元件(靶向盒)、侧接所述编辑元件(靶向盒)并且与位于所述目标基因座(目标位点)上游的区具有同源性的5'同源臂多核苷酸序列以及侧接所述编辑元件(靶向盒)并且与位于所述目标基因座(目标位点)下游的区具有同源性的3'同源臂多核苷酸序列，其中所述载体转导个体的细胞并且将所述一个或多个治疗性核苷酸序列整合到所述细胞的基因组中。编辑元件(靶向盒)可以含于包含如本文所描述的ITR的如本文所描述的校正基因组内。在某些实施例中，一个或多个分化体F衣壳或衣壳变异体可以包含选自以下群组的多肽序列：AAVF1(SEQ ID NO:2)、AAVF2(SEQ ID NO:3)、AAVF11(SEQ ID NO:4)、AAVF3(SEQ ID NO:5)、AAVF4(SEQ ID NO:6)、AAVF6(SEQ ID NO:7)、AAVF7(SEQ ID NO:8)、AAVF8(SEQ ID NO:9)、AAVF9(SEQ ID NO:10)、AAVF5(SEQ ID NO:11)、AAVF12(SEQ ID NO:12)、AAVF17(SEQ ID NO:13)、AAVF13(SEQ ID NO:14)、AAVF14(SEQ ID NO:15)、AAVF15(SEQ ID NO:16)、AAVF16(SEQ ID NO:17)、其变异体、片段、突变体以及任何组合。在某些实施例中，AAV分化体F载体或AAV载体变异体不含有用于一个或多个治疗性核苷酸序列的启动子。在某些实施例中，目标基因座(目标位点)可以是安全港位点。在某些实施例中，安全港位点可以是染色体19上的AAVS1基因座。在某些实施例中，目标基因座(目标位点)可以是与如本文所描述的疾病病况相关联的基因座。在某些实施例中，细胞可以是干细胞。在某些实施例中，干细胞可以是造血干细胞、多能干细胞、胚胎干细胞或间叶细胞干细胞。在某些实施例中，疾病或病症可以由细胞基因组中的一个或多个突变造成。在某些实施例中，疾病或病症可以选自遗传性代谢病、溶酶体贮积病、粘多糖病、免疫缺陷疾病以及血红蛋白病和感染。本文还公开包含一种或多种用于编辑细胞基因组的AAV分化体F载体或AAV载体变异体的试剂盒。

附图说明

[0053] 本申请含有至少一个彩图形式的图式。具有彩色图式的此申请的复本将在请求和支付必需费用之后由专利局提供。

[0054] 图1展示分化体F AAV衣壳变异体多肽序列与AAV9相比对。分化体F AAV衣壳变异体多核苷酸序列的相对应比对提供于美国专利公开第US20130096182A1号的图1中。

[0055] 图2是列出每一序列的衣壳中的一些核苷酸突变的图表，包括碱基变化、氨基酸变化以及其是处于VP1还是VP3中。

[0056] 图3展示用于基因组编辑的一组供体ITR-AAVS1-FP载体构筑体中的一部分的示意图。AAV载体含有5'同源臂和3'同源臂以及调节元件，所述调节元件包括2A序列、剪接受体序列和聚腺苷酸化序列。使用黄色荧光蛋白质(“YFP”或“FP”)作为转基因。AAV2ITR侧接同源臂，并且载体基因组包装于AAVF衣壳中以形成AAVF-AAVS1-FP供体载体。重要的是，含有

FP基因的载体不含有驱动表达的启动子。FP基因仅在其正确地整合到位于内源性染色体启动子下游的AAVS1中的情况下得到表示。

[0057] 图4展示目标染色体AAVS1基因座和经过编辑的AAVS1基因座(其是用于由AAVF载体介导的转基因整合的目标位点)的示意性图谱。顶部示意图(图4A)“野生型AAVS1基因座”示出含有5'同源臂和3'同源臂但不含有转基因的野生型AAVS1基因座。用位于同源区外侧的引物使用“外正向引物区”引物和“外反向引物区”引物进行扩增产生约1.9kb长的片段(参见标记为“片段1”的线),其指示不含有经过整合的转基因的片段。底部示意图(图4B)“经过编辑的AAVS1基因座”示出含有5'同源臂、调节元件、经过整合的转基因和3'同源臂的经过编辑的AAVS1基因座。用位于同源区外侧的引物使用“外正向引物区”引物和“外反向引物区”引物进行扩增产生约3.0kb长的片段(参见标记为“片段2”的线),其指示含有转基因的片段。使用“外正向引物区”引物和“内反向引物”对5'接合区(处于5'同源臂与转基因之间的接合点)进行的扩增产生约1.7kb长的片段(参见标记为“片段3”的线)。使用“外反向引物区”引物和“内正向引物”对3'接合区(处于转基因与3'同源臂之间的接合点)进行的扩增产生约1.2kb长的片段(参见标记为“片段4”的线)。如果没有整合转基因,那么在对5'接合区或3'接合区进行扩增不存在所得产物。

[0058] 图5展示在转导之后24小时时,对K562细胞中的YFP表达进行流式细胞测量分析得到的代表性散点图。用AAVF7FP载体以多种感染复数(MOI)转导细胞,(A)未用任何载体转导的细胞(未转导)、(B)50,000MOI、(C)100,000MOI、(D)200,000MOI以及(E)400,000MOI。展示来自代表性样品的数据。每一散布图内高于分界线的事件表示表达FP的细胞,指示在这些细胞中,来自供体ITR-AAVS1-FP载体的无启动子FP基因正确地整合到人类染色体19中位于内源性染色体启动子下游的AAVS1中。

[0059] 图6展示在转导之后72小时时,对K562细胞中的YFP表达进行流式细胞测量分析得到的代表性散点图。用AAVF7FP载体以多种感染复数(MOI)转导细胞,(A)未用任何载体转导的细胞(未转导)、(B)50,000MOI、(C)100,000MOI、(D)200,000MOI以及(E)400,000MOI。高于分界线的事件表示其中供体ITR-AAVS1-FP载体中的无启动子FP基因正确地进行靶向整合的细胞。

[0060] 图7展示在无启动子YFP转基因在CD34+K562白血病细胞中的AAVS1基因座中靶向整合后的平均YFP表达百分比。(A)展示在于细胞中用AAVF7载体以50,000;100,000;150,000;200,000;300,000;和400,000的MOI转导后24小时时的YFP表达的条形图。(B)展示在于细胞中用AAVF7载体以50,000;100,000;150,000;200,000;和400,000的MOI转导后72小时时的YFP表达的条形图。每一条形表示由至多7个样品编译的数据。

[0061] 图8展示PCR确认YFP转基因靶向整合到K562细胞中的AAVS1基因座中。A)展示来自不具有模板、未转导或用AAVF7FP载体以100,000的MOI转导的K562细胞的代表性样品的经过扩增的DNA的凝胶。色带1:DNA序列梯,色带2:无模板对照物,色带3:未转导的对照物,以及色带4:AAVF7FP转导的K562。箭头指向约3.1kb的FP整合的AAVS1片段或约1.9kb的非整合的AAVS1片段。B)展示来自不具有模板、未转导或用AAVF7FP载体以100,000的MOI转导的K562细胞的代表性样品的经过扩增的DNA的凝胶。色带1:DNA序列梯,色带2:无模板对照物,色带3:未转导的对照物,色带4:AAVF7FP转导的K562。箭头指向约3.1kb的经过扩增的FP整合的AAVS1片段或约1.9kb的经过扩增的非整合的AAVS1片段。

[0062] 图9展示在用AAVF FP载体转导后1天时，在原代CD34+细胞中靶向整合之后的YFP表达的代表性散点图。(A)没有用任何载体转导的细胞(未转导)，(B)用AAVF7FP载体转导的细胞，和(C)用AAVF17FP载体转导的细胞。与未转导的细胞相比较，用AAVF7或AAVF17载体转导的细胞展示大量YFP表达(将B和C分别与A比较)。(b)和(c)中的YFP表达指示由AAVF载体递送的无启动子FP基因精确地整合到染色体AAVS1基因座中。

[0063] 图10展示在用AAVF FP载体转导后4天时，在原代CD34+细胞中靶向整合之后的YFP表达的代表性散点图。(A)没有用任何载体转导的细胞(未转导)，(B)用AAVF7FP载体转导的细胞，和(C)用AAVF17FP载体转导的细胞。与未转导的细胞相比较，用AAVF7或AAVF17FP载体转导的细胞展示大量YFP表达(将B和C分别与A比较)，指示由AAVF载体递送的基因的精确靶向整合。

[0064] 图11展示在用AAVF FP载体转导后18天时，在原代CD34+细胞中靶向整合之后的YFP表达的代表性散点图。(A)没有用任何载体转导的细胞(未转导)，(B)用AAVF7FP载体转导的细胞，和(C)用AAVF17FP载体转导的细胞。与未转导的细胞相比较，用AAVF7或AAVF17FP载体转导的细胞展示大量YFP表达(将B和C分别与A比较)。

[0065] 图12展示在原代CD34+细胞中靶向整合之后的YFP表达。(A)展示未转导的细胞和用AAVF7FP载体或AAVF17FP载体转导的细胞在转导后4、18、20和39天时的YFP阳性细胞百分比的表。(B)展示在以100,000的MOI进行AAVF FP转导后4、18、20和39天时，表达YFP的原代CD34+细胞的出现频率的线图。菱形的线表示未转导的细胞，方形的线表示用AAVF7FP载体转导的细胞，并且三角形的线表示用AAVF17FP载体转导的细胞。

[0066] 图13展示PCR确认靶向整合到原代CD34+细胞中的AAVS1基因座中。展示来自不具有模板、未转导或用AAVF7FP载体以150,000的MOI转导的原代CD34+细胞的代表性样品的经过扩增的DNA的凝胶。色带1:DNA序列梯，色带2:无模板对照物，色带3:未转导的对照物，色带4:DNA标记物，以及色带5:AAVF7FP载体转导的。箭头指向展示5'接合区扩增产物的FP整合的AAVS1(约1.7kb片段)。左边的插图展示加载于色带1中的DNA序列梯。

[0067] 图14展示在外正向引物区处开始的序列确认YFP基因序列靶向整合到AAVS1基因座中。测序结果指示，存在YFP基因，并且其正确地整合到AAVS1基因座中。

[0068] 图15展示在接近5'同源臂处开始的序列确认YFP序列靶向整合到AAVS1基因座中。测序结果指示，存在YFP基因，并且其整合到AAVS1基因座中。

[0069] 图16展示在接近调节元件5'区处开始的序列确认YFP序列靶向整合到AAVS1基因座中。测序结果指示，存在YFP基因，并且其整合到AAVS1基因座中。

[0070] 图17展示在接近调节元件3'区处开始的序列确认YFP序列靶向整合到AAVS1基因座中。测序结果指示，存在YFP基因，并且其整合到AAVS1基因座中。

[0071] 图18展示在接近转基因5'区处开始的序列确认YFP序列靶向整合到AAVS1基因座中。测序结果指示，存在YFP基因，并且其整合到AAVS1基因座中。

[0072] 图19展示在接近“内反向引物”区处开始的序列确认YFP序列靶向整合到AAVS1基因座中。测序结果指示，存在YFP基因，并且其整合到AAVS1基因座中。

[0073] 图20展示在实例4中的实验中进行的步骤的示意图。获得一百万个人类脐血CD34+细胞(参见步骤1)，并且将其注射到经过亚致死剂量照射的免疫缺陷型NOD/SCID成年小鼠中(参见步骤2)。在用CD34+细胞注射之后两小时时，用AAVF-荧光素酶载体(即，AAVF7-荧光

素酶载体或AAVF17-荧光素酶载体)对小鼠进行注射。二到七天后,用AAVF-Venus载体(即,AAVF7-Venus载体或AAVF17-Venus载体)对小鼠进行注射(参见步骤3)。最后,在注射后4周时测量体内荧光素酶表达,并且在注射后6周时对Venus表达进行量化(参见步骤4)。

[0074] 图21展示代表性接受者中的体内比荧光素酶表达。图21A展示,先前用人类脐血CD34+HSC异种移植的接受AAVF-荧光素酶载体静脉内注射的免疫缺陷型成年小鼠在脊椎、脾脏、髋部和长骨中显示比荧光素酶表达,所述位置都是移植之后的造血位点。箭头指示脊椎、脾脏、肝脏、髋部和长骨中的荧光素酶表达。肝脏和脾脏的通量是4.08e9,并且尾部的通量是1.74e9。图21B展示,先前未用人类脐血CD34+HSC异种移植的接受AAVF-荧光素酶载体静脉内注射的免疫缺陷型成年小鼠并不显示高水平的比荧光素酶表达。肝脏和脾脏的通量是1.47e8,并且尾部的通量是2.22e8。

[0075] 图22展示示出表达Venus的人类CD34+或CD45+细胞在先前用人类脐血CD34+HSC异种移植的接受AAVF7-Venus或AAVF17-Venus载体静脉内注射的免疫缺陷型成年小鼠中的流式细胞测量术数据的频率分布图。图22A展示来自用AAVF7-Venus载体注射的异种移植小鼠的股骨CD34+细胞的流式细胞测量术数据。9.23%的植入的人类造血细胞表达Venus。图22B展示来自用AAVF7-Venus载体注射的异种移植小鼠的股骨CD45+细胞的流式细胞测量术数据。8.35%的植入的人类造血细胞表达Venus。图22A展示来自用AAVF17-Venus载体注射的异种移植小鼠的股骨CD34+细胞的流式细胞测量术数据。8.92%的植入的人类造血细胞表达Venus。图22D展示来自用AAVF17-Venus载体注射的异种移植小鼠的股骨CD45+细胞的流式细胞测量术数据。8.59%的植入的人类造血细胞表达Venus。图22E展示来自用AAVF7-Venus载体注射的异种移植小鼠的椎骨CD45+细胞的流式细胞测量术数据。15.3%的植入的人类造血细胞表达Venus。图22F展示来自用AAVF17-Venus载体注射的异种移植小鼠的椎骨CD45+细胞的流式细胞测量术数据。70.2%的植入的人类造血细胞表达Venus。图22G展示来自用AAVF7-Venus载体注射的异种移植小鼠的脾脏CD45+细胞的流式细胞测量术数据。10.3%的植入的人类造血细胞表达Venus。图22H展示来自用AAVF17-Venus载体注射的异种移植小鼠的脾脏CD45+细胞的流式细胞测量术数据。9.90%的植入的人类造血细胞表达Venus。来自图22的结果也提供于表5中。

[0076] 图23展示AAV分化体F病毒相对于彼此和其它AAV菌株的关系的系统发生谱图(phylogram)。所述系统发生谱图是基于AAVF病毒衣壳基因的核苷酸序列同源性(Smith等人,《分子治疗学(Mol Ther.)》2014年9月;22(9):1625-34)。

[0077] 图24展示用于插入较大DNA插入物的单链AAV载体基因组的图谱。单链AAV2基因组含有AAV2ITR、同源臂、调节序列和无启动子Venus开放阅读框(ORF)。Venus是荧光报告体蛋白质。含有Venus ORF的无启动子Venus位于剪接受体和2A序列下游。Venus ORF后接聚腺苷酸化信号。每一同源臂都是800bp长,并且靶向染色体19上PPP1R12C基因的内含子1

[0078] 图25展示AAVS1中编辑部分的插入位点的示意图。转基因盒由Venus开放阅读框组成,并且后接2A序列的剪接受体位点在任一侧由同源臂侧接。同源臂与染色体19上AAVS1基因座内人类PPP1R12C基因的内含子1互补,并且介导Venus插入到两个同源臂之间的位点中。

[0079] 图26A-F展示通过重组AAVF载体使较大蛋白质编码序列靶向基因组插入到人类细胞系中,以及展现AAVF介导的基因组编辑是稳健的并且在人类CD34+细胞和细胞系中存在

高效编辑的原代细胞。图26A:CD34+表示原代人类CD34+细胞激素预致敏的周边血液干细胞。图26B:K562是人类CD34+红白血病细胞系。图26C:HepG2是人类肝脏细胞系。对图26A-C展示显示Venus表达(指示精确插入)的细胞百分比。图26D展示代表性流动概况,其展示与未转导的细胞相比较,在用重组AAVF病毒转导之后不同的表达Venus的CD34+细胞群体。图26E展示AAVF7、AAVF12、AAVF15、AAVF17和AAV9与AAV6和AAV8相比较的在K562红白血病谱系中的编辑活性。图26F展示相同病毒在肝脏细胞系HepG2中的编辑活性。数据展示显示编辑和Venus表达的细胞百分比。

[0080] 图27展示用于检测插入到人类染色体19上AAVS1基因座中的较大和较小插入物的靶向整合分析。示意性图谱展示引物的位置。5'引物与染色体序列互补。3'引物对插入物具有特异性。预测特异性扩增子对较大插入物是1.7kb,并且对较小插入物是1kb。分裂的引物对(染色体和插入物特异性)使靶向整合分析具有特异性。

[0081] 图28A-E展示AAVF载体介导指定基因组位点处的核苷酸取代。图28A展示用于将10bp插入物插入到人类PPP1R12C基因的内含子1中的单链AAV载体基因组图谱。此载体编码含有Nhe1限制酶识别位点(GCTAGC)的野生型左同源臂(HA-L)。NS mut载体被设计成用于将染色体19上左同源臂中的TA序列变为AT。此变化引起Nhe1位点转化为Sph1位点,将序列从GCTAGC变为GCATGC。图28B展示,使用位于染色体序列上游的正向引物和位于染色体19上PPP1R12C基因的内含子1中10bp插入物中的反向引物来扩增左同源臂。上部示意图标示在使用野生型或NS Mut AAVF载体编辑来自K562细胞的基因组DNA时产生的预期片段的相对大小。图28C是展示来源于用野生型AAVF载体编辑的K562细胞的基因组DNA的实际扩增子的凝胶。色带展示未切割的扩增子(Un)、用Nhe 1切割(Nhe1)和用Sph 1切割(Sph 1)的扩增子。图28D展示在用编码野生型或NS Mut基因组的AAVF7或AAVF17载体编辑之后的K562DNA的凝胶电泳。图28E展示在用编码野生型或NS Mut基因组的AAVF7或AAVF17载体编辑之后的肝细胞癌细胞系HepG2的凝胶电泳。

[0082] 图29展示对来自用AAVF7和AAVF17野生型或NS Mut载体编辑的细胞的DNA进行的序列分析。

[0083] 图30是展示AAVF载体介导在分裂和非分裂细胞中进行编辑并且AAVF介导的基因编辑不需要DNA合成的表。所述图展示经过编辑的表达Venus的细胞在原代人类CD34+细胞的分裂和非分裂子组中的出现频率。Venus阳性并且BrdU阳性或阴性的所有CD34+细胞的百分比通过流式细胞测量术来测定。BrdU阳性细胞分裂细胞,并且BrdU阴性细胞表示非分裂细胞。

[0084] 图31A-C展示全身递送的AAVF载体在体内对植入的人类造血干细胞的高效编辑。图31A展示实验设计的图解。用人类脐血CD34+造血干细胞植入免疫缺陷型NOD/SCID小鼠。在静脉内注射AAVF17-Venus之前允许细胞植入7周。在AAVF注射之后12.5周时,从异种移植小鼠的椎骨和股骨骨髓和脾脏中采集造血细胞。通过多色流式细胞测量术分析细胞的Venus表达以及人类特异性表面标记物的存在。具体地说,在红细胞谱系的原始CD34+人类造血干细胞/祖细胞、CD45+人类分化单核造血细胞和血型糖蛋白A+细胞中分析Venus表达。图31B展示人类红细胞谱系从CD34+祖细胞到血型糖蛋白A+红细胞的分化路径的示意图。图31C展示在移植之后20周时,长期植入的人类细胞在异种移植小鼠的骨髓和脾脏细胞中的流式细胞测量概况。分析细胞的Venus(编辑标记物)以及特异性人类细胞表面标记物两者

的表达。

[0085] 图32A和B是向用人类脐血CD34+造血干细胞/祖细胞异种移植的免疫缺陷型小鼠中静脉内注射AAVF载体后的体内数据概述。Venus表达反映无启动子Venus盒靶向插入到体内植入的人类造血干细胞和其子代中的人类PPP1R12C基因的内含子1中。图32B是图32A中数据的概述。图32A和B展示编辑是长期的，编辑是稳定遗传的(插入物在分化的子代细胞中高效地表达)，体内编辑比离体转导后接移植高效得多，并且经过编辑的CD34+细胞的子代长期保留Venus表达。

[0086] 图33展示对无启动子SA/2A venus ORF靶向染色体插入K562红白血病细胞系、原代人类细胞激素预致敏的周边血液CD34+细胞(PBSC)和HepG2人类肝脏细胞系中进行的序列分析。使用染色体特异性引物和插入物特异性引物对位点特异性整合的序列进行扩增。将经过扩增的产物克隆到TOPO-TA载体中，并且使用桑格测序法(Sanger sequencing)进行测序。

[0087] 图34对10bp插入物靶向染色体插入到原代人类细胞激素预致敏的周边血液CD34+细胞和HepG2人类肝脏细胞系中进行的序列分析。使用染色体特异性引物和插入物特异性引物对位点特异性整合的序列进行扩增。将经过扩增的产物克隆到TOPO-TA载体中，并且使用桑格测序法进行测序。

[0088] 图35展示AAVF靶向较小插入物向AAVS1-CD34+细胞中的整合。

[0089] 图36展示AAV在HepG2(肝癌)细胞系中靶向无启动子Venus和RFLP。

[0090] 图37展示代表性股骨骨髓流式细胞测量术图，其展示分选的总群体、CD34和glycoA阳性细胞对Venus阳性和Venus阴性群体的背栅控、Venus阴性群体中的CD34/GlycoA阳性细胞以及glycoA阳性和CD34阳性群体。

[0091] 图38展示代表性脾脏流式细胞测量术图，其展示分选的总群体、CD34和glycoA阳性细胞对Venus阳性和Venus阴性群体的背栅控、Venus阴性群体中的CD34/GlycoA阳性细胞以及glycoA阳性和CD34阳性群体。

[0092] 图39展示用于测定AAV载体的相对转导对比编辑效率的CBA-mCherry和AAS1-Venus载体基因组的图谱。

[0093] 图40A和40B展示人类CD34+脐血细胞中mCherry(图40A)和Venus(图40B)的流式细胞测量概况。图40C展示在转导之后48小时时，CD34+细胞中mCherry和Venus表达的量化。图40D展示Venus与mCherry的相对表达的比较(编辑比率)。条形以具有相对应衣壳的表达mCherry的那些细胞的比率形式指示表达Venus的细胞比例的比率。黑色水平条形指示比率为1，其将指示Venus:mCherry表达效率相等。

具体实施方式

[0094] 使用具体实例、序列和附图详细地描述本发明的某些实施例。所列举实施例并不打算将本发明限制为那些实施例，因为本发明打算涵盖可以包括在如由权利要求书所界定的本发明范围内的所有替代方案、修改和等效物。所属领域的技术人员将认识到类似或等效于本文所描述的那些方法和材料的方法和材料，其可以用于本发明的实践中。除非另外定义，否则本文所用的所有技术和科学术语都具有与本发明所属领域的普通技术人员通常所理解的相同含义。所有公开和/或专利以引用的方式并入，如同充分阐述于本文中一般。

[0095] 本文提供腺相关病毒(AAV)分化体F载体(例如,包含包封于分化体F衣壳中的校正基因组的复制缺陷型AAV)或AAV载体变异体(例如,包含相对于AAV9衣壳的衣壳变异体的复制缺陷型AAV)以及研发用于在不需要添加外源性核酸酶的情况下使用同源重组来精确编辑细胞基因组的其相关方法。在某些实施例中,基因组编辑可以包括(但不限于)将插入、缺失、改变、点突变或其任何组合引入到细胞的基因组序列(例如哺乳动物染色体的目标基因座)中。在某些实施例中,本文所提供的AAV分化体F载体或AAV载体变异体以及其相关方法可以用于在于整合一个或多个核苷酸序列之前不需要添加外源性核酸酶的情况下将所述一个或多个核苷酸序列插入到细胞基因组的特定位置中。在某些实施例中,本文所提供的分化体F载体或AAV载体变异体以及其相关方法可以用于在于整合核苷酸间键之前不需要添加外源性核酸酶的情况下将所述核苷酸间键插入到细胞基因组的特定位置中。在某些实施例中还提供在个体中通过离体编辑所述个体的细胞基因组来治疗疾病或病症的方法,其经由以下步骤来进行:用如本文所描述的分化体F载体或AAV载体变异体转导所述细胞和将经过转导的细胞进一步移植到所述个体中以治疗所述个体的所述疾病或病症。本文还提供在个体中通过体内基因组编辑来治疗疾病或病症的方法,所述体内基因组编辑通过向所述个体直接投与如本文所描述的分化体F载体或AAV载体变异体来进行。本文还提供用于对细胞进行基因组编辑的试剂盒,其包含本文所描述的分化体F载体或AAV载体变异体中的一种或多种。

[0096] 先前已经报告使用各种AAV载体(例如AAV2、AAV6和AAV8)的同源重组;然而,所报告的效率极低,大致一百万细胞中1个。如下文实例1和2中所示,AAV分化体F载体(或AAV载体变异体)用于以比先前所见显著更大的频率可再现地靶向进入指定染色体位置处的基因插入。举例来说,靶向基因组编辑通过以下来实现:用AAV分化体F载体(或AAV载体变异体)转导原代细胞,以引起转基因以出乎意料地高的出现频率插入到所述原代细胞的基因组中,其中在转导后六周时,大致10%的原代细胞显示转基因插入。此出现频率的效率是先前所报告的1,000到100,000倍(参见例如Khan,2011)。如下文实例1和2中所示,高水平基因组编辑使用原代人类CD34+造血干细胞(K562)和CD34+原代周边血液源性人类造血干细胞(PBSC)来实现。在短期(一天)和长期(高达几乎六周)CD34+培养物中观察到靶向基因插入,并且通过转基因表达和序列分析加以证实。此外,如本文所描述的分化体F载体或AAV载体变异体靶向重组允许进行特异性基因组工程改造而无相关毒性。如下文实例3中所示,静脉内注射用AAVF7或AAVF17假型化的AAV载体在体内引起人类CD34+造血干细胞和祖细胞的转导。如下文实例4中所示,在细胞培养物中和体内对较小(约10bps)和较大插入(约800bps)实现基因组编辑,并且通过测序展示其精确地整合到目标基因座中。如下文实例5中所示,实现对各种人类细胞系(例如成纤维细胞、肝细胞癌细胞、乳癌细胞、成视网膜细胞瘤细胞、白血病细胞和B细胞)的基因组编辑,展现分化体F载体可以用于编辑若干不同细胞类型种的基因组,如成纤维细胞、肝细胞、乳房细胞、视网膜细胞和B细胞。因而,此技术具有用于在细胞中离体进行以及在特定器官中体内进行靶向基因组编辑的极大潜能。

[0097] 本文提供用于编辑细胞基因组的分化体F载体(例如,包含包封于分化体F衣壳中的校正基因组的复制缺陷型AAV)或AAV载体变异体(例如,包含相对于AAV9衣壳的衣壳变异体的复制缺陷型AAV)以及其方法(经由重组,并且优选地不使用外源性核酸酶)。在某些实施例中,基因组编辑可以包括(但不限于)在基因组中校正或插入一个或多个突变,使基因

组中的一个或多个核苷酸缺失，改变包括调节序列的基因组序列，在基因组中的安全港位点或其它特定位置处插入一个或多个包括转基因的核苷酸，或其任何组合。在某些实施例中，如本文所描述的使用分化体F载体或AAV载体变异体的基因组编辑和其方法可以在不插入外源性病毒序列或其它足迹的情况下诱导一个或多个基因组序列的精确改变。

[0098] 在一些方面，本发明提供一种复制缺陷型腺相关病毒(AAV)，其包含包封于如本文所描述的衣壳(例如AAV分化体F衣壳)中的校正基因组。在一些实施例中，“校正基因组”是含有如本文所描述编辑元件以及足以在如本文所描述的衣壳内进行衣壳化的其它元件(例如，5’反向末端重复(5’ITR)核苷酸序列或其片段，以及3’反向末端重复(3’ITR)核苷酸序列或其片段)。应理解，术语“校正基因组”不一定需要含于所述校正基因组内的编辑元件将在整合到基因组中的目标基因座中时立刻“校正”所述目标基因座(例如，通过用野生型序列替换来校正含有突变的目标基因座)。因此，在一些实施例中，校正基因组可以含有编辑元件，所述编辑元件可以包含向目标基因座加成的核苷酸序列(例如目标基因座是第一开放阅读框的3’端，并且编辑元件是第二开放阅读框，所述第二开放阅读框在整合到目标基因座中时将形成编码融合蛋白的基因)。

[0099] 在一些实施例中，复制缺陷型腺相关病毒(AAV)包含校正基因组，所述校正基因组包含(a)编辑元件，所述编辑元件选自用于整合到哺乳动物染色体的目标基因座中的核苷酸间键或核苷酸序列，(b)位于所述编辑元件5’端的5’同源臂核苷酸序列，所述核苷酸序列与所述哺乳动物染色体中相对于所述目标基因座的5’区具有同源性，以及(c)位于所述编辑元件3’端的3’同源臂核苷酸序列，所述核苷酸序列与所述哺乳动物染色体中相对于所述目标基因座的3’区具有同源性。在一些实施例中，复制缺陷型AAV包含校正基因组，所述校正基因组包含用于整合到哺乳动物染色体的目标基因座中的编辑元件核苷酸序列，所述校正基因组基本不存在可操作地连接于所述编辑元件核苷酸序列的启动子。在一些实施例中，复制缺陷型AAV包含校正基因组，所述校正基因组包含编辑元件，所述编辑元件选自用于整合到细胞中哺乳动物染色体的目标基因座中的核苷酸间键或核苷酸序列；所述AAV用于将所述编辑元件整合到所述细胞中所述哺乳动物染色体的所述目标基因座中的染色体整合效率为至少约1％(例如，至少约2％、至少约5％、至少约10％、至少约20％、至少约30％、至少约40％、至少约50％、至少约60％、至少约70％、至少约80％或至少约90％)。在一些实施例中，复制缺陷型AAV包含校正基因组，所述校正基因组包含编辑元件，所述编辑元件选自用于整合到细胞中哺乳动物染色体的目标基因座中的核苷酸间键或核苷酸序列；所述AAV在无外源性核酸酶存在下用于将所述编辑元件整合到所述细胞中所述哺乳动物染色体的所述目标基因座中的染色体整合效率为至少约1％(例如，至少约2％、至少约5％、至少约10％、至少约20％、至少约30％、至少约40％、至少约50％、至少约60％、至少约70％、至少约80％或至少约90％)。在校正基因组中的任一个的一些实施例中，所述校正基因组基本不存在可操作地连接于编辑元件核苷酸序列的启动子。在校正基因组中的任一个的一些实施例中，所述校正基因组进一步包含可操作地连接于编辑元件的外源性启动子。在复制缺陷型AAV中的任一个的一些实施例中，所述AAV用于将编辑元件整合到细胞中哺乳动物染色体的目标基因座中的染色体整合效率为至少约1％、至少约2％、至少约3％、至少约4％、至少约5％、至少约10％、至少约20％、至少约30％、至少约40％、至少约50％、至少约60％、至少约70％、至少约80％或至少约90％。

[0100] 本发明的其它方面涉及一种包含复制缺陷型腺相关病毒(AAV)的基因编辑载体，所述AAV包含包封于AAV衣壳中的校正基因组，所述校正基因组如本文所描述(例如，包含编辑元件，所述编辑元件选自用于整合到哺乳动物细胞染色体的目标基因座中的核苷酸间键或核苷酸序列；位于所述编辑元件5'端的5'同源臂核苷酸序列，所述核苷酸序列与所述染色体中相对于所述目标基因座的5'区具有同源性；位于所述编辑元件3'端的3'同源臂核苷酸序列，所述核苷酸序列与所述染色体中相对于所述目标基因座的3'区具有同源性)；其中所述AAV用于将如本文所描述的编辑元件整合到如本文所描述的目标基因座中的染色体整合效率为至少10％(例如，至少15％、至少20％、至少30％、至少40％、至少50％、至少60％、至少70％、至少80％或至少90％)。在一些实施例中，用于在无外源性核酸酶存在下将如本文所描述的编辑元件整合到如本文所描述的目标基因座中的染色体整合效率是至少10％(例如，至少15％、至少20％、至少30％、至少40％、至少50％、至少60％、至少70％、至少80％或至少90％)。

[0101] 如本文所描述的校正可以包含位于5'同源臂核苷酸序列5'端的5'反向末端重复(5'ITR)和位于3'同源臂核苷酸序列3'端的3'反向末端重复(3'ITR)。在一些实施例中，5'ITR核苷酸序列和3'ITR核苷酸序列分别与AAV2病毒5'ITR和AAV2病毒3'ITR基本上一致(例如，至少90％、至少95％、至少98％、至少99％一致或100％一致)。在一些实施例中，5'ITR核苷酸序列与SEQ ID NO:36具有至少95％(例如，至少96％、至少97％、至少98％、至少99％或100％)序列一致性，并且3'ITR核苷酸序列与SEQ ID NO:37具有至少95％(例如，至少96％、至少97％、至少98％、至少99％或100％)序列一致性。在一些实施例中，5'ITR核苷酸序列和3'ITR核苷酸序列分别与AAV5病毒5'ITR和AAV5病毒3'ITR基本上一致(例如，至少90％、至少95％、至少98％、至少99％一致或100％一致)。在一些实施例中，5'ITR核苷酸序列与SEQ ID NO:38具有至少95％(例如，至少96％、至少97％、至少98％、至少99％或100％)序列一致性，并且3'ITR核苷酸序列与SEQ ID NO:39具有至少95％(例如，至少96％、至少97％、至少98％、至少99％或100％)序列一致性。在一些实施例中，5'ITR核苷酸序列和3'ITR核苷酸序列基本上是彼此d镜像(例如，除在5'或3'ITR中的1、2、3、4或5个核苷酸位置处以外，是彼此的镜像)。

[0102] 示例性AAV2 5'ITR(SEQ ID NO:36)-

[0103] ttggccactccctctctgcgcgctcgctcgctcactgaggccgggcgaccaaaggtcgcccgacgcccgggctttgcccgggcggcctcagtgagcgagcgagcgcgcagagagggagtggccaactccatcactaggggttcct

[0104] 示例性AAV2 3'ITR(SEQ ID NO:37)-

[0105] aggaacccctagtgatggagttggccactccctctctgcgcgctcgctcgctcactgaggccgggcgaccaaaggtcgcccgacgcccgggctttgcccgggcggcctcagtgagcgagcgagcgcgcagagagggagtggccaa

[0106] 示例性AAV5 5'ITR(SEQ ID NO:38)-

[0107] ctctcccccctgtcgcgttcgctcgctcgctggctcgtttggggggtggcagctcaaagagctgccagacgacggccctctggccgtcgcccccccaaacgagccagcgagcgagcgaacgcgacagggggagagtgccacactctcaagcaagggggttttgta

[0108] 示例性AAV5 3'ITR(SEQ ID NO:39)-

[0109] tacaaaacctccttgcttgagagtgtggcactctccccctgtcgcgttcgctcgctcgctggctcgtttggggggtggcagctcaaagagctgccagacgacggccctctggccgtcgcccccccaaacgagccagcgagcg

agcgaacgcgacagggggagag

[0110] 在一些实施例中，如本文所描述的校正基因组的大小不超过7kb(千碱基)，不超过6kb，不超过5kb或不超过4kb。在一些实施例中，如本文所描述的校正基因组在4kb与7kb、4kb与6kb、4kb与5kb或4.1kb与4.9kb之间。

[0111] 在某些实施例中，用于编辑细胞基因组的AAV分化体F载体或AAV载体变异体包含一个或多个分化体F衣壳或衣壳变异体(相对于AAV9衣壳的变异体)。在某些实施例中，用于编辑细胞基因组的AAV分化体F载体或AAV载体变异体包含一个或多个AAV分化体F衣壳。在某些实施例中，供体载体可以根据标准AAV包装方法包装到本文所描述的分化体F衣壳或衣壳变异体中，导致形成AAV分化体F载体或AAV载体变异体(参见例如Chatterjee,1992)。在某些实施例中，一个或多个分化体F衣壳或衣壳变异体影响AAV分化体F载体或AAV载体变异体对特定细胞的向性。

[0112] 根据某些实施例，一个或多个分化体F衣壳或衣壳变异体可以来源于人类干细胞源性AAV。先前已经展示，来自健康供体的细胞激素预致敏的周边血液$CD34^+$干细胞在其基因组中具有内源性天然AAV序列(参见例如美国专利公开第US20130096182A1号和第US20110294218A1号)。先前已经展现AAV分离株变异体(相对于AAV9的变异体)的功效，包括个别衣壳核苷酸和蛋白质用于细胞转导中的功效(参见例如美国专利公开第US20130096182A1号和第US20110294218A1号)。

[0113] 从在其基因组中具有内源性天然AAV序列的供体中分离全长AAV衣壳变异体基因(相对于AAV9的变异体)，并且对其进行测序。衣壳变异体的多核苷酸和多肽序列提供于图1中，并且提供于2012年11月2日提交的以美国专利公开第US20130096182A1号公开的美国专利申请第13/668,120号和2011年4月28日提交的美国专利申请第13/097,046号、以美国专利公开第US20130096182A1号公开的在2014年1月14日以美国专利第8,628,966号授权的US20110294218A1中，其全部以全文引用的方式并入本文中，如同在本文中充分阐述一般。在某些实施例中，本文所描述的AAV分化体F载体或AAV载体变异体可以包含一个或多个包含选自以下群组的多核苷酸序列的分化体F衣壳或衣壳变异体：AAVF1(SEQ ID NO:20)、AAVF2(SEQ ID NO:21)、AAVF3(SEQ ID NO:22)、AAVF4(SEQ ID NO:23)、AAVF5(SEQ ID NO:25)、AAVF11(SEQ ID NO:26)、AAVF7(SEQ ID NO:27)、AAVF8(SEQ ID NO:28)、AAVF9(SEQ ID NO:29)、AAVF12(SEQ ID NO:30)、AAVF13(SEQ ID NO:31)、AAVF14(SEQ ID NO:32)、AAVF15(SEQ ID NO:33)、AAVF16(SEQ ID NO:34)、AAVF17(SEQ ID NO:35)、其变异体、片段、突变体以及任何组合。在某些实施例中，本文所描述的AAV分化体F载体或AAV载体变异体可以包含一个或多个包含选自以下群组的多肽序列的分化体F衣壳或衣壳变异体：AAVF1(SEQ ID NO:2)、AAVF2(SEQ ID NO:3)、AAVF11(SEQ ID NO:4)、AAVF3(SEQ ID NO:5)、AAVF4(SEQ ID NO:6)、AAVF6(SEQ ID NO:7)、AAVF7(SEQ ID NO:8)、AAVF8(SEQ ID NO:9)、AAVF9(SEQ ID NO:10)、AAVF5(SEQ ID NO:11)、AAVF12(SEQ ID NO:12)、AAVF17(SEQ ID NO:13)、AAVF13(SEQ ID NO:14)、AAVF14(SEQ ID NO:15)、AAVF15(SEQ ID NO:16)、AAVF16(SEQ ID NO:17)、其变异体、片段、突变体以及任何组合(参见例如图1)。

[0114] 根据某些实施例，分化体F衣壳或衣壳变异体的多核苷酸或多肽序列可以与在本发明书中所传授的序列具有至少约95％、96％、97％，更优选地约98％，并且最优选地约99％序列一致性。一致性百分比可以使用多种序列比较程序或方法中的任一个来计算，如

Pearson和Lipman,《美国国家科学院院刊(Proc.Natl.Acad.Sci.USA)》,85:2444(1988),以及实施比较算法的程序,所述比较算法如GAP、BESTFIT、FASTA或TFASTA(来自Wisconsin Genetics Software Package,Genetics Computer Group,575Science Drive,Madison,Wis.)或可由美国国家生物技术信息中心(National Center for Biotechnology Information)网站获得的BLAST。

[0115] 分化体F衣壳或衣壳变异体序列可以在V1和/或V3cap基因中的一个或多个处进行修饰,这些基因或所述基因的功能部分可以在本文所描述的AAV分化体F载体或AAV载体变异体和方法中分开地或一起使用。可以从多个突变序列中取代Cap基因V1、V2和V3,并且其通常以共线方式V1-V2-V3使用。然而,序列可以是截短,如部分V1-V2-V3或V1-V3或V1-V1-V2-V3。举例来说,一个序列可以是(AAVF8)的V1-(AAVF4)的V2-AAVF14的V3。优选地,分化体F衣壳或衣壳变异体以等于或高于AAV2的水平转导目标细胞。

[0116] 在某些实施例中,一个或多个衣壳变异体可以包含衣壳变异体(例如SEQ ID NO:20-35,相对于AAV9衣壳的变异体)、AAV9衣壳(SEQ ID NO:18)、AAV2衣壳(SEQ ID NO:19)、其变异体、片段或突变体的一个或多个V1、V2和V3多核苷酸序列的组合。在某些实施例中,一个或多个或衣壳变异体可以包含衣壳变异体(SEQ ID NO:20-35,相对于AAV9衣壳的变异体)、任何其它已知AAV衣壳、其变异体、片段或突变体的一个或多个V1、V2和V3多核苷酸序列的组合。

[0117] 在某些实施例中,一个或多个分化体F衣壳或衣壳变异体可以包含衣壳变异体(SEQ ID NO:2-17,相对于AAV9衣壳的变异体)、AAV9衣壳(SEQ ID NO:1)、其变异体、片段或突变体的V1、V2和V3多肽序列的组合。在某些实施例中,一个或多个衣壳变异体可以包含分化体F衣壳变异体(SEQ ID NO:2-17,相对于AAV9的变异体)、任何其它已知AAV衣壳、其变异体、片段或突变体的V1、V2和V3多肽序列的组合。

[0118] 在一些实施例中,用于编辑细胞基因组的AAV分化体F载体或AAV载体变异体包含AAV分化体F衣壳。在一些实施例中,“AAV分化体F衣壳”是指其AAV VP1、VP2和/或VP3序列分别与AAV9的AAV VP1、VP2和/或VP3序列具有至少86%(例如,至少86%、至少87%、至少88%、至少89%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%或至少99%)序列一致性的衣壳。示例性分化体F衣壳包括AAVF1-17(在本文中也称为AAVHSC1-17)、AAV9、AAVHU31、AAVHU32和AAVAnc110(参见例如Zinn等人病毒进化谱系的计算机模拟重构产生强效基因疗法载体(In Silico Reconstruction of the Viral Evolutionary Lineage Yields a Potent Gene Therapy Vector)(2015)《细胞报告(Cell Reports)》,第12卷,第1056-1068页)。

[0119] 在一些实施例中,AAV分化体F衣壳包含至少一个或至少两个选自分化体F VP1、分化体F VP2和分化体F VP3的蛋白质。在一些实施例中,AAV分化体F衣壳包含分化体F VP1、分化体F VP2和分化体F VP3蛋白质。

[0120] AAV分化体F衣壳的示例性AAV VP1、VP2和VP3蛋白质序列提供于下表中。

[0121]

AAV 衣壳	VP1	VP2	VP3
AAV9	SEQ ID NO: 1 的氨基酸 1 到 736	SEQ ID NO: 1 的氨基酸 138 到 736	SEQ ID NO: 1 的氨基酸 203 到 736
AAVF1	SEQ ID NO: 2 的氨基酸 1 到 736	SEQ ID NO: 2 的氨基酸 138 到 736	SEQ ID NO: 2 的氨基酸 203 到 736
AAVF2	SEQ ID NO: 3 的氨基酸 1 到 736	SEQ ID NO: 3 的氨基酸 138 到 736	SEQ ID NO: 3 的氨基酸 203 到 736
AAVF3	SEQ ID NO: 5 的氨基酸 1 到 736	SEQ ID NO: 5 的氨基酸 138 到 736	SEQ ID NO: 5 的氨基酸 203 到 736
AAVF4	SEQ ID NO: 6 的氨基酸 1 到 736	SEQ ID NO: 6 的氨基酸 138 到 736	SEQ ID NO: 6 的氨基酸 203 到 736
AAVF5	SEQ ID NO: 11 的氨基酸 1 到 736	SEQ ID NO: 11 的氨基酸 138 到 736	SEQ ID NO: 11 的氨基酸 203 到 736
AAVF6	SEQ ID NO: 7 的氨基酸 1 到 736	SEQ ID NO: 7 的氨基酸 138 到 736	SEQ ID NO: 7 的氨基酸 203 到 736
AAVF7	SEQ ID NO: 8 的氨基酸 1 到 736	SEQ ID NO: 8 的氨基酸 138 到 736	SEQ ID NO: 8 的氨基酸 203 到 736
AAVF8	SEQ ID NO: 9 的氨基酸 1 到 736	SEQ ID NO: 9 的氨基酸 138 到 736	SEQ ID NO: 9 的氨基酸 203 到 736
AAVF9	SEQ ID NO: 10 的氨基酸 1 到 736	SEQ ID NO: 10 的氨基酸 138 到 736	SEQ ID NO: 10 的氨基酸 203 到 736
AAVF10	SEQ ID NO: 3 的氨基酸 1 到 736	SEQ ID NO: 3 的氨基酸 138 到 736	SEQ ID NO: 3 的氨基酸 203 到 736
AAVF11	SEQ ID NO: 4 的氨基酸 1 到 736	SEQ ID NO: 4 的氨基酸 138 到 736	SEQ ID NO: 4 的氨基酸 203 到 736
AAVF12	SEQ ID NO: 12 的氨基酸 1 到 736	SEQ ID NO: 12 的氨基酸 138 到 736	SEQ ID NO: 12 的氨基酸 203 到 736
AAVF13	SEQ ID NO: 14 的氨基酸 1 到 736	SEQ ID NO: 14 的氨基酸 138 到 736	SEQ ID NO: 14 的氨基酸 203 到 736

[0122]

AAVF14	SEQ ID NO: 15 的氨基酸 1 到 736	SEQ ID NO: 15 的氨基酸 138 到 736	SEQ ID NO: 15 的氨基酸 203 到 736
AAVF15	SEQ ID NO: 16 的氨基酸 1 到 736	SEQ ID NO: 16 的氨基酸 138 到 736	SEQ ID NO: 16 的氨基酸 203 到 736
AAVF16	SEQ ID NO: 17 的氨基酸 1 到 736	SEQ ID NO: 17 的氨基酸 138 到 736	SEQ ID NO: 17 的氨基酸 203 到 736
AAVF17	SEQ ID NO: 13 的氨基酸 1 到 736	SEQ ID NO: 13 的氨基酸 138 到 736	SEQ ID NO: 13 的氨基酸 203 到 736

[0123] 在一些实施例中，AAV分化体F衣壳包含分别与SEQ ID NO:1的氨基酸1到736、氨基酸138到736或氨基酸203到736具有至少85％(例如，至少86％、至少88％、至少90％、至少92％、至少94％、至少96％、至少97％、至少98％、至少99％或100％)氨基酸序列一致性的VP1、VP2或VP3蛋白质，所述氨基酸分别对应于AAV9衣壳蛋白VP1、VP2和VP3的氨基酸序列。在一些实施例中，AAV分化体F衣壳包含分别与SEQ ID NO:1的氨基酸1到736和氨基酸138到736具有至少85％(例如，至少86％、至少88％、至少90％、至少92％、至少94％、至少96％、至少97％、至少98％、至少99％或100％)氨基酸序列一致性的VP1和VP2蛋白质，所述氨基酸分别对应于AAV9衣壳蛋白VP1和VP2的氨基酸序列；分别与SEQ ID NO:1的氨基酸1到736和氨基酸203到736具有至少85％(例如，至少86％、至少88％、至少90％、至少92％、至少94％、至少96％、至少97％、至少98％、至少99％或100％)氨基酸序列一致性的VP1和VP3蛋白质，所

述氨基酸分别对应于AAV9衣壳蛋白VP1和VP3的氨基酸序列;或分别与SEQ ID NO:1的氨基酸138到736和氨基酸203到736具有至少85%(例如,至少86%、至少88%、至少90%、至少92%、至少94%、至少96%、至少97%、至少98%、至少99%或100%)氨基酸序列一致性的VP2和VP3蛋白质,所述氨基酸分别对应于AAV9衣壳蛋白VP2和VP3的氨基酸序列。在一些实施例中,AAV分化体F衣壳包含分别与SEQ ID NO:1的氨基酸1到736、氨基酸138到736和氨基酸203到736具有至少85%(例如,至少86%、至少88%、至少90%、至少92%、至少94%、至少96%、至少97%、至少98%、至少99%或100%)氨基酸序列一致性的VP1、VP2和VP3蛋白质,所述氨基酸分别对应于AAV9衣壳蛋白VP1、VP2和VP3的氨基酸序列。

[0124] 在一些实施例中,AAV分化体F衣壳包含分别与SEQ ID NO:2、3、5、6、11、7、8、9、10、4、12、14、15、16、17或13中任一个的氨基酸1到736、氨基酸138到736或氨基酸203到736具有至少85%(例如,至少86%、至少88%、至少90%、至少92%、至少94%、至少96%、至少97%、至少98%、至少99%或100%)氨基酸序列一致性的VP1、VP2或VP3蛋白质,所述氨基酸分别对应于AAVF1到AAVF9和AAVF11到AAVF17衣壳蛋白VP1、VP2和VP3的氨基酸序列。在一些实施例中,AAV分化体F衣壳包含分别与SEQ ID NO:2、3、5、6、11、7、8、9、10、4、12、14、15、16、17或13中任一个的氨基酸1到736和氨基酸138到736具有至少85%(例如,至少86%、至少88%、至少90%、至少92%、至少94%、至少96%、至少97%、至少98%、至少99%或100%)氨基酸序列一致性的VP1和VP2蛋白质,所述氨基酸分别对应于AAVF1到AAVF9和AAVF11到AAVF17衣壳蛋白VP1和VP2的氨基酸序列;分别与SEQ ID NO:2、3、5、6、11、7、8、9、10、4、12、14、15、16、17或13中任一个的氨基酸1到736和氨基酸203到736具有至少85%(例如,至少86%、至少88%、至少90%、至少92%、至少94%、至少96%、至少97%、至少98%、至少99%或100%)氨基酸序列一致性的VP1和VP3蛋白质,所述氨基酸分别对应于AAVF1到AAVF9和AAVF11到AAVF17衣壳蛋白VP1和VP3的氨基酸序列;或分别与SEQ ID NO:2、3、5、6、11、7、8、9、10、4、12、14、15、16、17或13中任一个的氨基酸138到736或氨基酸203到736具有至少85%(例如,至少86%、至少88%、至少90%、至少92%、至少94%、至少96%、至少97%、至少98%、至少99%或100%)氨基酸序列一致性的VP2和VP3蛋白质,所述氨基酸分别对应于AAVF1到AAVF9和AAVF11到AAVF17衣壳蛋白VP2和VP3的氨基酸序列。在一些实施例中,AAV分化体F衣壳包含分别与SEQ ID NO:2、3、5、6、11、7、8、9、10、4、12、14、15、16、17或13中任一个的氨基酸1到736、氨基酸138到736和氨基酸203到736具有至少85%(例如,至少86%、至少88%、至少90%、至少92%、至少94%、至少96%、至少97%、至少98%、至少99%或100%)氨基酸序列一致性的VP1、VP2和VP3蛋白质,所述氨基酸分别对应于AAVF1到AAVF9和AAVF11到AAVF17衣壳蛋白VP1、VP2和VP3的氨基酸序列。

[0125] 在一些实施例中,AAV分化体F衣壳包含由与SEQ ID NO:18包含至少85%(例如,至少86%、至少88%、至少90%、至少92%、至少94%、至少96%、至少97%、至少98%、至少99%或100%)核苷酸序列一致性的核苷酸序列编码的VP1、VP2或VP3蛋白质,所述核苷酸序列分别对应于编码AAV9衣壳蛋白VP1、VP2和VP3的核苷酸序列。在一些实施例中,AAV分化体F衣壳包含由与SEQ ID NO:18包含至少85%(例如,至少86%、至少88%、至少90%、至少92%、至少94%、至少96%、至少97%、至少98%、至少99%或100%)核苷酸序列一致性的核苷酸序列编码的VP1和VP2蛋白质,所述核苷酸序列对应于编码AAV9衣壳蛋白VP1、VP2和VP3的核苷酸序列;由与SEQ ID NO:18包含至少85%(例如,至少86%、至少88%、至少90%、至

少92％、至少94％、至少96％、至少97％、至少98％、至少99％或100％)核苷酸序列一致性的核苷酸序列编码的VP1和VP3蛋白质;或由与SEQ ID NO:18包含至少85％(例如,至少86％、至少88％、至少90％、至少92％、至少94％、至少96％、至少97％、至少98％、至少99％或100％)核苷酸序列一致性的核苷酸序列编码的VP2和VP3蛋白质。在一些实施例中,AAV分化体F衣壳包含由与SEQ ID NO:18包含至少85％(例如,至少86％、至少88％、至少90％、至少92％、至少94％、至少96％、至少97％、至少98％、至少99％或100％)核苷酸序列一致性的核苷酸序列编码的VP1、VP2和VP3蛋白质,所述核苷酸序列对应于编码AAV9衣壳蛋白VP1、VP2和VP3的核苷酸序列。

[0126] 在一些实施例中,AAV分化体F衣壳包含由与SEQ ID NO:20、21、22、23、25、24、27、28、29、26、30、31、32、33、34或35中的任一个包含至少85％(例如,至少86％、至少88％、至少90％、至少92％、至少94％、至少96％、至少97％、至少98％、至少99％或100％)核苷酸序列一致性的核苷酸序列编码的VP1、VP2或VP3蛋白质,所述核苷酸序列分别对应于编码AAVF1到AAVF17衣壳蛋白VP1、VP2和VP3的核苷酸序列。在一些实施例中,AAV分化体F衣壳包含由与SEQ ID NO:20-35中的任一个包含至少85％(例如,至少86％、至少88％、至少90％、至少92％、至少94％、至少96％、至少97％、至少98％、至少99％或100％)核苷酸序列一致性的核苷酸序列编码的VP1和VP2蛋白质;由与SEQ ID NO:20-35中的任一个包含至少85％(例如,至少86％、至少88％、至少90％、至少92％、至少94％、至少96％、至少97％、至少98％、至少99％或100％)核苷酸序列一致性的核苷酸序列编码的VP1和VP3蛋白质;或由与SEQ ID NO:20-35中的任一个包含至少85％(例如,至少86％、至少88％、至少90％、至少92％、至少94％、至少96％、至少97％、至少98％、至少99％或100％)核苷酸序列一致性的核苷酸序列编码的VP2和VP3蛋白质。在一些实施例中,AAV分化体F衣壳包含由与SEQ ID NO:20、21、22、23、25、24、27、28、29、26、30、31、32、33、34或35中的任一个包含至少85％(例如,至少86％、至少88％、至少90％、至少92％、至少94％、至少96％、至少97％、至少98％、至少99％或100％)核苷酸序列一致性的核苷酸序列编码的VP1、VP2和VP3蛋白质,所述核苷酸序列分别对应于编码AAVF1到AAVF17衣壳蛋白VP1、VP2和VP3的核苷酸序列。

[0127] 在一些实施例中,AAV分化体F衣壳包含AAV9VP1、VP2或VP3衣壳蛋白,其分别对应于如SEQ ID NO:1中所阐述的氨基酸1到736、氨基酸138到736和氨基酸203到736。在一些实施例中,AAV分化体F衣壳包含AAV9VP1和VP2衣壳蛋白,所述衣壳蛋白分别对应于如SEQ ID NO:1中所阐述的氨基酸1到736和氨基酸138到736;AAV9VP1和VP3衣壳蛋白,所述衣壳蛋白分别对应于如SEQ ID NO:1中所阐述的氨基酸1到736和氨基酸203到736;或AAV9VP2和VP3衣壳蛋白,所述衣壳蛋白分别对应于如SEQ ID NO:1中所阐述的氨基酸138到736和氨基酸203到736。在一些实施例中,AAV分化体F衣壳包含AAV9VP1、VP2和VP3衣壳蛋白,其分别对应于如SEQ ID NO:1中所阐述的氨基酸1到736、氨基酸138到736和氨基酸203到736。

[0128] 在一些实施例中,AAV分化体F衣壳包含选自AAVF1到AAVF9和AAVF11到AAVF17中任一个的VP1衣壳蛋白的VP1衣壳蛋白,所述衣壳蛋白分别对应于如SEQ ID NO:2、3、5、6、11、7、8、9、10、4、12、14、15、16、17或13中所阐述的氨基酸1到736。在一些实施例中,AAV分化体F衣壳包含独立地选自AAVF1到AAVF9和AAVF11到AAVF17中任一个的VP1和VP2衣壳蛋白的VP1和VP2衣壳蛋白,所述衣壳蛋白分别对应于如SEQ ID NO:2、3、5、6、11、7、8、9、10、4、12、14、15、16、17或13中所阐述的氨基酸1到736和氨基酸138到736。在一些实施例中,AAV分化体F

衣壳包含独立地选自AAVF1到AAVF9和AAVF11到AAVF17中任一个的VP2和VP3衣壳蛋白的VP2和VP3衣壳蛋白，所述衣壳蛋白分别对应于如SEQ ID NO:2、3、5、6、11、7、8、9、10、4、12、14、15、16、17或13中所阐述的氨基酸138到736和氨基酸203到736。在一些实施例中，AAV分化体F衣壳包含AAVF1到AAVF9和AAVF11到AAVF17中任一个的VP1、VP2和VP3衣壳蛋白中的每一个，所述衣壳蛋白分别对应于如SEQ ID NO:2、3、5、6、11、7、8、9、10、4、12、14、15、16、17或13中所阐述的氨基酸1到736、氨基酸138到736和氨基酸203到736。

[0129] 如本文所用，多核苷酸序列的片段可以是编码提供与由全长多核苷酸序列编码的多肽基本上相同功能的多肽的多核苷酸部分。如本文所用，多核苷酸序列的突变体可以通过对特定多核苷酸序列进行一个或多个核苷酸的缺失、取代、添加和/或插入来获得。应理解，多核苷酸序列的此类片段和/或突变体编码具有与由全长多核苷酸序列编码的多肽基本上相同功能的多肽。

[0130] 如本文所用，多肽序列可以包括多肽序列的片段和/或突变体，而仍提供与全长多肽序列基本上相同的功能。多肽序列的片段意味着提供与全长多肽序列基本上相同功能的多肽序列部分。多肽序列突变体的实例包括对多肽序列进行的一个或多个氨基酸的缺失、取代、添加和/或插入。

[0131] 在某些实施例中，多核苷酸序列可以是重组或非天然存在的多核苷酸。在某些实施例中，多核苷酸序列可以是cDNA。

[0132] 在某些实施例中，本文所提供的AAV分化体F载体或AAV载体变异体可以包含本文所描述的AAVF(或AAVHSC)或任何其它AAV分化体F载体中的任一个。在某些实施例中，AAV分化体F载体或AAV载体变异体可以包含本文中所描述的AAVF(或AAVHSC)载体中的任一个，如AAVF1、AAVF2、AAVF3、AAVF4、AAVF5、AAVF6、AAVF7、AAVF8、AAVF9、AAVF10、AAVF11、AAVF12、AAVF13、AAVF14、AAVF15、AAVF16、AAVF17、其变异体、片段、突变体或任何组合。在某些实施例中，分化体F载体或AAV载体变异体可以包含AAV9、AAVF1、AAVF2、AAVF3、AAVF4、AAVF5、AAVF6、AAVF7、AAVF8、AAVF9、AAVF10、AAVF11、AAVF12、AAVF13、AAVF14、AAVF15、AAVF16、AAVF17、AAVHU31、AAVHU32、其变异体、片段、突变体或任何组合中的任一个。

[0133] 在某些实施例中，本文所提供的AAV分化体F载体或AAV载体变异体可以包含编辑元件(在本文中也称为靶向盒，意味着术语“编辑元件”和“靶向盒”在本文中可互换地使用)，所述编辑元件包含待整合到基因组的目标基因座(在本文中也称为目标位点，意味着所述术语在本文中可互换地使用)中的一个或多个核苷酸序列或核苷酸间键；位于所述编辑元件5’端的5’同源臂核苷酸序列，所述核苷酸序列与哺乳动物染色体中相对于所述目标基因座的5’区具有同源性(例如，侧接所述编辑元件(靶向盒)并且与位于所述目标基因座(目标位点)上游的区具有同源性的5’同源臂多核苷酸序列)，以及位于所述编辑元件3’端的3’同源臂核苷酸序列，所述核苷酸序列与所述哺乳动物染色体中相对于所述目标基因座的3’区具有同源性(例如，侧接所述编辑元件(靶向盒)并且与位于所述目标基因座(目标位点)下游的区具有同源性的3’同源臂多核苷酸序列)。

[0134] 在某些实施例中，待整合到基因组的目标位点中的一个或多个核苷酸序列可以是一个或多个治疗性核苷酸序列。如本文所用的术语“治疗性”是指引起对疾病或病症的治疗的物质或方法。“治疗性核苷酸序列”是提供治疗效果的核苷酸序列。治疗效果可以是直接的(例如，取代表达为蛋白质的基因的核酸或将cDNA插入到用于表达的内含子中)或间接的

(例如,校正调节元件,如启动子)。在某些实施例中,治疗性核苷酸序列可以包括一个或多个核苷酸。在某些实施例中,治疗性核苷酸序列可以是基因、其变异体、片段或突变体。在某些实施例中,当需要基因疗法时,治疗性核苷酸序列可以是编码治疗学上有效的蛋白质(包括治疗性抗体)的任何核苷酸序列。包含治疗性核苷酸序列的分化体F载体或AAV载体优选地以治疗有效量经由合适的投药途径,如注射、吸入、吸收、摄取或其它方法来投与。

[0135] 在一些实施例中,如本文所描述的编辑元件由一个核苷酸组成。在一些实施例中,如本文所描述的编辑元件由一个核苷酸组成,并且如本文所描述的目标基因座是由一个核苷酸组成的核苷酸序列,所述目标基因座表示点突变。在一些实施例中,如本文所描述的编辑元件包含至少1、2、10、100、200、500、1000、1500、2000、3000、4000或5000个核苷酸。在一些实施例中,如本文所描述的编辑元件包含以下或由以下组成:1到5500、1到5000、1到4500、1到4000、1到3000、1到2000、1到1000、1到500、1到200或1到100个核苷酸,或2到5500、2到5000、2到4500、2到4000、2到3000、2到2000、2到1000、2到500、2到200或2到100个核苷酸,或10到5500、10到5000、10到4500、10到4000、10到3000、10到2000、10到1000、10到500、10到200或10到100个核苷酸。在一些实施例中,如本文所描述的编辑元件包含以下或由以下组成:外显子、内含子、5'非翻译区(UTR)、3'UTR、启动子、剪接供体、剪接受体、编码或非编码RNA的序列、绝缘子、基因或其组合。在一些实施例中,如本文所描述的编辑元件是处于如本文所描述的目标基因座内或跨所述目标基因座的基因的编码序列的片段(例如不超过2kb、不超过1kb、不超过500bp、不超过250bp、不超过100bp、不超过50bp或不超过25bp)。在一些实施例中,如本文所描述的编辑元件是核苷酸间键(例如,连接两个相邻核苷酸的磷酸二酯键)。在一些实施例中,如本文所描述的编辑元件是核苷酸间键,如本文所描述的染色体中的目标基因座是包含一个或多个核苷酸的核苷酸序列,并且编辑元件包含对染色体中的目标基因座进行的缺失。

[0136] 在某些实施例中,AAV分化体F载体或AAV载体变异体的编辑元件(或靶向盒)可以包含一个或多个调节元件多核苷酸序列。举例来说,在某些实施例中,一个或多个调节元件多核苷酸序列可以选自2A序列、剪接受体序列、聚腺苷酸化序列以及其任何组合。在某些实施例中,靶向盒可以包含一个或多个侧接5'和3'同源臂多核苷酸序列的AAV反向末端重复(ITR)多核苷酸序列。在某些实施例中,编辑元件(或靶向盒)不含有驱动一个或多个核苷酸序列表达的启动子。在某些实施例中,如果编辑元件(或靶向盒)不含有启动子,那么一个或多个核苷酸序列在整合到细胞基因组中之后的表达可以由细胞的一个或多个调节元件控制。在某些实施例中,无启动子的一个或多个核苷酸序列的表达展现,所述一个或多个核苷酸序列正确地整合到细胞中。

[0137] 在某些实施例中,AAV分化体F载体或AAV载体变异体可以包含一个或多个同源臂多核苷酸序列。在某些实施例中,一个或多个同源臂多核苷酸序列可以与基因组的目标基因座(目标位点)区同源。在某些实施例中,一个或多个同源臂多核苷酸序列可以是5'同源臂多核苷酸序列。在某些实施例中,5'同源臂多核苷酸序列可以侧接编辑元件(或靶向盒)的5'端。在某些实施例中,侧接编辑元件(或靶向盒)的5'同源臂多核苷酸序列可以与位于基因组目标基因座(目标位点)上游的区同源。在某些实施例中,一个或多个同源臂多核苷酸序列可以是3'同源臂多核苷酸序列。在某些实施例中,3'同源臂多核苷酸序列可以侧接编辑元件(或靶向盒)的3'端。在某些实施例中,侧接编辑元件(或靶向盒)的3'同源臂多核

苷酸序列可以与位于基因组目标基因座(目标位点)下游的区同源。在某些实施例中,同源多核苷酸序列可以是大致500到1,000个核苷酸长。举例来说,在某些实施例中,同源臂多核苷酸序列可以是大致800个核苷酸长。在某些实施例中,同源臂多核苷酸序列可以是至多大致3,000个核苷酸长。在一些实施例中,5'和3'同源臂核苷酸序列中的每一个的核苷酸长度独立地介于约50到2000个核苷酸之间,如介于约500-1000、约600-1000或约700-900个核苷酸之间。在一些实施例中,5'和3'同源臂核苷酸序列中的每一个的核苷酸长度独立地介于约600、约800或约1000个核苷酸之间。

[0138] 在一些实施例中,5'和3'同源臂核苷酸序列具有基本上相等的核苷酸长度。在一些实施例中,5'和3'同源臂核苷酸序列具有不对称的核苷酸长度。在一些实施例中,核苷酸长度的不对称性由以下界定:介于5'与3'同源臂核苷酸序列长度之间的差异为长度的至多50%,如长度的至多40%、30%、20%或10%差异。在一些实施例中,核苷酸长度的不对称性由以下界定:5'和3'同源臂的一个臂的长度为约600个核苷酸,并且5'和3'同源臂的另一个臂的长度为约800或约900个核苷酸。

[0139] 在一些实施例中,5'同源臂核苷酸序列与哺乳动物染色体中相对于目标基因座的5'区具有至少约90%(例如,至少约90%、至少约95%、至少约96%、至少约97%、至少约98%、至少约99%或至少约99.5%)核苷酸序列一致性。在一些实施例中,3'同源臂核苷酸序列与哺乳动物染色体中相对于目标基因座的3'区具有至少约90%(例如,至少约90%、至少约95%、至少约96%、至少约97%、至少约98%、至少约99%或至少约99.5%)核苷酸序列一致性。在一些实施例中,5'同源臂或3'同源臂分别与相对应的哺乳动物染色体5'区或3'区的核苷酸序列差异可以包含核苷酸序列的非编码差异,基本上由其组成或由其组成。在一些实施例中,5'同源臂或3'同源臂分别与相对应的哺乳动物染色体5'区或3'区的核苷酸序列差异可以包含引起保守性氨基酸变化的核苷酸序列差异(例如,碱性氨基酸变为不同碱性氨基酸),基本上由其组成或由其组成。在一些实施例中,5'同源臂核苷酸序列与哺乳动物染色体中相对于目标基因座的5'区具有100%一致性,并且3'同源臂核苷酸序列与哺乳动物染色体中相对于目标基因座的3'区具有100%一致性。在一些实施例中,即使目标基因座与5'或3'同源臂相比较含有一个或多个突变,如一个或多个天然存在的SNP,但仍将5'同源臂核苷酸序列和3'同源臂核苷酸序列视为分别与哺乳动物染色体中相对于目标基因座的5'区和3'区同源。

[0140] 在某些实施例中,细胞基因组的目标基因座(目标位点)可以是所述基因组中需要进行细胞基因组编辑的任何区。举例来说,细胞基因组的目标基因座(目标位点)可以包含细胞中染色体的基因座(例如哺乳动物染色体的区)。在某些实施例中,染色体的基因座可以是安全港位点。安全港位点是基因组中核苷酸序列可以整合并且以可预测方式起作用而不扰乱内源性基因活性的位置。在某些实施例中,安全港位点可以是人类染色体19中的AAVS1基因座(也称为PPP1R12C基因座)。在某些实施例中,安全港位点可以是染色体19中AAVS1基因座中PPP1R12C的第一内含子。先前展示染色体19上的AAVS1基因座qter13.3-13.4为用于插入转基因的"安全港"位点,这是因为此处插入的基因进行表达而无病原性后果,其与在此基因座处整合而无病原性后果的野生型AAV类似(Giraud 1994;Linden,1996A;Linden 1996B)。在一些实施例中,目标基因座(目标位点)是与如本文所描述的疾病病况相关联的基因座。

[0141] 在某些实施例中，目标基因座是哺乳动物染色体中的相对于相对应野生型哺乳动物染色体包含一个或多个突变型核苷酸的突变型目标基因座。在一些实施例中，突变型目标基因座包含点突变、错义突变、无义突变、一个或多个核苷酸的插入、一个或多个核苷酸的缺失，或其组合。在一些实施例中，突变型目标基因座包含无效突变、新效突变或反效突变。在一些实施例中，突变型目标基因座包含常染色体显性突变、常染色体隐性突变、异型接合子突变、同型接合子突变，或其组合。在本文所描述的突变型目标基因座中的任一个的一些实施例中，突变型目标基因座选自哺乳动物染色体内的启动子、强化子、信号序列、内含子、外显子、剪接供体位点、剪接受体位点、内部核糖体进入位点、反向外显子、绝缘子、基因、染色体反转和染色体易位。

[0142] 在一些实施例中，如本文所描述的染色体中的目标基因座是包含n个核苷酸的核苷酸序列，其中n是大于或等于1的整数(例如1、2、3、4、5、10、20、30、40、50、100、500、1000、2000、3000、4000、5000或其间的任何整数)，如本文所描述的编辑元件包含m个核苷酸，其中m是等于n的整数，并且所述编辑元件表示对染色体目标基因座的取代。在一些实施例中，如本文所描述的染色体中的目标基因座是包含n个核苷酸的核苷酸序列，其中n是大于或等于1的整数(例如1、2、3、4、5、10、20、30、40、50、100、500、1000、2000、3000、4000、5000或其间的任何整数)，如本文所描述的编辑元件包含m个核苷酸，其中m是大于n的整数，并且所述编辑元件表示对染色体目标基因座的取代性添加。在一些实施例中，如本文所描述的染色体中的目标基因座是包含n个核苷酸的核苷酸序列，其中n是大于或等于2的整数(例如2、3、4、5、10、20、30、40、50、100、500、1000、2000、3000、4000、5000或其间的任何整数)，如本文所描述的编辑元件包含m个核苷酸，其中m是小于n的整数；并且所述编辑元件表示对染色体目标基因座的取代性缺失。在一些实施例中，如本文所描述的染色体中的目标基因座是核苷酸间键，如本文所描述的编辑元件包含m个核苷酸，其中m是大于或等于1的整数(例如1、2、3、4、5、10、20、30、40、50、100、500、1000、2000、3000、4000、5000或其间的任何整数)；并且所述编辑元件表示对染色体目标基因座的添加。

[0143] 在一些实施例中，染色体中的目标基因座是哺乳动物染色体(例如人类、小鼠、牛、马、犬、猫、大鼠或兔染色体)中的目标基因座。在一些实施例中，目标基因座可以包含哺乳动物染色体的内含子。在一些实施例中，目标基因座可以包含哺乳动物染色体的外显子。在一些实施例中，目标基因座可以包含哺乳动物染色体的非编码区。在一些实施例中，目标基因座可以包含哺乳动物染色体的调节区。在一些实施例中，哺乳动物染色体选自人类染色体1、2、3、4、5、6、7、8、9、10,11、12、13、14、15、16、17、18、19、20、21、22、X以及Y。在一些实施例中，哺乳动物染色体选自小鼠染色体1、2、3、4、5、6、7、8、9、10,11、12、13、14、15、16、17、18、19、X以及Y。在一些实施例中，哺乳动物染色体不是人类染色体19。在一些实施例中，哺乳动物染色体是体细胞染色体。示例性体细胞进一步描述于本文中。

[0144] 在某些实施例中，一个或多个核苷酸序列或编辑元件可以在于整合之前不需要DNA裂解的情况下经由同源重组整合到基因组中。在某些实施例中，一个或多个核苷酸序列或编辑元件可以在不需要添加外源性核酸酶的情况下经由同源重组整合到基因组中，所述外源性核酸酶如锌指核酸酶(ZFN)、转录激活子样效应核酸酶(TALEN)或RNA引导核酸酶(CRISPR/Cas)。

[0145] 在某些实施例中，由本文所描述的AAV分化体F载体或AAV载体变异体编辑的细胞

可以是任何类型的细胞。在某些实施例中，细胞可以是多种哺乳动物细胞，例如肝脏、肺、软骨和其它结缔组织、眼睛、中枢和外周神经系统、淋巴系统、骨骼、肌肉、血液、大脑、皮肤、心脏和消化道的细胞。当待由AAV分化体F载体或AAV载体变异体编辑的细胞是例如肝脏细胞时，所插入的核苷酸序列是有关治疗病状(改善或治愈病症或停止疾病或病症的进一步进程)或预防病状。当待由AAV分化体F载体或AAV载体变异体编辑的细胞是肝脏细胞时，所治疗或预防的肝脏病状包含血友病、酶递送、肝硬化、癌症或动脉粥样硬化以及其它肝脏病状。在某些实施例中，细胞可以是体细胞(例如哺乳动物体细胞)。在某些实施例中，细胞(例如，体细胞，如哺乳动物体细胞)可以来自选自由以下组成的群组的组织：结缔组织(包括血液)、肌肉组织、神经组织和上皮组织。在某些实施例中，细胞(例如，体细胞，如哺乳动物体细胞)可以来自选自由以下组成的群组的器官：肺、心脏、肝脏、肾脏、肌肉、大脑、眼睛、乳房、骨骼和软骨。在一些实施例中，细胞是CD34+细胞(例如，CD34+体细胞)。在一些实施例中，细胞(例如，体细胞，如哺乳动物体细胞)是肝脏细胞、成纤维细胞、乳房细胞、淋巴细胞或视网膜细胞。

[0146] 如本文所示，用本文所描述的分化体F衣壳或衣壳变异体包装的AAV展现对某些目标组织的特异性向性，所述目标组织如血液干细胞、肝脏、心脏、眼睛、乳房和关节组织，并且可以用于转导干细胞以用于将所关注的基因引入到目标组织中。所述载体中的某些能够穿过紧密受控的生物学接合点(如血-脑屏障)，其开创载体的其它新颖用途和目标器官，提供经由基因组编辑进行基因疗法的方法。因此，分化体F载体或AAV载体变异体可以基于其分化体F衣壳或衣壳变异体而展现对特定细胞的向性。举例来说，a)对于肌肉组织或细胞，AAV分化体F载体或AAV载体变异体可以选自以下群组：AAVF5、AAVF7、AAVF13、AAVF15和AAVF17；b)对于心脏或肺组织或细胞，载体可以选自以下群组：AAVF13、AAVF15和AAVF17；c)对于肝脏或CNS组织或细胞，载体可以选自AAVF5、AAVF13、AAVF17、AAVF7或AAVF15；d)对于干细胞，载体可以是AAVF17；e)对于B细胞祖细胞，载体可以是AAVF5；f)对于骨髓和红细胞祖细胞，载体可以是AAVF12；并且g)对于淋巴结、肾脏、脾脏、软骨和骨骼组织或细胞，载体可以选自选自以下群组的载体的群组：AAVF7、AAVF13、AAVF15和AAVF17。

[0147] 另外，分化体F载体或AAV载体变异体可以对含有各种标签的细胞具有向性，如六His标签或亲和力标签；或对干扰素反应具有向性，如响应于病原体或肿瘤细胞而投与的天然存在的抗体引起或引入的单株抗体。

[0148] 在某些实施例中，细胞可以是干细胞(例如哺乳动物干细胞)。在某些实施例中，干细胞可以是任何类型的干细胞，包括造血干细胞、多能干细胞、胚胎干细胞或间叶细胞干细胞。在某些实施例中，干细胞(例如哺乳动物干细胞)可以是造血干细胞、脐血干细胞、骨髓干细胞、胎儿肝脏干细胞或周边血液干细胞。在一些实施例中，干细胞可以是CD34+干细胞。在某些实施例中，干细胞(例如哺乳动物干细胞)可以是造血干细胞或周边血液干细胞。干细胞的转导可以是瞬时或永久的(也称为持续性的)。如果是瞬时的，那么一个实施例允许使用或表达治疗性核苷酸的时间长度受到载体、受到附接于载体的物质或受到外部因素或外力控制。

[0149] 在某些实施例中，细胞可以选自由以下组成的群组：CD34+造血干细胞系(HSC)、K562 CD34+白血病细胞系、HepG2人类肝脏细胞系、周边血液干细胞、脐血干细胞、CD34+周边血液干细胞、WI-38人类二倍体成纤维细胞细胞系、MCF7人类乳癌细胞系、Y79人类成视网

膜细胞瘤细胞系、SCID-X1 LBL人类EBV永生化B细胞系、原代人类肝细胞、原代肝窦状隙内皮细胞以及原代骨骼肌成肌细胞。

[0150] 本文还提供离体编辑个体细胞基因组的方法，其包含用如本文所描述的分化体F或AAV载体变异体转导所述细胞。在某些实施例中，用分化体F载体或AAV载体变异体转导细胞可以在没有其它外源性核酸酶的情况下进行，所述外源性核酸酶如锌指核酸酶(ZFN)、转录激活子样效应核酸酶(TALEN)或RNA引导核酸酶(CRISPR/Cas)。在某些实施例中，细胞可以是任何类型的细胞。在某些实施例中，细胞可以是如本文所描述的干细胞。举例来说，在某些实施例中，编辑干细胞基因组的方法可以包含用一种或多种分化体F载体或AAV载体变异体转导所述干细胞。在某些实施例中，干细胞的转导可以在不需要其它外源性核酸酶的情况下进行。在某些实施例中，细胞可以是如本文所描述的体细胞。举例来说，在某些实施例中，编辑体细胞基因组的方法可以包含用一种或多种分化体F载体或AAV载体变异体转导所述体细胞。在某些实施例中，体细胞的转导可以在不需要其它外源性核酸酶的情况下进行。在某些实施例中，分化体F载体或AAV载体变异体包含一个或多个分化体F衣壳或衣壳变异体(相对于AAV9的变异体)、选自用于整合到哺乳动物染色体的目标基因座中的核苷酸间键或核苷酸序列或包含一个或多个待整合到基因组的目标基因座(目标位点)中的治疗性核苷酸序列的编辑元件(靶向盒)、侧接所述编辑元件(靶向盒)并且与位于所述目标基因座(目标位点)上游的区具有同源性的5'同源臂多核苷酸序列以及侧接所述编辑元件(靶向盒)并且与位于所述目标基因座(目标位点)下游的区具有同源性的3'同源臂多核苷酸序列。在某些实施例中，核苷酸间键或核苷酸序列或一个或多个治疗性核苷酸序列可以在于整合之前不需要用于DNA裂解的其它外源性核酸酶的情况下整合到基因组中。

[0151] 本文还提供在个体中治疗疾病或病症的方法，其通过离体编辑所述个体的细胞的基因组来进行，包括用分化体F载体或AAV载体变异体转导所述细胞和将经过转导的细胞进一步移植到所述个体中以治疗所述疾病或病症。在某些实施例中，所述方法可以包含用本文所描述的分化体F载体或AAV载体变异体转导个体细胞。在某些实施例中，细胞可以在没有其它外源性核酸酶的情况下进行转导。在某些实施例中，用分化体F载体或AAV载体变异体转导细胞可以如本文所提供来进行或通过所属领域的普通技术人员已知的任何转导方法来进行。在某些实施例中，细胞可以用分化体F载体或AAV载体变异体以50,000;100,000;150,000;200,000;250,000;300,000;350,000;400,000;450,000;或500,000的感染复数(MOI)或以提供最佳细胞转导的任何MOI加以转导。在某些实施例中，将经过转导的细胞进一步移植到个体中，其中所述经过转导的细胞治疗疾病或病症。在某些实施例中，细胞可以是本文所描述的任何类型的细胞。

[0152] 本文还提供如本文所描述编辑哺乳动物基因组的目标基因座的方法。在一些实施例中，所述方法包含用如本文所描述的AAV(例如，包含包封于衣壳中的校正基因组的复制缺陷型AAV)转导包含哺乳动物基因组的细胞(如人类、小鼠、牛、马、犬、猫、大鼠或兔细胞)。在一些实施例中，所述方法包含(a)从哺乳动物(如人类、小鼠、牛、马、犬、猫、大鼠或兔)获得哺乳动物细胞；(b)离体培养所述哺乳动物细胞以形成离体培养物；(c)在所述离体培养物中用如本文所描述的AAV(例如，包含包封于衣壳中的校正基因组的复制缺陷型AAV)转导所述哺乳动物细胞以形成经过转导的哺乳动物细胞；以及(d)向所述哺乳动物投与所述经过转导的哺乳动物细胞。在一些实施例中，所述方法包含(a)从第一哺乳动物获得哺乳动物

