(19)国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 116390934 A
(43)申请公布日 2023.07.04

(21)申请号 202180058369.6

(22)申请日 2021.05.26

(30)优先权数据
63/030,038 2020.05.26 US
63/119,554 2020.11.30 US
63/134,885 2021.01.07 US
63/181,037 2021.04.28 US

(85)PCT国际申请进入国家阶段日
2023.01.30

(86)PCT国际申请的申请数据
PCT/US2021/034329 2021.05.26

(87)PCT国际申请的公布数据
WO2021/242909 EN 2021.12.02

(71)申请人 塑造治疗公司
地址 美国华盛顿州

(72)发明人 弗朗索瓦·维格纳尔特
托马斯·帕卡德
大卫·杰弗里·赫斯
凯文·克里斯多佛·斯坦

(74)专利代理机构 北京安信方达知识产权代理有限公司 11262
专利代理师 王玮玮 郑霞

(51)Int.Cl.
C07K 14/005(2006.01)

权利要求书5页 说明书553页 附图48页

(54)发明名称

功能性AAV衣壳的高通量工程化

(57)摘要

本文公开了能够组装成病毒颗粒并且具有提高的对例如CNS组织的组织嗜性的工程化AAV VP衣壳多肽。使用本文所述的高通量发现系统对所述衣壳进行工程化。在某些实施方案中,本文提供了在对应于SEQ ID NO:1中的残基581至残基589的残基中具有至少一个突变的重组腺相关病毒(AAV)VP衣壳多肽。

CN 116390934 A

1.一种工程化腺相关病毒(AAV)病毒蛋白(VP)衣壳多肽,所述工程化腺相关病毒(AAV)病毒蛋白(VP)衣壳多肽具有与SEQ ID NO:1至少70%同一的氨基酸序列,

其中所述工程化AAV VP衣壳多肽相较于SEQ ID NO:1在从对应于SEQ ID NO:1的残基581的残基至对应于SEQ ID NO:1的残基589的残基的区域,包括端点在内,中具有至少一个突变,

其中所述工程化AAV VP衣壳多肽能够组装成重组AAV病毒体(rAAV),

其中与具有SEQ ID NO:1的野生型AAV5 VP衣壳多肽的rAAV病毒体相比,所述至少一个突变赋予所述rAAV较高的对中枢神经系统(CNS)组织的嗜性,并且

其中所述工程化AAV VP衣壳多肽不具有SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7和SEQ ID NO:8中任一者的序列。

2.如权利要求1所述的工程化AAV VP衣壳多肽,其中所述工程化AAV VP衣壳多肽具有SEQ ID NO:2的序列,并且其中Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9各自独立地选自任何氨基酸。

3.一种工程化腺相关病毒(AAV)病毒蛋白(VP)衣壳多肽,所述工程化腺相关病毒(AAV)病毒蛋白(VP)衣壳多肽具有SEQ ID NO:2的氨基酸序列,

其中氨基酸残基Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9各自独立地选自A、R、N、D、C、E、Q、G、H、I、L、K、M、F、P、S、T、W、Y和V,

其中所述工程化AAV VP衣壳多肽能够组装成重组AAV病毒体(rAAV),并且

其中所述工程化AAV VP衣壳多肽不具有SEQ ID NO:1、SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7和SEQ ID NO:8中任一者的序列。

4.如权利要求3所述的工程化AAV VP衣壳多肽,其中与具有SEQ ID NO:1的野生型AAV5 VP衣壳多肽的rAAV病毒体相比,所述rAAV具有较高的对中枢神经系统(CNS)组织的嗜性。

5.如权利要求1-2中任一项所述的工程化AAV VP衣壳多肽,其中从对应于残基581的残基至对应于残基589的残基的区域,包括端点在内,具有与SEQ ID NO:7118-SEQ ID NO:10117中任一者至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%同一的序列。

6.如权利要求5所述的工程化AAV VP衣壳多肽,其中从对应于残基581的残基至对应于残基589的残基的区域,包括端点在内,具有SEQ ID NO:7118-SEQ ID NO:10117中任一者的序列。

7.如权利要求3-4中任一项所述的工程化AAV VP衣壳多肽,其中SEQ ID NO:2的Xaa1至Xaa9区域具有与SEQ ID NO:7118-SEQ ID NO:10117中任一者至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%同一的序列。

8.如权利要求7所述的工程化AAV VP衣壳多肽,其中SEQ ID NO:2的Xaa1至Xaa9区域具有SEQ ID NO:7118-SEQ ID NO:10117中任一者的序列。

9.如权利要求2-4中任一项所述的工程化AAV VP衣壳多肽,其中:

Xaa1选自由A、C、K、M、Q、R、T和W组成的组;或者

Xaa2选自由F、I、K、R、T和W组成的组;或者

Xaa3选自由A、H、N、R和W组成的组;或者

Xaa4选自由E、G、I、M、Q和R组成的组;或者

Xaa5选自由C、G、K、I、M和R组成的组;或者

Xaa6选自由I、K、L、P、Q、R和Y组成的组;或者

Xaa7选自由D、I、K、R、V和W组成的组;或者

Xaa8选自由C、G、H、K、L和V组成的组;或者

Xaa9选自由I、K、L、R和V组成的组;或者

它们的任何组合。

10.如权利要求9所述的工程化AAV VP衣壳多肽,其中Xaa1选自K、Q、R和W。

11.如权利要求9所述的工程化AAV VP衣壳多肽,其中Xaa2选自F、I、R和T。

12.如权利要求9所述的工程化AAV VP衣壳多肽,其中Xaa3选自A、R和W。

13.如权利要求9所述的工程化AAV VP衣壳多肽,其中Xaa4选自E、M和R。

14.如权利要求9所述的工程化AAV VP衣壳多肽,其中Xaa5选自K、I和R。

15.如权利要求9所述的工程化AAV VP衣壳多肽,其中Xaa6选自K、R和Y。

16.如权利要求9所述的工程化AAV VP衣壳多肽,其中Xaa7选自I、R和V。

17.如权利要求9所述的工程化AAV VP衣壳多肽,其中Xaa8选自H、K和V。

18.如权利要求9所述的工程化AAV VP衣壳多肽,其中Xaa9选自I、K和R。

19.如权利要求9所述的工程化AAV VP衣壳多肽,其中Xaa1为K。

20.如权利要求9所述的工程化AAV VP衣壳多肽,其中Xaa2为R。

21.如权利要求9所述的工程化AAV VP衣壳多肽,其中Xaa3为R。

22.如权利要求9所述的工程化AAV VP衣壳多肽,其中Xaa4为R。

23.如权利要求9所述的工程化AAV VP衣壳多肽,其中Xaa5为I。

24.如权利要求9所述的工程化AAV VP衣壳多肽,其中Xaa6为R。

25.如权利要求9所述的工程化AAV VP衣壳多肽,其中Xaa7为V。

26.如权利要求9所述的工程化AAV VP衣壳多肽,其中Xaa8为H。

27.如权利要求9所述的工程化AAV VP衣壳多肽,其中Xaa9为R。

28.如权利要求1-27中任一项所述的工程化AAV VP衣壳多肽,其中从对应于残基581的残基至对应于残基589的残基的区域,包括端点在内,具有与SEQ ID NO:7118-SEQ ID NO:8117中任一者至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%同一的序列。

29.如权利要求28所述的工程化AAV VP衣壳多肽,其中SEQ ID NO:1的从残基581至残基589的区域具有SEQ ID NO:7118-SEQ ID NO:8117中任一者的序列。

30.如权利要求2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa1具有低氨基酸溶解性。

31.如权利要求30所述的工程化AAV VP衣壳多肽,其中Xaa1选自K、R和Q。

32.如权利要求2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa1具有低氨基酸亲疏水性。

33.如权利要求32所述的工程化AAV VP衣壳多肽,其中Xaa1选自K和R。

34.如权利要求2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa1具有高平均氨基酸柔性指数。

35.如权利要求34所述的工程化AAV VP衣壳多肽,其中Xaa1选自D、E、R、K、G、I、N、Q和r

S。

36.如权利要求2-4中任一项所述的工程化AAV VP衣壳多肽，其中Xaa1具有高氢键供体。

37.如权利要求36所述的工程化AAV VP衣壳多肽，其中Xaa1选自K和R。

38.如权利要求2-4中任一项所述的工程化AAV VP衣壳多肽，其中Xaa1具有低氨基酸可突变性。

39.如权利要求38所述的工程化AAV VP衣壳多肽，其中Xaa1选自K、R、P和H。

40.如权利要求2-4中任一项所述的工程化AAV VP衣壳多肽，其中Xaa2具有低氨基酸溶解性。

41.如权利要求40所述的工程化AAV VP衣壳多肽，其中Xaa2选自R、K、Q和S。

42.如权利要求2-4中任一项所述的工程化AAV VP衣壳多肽，其中Xaa2具有低氨基酸亲疏水性。

43.如权利要求42所述的工程化AAV VP衣壳多肽，其中Xaa2选自R、K、D、E、N、Q、H、P、Y、W、S和r T。

44.如权利要求2-4中任一项所述的工程化AAV VP衣壳多肽，其中Xaa2具有高氨基酸电荷。

45.如权利要求44所述的工程化AAV VP衣壳多肽，其中Xaa2选自R、K和H。

46.如权利要求2-4中任一项所述的工程化AAV VP衣壳多肽，其中Xaa3具有高氨基酸溶解性。

47.如权利要求46所述的工程化AAV VP衣壳多肽，其中Xaa3选自A、M、V、W、L和I。

48.如权利要求2-4中任一项所述的工程化AAV VP衣壳多肽，其中Xaa5具有高氨基酸溶解性。

49.如权利要求48所述的工程化AAV VP衣壳多肽，其中Xaa5选自C、M、V、W、L和I。

50.如权利要求2-4中任一项所述的工程化AAV VP衣壳多肽，其中Xaa5具有高亲疏水性。

51.如权利要求50所述的工程化AAV VP衣壳多肽，其中Xaa5选自M、V和I。

52.如权利要求2-4中任一项所述的工程化AAV VP衣壳多肽，其中Xaa5具有低平均氨基酸柔性指数。

53.如权利要求52所述的工程化AAV VP衣壳多肽，其中Xaa5选自M、W、F和C。

54.如权利要求2-4中任一项所述的工程化AAV VP衣壳多肽，其中Xaa8具有高氨基酸溶解性。

55.如权利要求54所述的工程化AAV VP衣壳多肽，其中Xaa8选自H、V和I。

56.如权利要求30-55中任一项所述的工程化AAV VP衣壳多肽，其中从对应于残基581的残基至对应于残基589的残基的区域包括端点在内具有与SEQ ID NO:8118-SEQ ID NO:9117中任一者具有至少70％、至少75％、至少80％、至少85％、至少90％、至少92％、至少95％、至少97％或至少99％同一的序列。

57.如权利要求56所述的工程化AAV VP衣壳多肽，其中从对应于残基581的残基至对应于残基589的残基的区域包括端点在内具有SEQ ID NO:8118-SEQ ID NO:9117中任一者的序列。

58.如权利要求1-57中任一项所述的工程化AAV VP衣壳多肽，其中CNS组织选自前脑皮层、枕叶皮层、颞叶皮层、丘脑、下丘脑、黑质、海马DG、海马CA1、海马CA3、小脑以及它们的任何组合。

59.如权利要求1-58中任一项所述的工程化AAV VP衣壳多肽，其中对CNS组织的嗜性被测量为所述rAAV病毒体在CNS组织中相较于非CNS组织中的相对积聚，其中所述非CNS组织由肝脏、骨骼肌、心脏、肺、脾脏、淋巴结、骨髓、乳腺、皮肤、肾上腺、甲状腺、结肠、坐骨神经和脊髓共同组成。

60.如权利要求1-59中任一项所述的工程化AAV VP衣壳多肽，其中所述较高的组织嗜性是在CNS组织中相较于非CNS组织中1.0005倍至约1000倍增加的积聚。

61.如权利要求60所述的工程化AAV VP衣壳多肽，其中所述较高的组织嗜性是在CNS组织中相较于非CNS组织中至少约1.0005倍、至少约两倍、至少约三倍、至少约四倍、至少约五倍、至少约十倍、至少约二十倍、至少约50倍、至少约75倍、至少约100倍或至少约1000倍增加的积聚。

62.如权利要求1-61中任一项所述的工程化AAV VP衣壳多肽，其中所述工程化AAV VP衣壳多肽是工程化AAV5 VP衣壳多肽。

63.如权利要求1-61中任一项所述的工程化AAV VP衣壳多肽，其中所述突变是取代。

64.一种重组AAV病毒体(rAAV)，所述重组AAV病毒体(rAAV)包含：

组装成衣壳的权利要求1-63中任一项所述的工程化AAV VP衣壳多肽;和

有效载荷，

其中所述工程化AAV VP衣壳多肽衣壳包裹所述有效载荷。

65.如权利要求64所述的重组rAAV，其中所述有效载荷包含治疗性多核苷酸，任选地其中所述治疗性多核苷酸编码转基因或基因组修饰实体。

66.如权利要求65所述的rAAV，其中所述治疗性多核苷酸编码指导RNA、tRNA、阻抑tRNA、siRNA、miRNA、mRNA、shRNA、环状RNA或反义寡核苷酸(ASO)、核酶、DNA酶、适体或它们的任何组合。

67.如权利要求65所述的rAAV，其中所述治疗性多核苷酸编码线性治疗性多核苷酸或环状治疗性多核苷酸。

68.如权利要求65所述的rAAV，其中所述转基因选自表1。

69.如权利要求65所述的rAAV，其中所述基因组修饰实体包含CRISPR/Cas系统、作用于RNA的腺苷脱氨酶(ADAR)酶、转录激活物或转录阻抑物。

70.如权利要求69所述的rAAV，其中所述CRISPR/Cas系统包括Cas3、Cas8、Cas10、Cas9、Cas4、Cas12或Cas13。

71.如权利要求69所述的rAAV，其中所述ADAR是人ADAR1或人ADAR2。

72.如权利要求69所述的rAAV，其中所述转录激活物是VP64。

73.如权利要求69所述的rAAV，其中所述转录阻抑物是KRAB。

74.一种药物组合物，所述药物组合物包含权利要求64-73中任一项所述的rAAV病毒体和药学上可接受的载体。

75.一种治疗方法，所述方法包括向有需要的受试者施用治疗有效量的权利要求74所述的药物组合物。

76.如权利要求74所述的药物组合物，所述药物组合物用于治疗有需要的受试者，其中所述用途包括向所述受试者施用治疗有效量的如权利要求74所述的药物组合物。

77.如权利要求64-73中任一项所述的rAAV病毒体，所述rAAV病毒体用于制造用于治疗有需要的受试者的药物，所述治疗通过施用治疗有效量的所述药物来实现。

78.如权利要求75-77中任一项所述的方法，其中所述rAAV病毒体的所述治疗有效量小于衣壳包裹相同有效载荷的具有SEQ ID NO:1的野生型AAV5 VP衣壳多肽的rAAV病毒体的治疗有效量，其中具有所述SEQ ID NO:1的野生型AAV5 VP衣壳多肽的所述rAAV病毒体通过相同的施用途径施用。

79.如权利要求75-78中任一项所述的方法，其中与衣壳包裹相同有效载荷的具有SEQ ID NO:1的野生型AAV5 VP衣壳多肽的rAAV病毒体的治疗有效量相比，所述rAAV病毒体的所述治疗有效量在所述受试者中引起较低的毒性，其中具有所述SEQ ID NO:1的野生型AAV5 VP衣壳多肽的所述rAAV病毒体通过相同的施用途径施用。

80.如权利要求75-79中任一项所述的方法，其中所述治疗有效量为至少1x 10^{5}个病毒基因组/kg受试者体重至5x 10^{14}个病毒基因组/kg受试者体重。

81.如权利要求75-80中任一项所述的方法，其中通过全身施用来施用所述rAAV病毒体。

82.如权利要求81所述的方法，其中所述全身施用是静脉内施用、肌肉内施用、腹膜内施用或口服施用。

83.如权利要求75-82中任一项所述的方法，其中所述受试者是人类或非人类动物。

84.如权利要求75-83中任一项所述的方法，其中所述受试者患有疾病。

85.如权利要求84所述的方法，其中所述疾病是神经系统疾患。

86.如权利要求85所述的方法，其中所述神经系统疾患是AADC缺乏症、阿尔茨海默氏病、tau蛋白病、突触核蛋白病、贝登氏病、MPS-IIIB、伴有GBA1突变的额颞叶痴呆(PD-GBA)、神经性戈谢氏病、戈谢病2型、卡纳万病、帕金森氏病、泰萨克斯病、亨廷顿氏病、波托基-卢普斯基综合征、肌萎缩侧索硬化症、唐氏综合征、圣菲利波病A型、圣菲利波病B型或雷特综合征。

87.如权利要求86所述的方法，其中所述痴呆是额颞叶痴呆(FTD)。

功能性AAV衣壳的高通量工程化

[0001] 1.相关申请的交叉引用

[0002] 本申请要求2020年5月26日提交的美国临时专利申请第63/030,038号;2020年11月30日提交的美国临时专利申请第63/119,554号;2021年1月7日提交的美国临时专利申请第63/134,885号;和2021年4月28日提交的美国临时专利申请第63/181,037号的权益和优先权,所述美国临时专利申请以引用方式整体并入。

[0003] 2.序列表

[0004] 本申请含有具XX序列的序列表,所述序列表通过EFS-Web提交并且特此以引用方式整体并入本文。于XXXX创建的所述ASCII副本命名为45736WO_sequencelisting.txt并且大小为XXX字节。

3.背景技术

[0005] 重组腺相关病毒(rAAV)为基因疗法的体内递送提供了领先的平台。目前的临床试验采用数量有限的AAV衣壳,主要来自天然存在的人或灵长类动物血清型,如AAV1、AAV2、AAV5、AAV6、AAV8、AAV9、AAVrh.10、AAV4rh.74和AAVhu.67。这些衣壳通常提供对目标组织的次优靶向,这既是由于对目标组织的感染性差,也是由于竞争性的肝脏嗜性。增加剂量以确保感染所需组织可导致剂量依赖性肝脏毒性。此外,使用天然存在的衣壳对免疫记忆提出了挑战-免疫前的患者群体被排除在治疗之外,并且对先前未经免疫的患者重复给药往往是不可能的。因此,需要用于基因疗法的另外AAV衣壳,特别地是赋予rAAV对特定组织的高感染性和低肝脏嗜性的衣壳。

4.发明内容

[0006] 我们设计了一种用于高通量工程化对各种组织具有改变的嗜性的功能性AAV衣壳的系统,并使用该系统鉴定了具有增加或减少的肝脏嗜性,以及增加的对靶组织(如肝脏或中枢神经系统(CNS)组织)的嗜性的衣壳变体。

[0007] 本文提供了一种具有与SEQ ID NO:1至少70%同一的氨基酸序列的工程化腺相关病毒(AAV)病毒蛋白(VP)衣壳多肽,其中所述工程化AAV VP衣壳多肽相较于SEQ ID NO:1在从对应于SEQ ID NO:1的残基581的残基至对应于SEQ ID NO:1的残基589的残基的区域,包括端点在内,中具有至少一个突变,其中所述工程化AAV VP衣壳多肽能够组装成重组AAV病毒体(rAAV),其中与具有SEQ ID NO:1的野生型AAV5 VP衣壳多肽的rAAV病毒体相比,所述至少一个突变赋予所述rAAV较高的对中枢神经系统(CNS)组织的嗜性,并且其中所述工程化AAV VP衣壳多肽不具有SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7和SEQ ID NO:8中任一者的序列。在某些实施方案中,所述工程化AAV VP衣壳多肽具有SEQ ID NO:2的序列并且其中Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9各自独立地选自任何氨基酸。

[0008] 此外,本文提供了具有SEQ ID NO:2的氨基酸序列的工程化腺相关病毒(AAV)病毒蛋白(VP)衣壳多肽,其中氨基酸残基Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9

各自独立地选自A、R、N、D、C、E、Q、G、H、I、L、K、M、F、P、S、T、W、Y和V，其中所述工程化AAV VP衣壳多肽能够组装成重组AAV病毒体(rAAV)，并且其中所述工程化AAV VP衣壳多肽不具有SEQ ID NO:1、SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7和SEQ ID NO:8中任一者的序列。在某些实施方案中，与具有SEQ ID NO:1的野生型AAV5 VP衣壳多肽的rAAV病毒体相比，所述rAAV具有较高的对中枢神经系统(CNS)组织的嗜性。

5.附图说明

[0009] 参考以下描述和附图将更好地理解本发明的这些和其他特征、方面和优点，在附图中：

[0010] 图1是高通量AAV衣壳工程化系统的示意图。

[0011] 图2A提供了已显示与靶细胞相互作用的已知AAV衣壳的关键残基的侧视图(上图)和顶视图(下图)。图2B示出了实施例1中描述的AAV衣壳文库的突出特征，显示出引入多样性的区域，其中残基编号对应于AAV5 VP1中的氨基酸编号。

[0012] 图3是一张凝胶照片，显示了在用PS-DNA酶消化后有效去除线性插入片段和线性质粒骨架(“BB”)，富集“松弛形式”(“RF”)环状物种以进行高效的后续细菌转化。

[0013] 图4是高通量AAV衣壳工程化系统的示意图。左侧柱形图显示了AAV衣壳文库在不同制备阶段的大小，追踪了文库从理论多样性、合成衣壳基因、总克隆变体到所测序组装病毒的大小(误差条表示从NGS测序分析中预测的最小/最大多样性)。将汇集的病毒文库注射到非人灵长类动物中，并在足以确保稳定转导的一段时间后，对动物实施安乐死，并收获组织。从组织中纯化DNA(1)，并且(2)附加独特分子标识符(UMI)。添加核酸外切酶I以消化过量的含UMI的引物(3)。后续PCR扩增添加测序索引/组织条形码和衔接子用于下一代测序(NGS)测序。添加UMI可以对组织中的衣壳变体进行高分辨率频率分析。

[0014] 图5A至图5C示出了用于使来自文库的rAAV产量最大化的包装方法，其中图5A比较了标准三重转染(上图)与本文中使用的双质粒顺式包装方法(下图)，并且图5B显示了使用这两种转染方法的野生型AAV5的相对产量。图5C是肝脏标本中UMI分布的实例，显示一次(单个UMI)发现大多数衣壳变体。然而，衣壳变体的一个子集在给定组织中富集并且具有增加的UMI数，与它们在靶组织中增加的频率相对应。

[0015] 图6A至图6B是热图和统计分析，显示了通过病毒组装和组织转导的阶段进行的明确功能选择。图6A是在高通量系统从组装前、组装病毒到肝脏转导的各个阶段制备的热图。对于组装前热图，将衣壳变体文库克隆到质粒中并转化到电感受态细菌中。对所得文库进行测序，并测量每个位置处的氨基酸多样性，低水平的位置变异可能是由综合误差引起的。随后组装成病毒显示出明显的氨基酸残基/位置偏差，突出了病毒组装的选择。肝脏转导病毒显示出更大的AA残基/位置选择模式，与明显有利/不利的变体。请注意，对这些热图各自根据它们各自的输入频率进行归一化。图6B是统计分析表(ANOVA)，显示在三个肝脏样本中，氨基酸残基分布和残基位置显著改变，但样本间变异不显著。

[0016] 图7A至图7B显示了对多个样本之间重复观察到的(高UMI)衣壳变体的分析。图7A是示出两个肝脏样本中重叠变体鉴定计数的维恩图：在两个样本中观察到约38％的具有＞10个UMI的变体。图7B是显示位置氨基酸分布的热图，示出了重复观察到的衣壳变体的强选择残基/位置。

[0017] 图8是基于机器学习的衣壳变体聚类的示意图。该实例利用衣壳变体序列，对其使用机器学习算法来绘制感染肝脏的衣壳之间的相似性。然后，该输出可用于告知候选变体的选择，以选择性地靶向目标组织。

[0018] 图9示出了静脉内施用于非人灵长类动物(NHP)后，文库整体感染皮层的相对水平高于肝脏的性能实例。从肝脏和皮层分离DNA，并使用qPCR(左)或液滴数字PCR(ddPCR)(右)来量化从各自组织回收的病毒基因组。根据这些量化方法，文库整体具有相对于肝脏约2倍增加的CNS靶向性。

[0019] 图10示出了强调静电势的AAV5衣壳的野生型晶体结构，以及通过本文所述的方法获得的两个实例性重组VP1衣壳变体。

[0020] 图11A至图11D显示了维恩图和分析表，描绘了针对三种不同测序分析过滤器在仅肝脏组织、肝脏和脑皮层组织以及仅皮层组织中发现的独特序列变体的数量。对从肝脏或皮层组织样本回收的病毒基因组衣壳变体进行新一代测序。附加独特分子标识符(UMI)作为本文所述的分子回收过程的一部分。这增加了组织中存在给定衣壳变体的置信度，并且可与丰度相关联。图11A显示了维恩图，其中应用的测序分析过滤器为每各不同衣壳序列变体，4个或更多个不同UMI(在本文中也称为“计数”)。图11B显示了维恩图，其中应用的测序分析过滤器为50个或更多个UMI。图11C显示了维恩图，其中应用的测序分析过滤器为100个或更多个不同UMI。图11D显示了在仅皮层中发现的衣壳序列变体群体中UMI分布的直方图分析，其中应用的测序分析过滤器为如图11B所示的50个或更多个UMI。

[0021] 图12显示的热图显示了本公开rAAV5衣壳多肽序列中的归一化氨基酸残基频率，以及它们如何在变体从组装前到组装后到CNS转导子集的过程中富集。热图的x轴指示581-589区域中的位置，并且y轴包括所有氨基酸残基。

[0022] 图13显示了机器学习(ML)模型性能验证的实例。在CNS和非CNS组织中发现了不同的变体，也在两者中发现了一些共有变体(左图)。随机森林(RF)ML模型在预测CNS靶向性方面表现出良好的性能。在高预测类别概率值下，ML模型可以从非CNS靶向性变体序列解析CNS靶向性(中间图，右图)。

[0023] 图14示出了对模型输出概率有贡献的前20个位置特征(左图)。沙普利加和解释(SHAP)值可用于询问特征对模型预测的相对贡献。可以在组织靶向性和非靶向性变体之间进一步比较这些特征。右侧是具有KRLQQMETM(SEQ ID NO:1117)序列的CNS变体模型，代表预测增加CNS靶向性的一些特征。

[0024] 图15显示了从两个NHP跨所有组织类型回收的本公开的rAAV5变体，所述组织类型包括皮肤、肝脏、肺、心脏、脾脏、淋巴结、甲状腺、骨骼肌、骨髓、乳腺、肾上腺、结肠、坐骨神经(周围神经)和许多CNS组织(前脑皮层、枕叶皮层、颞叶皮层、丘脑、下丘脑、黑质、海马DG、海马CA1、海马CA3、小脑和脊髓)。组织样本的分析使得能够富集特定变体的有利特性(例如，在第一组织中富集，而在一个或多个其他组织中富集)。应用机器学习算法来发现组织嗜性的决定因素。

[0025] 图16显示rAAV5变体包括大多数存在于CNS中的变体，以及在黑质中发现的变体，黑质是大脑的一个亚区域，在帕金森病中尤其受影响。此外，图16显示机器学习模型在预测高CNS特异性的决定因素方面表现出良好的性能(右)。

[0026] 图17显示rAAV5变体包括存在于肌肉组织中的变体，包括靶向心脏(心脏组织)和

骨骼组织而对肝脏组织去靶向(de-target)的变体(左图)。此外,图17显示机器学习模型在预测肝脏去靶向性的决定因素方面表现出良好的性能(右图)。

[0027] 图18A显示了相较于所有其他分析组织在CNS组织中氨基酸位置频率的热图。图18B显示了相较于所有其他分析组织在肝脏组织中氨基酸位置频率的热图。图18C显示了相较于所有其他分析组织在骨骼肌组织中氨基酸位置频率的热图。图18D显示了相较于所有其他分析组织在心脏组织中氨基酸位置频率的热图。图18E显示了相较于所有其他分析组织在肺组织中氨基酸位置频率的热图。图18F显示了相较于所有其他分析组织在脾脏组织中氨基酸位置频率的热图。图18G显示了相较于所有其他分析组织在淋巴结组织中氨基酸位置频率的热图。图18H显示了相较于所有其他分析组织在骨髓组织中氨基酸位置频率的热图。图18I显示了相较于所有其他分析组织在乳腺组织中氨基酸位置频率的热图。图18J显示了相较于所有其他分析组织在皮肤组织中氨基酸位置频率的热图。图18K显示了相较于所有其他分析组织在肾上腺组织中氨基酸位置频率的热图。图18L显示了相较于所有其他分析组织在甲状腺组织中氨基酸位置频率的热图。图18M显示了相较于所有其他分析组织在结肠组织中氨基酸位置频率的热图。图18N显示了相较于所有其他分析组织在坐骨神经组织中氨基酸位置频率的热图。图180显示了相较于所有其他分析组织在脊髓组织中氨基酸位置频率的热图。

[0028] 图19A显示了NGS数据的生物信息学预处理的各个步骤的结果。显示了通过过滤器的48个单独样本的逐步读取计数。图19B显示了CNS/非CNS变体序列的位置比较。

[0029] 图20A显示了通过本文所述的方法分析的周围和CNS组织样本的实例。图20B显示了相较于非CNS变体从CNS变体回收的前1000个机器学习预测/过滤的变体中的氨基酸位置丰度。

[0030] 图21A显示了如下两种ML模型的平均预测CNS概率的集成的实例:随机森林(RF)和基于直方图的梯度提升树(HGB)。两种ML模型的CNS靶向概率输出显示出良好的一致性。图21B显示了所有CNS变体的平均CNS靶向概率的分布。

[0031] 图22显示了机器学习模型性能验证的实例。基于直方图的梯度提升树(HGB)(上图)和随机森林(RF)(下图)机器学习模型在预测CNS靶向性方面表现出良好的性能。在高预测类别概率值下,两种机器学习模型从非CNS靶向性中解析CNS靶向性。

[0032] 图23显示了一组前20个位置特征如何促成模型输出概率。使用沙普利加和解释(SHAP)值来询问特征对模型预测的相对贡献。可以在组织靶向性和非靶向性变体之间进一步比较这些特征(如图24A至图24C所示)。

[0033] 图24A、图24B和图24C显示了对一组排名靠前的预测位置特征的比较。如果发现特征对HGB和RF模型都重要,则选择这些特征。显示了相较于随机2%CNS变体,前1000个机器学习预测的CNS变体中的特征的概述。

[0034] 图25显示了在靶向肝脏组织的序列的机器学习分析中一组对模型输出概率有贡献的前20个位置特征。

[0035] 图26A、图26B和图26C显示了在靶向肝脏组织的序列的机器学习分析中排名靠前的预测位置特征的比较。在组织靶向性和非靶向性变体之间比较特征。

[0036] 图27显示了在为肝脏去靶向性的序列的机器学习分析中一组对模型输出概率有贡献的前20个位置特征。

[0037] 图28A和图28B显示了肝脏去靶向性变体的排名靠前的预测位置特征的比较。

[0038] 图29显示了在靶向肌肉组织的序列的机器学习分析中一组对模型输出概率有贡献的前20个位置特征。

[0039] 图30A至图30B显示了在靶向肌肉组织的序列的机器学习分析中排名靠前的预测位置特征的比较。在组织靶向性和非靶向性变体之间比较特征。

6.具体实施方式

[0040] 除非另有描述,否则本文中使用的所有技术和科学术语具有与本发明所属领域的普通技术人员通常所理解的含义。

[0041] 除非另有说明,否则无论何时叙述一个范围,该范围都包括所叙述的端点。例如,SEQ ID NO:1的从氨基酸残基581至氨基酸残基589的区域包括氨基酸残基581和589。

[0042] “同源性”或“同一性”或“相似性”可以指两个肽之间或两个核酸分子之间的序列相似性。同源性可以通过比较各序列中可为比较目的来比对的位置而确定。当所比较序列中的位置可以被相同的碱基或氨基酸占据时,那么这些分子在那个位置上可以是同源的。序列之间的同源性程度可随着序列共有的匹配或同源位置的数目而变化。“无关”或“非同源的”序列与本公开的序列中之一共有小于40%的同一性,或替代地小于25%的同一性。同源性可以指序列与参考序列的同一性%。实际上,任何特定序列是否可以与本文所述的任何序列(其可以对应于本文所述的特定核酸序列)至少50%、60%、70%、80%、85%、90%、92%、95%、96%、97%、98%或99%同一,这种特定的多肽序列可以常规地使用已知的计算机程序,如Bestfit程序(Wisconsin Sequence Analysis Package,Version 8for Unix,Genetics Computer Group,University Research Park,575Science Drive,Madison,Wis.53711)来确定。当使用Bestfit或任何其他序列比对程序来确定特定序列是否与参考序列例如95%同一时,可以设置参数,使得在参考序列全长上计算同一性百分比,并且可以允许序列同源性中的空位高达总参考序列的5%。在两个或更多个核酸或多肽序列的上下文中,术语“同一性”百分比或“同源性”百分比是指当如使用下述序列比较算法(例如,BLASTP和BLASTN或本领域技术人员可用的其他算法)之一或通过目测检查比较和比对最大对应性时两个或更多个序列或子序列具有特定百分比的相同核苷酸或氨基酸残基。根据应用,“同一性”百分比可以存在于被比较序列的一个区域中,例如,存在于功能结构域中,或者,存在于要比较的两个序列的全长中。对于序列比较,通常一个序列充当与测试序列进行比较的参考序列。当使用序列比较算法时,将测试序列和参考序列输入计算机中,指定子序列坐标,并在必要时指定和序列算法程序参数。然后序列比较算法基于指定程序参数计算测试序列相对于参考序列的序列同一性百分比。出于本文的目的,使用描述于Altschul等人,J.Mol.Biol.215:403-410(1990)中的BLAST算法执行同一性和序列相似性百分比。用于执行BLAST分析的软件可通过National Center for Biotechnology Information(www.ncbi.nlm.nih.gov/)公开获得。

[0043] 在一些情况下,参考序列(查询序列,例如,本公开的序列)和主题序列之间的同一性,也称为全局序列比对,可以使用FASTDB计算机程序来确定。在一些实施方案中,在FASTDB氨基酸比对中使用的可狭义地解释同一性的具体实施方案的参数可以包括:评分方案=PAM(接受的突变的百分比)0,k-tuple=2,错配罚分=1,连接罚分=20,随机分组长度

＝0,截止值评分＝1,窗口大小＝序列长度,空位罚分＝5,空位大小罚分＝0.05,窗口大小＝500或主题序列长度,以较短者为准。根据该实施方案,如果主题序列可以由于N末端或C末端缺失而不是由于内部缺失而比查询序列更短,则可以对结果进行手动校正,以将FASTDB程序在计算全局同一性百分比时不考虑主题序列的N末端和C末端截短的事实纳入考虑。对于在N末端和C末端处截短的主题序列,相对于查询序列,可以通过计算可以横向于主题序列的N末端和C末端的查询序列的残基的数量(残基可以不与相应的主题残基匹配/对齐)将其作为查询序列的总碱基的百分比来校正同一性百分比。残基是否可以匹配/对齐的确定可以通过FASTDB序列比对的结果来确定。然后,可以从FASTDB程序使用指定参数计算的同一性百分比中减去该百分比,以得到最终的同一性百分比评分。该最终同一性百分比评分可以用于本实施方案的目的。在一些情况下,为了手动调整同一性百分比评分的目的,可以只考虑主题序列的N末端和C末端的残基,这些残基可以不与查询序列匹配/对齐。也就是说,对于该手动校正,可以只考虑在主题序列的最远N末端和C末端残基之外的查询残基位置。例如,可以将90个残基的主题序列与100个残基的查询序列比对,以确定同一性百分比。缺失发生在主题序列的N末端处,并因此FASTDB比对不显示N末端处的前10个残基的匹配/比对。10个未配对的残基表示序列的10％(N末端和C末端处未匹配的残基数/查询序列中的残基总数),因此可以从FASTDB程序计算的同一性百分比评分中减去10％。如果剩下的90个残基完全匹配,那么最终同一性百分比可以是90％。在另一实例中,90个残基的主题序列可以与100个残基的查询序列进行比较。这一次,缺失可以是内部缺失,因此在主题序列的N末端或C末端可以没有与查询不匹配/对齐的残基。在这种情况下,不能手动校正FASTDB计算的同一性百分比。同样,只有主题序列的N末端和C末端之外的残基位置(如FASTDB比对中显示的)可以手动校正,这些位置可以不与查询序列匹配/对齐。

[0044] 如本文所用,rAAV对组织的“嗜性”定义为给定rAAV优先感染给定细胞或组织的能力。经改变或工程化的嗜性包括对所需组织的靶向能力增强或降低,相应地对靶组织的感染增加或减少。

[0045] 在本公开全文中为简单起见,病毒衣壳蛋白通常称为“VP”。当引用AAV5 VP1位置符号时,病毒衣壳蛋白称为VP1。在所有情况下,本文公开的病毒衣壳序列和突变应理解为与衣壳蛋白的所有同种型(VP1、VP2和VP3)有关,因为这些同种型的混合物组装形成病毒体。位置氨基酸残基名称“581-589”与VP1多肽的翻译起始点有关,应根据VP2和VP3的相对起始位点进行相应调整。应当理解,当描述在特定氨基酸残基位置具有突变的任何特定VP1序列时,本公开必然还涵盖VP2和VP3中的相应突变。例如,在581-589区的氨基酸残基中具有一个或多个突变的VP1衣壳蛋白的任何共有序列或特定序列也涵盖在对应于VP1 581-589区的氨基酸残基的VP2和VP3中氨基酸残基区域中具有所述一个或多个突变的VP2和VP3衣壳蛋白。例如,VP1(SEQ ID NO:1)的581至589区域的氨基酸残基对应于VP2(SEQ ID NO:1115)的445至453区域的氨基酸残基和VP3(SEQ ID NO:1116)的389至397区域的氨基酸残基。

[0046] 应当理解,本公开包括编码本文公开的任何序列的多核苷酸序列。例如,如果提供氨基酸序列,则本公开还涵盖编码所述氨基酸序列的多核苷酸序列。

[0047] 应当理解,另外的实施方案包括VP1、VP2、VP3或它们的任何组合中不改变所需性质(例如,特定的组织嗜性)或影响病毒组装的突变,如本文所述。

[0048] 如本文所用,“组织嗜性”是指具有本公开的工程化VP衣壳多肽的病毒感染给定组织或在给定组织中富集或积聚的偏好。当用作相对术语并取决于本文描述其的上下文时,组织嗜性是指具有第一衣壳多肽的给定rAAV病毒体在第一组织中的组织嗜性相较于第二组织增加或减少并且/或者是指具有第一衣壳多肽的给定rAAV病毒体相较于具有第二衣壳多肽的rAAV病毒体的组织嗜性增加或减少。在一些实施方案中,第一组织可以是一组组织。在一些实施方案中,第二组织可以是一组组织。例如,第一组织可以是CNS组织,包括前脑皮层、枕叶皮层、颞叶皮层、丘脑、下丘脑、黑质、海马DG、海马CA1、海马CA3和小脑,并且第二组织可以是非CNS组织,由肝脏、骨骼肌、心脏、肺、脾脏、淋巴结、骨髓、乳腺、皮肤、肾上腺、甲状腺、结肠、坐骨神经和脊髓组织共同组成。再如,第一组织可以是肝脏组织,并且第二组织可以是非肝脏组织,由CNS组织、骨骼肌、心脏、肺、脾脏、淋巴结、骨髓、乳腺、皮肤、肾上腺、甲状腺、结肠、坐骨神经和脊髓组织共同组成。

[0049] 如本文所用,AAV衣壳多肽上下文中的词语“重组”可互换地指“工程化”或“变体”AAV衣壳多肽。如本文所用,AAV病毒体(缩写为rAAV)上下文中的词语“重组”是指重组病毒颗粒。所述rAAV病毒体由可以包括本文公开的工程化AAV5 VP衣壳多肽的衣壳制成。

[0050] 6.1.衣壳工程化方法

[0051] 本文公开了一种用于高通量工程化具有修饰功能的工程化AAV衣壳的系统,所述修饰功能包括对所需组织的感染性增加或减少,如肝脏嗜性增加或减少,或对中枢神经系统(CNS)的靶向性增加。该过程的一般示意图示于图1中,然而,应当理解,本公开还涵盖对本领域普通技术人员所理解的方法的合理变化或扩展。如图1所示,该方法可以从产生理论多样性为$5x10^{11}$个独特序列变体的衣壳文库开始。较高或更低的理论多样性也涵盖在本文中。例如,衣壳文库的理论多样性可以为约1x10^3至约1x10^20。然后可以将该文库克隆到质粒中,转化到细菌中,随后筛选文库质粒以在生产细胞系中进行生产性病毒体组装。然后可以将组装的病毒体静脉内施用到非人灵长类动物(NHP)中。在足以分布、感染和稳定转导的一段时间后,可以牺牲NHP,收获器官,并且可以通过深度测序确定每个组织中AAV衣壳的序列。

[0052] 图2A提供了原型AAV病毒体表面的侧视图(上图)和俯视图(下图),鉴定了已知AAV衣壳一包括AAV2、AAV5、AAV6和AAV9-的残基,这些残基已在研究文献中显示与靶细胞相互作用。这些与靶标相互作用的残基对应于AAV5 VP1衣壳蛋白中的氨基酸581-589。

[0053] 图2B显示了文库质粒的突出元件,示出了位于AAV ITR之间的rep和cap编码序列。在实施例1中进一步描述的所示实施方案中,将变异引入AAV5cap蛋白的各残基581-589(“文库变体区域”)。将20种天然氨基酸各自引入9个位置中的每一者,得到理论文库多样性为20^9(20^9;大约$5x10^{11}$)个独特序列变体。

[0054] 工程化的目标区域最有可能与靶细胞受体相互作用,并且在不破坏病毒体组装的情况下对变化相对耐受。与添加突出于病毒体3倍对称轴上方的非结构化肽的早期方法不同,当前方法引入了改变结合口袋特征的序列多样性。此外,这种方法可能会改变受体结合三聚体的总体结构,从而改变结合口袋外的变构相互作用(例如,AAVR PKD1)。引入的多样性是非随机的,从而减少了随机文库的错义和移码。

[0055] 通过将编码衣壳变体的多核苷酸克隆到包装的病毒基因组中(在ITR之间),具有变体衣壳的重组病毒体携带具有其同源突变的多核苷酸,因此可以通过对包装的病毒或感

染细胞进行测序来鉴定提供所需功能的独特变体。

[0056] 在一些实施方案中，衣壳是选自以下的衣壳：AAV1、AAV2、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAV9、AAV10、AAV11、AAV12、AAV13、AAV14、AAV15和AAV16、AAV.rh8、AAV.rh10、AAV.rh20、AAV.rh39、AAV.Rh74、AAV.RHM4-1、AAV.hu37、AAV.Anc80、AAV.Anc80L65、AAV.7m8、AAV.PHP.B、AAV2.5、AAV2tYF、AAV3B、AAV.LK03、AAV.HSC 1、AAV.HSC2、AAV.HSC3、AAV.HSC4、AAV.HSC5、AAV.HSC6、AAV.HSC7、AAV.HSC8、AAV.HSC9、AAV.HSC10、AAV.HSC11、AAV.HSC12、AAV.HSC13、AAV.HSC14、AAV.HSC15、AAV.HSC16或AAVhu68(描述于WO2020/033842，其以引用方式整体并入本文)。hu68衣壳描述于WO2018/160582，其以引用方式整体并入本文。

[0057] 此类衣壳可以包含对应于AAV5 VP1的581-589区的区域，因此本文涵盖具有所需组织嗜性、具有组装能力以及表现出各种其他所需性状的类似工程化VP衣壳。因此，在与AAV5 VP1的581至589区相对应的区域中具有突变的本文公开的任何一种工程化AAV5 VP衣壳多肽可以插入本文所述的任何其他AAV衣壳的相应区域中并且本公开涵盖此类变体。

[0058] 在一些实施方案中，衣壳是AAV1、AAV2、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAV9、AAV 10、AAV11、AAV 12、AAV 13、AAV 14、AAV 15和AAV 16、AAV.rh8、AAV.rh10、AAV.rh20、AAV.rh39、AAV.Rh74、AAV.RHM4-1、AAV.hu37、AAV.Anc80、AAV.Anc80L65、AAV.7m8、AAV.PHP.B、AAV2.5、AAV2tYF、AAV3B、AAV.LK03、AAV.HSC 1、AAV.HSC2、AAV.HSC3、AAV.HSC4、AAV.HSC5、AAV.HSC6、AAV.HSC7、AAV.HSC8、AAV.HSC9、AAV.HSC10、AAV.HSC11、AAV.HSC12、AAV.HSC13、AAV.HSC14、AAV.HSC15、AAV.HSC16或AAVhu68的衍生物、修饰或假型。

[0059] 在一些实施方案中，衣壳蛋白是来自两种或更多种选自以下的血清型的衣壳蛋白的嵌合体：AAV1、AAV2、rAAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAV9、AAV 10、AAV11、AAV 12、AAV13、AAV 14、AAV 15和AAV 16、AAV.rh8、AAV.rh10、AAV.rh20、AAV.rh39、AAV.Rh74、AAV.RHM4-1、AAV.hu37、AAV.Anc80、AAV.Anc80L65、AAV.7m8、AAV.PHP.B、AAV2.5、AAV2tYF、AAV3B、AAV.LK03、AAV.HSC 1、AAV.HSC2、AAV.HSC3、AAV.HSC4、AAV.HSC5、AAV.HSC6、AAV.HSC7、AAV.HSC8、AAV.HSC9、AAV.HSC10、AAV.HSC11、AAV.HSC12、AAV.HSC13、AAV.HSC14、AAV.HSC15和AAV.HSC16(描述于WO2020/033842，其以引用方式整体并入本文)。在某些实施方案中，衣壳是rh32.33衣壳，其描述于美国专利第8,999,678号，该专利以引用方式整体并入本文。

[0060] 此类衣壳可以包含对应于AAV5 VP1的581-589的区域，因此本文涵盖具有所需组织嗜性、具有组装能力以及表现出各种其他所需性状的类似工程化VP衣壳。

[0061] 6.2.编码VP的多核苷酸、载体和载体文库

[0062] 因此，在第一方面，提供了多核苷酸。多核苷酸编码腺相关病毒(AAV)VP1衣壳多肽，所述多肽具有SEQ ID NO:2的氨基酸序列，其中Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9各自独立地选自20种天然存在的氨基酸-使用标准的单字母代码-A、R、N、D、C、E、Q、G、H、I、L、K、M、F、P、S、T、W、Y和V。所述多肽包括天然AAV5衣壳的至少一个突变，因此不具有SEQ ID NO:1的序列。此外，所述多肽不具有SEQ ID NO:3、SEQ ID NO:NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7或SEQ ID NO:8的序列。

[0063] 在一些实施方案中，多核苷酸编码AAV VP1衣壳多肽，所述多肽相对于SEQ ID NO:1在581-589区域之外的氨基酸残基处还包含一个或多个突变，其中所得重组衣壳能够形成

组装的病毒体，所述组装的病毒体表现出所需的组织靶向性。

[0064] 在另一方面，提供了一种能够在原核细胞中复制的载体，其中所述载体包含上文刚刚描述的多核苷酸。在典型的实施方案中，所述载体是编码具有复制能力的AAV基因组的质粒。

[0065] 在另一方面，提供了一种文库。所述文库包含多个包含编码AAV衣壳的多核苷酸的载体。在一些实施方案中，所述载体是质粒，并且所述多个质粒包含多个不同的编码AAV VP的多核苷酸。

[0066] 在各个文库实施方案中，SEQ ID NO:2的Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9中的至少一者是不变的。此类不变残基在本文中也称为“框架”残基。框架残基可能有助于衣壳组装成功能性病毒体或感染特定靶细胞或组织的能力。

[0067] 在一些文库实施方案中，SEQ ID NO:2的Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9中的一个残基是不变的。在具体实施方案中，不变残基是AAV5 VP1中VP1一级氨基酸序列内该位置处的天然氨基酸。在具体实施方案中，不变残基是AAV5 VP1中该位置处的天然氨基酸以外的氨基酸。在一些实施方案中，SEQ ID NO:2的Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9中的两者是不变的。在具体实施方案中，各不变残基是AAV5 VP1中对应位置处的天然氨基酸。在具体实施方案中，各不变残基是AAV5VP1中对应位置处的天然氨基酸以外的氨基酸。在一些实施方案中，SEQ ID NO:2的Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9中的三者是不变的。在具体实施方案中，各不变残基是AAV5VP1中对应位置处的天然氨基酸。在具体实施方案中，各不变残基是AAV5 VP1中对应位置处的天然氨基酸以外的氨基酸。在一些实施方案中，SEQ ID NO:2的Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9中的四者是不变的。在具体实施方案中，各不变残基是AAV5 VP1中对应位置处的天然氨基酸。在具体实施方案中，各不变残基是AAV5 VP1中对应位置处的天然氨基酸以外的氨基酸。在一些实施方案中，SEQ ID NO:2的Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9中的五者是不变的。在具体实施方案中，各不变残基是AAV5 VP1中对应位置处的天然氨基酸。在具体实施方案中，各不变残基是AAV5 VP1中对应位置处的天然氨基酸以外的氨基酸。在一些实施方案中，SEQ ID NO:2的Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9中的六者是不变的。在具体实施方案中，各不变残基是AAV5 VP1中对应位置处的天然氨基酸。在具体实施方案中，各不变残基是AAV5 VP1中对应位置处的天然氨基酸以外的氨基酸。

[0068] 在具体实施方案中，rAAV VP1衣壳在位置587(Xaa7)处不是A、C、D、E、F、G、H、I、K、M、P、Q、R、V、W或Y。在一些实施方案中，位置587可以是N、S或T。在具体实施方案中，rAAV VP1衣壳在位置582(Xaa2)处不是G、V、L或I。

[0069] 在各个实施方案中，文库编码至少$1x10^{9}$种不同的AAV VP衣壳多肽、至少$2.5x10^{9}$种不同的AAV VP衣壳多肽、至少$5x10^{9}$种不同的AAV VP衣壳多肽、至少$7.5x10^{9}$种不同的AAV VP衣壳多肽、至少$1x10^{10}$种不同的AAV VP衣壳多肽、至少$2.5x10^{10}$种不同的AAV VP衣壳多肽、至少$5x10^{10}$种不同的AAV VP衣壳多肽、至少$7.5x10^{10}$种不同的AAV VP衣壳多肽、至少$1x10^{11}$种不同的AAV VP衣壳多肽、至少$2.5x10^{11}$种不同的AAV VP衣壳多肽，或至少$5x10^{11}$种不同的AAV VP衣壳多肽。

[0070] 在另一方面，提供了包含所述载体的原核细胞。在一些实施方案中，原核细胞是大

肠杆菌(E.coli)细胞并且载体是质粒。

[0071] 在相关方面,提供了文库,所述文库包含多个大肠杆菌细胞,其中所述多个细胞包含多个质粒,其中所述多个质粒包含多个不同的编码AAV VP的多核苷酸。

[0072] 在各个文库实施方案中,SEQ ID NO:2的Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9中的至少一者是不变的。

[0073] 在一些文库实施方案中,SEQ ID NO:2的Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9中的一者是不变的。在具体实施方案中,不变残基是AAV5 VP1中VP1一级氨基酸序列内该位置处的天然氨基酸。在具体实施方案中,不变残基是AAV5 VP1中该位置处的天然氨基酸以外的氨基酸。在一些实施方案中,SEQ ID NO:2的Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9中的两者是不变的。在具体实施方案中,各不变残基是AAV5 VP1中对应位置处的天然氨基酸。在具体实施方案中,各不变残基是AAV5VP1中对应位置处的天然氨基酸以外的氨基酸。在一些实施方案中,SEQ ID NO:2的Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9中的三者是不变的。在具体实施方案中,各不变残基是AAV5VP1中对应位置处的天然氨基酸。在具体实施方案中,各不变残基是AAV5 VP1中对应位置处的天然氨基酸以外的氨基酸。在一些实施方案中,SEQ ID NO:2的Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9中的四者是不变的。在具体实施方案中,各不变残基是AAV5 VP1中对应位置处的天然氨基酸。在具体实施方案中,各不变残基是AAV5 VP1中对应位置处的天然氨基酸以外的氨基酸。在一些实施方案中,SEQ ID NO:2的Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9中的五者是不变的。在具体实施方案中,各不变残基是AAV5 VP1中对应位置处的天然氨基酸。在具体实施方案中,各不变残基是AAV5 VP1中对应位置处的天然氨基酸以外的氨基酸。在一些实施方案中,SEQ ID NO:2的Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9中的六者是不变的。在具体实施方案中,各不变残基是AAV5 VP1中对应位置处的天然氨基酸。在具体实施方案中,各不变残基是AAV5 VP1中对应位置处的天然氨基酸以外的氨基酸。

[0074] 在一些实施方案中,文库编码至少$1x10^9$种不同的AAV VP衣壳多肽、至少$2.5x10^9$种不同的AAV VP衣壳多肽、至少$5x10^9$种不同的AAV VP衣壳多肽、至少$7.5x10^9$种不同的AAV VP衣壳多肽、至少$1x10^{10}$种不同的AAV VP衣壳多肽、至少$5x10^{10}$种不同的AAV VP衣壳多肽、至少$7.5x10^{10}$种不同的AAV VP衣壳多肽、至少$1x10^{11}$种不同的AAV VP衣壳多肽、至少$2.5x10^{11}$种不同的AAV VP衣壳多肽,或至少$5x10^{11}$种不同的AAV VP衣壳多肽。

[0075] 6.3.VP多肽、肽库

[0076] 在另一方面,提供了AAV VP1衣壳多肽。所述多肽具有SEQ ID NO:2的氨基酸序列,其中Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9各自独立地选自A、R、N、D、C、E、Q、G、H、I、L、K、M、F、P、S、T、W、Y和V。所述多肽相较于天然AAV VP1包括至少一个突变,因此不具有SEQ ID NO:1的序列。此外,所述多肽不具有SEQ ID NO:3、SEQ ID NO:NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7或SEQ ID NO:8的序列。

[0077] 在一些实施方案中,所述多肽相对于SEQ ID NO:1在581-589区域之外的氨基酸残基处还包含一个或多个突变,其中所得重组衣壳能够形成组装的病毒体,所述组装的病毒体表现出所需的组织靶向性。

[0078] 在另一方面,提供了文库,所述文库包含多个如上文刚刚描述的多肽,所述多个多

肽具有不同的一级氨基酸序列。

[0079] 在各个文库实施方案中，SEQ ID NO:2的Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9中的至少一者是不变的。

[0080] 在一些文库实施方案中，SEQ ID NO:2的Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9中的一者是不变的。此类不变残基在本文中也称为“框架”残基。在具体实施方案中，不变残基是AAV5 VP1中VP1一级氨基酸序列内该位置处的天然氨基酸。在具体实施方案中，不变残基是AAV5 VP1中该位置处的天然氨基酸以外的氨基酸。在一些实施方案中，SEQ ID NO:2的Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9中的两者是不变的。在具体实施方案中，各不变残基是AAV5 VP1中对应位置处的天然氨基酸。在具体实施方案中，各不变残基是AAV5 VP1中对应位置处的天然氨基酸以外的氨基酸。在一些实施方案中，SEQ ID NO:2的Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9中的三者是不变的。在具体实施方案中，各不变残基是AAV5 VP1中对应位置处的天然氨基酸。在具体实施方案中，各不变残基是AAV5 VP1中对应位置处的天然氨基酸以外的氨基酸。在一些实施方案中，SEQ ID NO:2的Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9中的四者是不变的。在具体实施方案中，各不变残基是AAV5 VP1中对应位置处的天然氨基酸。在具体实施方案中，各不变残基是AAV5 VP1中对应位置处的天然氨基酸以外的氨基酸。在一些实施方案中，SEQ ID NO:2的Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9中的五者是不变的。在具体实施方案中，各不变残基是AAV5 VP1中对应位置处的天然氨基酸。在具体实施方案中，各不变残基是AAV5VP1中对应位置处的天然氨基酸以外的氨基酸。在一些实施方案中，SEQ ID NO:2的Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9中的六者是不变的。在具体实施方案中，各不变残基是AAV5 VP1中对应位置处的天然氨基酸。在具体实施方案中，各不变残基是AAV5 VP1中对应位置处的天然氨基酸以外的氨基酸。

[0081] 在一些实施方案中，文库包含至少$1x10^{9}$种不同的AAV VP衣壳多肽、至少$2.5x10^{9}$种不同的AAV VP衣壳多肽、至少$5x10^{9}$种不同的AAV VP衣壳多肽、至少$7.5x10^{9}$种不同的AAV VP衣壳多肽、至少$1x10^{10}$种不同的AAV VP衣壳多肽、至少$2.5x10^{10}$种不同的AAV VP衣壳多肽、至少$5x10^{10}$种不同的AAV VP衣壳多肽、至少$7.5x10^{10}$种不同的AAV VP衣壳多肽、至少$1x10^{11}$种不同的AAV VP衣壳多肽、至少$2.5x10^{11}$种不同的AAV VP衣壳多肽，或至少$5x10^{11}$种不同的AAV VP衣壳多肽。

[0082] 在某些实施方案中，文库包含至少约$1x10^{5}$至至少约$5x10^{11}$种不同的AAV VP衣壳多肽。在某些实施方案中，文库包含至少约$1x10^{5}$、至少约$2x10^{5}$、至少约$3x10^{5}$、至少约$4x10^{5}$、至少约$5x10^{5}$、至少约$6x10^{5}$、至少约$7x10^{5}$、至少约$8x10^{5}$、至少约$9x10^{5}$、至少约$1x10^{6}$、至少约$2x10^{6}$、至少约$3x10^{6}$、至少约$4x10^{6}$、至少约$5x10^{6}$、至少约$6x10^{6}$、至少约$7x10^{6}$、至少约$8x10^{6}$、至少约$9x10^{6}$、至少约$1x10^{7}$、至少约$2x10^{7}$、至少约$3x10^{7}$、至少约$4x10^{7}$、至少约$5x10^{7}$、至少约$6x10^{7}$、至少约$7x10^{7}$、至少约$8x10^{7}$、至少约$9x10^{7}$、至少约$1x10^{8}$、至少约$2x10^{8}$、至少约$3x10^{8}$、至少约$4x10^{8}$、至少约$5x10^{8}$、至少约$6x10^{8}$、至少约$7x10^{8}$、至少约$8x10^{8}$、至少约$9x10^{8}$、至少约$1x10^{9}$、至少约$2x10^{9}$、至少约$3x10^{9}$、至少约$4x10^{9}$、至少约$5x10^{9}$、至少约$6x10^{9}$、至少约$7x10^{9}$、至少约$8x10^{9}$、至少约$9x10^{9}$、至少约$1x10^{10}$、至少约$2x10^{10}$、至少约$3x10^{10}$、至少约$4x10^{10}$、至少约$5x10^{10}$、至少约$6x10^{10}$、至少约$7x10^{10}$、至少约$8x10^{10}$、至少约$9x10^{10}$、至少约$1x10^{11}$、至少约$2x10^{11}$、至少约$3x10^{11}$、至少约$4x10^{11}$，或至少约$5x10^{11}$种AAV VP衣壳多肽。

[0083] 在某些实施方案中，本文提供了一种重组腺相关病毒AAV VP1衣壳多肽，所述多肽在SEQ ID NO:1中的581至残基589区域，包括端点在内，的残基中具有至少一个突变，其中与具有野生型AAV5VP1衣壳多肽的野生型AAV病毒体相比，所述突变赋予具有所述AAV VP1衣壳多肽的AAV病毒体在非肝脏组织中相较于肝脏组织中至少约两倍增加的积聚，并且其中AAV VP1衣壳多肽不具有SEQ ID NO:1、SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7和SEQ ID NO:8中任一者的序列。

[0084] 6.4.rAAV病毒体，病毒体文库

[0085] 在另一方面，提供了重组AAV病毒体(rAAV)。病毒体包含如上文所述的AAV VP衣壳多肽。

[0086] 在一些实施方案中，与具有天然AAV5 VP1衣壳多肽(SEQ ID NO:1)的rAAV相比，所述rAAV具有增加的对灵长类动物和人肝脏的嗜性。在一些实施方案中，与具有天然AAV5 VP1衣壳多肽(SEQ ID NO:1)的rAAV相比，所述rAAV具有增强的组装能力，或者表现出更大的病毒体稳定性。

[0087] 在一些实施方案中，与具有天然AAV5 VP1衣壳多肽(SEQ ID NO:1)的rAAV相比，所述rAAV具有降低的对人肝脏的嗜性。

[0088] 在一些实施方案中，与具有天然AAV5 VP1衣壳多肽(SEQ ID NO:1)的rAAV相比，所述rAAV在静脉内施用后具有增强的穿过血脑屏障的能力。

[0089] 在这些实施方案的某些中，与具有天然AAV5 VP1衣壳多肽(SEQ ID NO:1)的rAAV相比，所述rAAV在静脉内、鞘内、脑室内或脑池内施用后具有增强的感染选自海马、齿状回、大脑皮层、颞叶皮层、枕叶皮层、丘脑、前脑、黑质、下丘脑和小脑的一个或多个脑区的能力。

[0090] 在一些实施方案中，与具有天然AAV5 VP1衣壳多肽(SEQ ID NO:1)的rAAV相比，所述rAAV在静脉内、鞘内、脑室内或脑池内施用后具有增强的感染选自海马、齿状回、大脑皮层、颞叶皮层、枕叶皮层、丘脑、前脑、黑质、下丘脑和小脑的一个或多个脑区的能力，并且还具有降低的对所有非CNS组织的嗜性，包括对心脏组织去靶向。

[0091] 在一些实施方案中，与具有天然AAV5 VP1衣壳多肽(SEQ ID NO:1)的rAAV相比，所述rAAV在玻璃体内注射后具有增强的感染人视网膜细胞的能力。

[0092] 在一些实施方案中，与具有VP1衣壳多肽的具有天然AAV5 VP1衣壳多肽(SEQ ID NO:1)的rAAV相比，所述rAAV在静脉内施用后具有增强的感染人骨骼肌的能力。

[0093] 在一些实施方案中，与具有天然AAV5衣壳多肽(SEQ ID NO:1)的rAAV相比，所述rAAV在静脉内施用后具有增强的感染选自以下的组织的能力:脂肪、肾上腺、主动脉、脑(包括海马:齿状回、CA1和CA3;小脑、尾状核、壳核、中脑、脑桥、下丘脑、皮层-包括枕叶、颞叶和前脑;黑质和丘脑)、骨髓、盲肠、结肠、背根神经节、十二指肠、附睾、食管、眼睛、胆囊、心脏、回肠、空肠、肾脏、肺、淋巴结、乳腺、卵巢、胰腺、甲状旁腺、周围神经、垂体、前列腺、唾液腺、精囊、骨骼肌、皮肤、脊髓、脾脏、胃、睾丸、胸腺、甲状腺、气管、膀胱、子宫和阴道。

[0094] 此外，提供了编码本文所述的rAAV衣壳VP蛋白的多核苷酸序列。

[0095] 在另一方面，提供了包含多个如上文所述的rAAV的文库。所述多个rAAV包含多个具有不同一级氨基酸序列的VP衣壳多肽。

[0096] 在各个文库实施方案中，SEQ ID NO:2的Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9中的至少一者是不变的。

[0097] 在一些文库实施方案中，SEQ ID NO:2的Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9中的一者是不变的。此类不变残基在本文中也称为“框架”残基。在具体实施方案中，不变残基是AAV5 VP1中VP1一级氨基酸序列内该位置处的天然氨基酸。在具体实施方案中，不变残基是AAV5 VP1中该位置处的天然氨基酸以外的氨基酸。在一些实施方案中，SEQ ID NO:2的Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9中的两者是不变的。在具体实施方案中，各不变残基是AAV5 VP1中对应位置处的天然氨基酸。在具体实施方案中，各不变残基是AAV5 VP1中对应位置处的天然氨基酸以外的氨基酸。在一些实施方案中，SEQ ID NO:2的Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9中的三者是不变的。在具体实施方案中，各不变残基是AAV5 VP1中对应位置处的天然氨基酸。在具体实施方案中，各不变残基是AAV5 VP1中对应位置处的天然氨基酸以外的氨基酸。在一些实施方案中，SEQ ID NO:2的Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9中的四者是不变的。在具体实施方案中，各不变残基是AAV5 VP1中对应位置处的天然氨基酸。在具体实施方案中，各不变残基是AAV5 VP1中对应位置处的天然氨基酸以外的氨基酸。在一些实施方案中，SEQ ID NO:2的Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9中的五者是不变的。在具体实施方案中，各不变残基是AAV5 VP1中对应位置处的天然氨基酸。在具体实施方案中，各不变残基是AAV5VP1中对应位置处的天然氨基酸以外的氨基酸。在一些实施方案中，SEQ ID NO:2的Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9中的六者是不变的。在具体实施方案中，各不变残基是AAV5VP1中对应位置处的天然氨基酸。在具体实施方案中，各不变残基是AAV5 VP1中对应位置处的天然氨基酸以外的氨基酸。

[0098] 在一些实施方案中，文库包含至少$1x10^{9}$种不同的AAV VP衣壳多肽、至少$2.5x10^{9}$种不同的AAV VP衣壳多肽、至少$5x10^{9}$种不同的AAV VP衣壳多肽、至少$7.5x10^{9}$种不同的AAV VP衣壳多肽、至少$1x10^{10}$种不同的AAV VP衣壳多肽、至少$2.5x10^{10}$种不同的AAV VP衣壳多肽、至少$5x10^{10}$种不同的AAV VP衣壳多肽、至少$7.5x10^{10}$种不同的AAV VP衣壳多肽、至少$1x10^{11}$种不同的AAV VP衣壳多肽、至少$2.5x10^{11}$种不同的AAV VP衣壳多肽，或至少$5x10^{11}$种不同的AAV VP衣壳多肽。

[0099] 6.5.药物组合物

[0100] 在另一方面，提供了药物组合物。所述药物组合物包含如上文所述的rAAV和药学上可接受的载体。

[0101] 药物组合物可以包含第一活性成分。第一活性成分可以包含如本文所述的病毒载体和/或如本文所述的任何有效载荷。药物组合物可以配制成单位剂型。药物组合物可以包含药学上可接受的赋形剂、稀释剂或载体。药物组合物可以包含第二、第三或第四活性成分-例如用以促进增强的基因置换、RNA编辑、DNA编辑或成像。

[0102] 本文所述的药物组合物可以包含赋形剂。赋形剂可以包括冷冻防腐剂，如DMSO、甘油、聚乙烯吡咯烷酮(PVP)或它们的任何组合。赋形剂可以包括冷冻防腐剂，如蔗糖、海藻糖、淀粉，这些中的任一者的盐，这些中的任一者的衍生物，或它们的任何组合。赋形剂可以包括pH剂(以最大程度地减少组合物组分的氧化或降解)、稳定剂(以防止组合物组分的改性或降解)、缓冲剂(以增强温度稳定性)、增溶剂(以增加蛋白溶解度)，或它们的任何组合。赋形剂可以包括表面活性剂、糖、氨基酸、抗氧化剂、盐、非离子表面活性剂、增溶剂、甘油三酸酯、醇或它们的任何组合。赋形剂可以包括碳酸钠、乙酸盐、柠檬酸盐、磷酸盐，聚乙二醇

(PEG)、人血清白蛋白(HSA)、山梨糖醇、蔗糖、海藻糖、聚山梨醇酯80、磷酸钠、蔗糖、磷酸二钠、甘露醇、聚山梨醇酯20、组氨酸、柠檬酸盐、白蛋白、氢氧化钠、甘氨酸、柠檬酸钠、海藻糖、精氨酸、乙酸钠、乙酸盐、HCl、依地酸二钠、卵磷脂、甘油、黄原胶、大豆异黄酮、聚山梨醇酯80、乙醇、水，替普瑞酮(teprenone)或它们的任何组合。

[0103] 本文提供的组合物和方法可以利用药物组合物。通篇描述的组合物可以配制成药物并用于治疗经诊断患有疾病的有需要的人或哺乳动物。在一些情况下，可以预防性地使用药物组合物。

[0104] 本文提供的组合物可用于本文提供的方法中。本文提供的任何所提供组合物可用于本文提供的方法中。在一些情况下，一种方法包括至少部分地预防、减少、改善并且/或者治疗疾病或疾患，或疾病或疾患的症状。受试者可以是人或非人。受试者可以是哺乳动物(例如，大鼠、小鼠、牛、狗、猪、羊、马)。受试者可以是脊椎动物或无脊椎动物。受试者可以是实验动物。受试者可以是患者。受试者可能罹患疾病。受试者可能展现疾病的症状。受试者可能没有展现疾病的症状，但仍然患有疾病。受试者可以接受照护者的医疗护理(例如，受试者住院并由医师治疗)。

[0105] 6.6.治疗或检测方法

[0106] 在一些方面，本公开提供了使用具有本文公开的任何一种工程化AAV VP衣壳多肽序列的rAAV病毒体的治疗方法。在一些方面，本公开提供了使用具有本文公开的任何一种工程化AAV VP衣壳多肽序列的rAAV病毒体的检测方法。所述方法包括向有需要的受试者施用有效量的药物组合物，所述药物组合物包含具有本文公开的任何一种AAV VP衣壳多肽序列的rAAV病毒体。rAAV病毒体衣壳包裹(encapsidate)任何有效载荷，包括本文公开的那些有效载荷。

[0107] 在一些实施方案中，有效量为至少$1x10^8$个病毒基因组/剂量。在一些实施方案中，有效量为至少$5x10^8$个病毒基因组/剂量、$7.5x10^8$个病毒基因组/剂量、至少$1x10^9$个病毒基因组/剂量、至少$2.5x10^9$个病毒基因组/剂量、至少$5x10^9$个病毒基因组/剂量。

[0108] 在一些实施方案中，有效量为至少$1x10^{11}$个病毒基因组/kg患者体重、至少$5x10^{11}$个病毒基因组/kg、至少$1x10^{12}$个病毒基因组/kg、至少$5x10^{12}$个病毒基因组/kg、至少$1x10^{13}$个病毒基因组/kg、至少$1x10^{14}$个病毒基因组/kg，或至少$5x10^{14}$。

[0109] 在一些实施方案中，rAAV病毒体通过全身施用途径施用，包括肠内施用途径和肠胃外施用途径。可以静脉内施用rAAV病毒体。在一些实施方案中，可以肌肉内施用rAAV。在一些实施方案中，可以腹膜内施用rAAV。在一些实施方案中，可以局部施用rAAV。在一些实施方案中，可以口服施用rAAV。在具体实施方案中，静脉内施用rAAV病毒体。在一些实施方案中，鞘内施用rAAV。在一些实施方案中，通过脑室内注射来施用rAAV。在一些实施方案中，通过脑池内施用来施用rAAV。在一些实施方案中，通过玻璃体内注射来施用rAAV。

[0110] 在各个实施方案中，患者罹患下表1中列出的疾患之一。在具体实施方案中，患者罹患表1中列出的疾患之一，并且rAAV包括与表1中所列疾患相关的转基因产物。

[0111] 在一些实施方案中，具有本文公开的任何工程化AAV VP衣壳多肽序列的本公开的rAAV病毒体包含载体基因组，所述载体基因组包含治疗性多核苷酸或有效载荷。在另外的实施方案中，所述有效载荷可以处于引导受感染人细胞中的表达的调控序列的控制下。在一些实施方案中，有效载荷包含编码任何可遗传编码的有效载荷，如RNA(例如，指导RNA)、

阻抑tRNA、转基因或基因组修饰实体的治疗性多核苷酸。

[0112] 在一些实施方案中，治疗性多核苷酸编码指导RNA、tRNA、阻抑tRNA、siRNA、miRNA、mRNA、shRNA、环状RNA或反义寡核苷酸(ASO)、核酶、DNA酶(DNAzyme)、适体或它们的任何组合。在一些实施方案中，治疗性多核苷酸编码线性治疗性多核苷酸或环状治疗性多核苷酸。

[0113] 在一些实施方案中，治疗性多核苷酸编码治疗性蛋白质(转基因)。在具体实施方案中，转基因编码选自表1的适合修饰的靶标或转基因产物的蛋白质。

[0114]

表1
合适的治疗靶标

主要基因递送靶标	疾患	治疗性多核苷酸的靶标
脑/CNS	AADC 缺乏症	AADC
	阿尔茨海默氏病	多种，包括 APP、SNCA、MAPT、ApoE、NGF、TERT
	Tau 蛋白病	MAPT
	突触核蛋白病	SNCA
	贝登氏病 (Batten disease) (CLN2)	CLN2

[0115]

贝登氏病(CLN3)	CLN3
贝登氏病(CLN6)	CLN6
MPS-IIIB	NAGLU
伴有 GRN 突变的额颞叶痴呆(FTD-GRN)	GRN
伴有 GBA1 突变的帕金森氏病(PD-GBA)和神经性戈谢氏病(Gaucher's disease)	GBA1
突触核蛋白病	GBA1 + α-突触核蛋白
戈谢病 2 型	GBA
卡纳万病 (Canavan Disease)	ASPA
帕金森病	AADC
帕金森病	GDNF
帕金森病	神经秩蛋白(Neurturin)
帕金森病	GAD
帕金森病	NTN
帕金森病	hFOXG1
帕金森病	hKCNQ2
帕金森病	hFMR1
帕金森病	抗 Tau/miRNA
帕金森病	EPM2A 或 EPM2B
帕金森病	LRRK2
帕金森氏病	LRRK2
帕金森氏病	SNCA
泰萨克斯病(Tay-Sachs Disease)	HEXA
亨廷顿氏病(Huntington's disease)	IT15
亨廷顿氏病	CYP46A1
亨廷顿氏病	HTT
波托基-卢普斯基综合征 (Protocki-Lupski Syndrome)	IT15
肌萎缩侧索硬化症	C9orf72

[0116]

	肌萎缩侧索硬化症	SOD1
	唐氏综合症	DYRK1A
	圣菲利波病(Sanfilippo disease) A 型	SGSH
	圣菲利波病 B 型	hNAGLU
	(神经系统)	HEXB 和 HEXA
	(神经系统)	人密码子优化的 CLN1 互补 DNA
	(神经系统)	SURF1
	(神经系统)	抗 UBE3A-ATS shRNA
	(神经系统)	hSLC6A1
	雷特综合征(Rett syndrome)	MECP2
脊髓	脊髓性肌萎缩症(SMA)	SMN
	巨轴索神经病变	GAN
	慢性疼痛	Nav1.7
	脊髓小脑共济失调(SCA)，	ATXN3
眼睛	色盲	CNGB3
	脉络膜血症	REP1
	ad 色素性视网膜炎	NRL、RDH12、PRPH2(RDS)、RHO、RPGR、SNRNP200、NR2E3、IMPDH1、CRX、HK1、IMPDH2、SNRNP200
	斯塔加特病(Stargardt disease)	ABCA4
	乌谢尔综合征(Usher Syndrome) 2A	USH2A
	湿性 AMD，干性 AMD	NRP1
	莱伯先天性黑朦(LCA)	RPE65
	莱伯遗传性视神经病变(LHON)	ND4
	色素性视网膜炎(RP，包括 RLBP1)	RLBP1
	湿性 AMD	抗 VEGF 抗体

[0117]

		X 连锁色素性视网膜炎(X 连锁 RP)	RPGR
		X 连锁视网膜劈裂症	RS1
	肝脏	克 - 纳二氏综合症(Crigler–Najjar syndrome)	UGT1A1
		家族性高胆固醇血症(FH 纯合子)	LDLR
		糖原贮积病 1A 型(GSD1a)	G6PC
		甲型血友病	FVIII
		乙型血友病	FIX
		粘多糖贮积症 I (MPS-I)	ZFN1、ZFN2 和 IDUA 供体
		粘多糖贮积症 II (MPS-II)	ZFN1、ZFN2 和 IDS 供体
		粘多糖贮积症 IIIA (MPS-IIIA)	SGSH
		粘多糖贮积症 IIIB (MPS-IIIB)	NAGLU
		粘多糖贮积症 VI (MPS-VI)	ARSB
		羟化酶缺乏症	CYP21A2
		心血管疾病	PCSK9
		卟啉症和急性肝卟啉症	ALAS1
		血色素沉着症	HFE
		胆固醇酯贮积病	LIPA
		威尔逊病 (Wilson disease)	ATP7B
		成人多糖体病	GBE1 (也是肌肉细胞)
		肝脂肪变性	hSLC13A5
		α-1 抗胰蛋白酶缺乏症	SERPINA1
		鸟氨酸转氨甲酰酶缺乏症(OTC 缺乏症)	OTC
		α-1 抗胰蛋白酶缺乏症	A1AT

[0118]

	(A1AT 缺乏症)	
肌肉	恰克-马利-杜斯氏病(Charcot-Marie-Tooth Disease) 1A 型(CMT1A)	NTF3
	杜兴氏肌营养不良(Duchenne muscular dystrophy) (DMD)	小肌营养不良蛋白
	杜兴氏肌营养不良(DMD)	微肌营养不良蛋白
	Dysferlinopathy	DYSF
	庞贝病(Pompe disease)	GAA
	肢带型肌营养不良症(LGMD) (2i/R9)	FKRP
	杜兴氏肌营养不良(DMD)	DMD
	面肩肱型营养不良	DUX4
	肌强直性营养不良	DMPK
	糖原贮积症	抗 GYS1 miRNA
	X 连锁肌管性肌病(X 连锁 MTM)	MTM1
	常染色质组蛋白-赖氨酸 N-甲基转移酶 2	抗 EHMT2 shRNA
其他	(与听力损失相关)	TMC1
	肥胖(脂肪组织)	CIDEC、SCD1、GNB3
	骨(破骨细胞)	CLCN7
	软骨细胞	FGFR3
	原发性高草酸尿症 1 型(肾脏)	HAO1
	原发性高草酸尿症(肾脏)	LDHA
	肢端肥大症(多器官)	GHR
	哮喘(WBC; 中性粒细胞、嗜酸性粒细胞)	Mex3B
	奥尔波特综合征(Alport syndrome) (肾脏)	COL4A5

[0119]

	转甲状腺素蛋白淀粉样变性(家族性)(多器官)	TTR
	恰克-马利-杜斯氏综合征(PNS/坐骨神经;雪旺细胞(Schwann Cell))	PMP22
	安琪儿综合征(Angelman syndrome)(神经系统)	UBE3A
	充血性心力衰竭(心脏)	I-1c
	甲基丙二酸血症(MMA)(肾脏)	MMUT、MMAA、MMAB、MMADHC、MCEE
	囊性纤维化(肺)	CFTR
	HIV感染	PG9抗体
	HIV感染	VRC07抗体
贫血相关病症	血友病	F8(因子VIII)、F9(因子IX)
镰状细胞相关病症	镰状细胞性贫血	HBB
	镰状细胞血红蛋白C病	血红蛋白
	镰状细胞地中海贫血病	β地中海贫血

[0120] 在一些实施方案中,治疗性多核苷酸编码治疗性RNA。在一些实施方案中,治疗性多核苷酸编码RNA,如用于基因组编辑或RNA编辑的指导RNA(包括工程化或合成的指导RNA)。

[0121] 在一些实施方案中,治疗性多核苷酸编码tRNA或修饰的tRNA(工程化或合成的tRNA)。例如,tRNA或修饰的tRNA可以是阻抑tRNA。阻抑tRNA可以被设计成具有识别终止密码子如任何提前终止密码子(蛋白石、赭石或琥珀终止密码子)的反密码子区域。

[0122] 在一些实施方案中,治疗性多核苷酸(例如,治疗性RNA、tRNA或基因组修饰实体)可以靶向表1中列出的基因或与神经疾病、帕金森氏病、阿尔茨海默氏病、Tau蛋白病、斯塔加特病、α-1抗胰蛋白酶缺乏症、杜兴氏肌营养不良、雷特综合征、囊性纤维化或任何遗传疾病相关的任何基因。在一些实施方案中,靶向基因可以是ABCA4、AAT、SERPINA1、SERPINA1E342K、HEXA、LRRK2、SNCA、DMD、APP、Tau、GBA、PINK1、RAB7A、CFTR、ALAS 1、ATP7B、ATP7B G1226R、HFE C282Y、LIPAc.894G>A、PCSK9起始位点或SCNN1A起始位点,这些中任一者的片段,或它们的任何组合。在一些实施方案中,治疗性多核苷酸是基因疗法有效载荷(例如,转基因),因此,其本身可以是表1中列出的基因之一或与神经疾病、帕金森氏病、阿尔茨海默氏病、Tau蛋白病、斯塔加特病、α-1抗胰蛋白酶缺乏症、杜兴氏肌营养不良、雷特综合征、囊性纤维化或任何遗传疾病相关的任何基因。在一些实施方案中,转基因可以是

ABCA4、AAT、SERPINA1、SERPINA1E342K、HEXA、LRRK2、SNCA、DMD、APP、Tau、GBA、PINK1、RAB7A、CFTR、ALAS1、ATP7B、ATP7B G1226R、HFE C282Y、LIPA c.894G＞A、PCSK9起始位点或SCNN1A起始位点，这些中任一者的片段，或它们的任何组合。

[0123] 在一些实施方案中，治疗性多核苷酸编码基因组修饰实体。例如，基因组修饰实体可以是DNA编辑酶、RNA编辑酶、转录激活物或转录阻抑物。DNA编辑酶可以是任何DNA编辑酶，包括任何CRISPR/Cas系统、大范围核酸酶、锌指核酸酶(ZFN)、TALE核酸酶(TALEN和megaTALEN)。CRISPR/Cas系统可以是Cas3、Cas8、Cas10、Cas9、Cas4、Cas12或Cas13。RNA编辑酶可以是ADAR。在一些实施方案中，ADAR是人ADAR1或人ADAR2。转录激活物可以是VP64。转录阻抑物可以是KRAB。此类基因组修饰实体可以靶向表1中列出的任何基因进行编辑。

[0124] 在一些实施方案中，本公开提供了具有工程化AAV VP衣壳多肽的rAAV病毒体，其中所述病毒体衣壳包裹本文公开的治疗性有效载荷中的任一者或它们的组合。在一些实施方案中，治疗性有效载荷的多个拷贝被衣壳包裹。

[0125] 在一些实施方案中，治疗性多核苷酸是能够充当同源性模板用于同源定向修复的多核苷酸。

[0126] 在一些实施方案中，具有本文公开的任何工程化AAV VP衣壳多肽序列的本公开的rAAV病毒体包含载体基因组，所述载体基因组包含可检测多核苷酸或有效载荷。在另外的实施方案中，所述有效载荷可以处于引导受感染人细胞中的表达的调控序列的控制下。可检测多核苷酸的实例包括但不限于任何可遗传编码的可检测部分。例如，可遗传编码的可检测部分可以是荧光蛋白，如EGFP、GFP、YFP、RFP、CFP或它们的任何变体。在一些实施方案中，本公开提供了具有工程化AAV VP衣壳多肽的rAAV病毒体，其中所述病毒体衣壳包裹本文公开的可检测有效载荷中的任一者或它们的组合。在一些实施方案中，可检测有效载荷的多个拷贝被衣壳包裹。

[0127] 在一些实施方案中，本公开提供了具有工程化AAV VP衣壳多肽的rAAV病毒体，其中所述病毒体衣壳包裹本文公开的治疗性有效载荷和可检测有效载荷中的任一者或它们的组合。例如，具有工程化AAV VP衣壳多肽的本公开的rAAV可以衣壳包裹转基因和荧光蛋白。再如，具有工程化AAV VP衣壳多肽的本公开的rAAV可以衣壳包裹治疗性RNA(例如，指导RNA)和荧光蛋白。

[0128] 6.7.体内选定的VP多肽

[0129] 在另一方面，提供了使用本文所述的方法鉴定的工程化(同义地，重组)腺相关病毒(AAV)VP衣壳多肽。

[0130] 在一些实施方案中，工程化腺相关病毒(AAV)病毒蛋白(VP)衣壳多肽具有与SEQ ID NO:1至少70％同一的氨基酸序列，其中工程化AAV VP衣壳多肽相较于SEQ ID NO:1在SEQ ID NO:1的从残基581至残基589的区域，包括端点在内，中具有至少一个取代，其中衣壳多肽能够组装成重组AAV病毒体(rAAV)，并且其中VP衣壳多肽不具有SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7和SEQ ID NO:8中任一者的序列。

[0131] 在一些实施方案中，工程化腺相关病毒(AAV)病毒蛋白(VP)衣壳多肽具有与SEQ ID NO:1至少70％同一的氨基酸序列，其中工程化AAV VP衣壳多肽相较于SEQ ID NO:1在SEQ ID NO:1的从残基581至残基589的区域，包括端点在内，中具有至少一个取代，其中衣壳多肽能够组装成重组AAV病毒体(rAAV)，其中与具有SEQ ID NO:1的AAV5 VP衣壳多肽的

rAAV病毒体相比，所述至少一个取代赋予所述rAAV较高的对中枢神经系统(CNS)组织的嗜性，并且其中VP衣壳多肽不具有SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7和SEQ ID NO:8中任一者的序列。

[0132] 在具体实施方案中，工程化腺相关病毒(AAV)病毒蛋白(VP)衣壳多肽具有与SEQ ID NO:1的序列至少75％、80％、85％、90％、91％、92％、93％、94％、95％、96％、97％、98％同一的氨基酸序列。

[0133] 在一些实施方案中，AAV VP衣壳多肽具有SEQ ID NO:2的氨基酸序列，其中Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9各自独立地选自任何氨基酸，其中衣壳多肽能够组装成重组AAV病毒体(rAAV)，并且其中VP衣壳多肽不具有SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7和SEQ ID NO:8中任一者的序列，任选地在VP衣壳多肽的别处具有另外的突变。

[0134] 在一些实施方案中，AAV VP衣壳多肽具有SEQ ID NO:2的氨基酸序列，其中Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9各自独立地选自任何氨基酸，其中衣壳多肽能够组装成重组AAV病毒体(rAAV)，其中与具有SEQ ID NO:1的AAV5 VP衣壳多肽的rAAV病毒体相比，所述至少一个取代赋予所述rAAV较高的对中枢神经系统(CNS)组织的嗜性，并且其中VP衣壳多肽不具有SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7和SEQ ID NO:8中任一者的序列，任选地在VP衣壳多肽的别处具有另外的突变。

[0135] 在一些实施方案中，工程化腺相关病毒(AAV)病毒蛋白(VP)衣壳多肽具有SEQ ID NO:2的氨基酸序列，其中氨基酸残基Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9各自独立地选自A、R、N、D、C、E、Q、G、H、I、L、K、M、F、P、S、T、W、Y和V；其中工程化AAV VP衣壳多肽能够组装成重组AAV病毒体(rAAV)；并且其中rAAV VP衣壳多肽不具有SEQ ID NO:1、SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7和SEQ ID NO:8中任一者的序列。

[0136] 在一些实施方案中，工程化VP衣壳多肽的残基581至残基589的区域，包括端点在内，具有与SEQ ID NO:7118-SEQ ID NO:10,117中任一者至少70％、至少75％、至少80％、至少85％、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％同一的序列。

[0137] 在具体实施方案中，工程化VP衣壳多肽的残基581至残基589的区域，包括端点在内，具有与SEQ ID NO:7118-SEQ ID NO:10,117中任一者同一的序列。

[0138] 在一些实施方案中，工程化腺相关病毒(AAV)病毒蛋白(VP)衣壳多肽是工程化AAV5衣壳蛋白，其中工程化AAV VP5衣壳多肽相较于SEQ ID NO:1在SEQ ID NO:1的从残基581至残基589的区域，包括端点在内，中具有至少一个取代；其中衣壳多肽能够组装成重组AAV病毒体(rAAV)；其中与具有SEQ ID NO:1的AAV5 VP衣壳多肽的rAAV病毒体相比，所述至少一个取代赋予所述rAAV较高的对中枢神经系统(CNS)组织的嗜性，并且其中VP衣壳多肽不具有SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7和SEQ ID NO:8中任一者的序列，任选地在VP蛋白的别处具有另外的突变。

[0139] 在一些实施方案中，AAV VP衣壳多肽具有SEQ ID NO:2的氨，基酸序列，其中Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9各自独立地选自A、R、N、D、C、E、Q、G、H、I、L、K、M、F、P、S、T、W、Y和V；并且其中多肽不具有SEQ ID NO:1、SEQ ID NO:3、SEQ ID NO:4、SEQ

ID NO:5、SEQ ID NO:6、SEQ ID NO:7和SEQ ID NO:8中任一者的序列。

[0140] 在一些实施方案中,工程化AAV VP衣壳多肽包含由下式表示的多肽序列:(A)-(X)-(B)

[0141] 其中:

[0142] (A)是SEQ ID NO:47438(VAYNVGGQMATNNQSSTTAP SEQ ID NO:2的残基561-580)的多肽序列;

[0143] (X)是包含SEQ ID NO:2的氨基酸残基Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9的多肽序列;并且

[0144] (B)是SEQ ID NO:47439(SEQ ID NO:2的残基590-609IVPGSVWMERDVYLQGPIWA)的多肽序列;

[0145] 其中Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9各自独立地选自A、R、N、D、C、E、Q、G、H、I、L、K、M、F、P、S、T、W、Y和V;并且其中衣壳多肽能够组装成重组AAV病毒体(rAAV);并且;

[0146] 其中多肽不具有SEQ ID NO:1、SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7和SEQ ID NO:8中任一者的序列。

[0147] 在一些实施方案中,工程化AAV VP衣壳多肽包含由下式表示的多肽序列:(A)-(X)-(B),其中:

[0148] (A)是SEQ ID NO:47438(SEQ ID NO:2的残基561-580VAYNVGGQMATNNQSSTTAP)的多肽序列;

[0149] (X)是选自表8(SEQ ID NO:115-1114)或表1 0(SEQ ID NO:7118-8117)中的多肽的列表的多肽序列,所述多肽序列赋予重组AAV病毒体(rAAV)CNS组织嗜性;并且

[0150] (B)是SEQ ID NO:47439(SEQ ID NO:2的残基590-609(IVPGSVWMERDVYLQGPIWA))的多肽序列;并且

[0151] 其中衣壳多肽能够组装成rAAV,并且,

[0152] 衣壳不具有SEQ ID NO:1、SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7和SEQ ID NO:8中任一者的序列。

[0153] 在一些实施方案中,工程化AAV VP衣壳多肽赋予CNS组织嗜性,其中CNS组织选自由以下组成的组:海马:(齿状回,CA1和CA3);小脑、下丘脑、皮层:(枕叶、颞叶和前脑);黑质、丘脑,以及它们的任何组合。

[0154] 在一些实施方案中,工程化AAV VP衣壳多肽包含由下式表示的多肽序列:(A)-(X)-(B),其中:

[0155] (A)是SEQ ID NO:47438(SEQ ID NO:2的残基:(561-580VAYNVGGQMATNNQSSTTAP))的多肽序列;

[0156] (X)是选自SEQ ID NO:115-1114或SEQ ID NO:1118-47437的多肽的多肽序列,所述多肽序列赋予重组AAV病毒体(rAAV)相应的组织嗜性;并且

[0157] (B)是SEQ ID NO:47439(SEQ ID NO:2的残基590-609(IVPGSVWMERDVYLQGPIWA))的多肽序列;并且

[0158] 其中衣壳多肽能够组装成rAAV,并且,

[0159] 衣壳不具有SEQ ID NO:1、SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:

6、SEQ ID NO:7和SEQ ID NO:8中任一者的序列。

[0160] 下文描述的是工程突变的AAV5 VP1多肽序列,其赋予稳定或改进的病毒体组装、组织嗜性或两者。在一些实施方案中,本公开提供了与SEQ ID NO:1具有不超过98.7%序列同源性的AAV5 VP1衣壳多肽,其中所述AAV5衣壳多肽在从对应于SEQ ID NO:1的581的位置至对应于589的位置的区域中具有至少一个突变。

[0161] 本文还涵盖由工程化AAV5 VP2衣壳多肽和工程化AAV5 VP3衣壳多肽组成的rAAV,所述多肽在对应于AAV5 VP1581至589区域中的氨基酸的AAV5 VP2(氨基酸残基445至453)和AAV5 VP3(氨基酸残基389-397)中的区域处具有实施例的表(例如,表7、10、12、14、16、18、20、22、24、26、28、30、32、34、36、39、41、43、45、47、49、51、53、55、57、59、61、63、65、67、69、71、72、73、74、75、76、77、78、79、80、81、82、83、84、85和86)中公开的序列。

[0162] 6.7.1.体内选定的赋予增加的肝脏嗜性的VP多肽

[0163] 在各个实施方案中,本公开提供了一种能够形成组装的病毒体的突变VP多肽,与包含SEQ ID NO:1的AAV5 VP1衣壳多肽的病毒体相比,所述病毒体表现出增加的对肝脏组织的嗜性。肝脏组织嗜性由相对于给定氨基酸残基在施用于NHP的总病毒文库的频率而言该给定氨基酸残基出现在对应于SEQ ID NO:1(推及于SEQ ID NO:2)的位置581至位置589的指定位置的频率决定。

[0164] 在一些实施方案中,Xaa1选自A、G、K、M、N、Q、R、S或T。

[0165] 在一些实施方案中,Xaa1选自A、K、M或T。

[0166] 在一些实施方案中,Xaa1为K。

[0167] 在另外的实施方案中,Xaa2选自A、C、H、I、K、S、T或V。

[0168] 在一些实施方案中,Xaa2选自A、S、T或V。

[0169] 在一些实施方案中,Xaa2为T。

[0170] 在另外的实施方案中,Xaa3选自A、G、H、K、M,N、Q、R、S、T或V。

[0171] 在一些实施方案中,Xaa3选自A、M或T。

[0172] 在一些实施方案中,Xaa3为A或T。

[0173] 在另外的实施方案中,Xaa4选自L、M,P、Q、R、T或W。

[0174] 在一些实施方案中,Xaa4选自L、P、Q或T。

[0175] 在一些实施方案中,Xaa4为P。

[0176] 在另外的实施方案中,Xaa5选自F、H、I、K、M,T或Y。

[0177] 在一些实施方案中,Xaa5选自H、I或Y。

[0178] 在一些实施方案中,Xaa5为Y。

[0179] 在另外的实施方案中,Xaa6选自E、G、H、L、M,N、Q、T或W。

[0180] 在一些实施方案中,Xaa6选自N或Q。

[0181] 在一些实施方案中,Xaa6为N。

[0182] 在另外的实施方案中,Xaa7选自A、C、G、H、L、M,R或S。

[0183] 在一些实施方案中,Xaa7选自A、C、H或M。

[0184] 在一些实施方案中,Xaa7为A。

[0185] 在另外的实施方案中,Xaa8选自A、C、D、F、G、H、M,Q、S、V、W或Y。

[0186] 在一些实施方案中,Xaa8选自G、M,Q或S。

[0187] 在一些实施方案中,Xaa8为G。

[0188] 在另外的实施方案中,Xaa9选自A、C、E、G、H、M,N、P、Q、S、V或W。

[0189] 在一些实施方案中,Xaa9选自E、G或P。

[0190] 在一些实施方案中,Xaa9为G。

[0191] 在具体实施方案中,工程化(重组)衣壳多肽的Xaa1-Xaa9序列选自表2中提供的氨基酸序列。

[0192] 在一些实施方案中,当与AAV5野生型衣壳和相应病毒体相比时,工程化AAV衣壳和相应病毒体表现出增加的肝脏嗜性。当与包含SEQ ID NO:1的AAV5 VP1衣壳多肽的病毒体相比时,此增加的嗜性可以在约1.5、2.0、2.5、3.0、3.5、4.0、4.5、5.0、5.5、6.0、6.5、7.0、7.5、8.0、8.5、9.0、9.5至约10.0倍的范围内。

[0193]

表2
驱动肝脏嗜性的AAV5 VP1衣壳多肽中581至589区域的序列

变体编号	翻译
SEQ ID NO: 10	IAVASHAHG
SEQ ID NO: 11	SAYPKSEDV
SEQ ID NO: 12	SSMQCAMRP
SEQ ID NO: 13	ITVVVGEVN
SEQ ID NO: 14	SEAGDRCSD
SEQ ID NO: 15	ECACGHVHL
SEQ ID NO: 16	TFTNAGKMC
SEQ ID NO: 17	LASPYLIDA
SEQ ID NO: 18	TTQLYEVGS
SEQ ID NO: 19	TVHPIDHAT
SEQ ID NO: 20	TWMMVHRYV
SEQ ID NO: 21	VWYRCDVRH
SEQ ID NO: 22	HWADKALDW
SEQ ID NO: 23	MHTMKNCGW
SEQ ID NO: 24	LCNRMQADQ
SEQ ID NO: 25	WKATPGGQC
SEQ ID NO: 26	ARHTFPYQT
SEQ ID NO: 27	MHMSYHWRE
SEQ ID NO: 28	MPKAEIFVP
SEQ ID NO: 29	DKTICQHQL
SEQ ID NO: 30	TTRGSERCS
SEQ ID NO: 31	HTPLAWHGD
SEQ ID NO: 32	SCTPFQQHG
SEQ ID NO: 33	AHQLAMVMN
SEQ ID NO: 34	AAAQRSRYV
SEQ ID NO: 35	DCMTSMLGP
SEQ ID NO: 36	RYLFAVWRY
SEQ ID NO: 37	MMNMQMDRM
SEQ ID NO: 38	AVQWIQWVG
SEQ ID NO: 39	QSQPHVQYN
SEQ ID NO: 40	THTFAYHNN
SEQ ID NO: 41	GWQDEQMHL
SEQ ID NO: 42	TSKQETPQQ
SEQ ID NO: 43	CSNSPACLC
SEQ ID NO: 44	VDAHTETNG
SEQ ID NO: 45	CNEKGYHWH
SEQ ID NO: 46	IADIGRMQW
SEQ ID NO: 47	TEWPYGVAF
SEQ ID NO: 48	QCQYQTAWW
SEQ ID NO: 49	GPSGIFHCG
SEQ ID NO: 50	MAHAIDALQ
SEQ ID NO: 51	MHHCYQTFA
SEQ ID NO: 52	GQTVFHYDG
SEQ ID NO: 53	CNAIIAMPC
SEQ ID NO: 54	GTITLMPTQ
SEQ ID NO: 55	TVQCEMDVC
SEQ ID NO: 56	CQHNTPFVQ
SEQ ID NO: 57	RWTDTNFRG
SEQ ID NO: 58	KYEQKWMMD
SEQ ID NO: 59	MLYSTTCWK
SEQ ID NO: 60	ASWVGSFAQ
SEQ ID NO: 61	TTKQYHTME
SEQ ID NO: 62	PGGCYRMME
SEQ ID NO: 63	RNAVCLYRH
SEQ ID NO: 64	VIHEMLSAC
SEQ ID NO: 65	YGEQKYNHS
SEQ ID NO: 66	ANQWENAVK
SEQ ID NO: 67	TSAWWWCSP
SEQ ID NO: 68	EQTMMDLVY
SEQ ID NO: 69	YAMPRNPNV
SEQ ID NO: 70	MHKVWWMKN
SEQ ID NO: 71	AILQMYTAQ
SEQ ID NO: 72	CSTPANSCP
SEQ ID NO: 73	KVVEAELWQ
SEQ ID NO: 74	IANIDPSRV
SEQ ID NO: 75	QLTTTKPLR
SEQ ID NO: 76	ACWRFECDD
SEQ ID NO: 77	QINPEHGGC
SEQ ID NO: 78	MAPPFHVYE
SEQ ID NO: 79	DNWGTWTWT
SEQ ID NO: 80	DESSCVAHG
SEQ ID NO: 81	VAATFNNKV
SEQ ID NO: 82	HVVPEPWAF
SEQ ID NO: 83	TQMPALAEW
SEQ ID NO: 84	NVTNQQYDH
SEQ ID NO: 85	NMTEYVAYR
SEQ ID NO: 86	NVSTVHQPL
SEQ ID NO: 87	QTGPMLHSW
SEQ ID NO: 88	ECMHKHASY
SEQ ID NO: 89	LQAHVVRCT

[0194]

SEQ ID NO: 90	VGDRYSSMG
SEQ ID NO: 91	PQGLIPMWA
SEQ ID NO: 92	MYVHKGYRS
SEQ ID NO: 93	AVPQYQKAE
SEQ ID NO: 94	ERMMILCSP
SEQ ID NO: 95	NFGFTCPVY
SEQ ID NO: 96	SQIWNVAAY
SEQ ID NO: 97	MWGQQGTWA
SEQ ID NO: 98	QAMMMTMMN
SEQ ID NO: 99	AHTANEFSP
SEQ ID NO: 100	DAHYVYEKG

SEQ ID NO: 101	CNNWIWAHE
SEQ ID NO: 102	NHNLMWVVS
SEQ ID NO: 103	ATMWGDCDY
SEQ ID NO: 104	EWMQEFAGP
SEQ ID NO: 105	QDGSVEWAF
SEQ ID NO: 106	WCPQPPGGN
SEQ ID NO: 107	AECQIWYDW
SEQ ID NO: 108	NAVKFVCED
SEQ ID NO: 109	TQCFASCVA
SEQ ID NO: 110	TVNNHDIGY

[0195] 6.7.2.体内选定的能够组装成rAAV的工程化VP多肽

[0196] 在各个优选的实施方案中,本公开的突变(工程化,重组)VP衣壳多肽能够形成组装的病毒体,并且在一些情况下在与包含SEQ ID NO:1的AAV5 VP1衣壳多肽的病毒体相比时表现出相似或提高的稳定性。

[0197] 分析相对于给定氨基酸残基出现在整个质粒文库中的指定位置的频率而言该给定氨基酸残基出现在组装的纯化病毒中对应于SEQ ID NO:1(推及于SEQ ID NO:2)的位置581至位置589的所述指定位置的频率,以确定优先病毒组装的衣壳的序列规则。

[0198] 本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽,所述多肽与野生型AAV5 VP衣壳多肽相比可以表现出相似或提高的稳定性,其中工程化变体AAV5 VP衣壳多肽序列具有一个或多个突变,其中VP1多肽序列在从对应于SEQ ID NO:2中的581的位置至对应于SEQ ID NO:2中的589的位置的区域中具有所述一个或多个突变,并且其中Xaa1选自A、D、E、G、L、M,N、Q、S、T或V,或者Xaa1选自A、D、E、M或T。在一些实施方案中,Xaa1为E;或Xaa2选自A、C、D、E、G、H、I、N、P、Q、S、T或V,或Xaa2选自A、S、T或V,或Xaa2为A;或者其中Xaa3选自A、D、E、G、H、M,N、Q、S、T或V,或Xaa3选自D、E、N、Q或T,或Xaa3为D或T;或者其中Xaa4选自A、D、E、G、H、N、P、Q、S或T,或Xaa4选自D、E、P或Q,或Xaa4为E;或者其中Xaa5选自A、C、D、E、G、H、N、Q、S、T或Y,或Xaa5选自D、E、N、Q或T,或Xaa5为N;或者其中Xaa6选自A、D、E、G、H、N、P、Q、S或T,或Xaa6选自D、N或Q,或Xaa6为D;或者其中Xaa7选自A、C、D、E、G、H、N、Q、S或T,或Xaa7选自A、D、E或G,或Xaa7为A;或者其中Xaa8选自A、C、D、E、G、H、N、Q、S或T,或Xaa8包括A、D、G或S,或Xaa8为G;或者其中Xaa9选自A、D、E、G、H、N、P、Q、S或T,或Xaa9选自A、D、G或P,或Xaa9为G。

[0199] 在各个实施方案中,VP多肽能够形成组装的病毒体,并且在一些情况下在与包含SEQ ID NO:1的AAV5 VP1衣壳多肽的病毒体相比时表现出相似或提高的稳定性。

[0200] 在一些实施方案中,Xaa1选自A、D、E、G、L、M,N、Q、S、T或V。

[0201] 在一些实施方案中,Xaa1选自A、D、E、M或T。在一些实施方案中,Xaa1为E。

[0202] 在一些实施方案中,Xaa2选自A、C、D、E、G、H、I、N、P、Q、S、T或V。在一些实施方案中,Xaa2选自A、S、T或V。在一些实施方案中,Xaa2为A。

[0203] 在一些实施方案中,Xaa3选自A、D、E、G、H、M,N、Q、S、T或V。在一些实施方案中,Xaa3选自D、E、N、Q或T。在一些实施方案中,Xaa3为D或T。

[0204] 在一些实施方案中,Xaa4选自A、D、E、G、H、N、P、Q、S或T。在一些实施方案中,Xaa4选自D、E、P或Q。在一些实施方案中,Xaa4为E。

[0205] 在一些实施方案中，Xaa5选自A、C、D、E、G、H、N、Q、S、T或Y。在一些实施方案中，Xaa5选自D、E、N、Q或T。在一些实施方案中，Xaa5为N。

[0206] 在一些实施方案中，Xaa6选自A、D、E、G、H、N、P、Q、S或T。在一些实施方案中，Xaa6选自D、N或Q。在一些实施方案中，Xaa6为D。

[0207] 在一些实施方案中，Xaa7选自A、C、D、E、G、H、N、Q、S或T。在一些实施方案中，Xaa7选自A、D、E或G。在一些实施方案中，Xaa7为A。

[0208] 在一些实施方案中，Xaa8选自A、C、D、E、G、H、N、Q、S或T。在一些实施方案中，Xaa8包括A、D、G或S。在一些实施方案中，Xaa8为G。

[0209] 在一些实施方案中，Xaa9选自A、D、E、G、H、N、P、Q、S或T。在一些实施方案中，Xaa9选自A、D、G或P。在一些实施方案中，Xaa9为G。

[0210] 6.7.3.体内选定的能够组装成rAAV病毒体并且表现出降低的肝脏嗜性的突变VP多肽

[0211] 本公开提供了具有VP衣壳多肽的AAV5病毒体，所述VP衣壳多肽在具有与靶细胞相互作用的残基的区域中具有至少一个突变，其中与野生型VP衣壳多肽相比，所述至少一个突变赋予降低的肝脏组织嗜性。在一些实施方案中，本文提供了与SEQ ID NO:1具有至少80％序列同源性的AAV5 VP1衣壳多肽，其中所述AAV5 VP1衣壳多肽在从对应于SEQ ID NO:1的581的位置至对应于589的位置的区域中具有至少一个突变，并且其中所述至少一个突变驱动降低的肝脏嗜性。

[0212] 分析相对于给定氨基酸残基出现在形成组装病毒的变体中的指定位置的频率而言该给定氨基酸残基出现在肝脏中未鉴定的变体中对应于SEQ ID NO:1(推及于SEQ ID NO:2)的位置581至位置589的所述指定位置的频率，以确定优先对肝脏组织去靶向的衣壳的一组序列规则。

[0213] 本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽，所述多肽与野生型AAV5 VP衣壳多肽相比表现出降低的肝脏组织嗜性，其中工程化变体AAV5 VP衣壳多肽序列具有一个或多个突变，其中VP1多肽序列在从对应于SEQ ID NO:2中的581的位置至对应于SEQ ID NO:2中的589的位置的区域中具有所述一个或多个突变，并且其中Xaa1不是K，或Xaa1不是A、K、M或T，或Xaa1不是A、G、K、M、N、Q、R、S或T；或者其中Xaa2不是T，或Xaa2不是A、S、T或V，或Xaa2不是A、C、H、I、K、S、T或V；或者其中Xaa3不是A或T，或Xaa3不是A、M或T，或Xaa3不是A、G、H、K、M、N、Q、R、S、T或V；或者其中Xaa4不是P，或者其中Xaa4不是L、P、Q或T，或Xaa4不是L、M、P、Q、R、T或W；或者其中Xaa5不是Y，或Xaa5不是H、I或Y，或Xaa5不是F、H、I、K、M、T或Y；或者其中Xaa6不是N，或Xaa6不是N或Q，或Xaa6不是E、G、H、L、M、N、Q、T或W；或者其中Xaa7不是A，或Xaa7不是A、C、H或M，或Xaa7不是A、C、G、H、L、M、R或S；或者其中Xaa8不是G，或Xaa8不是G、M、Q或S，或Xaa8不是A、C、D、F、G、H、M、Q、S、V、W或Y；或者其中Xaa9不是G，或Xaa9不是E、G或P，或Xaa9不是A、C、E、G、H、M、N、P、Q、S、V或W。

[0214] 在某些实施方案中，Xaa1不是K，并且其中VP衣壳能够形成组装的病毒体，与包含SEQ ID NO:1的AAV5 VP1衣壳的病毒体相比，所述组装的病毒体表现出降低的对肝脏组织的嗜性。

[0215] 在某些实施方案中，Xaa1不是A、K、M或T，并且其中VP衣壳能够形成组装的病毒体，与包含SEQ ID NO:1的AAV5 VP1衣壳的病毒体相比，所述组装的病毒体表现出降低的对肝

脏组织的嗜性。

[0216] 在某些实施方案中，Xaa1不是A、G、K、M、N、Q、R、S或T，并且其中VP衣壳能够形成组装的病毒体，与包含SEQ ID NO:1的AAV5 VP1衣壳的病毒体相比，所述组装的病毒体表现出降低的对肝脏组织的嗜性。

[0217] 在某些实施方案中，Xaa2不是T，并且其中VP衣壳能够形成组装的病毒体，与包含SEQ ID NO:1的AAV5 VP1衣壳的病毒体相比，所述组装的病毒体表现出降低的对肝脏组织的嗜性。

[0218] 在某些实施方案中，Xaa2不是A、S、T或V，并且其中VP衣壳能够形成组装的病毒体，与包含SEQ ID NO:1的AAV5 VP1衣壳的病毒体相比，所述组装的病毒体表现出降低的对肝脏组织的嗜性。

[0219] 在某些实施方案中，Xaa2不是A、C、H、I、K、S、T或V，并且其中VP衣壳能够形成组装的病毒体，与包含SEQ ID NO:1的AAV5 VP1衣壳的病毒体相比，所述组装的病毒体表现出降低的对肝脏组织的嗜性。

[0220] 在某些实施方案中，Xaa3不是A或T，并且其中VP衣壳能够形成组装的病毒体，与包含SEQ ID NO:1的AAV5 VP1衣壳的病毒体相比，所述组装的病毒体表现出降低的对肝脏组织的嗜性。

[0221] 在某些实施方案中，Xaa3不是A、M或T，并且其中VP衣壳能够形成组装的病毒体，与包含SEQ ID NO:1的AAV5 VP1衣壳的病毒体相比，所述组装的病毒体表现出降低的对肝脏组织的嗜性。

[0222] 在某些实施方案中，Xaa3不是A、G、H、K、M、N、Q、R、S、T或V，并且其中VP衣壳能够形成组装的病毒体，与包含SEQ ID NO:1的AAV5 VP1衣壳的病毒体相比，所述组装的病毒体表现出降低的对肝脏组织的嗜性。

[0223] 在某些实施方案中，Xaa4不是P，并且其中VP衣壳能够形成组装的病毒体，与包含SEQ ID NO:1的AAV5 VP1衣壳的病毒体相比，所述组装的病毒体表现出降低的对肝脏组织的嗜性。

[0224] 在某些实施方案中，其中Xaa4不是L、P、Q或T，并且其中VP衣壳能够形成组装的病毒体，与包含SEQ ID NO:1的AAV5 VP1衣壳的病毒体相比，所述组装的病毒体表现出降低的对肝脏组织的嗜性。

[0225] 在某些实施方案中，Xaa4不是L、M、P、Q、R、T或W，并且其中VP衣壳能够形成组装的病毒体，与包含SEQ ID NO:1的AAV5VP1衣壳的病毒体相比，所述组装的病毒体表现出降低的对肝脏组织的嗜性。

[0226] 在某些实施方案中，Xaa5不是Y，并且其中VP衣壳能够形成组装的病毒体，与包含SEQ ID NO:1的AAV5 VP1衣壳的病毒体相比，所述组装的病毒体表现出降低的对肝脏组织的嗜性。

[0227] 在某些实施方案中，Xaa5不是H、I或Y，并且其中VP衣壳能够形成组装的病毒体，与包含SEQ ID NO:1的AAV5 VP1衣壳的病毒体相比，所述组装的病毒体表现出降低的对肝脏组织的嗜性。

[0228] 在某些实施方案中，Xaa5不是F、H、I、K、M、T或Y，并且其中VP衣壳能够形成组装的病毒体，与包含SEQ ID NO:1的AAV5VP1衣壳的病毒体相比，所述组装的病毒体表现出降低

的对肝脏组织的嗜性。

[0229] 在某些实施方案中,Xaa6不是N,并且其中VP衣壳能够形成组装的病毒体,与包含SEQ ID NO:1的AAV5 VP1衣壳的病毒体相比,所述组装的病毒体表现出降低的对肝脏组织的嗜性。

[0230] 在某些实施方案中,Xaa6不是N或Q,并且其中VP衣壳能够形成组装的病毒体,与包含SEQ ID NO:1的AAV5 VP1衣壳的病毒体相比,所述组装的病毒体表现出降低的对肝脏组织的嗜性。

[0231] 在某些实施方案中,Xaa6不是E、G、H、L、M,N、Q、T或W,并且其中VP衣壳能够形成组装的病毒体,与包含SEQ ID NO:1的AAV5 VP1衣壳的病毒体相比,所述组装的病毒体表现出降低的对肝脏组织的嗜性。

[0232] 在某些实施方案中,Xaa7不是A,并且其中VP衣壳能够形成组装的病毒体,与包含SEQ ID NO:1的AAV5 VP1衣壳的病毒体相比,所述组装的病毒体表现出降低的对肝脏组织的嗜性。

[0233] 在某些实施方案中,Xaa7不是A、C、H或M,并且其中VP衣壳能够形成组装的病毒体,与包含SEQ ID NO:1的AAV5 VP1衣壳的病毒体相比,所述组装的病毒体表现出降低的对肝脏组织的嗜性。

[0234] 在某些实施方案中,Xaa7不是A、C、G、H、L、M、R或S,并且其中VP衣壳能够形成组装的病毒体,与包含SEQ ID NO:1的AAV5 VP1衣壳的病毒体相比,所述组装的病毒体表现出降低的对肝脏组织的嗜性。

[0235] 在某些实施方案中,Xaa8不是G,并且其中VP衣壳能够形成组装的病毒体,与包含SEQ ID NO:1的AAV5 VP1衣壳的病毒体相比,所述组装的病毒体表现出降低的对肝脏组织的嗜性。

[0236] 在某些实施方案中,Xaa8不是G、M、Q或S,并且其中VP衣壳能够形成组装的病毒体,与包含SEQ ID NO:1的AAV5 VP1衣壳的病毒体相比,所述组装的病毒体表现出降低的对肝脏组织的嗜性。

[0237] 在某些实施方案中,Xaa8不是A、C、D、F、G、H、M、Q、S、V、W或Y,并且其中VP衣壳能够形成组装的病毒体,与包含SEQ ID NO:1的AAV5 VP1衣壳的病毒体相比,所述组装的病毒体表现出降低的对肝脏组织的嗜性。

[0238] 在某些实施方案中,Xaa9不是G,并且其中VP衣壳能够形成组装的病毒体,与包含SEQ ID NO:1的AAV5 VP1衣壳的病毒体相比,所述组装的病毒体表现出降低的对肝脏组织的嗜性。

[0239] 在某些实施方案中,Xaa9不是E、G或P,并且其中VP衣壳能够形成组装的病毒体,与包含SEQ ID NO:1的AAV5 VP1衣壳的病毒体相比,所述组装的病毒体表现出降低的对肝脏组织的嗜性。

[0240] 在某些实施方案中,Xaa9不是A、C、E、G、H、M、N、P、Q、S、V或W,并且其中VP衣壳能够形成组装的病毒体,与包含SEQ ID NO:1的AAV5 VP1衣壳的病毒体相比,所述组装的病毒体表现出降低的对肝脏组织的嗜性。

[0241] 本公开涵盖变体VP衣壳,与SEQ ID NO:1的AAV5 VP1衣壳相比,所述变体VP衣壳具有增加的对以下组织中任一者的组织嗜性:脂肪、肾上腺、主动脉、脑(包括海马:齿状回、

CA1和CA3;小脑、尾状核、壳核、中脑、脑桥、下丘脑、皮层-包括枕叶、颞叶和前脑;黑质和丘脑)、骨髓、盲肠、结肠、背根神经节、十二指肠、附睾、食管、眼睛、胆囊、心脏、回肠、空肠、肾脏、肺、淋巴结、乳腺、卵巢、胰腺、甲状旁腺、周围神经、垂体、前列腺、唾液腺、精囊、骨骼肌、皮肤、脊髓、脾脏、胃、睾丸、胸腺、甲状腺、气管、膀胱、子宫和阴道的组织。

[0242] 6.7.4.体内选定的对肝脏组织去靶向的突变VP多肽

[0243] A.位置频率规则

[0244] 在本章节中,除非另外说明,否则分析相对于给定氨基酸残基出现在所有其他所收获组织(CNS(前脑皮层、枕叶皮层、颞叶皮层、丘脑、下丘脑、黑质、海马DG、海马CA1、海马CA3、小脑)、骨骼肌、心脏、肺、脾脏、淋巴结、骨髓、乳腺、皮肤、肾上腺、甲状腺、结肠、坐骨神经和脊髓组织、肝脏)中鉴定的变体中指定位置的频率而言该给定氨基酸残基出现在非肝脏中鉴定的变体中对应于SEQ ID NO:1(推及于SEQ ID NO:2)的位置581至位置589的所述指定位置的频率,以确定优先对肝脏组织去靶向的衣壳的一组序列规则。来自体内数据的位置频率规则的确定详细描述于实施例4中。参考实施例4中的表6B,和SEQ ID NO:2(AAV5 VP1),以下氨基酸可以由此以任何组合在Xaa1-Xaa9任何一个或多个位置处独立地突变,以提供与野生型AAV5 VP1衣壳(SEQ ID NO:1)相比具有降低的肝脏嗜性的VP1衣壳,其中肝脏嗜性在此是指对肝脏转导具有确定性的特性,而不是对所有其他所收获组织转导具有确定性的特性。

[0245] 本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽,所述多肽与野生型AAV5 VP衣壳多肽相比表现出降低的肝脏组织嗜性,其中工程化变体AAV5 VP衣壳多肽序列具有一个或多个突变,其中VP1多肽序列在从对应于SEQ ID NO:2中的581的位置至对应于SEQ ID NO:2中的589的位置的区域中具有所述一个或多个突变,并且其中Xaa1不包括K;或Xaa1不包括A、K、M或T;或Xaa1不包括A、G、K、M、N、Q、R、S或T;或Xaa2不包括T;或Xaa2不包括A、S、T或V;或Xaa2不包括A、C、H、I、K、S、T或V;或Xaa3不包括A或T;或Xaa3不包括A、M或T;或Xaa3不包括A、G、H、K、M、N、Q、R、S、T或V;或Xaa4不包括P;或Xaa4不包括L、P、Q或T;或Xaa4不包括L、M、P、Q、R、T或W;或Xaa5不包括Y;或Xaa5不包括H、I或Y;或Xaa5不包括F、H、I、K、M、T或Y;或Xaa6不包括N;或Xaa6不包括N或Q;或Xaa6不包括E、G、H、L、M、N、Q、T或W;或Xaa7不包括A;或Xaa7不包括A、C、H或M;或Xaa7不包括A、C、G、H、L、M、R或S;或Xaa8不包括G;或Xaa8不包括G、M、Q或S;或Xaa8不包括A、C、D、F、G、H、M、Q、S、V、W或Y;或Xaa9不包括G;或Xaa9不包括E、G或P;或Xaa9不包括A、C、E、G、H、M、N、P、Q、S、V或W。

[0246] 本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽,与野生型AAV5 VP衣壳多肽相比,所述组装的病毒体表现出降低的肝脏组织嗜性,其中工程化变体AAV5 VP衣壳多肽序列具有一个或多个突变,其中VP1多肽序列在从对应于SEQ ID NO:2中的581的位置至对应于SEQ ID NO:2中的589的位置的区域中具有所述一个或多个突变,并且其中所述一个或多个突变选自以下规则。在一些实施方案中,Xaa1不包括K。在一些实施方案中,Xaa1不包括A、K、M或T。在一些实施方案中,Xaa1不包括A、G、K、M、N、Q、R、S或T。在一些实施方案中,Xaa2不包括T。在一些实施方案中,Xaa2不包括A、S、T或V。在一些实施方案中,Xaa2不包括A、C、H、I、K、S、T或V。在一些实施方案中,Xaa3不包括A或T。在一些实施方案中,Xaa3不包括A、M或T。在一些实施方案中,Xaa3不包括A、G、H、K、M、N、Q、R、S、T或V。在一些实施方案中,Xaa4不包括P。在一些实施方案中,Xaa4不包括L、P、Q或T。在一些实施方案中,Xaa4不包括L、

M、P、Q、R、T或W。在一些实施方案中，Xaa5不包括Y。在一些实施方案中，Xaa5不包括H、I或Y。在一些实施方案中，Xaa5不包括F、H、I、K、M、T或Y。在一些实施方案中，Xaa6不包括N。在一些实施方案中，Xaa6不包括N或Q。在一些实施方案中，Xaa6不包括E、G、H、L、M、N、Q、T或W。在一些实施方案中，Xaa7不包括A。在一些实施方案中，Xaa7不包括A、C、H或M。在一些实施方案中，Xaa7不包括A、C、G、H、L、M、R或S。在一些实施方案中，Xaa8不包括G。在一些实施方案中，Xaa8不包括G、M、Q或S。在一些实施方案中，Xaa8不包括A、C、D、F、G、H、M、Q、S、V、W或Y。在一些实施方案中，Xaa9不包括G。在一些实施方案中，Xaa9不包括E、G或P。在一些实施方案中，Xaa9不包括A、C、E、G、H、M、N、P、Q、S、V或W。

[0247] 本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽，与野生型AAV5 VP衣壳多肽相比，所述组装的病毒体表现出降低的肝脏组织嗜性，其中工程化变体AAV5 VP衣壳多肽序列具有一个或多个突变，其中VP1多肽序列在从对应于SEQ ID NO:2中的581的位置至对应于SEQ ID NO:2中的589的位置的区域中具有所述一个或多个突变，并且其中所述一个或多个突变选自以下规则:Xaa1不包括A、K、M或T，Xaa2不包括，Xaa3不包括A或T，Xaa4不包括P，Xaa5不包括Y，Xaa6不包括N，Xaa7不包括A，Xaa8不包括G，并且Xaa9不包括G。

[0248] 本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽，与野生型AAV5 VP衣壳多肽相比，所述组装的病毒体表现出降低的肝脏组织嗜性，其中工程化变体AAV5 VP衣壳多肽序列具有一个或多个突变，其中VP1多肽序列在从对应于SEQ ID NO:2中的581的位置至对应于SEQ ID NO:2中的589的位置的区域中具有所述一个或多个突变，并且其中所述一个或多个突变如下:Xaa1不包括A、K、M或T，或Xaa2不包括，或Xaa3不包括A或T，或Xaa4不包括P，或Xaa5不包括Y，或Xaa6不包括N，或Xaa7不包括A，或Xaa8不包括G，或Xaa9不包括G，或它们的任何组合。

[0249] B.ML规则

[0250] 对于本段中描述的以下一组规则，使用体内数据和两个ML模型(描述于实施例21)确定工程化AAV5 VP1衣壳多肽的581至589区域中每个位置处的有利生物物理特性和有利氨基酸残基。本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽，与具有野生型AAV5 VP衣壳多肽的rAAV病毒体相比，所述组装的病毒体表现出降低的肝脏组织嗜性，其中工程化AAV5 VP衣壳多肽序列具有一个或多个突变，其中VP1多肽序列在从对应于SEQ ID NO:2中的581的位置至对应于SEQ ID NO:2中的589的位置的区域中具有所述一个或多个突变，并且其中Xaa1选自位置Xaa1处低溶解性的氨基酸(例如，Xaa1选自D或P)；或者其中Xaa1选自位置Xaa1处低可突变性的氨基酸(例如，Xaa1选自C、K或L)；或者其中Xaa2选自位置Xaa2处低溶解性的氨基酸(例如，Xaa2选自N、K、P、E或D)；或者其中Xaa2选自位置Xaa2处低亲疏水性的氨基酸(例如，Xaa2选自D、E、R、K、H、N或Q)；或者其中Xaa2选自位置Xaa2处低电荷的氨基酸(例如，Xaa2选自D或E)；或者其中Xaa2选自位置Xaa2处高总潜在氢键数的氨基酸(例如，Xaa2选自H、N、Q、D、E或R)；或者其中Xaa2选自位置Xaa2处中等体积的氨基酸(例如，Xaa2选自D、E、V、P、N或T)；或者其中Xaa3选自位置Xaa3处低溶解性的氨基酸(例如，Xaa3选自P或D)；或者其中Xaa4选自位置Xaa4处中等体积的氨基酸(例如，Xaa4选自D、E、V、P、N或T)；或者其中Xaa5选自位置Xaa5处低溶解性的氨基酸(例如，Xaa5选自N、P、E或D)；或者其中Xaa8选自位置Xaa8处低溶解性的氨基酸(例如，Xaa8选自K或Q)；或者其中Xaa8选自位置Xaa8处低亲疏水性的氨基酸(例如，Xaa8选自K或R)；或者其中Xaa8选自位置Xaa8处高表面

可接近性的氨基酸(例如,Xaa8选自E、R或K);或它们的任何组合。

[0251] 在一些实施方案中,Xaa1选自位置Xaa1处低溶解性的氨基酸。在一些实施方案中,Xaa1选自D或P。在一些实施方案中,Xaa1选自位置Xaa1处低可突变性的氨基酸。在一些实施方案中,Xaa1选自C、K或L。在一些实施方案中,Xaa2选自位置Xaa2处低溶解性的氨基酸。在一些实施方案中,Xaa2选自N、K、P、E或D。在一些实施方案中,Xaa2选自位置Xaa2处低亲疏水性的氨基酸。在一些实施方案中,Xaa2选自D、E、R、K、H、N或Q。在一些实施方案中,Xaa2选自位置Xaa2处低电荷的氨基酸。在一些实施方案中,Xaa2选自D、E。在一些实施方案中,Xaa2选自位置Xaa2处高总潜在氢键数的氨基酸。在一些实施方案中,Xaa2选自H、N、Q、D、E或R。在一些实施方案中,Xaa2选自位置Xaa2处中等体积的氨基酸。在一些实施方案中,Xaa2选自D、E、V、P、N或T。在一些实施方案中,Xaa3选自位置Xaa3处低溶解性的氨基酸。在一些实施方案中,Xaa3选自P或D。在一些实施方案中,Xaa4选自位置Xaa4处中等体积的氨基酸。在一些实施方案中,Xaa4选自D、E、V、P、N或T。在一些实施方案中,Xaa5选自位置Xaa5处低溶解性的氨基酸。在一些实施方案中,Xaa5选自N、P、E或D。在一些实施方案中,Xaa8选自位置Xaa8处低溶解性的氨基酸。在一些实施方案中,Xaa8选自K或Q。在一些实施方案中,Xaa8选自位置Xaa8处低亲疏水性的氨基酸。在一些实施方案中,Xaa8选自K或R。在一些实施方案中,Xaa8选自位置Xaa8处高表面可接近性的氨基酸。在一些实施方案中,Xaa8选自E、R或K。

[0252] 在一些实施方案中,本文提供了AAV5 VP衣壳多肽,所述多肽在从对应于AAV5 VP1的581的位置至对应于589的位置的区域中具有至少一个突变并且与选自SEQ ID NO:46438-SEQ ID NO:47437的任何序列具有至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%、至少99%或100%序列同一性,其中所述至少一个突变驱动肝脏去靶向组织嗜性。

[0253] 6.7.5.体内选定的赋予增加的肝脏嗜性的突变VP多肽

[0254] 本公开提供了具有VP衣壳多肽的AAV5病毒体,所述VP衣壳多肽在具有与靶细胞(例如,目标靶肝脏组织中的靶肝脏细胞)相互作用的残基的区域中具有至少一个突变,其中与野生型VP衣壳多肽相比,所述至少一个突变赋予增加的肝脏组织嗜性。在一些实施方案中,本文提供了与SEQ ID NO:1具有至少80%序列同源性的AAV5 VP1衣壳多肽,其中所述AAV5 VP1衣壳多肽在从对应于SEQ ID NO:1的581的位置至对应于589的位置的区域中具有至少一个突变,并且其中所述至少一个突变驱动增加的肝脏嗜性。以下序列规则和序列也适用于对应于AAV5 VP1氨基酸残基581至589的AAV5 VP2(氨基酸残基445至453;VP2序列显示于SEQ ID NO:1115)和AAV5 VP3(氨基酸残基389至397;VP3序列显示于SEQ ID NO:1116)中的区域。因此,本公开涵盖在对应于581至589区域的AAV5 VP1氨基酸残基的VP2和VP3区域中具有一个或多个突变的AAV5 VP2衣壳多肽和AAV5 VP3衣壳多肽,其中所述一个或多个突变与以下章节中的规则或顺序相称。

[0255] A.位置频率规则

[0256] 在本章节中,除非另外说明,否则分析相对于给定氨基酸残基出现在所有其他所收获组织(CNS(前脑皮层、枕叶皮层、颞叶皮层、丘脑、下丘脑、黑质、海马DG、海马CA1、海马CA3、小脑)、骨骼肌、心脏、肺、脾脏、淋巴结、骨髓、乳腺、皮肤、肾上腺、甲状腺、结肠、坐骨神经和脊髓组织)中鉴定的变体中指定位置的频率而言该给定氨基酸残基出现在肝脏中鉴定的变体中对应于SEQ ID NO:1(推及于SEQ ID NO:2)的位置581至位置589的所述指定位置

的频率,以确定优先靶向肝脏组织的衣壳的一组序列规则。来自体内数据的位置频率规则的确定详细描述于实施例4中。

[0257] 本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽,与野生型AAV5 VP衣壳多肽相比,所述组装的病毒体表现出增加的肝脏组织嗜性,其中工程化变体AAV5 VP衣壳多肽序列具有一个或多个突变,其中VP1多肽序列在从对应于SEQ ID NO:2中的581的位置至对应于SEQ ID NO:2中的589的位置的区域中具有所述一个或多个突变,并且其中Xaa1选自A、G、K、M、N、Q、R、S或T,或者

[0258] Xaa1选自A、K、Q或R,或Xaa1为K;或者其中Xaa2选自A、C、I、K、S、T或V,或Xaa2选自A、K、S或T,或Xaa2为A;或者其中Xaa3选自A、G、I、K、M、Q、R、S、T或V,或Xaa3选自A、K、Q、S或T,或Xaa3选自K、Q或T,或Xaa3为K;或者其中Xaa4选自A、I、K、L、P、Q、R、S、T或V,或Xaa4选自K、I、S或V,或Xaa4为K;或Xaa5选自F、I、L、M、T、V或Y,或者其中Xaa5选自F、L或Y,或Xaa5为F;或Xaa6选自F、H、M、N、Q、S或Y,或者其中Xaa6选自M或N,或Xaa6为N;或Xaa7选自A、C、K、M、Q或S,或者其中Xaa7选自A、C或S,或Xaa7为S;或者其中Xaa8选自A、C、F、G、M、Q或S,或Xaa8选自A、C、M或S,或Xaa8为C;或者其中Xaa9选自E、F、L、Q、R或Y,或Xaa9选自L、Q或R,或Xaa9为R。

[0259] 本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽,与野生型AAV5 VP衣壳多肽相比,所述组装的病毒体表现出增加的肝脏组织嗜性,其中工程化变体AAV5 VP衣壳多肽序列具有一个或多个突变,其中VP1多肽序列在从对应于SEQ ID NO:2中的581的位置至对应于SEQ ID NO:2中的589的位置的区域中具有所述一个或多个突变,并且其中所述一个或多个突变选自以下规则。在一些实施方案中,Xaa1选自A、G、K、M、N、Q、R、S或T。在一些实施方案中,Xaa1选自A、K、Q或R。在一些实施方案中,Xaa1为K。在一些实施方案中,Xaa2选自A、C、I、K、S、T或V。在一些实施方案中,选自A、K、S或T,或Xaa2为A。在一些实施方案中,其中Xaa3选自A、G、I、K、M、Q、R、S、T或V,或Xaa3选自A、K、Q、S或T。在一些实施方案中,Xaa3选自K、Q或T。在一些实施方案中,Xaa3为K。在一些实施方案中,Xaa4选自A、I、K、L、P、Q、R、S、T或V。在一些实施方案中,Xaa4选自K、I、S或V。在一些实施方案中,Xaa4为K。在一些实施方案中,Xaa5选自F、I、L、M、T、V或Y。在一些实施方案中,Xaa5选自F、L或Y,或Xaa5为F。在一些实施方案中,Xaa6选自F、H、M、N、Q、S或Y。在一些实施方案中,其中Xaa6选自M或N,或Xaa6为N。在一些实施方案中,Xaa7选自A、C、K、M、Q或S。在一些实施方案中,Xaa7选自A、C或S,或Xaa7为S。在一些实施方案中,Xaa8选自A、C、F、G、M、Q或S。在一些实施方案中,Xaa8选自A、C、M或S,或Xaa8为C。在一些实施方案中,Xaa9选自E、F、L、Q、R或Y。在一些实施方案中,Xaa9选自L、Q或R,或Xaa9为R。

[0260] 本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽,与野生型AAV5 VP衣壳多肽相比,所述组装的病毒体表现出增加的肝脏组织嗜性,其中工程化变体AAV5 VP衣壳多肽序列具有一个或多个突变,其中VP1多肽序列在从对应于SEQ ID NO:2中的581的位置至对应于SEQ ID NO:2中的589的位置的区域中具有所述一个或多个突变,并且其中所述一个或多个突变选自以下规则:Xaa1选自A、G、K、M、N、Q、R、S或T,Xaa2选自A、C、I、K、S、T或V,Xaa3选自A、G、I、K、M、Q、R、S、T或V,Xaa4选自A、I、K、L、P、Q、R、S、T或V,Xaa5选自F、I、L、M、T、V或Y,Xaa6选自F、H、M、N、Q、S或Y,Xaa7选自A、C或S,Xaa8选自A、C、F、G、M、Q或S,并且Xaa9选自E、F、L、Q、R或Y。

[0261] 本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽,与野生型AAV5 VP

衣壳多肽相比，所述组装的病毒体表现出增加的肝脏组织嗜性，其中工程化变体AAV5 VP衣壳多肽序列具有一个或多个突变，其中VP1多肽序列在从对应于SEQ ID NO:2中的581的位置至对应于SEQ ID NO:2中的589的位置的区域中具有所述一个或多个突变，并且其中所述一个或多个突变如下:Xaa1选自A、G、K、M、N、Q、R、S或T，或Xaa2选自A、C、I、K、S、T或V，或Xaa3选自A、G、I、K、M、Q、R、S、T或V，或Xaa4选自A、I、K、L、P、Q、R、S、T或V，或Xaa5选自F、I、L、M、T、V或Y，或Xaa6选自F、H、M、N、Q、S或Y，或Xaa7选自A、C或S，或Xaa8选自A、C、F、G、M、Q或S，或Xaa9选自E、F、L、Q、R或Y，或它们的任何组合。

[0262] 在一些实施方案中，本文提供了AAV5 VP衣壳多肽，所述多肽在从对应于AAV5 VP1的581的位置至对应于589的位置的区域中具有至少一个突变并且与选自SEQ ID NO:43438-SEQ ID NO:44437的任何序列具有至少70％、至少75％、至少80％、至少85％、至少90％、至少92％、至少95％、至少97％、至少99％或100％序列同一性，其中所述至少一个突变驱动增加的肝脏组织嗜性。

[0263] B.ML规则

[0264] 对于本段中描述的以下一组规则，使用体内数据和两个ML模型(描述于实施例20)确定工程化AAV5 VP1衣壳多肽的581至589区域中每个位置处的有利生物物理特性和有利氨基酸残基。本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽，与具有野生型AAV5 VP衣壳多肽的rAAV病毒体相比，所述组装的病毒体表现出增加的肝脏组织嗜性，其中工程化AAV5 VP衣壳多肽序列具有一个或多个突变，其中VP1多肽序列在从对应于SEQ ID NO:2中的581的位置至对应于SEQ ID NO:2中的589的位置的区域中具有所述一个或多个突变，并且其中Xaa1选自高表面可接近性的氨基酸(例如，Xaa1选自K、R或E)；或者其中Xaa1选自低亲疏水性的氨基酸(例如，Xaa1选自K、R)；或者其中Xaa1选自低氨基酸可突变性的氨基酸(例如，Xaa1选自H、P、K或R)；或者其中Xaa1选自低氨基酸溶解性的氨基酸(例如，Xaa1选自Q、K、R)；或者其中Xaa2选自高表面可接近性的氨基酸(例如，Xaa2选自E、R或K)；或者其中Xaa2选自低亲疏水性的氨基酸(例如，Xaa2选自K、R)；或者其中Xaa2选自高氨基酸体积的氨基酸(例如，Xaa2选自S、L、I、A、R或K)；或者其中Xaa3选自高可突变性的氨基酸(例如，Xaa3选自N、I、A、M、E或D)；或者其中Xaa3选自低溶解性的氨基酸(例如，Xaa3选自N、K、R或E)；或者其中Xaa4选自低亲疏水性的氨基酸(例如，Xaa4选自K或R)；或者其中Xaa4选自高氨基酸体积的氨基酸(例如，Xaa4选自K、R、I或L)；或者其中Xaa5选自中等氨基酸溶解性的氨基酸(例如，Xaa5选自H或T)；或者其中Xaa8选自低表面可接近性的氨基酸(例如，Xaa8选自V或C)；或者其中Xaa8选自低平均柔性指数的氨基酸(例如，Xaa8选自W、V、M、A、F、L、H或C)；或它们的任何组合。

[0265] 在一些实施方案中，Xaa1选自高表面可接近性的氨基酸。在一些实施方案中，Xaa1选自K、R或E。在一些实施方案中，Xaa1选自低亲疏水性的氨基酸。在一些实施方案中，Xaa1选自K或R。在一些实施方案中，Xaa1选自低氨基酸可突变性的氨基酸。在一些实施方案中，Xaa1选自H、P、K或R。在一些实施方案中，Xaa1选自低氨基酸溶解性的氨基酸。在一些实施方案中，Xaa1选自Q、K或R。在一些实施方案中，Xaa2选自高表面可接近性的氨基酸。在一些实施方案中，Xaa2选自E、R或K。在一些实施方案中，Xaa2选自低亲疏水性的氨基酸。在一些实施方案中，Xaa2选自K或R。在一些实施方案中，Xaa2选自高氨基酸体积的氨基酸。在一些实施方案中，Xaa2选自S、L、I、A、R或K。在一些实施方案中，Xaa3选自高可突变性的氨基酸。在

一些实施方案中，Xaa3选自N、I、A、M、E或D。在一些实施方案中，Xaa3选自低溶解性的氨基酸。在一些实施方案中，Xaa3选自N、K、R或E。在一些实施方案中，Xaa4选自低亲疏水性的氨基酸。在一些实施方案中，Xaa4选自K、R。在一些实施方案中，Xaa4选自高氨基酸体积的氨基酸。在一些实施方案中，Xaa4选自K、R、I或L。在一些实施方案中，Xaa5选自中等氨基酸溶解性的氨基酸。在一些实施方案中，Xaa5选自H、T。在一些实施方案中，Xaa8选自低表面可接近性的氨基酸。在一些实施方案中，Xaa8选自V或C。在一些实施方案中，Xaa8选自低平均柔性指数的氨基酸。在一些实施方案中，Xaa8选自W、V、M、A、F、L、H或C。

[0266] 在一些实施方案中，本文提供了AAV5 VP衣壳多肽，所述多肽在从对应于AAV5 VP1的581的位置至对应于589的位置的区域中具有至少一个突变并且与选自SEQ ID NO:44438-SEQ ID NO:45437的任何序列具有至少70％、至少75％、至少80％、至少85％、至少90％、至少92％、至少95％、至少97％、至少99％或100％序列同一性，其中所述至少一个突变驱动增加的肝脏组织嗜性。

[0267] C.富集的肝脏序列

[0268] 在一些实施方案中，本文提供了AAV5 VP衣壳多肽，所述多肽在从对应于AAV5 VP1的581的位置至对应于589的位置的区域中具有至少一个突变并且与选自SEQ ID NO:45438-SEQ ID NO:46437的任何序列具有至少70％、至少75％、至少80％、至少85％、至少90％、至少92％、至少95％、至少97％、至少99％或100％序列同一性，其中所述至少一个突变驱动增加的肝脏组织嗜性。

[0269] 6.7.6.体内选定的赋予增加的中枢神经系统嗜性的突变VP多肽，基于位置频率的规则和ML规则

[0270] 本公开提供了具有VP衣壳多肽的AAV5病毒体，所述VP衣壳多肽在具有与靶细胞(例如，目标靶CNS组织中的靶CNS细胞)相互作用的残基的区域中具有至少一个突变，其中与野生型VP衣壳多肽相比，所述至少一个突变赋予增加的CNS组织嗜性。在一些实施方案中，本文提供了与SEQ ID NO:1具有至少80％序列同源性的AAV5VP1衣壳多肽，其中所述AAV5 VP1衣壳多肽在从对应于SEQ ID NO:1的581的位置至对应于589的位置的区域中具有至少一个突变，并且其中所述至少一个突变驱动增加的中枢神经系统嗜性。以下序列规则和序列也适用于对应于AAV5 VP1氨基酸残基581至589的AAV5VP2(氨基酸残基445至453；VP2序列显示于SEQ ID NO:1115)和AAV5 VP3(氨基酸残基389至397;VP3序列显示于SEQ ID NO:1116)中的区域。因此，本公开涵盖在对应于581至589区域的AAV5 VP1氨基酸残基的VP2和VP3区域中具有一个或多个突变的AAV5 VP2衣壳多肽和AAV5 VP3衣壳多肽，其中所述一个或多个突变与以下章节中的规则或顺序相称。

[0271] A.位置频率规则

[0272] 在本章节中，除非另外说明，否则分析相对于给定氨基酸残基出现在所有其他所收获组织(肝脏、骨骼肌、心脏、肺、脾脏、淋巴结、骨髓、乳腺、皮肤、肾上腺、甲状腺、结肠、坐骨神经和脊髓组织)中鉴定的变体中指定位置的频率而言该给定氨基酸残基出现在中枢神经系统(前脑皮层、枕叶皮层、颞叶皮层、丘脑、下丘脑、黑质、海马DG、海马CA1、海马CA3、小脑)中鉴定的变体中对应于SEQ ID NO:1(推及于SEQ ID NO:2)的位置581至位置589的所述指定位置的频率，以确定优先靶向中枢神经系统组织的衣壳的一组序列规则。来自体内数据的位置频率规则的确定详细描述于实施例5中。

[0273] 本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽，与野生型AAV5 VP衣壳多肽相比，所述组装的病毒体表现出增加的中枢神经系统组织嗜性，其中工程化变体AAV5 VP衣壳多肽序列具有一个或多个突变，其中VP1多肽序列在从对应于SEQ ID NO:2中的581的位置至对应于SEQ ID NO:2中的589的位置的区域中具有所述一个或多个突变，并且其中Xaa1选自A、C、K、M、Q、R、T或W，或Xaa1选自K、Q、R或W，或Xaa1为K；或Xaa2选自F、I、K、R、T或W，或Xaa2选自F、I、R或T，或Xaa2为R；或Xaa3选自A、H、N、R或W，或Xaa3选自A、R或W，或Xaa3为R；或Xaa4选自E、G、I、M、Q或R，或Xaa4选自E、M或R，或Xaa4为R；或Xaa5选自C、G、K、I、M或R，或Xaa5选自K、I或R，或Xaa5为I；或Xaa6选自I、K、L、P、Q、R、Y，或Xaa6选自K、R或Y，或Xaa6为R；或Xaa7选自D、I、K、R、V或W，或Xaa7选自I、R或V，或Xaa7为V；或Xaa8选自C、G、H、K、L或V，或Xaa8选自H、K或V，或Xaa8为H；或Xaa9选自I、K、L、R或V，或Xaa9选自I、K或R，或Xaa9为R。

[0274] 本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽，与野生型AAV5 VP衣壳多肽相比，所述组装的病毒体表现出增加的中枢神经系统组织嗜性，其中工程化变体AAV5 VP衣壳多肽序列具有一个或多个突变，其中VP1多肽序列在从对应于SEQ ID NO:2中的581的位置至对应于SEQ ID NO:2中的589的位置的区域中具有所述一个或多个突变，并且其中所述一个或多个突变选自以下规则。在一些实施方案中，Xaa1选自A、C、K、M、Q、R、T或W。在一些实施方案中，Xaa1选自K、Q、R或W。在一些实施方案中，Xaa1为K。在一些实施方案中，Xaa2选自F、I、K、R、T或W。在一些实施方案中，Xaa2选自F、I、R或T。在一些实施方案中，Xaa2为R。在一些实施方案中，Xaa3选自A、H、N、R或W。在一些实施方案中，Xaa3选自A、R或W。在一些实施方案中，Xaa3为R。在一些实施方案中，Xaa4选自E、G、I、M、Q或R。在一些实施方案中，Xaa4选自E、M或R。在一些实施方案中，Xaa4为R。在一些实施方案中，Xaa5选自C、G、K、I、M或R。在一些实施方案中，Xaa5选自K、I或R。在一些实施方案中，Xaa5为I。在一些实施方案中，Xaa6选自I、K、L、P、Q、R、Y。在一些实施方案中，Xaa6选自K、R或Y。在一些实施方案中，Xaa6为R。在一些实施方案中，Xaa7选自D、I、K、R、V或W。在一些实施方案中，Xaa7选自I、R或V。在一些实施方案中，Xaa7为V。在一些实施方案中，Xaa8选自C、G、H、K、L或V。在一些实施方案中，Xaa8选自H、K或V。在一些实施方案中，Xaa8为H。在一些实施方案中，Xaa9选自I、K、L、R或V。在一些实施方案中，Xaa9选自I、K或R。在一些实施方案中，Xaa9为R。

[0275] 本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽，与野生型AAV5 VP衣壳多肽相比，所述组装的病毒体表现出增加的中枢神经系统组织嗜性，其中工程化变体AAV5 VP衣壳多肽序列具有一个或多个突变，其中VP1多肽序列在从对应于SEQ ID NO:2中的581的位置至对应于SEQ ID NO:2中的589的位置的区域中具有所述一个或多个突变，并且其中所述一个或多个突变选自以下规则：Xaa1选自A、C、K、M、Q、R、T或W，Xaa2选自F、I、K、R、T或W，Xaa3选自A、H、N、R或W，Xaa4选自E、G、I、M、Q或R，Xaa5选自C、G、K、I、M或R，Xaa6选自I、K、L、P、Q、R、Y，Xaa7选自D、I、K、R、V或W，Xaa8选自C、G、H、K、L或V，并且Xaa9选自I、K、L、R或V。

[0276] 本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽，与野生型AAV5 VP衣壳多肽相比，所述组装的病毒体表现出增加的中枢神经系统组织嗜性，其中工程化变体AAV5 VP衣壳多肽序列具有一个或多个突变，其中VP1多肽序列在从对应于SEQ ID NO:2中的581的位置至对应于SEQ ID NO:2中的589的位置的区域中具有所述一个或多个突变，并

且其中所述一个或多个突变如下:Xaa1选自A、C、K、M、Q、R、T或W,Xaa2选自F、I、K、R、T或W,Xaa3选自A、H、N、R或W,Xaa4选自E、G、I、M、Q或R,Xaa5选自C、G、K、I、M或R,Xaa6选自I、K、L、P、Q、R、Y,Xaa7选自D、I、K、R、V或W,Xaa8选自C、G、H、K、L或V,Xaa9选自I、K、L、R或V,或它们的任何组合。

[0277] 在一些实施方案中,本文提供了AAV5 VP衣壳多肽,所述多肽在从对应于AAV5 VP1的581的位置至对应于589的位置的区域中具有至少一个突变并且与选自SEQ ID NO:7118-SEQ ID NO:8117的任何序列具有至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%、至少99%或100%序列同一性,其中所述至少一个突变驱动增加的中枢神经系统组织嗜性。

[0278] B.ML规则

[0279] 对于本章节后续段落中描述的以下一组规则,使用体内数据和两个ML模型(描述于实施例19)确定工程化AAV5 VP1衣壳多肽的581至589区域中每个位置处的有利生物物理特性和有利氨基酸残基。本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽,与具有野生型AAV5 VP衣壳多肽的rAAV病毒体相比,所述组装的病毒体表现出增加的中枢神经系统组织嗜性,其中工程化变体AAV5 VP衣壳多肽序列具有一个或多个突变,其中VP1多肽序列在从对应于SEQ ID NO:2中的581的位置至对应于SEQ ID NO:2中的589的位置的区域中具有所述一个或多个突变,并且其中Xaa1选自低氨基酸溶解性的氨基酸(例如,Xaa1选自K、R或Q);或者其中Xaa1选自低氨基酸亲疏水性的氨基酸(例如,Xaa1选自K或R);或者其中Xaa1选自高平均氨基酸柔性指数的氨基酸(例如,Xaa1选自D、E、R、K、G、I、N、Q或S);或者其中Xaa1选自氢键供体的氨基酸(例如,Xaa1选自K或R);或者其中Xaa1选自氨基酸可突变性的氨基酸(例如,Xaa1选自K、R、P或H);或者其中Xaa2选自低氨基酸溶解性的氨基酸(例如,Xaa2选自R、K、Q或S);或者其中Xaa2选自低氨基酸亲疏水性的氨基酸(例如,Xaa2选自R、K、D、E、N、Q、H、P、Y、W、S或T);或者其中Xaa2选自高氨基酸电荷的氨基酸(例如,Xaa2选自R、K或H);或者其中Xaa3选自高氨基酸溶解性的氨基酸(例如,Xaa3选自A、M、V、W、L或I);或者其中Xaa5选自高氨基酸溶解性的氨基酸(例如,Xaa5选自C、M、V、W、L或I);或者其中Xaa5选自高亲疏水性的氨基酸(例如,Xaa5选自M、V或I);或者其中Xaa5选自低平均氨基酸柔性指数的氨基酸(例如,Xaa5选自M、W、F或C);或者其中Xaa8选自高氨基酸溶解性的氨基酸(例如,Xaa8选自H、V或I);或它们的任何组合。

[0280] 在一些实施方案中,Xaa1选自低氨基酸溶解性的氨基酸。在一些实施方案中,Xaa1选自K、R、Q。在一些实施方案中,Xaa1选自低氨基酸亲疏水性的氨基酸。在一些实施方案中,Xaa1选自K或R。在一些实施方案中,Xaa1选自高平均氨基酸柔性指数的氨基酸。在一些实施方案中,Xaa1选自D、E、R、K、G、I、N、Q或S。在一些实施方案中,Xaa1选自高氢键供体的氨基酸。在一些实施方案中,Xaa1选自K或R。在一些实施方案中,Xaa1选自氨基酸可突变性的氨基酸。在一些实施方案中,Xaa1选自K、R、P或H。在一些实施方案中,Xaa2选自低氨基酸溶解性的氨基酸。在一些实施方案中,Xaa2选自R、K、Q或S。在一些实施方案中,Xaa2选自低氨基酸亲疏水性的氨基酸。在一些实施方案中,Xaa2选自R、K、D、E、N、Q、H、P、Y、W、S或T。在一些实施方案中,Xaa2选自高氨基酸电荷的氨基酸。在一些实施方案中,Xaa2选自R、K、H。在一些实施方案中,Xaa3选自高氨基酸溶解性的氨基酸。在一些实施方案中,Xaa3选自A、M、V、W、L或I。在一些实施方案中,Xaa5选自高氨基酸溶解性的氨基酸。在一些实施方案中,Xaa5选自C、

M、V、W、L或I。在一些实施方案中，Xaa5选自高亲疏水性的氨基酸。在一些实施方案中，Xaa5选自M、V或I。在一些实施方案中，Xaa5选自低平均氨基酸柔性指数的氨基酸。在一些实施方案中，Xaa5选自M、W、F或C。在一些实施方案中，Xaa8选自高氨基酸溶解性的氨基酸。在一些实施方案中，Xaa8选自H、V或I。

[0281] 在一些实施方案中，本文提供了AAV5 VP衣壳多肽，所述多肽在从对应于AAV5 VP1的581的位置至对应于589的位置的区域中具有至少一个突变并且与选自SEQ ID NO:8118-SEQ ID NO:9117的任何序列具有至少70％、至少75％、至少80％、至少85％、至少90％、至少92％、至少95％、至少97％、至少99％或100％序列同一性，其中所述至少一个突变驱动增加的CNS组织嗜性。

[0282] C.富集的CNS序列

[0283] 在一些实施方案中，本文提供了AAV5 VP衣壳多肽，所述多肽在从对应于AAV5 VP1的581的位置至对应于589的位置的区域中具有至少一个突变并且与选自SEQ ID NO:9118-SEQ ID NO:10117的任何序列具有至少70％、至少75％、至少80％、至少85％、至少90％、至少92％、至少95％、至少97％、至少99％或100％序列同一性，其中所述至少一个突变驱动增加的CNS组织嗜性。

[0284] 6.7.7.体内选定的赋予增加的脾脏嗜性的突变VP多肽

[0285] 本公开提供了具有VP衣壳多肽的AAV5病毒体，所述VP衣壳多肽在具有与靶细胞(例如，目标靶脾脏组织中的靶脾脏细胞)相互作用的残基的区域中具有至少一个突变，其中与野生型VP衣壳多肽相比，所述至少一个突变赋予增加的脾脏组织嗜性。在一些实施方案中，本文提供了与SEQ ID NO:1具有至少80％序列同源性的AAV5 VP1衣壳多肽，其中所述AAV5 VP1衣壳多肽在从对应于SEQ ID NO:1的581的位置至对应于589的位置的区域中具有至少一个突变，并且其中所述至少一个突变驱动增加的脾脏嗜性。以下序列规则和序列也适用于对应于AAV5 VP1氨基酸残基581至589的AAV5 VP2(氨基酸残基445至453;VP2序列显示于SEQ ID NO:1115)和AAV5 VP3(氨基酸残基389至397;VP3序列显示于SEQ ID NO:1116)中的区域。因此，本公开涵盖在对应于581至589区域的AAV5 VP1氨基酸残基的VP2和VP3区域中具有一个或多个突变的AAV5 VP2衣壳多肽和AAV5 VP3衣壳多肽，其中所述一个或多个突变与以下章节中的规则或顺序相称。

[0286] A.位置频率规则

[0287] 在本章节中，除非另外说明，否则分析相对于给定氨基酸残基出现在所有其他所收获组织(CNS(前脑皮层、枕叶皮层、颞叶皮层、丘脑、下丘脑、黑质、海马DG、海马CA1、海马CA3、小脑)、肝脏、骨骼肌、心脏、肺、淋巴结、骨髓、乳腺、皮肤、肾上腺、甲状腺、结肠、坐骨神经和脊髓组织)中鉴定的变体中指定位置的频率而言该给定氨基酸残基出现在脾脏中鉴定的变体中对应于SEQ ID NO:1(推及于SEQ ID NO:2)的位置581至位置589的所述指定位置的频率，以确定优先靶向脾脏组织的衣壳的一组序列规则。来自体内数据的位置频率规则的确定详细描述于实施例6中。

[0288] 本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽，与野生型AAV5 VP衣壳多肽相比，所述组装的病毒体表现出增加的脾脏组织嗜性，其中工程化变体AAV5 VP1衣壳多肽序列具有一个或多个突变，其中Xaa1选自C、F、H、I、L、P、W或Y，或Xaa1选自C、F、P、W或Y，或Xaa1选自P、W或Y，或Xaa1为P;或Xaa2选自D、E、L、N、P、R或W，或Xaa2选自D、E或W，或

Xaa2为D;或Xaa3选自C、D、E、P或W,或Xaa3选自D、P或W,或Xaa3为P;或Xaa4选自C、F、G、H、R、W或Y,或Xaa4选自C、H或W,或Xaa4为C;或Xaa5选自A、D、E、G、P、R或W,或Xaa5选自D、E、G或P,或Xaa5为D;或Xaa6选自A、C、D、E、K、R、W,或Xaa6选自C、K或R,或Xaa6为K;或Xaa7选自F、L、P、R、W、Y,或Xaa7选自L、P或W,或Xaa7为P;或Xaa8选自E、I、K、L、P、R或T,或Xaa8选自P、R或K,或Xaa8为K;或Xaa9选自C、H、M、T、V或W,或Xaa9选自C、T或V,或Xaa9为V。

[0289] 本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽,与野生型AAV5 VP衣壳多肽相比,所述组装的病毒体表现出增加的脾脏组织嗜性,其中工程化变体AAV5 VP衣壳多肽序列具有一个或多个突变,其中VP1多肽序列在从对应于SEQ ID NO:2中的581的位置至对应于SEQ ID NO:2中的589的位置的区域中具有所述一个或多个突变,并且其中所述一个或多个突变选自以下规则。在一些实施方案中,Xaa1选自C、F、H、I、L、P、W或Y。在一些实施方案中,Xaa1选自C、F、P、W或Y。在一些实施方案中,Xaa1选自P、W或Y。在一些实施方案中,Xaa1为P。在一些实施方案中,Xaa2选自D、E、L、N、P、R或W。在一些实施方案中,Xaa2选自D、E或W。在一些实施方案中,Xaa2为D。在一些实施方案中,Xaa3选自C、D、E、P或W。在一些实施方案中,Xaa3选自D、P或W。在一些实施方案中,Xaa3为P。在一些实施方案中,Xaa4选自C、F、G、H、R、W或Y。在一些实施方案中,Xaa4选自C、H或W。在一些实施方案中,Xaa4为C。在一些实施方案中,Xaa5选自A、D、E、G、P、R或W。在一些实施方案中,Xaa5选自D、E、G或P。在一些实施方案中,Xaa5为D。在一些实施方案中,Xaa6选自A、C、D、E、K、R、W。在一些实施方案中,Xaa6选自C、K或R。在一些实施方案中,Xaa6为K。在一些实施方案中,Xaa7选自F、L、P、R、W、Y。在一些实施方案中,Xaa7选自L、P或W。在一些实施方案中,Xaa7为P。在一些实施方案中,Xaa8选自E、I、K、L、P、R或T。在一些实施方案中,Xaa8选自P、R或K。在一些实施方案中,Xaa8为K。在一些实施方案中,Xaa9选自C、H、M、T、V或W。在一些实施方案中,Xaa9选自C、T或V。在一些实施方案中,Xaa9为V。

[0290] 本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽,与野生型AAV5 VP衣壳多肽相比,所述组装的病毒体表现出增加的脾脏组织嗜性,其中工程化变体AAV5 VP衣壳多肽序列具有一个或多个突变,其中VP1多肽序列在从对应于SEQ ID NO:2中的581的位置至对应于SEQ ID NO:2中的589的位置的区域中具有所述一个或多个突变,并且其中所述一个或多个突变选自以下规则:Xaa1选自C、F、H、I、L、P、W或Y,Xaa2选自D、E、L、N、P、R或W,Xaa3选自C、D、E、P或W,Xaa4选自C、F、G、H、R、W或Y,Xaa5选自A、D、E、G、P、R或W,Xaa6选自A、C、D、E、K、R、W,Xaa7选自F、L、P、R、W、Y,Xaa8选自E、I、K、L、P、R或T,并且Xaa9选自C、H、M、T、V或W。

[0291] 本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽,与野生型AAV5 VP衣壳多肽相比,所述组装的病毒体表现出增加的脾脏组织嗜性,其中工程化变体AAV5 VP衣壳多肽序列具有一个或多个突变,其中VP1多肽序列在从对应于SEQ ID NO:2中的581的位置至对应于SEQ ID NO:2中的589的位置的区域中具有所述一个或多个突变,并且其中所述一个或多个突变如下:Xaa1选自C、F、H、I、L、P、W或Y,Xaa2选自D、E、L、N、P、R或W,Xaa3选自C、D、E、P或W,Xaa4选自C、F、G、H、R、W或Y,Xaa5选自A、D、E、G、P、R或W,Xaa6选自A、C、D、E、K、R、W,Xaa7选自F、L、P、R、W、Y,Xaa8选自E、I、K、L、P、R或T,Xaa9选自C、H、M、T、V或W,或它们的任何组合。

[0292] 在一些实施方案中,本文提供了AAV5 VP衣壳多肽,所述多肽在从对应于AAV5 VP1

的581的位置至对应于589的位置的区域中具有至少一个突变并且与选自SEQ ID NO:37438-SEQ ID NO:38437的任何序列具有至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%、至少99%或100%序列同一性,其中所述至少一个突变驱动增加的脾脏组织嗜性。

[0293] B.ML规则

[0294] 对于本段中描述的以下一组规则,使用体内数据和两个ML模型(描述于实施例42)确定工程化AAV5 VP1衣壳多肽的581至589区域中每个位置处的有利生物物理特性和有利氨基酸残基。本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽,与具有野生型AAV5 VP衣壳多肽的rAAV病毒体相比,所述组装的病毒体表现出增加的脾脏组织嗜性,其中工程化AAV5 VP衣壳多肽序列具有一个或多个突变,其中工程化AAV5 VP1衣壳多肽序列在从对应于SEQ ID NO:2中的581的位置至对应于SEQ ID NO:2中的589的位置的区域中具有所述一个或多个突变,并且其中Xaa1选自低溶解性的氨基酸(例如,Xaa1选自D或P);或者其中Xaa1选自高溶解性的氨基酸(例如,Xaa1选自F、I、L);或者其中Xaa1选自低亲疏水性的氨基酸(例如,Xaa1选自Y或P);或者其中Xaa1选自低可突变性的氨基酸(例如,Xaa1选自C、K或P);或者其中Xaa2选自低溶解性的氨基酸(例如,Xaa2选自D、Q或R);或者其中Xaa2选自低亲疏水性的氨基酸(例如,Xaa2选自D、E、R、K、H、N或Q);或者其中Xaa2选自低电荷的氨基酸(例如,Xaa2选自D或E);或者其中Xaa2选自低体积的氨基酸(例如,Xaa2选自T、N、P或D);或者其中Xaa2选自高平均柔性的氨基酸(例如,Xaa2选自D、E、R、P、G、Q或S);或者其中Xaa3选自低溶解性的氨基酸(例如,Xaa3选自D、E、P或N);或者其中Xaa3选自低亲疏水性的氨基酸(例如,Xaa3选自D、E、H、N、Q或P);或者其中Xaa4选自低亲疏水性的氨基酸(例如,Xaa4选自K或R);或者其中Xaa5选自低溶解性的氨基酸(例如,Xaa5选自D、E、P或N);或者其中Xaa5选自高平均柔性的氨基酸(例如,Xaa5选自D、E、R、P、G、Q或S);或者其中Xaa6选自低可突变性的氨基酸(例如,Xaa6选自C);或者其中Xaa8选自高表面可接近性的氨基酸(例如,Xaa8选自E、R或K);或者其中Xaa8选自低溶解性的氨基酸(例如,Xaa8选自E、P、R、K、N或Q);或者其中Xaa8选自中等体积的氨基酸(例如,Xaa8选自E、D、R、K、V、P、M、I、L、H、N、Q或T);或者其中Xaa9选自中等摩尔质量的氨基酸(例如,Xaa9选自E、D、K、M、I、L、H或N);或它们的任何组合。

[0295] 在一些实施方案中,Xaa1选自低溶解性的氨基酸。在一些实施方案中,Xaa1选自D或P。在一些实施方案中,Xaa1选自高溶解性的氨基酸。在一些实施方案中,Xaa1选自F、I或L。在一些实施方案中,Xaa1选自低亲疏水性的氨基酸。在一些实施方案中,Xaa1选自Y或P。在一些实施方案中,Xaa1选自低可突变性的氨基酸。在一些实施方案中,Xaa1选自C、K或P。在一些实施方案中,Xaa2选自低溶解性的氨基酸。在一些实施方案中,Xaa2选自D、Q或R。在一些实施方案中,Xaa2选自低亲疏水性的氨基酸。在一些实施方案中,Xaa2选自D、E、R、K、H、N或Q。在一些实施方案中,Xaa2选自低电荷的氨基酸。在一些实施方案中,Xaa2选自D或E。在一些实施方案中,Xaa2选自低体积的氨基酸。在一些实施方案中,Xaa2选自T、N、P或D。在一些实施方案中,Xaa2选自高平均柔性的氨基酸。在一些实施方案中,Xaa2选自D、E、R、P、G、Q或S。在一些实施方案中,Xaa3选自低溶解性的氨基酸。在一些实施方案中,Xaa3选自D、E、P或N。在一些实施方案中,Xaa3选自低亲疏水性的氨基酸。在一些实施方案中,Xaa3选自D、E、H、N、Q或P。在一些实施方案中,Xaa4选自低亲疏水性的氨基酸。在一些实施方案中,Xaa4选自K或R。在一些实施方案中,Xaa5选自低溶解性的氨基酸。在一些实施方案中,Xaa5选自D、

E、P或N。在一些实施方案中，Xaa5选自高平均柔性的氨基酸。在一些实施方案中，Xaa5选自D、E、R、P、G、Q或S。在一些实施方案中，Xaa6选自低可突变性的氨基酸。在一些实施方案中，Xaa6选自C。在一些实施方案中，Xaa8选自高表面可接近性的氨基酸。在一些实施方案中，Xaa8选自E、R或K。在一些实施方案中，Xaa8选自低溶解性的氨基酸。在一些实施方案中，Xaa8选自E、P、R、K、N或Q。在一些实施方案中，Xaa8选自中等体积的氨基酸。在一些实施方案中，Xaa8选自E、D、R、K、V、P、M、I、L、H、N、Q或T。在一些实施方案中，Xaa9选自中等摩尔质量的氨基酸。在一些实施方案中，Xaa9选自E、D、K、M、I、L、H或N。

[0296] 在一些实施方案中，本文提供了AAV5 VP衣壳多肽，所述多肽在从对应于AAV5 VP1的581的位置至对应于589的位置的区域中具有至少一个突变并且与选自SEQ ID NO:38438-SEQ ID NO:39437的任何序列具有至少70％、至少75％、至少80％、至少85％、至少90％、至少92％、至少95％、至少97％、至少99％或100％序列同一性，其中所述至少一个突变驱动增加的脾脏组织嗜性。

[0297] C.富集的脾脏序列

[0298] 在一些实施方案中，本文提供了AAV5 VP衣壳多肽，所述多肽在从对应于AAV5 VP1的581的位置至对应于589的位置的区域中具有至少一个突变并且与选自SEQ ID NO:39438-SEQ ID NO:40437的任何序列具有至少70％、至少75％、至少80％、至少85％、至少90％、至少92％、至少95％、至少97％、至少99％或100％序列同一性，其中所述至少一个突变驱动增加的脾脏组织嗜性。

[0299] 6.7.8.体内选定的赋予增加的肾上腺嗜性的突变VP多肽

[0300] 本公开提供了具有VP衣壳多肽的AAV5病毒体，所述VP衣壳多肽在具有与靶细胞(例如，目标靶肾上腺组织中的靶肾上腺细胞)相互作用的残基的区域中具有至少一个突变，其中与野生型VP衣壳多肽相比，所述至少一个突变赋予增加的肾上腺组织嗜性。在一些实施方案中，本文提供了与SEQ ID NO:1具有至少80％序列同源性的AAV5 VP1衣壳多肽，其中所述AAV5 VP1衣壳多肽在从对应于SEQ ID NO:1的581的位置至对应于589的位置的区域中具有至少一个突变，并且其中所述至少一个突变驱动增加的肾上腺嗜性。以下序列规则和序列也适用于对应于AAV5 VP1氨基酸残基581至589的AAV5 VP2(氨基酸残基445至453；VP2序列显示于SEQ ID NO:111 5)和AAV5 VP3(氨基酸残基389至397；VP3序列显示于SEQ ID NO:1116)中的区域。因此，本公开涵盖在对应于581至589区域的AAV5 VP1氨基酸残基的VP2和VP3区域中具有一个或多个突变的AAV5 VP2衣壳多肽和AAV5 VP3衣壳多肽，其中所述一个或多个突变与以下章节中的规则或顺序相称。

[0301] A.位置频率规则

[0302] 在本章节中，除非另外说明，否则分析相对于给定氨基酸残基出现在所有其他所收获组织(CNS(前脑皮层、枕叶皮层、颞叶皮层、丘脑、下丘脑、黑质、海马DG、海马CA1、海马CA3、小脑)、肝脏、骨骼肌、心脏、肺、脾脏、淋巴结、骨髓、乳腺、皮肤、甲状腺、结肠、坐骨神经和脊髓组织)中鉴定的变体中指定位置的频率而言该给定氨基酸残基出现在肾上腺中鉴定的变体中对应于SEQ ID NO:1(推及于SEQ ID NO:2)的位置581至位置589的所述指定位置的频率，以确定优先靶向肾上腺组织的衣壳的一组序列规则。来自体内数据的位置频率规则的确定详细描述于实施例7中。

[0303] 本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽，与野生型AAV5 VP

衣壳多肽相比，所述组装的病毒体表现出增加的肾上腺组织嗜性，其中工程化变体AAV5 VP1衣壳多肽序列具有一个或多个突变，其中Xaa1选自A、C、K、Q、R、S或T，或Xaa1选自C、K或R，或Xaa1为C；或Xaa2选自A、C、I、S、T或V，或Xaa2选自A、V或T，或Xaa2为V；或Xaa3选自A、F、G、K、M、Q、R、T或V，或Xaa3选自A、G或M，或Xaa3为M；或Xaa4选自A、K、M、Q、R或V，或Xaa4选自A、R或K，或Xaa4为K；或Xaa5选自F、I、L、M、R、T、V或Y，或Xaa5选自R、V或Y，或Xaa5为V；或Xaa6选自G、H、M、N、R或S，或Xaa6选自H或N，或Xaa6为N；或Xaa7选自A、H、K、Q、R、S或V，或Xaa7选自H、Q或V，或Xaa7为H；或Xaa8选自A、G、H、M、Q或S，或Xaa8选自A、G、M或S，或Xaa8为S；或Xaa9选自A、E、N、P、R、S或Y，或Xaa9选自P或E，或Xaa9为P。

[0304] 本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽，与野生型AAV5 VP衣壳多肽相比，所述组装的病毒体表现出增加的肾上腺组织嗜性，其中工程化变体AAV5 VP衣壳多肽序列具有一个或多个突变，其中VP1多肽序列在从对应于SEQ ID NO:2中的581的位置至对应于SEQ ID NO:2中的589的位置的区域中具有所述一个或多个突变，并且其中所述一个或多个突变选自以下规则。在一些实施方案中，Xaa1选自A、C、K、Q、R、S或T。在一些实施方案中，Xaa1选自C、K或R。在一些实施方案中，Xaa1为C。在一些实施方案中，Xaa2选自A、C、I、S、T或V。在一些实施方案中，Xaa2选自A、V或T。在一些实施方案中，Xaa2为V。在一些实施方案中，Xaa3选自A、F、G、K、M、Q、R、T或V。在一些实施方案中，Xaa3选自A、G或M。在一些实施方案中，Xaa3为M。在一些实施方案中，Xaa4选自A、K、M、Q、R或V。在一些实施方案中，Xaa4选自A、R或K。在一些实施方案中，Xaa4为K。在一些实施方案中，Xaa5选自F、I、L、M、R、T、V或Y。在一些实施方案中，Xaa5选自R、V或Y。在一些实施方案中，Xaa5为V。在一些实施方案中，Xaa6选自G、H、M、N、R或S。在一些实施方案中，Xaa6选自H或N。在一些实施方案中，Xaa6为N。在一些实施方案中，Xaa7选自A、H、K、Q、R、S或V。在一些实施方案中，Xaa7选自H、Q或V。在一些实施方案中，Xaa7为H。在一些实施方案中，Xaa8选自A、G、H、M、Q或S。在一些实施方案中，Xaa8选自A、G、M或S。在一些实施方案中，Xaa8为S。在一些实施方案中，Xaa9选自A、E、N、P、R、S或Y。在一些实施方案中，Xaa9选自P或E。在一些实施方案中，Xaa9为P。

[0305] 本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽，与野生型AAV5 VP衣壳多肽相比，所述组装的病毒体表现出增加的肾上腺组织嗜性，其中工程化变体AAV5 VP衣壳多肽序列具有一个或多个突变，其中VP1多肽序列在从对应于SEQ ID NO:2中的581的位置至对应于SEQ ID NO:2中的589的位置的区域中具有所述一个或多个突变，并且其中所述一个或多个突变选自以下规则:Xaa1选自A、C、K、Q、R、S或T，Xaa2选自A、C、I、S、T或V，Xaa3选自A、F、G、K、M、Q、R、T或V，Xaa4选自A、K、M、Q、R或V，Xaa5选自F、I、L、M、R、T、V或Y，Xaa6选自G、H、M、N、R或S，Xaa7选自A、H、K、Q、R、S或V，Xaa8选自A、G、H、M、Q或S，并且Xaa9选自A、E、N、P、R、S或Y。

[0306] 本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽，与野生型AAV5 VP衣壳多肽相比，所述组装的病毒体表现出增加的肾上腺组织嗜性，其中工程化变体AAV5 VP衣壳多肽序列具有一个或多个突变，其中VP1多肽序列在从对应于SEQ ID NO:2中的581的位置至对应于SEQ ID NO:2中的589的位置的区域中具有所述一个或多个突变，并且其中所述一个或多个突变如下:Xaa1选自A、C、K、Q、R、S或T，Xaa2选自A、C、I、S、T或V，Xaa3选自A、F、G、K、M、Q、R、T或V，Xaa4选自A、K、M、Q、R或V，Xaa5选自F、I、L、M、R、T、V或Y，Xaa6选自G、H、M、N、R或S，Xaa7选自A、H、K、Q、R、S或V，Xaa8选自A、G、H、M、Q或S，Xaa9选自A、E、N、P、R、S或Y，或它

们的任何组合。

[0307] 在一些实施方案中，本文提供了AAV5 VP衣壳多肽，所述多肽在从对应于581的位置至对应于589的位置的区域中具有至少一个突变并且与选自SEQ ID NO:1118-SEQ ID NO:2117的任何序列具有至少70％、至少75％、至少80％、至少85％、至少90％、至少92％、至少95％、至少97％、至少99％或100％序列同一性，其中所述至少一个突变驱动增加的肾上腺组织嗜性。

[0308] B.ML规则

[0309] 对于本段中描述的以下一组规则，使用体内数据和两个ML模型(描述于实施例31)确定工程化AAV5 VP1衣壳多肽的581至589区域中每个位置处的有利生物物理特性和有利氨基酸残基。本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽，与具有野生型AAV5 VP衣壳多肽的rAAV病毒体相比，所述组装的病毒体表现出增加的肾上腺组织嗜性，其中工程化AAV5 VP衣壳多肽序列具有一个或多个突变，其中工程化AAV5 VP1衣壳多肽序列在从对应于SEQ ID NO:2中的581的位置至对应于SEQ ID NO:2中的589的位置的区域中具有所述一个或多个突变，并且其中Xaa1选自Xaa1处低摩尔质量的氨基酸(例如，Xaa1选自V、P、S或C)；或者其中Xaa1选自低亲疏水性的氨基酸(例如，Xaa1选自T、S、W或Y)；或者其中Xaa2选自低亲疏水性的氨基酸(例如，Xaa2选自R)；或者其中Xaa2选自低可突变性的氨基酸(例如，Xaa2选自C)；或者其中Xaa2选自低溶解性的氨基酸(例如，Xaa2选自K)；或者其中Xaa3选自低平均柔性的氨基酸(例如，Xaa3选自W、M或F)；或者其中Xaa3选自高溶解性的氨基酸(例如，Xaa3选自M)；或者其中Xaa4选自高表面可接近性的氨基酸(例如，Xaa4选自K或R)；或者其中Xaa4选自高平均柔性的氨基酸(例如，Xaa4选自K、I或N)；或者其中Xaa5选自中等可突变性的氨基酸(例如，Xaa5选自R或H)；或者其中Xaa5选自高goldman engelman steitz的氨基酸(例如，Xaa5选自V或L)；或者其中Xaa5选自低亲疏水性的氨基酸(例如，Xaa5选自R)；或者其中Xaa5选自高体积的氨基酸(例如，Xaa5选自Y、R或F)；或者其中Xaa6选自高溶解性的氨基酸(例如，Xaa6选自Y、V、M、A或C)；或者其中Xaa7选自中等可突变性的氨基酸(例如，Xaa7选自V、H或R)；或者其中Xaa7选自低溶解性的氨基酸(例如，Xaa7选自R)；或者其中Xaa8选自高平均柔性的氨基酸(例如，Xaa8选自K、I或N)；或者其中Xaa8选自高摩尔质量的氨基酸(例如，Xaa8选自R或Y)；或者其中Xaa9选自高可突变性的氨基酸(例如，Xaa9选自N)；或它们的任何组合。

[0310] 在一些实施方案中，Xaa1选自低摩尔质量的氨基酸。在一些实施方案中，Xaa1选自V、P、S或C。在一些实施方案中，Xaa1选自低亲疏水性的氨基酸。在一些实施方案中，Xaa1选自T、S、W或Y。在一些实施方案中，Xaa2选自低亲疏水性的氨基酸。在一些实施方案中，Xaa2选自R。在一些实施方案中，Xaa2选自低可突变性的氨基酸。在一些实施方案中，Xaa2选自C。在一些实施方案中，Xaa2选自低溶解性的氨基酸。在一些实施方案中，Xaa2选自K。在一些实施方案中，Xaa3选自低平均柔性的氨基酸。在一些实施方案中，Xaa3选自W、M或F。在一些实施方案中，Xaa3选自高溶解性的氨基酸。在一些实施方案中，Xaa3选自M。在一些实施方案中，Xaa4选自高表面可接近性的氨基酸。在一些实施方案中，Xaa4选自K或R。在一些实施方案中，Xaa4选自高平均柔性的氨基酸。在一些实施方案中，Xaa4选自K、I或N。在一些实施方案中，Xaa5选自中等可突变性的氨基酸。在一些实施方案中，Xaa5选自R、H。在一些实施方案中，Xaa5选自高goldman engelman steitz的氨基酸。在一些实施方案中，Xaa5选自V、L。在

一些实施方案中,Xaa5选自低亲疏水性的氨基酸。在一些实施方案中,Xaa5选自R。在一些实施方案中,Xaa5选自高体积的氨基酸。在一些实施方案中,Xaa5选自Y、R或F。在一些实施方案中,Xaa6选自高溶解性的氨基酸。在一些实施方案中,Xaa6选自Y、V、M、A或C。在一些实施方案中,Xaa7选自中等可突变性的氨基酸。在一些实施方案中,Xaa7选自V、H或R。在一些实施方案中,Xaa7选自低溶解性的氨基酸。在一些实施方案中,Xaa7选自R。在一些实施方案中,Xaa8选自高平均柔性的氨基酸。在一些实施方案中,Xaa8选自K、I或N。在一些实施方案中,Xaa8选自高摩尔质量的氨基酸。在一些实施方案中,Xaa8选自R或Y。在一些实施方案中,Xaa9选自高可突变性的氨基酸。在一些实施方案中,Xaa9选自N。

[0311] 在一些实施方案中,本文提供了AAV5 VP衣壳多肽,所述多肽在从对应于AAV5 VP1的581的位置至对应于589的位置的区域中具有至少一个突变并且与选自SEQ ID NO:2118-SEQ ID NO:3117的任何序列具有至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%、至少99%或100%序列同一性,其中所述至少一个突变驱动增加的肾上腺组织嗜性。

[0312] C.富集的肾上腺序列

[0313] 在一些实施方案中,本文提供了AAV5 VP衣壳多肽,所述多肽在从对应于AAV5 VP1的581的位置至对应于589的位置的区域中具有至少一个突变并且与选自SEQ ID NO:3118-SEQ ID NO:4117的任何序列具有至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%、至少99%或100%序列同一性,其中所述至少一个突变驱动增加的肾上腺组织嗜性。

[0314] 6.7.9.体内选定的赋予增加的坐骨神经嗜性的突变VP多肽

[0315] 本公开提供了具有VP衣壳多肽的AAV5病毒体,所述VP衣壳多肽在具有与靶细胞(例如,目标靶坐骨神经组织中的靶坐骨神经细胞)相互作用的残基的区域中具有至少一个突变,其中与野生型VP衣壳多肽相比,所述至少一个突变赋予增加的坐骨神经组织嗜性。在一些实施方案中,本文提供了与SEQ ID NO:1具有至少80%序列同源性的AAV5 VP1衣壳多肽,其中所述AAV5 VP1衣壳多肽在从对应于SEQ ID NO:1的581的位置至对应于589的位置的区域中具有至少一个突变,并且其中所述至少一个突变驱动增加的坐骨神经嗜性。以下序列规则和序列也适用于对应于AAV5 VP1氨基酸残基581至589的AAV5 VP2(氨基酸残基445至453;VP2序列显示于SEQ ID NO:111 5)和AAV5 VP3(氨基酸残基389至397;VP3序列显示于SEQ ID NO:1116)中的区域。因此,本公开涵盖在对应于581至589区域的AAV5 VP1氨基酸残基的VP2和VP3区域中具有一个或多个突变的AAV5 VP2衣壳多肽和AAV5 VP3衣壳多肽,其中所述一个或多个突变与以下章节中的规则或顺序相称。

[0316] A.位置频率规则

[0317] 在本章节中,除非另外说明,否则分析相对于给定氨基酸残基出现在所有其他所收获组织(CNS(前脑皮层、枕叶皮层、颞叶皮层、丘脑、下丘脑、黑质、海马DG、海马CA1、海马CA3、小脑)、肝脏、骨骼肌、心脏、肺、脾脏、淋巴结、骨髓、乳腺、皮肤、肾上腺、甲状腺、结肠和脊髓组织)中鉴定的变体中指定位置的频率而言该给定氨基酸残基出现在坐骨神经中鉴定的变体中对应于SEQ ID NO:1(推及于SEQ ID NO:2)的位置581至位置589的所述指定位置的频率,以确定优先靶向坐骨神经组织的衣壳的一组序列规则。来自体内数据的位置频率规则的确定详细描述于实施例8中。

[0318] 本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽，与野生型AAV5 VP衣壳多肽相比，所述组装的病毒体表现出增加的坐骨神经组织嗜性，其中工程化变体AAV5 VP1衣壳多肽序列具有一个或多个突变，其中Xaa1选自C、G、K、M、Q、R或Y，或Xaa1选自C、R或Q，或Xaa1为C；或Xaa2选自A、C、F、I、Q、T或V，或Xaa2选自A、C或I，或Xaa2为A；或Xaa3选自A、F、I、M、R、S或T，或Xaa3选自F、M、R或S，或Xaa3为R；或Xaa4选自E、N、T、Q或V，或Xaa4选自E、T或V，或Xaa4为T；或Xaa5选自F、H、Q、S、V或Y，或Xaa5选自F、V或Y，或Xaa5为V；或Xaa6选自K、M、N、Q、S或V，或Xaa6选自M、N或S，或Xaa6为N；或Xaa7选自K、M、Q、R或T，或Xaa7选自M、Q或T，或Xaa7为M；或Xaa8选自A、G、H、Q、S或V，或Xaa8选自H或S，或Xaa8为H；或Xaa9选自C、E、I、K或R，或Xaa9选自C、I或K，或Xaa9为I。

[0319] 本文是能够形成组装的病毒体的工程化AAV5 VP衣壳多肽，与野生型AAV5 VP衣壳多肽相比，所述组装的病毒体表现出增加的坐骨神经组织嗜性，其中工程化变体AAV5 VP衣壳多肽序列具有一个或多个突变，其中VP1多肽序列在从对应于SEQ ID NO:2中的581的位置至对应于SEQ ID NO:2中的589的位置的区域中具有所述一个或多个突变，并且其中所述一个或多个突变选自以下规则。在一些实施方案中，Xaa1选自C、G、K、M、Q、R或Y。在一些实施方案中，Xaa1选自C、R或Q。在一些实施方案中，Xaa1为C。在一些实施方案中，Xaa2选自A、C、F、I、Q、T或V。在一些实施方案中，Xaa2选自A、C或I。在一些实施方案中，Xaa2为A。在一些实施方案中，Xaa3选自A、F、I、M、R、S或T。在一些实施方案中，Xaa3选自F、M、R或S。在一些实施方案中，Xaa3为R。在一些实施方案中，Xaa4选自E、N、T、Q或V。在一些实施方案中，Xaa4选自E、T或V。在一些实施方案中，Xaa4为T。在一些实施方案中，Xaa5选自F、H、Q、S、V或Y。在一些实施方案中，Xaa5选自F、V或Y。在一些实施方案中，Xaa5为V。在一些实施方案中，Xaa6选自K、M、N、Q、S或V。在一些实施方案中，Xaa6选自M、N或S。在一些实施方案中，Xaa6为N。在一些实施方案中，Xaa7选自K、M、Q、R或T。在一些实施方案中，Xaa7选自M、Q或T。在一些实施方案中，Xaa7为M。在一些实施方案中，Xaa8选自A、G、H、Q、S或V。在一些实施方案中，Xaa8选自H或S。在一些实施方案中，Xaa8为H。在一些实施方案中，Xaa9选自C、E、I、K或R。在一些实施方案中，Xaa9选自C、I或K。在一些实施方案中，Xaa9为I。

[0320] 本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽，与野生型AAV5 VP衣壳多肽相比，所述组装的病毒体表现出增加的坐骨神经组织嗜性，其中工程化变体AAV5 VP衣壳多肽序列具有一个或多个突变，其中VP1多肽序列在从对应于SEQ ID NO:2中的581的位置至对应于SEQ ID NO:2中的589的位置的区域中具有所述一个或多个突变，并且其中所述一个或多个突变选自以下规则:Xaa1选自C、G、K、M、Q、R或Y，Xaa2选自A、C、F、I、Q、T或V，Xaa3选自A、F、I、M、R、S或T，Xaa4选自E、N、T、Q或V，Xaa5选自F、H、Q、S、V或Y，Xaa6选自K、M、N、Q、S或V，Xaa7选自K、M、Q、R或T，Xaa8选自A、G、H、Q、S或V和Xaa9选自C、E、I、K或R。

[0321] 本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽，与野生型AAV5 VP衣壳多肽相比，所述组装的病毒体表现出增加的坐骨神经组织嗜性，其中工程化变体AAV5 VP衣壳多肽序列具有一个或多个突变，其中VP1多肽序列在从对应于SEQ ID NO:2中的581的位置至对应于SEQ ID NO:2中的589的位置的区域中具有所述一个或多个突变，并且其中所述一个或多个突变如下:Xaa1选自C、G、K、M、Q、R或Y，Xaa2选自A、C、F、I、Q、T或V，Xaa3选自A、F、I、M、R、S或T，Xaa4选自E、N、T、Q或V，Xaa5选自F、H、Q、S、V或Y，Xaa6选自K、M、N、Q、S或V，Xaa7选自K、M、Q、R或T，Xaa8选自A、G、H、Q、S或V，Xaa9选自C、E、I、K或R，或它们的任何组合。

[0322] 在一些实施方案中，本文提供了AAV5 VP衣壳多肽，所述多肽在从对应于AAV5 VP1的581的位置至对应于589的位置的区域中具有至少一个突变并且与选自SEQ ID NO:26118-SEQ ID NO:26990的任何序列具有至少70％、至少75％、至少80％、至少85％、至少90％、至少92％、至少95％、至少97％、至少99％或100％序列同一性，其中所述至少一个突变驱动增加的坐骨神经组织嗜性。

[0323] B.ML规则

[0324] 对于本段中描述的以下一组规则，使用体内数据和两个ML模型(描述于实施例38)确定工程化AAV5 VP1衣壳多肽的581至589区域中每个位置处的有利生物物理特性和有利氨基酸残基。本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽，与具有野生型AAV5 VP衣壳多肽的rAAV病毒体相比，所述组装的病毒体表现出增加的坐骨神经组织嗜性，其中工程化AAV5 VP衣壳多肽序列具有一个或多个突变，其中工程化AAV5 VP1衣壳多肽序列在从对应于SEQ ID NO:2中的581的位置至对应于SEQ ID NO:2中的589的位置的区域中具有所述一个或多个突变，并且其中Xaa1选自高平均柔性的氨基酸(例如，Xaa1选自G或R)；或者其中Xaa1选自低溶解性的氨基酸(例如，Xaa1选自R或Q)；或者其中Xaa1选自低可突变性的氨基酸(例如，Xaa1选自C、L、F、Y、R、K、P或H)；或者其中Xaa1选自高体积的氨基酸(例如，Xaa1选自Y、F)；或者其中Xaa2选自高表面可接近性的氨基酸(例如，Xaa2选自E、R或K)；或者其中Xaa3选自中等可突变性的氨基酸(例如，Xaa3选自H或R)；或者其中Xaa3选自中等平均柔性的氨基酸(例如，Xaa3选自V或Y)；或者其中Xaa4选自高可突变性的氨基酸(例如，Xaa4选自N)；或者其中Xaa4选自高平均柔性的氨基酸(例如，Xaa4选自I、N、G或R)；或者其中Xaa4选自低溶解性的氨基酸(例如，Xaa4选自N)；或者其中Xaa6选自低可突变性的氨基酸(例如，Xaa6选自C、L、F或Y)；或者其中Xaa6选自高体积的氨基酸(例如，Xaa6选自K、M、I或L)；或者其中Xaa7选自低可突变性的氨基酸(例如，Xaa7选自L、F或Y)；或者其中Xaa7选自中等摩尔质量的氨基酸(例如，Xaa7选自D、I、L或N)；或者其中Xaa8选自高表面可接近性的氨基酸(例如，Xaa8选自S、Y、T、D、P、H或N)；或者其中Xaa9选自低可突变性的氨基酸(例如，Xaa9选自C、H、R)；或者其中Xaa9选自中等溶解性的氨基酸(例如，Xaa9选自Q、T或C)；或者其中Xaa9选自低表面可接近性的氨基酸(例如，Xaa9选自C)；或它们的任何组合。

[0325] 在一些实施方案中，Xaa1选自高平均柔性的氨基酸。在一些实施方案中，Xaa1选自G或R。在一些实施方案中，Xaa1选自低溶解性的氨基酸。在一些实施方案中，Xaa1选自R或Q。在一些实施方案中，Xaa1选自低可突变性的氨基酸。在一些实施方案中，Xaa1选自C、L、F、Y、R、K、P或H。在一些实施方案中，Xaa1选自高体积的氨基酸。在一些实施方案中，Xaa1选自Y或F。在一些实施方案中，Xaa2选自高表面可接近性的氨基酸。在一些实施方案中，Xaa2选自E、R或K。在一些实施方案中，Xaa3选自中等可突变性的氨基酸。在一些实施方案中，Xaa3选自H或R。在一些实施方案中，Xaa3选自中等平均柔性的氨基酸。在一些实施方案中，Xaa3选自V或Y。在一些实施方案中，Xaa4选自高可突变性的氨基酸。在一些实施方案中，Xaa4选自N。在一些实施方案中，Xaa4选自高平均柔性的氨基酸。在一些实施方案中，Xaa4选自I、N、G或R。在一些实施方案中，Xaa4选自低溶解性的氨基酸。在一些实施方案中，Xaa4选自N。在一些实施方案中，Xaa6选自低可突变性的氨基酸。在一些实施方案中，Xaa6选自C、L、F或Y。在一些实施方案中，Xaa6选自高体积的氨基酸。在一些实施方案中，Xaa6选自K、M、I或L。在一些实施方案中，Xaa7选自低可突变性的氨基酸。在一些实施方案中，Xaa7选自L、F或Y。在一些实

施方案中,Xaa7选自中等摩尔质量的氨基酸。在一些实施方案中,Xaa7选自D、I、L或N。在一些实施方案中,Xaa8选自高表面可接近性的氨基酸。在一些实施方案中,Xaa8选自S、Y、T、D、P、H或N。在一些实施方案中,Xaa9选自低可突变性的氨基酸。在一些实施方案中,Xaa9选自C、H或R。在一些实施方案中,Xaa9选自中等溶解性的氨基酸。在一些实施方案中,Xaa9选自Q、T或C。在一些实施方案中,Xaa9选自低表面可接近性的氨基酸。在一些实施方案中,Xaa9选自C。

[0326] 在一些实施方案中,本文提供了AAV5 VP衣壳多肽,所述多肽在从对应于AAV5 VP1的581的位置至对应于589的位置的区域中具有至少一个突变并且与选自SEQ ID NO:26991-SEQ ID NO:27990的任何序列具有至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%、至少99%或100%序列同一性,其中所述至少一个突变驱动增加的坐骨神经组织嗜性。

[0327] C.富集的坐骨神经序列

[0328] 在一些实施方案中,本文提供了AAV5 VP衣壳多肽,所述多肽在从对应于AAV5 VP1的581的位置至对应于589的位置的区域中具有至少一个突变并且与选自SEQ ID NO:27991-SEQ ID NO:28990的任何序列具有至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%、至少99%或100%序列同一性,其中所述至少一个突变驱动增加的坐骨神经组织嗜性。

[0329] 6.7.10.体内选定的赋予增加的骨骼肌嗜性的突变VP多肽

[0330] 本公开提供了具有VP衣壳多肽的AAV5病毒体,所述VP衣壳多肽在具有与靶细胞(例如,目标靶骨骼肌组织中的靶骨骼肌细胞)相互作用的残基的区域中具有至少一个突变,其中与野生型VP衣壳多肽相比,所述至少一个突变赋予增加的骨骼肌组织嗜性。在一些实施方案中,本文提供了与SEQ ID NO:1具有至少80%序列同源性的AAV5 VP1衣壳多肽,其中所述AAV5 VP1衣壳多肽在从对应于SEQ ID NO:1的581的位置至对应于589的位置的区域中具有至少一个突变,并且其中所述至少一个突变驱动增加的骨骼肌嗜性。以下序列规则和序列也适用于对应于AAV5 VP1氨基酸残基581至589的AAV5 VP2(氨基酸残基445至453;VP2序列显示于SEQ ID NO:111 5)和AAV5 VP3(氨基酸残基389至397;VP3序列显示于SEQ ID NO:1116)中的区域。因此,本公开涵盖在对应于581至589区域的AAV5 VP1氨基酸残基的VP2和VP3区域中具有一个或多个突变的AAV5 VP2衣壳多肽和AAV5 VP3衣壳多肽,其中所述一个或多个突变与以下章节中的规则或顺序相称。

[0331] A.位置频率规则

[0332] 在本章节中,除非另外说明,否则分析相对于给定氨基酸残基出现在所有其他所收获组织(CNS(前脑皮层、枕叶皮层、颞叶皮层、丘脑、下丘脑、黑质、海马DG、海马CA1、海马CA3、小脑)、肝脏、心脏、肺、脾脏、淋巴结、骨髓、乳腺、皮肤、肾上腺、甲状腺、结肠、坐骨神经和脊髓组织)中鉴定的变体中指定位置的频率而言该给定氨基酸残基出现在骨骼肌中鉴定的变体中对应于SEQ ID NO:1(推及于SEQ ID NO:2)的位置581至位置589的所述指定位置的频率,以确定优先靶向骨骼肌组织的衣壳的一组序列规则。来自体内数据的位置频率规则的确定详细描述于实施例9中。

[0333] 本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽,与野生型AAV5 VP衣壳多肽相比,所述组装的病毒体表现出增加的骨骼肌组织嗜性,其中工程化变体AAV5

VP1衣壳多肽序列具有一个或多个突变，其中Xaa1选自A、E、H、M、P、Q或S，或Xaa1选自P或Q，或Xaa1为Q；或Xaa2选自F、H、I、T或V，或Xaa2选自T或V，或Xaa2为V；或Xaa3选自A、G、I、K、M、Q、R、S、T或V，或Xaa3选自A、L、P、R或T，或Xaa3选自L、P或T，或Xaa3为P；或Xaa4选自D、E、G、P或S，或Xaa4选自D、E或S，或Xaa4为E；或Xaa5选自H、L、M、P或V，或Xaa5选自L、M或V，或Xaa5为L；或Xaa6选自E、H、N或P，或Xaa6为P；或Xaa7选自A、H、N、Q或T，或Xaa7为H；或Xaa8选自I、K、M、P或W，或Xaa8选自I、P或W，或Xaa8为P；或Xaa9选自A、I、M、P或V，或Xaa9选自A、M或P，或Xaa9为M。

[0334] 本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽，与野生型AAV5 VP衣壳多肽相比，所述组装的病毒体表现出增加的骨骼肌组织嗜性，其中工程化变体AAV5 VP衣壳多肽序列具有一个或多个突变，其中VP1多肽序列在从对应于SEQ ID NO:2中的581的位置至对应于SEQ ID NO:2中的589的位置的区域中具有所述一个或多个突变，并且其中所述一个或多个突变选自以下规则。在一些实施方案中，Xaa1选自A、E、H、M、P、Q或S。在一些实施方案中，Xaa1选自P或Q。在一些实施方案中，Xaa1为Q。在一些实施方案中，Xaa2选自F、H、I、T或V。在一些实施方案中，Xaa2选自T或V。在一些实施方案中，Xaa2为V。在一些实施方案中，Xaa3选自A、G、I、K、M、Q、R、S、T或V。在一些实施方案中，Xaa3选自A、L、P、R或T。在一些实施方案中，Xaa3选自L、P或T。在一些实施方案中，Xaa3为P。在一些实施方案中，Xaa4选自D、E、G、P或S。在一些实施方案中，Xaa4选自D、E或S。在一些实施方案中，Xaa4为E。在一些实施方案中，Xaa5选自H、L、M、P或V。在一些实施方案中，Xaa5选自L、M或V。在一些实施方案中，Xaa5为L。在一些实施方案中，Xaa6选自E、H、N或P。在一些实施方案中，Xaa6为P。在一些实施方案中，Xaa7选自A、H、N、Q或T。在一些实施方案中，Xaa7为H。在一些实施方案中，Xaa8选自I、K、M、P或W。在一些实施方案中，Xaa8选自I、P或W。在一些实施方案中，Xaa8为P。在一些实施方案中，Xaa9选自A、I、M、P或V。在一些实施方案中，Xaa9选自A、M或P。在一些实施方案中，Xaa9为M。

[0335] 本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽，与野生型AAV5 VP衣壳多肽相比，所述组装的病毒体表现出增加的骨骼肌组织嗜性，其中工程化变体AAV5 VP衣壳多肽序列具有一个或多个突变，其中VP1多肽序列在从对应于SEQ ID NO:2中的581的位置至对应于SEQ ID NO:2中的589的位置的区域中具有所述一个或多个突变，并且其中所述一个或多个突变选自以下规则：Xaa1选自A、E、H、M、P、Q或S，Xaa2选自F、H、I、T或V，Xaa3选自A、G、I、K、M、Q、R、S、T或V，Xaa4选自D、E、G、P或S，Xaa5选自H、L、M、P或V，Xaa6选自E、H、N或P，Xaa7选自A、H、N、Q或T，Xaa8选自I、K、M、P或W，并且Xaa9选自A、I、M、P或V。

[0336] 本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽，与野生型AAV5 VP衣壳多肽相比，所述组装的病毒体表现出增加的骨骼肌组织嗜性，其中工程化变体AAV5 VP衣壳多肽序列具有一个或多个突变，其中VP1多肽序列在从对应于SEQ ID NO:2中的581的位置至对应于SEQ ID NO:2中的589的位置的区域中具有所述一个或多个突变，并且其中所述一个或多个突变如下：Xaa1选自A、E、H、M、P、Q或S，Xaa2选自F、H、I、T或V，Xaa3选自A、G、I、K、M、Q、R、S、T或V，Xaa4选自D、E、G、P或S，Xaa5选自H、L、M、P或V，Xaa6选自E、H、N或P，Xaa7选自A、H、N、Q或T，Xaa8选自I、K、M、P或W，Xaa9选自A、I、M、P或V，或它们的任何组合。

[0337] 在一些实施方案中，本文提供了AAV5 VP衣壳多肽，所述多肽在从对应于AAV5 VP1的581的位置至对应于589的位置的区域中具有至少一个突变并且与选自SEQ ID NO:

28991-SEQ ID NO:29990的任何序列具有至少70％、至少75％、至少80％、至少85％、至少90％、至少92％、至少95％、至少97％、至少99％或100％序列同一性，其中所述至少一个突变驱动增加的骨骼肌组织嗜性。

[0338] B.ML规则

[0339] 对于本段中描述的以下一组规则，使用体内数据和两个ML模型(描述于实施例39)确定工程化AAV5 VP1衣壳多肽的581至589区域中每个位置处的有利生物物理特性和有利氨基酸残基。本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽，与具有野生型AAV5 VP衣壳多肽的rAAV病毒体相比，所述组装的病毒体表现出增加的骨骼肌组织嗜性，其中工程化AAV5 VP衣壳多肽序列具有一个或多个突变，其中工程化AAV5 VP1衣壳多肽序列在从对应于SEQ ID NO:2中的581的位置至对应于SEQ ID NO:2中的589的位置的区域中具有所述一个或多个突变，并且其中Xaa1选自高平均柔性的氨基酸(例如，Xaa1选自G或R)；或者其中Xaa1选自低平均柔性的氨基酸(例如，Xaa1选自W、M、F或H)；或者其中Xaa1选自高摩尔质量的氨基酸(例如，Xaa1选自R、F或W)；或者其中Xaa2选自低亲疏水性的氨基酸(例如，Xaa2选自K或R)；或者其中Xaa2选自低可突变性的氨基酸(例如，Xaa2选自C、R或H)；或者其中Xaa2选自高平均柔性的氨基酸(例如，Xaa2选自G或R)；或者其中Xaa3选自高平均柔性的氨基酸(例如，Xaa3选自G或R)；或者其中Xaa4选自高亲水性的氨基酸(例如，Xaa4选自D、E、R、K或N)；或者其中Xaa4选自低可突变性的氨基酸(例如，Xaa4选自C、R、H)；或者其中Xaa5选自低摩尔质量的氨基酸(例如，Xaa5选自A)；或者其中Xaa5选自低平均柔性的氨基酸(例如，Xaa5选自A或L)；或者其中Xaa5选自高可突变性的氨基酸(例如，Xaa5选自D、A或E)；或者其中Xaa6选自低平均柔性的氨基酸(例如，Xaa6选自W、M或F)；或者其中Xaa6选自低可突变性的氨基酸(例如，Xaa6选自C)；或者其中Xaa6选自高摩尔质量的氨基酸(例如，Xaa6选自W)；或者其中Xaa7选自低goldman engelman steitz的氨基酸(例如，Xaa7选自R)；或者其中Xaa7选自高平均柔性的氨基酸(例如，Xaa7选自D、R、P、G或S)；或者其中Xaa7选自高可突变性的氨基酸(例如，Xaa7选自R、H或N)；或者其中Xaa7选自低溶解性的氨基酸(例如，Xaa7选自R或Q)；或者其中Xaa8选自高亲水性的氨基酸(例如，Xaa8选自D、E、R、K或N)；或者其中Xaa9选自低可突变性的氨基酸(例如，Xaa9选自Y、F或L)；或它们的任何组合。

[0340] 在一些实施方案中，Xaa1选自高平均柔性的氨基酸。在一些实施方案中，Xaa1选自G或R。在一些实施方案中，Xaa1选自低平均柔性的氨基酸。在一些实施方案中，Xaa1选自W、M、F或H。在一些实施方案中，Xaa1选自高摩尔质量的氨基酸。在一些实施方案中，Xaa1选自R、F或W。在一些实施方案中，Xaa2选自低亲疏水性的氨基酸。在一些实施方案中，Xaa2选自K或R。在一些实施方案中，Xaa2选自低可突变性的氨基酸。在一些实施方案中，Xaa2选自C、R或H。在一些实施方案中，Xaa2选自高平均柔性的氨基酸。在一些实施方案中，Xaa2选自G或R。在一些实施方案中，Xaa3选自高平均柔性的氨基酸。在一些实施方案中，Xaa3选自G或R。在一些实施方案中，Xaa4选自高亲水性的氨基酸。在一些实施方案中，Xaa4选自D、E、R、K或N。在一些实施方案中，Xaa4选自低可突变性的氨基酸。在一些实施方案中，Xaa4选自C、R或H。在一些实施方案中，Xaa5选自低摩尔质量的氨基酸。在一些实施方案中，Xaa5选自A。在一些实施方案中，Xaa5选自低平均柔性的氨基酸。在一些实施方案中，Xaa5选自A或L。在一些实施方案中，Xaa5选自高可突变性的氨基酸。在一些实施方案中，Xaa5选自D、A或E。在一些实施方案中，Xaa6选自低平均柔性的氨基酸。在一些实施方案中，Xaa6选自W、M或F。在一些

实施方案中,Xaa6选自低可突变性的氨基酸。在一些实施方案中,Xaa6选自C。在一些实施方案中,Xaa6选自高摩尔质量的氨基酸。在一些实施方案中,Xaa6选自W。在一些实施方案中,Xaa7选自低goldman engelman steitz的氨基酸。在一些实施方案中,Xaa7选自R。在一些实施方案中,Xaa7选自高平均柔性的氨基酸。在一些实施方案中,Xaa7选自D、R、P、G或S。在一些实施方案中,Xaa7选自高可突变性的氨基酸。在一些实施方案中,Xaa7选自R、H或N。在一些实施方案中,Xaa7选自低溶解性的氨基酸。在一些实施方案中,Xaa7选自R或Q。在一些实施方案中,Xaa8选自高亲水性的氨基酸。在一些实施方案中,Xaa8选自D、E、R、K或N。在一些实施方案中,Xaa9选自低可突变性的氨基酸。在一些实施方案中,Xaa9选自Y、F或L。

[0341] 在一些实施方案中,本文提供了AAV5 VP衣壳多肽,所述多肽在从对应于AAV5 VP1的581的位置至对应于589的位置的区域中具有至少一个突变并且与选自SEQ ID NO:29991-SEQ ID NO:30990的任何序列具有至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%、至少99%或100%序列同一性,其中所述至少一个突变驱动增加的骨骼肌组织嗜性。

[0342] C.富集的骨骼肌序列

[0343] 在一些实施方案中,本文提供了AAV5 VP衣壳多肽,所述多肽在从对应于AAV5 VP1的581的位置至对应于589的位置的区域中具有至少一个突变并且与选自SEQ ID NO:30991-SEQ ID NO:31990的任何序列具有至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%、至少99%或100%序列同一性,其中所述至少一个突变驱动增加的骨骼肌组织嗜性。

[0344] 6.7.11.体内选定的赋予增加的脊髓嗜性的突变VP多肽

[0345] 本公开提供了具有VP衣壳多肽的AAV5病毒体,所述VP衣壳多肽在具有与靶细胞(例如,目标靶脊髓组织中的靶脊髓细胞)相互作用的残基的区域中具有至少一个突变,其中与野生型VP衣壳多肽相比,所述至少一个突变赋予增加的CNS组织嗜性。在一些实施方案中,本文提供了与SEQ ID NO:1具有至少80%序列同源性的AAV5VP1衣壳多肽,其中所述AAV5 VP1衣壳多肽在从对应于SEQ ID NO:1的581的位置至对应于589的位置的区域中具有至少一个突变,并且其中所述至少一个突变驱动增加的脊髓嗜性。以下序列规则和序列也适用于对应于AAV5 VP1氨基酸残基581至589的AAV5 VP2(氨基酸残基445至453;VP2序列显示于SEQ ID NO:1115)和AAV5 VP3(氨基酸残基389至397;VP3序列显示于SEQ ID NO:1116)中的区域。因此,本公开涵盖在对应于581至589区域的AAV5 VP1氨基酸残基的VP2和VP3区域中具有一个或多个突变的AAV5 VP2衣壳多肽和AAV5 VP3衣壳多肽,其中所述一个或多个突变与以下章节中的规则或顺序相称。

[0346] A.位置频率规则

[0347] 在本章节中,除非另外说明,否则分析相对于给定氨基酸残基出现在所有其他所收获组织(CNS(前脑皮层、枕叶皮层、颞叶皮层、丘脑、下丘脑、黑质、海马DG、海马CA1、海马CA3、小脑)、肝脏、骨骼肌、心脏、肺、脾脏、淋巴结、骨髓、乳腺、皮肤、肾上腺、甲状腺、结肠和坐骨神经组织)中鉴定的变体中指定位置的频率而言该给定氨基酸残基出现在脊髓中鉴定的变体中对应于SEQ ID NO:1(推及于SEQ ID NO:2)的位置581至位置589的所述指定位置的频率,以确定优先靶向脊髓组织的衣壳的一组序列规则。来自体内数据的位置频率规则的确定详细描述于实施例10中。

[0348] 本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽，与野生型AAV5 VP衣壳多肽相比，所述组装的病毒体表现出增加的脊髓组织嗜性，其中工程化变体AAV5 VP1衣壳多肽序列具有一个或多个突变，其中Xaa1选自A、C、K、Q、R、S或W，或Xaa1选自K、R或W，或Xaa1为K；或Xaa2选自H、I、K、L、T、V或W，或Xaa2选自H、I或T，或Xaa2为I；或Xaa3选自C、F、G、H、I、K、N或R，或Xaa3选自F、I或R，或Xaa3为I；或Xaa4选自I、M、Q、S或V，或Xaa4选自I、M或V，或Xaa4为V；或Xaa5选自H、K、Q、T、W或Y，或Xaa5选自T、W或Y，或Xaa5为Y；或Xaa6选自H、L、N、Q、R、W或Y，或Xaa6选自L、N、R或Y，或Xaa6为Y；或Xaa7选自D、H、P、Q或R，或Xaa7为R；或Xaa8选自D、F、L、S、T或Y，或Xaa8选自S、T或Y，或Xaa8为T；或Xaa9选自C、I、N、P、R、S或Y，或Xaa9选自I、P或R，或Xaa9为I。

[0349] 本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽，与野生型AAV5 VP衣壳多肽相比，所述组装的病毒体表现出增加的脊髓组织嗜性，其中工程化变体AAV5 VP衣壳多肽序列具有一个或多个突变，其中VP1多肽序列在从对应于SEQ ID NO:2中的581的位置至对应于SEQ ID NO:2中的589的位置的区域中具有所述一个或多个突变，并且其中所述一个或多个突变选自以下规则。在一些实施方案中，Xaa1选自A、C、K、Q、R、S或W。在一些实施方案中，Xaa1选自K、R或W。在一些实施方案中，Xaa1为K。在一些实施方案中，Xaa2选自H、I、K、L、T、V或W。在一些实施方案中，Xaa2选自H、I或T。在一些实施方案中，Xaa2为I。在一些实施方案中，Xaa3选自C、F、G、H、I、K、N或R。在一些实施方案中，Xaa3选自F、I或R。在一些实施方案中，Xaa3为I。在一些实施方案中，Xaa4选自I、M、Q、S或V。在一些实施方案中，Xaa4选自I、M或V。在一些实施方案中，Xaa4为V。在一些实施方案中，Xaa5选自H、K、Q、T、W或Y。在一些实施方案中，Xaa5选自T、W或Y。在一些实施方案中，Xaa5为Y。在一些实施方案中，Xaa6选自H、L、N、Q、R、W或Y。在一些实施方案中，Xaa6选自L、N、R或Y。在一些实施方案中，Xaa6为Y。在一些实施方案中，Xaa7选自D、H、P、Q或R。在一些实施方案中，Xaa7为R。在一些实施方案中，Xaa8选自D、F、L、S、T或Y。在一些实施方案中，Xaa8选自S、T或Y。在一些实施方案中，Xaa8为T。在一些实施方案中，Xaa9选自C、I、N、P、R、S或Y。在一些实施方案中，Xaa9选自I、P或R。在一些实施方案中，Xaa9为I。

[0350] 本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽，与野生型AAV5 VP衣壳多肽相比，所述组装的病毒体表现出增加的脊髓组织嗜性，其中工程化变体AAV5 VP衣壳多肽序列具有一个或多个突变，其中VP1多肽序列在从对应于SEQ ID NO:2中的581的位置至对应于SEQ ID NO:2中的589的位置的区域中具有所述一个或多个突变，并且其中所述一个或多个突变选自以下规则:Xaa1选自A、C、K、Q、R、S或W，Xaa2选自H、I、K、L、T、V或W，Xaa3选自C、F、G、H、I、K、N或R，Xaa4选自I、M、Q、S或V，Xaa5选自H、K、Q、T、W或Y，Xaa6选自H、L、N、Q、R、W或Y，Xaa7选自D、H、P、Q或R，Xaa8选自D、F、L、S、T或Y，并且Xaa9选自C、I、N、P、R、S或Y。

[0351] 本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽，与野生型AAV5 VP衣壳多肽相比，所述组装的病毒体表现出增加的脊髓组织嗜性，其中工程化变体AAV5 VP衣壳多肽序列具有一个或多个突变，其中VP1多肽序列在从对应于SEQ ID NO:2中的581的位置至对应于SEQ ID NO:2中的589的位置的区域中具有所述一个或多个突变，并且其中所述一个或多个突变如下:Xaa1选自A、C、K、Q、R、S或W，Xaa2选自H、I、K、L、T、V或W，Xaa3选自C、F、G、H、I、K、N或R，Xaa4选自I、M、Q、S或V，Xaa5选自H、K、Q、T、W或Y，Xaa6选自H、L、N、Q、R、W或Y，Xaa7选自D、H、P、Q或R，Xaa8选自D、F、L、S、T或Y，Xaa9选自C、I、N、P、R、S或Y，或它们的任何组

合。

[0352] 在一些实施方案中,本文提供了AAV5 VP衣壳多肽,所述多肽在从对应于AAV5 VP1的581的位置至对应于589的位置的区域中具有至少一个突变并且与选自SEQ ID NO:34991-SEQ ID NO:35437的任何序列具有至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%、至少99%或100%序列同一性,其中所述至少一个突变驱动增加的脊髓组织嗜性。

[0353] B.ML规则

[0354] 对于本段中描述的以下一组规则,使用体内数据和两个ML模型(描述于实施例41)确定工程化AAV5 VP1衣壳多肽的581至589区域中每个位置处的有利生物物理特性和有利氨基酸残基。本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽,与具有野生型AAV5 VP衣壳多肽的rAAV病毒体相比,所述组装的病毒体表现出增加的脊髓组织嗜性,其中工程化AAV5 VP衣壳多肽序列具有一个或多个突变,其中工程化AAV5 VP1衣壳多肽序列在从对应于SEQ ID NO:2中的581的位置至对应于SEQ ID NO:2中的589的位置的区域中具有所述一个或多个突变,并且其中Xaa1选自高体积的氨基酸(例如,Xaa1选自F、W或Y);或者其中Xaa1选自低可突变性的氨基酸(例如,Xaa1选自Y、F、L或C);或者其中Xaa1选自高溶解性的氨基酸(例如,Xaa1选自W、F、I或L);或者其中Xaa1选自低平均柔性的氨基酸(例如,Xaa1选自F、M或W);或者其中Xaa2选自低亲疏水性的氨基酸(例如,Xaa2选自P或Y);或者其中Xaa3选自低亲水性的氨基酸(例如,Xaa3选自Y、W、V、M、F、I或L);或者其中Xaa3选自高溶解性的氨基酸(例如,Xaa3选自W、F、I或L);或者其中Xaa6选自高体积的氨基酸(例如,Xaa6选自W、R、K、M、I或L);或者其中Xaa6选自高摩尔质量的氨基酸(例如,Xaa6选自W);或者其中Xaa8选自高摩尔质量的氨基酸(例如,Xaa8选自W、E、K、M、H或Q);或者其中Xaa8选自高体积的氨基酸(例如,Xaa8选自W、K、M、I或L);或者其中Xaa8选自高goldman engelman steitz的氨基酸(例如,Xaa8选自V或L);或者其中Xaa9选自高亲疏水性的氨基酸(例如,Xaa9选自V或I);或者其中Xaa9选自高溶解性的氨基酸(例如,Xaa9选自W、F、I或L);或它们的任何组合。

[0355] 在一些实施方案中,Xaa1选自高体积的氨基酸。在一些实施方案中,Xaa1选自F、W或Y。在一些实施方案中,Xaa1选自低可突变性的氨基酸。在一些实施方案中,Xaa1选自Y、F、L或C。在一些实施方案中,Xaa1选自高溶解性的氨基酸。在一些实施方案中,Xaa1选自W、F、I或L。在一些实施方案中,Xaa1选自低平均柔性的氨基酸。在一些实施方案中,Xaa1选自F、M或W。在一些实施方案中,Xaa2选自低亲疏水性的氨基酸。在一些实施方案中,Xaa2选自P或Y。在一些实施方案中,Xaa3选自低亲水性的氨基酸。在一些实施方案中,Xaa3选自Y、W、V、M、F、I或L。在一些实施方案中,Xaa3选自高溶解性的氨基酸。在一些实施方案中,Xaa3选自W、F、I、L。在一些实施方案中,Xaa6选自高体积的氨基酸。在一些实施方案中,Xaa6选自W、R、K、M、I或L。在一些实施方案中,Xaa6选自高摩尔质量的氨基酸。在一些实施方案中,Xaa6选自W。在一些实施方案中,Xaa8选自高摩尔质量的氨基酸。在一些实施方案中,Xaa8选自W、E、K、M、H或Q。在一些实施方案中,Xaa8选自高体积的氨基酸。在一些实施方案中,Xaa8选自W、K、M、I、L。在一些实施方案中,Xaa8选自高goldman engelman steitz的氨基酸。在一些实施方案中,Xaa8选自V或L。在一些实施方案中,Xaa9选自高亲疏水性的氨基酸。在一些实施方案中,Xaa9选自V或I。在一些实施方案中,Xaa9选自高溶解性的氨基酸。在一些实施方案中,Xaa9选自W、F、I或L。

[0356] 在一些实施方案中，本文提供了AAV5 VP衣壳多肽，所述多肽在从对应于AAV5 VP1的581的位置至对应于589的位置的区域中具有至少一个突变并且与选自SEQ ID NO:35438-SEQ ID NO:36437的任何序列具有至少70％、至少75％、至少80％、至少85％、至少90％、至少92％、至少95％、至少97％、至少99％或100％序列同一性，其中所述至少一个突变驱动增加的脊髓组织嗜性。

[0357] C.富集的脊髓序列

[0358] 在一些实施方案中，本文提供了AAV5 VP衣壳多肽，所述多肽在从对应于AAV5 VP1的581的位置至对应于589的位置的区域中具有至少一个突变并且与选自SEQ ID NO:36438-SEQ ID NO:37437的任何序列具有至少70％、至少75％、至少80％、至少85％、至少90％、至少92％、至少95％、至少97％、至少99％或100％序列同一性，其中所述至少一个突变驱动增加的脊髓组织嗜性。

[0359] 6.7.12.体内选定的赋予增加的乳腺嗜性的突变VP多肽

[0360] 本公开提供了具有VP衣壳多肽的AAV5病毒体，所述VP衣壳多肽在具有与靶细胞(例如，目标靶乳腺组织中的靶乳腺细胞)相互作用的残基的区域中具有至少一个突变，其中与野生型VP衣壳多肽相比，所述至少一个突变赋予增加的乳腺组织嗜性。在一些实施方案中，本文提供了与SEQ ID NO:1具有至少80％序列同源性的AAV5 VP1衣壳多肽，其中所述AAV5 VP1衣壳多肽在从对应于SEQ ID NO:1的581的位置至对应于589的位置的区域中具有至少一个突变，并且其中所述至少一个突变驱动增加的乳腺嗜性。以下序列规则和序列也适用于对应于AAV5 VP1氨基酸残基581至589的AAV5 VP2(氨基酸残基445至453;VP2序列显示于SEQ ID NO:1115)和AAV5 VP3(氨基酸残基389至397;VP3序列显示于SEQ ID NO:1116)中的区域。因此，本公开涵盖在对应于581至589区域的AAV5 VP1氨基酸残基的VP2和VP3区域中具有一个或多个突变的AAV5 VP2衣壳多肽和AAV5 VP3衣壳多肽，其中所述一个或多个突变与以下章节中的规则或顺序相称。

[0361] A.位置频率规则

[0362] 在本章节中，除非另外说明，否则分析相对于给定氨基酸残基出现在所有其他所收获组织(CNS(前脑皮层、枕叶皮层、颞叶皮层、丘脑、下丘脑、黑质、海马DG、海马CA1、海马CA3、小脑)、肝脏、骨骼肌、心脏、肺、脾脏、淋巴结、骨髓、皮肤、肾上腺、甲状腺、结肠、坐骨神经和脊髓组织)中鉴定的变体中指定位置的频率而言该给定氨基酸残基出现在乳腺中鉴定的变体中对应于SEQ ID NO:1(推及于SEQ ID NO:2)的位置581至位置589的所述指定位置的频率，以确定优先靶向乳腺组织的衣壳的一组序列规则。来自体内数据的位置频率规则的确定详细描述于实施例11中。

[0363] 本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽，与野生型AAV5 VP衣壳多肽相比，所述组装的病毒体表现出增加的乳腺组织嗜性，其中工程化变体AAV5 VP1衣壳多肽序列具有一个或多个突变，其中Xaa1选自C、K、M、Q、R或Y，或Xaa1选自C、Q或R，或Xaa1为C;或Xaa2选自A、F、I、K、S、T或V，或Xaa2选自A、S或V，或Xaa2为V;或Xaa3选自A、F、G、I、K、L、R、T或Y，或Xaa3选自F、G、K、R或Y，或Xaa3选自F、K或Y，或Xaa3为F;或Xaa4选自A、I、K、Q、R或T，或Xaa4选自A、I或R，或Xaa4为I;或Xaa5选自I、L、M、Q、R、T、V或Y，或Xaa5选自I、M或Y，或Xaa5为Y;或Xaa6选自H、N、S或V，或Xaa6为H;或Xaa7选自A、H、I、N、S或Y，或Xaa7为N或S，或Xaa7为N;或Xaa8选自A、C、D、G、H、M、Q或S，或Xaa8选自G、M或Q，或Xaa8为G;或Xaa9选自A、

E、L、W或Y，或Xaa9选自A、L或W，或Xaa9为A。

[0364] 本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽，与野生型AAV5 VP衣壳多肽相比，所述组装的病毒体表现出增加的乳腺组织嗜性，其中工程化变体AAV5 VP衣壳多肽序列具有一个或多个突变，其中VP1多肽序列在从对应于SEQ ID NO:2中的581的位置至对应于SEQ ID NO:2中的589的位置的区域中具有所述一个或多个突变，并且其中所述一个或多个突变选自以下规则。在一些实施方案中，Xaa1选自C、K、M、Q、R或Y。在一些实施方案中，Xaa1选自C、Q或R。在一些实施方案中，Xaa1为C。在一些实施方案中，Xaa2选自A、F、I、K、S、T或V。在一些实施方案中，Xaa2选自A、S或V。在一些实施方案中，Xaa2为V。在一些实施方案中，Xaa3选自A、F、G、I、K、L、R、T或Y。在一些实施方案中，Xaa3选自F、G、K、R或Y。在一些实施方案中，Xaa3选自F、K或Y。在一些实施方案中，Xaa3为F。在一些实施方案中，Xaa4选自A、I、K、Q、R或T。在一些实施方案中，Xaa4选自A、I或R。在一些实施方案中，Xaa4为I。在一些实施方案中，Xaa5选自I、L、M、Q、R、T、V或Y。在一些实施方案中，Xaa5选自I、M或Y。在一些实施方案中，Xaa5为Y。在一些实施方案中，Xaa6选自H、N、S或V。在一些实施方案中，Xaa6为H。在一些实施方案中，Xaa7选自A、H、I、N、S或Y。在一些实施方案中，Xaa7为N或S。在一些实施方案中，Xaa7为N。在一些实施方案中，Xaa8选自A、C、D、G、H、M、Q或S。在一些实施方案中，Xaa8选自G、M或Q。在一些实施方案中，Xaa8为G。在一些实施方案中，Xaa9选自A、E、L、W或Y。在一些实施方案中，Xaa9选自A、L或W。在一些实施方案中，Xaa9为A。

[0365] 本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽，与野生型AAV5 VP衣壳多肽相比，所述组装的病毒体表现出增加的乳腺组织嗜性，其中工程化变体AAV5 VP衣壳多肽序列具有一个或多个突变，其中VP1多肽序列在从对应于SEQ ID NO:2中的581的位置至对应于SEQ ID NO:2中的589的位置的区域中具有所述一个或多个突变，并且其中所述一个或多个突变选自以下规则:Xaa1选自C、K、M、Q、R或Y，Xaa2选自A、F、I、K、S、T或V，Xaa3选自A、F、G、I、K、L、R、T或Y，Xaa4选自A、I、K、Q、R或T，Xaa5选自I、L、M、Q、R、T、V或Y，Xaa6选自H、N、S或V，Xaa7选自A、H、I、N、S或Y，Xaa8选自A、C、D、G、H、M、Q或S，并且Xaa9选自A、E、L、W或Y。

[0366] 本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽，与野生型AAV5 VP衣壳多肽相比，所述组装的病毒体表现出增加的乳腺组织嗜性，其中工程化变体AAV5 VP衣壳多肽序列具有一个或多个突变，其中VP1多肽序列在从对应于SEQ ID NO:2中的581的位置至对应于SEQ ID NO:2中的589的位置的区域中具有所述一个或多个突变，并且其中所述一个或多个突变如下:Xaa1选自C、K、M、Q、R或Y，Xaa2选自A、F、I、K、S、T或V，Xaa3选自A、F、G、I、K、L、R、T或Y，Xaa4选自A、I、K、Q、R或T，Xaa5选自I、L、M、Q、R、T、V或Y，Xaa6选自H、N、S或V，Xaa7选自A、H、I、N、S或Y，Xaa8选自A、C、D、G、H、M、Q或S，Xaa9选自A、E、L、W或Y，或它们的任何组合。

[0367] 在一些实施方案中，本文提供了AAV5 VP衣壳多肽，所述多肽在从对应于AAV5 VP1的581的位置至对应于589的位置的区域中具有至少一个突变并且与选自SEQ ID NO:22118-SEQ ID NO:23117的任何序列具有至少70％、至少75％、至少80％、至少85％、至少90％、至少92％、至少95％、至少97％、至少99％或100％序列同一性，其中所述至少一个突变驱动增加的乳腺组织嗜性。

[0368] B.ML规则

[0369] 对于本段中描述的以下一组规则，使用体内数据和两个ML模型(描述于实施例36)

确定工程化AAV5 VP1衣壳多肽的581至589区域中每个位置处的有利生物物理特性和有利氨基酸残基。本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽，与具有野生型AAV5 VP衣壳多肽的rAAV病毒体相比，所述组装的病毒体表现出增加的乳腺组织嗜性，其中工程化AAV5 VP衣壳多肽序列具有一个或多个突变，其中工程化AAV5 VP1衣壳多肽序列在从对应于SEQ ID NO:2中的581的位置至对应于SEQ ID NO:2中的589的位置的区域中具有所述一个或多个突变，并且其中Xaa1选自低表面可接近性的氨基酸(例如，Xaa1选自C)；或者其中Xaa1选自中等摩尔质量的氨基酸(例如，Xaa1选自C)；或者其中Xaa2选自高表面可接近性的氨基酸(例如，Xaa2选自D、N或Q)；或者其中Xaa2选自低亲疏水性的氨基酸(例如，Xaa2选自D、E、R、K、H、N或Q)；或者其中Xaa3选自高平均柔性的氨基酸(例如，Xaa3选自D、E、R、P、G或S)；或者其中Xaa3选自中等可突变性的氨基酸(例如，Xaa3选自R或H)；或者其中Xaa4选自高可突变性的氨基酸(例如，Xaa4选自M、I、Q或T)；或者其中Xaa4选自高溶解性的氨基酸(例如，Xaa4选自W、F、I或L)；或者其中Xaa4选自高表面可接近性的氨基酸(例如，Xaa4选自E、R或K)；或者其中Xaa5选自高溶解性的氨基酸(例如，Xaa5选自W、F、I或L)；或者其中Xaa5选自低可突变性的氨基酸(例如，Xaa5选自Y、F或L)；或者其中Xaa6选自高亲疏水性的氨基酸(例如，Xaa6选自V、I或L)；或者其中Xaa6选自中等摩尔质量的氨基酸(例如，Xaa6选自D、I、L或N)；或者其中Xaa8选自低表面可接近性的氨基酸(例如，Xaa8选自C)；或者其中Xaa8选自低可突变性的氨基酸(例如，Xaa8选自C、R或H)；或者其中Xaa9选自中等可突变性的氨基酸(例如，Xaa9选自R或H)；或它们的任何组合。

[0370] 在一些实施方案中，Xaa1选自低表面可接近性的氨基酸。在一些实施方案中，Xaa1选自C。在一些实施方案中，Xaa1选自中等摩尔质量的氨基酸。在一些实施方案中，Xaa1选自C。在一些实施方案中，Xaa2选自高表面可接近性的氨基酸。在一些实施方案中，Xaa2选自D、N或Q。在一些实施方案中，Xaa2选自低亲疏水性的氨基酸。在一些实施方案中，Xaa2选自D、E、R、K、H、N或Q。在一些实施方案中，Xaa3选自高平均柔性的氨基酸。在一些实施方案中，Xaa3选自D、E、R、P、G或S。在一些实施方案中，Xaa3选自中等可突变性的氨基酸。在一些实施方案中，Xaa3选自R或H。在一些实施方案中，Xaa4选自高可突变性的氨基酸。在一些实施方案中，Xaa4选自M、I、Q或T。在一些实施方案中，Xaa4选自高溶解性的氨基酸。在一些实施方案中，Xaa4选自W、F、I或L。在一些实施方案中，Xaa4选自高表面可接近性的氨基酸。在一些实施方案中，Xaa4选自E、R或K。在一些实施方案中，Xaa5选自高溶解性的氨基酸。在一些实施方案中，Xaa5选自W、F、I或L。在一些实施方案中，Xaa5选自低可突变性的氨基酸。在一些实施方案中，Xaa5选自Y、F或L。在一些实施方案中，Xaa6选自高亲疏水性的氨基酸。在一些实施方案中，Xaa6选自V、I或L。在一些实施方案中，Xaa6选自中等摩尔质量的氨基酸。在一些实施方案中，Xaa6选自D、I、L或N。在一些实施方案中，Xaa8选自低表面可接近性的氨基酸。在一些实施方案中，Xaa8选自C。在一些实施方案中，Xaa8选自低可突变性的氨基酸。在一些实施方案中，Xaa8选自C、R或H。在一些实施方案中，Xaa9选自中等可突变性的氨基酸。在一些实施方案中，Xaa9选自R或H。

[0371] 在一些实施方案中，本文提供了AAV5 VP衣壳多肽，所述多肽在从对应于AAV5 VP1的581的位置至对应于589的位置的区域中具有至少一个突变并且与选自SEQ ID NO:23118-SEQ ID NO:24117的任何序列具有至少70％、至少75％、至少80％、至少85％、至少90％、至少92％、至少95％、至少97％、至少99％或100％序列同一性，其中所述至少一个突

变驱动增加的乳腺组织嗜性。

[0372] C.富集的乳腺序列

[0373] 在一些实施方案中,本文提供了AAV5 VP衣壳多肽,所述多肽在从对应于AAV5 VP1的581的位置至对应于589的位置的区域中具有至少一个突变并且与选自SEQ ID NO:24118-SEQ ID NO:25117的任何序列具有至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%、至少99%或100%序列同一性,其中所述至少一个突变驱动增加的乳腺组织嗜性。

[0374] 6.7.13.体内选定的赋予增加的肺嗜性的突变VP多肽

[0375] 本公开提供了具有VP衣壳多肽的AAV5病毒体,所述VP衣壳多肽在具有与靶细胞(例如,目标靶肺组织中的靶肺细胞)相互作用的残基的区域中具有至少一个突变,其中与野生型VP衣壳多肽相比,所述至少一个突变赋予增加的肺组织嗜性。在一些实施方案中,本文提供了与SEQ ID NO:1具有至少80%序列同源性的AAV5 VP1衣壳多肽,其中所述AAV5 VP1衣壳多肽在从对应于SEQ ID NO:1的581的位置至对应于589的位置的区域中具有至少一个突变,并且其中所述至少一个突变驱动增加的肺嗜性。以下序列规则和序列也适用于对应于AAV5 VP1氨基酸残基581至589的AAV5 VP2(氨基酸残基445至453;VP2序列显示于SEQ ID NO:1115)和AAV5 VP3(氨基酸残基389至397;VP3序列显示于SEQ ID NO:1116)中的区域。因此,本公开涵盖在对应于581至589区域的AAV5 VP1氨基酸残基的VP2和VP3区域中具有一个或多个突变的AAV5 VP2衣壳多肽和AAV5VP3衣壳多肽,其中所述一个或多个突变与以下章节中的规则或顺序相称。

[0376] A.位置频率规则

[0377] 在本章节中,除非另外说明,否则分析相对于给定氨基酸残基出现在所有其他所收获组织(CNS(前脑皮层、枕叶皮层、颞叶皮层、丘脑、下丘脑、黑质、海马DG、海马CA1、海马CA3、小脑)、肝脏、骨骼肌、心脏、脾脏、淋巴结、骨髓、乳腺、皮肤、肾上腺、甲状腺、结肠、坐骨神经和脊髓组织)中鉴定的变体中指定位置的频率而言该给定氨基酸残基出现在肺中鉴定的变体中对应于SEQ ID NO:1(推及于SEQ ID NO:2)的位置581至位置589的所述指定位置的频率,以确定优先靶向肺组织的衣壳的一组序列规则。来自体内数据的位置频率规则的确定详细描述于实施例12中。

[0378] 本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽,与野生型AAV5 VP衣壳多肽相比,所述组装的病毒体表现出增加的肺组织嗜性,其中工程化变体AAV5 VP1衣壳多肽序列具有一个或多个突变,其中Xaa1选自A、E、K、M、Q、R、S或T,或Xaa1选自A、E或Q,或Xaa1为E;或Xaa2选自A、I、K、S、T或V,或Xaa2选自S、T或V,或Xaa2为T;或Xaa3选自A、E、K、M、Q、R、S、T或V,或Xaa3选自A、K、R或S,或Xaa3为R;或Xaa4选自M、P、R、S或T,或Xaa4选自P、Q或T,或Xaa4为Q;或Xaa5选自I、K、L、M、T、V或Y,或Xaa5选自L、M或Y,或Xaa5为L;或Xaa6选自D、G、H、M、N、R或S,或Xaa6选自H或N,或Xaa6为N;或Xaa7选自A、K、M、Q或R,或Xaa7选自A、K或R,或Xaa7为R;或Xaa8选自A、F、G、S、W或Y,或Xaa8选自A、F或G,或Xaa8为F;或Xaa9选自A、E、G、P、R或Y,或Xaa9选自G、P或R,或Xaa9为G。

[0379] 本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽,与野生型AAV5 VP衣壳多肽相比,所述组装的病毒体表现出增加的肺组织嗜性,其中工程化变体AAV5 VP衣壳多肽序列具有一个或多个突变,其中VP1多肽序列在从对应于SEQ ID NO:2中的581的位置

至对应于SEQ ID NO:2中的589的位置的区域中具有所述一个或多个突变,并且其中所述一个或多个突变选自以下规则。在一些实施方案中,Xaa1选自A、E、K、M、Q、R、S或T。在一些实施方案中,Xaa1选自A、E或Q。在一些实施方案中,Xaa1为E。在一些实施方案中,Xaa2选自A、I、K、S、T或V。在一些实施方案中,Xaa2选自S、T或V。在一些实施方案中,Xaa2为T。在一些实施方案中,Xaa3选自A、E、K、M、Q、R、S、T或V。在一些实施方案中,Xaa3选自A、K、R或S。在一些实施方案中,Xaa3为R。在一些实施方案中,Xaa4选自M、P、R、S或T。在一些实施方案中,Xaa4选自P、Q或T。在一些实施方案中,Xaa4为Q。在一些实施方案中,Xaa5选自I、K、L、M、T、V或Y。在一些实施方案中,Xaa5选自L、M或Y。在一些实施方案中,Xaa5为L。在一些实施方案中,Xaa6选自D、G、H、M、N、R或S。在一些实施方案中,Xaa6选自H或N。在一些实施方案中,Xaa6为N。在一些实施方案中,Xaa7选自A、K、M、Q或R。在一些实施方案中,Xaa7选自A、K或R。在一些实施方案中,Xaa7为R。在一些实施方案中,Xaa8选自A、F、G、S、W或Y。在一些实施方案中,Xaa8选自A、F或G。在一些实施方案中,Xaa8为F。在一些实施方案中,Xaa9选自A、E、G、P、R或Y。在一些实施方案中,Xaa9选自G、P或R。在一些实施方案中,Xaa9为G。

[0380] 本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽,与野生型AAV5 VP衣壳多肽相比,所述组装的病毒体表现出增加的肺组织嗜性,其中工程化变体AAV5 VP衣壳多肽序列具有一个或多个突变,其中VP1多肽序列在从对应于SEQ ID NO:2中的581的位置至对应于SEQ ID NO:2中的589的位置的区域中具有所述一个或多个突变,并且其中所述一个或多个突变选自以下规则:Xaa1选自A、E、K、M、Q、R、S或T,Xaa2选自A、I、K、S、T或V,Xaa3选自A、E、K、M、Q、R、S、T或V,Xaa4选自M、P、R、S或T,Xaa5选自I、K、L、M、T、V或Y,Xaa6选自D、G、H、M、N、R或S,Xaa7选自A、K、M、Q或R,Xaa8选自A、F、G、S、W或Y,并且Xaa9选自A、E、G、P、R或Y。

[0381] 本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽,与野生型AAV5 VP衣壳多肽相比,所述组装的病毒体表现出增加的肺组织嗜性,其中工程化变体AAV5 VP衣壳多肽序列具有一个或多个突变,其中VP1多肽序列在从对应于SEQ ID NO:2中的581的位置至对应于SEQ ID NO:2中的589的位置的区域中具有所述一个或多个突变,并且其中所述一个或多个突变如下:Xaa1选自A、E、K、M、Q、R、S或T,Xaa2选自A、I、K、S、T或V,Xaa3选自A、E、K、M、Q、R、S、T或V,Xaa4选自M、P、R、S或T,Xaa5选自I、K、L、M、T、V或Y,Xaa6选自D、G、H、M、N、R或S,Xaa7选自A、K、M、Q或R,Xaa8选自A、F、G、S、W或Y,Xaa9选自A、E、G、P、R或Y,或它们的任何组合。

[0382] 在一些实施方案中,本文提供了AAV5 VP衣壳多肽,所述多肽在从对应于AAV5 VP1的581的位置至对应于589的位置的区域中具有至少一个突变并且与选自SEQ ID NO:16118-SEQ ID NO:17117的任何序列具有至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%、至少99%或100%序列同一性,其中所述至少一个突变驱动增加的肺组织嗜性。

[0383] B.ML规则

[0384] 对于本段中描述的以下一组规则,使用体内数据和两个ML模型(描述于实施例37)确定工程化AAV5 VP1衣壳多肽的581至589区域中每个位置处的有利生物物理特性和有利氨基酸残基。本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽,与具有野生型AAV5 VP衣壳多肽的rAAV病毒体相比,所述组装的病毒体表现出增加的肺组织嗜性,其中工程化AAV5 VP衣壳多肽序列具有一个或多个突变,其中工程化AAV5 VP1衣壳多肽序列在从

对应于SEQ ID NO:2中的581的位置至对应于SEQ ID NO:2中的589的位置的区域中具有所述一个或多个突变,并且其中Xaa1选自高可突变性的氨基酸(例如,Xaa1选自D、E、M、A、I、Q或T);或者其中Xaa2选自高摩尔质量的氨基酸(例如,Xaa2选自F);或者其中Xaa2选自低可突变性的氨基酸(例如,Xaa2选自Y、F或L);或者其中Xaa3选自低可突变性的氨基酸(例如,Xaa3选自K、V、P或H);或者其中Xaa3选自低亲疏水性的氨基酸(例如,Xaa3选自K或R);或者其中Xaa4选自低可突变性的氨基酸(例如,Xaa4选自K或P);或者其中Xaa4选自高平均柔性的氨基酸(例如,Xaa4选自D、E、P或S);或者其中Xaa5选自低平均柔性的氨基酸(例如,Xaa5选自W、M或F);或者其中Xaa5选自高溶解性的氨基酸(例如,Xaa5选自W、F、I或L);或者其中Xaa6选自中等可突变性的氨基酸(例如,Xaa6选自R或H);或者其中Xaa6选自高表面可接近性的氨基酸(例如,Xaa6选自T);或者其中Xaa7选自低可突变性的氨基酸(例如,Xaa7选自C);或者其中Xaa7选自高溶解性的氨基酸(例如,Xaa7选自W、V、M、F、I或L);或者其中Xaa8选自高可突变性的氨基酸(例如,Xaa8选自D、E、M、A、I、Q或T);或者其中Xaa8选自低亲疏水性的氨基酸(例如,Xaa8选自R或K);或者其中Xaa9选自高平均柔性的氨基酸(例如,Xaa9选自R或G);或它们的任何组合。

[0385] 在一些实施方案中,Xaa1选自高可突变性的氨基酸。在一些实施方案中,Xaa1选自D、E、M、A、I、Q或T。在一些实施方案中,Xaa2选自高摩尔质量的氨基酸。在一些实施方案中,Xaa2选自F。在一些实施方案中,Xaa2选自低可突变性的氨基酸。在一些实施方案中,Xaa2选自Y、F或L。在一些实施方案中,Xaa3选自低可突变性的氨基酸。在一些实施方案中,Xaa3选自K、V、P或H。在一些实施方案中,Xaa3选自低亲疏水性的氨基酸。在一些实施方案中,Xaa3选自K或R。在一些实施方案中,Xaa4选自低可突变性的氨基酸。在一些实施方案中,Xaa4选自K或P。在一些实施方案中,Xaa4选自高平均柔性的氨基酸。在一些实施方案中,Xaa4选自D、E、P或S。在一些实施方案中,Xaa5选自低平均柔性的氨基酸。在一些实施方案中,Xaa5选自W、M或F。在一些实施方案中,Xaa5选自高溶解性的氨基酸。在一些实施方案中,Xaa5选自W、F、I或L。在一些实施方案中,Xaa6选自中等可突变性的氨基酸。在一些实施方案中,Xaa6选自R或H。在一些实施方案中,Xaa6选自高表面可接近性的氨基酸。在一些实施方案中,Xaa6选自T。在一些实施方案中,Xaa7选自低可突变性的氨基酸。在一些实施方案中,Xaa7选自C。在一些实施方案中,Xaa7选自高溶解性的氨基酸。在一些实施方案中,Xaa7选自W、V、M、F、I或L。在一些实施方案中,Xaa8选自高可突变性的氨基酸。在一些实施方案中,Xaa8选自D、E、M、A、I、Q或T。在一些实施方案中,Xaa8选自低亲疏水性的氨基酸。在一些实施方案中,Xaa8选自R或K。在一些实施方案中,Xaa9选自高平均柔性的氨基酸。在一些实施方案中,Xaa9选自R或G。

[0386] 在一些实施方案中,本文提供了AAV5 VP衣壳多肽,所述多肽在从对应于AAV5 VP1的581的位置至对应于589的位置的区域中具有至少一个突变并且与选自SEQ ID NO:17118-SEQ ID NO:18117的任何序列具有至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%、至少99%或100%序列同一性,其中所述至少一个突变驱动增加的肺组织嗜性。

[0387] C.富集的肺序列

[0388] 在一些实施方案中,本文提供了AAV5 VP衣壳多肽,所述多肽在从对应于AAV5 VP1的581的位置至对应于589的位置的区域中具有至少一个突变并且与选自SEQ ID NO:

18118-SEQ ID NO:19117的任何序列具有至少70％、至少75％、至少80％、至少85％、至少90％、至少92％、至少95％、至少97％、至少99％或100％序列同一性，其中所述至少一个突变驱动增加的肺组织嗜性。

[0389] 6.7.14.体内选定的赋予增加的心脏嗜性的突变VP多肽

[0390] 本公开提供了具有VP衣壳多肽的AAV5病毒体，所述VP衣壳多肽在具有与靶细胞(例如，目标靶心脏组织中的靶心脏细胞)相互作用的残基的区域中具有至少一个突变，其中与野生型VP衣壳多肽相比，所述至少一个突变赋予增加的心脏组织嗜性。在一些实施方案中，本文提供了与SEQ ID NO:1具有至少80％序列同源性的AAV5 VP1衣壳多肽，其中所述AAV5 VP1衣壳多肽在从对应于SEQ ID NO:1的581的位置至对应于589的位置的区域中具有至少一个突变，并且其中所述至少一个突变驱动增加的心脏嗜性。以下序列规则和序列也适用于对应于AAV5 VP1氨基酸残基581至589的AAV5 VP2(氨基酸残基445至453;VP2序列显示于SEQ ID NO:1115)和AAV5 VP3(氨基酸残基389至397;VP3序列显示于SEQ ID NO:1116)中的区域。因此，本公开涵盖在对应于581至589区域的AAV5 VP1氨基酸残基的VP2和VP3区域中具有一个或多个突变的AAV5 VP2衣壳多肽和AAV5 VP3衣壳多肽，其中所述一个或多个突变与以下章节中的规则或顺序相称。在一些实施方案中，“心脏”和“心脏的”在本文中可以互换使用。

[0391] A.位置频率规则

[0392] 在本章节中，除非另外说明，否则分析相对于给定氨基酸残基出现在所有其他所收获组织(CNS(前脑皮层、枕叶皮层、颞叶皮层、丘脑、下丘脑、黑质、海马DG、海马CA1、海马CA3、小脑)、肝脏、骨骼肌、肺、脾脏、淋巴结、骨髓、乳腺、皮肤、肾上腺、甲状腺、结肠、坐骨神经和脊髓组织)中鉴定的变体中指定位置的频率而言该给定氨基酸残基出现在心脏中鉴定的变体中对应于SEQ ID NO:1(推及于SEQ ID NO:2)的位置581至位置589的所述指定位置的频率，以确定优先靶向心脏组织的衣壳的一组序列规则。来自体内数据的位置频率规则的确定详细描述于实施例13中。

[0393] 本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽，与野生型AAV5 VP衣壳多肽相比，所述组装的病毒体表现出增加的心脏组织嗜性，其中工程化变体AAV5 VP衣壳多肽序列具有一个或多个突变，其中Xaa1选自I、K、L、M、T或V，或Xaa1选自K或L，或Xaa1为K;或Xaa2选自A、C、G、I、K或S，或Xaa2选自A、C或S，或Xaa2为A;或Xaa3选自A、D、E、G、K、M或V，或Xaa3选自E或V，或Xaa3为E;或Xaa4选自F、H、R、T、W或Y，或Xaa4选自F、R或T，或Xaa4为R;或Xaa5选自F、L、M或R，或Xaa5为L;或Xaa6选自A、H、N、W或Y，或Xaa6选自H、N或Y，或Xaa6为H;或Xaa7选自A、C、E、F、K或T，或Xaa7选自C、F或T，或Xaa7为F;或Xaa8选自A、C、M、S或T，或Xaa8选自C、M或S，或Xaa8为C;或Xaa9选自A、D、G或P，或Xaa9选自A或G，或Xaa9为A。

[0394] 本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽，与野生型AAV5 VP衣壳多肽相比，所述组装的病毒体表现出增加的心脏组织嗜性，其中工程化变体AAV5 VP衣壳多肽序列具有一个或多个突变，其中VP1多肽序列在从对应于SEQ ID NO:2中的581的位置至对应于SEQ ID NO:2中的589的位置的区域中具有所述一个或多个突变，并且其中所述一个或多个突变选自以下规则。在一些实施方案中，Xaa1选自I、K、L、M、T或V。在一些实施方案中，Xaa1选自K或L。在一些实施方案中，Xaa1为K。在一些实施方案中，Xaa2选自A、C、G、I、K或S。在一些实施方案中，Xaa2选自A、C或S。在一些实施方案中，Xaa2为A。在一些实施方案

中,Xaa3选自A、D、E、G、K、M或V。在一些实施方案中,Xaa3选自E或V。在一些实施方案中,Xaa3为E。在一些实施方案中,Xaa4选自F、H、R、T、W或Y。在一些实施方案中,Xaa4选自F、R或T。在一些实施方案中,Xaa4为R。在一些实施方案中,Xaa5选自F、L、M或R。在一些实施方案中,Xaa5为L。在一些实施方案中,Xaa6选自A、H、N、W或Y。在一些实施方案中,Xaa6选自H、N或Y。在一些实施方案中,Xaa6为H。在一些实施方案中,Xaa7选自A、C、E、F、K或T。在一些实施方案中,Xaa7选自C、F或T。在一些实施方案中,Xaa7为F。在一些实施方案中,Xaa8选自A、C、M、S或T。在一些实施方案中,Xaa8选自C、M或S。在一些实施方案中,Xaa8为C。在一些实施方案中,Xaa9选自A、D、G或P。在一些实施方案中,Xaa9选自A或G。在一些实施方案中,Xaa9为A。