细胞；(b)离体培养所述哺乳动物细胞以形成离体培养物；(c)在所述离体培养物中用如本文所描述的AAV(例如，包含包封于衣壳中的校正基因组的复制缺陷型AAV)转导所述哺乳动物细胞以形成经过转导的哺乳动物细胞；以及(d)向第二哺乳动物投与所述经过转导的哺乳动物细胞。在一些实施例中，第一哺乳动物和第二哺乳动物是不同物种(例如，第一哺乳动物是人类、小鼠、牛、马、犬、猫、大鼠或兔，并且第二哺乳动物是不同物种)。在一些实施例中，第一哺乳动物和第二哺乳动物是相同物种(例如，两者都是人类、小鼠、牛、马、犬、猫、大鼠或兔)。在一些实施例中，所述方法包含以有效用AAV体内转导哺乳动物细胞的量向哺乳动物(如人类、小鼠、牛、马、犬、猫、大鼠或兔)投与如本文所描述的AAV(例如，包含包封于衣壳中的校正基因组的复制缺陷型AAV)。

[0153] 在所述方法中的任一个的一些实施例中，哺乳动物细胞来自选自由以下组成的群组的组织：结缔组织(包括血液)、肌肉组织、神经组织和上皮组织。在所述方法中的任一个的一些实施例中，哺乳动物细胞来自选自由以下组成的群组的器官：肺、心脏、肝脏、肾脏、肌肉、大脑、眼睛、乳房、骨骼和软骨。在所述方法中的任一个的一些实施例中，哺乳动物细胞是干细胞。在一些实施例中，干细胞是造血干细胞或周边血液干细胞。在所述方法中的任一个的一些实施例中，哺乳动物细胞是CD34+细胞。

[0154] 在所述方法中的任一个的一些实施例中，在不共转导或共投与外源性核酸酶或编码外源性核酸酶的核苷酸序列的情况下转导或投与AAV(例如分化体F AAV)。示例性外源性核酸酶包括锌指核酸酶(ZFN)、转录激活子样效应核酸酶(TALEN)或RNA引导核酸酶(CRISPR/Cas)。在所述方法中的任一个的一些实施例中，在不共转导或共投与外源性锌指核酸酶或编码外源性锌指核酸酶的核苷酸序列的情况下转导或投与AAV。在一些实施例中，锌指核酸酶是包含靶向AAVS1基因座的DNA结合域(例如靶向AAVS1基因座中PPP1R12C第一内含子的DNA结合域)的锌指核酸酶。

[0155] 在所述方法中的任一个的一些实施例中，AAV(例如分化体F AAV)用于将编辑元件整合到哺乳动物染色体的目标基因座中的染色体整合效率为至少约1％(例如，至少约2％、至少约3％、至少约4％、至少约5％、至少约10％、至少约20％、至少约30％、至少约40％、至少约50％、至少约60％、至少约70％、至少约80％、至少约90％、至少约95％、至少约98％或约100％)。在所述方法中的任一个的一些实施例中，AAV(例如分化体F AAV)用于在无外源性核酸酶存在下将编辑元件整合到哺乳动物染色体的目标基因座中的染色体整合效率为至少约1％(例如，至少约2％、至少约3％、至少约4％、至少约5％、至少约10％、至少约20％、至少约30％、至少约40％、至少约50％、至少约60％、至少约70％、至少约80％、至少约90％、至少约95％、至少约98％或约100％)。在所述方法中的任一个的一些实施例中，校正基因组的编辑元件以介于哺乳动物细胞10％到70％、20％到70％、40％到70％、50％到70％、10％到80％、20％到80％、40％到80％、50％到80％、10％到90％、20％到90％、40％到90％、50％到90％、10％到100％、20％到100％、40％到100％或50％到100％范围内的染色体整合效率整合到哺乳动物染色体的目标基因座中。在所述方法中的任一个的一些实施例中，校正基因组的编辑元件在无外源性核酸酶存在下以介于哺乳动物细胞10％到70％、20％到70％、40％到70％、50％到70％、10％到80％、20％到80％、40％到80％、50％到80％、10％到90％、20％到90％、40％到90％、50％到90％、10％到100％、20％到100％、40％到100％或50％到100％范围内的染色体整合效率整合到哺乳动物染色体的目标基因座中。

[0156] 在所述方法中的任一个的一些实施例中，AAV(例如分化体F AAV)的染色体整合效率的另一个特征为，对于包含整合到哺乳动物染色体的目标基因座中的编辑元件的等位基因，细胞群体中的等位基因出现频率为至少约10％(例如，至少约20％、至少约30％、至少约40％、至少约50％、至少约60％、至少约75％、至少约85％、至少约90％或至少约95％)。在一些实施例中，细胞群体中的等位基因出现频率是体外细胞群体中的等位基因出现频率，如体外的本文所提供的细胞类型群体(例如，CD34+造血干细胞系(HSC)、K562CD34+白血病细胞系、HepG2人类肝脏细胞系、周边血液干细胞、脐血干细胞、CD34+周边血液干细胞、WI-38人类二倍体成纤维细胞细胞系、MCF7人类乳癌细胞系、Y79人类成视网膜细胞瘤细胞系、SCID-X1LBL人类EBV永生化B细胞系、原代人类肝细胞、原代肝窦状隙内皮细胞以及原代骨骼肌成肌细胞)中的等位基因出现频率。

[0157] 根据某些实施例，提供在个体中治疗疾病或病症的方法，其通过以下步骤来进行：离体编辑所述个体的干细胞的基因组和将经过编辑的细胞进一步移植到所述个体中以治疗所述疾病或病症。在某些实施例中，在个体中通过编辑所述个体的干细胞的基因组来治疗疾病或病症的方法可以包含以下步骤：用如本文所描述的AAV分化体F载体或AAV载体变异体转导所述个体的所述干细胞和将经过转导的干细胞移植到所述个体中，其中所述经过转导的干细胞治疗所述疾病或病症。在某些实施例中，AAV分化体F载体或AAV载体变异体可以包含一个或多个分化体F衣壳或衣壳变异体、包含一个或多个待整合到干细胞基因组的目标基因座(目标位点)中的编辑元件(靶向盒)、侧接所述编辑元件(靶向盒)并且与位于所述目标基因座(目标位点)上游的区具有同源性的5'同源臂多核苷酸序列以及侧接所述编辑元件(靶向盒)并且与位于所述目标基因座(目标位点)下游的区具有同源性的3'同源臂多核苷酸序列。在某些实施例中，转导干细胞可以在没有其它外源性核酸酶的情况下进行。在某些实施例中，一个或多个治疗性核苷酸序列可以在于整合之前不需要用于DNA裂解的其它外源性核酸酶的情况下整合到基因组中。

[0158] 在某些实施例中，当细胞是干细胞时，所治疗的疾病或病症可以是由基因组的一个或多个突变造成的任何疾病或病症。在某些实施例中，所治疗的疾病或病症选自遗传性代谢病、溶酶体贮积病、粘多糖病、免疫缺陷疾病以及血红蛋白病和感染。在某些实施例中，当待编辑的细胞是干细胞时，AAV分化体F载体或AAV载体变异体可以选自以下群组：AAVF7AAVF12、AAVF15、AAVF17、其变异体、突变体以及组合。在某些实施例中，当待编辑的细胞是干细胞时，分化体F载体或AAV载体变异体可以选自以下群组：AAVF5、AAVF7、AAVF12、AAVF15、AAVF17、其变异体、突变体以及组合。在某些实施例中，分化体F载体或AAV载体变异体可以包含一个或多个包含选自以下群组的多核苷酸序列的分化体F衣壳或衣壳变异体：AAVF7(SEQ ID NO:27)、AAVF12(SEQ ID NO:30)、AAVF15(SEQ ID NO:33)、AAVF17(SEQ ID NO:35)、其变异体、片段、突变体以及组合。在某些实施例中，分化体F载体或AAV载体变异体可以包含一个或多个包含选自以下群组的多核苷酸序列的分化体F衣壳或衣壳变异体：AAVF5(SEQ ID NO:25)、AAVF7(SEQ ID NO:27)、AAVF12(SEQ ID NO:30)、AAVF15(SEQ ID NO:33)、AAVF17(SEQ ID NO:35)、其变异体、片段、突变体以及组合。在某些实施例中，AAV分化体F载体或AAV载体变异体可以包含一个或多个包含选自以下群组的多肽序列的分化体F衣壳或衣壳变异体：AAVF7(SEQ ID NO:8)、AAVF12(SEQ ID NO:12)、AAVF15(SEQ ID NO:16)、AAVF17(SEQ ID NO:13)、其变异体、片段、突变体以及组合。在某些实施例中，AAV分化

体F载体或AAV载体变异体可以包含一个或多个包含选自以下群组的多肽序列的分化体F衣壳或衣壳变异体：AAVF5(SEQ ID NO:11)、AAVF7(SEQ ID NO:8)、AAVF12(SEQ ID NO:12)、AAVF15(SEQ ID NO:16)、AAVF17(SEQ ID NO:13)、其变异体、片段、突变体以及组合。

[0159] 在另一个实施例中，来自$CD34^+$HSC或来自另一个来源的能够进行基因组编辑的AAV分化体F载体或AAV载体变异体可以用于高效转导干细胞(包括HSC和iPSC)以及其它细胞，如心脏、关节、中枢神经系统(包括大脑)、肌肉和肝脏的那些细胞。如果体外使用AAV分化体F载体或AAV载体变异体，那么其可以用于研究和调查目的或用以制备将稍后植入到个体中的细胞或组织。优选地，个体是哺乳动物，如人类，但可以是其组织可以由本发明载体和使用那些载体的方法转导的任何其它动物。本发明AAV分化体F载体或AAV载体变异体极适合于人用和兽用两者。AAV分化体F载体或AAV载体变异体也可以体外用于瞬时转导干细胞，如HSC。转导时长可以受培养条件控制。如果体内使用AAV分化体F载体或AAV载体变异体，那么其可以向接受疗法的个体直接投与以用于在目标细胞(如肝脏或软骨细胞)中吸收或使用。如果AAV分化体F载体或AAV载体变异体用于转导中枢神经系统的细胞，那么其优选地能够穿越血-脑屏障并且维持其功效。

[0160] 本文还提供在个体中通过对所述个体的细胞进行体内基因组编辑来治疗疾病或病症的方法，其通过向所述个体直接投与AAV分化体F载体或AAV载体变异体来进行。在某些实施例中，AAV分化体F载体或AAV载体变异体可以是本文所描述的任何AAV分化体F载体或AAV载体变异体。在某些实施例中，AAV分化体F载体或AAV载体变异体可以包含一个或多个分化体F衣壳或衣壳变异体、包含一个或多个待整合到干细胞基因组的目标基因座(目标位点)中的编辑元件(靶向盒)、侧接所述编辑元件(靶向盒)并且与位于所述目标基因座(目标位点)上游的区具有同源性的5'同源臂多核苷酸序列以及侧接所述编辑元件(靶向盒)并且与位于所述目标基因座(目标位点)下游的区具有同源性的3'同源臂多核苷酸序列。在某些实施例中，所投与的AAV分化体F载体或AAV载体变异体通过对个体的细胞进行基因组编辑来治疗疾病或病症。在某些实施例中，体内基因组编辑可以在没有其它外源性核酸酶的情况下进行。在某些实施例中，一个或多个分化体F衣壳或衣壳变异体包含如本文所提供的多核苷酸或多肽序列。在某些实施例中，多核苷酸或多肽序列可以选自美国专利公开第20130096182A1号的图1或本文的图1中所提供的序列、其变异体、片段、突变体以及组合。在某些实施例中，AAV分化体F载体或AAV载体变异体优选地以治疗有效量经由合适的投药途径，如注射、吸入、吸收、摄取或其它方法来投与。

[0161] 包括Xu等人，Wang等人和Carbonaro等人的先前研究已经展示在体内递送病毒载体后的HSC转导(参见Xu 2004;Wang 2014;和Carbonaro 2006)。然而，这些研究中的所有三个都涉及逆转录病毒(Xu 2004)或慢病毒(Wang 2014和Carbonaro 2006)。此外，Xu等人和Carbonaro等人中的注射在新生小鼠中进行，并且在Wang等人中需要雷帕霉素(rapamycin)和股骨内注射以用于高效转导。然而，这些论文都未报告通过将分化体F载体或AAV载体变异体体内转导到成年小鼠中来转导HSC。实例3中所提供的新颖结果初次展示通过静脉内注射对HSC进行的AAV载体转导。

[0162] 如下文实例3中所示，静脉内注射用AAVF7或AAVF17假型化的分化体F载体(或AAV载体变异体)引起人类CD34+造血干细胞和祖细胞的体内转导。静脉内注射的分化体F载体或AAV载体运输到人类造血位点和经过转导的人类细胞处。静脉内注射分化体F载体或AAV

载体变异体在人类CD34+干祖细胞以及其CD45+子代中引起Venus表达。这些数据展示，静脉内注射分化体F载体或AAV载体变异体可以用于在不需要干细胞采集、离体转导、接受者适应和后续移植经过转导的细胞的情况下进行体内基因组工程改造。此途径使干细胞基因疗法显著地更安全，世界范围的患者更可获得，更便宜，并且避免住院需要。

[0163] 在某些实施例中，公开在个体中通过对所述个体的细胞进行体内基因组编辑来治疗疾病或病症的方法，其通过向所述个体直接投与AAV分化体F载体或AAV载体变异体来进行。在某些实施例中，AAV分化体F载体或AAV载体变异体可以包含一个或多个分化体F衣壳或衣壳变异体、包含一个或多个待整合到基因组的目标基因座(目标位点)中的治疗性核苷酸序列的编辑元件(靶向盒)、侧接所述编辑元件(靶向盒)并且与位于所述目标基因座(目标位点)上游的区具有同源性的5'同源臂多核苷酸序列以及侧接所述编辑元件(靶向盒)并且与位于所述目标基因座(目标位点)下游的区具有同源性的3'同源臂多核苷酸序列，其中所述载体转导个体的细胞并且将所述一个或多个治疗性核苷酸序列整合到所述细胞的基因组中。在某些实施例中，一个或多个分化体F衣壳或衣壳变异体可以包含选自以下群组的多肽序列：AAVF1(SEQ ID NO:2)、AAVF2(SEQ ID NO:3)、AAVF11(SEQ ID NO:4)、AAVF3(SEQ ID NO:5)、AAVF4(SEQ ID NO:6)、AAVF6(SEQ ID NO:7)、AAVF7(SEQ ID NO:8)、AAVF8(SEQ ID NO:9)、AAVF9(SEQ ID NO:10)、AAVF5(SEQ ID NO:11)、AAVF12(SEQ ID NO:12)、AAVF17(SEQ ID NO:13)、AAVF13(SEQ ID NO:14)、AAVF14(SEQ ID NO:15)、AAVF15(SEQ ID NO:16)、AAVF16(SEQ ID NO:17)、其变异体、片段、突变体以及任何组合。在某些实施例中，一个或多个分化体F衣壳或衣壳变异体可以包含AAVF7(SEQ ID NO:8)或AAVF17(SEQ ID NO:13)的多肽序列。在某些实施例中，一个或多个分化体F衣壳或衣壳变异体可以包含AAVF5(SEQ ID NO:11)、AAVF(SEQ ID NO:8)或AAVF17(SEQ ID NO:13)的多肽序列。在某些实施例中，AAV分化体F载体或AAV载体变异体不含有用于一个或多个治疗性核苷酸序列的启动子。在某些实施例中，目标基因座(目标位点)可以是安全港位点。在某些实施例中，安全港位点可以是染色体19上的AAVS1基因座。在某些实施例中，细胞可以是干细胞。在某些实施例中，干细胞可以是造血干细胞、多能干细胞、胚胎干细胞或间叶细胞干细胞。在某些实施例中，疾病或病症可以由细胞基因组中的一个或多个突变造成。在某些实施例中，疾病或病症可以选自遗传性代谢病、溶酶体贮积病、粘多糖病、免疫缺陷疾病以及血红蛋白病和感染。

[0164] 进一步展现体内应用的功效，用分离株变异体(相对于AAV9)向免疫缺陷型小鼠中移植经过转导的细胞产生长期和持续的转基因表达，并且可以用于基因疗法。在某些实施例中，当全身递送时，这些载体显示对肝脏和软骨的向性，并且暗示遗传性、获得性、感染性和肿瘤学疾病的疗法。关于肝脏转导，本发明AAV分离株的肝脏转导水平是目前向肝脏进行全身基因递送的最高准则AAV8的高达大致10倍。此特性可以用于基于基因的由肝脏的酶替换疗法以用于如血友病、酶缺陷疾病和动脉粥样硬化的疾病。本发明AAV分离株对关节中软骨组织的额外向性可以用于治疗骨骼病症，如关节炎、骨质疏松或其它基于软骨/骨骼的疾病。因此，在不需要长期整合的情况下，变异体序列和方法可以用于瞬时转导。

[0165] AAV分化体F衣壳家族或AAV衣壳变异体家族的成员转导HSC(例如AAVF 15和AAVF 17)，从而产生具有持续性基因表达的长期植入，并且因此是用于干细胞基因疗法载体的强候选物。AAVF17和AAVF15(也以缩写形式称为“HSC17”和“HSC15”)在移植后高达22周时仍支持最高水平的长期体内转导。在静脉内注射AAV变异体后的连续生物发光成像揭示，AAVF15

一般支持最高水平的体内长期转基因表达。包括AAVF13和17的其它AAV变异体也支持强体内转导。

[0166] 发现AAVF15具有高度肝脏向性，约为AAV9的5-10倍。AAVF13和AAVF15对心脏和骨骼肌的转导也是AAV9的至少10倍。体外中和效价揭示，在合并的人类IVIG中针对AAVF 1-9衣壳的抗体的普遍性与AAV9类似，而针对AAVF13、AAVF15、AAVF16和AAVF17的抗体在一定程度上较不普遍。体内中和分析确认，在用AAVF15进行IVIG投药后，与AAV9相比较，在肝脏和肌肉中发现超过100的载体基因组复本/细胞，表明预先存在的抗体不完全中和AAVF15。肌肉疾病或病症可以包含具有疾病或病症的任何细胞、含有肌肉细胞的组织、器官或系统，包括心脏，如冠心病或心肌病。

[0167] 另外，位点特异性突变诱发实验指示，AAVF15中的R505G突变造成肝脏向性增强。AAV分化体F载体或AAV载体变异体可以用于治疗遗传疾病的整个宿主，所述遗传疾病如血友病、动脉粥样硬化和多种先天性代谢错误。在一个例子中，AAVF15有效治疗血友病B。此家族的一些成员也在全身注射之后靶向关节，其可以用于治疗关节和软骨疾病，如关节炎。家族的其它成员在静脉内注射后靶向心脏。家族的又其它成员靶向大脑。在一些实施例中，由于展示AAVF5转导多种细胞类型，提供包含AAVF5衣壳蛋白的载体作为本文所提供的方法、试剂盒或组合物的一部分(参见图4)。

[0168] 在某些实施例中，在个体中通过基因组编辑治疗神经疾病或病症的方法可以包含投与能够穿过血-脑屏障、血-眼屏障或血-神经屏障的AAV分化体F载体或AAV载体变异体。本文所公开的AAV分化体F载体或AAV载体变异体中的某些具有以下独特能力：使用经过修饰的病毒载体穿越先前不知道可由用于基因疗法或其它诊断或治疗目的的任何载体接近的生物学接合点。这些接合点具有共同特征。血-脑屏障是在身体中循环的血液与中枢神经系统中的脑细胞外液之间的隔离，并且由围绕毛细管的紧密接合点产生。血-脑屏障一般仅允许较小疏水性分子扩散通过。血-脑屏障是在局部血管与眼睛的大部分之间进行的隔离，并且由视网膜和虹膜的毛细管的内皮产生。血-神经屏障是其中轴突、许旺细胞(Schwann cell)和外周神经的其它相关细胞起作用的生理学空间，并且由神经束和膜化神经束膜(investing perineurium)内的神经内膜微血管产生。如同这些屏障中的三个一般，渗透性受限制以在内部环境(此处为神经)中保护免受血管和其它细胞外空间中的剧烈浓度变化影响。穿越这些屏障中任一个的载体具有以下独特能力：递送一个或多个治疗性核苷酸序列以用于治疗神经疾病或病症或充当标记剂和或诊断剂。已经以实验方式验证尤其适合于穿过这些生物学屏障的AAV分化体F载体或AAV载体变异体中的某些包括AAVF15、AAVF15A346T和AAVF15R505G。

[0169] 存在所属领域的技术人员熟知的许多神经疾病或病症，其可以一般分类为细胞型或器官型，如大脑、脊髓、神经节、运动神经、感觉神经、自主神经、视神经、视网膜神经和听觉神经的疾病或病症。借助于实例，大脑疾病或病症可以包括癌症或其它大脑肿瘤、炎症、细菌感染、病毒感染(包括狂犬病)、阿米巴原虫或寄生虫感染、中风、麻痹、神经退化性病症(如阿兹海默病(Alzheimer's Disease)、帕金森病(Parkinson's Disease)或其它痴呆或认知功能降低)、斑块、脑病、亨廷顿病(Huntington's Disease)、动脉瘤、遗传性或获得性畸形、获得性大脑损伤、妥瑞综合症(Tourette Syndrome)、发作性睡病、肌肉营养不良、震颤、大脑性麻痹、孤独症、唐氏综合症(Down Syndrome)、注意力缺陷和注意缺陷多动障碍、

慢性炎症、癫痫症、昏迷、脑膜炎、多发性硬化症、重症肌无力、各种神经病、不宁腿综合症以及泰伊-萨克斯二氏病(Tay-Sachs disease)。

[0170] 仅借助于实例,肌肉疾病或病症包括肌病、慢性疲乏综合症、肌肉纤维疼痛、肌肉营养不良、多发性硬化症、萎缩症、抽筋、痉挛、僵硬、各种炎症(如皮肌炎)、横纹肌溶解症、肌筋膜疼痛综合症、肿胀、隔室综合症、嗜伊红血球增多症-肌痛综合症、粒线体肌病、肌紧张性病症、麻痹、肌腱炎、风湿病性多肌痛、癌症以及肌腱病症(如肌腱炎和腱鞘炎)。

[0171] 仅借助于实例心脏疾病或病症包括冠状动脉疾病、冠心病、充血性心脏衰竭、心肌病、心肌炎、心包疾病、先天性心脏病、癌症、心内膜炎以及瓣膜疾病。

[0172] 仅借助于实例,肺疾病或病症包括哮喘、过敏、慢性阻塞性肺病、支气管炎、肺气肿、囊肿性纤维化、肺炎、肺结核、肺部水肿、癌症、急性呼吸窘迫综合症、尘肺病以及间质性肺病。

[0173] 仅借助于实例,肝脏疾病或病症包括癌症,A型、B型和C型肝炎,肝硬化,黄疸以及肝病。仅借助于实例,肾脏疾病或病症包括癌症、糖尿病、肾病综合症、肾结石、获得性肾病、先天性疾病、多囊性肾病、肾炎、原发性高草酸盐尿症以及胱氨酸尿症。仅借助于实例,脾脏疾病或病症包括癌症、脾梗塞、结节病以及高雪氏病(Gaucher's disease)。仅借助于实例,骨骼疾病或病症包括骨质疏松、癌症、低骨密度、佩吉特氏病(Paget's disease)以及感染。

[0174] 对于使用由或用AAV分化体F载体或AAV载体变异体输送的治疗性核苷酸序列或甚至小分子治疗的这些疾病或病症中的任一个,借助于实例,所述治疗性核苷酸序列可以是编码蛋白质治疗剂(如对于癌症,凋亡蛋白质)的核酸、miRNA、、shRNA、siRNA、其它RNA亚型或其组合。在一些实施例中,载体如本文所描述进行分离和纯化。分离和纯化是增加功效和减少污染的优选体内投药方式。载体可以永久或瞬时转导转基因,所述转基因是已经从一个生物体中分离并且引入到另一个中的基因或其它基因材料。此处,其它生物体可以是接受载体的个体。

[0175] 在某些实施例中,用于基因组编辑的AAV分化体F载体或AAV载体变异体可以基于在既定目标细胞或组织中对于如本文所示的既定疾病或病症具有最高功效的实验结果来选择。举例来说,a)对于肌肉疾病或病症和对于经由肌肉向个体投与的抗体基因或其它疫苗治疗,AAV分化体F载体或AAV载体变异体选自以下群组:AAVF5、AAVF7、AAVF13、AAVF15和AAVF17;b)对于心脏和肺疾病或病症,载体选自以下群组:AAVF13、AAVF15和AAVF17;c)对于肝脏或神经疾病或病症,载体选自AAVF5和AAVF15;d)对于通过植入干细胞来治疗的病状,载体AAVF17;e)对于通过转导B细胞祖细胞来治疗的病状,载体AAVF5;f)对于通过转导骨髓和红细胞祖细胞来治疗的病状,载体AAVF12;并且g)对于淋巴结、肾脏、脾脏、软骨和骨骼疾病或病症,载体选自选自以下群组的载体的群组:AAVF7、AAVF13、AAVF15和AAVF17;其中所述AAV分化体F载体或AAV载体变异体转导细胞或组织,并且一个或多个治疗性核苷酸序列整合到所述细胞的基因组中并且治疗所述疾病或病症。在某些实施例中,AAV分化体F载体或AAV载体变异体可以包含一个或多个对如本文所描述的细胞展现向性的分化体F衣壳或衣壳变异体(相对于AAV9)。

[0176] 个体是所述方法对其起作用的任何动物,但优选地是哺乳动物,所述哺乳动物可以是人类。如果载体含有抗体基因或其它疫苗治疗,那么其可以经由在肌肉中注射来投与,并且可以提供针对疾病的免疫预防,包括HIV、流感、疟疾、破伤风、麻疹、腮腺炎、风疹、HPV、

百日咳或任何其它疫苗。载体可以经过包装、分离和纯化，并且可以用至少一个治疗性核苷酸序列转导任何类型的干细胞。载体也可以将转基因转导或将内源性校正基因运载到个体和/或相同物种的其它个体中。

[0177] “AAV”是腺相关病毒。除非另外指明，否则所述术语可以用于指病毒或其衍生物、病毒亚型以及天然存在和重组的形式。AAV具有超过100种不同亚型，其被称作AAV-1、AAV-2等，并且包括人类和非人类源性AAV两者。存在约十二种AAV血清型。各种亚型的AAV可以用作重组基因转移病毒以转导许多不同细胞类型。

[0178] 如应用于多核苷酸的“重组”意味着所述多核苷酸是克隆、限制和/或接合步骤以及其它产生与天然存在的多核苷酸不同的构筑体的程序的多种组合的产物。重组病毒包含重组多核苷酸的病毒粒子，包括原始多核苷酸构筑体的复制品和原始病毒构筑体的子代。“rAAV载体”是指包含非AAV来源的多核苷酸序列(即，与AAV异源多核苷酸)的重组AAV载体，其通常是用于细胞遗传转化的所关注的序列。

[0179] 如本文所用的用于AAV的“辅助病毒”是允许AAV由哺乳动物细胞复制和包装的病毒。用于AAV的辅助病毒是所属领域中已知的，并且包括例如腺病毒(如子组C的5型腺病毒)、疱疹病毒(如单纯性疱疹病毒、艾伯斯坦-巴尔病毒(Epstein-Bar virus)和巨细胞病毒)以及痘病毒。

[0180] “关节组织”包含多种组织，包括软骨、滑液以及产生或本身为以下细胞的成熟细胞、祖细胞和干细胞：(i)软骨产生细胞；(ii)I型滑膜细胞；(iii)II型滑膜细胞；(iv)定居(resident)或循环白血球；(v)成纤维细胞；(vi)血管内皮细胞；以及(vii)周细胞。

[0181] “有复制能力的(replication-competent)”病毒是指具有感染性并且能够在受感染的细胞中复制的病毒。在AAV的情况下，复制能力一般需要存在功能性AAV包装基因以及辅助病毒基因，如腺病毒和单纯性疱疹病毒。一般来说，rAAV载体是没有复制能力的(在本文中也称为复制缺陷的)，这是因为其缺乏一个或多个AAV包装基因。在一些实施例中，如果AAV基本不存在AAV rep基因和/或AAV cap基因，那么AAV可以视为复制缺陷(或没有复制能力的)。在一些实施例中，如果AAV缺乏AAV rep基因和/或AAV cap基因，那么AAV可以视为复制缺陷(或没有复制能力的)。在一些实施例中，包含AAV分化体F载体或AAV变异体分离株的组合物是无细胞组合物。组合物一般不含细胞蛋白质和/或其它污染物，并且可以包含其它要素，如缓冲液(例如磷酸盐缓冲液、Tris缓冲液)、盐(例如NaCl、MgCl2)、离子(例如镁离子、锰离子、锌离子)、防腐剂、增溶剂或清洁剂(例如非离子清洁剂；二甲亚砜)。

[0182] 在另一个实施例中，表达盒包含编码包含分化体F衣壳或AAV变异体分离株中的一或多个的多肽的多核苷酸序列，其中编码所述多肽的所述多核苷酸序列包含与在本发明书中所传授的序列具有至少约95％、96％、97％，更优选地约98％，并且最优选地约99％序列一致性的序列。一致性百分比可以使用多种序列比较程序或方法中的任一个来计算，如Pearson和Lipman，《美国国家科学院院刊》，85:2444(1988)，以及实施比较算法的程序，所述比较算法如GAP、BESTFIT、FASTA或TFASTA(来自Wisconsin Genetics Software Package，Genetics Computer Group，575Science Drive，Madison，Wis.)或可由美国国家生物技术信息中心(National Center for Biotechnology Information)网站获得的BLAST。

[0183] 在另一方面，表达盒包含编码包含分化体F衣壳或AAV变异体分离株中的一或多个

的多肽的多核苷酸序列，其中所述序列包含有包含衣壳蛋白的三个基因的部分V1-V3(也被称作VP1-VP3)。举例来说，所述盒可以包含来自衣壳AAVF1的V1、与AAV9hu.14相比的标准V2以及来自AAVF17衣壳的V3。在又另一个实施例中，衣壳可以包含超过一个的衣壳基因组分中的每一个。举例来说，分化体F衣壳或衣壳变异体可以选自用于本文所阐述衣壳序列的VP1-VP3(V1-V3)中的任一个，并且可以以任何次序和任何组合合并，只要实现所需的转导增加特性即可。举例来说衣壳序列可以是VP1A-VP1B-VP2-VP3(V1A-V1B-V2-V3)、P3-VP1-VP2(V3-V1-V2)或VP1-VP2-VP3A-VP3B(V1-V2-V3A-V3B)。

[0184] 另一个实施例包括对个体进行免疫接种的方法。可以用引起在个体中产生免疫性的免疫反应的方式将包含分化体F衣壳或衣壳变异体的组合物引入到所述个体中。分化体F衣壳或衣壳变异体可以单独或作为表达盒的一部分处于组合物中。在一个实施例中，可以使用基因递送载体来引入表达盒(或多核苷酸)。基因递送载体可以例如是非病毒载体或病毒载体。示例性病毒载体包括(但不限于)辛德毕斯病毒(Sindbis-virus)源性载体、反转录病毒载体和慢病毒载体。适用于产生免疫反应的组合物也可以使用颗粒状载剂来递送。此外，此类组合物可以涂布在例如金或钨粒子上，并且使用例如基因枪将经过涂布的粒子递送到个体中。组合物也可以调配为脂质体。在此方法的一个实施例中，个体是哺乳动物，并且可以例如是人类。

[0185] 术语“亲和力标签”在本文中用于指示可以附接于第二多肽以提供对所述第二多肽的纯化或检测或提供用于将第二多肽附接于底物的位点的多肽区段。原则上，可获得对其使用的抗体或其它特异性结合剂的任何肽或蛋白质可以用作亲和力标签。亲和力标签包括多组氨酸束蛋白A(Nilsson等人，EMBO J.4:1075,1985;Nilsson等人，《酶学方法(Methods Enzymol.)》198:3,1991)、谷胱甘肽S转移酶(Smith和Johnson，《基因(Gene)》67:31,1988)、Glu-Glu亲和力标签(Grussenmeyer等人，《美国国家科学院院刊》82:7952-4,1985)、物质P、Flag.TM.肽(Hopp等人，《生物技术(Biotechnology)》6:1204-10,1988)、抗生蛋白链菌素结合肽或其它抗原的抗原决定基或结合域。一般参见Ford等人，《蛋白质表达和纯化(Protein Expression and Purification)》2:95-107,1991，编码亲和力标签的DNA可购自商业供应商(例如Pharmacia Biotech，新泽西州皮斯卡塔韦(Piscataway,N.J.))。

[0186] 在多种可获得的亲和力标签纯化系统中，最常采用的系统利用多组氨酸(His)或谷胱甘肽S-转移酶(GST)标签。His以良好选择性结合于并有Ni+2离子的基质，所述基质通常用亚氨基二乙酸或次氮基三乙酸螯合基团固定化。所述技术称为固定化金属亲和色谱。在中性到微碱性pH下进行His标记的蛋白质的吸附以防止弱碱性组氨酸咪唑基团的质子化和结合能力损失。所结合蛋白质的洗脱由用咪唑或低pH条件进行的置换引起。

[0187] 还描述在不永久引入外来DNA的情况下由体细胞产生诱导多能干细胞的方法。所述方法涉及用包含如本文所描述编码多肽序列或其VP1(V1)或VP3(V3)部分的分化体F衣壳或衣壳变异体核苷酸序列的载体瞬时转导干细胞。

[0188] 对于这些和其它实验，所属领域的技术人员知道如何修饰和繁殖AAV。举例来说，AAV-2可以作为裂解病毒和作为原病毒两者繁殖。对于裂解生长，AAV需要用辅助病毒共感染。腺病毒或单纯性疱疹可以提供辅助功能。当无辅助可获得时，AAV可以保持为整合的原病毒，其涉及在AAV端与宿主序列之间的重组，并且大部分AAV序列在原病毒中保持完整。AAV整合到宿主DNA中的能力允许在不存在辅助病毒的情况下繁殖。当随后用辅助感染运载

AAV原病毒的细胞时，修复经过整合的AAV基因组，并且进行生产性裂解循环。运载特定修饰的rAAV载体的构筑和例如用经过修饰的衣壳进行的rAAV粒子生产描述于例如Shi等人(2001)，《人类基因疗法(Human Gene Therapy)》12:1697-1711;Rabinowitz等人(1999)，《病毒学(Virology)》265:274-285;Nicklin等人(2001)，《分子疗法(Molecular Therapy)》4:174-181;Wu等人(2000)，J.《病毒学》74:8635-8647;以及Grifman等人(2001)，《分子疗法》3:964-974中。

[0189] 又另一方面是关于含有如本文所描述的AAV分化体F载体或AAV载体变异体或AAV粒子的医药组合物。含有AAV分化体F载体或AAV载体变异体或粒子的医药组合物优选地含有药学上可接受的赋形剂、佐剂、稀释剂、媒剂或载剂，或其组合。“药学上可接受的载剂”包括在与组合物的活性成分组合时允许所述成分保留生物活性并且不引起破坏性生理学反应(如不希望的免疫反应)的任何材料。药学上可接受的载剂包括水、磷酸盐缓冲盐水、乳液(如油/水乳液)和润湿剂。通过熟知的常规方法调配包含此类载剂的组合物，如以下中所阐述的那些方法:《雷明顿氏药物科学(Remington's Pharmaceutical Sciences)》，当前版，Mack Publishing Co.,Easton Pa.18042,USA;A.Gennaro(2000)《雷明顿:医药科学和实践(Remington:The Science and Practice of Pharmacy)》，第20版，Lippincott,Williams,&Wilkins;《医药剂型和药物递送系统(Pharmaceutical Dosage Forms and Drug Delivery Systems)》(1999)H.C.Ansel等人，第7版，Lippincott,Williams,&Wilkins;以及《医药赋形剂手册(Handbook of Pharmaceutical Excipients)》(2000)A.H.Kibbe等人，第3版Amer.Pharmaceutical Assoc。此类载剂可以通过常规方法来调配，并且可以以合适剂量向个体投与。合适的组合物的投与可以通过不同方式来实现，例如通过静脉内、腹膜内、皮下、肌肉内、局部或皮内投药。在一些实施例中，调配组合物以用于向哺乳动物投与。在一些实施例中，调配组合物以用于经由静脉内注射、皮下注射、肌肉内注射、自体细胞转移或同种异体细胞转移向哺乳动物投与。当然，投药途径尤其取决于医药组合物中所含有的载体的种类。给药方案将由主治医生和其它临床因素决定。如在医学领域中所熟知的，用于任一患者的剂量取决于许多因素，包括患者的大小、体表面积、年龄、性别、待投与的特定化合物、投药的时间和途径、感染或疾病的种类和阶段、一般健康状况以及并行投与的其它药物。

[0190] AAV分化体F载体或衣壳变异体中的一些能够在移植经过转导的造血干细胞之后或在直接全身递送rAAV之后支持长期的稳定的体内转基因表达。

[0191] 在某些实施例中，包含分化体F衣壳或AAV衣壳分离株变异体的核酸可以插入到新病毒的基因组中，其中在分化体F衣壳或衣壳分离株变异体的添加中，基因将与分化体F衣壳或AAV衣壳分离株相同或类似组织或器官向性传输到新病毒中。此类基因疗法可以使用体内和离体基因疗法程序来实现;参见例如美国专利第5,474,935号;Okada,《基因疗法(Gene Ther.)》3:957-964,1996。使用AAV分化体F衣壳或AAV衣壳变异体基因的基因疗法通常涉及将目标基因单独或与打算用于治疗性目的的另一个基因一起体外引入到新病毒中。如果将向性基因与一种或多种其它基因一起引入，那么优选地出于治疗性目的在分化体F衣壳或AAV分离株对其具有向性的组织中投与所得多肽。可以然后向需要此类疗法的患者投与病毒，或可以如向等待移植的器官离体投与病毒。病毒可以是反转录病毒、RNA病毒、DNA病毒，如腺病毒载体、腺相关病毒载体、牛痘病毒载体、疱疹病毒载体等等。还涵盖使用

病毒载体的转染方法，所述转染方法使用其中囊封有新病毒载体的脂质体用于投药。

[0192] 根据本文所提供的某些实施例，提供试剂盒，其包含一种或多种本文所描述的AAV分化体F载体或AAV载体变异体或其组合物或调配物。在某些实施例中，试剂盒中的一种或多种AAV分化体F载体或AAV载体变异体可以用于对细胞进行基因组编辑。在某些实施例中，试剂盒可以用做调查一种或多种AAV分化体F载体或AAV载体变异体进行基因组编辑的效果的研究工具。

[0193] 本发明的其它方面涉及一种用于重组制备如本文所描述的AAV(例如AAV分化体F载体或AAV变异体载体)的包装系统和其使用方法。在一些实施例中，包装系统包含编码如本文所描述的AAV分化体F衣壳的一个或多个AAV Rep蛋白质的Rep核苷酸序列；编码其一个或多个AAV Cap蛋白质的Cap核苷酸序列；以及如本文所描述校正基因组，其中所述包装系统在细胞中可操作地用于将校正基因组包封于衣壳中以形成腺相关病毒。

[0194] 在一些实施例中，包装系统包含有包含Rep核苷酸序列和Cap核苷酸序列的第一载体以及包含校正基因组的第二载体。如在如本文所描述的包装系统的情形下所用，“载体”是指本身为用于将核酸引入到细胞中的媒剂的核酸分子(例如，质体、病毒、粘质体、人工染色体等)。

[0195] 在包装系统的一些实施例中，AAV分化体F衣壳包含至少一个或至少两个选自分化体F VP1、分化体F VP2和分化体F VP3的蛋白质。在包装系统的一些实施例中，AAV分化体F衣壳包含分化体F VP1、分化体F VP2和分化体F VP3。在包装系统的一些实施例中，AAV分化体F衣壳选自由以下组成的群组：AAV9、AAVHSC1、AAVHSC2、AAVHSC3、AAVHSC4、AAVHSC5、AAVHSC6、AAVHSC7、AAVHSC8、AAVHSC9、AAVHSC10、AAVHSC11、AAVHSC12、AAVHSC13、AAVHSC14、AAVHSC15、AAVHSC16、AAVHSC17、AAVHU31以及AAVHU32。

[0196] 在包装系统的一些实施例中，Rep核苷酸序列编码AAV2Rep蛋白质。在包装系统的一些实施例中，所编码的AAV2Rep蛋白质是Rep 78/68或Rep 68/52中的至少一个。在包装系统的一些实施例中，编码AAV2Rep蛋白质的核苷酸序列包含编码与SEQ ID NO:40的AAV2Rep氨基酸序列具有最小序列一致性百分比的蛋白质的核苷酸序列，其中所述最小序列一致性百分比是跨AAV2Rep蛋白质的氨基酸序列的长度至少70％(例如，至少75％、至少80％、至少85％、至少90％、至少95％、至少98％、至少99％或100％)。

[0197] 示例性AAV2Rep氨基酸序列(SEQ ID NO:40)-

[0198] mpgfyeivikvpsdldehlpgisdsfvnwvaekewelppdsdmdlnlieqapltvaeklqrdfltewrrvskapealffvqfekgesyfhmhvlvettgvksmvlgrflsqirekliqriyrgieptlpnwfavtktrngagggnkvvdecyipnyllpktqpelqwawtnmeqylsaclnlterkrlvaqhlthvsqtqeqnkenqnpnsdapvirsktsarymelvgwlvdkgitsekqwiqedqasyisfnaasnsrsqikaaldnagkimsltktapdylvgqqpvedissnriykilelngydpqyaasvflgwatkkfgkrntiwlfgpattgktniaeaiahtvpfygcvnwtnenfpfndcvdkmviwweegkmtakvvesakailggskvrvdqkcckssaqidptpvivtsntnmcavidgnsttfehqqplqdrmfkfeltrrldhdfgkvtkqevkdffrwakdhvvevehefyvkkggakkrpapsdadisepkrvresvaqpstsdaeasinyadryqnkcsrhvgmnlmlfpcrqcermnqnsnicfthgqkdclecfpvsesqpvsvvkkayqklcyihhimgkvpdactacdlvnvdlddcifeq

[0199] 在包装系统的一些实施例中，所述包装系统进一步包含第三载体，例如辅助病毒载体。第三载体可以是独立的第三载体，与第一载体整合，或与第二载体整合。在一些实施

例中，第三载体包含编码辅助病毒蛋白质的基因。

[0200] 在包装系统的一些实施例中，辅助病毒选自由以下组成的群组：腺病毒、疱疹病毒(包括单纯性疱疹病毒(HSV))、痘病毒(如牛痘病毒)、巨细胞病毒(CMV)和杆状病毒。在其中辅助病毒是腺病毒的包装系统的一些实施例中，腺病毒基因组包含一个或多个选自由以下组成的群组的腺病毒RNA基因：E1、E2、E4和VA。在其中辅助病毒是HSV的包装系统的一些实施例中，HSV基因组包含选自由以下组成的群组的HSV基因中的一或多个：UL5/8/52、ICPO、ICP4、ICP22和UL30/UL42。

[0201] 在包装系统的一些实施例中，第一、第二和/或第三载体含于一个或多个转染质体内。在一些实施例中，第一载体和第三载体含于第一转染质体内。在一些实施例中，第二载体和第三载体含于第二转染质体内。

[0202] 在包装系统的一些实施例中，第一、第二和/或第三载体含于一个或多个重组辅助病毒内。在一些实施例中，第一载体和第三载体含于重组辅助病毒内。在一些实施例中，第二载体和第三载体含于重组辅助病毒内。

[0203] 在一些方面，本发明提供一种用于重组制备如本文所描述的AAV(例如，AAV分化体F载体或AAV变异体载体)的方法，其中所述方法包含用如所描述的包装系统在可操作地用于将校正基因组包封于衣壳中的条件下转染或转导细胞以形成如本文所描述的AAV(例如，AAV分化体F载体或AAV变异体载体)。用于重组制备AAV的示例性方法包括瞬时转染(例如，用一个或多个含有如本文所描述的第一和第二载体以及任选地第三载体的转染质体)、病毒感染(例如，用一个或多个重组辅助病毒，如腺病毒、痘病毒(如牛痘病毒)、疱疹病毒(包括单纯性疱疹病毒(HSV)、巨细胞病毒或杆状病毒，其含有如本文所描述的第一和第二载体以及任选地第三载体)以及稳定的生产细胞系转染或感染(例如，用稳定的生产细胞，如哺乳动物或昆虫细胞，其含有编码如本文所描述的AAV分化体F衣壳的一个或多个AAV Rep蛋白质的Rep核苷酸序列和/或编码其一个或多个AAV Cap蛋白质的Cap核苷酸序列，并且其中如本文所描述的校正基因组以转染质体或重组辅助病毒的形式递送)。

[0204] 下文提供本发明的其它示例性非限制性实施例。

[0205] 实施例1.一种用于编辑干细胞基因组的腺相关病毒(AAV)载体变异体，其包含一个或多个衣壳变异体；包含一个或多个待整合到所述基因组的目标位点中的治疗性核苷酸序列的靶向盒；侧接所述靶向盒并且与位于所述目标位点上游的区具有同源性的5'同源臂多核苷酸序列；以及侧接所述靶向盒并且与位于所述目标位点下游的区具有同源性的3'同源臂多核苷酸序列。

[0206] 实施例2.根据实施例1所述的AAV载体变异体，其中所述一个或多个衣壳变异体包含选自以下群组的多肽序列：HSC7(SEQ ID NO:8)、HSC12(SEQ ID NO:12)、HSC15(SEQ ID NO:16)、HSC17(SEQ ID NO:13)、其变异体、片段、突变体以及任何组合。

[0207] 实施例3.根据实施例2所述的AAV载体变异体，其中所述一个或多个衣壳变异体包含与选自以下群组的多肽序列具有至少95％序列一致性百分比的多肽序列：HSC7(SEQ ID NO:8)、HSC12(SEQ ID NO:12)、HSC15(SEQ ID NO:16)、HSC17(SEQ ID NO:13)、其变异体、片段、突变体以及任何组合。

[0208] 实施例4.根据实施例1所述的AAV载体变异体，其中所述目标位点是安全港位点。

[0209] 实施例5.根据实施例4所述的AAV载体变异体，其中所述安全港位点是染色体19上

的AAVS1基因座。

[0210] 实施例6.根据实施例1所述的AAV载体变异体，其中所述干细胞是造血干细胞、多能干细胞、胚胎干细胞或间叶细胞干细胞。

[0211] 实施例7.一种编辑干细胞基因组的方法，其包含在没有其它外源性核酸酶的情况下用一种或多种腺相关病毒(AAV)载体变异体转导所述干细胞，所述载体变异体包含一个或多个衣壳变异体；包含一个或多个待整合到所述基因组的目标位点中的治疗性核苷酸序列的靶向盒；侧接所述靶向盒并且与位于所述目标位点上游的区具有同源性的5'同源臂多核苷酸序列；以及侧接所述靶向盒并且与位于所述目标位点下游的区具有同源性的3'同源臂多核苷酸序列。

[0212] 实施例8.根据实施例7所述的方法，其中所述一个或多个衣壳变异体包含选自以下群组的多肽序列：HSC7(SEQ ID NO:8)、HSC12(SEQ ID NO:12)、HSC15(SEQ ID NO:16)、HSC17(SEQ ID NO:13)、其变异体、片段、突变体以及任何组合。

[0213] 实施例9.根据实施例7所述的方法，其中所述AAV载体变异体不含有用于一个或多个治疗性核苷酸序列的启动子。

[0214] 实施例10.根据实施例7所述的方法，其中所述目标位点是安全港位点。

[0215] 实施例11.根据实施例10所述的方法，其中所述安全港位点是染色体19上的AAVS1基因座。

[0216] 实施例12.根据实施例7所述的方法，其中所述干细胞是造血干细胞、多能干细胞、胚胎干细胞或间叶细胞干细胞。

[0217] 实施例13.一种在个体中通过编辑所述个体的干细胞的基因组来治疗疾病或病症的方法，其包含在没有其它外源性核酸酶的情况下用腺相关病毒(AAV)载体变异体转导所述个体的所述干细胞，所述载体变异体包含一个或多个衣壳变异体；包含一个或多个待整合到所述干细胞的所述基因组中的目标位点中的治疗性核苷酸序列的靶向盒；侧接所述靶向盒并且与位于所述目标位点上游的区具有同源性的5'同源臂多核苷酸序列；侧接所述靶向盒并且与位于所述目标位点下游的区具有同源性的3'同源臂多核苷酸序列；和将经过转导的干细胞移植到所述个体中，其中所述经过转导的干细胞治疗所述疾病或病症。

[0218] 实施例14.根据实施例13所述的方法，其中所述一个或多个衣壳变异体包含来自以下群组的多肽序列：HSC7(SEQ ID NO:8)、HSC12(SEQ ID NO:12)、HSC15(SEQ ID NO:16)、HSC17(SEQ ID NO:13)、其变异体、片段、突变体以及任何组合。

[0219] 实施例15.根据实施例13所述的方法，其中所述AAV载体变异体不含有用于一个或多个治疗性核苷酸序列的启动子。

[0220] 实施例16.根据实施例13所述的方法，其中所述目标位点是安全港位点。

[0221] 实施例17.根据实施例16所述的方法，其中所述安全港位点是染色体19上的AAVS1基因座。

[0222] 实施例18.根据实施例13所述的方法，其中所述干细胞是造血干细胞、多能干细胞、胚胎干细胞或间叶细胞干细胞。

[0223] 实施例19.根据实施例13所述的方法，其中所述疾病或病症由细胞基因组中的一个或多个突变造成。

[0224] 实施例20.根据实施例19所述的方法，其中所述疾病或病症选自遗传性代谢病、溶

酶体贮积病、粘多糖病、免疫缺陷疾病以及血红蛋白病和感染。

[0225] 实施例21.一种在个体中通过对所述个体的细胞进行体内基因组编辑来治疗疾病或病症的方法,其通过向所述个体直接投与AAV载体变异体来进行,所述载体包含一个或多个衣壳变异体;包含一个或多个待整合到所述基因组中的目标位点中的治疗性核苷酸序列的靶向盒;侧接所述靶向盒并且与位于所述目标位点上游的区具有同源性的5'同源臂多核苷酸序列;以及侧接所述靶向盒并且与位于所述目标位点下游的区具有同源性的3'同源臂多核苷酸序列,其中所述载体转导所述个体的所述细胞并且将所述一个或多个治疗性核苷酸序列整合到所述细胞的所述基因组中。

[0226] 实施例22.根据实施例21所述的方法,其中所述一个或多个衣壳变异体包含选自以下群组的多肽序列:HSC1(SEQ ID NO:2)、HSC2(SEQ ID NO:3)、HSC11(SEQ ID NO:4)、HSC3(SEQ ID NO:5)、HSC4(SEQ ID NO:6)、HSC6(SEQ ID NO:7)、HSC7(SEQ ID NO:8)、HSC8(SEQ ID NO:9)、HSC9(SEQ ID NO:10)、HSC5(SEQ ID NO:11)、HSC12(SEQ ID NO:12)、HSC17(SEQ ID NO:13)、HSC13(SEQ ID NO:14)、HSC14(SEQ ID NO:15)、HSC15(SEQ ID NO:16)、HSC16(SEQ ID NO:17)、其变异体、片段、突变体以及任何组合。

[0227] 实施例23.根据实施例21所述的方法,其中所述细胞是干细胞。

[0228] 实施例24.根据实施例23所述的方法,其中所述干细胞是造血干细胞、多能干细胞、胚胎干细胞或间叶细胞干细胞。

[0229] 实施例25.根据实施例24所述的方法,其中所述疾病或病症由细胞基因组中的一个或多个突变造成。

[0230] 实施例26.根据实施例25所述的方法,其中所述疾病或病症选自遗传性代谢病、溶酶体贮积病、粘多糖病、免疫缺陷疾病以及血红蛋白病和感染。

[0231] 实施例27.根据实施例21所述的方法,其中所述AAV载体变异体不含有用于一个或多个治疗性核苷酸序列的启动子。

[0232] 实施例28.根据实施例22所述的方法,其中所述目标位点是安全港位点。

[0233] 实施例29.根据实施例23所述的方法,其中所述安全港位点是染色体19上的AAVS1基因座。

[0234] 以下实例打算说明本发明的各种实施例。因而,所论述的具体实施例不应解释为对本发明范围的限制。对所属领域的技术人员将显而易见的是,可以在不脱离本发明范围的情况下得到各种等效物、变化和修改,并且应理解,此类等效实施例将包括于本文中。此外,本发明中所引用的所有参考文献都以全文引用的方式并入本文中,如同在本文中充分阐述一般。

[0235] 实例

[0236] 实例1:AAV分化体F载体变异体在CD34+人类造血干细胞系或K562细胞系中介导的基因组编辑

[0237] 在来自健康供体人类的CD34+造血细胞系或K562细胞系(其是CD34+红白血病细胞系)中进行在不使用外源性核酸酶的情况下使用AAVF(AAVHSC)载体经由位点特异性插入或特定DNA序列靶向整合的基因组编辑。构筑一组供体重组AAV载体ITR-AAVS1-FP载体,并且将其用于将转基因整合到染色体19上的AAVS1基因座(天然野生型AAV整合位点)中(Kotin,1992;Giraud,1994)。先前展示染色体19上的AAVS1基因座qter13.3-13.4为用于插入转基

因的“安全港”位点，这是因为此处插入的基因进行表达而无病原性后果，其与在此基因座处整合而无病原性后果的野生型AAV类似(Giraud,1994;Linden,1996A;Linden 1996B)。待整合的转基因是Venus黄色荧光蛋白质(“YFP”或“FP”)基因，其在每一侧由与人类染色体19上的AAVS1基因座具有同源性的大致800个核苷酸侧接(参见图3中的示意图)。供体AAV载体被设计成使得转基因无启动子并且仅将在其整合在校正基因座处的情况下进行表达，所述校正基因座位于染色体地编码的调节序列的下游(参见图4)。因此，所进行的任何Venus YFP转基因表达都处于位于AAVS1中或接近AAVS1的染色体启动子的控制下。

[0238] 根据在Chatterjee等人，1992中所描述的标准AAV包装方法将供体载体ITR-AAVS1-FP包装到AAVHSC衣壳中。具体地说，将ITR-AAVS1-FP包装到AAVHSC7、AAVHSC15或AAVHSC17衣壳中，从而形成假型化的AAVHSC-AAVS1-FP载体(即，AAV载体变异体)。用假型化的AAVHSC-AAVS1-FP以不同感染复数(MOI)(即，50,000MOI;100,000MOI;200,000MOI;300,000MOI;以及400,000MOI)转导细胞人类CD34+造血干细胞系或K562细胞。

[0239] 最初，通过对经过转导的K562细胞中的YFP表达进行细胞荧光测定分析来分析YFP转基因通过同源重组向AAVS1基因座中的整合。使用AAVHSC7FP载体的靶向整合在转导后24小时时(图5和7A)和转导后72小时时(图6和7B)引起YFP转基因的表达。此外，随着AAVHSC7FP载体的MOI增加，平均YFP表达百分比也增加(图7A和B)。

[0240] YFP转基因的靶向整合进一步通过使用位于同源区外侧的引物对经过编辑的基因组进行PCR扩增来确认。简言之，从以100,000个载体基因组/细胞的MOI用AAVHSC7FP载体转导的K562细胞中提取DNA。使用“外正向引物区”和“外反向引物区”引物来进行PCR扩增(参见图4)。YFP转基因的整合引起AAVS1基因座的大小从约1.9kb的野生型大小增加为含有YFP转基因的约3.1kb片段(参见图4，将标记为“片段1”的线与标记为“片段2”的线相比较)。在染色体19AAVS1基因座内含有YFP转基因的约3.1kb片段的扩增指示，YFP转基因有效地整合到用AAVHSC7FP载体转导的细胞中的AAVS1基因座中(参见图8A和8B，色带4)。

[0241] 实例2:AAV载体变异体在原代人类CD34+周边血液干细胞中介导的基因组编辑

[0242] 在人类CD34+原代周边血液源性人类造血干细胞(PBSC)中还进行在不使用外源性核酸酶的情况下使用AAVHSC载体经由位点特异性插入或特定DNA序列靶向整合的基因组编辑。简言之，将载体ITR-AAVS1-FP包装在AAVHSC衣壳(包括AAVHSC7、AAVHSC12、AAVHSC15和AAVHSC17)中(标准AAV包装方法参见Chatterjee,1992)。用假型化的AAVHSC-AAVS1-FP载体(即，AAV载体变异体)以100,000和150,000的MOI转导原代CD34+细胞。

[0243] 通过对YFP表达进行细胞荧光测定分析来分析YFP转基因通过同源重组向AAVS1基因座中的整合。在原代CD34+细胞中使用AAVHSC7FP和AAVHSC17FP载体的靶向整合在以150,000的MOI转导后1天时(图9)、以100,000的MOI转导后4天时(图10)和以100,000的MOI转导后18天时(图11)引起YFP转基因的表达。在以100,000的MOI转导后5.5周(39天)时，表达YFP的阳性细胞的百分比未衰减(参见图12A和B，将20天结果与39天结果相比较)。无启动子YFP转基因在分裂细胞群体中的此长期表达指示转基因的精确整合。

[0244] 通过PCR分析进一步确认YFP转基因的靶向整合。使用对处于所插入转基因序列与原生染色体5’同源臂序列之间的5’接合区进行扩增的引物来对经过编辑的基因组进行扩增(参见图4，参见标记为“片段3”的线)。简言之，从以150,000个载体基因组/细胞的MOI用AAVHSC7FP载体转导的原代CD34+细胞中提取DNA。使用“外正向引物区”和“内反向引物”引

物来进行PCR扩增(参见图4)。用于经过转导的原代CD34+细胞的5'接合区中约1.7kb片段的扩增指示YFP转基因成功地整合到AAVS1基因座中(参见图13,色带5)。然而,未用AAVHSC7FP载体转导的那些细胞不存在经过扩增的产物(参见图13,色带3)。

[0245] 通过对经过编辑的AAVS1基因座的不同部分进行序列分析来进一步确认YFP转基因的靶向整合。在接近"外正向引物区"处(参见图14)、接近5'同源臂处(参见图15)、接近调节元件5'区处(参见图16)、接近调节元件3'区处(参见图17)、接近转基因5'区处(参见图18)和接近"内反向引物"区处(参见图19)开始进行测序。测序结果指示,存在YFP基因,并且其成功地整合到AAVS1基因座中。

[0246] 如在实例1和2中所提供,AAVHSC载体在不需要添加外源性核酸内切酶的情况下,在人类CD34+造血细胞系和CD34+PBSC中成功地介导高效靶向基因组编辑。AAVHSC载体能够引导YFP转基因基于对应于AAVS1基因座的侧接同源臂而整合到染色体19上的AAVS1基因座中。先前已经报告AAV介导的基因转移;然而,所报告的出现频率一直较低,数量级通常为1e6个细胞之一到1e4个细胞之一。如本文所示,使用AAVHSC载体的靶向基因组编辑以原代细胞大致10%的出现频率长期进行,其效率是先前所报告的1,000到100,000倍(参见例如Khan,2011)。早在转导后第一天就观察到YFP转基因在人类CD34+造血细胞系中的表达,并且在第三天确认。从第一天开始就观察到YFP转基因在PBSC中的表达并且长期持续(高达几乎6周),6周是所分析的最晚时间点。未观察到由于AAVHSC载体转导的明显毒性。基于所观察到的高频率插入、易用性和不具有毒性,基于使用AAVHSC载体的靶向基因组编辑的疗法是实际并且可行的。

[0247] 实例3:使用AAV载体变异体的体内基因组工程改造

[0248] 将编码荧光素酶的AAVHSC载体和编码Venus的AAVHSC载体注射到先前用人类脐血CD34+HSC异种移植的成年免疫缺陷型小鼠中。如下文所示,AAVHSC载体的静脉内注射引起人类CD34+造血干细胞和祖细胞的体内转导,并且静脉内注射的AAVHSC载体运输到人类造血位点和经过转导的人类细胞处。

[0249] 方法.简言之,首先用^{137}Cs源的350cGy亚致死剂量照射免疫缺陷型NOD/SCID成年小鼠。其次,将一百万个人类脐血CD34+细胞注射到经过亚致死剂量照射的免疫缺陷型NOD/SCID小鼠中。接下来,在CD34+HSC移植之后两小时时,用大致1E11-5e11个AAVHSC-荧光素酶载体(AAVHSC7-荧光素酶载体或AAVHSC17-荧光素酶载体)粒子对小鼠进行静脉内注射。在无外源性核酸酶存在下使用这些载体。AAVHSC-荧光素酶载体在普遍存在的CBA启动子的控制下编码单链萤火虫荧光素酶基因(ssLuc)以允许对转基因表达进行连续体内生物发光监测。这些载体具体描述于美国申请第13/668,120号(以美国专利公开第20130096182A1号公开)和以全文引用的方式并入本文中,如同在本文中充分阐述一般的Smith等人(参见Smith,2014)中。如美国申请第US 13/668,120号(以美国专利公开第20130096182A1号公开)和Smith等人中所描述,在HSC7衣壳变异体中(HSC7衣壳的多核苷酸序列提供为SEQ ID NO:27,并且HSC7衣壳的多肽序列提供为SEQ ID NO:8(参见图1))或在HSC17衣壳变异体中(HSC17衣壳的多核苷酸序列提供为SEQ ID NO:35,并且HSC17衣壳的多肽序列提供为SEQ ID NO:13(参见图1))对AAVHSC-荧光素酶载体进行假型化,以形成AAVHSC7-荧光素酶载体和AAVHSC17-荧光素酶载体(参见Smith,2014)(标准AAV包装方法参见Chatterjee,1992)。应注意,AAVHSC-荧光素酶载体可以转导小鼠和人类细胞两者。然而,相比于在下文所描述

的AAVHSC-Venus载体中编码的Venus，荧光素酶将不整合到AAVS1中。相反，荧光素酶可以随机整合或保持游离。

[0250] 在用AAVHSC-荧光素酶载体注射之后二到七天时，用大致1E11-5e11个AAVHSC-Venus载体粒子对小鼠进行注射。具体地说，对首先用AAVHSC7-荧光素酶载体注射的小鼠注射AAVHSC7-Venus载体，并且对首先用AAVHSC17-荧光素酶载体注射的小鼠注射AAVHSC17-Venus载体。所用Venus供体载体具体描述于上文实例1和2中。供体AAV载体被设计成使得Venus转基因无启动子并且仅将在其整合在校正AAVS1基因座处的情况下进行表达，所述校正AAVS1基因座位于染色体地编码的调节序列的下游(参见图3和4)。因此，所进行的任何Venus转基因表达都处于位于AAVS1中或接近AAVS1的染色体启动子的控制下。重要的是，含有Venus基因的载体不含有驱动表达的启动子。Venus基因仅将在其正确地整合到人类细胞中AAVS1中的情况下表达，所述AAVS1位于内源性染色体启动子下游。根据Chatterjee等人，1992中所描述的标准AAV包装方法将供体载体ITR-AAVS1-Venus(参见图3)包装到AAVHSC衣壳中。在HSC7衣壳变异体中(HSC7衣壳的多核苷酸序列提供为SEQ ID NO:27，并且HSC7衣壳的多肽序列提供为SEQ ID NO:8(参见图1))或在HSC17衣壳变异体中(HSC17衣壳的多核苷酸序列提供为SEQ ID NO:35，并且HSC17衣壳的多肽序列提供为SEQ ID NO:13(参见图1))对载体进行假型化，以形成AAVHSC7-Venus载体和AAVHSC17-Venus载体。

[0251] 最后，在注射后4周时测量体内荧光素酶表达。在注射后六周时，测量人类CD34+和CD45+细胞的植入，并且对Venus表达进行量化。在此实例中进行的实验整体示意图参见图20。

[0252] 结果.在注射后四周，对接受AAVHSC-荧光素酶载体静脉内注射的异种移植和非异种移植的小鼠进行体内成像。结果展示在脊椎、脾脏、髋部和长骨中的特异性荧光素酶表达，其全部是移植之后的造血位点(参见图21A)。然而，在先前未用人类脐血CD34+HSC异种移植的小鼠中未观察到造血器官中的特异性荧光素酶表达(参见图21B)。这些结果指示，静脉内注射的AAVHSC载体运输到体内的人类造血位点处并且优先转导干细胞和祖细胞。

[0253] 在用AAVHSC-Venus载体注射之后六周时，使用流式细胞测量术来分析人类CD34+和CD45+细胞。结果指示所注射的人类脐血CD34+细胞植入到小鼠中，并且得到更成熟的血细胞。具体地说，在骨髓中观察到原始人类血液祖细胞(即，CD34+细胞)(参见表1，CD34+细胞和股骨骨髓)。此外，如表1中所示，人类单核血细胞(即，CD45+细胞)在股骨骨髓、椎骨骨髓和脾脏中显而易见。

[0254] 表1：人类血细胞在免疫缺陷型小鼠中的植入

[0255]

AAVHSC7			
细胞类型	股骨骨髓	椎骨骨髓	脾脏
CD45+	68.4	30.4	24.3
CD34+	22	NT*	NT
AAVHSC17			
细胞类型	股骨骨髓	椎骨骨髓	脾脏
CD45+	46.6	24	18.3
CD34+	13.1	NT	NT
*NT=未经测试 CD45+细胞：人类单核血细胞 CD34+细胞：人类造血祖细胞			

[0256] 在注射后六周时，使用流式细胞测量术来分析接受AAVHSC7-Venus或AAVHSC17-Venus载体静脉内注射的异种移植小鼠的人类HSC中的Venus表达。结果揭示，在来自股骨和椎骨骨髓的CD45+人类HSC以及CD34+人类HSC中易于观察到AAVHSC转导(参见表2和图22A-F)。

[0257] 表2:表达Venus的所植入人类造血细胞的百分比

[0258]

AAVHSC7			
细胞类型	股骨骨髓	椎骨骨髓	脾脏
CD45+	8.35	15.3	10.3
CD34+	9.23	NT*	NT
AAVHSC17			
细胞类型	股骨骨髓	椎骨骨髓	脾脏
CD45+	8.59	70.2	9.9
CD34+	8.92	NT	NT
*NT=未经测试			

[0259]

CD45+细胞：人类单核血细胞 CD34+细胞：人类造血祖细胞

[0260] 此外，由于Venus在脾脏中的人类CD45+细胞中表达，故这些细胞易于展示转导证据(参见表2和图22G和H)。这展现由所移植人类脐血CD34+细胞产生的CD45+细胞表达Venus。这些结果指示，体内的AAVHSC载体静脉内注射引起人类造血细胞的转导。

[0261] 实例4:使用AAV分化体F载体将较大和较小编辑元件插入到基因组中

[0262] 方法.rAAV产生、纯化和滴定.使用New England Biolabs Gibson Assembly克隆试剂盒用使用NEBuilder v.1.6.2(马萨诸塞州伊普威治(Ipswich,MA))设计的引物将所有靶向基因组克隆到AAV2主链中。对所有靶向基因组进行测序，并且使用限制消化和测序来确认AAV2ITR完整性。将单链靶向基因组包装到单纯性疱疹病毒(HSV)感染的293细胞的AAVF衣壳中。经由两轮CsCl2密度离心梯度来纯化所得重组AAV载体，并且使用qPCR用转基因特异性引物和探针来测定效价。

[0263] K562、HepG2和PBSC转导.从美国典型培养物保藏中心(ATCC)(弗吉尼亚州马纳萨斯(Manassas,VA))获得慢性骨髓性白血病(CML)细胞系K562和肝细胞癌细胞系HepG2，并且根据ATCC指导原则进行培养。使用CD34+间接分离试剂盒(美天旎生物技术公司(Miltenyi Biotech))由来自细胞激素预致敏的健康供体PB的单核细胞纯化周边血液干细胞(PBSC)，

并且紧接在分离之后进行转导。在含有20%FCS、100μg/mL链霉素、100U/ml青霉素、2mmol/L L-谷氨酰胺、IL-3(10ng/mL;R&D Systems)、IL-6(10ng/mL;R&D Systems)和干细胞因子(1ng/mL;R&D Systems)的伊斯寇氏修饰的杜尔贝科氏培养基(Iscove's Modified Dulbecco's Medium,IMDM)(Irvine Scientific)中培养PBSC。在转导之前大致24小时时分裂和接种HepG2细胞。在即将转导之前接种K562细胞。用靶向AAVF的载体以介于5E4到4E5范围内的MOI转导K562、HepG2或PBSC。在无外源性核酸酶存在下转导细胞并实现同源重组。在介于转导后1到39天之间的时间点采集细胞用于流动和分子分析。在如所指示使用APC BrdU流动试剂盒(BD Biosciences)采集之前进行体外转导的BrdU标记。

[0264] TI PCR和测序.从用靶向AAVF的载体转导的K562、HepG2或PBSC中分离出高分子量DNA。使用粘接到同源臂外侧染色体区的引物,即Sigma AAVS1正向引物(5'-GGC CCT GGC CAT TGT CAC TT-3')和粘接到所插入盒的引物,即Venus反向引物(5'-AAC GAG AAG CGC GAT CAC A-3')或RFLP HindIII反向引物(5'-CCAATCCTGTCCCTAGTAAAGCTT-3')来进行TI特异性PCR。使用罗氏Expand Hifidelity PCR系统(印第安纳州印第安纳波利斯(Indianapolis,IN)),并且循环条件如下:1个循环,5分钟-95℃;15个循环,30秒-95℃,30秒-在62℃下开始并且每循环降低0.5℃,2分钟-68℃;20个循环,30秒-95℃,30秒-53℃,2分钟-68℃;1个循环,5分钟-68℃。对PCR产物进行PCR纯化以用于使用Qiaquick PCR纯化试剂盒(Qiagen)进行直接测序,或使用TOPO TA测序用克隆试剂盒进行克隆并且通过桑格测序(Sanger Sequencing)(Life Technologies)进行克隆测序。

[0265] CD34+细胞移植.在由机构动物照护与使用委员会(Institutional Animal Care and Use Committee)批准的方案下进行所有动物照护和实验。在无特定病原体的设施中养护6-8周龄雄性NOD.CB17-Prkdcscid/NCrCrl(NOD/SCID)小鼠。在移植之前,对小鼠使用磺胺甲基异噁唑(sulfamethoxazole)和三甲氧苄二氨嘧啶(trimethoprim)口服儿童抗生素(Hi-Tech Pharmacal(纽约州艾米提维尔(Amityville,NY)),10ml/500ml H2O)持续至少48小时。用137Cs源的350cGy亚致死剂量照射小鼠,并且在移植之前允许其恢复最少4小时。使用CD34+间接分离试剂盒(美天旎生物技术公司)来分离脐带血(CB)CD34+细胞。将1×10^6个CB CD34+细胞再悬浮于大致200μL中,并且通过尾部静脉注射移植。在CB CD34+移植后1周或7周时,经由尾部静脉进行静脉内注射2.4E11到6.0E11个靶向AAVF的载体的粒子。在移植后6、7或19周时采集股骨骨髓(BM)、椎骨BM和脾脏。

[0266] 流式细胞测量分析.使用Cyan ADP流式细胞仪(Dako)来分析体外转导中靶向AAVF的载体所介导的整合以及BrdU和7-AAD。在扣除自身荧光后对比荧光进行量化。在FACS Aria SORP(BD Biosciences)上通过用人类特异性抗体、APC结合的抗CD34和PE结合的抗血型糖蛋白A以及PE和APC结合的IgG对照物(BD Biosciences)染色来在所采集的异种移植小鼠的椎骨BM、股骨BM和脾脏中分析经过整合的荧光盒在CD34+细胞和红细胞中的体内表达。使用FlowJo软体(Treestar)来分析流式细胞测量术数据。

[0267] 结果.展示干细胞源性AAV基于衣壳基因的核苷酸序列同源性而按图谱变为AAV分化体F(图23,Smith等人,《分子治疗学》2014年9月;22(9):1625-34)。这些干细胞源性AAV称为AAVHSC1-17。这些AAV在本文中也相应称为AAVF1-17。

[0268] 单链AAV载体基因组用于设计含有同源臂和较大插入物的校正基因组(图24)。所述插入物含有位于剪接受体(SA)和2A序列(2A)下游的无启动子Venus开放阅读框(ORF)以

允许独立蛋白质表达。Venus是黄色荧光蛋白质变异体(参见例如Nagai等人《自然·生物技术(Nat Biotechnol)》,2002,20(1):87-90)。左和右同源臂(HA)各自是800bp长,并且与位于染色体19上AAVS1基因座中的人类PPP1R12C内含子1的序列互补(图25)。类似单链AAV载体基因组被设计成在两个同源臂之间具有10bp插入物(图35)。AAVS1基因座被视为用于异源转基因插入的安全港位点。

[0269] 在AAV2反向末端重复(ITR)之间克隆同源臂、开放阅读框和调节序列。然后将此校正基因组在不同AAV衣壳中包装(假型化),所述AAV衣壳包括AAVHSC、AAV8、9、6和2。然后使用重组病毒将经过编辑的基因组递送到目标细胞的原子核中。所测试的目标细胞包括CD34+红白血病细胞系、肝脏细胞系和原代人类CD34+造血干细胞/祖细胞以及其造血子代。

[0270] 含有由与人类PPP1R12C基因内含子1互补的同源臂前接和侧接的Venus ORF的AAVF载体用于将经过编辑的基因组递送到原代人类CD34+细胞激素预致敏的周边血液干细胞(图26A)、人类CD34+红白血病细胞系K562(图26B)和人类肝脏细胞系HepG2(图26C)中。原代CD34+细胞支持最高水平的编辑,高达60%(图26A-F)。永生化细胞系(包括K562和HepG2)也展示显著编辑水平。在所有情况下,所实现的编辑水平一致地显著高于用非分化体F病毒(包括AAV6和AAV8)实现的编辑水平(图26A-F和图36)。

[0271] 在另一个实验中,用染色体特异性引物和插入物特异性引物对从用AAVF7、AAVF15或AAVF17载体转导的细胞激素预致敏的CD34+周边血液干细胞(PBSC)中提取的DNA进行扩增。所述载体包括较大插入物(Venus)或较短插入物(10bp,RFLP)。凝胶展示由经过编辑的CD34细胞对大小正确的扩增子进行扩增(图27、图35和图36)。1.7kb和1kb条带的存在分别反映较大和较小插入物的正确靶向整合。在用AAVF载体编辑之后,展示短期和长期时间点的靶向整合(图27)。

[0272] 在另一个实验中,单链AAV载体基因组被设计成用于将10bp插入物插入在人类PPP1R12C基因的内含子1中(图28A)。这些载体包括含有Nhe1限制酶识别位点(GCTAGC)的野生型左同源臂(HA-L)。NS mut载体被设计成用于将染色体19上左同源臂中的TA序列变为AT。此变化引起Nhe1位点转化为Sph1位点,将序列从GCTAGC变为GCATGC。图28B展示当使用野生型或NS Mut AAVF载体编辑来自K562细胞的基因组DNA时,通过用Nhe1或Sph1切割而产生的预期片段的相对大小。如所预测的,在用Nhe 1而非用Sph1的情况下消化来源于用野生型AAVF载体编辑的K562细胞的基因组DNA中的实际扩增子(图28C)。经过扩增的K562DNA在用编码野生型或NS Mut基因组的AAVF7或AAVF17载体编辑之后的结果展示,用Nhe1消化不再引起扩增子裂解,与来自未经编辑的细胞的扩增子类似。用Sph1消化引起裂解,展现染色体左同源臂中的Nhe1位点由Sph1位点替换(图28D)。来自肝细胞癌细胞系HepG2的经过扩增的DNA在用编码野生型或NS Mut基因组的AAVF7或AAVF17载体编辑之后的电泳展示,用Nhe1消化不再引起扩增子裂解,与来自未经编辑的细胞的扩增子类似(图28E)。用Sph1消化引起裂解,展现染色体左同源臂中的Nhe1位点由Sph1位点替换。序列分析确认使用AAVF7和AAVF17野生型或NS Mut载体编辑(图29)。这些结果展现,AAVF7以及AAVF17两者都在两个不同细胞系中的染色体序列中成功地介导2个核苷酸取代。这些结果还展现AAVF载体在人类细胞的基因组DNA中介导2歌碱基对取代的能力,表明其用于在基因组中校正疾病造成的突变或诱导新突变的用途。

[0273] 在另一个实验中,对健康人类CD34+PBSC测试细胞分裂对AAVF载体编辑能力的潜

在要求。经由用10μM BrdU脉冲2小时来将BrdU并入到经过转导的人类CD34+PBSC中。在DNA酶处理之前，采集经过AAVF转导CD34+细胞，对其进行渗透并且固定。在DNA酶处理之后，用抗BrdU APC抗体对经过处理的细胞进行染色持续20分钟。在使用APC BrdU流动试剂盒(BD Biosciences)采集之前按照指示进行经过体外编辑的细胞的BrdU标记。接着通过流式细胞测量术来分析细胞的Venus表达以及BrdU标记。结果揭示，Venus表达在BrdU阳性和阴性群体两者中的出现频率类似，表明细胞分裂不是AAVF介导的编辑所需的(图30)。

[0274] 在另一个实验中，通过对植入有人类脐血CD34+造血干细胞的免疫缺陷型NOD/SCID小鼠全身递送AAVF载体来测试所植入人类造血干细胞的体内编辑(图31A和B)。在骨髓以及脾脏中，发现大部分人类细胞表达Venus，而在小鼠细胞中未观察到Venus表达(图31C、图37和图38)。因为小鼠基因组不含有与同源臂互补的AAVS1基因座，所以这些发现展现基因靶向的特异性。在所分析的人类细胞中，在原始CD34+祖细胞以及骨髓和脾脏两者中的成熟血型糖蛋白A+红细胞中观察到Venus表达。

[0275] 在另一个实验中，小鼠植入有人类脐血CD34+细胞，并且在1周或7周后通过静脉内途径注射AAVF-Venus。在静脉内注射Venus之后5、6或13周时采集椎骨或股骨骨髓或脾脏。这些表示6、7或20周的移植后累积时间。结果揭示，AAVF-Venus的静脉内注射引起体内植入的原始(CD34+)以及更成熟的分化(CD45+)人类造血细胞两者的编辑。在移植和注射之后，人类红细胞谱系的细胞展现极高效的长期编辑(图32)。发现AAVF介导的编辑长期稳定，并且由体内植入的人类CD34+细胞的分化子代稳定遗传(图32)。经过编辑的CD34+细胞的分化子代长期表达Venus(图32)。

[0276] 在另一个实验中，对无启动子SA/2A venus ORF在K562红白血病细胞系、原代人类细胞激素预致敏的周边血液CD34+细胞和HepG2人类肝脏细胞系中的靶向染色体插入进行序列分析(图33)。使用染色体特异性引物和插入物特异性引物对位点特异性整合的序列进行扩增。结果揭示SA/2A Venus在每一情况下在处于左与右同源臂之间的接合点处的精确插入(图33)。

[0277] 在另一个实验中，对10bp插入物在原代人类细胞激素预致敏的周边血液CD34+细胞和HepG2人类肝脏细胞系中的靶向染色体插入进行序列分析(图34)。使用染色体特异性引物和插入物特异性引物对位点特异性整合的序列进行扩增。结果揭示10bp插入物在每一情况下在处于左与右同源臂之间的接合点处的精确插入(图34)。

[0278] 这些数据展示，较大和较短插入物两者都可以使用AAV分化体F载体成功地编辑到基因组中，并且向基因组中的整合是精确的。这些数据还展示，AAV分化体F载体可以用于在无外源性核酸酶存在下进行高效基因组编辑。

[0279] 实例5：编辑人类细胞系中的PPP1R12c基因座

[0280] 方法

[0281] 为了评定AAVF载体对人类细胞类型的编辑，使用以下人类细胞系：

[0282]

细胞系	组织类型
WI-38	正常人类二倍体成纤维细胞
MCF7	人类乳癌细胞系
Hep-G2	人类肝细胞癌细胞系
K562	CD34+红白血病细胞系

Y79	人类成视网膜细胞瘤细胞系
SCID-X1LBL	来自SCID-X1患者的人类EBV永生化B细胞系

[0283] 所用AAVF载体各自含有含有编码无启动子Venus报告体的编辑元件的载体基因组。无启动子Venus含有位于剪接受体(SA)和2A序列(2A)下游的Venus开放阅读框(ORF)以允许独立蛋白质表达。左和右同源臂(HA)各自是800bp长,并且与位于染色体19上AAVS1基因座中的人类PPP1R12C内含子1的序列互补。AAVS1基因座被视为用于异源转基因插入的安全港位点。在AAV2反向末端重复(ITR)之间克隆含有同源臂、Venus ORF和调节序列的编辑元件。

[0284] 细胞培养.所有细胞系都在5%CO2的潮湿氛围中在37℃下生长,并且如下培养:SCID-X1淋巴母细胞(Coriell)在补充有15%胎牛血清(FBS)(Gibco,目录号26140)的RPMI1640(Gibco,目录号21875)中培养;K562细胞(ATCC)在补充有10%FBS(Gibco,目录号26140)和1%L-谷氨酰胺(Gibco,目录号25030)的DMEM(Corning,目录号15-017-CVR)中培养;HepG2细胞(ATCC)在补充有10%FBS(Gibco,目录号26140)的EMEM(ATCC,目录号30-2003)中培养;MCF-7细胞(ATCC)在补充有10%FBS(Gibco,目录号26140)、1%MEM非必需氨基酸(Gibco,目录号11140)、1%丙酮酸钠(Gibco,目录号11360)和10μg/ml人类重组胰岛素(Gibco,目录号12585-014)的MEM(Gibco,目录号11095)中培养;WI-38成纤维细胞(ATCC)和HEK293细胞(ATCC)在补充有10%FBS(Gibco,目录号26140)、1%MEM非必需氨基酸(Gibco,目录号11140)和1%丙酮酸钠(Gibco,目录号11360)的MEM(Gibco,目录号11095)中培养;并且Y79细胞(ATCC)在补充有20%FBS(Gibco,目录号26140)的RPMI1640(Gibco,目录号A10491)中培养。

[0285] AAVF介导的人类细胞系编辑.在第0天,对96孔形式以20,000个细胞/0.1毫升或对24孔形式以20,000个细胞/0.5毫升接种黏附细胞。24小时后(在第1天),在添加AAVF载体之前测量细胞计数。在第1天,对96孔形式以20,000个细胞/0.1毫升或对24孔形式以20,000个细胞/0.5毫升接种悬浮细胞。在第1天,使用150,000个载体基因组(VG)/细胞的感染复数(MOI)向细胞中添加AAVF。在AAVF添加之前,用硫酸鱼精蛋白(10mg/ml)涂布用于转移的移液管尖端。在即将转导之前,在涡流混合器上以全速持续30秒来使AAVF充分地悬浮。向细胞中添加的AAVF的体积不超过孔中总体积的5%。在第3天,采集细胞(使黏附细胞轻度胰蛋白酶化以将其从组织培养板去除),洗涤,并且使用装配有Hypercyt自动进样器的Intellicyt流式细胞仪通过流式细胞测量术分析其Venus表达。编辑表示为Venus阳性的总细胞群体减去未转导的细胞中所观察到的背景荧光(通常小于1%Venus阳性细胞)的百分比。在不使用外源性核酸酶的情况下进行AAVF编辑实验。

[0286] 结果

[0287] 所测试的所有AAVF都观察到人类PPP1R12c基因座的编辑,并且展示细胞类型选择性(表3)。一般来说,在感染48小时之后,AAVF5在每一个细胞系中产生最高水平的编辑,为总细胞群体的12%-45%。AAVF9也在B淋巴母细胞细胞系(SCID-X1LBL和K562)中产生高水平的基因编辑。AAVF17在正常人类二倍体成纤维细胞(WI-38细胞)中在这些条件下产生最高水平的基因编辑。AAVF1、AAVF4、AAVF7载体产生的编辑水平一致地大于未转导的细胞中所见但一般低于在用AAVF5、AAVF9和AAVF17的情况下所观察到的最大水平。这些数据展现,AAVF具有用于人类组织的宽向性,并且可以尤其适用于在肝脏、CNS(例如视网膜)、组织成

纤维细胞、乳房和淋巴细胞中进行基因编辑。

[0288] 表3.通过AAVF编辑人类细胞系

[0289]

	人类细胞系					
载体	K562	Hep-G2	WI-38	Y79	MCF7	SCID-X1 LBL
	Venus 阳性细胞百分比					
AAVF1	2.06	1.92	1.39	0.7	2.71	0.57
AAVF4	1.36	2.95	6.29	0.35	3.42	0.12
AAVF5	45.32	-	11.85	12.5	20.47	45.02
AAVF7	5.78	3.83	4.87	1.42	3.6	0.94
AAVF9	18.06	-	-	2.46	6.14	6.79
AAVF17	3.18	3.71	17.6	1.29	5.66	0.44

[0290] 实例6:原代人类细胞中的AAVF编辑

[0291] 为了评定AAVF载体对原代人类细胞的编辑,使用人类肝细胞、肝窦状隙内皮细胞和骨骼肌成肌细胞的原代培养物。

[0292] 所用AAVF和AAV载体各自包装编码无启动子Venus报告体的载体基因组。插入物包含位于剪接受体(SA)和2A序列(2A)下游的Venus开放阅读框(ORF)以允许独立蛋白质表达。左和右同源臂(HA)各自是800bp长,并且与位于染色体19上AAVS1基因座中的人类PPP1R12C内含子1的序列互补。AAVS1基因座被视为用于异源转基因插入的安全港位点。在AAV2反向末端重复(ITR)之间克隆由同源臂、开放阅读框和调节序列组成的校正基因组。

[0293] 方法

[0294] 细胞培养.所有原代人类细胞都在37℃下,在5%潮湿CO2下,在组织培养恒温箱中培养。除非另外规定,否则所有材料和培养基组分都是从Life Technologies获得。

[0295] 从Invitrogen获得人类肝细胞的原代培养物,并且如由制造商所建议,在经过I型胶原蛋白涂布的板上进行培养。在解冻/接种培养基[每35.0mL毫升体积中,32.5mL威廉氏E培养基(William's E Medium)(#A12176),1.6mL胎牛血清,3.2μL的10mM地塞米松(Dexamethazone)和0.9mL接种混合液A]中由液氮中的储备回收细胞。接种混合液A由以下组成:每1.8mL最终体积中,0.5mL青霉素(10,000U/mL)/链霉素(10,000μg/mL)溶液(目录号15140)、0.05mL的4.0mg人类重组胰岛素/mL(产品目录号12585-014)、0.5mL的200mM GlutaMAXTM溶液(产品目录号35050)和0.75mL的1.0M Hepes,pH7.4(产品目录号15630)。将人类肝细胞维持在维持培养基中,所述维持培养基每103.6mL最终体积含有100mL威廉姆斯E培养基、0.001mL地塞米松和3.6mL维持混合液B(产品目录号A13448)。

[0296] 人类骨骼肌成肌细胞的原代培养物从Lonza获得,并且如制造商所描述在SkGMTM培养基中培养。

[0297] SkGM™-2Bullit™试剂盒(Lonza,目录号CC-3245)在500mL SkGM-2培养基(#0000482653)中含有0.5mL人类表皮生长因子[hEGF](#0000482653)、0.5mL地塞米松(#0000474738)、10mLL-谷氨酰胺(#0000474740)、50mL胎牛血清(#0000474741)、0.5mL庆大霉素(Gentamicin)/两性霉素-B(Amphotericin-B)[GA](#0000474736)。

[0298] 人类肝窦状隙内皮细胞的原代培养物购自Creative Bioarray,并且如制造商所描述,在SuperCult®内皮细胞培养基中培养并在基于明胶的包覆层上生长。SuperCult®内皮细胞培养基补充剂试剂盒(产品目录号ECM-500Kit)含有0.5mL VEGF(#15206)、0.5mL肝

素(#15250)、0.5mL EGF(#15217)、0.5mL FGF(#15204)、0.5mL氢皮质酮(#15318)、5.0mL抗生-抗霉菌溶液(#15179)、5.0mL L-谷氨酰胺(#15409)、10.0mL内皮细胞补充剂(#15604)、50.0mL FBS(#15310)以及500.0mLSuperCult®内皮细胞培养基(#15517)。

[0299] AAVF介导的原代人类细胞编辑.在第0天,对96孔形式以2×10^4个细胞/0.1毫升或对24孔形式以2×10^4个细胞/0.5毫升接种人类原代肝细胞。48小时后,在第2天如下文所描述添加病毒载体。在第1天,对96孔形式以2×10^4个细胞/0.1毫升或对24孔形式以2×10^4个细胞/0.5毫升接种人类骨骼肌成肌细胞和人类肝窦状隙内皮细胞。在第2天,以150,000VG(载体基因组)/细胞的感染复数(MOI)向这些细胞中添加病毒载体(对AAVF5,使用5×10^4VG/细胞的MOI,并且对AAVF17,使用2.5×10^4VG/细胞的MOI)。在向细胞中添加载体之前,用硫酸鱼精蛋白(10mg/mL)涂布用于载体转移的所有移液管尖端,并且在即将转导之前,通过涡流30秒使载体充分地混合。向细胞中添加的载体的体积不超过孔中总体积的5%。

[0300] 在第3天更新用于人类原代肝细胞的培养基。在第4天,采集细胞(使黏附细胞胰蛋白酶化),并且使用装配有Hypercyt自动进样器的Intellicyt流式细胞仪通过流式细胞测量术分析其Venus表达。编辑表示为Venus阳性的总细胞群体减去未转导的细胞中所观察到的背景荧光(通常小于1%Venus阳性细胞)的百分比。

[0301] 结果

[0302] 所测试的所有AAVF都观察到人类PPP1R12c基因座的编辑,并且展示细胞类型选择性(表4)。一般来说,在感染48小时之后,AAVF5在每一个原代细胞群体中产生最高水平的编辑,分别在人类肝细胞中为较低的2%,在原代骨骼成肌细胞和肝窦状隙内皮细胞中为24%到35%。AAVF7和AAVF17也在原代肝窦状隙内皮细胞和骨骼成肌细胞中产生高水平的基因编辑。在这些细胞中使用AAVF载体的编辑水平是用AAV2或AAV6所观察到的10到50倍。AAVF载体产生的编辑水平一致地大于未转导的细胞中或用包装无启动子Venus载体基因组的AAV2或AAV6转导的细胞中所见。AAV2或AAV6在这些细胞中极少或未观察到编辑。原代人类细胞中的这些数据展现,AAVF5、AAVF7和AAVF17具有用于人类组织的宽向性,并且可以尤其适用于针对肝脏、骨骼肌和内皮细胞群体的基因疗法应用。

[0303] 对于比较,原代人类细胞群体中的每一个也都用包装mCherry的AAVF基因转移载体在鸡β肌动蛋白(CBA)启动子的控制下转导。然后对于AAV6、AAVF5和AAVF7,将Venus/mCherry的蛋白质表达比率用于估计编辑比率,其反映在各种细胞类型中具有可使用实验基因编辑载体(Venus)检测的蛋白质表达的细胞数目相对于具有可使用上述基因转移载体(mCherry)检测的蛋白质表达的细胞数目的比率。如表5中所示,在所研究的原代人类细胞中,AAVF介导的基因编辑对于蛋白质表达一般比相对应的AAVF介导的基因转移途径更有效,而在此类细胞中,AAV6介导的基因编辑的有效性与AAV6介导的基因转移基本上相同或略小。值得注意地,如也在表5中所示,对于所研究的原代人类细胞,AAVF介导的基因编辑高于AAV6介导的基因编辑所观察到的。这些数据展现,AAVF介导的基因编辑在多种原代人类细胞中比AAV6更高效,并且AAVF介导的基因编辑比此类细胞中相对应的基因转移有利。

[0304] 包装mCherry报告体的AAVF载体也有效转导原代人类细胞并且展示细胞类型特异性(表6)。对于人类脐静脉内皮细胞(HUVEC)和肝窦状隙内皮细胞(HSEC),用AAVF9转导在感染48小时之后产生38%-50%mCherry阳性细胞。AAVF9也高效地转导人类骨骼肌成肌细胞,并且在较小程度上,所测试的所有AAVF在这些条件下都有效转导HSEC(表6)。在不受理论束

缚的情况下，由于在mCherry载体基因组中不存在同源臂并且表达由CBA启动子驱动，故这些研究中的大部分(如果不是全部)mCherry表达很可能表示报告体的游离表达。

[0305] 表4.AAVF介导的原代人类细胞编辑

[0306]

载体	HSEC	肝细胞	SMM
	Venus 阳性细胞百分比		
AAV2	0.34	0.06	0.52
AAV6	0.00	0.00	0.00
AAVF5	34.70	2.00	24.20
AAVF7	18.40	0.62	18.65
AAVF17	21.40	1.37	19.60

[0307] 表5.AAVF在原代人类细胞中的编辑比率

[0308]

载体	肝细胞		SMM	
	Venus/mCherry	AAVF/AAV6	Venus/mCherry	AAVF/AAV6
AAV6	0.84		0.80	
AAVF5	9.42	11.21	14.32	17.90
AAVF7	0.95	1.10	13.05	16.30

[0309]

载体	HSEC	
	Venus/mCherry	AAVF/AAV6
AAV6	0.60	
AAVF5	5.30	8.80
AAVF7	2.70	4.50

[0310] HSEC＝肝窦状隙内皮细胞

[0311] SMM＝骨骼肌成肌细胞

[0312] 表6.用包装mCherry的AAVF载体转导原代人类细胞

[0313]

载体	HSEC	肝细胞	SMM	HUVEC
	mCherry 阳性细胞百分比			
AAVF1	7.04	0.13	1.65	1.47
AAVF4	9.10	0.05	1.72	0.61
AAVF5	4.00	0.01	1.03	0.71
AAVF7	4.43	0.70	0.77	2.15
AAVF9	38.24	1.05	29.27	50.13

[0314] HSEC＝肝窦状隙内皮细胞

[0315] SMM＝骨骼肌成肌细胞

[0316] HUVEC＝人类脐静脉内皮细胞

[0317] 实例7：AAVF相对基因编辑(HDR)对比基因转移(转导)效率的研究

[0318] 如由相对蛋白质表达所测量，两种类型的基于AAV的载体(CBA-mCherry基因转移载体和AAVS1-Venus基因编辑载体，图39)用于评定AAVF载体以及AAV2和AAV6的相对基因编辑效率对比基因转移转导效率。对于基因转移载体，CBA-mCherry构筑体包括处于鸡β肌动蛋白(CBA)启动子的控制下的mCherry基因，并且利用聚腺苷酸化信号，但不含有任何同源臂。对于基因编辑性载体，AAVS1-Venus构筑体包括无启动子Venus开放阅读框(ORF)，其中(HA-L)和右(HA-R)同源臂将所述Venus ORF靶向到染色体19AAVS1区内PPP1R12C基因的内含子1上。还存在位于Venus ORF上游的剪接受体(SA)和2A序列，其允许Venus转录物被剪接

出并且不依赖于PPP1R12C基因而表达。

[0319] 在原代人类脐血CD34+造血干细胞/祖细胞中比较AAVF、AAV2和AAV6载体介导基因编辑对比基因转移的能力。由多个供体合并所用细胞。用基因转移载体(AAV*-CBA-mCherry)或基因编辑载体(AAV*-Venus)以每细胞150,000个载体基因组(VG)的复数转导经过纯化的CD34+细胞。四十八小时后,采集细胞,并且通过流式细胞测量术来分析(图40A和40B)。所示数据包括扣除背景(未转导的细胞)。

[0320] 假设mCherry表达将表示转导效率,而Venus表达将表示基因编辑效率。在不受理论束缚的情况下,以下是AAV转导对比编辑的潜在机制的概述。在感染后,AAV结合于细胞表面受体并且内化,之后进行核转位并进入。在原子核中,AAV进行脱壳,并且释放载体基因组。不论载体基因组如何,对于相同细胞群体中的每个既定衣壳,这些过程很可能以相同速率进行。在脱壳后,单链CBA-mCherry基因组进行第二链合成,之后进行mCherry表达。另一方面,无启动子Venus编辑载体是针对染色体19上的互补基因组区。然后,由同源臂结合的SA/2A-Venus盒可以经由同源性依赖性修复(HDR)机制而重组到染色体中同源臂的内部接合点处。在此重组事件后,在经过成功编辑的细胞中表达Venus。

[0321] 对于除了AAVF9之外的所有载体,其中表达Venus的细胞比例比表达mCherry高得多(图40A和40C)。在用AAVF1、AAVF5、AAVF7、AAVF16和AAVF17转导后,Venus表达尤其高。相同衣壳导致低得多的mCherry表达水平(图40A和40C)。在用AAV2和AAV6转导后,观察到最小量的mCherry和Venus表达。也进行Venus与mCherry相对表达的比较。具体地说,测定编辑比率,其反映在各种细胞类型中具有可使用实验基因编辑载体(Venus)检测的蛋白质表达的细胞数目相对于具有可使用上述基因转移载体(mCherry)检测的蛋白质表达的细胞数目的比率。在针对病毒经由脱壳进入原子核内的过程归一化之后,编辑比率提供对每一AAV衣壳所介导编辑的相对效率的估计。除AAVF9以外,所测试的所有AAV载体都展现基因编辑(Venus)比基因转移/转基因表达(mCherry)更高效(图40D和表7)。AAVF5和AAVF7显示最高编辑比率(图40D和表7)。也将AAVF介导的基因编辑(Venus)相对于AAV2和AAV6介导的基因编辑(Venus)进行比较(表7,其展示AAVF的Venus:mCherry比率除以AAV2或AAV6的相同比率)。AAVF基因编辑性构筑体的基因编辑有效性相对于AAV2和AAV6基因编辑性构筑体有利。AAVF5和AAVF7的编辑比率是所比较载体中最高的,指示这些载体尤其介导高效编辑。

[0322] 表7.CD34+CB细胞中的编辑与转导比率

[0323]

载体	比率 编辑:转导 (Venus:mCherry)	Ratio AAVF 编辑:AAV6 编辑	Ratio AAVF 编辑:AAV2 编辑
AAVF1	4.32	1.23	3.02
AAVF4	3.42	0.98	0.98
AAVF5	16.60	4.74	11.62
AAVF7	14.66	4.19	10.26
AAVF9	0.73	0.21	0.51
AAV2	1.43	0.41	1.00
AAV6	3.50	1.00	2.45

[0324] 本发明的范围不受打算作为本发明几个方面的说明的实例中所公开的具体实施例限制,并且功能上等效的任何实施例都处于本发明的范围内。实际上,除本文所展示和所描述的那些修改以外,本发明的各种修改对于所属领域的技术人员将是显而易见的,并且

打算属于所附权利要求书的范围内。

[0325] 参考文献

[0326] 1.BAINBRIDGE,J.W.,SMITH,A.J.,BARKER,S.S.,ROBBIE,S.,HENDERSON,R.,BALAGGAN,K.,VISWANATHAN,A.,HOLDER,G.E.,STOCKMAN,A.,TYLER,N.,PETERSEN-JONES,S.,BHATTACHARYA,S.S.,THRASHER,A.J.,FITZKE,F.W.,CARTER,B.J.,RUBIN,G.S.,MOORE,A.T.以及ALI,R.R.(2008).基因疗法对雷伯氏先天性黑蒙中视觉功能的影响(Effect of gene therapy on visual function in Leber's congenital amaurosis).《新英格兰医学杂志(N Engl J Med)》358,2231-2239.

[0327] 2.BATCHU,R.B.,SHAMMAS,M.A.,WANG,J.Y.,FREEMAN,J.,ROSEN,N.以及MUNSHI,N.C.(2002).腺相关病毒保护成视网膜细胞瘤蛋白质家族免受腺病毒诱导的功能失活影响(Adeno-associated virus protects the retinoblastoma family of proteins from adenoviral-induced functional inactivation).《癌症研究(Cancer Res)》62,2982-2985.

[0328] 3.BELL,P.,WANG,L.,LEBHERZ,C.,FLIEDER,D.B.,BOVE,M.S.,WU,D.,GAO,G.P.,WILSON,J.M.以及WIVEL,N.A.(2005).在大规模小鼠研究中没有AAV载体致瘤作用的证据(No evidence for tumorigenesis of AAV vectors in a large-scale study in mice).《分子治疗学(Mol Ther)》12,299-306.

[0329] 4.BERNS,K.I.和GIRAUD,C.(1996).腺相关病毒的生物学(Biology of adeno-associated virus).《微生物学与免疫学最新课题(Curr Top Microbiol Immunol)》218,1-23.

[0330] 5.BIFFI,A.和CESANI,M.(2008).基因疗法中的人类造血干细胞:临床前和临床问题(Human hematopoietic stem cells in gene therapy:preclinical and clinical issues).《当今基因疗法(Curr Gene Ther)》8,135-146.

[0331] 6.BRANTLY,M.L.,CHULAY,J.D.,WANG,L.,MUELLER,C.,HUMPHRIES,M.,SPENCER,L.T.,ROUHANI,F.,CONLON,T.J.,CALCEDO,R.,BETTS,M.R.,SPENCER,C.,BYRNE,B.J.,WILSON,J.M.以及FLOTTE,T.R.(2009).在rAAV1-AAT基因疗法临床试验中尽管存在T淋巴细胞反应但仍持续的转基因表达(Sustained transgene expression despite T lymphocyte responses in a clinical trial of rAAV1-AAT gene therapy).《美国国家科学院院刊(Proc Natl Acad Sci U S A)》.

[0332] 7.CHATTERJEE,S.,JOHNSON,P.R.以及WONG,K.K.,JR.(1992).借助于腺相关病毒反义载体的体外HIV-1双目标抑制(Dual-target inhibition of HIV-1in vitro by means of an adeno-associated virus antisense vector).《科学(Science)》258,1485-1488.

[0333] 8.CHATTERJEE,S.,LI,W.,WONG,C.A.,FISHER-ADAMS,G.,LU,D.,GUHA,M.,MACER,J.A.,FORMAN,S.J.以及WONG,K.K.,JR.(1999).用腺相关病毒载体对原始人类骨髓和脐血源性造血祖细胞进行转导(Transduction of primitive human marrow and cord blood-derived hematopoietic progenitor cells with adeno-associated virus vectors).《血液(Blood)》93,1882-1894.

[0334] 9.CHATTERJEE,S.,WONG,KK.(1993).用于反义RNA递送的腺相关病毒载体(Adeno-

Associated Viral Vectors for the Delivery of Antisense RNA).《方法-伦敦-酶学方法伴随刊物(METHODS-LONDON-A COMPANION TO METHODS IN ENZYMOLOGY)》-5,1.

[0335] 10.CIDECIYAN,A.V.,HAUSWIRTH,W.W.,ALEMAN,T.S.,KAUSHAL,S.,SCHWARTZ,S.B.,BOYE,S.L.,WINDSOR,E.A.,CONLON,T.J.,SUMAROKA,A.,PANG,J.J.,ROMAN,A.J.,BYRNE,B.J.以及JACOBSON,S.G.(2009).用于雷伯氏先天性黑蒙的人类RPE65基因疗法:1年时的早期视力改善和安全性持续性(Human RPE65gene therapy for Leber congenital amaurosis:persistence of early visual improvements and safety at 1year).《人类基因疗法(Hum Gene Ther)》20,999-1004.

[0336] 11.EINERHAND,M.P.,ANTONIOU,M.,ZOLOTUKHIN,S.,MUZYCZKA,N.,BERNS,K.I.,GROSVELD,F.以及VALERIO,D.(1995).在重组腺相关病毒介导的基因转移后在红细胞中调节的高水平人类β-血球蛋白基因表达(Regulated high-level human beta-globin gene expression in erythroid cells following recombinant adeno-associated virus-mediated gene transfer).《基因疗法》2,336-343.

[0337] 12.FISHER-ADAMS,G.,WONG,K.K.,JR.,PODSAKOFF,G.,FORMAN,S.J.以及CHATTERJEE,S.(1996).腺相关病毒载体在转导之后在CD34+人类造血祖细胞中的整合(Integration of adeno-associated virus vectors in CD34+human hematopoietic progenitor cells after transduction).《血液》88,492-504.

[0338] 13.FLOTTE,T.R.,BRANTLY,M.L.,SPENCER,L.T.,BYRNE,B.J.,SPENCER,C.T.,BAKER,D.J.以及HUMPHRIES,M.(2004).将重组腺相关病毒α1-抗胰蛋白酶(rAAV2-CB-hAAT)基因载体肌肉内注射到AAT缺陷型成人中的I期试验(Phase I trial of intramuscular injection of a recombinant adeno-associated virus alpha 1-antitrypsin(rAAV2-CB-hAAT)gene vector to AAT-deficient adults).《人类基因疗法》15,93-128.

[0339] 14.GAO,G.,VANDENBERGHE,L.H.,ALVIRA,M.R.,LU,Y.,CALCEDO,R.,ZHOU,X.以及WILSON,J.M.(2004).腺相关病毒的分化体在人类组织中广泛散播(Clades of Adeno-associated viruses are widely disseminated in human tissues).《病毒学杂志(J Virol)》78,6381-6388.

[0340] 15.HACEIN-BEY-ABINA,S.,VON KALLE,C.,SCHMIDT,M.,LE DEIST,F.,WULFFRAAT,N.,MCINTYRE,E.,RADFORD,I.,VILLEVAL,J.L.,FRASER,C.C.,CAVAZZANA-CALVO,M.以及FISCHER,A.(2003).在针对X连锁重症综合性免疫缺陷的成功基因疗法之后的严重不良事件(A serious adverse event after successful gene therapy for X-linked severe combined immunodeficiency).《新英格兰医学杂志》348,255-256.

[0341] 16.HAN,Z.,ZHONG,L.,MAINA,N.,HU,Z.,LI,X.,CHOUTHAI,N.S.,BISCHOF,D.,WEIGEL-VAN AKEN,K.A.,SLAYTON,W.B.,YODER,M.C.以及SRIVASTAVA,A.(2008).重组腺相关病毒载体基因组在鼠类造血干细胞转导之后的稳定整合(Stable integration of recombinant adeno-associated virus vector genomes after transduction of murine hematopoietic stem cells).《人类基因疗法》19,267-278.

[0342] 17.JAYANDHARAN,G.R.,ZHONG,L.,LI,B.,KACHNIARZ,B.以及SRIVASTAVA,A.(2008).用于改善单链腺相关病毒载体的体外和体内转导效率的策略(Strategies for improving the transduction efficiency of single-stranded adeno-associated

virus vectors in vitro and in vivo).《基因疗法》15,1287-1293.

[0343] 18.KAPLITT,M.G.,FEIGIN,A.,TANG,C.,FITZSIMONS,H.L.,MATTIS,P.,LAWLOR,P.A.,BLAND,R.J.,YOUNG,D.,STRYBING,K.,EIDELBERG,D.以及DURING,M.J.(2007).使用腺相关病毒(AAV)所携带GAD基因用于帕金森氏病的基因疗法的安全性和可容忍性:开放标签的I期试验(Safety and tolerability of gene therapy with an adeno-associated virus(AAV)borne GAD gene for Parkinson's disease:an open label,phase I trial).《柳叶刀(Lancet)》369,2097-2105.

[0344] 19.KELLS,A.P.,HADACZEK,P.,YIN,D.,BRINGAS,J.,VARENIKA,V.,FORSAYETH,J.以及BANKIEWICZ,K.S.(2009).基于高效基因疗法的用于向灵长类动物皮质递送治疗剂的方法(Efficient gene therapy-based method for the delivery of therapeutics to primate cortex).《美国国家科学院院刊》106,2407-2411.

[0345] 20.KESSLER,P.D.,PODSAKOFF,G.M.,CHEN,X.,MCQUISTON,S.A.,COLOSI,P.C.,MATELIS,L.A.,KURTZMAN,G.J.以及BYRNE,B.J.(1996).向骨骼肌的基因递送引起治疗性蛋白质的持续表达和全身递送(Gene delivery to skeletal muscle results in sustained expression and systemic delivery of a therapeutic protein).《美国国家科学院院刊》93,14082-14087.

[0346] 21.MANNO,C.S.,CHEW,A.J.,HUTCHISON,S.,LARSON,P.J.,HERZOG,R.W.,ARRUDA,V.R.,TAI,S.J.,RAGNI,M.V.,THOMPSON,A.,OZELO,M.,COUTO,L.B.,LEONARD,D.G.,JOHNSON,F.A.,MCCLELLAND,A.,SCALLAN,C.,SKARSGARD,E.,FLAKE,A.W.,KAY,M.A.,HIGH,K.A.以及GLADER,B.(2003).在患有严重B型血友病的患者中AAV介导的向骨骼肌中的因子IX基因转移(AAV-mediated factor IX gene transfer to skeletal muscle in patients with severe hemophilia B).《血液》101,2963-2972.

[0347] 22.MCCORMACK,M.P.和RABBITTS,T.H.(2004).T细胞致癌基因LMO2在用于X连锁重症综合性免疫缺陷的基因疗法之后的活化(Activation of the T-cell oncogene LMO2after gene therapy for X-linked severe combined immunodeficiency).《新英格兰医学杂志》350,913-922.

[0348] 23.MILLER,D.G.,ADAM,M.A.以及MILLER,A.D.(1990).反转录病毒载体的基因转移仅在感染时主动复制的细胞中进行(Gene transfer by retrovirus vectors occurs only in cells that are actively replicating at the time of infection).《分子与细胞生物学(Mol Cell Biol)》10,4239-4242.

[0349] 24.PAZ,H.,WONG,C.A.,LI,W.,SANTAT,L.,WONG,K.K.以及CHATTERJEE,S.(2007).人类CD34阳性造血干细胞的静止亚群是用于稳定重组2型腺相关病毒转导的优选目标(Quiescent subpopulations of human CD34-positive hematopoietic stem cells are preferred targets for stable recombinant adeno-associated virus type 2transduction).《人类基因疗法》18,614-626.

[0350] 25.PETRS-SILVA,H.,DINCULESCU,A.,LI,Q.,MIN,S.H.,CHIODO,V.,PANG,J.J.,ZHONG,L.,ZOLOTUKHIN,S.,SRIVASTAVA,A.,LEWIN,A.S.以及HAUSWIRTH,W.W.(2009).酪氨酸突变型AAV血清型载体对小鼠视网膜的高效转导(High-efficiency transduction of the mouse retina by tyrosine-mutant AAV serotype vectors).《分子治疗学》17,463-

471.

[0351] 26.PODSAKOFF,G.,WONG,K.K.,JR.以及CHATTERJEE,S.(1994).基于腺相关病毒的载体向非分裂细胞中的高效基因转移(Efficient gene transfer into nondividing cells by adeno-associated virus-based vectors).《病毒学杂志》68,5656-5666.

[0352] 27.PONNAZHAGAN,S.,YODER,M.C.以及SRIVASTAVA,A.(1997).具有长期再生能力和体内人类血球蛋白基因持续表达的2型腺相关病毒介导的鼠类造血细胞转导(Adeno-associated virus type 2-mediated transduction of murine hematopoietic cells with long-term repopulating ability and sustained expression of a human globin gene in vivo).《病毒学杂志》71,3098-3104.

[0353] 28.RAJ,K.,OGSTON,P.以及BEARD,P.(2001).病毒介导杀灭不具有p53活性的细胞(Virus-mediated killing of cells that lack p53 activity).《自然(Nature)》412,914-917.

[0354] 29.SANTAT,L.,PAZ,H.,WONG,C.,LI,L.,MACER,J.,FORMAN,S.,WONG,K.K.以及CHATTERJEE,S.(2005).能够连续植入免疫缺陷型小鼠中的原始人类造血干细胞的重组AAV2转导(Recombinant AAV2transduction of primitive human hematopoietic stem cells capable of serial engraftment in immune-deficient mice).《美国国家科学院院刊》102,11053-11058.

[0355] 30.SRIVASTAVA,A.(2004).AAV2载体向人类和鼠类原始造血干细胞和祖细胞中的基因递送(Gene delivery to human and murine primitive hematopoietic stem and progenitor cells by AAV2vectors).《分子生物学方法(分子生物学方法)》246,245-254.

[0356] 31.TOWNE,C.,SCHNEIDER,B.L.,KIERAN,D.,REDMOND,D.E.,JR.以及AEBISCHER,P.(2009).非人类灵长类动物运动神经元在肌肉内递送血清6型重组AAV之后的高效转导(Efficient transduction of non-human primate motor neurons after intramuscular delivery of recombinant AAV serotype 6).《基因疗法》.

[0357] 32.ZHONG,L.,CHEN,L.,LI,Y.,QING,K.,WEIGEL-KELLEY,K.A.,CHAN,R.J.,YODER,M.C.以及SRIVASTAVA,A.(2004a).自互补腺相关病毒2(AAV)-T细胞蛋白质酪氨酸磷酸酶载体作为辅助病毒以改良常规单链AAV载体的体外和体内转导效率(Self-complementary adeno-associated virus 2(AAV)-T cell protein tyrosine phosphatase vectors as helper viruses to improve transduction efficiency of conventional single-stranded AAV vectors in vitro and in vivo).《分子治疗学》10,950-957.

[0358] 33.ZHONG,L.,LI,B.,JAYANDHARAN,G.,MAH,C.S.,GOVINDASAMY,L.,AGBANDJEMCKENNA,M.,HERZOG,R.W.,WEIGEL-VAN AKEN,K.A.,HOBBS,J.A.,ZOLOTUKHIN,S.,MUZYCZKA,N.以及SRIVASTAVA,A.(2008a).AAV2载体的酪氨酸磷酸化和其对病毒细胞内运输和转基因表达的影响(Tyrosine phosphorylation of AAV2vectors and its consequences on viral intracellular trafficking and transgene expression).《病毒学(Virology)》381,194-202.

[0359] 34.ZHONG,L.,LI,B.,MAH,C.S.,GOVINDASAMY,L.,AGBANDJE-MCKENNA,M.,COOPER,M.,HERZOG,R.W.,ZOLOTUKHIN,I.,WARRINGTON,K.H.,JR.,WEIGEL-VAN AKEN,K.A.,HOBBS,J.A.,ZOLOTUKHIN,S.,MUZYCZKA,N.以及SRIVASTAVA,A.(2008b).下一代腺相关病毒2载体：

酪氨酸中的点突变以较低剂量引起高效转导(Next generation of adeno-associated virus 2vectors:point mutations in tyrosines lead to high-efficiency transduction at lower doses).《美国国家科学院院刊》105,7827-7832.

[0360] 35.ZHONG,L.,LI,W.,YANG,Z.,QING,K.,TAN,M.,HANSEN,J.,LI,Y.,CHEN,L.,CHAN,R.J.,BISCHOF,D.,MAINA,N.,WEIGEL-KELLEY,K.A.,ZHAO,W.,LARSEN,S.H.,YODER,M.C.,SHOU,W.以及SRIVASTAVA,A.(2004b).核转运和脱壳受损限制重组腺相关病毒2载体介导的原代鼠类造血细胞转导(Impaired nuclear transport and uncoating limit recombinant adeno-associated virus 2vector-mediated transduction of primary murine hematopoietic cells).《人类基因疗法》15,1207-1218.

[0361] 36.ZHONG,L.,ZHAO,W.,WU,J.,LI,B.,ZOLOTUKHIN,S.,GOVINDASAMY,L.,AGBANDJEMCKENNA,M.以及SRIVASTAVA,A.(2007).EGFR蛋白质酪氨酸激酶信号传导在AAV2衣壳泛素化和病毒第二链DNA合成中的双重作用(A dual role of EGFR protein tyrosine kinase signaling in ubiquitination of AAV2capsids and viral second-strand DNA synthesis).《分子治疗学》15,1323-1330.

[0362] 37.ZHOU,S.Z.,BROXMEYER,H.E.,COOPER,S.,HARRINGTON,M.A.以及SRIVASTAVA,A.(1993).腺相关病毒2在鼠类造血祖细胞中介导的基因转移(Adeno-associated virus 2-mediated gene transfer in murine hematopoietic progenitor cells).《实验血液学(Exp Hematol)》21,928-933.

[0363] 38.KHAN,I.G.,HIRATA,R.K.,RUSSELL,D.W.(2011)用于人类细胞的AAV介导的基因靶向方法(AAV-mediated gene targeting methods for human cells).《自然·方案(Nat Protoc)》.4,482-501.

[0364] 39.KHOTIN,R.M.,LINDEN,R.M.以及BERNS,K.I.(1992).用于通过非同源重组整合腺相关病毒DNA的人类染色体19q优选位点的表征(Characterization of a preferred site on human chromosome 19q for integration of adeno-associated virus DNA by non-homologous recombination).EMBO 13,5071-8.

[0365] 40.GIRAUD,C.,WINOCOUR,E.以及BERNS,K.I.(1994).腺相关病毒的位点特异性整合由细胞DNA序列引导(Site-specific integration by adeno-associated virus is directed by a cellular DNA sequence).《美国国家科学院院刊》21,10039-43.

[0366] 41.LINDEN,R.M.,WARD,P.,GIRAUD,C.,WINOCOUR,E.以及BERNS,K.I.(1996A).腺相关病毒的位点特异性整合(Site-specific integration by adeno-associated virus).《美国国家科学院院刊》21,11288-94.

[0367] 42.LINDEN,R.M.,WINOCOUR,E.以及BERNS,K.I.(1996B).用于腺相关病毒位点特异性整合的重组信号(The recombination signals for adeno-associated virus site-specific integration).《美国国家科学院院刊》15,7966-72.

[0368] 43.XU,L.,O'MALLEY,T.,SANDS,M.S.,WANG,B.,MEYERROSE,T.,HASKINS,M.E.以及PONDER,K.P(2004).在小鼠中新生静脉内注射两性反转录病毒载体之后的造血干细胞体内转导(In Vivo Transduction of Hematopoietic Stem Cells After Neonatal Intravenous Injection of an Amphotrophic Retroviral Vector in Mice).《分子治疗学》7月;10(1):37-44.

[0369] 44.WANG,C.X.,SATHER,B.D.,WANG,X.,ADAIR,J.,KHAN,I.,SINGH,S.,LANG,S.,ADAMS,A.,CURINGA,G.,KIEM,H-P.,MIAO,C.H.,RAWLINGS,D.J.以及TORBETT,B.E.(2014).雷帕霉素在人类和小鼠造血干细胞中减轻豆状病毒载体转导抗性(Rapamycin relieves lentiviral vector transduction resistance in human and mouse hematopoietic stem cells).《血液》.8月;124(6):913-23.

[0370] 45.CARBONARO,D.A.,JIN,X.,PETERSEN,D.,WANG,X.,DOREY,F.,KIL,K.S.,ALDRICH,M.,BLACKBURN,M.R.,KELLEMS,R.E.以及KOHN,D.B.(2006)通过向新生ADA基因剔除小鼠中静脉内注射表达人类ADA的豆状病毒载体来进行的体内转导:用于ADA不足的新颖酶替换疗法形式(In Vivo Transduction by Intravenous Injection of a Lentiviral Vector Expressing Human ADA into Neonatal ADA Gene Knockout Mice:A Novel Form of Enzyme Replacement Therapy for ADA Deficiency).《分子治疗学》6月;13(6):1110-20.

[0371] 46.SMITH,L.J.,UL-HASAN,T.,CARVAINES,S.K.,VAN VLIET,K.,YANG,E.,WONG,JR,K.K.,AGBANDJE-MCKENNA,M.以及CHATTERJEE,S.(2014)人类干细胞源性AAV的基因转移特性和结构模型化(Gene Transfer Properties and Structural Modeling of Human Stem Cell-derived AAV).《分子治疗学》9月;22(9):1625-1634.

序列表

<110> CITY OF HOPE
CHATTERJEE, Saswati
SMITH, Laura
WONG, Kamehameha

<120> 用于高效基因组编辑的腺相关病毒载体变异体和其方法

<130> 054435-8138.W000

<150> US 62/209,862
<151> 2015-08-25

<150> US 62/063,587
<151> 2014-10-14

[0001] <150> US 62/054,899
<151> 2014-09-24

<160> 40

<170> PatentIn version 3.5

<210> 1
<211> 736
<212> PRT
<213> 人工序列

<220>
<223> 腺相关 AAV9

<400> 1

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

1 5 10 15

Glu Gly Ile Arg Glu Trp Trp Ala Leu Lys Pro Gly Ala Pro Gln Pro

20 25 30

Lys Ala Asn Gln Gln His Gln Asp Asn Ala Arg Gly Leu Val Leu Pro

35 40 45

Gly Tyr Lys Tyr Leu Gly Pro Gly Asn Gly Leu Asp Lys Gly Glu Pro

50 55 60

Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp

65 70 75 80

Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala

85 90 95

Asp Ala Glu Phe Gln Glu Arg Leu Lys Glu Asp Thr Ser Phe Gly Gly

100 105 110

Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Leu Leu Glu Pro

115 120 125

Leu Gly Leu Val Glu Glu Ala Ala Lys Thr Ala Pro Gly Lys Lys Arg

130 135 140

[0002] Pro Val Glu Gln Ser Pro Gln Glu Pro Asp Ser Ser Ala Gly Ile Gly

145 150 155 160

Lys Ser Gly Ala Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr

165 170 175

Gly Asp Thr Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro

180 185 190

Ala Ala Pro Ser Gly Val Gly Ser Leu Thr Met Ala Ser Gly Gly Gly

195 200 205

Ala Pro Val Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser Ser

210 215 220

Ser Gly Asn Trp His Cys Asp Ser Gln Trp Leu Gly Asp Arg Val Ile

225 230 235 240

Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu

245 250 255

Tyr Lys Gln Ile Ser Asn Ser Thr Ser Gly Gly Ser Ser Asn Asp Asn

260 265 270

Ala Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg

275 280 285

Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn
290 295 300
Asn Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile
305 310 315 320
Gln Val Lys Glu Val Thr Asp Asn Asn Gly Val Lys Thr Ile Ala Asn
325 330 335
Asn Leu Thr Ser Thr Val Gln Val Phe Thr Asp Ser Asp Tyr Gln Leu
340 345 350
Pro Tyr Val Leu Gly Ser Ala His Glu Gly Cys Leu Pro Pro Phe Pro
355 360 365
Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asp
370 375 380
Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe
385 390 395 400
Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Gln Phe Ser Tyr Glu
405 410 415
Phe Glu Asn Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu
[0003] 420 425 430
Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser
435 440 445
Lys Thr Ile Asn Gly Ser Gly Gln Asn Gln Gln Thr Leu Lys Phe Ser
450 455 460
Val Ala Gly Pro Ser Asn Met Ala Val Gln Gly Arg Asn Tyr Ile Pro
465 470 475 480
Gly Pro Ser Tyr Arg Gln Gln Arg Val Ser Thr Thr Val Thr Gln Asn
485 490 495
Asn Asn Ser Glu Phe Ala Trp Pro Gly Ala Ser Ser Trp Ala Leu Asn
500 505 510
Gly Arg Asn Ser Leu Met Asn Pro Gly Pro Ala Met Ala Ser His Lys
515 520 525
Glu Gly Glu Asp Arg Phe Phe Pro Leu Ser Gly Ser Leu Ile Phe Gly
530 535 540
Lys Gln Gly Thr Gly Arg Asp Asn Val Asp Ala Asp Lys Val Met Ile
545 550 555 560
Thr Asn Glu Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr Glu Ser

565 570 575

Tyr Gly Gln Val Ala Thr Asn His Gln Ser Ala Gln Ala Gln Ala Gln

580 585 590

Thr Gly Trp Val Gln Asn Gln Gly Ile Leu Pro Gly Met Val Trp Gln

595 600 605

Asp Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His

610 615 620

Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Met

625 630 635 640

Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala

645 650 655

Asp Pro Pro Thr Ala Phe Asn Lys Asp Lys Leu Asn Ser Phe Ile Thr

660 665 670

Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln

675 680 685

Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn

690 695 700

[0004] Tyr Tyr Lys Ser Asn Asn Val Glu Phe Ala Val Asn Thr Glu Gly Val

705 710 715 720

Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu

725 730 735

<210> 2

<211> 736

<212> PRT

<213> 人工序列

<220>

<223> 新颖AAV分离株

<400> 2

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

1 5 10 15

Glu Gly Ile Arg Glu Trp Trp Ala Leu Lys Pro Gly Ala Pro Gln Pro

20 25 30

Lys Ala Asn Gln Gln His Gln Asp Asn Ala Arg Gly Leu Val Leu Pro
35 40 45
Gly Tyr Lys Tyr Leu Gly Pro Gly Asn Gly Leu Asp Lys Gly Glu Pro
50 55 60
Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp
65 70 75 80
Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala
85 90 95
Asp Ala Glu Phe Gln Glu Arg Leu Lys Glu Asp Thr Ser Phe Gly Gly
100 105 110
Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Leu Leu Glu Pro
115 120 125
Leu Gly Leu Val Glu Glu Ala Ala Lys Thr Ala Pro Gly Lys Lys Arg
130 135 140
Pro Val Glu Gln Ser Pro Gln Glu Pro Asp Ser Ser Ala Gly Ile Gly
145 150 155 160
Lys Ser Gly Ala Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr
[0005] 165 170 175
Gly Asp Thr Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro
180 185 190
Ala Ala Pro Ser Gly Val Gly Ser Leu Thr Met Ala Ser Gly Gly Gly
195 200 205
Ala Pro Val Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser Ser
210 215 220
Ser Gly Asn Trp His Cys Asp Ser Gln Trp Leu Gly Asp Arg Val Ile
225 230 235 240
Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu
245 250 255
Tyr Lys Gln Ile Ser Asn Ser Thr Ser Gly Gly Ser Ser Asn Asp Asn
260 265 270
Ala Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg
275 280 285
Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn
290 295 300
Asn Trp Gly Phe Arg Pro Lys Gln Leu Asn Phe Lys Leu Phe Asn Ile

305 310 315 320

Gln Val Lys Glu Val Thr Asp Asn Asn Gly Val Lys Thr Ile Ala Asn

325 330 335

Asn Leu Thr Ser Thr Val Gln Val Phe Thr Asp Ser Asp Tyr Gln Leu

340 345 350

Pro Tyr Val Leu Gly Ser Ala His Glu Gly Cys Leu Pro Pro Phe Pro

355 360 365

Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asp

370 375 380

Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe

385 390 395 400

Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Gln Phe Ser Tyr Glu

405 410 415

Phe Glu Asn Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu

420 425 430

Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser

435 440 445

[0006] Lys Thr Ile Asn Gly Ser Gly Gln Asn Gln Gln Thr Leu Lys Phe Ser

450 455 460

Val Ala Gly Pro Ser Asn Met Ala Val Gln Gly Arg Asn Tyr Ile Pro

465 470 475 480

Gly Pro Ser Tyr Arg Gln Gln Arg Val Ser Thr Thr Val Thr Gln Asn

485 490 495

Asn Asn Ser Glu Phe Ala Trp Pro Gly Ala Ser Ser Trp Ala Leu Asn

500 505 510

Gly Arg Asn Ser Leu Met Asn Pro Gly Pro Ala Met Ala Ser His Lys

515 520 525

Glu Gly Glu Asp Arg Phe Phe Pro Leu Ser Gly Ser Leu Ile Phe Gly

530 535 540

Lys Gln Gly Thr Gly Arg Asp Asn Val Asp Ala Asp Lys Val Met Ile

545 550 555 560

Thr Asn Glu Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr Glu Ser

565 570 575

Tyr Gly Gln Val Ala Thr Asn His Gln Ser Ala Gln Ala Gln Ala Gln

580 585 590

Thr Gly Trp Val Gln Asn Gln Gly Ile Leu Pro Gly Met Val Trp Gln
595 600 605
Asp Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His
610 615 620
Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Met
625 630 635 640
Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala
645 650 655
Asp Pro Pro Thr Ala Phe Asn Lys Asp Lys Leu Asn Ser Phe Ile Thr
660 665 670
Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln
675 680 685
Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn
690 695 700
Tyr Tyr Lys Ser Asn Asn Val Glu Phe Ala Val Asn Thr Glu Gly Val
705 710 715 720
Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu
[0007] 725 730 735

<210> 3
<211> 736
<212> PRT
<213> 人工序列

<220>
<223> 新颖AAV分离株

<400> 3

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

50 55 60

Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp

65 70 75 80

Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala

85 90 95

Asp Ala Glu Phe Gln Glu Arg Leu Lys Glu Asp Thr Ser Phe Gly Gly

100 105 110

Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Leu Leu Glu Pro

115 120 125

Leu Gly Leu Val Glu Glu Ala Ala Lys Thr Ala Pro Gly Lys Lys Arg

130 135 140

Pro Val Glu Gln Ser Pro Gln Glu Pro Asp Ser Ser Ala Gly Ile Gly

145 150 155 160

Lys Ser Gly Ala Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr

165 170 175

Gly Asp Thr Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro

180 185 190

[0008] Ala Ala Pro Ser Gly Val Gly Ser Leu Thr Met Ala Ser Gly Gly Gly

195 200 205

Ala Pro Val Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser Ser

210 215 220

Ser Gly Asn Trp His Cys Asp Ser Gln Trp Leu Gly Asp Arg Val Ile

225 230 235 240

Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu

245 250 255

Tyr Lys Gln Ile Ser Asn Ser Thr Ser Gly Gly Ser Ser Asn Asp Asn

260 265 270

Ala Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg

275 280 285

Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn

290 295 300

Asn Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile

305 310 315 320

Gln Val Lys Glu Val Thr Asp Asn Asn Gly Val Lys Thr Ile Ala Asn

325 330 335

Asn Leu Thr Ser Thr Val Gln Val Phe Thr Asp Ser Asp Tyr Gln Leu
340 345 350
Pro Tyr Val Leu Gly Ser Ala His Glu Gly Cys Leu Pro Pro Phe Pro
355 360 365
Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asp
370 375 380
Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe
385 390 395 400
Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Gln Phe Ser Tyr Glu
405 410 415
Phe Glu Asn Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu
420 425 430
Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser
435 440 445
Lys Thr Ile Asn Gly Ser Gly Gln Asn Gln Gln Thr Leu Lys Phe Ser
450 455 460
Val Ala Gly Pro Ser Asn Met Ala Val Gln Gly Arg Asn Tyr Ile Pro
[0009] 465 470 475 480
Gly Pro Ser Tyr Arg Gln Gln Arg Val Ser Thr Thr Val Thr Gln Asn
485 490 495
Asn Asn Ser Glu Phe Ala Trp Pro Gly Ala Ser Ser Trp Ala Leu Asn
500 505 510
Gly Arg Asn Ser Leu Met Asn Pro Gly Pro Ala Met Ala Ser His Lys
515 520 525
Glu Gly Glu Asp Arg Phe Phe Pro Leu Ser Gly Ser Leu Ile Phe Gly
530 535 540
Lys Gln Gly Thr Gly Arg Asp Asn Val Asp Ala Asp Lys Val Met Ile
545 550 555 560
Thr Asn Glu Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr Glu Ser
565 570 575
Tyr Gly Gln Val Ala Thr Asn His Gln Ser Ala Gln Ala Gln Ala Gln
580 585 590
Thr Gly Trp Val Gln Asn Gln Gly Ile Leu Pro Gly Met Val Trp Gln
595 600 605
Asp Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His

610 615 620

Thr Gly Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Met

625 630 635 640

Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala

645 650 655

Asp Pro Pro Thr Ala Phe Asn Lys Asp Lys Leu Asn Ser Phe Ile Thr

660 665 670

Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln

675 680 685

Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn

690 695 700

Tyr Tyr Lys Ser Asn Asn Val Glu Phe Ala Val Asn Thr Gly Gly Val

705 710 715 720

Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu

725 730 735

<210> 4

<211> 736

[0010] <212> PRT

<213> 人工序列

<220>

<223> 新颖AAV分离株

<400> 4

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

1 5 10 15

Glu Gly Ile Arg Glu Trp Trp Ala Leu Lys Pro Gly Ala Pro Gln Pro

20 25 30

Lys Ala Asn Gln Gln His Gln Asp Asn Ala Arg Gly Leu Val Leu Pro

35 40 45

Gly Tyr Lys Tyr Leu Gly Pro Gly Asn Gly Leu Asp Lys Gly Glu Pro

50 55 60

Ile Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp

65 70 75 80

Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala
85 90 95
Asp Ala Glu Phe Gln Glu Arg Leu Lys Glu Asp Thr Ser Phe Gly Gly
100 105 110
Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Leu Leu Glu Pro
115 120 125
Leu Gly Leu Val Glu Glu Ala Ala Lys Thr Ala Pro Gly Lys Lys Arg
130 135 140
Pro Val Glu Gln Ser Pro Gln Glu Pro Asp Ser Ser Ala Gly Ile Gly
145 150 155 160
Lys Ser Gly Ala Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr
165 170 175
Gly Asp Thr Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro
180 185 190
Ala Ala Pro Ser Gly Val Gly Ser Leu Thr Met Ala Ser Gly Gly Gly
195 200 205
Ala Pro Val Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser Ser
[0011] 210 215 220
Ser Gly Asn Trp His Cys Asp Ser Gln Trp Leu Gly Asp Arg Val Ile
225 230 235 240
Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu
245 250 255
Tyr Lys Gln Ile Ser Asn Ser Thr Ser Gly Gly Ser Ser Asn Asp Asn
260 265 270
Ala Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg
275 280 285
Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn
290 295 300
Asn Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile
305 310 315 320
Gln Val Lys Glu Val Thr Asp Asn Asn Gly Val Lys Thr Ile Ala Asn
325 330 335
Asn Leu Thr Ser Thr Val Gln Val Phe Thr Asp Ser Asp Tyr Gln Leu
340 345 350
Pro Tyr Val Leu Gly Ser Ala His Glu Gly Cys Leu Pro Pro Phe Pro

355 360 365

Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asp

370 375 380

Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe

385 390 395 400

Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Gln Phe Ser Tyr Glu

405 410 415

Phe Glu Asn Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu

420 425 430

Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser

435 440 445

Lys Thr Ile Asn Gly Ser Gly Gln Asn Gln Gln Thr Leu Lys Phe Ser

450 455 460

Val Ala Gly Pro Ser Asn Met Ala Val Gln Gly Arg Asn Tyr Ile Pro

465 470 475 480

Gly Pro Ser Tyr Arg Gln Gln Arg Val Ser Thr Thr Val Thr Gln Asn

485 490 495

[0012] Asn Asn Ser Glu Phe Ala Trp Pro Gly Ala Ser Ser Trp Ala Leu Asn

500 505 510

Gly Arg Asn Ser Leu Met Asn Pro Gly Pro Ala Met Ala Ser His Lys

515 520 525

Glu Gly Glu Asp Arg Phe Phe Pro Leu Ser Gly Ser Leu Ile Phe Gly

530 535 540

Lys Gln Gly Thr Gly Arg Asp Asn Val Asp Ala Asp Lys Val Met Ile

545 550 555 560

Thr Asn Glu Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr Glu Ser

565 570 575

Tyr Gly Gln Val Ala Thr Asn His Gln Ser Ala Gln Ala Gln Ala Gln

580 585 590

Thr Gly Trp Val Gln Asn Gln Gly Ile Leu Pro Gly Met Val Trp Gln

595 600 605

Asp Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His

610 615 620

Thr Tyr Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Met

625 630 635 640

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

<210> 5

<211> 736

<212> PRT

<213> 人工序列

[0013] <220>

<223> 新颖AAV分离株

<400> 5

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

100 105 110

Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Leu Leu Glu Pro

115 120 125

Leu Gly Leu Val Glu Glu Ala Ala Lys Thr Ala Pro Gly Lys Lys Arg

130 135 140

Pro Val Glu Gln Ser Pro Gln Glu Pro Asp Ser Ser Ala Gly Ile Asp

145 150 155 160

Lys Ser Gly Ala Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr

165 170 175

Gly Asp Thr Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro

180 185 190

Ala Ala Pro Ser Gly Val Gly Ser Leu Thr Met Ala Ser Gly Gly Gly

195 200 205

Ala Pro Val Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser Ser

210 215 220

Ser Gly Asn Trp His Cys Asp Ser Gln Trp Leu Gly Asp Arg Val Ile

225 230 235 240

[0014] Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu

245 250 255

Tyr Lys Gln Ile Ser Asn Ser Thr Ser Gly Gly Ser Ser Asn Asp Asn

260 265 270

Ala Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg

275 280 285

Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn

290 295 300

Asn Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile

305 310 315 320

Gln Val Lys Glu Val Thr Asp Asn Asn Gly Val Lys Thr Ile Ala Asn

325 330 335

Asn Leu Thr Ser Thr Val Gln Val Phe Thr Asp Ser Asp Tyr Gln Leu

340 345 350

Pro Tyr Val Leu Gly Ser Ala His Glu Gly Cys Leu Pro Pro Phe Pro

355 360 365

Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asp

370 375 380

Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe
385 390 395 400
Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Gln Phe Ser Tyr Glu
405 410 415
Phe Glu Asn Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu
420 425 430
Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser
435 440 445
Lys Thr Ile Asn Gly Ser Gly Gln Asn Gln Gln Thr Leu Lys Phe Ser
450 455 460
Val Ala Gly Pro Ser Asn Met Ala Val Gln Gly Arg Asn Tyr Ile Pro
465 470 475 480
Gly Pro Ser Tyr Arg Gln Gln Arg Val Ser Thr Thr Val Thr Gln Asn
485 490 495
Asn Asn Ser Glu Phe Ala Trp Pro Gly Ala Ser Ser Trp Ala Leu Asn
500 505 510
Gly Arg Asn Ser Leu Met Asn Pro Gly Pro Ala Met Ala Ser His Lys
[0015] 515 520 525
Glu Gly Glu Asp Arg Phe Phe Pro Leu Ser Gly Ser Leu Ile Phe Gly
530 535 540
Lys Gln Gly Thr Gly Arg Asp Asn Val Asp Ala Asp Lys Val Met Ile
545 550 555 560
Thr Asn Glu Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr Glu Ser
565 570 575
Tyr Gly Gln Val Ala Thr Asn His Gln Ser Ala Gln Ala Gln Ala Gln
580 585 590
Thr Gly Trp Val Gln Asn Gln Gly Ile Leu Pro Gly Met Val Trp Gln
595 600 605
Asp Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His
610 615 620
Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Met
625 630 635 640
Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala
645 650 655
Asp Pro Pro Thr Ala Phe Asn Lys Asp Lys Leu Asn Ser Phe Ile Thr

660 665 670

Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln

675 680 685

Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn

690 695 700

Tyr Tyr Lys Ser Asn Asn Val Glu Phe Ala Val Asn Thr Glu Gly Val

705 710 715 720

Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu

725 730 735

<210> 6

<211> 736

<212> PRT

<213> 人工序列

<220>

<223> 新颖AAV分离株

[0016] <400> 6

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

1 5 10 15

Glu Gly Ile Arg Glu Trp Trp Ala Leu Lys Pro Gly Ala Pro Gln Pro

20 25 30

Lys Ala Asn Gln Gln His Gln Asp Asn Ala Arg Gly Leu Val Leu Pro

35 40 45

Gly Tyr Lys Tyr Leu Gly Pro Gly Asn Gly Leu Asp Lys Gly Glu Pro

50 55 60

Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp

65 70 75 80

Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala

85 90 95

Asp Ala Glu Phe Gln Glu Arg Leu Lys Glu Asp Thr Ser Phe Gly Gly

100 105 110

Asn Leu Gly Arg Ala Val Leu Gln Ala Lys Lys Arg Leu Leu Glu Pro

115 120 125

Leu Gly Leu Val Glu Glu Ala Ala Lys Thr Ala Pro Gly Lys Lys Arg
130 135 140
Pro Val Glu Gln Ser Pro Gln Glu Pro Asp Ser Ser Ala Gly Ile Gly
145 150 155 160
Lys Ser Gly Ala Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr
165 170 175
Gly Asp Thr Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro
180 185 190
Ala Ala Pro Ser Gly Val Gly Ser Leu Thr Met Ala Ser Gly Gly Gly
195 200 205
Ala Pro Val Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser Ser
210 215 220
Ser Gly Asn Trp His Cys Asp Ser Gln Trp Leu Gly Asp Arg Val Ile
225 230 235 240
Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu
245 250 255
Tyr Lys Gln Ile Ser Asn Ser Thr Ser Gly Gly Ser Ser Asn Asp Asn
[0017] 260 265 270
Ala Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg
275 280 285
Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn
290 295 300
Asn Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile
305 310 315 320
Gln Val Lys Glu Val Thr Asp Asn Asn Gly Val Lys Thr Ile Ala Asn
325 330 335
Asn Leu Thr Ser Thr Val Gln Val Phe Thr Asp Ser Asp Tyr Gln Leu
340 345 350
Pro Tyr Val Leu Gly Ser Ala His Glu Gly Cys Leu Pro Pro Phe Pro
355 360 365
Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asp
370 375 380
Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe
385 390 395 400
Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Gln Phe Ser Tyr Glu

405 410 415

Phe Glu Asn Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu

420 425 430

Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser

435 440 445

Lys Thr Ile Asn Gly Ser Gly Gln Asn Gln Gln Thr Leu Lys Phe Ser

450 455 460

Val Ala Gly Ser Ser Asn Met Ala Val Gln Gly Arg Asn Tyr Ile Pro

465 470 475 480

Gly Pro Ser Tyr Arg Gln Gln Arg Val Ser Thr Thr Val Thr Gln Asn

485 490 495

Asn Asn Ser Glu Phe Ala Trp Pro Gly Ala Ser Ser Trp Ala Leu Asn

500 505 510

Gly Arg Asn Ser Leu Met Asn Pro Gly Pro Ala Met Ala Ser His Lys

515 520 525

Glu Gly Glu Asp Arg Phe Phe Pro Leu Ser Gly Ser Leu Ile Phe Gly

530 535 540

[0018] Lys Gln Gly Thr Gly Arg Asp Asn Val Asp Ala Asp Lys Val Met Ile

545 550 555 560

Thr Asn Glu Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr Glu Ser

565 570 575

Tyr Gly Gln Val Ala Thr Asn His Gln Ser Ala Gln Ala Gln Ala Gln

580 585 590

Thr Gly Trp Val Gln Asn Gln Gly Ile Leu Pro Gly Met Val Trp Gln

595 600 605

Asp Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His

610 615 620

Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Met

625 630 635 640

Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala

645 650 655

Asp Pro Pro Thr Ala Phe Asn Lys Asp Lys Leu Asn Ser Phe Ile Thr

660 665 670

Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln

675 680 685

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

<210> 7

<211> 736

<212> PRT

<213> 人工序列

<220>

<223> 新颖AAV分离株

<400> 7

[0019]

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

145 150 155 160

Lys Ser Gly Ala Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr

165 170 175

Gly Asp Thr Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro

180 185 190

Ala Ala Pro Ser Gly Val Gly Ser Leu Thr Met Ala Ser Gly Gly Gly

195 200 205

Ala Pro Val Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser Ser

210 215 220

Ser Gly Asn Trp His Cys Asp Ser Gln Trp Leu Gly Asp Arg Val Ile

225 230 235 240

Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu

245 250 255

Tyr Lys Gln Ile Ser Asn Ser Thr Ser Gly Gly Ser Ser Asn Asp Asn

260 265 270

Ala Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg

275 280 285

[0020] Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn

290 295 300

Asn Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile

305 310 315 320

Gln Val Lys Glu Val Thr Asp Asn Asn Gly Val Lys Thr Ile Ala Asn

325 330 335

Asn Leu Thr Ser Thr Val Gln Val Phe Thr Asp Ser Asp Tyr Gln Leu

340 345 350

Pro Tyr Val Leu Gly Ser Ala His Glu Gly Cys Leu Pro Pro Phe Pro

355 360 365

Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asp

370 375 380

Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe

385 390 395 400

Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Gln Phe Ser Tyr Glu

405 410 415

Phe Glu Asn Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu

420 425 430

Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser
435 440 445
Lys Thr Ile Asn Gly Ser Gly Gln Asn Gln Gln Thr Leu Lys Phe Ser
450 455 460
Val Ala Gly Pro Ser Asn Met Ala Val Gln Gly Arg Asn Tyr Ile Pro
465 470 475 480
Gly Pro Ser Tyr Arg Gln Gln Arg Val Ser Thr Thr Val Thr Gln Asn
485 490 495
Asn Asn Ser Glu Phe Ala Trp Pro Gly Ala Ser Ser Trp Ala Leu Asn
500 505 510
Gly Arg Asn Ser Leu Met Asn Pro Gly Pro Ala Met Ala Ser His Lys
515 520 525
Glu Gly Glu Asp Arg Phe Phe Pro Leu Ser Gly Ser Leu Ile Phe Gly
530 535 540
Lys Gln Gly Thr Gly Arg Asp Asn Val Asp Ala Asp Lys Val Met Ile
545 550 555 560
Thr Asn Glu Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr Glu Ser
[0021] 565 570 575
Tyr Gly Gln Val Ala Thr Asn His Gln Ser Ala Gln Ala Arg Ala Gln
580 585 590
Thr Gly Trp Val Gln Asn Gln Gly Ile Leu Pro Gly Met Val Trp Gln
595 600 605
Asp Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His
610 615 620
Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Met
625 630 635 640
Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala
645 650 655
Asp Pro Pro Thr Ala Phe Asn Lys Asp Lys Leu Asn Ser Phe Ile Thr
660 665 670
Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln
675 680 685
Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn
690 695 700
Tyr Tyr Lys Ser Asn Asn Val Glu Phe Ala Val Asn Thr Glu Gly Val

705 710 715 720

Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu

725 730 735

<210> 8

<211> 736

<212> PRT

<213> 人工序列

<220>

<223> 新颖AAV分离株

<400> 8

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

1 5 10 15

Glu Gly Ile Arg Glu Trp Trp Ala Leu Lys Pro Gly Ala Pro Gln Pro

20 25 30

[0022] Lys Ala Asn Gln Gln His Gln Asp Asn Ala Arg Gly Leu Val Leu Pro

35 40 45

Gly Tyr Lys Tyr Leu Gly Pro Gly Asn Gly Leu Asp Lys Gly Glu Pro

50 55 60

Val Asn Ala Val Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp

65 70 75 80

Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala

85 90 95

Asp Ala Glu Phe Gln Glu Arg Leu Lys Glu Asp Thr Ser Phe Gly Gly

100 105 110

Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Leu Leu Glu Pro

115 120 125

Leu Gly Leu Val Glu Glu Ala Ala Lys Thr Ala Pro Gly Lys Lys Arg

130 135 140

Pro Val Glu Gln Ser Pro Gln Glu Pro Asp Ser Ser Ala Gly Ile Gly

145 150 155 160

Lys Ser Gly Ala Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr

165 170 175

Gly Asp Thr Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro
180 185 190
Ala Ala Pro Ser Gly Val Gly Ser Leu Thr Met Ala Ser Gly Gly Gly
195 200 205
Ala Pro Val Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser Ser
210 215 220
Ser Gly Asn Trp His Cys Asp Ser Gln Trp Leu Gly Asp Arg Val Ile
225 230 235 240
Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu
245 250 255
Tyr Lys Gln Ile Ser Asn Ser Thr Ser Gly Gly Ser Ser Asn Asp Asn
260 265 270
Ala Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg
275 280 285
Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn
290 295 300
Asn Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile
[0023] 305 310 315 320
Gln Val Lys Glu Val Thr Asp Asn Asn Gly Val Lys Thr Ile Ala Asn
325 330 335
Asn Leu Thr Ser Thr Val Gln Val Phe Thr Asp Ser Asp Tyr Gln Leu
340 345 350
Pro Tyr Val Leu Gly Ser Ala His Glu Gly Cys Leu Pro Pro Phe Pro
355 360 365
Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asp
370 375 380
Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe
385 390 395 400
Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Gln Phe Ser Tyr Glu
405 410 415
Phe Glu Asn Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu
420 425 430
Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser
435 440 445
Lys Thr Ile Asn Gly Ser Gly Gln Asn Gln Gln Thr Leu Lys Phe Ser

450 455 460

Val Ala Gly Pro Ser Asn Met Ala Val Gln Gly Arg Asn Tyr Ile Pro

465 470 475 480

Gly Pro Ser Tyr Arg Gln Gln Arg Val Ser Thr Thr Val Thr Gln Asn

485 490 495

Asn Asn Ser Glu Phe Ala Trp Pro Gly Ala Ser Ser Trp Ala Leu Asn

500 505 510

Gly Arg Asn Ser Leu Met Asn Pro Gly Pro Ala Met Ala Ser His Lys

515 520 525

Glu Gly Glu Asp Arg Phe Phe Pro Leu Ser Gly Ser Leu Ile Phe Gly

530 535 540

Lys Gln Gly Thr Gly Arg Asp Asn Val Asp Ala Asp Lys Val Met Ile

545 550 555 560

Thr Asn Glu Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr Glu Ser

565 570 575

Tyr Gly Gln Val Ala Thr Asn His Gln Ser Ala Gln Ala Gln Ala Gln

580 585 590

[0024] Thr Gly Trp Val Gln Asn Gln Gly Ile Leu Pro Gly Met Val Trp Gln

595 600 605

Asp Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His

610 615 620

Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Met

625 630 635 640

Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala

645 650 655

Asp Pro Pro Thr Ala Phe Asn Lys Asp Lys Leu Asn Ser Phe Ile Thr

660 665 670

Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln

675 680 685

Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn

690 695 700

Tyr Tyr Lys Ser Asn Asn Val Glu Phe Ala Val Asn Thr Glu Gly Val

705 710 715 720

Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu

725 730 735

<210> 9

<211> 736

<212> PRT

<213> 人工序列

<220>

<223> 新颖AAV分离株

<400> 9

[0025]

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

195 200 205

Ala Pro Val Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser Ser

210 215 220

Ser Gly Asn Trp His Cys Asp Ser Gln Trp Leu Gly Asp Arg Val Ile

225 230 235 240

Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu

245 250 255

Tyr Lys Gln Ile Ser Asn Ser Thr Ser Gly Gly Ser Ser Asn Asp Asn

260 265 270

Ala Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg

275 280 285

Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn

290 295 300

Asn Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile

305 310 315 320

Gln Val Lys Glu Val Thr Asp Asn Asn Gly Val Lys Thr Ile Ala Asn

325 330 335

[0026] Asn Leu Thr Ser Thr Val Gln Val Phe Thr Asp Ser Asp Tyr Gln Leu

340 345 350

Pro Tyr Val Leu Gly Ser Ala His Glu Gly Cys Leu Pro Pro Phe Pro

355 360 365

Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asp

370 375 380

Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe

385 390 395 400

Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Gln Phe Ser Tyr Glu

405 410 415

Phe Glu Asn Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu

420 425 430

Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser

435 440 445

Lys Thr Ile Asn Gly Ser Gly Gln Asn Gln Gln Thr Leu Lys Phe Ser

450 455 460

Val Ala Gly Pro Ser Asn Met Ala Val Gln Gly Arg Asn Tyr Ile Pro

465 470 475 480

Gly Pro Ser Tyr Arg Gln Gln Arg Val Ser Thr Thr Val Thr Gln Asn
485 490 495
Asn Asn Ser Glu Phe Ala Trp Pro Gly Ala Ser Ser Trp Ala Leu Asn
500 505 510
Gly Arg Asn Ser Leu Met Asn Pro Gly Pro Ala Met Ala Ser His Lys
515 520 525
Glu Gly Glu Asp Arg Phe Phe Pro Leu Ser Gly Ser Leu Ile Phe Gly
530 535 540
Lys Gln Gly Thr Gly Arg Asp Asn Val Asp Ala Asp Lys Val Met Ile
545 550 555 560
Thr Asn Glu Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr Glu Ser
565 570 575
Tyr Gly Gln Val Ala Thr Asn His Gln Ser Ala Gln Ala Gln Ala Gln
580 585 590
Thr Gly Trp Val Gln Asn Gln Gly Ile Leu Pro Gly Met Val Trp Gln
595 600 605
Asp Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His
[0027] 610 615 620
Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Met
625 630 635 640
Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala
645 650 655
Asp Pro Pro Thr Ala Phe Asn Lys Asp Lys Leu Asn Ser Phe Ile Thr
660 665 670
Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln
675 680 685
Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn
690 695 700
Tyr Tyr Lys Ser Asn Asn Val Glu Phe Ala Val Asn Thr Glu Gly Val
705 710 715 720
Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu
725 730 735

<210> 10

<211> 736

<212> PRT

<213> 人工序列

<220>

<223> 新颖AAV分离株

<400> 10

Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser
1 5 10 15
Glu Gly Ile Arg Glu Trp Trp Ala Leu Lys Pro Gly Ala Pro Gln Pro
20 25 30
Lys Ala Asn Gln Gln His Gln Asp Asn Ala Arg Gly Leu Val Leu Pro
35 40 45
Gly Tyr Lys Tyr Leu Gly Pro Gly Asn Gly Leu Asp Lys Gly Glu Pro
50 55 60
Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp
65 70 75 80
[0028] Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala
85 90 95
Asp Ala Glu Phe Gln Glu Arg Leu Lys Glu Asp Thr Ser Phe Gly Gly
100 105 110
Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Leu Leu Glu Pro
115 120 125
Leu Gly Leu Val Glu Glu Ala Ala Lys Thr Ala Pro Gly Lys Lys Arg
130 135 140
Pro Val Glu Gln Ser Pro Gln Glu Pro Asp Ser Ser Ala Gly Ile Gly
145 150 155 160
Lys Ser Gly Ala Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr
165 170 175
Gly Asp Thr Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro
180 185 190
Ala Ala Pro Ser Gly Val Gly Ser Leu Thr Met Ala Ser Cys Gly Gly
195 200 205
Ala Pro Val Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser Ser
210 215 220

Ser Gly Asn Trp His Cys Asp Ser Gln Trp Leu Gly Asp Arg Val Ile
225 230 235 240
Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu
245 250 255
Tyr Lys Gln Ile Ser Asn Ser Thr Ser Gly Gly Ser Ser Asn Asp Asn
260 265 270
Ala Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg
275 280 285
Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn
290 295 300
Asn Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile
305 310 315 320
Gln Val Lys Glu Val Thr Asp Asn Asn Gly Val Lys Thr Ile Ala Asn
325 330 335
Asn Leu Thr Ser Thr Val Gln Val Phe Thr Asp Ser Asp Tyr Gln Leu
340 345 350
Pro Tyr Val Leu Gly Ser Ala His Glu Gly Cys Leu Pro Pro Phe Pro
[0029] 355 360 365
Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asp
370 375 380
Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe
385 390 395 400
Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Gln Phe Ser Tyr Glu
405 410 415
Phe Glu Asn Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu
420 425 430
Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser
435 440 445
Lys Thr Ile Asn Gly Ser Gly Gln Asn Gln Gln Thr Leu Lys Phe Ser
450 455 460
Val Ala Gly Pro Ser Asn Met Ala Val Gln Gly Arg Asn Tyr Ile Pro
465 470 475 480
Gly Pro Ser Tyr Arg Gln Gln Arg Val Ser Thr Thr Val Thr Gln Asn
485 490 495
Asn Asn Ser Glu Phe Ala Trp Pro Gly Ala Ser Ser Trp Ala Leu Asn

500 505 510

Gly Arg Asn Ser Leu Met Asn Pro Gly Pro Ala Met Ala Ser His Lys

515 520 525

Glu Gly Glu Asp Arg Phe Phe Pro Leu Ser Gly Ser Leu Ile Phe Gly

530 535 540

Lys Gln Gly Thr Gly Arg Asp Asn Val Asp Ala Asp Lys Val Met Ile

545 550 555 560

Thr Asn Glu Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr Glu Ser

565 570 575

Tyr Gly Gln Val Ala Thr Asn His Gln Ser Ala Gln Ala Gln Ala Gln

580 585 590

Thr Gly Trp Val Gln Asn Gln Gly Ile Leu Pro Gly Met Val Trp Gln

595 600 605

Asp Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His

610 615 620

Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Met

625 630 635 640

[0030] Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala

645 650 655

Asp Pro Pro Thr Ala Phe Asn Lys Asp Lys Leu Asn Ser Phe Ile Thr

660 665 670

Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln

675 680 685

Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn

690 695 700

Tyr Tyr Lys Ser Asn Asn Val Glu Phe Ala Val Asn Thr Glu Gly Val

705 710 715 720

Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu

725 730 735

<210> 11

<211> 736

<212> PRT

<213> 人工序列

<220>

<223> 新颖AAV分离株

<400> 11

Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser
1 5 10 15
Glu Gly Ile Arg Glu Trp Trp Ala Leu Lys Pro Gly Ala Pro Gln Pro
20 25 30
Lys Ala Asn Gln Gln His Gln Asp Asn Ala Arg Gly Leu Val Leu Pro
35 40 45
Gly Tyr Lys Tyr Leu Gly Pro Gly Asn Gly Leu Asp Lys Gly Glu Pro
50 55 60
Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Arg Ala Tyr Asp
65 70 75 80
Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala
85 90 95
Asp Ala Glu Phe Gln Glu Arg Leu Lys Glu Asp Thr Ser Phe Gly Gly
[0031] 100 105 110
Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Leu Leu Glu Pro
115 120 125
Leu Gly Leu Val Glu Glu Ala Ala Lys Thr Ala Pro Gly Lys Lys Arg
130 135 140
Pro Val Glu Gln Ser Pro Gln Glu Pro Asp Ser Ser Ala Gly Ile Gly
145 150 155 160
Lys Ser Gly Ala Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr
165 170 175
Gly Asp Thr Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro
180 185 190
Ala Ala Pro Ser Gly Val Gly Ser Leu Thr Met Ala Ser Gly Gly Gly
195 200 205
Ala Pro Val Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser Ser
210 215 220
Ser Gly Asn Trp His Cys Asp Ser Gln Trp Leu Gly Asp Arg Val Ile
225 230 235 240
Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu

245 250 255

Tyr Lys Gln Ile Ser Asn Ser Thr Ser Gly Gly Ser Ser Asn Asp Asn

260 265 270

Ala Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg

275 280 285

Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn

290 295 300

Asn Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile

305 310 315 320

Gln Val Lys Glu Val Thr Asp Asn Asn Gly Val Lys Thr Ile Ala Asn

325 330 335

Asn Leu Thr Ser Thr Val Gln Val Phe Thr Asp Ser Asp Tyr Gln Leu

340 345 350

Pro Tyr Val Leu Gly Ser Ala His Glu Gly Cys Leu Pro Pro Phe Pro

355 360 365

Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asp

370 375 380

[0032] Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe

385 390 395 400

Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Gln Phe Ser Tyr Glu

405 410 415

Phe Glu Asn Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu

420 425 430

Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser

435 440 445

Lys Thr Ile Asn Gly Ser Gly Gln Asn Gln Gln Thr Leu Lys Phe Ser

450 455 460

Val Ala Gly Pro Ser Asn Met Ala Val Gln Gly Arg Asn Tyr Ile Pro

465 470 475 480

Gly Pro Ser Tyr Arg Gln Gln Arg Val Ser Thr Thr Val Thr Gln Asn

485 490 495

Asn Asn Ser Glu Phe Ala Trp Pro Gly Ala Ser Ser Trp Ala Leu Asn

500 505 510

Gly Arg Asn Ser Leu Met Asn Pro Gly Pro Ala Met Ala Ser His Lys

515 520 525

Glu Gly Glu Asp Arg Phe Phe Pro Leu Ser Gly Ser Leu Ile Phe Gly
530 535 540
Lys Gln Gly Thr Gly Arg Asp Asn Val Asp Ala Asp Lys Val Met Ile
545 550 555 560
Thr Asn Glu Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr Glu Ser
565 570 575
Tyr Gly Gln Val Ala Thr Asn His Gln Ser Ala Gln Ala Gln Ala Gln
580 585 590
Thr Gly Trp Val Gln Asn Gln Gly Ile Leu Pro Gly Met Val Trp Gln
595 600 605
Asp Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His
610 615 620
Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Met
625 630 635 640
Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala
645 650 655
Asp Pro Pro Thr Ala Phe Asn Lys Asp Lys Leu Asn Ser Phe Ile Thr
[0033] 660 665 670
Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln
675 680 685
Lys Lys Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn
690 695 700
Tyr Tyr Lys Ser Asn Asn Val Glu Phe Ala Val Asn Thr Glu Gly Val
705 710 715 720
Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu
725 730 735

<210> 12

<211> 736

<212> PRT

<213> 人工序列

<220>

<223> 新颖AAV分离株

<400> 12

[0034]

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

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

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

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

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

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

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

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

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

Pro Val Glu Gln Ser Pro Gln Glu Pro Asp Ser Ser Ala Gly Ile Gly
145 150 155 160

Lys Ser Gly Ala Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr
165 170 175

Gly Asp Thr Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro
180 185 190

Ala Ala Pro Ser Gly Val Gly Ser Leu Thr Met Ala Ser Gly Gly Gly
195 200 205

Ala Pro Val Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser Ser
210 215 220

Ser Gly Asn Trp His Cys Asp Ser Gln Trp Leu Gly Asp Arg Val Ile
225 230 235 240

Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu
245 250 255

Tyr Lys Gln Ile Ser Asn Ser Thr Ser Gly Gly Ser Ser Asn Asp Asn
260 265 270

Ala Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg
275 280 285
Phe His Cys His Phe Ser Pro His Asp Trp Gln Arg Leu Ile Asn Asn
290 295 300
Asn Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile
305 310 315 320
Gln Val Lys Glu Val Thr Asp Asn Asn Gly Val Lys Thr Ile Ala Asn
325 330 335
Asn Leu Thr Ser Thr Val Gln Val Phe Thr Asp Ser Asp Tyr Gln Leu
340 345 350
Pro Tyr Val Leu Gly Ser Ala His Glu Gly Cys Leu Pro Pro Phe Pro
355 360 365
Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asp
370 375 380
Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe
385 390 395 400
Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Gln Phe Ser Tyr Glu
[0035] 405 410 415
Phe Glu Asn Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu
420 425 430
Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser
435 440 445
Lys Thr Ile Asn Gly Ser Gly Gln Asn Gln Gln Thr Leu Lys Phe Asn
450 455 460
Val Ala Gly Pro Ser Asn Met Ala Val Gln Gly Arg Asn Tyr Ile Pro
465 470 475 480
Gly Pro Ser Tyr Arg Gln Gln Arg Val Ser Thr Thr Val Thr Gln Asn
485 490 495
Asn Asn Ser Glu Phe Ala Trp Pro Arg Ala Ser Ser Trp Ala Leu Asn
500 505 510
Gly Arg Asn Ser Leu Met Asn Pro Gly Pro Ala Met Ala Ser His Lys
515 520 525
Glu Gly Glu Asp Arg Phe Phe Pro Leu Ser Gly Ser Leu Ile Phe Gly
530 535 540
Lys Gln Gly Thr Gly Arg Asp Asn Val Asp Ala Asp Lys Val Met Ile