[0395] 本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽,与野生型AAV5 VP衣壳多肽相比,所述组装的病毒体表现出增加的心脏组织嗜性,其中工程化变体AAV5 VP衣壳多肽序列具有一个或多个突变,其中VP1多肽序列在从对应于SEQ ID NO:2中的581的位置至对应于SEQ ID NO:2中的589的位置的区域中具有所述一个或多个突变,并且其中所述一个或多个突变选自以下规则:Xaa1选自I、K、L、M、T或V,Xaa2选自A、C、G、I、K或S,Xaa3选自A、D、E、G、K、M或V,Xaa4选自F、H、R、T、W或Y,Xaa5选自F、L、M或R,Xaa6选自A、H、N、W或Y,Xaa7选自A、C、E、F、K或T,Xaa8选自A、C、M、S或T,并且Xaa9选自A、D、G或P。

[0396] 本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽,与野生型AAV5 VP衣壳多肽相比,所述组装的病毒体表现出增加的心脏组织嗜性,其中工程化变体AAV5 VP衣壳多肽序列具有一个或多个突变,其中VP1多肽序列在从对应于SEQ ID NO:2中的581的位置至对应于SEQ ID NO:2中的589的位置的区域中具有所述一个或多个突变,并且其中所述一个或多个突变如下:Xaa1选自I、K、L、M、T或V,Xaa2选自A、C、G、I、K或S,Xaa3选自A、D、E、G、K、M或V,Xaa4选自F、H、R、T、W或Y,Xaa5选自F、L、M或R,Xaa6选自A、H、N、W或Y,Xaa7选自A、C、E、F、K或T,Xaa8选自A、C、M、S或T,Xaa9选自A、D、G或P,或它们的任何组合。

[0397] 在一些实施方案中,本文提供了AAV5 VP衣壳多肽,所述多肽在从对应于AAV5 VP1的581的位置至对应于589的位置的区域中具有至少一个突变并且与选自SEQ ID NO:13118-SEQ ID NO:14117的任何序列具有至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%、至少99%或100%序列同一性,其中所述至少一个突变驱动增加的心脏组织嗜性。

[0398] B.ML规则

[0399] 对于本段中描述的以下一组规则,使用体内数据和两个ML模型(描述于实施例34)确定工程化AAV5 VP1衣壳多肽的581至589区域中每个位置处的有利生物物理特性和有利氨基酸残基。本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽,与具有野生型AAV5 VP衣壳多肽的rAAV病毒体相比,所述组装的病毒体表现出增加的心脏组织嗜性,其中工程化AAV5 VP衣壳多肽序列具有一个或多个突变,其中工程化AAV5 VP1衣壳多肽序列在从对应于SEQ ID NO:2中的581的位置至对应于SEQ ID NO:2中的589的位置的区域中具有所述一个或多个突变,并且其中Xaa1选自低溶解性的氨基酸(例如,Xaa1选自N或E);或者其中Xaa1选自低亲疏水性的氨基酸(例如,Xaa1选自H、N、Q、P、Y、D或E);或者其中Xaa1选自高可突变性的氨基酸(例如,Xaa1选自A或E);或者其中Xaa2选自高亲疏水性的氨基酸(例如,Xaa2选自V或I);或者其中Xaa2选自中等可突变性的氨基酸(例如,Xaa2选自V);或者其中Xaa2选自中等体积的氨基酸(例如,Xaa2选自V、E或Q);或者其中Xaa2选自高溶解性的氨基

酸(例如,Xaa2选自V或M);或者其中Xaa3选自低溶解性的氨基酸(例如,Xaa3选自R或Q);或者其中Xaa4选自低表面可接近性的氨基酸(例如,Xaa4选自C);或者其中Xaa4选自高溶解性的氨基酸(例如,Xaa4选自C);或者其中Xaa4选自低电荷的氨基酸(例如,Xaa4选自D、E、Y、W、V、P、M、A、G、F、I、L、N、Q、S、T或C);或者其中Xaa4选自高亲疏水性的氨基酸(例如,Xaa4选自C);或者其中Xaa5选自高表面可接近性的氨基酸(例如,Xaa5选自D、E、R、K、N或Q);或者其中Xaa5选自低溶解性的氨基酸(例如,Xaa5选自D);或者其中Xaa6选自低可突变性的氨基酸(例如,Xaa6选自C);或者其中Xaa6选自低溶解性的氨基酸(例如,Xaa6选自D);或者其中Xaa8选自高表面可接近性的氨基酸(例如,Xaa8选自D或N);或者其中Xaa8选自高平均柔性的氨基酸(例如,Xaa8选自D、R、P、G或S);或者其中Xaa9选自中等摩尔质量的氨基酸(例如,Xaa9选自N、D、L或I);或它们的任何组合。

[0400] 在一些实施方案中,Xaa1选自低溶解性的氨基酸。在一些实施方案中,Xaa1选自N或E。在一些实施方案中,Xaa1选自低亲疏水性的氨基酸。在一些实施方案中,Xaa1选自H、N、Q、P、Y、D或E。在一些实施方案中,Xaa1选自高可突变性的氨基酸。在一些实施方案中,Xaa1选自A或E。在一些实施方案中,Xaa2选自高亲疏水性的氨基酸。在一些实施方案中,Xaa2选自V或I。在一些实施方案中,Xaa2选自中等可突变性的氨基酸。在一些实施方案中,Xaa2选自V。在一些实施方案中,Xaa2选自中等体积的氨基酸。在一些实施方案中,Xaa2选自V、E或Q。在一些实施方案中,Xaa2选自高溶解性的氨基酸。在一些实施方案中,Xaa2选自V或M。在一些实施方案中,Xaa3选自低溶解性的氨基酸。在一些实施方案中,Xaa3选自R或Q。在一些实施方案中,Xaa4选自低表面可接近性的氨基酸。在一些实施方案中,Xaa4选自C。在一些实施方案中,Xaa4选自高溶解性的氨基酸。在一些实施方案中,Xaa4选自C。在一些实施方案中,Xaa4选自低电荷的氨基酸。在一些实施方案中,Xaa4选自D、E、Y、W、V、P、M、A、G、F、I、L、N、Q、S、T或C。在一些实施方案中,Xaa4选自高亲疏水性的氨基酸。在一些实施方案中,Xaa4选自C。在一些实施方案中,Xaa5选自高表面可接近性的氨基酸。在一些实施方案中,Xaa5选自D、E、R、K、N或Q。在一些实施方案中,Xaa5选自低溶解性的氨基酸。在一些实施方案中,Xaa5选自D。在一些实施方案中,Xaa6选自低可突变性的氨基酸。在一些实施方案中,Xaa6选自C。在一些实施方案中,Xaa6选自低溶解性的氨基酸。在一些实施方案中,Xaa6选自D。在一些实施方案中,Xaa8选自高表面可接近性的氨基酸。在一些实施方案中,Xaa8选自D或N。在一些实施方案中,Xaa8选自高平均柔性的氨基酸。在一些实施方案中,Xaa8选自D、R、P、G或S。在一些实施方案中,Xaa9选自中等摩尔质量的氨基酸。在一些实施方案中,Xaa9选自N、D、L或I。

[0401] 在一些实施方案中,本文提供了AAV5 VP衣壳多肽,所述多肽在从对应于AAV5 VP1的581的位置至对应于589的位置的区域中具有至少一个突变并且与选自SEQ ID NO:14118-SEQ ID NO:15117的任何序列具有至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%、至少99%或100%序列同一性,其中所述至少一个突变驱动增加的心脏组织嗜性。

[0402] C.富集的心脏序列

[0403] 在一些实施方案中,本文提供了AAV5 VP衣壳多肽,所述多肽在从对应于AAV5 VP1的581的位置至对应于589的位置的区域中具有至少一个突变并且与选自SEQ ID NO:15118-SEQ ID NO:16117的任何序列具有至少70%、至少75%、至少80%、至少85%、至少

90％、至少92％、至少95％、至少97％、至少99％或100％序列同一性，其中所述至少一个突变驱动增加的心脏组织嗜性。

[0404] 6.7.15.体内选定的赋予增加的结肠嗜性的突变VP多肽

[0405] 本公开提供了具有VP衣壳多肽的AAV5病毒体，所述VP衣壳多肽在具有与靶细胞(例如，目标靶结肠组织中的靶结肠细胞)相互作用的残基的区域中具有至少一个突变，其中与野生型VP衣壳多肽相比，所述至少一个突变赋予增加的结肠组织嗜性。在一些实施方案中，本文提供了与SEQ ID NO:1具有至少80％序列同源性的AAV5 VP1衣壳多肽，其中所述AAV5 VP1衣壳多肽在从对应于SEQ ID NO:1的581的位置至对应于589的位置的区域中具有至少一个突变，并且其中所述至少一个突变驱动增加的结肠嗜性。以下序列规则和序列也适用于对应于AAV5 VP1氨基酸残基581至589的AAV5 VP2(氨基酸残基445至453;VP2序列显示于SEQ ID NO:1115)和AAV5 VP3(氨基酸残基389至397;VP3序列显示于SEQ ID NO:1116)中的区域。因此，本公开涵盖在对应于581至589区域的AAV5 VP1氨基酸残基的VP2和VP3区域中具有一个或多个突变的AAV5 VP2衣壳多肽和AAV5 VP3衣壳多肽，其中所述一个或多个突变与以下章节中的规则或顺序相称。

[0406] A.位置频率规则

[0407] 在本章节中，除非另外说明，否则分析相对于给定氨基酸残基出现在所有其他所收获组织(CNS(前脑皮层、枕叶皮层、颞叶皮层、丘脑、下丘脑、黑质、海马DG、海马CA1、海马CA3、小脑)、肝脏、骨骼肌、心脏、肺、脾脏、淋巴结、骨髓、乳腺、皮肤、肾上腺、甲状腺、坐骨神经和脊髓组织)中鉴定的变体中指定位置的频率而言该给定氨基酸残基出现在结肠中鉴定的变体中对应于SEQ ID NO:1(推及于SEQ ID NO:2)的位置581至位置589的所述指定位置的频率，以确定优先靶向结肠组织的衣壳的一组序列规则。来自体内数据的位置频率规则的确定详细描述于实施例14中。

[0408] 本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽，与野生型AAV5 VP衣壳多肽相比，所述组装的病毒体表现出增加的结肠组织嗜性，其中工程化变体AAV5 VP1衣壳多肽序列具有一个或多个突变，其中Xaa1选自C、F、H、N、P、W或Y，或Xaa1选自F、P或W，或Xaa1为P;或Xaa2选自D、E、F、L或P，或Xaa2选自D、E、L或P，或Xaa2为P;或Xaa3选自C、F、H、I、L、P或Y，或Xaa3选自C、H或P，或Xaa3为P;或Xaa4选自C、D、E、N或P，或Xaa4选自C、D或E，或Xaa4为C;或Xaa5选自D、E、G、P或W，或Xaa5选自G、P或W，或Xaa5为P;或Xaa6选自C、K、R或V，或Xaa6选自K或R，或Xaa6为R;或Xaa7选自D、M、P或V，或Xaa7为P;或Xaa8选自D、I、K、L、P、R或V，或Xaa8选自K、P或R，或Xaa8为P;或Xaa9选自C、H、I、K、L、M或W，或Xaa9选自I、L或M，或Xaa9为I。

[0409] 本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽，与野生型AAV5 VP衣壳多肽相比，所述组装的病毒体表现出增加的结肠组织嗜性，其中工程化变体AAV5 VP衣壳多肽序列具有一个或多个突变，其中VP1多肽序列在从对应于SEQ ID NO:2中的581的位置至对应于SEQ ID NO:2中的589的位置的区域中具有所述一个或多个突变，并且其中所述一个或多个突变选自以下规则。在一些实施方案中，Xaa1选自C、F、H、N、P、W或Y。在一些实施方案中，Xaa1选自F、P或W。在一些实施方案中，Xaa1为P。在一些实施方案中，Xaa2选自D、E、F、L或P。在一些实施方案中，Xaa2选自D、E、L或P。在一些实施方案中，Xaa2为P。在一些实施方案中，Xaa3选自C、F、H、I、L、P或Y。在一些实施方案中，Xaa3选自C、H或P。在一些实施方案

中,Xaa3为P。在一些实施方案中,Xaa4选自C、D、E、N或P。在一些实施方案中,Xaa4选自C、D或E。在一些实施方案中,Xaa4为C。在一些实施方案中,Xaa5选自D、E、G、P或W。在一些实施方案中,Xaa5选自G、P或W。在一些实施方案中,Xaa5为P。在一些实施方案中,Xaa6选自C、K、R或V。在一些实施方案中,Xaa6选自K或R。在一些实施方案中,Xaa6为R。在一些实施方案中,Xaa7选自D、M、P或V。在一些实施方案中,Xaa7为P。在一些实施方案中,Xaa8选自D、I、K、L、P、R或V。在一些实施方案中,Xaa8选自K、P或R。在一些实施方案中,Xaa8为P。在一些实施方案中,Xaa9选自C、H、I、K、L、M或W。在一些实施方案中,Xaa9选自I、L或M。在一些实施方案中,Xaa9为I。

[0410] 本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽,与野生型AAV5 VP衣壳多肽相比,所述组装的病毒体表现出增加的结肠组织嗜性,其中工程化变体AAV5 VP衣壳多肽序列具有一个或多个突变,其中VP1多肽序列在从对应于SEQ ID NO:2中的581的位置至对应于SEQ ID NO:2中的589的位置的区域中具有所述一个或多个突变,并且其中所述一个或多个突变选自以下规则:Xaa1选自C、F、H、N、P、W或Y,Xaa2选自D、E、F、L或P,Xaa3选自C、F、H、I、L、P或Y,Xaa4选自C、D、E、N或P,Xaa5选自D、E、G、P或W,Xaa6选自C、K、R或V,Xaa7选自D、M、P或V,Xaa8选自D、I、K、L、P、R或V,并且Xaa9选自C、H、I、K、L、M或W。

[0411] 本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽,与野生型AAV5 VP衣壳多肽相比,所述组装的病毒体表现出增加的结肠组织嗜性,其中工程化变体AAV5 VP衣壳多肽序列具有一个或多个突变,其中VP1多肽序列在从对应于SEQ ID NO:2中的581的位置至对应于SEQ ID NO:2中的589的位置的区域中具有所述一个或多个突变,并且其中所述一个或多个突变如下:Xaa1选自C、F、H、N、P、W或Y,Xaa2选自D、E、F、L或P,Xaa3选自C、F、H、I、L、P或Y,Xaa4选自C、D、E、N或P,Xaa5选自D、E、G、P或W,Xaa6选自C、K、R或V,Xaa7选自D、M、P或V,Xaa8选自D、I、K、L、P、R或V,Xaa9选自C、H、I、K、L、M或W,或它们的任何组合。

[0412] 在一些实施方案中,本文提供了AAV5 VP衣壳多肽,所述多肽在从对应于AAV5 VP1的581的位置至对应于589的位置的区域中具有至少一个突变并且与选自SEQ ID NO:10118-SEQ ID NO:11117的任何序列具有至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%、至少99%或100%序列同一性,其中所述至少一个突变驱动增加的结肠组织嗜性。

[0413] B.ML规则

[0414] 对于本段中描述的以下一组规则,使用体内数据和两个ML模型(描述于实施例33)确定工程化AAV5 VP1衣壳多肽的581至589区域中每个位置处的有利生物物理特性和有利氨基酸残基。本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽,与具有野生型AAV5 VP衣壳多肽的rAAV病毒体相比,所述组装的病毒体表现出增加的结肠组织嗜性,其中工程化AAV5 VP衣壳多肽序列具有一个或多个突变,其中工程化AAV5 VP1衣壳多肽序列在从对应于SEQ ID NO:2中的581的位置至对应于SEQ ID NO:2中的589的位置的区域中具有所述一个或多个突变,并且其中Xaa1选自高摩尔质量的氨基酸(例如,Xaa1选自Y或W);或者其中Xaa1选自高溶解性的氨基酸(例如,Xaa1选自W、F、I或L);或者其中Xaa2选自低溶解性的氨基酸(例如,Xaa2选自D);或者其中Xaa2选自低可突变性的氨基酸(例如,Xaa2选自P或K);或者其中Xaa2选自中等摩尔质量的氨基酸(例如,Xaa2选自D、E、N、K、M、Q、I或L);或者其中Xaa2选自低亲疏水性的氨基酸(例如,Xaa2选自D、E、R、K、H、N或Q);或者其中Xaa3选自低

可突变性的氨基酸(例如,Xaa3选自K、V、P或C);或者其中Xaa3选自高溶解性的氨基酸(例如,Xaa3选自W、F、I或L);或者其中Xaa5选自高平均柔性的氨基酸(例如,Xaa5选自S、P、G、R、E或D);或者其中Xaa5选自高表面可接近性的氨基酸(例如,Xaa5选自D或N);或者其中Xaa6选自低亲疏水性的氨基酸(例如,Xaa6选自R);或者其中Xaa6选自低可突变性的氨基酸(例如,Xaa6选自Y、R、F或L);或者其中Xaa6选自低溶解性的氨基酸(例如,Xaa6选自R或Q);或者其中Xaa6选自高表面可接近性的氨基酸(例如,Xaa6选自E、R或K);或者其中Xaa6选自高平均柔性的氨基酸(例如,Xaa6选自G或R);或者其中Xaa8选自低溶解性的氨基酸(例如,Xaa8选自D);或它们的任何组合。

[0415] 在一些实施方案中,Xaa1选自高摩尔质量的氨基酸。在一些实施方案中,Xaa1选自Y或W。在一些实施方案中,Xaa1选自高溶解性的氨基酸。在一些实施方案中,Xaa1选自W、F、I或L。在一些实施方案中,Xaa2选自低溶解性的氨基酸。在一些实施方案中,Xaa2选自D。在一些实施方案中,Xaa2选自低可突变性的氨基酸。在一些实施方案中,Xaa2选自P、K。在一些实施方案中,Xaa2选自中等摩尔质量的氨基酸。在一些实施方案中,Xaa2选自D、E、N、K、M、Q、I或L。在一些实施方案中,Xaa2选自低亲疏水性的氨基酸。在一些实施方案中,Xaa2选自D、E、R、K、H、N或Q。在一些实施方案中,Xaa3选自低可突变性的氨基酸。在一些实施方案中,Xaa3选自K、V、P或C。在一些实施方案中,Xaa3选自高溶解性的氨基酸。在一些实施方案中,Xaa3选自W、F、I或L。在一些实施方案中,Xaa5选自高平均柔性的氨基酸。在一些实施方案中,Xaa5选自S、P、G、R、E或D。在一些实施方案中,Xaa5选自高表面可接近性的氨基酸。在一些实施方案中,Xaa5选自D或N。在一些实施方案中,Xaa6选自低亲疏水性的氨基酸。在一些实施方案中,Xaa6选自R。在一些实施方案中,Xaa6选自低可突变性的氨基酸。在一些实施方案中,Xaa6选自Y、R、F或L。在一些实施方案中,Xaa6选自低溶解性的氨基酸。在一些实施方案中,Xaa6选自R或Q。在一些实施方案中,Xaa6选自高表面可接近性的氨基酸。在一些实施方案中,Xaa6选自E、R或K。在一些实施方案中,Xaa6选自高平均柔性的氨基酸。在一些实施方案中,Xaa6选自G或R。在一些实施方案中,Xaa8选自低溶解性的氨基酸。在一些实施方案中,Xaa8选自D。

[0416] 在一些实施方案中,本文提供了AAV5 VP衣壳多肽,所述多肽在从对应于AAV5 VP1的581的位置至对应于589的位置的区域中具有至少一个突变并且与选自SEQ ID NO:11118-SEQ ID NO:12117的任何序列具有至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%、至少99%或100%序列同一性,其中所述至少一个突变驱动增加的结肠组织嗜性。

[0417] C.富集的结肠序列

[0418] 在一些实施方案中,本文提供了AAV5 VP衣壳多肽,所述多肽在从对应于AAV5 VP1的581的位置至对应于589的位置的区域中具有至少一个突变并且与选自SEQ ID NO:12118-SEQ ID NO:13117的任何序列具有至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%、至少99%或100%序列同一性,其中所述至少一个突变驱动增加的结肠组织嗜性。

[0419] 6.7.16.体内选定的赋予增加的甲状腺嗜性的突变VP多肽

[0420] 本公开提供了具有VP衣壳多肽的AAV5病毒体,所述VP衣壳多肽在具有与靶细胞(例如,目标靶甲状腺组织中的靶甲状腺细胞)相互作用的残基的区域中具有至少一个突

变,其中与野生型VP衣壳多肽相比,所述至少一个突变赋予增加的甲状腺组织嗜性。在一些实施方案中,本文提供了与SEQ ID NO:1具有至少80%序列同源性的AAV5 VP1衣壳多肽,其中所述AAV5 VP1衣壳多肽在从对应于SEQ ID NO:1的581的位置至对应于589的位置的区域中具有至少一个突变,并且其中所述至少一个突变驱动增加的甲状腺嗜性。以下序列规则和序列也适用于对应于AAV5 VP1氨基酸残基581至589的AAV5 VP2(氨基酸残基445至453;VP2序列显示于SEQ ID NO:1115)和AAV5 VP3(氨基酸残基389至397;VP3序列显示于SEQ ID NO:1116)中的区域。因此,本公开涵盖在对应于581至589区域的AAV5 VP1氨基酸残基的VP2和VP3区域中具有一个或多个突变的AAV5 VP2衣壳多肽和AAV5 VP3衣壳多肽,其中所述一个或多个突变与以下章节中的规则或顺序相称。

[0421] A.位置频率规则

[0422] 在本章节中,除非另外说明,否则分析相对于给定氨基酸残基出现在所有其他所收获组织(CNS(前脑皮层、枕叶皮层、颞叶皮层、丘脑、下丘脑、黑质、海马DG、海马CA1、海马CA3、小脑)、肝脏、骨骼肌、心脏、肺、脾脏、淋巴结、骨髓、乳腺、皮肤、肾上腺、结肠、坐骨神经和脊髓组织)中鉴定的变体中指定位置的频率而言该给定氨基酸残基出现在甲状腺中鉴定的变体中对应于SEQ ID NO:1(推及于SEQ ID NO:2)的位置581至位置589的所述指定位置的频率,以确定优先靶向甲状腺组织的衣壳的一组序列规则。来自体内数据的位置频率规则的确定详细描述于实施例15中。

[0423] 本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽,与野生型AAV5 VP衣壳多肽相比,所述组装的病毒体表现出增加的甲状腺组织嗜性,其中工程化变体AAV5 VP1衣壳多肽序列具有一个或多个突变,其中Xaa1选自A、K、M、N、Q或R,或Xaa1选自K、N或Q,或Xaa1为K;或Xaa2选自A、F、K、L、M、T、V或W,或Xaa2选自F、V或W,或Xaa2为W;或Xaa3选自A、I、K、R、S、T、V或W,或Xaa3选自A、R或T,或Xaa3为R;或Xaa4选自A、D、E、I、P或V,或Xaa4选自A、E或I,或Xaa4为A;或Xaa5选自F、I、M、Q、V或Y,或Xaa5为M、V、Y,或Xaa5为M;或Xaa6选自H、M、N或Y,或Xaa6为N;或Xaa7选自H、I、N、Q、S或W,或Xaa7选自H、I或N,或Xaa7为H;或Xaa8选自A、D、F、Q、S或Y,或Xaa8选自A、F或S,或Xaa8为F;或Xaa9选自A、Q、S或Y,或Xaa9选自A或S,或Xaa9为A。

[0424] 本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽,与野生型AAV5 VP衣壳多肽相比,所述组装的病毒体表现出增加的甲状腺组织嗜性,其中工程化变体AAV5 VP衣壳多肽序列具有一个或多个突变,其中VP1多肽序列在从对应于SEQ ID NO:2中的581的位置至对应于SEQ ID NO:2中的589的位置的区域中具有所述一个或多个突变,并且其中所述一个或多个突变选自以下规则。在一些实施方案中,Xaa1选自A、K、M、N、Q或R。在一些实施方案中,Xaa1选自K、N或Q。在一些实施方案中,Xaa1为K。在一些实施方案中,Xaa2选自A、F、K、L、M、T、V或W。在一些实施方案中,Xaa2选自F、V或W。在一些实施方案中,Xaa2为W。在一些实施方案中,Xaa3选自A、I、K、R、S、T、V或W。在一些实施方案中,Xaa3选自A、R或T。在一些实施方案中,Xaa3为R。在一些实施方案中,Xaa4选自A、D、E、I、P或V。在一些实施方案中,Xaa4选自A、E或I。在一些实施方案中,Xaa4为A。在一些实施方案中,Xaa5选自F、I、M、Q、V或Y。在一些实施方案中,Xaa5为M、V、Y。在一些实施方案中,Xaa5为M。在一些实施方案中,Xaa6选自H、M、N或Y。在一些实施方案中,Xaa6为N。在一些实施方案中,Xaa7选自H、I、N、Q、S或W。在一些实施方案中,Xaa7选自H、I或N。在一些实施方案中,Xaa7为H。在一些实施方案中,Xaa8选

自A、D、F、Q、S或Y。在一些实施方案中，Xaa8选自A、F或S。在一些实施方案中，Xaa8为F。在一些实施方案中，Xaa9选自A、Q、S或Y。在一些实施方案中，Xaa9选自A或S。在一些实施方案中，Xaa9为A。

[0425] 本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽，与野生型AAV5 VP衣壳多肽相比，所述组装的病毒体表现出增加的甲状腺组织嗜性，其中工程化变体AAV5 VP衣壳多肽序列具有一个或多个突变，其中VP1多肽序列在从对应于SEQ ID NO:2中的581的位置至对应于SEQ ID NO:2中的589的位置的区域中具有所述一个或多个突变，并且其中所述一个或多个突变选自以下规则:Xaa1选自A、K、M、N、Q或R，Xaa2选自A、F、K、L、M、T、V或W，Xaa3选自A、I、K、R、S、T、V或W，Xaa4选自A、D、E、I、P或V，Xaa5选自F、I、M、Q、V或Y，Xaa6选自H、M、N或Y，Xaa7选自H、I、N、Q、S或W，Xaa8选自A、D、F、Q、S或Y，并且Xaa9选自A、Q、S或Y。

[0426] 本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽，与野生型AAV5 VP衣壳多肽相比，所述组装的病毒体表现出增加的甲状腺组织嗜性，其中工程化变体AAV5 VP衣壳多肽序列具有一个或多个突变，其中VP1多肽序列在从对应于SEQ ID NO:2中的581的位置至对应于SEQ ID NO:2中的589的位置的区域中具有所述一个或多个突变，并且其中所述一个或多个突变如下:Xaa1选自A、K、M、N、Q或R，Xaa2选自A、F、K、L、M、T、V或W，Xaa3选自A、I、K、R、S、T、V或W，Xaa4选自A、D、E、I、P或V，Xaa5选自F、I、M、Q、V或Y，Xaa6选自H、M、N或Y，Xaa7选自H、I、N、Q、S或W，Xaa8选自A、D、F、Q、S或Y，Xaa9选自A、Q、S或Y，或它们的任何组合。

[0427] 在一些实施方案中，本文提供了AAV5 VP衣壳多肽，所述多肽在从对应于AAV5 VP1的581的位置至对应于589的位置的区域中具有至少一个突变并且与选自SEQ ID NO:40438-SEQ ID NO:41437的任何序列具有至少70％、至少75％、至少80％、至少85％、至少90％、至少92％、至少95％、至少97％、至少99％或100％序列同一性，其中所述至少一个突变驱动增加的甲状腺组织嗜性。

[0428] B.ML规则

[0429] 对于本段中描述的以下一组规则，使用体内数据和两个ML模型(描述于实施例43)确定工程化AAV5 VP1衣壳多肽的581至589区域中每个位置处的有利生物物理特性和有利氨基酸残基。本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽，与具有野生型AAV5 VP衣壳多肽的rAAV病毒体相比，所述组装的病毒体表现出增加的甲状腺组织嗜性，其中工程化AAV5 VP衣壳多肽序列具有一个或多个突变，其中工程化AAV5 VP1衣壳多肽序列在从对应于SEQ ID NO:2中的581的位置至对应于SEQ ID NO:2中的589的位置的区域中具有所述一个或多个突变，并且其中Xaa1选自高可突变性的氨基酸(例如，Xaa1选自N)；或者其中Xaa2选自低表面可接近性的氨基酸(例如，Xaa2选自F、G或M)；或者其中Xaa3选自高溶解性的氨基酸(例如，Xaa3选自F)；或者其中Xaa3选自低可突变性的氨基酸(例如，Xaa3选自Y、F、L或C)；或者其中Xaa3选自中等摩尔质量的氨基酸(例如，Xaa3选自D、E、R、K、V、P、M、I、L、N、Q、T或C)；或者其中Xaa3选自低表面可接近性的氨基酸(例如，Xaa3选自V、I、L或C)；或者其中Xaa4选自高goldman engelman steitz的氨基酸(例如，Xaa4选自L或V)；或者其中Xaa4选自低表面可接近性的氨基酸(例如，Xaa4选自V、M、A、G、F、I或L)；或者其中Xaa4选自低摩尔质量的氨基酸(例如，Xaa4选自D、A、G、I、L或N)；或者其中Xaa5选自高溶解性的氨基酸(例如，Xaa5选自C、L、F、M、V或Y)；或者其中Xaa5选自低溶解性的氨基酸(例如，Xaa5选自D)；或者其中Xaa5选自低平均柔性的氨基酸(例如，Xaa5选自F、M或W)；或者其中Xaa6选自低

平均柔性的氨基酸(例如,Xaa6选自F、M或W);或者其中Xaa7选自高可突变性的氨基酸(例如,Xaa7选自N);或者其中Xaa7选自低体积的氨基酸(例如,Xaa7选自P、N或T);或者其中Xaa8选自低平均柔性的氨基酸(例如,Xaa8选自F、M或W);或者其中Xaa8选自低表面可接近性的氨基酸(例如,Xaa8选自M、G或F);或者其中Xaa9选自低可突变性的氨基酸(例如,Xaa9选自R、K、P、H或C);或者其中Xaa9选自低亲疏水性的氨基酸(例如,Xaa9选自R);或它们的任何组合。

[0430] 在一些实施方案中,Xaa1选自高可突变性的氨基酸。在一些实施方案中,Xaa1选自N。在一些实施方案中,Xaa2选自低表面可接近性的氨基酸。在一些实施方案中,Xaa2选自F、G或M。在一些实施方案中,Xaa3选自高溶解性的氨基酸。在一些实施方案中,Xaa3选自F。在一些实施方案中,Xaa3选自低可突变性的氨基酸。在一些实施方案中,Xaa3选自Y、F、L或C。在一些实施方案中,Xaa3选自中等摩尔质量的氨基酸。在一些实施方案中,Xaa3选自D、E、R、K、V、P、M、I、L、N、Q、T或C。在一些实施方案中,Xaa3选自低表面可接近性的氨基酸。在一些实施方案中,Xaa3选自V、I、L或C。在一些实施方案中,Xaa4选自高goldman engelman steitz的氨基酸。在一些实施方案中,Xaa4选自L或V。在一些实施方案中,Xaa4选自低表面可接近性的氨基酸。在一些实施方案中,Xaa4选自V、M、A、G、F、I或L。在一些实施方案中,Xaa4选自低摩尔质量的氨基酸。在一些实施方案中,Xaa4选自D、A、G、I、L或N。在一些实施方案中,Xaa5选自高溶解性的氨基酸。在一些实施方案中,Xaa5选自C、L、F、M、V或Y。在一些实施方案中,Xaa5选自低溶解性的氨基酸。在一些实施方案中,Xaa5选自D。在一些实施方案中,Xaa5选自低平均柔性的氨基酸。在一些实施方案中,Xaa5选自F、M或W。在一些实施方案中,Xaa6选自低平均柔性的氨基酸。在一些实施方案中,Xaa6选自F、M或W。在一些实施方案中,Xaa7选自高可突变性的氨基酸。在一些实施方案中,Xaa7选自N。在一些实施方案中,Xaa7选自低体积的氨基酸。在一些实施方案中,Xaa7选自P、N或T。在一些实施方案中,Xaa8选自低平均柔性的氨基酸。在一些实施方案中,Xaa8选自F、M或W。在一些实施方案中,Xaa8选自低表面可接近性的氨基酸。在一些实施方案中,Xaa8选自M、G或F。在一些实施方案中,Xaa9选自低可突变性的氨基酸。在一些实施方案中,Xaa9选自R、K、P、H或C。在一些实施方案中,Xaa9选自低亲疏水性的氨基酸。在一些实施方案中,Xaa9选自R。

[0431] 在一些实施方案中,本文提供了AAV5 VP衣壳多肽,所述多肽在从对应于AAV5 VP1的581的位置至对应于589的位置的区域中具有至少一个突变并且与选自SEQ ID NO:41438-SEQ ID NO:42437的任何序列具有至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%、至少99%或100%序列同一性,其中所述至少一个突变驱动增加的甲状腺组织嗜性。

[0432] C.富集的甲状腺序列

[0433] 在一些实施方案中,本文提供了AAV5 VP衣壳多肽,所述多肽在从对应于AAV5 VP1的581的位置至对应于589的位置的区域中具有至少一个突变并且与选自SEQ ID NO:42438-SEQ ID NO:43437的任何序列具有至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%、至少99%或100%序列同一性,其中所述至少一个突变驱动增加的甲状腺组织嗜性。

[0434] 6.7.17.体内选定的赋予增加的淋巴结嗜性的突变VP多肽

[0435] 本公开提供了具有VP衣壳多肽的AAV5病毒体,所述VP衣壳多肽在具有与靶细胞

(例如,目标靶淋巴结组织中的靶淋巴结细胞)相互作用的残基的区域中具有至少一个突变,其中与野生型VP衣壳多肽相比,所述至少一个突变赋予增加的淋巴结组织嗜性。在一些实施方案中,本文提供了与SEQ ID NO:1具有至少80%序列同源性的AAV5 VP1衣壳多肽,其中所述AAV5 VP1衣壳多肽在从对应于SEQ ID NO:1的581的位置至对应于589的位置的区域中具有至少一个突变,并且其中所述至少一个突变驱动增加的淋巴结嗜性。以下序列规则和序列也适用于对应于AAV5 VP1氨基酸残基581至589的AAV5 VP2(氨基酸残基445至453;VP2序列显示于SEQ ID NO:1115)和AAV5 VP3(氨基酸残基389至397;VP3序列显示于SEQ ID NO:1116)中的区域。因此,本公开涵盖在对应于581至589区域的AAV5 VP1氨基酸残基的VP2和VP3区域中具有一个或多个突变的AAV5 VP2衣壳多肽和AAV5 VP3衣壳多肽,其中所述一个或多个突变与以下章节中的规则或顺序相称。

[0436] A.位置频率规则

[0437] 在本章节中,除非另外说明,否则分析相对于给定氨基酸残基出现在所有其他所收获组织(CNS(前脑皮层、枕叶皮层、颞叶皮层、丘脑、下丘脑、黑质、海马DG、海马CA1、海马CA3、小脑)、肝脏、骨骼肌、心脏、肺、脾脏、骨髓、乳腺、皮肤、肾上腺、甲状腺、结肠、坐骨神经和脊髓组织)中鉴定的变体中指定位置的频率而言该给定氨基酸残基出现在淋巴结中鉴定的变体中对应于SEQ ID NO:1(推及于SEQ ID NO:2)的位置581至位置589的所述指定位置的频率,以确定优先靶向淋巴结组织的衣壳的一组序列规则。来自体内数据的位置频率规则的确定详细描述于实施例16中。

[0438] 本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽,与野生型AAV5 VP衣壳多肽相比,所述组装的病毒体表现出增加的淋巴结组织嗜性,其中工程化变体AAV5 VP1衣壳多肽序列具有一个或多个突变,其中Xaa1选自A、D、E、Q、S或T,或Xaa1选自D、E或T,或Xaa1为E;或Xaa2选自A、H、I、S、T或V,或Xaa2选自I、T或V,或Xaa2为V;或Xaa3选自A、E、H、I、T或V,或Xaa3选自A、I、T或V,或Xaa3为T;或Xaa4选自A、D、E或P,或Xaa4选自D或E,或Xaa4为E;或Xaa5选自I、L、M、V或Y,或Xaa5选自I、L、V或Y,或Xaa5为L;或Xaa6选自D、E、I、N或Q,或Xaa6选自D、E或I,或Xaa6为D;或Xaa7选自A、E、G、Q或V,或Xaa7为A、Q或V,或Xaa7为V;或Xaa8选自F、G、M或W,或Xaa8选自F或W,或Xaa8为W;或Xaa9选自I、P、T或Y,或Xaa9为I或P。

[0439] 本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽,与野生型AAV5 VP衣壳多肽相比,所述组装的病毒体表现出增加的淋巴结组织嗜性,其中工程化变体AAV5 VP衣壳多肽序列具有一个或多个突变,其中VP1多肽序列在从对应于SEQ ID NO:2中的581的位置至对应于SEQ ID NO:2中的589的位置的区域中具有所述一个或多个突变,并且其中所述一个或多个突变选自以下规则。在一些实施方案中,Xaa1选自A、D、E、Q、S或T。在一些实施方案中,Xaa1选自D、E或T。在一些实施方案中,Xaa1为E。在一些实施方案中,Xaa2选自A、H、I、S、T或V。在一些实施方案中,Xaa2选自I、T或V。在一些实施方案中,Xaa2为V。在一些实施方案中,Xaa3选自A、E、H、I、T或V。在一些实施方案中,Xaa3选自A、I、T或V。在一些实施方案中,Xaa3为T。在一些实施方案中,Xaa4选自A、D、E或P。在一些实施方案中,Xaa4选自D或E。在一些实施方案中,Xaa4为E。在一些实施方案中,Xaa5选自I、L、M、V或Y。在一些实施方案中,Xaa5选自I、L、V或Y。在一些实施方案中,Xaa5为L。在一些实施方案中,Xaa6选自D、E、I、N或Q。在一些实施方案中,Xaa6选自D、E或I。在一些实施方案中,Xaa6为D。在一些实施方案中,Xaa7选自A、E、G、Q或V。在一些实施方案中,Xaa7为A、Q或V。在一些实施方案中,Xaa7为V。在

一些实施方案中，Xaa8选自F、G、M或W。在一些实施方案中，Xaa8选自F或W。在一些实施方案中，Xaa8为W。在一些实施方案中，Xaa9选自I、P、T或Y。在一些实施方案中，Xaa9为I或P。

[0440] 本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽，与野生型AAV5 VP衣壳多肽相比，所述组装的病毒体表现出增加的淋巴结组织嗜性，其中工程化变体AAV5 VP衣壳多肽序列具有一个或多个突变，其中VP1多肽序列在从对应于SEQ ID NO:2中的581的位置至对应于SEQ ID NO:2中的589的位置的区域中具有所述一个或多个突变，并且其中所述一个或多个突变选自以下规则:Xaa1选自A、D、E、Q、S或T，Xaa2选自A、H、I、S、T或V，Xaa3选自A、E、H、I、T或V，Xaa4选自A、D、E或P，Xaa5选自I、L、M、V或Y，Xaa6选自D、E、I、N或Q，Xaa7选自A、E、G、Q或V，Xaa8选自F、G、M或W，并且Xaa9选自I、P、T或Y。

[0441] 本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽，与野生型AAV5 VP衣壳多肽相比，所述组装的病毒体表现出增加的淋巴结组织嗜性，其中工程化变体AAV5 VP衣壳多肽序列具有一个或多个突变，其中VP1多肽序列在从对应于SEQ ID NO:2中的581的位置至对应于SEQ ID NO:2中的589的位置的区域中具有所述一个或多个突变，并且其中所述一个或多个突变如下:Xaa1选自A、D、E、Q、S或T，Xaa2选自A、H、I、S、T或V，Xaa3选自A、E、H、I、T或V，Xaa4选自A、D、E或P，Xaa5选自I、L、M、V或Y，Xaa6选自D、E、I、N或Q，Xaa7选自A、E、G、Q或V，Xaa8选自F、G、M或W，Xaa9选自I、P、T或Y，或它们的任何组合。

[0442] 在一些实施方案中，本文提供了AAV5 VP衣壳多肽，所述多肽在从对应于AAV5 VP1的581的位置至对应于589的位置的区域中具有至少一个突变并且与选自SEQ ID NO:19118-SEQ ID NO:20117的任何序列具有至少70％、至少75％、至少80％、至少85％、至少90％、至少92％、至少95％、至少97％、至少99％或100％序列同一性，其中所述至少一个突变驱动增加的淋巴结组织嗜性。

[0443] B.ML规则

[0444] 对于本段中描述的以下一组规则，使用体内数据和两个ML模型(描述于实施例35)确定工程化AAV5 VP1衣壳多肽的581至589区域中每个位置处的有利生物物理特性和有利氨基酸残基。本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽，与具有野生型AAV5 VP衣壳多肽的rAAV病毒体相比，所述组装的病毒体表现出增加的淋巴结组织嗜性，其中工程化AAV5 VP衣壳多肽序列具有一个或多个突变，其中工程化AAV5 VP1衣壳多肽序列在从对应于SEQ ID NO:2中的581的位置至对应于SEQ ID NO:2中的589的位置的区域中具有所述一个或多个突变，并且其中Xaa1选自高平均柔性的氨基酸(例如，Xaa1选自D、E、P、G、Q、S或R)；或者其中Xaa1选自高氢键供体的氨基酸(例如，Xaa1选自R)；或者其中Xaa1选自高摩尔质量的氨基酸(例如，Xaa1选自Y、W、R或F)；或者其中Xaa2选自低溶解性的氨基酸(例如，Xaa2选自N或E)；或者其中Xaa3选自低平均柔性的氨基酸(例如，Xaa3选自W、M或F)；或者其中Xaa3选自低可突变性的氨基酸(例如，Xaa3选自R、H、K、P、Y、F、L或C)；或者其中Xaa4选自低可突变性的氨基酸(例如，Xaa4选自C)；或者其中Xaa5选自高可突变性的氨基酸(例如，Xaa5选自N)；或者其中Xaa5选自中等摩尔质量的氨基酸(例如，Xaa5选自D、I、L或N)；或者其中Xaa6选自高摩尔质量的氨基酸(例如，Xaa6选自Y、W、R或F)；或者其中Xaa6选自高平均柔性的氨基酸(例如，Xaa6选自G或R)；或者其中Xaa7选自高平均柔性的氨基酸(例如，Xaa7选自D、E、K、P、I、N、Q或S)；或者其中Xaa7选自低溶解性的氨基酸(例如，Xaa7选自N、E)；或者其中Xaa8选自低溶解性的氨基酸(例如，Xaa8选自N、E或D)；或者其中Xaa8选自中等可突变性

的氨基酸(例如,Xaa8选自R或H);或者其中Xaa9选自低可突变性的氨基酸(例如,Xaa9选自P或K);或者其中Xaa9选自高平均柔性的氨基酸(例如,Xaa9选自D、E、P或S);或者其中Xaa9选自高溶解性的氨基酸(例如,Xaa9选自M或V);或它们的任何组合。

[0445] 在一些实施方案中,Xaa1选自高平均柔性的氨基酸。在一些实施方案中,Xaa1选自D、E、P、G、Q、S或R。在一些实施方案中,Xaa1选自高氢键供体的氨基酸。在一些实施方案中,Xaa1选自R。在一些实施方案中,Xaa1选自高摩尔质量的氨基酸。在一些实施方案中,Xaa1选自Y、W、R或F。在一些实施方案中,Xaa2选自低溶解性的氨基酸。在一些实施方案中,Xaa2选自N或E。在一些实施方案中,Xaa3选自低平均柔性的氨基酸。在一些实施方案中,Xaa3选自W、M或F。在一些实施方案中,Xaa3选自低可突变性的氨基酸。在一些实施方案中,Xaa3选自R、H、K、P、Y、F、L或C。在一些实施方案中,Xaa4选自低可突变性的氨基酸。在一些实施方案中,Xaa4选自C。在一些实施方案中,Xaa5选自高可突变性的氨基酸。在一些实施方案中,Xaa5选自N。在一些实施方案中,Xaa5选自中等摩尔质量的氨基酸。在一些实施方案中,Xaa5选自D、I、L或N。在一些实施方案中,Xaa6选自高摩尔质量的氨基酸。在一些实施方案中,Xaa6选自Y、W、R或F。在一些实施方案中,Xaa6选自高平均柔性的氨基酸。在一些实施方案中,Xaa6选自G、R。在一些实施方案中,Xaa7选自高平均柔性的氨基酸。在一些实施方案中,Xaa7选自D、E、K、P、I、N、Q或S。在一些实施方案中,Xaa7选自低溶解性的氨基酸。在一些实施方案中,Xaa7选自N或E。在一些实施方案中,Xaa8选自低溶解性的氨基酸。在一些实施方案中,Xaa8选自N、E或D。在一些实施方案中,Xaa8选自中等可突变性的氨基酸。在一些实施方案中,Xaa8选自R或H。在一些实施方案中,Xaa9选自低可突变性的氨基酸。在一些实施方案中,Xaa9选自P或K。在一些实施方案中,Xaa9选自高平均柔性的氨基酸。在一些实施方案中,Xaa9选自D、E、P或S。在一些实施方案中,Xaa9选自高溶解性的氨基酸。在一些实施方案中,Xaa9选自M或V。

[0446] 在一些实施方案中,本文提供了AAV5 VP衣壳多肽,所述多肽在从对应于AAV5 VP1的581的位置至对应于589的位置的区域中具有至少一个突变并且与选自SEQ ID NO:20118-SEQ ID NO:21117的任何序列具有至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%、至少99%或100%序列同一性,其中所述至少一个突变驱动增加的淋巴结组织嗜性。

[0447] C.富集的淋巴结序列

[0448] 在一些实施方案中,本文提供了AAV5 VP衣壳多肽,所述多肽在从对应于AAV5 VP1的581的位置至对应于589的位置的区域中具有至少一个突变并且与选自SEQ ID NO:21118-SEQ ID NO:22117的任何序列具有至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%、至少99%或100%序列同一性,其中所述至少一个突变驱动增加的淋巴结组织嗜性。

[0449] 6.7.18.体内选定的赋予增加的皮肤嗜性的突变VP多肽

[0450] 本公开提供了具有VP衣壳多肽的AAV5病毒体,所述VP衣壳多肽在具有与靶细胞(例如,目标靶皮肤组织中的靶皮肤细胞)相互作用的残基的区域中具有至少一个突变,其中与野生型VP衣壳多肽相比,所述至少一个突变赋予增加的皮肤组织嗜性。在一些实施方案中,本文提供了与SEQ ID NO:1具有至少80%序列同源性的AAV5 VP1衣壳多肽,其中所述AAV5 VP1衣壳多肽在从对应于SEQ ID NO:1的581的位置至对应于589的位置的区域中具有

至少一个突变，并且其中所述至少一个突变驱动增加的皮肤嗜性。以下序列规则和序列也适用于对应于AAV5 VP1氨基酸残基581至589的AAV5 VP2(氨基酸残基445至453;VP2序列显示于SEQ ID NO:1115)和AAV5 VP3(氨基酸残基389至397;VP3序列显示于SEQ ID NO:1116)中的区域。因此，本公开涵盖在对应于581至589区域的AAV5 VP1氨基酸残基的VP2和VP3区域中具有一个或多个突变的AAV5 VP2衣壳多肽和AAV5 VP3衣壳多肽，其中所述一个或多个突变与以下章节中的规则或顺序相称。

[0451] A.位置频率规则

[0452] 在本章节中，除非另外说明，否则分析相对于给定氨基酸残基出现在所有其他所收获组织(CNS(前脑皮层、枕叶皮层、颞叶皮层、丘脑、下丘脑、黑质、海马DG、海马CA1、海马CA3、小脑)、肝脏、骨骼肌、心脏、肺、脾脏、淋巴结、骨髓、乳腺、肾上腺、甲状腺、结肠、坐骨神经和脊髓组织)中鉴定的变体中指定位置的频率而言该给定氨基酸残基出现在皮肤中鉴定的变体中对应于SEQ ID NO:1(推及于SEQ ID NO:2)的位置581至位置589的所述指定位置的频率，以确定优先靶向皮肤组织的衣壳的一组序列规则。来自体内数据的位置频率规则的确定详细描述于实施例17中。

[0453] 本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽，与野生型AAV5 VP衣壳多肽相比，所述组装的病毒体表现出增加的皮肤组织嗜性，其中工程化变体AAV5 VP1衣壳多肽序列具有一个或多个突变，其中Xaa1选自A、C、K、Q、R或T，或Xaa1选自C、K或R，或Xaa1为C;或Xaa2选自A、C、I、S、T或V，或Xaa2选自A、S、T或V，或Xaa2为V;或Xaa3选自A、C、F、G、M、Q、S或V，或Xaa3选自A、C、F、M或Q，或Xaa3为C;或Xaa4选自C、K、L、P、R或W，或Xaa4选自L、P或R，或Xaa4为R;或Xaa5选自F、H、I、M、V或Y，或Xaa5选自M、V或Y，或Xaa5为Y;或Xaa6选自F、H、I、M、N、Q或S，或Xaa6选自M、N或Q，或Xaa6为N;或Xaa7选自A、H、K、M、N、R或V，或Xaa7为A、H、K或R，或Xaa7为K;或Xaa8选自A、F、G、H、S或Y，或Xaa8选自A、F或S，或Xaa8为S;或Xaa9选自A、E、G、P、Q、R或S，或Xaa9选自A、Q或S，或Xaa9为A。

[0454] 本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽，与野生型AAV5 VP衣壳多肽相比，所述组装的病毒体表现出增加的皮肤组织嗜性，其中工程化变体AAV5 VP衣壳多肽序列具有一个或多个突变，其中VP1多肽序列在从对应于SEQ ID NO:2中的581的位置至对应于SEQ ID NO:2中的589的位置的区域中具有所述一个或多个突变，并且其中所述一个或多个突变选自以下规则。在一些实施方案中，Xaa1选自A、C、K、Q、R或T。在一些实施方案中，Xaa1选自C、K或R。在一些实施方案中，Xaa1为C。在一些实施方案中，Xaa2选自A、C、I、S、T或V。在一些实施方案中，Xaa2选自A、S、T或V。在一些实施方案中，Xaa2为V。在一些实施方案中，Xaa3选自A、C、F、G、M、Q、S或V。在一些实施方案中，Xaa3选自A、C、F、M或Q。在一些实施方案中，Xaa3为C。在一些实施方案中，Xaa4选自C、K、L、P、R或W。在一些实施方案中，Xaa4选自L、P或R。在一些实施方案中，Xaa4为R。在一些实施方案中，Xaa5选自F、H、I、M、V或Y。在一些实施方案中，Xaa5选自M、V或Y。在一些实施方案中，Xaa5为Y。在一些实施方案中，Xaa6选自F、H、I、M、N、Q或S。在一些实施方案中，Xaa6选自M、N或Q。在一些实施方案中，Xaa6为N。在一些实施方案中，Xaa7选自A、H、K、M、N、R或V。在一些实施方案中，Xaa7为A、H、K或R。在一些实施方案中，Xaa7为K。在一些实施方案中，Xaa8选自A、F、G、H、S或Y。在一些实施方案中，Xaa8选自A、F或S。在一些实施方案中，Xaa8为S。在一些实施方案中，Xaa9选自A、E、G、P、Q、R或S。在一些实施方案中，Xaa9选自A、Q或S。在一些实施方案中，Xaa9为A。

[0455] 本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽，与野生型AAV5 VP衣壳多肽相比，所述组装的病毒体表现出增加的皮肤组织嗜性，其中工程化变体AAV5 VP衣壳多肽序列具有一个或多个突变，其中VP1多肽序列在从对应于SEQ ID NO:2中的581的位置至对应于SEQ ID NO:2中的589的位置的区域中具有所述一个或多个突变，并且其中所述一个或多个突变选自以下规则:Xaa1选自A、C、K、Q、R或T，Xaa2选自A、C、I、S、T或V，Xaa3选自A、C、F、G、M、Q、S或V，Xaa4选自C、K、L、P、R或W，Xaa5选自F、H、I、M、V或Y，Xaa6选自F、H、I、M、N、Q或S，Xaa7选自A、H、K、M、N、R或V，Xaa8选自A、F、G、H、S或Y，并且Xaa9选自A、E、G、P、Q、R或S。

[0456] 本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽，与野生型AAV5 VP衣壳多肽相比，所述组装的病毒体表现出增加的皮肤组织嗜性，其中工程化变体AAV5 VP衣壳多肽序列具有一个或多个突变，其中VP1多肽序列在从对应于SEQ ID NO:2中的581的位置至对应于SEQ ID NO:2中的589的位置的区域中具有所述一个或多个突变，并且其中所述一个或多个突变如下:Xaa1选自A、C、K、Q、R或T，Xaa2选自A、C、I、S、T或V，Xaa3选自A、C、F、G、M、Q、S或V，Xaa4选自C、K、L、P、R或W，Xaa5选自F、H、I、M、V或Y，Xaa6选自F、H、I、M、N、Q或S，Xaa7选自A、H、K、M、N、R或V，Xaa8选自A、F、G、H、S或Y，Xaa9选自A、E、G、P、Q、R或S，或它们的任何组合。

[0457] 在一些实施方案中，本文提供了AAV5 VP衣壳多肽，所述多肽在从对应于AAV5 VP1的581的位置至对应于589的位置的区域中具有至少一个突变并且与选自SEQ ID NO:31991-SEQ ID NO:32990的任何序列具有至少70％、至少75％、至少80％、至少85％、至少90％、至少92％、至少95％、至少97％、至少99％或100％序列同一性，其中所述至少一个突变驱动增加的皮肤组织嗜性。

[0458] B.ML规则

[0459] 对于本段中描述的以下一组规则，使用体内数据和两个ML模型(描述于实施例40)确定工程化AAV5 VP1衣壳多肽的581至589区域中每个位置处的有利生物物理特性和有利氨基酸残基。本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽，与具有野生型AAV5 VP衣壳多肽的rAAV病毒体相比，所述组装的病毒体表现出增加的皮肤组织嗜性，其中工程化AAV5 VP衣壳多肽序列具有一个或多个突变，其中工程化AAV5 VP1衣壳多肽序列在从对应于SEQ ID NO:2中的581的位置至对应于SEQ ID NO:2中的589的位置的区域中具有所述一个或多个突变，并且其中Xaa1选自低表面可接近性的氨基酸(例如，Xaa1选自C)；或者其中Xaa1选自低体积的氨基酸(例如，Xaa1选自C)；或者其中Xaa1选自低可突变性的氨基酸(例如，Xaa1选自C)；或者其中Xaa2选自高表面可接近性的氨基酸(例如，Xaa2选自R或K)；或者其中Xaa2选自高平均柔性的氨基酸(例如，Xaa2选自K、I或N)；或者其中Xaa2选自低可突变性的氨基酸(例如，Xaa2选自P或K)；或者其中Xaa3选自高亲疏水性的氨基酸(例如，Xaa3选自I或V)；或者其中Xaa4选自低可突变性的氨基酸(例如，Xaa4选自L、F或Y)；或者其中Xaa4选自低平均柔性的氨基酸(例如，Xaa4选自W、H、F或M)；或者其中Xaa5选自高平均柔性的氨基酸(例如，Xaa5选自G、R、K、I或N)；或者其中Xaa6选自高平均柔性的氨基酸(例如，Xaa6选自G、R、K、I或N)；或者其中Xaa8选自高表面可接近性的氨基酸(例如，Xaa8选自M、G或F)；或者其中Xaa8选自低平均柔性的氨基酸(例如，Xaa8选自H、F、M或W)；或者其中Xaa8选自低可突变性的氨基酸(例如，Xaa8选自L、F、Y)；或者其中Xaa9选自高平均柔性的氨基酸(例如，Xaa9选自D、E、R、K、P或G)；或者其中Xaa9选自高可突变性的氨基酸(例如，Xaa9选自D、E、

R、V、A或H);或它们的任何组合。

[0460] 在一些实施方案中,Xaa1选自低表面可接近性的氨基酸。在一些实施方案中,Xaa1选自C。在一些实施方案中,Xaa1选自低体积的氨基酸。在一些实施方案中,Xaa1选自C。在一些实施方案中,Xaa1选自低可突变性的氨基酸。在一些实施方案中,Xaa1选自C。在一些实施方案中,Xaa2选自高表面可接近性的氨基酸。在一些实施方案中,Xaa2选自R或K。在一些实施方案中,Xaa2选自高平均柔性的氨基酸。在一些实施方案中,Xaa2选自K、I或N。在一些实施方案中,Xaa2选自低可突变性的氨基酸。在一些实施方案中,Xaa2选自P或K。在一些实施方案中,Xaa3选自高亲疏水性的氨基酸。在一些实施方案中,Xaa3选自I或V。在一些实施方案中,Xaa4选自低可突变性的氨基酸。在一些实施方案中,Xaa4选自L、F或Y。在一些实施方案中,Xaa4选自低平均柔性的氨基酸。在一些实施方案中,Xaa4选自W、H、F或M。在一些实施方案中,Xaa5选自高平均柔性的氨基酸。在一些实施方案中,Xaa5选自G、R、K、I或N。在一些实施方案中,Xaa6选自高平均柔性的氨基酸。在一些实施方案中,Xaa6选自G、R、K、I或N。在一些实施方案中,Xaa8选自高表面可接近性的氨基酸。在一些实施方案中,Xaa8选自M、G或F。在一些实施方案中,Xaa8选自低平均柔性的氨基酸。在一些实施方案中,Xaa8选自H、F、M或W。在一些实施方案中,Xaa8选自低可突变性的氨基酸。在一些实施方案中,Xaa8选自L、F或Y。在一些实施方案中,Xaa9选自高平均柔性的氨基酸。在一些实施方案中,Xaa9选自D、E、R、K、P或G。在一些实施方案中,Xaa9选自高可突变性的氨基酸。在一些实施方案中,Xaa9选自D、E、R、V、A或H。

[0461] 在一些实施方案中,本文提供了AAV5 VP衣壳多肽,所述多肽在从对应于AAV5 VP1的581的位置至对应于589的位置的区域中具有至少一个突变并且与选自SEQ ID NO:32991-SEQ ID NO:33990的任何序列具有至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%、至少99%或100%序列同一性,其中所述至少一个突变驱动增加的皮肤组织嗜性。

[0462] C.富集的皮肤序列

[0463] 在一些实施方案中,本文提供了AAV5 VP衣壳多肽,所述多肽在从对应于AAV5 VP1的581的位置至对应于589的位置的区域中具有至少一个突变并且与选自SEQ ID NO:33991-SEQ ID NO:34990的任何序列具有至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%、至少99%或100%序列同一性,其中所述至少一个突变驱动增加的皮肤组织嗜性。

[0464] 6.7.19.体内选定的赋予增加的骨髓嗜性的突变VP多肽

[0465] 本公开提供了具有VP衣壳多肽的AAV5病毒体,所述VP衣壳多肽在具有与靶细胞(例如,目标靶骨髓组织中的靶骨髓细胞)相互作用的残基的区域中具有至少一个突变,其中与野生型VP衣壳多肽相比,所述至少一个突变赋予增加的骨髓组织嗜性。在一些实施方案中,本文提供了与SEQ ID NO:1具有至少80%序列同源性的AAV5 VP1衣壳多肽,其中所述AAV5 VP1衣壳多肽在从对应于SEQ ID NO:1的581的位置至对应于589的位置的区域中具有至少一个突变,并且其中所述至少一个突变驱动增加的骨髓嗜性。以下序列规则和序列也适用于对应于AAV5 VP1氨基酸残基581至589的AAV5 VP2(氨基酸残基445至453;VP2序列显示于SEQ ID NO:1115)和AAV5 VP3(氨基酸残基389至397;VP3序列显示于SEQ ID NO:1116)中的区域。因此,本公开涵盖在对应于581至589区域的AAV5 VP1氨基酸残基的VP2和VP3区

域中具有一个或多个突变的AAV5 VP2衣壳多肽和AAV5 VP3衣壳多肽，其中所述一个或多个突变与以下章节中的规则或顺序相称。

[0466] A.位置频率规则

[0467] 在本章节中，除非另外说明，否则分析相对于给定氨基酸残基出现在所有其他所收获组织(CNS(前脑皮层、枕叶皮层、颞叶皮层、丘脑、下丘脑、黑质、海马DG、海马CA1、海马CA3、小脑)、肝脏、骨骼肌、心脏、肺、脾脏、淋巴结、乳腺、皮肤、肾上腺、甲状腺、结肠、坐骨神经和脊髓组织)中鉴定的变体中指定位置的频率而言该给定氨基酸残基出现在骨髓中鉴定的变体中对应于SEQ ID NO:1(推及于SEQ ID NO:2)的位置581至位置589的所述指定位置的频率，以确定优先靶向骨髓组织的衣壳的一组序列规则。来自体内数据的位置频率规则的确定详细描述于实施例18中。

[0468] 本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽，与野生型AAV5 VP衣壳多肽相比，所述组装的病毒体表现出增加的骨髓组织嗜性，其中工程化变体AAV5 VP1衣壳多肽序列具有一个或多个突变，其中Xaa1选自A、E、G、Q、S或T，或Xaa1选自A、E或T，或Xaa1为E;或Xaa2选自A、I、Q、S、T、V或Y，或Xaa2选自A、S、T，或Xaa2为A;或Xaa3选自A、G、I、M、Q、S或T，或Xaa3选自A、Q或T，或Xaa3为Q;或Xaa4选自A、E、P、Q、T或V，或Xaa4选自A、P或Q，或Xaa4为Q;或Xaa5选自F、I、L、M、Q、V或Y，或Xaa5选自F、V或Y，或Xaa5为V;或Xaa6选自F、I、N、Q、S或V，或Xaa6选自I、N、Q或S，或Xaa6为S;或Xaa7选自A、C、M、S或V，或Xaa7为A、C或V，或Xaa7为C;或Xaa8选自A、C、D、G、M、S或Y，或Xaa8选自A、M、S或Y，或Xaa8为M;或Xaa9选自D、E、G、L、P、S或Y，或Xaa9选自D、E或P，或Xaa9为P。

[0469] 本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽，与野生型AAV5 VP衣壳多肽相比，所述组装的病毒体表现出增加的骨髓组织嗜性，其中工程化变体AAV5 VP衣壳多肽序列具有一个或多个突变，其中VP1多肽序列在从对应于SEQ ID NO:2中的581的位置至对应于SEQ ID NO:2中的589的位置的区域中具有所述一个或多个突变，并且其中所述一个或多个突变选自以下规则。在一些实施方案中，Xaa1选自A、E、G、Q、S或T。在一些实施方案中，Xaa1选自A、E或T。在一些实施方案中，Xaa1为E。在一些实施方案中，Xaa2选自A、I、Q、S、T、V或Y。在一些实施方案中，Xaa2选自A、S、T。在一些实施方案中，Xaa2为A。在一些实施方案中，Xaa3选自A、G、I、M、Q、S或T。在一些实施方案中，Xaa3选自A、Q或T。在一些实施方案中，Xaa3为Q。在一些实施方案中，Xaa4选自A、E、P、Q、T或V。在一些实施方案中，Xaa4选自A、P或Q。在一些实施方案中，Xaa4为Q。在一些实施方案中，Xaa5选自F、I、L、M、Q、V或Y。在一些实施方案中，Xaa5选自F、V或Y。在一些实施方案中，Xaa5为V。在一些实施方案中，Xaa6选自F、I、N、Q、S或V。在一些实施方案中，Xaa6选自I、N、Q或S。在一些实施方案中，Xaa6为S。在一些实施方案中，Xaa7选自A、C、M、S或V。在一些实施方案中，Xaa7为A、C或V。在一些实施方案中，Xaa7为C。在一些实施方案中，Xaa8选自A、C、D、G、M、S或Y。在一些实施方案中，Xaa8选自A、M、S或Y。在一些实施方案中，Xaa8为M。在一些实施方案中，Xaa9选自D、E、G、L、P、S或Y。在一些实施方案中，Xaa9选自D、E或P。在一些实施方案中，Xaa9为P。

[0470] 本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽，与野生型AAV5 VP衣壳多肽相比，所述组装的病毒体表现出增加的骨髓组织嗜性，其中工程化变体AAV5 VP衣壳多肽序列具有一个或多个突变，其中VP1多肽序列在从对应于SEQ ID NO:2中的581的位置至对应于SEQ ID NO:2中的589的位置的区域中具有所述一个或多个突变，并且其中所述

一个或多个突变选自以下规则:Xaa1选自A、E、G、Q、S或T,Xaa2选自A、I、Q、S、T、V或Y,Xaa3选自A、G、I、M、Q、S或T,Xaa4选自A、E、P、Q、T或V,Xaa5选自F、I、L、M、Q、V或Y,Xaa6选自F、I、N、Q、S或V,Xaa7选自A、C、M、S或V,Xaa8选自A、C、D、G、M、S或Y,并且Xaa9选自D、E、G、L、P、S或Y。

[0471] 本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽,与野生型AAV5 VP衣壳多肽相比,所述组装的病毒体表现出增加的骨髓组织嗜性,其中工程化变体AAV5 VP衣壳多肽序列具有一个或多个突变,其中VP1多肽序列在从对应于SEQ ID NO:2中的581的位置至对应于SEQ ID NO:2中的589的位置的区域中具有所述一个或多个突变,并且其中所述一个或多个突变如下:Xaa1选自A、E、G、Q、S或T,Xaa2选自A、I、Q、S、T、V或Y,Xaa3选自A、G、I、M、Q、S或T,Xaa4选自A、E、P、Q、T或V,Xaa5选自F、I、L、M、Q、V或Y,Xaa6选自F、I、N、Q、S或V,Xaa7选自A、C、M、S或V,Xaa8选自A、C、D、G、M、S或Y,Xaa9选自D、E、G、L、P、S或Y,或它们的任何组合。

[0472] 在一些实施方案中,本文提供了AAV5 VP衣壳多肽,所述多肽在从对应于AAV5 VP1的581的位置至对应于589的位置的区域中具有至少一个突变并且与选自SEQ ID NO:4118-SEQ ID NO:5117的任何序列具有至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%、至少99%或100%序列同一性,其中所述至少一个突变驱动增加的骨髓组织嗜性。

[0473] B.ML规则

[0474] 对于本段中描述的以下一组规则,使用体内数据和两个ML模型(描述于实施例32)确定工程化AAV5 VP1衣壳多肽的581至589区域中每个位置处的有利生物物理特性和有利氨基酸残基。本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽,与具有野生型AAV5 VP衣壳多肽的rAAV病毒体相比,所述组装的病毒体表现出增加的骨髓组织嗜性,其中工程化AAV5 VP衣壳多肽序列具有一个或多个突变,其中工程化AAV5 VP1衣壳多肽序列在从对应于SEQ ID NO:2中的581的位置至对应于SEQ ID NO:2中的589的位置的区域中具有所述一个或多个突变,并且其中Xaa1选自高亲疏水性的氨基酸(例如,Xaa1选自V、I或L);或者其中Xaa1选自低可突变性的氨基酸(例如,Xaa1选自Y、L、F或C);或者其中Xaa2选自低亲疏水性的氨基酸(例如,Xaa2选自Y或W);或者其中Xaa2选自高摩尔质量的氨基酸(例如,Xaa2选自W);或者其中Xaa2选自低表面可接近性的氨基酸(例如,Xaa2选自W或A);或者其中Xaa2选自低亲水性的氨基酸(例如,Xaa2选自W);或者其中Xaa2选自低可突变性的氨基酸(例如,Xaa2选自C);或者其中Xaa2选自低平均柔性的氨基酸(例如,Xaa2选自W、M或F);或者其中Xaa5选自低平均柔性的氨基酸(例如,Xaa5选自W、M或F);或者其中Xaa6选自低平均柔性的氨基酸(例如,Xaa6选自W、M或F);或者其中Xaa6选自低可突变性的氨基酸(例如,Xaa6选自Y、F、L或C);或者其中Xaa6选自高溶解性的氨基酸(例如,Xaa6选自W、F、I或L);或者其中Xaa7选自低表面可接近性的氨基酸(例如,Xaa7选自C);或者其中Xaa7选自高表面可接近性的氨基酸(例如,Xaa7选自D或N);或者其中Xaa7选自低可突变性的氨基酸(例如,Xaa7选自C);或者其中Xaa7选自高溶解性的氨基酸(例如,Xaa7选自C);或者其中Xaa7选自低溶解性的氨基酸(例如,Xaa7选自D);或者其中Xaa8选自低电荷的氨基酸(例如,Xaa8选自D或E);或者其中Xaa8选自高可突变性的氨基酸(例如,Xaa8选自D、E、A或T);或者其中Xaa9选自高摩尔质量的氨基酸(例如,Xaa9选自H或F);或者其中Xaa9选自低可突变性的氨基酸(例如,Xaa9选自Y、F或L);或它们的任何组合。

[0475] 在一些实施方案中,Xaa1选自高亲疏水性的氨基酸。在一些实施方案中,Xaa1选自V、I或L。在一些实施方案中,Xaa1选自低可突变性的氨基酸。在一些实施方案中,Xaa1选自Y、L、F或C。在一些实施方案中,Xaa2选自低亲疏水性的氨基酸。在一些实施方案中,Xaa2选自Y或W。在一些实施方案中,Xaa2选自高摩尔质量的氨基酸。在一些实施方案中,Xaa2选自W。在一些实施方案中,Xaa2选自低表面可接近性的氨基酸。在一些实施方案中,Xaa2选自W或A。在一些实施方案中,Xaa2选自低亲水性的氨基酸。在一些实施方案中,Xaa2选自W。在一些实施方案中,Xaa2选自低可突变性的氨基酸。在一些实施方案中,Xaa2选自C。在一些实施方案中,Xaa2选自低平均柔性的氨基酸。在一些实施方案中,Xaa2选自W、M或F。在一些实施方案中,Xaa5选自低平均柔性的氨基酸。在一些实施方案中,Xaa5选自W、M或F。在一些实施方案中,Xaa6选自低平均柔性的氨基酸。在一些实施方案中,Xaa6选自W、M或F。在一些实施方案中,Xaa6选自低可突变性的氨基酸。在一些实施方案中,Xaa6选自Y、F、L或C。在一些实施方案中,Xaa6选自高溶解性的氨基酸。在一些实施方案中,Xaa6选自W、F、I或L。在一些实施方案中,Xaa7选自低表面可接近性的氨基酸。在一些实施方案中,Xaa7选自C。在一些实施方案中,Xaa7选自高表面可接近性的氨基酸。在一些实施方案中,Xaa7选自D或N。在一些实施方案中,Xaa7选自低可突变性的氨基酸。在一些实施方案中,Xaa7选自C。在一些实施方案中,Xaa7选自高溶解性的氨基酸。在一些实施方案中,Xaa7选自C。在一些实施方案中,Xaa7选自低溶解性的氨基酸。在一些实施方案中,Xaa7选自D。在一些实施方案中,Xaa8选自低电荷的氨基酸。在一些实施方案中,Xaa8选自D或E。在一些实施方案中,Xaa8选自高可突变性的氨基酸。在一些实施方案中,Xaa8选自D、E、A或T。在一些实施方案中,Xaa9选自高摩尔质量的氨基酸。在一些实施方案中,Xaa9选自H或F。在一些实施方案中,Xaa9选自低可突变性的氨基酸。在一些实施方案中,Xaa9选自Y、F或L。

[0476] 在一些实施方案中,本文提供了AAV5 VP衣壳多肽,所述多肽在从对应于AAV5 VP1的581的位置至对应于589的位置的区域中具有至少一个突变并且与选自SEQ ID NO:5118-SEQ ID NO:6117的任何序列具有至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%、至少99%或100%序列同一性,其中所述至少一个突变驱动增加的骨髓组织嗜性。

[0477] C.富集的骨髓序列

[0478] 在一些实施方案中,本文提供了AAV5 VP衣壳多肽,所述多肽在从对应于AAV5 VP1的581的位置至对应于589的位置的区域中具有至少一个突变并且与选自SEQ ID NO:6118-SEQ ID NO:7117的任何序列具有至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%、至少99%或100%序列同一性,其中所述至少一个突变驱动增加的骨髓组织嗜性。

[0479] 6.7.20.体内选定的赋予增加的骨骼肌嗜性或心肌嗜性的突变VP多肽

[0480] A.ML规则

[0481] 对于本段中描述的以下一组规则,使用体内数据和两个ML模型(描述于实施例22)确定工程化AAV5 VP1衣壳多肽的581至589区域中每个位置处的有利生物物理特性和有利氨基酸残基。本文公开了能够形成组装的病毒体的工程化AAV5 VP衣壳多肽,与具有野生型AAV5 VP衣壳多肽的rAAV病毒体相比,所述组装的病毒体表现出增加的骨骼肌组织嗜性或心肌组织嗜性,其中工程化AAV5 VP衣壳多肽序列具有一个或多个突变,其中工程化AAV5

VP1衣壳多肽序列在从对应于SEQ ID NO:2中的581的位置至对应于SEQ ID NO:2中的589的位置的区域中具有所述一个或多个突变,并且其中Xaa1选自低溶解性的氨基酸(例如,Xaa1选自D、E、R、K、P、N或Q);或者其中Xaa1选自低亲疏水性的氨基酸(例如,Xaa1选自D、E、R、K、Q、N、Y或P);或者其中Xaa1选自高表面可接近性的氨基酸(例如,Xaa1选自E、R或K);或者其中Xaa2选自高亲疏水性的氨基酸(例如,Xaa2选自V、I、F、L或C);或者其中Xaa2选自低可突变性的氨基酸(例如,Xaa2选自R、V、I、H或C);或者其中Xaa2选自中等体积的氨基酸(例如,Xaa2选自E、V或Q);或者其中Xaa3选自低溶解性的氨基酸(例如,Xaa3选自D、R或Q);或者其中Xaa4选自低溶解性的氨基酸(例如,Xaa4选自D、E、P或N);或者其中Xaa4选自低电荷的氨基酸(例如,Xaa4选自D或E);或者其中Xaa5选自低氨基酸溶解性的氨基酸(例如,Xaa5选自D、E、R、K、N或Q);或者其中Xaa8选自低溶解性的氨基酸(例如,Xaa8选自D、E、K、P或N);或者其中Xaa8选自高柔性指数的氨基酸(例如,Xaa8选自Q、S、P、E或D);或者其中Xaa8选自高表面可接近性的氨基酸(例如,Xaa8选自S、D、P、N、E、R或K);或它们的任何组合。

[0482] 在一些实施方案中,Xaa1选自低溶解性的氨基酸。在一些实施方案中,Xaa1选自D、E、R、K、P、N或Q。在一些实施方案中,Xaa1选自低亲疏水性的氨基酸。在一些实施方案中,Xaa1选自D、E、R、K、Q、N、Y或P。在一些实施方案中,Xaa1选自高表面可接近性的氨基酸。在一些实施方案中,Xaa1选自E、R或K。在一些实施方案中,Xaa2选自高亲疏水性的氨基酸。在一些实施方案中,Xaa2选自V、I、F、L或C。在一些实施方案中,Xaa2选自低可突变性的氨基酸。在一些实施方案中,Xaa2选自R、V、I、H或C。在一些实施方案中,Xaa2选自中等体积的氨基酸。在一些实施方案中,Xaa2选自E、V或Q。在一些实施方案中,Xaa3选自低溶解性的氨基酸。在一些实施方案中,Xaa3选自D、R或Q。在一些实施方案中,Xaa4选自低溶解性的氨基酸。在一些实施方案中,Xaa4选自D、E、P或N。在一些实施方案中,Xaa4选自低电荷的氨基酸。在一些实施方案中,Xaa4选自D或E。在一些实施方案中,Xaa5选自低氨基酸溶解性的氨基酸。在一些实施方案中,Xaa5选自D、E、R、K、N或Q。在一些实施方案中,Xaa8选自低溶解性的氨基酸。在一些实施方案中,Xaa8选自D、E、K、P或N。在一些实施方案中,Xaa8选自高柔性指数的氨基酸。在一些实施方案中,Xaa8选自Q、S、P、E或D。在一些实施方案中,Xaa8选自高表面可接近性的氨基酸。在一些实施方案中,Xaa8选自S、D、P、N、E、R或K。

[0483] 在一些实施方案中,本文提供了AAV5 VP衣壳多肽,所述多肽在从对应于AAV5 VP1的581的位置至对应于589的位置的区域中具有至少一个突变并且与选自SEQ ID NO:25118-SEQ ID NO:26117的任何序列具有至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%、至少99%或100%序列同一性,其中所述至少一个突变驱动增加的骨骼肌组织嗜性和心肌组织嗜性。

[0484] 1.1.编号的实施方案

[0485] 本文公开了许多组合物和方法。这些组合物和方法的具体实例性实施方案在下文公开。以下实施方案列举了本文公开的特征的组合的非限制性排列。还设想了特征组合的其他排列。特别地,这些编号的实施方案中的每一者被设想为依赖于每个先前或随后的编号实施方案或与之相关,而与它们所列出的顺序无关。

[0486] 在另一方面,提供了以下实施方案。对前述实施方案的所有数字参考是指在同一小节内如此编号的实施方案。在又一方面,提供了包含以下编号实施方案的重组或工程化VP衣壳多肽的rAAV,以及使用包含所述rAAV的药物组合物治疗有需要的受试者的方法。

[0487] A系列实施方案

[0488] 在A系列实施方案中,“重组”腺相关(AAV)VP1衣壳多肽与“工程化”腺相关(AAV)VP1衣壳多肽同义。

[0489] 1.一种具有SEQ ID NO:2的氨基酸序列的重组腺相关病毒(AAV)VP1衣壳多肽,其中Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9各自独立地选自A、R、N、D、C、E、Q、G、H、I、L、K、M、F、P、S、T、W、Y和V;并且其中所述多肽不具有SEQ ID NO:1、SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7和SEQ ID NO:8中任一者的序列。2.一种重组腺相关病毒(AAV)VP1衣壳多肽,所述多肽在对应于SEQ ID NO:1中残基581至残基589的残基中具有至少一个突变,其中所述突变赋予相较于第二组织对第一组织的组织嗜性,并且其中AAV VP1衣壳多肽不具有SEQ ID NO:1、SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7和SEQ ID NO:8中任一者的序列。3.如实施方案2所述的重组AAV VP1衣壳多肽,其中所述AAV VP1衣壳多肽是AAV5 VP1衣壳多肽。4.如实施方案1-3中任一项所述的重组AAV VP1衣壳多肽,其中对应于SEQ ID NO:1的残基581-589处的残基的特定顺序导致特定的组织嗜性。5.如实施方案1-4中任一项所述的重组AAV VP1衣壳多肽,其中所述第一组织选自脂肪、肾上腺、主动脉、脑(包括海马:齿状回、CA1和CA3;小脑、尾状核、壳核、中脑、脑桥、下丘脑、皮层-包括枕叶、颞叶和前脑;黑质和丘脑)、骨髓、盲肠、结肠、背根神经节、十二指肠、附睾、食管、眼睛、胆囊、心脏、回肠、空肠、肾脏、肺、淋巴结、乳腺、卵巢、胰腺、甲状旁腺、周围神经、垂体、前列腺、唾液腺、精囊、骨骼肌、皮肤、脊髓、脾脏、胃、睾丸、胸腺、甲状腺、气管、膀胱、子宫和阴道;其中所述第二组织选自:脂肪、肾上腺、主动脉、脑(包括海马:齿状回、CA1和CA3;小脑、尾状核、壳核、中脑、脑桥、下丘脑、皮层-包括枕叶、颞叶和前脑;黑质和丘脑)、骨髓、盲肠、结肠、背根神经节、十二指肠、附睾、食管、眼睛、胆囊、心脏、回肠、空肠、肾脏、肺、淋巴结、乳腺、卵巢、胰腺、甲状旁腺、周围神经、垂体、前列腺、唾液腺、精囊、骨骼肌、皮肤、脊髓、脾脏、胃、睾丸、胸腺、甲状腺、气管、膀胱、子宫和阴道;并且其中所述第一组织和所述第二组织是不同的。6.如实施方案2-4中任一项所述的重组AAV VP1衣壳多肽,其中与具有SEQ ID NO:1的VP1衣壳多肽的rAAV相比,所述rAAV在静脉内施用后具有增强的感染选自以下的组织的能力:脂肪、肾上腺、主动脉、脑(包括海马:齿状回、CA1和CA3;小脑、尾状核、壳核、中脑、脑桥、下丘脑、皮层-包括枕叶、颞叶和前脑;黑质和丘脑)、骨髓、盲肠、结肠、背根神经节、十二指肠、附睾、食管、眼睛、胆囊、心脏、回肠、空肠、肾脏、肺、淋巴结、乳腺、卵巢、胰腺、甲状旁腺、周围神经、垂体、前列腺、唾液腺、精囊、骨骼肌、皮肤、脊髓、脾脏、胃、睾丸、胸腺、甲状腺、气管、膀胱、子宫和阴道。7.如实施方案2-6中任一项所述的重组AAV VP1衣壳多肽,其中所述rAAV表现出在第一组织中相较于第二组织约1.0005倍至约1000倍增加的积聚。8.如实施方案2-6中任一项所述的重组AAV VP1衣壳多肽,其中所述rAAV表现出在第一组织中相较于第二组织至少约1.0005倍、至少约两倍、至少约三倍、至少约四倍、至少约五倍、至少约十倍、至少约二十倍、至少约50倍、至少约75倍、至少约100倍、至少约1000倍增加的积聚。9.如前述实施方案中任一项所述的重组衣壳多肽,所述重组衣壳多肽在581-589区域之外的氨基酸残基处还包含一个或多个突变,其中在581-589区域之外的氨基酸残基处的所述一个或多个突变赋予改进的可制造性、改进的病毒组装、提高的组织靶向性/嗜性,或它们的任何组合。10.如实施方案1-9中任一项所述的重组AAV VP1衣壳多肽,其中Xaa1选自A、G、K、M、N、Q、R、S或T。11.如前述实施方案中任一项

所述的重组衣壳多肽，其中Xaa1选自A、K、M或T。12.如前述实施方案中任一项所述的重组衣壳多肽，其中Xaa1为K。13.如前述实施方案中任一项所述的重组衣壳多肽，其中Xaa2选自A、C、H、I、K、S、T或V。14.如前述实施方案中任一项所述的重组衣壳多肽，其中Xaa2选自A、S、T或V。15.如前述实施方案中任一项所述的重组衣壳多肽，其中Xaa2为T。16.如前述实施方案中任一项所述的重组衣壳多肽，其中Xaa3选自A、G、H、K、M、N、Q、R、S、T或V。17.如前述实施方案中任一项所述的重组衣壳多肽，其中Xaa3选自A、M或T。18.如前述实施方案中任一项所述的重组衣壳多肽，其中Xaa3包括A或T。19.如前述实施方案中任一项所述的重组衣壳多肽，其中Xaa4选自L、M、P、Q、R、T或W。20.如前述实施方案中任一项所述的重组衣壳多肽，其中Xaa4选自L、P、Q或T。21.如前述实施方案中任一项所述的重组衣壳多肽，其中Xaa4为P；22.如前述实施方案中任一项所述的重组衣壳多肽，其中Xaa5选自F、H、I、K、M、T或Y。23.如前述实施方案中任一项所述的重组衣壳多肽，其中Xaa5选自H、I或Y。24.如前述实施方案中任一项所述的重组衣壳多肽，其中Xaa5为Y。25.如前述实施方案中任一项所述的重组衣壳多肽，其中Xaa6选自E、G、H、L、M、N、Q、T或W。26.如前述实施方案中任一项所述的重组衣壳多肽，其中Xaa6选自N或Q。27.如前述实施方案中任一项所述的重组衣壳多肽，其中Xaa6为N。28.如前述实施方案中任一项所述的重组衣壳多肽，其中Xaa7选自A、C、G、H、L、M、R或S。29.如前述实施方案中任一项所述的重组衣壳多肽，其中Xaa7选自A、C、H或M。30.如前述实施方案中任一项所述的重组衣壳多肽，其中Xaa7为A。31.如前述实施方案中任一项所述的重组衣壳多肽，其中Xaa8选自A、C、D、F、G、H、M、Q、S、V、W或Y。32.如前述实施方案中任一项所述的重组衣壳多肽，其中Xaa8选自G、M、Q或S。33.如前述实施方案中任一项所述的重组衣壳多肽，其中Xaa8为G。34.如前述实施方案中任一项所述的重组衣壳多肽，其中Xaa9选自A、C、E、G、H、M、N、P、Q、S、V或W。35.如前述实施方案中任一项所述的重组衣壳多肽，其中Xaa9选自E、G或P。36.如前述实施方案中任一项所述的重组衣壳多肽，其中Xaa9为G。37.如实施方案1-9中任一项所述的重组衣壳多肽，其中Xaa1选自A、D、E、G、L、M、N、Q、S、T或V。38.如实施方案37所述的重组衣壳多肽，其中Xaa1选自A、D、E、M或T。39.如实施方案37所述的重组衣壳多肽，其中Xaa1为E。40.如实施方案37-39中任一项所述的重组衣壳多肽，其中Xaa2选自A、C、D、E、G、H、I、N、P、Q、S、T或V。41.如实施方案37-40中任一项所述的重组衣壳多肽，其中Xaa2选自A、S、T或V。42.如实施方案37-41中任一项所述的重组衣壳多肽，其中Xaa2为A。43.如实施方案37-42中任一项所述的重组衣壳多肽，其中Xaa3选自A、D、E、G、H、M、N、Q、S、T或V。44.如实施方案37-43中任一项所述的重组衣壳多肽，其中Xaa3选自D、E、N、Q或T。45如实施方案37-44中任一项所述的重组衣壳多肽，其中Xaa3为D或T。46.如实施方案37-45中任一项所述的重组衣壳多肽，其中Xaa4选自A、D、E、G、H、N、P、Q、S或T。47.如实施方案37-46中任一项所述的重组衣壳多肽，其中Xaa4选自D、E、P或Q。48.如实施方案37-47中任一项所述的重组衣壳多肽，其中Xaa4为E。49.如实施方案37-48中任一项所述的重组衣壳多肽，其中Xaa5选自A、C、D、E、G、H、N、Q、S、T或Y。50.如实施方案37-49中任一项所述的重组衣壳多肽，其中Xaa5选自D、E、N、Q或T。51.如实施方案37-50中任一项所述的重组衣壳多肽，其中Xaa5为N。52.如实施方案37-51中任一项所述的重组衣壳多肽，其中Xaa6选自A、D、E、G、H、N、P、Q、S或T。53.如实施方案37-52中任一项所述的重组衣壳多肽，其中Xaa6选自D、N或Q。54.如实施方案37-53中任一项所述的重组衣壳多肽，其中Xaa6为D。55.如实施方案37-54中任一项所述的重组衣壳多肽，其中Xaa7选自A、C、D、E、G、H、N、Q、S或T。56.如实施方案37-55中任一项所述的重组衣壳多

肽，其中Xaa7选自A、D、E或G。57.如实施方案37-56中任一项所述的重组衣壳多肽，其中Xaa7为A。58.如实施方案37-57中任一项所述的重组衣壳多肽，其中Xaa8选自A、C、D、E、G、H、N、Q、S或T。59.如实施方案37-58中任一项所述的重组衣壳多肽，其中Xaa8包括A、D、G或S。60.如实施方案37-59中任一项所述的重组衣壳多肽，其中Xaa8为G。61.如实施方案37-60中任一项所述的重组衣壳多肽，其中Xaa9选自A、D、E、G、H、N、P、Q、S或T。62.如实施方案37-61中任一项所述的重组衣壳多肽，其中Xaa9选自A、D、G或P。63.如实施方案37-62中任一项所述的重组衣壳多肽，其中Xaa9为G。64.一种实施方案1-36中任一项所述的重组衣壳多肽与实施方案37-63中任一项所述的重组衣壳多肽组合的重组衣壳多肽，其中所述VP1衣壳能够形成组装的病毒体，与包含SEQ ID NO:1的AAV5VP1衣壳多肽的病毒体相比，所述组装的病毒体表现出增加的对肝脏组织的嗜性。65.如实施方案1所述的重组衣壳多肽，其中Xaa1不是K，并且其中所述VP1衣壳能够形成组装的病毒体，与包含SEQ ID NO:1的AAV5 VP1衣壳的病毒体相比，所述组装的病毒体表现出降低的对肝脏组织的嗜性。66.如实施方案1所述的重组衣壳多肽，其中Xaa1不是A、K、M或T，并且其中所述VP1衣壳能够形成组装的病毒体，与包含SEQ ID NO:1的AAV5 VP1衣壳的病毒体相比，所述组装的病毒体表现出降低的对肝脏组织的嗜性。67.如实施方案1所述的重组衣壳多肽，其中Xaa1不是A、G、K、M、N、Q、R、S或T，并且其中所述VP1衣壳能够形成组装的病毒体，与包含SEQ ID NO:1的AAV5 VP1衣壳的病毒体相比，所述组装的病毒体表现出降低的对肝脏组织的嗜性。68.如实施方案1或实施方案65-67中任一项所述的重组衣壳多肽，其中Xaa2不是T，并且其中所述VP1衣壳能够形成组装的病毒体，与包含SEQ ID NO:1的AAV5 VP1衣壳的病毒体相比，所述组装的病毒体表现出降低的对肝脏组织的嗜性。69.如实施方案1或实施方案65-68中任一项所述的重组衣壳多肽，其中Xaa2不是A、S、T或V，并且其中所述VP1衣壳能够形成组装的病毒体，与包含SEQ ID NO:1的AAV5 VP1衣壳的病毒体相比，所述组装的病毒体表现出降低的对肝脏组织的嗜性。70.如实施方案1或实施方案65-69中任一项所述的重组衣壳多肽，其中Xaa2不是A、C、H、I、K、S、T或V，并且其中所述VP1衣壳能够形成组装的病毒体，与包含SEQ ID NO:1的AAV5 VP1衣壳的病毒体相比，所述组装的病毒体表现出降低的对肝脏组织的嗜性。71.如实施方案1或实施方案65-70中任一项所述的重组衣壳多肽，其中Xaa3不是A或T，并且其中所述VP1衣壳能够形成组装的病毒体，与包含SEQ ID NO:1的AAV5 VP1衣壳的病毒体相比，所述组装的病毒体表现出降低的对肝脏组织的嗜性。72.如实施方案1或实施方案65-71中任一项所述的重组衣壳多肽，Xaa3不是A、M或T，并且其中所述VP1衣壳能够形成组装的病毒体，与包含SEQ ID NO:1的AAV5 VP1衣壳的病毒体相比，所述组装的病毒体表现出降低的对肝脏组织的嗜性。73.如实施方案1或实施方案65-72中任一项所述的重组衣壳多肽，其中Xaa3不是A、G、H、K、M、N、Q、R、S、T或V，并且其中所述VP1衣壳能够形成组装的病毒体，与包含SEQ ID NO:1的AAV5 VP1衣壳的病毒体相比，所述组装的病毒体表现出降低的对肝脏组织的嗜性。74.如实施方案1或实施方案65-73中任一项所述的重组衣壳多肽，其中Xaa4不是P，并且其中所述VP1衣壳能够形成组装的病毒体，与包含SEQ ID NO:1的AAV5 VP1衣壳的病毒体相比，所述组装的病毒体表现出降低的对肝脏组织的嗜性。75.如实施方案1或实施方案65-74中任一项所述的重组衣壳多肽，其中Xaa4不是L、P、Q或T，并且其中所述VP1衣壳能够形成组装的病毒体，与包含SEQ ID NO:1的AAV5 VP1衣壳的病毒体相比，所述组装的病毒体表现出降低的对肝脏组织的嗜性。76.如实施方案1或实施方案65-75中任一项所述的重组衣壳多肽，其中

Xaa4不是L、M、P、Q、R、T或W，并且其中所述VP1衣壳能够形成组装的病毒体，与包含SEQ ID NO:1的AAV5 VP1衣壳的病毒体相比，所述组装的病毒体表现出降低的对肝脏组织的嗜性。77.如实施方案1或实施方案65-76中任一项所述的重组衣壳多肽，其中Xaa5不是Y，并且其中所述VP1衣壳能够形成组装的病毒体，与包含SEQ ID NO:1的AAV5 VP1衣壳的病毒体相比，所述组装的病毒体表现出降低的对肝脏组织的嗜性。78.如实施方案1或实施方案65-77中任一项所述的重组衣壳多肽，其中Xaa5不是H、I或Y，并且其中所述VP1衣壳能够形成组装的病毒体，与包含SEQ ID NO:1的AAV5 VP1衣壳的病毒体相比，所述组装的病毒体表现出降低的对肝脏组织的嗜性。79.如实施方案1或实施方案65-78中任一项所述的重组衣壳多肽，其中Xaa5不是F、H、I、K、M、T或Y，并且其中所述VP1衣壳能够形成组装的病毒体，与包含SEQ ID NO:1的AAV5 VP1衣壳的病毒体相比，所述组装的病毒体表现出降低的对肝脏组织的嗜性。80.如实施方案1或实施方案65-79中任一项所述的重组衣壳多肽，其中Xaa6不是N，并且其中所述VP1衣壳能够形成组装的病毒体，与包含SEQ ID NO:1的AAV5 VP1衣壳的病毒体相比，所述组装的病毒体表现出降低的对肝脏组织的嗜性。81.如实施方案1或实施方案65-80中任一项所述的重组衣壳多肽，其中Xaa6不是N或Q，并且其中所述VP1衣壳能够形成组装的病毒体，与包含SEQ ID NO:1的AAV5 VP1衣壳的病毒体相比，所述组装的病毒体表现出降低的对肝脏组织的嗜性。82.如实施方案1或实施方案65-81中任一项所述的重组衣壳多肽，其中Xaa6不是E、G、H、L、M、N、Q、T或W，并且其中所述VP1衣壳能够形成组装的病毒体，与包含SEQ ID NO:1的AAV5 VP1衣壳的病毒体相比，所述组装的病毒体表现出降低的对肝脏组织的嗜性。83.如实施方案1或实施方案65-82中任一项所述的重组衣壳多肽，其中Xaa7不是A，并且其中所述VP1衣壳能够形成组装的病毒体，与包含SEQ ID NO:1的AAV5 VP1衣壳的病毒体相比，所述组装的病毒体表现出降低的对肝脏组织的嗜性。84.如实施方案1或实施方案65-83中任一项所述的重组衣壳多肽，其中Xaa7不是A、C、H或M，并且其中所述VP1衣壳能够形成组装的病毒体，与包含SEQ ID NO:1的AAV5 VP1衣壳的病毒体相比，所述组装的病毒体表现出降低的对肝脏组织的嗜性。85.如实施方案1或实施方案65-84中任一项所述的重组衣壳多肽，其中Xaa7不是A、C、G、H、L、M、R或S，并且其中所述VP1衣壳能够形成组装的病毒体，与包含SEQ ID NO:1的AAV5 VP1衣壳的病毒体相比，所述组装的病毒体表现出降低的对肝脏组织的嗜性。86.如实施方案1或实施方案65-85中任一项所述的重组衣壳多肽，其中Xaa8不是G，并且其中所述VP1衣壳能够形成组装的病毒体，与包含SEQ ID NO:1的AAV5 VP1衣壳的病毒体相比，所述组装的病毒体表现出降低的对肝脏组织的嗜性。87.如实施方案1或实施方案65-86中任一项所述的重组衣壳多肽，其中Xaa8不是G、M、Q或S，并且其中所述VP1衣壳能够形成组装的病毒体，与包含SEQ ID NO:1的AAV5 VP1衣壳的病毒体相比，所述组装的病毒体表现出降低的对肝脏组织的嗜性。88.如实施方案1或实施方案65-87中任一项所述的重组衣壳多肽，其中Xaa8不是A、C、D、F、G、H、M、Q、S、V、W或Y，并且其中所述VP1衣壳能够形成组装的病毒体，与包含SEQ ID NO:1的AAV5 VP1衣壳的病毒体相比，所述组装的病毒体表现出降低的对肝脏组织的嗜性。89.如实施方案1或实施方案65-88中任一项所述的重组衣壳多肽，Xaa9不是G，并且其中所述VP1衣壳能够形成组装的病毒体，与包含SEQ ID NO:1的AAV5 VP1衣壳的病毒体相比，所述组装的病毒体表现出降低的对肝脏组织的嗜性。90.如实施方案1或实施方案65-89中任一项所述的重组衣壳多肽，其中Xaa9不是E、G或P，并且其中所述VP1衣壳能够形成组装的病毒体，与包含SEQ ID NO:1的AAV5 VP1衣壳的病毒体

相比，所述组装的病毒体表现出降低的对肝脏组织的嗜性。91.如实施方案1或实施方案65-90中任一项所述的重组衣壳多肽，其中Xaa9不是A、C、E、G、H、M、N、P、Q、S、V或W，并且其中所述VP1衣壳能够形成组装的病毒体，与包含SEQ ID NO:1的AAV5 VP1衣壳的病毒体相比，所述组装的病毒体表现出降低的对肝脏组织的嗜性。92.一种实施方案65-91中任一项所述的重组衣壳多肽与实施方案37-63中任一项所述的重组衣壳多肽组合的重组衣壳多肽，其中所述VP1衣壳能够形成组装的病毒体，与包含SEQ ID NO:1的AAV5 VP1衣壳多肽的病毒体相比，所述组装的病毒体表现出降低的对肝脏组织的嗜性。93.一种重组腺相关病毒(AAV)VP1衣壳多肽，所述多肽具有对应于SEQ ID NO:2中的残基581至残基589的至少一个残基，其中所述至少一个残基为:Xaa1，并且Xaa1选自A、G、K、M、N、Q、R、S或T;Xaa2，并且Xaa2选自A、C、H、I、K、S、T或V;Xaa3，并且Xaa3选自A、G、H、K、M、N、Q、R、S、T或V;Xaa4，并且Xaa4选自L、M、P、Q、R、T或W;Xaa5，并且Xaa5选自F、H、I、K、M、T或Y;Xaa6，并且Xaa6选自E、G、H、L、M、N、Q、T或W;Xaa7，并且Xaa7选自A、C、G、H、L、M、R或S;Xaa8，并且Xaa8选自A、C、D、F、G、H、M、Q、S、V、W或Y;Xaa9，并且Xaa9选自A、C、E、G、H、M、N、P、Q、S、V或W;或它们的任何组合，其中所述AAV VP1衣壳多肽能够表现出对肝脏组织的组织嗜性。94.一种重组腺相关病毒AAV VP1衣壳多肽，所述多肽在SEQ ID NO:1中的581至残基589区域的残基中具有至少一个突变，其中与AAV5VP1相比，所述突变赋予在非肝脏组织中相较于肝脏组织中至少约两倍增加的积聚，并且其中所述AAV VP1衣壳多肽不具有SEQ ID NO:1、SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7和SEQ ID NO:8中任一者的序列。95.如实施方案94所述的重组AAV VP1衣壳多肽，其中所述突变赋予在非肝脏组织中相较于肝脏组织中至少约三倍、至少约四倍、至少约五倍、至少约十倍、至少约二十倍、至少约五十倍、至少约75倍、至少约100倍增加的积聚。96.如实施方案94所述的重组AAV VP1衣壳多肽，其中所述突变赋予在非肝脏组织中相较于肝脏组织中约1.0005倍至约1000倍增加的积聚。97.如前述实施方案中任一项所述的重组衣壳多肽，所述重组衣壳多肽相对于SEQ ID NO:1在581-589区域之外的氨基酸残基处还包含一个或多个突变，其中所得重组衣壳能够形成组装的病毒体，所述组装的病毒体表现出所需的组织靶向性/嗜性。98.如实施方案97所述的重组衣壳多肽，其中在581-589区域之外的氨基酸残基处的所述一个或多个突变赋予改进的可制造性、改进的病毒组装、提高的组织靶向性/嗜性，或它们的任何组合。99.一种能够在原核细胞中复制的载体，其中所述载体包含编码前述实施方案中任一项所述的重组衣壳多肽的多核苷酸。100.如实施方案99所述的载体，其中所述载体是质粒。101.一种包含多个实施方案100所述的质粒的文库，所述多个质粒包含多个不同的编码AAV VP1的多核苷酸。102.如实施方案101所述的质粒文库，其中所述文库编码至少1x109个不同的AAV VP1衣壳多肽。103.如实施方案102所述的质粒文库，其中所述文库编码至少5x109个不同的AAV VP1衣壳多肽。104.如实施方案103所述的质粒文库，其中所述文库编码至少1x1010个不同的AAV VP1衣壳多肽。105.如实施方案104所述的质粒文库，其中所述文库编码至少5x1 010个不同的AAV VP1衣壳多肽。106.如实施方案105所述的质粒文库，其中所述文库编码至少7.5x1010个不同的AAV VP1衣壳多肽。107.如实施方案106所述的质粒文库，其中所述文库编码至少1x1 011个不同的AAV VP1衣壳多肽。108.如实施方案107所述的质粒文库，其中所述文库编码至少2.5x1011个不同的AAV VP1衣壳多肽。109.如实施方案108所述的质粒文库，其中所述文库编码至少5x1 011个不同的AAV VP1衣壳多肽。110.一种包含实施方案100所述的载体的原核细胞。111.如实施

方案110所述的原核细胞，其中所述原核细胞是大肠杆菌细胞并且所述载体是质粒。112.一种包含多个实施方案111所述的大肠杆菌细胞的文库，其中所述多个细胞包含多个质粒，其中所述多个质粒包含多个不同的编码AAV VP1的多核苷酸。113.一种包含多个实施方案1-98中任一项所述的多肽的文库，所述多个多肽具有不同的一级氨基酸序列。114.实施方案113的文库，其中所述文库包含至少约1×105至至少约5×1011个不同的AAV VP1衣壳多肽。115.一种重组AAV病毒体(rAAV)，所述病毒体包含实施方案1-98中任一项所述的AAV VP1衣壳多肽。116.如实施方案115所述的rAAV病毒体，其中与具有含SEQ ID NO:1序列的VP1衣壳多肽的rAAV相比，所述rAAV具有降低的对人肝脏的嗜性。117.如实施方案115或实施方案116所述的rAAV病毒体，其中与具有含SEQ ID NO:1序列的VP1衣壳多肽的rAAV相比，所述rAAV在静脉内施用后具有增强的穿过血脑屏障的能力。118.如实施方案115-117中任一项所述的rAAV病毒体，其中与具有含SEQ ID NO:1序列的VP1衣壳多肽的rAAV相比，所述rAAV在静脉内、鞘内、脑室内或脑池内施用后具有增强的感染选自海马、齿状回、大脑皮层、颞叶皮层、枕叶皮层、丘脑、前脑、黑质、下丘脑和小脑的一个或多个脑区的能力。119.如实施方案115-118中任一项所述的rAAV病毒体，其中与具有含SEQ ID NO:1序列的VP1衣壳多肽的rAAV相比，所述rAAV在玻璃体内注射后具有增强的感染人视网膜细胞的能力。120.如实施方案115-119中任一项所述的rAAV病毒体，其中与具有含SEQ ID NO:1序列的VP1衣壳多肽的rAAV相比，所述rAAV在静脉内施用后具有增强的感染人骨骼肌的能力。121.如实施方案115和117-120中任一项所述的rAAV病毒体，其中与具有含SEQ ID NO:1序列的VP1衣壳多肽的rAAV相比，所述rAAV具有增加的对人肝脏的嗜性。122.如实施方案115、116或121所述的rAAV病毒体，其中与具有SEQ ID NO:1的VP1衣壳多肽的rAAV相比，所述rAAV在静脉内施用后具有增强的感染选自以下的组织的能力:脂肪、肾上腺、主动脉、脑(包括海马:齿状回、CA1和CA3;小脑、尾状核、壳核、中脑、脑桥、下丘脑、皮层-包括枕叶、颞叶和前脑;黑质和丘脑)、骨髓、盲肠、结肠、背根神经节、十二指肠、附睾、食管、眼睛、胆囊、心脏、回肠、空肠、肾脏、肺、淋巴结、乳腺、卵巢、胰腺、甲状旁腺、周围神经、垂体、前列腺、唾液腺、精囊、骨骼肌、皮肤、脊髓、脾脏、胃、睾丸、胸腺、甲状腺、气管、膀胱、子宫和阴道。123.如实施方案115-122中任一项所述的rAAV病毒体，其中所述病毒体还包含载体基因组，所述载体基因组包含编码以下任一项的治疗性多核苷酸:选自指导RNA或tRNA的治疗性RNA，或编码处于引导受感染人细胞中转基因表达的调控序列控制下的蛋白质。124.如实施方案123所述的rAAV病毒体，其中所述转基因编码选自表1的转基因产物的蛋白质。125一种包含多个实施方案115-124中任一项所述的rAAV病毒体的文库，其中所述多个rAAV病毒体包含多个具有不同一级氨基酸序列的VP1衣壳多肽。126.如实施方案125所述的文库，其中所述文库包含至少约$1x10^{5}$至至少约$5x10^{11}$个不同的AAV VP1衣壳多肽不同的AAV VP1衣壳多肽。127.一种药物组合物，所述药物组合物包含实施方案123或实施方案124所述的rAAV和药学上可接受的载体。128.一种治疗方法，所述方法包括:向有需要的患者施用有效量的实施方案127所述的药物组合物。129.如实施方案128所述的方法，其中所述rAAV的所述有效量小于野生型rAAV的有效量。130.如实施方案128所述的方法，其中所述rAAV的所述有效量小于在对应于SEQ ID NO:1的残基581至残基589的位置处缺乏一个或多于一个突变的其他方面相当的rAAV的有效量。131.如实施方案128-130中任一项所述的方法，其中与所述野生型rAAV、所述其他方面相当的rAAV或两者的所述有效量相比，所述的所述有效量在所述患者中引起更低的毒

性。132.如实施方案128所述的方法，其中所述有效量为至少1x10^{5}个病毒基因组/kg患者体重至5x10^{14}个病毒基因组/kg。133.如实施方案128-132中任一项所述的方法，其中静脉内施用所述rAAV。134.如实施方案128-132中任一项所述的方法，其中鞘内施用所述rAAV。135.如实施方案128-132中任一项所述的方法，其中通过脑池内施用来施用所述rAAV。136.如实施方案128-132中任一项所述的方法，其中通过玻璃体内注射来施用所述rAAV。137.一种鉴定赋予对所需组织的嗜性的AAV VP1衣壳多肽的方法，所述方法包括：将实施方案101-109或112-144或125-126中任一项所述的文库的等分试样施用于非人灵长类动物；以及鉴定已感染所述所需组织的rAAV的AAV衣壳序列的序列。138.如实施方案137所述的方法，其中静脉内施用所述文库等分试样。139.如实施方案137所述的方法，其中鞘内施用所述文库等分试样。140.如实施方案137所述的方法，其中通过脑池内施用来施用所述文库等分试样。141.如实施方案137所述的方法，其中通过脑室内注射来施用所述文库等分试样。142.一种配制在病毒体中的实施方案123-124中任一项所述的治疗性多核苷酸的方法，所述方法包括：用编码实施方案1-98中任一项所述的重组衣壳多肽的质粒转染细胞以及用编码所述治疗性多核苷酸的质粒转染所述细胞，其中在转染后，所述细胞产生所述病毒体，所述病毒体内包装有所述治疗性多核苷酸。143.一种的组合物，所述组合物包含含有实施方案1-98中任一项所述的重组衣壳多肽的AAV病毒体，所述病毒体内包装有实施方案123-124中任一项所述的治疗性多核苷酸。144.一种重组AAV VP1衣壳多肽，所述重组AAV VP1衣壳多肽具有表8(SEQ ID NO:115-1114)中列举的VP1衣壳突变中的任一者，并且其中所述VP1衣壳多肽不具有SEQ ID NO:1、SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7和SEQ ID NO:8中任一者的序列。145.一种重组AAV VP1衣壳多肽，所述重组AAV VP1衣壳多肽具有表8(SEQ ID NO:115-1114)中列举的VP1衣壳突变中的任一者，其中所述突变赋予相较于第二组织对第一组织的组织嗜性，并且其中所述AAV VP1衣壳多肽不具有SEQ ID NO:1、SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7和SEQ ID NO:8中任一者的序列。146.如实施方案145所述的重组AAV VP1衣壳多肽，其中与包含AAV5 VP1(SEQ ID NO:1)的rAAV的积聚相比，所述突变赋予包含所述突变VP1蛋白的rAAV在非肝脏组织中相较于肝脏组织中至少约两倍增加的积聚，其中所述突变rAAV和所述AAV5rAAV各自以相同的滴度静脉内施用，并且其中所述VP1衣壳多肽不具有SEQ ID NO:1、SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7和SEQ ID NO:8中任一者的序列。147.一种组合物，所述组合物包含含有实施方案144-146中任一项所述的重组衣壳多肽的AAV病毒体，所述AAV病毒体内包装有编码以下任一项的治疗性多核苷酸：选自指导RNA或tRNA的治疗性RNA，或编码处于引导受感染人细胞中转基因表达的调控序列控制下的蛋白质。

[0490] [403]B系列实施方案-CNS嗜性衣壳

[0491] 1.一种工程化腺相关病毒(AAV)病毒蛋白(VP)衣壳多肽，所述工程化腺相关病毒(AAV)病毒蛋白(VP)衣壳多肽具有与SEQ ID NO:1至少70％同一的氨基酸序列，

[0492] 其中所述工程化AAV VP衣壳多肽相较于SEQ ID NO:1在从对应于SEQ ID NO:1的残基581的残基至对应于SEQ ID NO:1的残基589的残基的区域，包括端点在内，中具有至少一个突变，

[0493] 其中所述工程化AAV VP衣壳多肽能够组装成重组AAV病毒体(rAAV)，

[0494] 其中与具有SEQ ID NO:1的野生型AAV5 VP衣壳多肽的rAAV病毒体相比，所述至少

一个突变赋予所述rAAV较高的对中枢神经系统(CNS)组织的嗜性,并且

[0495] 其中所述工程化AAV VP衣壳多肽不具有SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7和SEQ ID NO:8中任一者的序列。2.如实施方案1所述的工程化AAV VP衣壳多肽,其中所述工程化AAV VP衣壳多肽具有SEQ ID NO:2的序列,并且其中Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9各自独立地选自任何氨基酸。3.一种工程化腺相关病毒(AAV)病毒蛋白(VP)衣壳多肽,所述工程化腺相关病毒(AAV)病毒蛋白(VP)衣壳多肽具有SEQ ID NO:2的氨基酸序列,

[0496] 其中氨基酸残基Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9各自独立地选自A、R、N、D、C、E、Q、G、H、I、L、K、M、F、P、S、T、W、Y和V,

[0497] 其中所述工程化AAV VP衣壳多肽能够组装成重组AAV病毒体(rAAV),并且

[0498] 其中所述工程化AAV VP衣壳多肽不具有SEQ ID NO:1、SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7和SEQ ID NO:8中任一者的序列。4.如实施方案3所述的工程化AAV VP衣壳多肽,其中与具有SEQ ID NO:1的野生型AAV5 VP衣壳多肽的rAAV病毒体相比,所述rAAV具有较高的对中枢神经系统(CNS)组织的嗜性。5.如实施方案1-2中任一项所述的工程化AAV VP衣壳多肽,其中从对应于残基581的残基至对应于残基589的残基的区域,包括端点在内,具有与SEQ ID NO:7118-SEQ ID NO:10117中任一者至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%同一的序列。6.如实施方案4所述的工程化AAV VP衣壳多肽,其中从对应于残基581的残基至对应于残基589的残基的区域,包括端点在内,具有SEQ ID NO:7118-SEQ ID NO:10117中任一者的序列。7.如实施方案3所述的工程化AAV VP衣壳多肽,其中SEQ ID NO:2的Xaa1至Xaa9区具有与SEQ ID NO:7118-SEQ ID NO:10117中任一者至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%同一的序列。8.如实施方案6所述的工程化AAV VP衣壳多肽,其中SEQ ID NO:2的Xaa1至Xaa9区域具有SEQ ID NO:7118-SEQ ID NO:10117中任一者的序列。9.如实施方案2-3中任一项所述的工程化AAV VP衣壳多肽,其中:

[0499] Xaa1选自由A、C、K、M、Q、R、T和W组成的组;或者

[0500] Xaa2选自由F、I、K、R、T和W组成的组;或者

[0501] Xaa3选自由A、H、N、R和W组成的组;或者

[0502] Xaa4选自由E、G、I、M、Q和R组成的组;或者

[0503] Xaa5选自由C、G、K、I、M和R组成的组;或者

[0504] Xaa6选自由I、K、L、P、Q、R和Y组成的组;或者

[0505] Xaa7选自由D、I、K、R、V和W组成的组;或者

[0506] Xaa8选自由C、G、H、K、L和V组成的组;或者

[0507] Xaa9选自由I、K、L、R和V组成的组;或者

[0508] 它们的任何组合。

[0509] 10.如实施方案5所述的工程化AAV VP衣壳多肽,其中Xaa1选自K、Q、R或W。11.如实施方案5所述的工程化AAV VP衣壳多肽,其中Xaa2选自F、I、R或T。12.如实施方案5所述的工程化AAV VP衣壳多肽,其中Xaa3选自A、R或W。13.如实施方案5所述的工程化AAV VP衣壳多肽,其中Xaa4选自E、M或R。14.如实施方案5所述的工程化AAV VP衣壳多肽,其中Xaa5选自K、

I或R。15.如实施方案5所述的工程化AAV VP衣壳多肽，其中Xaa6选自K、R或Y。16.如实施方案5所述的工程化AAV VP衣壳多肽，其中Xaa7选自I、R或V。17.如实施方案5所述的工程化AAV VP衣壳多肽，其中Xaa8选自H、K或V。18.如实施方案5所述的工程化AAV VP衣壳多肽，其中Xaa9选自I、K或R。19.如实施方案5所述的工程化AAV VP衣壳多肽，其中Xaa1为K。20.如实施方案5所述的工程化AAV VP衣壳多肽，其中Xaa2为R。21.如实施方案5所述的工程化AAV VP衣壳多肽，其中Xaa3为R。22.如实施方案5所述的工程化AAV VP衣壳多肽，其中Xaa4为R。23.如实施方案5所述的工程化AAV VP衣壳多肽，其中Xaa5为I。24.如实施方案5所述的工程化AAV VP衣壳多肽，其中Xaa6为R。25.如实施方案5所述的工程化AAV VP衣壳多肽，其中Xaa7为V。26.如实施方案5所述的工程化AAV VP衣壳多肽，其中Xaa8为H。27.如实施方案5所述的工程化AAV VP衣壳多肽，其中Xaa9为R。28.如实施方案1-8中任一项所述的工程化AAV VP衣壳多肽，其中从对应于残基581的残基至对应于残基589的残基的区域，包括端点在内，具有与SEQ ID NO:7118-SEQ ID NO:8117中任一者至少70％、至少75％、至少80％、至少85％、至少90％、至少92％、至少95％、至少97％、至少98％或至少99％同一的序列。29.如实施方案27所述的工程化AAV VP衣壳多肽，其中SEQ ID NO:1的从残基581至残基589的区域具有SEQ ID NO:7118-SEQ ID NO:8117中任一者的序列。30.如实施方案2-3中任一项所述的工程化AAV VP衣壳多肽，其中Xaa1具有低氨基酸溶解性。31.如实施方案25所述的工程化AAV VP衣壳多肽，其中Xaa1选自K、R或Q。32.如实施方案2-3中任一项所述的工程化AAV VP衣壳多肽，其中Xaa1具有低氨基酸亲疏水性。33.如实施方案27所述的工程化AAV VP衣壳多肽，其中Xaa1选自K或R。34.如实施方案2-3中任一项所述的工程化AAV VP衣壳多肽，其中Xaa1具有高平均氨基酸柔性指数。35.如实施方案29所述的工程化AAV VP衣壳多肽，其中Xaa1选自D、E、R、K、G、I、N、Q或S。36.如实施方案2-3中任一项所述的工程化AAV VP衣壳多肽，其中Xaa1具有高氢键供体。37.如实施方案31所述的工程化AAV VP衣壳多肽，其中Xaa1选自K、R。38.如实施方案2-3中任一项所述的工程化AAV VP衣壳多肽，其中Xaa1具有低氨基酸可突变性。39.如实施方案33所述的工程化AAV VP衣壳多肽，其中Xaa1选自K、R、P或H。40.如实施方案2-3中任一项所述的工程化AAV VP衣壳多肽，其中Xaa2具有低氨基酸溶解性。41.如实施方案35所述的工程化AAV VP衣壳多肽，其中Xaa2选自R、K、Q或S。42.如实施方案2-3中任一项所述的工程化AAV VP衣壳多肽，其中Xaa2具有低氨基酸亲疏水性。43.如实施方案37所述的工程化AAV VP衣壳多肽，其中Xaa2选自R、K、D、E、N、Q、H、P、Y、W、S或T。44.如实施方案2-3中任一项所述的工程化AAV VP衣壳多肽，其中Xaa2具有高氨基酸电荷。45.如实施方案39所述的工程化AAV VP衣壳多肽，其中Xaa2选自R、K或H。46.如实施方案2-3中任一项所述的工程化AAV VP衣壳多肽，其中Xaa3具有高氨基酸溶解性。47.如实施方案41所述的工程化AAV VP衣壳多肽，其中Xaa3选自A、M、V、W、L或I。48.如实施方案2-3中任一项所述的工程化AAV VP衣壳多肽，其中Xaa5具有高氨基酸溶解性。49.如实施方案43所述的工程化AAV VP衣壳多肽，其中Xaa5选自C、M、V、W、L或I。50.如实施方案2-3中任一项所述的工程化AAV VP衣壳多肽，其中Xaa5具有高亲疏水性。51.如实施方案45所述的工程化AAV VP衣壳多肽，其中Xaa5选自M、V或I。52.如实施方案2-3中任一项所述的工程化AAV VP衣壳多肽，其中Xaa5具有低平均氨基酸柔性指数。53.如实施方案47所述的工程化AAV VP衣壳多肽，其中Xaa5选自M、W、F或C。54.如实施方案2-3中任一项所述的工程化AAV VP衣壳多肽，其中Xaa8具有高氨基酸溶解性。55.如实施方案49所述的工程化AAV VP衣壳多肽，其中Xaa8选自H、V

或I。56.如实施方案29-54中任一项所述的工程化AAV VP衣壳多肽，其中从对应于残基581的残基至对应于残基589的残基的区域包括端点在内具有与SEQ ID NO:8118-SEQ ID NO:9117中任一者具有至少70％、至少75％、至少80％、至少85％、至少90％、至少92％、至少95％、至少97％或至少99％同一的序列。57.如实施方案25-50中任一项所述的工程化AAV VP衣壳多肽，其中从对应于残基581的残基至对应于残基589的残基的区域包括端点在内具有SEQ ID NO:8118-SEQ ID NO:9117中任一者的序列。58.如实施方案1-56中任一项所述的工程化AAV VP衣壳多肽，其中所述CNS组织选自前脑皮层、枕叶皮层、颞叶皮层、丘脑、下丘脑、黑质、海马DG、海马CA1、海马CA3、小脑以及它们的任何组合。59.如实施方案1-57中任一项所述的工程化AAV VP衣壳多肽，其中对CNS组织的嗜性被测量为所述rAAV病毒体在CNS组织中相较于非CNS组织中的相对积聚，其中所述非CNS组织由肝脏、骨骼肌、心脏、肺、脾脏、淋巴结、骨髓、乳腺、皮肤、肾上腺、甲状腺、结肠、坐骨神经和脊髓共同组成。60.如实施方案1-58中任一项所述的工程化AAV VP衣壳多肽，其中所述较高的组织嗜性是在CNS组织中相较于非CNS组织中1.0005倍至约1000倍增加的积聚。61.如实施方案59所述的工程化AAV VP衣壳多肽，其中所述较高的组织嗜性是在CNS组织中相较于非CNS组织中至少约1.0005倍、至少约两倍、至少约三倍、至少约四倍、至少约五倍、至少约十倍、至少约二十倍、至少约50倍、至少约75倍、至少约100倍或至少约1000倍增加的积聚。

[0510] [404]C系列实施方案-肝脏去靶向性衣壳

[0511] 1.一种工程化腺相关病毒(AAV)病毒蛋白(VP)衣壳多肽，所述工程化腺相关病毒(AAV)病毒蛋白(VP)衣壳多肽具有与SEQ ID NO:1至少70％同一的氨基酸序列，

[0512] 其中所述工程化AAV VP衣壳多肽相较于SEQ ID NO:1在从对应于SEQ ID NO:1的残基581的残基至对应于SEQ ID NO:1的残基589的残基的区域，包括端点在内，中具有至少一个突变，

[0513] 其中所述工程化AAV VP衣壳多肽能够组装成重组AAV病毒体(rAAV)，

[0514] 其中与具有SEQ ID NO:1的野生型AAV5 VP衣壳多肽的rAAV病毒体相比，所述至少一个突变赋予所述rAAV较高的对非肝脏组织的嗜性，并且

[0515] 其中所述工程化AAV VP衣壳多肽不具有SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7和SEQ ID NO:8中任一者的序列。2.如实施方案1所述的工程化AAV VP衣壳多肽，其中所述工程化AAV VP衣壳多肽具有SEQ ID NO:2的序列，并且其中Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9各自独立地选自任何氨基酸。3.一种工程化腺相关病毒(AAV)病毒蛋白(VP)衣壳多肽，所述工程化腺相关病毒(AAV)病毒蛋白(VP)衣壳多肽具有SEQ ID NO:2的氨基酸序列，

[0516] 其中氨基酸残基Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9各自独立地选自A、R、N、D、C、E、Q、G、H、I、L、K、M、F、P、S、T、W、Y和V，

[0517] 其中所述工程化AAV VP衣壳多肽能够组装成重组AAV病毒体(rAAV)，并且

[0518] 其中所述工程化AAV VP衣壳多肽不具有SEQ ID NO:1、SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7和SEQ ID NO:8中任一者的序列。4.如实施方案3所述的工程化AAV VP衣壳多肽，其中与具有SEQ ID NO:1的野生型AAV5 VP衣壳多肽的rAAV病毒体相比，所述rAAV具有较高的对非肝脏组织的嗜性。5.如实施方案1所述的工程化AAV VP衣壳多肽，其中从对应于残基581的残基至对应于残基589的残基的区域，包括端点

在内，具有与SEQ ID NO:46438-SEQ ID NO:47437中任一者至少70％、至少75％、至少80％、至少85％、至少90％、至少92％、至少95％、至少97％、至少98％、或至少99％同一的序列。6.如实施方案5所述的工程化AAV VP衣壳多肽，其中从对应于残基581的残基至对应于残基589的残基的区域，包括端点在内，具有SEQ ID NO:46438-SEQ ID NO:47437中任一者的序列。7.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽，其中SEQ ID NO:2的Xaa1至Xaa9区的序列具有与SEQ ID NO:46438-SEQ ID NO:47437中任一者至少70％、至少75％、至少80％、至少85％、至少90％、至少92％、至少95％、至少97％或至少99％同一的序列。8.如实施方案7所述的工程化AAV VP衣壳多肽，其中SEQ ID NO:2的Xaa1至Xaa9区域具有SEQ ID NO:46438-SEQ ID NO:47437中任一者的序列。9.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽，其中：

[0519] Xaa1不包括A、G、K、M、N、Q、R、S或T；或者

[0520] Xaa2不包括A、C、I、K、S、T或V；或者

[0521] Xaa3不包括A、G、I、K、M、Q、R、S、T或V；或者

[0522] Xaa4不包括A、I、K、L、P、Q、R、S、T或V；或者

[0523] Xaa5不包括F、I、L、M、T、V或Y；或者

[0524] Xaa6不包括F、H、M、N、Q、S或Y；或者

[0525] Xaa7不包括A、C、K、M、Q或S；或者

[0526] Xaa8不包括A、C、F、G、M、Q或S；或者

[0527] Xaa9不包括E、F、L、Q、R或Y；或者

[0528] 它们的任何组合。

[0529] 10.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa1不包括A、K、Q或R。11.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa2不包括A、K、S或T。12.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa3不包括A、K、Q、S或T。13.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa4不包括K、I、S或V。14.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa5不包括F、L或Y。15.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa6不包括M或N。16.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa7不包括A、C或S。17.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa8不包括A、C、M或S。18.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa9不包括L、Q或R。19.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa1不包括K。20.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa2不包括A。21.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa3不包括K、Q或T。22.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa3不包括K。23.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa4不包括K。24.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa5不包括F。25.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa6不包括N。26.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa7不包括S。27.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa8不包括C。28.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa9不包括R。29.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽，其中Xaa1具有低溶解性。30.如实施方案29所述的工程化AAV VP衣壳多肽，其中Xaa1选自D和P。31.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽，其中Xaa1具有低可突变性。32.如实施方案31所述的工程化AAV VP衣壳多肽，其中Xaa1选自Xaa1选自C、K和L。33.如实施方案2-4中任一项所述的工