545 550 555 560

Thr Asn Glu Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr Glu Ser

565 570 575

Tyr Gly Gln Val Ala Thr Asn His Gln Ser Ala Gln Ala Gln Ala Gln

580 585 590

Thr Gly Trp Val Gln Asn Gln Gly Ile Leu Pro Gly Met Val Trp Gln

595 600 605

Asp Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His

610 615 620

Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Met

625 630 635 640

Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala

645 650 655

Asp Pro Pro Thr Ala Phe Asn Lys Asp Lys Leu Asn Ser Phe Ile Thr

660 665 670

Gln Tyr Ser Thr Gly Gln Val Ser Met Glu Ile Glu Trp Glu Leu Gln

675 680 685

[0036] Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn

690 695 700

Tyr Tyr Lys Ser Asn Asn Val Glu Phe Ala Val Asn Thr Glu Gly Val

705 710 715 720

Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu

725 730 735

<210> 13

<211> 736

<212> PRT

<213> 人工序列

<220>

<223> 新颖AAV分离株

<400> 13

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

1 5 10 15

Glu Gly Ile Arg Glu Trp Trp Ala Leu Lys Pro Gly Ala Pro Gln Pro
20 25 30
Lys Ala Asn Gln Gln His Gln Asp Asn Ala Arg Gly Leu Val Leu Pro
35 40 45
Gly Tyr Lys Tyr Leu Gly Pro Gly Asn Gly Leu Asp Lys Gly Glu Pro
50 55 60
Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp
65 70 75 80
Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala
85 90 95
Asp Ala Glu Phe Gln Glu Arg Leu Lys Glu Asp Thr Ser Phe Gly Gly
100 105 110
Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Leu Leu Glu Pro
115 120 125
Leu Gly Leu Val Glu Glu Ala Ala Lys Thr Ala Pro Gly Lys Lys Arg
130 135 140
Pro Val Glu Gln Ser Pro Gln Glu Pro Asp Ser Ser Ala Gly Ile Gly
[0037] 145 150 155 160
Lys Ser Gly Ala Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr
165 170 175
Gly Asp Thr Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro
180 185 190
Ala Ala Pro Ser Gly Val Gly Ser Leu Thr Met Ala Ser Gly Gly Gly
195 200 205
Ala Pro Val Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser Ser
210 215 220
Ser Gly Asn Trp His Cys Asp Ser Gln Trp Leu Gly Asp Arg Val Ile
225 230 235 240
Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu
245 250 255
Tyr Lys Gln Ile Ser Asn Ser Thr Ser Gly Gly Ser Ser Asn Asp Asn
260 265 270
Ala Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg
275 280 285
Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn

290 295 300

Asn Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile

305 310 315 320

Gln Val Lys Glu Val Thr Asp Asn Asn Gly Val Lys Thr Ile Ala Asn

325 330 335

Asn Leu Thr Ser Thr Val Gln Val Phe Thr Asp Ser Asp Tyr Gln Leu

340 345 350

Pro Tyr Val Leu Gly Ser Ala His Glu Gly Cys Leu Pro Pro Phe Pro

355 360 365

Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asp

370 375 380

Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe

385 390 395 400

Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Gln Phe Ser Tyr Glu

405 410 415

Phe Glu Asn Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu

420 425 430

[0038] Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser

435 440 445

Lys Thr Ile Asn Gly Ser Gly Gln Asn Gln Gln Thr Leu Lys Phe Ser

450 455 460

Val Ala Gly Pro Ser Asn Met Ala Val Gln Gly Arg Asn Tyr Ile Pro

465 470 475 480

Gly Pro Ser Tyr Arg Gln Gln Arg Val Ser Thr Thr Val Thr Gln Asn

485 490 495

Asn Asn Ser Glu Phe Ala Trp Pro Arg Ala Ser Ser Trp Ala Leu Asn

500 505 510

Gly Arg Asn Ser Leu Met Asn Pro Gly Pro Ala Met Ala Ser His Lys

515 520 525

Glu Gly Glu Asp Arg Phe Phe Pro Leu Ser Gly Ser Leu Ile Phe Gly

530 535 540

Lys Gln Gly Thr Gly Arg Asp Asn Val Asp Ala Asp Lys Val Met Ile

545 550 555 560

Thr Asn Glu Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr Glu Ser

565 570 575

Tyr Gly Gln Val Ala Thr Asn His Gln Ser Ala Gln Ala Gln Ala Gln
580 585 590
Thr Gly Trp Val Gln Asn Gln Gly Ile Leu Pro Gly Met Val Trp Gln
595 600 605
Asp Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His
610 615 620
Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Met
625 630 635 640
Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala
645 650 655
Asp Pro Pro Thr Ala Phe Asn Lys Asp Lys Leu Asn Ser Phe Ile Thr
660 665 670
Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln
675 680 685
Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn
690 695 700
Tyr Tyr Lys Ser Asn Asn Val Glu Phe Ala Val Asn Thr Glu Gly Val
[0039] 705 710 715 720
Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu
725 730 735

<210> 14
<211> 736
<212> PRT
<213> 人工序列

<220>
<223> 新颖AAV分离株

<400> 14

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

35 40 45

Gly Tyr Lys Tyr Leu Gly Pro Gly Asn Gly Leu Asp Lys Gly Glu Pro

50 55 60

Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp

65 70 75 80

Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala

85 90 95

Asp Ala Glu Phe Gln Glu Arg Leu Lys Glu Asp Thr Ser Phe Gly Gly

100 105 110

Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Leu Leu Glu Pro

115 120 125

Leu Gly Leu Val Glu Glu Ala Ala Lys Thr Ala Pro Gly Lys Lys Arg

130 135 140

Pro Val Glu Gln Ser Pro Gln Glu Pro Asp Ser Ser Ala Gly Ile Gly

145 150 155 160

Lys Ser Gly Ala Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr

165 170 175

[0040] Gly Asp Thr Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro

180 185 190

Ala Ala Pro Ser Gly Val Gly Ser Leu Thr Met Ala Ser Gly Gly Gly

195 200 205

Ala Pro Val Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser Ser

210 215 220

Ser Gly Asn Trp His Cys Asp Ser Gln Trp Leu Gly Asp Arg Val Ile

225 230 235 240

Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu

245 250 255

Tyr Lys Gln Ile Ser Asn Ser Thr Ser Gly Gly Ser Ser Asn Asp Asn

260 265 270

Ala Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg

275 280 285

Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn

290 295 300

Asn Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile

305 310 315 320

Gln Val Lys Glu Val Thr Asp Asn Asn Gly Val Lys Thr Ile Ala Asn
325 330 335
Asn Leu Thr Ser Thr Val Gln Val Phe Thr Asp Ser Asp Tyr Gln Leu
340 345 350
Pro Tyr Val Leu Gly Ser Ala His Glu Gly Cys Leu Pro Pro Phe Pro
355 360 365
Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asp
370 375 380
Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe
385 390 395 400
Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Gln Phe Ser Tyr Glu
405 410 415
Phe Glu Asn Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu
420 425 430
Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser
435 440 445
Lys Thr Ile Asn Gly Ser Gly Gln Asn Gln Gln Thr Leu Lys Phe Ser
[0041] 450 455 460
Val Ala Gly Pro Ser Asn Met Ala Val Gln Gly Arg Asn Tyr Ile Pro
465 470 475 480
Gly Pro Ser Tyr Arg Gln Gln Arg Val Ser Thr Thr Val Thr Gln Asn
485 490 495
Asn Asn Ser Glu Phe Ala Trp Pro Arg Ala Ser Ser Trp Ala Leu Asn
500 505 510
Gly Arg Asn Ser Leu Met Asn Pro Gly Pro Ala Met Ala Ser His Lys
515 520 525
Glu Gly Glu Asp Arg Phe Phe Pro Leu Ser Gly Ser Leu Ile Phe Gly
530 535 540
Lys Gln Gly Thr Gly Arg Asp Asn Val Asp Ala Asp Lys Val Met Ile
545 550 555 560
Thr Asn Glu Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr Glu Ser
565 570 575
Tyr Gly Gln Val Ala Thr Asn His Gln Ser Ala Gln Ala Gln Ala Gln
580 585 590
Thr Gly Trp Val Gln Asn Gln Gly Ile Leu Pro Gly Met Val Trp Gln

595 600 605

Asp Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His

610 615 620

Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Met

625 630 635 640

Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala

645 650 655

Asp Pro Pro Thr Ala Phe Asn Lys Asp Lys Leu Asn Ser Phe Ile Thr

660 665 670

Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln

675 680 685

Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn

690 695 700

Tyr Tyr Lys Ser Asn Asn Val Glu Phe Ala Val Asn Thr Glu Gly Val

705 710 715 720

Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu

725 730 735

[0042] <210> 15

<211> 736

<212> PRT

<213> 人工序列

<220>

<223> 新颖AAV分离株

<400> 15

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

1 5 10 15

Glu Gly Ile Arg Glu Trp Trp Ala Leu Lys Pro Gly Ala Pro Gln Pro

20 25 30

Lys Ala Asn Gln Gln His Gln Asp Asn Ala Arg Gly Leu Val Leu Pro

35 40 45

Gly Tyr Lys Tyr Leu Gly Pro Gly Asn Gly Leu Asp Lys Gly Glu Pro

50 55 60

Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp
65 70 75 80
Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala
85 90 95
Asp Ala Glu Phe Gln Glu Arg Leu Lys Glu Asp Thr Ser Phe Gly Gly
100 105 110
Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Leu Leu Glu Pro
115 120 125
Leu Gly Leu Val Glu Glu Ala Ala Lys Thr Ala Pro Gly Lys Lys Arg
130 135 140
Pro Val Glu Gln Ser Pro Gln Glu Pro Asp Ser Ser Ala Gly Ile Gly
145 150 155 160
Lys Ser Gly Ala Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr
165 170 175
Gly Asp Thr Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro
180 185 190
Ala Ala Pro Ser Gly Val Gly Ser Leu Thr Met Ala Ser Gly Gly Gly
[0043] 195 200 205
Ala Pro Val Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser Ser
210 215 220
Ser Gly Asn Trp His Cys Asp Ser Gln Trp Leu Gly Asp Arg Val Ile
225 230 235 240
Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu
245 250 255
Tyr Lys Gln Ile Ser Asn Ser Thr Ser Gly Gly Ser Ser Asn Asp Asn
260 265 270
Ala Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg
275 280 285
Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn
290 295 300
Asn Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile
305 310 315 320
Gln Val Lys Glu Val Thr Asp Asn Asn Gly Val Lys Thr Ile Ala Asn
325 330 335
Asn Leu Thr Ser Thr Val Gln Val Phe Thr Asp Ser Asp Tyr Gln Leu

340 345 350

Pro Tyr Val Leu Gly Ser Ala His Glu Gly Cys Leu Pro Pro Phe Pro

355 360 365

Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asp

370 375 380

Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe

385 390 395 400

Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Gln Phe Ser Tyr Glu

405 410 415

Phe Glu Asn Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu

420 425 430

Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser

435 440 445

Lys Thr Ile Asn Gly Ser Gly Gln Asn Gln Gln Thr Leu Lys Phe Ser

450 455 460

Val Ala Gly Pro Ser Asn Met Ala Val Gln Gly Arg Asn Tyr Ile Pro

465 470 475 480

[0044] Gly Pro Ser Tyr Arg Gln Gln Arg Val Ser Thr Thr Val Thr Gln Asn

485 490 495

Asn Asn Ser Glu Phe Ala Trp Pro Arg Ala Ser Ser Trp Ala Leu Asn

500 505 510

Gly Arg Asn Ser Leu Met Asn Pro Gly Pro Ala Met Ala Ser His Lys

515 520 525

Glu Gly Glu Asp Arg Phe Phe Pro Leu Ser Gly Ser Leu Ile Phe Gly

530 535 540

Lys Gln Gly Thr Gly Arg Asp Asn Val Asp Ala Asp Lys Val Met Ile

545 550 555 560

Thr Asn Glu Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr Glu Ser

565 570 575

Tyr Gly Gln Val Ala Thr Asn His Gln Ser Ala Gln Ala Gln Ala Gln

580 585 590

Thr Gly Trp Val Gln Asn Gln Gly Ile Leu Pro Gly Met Val Trp Gln

595 600 605

Asp Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His

610 615 620

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

<210> 16

<211> 736

<212> PRT

[0045] <213> 人工序列

<220>

<223> 新颖AAV分离株

<400> 16

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

85 90 95

Asp Ala Glu Phe Gln Glu Arg Leu Lys Glu Asp Thr Ser Phe Gly Gly

100 105 110

Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Leu Leu Glu Pro

115 120 125

Leu Gly Leu Val Glu Glu Ala Ala Lys Thr Ala Pro Gly Lys Lys Arg

130 135 140

Pro Val Glu Gln Ser Pro Gln Glu Pro Asp Ser Ser Ala Gly Ile Gly

145 150 155 160

Lys Ser Gly Ala Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr

165 170 175

Gly Asp Thr Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro

180 185 190

Ala Ala Pro Ser Gly Val Gly Ser Leu Thr Met Ala Ser Gly Gly Gly

195 200 205

Ala Pro Val Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser Ser

210 215 220

[0046] Ser Gly Asn Trp His Cys Asp Ser Gln Trp Leu Gly Asp Arg Val Ile

225 230 235 240

Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu

245 250 255

Tyr Lys Gln Ile Ser Asn Ser Thr Ser Gly Gly Ser Ser Asn Asp Asn

260 265 270

Ala Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg

275 280 285

Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn

290 295 300

Asn Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile

305 310 315 320

Gln Val Lys Glu Val Thr Asp Asn Asn Gly Val Lys Thr Ile Ala Asn

325 330 335

Asn Leu Thr Ser Thr Val Gln Val Phe Ala Asp Ser Asp Tyr Gln Leu

340 345 350

Pro Tyr Val Leu Gly Ser Ala His Glu Gly Cys Leu Pro Pro Phe Pro

355 360 365

Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asp
370 375 380
Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe
385 390 395 400
Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Gln Phe Ser Tyr Glu
405 410 415
Phe Glu Asn Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu
420 425 430
Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser
435 440 445
Lys Thr Ile Asn Gly Ser Gly Gln Asn Gln Gln Thr Leu Lys Phe Ser
450 455 460
Val Ala Gly Pro Ser Asn Met Ala Val Gln Gly Arg Asn Tyr Ile Pro
465 470 475 480
Gly Pro Ser Tyr Arg Gln Gln Arg Val Ser Thr Thr Val Thr Gln Asn
485 490 495
Asn Asn Ser Glu Phe Ala Trp Pro Arg Ala Ser Ser Trp Ala Leu Asn
[0047] 500 505 510
Gly Arg Asn Ser Leu Met Asn Pro Gly Pro Ala Met Ala Ser His Lys
515 520 525
Glu Gly Glu Asp Arg Phe Phe Pro Leu Ser Gly Ser Leu Ile Phe Gly
530 535 540
Lys Gln Gly Thr Gly Arg Asp Asn Val Asp Ala Asp Lys Val Met Ile
545 550 555 560
Thr Asn Glu Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr Glu Ser
565 570 575
Tyr Gly Gln Val Ala Thr Asn His Gln Ser Ala Gln Ala Gln Ala Gln
580 585 590
Thr Gly Trp Val Gln Asn Gln Gly Ile Leu Pro Gly Met Val Trp Gln
595 600 605
Asp Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His
610 615 620
Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Met
625 630 635 640
Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala

645 650 655

Asp Pro Pro Thr Ala Phe Asn Lys Asp Lys Leu Asn Ser Phe Ile Thr

660 665 670

Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln

675 680 685

Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn

690 695 700

Tyr Tyr Lys Ser Asn Asn Val Glu Phe Ala Val Asn Thr Glu Gly Val

705 710 715 720

Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu

725 730 735

<210> 17

<211> 736

<212> PRT

<213> 人工序列

<220>

[0048] <223> 新颖AAV分离株

<400> 17

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

1 5 10 15

Glu Gly Ile Arg Glu Trp Trp Ala Leu Lys Pro Gly Ala Pro Gln Pro

20 25 30

Lys Ala Asn Gln Gln His Gln Asp Asn Ala Arg Gly Leu Val Leu Pro

35 40 45

Gly Tyr Lys Tyr Leu Gly Pro Gly Asn Gly Leu Asp Lys Gly Glu Pro

50 55 60

Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp

65 70 75 80

Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala

85 90 95

Asp Ala Glu Phe Gln Glu Arg Leu Lys Glu Asp Thr Ser Phe Gly Gly

100 105 110

Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Leu Leu Glu Pro
115 120 125
Leu Gly Leu Val Glu Glu Ala Ala Lys Thr Ala Pro Gly Lys Lys Arg
130 135 140
Pro Val Glu Gln Ser Pro Gln Glu Pro Asp Ser Ser Ala Gly Ile Gly
145 150 155 160
Lys Ser Gly Ala Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr
165 170 175
Gly Asp Thr Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro
180 185 190
Ala Ala Pro Ser Gly Val Gly Ser Leu Thr Met Ala Ser Gly Gly Gly
195 200 205
Ala Pro Val Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser Ser
210 215 220
Ser Gly Asn Trp His Cys Asp Ser Gln Trp Leu Gly Asp Arg Val Ile
225 230 235 240
Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu
[0049] 245 250 255
Tyr Lys Gln Ile Ser Asn Ser Thr Ser Gly Gly Ser Ser Asn Asp Asn
260 265 270
Ala Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg
275 280 285
Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn
290 295 300
Asn Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile
305 310 315 320
Gln Val Lys Glu Val Thr Asp Asn Asn Gly Val Lys Thr Ile Ala Asn
325 330 335
Asn Leu Thr Ser Thr Val Gln Val Phe Thr Asp Ser Asp Tyr Gln Leu
340 345 350
Pro Tyr Val Leu Gly Ser Ala His Glu Gly Cys Leu Pro Pro Phe Pro
355 360 365
Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asp
370 375 380
Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe

385 390 395 400

Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Gln Phe Ser Tyr Glu

405 410 415

Phe Glu Asn Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu

420 425 430

Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser

435 440 445

Lys Thr Ile Asn Gly Ser Gly Gln Asn Gln Gln Thr Leu Lys Phe Ser

450 455 460

Val Ala Gly Pro Ser Asn Met Ala Val Gln Gly Arg Asn Tyr Ile Pro

465 470 475 480

Gly Pro Ser Tyr Arg Gln Gln Arg Val Ser Thr Thr Val Thr Gln Asn

485 490 495

Asn Asn Ser Glu Ile Ala Trp Pro Arg Ala Ser Ser Trp Ala Leu Asn

500 505 510

Gly Arg Asn Ser Leu Met Asn Pro Gly Pro Ala Met Ala Ser His Lys

515 520 525

[0050] Glu Gly Glu Asp Arg Phe Phe Pro Leu Ser Gly Ser Leu Ile Phe Gly

530 535 540

Lys Gln Gly Thr Gly Arg Asp Asn Val Asp Ala Asp Lys Val Met Ile

545 550 555 560

Thr Asn Glu Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr Glu Ser

565 570 575

Tyr Gly Gln Val Ala Thr Asn His Gln Ser Ala Gln Ala Gln Ala Gln

580 585 590

Thr Gly Trp Val Gln Asn Gln Gly Ile Leu Pro Gly Met Val Trp Gln

595 600 605

Asp Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His

610 615 620

Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Met

625 630 635 640

Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala

645 650 655

Asp Pro Pro Thr Ala Phe Asn Lys Asp Lys Leu Asn Ser Phe Ile Thr

660 665 670

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

<210> 18

<211> 2211

<212> DNA

<213> 人工序列

<220>

<223> 新颖AAV分离株

<400> 18

[0051] atggctgccg atggttatct tccagattgg ctcgaggaca accttagtga aggaattcgc 60
gagtggtggg ctttgaaacc tggagcccct caacccaagg caaatcaaca acatcaagac 120
aacgctcgag gtcttgtgct tccgggttac aaataccttg gacccggcaa cggactcgac 180
aagggggagc cggtcaacgc agcagacgcg gcggccctcg agcacgacaa ggcctacgac 240
cagcagctca aggccggaga caacccgtac ctcaagtaca accacgccga cgccgagttc 300
caggagcggc tcaaagaaga tacgtctttt gggggcaacc tcgggcgagc agtcttccag 360
gccaaaaaga ggcttcttga acctcttggt ctggttgagg aagcggctaa gacggctcct 420
ggaaagaaga ggcctgtaga gcagtctcct caggaaccgg actcctccgc gggtattggc 480
aaatcgggtg cacagcccgc taaaaagaga ctcaatttcg gtcagactgg cgacacagag 540
tcagtcccag accctcaacc aatcggagaa cctcccgcag ccccctcagg tgtgggatct 600
cttacaatgg cttcaggtgg tggcgcacca gtggcagaca ataacgaagg tgccgatgga 660
gtgggtagtt cctcgggaaa ttggcattgc gattcccaat ggctggggga cagagtcatc 720
accaccagca cccgaacctg ggccctgccc acctacaaca atcacctcta caagcaaatc 780
tccaacagca catctggagg atcttcaaat gacaacgcct acttcggcta cagcaccccc 840
tgggggtatt ttgacttcaa cagattccac tgccacttct caccacgtga ctggcagcga 900
ctcatcaaca acaactgggg attccggcct aagcgactca acttcaagct cttcaacatt 960
caggtcaaag aggttacgga caacaatgga gtcaagacca tcgccaataa ccttaccagc 1020
acggtccagg tcttcacgga ctcagactat cagctcccgt acgtgctcgg gtcggctcac 1080

gagggctgcc tcccgccgtt cccagcggac gttttcatga ttcctcagta cgggtatctg 1140
acgcttaatg atggaagcca ggccgtgggt cgttcgtcct tttactgcct ggaatatttc 1200
ccgtcgcaaa tgctaagaac gggtaacaac ttccagttca gctacgagtt tgagaacgta 1260
cctttccata gcagctacgc tcacagccaa agcctggacc gactaatgaa tccactcatc 1320
gaccaatact tgtactatct ctcaaagact attaacggtt ctggacagaa tcaacaaacg 1380
ctaaaattca gtgtggccgg acccagcaac atggctgtcc agggaagaaa ctacatacct 1440
ggacccagct accgacaaca acgtgtctca accactgtga ctcaaaacaa caacagcgaa 1500
tttgcttggc ctggagcttc ttcttgggct ctcaatggac gtaatagctt gatgaatcct 1560
ggacctgcta tggccagcca caaagaagga gaggaccgtt tctttccttt gtctggatct 1620
ttaatttttg gcaaacaagg aactggaaga gacaacgtgg atgcggacaa agtcatgata 1680
accaacgaag aagaaattaa aactactaac ccggtagcaa cggagtccta tggacaagtg 1740
gccacaaacc accagagtgc ccaagcacag gcgcagaccg gctgggttca aaaccaagga 1800
atacttccgg gtatggtttg gcaggacaga gatgtgtacc tgcaaggacc catttgggcc 1860
aaaattcctc acacggacgg caactttcac ccttctccgc tgatgggagg gtttggaatg 1920
aagcacccgc ctcctcagat cctcatcaaa aacacacctg tacctgcgga tcctccaacg 1980
gccttcaaca aggacaagct gaactctttc atcacccagt attctactgg ccaagtcagc 2040
gtggagatcg agtgggagct gcagaaggaa aacagcaagc gctggaaccc ggagatccag 2100
[0052] tacacttcca actattacaa gtctaataat gttgaatttg ctgttaatac tgaaggtgta 2160
tatagtgaac cccgccccat tggcaccaga tacctgactc gtaatctgta a 2211

<210> 19
<211> 4679
<212> DNA
<213> 人工序列

<220>
<223> 新颖 AAV 分离株

<400> 19
ttggccactc cctctctgcg cgctcgctcg ctcactgagg ccgggcgacc aaaggtcgcc 60
cgacgcccgg gctttgcccg ggcggcctca gtgagcgagc gagcgcgcag agagggagtg 120
gccaactcca tcactagggg ttcctggagg ggtggagtcg tgacgtgaat tacgtcatag 180
ggttagggag gtcctgtatt agaggtcacg tgagtgtttt gcgacatttt gcgacaccat 240
gtggtcacgc tgggtattta agcccgagtg agcacgcagg gtctccattt tgaagcggga 300
ggtttgaacg cgcagccgcc atgccggggt tttacgagat tgtgattaag gtccccagcg 360

accttgacga gcatctgccc ggcatttctg acagctttgt gaactgggtg gccgagaagg 420
aatgggagtt gccgccagat tctgacatgg atctgaatct gattgagcag gcacccctga 480
ccgtggccga gaagctgcag cgcgactttc tgacggaatg gcgccgtgtg agtaaggccc 540
cggaggccct tttctttgtg caatttgaga agggagagag ctacttccac atgcacgtgc 600
tcgtggaaac caccggggtg aaatccatgg ttttgggacg tttcctgagt cagattcgcg 660
aaaaactgat tcagagaatt taccgcggga tcgagccgac tttgccaaac tggttcgcgg 720
tcacaaagac cagaaatggc gccggaggcg ggaacaaggt ggtggatgag tgctacatcc 780
ccaattactt gctccccaaa acccagcctg agctccagtg ggcgtggact aatatggaac 840
agtatttaag cgcctgtttg aatctcacgg agcgtaaacg gttggtggcg cagcatctga 900
cgcacgtgtc gcagacgcag gagcagaaca aagagaatca gaatcccaat tctgatgcgc 960
cggtgatcag atcaaaaact tcagccaggt acatggagct ggtcgggtgg ctcgtggaca 1020
aggggattac ctcggagaag cagtggatcc aggaggacca ggcctcatac atctccttca 1080
atgcggcctc caactcgcgg tcccaaatca aggctgcctt ggacaatgcg ggaaagatta 1140
tgagcctgac taaaaccgcc cccgactacc tggtgggcca gcagcccgtg gaggacattt 1200
ccagcaatcg gatttataaa attttggaac taaacgggta cgatccccaa tatgcggctt 1260
ccgtctttct gggatgggcc acgaaaaagt tcggcaagag gaacaccatc tggctgtttg 1320
ggcctgcaac taccgggaag accaacatcg cggaggccat agcccacact gtgcccttct 1380
[0053] acgggtgcgt aaactggacc aatgagaact ttcccttcaa cgactgtgtc gacaagatgg 1440
tgatctggtg ggaggagggg aagatgaccg ccaaggtcgt ggagtcggcc aaagccattc 1500
tcggaggaag caaggtgcgc gtggaccaga aatgcaagtc ctcggcccag atagacccga 1560
ctcccgtgat cgtcacctcc aacaccaaca tgtgcgccgt gattgacggg aactcaacga 1620
ccttcgaaca ccagcagccg ttgcaagacc ggatgttcaa atttgaactc acccgccgtc 1680
tggatcatga ctttgggaag gtcaccaagc aggaagtcaa agactttttc cggtgggcaa 1740
aggatcacgt ggttgaggtg gagcatgaat tctacgtcaa aaagggtgga gccaagaaaa 1800
gacccgcccc cagtgacgca gatataagtg agcccaaacg ggtgcgcgag tcagttgcgc 1860
agccatcgac gtcagacgcg gaagcttcga tcaactacgc agacaggtac caaaacaaat 1920
gttctcgtca cgtgggcatg aatctgatgc tgtttccctg cagacaatgc gagagaatga 1980
atcagaattc aaatatctgc ttcactcacg gacagaaaga ctgtttagag tgctttcccg 2040
tgtcagaatc tcaacccgtt tctgtcgtca aaaaggcgta tcagaaactg tgctacattc 2100
atcatatcat gggaaaggtg ccagacgctt gcactgcctg cgatctggtc aatgtggatt 2160
tggatgactg catctttgaa caataaatga tttaaatcag gtatggctgc cgatggttat 2220
cttccagatt ggctcgagga cactctctct gaaggaataa gacagtggtg gaagctcaaa 2280
cctggcccac caccaccaaa gcccgcagag cggcataagg acgacagcag gggtcttgtg 2340
cttcctgggt acaagtacct cggacccttc aacggactcg acaagggaga gccggtcaac 2400
gaggcagacg ccgcggccct cgagcacgac aaagcctacg accggcagct cgacagcgga 2460

gacaacccgt acctcaagta caaccacgcc gacgcggagt ttcaggagcg ccttaaagaa 2520
gatacgtctt ttgggggcaa cctcggacga gcagtcttcc aggcgaaaaa gagggttctt 2580
gaacctctgg gcctggttga ggaacctgtt aagacggctc cgggaaaaaa gaggccggta 2640
gagcactctc ctgtggagcc agactcctcc tcgggaaccg gaaaggcggg ccagcagcct 2700
gcaagaaaaa gattgaattt tggtcagact ggagacgcag actcagtacc tgacccccag 2760
cctctcggac agccaccagc agccccctct ggtctgggaa ctaatacgat ggctacaggc 2820
agtggcgcac caatggcaga caataacgag ggcgccgacg gagtgggtaa ttcctcggga 2880
aattggcatt gcgattccac atggatgggc gacagagtca tcaccaccag cacccgaacc 2940
tgggccctgc ccacctacaa caaccacctc tacaaacaaa tttccagcca atcaggagcc 3000
tcgaacgaca atcactactt tggctacagc accccttggg ggtattttga cttcaacaga 3060
ttccactgcc acttttcacc acgtgactgg caaagactca tcaacaacaa ctggggattc 3120
cgacccaaga gactcaactt caagctcttt aacattcaag tcaaagaggt cacgcagaat 3180
gacggtacga cgacgattgc caataacctt accagcacgg ttcaggtgtt tactgactcg 3240
gagtaccagc tcccgtacgt cctcggctcg gcgcatcaag gatgcctccc gccgttccca 3300
gcagacgtct tcatggtgcc acagtatgga tacctcaccc tgaacaacgg gagtcaggca 3360
gtaggacgct cttcatttta ctgcctggag tactttcctt ctcagatgct gcgtaccgga 3420
aacaacttta ccttcagcta cacttttgag gacgttcctt tccacagcag ctacgctcac 3480
[0054] agccagagtc tggaccgtct catgaatcct ctcatcgacc agtacctgta ttacttgagc 3540
agaacaaaca ctccaagtgg aaccaccacg cagtcaaggc ttcagttttc tcaggccgga 3600
gcgagtgaca ttcgggacca gtctaggaac tggcttcctg gaccctgtta ccgccagcag 3660
cgagtatcaa agacatctgc ggataacaac aacagtgaat actcgtggac tggagctacc 3720
aagtaccacc tcaatggcag agactctctg gtgaatccgg gcccggccat ggcaagccac 3780
aaggacgatg aagaaaagtt ttttcctcag agcggggttc tcatctttgg gaagcaaggc 3840
tcagagaaaa caaatgtgga cattgaaaag gtcatgatta cagacgaaga ggaaatcagg 3900
acaaccaatc ccgtggctac ggagcagtat ggttctgtat ctaccaacct ccagagaggc 3960
aacagacaag cagctaccgc agatgtcaac acacaaggcg ttcttccagg catggtctgg 4020
caggacagag atgtgtacct tcaggggccc atctgggcaa agattccaca cacggacgga 4080
cattttcacc cctctcccct catgggtgga ttcggactta aacaccctcc tccacagatt 4140
ctcatcaaga acaccccggt acctgcgaat ccttcgacca ccttcagtgc ggcaaagttt 4200
gcttccttca tcacacagta ctccacggga caggtcagcg tggagatcga gtgggagctg 4260
cagaaggaaa acagcaaacg ctggaatccc gaaattcagt acacttccaa ctacaacaag 4320
tctgttaatg tggactttac tgtggacact aatggcgtgt attcagagcc tcgccccatt 4380
ggcaccagat acctgactcg taatctgtaa ttgcttgtta atcaataaac cgtttaattc 4440
gtttcagttg aactttggtc tctgcgtatt tctttcttat ctagtttcca tggctacgta 4500
gataagtagc atggcgggtt aatcattaac tacaaggaac ccctagtgat ggagttggcc 4560

actccctctc tgcgcgctcg ctcgctcact gaggccgggc gaccaaaggt cgcccgacgc 4620

ccgggctttg cccgggcggc ctcagtgagc gagcgagcgc gcagagaggg agtggccaa 4679

<210> 20

<211> 2601

<212> DNA

<213> 人工序列

<220>

<223> 新颖 AAV 分离株

<400> 20

ccatcgacgt cagacgcgga agcttcgatc aactacgcgg acaggtacca aaacaaatgt 60

tctcgtcacg tgggcatgaa tctgatgctg tttccctgca gacaatgcga gagactgaat 120

cagaattcaa atatctgctt cactcacggt gtcaaagact gtttagagtg ctttcccgtg 180

tcagaatctc aacccgtttc tgtcgtcaaa aaggcgtatc agaaactgtg ctacattcat 240

cacatcatgg gaaaggtgcc agacgcttgc actgcttgcg acctggtcaa tgtggacttg 300

[0055] gatgactgtg tttctgaaca ataaatgact taaaccaggt atgactgccg atggttatct 360

tccagattgg ctcgaggaca accttagtga aggaattcgc gagtggtggg ctttgaaacc 420

tggagcccct caacccaagg caaatcaaca acatcaagac aacgctcgag gtcttgtgct 480

tccgggttac aaataccttg gacccggcaa cggactcgac aagggggagc cggtcaacgc 540

agcagacgcg gcggccctcg agcacgacaa ggcctacgac cagcagctca aggccggaga 600

caacccgtac ctcaagtaca accacgccga cgccgagttc caggagcggc tcaaagaaga 660

tacgtctttt gggggcaacc tcgggcgagc agtcttccag gccaaaaaga ggcttcttga 720

acctcttggt ctggttgagg aagcggctaa gacggctcct ggaaagaaga ggcctgtaga 780

gcagtctcct caggaaccgg actcctccgc gggtattggc aaatcgggtg cacagcccgc 840

taaaaagaga ctcaatttcg gtcagactgg cgacacagag tcagtcccag accctcaacc 900

aatcggagaa cctcccgcag ccccctcagg tgtgggatct cttacaatgg cttcaggtgg 960

tggcgcacca gtggcagaca ataacgaagg tgccgatgga gtgggtagtt cctcgggaaa 1020

ttggcattgc gattcccaat ggctggggga cagagtcatc accaccagca cccgaacctg 1080

ggccctgccc acctacaaca atcacctcta caagcaaatc tccaacagca catctggagg 1140

atcttcaaat gacaacgcct acttcggcta cagcaccccc tgggggtatt ttgacttcaa 1200

cagattccac tgccacttct caccacgtga ctggcagcga ctcatcaaca acaactgggg 1260

attccggcct aagcaactca acttcaagct cttcaacatt caggtcaaag aggttacgga 1320

caacaatgga gtcaagacca tcgccaataa ccttaccagc acggtccagg tcttcacgga 1380

ctcagactat cagctcccgt acgtgctcgg gtcggctcac gagggctgcc tcccgccgtt 1440

cccagcggac gttttcatga ttcctcagta cgggtatctg acgcttaatg atggaagcca 1500

ggccgtgggt cgttcgtcct tttactgcct ggaatatttc ccgtcgcaaa tgctaagaac 1560

gggtaacaac ttccagttca gctacgagtt tgagaacgta cctttccata gcagctacgc 1620

tcacagccaa agcctggacc gactaatgaa tccactcatc gaccaatact tgtactatct 1680

ctcaaagact attaacggtt ctggacagaa tcaacaaacg ctaaaattca gtgtggccgg 1740

acccagcaac atggctgtcc agggaagaaa ctacatacct ggacccagct accgacaaca 1800

acgtgtctca accactgtga ctcaaaacaa caacagcgaa tttgcttggc ctggagcttc 1860

ttcttgggct ctcaatggac gtaatagctt gatgaatcct ggacctgcta tggccagcca 1920

caaagaagga gaggaccgtt tctttccttt gtctggatct ttaatttttg gcaaacaagg 1980

aactggaaga gacaacgtgg atgcggacaa agtcatgata accaacgaag aagaaattaa 2040

aactactaac ccggtagcaa cggagtccta tggacaagtg gccacaaacc accagagtgc 2100

ccaagcacag gcgcagaccg gctgggttca aaaccaagga atacttccgg gtatggtttg 2160

gcaggacaga gatgtgtacc tgcaaggacc catttgggcc aaaattcctc acacggacgg 2220

caactttcac ccttctccgc tgatgggagg gtttggaatg aagcacccgc ctcctcagat 2280

cctcatcaaa aacacacctg tacctgcgga tcctccaacg gccttcaaca aggacaagct 2340

gaactctttc atcacccagt attctactgg ccaagtcagc gtggagatcg agtgggagct 2400

[0056] gcagaaggaa aacagcaagc gctggaaccc ggagatccag tacacttcca actattacaa 2460

gtctaataat gttgaatttg ctgttaatac tgaaggtgta tatagtgaac cccgccccat 2520

tggcaccaga tacctgactc gtaatctgta attgcttgtt aatcaataaa ccgtttaatt 2580

cgtttcagtt gaactgcggc c 2601

<210> 21

<211> 2601

<212> DNA

<213> 人工序列

<220>

<223> 新颖AAV分离株

<400> 21

ccatcgacgt cagacgcgga agcttcgatc aactacgcgg acaggtacca aaacaaatgt 60

tctcgtcacg tgggcatgaa tctgatgctg tttccctgca gacaatgcga gagactgaat 120

cagaattcaa atatctgctt cactcacggt gtcaaagact gtttagagtg ctttcccgtg 180

tcagaatctc aacccgtttc tgtcgtcaaa aaggcgtatc agaaactgtg ctacattcat 240

cacatcatgg gaaaggtgcc agacgcttgc actgcttgcg acctggtcaa tgtggacttg 300

gatgactgtg tttctgaaca ataaatgact taaaccaggt atggctgccg atggttatct 360

tccagattgg ctcgaggaca accttagtga aggaattcgc gagtggtggg ctttgaaacc 420

tggagcccct caacccaagg caaatcaaca acatcaagac aacgctcgag gtcttgtgct 480

tccgggttac aaataccttg gacccggcaa cggactcgac aagggggagc cggtcaacgc 540

agcagacgcg gcggccctcg agcacgacaa ggcctacgac cagcagctca aggccggaga 600

caacccgtac ctcaagtaca accacgccga cgccgagttc caggagcggc tcaaagaaga 660

tacgtctttt gggggcaacc tcgggcgagc agtcttccag gccaaaaaga ggcttcttga 720

acctcttggt ctggttgagg aagcggctaa gacggctcct ggaaagaaga ggcctgtaga 780

gcagtctcct caggaaccgg actcctccgc gggtattggc aaatcgggtg cacagcccgc 840

taaaaagaga ctcaatttcg gtcagactgg cgacacagag tcagtcccag accctcaacc 900

aatcggagaa cctcccgcag ccccctcagg tgtgggatct cttacaatgg cttcaggtgg 960

tggcgcacca gtggcagaca ataacgaagg tgccgatgga gtgggtagtt cctcgggaaa 1020

ttggcattgc gattcccaat ggctggggga cagagtcatc accaccagca cccgaacctg 1080

ggccctgccc acctacaaca atcacctcta caagcaaatc tccaacagca catctggagg 1140

atcttcaaat gacaacgcct acttcggcta cagcaccccc tgggggtatt ttgacttcaa 1200

cagattccac tgccacttct caccacgtga ctggcagcga ctcatcaaca acaactgggg 1260

[0057] attccggcct aagcgactca acttcaagct cttcaacatt caggtcaaag aggttacgga 1320

caacaatgga gtcaagacca tcgccaataa ccttaccagc acggtccagg tcttcacgga 1380

ctcagactat cagctcccgt acgtgctcgg gtcggctcac gagggctgcc tcccgccgtt 1440

cccagcggac gttttcatga ttcctcagta cgggtatctg acgcttaatg atggaagcca 1500

ggccgtgggt cgttcgtcct tttactgcct ggaatatttc ccgtcgcaaa tgctaagaac 1560

gggtaacaac ttccagttca gctacgagtt tgagaacgta cctttccata gcagctacgc 1620

tcacagccaa agcctggacc gactaatgaa tccactcatc gaccaatact tgtactatct 1680

ctcaaagact attaacggtt ctggacagaa tcaacaaacg ctaaaattca gtgtggccgg 1740

acccagcaac atggctgtcc agggaagaaa ctacatacct ggacccagct accgacaaca 1800

acgtgtctca accactgtga ctcaaaacaa caacagcgaa tttgcttggc ctggggcttc 1860

ttcttgggct ctcaatggac gtaatagctt gatgaatcct ggacctgcta tggccagcca 1920

caaagaagga gaggaccgtt tctttccttt gtctggatct ttaatttttg gcaaacaagg 1980

aactggaaga gacaacgtgg atgcggacaa agtcatgata accaacgaag aagaaattaa 2040

aactactaac ccggtagcaa cggagtccta tggacaagtg gccacaaacc accagagtgc 2100

ccaagcacag gcgcagaccg gctgggttca aaaccaagga atacttccgg gtatggtttg 2160

gcaggacaga gatgtgtacc tgcaaggacc catttgggcc aaaattcctc acacgggcgg 2220

caactttcac ccttctccgc tgatgggagg gtttggaatg aagcacccgc ctcctcagat 2280

cctcatcaaa aacacacctg tacctgcgga tcctccaacg gccttcaaca aggacaagct 2340

gaactctttc atcacccagt attctactgg ccaagtcagc gtggagatcg agtgggagct 2400
gcagaaggaa aacagcaagc gctggaaccc ggagatccag tacacttcca actattacaa 2460
gtctaataat gttgaatttg ctgttaatac tggaggtgta tatagtgaac cccgccccat 2520
tggcaccaga tacctgactc gtaatctgta attgcttgtt aatcaataaa ccgtttaatt 2580
cgtttcagtt gaactgcggc c 2601

<210> 22
<211> 2601
<212> DNA
<213> 人工序列

<220>
<223> 新颖 AAV 分离株

<400> 22
ccatcgacgt cagacgcgga agcttcgatc aactacgcgg acaggtacca aaacaaatgt 60
tctcgtcacg tgggcatgaa tctgatgctg tttccctgca gacaatgcga gagactgaat 120
[0058] cagaattcaa atatctgctt cactcacggt gtcaaagact gtttagagtg ctttcccgtg 180
tcagaatctc aacccgtttc tgtcgtcaaa aaggcgtatc agaaactgtg ctacattcat 240
cacatcatgg gaaaggtgcc agacgcttgc actgcttgcg acctggtcaa tgtggactcg 300
gatgactgtg tttctgaaca ataaatgact taaaccaggt atggctgccg atggttatct 360
tccagattgg ctcgaggaca accttagtga aggaattcgc gagtggtggg ctttgaaacc 420
tggagcccct caacccaagg caaatcaaca acatcaagac aacgctcgag gtcttgtgct 480
tccgggttac aaataccttg gacccggcaa cggactcgac aagggggagc cggtcaacgc 540
agcagacgcg gcggccctcg agcacgacaa ggcctacgac cagcagctca aggccggaga 600
caacccgtac ctcaagtaca accacgccga cgccgagttc caggagcggc tcaaagaaga 660
tacgtctttt gggggcaacc tcgggcgagc agtcttccag gccaaaaaga ggcttcttga 720
acctcttggt ctggttgagg aagcggctaa gacggctcct ggaaagaaga ggcctgtaga 780
gcagtctcct caggaaccgg actcctccgc gggtattgac aaatcgggtg cacagcccgc 840
taaaaagaga ctcaatttcg gtcagactgg cgacacagag tcagtcccag accctcaacc 900
aatcggagaa cctcccgcag ccccctcagg tgtgggatct cttacaatgg cttcaggtgg 960
tggcgcacca gtggcagaca ataacgaagg tgccgatgga gtgggtagtt cctcgggaaa 1020
ttggcattgc gattcccaat ggctggggga cagagtcatc accaccagca cccgaacctg 1080
ggccctgccc acctacaaca atcacctcta caagcaaatc tccaacagca catctggagg 1140
atcttcaaat gacaacgcct acttcggcta cagcaccccc tgggggtatt ttgacttcaa 1200

cagattccac tgccacttct caccacgtga ctggcagcga ctcatcaaca acaactgggg 1260

attccggcct aagcgactca acttcaagct cttcaacatt caggtcaaag aggttacgga 1320

caacaatgga gtcaagacca tcgccaataa ccttaccagc acggtccagg tcttcacgga 1380

ctcggactat cagctcccgt acgtgctcgg gtcggctcac gagggctgcc tcccgccgtt 1440

cccagcggac gttttcatga ttcctcagta cgggtatctg acgcttaatg atggaagcca 1500

ggccgtgggt cgttcgtcct tttactgcct ggaatatttc ccgtcgcaaa tgctaagaac 1560

gggtaacaac ttccagttca gctacgagtt tgagaacgta cctttccata gcagctacgc 1620

tcacagccaa agcctggacc gactaatgaa tccactcatc gaccaatact tgtactatct 1680

ctcaaagact attaacggtt ctggacagaa tcaacaaacg ctaaaattca gtgtggccgg 1740

acccagcaac atggctgtcc agggaagaaa ctacatacct ggacccagct accgacaaca 1800

acgtgtctca accactgtga ctcaaaacaa caacagcgaa tttgcttggc ctggagcttc 1860

ttcttgggct ctcaatggac gtaatagctt gatgaatcct ggacctgcta tggccagcca 1920

caaagaagga gaggaccgtt tctttccttt gtctggatct ttaatttttg gcaaacaagg 1980

aactggaaga gacaacgtgg atgcggacaa agtcatgata accaacgaag aagaaattaa 2040

aactactaac ccggtagcaa cggagtccta tggacaagtg gccacaaacc accagagtgc 2100

ccaagcacag gcgcagaccg gctgggttca aaaccaagga atacttccgg gtatggtttg 2160

gcaggacaga gatgtgtacc tgcaaggacc catttgggcc aaaattcctc acacggacgg 2220

[0059] caactttcac ccttctccgc tgatgggagg gtttggaatg aagcacccgc ctcctcagat 2280

cctcatcaaa aacacacctg tacctgcgga tcctccaacg gccttcaaca aggacaagct 2340

gaactctttc atcacccagt attctactgg ccaagtcagc gtggagatcg agtgggagct 2400

gcagaaggaa aacagcaagc gctggaaccc ggagatccag tacacttcca actattacaa 2460

gtctaataat gttgaatttg ctgttaatac tgaaggtgta tatagtgaac cccgccccat 2520

tggcaccaga tacctgactc gtaatctgta attgcttgtt aatcaataaa ccgtttaatt 2580

cgtttcagtt gaactgcggc c 2601

<210> 23

<211> 2601

<212> DNA

<213> 人工序列

<220>

<223> 新颖 AAV 分离株

<400> 23

ccatcgacgt cagacgcgga agcttcgatc aactacgcgg acaggtacca aaacaaatgt 60

tctcgtcacg tgggcatgaa tctgatgctg tttccctgca gacaatgcga gagactgaat 120
cagaattcaa atatctgctt cactcacggt gtcaaagact gtttagagtg ctttcccgtg 180
tcagaatctc aacccgtttc tgtcgtcaaa aaggcgtatc agaaactgtg ctacattcat 240
cacatcatgg gaaaggtgcc agacgcttgc actgcttgcg acctggtcaa tgtggacttg 300
gatgactgtg tttctgaaca ataaatgact taaaccaggt atggctgccg atggttatct 360
tccagattgg ctcgaggaca accttagtga aggaattcgc gagtggtggg ctttgaaacc 420
tggagcccct caacccaagg caaatcaaca acatcaagac aacgctcgag gtcttgtgct 480
tccgggttac aaataccttg gacccggcaa cggactcgac aagggggagc cggtcaacgc 540
agcagacgcg gcggccctcg agcacgacaa ggcctacgac cagcagctca aggccggaga 600
caacccgtac ctcaagtaca accacgccga cgccgagttc caggagcggc tcaaagaaga 660
tacgtctttt gggggcaacc tcgggcgagc agtcctccag gccaaaaaga ggcttcttga 720
acctcttggt ctggttgagg aagcggctaa gacggctcct ggaaagaaga ggcctgtaga 780
gcagtctcct caggaaccgg actcctccgc gggtattggc aaatcgggtg cacagcccgc 840
taaaaagaga ctcaatttcg gtcagactgg cgacacagag tcagtcccag accctcaacc 900
aatcggagaa cctcccgcag ccccctcagg tgtgggatct cttacaatgg cttcaggtgg 960
tggcgcacca gtggcagaca ataacgaagg tgccgatgga gtgggtagtt cctcgggaaa 1020
ttggcattgc gattcccaat ggctggggga cagagtcatc accaccagca cccgaacctg 1080
[0060] ggccctgccc acctacaaca atcacctcta caagcaaatc tccaacagca catctggagg 1140
atcttcaaat gacaacgcct acttcggcta cagcaccccc tgggggtatt ttgacttcaa 1200
cagattccac tgccacttct caccacgtga ctggcagcga ctcatcaaca acaactgggg 1260
attccggcct aagcgactca acttcaagct cttcaacatt caggtcaaag aggttacgga 1320
caacaatgga gtcaagacca tcgccaataa ccttaccagc acggtccagg tcttcacgga 1380
ctcagactat cagctcccgt acgtgctcgg gtcggctcac gagggctgcc tcccgccgtt 1440
cccagcggac gttttcatga ttcctcagta cgggtatctg acgcttaatg atggaagcca 1500
ggccgtgggt cgttcgtcct tttactgcct ggaatatttc ccgtcgcaaa tgctaagaac 1560
gggtaacaac ttccagttca gctacgagtt tgagaacgta cctttccata gcagctacgc 1620
tcacagccaa agcctggacc gactaatgaa tccactcatc gaccaatact tgtactatct 1680
ctcaaagact attaacggtt ctggacagaa tcaacaaacg ctaaaattca gtgtggccgg 1740
atccagcaac atggctgtcc agggaagaaa ctacatacct ggacccagct accgacaaca 1800
acgtgtctca accactgtga ctcaaaacaa caacagcgaa tttgcttggc ctggagcttc 1860
ttcttgggct ctcaatggac gtaatagctt gatgaatcct ggacctgcta tggccagcca 1920
caaagaagga gaggaccgtt tctttccttt gtctggatct ttaatttttg gcaaacaagg 1980
aactggaaga gacaacgtgg atgcggacaa agtcatgata accaacgaag aagaaattaa 2040
aactactaac ccggtagcaa cggagtccta tggacaagtg gccacaaacc accagagtgc 2100
ccaagcacag gcgcagaccg gctgggttca aaaccaagga atacttccag gtatggtttg 2160

gcaggacaga gatgtgtacc tgcaaggacc catttgggcc aaaattcctc acacggacgg 2220
caactttcac ccttctccgc tgatgggagg gtttggaatg aagcacccgc ctcctcagat 2280
cctcatcaaa aacacacctg tacctgcgga tcctccaacg gccttcaaca aggacaagct 2340
gaactctttc atcacccagt attctactgg ccaagtcagc gtggagatcg agtgggagct 2400
gcagaaggaa aacagcaagc gctggaaccc ggagatccag tacacttcca actattacaa 2460
gtctaataat gttgaatttg ctgttaatac tgaaggtgta tatagtgaac cccgccccat 2520
tggcaccaga tacctgactc gtaatctgta attgcttgtt aatcaataaa ccgtttaatt 2580
cgtttcagtt gaactgcggc c 2601

<210> 24
<211> 2601
<212> DNA
<213> 人工序列

<220>
<223> 新颖 AAV 分离株

[0061] <400> 24
ccatcgacgt cagacgcgga agcttcgatc aactacgcgg acaggtacca aaacaaatgt 60
tctcgtcacg tgggcatgaa tctgatgctg tttccctgca gacaatgcga gagactgaat 120
cagaattcaa atatctgctt cactcacggt gtcaaagact gtttagagtg ctttcccgtg 180
tcagaatctc aacccgtttc tgtcgtcaaa aaggcgtatc agaaactgtg ctacattcat 240
cacatcatgg gaaaggtgcc agacgcttgc actgcttgcg acctggtcaa tgtggacttg 300
gatgactgtg tttctgaaca ataaatgact taaaccaggt atggctgccg atggttatct 360
tccagattgg ctcgaggaca accttagtga aggaattcgc gagtggtggg ctttgaaacc 420
tggagcccct caacccaagg caaatcaaca acatcaagac aacgctcgag gtcttgtgct 480
tccgggttac aaataccttg gacccggcaa cggactcgac aagggggagc cggtcaacgc 540
agcagacgcg gcggccctcg agcacgacaa ggcctacgac cagcagctca aggccggaga 600
caacccgtac ctcaagtaca accacgccga cgccgagttc caggagcggc tcaaagaaga 660
tacgtctttt gggggcaacc tcgggcgagc agtcttccag gccaaaaaga ggcttcttga 720
acctcttggt ctggttgagg aagcggctaa gacggctcct ggaaagaaga ggcctgtaga 780
gcagtctcct caggaaccgg actcctccgc gggtattggc aaatcgggtg cacagcccgc 840
taaaaagaga ctcaatttcg gtcagactgg cgacacagag tcagtcccag accctcaacc 900
aatcggagaa cctcccgcag ccccctcagg tgtgggatct cttacaatgg cttcaggtgg 960
tggcgcacca gtggcagaca ataacgaagg tgccgatgga gtgggtagtt cctcgggaaa 1020

ttggcattgc gattcccaat ggctggggga cagagtcatc accaccagca cccgaacctg 1080

ggccctgccc acctacaaca atcacctcta caagcaaatc tccaacagca catctggagg 1140

atcttcaaat gacaacgcct acttcggcta cagcaccccc tgggggtatt ttgacttcaa 1200

cagattccac tgccacttct caccacgtga ctggcagcga ctcatcaaca acaactgggg 1260

attccggcct aagcgactca acttcaagct cttcaacatt caggtcaaag aggttacgga 1320

caacaatgga gtcaagacca tcgccaataa ccttaccagc acggtccagg tcttcacgga 1380

ctcagactat cagctcccgt acgtgctcgg gtcggctcac gagggctgcc tcccgccgtt 1440

cccagcggac gttttcatga ttcctcagta cgggtatctg acgcttaatg atggaagcca 1500

ggccgtgggt cgttcgtcct tttactgcct ggaatatttc ccgtcgcaaa tgctaagaac 1560

gggtaacaac ttccagttca gctacgagtt tgagaacgta cctttccata gcagctacgc 1620

tcacagccaa agcctggacc gactaatgaa tccactcatc gaccaatact tgtactatct 1680

ctcaaagact attaacggtt ctggacagaa tcaacaaacg ctaaaattca gtgtggccgg 1740

acccagcaac atggctgtcc agggaagaaa ctacatacct ggacccagct accgacaaca 1800

acgtgtctca accactgtga ctcaaaacaa caacagcgaa tttgcttggc ctggagcttc 1860

ttcttgggct ctcaatggac gtaatagctt gatgaatcct ggacctgcta tggccagcca 1920

caaagaagga gaggaccgtt tctttccttt gtctggatct ttaatttttg gcaaacaagg 1980

aactggaaga gacaacgtgg atgcggacaa agtcatgata accaacgaag aagaaattaa 2040

[0062] aactactaac ccggtagcaa cggagtccta tggacaagtg gccacaaacc accagagtgc 2100

ccaagcacgg gcgcagaccg gctgggttca aaaccaagga atacttccgg gtatggtttg 2160

gcaggacaga gatgtgtacc tgcaaggacc catttgggcc aaaattcctc acacggacgg 2220

caactttcac ccttctccgc tgatgggagg gtttggaatg aagcacccgc ctcctcagat 2280

cctcatcaaa aacacacctg tacctgcgga tcctccaacg gccttcaaca aggacaagct 2340

gaactctttc atcacccagt attctactgg ccaagtcagc gtggagatcg agtgggagct 2400

gcagaaggaa aacagcaagc gctggaaccc ggagatccag tacacttcca actattacaa 2460

gtctaataat gttgaatttg ctgttaatac tgaaggtgta tatagtgaac cccgccccat 2520

tggcaccaga tacctgactc gtaatctgta attgcttgtt aatcaataaa ccgtttaatt 2580

cgtttcagtt gaactgcggc c 2601

<210> 25

<211> 2601

<212> DNA

<213> 人工序列

<220>

<223> 新颖 AAV 分离株

<400> 25

[0063]

ccatcgacgt cagacgcgga agcttcgatc aactacgcgg acaggtacca aaacaaatgt 60
tctcgtcacg tgggcatgaa tctgatgctg tttccctgca gacaatgcga gagactgaat 120
cagaattcaa atatctgctt cactcacggt gtcaaagact gtttagagtg ctttcccgtg 180
tcagaatctc aacccgtttc tgtcgtcaaa aaggcgtatc agaaactgtg ctacattcat 240
cacatcatgg gaaaggtgcc agacgcttgc actgcttgcg acctggtcaa tgtggacttg 300
gatgactgtg tttctgaaca ataaatgact taaaccaggt atggctgccg atggttatct 360
tccagattgg ctcgaggaca accttagtga aggaattcgc gagtggtggg ctttgaaacc 420
tggagcccct caacccaagg caaatcaaca acatcaagac aacgctcgag gtcttgtgct 480
tccgggttac aaataccttg gacccggcaa cggactcgac aagggggagc cggtcaacgc 540
agcagacgcg gcggccctcg agcacgacag ggcctacgac cagcagctca aggccggaga 600
caacccgtac ctcaagtaca accacgccga cgccgagttc caggagcggc tcaaagaaga 660
tacgtctttt gggggcaacc tcgggcgagc agtcttccag gccaaaaaga ggcttcttga 720
acctcttggt ctggttgagg aagcggctaa gacggctcct ggaaagaaga ggcctgtaga 780
gcagtctcct caggaaccgg actcctccgc gggtattggc aaatcgggtg cacagcccgc 840
taaaaagaga ctcaatttcg gtcagactgg cgacacagag tcagtcccag accctcaacc 900
aatcggagaa cctcccgcag ccccctcagg tgtgggatct cttacaatgg cttcaggtgg 960
tggcgcacca gtggcagaca ataacgaagg tgccgatgga gtgggtagtt cctcgggaaa 1020
ttggcattgc gattcccaat ggctggggga cagagtcatc accaccagca cccgaacctg 1080
ggccctgccc acctacaaca atcacctcta caagcaaatc tccaacagca catctggagg 1140
atcttcaaat gacaacgcct acttcggcta cagcaccccc tgggggtatt ttgacttcaa 1200
cagattccac tgccacttct caccacgtga ctggcagcga ctcatcaaca acaactgggg 1260
attccggcct aagcgactca acttcaagct cttcaacatt caggtcaaag aggttacgga 1320
caacaatgga gtcaagacca tcgccaataa ccttaccagc acggtccagg tcttcacgga 1380
ctcagactat cagctcccgt acgtgctcgg gtcggctcac gagggctgcc tcccgccgtt 1440
cccagcggac gttttcatga ttcctcagta cgggtatctg acgcttaatg atggaagcca 1500
ggccgtgggt cgttcgtcct tttactgcct ggaatatttc ccgtcgcaaa tgctaagaac 1560
gggtaacaac ttccagttca gctacgagtt tgagaacgta cctttccata gcagctacgc 1620
tcacagccaa agcctggacc gactaatgaa tccactcatc gaccaatact tgtactatct 1680
ctcaaagact attaacggtt ctggacagaa tcaacaaacg ctaaaattca gtgtggccgg 1740
acctagcaac atggctgtcc agggaagaaa ctacatacct ggacccagct accgacaaca 1800
acgtgtctca accactgtga ctcaaaacaa caacagcgaa tttgcttggc ctggagcttc 1860
ttcttgggct ctcaatggac gtaatagctt gatgaatcct ggacctgcta tggccagcca 1920
caaagaagga gaggaccgtt tctttccttt gtctggatct ttaatttttg gcaaacaagg 1980

aactggaaga gacaacgtgg atgcggacaa agtcatgata accaacgaag aagaaattaa 2040

aactactaac ccggtagcaa cggagtccta tggacaagtg gccacaaacc accagagtgc 2100

ccaagcacag gcgcagaccg gctgggttca aaaccaagga atacttccgg gtatggtttg 2160

gcaggacaga gatgtgtacc tgcaaggacc catttgggcc aaaattcctc acacggacgg 2220

caactttcac ccttctccgc tgatgggagg gtttggaatg aagcacccgc ctcctcagat 2280

cctcatcaaa aacacacctg tacctgcgga tcctccaacg gccttcaaca aggacaagct 2340

gaactctttc atcacccagt attctactgg ccaagtcagc gtggagatcg agtgggagct 2400

gcagaagaaa aacagcaagc gctggaaccc ggagatccag tacacttcca actattacaa 2460

gtctaataat gttgaatttg ctgttaacac tgaaggtgta tatagtgaac cccgccccat 2520

tggcaccaga tacctgactc gtaatctgta attgcttgtt aatcaataaa ccgtttaatt 2580

cgtttcagtt gaactgcggc c 2601

<210> 26

<211> 2601

<212> DNA

<213> 人工序列

[0064] <220>

<223> 新颖AAV分离株

<400> 26

ccatcgacgt cagacgcgga agcttcgatc aactacgcgg acaggtacca aaacaaatgt 60

tctcgtcacg tgggcatgaa tctgatgctg tttccctgca gacaatgcga gagactgaat 120

cagaattcaa atatctgctt cactcacggt gtcaaagact gtttagagtg ctttcccgtg 180

tcagaatctc aacccgtttc tgtcgtcaaa aaggcgtatc agaaactgtg ctacattcat 240

cacatcatgg gaaaggtgcc agacgcttgc actgcttgcg acctggtcaa tgtggacttg 300

gatgactgtg tttctgaaca ataaatgact taaaccaggt atggctgccg atggttatct 360

tccagattgg ctcgaggaca accttagtga aggaattcgc gagtggtggg ctttgaaacc 420

tggagcccct caacccaagg caaatcaaca acatcaagac aacgctcgag gtcttgtgct 480

tccgggttac aaataccttg gacccggcaa cggactcgac aagggggagc cgatcaacgc 540

agcagacgcg gcggccctcg agcacgacaa ggcctacgac cagcagctca aggccggaga 600

caacccgtac ctcaagtaca accacgccga cgccgagttc caggagcggc tcaaagaaga 660

tacgtctttt gggggcaacc tcgggcgagc agtcttccag gccaaaaaga ggcttcttga 720

acctcttggt ctggttgagg aagcggctaa gacggctcct ggaaagaaga ggcctgtaga 780

gcagtctcct caggaaccgg actcctccgc gggtattggc aaatcgggtg cacagcccgc 840

taaaaagaga ctcaatttcg gtcagactgg cgacacagag tcagtcccag accctcaacc 900

aatcggagaa cctcccgcag ccccctcagg tgtgggatct cttacaatgg cttcaggtgg 960

tggcgcacca gtggcagaca ataacgaagg tgccgatgga gtgggtagtt cctcgggaaa 1020

ttggcattgc gattcccaat ggctggggga cagagtcatc accaccagca cccgaacctg 1080

ggccctgccc acctacaaca atcacctcta caagcaaatc tccaacagca catctggagg 1140

atcttcaaat gacaacgcct acttcggcta cagcaccccc tgggggtatt ttgacttcaa 1200

cagattccac tgccacttct caccacgtga ctggcagcga ctcatcaaca acaactgggg 1260

attccggcct aagcgactca acttcaagct cttcaacatt caggtcaaag aggttacgga 1320

caacaatgga gtcaagacca tcgccaataa ccttaccagc acggtccagg tcttcacgga 1380

ctcagactat cagctcccgt acgtgctcgg gtcggctcac gagggctgcc tcccgccgtt 1440

cccagcggac gttttcatga ttcctcagta cgggtatctg acgcttaatg atggaagcca 1500

ggccgtgggt cgttcgtcct tttactgcct ggaatatttc ccgtcgcaaa tgctaagaac 1560

gggtaacaac ttccagttca gctacgagtt tgagaacgta cctttccata gcagctacgc 1620

tcacagccaa agcctggacc gactaatgaa tccactcatc gaccaatact tgtactatct 1680

ctcaaagact attaacggtt ctggacagaa tcaacaaacg ctaaaattca gtgtggccgg 1740

acccagcaac atggctgtcc agggaagaaa ctacatacct ggacccagct accgacaaca 1800

acgtgtctca accactgtga ctcaaaacaa caacagcgaa tttgcttggc ctggagcttc 1860

[0065] ttcttgggct ctcaatggac gtaatagctt gatgaatcct ggacctgcta tggccagcca 1920

caaagaagga gaggaccgtt tctttccttt gtctggatct ttaatttttg gcaaacaagg 1980

aactggaaga gacaacgtgg atgcggacaa agtcatgata accaacgaag aagaaattaa 2040

aactactaac ccggtagcaa cggagtccta tggacaagtg gccacaaacc accagagtgc 2100

ccaagcacag gcgcagaccg gctgggttca aaaccaagga atacttccgg gtatggtttg 2160

gcaggacaga gatgtgtacc tgcaaggacc catttgggcc aaaattcctc acacgtacgg 2220

caactttcac ccttctccgc tgatgggagg gtttggaatg aagcacccgc ctcctcagat 2280

cctcatcaaa aacacacctg tacctgcgga tcctccaacg gccttcaaca aggacaagct 2340

gaactctttc atcacccagt attctactgg ccaagtcagc gtggagatcg agtgggagct 2400

gcagaaggaa aacagcaagc gctggaaccc ggagatccag tacacttcca actattacaa 2460

gtctaataat gttgaatttg ctgttaatac tgaaggtgta tatagtgaac cccgccccat 2520

tggcaccaga tacctgactc gtaatctgta attgcttgtt aatcaataaa ccgtttaatt 2580

cgtttcagtt gaactgcggc c 2601

〈210〉 27

〈211〉 2601

〈212〉 DNA

〈213〉 人工序列

<220>

<223> 新颖 AAV 分离株

<400> 27

[0066]

ccatcgacgt cagacgcgga agcttcgatc aactacgcgg acaggtacca aaacaaatgt 60

tctcgtcacg tgggcatgaa tctgatgctg tttccctgca gacaatgcga gagactgaat 120

cagaattcaa atatctgctt cactcacggt gtcaaagact gtttagagtg ctttcccgtg 180

tcagaatctc aacccgtttc tgtcgtcaaa aaggcgtatc agaaactgtg ctacattcat 240

cacatcatgg gaaaggtgcc agacgcttgc actgcttgcg acctggtcaa tgtggacttg 300

gatgactgtg tttctgaaca ataaatgact taaaccaggt atggctgccg atggttatct 360

tccagattgg ctcgaggaca accttagtga aggaattcgc gagtggtggg ctttgaaacc 420

tggagcccct caacccaagg caaatcaaca acatcaagac aacgctcgag gtcttgtgct 480

tccgggttac aaataccttg gacccggcaa cggactcgac aagggggagc cggtcaacgc 540

agtagacgcg gcggccctcg agcacgacaa ggcctacgac cagcagctca aggccggaga 600

caacccgtac ctcaagtaca accacgccga cgccgagttc caggagcggc tcaaagaaga 660

tacgtctttt gggggcaacc tcgggcgagc agtcttccag gccaaaaaga ggcttcttga 720

acctcttggt ctggttgagg aagcggctaa gacggctcct ggaaagaaga ggcctgtaga 780

gcagtctcct caggaaccgg actcctccgc gggtattggc aaatcgggtg cacagcccgc 840

taaaaagaga ctcaatttcg gtcagactgg cgacacagag tcagtcccag accctcaacc 900

aatcggagaa cctcccgcag ccccctcagg tgtgggatct cttacaatgg cttcaggtgg 960

tggcgcacca gtggcagaca ataacgaagg tgccgatgga gtgggtagtt cctcgggaaa 1020

ttggcattgc gattcccaat ggctggggga cagagtcatc accaccagca cccgaacctg 1080

ggccctgccc acctacaaca atcacctcta caagcaaatc tccaacagca catctggagg 1140

atcttcaaat gacaacgcct acttcggcta cagcaccccc tgggggtatt ttgacttcaa 1200

cagattccac tgccacttct caccacgtga ctggcagcga ctcatcaaca acaactgggg 1260

atttcggcct aagcgactca acttcaagct cttcaacatt caggtcaaag aggttacgga 1320

caacaatgga gtcaagacca tcgccaataa ccttaccagc acggtccagg tcttcacgga 1380

ctcagactat cagctcccgt acgtgctcgg gtcggctcac gagggctgcc tcccgccgtt 1440

cccagcggac gttttcatga ttcctcagta cgggtatctg acgcttaatg atggaagcca 1500

ggccgtgggt cgttcgtcct tttactgcct ggaatatttc ccgtcgcaaa tgctaagaac 1560

gggtaacaac ttccagttca gctacgagtt tgagaacgta cctttccata gcagctacgc 1620

tcacagccaa agcctggacc gactaatgaa tccactcatc gaccaatact tgtactatct 1680

ctcaaagact attaacggtt ctggacagaa tcaacaaacg ctaaaattca gtgtggccgg 1740

acccagcaac atggctgtcc agggaagaaa ctacatacct ggacccagct accgacaaca 1800

acgtgtctca accactgtga ctcaaaacaa caacagcgaa tttgcttggc ctggagcttc 1860

ttcttgggct ctcaatggac gtaatagctt gatgaatcct ggacctgcta tggccagcca 1920

caaagaagga gaggaccgtt tctttccttt gtctggatct ttaatttttg gcaaacaagg 1980

aactggaaga gacaacgtgg atgcggacaa agtcatgata accaacgaag aagaaattaa 2040

aactactaac ccggtagcaa cggagtccta tggacaagtg gccacaaacc accagagtgc 2100

ccaagcacag gcgcagaccg gctgggttca aaaccaagga atacttccgg gtatggtttg 2160

gcaggacaga gatgtgtacc tgcaaggacc catttgggcc aaaattcctc acacggacgg 2220

caactttcac ccttctccgc tgatgggagg gtttggaatg aagcacccgc ctcctcagat 2280

cctcatcaaa aacacacctg tacctgcgga tcctccaacg gccttcaaca aggacaagct 2340

gaactctttc atcacccagt attctactgg ccaagtcagc gtggagatcg agtgggagct 2400

gcagaaggaa aacagcaagc gctggaaccc ggagatccag tacacttcca actattacaa 2460

gtctaataat gttgaatttg ctgttaatac tgaaggtgta tatagtgaac cccgccccat 2520

tggcaccaga tacctgactc gtaatctgta attgcttgtt aatcaataaa ccgtttaatt 2580

cgtttcagtt gaactgcggc c 2601

<210> 28

<211> 2601

[0067] <212> DNA

<213> 人工序列

<220>

<223> 新颖AAV分离株

<400> 28

ccatcgacgt cagacgcgga agcttcgatc aactacgcgg acaggtacca aaacaaatgt 60

tctcgtcacg tgggcatgaa tctgatgctg tttccctgca gacaatgcga gagactgaat 120

cagaattcaa atatctgctt cactcacggt gtcaaagact gtttagagtg ctttcccgtg 180

tcagaatctc aacccgtttc tgtcgtcaaa aaggcgtatc agaaactgtg ctacattcat 240

cacatcatgg gaaaggtgcc agacgcttgc actgcttgcg acctggtcaa tgtggacttg 300

gatgactgtg tttctgaaca ataaatgact taaaccaggt atggctgccg atggttatct 360

tccagattgg ctcgaggaca accttagtga aggaattcgc gagtggtggg ctttgaaacc 420

tggagcccct caacccaagg caaatcaaca acatcaagac aacgctcgag gtcttgtgct 480

tccgggttac aaataccttg gacccggcaa cggactcgac aagggggagc cggtcaacgc 540

agcagacgcg gcggccctcg agcacgacaa ggcctacgac cagcagctca aggccggaga 600

caacccgtac ctcaagtaca accacgccga cgccgagttc caggagcggc tcaaagaaga 660

tacgtctttt gggggcaacc tcgggcgagc agtcttccag gccaaaaaga ggcttcttga 720

acctcttggt ctggttgagg aagcggctaa gacggctcct ggaaagaaga ggcctgtaga 780

gcagtctcct cgggaaccgg actcctccgc gggtattggc aaatcgggtg cacagcccgc 840

taaaaagaga ctcaatttcg gtcagactgg cgacacagag tcagtcccag accctcaacc 900

aatcggagaa cctcccgcag ccccctcagg tgtgggatct cttacaatgg cttcaggtgg 960

tggcgcacca gtggcagaca ataacgaagg tgccgatgga gtgggtagtt cctcgggaaa 1020

ttggcattgc gattcccaat ggctggggga cagagtcatc accaccagca cccgaacctg 1080

ggccctgccc acctacaaca atcacctcta caagcaaatc tccaacagca catctggagg 1140

atcttcaaat gacaacgcct acttcggcta cagcaccccc tgggggtatt ttgacttcaa 1200

cagattccac tgccacttct caccacgtga ctggcagcga ctcatcaaca acaactgggg 1260

attccggcct aagcgactca acttcaagct cttcaacatt caggtcaaag aggttacgga 1320

caacaatgga gtcaagacca tcgccaataa ccttaccagc acggtccagg tcttcacgga 1380

ctcagactat cagctcccgt acgtgctcgg gtcggctcac gagggctgcc tcccgccgtt 1440

cccagcggac gttttcatga ttcctcagta cgggtatctg acgcttaatg atggaagcca 1500

ggccgtgggt cgttcgtcct tttactgcct ggaatatttc ccgtcgcaaa tgctaagaac 1560

gggtaacaac ttccagttca gctacgagtt tgagaacgta cctttccata gcagctacgc 1620

tcacagccaa agcctggacc gactaatgaa tccactcatc gaccaatact tgtactatct 1680

[0068] ctcaaagact attaacggtt ctggacagaa tcaacaaacg ctaaaattca gtgtggccgg 1740

acccagcaac atggctgtcc agggaagaaa ctacatacct ggacccagct accgacaaca 1800

acgtgtctca accactgtga ctcaaaacaa caacagcgaa tttgcttggc ctggagcttc 1860

ttcttgggct ctcaatggac gtaatagctt gatgaatcct ggacctgcta tggccagcca 1920

caaagaagga gaggaccgtt tctttccttt gtctggatct ttaatttttg gcaaacaagg 1980

aactggaaga gacaacgtgg atgcggacaa agtcatgata accaacgaag aagaaattaa 2040

aactactaac ccggtagcaa cggagtccta tggacaagtg gccacaaacc accagagtgc 2100

ccaagcacag gcgcagaccg gctgggttca aaaccaagga atacttccgg gtatggtttg 2160

gcaggacaga gatgtgtacc tgcaaggacc catttgggcc aaaattcctc acacggacgg 2220

caactttcac ccttctccgc tgatgggagg gtttggaatg aagcacccgc ctcctcagat 2280

cctcatcaaa aacacacctg tacctgcgga tcctccaacg gccttcaaca aggacaagct 2340

gaactctttc atcacccagt attctactgg ccaagtcagc gtggagatcg agtgggaact 2400

gcagaaggaa aacagcaagc gctggaaccc ggagatccag tacacttcca actattacaa 2460

gtctaataat gttgaatttg ctgttaatac tgaaggtgta tatagtgaac cccgccccat 2520

tggcaccaga tacctgactc gtaatctgta attgcttgtt aatcaataaa ccgtttaatt 2580

cgtttcagtt gaactgcggc c 2601

<210> 29

<211> 2601

<212> DNA

<213> 人工序列

<220>

<223> 新颖 AAV 分离株

<400> 29

ccatcgacgt cagacgcgga agcttcgatc aactacgcgg acaggtacca aaacaaatgt 60

tctcgtcacg tgggcatgaa tctgatgctg tttccctgca gacaatgcga gagactgaat 120

cagaattcaa atatctgctt cactcacggt gtcaaagact gtttagagtg ctttcccgtg 180

tcagaatctc aacccgtttc tgtcgtcaaa aaggcgtatc agaaactgtg ctacattcat 240

cacatcatgg gaaaggtgcc agacgcttgc actgcttgcg acctggtcaa tgtggacttg 300

gatgactgtg tttctgaaca ataaatgact taaaccaggt atggctgccg atggttatct 360

tccagattgg ctcgaggaca accttagtga aggaattcgc gagtggtggg ctttgaaacc 420

tggagcccct caacccaagg caaatcaaca acatcaagac aacgctcgag gtcttgtgct 480

tccgggttac aaataccttg gacccggcaa cggactcgac aagggggagc cggtcaacgc 540

[0069] agcagacgcg gcggccctcg agcacgacaa ggcctacgac cagcagctca aggccggaga 600

caacccgtac ctcaagtaca accacgccga cgccgagttc caggagcggc tcaaagaaga 660

tacgtctttt gggggcaacc tcgggcgagc agtctttcag gccaaaaaga ggcttcttga 720

acctcttggt ctggttgagg aagcggctaa gacggctcct ggaaagaaga ggcctgtaga 780

gcagtctcct caggaaccgg actcctccgc gggtattggc aaatcgggtg cacagcccgc 840

taaaaagaga ctcaatttcg gtcagactgg cgacacagag tcagtcccag accctcaacc 900

aatcggagaa cctcccgcag ccccctcagg tgtgggatct cttacaatgg cttcatgtgg 960

tggcgcacca gtggcagaca ataacgaagg tgccgatgga gtgggtagtt cctcgggaaa 1020

ttggcattgc gattcccaat ggctggggga cagagtcatc accaccagca cccgaacctg 1080

ggccctgccc acctacaaca atcacctcta caagcaaatc tccaacagca catctggagg 1140

atcttcaaat gacaacgcct acttcggcta cagcaccccc tgggggtatt ttgacttcaa 1200

cagattccac tgccacttct caccacgtga ctggcagcga ctcatcaaca acaactgggg 1260

attccggcct aagcgactca acttcaagct cttcaacatt caggtcaaag aggttacgga 1320

caacaatgga gtcaagacca tcgccaataa ccttaccagc acggtccagg tcttcacgga 1380

ctcagactat cagctcccgt acgtgctcgg gtcggctcac gagggctgcc tcccgccgtt 1440

cccagcggac gttttcatga ttcctcagta cgggtatctg acgcttaatg atggaagcca 1500

ggccgtgggt cgttcgtcct tttactgcct ggaatatttc ccgtcgcaaa tgctaagaac 1560

gggtaacaac ttccagttca gctacgagtt tgagaacgta cctttccata gcagctacgc 1620

tcacagccaa agcctggacc gactaatgaa tccactcatc gaccaatact tgtactatct 1680

ctcaaagact attaacggtt ctggacagaa tcaacaaacg ctaaaattca gtgtggccgg 1740

acccagcaac atggctgtcc agggaagaaa ctacatacct ggacccagct accgacaaca 1800

acgtgtctca accactgtga ctcaaaacaa caacagcgaa tttgcttggc ctggagcttc 1860

ttcttgggct ctcaatggac gtaatagctt gatgaatcct ggacctgcta tggccagcca 1920

caaagaagga gaggaccgtt tctttccttt gtctggatct ttaatttttg gcaaacaagg 1980

aactggaaga gacaacgtgg atgcggacaa agtcatgata accaacgaag aagaaattaa 2040

aactactaac ccggtagcaa cggagtccta tggacaagtg gccacaaacc accagagtgc 2100

ccaagcacag gcgcagaccg gctgggttca aaaccaagga atacttccgg gtatggtttg 2160

gcaggacaga gatgtgtacc tgcaaggacc catttgggcc aaaattcctc acacggacgg 2220

caactttcac ccttctccgc tgatgggagg gtttggaatg aagcacccgc ctcctcagat 2280

cctcatcaaa aacacacctg tacctgcgga tcctccaacg gccttcaaca aggacaagct 2340

gaactctttc atcacccagt attctactgg ccaagtcagc gtggagatcg agtgggagct 2400

gcagaaggaa aacagcaagc gctggaaccc ggagatccag tacacttcca actattacaa 2460

gtctaataat gttgaatttg ctgttaatac tgaaggtgta tatagtgaac cccgccccat 2520

tggcaccaga tacctgactc gtaatctgta attgcttgtt aatcaataaa ccgtttaatt 2580

cgtttcagtt gaactgcggc c 2601

[0070]

<210> 30

<211> 2601

<212> DNA

<213> 人工序列

<220>

<223> 新颖 AAV 分离株

<400> 30

ccatcgacgt cagacgcgga agcttcgatc aactacgcgg acaggtacca aaacaaatgt 60

tctcgccacg tgggcatgaa tctgatgctg tttccctgca gacaatgcga gagactgaat 120

cagaattcaa atatctgctt cactcacggt gtcaaagact gtttagagtg ctttcccgtg 180

tcagaatctc aacccgtttc tgtcgtcaaa aaggcgtatc agaaactgtg ctacattcat 240

cacatcatgg gaaaggtgcc agacgcttgc actgcttgcg acctggtcaa tgtggacttg 300

gatgactgtg tttctgaaca ataaatgact taaaccaggt atggctgccg atggttatct 360

tccagattgg ctcgaggaca accttagtga aggaattcgc gagtggtggg ctttgaaacc 420

tggagcccct caacccaagg caaatcaaca acatcaagac aacgctcgag gtcttgtgct 480

tccgggttac aaataccttg gacccggcaa cggactcgac aagggggagc cggtcaacgc 540

agcagacgcg gcggccctcg agcacgacaa ggcctacgac cagcagctca aggccggaga 600

caacccgtac ctcaagtaca accacgccga cgccgagttc caggagcggc tcaaagaaga 660

tacgtctttt gggggcaacc tcgggcgagc agtcttccag gccaaaaaga ggcttcttga 720

acctcttggt ctggttgagg aagcggctaa gacggctcct ggaaagaaga ggcctgtaga 780

gcagtctcct caggaaccgg actcctccgc gggtattggc aaatcgggtg cacagcccgc 840

taaaaagaga ctcaatttcg gtcagactgg cgacacagag tcagtcccag accctcaacc 900

aatcggagaa cctcccgctg ccccctcagg tgtgggatct cttacaatgg cttcaggtgg 960

tggcgcacca gtggcagaca ataacgaagg tgccgatgga gtgggtagtt cctcgggaaa 1020

ttggcattgc gattcccaat ggctggggga cagagtcatc accaccagca cccgaacctg 1080

ggccctgccc acctacaaca atcacctcta caagcaaatc tccaacagca catctggagg 1140

atcttcaaat gacaacgcct acttcggcta cagcaccccc tgggggtatt ttgacttcaa 1200

cagattccac tgccacttct caccacatga ctggcagcga ctcatcaaca acaactgggg 1260

attccggcct aagcgactca acttcaagct cttcaacatt caggtcaaag aggttacgga 1320

caacaatgga gtcaagacca tcgccaataa ccttaccagc acggtccagg tcttcacgga 1380

ctcagactat cagctcccgt acgtgctcgg gtcggctcac gagggctgcc tcccgccgtt 1440

cccagcggac gttttcatga ttcctcagta cgggtatctg acgcttaatg atggaagcca 1500

[0071] ggccgtgggt cgttcgtcct tttactgcct ggaatatttc ccgtcgcaaa tgctaagaac 1560

gggtaacaac ttccagttca gctacgagtt tgagaacgta cctttccata gcagctacgc 1620

tcacagccaa agcctggacc gactaatgaa tccactcatc gaccaatact tgtactatct 1680

ctcaaagact attaacggtt ctggacagaa tcaacaaacg ctaaaattca atgtggccgg 1740

acccagcaac atggctgtcc agggaagaaa ctacatacct ggacccagct accgacaaca 1800

acgtgtctca accactgtga ctcaaaacaa caacagcgaa tttgcctggc ctagagcttc 1860

ttcttgggct ctcaatggac gtaatagctt gatgaatcct ggacctgcta tggccagcca 1920

caaagaagga gaggaccgtt tctttccttt gtctggatct ttaatttttg gcaaacaagg 1980

aactggaaga gacaacgtgg atgcggacaa agtcatgata accaacgaag aagaaattaa 2040

aactactaac ccggtagcaa cggagtccta tggacaagtg gccacaaacc accagagtgc 2100

ccaagcacag gcgcagaccg gctgggttca aaaccaagga atacttccgg gtatggtttg 2160

gcaggacaga gatgtgtacc tgcaaggacc catttgggcc aaaattcctc acacggacgg 2220

caactttcac ccttctccgc tgatgggagg gtttggaatg aagcacccgc ctcctcagat 2280

cctcatcaaa aacacacctg tacctgcgga tcctccaacg gccttcaaca aggacaagct 2340

gaactctttc atcacccagt attctactgg ccaagtcagc atggagatcg agtgggagct 2400

gcagaaggaa aacagcaagc gctggaaccc ggagatccag tacacttcca actattacaa 2460

gtctaataat gttgaatttg ctgttaatac tgaaggtgta tatagtgaac cccgccccat 2520

tggcaccaga tacctgactc gtaatctgta attgcttgtt aatcaataaa ccgtttaatt 2580

cgtttcagtt gaactgcggc c 2601

<210> 31

<211> 2601

<212> DNA

<213> 人工序列

<220>

<223> 新颖AAV分离株

<400> 31

ccatcgacgt cagacgcgga agcttcgatc aactacgcgg acaggtacca aaacaaatgt 60

tctcgtcacg tgggcatgaa tctgatgctg tttccctgca gacaatgcga gagactgaat 120

cagaattcaa atatctgctt cactcacggt gtcaaagact gtttagagtg ctttcccgtg 180

tcagaatctc aacccgtttc tgtcgtcaaa aaggcgtatc agaaactgtg ctacattcat 240

cacatcatgg gaaaggtgcc agacgcttgc actgcttgcg acctggtcaa tgtggacttg 300

gatgactgtg tttctgaaca ataaatgact taaaccaggt atggctgccg atggttatct 360

[0072] tccagattgg ctcgaggaca accttagtga aggaattcgc gagtggtggg ctttgaaacc 420

tggagcccct caacccaagg caaatcaaca acatcaagac aacgctcgag gtcttgtgct 480

tccgggttac aaataccttg gacccggcaa cggactcgac aagggggagc cggtcaacgc 540

agcagacgcg gcggccctcg agcacgacaa ggcctacgac cagcagctca aggccggaga 600

caacccgtac ctcaagtaca accacgccga cgccgagttc caggagcggc tcaaagaaga 660

tacgtctttt gggggcaacc tcgggcgagc agtcttccag gccaaaaaga ggcttcttga 720

acctcttggt ctggttgagg aagcggctaa gacggctcct ggaaagaaga ggcctgtaga 780

gcagtctcct caggaaccgg actcctccgc gggtattggc aaatcgggtg cacagcccgc 840

taaaaagaga ctcaatttcg gtcagactgg cgacacagag tcagtcccag accctcaacc 900

aatcggagaa cctcccgcag ccccctcagg tgtgggatct cttacaatgg cttcaggtgg 960

tggcgcacca gtggcagaca ataacgaagg tgccgatgga gtgggtagtt cctcgggaaa 1020

ttggcattgc gattcccaat ggctggggga cagagtcatc accaccagca cccgaacctg 1080

ggccctgccc acctacaaca atcacctcta caagcaaatc tccaacagca catctggagg 1140

atcttcaaat gacaacgcct acttcggcta cagcaccccc tgggggtatt ttgacttcaa 1200

cagattccac tgccacttct caccacgtga ctggcagcga ctcatcaaca acaactgggg 1260

attccggcct aagcgactca acttcaagct cttcaacatt caggtcaaag aggttacgga 1320

caacaatgga gtcaagacca tcgccaataa ccttaccagc acggtccagg tcttcacgga 1380

ctcagactat cagctcccgt acgtgctcgg gtcggctcac gagggctgcc tcccgccgtt 1440

cccagcggac gttttcatga ttcctcagta cgggtatctg acgcttaatg atggaagcca 1500
ggccgtgggt cgttcgtcct tttactgcct ggaatatttc ccgtcgcaaa tgctaagaac 1560
gggtaacaac ttccagttca gctacgagtt tgagaacgta cctttccata gcagctacgc 1620
tcacagccaa agcctggacc gactaatgaa tccactcatc gaccaatact tgtactatct 1680
ctcaaagact attaacggtt ctggacagaa tcaacaaacg ctaaaattca gtgtggccgg 1740
acccagcaac atggctgtcc agggaagaaa ctacatacct ggacccagct accgacaaca 1800
acgtgtctca accactgtga ctcaaaacaa caacagcgaa tttgcttggc ctagagcttc 1860
ttcttgggct ctcaatggac gtaatagctt gatgaatcct ggacctgcta tggccagcca 1920
caaagaagga gaggaccgtt tctttccttt gtctggatct ttaatttttg gcaaacaagg 1980
aactggaaga gacaacgtgg atgcggacaa agtcatgata accaacgaag aagaaattaa 2040
aactactaac ccggtagcaa cggagtccta tggacaagtg gccacaaacc accagagtgc 2100
ccaagcacag gcgcagaccg gctgggttca aaaccaagga atacttccgg gtatggtttg 2160
gcaggacaga gatgtgtacc tgcaaggacc catttgggcc aaaattcctc acacggacgg 2220
caactttcac ccttctccgc tgatgggagg gtttggaatg aagcacccgc ctcctcagat 2280
cctcatcaaa aacacacctg tacctgcgga tcctccaacg gccttcaaca aggacaagct 2340
gaactctttc atcacccagt attctactgg ccaagtcagc gtggagatcg agtgggagct 2400
gcagaaggaa aacagcaagc gctggaaccc ggagatccag tacacttcca actattacaa 2460
[0073] gtctaataat gttgaatttg ctgttaatac tgaaggtgta tatagtgaac cccgccccat 2520
tggcaccaga tacctgactc gtaatctgta attgcttgtt aatcaataaa ccgtttaatt 2580
cgtttcagtt gaactgcggc c 2601

<210> 32
<211> 2601
<212> DNA
<213> 人工序列

<220>
<223> 新颖AAV分离株

<400> 32
ccatcgacgt cagacgcgga agcttcgatc aactacgcgg acaggtacca aaacaaatgt 60
tctcgtcacg tgggcatgaa tctgatgctg tttccctgca gacaatgcga gagactgaat 120
cagaattcaa atatctgctt cactcacggt gtcaaagact gtttagagtg ctttcccgtg 180
tcagaatctc aacccgtttc tgtcgtcaaa aaggcgtatc agaaactgtg ctacattcat 240
cacatcatgg gaaaggtgcc agacgcttgc actgcttgcg acctggtcaa tgtggacttg 300

gatgactgtg tttctgaaca ataaatgact taaaccaggt atggctgccg atggttatct 360

tccagattgg ctcgaggaca accttagtga aggaattcgc gagtggtggg ctttgaaacc 420

tggagcccct caacccaagg caaatcaaca acatcaagac aacgctcgag gtcttgtgct 480

tccgggttac aaataccttg gacccggcaa cggactcgac aagggggagc cggtcaacgc 540

agcagacgcg gcggccctcg agcacgacaa ggcctacgac cagcagctca aggccggaga 600

caacccgtac ctcaagtaca accacgccga cgccgagttc caggagcggc tcaaagaaga 660

tacgtctttt gggggcaacc tcgggcgagc agtcttccag gccaaaaaga ggcttcttga 720

acctcttggt ctggttgagg aagcggctaa gacggctcct ggaaagaaga ggcctgtaga 780

gcagtctcct caggaaccgg actcctccgc gggtattggc aaatcgggtg cacagcccgc 840

taaaaagaga ctcaatttcg gtcagactgg cgacacagag tcagtcccag accctcaacc 900

aatcggagaa cctcccgcag ccccctcagg tgtgggatct cttacaatgg cttcaggtgg 960

tggcgcacca gtggcagaca ataacgaagg tgccgatgga gtgggtagtt cctcgggaaa 1020

ttggcattgc gattcccaat ggctggggga cagagtcatc accaccagca cccgaacctg 1080

ggccctgccc acctacaaca atcacctcta caagcaaatc tccaacagca catctggagg 1140

atcttcaaat gacaacgcct acttcggcta cagcaccccc tgggggtatt ttgacttcaa 1200

cagattccac tgccacttct caccacgtga ctggcagcga ctcatcaaca acaactgggg 1260

attccggcct aagcgactca acttcaagct cttcaacatt caggtcaaag aggttacgga 1320

[0074] caacaatgga gtcaagacca tcgccaataa ccttaccagc acggtccagg tcttcacgga 1380

ctcagactat cagctcccgt acgtgctcgg gtcggctcac gagggctgcc tcccgccgtt 1440

cccagcggac gttttcatga ttcctcagta cgggtatctg acgcttaatg atggaagcca 1500

ggccgtgggt cgttcgtcct tttactgcct ggaatatttc ccgtcgcaaa tgctaagaac 1560

gggtaacaac ttccagttca gctacgagtt tgagaacgta cctttccata gcagctacgc 1620

tcacagccaa agcctggacc gactaatgaa tccactcatc gaccaatact tgtactatct 1680

ctcaaagact attaacggtt ctggacagaa tcaacaaacg ctaaaattca gtgtggccgg 1740

acccagcaac atggctgtcc agggaagaaa ctacatacct ggacccagct accgacaaca 1800

acgtgtctca accactgtga ctcaaaacaa caacagcgaa tttgcttggc ctagagcttc 1860

ttcttgggct ctcaatggac gtaatagctt gatgaatcct ggacctgcta tggccagcca 1920

caaagaagga gaggaccgtt tctttccttt gtctggatct ttaatttttg gcaaacaagg 1980

aactggaaga gacaacgtgg atgcggacaa agtcatgata accaacgaag aagaaattaa 2040

aactactaac ccggtagcaa cggagtccta tggacaagtg gccacaaacc accagagtgc 2100

ccaagcacag gcgcagaccg gctgggttca aaaccaagga atacttccgg gtatggtttg 2160

gcaggacaga gatgtgtacc tgcaaggacc catttgggcc aaaattcctc acacggacgg 2220

caactttcac ccttctccgc tgatgggagg gtttggaatg aagcacccgc ctcctcagat 2280

cctcatcaaa aacacacctg tacctgcgga tcctccaacg gccttcaaca aggacaagct 2340

gaactctttc atcacccagt attctactgg ccaagtcagc gtggagatcg agtgggagcg 2400

gcagaaggaa aacagcaagc gctggaaccc ggagatccag tacacttcca actattacaa 2460

gtctaataat gttgaatttg ctgttaatac tgaaggtgta tatagtgaac cccgccccat 2520

tggcaccaga tacctgactc gtaatctgta attgcttgtt aatcaataaa ccgtttaatt 2580

cgtttcagtt gaactgcggc c 2601

<210> 33

<211> 2601

<212> DNA

<213> 人工序列

<220>

<223> 新颖 AAV 分离株

<400> 33

ccatcgacgt cagacgcgga agcttcgatc aactacgcgg acaggtacca aaacaaatgt 60

tctcgtcacg tgggcatgaa tctgatgctg tttccctgca gacaatgcga gagactgaat 120

cagaattcaa atatctgctt cactcacggt gtcaaagact gtttagagtg ctttcccgtg 180

[0075] tcagaatctc aacccgtttc tgtcgtcaaa aaggcgtatc agaaactgtg ctacattcat 240

cacatcatgg gaaaggtgcc agacgcttgc actgcttgcg acctggtcaa tgtggacttg 300

gatgactgtg tttctgaaca ataaatgact taaaccaggt atggctgccg atggttatct 360

tccagattgg ctcgaggaca accttagtga aggaattcgc gagtggtggg ctttgaaacc 420

tggagcccct caacccaagg caaatcaaca acatcaagac aacgctcgag gtcttgtgct 480

tccgggttac aaataccttg gacccggcaa cggactcgac aagggggagc cggtcaacgc 540

agcagacgcg gcggccctcg agcacgacaa ggcctacgac cagcagctca aggccggaga 600

caacccgtac ctcaagtaca accacgccga cgccgagttc caggagcggc tcaaagaaga 660

tacgtctttt gggggcaacc tcgggcgagc agtcttccag gccaaaaaga ggcttcttga 720

acctcttggt ctggttgagg aagcggctaa gacggctcct ggaaagaaga ggcctgtaga 780

gcagtctcct caggaaccgg actcctccgc gggtattggc aaatcgggtg cacagcccgc 840

taaaaagaga ctcaatttcg gtcagactgg cgacacagag tcagtcccag accctcaacc 900

aatcggagaa cctcccgcag ccccctcagg tgtgggatct cttacaatgg cttcaggtgg 960

tggcgcacca gtggcagaca ataacgaagg tgccgatgga gtgggtagtt cctcgggaaa 1020

ttggcattgc gattcccaat ggctggggga cagagtcatc accaccagca cccgaacctg 1080

ggccctgccc acctacaaca atcacctcta caagcaaatc tccaacagca catctggagg 1140

atcttcaaat gacaacgcct acttcggcta cagcaccccc tgggggtatt ttgacttcaa 1200

cagattccac tgccacttct caccacgtga ctggcagcga ctcatcaaca acaactgggg 1260

attccggcct aagcgactca acttcaagct cttcaacatt caggtcaaag aggttacgga 1320

caacaatgga gtcaagacca tcgccaataa ccttaccagc acggtccagg tcttcgcgga 1380

ctcagactat cagctcccgt acgtgctcgg gtcggctcac gagggctgcc tcccgccgtt 1440

cccagcggac gttttcatga ttcctcagta cgggtatctg acgcttaatg atggaagcca 1500

ggccgtgggt cgttcgtcct tttactgcct ggaatatttc ccgtcgcaaa tgctaagaac 1560

gggtaacaac ttccagttca gctacgagtt tgagaacgta cctttccata gcagctacgc 1620

tcacagccaa agcctggacc gactaatgaa tccactcatc gaccaatact tgtactatct 1680

ctcaaagact attaacggtt ctggacagaa tcaacaaacg ctaaaattca gtgtggccgg 1740

acccagcaac atggctgtcc agggaagaaa ctacatacct ggacccagct accgacaaca 1800

acgtgtctca accactgtga ctcaaaacaa caacagcgaa tttgcttggc ctagagcttc 1860

ttcttgggct ctcaatggac gtaatagctt gatgaatcct ggacctgcta tggccagcca 1920

caaagaagga gaggaccgtt tctttccttt gtctggatct ttaatttttg gcaaacaagg 1980

aactggaaga gacaacgtgg atgcggacaa agtcatgata accaacgaag aagaaattaa 2040

aactactaac ccggtagcaa cggagtccta tggacaagtg gccacaaacc accagagtgc 2100

ccaagcacag gcgcagaccg gctgggttca aaaccaagga atacttccgg gtatggtttg 2160

gcaggacaga gatgtgtacc tgcaaggacc catttgggcc aaaattcctc acacggacgg 2220

caactttcac ccttctccgc tgatgggagg gtttggaatg aagcacccgc ctcctcagat 2280

[0076] cctcatcaaa aacacacctg tacctgcgga tcctccaacg gccttcaaca aggacaagct 2340

gaactctttc atcacccagt attctactgg ccaagtcagc gtggagatcg agtgggagct 2400

gcagaaggaa aacagcaagc gctggaaccc ggagatccag tacacttcca actattacaa 2460

gtctaataat gttgaatttg ctgttaatac tgaaggtgta tatagtgaac cccgccccat 2520

tggcaccaga tacctgactc gtaatctgta attgcttgtt aatcaataaa ccgtttaatt 2580

cgtttcagtt gaactgcggc c 2601

<210> 34

<211> 2601

<212> DNA

<213> 人工序列

<220>

<223> 新颖 AAV 分离株

<400> 34

ccatcgacgt cagacgcgga agcttcgatc aactacgcgg acaggtacca aaacaaatgt 60

tctcgtcacg tgggcatgaa tctgatgctg tttccctgca gacaatgcga gagactgaat 120

cagaattcaa atatctgctt cactcacggt gtcaaagact gtttagagtg ctttcccgtg 180

tcagaatctc aacccgtttc tgtcgtcaaa aaggcgtatc agaaactgtg ctacattcat 240

cacatcatgg gaaaggtgcc agacgcttgc actgcttgcg acctggtcaa tgtggacttg 300

gatgactgtg tttctgaaca ataaatgact taaaccaggt atggctgccg atggttatct 360

tccagattgg ctcgaggaca accttagtga aggaattcgc gagtggtggg ctttgaaacc 420

tggagcccct caacccaagg caaatcaaca acatcaagac aacgctcgag gtcttgtgct 480

tccgggttac aaataccttg gacccggcaa cggactcgac aagggggagc cggtcaacgc 540

agcagacgcg gcggccctcg agcacgacaa ggcctacgac cagcagctca aggccggaga 600

caacccgtac ctcaagtaca accacgccga cgccgagttc caggagcggc tcaaagaaga 660

tacgtctttt gggggcaacc tcgggcgagc agtcttccag gccaaaaaga ggcttcttga 720

acctcttggt ctggttgagg aagcggctaa gacggctcct ggaaagaaga ggcctgtaga 780

gcagtctcct caggaaccgg actcctccgc gggtattggc aaatcgggtg cacagcccgc 840

taaaaagaga ctcaatttcg gtcagactgg cgacacagag tcagtcccag accctcaacc 900

aatcggagaa cctcccgcag ccccctcagg tgtgggatct cttacaatgg cttcaggtgg 960

tggcgcacca gtggcagaca ataacgaagg tgccgatgga gtgggtagtt cctcgggaaa 1020

ttggcattgc gattcccaat ggctggggga cagagtcatc accaccagca cccgaacctg 1080

ggccctgccc acctacaaca atcacctcta caagcaaatc tccaacagca catctggagg 1140

[0077] atcttcaaat gacaacgcct acttcggcta cagcaccccc tgggggtatt ttgacttcaa 1200

cagattccac tgccacttct caccacgtga ctggcagcga ctcatcaaca acaactgggg 1260

attccggcct aagcgactca acttcaagct cttcaacatt caggtcaaag aggttacgga 1320

caacaatgga gtcaagacca tcgccaataa ccttaccagc acggtccagg tcttcacgga 1380

ctcagactat cagctcccgt acgtgctcgg gtcggctcac gagggctgcc tcccgccgtt 1440

cccagcggac gttttcatga ttcctcagta cgggtatctg acgcttaatg atggaagcca 1500

ggccgtgggt cgttcgtcct tttactgcct ggaatatttc ccgtcgcaaa tgctaagaac 1560

gggtaacaac ttccagttca gctacgagtt tgagaacgta cctttccata gcagctacgc 1620

tcacagccaa agcctggacc gactaatgaa tccactcatc gaccaatact tgtactatct 1680

ctcaaagact attaacggtt ctggacagaa tcaacaaacg ctaaaattca gtgtggccgg 1740

acccagcaac atggctgtcc agggaagaaa ctacatacct ggacccagct accgacaaca 1800

acgtgtctca accactgtga ctcaaaacaa caacagcgaa attgcttggc ctagagcttc 1860

ttcttgggct ctcaatggac gtaatagctt gatgaatcct ggacctgcta tggccagcca 1920

caaagaagga gaggaccgtt tctttccttt gtctggatct ttaatttttg gcaaacaagg 1980

aactggaaga gacaacgtgg atgcggacaa agtcatgata accaacgaag aagaaattaa 2040

aactactaac ccagtagcaa cggagtccta tggacaagtg gccacaaacc accagagtgc 2100

ccaagcacag gcgcagaccg gctgggttca aaaccaagga atacttccgg gtatggtttg 2160

gcaggacaga gatgtgtacc tgcaaggacc catttgggcc aaaattcctc acacggacgg 2220

caactttcac ccttctccgc tgatgggagg gtttggaatg aagcacccgc ctcctcagat 2280
cctcatcaaa aacacacctg tacctgcgga tcctccaacg gccttcaaca aggacaagct 2340
gaactctttc atcacccagt attctactgg ccaagtcagc gtggagatcg agtgggagct 2400
gcagaaggaa aacagcaagc gctggaaccc ggagatccag tacacttcca actattgcaa 2460
gtctaataat gttgaatttg ctgttaatac tgaaggtgta tatagtgaac cccgccccat 2520
tggcaccaga tacctgactc gtaatctgta attgcttgtt aatcaataaa ccgtttaatt 2580
cgtttcagtt gaactgcggc c 2601

<210> 35
<211> 2601
<212> DNA
<213> 人工序列

<220>
<223> 新颖 AAV 分离株

<400> 35
[0078] ccatcgacgt cagacgcgga agcttcgatc aactacgcgg acaggtacca aaacaaatgt 60
tctcgtcacg tgggcatgaa tctgatgctg tttccctgca gacaatgcga gagactgaat 120
cagaattcaa atatctgctt cactcacggt gtcaaagact gtttagagtg ctttcccgtg 180
tcagaatctc aacccgtttc tgtcgtcaaa aaggcgtatc agaaactgtg ctacattcat 240
cacatcatgg gaaaggtgcc agacgcttgc actgcttgcg acctggtcaa tgtggacttg 300
gatgtctgtg tttctgaaca ataaatgact taaaccaggt atggctgccg atggttatct 360
tccagattgg ctcgaggaca accttagtga aggaattcgc gagtggtggg ctttgaaacc 420
tggagcccct caacccaagg caaatcaaca acatcaagac aacgctcgag gtcttgtgct 480
tccgggttac aaataccttg gacccggcaa cggactcgat aagggggagc cggtcaacgc 540
agcagacgcg gcggccctcg agcacgacaa ggcctacgac cagcagctca aggccggaga 600
caacccgtac ctcaagtaca accacgccga cgccgagttc caggagcggc tcaaagaaga 660
tacgtctttt gggggcaacc tcgggcgagc agtcttccag gccaaaaaga ggcttcttga 720
acctcttggt ctggttgagg aagcggctaa gacggctcct ggaaagaaga ggcctgtaga 780
gcagtctcct caggaaccgg actcctccgc gggtattggc aaatcgggtg cacagcccgc 840
taaaaagaga ctcaatttcg gtcagactgg cgacacagag tcagtcccag accctcaacc 900
aatcggagaa cctcccgcag ccccctcagg tgtgggatct cttacaatgg cttcaggtgg 960
tggcgcacca gtggcagaca ataacgaagg tgccgatgga gtgggtagtt cctcgggaaa 1020
ttggcattgc gattcccaat ggctggggga cagagtcatc accaccagca cccgaacctg 1080

[0079]

ggccctgccc acctacaaca atcacctcta caagcaaatc tccaacagca catctggagg 1140
atcttcaaat gacaacgcct acttcggcta cagcaccccc tgggggtatt ttgacttcaa 1200
cagattccac tgccacttct caccacgtga ctggcagcga ctcatcaaca acaactgggg 1260
attccggcct aagcgactca acttcaagct cttcaacatt caggtcaaag aggttacgga 1320
caacaatgga gtcaagacca tcgccaataa ccttaccagc acggtccagg tcttcacgga 1380
ctcagactat cagctcccgt acgtgctcgg gtcggctcac gagggctgcc tcccgccgtt 1440
cccagcggac gttttcatga ttcctcagta cgggtatctg acgcttaatg atggaagcca 1500
ggccgtgggt cgttcgtcct tttactgcct ggaatatttc ccgtcgcaaa tgctaagaac 1560
gggtaacaac ttccagttca gctacgagtt tgagaacgta cctttccata gcagctacgc 1620
tcacagccaa agcctggacc gactaatgaa tccactcatc gaccaatact tgtactatct 1680
ctcaaagact attaacggtt ctggacagaa tcaacaaacg ctaaaattca gtgtggccgg 1740
acccagcaac atggctgtcc agggaagaaa ctacatacct ggacccagct accgacaaca 1800
acgtgtctca accactgtga ctcaaaacaa caacagcgaa tttgcttggc ctagagcttc 1860
ttcttgggct ctcaatggac gtaatagctt gatgaatcct ggacctgcta tggccagcca 1920
caaagaagga gaggaccgtt tctttccttt gtctggatct ttaatttttg gcaaacaagg 1980
aactggaaga gacaacgtgg atgcggacaa agtcatgata accaacgaag aagaaattaa 2040
aactactaac ccggtagcaa cggagtccta tggacaagtg gccacaaacc accagagtgc 2100
ccaagcacag gcgcagaccg gctgggttca aaaccaagga atacttccgg gtatggtttg 2160
gcaggacaga gatgtgtacc tgcaaggacc catttgggcc aaaattcctc acacggacgg 2220
caactttcac ccttctccgc tgatgggagg gtttggaatg aagcacccgc ctcctcagat 2280
cctcatcaaa aacacacctg tacctgcgga tcctccaacg gccttcaaca aggacaagct 2340
gaactctttc atcacccagt attctactgg ccaagtcagc gtggagatcg agtgggagct 2400
gcagaaggaa aacagcaagc gctggaaccc ggagatccag tacacttcca actattacaa 2460
gtctaataat gttgaatttg ctgttaatac tgaaggtgta tatagtgaac cccgccccat 2520
tggcaccaga tacctgactc gtaatctgta attgcttgtt aatcaataaa ccgtttaatt 2580
cgtttcagtt gaactgcggc c 2601

<210> 36
<211> 145
<212> DNA
<213> 人工序列

<220>
<223> AAV2 5' ITR

<400> 36

ttggccactc cctctctgcg cgctcgctcg ctcactgagg ccgggcgacc aaaggtcgcc 60

cgacgcccgg gctttgcccg ggcggcctca gtgagcgagc gagcgcgcag agagggagtg 120

gccaactcca tcactagggg ttcct 145

<210> 37

<211> 145

<212> DNA

<213> 人工序列

<220>

<223> AAV2 3' ITR

<400> 37

aggaacccct agtgatggag ttggccactc cctctctgcg cgctcgctcg ctcactgagg 60

ccgggcgacc aaaggtcgcc cgacgcccgg gctttgcccg ggcggcctca gtgagcgagc 120

gagcgcgcag agagggagtg gccaa 145

[0080]

<210> 38

<211> 167

<212> DNA

<213> 人工序列

<220>

<223> AAV5 5' ITR

<400> 38

ctctcccccc tgtcgcgttc gctcgctcgc tggctcgttt ggggggggtgg cagctcaaag 60

agctgccaga cgacggccct ctggccgtcg cccccccaaa cgagccagcg agcgagcgaa 120

cgcgacaggg gggagagtgc cacactctca agcaaggggg ttttgta 167

<210> 39

<211> 167

<212> DNA

<213> 人工序列

<220>

<223> AAV5 3' ITR

<400> 39

tacaaaacct ccttgcttga gagtgtggca ctctcccccc tgtcgcgttc gctcgctcgc 60

tggctcgttt ggggggtgg cagctcaaag agctgccaga cgacggccct ctggccgtcg 120

ccccccccaaa cgagccagcg agcgagcgaa cgcgacaggg gggagag 167

<210> 40

<211> 621

<212> PRT

<213> 人工序列

<220>

<223> AAV2 Rep

<400> 40

[0081]

Met Pro Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp

1 5 10 15

Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu

20 25 30

Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile

35 40 45

Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu

50 55 60

Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val

65 70 75 80

Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu

85 90 95

Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile

100 105 110

Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu

115 120 125

Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly
130 135 140
Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys
145 150 155 160
Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu
165 170 175
Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His
180 185 190
Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn
195 200 205
Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr
210 215 220
Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys
225 230 235 240
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala
245 250 255
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys
[0082] 260 265 270
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln
275 280 285
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu
290 295 300
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala
305 310 315 320
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala
325 330 335
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro
340 345 350
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp
355 360 365
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala
370 375 380
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg
385 390 395 400
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val

405 410 415

Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser

420 425 430

Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe

435 440 445

Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln

450 455 460

Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val

465 470 475 480

Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala

485 490 495

Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val

500 505 510

[0083] Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp

515 520 525

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

530 535 540

Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys

545 550 555 560

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

565 570 575

Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr

580 585 590

Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp

595 600 605

Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln

610 615 620

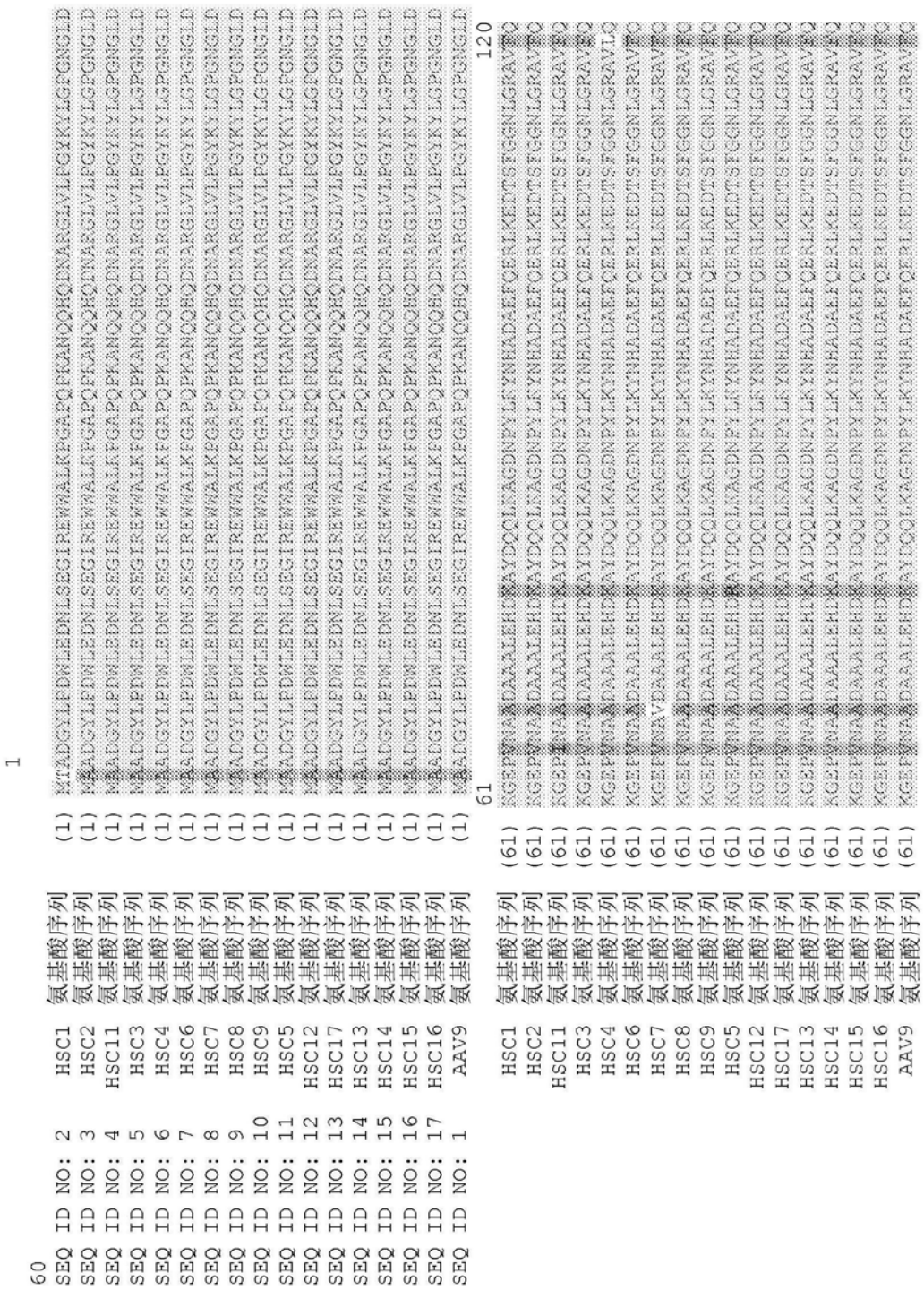

1
60
SEQ ID NO: 2 HSC1 氨基酸序列 (1) MTADGYLPDWLEDNLSEGIREWWALKPGAPQPKANQQHQDNARGLVLPGYKYLGPGNGLD
SEQ ID NO: 3 HSC2 氨基酸序列 (1) MAADGYLPDWLEDNLSEGIREWWALKPGAPQPKANQQHQDNARGLVLPGYKYLGPGNGLD
SEQ ID NO: 4 HSC11 氨基酸序列 (1) MADADGYLPDWLEDNLSEGIREWWALKPGAPQPKANQQHQDNARGLVLPGYKYLGPGNGLD
SEQ ID NO: 5 HSC3 氨基酸序列 (1) MADADGYLPDWLEDNLSEGIREWWALKPGAPQPKANQQHQDNARGLVLPGYKYLGPGNGLD
SEQ ID NO: 6 HSC4 氨基酸序列 (1) MADADGYLPDWLEDNLSEGIREWWALKPGAPQPKANQQHQDNARGLVLPGYKYLGPGNGLD
SEQ ID NO: 7 HSC6 氨基酸序列 (1) MADADGYLPDWLEDNLSEGIREWWALKPGAPQPKANQQHQDNARGLVLPGYKYLGPGNGLD
SEQ ID NO: 8 HSC7 氨基酸序列 (1) MADADGYLPDWLEDNLSEGIREWWALKPGAPQPKANQQHQDNARGLVLPGYKYLGPGNGLD
SEQ ID NO: 9 HSC8 氨基酸序列 (1) MADADGYLPDWLEDNLSEGIREWWALKPGAPQPKANQQHQDNARGLVLPGYKYLGPGNGLD
SEQ ID NO: 10 HSC9 氨基酸序列 (1) MADADGYLPDWLEDNLSEGIREWWALKPGAPQPKANQQHQDNARGLVLPGYKYLGPGNGLD
SEQ ID NO: 11 HSC5 氨基酸序列 (1) MADADGYLPDWLEDNLSEGIREWWALKPGAPQPKANQQHQDNARGLVLPGYKYLGPGNGLD
SEQ ID NO: 12 HSC12 氨基酸序列 (1) MADADGYLPDWLEDNLSEGIREWWALKPGAPQPKANQQHQTNARGLVLPGYKYLGPGNGLD
SEQ ID NO: 13 HSC17 氨基酸序列 (1) MADADGYLPTWLEDNLSEGIREWWALKPGAPQPKANQQHQTNARGLVLPGYKYLGPGNGLD
SEQ ID NO: 14 HSC13 氨基酸序列 (1) MADADGYLPTWLEDNLSEGIREWWALKPGAPQPKANQQHQDNARGLVLPGYNYLGPGNGLD
SEQ ID NO: 15 HSC14 氨基酸序列 (1) MADADGYLPTWLEDNLSEGIREWWALKPGAPQPKANQQHQDNARGLVLPGYNYLGPGNGLD
SEQ ID NO: 16 HSC15 氨基酸序列 (1) MADADGYLPDWLEDNLSEGIREWWALKPGAPQPKANQQHQDNARGLVLPGYKYLGPGNGLD
SEQ ID NO: 17 HSC16 氨基酸序列 (1) MADADGYLPDWLEDNLSEGIREWWALKPGAPQPKANQQHQDNARGLVLPGYKYLGPGNGLD
SEQ ID NO: 1 AAV9 氨基酸序列 (1) MAADGYLPDWLEDNLSEGIREWWALKPGAPQPKANQQHQDNARGLVLPGYKYLGPGNGLE
61
120
HSC1 氨基酸序列 (61) KGEPVNAADAAALEHDKAYDQQLKAGDNPYLKYNHADAEFQERLKEDTSFGGNLGRAVFQ
HSC2 氨基酸序列 (61) KGEPVNAADAAALEHDKAYDQQLKAGDNPYLKYNHADAEFQERLKEDTSFGGNLGRAVFQ
HSC11 氨基酸序列 (61) KGEPVNAADAAALEHDKAYDQQLKAGDNPYLKYNHADAEFQERLKEDTSFGGNLGRAVFQ
HSC3 氨基酸序列 (61) KGEPVNAATAAALEHDKAYDQQLKAGDNPYLKYNHADAEFQERLKEDTSFGGNLGRAVFQ
HSC4 氨基酸序列 (61) KGEPVNAADAAALEHDKAYDQQLKAGDNPYLKYNHADAEFQERLKEDTSFGGNLGRAVLQ
HSC6 氨基酸序列 (61) KGEPVNAADAAALEHDKAYDQQLKAGDNPYLKYNHADAEFQERLKEDTSFGGNLGRAVFQ
HSC7 氨基酸序列 (61) KGEPVNAVDAAALEHDKAYDQQLKAGDNPYLKYNHADAEFQERLKEDTSFGGNLGRAVFQ
HSC8 氨基酸序列 (61) KGEPVNAADAAALEHDKAYDQQLKAGDNPYLKYNHADAEFQERLKEDTSFGGNLGRAVFQ
HSC9 氨基酸序列 (61) KGEPVNAADAAALEHDKAYDQQLKAGDNPYLKYNHADAEFQERLKEDTSFGGNLGRAVFQ
HSC5 氨基酸序列 (61) KGEPVNAADAAALEHDKAYDQQLKAGDNPYLKYNHADAEFQERLKEDTSFGGNLGRAVFQ
HSC12 氨基酸序列 (61) KGEPVNAADAAALEHDKAYDQQLKAGDNPYLKYNHADAEFQERLKEDTSFGGNLGRAVFQ
HSC17 氨基酸序列 (61) KGEPVNAADAAALEHDKAYDQQLKAGDNPYLKYNHADAEFQERLKEDTSFGGNLGRAVFQ
HSC13 氨基酸序列 (61) KGEPVNAADAAALEHDKAYDQQLKAGDNPYLKYNHADAEFQERLKEDTSFGGNLGRAVFQ
HSC14 氨基酸序列 (61) KGEPVNAATAAALEHDKAYDQQLKAGDNPYLKYNHADAEFQERLKEDTSFGGNLGRAVFQ
HSC15 氨基酸序列 (61) KGEPVNAADAAALEHDKAYDQQLKAGDNPYLKYNHADAEFQERLKEDTSFGGNLGRAVFQ
HSC16 氨基酸序列 (61) KGEPVNAADAAALEHDKAYDQQLKAGDNPYLKYNHADAEFQERLKEDTSFGGNLGRAVFQ
AAV9 氨基酸序列 (61) KGEPVNAADAAALEHDKAYDQQLKAGDNPYLKYNHADAEFQERLKEDTSFGGNLGRAVFQ

图1

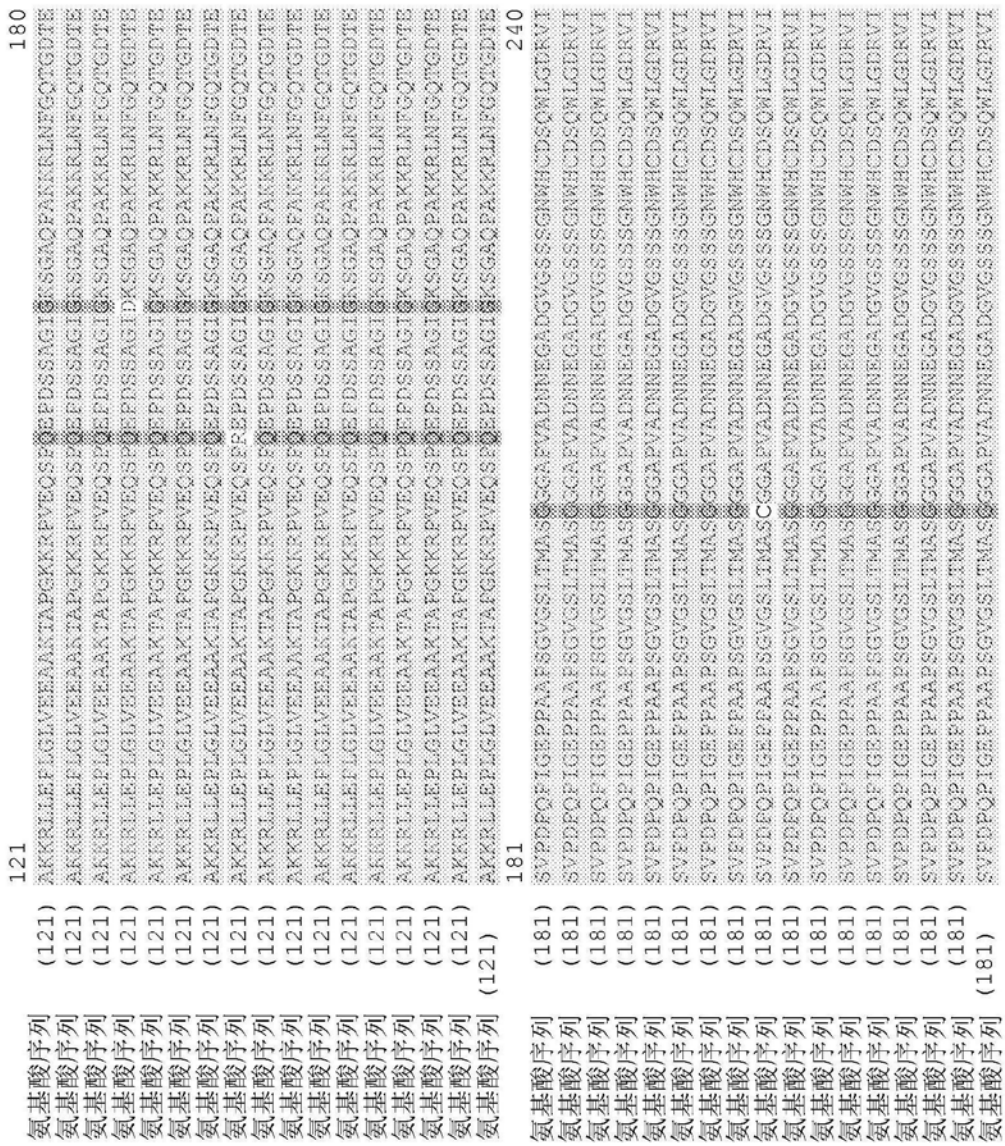

121
180
HSC1 氨基酸序列 (121) AKKRLLEPLGLVEEAAKTAPGKKRPVEQSPQEPDSSAGIGKSGAQPAKKRLNFGQTGDTE
HSC2 氨基酸序列 (121) AKKRLLEPLGLVEEAAKTAPGKKRPVEQSPQEPDSSAGIGKSGAQPAKKRLNFGQTGDTE
HSC11 氨基酸序列 (121) AKKRLLEPLGLVEEAAKTAPGKKRPVEQSPQEPDSSAGIGKSGAQPAKKRLNFGQTGDTE
HSC3 氨基酸序列 (121) AKKRLLEPLGLVEEAAKTAPGKKRPVEQSPQEPDSSAGIDKSGAQPAKKRLNFGQTGDTE
HSC4 氨基酸序列 (121) AKKRLLEPLGLVEEAAKTAPGKKRPVEQSPQEPDSSAGIGKSGAQPAKKRLNFGQTGDTE
HSC6 氨基酸序列 (121) AKKRLLEPLGLVEEAAKTAPGKKRPVEQSPQEPDSSAGIGKSGAQPAKKRLNFGQTGDTE
HSC7 氨基酸序列 (121) AKKRLLEPLGLVEEAAKTAPGKKRPVEQSPQEPDSSAGIGKSGAQPAKKRLNFGQTGDTE
HSC8 氨基酸序列 (121) AKKRLLEPLGLVEEAAKTAPGKKRPVEQSPREPDSSAGIGKSGAQPAKKRLNFGQTGDTE
HSC9 氨基酸序列 (121) AKKRLLEPLGLVEEAAKTAPGKKRPVEQSPQEPDSSAGIGKSGAQPAKKRLNFGQTGDTE
HSC5 氨基酸序列 (121) AKKRLLEPLGLVEEAAKTAPGKKRPVEQSPQEPDSSAGIGKSGAQPAKKRLNFGQTGDTE
HSC12 氨基酸序列 (121) AKKRLLEPLGLVEEAAKTAPGKKRPVEQSPQEPDSSAGIGKSGAQPAKKRLNFGQTGDTE
HSC17 氨基酸序列 (121) AKKRLLEPLGLVEEAAKTAPGKKRPVEQSPQEPDSSAGIGKSGAQPAKKRLNFGQTGDTE
HSC13 氨基酸序列 (121) AKKRLLEPLGLVEEAAKTAPGKKRPVEQSPQEPDSSAGIGKSGAQPAKKRLNFGQTGDTE
HSC14 氨基酸序列 (121) AKKRLLEPLGLVEEAAKTAPGKKRPVEQSPQEPDSSAGIGKSGAQPAKKRLNFGQTGDTE
HSC15 氨基酸序列 (121) AKKRLLEPLGLVEEAAKTAPGKKRPVEQSPQEPDSSAGIGKSGAQPAKKRLNFGQTGDTE
HSC16 氨基酸序列 (121) AKKRLLEPLGLVEEAAKTAPGKKRPVEQSPQEPDSSAGIGKSGAQPAKKRLNFGQTGDTE
AAV9 氨基酸序列 (121) AKKRLLEPLGLVEEAAKTAPGKKRPVEQSPQEPDSSAGIGKSGAQPAKKRLNFGQTGDTE
181
240
HSC1 氨基酸序列 (181) SVPDPQPIGEPPAAPSGVGSLTMASGGAPVADNNEGADGVGSSSGNWHCDSQWLGDRVI
HSC2 氨基酸序列 (181) SVPDPQPIGEPPAAPSGVGSLTMASGGAPVADNNEGADGVGSSSGNWHCDSQWLGDRVI
HSC11 氨基酸序列 (181) SVPDPQPIGEPPAAPSGVGSLTMASGGAPVADNNEGADGVGSSSGNWHCDSQWLGDRVI
HSC3 氨基酸序列 (181) SVPDPQPIGEPPAAPSGVGSLTMASGGAPVADNNEGAIGVGSSSGNWHCDSQWLGDRVI
HSC4 氨基酸序列 (181) SVPDPQPIGEPPAAPSGVGSLTMASGGAPVADNNEGADGVGSSSGNWHCDSQWLGDRVI
HSC6 氨基酸序列 (181) SVPDPQPIGEPPAAPSGVGSLTMASGGAPVADNNEGADGVGSSSGNWHCDSQWLGDRVI
HSC7 氨基酸序列 (181) SVPDPQPIGEPPAAPSGVGSLTMASGGAPVADNNEGADGVGSSSGNWHCDSQWLGDRVI
HSC8 氨基酸序列 (181) SVPDPQPIGEPPAAPSGVGSLTMASGGAPVADNNEGADGVGSSSGNWHCDSQWLGDRVI
HSC9 氨基酸序列 (181) SVPDPQPIGEPPAAPSGVGSLTMASCGGAPVADNNEGADGVGSSSGNWHCDSQWLGDRVI
HSC5 氨基酸序列 (181) SVPDPQPIGEPPAAPSGVGSLTMASGGAPVADNNEGADGVGSSSGNWHCDSQWLGDRVI
HSC12 氨基酸序列 (181) SVPDPQPIGEPPNAPSGVGSLTMASGGAPVADNNEGADGVGSSSGNWHCDSQWLGDRVI
HSC17 氨基酸序列 (181) SVPDPQPIGEPPAAPSGVGSLTMASGGAPVADNNEGAIGVGSSSGNWHCDSQWLGDRVI
HSC13 氨基酸序列 (181) SVPDPQPIGEPPAAPSGVGSLTMASGGAPVADNNEGAIGVGSSSGNWHCDSQWLGDRVI
HSC14 氨基酸序列 (181) SVPDPQPIGEPPAAPSGVGSLTMASGGAPVADNNEGAIGVGSSSGNWHCDSQWLGDRVI
HSC15 氨基酸序列 (181) SVPDPQPIGEPPAAPSGVGSLTMASGGAPVADNNEGADGVGSSSGNWHCDSQWLGDRVI
HSC16 氨基酸序列 (181) SVPDPQPIGEPPAAPSGVGSLTMASGGAPVADNNEGADGVGSSSGNWHCDSQWLGDRVI
AAV9 氨基酸序列 (181) SVPDPQPIGEPPAAPSGVGSLTMASGGAPVADNNEGADGVGSSSGNWHCDSQWLGDRVI

图1(续)

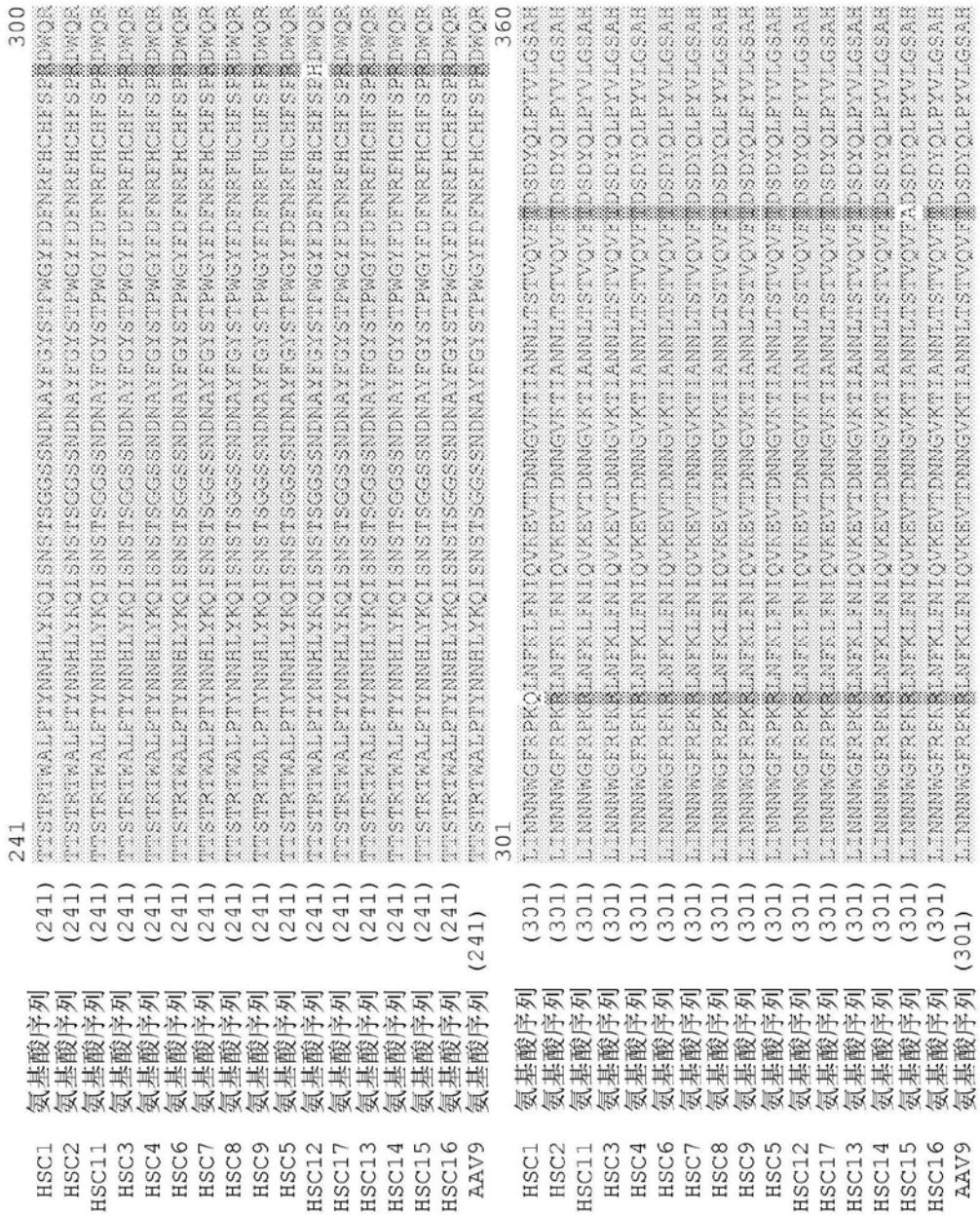

241
300
HSC1 氨基酸序列 (241)
HSC2 氨基酸序列 (241)
HSC11 氨基酸序列 (241)
HSC3 氨基酸序列 (241)
HSC4 氨基酸序列 (241)
HSC6 氨基酸序列 (241)
HSC7 氨基酸序列 (241)
HSC8 氨基酸序列 (241)
HSC9 氨基酸序列 (241)
HSC5 氨基酸序列 (241)
HSC12 氨基酸序列 (241)
HSC17 氨基酸序列 (241)
HSC13 氨基酸序列 (241)
HSC14 氨基酸序列 (241)
HSC15 氨基酸序列 (241)
HSC16 氨基酸序列 (241)
AAV9 氨基酸序列 (241)
301
360
HSC1 氨基酸序列 (301)
HSC2 氨基酸序列 (301)
HSC11 氨基酸序列 (301)
HSC3 氨基酸序列 (301)
HSC4 氨基酸序列 (301)
HSC6 氨基酸序列 (301)
HSC7 氨基酸序列 (301)
HSC8 氨基酸序列 (301)
HSC9 氨基酸序列 (301)
HSC5 氨基酸序列 (301)
HSC12 氨基酸序列 (301)
HSC17 氨基酸序列 (301)
HSC13 氨基酸序列 (301)
HSC14 氨基酸序列 (301)
HSC15 氨基酸序列 (301)
HSC16 氨基酸序列 (301)
AAV9 氨基酸序列 (301)

图1(续)

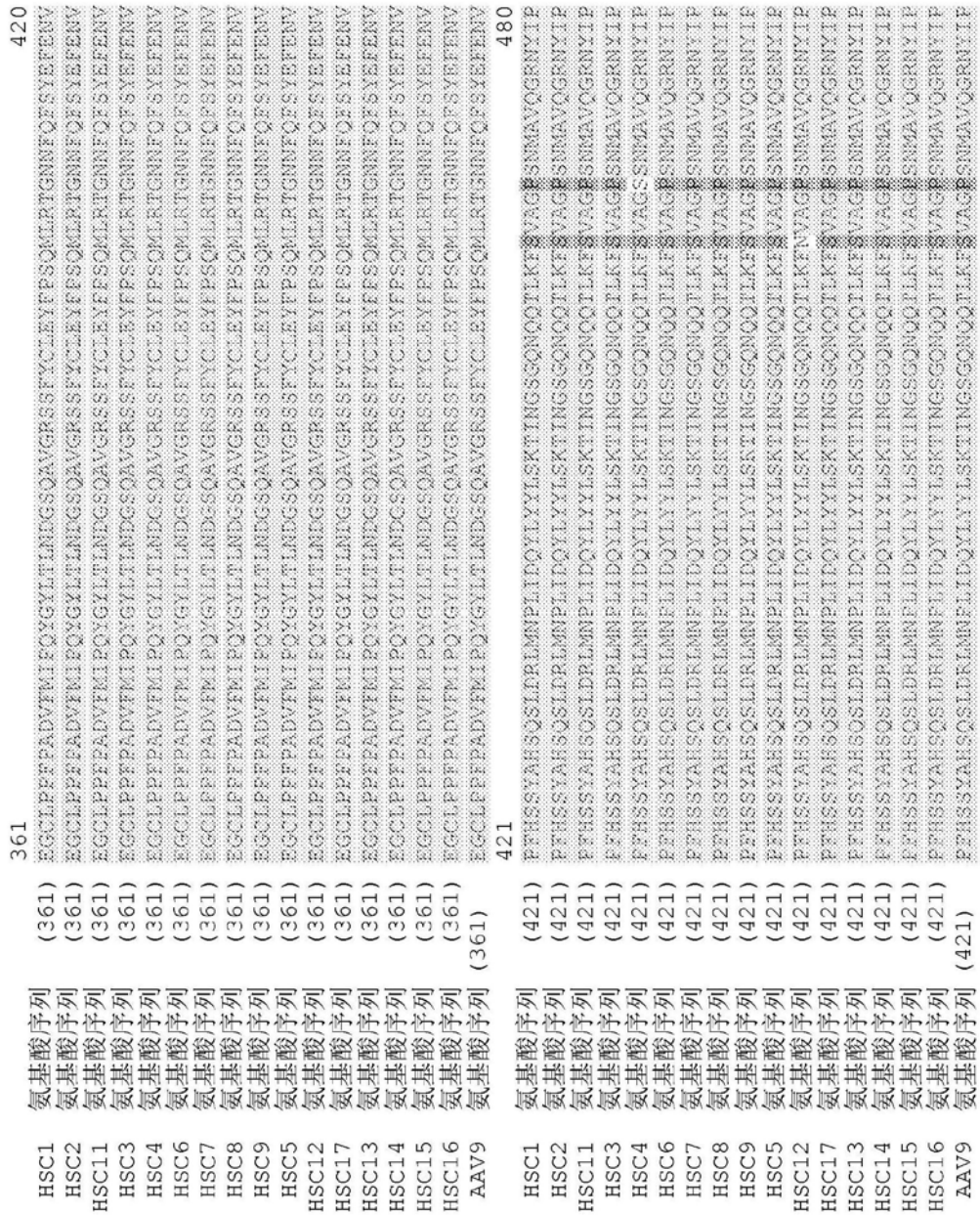

HSC1 氨基酸序列 (361) EGCLPFPADVFMIPQYGYLTLNDGSQAVGRSSFYCLEYFPSQMLRTGNNFQFSYEFENV
HSC2 氨基酸序列 (361) EGCLPFPADVFMIPQYGYLTLNDGSQAVGRSSFYCLEYFPSQMLRTGNNFQFSYEFENV
HSC11 氨基酸序列 (361) EGCLPFPADVFMIPQYGYLTLNDGSQAVGRSSFYCLEYFPSQMLRTGNNFQFSYEFENV
HSC3 氨基酸序列 (361) EGCLPFPADVFMIPQYGYLTLNDGSQAVGRSSFYCLEYFPSQMLRTGNNFQFSYEFENV
HSC4 氨基酸序列 (361) EGCLPFPADVFMIPQYGYLTLNDGSQAVGRSSFYCLEYFPSQMLRTGNNFQFSYEFENV
HSC6 氨基酸序列 (361) EGCLPFPADVFMIPQYGYLTLNDGSQAVGRSSFYCLEYFPSQMLRTGNNFQFSYEFENV
HSC7 氨基酸序列 (361) EGCLPFPADVFMIPQYGYLTLNDGSQAVGRSSFYCLEYFPSQMLRTGNNFQFSYEFENV
HSC8 氨基酸序列 (361) EGCLPFPADVFMIPQYGYLTLNDGSQAVGRSSFYCLEYFPSQMLRTGNNFQFSYEFENV
HSC9 氨基酸序列 (361) EGCLPFPADVFMIPQYGYLTLNDGSQAVGRSSFYCLEYFPSQMLRTGNNFQFSYEFENV
HSC5 氨基酸序列 (361) EGCLPFPADVFMIPQYGYLTLNDGSQAVGRSSFYCLEYFPSQMLRTGNNFQFSYEFENV
HSC12 氨基酸序列 (361) EGCLPFPADVFMIPQYGYLTLNDGSQAVGRSSFYCLEYFPSQMLRTGNNFQFSYEFENV
HSC17 氨基酸序列 (361) EGCLPFPADVFMIPQYGYLTLNDGSQAVGRSSFYCLEYFPSQMLRTGNNFQFSYEFENV
HSC13 氨基酸序列 (361) EGCLPFPADVFMIPQYGYLTLNDGSQAVGRSSFYCLEYFPSQMLRTGNNFQFSYEFENV
HSC14 氨基酸序列 (361) EGCLPFPADVFMIPQYGYLTLNDGSQAVGRSSFYCLEYFPSQMLRTGNNFQFSYEFENV
HSC15 氨基酸序列 (361) EGCLPFPADVFMIPQYGYLTLNDGSQAVGRSSFYCLEYFPSQMLRTGNNFQFSYEFENV
HSC16 氨基酸序列 (361) EGCLPFPADVFMIPQYGYLTLNDGSQAVGRSSFYCLEYFPSQMLRTGNNFQFSYEFENV
AAV9 氨基酸序列 (361) EGCLPFPADVFMIPQYGYLTLNDGSQAVGRSSFYCLEYFPSQMLRTGNNFQFSYEFENV
361
420
HSC1 氨基酸序列 (421) FTHSSYAHSQSLDRLMNPLIDQYLYYLSKTINGSGQNQQTLKF VAGPSMMAVQGRNYIP
HSC2 氨基酸序列 (421) FTHSSYAHSQSLDRLMNPLIDQYLYYLSKTINGSGQNQQTLKF VAGPSMMAVQGRNYIP
HSC11 氨基酸序列 (421) FTHSSYAHSQSLDRLMNPLIDQYLYYLSKTINGSGQNQQTLKF VAGPSMMAVQGRNYIP
HSC3 氨基酸序列 (421) FTHSSYAHSQSLDRLMNPLIDQYLYYLSKTINGSGQNQQTLKF VAGPSMMAVQGRNYIP
HSC4 氨基酸序列 (421) FTHSSYAHSQSLDRLMNPLIDQYLYYLSKTINGSGQNQQTLKF VAGSSMMAVQGRNYIP
HSC6 氨基酸序列 (421) FTHSSYAHSQSLDRLMNPLIDQYLYYLSKTINGSGQNQQTLKF VAGPSMMAVQGRNYIP
HSC7 氨基酸序列 (421) FTHSSYAHSQSLDRLMNPLIDQYLYYLSKTINGSGQNQQTLKF VAGPSMMAVQGRNYIP
HSC8 氨基酸序列 (421) FTHSSYAHSQSLDRLMNPLIDQYLYYLSKTINGSGQNQQTLKF VAGPSMMAVQGRNYIP
HSC9 氨基酸序列 (421) FTHSSYAHSQSLDRLMNPLIDQYLYYLSKTINGSGQNQQTLKF VAGPSMMAVQGRNYIP
HSC5 氨基酸序列 (421) FTHSSYAHSQSLDRLMNPLIDQYLYYLSKTINGSGQNQQTLKF VAGPSMMAVQGRNYIP
HSC12 氨基酸序列 (421) FTHSSYAHSQSLDRLMNPLIDQYLYYLSKTINGSGQNQQTLKFNVAGPSMMAVQGRNYIP
HSC17 氨基酸序列 (421) FTHSSYAHSQSLDRLMNPLIDQYLYYLSKTINGSGQNQQTLKF VAGPSMMAVQGRNYIP
HSC13 氨基酸序列 (421) FTHSSYAHSQSLDRLMNPLIDQYLYYLSKTINGSGQNQQTLKF VAGPSMMAVQGRNYIP
HSC14 氨基酸序列 (421) FTHSSYAHSQSLDRLMNPLIDQYLYYLSKTINGSGQNQQTLKF VAGPSMMAVQGRNYIP
HSC15 氨基酸序列 (421) FTHSSYAHSQSLDRLMNPLIDQYLYYLSKTINGSGQNQQTLKF VAGPSMMAVQGRNYIP
HSC16 氨基酸序列 (421) FTHSSYAHSQSLDRLMNPLIDQYLYYLSKTINGSGQNQQTLKF VAGPSMMAVQGRNYIP
AAV9 氨基酸序列 (421) FTHSSYAHSQSLDRLMNPLIDQYLYYLSKTINGSGQNQQTLKF VAGPSMMAVQGRNYIP
421
480

图1(续)

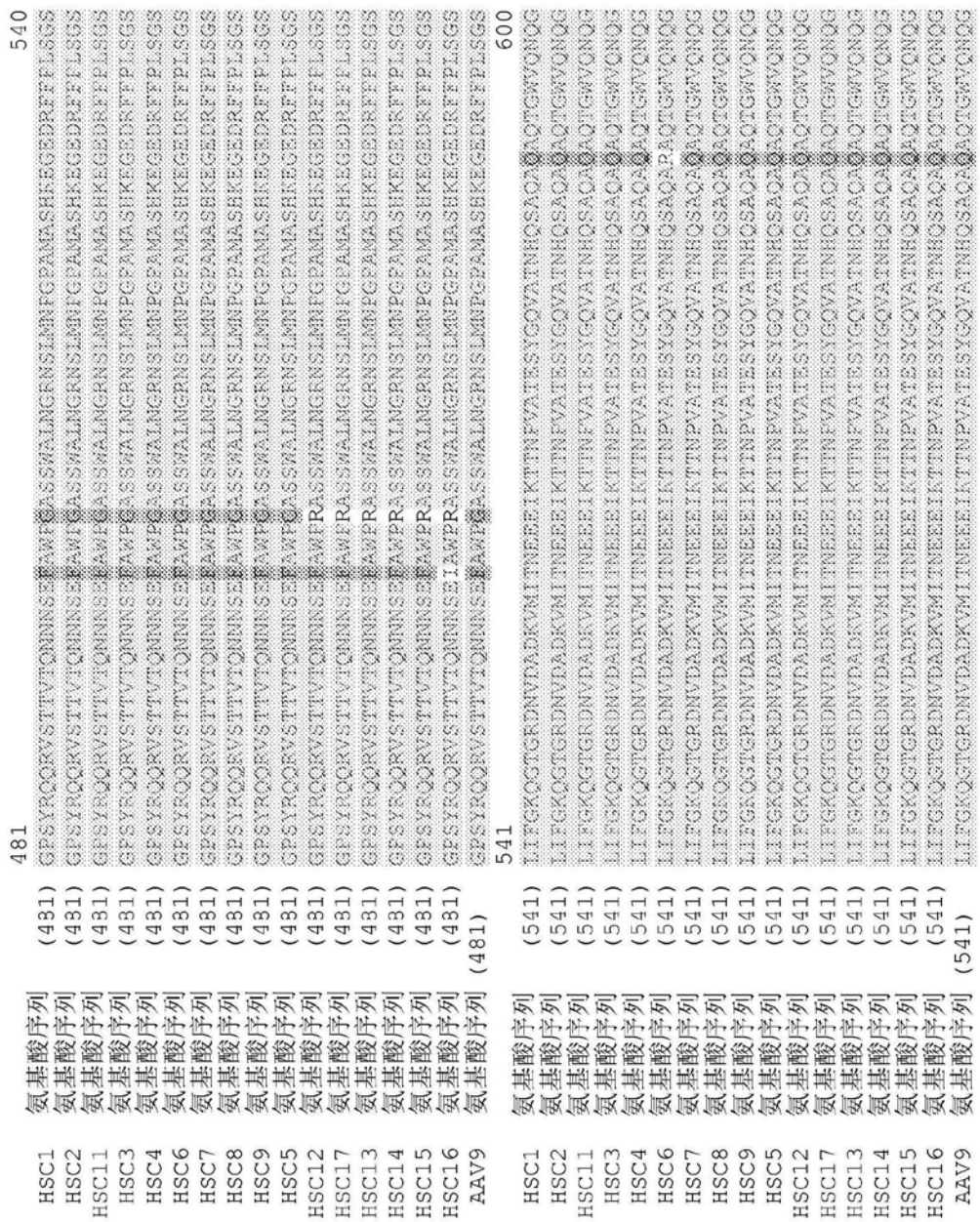

481
540
541
600
HSC1 氨基酸序列 (431) GPSYRQQRVSTTVTQNNNSEFAWPGASSWALNGRNSLMNPGPAMASHKEGEDRFFPLSGS
HSC2 氨基酸序列 (431) GPSYRQQRVSTTVTQNNNSEFAWIGASSWALNGRNSLMNIGPAMASHKEGEDRFTTLSGS
HSC11 氨基酸序列 (431) GPSYRQQRVSTTVTQNNNSEFAWPGASSWALNGRNSLMNPGPAMASHKEGEDRFFPLSGS
HSC3 氨基酸序列 (431) GFSYRQQRVSTTVTQNNNSEFAWPGASSWALNGRNSLMNPGPAMASHKEGEDRFFPLSGS
HSC4 氨基酸序列 (431) GPSYRQQRVSTTVTQNNNSEFAWPGASSWALNGRNSLMNPGPAMASHKEGEDRFFPLSGS
HSC6 氨基酸序列 (431) GPSYRQQRVSTTVTQNNNSEFAWPGASSWALNGRNSLMNPGPAMASHKEGEDRFFPLSGS
HSC7 氨基酸序列 (431) GPSYRQQRVSTTVTQNNNSEFAWPGASSWALNGRNSLMNPGPAMASHKEGEDRFFPLSGS
HSC8 氨基酸序列 (431) GPSYRQQRVSTTVTQNNNSEFAWPGASSWALNGRNSLMNPGPAMASHKEGEDRFFPLSGS
HSC9 氨基酸序列 (431) GPSYRQQRVSTTVTQNNNSEFAWPGASSWALNGRNSLMNPGPAMASHKEGEDRFFPLSGS
HSC5 氨基酸序列 (431) GPSYRQQRVSTTVTQNNNSEFAWPGASSWALNGRNSLMNPGPAMASHKEGEDRFFPLSGS
HSC12 氨基酸序列 (431) GPSYRQQRVSTTVTQNNNSEFAWPRASSWALNGRNSLMNPGPAMASHKEGEDRFFPLSGS
HSC17 氨基酸序列 (431) GPSYRQQRVSTTVTQNNNSEFAWPRASSWALNGRNSLMNPGPAMASHKEGEDRFFPLSGS
HSC13 氨基酸序列 (431) GFSYRQQRVSTTVTQNNNSEFAMPRASSWALNGRNSLMNPGPAMASHKEGEDRFFPLSGS
HSC14 氨基酸序列 (431) GFSYRQQRVSTTVTQNNNSEFAMPRASSWALNGRNSLMNPGPAMASHKEGEDRFFPLSGS
HSC15 氨基酸序列 (431) GFSYRQQRVSTTVTQNNNSEFAMPRASSWALNGRNSLMNPGPAMASHKEGEDRFFPLSGS
HSC16 氨基酸序列 (431) GFSYRQQRVSTTVTQNNNSEIAMPRASSWALNGPNSLMNPGPAMASHKEGEDRFFPLSGS
AAV9 氨基酸序列 (481) GPSYRQQRVSTTVTQNNNSEFAMPGASSWALNGPNSLMNPGPAMASHKEGEDRFFPLSGS
HSC1 氨基酸序列 (541) LIFGKQGTGRDNVDADKVMITNEEEIKTTNPVATESYGQVATNHQSAQAQAQTGWVQNQG
HSC2 氨基酸序列 (541) LIFGKQGTGRDNVDADKVMITNEEEIKTTNPVATESYGQVATNHQSAQAQAQTGWVQNQG
HSC11 氨基酸序列 (541) LIFGKQGTGRDNVDADKVMITNEEEIKTTNPVATESYGQVATNHQSAQAQAQTGWVQNQG
HSC3 氨基酸序列 (541) LIFGKQGTGRDNVDADKVMITNEEEIKTTNPVATESYGQVATNHQSAQAQAQTGWVQNQG
HSC4 氨基酸序列 (541) LIFGKQGTGRDNVDADKVMITNEEEIKTTNPVATESYGQVATNHQSAQAQAQTGWVQNQG
HSC6 氨基酸序列 (541) LIFGKQGTGRDNVDADKVMITNEEEIKTTNPVATESYGQVATNHQSAPAQTGWVQNQG
HSC7 氨基酸序列 (541) LIFGKQGTGRDNVDADKVMITNEEEIKTTNPVATESYGQVATNHQSAQAQAQTGWVQNQG
HSC8 氨基酸序列 (541) LIFGKQGTGRDNVDADKVMITNEEEIKTTNPVATESYGQVATNHQSAQAQAQTGWVQNQG
HSC9 氨基酸序列 (541) LIFGKQGTGRDNVDADKVMITNEEEIKTTNPVATESYGQVATNHQSAQAQAQTGWVQNQG
HSC5 氨基酸序列 (541) LIFGKQGTGRDNVDADKVMITNEEEIKTTNPVATESYGQVATNHQSAQAQAQTGWVQNQG
HSC12 氨基酸序列 (541) LIFGKQGTGRDNVDADKVMITNEEEIKTTNPVATESYGQVATNHQSAQAQAQTGWVQNQG
HSC17 氨基酸序列 (541) LIFGKQGTGRDNVDADKVMITNEEEIKTTNPVATESYGQVATNHQSAQAQAQTGWVQNQG
HSC13 氨基酸序列 (541) LIFGKQGTGRDNVDADKVMITNEEEIKTTNPVATESYGQVATNHQSAQAQAQTGWVQNQG
HSC14 氨基酸序列 (541) LIFGKQGTGRDNVDADKVMITNEEEIKTTNPVATESYGQVATNHQSAQAQAQTGWVQNQG
HSC15 氨基酸序列 (541) LIFGKQGTGRDNVDADKVMITNEEEIKTTNPVATESYGQVATNHQSAQAQAQTGWVQNQG
HSC16 氨基酸序列 (541) LIFGKQGTGRDNVDADKVMITNEEEIKTTNPVATESYGQVATNHQSAQAQAQTGWVQNQG
AAV9 氨基酸序列 (541) LIFGKQGTGRDNVDADKVMITNEEEIKTTNPVATESYGQVATNHQSAQAQAQTGWVQNQG

图1(续)

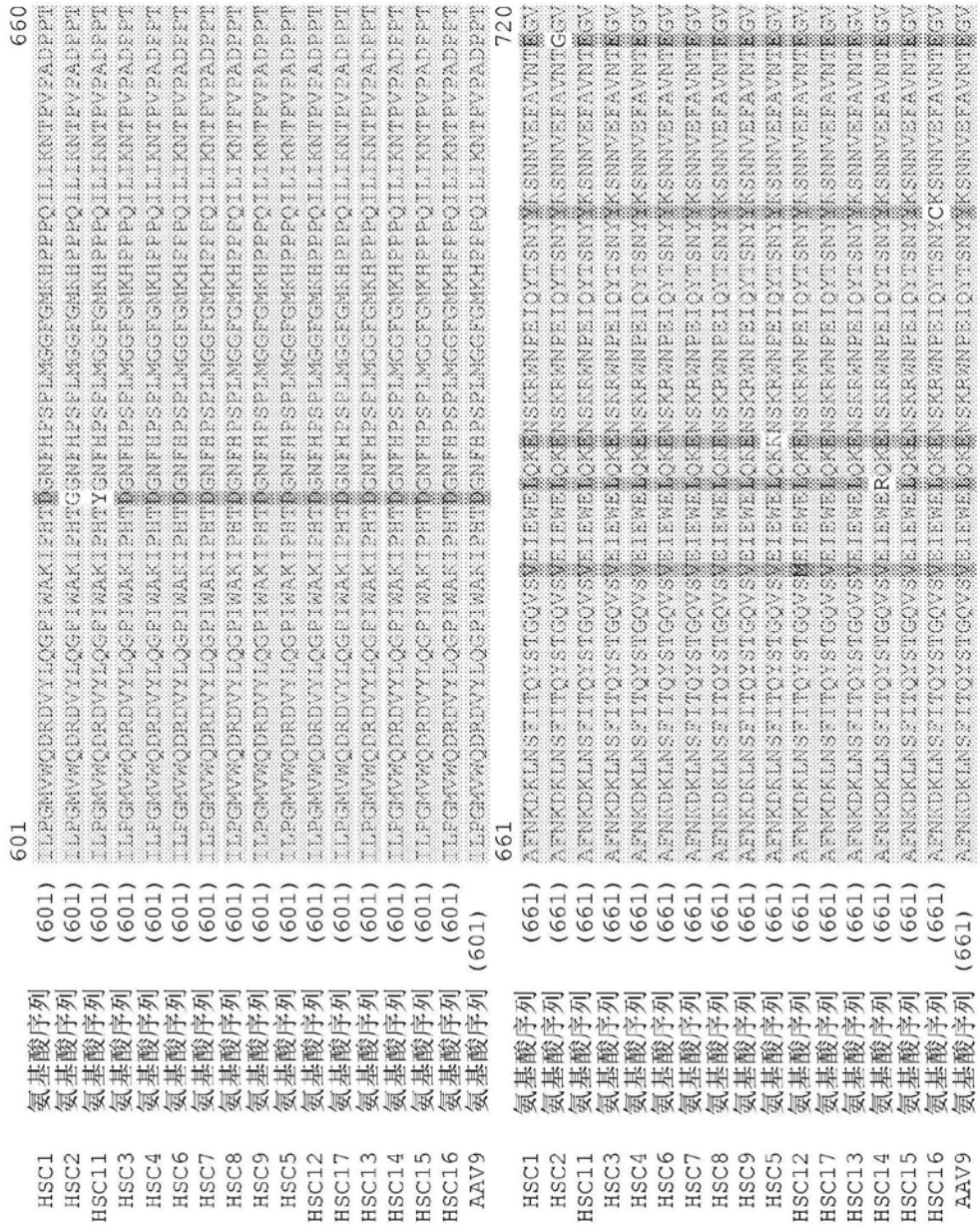

601
660
661
720
HSC1 氨基酸序列 (601) ILPGMVWQDRDVYYLQGPIWAKIPHTDGNFHPSPLMGGFGMKHPPPQILIKNTPVPADPPT
HSC2 氨基酸序列 (601) LPGMVWQDRDVYYLQGPIWAKIPHTGGNFHPSPLMGGFGMKHPPPQILIKNTPVPADPPT
HSC11 氨基酸序列 (601) ILPGMVWQDRDVYLQGPIWAKIPHTYGNFHPSPLMGGFGMKHPPPQILIKNTPVPADPPT
HSC3 氨基酸序列 (601) ILPGMVWQDRDVYLQGPIWAKIPHTDGNFHPSPLMGGFGMKHPPPQILIKNTPVPADPPT
HSC4 氨基酸序列 (601) ILPGMVWQDRDVYLQGPIWAKIPHTDGNFHPSPLMGGFGMKHPPPQILIKNTPVPADPPT
HSC6 氨基酸序列 (601) ILPGMVWQDRDVYLQGPIWAKIPHTDGNFHPSPLMGGFGMKHPPPQILIKNTPVPADPPT
HSC7 氨基酸序列 (601) ILPGMVWQDRDVYLQGPIWAKIPHTDGNFHPSPLMGGFGMKHPPPQILIKNTPVPADPPT
HSC8 氨基酸序列 (601) ILPGMVWQDRDVYLQGPIWAKIPHTDGNFHPSPLMGGFGMKHPPPQILIKNTPVPADPPT
HSC9 氨基酸序列 (601) ILPGMVWQDRDVYLQGPIWAKIPHTDGNFHPSPLMGGFGMKHPPPQILIKNTPVPADPPT
HSC5 氨基酸序列 (601) ILPGMVWQDRDVYLQGPIWAKIPHTDGNFHPSPLMGGFGMKHPPPQILIKNTPVPADPPT
HSC12 氨基酸序列 (601) ILPGMVWQDRDVYLQGPIWAKIPHTDGNFHPSPLMGGFGMKHPPPQILIKNTPVPADPPT
HSC17 氨基酸序列 (601) ILPGMVWQDRDVYLQGPIWAKIPHTDGNFHPSPLMGGFGMKHPPPQILIKNTPVPADPPT
HSC13 氨基酸序列 (601) ILPGMVWQDRDVYLQGPIWAKIPHTDGNFHPSPLMGGFGMKHPPPQILIKNTPVPADPPT
HSC14 氨基酸序列 (601) ILPGMVWQDRDVYLQGPIWAKIPHTDGNFHPSPLMGGFGMKHPPPQILIKNTPVPADPPT
HSC15 氨基酸序列 (601) ILPGMVWQDRDVYLQGPIWAKIPHTDGNFHPSPLMGGFGMKHPPPQILIKNTPVPADPPT
HSC16 氨基酸序列 (601) ILPGMVWQDRDVYLQGPIWAKIPHTDGNFHPSPLMGGFGMKHPPPQILIKNTPVPADPPT
AAV9 氨基酸序列 (601) ILPGMVWQDRDVYLQGPIWAKIPHTDGNFHPSPLMGGFGMKHPPPQILIKNTPVPADPPT
HSC1 氨基酸序列 (661) AFNKDKLNSFITQYSTGQVSVEIEWELQKENSKRWNPEIQYTSNYKKSNNVEFAVNTEGV
HSC2 氨基酸序列 (661) AFNKDKLNSFITQYSTGQVSVEIEWELQKENSKRWNPEIQYTSNYKKSNNVEFAVNTGGV
HSC11 氨基酸序列 (661) AFNKDKLNSFITQYSTGQVSVEIEWELQKENSKRWNPEIQYTSNYKKSNNVEFAVNTEGV
HSC3 氨基酸序列 (661) AFNKDKLNSFITQYSTGQVSVEIEWELQKENSKRWNPEIQYTSNYKKSNNVEFAVNTEGV
HSC4 氨基酸序列 (661) AFNKDKLNSFITQYSTGQVSVEIEWELQKENSKRWNPEIQYTSNYKKSNNVEFAVNTEGV
HSC6 氨基酸序列 (661) AFNKDKLNSFITQYSTGQVSVEIEWELQKENSKRWNPEIQYTSNYKKSNNVEFAVNTEGV
HSC7 氨基酸序列 (661) AFNKDKLNSFITQYSTGQVSVEIEWELQKENSKRWNPEIQYTSNYKKSNNVEFAVNTEGV
HSC8 氨基酸序列 (661) AFNKDKLNSFITQYSTGQVSVEIEWELQKENSKRWNPEIQYTSNYKKSNNVEFAVNTEGV
HSC9 氨基酸序列 (661) AFNKDKLNSFITQYSTGQVSVEIEWELQKENSKRWNPEIQYTSNYKKSNNVEFAVNTEGV
HSC5 氨基酸序列 (661) AFNKDKLNSFITQYSTGQVSVEIEWELQKENSKRWNPEIQYTSNYKKSNNVEFAVNTEGV
HSC12 氨基酸序列 (661) AFNKDKLNSFITQYSTGQVSMEIEWELQKENSKRWNPEIQYTSNYMKSNNVEFAVNTEGV
HSC17 氨基酸序列 (661) AFNKDKLNSFITQYSTGQVSVEIEWELQKENSKRWNPEIQYTSNYMKSNNVEFAVNTEGV
HSC13 氨基酸序列 (661) AFNKDKLNSFITQYSTGQVSVEIEWELQKENSKRWNPEIQYTSNYMKSNNVEFAVNTEGV
HSC14 氨基酸序列 (661) AFNKDKLNSFITQYSTGQVSVEIEWERQKENSKRWNPEIQYTSNYKKSNNVEFAVNTEGV
HSC15 氨基酸序列 (661) AFNKDKLNSFITQYSTGQVSVEIEWELQKENSKRWNPEIQYTSNYKKSNNVEFAVNTEGV
HSC16 氨基酸序列 (661) AFNKDKLNSFITQYSTGQVSVEIEWELQKENSKRWNPEIQYTSNYCKSNNVEFAVNTEGV
AAV9 氨基酸序列 (661) AFNKDKLNSFITQYSTGQVSVEIEWELQKENSKRWNPEIQYTSNYKKSNNVEFAVNTEGV