程化AAV VP衣壳多肽，其中Xaa2具有低溶解性。34.如实施方案33所述的工程化AAV VP衣壳多肽，其中Xaa2选自Xaa2选自N、K、P、E和D。35.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽，其中Xaa2具有低亲疏水性。36.如实施方案35所述的工程化AAV VP衣壳多肽，其中Xaa2选自Xaa2选自D、E、R、K、H、N和Q。37.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽，其中Xaa2具有低电荷。38.如实施方案37所述的工程化AAV VP衣壳多肽，其中Xaa2选自D和E。39.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽，其中Xaa2具有高总潜在氢键数。40.如实施方案39所述的工程化AAV VP衣壳多肽，其中Xaa2选自H、N、Q、D、E和R。41.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽，其中Xaa2具有中等体积。42.如实施方案41所述的工程化AAV VP衣壳多肽，其中Xaa2选自D、E、V、P、N和T。43.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽，其中Xaa3具有低溶解性。44.如实施方案43所述的工程化AAV VP衣壳多肽，其中Xaa3选自P和D。45.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽，其中Xaa4具有中等体积。46.如实施方案45所述的工程化AAV VP衣壳多肽，其中Xaa4选自D、E、V、P、N和T。47.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽，其中Xaa5具有低溶解性。48.如实施方案47所述的工程化AAV VP衣壳多肽，其中Xaa5选自N、P、E和D。49.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽，其中Xaa8具有低溶解性。50.如实施方案49所述的工程化AAV VP衣壳多肽，其中Xaa8选自K和Q。51.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽，其中Xaa8具有低亲疏水性。52.如实施方案51所述的工程化AAV VP衣壳多肽，其中Xaa8选自K和R。53.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽，其中Xaa8具有高表面可接近性。54.如实施方案53所述的工程化AAV VP衣壳多肽，其中Xaa8选自E、R和K。55.如实施方案1-54中任一项所述的工程化AAV VP衣壳多肽，其中从对应于残基581的残基至对应于残基589的残基的区域包括端点在内具有与SEQ ID NO:46438-SEQ ID NO:47437中任一者具有至少70％、至少75％、至少80％、至少85％、至少90％、至少92％、至少95％、至少97％或至少99％同一的序列。56.如实施方案55所述的工程化AAV VP衣壳多肽，其中从对应于残基581的残基至对应于残基589的残基的区域包括端点在内具有SEQ ID NO:46438-SEQ ID NO:47437中任一者的序列。57.如实施方案1-56中任一项所述的工程化AAV VP衣壳多肽，其中对非肝脏组织的嗜性被测量为所述rAAV病毒体在非肝脏组织中相较于肝脏组织中的相对积聚，其中所述非肝脏组织组织由CNS组织、骨骼肌、心脏、肺、脾脏、淋巴结、骨髓、乳腺、皮肤、肾上腺、甲状腺、结肠、坐骨神经和脊髓共同组成。58.如实施例57所述的工程化AAV VP衣壳多肽，其中所述CNS组织选自前脑皮层、枕叶皮层、颞叶皮层、丘脑、下丘脑、黑质、海马DG、海马CA1、海马CA3、小脑以及它们的任何组合。59.如实施例57-58中任一项所述的工程化AAV VP衣壳多肽，其中所述较高的组织嗜性是在非肝脏组织中相较于肝脏组织中1.0005倍至约1000倍增加的积聚。60.如实施方案59所述的工程化AAV VP衣壳多肽，其中所述较高的组织嗜性是在非肝脏组织中相较于肝脏组织中至少约1.0005倍、至少约两倍、至少约三倍、至少约四倍、至少约五倍、至少约十倍、至少约二十倍、至少约50倍、至少约75倍、至少约100倍或至少约1000倍增加的积聚。

[0530] [405]D系列实施方案-肝脏嗜性衣壳

[0531] 1.一种工程化腺相关病毒(AAV)病毒蛋白(VP)衣壳多肽，所述工程化腺相关病毒(AAV)病毒蛋白(VP)衣壳多肽具有与SEQ ID NO:1至少70％同一的氨基酸序列，

[0532] 其中所述工程化AAV VP衣壳多肽相较于SEQ ID NO:1在从对应于SEQ ID NO:1的残基581的残基至对应于SEQ ID NO:1的残基589的残基的区域,包括端点在内,中具有至少一个突变,

[0533] 其中所述工程化AAV VP衣壳多肽能够组装成重组AAV病毒体(rAAV),

[0534] 其中与具有SEQ ID NO:1的野生型AAV5 VP衣壳多肽的rAAV病毒体相比,所述至少一个突变赋予所述rAAV较高的对肝脏组织的嗜性,并且

[0535] 其中所述工程化AAV VP衣壳多肽不具有SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7和SEQ ID NO:8中任一者的序列。2.如实施方案1所述的工程化AAV VP衣壳多肽,其中所述工程化AAV VP衣壳多肽具有SEQ ID NO:2的序列,并且其中Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9各自独立地选自任何氨基酸。3.一种工程化腺相关病毒(AAV)病毒蛋白(VP)衣壳多肽,所述工程化腺相关病毒(AAV)病毒蛋白(VP)衣壳多肽具有SEQ ID NO:2的氨基酸序列,

[0536] 其中氨基酸残基Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9各自独立地选自A、R、N、D、C、E、Q、G、H、I、L、K、M、F、P、S、T、W、Y和V,

[0537] 其中所述工程化AAV VP衣壳多肽能够组装成重组AAV病毒体(rAAV),并且

[0538] 其中所述工程化AAV VP衣壳多肽不具有SEQ ID NO:1、SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7和SEQ ID NO:8中任一者的序列。4.如实施方案3所述的工程化AAV VP衣壳多肽,其中与具有SEQ ID NO:1的野生型AAV5 VP衣壳多肽的rAAV病毒体相比,所述rAAV具有较高的对肝脏组织的嗜性。5.如实施方案1所述的工程化AAV VP衣壳多肽,其中从对应于残基581的残基至对应于残基589的残基的区域,包括端点在内,具有与SEQ ID NO:43438-SEQ ID NO:46437中任一者至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%同一的序列。6.如实施方案5所述的工程化AAV VP衣壳多肽,其中从对应于残基581的残基至对应于残基589的残基的区域,包括端点在内,具有SEQ ID NO:43438-SEQ ID NO:46437中任一者的序列。7.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中SEQ ID NO:2的Xaa1至Xaa9区具有与SEQ ID NO:43438-SEQ ID NO:46437中任一者至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%同一的序列。8.如实施方案7所述的工程化AAV VP衣壳多肽,其中SEQ ID NO:2的Xaa1至Xaa9区域具有SEQ ID NO:43438-SEQ ID NO:46437中任一者的序列。9.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中:

[0539] Xaa1选自由A、G、K、M、N、Q、R、S和T组成的组;或者

[0540] Xaa2选自由A、C、I、K、S、T和V组成的组;或者

[0541] Xaa3选自由A、G、I、K、M、Q、R、S、T和V组成的组;或者

[0542] Xaa4选自由A、I、K、L、P、Q、R、S、T和V组成的组;或者

[0543] Xaa5选自由F、I、L、M、T、V和Y组成的组;或者

[0544] Xaa6选自由F、H、M、N、Q、S和Y组成的组;或者

[0545] Xaa7选自由A、C、K、M、Q和S组成的组;或者

[0546] Xaa8选自由A、C、F、G、M、Q和S组成的组;或者

[0547] Xaa9选自由E、F、L、Q、R和Y组成的组;或者

[0548] 它们的任何组合。

[0549] 10.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa1选自A、K、Q和R。11.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa1为K。12.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa2选自A、K、S和T。13.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa2为A。14.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa3选自A、K、Q、S和T。15.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa3选自K、Q和T。16.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa3为K。17.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa4选自K、I、S和V。18.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa4为K。19.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa5选自F、L和Y。20.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa5为F。21.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa6选自M和N。22.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa6为N。23.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa7选自A、C和S。24.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa7为S。25.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa8选自A、C、M和S。26.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa8为C。27.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa9选自L、Q和R。28.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa9为R。29.如实施方案1-28中任一项所述的工程化AAV VP衣壳多肽,其中从对应于残基581的残基至对应于残基589的残基的区域,包括端点在内,具有与SEQ ID NO:43438-SEQ ID NO:44437中任一者至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%同一的序列。30.如实施方案29所述的工程化AAV VP衣壳多肽,其中SEQ ID NO:1的从残基581至残基589的区域具有SEQ ID NO:43438-SEQ ID NO:44437中任一者的序列。31.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa1具有高表面可接近性。32.如实施方案31所述的工程化AAV VP衣壳多肽,其中Xaa1选自K、R和E。33.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa1具有低亲疏水性(＜-3.5)。34.如实施方案33所述的工程化AAV VP衣壳多肽,其中Xaa1选自K和R。35.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa1具有低氨基酸可突变性。36.如实施方案35所述的工程化AAV VP衣壳多肽,其中Xaa1选自H、P、K和R。37.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa1具有低氨基酸溶解性。38.如实施方案37所述的工程化AAV VP衣壳多肽,其中Xaa1选自Q、K和R。39.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa2具有高表面可接近性。

[0550] 40.如实施方案39所述的工程化AAV VP衣壳多肽,其中Xaa2选自E、R和K。41.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa2具有低亲疏水性。42.如实施方案41所述的工程化AAV VP衣壳多肽,其中Xaa2选自K和R。43.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa2具有高氨基酸体积。44.如实施方案43所述的工程化AAV VP衣壳多肽,其中Xaa2选自S、L、I、A、R和K。45.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa3具有高可突变性。46.如实施方案45所述的工程化AAV VP衣壳多肽,其中Xaa3选自N、I、A、M、E和D。47.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa3具有低溶解性。48.如实施方案47所述的工程化AAV VP衣壳多肽,其中Xaa3选自N、K、R和E。49.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa4具有低亲疏水性。50.如实施方案49所述的工程化AAV VP衣壳多肽,其中Xaa4选自K和R。51.如实施

方案2-4中任一项所述的工程化AAV VP衣壳多肽，其中Xaa4具有高氨基酸体积。52.如实施方案51所述的工程化AAV VP衣壳多肽，其中Xaa4选自K、R、I和L。53.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽，其中Xaa5具有中等氨基酸溶解性。54.如实施方案53所述的工程化AAV VP衣壳多肽，其中Xaa5选自H和T。55.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽，其中Xaa8具有低表面可接近性。56.如实施方案55所述的工程化AAV VP衣壳多肽，其中Xaa8选自V和C。57.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽，其中Xaa8具有低平均柔性指数。58.如实施方案57所述的工程化AAV VP衣壳多肽，其中Xaa8选自W、V、M、A、F、L、H和C。59.如实施方案1-58中任一项所述的工程化AAV VP衣壳多肽，其中从对应于残基581的残基至对应于残基589的残基的区域包括端点在内具有与SEQ ID NO:44438-SEQ ID NO:45437中任一者具有至少70％、至少75％、至少80％、至少85％、至少90％、至少92％、至少95％、至少97％或至少99％同一的序列。60.如实施方案59所述的工程化AAV VP衣壳多肽，其中从对应于残基581的残基至对应于残基589的残基的区域包括端点在内具有SEQ ID NO:44438-SEQ ID NO:45437中任一者的序列。61.如实施方案1-60中任一项所述的工程化AAV VP衣壳多肽，其中对肝脏组织的嗜性被测量为所述rAAV病毒体在肝脏组织中相较于非肝脏组织中的相对积聚，其中所述非肝脏组织由CNS、骨骼肌、心脏、肺、脾脏、淋巴结、骨髓、乳腺、皮肤、肾上腺、甲状腺、结肠、坐骨神经和脊髓共同组成。62.如实施方案61所述的工程化AAV VP衣壳多肽，其中所述CNS组织选自前脑皮层、枕叶皮层、颞叶皮层、丘脑、下丘脑、黑质、海马DG、海马CA1、海马CA3、小脑以及它们的任何组合。63.如实施方案61-62中任一项所述的工程化AAV VP衣壳多肽，其中所述较高的组织嗜性是在肝脏组织中相较于非肝脏组织中1.0005倍至约1000倍增加的积聚。64.如实施方案63所述的工程化AAV VP衣壳多肽，其中所述较高的组织嗜性是在肝脏组织中相较于非肝脏组织中至少约1.0005倍、至少约两倍、至少约三倍、至少约四倍、至少约五倍、至少约十倍、至少约二十倍、至少约50倍、至少约75倍、至少约100倍或至少约1000倍增加的积聚。

[0551] [406]E系列实施方案-肾上腺嗜性衣壳

[0552] 1.一种工程化腺相关病毒(AAV)病毒蛋白(VP)衣壳多肽，所述工程化腺相关病毒(AAV)病毒蛋白(VP)衣壳多肽具有与SEQ ID NO:1至少70％同一的氨基酸序列，

[0553] 其中所述工程化AAV VP衣壳多肽相较于SEQ ID NO:1在从对应于SEQ ID NO:1的残基581的残基至对应于SEQ ID NO:1的残基589的残基的区域，包括端点在内，中具有至少一个突变，

[0554] 其中所述工程化AAV VP衣壳多肽能够组装成重组AAV病毒体(rAAV)，

[0555] 其中与具有SEQ ID NO:1的野生型AAV5 VP衣壳多肽的rAAV病毒体相比，所述至少一个突变赋予所述rAAV较高的对肾上腺组织的嗜性，并且

[0556] 其中所述工程化AAV VP衣壳多肽不具有SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7和SEQ ID NO:8中任一者的序列。2.如实施方案1所述的工程化AAV VP衣壳多肽，其中所述工程化AAV VP衣壳多肽具有SEQ ID NO:2的序列，并且其中Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9各自独立地选自任何氨基酸。3.一种工程化腺相关病毒(AAV)病毒蛋白(VP)衣壳多肽，所述工程化腺相关病毒(AAV)病毒蛋白(VP)衣壳多肽具有SEQ ID NO:2的氨基酸序列，

[0557] 其中氨基酸残基Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9各自独立地

选自A、R、N、D、C、E、Q、G、H、I、L、K、M、F、P、S、T、W、Y和V，

[0558] 其中所述工程化AAV VP衣壳多肽能够组装成重组AAV病毒体(rAAV)，并且

[0559] 其中所述工程化AAV VP衣壳多肽不具有SEQ ID NO:1、SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7和SEQ ID NO:8中任一者的序列。4.如实施方案3所述的工程化AAV VP衣壳多肽，其中与具有SEQ ID NO:1的野生型AAV5 VP衣壳多肽的rAAV病毒体相比，所述rAAV具有较高的对肾上腺组织的嗜性。5.如实施方案1所述的工程化AAV VP衣壳多肽，其中从对应于残基581的残基至对应于残基589的残基的区域，包括端点在内，具有与SEQ ID NO:1118-SEQ ID NO:4117中任一者至少70％、至少75％、至少80％、至少85％、至少90％、至少92％、至少95％、至少97％或至少99％同一的序列。6.如实施方案5所述的工程化AAV VP衣壳多肽，其中从对应于残基581的残基至对应于残基589的残基的区域，包括端点在内，具有SEQ ID NO:1118-SEQ ID NO:4117中任一者的序列。7.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽，其中SEQ ID NO:2的Xaa1至Xaa9区域具有与SEQ ID NO:1118-SEQ ID NO:4117中任一者至少70％、至少75％、至少80％、至少85％、至少90％、至少92％、至少95％、至少97％或至少99％同一的序列。8.如实施方案7所述的工程化AAV VP衣壳多肽，其中SEQ ID NO:2的Xaa1至Xaa9区域具有SEQ ID NO:1118-SEQ ID NO:4117中任一者的序列。9.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽，其中：

[0560] Xaa1选自由A、C、K、Q、R、S和T组成的组；或者

[0561] Xaa2选自由A、C、I、S、T和V组成的组；或者

[0562] Xaa3选自由A、F、G、K、M、Q、R、T和V组成的组；或者

[0563] Xaa4选自由A、K、M、Q、R和V组成的组；或者

[0564] Xaa5选自由F、I、L、M、R、T、V和Y组成的组；或者

[0565] Xaa6选自由G、H、M、N、R和S组成的组；或者

[0566] Xaa7选自由A、H、K、Q、R、S和V组成的组；或者

[0567] Xaa8选自由A、G、H、M、Q和S组成的组；或者

[0568] Xaa9选自由A、E、N、P、R、S和Y组成的组；或者

[0569] 它们的任何组合。10.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa1选自C、K和R。11.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa1为C。12.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa2选自A、V和T。13.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa2为V。14.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa3选自A、G和M。15.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa3为M。16.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa4选自A、R和K。17.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa4为K。18.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa5选自R、V和Y。19.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa5为V。20.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa6选自H和N。21.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa6为N。22.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa7选自H、Q和V。23.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa7为H。24.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa8选自A、G、M和S。25.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa8为S。26.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa9选自P和E。27.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa9为P。28.如实施方案1-27中任一项所述的工程

化AAV VP衣壳多肽,其中从对应于残基581的残基至对应于残基589的残基的区域,包括端点在内,具有与SEQ ID NO:1118-SEQ ID NO:2117中任一者至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%同一的序列。29.如实施方案28所述的工程化AAV VP衣壳多肽,其中SEQ ID NO:1的从残基581至残基589的区域具有SEQ ID NO:1118-SEQ ID NO:2117中任一者的序列。30.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa1具有低摩尔质量。31.如实施方案30所述的工程化AAV VP衣壳多肽,其中Xaa1选自V、P、S和C。32.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa1具有低亲疏水性。33.如实施方案X所述的工程化AAV VP衣壳多肽,其中Xaa1选自T、S、W和Y。34.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa2具有低亲疏水性。35.如实施方案34所述的工程化AAV VP衣壳多肽,其中Xaa2为R。36.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa2具有低可突变性。37.如实施方案36所述的工程化AAV VP衣壳多肽,其中Xaa2为C。38.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa2具有低溶解性。39.如实施方案38所述的工程化AAV VP衣壳多肽,其中Xaa2为K。40.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa3具有低平均柔性。41.如实施方案40所述的工程化AAV VP衣壳多肽,其中Xaa3选自W、M和F。42.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa3具有高溶解性。43.如实施方案42所述的工程化AAV VP衣壳多肽,其中Xaa3为M。44.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa4具有高表面可接近性。45.如实施方案44所述的工程化AAV VP衣壳多肽,其中Xaa4选自K和R。46.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa4具有高平均柔性。47.如实施方案46所述的工程化AAV VP衣壳多肽,其中Xaa4选自K、I和N。

[0570] 48.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa5具有中等可突变性。49.如实施方案48所述的工程化AAV VP衣壳多肽,其中Xaa5选自R和H。50.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa5具有高goldman engelman steitz。51.如实施方案50所述的工程化AAV VP衣壳多肽,其中Xaa5选自V和L。52.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa5具有低亲疏水性。53.如实施方案52所述的工程化AAV VP衣壳多肽,其中Xaa5为R。54.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa5具有高体积。

[0571] 55.如实施方案54所述的工程化AAV VP衣壳多肽,其中Xaa5选自Y、R和F。56.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa6具有高溶解性。57.如实施方案56所述的工程化AAV VP衣壳多肽,其中Xaa6选自Y、V、M、A和C。58.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa7具有中等可突变性。59.如实施方案58所述的工程化AAV VP衣壳多肽,其中Xaa7选自V、H和R。60.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa7具有低溶解性。61.如实施方案59所述的工程化AAV VP衣壳多肽,其中Xaa7为R。62.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa8具有高平均柔性。64.如实施方案62所述的工程化AAV VP衣壳多肽,其中Xaa8选自K、I和N。65.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa8具有高摩尔质量。66.如实施方案65所述的工程化AAV VP衣壳多肽,其中Xaa8选自R和Y。67.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa9具有高可突变性。68.如实施方案X所述的工程化AAV

VP衣壳多肽，其中Xaa9为N。69.如实施方案1-68中任一项所述的工程化AAV VP衣壳多肽，其中从对应于残基581的残基至对应于残基589的残基的区域包括端点在内具有与SEQ ID NO:2118-SEQ ID NO:3117中任一者具有至少70％、至少75％、至少80％、至少85％、至少90％、至少92％、至少95％、至少97％或至少99％同一的序列。70.如实施方案69所述的工程化AAV VP衣壳多肽，其中从对应于残基581的残基至对应于残基589的残基的区域包括端点在内具有SEQ ID NO:2118-SEQ ID NO:3117中任一者的序列。71.如实施方案1-70中任一项所述的工程化AAV VP衣壳多肽，其中对肾上腺组织的嗜性被测量为所述rAAV病毒体在肾上腺组织中相较于非肾上腺组织中的相对积聚，其中所述非肾上腺组织由CNS、肝脏、骨骼肌、心脏、肺、脾脏、淋巴结、骨髓、乳腺、皮肤、甲状腺、结肠、坐骨神经和脊髓共同组成。72.如实施例71所述的工程化AAV VP衣壳多肽，其中所述CNS组织选自前脑皮层、枕叶皮层、颞叶皮层、丘脑、下丘脑、黑质、海马DG、海马CA1、海马CA3、小脑以及它们的任何组合。

[0572] 73.如实施方案71-72中任一项所述的工程化AAV VP衣壳多肽，其中所述较高的组织嗜性是在肾上腺组织中相较于非肾上腺组织中1.0005倍至约1000倍增加的积聚。74.如实施方案73所述的工程化AAV VP衣壳多肽，其中所述较高的组织嗜性是在肾上腺组织中相较于非肾上腺组织中至少约1.0005倍、至少约两倍、至少约三倍、至少约四倍、至少约五倍、至少约十倍、至少约二十倍、至少约50倍、至少约75倍、至少约100倍或至少约1000倍增加的积聚。

[0573] [407]F系列实施方案-骨髓嗜性衣壳

[0574] 1.一种工程化腺相关病毒(AAV)病毒蛋白(VP)衣壳多肽，所述工程化腺相关病毒(AAV)病毒蛋白(VP)衣壳多肽具有与SEQ ID NO:1至少70％同一的氨基酸序列，

[0575] 其中所述工程化AAV VP衣壳多肽相较于SEQ ID NO:1在从对应于SEQ ID NO:1的残基581的残基至对应于SEQ ID NO:1的残基589的残基的区域，包括端点在内，中具有至少一个突变，

[0576] 其中所述工程化AAV VP衣壳多肽能够组装成重组AAV病毒体(rAAV)，

[0577] 其中与具有SEQ ID NO:1的野生型AAV5 VP衣壳多肽的rAAV病毒体相比，所述至少一个突变赋予所述rAAV较高的对骨髓组织的嗜性，并且

[0578] 其中所述工程化AAV VP衣壳多肽不具有SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7和SEQ ID NO:8中任一者的序列。2.如实施方案1所述的工程化AAV VP衣壳多肽，其中所述工程化AAV VP衣壳多肽具有SEQ ID NO:2的序列，并且其中Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9各自独立地选自任何氨基酸。3.一种工程化腺相关病毒(AAV)病毒蛋白(VP)衣壳多肽，所述工程化腺相关病毒(AAV)病毒蛋白(VP)衣壳多肽具有SEQ ID NO:2的氨基酸序列，

[0579] 其中氨基酸残基Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9各自独立地选自A、R、N、D、C、E、Q、G、H、I、L、K、M、F、P、S、T、W、Y和V，

[0580] 其中所述工程化AAV VP衣壳多肽能够组装成重组AAV病毒体(rAAV)，并且

[0581] 其中所述工程化AAV VP衣壳多肽不具有SEQ ID NO:1、SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7和SEQ ID NO:8中任一者的序列。4.如实施方案3所述的工程化AAV VP衣壳多肽，其中与具有SEQ ID NO:1的野生型AAV5 VP衣壳多肽的rAAV病毒体相比，所述rAAV具有较高的对骨髓组织的嗜性。5.如实施方案1所述的工程化

AAV VP衣壳多肽,其中从对应于残基581的残基至对应于残基589的残基的区域,包括端点在内,具有与SEQ ID NO:4118-SEQ ID NO:7117中任一者至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%同一的序列。6.如实施方案5所述的工程化AAV VP衣壳多肽,其中从对应于残基581的残基至对应于残基589的残基的区域,包括端点在内,具有SEQ ID NO:4118-SEQ ID NO:7117中任一者的序列。7.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中SEQ ID NO:2的Xaa1至Xaa9区域具有与SEQ ID NO:4118-SEQ ID NO:7117中任一者至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%、至少98%或至少99%同一的序列。8.如实施方案7所述的工程化AAV VP衣壳多肽,其中SEQ ID NO:2的Xaa1至Xaa9区域具有SEQ ID NO:4118-SEQ ID NO:7117中任一者的序列。9.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中:

[0582] Xaa1选自由A、E、G、Q、S和T组成的组;或者

[0583] Xaa2选自由A、I、Q、S、T、V和Y组成的组;或者

[0584] Xaa3选自由A、G、I、M、Q、S和T组成的组;或者

[0585] Xaa4选自由A、E、P、Q、T和V组成的组;或者

[0586] Xaa5选自由F、I、L、M、Q、V和Y组成的组;或者

[0587] Xaa6选自由F、I、N、Q、S和V组成的组;或者

[0588] Xaa7选自由A、C、M、S和V组成的组;或者

[0589] Xaa8选自由A、C、D、G、M、S和Y组成的组;或者

[0590] Xaa9选自由D、E、G、L、P、S和Y组成的组;或者

[0591] 它们的任何组合。10.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa1选自A、E和T。11.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa1为E。12.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa2选自A、S和T。13.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa2为A。14.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa3选自A、Q和T。15.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa3为Q。16.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa4选自A、P和Q。17.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa4为Q。18.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa5选自F、V和Y。19.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa5为V。20.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa6选自I、N、Q和S。21.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa6为S。22.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa7为A、C和V。23.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa7为C。24.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa8选自A、M、S和Y。25.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa8为M。26.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa9选自D、E和P。27.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa9为P。28.如实施方案1-27中任一项所述的工程化AAV VP衣壳多肽,其中从对应于残基581的残基至对应于残基589的残基的区域,包括端点在内,具有与SEQ ID NO:4118-SEQ ID NO:5117中任一者至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%同一的序列。29.如实施方案28所述的工程化AAV VP衣壳多肽,其中SEQ ID NO:1的从残基581至残基589的区域具有SEQ ID NO:4118-SEQ ID NO:5117中任一者的序列。30.如实施方案2-4中任一项所

述的工程化AAV VP衣壳多肽，其中Xaa1具有高亲疏水性。31.如实施方案30所述的工程化AAV VP衣壳多肽，其中Xaa1选自V、I和L。32.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽，其中Xaa1具有低可突变性。33.如实施方案32所述的工程化AAV VP衣壳多肽，其中Xaa1选自Y、L、F和C。34.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽，其中Xaa2具有低亲疏水性。35.如实施方案34所述的工程化AAV VP衣壳多肽，其中Xaa2选自Y和W。36.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽，其中Xaa2具有高摩尔质量。37.如实施方案36所述的工程化AAV VP衣壳多肽，其中Xaa2为W。38.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽，其中Xaa2具有低表面可接近性。39.如实施方案38所述的工程化AAV VP衣壳多肽，其中Xaa2选自W和A。40.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽，其中Xaa2具有低亲水性。41.如实施方案40所述的工程化AAV VP衣壳多肽，其中Xaa2为W。42.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽，其中Xaa2具有低可突变性。43.如实施方案42所述的工程化AAV VP衣壳多肽，其中Xaa2为C。44.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽，其中Xaa2具有低平均柔性。45.如实施方案44所述的工程化AAV VP衣壳多肽，其中Xaa2选自W、M和F。46.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽，其中Xaa5具有低平均柔性。47.如实施方案46所述的工程化AAV VP衣壳多肽，其中Xaa5选自W、M和F。48.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽，其中Xaa6具有低平均柔性。

[0592] 49.如实施方案48所述的工程化AAV VP衣壳多肽，其中Xaa6选自W、M和F。50.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽，其中Xaa6具有低可突变性。51.如实施方案50所述的工程化AAV VP衣壳多肽，其中Xaa6选自Y、F、L和C。52.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽，其中Xaa6具有高溶解性。53.如实施方案52所述的工程化AAV VP衣壳多肽，其中Xaa6选自W、F、I和L。54.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽，其中Xaa7具有低表面可接近性。55.如实施方案54所述的工程化AAV VP衣壳多肽，其中Xaa7为C。56.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽，其中Xaa7具有高表面可接近性。57.如实施方案56所述的工程化AAV VP衣壳多肽，其中Xaa7选自D和N。

[0593] 58.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽，其中Xaa7具有低可突变性。

[0594] 59.如实施方案58所述的工程化AAV VP衣壳多肽，其中Xaa7为C。

[0595] 60.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽，其中Xaa7具有高溶解性。

[0596] 61.如实施方案60所述的工程化AAV VP衣壳多肽，其中Xaa7为C。

[0597] 62.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽，其中Xaa7具有低溶解性。

[0598] 63.如实施方案X所述的工程化AAV VP衣壳多肽，其中Xaa7为D。

[0599] 64.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽，其中Xaa8具有低电荷。

[0600] 65.如实施方案64所述的工程化AAV VP衣壳多肽，其中Xaa8选自D和E。

[0601] 66.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽，其中Xaa8具有高可突变性。

[0602] 67.如实施方案66所述的工程化AAV VP衣壳多肽,其中Xaa8选自D、E、A和T。

[0603] 68.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa9具有高摩尔质量。

[0604] 69.如实施方案68所述的工程化AAV VP衣壳多肽,其中Xaa9选自H和F。

[0605] 70.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa9具有低可突变性。

[0606] 71.如实施方案X所述的工程化AAV VP衣壳多肽,其中Xaa9选自Y、F和L。

[0607] 72.如实施方案1-71中任一项所述的工程化AAV VP衣壳多肽,其中从对应于残基581的残基至对应于残基589的残基的区域包括端点在内具有与SEQ ID NO:5118-SEQ ID NO:6117中任一者具有至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%同一的序列。

[0608] 73.如实施方案72所述的工程化AAV VP衣壳多肽,其中从对应于残基581的残基至对应于残基589的残基的区域包括端点在内具有SEQ ID NO:5118-SEQ ID NO:6117中任一者的序列。

[0609] 74.如实施方案1-73中任一项所述的工程化AAV VP衣壳多肽,其中对骨髓组织的嗜性被测量为所述rAAV病毒体在骨髓组织中相较于非骨髓组织中的相对积聚,其中所述非骨髓组织由CNS、肝脏、骨骼肌、心脏、肺、脾脏、淋巴结、乳腺、皮肤、肾上腺、甲状腺、结肠、坐骨神经和脊髓共同组成。

[0610] 75.如实施方案74所述的工程化AAV VP衣壳多肽,其中所述CNS组织选自前脑皮层、枕叶皮层、颞叶皮层、丘脑、下丘脑、黑质、海马DG、海马CA1、海马CA3、小脑以及它们的任何组合。

[0611] 76.如实施方案74-75中任一项所述的工程化AAV VP衣壳多肽,其中所述较高的组织嗜性是在骨髓组织中相较于非骨髓组织中1.0005倍至约1000倍增加的积聚。

[0612] 77.如实施方案76所述的工程化AAV VP衣壳多肽,其中所述较高的组织嗜性是在骨髓组织中相较于非骨髓组织中至少约1.0005倍、至少约两倍、至少约三倍、至少约四倍、至少约五倍、至少约十倍、至少约二十倍、至少约50倍、至少约75倍、至少约100倍或至少约1000倍增加的积聚。

[0613] [408]G系列实施方案-结肠嗜性衣壳

[0614] 1.一种工程化腺相关病毒(AAV)病毒蛋白(VP)衣壳多肽,所述工程化腺相关病毒(AAV)病毒蛋白(VP)衣壳多肽具有与SEQ ID NO:1至少70%同一的氨基酸序列,

[0615] 其中所述工程化AAV VP衣壳多肽相较于SEQ ID NO:1在从对应于SEQ ID NO:1的残基581的残基至对应于SEQ ID NO:1的残基589的残基的区域,包括端点在内,中具有至少一个突变,

[0616] 其中所述工程化AAV VP衣壳多肽能够组装成重组AAV病毒体(rAAV),

[0617] 其中与具有SEQ ID NO:1的野生型AAV5 VP衣壳多肽的rAAV病毒体相比,所述至少一个突变赋予所述rAAV较高的对结肠组织的嗜性,并且

[0618] 其中所述工程化AAV VP衣壳多肽不具有SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7和SEQ ID NO:8中任一者的序列。2.如实施方案1所述的工程化AAV VP衣壳多肽,其中所述工程化AAV VP衣壳多肽具有SEQ ID NO:2的序列,并且其中

Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9各自独立地选自任何氨基酸。3.一种工程化腺相关病毒(AAV)病毒蛋白(VP)衣壳多肽，所述工程化腺相关病毒(AAV)病毒蛋白(VP)衣壳多肽具有SEQ ID NO:2的氨基酸序列，

[0619] 其中氨基酸残基Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9各自独立地选自A、R、N、D、C、E、Q、G、H、I、L、K、M、F、P、S、T、W、Y和V，

[0620] 其中所述工程化AAV VP衣壳多肽能够组装成重组AAV病毒体(rAAV)，并且

[0621] 其中所述工程化AAV VP衣壳多肽不具有SEQ ID NO:1、SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7和SEQ ID NO:8中任一者的序列。4.如实施方案3所述的工程化AAV VP衣壳多肽，其中与具有SEQ ID NO:1的野生型AAV5 VP衣壳多肽的rAAV病毒体相比，所述rAAV具有较高的对结肠组织的嗜性。

[0622] 5.如实施方案1所述的工程化AAV VP衣壳多肽，其中从对应于残基581的残基至对应于残基589的残基的区域，包括端点在内，具有与SEQ ID NO:10118-SEQ ID NO:13117中任一者至少70％、至少75％、至少80％、至少85％、至少90％、至少92％、至少95％、至少97％或至少99％同一的序列。

[0623] 6.如实施方案5所述的工程化AAV VP衣壳多肽，其中从对应于残基581的残基至对应于残基589的残基的区域，包括端点在内，具有SEQ ID NO:10118-SEQ ID NO:13117中任一者的序列。

[0624] 7.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽，其中SEQ ID NO:2的Xaa1至Xaa9区域具有与SEQ ID NO:10118-SEQ ID NO:13117中任一者至少70％、至少75％、至少80％、至少85％、至少90％、至少92％、至少95％、至少97％或至少99％同一的序列。

[0625] 8.如实施方案7所述的工程化AAV VP衣壳多肽，其中SEQ ID NO:2的Xaa1至Xaa9区域具有SEQ ID NO:10118-SEQ ID NO:13117中任一者的序列。

[0626] 9.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽，其中：

[0627] Xaa1选自由C、F、H、N、P、W和Y组成的组；或者

[0628] Xaa2选自由D、E、F、L和P组成的组；或者

[0629] Xaa3选自由C、F、H、I、L、P和Y组成的组；或者

[0630] Xaa4选自由C、D、E、N和P组成的组；或者

[0631] Xaa5选自由D、E、G、P和W组成的组；或者

[0632] Xaa6选自由C、K、R和V组成的组；或者

[0633] Xaa7选自由D、M、P和V组成的组；或者

[0634] Xaa8选自由D、I、K、L、P、R和V组成的组；或者

[0635] Xaa9选自由C、H、I、K、L、M和W组成的组；或者

[0636] 它们的任何组合。

[0637] 10.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa1选自F、P和W。

[0638] 11.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa1为P。

[0639] 12.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa2选自D、E、L和P。

[0640] 13.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa2为P。

[0641] 14.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa3选自C、H和P。

[0642] 15.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa3为P。

[0643] 16.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa4选自C、D和E。

[0644] 17.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa4为C。

[0645] 18.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa5选自G、P和W。

[0646] 19.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa5为P。

[0647] 20.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa6选自K和R。

[0648] 21.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa6为R。

[0649] 22.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa7为P。

[0650] 23.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa8选自K、P和R。

[0651] 24.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa8为P。

[0652] 25.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa9选自I、L和M。

[0653] 26.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa9为I。

[0654] 27.如实施方案1-26中任一项所述的工程化AAV VP衣壳多肽,其中从对应于残基581的残基至对应于残基589的残基的区域,包括端点在内,具有与SEQ ID NO:10118-SEQ ID NO:11117中任一者至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%同一的序列。

[0655] 28.如实施方案27所述的工程化AAV VP衣壳多肽,其中SEQ ID NO:1的从残基581至残基589的区域具有SEQ ID NO:10118-SEQ ID NO:11117中任一者的序列。

[0656] 29.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa1具有高摩尔质量。

[0657] 30.如实施方案29所述的工程化AAV VP衣壳多肽,其中Xaa1选自Y、W。

[0658] 31.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa1具有高溶解性。

[0659] 32.如实施方案31所述的工程化AAV VP衣壳多肽,其中Xaa1选自W、F、I和L。

[0660] 33.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa2具有低溶解性。

[0661] 34.如实施方案33所述的工程化AAV VP衣壳多肽,其中Xaa2为D。

[0662] 35.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa2具有低可突变性。

[0663] 36.如实施方案35所述的工程化AAV VP衣壳多肽,其中Xaa2选自P和K。

[0664] 37.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa2具有中等摩尔质量。

[0665] 38.如实施方案37所述的工程化AAV VP衣壳多肽,其中Xaa2选自D、E、N、K、M、Q、I和L。

[0666] 39.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa2具有低亲疏水性。

[0667] 40.如实施方案39所述的工程化AAV VP衣壳多肽,其中Xaa2选自D、E、R、K、H、N和Q。

[0668] 41.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa3具有低可突变性。

[0669] 42.如实施方案41所述的工程化AAV VP衣壳多肽,其中Xaa3选自K、V、P和C。

[0670] 43.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa3具有高溶解性。

[0671] 44.如实施方案43所述的工程化AAV VP衣壳多肽,其中Xaa3选自W、F、I和L。

[0672] 45.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa5具有高平均柔性。

[0673] 46.如实施方案45所述的工程化AAV VP衣壳多肽,其中Xaa5选自S、P、G、R、E和D。

[0674] 47.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa5具有高表面可接近性。

[0675] 48.如实施方案47所述的工程化AAV VP衣壳多肽,其中Xaa5选自D和N。

[0676] 49.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa6具有低亲疏水性。

[0677] 50.如实施方案49所述的工程化AAV VP衣壳多肽,其中Xaa6选自R。

[0678] 51.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa6具有低可突变性。

[0679] 52.如实施方案51所述的工程化AAV VP衣壳多肽,其中Xaa6选自Y、R、F和L。

[0680] 53.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa6具有低溶解性。

[0681] 54.如实施方案53所述的工程化AAV VP衣壳多肽,其中Xaa6选自R和Q。

[0682] 55.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa6具有高表面可接近性。

[0683] 56.如实施方案55所述的工程化AAV VP衣壳多肽,其中Xaa6选自E、R、K。

[0684] 57.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa6具有高平均柔性。

[0685] 58.如实施方案57所述的工程化AAV VP衣壳多肽,其中Xaa6选自G和R。

[0686] 59.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa8具有低溶解性。

[0687] 60.如实施方案59所述的工程化AAV VP衣壳多肽,其中Xaa8为D。

[0688] 61.如实施方案1-60中任一项所述的工程化AAV VP衣壳多肽,其中从对应于残基581的残基至对应于残基589的残基的区域包括端点在内具有与SEQ ID NO:11118-SEQ ID NO:12117中任一者具有至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%同一的序列。

[0689] 62.如实施方案61所述的工程化AAV VP衣壳多肽,其中从对应于残基581的残基至对应于残基589的残基的区域包括端点在内具有SEQ ID NO:11118-SEQ ID NO:12117中任一者的序列。

[0690] 63.如实施方案1-62中任一项所述的工程化AAV VP衣壳多肽,其中对结肠组织的嗜性被测量为所述rAAV病毒体在结肠组织中相较于非结肠组织中的相对积聚,其中所述非结肠组织由CNS、肝脏、骨骼肌、心脏、肺、脾脏、淋巴结、骨髓、乳腺、皮肤、肾上腺、甲状腺、坐骨神经和脊髓共同组成。

[0691] 64.如实施方案63所述的工程化AAV VP衣壳多肽,其中所述CNS组织选自前脑皮

层、枕叶皮层、颞叶皮层、丘脑、下丘脑、黑质、海马DG、海马CA1、海马CA3、小脑以及它们的任何组合。

[0692] 65.如实施方案63-64中任一项所述的工程化AAV VP衣壳多肽，其中所述较高的组织嗜性是在结肠组织中相较于非结肠组织中1.0005倍至约1000倍增加的积聚。

[0693] 66.如实施方案65所述的工程化AAV VP衣壳多肽，其中所述较高的组织嗜性是在结肠组织中相较于非结肠组织中至少约1.0005倍、至少约两倍、至少约三倍、至少约四倍、至少约五倍、至少约十倍、至少约二十倍、至少约50倍、至少约75倍、至少约100倍或至少约1000倍增加的积聚。

[0694] [409]H系列实施方案-心脏嗜性衣壳

[0695] 1.一种工程化腺相关病毒(AAV)病毒蛋白(VP)衣壳多肽，所述工程化腺相关病毒(AAV)病毒蛋白(VP)衣壳多肽具有与SEQ ID NO:1至少70％同一的氨基酸序列，

[0696] 其中所述工程化AAV VP衣壳多肽相较于SEQ ID NO:1在从对应于SEQ ID NO:1的残基581的残基至对应于SEQ ID NO:1的残基589的残基的区域，包括端点在内，中具有至少一个突变，

[0697] 其中所述工程化AAV VP衣壳多肽能够组装成重组AAV病毒体(rAAV)，

[0698] 其中与具有SEQ ID NO:1的野生型AAV5 VP衣壳多肽的rAAV病毒体相比，所述至少一个突变赋予所述rAAV较高的对心脏组织的嗜性，并且

[0699] 其中所述工程化AAV VP衣壳多肽不具有SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7和SEQ ID NO:8中任一者的序列。2.如实施方案1所述的工程化AAV VP衣壳多肽，其中所述工程化AAV VP衣壳多肽具有SEQ ID NO:2的序列，并且其中Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9各自独立地选自任何氨基酸。3.一种工程化腺相关病毒(AAV)病毒蛋白(VP)衣壳多肽，所述工程化腺相关病毒(AAV)病毒蛋白(VP)衣壳多肽具有SEQ ID NO:2的氨基酸序列，

[0700] 其中氨基酸残基Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9各自独立地选自A、R、N、D、C、E、Q、G、H、I、L、K、M、F、P、S、T、W、Y和V，

[0701] 其中所述工程化AAV VP衣壳多肽能够组装成重组AAV病毒体(rAAV)，并且

[0702] 其中所述工程化AAV VP衣壳多肽不具有SEQ ID NO:1、SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7和SEQ ID NO:8中任一者的序列。4.如实施方案3所述的工程化AAV VP衣壳多肽，其中与具有SEQ ID NO:1的野生型AAV5 VP衣壳多肽的rAAV病毒体相比，所述rAAV具有较高的对心脏组织的嗜性。

[0703] 5.如实施方案1所述的工程化AAV VP衣壳多肽，其中从对应于残基581的残基至对应于残基589的残基的区域，包括端点在内，具有与SEQ ID NO:13118-SEQ ID NO:16117中任一者至少70％、至少75％、至少80％、至少85％、至少90％、至少92％、至少95％、至少97％或至少99％同一的序列。

[0704] 6.如实施方案5所述的工程化AAV VP衣壳多肽，其中从对应于残基581的残基至对应于残基589的残基的区域，包括端点在内，具有SEQ ID NO:13118-SEQ ID NO:16117中任一者的序列。

[0705] 7.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽，其中SEQ ID NO:2的Xaa1至Xaa9区域具有与SEQ ID NO:13118-SEQ ID NO:16117中任一者至少70％、至少75％、

至少80％、至少85％、至少90％、至少92％、至少95％、至少97％或至少99％同一的序列。

[0706] 8.如实施方案7所述的工程化AAV VP衣壳多肽，其中SEQ ID NO:2的Xaa1至Xaa9区域具有SEQ ID NO:13118-SEQ ID NO:16117中任一者的序列。

[0707] 9.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽，其中：

[0708] Xaa1选自由I、K、L、M、T和V组成的组；或者

[0709] Xaa2选自由A、C、G、I、K和S组成的组；或者

[0710] Xaa3选自由A、D、E、G、K、M和V组成的组；或者

[0711] Xaa4选自由F、H、R、T、W和Y组成的组；或者

[0712] Xaa5选自由F、L、M和R组成的组；或者

[0713] Xaa6选自由A、H、N、W和Y组成的组；或者

[0714] Xaa7选自由A、C、E、F、K和T组成的组；或者

[0715] Xaa8选自由A、C、M、S和T组成的组；或者

[0716] Xaa9选自由A、D、G和P组成的组；或者

[0717] 它们的任何组合。

[0718] 10.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa1选自K和L。

[0719] 11.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa1为K。

[0720] 12.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa2选自A、C和S。

[0721] 13.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa2为A。

[0722] 14.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa3选自E和V。

[0723] 15.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa3为E。

[0724] 16.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa4选自F、R和T。

[0725] 17.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa4为R。

[0726] 18.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa5为L。

[0727] 19.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa6选自H、N和Y。

[0728] 20.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa6为H。

[0729] 21.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa7选自C、F和T。

[0730] 22.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa7为F。

[0731] 23.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa8选自C、M和S。

[0732] 24.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa8为C。

[0733] 25.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa9选自A和G。

[0734] 26.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa9为A。

[0735] 27.如实施方案1-26中任一项所述的工程化AAV VP衣壳多肽，其中从对应于残基581的残基至对应于残基589的残基的区域，包括端点在内，具有与SEQ ID NO:13118-SEQ ID NO:14117中任一者至少70％、至少75％、至少80％、至少85％、至少90％、至少92％、至少95％、至少97％或至少99％同一的序列。

[0736] 28.如实施方案27所述的工程化AAV VP衣壳多肽，其中SEQ ID NO:1的从残基581至残基589的区域具有SEQ ID NO:13118-SEQ ID NO:14117中任一者的序列。

[0737] 29.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽，其中Xaa1具有低溶解性。

[0738] 30.如实施方案29所述的工程化AAV VP衣壳多肽,其中Xaa1选自N和E。

[0739] 31.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa1具有低亲疏水性。

[0740] 32.如实施方案31所述的工程化AAV VP衣壳多肽,其中Xaa1选自H、N、Q、P、Y、D和E。

[0741] 33.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa1具有高可突变性。

[0742] 34.如实施方案33所述的工程化AAV VP衣壳多肽,其中Xaa1选自A和E。

[0743] 35.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa2具有高亲疏水性。

[0744] 36.如实施方案35所述的工程化AAV VP衣壳多肽,其中Xaa2选自V和I。

[0745] 37.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa2具有中等可突变性。

[0746] 38.如实施方案37所述的工程化AAV VP衣壳多肽,其中Xaa2为V。

[0747] 39.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa2具有中等体积。

[0748] 40.如实施方案39所述的工程化AAV VP衣壳多肽,其中Xaa2选自V、E和Q。

[0749] 41.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa2具有高溶解性。

[0750] 42.如实施方案41所述的工程化AAV VP衣壳多肽,其中Xaa2选自V和M。

[0751] 43.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa3具有低溶解性。

[0752] 44.如实施方案43所述的工程化AAV VP衣壳多肽,其中Xaa3选自R和Q。

[0753] 45.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa4具有低表面可接近性。

[0754] 46.如实施方案45所述的工程化AAV VP衣壳多肽,其中Xaa4为C。

[0755] 47.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa4具有高溶解性。

[0756] 48.如实施方案47所述的工程化AAV VP衣壳多肽,其中Xaa4为C。

[0757] 49.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa4具有低电荷。

[0758] 50.如实施方案49所述的工程化AAV VP衣壳多肽,其中Xaa4选自D、E、Y、W、V、P、M、A、G、F、I、L、N、Q、S、T和C。

[0759] 51.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa4具有高亲疏水性。

[0760] 52.如实施方案51所述的工程化AAV VP衣壳多肽,其中Xaa4为C。

[0761] 53.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa5具有高表面可接近性。

[0762] 54.如实施方案53所述的工程化AAV VP衣壳多肽,其中Xaa5选自D、E、R、K、N和Q。

[0763] 55.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa5具有低溶解

性。

[0764] 56.如实施方案55所述的工程化AAV VP衣壳多肽,其中Xaa5为D。

[0765] 57.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa6具有低可突变性。

[0766] 58.如实施方案57所述的工程化AAV VP衣壳多肽,其中Xaa6为C。

[0767] 59.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa6具有低溶解性。

[0768] 60.如实施方案59所述的工程化AAV VP衣壳多肽,其中Xaa6为D。

[0769] 61.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa8具有高表面可接近性。

[0770] 62.如实施方案61所述的工程化AAV VP衣壳多肽,其中Xaa8选自D和N。

[0771] 63.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa8具有高平均柔性。

[0772] 64.如实施方案63所述的工程化AAV VP衣壳多肽,其中Xaa8选自D、R、P、G和S。

[0773] 65.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa9具有中等摩尔质量。

[0774] 66.如实施方案65所述的工程化AAV VP衣壳多肽,其中Xaa9选自N、D、L和I。

[0775] 67.如实施方案1-66中任一项所述的工程化AAV VP衣壳多肽,其中从对应于残基581的残基至对应于残基589的残基的区域包括端点在内具有与SEQ ID NO:14118-SEQ ID NO:15117中任一者具有至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%同一的序列。

[0776] 68.如实施方案67所述的工程化AAV VP衣壳多肽,其中从对应于残基581的残基至对应于残基589的残基的区域包括端点在内具有SEQ ID NO:14118-SEQ ID NO:15117中任一者的序列。

[0777] 69.如实施方案1-68中任一项所述的工程化AAV VP衣壳多肽,其中对心脏组织的嗜性被测量为所述rAAV病毒体在心脏组织中相较于非心脏组织中的相对积聚,其中所述非心脏组织由CNS、肝脏、骨骼肌、肺、脾脏、淋巴结、骨髓、乳腺、皮肤、肾上腺、甲状腺、结肠、坐骨神经和脊髓共同组成。

[0778] 70.如实施方案69所述的工程化AAV VP衣壳多肽,其中所述CNS组织选自前脑皮层、枕叶皮层、颞叶皮层、丘脑、下丘脑、黑质、海马DG、海马CA1、海马CA3、小脑以及它们的任何组合。

[0779] 71.如实施方案69-70中任一项所述的工程化AAV VP衣壳多肽,其中所述较高的组织嗜性是在心脏组织中相较于非心脏组织中1.0005倍至约1000倍增加的积聚。

[0780] 72.如实施方案71所述的工程化AAV VP衣壳多肽,其中所述较高的组织嗜性是在心脏组织中相较于非心脏组织中至少约1.0005倍、至少约两倍、至少约三倍、至少约四倍、至少约五倍、至少约十倍、至少约二十倍、至少约50倍、至少约75倍、至少约100倍或至少约1000倍增加的积聚。

[0781] [410]I系列实施方案-肺嗜性衣壳

[0782] 1.一种工程化腺相关病毒(AAV)病毒蛋白(VP)衣壳多肽,所述工程化腺相关病毒

(AAV)病毒蛋白(VP)衣壳多肽具有与SEQ ID NO:1至少70%同一的氨基酸序列，

[0783] 其中所述工程化AAV VP衣壳多肽相较于SEQ ID NO:1在从对应于SEQ ID NO:1的残基581的残基至对应于SEQ ID NO:1的残基589的残基的区域，包括端点在内，中具有至少一个突变，

[0784] 其中所述工程化AAV VP衣壳多肽能够组装成重组AAV病毒体(rAAV)，

[0785] 其中与具有SEQ ID NO:1的野生型AAV5 VP衣壳多肽的rAAV病毒体相比，所述至少一个突变赋予所述rAAV较高的对肺组织的嗜性，并且

[0786] 其中所述工程化AAV VP衣壳多肽不具有SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7和SEQ ID NO:8中任一者的序列。2.如实施方案1所述的工程化AAV VP衣壳多肽，其中所述工程化AAV VP衣壳多肽具有SEQ ID NO:2的序列，并且其中Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9各自独立地选自任何氨基酸。3.一种工程化腺相关病毒(AAV)病毒蛋白(VP)衣壳多肽，所述工程化腺相关病毒(AAV)病毒蛋白(VP)衣壳多肽具有SEQ ID NO:2的氨基酸序列，

[0787] 其中氨基酸残基Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9各自独立地选自A、R、N、D、C、E、Q、G、H、I、L、K、M、F、P、S、T、W、Y和V，

[0788] 其中所述工程化AAV VP衣壳多肽能够组装成重组AAV病毒体(rAAV)，并且

[0789] 其中所述工程化AAV VP衣壳多肽不具有SEQ ID NO:1、SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7和SEQ ID NO:8中任一者的序列。

[0790] 4.如实施方案3所述的工程化AAV VP衣壳多肽，其中与具有SEQ ID NO:1的野生型AAV5 VP衣壳多肽的rAAV病毒体相比，所述rAAV具有较高的对肺组织的嗜性。

[0791] 5.如实施方案1所述的工程化AAV VP衣壳多肽，其中从对应于残基581的残基至对应于残基589的残基的区域，包括端点在内，具有与SEQ ID NO:16118-SEQ ID NO:19117中任一者至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%同一的序列。

[0792] 6.如实施方案5所述的工程化AAV VP衣壳多肽，其中从对应于残基581的残基至对应于残基589的残基的区域，包括端点在内，具有SEQ ID NO:16118-SEQ ID NO:19117中任一者的序列。

[0793] 7.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽，其中SEQ ID NO:2的Xaa1至Xaa9区域具有与SEQ ID NO:1 6118-SEQ ID NO:19117中任一者至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%同一的序列。

[0794] 8.如实施方案7所述的工程化AAV VP衣壳多肽，其中SEQ ID NO:2的Xaa1至Xaa9区域具有SEQ ID NO:16118-SEQ ID NO:19117中任一者的序列。

[0795] 9.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽，其中：

[0796] Xaa1选自由A、E、K、M、Q、R、S和T组成的组；或者

[0797] Xaa2选自由A、I、K、S、T和V组成的组；或者

[0798] Xaa3选自由A、E、K、M、Q、R、S、T和V组成的组；或者

[0799] Xaa4选自由M、P、R、S和T组成的组；或者

[0800] Xaa5选自由I、K、L、M、T、V和Y组成的组；或者

[0801] Xaa6选自由D、G、H、M、N、R和S组成的组;或者

[0802] Xaa7选自由A、K、M、Q和R组成的组;或者

[0803] Xaa8选自由A、F、G、S、W和Y组成的组;或者

[0804] Xaa9选自由A、E、G、P、R和Y组成的组;或者

[0805] 它们的任何组合。

[0806] 10.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa1选自A、E和Q。

[0807] 11.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa1为E。

[0808] 12.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa2选自S、T和V。

[0809] 13.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa2为T。

[0810] 14.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa3选自A、K、R和S。

[0811] 15.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa3为R。

[0812] 16.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa4选自P、Q和T。

[0813] 17.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa4为Q。

[0814] 18.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa5选自L、M和Y。

[0815] 19.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa5为L。

[0816] 20.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa6选自H和N。

[0817] 21.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa6为N。

[0818] 22.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa7选自A、K和R。

[0819] 23.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa7为R。

[0820] 24.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa8选自A、F和G。

[0821] 25.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa8为F。

[0822] 26.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa9选自G、P和R。

[0823] 27.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa9为G。

[0824] 28.如实施方案1-27中任一项所述的工程化AAV VP衣壳多肽,其中从对应于残基581的残基至对应于残基589的残基的区域,包括端点在内,具有与SEQ ID NO:16118-SEQ ID NO:17117中任一者至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%同一的序列。

[0825] 29.如实施方案28所述的工程化AAV VP衣壳多肽,其中SEQ ID NO:1的从残基581至残基589的区域具有SEQ ID NO:16118-SEQ ID NO:17117中任一者的序列。

[0826] 30.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa1具有高可突变性。

[0827] 31.如实施方案30所述的工程化AAV VP衣壳多肽,其中Xaa1选自D、E、M、A、I、Q和T。

[0828] 32.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa2具有高摩尔质量。

[0829] 33.如实施方案32所述的工程化AAV VP衣壳多肽,其中Xaa2为F。

[0830] 34.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa2具有低可突变性。

[0831] 35.如实施方案34所述的工程化AAV VP衣壳多肽,其中Xaa2选自Y、F和L。

[0832] 36.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa3具有低可突

变性。

[0833] 37.如实施方案36所述的工程化AAV VP衣壳多肽,其中Xaa3选自K、V、P和H。

[0834] 38.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa3具有低亲疏水性。

[0835] 39.如实施方案38所述的工程化AAV VP衣壳多肽,其中Xaa3选自K和R。

[0836] 40.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa4具有低可突变性。

[0837] 41.如实施方案40所述的工程化AAV VP衣壳多肽,其中Xaa4选自K和P。

[0838] 42.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa4具有高平均柔性。

[0839] 43.如实施方案42所述的工程化AAV VP衣壳多肽,其中Xaa4选自D、E、P和S。

[0840] 44.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa5具有低平均柔性。

[0841] 45.如实施方案44所述的工程化AAV VP衣壳多肽,其中Xaa5选自W、M和F。

[0842] 46.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa5具有高溶解性。

[0843] 47.如实施方案46所述的工程化AAV VP衣壳多肽,其中Xaa5选自W、F、I和L。

[0844] 48.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa6具有中等可突变性。

[0845] 49.如实施方案48所述的工程化AAV VP衣壳多肽,其中Xaa6选自R和H。

[0846] 50.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa6具有高表面可接近性。

[0847] 51.如实施方案50所述的工程化AAV VP衣壳多肽,其中Xaa6选自T。

[0848] 52.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa7具有低可突变性。

[0849] 53.如实施方案52所述的工程化AAV VP衣壳多肽,其中Xaa7为C。

[0850] 54.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa7具有高溶解性。

[0851] 55.如实施方案54所述的工程化AAV VP衣壳多肽,其中Xaa7选自W、V、M、F、I和L。

[0852] 56.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa8具有高可突变性。

[0853] 57.如实施方案56所述的工程化AAV VP衣壳多肽,其中Xaa8选自D、E、M、A、I、Q和T。

[0854] 58.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa8具有低亲疏水性。

[0855] 59.如实施方案58所述的工程化AAV VP衣壳多肽,其中Xaa8选自R和K。

[0856] 60.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa9具有高平均柔性。

[0857] 61.如实施方案60所述的工程化AAV VP衣壳多肽,其中Xaa9选自R和G。

[0858] 62.如实施方案1-61中任一项所述的工程化AAV VP衣壳多肽,其中从对应于残基

581的残基至对应于残基589的残基的区域包括端点在内具有与SEQ ID NO:17118-SEQ ID NO:18117中任一者具有至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%同一的序列。

[0859] 63.如实施方案62所述的工程化AAV VP衣壳多肽,其中从对应于残基581的残基至对应于残基589的残基的区域包括端点在内具有SEQ ID NO:17118-SEQ ID NO:18117中任一者的序列。

[0860] 64.如实施方案1-63中任一项所述的工程化AAV VP衣壳多肽,其中对肺组织的嗜性被测量为所述rAAV病毒体在肺组织中相较于非肺组织中的相对积聚,其中所述非肺组织由CNS、肝脏、骨骼肌、脾脏、淋巴结、骨髓、乳腺、皮肤、肾上腺、甲状腺、结肠、坐骨神经和脊髓共同组成。

[0861] 65.如实施方案64所述的工程化AAV VP衣壳多肽,其中所述CNS组织选自前脑皮层、枕叶皮层、颞叶皮层、丘脑、下丘脑、黑质、海马DG、海马CA1、海马CA3、小脑以及它们的任何组合。

[0862] 66.如实施方案64-65中任一项所述的工程化AAV VP衣壳多肽,其中所述较高的组织嗜性是在肺组织中相较于非肺组织中1.0005倍至约1000倍增加的积聚。

[0863] 67.如实施方案66所述的工程化AAV VP衣壳多肽,其中所述较高的组织嗜性是在肺组织中相较于非肺组织中至少约1.0005倍、至少约两倍、至少约三倍、至少约四倍、至少约五倍、至少约十倍、至少约二十倍、至少约50倍、至少约75倍、至少约100倍或至少约1000倍增加的积聚。

[0864] [411]J系列实施方案-淋巴结嗜性衣壳

[0865] 1.一种工程化腺相关病毒(AAV)病毒蛋白(VP)衣壳多肽,所述工程化腺相关病毒(AAV)病毒蛋白(VP)衣壳多肽具有与SEQ ID NO:1至少70%同一的氨基酸序列,

[0866] 其中所述工程化AAV VP衣壳多肽相较于SEQ ID NO:1在从对应于SEQ ID NO:1的残基581的残基至对应于SEQ ID NO:1的残基589的残基的区域,包括端点在内,中具有至少一个突变,

[0867] 其中所述工程化AAV VP衣壳多肽能够组装成重组AAV病毒体(rAAV),

[0868] 其中与具有SEQ ID NO:1的野生型AAV5 VP衣壳多肽的rAAV病毒体相比,所述至少一个突变赋予所述rAAV较高的对淋巴结组织的嗜性,并且

[0869] 其中所述工程化AAV VP衣壳多肽不具有SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7和SEQ ID NO:8中任一者的序列。2.如实施方案1所述的工程化AAV VP衣壳多肽,其中所述工程化AAV VP衣壳多肽具有SEQ ID NO:2的序列,并且其中Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9各自独立地选自任何氨基酸。3.一种工程化腺相关病毒(AAV)病毒蛋白(VP)衣壳多肽,所述工程化腺相关病毒(AAV)病毒蛋白(VP)衣壳多肽具有SEQ ID NO:2的氨基酸序列,

[0870] 其中氨基酸残基Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9各自独立地选自A、R、N、D、C、E、Q、G、H、I、L、K、M、F、P、S、T、W、Y和V,

[0871] 其中所述工程化AAV VP衣壳多肽能够组装成重组AAV病毒体(rAAV),并且

[0872] 其中所述工程化AAV VP衣壳多肽不具有SEQ ID NO:1、SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7和SEQ ID NO:8中任一者的序列。4.如实施方案

3所述的工程化AAV VP衣壳多肽，其中与具有SEQ ID NO:1的野生型AAV5 VP衣壳多肽的rAAV病毒体相比，所述rAAV具有较高的对淋巴结组织的嗜性。5.如实施方案1所述的工程化AAV VP衣壳多肽，其中从对应于残基581的残基至对应于残基589的残基的区域，包括端点在内，具有与SEQ ID NO:19118-SEQ ID NO:22117中任一者至少70％、至少75％、至少80％、至少85％、至少90％、至少92％、至少95％、至少97％或至少99％同一的序列。6.如实施方案5所述的工程化AAV VP衣壳多肽，其中从对应于残基581的残基至对应于残基589的残基的区域，包括端点在内，具有SEQ ID NO:19118-SEQ ID NO:22117中任一者的序列。7.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽，其中SEQ ID NO:2的Xaa1至Xaa9区域具有与SEQ ID NO:19118-SEQ ID NO:22117中任一者至少70％、至少75％、至少80％、至少85％、至少90％、至少92％、至少95％、至少97％或至少99％同一的序列。8.如实施方案7所述的工程化AAV VP衣壳多肽，其中SEQ ID NO:2的Xaa1至Xaa9区域具有SEQ ID NO:19118-SEQ ID NO:22117中任一者的序列。9.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽，其中：

[0873] Xaa1选自由A、D、E、Q、S和T组成的组；或者

[0874] Xaa2选自由A、H、I、S、T和V组成的组；或者

[0875] Xaa3选自由A、E、H、I、T和V组成的组；或者

[0876] Xaa4选自由A、D、E和P组成的组；或者

[0877] Xaa5选自由I、L、M、V和Y组成的组；或者

[0878] Xaa6选自由D、E、I、N和Q组成的组；或者

[0879] Xaa7选自由A、E、G、Q和V组成的组；或者

[0880] Xaa8选自由F、G、M和W组成的组；或者

[0881] Xaa9选自由I、P、T和Y组成的组；或者

[0882] 它们的任何组合。

[0883] 10.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa1选自D、E和T。

[0884] 11.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa1为E。

[0885] 12.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa2选自I、T和V。

[0886] 13.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa2为V。

[0887] 14.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa3选自A、I、T和V。

[0888] 15.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa3为T。

[0889] 16.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa4选自D和E。

[0890] 17.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa4为E。

[0891] 18.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa5选自I、L、V和Y。

[0892] 19.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa5为L。

[0893] 20.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa6选自D、E和I。

[0894] 21.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa6为D。

[0895] 22.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa7为A、Q或V。

[0896] 23.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa7为V。

[0897] 24.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa8选自F和W。

[0898] 25.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa8为W。

[0899] 26.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa9为I或P。

[0900] 27.如实施方案1-26中任一项所述的工程化AAV VP衣壳多肽,其中从对应于残基581的残基至对应于残基589的残基的区域,包括端点在内,具有与SEQ ID NO:19118-SEQ ID NO:20117中任一者至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%同一的序列。

[0901] 28.如实施方案27所述的工程化AAV VP衣壳多肽,其中SEQ ID NO:1的从残基581至残基589的区域具有SEQ ID NO:19118-SEQ ID NO:20117中任一者的序列。

[0902] 29.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa1具有高平均柔性。

[0903] 30.如实施方案29所述的工程化AAV VP衣壳多肽,其中Xaa1选自D、E、P、G、Q、S和R。

[0904] 31.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa1具有高氢键供体。

[0905] 32.如实施方案31所述的工程化AAV VP衣壳多肽,其中Xaa1为R。

[0906] 33.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa1具有高摩尔质量。

[0907] 34.如实施方案33所述的工程化AAV VP衣壳多肽,其中Xaa1选自Y、W、R和F。

[0908] 35.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa2具有低溶解性。

[0909] 36.如实施方案35所述的工程化AAV VP衣壳多肽,其中Xaa2选自N和E。

[0910] 37.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa3具有低平均柔性。

[0911] 38.如实施方案37所述的工程化AAV VP衣壳多肽,其中Xaa3选自W、M和F。

[0912] 39.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa3具有低可突变性。

[0913] 40.如实施方案39所述的工程化AAV VP衣壳多肽,其中Xaa3选自R、H、K、P、Y、F、L和C。

[0914] 41.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa4具有低可突变性。

[0915] 42.如实施方案41所述的工程化AAV VP衣壳多肽,其中Xaa4为C。

[0916] 43.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa5具有高可突变性。

[0917] 44.如实施方案43所述的工程化AAV VP衣壳多肽,其中Xaa5为N。

[0918] 45.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa5具有中等摩尔质量。

[0919] 46.如实施方案45所述的工程化AAV VP衣壳多肽,其中Xaa5选自D、I、L和N。

[0920] 47.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa6具有高摩尔质量。

[0921] 48.如实施方案47所述的工程化AAV VP衣壳多肽,其中Xaa6选自Y、W、R和F。

[0922] 49.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa6具有高平均

柔性。

[0923] 50.如实施方案49所述的工程化AAV VP衣壳多肽,其中Xaa6选自G和R。

[0924] 51.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa7具有高平均柔性。

[0925] 52.如实施方案51所述的工程化AAV VP衣壳多肽,其中Xaa7选自D、E、K、P、I、N、Q和S。

[0926] 53.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa7具有低溶解性。

[0927] 54.如实施方案53所述的工程化AAV VP衣壳多肽,其中Xaa7选自N和E。

[0928] 55.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa8具有低溶解性。

[0929] 56.如实施方案55所述的工程化AAV VP衣壳多肽,其中Xaa8选自N、E和D。

[0930] 57.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa8具有中等可突变性。

[0931] 58.如实施方案57所述的工程化AAV VP衣壳多肽,其中Xaa8选自R和H。

[0932] 59.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa9具有低可突变性。

[0933] 60.如实施方案59所述的工程化AAV VP衣壳多肽,其中Xaa9选自P和K。

[0934] 61.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa9具有高平均柔性。

[0935] 62.如实施方案61所述的工程化AAV VP衣壳多肽,其中Xaa9选自D、E、P和S。

[0936] 63.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa9具有高溶解性。

[0937] 64.如实施方案63所述的工程化AAV VP衣壳多肽,其中Xaa9选自M和V。

[0938] 65.如实施方案1-64中任一项所述的工程化AAV VP衣壳多肽,其中从对应于残基581的残基至对应于残基589的残基的区域包括端点在内具有与SEQ ID NO:20118-SEQ ID NO:21117中任一者具有至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%同一的序列。

[0939] 66.如实施方案65所述的工程化AAV VP衣壳多肽,其中从对应于残基581的残基至对应于残基589的残基的区域包括端点在内具有SEQ ID NO:20118-SEQ ID NO:21117中任一者的序列。

[0940] 67.如实施方案1-66中任一项所述的工程化AAV VP衣壳多肽,其中对淋巴结组织的嗜性被测量为所述rAAV病毒体在淋巴结组织中相较于非淋巴结组织中的相对积聚,其中所述非淋巴结组织由CNS、肝脏、骨骼肌、心脏、肺、脾脏、骨髓、乳腺、皮肤、肾上腺、甲状腺、结肠、坐骨神经和脊髓共同组成。

[0941] 68.如实施方案67所述的工程化AAV VP衣壳多肽,其中所述CNS组织选自前脑皮层、枕叶皮层、颞叶皮层、丘脑、下丘脑、黑质、海马DG、海马CA1、海马CA3、小脑以及它们的任何组合。

[0942] 69.如实施方案67-68中任一项所述的工程化AAV VP衣壳多肽,其中所述较高的组

织嗜性是在淋巴结组织中相较于非淋巴结组织中1.0005倍至约1000倍增加的积聚。

[0943] 70.如实施方案69所述的工程化AAV VP衣壳多肽,其中所述较高的组织嗜性是在淋巴结组织中相较于非淋巴结组织中至少约1.0005倍、至少约两倍、至少约三倍、至少约四倍、至少约五倍、至少约十倍、至少约二十倍、至少约50倍、至少约75倍、至少约100倍或至少约1000倍增加的积聚。

[0944] [412]K系列实施方案-乳腺嗜性衣壳

[0945] 1.一种工程化腺相关病毒(AAV)病毒蛋白(VP)衣壳多肽,所述工程化腺相关病毒(AAV)病毒蛋白(VP)衣壳多肽具有与SEQ ID NO:1至少70%同一的氨基酸序列,

[0946] 其中所述工程化AAV VP衣壳多肽相较于SEQ ID NO:1在从对应于SEQ ID NO:1的残基581的残基至对应于SEQ ID NO:1的残基589的残基的区域,包括端点在内,中具有至少一个突变,

[0947] 其中所述工程化AAV VP衣壳多肽能够组装成重组AAV病毒体(rAAV),

[0948] 其中与具有SEQ ID NO:1的野生型AAV5 VP衣壳多肽的rAAV病毒体相比,所述至少一个突变赋予所述rAAV较高的对乳腺组织的嗜性,并且

[0949] 其中所述工程化AAV VP衣壳多肽不具有SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7和SEQ ID NO:8中任一者的序列。2.如实施方案1所述的工程化AAV VP衣壳多肽,其中所述工程化AAV VP衣壳多肽具有SEQ ID NO:2的序列,并且其中Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9各自独立地选自任何氨基酸。3.一种工程化腺相关病毒(AAV)病毒蛋白(VP)衣壳多肽,所述工程化腺相关病毒(AAV)病毒蛋白(VP)衣壳多肽具有SEQ ID NO:2的氨基酸序列,

[0950] 其中氨基酸残基Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9各自独立地选自A、R、N、D、C、E、Q、G、H、I、L、K、M、F、P、S、T、W、Y和V,

[0951] 其中所述工程化AAV VP衣壳多肽能够组装成重组AAV病毒体(rAAV),并且

[0952] 其中所述工程化AAV VP衣壳多肽不具有SEQ ID NO:1、SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7和SEQ ID NO:8中任一者的序列。4.如实施方案3所述的工程化AAV VP衣壳多肽,其中与具有SEQ ID NO:1的野生型AAV5 VP衣壳多肽的rAAV病毒体相比,所述rAAV具有较高的对乳腺组织的嗜性。

[0953] 5.如实施方案1所述的工程化AAV VP衣壳多肽,其中从对应于残基581的残基至对应于残基589的残基的区域,包括端点在内,具有与SEQ ID NO:22118-SEQ ID NO:25117中任一者至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%同一的序列。

[0954] 6.如实施方案5所述的工程化AAV VP衣壳多肽,其中从对应于残基581的残基至对应于残基589的残基的区域,包括端点在内,具有SEQ ID NO:22118-SEQ ID NO:25117中任一者的序列。

[0955] 7.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中SEQ ID NO:2的Xaa1至Xaa9区域具有与SEQ ID NO:22118-SEQ ID NO:25117中任一者至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%同一的序列。

[0956] 8.如实施方案7所述的工程化AAV VP衣壳多肽,其中SEQ ID NO:2的Xaa1至Xaa9区域具有SEQ ID NO:22118-SEQ ID NO:25117中任一者的序列。

[0957] 9.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中:

[0958] Xaa1选自由C、K、M、Q、R和Y组成的组;或者

[0959] Xaa2选自由A、F、I、K、S、T和V组成的组;或者

[0960] Xaa3选自由A、F、G、I、K、L、R、T和Y组成的组;或者

[0961] Xaa4选自由A、I、K、Q、R和T组成的组;或者

[0962] Xaa5选自由I、L、M、Q、R、T、V和Y组成的组;或者

[0963] Xaa6选自由H、N、S和V组成的组;或者

[0964] Xaa7选自由A、H、I、N、S和Y组成的组;或者

[0965] Xaa8选自由A、C、D、G、H、M、Q和S组成的组;或者

[0966] Xaa9选自由A、E、L、W和Y组成的组;或者

[0967] 它们的任何组合。

[0968] 10.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa1选自C、Q和R。

[0969] 11.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa1为C。

[0970] 12.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa2选自A、S和V。

[0971] 13.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa2为V。

[0972] 14.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa3选自F、G、K、R和Y。

[0973] 15.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa3选自F、K和Y。

[0974] 16.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa3为F。

[0975] 17.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa4选自A、I和R。

[0976] 18.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa4为I。

[0977] 19.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa5选自I、M和Y。

[0978] 20.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa5为Y。

[0979] 21.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa6为H。

[0980] 22.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa7为N或S。

[0981] 23.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa7为N。

[0982] 24.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa8选自G、M和Q。

[0983] 25.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa8为G。

[0984] 26.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa9选自A、L和W。

[0985] 27.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa9为A。

[0986] 28.如实施方案1-27中任一项所述的工程化AAV VP衣壳多肽,其中从对应于残基581的残基至对应于残基589的残基的区域,包括端点在内,具有与SEQ ID NO:22118-SEQ ID NO:23117中任一者至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%同一的序列。

[0987] 29.如实施方案28所述的工程化AAV VP衣壳多肽,其中SEQ ID NO:1的从残基581至残基589的区域具有SEQ ID NO:22118-SEQ ID NO:23117中任一者的序列。

[0988] 30.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa1具有低表面可接近性。

[0989] 31.如实施方案30所述的工程化AAV VP衣壳多肽,其中Xaa1为C。

[0990] 32.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa1具有中等摩

尔质量。

[0991] 33.如实施方案32所述的工程化AAV VP衣壳多肽,其中Xaa1为C。

[0992] 34.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa2具有高表面可接近性。

[0993] 35.如实施方案34所述的工程化AAV VP衣壳多肽,其中Xaa2选自D、N和Q。

[0994] 36.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa2具有低亲疏水性。

[0995] 37.如实施方案36所述的工程化AAV VP衣壳多肽,其中Xaa2选自D、E、R、K、H、N和Q。

[0996] 38.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa3具有高平均柔性。

[0997] 39.如实施方案38所述的工程化AAV VP衣壳多肽,其中Xaa3选自D、E、R、P、G和S。

[0998] 40.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa3具有中等可突变性。

[0999] 41.如实施方案40所述的工程化AAV VP衣壳多肽,其中Xaa3选自R和H。

[1000] 42.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa4具有高可突变性。

[1001] 43.如实施方案42所述的工程化AAV VP衣壳多肽,其中Xaa4选自M、I、Q和T。

[1002] 44.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa4具有高溶解性。

[1003] 45.如实施方案44所述的工程化AAV VP衣壳多肽,其中Xaa4选自W、F、I和L。

[1004] 46.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa4具有高表面可接近性。

[1005] 47.如实施方案46所述的工程化AAV VP衣壳多肽,其中Xaa4选自E、R和K。

[1006] 48.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa5具有高溶解性。

[1007] 49.如实施方案48所述的工程化AAV VP衣壳多肽,其中Xaa5选自W、F、I和L。

[1008] 50.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa5具有低可突变性。

[1009] 51.如实施方案50所述的工程化AAV VP衣壳多肽,其中Xaa5选自Y、F和L。

[1010] 52.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa6具有高亲疏水性。

[1011] 53.如实施方案52所述的工程化AAV VP衣壳多肽,其中Xaa6选自V、I和L。

[1012] 54.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa6具有中等摩尔质量。

[1013] 55.如实施方案54所述的工程化AAV VP衣壳多肽,其中Xaa6选自D、I、L和N。

[1014] 56.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa8具有低表面可接近性。

[1015] 57.如实施方案56所述的工程化AAV VP衣壳多肽,其中Xaa8为C。

[1016] 58.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa8具有低可突

变性。

[1017] 59.如实施方案58所述的工程化AAV VP衣壳多肽,其中Xaa8选自C、R和H。

[1018] 60.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa8具有中等可突变性。

[1019] 61.如实施方案60所述的工程化AAV VP衣壳多肽,其中Xaa9选自R和H。

[1020] 62.如实施方案1-61中任一项所述的工程化AAV VP衣壳多肽,其中从对应于残基581的残基至对应于残基589的残基的区域包括端点在内具有与SEQ ID NO:23118-SEQ ID NO:24117中任一者具有至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%同一的序列。

[1021] 63.如实施方案62所述的工程化AAV VP衣壳多肽,其中从对应于残基581的残基至对应于残基589的残基的区域包括端点在内具有SEQ ID NO:23118—SEQ ID NO:24117中任一者的序列。

[1022] 64.如实施方案1-63中任一项所述的工程化AAV VP衣壳多肽,其中对乳腺组织的嗜性被测量为所述rAAV病毒体在乳腺组织中相较于非乳腺组织中的相对积聚,其中所述非乳腺组织由CNS、肝脏、骨骼肌、心脏、肺、脾脏、淋巴结、骨髓、皮肤、肾上腺、甲状腺、结肠、坐骨神经和脊髓共同组成。

[1023] 65.如实施方案65所述的工程化AAV VP衣壳多肽,其中所述CNS组织选自前脑皮层、枕叶皮层、颞叶皮层、丘脑、下丘脑、黑质、海马DG、海马CA1、海马CA3、小脑以及它们的任何组合。

[1024] 66.如实施方案64-65中任一项所述的工程化AAV VP衣壳多肽,其中所述较高的组织嗜性是在乳腺组织中相较于非乳腺组织中1.0005倍至约1000倍增加的积聚。

[1025] 67.如实施方案66所述的工程化AAV VP衣壳多肽,其中所述较高的组织嗜性是在乳腺组织中相较于非乳腺组织中至少约1.0005倍、至少约两倍、至少约三倍、至少约四倍、至少约五倍、至少约十倍、至少约二十倍、至少约50倍、至少约75倍、至少约100倍或至少约1000倍增加的积聚。

[1026] [413]L系列实施方案-骨骼肌或心肌嗜性衣壳

[1027] 1.一种工程化腺相关病毒(AAV)病毒蛋白(VP)衣壳多肽,所述工程化腺相关病毒(AAV)病毒蛋白(VP)衣壳多肽具有与SEQ ID NO:1至少70%同一的氨基酸序列,

[1028] 其中所述工程化AAV VP衣壳多肽相较于SEQ ID NO:1在从对应于SEQ ID NO:1的残基581的残基至对应于SEQ ID NO:1的残基589的残基的区域,包括端点在内,中具有至少一个突变,

[1029] 其中所述工程化AAV VP衣壳多肽能够组装成重组AAV病毒体(rAAV),

[1030] 其中与具有SEQ ID NO:1的野生型AAV5 VP衣壳多肽的rAAV病毒体相比,所述至少一个突变赋予所述rAAV较高的对骨骼肌组织或心肌组织的嗜性,并且

[1031] 其中所述工程化AAV VP衣壳多肽不具有SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7和SEQ ID NO:8中任一者的序列。2.如实施方案1所述的工程化AAV VP衣壳多肽,其中所述工程化AAV VP衣壳多肽具有SEQ ID NO:2的序列,并且其中Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9各自独立地选自任何氨基酸。3.一种工程化腺相关病毒(AAV)病毒蛋白(VP)衣壳多肽,所述工程化腺相关病毒(AAV)病毒蛋白

(VP)衣壳多肽具有SEQ ID NO:2的氨基酸序列，

[1032] 其中氨基酸残基Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9各自独立地选自A、R、N、D、C、E、Q、G、H、I、L、K、M、F、P、S、T、W、Y和V，

[1033] 其中所述工程化AAV VP衣壳多肽能够组装成重组AAV病毒体(rAAV)，并且

[1034] 其中所述工程化AAV VP衣壳多肽不具有SEQ ID NO:1、SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7和SEQ ID NO:8中任一者的序列。4.如实施方案3所述的工程化AAV VP衣壳多肽，其中与具有SEQ ID NO:1的野生型AAV5 VP衣壳多肽的rAAV病毒体相比，所述rAAV具有较高的对骨骼肌组织或心肌组织的嗜性。

[1035] 5.如实施方案1所述的工程化AAV VP衣壳多肽，其中从对应于残基581的残基至对应于残基589的残基的区域，包括端点在内，具有与SEQ ID NO:25118-SEQ ID NO:26117中任一者至少70％、至少75％、至少80％、至少85％、至少90％、至少92％、至少95％、至少97％或至少99％同一的序列。

[1036] 6.如实施方案5所述的工程化AAV VP衣壳多肽，其中从对应于残基581的残基至对应于残基589的残基的区域，包括端点在内，具有SEQ ID NO:25118-SEQ ID NO:26117中任一者的序列。

[1037] 7.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽，其中SEQ ID NO:2的Xaa1至Xaa9区域具有与SEQ ID NO:25118-SEQ ID NO:26117中任一者至少70％、至少75％、至少80％、至少85％、至少90％、至少92％、至少95％、至少97％或至少99％同一的序列。

[1038] 8.如实施方案7所述的工程化AAV VP衣壳多肽，其中SEQ ID NO:2的Xaa1至Xaa9区域具有SEQ ID NO:25118-SEQ ID NO:26117中任一者的序列。

[1039] 9.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽，其中Xaa1具有低溶解性。

[1040] 10.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa1选自D、E、R、K、P、N和Q。

[1041] 11.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽，其中Xaa1具有低亲疏水性。

[1042] 12.如实施方案11所述的工程化AAV VP衣壳多肽，其中Xaa1选自D、E、R、K、Q、N、Y、P。

[1043] 13.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽，其中Xaa1具有高表面可接近性。

[1044] 14.如实施方案13所述的工程化AAV VP衣壳多肽，其中Xaa1选自E、R和K。

[1045] 15.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽，其中Xaa2具有高亲疏水性。

[1046] 16.如实施方案15所述的工程化AAV VP衣壳多肽，其中Xaa2选自V、I、F、L和C。

[1047] 17.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽，其中Xaa2具有低可突变性。

[1048] 18.如实施方案17所述的工程化AAV VP衣壳多肽，其中Xaa2选自R、V、I、H和C。

[1049] 19.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽，其中Xaa2具有中等体积。

[1050] 20.如实施方案19所述的工程化AAV VP衣壳多肽，其中Xaa2选自E、V和Q。

[1051] 21.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa3具有低溶解性。

[1052] 22.如实施方案21所述的工程化AAV VP衣壳多肽,其中Xaa3选自D、R和Q。

[1053] 23.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa4具有低溶解性。

[1054] 24.如实施方案23所述的工程化AAV VP衣壳多肽,其中Xaa4选自D、E、P和N。

[1055] 25.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa4具有低电荷。

[1056] 26.如实施方案25所述的工程化AAV VP衣壳多肽,其中Xaa4选自D和E。

[1057] 27.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa5具有低氨基酸溶解性。

[1058] 28.如实施方案27所述的工程化AAV VP衣壳多肽,其中Xaa5选自D、E、R、K、N、Q。

[1059] 29.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa8具有低溶解性。

[1060] 30.如实施方案29所述的工程化AAV VP衣壳多肽,其中Xaa8选自D、E、K、P和N。

[1061] 31.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa8具有高柔性指数。

[1062] 32.如实施方案31所述的工程化AAV VP衣壳多肽,其中Xaa8选自Q、S、P、E和D。

[1063] 33.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa8具有高表面可接近性。

[1064] 34.如实施方案33所述的工程化AAV VP衣壳多肽,其中Xaa8选自S、D、P、N、E、R和K。

[1065] 35.如实施方案1-34中任一项所述的工程化AAV VP衣壳多肽,其中从对应于残基581的残基至对应于残基589的残基的区域包括端点在内具有与SEQ ID NO:25118-SEQ ID NO:26117中任一者具有至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%同一的序列。

[1066] 36.如实施方案35所述的工程化AAV VP衣壳多肽,其中从对应于残基581的残基至对应于残基589的残基的区域包括端点在内具有SEQ ID NO:25118—SEQ ID NO:26117中任一者的序列。

[1067] 37.如实施方案1-36中任一项所述的工程化AAV VP衣壳多肽,其中对骨骼肌组织或心肌组织的嗜性被测量为所述rAAV病毒体在骨骼肌组织或心肌组织中相较于非骨骼肌组织或非心肌组织中的相对积聚,其中所述非骨骼肌组织或非心肌组织由CNS、肝脏、肺、脾脏、淋巴结、骨髓、乳腺、皮肤、肾上腺、甲状腺、结肠、坐骨神经和脊髓共同组成。

[1068] 38.如实施方案37所述的工程化AAV VP衣壳多肽,其中所述CNS组织选自前脑皮层、枕叶皮层、颞叶皮层、丘脑、下丘脑、黑质、海马DG、海马CA1、海马CA3、小脑以及它们的任何组合。

[1069] 39.如实施方案37-38中任一项所述的工程化AAV VP衣壳多肽,其中所述较高的组织嗜性是在骨骼肌组织或心肌组织中相较于非骨骼肌组织或非心肌组织中1.0005倍至约1000倍增加的积聚。

[1070] 40.如实施方案39所述的工程化AAV VP衣壳多肽,其中所述较高的组织嗜性是在

骨骼肌组织或心肌组织中相较于非骨骼肌组织或非心肌组织中至少约1.0005倍、至少约两倍、至少约三倍、至少约四倍、至少约五倍、至少约十倍、至少约二十倍、至少约50倍、至少约75倍、至少约100倍或至少约1000倍增加的积聚。

[1071] [414]M系列实施方案-坐骨神经嗜性衣壳

[1072] 1.一种工程化腺相关病毒(AAV)病毒蛋白(VP)衣壳多肽,所述工程化腺相关病毒(AAV)病毒蛋白(VP)衣壳多肽具有与SEQ ID NO:1至少70%同一的氨基酸序列,

[1073] 其中所述工程化AAV VP衣壳多肽相较于SEQ ID NO:1在从对应于SEQ ID NO:1的残基581的残基至对应于SEQ ID NO:1的残基589的残基的区域,包括端点在内,中具有至少一个突变,

[1074] 其中所述工程化AAV VP衣壳多肽能够组装成重组AAV病毒体(rAAV),

[1075] 其中与具有SEQ ID NO:1的野生型AAV5 VP衣壳多肽的rAAV病毒体相比,所述至少一个突变赋予所述rAAV较高的对坐骨神经组织的嗜性,并且

[1076] 其中所述工程化AAV VP衣壳多肽不具有SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7和SEQ ID NO:8中任一者的序列。2.如实施方案1所述的工程化AAV VP衣壳多肽,其中所述工程化AAV VP衣壳多肽具有SEQ ID NO:2的序列,并且其中Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9各自独立地选自任何氨基酸。3.一种工程化腺相关病毒(AAV)病毒蛋白(VP)衣壳多肽,所述工程化腺相关病毒(AAV)病毒蛋白(VP)衣壳多肽具有SEQ ID NO:2的氨基酸序列,

[1077] 其中氨基酸残基Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9各自独立地选自A、R、N、D、C、E、Q、G、H、I、L、K、M、F、P、S、T、W、Y和V,

[1078] 其中所述工程化AAV VP衣壳多肽能够组装成重组AAV病毒体(rAAV),并且

[1079] 其中所述工程化AAV VP衣壳多肽不具有SEQ ID NO:1、SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7和SEQ ID NO:8中任一者的序列。4.如实施方案3所述的工程化AAV VP衣壳多肽,其中与具有SEQ ID NO:1的野生型AAV5 VP衣壳多肽的rAAV病毒体相比,所述rAAV具有较高的对坐骨神经组织的嗜性。5.如实施方案1所述的工程化AAV VP衣壳多肽,其中从对应于残基581的残基至对应于残基589的残基的区域,包括端点在内,具有与SEQ ID NO:26118-SEQ ID NO:28990中任一者至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%同一的序列。

[1080] 6.如实施方案5所述的工程化AAV VP衣壳多肽,其中从对应于残基581的残基至对应于残基589的残基的区域,包括端点在内,具有SEQ ID NO:26118-SEQ ID NO:28990中任一者的序列。

[1081] 7.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中SEQ ID NO:2的Xaa1至Xaa9区域具有与SEQ ID NO:26118-SEQ ID NO:28990中任一者至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%同一的序列。

[1082] 8.如实施方案7所述的工程化AAV VP衣壳多肽,其中SEQ ID NO:2的Xaa1至Xaa9区域具有SEQ ID NO:26118-SEQ ID NO:28990中任一者的序列。

[1083] 9.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中:

[1084] Xaa1选自由C、G、K、M、Q、R和Y组成的组;或者

[1085] Xaa2选自由A、C、F、I、Q、T和V组成的组;或者

[1086] Xaa3选自由A、F、I、M、R、S和T组成的组;或者

[1087] Xaa4选自由E、N、T、Q和V组成的组;或者

[1088] Xaa5选自由F、H、Q、S、V和Y组成的组;或者

[1089] Xaa6选自由K、M、N、Q、S和V组成的组;或者

[1090] Xaa7选自由K、M、Q、R和T组成的组;或者

[1091] Xaa8选自由A、G、H、Q、S和V组成的组;或者

[1092] Xaa9选自由C、E、I、K和R组成的组;或者

[1093] 它们的任何组合。

[1094] 10.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa1选自C、R和Q。

[1095] 11.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa1为C。

[1096] 12.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa2选自A、C和I。

[1097] 13.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa2为A。

[1098] 14.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa3选自F、M、R和S。

[1099] 15.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa3为R。

[1100] 16.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa4选自E、T和V。

[1101] 17.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa4为T。

[1102] 18.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa5选自F、V和Y。

[1103] 19.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa5为V。

[1104] 20.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa6选自M、N和S。

[1105] 21.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa6为N。

[1106] 22.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa7选自M、Q和T。

[1107] 23.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa7为M。

[1108] 24.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa8选自H和S。

[1109] 25.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa8为H。

[1110] 26.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa9选自C、I和K。

[1111] 27.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa9为I。

[1112] 28.如实施方案1-27中任一项所述的工程化AAV VP衣壳多肽,其中从对应于残基581的残基至对应于残基589的残基的区域,包括端点在内,具有与SEQ ID NO:26118-SEQ ID NO:26990中任一者至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%同一的序列。

[1113] 29.如实施方案28所述的工程化AAV VP衣壳多肽,其中SEQ ID NO:1的从残基581至残基589的区域具有SEQ ID NO:26118-SEQ ID NO:26990中任一者的序列。

[1114] 30.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa1具有高平均柔性。

[1115] 31.如实施方案30所述的工程化AAV VP衣壳多肽,其中Xaa1选自G和R。

[1116] 32.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa1具有低溶解性。

[1117] 33.如实施方案32所述的工程化AAV VP衣壳多肽,其中Xaa1选自R和Q。

[1118] 34.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa1具有低可突

变性。

[1119] 35.如实施方案34所述的工程化AAV VP衣壳多肽,其中Xaa1选自C、L、F、Y、R、K、P和H。

[1120] 36.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa1具有高体积。

[1121] 37.如实施方案36所述的工程化AAV VP衣壳多肽,其中Xaa1选自Y和F。

[1122] 38.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa2具有高表面可接近性。

[1123] 39.如实施方案38所述的工程化AAV VP衣壳多肽,其中Xaa2选自E、R和K。

[1124] 40.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa3具有中等可突变性。

[1125] 41.如实施方案40所述的工程化AAV VP衣壳多肽,其中Xaa3选自H和R。

[1126] 42.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa3具有中等平均柔性。

[1127] 43.如实施方案42所述的工程化AAV VP衣壳多肽,其中Xaa3选自V和Y。

[1128] 44.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa4具有高可突变性。

[1129] 45.如实施方案44所述的工程化AAV VP衣壳多肽,其中Xaa4为N。

[1130] 46.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa4具有高平均柔性。

[1131] 47.如实施方案46所述的工程化AAV VP衣壳多肽,其中Xaa4选自I、N、G和R。

[1132] 48.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa4具有低溶解性。

[1133] 49.如实施方案48所述的工程化AAV VP衣壳多肽,其中Xaa4为N。

[1134] 50.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa6具有低可突变性。

[1135] 51.如实施方案50所述的工程化AAV VP衣壳多肽,其中Xaa6选自C、L、F和Y。

[1136] 52.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa6具有高体积。

[1137] 53.如实施方案52所述的工程化AAV VP衣壳多肽,其中Xaa6选自K、M、I和L。

[1138] 54.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa7具有低可突变性。

[1139] 55.如实施方案54所述的工程化AAV VP衣壳多肽,其中Xaa7选自L、F和Y。

[1140] 56.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa7具有中等摩尔质量。

[1141] 57.如实施方案56所述的工程化AAV VP衣壳多肽,其中Xaa7选自D、I、L和N。

[1142] 58.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa8具有高表面可接近性。

[1143] 59.如实施方案58所述的工程化AAV VP衣壳多肽,其中Xaa8选自S、Y、T、D、P、H和N。

[1144] 60.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa9具有低可突变性。

[1145] 61.如实施方案60所述的工程化AAV VP衣壳多肽,其中Xaa9选自C、H和R。

[1146] 62.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa9具有中等溶解性。

[1147] 63.如实施方案62所述的工程化AAV VP衣壳多肽,其中Xaa9选自Q、T和C。

[1148] 64.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa9具有低表面可接近性。

[1149] 65.如实施方案64所述的工程化AAV VP衣壳多肽,其中Xaa9为C。

[1150] 66.如实施方案1-65中任一项所述的工程化AAV VP衣壳多肽,其中从对应于残基581的残基至对应于残基589的残基的区域包括端点在内具有与SEQ ID NO:26991-SEQ ID NO:27990中任一者具有至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%同一的序列。

[1151] 67.如实施方案66所述的工程化AAV VP衣壳多肽,其中从对应于残基581的残基至对应于残基589的残基的区域包括端点在内具有SEQ ID NO:26991-SEQ ID NO:27990中任一者的序列。

[1152] 68.如实施方案1-67中任一项所述的工程化AAV VP衣壳多肽,其中对坐骨神经组织的嗜性被测量为所述rAAV病毒体在坐骨神经组织中相较于非坐骨神经组织中的相对积聚,其中所述非坐骨神经组织由CNS、肝脏、骨骼肌、心脏、肺、脾脏、淋巴结、骨髓、乳腺、皮肤、肾上腺、甲状腺、结肠和脊髓共同组成。

[1153] 69.如实施方案68中任一项所述的工程化AAV VP衣壳多肽,其中所述CNS组织选自前脑皮层、枕叶皮层、颞叶皮层、丘脑、下丘脑、黑质、海马DG、海马CA1、海马CA3、小脑以及它们的任何组合。

[1154] 70.如实施方案68-69中任一项所述的工程化AAV VP衣壳多肽,其中所述较高的组织嗜性是在坐骨神经组织中相较于非坐骨神经组织中1.0005倍至约1000倍增加的积聚。

[1155] 71.如实施方案70所述的工程化AAV VP衣壳多肽,其中所述较高的组织嗜性是在坐骨神经组织中相较于非坐骨神经组织中至少约1.0005倍、至少约两倍、至少约三倍、至少约四倍、至少约五倍、至少约十倍、至少约二十倍、至少约50倍、至少约75倍、至少约100倍或至少约1000倍增加的积聚。

[1156] [415]N系列实施方案-骨骼肌嗜性衣壳

[1157] 1.一种工程化腺相关病毒(AAV)病毒蛋白(VP)衣壳多肽,所述工程化腺相关病毒(AAV)病毒蛋白(VP)衣壳多肽具有与SEQ ID NO:1至少70%同一的氨基酸序列,

[1158] 其中所述工程化AAV VP衣壳多肽相较于SEQ ID NO:1在从对应于SEQ ID NO:1的残基581的残基至对应于SEQ ID NO:1的残基589的残基的区域,包括端点在内,中具有至少一个突变,

[1159] 其中所述工程化AAV VP衣壳多肽能够组装成重组AAV病毒体(rAAV),

[1160] 其中与具有SEQ ID NO:1的野生型AAV5 VP衣壳多肽的rAAV病毒体相比,所述至少一个突变赋予所述rAAV较高的对骨骼肌组织的嗜性,并且

[1161] 其中所述工程化AAV VP衣壳多肽不具有SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:

5、SEQ ID NO:6、SEQ ID NO:7和SEQ ID NO:8中任一者的序列。2.如实施方案1所述的工程化AAV VP衣壳多肽，其中所述工程化AAV VP衣壳多肽具有SEQ ID NO:2的序列，并且其中Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9各自独立地选自任何氨基酸。3.一种工程化腺相关病毒(AAV)病毒蛋白(VP)衣壳多肽，所述工程化腺相关病毒(AAV)病毒蛋白(VP)衣壳多肽具有SEQ ID NO:2的氨基酸序列，

[1162] 其中氨基酸残基Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9各自独立地选自A、R、N、D、C、E、Q、G、H、I、L、K、M、F、P、S、T、W、Y和V，

[1163] 其中所述工程化AAV VP衣壳多肽能够组装成重组AAV病毒体(rAAV)，并且

[1164] 其中所述工程化AAV VP衣壳多肽不具有SEQ ID NO:1、SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7和SEQ ID NO:8中任一者的序列。4.如实施方案3所述的工程化AAV VP衣壳多肽，其中与具有SEQ ID NO:1的野生型AAV5 VP衣壳多肽的rAAV病毒体相比，所述rAAV具有较高的对骨骼肌组织的嗜性。5.如实施方案1中任一项所述的工程化AAV VP衣壳多肽，其中从对应于残基581的残基至对应于残基589的残基的区域，包括端点在内，具有与SEQ ID NO:28991-SEQ ID NO:31990中任一者至少70％、至少75％、至少80％、至少85％、至少90％、至少92％、至少95％、至少97％或至少99％同一的序列。6.如实施方案5所述的工程化AAV VP衣壳多肽，其中从对应于残基581的残基至对应于残基589的残基的区域，包括端点在内，具有SEQ ID NO:28991-SEQ ID NO:31990中任一者的序列。7.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽，其中SEQ ID NO:2的Xaa1至Xaa9区域具有与SEQ ID NO:28991-SEQ ID NO:31990中任一者至少70％、至少75％、至少80％、至少85％、至少90％、至少92％、至少95％、至少97％或至少99％同一的序列。

[1165] 8.如实施方案7所述的工程化AAV VP衣壳多肽，其中SEQ ID NO:2的Xaa1至Xaa9区域具有SEQ ID NO:28991-SEQ ID NO:31990中任一者的序列。

[1166] 9.如实施方案X-Y中任一项所述的工程化AAV VP衣壳多肽，其中：

[1167] Xaa1选自由A、E、H、M、P、Q和S组成的组；或者

[1168] Xaa2选自由F、H、I、T和V组成的组；或者

[1169] Xaa3选自由A、G、I、K、M、Q、R、S、T和V组成的组；或者

[1170] Xaa4选自由D、E、G、P和S组成的组；或者

[1171] Xaa5选自由H、L、M、P和V组成的组；或者

[1172] Xaa6选自由E、H、N和P组成的组；或者

[1173] Xaa7选自由A、H、N、Q和T组成的组；或者

[1174] Xaa8选自由I、K、M、P和W组成的组；或者

[1175] Xaa9选自由A、I、M、P和V组成的组；或者

[1176] 它们的任何组合。

[1177] 10.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa1选自P和Q。

[1178] 11.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa1为Q。

[1179] 12.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa2选自T和V。

[1180] 13.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa2为V。

[1181] 14.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa3选自A、L、P、R和T。

[1182] 15.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa3选自L、P和T。

[1183] 16.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa3为P。

[1184] 17.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa4选自D、E和S。

[1185] 18.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa4为E。

[1186] 19.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa5选自L、M和V。

[1187] 20.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa5为L。

[1188] 21.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa6为P。

[1189] 22.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa7为H。

[1190] 23.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa8选自I、P和W。

[1191] 24.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa8为P。

[1192] 25.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa9选自A、M和P。

[1193] 26.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa9为M。

[1194] 27.如实施方案1-26中任一项所述的工程化AAV VP衣壳多肽,其中从对应于残基581的残基至对应于残基589的残基的区域,包括端点在内,具有与SEQ ID NO:28991-SEQ ID NO:29990中任一者至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%同一的序列。

[1195] 28.如实施方案27所述的工程化AAV VP衣壳多肽,其中SEQ ID NO:1的从残基581至残基589的区域具有SEQ ID NO:28991-SEQ ID NO:29990中任一者的序列。

[1196] 29.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa1具有高平均柔性。

[1197] 30.如实施方案29所述的工程化AAV VP衣壳多肽,其中Xaa1选自G和R。

[1198] 31.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa1具有低平均柔性。

[1199] 32.如实施方案31所述的工程化AAV VP衣壳多肽,其中Xaa1选自W、M、F和H。

[1200] 33.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa1具有高摩尔质量。

[1201] 34.如实施方案33所述的工程化AAV VP衣壳多肽,其中Xaa1选自R、F和W。

[1202] 35.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa2具有低亲疏水性。

[1203] 36.如实施方案35所述的工程化AAV VP衣壳多肽,其中Xaa2选自K和R。

[1204] 37.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa2具有低可突变性。

[1205] 38.如实施方案37所述的工程化AAV VP衣壳多肽,其中Xaa2选自C、R和H。

[1206] 39.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa2具有高平均柔性。

[1207] 40.如实施方案39所述的工程化AAV VP衣壳多肽,其中Xaa2选自G和R。

[1208] 41.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa3具有高平均柔性。

[1209] 42.如实施方案41所述的工程化AAV VP衣壳多肽,其中Xaa3选自G和R。

[1210] 43.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa4具有高亲水

性。

[1211] 44.如实施方案43所述的工程化AAV VP衣壳多肽，其中Xaa4选自D、E、R、K和N。

[1212] 45.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽，其中Xaa4具有低可突变性。

[1213] 46.如实施方案45所述的工程化AAV VP衣壳多肽，其中Xaa4选自C、R和H。

[1214] 47.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽，其中Xaa5具有低摩尔质量。

[1215] 48.如实施方案47所述的工程化AAV VP衣壳多肽，其中Xaa5为A。

[1216] 49.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽，其中Xaa5具有低平均柔性。

[1217] 50.如实施方案49所述的工程化AAV VP衣壳多肽，其中Xaa5选自A和L。

[1218] 51.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽，其中Xaa5具有高可突变性。

[1219] 52.如实施方案51所述的工程化AAV VP衣壳多肽，其中Xaa5选自D、A和E。

[1220] 53.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽，其中Xaa6具有低平均柔性。

[1221] 54.如实施方案53所述的工程化AAV VP衣壳多肽，其中Xaa6选自W、M和F。

[1222] 55.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽，其中Xaa6具有低可突变性。

[1223] 56.如实施方案55所述的工程化AAV VP衣壳多肽，其中Xaa6为C。

[1224] 57.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽，其中Xaa6具有高摩尔质量。

[1225] 58.如实施方案57所述的工程化AAV VP衣壳多肽，其中Xaa6为W。

[1226] 59.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽，其中Xaa7具有低goldman engelman steitz。

[1227] 60.如实施方案59所述的工程化AAV VP衣壳多肽，其中Xaa7为R。

[1228] 61.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽，其中Xaa7具有高平均柔性。

[1229] 62.如实施方案61所述的工程化AAV VP衣壳多肽，其中Xaa7选自D、R、P、G和S。

[1230] 63.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽，其中Xaa7具有高可突变性。

[1231] 64.如实施方案63所述的工程化AAV VP衣壳多肽，其中Xaa7选自R、H和N。

[1232] 65.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽，其中Xaa7具有低溶解性。

[1233] 66.如实施方案65所述的工程化AAV VP衣壳多肽，其中Xaa7选自R和Q。

[1234] 67.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽，其中Xaa8具有高亲水性。

[1235] 68.如实施方案67所述的工程化AAV VP衣壳多肽，其中Xaa8选自D、E、R、K和N。

[1236] 69.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽，其中Xaa9具有低可突

变性。

[1237] 70.如实施方案69所述的工程化AAV VP衣壳多肽,其中Xaa9选自Y、F和L。

[1238] 71.如实施方案1-70中任一项所述的工程化AAV VP衣壳多肽,其中从对应于残基581的残基至对应于残基589的残基的区域包括端点在内具有与SEQ ID NO:29991-SEQ ID NO:30990中任一者具有至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%同一的序列。

[1239] 72.如实施方案71所述的工程化AAV VP衣壳多肽,其中从对应于残基581的残基至对应于残基589的残基的区域包括端点在内具有SEQ ID NO:29991-SEQ ID NO:30990中任一者的序列。

[1240] 73.如实施方案1-72中任一项所述的工程化AAV VP衣壳多肽,其中对骨骼肌组织的嗜性被测量为所述rAAV病毒体在骨骼肌组织中相较于非骨骼肌组织中的相对积聚,其中所述非骨骼肌组织由CNS、肝脏、心脏、肺、脾脏、淋巴结、骨髓、乳腺、皮肤、肾上腺、甲状腺、结肠、坐骨神经和脊髓共同组成。

[1241] 74.如实施方案73所述的工程化AAV VP衣壳多肽,其中所述CNS组织选自前脑皮层、枕叶皮层、颞叶皮层、丘脑、下丘脑、黑质、海马DG、海马CA1、海马CA3、小脑以及它们的任何组合。

[1242] 75.如实施方案73和74中任一项所述的工程化AAV VP衣壳多肽,其中所述较高的组织嗜性是在骨骼肌组织中相较于非骨骼肌组织中1.0005倍至约1000倍增加的积聚。

[1243] 76.如实施方案75所述的工程化AAV VP衣壳多肽,其中所述较高的组织嗜性是在骨骼肌组织中相较于非骨骼肌组织中至少约1.0005倍、至少约两倍、至少约三倍、至少约四倍、至少约五倍、至少约十倍、至少约二十倍、至少约50倍、至少约75倍、至少约100倍或至少约1000倍增加的积聚。

[1244] [416]0系列实施方案-皮肤嗜性衣壳

[1245] 1.一种工程化腺相关病毒(AAV)病毒蛋白(VP)衣壳多肽,所述工程化腺相关病毒(AAV)病毒蛋白(VP)衣壳多肽具有与SEQ ID NO:1至少70%同一的氨基酸序列,

[1246] 其中所述工程化AAV VP衣壳多肽相较于SEQ ID NO:1在从对应于SEQ ID NO:1的残基581的残基至对应于SEQ ID NO:1的残基589的残基的区域,包括端点在内,中具有至少一个突变,

[1247] 其中所述工程化AAV VP衣壳多肽能够组装成重组AAV病毒体(rAAV),

[1248] 其中与具有SEQ ID NO:1的野生型AAV5 VP衣壳多肽的rAAV病毒体相比,所述至少一个突变赋予所述rAAV较高的对皮肤组织的嗜性,并且

[1249] 其中所述工程化AAV VP衣壳多肽不具有SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7和SEQ ID NO:8中任一者的序列。2.如实施方案1所述的工程化AAV VP衣壳多肽,其中所述工程化AAV VP衣壳多肽具有SEQ ID NO:2的序列,并且其中Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9各自独立地选自任何氨基酸。3.一种工程化腺相关病毒(AAV)病毒蛋白(VP)衣壳多肽,所述工程化腺相关病毒(AAV)病毒蛋白(VP)衣壳多肽具有SEQ ID NO:2的氨基酸序列,

[1250] 其中氨基酸残基Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9各自独立地选自A、R、N、D、C、E、Q、G、H、I、L、K、M、F、P、S、T、W、Y和V,

[1251] 其中所述工程化AAV VP衣壳多肽能够组装成重组AAV病毒体(rAAV),并且

[1252] 其中所述工程化AAV VP衣壳多肽不具有SEQ ID NO:1、SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7和SEQ ID NO:8中任一者的序列。4.如实施方案3所述的工程化AAV VP衣壳多肽,其中与具有SEQ ID NO:1的野生型AAV5 VP衣壳多肽的rAAV病毒体相比,所述rAAV具有较高的对皮肤组织的嗜性。5.如实施方案1所述的工程化AAV VP衣壳多肽,其中从对应于残基581的残基至对应于残基589的残基的区域,包括端点在内,具有与SEQ ID NO:31991-SEQ ID NO:34990中任一者至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%同一的序列。

[1253] 6.如实施方案5所述的工程化AAV VP衣壳多肽,其中从对应于残基581的残基至对应于残基589的残基的区域,包括端点在内,具有SEQ ID NO:31991-SEQ ID NO:34990中任一者的序列。

[1254] 7.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中SEQ ID NO:2的Xaa1至Xaa9区域具有与SEQ ID NO:31 991-SEQ ID NO:34990中任一者至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%同一的序列。

[1255] 8.如实施方案7所述的工程化AAV VP衣壳多肽,其中SEQ ID NO:2的Xaa1至Xaa9区域具有SEQ ID NO:31991-SEQ ID NO:34990中任一者的序列。

[1256] 9.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中:

[1257] Xaa1选自由A、C、K、Q、R和T组成的组;或者

[1258] Xaa2选自由A、C、I、S、T和V组成的组;或者

[1259] Xaa3选自由A、C、F、G、M、Q、S和V组成的组;或者

[1260] Xaa4选自由C、K、L、P、R和W组成的组;或者

[1261] Xaa5选自由F、H、I、M、V和Y组成的组;或者

[1262] Xaa6选自由F、H、I、M、N、Q和S组成的组;或者

[1263] Xaa7选自由A、H、K、M、N、R和V组成的组;或者

[1264] Xaa8选自由A、F、G、H、S和Y组成的组;或者

[1265] Xaa9选自由A、E、G、P、Q、R和S组成的组;或者

[1266] 它们的任何组合。

[1267] 10.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa1选自C、K和R。

[1268] 11.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa1为C。

[1269] 12.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa2选自A、S、T和V。

[1270] 13.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa2为V。

[1271] 14.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa3选自A、C、F、M和Q。

[1272] 15.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa3为C。

[1273] 16.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa4选自L、P和R。

[1274] 17.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa4为R。

[1275] 18.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa5选自M、V和Y。

[1276] 19.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa5为Y。

[1277] 20.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa6选自M、N和Q。

[1278] 21.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa6为N。

[1279] 22.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa7为A、H、K或R。

[1280] 23.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa7为K。

[1281] 24.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa8选自A、F和S。

[1282] 25.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa8为S。

[1283] 26.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa9选自A、Q和S。

[1284] 27.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa9为A。

[1285] 28.如实施方案1-27中任一项所述的工程化AAV VP衣壳多肽,其中从对应于残基581的残基至对应于残基589的残基的区域,包括端点在内,具有与SEQ ID NO:31991-SEQ ID NO:32990中任一者至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%同一的序列。

[1286] 29.如实施方案28所述的工程化AAV VP衣壳多肽,其中SEQ ID NO:1的从残基581至残基589的区域具有SEQ ID NO:31991-SEQ ID NO:32990中任一者的序列。

[1287] 30.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa1具有低表面可接近性。

[1288] 31.如实施方案30所述的工程化AAV VP衣壳多肽,其中Xaa1为C。

[1289] 32.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa1具有低体积。

[1290] 33.如实施方案32所述的工程化AAV VP衣壳多肽,其中Xaa1为C。

[1291] 34.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa1具有低可突变性。

[1292] 35.如实施方案34所述的工程化AAV VP衣壳多肽,其中Xaa1为C。

[1293] 36.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa2具有高表面可接近性。

[1294] 37.如实施方案36所述的工程化AAV VP衣壳多肽,其中Xaa2选自R和K。

[1295] 38.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa2具有高平均柔性。

[1296] 39.如实施方案38所述的工程化AAV VP衣壳多肽,其中Xaa2选自K、I和N。

[1297] 40.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa2具有低可突变性。

[1298] 41.如实施方案40所述的工程化AAV VP衣壳多肽,其中Xaa2选自P和K。

[1299] 42.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa3具有高亲疏水性。

[1300] 43.如实施方案42所述的工程化AAV VP衣壳多肽,其中Xaa3选自I和V。

[1301] 44.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa4具有低可突变性。

[1302] 45.如实施方案44所述的工程化AAV VP衣壳多肽,其中Xaa4选自L、F和Y。

[1303] 46.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa4具有低平均柔性。

[1304] 47.如实施方案46所述的工程化AAV VP衣壳多肽，其中Xaa4选自W、H、F和M。

[1305] 48.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽，其中Xaa5具有高平均柔性。

[1306] 49.如实施方案48所述的工程化AAV VP衣壳多肽，其中Xaa5选自G、R、K、I和N。

[1307] 50.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽，其中Xaa6具有高平均柔性。

[1308] 51.如实施方案50所述的工程化AAV VP衣壳多肽，其中Xaa6选自G、R、K、I和N。

[1309] 52.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽，其中Xaa8具有高表面可接近性。

[1310] 53.如实施方案52所述的工程化AAV VP衣壳多肽，其中Xaa8选自M、G和F。

[1311] 54.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽，其中Xaa8具有低平均柔性。

[1312] 55.如实施方案54所述的工程化AAV VP衣壳多肽，其中Xaa8选自H、F、M和W。

[1313] 56.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽，其中Xaa8具有低可突变性。

[1314] 57.如实施方案56所述的工程化AAV VP衣壳多肽，其中Xaa8选自L、F和Y。

[1315] 58.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽，其中Xaa9具有高平均柔性。

[1316] 59.如实施方案58所述的工程化AAV VP衣壳多肽，其中Xaa9选自D、E、R、K、P和G。

[1317] 60.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽，其中Xaa9具有高可突变性。

[1318] 61.如实施方案60所述的工程化AAV VP衣壳多肽，其中Xaa9选自D、E、R、V、A和H。

[1319] 62.如实施方案1-61中任一项所述的工程化AAV VP衣壳多肽，其中从对应于残基581的残基至对应于残基589的残基的区域包括端点在内具有与SEQ ID NO:32991-SEQ ID NO:33990中任一者具有至少70％、至少75％、至少80％、至少85％、至少90％、至少92％、至少95％、至少97％或至少99％同一的序列。

[1320] 63.如实施方案62所述的工程化AAV VP衣壳多肽，其中从对应于残基581的残基至对应于残基589的残基的区域包括端点在内具有SEQ ID NO:32991-SEQ ID NO:33990中任一者的序列。

[1321] 64.如实施方案1-61中任一项所述的工程化AAV VP衣壳多肽，其中对皮肤组织的嗜性被测量为所述rAAV病毒体在皮肤组织中相较于非皮肤组织中的相对积聚，其中所述非皮肤组织由CNS、肝脏、骨骼肌、心脏、肺、脾脏、淋巴结、骨髓、乳腺、肾上腺、甲状腺、结肠、坐骨神经和脊髓共同组成。

[1322] 65.如实施方案64所述的工程化AAV VP衣壳多肽，其中所述CNS组织选自前脑皮层、枕叶皮层、颞叶皮层、丘脑、下丘脑、黑质、海马DG、海马CA1、海马CA3、小脑以及它们的任何组合。

[1323] 66.如实施方案64-65中任一项所述的工程化AAV VP衣壳多肽，其中所述较高的组织嗜性是在皮肤组织中相较于非皮肤组织中1.0005倍至约1000倍增加的积聚。

[1324] 67.如实施方案66所述的工程化AAV VP衣壳多肽，其中所述较高的组织嗜性是在

皮肤组织中相较于非皮肤组织中至少约1.0005倍、至少约两倍、至少约三倍、至少约四倍、至少约五倍、至少约十倍、至少约二十倍、至少约50倍、至少约75倍、至少约100倍或至少约1000倍增加的积聚。

[1325] [417]P系列实施方案-脊髓嗜性衣壳

[1326] 1.一种工程化腺相关病毒(AAV)病毒蛋白(VP)衣壳多肽,所述工程化腺相关病毒(AAV)病毒蛋白(VP)衣壳多肽具有与SEQ ID NO:1至少70%同一的氨基酸序列,

[1327] 其中所述工程化AAV VP衣壳多肽相较于SEQ ID NO:1在从对应于SEQ ID NO:1的残基581的残基至对应于SEQ ID NO:1的残基589的残基的区域,包括端点在内,中具有至少一个突变,

[1328] 其中所述工程化AAV VP衣壳多肽能够组装成重组AAV病毒体(rAAV),

[1329] 其中与具有SEQ ID NO:1的野生型AAV5 VP衣壳多肽的rAAV病毒体相比,所述至少一个突变赋予所述rAAV较高的对脊髓组织的嗜性,并且

[1330] 其中所述工程化AAV VP衣壳多肽不具有SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7和SEQ ID NO:8中任一者的序列。2.如实施方案1所述的工程化AAV VP衣壳多肽,其中所述工程化AAV VP衣壳多肽具有SEQ ID NO:2的序列,并且其中Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9各自独立地选自任何氨基酸。3.一种工程化腺相关病毒(AAV)病毒蛋白(VP)衣壳多肽,所述工程化腺相关病毒(AAV)病毒蛋白(VP)衣壳多肽具有SEQ ID NO:2的氨基酸序列,

[1331] 其中氨基酸残基Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9各自独立地选自A、R、N、D、C、E、Q、G、H、I、L、K、M、F、P、S、T、W、Y和V,

[1332] 其中所述工程化AAV VP衣壳多肽能够组装成重组AAV病毒体(rAAV),并且

[1333] 其中所述工程化AAV VP衣壳多肽不具有SEQ ID NO:1、SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7和SEQ ID NO:8中任一者的序列。4.如实施方案3所述的工程化AAV VP衣壳多肽,其中与具有SEQ ID NO:1的野生型AAV5 VP衣壳多肽的rAAV病毒体相比,所述rAAV具有较高的对脊髓组织的嗜性。5.如实施方案1所述的工程化AAV VP衣壳多肽,其中从对应于残基581的残基至对应于残基589的残基的区域,包括端点在内,具有与SEQ ID NO:34991-SEQ ID NO:37437中任一者至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%同一的序列。6.如实施方案5所述的工程化AAV VP衣壳多肽,其中从对应于残基581的残基至对应于残基589的残基的区域,包括端点在内,具有SEQ ID NO:34991-SEQ ID NO:37437中任一者的序列。7.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中SEQ ID NO:2的Xaa1至Xaa9区具有与SEQ ID NO:34991-SEQ ID NO:37437中任一者至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%同一的序列。

[1334] 8.如实施方案7所述的工程化AAV VP衣壳多肽,其中SEQ ID NO:2的Xaa1至Xaa9区域具有SEQ ID NO:34991-SEQ ID NO:37437中任一者的序列。

[1335] 9.如实施方案X-Y中任一项所述的工程化AAV VP衣壳多肽,其中:

[1336] Xaa1选自由A、C、K、Q、R、S和W组成的组;或者

[1337] Xaa2选自由H、I、K、L、T、V和W组成的组;或者

[1338] Xaa3选自由C、F、G、H、I、K、N和R组成的组;或者

[1339] Xaa4选自由I、M、Q、S和V组成的组;或者

[1340] Xaa5选自由H、K、Q、T、W和Y组成的组;或者

[1341] Xaa6选自由H、L、N、Q、R、W和Y组成的组;或者

[1342] Xaa7选自由D、H、P、Q和R组成的组;或者

[1343] Xaa8选自由D、F、L、S、T和Y组成的组;或者

[1344] Xaa9选自由C、I、N、P、R、S和Y组成的组;或者

[1345] 它们的任何组合。

[1346] 10.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa1选自K、R和W。

[1347] 11.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa1为K。

[1348] 12.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa2选自H、I和T。

[1349] 13.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa2为I。

[1350] 14.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa3选自F、I和R。

[1351] 15.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa3为I。

[1352] 16.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa4选自I、M和V。

[1353] 17.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa4为V。

[1354] 18.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa5选自T、W和Y。

[1355] 19.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa5为Y。

[1356] 20.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa6选自L、N、R和Y。

[1357] 21.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa6为Y。

[1358] 22.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa7为R。

[1359] 23.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa8选自S、T和Y。

[1360] 24.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa8为T。

[1361] 25.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa9选自I、P和R。

[1362] 26.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa9为I。

[1363] 27.如实施方案1-26中任一项所述的工程化AAV VP衣壳多肽,其中从对应于残基581的残基至对应于残基589的残基的区域,包括端点在内,具有与SEQ ID NO:34991-SEQ ID NO:35437中任一者至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%同一的序列。

[1364] 28.如实施方案27所述的工程化AAV VP衣壳多肽,其中SEQ ID NO:1的从残基581至残基589的区域具有SEQ ID NO:34991-SEQ ID NO:35437中任一者的序列。

[1365] 29.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa1具有高体积。

[1366] 30.如实施方案29所述的工程化AAV VP衣壳多肽,其中Xaa1选自F、W和Y。

[1367] 31.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa1具有低可突变性。

[1368] 32.如实施方案31所述的工程化AAV VP衣壳多肽,其中Xaa1选自Y、F、L和C。

[1369] 33.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa1具有高溶解性。

[1370] 34.如实施方案33所述的工程化AAV VP衣壳多肽,其中Xaa1选自W、F、I和L。

[1371] 35.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa1具有低平均柔性。

[1372] 36.如实施方案35所述的工程化AAV VP衣壳多肽,其中Xaa1选自F、M和W。

[1373] 37.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa2具有低亲疏水性。

[1374] 38.如实施方案37所述的工程化AAV VP衣壳多肽,其中Xaa2选自P和Y。

[1375] 39.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa3具有低亲水性。

[1376] 40.如实施方案39所述的工程化AAV VP衣壳多肽,其中Xaa3选自Y、W、V、M、F、I和L。

[1377] 41.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa3具有高溶解性。

[1378] 42.如实施方案41所述的工程化AAV VP衣壳多肽,其中Xaa3选自W、F、I和L。

[1379] 43.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa6具有高体积。

[1380] 44.如实施方案43所述的工程化AAV VP衣壳多肽,其中Xaa6选自W、R、K、M、I和L。

[1381] 45.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa6具有高摩尔质量。

[1382] 46.如实施方案45所述的工程化AAV VP衣壳多肽,其中Xaa6选自W。

[1383] 47.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa8具有高摩尔质量。

[1384] 48.如实施方案47所述的工程化AAV VP衣壳多肽,其中Xaa8选自W、E、K、M、H和Q。

[1385] 49.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa8具有高体积。

[1386] 50.如实施方案49所述的工程化AAV VP衣壳多肽,其中Xaa8选自W、K、M、I和L。

[1387] 51.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa8具有高goldman engelman steitz。

[1388] 52.如实施方案51所述的工程化AAV VP衣壳多肽,其中Xaa8选自V和L。

[1389] 53.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa9具有高亲疏水性。

[1390] 54.如实施方案53所述的工程化AAV VP衣壳多肽,其中Xaa9选自V和I。

[1391] 55.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa9具有高溶解性。

[1392] 56.如实施方案55所述的工程化AAV VP衣壳多肽,其中Xaa9选自W、F、I和L。

[1393] 57.如实施方案1-56中任一项所述的工程化AAV VP衣壳多肽,其中从对应于残基581的残基至对应于残基589的残基的区域包括端点在内具有与SEQ ID NO:35438-SEQ ID NO:36437中任一者具有至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%同一的序列。

[1394] 58.如实施方案57所述的工程化AAV VP衣壳多肽,其中从对应于残基581的残基至对应于残基589的残基的区域包括端点在内具有SEQ ID NO:35438-SEQ ID NO:36437中任

一者的序列。

[1395] 59.如实施方案1-58中任一项所述的工程化AAV VP衣壳多肽,其中对脊髓组织的嗜性被测量为所述rAAV病毒体在脊髓组织中相较于非脊髓组织中的相对积聚,其中所述非脊髓组织由CNS、肝脏、骨骼肌、心脏、肺、脾脏、淋巴结、骨髓、乳腺、皮肤、肾上腺、甲状腺、结肠和坐骨神经共同组成。