图1(续)

			721	736
HSC1	氨基酸序列	(721)	YSEPRPIGTRYLTRNL	
HSC2	氨基酸序列	(721)	YSEPRPIGTRYLTRNL	
HSC11	氨基酸序列	(721)	YSEPRPIGTRYLTRNL	
HSC3	氨基酸序列	(721)	YSEPRPIGTRYLTRNL	
HSC4	氨基酸序列	(721)	YSEPRPIGTRYLTRNL	
HSC6	氨基酸序列	(721)	YSEPRPIGTRYLTRNL	
HSC7	氨基酸序列	(721)	YSEPRPIGTRYLTRNL	
HSC8	氨基酸序列	(721)	YSEPRPIGTRYLTRNL	
HSC9	氨基酸序列	(721)	YSEPRPIGTRYLTRNL	
HSC5	氨基酸序列	(721)	YSEPRPIGTRYLTRNL	
HSC12	氨基酸序列	(721)	YSEPRPIGTRYLTRNL	
HSC17	氨基酸序列	(721)	YSEPRPIGTRYLTRNL	
HSC13	氨基酸序列	(721)	YSEPRPIGTRYLTRNL	
HSC14	氨基酸序列	(721)	YSEPRPIGTRYLTRNL	
HSC15	氨基酸序列	(721)	YSEPRPIGTRYLTRNL	
HSC16	氨基酸序列	(721)	YSEPRPIGTRYLTRNL	
AAV9	氨基酸序列	(721)	YSEPRPIGTRYLTRNL	

图1(续)

衣壳	碱基变化	AAV 9 hu.分离株氨基酸	克隆体新氨基酸	氨基酸差异	氨基酸位置（HVR =高变区）
HSC1 SEQ ID NO: 2	G变为A	GCT -丙氨酸	ACT -苏氨酸	丙氨酸是非极性并且中性的；苏氨酸是极性并且中性的	4个碱基进入衣壳中；第2个氨基酸；不在HVR，VP1中
HSC1 SEQ ID NO: 2	G变为A	CGA -精氨酸	CAA -谷氨酰胺	精氨酸是极性并且强碱性的；谷氨酰胺是极性并且中性的	935个碱基进入衣壳中；第312个氨基酸；不在HVR，VP3中
HSC2 SEQ ID NO: 3	A变为G	GGA -甘氨酸	GGG -甘氨酸	无氨基酸差异	1515个碱基进入衣壳中；第505个氨基酸；HVR 7之后的第1个氨基酸，VP3
HSC2 SEQ ID NO: 3	A变为G	GAC -天冬氨酸	GGC -甘氨酸	天冬氨酸是极性并且酸性的；甘氨酸是非极性并且中性的	1877个碱基进入衣壳中；第626个氨基酸；不在HVR，VP3中
HSC2 SEQ ID NO: 3	A变为G	GAA -谷氨酸	GGA -甘氨酸	谷氨酸是极性并且酸性的；甘氨酸是非极性并且中性的	2153个碱基进入衣壳中；第718个氨基酸；在HVR 12最后一个氨基酸中，VP3
HSC3 SEQ ID NO: 5	G变为A	GGC- 甘氨酸	GAC -天冬氨酸	甘氨酸是非极性并且中性的；天冬氨酸是极性并且酸性的	479个碱基进入衣壳中；第160个氨基酸；不在HVR，VP1中
HSC3 SEQ ID NO: 5	A变为G	TCA -丝氨酸	TCG -丝氨酸	无氨基酸差异	1044个碱基进入衣壳中，第348个氨基酸；不在HVR VP3中
HSC4 SEQ ID NO: 6	T变为C	TTC -苯丙氨酸	CTC -亮氨酸	苯丙氨酸是非极性并且中性的；亮氨酸是非极性并且中性的	355个碱基进入衣壳中，第119个氨基酸；不在HVR；VP1中
HSC4 SEQ ID NO: 6	C变为T	CCC -脯氨酸	TCC -丝氨酸	脯氨酸是非极性并且中性的；丝氨酸是极性并且中性的	1402个碱基进入衣壳中；第468个氨基酸；在HVR 5，VP3中
HSC4 SEQ ID NO: 6	G变为A	CCG -脯氨酸	CCA -脯氨酸	无氨基酸差异	1809个碱基进入衣壳中；第603个氨基酸；不在HVR，VP3中

图2

衣壳	碱基变化	AAV 9 hu.分离株氨基酸	克隆体新氨基酸	氨基酸差异	氨基酸位置（HVR =高变区）
HSC6 SEQ ID NO: 8	A变为G	CAG -谷氨酰胺	CGG -精氨酸	谷氨酰胺是极性并且中性的；精氨酸是极性并且强碱性的	1769个碱基进入衣壳中；第590个氨基酸；在HVR 10中；VP3
HSC5 SEQ ID NO: 11	A变为G	AAG -赖氨酸	AGG -精氨酸	赖氨酸是极性并且碱性的；精氨酸是极性并且强碱性的	230个碱基进入衣壳中；第77个氨基酸；不在HVR，VP1中
HSC5 SEQ ID NO: 11	C变为T	CCC -脯氨酸	CCT -脯氨酸	无氨基酸差异	1404个碱基进入衣壳中；第468个氨基酸；在HVR 5，VP3中
HSC5 SEQ ID NO: 11	G变为A	GAA -谷氨酸	AAA -赖氨酸	谷氨酸是极性并且酸性的；赖氨酸是极性并且碱性的	2068个碱基进入衣壳中；第690个氨基酸；不在HVR，VP3中
HSC5 SEQ ID NO: 11	T变为C	AAT -天冬酰胺	AAC -天冬酰胺	无氨基酸差异	2148个碱基进入衣壳中；第716个氨基酸；在HVR 13，VP3中
HSC11 SEQ ID NO: 5	G变为A	GTC -缬氨酸	ATC -异亮氨酸	异亮氨酸是非极性并且中性的；缬氨酸是非极性并且中性的	193个碱基进入衣壳中；第65个氨基酸；不在HVR中；VP1
HSC11 SEQ ID NO: 4	G变为T	GAC -天冬氨酸	TAC -酪氨酸	天冬氨酸是极性并且酸性的；酪氨酸是极性并且中性的	1876个碱基进入衣壳中；第626个氨基酸；不在HVR中；VP3
HSC7 SEQ ID NO: 8	C变为T	GCA -丙氨酸	GTA -缬氨酸	丙氨酸是非极性并且中性的；缬氨酸是非极性并且中性的	203个碱基进入衣壳中；第68个氨基酸；不在HVR中；VP1
HSC7 SEQ ID NO: 8	C变为T	TTC -苯丙氨酸	TTT -苯丙氨酸	无氨基酸差异	924个碱基进入衣壳中；第308个氨基酸；不在HVR中；VP3
HSC8 SEQ ID NO: 9	A变为G	CAG -谷氨酰胺	CGG -精氨酸	谷氨酰胺是极性并且中性的；精氨酸是极性并且强碱性的	452个碱基进入衣壳中；第151个氨基酸；在HVR 1中；VP1

图2(续)

衣壳	碱基变化	AAV 9 hu.分离株氨基酸	克隆体新氨基酸	氨基酸差异	氨基酸位置（HVR =高变区）
HSC8 SEQ ID NO: 9	G变为A	GAG -谷氨酰胺酸	GAA -谷氨酰胺酸	无氨基酸差异	2058个碱基进入衣壳中；第686个氨基酸；不在HVR中；VP3
HSC9 SEQ ID NO: 10	T变为C	TTT -苯丙氨酸	TTC -苯丙氨酸	无氨基酸差异	357个碱基进入衣壳中；第119个氨基酸；不在HVR中；VP1
HSC9 SEQ ID NO: 10	T变为G	TGT- 半胱氨酸	GGT -甘氨酸	半胱氨酸是极性并且中性的；甘氨酸是非极性并且中性的	616个碱基进入衣壳中；第206个氨基酸；不在HVR中；VP3极开始处
HSC12 SEQ ID NO: 12	A变为T	GCA -丙氨酸	GCT -丙氨酸	无氨基酸差异	579个碱基进入衣壳中；第193个氨基酸进入衣壳中，接近HVR 2的VP2末端中
HSC12 SEQ ID NO: 12	G变为A	CGT -精氨酸	CAT -组氨酸	精氨酸是极性并且强碱性的；组氨酸是极性并且弱碱性的	887个碱基进入衣壳中；第296个氨基酸进入衣壳中；在VP3中，不在HVR中
HSC12 SEQ ID NO: 12	G变为A	AGT -丝氨酸	AAT -天冬酰胺	丝氨酸是极性并且中性的；天冬酰胺是极性并且中性的	1391个碱基进入衣壳中；第464个氨基酸进入衣壳中；在HVR 5中的VP3中
HSC12 SEQ ID NO: 12	T变为C	GCT -丙氨酸	GCC -丙氨酸	无氨基酸差异	1506个碱基进入衣壳中；第502个氨基酸进入衣壳中；在HVR 7中的VP3中
HSC12 SEQ ID NO: 12	G变为A	GGA -甘氨酸	AGA -精氨酸	甘氨酸是非极性并且中性的；精氨酸是极性并且强碱性的	1513个碱基进入衣壳中；第505个氨基酸进入衣壳中；紧接在HVR 7之后的VP3中
HSC12 SEQ ID NO: 12	G变为A	GTG -缬氨酸	ATG -甲硫氨酸	缬氨酸是非极性并且中性的；甲硫氨酸是非极性并且中性的	2041个碱基进入衣壳中；第681个氨基酸进入衣壳中；在VP3中；不在HVR中
HSC13 SEQ ID NO: 14	G变为A	GGA -甘氨酸	AGA -精氨酸	甘氨酸是非极性并且中性的；精氨酸是极性并且强碱性的	1513个碱基进入衣壳中；第505个氨基酸进入衣壳中；紧接在HVR 7之后的VP3中

图2(续)

衣壳	碱基变化	AAV 9 hu.分离株氨基酸	克隆体新氨基酸	氨基酸差异	氨基酸位置（HVR =高变区）
HSC14 SEQ ID NO: 15	G变为A	GGA -甘氨酸	AGA -精氨酸	甘氨酸是非极性并且中性的；精氨酸是极性并且强碱性的	1513个碱基进入衣壳中；第505个氨基酸进入衣壳中；紧接在HVR 7之后的VP3中
HSC14 SEQ ID NO: 15	T变为G	CTG -亮氨酸	CGG -精氨酸	亮氨酸是非极性并且中性的；精氨酸是极性并且强碱性的	2060个碱基进入衣壳中；第687个氨基酸进入衣壳中；在VP3中；不在HVR中。
HSC15 SEQ ID NO: 16	A变为G	ACG -苏氨酸	GCG -丙氨酸	苏氨酸是极性并且中性的；精氨酸是极性并且强碱性的	1036个碱基进入衣壳中；第346个氨基酸进入衣壳中；在VP3中；不在HVR中
HSC15 SEQ ID NO: 16	G变为A	GGA -甘氨酸	AGA -精氨酸	甘氨酸是非极性并且中性的；精氨酸是极性并且强碱性的	1513个碱基进入衣壳中；第505个氨基酸进入衣壳中；紧接在HVR 7之后的VP3中
HSC16 SEQ ID NO: 17	T变为A	TTT -苯丙氨酸	ATT -异亮氨酸	苯丙氨酸是非极性并且中性的；异亮氨酸是非极性并且中性的	1501个碱基进入衣壳中；第501个氨基酸进入衣壳中；在VP3中，在HVR7中
HSC16 SEQ ID NO: 17	G变为A	GGA -甘氨酸	AGA -精氨酸	甘氨酸是非极性并且中性的；精氨酸是极性并且强碱性的	1513个碱基进入衣壳中；第505个氨基酸进入衣壳中；紧接在HVR 7之后的VP3中
HSC16 SEQ ID NO: 17	G变为A	CCG -脯氨酸	CCA -脯氨酸	无氨基酸差异	1713个碱基进入衣壳中；第571个氨基酸进入衣壳中；在VP3中，不在HVR中
HSC16 SEQ ID NO: 17	A变为G	TAC -酪氨酸	TGC -半胱氨酸	酪氨酸是极性并且中性的；半胱氨酸是极性并且中性的	2117个碱基进入衣壳中；第706个氨基酸进入衣壳中；在VP3中；在HVR 12中并且紧邻YF突变
HSC17 SEQ ID NO: 13	C变为T	GAC -天冬氨酸	GAT -天冬氨酸	无氨基酸差异	180个碱基进入衣壳中；第60个氨基酸进入衣壳中；在VP1中；不在HVR中
HSC17 SEQ ID NO: 13	G变为A	GGA -甘氨酸	AGA -精氨酸	甘氨酸是非极性并且中性的；精氨酸是极性并且强碱性的	1513个碱基进入衣壳中；第505个氨基酸进入衣壳中；紧接在HVR 7之后的VP3中

图2(续)

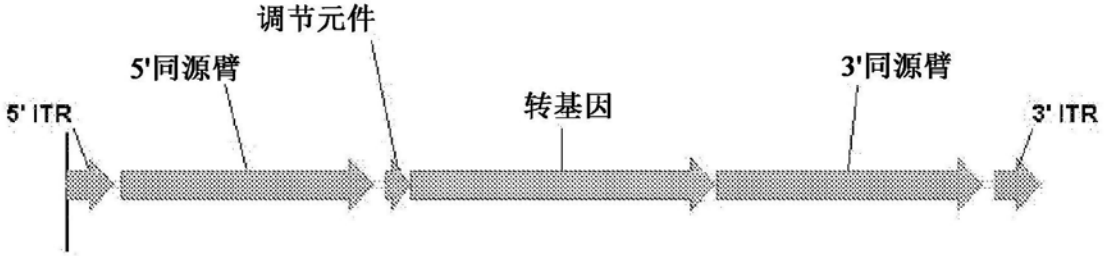

调节元件
5'同源臂
3'同源臂
5' ITR
转基因
3' ITR

图3

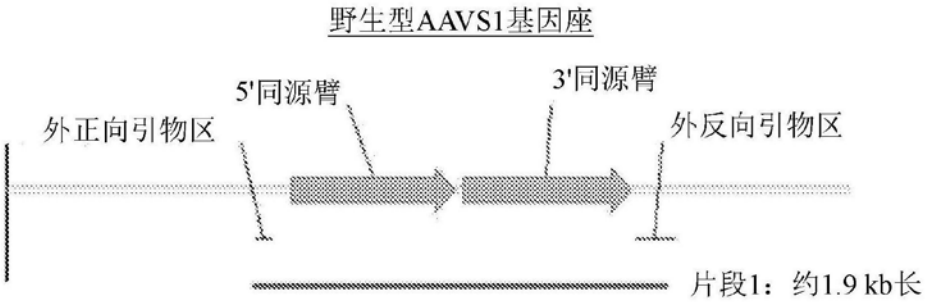

野生型AAVS1基因座
5'同源臂
3'同源臂
外正向引物区
外反向引物区
片段1：约1.9 kb长

图4A

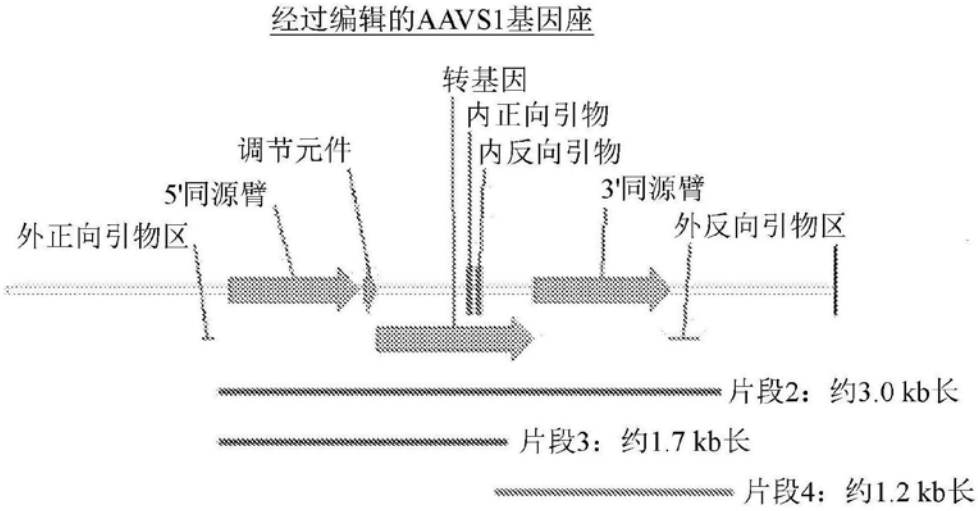

经过编辑的AAVS1基因座
转基因
内正向引物
内反向引物
调节元件
5'同源臂
3'同源臂
外正向引物区
外反向引物区
片段2：约3.0 kb长
片段3：约1.7 kb长
片段4：约1.2 kb长

图4B

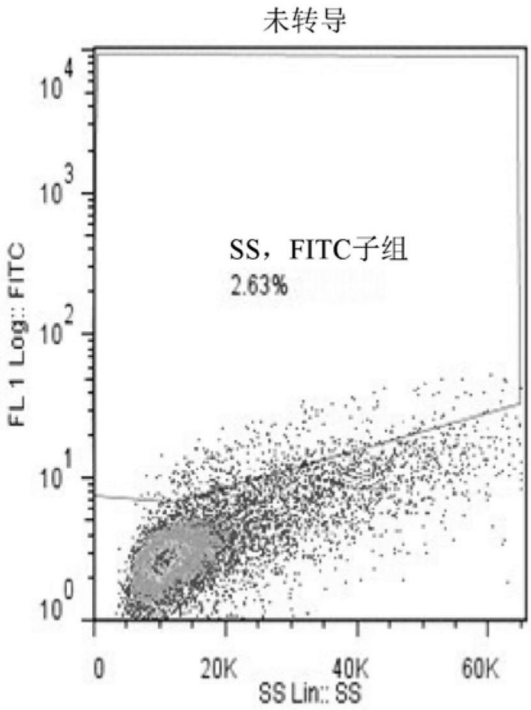

未转导
SS，FITC子组
2.63%
FL 1 Log:: FITC
SS Lin:: SS
0
20K
40K
60K

图5A

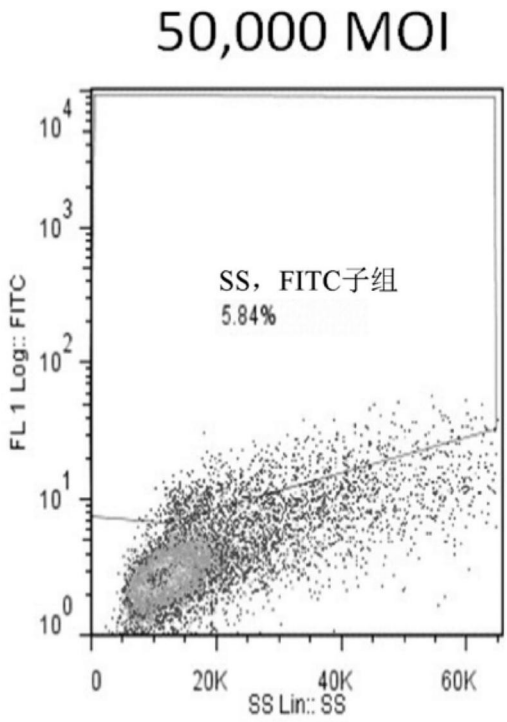

50,000 MOI
SS，FITC子组
5.84%
FL 1 Log:: FITC
SS Lin:: SS
0
20K
40K
60K

图5B

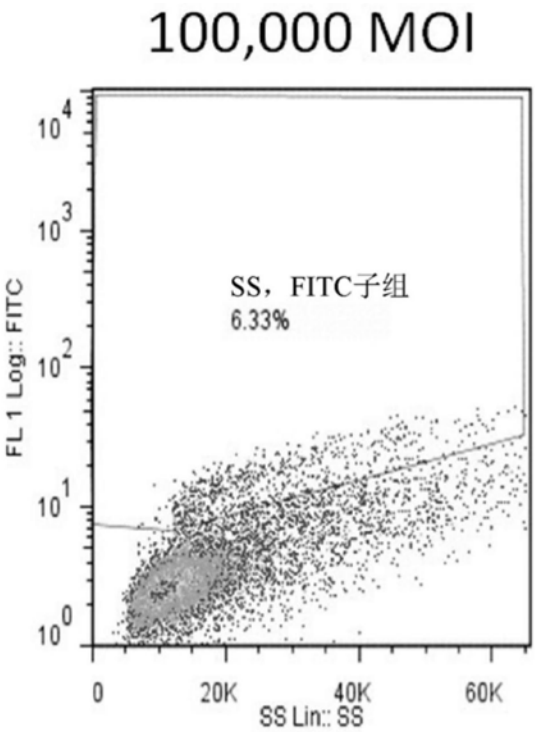

100,000 MOI
SS，FITC子组
6.33%
FL 1 Log:: FITC
SS Lin:: SS
0
20K
40K
60K

图5C

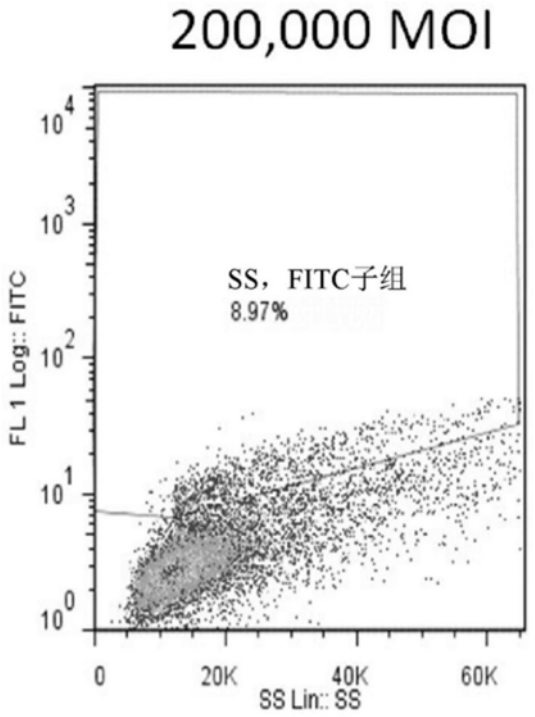

200,000 MOI
SS，FITC子组
8.97%
FL 1 Log:: FITC
SS Lin:: SS
0
20K
40K
60K

图5D

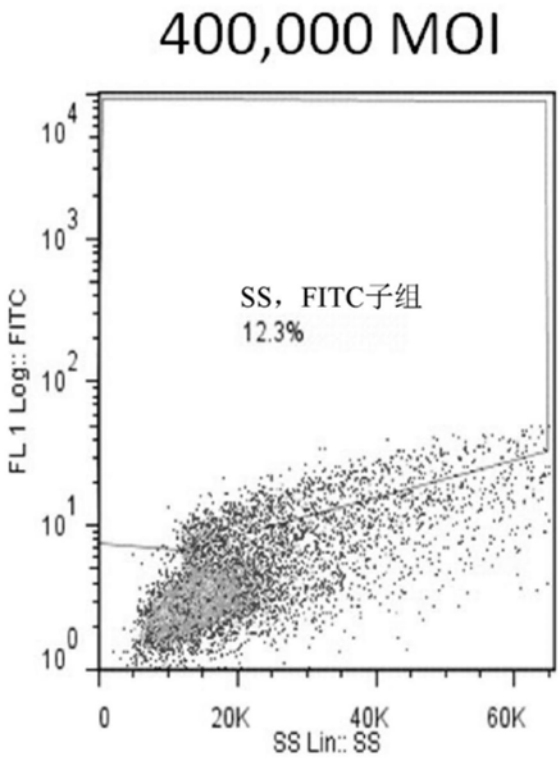

400,000 MOI
SS，FITC子组
12.3%
FL 1 Log:: FITC
SS Lin:: SS
0
20K
40K
60K

图5E

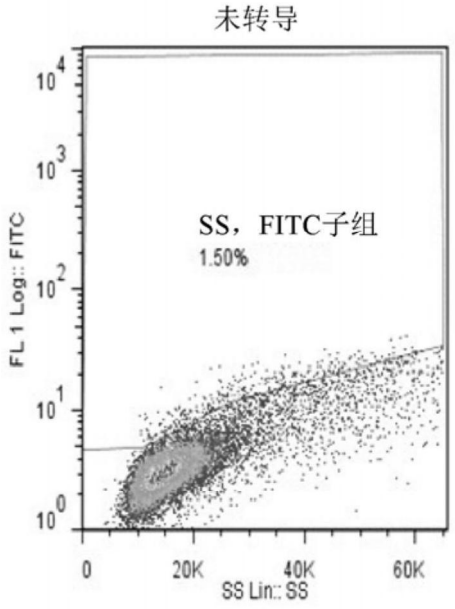

未转导
SS，FITC子组
1.50%
FL 1 Log:: FITC
SS Lin:: SS
0
20K
40K
60K

图6A

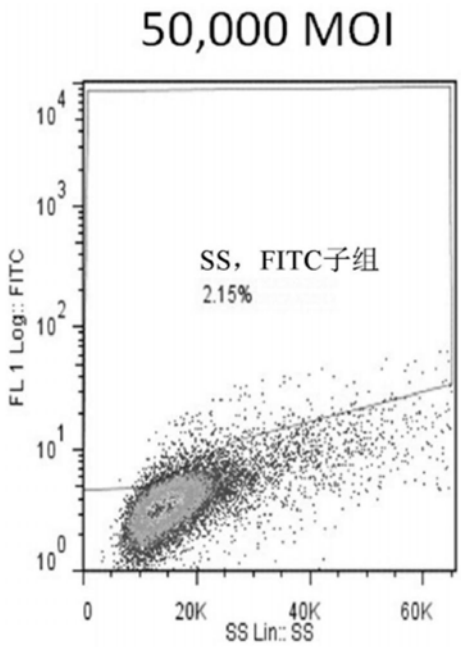

50,000 MOI
SS，FITC子组
2.15%
FL 1 Log:: FITC
SS Lin:: SS
0
20K
40K
60K

图6B

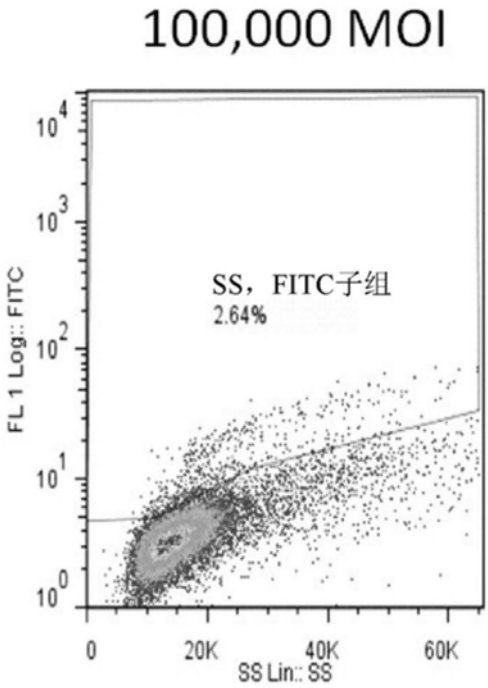

100,000 MOI
SS，FITC子组
2.64%
FL 1 Log:: FITC
SS Lin:: SS
0
20K
40K
60K

图6C

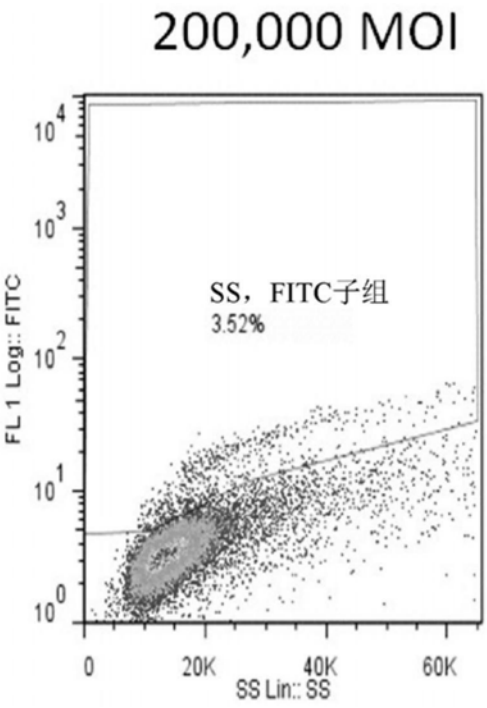

200,000 MOI
SS，FITC子组
3.52%
FL 1 Log:: FITC
SS Lin:: SS
0
20K
40K
60K

图6D

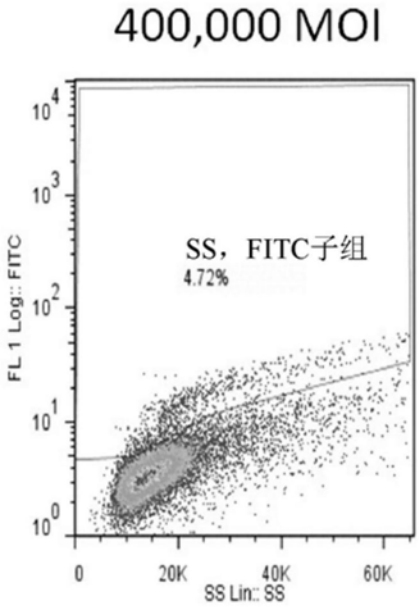

400,000 MOI
SS，FITC子组
4.72%
FL 1 Log:: FITC
SS Lin:: SS
0
20K
40K
60K

图6E

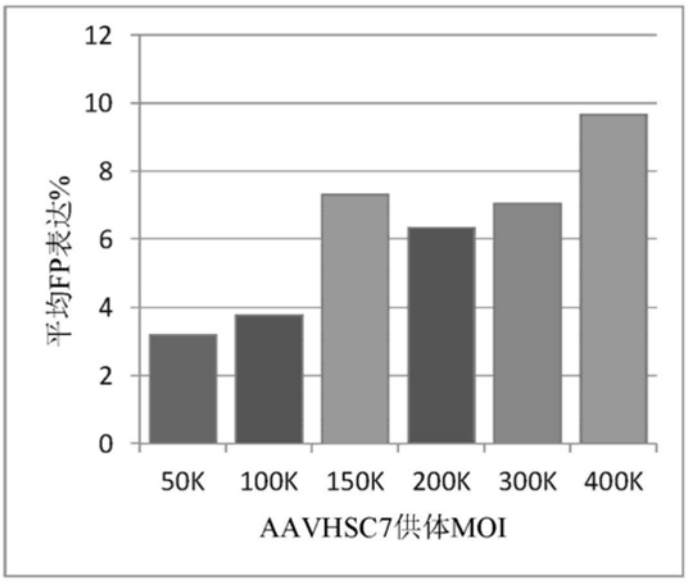

12
10
8
6
4
2
0
平均FP表达%
50K
100K
150K
200K
300K
400K
AAVHSC7供体MOI

图7A

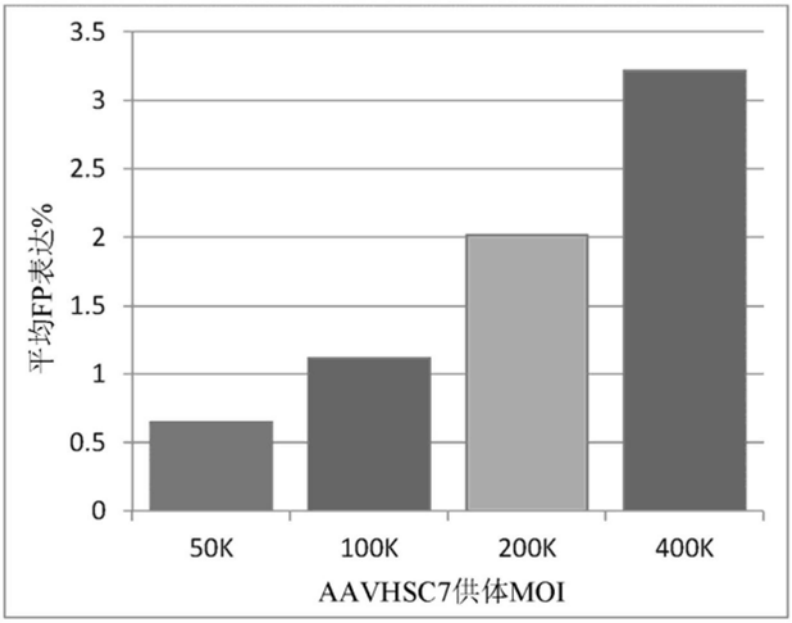

3.5
3
2.5
2
1.5
1
0.5
0
平均FP表达%
50K
100K
200K
400K
AAVHSC7供体MOI

图7B

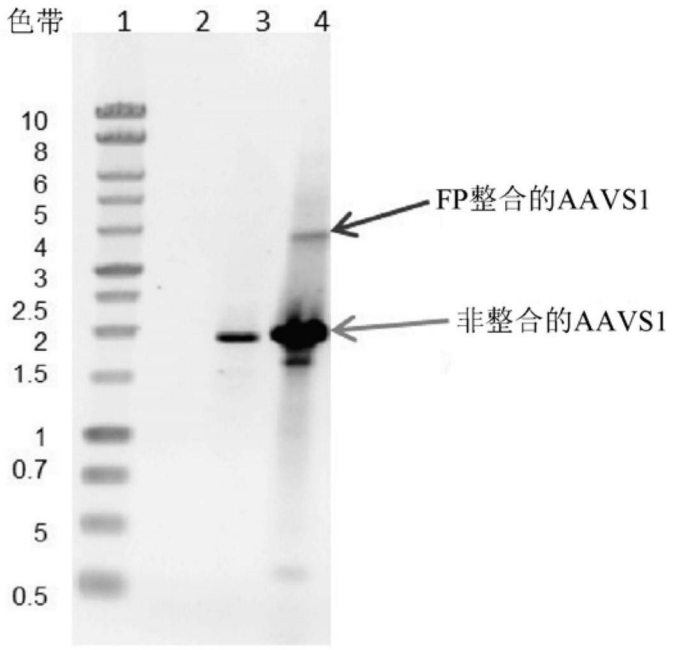

色带
1
2
3
4
10
8
6
5
4
3
2.5
2
1.5
1
0.7
5
0.5
FP整合的AAVS1
非整合的AAVS1

图8A

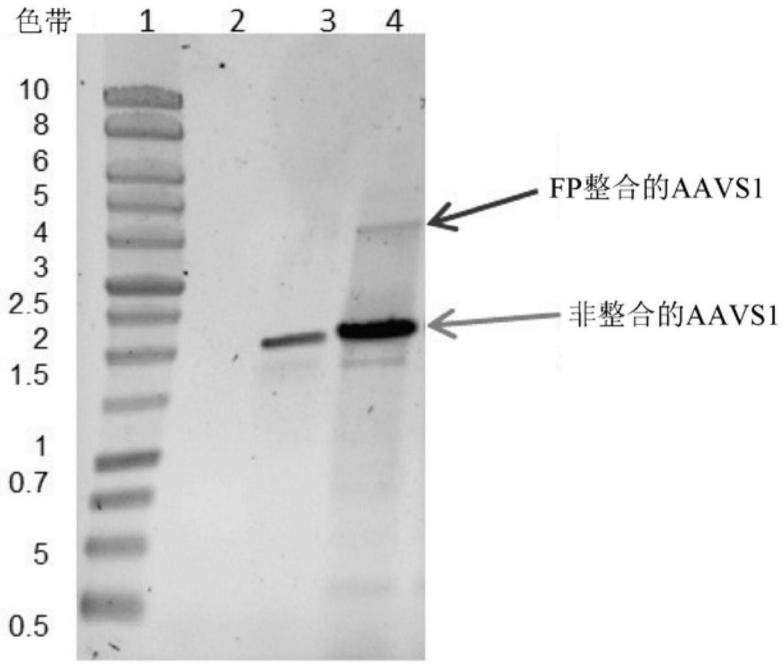

色带
1
2
3
4
10
8
6
5
4
3
2.5
2
1.5
1
0.7
5
0.5
FP整合的AAVS1
非整合的AAVS1

图8B

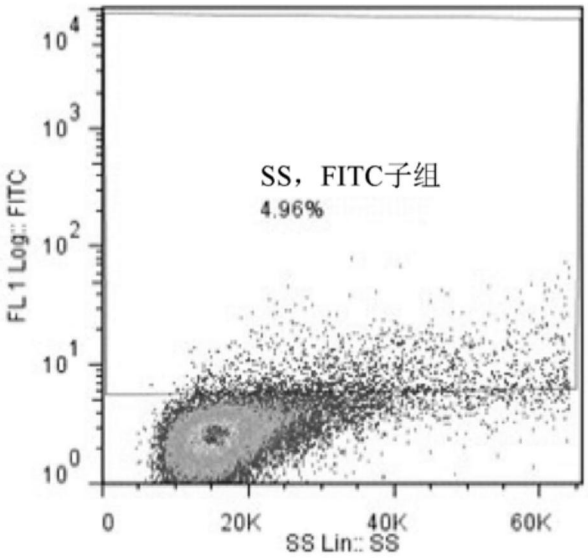

SS，FITC子组
4.96%
FL 1 Log:: FITC
SS Lin:: SS
0
20K
40K
60K

图9A

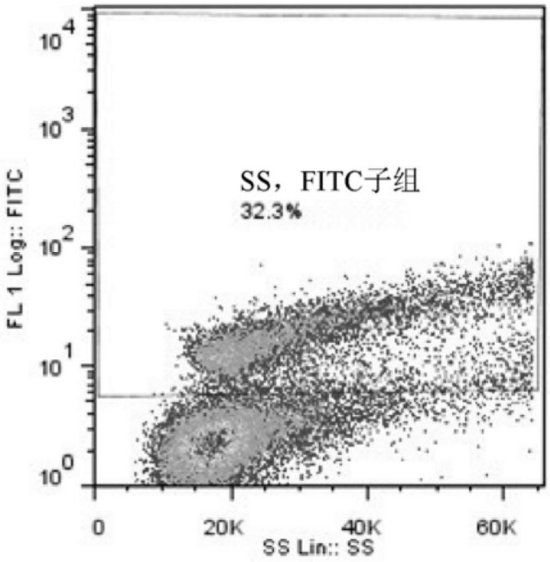

SS，FITC子组
32.3%
FL 1 Log:: FITC
SS Lin:: SS
0
20K
40K
60K

图9B

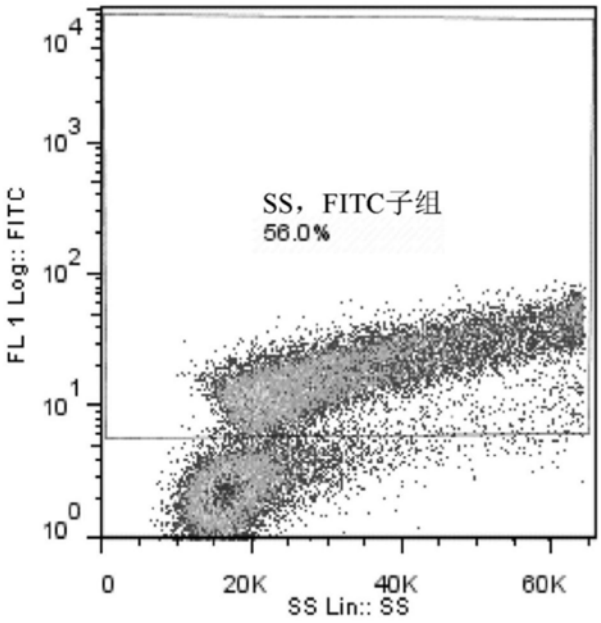

SS，FITC子组
56.0%
FL 1 Log:: FITC
SS Lin:: SS
0
20K
40K
60K

图9C

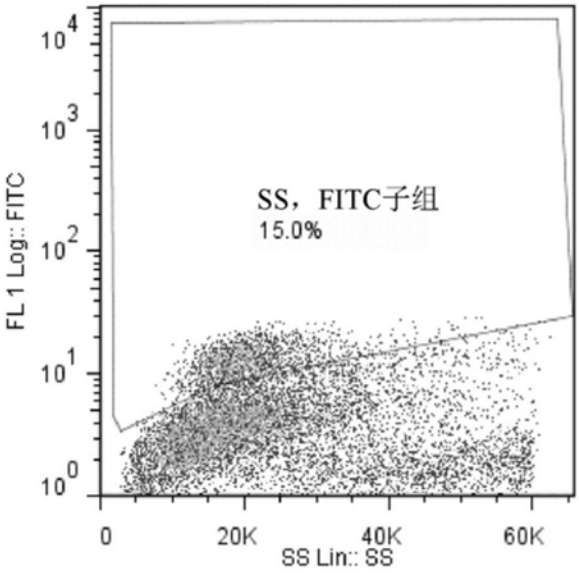

SS，FITC子组
15.0%
FL 1 Log:: FITC
SS Lin:: SS
0
20K
40K
60K

图10A

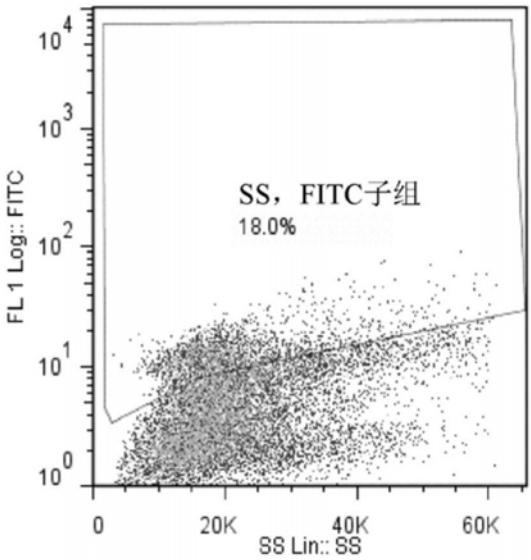

SS，FITC子组
18.0%
FL 1 Log:: FITC
SS Lin:: SS
0
20K
40K
60K

图10B

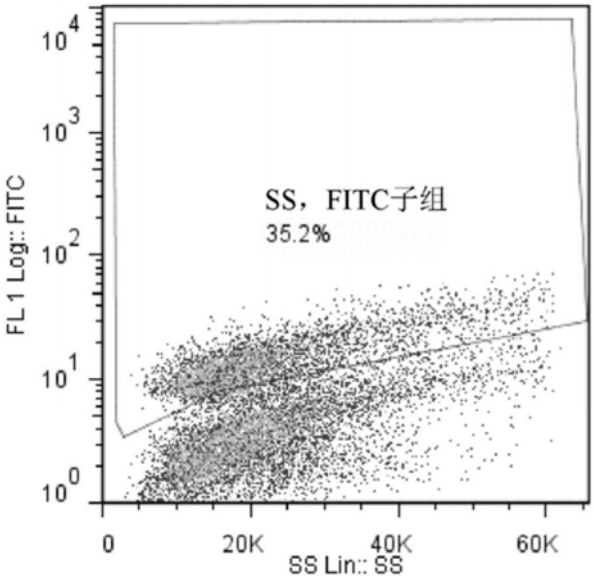

SS，FITC子组
35.2%
FL 1 Log:: FITC
SS Lin:: SS
0
20K
40K
60K

图10C

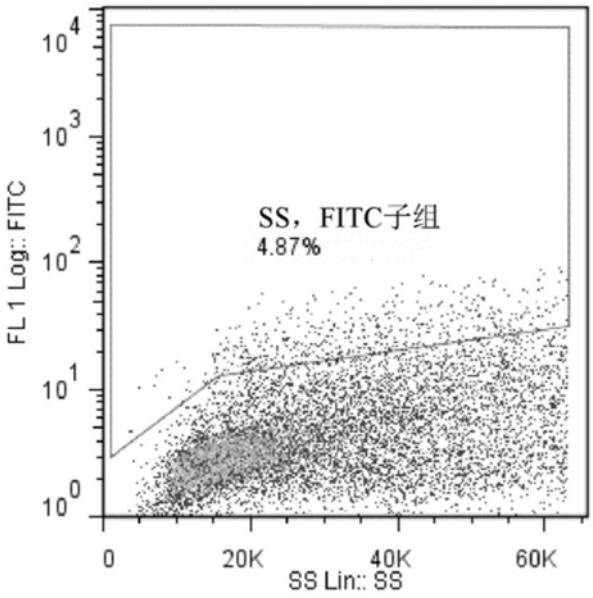

SS，FITC子组
4.87%
FL 1 Log:: FITC
SS Lin:: SS
0
20K
40K
60K

图11A

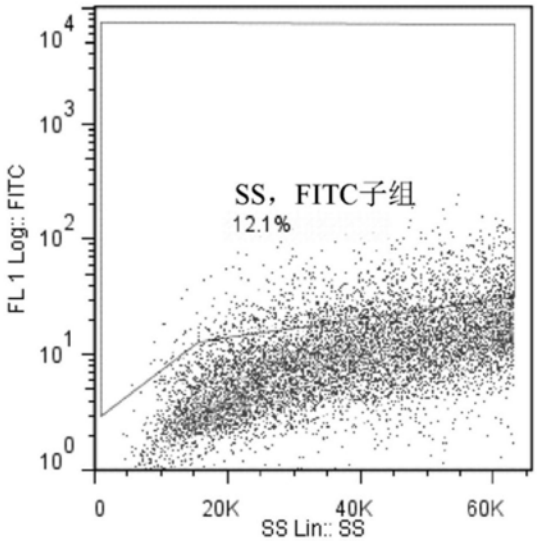

SS，FITC子组
12.1%
FL 1 Log:: FITC
SS Lin:: SS
0
20K
40K
60K

图11B

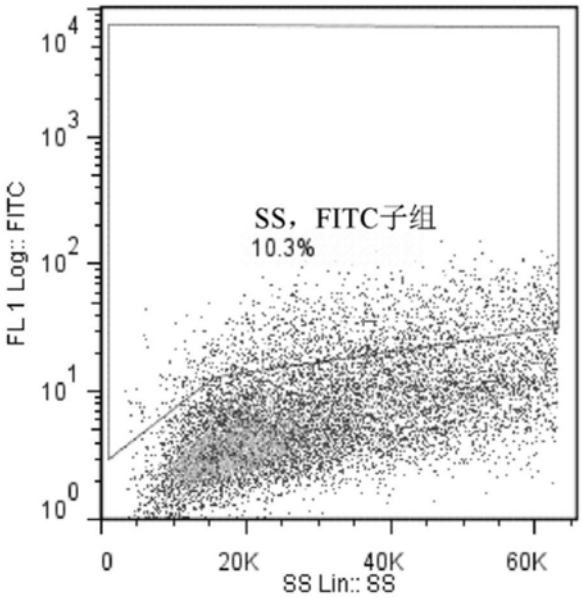


图11C

FP阳性细胞%

	4天	18天	20天	39天
未转导	0	0	0	0
AAVHSC7 FP供体	3	7.23	4.07	4.61
AAVHSC17 FP供体	20.2	5.43	2.85	9.18

图12A

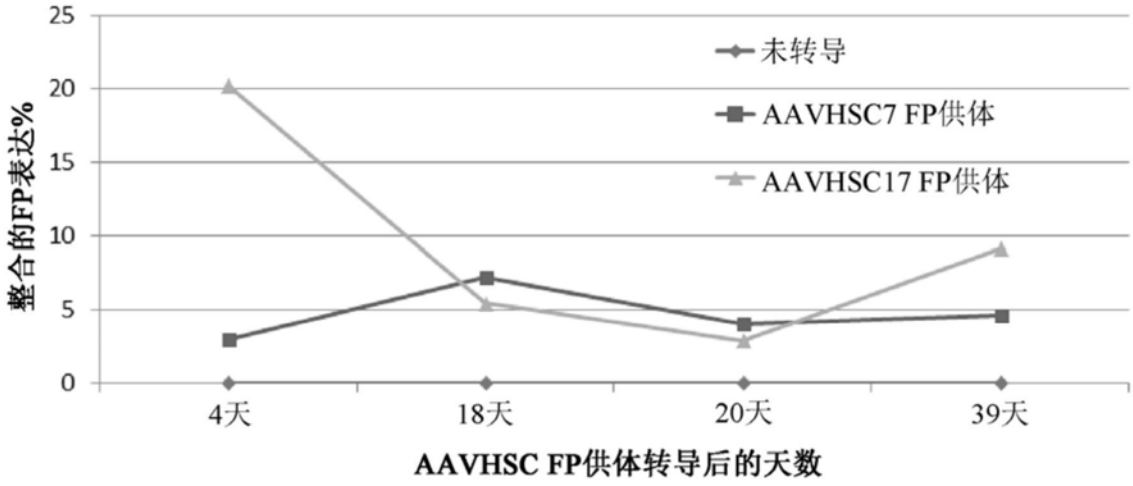
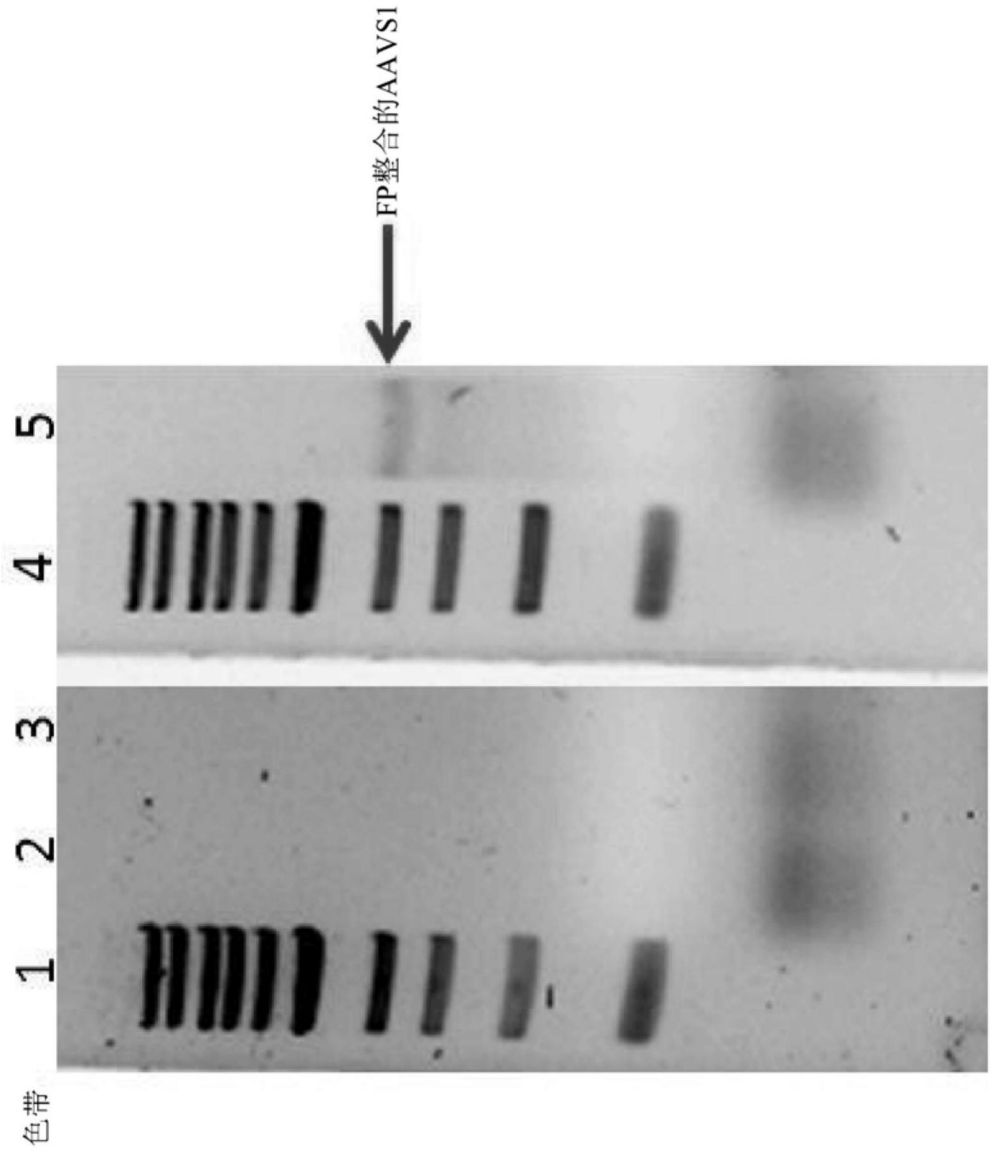


图12B


FP整合的AAVS1
5
4
3
2
1
色带

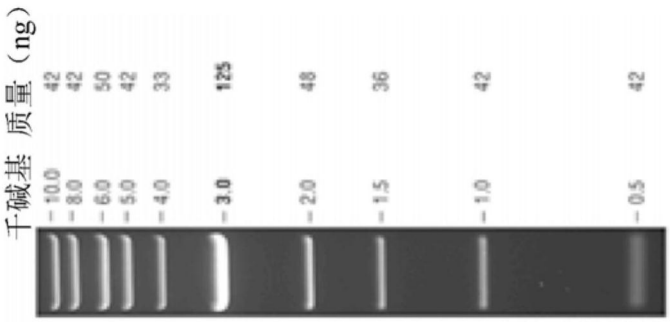

千碱基 质量（ng）
10.0 42
8.0 42
6.0 50
5.0 42
4.0 33
3.0 125
2.0 48
1.5 36
1.0 42
0.5 42

图13

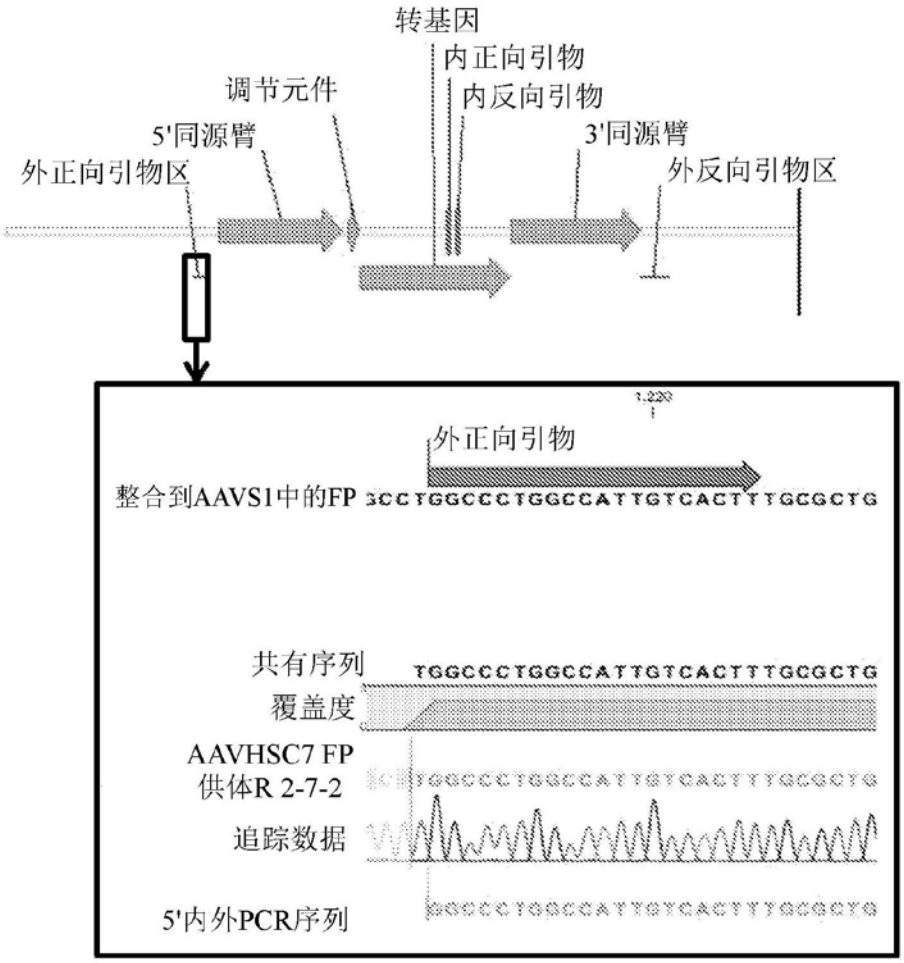

转基因
内正向引物
内反向引物
调节元件
5'同源臂
3'同源臂
外正向引物区
外反向引物区
外正向引物
整合到AAVS1中的FP
共有序列
覆盖度
AAVHSC7 FP
供体R 2-7-2
追踪数据
5'内外PCR序列

图14

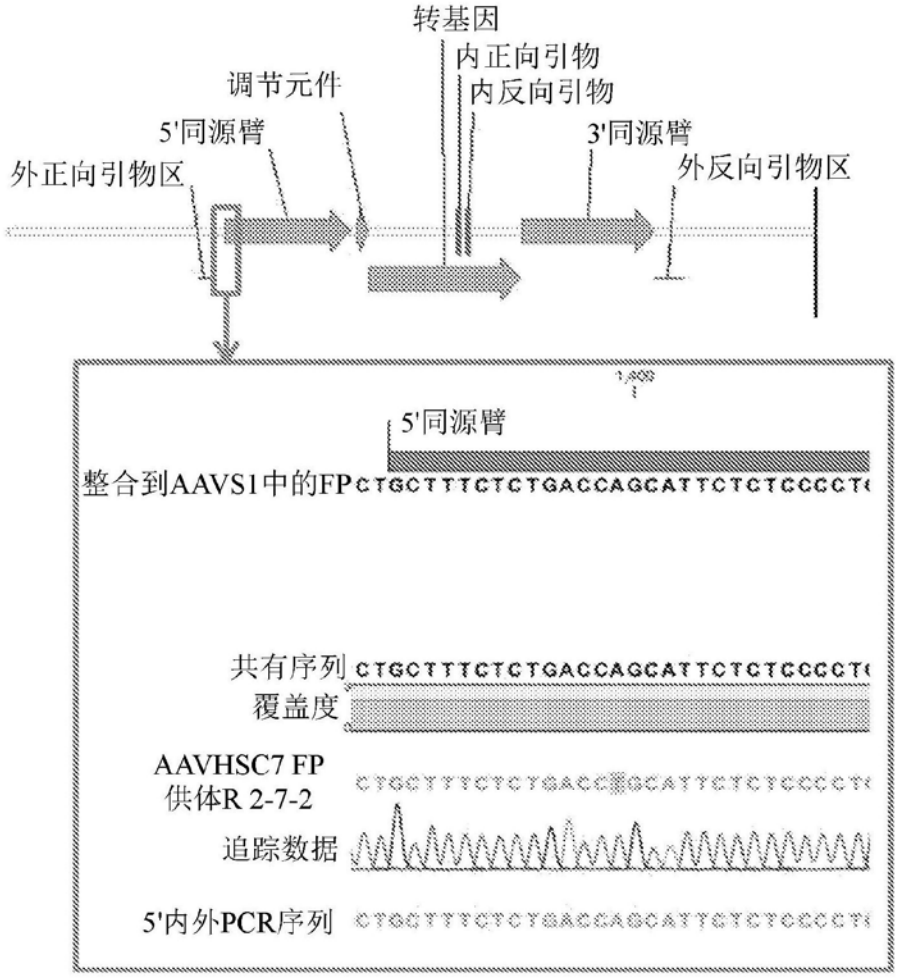

转基因
内正向引物
内反向引物
调节元件
5'同源臂
3'同源臂
外正向引物区
外反向引物区
5'同源臂
整合到AAVS1中的FP
共有序列
覆盖度
AAVHSC7 FP
供体R 2-7-2
追踪数据
5'内外PCR序列

图15

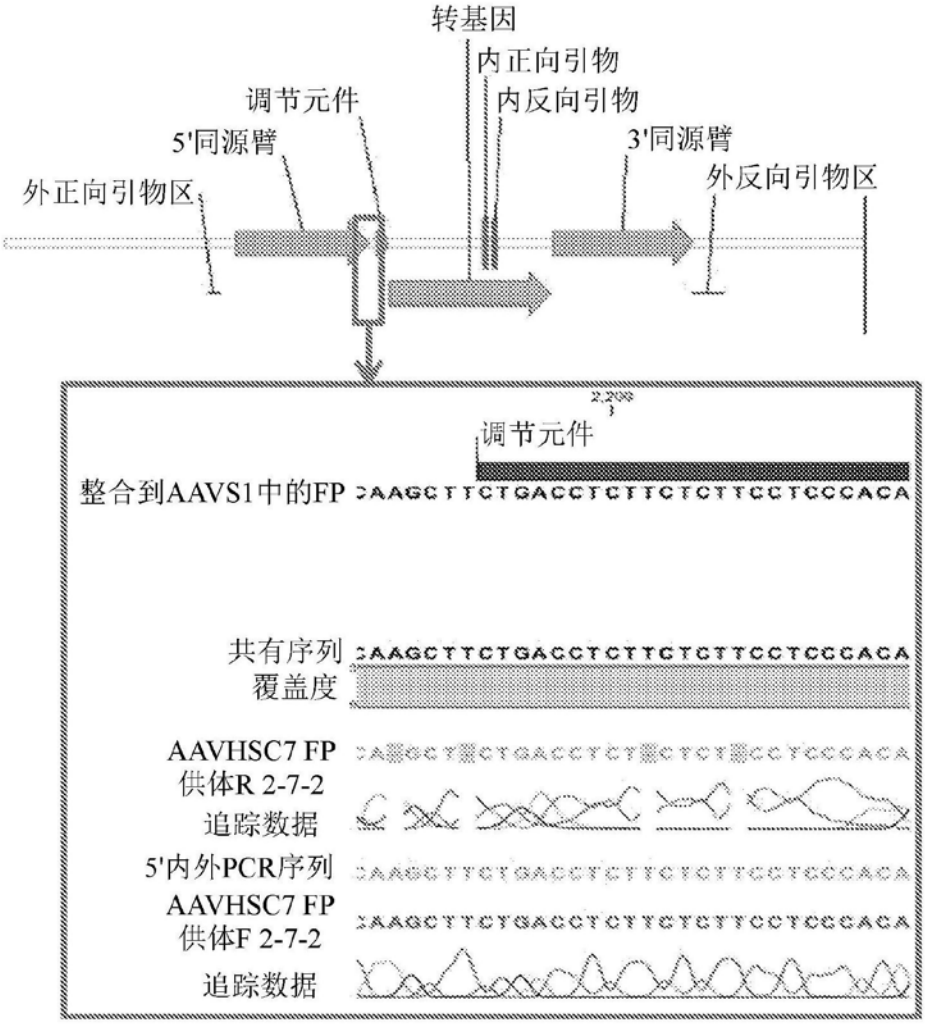

转基因
内正向引物
内反向引物
调节元件
5'同源臂
3'同源臂
外正向引物区
外反向引物区
2,200
调节元件
整合到AAVS1中的FP
共有序列
覆盖度
AAVHSC7 FP
供体R 2-7-2
追踪数据
5'内外PCR序列
AAVHSC7 FP
供体F 2-7-2
追踪数据

图16

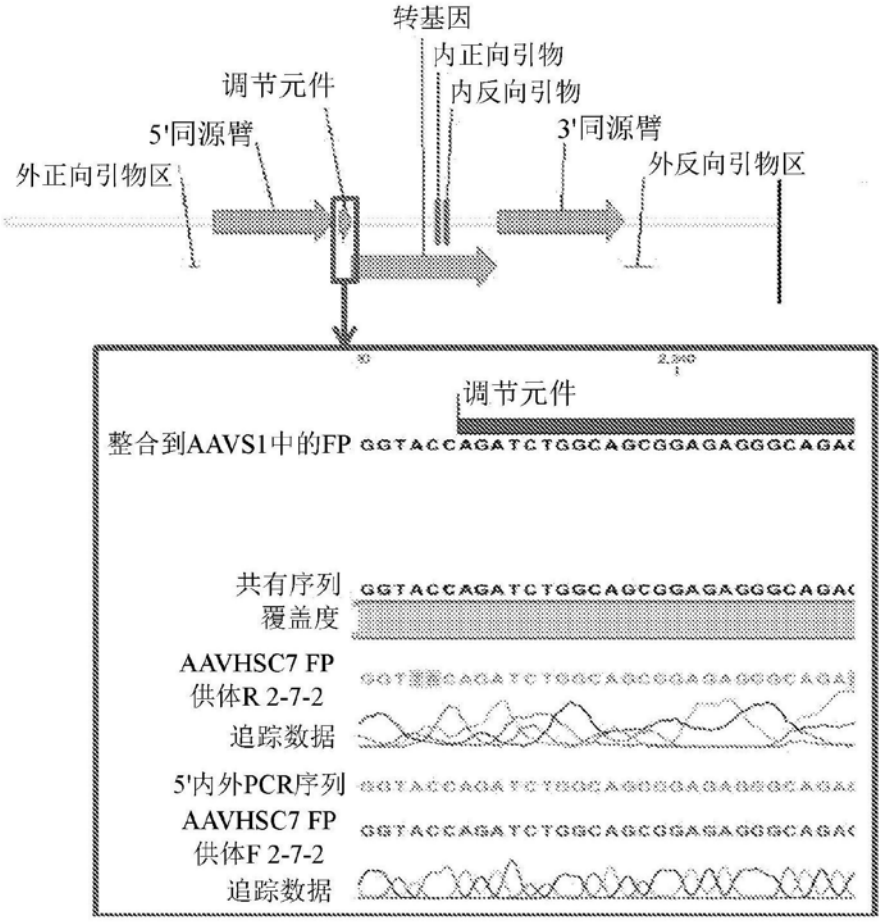

转基因
内正向引物
内反向引物
调节元件
5'同源臂
3'同源臂
外正向引物区
外反向引物区
调节元件
整合到AAVS1中的FP
共有序列
覆盖度
AAVHSC7 FP
供体R 2-7-2
追踪数据
5'内外PCR序列
AAVHSC7 FP
供体F 2-7-2
追踪数据

图17

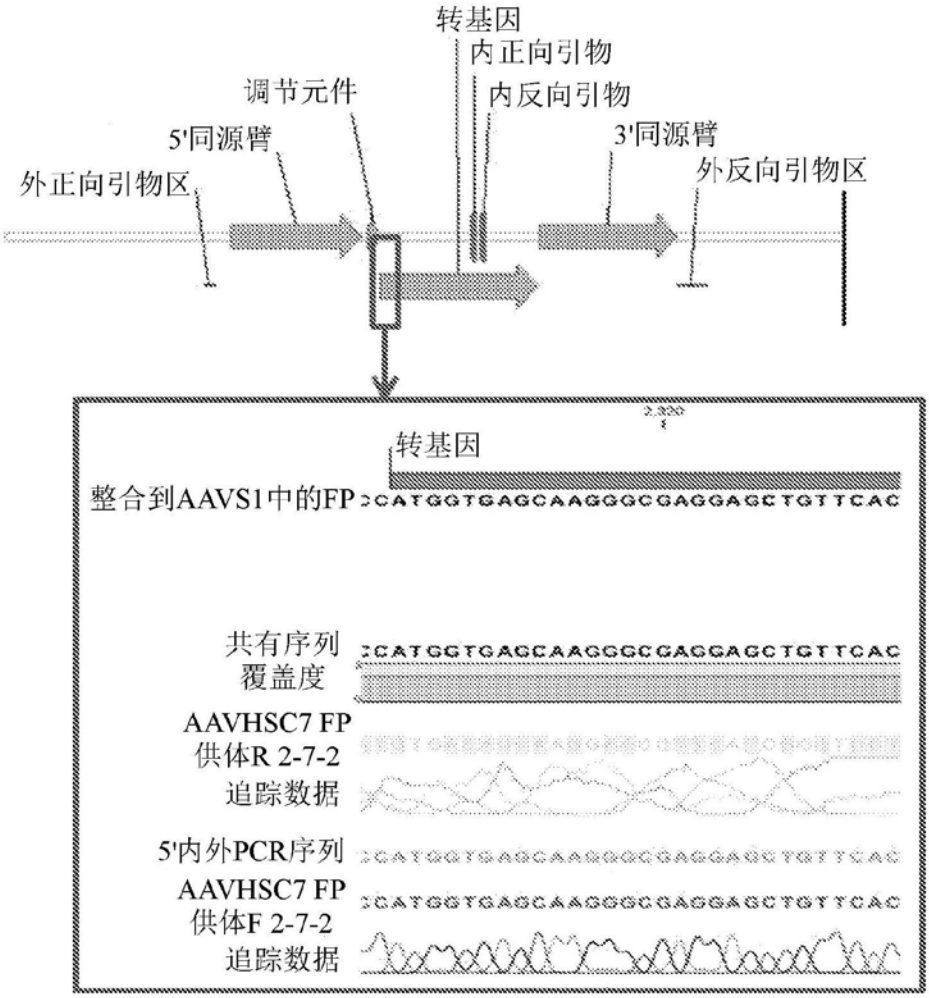

转基因
内正向引物
调节元件
内反向引物
5'同源臂
3'同源臂
外正向引物区
外反向引物区
转基因
整合到AAVS1中的FP
共有序列
覆盖度
AAVHSC7 FP
供体R 2-7-2
追踪数据
5'内外PCR序列
AAVHSC7 FP
供体F 2-7-2
追踪数据

图18

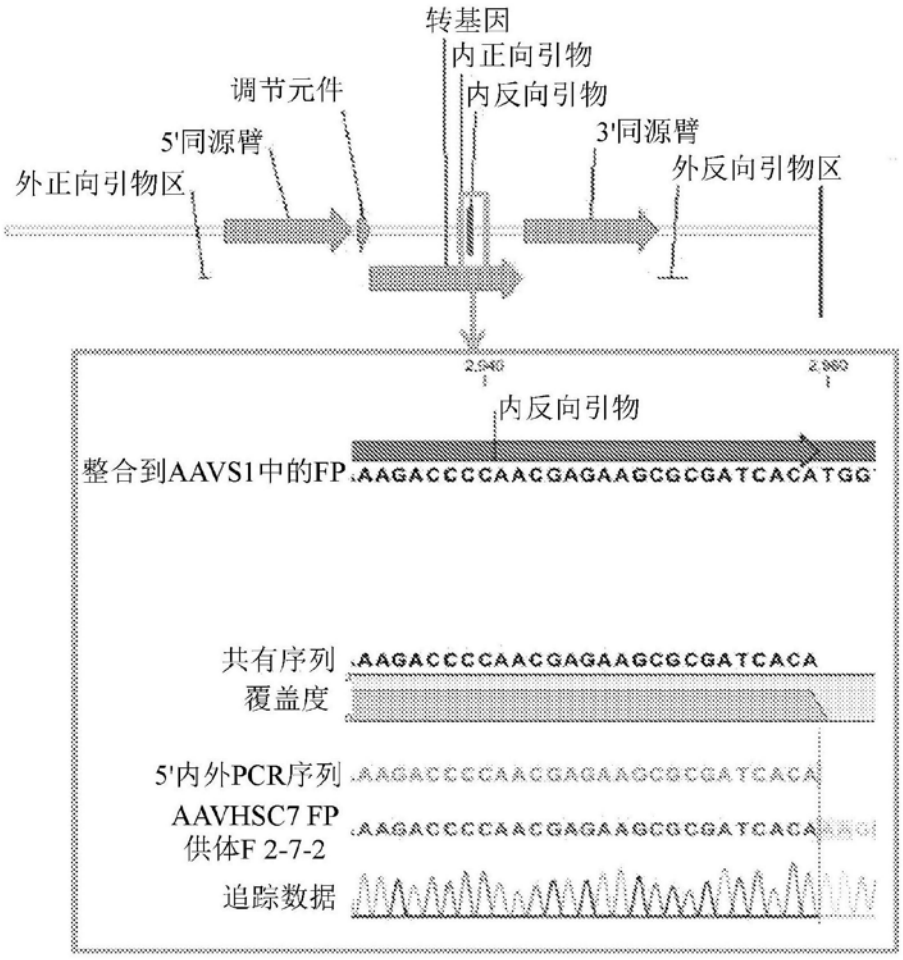

转基因
内正向引物
内反向引物
调节元件
5'同源臂
3'同源臂
外正向引物区
外反向引物区
2,940
2,960
内反向引物
整合到AAVS1中的FP
AAGACCCCAACGAGAAGCGCGATCACATGG
共有序列
AAGACCCCAACGAGAAGCGCGATCACA
覆盖度
5'内外PCR序列
AAGACCCCAACGAGAAGCGCGATCACA
AAVHSC7 FP
供体F 2-7-2
AAGACCCCAACGAGAAGCGCGATCACA
追踪数据