[1396] 60.如实施方案59所述的工程化AAV VP衣壳多肽,其中所述CNS组织选自前脑皮层、枕叶皮层、颞叶皮层、丘脑、下丘脑、黑质、海马DG、海马CA1、海马CA3、小脑以及它们的任何组合。

[1397] 61.如实施方案59-60中任一项所述的工程化AAV VP衣壳多肽,其中所述较高的组织嗜性是在脊髓组织中相较于非脊髓组织中1.0005倍至约1000倍增加的积聚。

[1398] 62.如实施方案61所述的工程化AAV VP衣壳多肽,其中所述较高的组织嗜性是在脊髓组织中相较于非脊髓组织中至少约1.0005倍、至少约两倍、至少约三倍、至少约四倍、至少约五倍、至少约十倍、至少约二十倍、至少约50倍、至少约75倍、至少约100倍或至少约1000倍增加的积聚。

[1399] [418]Q系列实施方案-脾脏嗜性衣壳

[1400] 1.一种工程化腺相关病毒(AAV)病毒蛋白(VP)衣壳多肽,所述工程化腺相关病毒(AAV)病毒蛋白(VP)衣壳多肽具有与SEQ ID NO:1至少70%同一的氨基酸序列,

[1401] 其中所述工程化AAV VP衣壳多肽相较于SEQ ID NO:1在从对应于SEQ ID NO:1的残基581的残基至对应于SEQ ID NO:1的残基589的残基的区域,包括端点在内,中具有至少一个突变,

[1402] 其中所述工程化AAV VP衣壳多肽能够组装成重组AAV病毒体(rAAV),

[1403] 其中与具有SEQ ID NO:1的野生型AAV5 VP衣壳多肽的rAAV病毒体相比,所述至少一个突变赋予所述rAAV较高的对脾脏组织的嗜性,并且

[1404] 其中所述工程化AAV VP衣壳多肽不具有SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7和SEQ ID NO:8中任一者的序列。2.如实施方案1所述的工程化AAV VP衣壳多肽,其中所述工程化AAV VP衣壳多肽具有SEQ ID NO:2的序列,并且其中Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9各自独立地选自任何氨基酸。3.一种工程化腺相关病毒(AAV)病毒蛋白(VP)衣壳多肽,所述工程化腺相关病毒(AAV)病毒蛋白(VP)衣壳多肽具有SEQ ID NO:2的氨基酸序列,

[1405] 其中氨基酸残基Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9各自独立地选自A、R、N、D、C、E、Q、G、H、I、L、K、M、F、P、S、T、W、Y和V,

[1406] 其中所述工程化AAV VP衣壳多肽能够组装成重组AAV病毒体(rAAV),并且

[1407] 其中所述工程化AAV VP衣壳多肽不具有SEQ ID NO:1、SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7和SEQ ID NO:8中任一者的序列。4.如实施方案3所述的工程化AAV VP衣壳多肽,其中与具有SEQ ID NO:1的野生型AAV5 VP衣壳多肽的rAAV病毒体相比,所述rAAV具有较高的对脾脏组织的嗜性。5.如实施方案1所述的工程化AAV VP衣壳多肽,其中从对应于残基581的残基至对应于残基589的残基的区域,包括端点在内,具有与SEQ ID NO:37438-SEQ ID NO:40437中任一者至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%同一的序列。6.如实施方案

5所述的工程化AAV VP衣壳多肽，其中从对应于残基581的残基至对应于残基589的残基的区域，包括端点在内，具有SEQ ID NO:37438-SEQ ID NO:40437中任-者的序列。

[1408] 7.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽，其中SEQ ID NO:2的Xaa1至Xaa9区域具有与SEQ ID NO:37438-SEQ ID NO:40437中任一者至少70％、至少75％、至少80％、至少85％、至少90％、至少92％、至少95％、至少97％或至少99％同一的序列。

[1409] 8.如实施方案7所述的工程化AAV VP衣壳多肽，其中SEQ ID NO:2的Xaa1至Xaa9区域具有SEQ ID NO:37438-SEQ ID NO:40437中任一者的序列。

[1410] 9.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽，其中：

[1411] Xaa1选自由C、F、H、I、L、P、W和Y组成的组；或者

[1412] Xaa2选自由D、E、L、N、P、R和W组成的组；或者

[1413] Xaa3选自由C、D、E、P和W组成的组；或者

[1414] Xaa4选自由C、F、G、H、R、W和Y组成的组；或者

[1415] Xaa5选自由以下组成的组：A、D、E、G、P、R和W；或者

[1416] Xaa6选自由A、C、D、E、K、R和W组成的组；或者

[1417] Xaa7选自由F、L、P、R、W和Y组成的组；或者

[1418] Xaa8选自由E、I、K、L、P、R和T组成的组；或者

[1419] Xaa9选自由C、H、M、T、V和W组成的组；或者

[1420] 它们的任何组合。

[1421] 10.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa1选自C、F、P、W和Y。

[1422] 11.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa1选自P、W和Y。

[1423] 12.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa1为P。

[1424] 13.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa2选自D、E和W。

[1425] 14.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa2为D。

[1426] 15.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa3选自D、P和W。

[1427] 16.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa3为P。

[1428] 17.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa4选自C、H和W。

[1429] 18.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa4为C。

[1430] 19.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa5选自D、E、G和P。

[1431] 20.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa5为D。

[1432] 21.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa6选自C、K和R。

[1433] 22.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa6为K。

[1434] 23.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa7选自L、P和W。

[1435] 24.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa7为P。

[1436] 25.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa8选自P、R和K。

[1437] 26.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa8为K。

[1438] 27.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa9选自C、T和V。

[1439] 28.如实施方案9所述的工程化AAV VP衣壳多肽，其中Xaa9为V。

[1440] 29.如实施方案1-28中任一项所述的工程化AAV VP衣壳多肽，其中从对应于残基581的残基至对应于残基589的残基的区域，包括端点在内，具有与SEQ ID NO:37438-SEQ

ID NO:38437中任一者至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%同一的序列。

[1441] 30.如实施方案29所述的工程化AAV VP衣壳多肽,其中SEQ ID NO:1的从残基581至残基589的区域具有SEQ ID NO:37438-SEQ ID NO:38437中任一者的序列。

[1442] 31.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa1具有低溶解性。

[1443] 32.如实施方案31所述的工程化AAV VP衣壳多肽,其中Xaa1选自D和P。

[1444] 33.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa1具有高溶解性。

[1445] 34.如实施方案33所述的工程化AAV VP衣壳多肽,其中Xaa1选自F、I和L。

[1446] 35.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa1具有低亲疏水性。

[1447] 36.如实施方案35所述的工程化AAV VP衣壳多肽,其中Xaa1选自Y和P。

[1448] 37.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa1具有低可突变性。

[1449] 38.如实施方案37所述的工程化AAV VP衣壳多肽,其中Xaa1选自C、K和P。

[1450] 39.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa2具有低溶解性。

[1451] 40.如实施方案39所述的工程化AAV VP衣壳多肽,其中Xaa2选自D、Q和R。

[1452] 41.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa2具有低亲疏水性。

[1453] 42.如实施方案41所述的工程化AAV VP衣壳多肽,其中Xaa2选自D、E、R、K、H、N和Q。

[1454] 43.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa2具有低电荷。

[1455] 44.如实施方案43所述的工程化AAV VP衣壳多肽,其中Xaa2选自D和E。

[1456] 45.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa2具有低体积。

[1457] 46.如实施方案45所述的工程化AAV VP衣壳多肽,其中Xaa2选自T、N、P和D。

[1458] 47.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa2具有高平均柔性。

[1459] 48.如实施方案47所述的工程化AAV VP衣壳多肽,其中Xaa2选自D、E、R、P、G、Q和S。

[1460] 49.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa3具有低溶解性。

[1461] 50.如实施方案49所述的工程化AAV VP衣壳多肽,其中Xaa3选自D、E、P和N。

[1462] 51.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa3具有低亲疏水性。

[1463] 52.如实施方案51所述的工程化AAV VP衣壳多肽,其中Xaa3选自D、E、H、N、Q和P。

[1464] 53.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa4具有低亲疏水性。

[1465] 54.如实施方案53所述的工程化AAV VP衣壳多肽,其中Xaa4选自K和R。

[1466] 55.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa5具有低溶解性。

[1467] 56.如实施方案55所述的工程化AAV VP衣壳多肽,其中Xaa5选自D、E、P和N。

[1468] 57.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa5具有高平均柔性。

[1469] 58.如实施方案57所述的工程化AAV VP衣壳多肽,其中Xaa5选自D、E、R、P、G、Q和S。

[1470] 59.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa6具有低可突变性。

[1471] 60.如实施方案59所述的工程化AAV VP衣壳多肽,其中Xaa6选自C。

[1472] 61.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa8具有高表面可接近性。

[1473] 62.如实施方案61所述的工程化AAV VP衣壳多肽,其中Xaa8选自E、R和K。

[1474] 63.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa8具有低溶解性。

[1475] 64.如实施方案63所述的工程化AAV VP衣壳多肽,其中Xaa8选自E、P、R、K、N、Q。

[1476] 65.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa8具有中等体积。

[1477] 66.如实施方案65所述的工程化AAV VP衣壳多肽,其中Xaa8选自E、D、R、K、V、P、M、I、L、H、N、Q和T。

[1478] 67.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa9具有中等摩尔质量。

[1479] 68.如实施方案67所述的工程化AAV VP衣壳多肽,其中Xaa9选自E、D、K、M、I、L、H、N。

[1480] 69.如实施方案1-68中任一项所述的工程化AAV VP衣壳多肽,其中从对应于残基581的残基至对应于残基589的残基的区域包括端点在内具有与SEQ ID NO:38438-SEQ ID NO:39437中任一者具有至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%同一的序列。

[1481] 70.如实施方案69所述的工程化AAV VP衣壳多肽,其中从对应于残基581的残基至对应于残基589的残基的区域包括端点在内具有SEQ ID NO:38438-SEQ ID NO:39437中任一者的序列。

[1482] 71.如实施方案1-70中任一项所述的工程化AAV VP衣壳多肽,其中对脾脏组织的嗜性被测量为所述rAAV病毒体在脾脏组织中相较于非脾脏组织中的相对积聚,其中所述非脾脏组织由CNS、肝脏、骨骼肌、心脏、肺、淋巴结、骨髓、乳腺、皮肤、肾上腺、甲状腺、结肠、坐骨神经和脊髓共同组成。

[1483] 72.如实施方案71所述的工程化AAV VP衣壳多肽,其中所述CNS组织选自前脑皮层、枕叶皮层、颞叶皮层、丘脑、下丘脑、黑质、海马DG、海马CA1、海马CA3、小脑以及它们的任何组合。

[1484] 73.如实施方案71-72中任一项所述的工程化AAV VP衣壳多肽,其中所述较高的组

织嗜性是在脾脏组织中相较于非脾脏组织中1.0005倍至约1000倍增加的积聚。

[1485] 74.如实施方案73所述的工程化AAV VP衣壳多肽,其中所述较高的组织嗜性是在脾脏组织中相较于非脾脏组织中至少约1.0005倍、至少约两倍、至少约三倍、至少约四倍、至少约五倍、至少约十倍、至少约二十倍、至少约50倍、至少约75倍、至少约100倍或至少约1000倍增加的积聚。

[1486] [419]R系列实施方案-甲状腺嗜性衣壳

[1487] 1.一种工程化腺相关病毒(AAV)病毒蛋白(VP)衣壳多肽,所述工程化腺相关病毒(AAV)病毒蛋白(VP)衣壳多肽具有与SEQ ID NO:1至少70%同一的氨基酸序列,

[1488] 其中所述工程化AAV VP衣壳多肽相较于SEQ ID NO:1在从对应于SEQ ID NO:1的残基581的残基至对应于SEQ ID NO:1的残基589的残基的区域,包括端点在内,中具有至少一个突变,

[1489] 其中所述工程化AAV VP衣壳多肽能够组装成重组AAV病毒体(rAAV),

[1490] 其中与具有SEQ ID NO:1的野生型AAV5 VP衣壳多肽的rAAV病毒体相比,所述至少一个突变赋予所述rAAV较高的对甲状腺组织的嗜性,并且

[1491] 其中所述工程化AAV VP衣壳多肽不具有SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7和SEQ ID NO:8中任一者的序列。2.如实施方案1所述的工程化AAV VP衣壳多肽,其中所述工程化AAV VP衣壳多肽具有SEQ ID NO:2的序列,并且其中Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9各自独立地选自任何氨基酸。3.一种工程化腺相关病毒(AAV)病毒蛋白(VP)衣壳多肽,所述工程化腺相关病毒(AAV)病毒蛋白(VP)衣壳多肽具有SEQ ID NO:2的氨基酸序列,

[1492] 其中氨基酸残基Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9各自独立地选自A、R、N、D、C、E、Q、G、H、I、L、K、M、F、P、S、T、W、Y和V,

[1493] 其中所述工程化AAV VP衣壳多肽能够组装成重组AAV病毒体(rAAV),并且

[1494] 其中所述工程化AAV VP衣壳多肽不具有SEQ ID NO:1、SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7和SEQ ID NO:8中任一者的序列。4.如实施方案3所述的工程化AAV VP衣壳多肽,其中与具有SEQ ID NO:1的野生型AAV5 VP衣壳多肽的rAAV病毒体相比,所述rAAV具有较高的对甲状腺组织的嗜性。5.如实施方案1所述的工程化AAV VP衣壳多肽,其中从对应于残基581的残基至对应于残基589的残基的区域,包括端点在内,具有与SEQ ID NO:40438-SEQ ID NO:43437中任一者至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%同一的序列。6.如实施方案5所述的工程化AAV VP衣壳多肽,其中从对应于残基581的残基至对应于残基589的残基的区域,包括端点在内,具有SEQ ID NO:40438-SEQ ID NO:43437中任一者的序列。7.如实施方案2-4所述的工程化AAV VP衣壳多肽,其中SEQ ID NO:2的Xaa1至Xaa9区域具有与SEQ ID NO:40438-SEQ ID NO:43437中任一者至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%同一的序列。8.如实施方案7所述的工程化AAV VP衣壳多肽,其中SEQ ID NO:2的Xaa1至Xaa9区域具有SEQ ID NO:40438-SEQ ID NO:43437中任一者的序列。

[1495] 9.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中:

[1496] Xaa1选自由A、K、M、N、Q或R组成的组;或者

[1497] Xaa2选自由A、F、K、L、M、T、V或W组成的组;或者

[1498] Xaa3选自由A、I、K、R、S、T、V或W组成的组;或者

[1499] Xaa4选自由A、D、E、I、P或V组成的组;或者

[1500] Xaa5选自由F、I、M、Q、V或Y组成的组;或者

[1501] Xaa6选自由H、M、N或Y组成的组;或者

[1502] Xaa7选自由H、I、N、Q、S或W组成的组;或者

[1503] Xaa8选自由A、D、F、Q、S或Y组成的组;或者

[1504] Xaa9选自由A、Q、S或Y组成的组;或者

[1505] 它们的任何组合。

[1506] 10.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa1选自K、N和Q。

[1507] 11.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa1为K。

[1508] 12.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa2选自F、V和W。

[1509] 13.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa2为W。

[1510] 14.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa3选自A、R和T。

[1511] 15.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa3为R。

[1512] 16.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa4选自A、E和I。

[1513] 17.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa4为A。

[1514] 18.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa5为M、V或Y。

[1515] 19.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa5为M。

[1516] 20.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa6为N。

[1517] 21.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa7选自H、I和N。

[1518] 22.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa7为H。

[1519] 23.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa8选自A、F和S。

[1520] 24.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa8为F。

[1521] 25.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa9选自A和S。

[1522] 26.如实施方案9所述的工程化AAV VP衣壳多肽,其中Xaa9为A。

[1523] 27.如实施方案1-26中任一项所述的工程化AAV VP衣壳多肽,其中从对应于残基581的残基至对应于残基589的残基的区域,包括端点在内,具有与SEQ ID NO:40438-SEQ ID NO:41437中任一者至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%同一的序列。

[1524] 28.如实施方案27所述的工程化AAV VP衣壳多肽,其中SEQ ID NO:1的从残基581至残基589的区域具有SEQ ID NO:40438-SEQ ID NO:41437中任一者的序列。

[1525] 29.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa1具有高可突变性。

[1526] 30.如实施方案29所述的工程化AAV VP衣壳多肽,其中Xaa1为N。

[1527] 31.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa2具有低表面可接近性。

[1528] 32.如实施方案31所述的工程化AAV VP衣壳多肽,其中Xaa2选自F、G和M。

[1529] 33.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa3具有高溶解

性。

[1530] 34.如实施方案33所述的工程化AAV VP衣壳多肽,其中Xaa3选自F。

[1531] 35.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa3具有低可突变性。

[1532] 36.如实施方案35所述的工程化AAV VP衣壳多肽,其中Xaa3选自Y、F、L和C。

[1533] 37.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa3具有中等摩尔质量。

[1534] 38.如实施方案37所述的工程化AAV VP衣壳多肽,其中Xaa3选自D、E、R、K、V、P、M、I、L、N、Q、T和C。

[1535] 39.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa3具有低表面可接近性。

[1536] 40.如实施方案39所述的工程化AAV VP衣壳多肽,其中Xaa3选自V、I、L和C。

[1537] 41.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa4具有高goldman engelman steitz。

[1538] 42.如实施方案41所述的工程化AAV VP衣壳多肽,其中Xaa4选自L和V。

[1539] 43.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa4具有低表面可接近性。

[1540] 44.如实施方案43所述的工程化AAV VP衣壳多肽,其中Xaa4选自V、M、A、G、F、I和L。

[1541] 45.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa4具有低摩尔质量。

[1542] 46.如实施方案45所述的工程化AAV VP衣壳多肽,其中Xaa4选自D、A、G、I、L和N。

[1543] 47.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa5具有高溶解性。

[1544] 48.如实施方案47所述的工程化AAV VP衣壳多肽,其中Xaa5选自C、L、F、M、V和Y。

[1545] 49.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa5具有低溶解性。

[1546] 50.如实施方案49所述的工程化AAV VP衣壳多肽,其中Xaa5选自D。

[1547] 51.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa5具有低平均柔性。

[1548] 52.如实施方案51所述的工程化AAV VP衣壳多肽,其中Xaa5选自F、M和W。

[1549] 53.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa6具有低平均柔性。

[1550] 54.如实施方案53所述的工程化AAV VP衣壳多肽,其中Xaa6选自F、M和W。

[1551] 55.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa7具有高可突变性。

[1552] 56.如实施方案55所述的工程化AAV VP衣壳多肽,其中Xaa7选自N。

[1553] 57.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa7具有低体积。

[1554] 58.如实施方案57所述的工程化AAV VP衣壳多肽,其中Xaa7选自P、N和T。

[1555] 59.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa8具有低平均柔性。

[1556] 60.如实施方案59所述的工程化AAV VP衣壳多肽,其中Xaa8选自F、M和W。

[1557] 61.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa8具有低表面可接近性。

[1558] 62.如实施方案61所述的工程化AAV VP衣壳多肽,其中Xaa8选自M、G和F。

[1559] 63.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa9具有低可突变性。

[1560] 64.如实施方案63所述的工程化AAV VP衣壳多肽,其中Xaa9选自R、K、P、H和C。

[1561] 65.如实施方案2-4中任一项所述的工程化AAV VP衣壳多肽,其中Xaa9具有低亲疏水性。

[1562] 66.如实施方案65所述的工程化AAV VP衣壳多肽,其中Xaa9选自R。

[1563] 67.如实施方案1-66中任一项所述的工程化AAV VP衣壳多肽,其中从对应于残基581的残基至对应于残基589的残基的区域包括端点在内具有与SEQ ID NO:41438-SEQ ID NO:42437中任一者具有至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少95%、至少97%或至少99%同一的序列。

[1564] 68.如实施方案67所述的工程化AAV VP衣壳多肽,其中从对应于残基581的残基至对应于残基589的残基的区域包括端点在内具有SEQ ID NO:41438-SEQ ID NO:42437中任一者的序列。

[1565] 69.如实施方案1-68中任一项所述的工程化AAV VP衣壳多肽,其中对甲状腺组织的嗜性被测量为所述rAAV病毒体在甲状腺组织中相较于非甲状腺组织中的相对积聚,其中所述非甲状腺组织由CNS、肝脏、骨骼肌、心脏、肺、脾脏、淋巴结、骨髓、乳腺、皮肤、肾上腺、结肠、坐骨神经和脊髓共同组成。

[1566] 70.如实施方案69所述的工程化AAV VP衣壳多肽,其中所述CNS组织选自前脑皮层、枕叶皮层、颞叶皮层、丘脑、下丘脑、黑质、海马DG、海马CA1、海马CA3、小脑以及它们的任何组合。

[1567] 71.如实施方案69-70中任一项所述的工程化AAV VP衣壳多肽,其中所述较高的组织嗜性是在甲状腺组织中相较于非甲状腺组织中1.0005倍至约1000倍增加的积聚。

[1568] 72.如实施方案71所述的工程化AAV VP衣壳多肽,其中所述较高的组织嗜性是在甲状腺组织中相较于非甲状腺组织中至少约1.0005倍、至少约两倍、至少约三倍、至少约四倍、至少约五倍、至少约十倍、至少约二十倍、至少约50倍、至少约75倍、至少约100倍或至少约1000倍增加的积聚。

[1569] S系列实施方案

[1570] 1.一种工程化腺相关病毒(AAV)VP衣壳多肽,所述工程化腺相关病毒(AAV)VP衣壳多肽具有SEQ ID NO:2的氨基酸序列,

[1571] 其中氨基酸残基Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9各自独立地选自A、R、N、D、C、E、Q、G、H、I、L、K、M、F、P、S、T、W、Y和V;

[1572] 其中所述衣壳多肽能够组装成重组AAV病毒体(rAAV);并且

[1573] 其中多肽不具有SEQ ID NO:1、SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID

NO:6、SEQ ID NO:7和SEQ ID NO:8中任一者的序列。

[1574] 2.一种工程化腺相关病毒(AAV)VP衣壳多肽,所述多肽在对应于SEQ ID NO:1中的残基581至残基589的残基中具有至少一个突变,其中所述衣壳多肽能够组装成重组AAV病毒体(rAAV),并且其中所述多肽突变赋予所述重组rAAV对第一组织相较于第二组织的组织嗜性,并且其中所述AAV VP衣壳多肽不具有SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7和SEQ ID NO:8中任一者的序列。

[1575] 3.一种工程化AAV VP衣壳多肽,所述工程化AAV VP衣壳多肽包含由下式表示的多肽序列:(A)-(X)-(B)

[1576] 其中:

[1577] (A)是SEQ ID NO:47438(VAYNVGGQMATNNQSSTTAP SEQ ID NO:2的残基561-580)的多肽序列;

[1578] (X)是包含SEQ ID NO:2的氨基酸残基Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9的多肽序列;并且

[1579] (B)是SEQ ID NO:47439(SEQ ID NO:2的残基590-609IVPGSVWMERDVYLQGPIWA)的多肽序列;

[1580] 其中Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9各自独立地选自A、R、N、D、C、E、Q、G、H、I、L、K、M、F、P、S、T、W、Y和V;并且

[1581] 其中所述衣壳多肽能够组装成重组AAV病毒体(rAAV);并且;

[1582] 其中多肽不具有SEQ ID NO:1、SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7和SEQ ID NO:8中任一者的序列。

[1583] 4.如实施方案3所述的工程化AAV VP衣壳多肽,所述工程化AAV VP衣壳多肽包含由下式表示的多肽序列:(A)-(X)-(B),其中:

[1584] (A)是SEQ ID NO:47438(SEQ ID NO:2的残基561-580VAYNVGGQMATNNQSSTTAP)的多肽序列;

[1585] (X)是选自表8(SEQ ID NO:115-1114)或表10(SEQ ID NO:7118-8117)中的多肽的列表的多肽序列,所述多肽序列赋予重组AAV病毒体(rAAV)CNS组织嗜性;并且1

[1586] (B)是SEQ ID NO:47439(SEQ ID NO:2的残基590-609(IVPGSVWMERDVYLQGPIWA))的多肽序列;并且

[1587] 其中衣壳多肽能够组装成rAAV,并且,

[1588] 衣壳不具有SEQ ID NO:1、SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7和SEQ ID NO:8中任一者的序列。

5.如实施方案4所述的工程化AAV VP衣壳多肽,其中CNS组织选自由以下组成的组:海马:(齿状回,CA1和CA3);小脑、下丘脑、皮层:(枕叶、颞叶和前脑);黑质、丘脑,以及它们的任何组合。

[1589] 6.如实施方案3所述的重组AAV VP衣壳多肽,所述工程化AAV VP衣壳多肽包含由下式表示的多肽序列:(A)-(X)-(B),其中:

[1590] (A)是SEQ ID NO:47438(SEQ ID NO:2的残基:(561-580VAYNVGGQMATNNQSSTTAP))的多肽序列;

[1591] (X)是选自SEQ ID NO:115-1114或1118-47437的多肽中任一者的多肽序列,所述

多肽序列赋予重组AAV病毒体(rAAV)相应的组织嗜性;并且

[1592] (B)是SEQ ID NO:47439(SEQ ID NO:2的残基590-609(IVPGSVWMERDVYLQGPIWA))的多肽序列;并且

[1593] 其中衣壳多肽能够组装成rAAV,并且,

[1594] 所述衣壳不具有SEQ ID NO:1、SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7和SEQ ID NO:8中任一者的序列。7.如前述实施方案中任一项所述的工程化AAV VP衣壳多肽,其中所述AAV VP衣壳多肽为AAV5 VP1衣壳多肽。8.如前述实施方案中任一项所述的工程化AAV VP衣壳多肽,其中所述衣壳赋予重组AAV病毒体(rAAV)增加的对选自以下组织的任何组织的组织靶向性:

[1595] 脂肪、肾上腺、主动脉、脑(包括海马:齿状回、CA1和CA3;小脑、尾状核、壳核、中脑、脑桥、下丘脑、皮层-包括枕叶、颞叶和前脑;黑质、丘脑,以及它们的任何组合)、骨髓、盲肠、结肠、背根神经节、十二指肠、附睾、食管、眼睛、胆囊、心脏、回肠、空肠、肾脏、肺、淋巴结、乳腺、卵巢、胰腺、甲状旁腺、周围神经、垂体、前列腺、唾液腺、精囊、骨骼肌、皮肤、脊髓、脾脏、胃、睾丸、胸腺、甲状腺、气管、膀胱、子宫和阴道的组织,

[1596] 相较于包含具有以下序列中任一者的衣壳的对照病毒体而言:SEQ ID NO:1、SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7和SEQ ID NO:8。9.如实施方案2所述的工程化AAV VP衣壳多肽,其中所述rAAV表现出在第一组织中相较于第二组织中约1.0005倍至约1000倍增加的积聚。10.如实施方案9所述的工程化AAV VP衣壳多肽,其中所述rAAV表现出在第一组织中相较于第二组织中至少约1.0005倍、至少约两倍、至少约三倍、至少约四倍、至少约五倍、至少约十倍、至少约二十倍、至少约50倍、至少约75倍、至少约100倍或至少约1000倍增加的积聚。11.如实施方案1或3-8中任一项所述的工程化AAV VP衣壳多肽,其中所述rAAV表现出在组织中约1.0005倍至约1000倍增加的积聚。12.如实施方案11所述的工程化AAV VP衣壳多肽,其中所述rAAV表现出在组织中至少约1.0005倍、至少约两倍、至少约三倍、至少约四倍、至少约五倍、至少约十倍、至少约二十倍、至少约50倍、至少约75倍、至少约100倍或至少约1000倍增加的积聚。13.如实施方案1-3或6-12中任一项所述的工程化AAV VP衣壳多肽,其中至少一个突变残基是:

[1597] Xaa1和Xaa1选自A、G、K、M、N、Q、R、S或T;

[1598] Xaa2和Xaa2选自A、C、H、I、K、S、T或V;

[1599] Xaa3和Xaa3选自A、G、H、K、M、N、Q、R、S、T或V;

[1600] Xaa4和Xaa4选自L、M、P、Q、R、T或W;

[1601] Xaa5和Xaa5选自F、H、I、K、M、T或Y;

[1602] Xaa6和Xaa6选自E、G、H、L、M、N、Q、T或W;

[1603] Xaa7和Xaa7选自A、C、G、H、L、M、R或S;

[1604] Xaa8和Xaa8选自A、C、D、F、G、H、M、Q、S、V、W或Y;

[1605] Xaa9和Xaa9选自A、C、E、G、H、M、N、P、Q、S、V或W;

[1606] 或它们的任何组合,

[1607] 其中所述衣壳多肽能够组装成重组AAV病毒体(rAAV),并且其中所述rAAV能够表现出对肝脏组织的组织嗜性。

[1608] 14.如实施方案1-12中任一项所述的工程化AAV VP衣壳多肽,其中:

[1609] Xaa1不包括A、G、K、M、N、Q、R、S或T；

[1610] Xaa2不包括A、C、H、I、K、S、T或V；

[1611] Xaa3不包括A、G、H、K、M、N、Q、R、S、T或V；

[1612] Xaa4不包括L、M、P、Q、R、T或W；

[1613] Xaa5不包括F、H、I、K、M、T或Y；

[1614] Xaa6不包括E、G、H、L、M、N、Q、T或W；

[1615] Xaa7不包括A、C、G、H、L、M、R或S。

[1616] Xaa8不包括A、C、D、F、G、H、M、Q、S、V、W或Y；

[1617] Xaa9不包括A、C、E、G、H、M、N、P、Q、S、V或W。

[1618] 或它们的任何组合，

[1619] 其中与具有含SEQ ID NO:1序列的VP1衣壳多肽的rAAV相比，所述衣壳多肽能够组装成重组AAV病毒体(rAAV)并且表现出更少的对肝脏组织的靶向性或嗜性。15.如前述实施方案中任一项所述的工程化AAV VP衣壳多肽，所述多肽具有至少一个对应于SEQ ID NO:2中残基581至残基589的残基，其中：

[1620] Xaa1为A、D、E、G、L、M、N、Q、S、T或V；

[1621] Xaa2为A、C、D、E、G、H、I、N、P、Q、S、T或V；

[1622] Xaa3为A、D、E、G、H、M、N、Q、S、T或V；

[1623] Xaa4为A、D、E、G、H、N、P、Q、S或T；

[1624] Xaa5为A、C、D、E、G、H、N、Q、S、T或Y；

[1625] Xaa6为A、D、E、G、H、N、P、Q、S或T；

[1626] Xaa7为A、C、D、E、G、H、N、Q、S或T；

[1627] Xaa8为A、C、D、E、G、H、N、Q、S或T；

[1628] Xaa9为A、D、E、G、H、N、P、Q、S或T；

[1629] 或它们的任何组合，

[1630] 其中所述衣壳多肽能够组装成重组AAV病毒体(rAAV)。

[1631] 16.如前述实施方案中任一项所述的重组衣壳多肽，所述重组衣壳多肽还在581-589区域之外的氨基酸残基处包含一个或多个突变或取代，其中在581-589区域之外的氨基酸残基处的一个或多个突变或取代赋予改进的可制造性、改进的病毒组装、提高的组织靶向性/嗜性，或它们的任何组合。17.一种编码实施方案1-16所述的任何重组腺相关病毒(AAV)VP衣壳多肽的多核苷酸。18.一种重组AAV病毒体(rAAV)，所述病毒体包含实施方案1-15中任一项所述的AAV VP衣壳多肽。19.如实施方案18所述的rAAV病毒体，其中与具有含SEQ ID NO:1序列的VP1衣壳多肽的rAAV相比，所述rAAV具有降低的对人肝脏的嗜性。20.如实施方案18或实施方案19所述的rAAV病毒体，其中与具有含SEQ ID NO:1序列的VP1衣壳多肽的rAAV相比，所述rAAV在静脉内施用后具有增强的穿过血脑屏障的能力。21.如实施方案18-20中任一项所述的rAAV病毒体，其中与具有含SEQ ID NO:1序列的VP1衣壳多肽的rAAV相比，所述rAAV在静脉内、鞘内、脑室内或脑池内施用后具有增强的感染选自海马、齿状回、大脑皮层、颞叶皮层、枕叶皮层、丘脑、前脑、黑质、下丘脑和小脑的一个或多个脑区的能力。

[1632] 22.如实施方案18-21中任一项所述的rAAV病毒体，其中与具有含SEQ ID NO:1序列的VP1衣壳多肽的rAAV相比，所述rAAV在玻璃体内注射后具有增强的感染人视网膜细胞

的能力。

[1633] 23.如实施方案18中任一项所述的rAAV病毒体,其中与具有含SEQ ID NO:1序列的VP1衣壳多肽的rAAV相比,所述rAAV在静脉内施用后具有增强的感染人骨骼肌的能力。

[1634] 24.如实施方案18中任一项所述的rAAV病毒体,其中与具有含SEQ ID NO:1序列的VP1衣壳多肽的rAAV相比,所述rAAV具有增加的对人肝脏的嗜性。25.如实施方案18-24中任一项所述的rAAV病毒体,其中所述病毒体还包含载体基因组,所述载体基因组包含编码以下任一项的治疗性多核苷酸:选自指导RNA或tRNA的治疗性RNA,或编码处于引导受感染人细胞中转基因表达的调控序列控制下的蛋白质。26.如实施方案25所述的rAAV病毒体,其中所述转基因编码选自表1的转基因产物的蛋白质。27.一种组合物,所述组合物包含含有实施方案1-16中任一项所述的工程化衣壳多肽的AAV病毒体,所述AAV病毒体内包装有实施方案25所述的治疗性多核苷酸。28.一种组合物,所述组合物包含含有实施方案1-16中任一项所述的工程化衣壳多肽的AAV病毒体,所述AAV病毒体内包装有编码以下任一项的治疗性多核苷酸:选自指导RNA或tRNA的治疗性RNA,或编码处于引导受感染人细胞中转基因表达的调控序列控制下的蛋白质。29.如实施方案18-26中任一项所述的rAAV病毒体,其中所述病毒体还包含载体基因组,所述载体基因组包含编码以下任一项的治疗性多核苷酸:选自指导RNA或tRNA的治疗性RNA,或编码处于引导受感染人细胞中转基因表达的调控序列控制下的蛋白质。30.如实施方案29所述的rAAV病毒体,其中所述转基因编码选自表1的转基因产物的蛋白质。

[1635] 31.一种药物组合物,所述药物组合物包含实施方案29或实施方案30所述的rAAV和药学上可接受的载体。32.一种治疗方法,所述方法包括:

[1636] 向有需要的患者施用有效量的实施方案31所述的药物组合物。33.如实施方案32所述的方法,其中所述rAAV的所述有效量小于野生型rAAV的有效量。34.如实施方案32所述的方法,其中所述rAAV的所述有效量小于在对应于SEQ ID NO:1的残基581至残基589的位置处缺乏一个或多于一个突变或取代的其他方面相当的rAAV的有效量。35.如实施方案34所述的方法,其中与所述野生型rAAV、所述其他方面相当的rAAV或两者的所述有效量相比,所述rAAV的所述有效量在所述患者中引起更低的毒性。

[1637] 36.如实施方案32所述的方法,其中所述有效量为至少$1x10^5$个病毒基因组/kg患者体重至$5x10^{14}$个病毒基因组/kg。

[1638] 37.如实施方案32-36中任一项所述的方法,其中静脉内施用所述rAAV。

[1639] 38.如实施方案32-36中任一项所述的方法,其中鞘内施用所述rAAV。

[1640] 39.如实施方案32-36中任一项所述的方法,其中通过脑池内施用来施用所述rAAV。

[1641] 40.如实施方案32-36中任一项所述的方法,其中通过玻璃体内注射来施用所述rAAV。

[1642] 41.如实施方案18-26或29-30中任一项所述的rAAV病毒体,所述rAAV病毒体用于治疗患者的疾病。

[1643] 42.根据实施方案41使用的rAAV病毒体,所述rAAV病毒体用于治疗表1中的任何疾病。

[1644] 43.根据实施方案41使用的rAAV病毒体,所述rAAV病毒体用于治疗CNS疾病。

[1645] 44.根据实施方案43使用的rAAV病毒体,其中所述CNS疾病选自表1中列出的疾患。

[1646] 45.根据实施方案43或44使用的rAAV病毒体,其中所述工程化AAV VP衣壳多肽包含表8(SEQ ID NO:115-1114)、表10(SEQ ID NO:7118-8117)、表39(SEQ ID NO:8118-9117)或表74(SEQ ID NO:9118-10117)中列举的突变中的任一者。

[1647] 46.根据实施方案45使用的rAAV病毒体,其中所述工程化VP1衣壳多肽不具有SEQ ID NO:1、SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7和SEQ ID NO:8中任一者的序列。

[1648] 47.根据实施方案42使用的rAAV病毒体,其中所述病毒体包含编码来自表1中列出的任何靶标的靶标的治疗性多核苷酸。

[1649] T系列实施方案-AAV5

[1650] 1.如本文描述的实施方案中任一项所述的工程化AAV VP衣壳多肽,其中所述工程化AAV VP衣壳多肽是工程化AAV5 VP衣壳多肽。

[1651] 2.如实施方案1所述的工程化AAV VP衣壳多肽,其中所述突变是取代。

[1652] U系列实施方案-组装

[1653] 1.如本文描述的实施方案中任一项所述的工程化AAV VP衣壳多肽,其中:

[1654] Xaa1选自由A、D、E、G、L、M、N、Q、S、T或V组成的组;或者

[1655] Xaa2选自由A、C、D、E、G、H、I、N、P、Q、S、T或V组成的组;或者

[1656] Xaa3选自由A、D、E、G、H、M、N、Q、S、T或V组成的组;或者

[1657] Xaa4选自由A、D、E、G、H、N、P、Q、S或T组成的组;或者

[1658] Xaa5选自由A、C、D、E、G、H、N、Q、S、T或Y组成的组;或者

[1659] Xaa6选自由A、D、E、G、H、N、P、Q、S或T组成的组;或者

[1660] Xaa7选自由A、C、D、E、G、H、N、Q、S或T组成的组;或者

[1661] Xaa8选自由A、C、D、E、G、H、N、Q、S或T组成的组;或者

[1662] Xaa9选自由A、D、E、G、H、N、P、Q、S或T组成的组;或者

[1663] 它们的任何组合。

[1664] 2.如实施方案1所述的工程化AAV VP衣壳多肽,其中Xaa1为A、D、E、M或T。

[1665] 3.如实施方案1所述的工程化AAV VP衣壳多肽,其中Xaa1为E。

[1666] 4.如实施方案1所述的工程化AAV VP衣壳多肽,其中Xaa2为A、S、T或V。

[1667] 5.如实施方案1所述的工程化AAV VP衣壳多肽,其中Xaa2为A。

[1668] 6.如实施方案1所述的工程化AAV VP衣壳多肽,其中Xaa3为D、E、N、Q或T。

[1669] 7.如实施方案1所述的工程化AAV VP衣壳多肽,其中Xaa3为D或T。

[1670] 8.如实施方案1所述的工程化AAV VP衣壳多肽,其中Xaa4为D、E、P或Q。

[1671] 9.如实施方案1所述的工程化AAV VP衣壳多肽,其中Xaa4为E。

[1672] 10.如权利要求X-Y中任一项所述的工程化AAV VP衣壳多肽,其中Xaa5为D、E、N、Q或T。

[1673] 11.如实施方案1所述的工程化AAV VP衣壳多肽,其中Xaa5为N。

[1674] 12.如实施方案1所述的工程化AAV VP衣壳多肽,其中Xaa6为D、N或Q。

[1675] 13.如实施方案1所述的工程化AAV VP衣壳多肽,其中Xaa6为D。

[1676] 14.如权利要求X-Y中任一项所述的工程化AAV VP衣壳多肽,其中Xaa7为A、D、E或

G。

[1677] 15.如实施方案1所述的工程化AAV VP衣壳多肽,其中Xaa7为A。

[1678] 16.如权利要求X-Y中任一项所述的工程化AAV VP衣壳多肽,其中Xaa8为A、D、G或S。

[1679] 17.如实施方案1所述的工程化AAV VP衣壳多肽,其中Xaa8为G。

[1680] 18.如实施方案1所述的工程化AAV VP衣壳多肽,其中Xaa9为A、D、G或P。

[1681] 19.如实施方案1所述的工程化AAV VP衣壳多肽,其中Xaa9为G。

[1682] V系列实施方案-重组AAV病毒体

[1683] 1.一种重组AAV病毒体(rAAV),所述重组AAV病毒体(rAAV)包含:

[1684] 本文描述的实施方案所述的组装成衣壳的工程化AAV VP衣壳多肽;和

[1685] 有效载荷,

[1686] 其中所述工程化AAV VP衣壳多肽衣壳包裹所述有效载荷。

[1687] 2.如实施方案1所述的重组rAAV,其中所述有效载荷包含治疗性多核苷酸,任选地其中所述治疗性多核苷酸编码转基因或基因组修饰实体。

[1688] 3.如实施方案2所述的rAAV,其中所述治疗性多核苷酸编码指导RNA、tRNA、阻抑tRNA、siRNA、miRNA、mRNA、shRNA、环状RNA或反义寡核苷酸(ASO)、核酶、DNA酶、适体或它们的任何组合。

[1689] 4.如实施方案3所述的rAAV,其中所述治疗性多核苷酸编码线性治疗性多核苷酸或环状治疗性多核苷酸。

[1690] 5.如实施方案2所述的rAAV,其中所述转基因选自表1。

[1691] 6.如实施方案2所述的rAAV,其中所述基因组修饰实体包含CRISPR/Cas系统、作用于RNA的腺苷脱氨酶(ADAR)酶、转录激活物或转录阻抑物。

[1692] 7.如实施方案6所述的rAAV,其中所述CRISPR/Cas系统包括Cas3、Cas8、Cas10、Cas9、Cas4、Cas12或Cas13。

[1693] 8.如实施方案6所述的rAAV,其中所述ADAR是人ADAR1或人ADAR2。

[1694] 9.如实施方案6所述的rAAV,其中所述转录激活物是VP64。

[1695] 10.如实施方案6所述的rAAV,其中所述转录阻抑物是KRAB。

[1696] 11.一种药物组合物,所述药物组合物包含实施方案1-10中任一项所述的rAAV病毒体和药学上可接受的载体。

[1697] 12.一种治疗方法,所述方法包括向有需要的受试者施用治疗有效量的实施方案11所述的药物组合物。

[1698] 13.如实施方案12所述的药物组合物,所述药物组合物用于治疗有需要的受试者,其中所述用途包括向受试者施用治疗有效量的如实施方案11所述的药物组合物。

[1699] 14.如实施方案1-10中任一项所述的rAAV病毒体,所述rAAV病毒体用于制造用于治疗有需要的受试者的药物,所述治疗通过施用治疗有效量的所述药物来实现。

[1700] 15.如实施方案12-14中任一项所述的方法或用途,其中所述rAAV病毒体的所述治疗有效量小于衣壳包裹相同有效载荷的具有SEQ ID NO:1的野生型AAV5 VP衣壳多肽的rAAV病毒体的治疗有效量,其中具有所述SEQ ID NO:1的野生型AAV5 VP衣壳多肽的所述rAAV病毒体通过相同的施用途径施用。

[1701] 16.如实施方案12-15中任一项所述的方法或用途,其中与衣壳包裹相同有效载荷的具有SEQ ID NO:1的野生型AAV5 VP衣壳多肽的rAAV病毒体的治疗有效量相比,所述rAAV病毒体的所述治疗有效量在所述受试者中引起较低的毒性,其中具有所述SEQ ID NO:1的野生型AAV5 VP衣壳多肽的所述rAAV病毒体通过相同的施用途径施用。

[1702] 17.如实施方案12-16中任一项所述的方法或用途,其中所述治疗有效量为至少$1x10^{5}$个病毒基因组/kg受试者体重至$5x10^{14}$个病毒基因组/kg受试者体重。

[1703] 18.如实施方案12-17中任一项所述的方法或用途,其中通过全身施用来施用所述rAAV病毒体。

[1704] 19.如实施方案18所述的方法或用途,其中所述全身施用是静脉内施用、肌肉内施用、腹膜内施用或口服施用。

[1705] 20.如实施方案12-19中任一项所述的方法或用途,其中所述受试者是人类或非人类动物。

[1706] 21.如实施方案12-20中任一项所述的方法或用途,其中所述受试者患有疾病。

[1707] 22.如实施方案21所述的方法或用途,其中所述疾病是神经系统疾患。

[1708] 23.如实施方案22所述的方法,其中所述神经系统疾患是AADC缺乏症、阿尔茨海默氏病、tau蛋白病、突触核蛋白病、贝登氏病、MPS-IIIB、伴有GBA1突变的额颞叶痴呆(PD-GBA)、神经性戈谢氏病、戈谢病2型、卡纳万病、帕金森氏病、泰萨克斯病、亨廷顿氏病、波托基-卢普斯基综合征、肌萎缩侧索硬化症、唐氏综合征、圣菲利波病A型、圣菲利波病B型或雷特综合征。

[1709] 24.如实施方案23所述的方法,其中所述痴呆是额颞叶痴呆(FTD)。

[1710] W系列实施方案

[1711] 1.如本文公开的实施方案中任一项所述的方法或用途,其中所述疾病选自表1。

[1712] 2.如本文公开的实施方案中任一项所述的方法或用途,其中所述rAAV VP1衣壳的位置587处(Xaa7)不是A、C、D、E、F、G、H、I、K、M、P、Q、R、V、W或Y。

[1713] 3.如本文公开的实施方案中任一项所述的方法或用途,其中所述rAAV VP1衣壳的位置587处可以是N、S或T。

[1714] 4.如本文公开的实施方案中任一项所述的方法或用途,其中所述rAAV VP1衣壳的位置582处(Xaa2)不是G、V、L或I。

[1715] 5.如本文公开的实施方案中任一项所述的方法或用途,其中所述rAAV VP1衣壳包含581-589(Xaa1-Xaa9)区域之外的一个或多个突变,其中所述rAAV VP1衣壳能够组装并表现出所需的组织嗜性。

[1716] 1.2.实施例

[1717] 以下是用于实施本发明的具体实施方案的实例。这些实施例仅为了说明目的提供而并非旨在以任何方式限制本发明的范围。已经努力确保关于所使用的数值(例如用量、温度等)的准确性,但当然也应允许一些实验误差和偏差。

[1718] 除非另外指示,否则将采用本领域技术范围内的蛋白质化学、生物化学、重组DNA技术和药理学的常规方法来实施本发明。文献中充分说明了此类技术。

[1719] 实施例1

[1720] 高通量衣壳工程化系统

[1721] 此实施例描述了本文公开的高通量衣壳工程化系统和方法以发现工程化组织嗜性AAV变体。所述高通量衣壳工程化系统在图1中示意。如图1所示，在组装前阶段，本文所述的方法允许产生含有AAV衣壳基因变体的质粒文库，所述变体在衣壳基因编码序列的特定位置具有大量多样性。使用本文所述的2质粒系统将此变体文库组装成病毒，然后将组装的病毒直接注射到非人灵长类动物(NHP)中以进行体内选择。收获组织并将组织转导衣壳基因用独特分子标识符和条形码标记。使这些变体序列/UMI/条形码参数化并使用机器学习算法鉴定特定组织靶向性/去靶向性衣壳的确定性特征。“去靶向”在本文中可以称为相对于特定组织而言的“保留”。例如，肝脏去靶向或肝脏保留在本文中都可用于指优先表现出对肝脏组织的归巢减少的变体或变体特性。

[1722] 本文所述的方法允许克隆具有大量多样性的AAV衣壳基因的合成寡核苷酸变体。使用本文所述的2质粒系统将此变体文库组装成病毒，然后将组装的病毒直接注射到非人灵长类动物中以进行体内选择。该过程如下进行，但应当理解，对这些方法的合理且科学标准的修改涵盖在本文中。对文库模板进行初次PCR扩增(8-10个循环)，并使用柱纯化DNA(池文库)。任选地，从文库池中进行第二次8循环扩增。使用合适的柱纯化DNA。接下来，使用NEB BsaI HF-V2进行消化(文库-3小时；主链-2小时+1小时小牛肠碱性磷酸酶)并使用柱纯化DNA。将所得消化物以3:1比率合并，在16℃下用NEB T4连接酶连接，并在第二天早上添加更多的ATP(1μM)和T4连接酶。“连接酶循环”进行20次。连接酶循环是指每个“循环”在16℃下10秒，随后在25℃下30秒。将连接酶热灭活，并添加更多的ATP(1μM)。在37℃下添加PS-DNA酶持续2小时以消化线性(未连接)物种，富集松弛形式(rf)环状DNA(图3)。使用柱纯化DNA，并在水中洗脱DNA。通过执行许多(>100)电穿孔反应来大规模转化电感受态重组酶缺陷细菌以整合文库质粒。通过在37℃下液体振荡培养过夜，使用抗生素选择来选择转化的细菌。通过在抗生素选择性琼脂平板上铺板有限稀释的转化细菌来确认转化效率并测量文库覆盖率。使用Megaprep试剂盒制备文库质粒DNA用于测序和转染从而进行病毒生产。通过“双重转染”生产病毒体：将病毒生产者真核细胞系(例如，HEK293)与腺病毒源性“辅助”基因质粒和文库质粒的混合物共转染。在接下来的3-7天使用密度离心或FPLC方法纯化来自细胞/悬浮液的病毒。将来自病毒体的DNA扩增用于组装体的NGS测序分析。任选地，可以用具有组装能力的变体重复克隆。将纯化的文库病毒以1-5x10^{13}个病毒基因组/kg的剂量范围静脉内施用于NHP。4-6周后收获组织并对文库变体进行测序。

[1723] 收获组织并将转导衣壳基因用独特分子标识符和条形码标记。使这些变体序列/UMI/条形码参数化并使用机器学习算法鉴定特定组织靶向性/去靶向性衣壳的确定性特征。

[1724] 图2A提供了已显示与靶细胞相互作用的已知AAV衣壳的关键残基的侧视图(上图)和俯视图(下图)。图2B示出了实施例1中描述的AAV衣壳文库的突出特征，显示出引入多样性的区域，其中残基编号对应于AAV5 VP1中的氨基酸编号。文库质粒包括AAV ITR之间的不变的rep和多样化的cap编码序列(顺式包装方法)。

[1725] 如上文所述，质粒安全DNA酶(PS-DNA酶)用于在连接后去除线性DNA。使用这种线性选择性DNA酶酶降解未连接的DNA可最大限度地提高转化效率，从而最大限度地提高文库多样性。图3是显示在用PS-DNA酶消化后有效去除线性插入片段和线性质粒骨架(“BB”)的凝胶照片。在该方法的一个实例中，质粒文库产生的总效率为4.23x10^{9}个菌落形成单位/μg

质粒DNA。

[1726] 图4是实施例1中描述的高通量AAV衣壳工程化系统的示意图。直方图显示了AAV衣壳文库在不同制备阶段的大小,追踪了文库从理论多样性、合成衣壳基因、总克隆变体到所测序组装病毒的大小(误差条表示从NGS测序分析中预测的最小/最大多样性)。将汇集的病毒文库注射到非人灵长类动物体内,经过一段足以确保稳定转导的时间后,将动物安乐死,并收获组织。从组织中纯化DNA(1),并且(2)附加独特分子标识符(UMI)。添加核酸外切酶I以消化过量的含UMI的引物(3)。后续PCR扩增添加测序索引/组织条形码和衔接子用于测序。添加UMI可以对组织中的衣壳变体进行高分辨率频率分析。

[1727] 图5A示出了用于使来自文库的rAAV产量最大化的包装方法,其中图5A比较了标准三重转染(上图)与实施例1中使用的双质粒顺式包装方法(下图),并且图5B显示了使用这两种转染方法的野生型AAV5的相对产量。图5C是肝脏标本中UMI分布的实例,显示一次(单个UMI)发现大多数衣壳变体。然而,衣壳变体的一个子集在给定组织中富集并且可能具有增加的UMI数,与它们在靶组织中增加的频率相对应。

[1728] 实施例2

[1729] NHP中AAV5变体的鉴定和分析

[1730] 此实施例描述了使用实施例1中描述的一般方法发现的组织靶向性AAV5衣壳变体的鉴定和分析。为了鉴定优先感染特定组织(嗜性)的变体,在非人灵长类动物(NHP)中评估基于AAV5的变体衣壳文库的生物分布。根据Primate Products,LLC标准操作程序和授权的兽医标准进行生物分布调查。在文库剂量注射之前,进行了一项研究前活动计划,以对NHP临床病理学、一般健康和AAV5血清阴性进行基准测试。基于针对AAV5的抗体的血清阴性选择研究动物。

[1731] NHP接受1-5x10^{13}个病毒基因组/kg之间的剂量范围内的基于AAV5的衣壳文库的IV推注注射。将最终稀释的基于AAV5的衣壳文库样本混合并转移到无菌注射器中进行静脉内施用,剂量体积为5mL/kg。使用基于AAV5的衣壳文库的常见移液和转移技术进行剂量准备。从基于AAV5的衣壳文库的稀释储备液的普通储备液中,根据每只动物各自的体重分别制备剂量体积,以便以每公斤为基础向每只动物递送等量的载体。

[1732] 第1天文库注射后,每天进行两次临床观察(第1-29天)并每周检查体重。在第29天,根据Primate Products,LLC标准安乐死程序对所有研究动物实施安乐死并进行组织收集。

[1733] 收集了各种各样的组织,包括:肾上腺、主动脉、骨与骨髓、脑(小脑)、脑(海马,齿状回)、脑(海马,CA1)、脑(海马,CA3)、脑(下丘脑)、脑(皮层,颞叶)、脑(皮层,前脑)、脑(皮层,枕叶)、脑(黑质)、脑(丘脑)、盲肠、结肠、十二指肠、附睾、食管、眼睛、胆囊、心脏、回肠、空肠、肾脏、肝脏、肺、淋巴结、乳腺、卵巢、胰腺、甲状旁腺、周围神经(坐骨神经)、垂体、前列腺、唾液腺、精囊、骨骼肌、皮肤、脊髓、脾脏、胃、睾丸、胸腺、甲状腺、气管、膀胱、子宫和阴道。将组织作为两个1cm^3的块(或整个器官)收集并保存在液氮中。肝脏专门以五个1cm^3块收集。对于大脑,从如上述列表中所指出的每个区域收集两个1cm^3块,但较小的区域除外,它们是完整收集的。

[1734] 图8是基于机器学习的衣壳变体聚类的示意图。该实例利用衣壳变体序列,对其使用机器学习算法来绘制感染肝脏的衣壳之间的相似性。然后,该输出可用于告知候选变体

的选择,以选择性地靶向目标组织。

[1735] 图9示出了静脉内施用于NHP后,文库整体感染皮层的相对水平高于肝脏的性能实例。从肝脏和皮层分离DNA,并使用qPCR(左)或液滴数字PCR(ddPCR)(右)来量化从各自组织回收的病毒基因组。根据这些量化方法,文库整体具有相对于肝脏约2倍增加的CNS靶向性。

[1736] 图11A至图11D显示了维恩图和分析表,描绘了针对三种不同测序分析过滤器在仅肝脏组织、肝脏和皮层组织以及仅皮层组织中发现的独特序列变体的数量。对从肝脏或皮层组织样本回收的病毒基因组衣壳变体进行新一代测序。附加独特分子标识符(UMI)作为本文所述的分子回收过程的一部分。这增加了组织中存在给定衣壳变体的置信度,并且可与丰度相关联。图11A显示了维恩图,其中应用的测序分析过滤器为每各不同衣壳序列变体4个或更多个不同UMI--在本文中也称为"计数"。图11B显示了维恩图,其中应用的测序分析过滤器为50个或更多个UMI。图11C显示了维恩图,其中应用的测序分析过滤器为100个或更多个不同UMI。图11D显示了在仅皮层中发现的衣壳序列变体群体中UMI分布的直方图分析,其中应用的测序分析过滤器为如图11B所示的50个或更多个UMI。

[1737] 实施例3

[1738] 在病毒中功能性组装的AAV5变体

[1739] 此实施例描述了使用实施例1-实施例2中描述的方法和系统发现的组装成病毒体的工程化AAV5变体。在如实施例1中描述的文库病毒的产生和纯化之后,评估AAV5文库病毒的组装。通过对跨越变体区域的扩增子的NGS测量衣壳变体的组成。分析相对于给定氨基酸残基出现在文库质粒中的指定位置的频率而言该给定氨基酸残基出现在对应于SEQ ID NO:1(推及于SEQ ID NO:2)的位置581至位置589的所述指定位置的频率,以确定优选病毒组装的序列规则。参考下文表3,和SEQ ID NO:2,以下氨基酸可以在Xaa1-Xaa9任何一个或多个位置处以任何组合独立地突变,以提供能够组装的AAV VP1衣壳。此外,可以允许Xaa1-Xaa9区域之外的一个或多个突变,只要衣壳仍然能够组装即可。

[1740]

表 3 病毒组装选项
Xaa1 为 A、D、E、G、L、M、N、Q、S、T 或 V
Xaa1 为 A、D、E、M 或 T
Xaa1 为 E
Xaa2 为 A、C、D、E、G、H、I、N、P、Q、S、T 或 V
Xaa2 为 A、S、T 或 V
Xaa2 为 A
Xaa3 为 A、D、E、G、H、M、N、Q、S、T 或 V
Xaa3 为 D、E、N、Q 或 T
Xaa3 为 D 或 T
Xaa4 为 A、D、E、G、H、N、P、Q、S 或 T
Xaa4 为 D、E、P 或 Q
Xaa4 为 E
Xaa5 为 A、C、D、E、G、H、N、Q、S、T 或 Y
Xaa5 为 D、E、N、Q 或 T
Xaa5 为 N。
Xaa6 为 A、D、E、G、H、N、P、Q、S 或 T;
Xaa6 为 D、N 或 Q; 或者
Xaa6 为 D。
Xaa7 为 A、C、D、E、G、H、N、Q、S 或 T;
Xaa7 为 A、D、E 或 G; 或者
Xaa7 为 A。
Xaa8 为 A、C、D、E、G、H、N、Q、S 或 T;
Xaa8 为 A、D、G 或 S; 或者
Xaa8 为 G。

[1741]

Xaa9为A、D、E、G、H、N、P、Q、S或T;
Xaa9为A、D、G或P; 或者
Xaa9为G。

[1742] 实施例4

[1743] 具有肝脏组织靶向性的AAV5变体和具有肝脏去靶向特性的AAV5变体

[1744] 此实施例描述了使用实施例1-实施例2中描述的方法和系统发现的在肝脏中具有组织嗜性的工程化AAV5变体和优先对肝脏组织去靶向的AAV5变体。

[1745] 图6A至图6B是热图和统计分析,显示了通过病毒组装和肝脏中的组织转导进行的明确功能选择。图6A是在实施例1-2中描述的高通量系统从组装前、组装病毒到肝脏转导的各个阶段制备的热图。因此,图6A(右)示出了除以所有组装病毒所有肝脏转导变体的有利位置残基的热图。对于组装前热图,将衣壳变体文库克隆到质粒中并转化到电感受态细菌中。对所得文库进行测序,并测量每个位置处的氨基酸多样性,低水平的位置变异可能是由综合误差引起的。随后组装成病毒显示出明显的氨基酸残基/位置偏差,突出了病毒组装的选择。肝脏转导病毒显示出更大的AA残基/位置选择模式,具有明显有利/不利的变体。对这些热图各自根据它们各自的输入频率进行了归一化。图6B是统计分析表(ANOVA),显示在三个肝脏样本中,氨基酸残基分布和残基位置显著改变,但样本间变异不显著。

[1746] 图7A至图7B示出了对多个样本之间重复观察到的(高UMI)衣壳变体的分析。图7A是示出两个肝脏样本中重叠变体鉴定计数的维恩图:在两个样本中观察到约38%的具有>10个UMI的变体。图7B是显示位置氨基酸分布的热图,示出重复观察到的衣壳变体的强选择残基/位置。

[1747] 图10提供了结构图解,左边的图解代表野生型AAV5[SEQ NO:1]的晶体结构(PDB:6JCT,来自Zhang等人Nat Commun.2019年8月21日;10(1):3760),指向3重对称轴。中间和右边的图解代表从NHP肝脏回收的变体,所述变体的581-589残基用NPAMFNCDY(中图,氨基酸SEQ ID NO:112)和KLTAHIYAL(右图,氨基酸SEQ ID NO:114)置换。这些图解使用PyMOL产生以置换变体序列,并根据预测的静电电荷进行着色,以突出从同一组织回收的肝脏变体之间的电荷差异。

[1748] 分析相对于给定氨基酸残基出现在形成组装病毒的变体中的指定位置的频率而言该给定氨基酸残基出现在肝脏中鉴定的变体中对应于SEQ ID NO:1(推及于SEQ ID NO:2)的位置581至位置589的所述指定位置的频率,以确定优先靶向肝脏组织的衣壳的一组序列规则。参考下文表4,和SEQ ID NO:2(AAV5 VP1),以下氨基酸可以在Xaa1-Xaa9任何一个或多个位置处以任何组合独立地突变,以提供与野生型AAV5 VP1衣壳(SEQ ID NO:1)相比具有增强的肝脏嗜性的VP1衣壳,其中肝脏嗜性在此是指优选肝脏转导的特性,而不是优选病毒体组装的特性。此外,可以允许Xaa1-Xaa9区域之外的一个或多个突变,只要衣壳仍然能够组装并表现出所需的肝脏靶向性特性即可。

[1749]

表 4 肝脏嗜性选项和规则
Xaa1 为 A、G、K、M、N、Q、R、S 或 T;
Xaa1 为 A、K、M 或 T; 或者
Xaa1 为 K。
Xaa2 为 A、C、H、I、K、S、T 或 V;
Xaa2 为 A、S、T 或 V; 或者
Xaa2 为 T。
Xaa3 为 A、G、H、K、M、N、Q、R、S、T 或 V;
Xaa3 为 A、M 或 T; 或者
Xaa3 为 A 或 T。
Xaa4 为 L、M、P、Q、R、T 或 W;
Xaa4 为 L、P、Q 或 T; 或者
Xaa4 为 P。
Xaa5 为 F、H、I、K、M、T 或 Y;
Xaa5 为 H、I 或 Y; 或者
Xaa5 为 Y。
Xaa6 为 E、G、H、L、M、N、Q、T 或 W;
Xaa6 为 N 或 Q; 或者
Xaa6 为 N。

[1750]

Xaa7 为 A、C、G、H、L、M、R 或 S;
Xaa7 为 A、C、H 或 M; 或者
Xaa7 为 A。
Xaa8 为 A、C、D、F、G、H、M、Q、S、V、W 或 Y;
Xaa8 为 G、M、Q 或 S; 或者
Xaa8 为 G。
Xaa9 为 A、C、E、G、H、M、N、P、Q、S、V 或 W;
Xaa9 为 E、G 或 P; 或者
Xaa9 为 G。

[1751] 分析相对于给定氨基酸残基出现在形成组装病毒的变体中的指定位置的频率而

言该给定氨基酸残基出现在肝脏中未鉴定的变体中对应于SEQ ID NO:1(推及于SEQ ID NO:2)的位置581至位置589的所述指定位置的频率,以确定优选肝脏去靶向的一组序列规则。参考下文表5和AAV5 VP1(SEQ ID NO:2),以下氨基酸可以以任何组合独立地排除,以提供能够组装并且与野生型AAV5 VP1衣壳(SEQ ID NO:1)相比较少靶向肝脏组织(“肝脏保留”或“肝脏去靶向”)的AAV VP1衣壳,其中肝脏去靶向性在此是指在具有降低的肝脏转导的病毒中优选的特性,而不是优选病毒体组装的特性。此外,可以允许Xaa1-Xaa9区域之外的一个或多个突变,只要衣壳仍然能够组装并表现出所需特征即可。

表 5 肝脏去靶向规则
Xaa1 不包括 K;
Xaa1 不包括 A、K、M 或 T; 或者
Xaa1 不包括 A、G、K、M、N、Q、R、S 或 T。
Xaa2 不包括 T;
Xaa2 不包括 A、S、T 或 V; 或者
Xaa2 不包括 A、C、H、I、K、S、T 或 V。
Xaa3 不包括 A 或 T;
Xaa3 不包括 A、M 或 T; 或者
Xaa3 不包括 A、G、H、K、M、N、Q、R、S、T 或 V。
Xaa4 不包括 P;
Xaa4 不包括 L、P、Q 或 T; 或者
Xaa4 不包括 L、M、P、Q、R、T 或 W。
[1752] Xaa5 不包括 Y;
Xaa5 不包括 H、I 或 Y; 或者
Xaa5 不包括 F、H、I、K、M、T 或 Y。
Xaa6 不包括 N;
Xaa6 不包括 N 或 Q; 或者
Xaa6 不包括 E、G、H、L、M、N、Q、T 或 W。
Xaa7 不包括 A;
Xaa7 不包括 A、C、H 或 M; 或者
Xaa7 不包括 A、C、G、H、L、M、R 或 S。
Xaa8 不包括 G;
Xaa8 不包括 G、M、Q 或 S; 或者
Xaa8 不包括 A、C、D、F、G、H、M、Q、S、V、W 或 Y。
Xaa9 不包括 G;
Xaa9 不包括 E、G 或 P; 或者
Xaa9 不包括 A、C、E、G、H、M、N、P、Q、S、V 或 W。

[1753] 分析肝脏、CNS、骨骼肌、心脏、肺、脾脏、淋巴结、骨髓、乳腺、皮肤、肾上腺、甲状腺、结肠、坐骨神经和脊髓组织以鉴定AAV5 VP1 581-589区域中驱动肝脏嗜性的氨基酸残基。

结果显示于图18B中，该图显示了相较于所有其他分析组织在肝脏组织中AAV5 VP1的581至589区域中各氨基酸的位置频率，其中衣壳多肽具有SEQ ID NO:2的序列。位置(“Xaan”，n＝1-9)在x轴上表示并且氨基酸残基在y轴上表示。特定位置的特定氨基酸残基的频率在热图中作为强度表示。较深的方块表示特定氨基酸在特定位置的频率较高，较浅的方块表示特定氨基酸在特定位置的频率较低。

[1754] 比较相对于给定氨基酸残基出现在所有其他分析组织中鉴定的变体中的指定位置的频率而言该给定氨基酸残基出现在肝脏中鉴定的变体中对应于SEQ ID NO:1(推及于SEQ ID NO:2)的位置581至位置589的所述指定位置的频率，以确定优选肝脏靶向的一组序列规则。参考下文表6A，和SEQ ID NO:2(AAV5 VP1)，以下氨基酸由此可以在Xaa1-Xaa9任何一个或多个位置处以任何组合独立地突变，以提供与野生型AAV5 VP1衣壳(SEQ ID NO:1)相比具有增强的肝脏嗜性的VP1衣壳，其中肝脏嗜性在此是指优选肝脏转导的特性，而不是优选转导所有其他所收获组织的特性。

[1755]

表 6A
肝脏嗜性规则

Xaa1 选自 A、G、K、M、N、Q、R、S 或 T
Xaa1 选自 A、K、Q 或 R
Xaa1 为 K
Xaa2 选自 A、C、I、K、S、T 或 V
Xaa2 选自 A、K、S 或 T
Xaa2 为 A
Xaa3 选自 A、G、I、K、M、Q、R、S、T 或 V
Xaa3 选自 A、K、Q、S 或 T
Xaa3 选自 K、Q 或 T
Xaa3 为 K

[1756]

Xaa4 选自 A、I、K、L、P、Q、R、S、T 或 V
Xaa4 选自 K、I、S 或 V
Xaa4 为 K
Xaa5 选自 F、I、L、M、T、V 或 Y。
Xaa5 选自 F、L 或 Y
Xaa5 为 F
Xaa6 选自 F、H、M、N、Q、S 或 Y
Xaa6 选自 M 或 N
Xaa6 为 N
Xaa7 选自 A、C、K、M、Q 或 S
Xaa7 选自 A、C 或 S
Xaa7 为 S
Xaa8 选自 A、C、F、G、M、Q 或 S
Xaa8 选自 A、C、M 或 S
Xaa8 为 C
Xaa9 选自 E、F、L、Q、R 或 Y
Xaa9 选自 L、Q 或 R
Xaa9 为 R

[1757] 参考下文表6B,和SEQ ID NO:2(AAV5 VP1),以下氨基酸可以由此以任何组合在Xaa1-Xaa9任何一个或多个位置处独立地突变,以提供与野生型AAV5 VP1衣壳(SEQ ID NO:1)相比具有降低的肝脏嗜性的VP1衣壳,其中肝脏嗜性在此是指对肝脏转导具有确定性的特性,而不是对所有其他所收获组织转导具有确定性的特性。

[1758]

表 6B **肝脏去靶向规则**
Xaa1 不选自 A、G、K、M、N、Q、R、S 或 T
Xaa1 不选自 A、K、Q 或 R
Xaa1 不是 K
Xaa2 不选自 A、C、I、K、S、T 或 V
Xaa2 不选自 A、K、S 或 T
Xaa2 不是 A

[1759]

Xaa3 不选自 A、G、I、K、M、Q、R、S、T 或 V
Xaa3 不选自 A、K、Q、S 或 T
Xaa3 不选自 K、Q 或 T
Xaa3 不是 K
Xaa4 不选自 A、I、K、L、P、Q、R、S、T 或 V
Xaa4 不选自 K、I、S 或 V
Xaa4 不是 K
Xaa5 不选自 F、I、L、M、T、V 或 Y。
Xaa5 不选自 F、L 或 Y
Xaa5 不是 F
Xaa6 不选自 F、H、M、N、Q、S 或 Y
Xaa6 不选自 M 或 N
Xaa6 不是 N
Xaa7 不选自 A、C、K、M、Q 或 S
Xaa7 不选自 A、C 或 S
Xaa7 不是 S
Xaa8 不选自 A、C、F、G、M、Q 或 S
Xaa8 不选自 A、C、M 或 S
Xaa8 不是 C
Xaa9 不选自 E、F、L、Q、R 或 Y
Xaa9 不选自 L、Q 或 R
Xaa9 不是 R

[1760] 下文表7提供了在肝脏组织中发现并且符合表6中提供的一种或多种规则的变体AAV VP1衣壳多肽的581至589区域的序列。因此,本公开提供了由具有SEQ ID NO:2的工程化AAV5 VP1衣壳多肽组成的rAAV,其中Xaa1至Xaa9区域是SEQ ID NO:43438-SEQ ID NO:44437中的任一者,如表7中所公开。本文还涵盖由工程化AAV5 VP2衣壳多肽和工程化AAV5 VP3衣壳多肽组成的rAAV,所述多肽在对应于AAV5 VP1 581至589区域中的氨基酸的AAV5 VP2(氨基酸残基445至453)和AAV5 VP3(氨基酸残基389-397)中的区域处具有下文表中公开的序列。

[1761] 表7

[1762] 驱动肝脏组织嗜性的AAV5 VP1衣壳多肽中581至589区域的序列

[1763]

SEQ ID NO	581-589 序列
43438	KADPATSIG
43439	KAEAMRNDR
43440	KAECYYGYE
43441	KAEGENVAK
43442	KAFPAPQGP
43443	KAGCMFWCE
43444	KAQHDTAVG
43445	KARVNCHGS
43446	KASDLNTQP
43447	KCCEYCDIA
43448	KCCTCVSFP
43449	KCDFWMGWM
43450	KCICSDDLK
43451	KCLSESFWG
43452	KCLWPFDWG
43453	KCLWPFNWG
43454	KCMDAKEHA
43455	KCQPCIVKC
43456	KCQSKDTHF
43457	KCSTDKAVE
43458	KCVVLENTQ
43459	KDHWECAGY
43460	KDNCVMWHW
43461	KDPKEQFFQ
43462	KDQQDQCND
43463	KEDERKSGT
43464	KETCTNMDL
43465	KFDLCQDSM
43466	KFWMPEMEF
43467	KGDLLDCLE
43468	KGEITPSGD
43469	KGGDGNYPF
43470	KGLDVHALN
43471	KGWKPDMTN
43472	KHAAYNACE
43473	KHQDIMHRG
43474	KIAQVDTWV
43475	KIDSQQSNY
43476	KIDSYGTSM
43477	KIEMQYRQS
43478	KIGYWWPQA
43479	KITPPLGNY
43480	KIYPECHYS
43481	KLEPNMADR
43482	KLLIHGPYE
43483	KMCGNCEAG
43484	KMEHMQHQS
43485	KMESRQLYN
43486	KMQHYDAEA
43487	KNDEYRMQN
43488	KNMLGDWPN
43489	KNMYYMYEA
43490	KPATQLDCA
43491	KPCKHMGGR
43492	KPDHMRWAL
43493	KPDQGGICK
43494	KPEGCLNSR
43495	KPFQALMMS
43496	KPGVHSHHG
43497	KPIAETVNW
43498	KPMGCHASS
43499	KPTPHAYDA
43500	KPWTQIASH
43501	KPWYERYQT
43502	KQATYPHAW
43503	KQDTKNLDE
43504	KQEMFGRFL
43505	KQIPFVSGD
43506	KQQMDHCAM
43507	KQTPYEMFH
43508	KRWVAEWIV
43509	KSDCQFRDH
43510	KSDRWAWGC
43511	KSELSTWGI
43512	KSEPPQVVT
43513	KSGYIDLYH
43514	KSIFLDSWE
43515	KSILKEDDG
43516	KSNGDMSGL
43517	KSNRDMNFA
43518	KSQMEMCIG
43519	KSRDNDGVL
43520	KSTTDSGCQ
43521	KSVQVPDAM
43522	KSWINEWLS
43523	KTDGPGTSW
43524	KTKPLDTVD
43525	KTLSDLAVQ
43526	KTLWALSCN
43527	KTNAEENFI
43528	KTNAWDTPC
43529	KTNPVKYVK
43530	KTQCTAQQN
43531	KTQNEKVWL
43532	KTQTPCFQH
43533	KTSWMMEAV
43534	KTTDGKMID
43535	KTTQWHEWA
43536	KVAQGLGPE
43537	KVCEIGTQG
43538	KVDAFEGRM
43539	KVECAQFDV
43540	KVEPSMLLD
43541	KVGINELFT
43542	KVHLSLWPQ
43543	KVIMYQHNG
43544	KVMWAGSKR
43545	KVNADGSRC
43546	KVPFACCSS
43547	KVRNKDPTM
43548	KVSVQGQVF
43549	KVTHCEAMN
43550	KWELYNNDP
43551	KWHTEDRNC
43552	KYNYKTYKT
43553	KACSAMDFC
43554	KALHQINED
43555	KAPMGWMNE
43556	KAQCALDTM
43557	KCENKDVGY
43558	KCNMNQEAE
43559	KCPQLNDNQ
43560	KCQQEQQYN
43561	KEGVENTGW
43562	KEMWNCDPA
43563	KETPFLMQS
43564	KFHVKPVDM
43565	KHMQSDHFT
43566	KHVYSEMGL
43567	KIHKNEDLF
43568	KKLDGHDKY
43569	KKVMGMGIY
43570	KLVQRCEVD
43571	KMLDWKNCS

[1764]

43572	KNGVDDAHN	43614	LASFTSPDP	43656	RADLWHIQN	43698	TANYGAAFA
43573	KNVVHEKQA	43615	LASTYWYSQ	43657	RAEDKWRSC	43699	TAQKSFMNC
43574	KQALMENFF	43616	LATLNAFQN	43658	RAERNSAEQ	43700	TATQCKAKE
43575	KQTTAEIGY	43617	LAVVRATLD	43659	RAMAMDSSH	43701	TAVTTWNMP
43576	KRDDCHNDN	43618	MAASDWHVM	43660	RAMDDQHYG	43702	VADSWAYNP
43577	KRTVIHEMM	43619	MAAWLEWGL	43661	RAMLVEQWG	43703	VAEWIMEKC
43578	KSHYGAGVT	43620	MAEIRTRHV	43662	RAQLFHQKT	43704	VAKQSADPI
43579	KSIDNYLCE	43621	MAHENTKAG	43663	RAQPTSCCG	43705	VALNVAQKG
43580	KSKRWNWDV	43622	MAMTLEISD	43664	RASVMPEAV	43706	VAMKKMTDD
43581	KSVCCQINN	43623	MASTYIERE	43665	SAAFQARLG	43707	VAMPTMKYK
43582	KTFCTQYMD	43624	MAVPMKMHC	43666	SAAFSAFNV	43708	VAQLALDEK
43583	KTFLVKTFI	43625	MAYLVVRQH	43667	SAAWQLQMT	43709	VARFTELCP
43584	KTRNCPKEA	43626	NAARMEQAT	43668	SACQKEWCP	43710	VARTFGEEK
43585	KTSLVVRRG	43627	NACTTKVGS	43669	SAEHKTHTM	43711	VASKFAGSV
43586	KTTLCVQWG	43628	NADYCSFPI	43670	SAGRRNIGM	43712	VATAQWLSC
43587	KTVQQVQMG	43629	NAENHTKTC	43671	SAHAGMHSN	43713	VAVQFREWL
43588	KTWWTEQSA	43630	NAFSMDIHV	43672	SAHKEDISG	43714	WAEHRDMRN
43589	KVAADLGHL	43631	NAQSDPHQG	43673	SAKSHLSNV	43715	WALTGYNIT
43590	KVEHYYDIG	43632	NASKYGDFL	43674	SAKYFSFGV	43716	WAQVCHCGE
43591	KVISHCDHY	43633	NASSTVINT	43675	SANMFVFGK	43717	YAAGAMEAS
43592	KWRVSVYVM	43634	NATANYCLG	43676	SANYYMTNM	43718	YADMPSERG
43593	KYEEFECMY	43635	NATTHSHTS	43677	SARVCDCPW	43719	YAELLQKMF
43594	KYICTRWHC	43636	NAVDRGIAT	43678	SASDPKNCY	43720	YAGCMLAIQ
43595	KYSPTPFMY	43637	NAVPIKYPY	43679	SASKFGWFS	43721	YAGSHEENC
43596	KYTPWDATF	43638	NAWSSNEEA	43680	SASWLDDQS	43722	YAHAIMGCQ
43597	IACRCNCFD	43639	PAQRVIANL	43681	SATGMRDYA	43723	YAHENQQTS
43598	IADSKRLQK	43640	QAALMTEQF	43682	SAWFTIVSG	43724	YANMKMICE
43599	IADVQRCHY	43641	QADRMLKTG	43683	TAAGYHFCK	43725	YAQTYYNKD
43600	IAGHYDSWK	43642	QADRWTGII	43684	TAAPWRHQS	43726	YARRFWTCV
43601	IAGMKVITP	43643	QAEWPGTMT	43685	TAAQFMDEL	43727	YARTTLANK
43602	IAHSVFTDH	43644	QAGLGSDAP	43686	TACLCLERP	43728	YASILTNQS
43603	IAQRTRQQQ	43645	QAHCVNLSS	43687	TADPLCRGA	43729	YASQSWQQE
43604	IAQTTLLPG	43646	QAMNSARTH	43688	TAGILDHYL	43730	YATELGEYG
43605	IATIKVHAS	43647	QAMTMLMQH	43689	TAHCVMNKQ	43731	YAVFGHQSR
43606	LAAPKTMCH	43648	QANNYNRMD	43690	TAHGIYTQF	43732	YAVLQNHYG
43607	LACFSYTNV	43649	QANTGRVYS	43691	TAHPGWEGY	43733	AALIGLIGD
43608	LAEPKKWHT	43650	QAQCAMGAR	43692	TAKFFADWP	43734	AANPTQWMT
43609	LAFQYCYPP	43651	QAQWMGCKW	43693	TALIISQNL	43735	AATMQTVPT
43610	LAGLKHAYK	43652	QARIENNQF	43694	TALLLDANP	43736	AAYMAKLRA
43611	LAHWQGLTY	43653	QATLQPVCI	43695	TAMPITAIP	43737	CAEQIMRNQ
43612	LALCIRHSW	43654	QATQTYDWH	43696	TAMQSGKVS	43738	CAQRSMRSM
43613	LAPPHKLDL	43655	QATYPITSE	43697	TANPEALVD	43739	CAQSFTVKG

[1765]

43740	CASHRKGAV	43782	SAYWFQMAG	43824	MQKLMALQA	43866	VFKLPEPVV
43741	DAENFVFKG	43783	TAEHIKVDL	43825	MQKTIQQWM	43867	VSKCGVSHI
43742	DANMTRTRP	43784	TAISEIPAP	43826	MWKVYEKSE	43868	VTKYKYEGM
43743	DANVIAHNP	43785	TAKMVAAME	43827	NHKWICLTR	43869	WKKMKYTWK
43744	EASFVVEPV	43786	TALMLNMAS	43828	NIKSRMIEF	43870	WMKPEFHYD
43745	FAIWEAHGR	43787	TAVTTNKWA	43829	NMKSQWQPS	43871	YSKYKRLWI
43746	FASHNYAPV	43788	VAAKNYKER	43830	NNKFWDYDP	43872	AEKPWNAMN
43747	FASRVCVDG	43789	VAMMVAGFQ	43831	NQKATTASY	43873	AIKFKYAPD
43748	GAEQNNECT	43790	VAMTDAKMC	43832	NQKWSTWNN	43874	ANKVGTAGN
43749	GAFTVEFSP	43791	VANVAIADM	43833	NVKCMQHMV	43875	AQKEYSKYL
43750	HAMQEELPP	43792	VASGSTEVC	43834	PRKNQVDLG	43876	CKKMGLHVM
43751	HAQFRGSAT	43793	VATMIATQL	43835	QCKQPLVIN	43877	CSKFRECCG
43752	IACTQQISV	43794	WAHKDWSNW	43836	QDKTKSCMY	43878	DFKTMNYMV
43753	IADWFDYEC	43795	WAMEDGKSA	43837	QGKYSANWY	43879	DFKWIQDQG
43754	IAMQYPEQY	43796	YAALTYYYS	43838	QKKFIQCYG	43880	DHKILGAIP
43755	IAPSVNQAT	43797	ICKQNYSRT	43839	QLKAPPAAH	43881	ELKVAISSG
43756	IAPWWSTTT	43798	IFKEGIHYQ	43840	QMKMLWFNV	43882	ENKTQFAYL
43757	LAEAGQNQV	43799	IPKNMGGCE	43841	QNKPASMHP	43883	FRKMGCGKT
43758	LAEGLKRTG	43800	IRKEVAVVP	43842	QQKPQECYS	43884	GLKLNLFQA
43759	LAGKSGQQV	43801	IRKNYRMKQ	43843	QVKDVVGNV	43885	GSKTCQQVM
43760	LAIEMVDRS	43802	ITKNRDRWP	43844	RGKTYEHGN	43886	LNKGQPLFM
43761	LAISFKGKW	43803	IVKLVLPVG	43845	RPKTFAALR	43887	LTKETGQVV
43762	LARGKDVAE	43804	IVKQEHESY	43846	RSKQARISH	43888	MLKTARWRD
43763	MAKIGYYMI	43805	IYKPNTQLQ	43847	SCKLNLHEG	43889	MSKPRVCCA
43764	MAKPNPPLG	43806	IYKTRHGWW	43848	SFKFYAEEA	43890	MVKTGVEEM
43765	NAAKYFWKQ	43807	IYKVLKPVH	43849	SGKQGNWRL	43891	MWKFAQLRL
43766	NADLQIFFR	43808	LCKQLGKQE	43850	SIKELVGCG	43892	MYKPEQWCM
43767	NALPESMTC	43809	LGKQEAVTN	43851	SIKMIDCFG	43893	QIKSCKNTD
43768	NAYPFVDHR	43810	LGKWKIQGT	43852	SMKWFITDD	43894	STKQHPHLN
43769	PAAMDYLMD	43811	LIKPLEISD	43853	SVKYCQSKD	43895	STKVPWWFP
43770	PACQETGMF	43812	LMKCLNWCC	43854	TFKSYEMET	43896	TMKFSHESA
43771	PAEAEKIGR	43813	LMKCNEAFT	43855	TFKYHYNAC	43897	TYKYKESSE
43772	QAEFRCGNH	43814	LQKSWNMLQ	43856	THKWGWWIA	43898	VQKSWEYDN
43773	QAFCEPNTP	43815	LVKRKELET	43857	TIKPKNEWC	43899	WSKKDIWAD
43774	QAGKWAVCE	43816	LVKWTNVLP	43858	TIKPYYPYQ	43900	IIPKDNKWR
43775	QALWYEDMT	43817	LYKMMYAFR	43859	TKKKRNGHL	43901	IPAKFRSWW
43776	QASCTHVWD	43818	MCKLKLEIP	43860	TRKDTLVAS	43902	ISVKGLTTH
43777	RAEWKQWSC	43819	MCKNQNKPA	43861	TSKTCFAEK	43903	IYCKIKEWG
43778	SAEWMQSWH	43820	MHKPEGMNG	43862	TTKEDFVCS	43904	LDHKCAAGM
43779	SAEWTHNNQ	43821	MIKPSVEGM	43863	TTKYTQSIF	43905	LHLKLINHC
43780	SAGMYARVA	43822	MLKCTHDLP	43864	TVKPAVPGE	43906	LLEKKQASG
43781	SATMCDSQS	43823	MNKTMINYA	43865	VEKSSIMNR	43907	LSQKPSNLI

[1766]

43908	LYAKKMAGY	43950	CVLKWATMW	43992	LSRDFCNWF	44034	YMSSFTCHW
43909	LYAKKMVGY	43951	DQMKWSQMG	43993	LVTSFASML	44035	YPMTFRAAC
43910	MGSKIHFMA	43952	DWVKATQMC	43994	MCDHFQMAY	44036	YYNNFICDK
43911	MQGKNMAVT	43953	ESDKEERMT	43995	MPRCFDAAA	44037	ACVAFSQQM
43912	NCVKAMDWY	43954	EVVKMPNHR	43996	MPVMFGLHC	44038	AHEIFNLQV
43913	NHEKTIASF	43955	EYIKWDGTP	43997	NMYPFMKYQ	44039	AHTQFLSAD
43914	NIDKRGPMM	43956	FCQKIEYYH	43998	NWITFSGAV	44040	AMQSFDMPG
43915	NMLKHMFGQ	43957	FGPKLTWHT	43999	QCDLFSEGY	44041	ARFLFYFDV
43916	NMVKYIHMG	43958	GCGKCECWN	44000	QHSVFMQIP	44042	CCVYFNRQM
43917	NRAKEHWHQ	43959	GTFKNPVHQ	44001	QNIRFTTDM	44043	CNIDFTADY
43918	NSYKLDWGW	43960	GVTKLEVWP	44002	QSNHFEQEK	44044	CSQIFWERG
43919	NTGKMEEKR	43961	GWDKKRDFR	44003	QYGSFNQNV	44045	CTPQFNRGF
43920	PFEKWNSQS	43962	IPIKLPPQG	44004	REEIFSYMV	44046	DEIAFWDGE
43921	PFRKSAGYM	43963	LEAKVAQYY	44005	RGIVFSQDN	44047	DGLMFHQQR
43922	QHTKPNCRD	43964	MDCKYPRDQ	44006	RKADFMHIT	44048	DIDQFASVN
43923	QLCKVGRDN	43965	NDGKIQFSG	44007	RPIWFIFVL	44049	DREAFDMAS
43924	QTEKLMRLD	43966	NLSKLAQMP	44008	RQQEFDRCR	44050	DSQWFEHNA
43925	QTNKWDVVT	43967	NLSKLTQMP	44009	RTDHFGHWE	44051	DTGPFMLFQ
43926	QTWKLITHG	43968	NYAKLFQDI	44010	RVEMFKWRT	44052	ETRPFEQVW
43927	QVHKGQKFS	43969	QKRKLQVYS	44011	SCHGFQNWP	44053	FTNRFQLMA
43928	QVNKITWVN	43970	QPAKKFWDF	44012	SIGLFMKDS	44054	GCCHFCDHL
43929	QVTKFVTEG	43971	QSSKWMWPL	44013	SKEWFSGTS	44055	HRMYFLRWP
43930	QWCKLINMC	43972	QWYKMGYNR	44014	SNAPFEQCN	44056	IICGFFAWS
43931	QWVKYHDIW	43973	RGPKDTNVK	44015	SNLSFGLTA	44057	LGHQFNNWT
43932	QYTKKYWNH	43974	RSLKYEGQQ	44016	SQGYFDANL	44058	MDGYFGFAV
43933	RVHKGQKFS	43975	SPNKSPFPG	44017	STQMFANVQ	44059	MGSNFVKQF
43934	SCAKFDVVE	43976	SSIKPMLSP	44018	TCMYFSCAL	44060	NICGFFAWS
43935	SDSKLHQVA	43977	SYHKQERCQ	44019	TEAWFQSAF	44061	QHCQFEQAI
43936	SGTKYCCWW	43978	TCVKAGLCL	44020	TIRNFWLQN	44062	QKDIFPMQM
43937	SIGKWPAYE	43979	TGLKVIGQH	44021	TNMLFLVWP	44063	QMELFCQQD
43938	SMMKCPTQY	43980	TIMKPLCFA	44022	TPCRFNSQR	44064	QNTEFMDHY
43939	SVWKCERIS	43981	TNHKFRCDR	44023	TQSCFHLEN	44065	QTSDFAMSS
43940	TCTKAVDHC	43982	TSSKCTMDA	44024	TRQFFHHGD	44066	QVMPFVIEG
43941	TYEKFNKHQ	43983	VHPKVQRVT	44025	TVEQFTNSV	44067	SKMNFGTAN
43942	TYIKSYQMA	43984	YSDKMESSN	44026	TVRMFCRAI	44068	SKTEFSLKM
43943	VRDKGVQCA	43985	IHCMFGAGA	44027	TWHSFPNQE	44069	TGDLFCCHL
43944	VSQKSPTAW	43986	IQQYFNWCE	44028	TYISFQDEK	44070	TQGLFAQVG
43945	WCTKCNYNQ	43987	IRDMFEVGR	44029	VSQSFDHVN	44071	VCFHFRDAA
43946	WCVKRMENH	43988	LEFEFYWDE	44030	VWQPFSGDH	44072	VCMLFQHEA
43947	WVEKHLMLG	43989	LIHPFARGN	44031	YCCWFEHRL	44073	VHSRFGQYT
43948	YSQKPEEAN	43990	LIQAFPRDT	44032	YGCHFQRQE	44074	WQPRFYHIW
43949	CTYKTGSPA	43991	LPTPFMAHL	44033	YHMGFTGMG	44075	YEDPFRHYS

[1767]

44076	ICHSGNRLL	44118	PRCYINWRY	44160	GPDRYNEES	44202	MGDDSMSEH
44077	ICQGSNSMA	44119	PWQFTNLEK	44161	GSLALNVRQ	44203	MPRTCMSTN
44078	IGTVRNWCW	44120	QCCHPNGYQ	44162	GTQRDNTGF	44204	NCEQWKSMF
44079	IICPTNEYM	44121	QNCWYNFFR	44163	HCICNNSGA	44205	NDQTSSSAW
44080	IKILLNEMT	44122	QQCNTNTDS	44164	HMTLGNLCP	44206	NIQFHTSGQ
44081	IQCFHNMEE	44123	QVVTVNASD	44165	ICGWENCSE	44207	NKDVDRSHV
44082	IVMCHNVYP	44124	RIGIANDWV	44166	LCINVNPSI	44208	NLVRWESGW
44083	IVNLNNHHW	44125	SDGTINRMN	44167	LFTEINGCY	44209	NMSPWMSDR
44084	LCAQHNIVC	44126	SHLQCNYMR	44168	LMCQMNNRS	44210	NPCIVISSV
44085	LETEGNPFY	44127	SLTVWNQNG	44169	MPGFENKMP	44211	NSYHLESMD
44086	LGIGNNWEQ	44128	STMWRNRES	44170	MPRDENHIG	44212	NVAMYASSE
44087	LITPPNWWH	44129	STVYWNMET	44171	NFDCSNDQT	44213	NVLCVKSVQ
44088	LLEIMNEQC	44130	SYTSGNESW	44172	NLLVSNVTT	44214	NVMICHSPA
44089	LLFWYNHMS	44131	TGAPGNWFH	44173	PDGPMNLLG	44215	NYCTLMSAC
44090	LMILRNTHQ	44132	TRNAYNCHQ	44174	PMMMANRIA	44216	PILITHSVE
44091	LQMPYNWQT	44133	TSSWVNWDP	44175	QHRQANGIH	44217	QIGWEASSL
44092	LRHICNCGQ	44134	TTMPENIPL	44176	QMSYVNQYR	44218	QIVELHSLG
44093	LSLSKNCAG	44135	TTRFHNNVM	44177	QSVSNNRRA	44219	QIVQAMSSE
44094	LTPLANQYC	44136	TVWNTNTGW	44178	QTFMTNFDT	44220	QLADNHSRP
44095	LTRFNNDLP	44137	VCRNGNRQA	44179	QYSPANQWD	44221	QSLYDRSGT
44096	LVPRINMLW	44138	VIFPENSHY	44180	RTFLTNPKD	44222	QSTCYFSNC
44097	LYELHNTWD	44139	VLTWDNPDE	44181	SMHDYNFNV	44223	QTREGVSYY
44098	MFIQDNAPT	44140	VTHATNVLE	44182	STELKNMER	44224	QVDMARSSF
44099	MLLYPNNTN	44141	WNGFMNQGY	44183	SVLYSNYFA	44225	QVMENSSWA
44100	MMCSNNSHR	44142	WPVRLNEYK	44184	TGAMTNYCY	44226	RECVHLSCP
44101	MQSPENTVY	44143	YIAVPNKMS	44185	TMMQNNTEV	44227	REGLLWSNS
44102	MTCLQNQSE	44144	YYECRNYSD	44186	TRDTHNSVD	44228	RISQQWSRS
44103	MTIANNIRN	44145	AKVYSNLYG	44187	TRIHVNRYG	44229	RTIQKESHV
44104	MTYNYNSKD	44146	ATILVNRDG	44188	TTWLDNEDA	44230	RWCMKISAN
44105	MVWICNAPR	44147	AWFTHNALP	44189	VIRPINQEW	44231	SCTHVMSNA
44106	NCGYKNNKA	44148	CIHHRNDKV	44190	WQAPHNQSY	44232	SHELQLSGL
44107	NICFVNDAL	44149	CNYDTNLRP	44191	YVITANNHQ	44233	SHMNQASMP
44108	NILSPNCYE	44150	CTQSINGLI	44192	ICQHIGSWG	44234	SITRWVSAE
44109	NKVGMNYMQ	44151	CTVLINCRD	44193	IDHDYASDS	44235	SLYMEMSWQ
44110	NMCHHNWRD	44152	CVTWENVMT	44194	ILIMILSHS	44236	SWCSIHSPN
44111	NPAYKNWGW	44153	DLVHCNNHM	44195	IPRFVYSPV	44237	SYIEWGSWM
44112	NQIYTNPWG	44154	DRGYYNSMM	44196	ITVADRSSH	44238	TCERKRSDK
44113	NSLISNAWT	44155	DTIHENEGV	44197	LEGATESLW	44239	TCTYDKSQI
44114	NTMTPNFRK	44156	EPHIMNSHE	44198	LLPLHGSWH	44240	THIQCRSMM
44115	NYVAKNYMV	44157	EQRVYNHQY	44199	MCLAPTSMH	44241	THWWQPSSG
44116	PKFIANRQE	44158	FTVTMNCTA	44200	MCSPEPSLQ	44242	TKHMIQSYG
44117	PNYWINYQN	44159	GGALDNGQC	44201	MEHQQDSRF	44243	TSARCISAE

[1768]

44244	TSTHVKSYQ	44286	QQTRHRSCV	44328	RQYIQDVCD	44370	NCHRNDACD
44245	TTHFAVSGI	44287	RTEWHQSQA	44329	RREPTKCCW	44371	NENQMMHCA
44246	TTSHCVSPS	44288	SMQPTGSGW	44330	RVTEWFQCG	44372	NTAQQLCCE
44247	TVHMCQSFV	44289	STQWEQSMQ	44331	RWCGKGYCV	44373	NTTQRALCG
44248	VGTAAHSQR	44290	SWVSKFSEI	44332	SFSYAECCA	44374	PTRTRRTCR
44249	VIAQELSAK	44291	TNMPMPSYE	44333	SHENQWACI	44375	QINLDSQCM
44250	VSDWGHSDS	44292	TQGRMYSCW	44334	SHMFQHTCG	44376	QSAIEEVCR
44251	VWWREDSTN	44293	TWGFNRSYC	44335	SNTLYDTCY	44377	QSHPRITCA
44252	WLEPCVSSL	44294	VQHCSHSLD	44336	SQFMDDDCA	44378	QVQMGRDCG
44253	WPETVGSSE	44295	VTYEAKSQW	44337	SRNAMQYCP	44379	RCITPTGCI
44254	WYAPAESSN	44296	WDSQPQSGV	44338	SSHIDGICN	44380	SCCYCPTCC
44255	YCCRPVSDF	44297	WELNNPSVM	44339	TELVYMTCQ	44381	SMCFKGTCE
44256	YHMDGDSRH	44298	WLTWHFSSF	44340	TGDSRMDCQ	44382	SRNANDGCF
44257	ADETQYSCH	44299	IIATGWECD	44341	TKGFMSVCR	44383	SVFMRDECC
44258	AFNQHTSMD	44300	ISLPMLECI	44342	TNYPNWICH	44384	TCWQWLKCE
44259	AISPWMSCA	44301	LKIYTKCCD	44343	TQEAISDCF	44385	TMDRQPMCI
44260	AMYNVSSAA	44302	LSSHSEICQ	44344	TQQFCHDCQ	44386	TSSTEQMCV
44261	AMYQNPSPC	44303	MCFDIATCA	44345	TTGMLENCN	44387	IGRITMGDR
44262	ARAEWKSKD	44304	MEHFLQVCN	44346	TWTQMIDCY	44388	IHTLKHDYR
44263	CVCADTSVF	44305	MKWPAFACA	44347	TYVRRWMCL	44389	IMFMVAFDR
44264	DIELVTSEQ	44306	MMNVHQNCE	44348	VIHGSLQCQ	44390	LHVPPSWAR
44265	ELAAPHSQV	44307	MPVRLTKCS	44349	VIMMDGYCT	44391	LMCTMLTDR
44266	ERATLKSMC	44308	MWVDCCECH	44350	VIMTQMNCW	44392	LNMELMHLR
44267	ERMWKTSDN	44309	NGGQKATCP	44351	VPNTKDACF	44393	MDFLSCGLR
44268	ESIRLRSSQ	44310	NGWFVRICG	44352	VWCGNQWCN	44394	MELDVSNDR
44269	FNAVHMSLE	44311	NNASPSYCG	44353	WFTNVLWCG	44395	MFMMMHDIR
44270	GGRFNQSWA	44312	NRPHQWLCC	44354	YPIMERWCH	44396	MNHWNMHMR
44271	GHFMTLSKG	44313	NSHACSECS	44355	YSEICDRCI	44397	NDPHQHDAR
44272	GNERMFSSA	44314	PNNQCSYCC	44356	AFAQLGQCH	44398	NFLDCSERR
44273	HLTMWESFD	44315	PRTEDLQCI	44357	ALVLRCICL	44399	NPNLIEKHR
44274	LCTFNGSNG	44316	QCMAHIQCD	44358	CHPDKMECP	44400	QCLLTLQNR
44275	LKGTMMSPI	44317	QEMDGWYCC	44359	CVMVIWLCS	44401	QEHHVFGVR
44276	LTYLKRSIG	44318	QGAIYEVCL	44360	DGSCKWNCH	44402	QITHLQVGR
44277	MLNLHGSDM	44319	QIFLWFYCG	44361	DSMRKYQCR	44403	QSSFNPDMR
44278	MSIPASSEV	44320	QKQSTVICE	44362	DTVMQAACW	44404	QYEVKMWSR
44279	NIGYRKSSG	44321	QLNVTPICC	44363	GGPHREFCI	44405	RNDDNLERR
44280	NRESDVSTG	44322	QMAPYSICQ	44364	HTIDKMNCP	44406	RTFYKQVAR
44281	PCQWPVSRA	44323	QNGNIEKCQ	44365	IPYQQDCCA	44407	SCSQIQEIR
44282	PCVYECSLH	44324	QQNPNGCCH	44366	LILHMLECQ	44408	SEGCRMPPR
44283	PTIWLHSYL	44325	QVSSHVECV	44367	MHTSMLICC	44409	SFAVMMEDR
44284	PVCPWGSFD	44326	RLSAGDACY	44368	MKAATMVCL	44410	SPNLIEKHR
44285	PWVAVESAE	44327	RPVLECNCE	44369	MKFTLPQCN	44411	SSSSRRDYR

[1769]

44412	SVYLKLRIR	44419	WEHHCAEWR	44426	DHAHYWYGR	44433	HWCTIGKLR
44413	SYLQWHVTR	44420	WLPFAMTNR	44427	DSVSQQVER	44434	MGADYGVWR
44414	TGYFYMAHR	44421	YCYRAKEQR	44428	ECCNYVWER	44435	NFNQCTIYR
44415	TPAFTHWER	44422	YIEPPALGR	44429	ESQYSQGTR	44436	QKQGTICER
44416	TPMSTQNVR	44423	YINMNAQER	44430	FMSMQGDQR	44437	QSAIEEVYR
44417	TTPQYMRER	44424	YKLCRVKGR	44431	GKNCWPIKR		
44418	VNMPLYKFR	44425	YMLVDMIFR	44432	HTGPMLGDR		

[1770] 实施例5

[1771] 在CNS中具有组织嗜性的AAV5变体

[1772] 此实施例描述了使用实施例1-实施例2中描述的方法和系统发现的在CNS中具有组织嗜性的工程化AAV5变体。

[1773] 表8是1000个变体(CNS变体)的列表,这些变体在两个皮层测序样本中具有最高的归一化UMI计数,但在两个肝脏测序样本中不存在。该列表的最低纳入标准包括在两个质量过滤的皮层NGS数据集中存在;并且排除标准是在两个独立的肝脏测序样本中检测到给定的变体。在对四个数据集进行分析后,符合这些标准的变体总数为2108。然后对两个单独的皮层NGS数据集的UMI计数进行独立归一化,接着根据两个独立归一化的UMI计数的总和对变体序列进行排序(最大可能=200%),前1000名在此处列出(SEQ ID NO:115-SEQ ID NO:1114)。变体残基在位置581-589处。

[1774] 表8

[1775]

SEQ ID NO	var_aa	归一化的2个样本(最大可能评分200%)
SEQ ID NO: 115	RVTIMFTGT	112.28%
SEQ ID NO: 116	YTGCIDGFL	100.01%
SEQ ID NO: 117	TNMSTQPCW	97.08%
SEQ ID NO: 118	VWTTVCNDY	96.78%
SEQ ID NO: 119	DYANTLKQS	91.84%
SEQ ID NO: 120	EFVLGCQGI	76.88%
SEQ ID NO: 121	TIMWTMMRE	76.32%
SEQ ID NO: 122	SPMVVTHSD	76.07%

[1776]

SEQ ID NO: 123	DWYDIWTLV	69.59%
SEQ ID NO: 124	REFAISGGN	69.44%
SEQ ID NO: 125	SITNCNHDR	69.01%
SEQ ID NO: 126	HIQEELEEQ	64.92%
SEQ ID NO: 127	DINGKKNIC	64.05%
SEQ ID NO: 128	TLAEHFTSH	63.45%
SEQ ID NO: 129	QCQSQEWFH	60.94%
SEQ ID NO: 130	ARSGINGHE	60.24%
SEQ ID NO: 131	CIRQYDAID	59.65%
SEQ ID NO: 132	ILIFESRIG	58.77%
SEQ ID NO: 133	HFADNQVVQ	56.73%
SEQ ID NO: 134	NESHVERHC	56.14%
SEQ ID NO: 135	TMVTSVTLP	55.86%
SEQ ID NO: 136	PTLSCPYHK	54.39%
SEQ ID NO: 137	TVAKVLSGH	53.82%
SEQ ID NO: 138	KIFITTDEG	52.65%
SEQ ID NO: 139	KSAQVYWHP	52.63%
SEQ ID NO: 140	CSALFAQHD	52.05%
SEQ ID NO: 141	AGSACMPAH	50.59%
SEQ ID NO:142	IVIWQGEWI	50.29%
SEQ ID NO:143	SYCEVEDAP	50.00%
SEQ ID NO:144	DCSWWLQLG	49.42%
SEQ ID NO:145	NFCKQGNTC	47.08%
SEQ ID NO:146	HIHDTMSDY	46.81%
SEQ ID NO:147	QYVHMLGEC	46.50%
SEQ ID NO:148	RQHHCLPVN	46.22%
SEQ ID NO:149	AHVDHQCSG	45.43%
SEQ ID NO:150	SFSDTHLGW	45.40%
SEQ ID NO:151	AASQMWQRG	45.04%

SEQ ID NO:152	NDCTCRAIM	44.16%
SEQ ID NO:153	NSHEYLDRH	43.86%
SEQ ID NO:154	DWIWGTDMW	43.60%
SEQ ID NO:155	MIFYRTYGF	42.40%
SEQ ID NO:156	HPAPLNMPH	42.11%
SEQ ID NO:157	SQQHCQHDQ	41.58%
SEQ ID NO:158	HEVMMCPAP	41.01%
SEQ ID NO:159	KRTDLHAAD	40.95%
SEQ ID NO:160	EYFKPKSQD	40.94%
SEQ ID NO:161	KLAHSWCIW	40.94%
SEQ ID NO:162	NAMANFSVR	40.64%
SEQ ID NO:163	EVTTPRPPT	39.50%
SEQ ID NO:164	YFHDEQIDY	38.90%
SEQ ID NO:165	LTHQGGRQR	38.31%
SEQ ID NO:166	AQWYMWCWD	37.73%
SEQ ID NO:167	RMMWQSHCD	37.44%
SEQ ID NO:168	EIHKRIAMT	36.55%
SEQ ID NO:169	CMSCCKWPH	36.42%
SEQ ID NO:170	DFQRPTETQ	35.67%
SEQ ID NO:171	SFMPCHCRG	35.67%
SEQ ID NO:172	LTHNTPFPK	35.40%
SEQ ID NO:173	HYVVRHEDP	35.14%
SEQ ID NO:174	FQEAPPHVH	34.80%
SEQ ID NO:175	YAELEWKVC	33.92%
SEQ ID NO:176	RCVAEPNEP	33.37%
SEQ ID NO:177	GLIFHQSLS	33.34%
SEQ ID NO:178	ASHGMDEYN	33.33%
SEQ ID NO:179	PIACRAINW	33.33%
SEQ ID NO:180	NCKPDQIAM	32.69%

[1777]

SEQ ID NO:181	FRMWFTECI	31.87%
SEQ ID NO:182	MGYAKHDSH	31.59%
SEQ ID NO:183	DTSCLELMQ	31.29%
SEQ ID NO:184	KVAGGHETT	31.16%
SEQ ID NO:185	YSRYHHSLL	30.99%
SEQ ID NO:186	VADTWQVHC	30.43%
SEQ ID NO:187	LWGYGAVNM	29.25%
SEQ ID NO:188	FGRQAYVAW	29.24%
SEQ ID NO:189	EEVQFTQQD	29.24%
SEQ ID NO:190	GTGHIGNAT	29.24%
SEQ ID NO:191	AYTYNSKLY	28.95%
SEQ ID NO:192	FIFRKCDHW	28.37%
SEQ ID NO:193	QCVDCPPYC	27.79%
SEQ ID NO:194	ETPPLCHGP	27.78%
SEQ ID NO:195	RMTPNNCQR	27.28%
SEQ ID NO:196	NSTPMACME	26.91%
SEQ ID NO:197	STHPLNLYH	26.90%
SEQ ID NO:198	TTVPDPGPR	26.32%
SEQ ID NO:199	LIWSVYQDE	26.09%
SEQ ID NO:200	YDNHQKARN	26.02%
SEQ ID NO:201	SIYQKWVTN	25.92%
SEQ ID NO:202	MQSMQHWHP	25.74%
SEQ ID NO:203	TLVSCDQGN	25.73%
SEQ ID NO:204	WSITGQPWF	25.73%
SEQ ID NO:205	YQKVEYWNY	25.73%
SEQ ID NO:206	AHYQQSTDR	25.44%
SEQ ID NO:207	QVRVITHKS	25.17%
SEQ ID NO:208	WAAKPGDPK	25.15%
SEQ ID NO:209	RSDNDWCFL	25.15%

SEQ ID NO:210	ETHRGTVGL	24.89%
SEQ ID NO:211	KCSVQEKKA	24.89%
SEQ ID NO:212	WVEPEQCFG	24.56%
SEQ ID NO:213	NHQQCPKCP	24.28%
SEQ ID NO:214	AAHSAPFFD	24.27%
SEQ ID NO:215	KGQRRDSVS	24.07%
SEQ ID NO:216	DVMQNQCKS	23.98%
SEQ ID NO:217	ALVYSGWIC	23.72%
SEQ ID NO:218	FAKARHNWR	23.69%
SEQ ID NO:219	TYIQYPTNA	23.69%
SEQ ID NO:220	VSKFIAASW	23.69%
SEQ ID NO:221	RQREGKHGN	23.68%
SEQ ID NO:222	MAAHTKYRF	23.40%
SEQ ID NO:223	FVGLMNQCG	23.39%
SEQ ID NO:224	TWQSMVNSE	23.15%
SEQ ID NO:225	SAITCRMLK	23.10%
SEQ ID NO:226	SRHYEQFEC	23.10%
SEQ ID NO:227	VTAGQLGLN	23.10%
SEQ ID NO:228	DSQYHVKCG	22.82%
SEQ ID NO:229	LSECSDRAN	22.51%
SEQ ID NO:230	YAYPTEGQI	22.22%
SEQ ID NO:231	KRFVDMPEM	21.95%
SEQ ID NO:232	VHKFQGNFG	21.94%
SEQ ID NO:233	RGHEPPMWL	21.93%
SEQ ID NO:234	QKCYCDEQC	21.64%
SEQ ID NO:235	RHAPTDISP	21.60%
SEQ ID NO:236	MTGIGKCWN	21.05%
SEQ ID NO:237	QAYCCMHQY	20.48%
SEQ ID NO:238	KHDDCCTGT	20.47%

[1778]

SEQ ID NO:239	KWYQFGMVA	20.42%
SEQ ID NO:240	LTWECKFGG	20.18%
SEQ ID NO:241	AIVQQHYYP	19.88%
SEQ ID NO:242	EENVWMQYQ	19.88%
SEQ ID NO:243	NSQPMDFLP	19.88%
SEQ ID NO:244	FDTAFQGDI	19.59%
SEQ ID NO:245	QMEARMDCE	19.49%
SEQ ID NO:246	NWVPQIEHW	19.33%
SEQ ID NO:247	VAILGLNPT	19.31%
SEQ ID NO:248	ANVPMCSLQ	19.31%
SEQ ID NO:249	LIPMDWIGM	19.30%
SEQ ID NO:250	IYCGNSDMQ	18.73%
SEQ ID NO:251	PAETLMDGF	18.71%
SEQ ID NO:252	FSYQIQCQN	18.43%
SEQ ID NO:253	TMRLFWDMG	18.42%
SEQ ID NO:254	QRVPMQLVE	18.42%
SEQ ID NO:255	AMTRQNGWM	18.16%
SEQ ID NO:256	GMRTHSVMN	18.13%
SEQ ID NO:257	LITVQHIRN	18.13%
SEQ ID NO:258	TTYPHNTHG	17.88%
SEQ ID NO:259	RIEMDEWEA	17.85%
SEQ ID NO:260	DNTPEHCYM	17.57%
SEQ ID NO:261	SSVVMAKQP	17.55%
SEQ ID NO:262	CLVRIQKDH	17.54%
SEQ ID NO:263	TLPKMHMWP	17.54%
SEQ ID NO:264	MPHHSPDCW	17.32%
SEQ ID NO:265	QELECHACS	17.31%
SEQ ID NO:266	HMHNLPVKP	17.31%
SEQ ID NO:267	GIAVWFDFQ	17.25%

SEQ ID NO:268	ANRSRPTWQ	17.25%
SEQ ID NO:269	KTFPILEAW	17.25%
SEQ ID NO:270	ALQYCQGKD	16.97%
SEQ ID NO:271	YHFEYMTSF	16.97%
SEQ ID NO:272	VCGHRNDGN	16.96%
SEQ ID NO:273	EICAAMDSF	16.68%
SEQ ID NO:274	YWFVLNPCV	16.41%
SEQ ID NO:275	QQNPCLGGM	16.32%
SEQ ID NO:276	FNWFVRAYI	16.14%
SEQ ID NO:277	DCDPYNMVD	16.12%
SEQ ID NO:278	RTGDAGMSP	16.10%
SEQ ID NO:279	LEFNGHLFN	16.08%
SEQ ID NO:280	VYVAAGGGC	15.92%
SEQ ID NO:281	MTTGLMAYE	15.87%
SEQ ID NO:282	MAIDNLECK	15.83%
SEQ ID NO:283	QKAIVAQND	15.81%
SEQ ID NO:284	GGDGRVHNK	15.52%
SEQ ID NO:285	THDAARKTA	15.50%
SEQ ID NO:286	TWNEPHQNQ	15.20%
SEQ ID NO:287	ANRADWQHL	14.96%
SEQ ID NO:288	CTKAYRKPG	14.93%
SEQ ID NO:289	RTAVIQCIE	14.92%
SEQ ID NO:290	WKVSPIQGF	14.92%
SEQ ID NO:291	MCGEYQPGS	14.91%
SEQ ID NO:292	NHHPYTFHQ	14.91%
SEQ ID NO:293	NPRYCLSLS	14.64%
SEQ ID NO:294	SRASQQSMW	14.64%
SEQ ID NO:295	VLKSCDRHG	14.64%
SEQ ID NO:296	KEARSQTGD	14.63%

[1779]

SEQ ID NO:297	FANGCTGSI	14.63%
SEQ ID NO:298	GAMAAPYHD	14.62%
SEQ ID NO:299	YEHPIMWPV	14.62%
SEQ ID NO:300	CYGWTGCKC	14.33%
SEQ ID NO:301	AAAMRDKTD	14.33%
SEQ ID NO:302	KLRIGVVCN	14.33%
SEQ ID NO:303	TPHSIHTDK	14.33%
SEQ ID NO:304	QTETWEFGA	14.26%
SEQ ID NO:305	KAEMKMGCN	14.16%
SEQ ID NO:306	VPICCNLPM	14.11%
SEQ ID NO:307	LPSHCNLGS	14.05%
SEQ ID NO:308	SMNRTHLER	14.04%
SEQ ID NO:309	FGEFTMDEN	13.84%
SEQ ID NO:310	YTAHNANVV	13.78%
SEQ ID NO:311	GTMKHNGRA	13.76%
SEQ ID NO:312	EHYCKDNVL	13.75%
SEQ ID NO:313	TKRMMEDSG	13.75%
SEQ ID NO:314	CERDCNCTA	13.74%
SEQ ID NO:315	TGRVTADMA	13.74%
SEQ ID NO:316	YNGDRLFCF	13.53%
SEQ ID NO:317	NHEVVAFMQ	13.47%
SEQ ID NO:318	QNEICVHHE	13.47%
SEQ ID NO:319	LRSGMIVSG	13.46%
SEQ ID NO:320	QPKMWMMAS	13.45%
SEQ ID NO:321	DVRHSWSEK	13.45%
SEQ ID NO:322	EILQDVYYL	13.45%
SEQ ID NO:323	SVVNNLKAS	13.45%
SEQ ID NO:324	RCTDCSYFY	13.29%
SEQ ID NO:325	GSTSRNQRD	13.17%

SEQ ID NO:326	LDCNAMPST	13.16%
SEQ ID NO:327	QGRWACWYS	13.16%
SEQ ID NO:328	TIEWYPDSQ	13.16%
SEQ ID NO:329	KAVSYHHDG	13.16%
SEQ ID NO:330	TMEHCQRPQ	13.16%
SEQ ID NO:331	RMFWDTSDR	12.89%
SEQ ID NO:332	MCNPKMMSS	12.88%
SEQ ID NO:333	DSHDMFVYK	12.87%
SEQ ID NO:334	LTCYHNELS	12.87%
SEQ ID NO:335	AVVEHAVNR	12.87%
SEQ ID NO:336	CGQMVICSE	12.87%
SEQ ID NO:337	NLRDAHWCM	12.87%
SEQ ID NO:338	ESSQEIKTC	12.77%
SEQ ID NO:339	AKMWQLENP	12.73%
SEQ ID NO:340	GTIPKQHEF	12.59%
SEQ ID NO:341	ARDCYWKAR	12.58%
SEQ ID NO:342	AADREQGWW	12.57%
SEQ ID NO:343	DSVQQRWYV	12.57%
SEQ ID NO:344	KCGQDWIHG	12.33%
SEQ ID NO:345	QVYAPEIEG	12.32%
SEQ ID NO:346	CCCVPRSIY	12.29%
SEQ ID NO:347	AAHGEMKTA	12.28%
SEQ ID NO:348	QFWLMHEWP	12.28%
SEQ ID NO:349	KFAMDWVSD	12.00%
SEQ ID NO:350	TSDSVINWL	12.00%
SEQ ID NO:351	QIIYRCRST	12.00%
SEQ ID NO:352	GYDLRGADE	11.99%
SEQ ID NO:353	CIRFMMREG	11.99%
SEQ ID NO:354	HSNTHGYPL	11.73%

[1780]

SEQ ID NO:355	HGNMPHCYP	11.72%
SEQ ID NO:356	VGLSIKMYG	11.71%
SEQ ID NO:357	ESWRYKHQA	11.70%
SEQ ID NO:358	PKRYQAWTW	11.70%
SEQ ID NO:359	SCSQKLMMQ	11.70%
SEQ ID NO:360	ALVTYRSMQ	11.70%
SEQ ID NO:361	ALNSDMAAW	11.60%
SEQ ID NO:362	NHSLCWDSK	11.60%
SEQ ID NO:363	KKEDYAKFF	11.41%
SEQ ID NO:364	RQYLYKRQQ	11.41%
SEQ ID NO:365	TAMPEHRWD	11.41%
SEQ ID NO:366	GFCTIPVAD	11.40%
SEQ ID NO:367	IAVQKRLFC	11.40%
SEQ ID NO:368	LAEIKDWVP	11.40%
SEQ ID NO:369	MLPMFMGQN	11.40%
SEQ ID NO:370	NVRTYIDDS	11.40%
SEQ ID NO:371	YFWISKALY	11.13%
SEQ ID NO:372	ISYQPNPME	11.12%
SEQ ID NO:373	ACDDWCWMC	11.11%
SEQ ID NO:374	QTEMKKCAR	11.11%
SEQ ID NO:375	ATTHKVKDN	11.11%
SEQ ID NO:376	DTFKCMNVS	11.11%
SEQ ID NO:377	QSFSYDAKE	11.11%
SEQ ID NO:378	VIGCQQGLG	11.11%
SEQ ID NO:379	YGPEILKLT	11.11%
SEQ ID NO:380	WAAQMYQQE	11.09%
SEQ ID NO:381	DIACLDQWQ	10.93%
SEQ ID NO:382	TSLVMPSLP	10.89%
SEQ ID NO:383	MHSQQRAIK	10.84%

SEQ ID NO:384	VRQSVPHWA	10.83%
SEQ ID NO:385	DTRPSRDSS	10.82%
SEQ ID NO:386	RSGAKCYCD	10.82%
SEQ ID NO:387	ECKTHRDVY	10.63%
SEQ ID NO:388	DFLHVTSIM	10.55%
SEQ ID NO:389	QTKFMDDMH	10.54%
SEQ ID NO:390	TSMTERRTV	10.54%
SEQ ID NO:391	WTTWAIPCC	10.53%
SEQ ID NO:392	LPTLRVADK	10.53%
SEQ ID NO:393	KTGTTMEPM	10.52%
SEQ ID NO:394	TKQFVHNED	10.50%
SEQ ID NO:395	TCQEPHNIT	10.28%
SEQ ID NO:396	QKVASMGCW	10.27%
SEQ ID NO:397	CKGHDAGEY	10.27%
SEQ ID NO:398	FPWSRPKCW	10.26%
SEQ ID NO:399	HAMSLSTFQ	10.25%
SEQ ID NO:400	MTCVWHSDF	10.25%
SEQ ID NO:401	VKAWWHDHQ	10.24%
SEQ ID NO:402	DVLKCFINE	10.24%
SEQ ID NO:403	PIKGVAVQE	10.24%
SEQ ID NO:404	LSRLYPNTY	10.23%
SEQ ID NO:405	KQTVMTDTS	9.96%
SEQ ID NO:406	SATPHWYCH	9.95%
SEQ ID NO:407	KTTDLFQYE	9.94%
SEQ ID NO:408	VTWINLRSG	9.92%
SEQ ID NO:409	IHQLVGAFW	9.66%
SEQ ID NO:410	CVNLMAVNN	9.66%
SEQ ID NO:411	EKQKLTVSS	9.66%
SEQ ID NO:412	RLDHWATCD	9.65%

[1781]

SEQ ID NO:413	AVEFWLERM	9.65%
SEQ ID NO:414	KTNNTVCHS	9.65%
SEQ ID NO:415	HHLRNVHHC	9.44%
SEQ ID NO:416	ARGTCNYCY	9.37%
SEQ ID NO:417	STEENEALY	9.37%
SEQ ID NO:418	KKQQISIRL	9.36%
SEQ ID NO:419	SAAEKAMDQ	9.36%
SEQ ID NO:420	TAPGGSTPQ	9.36%
SEQ ID NO:421	QTNQVVVVE	9.36%
SEQ ID NO:422	DSTPNLFDP	9.08%
SEQ ID NO:423	DVASYCQLQ	9.08%
SEQ ID NO:424	TQCKQLTVM	9.08%
SEQ ID NO:425	TMEHHECMW	9.07%
SEQ ID NO:426	DVLVCEILW	9.07%
SEQ ID NO:427	NIYSGINNP	9.07%
SEQ ID NO:428	MFPKKMCLH	8.83%
SEQ ID NO:429	THQNTWLFY	8.83%
SEQ ID NO:430	NHGYRNNSM	8.78%
SEQ ID NO:431	QMHICMYEG	8.74%
SEQ ID NO:432	KWEVMSAAY	8.73%
SEQ ID NO:433	QRMIVDALQ	8.50%
SEQ ID NO:434	TMGVSSYSM	8.50%
SEQ ID NO:435	TGAGIKEAH	8.49%
SEQ ID NO:436	VPWVEFRLQ	8.49%
SEQ ID NO:437	QPWPHLINK	8.48%
SEQ ID NO:438	SSLCRAYDN	8.22%
SEQ ID NO:439	CVKSRQDSD	8.21%
SEQ ID NO:440	NYILTQLGM	8.20%
SEQ ID NO:441	FGGVEDSWM	8.19%

SEQ ID NO:442	SVKDPMDWG	8.19%
SEQ ID NO:443	YSDRFQFQF	8.19%
SEQ ID NO:444	NCPMYDMIS	8.12%
SEQ ID NO:445	IWGIYVHYF	8.08%
SEQ ID NO:446	FTPELWENY	7.91%
SEQ ID NO:447	AHSRMDWTP	7.91%
SEQ ID NO:448	QPFQRKVEE	7.91%
SEQ ID NO:449	DYENCFESY	7.91%
SEQ ID NO:450	IEVYHFTNS	7.89%
SEQ ID NO:451	KALTHDARN	7.89%
SEQ ID NO:452	KTQMKRLRM	7.70%
SEQ ID NO:453	KAPWAPMDS	7.67%
SEQ ID NO:454	ETVMMPMSI	7.64%
SEQ ID NO:455	DYVDQCQVS	7.63%
SEQ ID NO:456	RACVGSPLW	7.62%
SEQ ID NO:457	KIMHSWYND	7.62%
SEQ ID NO:458	RHEPTQSVS	7.62%
SEQ ID NO:459	DPKGQSIQA	7.61%
SEQ ID NO:460	RTTSRDVFD	7.61%
SEQ ID NO:461	QCMVMTEAD	7.61%
SEQ ID NO:462	VQHNSESNI	7.61%
SEQ ID NO:463	WHESTTIGS	7.61%
SEQ ID NO:464	RHEQYVVAT	7.60%
SEQ ID NO:465	VVCTDTRAL	7.49%
SEQ ID NO:466	EHHAMMLHE	7.47%
SEQ ID NO:467	KACMSHQGR	7.42%
SEQ ID NO:468	DSPAYTTQY	7.35%
SEQ ID NO:469	ICELTHCNY	7.33%
SEQ ID NO:470	KATDLMVNG	7.32%

[1782]

SEQ ID NO:471	PQSHEKQLL	7.32%
SEQ ID NO:472	CHCHYNPFC	7.32%
SEQ ID NO:473	EAECVAGER	7.32%
SEQ ID NO:474	EFWTTHSKG	7.31%
SEQ ID NO:475	GHVLLGWYP	7.31%
SEQ ID NO:476	YCQQLFLCD	7.05%
SEQ ID NO:477	AKEPVRLYP	7.04%
SEQ ID NO:478	EERLFTGCH	7.04%
SEQ ID NO:479	QTSPKFCMC	7.04%
SEQ ID NO:480	WMNQCGCKD	7.04%
SEQ ID NO:481	NYSRGKENE	7.03%
SEQ ID NO:482	DIQRWEHEG	7.02%
SEQ ID NO:483	MMAMKQGFY	7.02%
SEQ ID NO:484	KVNACGNLT	6.89%
SEQ ID NO:485	ETTHRREWS	6.75%
SEQ ID NO:486	PTVEEVVVQ	6.75%
SEQ ID NO:487	SIEPRKTVG	6.75%
SEQ ID NO:488	QTIMRHAQT	6.74%
SEQ ID NO:489	RIETVSVPL	6.74%
SEQ ID NO:490	SRTICVHGN	6.74%
SEQ ID NO:491	GACKYSMCL	6.73%
SEQ ID NO:492	GWDQNCKDS	6.73%
SEQ ID NO:493	ASINCHLTE	6.44%
SEQ ID NO:494	HYMELLCGC	6.44%
SEQ ID NO:495	ITRTYQEMQ	6.44%
SEQ ID NO:496	RANENHPFT	6.44%
SEQ ID NO:497	LNKMQSYSA	6.42%
SEQ ID NO:498	NGYLVHACQ	6.21%
SEQ ID NO:499	LMAYCYANN	6.18%

SEQ ID NO:500	KESKEHRWA	6.16%
SEQ ID NO:501	FTTMCEHTA	6.15%
SEQ ID NO:502	ATMWVQAYG	6.15%
SEQ ID NO:503	RGQFHLIMD	6.15%
SEQ ID NO:504	DLVDLRSWM	6.14%
SEQ ID NO:505	LGAVLNNMV	6.14%
SEQ ID NO:506	MYRQKEMQT	6.14%
SEQ ID NO:507	RLEDTVGGS	6.14%
SEQ ID NO:508	ACHYRLDNC	6.06%
SEQ ID NO:509	LNNAGYRPC	6.02%
SEQ ID NO:510	SFSHTDLEA	5.89%
SEQ ID NO:511	ATEPLKFQY	5.87%
SEQ ID NO:512	NQNPYDMMG	5.87%
SEQ ID NO:513	FKIEWSQDI	5.86%
SEQ ID NO:514	QIGHQYKWD	5.86%
SEQ ID NO:515	SWSQLQHTD	5.86%
SEQ ID NO:516	QMMDFQHPA	5.86%
SEQ ID NO:517	LFTQRGWGN	5.85%
SEQ ID NO:518	TATQREAFV	5.59%
SEQ ID NO:519	KENPSEYEM	5.58%
SEQ ID NO:520	KEFQLQMEP	5.57%
SEQ ID NO:521	KSAGIRTLL	5.57%
SEQ ID NO:522	MSEKGYAWV	5.57%
SEQ ID NO:523	VYCGKDFQA	5.57%
SEQ ID NO:524	CVKLVMANC	5.56%
SEQ ID NO:525	ALRPYMDDR	5.56%
SEQ ID NO:526	CSIQINNCR	5.56%
SEQ ID NO:527	GTRINTCNG	5.56%
SEQ ID NO:528	FGAVELDAP	5.56%

[1783]

SEQ ID NO:529	YCQRGHLSC	5.56%
SEQ ID NO:530	HRMNNHAWF	5.54%
SEQ ID NO:531	NADEKYNNT	5.52%
SEQ ID NO:532	MGATKSDDF	5.32%
SEQ ID NO:533	SMHDYNFNV	5.31%
SEQ ID NO:534	GNECFHNVE	5.29%
SEQ ID NO:535	GVWMPHGWR	5.29%
SEQ ID NO:536	EIWPAAWMD	5.28%
SEQ ID NO:537	APKWQIVDP	5.27%
SEQ ID NO:538	HAMHWWSPM	5.27%
SEQ ID NO:539	QCKYSERGT	5.21%
SEQ ID NO:540	TCCMWTFIW	5.01%
SEQ ID NO:541	ETGYRENPC	5.00%
SEQ ID NO:542	MEMFTMKSC	4.99%
SEQ ID NO:543	PQGFCKPQC	4.98%
SEQ ID NO:544	KTTFDVACY	4.98%
SEQ ID NO:545	WCRTPQREN	4.98%
SEQ ID NO:546	GSQQMQASQ	4.97%
SEQ ID NO:547	ITGSPQNGC	4.97%
SEQ ID NO:548	GTVEAPQSF	4.93%
SEQ ID NO:549	LNCPSKDET	4.85%
SEQ ID NO:550	CSNTETHGV	4.84%
SEQ ID NO:551	QMEQWHRGG	4.78%
SEQ ID NO:552	IPFNSMIYC	4.72%
SEQ ID NO:553	RVEQWRGQQ	4.70%
SEQ ID NO:554	WNMPWGFYH	4.70%
SEQ ID NO:555	MISCMDSWN	4.69%
SEQ ID NO:556	TSRALDVAA	4.69%
SEQ ID NO:557	FFRWRWGLG	4.69%
SEQ ID NO:558	FPKCAIWHY	4.69%
SEQ ID NO:559	NMIKRCSQG	4.69%
SEQ ID NO:560	GTSVNTLMM	4.68%
SEQ ID NO:561	MRQRTYICQ	4.68%
SEQ ID NO:562	RGCMAECMP	4.68%
SEQ ID NO:563	MPMPVYQRD	4.59%
SEQ ID NO:564	ALNKTLAES	4.55%
SEQ ID NO:565	EAGMWPSKA	4.41%
SEQ ID NO:566	AAVKEQQVM	4.40%
SEQ ID NO:567	SAQYSGFDT	4.40%
SEQ ID NO:568	LILDFNKID	4.40%
SEQ ID NO:569	MSAVSLGRS	4.39%
SEQ ID NO:570	RWGSPMAVP	4.39%
SEQ ID NO:571	KTVTNGDAG	4.39%
SEQ ID NO:572	YTMSMQMQR	4.39%
SEQ ID NO:573	AHEHYRIFP	4.11%
SEQ ID NO:574	GKFPMPTHS	4.11%
SEQ ID NO:575	KIENRGQCR	4.10%
SEQ ID NO:576	MWEMRRAMG	4.10%
SEQ ID NO:577	HGSSYNTVQ	4.09%
SEQ ID NO:578	VIKAFQHFN	4.09%
SEQ ID NO:579	EVRCNKCRS	3.88%
SEQ ID NO:580	LALQFQYNS	3.85%
SEQ ID NO:581	GVITVDAKG	3.84%
SEQ ID NO:582	CLHHHYPVT	3.83%
SEQ ID NO:583	AFCRGHYQA	3.82%
SEQ ID NO:584	KTGWDSHWS	3.82%
SEQ ID NO:585	RYEMMDTSD	3.82%
SEQ ID NO:586	DIFWSSMDF	3.81%

[1784]

SEQ ID NO:587	MSTHVLKWV	3.81%
SEQ ID NO:588	CDVWWFPNL	3.81%
SEQ ID NO:589	IVFHNKFQH	3.81%
SEQ ID NO:590	SHIQSASNL	3.81%
SEQ ID NO:591	VKQLRQDGF	3.81%
SEQ ID NO:592	DKCLNNKFY	3.80%
SEQ ID NO:593	RNDWQQVFS	3.80%
SEQ ID NO:594	QRREQCTST	3.77%
SEQ ID NO:595	YHWSYEGAT	3.75%
SEQ ID NO:596	TDFDKRQVP	3.67%
SEQ ID NO:597	QISMKGEHG	3.55%
SEQ ID NO:598	LVGERKNAY	3.53%
SEQ ID NO:599	DPMKIFVSH	3.53%
SEQ ID NO:600	KQTRTDYAY	3.52%
SEQ ID NO:601	AVAMIRPNQ	3.51%
SEQ ID NO:602	ESTHMLYEM	3.51%
SEQ ID NO:603	VDADYCSCS	3.51%
SEQ ID NO:604	EMVRMQRGT	3.51%
SEQ ID NO:605	NCDHRFGTL	3.50%
SEQ ID NO:606	SVVFSAAEV	3.48%
SEQ ID NO:607	NARPMCGQV	3.40%
SEQ ID NO:608	TEATEMTYY	3.33%
SEQ ID NO:609	AYGSFEKDW	3.24%
SEQ ID NO:610	MNSFYRAEW	3.24%
SEQ ID NO:611	MIDNCKICL	3.23%
SEQ ID NO:612	VCRGSPLAE	3.23%
SEQ ID NO:613	DTRMDQCCY	3.22%
SEQ ID NO:614	LVDCSNTCQ	3.22%
SEQ ID NO:615	INKNMDAAY	3.22%
SEQ ID NO:616	TGMTFNSFR	3.22%
SEQ ID NO:617	AWKIAHNSF	3.22%
SEQ ID NO:618	NFMQMSIMG	3.22%
SEQ ID NO:619	WIKHCWPPH	3.22%
SEQ ID NO:620	NLWNIAMQQ	3.18%
SEQ ID NO:621	EHMDRADDA	3.11%
SEQ ID NO:622	AQDVYYPGS	3.08%
SEQ ID NO:623	FPQTMHAQA	2.96%
SEQ ID NO:624	TFGCIESMQ	2.95%
SEQ ID NO:625	KFFYTELMH	2.94%
SEQ ID NO:626	GRLGPGTWK	2.93%
SEQ ID NO:627	ASTDGTWKD	2.92%
SEQ ID NO:628	DARPFLAWM	2.92%
SEQ ID NO:629	KFHSKERMN	2.91%
SEQ ID NO:630	SLQHFPECA	2.87%
SEQ ID NO:631	VCVSAHHHF	2.87%
SEQ ID NO:632	AVIFELDAT	2.84%
SEQ ID NO:633	EQADQMQVD	2.83%
SEQ ID NO:634	MSQFKAYLF	2.82%
SEQ ID NO:635	EMSIQSSLG	2.76%
SEQ ID NO:636	GTGYRTFIG	2.67%
SEQ ID NO:637	KASHTVTGT	2.65%
SEQ ID NO:638	NSVALENRG	2.64%
SEQ ID NO:639	NEGPLMCWF	2.64%
SEQ ID NO:640	NNELQSRPK	2.63%
SEQ ID NO:641	NGTQCLNMD	2.52%
SEQ ID NO:642	SSTWQNVDK	2.51%
SEQ ID NO:643	YMVPYSNYP	2.50%
SEQ ID NO:644	WFDRPMCMF	2.48%

[1785]

SEQ ID NO:645	LECTALVAM	2.38%
SEQ ID NO:646	QVQSGQFIF	2.37%
SEQ ID NO:647	KMETMICFD	2.37%
SEQ ID NO:648	QSDAVHEPA	2.36%
SEQ ID NO:649	GYMCEDAGI	2.35%
SEQ ID NO:650	EIHPGDLND	2.35%
SEQ ID NO:651	EFSKACTNE	2.35%
SEQ ID NO:652	KSTYYNLKW	2.34%
SEQ ID NO:653	TSSRTNPWG	2.34%
SEQ ID NO:654	SFKGFHLQN	2.28%
SEQ ID NO:655	CDEWDNLTH	2.20%
SEQ ID NO:656	DAGIGMKFY	2.16%
SEQ ID NO:657	LGDQAGYHV	2.14%
SEQ ID NO:658	IIVWMMMYR	2.13%
SEQ ID NO:659	GADEWLLTM	2.10%
SEQ ID NO:660	QSYPTMYQC	2.09%
SEQ ID NO:661	DRTWTMTQM	2.08%
SEQ ID NO:662	TSWTTEHRM	2.06%
SEQ ID NO:663	TLRICELHG	2.06%
SEQ ID NO:664	CQASHEHAH	2.06%
SEQ ID NO:665	ESAEPRMIP	2.06%
SEQ ID NO:666	FWMQLWCLP	2.03%
SEQ ID NO:667	QGYCRNCTG	2.02%
SEQ ID NO:668	ESCLVRRGW	2.00%
SEQ ID NO:669	QIRCTCGWP	1.98%
SEQ ID NO:670	HAYQHDHSS	1.97%
SEQ ID NO:671	DMAVSEAHC	1.97%
SEQ ID NO:672	WAGMAYHAQ	1.94%
SEQ ID NO:673	FQQDIQIGQ	1.92%

SEQ ID NO:674	VVHKGRWKE	1.91%
SEQ ID NO:675	MQLRQSDKW	1.89%
SEQ ID NO:676	YVRQFPIEL	1.87%
SEQ ID NO:677	IIHRTEGAA	1.87%
SEQ ID NO:678	GQYWGQGWL	1.85%
SEQ ID NO:679	GSTHMNQEG	1.85%
SEQ ID NO:680	INAGDKPNT	1.83%
SEQ ID NO:681	LGTAMREMQ	1.81%
SEQ ID NO:682	PGGGDRPAC	1.81%
SEQ ID NO:683	QDVQMWYQI	1.80%
SEQ ID NO:684	QARAVMTAD	1.79%
SEQ ID NO:685	GAHDPKVWY	1.78%
SEQ ID NO:686	GTLGYHITG	1.77%
SEQ ID NO:687	SMLMMITGM	1.77%
SEQ ID NO:688	RDDYGIYDN	1.77%
SEQ ID NO:689	HLVTTRHSH	1.76%
SEQ ID NO:690	ITDLRCWLK	1.76%
SEQ ID NO:691	IMVRELMSS	1.75%
SEQ ID NO:692	NAEHCFMGQ	1.73%
SEQ ID NO:693	NNLQWMYAK	1.73%
SEQ ID NO:694	SPMPFLNVL	1.72%
SEQ ID NO:695	CEQQLIAEG	1.72%
SEQ ID NO:696	NDKYAFADR	1.71%
SEQ ID NO:697	WIKRHIPSG	1.71%
SEQ ID NO:698	REGAKHCVN	1.69%
SEQ ID NO:699	CDDWTFERQ	1.69%
SEQ ID NO:700	GVGPSHGER	1.69%
SEQ ID NO:701	ESTFWQHTF	1.69%
SEQ ID NO:702	MFVCCTIIV	1.67%

[1786]

SEQ ID NO:703	QKYEPRFFG	1.67%
SEQ ID NO:704	AHVKIQLSW	1.66%
SEQ ID NO:705	TFSHMLSVE	1.66%
SEQ ID NO:706	LRPLQHIQL	1.65%
SEQ ID NO:707	MANGNWGVW	1.65%
SEQ ID NO:708	VTDAAMFMG	1.64%
SEQ ID NO:709	WHEEKPANS	1.62%
SEQ ID NO:710	KCQRMSVQQ	1.62%
SEQ ID NO:711	DDRIGHCRN	1.62%
SEQ ID NO:712	IAPEARCGT	1.62%
SEQ ID NO:713	GCSSSVNMR	1.61%
SEQ ID NO:714	RKRVRAYSE	1.61%
SEQ ID NO:715	QEPKEDHIV	1.60%
SEQ ID NO:716	SILPYPIDV	1.59%
SEQ ID NO:717	ILDIINTET	1.59%
SEQ ID NO:718	RFMLESWLH	1.59%
SEQ ID NO:719	NSAHVHYFH	1.58%
SEQ ID NO:720	LVKGHMTTS	1.58%
SEQ ID NO:721	DKYANVHNE	1.57%
SEQ ID NO:722	RYCTPDHEV	1.57%
SEQ ID NO:723	RYHRRHDKP	1.57%
SEQ ID NO:724	SMRMFEHSC	1.56%
SEQ ID NO:725	SRCNKPTVR	1.56%
SEQ ID NO:726	TFHQSDYWE	1.55%
SEQ ID NO:727	VLQMPWLYT	1.55%
SEQ ID NO:728	VCTPPLNMQ	1.55%
SEQ ID NO:729	LNSDRLIEY	1.52%
SEQ ID NO:730	CETYVPVIY	1.52%
SEQ ID NO:731	WVPTFHARY	1.52%
SEQ ID NO:732	CEIRTYEWS	1.52%
SEQ ID NO:733	KHDGFSVVG	1.51%
SEQ ID NO:734	AWYHTQDDM	1.50%
SEQ ID NO:735	WMCEFSCIQ	1.50%
SEQ ID NO:736	AMVMSKTES	1.49%
SEQ ID NO:737	NFCCLECCI	1.49%
SEQ ID NO:738	RTSPALLCA	1.49%
SEQ ID NO:739	DITLWMATE	1.48%
SEQ ID NO:740	AVAEIRPEP	1.47%
SEQ ID NO:741	SARVTTMLY	1.47%
SEQ ID NO:742	TLRAEMSFH	1.47%
SEQ ID NO:743	ASKQNGINC	1.47%
SEQ ID NO:744	RRCPALTIR	1.47%
SEQ ID NO:745	ARPLVCQAL	1.46%
SEQ ID NO:746	MMLGYMGQD	1.46%
SEQ ID NO:747	QEHIYKHSN	1.46%
SEQ ID NO:748	DQNTFPETG	1.46%
SEQ ID NO:749	KSEGWLLGD	1.45%
SEQ ID NO:750	SPCFDVFEE	1.44%
SEQ ID NO:751	TINKFPWVC	1.44%
SEQ ID NO:752	IRFTTQIVD	1.44%
SEQ ID NO:753	SVWEKTQMR	1.43%
SEQ ID NO:754	CPCHKWASC	1.43%
SEQ ID NO:755	HDHAAKQLD	1.43%
SEQ ID NO:756	RPVGKLYMI	1.43%
SEQ ID NO:757	MWGCKLFVC	1.42%
SEQ ID NO:758	KHNPSRHDI	1.42%
SEQ ID NO:759	NPYQCIVAG	1.39%
SEQ ID NO:760	CMRGGQKLT	1.36%

[1787]

SEQ ID NO:761	DVDTSQFDR	1.34%
SEQ ID NO:762	VCVKWRNVN	1.34%
SEQ ID NO:763	WASPSVWRR	1.34%
SEQ ID NO:764	YFHYRYDYG	1.33%
SEQ ID NO:765	KVYYIISHQ	1.32%
SEQ ID NO:766	QTYRGDWQK	1.31%
SEQ ID NO:767	RVDKAGAIF	1.28%
SEQ ID NO:768	VFNLMLQDK	1.27%
SEQ ID NO:769	VKKPQYDMH	1.26%
SEQ ID NO:770	QYVQCAAKD	1.23%
SEQ ID NO:771	EGCYFWKQV	1.22%
SEQ ID NO:772	TMKTWIYRN	1.22%
SEQ ID NO:773	GWESFPHCG	1.21%
SEQ ID NO:774	QKAEHLFVG	1.21%
SEQ ID NO:775	PDSIPNSWC	1.21%
SEQ ID NO:776	CHGRMLCYI	1.20%
SEQ ID NO:777	NQQRTYNSG	1.19%
SEQ ID NO:778	LDTKTLANS	1.19%
SEQ ID NO:779	FAPEFDNHC	1.19%
SEQ ID NO:780	REFAISGGS	1.19%
SEQ ID NO:781	AFKTFGHVN	1.18%
SEQ ID NO:782	ASSRLAEVI	1.18%
SEQ ID NO:783	RMSAEKERS	1.17%
SEQ ID NO:784	TMGLGVNAI	1.17%
SEQ ID NO:785	PADWDREHG	1.17%
SEQ ID NO:786	PELRPCAPR	1.17%
SEQ ID NO:787	PDTQSAFWI	1.17%
SEQ ID NO:788	KFKWLVRMT	1.16%
SEQ ID NO:789	MTNRCAKYK	1.15%
SEQ ID NO:790	AQPRTCQII	1.13%
SEQ ID NO:791	IVVFGQMGY	1.13%
SEQ ID NO:792	KVVEMFRWY	1.12%
SEQ ID NO:793	CQTHWEKGD	1.12%
SEQ ID NO:794	SGTEHVALH	1.12%
SEQ ID NO:795	LCDLIMTDE	1.11%
SEQ ID NO:796	PTLDRWEYH	1.10%
SEQ ID NO:797	QLPIYIQKS	1.08%
SEQ ID NO:798	HLAWIVNED	1.00%
SEQ ID NO:799	HYFKWLFVE	0.98%
SEQ ID NO:800	KFTWKKYFA	0.97%
SEQ ID NO:801	HTWPPLFYW	0.94%
SEQ ID NO:802	TWMPKVEYK	0.91%
SEQ ID NO:803	ALTAVHAGT	0.90%
SEQ ID NO:804	TLNSWHPEC	0.90%
SEQ ID NO:805	AHMSYCLSR	0.89%
SEQ ID NO:806	TDYPWVLHG	0.89%
SEQ ID NO:807	WFSIIMSKL	0.89%
SEQ ID NO:808	SVKVKNGFT	0.89%
SEQ ID NO:809	NHIFKKVTW	0.88%
SEQ ID NO:810	KPPACIGGK	0.88%
SEQ ID NO:811	PPDARQSLR	0.88%
SEQ ID NO:812	RQTGTGNME	0.88%
SEQ ID NO:813	EQFADQAPR	0.88%
SEQ ID NO:814	ISISAQDID	0.88%
SEQ ID NO:815	KFTLQQMAS	0.88%
SEQ ID NO:816	SFSGTHLGW	0.88%
SEQ ID NO:817	NHSDTDTAQ	0.86%
SEQ ID NO:818	AFCYSYYQN	0.80%

[1788]

SEQ ID NO:819	EGKITKAAV	0.80%
SEQ ID NO:820	VQCLWWHKE	0.77%
SEQ ID NO:821	QNLMYCSSV	0.72%
SEQ ID NO:822	RRWLNPAYN	0.69%
SEQ ID NO:823	PWFYKDWPH	0.66%
SEQ ID NO:824	QALWLESDN	0.62%
SEQ ID NO:825	TLNLRHPEC	0.62%
SEQ ID NO:826	APEFRNRLT	0.60%
SEQ ID NO:827	DVMRMGHYL	0.60%
SEQ ID NO:828	ERRAHNHYA	0.60%
SEQ ID NO:829	TSHKMMVLS	0.59%
SEQ ID NO:830	EVDRNTNWC	0.59%
SEQ ID NO:831	MALPLRVLS	0.59%
SEQ ID NO:832	PCKSTRSIL	0.59%
SEQ ID NO:833	TTDGMMGDH	0.59%
SEQ ID NO:834	YPVHLMPPL	0.59%
SEQ ID NO:835	ECWMAVQSC	0.59%
SEQ ID NO:836	ERTACKTKS	0.59%
SEQ ID NO:837	GTARHNMQL	0.59%
SEQ ID NO:838	GTVDAQDRE	0.59%
SEQ ID NO:839	RQIGTGNTE	0.59%
SEQ ID NO:840	SPDGRGLCG	0.59%
SEQ ID NO:841	TLNSRHLEC	0.59%
SEQ ID NO:842	VLTDGACAA	0.59%
SEQ ID NO:843	AHMYYCLSR	0.58%
SEQ ID NO:844	AMLMSDGGG	0.58%
SEQ ID NO:845	AQGFKPSGG	0.58%
SEQ ID NO:846	ATGTYNLEE	0.58%
SEQ ID NO:847	AYAASKNAM	0.58%

SEQ ID NO:848	CAWYMKQGN	0.58%
SEQ ID NO:849	CCRQIWIQY	0.58%
SEQ ID NO:850	CMKTFPMNA	0.58%
SEQ ID NO:851	DRINLMWKD	0.58%
SEQ ID NO:852	DRVDLMWKD	0.58%
SEQ ID NO:853	DSATFESAD	0.58%
SEQ ID NO:854	EFEMRDSGY	0.58%
SEQ ID NO:855	EKYQIHRDR	0.58%
SEQ ID NO:856	HNQRTYPAL	0.58%
SEQ ID NO:857	KMTNSSENS	0.58%
SEQ ID NO:858	KQPTFPMLA	0.58%
SEQ ID NO:859	LMRWYNRCW	0.58%
SEQ ID NO:860	LRCTVMDPG	0.58%
SEQ ID NO:861	LYQKDQRFG	0.58%
SEQ ID NO:862	MHVNRRIDQ	0.58%
SEQ ID NO:863	NCKSDQIAM	0.58%
SEQ ID NO:864	NNHRLNDAN	0.58%
SEQ ID NO:865	QPKVHGNYL	0.58%
SEQ ID NO:866	QVEDRYWSY	0.58%
SEQ ID NO:867	RRGYVNIWE	0.58%
SEQ ID NO:868	SPCPTWDGL	0.58%
SEQ ID NO:869	SRLLVHECP	0.58%
SEQ ID NO:870	SSMQCVMRP	0.58%
SEQ ID NO:871	SVTYHEWHL	0.58%
SEQ ID NO:872	TCICSHTYP	0.58%
SEQ ID NO:873	TGMAVSEQS	0.58%
SEQ ID NO:874	TIEHGAWKC	0.58%
SEQ ID NO:875	TPLTKWLPI	0.58%
SEQ ID NO:876	VSIPNLGWP	0.58%

[1789]

SEQ ID NO:877	YCHLEDRIT	0.58%
SEQ ID NO:878	EWMVIMAKN	0.40%
SEQ ID NO:879	RVAIMFTGT	0.35%
SEQ ID NO:880	IPDIMRRCP	0.33%
SEQ ID NO:881	ERTACKMRS	0.33%
SEQ ID NO:882	YTTVEGRIT	0.33%
SEQ ID NO:883	DLEFCVMER	0.32%
SEQ ID NO:884	YAQRTCWSE	0.31%
SEQ ID NO:885	DRVNLTWKD	0.31%
SEQ ID NO:886	DSTPNLLDP	0.31%
SEQ ID NO:887	EPLYSWVGV	0.31%
SEQ ID NO:888	KQLFMKFWS	0.31%
SEQ ID NO:889	FVGLMSQCG	0.31%
SEQ ID NO:890	KPCSSMSIQ	0.31%
SEQ ID NO:891	SIYQKWVTS	0.31%
SEQ ID NO:892	TLSSRHPEC	0.31%
SEQ ID NO:893	AQTNNTCSI	0.31%
SEQ ID NO:894	IPSHVHLGE	0.31%
SEQ ID NO:895	MALPPHVLS	0.31%
SEQ ID NO:896	MGCIGNCIL	0.31%
SEQ ID NO:897	QTQHYGSTT	0.31%
SEQ ID NO:898	RVDRNTNWC	0.31%
SEQ ID NO:899	TYDTRSHCL	0.31%
SEQ ID NO:900	ARVDHQCSG	0.30%
SEQ ID NO:901	NGSHYNAMR	0.30%
SEQ ID NO:902	QTQLYQNDE	0.30%
SEQ ID NO:903	SPYGRGLCG	0.30%
SEQ ID NO:904	TYSLAMSLG	0.30%
SEQ ID NO:905	WVGNANRPA	0.30%

SEQ ID NO:906	AHVDHQCGG	0.30%
SEQ ID NO:907	APEFNVARR	0.30%
SEQ ID NO:908	CCQQIWVQY	0.30%
SEQ ID NO:909	CITARSTAA	0.30%
SEQ ID NO:910	CTIARSTAA	0.30%
SEQ ID NO:911	DVIRTTHSG	0.30%
SEQ ID NO:912	GSQQTTPSC	0.30%
SEQ ID NO:913	GYATFPDSD	0.30%
SEQ ID NO:914	ISSTMVEHG	0.30%
SEQ ID NO:915	KLLFMKFWN	0.30%
SEQ ID NO:916	KVDRNTYWC	0.30%
SEQ ID NO:917	LFRAYAWSQ	0.30%
SEQ ID NO:918	LITRDTGLA	0.30%
SEQ ID NO:919	LSECSDRTN	0.30%
SEQ ID NO:920	NCKPNQIAM	0.30%
SEQ ID NO:921	PAVPADEFF	0.30%
SEQ ID NO:922	QATDMHRWQ	0.30%
SEQ ID NO:923	STARHNTQL	0.30%
SEQ ID NO:924	THMNYCLSR	0.30%
SEQ ID NO:925	TSHRMMVFS	0.30%
SEQ ID NO:926	AMTRCMAGH	0.30%
SEQ ID NO:927	CPEHNLVGV	0.30%
SEQ ID NO:928	CPEHSQVGV	0.30%
SEQ ID NO:929	DRANLMWKD	0.30%
SEQ ID NO:930	DRVNLMWKG	0.30%
SEQ ID NO:931	DWGYCHSKA	0.30%
SEQ ID NO:932	EFVLGYQGI	0.30%
SEQ ID NO:933	EKYQIHWDH	0.30%
SEQ ID NO:934	EQFVDRAPR	0.30%

[1790]

SEQ ID NO:935	ERQAHSHYA	0.30%
SEQ ID NO:936	FIYQKWVTN	0.30%
SEQ ID NO:937	GQFRSGAHR	0.30%
SEQ ID NO:938	HTNEQYLQH	0.30%
SEQ ID NO:939	HVNTSCKQS	0.30%
SEQ ID NO:940	KYSPIQEHN	0.30%
SEQ ID NO:941	LCLQGAKMD	0.30%
SEQ ID NO:942	LMSITSAVE	0.30%
SEQ ID NO:943	LTHNIPFPK	0.30%
SEQ ID NO:944	MALPLHALS	0.30%
SEQ ID NO:945	MCQSVEKEY	0.30%
SEQ ID NO:946	NGNRYNAMR	0.30%
SEQ ID NO:947	NHRSREIDI	0.30%
SEQ ID NO:948	NIVRSRSIT	0.30%
SEQ ID NO:949	NKTCYLTCP	0.30%
SEQ ID NO:950	PQARVDMEF	0.30%
SEQ ID NO:951	QAAGMHRWQ	0.30%
SEQ ID NO:952	REFVISGGN	0.30%
SEQ ID NO:953	SMREYQKHI	0.30%
SEQ ID NO:954	SPHSCQHMG	0.30%
SEQ ID NO:955	SRLSVHECP	0.30%
SEQ ID NO:956	STTRHNMQL	0.30%
SEQ ID NO:957	TFHFSILLT	0.30%
SEQ ID NO:958	TPQPMNVAA	0.30%
SEQ ID NO:959	TYIQHPTNA	0.30%
SEQ ID NO:960	VAGRFGSNT	0.30%
SEQ ID NO:961	VSMPSLGWP	0.30%
SEQ ID NO:962	YDNHQKARI	0.30%
SEQ ID NO:963	AHMIYCLSR	0.29%
SEQ ID NO:964	AHMNYCLSH	0.29%
SEQ ID NO:965	AQWIQQKHG	0.29%
SEQ ID NO:966	ARTLREGWG	0.29%
SEQ ID NO:967	ASKLEVQEQ	0.29%
SEQ ID NO:968	AVEHGAWKC	0.29%
SEQ ID NO:969	AVETGESAV	0.29%
SEQ ID NO:970	CCRQIWVQH	0.29%
SEQ ID NO:971	CGGQETGLA	0.29%
SEQ ID NO:972	CNSWCTGSL	0.29%
SEQ ID NO:973	DCRHTTLYV	0.29%
SEQ ID NO:974	DESTCVAHG	0.29%
SEQ ID NO:975	DGMHRELHS	0.29%
SEQ ID NO:976	DKYGLGIFH	0.29%
SEQ ID NO:977	DRVYLMWKD	0.29%
SEQ ID NO:978	DTDYKPHFA	0.29%
SEQ ID NO:979	DVHDGQLCY	0.29%
SEQ ID NO:980	DVYDSQLCY	0.29%
SEQ ID NO:981	DYAPTNTMM	0.29%
SEQ ID NO:982	DYSQVNCER	0.29%
SEQ ID NO:983	DYSQVNREK	0.29%
SEQ ID NO:984	EEYYTWECH	0.29%
SEQ ID NO:985	ETHHEYRDD	0.29%
SEQ ID NO:986	EYACMLMTR	0.29%
SEQ ID NO:987	FLRPKRHFS	0.29%
SEQ ID NO:988	GCQLGFQIT	0.29%
SEQ ID NO:989	GQFRRGAYR	0.29%
SEQ ID NO:990	GSLSKDTPL	0.29%
SEQ ID NO:991	GSQQTTLSR	0.29%
SEQ ID NO:992	HCHQEDRIT	0.29%

[1791]

SEQ ID NO:993	HHECLYQIY	0.29%
SEQ ID NO:994	HVDQHRIQA	0.29%
SEQ ID NO:995	IADIGRKQW	0.29%
SEQ ID NO:996	IAWPTDCSF	0.29%
SEQ ID NO:997	IYCPCQDSD	0.29%
SEQ ID NO:998	KAALFLKDC	0.29%
SEQ ID NO:999	KAGMDQTAV	0.29%
SEQ ID NO:1000	KCRMAVQSC	0.29%
SEQ ID NO:1001	KPLASIGQH	0.29%
SEQ ID NO:1002	KQLFMEFWN	0.29%
SEQ ID NO:1003	KSLQEQGGA	0.29%
SEQ ID NO:1004	KTQLSDHKH	0.29%
SEQ ID NO:1005	LDHTNLMKL	0.29%
SEQ ID NO:1006	LIQWYNRCW	0.29%
SEQ ID NO:1007	LMKTSPCYW	0.29%
SEQ ID NO:1008	LMLGYMEQD	0.29%
SEQ ID NO:1009	LSECSDQAN	0.29%
SEQ ID NO:1010	LVVIDPAHY	0.29%
SEQ ID NO:1011	MALPLYVLS	0.29%
SEQ ID NO:1012	MCQMSRRSI	0.29%
SEQ ID NO:1013	MMLGYMERD	0.29%
SEQ ID NO:1014	MTQVHQALV	0.29%
SEQ ID NO:1015	NACPQHDFG	0.29%
SEQ ID NO:1016	NALWMTKAP	0.29%
SEQ ID NO:1017	NCRHTILYV	0.29%
SEQ ID NO:1018	NDCTCRAIL	0.29%
SEQ ID NO:1019	NFANHQPLP	0.29%
SEQ ID NO:1020	NQRGLCKDE	0.29%
SEQ ID NO:1021	PACLLRRML	0.29%

SEQ ID NO:1022	PANPSFSYL	0.29%
SEQ ID NO:1023	PCQQLMCFT	0.29%
SEQ ID NO:1024	PDARQSLRH	0.29%
SEQ ID NO:1025	PIYQKWVTN	0.29%
SEQ ID NO:1026	PNGCRPFCC	0.29%
SEQ ID NO:1027	PQHAGRAEH	0.29%
SEQ ID NO:1028	QCFLWNDYQ	0.29%
SEQ ID NO:1029	QHLMTLSHE	0.29%
SEQ ID NO:1030	QPLSRNGAR	0.29%
SEQ ID NO:1031	QPQLRNGAR	0.29%
SEQ ID NO:1032	QRQSQEWFH	0.29%
SEQ ID NO:1033	QTGVMAAIG	0.29%
SEQ ID NO:1034	QVVIIECKN	0.29%
SEQ ID NO:1035	RDRRVRDEC	0.29%
SEQ ID NO:1036	RLETTRYIW	0.29%
SEQ ID NO:1037	RMARVSVHG	0.29%
SEQ ID NO:1038	RQHHCLPVI	0.29%
SEQ ID NO:1039	RQVGTGNME	0.29%
SEQ ID NO:1040	RSTRFTGIA	0.29%
SEQ ID NO:1041	RTNEQYPQH	0.29%
SEQ ID NO:1042	RTRISQMSG	0.29%
SEQ ID NO:1043	RTRTSQMFG	0.29%
SEQ ID NO:1044	RVNGSTCGH	0.29%
SEQ ID NO:1045	SATTSGTAF	0.29%
SEQ ID NO:1046	SATYHEWHP	0.29%
SEQ ID NO:1047	SEAKYDTAT	0.29%
SEQ ID NO:1048	SKRWYGHSA	0.29%
SEQ ID NO:1049	SMQEYQKHV	0.29%
SEQ ID NO:1050	SRCTVVDPG	0.29%

[1792]

SEQ ID NO:1051	SSWMQMQGP	0.29%
SEQ ID NO:1052	TAMVSECID	0.29%
SEQ ID NO:1053	TEFHMMIQT	0.29%
SEQ ID NO:1054	TGRWLVYQA	0.29%
SEQ ID NO:1055	TLNSRHPEY	0.29%
SEQ ID NO:1056	TSEPHCSFA	0.29%
SEQ ID NO:1057	TVDGMMGDH	0.29%
SEQ ID NO:1058	TVPKWDNIS	0.29%
SEQ ID NO:1059	TYHCYMSWM	0.29%
SEQ ID NO:1060	TYIQCPTNA	0.29%
SEQ ID NO:1061	TYIQYPANA	0.29%
SEQ ID NO:1062	VHGQTDLSL	0.29%
SEQ ID NO:1063	VPVPFLSLR	0.29%
SEQ ID NO:1064	VSFQRMNNM	0.29%
SEQ ID NO:1065	VSTPNLGWP	0.29%
SEQ ID NO:1066	VSVQGRAQL	0.29%
SEQ ID NO:1067	VTEVLNQDV	0.29%
SEQ ID NO:1068	YATHMQHLE	0.29%
SEQ ID NO:1069	YCHQEDRII	0.29%
SEQ ID NO:1070	YETDSEGYD	0.29%
SEQ ID NO:1071	YPAHMMPPL	0.29%
SEQ ID NO:1072	YWCPSQRWR	0.29%
SEQ ID NO:1073	QLKYAVGAN	0.23%
SEQ ID NO:1074	MNHQCMLWD	0.07%
SEQ ID NO:1075	RHQRVMPTY	0.05%
SEQ ID NO:1076	NSGIATADH	0.05%
SEQ ID NO:1077	ACHTGQSGM	0.04%
SEQ ID NO:1078	SHWFFHFDL	0.03%
SEQ ID NO:1079	DVTQTRYCY	0.03%
SEQ ID NO:1080	FSEQNVSAA	0.03%
SEQ ID NO:1081	GPLYSWVGI	0.03%
SEQ ID NO:1082	NAGMKGVDV	0.03%
SEQ ID NO:1083	CTEAGMTFG	0.02%
SEQ ID NO:1084	PMNDMQFVP	0.02%
SEQ ID NO:1085	SGRRSGGRP	0.02%
SEQ ID NO:1086	DVAQYGLGS	0.02%
SEQ ID NO:1087	NPTDLLQHE	0.02%
SEQ ID NO:1088	QAQCIAMQR	0.02%
SEQ ID NO:1089	DGMHRELCS	0.02%
SEQ ID NO:1090	DGTHRELYS	0.02%
SEQ ID NO:1091	KWCQFGMVA	0.02%
SEQ ID NO:1092	LTHNTPFPE	0.02%
SEQ ID NO:1093	PMPTADVLH	0.02%
SEQ ID NO:1094	QGEEVWWEA	0.02%
SEQ ID NO:1095	RMVTMERDR	0.02%
SEQ ID NO:1096	RVNGSTCWH	0.02%
SEQ ID NO:1097	RWYQFGMVA	0.02%
SEQ ID NO:1098	TFLEYQALG	0.02%
SEQ ID NO:1099	TTGVCSFAQ	0.02%
SEQ ID NO:1100	YHKQMAVWW	0.02%
SEQ ID NO:1101	AVSVACLNT	0.01%
SEQ ID NO:1102	CDSPSLLQS	0.01%
SEQ ID NO:1103	CMYLPALQY	0.01%
SEQ ID NO:1104	CSSWANGCD	0.01%
SEQ ID NO:1105	EDCPKVFTI	0.01%
SEQ ID NO:1106	EFALGCQGI	0.01%
SEQ ID NO:1107	ERGEMRATG	0.01%
SEQ ID NO:1108	ERGGMKATG	0.01%

[1793]

SEQ ID NO:1109	GRVHGNVCI	0.01%
SEQ ID NO:1110	HVWSLEWHL	0.01%
SEQ ID NO:1111	IPGHVHFGE	0.01%
SEQ ID NO:1112	IWGVYVHYF	0.01%
SEQ ID NO:1113	NHDEKLWAP	0.01%
SEQ ID NO:1114	NVWSLAVRD	0.01%

[1794] 分析CNS(前脑皮层、枕叶皮层、颞叶皮层、丘脑、下丘脑、黑质、海马DG、海马CA1、海马CA3、小脑)、肝脏、骨骼肌、心脏、肺、脾脏、淋巴结、骨髓、乳腺、皮肤、肾上腺、甲状腺、结肠、坐骨神经和脊髓组织,以鉴定AAV5 VP1 581-589区域中驱动CNS嗜性的氨基酸残基。结果显示于图18B中,该图显示了相较于所有其他分析组织在CNS组织中AAV5 VP1的581至589区域中各氨基酸的位置频率,其中衣壳多肽具有SEQ ID NO:2的序列。位置("Xaan",n=1-9)在x轴上表示并且氨基酸残基在y轴上表示。特定位置的特定氨基酸残基的频率在热图中作为强度表示。较深的方块表示特定氨基酸在特定位置的频率较高,较浅的方块表示特定氨基酸在特定位置的频率较低。

[1795] 分析相对于给定氨基酸残基出现在所有其他所收获组织中鉴定的变体中的指定位置的频率而言该给定氨基酸残基出现在CNS中鉴定的变体中对应于SEQ ID NO:1(推及于SEQ ID NO:2)的位置581至位置589的所述指定位置的频率,以确定优先靶向CNS组织的衣壳的一组序列规则。参考下文表9,和SEQ ID NO:2(AAV5 VP1),以下氨基酸由此可以在Xaa1-Xaa9任何一个或多个位置处以任何组合独立地突变,以提供与野生型AAV5 VP1衣壳(SEQ ID NO:1)相比具有增强的CNS嗜性的VP1衣壳,其中CNS嗜性在此是指优选CNS转导的特性,而不是优选转导所有其他所收获组织的特性。

[1796]

表 9
CNS 嗜性规则

[1797]

Xaa1选自A、C、K、M、Q、R、T或W
Xaa1选自K、Q、R或W
Xaa1为K
Xaa2选自F、I、K、R、T或W
Xaa2选自F、I、R或T
Xaa2为R
Xaa3选自A、H、N、R或W
Xaa3选自A、R或W
Xaa3为R
Xaa4选自E、G、I、M、Q或R
Xaa4选自E、M或R
Xaa4为R
Xaa5选自C、G、K、I 、M或R
Xaa5选自K、I或R
Xaa5为I
Xaa6选自I、K、L、P、Q、R、Y
Xaa6选自K、R或Y
Xaa6为R
Xaa7选自D、I、K、R、V或W
Xaa7选自I、R或V
Xaa7为V
Xaa8选自C、G、H、K、L或V
Xaa8选自H、K或V
Xaa8为H
Xaa9选自I、K、L、R或V
Xaa9选自I、K或R
Xaa9为R

[1798] 下文表10提供了在CNS组织中发现并且符合表9中提供的一种或多种规则的变体AAV VP1衣壳多肽的581至589区域的序列。因此，本公开提供了由具有SEQ ID NO:2的工程化AAV5 VP1衣壳多肽组成的rAAV，其中Xaa1至Xaa9区域是SEQ ID NO:7118-SEQ ID NO:8117中的任一者，如表10中所公开。本文还涵盖由工程化AAV5VP2衣壳多肽和工程化AAV5 VP3衣壳多肽组成的rAAV，所述多肽在对应于AAV5 VP1 581至589区域中的氨基酸的AAV5 VP2(氨基酸残基445至453)和AAV5 VP3(氨基酸残基389-397)中的区域处具有下文表中公开的序列。

[1799] 表10

[1800] 驱动CNS组织嗜性的AAV5 VP1衣壳多肽中581至589区域的序列

[1801]

SEQ ID NO	581-589 序列
7118	KTTFDVACY
7119	KDGYREAWM
7120	KTSNVMDQM
7121	KTLHMLESK
7122	KSHDYHGRA
7123	KSEWTCCPQ
7124	KMHGNERCF
7125	KQTVFVDQE
7126	KVAHEMMSA
7127	KMNTWDQAR
7128	KFNTKDQLM
7129	KCQSKECGG
7130	KGLITENNY
7131	KVEWRWGSA
7132	KTFQYMGPT
7133	KCGLTPGVE
7134	KQRIAYEWQ
7135	KGQRRDSVS
7136	KQCSLAPTI
7137	KEWCVQQQC
7138	KKDPLHLGD
7139	KQTKSMVYM
7140	KISGLNDAS
7141	KPDILKTKT
7142	KWSAYFLIE
7143	KGESQCTSQ
7144	KQEYIMRQR
7145	KLDPRQADQ
7146	KIYMQCDQE
7147	KWAASHKHD
7148	KENKIMQAS
7149	KVAGGHETT
7150	KTDPIAEYC
7151	KPDHAMAQQ
7152	KSDKVWYVL
7153	KIEAEEMHA
7154	KKTECEMIV
7155	KGTPNWWGP
7156	KPAGKNTCW
7157	KSVHYETCE
7158	KCNVITYHG
7159	KQEQFHTGS
7160	KTMADMYNR
7161	KTVETYAAH
7162	KSWVWVDTE
7163	KAIWQGTEN
7164	KPLTYTVGS
7165	KYASMDIYS
7166	KCTMITPCF
7167	KEAYTDMSC
7168	KHMPTLQWQ
7169	KIEYGHWIH
7170	KYSPFQEHN
7171	KHTLAEMYG
7172	KIWAGVNLG
7173	KCDDRNTQH
7174	KSIMMSQLD
7175	KKIFCSWDY
7176	KQAGYDFFH
7177	KSDAQRNVA
7178	KPPSWPMSL
7179	KAHDITWNL
7180	KQATNNNGW
7181	KDPDCYDWP
7182	KHTEYCSHE
7183	KWYHYNEMT
7184	KGQMISETR
7185	KHEPMCVQL
7186	KVDTVRNPK
7187	KRVQQMETM
7188	KIIFELGMD
7189	KSGWGAPYE
7190	KIQMAVSLD
7191	KVPSQLVFE
7192	KSPDMTLLL
7193	KSMTCVYYP
7194	KPDKPNDYA
7195	KFICTTIAC
7196	KPAYEHRGF
7197	KNNTEDGGR
7198	KATREYDIP
7199	KCETKKRFN
7200	KTEKWASHE
7201	KATVEGSWR
7202	KRRDGTRSY
7203	KALHYAMIK
7204	KWEVMSAAY
7205	KALYLHTIG
7206	KLNPFSTGT
7207	KVEIINWHP
7208	KNDPIDNFG
7209	KTFTKEVTP
7210	KCGNEGRMS
7211	KNETFIAST
7212	KDNQSCEWN
7213	KAQVVPVEC
7214	KMMAFEKQC

[1802]

7215	KWQEMPTFQ	7256	KAPWAPMDS	7297	IRMIGKACN	7338	HRMNNHAWF
7216	KYCPPQTPN	7257	KRAIVGYAL	7298	NRLDQITRR	7339	QRTQCNFSI
7217	KCAMFNAGE	7258	KSGMIQEET	7299	SRSWFDSCG	7340	TRILNEFFC
7218	KSTIPQTYP	7259	KSYQMIMTW	7300	VRCQQHQGN	7341	ARLEIEYMG
7219	KFTEALSCG	7260	KFELHGTVG	7301	YRSWMCWDI	7342	WRFCDMRMS
7220	KGEIEGGAK	7261	KSGTCGYGQ	7302	CRNTQYEQH	7343	DRLYCKLAH
7221	KQLTFPMLA	7262	KSLQCASRH	7303	ARKNCDHLN	7344	ERKTTADWK
7222	KEFVGFQAM	7263	KTGTTMEPM	7304	ARCQNHWSH	7345	GRSMVLHFV
7223	KHDGFSVVG	7264	KTDYSRLSD	7305	MRQRTYICQ	7346	TRHDNGPSD
7224	KATRDPKFP	7265	KHYKWGHCQ	7306	SRCTVMDPG	7347	PRAPIRPPQ
7225	KRWLHMYQG	7266	KLYGLNHSC	7307	PRNLVAMYW	7348	MRRKTPEDD
7226	KHLDMTSCH	7267	KLKESMESA	7308	TRMQAVDHE	7349	ERFRVQEFQ
7227	KDEQFARCK	7268	KAGYQEGSC	7309	FRAVYYYHS	7350	ERYYPISSF
7228	KTLPCSNSE	7269	KCWDCHEAG	7310	RREAMDDSP	7351	GRAHGNVCI
7229	KYMQENVLS	7270	KPVQCACQN	7311	GRQSINEIA	7352	ARSGINGHE
7230	KSAPRENFN	7271	KDYWAGYQM	7312	TRHPTIAMH	7353	ARVALVKPG
7231	KQRMTRQGE	7272	KSEHHWINA	7313	YRLHVWMKV	7354	SRYVNNGTV
7232	KARQDDVNT	7273	KGGPWAYNC	7314	DRREKQLMK	7355	ARAYNSYQA
7233	KCGVEHEHR	7274	KYAYSLDCA	7315	ERKAMRWSL	7356	PRFKACHDP
7234	KSELFMKLH	7275	KLWIDQRGH	7316	LRMQKEGQE	7357	QRQGQLHYW
7235	KNELENVIN	7276	KLWIDQRGY	7317	IRPQYAKMD	7358	SRCTIDVPI
7236	KYWCKNEWD	7277	QRVTKMHFE	7318	SRTICVHGN	7359	DRTMCKVQG
7237	KASLGACVE	7278	ERTACKMKS	7319	NRCGVSKCC	7360	LRVSQSFSN
7238	KSCNDGQQN	7279	ERGEMKATG	7320	CRCNDAWNM	7361	NRTNSPYVH
7239	KVTPILEEV	7280	ARGTCNYCY	7321	FRKQGHGEK	7362	CRGSLSNAN
7240	KQEWHVHTE	7281	QRNEMSASR	7322	YRMEYGTCE	7363	VRQSVPHWA
7241	KPKCDEQKS	7282	FRDVCMHDV	7323	SRPFNYFQR	7364	HRQESPNNK
7242	KWHGGPGVV	7283	ERMMWEEYP	7324	QRGPLQHQQ	7365	TRNIGGVFQ
7243	KGNSVTHVQ	7284	HRMSYKWGW	7325	MRQCNYDFL	7366	RRWMNAKAT
7244	KWDDMHTQM	7285	IRVADDMTC	7326	MRQCNYDFW	7367	YRAHSAGYC
7245	KQAHLEVAW	7286	MRSICEHAM	7327	MRQCNYGFW	7368	YRAHSAGYY
7246	KVNACGNLT	7287	NRGIWLINF	7328	SRTRMYMDH	7369	PPRTVARFK
7247	KAIPQGSSC	7288	NRQFWYLNT	7329	ERTLTKAGS	7370	MTRSCQTEG
7248	KINCRMDGC	7289	PRLLRPNRV	7330	WRLAALYQS	7371	SMRNSKGSG
7249	KQRDWENFA	7290	TRRVATHMP	7331	CRCCPVMKD	7372	GGRSLHAYL
7250	KMCWWPESM	7291	VRHSRNNTV	7332	YRDRCWNRI	7373	GTRIYGKVS
7251	KQNLMCCHN	7292	SRWFQIGTE	7333	LRNMRGDWD	7374	EVRSTRKNH
7252	KISPWDGFY	7293	TRQWMCVHC	7334	CRKCEISAR	7375	GLRHGFYAV
7253	KGMQMWNCE	7294	HRVHSEKTD	7335	ARTQFHSDE	7376	LARTTTNSE
7254	KAILNAMME	7295	GRHTAFATG	7336	ERSMKDADE	7377	VVREQFGIH
7255	KIDAHSNME	7296	LRSRVIKDH	7337	YRRWQSNTT	7378	ANRADWQHL

[1803]

7379	WCRTPQREN	7420	AIRQFMMHL	7461	TIRYIGMSG	7502	EYSRMNKAS
7380	WKRTAGVHF	7421	LMRWSKDSN	7462	FFRICGSYI	7503	DKVRQGQKD
7381	GYRNYFIWP	7422	LMRGSKDSN	7463	NTRSDLRAG	7504	EFNRYPHGN
7382	TARNFGDPP	7423	LMRWSKDSS	7464	ANRMIWYPY	7505	HEMRRDYKS
7383	EARISGMHD	7424	LMRWSRDSN	7465	AQRCFHRWG	7506	QVYRMMMQP
7384	MARLYHPQM	7425	ASRQFETCW	7466	TCRYEELAD	7507	TVNRAMREE
7385	SNRYVSLFI	7426	SDRMVSISA	7467	STRTTDWNM	7508	YKLRSSFRW
7386	PWRWHKSAC	7427	NTRIVCDLS	7468	DARKTYICY	7509	FANRCPPDD
7387	TLRAEMSFH	7428	EFRCRLIGS	7469	SSRGQDGVM	7510	NKIRDVAVI
7388	EQRHISRAP	7429	DCRCYDCHD	7470	RTREPFWWW	7511	PVARKMSTC
7389	TKRIKCCWE	7430	TMRDSDYNW	7471	IERWWVIDV	7512	VIQRKFDSG
7390	MKRAPTLDY	7431	DVRACMEDL	7472	ENRSPNNPA	7513	ALDRERLIF
7391	ENRIAVWFN	7432	WCRCKRFIE	7473	DWRYFCVSL	7514	AGSRMQWAM
7392	AERDHKTIY	7433	WIRSAQRPG	7474	NPRQGFVPV	7515	WSSREYWHK
7393	ELRMSFTKV	7434	MGRPHAWWG	7475	NSRLLIKHL	7516	THARFKMKS
7394	CVRIGMREP	7435	NFRMQTGCH	7476	HTRVVKSTG	7517	NCARGPMEK
7395	TARSTSDKC	7436	SMRDLIAEA	7477	EIRQFVKGA	7518	LWSRFISMD
7396	GKRVEDTSC	7437	TARDREHAM	7478	QERSMNAWW	7519	IAMRCWGNE
7397	NQRGFCKDE	7438	CSRNRCWSG	7479	CARLEGRGN	7520	LVDRSGRTA
7398	TPRCIRADG	7439	DYREKYKCQ	7480	FMRDVFAWD	7521	SQSRTIIQE
7399	DQRMLLKEL	7440	VCRQKTDSW	7481	MLRTQIHCA	7522	HEVRRYGQN
7400	SNRYYSRIF	7441	YSRMEENDR	7482	QLRCEYQHT	7523	NSTRNTSYY
7401	EDRMGLPHN	7442	PIRYVKPAR	7483	CCRMDNKKS	7524	GPARSFEAC
7402	EGRTISYGN	7443	QPRVYEGFY	7484	YIRVGGNNE	7525	YCPRLVEAW
7403	ASRSCMTQE	7444	MDRMMYMLG	7485	NVRNFFDHC	7526	VIGRTVDDP
7404	QSRQMMLLH	7445	VNRGDLSLF	7486	ATRWVQMGQ	7527	WFDRSWHIL
7405	MTRVSMTWR	7446	HTRKEGFNS	7487	MDRDIPHCR	7528	CYIRNMVNL
7406	TFRHWEGHH	7447	LFRTIPIWQ	7488	CYRRSDSFC	7529	LIGRDSKVW
7407	VARQMQENQ	7448	APRREEQNC	7489	EYSRCCHQE	7530	QNIRLGVPT
7408	INRPMGHGM	7449	NTRQINGVD	7490	VHLRSHTWF	7531	VICRMKKHS
7409	CIRFMMREG	7450	DARHKWCLE	7491	QQLRCHCFS	7532	VEDRVFACC
7410	FFRWRWGLG	7451	AERQACTMG	7492	AAIRYDPGF	7533	GHLRDHPLY
7411	QGRWACWYS	7452	ISRMLRQSS	7493	MSNRIIDIQ	7534	FIFRKCDHW
7412	HHRVNYEMY	7453	ETRVREDAV	7494	VTQRPLHHS	7535	RKTRCFCNY
7413	SVRTCHQDE	7454	AMRAGLENS	7495	NQQRTYNNG	7536	TSERHSCMV
7414	DDRVKMFHY	7455	TCRWFAHYS	7496	SLGRIECDF	7537	SMSRHMMHQ
7415	GMRYSEWHW	7456	QVRFEQNPA	7497	SADRTAMFH	7538	DIGRFTAYY
7416	ETREHHLLV	7457	NERVYGWGM	7498	NILRYLICD	7539	GVSRLPYHA
7417	MFRQFYIQG	7458	EKRLFDRTY	7499	TFMRKICGM	7540	QVDRTMWMH
7418	TTRQISNID	7459	VSRCGIDRS	7500	CWARSVAQN	7541	CILRKPSYH
7419	DCRHTILYV	7460	DARKTVNEW	7501	ELPRYYGQM	7542	NNDRHQHLK

[1804]

7543	TVDRKVEIE	7584	YAQRTCWPE	7625	VQPTISYSD	7666	ACSQIGQDD
7544	NSARYAWVY	7585	WCMRDWYQE	7626	YSPLIGRWK	7667	DSQEIGSES
7545	AITRHVADT	7586	ISQRGTLWC	7627	IHSNIPVAE	7668	NWEVIGRWH
7546	TDKRAHVGA	7587	RMVRVSVHG	7628	NHDLIQFVK	7669	HSGFIDRQA
7547	LLWRRVKRT	7588	EAWRNSSSG	7629	WSPSIPGGE	7670	PVKPIIHAS
7548	YVIRKTWFN	7589	LWDRFCCPV	7630	CNYPIIWDK	7671	IQAVIASGW
7549	FAERANYCC	7590	NHTRGKQCV	7631	VSAYIERQW	7672	DHLSICCQN
7550	SYMRHPNFD	7591	NCERWFQIE	7632	QVSHISEHW	7673	CTKPIVMDW
7551	GGDRMATDR	7592	QAVRTRTFR	7633	MMTTIQYHD	7674	LNHPIADYD
7552	FQTRPSCSF	7593	AYLRRNVGN	7634	WAEAIKHFT	7675	TQAKIECCS
7553	MMVRREFAD	7594	LTNRTQESW	7635	VSKFIAASW	7676	GTEHIPMQA
7554	AVMRNDQWL	7595	NVDRQHAQS	7636	CSNDINGRA	7677	SFPMILVNC
7555	LQERDHHMQ	7596	DYERSIENH	7637	LNSMIMWGN	7678	GCQQIQLQT
7556	LTPRSEGLS	7597	MPTRMQKGA	7638	HACVITLDG	7679	ATDMIFVSN
7557	LVGRTYQHQ	7598	MSARINTPD	7639	GTEPIMWRD	7680	RQCVIGDYD
7558	DATRAQTGE	7599	CSIRGFSAF	7640	LSTLITMMD	7681	RLAWIEQSP
7559	MMVRELMSS	7600	MTERYYWCQ	7641	GWALIVNGL	7682	YEHPIMWPV
7560	CGHRHSSMG	7601	CHIRDCSPI	7642	MVNHILPNN	7683	HEGLIMRLE
7561	HHLRNVHHC	7602	NNVRKVGNT	7643	RSLDIQYDI	7684	TSYCIQTVD
7562	NSTRWYHTG	7603	VEMRANKGL	7644	RVVPIEDSM	7685	AMEMIVMHE
7563	NHMRKQSSH	7604	WDGRPDLST	7645	LVNTIWKAA	7686	MGCLIQFCE
7564	DDCRLNVRT	7605	FSYQIQCQN	7646	GYKFIGRSR	7687	FSWPIHEPD
7565	MCARGAMCA	7606	TSSEITSPP	7647	QYLLIPHLP	7688	ESTHIMKTP
7566	SWGRIPHES	7607	CTKPIPCWM	7648	NNGPIWISS	7689	HPDKIEYWR
7567	QLKRNEQTS	7608	IMHEIPEID	7649	YGENISPSP	7690	TFGCIESMQ
7568	RNQRTYPAL	7609	HQSPIHTNT	7650	QSDWISIPC	7691	EDGLIVYGP
7569	SFFREDYSF	7610	REMMIVNCD	7651	CMAHIWEGH	7692	DHIHIERAN
7570	DSHRCLKAD	7611	IWPQIQYRM	7652	NWFGINSWH	7693	GCMCIRHAY
7571	SAPRNEITR	7612	PEKWIHMHE	7653	FQTSINYYG	7694	APIPINMEG
7572	DIQRWEHEG	7613	ILIFIQNCP	7654	TMEYIGLDQ	7695	QMNAIQLCV
7573	EQLRDLNIE	7614	SACNIIYHT	7655	GKIPIDTFR	7696	SQSQIRQAD
7574	NYARHDILA	7615	SGNQIVQAP	7656	WAYYIRLQP	7697	PIVLIHYDA
7575	MGQRSAANY	7616	HWELIRCHD	7657	VAYPIQTDV	7698	VADWIGHTP
7576	STARHNMQL	7617	NSHTIMNGM	7658	LSIQINNQI	7699	GSNSITRDT
7577	PAYRMECCE	7618	DVDIIGLPA	7659	EQTYISEWG	7700	QWEDITKCS
7578	ETHRGTVGL	7619	QKHMILCHF	7660	TCGDIQLRA	7701	EDQWISNPF
7579	DVCRKPVSH	7620	ACEWINMYN	7661	MAGHIGHHS	7702	QCGFIRLNE
7580	DVMRMEHYL	7621	ATYQIQWPG	7662	GFKDILMVQ	7703	DSDPITVDL
7581	PAWRRYCWT	7622	FIIWIEHRH	7663	VWWMIIHTS	7704	EIGIIRHGH
7582	DMHRGGMWT	7623	GIEQIMGCP	7664	NAAMIATVC	7705	FQTLILSGL
7583	CTLRQIKMH	7624	VIMGIQTDD	7665	SMHLICGEA	7706	DNNMIVRCA

[1805]

7707	MAPDIPHCR	7748	LNYMRRAMW	7789	VADTYRMWH	7830	HFADNQVVQ
7708	MDADIPHCR	7749	FAIWKRNES	7790	DYYLNRASH	7831	MHQFLQVAQ
7709	MDPAIPHCR	7750	WKDDLRSTY	7791	NSKACRGCT	7832	NAGMKGVDV
7710	MDPDIAHCR	7751	ENQQRRNFG	7792	LHNSCRTLE	7833	STQKVPVLA
7711	MDPDILHCR	7752	IMPPKRYFV	7793	TVGSKRDLY	7834	GWDMLSVAK
7712	MDPDIPHCC	7753	VAADSRMMI	7794	FNWFVRAYI	7835	RNLPTSVEA
7713	MDPDIPHCG	7754	IMCHARMMA	7795	TNIKERTGS	7836	GMSDCEVSF
7714	MDPDIPHCH	7755	RDIWERPCV	7796	TTYHHRNTL	7837	GKLTVQVNH
7715	MDPDIPHCL	7756	EHSSPREDN	7797	MWTVARMFP	7838	LILQRNVKD
7716	MDPDIPHCR	7757	EVHLFRGEG	7798	EQFADRAPR	7839	NAKEQVVNN
7717	MDPDIPHCS	7758	MYTDMRCWR	7799	LNELPRAEY	7840	GDHKCHVEW
7718	MDPDIPHFR	7759	LIHYCRCIN	7800	RAEWFRCQH	7841	GFKYCEVDH
7719	MDPDIPHGR	7760	HQCQSRSYI	7801	PANHARNPK	7842	MFDMMSVLK
7720	MDPDIPHWR	7761	GTAVSRQSK	7802	SWQQHRKMS	7843	IVGINHVAW
7721	MDPDIPNCR	7762	VMFYWRYNS	7803	LPDHWRKFD	7844	MMNQCQVWN
7722	MDPDIPPCR	7763	FGDHLRVMY	7804	EDWLRRNIW	7845	QGQHVMVHS
7723	MDPDIPRCR	7764	NSEWLRYTF	7805	DAAMTRTVP	7846	NNCNMNVMK
7724	MDPDIPYCR	7765	AVHWHRKCE	7806	LLWMGRRRR	7847	EKHFMQVDR
7725	MDPDIRHCR	7766	PMELWRTWY	7807	ATCWVRMGQ	7848	PIAQDQVWS
7726	MDPGIPHCR	7767	SQDCGRPES	7808	WVKQKDVLC	7849	ECHKCEVCV
7727	MDPNIPHCR	7768	AAKKGRDLH	7809	MSLDTSVPF	7850	PHAYMQVWR
7728	MDPVIPHCR	7769	QNLWYRPLT	7810	TSHKMMVFS	7851	MQVNLQVIE
7729	MDPYIPHCR	7770	TCYCHRAHE	7811	NMKPKDVFR	7852	TPQPMNVAT
7730	MDTDIPHCR	7771	LACTPRCIF	7812	LLAKYMVHT	7853	AWMCRNVYC
7731	MNPDIPHCR	7772	VAFGMRWLL	7813	QDCERQVFA	7854	CGYWLVVPW
7732	MYPDIPHCR	7773	WAQCDRAVC	7814	PTMLPHVKH	7855	QHTQQLVQN
7733	VDPDIPHCR	7774	FCDIGRSGG	7815	QGFSTDVKP	7856	CPEHNQVGI
7734	HVDQHRVQA	7775	GHQWDRLLA	7816	EPLLPAVGN	7857	HDEKQKVYG
7735	MYLNYRHFA	7776	WLTEKRHFE	7817	IAKDDAVRL	7858	EPFLKEVNV
7736	AVASERQLR	7777	MTGQCRNSQ	7818	MENLKGVYD	7859	AVTQLSVER
7737	LEQLDRCMY	7778	IHPMMRENG	7819	QTNQVVVVE	7860	QHKTKCVDN
7738	LFHQGRFPM	7779	IFQLSRPAW	7820	GTTGHAVNE	7861	IHCNVLVNY
7739	CVCEMRTVD	7780	IKETWRVYF	7821	AFLQTPVFS	7862	MDGHFSVGC
7740	NDCTCRAIM	7781	NTNPGRYGC	7822	ETAFQQVQV	7863	YHQLVQVNE
7741	TIDVKRWGQ	7782	DLVDLRSWM	7823	TWHKQLVGV	7864	GQPVWCVCM
7742	EATYTRYMT	7783	CNEVCRNYL	7824	TKYHYHVKV	7865	MIKPLLVVG
7743	HSGYKRQMK	7784	GFTPDRCRN	7825	RNDWQQVFS	7866	GLGWWNVHT
7744	TYMEYRQHK	7785	DVPIHRVQL	7826	QVGVTVVQR	7867	TMAQWTVHN
7745	VDCWLRDPI	7786	TMEAHRFGS	7827	DSHDMFVYK	7868	EPMPMPVWN
7746	YKTEFRTKL	7787	GSATNRCGN	7828	MYLSFPVHE	7869	IPTFDGVEP
7747	DPYNTRFRT	7788	NSAECREFQ	7829	PWEIEWVYW	7870	ELPDTEVPW

[1806]

7871	QAMLHHVWV	7912	NVYATPVVC	7953	QCLNMMMHG	7994	NAMTCSMHN
7872	FTEWKPVFF	7913	NCHIDHVQD	7954	QVGTGMYHS	7995	TIQHAFKHN
7873	QAVTLEVVC	7914	VDMGVHVGS	7955	WDHYLPEHV	7996	NVFPAETHV
7874	HECDTVVVN	7915	VIKVNTVNP	7956	TVQDNDEHA	7997	TLEPAYRHA
7875	CVNLMAVNN	7916	DFKALDVQC	7957	TKQFFTAHC	7998	YEICKHGHF
7876	QHQAFFVFT	7917	MSAQRYVQT	7958	NFSVCPRHL	7999	HSLVQQLHP
7877	ATEKQMVDH	7918	FCKCFMKHC	7959	ITTDATGHF	8000	VPVATMIHH
7878	TGMHPFVRG	7919	VIPEFMNHA	7960	STFHFQSHD	8001	RVDGADRHY
7879	QVMQEKVRC	7920	AVIHEKPHG	7961	HMDHCQCHE	8002	CHILDGLHD
7880	DFPHMHVGT	7921	MLFDFEEHN	7962	QYWWPATHD	8003	MLGVAPGHL
7881	NAGVVKVWP	7922	RCVQAHFHK	7963	LGEPYKAHD	8004	ANLISQGHC
7882	EITQHFVKH	7923	NPTDLFQHE	7964	ISNTMVEHG	8005	GSYIVQNHC
7883	VADTWQVHC	7924	GEMTRCFHN	7965	VTAFTACHY	8006	WPEAFHLHQ
7884	SIYQKWVTN	7925	IAVASHAHG	7966	EWLSLYCHW	8007	CDFMFTRHI
7885	PIKGVAVQE	7926	WKWYGLPHH	7967	PIHTYKRHP	8008	HVWNDVFHM
7886	IFSHPSVSG	7927	SVTYHEWHP	7968	GYNYYWQHA	8009	ADYSMMWHM
7887	ECSVHPVNC	7928	TDHPWVLHG	7969	ATVHMPKHE	8010	WNWLKVMHE
7888	DKMSLQVQI	7929	AKEHMPHHM	7970	MFSAFSMHD	8011	PVTTKNCHE
7889	EGEWLIVPV	7930	GIHQLIHHC	7971	TCTNSEPHP	8012	FCHEVFSHT
7890	TQNACKVRT	7931	EACMPWAHH	7972	TIDWCEPHC	8013	IHKMNFDHL
7891	TGSLTDVIW	7932	DVMQYQNHM	7973	LMHGVMDHV	8014	WAIFGDIHH
7892	GSAMLMVIR	7933	DVDFLKMHC	7974	YITCQYSHH	8015	QSGTCHYHD
7893	TAGEWIVQH	7934	MYAGWLNHG	7975	LATTTSYHC	8016	WVYSHEKHY
7894	TGMSRYVQD	7935	TNGVQIMHV	7976	GTCLFKTHK	8017	DVKYATLHD
7895	FAGFTPVWM	7936	DGLSYEAHD	7977	DSAKQWHHQ	8018	GTEMMGDHA
7896	NMWYSFVMI	7937	EHCYYKEHE	7978	SYAAFLHHG	8019	QTKVHQQHT
7897	ETEHGSVYT	7938	HPSWPTMHK	7979	PNCDPAHHD	8020	ASSSQYPHN
7898	PSILPVVMF	7939	ISCIMKKHL	7980	FAPEFDNHC	8021	TSLCAKPHF
7899	PEYVKWVAD	7940	IWGTRSAHR	7981	QISMKGEHG	8022	GYTQQDEHF
7900	TIHPAWVTL	7941	NQCWCPKHY	7982	QLTTNDAHG	8023	DHCHCIAHP
7901	QWFTKGVDE	7942	SFKWCPGHE	7983	CMAQMVYHH	8024	IPACGIMHE
7902	CYAEEPVWP	7943	ADCFRPFHI	7984	CIDVLFKHL	8025	TCNVSWKHE
7903	QIMSAKVFE	7944	IQDCSTLHW	7985	EAGKEGAHF	8026	TKGMWQQHS
7904	MTTFCDVEA	7945	LQHFDLIHL	7986	QYCNTDNHI	8027	SAGPFFDHY
7905	FTGTPPVDM	7946	PQQCGGFHR	7987	YLESKTIHA	8028	HTVNAESHM
7906	ESWMCPVDY	7947	TWGLHHYHT	7988	PHQYGDEHQ	8029	DDGSNTMHG
7907	DCTYCCVKA	7948	VNPPDMIHT	7989	WPTMLSAHM	8030	GVAQYLHHC
7908	AVVEHAVNR	7949	DKANKMMHG	7990	DQFCPFDHY	8031	EFADYNFHF
7909	MDYHCQVED	7950	ETMSEGIHC	7991	LPTQVYGHN	8032	NTCEGWSHF
7910	QPGVVPVPC	7951	FYPYTNCHH	7992	DTHPMSNHR	8033	FPKCAIWHY
7911	MPYPLNVKE	7952	PAAPLILHY	7993	MGVITSIHQ	8034	AFGLTTGHT

[1807]

8035	STGSFSIHC	8056	FAQQDKQDR	8077	HTLQWEAVR	8098	AHMNYCLSR
8036	HHEFTNSHA	8057	DANWTVRRR	8078	QCAHFYQTR	8099	NYCKYNKMR
8037	QVVIDPAHY	8058	MCNGGIPAR	8079	EWEGVSQFR	8100	GVGPSHGER
8038	CAGMAWSHS	8059	EGATAWQDR	8080	QAVIAVDGR	8101	AKISATWDR
8039	TVQWNFAHM	8060	SHIQLNAWR	8081	GITEQPLAR	8102	INMEYLTDR
8040	RCVQDHFHK	8061	IPAEWDSWR	8082	GIVMLKCGR	8103	YTGHTNSGR
8041	LLWMGQRHR	8062	DAHNDALAR	8083	TLLLVGKDR	8104	MHCLHQTMR
8042	QQTWEQMQR	8063	QPQSRNGAR	8084	GLWLNVAPR	8105	YWCPSQWWR
8043	CAMPHYQVR	8064	NWNHCHQWR	8085	QTEMKKCAR	8106	MLAFDPMGR
8044	LGKVQTHIR	8065	CWHFFGADR	8086	GHIDLEIMR	8107	IIAVSMRGR
8045	CSVWMQRDR	8066	CAMDNTMCR	8087	RMTPNNCQR	8108	VWEDNPPFR
8046	VFPKDEQFR	8067	TFVYGARDR	8088	TCGPKLTQR	8109	CTLMHSHTR
8047	EHITEQDLR	8068	QLDTYMEGR	8089	HIGWMGWLR	8110	LLWMEQRRR
8048	DKTVQTSAR	8069	AATTHHSAR	8090	QETHWAMMR	8111	LLWMGQQRR
8049	GGMWQCEVR	8070	YTNHYYNAR	8091	PYHNPWECR	8112	LLWMGQRCR
8050	LEKNSSQMR	8071	VGIWFIPCR	8092	YDLWGWSTR	8113	LLWMGQRRR
8051	CDAEAEFRR	8072	WSEDMSQYR	8093	QHYNSMIDR	8114	LLWMGQWRR
8052	MGGSCKQER	8073	MCSQTISMR	8094	TYMTNLGDR	8115	MDPDMPHCR
8053	TCVIGKGRR	8074	RGLAQDCCR	8095	GKLTFGTER	8116	MDPDSPHCR
8054	VAFFWQRCR	8075	FCNKLTSVR	8096	GVHDPLHDR	8117	MDPDVPHCR
8055	FAESMWCDR	8076	PTNYGWMWR	8097	TTYLSMRDR		

[1808] 实施例6

[1809] 在脾脏中具有组织嗜性的AAV5变体

[1810] 此实施例描述了使用实施例1-实施例2中描述的方法和系统发现的在脾脏中具有组织嗜性的工程化AAV5变体。

[1811] 分析CNS、肝脏、骨骼肌、心脏、肺、脾脏、淋巴结、骨髓、乳腺、皮肤、肾上腺、甲状腺、结肠、坐骨神经和脊髓组织以鉴定AAV5VP1 581-589区域中驱动脾脏嗜性的氨基酸残基。结果显示于图18F中，该图显示了相较于所有其他分析组织在脾脏组织中AAV5 VP1的581至589区域中每个氨基酸的位置频率，其中衣壳多肽具有SEQ ID NO:2的序列。位置(“Xaan”，n＝1-9)在x轴上表示并且氨基酸残基在y轴上表示。特定位置的特定氨基酸残基的频率在热图中作为强度表示。较深的方块表示特定氨基酸在特定位置的频率较高，较浅的方块表示特定氨基酸在特定位置的频率较低。

[1812] 分析相对于给定氨基酸残基出现在所有其他所收获组织中鉴定的变体中的指定位置的频率而言该给定氨基酸残基出现在脾脏中鉴定的变体中对应于SEQ ID NO:1(推及于SEQ ID NO:2)的位置581至位置589的所述指定位置的频率，以确定优先靶向脾脏组织的衣壳的一组序列规则。参考下文表11，和SEQ ID NO:2(AAV5 VP1)，以下氨基酸由此可以在Xaa1-Xaa9任何一个或多个位置处以任何组合独立地突变，以提供与野生型AAV5 VP1衣壳(SEQ ID NO:1)相比具有增强的脾脏嗜性的VP1衣壳，其中脾脏嗜性在此是指优选脾脏转导的特性，而不是优选转导所有其他所收获组织的特性。

[1813]

表 11 脾脏嗜性规则
Xaa1 选自 C、F、H、I、L、P、W 或 Y
Xaa1 选自 C、F、P、W 或 Y
Xaa1 选自 P、W 或 Y
Xaa1 为 P
Xaa2 选自 D、E、L、N、P、R 或 W
Xaa2 选自 D、E 或 W
Xaa2 为 D
Xaa3 选自 C、D、E、P 或 W
Xaa3 选自 D、P 或 W
Xaa3 为 P
Xaa4 选自 C、F、G、H、R、W 或 Y
Xaa4 选自 C、H 或 W
Xaa4 为 C
Xaa5 选自 A、D、E、G 、P、 R 或 W
Xaa5 选自 D、E、G 或 P
Xaa5 为 D
Xaa6 选自 A、C、D、E、K、R、W
Xaa6 选自 C、K 或 R
Xaa6 为 K
Xaa7 选自 F、L、P、R、W、Y
Xaa7 选自 L、P 或 W
Xaa7 为 P

[1814]

Xaa8 选自 E、I、K、L、P、R 或 T
Xaa8 选自 P、R 或 K
Xaa8 为 K
Xaa9 选自 C、H、M、T、V 或 W
Xaa9 选自 C、T 或 V
Xaa9 为 V

[1815] 下文表12提供了在脾脏组织中发现并且符合表11中提供的一种或多种规则的变体AAV VP1衣壳多肽的581至589区域的序列。因此，本公开提供了由具有SEQ ID NO:2的工程化AAV5 VP1衣壳多肽组成的rAAV，其中Xaa1至Xaa9区域是SEQ ID NO:37438-SEQ ID NO:38437中的任一者，如表12中所公开。本文还涵盖由工程化AAV5 VP2衣壳多肽和工程化AAV5 VP3衣壳多肽组成的rAAV，所述多肽在对应于AAV5 VP1 581至589区域中的氨基酸的AAV5

VP2(氨基酸残基445至453)和AAV5 VP3(氨基酸残基389-397)中的区域处具有下文表中公开的序列。

[1816] 表12驱动脾脏组织嗜性的AAV5 VP1衣壳多肽中581至589区域的序列

[1817]

SEQ ID NO	581-589 序列						
		37438	PANFGQAYP	37440	PCYGQTDWN	37442	PDPRPATND
		37439	PCVQDLRAI	37441	PDAQNEMWH	37443	PDTGHEWWW

[1818]

37444	PDWREATVV	37485	PACLGRTDA	37526	PAQQLTCIK	37567	PCHAFEPDH
37445	PEDPTVTNH	37486	PACNCQIAS	37527	PARCAQFVA	37568	PCHMMVIAN
37446	PEKHSCSTT	37487	PACRKGYWN	37528	PAREDGNGW	37569	PCHQQELDK
37447	PFILKQRNQ	37488	PADEAFVEC	37529	PARKAYCLP	37570	PCHQYWIED
37448	PGMTTTPSF	37489	PADKSQLGA	37530	PARLISDYE	37571	PCHRYSKFM
37449	PGSLECAAL	37490	PADLNLSIP	37531	PARSLWQTY	37572	PCKCYPRKC
37450	PHVANMGPG	37491	PADNLHCHG	37532	PARSTYGAA	37573	PCLHMQYNP
37451	PHVNYDTLC	37492	PADSAKQRK	37533	PASCMWFSH	37574	PCLHTCHRG
37452	PIWGNCAKT	37493	PADYGWCVG	37534	PASLGIPKM	37575	PCLWPMGCY
37453	PKGSQYGEQ	37494	PADYRAYGD	37535	PATEVQQFM	37576	PCMHVQQWE
37454	PKGVLLGGA	37495	PAECKYIYA	37536	PATNQCKRI	37577	PCMKKHSST
37455	PKLQYHSES	37496	PAEFPFMIH	37537	PATQERGEK	37578	PCMQNFDDG
37456	PLARIEPLW	37497	PAEFRCTPF	37538	PAVADDMYK	37579	PCNKWNPKQ
37457	PLECKMWYN	37498	PAEPGQSMG	37539	PAVHMQLDP	37580	PCNLMRDQN
37458	PMPFCKHQF	37499	PAEVTNFGV	37540	PAVNANATN	37581	PCNLVWNCE
37459	PMQQKLWME	37500	PAEWCLQHV	37541	PAVNVPMST	37582	PCNTDPSIH
37460	PMRWNKFVP	37501	PAFGCMLNH	37542	PAVQCSWHG	37583	PCPDWYVMS
37461	PMVKLEETS	37502	PAGCHGKTH	37543	PAVQGLPKM	37584	PCPHDRLNA
37462	PNCKHSIAV	37503	PAGCVCASS	37544	PAVSACPYW	37585	PCPIIWESG
37463	PNTCTKDGL	37504	PAGMAMPGT	37545	PAYGSPPWA	37586	PCPLSQYQT
37464	PPHRFRHIP	37505	PAGNAAWTT	37546	PCADTMSVW	37587	PCPSEDWFW
37465	PPWPPETAE	37506	PAGWNECAC	37547	PCAFETTQH	37588	PCQSIKCCW
37466	PQASKAFSK	37507	PAHKVDLCF	37548	PCAGKDICN	37589	PCQSSWLMI
37467	PQQQYLSVC	37508	PAHQCTVHD	37549	PCAPYARSK	37590	PCRVGFSYE
37468	PQTAYMQAY	37509	PAHYVPGIE	37550	PCAQWYCDC	37591	PCSESWSNT
37469	PRYFVEACE	37510	PAISNKFAG	37551	PCCEDVSNI	37592	PCSHYCART
37470	PSHWFCGGA	37511	PAKLYITVR	37552	PCCEGTMDN	37593	PCSWFYHGG
37471	PTDFIRVNE	37512	PAKQAADYG	37553	PCCQICNTW	37594	PCSYTIFDT
37472	PTRATMFID	37513	PAKYLPWQY	37554	PCCYTQRPT	37595	PCTCCFNRQ
37473	PVAWRENSL	37514	PALRETPWT	37555	PCDRTCPDM	37596	PCTCVWYRC
37474	PVTRFECCK	37515	PALWRYYKQ	37556	PCEFKSTET	37597	PCTGQLFMA
37475	PYCICTQPY	37516	PAMMMPTKQ	37557	PCEGHEKPQ	37598	PCTQKQKPF
37476	PYHDNSAPQ	37517	PAMRYDSLY	37558	PCERALLQH	37599	PCTWAHSHD
37477	PYHLNQDNN	37518	PANNCRIHT	37559	PCEVQNAQT	37600	PCVHGTQTW
37478	PWAIEYSLV	37519	PAPCWWNVT	37560	PCFAVMAGC	37601	PCVKKDCEE
37479	PYIMEFHPF	37520	PAPDNSDTC	37561	PCFCHYTNC	37602	PCVNKVGTM
37480	PKEHFPTCM	37521	PAPHAWDVH	37562	PCFGPSYAD	37603	PCVWQWHCT
37481	PAAAMFLPS	37522	PAPIWQVLD	37563	PCFVKDSTA	37604	PCWLKPTSW
37482	PAAHKEWEW	37523	PAPKHSYQN	37564	PCGGLTHPH	37605	PCWRRDIHP
37483	PAAHNACHC	37524	PAPPLAQSN	37565	PCGLMDDYP	37606	PCWVNDHPM
37484	PAALDLYNG	37525	PAQNAPMCG	37566	PCGQWKIQN	37607	PCWWKWMCG

[1819]

37608	PCYLCMLIL	37649	PDMIVDPDN	37690	PECFLDSNV	37731	PEMWMILEG
37609	PCYWAIAGV	37650	PDMLTDSTC	37691	PECFSPENM	37732	PENHKMGPW
37610	PCYYFGMPH	37651	PDMPHDRHC	37692	PECLINYTC	37733	PENTFCSTY
37611	PDACSPFGA	37652	PDMQWNGEW	37693	PECQSIDSV	37734	PENVGASGA
37612	PDADQFVPQ	37653	PDMTCTINI	37694	PECSNWRNH	37735	PENWRPWVT
37613	PDAGWQNSP	37654	PDNIFWGEA	37695	PECVIEYED	37736	PEPKLIGLF
37614	PDAMKDCVT	37655	PDNLAPYVP	37696	PECVRYSKC	37737	PEPLGQGTW
37615	PDANIEFLC	37656	PDNMITMQH	37697	PEDCCRQDS	37738	PEPQTQDDG
37616	PDAPWEGWH	37657	PDNNFLSNQ	37698	PEDMTLDGS	37739	PEPRNVMKS
37617	PDCCDMQCF	37658	PDPGHRMYA	37699	PEDNNIRRN	37740	PEQCSERIV
37618	PDCTERFME	37659	PDPHCEDRK	37700	PEDSGYHWH	37741	PEQNDVNGM
37619	PDDEQVRVN	37660	PDPHIVWCG	37701	PEEANQQVT	37742	PEQRRRYDG
37620	PDDEVCSKV	37661	PDPLARRFE	37702	PEEAPKQMF	37743	PEQWSWGVC
37621	PDDGHKHPF	37662	PDQHTHLAM	37703	PEEETMYAC	37744	PEQWWRYQA
37622	PDDMWQYEA	37663	PDQRGQTGQ	37704	PEEHEVEQT	37745	PEQYKYYQC
37623	PDEEYAQYW	37664	PDRPCEDVN	37705	PEEINDHCY	37746	PEQYNAQAE
37624	PDEGFLYDN	37665	PDSCLHRTA	37706	PEEYEHMNC	37747	PERINHLGC
37625	PDELHLEDG	37666	PDSVPVQHN	37707	PEFFHWPHS	37748	PERLFTCFV
37626	PDEPRSKCQ	37667	PDTAKQDMS	37708	PEFMSKGDS	37749	PERNKVWVA
37627	PDEQHMGYW	37668	PDTCWNWPR	37709	PEFVQYGSC	37750	PERYNQNQL
37628	PDFQRENFE	37669	PDTSMECDH	37710	PEGCTYPQL	37751	PESCGAAAP
37629	PDGFGVSMP	37670	PDTTMKNTL	37711	PEGHEEHPS	37752	PESDRIYRP
37630	PDGRRYVCS	37671	PDWATLMGN	37712	PEGLVIHMT	37753	PESIKTIQR
37631	PDHAMEYQC	37672	PDWAVMKAP	37713	PEHGIHDGE	37754	PESNKWEDH
37632	PDHEPNGDA	37673	PDWFEENTV	37714	PEHRAANVG	37755	PESNWRCPE
37633	PDHGIQFGW	37674	PDWGSHYCN	37715	PEHSNLFSY	37756	PETAAAKAM
37634	PDHIRMVTP	37675	PDWHASYGN	37716	PEHTVIWES	37757	PETFDTYPH
37635	PDHKLIVYD	37676	PDWNMRMYP	37717	PEHVETHGQ	37758	PETHSESCN
37636	PDHMGLPYH	37677	PDWRYGWYS	37718	PEIFGQMSW	37759	PETWNDRCH
37637	PDHVHCIHA	37678	PDWSGKYDI	37719	PEIIWQIAT	37760	PETYLGIPD
37638	PDHVREHCM	37679	PDWSPPIQN	37720	PEITCKIAD	37761	PEVIQNESM
37639	PDIHGPQSE	37680	PDWWAIEMG	37721	PEIVNPARF	37762	PEVPSTNHA
37640	PDKACSHHP	37681	PDYHIGCNH	37722	PEKVLLGMV	37763	PEVTSRDIM
37641	PDKDNHVAY	37682	PDYMDCAPM	37723	PELYTMFFT	37764	PEVWGGRTQ
37642	PDKSEAQRQ	37683	PDYWKSAFA	37724	PEMALTKGL	37765	PEWGHIGNH
37643	PDKVSDYQE	37684	PEAKWKEHD	37725	PEMCDPADQ	37766	PEWIDGYVM
37644	PDLQPRDTT	37685	PEAMIGQEN	37726	PEMDKFKNM	37767	PEWWGQNIH
37645	PDLVCDYCQ	37686	PEANEVQTR	37727	PEMGEATKV	37768	PEYHWDHYY
37646	PDLVPIVGN	37687	PEARSWWAQ	37728	PEMPRHICR	37769	PEYSTSFEQ
37647	PDLWSPFWY	37688	PEAVWCAFA	37729	PEMTYEPMS	37770	PEYTEFEHQ
37648	PDMFKERSQ	37689	PECAESPTT	37730	PEMVEEMNV	37771	PFAKPNIEE

[1820]

37772	PFCCEMHIP	37813	PFTHPICWH	37854	PGHYAIMKR	37895	PHAEGLFCT
37773	PFCLQDVCH	37814	PFTHQPAGM	37855	PGIHSVMNQ	37896	PHAHPMFMT
37774	PFDTCFICT	37815	PFVGLLEHS	37856	PGIRAQCGE	37897	PHASSPMQP
37775	PFDVDGPIY	37816	PFVKPLDIA	37857	PGKLQTMGP	37898	PHCFPAYCC
37776	PFEAQINFD	37817	PFVQYKWDR	37858	PGKMINPQM	37899	PHCKICRWN
37777	PFECHKSSS	37818	PFWCEAISY	37859	PGKNLKWTD	37900	PHCLADYFM
37778	PFESHDQYT	37819	PFWIHRMIE	37860	PGKVAWSQG	37901	PHCPMIRQD
37779	PFFCSGVKC	37820	PFWLFHAKC	37861	PGLICSIDW	37902	PHCREGEVQ
37780	PFFEPDDEM	37821	PFWVRITPE	37862	PGLPYRNHH	37903	PHCWWKWSG
37781	PFGLVEINT	37822	PFYDQKLSF	37863	PGMEKNGDL	37904	PHDMSMRFY
37782	PFGRMIVDD	37823	PFYKKYTAF	37864	PGNCIPAMG	37905	PHDTNNAKY
37783	PFGWESQQP	37824	PFYPLWENF	37865	PGNEHMQMT	37906	PHEEGNKWG
37784	PFHEDCPTE	37825	PFYQHLSGY	37866	PGNGRLYWT	37907	PHEGHDRKQ
37785	PFHEPWAYA	37826	PFYTSDGCP	37867	PGNRFMTYY	37908	PHEHKEGVC
37786	PFHGVILEN	37827	PFYYFTCCQ	37868	PGNWAKPGG	37909	PHENEAIHW
37787	PFHLPRFPS	37828	PGADTHGPQ	37869	PGPHDMGSL	37910	PHEYIVVKS
37788	PFHVHGEHY	37829	PGAGLGPPS	37870	PGPSDNTPN	37911	PHFKHHKTQ
37789	PFHYNKTLC	37830	PGANYWGCM	37871	PGQGSDTEE	37912	PHFKMEPHD
37790	PFKAHKPSL	37831	PGAYKNQQH	37872	PGQLTSRHF	37913	PHFSRQNGT
37791	PFKHENLRH	37832	PGCLEIFHM	37873	PGQSDSDMV	37914	PHGGQRISN
37792	PFKPCPIEI	37833	PGCPNSNPH	37874	PGQTEERVW	37915	PHGLNEAAI
37793	PFKSDFWAI	37834	PGDCEEAII	37875	PGRIDFWKP	37916	PHHAGTQRS
37794	PFLAGDRPV	37835	PGDCHEAWC	37876	PGRPMWMTN	37917	PHHDVSIPC
37795	PFLMEENYR	37836	PGDQYREWC	37877	PGRYCKTHS	37918	PHHFTVYPF
37796	PFLYVIDRR	37837	PGDRGPIGR	37878	PGSQALQWA	37919	PHHIMWHHP
37797	PFMAHGMYQ	37838	PGDWQTTPT	37879	PGSSNKPYA	37920	PHHLMSCIG
37798	PFMISNALQ	37839	PGDYRQTLP	37880	PGSSSACLC	37921	PHIEVAIGH
37799	PFMQCCYSW	37840	PGEAAFEKQ	37881	PGSTYQAML	37922	PHIIPLEGL
37800	PFNFMGMVQ	37841	PGEAVCHRW	37882	PGSWYMHLP	37923	PHIKDWPRA
37801	PFNPDWGRF	37842	PGEFKGLVS	37883	PGTFWHEQY	37924	PHISCAHAW
37802	PFNRMTDYY	37843	PGFGMNSIR	37884	PGTGVADRN	37925	PHKFIFMTS
37803	PFNSDSYIR	37844	PGGAYVWTR	37885	PGTKACMKH	37926	PHKPGTTKD
37804	PFNYMFLIM	37845	PGGCMWARD	37886	PGTSWIPWM	37927	PHKWIKFAG
37805	PFPAMKIKL	37846	PGGGFCLNG	37887	PGVTNCDVG	37928	PHKWWNVHW
37806	PFPWEWLRL	37847	PGGINNMTV	37888	PGVTVWGYH	37929	PHMCYLLWR
37807	PFQQFLERC	37848	PGGWLEHPA	37889	PGVWLSYNC	37930	PHMWRDNHN
37808	PFQWVWREQ	37849	PGGWQVNPS	37890	PGWGKSFLT	37931	PHMYKWTRM
37809	PFRECKNCT	37850	PGHCYRMGW	37891	PGWHNFTAI	37932	PHNCELNSA
37810	PFSFSQPGS	37851	PGHFWWQHA	37892	PGWLEAWSP	37933	PHNQREHCA
37811	PFSKGAYLA	37852	PGHKFRSMD	37893	PGWPAPTGH	37934	PHNYDWGCS
37812	PFSQWIHSA	37853	PGHWNLYHQ	37894	PGYKTPWAK	37935	PHPENMSKQ

[1821]

37936	PHPFGHREA	37977	PIFMTEQNL	38018	QDHTNAWHQ	38059	GCPPDPWHN
37937	PHPHKQHMF	37978	PIGPSSTSC	38019	QDQDLNAWA	38060	GMPVMAFQT
37938	PHPQCTHMM	37979	PIHHKQTYM	38020	QDRDQSYHR	38061	GPPECHMML
37939	PHPRMGIQQ	37980	PIHRESPRI	38021	RDAANLEFP	38062	GSPQWLEYT
37940	PHQCKQLTS	37981	PIIPQAQLF	38022	RDDLMVKIV	38063	HSPLEFSSP
37941	PHRANSTQY	37982	PIITNCAHS	38023	RDIDHKNDE	38064	IAPLQCPVS
37942	PHRFPDRKM	37983	PIIWELGWA	38024	SDANDRADW	38065	IFPMADCRV
37943	PHSCCCMKC	37984	FDAGEEAYF	38025	SDCAMTIHV	38066	IQPYYWNDT
37944	PHSCLFMWA	37985	FDLDQQVHT	38026	SDCSDFTQF	38067	KTPIKCEQH
37945	PHSEQIKCG	37986	FDPQDDWCY	38027	SDDHAMQMN	38068	KVPPYRTCP
37946	PHSFNTIMS	37987	FDRISVLGV	38028	SDDVKLPGL	38069	KWPGMMGEV
37947	PHSIVATRD	37988	FDTDHRWCW	38029	SDHTCCWTF	38070	LEPNLMNTE
37948	PHSNQAEFN	37989	GDDLYVHGM	38030	SDMDPCSDK	38071	LMPHRTWKC
37949	PHSWAYHTS	37990	GDECFCKLI	38031	SDMWMERSC	38072	LNPHFVFGR
37950	PHTMRVCLP	37991	GDFCFYKMT	38032	SDNCMEADN	38073	LPPTCANGT
37951	PHTPQGTKA	37992	GDFHPDPLC	38033	SDRNALMPN	38074	LVPMAHRAC
37952	PHTREGLAQ	37993	GDGKSTKVT	38034	SDTKWNVCY	38075	MAPQPMSWG
37953	PHTWIKQEG	37994	GDVHHLHSK	38035	TDDFGTPWD	38076	MCPMFFWCE
37954	PHVMGGEWC	37995	GDVPTWPKC	38036	TDEHQGVFQ	38077	MLPYVCDPG
37955	PHVPHEIGI	37996	HDTSQAHNI	38037	TDHDHHQEF	38078	MPPPNHLDV
37956	PHVRCPMDL	37997	IDEDIFFAA	38038	TDKPMMGGH	38079	MTPSWIVED
37957	PHWDKGYDL	37998	IDIHMSANI	38039	TDPTRSTVP	38080	NCPAMPLKI
37958	PHWFFCNMN	37999	IDKNWKGLW	38040	TDRVGMPPY	38081	NRPIRAWKR
37959	PHWMPWNQV	38000	IDQCAKFHR	38041	VDGNEEEKP	38082	NTPERPHRK
37960	PHWQYETGS	38001	IDTFNCDTY	38042	VDHDVDTMQ	38083	NTPITIQPL
37961	PHWRPPLGF	38002	IDWFAGACG	38043	VDNHGFNDW	38084	NYPSAHQCG
37962	PHYDSLAHS	38003	KDAAYVTPD	38044	VDWGVDPWH	38085	QFPEIHRGE
37963	PHYLQVCCA	38004	KDHCYTCID	38045	VDYCKEFVW	38086	QMPTTTDRC
37964	PIAEDFVAV	38005	KDPRITWYV	38046	WDKAQDVRM	38087	QQPMIAEHV
37965	PIANDMHTY	38006	KDQFWGCLP	38047	WDSHCQKIQ	38088	RNPSIQYWS
37966	PIANKHPSI	38007	LDHPMHMMC	38048	WDSPFNYSF	38089	RVPYVETYD
37967	PIAPSGFVG	38008	MDAHGRNHG	38049	WDSRYKCPI	38090	SCPEMDLHR
37968	PIAWVEQFP	38009	MDHPRVEMQ	38050	YDAEHQQLA	38091	SNPIMDSDC
37969	PICFEDCCH	38010	MDLYYTLAN	38051	YDDCTTNQG	38092	SPPISYPRM
37970	PICMIGRWT	38011	MDNFDRPAA	38052	YDFPIMTCE	38093	SQPVMELSE
37971	PIDEAINDS	38012	MDNRELVVK	38053	YDPDKMKPG	38094	SVPFTNMQS
37972	PIDPRQSGE	38013	MDQCQQHTC	38054	CDWMANVSV	38095	SVPMWWCAP
37973	PIDPVYHLI	38014	NDEKAAIDI	38055	LDTIQINEV	38096	TRPCICCDV
37974	PIEMHVACS	38015	NDGSSVHMN	38056	EQPHADQKI	38097	TSPQNDHYS
37975	PIESPIHHS	38016	NDKHTTMEN	38057	FAPSGKFRD	38098	TTPFEKCCA
37976	PIESTPVLV	38017	NDPYNKQSQ	38058	FVPFHWWSH	38099	TVPLQFQDS

[1822]

38100	TVPRYSRNQ	38141	KGMCLSQGD	38182	VWCCSVNPG	38223	MRCADMLRY
38101	TYPRMITPA	38142	KGVCKGVLP	38183	WRNCMQDSK	38224	MTESDDRCV
38102	VGPCVRSDW	38143	KYECPFYYV	38184	WSICHKDSY	38225	NCGLDITMR
38103	VGPPYDTQD	38144	LHDCQMWGD	38185	QNWCGKNLE	38226	NIQADMSQQ
38104	VLPFGAVGT	38145	LHSCDTPWT	38186	SQDCGRPES	38227	NTVMDRMNF
38105	VMPANWFFS	38146	LKECCFEDA	38187	KVGCFVDVW	38228	QCDSDTQSR
38106	VQPCMQCWC	38147	LNECYYNRP	38188	FCKCFMKHC	38229	QTRKDGESH
38107	VRPMHMNCT	38148	LPDCTDYRK	38189	ESQRDDGWL	38230	QTRWDLIGL
38108	VTPRFQWLT	38149	LPRCCTINY	38190	ESRADIEVA	38231	QVCIDWSGG
38109	VYPTHNEVV	38150	LVGCSYCQS	38191	ETCIDMQGK	38232	QVNEDDAAK
38110	WNPVQWTHS	38151	MAGCPPHDS	38192	ETHTDYQGP	38233	REYMDDEKQ
38111	WTPVPSSSN	38152	MASCIPQTS	38193	ETNPDWEGR	38234	RHHIDIPTY
38112	WWPMRKWVS	38153	MFACTQFFG	38194	FCTSDQKAL	38235	RVLTDERMD
38113	YLPYQSICG	38154	MGDCHRSPD	38195	FGDHDNMPE	38236	SAREDDNPC
38114	YNPQQRSLS	38155	MMRCNVDFE	38196	FRHHDVNNH	38237	SCRIDWTDW
38115	IVPQQKIDE	38156	MMSCIAAVQ	38197	GEYYDEVNH	38238	SFNYDMVSR
38116	HGPFCDHAY	38157	MVSCRSAAC	38198	GMEPDTQLY	38239	SGRGDVNGA
38117	KGPIFHSAY	38158	NASCHWHFD	38199	GMVPDLLPH	38240	SMQVDVLQR
38118	ENKCFASTT	38159	NGLCVESTA	38200	GVNMDMEVS	38241	SNQWDWHIF
38119	EQLCGKKKQ	38160	NPACNKMPT	38201	HGELDCASC	38242	SVMLDEWIN
38120	ERTCMPSAY	38161	NQACNGWHK	38202	HHFTDVDAH	38243	SWCIDCEAD
38121	ESDCFMWLG	38162	NSHCTWRKF	38203	HHHADFHLA	38244	TIAVDGHHS
38122	EYECLDCNI	38163	NVECLPMLP	38204	HMHRDMHEE	38245	TILPDSAEF
38123	FCYCSHFPW	38164	QAMCCSYPR	38205	HPNRDSAHM	38246	TIVYDNWSR
38124	FHMCSCTRY	38165	QAMCVHWHM	38206	HRDGDILIE	38247	TTTADKPAS
38125	FSSCFNAVE	38166	QCYCAKYWS	38207	HYTSDWFQA	38248	TTVLDHGLQ
38126	FWRCDPNKV	38167	QGDCYEMWE	38208	IAKEDICQP	38249	TWKEDWALL
38127	FYNCCDNPL	38168	QMTCPSDLW	38209	ICVNDNFHN	38250	TYKRDIMKT
38128	GCMCEGAEF	38169	QQICPVGEF	38210	ISAPDNQLS	38251	VILKDRVNW
38129	GGFCPCDQK	38170	QYACHQIGA	38211	IVTADNTME	38252	VKKSDAKQC
38130	GGSCSWDKP	38171	RGACDEFET	38212	KANWDVNHS	38253	VQDGDYSPS
38131	GIICYDATN	38172	RLMCYFRTW	38213	KENGDSITL	38254	VSMEDHCKH
38132	HRGCCESGR	38173	RPACTVPGK	38214	KVDPDCHMF	38255	WEKFDRPWR
38133	HSFCVEGIP	38174	SGDCRDWHD	38215	KVLQDQFTI	38256	WPSADCDIP
38134	IHFCMCCDF	38175	SGTCYHYAL	38216	KVVTDRKWD	38257	YMSRDQIKS
38135	IHRCIAEAA	38176	STGCGLMWG	38217	KYEPDGFIP	38258	YPTRDNLFM
38136	IHTCPKNTQ	38177	SVYCEEFYT	38218	LVGLDLSCH	38259	YWNFDDYQQ
38137	ILDCHKHYS	38178	TIMCVASLA	38219	MHYSDGFVS	38260	YYTGDLTYK
38138	ITECHPRKT	38179	TLYCYDDYH	38220	MIEYDVGYK	38261	SVKQDCVED
38139	ITYCNVVYW	38180	TRICCKNRQ	38221	MLKTDCPLH	38262	TVREDCCIN
38140	IVVCMHCEQ	38181	VKQCTKFDE	38222	MPCMDRRRP	38263	YFAADEQWG

[1823]

38264	NQIKDVAAS	38305	LMMMCKQQY	38346	GLDWVSVKA	38387	FHGERHQDV
38265	DARFDQWHH	38306	ESSWLYPHN	38347	HGNAETHKD	38388	FVRVQQRAV
38266	ALTHDDGRT	38307	FSITEEPFW	38348	HITNPTSKP	38389	HACIHTDEV
38267	FAYFQKDLF	38308	GVVDYWPTF	38349	HPEWKMSKC	38390	HLSEFNMQV
38268	FNADHKCPS	38309	HHTLNYPNE	38350	IAKQEHVKS	38391	HTFYCTCRV
38269	GRDALKNGD	38310	HWESFIPRV	38351	IVDALHHKW	38392	HVSIATGSV
38270	GVATLKWEQ	38311	IGGWIWPGP	38352	IWTHLCVKH	38393	HYDFQLTQV
38271	GVCSKKSGA	38312	IVSELCPSF	38353	LFIQYYHKS	38394	ICVNKQQSV
38272	HFAEIKYET	38313	KKQMPFPRD	38354	LQNDWDAKL	38395	IFHSHATCV
38273	HMRMCKATW	38314	LGRAPCPCW	38355	LVIPRDTKK	38396	IFSYSQRMV
38274	HPRSTKWCQ	38315	LSNVPMPAQ	38356	LYHETHKKC	38397	IRYPYGKMV
38275	IEEHLKVCY	38316	MAAEAYPCI	38357	MAIPTGVKE	38398	KSSEIVKWV
38276	IVATPKQYG	38317	MHAEFGPLP	38358	MAVEYDRKG	38399	KSYQRTFLV
38277	KLNHVKRTS	38318	MKFQSGPKS	38359	MMNWWMYKA	38400	KYTGGIVVV
38278	MCNYHKLNK	38319	NQFRAEPIS			38401	LTNWNGMHV
38279	MGGPYKLMC	38320	NVTNFDPHP	38360	MVTQTYFKN	38402	MIETTGQQV
38280	MGNIVKPAD	38321	QAAPGPPYE	38361	NCMAVWAKN	38403	MMGPPQHVV
38281	MIDRMKESY	38322	QPNGWAPPC	38362	NLDQWNVKV	38404	MPKWGREPV
38282	MRTNYKLEM	38323	QQFQTDPQT	38363	NTEGCRVKK	38405	MVEATDWWV
38283	NFWLAKPFK	38324	QVEIQEPAI	38364	QIGTPSQKA	38406	MWTTETMGV
38284	NLHGTKFGY	38325	RMLVRVPDG	38365	QIQMTDDKQ	38407	MWVMGPLCV
38285	QMMQTKLHD	38326	RPDMLTPYR	38366	QSAQWNVKA	38408	NHCMWEGYV
38286	QMTIPKWGY	38327	SQCTPFPFN	38367	RTFNRATKT	38409	NHHPFMYYV
38287	QSCWTKCEA	38328	SRDRRQPKK	38368	SLVHFHAKT	38410	NIMLGQNSV
38288	SGILSKDWK	38329	SSFHEDPCN	38369	TFGFFCTKG	38411	NWVHPITRV
38289	SIDEPKFTV	38330	THHGLQPND	38370	TFSTGLSKK	38412	QGCHFNMSV
38290	TEHENKCVG	38331	TKVVIEPNC	38371	TKGFMQKKE	38413	QVVKFEIPV
38291	TGLLRKEQS	38332	TLYGHDPAH	38372	TMDEHHMKI	38414	RRQWQDFMV
38292	TKNEFKTND	38333	TMFSFQPVE	38373	TSVDLQMKA	38415	SAGYMCYDV
38293	TMAQCKYPH	38334	VINNRVPPT	38374	VEMWVWEKK	38416	SGDGFHYGV
38294	TMFSHKPSD	38335	VQNYADPTG	38375	VKQFNAHKG	38417	SIMHLCEHV
38295	VMCVPKYKD	38336	YGTIPVPIT	38376	VNLTTSVKF	38418	SNCSVVKCV
38296	VNEEMKFQP	38337	YIKAVTPPF	38377	WCYQCILKF	38419	SPVILNADV
38297	WNGHHKDML	38338	IREPTEPYC	38378	WSMQLSHKM	38420	STRMCQMLV
38298	YWFLHKGCN	38339	IFQLSRPAW	38379	YAWTTVCKT	38421	SWQQSPYFV
38299	EVARTKFGG	38340	ISGWRMPCD	38380	NMVWKCRKS	38422	TARYYDGPV
38300	QMSWGKYHQ	38341	QAEVFAPYH	38381	VMWTWESKY	38423	TCGYFTRDV
38301	SYYQSKNTD	38342	DAMWKFPST	38382	EVLDSMSEV	38424	TWASRRKCV
38302	LWFFMKFAK	38343	TVNPCIPDN	38383	EVTHPIGAV	38425	VKCLGAMNV
38303	GACLWKRHL	38344	KCLQMCPNG	38384	FASRCPRTV	38426	VNCKHMNNV
38304	THARFKMKS	38345	FCNTHCGKP	38385	FFFTRWMRV	38427	VNYSWNIQV
				38386	FGVEAWLPV		

38428	VVMPFWSAV	38431	YHEHAVESV	38434	YTTFYRFQV	38437	TCGQYVNGV
38429	WTVELANGV	38432	YLEANMCHV	38435	QAMLHHVWV		
38430	WVQPRMSIV	38433	YQHWTTMGV	38436	QNLMYCSSV		

[1824] 实施例7

[1825] 在肾上腺中具有组织嗜性的AAV5变体

[1826] 此实施例描述了使用实施例1-实施例2中描述的方法和系统发现的在肾上腺中具有组织嗜性的工程化AAV5变体。

[1827] 分析CNS、肝脏、骨骼肌、心脏、肺、脾脏、淋巴结、骨髓、乳腺、皮肤、肾上腺、甲状腺、结肠、坐骨神经和脊髓组织以鉴定AAV5 VP1 581-589区域中驱动肾上腺嗜性的氨基酸残基。结果显示于图18K中，该图显示了相较于所有其他分析组织在肾上腺组织中AAV5 VP1的581至589区域中每个氨基酸的位置频率，其中衣壳多肽具有SEQ ID NO:2的序列。位置("Xaan"，n＝1-9)在x轴上表示并且氨基酸残基在y轴上表示。特定位置的特定氨基酸残基的频率在热图中作为强度表示。较深的方块表示特定氨基酸在特定位置的频率较高，较浅的方块表示特定氨基酸在特定位置的频率较低。

[1828] 分析相对于给定氨基酸残基出现在所有其他所收获组织中鉴定的变体中的指定位置的频率而言该给定氨基酸残基出现在肾上腺中鉴定的变体中对应于SEQ ID NO:1(推及于SEQ ID NO:2)的位置581至位置589的所述指定位置的频率，以确定优先靶向肾上腺组织的衣壳的一组序列规则。参考下文表13，和SEQ ID NO:2(AAV5 VP1)，以下氨基酸由此可以在Xaa1-Xaa9任何一个或多个位置处以任何组合独立地突变，以提供与野生型AAV5 VP1衣壳(SEQ ID NO:1)相比具有增强的肾上腺嗜性的VP1衣壳，其中肾上腺嗜性在此是指优选肾上腺转导的特性，而不是优选转导所有其他所收获组织的特性。

[1829]

表13 肾上腺嗜性规则
Xaa1选自A、C、K、Q、R、S或T
Xaa1选自C、K或R
Xaa1为C
Xaa2选自A、C、I、S、T或V

[1830]

Xaa2选自A、V或T
Xaa2为V
Xaa3选自A、F、G、K、M、Q、R、T或V
Xaa3选自A、G或M
Xaa3为M
Xaa4选自A、K、M、Q、R或V
Xaa4选自A、R或K
Xaa4为K
Xaa5选自F、I、L、M、R、T、V或Y
Xaa5选自R、V或Y
Xaa5为V
Xaa6选自G、H、M、N、R或S
Xaa6选自H或N
Xaa6为N
Xaa7选自A、H、K、Q、R、S或V
Xaa7选自H、Q或V
Xaa7为H
Xaa8选自A、G、H、M、Q或S
Xaa8选自A、G、M或S
Xaa8为S
Xaa9选自A、E、N、P、R、S或Y
Xaa9选自P或E
Xaa9为P

[1831] 下文表14提供了在肾上腺组织中发现并且符合表13中提供的一种或多种规则的变体AAV VP1衣壳多肽的581至589区域的序列。因此,本公开提供了由具有SEQ ID NO:2的工程化AAV5 VP1衣壳多肽组成的rAAV,其中Xaa1至Xaa9区域是SEQ ID NO:1118-SEQ ID NO:2117中的任一者,如表14中所公开。本文还涵盖由工程化AAV5 VP2衣壳多肽和工程化AAV5 VP3衣壳多肽组成的rAAV,所述多肽在对应于AAV5 VP1 581至589区域中的氨基酸的AAV5VP2(氨基酸残基445至453)和AAV5 VP3(氨基酸残基389-397)中的区域处具有下文表中公开的序列。

[1832] 表14驱动肾上腺组织嗜性的AAV5 VP1衣壳多肽中581至589区域的序列

[1833]

SEQ ID NO	581-589 序列
1118	CAAQYWLTP
1119	CARFVERWP
1120	CCAKFQMCT
1121	CEADGNWMR
1122	CIDMVQSWS
1123	CMLLAEKMF
1124	CMYPKSTDR
1125	CSLPYVREG
1126	CSVKMPGFR
1127	CTANVWTAQ
1128	CTVVMMRQP
1129	CVIPLRMYM
1130	CVIQTSRGI
1131	CAAYRGLLE
1132	CAFQPVRWF
1133	CAFQRMMYQ
1134	CAHQPQRDR
1135	CANYAFMTE
1136	CAPLVNSFC
1137	CARMFAAQW
1138	CCHFQHKAG
1139	CCLQPELTM
1140	CCRPFVLHQ
1141	CDVKCKKED
1142	CEMIANTNN
1143	CFNIIQRYD
1144	CFRQVQAMS
1145	CGDWLWRNL
1146	CHGHRPFDA
1147	CHIDYDLGF
1148	CHMKFNKFG
1149	CIGKCWPPM
1150	CIGRVNGTR
1151	CIKLTMDDR
1152	CIKSKFSWQ
1153	CINEKQIRW
1154	CIVKFSNTQ
1155	CKMATNGMT
1156	CKMYMHESR
1157	CLMALKRSG
1158	CLSSYSRYQ
1159	CLVKYPQFG
1160	CMYFIAEWH
1161	CPDEPRWFP
1162	CPFKVFAHG
1163	CPMIDEFPT
1164	CPTWSCQQW
1165	CPWQIKVAA
1166	CQMEQHYSP
1167	CSAPLRHWY
1168	CSIFGPTYP
1169	CSKFIVDWE
1170	CSLRYFMAR
1171	CSRTMEVGL
1172	CSSVRWMSP
1173	CSSYSFQHS
1174	CTFMANWGF
1175	CTGQRGFAN
1176	CTMDQWQTA
1177	CTTHFYRAR
1178	CVAPVRLQD
1179	CVAYRLVNP
1180	CVFMRLTSG
1181	CVFVRNNMM
1182	CVGSLMRIQ
1183	CVHLTSSSA
1184	CVKQYMQPP
1185	CVRHTENWC
1186	CVTTRHKQC
1187	CVWSAPVEC
1188	CVYQSRFSC
1189	CWNENTYEA
1190	CWNGIQRQA
1191	CYTRFHYSE
1192	CYVFRHHYC
1193	HVASTSYVT
1194	HVFDTDRWM
1195	HVFHLDEAN
1196	HVKDTDVQI
1197	HVQARQNPG
1198	HVSDKNHAK
1199	HVVPQNWNW
1200	HVVQRLTIP
1201	IVEARLEFN
1202	IVEDFEIVI

[1834]

1203	IVFYRFEFV	1244	NVYPDSMHG	1285	TVEHKMDAP	1326	DVASYEQAY
1204	IVLPTMLLA	1245	PVNFFEHPM	1286	TVISDYEDC	1327	DVEWQQSWT
1205	IVVIAKMTT	1246	QVDIDTQNA	1287	TVKMLQHWT	1328	DVGGLNREY
1206	IVVSTWHEQ	1247	QVDPHGNMH	1288	TVMKEQANM	1329	DVHEVMSCG
1207	IVWLQRVGD	1248	QVDRWCTAA	1289	TVNDHGPCI	1330	DVHICTVHI
1208	KVDHTINQM	1249	QVEYAHSRA	1290	TVNKWHQTA	1331	DVKGMLDGF
1209	KVDTEKNMM	1250	QVFMQHDIE	1291	TVNQCMTHC	1332	DVKMQQKYF
1210	KVTGHEQDV	1251	QVGNPLKEF	1292	TVYFVWDSQ	1333	DVMGHAKLA
1211	KVVLWEHFQ	1252	QVIEWGRSG	1293	VVGATATGM	1334	DVRAYLNKE
1212	LVCTHTKAA	1253	QVIIRKWFA	1294	VVGPPGHHY	1335	DVRYNQNAW
1213	LVDDDYQMD	1254	QVLMAQFSR	1295	VVMFWQEGD	1336	DVSFMGWDC
1214	LVEIPCSNW	1255	QVQKCMQYV	1296	VVMQQYRSH	1337	DVTHHFAKQ
1215	LVHSCQHTE	1256	QVQKGACPS	1297	VVNKDKWRI	1338	EVAHFLANE
1216	LVLQSFHHS	1257	QVQSGPYWG	1298	VVQWPTSLL	1339	EVDNMQMGE
1217	LVNCYTYRA	1258	QVQTENMCD	1299	VVSAAKSSG	1340	EVGVSAPMM
1218	LVQVMDMKP	1259	QVRCGIKTG	1300	VVSHKQRCS	1341	EVSGMFHPA
1219	LVVIFEGRH	1260	QVTFHAYVC	1301	VVSKYQIVP	1342	EVTTMLKAC
1220	LVWMDPHWG	1261	QVTIQNDYD	1302	VVTCRHQGS	1343	EVWINMAPE
1221	MVAAKEKHD	1262	QVWNPHLCM	1303	VVVPMSNCE	1344	EVYMEPANW
1222	MVAHNGLGF	1263	RVATEHFRE	1304	WVMCISGHY	1345	GVFFRYEHG
1223	MVCMLVHHC	1264	RVDNPFVSY	1305	YVMFNTWRE	1346	GVFSTGVVT
1224	MVDHSYQQE	1265	RVSECTENY	1306	YVTMSDERP	1347	GVGVGLGTQ
1225	MVEVIEGGQ	1266	RVTEEVGSR	1307	AVGFMECAN	1348	GVHQESGQN
1226	MVHGNHECW	1267	SVDQHDHLL	1308	DVKDRNELH	1349	GVNKLSYEN
1227	MVNPMHCCN	1268	SVERYPLTA	1309	EVSAEEGQN	1350	GVRIPMAAD
1228	MVRKYEKMH	1269	SVFDVVHEE	1310	GVMPRWGLF	1351	GVSQFMAET
1229	MVSCNQHDS	1270	SVHVRIVEP	1311	GVVNLVKMQ	1352	HVGAHTRSD
1230	MVSHVVASE	1271	SVKQHHVDW	1312	IVMYDHFNQ	1353	HVGKMPGGS
1231	MVSVDQMSV	1272	SVLGMPAYR	1313	KVILDSFRQ	1354	HVNSFQPDR
1232	MVWSRADWS	1273	SVMDANTWR	1314	KVMPEPVTE	1355	IVCTTFEDS
1233	NVAIILSTD	1274	SVNKDVRHC	1315	LVHLFQYHL	1356	IVDQTCMHA
1234	NVAVEHNMG	1275	SVNSYMDSV	1316	QVVACMHHT	1357	IVGIWQWMG
1235	NVEDTRKGS	1276	SVPRKSKCS	1317	RVDYSAIHM	1358	IVKNLEQYM
1236	NVFLAHQRA	1277	SVRSYHSQP	1318	SVIMHHDQH	1359	IVKTLICYN
1237	NVHNSGRFD	1278	SVTQPLYCQ	1319	VVEMNGDDD	1360	KVDVITVKD
1238	NVIGTENMS	1279	SVVKRNDDV	1320	VVEYAPWAW	1361	KVGAHNDEV
1239	NVKSEHRES	1280	TVAHFLNGG	1321	AVAKGVGVT	1362	KVTWLDLED
1240	NVLLEKAFW	1281	TVALSEQYR	1322	AVMATGVTQ	1363	LVEAVMNAT
1241	NVNRGHLLE	1282	TVAVRYETL	1323	AVSGNWMMT	1364	LVKACTDQH
1242	NVNWQARND	1283	TVCSTIVDW	1324	AVSKSGCGM	1365	LVKQYGLIH
1243	NVTHKQGCS	1284	TVDWRWGDF	1325	DVAAVQVLE	1366	LVLMDRQAY

[1835]

1367	LVSQTPRKV	1408	YVIVHMKGM	1449	SEMKLVNLP	1490	EAMRAMTPI
1368	LVVDHAEWS	1409	YVTPNIAVH	1450	SEMVKHNFE	1491	EAMYWQCAI
1369	MVALTDQLG	1410	HCMRQSESG	1451	SGMKQVGAA	1492	EIMWEGVGS
1370	MVDSNACLC	1411	HSMPEAVAY	1452	SGMLVGIYN	1493	ENMHKEYVQ
1371	MVHPSQNVD	1412	HTMEILPGL	1453	SHMVSNSAV	1494	EPMEHYLRS
1372	MVKEAPYYR	1413	HTMHDSCFN	1454	SSMSKHFVA	1495	GAMTMCCEW
1373	MVSEYATNA	1414	HTMLWAKDE	1455	STMPGTDQK	1496	GKMMGLHSH
1374	MVTGISKSS	1415	ICMSYLSWD	1456	TAMHQTMLN	1497	ILMEYDHDK
1375	NVESNYNLI	1416	IPMFSIHGN	1457	TAMMSGHLV	1498	IMMHRVMSD
1376	NVEWGTVQS	1417	KAMLEGPYL	1458	TAMVVEYDQ	1499	KAMEVHFQK
1377	NVFMTSFQC	1418	KNMINEAAP	1459	TCMKNINYQ	1500	KAMGLPSWY
1378	NVFQPEDHC	1419	KSMWEHAIF	1460	TCMLGAEAA	1501	KSMWEDRIC
1379	NVTWEDVQE	1420	KTMEDVIHA	1461	TCMQLATMQ	1502	LAMPWDTLQ
1380	PVAAAEPRN	1421	KYMSYERES	1462	TCMTLRHIG	1503	LCMLRMMGH
1381	QVAEPGGRF	1422	LAMTLAVGQ	1463	TCMYYWKQP	1504	LIMAHWLDM
1382	QVDSWPVLS	1423	LCMDYALVD	1464	THMLCPRMD	1505	LIMDSKMGQ
1383	QVGCTVGPG	1424	LCMLLASQQ	1465	TNMSYQRSD	1506	LLMPVLHQM
1384	RVQHSGAHW	1425	LSMCFRDDY	1466	TRMPVNQSC	1507	MGMKREGSG
1385	RVTNTWECM	1426	MCMVERRGS	1467	TWMEWHANA	1508	MQMSKNKHP
1386	RVYTTGKCS	1427	MFMSQHVED	1468	TYMGTHNHS	1509	MRMIHWMHC
1387	SVHHSRWGY	1428	MNMHTDCTT	1469	VFMPDVYDK	1510	NKMEWHKYC
1388	SVNQAMARK	1429	NAMWTTEYE	1470	VKMPYHYEG	1511	NWMQNTEGH
1389	SVSVMTYSE	1430	NEMKFSYSH	1471	VQMQEVPHY	1512	PEMSFEFTT
1390	TVAMDPDYA	1431	NIMNWQNGD	1472	WFMHYSWKH	1513	PGMNMVRKE
1391	TVATEPAQQ	1432	NLMPTNHAM	1473	WKMEGDCFK	1514	PTMWWYLLV
1392	TVDFFVMDC	1433	NSMQHCVCP	1474	EIMCAYVKA	1515	QCMIVSGMT
1393	TVDHCVGTV	1434	NTMHPVGMM	1475	GAMPNPYHE	1516	QGMHAMFYA
1394	TVGLNERNG	1435	NTMMVGRCE	1476	GTMWSTKSC	1517	QMMQLYCGA
1395	TVHWWECVQ	1436	PIMFEGECK	1477	HPMFMIPGY	1518	SAMVHSNKS
1396	TVQTNFTSK	1437	QCMMKSSYE	1478	HYMLKEYFA	1519	SHMSYLMSH
1397	TVTQEWMMK	1438	QFMPYLGFA	1479	IHMNTIVGE	1520	TCMPNIQCP
1398	TVVGLITNA	1439	QGMENQWHQ	1480	IMMAQWFAH	1521	TFMHQFGTY
1399	VVEKRMDRS	1440	QHMYANCLP	1481	PSMPNIKPS	1522	TQMMTPVGA
1400	VVHGYVCKQ	1441	QKMTRQHWG	1482	PYMQYWYWE	1523	TWMADWHQE
1401	VVHPLAYCT	1442	QSMEQEFVT	1483	REMLRDKMP	1524	VAMQSMQEH
1402	VVNKQQISE	1443	QSMMALKDY	1484	SAMQVMTMC	1525	VAMQTHQCY
1403	VVTSPDAFQ	1444	QSMMTDRTY	1485	TAMSICMYG	1526	WWMHGRVGR
1404	VVVHTQCNP	1445	REMTKVGGR	1486	TFMDKGLIR	1527	YCMLKHKAG
1405	WVTPTNHKP	1446	RGMYRCGWD	1487	ANMFRNGMQ	1528	YCMWQQTSN
1406	YVAEYGLGA	1447	RNMYISQWE	1488	ASMFWNYQV	1529	YKMDHNNTY
1407	YVEERGSQR	1448	RSMGNEQDP	1489	DRMVIKSAE	1530	YSMTWNQQT