图19

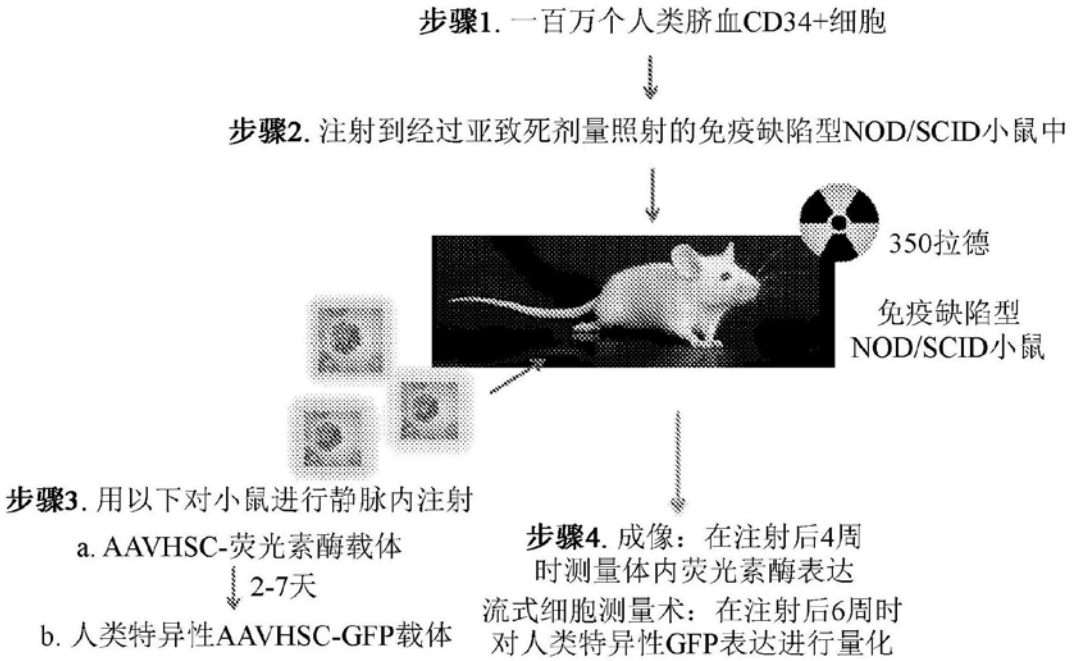


图20

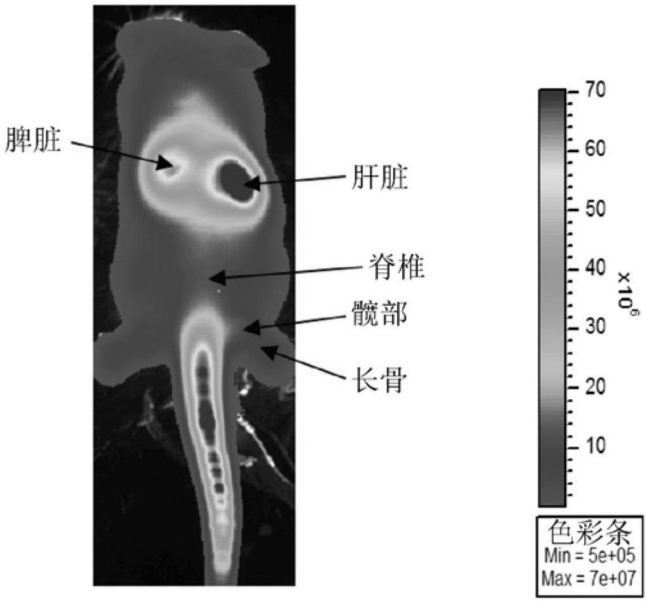


图21A

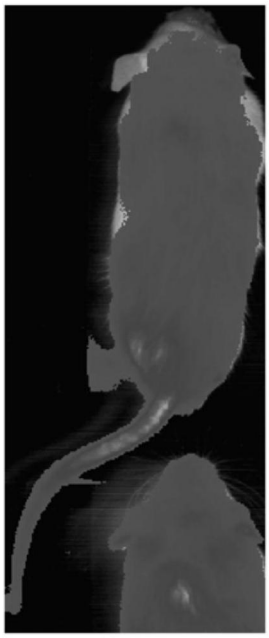


图21B

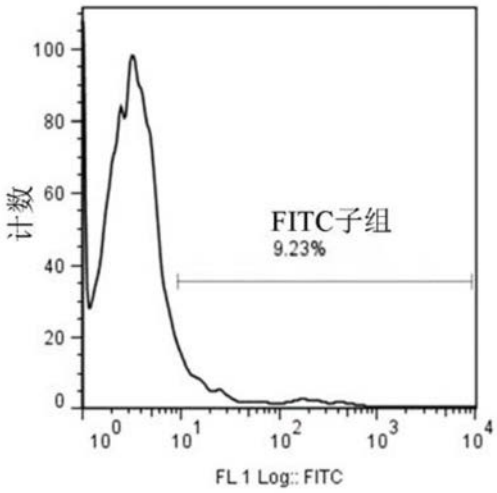

100
80
60
40
20
0
计数
FITC子组
9.23%
10^0
10^1
10^2
10^3
10^4
FL 1 Log:: FITC

图22A

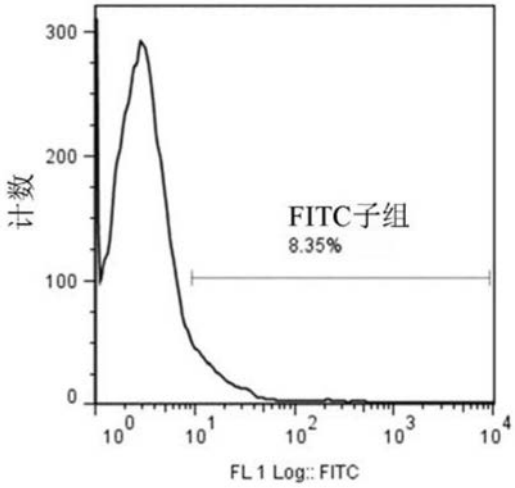

300
200
100
0
计数
FITC子组
8.35%
10^0
10^1
10^2
10^3
10^4
FL 1 Log:: FITC

图22B

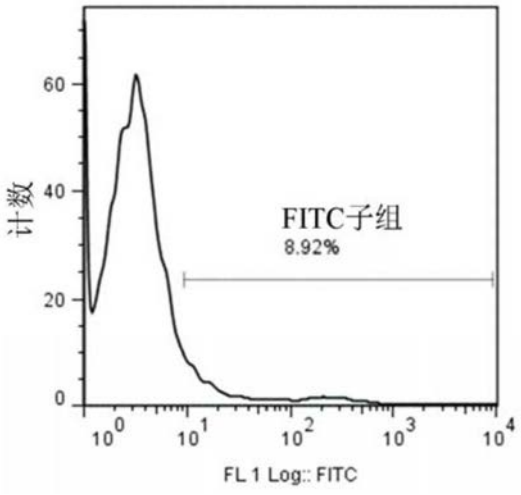

计数
60
40
20
0
FITC子组
8.92%
10^0
10^1
10^2
10^3
10^4
FL 1 Log:: FITC

图22C

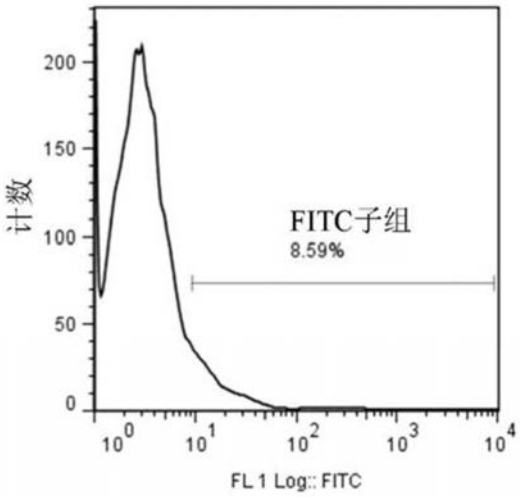

计数
200
150
100
50
0
FITC子组
8.59%
10^0
10^1
10^2
10^3
10^4
FL 1 Log:: FITC

图22D

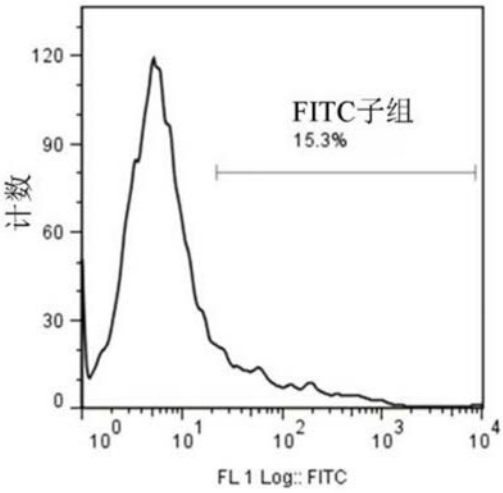

FITC子组
15.3%
计数
120
90
60
30
0
10^0
10^1
10^2
10^3
10^4
FL 1 Log:: FITC

图22E

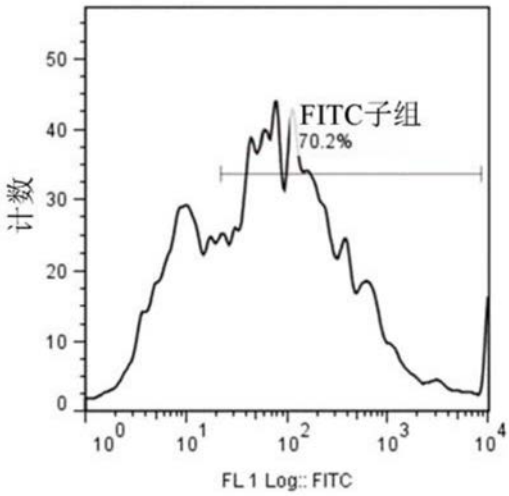

FITC子组
70.2%
计数
50
40
30
20
10
0
10^0
10^1
10^2
10^3
10^4
FL 1 Log:: FITC

图22F

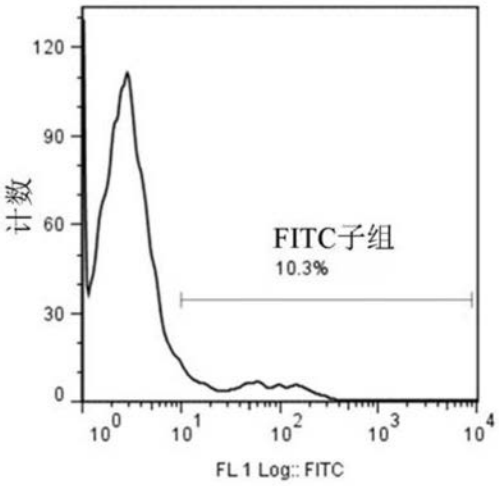

120
90
60
30
0
计数
FITC子组
10.3%
10^0
10^1
10^2
10^3
10^4
FL 1 Log:: FITC

图22G

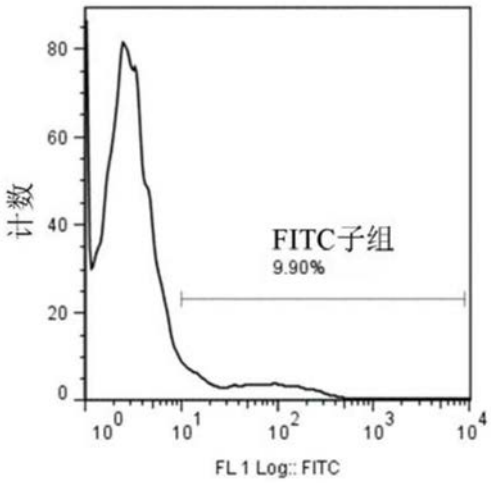

80
60
40
20
0
计数
FITC子组
9.90%
10^0
10^1
10^2
10^3
10^4
FL 1 Log:: FITC

图22H

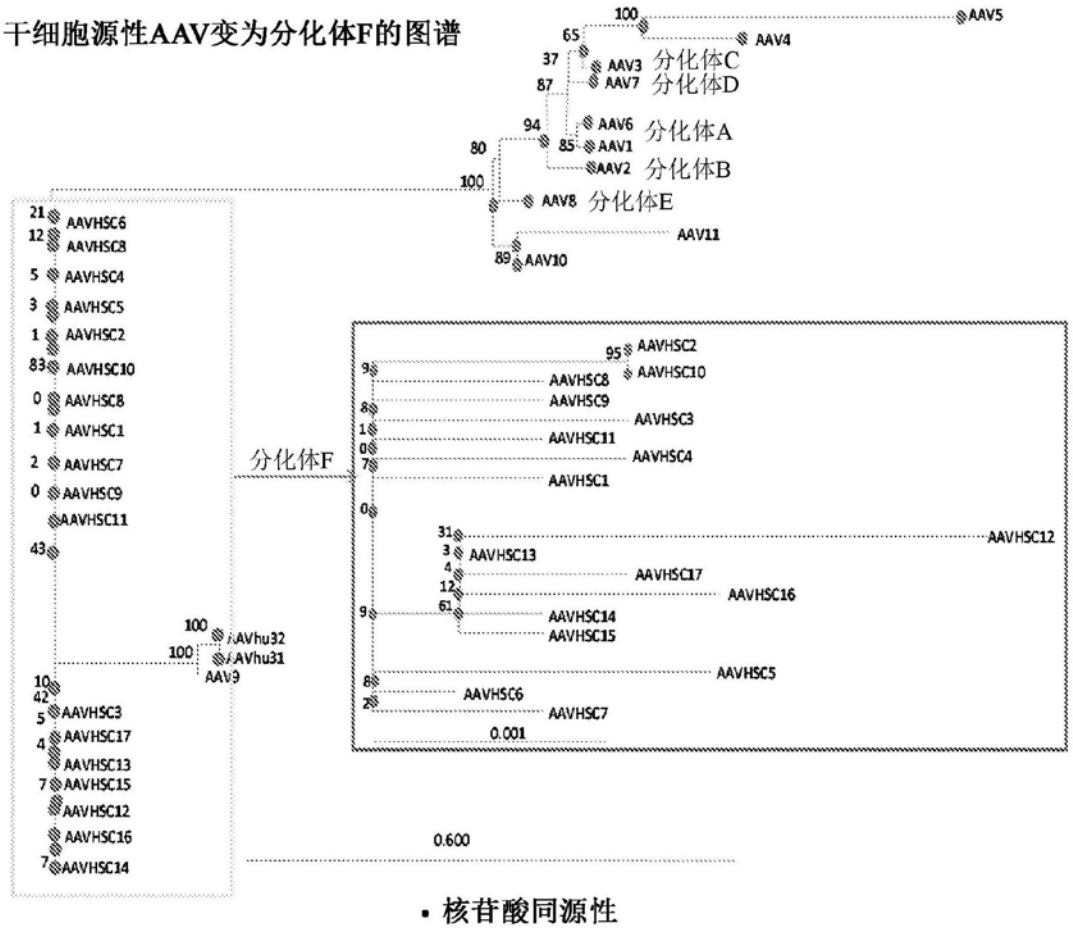

干细胞源性AAV变为分化体F的图谱
AAV5
AAV4
AAV3
AAV7
分化体C
分化体D
AAV6
AAV1
AAV2
分化体A
分化体B
AAV8
分化体E
AAV11
AAV10
AAVHSC6
AAVHSC8
AAVHSC4
AAVHSC5
AAVHSC2
AAVHSC10
AAVHSC8
AAVHSC1
AAVHSC7
AAVHSC9
AAVHSC11
AAVhu32
AAVhu31
AAV9
AAVHSC3
AAVHSC17
AAVHSC13
AAVHSC15
AAVHSC12
AAVHSC16
AAVHSC14
分化体F
AAVHSC2
AAVHSC10
AAVHSC8
AAVHSC9
AAVHSC3
AAVHSC11
AAVHSC4
AAVHSC1
AAVHSC12
AAVHSC13
AAVHSC17
AAVHSC16
AAVHSC14
AAVHSC15
AAVHSC5
AAVHSC6
AAVHSC7
0.001
0.600
• 核苷酸同源性

图23

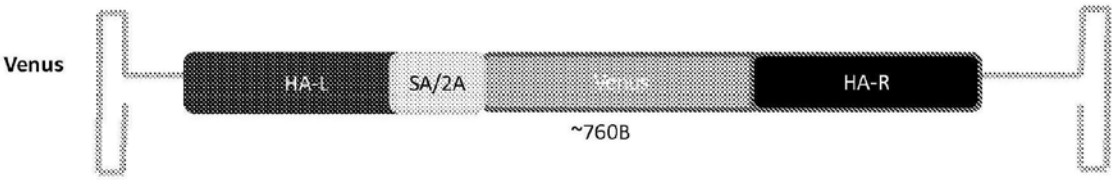

Venus
HA-L
SA/2A
Venus
HA-R
~760B

图24

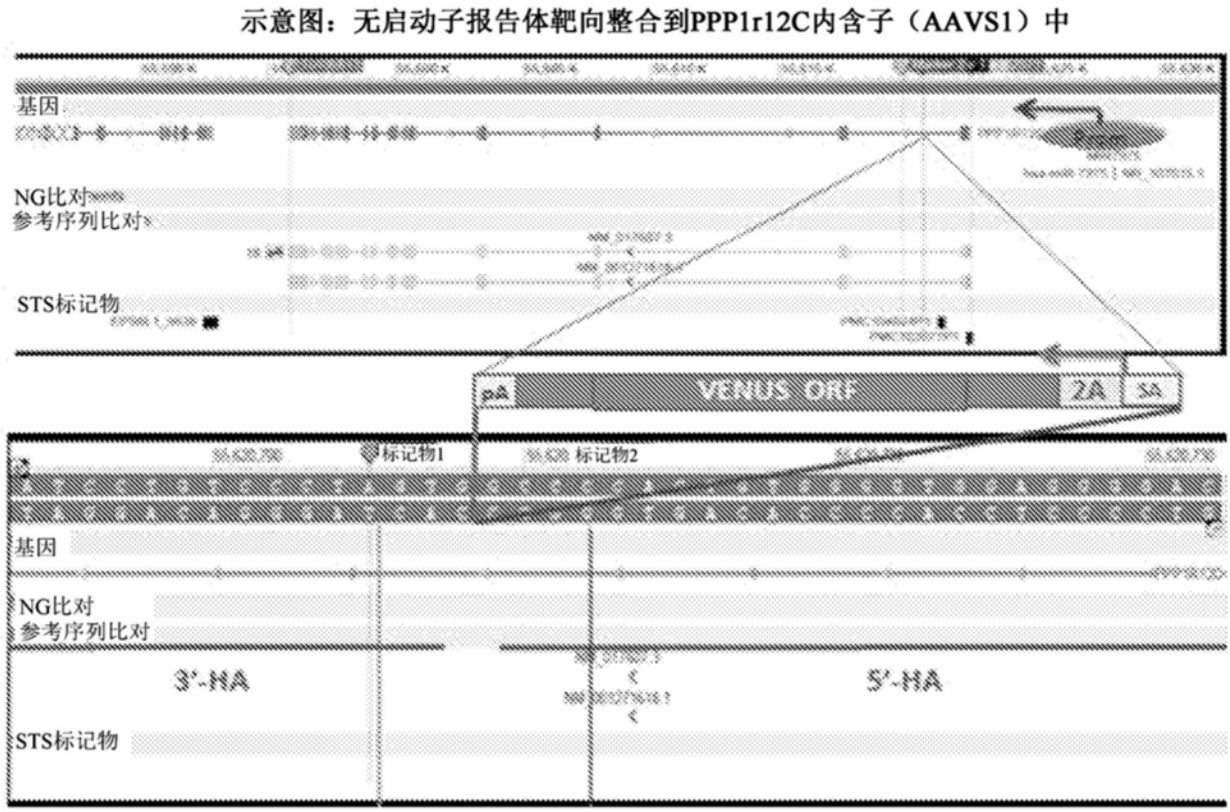

示意图：无启动子报告体靶向整合到PPP1r12C内含子（AAVS1）中
基因
NG比对
参考序列比对
STS标记物
pA
VENUS ORF
2A
SA
标记物1
标记物2
基因
NG比对
参考序列比对
3'-HA
5'-HA
STS标记物

图25

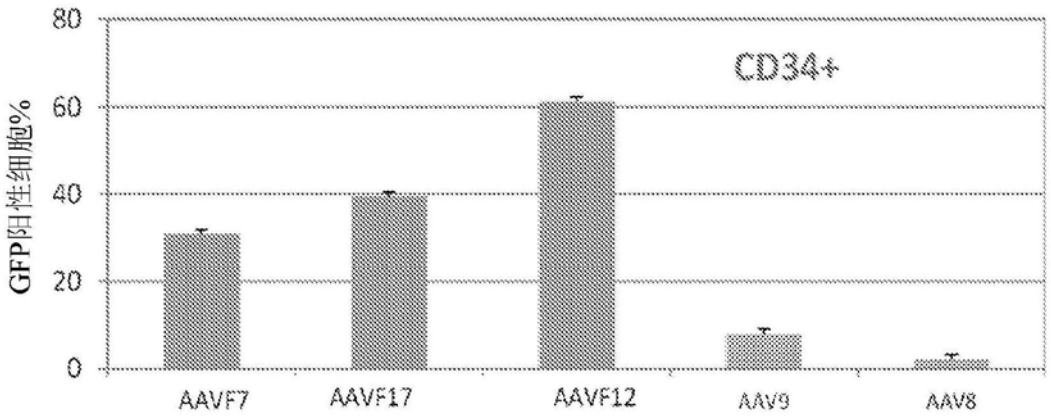

CD34+
GFP阳性细胞%
80
60
40
20
0
AAVF7
AAVF17
AAVF12
AAV9
AAV8

图26A

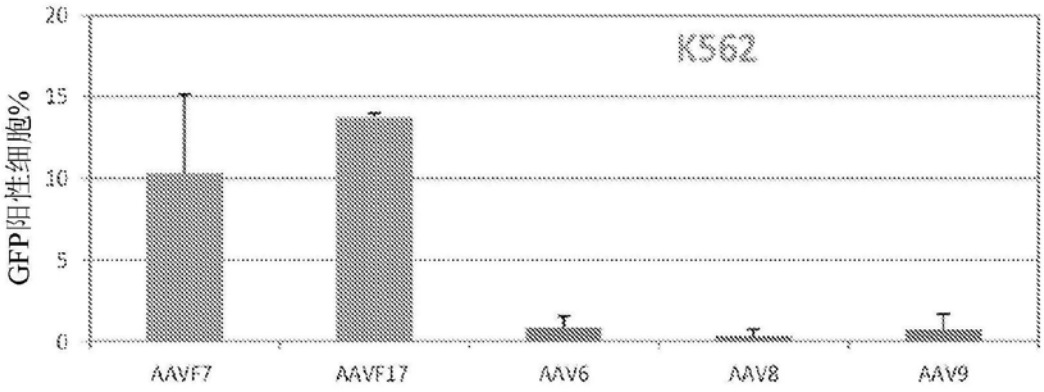

K562
GFP阳性细胞%
20
15
10
5
0
AAVF7
AAVF17
AAV6
AAV8
AAV9

图26B

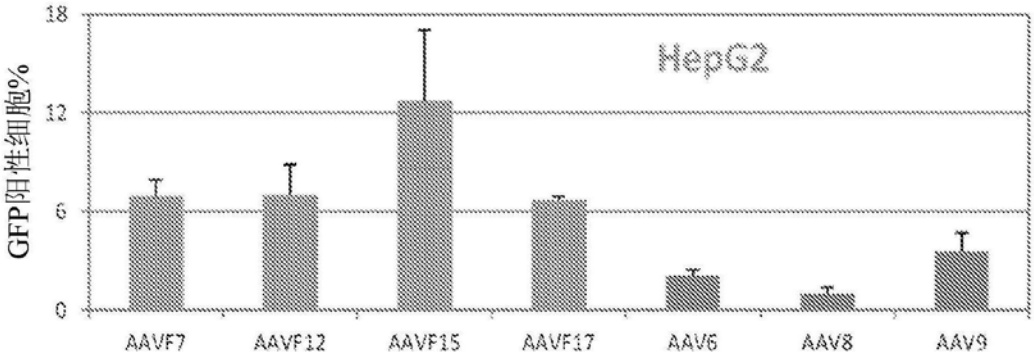
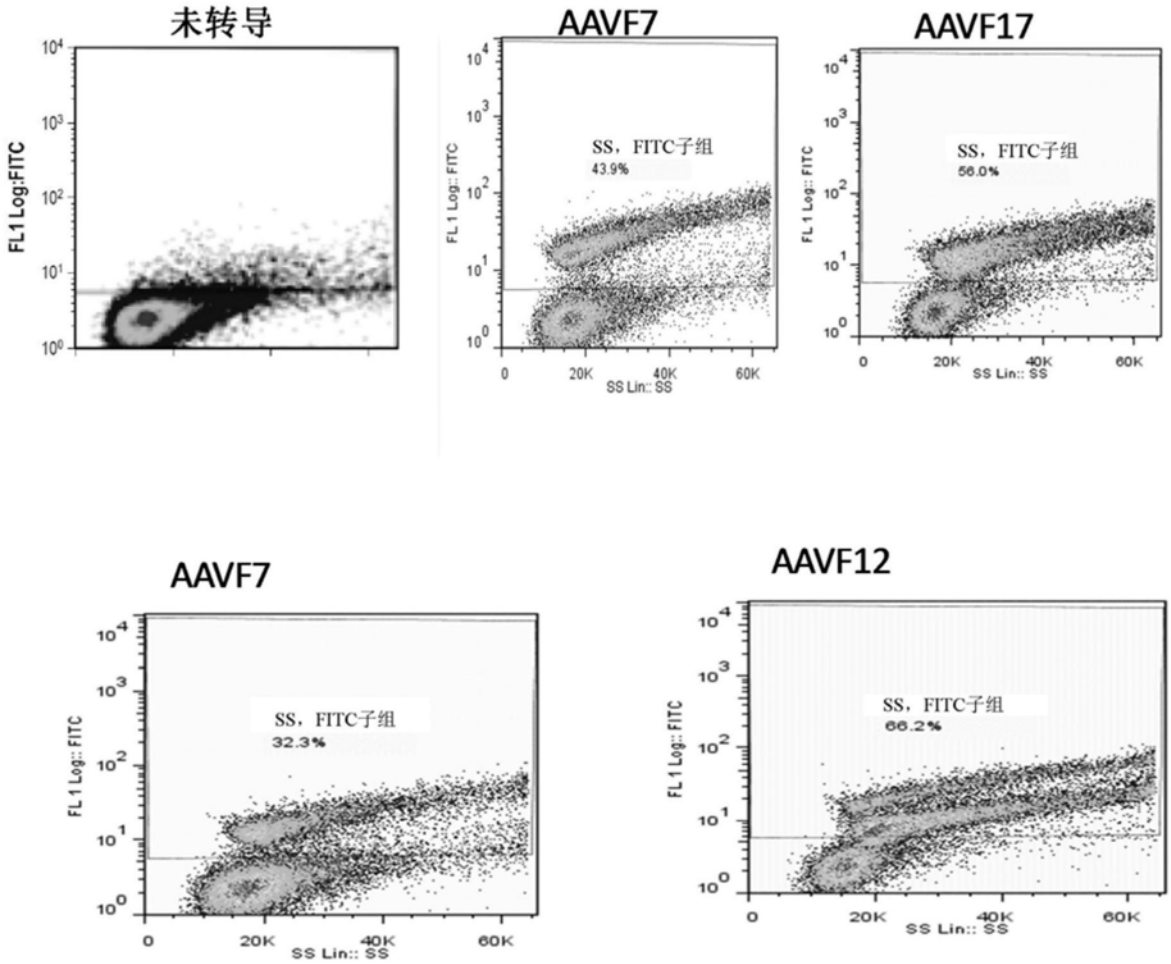

HepG2
GFP阳性细胞%
18
12
6
0
AAVF7
AAVF12
AAVF15
AAVF17
AAV6
AAV8
AAV9

图26C



- AAVF介导在原代CD34+细胞中的高效基因靶向
- 在人类白血病和肝脏细胞系中观察到高效靶向
- AAVF编辑不需要DNA合成

图26D

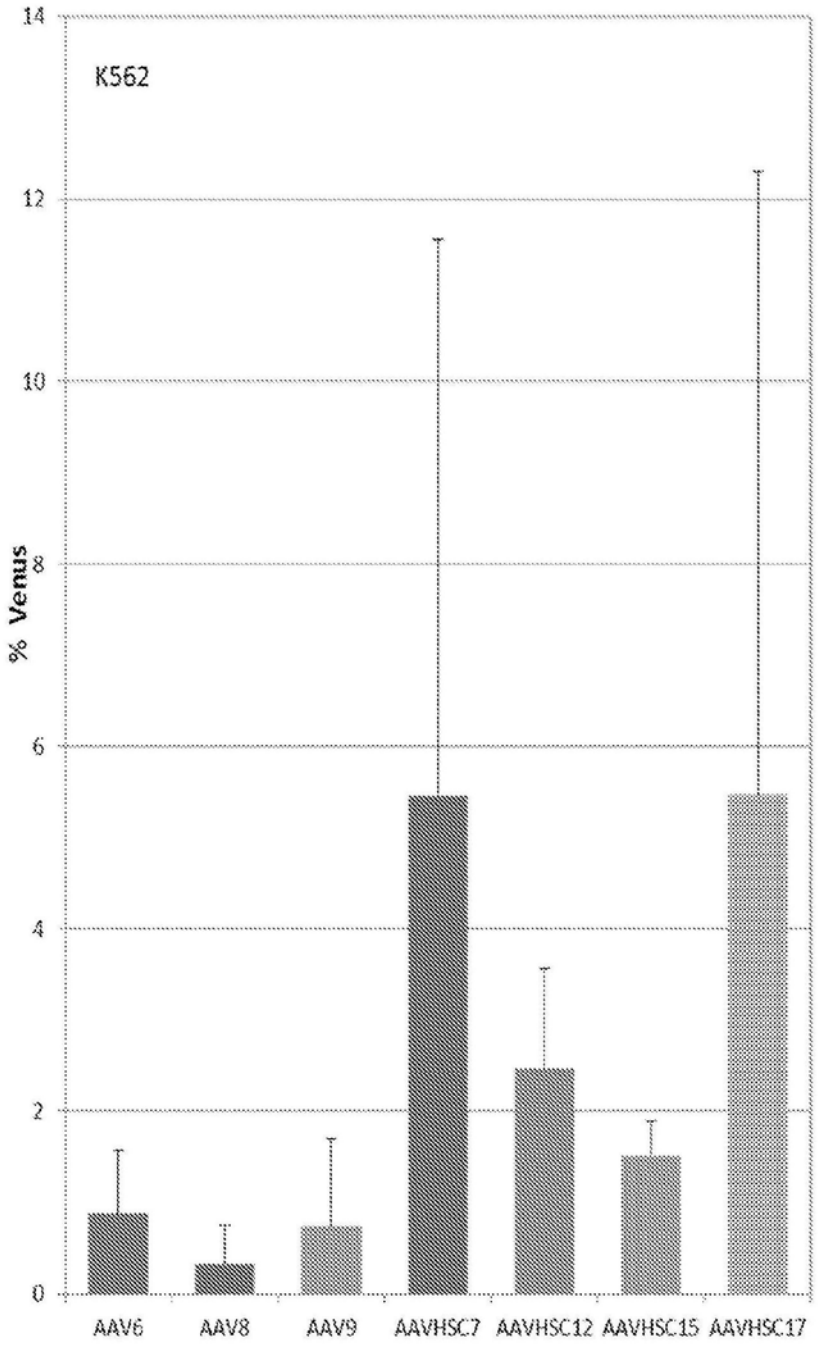

K562
% Venus
14
12
10
8
6
4
2
0
AAV6
AAV8
AAV9
AAVHSC7
AAVHSC12
AAVHSC15
AAVHSC17

图26E

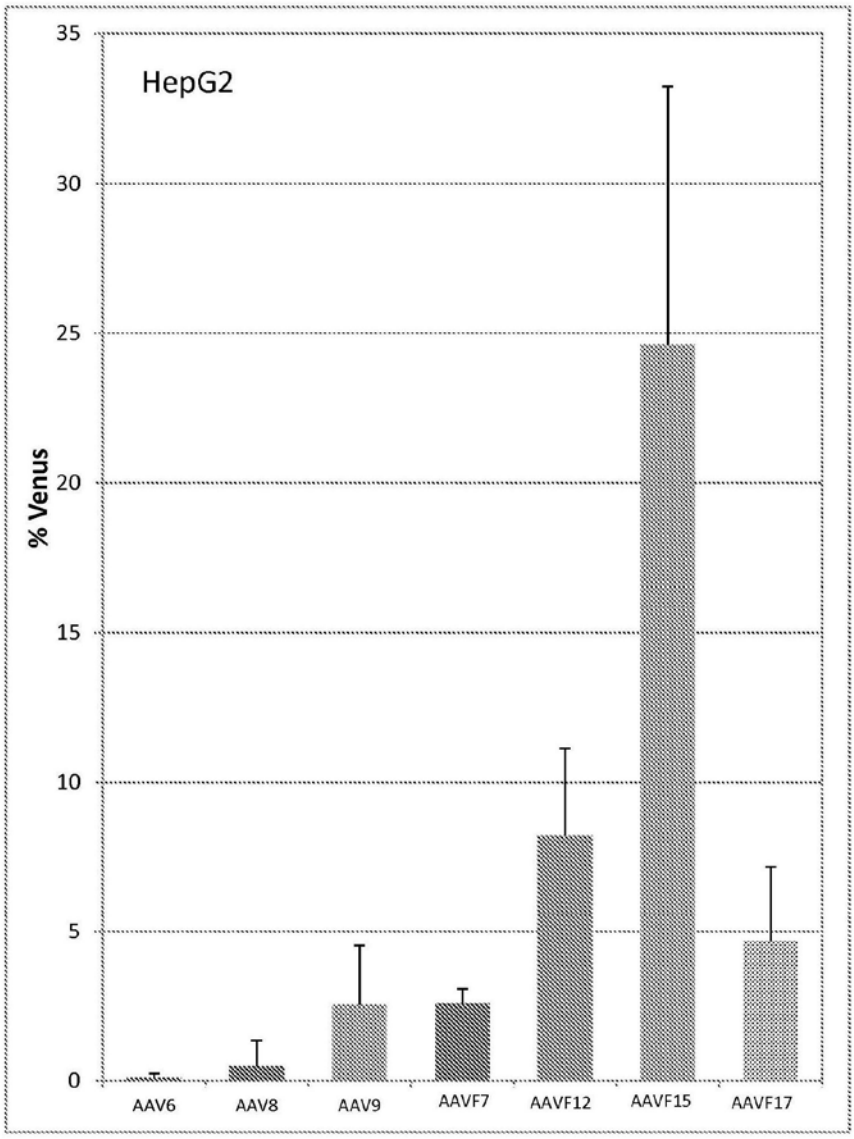
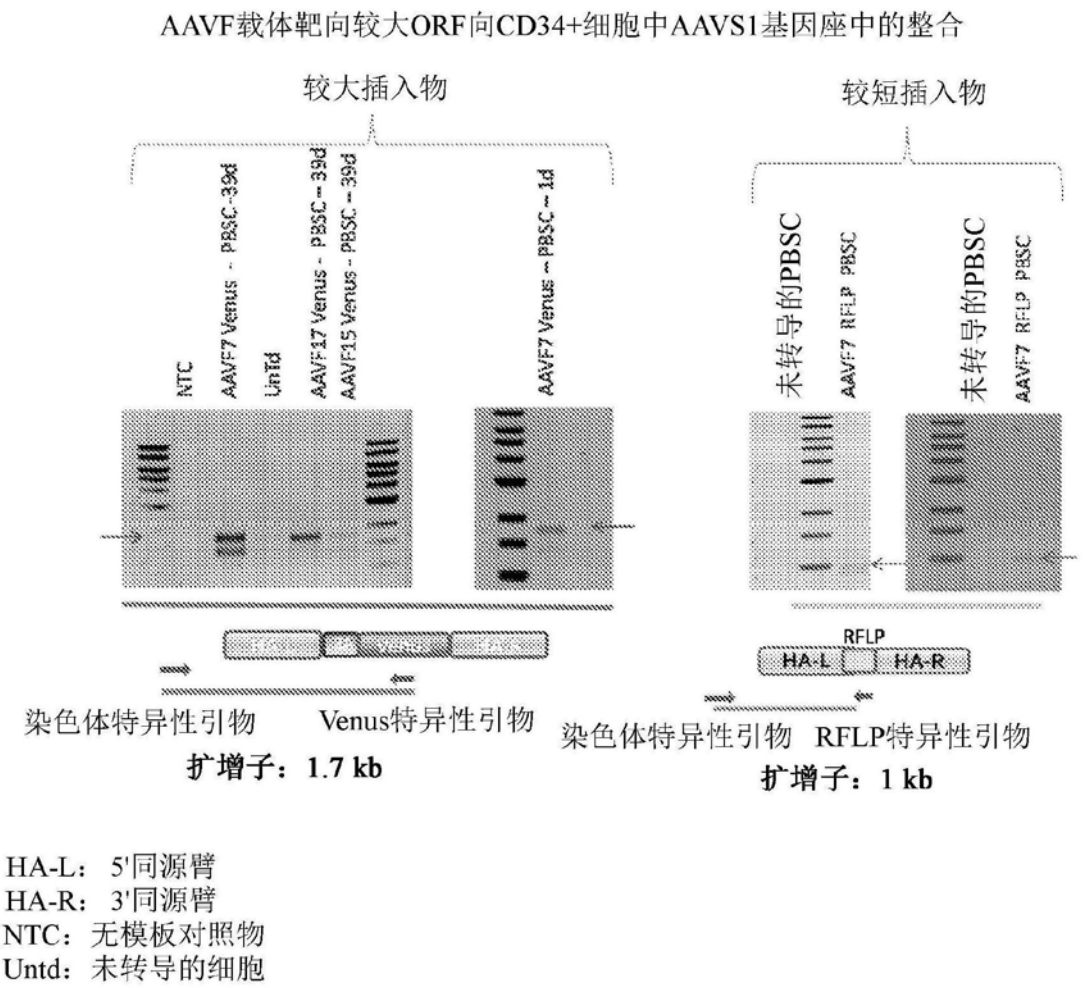

HepG2
% Venus
35
30
25
20
15
10
5
0
AAV6
AAV8
AAV9
AAVF7
AAVF12
AAVF15
AAVF17

图26F



HA-L： 5'同源臂
HA-R： 3'同源臂
NTC：无模板对照物
Untd：未转导的细胞

图27

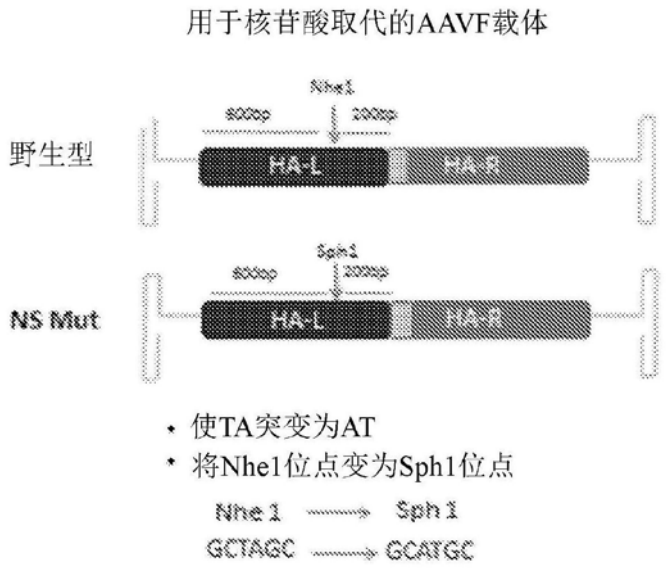


图28A

图28B

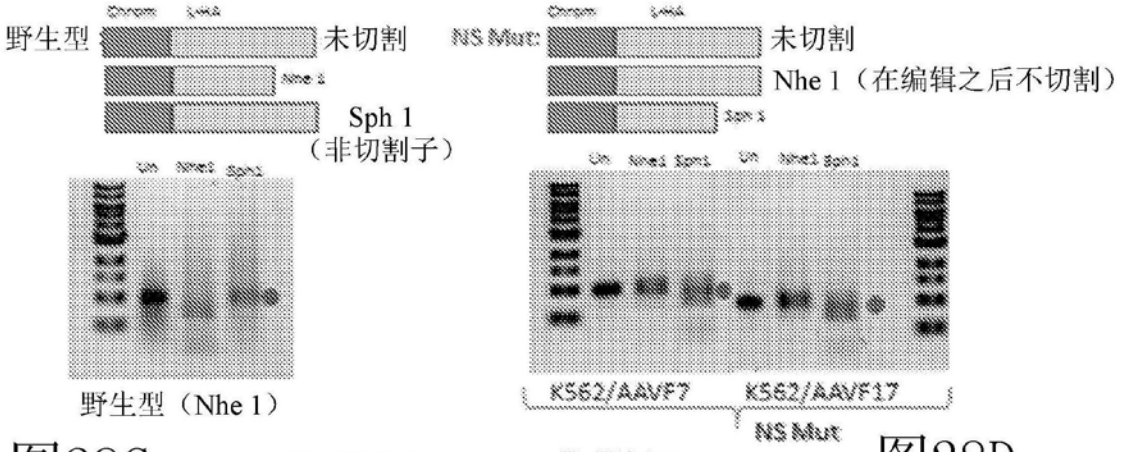


图28C

图28D

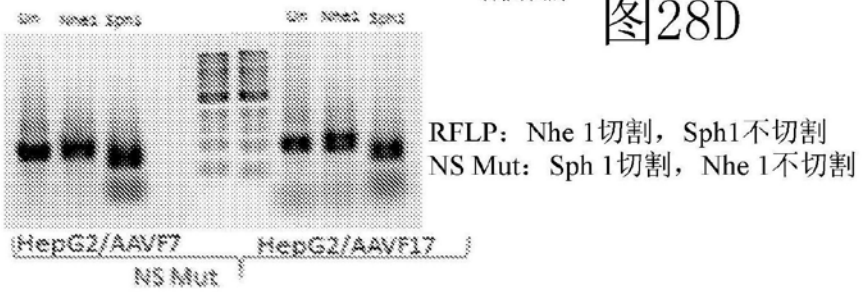


图28E

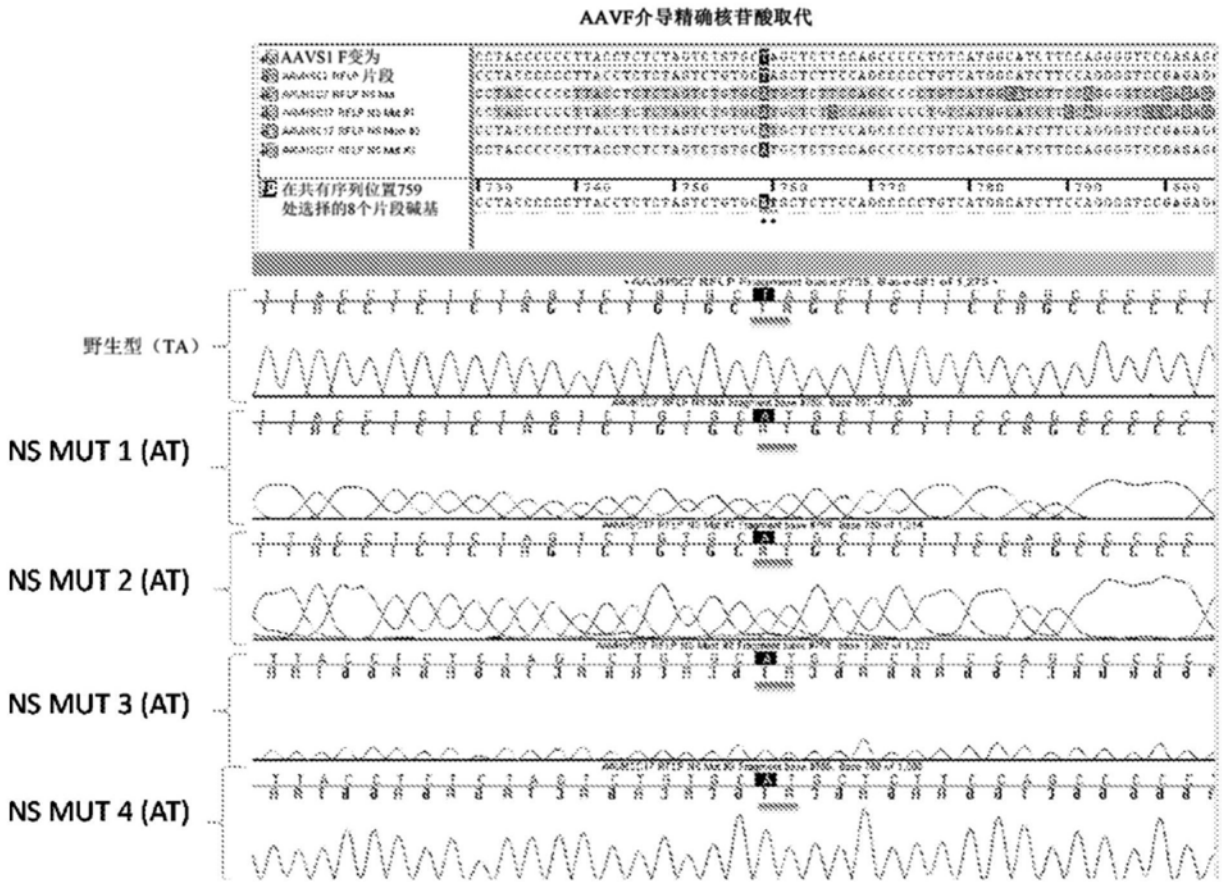


图29

	GFP+ BrdU- （非分裂细胞）	GFP+BrdU+ （分裂细胞）
AAVHSC7	**12**	14.6
AAVHSC17	**6.15**	6.9
AAV9	**4.31**	3.73
AAV8	**1.37**	0.839

图30

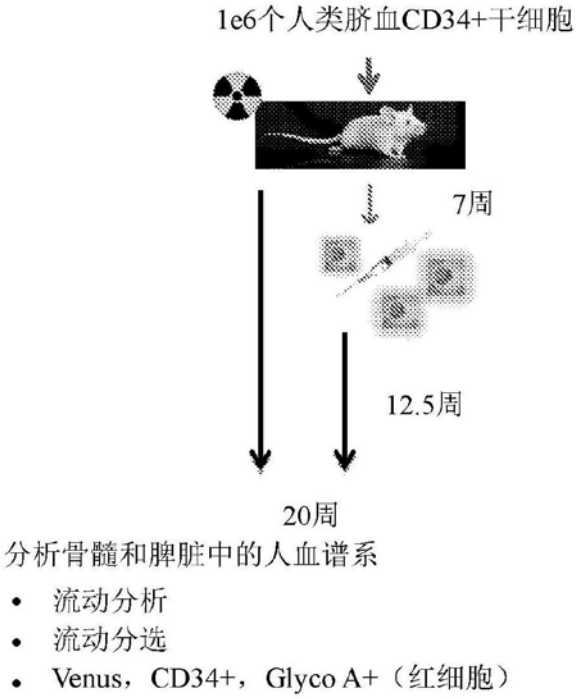

1e6个人类脐血CD34+干细胞
7周
12.5周
20周
分析骨髓和脾脏中的人血谱系
• 流动分析
• 流动分选
• Venus，CD34+，Glyco A+（红细胞）

图31A

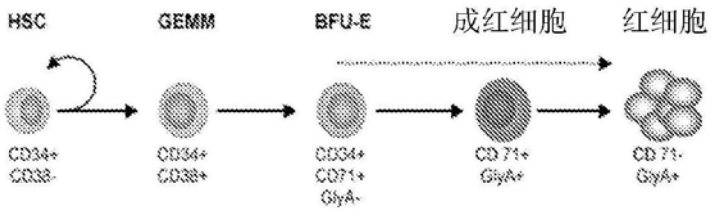

HSC
GEMM
BFU-E
成红细胞
红细胞
CD34+
CD38-
CD34+
CD38+
CD34+
CD71+
GlyA-
CD 71+
GlyA+
CD 71-
GlyA+

图31B

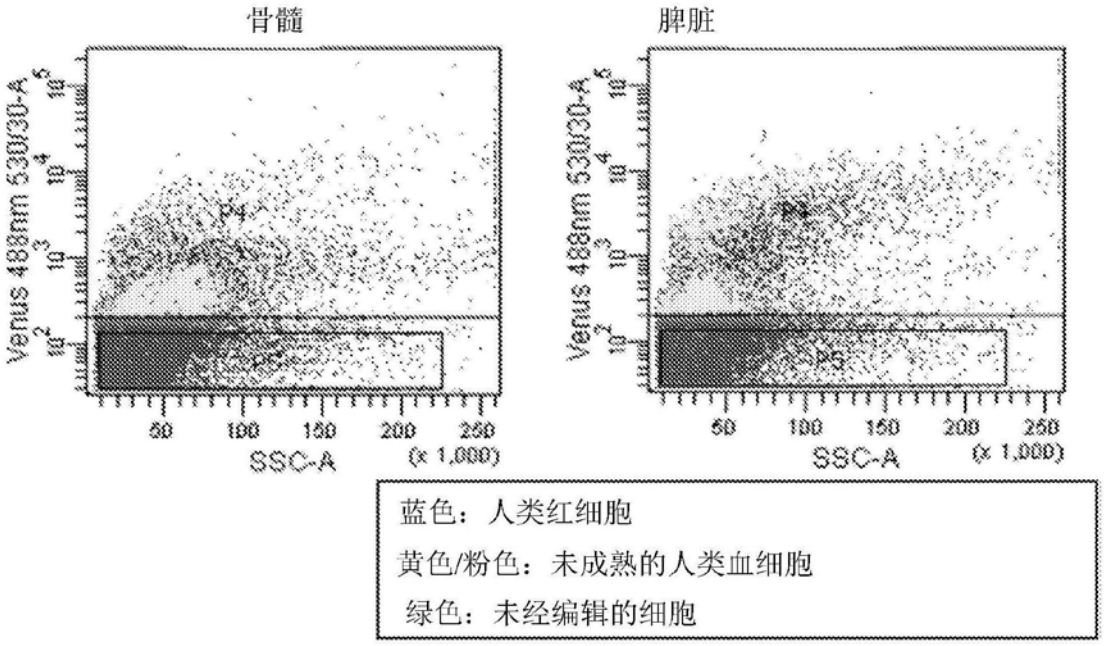

骨髓
脾脏
Venus 488nm 530/30-A
SSC-A
(x 1,000)
P4
P5
蓝色：人类红细胞
黄色/粉色：未成熟的人类血细胞
绿色：未经编辑的细胞

图31C

	实验	1A	1B	2A	2A	2B	2B	3	3
		AAVF7	AAVF17	AAVF7	AAVF7	AAVF17	AAVF17	AAVF13	AAVF17
		CD34+VEN+*	CD34+VEN+*	CD34+VEN+*	CD34+VEN+*	CD34+VEN+*	CD34+VEN+*	GATCO A+ VEN+*	CD34+VEN+*
TX变为 VENUS后的周数		1	1	1	1	1	1	7	7
VENUS变为HARVEST后的周数		5	5	6	6	6	6	12.5	12.5
TX变为 HARVEST后的周数		6	6	7	7	7	7	19	19
	VBN#	22.4	24.7	22.4	54.9	24.7	49.9	ND	ND
	细胞	11.2	9.49	9.49	20.9	11.1	16.8	99.4	79.9
	SPLN	15.9	26.8	35.9	71.3	24.8	82	99.2	98.1

图32A

Venus阳性%	中期	中期	中期	中期	长期	长期
	AAVF7	AAVF7	AAVF17	AAVF17	AAVF17	AAVF17
区室	CD34+	CD45+	CD34+	CD45+	CD34+	GlycoA+
脊椎骨髓	54.9	32.4	49.9	14.7	ND*	ND*
股骨骨髓	20.3	9.49	26.8	11.2	75.5	99.4
脾脏	71.1	15.9	82	24.8	58.1	99.2

* ND：未进行

图32B

展示VENUS通过AAVF载体而染色体插入到PPP1R12C内含子1中的序列分析

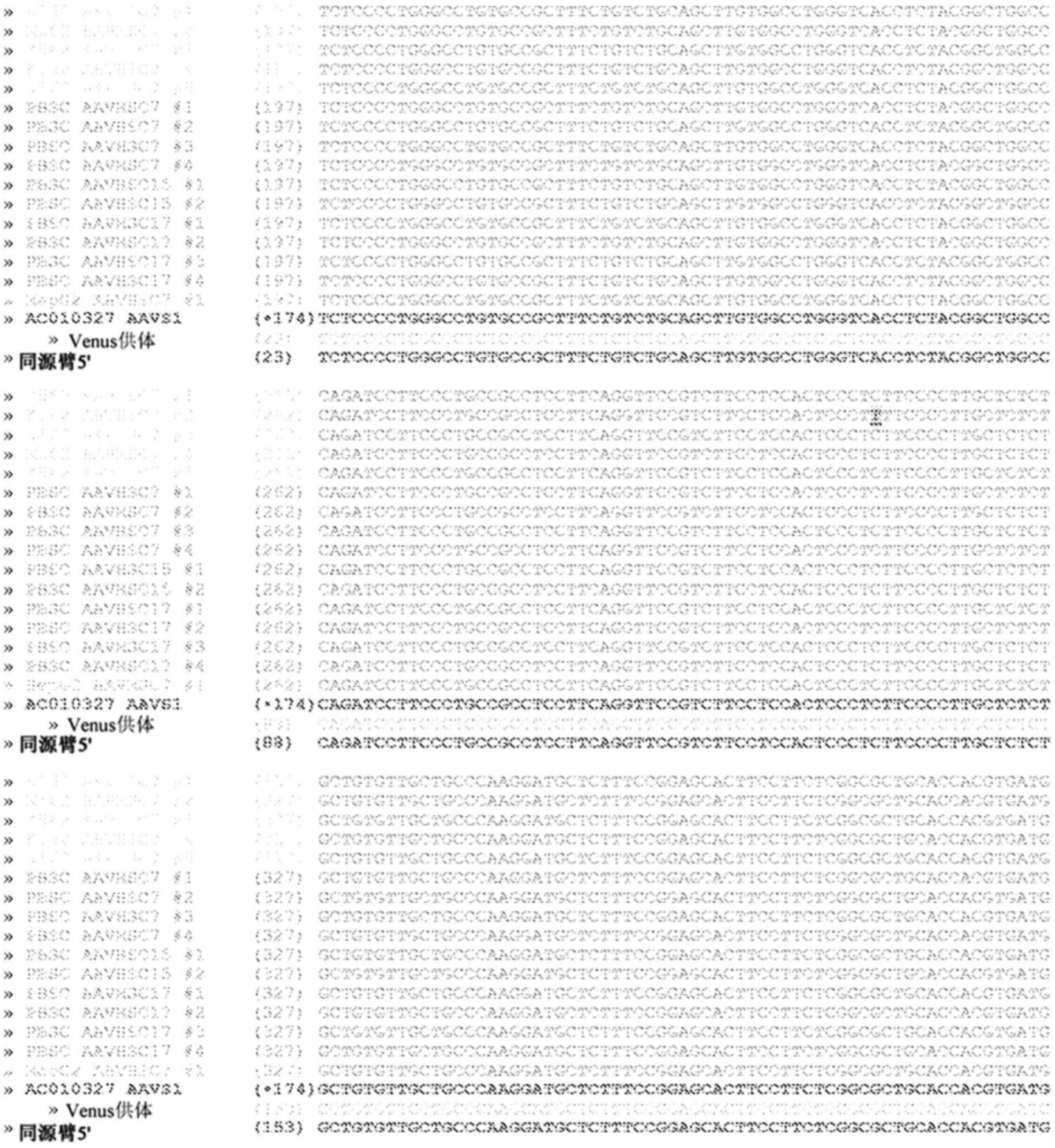
AAVS1中AAV靶向序列的比对

```
» K562 AAVESC7 #1      {2} GGCCCTGGCCATTGTCACTTTGCGCTGCCCTCCTCTCGCCCCCGAGTGCCCTTGCTGTGCCGCCG
» K562 AAVESC7 #2      {2} GGCCCTGGCCATTGTCACTTTGCGCTGCCCTCCTCTCGCCCCCGAGTGCCCTTGCTGTGCCGCCG
» K562 AAVESC7 #3      {2} GGCCCTGGCCATTGTCACTTTGCGCTGCCCTCCTCTCGCCCCCGAGTGCCCTTGCTGTGCCGCCG
» K562 AAVESC7 #4      {2} GGCCCTGGCCATTGTCACTTTGCGCTGCCCTCCTCTCGCCCCCGAGTGCCCTTGCTGTGCCGCCG
» K562 AAVESC7 #5      {2} GGCCCTGGCCATTGTCACTTTGCGCTGCCCTCCTCTCGCCCCCGAGTGCCCTTGCTGTGCCGCCG
» PBSC AAVESC7 #1      {2} GGCCCTGGCCATTGTCACTTTGCGCTGCCCTCCTCTCGCCCCCGAGTGCCCTTGCTGTGCCGCCG
» PBSC AAVESC7 #2      {2} GGCCCTGGCCATTGTCACTTTGCGCTGCCCTCCTCTCGCCCCCGAGTGCCCTTGCTGTGCCGCCG
» PBSC AAVESC7 #3      {2} GGCCCTGGCCATTGTCACTTTGCGCTGCCCTCCTCTCGCCCCCGAGTGCCCTTGCTGTGCCGCCG
» PBSC AAVESC7 #4      {2} GGCCCTGGCCATTGTCACTTTGCGCTGCCCTCCTCTCGCCCCCGAGTGCCCTTGCTGTGCCGCCG
» PBSC AAVESC15 #1     {2} GGCCCTGGCCATTGTCACTTTGCGCTGCCCTCCTCTCGCCCCCGAGTGCCCTTGCTGTGCCGCCG
» PBSC AAVESC15 #2     {2} GGCCCTGGCCATTGTCACTTTGCGCTGCCCTCCTCTCGCCCCCGAGTGCCCTTGCTGTGCCGCCG
» PBSC AAVESC17 #1     {2} GGCCCTGGCCATTGTCACTTTGCGCTGCCCTCCTCTCGCCCCCGAGTGCCCTTGCTGTGCCGCCG
» PBSC AAVESC17 #2     {2} GGCCCTGGCCATTGTCACTTTGCGCTGCCCTCCTCTCGCCCCCGAGTGCCCTTGCTGTGCCGCCG
» PBSC AAVESC17 #3     {2} GGCCCTGGCCATTGTCACTTTGCGCTGCCCTCCTCTCGCCCCCGAGTGCCCTTGCTGTGCCGCCG
» PBSC AAVESC17 #4     {2} GGCCCTGGCCATTGTCACTTTGCGCTGCCCTCCTCTCGCCCCCGAGTGCCCTTGCTGTGCCGCCG
  BspG2 AAVESC7 #1     {2} GGCCCTGGCCATTGTCACTTTGCGCTGCCCTCCTCTCGCCCCCGAGTGCCCTTGCTGTGCCGCCG
» AC010327 AAVS1       {2} GGCCCTGGCCATTGTCACTTTGCGCTGCCCTCCTCTCGCCCCCGAGTGCCCTTGCTGTGCCGCCG
    » Venus供体        {1}
» 同源臂5'             {1}

» K562 AAVESC7 #1      {67} GAACTCTGCCCTCTAACGCTGCCGTCTCTCTCCTGAGTCCGGACCACTTTGAGCTCTACTGGCTT
» K562 AAVESC7 #2      {67} GAACTCTGCCCTCTAACGCTGCCGTCTCTCTCCTGAGTCCGGACCACTTTGAGCTCTACTGGCTT
» K562 AAVESC7 #3      {67} GAACTCTGCCCTCTAACGCTGCCGTCTCTCTCCTGAGTCCGGACCACTTTGAGCTCTACTGGCTT
» K562 AAVESC7 #4      {67} GAACTCTGCCCTCTAACGCTGCCGTCTCTCTCCTGAGTCCGGACCACTTTGAGCTCTACTGGCTT
» K562 AAVESC7 #5      {67} GAACTCTGCCCTCTAACGCTGCCGTCTCTCTCCTGAGTCCGGACCACTTTGAGCTCTACTGGCTT
» PBSC AAVESC7 #1      {67} GAACTCTGCCCTCTAACGCTGCCGTCTCTCTCCTGAGTCCGGACCACTTTGAGCTCTACTGGCTT
» PBSC AAVESC7 #2      {67} GAACTCTGCCCTCTAACGCTGCCGTCTCTCTCCTGAGTCCGGACCACTTTGAGCTCTACTGGCTT
» PBSC AAVESC7 #3      {67} GAACTCTGCCCTCTAACGCTGCCGTCTCTCTCCTGAGTCCGGACCACTTTGAGCTCTACTGGCTT
» PBSC AAVESC7 #4      {67} GAACTCTGCCCTCTAACGCTGCCGTCTCTCTCCTGAGTCCGGACCACTTTGAGCTCTACTGGCTT
» PBSC AAVESC15 #1     {67} GAACTCTGCCCTCTAACGCTGCCGTCTCTCTCCTGAGTCCGGACCACTTTGAGCTCTACTGGCTT
» PBSC AAVESC15 #2     {67} GAACTCTGCCCTCTAACGCTGCCGTCTCTCTCCTGAGTCCGGACCACTTTGAGCTCTACTGGCTT
» PBSC AAVESC17 #1     {67} GAACTCTGCCCTCTAACGCTGCCGTCTCTCTCCTGAGTCCGGACCACTTTGAGCTCTACTGGCTT
» PBSC AAVESC17 #2     {67} GAACTCTGCCCTCTAACGCTGCCGTCTCTCTCCTGAGTCCGGACCACTTTGAGCTCTACTGGCTT
» PBSC AAVESC17 #3     {67} GAACTCTGCCCTCTAACGCTGCCGTCTCTCTCCTGAGTCCGGACCACTTTGAGCTCTACTGGCTT
» PBSC AAVESC17 #4     {67} GAACTCTGCCCTCTAACGCTGCCGTCTCTCTCCTGAGTCCGGACCACTTTGAGCTCTACTGGCTT
  BspG2 AAVESC7 #1     {67} GAACTCTGCCCTCTAACGCTGCCGTCTCTCTCCTGAGTCCGGACCACTTTGAGCTCTACTGGCTT
» AC010327 AAVS1       {67} GAACTCTGCCCTCTAACGCTGCCGTCTCTCTCCTGAGTCCGGACCACTTTGAGCTCTACTGGCTT
    » Venus供体        {1}
» 同源臂5'             {1}

» K562 AAVESC7 #1      {132} CTGCGCCGCCTCTGGCCCACTGTTTCCCCTTCCCAGGCAGGTCCTGCTTTCTCTGACCTGCATTC
» K562 AAVESC7 #2      {132} CTGCGCCGCCTCTGGCCCACTGTTTCCCCTTCCCAGGCAGGTCCTGCTTTCTCTGACCTGCATTC
» K562 AAVESC7 #3      {132} CTGCGCCGCCTCTGGCCCACTGTTTCCCCTTCCCAGGCAGGTCCTGCTTTCTCTGACCTGCATTC
» K562 AAVESC7 #4      {132} CTGCGCCGCCTCTGGCCCACTGTTTCCCCTTCCCAGGCAGGTCCTGCTTTCTCTGACCTGCATTC
» K562 AAVESC7 #5      {132} CTGCGCCGCCTCTGGCCCACTGTTTCCCCTTCCCAGGCAGGTCCTGCTTTCTCTGACCTGCATTC
» PBSC AAVESC7 #1      {132} CTGCGCCGCCTCTGGCCCACTGTTTCCCCTTCCCAGGCAGGTCCTGCTTTCTCTGACCTGCATTC
» PBSC AAVESC7 #2      {132} CTGCGCCGCCTCTGGCCCACTGTTTCCCCTTCCCAGGCAGGTCCTGCTTTCTCTGACCTGCATTC
» PBSC AAVESC7 #3      {132} CTGCGCCGCCTCTGGCCCACTGTTTCCCCTTCCCAGGCAGGTCCTGCTTTCTCTGACCTGCATTC
» PBSC AAVESC7 #4      {132} CTGCGCCGCCTCTGGCCCACTGTTTCCCCTTCCCAGGCAGGTCCTGCTTTCTCTGACCTGCATTC
» PBSC AAVESC15 #1     {132} CTGCGCCGCCTCTGGCCCACTGTTTCCCCTTCCCAGGCAGGTCCTGCTTTCTCTGACCTGCATTC
» PBSC AAVESC15 #2     {132} CTGCGCCGCCTCTGGCCCACTGTTTCCCCTTCCCAGGCAGGTCCTGCTTTCTCTGACCTGCATTC
» PBSC AAVESC17 #1     {132} CTGCGCCGCCTCTGGCCCACTGTTTCCCCTTCCCAGGCAGGTCCTGCTTTCTCTGACCTGCATTC
» PBSC AAVESC17 #2     {132} CTGCGCCGCCTCTGGCCCACTGTTTCCCCTTCCCAGGCAGGTCCTGCTTTCTCTGACCTGCATTC
» PBSC AAVESC17 #3     {132} CTGCGCCGCCTCTGGCCCACTGTTTCCCCTTCCCAGGCAGGTCCTGCTTTCTCTGACCTGCATTC
» PBSC AAVESC17 #4     {132} CTGCGCCGCCTCTGGCCCACTGTTTCCCCTTCCCAGGCAGGTCCTGCTTTCTCTGACCTGCATTC
  BspG2 AAVESC7 #1     {132} CTGCGCCGCCTCTGGCCCACTGTTTCCCCTTCCCAGGCAGGTCCTGCTTTCTCTGACCTGCATTC
» AC010327 AAVS1       {132} CTGCGCCGCCTCTGGCCCACTGTTTCCCCTTCCCAGGCAGGTCCTGCTTTCTCTGACCTGCATTC
    » Venus供体        {1}                                                   TGCTTTCTCTGACCTGCATTC
» 同源臂5'             {1}                                                   TGCTTTCTCTGACCTGCATTC
```

图33



```
» [illegible]              [illegible] TCTCCCCTGGGCCTGTGCCGCTTTCTGTCTGCAGCTTGTGGCCTGGGTCACCTCTACGGCTGGCC
» [illegible]              [illegible] TCTCCCCTGGGCCTGTGCCGCTTTCTGTCTGCAGCTTGTGGCCTGGGTCACCTCTACGGCTGGCC
» [illegible]              [illegible] TCTCCCCTGGGCCTGTGCCGCTTTCTGTCTGCAGCTTGTGGCCTGGGTCACCTCTACGGCTGGCC
» [illegible]              [illegible] TCTCCCCTGGGCCTGTGCCGCTTTCTGTCTGCAGCTTGTGGCCTGGGTCACCTCTACGGCTGGCC
» [illegible]              [illegible] TCTCCCCTGGGCCTGTGCCGCTTTCTGTCTGCAGCTTGTGGCCTGGGTCACCTCTACGGCTGGCC
» PBSC AAVHSC7 #1          (197)       TCTCCCCTGGGCCTGTGCCGCTTTCTGTCTGCAGCTTGTGGCCTGGGTCACCTCTACGGCTGGCC
» PBSC AAVHSC7 #2          (197)       TCTCCCCTGGGCCTGTGCCGCTTTCTGTCTGCAGCTTGTGGCCTGGGTCACCTCTACGGCTGGCC
» PBSC AAVHSC7 #3          (197)       TCTCCCCTGGGCCTGTGCCGCTTTCTGTCTGCAGCTTGTGGCCTGGGTCACCTCTACGGCTGGCC
» PBSC AAVHSC7 #4          (197)       TCTCCCCTGGGCCTGTGCCGCTTTCTGTCTGCAGCTTGTGGCCTGGGTCACCTCTACGGCTGGCC
» PBSC AAVHSC15 #1         (197)       TCTCCCCTGGGCCTGTGCCGCTTTCTGTCTGCAGCTTGTGGCCTGGGTCACCTCTACGGCTGGCC
» PBSC AAVHSC15 #2         (197)       TCTCCCCTGGGCCTGTGCCGCTTTCTGTCTGCAGCTTGTGGCCTGGGTCACCTCTACGGCTGGCC
» PBSC AAVHSC17 #1         (197)       TCTCCCCTGGGCCTGTGCCGCTTTCTGTCTGCAGCTTGTGGCCTGGGTCACCTCTACGGCTGGCC
» PBSC AAVHSC17 #2         (197)       TCTCCCCTGGGCCTGTGCCGCTTTCTGTCTGCAGCTTGTGGCCTGGGTCACCTCTACGGCTGGCC
» PBSC AAVHSC17 #3         (197)       TCTCCCCTGGGCCTGTGCCGCTTTCTGTCTGCAGCTTGTGGCCTGGGTCACCTCTACGGCTGGCC
» PBSC AAVHSC17 #4         (197)       TCTCCCCTGGGCCTGTGCCGCTTTCTGTCTGCAGCTTGTGGCCTGGGTCACCTCTACGGCTGGCC
» HepG2 AAVHSC7 #1         (197)       TCTCCCCTGGGCCTGTGCCGCTTTCTGTCTGCAGCTTGTGGCCTGGGTCACCTCTACGGCTGGCC
» AC010327 AAVS1           (•174)      TCTCCCCTGGGCCTGTGCCGCTTTCTGTCTGCAGCTTGTGGCCTGGGTCACCTCTACGGCTGGCC
   » Venus供体             [illegible] [illegible]
» 同源臂5'                 (23)        TCTCCCCTGGGCCTGTGCCGCTTTCTGTCTGCAGCTTGTGGCCTGGGTCACCTCTACGGCTGGCC

» [illegible]              [illegible] CAGATCCTTCCCTGCCGCCTCCTTCAGGTTCCGTCTTCCTCCACTCCCTCTTCCCCTTGCTCTCT
» [illegible]              [illegible] CAGATCCTTCCCTGCCGCCTCCTTCAGGTTCCGTCTTCCTCCACTCCCT[illegible]TTCCCCTTGCTCTCT
» [illegible]              [illegible] CAGATCCTTCCCTGCCGCCTCCTTCAGGTTCCGTCTTCCTCCACTCCCTCTTCCCCTTGCTCTCT
» [illegible]              [illegible] CAGATCCTTCCCTGCCGCCTCCTTCAGGTTCCGTCTTCCTCCACTCCCTCTTCCCCTTGCTCTCT
» [illegible]              [illegible] CAGATCCTTCCCTGCCGCCTCCTTCAGGTTCCGTCTTCCTCCACTCCCTCTTCCCCTTGCTCTCT
» PBSC AAVHSC7 #1          (262)       CAGATCCTTCCCTGCCGCCTCCTTCAGGTTCCGTCTTCCTCCACTCCCTCTTCCCCTTGCTCTCT
» PBSC AAVHSC7 #2          (262)       CAGATCCTTCCCTGCCGCCTCCTTCAGGTTCCGTCTTCCTCCACTCCCTCTTCCCCTTGCTCTCT
» PBSC AAVHSC7 #3          (262)       CAGATCCTTCCCTGCCGCCTCCTTCAGGTTCCGTCTTCCTCCACTCCCTCTTCCCCTTGCTCTCT
» PBSC AAVHSC7 #4          (262)       CAGATCCTTCCCTGCCGCCTCCTTCAGGTTCCGTCTTCCTCCACTCCCTCTTCCCCTTGCTCTCT
» PBSC AAVHSC15 #1         (262)       CAGATCCTTCCCTGCCGCCTCCTTCAGGTTCCGTCTTCCTCCACTCCCTCTTCCCCTTGCTCTCT
» PBSC AAVHSC15 #2         (262)       CAGATCCTTCCCTGCCGCCTCCTTCAGGTTCCGTCTTCCTCCACTCCCTCTTCCCCTTGCTCTCT
» PBSC AAVHSC17 #1         (262)       CAGATCCTTCCCTGCCGCCTCCTTCAGGTTCCGTCTTCCTCCACTCCCTCTTCCCCTTGCTCTCT
» PBSC AAVHSC17 #2         (262)       CAGATCCTTCCCTGCCGCCTCCTTCAGGTTCCGTCTTCCTCCACTCCCTCTTCCCCTTGCTCTCT
» PBSC AAVHSC17 #3         (262)       CAGATCCTTCCCTGCCGCCTCCTTCAGGTTCCGTCTTCCTCCACTCCCTCTTCCCCTTGCTCTCT
» PBSC AAVHSC17 #4         (262)       CAGATCCTTCCCTGCCGCCTCCTTCAGGTTCCGTCTTCCTCCACTCCCTCTTCCCCTTGCTCTCT
» HepG2 AAVHSC7 #1         (262)       CAGATCCTTCCCTGCCGCCTCCTTCAGGTTCCGTCTTCCTCCACTCCCTCTTCCCCTTGCTCTCT
» AC010327 AAVS1           (•174)      CAGATCCTTCCCTGCCGCCTCCTTCAGGTTCCGTCTTCCTCCACTCCCTCTTCCCCTTGCTCTCT
   » Venus供体             [illegible] [illegible]
» 同源臂5'                 (88)        CAGATCCTTCCCTGCCGCCTCCTTCAGGTTCCGTCTTCCTCCACTCCCTCTTCCCCTTGCTCTCT

» [illegible]              [illegible] GCTGTGTTGCTGCCCAAGGATGCTCTTTCCGGAGCACTTCCTTCTCGGCGCTGCACCACGTGATG
» [illegible]              [illegible] GCTGTGTTGCTGCCCAAGGATGCTCTTTCCGGAGCACTTCCTTCTCGGCGCTGCACCACGTGATG
» [illegible]              [illegible] GCTGTGTTGCTGCCCAAGGATGCTCTTTCCGGAGCACTTCCTTCTCGGCGCTGCACCACGTGATG
» [illegible]              [illegible] GCTGTGTTGCTGCCCAAGGATGCTCTTTCCGGAGCACTTCCTTCTCGGCGCTGCACCACGTGATG
» [illegible]              [illegible] GCTGTGTTGCTGCCCAAGGATGCTCTTTCCGGAGCACTTCCTTCTCGGCGCTGCACCACGTGATG
» PBSC AAVHSC7 #1          (327)       GCTGTGTTGCTGCCCAAGGATGCTCTTTCCGGAGCACTTCCTTCTCGGCGCTGCACCACGTGATG
» PBSC AAVHSC7 #2          (327)       GCTGTGTTGCTGCCCAAGGATGCTCTTTCCGGAGCACTTCCTTCTCGGCGCTGCACCACGTGATG
» PBSC AAVHSC7 #3          (327)       GCTGTGTTGCTGCCCAAGGATGCTCTTTCCGGAGCACTTCCTTCTCGGCGCTGCACCACGTGATG
» PBSC AAVHSC7 #4          (327)       GCTGTGTTGCTGCCCAAGGATGCTCTTTCCGGAGCACTTCCTTCTCGGCGCTGCACCACGTGATG
» PBSC AAVHSC15 #1         (327)       GCTGTGTTGCTGCCCAAGGATGCTCTTTCCGGAGCACTTCCTTCTCGGCGCTGCACCACGTGATG
» PBSC AAVHSC15 #2         (327)       GCTGTGTTGCTGCCCAAGGATGCTCTTTCCGGAGCACTTCCTTCTCGGCGCTGCACCACGTGATG
» PBSC AAVHSC17 #1         (327)       GCTGTGTTGCTGCCCAAGGATGCTCTTTCCGGAGCACTTCCTTCTCGGCGCTGCACCACGTGATG
» PBSC AAVHSC17 #2         (327)       GCTGTGTTGCTGCCCAAGGATGCTCTTTCCGGAGCACTTCCTTCTCGGCGCTGCACCACGTGATG
» PBSC AAVHSC17 #3         (327)       GCTGTGTTGCTGCCCAAGGATGCTCTTTCCGGAGCACTTCCTTCTCGGCGCTGCACCACGTGATG
» PBSC AAVHSC17 #4         (327)       GCTGTGTTGCTGCCCAAGGATGCTCTTTCCGGAGCACTTCCTTCTCGGCGCTGCACCACGTGATG
» HepG2 AAVHSC7 #1         (327)       GCTGTGTTGCTGCCCAAGGATGCTCTTTCCGGAGCACTTCCTTCTCGGCGCTGCACCACGTGATG
» AC010327 AAVS1           (•174)      GCTGTGTTGCTGCCCAAGGATGCTCTTTCCGGAGCACTTCCTTCTCGGCGCTGCACCACGTGATG
   » Venus供体             [illegible] [illegible]
» 同源臂5'                 (153)       GCTGTGTTGCTGCCCAAGGATGCTCTTTCCGGAGCACTTCCTTCTCGGCGCTGCACCACGTGATG
```

图33(续)

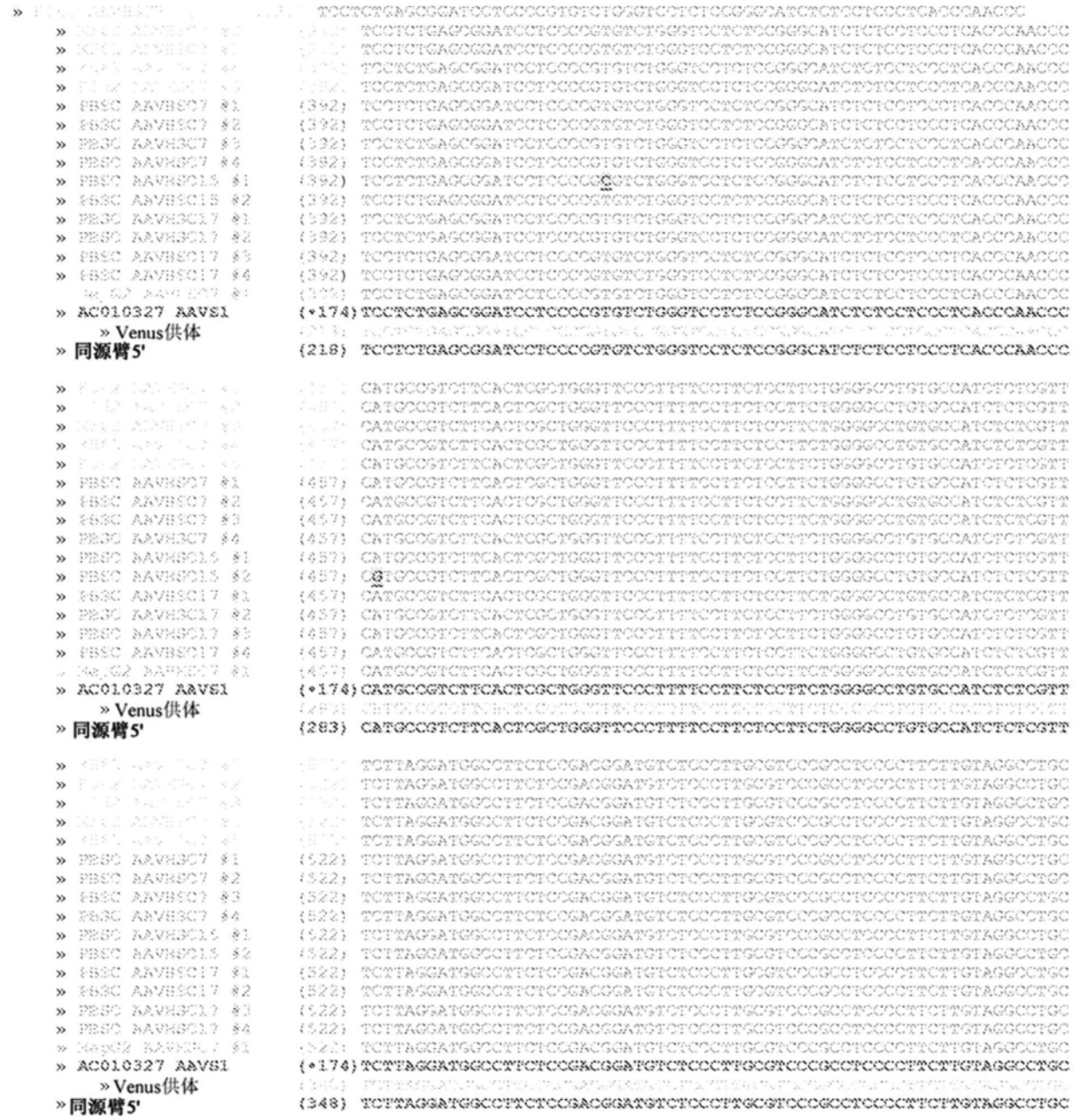

» AC010327 AAVS1 (+174) TCCTCTGAGCGGATCCTCCCCGTGTCTGGGTCCTCTCCGGGCATCTCTCCTCCCTCACCCAACCC
» Venus供体
» 同源臂5' (218) TCCTCTGAGCGGATCCTCCCCGTGTCTGGGTCCTCTCCGGGCATCTCTCCTCCCTCACCCAACCC
» AC010327 AAVS1 (+174) CATGCCGTCTTCACTCGCTGGGTTCCCTTTTCCTTCTCCTTCTGGGGCCTGTGCCATCTCTCGTT
» Venus供体
» 同源臂5' (283) CATGCCGTCTTCACTCGCTGGGTTCCCTTTTCCTTCTCCTTCTGGGGCCTGTGCCATCTCTCGTT
» AC010327 AAVS1 (+174) TCTTAGGATGGCCTTCTCCGACGGATGTCTCCCTTGCGTCCCGCCTCCCCTTCTTGTAGGCCTGC
» Venus供体
» 同源臂5' (348) TCTTAGGATGGCCTTCTCCGACGGATGTCTCCCTTGCGTCCCGCCTCCCCTTCTTGTAGGCCTGC

图33(续)

```
» PBSC AAVHSC7 #1      {587} ATCATCACCGTTTTTCTGGACAACCCCAAAGTACCCCGTCTCCCTGGCTTTAGCCACCTCTCCAT
» PBSC AAVHSC7 #2      {587} ATCATCACCGTTTTTCTGGACAACCCCAAAGTACCCCGTCTCCCTGGCTTTAGCCACCTCTCCAT
» PBSC AAVHSC7 #3      {587} ATCATCACCGTTTTTCTGGACAACCCCAAAGTACCCCGTCTCCCTGGCTTTAGCCACCTCTCCAT
» PBSC AAVHSC7 #4      {587} ATCATCACCGTTTTTCTGGACAACCCCAAAGTACCCCGTCTCCCTGGCTTTAGCCACCTCTCCAT
» PBSC AAVHSC15 #1     {587} ATCATCACCGTTTTTCTGGACAACCCCAAAGTACCCCGTCTCCCTGGCTTTAGCCACCTCTCCAT
» PBSC AAVHSC15 #2     {587} ATCATCACCGTTTTTCTGGACAACCCCAAAGTACCCCGTCTCCCTGGCTTTAGCCACCTCTCCAT
» PBSC AAVHSC17 #1     {587} ATCATCACCGTTTTTCTGGACAACCCCAAAGTACCCCGTCTCCCTGGCTTTAGCCACCTCTCCAT
» PBSC AAVHSC17 #2     {587} ATCATCACCGTTTTTCTGGACAACCCCAAAGTACCCCGTCTCCCTGGCTTTAGCCACCTCTCCAT
» PBSC AAVHSC17 #3     {587} ATCATCACCGTTTTTCTGGACAACCCCAAAGTACCCCGTCTCCCTGGCTTTAGCCACCTCTCCAT
» PBSC AAVHSC17 #4     {587} ATCATCACCGTTTTTCTGGACAACCCCAAAGTACCCCGTCTCCCTGGCTTTAGCCACCTCTCCAT
» PBSC AAVHSC7 #1      {587} ATCATCACCGTTTTTCTGGACAACCCCAAAGTACCCCGTCTCCCTGGCTTTAGCCACCTCTCCAT
» AC010327 AAVS1       {*174}ATCATCACCGTTTTTCTGGACAACCCCAAAGTACCCCGTCTCCCTGGCTTTAGCCACCTCTCCAT
   »Venus 供体         {413} ATCATCACCGTTTTTCTGGACAACCCCAAAGTACCCCGTCTCCCTGGCTTTAGCCACCTCTCCAT
» 同源臂 5′

» PBSC AAVHSC7 #1      {652} CCTCTTGCTTTCTTTGCCTGGACACCCCGTTCTCCTGTGGATTCGGGTCACCTCTCACTCCTTTC
» PBSC AAVHSC7 #2      {652} CCTCTTGCTTTCTTTGCCTGGACACCCCGTTCTCCTGTGGATTCGGGTCACCTCTCACTCCTTTC
» PBSC AAVHSC7 #3      {652} CCTCTTGCTTTCTTTGCCTGGACACCCCGTTCTCCTGTGGATTCGGGTCACCTCTCACTCCTTTC
» PBSC AAVHSC7 #4      {652} CCTCTTGCTTTCTTTGCCTGGACACCCCGTTCTCCTGTGGATTCGGGTCACCTCTCACTCCTTTC
» PBSC AAVHSC15 #1     {652} CCTCTTGCTTTCTTTGCCTGGACACCCCGTTCTCCTGTGGATTCGGGTCACCTCTCACTCCTTTC
» PBSC AAVHSC15 #2     {652} CCTCTTGCTTTCTTTGCCTGGACACCCCGTTCTCCTGTGGATTCGGGTCACCTCTCACTCCTTTC
» PBSC AAVHSC17 #1     {652} CCTCTTGCTTTCTTTGCCTGGACACCCCGTTCTCCTGTGGATTCGGGTCACCTCTCACTCCTTTC
» PBSC AAVHSC17 #2     {652} CCTCTTGCTTTCTTTGCCTGGACACCCCGTTCTCCTGTGGATTCGGGTCACCTCTCACTCCTTTC
» PBSC AAVHSC17 #3     {652} CCTCTTGCTTTCTTTGCCTGGACACCCCGTTCTCCTGTGGATTCGGGTCACCTCTCACTCCTTTC
» PBSC AAVHSC17 #4     {652} CCTCTTGCTTTCTTTGCCTGGACACCCCGTTCTCCTGTGGATTCGGGTCACCTCTCACTCCTTTC
» PBSC AAVHSC7 #1      {652} CCTCTTGCTTTCTTTGCCTGGACACCCCGTTCTCCTGTGGATTCGGGTCACCTCTCACTCCTTTC
» AC010327 AAVS1       {*174}CCTCTTGCTTTCTTTGCCTGGACACCCCGTTCTCCTGTGGATTCGGGTCACCTCTCACTCCTTTC
   » Venus 供体         {478} CCTCTTGCTTTCTTTGCCTGGACACCCCGTTCTCCTGTGGATTCGGGTCACCTCTCACTCCTTTC
» 同源臂 5′

» PBSC AAVHSC7 #1      {717} ATTTGGGCAGCTCCCCTACCCCCCTTACCTCTCTAGTCTGTGCTAGCTCTTCCAGCCCCCTGTCA
» PBSC AAVHSC7 #2      {717} ATTTGGGCAGCTCCCCTACCCCCCTTACCTCTCTAGTCTGTGCTAGCTCTTCCAGCCCCCTGTCA
» PBSC AAVHSC7 #3      {717} ATTTGGGCAGCTCCCCTACCCCCCTTACCTCTCTAGTCTGTGCTAGCTCTTCCAGCCCCCTGTCA
» PBSC AAVHSC7 #4      {717} ATTTGGGCAGCTCCCCTACCCCCCTTACCTCTCTAGTCTGTGCTAGCTCTTCCAGCCCCCTGTCA
» PBSC AAVHSC15 #1     {717} ATTTGGGCAGCTCCCCTACCCCCCTTACCTCTCTAGTCTGTGCTAGCTCTTCCAGCCCCCTGTCA
» PBSC AAVHSC15 #2     {717} ATTTGGGCAGCTCCCCTACCCCCCTTACCTCTCTAGTCTGTGCTAGCTCTTCCAGCCCCCTGTCA
» PBSC AAVHSC17 #1     {717} ATTTGGGCAGCTCCCCTACCCCCCTTACCTCTCTAGTCTGTGCTAGCTCTTCCAGCCCCCTGTCA
» PBSC AAVHSC17 #2     {717} ATTTGGGCAGCTCCCCTACCCCCCTTACCTCTCTAGTCTGTGCTAGCTCTTCCAGCCCCCTGTCA
» PBSC AAVHSC17 #3     {717} ATTTGGGCAGCTCCCCTACCCCCCTTACCTCTCTAGTCTGTGCTAGCTCTTCCAGCCCCCTGTCA
» PBSC AAVHSC17 #4     {717} ATTTGGGCAGCTCCCCTACCCCCCTTACCTCTCTAGTCTGTGCTAGCTCTTCCAGCCCCCTGTCA
» PBSC AAVHSC7 #1      {717} ATTTGGGCAGCTCCCCTACCCCCCTTACCTCTCTAGTCTGTGCTAGCTCTTCCAGCCCCCTGTCA
» AC010327 AAVS1       {*174}ATTTGGGCAGCTCCCCTACCCCCCTTACCTCTCTAGTCTGTGCTAGCTCTTCCAGCCCCCTGTCA
   »Venus 供体         {543} ATTTGGGCAGCTCCCCTACCCCCCTTACCTCTCTAGTCTGTGCTAGCTCTTCCAGCCCCCTGTCA
» 同源臂 5′
```