[1836]

1531	HTGKADVSV	1572	WYQKQAEHG	1613	KCLMVDWGV	1654	SQNMVDSQC
1532	IMEKLRRSS	1573	YCSKSVRCC	1614	KHGQVSIHD	1655	SSCIVNIAS
1533	IPFKAVMVG	1574	NFSKCAKYD	1615	KMDRVGYSA	1656	SSNLVGAGE
1534	IRVKPGCRV	1575	VCTKGFLNS	1616	KWNNVENNE	1657	STELVKYID
1535	ITGKYPMTA	1576	AAHKEMKDT	1617	KYSPVEWMY	1658	TMHPVCCPM
1536	KLEKTVVYS	1577	ACGKMHGPN	1618	LGQHVQCWQ	1659	TMREVQCEW
1537	LDIKMDFLC	1578	AFLKQSIIH	1619	LKSWVCATS	1660	TPAHVCRGY
1538	LMHKDQVSP	1579	AGSKRFDWF	1620	LMKCVMCRD	1661	VNDMVLAVV
1539	LNVKTGNAW	1580	AILKMTKVL	1621	MASPVNSEF	1662	VSFCVRKQS
1540	LTTKQPDMP	1581	DTGKFMCKW	1622	MCLFVRVSQ	1663	YALTVPHSE
1541	MASKLHMAV	1582	EHQKNINMH	1623	MHGMVHAFE	1664	YTPNVHYMV
1542	NSLKYEFSE	1583	ESEKVWMMT	1624	MLIHVNHKS	1665	ECANVIKFN
1543	NTGKTGQFN	1584	EWRKKKKRT	1625	MQDQVCKIG	1666	GACTVQNMT
1544	PCAKSYDQS	1585	FCRKIEQAN	1626	MQQPVNKHM	1667	HIEWVSMHR
1545	PKDKGVICE	1586	FICKTLHQH	1627	MRQTVDRPE	1668	MGQNVNHAG
1546	QEVKGQWWP	1587	GCDKTRFEN	1628	MYHAVQCMP	1669	QAAFVNYGP
1547	QIFKQTEVQ	1588	GILKFRDFM	1629	MYYQVSKQD	1670	RGFHVIVDQ
1548	QITKPILDR	1589	ICWKFLVSN	1630	NAENVDNGV	1671	RYDLVSYHQ
1549	QLAKAMTQS	1590	IEVKYWNHH	1631	NFEFVERYR	1672	VFDTVFAYC
1550	QPVKYPQTN	1591	IWYKWWVMF	1632	NGHEVPVEE	1673	VKKYVDVHF
1551	RISKDGSDR	1592	KCDKPMLKP	1633	NKQVVQMTA	1674	VQSAVSMCS
1552	RIVKRYHNG	1593	KMEKVERMP	1634	NLPLVHSLE	1675	APVFVEQMG
1553	SLCKMLTAK	1594	LDLKHNMCD	1635	NLQRVEMGV	1676	ASTAVLVQR
1554	SYKKPVKYM	1595	MCIKCDAND	1636	NMQNVARCY	1677	DAGFVREPV
1555	SYNKVMTQH	1596	NCNKCKNRT	1637	NSWDVGSFN	1678	EAVPVESGD
1556	TACKTEGGP	1597	NQCKINVGT	1638	NTYPVLSDW	1679	EEKEVICMC
1557	TCVKTTHGA	1598	QQLKTNINF	1639	PEYFVVYMY	1680	EFEQVGFMM
1558	TDGKTRMQQ	1599	RIEKLQCDY	1640	PMVSVSGYY	1681	EGTMVKQGE
1559	TETKWQHLE	1600	STLKTYAYN	1641	QGGAVMHEE	1682	EHNCVYLTE
1560	TIHKYCHHQ	1601	TGTKITMQT	1642	QIQPVNALF	1683	EMHYVCCQT
1561	TSTKHHWWC	1602	TIHKDGEFS	1643	QQLSVHMWH	1684	EPFRVTQCG
1562	TSVKKPHLA	1603	VLQKFGHSF	1644	QSGWVQIAA	1685	FASIVFGGE
1563	TTLKCTNTW	1604	VTGKHIAQN	1645	QWQYVMTDK	1686	FEYMVLNWH
1564	TWIKPHQQL	1605	YCYKTAVAA	1646	RCWSVEASQ	1687	FQAPVTKAS
1565	VCEKSNDYA	1606	HCALVNRLR	1647	RDQRVLKVM	1688	GLGYVHRSN
1566	VCFKMLLGN	1607	HFAMVWAGE	1648	RTTGVGDIS	1689	GWQSVWNAP
1567	VECKTMAHQ	1608	HIGYVEAMR	1649	SAWCVPGQL	1690	GYNWVQEQR
1568	VNAKWQTAS	1609	HTKPVNAQP	1650	SCHLVIKHH	1691	HAVMVYSDY
1569	VSCKIYHYA	1610	IHQYVNFVH	1651	SFAPVAHIE	1692	HTKGVVHQP
1570	VSGKECSED	1611	INNTVAHQE	1652	SGRHVLIEI	1693	IARHVDHDL
1571	VTYKYYCDF	1612	KALMVQQFD	1653	SITAVWSDE	1694	IQTEVHTVY

[1837]

1695	KQQDVTMNA	1736	NSTQPNLWF	1777	WCNYLNQGM	1818	NNQWMNSLE
1696	KYIDVWDAR	1737	NYTIYNPAE	1778	WQCTENWQN	1819	NQIRNNMER
1697	LGVDVLADY	1738	PIETTNHMD	1779	FYWAHNICE	1820	NTYMHNREG
1698	MQVLVSMTP	1739	PMYANNYFC	1780	KYEILNCNH	1821	NWNDSNDCC
1699	NTIDVQLGG	1740	PTQPLNQHG	1781	LYAMLNKFF	1822	QCYTFNDTK
1700	QKAYVSVSG	1741	QAWCANTFH	1782	MGNYQNTHY	1823	QRDLANGQF
1701	QSPAVNCAW	1742	QCFYANRGD	1783	RKDEANGQT	1824	QTASWNMEM
1702	RIEQVDVFN	1743	QHQQFNLTT	1784	YFIIQNHLD	1825	RCEPFNWRE
1703	RIGGVADAD	1744	QIDMMNCNR	1785	YTRSNNWDY	1826	RMAQINQQG
1704	RWECVTGGY	1745	QQLPENLYQ	1786	AAFMPNNYY	1827	STALMNEYQ
1705	TCIRVWQWA	1746	QTDVINKFP	1787	AALFPNWSN	1828	SWWLTNTAV
1706	TMRYVNTCV	1747	QTGMENHQN	1788	AEECSNHNV	1829	TADFYNDQM
1707	TSHNVANFG	1748	QYSWSNINT	1789	AHSMTNLVG	1830	TASSANLSW
1708	VEQMVMPPH	1749	RIEANNWGV	1790	AWFDHNEHQ	1831	TFEICNSPQ
1709	VYTVVDEHS	1750	RIICQNAVA	1791	AYARSNRQD	1832	TFLRLNDFV
1710	WFPVVWSHN	1751	SEKVINASC	1792	DATIANAQV	1833	TFSQHNHKA
1711	YHVLVEGDR	1752	SIAVKNMCH	1793	DIAEKNRMM	1834	TKCTTNQQC
1712	HGIALNLGT	1753	SKRDNNCEV	1794	DWKTSNAEY	1835	TSALNNNSK
1713	HRDYANSWN	1754	SMQRYNSKP	1795	EGDWYNYGT	1836	TSHDNNAQI
1714	IGTRFNAQT	1755	SSLCRNMWP	1796	GIIFSNTCM	1837	TTFHINEYP
1715	ILDDCNCKC	1756	STYGRNLYP	1797	GPPPYNAWF	1838	VCVPKNGIH
1716	IPRMLNHVE	1757	SWSNINHQT	1798	GSQRTNTSC	1839	VGAQTNCLQ
1717	ISYQPNPME	1758	TCGEYNHRC	1799	GTSRYNAFI	1840	VYNLLNHYQ
1718	KASELNQVL	1759	TFQEKNVMH	1800	IALLTNDLA	1841	VYRIANTQY
1719	KHLDCNERE	1760	THQMTNTYS	1801	ICGRLNNNP	1842	YAKSRNIVG
1720	KLSNMNHPI	1761	TIAQDNKEA	1802	INAPQNMSW	1843	HKESLGHHC
1721	KMEQGNMDS	1762	TLITTNRQY	1803	ISGYHNVDE	1844	HRHPCGHIP
1722	KMFHSNQEA	1763	TMVQNNMCF	1804	IYHTLNHCM	1845	HSAPSFHTQ
1723	KNWVCNLLT	1764	TRQPFNMSL	1805	KAEMSNAIT	1846	IDKLYWHWG
1724	KTEVTNGPT	1765	TSSEFNRQR	1806	KAIPENSHR	1847	IDQPIQHAS
1725	LCIRKNYTH	1766	TSVACNSTD	1807	KIHSPNQDE	1848	IGKMNDHGM
1726	LCQFNNATD	1767	TTVFWNNFK	1808	LKSPHNENG	1849	IIQYNGHQA
1727	LEAECNSMQ	1768	TWVPDNADS	1809	LLEVKNHDH	1850	IKTELAHQY
1728	LGDERNCGQ	1769	TYQVNNWSD	1810	LSDLQNKGA	1851	KAAGDSHTP
1729	LHDNWNRYT	1770	VANHHNEIQ	1811	LSHTSNMGH	1852	KADAPIHWS
1730	LPFRSNNAQ	1771	VIGTANAAP	1812	LTYLANQSD	1853	KAECKPHHE
1731	LTVQNNGIR	1772	VMGMINGAW	1813	MGKSRNGSE	1854	KATELGHCG
1732	MCAYMNKKE	1773	VPLVINKEQ	1814	MISMWNRSN	1855	KCTAMRHKE
1733	MPNLSNSFQ	1774	VSSQFNGPS	1815	MMDPTNFGS	1856	KTEQATHTC
1734	NGAEMNWHE	1775	VTHCKNVAC	1816	MNNITNWNP	1857	KTSEWHHVG
1735	NGNQFNNCT	1776	VWYPQNAVE	1817	MTGRANKFD	1858	LATSRVHFG

[1838]

1859	LGGWMEHSF	1900	DSFLNRHTE	1941	HSQIFMLSW	1982	TKENLAASY
1860	LSDHLHHYT	1901	DTCSRKHWM	1942	IGSQIFTST	1983	TMLVHHGSM
1861	MCHHRHHAE	1902	ESFYQRHVT	1943	IGSSIMSST	1984	TSFLLGVSS
1862	MFRERIHHD	1903	ESVVGPHWD	1944	IIHQQHNST	1985	TTHDEGVSI
1863	NCIFIEHPT	1904	ETDVCRHGV	1945	ITSRRGMSE	1986	VAAHPLTSD
1864	NKVGCQHSY	1905	HQNASMHMP	1946	KAEPLRLSP	1987	VFRMATESE
1865	NMKVDAHSN	1906	HYIQIQHDY	1947	KAFIRDESY	1988	VGQVIDRSW
1866	NSSYEEHDQ	1907	IMVALMHED	1948	KGDWTVKSY	1989	VTTRNPFST
1867	NTCFIGHNE	1908	IYNQFPHYG	1949	KGGHKDESP	1990	VYQVCTISD
1868	PGSQYVHWI	1909	KHSPDHHFE	1950	KIWAHMCSP	1991	WCQRFQTSS
1869	QASYRTHHD	1910	KMAMLRHAT	1951	KMFLLEQSD	1992	WITEIELSG
1870	QCAIMRHYC	1911	KSVGWCHAF	1952	LASCNWRSQ	1993	YFALSGFSS
1871	QIAEIIHFG	1912	LQQPIVHIE	1953	LAYPCYASA	1994	EIHHLPSSH
1872	QIYPEHHKE	1913	LTETEHHDQ	1954	LFSFLWRSA	1995	GTAETYISA
1873	QRKHYVHVF	1914	MATHCLHWR	1955	LIDQYTNST	1996	HSEQMHQSD
1874	QWDTLRHTP	1915	MMPEIMHVL	1956	LIEQTCVSN	1997	LAKSCSSSQ
1875	RICFHQHWG	1916	NMGLLYHTN	1957	LKVDFLLSN	1998	NASLDMISV
1876	RKECEDHGG	1917	NTEPMTHWN	1958	LMEQFDYSQ	1999	SSEPFTKSN
1877	RTLEWFHRG	1918	PMKPMHHGY	1959	LQGEQSASD	2000	VADMGPWSG
1878	SFSWTYHQD	1919	QCTFYFHWR	1960	LYEQYKASY	2001	APHSYAMSY
1879	TAVMYAHDC	1920	QHVQEIHGA	1961	MQQYAYISH	2002	ASHIKWNSH
1880	TREINDHGI	1921	QIYHNHHTA	1962	MSIRHMISY	2003	DAQQAPNSS
1881	TRGSMYHHC	1922	QSCVGLHQI	1963	NCGVSMKSN	2004	EDWIMHASF
1882	TTVRNFHAE	1923	SATSKSHTF	1964	NCHSKCYSS	2005	EPSFESDSN
1883	TYCCYHHQG	1924	SCDEYMHHP	1965	NTAHCWESD	2006	ESKWWIQSD
1884	TYECAAHYP	1925	SCVLKVHFP	1966	QGSFAHVSQ	2007	FGKDTFVSS
1885	VAQPYIHSH	1926	SLCMHRHDF	1967	QHKVMGRSW	2008	GANNIYVSM
1886	VNKPAGHIQ	1927	SMGPTVHQM	1968	QTRCDFLSH	2009	GKSLMMTSF
1887	VRDLLHHDM	1928	TAGHEGHIC	1969	RADHYRASK	2010	GNTQQPMSF
1888	VYETLKHRG	1929	TCEWWIHMD	1970	RCEAEGGSK	2011	GTAAMEQSC
1889	VYETMGHEE	1930	TFGWFMHIP	1971	RFHVMETSM	2012	IDASLDMSW
1890	WYYYPAHHK	1931	TFSQHSHKA	1972	RSNIEHLSE	2013	ISEPCYVSI
1891	YEVTIPHAI	1932	TRQQLFHTQ	1973	RTDSFYESD	2014	LMNSYCVSH
1892	ETRWWPHIE	1933	TTGSTKHIG	1974	RTIDPVQSS	2015	LNVEAELSF
1893	MSCMKEHSP	1934	VMPCYWHTT	1975	RTVEHQASA	2016	LSLPGVPSM
1894	NCAPMEHHV	1935	WDGTKVHWL	1976	SLGCREISL	2017	NMTGIHVSP
1895	QGLNYEHQE	1936	YHVILVHQA	1977	SLHAWAISA	2018	NTFATGNSP
1896	SCRHIYHEA	1937	YPTVFEHSE	1978	SNFEFQKSH	2019	QCYESMTSW
1897	SMNQAIHLE	1938	HIGYMPRSY	1979	TCGLFSGSY	2020	QICQTDASA
1898	AAYMTDHMN	1939	HLCPNQFSP	1980	TFTRPEQSQ	2021	SCVFRLASG
1899	ASTSIVHAH	1940	HSNGLGPSC	1981	TGKWISPSN	2022	SDPHAGPSD

[1839]

2023	SLWILRMSN	2047	LHEFHKSYP	2071	STNLIGQNP	2095	GMFWLTQAP
2024	SSCGNMVSN	2048	LQLRSDWCP	2072	TCGWSHAMP	2096	GWWVATMLP
2025	THWQYEASH	2049	LTIHSQREP	2073	TGNLWQSMP	2097	HAELTMNVP
2026	TNAQFSNSW	2050	MACLFRACP	2074	TSQAPMCNP	2098	IACCSEAIP
2027	TTGVSEGSA	2051	MEGAIGQNP	2075	YCDRSGIWP	2099	KGDYQADAP
2028	VANYHPYSM	2052	MGANIAQMP	2076	YFYEAKLMP	2100	LEKSMMTAP
2029	VIRYGDLSQ	2053	MGSSRMIYP	2077	YNGYTSVCP	2101	LRYLSDADP
2030	VMGPQPVSV	2054	MHIFHSLCP	2078	YRDRFFAVP	2102	MCRANGDMP
2031	VSVMLAASD	2055	MKSECMSGP	2079	YSHPRDGAP	2103	MCTLRHWQP
2032	HFYSQSTLP	2056	MWGHISEMP	2080	ACQLLHWQP	2104	MFELFVMGP
2033	HGAIHAVLP	2057	NFHMKSMCP	2081	DWDSMCWWP	2105	MSTPNMTQP
2034	HHECMWMFP	2058	NLHDFIFAP	2082	EDTMAQMWP	2106	NAEQWCCGP
2035	HTEARMVEP	2059	NWNLWDTYP	2083	LCHFRLVGP	2107	NMTPLMTLP
2036	IGNYYHKWP	2060	PEFEFFTTP	2084	NQTGHMPAP	2108	QAQPNVNYP
2037	IIRHHEFAP	2061	QCDEQEYNP	2085	QAPAYMWGP	2109	QAYTDDTGP
2038	IIWQQDIGP	2062	QGQRFYNYP	2086	WHSRKQWPP	2110	QWGGNQQTP
2039	IQSECVVVP	2063	QYARLDCLP	2087	YTKYLMTEP	2111	QYEEMAWGP
2040	KNESAKQCP	2064	RADIGSLCP	2088	AFATCSMMP	2112	SAYMLRYQP
2041	KSVNLCVDP	2065	RCILKREQP	2089	APVCGQDPP	2113	SGQQFIFCP
2042	KTSFSEMPP	2066	RETVAMDMP	2090	DADERQILP	2114	SPGHNMVGP
2043	KTSMTFYNP	2067	RFEFLMAPP	2091	DERNSFWFP	2115	SSVHTLYFP
2044	LARSDCRQP	2068	SAYIGEQGP	2092	DIHTNTIEP	2116	SYNYEHAWP
2045	LFYPEWDGP	2069	SLRVCIGQP	2093	DSGWWATMP	2117	TCNMQRAVP
2046	LGCDTEQMP	2070	STHLKLSDP	2094	GLTMNRCYP		

[1840] 实施例8

[1841] 在坐骨神经中具有组织嗜性的AAV5变体

[1842] 此实施例描述了使用实施例1-实施例2中描述的方法和系统发现的在坐骨神经中具有组织嗜性的工程化AAV5变体。

[1843] 分析CNS、肝脏、骨骼肌、心脏、肺、脾脏、淋巴结、骨髓、乳腺、皮肤、肾上腺、甲状腺、结肠、坐骨神经和脊髓组织以鉴定AAV5VP1 581-589区域中驱动坐骨神经嗜性的氨基酸残基。结果显示于图18N中，该图显示了相较于所有其他分析组织在坐骨神经组织中AAV5 VP1的581至589区域中每个氨基酸的位置频率，其中衣壳多肽具有SEQ ID NO:2的序列。位置(“Xaan”，n＝1-9)在x轴上表示并且氨基酸残基在y轴上表示。特定位置的特定氨基酸残基的频率在热图中作为强度表示。较深的方块表示特定氨基酸在特定位置的频率较高，较浅的方块表示特定氨基酸在特定位置的频率较低。

[1844] 分析相对于给定氨基酸残基出现在所有其他所收获组织中鉴定的变体中的指定位置的频率而言该给定氨基酸残基出现在坐骨神经中鉴定的变体中对应于SEQ ID NO:1(推及于SEQ ID NO:2)的位置581至位置589的所述指定位置的频率，以确定优先靶向坐骨神经组织的衣壳的一组序列规则。参考下文表15，和SEQ ID NO:2(AAV5VP1)，以下氨基酸由此可以在Xaa1-Xaa9任何一个或多个位置处以任何组合独立地突变，以提供与野生型AAV5 VP1衣壳(SEQ ID NO:1)相比具有增强的坐骨神经嗜性的VP1衣壳，其中坐骨神经嗜性在此是指优选坐骨神经转导的特性，而不是优选转导所有其他所收获组织的特性。

[1845]

表 15 坐骨神经嗜性规则
Xaa1 选自 C、G、K、M、Q、R 或 Y。
Xaa1 选自 C、R 或 Q。
Xaa1 为 C。
Xaa2 选自 A、C、F、I、Q、T 或 V。
Xaa2 选自 A、C 或 I。
Xaa2 为 A。
Xaa3 选自 A、F、I、M、R、S 或 T。
Xaa3 选自 F、M、R 或 S。
Xaa3 为 R。
Xaa4 选自 E、N、T、Q 或 V。
Xaa4 选自 E、T 或 V。
Xaa4 为 T。
Xaa5 选自 F、H、Q、S、V 或 Y。
Xaa5 选自 F、V 或 Y。
Xaa5 为 V。
Xaa6 选自 K、M、N、Q、S 或 V。
Xaa6 选自 M、N 或 S。
Xaa6 为 N。
Xaa7 选自 K、M、Q、R 或 T。

[1846]

Xaa7 选自 M、Q 或 T。
Xaa7 为 M。
Xaa8 选自 A、G、H、Q、S 或 V。
Xaa8 选自 H 或 S。
Xaa8 为 H。
Xaa9 选自 C、E、I、K 或 R。
Xaa9 选自 C、I 或 K。
Xaa9 为 I。

[1847] 下文表16提供了在坐骨神经组织中发现并且符合表15中提供的一种或多种规则的变体AAV VP1衣壳多肽的581至589区域的序列。因此,本公开提供了由具有SEQ ID NO:2的工程化AAV5 VP1衣壳多肽组成的rAAV,其中Xaa1至Xaa9区域是SEQ ID NO:26118-SEQ ID NO:26990中的任一者,如表16中所公开。本文还涵盖由工程化AAV5 VP2衣壳多肽和工程化AAV5 VP3衣壳多肽组成的rAAV,所述多肽在对应于AAV5 VP1 581至589区域中的氨基酸的AAV5VP2(氨基酸残基445至453)和AAV5 VP3(氨基酸残基389-397)中的区域处具有下文表中公开的序列。

[1848] 表16

[1849] 驱动坐骨神经组织嗜性的AAV5 VP1衣壳多肽中581至589区域的序列

[1850]

SEQ ID NO	581-589 序列
26118	CQGYEHSQE
26119	CAHAMKWGS
26120	CGGQVEFGW
26121	CSSCYHRVF
26122	CCCEALMPQ
26123	CENPLCQSS
26124	CFGPEGYAG
26125	CHPQHHHSH
26126	CSSFYHRVF
26127	CKEPRKMRE
26128	CPVHQMDFT
26129	CIVKSNGSK
26130	CPGLQFDTP
26131	CHHFNMPSQ
26132	CCRQIWVQY
26133	CIIARSTAA
26134	CAEKQEYVQ
26135	CAHMQYAGP
26136	CANPLWAQT
26137	CAVKTHTCT
26138	CCNRNWNQQ
26139	CDYMHKLNE
26140	CEGRHTWGK
26141	CFHRHSTVP
26142	CFIAADNRN
26143	CFKHQMMMH
26144	CFSDTDNAM
26145	CFTNHWEMG
26146	CFVPMFNHN
26147	CFWLDWWCW
26148	CGYMHKLNE
26149	CHSIFNVKC
26150	CISAGMKND
26151	CKDLRCGCA
26152	CLGQMWAQW
26153	CLHQWLMPT
26154	CLIAPPHMD
26155	CLTSTSHCI
26156	CMCPGKEYP
26157	CMLAWQNDF
26158	CNPKCDAMM
26159	CNYDSHKMK
26160	CPLFLPALP
26161	CQNMKRIDW
26162	CQQPPRPEK
26163	CRIQSIHVH
26164	CSAGDEWGG
26165	CSKHATMAH
26166	CSKHVTMAH
26167	CSMLFHHTQ
26168	CTILNHQDA
26169	CTRPDEIKE
26170	CVAICQDGL
26171	CVHTISNCH
26172	CVTWTNRKL
26173	CVWSEDAGY
26174	CWCEKSHLG
26175	CWLCQCPRQ
26176	CYDHFQMKD
26177	CYQFHTGFW
26178	CHAAVANLY
26179	CTMCHDCSP
26180	CADPQERCD
26181	CMEVTCFVV
26182	CIGPFHCAN
26183	CWGPIMHPF
26184	CSESYRGPD
26185	CQQHFMFAP
26186	CGHLLRRAE
26187	CRRDGYAAM
26188	CPEHNQVGV
26189	CIHRFVSSP
26190	CDWMANVSV
26191	CCRMDNKKS
26192	CLNNSTCRG
26193	CDHVTKHPT
26194	CKYAFKMEK
26195	CDAEAEFRR
26196	CPFKLYSKG
26197	CSIQINNCR
26198	CMVSTQKLA
26199	CFCTYKIEH
26200	CVCWFWFFW
26201	CHMLNDCMS
26202	CMAHIWEGH
26203	CGCQVKCHT
26204	CSMPQQCME
26205	CWECCTDSG
26206	CIMLNGCFQ
26207	CIDVLFKHL
26208	CQMLCRPPV
26209	CAGIHLNEA
26210	CYMNFIVDG
26211	CYIRNMVNL
26212	CPKQANRSG
26213	CVLSRSTMR
26214	CEAGWCRSD
26215	CIQPTWDSY
26216	CYWPMTNCT
26217	CNGYWMCMS
26218	CSVWMQRDR
26219	CWILHHRCS
26220	CYNQFHTDN
26221	YAVMSPPAG
26222	YAVMSQPAG
26223	YAVMSRPAG
26224	YAVVSHPAG
26225	RAVEMQPCT
26226	IAEEGWACP
26227	QAPLATAIE
26228	FAPEFDNHC
26229	SAHMPKSVT
26230	EAYLYNIPM
26231	NAEQCMIGA
26232	RASMCDCFE
26233	SACPLWCGS
26234	DAMHKWIFA
26235	HAPCEQWAW
26236	MASIDKQHV
26237	AAQWPMEYT
26238	TAEWKLCGE
26239	QAEGGPNMP
26240	IATEHPDEW
26241	NALDEDCQH
26242	AAMLYNACM
26243	VADWIGHTP
26244	VAIQCNRSY
26245	GATTLTKCL
26246	YASTMHSIP
26247	QALVGNAAE
26248	SAICQNYDA
26249	DAHNDALAR
26250	AADECMLHT
26251	AAEKLFAWE
26252	AAFLGQLSV
26253	AANSRENDW
26254	AATNMSNVH

[1851]

26255	DAEANGCDA	26296	PAYYGPGSE	26337	KAGYQEGSC	26378	MAHAPPNSD
26256	DAGKLGQHD	26297	QAAPPCAMS	26338	AAMWWQWDR	26379	MAMWAVWNK
26257	DAGWMNSWV	26298	QAIRGTHDQ	26339	GAYSLHDWR	26380	MAEQMERDF
26258	DALPMHGQR	26299	QAMRFSENG	26340	HAETDHVGL	26381	SATNCDLYP
26259	DAMCSIRKD	26300	QATAEYACD	26341	FAQQDKQDR	26382	HAKQANGMV
26260	DARANCDAA	26301	QAWMPVCHT	26342	YAYHPQCKP	26383	AADIFAIND
26261	EAEIAYGFM	26302	RAIMSDWEC	26343	QATGMHRWQ	26384	TATCGIEVE
26262	EAGCINVPS	26303	RALECNWFL	26344	IADGFAAVA	26385	AAIGNQTLA
26263	EAQLYSTKA	26304	RALTFDCHK	26345	QAGNSMWAK	26386	AAQYMGIMK
26264	EASTYFIWD	26305	SAEYSNDES	26346	MAMWIADTH	26387	KAPWAPMDS
26265	FADTCPRHA	26306	SAHTFCAPR	26347	MATCYMYAD	26388	NAENESSIG
26266	FAELHGMQG	26307	SAHVVLGEM	26348	IAKIGNCVW	26389	QAKMILDGT
26267	FAHDSCESG	26308	SALENVPKY	26349	NANYWFYYD	26390	DMRKTWEAC
26268	FANWLKSFP	26309	SAMTSSHQK	26350	EASWYYPAI	26391	TGRIKDAGP
26269	FAQIIVHEH	26310	SAMYTENIS	26351	KACMSHQGR	26392	PWRCMQGYY
26270	FATGCKQGT	26311	SASTKVASE	26352	QAEENSHWC	26393	MVRQCDYKM
26271	GANKTCCEG	26312	TADLNHGEQ	26353	SAAWQFGGS	26394	AIRTTMCDW
26272	GASPKLGPV	26313	TADVNLALD	26354	RACGMRGDY	26395	KSRRGKLTW
26273	HADEPYEWR	26314	TADVTAQWL	26355	DADAWWMMH	26396	TDRQIDQYH
26274	HAEYQQQMG	26315	TAGHFKWEQ	26356	NACDWKNDY	26397	TSRSEAADL
26275	HAGYIASPN	26316	TAIRADWQS	26357	SAIGKTKKM	26398	IMRHCYSNG
26276	HAIPFNDSQ	26317	TANRMTNQP	26358	QAVRRWWFF	26399	KRRDGTRSY
26277	IAFHRGADH	26318	TASEAHMEK	26359	AAIRYDPGF	26400	YIRVGGNNE
26278	IAGPQQQMT	26319	TATLWARGY	26360	NAGWQGQVD	26401	QERSMNAWW
26279	IAHHRGCHT	26320	TATQRGMFQ	26361	NALPAQQHY	26402	EKRESFDCG
26280	IAHVMDCLE	26321	TAVHNTHVQ	26362	IAYHYDQQN	26403	DYREKYKCQ
26281	IALVMTSWE	26322	VAFPAQLGA	26363	DATTSTECH	26404	DHRFKYYDH
26282	IAMMENRQQ	26323	VAISTSPTT	26364	HAVNFDCKK	26405	DIRTTGCQT
26283	KADMCPYII	26324	VASFNQICA	26365	EASIQASWN	26406	DSRTETFFG
26284	KADMKNPCD	26325	VATDRMWMC	26366	QAIPRENPM	26407	EDRITECGA
26285	KAELGKLIE	26326	VAWMMLAKD	26367	DAHAFHCQP	26408	ESRNWDDTD
26286	LAMPKPDKQ	26327	YAEVFRFKG	26368	LATTTSYHC	26409	ESRTMNMYE
26287	LANEQIDFT	26328	YAIDSSRME	26369	DASQIPITF	26410	EVRSTRKNH
26288	LATMHSGQC	26329	YANWLYDCT	26370	VAADSRMMI	26411	FPRQASSQQ
26289	MAAWCHKSQ	26330	NAELQQADI	26371	PACKEWCFC	26412	FWRQSGKLQ
26290	NADPCMWMQ	26331	TAAGYIVDA	26372	IAHWKCHYP	26413	HCRWVKYHK
26291	NAFQKEACI	26332	SAFQSSDVP	26373	AAMHNLFET	26414	HTRWPHIWH
26292	NAHMYTKHS	26333	EAYITYNKA	26374	FAQQIINAS	26415	IERVEYAVN
26293	NAMWHTKVG	26334	TAEVRWPCP	26375	GAKYIQNTD	26416	INRLLMTCD
26294	NANQKWYWS	26335	YAQRTCWPE	26376	EAGPRKRDV	26417	INRVEFTTS
26295	NASFSNHQA	26336	YAMPLYFCE	26377	PAHINDQYM	26418	ITRMNSETV

[1852]

26419	KIRVYCTPE	26460	SYRALFNGW	26501	GTGTHEDYD	26542	MEHTMDFGV
26420	LPRWADVQQ	26461	DLRWLPYDL	26502	IEGTCDRAG	26543	EHITEQDLR
26421	MERWWLVEG	26462	IIRQSEQYH	26503	IQTTEMTSQ	26544	ETSTHCCKV
26422	NCRPMEDMR	26463	IHRVVCRDH	26504	IWCTSCEIA	26545	TYDTRNHCL
26423	QERITGMVM	26464	EVRIQCKID	26505	KICTFGRES	26546	PSNTGKFCQ
26424	QSRTQGNDH	26465	WERCLELNQ	26506	KIITFISPH	26547	KSQTVMYDF
26425	SDRTAMCAY	26466	EQRHISRAP	26507	LCCTEECVN	26548	AECTEDNDD
26426	TCRWFAHYS	26467	DSRTQGGLC	26508	LESTSMWMY	26549	QNLTKEDQY
26427	TGREVDHTS	26468	TQREVETSY	26509	LGITMCTTM	26550	WVPTPWRTH
26428	VTRCIPTQW	26469	KQRIAYEWQ	26510	LQSTGIFYA	26551	NWCTMYRYW
26429	WCRPQQPDF	26470	NTRIVCDLS	26511	LSQTHQKLH	26552	AKPTHDVGS
26430	YVRVIDDKW	26471	KTYTHECVA	26512	LTPTTNAWH	26553	LFMTYSARN
26431	GKRVEDTSC	26472	AVATTVMMC	26513	MFCTSALRG	26554	DRTTMVGWY
26432	WCRCKRFIE	26473	EKMTHTCLQ	26514	MQETNQFMR	26555	GEMTRCFHN
26433	EIRQFVKGA	26474	MQMTQEWMA	26515	MVATCFNND	26556	NDCTCRAIM
26434	DVRSHDEMT	26475	YTKTYVDCF	26516	NMNTGQQPG	26557	IHFTSQSCL
26435	WKRTAGVHF	26476	KSGTCGYGQ	26517	NQGTCTQYS	26558	NSHTIMNGM
26436	GHRPKLCDQ	26477	SKNTVGIGH	26518	NQQTQVCMT	26559	YPDTADIFW
26437	GPRHIEEDG	26478	TLTTDKTYS	26519	PEATKPIKM	26560	SSMTAERFT
26438	ENRSPNNPA	26479	ACGTGVMAL	26520	RPCTYEYQD	26561	SVNTMWYPD
26439	ASRSCMTQE	26480	YYDTISPDG	26521	RWDTLYFDI	26562	GRHTAFATG
26440	ECRMAVQSC	26481	QMLTTAHNL	26522	SDETGVPAK	26563	PWTTPENYT
26441	NNRRQFGYV	26482	IKETWRVYF	26523	SILTYQNVG	26564	IWITCPYAA
26442	TCRYEELAD	26483	QLDTYMEGR	26524	SMATNHWAE	26565	ACHTGQSGM
26443	APRDMDACV	26484	ACDTPFQYH	26525	SQFTQLTQA	26566	QMVTQLNWE
26444	PPRTVARFK	26485	AEGTGMGNF	26526	THYTDVLVV	26567	QWFTKGVGE
26445	KERMQVSKA	26486	ALNTQEMCN	26527	VCFTANLMG	26568	QIKTHSNVF
26446	TNRCHDDRG	26487	ANFTFVRSE	26528	VGTTTMQSQ	26569	QTGTGHWFE
26447	EFRCRLIGS	26488	ATGTGWLDN	26529	VHHTNTMAN	26570	NHWTTCDYS
26448	GTRIYGKVS	26489	DFSTHPGTE	26530	VMSTDISQI	26571	LMKTSPSYW
26449	KIRDCDGNV	26490	DHMTPWERQ	26531	VQTTYDRKG	26572	NVWTNIAEM
26450	TLRAEMSFH	26491	DMITGMQGS	26532	VRWTYMTNA	26573	RQTTYDCLD
26451	VDRLFRMEY	26492	DPSTCCCFM	26533	VTPTLNRCQ	26574	GQNTTDRSN
26452	KQRDWENFA	26493	DSDTMMNHN	26534	YTHTPHLRE	26575	HVNTSCKQF
26453	SCRMVYGRA	26494	DWSTQEVNY	26535	DMVTVNCID	26576	NTHTGNDRL
26454	FFRICGSYI	26495	EFHTKNRDM	26536	DSATFKSAD	26577	TYMTNLGDR
26455	EKRLFDRTY	26496	EMVTEDFWQ	26537	AVETGESAM	26578	KQKTQNRNA
26456	ESRQVQDCY	26497	EWMTQCGWH	26538	QTNTMQMLS	26579	WKATPGGQC
26457	PWRWHKSAC	26498	GSHTCHLEV	26539	EIHTGRPGS	26580	VTMTTKSVW
26458	MKRAPTLDY	26499	GSLTQNDNL	26540	EPETLWYVG	26581	TCLSVMNQY
26459	AKRECTHIP	26500	GSNTIMTCL	26541	ATDTYNLQE	26582	QSDAVHEPA

[1853]

26583	YQERVRKDV	26624	TVIDVVRWH	26665	YRLHVWMKV	26706	EYHGLNLMA
26584	DDLDVWQTS	26625	VFPDVSFWQ	26666	YWEKVPCMK	26707	EYVWLNCGE
26585	GCNHVHIGQ	26626	VIAQVFKYH	26667	SGQPVDCMF	26708	FMHAGNLDA
26586	GIANVDDWE	26627	VIGKVHHDN	26668	FNWFVRAYI	26709	FVGCLNHYL
26587	NQIDVQWLQ	26628	VTNRVAMPI	26669	ACDQVNTVQ	26710	GHIKLNVLP
26588	REVMVCILR	26629	VTSLVTNMG	26670	QTNQVVVVE	26711	GIMQLNTEP
26589	LCHNVLDYY	26630	WYVPVNHLY	26671	FMMPVVDNY	26712	GNCFQNKSN
26590	NQPCVNTYW	26631	WYVPVNQLY	26672	WCEAVGDYA	26713	HPHVHNQKP
26591	TMAEVVINL	26632	RWDFVWYIQ	26673	SGFFVHTSN	26714	HTVDGNVFP
26592	QRMIVDALQ	26633	GVFYVGHWQ	26674	HFSHVDQCM	26715	IDAVKNRND
26593	RMVRVSVHG	26634	HGNMVLVIG	26675	NRCGVSKCC	26716	KEMAKNYEP
26594	FTVYVGARE	26635	NSVHVHASN	26676	TSTFANEFT	26717	KQDEWNCDR
26595	ACAGVTWGI	26636	NKEAVTGSS	26677	QRPIANDAM	26718	KQNCANLCE
26596	AFTNVTRWC	26637	HPVKVDDFT	26678	VQDKINYSY	26719	LLSDENTEC
26597	ARNCVGSFC	26638	SRLPVHECP	26679	QTGHFNQRG	26720	MKMEENSAS
26598	AVQSVLGCR	26639	WTLQVMKIN	26680	EGTLWNVCV	26721	MLVQENNSY
26599	DHCAVWNHE	26640	MRVQVFSDA	26681	GVIPLNMPS	26722	NLERCNTMM
26600	DISPVNRWH	26641	SDSKVQWGM	26682	MWTEFNEDG	26723	NQFFYNLQN
26601	DSYRVLCLS	26642	KSAQVYWHP	26683	KMTNSNENS	26724	NSCYMNSIE
26602	DVIPVPKSC	26643	QFMGVSSHS	26684	GRAHGNVCI	26725	NYTQCNFSN
26603	EQIHVTMAN	26644	LFIKVKDTS	26685	QCEYSNCCY	26726	QMAWENWFG
26604	EVYNVKGSA	26645	HFHEVAWGS	26686	DVISRNGRC	26727	QSCSDNYSG
26605	FLEFVHHGT	26646	ALDRVSMAC	26687	NMCEANHES	26728	RCDQMNLHF
26606	FYSSVQRDQ	26647	WCEHVIMNP	26688	EVAGENGNY	26729	RGDALNMVP
26607	GSHQVDILH	26648	TKQPVHFLG	26689	YTGHTNSGR	26730	SSFLLNHQY
26608	MGDSVLTAH	26649	YGNSVTCTS	26690	ACDFGNLGR	26731	SWVFCNVGQ
26609	MMSNVPGEY	26650	ESCLVRRGW	26691	ADVVTNINL	26732	TTLNTNCSH
26610	NHVPVVGAM	26651	EWEGVSQFR	26692	AIEASNGGG	26733	YVYRMNGNC
26611	NTLLVDFDF	26652	AEHMVVDHE	26693	ASESPNESG	26734	LMQWYNRCW
26612	PFTNVTRWC	26653	TCNKVANMY	26694	ATCQSNKCP	26735	VTEVFNQDV
26613	QNTCVIGAQ	26654	GSYIVQNHC	26695	ATTRHNTEA	26736	GNKLKNDQM
26614	QQDAVCVSH	26655	HHFHVNLDI	26696	DGWSGNKMA	26737	YMGSTNIWM
26615	QSFVVQGTY	26656	KTSNVMDQM	26697	DMKLQNDHP	26738	QLEECNPLS
26616	QWHAVTTNE	26657	NVEDVGMIE	26698	DMLKHNNAR	26739	LVDCSNTCQ
26617	RLDCVKESP	26658	VDMGVHVGS	26699	DNNAMNDQQ	26740	KLYGLNHSC
26618	RVVEVHAQV	26659	GGSYVSAPD	26700	DQQYSNVCC	26741	SGALKNREE
26619	SDNFVITHS	26660	VEDRVFACC	26701	DRIWINWVG	26742	MIKLINCVN
26620	SQQFVHHIY	26661	WKAIVLEYW	26702	DVQKKNDKS	26743	QPSRFNTAD
26621	SYYWVWMAG	26662	ASVMVAITG	26703	EKDEINRAW	26744	MQEHKNFYG
26622	TNNVVIEWG	26663	DSQSVGDSF	26704	ERFMANTGL	26745	TTYPHNTHG
26623	TSMCVETQY	26664	INMPVSKDQ	26705	ESSDTNYRQ	26746	IYKWDNKFH

[1854]

26747	QRTQCNFSI	26788	AGSQMHMLS	26829	YISEYVMMV	26870	AILVEFPHG
26748	KYWCKNEWD	26789	ASWSPQMNT	26830	TTTNQFMFQ	26871	AWTPSHGHP
26749	IEHCSNRYA	26790	ATGNKTMSY	26831	TVGIRHMKD	26872	DGCKMFNHC
26750	MINWGNKCT	26791	DLNQCQMYC	26832	QCFSTEMCQ	26873	DGYIYHTHL
26751	VPLGFNFAG	26792	EEVHSAMEG	26833	ELEPDGMHV	26874	DIGNQDRHS
26752	RENGYNATM	26793	EGPALWMNT	26834	THARFKMKS	26875	EDMPAVDHN
26753	QKSMLNSIQ	26794	EGVHLPMGN	26835	YSVQAYMAN	26876	EFDQYMTHC
26754	EGQYWNYMT	26795	EMEMEEMVG	26836	FQWRTEMNN	26877	ERVPRYVHH
26755	TFIYLNHFG	26796	EQGYRVMSF	26837	PSIKHGMPK	26878	ESVMQANHT
26756	DYFQYNEAP	26797	EVFHDRMMG	26838	DVDFLKMHC	26879	EVCIYCSHV
26757	SCQLLNDGG	26798	FCSEMVMNQ	26839	IHSFMAMAQ	26880	GEDEAWQHT
26758	AHEDPNTWC	26799	GNSKNDMHC	26840	ENCMSRMKP	26881	GIVVADRHV
26759	KNELENVIN	26800	GQCNTMMWQ	26841	MLMHADMTQ	26882	GSLPWKNHV
26760	VGGYSNQWV	26801	HINAAVMHE	26842	FTEPMCMQG	26883	HIADKSTHR
26761	WDNWSNTQI	26802	IEPDCPMIP	26843	YHCCFAMQS	26884	HILYAWWHS
26762	HQYVHNWRQ	26803	KIEAEEMHA	26844	PGGCYRMME	26885	HSPFFMNHA
26763	VHHKTNRSY	26804	KLVSKTMHH	26845	SQKQFSMQN	26886	HTYPLDHHF
26764	TMWEDNQNM	26805	MMESALMDR	26846	HMLHAIMES	26887	IYSDNQYHH
26765	DCDPYNMVD	26806	NQDIMVMGQ	26847	TNGVQIMHV	26888	KIVECVSHI
26766	QISAWNEFF	26807	NWACMGMGI	26848	ETEIRQMDG	26889	KSMWSERHC
26767	GNEDCNLAC	26808	PDWYNWMED	26849	YSNSTIMNL	26890	KWWWAYTHK
26768	MIMGYNFEF	26809	PMKSPVMKD	26850	YVKMCSMRN	26891	LIAMRIVHD
26769	LPSHCNLGS	26810	PQEEFMMTQ	26851	GTVECEMNA	26892	LYCSPPNHC
26770	TIDLSNKVT	26811	QELYLPMGC	26852	FWGMESMSC	26893	MEKRCGVHQ
26771	TTTITNRLC	26812	RFKHQMMMH	26853	ASGHMTMCM	26894	MIACDKVHY
26772	QCTAHNTCM	26813	RMDMKEMWG	26854	NTSELLMAQ	26895	MMAKHTNHE
26773	DSMVYNRPQ	26814	SQQLECMGP	26855	SEDCMDMYF	26896	NCVQWKYHQ
26774	KIQPTNHHE	26815	SRAEEYMPW	26856	STEFWVQHT	26897	NGCCLMTHL
26775	KGLIYNHHT	26816	TRVPSMMYK	26857	WSSREYWHK	26898	NYICTRKHP
26776	NYGCHNCKY	26817	TYAANIMHV	26858	TVMSAWSHM	26899	NYQNKLLHG
26777	QIGRFAMSH	26818	VCTHYVMQR	26859	ETHYPYNHI	26900	QHMIREAHG
26778	RTGDAGMSP	26819	VMKPCFMTP	26860	FQIRFDRHH	26901	QLSMLGVHF
26779	ILLPDDMLK	26820	SKNLYTMAG	26861	TSSHDEAHK	26902	SCNMCVSHE
26780	HLDCLKMEM	26821	LDGPIGMGT	26862	TSYLRKIHS	26903	SHKFMHYHW
26781	YSNSIQMSF	26822	FCNYAVMSP	26863	TRYRMRYHG	26904	SHLNNITHQ
26782	YREHRTMSD	26823	TLPKMHMWP	26864	GIWRTREHY	26905	SRNGSKRHA
26783	GWYQKCMGD	26824	TKPLEDMPT	26865	RGSLMSFHE	26906	STKPWPEHN
26784	PTNYGWMWR	26825	ESAEPRMIP	26866	SYAAFLHHG	26907	SWFECFHHE
26785	GCYWNMMQP	26826	AEVPYVMQQ	26867	LQFLTVVHL	26908	TIMMGVEHM
26786	PNGHYWMIP	26827	QHSRAYMWY	26868	GVSRLPYHA	26909	VMFYTVHHE
26787	ADYNNEMIQ	26828	TNFPTSMHG	26869	AHFFEIPHS	26910	YSMEEHFHN

[1855]

26911	YYVFMQLHA	26931	NKPAMMQHE	26951	ILANPIRQI	26971	EPLYSWVGI
26912	HWELIRCHD	26932	ITTDATGHF	26952	KHNPSRHDI	26972	ASIATMDRI
26913	YEICKHGHF	26933	RDQKGLGHM	26953	NFYLSGFLI	26973	TCGHTRFFI
26914	TCNVSWEHE	26934	MYLSFPVHE	26954	KQCSLAPTI	26974	EHKGCMQAI
26915	HGNSFLVHW	26935	DVKYATLHD	26955	GFHNTKLPI	26975	VNDDYECGI
26916	YITCQYSHH	26936	MYAGWLNHG	26956	TILQQMEEI	26976	NMPSLRTRI
26917	QVVIDPAHY	26937	AVIHEKPHG	26957	ACNPFWHSI	26977	FKIEWSQDI
26918	RVWSLEWHL	26938	GTIRMPEHR	26958	AFTETKCSI	26978	EFVLGCQGI
26919	KYSPFQEHN	26939	QKHMILCHF	26959	AGSHFGRRI	26979	WDAYYTNGI
26920	YPQPCSCHS	26940	QEAGYTRHN	26960	EVTRTQVQI	26980	QSMVACYTI
26921	QVWSDGQHV	26941	AKEHMPHHM	26961	FYLCNYGAI	26981	ATNKYANAI
26922	SVTYHEWHP	26942	TKQFFTAHC	26962	LEAKWQCSI	26982	WHFLMIFYI
26923	TCNVSWKHE	26943	AFGLTTGHT	26963	LWMWQLSCI	26983	QQEAYILRI
26924	MLFDFEEHN	26944	YLESKTIHA	26964	MKDRMHQVI	26984	GGALARSSI
26925	QTMCITIHP	26945	TRDPFWHHD	26965	NVAQPDFQI	26985	EFDWLLGTI
26926	KTEKWASHE	26946	GVWMQGFHP	26966	QITSGCHCI	26986	RRTDHGEPI
26927	FCKCFMKHC	26947	AVQMQQQHI	26967	VTCPIQRTI	26987	WPNIKHQPI
26928	QMSWGKYHQ	26948	VKAWWHDHQ	26968	VTHKLVDAI	26988	HQCQSRSYI
26929	KYINFDHHQ	26949	GIHQLIHHC	26969	YPHKKCRCI	26989	EICMRSGFI
26930	MKGYFVHHS	26950	TRQWMCVHC	26970	YSLEKRFQI	26990	IECSPRESI

[1856] 实施例9

[1857] 在骨骼肌中具有组织嗜性的AAV5变体

[1858] 此实施例描述了使用实施例1-实施例2中描述的方法和系统发现的在骨骼肌中具有组织嗜性的工程化AAV5变体。

[1859] 分析CNS、肝脏、骨骼肌、心脏、肺、脾脏、淋巴结、骨髓、乳腺、皮肤、肾上腺、甲状腺、结肠、坐骨神经和脊髓组织以鉴定AAV5VP1 581一589区域中驱动骨骼肌嗜性的氨基酸残基。结果显示于图18C中，该图显示了相较于所有其他分析组织在骨骼肌组织中AAV5VP1的581至589区域中每个氨基酸的位置频率，其中衣壳多肽具有SEQ ID NO:2的序列。位置(“Xaan”，n＝1-9)在x轴上表示并且氨基酸残基在y轴上表示。特定位置的特定氨基酸残基的频率在热图中作为强度表示。较深的方块表示特定氨基酸在特定位置的频率较高，较浅的方块表示特定氨基酸在特定位置的频率较低。

[1860] 分析相对于给定氨基酸残基出现在所有其他所收获组织中鉴定的变体中的指定位置的频率而言该给定氨基酸残基出现在骨骼肌中鉴定的变体中对应于SEQ ID NO:1(推及于SEQ ID NO:2)的位置581至位置589的所述指定位置的频率，以确定优先靶向骨骼肌组织的衣壳的一组序列规则。参考下文表17，和SEQ ID NO:2(AAV5 VP1)，以下氨基酸由此可以在Xaa1-Xaa9任何一个或多个位置处以任何组合独立地突变，以提供与野生型AAV5 VP1衣壳(SEQ ID NO:1)相比具有增强的骨骼肌嗜性的VP1衣壳，其中骨骼肌嗜性在此是指优选骨骼肌转导的特性，而不是优选转导所有其他所收获组织的特性。

[1861]

表17 骨骼肌嗜性规则
Xaa1选自A、E、H、M、P、Q或S
Xaa1选自P或Q
Xaa1为Q
Xaa2选自F、H、I、T或V
Xaa2选自T或V
Xaa2为V
Xaa3选自A、G、I、K、M、Q、R、S、T或V
Xaa3选自A、L、P、R或T
Xaa3选自L、P或T
Xaa3为P
Xaa4选自D、E、G、P或S
Xaa4选自D、E或S
Xaa4为E
Xaa5选自H、L、M、P或V
Xaa5选自L、M或V
Xaa5为L
Xaa6选自E、H、N或P
Xaa6为P
Xaa7选自A、H、N、Q或T
Xaa7为H
Xaa8选自I、K、M、P或W
Xaa8选自I、P或W
Xaa8为P

[1862]

Xaa9选自A、I、M、P或V
Xaa9选自A、M或P
Xaa9为M

[1863] 下文表18提供了在骨骼肌组织中发现并且符合表17中提供的一种或多种规则的变体AAV VP1衣壳多肽的581至589区域的序列。因此，本公开提供了由具有SEQ ID NO:2的工程化AAV5 VP1衣壳多肽组成的rAAV，其中Xaa1至Xaa9区域是SEQ ID NO:28991-SEQ ID NO:29990中的任一者，如表18中所公开。本文还涵盖由工程化AAV5 VP2衣壳多肽和工程化AAV5 VP3衣壳多肽组成的rAAV，所述多肽在对应于AAV5 VP1 581至589区域中的氨基酸的AAV5VP2(氨基酸残基445至453)和AAV5 VP3(氨基酸残基389-397)中的区域处具有下文表中公开的序列。

[1864] 表18

[1865] 驱动骨骼肌组织嗜性的AAV5 VP1衣壳多肽中581至589区域的序列

[1866]

SEQ ID NO	581-589 序列	29027	QPMSFNWKN	29065	QFAQLWCAK	29103	QLCNPLAPM
28991	QAEIVCNSK	29028	QPNYVEHCG	29066	QCTAHNTCM	29104	QLPGIKMWD
28992	QANGSMDHH	29029	QPTMHMPRG	29067	QAADSASAS	29105	QMFGFQPGN
28993	QANWWVDAG	29030	QPVVTWQSH	29068	QADRDGCYA	29106	QMHPMIRFF
28994	QARTMDERP	29031	QQITTEHSG	29069	QAGWINTLN	29107	QMHSMTNCN
28995	QAWEGCAGE	29032	QQLCQLDRK	29070	QATFHATQK	29108	QMRIWRMMW
28996	QCAFIPNSA	29033	QQYLCMSSD	29071	QCADKNVLE	29109	QMVPATPCN
28997	QCGNVMLMP	29034	QSCATTSAQ	29072	QCFNTCPCD	29110	QNAFMDLLA
28998	QCHFLRTCK	29035	QSFQAPRLA	29073	QCKTVIMGA	29111	QNCFRYECP
28999	QCMCVEIAA	29036	QSHRVMDIP	29074	QDFQFDEVW	29112	QNETRQHCW
29000	QCNAQAELF	29037	QSHYCHTFS	29075	QDTMMTYMS	29113	QNMPVMQYR
29001	QCQLTVSMS	29038	QSMSYQKEG	29076	QDWFNLAHW	29114	QNRPTECIG
29002	QCQWQKMEH	29039	QSNLKSMYC	29077	QEEHQKFGM	29115	QNVKEYSAS
29003	QECYEKMEC	29040	QSSSMACDN	29078	QEENENMGC	29116	QPADPNDGG
29004	QEHCMNYGP	29041	QSSWAWMAE	29079	QETQAAMRQ	29117	QPAPGHRGQ
29005	QEMFIQARG	29042	QTGNWWQYG	29080	QEVTTTDPQ	29118	QPILPNIWE
29006	QENLTNGDN	29043	QTIFWLNVG	29081	QFMVKNSNH	29119	QPQVTYDVR
29007	QFIMLGETN	29044	QTISGFQDK	29082	QFNLTDYFH	29120	QQLFWDIGV
29008	QHVQIMAGQ	29045	QTLHLHTNM	29083	QFRGLEHMV	29121	QQPLMRTEE
29009	QIEGVNHAN	29046	QTMRLEPYH	29084	QFVSTHHEA	29122	QQSTKMKWH
29010	QIERYWHAT	29047	QTNLLVDTS	29085	QGNTCHKDA	29123	QSDKKNRCA
29011	QIHGVTIWQ	29048	QTTAILKLD	29086	QGRIYDQAT	29124	QSDQLSRIQ
29012	QIHIWDGQS	29049	QTTLKMKED	29087	QGSWNFCPP	29125	QSELGHELP
29013	QIHLRECSF	29050	QVETPWSGA	29088	QHAPLHNFI	29126	QSLWQQLSM
29014	QIRAMLDSS	29051	QVQNVWCPK	29089	QHCPIVEQD	29127	QSQPEWNEK
29015	QISIADCYN	29052	QVVTMPDPL	29090	QHREIVKLQ	29128	QSSTDQEAW
29016	QIWNEVHAQ	29053	QVWDMNDNA	29091	QHSTDSGDK	29129	QSTENNWYS
29017	QKNEYWTAF	29054	QWFRQANYK	29092	QHVELQVMK	29130	QSVQPCDDP
29018	QKNLVPSYW	29055	QWLLLCHGN	29093	QIACLQKQG	29131	QSVSQLTCC
29019	QKSFPGDGI	29056	QYERLHTNM	29094	QIHDQLKGN	29132	QTAATMTED
29020	QLMVYMYES	29057	QYIPFERYW	29095	QIQEVQSSV	29133	QTAVKNIDN
29021	QMEYTWMGG	29058	QYMTQAKKD	29096	QISIGCNSQ	29134	QTCNAEVFT
29022	QMFQDHNQP	29059	QPSRFNTAD	29097	QISSSRDMY	29135	QTCNCGKAF
29023	QMHHWDVWG	29060	QKTANGWCH	29098	QITDWENFH	29136	QTGQYATLK
29024	QMQPTCGQF	29061	QMCVWMTRQ	29099	QKDPLHVHS	29137	QTIPREAEC
29025	QNENVVWMS	29062	QAKMILDGT	29100	QKDQVWHAW	29138	QTNLPVVKN
29026	QPMHQQVTF	29063	QTAEWYCGA	29101	QKEEQPKMG	29139	QTQTWERQD
		29064	QTNGHPYFP	29102	QKPELEQGP	29140	QTWCKRIHF

[1867]

29141	QTYMPHHAH	29182	AVNGNVYGR	29223	LVDPHRPMM	29264	RVNGSTCRH
29142	QVCSMFGNA	29183	CVEGMVFVP	29224	LVNHGDSYC	29265	AVESSKGYV
29143	QVINVHGRW	29184	CVGQYNNDM	29225	LVTGNSWEQ	29266	AVGCFGKEP
29144	QVIRHECQY	29185	CVLRQYDGG	29226	MVIGTSLLA	29267	AVMEYILYK
29145	QVSNTATCL	29186	CVNVRGRSS	29227	NVMGHDWVH	29268	AVNDQWQWG
29146	QVTQVQALG	29187	CVRQKAETE	29228	NVQWFEQNI	29269	AVQLGNSMF
29147	QVVEKHENR	29188	CVTAASPWN	29229	NVVWGMLYG	29270	AVSAPLSKE
29148	QVYHYCRRH	29189	DVAERHMHE	29230	NVYINLASS	29271	AVVQGMLYP
29149	QWHVGGTSA	29190	DVCMATIGQ	29231	PVPHNYQWW	29272	CVAHFDQCF
29150	QWKCDKQSN	29191	DVDPLPCEP	29232	PVRHCSVWF	29273	CVDYHRREQ
29151	QWRNPIYGP	29192	DVGYQMKKQ	29233	RVCTFCCCF	29274	CVESGTGHQ
29152	QWTQGYASP	29193	DVICYSDHR	29234	RVDFNVYCK	29275	CVIHRMDMP
29153	QYEWLNTCD	29194	DVISMSLMG	29235	RVFMLEMYA	29276	CVMNVIWDA
29154	QYFRFGFDP	29195	DVLIDHDFF	29236	RVNMNDIGE	29277	CVNETDTAA
29155	QYGNSLNAG	29196	DVMHFLDLA	29237	SVAMRQLDN	29278	CVSYNPSIR
29156	QYVPAGYFA	29197	DVQNMDIGV	29238	SVFWPEHNP	29279	CVTINTISG
29157	QADRIGQCL	29198	DVRDDEFGN	29239	SVHICEDGW	29280	DVNCCITCG
29158	QANAMAEFP	29199	EVCGMECSG	29240	SVLSFVAQK	29281	DVNTYANRD
29159	QCACTITMK	29200	EVDDKPSGH	29241	SVNPYPHGI	29282	DVRAMTNDW
29160	QCASTMSMT	29201	EVEYEAKTV	29242	SVRTHMHLP	29283	DVVEKCSWF
29161	QCELLNVLC	29202	EVFVADNNI	29243	SVTFDVGGP	29284	DVVHMWWWS
29162	QCIEYEATE	29203	EVGAWGASY	29244	SVYICNMDQ	29285	EVACYDMTG
29163	QDGMEAKNR	29204	EVKAANWQV	29245	TVAVNEHLA	29286	EVAMMINNN
29164	QDIKKWCSE	29205	EVQHHYVSG	29246	TVFVYDNPG	29287	EVGTECQWL
29165	QGFMEQQHT	29206	EVRMCNATH	29247	TVLDAWTAE	29288	EVGTNHKQH
29166	QHQAEMEGF	29207	EVVKKPWDG	29248	TVNSNRTGD	29289	EVKMFYKQY
29167	QHTEMDCSN	29208	FVFTHIDDY	29249	TVNVFPIHE	29290	EVLDHISAG
29168	QIGGFQYYY	29209	GVEPISRAA	29250	TVYDVNAFS	29291	EVREEGVLT
29169	QIKLQQWYD	29210	GVGQKCGHQ	29251	VVMDAHCNL	29292	EVTHQLKNA
29170	QKDEYLNWH	29211	GVLAQQAQW	29252	VVTDYMVGA	29293	EVTVFIAPW
29171	QKDMHAYPL	29212	GVLLPHHWM	29253	WVKFANDCQ	29294	EVWNEESMF
29172	QKPRWFCCP	29213	HVNQWGRYG	29254	WVRMRNLEK	29295	EVYIKQQYY
29173	QMSSGEWVR	29214	HVQSIIHAE	29255	YVAGLFNNR	29296	FVPPSCYNY
29174	QNELRESYE	29215	IVDWSNVLH	29256	YVKFSVNIN	29297	GVAAQPCSP
29175	QSKPMEHVQ	29216	IVGMANQEH	29257	YVQLRIADD	29298	GVAPKPDFE
29176	QTDWSMFLP	29217	IVHSLDAPG	29258	YVQTTAHSS	29299	GVASCNRFD
29177	QTYFPWHWF	29218	IVSFGYAQG	29259	MVTERKPWS	29300	GVMPKYFFG
29178	QVCSRWTKS	29219	IVSYLKTND	29260	SVLHVQALA	29301	GVNAKQQTR
29179	QWVEWPFPC	29220	IVYDMMADV	29261	WVHPYSYCA	29302	GVYFMCDFM
29180	AVEKVMEVQ	29221	KVDTQWKSV	29262	DVKSWAVCD	29303	HVQTAEWIP
29181	AVKLNQTYK	29222	KVRPPCIYI	29263	LVVTKFTYV	29304	IVDWPANRY

[1868]

29305	IVIGRWYSC	29346	TVAPQSSLW	29387	HHPKFAEKG	29428	SSPNVHDLD
29306	IVMRALANN	29347	TVCMDANFK	29388	HHPQVLRGM	29429	SWPFSWLIV
29307	IVMSFHRNH	29348	TVKDLRVVT	29389	HNPVWVQQE	29430	TKPDSANWD
29308	IVTDCRDGL	29349	TVLWKHDKA	29390	ICPLEQYEM	29431	VNPQEEFFD
29309	IVTTKQLCR	29350	TVSQCSFFP	29391	IGPMGQYGR	29432	VQPPYIAMV
29310	KVDDHICPG	29351	TVTLSEKWH	29392	IMPEAKNTP	29433	VSPRDLHHM
29311	KVGWIEWYA	29352	TVYGFPNEV	29393	KCPDEMNNI	29434	YFPQESENS
29312	KVYPLDFKM	29353	VVGVAPVMP	29394	KTPFTRWLR	29435	YQPNWRSQT
29313	LVDGCVWMH	29354	WVKAVTLEQ	29395	KTPLMVASI	29436	AHPWCLTPN
29314	LVDYTGNED	29355	WVVQTFTFD	29396	MLPSFGDEM	29437	DNPWLLTFR
29315	LVFYTSVRG	29356	YVDCRNDAG	29397	NGPQGLMAC	29438	KQPTSASPI
29316	LVTKEDCAY	29357	YVELYIVDQ	29398	NNPTNTEAR	29439	MDPLGPAAD
29317	MVAVDGTEP	29358	YVRDTALCF	29399	PCPDHHRWL	29440	NEPQICRKC
29318	MVDMKQLMF	29359	YVTSALDTF	29400	PLPCPRSGY	29441	NQPGRFRTV
29319	MVEANEHAM	29360	CVSFDACFA	29401	RIPPMMIML	29442	VPPRMTIEL
29320	MVEIHPDYD	29361	DVDRLCSIP	29402	SSPPWGHPF	29443	YSPTTNGGP
29321	MVHLRGDGE	29362	DVSQTNFSE	29403	SSPSHAEYL	29444	AISEHFWNV
29322	NVDSSPFLT	29363	DVVWSTKDQ	29404	WMPLYCMHV	29445	ANVELMTIK
29323	NVHGIMWDN	29364	DVWSSISDS	29405	WQPSHAYHF	29446	ATHEQGGQL
29324	NVQTKMHLP	29365	EVDGGGKMQ	29406	FCPCGQEMC	29447	CFCECSFCE
29325	NVTPHEFQA	29366	EVFLAGLMS	29407	SWPLPERPS	29448	CFDEFMLWG
29326	NVTTIEPAH	29367	EVLLAGIGP	29408	LSPWPQTDQ	29449	CITELINYT
29327	NVVQVTQNA	29368	EVTPMYCSD	29409	EQPNMGDYG	29450	CLAEQITCQ
29328	NVVSKDHGG	29369	FVRGNLQTP	29410	MLPLEMKNL	29451	CRSEGWIHT
29329	NVWHGYWLN	29370	GVCHNTGCT	29411	FPPREGPEW	29452	DGMEWDLCG
29330	PVAHVPFGM	29371	HVFILDRTC	29412	AHPHCCRHP	29453	DHVEIQHMS
29331	PVTAPGFAH	29372	HVIPDPRPA	29413	CAPQFNWQY	29454	DTTEQMQIY
29332	RVYNPMYWH	29373	IVETDGSLM	29414	CQPRYHWCY	29455	EACEGAMLE
29333	SVAQEMEST	29374	IVKFECCNH	29415	ESPNWGYMD	29456	ECEEPSLMK
29334	SVCEDAFFD	29375	LVYHSFVHE	29416	GQPQQIWQA	29457	EEVESGHFG
29335	SVDAQAFVY	29376	MVCREQENG	29417	HHPHQANIY	29458	EILEEQESC
29336	SVDQESAYR	29377	MVEVATMPH	29418	KDPLGESKW	29459	ELIEKVYRW
29337	SVERHGSDK	29378	RVDDYERTF	29419	KMPLKQADF	29460	FTREDKDRC
29338	SVHQWCDVT	29379	SVIHYSWAL	29420	MFPQAPADG	29461	GAVEKMVQE
29339	SVIAFHIHV	29380	SVSAWNPMS	29421	MFPVCQDRD	29462	GMEEVEREG
29340	SVLHTMMVK	29381	DKPFEWDLL	29422	MGPMKPMMD	29463	HGEEIHWKD
29341	SVNEVIDQV	29382	EQPHDPSST	29423	NPPLCQEYG	29464	IWTEAAENC
29342	SVNQESAYR	29383	FDPNRSQDV	29424	PGPEQITAG	29465	KGEEMYECG
29343	SVNSKNHTQ	29384	GEPKCVMME	29425	PPPWFQSRS	29466	LREEYKLAG
29344	SVSYGNHSV	29385	GGPDIKHYI	29426	PQPCSWPAQ	29467	LTCENQQIP
29345	SVTDRDHQY	29386	GGPSEQPRP	29427	SGPWCGSAW	29468	LWTESVMGN

[1869]

29469	MALERCDVD	29510	ESEEMHAYE	29551	AFHHLGEFQ	29592	KCADLIIRS
29470	MEEEYMFTA	29511	GKIEARETD	29552	ALMPLTRWN	29593	HATGLVNFM
29471	MMSEQCHLT	29512	HMFERPRTQ	29553	AMHSLHYMF	29594	ATMNLSWGL
29472	MNIERDRNH	29513	HPSEECSWR	29554	ASEYLASPM	29595	TYESLQNSW
29473	NTCEPVQMI	29514	KLVEMQESK	29555	ATTLLHYKN	29596	NLSNLVSCR
29474	PTLEFLNLM	29515	KWCEVLWQP	29556	AWWWLNAQG	29597	ALDALHYCS
29475	RPNEPLDYS	29516	LEGEAYAQE	29557	CNIQLHTHY	29598	ANIGLETGV
29476	SKREIENCH	29517	LFSEFFEYE	29558	DRVRLMMVA	29599	ARVDLMSHQ
29477	SLCEMMQPH	29518	LSNEMTIDA	29559	DWGYLRIWC	29600	AWTGLHVCN
29478	SMTEIVVAW	29519	LTIENPHDQ	29560	DYSKLDVHY	29601	CRDGLPHTV
29479	SSNEGGCWM	29520	LYKEIGENI	29561	ERSTLPMSR	29602	CTFWLYFFR
29480	STKEDFDMF	29521	NAEEWWSCD	29562	FADYLAAAS	29603	DFNYLNWFC
29481	STVEKMQTL	29522	NCVEAGSFA	29563	FPYHLWVGG	29604	DILDLVNSF
29482	TTDECHGHP	29523	NMMEYERFS	29564	GAEKLHLVH	29605	DIYQLHPSY
29483	TTVEMECGM	29524	NNMEDRWGT	29565	GGIDLTLMM	29606	DLFCLNYDF
29484	VEEERPFWS	29525	NSNEHGAMW	29566	HADCLELSC	29607	DQHKLTVEN
29485	VRGERLSDL	29526	PTCEVPENC	29567	HGEDLNEFL	29608	DTDTLMNSG
29486	VTIERNIWQ	29527	SFIEFPAMT	29568	HHDDLAPWQ	29609	DTELLLVQE
29487	VTRELWLPQ	29528	SIDEQTFPP	29569	IICTLSLDS	29610	DYMLLQGVY
29488	WCTEFSTRQ	29529	SKEEVQYIF	29570	ILNVLESWE	29611	ECHLLEEQR
29489	WNVETFLGY	29530	SWTEISSVK	29571	KERRLWSGV	29612	ECRSLCFQH
29490	YIAEMEFMN	29531	TAAEMIYEY	29572	LFDRLKLNE	29613	EDRDLSWTF
29491	KQHELMVRV	29532	TWDEQTAQV	29573	LRSMLASYD	29614	EGMILDHTP
29492	PKHEDMRWW	29533	TWRECQSSH	29574	MHGHLTQSI	29615	EKANLTYGV
29493	NQREWHGLA	29534	VGNEDKCDP	29575	MSISLWGNA	29616	EKRSLPACM
29494	SECEVLCSL	29535	VMIEKTMMH	29576	PENPLVTMS	29617	EPHSLDICL
29495	PHAETHWYT	29536	VPVEVHKQA	29577	RNTCLNTGD	29618	EYFVLHVYD
29496	AAMESCAEI	29537	WGAENHGPW	29578	SLSRLYWSN	29619	EYNSLLCHF
29497	AEGEIEMWT	29538	WKDEYTVDD	29579	TIRSLEMVW	29620	FRKLLTRRF
29498	AFVEFSDVT	29539	YAREFGNMY	29580	VGIPLQAHC	29621	FWCYLVLMV
29499	AGCETWNAP	29540	YTDERAMLN	29581	VMDGLGLVL	29622	GEDTLAAWG
29500	AMHEATDAQ	29541	CDAENSDGG	29582	VTGFLMIPT	29623	GGEFLAVAY
29501	AQSEWGSAT	29542	CLDEPRTDF	29583	VTRDLQKEQ	29624	HHMQLLRSG
29502	ARDESAEEH	29543	GHEEYWEHV	29584	VYDYLINGA	29625	HPGKLFETH
29503	ATHEMHLDA	29544	HHREQKDHS	29585	WHGRLMITC	29626	IGTALERQG
29504	CKKESGSYE	29545	HSAEQVLME	29586	WYQFLDNQL	29627	KIVVLAKGE
29505	DCDECQDDQ	29546	MIQEIKEDN	29587	YACLLDTLG	29628	KQTVLHVHH
29506	DKDERAEGM	29547	MWCEVMHCS	29588	YSYSLMSGW	29629	KSLPLHDHF
29507	DQWEYEVQF	29548	SCRECDECP	29589	DKINLAVSY	29630	LFTCLQEQE
29508	EELEQQMCT	29549	TPDEQSHYD	29590	HMEWLEEQA	29631	LSGWLSQFG
29509	EMHEWSRAN	29550	TTAEYGMSI	29591	ITVALNQGW	29632	LTTHLPEWN

[1870]

29633	MKTVLYVGD	29674	DPWHEPSMG	29715	DTGSTPVGT	29756	HHELDPDWE
29634	MNLHLPQWG	29675	DRHTKPNST	29716	EINPEPNWK	29757	IPHNEPKAG
29635	NFHSLESTP	29676	ETMYWPAMD	29717	FKLWNPDSI	29758	KECFIPCAF
29636	NTVTLVQHE	29677	FFRDCPYEY	29718	FWRSVPEYW	29759	MTRIRPDAN
29637	NYGKLGKSY	29678	GNNRNPTYW	29719	GISSWPSAH	29760	WKAPGPITR
29638	PMDSLTRVY	29679	GTSDFPEAV	29720	GPNTHPHMA	29761	AEHSPHHDF
29639	PWQNLAVNH	29680	HNLKIPMME	29721	GSSYVPICF	29762	ASALHCHNQ
29640	RWQDLLNTG	29681	IASHNPSKQ	29722	GTDACPLHL	29763	ATADNHHPR
29641	SQMHLFTSW	29682	IDFRPPNNT	29723	GTNGVPKVN	29764	CSSLYSHMP
29642	SSFILLDHQ	29683	IPNGPPNNQ	29724	HDYPTPKMY	29765	DMCKYSHFV
29643	TFNKLSCMW	29684	LGFRKPFTR	29725	HMTFKPTAQ	29766	ETHTRAHDC
29644	THAMLDRNF	29685	LQYTFPIGQ	29726	HQYNEPHVS	29767	ETIFVDHLN
29645	TQEALMDGQ	29686	LTEATPKWA	29727	ICTNYPKMD	29768	FNIHICHFI
29646	TQTNLAECA	29687	MFQYYPYCD	29728	IHLMKPAHI	29769	GEASMEHMQ
29647	TSQLLMHGN	29688	MGEDRPSMT	29729	IHNKFPLTM	29770	GRIFSKHPD
29648	TWGLLWKQY	29689	MHFKYPQGG	29730	LKDFDPAIH	29771	GSNQIQHGN
29649	VKLKLSEGQ	29690	MPDMDPAAL	29731	LKVLWPTEM	29772	GYQPDQHCP
29650	YTYALAHAT	29691	PQHISPDPT	29732	MHLRDPHQH	29773	HRLIEQHDT
29651	YYNNLNIHT	29692	PYDMIPCNR	29733	MMHGYPFYP	29774	IMMLMSHLE
29652	AHAMLDRNF	29693	SAAFQPRAS	29734	MNWLVPYSD	29775	ITCKFTHNG
29653	DYKYLSCQY	29694	SRQPVPLSK	29735	NAMAIPKEC	29776	ITQHMNHFQ
29654	EEMHLDNAL	29695	STARQPALE	29736	NDAGEPTRE	29777	ITSWMAHPQ
29655	ENHNLENHG	29696	SWYMPPMYC	29737	RAMMRPHGE	29778	KEAVMNHWT
29656	EQTVLHVHH	29697	TIFQPPFHY	29738	RMEYPPEQF	29779	KLFPWEHFG
29657	HNIDLQNWD	29698	VMDTTPLHF	29739	SCTHDPTQL	29780	KLMFWIHHT
29658	IDTLLYHAI	29699	WAVLGPQKF	29740	SSQPVPVPV	29781	KRMQPEHCA
29659	KCEWLQWVP	29700	WSGFHPPTM	29741	STYHTPFNN	29782	KTDHERHAY
29660	KPHPLDPFR	29701	WTVLGPQKF	29742	TCVMEPCIC	29783	KTWLPYHTW
29661	LCSILVHWS	29702	YANFSPLWS	29743	TIQRYPDYG	29784	KWLWNWHDW
29662	MHGPLYCAE	29703	SNTPMPYHG	29744	TNQPFPSSE	29785	LCLSACHFE
29663	MWGGLDMVY	29704	TFVRKPSLM	29745	TQNLEPHWN	29786	LHDNYCHET
29664	NRAQLCRCN	29705	KDGRNPHCL	29746	TSDQYPSSY	29787	LSQDVHHDT
29665	TAMWLHNMV	29706	TCEDDPSLR	29747	TTARRPTYD	29788	MAFFSMHQG
29666	TAVALFEMD	29707	CWMIVPAWH	29748	VAMAHPMTV	29789	MCTDYGHEG
29667	TNVGLIDGS	29708	AAEPIPTWP	29749	VAVDSPHEM	29790	MLKCNFHHQ
29668	VQEWLGLYC	29709	CHCMWPSGV	29750	VLGRSPQVM	29791	MPEPRIHTR
29669	CDGHDPFWY	29710	CSENEPFHR	29751	WTYAPPLNY	29792	MSCPHFHVG
29670	CEEKTPVWC	29711	DMGHCPHDM	29752	AHSHMPVCL	29793	NFEVCQHCE
29671	CIYLKPDYQ	29712	DSDNIPYCP	29753	CKDLSPYKF	29794	NHQNWMHMD
29672	CLQHMPRLQ	29713	DSWRFPGHP	29754	EEMQYPMFM	29795	RFILELHDS
29673	DARSSPDCL	29714	DSYVTPYFT	29755	ETCKMPLSD	29796	RPVNGVHLH

[1871]

29797	SAGQIMHNS	29838	RCGRTAHNH	29879	NFCKGMNPY	29920	IQIRANCPN
29798	SWTWDWHVA	29839	REWYHCHYM	29880	PATKTHDPF	29921	MEGHTGNPC
29799	THQADEHDA	29840	SKYPINHVQ	29881	SALNAAPPP	29922	TPCIWHSPF
29800	TTIQSNHCK	29841	SSEFAVHCG	29882	SWCRYHEPL	29923	AMEIQYMSM
29801	TTMQQGHDL	29842	TINVKHHFG	29883	TFMHESEPH	29924	ASTNEEKSM
29802	WSANQMHSH	29843	WDWFSYHAE	29884	WQYTQNCPA	29925	CIEQKHLCM
29803	YIERTNHTP	29844	WEAKCAHSV	29885	YLGWICAPC	29926	EAGIRAEDM
29804	FRDVCMHDV	29845	WPVLNDHEH	29886	YCECGHWPN	29927	EMSSSNMAM
29805	FANDFTHNM	29846	WPVLNDHGH	29887	MPWLDKPPC	29928	FFCKYNEVM
29806	GQLSTQHHM	29847	YHEKSCHFG	29888	CQMLCRPPV	29929	FHHNSRLDM
29807	ASKQGTHDY	29848	CQVSYQHYP	29889	PIWQGYWPW	29930	GGHSMTKAM
29808	AAEPYKHSK	29849	DGTVCVHMG	29890	YDYWMQAPT	29931	GIVRQCCIM
29809	ASNTYDHSL	29850	GAESCCHFG	29891	EWVVQQIPQ	29932	GIYYEQDGM
29810	CAMPGSHQI	29851	GELMHRHMD	29892	ADHWEKGPF	29933	GRRAELGLM
29811	CDVKDAHHD	29852	HTQHHSHHL	29893	AIDPINRPN	29934	GTMQVTGAM
29812	DGCDMTHTR	29853	ISSNFDHFE	29894	AMEKPTLPI	29935	GYSSTHVQM
29813	DGWLYQHTY	29854	KMQNCAHQH	29895	CTNFVNRPG	29936	IAVQTGSDM
29814	DTMSKNHCT	29855	LSSLHDHPE	29896	DIVNHESPG	29937	KCMPNVGWM
29815	DTYAAGHSA	29856	NTQWEAHSP	29897	GDESGLVPE	29938	KLKPICMKM
29816	EDTLTLHEN	29857	SFSSNAHLQ	29898	GNESGLVPE	29939	LCWMAGAQM
29817	EFTLFQHEN	29858	STECYYHGI	29899	GPKWGRSPL	29940	NAICTLEEM
29818	ENVDRQHHM	29859	TSTHIEHDT	29900	HDAPMFTPG	29941	NCEFWKMFM
29819	EPNMCCHHR	29860	YRVNCGHPN	29901	ICMICATPI	29942	NCVQTNSEM
29820	EYTDVMHDA	29861	AKDLISKPH	29902	KWEHGFMPF	29943	SADSEEGHM
29821	FAAVCHHPI	29862	AMCNDIAPW	29903	LCKVGMEPE	29944	SKMSNQDAM
29822	FDTFSVHRL	29863	DILPKVYPE	29904	MEETVTRPG	29945	SKSDVQSMM
29823	FTDCMGHTM	29864	DLEWQHCPP	29905	MEKRKGQPV	29946	SPLNRNCLM
29824	GAYQPCHYP	29865	DMEPQYWPA	29906	NDLSPAEPT	29947	VWLQSLDRM
29825	GSWHCVHCP	29866	DTLGPEAPH	29907	NMAMAKQPG	29948	WMGHKHKQM
29826	HAAPQCHYS	29867	EAMMMYQPN	29908	RCSKVKWPA	29949	MLTPVQATM
29827	KMEFQLHPG	29868	EIMAAWTPP	29909	RFVVASYPY	29950	KTSNVMDQM
29828	LDGAMSHMC	29869	EQGSVWTPP	29910	RHVSIMPPV	29951	AKCDVTGGM
29829	MHVTPRHPP	29870	EWFHQMSPH	29911	SAVNKCVPG	29952	AYMAYHTEM
29830	MQLIDHHGP	29871	FMEDNTKPR	29912	SETRNNFPY	29953	CTTASMSYM
29831	MSKVDRHCA	29872	HTIWSHDPW	29913	SIMRCKFPN	29954	EIGSRNRRM
29832	MTMLGGHCI	29873	LATQICFPT	29914	STGPILDPP	29955	ERSFHNSIM
29833	NCMYEMHWQ	29874	LRMDCVPPY	29915	TDCCVGPPQ	29956	ESRIGQECM
29834	NFYDSLHCF	29875	LTGPNLSPA	29916	VGNVIARPV	29957	EYFNILRNM
29835	NGNLTAHPN	29876	MGNRSGSPS	29917	YCVCIQRPF	29958	FATWHLYDM
29836	NIELRDHGQ	29877	MPSHWYRPI	29918	GLIKIVKPW	29959	FDCWANNNM
29837	PCHRKKHPK	29878	MTHKDIGPS	29919	HEHTAFRPE	29960	FIWDCGDIM

[1872]

29961	FKTFGSREM	29969	LQIRAWEYM	29977	RYNLQIMWM	29985	VSSSDGWNM
29962	GDDHRVLTM	29970	MLKVQCTHM	29978	SFMLKWWAM	29986	WTAHMYTNM
29963	GGYHKFPVM	29971	PQDFKHNYM	29979	SMFGDRQAM	29987	WTYHSMAQM
29964	GNDIFMVIM	29972	REFCGALMM	29980	SNCDNWSAM	29988	WWWGNFMFM
29965	GNTPCRGLM	29973	RFGREAGGM	29981	SPLHYLVQM	29989	YCNQQQWTM
29966	LIACRENEM	29974	RFTGHMGAM	29982	SYRPFRKMM	29990	YEQVITVMM
29967	LKATYEYVM	29975	RHESINRYM	29983	TLMQPGYQM		
29968	LNDIKWYVM	29976	RSDLKVCAM	29984	VPDHIVGMM		

[1873] 实施例10

[1874] 在脊髓中具有组织嗜性的AAV5变体

[1875] 此实施例描述了使用实施例1-实施例2中描述的方法和系统发现的在脊髓中具有组织嗜性的工程化AAV5变体。

[1876] 分析CNS、肝脏、骨骼肌、心脏、肺、脾脏、淋巴结、骨髓、乳腺、皮肤、肾上腺、甲状腺、结肠、坐骨神经和脊髓组织以鉴定AAV5VP1 581-589区域中驱动脊髓嗜性的氨基酸残基。结果显示于图180中，该图显示了相较于所有其他分析组织在脊髓组织中AAV5 VP1的581至589区域中每个氨基酸的位置频率，其中衣壳多肽具有SEQ ID NO:2的序列。位置(“Xaan”，n＝1-9)在x轴上表示并且氨基酸残基在y轴上表示。特定位置的特定氨基酸残基的频率在热图中作为强度表示。较深的方块表示特定氨基酸在特定位置的频率较高，较浅的方块表示特定氨基酸在特定位置的频率较低。

[1877] 分析相对于给定氨基酸残基出现在所有其他所收获组织中鉴定的变体中的指定位置的频率而言该给定氨基酸残基出现在脊髓中鉴定的变体中对应于SEQ ID NO:1(推及于SEQ ID NO:2)的位置581至位置589的所述指定位置的频率，以确定优先靶向脊髓组织的衣壳的一组序列规则。参考下文表19，和SEQ ID NO:2(AAV5 VP1)，以下氨基酸由此可以在Xaa1-Xaa9任何一个或多个位置处以任何组合独立地突变，以提供与野生型AAV5 VP1衣壳(SEQ ID NO:1)相比具有增强的脊髓嗜性的VP1衣壳，其中脊髓嗜性在此是指优选脊髓转导的特性，而不是优选转导所有其他所收获组织的特性。

表19 脊髓嗜性规则
Xaa1选自A、C、K、Q、R、S或W
Xaa1选自K、R或W
Xaa1为K
Xaa2选自H、I、K、L、T、V或W
Xaa2选自H、I或T
Xaa2为I
Xaa3选自C、F、G、H、I、K、N或R
Xaa3选自F、I或R
Xaa3为I
Xaa4选自I、M、Q、S或V
Xaa4选自I、M或V
Xaa4为V
Xaa5选自H、K、Q、T、W或Y
Xaa5选自T、W或Y
Xaa5为Y
Xaa6选自H、L、N、Q、R、W或Y
Xaa6选自L、N、R或Y
Xaa6为Y
Xaa7选自D、H、P、Q或R
Xaa7为R
Xaa8选自D、F、L、S、T或Y
Xaa8选自S、T或Y
Xaa8为T
Xaa9选自C、I、N、P、R、S或Y
Xaa9选自I、P或R
Xaa9为I

[1878]

[1879] 下文表20提供了在脊髓组织中发现并且符合表19中提供的一种或多种规则的变体AAV VP1衣壳多肽的581至589区域的序列。因此,本公开提供了由具有SEQ ID NO:2的工程化AAV5 VP1衣壳多肽组成的rAAV,其中Xaa1至Xaa9区域是SEQ ID NO:34991-SEQ ID NO:35437中的任一者,如表20中所公开。本文还涵盖由工程化AAV5 VP2衣壳多肽和工程化AAV5 VP3衣壳多肽组成的rAAV,所述多肽在对应于AAV5 VP1 581至589区域中的氨基酸的AAV5 VP2(氨基酸残基445至453)和AAV5 VP3(氨基酸残基389-397)中的区域处具有下文表中公开的序列。

[1880] 表20

[1881] 驱动脊髓组织嗜性的AAV5 VP1衣壳多肽中581至589区域的序列

[1882]

SEQ ID NO	581-589 序列
34991	KYNMLYHMD
34992	KGESQCTSQ
34993	KGQHKENRS
34994	KWHGGPGVV
34995	KRMLYPMCF
34996	KATNADGET
34997	KQSDKWGDD
34998	KIMHSWYND
34999	KKEQDHWDK
35000	KASLGACVE
35001	KGGFYGESA
35002	KQNPSMRQD
35003	KSFSWDSSA
35004	KFLIAFEVT
35005	KEFMPCCHR
35006	KEGITGLEL
35007	KTCTEWPGE
35008	KVNRSNRIW
35009	KVTNYYHSG
35010	KYSPTPFMY
35011	KCEGEEVGR
35012	KQEYIMRQR
35013	KWDDMHTQM
35014	KNETFIAST
35015	KAVTRADAS
35016	KVAEGQNML
35017	KHLDMTSCH
35018	KATRDPKFP
35019	KTLPCSNSE
35020	KFNTKDQLM
35021	KGGLSTTNA
35022	KIQPTNHHE
35023	KTGWDSHWS
35024	KGCVPWSKI
35025	KMEVNMSEL
35026	KTLHMLESK
35027	KTGATNLAW
35028	KCRYKDCQC
35029	KSIMMSQLD
35030	KSGMIQEET
35031	KAIPQGSSC
35032	KGQRRDSVS
35033	KIEYGHWIH
35034	KAHDITWNL
35035	KSAQVYWHP
35036	KIQPMMVKA
35037	KEFQLQMEP
35038	KSTMLEHHD
35039	KRFVDMPEM
35040	KYSPFQEHN
35041	KNCIMLASN
35042	KSFQEQGGA
35043	KTMGIGGGP

[1883]

35044	KVEGYGWTR	35085	GILVQGGSM	35126	CIQPTWDSY	35167	SQIQIRQAD
35045	KPLTYTVGS	35086	TIRPHHAKS	35127	QISKPSVQC	35168	NRIQMTDFQ
35046	KQATNNNGW	35087	AIFHACNSG	35128	SGITFPGSN	35169	FTIFPGNAG
35047	KVEIINWHP	35088	RIVDGNACG	35129	DFIPDWSST	35170	YVIGQKDEM
35048	KSEHHWINA	35089	CIHSQGMVC	35130	EMISNNCHR	35171	IWITCPYAA
35049	KGNPLDGDT	35090	QILQWDCPS	35131	EMINACFQA	35172	NAIPYFVQA
35050	KCDDRNTQH	35091	QIYSLASTW	35132	APIFVCDCC	35173	AVIHEKPHG
35051	KLMYKVMSC	35092	SIKFHERQH	35133	YSIIWRVYR	35174	ACIPQMHSA
35052	KQKTQNRNA	35093	TILQQMEEI	35134	RKIESYCIT	35175	TSIEGLKTM
35053	KCADLIIRS	35094	AIGPDRSNN	35135	EKIMVKRAF	35176	ASIATMDRI
35054	KHEPMCVQL	35095	TIHPAWVTL	35136	ATISKLECQ	35177	PVICVWHDN
35055	KIGCMYSSQ	35096	QISMKGEHG	35137	LKICFSPTV	35178	CWILHHRCS
35056	LIAAQKNFK	35097	EIIQAHLIN	35138	ELIYKIKTD	35179	VTWVHYDMY
35057	QIMQKYQYS	35098	CIIGRQEWL	35139	ARIWNDSMA	35180	YTMVAPDYE
35058	WIPAMTYWC	35099	TIWVHEYGM	35140	DQILPMNGV	35181	SMPVGTKID
35059	WIPAMTYWY	35100	SIAQDQVWS	35141	IHIHFNWFY	35182	MPKVHCASC
35060	WINMPHDRH	35101	PIWQGYWPW	35142	RGIYCYYVC	35183	WCLVNELPT
35061	DIVCLMNST	35102	IIKFLWRPT	35143	WRIYEEDPT	35184	VTSVAGGGE
35062	MIRPLTHCK	35103	DITSLGCIM	35144	YKIGAAECD	35185	QQTVVQFCL
35063	FIIRCMLIW	35104	WINEANCLM	35145	DPILQTYLS	35186	ADVVMQTCC
35064	TISSVFLLD	35105	GIKKPNEDN	35146	DQILSMNGV	35187	GVHVTESGA
35065	EICQHMGMA	35106	EIRQFVKGA	35147	GFIERYGDC	35188	NYCVNKRAV
35066	SIVNHNEAR	35107	TILPKKTLC	35148	DSIIKDQSD	35189	MTSVCNWDS
35067	GIELPVKHH	35108	DIGRFTAYY	35149	LFIKVKDTS	35190	REKVWSDLK
35068	EICSRGRTI	35109	TISQIPWVR	35150	MAIQVFMSD	35191	ICFVNLDHV
35069	GIKYKDVLS	35110	TIVVIPIKS	35151	ISIENEIMG	35192	MKHVSTMTK
35070	GIKFFVDNC	35111	LINSIVCGD	35152	APIQITKAY	35193	SGFVILEGI
35071	GIIERYGDC	35112	GIPLWVFID	35153	EKIRQNAAP	35194	VFTVAHCVH
35072	NICAKHRMS	35113	FIDEKQCHE	35154	ISICAQDID	35195	WFNVRQNGQ
35073	RIQKLKWCQ	35114	DICQKQSPN	35155	VGIMRNDTG	35196	QAPVIPIMA
35074	SIDVPPTAN	35115	NIGVIPHIQ	35156	GTIHMYCGW	35197	TNGVQIMHV
35075	TIVWSECAW	35116	CIDVLFKHL	35157	TRILNEFFC	35198	INMVSMYAY
35076	NIDAVERIW	35117	QIIYRCRST	35158	QNIRLGVPT	35199	QGFVVHSPS
35077	TIGWKELNE	35118	EIHMFGQSE	35159	DVIHTTHSG	35200	QFYVAEHLE
35078	TIMFGCLAE	35119	QIHSEHCAT	35160	SHIQLNAWR	35201	RQCVIGDYD
35079	GIFMEVAWC	35120	HIHMSDRGA	35161	ILIFIQNCP	35202	HTRVVKSTG
35080	RICPLFCIK	35121	AIGCFQNGG	35162	TRIVEPDAN	35203	GTAVSRQSK
35081	LIQSYPHQC	35122	HIKVEHEEV	35163	ATIAQWCVH	35204	QTDVKVNFD
35082	MISAPPDTC	35123	LIEKQRWML	35164	HVIGMGAGE	35205	NWTVHLIYG
35083	QIGPKAFGW	35124	GIWVRDTYN	35165	RAIYNDMTE	35206	TAMVSECKD
35084	DIACLDQWQ	35125	EICMRSGFI	35166	SAICQNYDA	35207	GHLVPFDCG

[1884]

35208	HYVVRHEDP	35249	SMHDYNFNV	35290	MYLNYRHFA	35331	WMPNQSRGW
35209	PTDVITTDQ	35250	GYRNYFIWP	35291	WVHPYSYCA	35332	SATARFRDN
35210	CMEVTCFVV	35251	FAMSYQRQN	35292	DYFQYNEAP	35333	EKRLFDRTY
35211	IQAVMASGW	35252	EVYCYQLQS	35293	RQTTYDCLD	35334	LGHHKGRWH
35212	LESVKMDCN	35253	YKNSYSQWP	35294	AVEDYWRYF	35335	YEWPWQRGL
35213	EMYVQRHSL	35254	SHMSYGQVQ	35295	TSCCYESSS	35336	YEWPWQRWL
35214	AAFVMNGDG	35255	IQCKYIHFT	35296	EPRLYCAQE	35337	YEWPWRRWL
35215	HCWVNEKCE	35256	ANHLYHAAN	35297	HTLLWYGIC	35338	GCLGTKRQI
35216	SKDVYGMYC	35257	NYHNYMFGM	35298	MDHPTYGIY	35339	HDSLFKRTE
35217	EDLVVYMGS	35258	CFNTYDQCF	35299	DGCQKYQEQ	35340	HEWPWQRWL
35218	VGNVMIKMC	35259	SDNMYHTIE	35300	FKFISYVGD	35341	SYLRQQRDT
35219	IHRVVCRDH	35260	RPVHYQGTN	35301	IYLKFYFSL	35342	ALANKRRNH
35220	MFVVNQNWA	35261	MEHFYDRES	35302	DCDMKYLYE	35343	HEGLIMRLE
35221	HEWVLDFFP	35262	PNGHYWMIP	35303	LGSPFYECG	35344	AQRCFHRWG
35222	RVEAYAPTT	35263	ATNKYANAI	35304	LLLNSYKWY	35345	LFCPTCRYA
35223	THCHYQGND	35264	DKTCYLTCP	35305	RLSLWYFCH	35346	HLYSDGRGS
35224	WAGCYTYAN	35265	TMSNYGCLQ	35306	MDRMMYMLG	35347	QCKTHPRSF
35225	WDHNYFTYL	35266	MGMAYSDSE	35307	AMGYQYLYE	35348	LFTDTKRPE
35226	IVMEYAFIP	35267	ELPRYYGQM	35308	SVMLDYHDM	35349	EFTHAIRQC
35227	NSCTYNKDA	35268	DKYLYMYSD	35309	IVKQCYDCH	35350	GATWLFRFP
35228	FKNWYINMH	35269	LLAKYMVHT	35310	QVEDGYWSY	35351	ASWLCGRNV
35229	ETNSYNSHH	35270	MTERYYWCQ	35311	EWLSLYCHW	35352	AMLQKKRQD
35230	LSSCYYMTY	35271	DFMWYHICS	35312	CTEIRYYGT	35353	DWYLSSRDQ
35231	GSTQYPCCK	35272	HGSSYNTVQ	35313	CLHHHYPVT	35354	LHSYQWRLG
35232	QAEGYRDCY	35273	SLTQYQWRG	35314	QYVGMYTQQ	35355	CGHLLRRAE
35233	IGQWYECMT	35274	TVNWYSGFG	35315	QHSRAYMWY	35356	IEFTHSRKV
35234	GNAGYQADL	35275	MEPHYWEWP	35316	TSSDHYALP	35357	PWYNWDRRC
35235	GCMPYQINV	35276	STTAYHQDY	35317	YAMPLYFCE	35358	LCTSTERQG
35236	SLCIYDHRH	35277	VVTQYVDAP	35318	ECVSDYACC	35359	LVDRSGRTA
35237	ETNHYLISD	35278	AVDHYSNGA	35319	GHHQCYKAD	35360	NLDMSRRDF
35238	EGSDYVGYN	35279	NGNHYNAMR	35320	VSLQKYDMT	35361	NAMCFGRAK
35239	NMMQYQWAI	35280	LHKAYVTDS	35321	NQQRTYNNG	35362	VHHKTNRSY
35240	LPKIYLSFC	35281	SGNMYESGE	35322	QEYHIYFPF	35363	TVNRAMREE
35241	QNLMYCSSV	35282	CSESYRGPD	35323	AEDTPYTIY	35364	EAGPRKRDV
35242	RPVTYPHTA	35283	YLEMYSKDW	35324	YSVQAYMAN	35365	CSVWMQRDR
35243	FDKTYFHDC	35284	DPGWYGLAP	35325	QTAEWYCGA	35366	TFVYGERDR
35244	LMEHYLREG	35285	MNSFYRAEW	35326	ECGCMYWGY	35367	LREEELMTR
35245	NVWYYFRTW	35286	NEDQYMHVE	35327	RKWSCHRGH	35368	WHRDEMLTT
35246	WPSFYANYR	35287	WHGFYKIMF	35328	RQTDTHRAN	35369	AMKPMILTN
35247	IACKYERVC	35288	YTWPYPETH	35329	NVDAHERYR	35370	RAAGIQGTA
35248	LMEHYLREW	35289	PGGCYRMME	35330	SRENLGRYN	35371	EQELKNSTQ

[1885]

35372	MATYSDPTS	35389	QQTKNNSTD	35406	HRVHSEKTD	35423	VEQMVWADI
35373	NPPHRNCTI	35390	DTFSNQNTV	35407	RFDQFWDTY	35424	CPWFIKWGI
35374	SVTPDSITT	35391	SAPRNEITR	35408	QSRCVDNTV	35425	PHVSGEGKI
35375	EACHPDLTP	35392	GSKIIMHTE	35409	FNESGWPTW	35426	QAWPKTEYI
35376	ASEHKANTT	35393	MTKMMCETI	35410	ESSQEIKTC	35427	FNWFVRAYI
35377	AYPLENDTD	35394	WCHITWPTN	35411	DQEAQVVTG	35428	DMGMDNKGI
35378	TYYNMWHTK	35395	CVGYKSNTP	35412	TCKNLTYYI	35429	GTMCREQMI
35379	DDYQQHWTG	35396	NCDHRFGTL	35413	ECLLFLVGI	35430	TSVIRLFWI
35380	DTLDMPETN	35397	SSQHASHTD	35414	EMKCIGAGI	35431	YMHFRQYPI
35381	NEQCHAKTS	35398	IMTMQMDTA	35415	QPHEAEFKI	35432	EFVLGCQGI
35382	NEQCHVKTS	35399	IKHITRNTY	35416	YTNFEDVLI	35433	WPNIKHQPI
35383	PYKIVMITR	35400	QFMMGQQTP	35417	PPHIGTQFI	35434	AAMESCAEI
35384	IKHITHNTY	35401	MAMWIADTH	35418	WDNWSNTQI	35435	HQCQSRSYI
35385	STNPNTETR	35402	RKVAVESTW	35419	WVTLCADLI	35436	YSDPKNMSI
35386	HMALVMQTA	35403	IGQHNVQTT	35420	RSLDIQYDI	35437	FFRICGSYI
35387	YDLWGWSTR	35404	MLMHADMTQ	35421	RSFYHPKLI		
35388	IYMSTGNTP	35405	MQPMESATW	35422	DHFYMSYGI		

[1886] 实施例11

[1887] 在乳腺中具有组织嗜性的AAV5变体

[1888] 此实施例描述了使用实施例1-实施例2中描述的方法和系统发现的在乳腺中具有组织嗜性的工程化AAV5变体。

[1889] 分析CNS、肝脏、骨骼肌、心脏、肺、脾脏、淋巴结、骨髓、乳腺、皮肤、肾上腺、甲状腺、结肠、坐骨神经和脊髓组织以鉴定AAV5VP1 581-589区域中驱动乳腺嗜性的氨基酸残基。结果显示于图18I中，该图显示了相较于所有其他分析组织在乳腺组织中AAV5 VP1的581至589区域中每个氨基酸的位置频率，其中衣壳多肽具有SEQ ID NO:2的序列。位置(“Xaan”，n＝1-9)在x轴上表示并且氨基酸残基在y轴上表示。特定位置的特定氨基酸残基的频率在热图中作为强度表示。较深的方块表示特定氨基酸在特定位置的频率较高，较浅的方块表示特定氨基酸在特定位置的频率较低。

[1890] 分析相对于给定氨基酸残基出现在所有其他所收获组织中鉴定的变体中的指定位置的频率而言该给定氨基酸残基出现在乳腺中鉴定的变体中对应于SEQ ID NO:1(推及于SEQ ID NO:2)的位置581至位置589的所述指定位置的频率，以确定优先靶向乳腺组织的衣壳的一组序列规则。参考下文表21，和SEQ ID NO:2(AAV5 VP1)，以下氨基酸由此可以在Xaa1-Xaa9任何一个或多个位置处以任何组合独立地突变，以提供与野生型AAV5 VP1衣壳(SEQ ID NO:1)相比具有增强的乳腺嗜性的VP1衣壳，其中乳腺嗜性在此是指优选乳腺转导的特性，而不是优选转导所有其他所收获组织的特性。

[1891]

表21 乳腺嗜性规则
Xaa1选自C、K、M、Q、R或Y
Xaa1选自C、Q或R
Xaa1为C
Xaa2选自A、F、I、K、S、T或V
Xaa2选自A、S或V
Xaa2为V
Xaa3选自A、F、G、I、K、L、R、T或Y
Xaa3选自F、G、K、R或Y
Xaa3选自F、K或Y
Xaa3为F
Xaa4选自A、I、K、Q、R或T
Xaa4选自A、I或R
Xaa4为I
Xaa5选自I、L、M、Q、R、T、V或Y
Xaa5选自I、M或Y
Xaa5为Y
Xaa6选自H、N、S或V
Xaa6为H
Xaa7选自A、H、I、N、S或Y
Xaa7为N或S
Xaa7为N
Xaa8选自A、C、D、G、H、M、Q或S
Xaa8选自G、M或Q
Xaa8为G
Xaa9选自A、E、L、W或Y
Xaa9选自A、L或W
Xaa9为A

[1892] 下文表22提供了在乳腺组织中发现并且符合表21中提供的一种或多种规则的变体AAV VP1衣壳多肽的581至589区域的序列。因此，本公开提供了由具有SEQ ID NO:2的工程化AAV5 VP1衣壳多肽组成的rAAV，其中Xaa1至Xaa9区域是SEQ ID NO:22118-SEQ ID NO:23117中的任一者，如表22中所公开。本文还涵盖由工程化AAV5 VP2衣壳多肽和工程化AAV5 VP3衣壳多肽组成的rAAV，所述多肽在对应于AAV5 VP1 581至589区域中的氨基酸的AAV5 VP2(氨基酸残基445至453)和AAV5 VP3(氨基酸残基389-397)中的区域处具有下文表中公开的序列。

[1893] 表22

[1894] 驱动乳腺组织嗜性的AAV5 VP1衣壳多肽中581至589区域的序列

[1895]

SEQ ID NO	581-589 序列
22118	CIKTDWQRI

[1896]

22119	CAVSGPECY	22123	CGYWLAVPW	22127	CQSTGAHME	22131	CYNSGDCGN
22120	CCEIWLASS	22124	CIPTYHHFG	22128	CQSTGSHMA	22132	CIMSHCRQH
22121	CDYWLVVPW	22125	CIRGQTHYN	22129	CSKATGGAE	22133	CSNKTCCKA
22122	CFTQKESAY	22126	CNTPYEVWA	22130	CSKTTWGAE	22134	CGYWLVVPW
22135	CHMLNDCMS	22136	CSKATWGAE	22137	CKNPSWYMD	22138	CFEMYYAMW
22139	CDPCMECGS	22140	CTDGKHASW	22141	CFCTYKIEH	22142	CIGPHWHIP
22143	CAHQQHNED	22144	CAHSYQMSR	22145	CAIHIFLQD	22146	CAIQISDTS
22147	CALSALTME	22148	CALVSRSQD	22149	CAMIHGKYD	22150	CANFAMQQF
22151	CASRSMLTE	22152	CAVRTFMGS	22153	CCDDCPPNR	22154	CCGIFNGIY
22155	CCHMQYRES	22156	CEMYREYFI	22157	CFAKMEKLQ	22158	CFAPYMNVD
22159	CFMHYGRIA	22160	CFSQFASGS	22161	CFYQKGAAY	22162	CGFWLDQNE
22163	CHRLVMGYH	22164	CIAPRPEWP	22165	CIKLKCSHS	22166	CILWKHAMP
22167	CISNTRFGV	22168	CITMPMNFG	22169	CIYIWFTRD	22170	CKSGMVHHA
22171	CKSMGLFLV	22172	CKYDYHQNT	22173	CLGYMWQNT	22174	CMKADGMQR
22175	CMSDRGLQS	22176	CMYMQNAMG	22177	CNNLSVCDE	22178	CNSLFTNQQ
22179	CPTMYKMQA	22180	CQKCYDHYQ	22181	CQYRTRMIT	22182	CSAPTLSGN
22183	CSFHIIMGE	22184	CSFHNNEHA	22185	CSICEHCMD	22186	CSMAEQKSQ
22187	CSMRYMAVH	22188	CSRPLHDSY	22189	CSTQYMKHM	22190	CSVNFYTSF
22191	CSWEWFRAE	22192	CSWQLWQSC	22193	CSYSTQIGH	22194	CTAITLRYP
22195	CTERTESNR	22196	CTHDIWAQA	22197	CTHQIHLPA	22198	CTKTYMHDF
22199	CTLNVQSDG	22200	CTLQWSAGP	22201	CTMRVQMKP	22202	CTVLRFDHQ
22203	CVAYSYNST	22204	CVCGKKQMG	22205	CVFAQSGSY	22206	CVFYQRKVL
22207	CVIAKYYQE	22208	CVIHTPYQY	22209	CVIRDGHYD	22210	CVKMKQDFN
22211	CVMEQQKER	22212	CVNTTHDYQ	22213	CVNVMAADR	22214	CVQITLDGF
22215	CVRMCVDHA	22216	CVSMIVPMV	22217	CVTCKFDEQ	22218	CVTLTHRDK
22219	CVTYSSAPK	22220	CVYPTMKAD	22221	CWICMCVSP	22222	CWKFRDHDF
22223	CWLHTRTHW	22224	CWTTANTSN	22225	CYAKVTANT	22226	CYEQWNGLA
22227	CYFLHNNHE	22228	CYYTIEKAE	22229	CAVGGPECY	22230	CKTQERANS
22231	CYNSGDCDN	22232	CTSTDPIAP	22233	CRGKSHDHT	22234	CPALNYDCV
22235	CMLTHIMEF	22236	CTFPSADMC	22237	CQSTGSHME	22238	CSFHMNITS
22239	CIPTYRHFG	22240	CYGRVAHHE	22241	CTEIRYYGT	22242	CVKPFLTGW
22243	CAWYMKQGN	22244	CAEPVYMWK	22245	CAGTSMAKE	22246	CASHVKWGP
22247	CAVAYMSCT	22248	CAYKPHVAP	22249	CCHNEHRHM	22250	CCVSFAYSC
22251	CDAVHNDAS	22252	CEYLPDMHK	22253	CFHWVNHMT	22254	CFQTANIGG
22255	CHDAQFKTD	22256	CHKSWWRLD	22257	CHYWVSTDG	22258	CIAHRHAQP
22259	CIFRLSALS	22260	CIGSRQAIA	22261	CIHIASRLG	22262	CISFYTSAG
22263	CITKLMVNA	22264	CIVTDSMAP	22265	CIYYPNFYN	22266	CKQMYLATE
22267	CKVPVFSLE	22268	CLRVCNCSP	22269	CLRVYNCSP	22270	CMAISPSFE
22271	CMLRYMLGR	22272	CMPQKYHKD	22273	CMSAMCMWS	22274	CMSKIVWSS
22275	CNDQAQYDW	22276	CNTSMAIMG	22277	CPWWSDNCM	22278	CRVTITYPT
22279	CSFQMQFMG	22280	CSFVLDNAF	22281	CSGPVSASA	22282	CSKYFNDWG
22283	CSKYSAEYP	22284	CSTPHHLQL	22285	CSVNIGQHA	22286	CSYKDQSQQ
22287	CSYTHWRSA	22288	CTAHPMLQA	22289	CTATKILHA	22290	CTFVALSMQ
22291	CTGSIVKQM	22292	CTHPRITGA	22293	CTIKQLCNG	22294	CVANCQQCI

[1897]

22295	CVARNINAN	22336	YVHEGGENP	22377	DVDNCYKGD	22418	SVIPLPHGD
22296	CVCMYQNMQ	22337	YVLESNYCQ	22378	DVICFGKWE	22419	SVITTEFQD
22297	CVFLAQKMK	22338	YVSASWMTP	22379	DVQTRNLSK	22420	SVVNHAGAC
22298	CVFMRSHDK	22339	YVTKFAECH	22380	DVYRDAMTP	22421	TVAHERSFQ
22299	CVFPVHSSN	22340	YVVTHVPWH	22381	EVALQWAST	22422	TVDNTNWLK
22300	CVHKICMHH	22341	YVYDEFLWS	22382	EVDVRVLFD	22423	TVHYSSQDP
22301	CVHQNSRAK	22342	FVMCYNDCK	22383	EVESWMWVC	22424	TVSGLWQEL
22302	CVHRGLTYT	22343	AVHWHRKWE	22384	EVHRVASCP	22425	VVGTYALTY
22303	CVHVAMHYP	22344	AVINPPCHP	22385	EVSKFNHYF	22426	VVHTCAMCT
22304	CVHYANMML	22345	AVSLQPQFG	22386	EVSTITFWS	22427	VVSAHMPGV
22305	CVISKMWCH	22346	DVKAFMGCV	22387	EVTPHWTGG	22428	VVTVAWTPG
22306	CVLRSNVQL	22347	GVEAYRNFK	22388	EVYYAMASW	22429	WVEKYLISY
22307	CVNEFFTCF	22348	GVIHRQYCP	22389	FVNVMKVTC	22430	YVLSAIPFG
22308	CVNMKTDGN	22349	KVEANPRVC	22390	GVKPGASCG	22431	YVWYRRTDL
22309	CVNQLCDSQ	22350	LVHAVSYIC	22391	GVQQFYRYD	22432	KVCNSEYWD
22310	CVNVENLTA	22351	MVAPPNCCV	22392	GVRLSSPWE	22433	QVHIKVTDH
22311	CVPDTIDTE	22352	NVNVRESEW	22393	GVYPMHEAC	22434	KVDRNTNWC
22312	CVQMIMKDW	22353	PVHAVSYIC	22394	HVCMMASCD	22435	GVEAYCNFK
22313	CVQTDNRRP	22354	RVTIMFTDT	22395	IVCAYSNAS	22436	HVWNDVFHM
22314	CVRHYGEYW	22355	SVHTADVCC	22396	IVEPAIVCA	22437	EVQRNIYSP
22315	CVVADQGRG	22356	SVPQQKIDE	22397	IVNKWDGFA	22438	GVAQYLHHC
22316	CVWHAPIQT	22357	TVKINSISP	22398	IVYYCGKRK	22439	DVPIHRVQL
22317	CVWHASIQT	22358	YVQCTHYAQ	22399	KVICHEFTS	22440	SVNLNCTEH
22318	CVYTKIIGH	22359	GVIHRQYCQ	22400	KVMHTGWHP	22441	FVFAMIPGF
22319	CWEFMFWHG	22360	QVPRPSMWW	22401	KVVQRIDHD	22442	AVHWHRKCE
22320	CWGVISKAP	22361	PVSPLRSYV	22402	LVCRYWDDE	22443	GVMIYEADE
22321	CWIRGTQDR	22362	QVTSDGKII	22403	LVFMQEYAA	22444	KVAGGHETT
22322	CYLITFREH	22363	VVGWDNYVA	22404	LVKCNTCME	22445	FVPWCHSFG
22323	TVADSSRLM	22364	NVHHEWCDW	22405	LVNHCGQMG	22446	NVTPGRTHS
22324	TVCMFQVLP	22365	MVDPMSVPF	22406	LVTEQVNSP	22447	LVSQWDYDM
22325	TVFPMTVVP	22366	DVDYRSEAW	22407	MVDRMKWHK	22448	RVDIGGFAQ
22326	TVGFSEQDA	22367	IVPQQKIDE	22408	MVYPIEWGA	22449	HVDQHRVQA
22327	TVIQWYWHH	22368	VVHPMAYCF	22409	NVGIITFQD	22450	AVEACYHSP
22328	TVLSWGLCN	22369	LVVTKFTYV	22410	NVLQNSIMC	22451	AVFTFATCG
22329	TVNQECNCQ	22370	AVKSRTTWV	22411	NVYTHHVLS	22452	AVQWSAENM
22330	TVRPQDNKN	22371	AVQGDPGVA	22412	QVDNYFNRL	22453	DVDYRPDAQ
22331	VVDHSCSRV	22372	AVSVETWWT	22413	QVNFGLHAQ	22454	DVELGGSNF
22332	VVDPFGHEQ	22373	AVVDCPLNI	22414	RVNEATHDD	22455	DVFQYEMLA
22333	VVHENQLKP	22374	AVWNVVSDG	22415	SVCPFSYNL	22456	DVGYASDPK
22334	WVTCRGTVF	22375	AVYQYVETP	22416	SVFSNQDGV	22457	DVKVTEDAE
22335	YVEIHHAQK	22376	DVDKLEHGQ	22417	SVGFEQSWY	22458	DVRLDDGAQ

[1898]

22459	DVRPDEFHN	22500	TVDPNVRVD	22541	KAFHIQREK	22582	VAFYQENTT
22460	DVRPDKFHN	22501	TVVPQGFDR	22542	LDFQIRFER	22583	WHFSSKWMD
22461	DVVSTESTT	22502	YVCIPAHDR	22543	LPFHLQSDT	22584	TPGITQQQW
22462	EVFREWMWM	22503	YVCMIHMDD	22544	LPFTQMHGS	22585	TSHICVSSL
22463	EVGSDCHQG	22504	YVDLIMSVD	22545	LSFPVVLTE	22586	TWLIDADMA
22464	EVKYMRHIC	22505	YVISKCAGQ	22546	NAFEYEKIM	22587	VAMIMEGQF
22465	EVRPSVGSD	22506	TSFCQIASR	22547	NAFTKTQDQ	22588	VFKIREDDN
22466	EVTSAGFDF	22507	VCFKLFHVE	22548	NSFSTNWSY	22589	WADIEYSHD
22467	FVIWRQESR	22508	AFFKFMTQM	22549	NTFCETVTF	22590	WTGIKCQSS
22468	FVRCLSTHN	22509	DKFPMPTHS	22550	QKFEGEWFT	22591	YFCILEKMT
22469	FVSEQMTAQ	22510	ETFGIRTFA	22551	RAFVGCSPE	22592	YMCIQMAGW
22470	GVDSSRPVQ	22511	GIFWASYGF	22552	RQFAHRFKM	22593	ESPIMLNQA
22471	HVCQKTHSI	22512	GKFPMPTHS	22553	SCFAILSAN	22594	RTPIELMHG
22472	HVNTSCKQF	22513	GMFCASYGF	22554	TSFELNDER	22595	AGQITTSVT
22473	HVQHHDAAT	22514	KDFWGHQCF	22555	TSFSPELFN	22596	ESDIFDGRQ
22474	IVAYGTCGA	22515	LLFTLMSSS	22556	WNFNCIILH	22597	HPSICKGGD
22475	IVHQKQLAA	22516	QTFYEGHAT	22557	STFISSGIQ	22598	QCHIIVARE
22476	LVDEMFSQW	22517	RDFRMNAGA	22558	FKFWDMETR	22599	QCHIIVARG
22477	LVKGSFAPD	22518	RTFYEGNAT	22559	KRFVDMPEM	22600	TAKIIMLAS
22478	LVKPQSGEL	22519	TTFQSLQQM	22560	VSFQRMNDM	22601	DMSIATWYP
22479	LVRQDSQIG	22520	NQFHWVRYD	22561	ERFKTFCEA	22602	ESDIFDGHQ
22480	LVVCSCPFP	22521	LWFFMKFAK	22562	LLFPLMSSS	22603	TCVIGKGRR
22481	MVEFYCSDE	22522	YTFWDYHRS	22563	EKFWFKTMN	22604	VACIEHAKC
22482	MVEFYYSDE	22523	KDFWEHQCF	22564	NWFQAIEIG	22605	QFDIDQCCT
22483	NVAYHMVNN	22524	VHFTVLNWG	22565	GIFCASYGF	22606	HPSICQGGD
22484	NVENAYDAM	22525	DYFQYNEAP	22566	IWFWWEKAE	22607	AWIISQEGR
22485	NVFLNLDLA	22526	QEFLEERYV	22567	RTFYEGHAT	22608	DMWILKDAS
22486	NVHTIENSA	22527	GTFHCCAWE	22568	TTFPQVIMN	22609	DYYIIDCFM
22487	NVVDILAGS	22528	ETFGIRTFD	22569	DAFINRLSG	22610	EHMIEFPHR
22488	NVYYKHAHP	22529	EQFADRAPR	22570	DIFDRPSSD	22611	FREIKEYDM
22489	PVIRTWVDI	22530	ARFMWNHTR	22571	ETFLPWVHT	22612	GLYITHSME
22490	QVASHTQPE	22531	DTFLEMSMG	22572	ETFSKTWFG	22613	HCNIFSREQ
22491	QVATWSMYH	22532	DTFPIAGSF	22573	KSFGSPFER	22614	IICICNRGH
22492	QVFMASTGE	22533	EAFRHDHFH	22574	KSFGSPFEW	22615	KEVIMKRCV
22493	RVGVQEAVC	22534	EGFFYHHGG	22575	MAFYKSQFN	22616	KLEIMCQAK
22494	RVHISLLKG	22535	EGFQFSNQG	22576	NAFGMEYNT	22617	KPGIMKFYN
22495	RVNMKTDGN	22536	ERFMTASMQ	22577	NSFKPSGVT	22618	LYNIPMGTT
22496	SVCLHNLVG	22537	FEFTVHCTH	22578	PWFMVIWWP	22619	MKDIRVTQR
22497	SVHGFDQSM	22538	GKFEPPVNA	22579	QCFFNHRCE	22620	MYTIGTCYG
22498	SVIKVAFFF	22539	HAFPTVRWG	22580	RAFQDHYAY	22621	NINIYWMID
22499	SVRLHNLVG	22540	IMFDPNIPN	22581	TIFSASDSK	22622	NSAIMSNYM

[1899]

22623	QEDILQTQY	22664	TTLIKMHVE	22705	AEEVYGYHQ	22746	WNSPYKKRY
22624	QFEICHSQE	22665	TWEIKHNQV	22706	AQMEYDCAM	22747	YQVHYASHG
22625	QHLIMPEPH	22666	VACIANLWT	22707	DDALYAVFY	22748	KCTEYVHTC
22626	QPLIQHAGG	22667	VFAIMTRTQ	22708	DSWMYHVQL	22749	PIHTYKRHP
22627	QQTIDSYGH	22668	WSNIYWRQD	22709	DYIHYKNGM	22750	RMKWYLPCV
22628	QSGINDLYG	22669	THGHYTMSA	22710	EFSYYHTHG	22751	FAMSYQRQN
22629	RCMIYHQNA	22670	TRAGYNEQD	22711	ESEKYTWSG	22752	EASWYYPAI
22630	RGGIFPQRP	22671	TSVKYSAWG	22712	FGATYNQTI	22753	MEPHYWEWP
22631	RMCIDYLRP	22672	TYWNYAEAK	22713	FGLWYRDHK	22754	MSTDYCKVN
22632	TACIQQKLL	22673	VFTDYWEQG	22714	FHITYLMDK	22755	TSKDYNEEI
22633	TNRIVTRII	22674	VFTDYWERG	22715	GAQQYADSI	22756	MTERYYWCQ
22634	TQQILHDAF	22675	VIDWYPHQQ	22716	GFCHYWKQG	22757	QTPKYFVHQ
22635	TRIINVILG	22676	VSDEYVMEQ	22717	GNCKYHVTQ	22758	YSKQYAEYY
22636	TSLITGQLY	22677	WELTYDCLA	22718	HIVTYEHHE	22759	HSDWYMKYP
22637	VEPIMFWPA	22678	WIICYIALQ	22719	HPCYYSWAP	22760	AEGYYVFQP
22638	VNYIICKMD	22679	WLIVYDHGN	22720	HWDNYWKMD	22761	AWGCYQKGY
22639	VNYIICKMG	22680	YAHSYGHFD	22721	HWREYPSKG	22762	DIKMYWEGA
22640	IKHITHNTY	22681	MAEQYKYGH	22722	ICHQYTRHG	22763	FGILYNNNE
22641	MYAIADELD	22682	TGAKYHDWL	22723	IPLAYPEYQ	22764	FNQLYMTFH
22642	NPSIQPKTM	22683	EAENYFPRH	22724	ISKAYVDWW	22765	GAEGYSKDQ
22643	PENIAEDHE	22684	GLMDYSLCA	22725	IYHPYDMLG	22766	GAEGYSKGQ
22644	MRKIKIACM	22685	KEIQYHSYE	22726	KQMEYDNFH	22767	GINDYKSGN
22645	LRAILDCGW	22686	MDTCYMYAD	22727	LMIQYVNKH	22768	GMDFYFNLE
22646	EIAIDQDIH	22687	QPPKYFVHQ	22728	LMREYMSVP	22769	HKMAYGRDQ
22647	SRTICVHGN	22688	RNTPYEFWA	22729	MCQWYATNL	22770	IAYEYWDYW
22648	IANITACAN	22689	SGNMYESGA	22730	MSDYYKEDS	22771	IKEGYQLGG
22649	EISIRLLDE	22690	TEPHYWEWP	22731	MYGGYAAGF	22772	ISQFYLLDQ
22650	EIYILPKFP	22691	TTERYYWCQ	22732	NFCGYQAMH	22773	KECNYQARG
22651	GSAIFVKEE	22692	KEIPYHSYE	22733	NGLQYMSSA	22774	KPLQYEAIQ
22652	GYTIKNGYN	22693	LMAWYENWY	22734	NYHYYDHAG	22775	LPNPYKTVV
22653	HHEIANLHQ	22694	ITRTYQEMQ	22735	PTVAYENLD	22776	LPVQYEDRD
22654	KHQINAVER	22695	VREYYPKFH	22736	QAIGYNYFW	22777	MCDNYKSLA
22655	NGEIFSCMS	22696	QQEAYILRI	22737	QGQYYYECG	22778	NHGNYEIAW
22656	NGEIFSYMS	22697	DMTFYLSRD	22738	QWPLYVMDR	22779	NTMPYGHSY
22657	QFKIYDCVG	22698	MMPEYFNKD	22739	REELYQCAP	22780	NYIPYGRLA
22658	QIGITYYLP	22699	FITAYHPCL	22740	REKNYELSE	22781	QTQTYAYMD
22659	QMGINHDDR	22700	GQKSYDGCH	22741	RWYWYESCN	22782	RFDHYNNEQ
22660	QPAINTQCA	22701	MFEKYFQTM	22742	SMTSYYRQQ	22783	RKEYYPNVQ
22661	SFKILNRGG	22702	NEAPYSTWQ	22743	SWVWYPNLC	22784	SAYQYLGCH
22662	SPDIDLINH	22703	SMSEYHYDE	22744	WATAYWQGI	22785	SCGPYLWGT
22663	STSITELAM	22704	AAVQYSLMG	22745	WCGFYFLAA	22786	SHGNYEIAW