图33(续)

```
» PBSC AAVHSC7 #1      (782) TGGCATCTTCCAGGGGTCCGAGAGCTCAGCTAGTCTTCTTCCTCCAACCCGGGCCCCTATGTCCA
» PBSC AAVHSC7 #2      (782) TGGCATCTTCCAGGGGTCCGAGAGCTCAGCTAGTCTTCTTCCTCCAACCCGGGCCCCTATGTCCA
» PBSC AAVHSC7 #3      (782) TGGCATCTTCCAGGGGTCCGAGAGCTCAGCTAGTCTTCTTCCTCCAACCCGGGCCCCTATGTCCA
» PBSC AAVHSC7 #4      (782) TGGCATCTTCCAGGGGTCCGAGAGCTCAGCTAGTCTTCTTCCTCCAACCCGGGCCCCTATGTCCA
» PBSC AAVHSC15 #1     (782) TGGCATCTTCCAGGGGTCCGAGAGCTCAGCTAGTCTTCTTCCTCCAACCCGGGCCCCTATGTCCA
» PBSC AAVHSC15 #2     (782) TGGCATCTTCCAGGGGTCCGAGAGCTCAGCTAGTCTTCTTCCTCCAACCCGGGCCCCTATGTCCA
» PBSC AAVHSC17 #1     (782) TGGCATCTTCCAGGGGTCCGAGAGCTCAGCTAGTCTTCTTCCTCCAACCCGGGCCCCTATGTCCA
» PBSC AAVHSC17 #2     (782) TGGCATCTTCCAGGGGTCCGAGAGCTCAGCTAGTCTTCTTCCTCCAACCCGGGCCCCTATGTCCA
» PBSC AAVHSC17 #3     (782) TGGCATCTTCCAGGGGTCCGAGAGCTCAGCTAGTCTTCTTCCTCCAACCCGGGCCCCTATGTCCA
» PBSC AAVHSC17 #4     (782) TGGCATCTTCCAGGGGTCCGAGAGCTCAGCTAGTCTTCTTCCTCCAACCCGGGCCCCTATGTCCA
» HepG2 AAVHSC7 #1     (782) TGGCATCTTCCAGGGGTCCGAGAGCTCAGCTAGTCTTCTTCCTCCAACCCGGGCCCCTATGTCCA
» AC010327 AAVS1       (*174)TGGCATCTTCCAGGGGTCCGAGAGCTCAGCTAGTCTTCTTCCTCCAACCCGGGCCCCTATGTCCA
    » Venus供体
» 同源臂5'             (608) TGGCATCTTCCAGGGGTCCGAGAGCTCAGCTAGTCTTCTTCCTCCAACCCGGGCCCCTATGTCCA

» PBSC AAVHSC7 #1      (847) CTTCAGGACAGCATGTTTGCTGCCTCCAGGGATCCTGTGTCCCCGAGCTGGGACCACCTTATATT
» PBSC AAVHSC7 #2      (847) CTTCAGGACAGCATGTTTGCTGCCTCCAGGGATCCTGTGTCCCCGAGCTGGGACCACCTTATATT
» PBSC AAVHSC7 #3      (847) CTTCAGGACAGCATGTTTGCTGCCTCCAGGGATCCTGTGTCCCCGAGCTGGGACCACCTTATATT
» PBSC AAVHSC7 #4      (847) CTTCAGGACAGCATGTTTGCTGCCTCCAGGGATCCTGTGTCCCCGAGCTGGGACCACCTTATATT
» PBSC AAVHSC15 #1     (847) CTTCAGGACAGCATGTTTGCTGCCTCCAGGGATCCTGTGTCCCCGAGCTGGGACCACCTTATATT
» PBSC AAVHSC15 #2     (847) CTTCAGGACAGCATGTTTGCTGCCTCCAGGGATCCTGTGTCCCCGAGCTGGGACCACCTTATATT
» PBSC AAVHSC17 #1     (847) CTTCAGGACAGCATGTTTGCTGCCTCCAGGGATCCTGTGTCCCCGAGCTGGGACCACCTTATATT
» PBSC AAVHSC17 #2     (847) CTTCAGGACAGCATGTTTGCTGCCTCCAGGGATCCTGTGTCCCCGAGCTGGGACCACCTTATATT
» PBSC AAVHSC17 #3     (847) CTTCAGGACAGCATGTTTGCTGCCTCCAGGGATCCTGTGTCCCCGAGCTGGGACCACCTTATATT
» PBSC AAVHSC17 #4     (847) CTTCAGGACAGCATGTTTGCTGCCTCCAGGGATCCTGTGTCCCCGAGCTGGGACCACCTTATATT
» HepG2 AAVHSC7 #1     (847) CTTCAGGACAGCATGTTTGCTGCCTCCAGGGATCCTGTGTCCCCGAGCTGGGACCACCTTATATT
» AC010327 AAVS1       (*174)CTTCAGGACAGCATGTTTGCTGCCTCCAGGGATCCTGTGTCCCCGAGCTGGGACCACCTTATATT
    » Venus供体
» 同源臂5'             (673) CTTCAGGACAGCATGTTTGCTGCCTCCAGGGATCCTGTGTCCCCGAGCTGGGACCACCTTATATT

» PBSC AAVHSC7 #1      (912) CCCAGGGCCGGTTAATGTGGCTCTGGTTCTGGGTACTTTTATCTGTCCCCTCCACCCCACAGTGG
» PBSC AAVHSC7 #2      (912) CCCAGGGCCGGTTAATGTGGCTCTGGTTCTGGGTACTTTTATCTGTCCCCTCCACCCCACAGTGG
» PBSC AAVHSC7 #3      (912) CCCAGGGCCGGTTAATGTGGCTCTGGTTCTGGGTACTTTTATCTGTCCCCTCCACCCCACAGTGG
» PBSC AAVHSC7 #4      (912) CCCAGGGCCGGTTAATGTGGCTCTGGTTCTGGGTACTTTTATCTGTCCCCTCCACCCCACAGTGG
» PBSC AAVHSC15 #1     (912) CCCAGGGCCGGTTAATGTGGCTCTGGTTCTGGGTACTTTTATCTGTCCCCTCCACCCCACAGTGG
» PBSC AAVHSC15 #2     (912) CCCAGGGCCGGTTAATGTGGCTCTGGTTCTGGGTACTTTTATCTGTCCCCTCCACCCCACAGTGG
» PBSC AAVHSC17 #1     (912) CCCAGGGCCGGTTAATGTGGCTCTGGTTCTGGGTACTTTTATCTGTCCCCTCCACCCCACAGTGG
» PBSC AAVHSC17 #2     (912) CCCAGGGCCGGTTAATGTGGCTCTGGTTCTGGGTACTTTTATCTGTCCCCTCCACCCCACAGTGG
» PBSC AAVHSC17 #3     (912) CCCAGGGCCGGTTAATGTGGCTCTGGTTCTGGGTACTTTTATCTGTCCCCTCCACCCCACAGTGG
» PBSC AAVHSC17 #4     (912) CCCAGGGCCGGTTAATGTGGCTCTGGTTCTGGGTACTTTTATCTGTCCCCTCCACCCCACAGTGG
» HepG2 AAVHSC7 #1     (912) CCCAGGGCCGGTTAATGTGGCTCTGGTTCTGGGTACTTTTATCTGTCCCCTCCACCCCACAGTGG
» AC010327 AAVS1       (*174)CCCAGGGCCGGTTAATGTGGCTCTGGTTCTGGGTACTTTTATCTGTCCCCTCCACCCCACAGTGG
    » Venus供体
» 同源臂5'             (738) CCCAGGGCCGGTTAATGTGGCTCTGGTTCTGGGTACTTTTATCTGTCCCCTCCACCCCACAGTGG
```

图33(续)

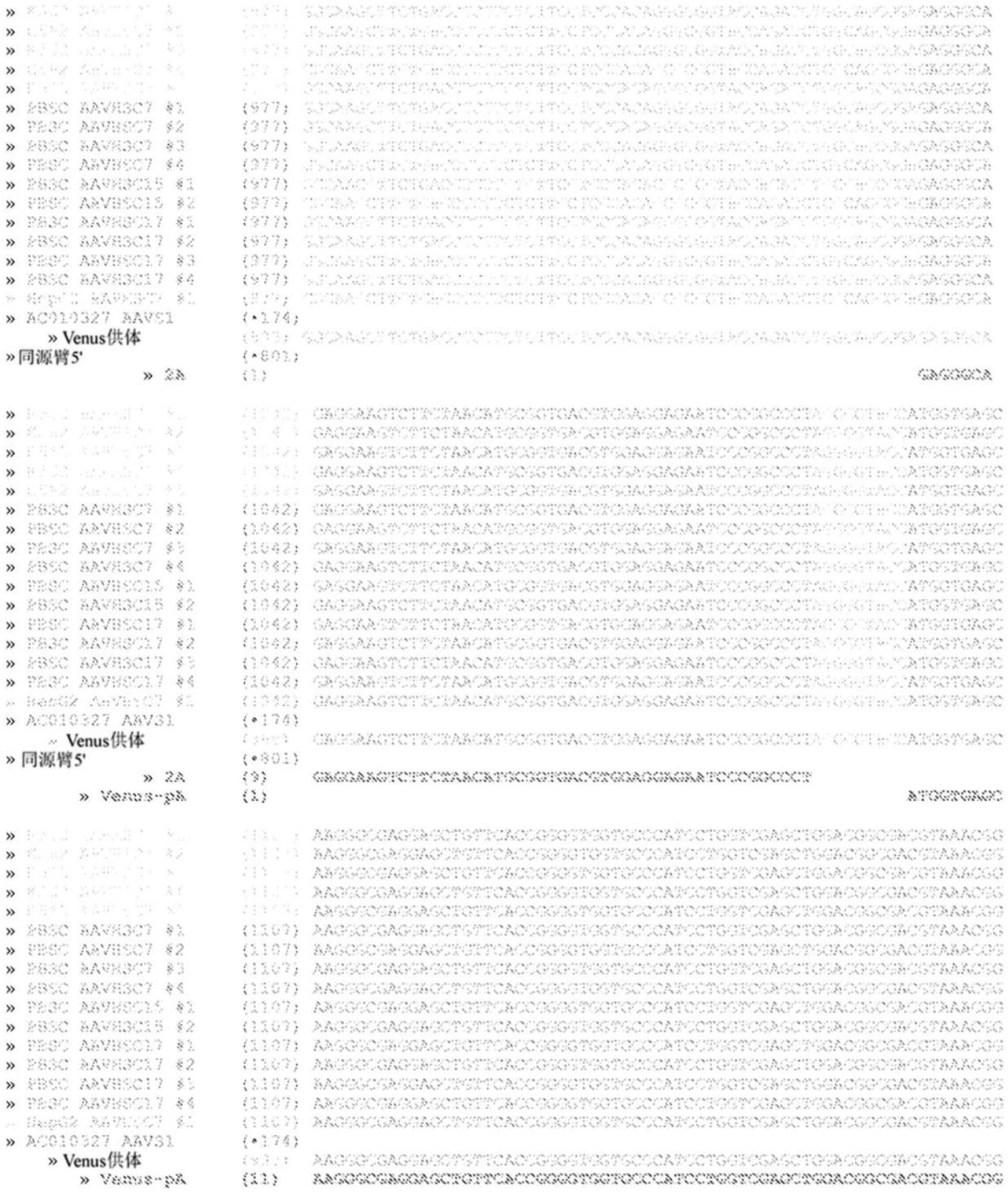

» PB3C AAVR3C7 #1 {977}
» PB3C AAVR3C7 #2 {977}
» PB3C AAVR3C7 #3 {977}
» PB3C AAVR3C7 #4 {977}
» PB3C AAVR3C15 #1 {977}
» PB3C AAVR3C15 #2 {977}
» PB3C AAVR3C17 #1 {977}
» PB3C AAVR3C17 #2 {977}
» PB3C AAVR3C17 #3 {977}
» PB3C AAVR3C17 #4 {977}
» AC010327 AAV31 {•174}
» Venus供体
» 同源臂5' {•801}
» 2A {1}
» PB3C AAVR3C7 #1 {1042}
» PB3C AAVR3C7 #2 {1042}
» PB3C AAVR3C7 #3 {1042}
» PB3C AAVR3C7 #4 {1042}
» PB3C AAVR3C15 #1 {1042}
» PB3C AAVR3C15 #2 {1042}
» PB3C AAVR3C17 #1 {1042}
» PB3C AAVR3C17 #2 {1042}
» PB3C AAVR3C17 #3 {1042}
» PB3C AAVR3C17 #4 {1042}
» AC010327 AAV31 {•174}
» Venus供体
» 同源臂5' {•801}
» 2A {9}
» Venus-pA {1}
» PB3C AAVR3C7 #1 {1107}
» PB3C AAVR3C7 #2 {1107}
» PB3C AAVR3C7 #3 {1107}
» PB3C AAVR3C7 #4 {1107}
» PB3C AAVR3C15 #1 {1107}
» PB3C AAVR3C15 #2 {1107}
» PB3C AAVR3C17 #1 {1107}
» PB3C AAVR3C17 #2 {1107}
» PB3C AAVR3C17 #3 {1107}
» PB3C AAVR3C17 #4 {1107}
» AC010327 AAV31 {•174}
» Venus供体
» Venus-pA {11}

图33(续)

```
» PBSC AAVHSC7 #1      (1172) CCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGC
» PBSC AAVHSC7 #2      (1172) CCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGC
» PBSC AAVHSC7 #3      (1172) CCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGC
» PBSC AAVHSC7 #4      (1172) CCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGC
» PBSC AAVHSC15 #1     (1172) CCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGC
» PBSC AAVHSC15 #2     (1172) CCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGC
» PBSC AAVHSC17 #1     (1172) CCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGC
» PBSC AAVHSC17 #2     (1172) CCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGC
» PBSC AAVHSC17 #3     (1172) CCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGC
» PBSC AAVHSC17 #4     (1172) CCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGC
» AC010327 AAVS1       (*174)
  » Venus供体
» 同源臂5'             (*601)
    » Venus-pA         (76)   CCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGC

» PBSC AAVHSC7 #1      (1237) TGATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGGGCTACGGC
» PBSC AAVHSC7 #2      (1237) TGATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGGGCTACGGC
» PBSC AAVHSC7 #3      (1237) TGATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGGGCTACGGC
» PBSC AAVHSC7 #4      (1237) TGATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGGGCTACGGC
» PBSC AAVHSC15 #1     (1237) TGATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGGGCTACGGC
» PBSC AAVHSC15 #2     (1237) TGATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGGGCTACGGC
» PBSC AAVHSC17 #1     (1237) TGATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGGGCTACGGC
» PBSC AAVHSC17 #2     (1237) TGATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGGGCTACGGC
» PBSC AAVHSC17 #3     (1237) TGATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGGGCTACGGC
» PBSC AAVHSC17 #4     (1237) TGATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGGGCTACGGC
» AC010327 AAVS1       (*174)
  » Venus供体
» 同源臂5'             (*601)
    » Venus-pA         (141)  TGATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGGGCTACGGC

» PBSC AAVHSC7 #1      (1302) CTGCAGTGCTTCGCCCGCTACCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCC
» PBSC AAVHSC7 #2      (1302) CTGCAGTGCTTCGCCCGCTACCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCC
» PBSC AAVHSC7 #3      (1302) CTGCAGTGCTTCGCCCGCTACCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCC
» PBSC AAVHSC7 #4      (1302) CTGCAGTGCTTCGCCCGCTACCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCC
» PBSC AAVHSC15 #1     (1302) CTGCAGTGCTTCGCCCGCTACCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCC
» PBSC AAVHSC15 #2     (1302) CTGCAGTGCTTCGCCCGCTACCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCC
» PBSC AAVHSC17 #1     (1302) CTGCAGTGCTTCGCCCGCTACCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCC
» PBSC AAVHSC17 #2     (1302) CTGCAGTGCTTCGCCCGCTACCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCC
» PBSC AAVHSC17 #3     (1302) CTGCAGTGCTTCGCCCGCTACCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCC
» PBSC AAVHSC17 #4     (1302) CTGCAGTGCTTCGCCCGCTACCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCC
» AC010327 AAVS1       (*174)
  » Venus供体
» 同源臂5'             (*601)
    » Venus-pA         (206)  CTGCAGTGCTTCGCCCGCTACCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCC
```

图33(续)

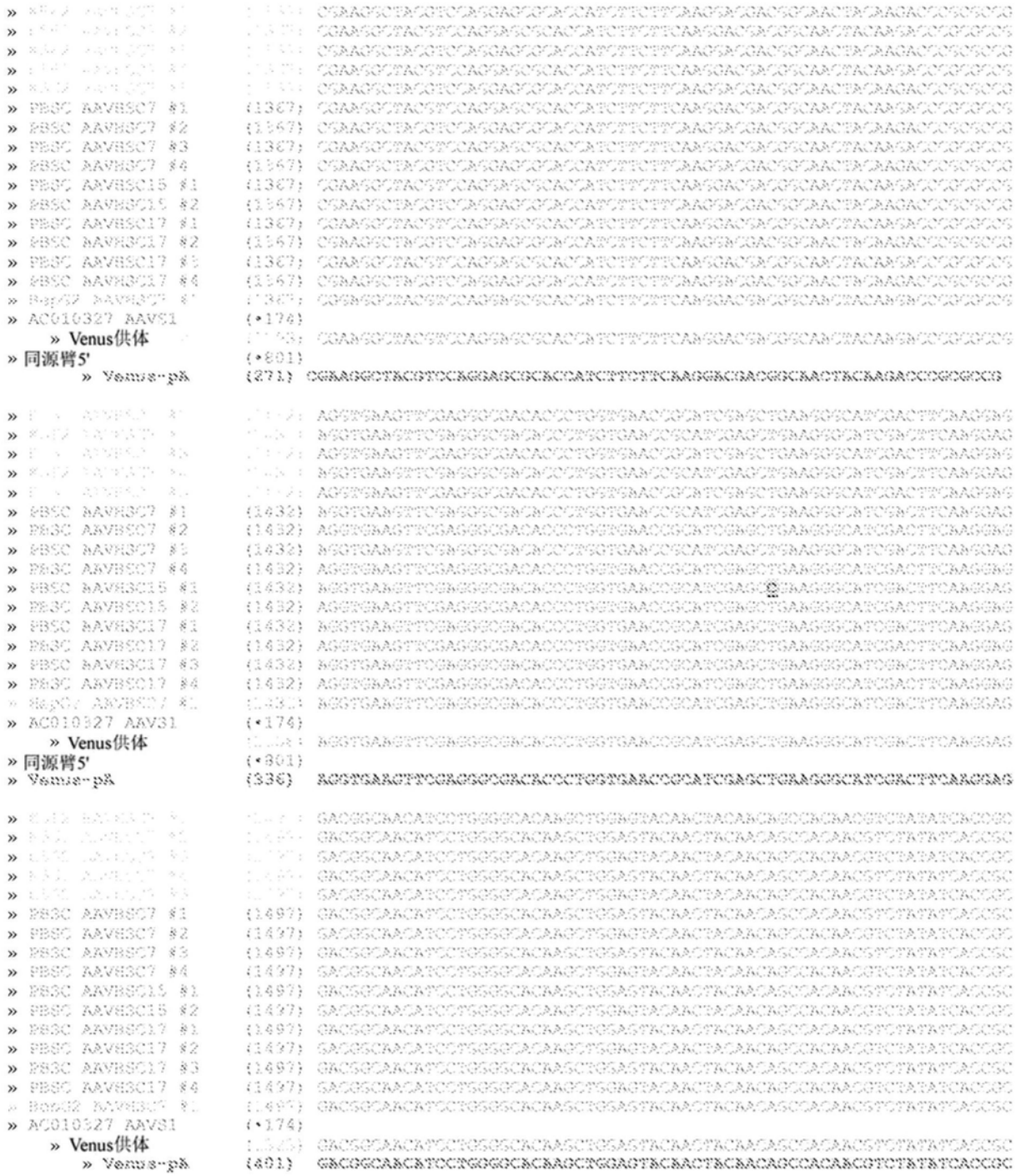

» PB3C AAVHSC7 #1 (1367)
» PB3C AAVHSC7 #2 (1367)
» PB3C AAVHSC7 #3 (1367)
» PB3C AAVHSC7 #4 (1367)
» PB3C AAVHSC15 #1 (1367)
» PB3C AAVHSC15 #2 (1367)
» PB3C AAVHSC17 #1 (1367)
» PB3C AAVHSC17 #2 (1367)
» PB3C AAVHSC17 #3 (1367)
» PB3C AAVHSC17 #4 (1367)
» AC010327 AAV31 (+174)
» Venus供体
» 同源臂5' (+801)
» Venus-pA (271)
» PB3C AAVHSC7 #1 (1432)
» PB3C AAVHSC7 #2 (1432)
» PB3C AAVHSC7 #3 (1432)
» PB3C AAVHSC7 #4 (1432)
» PB3C AAVHSC15 #1 (1432)
» PB3C AAVHSC15 #2 (1432)
» PB3C AAVHSC17 #1 (1432)
» PB3C AAVHSC17 #2 (1432)
» PB3C AAVHSC17 #3 (1432)
» PB3C AAVHSC17 #4 (1432)
» AC010327 AAV31 (+174)
» Venus供体
» 同源臂5' (+801)
» Venus-pA (336)
» PB3C AAVHSC7 #1 (1497)
» PB3C AAVHSC7 #2 (1497)
» PB3C AAVHSC7 #3 (1497)
» PB3C AAVHSC7 #4 (1497)
» PB3C AAVHSC15 #1 (1497)
» PB3C AAVHSC15 #2 (1497)
» PB3C AAVHSC17 #1 (1497)
» PB3C AAVHSC17 #2 (1497)
» PB3C AAVHSC17 #3 (1497)
» PB3C AAVHSC17 #4 (1497)
» AC010327 AAV31 (+174)
» Venus供体
» Venus-pA (401)

图33(续)

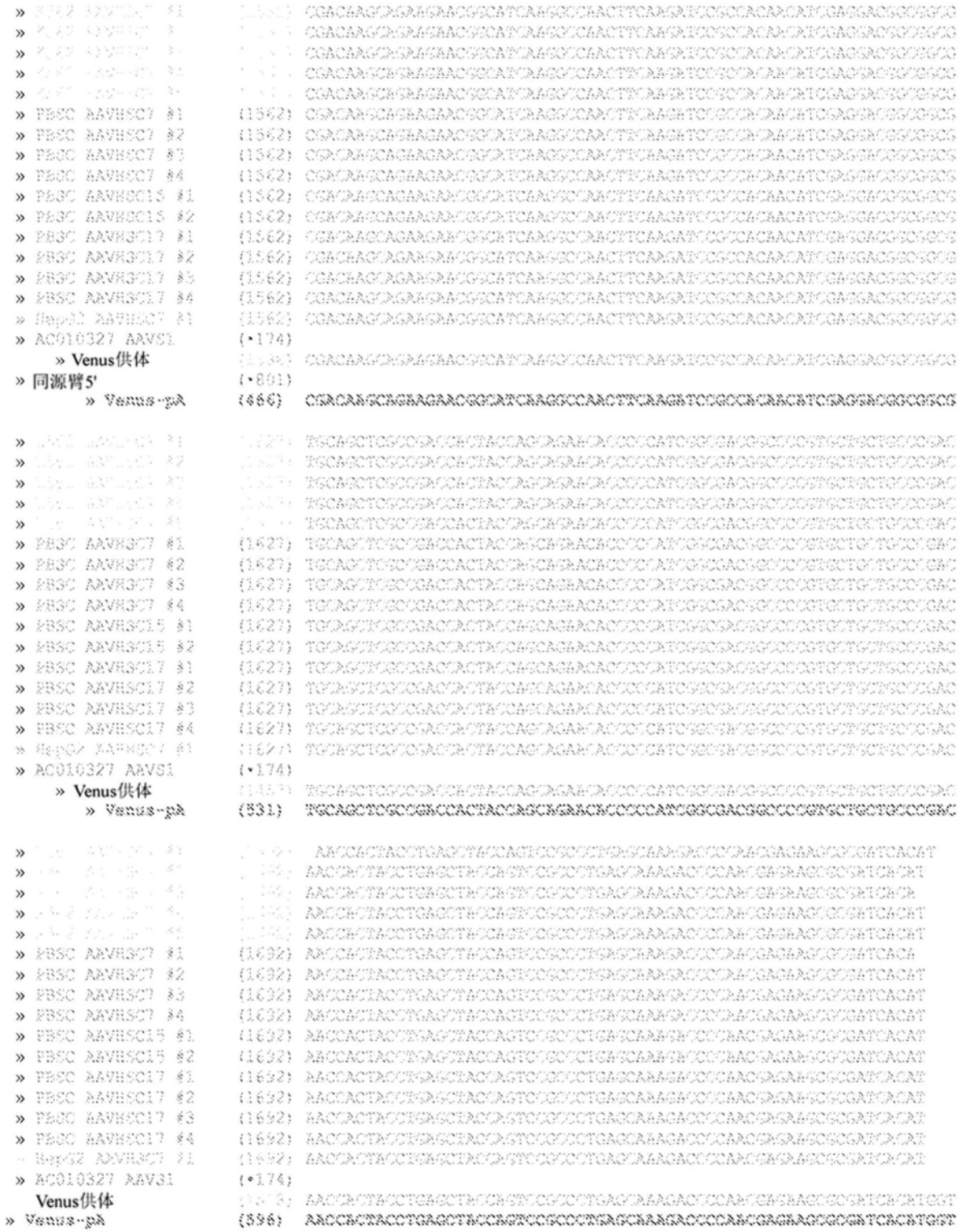

» PBSC AAVRSC7 #1 (1562)
» PBSC AAVRSC7 #2 (1562)
» PBSC AAVRSC7 #3 (1562)
» PBSC AAVRSC7 #4 (1562)
» PBSC AAVRSC15 #1 (1562)
» PBSC AAVRSC15 #2 (1562)
» PBSC AAVRSC17 #1 (1562)
» PBSC AAVRSC17 #2 (1562)
» PBSC AAVRSC17 #3 (1562)
» PBSC AAVRSC17 #4 (1562)
» AC010327 AAVS1 (•174)
» Venus供体
» 同源臂5'
» Venus-pA (466)
» PBSC AAVRSC7 #1 (1627)
» PBSC AAVRSC7 #2 (1627)
» PBSC AAVRSC7 #3 (1627)
» PBSC AAVRSC7 #4 (1627)
» PBSC AAVRSC15 #1 (1627)
» PBSC AAVRSC15 #2 (1627)
» PBSC AAVRSC17 #1 (1627)
» PBSC AAVRSC17 #2 (1627)
» PBSC AAVRSC17 #3 (1627)
» PBSC AAVRSC17 #4 (1627)
» AC010327 AAVS1 (•174)
» Venus供体
» Venus-pA (531)
» PBSC AAVRSC7 #1 (1692)
» PBSC AAVRSC7 #2 (1692)
» PBSC AAVRSC7 #3 (1692)
» PBSC AAVRSC7 #4 (1692)
» PBSC AAVRSC15 #1 (1692)
» PBSC AAVRSC15 #2 (1692)
» PBSC AAVRSC17 #1 (1692)
» PBSC AAVRSC17 #2 (1692)
» PBSC AAVRSC17 #3 (1692)
» PBSC AAVRSC17 #4 (1692)
» AC010327 AAVS1 (•174)
Venus供体
» Venus-pA (596)

图33(续)

展示较短序列通过AAVF载体而染色体插入到
PPP1R12C内含子1中的序列分析

```
                                    1                                                                   65
» AAVHSC7 RFLP PBSC #1      (2)  GGCCCTGGCCATTGTCACTTTGCGCTGCCCTCCTCTCGCCCCCGAGTGCCCTTGCTGTGCCGCCG
» AAVHSC7 RFLP PBSC #2      (2)  GGCCCTGGCCATTGTCACTTTGCGCTGCCCTCCTCTCGCCCCCGAGTGCCCTTGCTGTGCCGCCG
» AAVHSC7 RFLP HepG2 #1     (2)  GGCCCTGGCCATTGTCACTTTGCGCTGCCCTCCTCTCGCCCCCGAGTGCCCTTGCTGTGCCGCCG
» AAVHSC7 RFLP HepG2 #2     (2)  GGCCCTGGCCATTGTCACTTTGCGCTGCCCTCCTCTCGCCCCCGAGTGCCCTTGCTGTGCCGCCG
» AAVHSC7 RFLP HepG2 #3     (2)  GGCCCTGGCCATTGTCACTTTGCGCTGCCCTCCTCTCGCCCCCGAGTGCCCTTGCTGTGCCGCCG
» AAVHSC7 RFLP HepG2 #4     (2)  GGCCCTGGCCATTGTCACTTTGCGCTGCCCTCCTCTCGCCCCCGAGTGCCCTTGCTGTGCCGCCG
» RFLP载体序列              (1)
» AC010327 AAVS1             (2)  GGCCCTGGCCATTGTCACTTTGCGCTGCCCTCCTCTCGCCCCCGAGTGCCCTTGCTGTGCCGCCG
» 同源臂5'                  (1)
» 同源臂3'                  (1)
                                    66                                                                  130
» AAVHSC7 RFLP PBSC #1      (67) GAACTCTGCCCTCTAACGCTGCCGTCTCTCTCCTGAGTCCGGACCACTTTGAGCTCTACTGGCTT
» AAVHSC7 RFLP PBSC #2      (67) GAACTCTGCCCTCTAACGCTGCCGTCTCTCTCCTGAGTCCGGACCACTTTGAGCTCTACTGGCTT
» AAVHSC7 RFLP HepG2 #1     (67) GAACTCTGCCCTCTAACGCTGCCGTCTCTCTCCTGAGTCCGGACCACTTTGAGCTCTACTGGCTT
» AAVHSC7 RFLP HepG2 #2     (67) GAACTCTGCCCTCTAACGCTGCCGTCTCTCTCCTGAGTCCGGACCACTTTGAGCTCTACTGGCTT
» AAVHSC7 RFLP HepG2 #3     (67) GAACTCTGCCCTCTAACGCTGCCGTCTCTCTCCTGAGTCCGGACCACTTTGAGCTCTACTGGCTT
» AAVHSC7 RFLP HepG2 #4     (67) GAACTCTGCCCTCTAACGCTGCCGTCTCTCTCCTGAGTCCGGACCACTTTGAGCTCTACTGGCTT
» RFLP载体序列              (1)
» AC010327 AAVS1             (67) GAACTCTGCCCTCTAACGCTGCCGTCTCTCTCCTGAGTCCGGACCACTTTGAGCTCTACTGGCTT
» 同源臂5'                  (1)
» 同源臂3'                  (1)
                                    131                                                                 195
» AAVHSC7 RFLP PBSC #1      (132) CTGCGCCGCCTCTGGCCCACTGTTTCCCCTTCCCAGGCAGGTCCTGCTTTCTCTGACCTGCATTC
» AAVHSC7 RFLP PBSC #2      (132) CTGCGCCGCCTCTGGCCCACTGTTTCCCCTTCCCAGGCAGGTCCTGCTTTCTCTGACCTGCATTC
» AAVHSC7 RFLP HepG2 #1     (132) CTGCGCCGCCTCTGGCCCACTGTTTCCCCTTCCCAGGCAGGTCCTGCTTTCTCTGACCTGCATTC
» AAVHSC7 RFLP HepG2 #2     (132) CTGCGCCGCCTCTGGCCCACTGTTTCCCCTTCCCAGGCAGGTCCTGCTTTCTCTGACCTGCATTC
» AAVHSC7 RFLP HepG2 #3     (132) CTGCGCCGCCTCTGGCCCACTGTTTCCCCTTCCCAGGCAGGTCCTGCTTTCTCTGACCTGCATTC
» AAVHSC7 RFLP HepG2 #4     (132) CTGCGCCGCCTCTGGCCCACTGTTTCCCCTTCCCAGGCAGGTCCTGCTTTCTCTGACCTGCATTC
» RFLP载体序列              (1)                                                TGCTTTCTCTGACCAGCATTC
» AC010327 AAVS1             (132) CTGCGCCGCCTCTGGCCCACTGTTTCCCCTTCCCAGGCAGGTCC
» 同源臂5'                  (1)                                                TGCTTTCTCTGACCAGCATTC
» 同源臂3'                  (1)
                                    196                                                                 260
» AAVHSC7 RFLP PBSC #1      (197) TCTCCCCTGGGCCTGTGCCGCTTTCTGTCTGCAGCTTGTGGCCTGGGTCACCTCTACGGCTGGCC
» AAVHSC7 RFLP PBSC #2      (197) TCTCCCCTGGGCCTGTGCCGCTTTCTGTCTGCAGCTTGTGGCCTGGGTCACCTCTACGGCTGGCC
» AAVHSC7 RFLP HepG2 #1     (197) TCTCCCCTGGGCCTGTGCCGCTTTCTGTCTGCAGCTTGTGGCCTGGGTCACCTCTACGGCTGGCC
» AAVHSC7 RFLP HepG2 #2     (197) TCTCCCCTGGGCCTGTGCCGCTTTCTGTCTGCAGCTTGTGGCCTGGGTCACCTCTACGGCTGGCC
» AAVHSC7 RFLP HepG2 #3     (197) TCTCCCCTGGGCCTGTGCCGCTTTCTGTCTGCAGCTTGTGGCCTGGGTCACCTCTACGGCTGGCC
» AAVHSC7 RFLP HepG2 #4     (197) TCTCCCCTGGGCCTGTGCCGCTTTCTGTCTGCAGCTTGTGGCCTGGGTCACCTCTACGGCTGGCC
» RFLP载体序列              (23)  TCTCCCCTGGGCCTGTGCCGCTTTCTGTCTGCAGCTTGTGGCCTGGGTCACCTCTACGGCTGGCC
» AC010327 AAVS1            (•174)
» 同源臂5'                  (23) TCTCCCCTGGGCCTGTGCCGCTTTCTGTCTGCAGCTTGTGGCCTGGGTCACCTCTACGGCTGGCC
» 同源臂3'                  (1)
```

图34

```
                                    261                                                                                 325
» AAVHSC7 RFLP PBSC #1        (262) CAGATCCTTCCCTGCCGCCTCCTTCAGGTTCCGTCTTCCTCCACTCCCTCTTCCCCTTGCTCTCT
» AAVHSC7 RFLP PBSC #2        (262) CAGATCCTTCCCTGCCGCCTCCTTCAGGTTCCGTCTTCCTCCACTCCCTCTTCCCCTTGCTCTCT
» AAVHSC7 RFLP HepG2 #1       (262) CAGATCCTTCCCTGCCGCCTCCTTCAGGTTCCGTCTTCCTCCACTCCCTCTTCCCCTTGCTCTCT
» AAVHSC7 RFLP HepG2 #2       (262) CAGATCCTTCCCTGCCGCCTCCTTCAGGTTCCGTCTTCCTCCACTCCCTCTTCCCCTTGCTCTCT
» AAVHSC7 RFLP HepG2 #3       (262) CAGATCCTTCCCTGCCGCCTCCTTCAGGTTCCGTCTTCCTCCACTCCCTCTTCCCCTTGCTCTCT
» AAVHSC7 RFLP HepG2 #4       (262) CAGATCCTTCCCTGCCGCCTCCTTCAGGTTCCGTCTTCCTCCACTCCCTCTTCCCCTTGCTCTCT
» RFLP载体序列                (88)  CAGATCCTTCCCTGCCGCCTCCTTCAGGTTCCGTCTTCCTCCACTCCCTCTTCCCCTTGCTCTCT
» AC010327 AAVS1              (•174)
» 同源臂5'                    (88)  CAGATCCTTCCCTGCCGCCTCCTTCAGGTTCCGTCTTCCTCCACTCCCTCTTCCCCTTGCTCTCT
» 同源臂3'                    (1)
                                    326                                                                                 390
» AAVHSC7 RFLP PBSC #1        (327) GCTGTGTTGCTGCCCAAGGATGCTCTTTCCGGAGCACTTCCTTCTCGGCGCTGCACCACGTGATG
» AAVHSC7 RFLP PBSC #2        (327) GCTGTGTTGCTGCCCAAGGATGCTCTTTCCGGAGCACTTCCTTCTCGGCGCTGCACCACGTGATG
» AAVHSC7 RFLP HepG2 #1       (327) GCTGTGTTGCTGCCCAAGGATGCTCTTTCCGGAGCACTTCCTTCTCGGCGCTGCACCACGTGATG
» AAVHSC7 RFLP HepG2 #2       (327) GCTGTGTTGCTGCCCAAGGATGCTCTTTCCGGAGCACTTCCTTCTCGGCGCTGCACCACGTGATG
» AAVHSC7 RFLP HepG2 #3       (327) GCTGTGTTGCTGCCCAAGGATGCTCTTTCCGGAGCACTTCCTTCTCGGCGCTGCACCACGTGATG
» AAVHSC7 RFLP HepG2 #4       (327) GCTGTGTTGCTGCCCAAGGATGCTCTTTCCGGAGCACTTCCTTCTCGGCGCTGCACCACGTGATG
» RFLP载体序列                (153) GCTGTGTTGCTGCCCAAGGATGCTCTTTCCGGAGCACTTCCTTCTCGGCGCTGCACCACGTGATG
» AC010327 AAVS1              (•174)
» 同源臂5'                    (153) GCTGTGTTGCTGCCCAAGGATGCTCTTTCCGGAGCACTTCCTTCTCGGCGCTGCACCACGTGATG
» 同源臂3'                    (1)
                                    391                                                                                 455
» AAVHSC7 RFLP PBSC #1        (392) TCCTCTGAGCGGATCCTCCCCGTGTCTGGGTCCTCTCCGGGCATCTCTCCTCCCTCACCCAACCC
» AAVHSC7 RFLP PBSC #2        (392) TCCTCTGAGCGGATCCTCCCCGTGTCTGGGTCCTCTCCGGGCATCTCTCCTCCCTCACCCAACCC
» AAVHSC7 RFLP HepG2 #1       (392) TCCTCTGAGCGGATCCTCCCCGTGTCTGGGTCCTCTCCGGGCATCTCTCCTCCCTCACCCAACCC
» AAVHSC7 RFLP HepG2 #2       (392) TCCTCTGAGCGGATCCTCCCCGTGTCTGGGTCCTCTCCGGGCATCTCTCCTCCCTCACCCAACCC
» AAVHSC7 RFLP HepG2 #3       (392) TCCTCTGAGCGGATCCTCCCCGTGTCTGGGTCCTCTCCGGGCATCTCTCCTCCCTCACCCAACCC
» AAVHSC7 RFLP HepG2 #4       (392) TCCTCTGAGCGGATCCTCCCCGTGTCTGGGTCCTCTCCGGGCATCTCTCCTCCCTCACCCAACCC
» RFLP载体序列                (218) TCCTCTGAGCGGATCCTCCCCGTGTCTGGGTCCTCTCCGGGCATCTCTCCTCCCTCACCCAACCC
» AC010327 AAVS1              (•174)
» 同源臂5'                    (218) TCCTCTGAGCGGATCCTCCCCGTGTCTGGGTCCTCTCCGGGCATCTCTCCTCCCTCACCCAACCC
» 同源臂3'                    (1)
                                    456                                                                                 520
» AAVHSC7 RFLP PBSC #1        (457) CATGCCGTCTTCACTCGCTGGGTTCCCTTTTCCTTCTCCTTCTGGGGCCTGTGCCATCTCTCGTT
» AAVHSC7 RFLP PBSC #2        (457) CATGCCGTCTTCACTCGCTGGGTTCCCTTTTCCTTCTCCTTCTGGGGCCTGTGCCATCTCTCGTT
» AAVHSC7 RFLP HepG2 #1       (457) CATGCCGTCTTCACTCGCTGGGTTCCCTTTTCCTTCTCCTTCTGGGGCCTGTGCCATCTCTCGTT
» AAVHSC7 RFLP HepG2 #2       (457) CATGCCGTCTTCACTCGCTGGGTTCCCTTTTCCTTCTCCTTCTGGGGCCTGTGCCATCTCTCGTT
» AAVHSC7 RFLP HepG2 #3       (457) CATGCCGTCTTCACTCGCTGGGTTCCCTTTTCCTTCTCCTTCTGGGGCCTGTGCCATCTCTCGTT
» AAVHSC7 RFLP HepG2 #4       (457) CATGCCGTCTTCACTCGCTGGGTTCCCTTTTCCTTCTCCTTCTGGGGCCTGTGCCATCTCTCGTT
» RFLP载体序列                (283) CATGCCGTCTTCACTCGCTGGGTTCCCTTTTCCTTCTCCTTCTGGGGCCTGTGCCATCTCTCGTT
» AC010327 AAVS1              (•174)
» 同源臂5'                    (283) CATGCCGTCTTCACTCGCTGGGTTCCCTTTTCCTTCTCCTTCTGGGGCCTGTGCCATCTCTCGTT
» 同源臂3'                    (1)
```

图34(续)

```
                                      521                                                             585
» AAVHSC7 RFLP PBSC #1        (522) TCTTAGGATGGCCTTCTCCGACGGATGTCTCCCTTGCGTCCCGCCTCCCCTTCTTGTAGGCCTGC
» AAVHSC7 RFLP PBSC #2        (522) TCTTAGGATGGCCTTCTCCGACGGATGTCTCCCTTGCGTCCCGCCTCCCCTTCTTGTAGGCCTGC
» AAVHSC7 RFLP HepG2 #1       (522) TCTTAGGATGGCCTTCTCCGACGGATGTCTCCCTTGCGTCCCGCCTCCCCTTCTTGTAGGCCTGC
» AAVHSC7 RFLP HepG2 #2       (522) TCTTAGGATGGCCTTCTCCGACGGATGTCTCCCTTGCGTCCCGCCTCCCCTTCTTGTAGGCCTGC
» AAVHSC7 RFLP HepG2 #3       (522) TCTTAGGATGGCCTTCTCCGACGGATGTCTCCCTTGCGTCCCGCCTCCCCTTCTTGTAGGCCTGC
» AAVHSC7 RFLP HepG2 #4       (522) TCTTAGGATGGCCTTCTCCGACGGATGTCTCCCTTGCGTCCCGCCTCCCCTTCTTGTAGGCCTGC
» RFLP载体序列                (348) TCTTAGGATGGCCTTCTCCGACGGATGTCTCCCTTGCGTCCCGCCTCCCCTTCTTGTAGGCCTGC
» AC010327 AAVS1              (•174)
» 同源臂5'                    (348) TCTTAGGATGGCCTTCTCCGACGGATGTCTCCCTTGCGTCCCGCCTCCCCTTCTTGTAGGCCTGC
» 同源臂3'                    (1)
                                      586                                                             650
» AAVHSC7 RFLP PBSC #1        (587) ATCATCACCGTTTTTCTGGACAACCCCAAAGTACCCCGTCTCCCTGGCTTTAGCCACCTCTCCAT
» AAVHSC7 RFLP PBSC #2        (587) ATCATCACCGTTTTTCTGGACAACCCCAAAGTACCCCGTCTCCCTGGCTTTAGCCACCTCTCCAT
» AAVHSC7 RFLP HepG2 #1       (587) ATCATCACCGTTTTTCTGGACAACCCCAAAGTACCCCGTCTCCCTGGCTTTAGCCACCTCTCCAT
» AAVHSC7 RFLP HepG2 #2       (587) ATCATCACCGTTTTTCTGGACAACCCCAAAGTACCCCGTCTCCCTGGCTTTAGCCACCTCTCCAT
» AAVHSC7 RFLP HepG2 #3       (587) ATCATCACCGTTTTTCTGGACAACCCCAAAGTACCCCGTCTCCCCGGCTTTAGCCACCTCTCCAT
» AAVHSC7 RFLP HepG2 #4       (587) ATCATCACCGTTTTTCTGGACAACCCCAAAGTACCCCGTCTCCCTGGCTTTAGCCACCTCTCCAT
» RFLP载体序列                (413) ATCATCACCGTTTTTCTGGACAACCCCAAAGTACCCCGTCTCCCTGGCTTTAGCCACCTCTCCAT
» AC010327 AAVS1              (•174)
» 同源臂5'                    (413) ATCATCACCGTTTTTCTGGACAACCCCAAAGTACCCCGTCTCCCTGGCTTTAGCCACCTCTCCAT
» 同源臂3'                    (1)
                                      651                                                             715
» AAVHSC7 RFLP PBSC #1        (652) CCTCTTGCTTTCTTTGCCTGGACACCCCGTTCTCCTGTGGATTCGGGTCACCTCTCACTCCTTTC
» AAVHSC7 RFLP PBSC #2        (652) CCTCTTGCTTTCTTTGCCTGGACACCCCGTTCTCCTGTGGATTCGGGTCACCTCTCACTCCTTTC
» AAVHSC7 RFLP HepG2 #1       (652) CCTCTTGCTTTCTTTGCCTGGACACCCCGTTCTCCTGTGGATTCGGGTCACCCCTCACTCCTTTC
» AAVHSC7 RFLP HepG2 #2       (652) CCTCTTGCTTTCTTTGCCTGGACACCCCGTTCTCCTGTGGATTCGGGTCACCTCTCACTCCTTTC
» AAVHSC7 RFLP HepG2 #3       (652) CCTCTTGCTTTCTTTGCCTGGACACCCCGTTCTCCTGTGGATTCGGGTCACCTCTCACTCCTTTC
» AAVHSC7 RFLP HepG2 #4       (652) CCTCTTGCTTTCTTTGCCTGGACACCCCGTTCTCCTGTGGATTCGGGTCACCTCTCACTCCTTTC
» RFLP载体序列                (478) CCTCTTGCTTTCTTTGCCTGGACACCCCGTTCTCCTGTGGATTCGGGTCACCTCTCACTCCTTTC
» AC010327 AAVS1              (•174)
» 同源臂5'                    (478) CCTCTTGCTTTCTTTGCCTGGACACCCCGTTCTCCTGTGGATTCGGGTCACCTCTCACTCCTTTC
» 同源臂3'                    (1)
                                      716                                                             780
» AAVHSC7 RFLP PBSC #1        (717) ATTTGGGCAGCTCCCCTACCCCCCTTACCTCTCTAGTCTGTGCTAGCTCTTCCAGCCCCCTGTCA
» AAVHSC7 RFLP PBSC #2        (717) ATTTGGGCAGCTCCCCTACCCCCCTTACCTCTCTAGTCTGTGCTAGCTCTTCCAGCCCCCTGTCA
» AAVHSC7 RFLP HepG2 #1       (717) ATTTGGGCAGCTCCCCTACCCCCCTTACCTCTCTAGTCTGTGCTAGCTCTTCCAGCCCCCTGTCA
» AAVHSC7 RFLP HepG2 #2       (717) ATTTGGGCAGCTCCCCTACCCCCCTTACCTCTCTAGTCTGTGCTAGCTCTTCCAGCCCCCTGTCA
» AAVHSC7 RFLP HepG2 #3       (717) ATTTGGGCAGCTCCCCTACCCCCCTTACCTCTCTAGTCTGTGCTAGCTCTTCCAGCCCCCTGTCA
» AAVHSC7 RFLP HepG2 #4       (717) ATTTGGGCAGCTCCCCTACCCCCCTTACCTCTCTAGTCTGTGCTAGCTCTTCCAGCCCCCTGTCA
» RFLP载体序列                (543) ATTTGGGCAGCTCCCCTACCCCCCTTACCTCTCTAGTCTGTGCTAGCTCTTCCAGCCCCCTGTCA
» AC010327 AAVS1              (•174)
» 同源臂5'                    (543) ATTTGGGCAGCTCCCCTACCCCCCTTACCTCTCTAGTCTGTGCTAGCTCTTCCAGCCCCCTGTCA
» 同源臂3'                    (1)
```

图34(续)

```
                                  781                                                                                  845
» AAVHSC7 RFLP PBSC #1          (782) TGGCATCTTCCAGGGGTCCGAGAGCTCAGCTAGTCTTCTTCCTCCAACCCGGGCCCCTATGTCCA
» AAVHSC7 RFLP PBSC #2          (782) TGGCATCTTCCAGGGGTCCGAGAGCTCAGCTAGTCTTCTTCCTCCAACCCGGGCCCCTATGTCCA
» AAVHSC7 RFLP HepG2 #1         (782) TGGCATCTTCCAGGGGTCCGAGAGCTCAGCTAGTCTTCTTCCGCCAACCCGGGCCCCTATGTCCA
» AAVHSC7 RFLP HepG2 #2         (782) TGGCATCTTCCAGGGGTCCGAGAGCTCAGCTAGTCTTCTTCCTCCAACCCGGGCCCCTATGTCCA
» AAVHSC7 RFLP HepG2 #3         (782) TGGCATCTTCCAGGGGTCCGAGAGCTCAGCTAGTCTTCTTCCTCCAACCCGGGCCCCTATGTCCA
» AAVHSC7 RFLP HepG2 #4         (782) TGGCATCTTCCAGGGGTCCGAGAGCTCAGCTAGTCTTCTTCCTCCAACCCGGGCCCCTATGTCCA
» RFLP载体序列                  (608) TGGCATCTTCCAGGGGTCCGAGAGCTCAGCTAGTCTTCTTCCTCCAACCCGGGCCCCTATGTCCA
» AC010327 AAVS1                (•174)
» 同源臂5'                      (608) TGGCATCTTCCAGGGGTCCGAGAGCTCAGCTAGTCTTCTTCCTCCAACCCGGGCCCCTATGTCCA
» 同源臂3'                      (1)
                                      846                                                                                  910
» AAVHSC7 RFLP PBSC #1          (847) CTTCAGGACAGCATGTTTGCTGCCTCCAGGGATCCTGTGTCCCCGAGCTGGGACCACCTTATATT
» AAVHSC7 RFLP PBSC #2          (847) CTTCAGGACAGCATGTTTGCTGCCTCCAGGGATCCTGTGTCCCCGAGCTGGGACCACCTTATATT
» AAVHSC7 RFLP HepG2 #1         (847) CTTCAGGACAGCATGTTTGCTGCCTCCAGGGATCCTGTGTCCCCGAGCTGGGACCACCTTATATT
» AAVHSC7 RFLP HepG2 #2         (847) CTTCAGGACAGCATGTTTGCTGCCTCCAGGGATCCTGTGTCCCCGAGCTGGGACCACCTTATATT
» AAVHSC7 RFLP HepG2 #3         (847) CTTCAGGACAGCATGTTTGCTGCCTCCAGGGATCCTGTGTCCCCGAGCTGGGACCACCTTATATT
» AAVHSC7 RFLP HepG2 #4         (847) CTTCAGGACAGCATGTTTGCTGCCTCCAGGGATCCTGTGTCCCCGAGCTGGGACCACCTTATATT
» RFLP载体序列                  (673) CTTCAGGACAGCATGTTTGCTGCCTCCAGGGATCCTGTGTCCCCGAGCTGGGACCACCTTATATT
» AC010327 AAVS1                (•174)
» 同源臂5'                      (673) CTTCAGGACAGCATGTTTGCTGCCTCCAGGGATCCTGTGTCCCCGAGCTGGGACCACCTTATATT
» 同源臂3'                      (1)
                                      911                                                                                  975
» AAVHSC7 RFLP PBSC #1          (912) CCCAGGGCCGGTTAATGTGGCTCTGGTTCTGGGTACTTTTATCTGTCCCCTCCACCCCACAGTGG
» AAVHSC7 RFLP PBSC #2          (912) CCCAGGGCCGGTTAATGTGGCTCTGGTTCTGGGTACTTTTATCTGTCCCCTCCACCCCACAGTGG
» AAVHSC7 RFLP HepG2 #1         (912) CCCAGGGCCGGTTAATGTGGCTCTGGTTCTGGGTACTTTTATCTGTCCCCTCCACCCCACAGTGG
» AAVHSC7 RFLP HepG2 #2         (912) CCCAGGGCCGGTTAATGTGGCTCTGGTTCTGGGTACTTTTATCTGTCCCCTCCACCCCACAGTGG
» AAVHSC7 RFLP HepG2 #3         (912) CCCAGGGCCGGTTAATGTGGCTCTGGTTCTGGGTACTTTTATCTGTCCCCTCCACCCCACAGTGG
» AAVHSC7 RFLP HepG2 #4         (912) CCCAGGGCCGGTTAATGTGGCTCTGGTTCTGGGTACTTTTATCTGTCCCCTCCACCCCACAGTGG
» RFLP载体序列                  (738) CCCAGGGCCGGTTAATGTGGCTCTGGTTCTGGGTACTTTTATCTGTCCCCTCCACCCCACAGTGG
» AC010327 AAVS1                (•174)
» 同源臂5'                      (738) CCCAGGGCCGGTTAATGTGGCTCTGGTTCTGGGTACTTTTATCTGTCCCCTCCACCCCACAGTGG
» 同源臂3'                      (1)
                                      976                                                                                 1040
» AAVHSC7 RFLP PBSC #1          (977) GGCAAGCTTTACTAGGGACAGGATTGG
» AAVHSC7 RFLP PBSC #2          (977) GGCAAGCTTTACTAGGGACAGGATTGG
» AAVHSC7 RFLP HepG2 #1         (977) GGCAAGCTTTACTAGGGACAGGATTGG
» AAVHSC7 RFLP HepG2 #2         (977) GGCAAGCTTTACTAGGGACAGGATTGA
» AAVHSC7 RFLP HepG2 #3         (977) GGCAAGCTTTACTAGGGACAGGATTGG
» AAVHSC7 RFLP HepG2 #4         (977) GGCAAGCTTTACTAGGGACAGGATTGG
» RFLP载体序列                  (803) GGCAAGCTTTACTAGGGACAGGATTGGTGACAGAAAAGCCCCATCCTTAGGCCTCCTCCTTCCTA
» AC010327 AAVS1                (•174)
» 同源臂5'                      (•801)
» 同源臂3'                      (1)                TAGGGACAGGATTGGTGACAGAAAAGCCCCATCCTTAGGCCTCCTCCTTCCTA
>> Hind III                          AAGCTT
```

图34(续)

AAVF靶向较小插入物向AAVS1-CD34+细胞中的整合

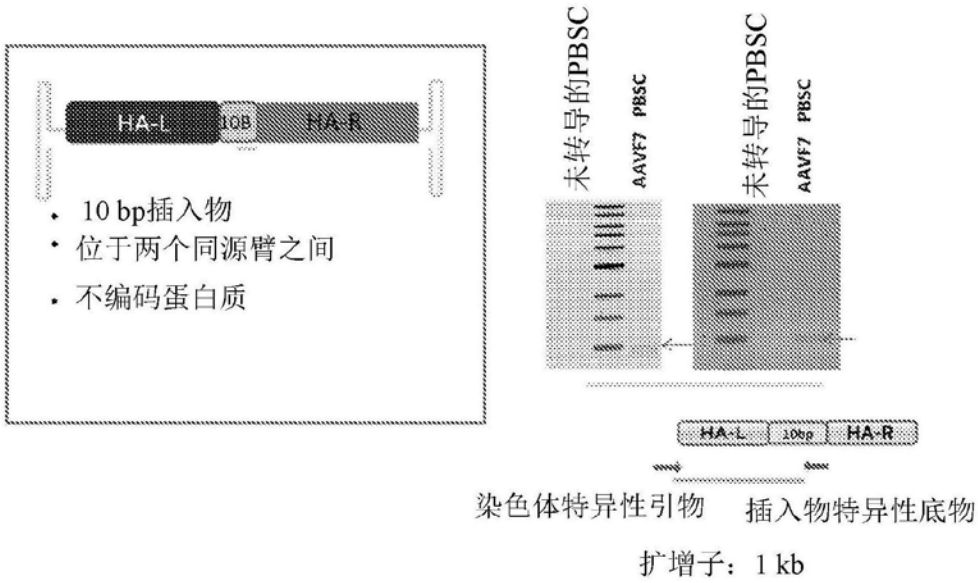


图35

GFP表达（细胞%）

	24H	48H
AAVF7	18.5	15.2
AAVF17	31.6	26.4
AAV6	0	ND
AAV8	0	ND
AAV9	2.81	ND

基因组整合

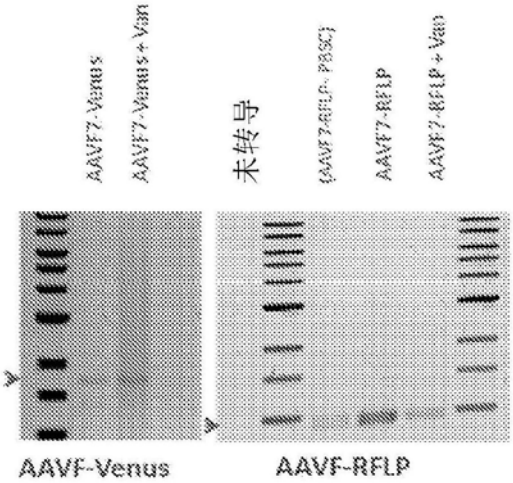


图36

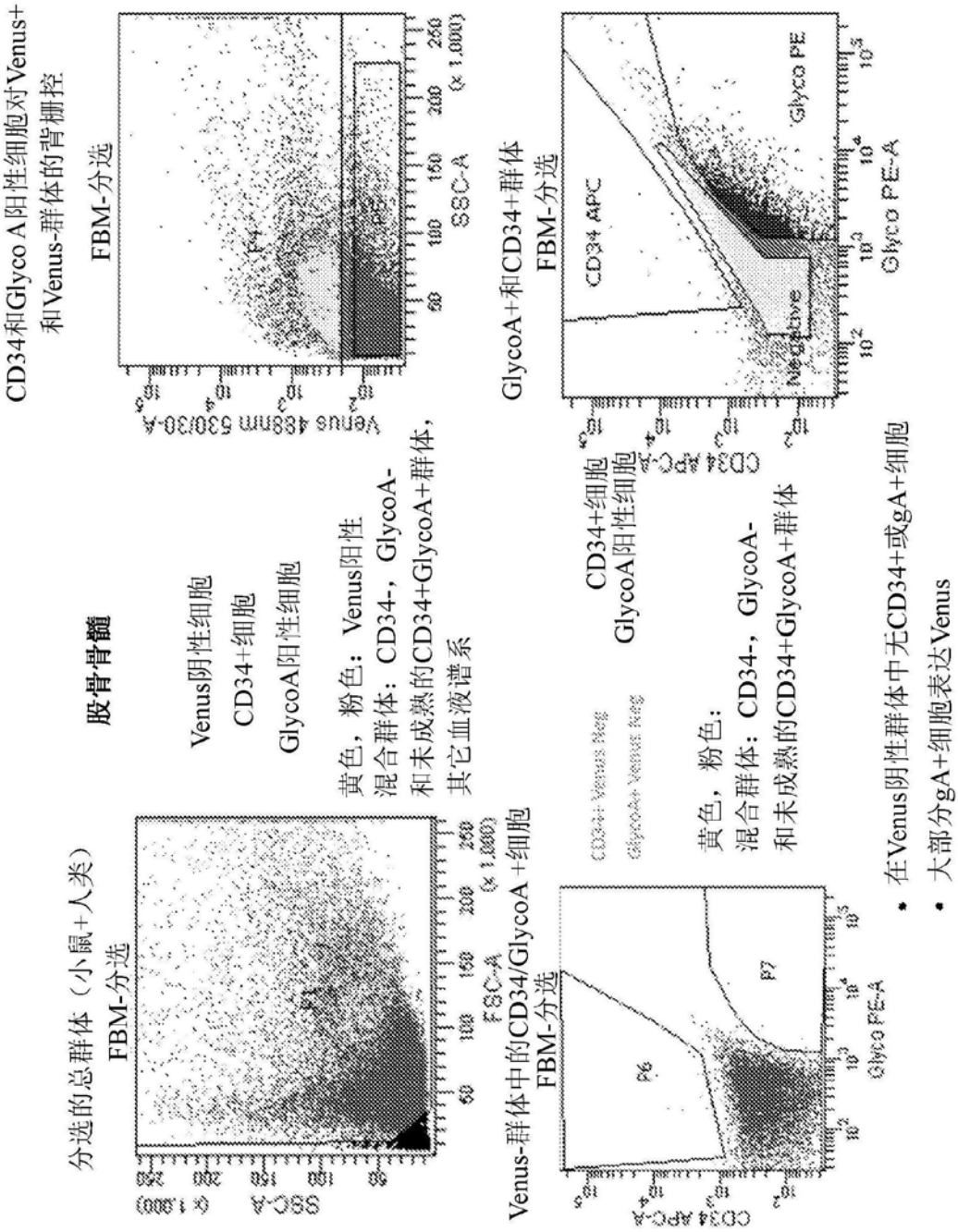

分选的总群体（小鼠+人类）
FBM-分选
SSC-A
FSC-A
股骨骨髓
Venus阴性细胞
CD34+细胞
GlycoA阳性细胞
黄色，粉色：Venus阳性
混合群体：CD34-，GlycoA-
和未成熟的CD34+GlycoA+群体，
其它血液谱系
CD34和Glyco A阳性细胞对Venus+
和Venus-群体的背栅控
FBM-分选
Venus 488nm 530/30-A
SSC-A
Venus-群体中的CD34/GlycoA +细胞
FBM-分选
CD34 APC-A
Glyco PE-A
CD34+细胞
GlycoA阳性细胞
黄色，粉色：
混合群体：CD34-，GlycoA-
和未成熟的CD34+GlycoA+群体
GlycoA+和CD34+群体
FBM-分选
CD34 APC
Negative
Glyco PE
• 在Venus阴性群体中无CD34+或gA+细胞
• 大部分gA+细胞表达Venus

图37

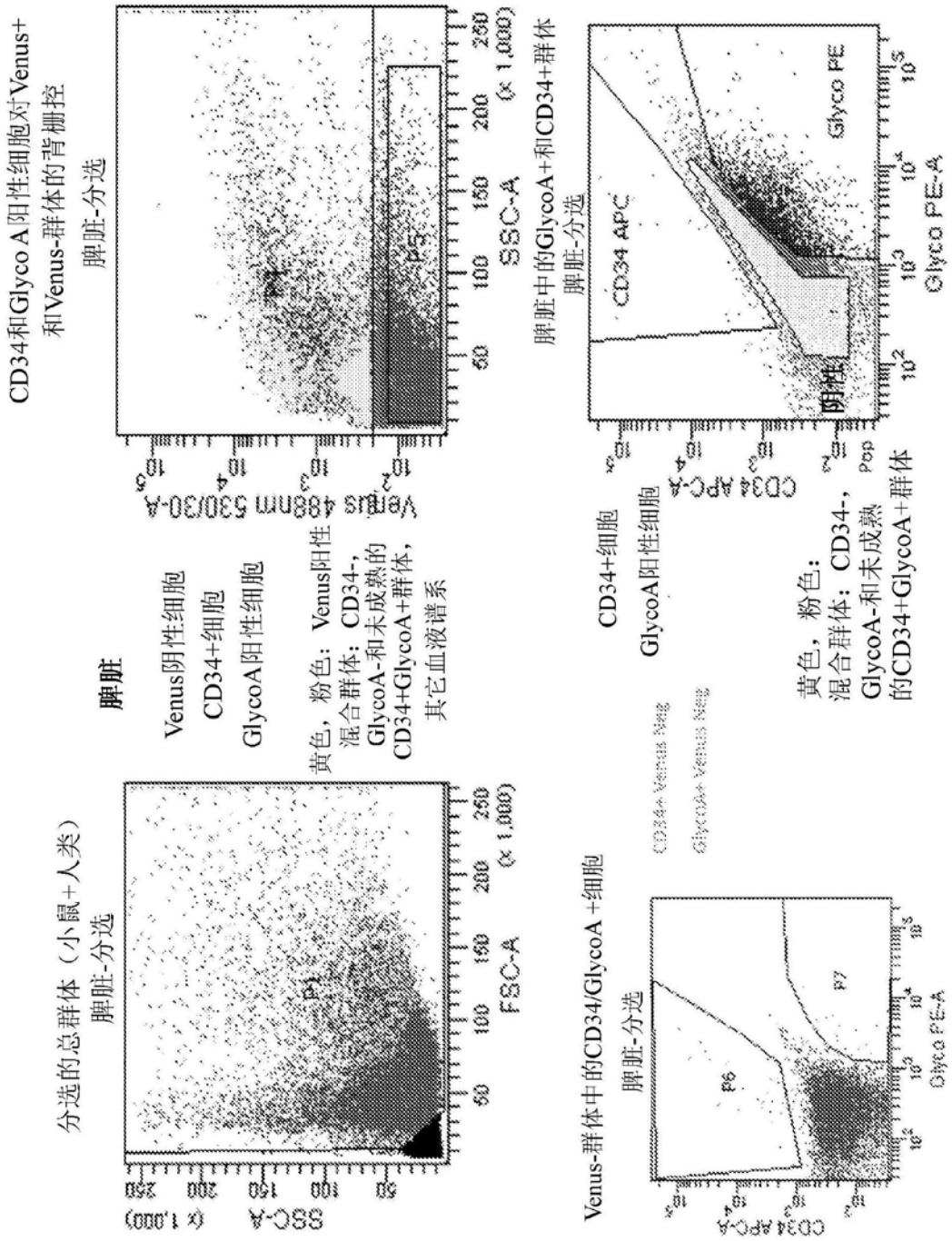

分选的总群体（小鼠+人类）
脾脏-分选
SSC-A (x 1,000)
FSC-A (x 1,000)
P1
CD34和Glyco A阳性细胞对Venus+
和Venus-群体的背栅控
脾脏-分选
Venus 488nm 530/30-A
SSC-A (x 1,000)
P4
P5
脾脏
Venus阴性细胞
CD34+细胞
GlycoA阳性细胞
黄色，粉色：Venus阳性
混合群体：CD34-，
GlycoA-和未成熟的
CD34+GlycoA+群体，
其它血液谱系
Venus-群体中的CD34/GlycoA +细胞
脾脏-分选
CD34 APC-A
Glyco PE-A
P6
P7
CD34+ Venus Neg
GlycoA+ Venus Neg
脾脏中的GlycoA+和CD34+群体
脾脏-分选
CD34 APC
Glyco PE
阴性
CD34 APC-A
Glyco PE-A
Pop
CD34+细胞
GlycoA阳性细胞
黄色，粉色：
混合群体：CD34-，
GlycoA-和未成熟
的CD34+GlycoA+群体

图38

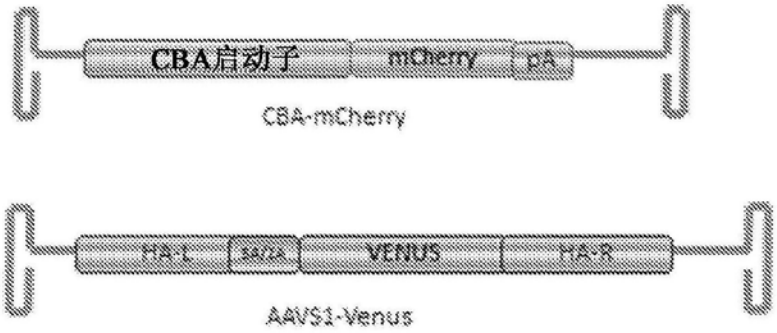

CBA启动子
mCherry
pA
CBA-mCherry
HA-L
SA/2A
VENUS
HA-R
AAVS1-Venus

图39

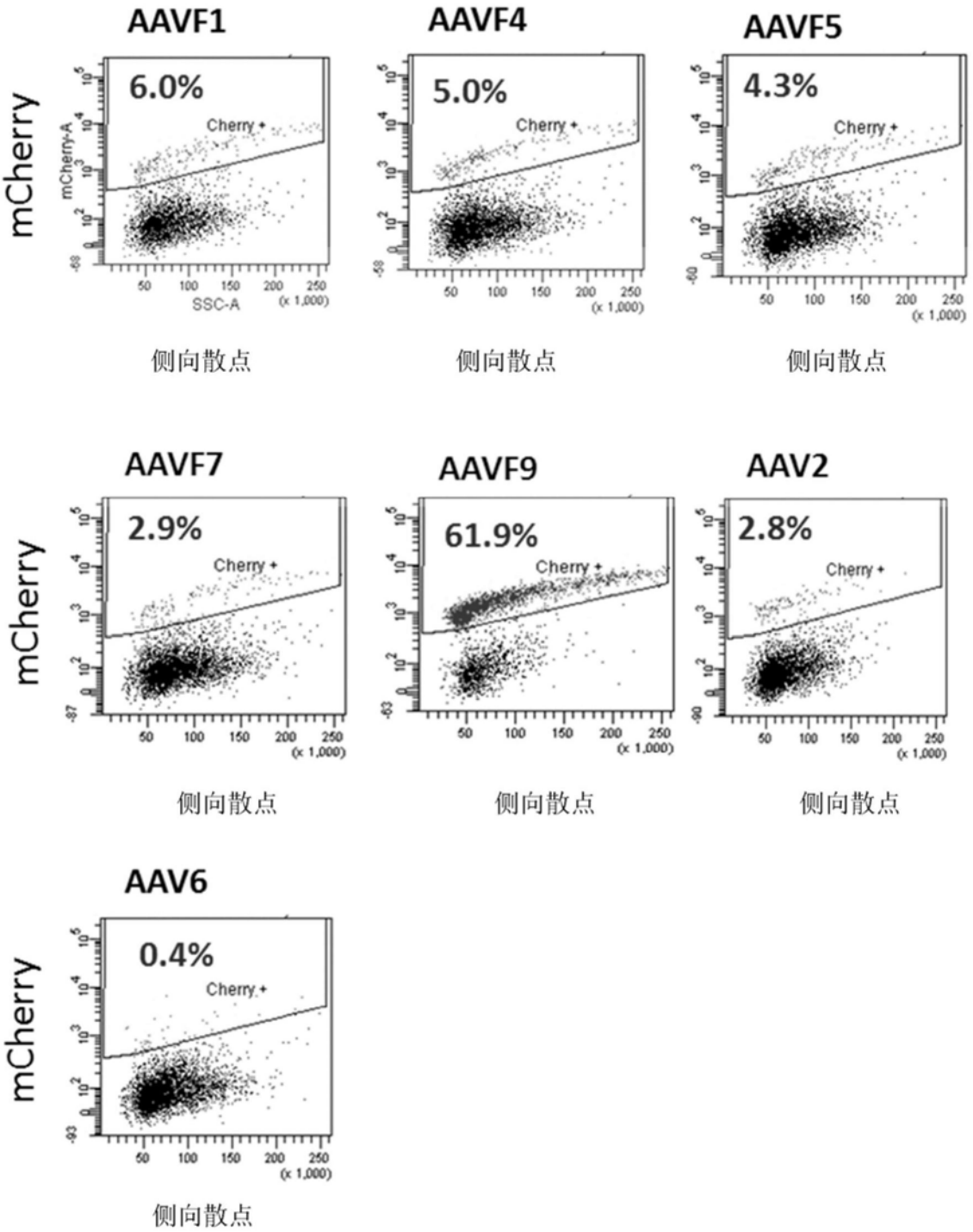

AAVF1
6.0%
Cherry +
mCherry
mCherry-A
SSC-A
(x 1,000)
侧向散点
AAVF4
5.0%
Cherry +
侧向散点
AAVF5
4.3%
Cherry +
侧向散点
AAVF7
2.9%
Cherry +
侧向散点
AAVF9
61.9%
Cherry +
侧向散点
AAV2
2.8%
Cherry +
侧向散点
AAV6
0.4%
Cherry +
侧向散点

图40A

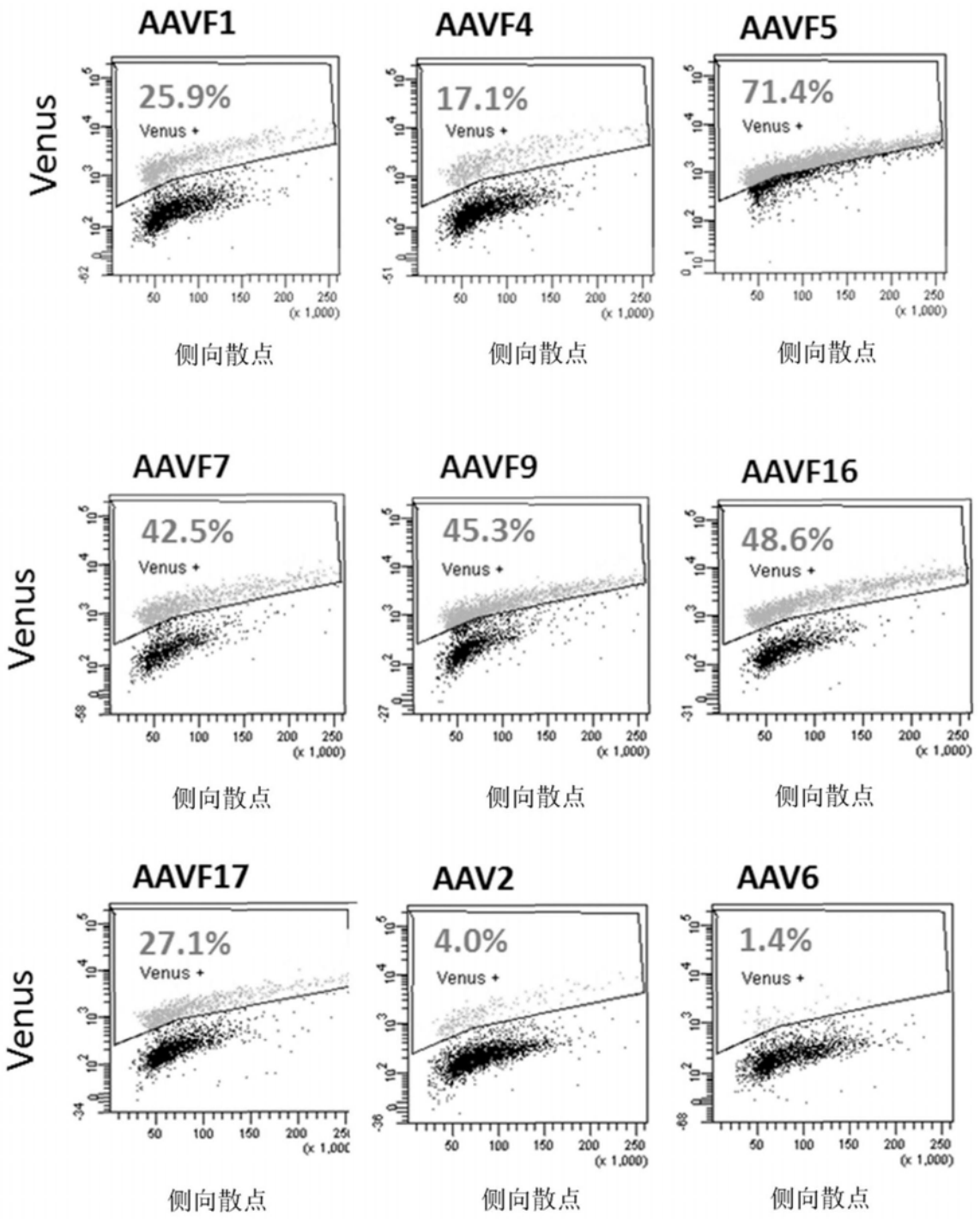

AAVF1
25.9%
Venus +
AAVF4
17.1%
Venus +
AAVF5
71.4%
Venus +
AAVF7
42.5%
Venus +
AAVF9
45.3%
Venus +
AAVF16
48.6%
Venus +
AAVF17
27.1%
Venus +
AAV2
4.0%
Venus +
AAV6
1.4%
Venus +
Venus
侧向散点
(x 1,000)

图40B

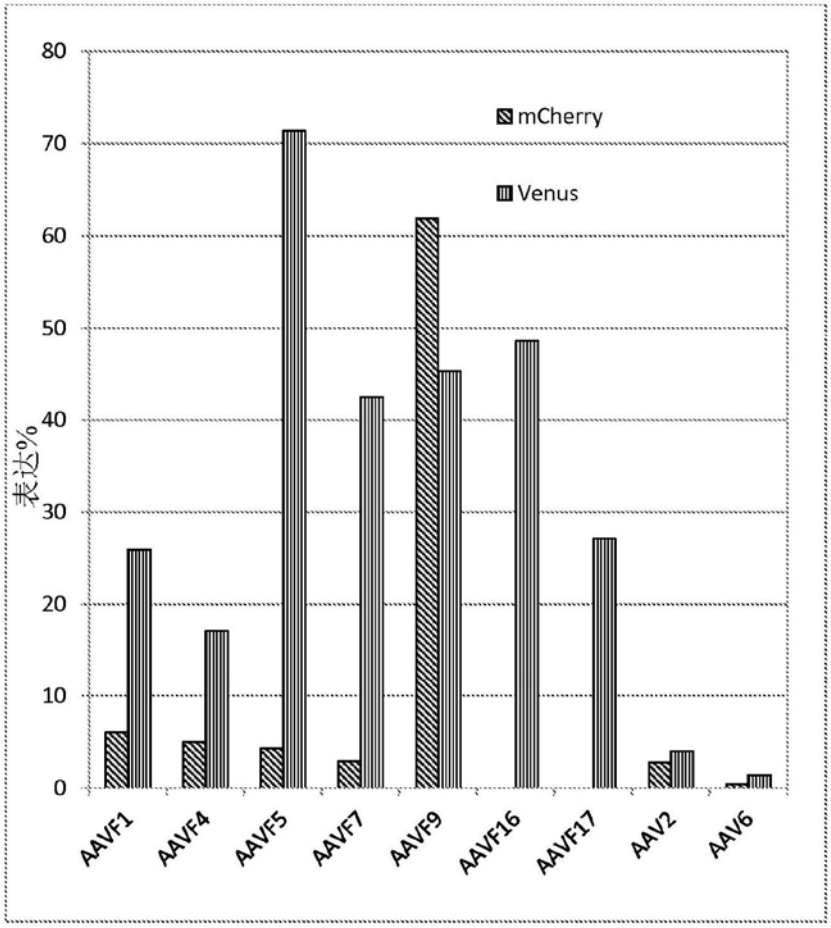

80
70
60
50
40
30
20
10
0
表达%
mCherry
Venus
AAVF1
AAVF4
AAVF5
AAVF7
AAVF9
AAVF16
AAVF17
AAV2
AAV6

图40C

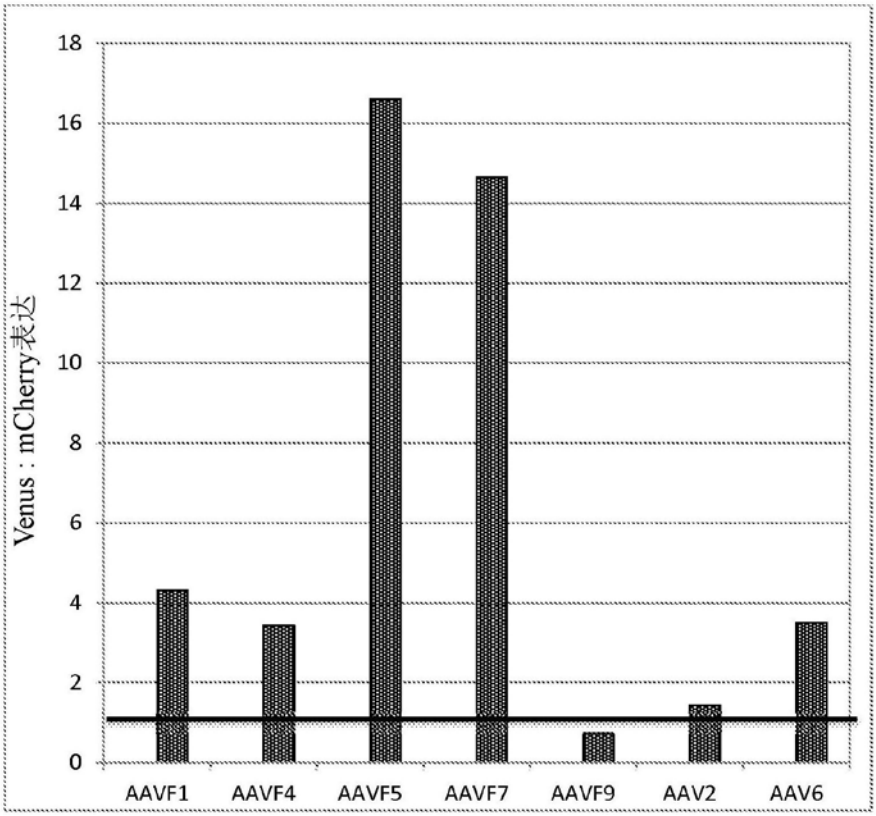

18
16
14
12
10
8
6
4
2
0
Venus : mCherry表达
AAVF1
AAVF4
AAVF5
AAVF7
AAVF9
AAV2
AAV6

图40D