[1900]

22787	SMNLYNRGV	22828	DHDTNHWSH	22869	NGYLVHACQ	22910	DITGHCNPA
22788	SMPPYRMLS	22829	EASTFHRQG	22870	SDGMVHNCW	22911	EAINRVNCQ
22789	SYQPYSKTV	22830	ECAETHFMK	22871	KCGVEHEHR	22912	EHEHRINQC
22790	TANCYPILG	22831	ECHVGHSGC	22872	QATNEHIFW	22913	EHRQLVNNN
22791	TSWMYYDYG	22832	EMCSVHQAW	22873	DTHHTHWQY	22914	EKIPHYNME
22792	TYCLYGISE	22833	FCKFMHETH	22874	FTQAIHHPD	22915	EKWDMFNTE
22793	VALHYDHPN	22834	GFCQLHTQE	22875	LEIVMHSLP	22916	FTYCVNNYQ
22794	YFQLYMKEG	22835	GPSLIHQAA	22876	GAKYRHYLQ	22917	GKTCAENWI
22795	YWAGYHMWP	22836	HYETEHDGC	22877	APNFCHCSF	22918	GYRPELNSA
22796	YYTCYPPSM	22837	IAKASHKDR	22878	ATHDMHRLC	22919	HFDVICNCY
22797	TTHSDHHHA	22838	IASGTHHCQ	22879	DDKCDHLQH	22920	IMALDQNTN
22798	TYNLEHQCS	22839	IIEQSHTLE	22880	DTKWMHQHC	22921	IMEPMNNAF
22799	TYRLEHECI	22840	KLDLTHRMQ	22881	DTLDWHMAY	22922	IRMYAYNWL
22800	VCQKVHGLF	22841	LFTKQHVYS	22882	EGVPLHRAF	22923	LESQQSNYT
22801	VLDVIHWHW	22842	LKHFVHAAS	22883	VSMFNQNSC	22924	LIGPQANLY
22802	VSEETHNFR	22843	LTLGNHGEY	22884	WAHYTINQQ	22925	LTSMQENHS
22803	VWNWEHKCT	22844	NAGCQHEYE	22885	YHRHIANNG	22926	MIQYNKNMG
22804	WYQCLHYQN	22845	NQRVEHAFS	22886	YLHNQGNGL	22927	MSEPRYNDL
22805	YAWCQHCKN	22846	NTEREHRYN	22887	YRGAACNPH	22928	NCEWEMNDR
22806	YHMDVHDSM	22847	QIVHDHRHR	22888	DLGMINNME	22929	NRSMFKNVR
22807	YNMLTHVSM	22848	RYDCMHCYA	22889	DEKAEKNES	22930	PTDFNENDC
22808	ETTLAHRWE	22849	SAGEIHEHF	22890	FASMQANDR	22931	QFYSMDNHD
22809	HSAARHMFV	22850	SLDGVHDGF	22891	GGHLEGNWQ	22932	QQEVKKNNA
22810	LAECKHSSA	22851	SLHYWHSAN	22892	HITMWYNGD	22933	RNCMDTNGM
22811	MFMMMHDIW	22852	SSNYWHASS	22893	HRQESPNDK	22934	RWADVTNMS
22812	MFMTMHDIW	22853	STKAHHWCQ	22894	NEIFNQNSY	22935	STIYAFNWA
22813	NLNNWHTDM	22854	TEAATHWHR	22895	NGVALENRG	22936	TCMQCSNMP
22814	QNNWGHYNK	22855	TEDLSHLWW	22896	NPWFLDNQF	22937	TWAPHNNGM
22815	QTLGGHMCT	22856	TFQQSHAAV	22897	QMTSSTNCV	22938	VAKTSWNHC
22816	SSNFNHHIY	22857	TISLVHLSE	22898	TKQQLINSN	22939	WICPKINAP
22817	MTAKEHAYL	22858	TMDVLHCDY	22899	NEISNQNSY	22940	WSILGTNWG
22818	HQVLDHQLQ	22859	TWCYRHNVE	22900	VECTDTNDE	22941	YDPHNNNGV
22819	DMRRAHQEK	22860	VSSQGHRFY	22901	HITVWYNGD	22942	YSVARCNMD
22820	VNQMMHTQA	22861	VWMTWHASY	22902	RNNWKVNSQ	22943	YSVCHMNAY
22821	MFSFKHVHW	22862	YHAPQHHMG	22903	YCCSDTNTE	22944	YIRVGGNNE
22822	AQASAHIMS	22863	TQKGNHCRW	22904	QMTSSTNCA	22945	FQIRQDNLA
22823	DAIPRHICR	22864	APRNVHCGT	22905	PACQKGNHA	22946	KIDAHSNME
22824	DCSQQHMTL	22865	LMEYCHDRF	22906	ACLWQRNDE	22947	NSVALENRG
22825	DEGVEHRPM	22866	NYDFLHHVE	22907	ATVQFANDG	22948	KNTMCDNDA
22826	DEMLQHSIM	22867	NCIAKHEVT	22908	DALTHGNFE	22949	ALKSWMNGS
22827	DFHNAHFGS	22868	NSAHVHYFH	22909	DHAKLNNPS	22950	ILIFIQNCP

[1901]

22951	QMKPLINDE	22992	DRSEFRQGT	23033	AGQFKSEGS	23074	DWKWKTMMA
22952	TISHHLNRE	22993	DTPDFNTGI	23034	TGACMVEGL	23075	DYRGQDHVA
22953	DHLFNNNEW	22994	ETEGQVAGA	23035	TLVSCDQGN	23076	EFTTCVELA
22954	ETGYRENPC	22995	ETNLHTKGH	23036	DGGLRVQGM	23077	EICFKEYWA
22955	TSDTMWEGD	22996	HALSLTWGD	23037	WKVSPIQGF	23078	EQIGVWWHA
22956	TTVRKQFGY	22997	HSLPHTIGL	23038	EGGKNIFGM	23079	ETMWNMWDA
22957	VGGSVNRGG	22998	IHSRVNAGH	23039	LDGPIGMGT	23080	IMITEVRSA
22958	VLYYSNYGH	22999	IQKVTKRGA	23040	EPLYSWVGI	23081	KQETRGSVA
22959	WAMFMYEGA	23000	KEGPTPVGH	23041	TNIKERTGS	23082	KSCTTSHCA
22960	YAQCFGQGI	23001	KFDFWVAGH	23042	TIDVKRWGQ	23083	LSHQECAAA
22961	YAQCFSQGI	23002	KHTQQDSGS	23043	TSHLAWCEA	23084	LTNDKSVTA
22962	YLRSLKCGN	23003	LATHQKDGR	23044	VAMMGCMQA	23085	MCRKNKILA
22963	NGCTTTKGE	23004	LMELTNEGW	23045	VCENEAGVA	23086	MEGSDNLSA
22964	DPKWEQYGM	23005	LSPKNDPGV	23046	VMCQTLVKA	23087	MHEGVAGIA
22965	EFVMGCQGI	23006	LSYSIVTGE	23047	VSLVRYVDA	23088	MNEYMIKCA
22966	FCQAINIGH	23007	MAVTLMHGC	23048	YPCFGTSLA	23089	MYGFVMWEA
22967	GCMTRLDGL	23008	MCVRWSVGD	23049	AHTWLQHWA	23090	NIYATGSQA
22968	HFHEVAGGS	23009	MDTHLQHGA	23050	DAGWQQFAA	23091	NQVHSVDLA
22969	HHMHIWLGL	23010	MGGSNVWGN	23051	DTEQWEATA	23092	NYTPFYKRA
22970	HTVGMAEGD	23011	MKWDCLIGP	23052	HQYELQMHA	23093	PEYWGQKIA
22971	IQAVMVSGW	23012	NFVCSDIGK	23053	LGYHCCQAA	23094	QFNFTESDA
22972	KSMPPGQGY	23013	NLMRRDHGG	23054	NSTPMACMA	23095	QTDWKVVVA
22973	SRDTWQIGR	23014	QAMYTTAGI	23055	QIEYCMFYA	23096	SAHDTIRDA
22974	TIGVKRWGQ	23015	QLKNAFEGS	23056	SDDDREREA	23097	SFMVMACYA
22975	VLMFNRGGS	23016	QQYAPVIGP	23057	SHDLVAAWA	23098	SICGTCQRA
22976	VLMFNRWGS	23017	QSREPRGGK	23058	THAWLQHWA	23099	SSSCVVIAA
22977	DPKWEKYGM	23018	SAVGWGVGQ	23059	SAQCAGHNA	23100	SSTTVWEYA
22978	TNHCRDAGT	23019	SYIFSEIGT	23060	VLPEGTMTA	23101	STGTNAEHA
22979	KWTDCIIGY	23020	SYMVIGWGG	23061	LGMQADRFA	23102	SYQWLCADA
22980	DTDKQGQGI	23021	TFTYAYDGQ	23062	LGYYCCQAA	23103	TSEGITREA
22981	MTRQPEDGF	23022	TKKRNCFGM	23063	KYAYSLDCA	23104	TSSHQTVKA
22982	ASDKMDVGS	23023	TSIHDNQGM	23064	GTEQWEATA	23105	VCNMTDLQA
22983	QPEQGGWGT	23024	VATSKDVGP	23065	WSKGTMLEA	23106	VKESSDACA
22984	KSMPPGQGD	23025	VNRPTNFGT	23066	KGIWQGMDA	23107	VMEQLQHTA
22985	AACHFTYGS	23026	WPQRVPLGM	23067	NIRTGACSA	23108	WCTPMNRVA
22986	AATCSLKGH	23027	YFPFHMAGQ	23068	KSEHHWINA	23109	YKQSHMVAA
22987	AHADMTTGF	23028	EIRMIISGC	23069	FSEQNVSAA	23110	MSWDMVEAA
22988	AMQRSVAGI	23029	GRHYIAAGY	23070	ASYYLVRNA	23111	QFRCMMDDA
22989	AYEHCEHGP	23030	MLAFDPMGR	23071	DGERFMMQA	23112	EIVQPQCYA
22990	DIADRNVGH	23031	HHMPIWLGL	23072	DIGLRNMNA	23113	ALQFVSSTA
22991	DRMCMMCGV	23032	HFHEVAWGS	23073	DTIQFVQNA	23114	TERRMFRTA

23115	QTPKSLLDA	23116	AEVMQWQAA	23117	EADTFTLIA

[1902] 实施例12

[1903] 在肺中具有组织嗜性的AAV5变体

[1904] 此实施例描述了使用实施例1-实施例2中描述的方法和系统发现的在肺中具有组织嗜性的工程化AAV5变体。

[1905] 分析CNS、肝脏、骨骼肌、心脏、肺、脾脏、淋巴结、骨髓、乳腺、皮肤、肾上腺、甲状腺、结肠、坐骨神经和脊髓组织以鉴定AAV5VP1 581-589区域中驱动肺嗜性的氨基酸残基。结果显示于图18E中，该图显示了相较于所有其他分析组织在肺组织中AAV5 VP1的581至589区域中每个氨基酸的位置频率，其中衣壳多肽具有SEQ ID NO:2的序列。位置(“Xaan”，n＝1-9)在x轴上表示并且氨基酸残基在y轴上表示。特定位置的特定氨基酸残基的频率在热图中作为强度表示。较深的方块表示特定氨基酸在特定位置的频率较高，较浅的方块表示特定氨基酸在特定位置的频率较低。

[1906] 分析相对于给定氨基酸残基出现在所有其他所收获组织中鉴定的变体中的指定位置的频率而言该给定氨基酸残基出现在肺中鉴定的变体中对应于SEQ ID NO:1(推及于SEQ ID NO:2)的位置581至位置589的所述指定位置的频率，以确定优先靶向肺组织的衣壳的一组序列规则。参考下文表23，和SEQ ID NO:2(AAV5 VP1)，以下氨基酸由此可以在Xaa1-Xaa9任何一个或多个位置处以任何组合独立地突变，以提供与野生型AAV5 VP1衣壳(SEQ ID NO:1)相比具有增强的肺嗜性的VP1衣壳，其中肺嗜性在此是指优选肺转导的特性，而不是优选转导所有其他所收获组织的特性。

[1907]

表 23
肺嗜性规则

Xaa1 选自 A、E、K、M、Q、R、S 或 T
Xaa1 选自 A、E 或 Q
Xaa1 为 E
Xaa2 选自 A、I、K、S、T 或 V
Xaa2 选自 S、T 或 V
Xaa2 为 T

Xaa3选自A、E、K、M、Q、R、S、T或V
Xaa3选自A、K、R或S
Xaa3为R
Xaa4选自M、P、R、S或T
Xaa4选自P、Q或T
Xaa4为Q
Xaa5选自I、K、L、M、T、V或Y
Xaa5选自L、M或Y
Xaa5为L
Xaa6选自D、G、H、M、N、R或S
Xaa6选自H或N
Xaa6为N
Xaa7选自A、K、M、Q或R
Xaa7选自A、K或R
Xaa7为R
Xaa8选自A、F、G、S、W或Y
Xaa8选自A、F或G
Xaa8为F
Xaa9选自A、E、G、P、R或Y
Xaa9选自G、P或R
Xaa9为G

[1908]

[1909] 下文表24提供了在肺组织中发现并且符合表23中提供的一种或多种规则的变体AAV VP1衣壳多肽的581至589区域的序列。因此,本公开提供了由具有SEQ ID NO:2的工程化AAV5 VP1衣壳多肽组成的rAAV,其中Xaa1至Xaa9区域是SEQ ID NO:16118-SEQ ID NO:17117中的任一者,如表24中所公开。本文还涵盖由工程化AAV5 VP2衣壳多肽和工程化AAV5 VP3衣壳多肽组成的rAAV,所述多肽在对应于AAV5 VP1 581至589区域中的氨基酸的AAV5 VP2(氨基酸残基445至453)和AAV5 VP3(氨基酸残基389-397)中的区域处具有下文表中公开的序列。

[1910] 表24

[1911] 驱动肺组织嗜性的AAV5 VP1衣壳多肽中581至589区域的序列

[1912]

SEQ ID NO	581-589 序列
16118	ETDYCYQCL
16119	EHVFGLLRE
16120	EDHISDSQL
16121	ENTEGIGPS
16122	EVTIHFVPQ
16123	EPRMSDMGD
16124	EPFCTQQLM
16125	EHPRQQDSA
16126	EVTRNHHFN
16127	ETRSGIAWQ
16128	EANLSHSWG
16129	EFLGTRLPD
16130	EGETVISQW
16131	EGRLKRHDF
16132	EGTEEGPTH
16133	ELRFVKTQR
16134	ENLARDAHY
16135	ERTIQHGEA
16136	ESTKAKKFP
16137	EVKLFKRNH
16138	EYKVYRHID
16139	EQKIKNGGH
16140	ERTCIATMW
16141	EVQRNIYSP
16142	EFSAMYSWY
16143	EHITEQDLR
16144	EACYAPSWA
16145	EAEYWNCGR
16146	EAGILMPAL
16147	EANSCEGFQ
16148	EANTWLNSA
16149	EARYSCWWG
16150	ECEQPNRWS
16151	ECFGKRFYD
16152	ECLPVLHND
16153	ECMGPGVCG
16154	ECQLCNWWP
16155	EDFWVFHYK
16156	EEEGQTAAG
16157	EEEQWGHMY
16158	EEVPCEKCF
16159	EEVRPKFGV
16160	EFCEMQNLC
16161	EFCHAMVCQ
16162	EFEDYGSWE
16163	EFEWECVQA
16164	EFFIIGAID
16165	EFKMSVFSN
16166	EFKQEFEAF
16167	EFTRKSCKF
16168	EGEVVDCDK
16169	EGGSQTYQG
16170	EGHQCQEGL
16171	EGRIHAQWM
16172	EGSPQESMS
16173	EGTGMQQAF
16174	EHADSWACP
16175	EHEGLNNHL

[1913]

16176	EHFFKREDR	16217	EQMGYVHKG	16258	EVPHFDVSQ	16299	TTMPKGGNL
16177	EHFTKITVE	16218	EQNTANLCD	16259	EVQCCMAAM	16300	VTESWSSWQ
16178	EHFYICVHD	16219	EQQFWNRLS	16260	EVQPVRRFT	16301	WTCLCYKFM
16179	EHNKEVYYG	16220	EQQTGGYGP	16261	EVSFKARCA	16302	YTEQEDNAA
16180	EHRQMHDSY	16221	EQRFWNRLS	16262	EVSMPEKTP	16303	KTYCQDGWT
16181	EHSCTHMGN	16222	EQVLMTYWF	16263	EVSMTGHIA	16304	LTTVYNWPH
16182	EHTRKDVCD	16223	ERAEFEMKE	16264	EVSPMRHIM	16305	ITPPYCDMV
16183	EIHSNIAAG	16224	ERAFSHNQS	16265	EVTNEHCCW	16306	LTMMAMCND
16184	EILQGTVVQ	16225	ERDHIKGCI	16266	EVTQFDNPV	16307	ATQTDPVAR
16185	EISMRPVTQ	16226	ERSDFHWYM	16267	EVTQGECFY	16308	HTLHSHQNT
16186	EIVYVDYPH	16227	ESDWMNVSM	16268	EVVHDPGYA	16309	STERGTPQY
16187	EKCFCEVCG	16228	ESFNYAHFK	16269	EVVKCPNSI	16310	TTGTFPALQ
16188	EKDHFDVMQ	16229	ESHPFTNFS	16270	EVVKIDMIP	16311	ATVMFIEIS
16189	EKLNHEAGY	16230	ESISRDQSW	16271	EVYAPRHMC	16312	FTNQLPEAN
16190	EKMHQLATP	16231	ESKMSHKEA	16272	EWDKMQKVP	16313	CTLRVEYVQ
16191	EKRMMIDTH	16232	ESLHALAHM	16273	EWGRRANKD	16314	VTWMVKTSN
16192	EKWMYDKDL	16233	ESMFWDSDK	16274	EWHAFMSHC	16315	ATQFSGENA
16193	ELDDICGIL	16234	ESMRRLGDT	16275	EWWRYMADY	16316	ATQLYAPIA
16194	ELEHYMSYQ	16235	ESNSNRWAY	16276	EYCWANKNG	16317	FTVYVGARE
16195	ELKWWWRIF	16236	ESRIMAPLW	16277	EYIREWAQD	16318	KTKCCKTHD
16196	ELPTILKHR	16237	ESWNILLTH	16278	EYKQQGVRR	16319	KTQIHMGDE
16197	ELQKWFRKS	16238	ETADRPSLD	16279	EYQGNYMAT	16320	RTEAHLPAP
16198	EMEGYRAIQ	16239	ETAHCPLNA	16280	EMMERQIHE	16321	TTERYEITP
16199	EMFWQVAGA	16240	ETCIANAWL	16281	EFITSGQCS	16322	VTVYVGARE
16200	EMKFFAYAA	16241	ETDDQMMPP	16282	EVVFMSGDI	16323	ITASIKDQW
16201	EMKNLKRTG	16242	ETHTSEANM	16283	EPGHPGYEM	16324	DTAPNESDR
16202	EMKVIGVDF	16243	ETKLLPDLG	16284	EVMGLGCDG	16325	HTRKEGFNS
16203	EMRQQKMSR	16244	ETMAQSCYG	16285	EAQWGMDDT	16326	GTWMEENLH
16204	EMRTVMAGV	16245	ETMGICHNS	16286	ECKTHRDVY	16327	RTDVMMKVM
16205	EMTTSWYMN	16246	ETMHPVVDG	16287	ENIPQSIWT	16328	ATEMFDCDV
16206	EMWSVKFGM	16247	ETMSTVYGQ	16288	ETCWTACDR	16329	ATNVLYLDC
16207	ENAQFLVWA	16248	ETNSCEGFQ	16289	EAYTDMPKT	16330	ATQQVFWNL
16208	ENDSVSAIG	16249	ETRQAHLYH	16290	NTRPMRLHM	16331	ATSHYAKGL
16209	ENGHCVSCT	16250	ETTAQKLKI	16291	NTRPMRPHM	16332	ATSIFCVFR
16210	ENHRITGQS	16251	ETVEYLSYE	16292	QTMQWMCKG	16333	ATTHWAGHH
16211	ENRKWTLAF	16252	EVASKFFEH	16293	QTMWLKEHR	16334	ATTMSAFYG
16212	ENSQTSAFG	16253	EVCRIGWMS	16294	RTEIRYYGT	16335	ATTSYNLTR
16213	EPGELSVHV	16254	EVCTIMEHP	16295	STRPAIQKS	16336	ATVDISNVK
16214	EPLTMDTGD	16255	EVGTSCRFE	16296	STTVWAQCS	16337	CTPPIVMDR
16215	EPRSCNFWC	16256	EVNDQYREL	16297	STVMESAMY	16338	CTSFACKYF
16216	EQASSGERG	16257	EVNMPTGQH	16298	TTGTFPALP	16339	DTAMMHKIA

[1914]

16340	DTGEIMTRS	16381	LTTGIWKTW	16422	RTHLMEVTQ	16463	ITAIMCRSS
16341	DTIPYCRGC	16382	LTWMGLSME	16423	STDLKFCAK	16464	CTESCQTGP
16342	DTKFWMQWN	16383	LTYVQGDWR	16424	STDNKFWLG	16465	MTESVENPG
16343	DTKIQTIHT	16384	MTEQSGRAY	16425	STEAAYENP	16466	STTAYHQDY
16344	DTLLPANWS	16385	MTETTDMAH	16426	STERMFKEA	16467	CTEIRYYGT
16345	DTNKLNDLS	16386	MTFFVNQAH	16427	STFTAGATM	16468	QERGSLWYG
16346	DTRQCLDTQ	16387	MTIGFHVNG	16428	STHGHISHN	16469	WWRAMEKDM
16347	DTTTSTMAL	16388	MTKQNKTHR	16429	STMNFLPCE	16470	LDRNQSYHA
16348	DTVNTCGQM	16389	MTTSSHWAK	16430	STMSCEILR	16471	SLRTPECDC
16349	FTELHCNGN	16390	MTVQWAQYG	16431	STMTCGEFA	16472	CYRCALWTV
16350	GTEYDRLSP	16391	NTAPKMSCT	16432	STSDHHCHS	16473	ANRFYDCSH
16351	GTFMDQRPN	16392	NTDVPQFMP	16433	STSSLDKLI	16474	DRREKQLMK
16352	GTMPMDAGG	16393	NTEPMWSCS	16434	STTNPYQLA	16475	DVRAMHSTC
16353	GTMTTLYAN	16394	NTEVMDTGQ	16435	TTAMYPVHT	16476	FCRMYANAD
16354	GTPHVVSQE	16395	NTLRPQHTF	16436	TTDCMWENH	16477	IGRLNASTP
16355	GTQVNRATE	16396	NTMEDWECY	16437	TTDQKRHGS	16478	IRRSMQQQN
16356	GTSVNMKDQ	16397	NTNFWNWSL	16438	TTEQWDFCR	16479	PSRWYKLGI
16357	GTSYWPQWN	16398	NTRSKDFWN	16439	TTERTRICN	16480	TIRYIGMSG
16358	HTAICQDAS	16399	NTSAHGYNQ	16440	TTMPHHEWQ	16481	ASRSCMTQE
16359	HTAYPYQCP	16400	NTTTFCHDE	16441	TTMQCNSYD	16482	AERQACTMG
16360	HTDCLFTDD	16401	NTVFMHEQT	16442	TTNPVEMFN	16483	VGRERVWDS
16361	HTIPFFEAE	16402	NTVSHFAYD	16443	TTTHMLTRC	16484	ALRPYMDDR
16362	HTKMTHGTS	16403	QTAANFCFD	16444	TTVPFQAWC	16485	ASRYDYNHG
16363	HTMQREYME	16404	QTDAQGIGH	16445	TTYCPPTVE	16486	ASRYDYSHG
16364	HTNCLFTDD	16405	QTELDTGIS	16446	VTDLVRHFA	16487	CERQALKMI
16365	HTTTQVKHP	16406	QTFLAMDGK	16447	VTEHQQRSV	16488	CLRQDRIIF
16366	HTVKIDQTC	16407	QTHDEYECA	16448	VTENLMMYH	16489	DARPSPISQ
16367	ITDQRRRAC	16408	QTIRDMSGF	16449	VTEQDDGQW	16490	DARVSNIQQ
16368	ITELRPLIL	16409	QTLLSNCCA	16450	VTKEMVWQC	16491	DCRCYGNNS
16369	ITNKQQLAR	16410	QTLWVHGTR	16451	VTNQCCKAM	16492	DKRDHHYFA
16370	ITNQFMSKH	16411	QTMPCMYVE	16452	VTTQSCKWM	16493	DSRFYEVQL
16371	ITRVDCSHD	16412	QTNVELHQE	16453	WTDNYFHES	16494	DSRQLHHIG
16372	KTEPSVRIA	16413	QTPQYHVGG	16454	WTLHNQSLK	16495	DSRYTQCYI
16373	KTLMWDCQT	16414	QTQLNCAAC	16455	WTVHNKDEA	16496	DVRMNQDLY
16374	KTTILEDKA	16415	QTSFAETDY	16456	YTKAEDSIC	16497	FGRNYTDHL
16375	KTYICASCF	16416	QTTVFQLEN	16457	YTKQIDNIE	16498	FGRYRLWKG
16376	LTAVSVFRR	16417	QTVQLGTPV	16458	STMMVGRQC	16499	GERWPHNML
16377	LTHDGWFYS	16418	QTWAAHCSR	16459	ITNSERATY	16500	GFRVGSLRH
16378	LTKCNEKGN	16419	QTYHVYTCC	16460	PTDVITTDQ	16501	GIRLKGNIM
16379	LTMLHMHCQ	16420	RTDRNTDFL	16461	DTITFKQQH	16502	GSRPRMKQP
16380	LTSGEGFEV	16421	RTGFEVPTS	16462	MTTEHITWW	16503	HCRLKEDEG

[1915]

16504	HERTAKFRY	16545	TGHQLPYNK	16586	DKMQCSSGG	16627	LICQSDCFP
16505	HRRLFNKMN	16546	LIDQNTKGN	16587	DMAQQQKSH	16628	LMTQEAQRA
16506	IERNSGQGS	16547	CGYQFSETM	16588	DVMQGNRAF	16629	LSVQAQWDV
16507	IGRDIDTSF	16548	SGCQYEPRN	16589	DWAQFCNHT	16630	LWMQHMCLD
16508	INRENRRGG	16549	IQAQMDQQQ	16590	DYAQFSVVP	16631	LYCQSPKDC
16509	ISRMCESAL	16550	CFNQRIQDS	16591	DYAQGAAMP	16632	MAHQSDIMH
16510	KLREVGRPC	16551	MGLQNDPTS	16592	FGMQRRHKE	16633	MAQQYHDEM
16511	KQRDYMRKR	16552	TATQREAFV	16593	FHKQFHSKF	16634	MGHQWSGNK
16512	LMRWPWDDT	16553	IAEQRLVEV	16594	GDEQTHCNG	16635	NFEQYVHNY
16513	LNRILDRMG	16554	QPSQQGRPG	16595	GFTQFYHHE	16636	NFKQPNHWD
16514	MGRTCADWT	16555	AVDQLRMFL	16596	GIHQETYCP	16637	NHQQQNLIG
16515	MIRNETQWA	16556	CDHQRHQQP	16597	GITQMSQHA	16638	QAEQENFDH
16516	MVRFAPAYP	16557	FRKQGHGEK	16598	GMVQDYVWL	16639	QMAQREDDY
16517	NMRYTEMGV	16558	IIGQWGQSD	16599	GSMQVMKKC	16640	QPAQMEDMG
16518	QCRNQDKSQ	16559	IIGQWGRSD	16600	GVCQMSEMR	16641	QSAQQPAGS
16519	QERWQIQTW	16560	NELQQGTSE	16601	GVFQEASHK	16642	QVFQLTMVK
16520	QSRLEHYGL	16561	NVSQWMITY	16602	GVTQGHEIT	16643	QVGQYTLPE
16521	SHREMNRAR	16562	QYVQCAAKD	16603	GWMQWHRQR	16644	RCWQADISG
16522	SPREKNHYW	16563	SLTQYQWRG	16604	GYEQPVKCE	16645	RPAQIDPWY
16523	TAREDCQHA	16564	VDMQKQTES	16605	GYIQCQMHA	16646	SAVQIQVLH
16524	TARLIEKYN	16565	VNVQMPSAP	16606	HASQEYTSL	16647	SLAQDIACF
16525	TVRWQADTM	16566	WAMQWWAKP	16607	HASQLGAMV	16648	SNDQNIHNG
16526	VERHKMKWQ	16567	YESQKGRRA	16608	HDAQFLPDQ	16649	SNEQIPRVP
16527	VSRSLERQA	16568	MGMQQDTQV	16609	HGLQVGQFY	16650	SQSQFCAQV
16528	WCRQTMMDW	16569	AAQQGLNPC	16610	HPEQFVACE	16651	SQTQWDHMD
16529	WFRRHNYFY	16570	ADIQISLCD	16611	HSNQMWYCY	16652	SSEQAGPHM
16530	WNRDHNDWT	16571	AFAQYSRGN	16612	HWAQLKVTY	16653	SVQQIQKQP
16531	DVRSHDEMT	16572	AFNQQEMWA	16613	IHKQFEMES	16654	TFHQCERSD
16532	TERRMFRTA	16573	ANHQVTSHP	16614	IHQQTHYQF	16655	THQQVALGN
16533	QVRFEQNPA	16574	ANMQPNAIM	16615	IMMQRALMG	16656	THTQYWKYA
16534	MDRMHKSGI	16575	ASGQFPKEG	16616	INQQYNSCW	16657	TISQMVYSK
16535	KNRDFEGFD	16576	ASLQATCNH	16617	ISTQMENTG	16658	TMGQICVHA
16536	NQDQARTEH	16577	AYAQKYRTA	16618	IVEQKPTVQ	16659	TQCQWNVAK
16537	QAFQFECHD	16578	CFSQTDCMS	16619	KDAQQWNCA	16660	VEAQFMDNA
16538	QATQAGANQ	16579	CIWQHNITQ	16620	KGNQQITSL	16661	VLSQYTQHA
16539	RMAQFFRAH	16580	CSQQVNCLS	16621	KMIQSTHDF	16662	VVSQTYQNR
16540	RNLQGGNHG	16581	CWMQSQENW	16622	KPNQSNRQE	16663	VWQQYILVI
16541	SCLQYKATE	16582	DCAQITYGG	16623	KSAQVYWNP	16664	WYIQVDRKG
16542	SIEQKLTFT	16583	DCAQLRQDG	16624	LAVQSIHSH	16665	YIKQIDNIE
16543	VGWQYCEEG	16584	DFVQSDTEF	16625	LHCQIWFQD	16666	YQTQYSFWH
16544	TRKQGEGPY	16585	DGLQIQIQE	16626	LIAQQMHIN	16667	LVGQCTSHW

[1916]

16668	GIKQTLVRQ	16709	DHCILPQTC	16750	THDSLIAHW	16791	NSYFINELY
16669	GCQQIQLQT	16710	DMFMLDEAP	16751	TIGNLTKPI	16792	TFAAVNCCQ
16670	KYMQENVLS	16711	DYFPLVVAP	16752	TLMGLWNQG	16793	TMKTFNYVF
16671	NCHQLGDNM	16712	GANFLVSQT	16753	TNDPLMRSD	16794	YCQLGNHNA
16672	MMVQWGAMA	16713	GKNPLMPSP	16754	TQCGLKMYP	16795	YHENWNMPM
16673	FSEQNVSAA	16714	GLDKLMMLA	16755	TRMLLEWTY	16796	GWMRHNRIQ
16674	TSGQVDHRT	16715	GVMMLGTAN	16756	TSEPLEGID	16797	QSGYMNKLW
16675	YYEQMDGEF	16716	HIDHLNAAT	16757	VDSWLSMRD	16798	VSMSSNYDK
16676	QPVQAPHCA	16717	HPLPLSVMS	16758	VIHYLTHEC	16799	KCAMFNAGE
16677	SIAQDQVWS	16718	HSMLLQEYM	16759	YCIHLKDDI	16800	AEYMTNASM
16678	PLKALIWEY	16719	INTALFVNA	16760	YCIPLQPMN	16801	AGYASNWAD
16679	QMEALMDCE	16720	KEYDLSLTM	16761	YSVPLHHCF	16802	AISPTNAER
16680	QSKTLCYQC	16721	KNFELDQFE	16762	YVGNLCRWK	16803	AKGTYNQGN
16681	QVSTLENVQ	16722	KPYSLAMTI	16763	YVVKLEDYK	16804	ASETNNGGY
16682	RQAKLATLH	16723	LGTPLEAGL	16764	YYADLNAHD	16805	AYQMFNNVQ
16683	SKESLPQWQ	16724	LIDRLPACY	16765	RNMALHHPT	16806	DAEHKNAIR
16684	TGVSLVNYS	16725	LRDCLHVLG	16766	DGCSLYAFN	16807	DAHYTNECR
16685	MLHNLWMQC	16726	LVDELNQIK	16767	ACDRLMFQG	16808	DCQPVNRIL
16686	AAQTLECHK	16727	MFDDLMVGR	16768	SLVDLYKTA	16809	DINFDNHAG
16687	YWFVLNPCV	16728	MINNLDECC	16769	FCNKLTSVR	16810	DMNSRNVQD
16688	CHVNLKAWW	16729	MVDGLTKEY	16770	IDAFLANRG	16811	DMVPCNLMH
16689	DFGMLEWQW	16730	NCQSLATNV	16771	NMMTLWSHL	16812	DWAANNWWP
16690	KYEILNCNH	16731	NEVHLEHNY	16772	QIVSKNSRQ	16813	FEAPINLSF
16691	MQYTLPTWP	16732	NLEMLAMFQ	16773	RGVHSNCAE	16814	FVHVINRCD
16692	TMANLCWRT	16733	NVMPLIQYR	16774	SAFLPNTVC	16815	GAMPGNGCT
16693	TNIPLGTNM	16734	PITILCCSQ	16775	SEYDNNPCQ	16816	GHKVCNCWE
16694	TWMPLAPCH	16735	PRQLLHFSQ	16776	SFTIYNHKW	16817	GVWPDNCPQ
16695	TWMPLAPCY	16736	QASLLDHYP	16777	SSMTRNNCK	16818	GYCGCNHMG
16696	HQNLLICNE	16737	QGQPLMRYS	16778	THDYFNGDY	16819	HHVWPNRQV
16697	TGKWLVYQA	16738	QQNRLNMNW	16779	TWANGNGGP	16820	HIEMRNAHM
16698	IPADLDVDH	16739	QQSPLLWER	16780	VNHHENWKK	16821	HMYVTNHDG
16699	FFLKLNDVP	16740	QRHDLEVDH	16781	IEWPTNSDH	16822	ICSYYNQTQ
16700	AHYFLLEYP	16741	QVVDLVDVY	16782	NAYVDNLRY	16823	IFYATNHSW
16701	AIIPLPMPT	16742	RGEILQLQQ	16783	QQCMGNDMT	16824	IKHDFNSAF
16702	AIVNLHTQQ	16743	SAQCLEMWP	16784	AFLHENMWG	16825	INAPPNLRT
16703	ANLPLQQQD	16744	SGLILQCNP	16785	ANGTYNLQE	16826	KGDEQNCYC
16704	AVEFLAHRY	16745	SINYLGEGH	16786	ANMFRNGMQ	16827	KIDVENCCC
16705	AVYPLFAPD	16746	SSKPLTNEC	16787	DSYFINELY	16828	KPGTVNPAA
16706	CHNLLHCVY	16747	SVAFLNQEW	16788	FFEMVNTEQ	16829	KQEPCNSGH
16707	CNSTLQYGH	16748	SWNRLNWCL	16789	HSWINNQEH	16830	LCAGTNLTP
16708	DFCMLQMGM	16749	TCESLYSRG	16790	KSVTQNRTR	16831	LEADNNDAD

[1917]

16832	LKIPFNMTC	16873	FREDMNATM	16914	LQEPRMRYE	16955	ILEAIVNFW
16833	LMAYANTDD	16874	SLWVGYRTD	16915	MCQSYERGA	16956	KISPWDGFY
16834	LRTPYNFWW	16875	TPKNDQRAE	16916	MFNRIPRFS	16957	AALENGWFN
16835	LVKVINERQ	16876	LNARTMRSA	16917	MICRMDRMK	16958	ACMVNPGFW
16836	MCTPRNEQE	16877	MEDFSDRWP	16918	MQKAHDRSA	16959	ACSTVMDFN
16837	MKGFDNYSP	16878	RMVIMERDR	16919	MSGFHIRCA	16960	ASIWRKLFD
16838	MLSYHNESW	16879	DGQLVERYD	16920	MSSFHIRCA	16961	AVHAEDIFM
16839	MQNEKNDNP	16880	APKNDQRAE	16921	NLIYQPRKS	16962	DASVPGVFE
16840	MRKDYNDTW	16881	TYYWDCRHV	16922	QATFTTRLV	16963	DCNWHWPFT
16841	MYQDANQIT	16882	LMEAASRSY	16923	QGDHFDRNS	16964	DKHFYQKFH
16842	NVHTFNYAF	16883	TQQTADRWA	16924	QGGMMKRGQ	16965	DSSLHGAFM
16843	QFSGYNYHY	16884	IVTCNGRWD	16925	QIDEVTRNL	16966	DSVNTSAFH
16844	QLHWPNNFG	16885	YDYYKMREG	16926	QVEPVEREY	16967	DVIRDDSFQ
16845	QMHVANINL	16886	ASKAKQRVA	16927	RCDGVMRSQ	16968	FGMVHHGFF
16846	QSNEFNPVA	16887	FQATTKRLP	16928	SAELVQREA	16969	FKYYIPEFC
16847	QVAPKNCWQ	16888	NAQLGKRGH	16929	SGSHCGRMA	16970	FPGWCASFP
16848	QYICFNCYH	16889	YLWMIARPD	16930	SLKMRPRQI	16971	HICDMDMFL
16849	SACEPNVSP	16890	YLWMITRPD	16931	SRGDYLRVG	16972	HYFYTQQFG
16850	SANKMNAER	16891	KQNPSMRQD	16932	SWEWQMRYP	16973	IAIPCASFY
16851	SFGTCNTCG	16892	SATARFRDN	16933	TAIHEHRGL	16974	KNVMHSEFR
16852	SIAMYNNMH	16893	AAEMHQRGE	16934	TIEPFWRSW	16975	KYDREWPFL
16853	SREAQNRET	16894	AAEMHQRGV	16935	TKYVDDRPG	16976	LCCDVCMFL
16854	SVKDNNRRL	16895	AFKATRRGV	16936	TSIPITRQE	16977	LGDTGLMFG
16855	SWTASNWRD	16896	ASGSHQREN	16937	VINMNRRWG	16978	LMSICAAFP
16856	SWTSRNAHH	16897	ASGSHQRKN	16938	VSINMWRAQ	16979	LQGWRWWFA
16857	TAQYVNREI	16898	ASHIYCRHE	16939	WDNLMQRTC	16980	LVSDHIEFG
16858	TQDVQNMYV	16899	DIKMQKRGV	16940	WIQPISRTG	16981	MASSEHCFF
16859	TSMLHNEEN	16900	DVGLALRNH	16941	YGSYYMRGW	16982	MHIEHVMFQ
16860	TVKEENAAC	16901	FQADYKRHE	16942	YMHEWMRLT	16983	MIWYNRKFM
16861	VFIAQNEGN	16902	FVKMKQRYG	16943	YSAPRIRDA	16984	MNEIFVEFP
16862	VGVGVNRFD	16903	GAGYRMRVN	16944	SMNLWCRGN	16985	MWTAQQTFM
16863	VHANANMSH	16904	GVSVEGRDE	16945	CDLFPIRWP	16986	MYIPVYTFN
16864	VPELGNGAY	16905	GYKNCHRHG	16946	TEGPMHRAW	16987	NVQPESTFF
16865	VQLPTNKQW	16906	HIDLQTRKE	16947	QGTGHSLFL	16988	PAMPFYWFE
16866	VSINTNEPG	16907	HVEWTMREH	16948	RHWRHTMFL	16989	QHQPFSMFS
16867	VVQYKNHAT	16908	IGGMFSRHA	16949	RMETWCGFS	16990	QNSHGPVFG
16868	YGTEVNSMI	16909	IQKLMRRMH	16950	PMWTQMPFP	16991	QNVMHSEFR
16869	YQVANNQQG	16910	KAISHERGQ	16951	DINASEWFC	16992	QQAPATIFP
16870	YYADMNAHD	16911	LAYTCHRGD	16952	MIVMGPHFT	16993	QQVEWSVFQ
16871	HMWMYNKCT	16912	LHWYWCRNW	16953	SLYNCMFFL	16994	QSISDMPFP
16872	ARTIANDTQ	16913	LKKKRGRAC	16954	WMKVRQGFG	16995	QVTPSMIFK

[1918]

16996	RDSCYINFP	17027	WDNMDQTQG	17058	DIMLRTCCG	17089	NSVSCGGQG
16997	SATCWLWFG	17028	AMLRQEKTG	17059	DRCMQTKHG	17090	PGDFTSGYG
16998	SCMGFTSFP	17029	DSQNNWPTG	17060	DSATQHMEG	17091	QCDEQQEAG
16999	SDEWEDVFG	17030	KCGWHGEHG	17061	DWTTISHEG	17092	QFGSFWPTG
17000	SGCHWGIFH	17031	VQQNNLKAG	17062	FEGSRQLWG	17093	QNARSSMVG
17001	SMYSTHCFC	17032	VAANTYCYG	17063	FREWRAMDG	17094	QYDTIQLNG
17002	SYPGHHWFY	17033	GEPWFQTHG	17064	GCTPSVCSG	17095	SCHHVGMCG
17003	TASGNTYFV	17034	GEPWFRTHG	17065	GVMLKMAGG	17096	SFKWTIVNG
17004	TILEGDNFA	17035	GMASMEHCG	17066	GYVPWLMEG	17097	SIMKYDQMG
17005	TITFIDLFN	17036	GQGTSRLMG	17067	HCHMIFKDG	17098	SMHHTQFTG
17006	TKNAMHNFF	17037	ICHLRHIAG	17068	HKDDYVQTG	17099	SVGFPYESG
17007	TVSRASYFD	17038	MMLGYMEQG	17069	ISEDIKENG	17100	TAQYDVWHG
17008	VAVTVTDFD	17039	NAKPKQVQG	17070	ISWLCEMNG	17101	TGYPKLKDG
17009	VMEKYAHFP	17040	PCYNAMNRG	17071	KNEWTYHQG	17102	TQEPWGCAG
17010	WGDRYGLFV	17041	QIHPYLEAG	17072	LEDMMSVEG	17103	TQGLPRGRG
17011	YSQWFTFFE	17042	SRDCVVHDG	17073	LIMEQPLAG	17104	TRLYNWPMG
17012	QVHAKLGFM	17043	VCPSCRDYG	17074	LNAECTSTG	17105	TSHHWAAQG
17013	SATIRFYFN	17044	WIKWKCSEG	17075	LSTCWAIAG	17106	TWWKDPQEG
17014	QHQAFFVFT	17045	CLNNSTCRG	17076	LVVNQYTLG	17107	VAHYYCWEG
17015	QMKGATEFC	17046	QHSKCHMHG	17077	MAAPTALNG	17108	VLMNWQPRG
17016	LVNMFQCFW	17047	AADFSGYYG	17078	MIGTTQSDG	17109	VLNHWEPIG
17017	ASTERAKFG	17048	AAGATMLRG	17079	MKASTMELG	17110	VVMHPYAAG
17018	NVKWHCDGG	17049	AEIPIMFMG	17080	MKVYHEMVG	17111	VVSHYDMIG
17019	PYFRFFPHG	17050	AFVNWPIHG	17081	MQVSADHHG	17112	WHTVHYMGG
17020	QCGAIYQTG	17051	ANQPKHGAG	17082	NAAWQSITG	17113	WLTFDRFEG
17021	QQNLADQAG	17052	APLIKASWG	17083	NCSHYDSRG	17114	AKCPWHHTG
17022	RQGKAMSMG	17053	AWDWCDMRG	17084	NGMVSGQYG	17115	VNPWYVKHG
17023	SCTHIKVPG	17054	AYNSRANGG	17085	NHPYQHIEG	17116	DVASRCWVG
17024	SCTMKDWWG	17055	CHGHFQHIG	17086	NKMHCDTMG	17117	SGTPHQQSG
17025	SCTPIKVPG	17056	DAKTCHSRG	17087	NKVMCLWQG		
17026	FLNNEELKG	17057	DCKLMRHQG	17088	NNAPVWLHG		

[1919] 实施例13

[1920] 在心脏中具有组织嗜性的AAV5变体

[1921] 此实施例描述了使用实施例1-实施例2中描述的方法和系统发现的在心脏中具有组织嗜性的工程化AAV5变体。

[1922] 分析CNS、肝脏、骨骼肌、心脏、肺、脾脏、淋巴结、骨髓、乳腺、皮肤、肾上腺、甲状腺、结肠、坐骨神经和脊髓组织以鉴定AAV5VP1 581-589区域中驱动心脏嗜性的氨基酸残基。结果显示于图18D中，该图显示了相较于所有其他分析组织在心脏组织中AAV5 VP1的581至589区域中每个氨基酸的位置频率，其中衣壳多肽具有SEQ ID NO:2的序列。位置(“Xaan”，n＝1-9)在x轴上表示并且氨基酸残基在y轴上表示。特定位置的特定氨基酸残基的频率在热图中作为强度表示。较深的方块表示特定氨基酸在特定位置的频率较高，较浅的方块表示特定氨基酸在特定位置的频率较低。

[1923] 分析相对于给定氨基酸残基出现在所有其他所收获组织中鉴定的变体中的指定位置的频率而言该给定氨基酸残基出现在心脏中鉴定的变体中对应于SEQ ID NO:1(推及

于SEQ ID NO:2)的位置581至位置589的所述指定位置的频率,以确定优先靶向心脏组织的衣壳的一组序列规则。参考下文表25,和SEQ ID NO:2(AAV5 VP1),以下氨基酸由此可以在Xaa1-Xaa9任何一个或多个位置处以任何组合独立地突变,以提供与野生型AAV5 VP1衣壳(SEQ ID NO:1)相比具有增强的心脏嗜性的VP1衣壳,其中心脏嗜性在此是指优选心脏转导的特性,而不是优选转导所有其他所收获组织的特性。

[1924]

表25 心脏嗜性规则
Xaa1选自I、K、L、M、T或V
Xaa1选自K或L
Xaa1为K
Xaa2选自A、C、G、I、K或S
Xaa2选自A、C或S
Xaa2为A
Xaa3选自A、D、E、G、K、M或V
Xaa3选自E或V
Xaa3为E
Xaa4选自F、H、R、T、W或Y
Xaa4选自F、R或T
Xaa4为R
Xaa5选自F、L、M或R
Xaa5为L
Xaa6选自A、H、N、W或Y

[1925]

Xaa6选自H、N或Y
Xaa6为H
Xaa7选自A、C、E、F、K或T
Xaa7选自C、F或T
Xaa7为F
Xaa8选自A、C、M、S或T
Xaa8选自C、M或S
Xaa8为C
Xaa9选自A、D、G或P
Xaa9选自A或G
Xaa9为A

[1926] 下文表26提供了在心脏组织中发现并且符合表25中提供的一种或多种规则的变体AAV VP1衣壳多肽的581至589区域的序列。因此,本公开提供了由具有SEQ ID NO:2的工程化AAV5 VP1衣壳多肽组成的rAAV,其中Xaa1至Xaa9区域是SEQ ID NO:13118-SEQ ID NO:

14117中的任一者，如表26中所公开。本文还涵盖由工程化AAV5 VP2衣壳多肽和工程化AAV5 VP3衣壳多肽组成的rAAV，所述多肽在对应于AAV5 VP1 581至589区域中的氨基酸的AAV5 VP2(氨基酸残基445至453)和AAV5 VP3(氨基酸残基389-397)中的区域处具有下文表中公开的序列。

[1927] 表26

[1928] 驱动心脏组织嗜性的AAV5 VP1衣壳多肽中581至589区域的序列

[1929]

SEQ ID NO	581-589 序列
13118	KAAPMDISI
13119	KADCPAYKA
13120	KADVWMAIA
13121	KAELSDHVL
13122	KATDALPSC
13123	KCSKQASYP
13124	KDGYALLSC
13125	KDNEAQYQP
13126	KDYMFNLPF
13127	KEAMMYNAN
13128	KEGLIFSTE
13129	KEHTNICRH
13130	KEIVEVDGN
13131	KEMAYQHMW
13132	KEMLAMWLQ
13133	KEWFVTWKA
13134	KFLPCITSY
13135	KFPYLPHRN
13136	KGCMSWEEG
13137	KGDRFCGEV
13138	KGEPGYMYY
13139	KGMFVANSH
13140	KGTKGPQMQ
13141	KHAGTRYKG
13142	KHFTSCDRT
13143	KHHDAAREM
13144	KHIFNECDC
13145	KHMQKYVHY
13146	KHPRFTGWH
13147	KIESDWHAQ
13148	KKHQFCEGT
13149	KLELHGCQN
13150	KLMSFRELE
13151	KLQLEKPLN
13152	KMDCHKYGV
13153	KMNETTNPR
13154	KMVDQSLAT
13155	KMVMIDTDV
13156	KNFMNHEPQ
13157	KNNRTIRWG
13158	KNNTWKWWY
13159	KNQYFITSQ
13160	KNWEKQNEV
13161	KNYNPVNYE
13162	KPCKACLNS
13163	KPCWDQAMQ
13164	KPDQKLWWD
13165	KPDWWIKVN
13166	KPSNSKQTV
13167	KPTQHGVGP
13168	KQAIGDDVG
13169	KQHIFQWHH
13170	KQIHEWEGQ
13171	KQLPEQQPL
13172	KQMLEFAPY
13173	KQNEINHSC
13174	KRIDHYGWG
13175	KRSFEWLER
13176	KSDYNSYDA
13177	KSLSSEMGV
13178	KSMDKDLTS
13179	KSSCVDDSS
13180	KSVAEGAND
13181	KTAQLSVME
13182	KTAYELVDG
13183	KTCDSQTMT
13184	KTEMYSDQC
13185	KTFEWLCPY
13186	KTMPVGAWD
13187	KTTWVDDFD
13188	KTYYVMAIV
13189	KVLDTLTEH
13190	KVMHACNAN
13191	KWDETMFMS
13192	KWEDFGWWK
13193	KWEQPNPQA
13194	KWEYVRCAS
13195	KWLIARMYL
13196	KWLKTLMCL
13197	KWQDAHYIL
13198	KWTLDAWLT
13199	KWWVDRLCQ
13200	KYAVSPCDP
13201	KYDWLDAGG
13202	KTDPIAEYC
13203	KGNPLDGDT
13204	VAHICDPET
13205	VASEHSCLP
13206	VAVEDRPAL
13207	WAREELGFA
13208	YAMCHRWMR
13209	YAPNNSHWP
13210	AACPHQFKD
13211	AADLGQSMC
13212	AAESAKFHE
13213	AAGIAVACT
13214	AAIRCMQTS
13215	AAKLYMDLS
13216	AAMKGEVIG
13217	AAMLRLLGK
13218	AANMLYCST
13219	AANWTHCSL
13220	AAPMHLHEI
13221	AAQFMEKNQ
13222	AASGLEVVA
13223	AATNWGVET
13224	AAVKWDQPL
13225	AAVPELHHS
13226	AAVQQPEHT
13227	AAYSQSAGS
13228	CADSCYWYP
13229	CAEAWFPQR
13230	CAFLDWDWC
13231	CAFRWVEWD
13232	CAHHVMWYP
13233	CAHQPDGYA
13234	CANAENATC
13235	CANCVAISP
13236	CAQPNTHNS
13237	CAVAWKGDY
13238	DAAQCTWAR
13239	DAHPEGPQR
13240	DALIYHSPF
13241	DAMRTHENT
13242	DAQKQQVMH
13243	DAQWIGCLE
13244	DARWRSQCP
13245	DARYSWPGA
13246	DATFNTSSL
13247	DAVLFTAMQ
13248	EAFGNGIDP
13249	EAGIHQKVI
13250	EAGRNYMDL
13251	EANQCHVHI
13252	EATWTAQGA
13253	EAYLWENQL
13254	FAEWKTEQQ

[1930]

13255	FAHCAMNES	13296	MAAVMIDQA	13337	RADDYQPVS	13378	YAANKDCFL
13256	FAHWRDCWY	13297	MACRAWIFV	13338	RAELFNNGK	13379	YAKYGQRWM
13257	FAKIRERHV	13298	MADKCANNA	13339	RAEQADWHE	13380	YATFEVWFR
13258	FASCMRGRA	13299	MAEQAVNSK	13340	RAILNEKRD	13381	YATMEYNNN
13259	FASPAGPWS	13300	MAGNVCLCV	13341	RAINFDKAS	13382	EAFQKCNWQ
13260	FAVNEHDVF	13301	MAHMNTGVP	13342	RAPELVMNE	13383	HATGLVNFM
13261	FAYATIRRD	13302	MAKWDDSFS	13343	RAPYEADSN	13384	AADIFAIND
13262	GAEIMVNGW	13303	MALMVEECD	13344	SAAYIVKGP	13385	EAELENVNM
13263	GAEPYVKGG	13304	MAMLCSICS	13345	SADGQQAAT	13386	IAYHYDQQN
13264	GAERQWFTP	13305	MANPEMYMQ	13346	SADLSLTYD	13387	TSEPLLAML
13265	GAHVNSMNA	13306	MAQFVWCNG	13347	SADQVPECP	13388	TTEDCPNRV
13266	GANSSELQQ	13307	MASMMNACN	13348	SAIYRNWFE	13389	TTEYTCEDQ
13267	GAVMWERDG	13308	MASMSKDDS	13349	SAKLSEQNV	13390	VFENCSRTF
13268	HADCYVADF	13309	MASQEHQES	13350	SAYALQSPR	13391	VGEWEDEHD
13269	HADRMEGPW	13310	MAVHRPTFC	13351	TAAIVGKES	13392	VIEHAHHVN
13270	HADRPMSAC	13311	MAWDKWFWP	13352	TACMDHHIR	13393	VSEGQVNMD
13271	HAEWTEPVD	13312	MAYQQVIAG	13353	TACNMMCAY	13394	VSEHDDCGL
13272	HAIPACKRD	13313	MAYQRHAAF	13354	TADWYNGGP	13395	VSEVYLAER
13273	HAKLETMWG	13314	NACIRRTTW	13355	TAEVWWVFH	13396	WYEKLMDAF
13274	HALAWCTKH	13315	NACIRWAFD	13356	TAGFAKHFP	13397	YCEVPQNPV
13275	HATYCQMAM	13316	NADALGALQ	13357	TANMCDEGI	13398	YDEYCESKS
13276	HAVSQERAG	13317	NADREMFEH	13358	TANSEAHYT	13399	YIEIVRGPM
13277	IACVHMKQW	13318	NAEHRGMTN	13359	TARWLDEHL	13400	YSEDATPKG
13278	IADERWCMA	13319	NAGAQWLGI	13360	TARYITSRA	13401	AFEFTDGCC
13279	IADYEQCSS	13320	NAGLEWKLQ	13361	TASEFPCGV	13402	AFESAADHW
13280	IAFETYWTN	13321	NAHCGAACC	13362	TAWKADSGM	13403	AIEVEGMWD
13281	IAGQWAAFG	13322	NAMAYENDY	13363	TAWWNDCWF	13404	ALELEAHRP
13282	IAKCIDVMD	13323	NAMLTLGAG	13364	VAAGPLFYG	13405	AMERIKMRN
13283	IAKWIEKIG	13324	NAMQRANIR	13365	VACAPEAMG	13406	ATEAYSKLQ
13284	IAMLSQHTV	13325	NAPCFRYYN	13366	VADGRQPME	13407	ATEKDHRVC
13285	IAQPYDNWG	13326	NARDDSIFN	13367	VAFYNKMRG	13408	AVEFQHGHC
13286	IATNHSAGP	13327	NASHYEKDG	13368	VAIQSVGYH	13409	AVEIHKCID
13287	IAVTGFAMS	13328	NATSYHKFM	13369	VAIVWMCSE	13410	AVEPGISPF
13288	LADGMGKVP	13329	PAEIKMPSQ	13370	VALPRCQLV	13411	AYEHDFKCS
13289	LAFSDITWS	13330	PARLMMWEF	13371	VAMYTRCHT	13412	CCEIKRHYY
13290	LAHSEDRTS	13331	QADQYYSLE	13372	VAPKDNQQM	13413	CIEKSISST
13291	LALQLMGAH	13332	QAMLKQGPR	13373	VAPQCTNEH	13414	CLEQMCRTN
13292	LAMWFNESC	13333	QASCISWSQ	13374	VAQPVQDMG	13415	CPEPFQKVD
13293	LAMWKMRQH	13334	QASQENRLT	13375	VASKTQFAM	13416	CREQNNCDR
13294	LATMEDQSN	13335	QAVECTQAP	13376	VAYYFDQAG	13417	CSEHGCDIP
13295	MAAPYQPYN	13336	QAYTGDKDI	13377	WAPPCYYDV	13418	CSEWIRAVR

[1931]

13419	CTEQYLVDK	13460	IRECGQHGS	13501	RVEVWYGEL	13542	VPCRPIYQH
13420	CTEVRNWYD	13461	ITETWDIFF	13502	SEESVVTHF	13543	VRWRGDVAN
13421	DCEFPMGAN	13462	LDEFGFWTH	13503	SHEQLAGCP	13544	WVLRCTYEY
13422	DDEGYIFNQ	13463	LGESEHCMA	13504	SMENRWEYL	13545	ADARVMTMM
13423	DGEFRYVVS	13464	LKEAWICKG	13505	SMEYYYGME	13546	ADLRPSVDY
13424	DGEQESGQT	13465	LKEGINHIC	13506	SNEQMMTMQ	13547	AGDRSQILG
13425	DHEVSVAGS	13466	LMEGSAKAM	13507	STECSGLSG	13548	AMMRAWSTF
13426	DIEMCPCYN	13467	LYEGRPWYN	13508	SYEVTYHMD	13549	AQDRFMLGC
13427	DKEIVNMVW	13468	MDEQMGKMS	13509	TCEKLISVC	13550	ATNRIMAAW
13428	DLEYVEFDW	13469	MIEKQYPTM	13510	TKELTRSEA	13551	AVVREHVES
13429	DQEVFQTHF	13470	MKEESKSWP	13511	TNEAIMQVC	13552	CIWRKYPEF
13430	EEEFWMHNA	13471	MLENYEHMH	13512	TNEIGEWWF	13553	CQCRFCHPL
13431	EIEPTIGPE	13472	MNEPQKMRQ	13513	TSELLVNGD	13554	CTIRVQNGE
13432	EKETMWAEP	13473	MPEYMKTMQ	13514	TSENREYMQ	13555	DEFRFLPAD
13433	ELEKYKLTN	13474	MPEYQARLM	13515	TWEKQSRNG	13556	DGGRFDDYY
13434	ELEVRHSND	13475	MREQMPVQY	13516	TWELIQTKF	13557	DHWRKLHKA
13435	EVEEETDEK	13476	MRETVFHEP	13517	VCEEAGELG	13558	DIARLQVLG
13436	EWEYERGAE	13477	MTESGMKDK	13518	VHEYYRSCA	13559	DKYRQQENS
13437	FCETEPSGA	13478	MVEVSRGEA	13519	VLEFCKDEE	13560	DQIRTLQMT
13438	FKEHVQNHM	13479	MWECGYGGN	13520	VMEKQGNGG	13561	DRLRDDYPC
13439	FQEACMACH	13480	MWERIGMGG	13521	VMEYMLIAW	13562	DRPRPMWGH
13440	FSEPSNNFC	13481	MYEDTHCVC	13522	VNENNELSA	13563	DVIRIVHDY
13441	FWESTSPFT	13482	MYEHINSQR	13523	VNEQWQIMC	13564	DYARNHMMW
13442	GEEWVEPTA	13483	MYEKQDGVL	13524	VNEWSADCS	13565	DYTRYHSYF
13443	GHEHGWQLS	13484	NGEGSQTCE	13525	VSEFHDLHL	13566	EGNRPEGLA
13444	GNERLSVCI	13485	NHEQKSYIF	13526	VSELCTNMA	13567	ERVRPLSTD
13445	GNETLEMHG	13486	NLEIYLQMN	13527	VTERWIGPM	13568	ESDREEQSR
13446	GRECKQGGH	13487	NVEFMRCNS	13528	VYEKFSYIT	13569	ESWRECMCR
13447	GTEHWLSAC	13488	NVEWWAPHV	13529	VYEREQMDV	13570	ETMRCQTMA
13448	GYENYNKRD	13489	NWEQEFVMI	13530	YHECKEGHI	13571	EVARGQDWA
13449	HCEYCQMMN	13490	NWEYMRRKC	13531	YSENRIFFR	13572	EVNREETIQ
13450	HFEAGYACS	13491	QDERMDLVQ	13532	YYECLGQDL	13573	EYLRESMMP
13451	HIEGYFRWP	13492	QIESGCPHC	13533	WIEGTLIGA	13574	FDDREPFNG
13452	HIEYTNKCL	13493	QRELITMGC	13534	NWEFNARHD	13575	FDGRPRMFE
13453	HPEKARPLA	13494	QSEVTHLKW	13535	YWEKVPCMK	13576	FGPRKNRHE
13454	HPEMKHQNH	13495	QTEQWMCSH	13536	GNEDCNLAC	13577	FPRRLDHHQ
13455	HVENALVFT	13496	QWEHIWCWH	13537	MNECFSGAQ	13578	FPYRIEGED
13456	ICEEQFLMG	13497	QYENINKVD	13538	TYESLQNSW	13579	FPYRKVECG
13457	IHEEQECGS	13498	RDEVVFAHC	13539	YCECGHWPN	13580	FTARSYQDA
13458	IHEGYKTAP	13499	RNEPNVWVN	13540	CWECCTDSG	13581	FVDRVDKHH
13459	IMEQSAQHT	13500	RTEQSRVWE	13541	VCPRDMTWE	13582	GNDRAEPDR

[1932]

13583	GTCREYGRN	13624	WNLRKANAP	13665	HSACLMTMS	13706	TERLLQQFY
13584	HGARCCWGV	13625	YHDRFERAH	13666	IDTALNVHC	13707	TFATLCDQQ
13585	HLSRNEYKS	13626	YHYREPSAD	13667	IENWLFVWL	13708	TIMGLLGNV
13586	HSLRDWFKH	13627	YPGRQHVAN	13668	IMTMLTSET	13709	TMYPLMHYS
13587	HSPRRATMV	13628	YRIRKTCAF	13669	INAVLWWRL	13710	TPAQLNGYG
13588	IHGRMINSI	13629	YVNRYGAGW	13670	ITHGLQTWD	13711	TQCMLHGPY
13589	IPPRYRHMA	13630	NNVRKVGNT	13671	ITTDLQMDH	13712	TQGWLTLCY
13590	IRLRDPELH	13631	TSDWLKHFC	13672	IWNELILGC	13713	TSDWLQNYI
13591	IWIRWPSMG	13632	TWSLLQQGD	13673	LDCELYQTR	13714	TTQCLDRWQ
13592	LFDRVGIEA	13633	VRAYLMGPN	13674	LGMYLQLGQ	13715	TTWILPHHD
13593	LGCRETRHD	13634	VSHALIDWH	13675	LGVQLSSAP	13716	TYVPLWAYW
13594	LIMRNASWC	13635	YPKQLHVWT	13676	LHLKLYGQL	13717	VEVQLKALQ
13595	LWIRWGSCT	13636	ACHTLDQQD	13677	LLTLLMYWV	13718	VMHCLRQFM
13596	MHTRIRWLV	13637	AHQVLHVRS	13678	LMYMLPPRH	13719	VRPWLVMDN
13597	MKCRIEKVN	13638	AKRELLDHD	13679	LRPWLPDST	13720	VTDGLGFRQ
13598	MVYRIKMTY	13639	ANCYLNHTT	13680	LTHALGTWE	13721	WVRDLKCPV
13599	NFHRWDMNT	13640	APKKLEHVD	13681	LWTHLMANH	13722	YRIFLMHLC
13600	NRDRSWCQH	13641	AQNALWERI	13682	MDQKLLTWQ	13723	CIDVLFKHL
13601	NSDRMHCME	13642	CFAALDYAS	13683	MEDMLMEGH	13724	VFKPLDKRY
13602	NYPRDHRQH	13643	CGNALPLIH	13684	MHMLLHDVK	13725	TQGSVHVMN
13603	NYYRALKWD	13644	CNDCLMFCD	13685	MITWLTDQV	13726	VEPTWHCDW
13604	PVTRWEAGS	13645	CNTMLVSAG	13686	MNVGLWALW	13727	VISNRHHLP
13605	QIDRAEKQE	13646	CSMYLMEFG	13687	MYACLAQPP	13728	WRVSAHYYI
13606	QNVRSNAYT	13647	CWYMLKAYT	13688	NHCMLHHMG	13729	YDHTIHRNP
13607	SDLRNLCAR	13648	DFVQLINRA	13689	NHDHLFERY	13730	YLLYPHSSG
13608	SFIRVLRSN	13649	DGYHLEGQC	13690	NQPHLWSCG	13731	ACDLQHNIA
13609	SIYRMVMQC	13650	DIKALMDGW	13691	NTRDLHTKR	13732	ADQQQHKVI
13610	SNDRFPIMC	13651	DKSKLDVSY	13692	PNGLLQMVI	13733	ADWHCHRND
13611	SQGRMASGA	13652	DTCKLGGWF	13693	QIAALHQNH	13734	AEAQFHTGE
13612	SQTRDTLNF	13653	EHMLLVEQL	13694	QLGMLPRAM	13735	ANHWKHEWE
13613	SVKRFQHNN	13654	EYNGLGQAY	13695	QMVMLKDQY	13736	ASGWGHCWR
13614	TQQRVYSIE	13655	FCCLLYARR	13696	QQGQLKTWS	13737	AVIKAHCAA
13615	TYDRMQKCN	13656	FFSGLEGRH	13697	QVSCLVRDD	13738	AYCHQHQHA
13616	VPCRGLVVL	13657	FTDQLFTNP	13698	RTWTLECDF	13739	AYMPYHKEQ
13617	VPLRHLTDH	13658	GGNILDIGP	13699	RVNILEKIA	13740	AYQAEHVGM
13618	VTDRAFSMN	13659	GIIPLHWQP	13700	SFAHLIDDR	13741	CINWKHVAS
13619	VTPRHKTLF	13660	GSDLLVYSA	13701	SIAVLCHGN	13742	DDHWHHRAK
13620	VTTRVEHFP	13661	GSMLLGGME	13702	SIDALKSAF	13743	DFLTAHCIE
13621	VVMRDEGGR	13662	GTMLLWRQN	13703	SIVKLNRSV	13744	DFLYMHRMG
13622	WCHRVGCTG	13663	GVANLIGLT	13704	SLNKLFHQE	13745	DGDAAHTKP
13623	WDGRFQAAY	13664	HELYLTTNT	13705	SLQFLYHGS	13746	DHQLCHVAL

[1933]

13747	DHVDNHEEA	13788	MNLTIHEQF	13829	WTTAFHKRT	13870	IFPWFKFHW
13748	DNAYFHIKG	13789	MQMFKHNWI	13830	YIVQTHCGG	13871	IQVHSLFGC
13749	DPCNHHWCI	13790	MTGVEHVFY	13831	YKCHDHCHD	13872	IRPFIEFLQ
13750	DVVYTHCIN	13791	NCFWTHVSG	13832	YNNFHHQMY	13873	LHQKPNFDC
13751	EENHSHDER	13792	NDYPIHCYS	13833	YRNEPHQGF	13874	LIHHCTFEQ
13752	EINFQHRHV	13793	NFTSFHTYS	13834	YTNYKHCCF	13875	LQLEDRFNG
13753	EIQWNHWYV	13794	NHMLTHHYH	13835	TGCNHHGRF	13876	MCSHYQFMS
13754	EKDMHHVAS	13795	NYAFNHQGV	13836	DVCDDHKPN	13877	MFYLEDFRQ
13755	ELLKVHGPK	13796	NYQQHHFFS	13837	DVAPMHQVE	13878	MITQFIFNI
13756	EMNGRHWMP	13797	QHSWKHFDS	13838	VFKPNSFQT	13879	MTVMQGFLS
13757	EPAVTHDSA	13798	QILHYHYHG	13839	VPIQFSFAS	13880	NPQWGLFNQ
13758	EQHIMHEQW	13799	QMNWKHVWN	13840	VPLWDVFQT	13881	NTGARPFAG
13759	ERMVCHWDA	13800	QNSERHDEP	13841	VTGVTIFEL	13882	PCLQTRFAQ
13760	EWPINHQKV	13801	QRMFTHQEY	13842	YTPTRYFGK	13883	PMYQDMFTL
13761	FDKAFHVHM	13802	QSFSPHFGV	13843	YVGWHMFLD	13884	QGDLMGFHY
13762	FKPIVHHCG	13803	QSVHTHANL	13844	IWKIQEFGT	13885	QIGMNTFMS
13763	GEDQIHVWH	13804	QTGNAHWQE	13845	AGQPVTFHF	13886	QMHVMIFGS
13764	GHQSEHRCS	13805	RCNEDHEQD	13846	AHASHMFGG	13887	REWDHYFQI
13765	GINSHHNDL	13806	RIIEQHAHG	13847	ARQMICFGG	13888	RLRISPFNI
13766	GKQNNHRIE	13807	RQCEMHTVD	13848	ATYYNQFHD	13889	RMGAWSFPS
13767	GNKYMHHHA	13808	SCGWKHDNP	13849	CCTWKCFCQ	13890	RVTTAEFLC
13768	GPKPVHFDV	13809	SCTEMHAGR	13850	DFPLFTFGN	13891	RWSQNKFQR
13769	GSVMGHCGS	13810	SLYVHHKDG	13851	DIRHMSFEF	13892	SDHDNVFED
13770	GWGCGHGVH	13811	STGLQHDFN	13852	DPSQPGFNA	13893	SHTTYDFFK
13771	HGYNWHHLG	13812	TEGQTHHTA	13853	DVAHKCFIS	13894	SNVCNSFVN
13772	HTLHRHWFW	13813	TERSAHVTY	13854	EGGKERFQQ	13895	SPHAGPFGC
13773	HTQFAHGVD	13814	TFKWMHEGW	13855	EKMTWVFAG	13896	SQFMMDFGP
13774	HVAIGHGGF	13815	TFNNCHMQS	13856	ENTTCNFQG	13897	SQVPGRFIF
13775	IERHEHNPY	13816	THTQQHVVN	13857	ESLGCEFML	13898	TCCLYEFIH
13776	IIPHFHLCY	13817	TMSDRHICH	13858	FGCCGRFVI	13899	TGGTAEFGL
13777	IKCIQHVHY	13818	TNATSHWCG	13859	FTGKRDFMK	13900	TSNWHEFTN
13778	IKDITHNHP	13819	TNGWMHHWD	13860	FTMWHRFAW	13901	TSPCCTFGI
13779	LCRQTHKWV	13820	TTMIRHHGH	13861	GDNAKDFRG	13902	VHDAQGFCA
13780	LFGVYHFSG	13821	TVFGTHRAL	13862	GHAPHLFAD	13903	VKMMWDFAD
13781	LICKYHYCQ	13822	TVGMFHCQG	13863	GIRSTDFYQ	13904	VTHITEFMA
13782	LQPMMHTHS	13823	TYTPRHRFS	13864	GMIVDVFWV	13905	VVSTDMFIH
13783	LTLCMHEAD	13824	VRKWQHHGG	13865	GMTTMAFQG	13906	WCAPAVFNE
13784	MCGLYHTEG	13825	VTTHCHHTY	13866	HGAKAPFTA	13907	YFREKPFDV
13785	MCHGCHRSY	13826	WEAKWHAWE	13867	HNDYPQFQP	13908	YQCKMAFHV
13786	MFDNCHVFP	13827	WGFNDHKRG	13868	HVNAGSFGS	13909	GVWMQGFHP
13787	MNDTYHRRA	13828	WRGIGHADL	13869	IFILRTFFA	13910	MWGCKLFVC

[1934]

13911	YLDSSECCW	13952	ITDYQSDCY	13993	VFGDGAHCI	14034	ELTIIMVVA
13912	YVKPCYDCI	13953	LCALHEQCQ	13994	VHDQRLVCD	14035	ERDGWNMAA
13913	AGLQHDTCV	13954	LSVSSDACY	13995	VNPCAQGCH	14036	ESSLFEHYA
13914	AITHVMHCH	13955	LVVAHCGCV	13996	VNPQFSQCT	14037	ETQWWDVIA
13915	ASVLQQLCV	13956	MDAYFNSCQ	13997	VPMYTQVCN	14038	FVIPSGAGA
13916	ATTGMLACH	13957	MGTFVAHCY	13998	VSALIDHCC	14039	GGHNYGMHA
13917	AVGATMMCH	13958	MIALNQNCT	13999	VTGPMPTCL	14040	GIAASCHLA
13918	AVHDNFWCL	13959	MMPSYTLCW	14000	VTMQVRQCA	14041	GSCSTLTSA
13919	CCKYNQLCH	13960	MPAWMMNCH	14001	VVMQKIGCI	14042	GSLPYSMYA
13920	CDKIDPYCD	13961	MQQCDGGCG	14002	VVNTVGCCV	14043	GTMIYSHMA
13921	CDTDESGCY	13962	MSWSVYWCI	14003	WHLFSACCS	14044	HHGQCGGMA
13922	CHRCKVNCQ	13963	MTNCIEYCA	14004	WHMLCDGCS	14045	HTFLCNITA
13923	CITPYRPCD	13964	MTTLYQSCL	14005	WMRTQCWCP	14046	IISAKQHDA
13924	CMYSWVPCG	13965	MVHGALMCR	14006	YGGLSCGCL	14047	IKCIPGWYA
13925	CVHYCFGCS	13966	MVRSREQCI	14007	YGPSTEPCQ	14048	ILAQANIHA
13926	CVKSYEKCF	13967	NGGKGQQCD	14008	LSAQRLVCD	14049	ILHSVTVRA
13927	DEHAHWSCS	13968	NLGYCEHCY	14009	WVHPYSYCA	14050	INRQCDMIA
13928	DEHQMMECG	13969	NLMTPWPCQ	14010	TYHDIIVTA	14051	IQKYGAEVA
13929	DIGFRIDCD	13970	NSNMMAHCE	14011	VPNVQNMKA	14052	IQMLVQGSA
13930	DMCEFMYCD	13971	NWVPVYDCH	14012	VWAWQNGYA	14053	IRMHMPRDA
13931	DMFYWQPCE	13972	PFILSNCCA	14013	AEAIWLPWA	14054	ISKPYQLTA
13932	DRRDHEACN	13973	QNGVWENCE	14014	AECLGYVDA	14055	ITKCDGNTA
13933	EFDSPNQCY	13974	QSDGFEDCS	14015	AHVVHTSGA	14056	IVTASVCQA
13934	EFHPHDWCL	13975	QTTNNQRCD	14016	ALLFAPARA	14057	LNSLVCSFA
13935	ERGFMEHCH	13976	RDTKRSCCH	14017	AMPNGLDVA	14058	LTKIVTTGA
13936	ESMSTRACG	13977	RFVDFGSCQ	14018	ARLCTWPYA	14059	LWCVSNSFA
13937	ETDPYKSCH	13978	RHYYWVLCP	14019	ATNKNCAHA	14060	MEIGIMKIA
13938	ETVSPVGCH	13979	RNMYFVDCM	14020	CCFQNNTQA	14061	MERIMLSLA
13939	EVAQTMRCA	13980	RYDKVCYCV	14021	CLNGWKQAA	14062	MGWYIVRHA
13940	EYQLSETCT	13981	SGFSEKCCK	14022	CSQSYAADA	14063	MHNWGGKIA
13941	FHKCYSACR	13982	SGRPKLVCN	14023	CTNHNDHDA	14064	MHWLKYNYA
13942	FKDMIARCP	13983	SHDPHNTCG	14024	DFIMMQDQA	14065	MMAHSGHMA
13943	GDDCKFDCL	13984	SIGYYCHCM	14025	DPLVHTMIA	14066	MMDPCNPPA
13944	GTGSTWECT	13985	SILPTWGCD	14026	DRDNVDSQA	14067	MNGMPTSFA
13945	GVAVTTCCR	13986	SKHDFDHCF	14027	DRSEMGAYA	14068	MPDSWVRVA
13946	HSNMRSACE	13987	THILYSWCT	14028	DRSVVDVDA	14069	MRIAQCNGA
13947	IKRTYPACW	13988	TIFYCRWCN	14029	DTDMIAQFA	14070	MSMQHLEHA
13948	IMCECQHCE	13989	TKNEYGLCF	14030	DTWKVMMAA	14071	MSTHEDHVA
13949	IPFPMCGCC	13990	TVQYGRGCS	14031	EGQSMMYQA	14072	MVNTGGNTA
13950	IRAPFNRCE	13991	TWGLKWMCP	14032	EIVGQKGWA	14073	MVTCVQSGA
13951	ITCTYRNCE	13992	VDMVVDNCC	14033	EKQLYMYEA	14074	MWMMQGMNA

[1935]

14075	NCDQMSISA	14086	QVMCMCAAA	14097	TRCTCERYA	14108	VTTQIKRIA
14076	NCKLESDLA	14087	RTNLQQVDA	14098	TTDACWQHA	14109	VWHVPRDQA
14077	NDMYVKGTA	14088	SCAQERHDA	14099	TTHHAQAGA	14110	VYCGNKTKA
14078	NINGIINQA	14089	SILTGGGFA	14100	TTSMGKHYA	14111	YDAKSSVNA
14079	NRTDRMMDA	14090	SLVCHRNFA	14101	TVGLDAHSA	14112	YGREKQAYA
14080	NTCQHVNFA	14091	SMDVHCNDA	14102	TYWPGCANA	14113	YVHADPEFA
14081	PCSTWLLAA	14092	SMKLISEAA	14103	VCTAINDNA	14114	YVSQKEAIA
14082	QKMFGDSYA	14093	SMNATRWKA	14104	VIDTIAGHA	14115	YWLEVDSKA
14083	QNMSNWLDA	14094	SQACSNHYA	14105	VLSLFKEPA	14116	DVQGSGQWA
14084	QQVCDIPSA	14095	SYMLKSDYA	14106	VNQQIESDA	14117	SVLHVQALA
14085	QSTKEYWQA	14096	THVNRARWA	14107	VTQLVTRLA		

[1936] 实施例14

[1937] 在结肠中具有组织嗜性的AAV5变体

[1938] 此实施例描述了使用实施例1-实施例2中描述的方法和系统发现的在结肠中具有组织嗜性的工程化AAV5变体。

[1939] 分析CNS、肝脏、骨骼肌、心脏、肺、脾脏、淋巴结、骨髓、乳腺、皮肤、肾上腺、甲状腺、结肠、坐骨神经和脊髓组织以鉴定AAV5VP1 581-589区域中驱动结肠嗜性的氨基酸残基。结果显示于图18M中，该图显示了相较于所有其他分析组织在结肠组织中AAV5 VP1的581至589区域中每个氨基酸的位置频率，其中衣壳多肽具有SEQ ID NO:2的序列。位置(“Xaan”，n＝1-9)在x轴上表示并且氨基酸残基在y轴上表示。特定位置的特定氨基酸残基的频率在热图中作为强度表示。较深的方块表示特定氨基酸在特定位置的频率较高，较浅的方块表示特定氨基酸在特定位置的频率较低。

[1940] 分析相对于给定氨基酸残基出现在所有其他所收获组织中鉴定的变体中的指定位置的频率而言该给定氨基酸残基出现在结肠中鉴定的变体中对应于SEQ ID NO:1(推及于SEQ ID NO:2)的位置581至位置589的所述指定位置的频率，以确定优先靶向结肠组织的衣壳的一组序列规则。参考下文表27，和SEQ ID NO:2(AAV5 VP1)，以下氨基酸由此可以在Xaa1-Xaa9任何一个或多个位置处以任何组合独立地突变，以提供与野生型AAV5 VP1衣壳(SEQ ID NO:1)相比具有增强的结肠嗜性的VP1衣壳，其中结肠嗜性在此是指优选结肠转导的特性，而不是优选转导所有其他所收获组织的特性。

[1941]

表27 结肠嗜性规则
Xaa1选自C、F、H、N、P、W或Y
Xaa1选自F、P或W
Xaa1为P
Xaa2选自D、E、F、L或P
Xaa2选自D、E、L或P
Xaa2为P
Xaa3选自C、F、H、I、L、P或Y
Xaa3选自C、H或P
Xaa3为P
Xaa4选自C、D、E、N或P
Xaa4选自C、D或E
Xaa4为C
Xaa5选自D、E、G、P或W
Xaa5选自G、P或W
Xaa5为P
Xaa6选自C、K、R或V
Xaa6选自K或R
Xaa6为R
Xaa7选自D、M、P或V
Xaa7为P
Xaa8选自D、I、K、L、P、R或V
Xaa8选自K、P或R
Xaa8为P
Xaa9选自C、H、I、K、L、M或W
Xaa9选自I、L或M
Xaa9为I

[1942] 下文表28提供了在结肠组织中发现并且符合表27中提供的一种或多种规则的变体AAV VP1衣壳多肽的581至589区域的序列。因此，本公开提供了由具有SEQ ID NO:2的工程化AAV5 VP1衣壳多肽组成的rAAV，其中Xaa1至Xaa9区域是SEQ ID NO:10118-SEQ ID NO:11117中的任一者，如表28中所公开。本文还涵盖由工程化AAV5 VP2衣壳多肽和工程化AAV5 VP3衣壳多肽组成的rAAV，所述多肽在对应于AAV5 VP1 581至589区域中的氨基酸的AAV5 VP2(氨基酸残基445至453)和AAV5 VP3(氨基酸残基389-397)中的区域处具有下文表中公开的序列。

[1943] 表28

[1944] 驱动结肠组织嗜性的AAV5 VP1衣壳多肽中581至589区域的序列

[1945]

SEQ ID NO	581-589 序列
10118	PENIAEDHE
10119	PYALAEGNP
10120	PLAGRFVAQ
10121	PAEAKQRFD
10122	PAAYDGRLL
10123	PACHCYESL
10124	PACRFEANS
10125	PACSPYGWW
10126	PADFVNSNY
10127	PAECWRMTF
10128	PAEFDYLFT
10129	PAHEIKNLE
10130	PAIDHLLHR
10131	PAMQGHMVD
10132	PAMRTVHMW
10133	PAQWAWADG
10134	PARCAVSEH
10135	PASLKEHCV
10136	PATCSQKKY
10137	PATPPTIVG
10138	PAYHTWQYN
10139	PCAPYHPNQ
10140	PCDCSYYFR
10141	PCGSDIGCM
10142	PCHMWMNFY
10143	PCIWNWETD
10144	PCKPYPKDL
10145	PCNNCGHIR
10146	PCPMSPGEW
10147	PCQLKSDGE
10148	PCRDTLEGI
10149	PCRPLPHLL
10150	PCSMPKAGG
10151	PCTVEGRQY

[1946]

10152	PCYMKSWVM	10193	PEQLLYSMK	10234	PHSQKLNSG	10275	PLHQTSEYK
10153	PDCQEFWKQ	10194	PEREDCHNP	10235	PHTDAGYIH	10276	PLLMTKTGS
10154	PDGFSCRFA	10195	PESEGRMTK	10236	PHTQALICP	10277	PLMMGPQRQ
10155	PDHCLQSTK	10196	PESKLMWGA	10237	PIASEMYED	10278	PLNHACYCC
10156	PDHRDEMEW	10197	PETKTFCQQ	10238	PICEFCTMN	10279	PLNWIIACR
10157	PDHTNYHIV	10198	PEWTTTTVF	10239	PIEFYTIKH	10280	PLQTSKWQA
10158	PDIIIEGNT	10199	PFACGMSTW	10240	PIEGYHIAD	10281	PLVKLEEIH
10159	PDINYSCAI	10200	PFAEFVGRT	10241	PIESKPRSS	10282	PMCICWASE
10160	PDLRTQLFS	10201	PFAGGEYDG	10242	PIFPPGQKP	10283	PMFYDEPKH
10161	PDMCGRMDM	10202	PFCIPWTEG	10243	PIHHQLMCF	10284	PMGCNGYVT
10162	PDMQTEAGA	10203	PFCQPPEFN	10244	PIIPCGGWQ	10285	PMGYSNLNC
10163	PDPRRTNWQ	10204	PFDRMDHLA	10245	PIISTFNNQ	10286	PMHPTCFVI
10164	PDPVDHRGD	10205	PFGLSQCKG	10246	PILCTMPFV	10287	PMIVPDVAY
10165	PDPWSTSHK	10206	PFGQLKFNM	10247	PILIPSLQP	10288	PMLKTEAMS
10166	PDQQYELSC	10207	PFHVPPPTH	10248	PILPNTHTG	10289	PMLPEKFPD
10167	PDRNLSQGP	10208	PFKCQDIDC	10249	PIMCKEPGV	10290	PMMISCNSV
10168	PDRWCQLYK	10209	PFMSTYQVE	10250	PIMMHQTLG	10291	PMMWSTNVK
10169	PDSHVPLRD	10210	PFPVQNVDS	10251	PIMPKYNAL	10292	PMNWAHSCA
10170	PDTKYLHRC	10211	PFSKWVYEW	10252	PINTSHSYF	10293	PMPNYYHPQ
10171	PDVQNRWEM	10212	PFVSQWEHM	10253	PIPHVMNEP	10294	PMRPHRDAE
10172	PDVTTNNVI	10213	PFWWVCKVI	10254	PIRQNAMMW	10295	PMSPFPFEH
10173	PDWLMWSHH	10214	PGGYKEHTW	10255	PITLRQPWH	10296	PMTSTEPME
10174	PDWYHCGFQ	10215	PGIPIVWMH	10256	PIVAWIWYD	10297	PMWYYYHRM
10175	PEAEFGQKS	10216	PGKQTLTTQ	10257	PIVHQNPMN	10298	PMYSQVHLG
10176	PEAEYEQMK	10217	PGMMPGCTA	10258	PIVVDGRKF	10299	PMYVLEAHY
10177	PECHNYNQI	10218	PGPFRGPER	10259	PIYSIVTKG	10300	PNAKPCCFE
10178	PEDENPLYE	10219	PGVTCFSQT	10260	PKCEILTGA	10301	PNCLKSIHG
10179	PEDQHNTWC	10220	PHCLKAGPL	10261	PKGADRYNY	10302	PNCPFYPHP
10180	PEDWKRHAN	10221	PHDARVSRM	10262	PKGVAMGGT	10303	PNDLNRRTK
10181	PEEQTCEGR	10222	PHEDNNHQT	10263	PKIEPNLCK	10304	PNEPMTPGQ
10182	PEFPTWGHY	10223	PHELCHISH	10264	PKIMPKSCS	10305	PNESGHQSS
10183	PEGLWSVAC	10224	PHENKFLDC	10265	PKLEDDEHC	10306	PNFYSDISR
10184	PEHKGYAGW	10225	PHHGANGFC	10266	PKLTCETWF	10307	PNIKYVYNG
10185	PEISHKNVY	10226	PHHWCIVES	10267	PKSDLCEWV	10308	PNIPLIDMC
10186	PEKESTVIS	10227	PHKCMFSWN	10268	PKSELQKSF	10309	PNMPTCNAP
10187	PEKYCHEGY	10228	PHNYRVFWE	10269	PKSIDELFE	10310	PNQYTANEI
10188	PELAMPWEM	10229	PHPPNPGEN	10270	PKTWMHGVP	10311	PNRHGPDAE
10189	PELTYAKSI	10230	PHQHPQIPE	10271	PLCLKYEGC	10312	PNSSTPYNW
10190	PEMCKMGSP	10231	PHRDVIWLG	10272	PLECNRGTD	10313	PNWQNLESD
10191	PENWFWPHQ	10232	PHRVNVERS	10273	PLEMSNTPW	10314	PPDAFRWCP
10192	PEPSAMTIP	10233	PHSNSVGKC	10274	PLGFAPEIT	10315	PPDIQPNCS

[1947]

10316	PPELEKCTP	10357	PSAGDLSQY	10398	PVGSHMTFW	10439	PYTNEQQYT
10317	PPFCKWYIV	10358	PSAQGNFSM	10399	PVIHGMHID	10440	PYVQAQAGR
10318	PPILLIQMR	10359	PSCVIFKPW	10400	PVIKPNTEV	10441	PYVRRVTEI
10319	PPKDKYLRE	10360	PSDHTWKNV	10401	PVIQQRRDQ	10442	MPTRQRHEK
10320	PPKPAEEGH	10361	PSEGDSWLC	10402	PVIRKIYEL	10443	APILLQASC
10321	PPKYPQQQI	10362	PSESLELCP	10403	PVKCEDHPE	10444	DPRRSMWYL
10322	PPPSVRMMM	10363	PSFLCGKYI	10404	PVKSPWFNP	10445	IPRIVYMCD
10323	PPQQCEWIA	10364	PSGCYKCMK	10405	PVLCEMKQY	10446	SPCNMDTHH
10324	PPRHSNSQG	10365	PSKHKDLNE	10406	PVNWPPFFC	10447	RPVTYPHTA
10325	PPTIGCSCN	10366	PSLYCCDGV	10407	PVPRRKCSS	10448	NPGWTLSMA
10326	PQAEVMLGE	10367	PSMYEDMTN	10408	PVPTFCSTC	10449	IPSHVHFGE
10327	PQCRMLLWE	10368	PSNPEQHLV	10409	PVPYETHMY	10450	APDNEAGMY
10328	PQDDKGVIP	10369	PSPRYPVCM	10410	PVRKAGWQK	10451	APDPSTANR
10329	PQECSHVHP	10370	PSVLQTTFN	10411	PVSPPCDAQ	10452	APFNANCCN
10330	PQKLNFTYC	10371	PSVTVDVGS	10412	PVVMIGHMS	10453	APGCYFYRN
10331	PQKTLAVGN	10372	PTCLGKIKW	10413	PWDLRNQYH	10454	APICEGEPK
10332	PQLVSHIGH	10373	PTCQTPPLH	10414	PWEAMYMTA	10455	APITIGAMS
10333	PQNPKMIPA	10374	PTEFDQEQT	10415	PWERRHAER	10456	APKWFNSLD
10334	PQNTFVRKQ	10375	PTEPREDAA	10416	PWETCRWKQ	10457	APLHNFPVT
10335	PQPCSYDSG	10376	PTETRWHTG	10417	PWIAMVAQE	10458	APMDMGRMT
10336	PQRYIFGAW	10377	PTFMFRRDL	10418	PWNQKCNYP	10459	APMHSFDAE
10337	PQSCENNCR	10378	PTGIYAMQC	10419	PWNWDGHMR	10460	APNDCMMQN
10338	PQSHCHMET	10379	PTHLFNFHW	10420	PWQCLRSPD	10461	APPHQDHNP
10339	PQTFWGETC	10380	PTINRYFNE	10421	PWVRWRTCQ	10462	APQCVIVFT
10340	PQVMQLAEA	10381	PTKEMMSMF	10422	PYAANHQQG	10463	APRQIWMHH
10341	PQYIQAIED	10382	PTKSSGGAA	10423	PYACKIIRG	10464	APVYWDHQM
10342	PQYRCIDEH	10383	PTMLAVKNA	10424	PYCNIWMES	10465	APWWEFPQY
10343	PQYTNRVIG	10384	PTMPWKKAE	10425	PYCPMQIDC	10466	APYDPKYVF
10344	PQYVPDRGT	10385	PTMYRPMHV	10426	PYCWMVDDA	10467	APYQAVMES
10345	PRCLKCRPH	10386	PTNGIASAH	10427	PYEHGKEHH	10468	CPALNQLGR
10346	PRGANSMMT	10387	PTPSIVLCQ	10428	PYGYNCVID	10469	CPAMPGYCH
10347	PRLFMREID	10388	PTQYIGDHR	10429	PYHRIMEYA	10470	CPDDSRLDQ
10348	PRLNVELDW	10389	PTRIQMTVD	10430	PYIRCKKDG	10471	CPDLPDWSG
10349	PRLQWMEPG	10390	PTVWANVAE	10431	PYMRCEHQH	10472	CPFFNYMIQ
10350	PRMCMSENE	10391	PTVWRTFEF	10432	PYMVTCYDF	10473	CPGYYETEI
10351	PRMWDKLDQ	10392	PTWYVLAHF	10433	PYNPQYCCC	10474	CPHAHYVWR
10352	PRRDKDVYM	10393	PTYPFISYS	10434	PYPWDRCWI	10475	CPIYYGPVS
10353	PRVCVTWVT	10394	PVAGDRYGG	10435	PYRKPWNAE	10476	CPKGHEYEN
10354	PRWMQTYQV	10395	PVAVPTESP	10436	PYRVEGEHN	10477	CPKQWLPAC
10355	PSACMVGPC	10396	PVDRLVHTN	10437	PYSFKDLLE	10478	CPLIDWCMY
10356	PSADCNPWA	10397	PVEMPDSCH	10438	PYTIQKGEY	10479	CPLPARWGD

[1948]

10480	CPNCWYDEV	10521	EPGIVHWYD	10562	GPHWYWHDM	10603	KPESKMLSE
10481	CPPQSENKV	10522	EPGYYDGGN	10563	GPINPYAHG	10604	KPFVSECTE
10482	CPPRRYVNG	10523	EPHKYNQPS	10564	GPLSKHFDV	10605	KPGFRYFAE
10483	CPQCMQGAS	10524	EPHLQPHQK	10565	GPMHNTGPP	10606	KPGTVKDCD
10484	CPQMPWLHF	10525	EPHLTAVPT	10566	GPPPSDLRS	10607	KPILMSEEA
10485	CPRQLYHWQ	10526	EPKNYSERH	10567	GPPWKHCDA	10608	KPITEGRRD
10486	CPSCEISVP	10527	EPLKYNHES	10568	GPSCYLDYW	10609	KPKCCQYNW
10487	CPSECDNLK	10528	EPLQFMELM	10569	GPTIWIHTY	10610	KPMCDGVAY
10488	CPSKQTRNP	10529	EPMRLGEWS	10570	GPTIYQYMQ	10611	KPMCLYVLT
10489	CPTQLYLHN	10530	EPPVCQGNE	10571	GPTKLNYNN	10612	KPNDCLREG
10490	CPVCEDMPL	10531	EPRRVDDQQ	10572	GPVVPCNSW	10613	KPPMPTNTG
10491	CPVLFTHME	10532	EPVMVRMQF	10573	GPVWEFEWM	10614	KPPYSCREQ
10492	CPWLELCYH	10533	FPALIKPHE	10574	GPWHACKIN	10615	KPQHENKQL
10493	CPYCDQEGH	10534	FPASPDILI	10575	HPCCDRYLA	10616	KPSMYPRNS
10494	DPAGDLMQG	10535	FPCGIVWTH	10576	HPDAPGPIQ	10617	KPTDMVGAW
10495	DPCNVDPGF	10536	FPCQNGPHN	10577	HPDVRNYEN	10618	KPTQIGGEH
10496	DPCTCHTAK	10537	FPDRMDYVE	10578	HPEVDSSGA	10619	KPVEPTTNP
10497	DPDPVDMYA	10538	FPGGQKPCL	10579	HPGCEYSWE	10620	LPAAYFWGA
10498	DPGAPGPKA	10539	FPHFTCWCD	10580	HPGIEVDRG	10621	LPAWYRARC
10499	DPHCIPTWK	10540	FPHGYAMIW	10581	HPGWTHAVF	10622	LPIEQQKSW
10500	DPHCVKNNC	10541	FPIQGNHFP	10582	HPHDVFHET	10623	LPIESEGEI
10501	DPICVVWTG	10542	FPKAPKYYL	10583	HPILPAQYF	10624	LPLGQADEA
10502	DPIHAHVEC	10543	FPKNNLTKC	10584	HPISPWQWL	10625	LPLVYAKNW
10503	DPKHYNETM	10544	FPLCVDPGP	10585	HPLSRGSES	10626	LPLYFHEHY
10504	DPKICWQER	10545	FPNIDNPKA	10586	HPMFEAYQN	10627	LPMDEKEGA
10505	DPKSCPRRE	10546	FPPPIHLAI	10587	HPNFIWCNH	10628	LPMQCDETY
10506	DPKSGFNGT	10547	FPPQQEFIH	10588	HPNGNEREH	10629	LPMVWGNPK
10507	DPMCFHAYQ	10548	FPQWGKEPI	10589	HPNPWGVFW	10630	LPMYIEHMC
10508	DPMQMHRSG	10549	FPRTTNTRV	10590	HPRFNRLQV	10631	LPNPVNATQ
10509	DPPKCYRYE	10550	FPSHDDFLG	10591	HPSFIKCQV	10632	LPNWFCRNY
10510	DPPNTSFFH	10551	FPVTCGTVF	10592	HPYLCTIRI	10633	LPPAHMFYC
10511	DPPRCDMQK	10552	FPWYSLILT	10593	IPDAWWANM	10634	LPQHYGIWH
10512	DPPSPVMCG	10553	FPYMAEQYR	10594	IPHRNIHPS	10635	LPQPYPDNY
10513	DPQLFCSDL	10554	FPYWDIDDT	10595	IPLWETMNM	10636	LPQRGTPIV
10514	DPRMGGSVN	10555	GPCMLYEMC	10596	IPNDRDRCA	10637	LPRHIYKST
10515	DPSYNERHR	10556	GPDPWSQWW	10597	IPRVHWKAE	10638	LPSHELWTW
10516	EPCHWMHEF	10557	GPDTEYRGC	10598	IPTFRTMKI	10639	LPTSDCRAE
10517	EPDQMQERL	10558	GPFADPPAT	10599	IPWFDFAME	10640	LPTYNKYVQ
10518	EPDQPRWDS	10559	GPFLSVPAH	10600	KPAPWMRME	10641	MPAHANSTC
10519	EPECTEFKV	10560	GPFRFEWEW	10601	KPCPPHAYV	10642	MPCTCSYPH
10520	EPFPHVMWG	10561	GPGTLRHLL	10602	KPEAWAVCG	10643	MPDNFGNGP

[1949]

10644	MPIMMNVSP	10685	QPDCEWPHN	10726	SPLQNYEER	10767	VPWMVLDNN
10645	MPMDLNYIQ	10686	QPDMYDCYA	10727	SPMDTTEPK	10768	WPALFVADA
10646	MPMFCQIEQ	10687	QPEFTYLQF	10728	SPMQKQKVS	10769	WPDPFIHAT
10647	MPPFHENES	10688	QPHSSVTLD	10729	SPQTCSFPT	10770	WPEATHRFA
10648	MPPPVQDWN	10689	QPPAAAKQD	10730	SPRCMTNGG	10771	WPKYDMPLH
10649	MPPVRLYSY	10690	QPTDDVIQG	10731	SPSGIMEGN	10772	WPMCPVCYF
10650	MPQDCGWNL	10691	QPTQRMYHC	10732	SPYQCPQAS	10773	WPMTRTNGH
10651	MPRDGTWWC	10692	QPVTFYEKH	10733	TPDHKTRAH	10774	WPNYHSWYE
10652	MPSHGYWYQ	10693	QPVWFPKCW	10734	TPDPDLNNR	10775	WPPNGKWFM
10653	MPSPYGEAY	10694	QPWMVWNSH	10735	TPGSDMSMV	10776	WPQRGTMDM
10654	MPSQGVIWH	10695	QPWMYVMGS	10736	TPHSSMKVG	10777	WPTGHCAYW
10655	MPSVWGPDM	10696	QPWTFWWNR	10737	TPKLEFPVD	10778	WPYYCCIEH
10656	MPTEYEAHG	10697	QPYCERTGI	10738	TPLLTSGHE	10779	YPANKCEWE
10657	MPTQAYMPQ	10698	RPAVDDEEC	10739	TPLMAYWIP	10780	YPASYDCHP
10658	MPVDIANQM	10699	RPCPNMTDG	10740	TPMDYEDQE	10781	YPDTVQKSP
10659	MPWWYGCIC	10700	RPDVRNELC	10741	TPMFDRHSL	10782	YPDWWEWHT
10660	NPCFMIWYT	10701	RPELGGQQT	10742	TPMLGMLWW	10783	YPEGEMNYE
10661	NPCMLQDHH	10702	RPETTAERP	10743	TPMWQASWQ	10784	YPFQPIEEP
10662	NPCTFWNNE	10703	RPHHFWVPF	10744	TPNCMGQRM	10785	YPFWAGWGD
10663	NPGHAWATM	10704	RPKECEQMW	10745	TPNVWEQQM	10786	YPGMKICGH
10664	NPGWWCRMW	10705	RPLDTSQED	10746	TPPHKCSVG	10787	YPIIMLPCR
10665	NPIMYPCLS	10706	RPMFGAVAG	10747	TPQCAMWNE	10788	YPMQGGIRW
10666	NPLIRDMWT	10707	RPMHGDYDD	10748	TPSQVASDD	10789	YPNDAKHGN
10667	NPLKLDYPE	10708	RPNCVMYDK	10749	TPSRNMVDL	10790	YPNLHHSMM
10668	NPPWLQPFS	10709	RPPDEIWGH	10750	TPYIRHHFE	10791	YPSDGLFEQ
10669	NPQFLHFQI	10710	RPRDNGVEG	10751	TPYTWCWPD	10792	YPSGHKPYW
10670	NPRAMDAAH	10711	RPWSSEFRV	10752	VPCPKWSMI	10793	YPSINCPTN
10671	NPRHDAGHS	10712	SPAEIPSIF	10753	VPEQGRWEA	10794	YPSPPERLE
10672	NPSWNACNV	10713	SPCPNGVRQ	10754	VPKLLDTDQ	10795	YPVNCTQKW
10673	NPTMPGIRG	10714	SPDVGWTES	10755	VPLLILNNN	10796	YPWENNAHL
10674	NPTQDFKVM	10715	SPEVTEYGF	10756	VPLVSYSQQ	10797	YPWMTACHD
10675	NPVTGEYSL	10716	SPFMWWRWW	10757	VPMMCQWSQ	10798	YPYICEFVW
10676	NPWHYTYAF	10717	SPGFILGGY	10758	VPQHKNEFQ	10799	YPYPRLWEC
10677	NPYLHMFNQ	10718	SPGKFWHPC	10759	VPQWIDERW	10800	NEPYCIWAY
10678	NPYNQHCIC	10719	SPHFQKHPS	10760	VPRHWAGVR	10801	GRPVMHRHS
10679	NPYRTFTDG	10720	SPIRIDWQG	10761	VPRYNSDEQ	10802	DSPSNHPAV
10680	QPAFEHAQV	10721	SPISDLRNC	10762	VPSCQCSES	10803	YYPFPGPTD
10681	QPAHMEVMY	10722	SPIWEGERT	10763	VPTIKVTRA	10804	AAPIPGKAF
10682	QPAHPGTGL	10723	SPKKAEQAI	10764	VPTQSVKGC	10805	AFPNVMYQN
10683	QPAQYAVMH	10724	SPLHVHWQS	10765	VPTSYWCTN	10806	AGPKMIEPW
10684	QPCDITYQS	10725	SPLISVLEG	10766	VPWFVMLMM	10807	AHPQQPHVT

[1950]

10808	AIPMIGSYC	10849	EYPSYWSDG	10890	HKPEQMEAY	10931	LHPWMPAGR
10809	AIPQEGVCC	10850	EYPYDVPNF	10891	HKPKMWMYG	10932	LKPSWQDKA
10810	AKPYKFSGE	10851	FAPEICGCC	10892	HNPLIRTMV	10933	LLPMDEMKT
10811	AMPIVMWHT	10852	FDPVCKCMK	10893	HNPLQENDC	10934	LMPCHGDIA
10812	ANPADWHWY	10853	FEPQNPTFQ	10894	HQPELMTDN	10935	LMPELGCTD
10813	ANPFVLAPW	10854	FFPTMHARF	10895	HRPDWNDGK	10936	LSPLIAGIY
10814	ARPWWMLEQ	10855	FIPCVISNQ	10896	HRPPCEMGD	10937	LTPDVRINA
10815	ASPPDYKAQ	10856	FLPFWKAFH	10897	HYPAYQKAC	10938	LTPSCRHKW
10816	AVPQMTGWC	10857	FLPTSKLHS	10898	HYPWTPWTP	10939	LVPTLINCL
10817	AYPHCEAVM	10858	FLPTYDGMH	10899	ICPTKGMEA	10940	LWPSSCQMK
10818	AYPSGIRAH	10859	FNPPIRQLW	10900	ICPVCSWDC	10941	MAPQPMGGG
10819	CCPVDNEFW	10860	FQPFTSNLL	10901	IDPEIYAWC	10942	MCPKTFRFE
10820	CDPPIDCYI	10861	FQPYMTSTS	10902	IDPLDGNHF	10943	MEPKKMKVW
10821	CDPSYELNI	10862	FRPIISMKR	10903	IHPQQWWDH	10944	MEPMSRDHL
10822	CDPWHKYNG	10863	FRPMITCWW	10904	IIPHQNIMW	10945	MEPWEQCHI
10823	CGPFRHQNG	10864	FSPPMVMKY	10905	IKPEPMWMT	10946	MGPRYDKEW
10824	CHPQEWHGQ	10865	FSPYFMVPH	10906	ILPPEKHQA	10947	MHPDSYNVT
10825	CIPDFISEY	10866	FTPDSLIVH	10907	IMPNNLEMM	10948	MKPDQNHHK
10826	CLPWVMSIA	10867	FTPECTRGM	10908	INPRKLSSK	10949	MLPLPNEFQ
10827	CMPPDCKCP	10868	FVPHDCNRD	10909	IQPMYVIYQ	10950	MLPMAWYQI
10828	CNPGWPKPG	10869	GAPFDELPH	10910	IQPWDYVRP	10951	MLPMRQLTQ
10829	CTPYPPRDQ	10870	GCPLIVPDL	10911	ISPLNICWF	10952	MLPWSSRSN
10830	CVPNSHIWQ	10871	GDPLVIWLC	10912	ITPEVLKTT	10953	MMPDILRSY
10831	CYPTLEKHA	10872	GFPREMYTE	10913	KCPLGKPAS	10954	MQPCWQPQC
10832	DFPLNPVQT	10873	GGPEKEAHR	10914	KCPRGGMNC	10955	MRPQQVWNG
10833	DFPLQVTGY	10874	GGPPIDEGN	10915	KDPCCGCHN	10956	MSPQPELSG
10834	DFPRNNAIM	10875	GHPFKDYHS	10916	KGPPHKDDQ	10957	MVPAAGPGE
10835	DGPGTQCIK	10876	GMPWEDISN	10917	KHPNIWCAN	10958	MVPKMMRSQ
10836	DLPAGQRMF	10877	GNPCSPKIN	10918	KLPDTYKNF	10959	MWPAMKVCW
10837	DLPLPKLMQ	10878	GNPKLLESF	10919	KMPPPAKDI	10960	NCPQPNCGW
10838	DNPGYLKEE	10879	GNPPTWLCP	10920	KTPHCIKHY	10961	NMPERKLTF
10839	DSPNYVWTH	10880	GNPYMYPWH	10921	KVPVLTAGG	10962	NMPIGGYCC
10840	DTPEYAPSM	10881	GRPWDPYFE	10922	KWPTFKWPK	10963	NNPQKQQSR
10841	DTPLQYKAQ	10882	GTPQWWNKG	10923	KYPEPGRGT	10964	NNPSNSLIW
10842	DYPKGEKCD	10883	GTPSMHMGI	10924	LAPQLGWRN	10965	NTPEFHWQQ
10843	EEPELQRLN	10884	GVPQDRIQE	10925	LCPGEWLVI	10966	NTPLDFTGE
10844	EHPPHAFAH	10885	HCPNQGRRS	10926	LEPKDLAAP	10967	NTPNNMEAR
10845	EKPLYYHHA	10886	HFPRSEAHR	10927	LEPPSTVNV	10968	NTPQQGQKG
10846	ELPATLFQF	10887	HGPMEVMHE	10928	LFPNQCNFV	10969	NVPTEPMDH
10847	ENPQFVRSI	10888	HHPFAKPII	10929	LFPQHIAQP	10970	NWPVSDEDM
10848	ESPCHQNQG	10889	HIPQHSRMC	10930	LGPHRWYGH	10971	NYPCTDKNH

[1951]

10972	NYPVGYHEL	11009	SMPHACYNG	11046	VKPYNGLVQ	11083	SCECPCCLA
10973	QAPMVQDTQ	11010	SMPTTSSSC	11047	VLPHFWNQG	11084	VCECCRQNI
10974	QAPWKHEGP	11011	SSPDNWKHM	11048	VLPKSHFFP	11085	ALCCFCCQG
10975	QCPDEKAEF	11012	STPLGNSAP	11049	VMPEESVRN	11086	TSYCIQTVD
10976	QCPIQSVVC	11013	STPMVCEMG	11050	VMPTNILEE	11087	DQECCQTLF
10977	QDPFWRGLH	11014	SVPHEVHCE	11051	VMPYAQGIS	11088	QTQVPIWNT
10978	QHPDQGQHC	11015	SVPLKFWTL	11052	VQPCFMMDH	11089	LATHPPYTY
10979	QIPYEMWWS	11016	SWPDEPSVA	11053	VQPFPYCCE	11090	LTYNPTMIG
10980	QLPCIKKCP	11017	SWPHCEIEI	11054	VQPHPFNPH	11091	TGIIPIRVC
10981	QLPGCRTDQ	11018	SYPLTRWPA	11055	VQPPILEQH	11092	TCQEPHNIT
10982	QLPKMWLPN	11019	SYPYMHMQD	11056	VQPWWDAMG	11093	QVYAPEIEG
10983	QMPNGEIVT	11020	TAPQCVSDA	11057	VRPLHNFID	11094	DYQSERLWS
10984	QNPDDVINL	11021	TCPNLTAYC	11058	VSPANADND	11095	KEQEQRPVR
10985	QNPSPMMHR	11022	TCPRNLCNY	11059	VTPNLQVQS	11096	ISMQNRCAD
10986	QNPSTYCHY	11023	TDPLCKHQT	11060	VTPYDWKFQ	11097	IFQLSRPAW
10987	QQPWPHQRG	11024	TDPLNPLCD	11061	WAPSPRQDR	11098	ECKTHRDVY
10988	QRPTMLQGY	11025	TEPPCPGQT	11062	WDPVYGERN	11099	YHNQDRHYL
10989	QSPPIFPTT	11026	TEPSIMNEV	11063	WEPDGIVPQ	11100	KVDTVRNPK
10990	QSPWKLWSH	11027	TFPCNRFGI	11064	WHPKCYDVF	11101	TVSRGRPVA
10991	QWPDFMVTI	11028	TFPSIQWSR	11065	WHPMDKNFR	11102	NQAYGRGHV
10992	QYPQLPHWA	11029	TGPNFQPAL	11066	WIPTEIYGH	11103	SVRLNEPHV
10993	REPFLHVNS	11030	TIPFCKKCE	11067	WKPEGQNFG	11104	CLVNDMPRK
10994	RFPVKLGAM	11031	TIPYEANCA	11068	WLPRMPSCD	11105	AHTFEAPKW
10995	RHPQLGLPH	11032	TKPQITEIS	11069	WMPCCPHGC	11106	IWNYCVPKH
10996	RIPWAELMK	11033	TQPCMPNRT	11070	WMPGMHWRC	11107	HWEQHNPIW
10997	RKPWREHNI	11034	TRPPSSLDE	11071	WMPSTTRWN	11108	CIDWAEMPQ
10998	RLPYYVQFC	11035	TTPTWQNSP	11072	WQPDWMYEH	11109	ACVMYDCPP
10999	RNPCSKHEV	11036	TVPHFYFYG	11073	WRPPVRVQD	11110	WVGNANRPA
11000	RNPRDYPMF	11037	TWPMSCDCD	11074	WSPERERFE	11111	SVEVGCSPG
11001	RQPANRFKE	11038	TYPFGWDKH	11075	WWPIWWIPQ	11112	CVFSTNHPA
11002	RTPFPHDSL	11039	VAPAILYRC	11076	WYPLDVLAS	11113	AMRQIIAPC
11003	RWPSEKEKF	11040	VAPIHDHRI	11077	WYPMLWELE	11114	TVVFSENEI
11004	SAPPSFAQK	11041	VDPAMMPDN	11078	LSICQPVMP	11115	WKGRWTTTI
11005	SAPVEWMRK	11042	VFPTASQLC	11079	TTRCFRHAK	11116	NSHYTEGVI
11006	SCPNRESGR	11043	VHPEQTLIG	11080	WATCDAPPV	11117	AWTPTEKQI
11007	SGPRTDCPL	11044	VHPFHKYEQ	11081	NYVCSTSGY		
11008	SGPVVMIQS	11045	VIPVHNMCG	11082	LTKCMNKGW		

[1952] 实施例15

[1953] 在甲状腺中具有组织嗜性的AAV5变体

[1954] 此实施例描述了使用实施例1一实施例2中描述的方法和系统发现的在甲状腺中具有组织嗜性的工程化AAV5变体。

[1955] 分析CNS、肝脏、骨骼肌、心脏、肺、脾脏、淋巴结、骨髓、乳腺、皮肤、肾上腺、甲状腺、结肠、坐骨神经和脊髓组织以鉴定AAV5 VP1 581-589区域中驱动甲状腺嗜性的氨基酸残基。结果显示于图18L中，该图显示了相较于所有其他分析组织在甲状腺组织中AAV5 VP1的581至589区域中每个氨基酸的位置频率，其中衣壳多肽具有SEQ ID NO:2的序列。位置

("Xaan",n＝1-9)在x轴上表示并且氨基酸残基在y轴上表示。特定位置的特定氨基酸残基的频率在热图中作为强度表示。较深的方块表示特定氨基酸在特定位置的频率较高,较浅的方块表示特定氨基酸在特定位置的频率较低。

[1956] 分析相对于给定氨基酸残基出现在所有其他所收获组织中鉴定的变体中的指定位置的频率而言该给定氨基酸残基出现在甲状腺中鉴定的变体中对应于SEQ ID NO:1(推及于SEQ ID NO:2)的位置581至位置589的所述指定位置的频率,以确定优先靶向甲状腺组织的衣壳的一组序列规则。参考下文表29,和SEQ ID NO:2(AAV5 VP1),以下氨基酸由此可以在Xaa1-Xaa9任何一个或多个位置处以任何组合独立地突变,以提供与野生型AAV5 VP1衣壳(SEQ ID NO:1)相比具有增强的甲状腺嗜性的VP1衣壳,其中甲状腺嗜性在此是指优选甲状腺转导的特性,而不是优选转导所有其他所收获组织的特性。

[1957]

表 29 甲状腺嗜性规则
Xaa1 选自 A、K、M、N、Q 或 R
Xaa1 选自 K、N 或 Q
Xaa1 为 K
Xaa2 选自 A、F、K、L、M、T、V 或 W
Xaa2 选自 F、V 或 W
Xaa2 为 W
Xaa3 选自 A、I、K、R、S、T、V 或 W
Xaa3 选自 A、R 或 T
Xaa3 为 R
Xaa4 选自 A、D、E、I、P 或 V
Xaa4 选自 A、E 或 I

Xaa4 为 A
Xaa5 选自 F、I、M、Q、V 或 Y
Xaa5 为 M、V、Y
Xaa5 为 M
Xaa6 选自 H、M、N 或 Y
Xaa6 为 N
Xaa7 选自 H、I、N、Q、S 或 W
Xaa7 选自 H、I 或 N
Xaa7 为 H
Xaa8 选自 A、D、F、Q、S 或 Y
Xaa8 选自 A、F 或 S
Xaa8 为 F
Xaa9 选自 A、Q、S 或 Y
Xaa9 选自 A 或 S
Xaa9 为 A

[1958]

[1959] 下文表30提供了在甲状腺组织中发现并且符合表29中提供的一种或多种规则的变体AAV VP1衣壳多肽的581至589区域的序列。因此，本公开提供了由具有SEQ ID NO:2的工程化AAV5 VP1衣壳多肽组成的rAAV，其中Xaa1至Xaa9区域是SEQ ID NO:40438-SEQ ID NO:41437中的任一者，如表30中所公开。本文还涵盖由工程化AAV5 VP2衣壳多肽和工程化AAV5 VP3衣壳多肽组成的rAAV，所述多肽在对应于AAV5 VP1 581至589区域中的氨基酸的AAV5VP2(氨基酸残基445至453)和AAV5 VP3(氨基酸残基389-397)中的区域处具有下文表中公开的序列。

[1960] 表30

[1961] 驱动甲状腺组织嗜性的AAV5 VP1衣壳多肽中581至589区域的序列

[1962]

SEQ ID NO	581-589 序列
40438	KAEDSQPDC
40439	KAMQVLENC
40440	KASWIPSNK
40441	KCEDNTFAS
40442	KCTDLGEYF
40443	KDHEEHGPK
40444	KERTAWSRE
40445	KFMQDPCNS
40446	KGELMCLDV
40447	KGPKECAAT
40448	KHILWGSLA
40449	KIGISEQDF
40450	KKYTDQDYG
40451	KMTPVEADA
40452	KNPTVFWKM
40453	KQAIYLDLD
40454	KSAMQEGVA
40455	KSFGSEVFF
40456	KSKWGHYWD
40457	KSMDLGQFM
40458	KSNQTYCAA
40459	KSYCLHMKQ
40460	KTGMACDIN
40461	KTLVKQMDY
40462	KTVWVYRVH
40463	KTYPHCFPV
40464	KVEDGLQWH
40465	KVEGMVTQS
40466	KVIEFCGEE
40467	KWLEANMTG
40468	KYAFRYEAS
40469	KCLQMCPNG
40470	KAIGWFHFD
40471	KAYTVEMDP
40472	KCCLPWQNI
40473	KCEQWQKVC
40474	KCHVGGNVY
40475	KCNPWSEPT
40476	KDDGDDNIP
40477	KDRDQVHWM
40478	KEGNMDACS
40479	KEYDPQHWW
40480	KFIYGQREV
40481	KFKQEDAQM
40482	KGIWQDHSC
40483	KHCENCEYF
40484	KIAPVAASP
40485	KIEMLCCQG
40486	KILNPNDSY
40487	KKRRDTQIY
40488	KLHIWNVPC
40489	KMQPNGEWP
40490	KNDQTTYPE
40491	KNEPDNLCE
40492	KPCTFAPFM
40493	KPEWAMMYD
40494	KPYQDIARM
40495	KQEIQGVCR
40496	KQETRTTDH
40497	KQFWMKVEV
40498	KQMLQQCFG
40499	KQWELYQID
40500	KRVVKFTST
40501	KRYTHGKNW
40502	KSADMPTTL
40503	KSFLHPAAH
40504	KSFYMCRCG
40505	KSYVGSYES
40506	KTCELTAMR
40507	KTEMSDMHW
40508	KTLVCVNWF
40509	KTYLWEHES
40510	KVSEAHYDL
40511	KVTTETGVP
40512	KVVAVNQSL
40513	KWQQAEYDE
40514	KYCHWWYKL
40515	KYQNAKPVA
40516	KKNSMDCQV
40517	KPYFKDCPR
40518	KSIFQATLQ
40519	KTAHDKDID
40520	KHSSWCCYT
40521	KSEHHWINA
40522	KVGMDQTAV
40523	KFFYTELMH
40524	KDFNFEAQG
40525	KDQQAERHV
40526	KDYANFGMN
40527	KHIHCTIQC
40528	KNPCLVWQK
40529	KSYEMVNQW
40530	KTGWDSHWS
40531	KVAYMWYNW
40532	KHGVLPMPA
40533	KTLHMLESK
40534	KCGLTPGVE
40535	KVDTVRNPK
40536	KAHDITWNL
40537	KAEMKMGCN
40538	KTTIVPAWN
40539	KAIPQGSSC
40540	KMTNSNENS
40541	KWYQFGMVA
40542	KFDVIRQAM
40543	KQLTFPMLA

[1963]

40544	KQEWHVHTE
40545	KIDQTQWQR
40546	KRFVDMPEM
40547	KLYGLNHSC
40548	KIQPMMVKA
40549	KMLHKVWID
40550	KGPIFHSAY
40551	KGMQMWNCE
40552	IWKFKNAHL
40553	IWNDQWDCP
40554	IWNWMTEDC
40555	IWTLNWAQL
40556	LWEQWSTLS
40557	LWQFCSICE
40558	LWVECHCNQ
40559	MWVELPSDD
40560	NWIHWNACF
40561	NWMERDHNH
40562	NWSYCNVSA
40563	NWTHQDMFD
40564	NWYFMSQTD
40565	QWAKWDWSS
40566	QWHQFNRMH
40567	QWTTITKEE
40568	RWKSQEEIQ
40569	SWAFANHTH
40570	SWEQNVVST
40571	TWTAMEIYG
40572	SWSPTGVAN
40573	AWDHTMPYG
40574	AWGNRYHYC
40575	AWHDYFRDN
40576	AWMDCSIMC
40577	CWQAALHSA
40578	CWWCALLST
40579	DWIMDTPKP
40580	DWIVDTPKP
40581	DWNYYACYT
40582	EWICVQSHN
40583	GWMIAMDGK
40584	GWNTFIWGQ
40585	GWTLNPSEG
40586	HWIPQWARS
40587	HWISFERHQ
40588	IWDFKCTQC
40589	IWKDYLHVG
40590	LWCNENAMD
40591	LWQQIPWGK
40592	MWAAMPVGH
40593	MWEQEPNRN
40594	MWFAENNGL
40595	MWGINGVAY
40596	MWRPARYKS
40597	NWEDSPWCY
40598	NWEPCWWAH
40599	NWGPQQSER
40600	NWNKTRFTC
40601	NWVSHDKEA
40602	PWAETEKPN
40603	PWIPNQIKQ
40604	PWTCNENQW
40605	QWPEMTCCD
40606	QWTRTGSSF
40607	RWHNGENCS
40608	RWNARWELG
40609	RWRCRMQGN
40610	SWIDMAMMQ
40611	SWTALGSVP
40612	SWTLRMWAS
40613	TWLPVGAAH
40614	TWVNFPTHC
40615	VWWFAPMEA
40616	WWAQPWHAL
40617	WWTDMSSWT
40618	YWWLDHANF
40619	GWTTNLHYH
40620	NWVVWPMEN
40621	RWVCANAQD
40622	CWECYQFKR
40623	CWSAIMTGC
40624	DWMQGGIYN
40625	EWFEDEVLW
40626	EWGMHTSNR
40627	FWSGYNAEK
40628	GWTHFHWQE
40629	LWKTILEGV
40630	RWGMYDSYY
40631	RWTDTNFRG
40632	TWAMMGQWW
40633	LWKLGNKVF
40634	TWTSRCELL
40635	VWNPDLKMN
40636	EWTKKTADH
40637	IWNYCVPKH
40638	YWAFPCCGF
40639	CWGPIMHPF
40640	TWKSPRGYM
40641	AWMCCARHD
40642	CWWNFGKDC
40643	SWCAMIYEM
40644	IVRFYFKSW
40645	LIRQGDCFI
40646	LTRDRFTEW
40647	LYRTSIALV
40648	MERYGHCHP
40649	MLRAESGLT
40650	NNRTYDWWN
40651	PVRMFTESG
40652	QGRQILKAI
40653	QHRYYTCQS
40654	QLRKHYTLL
40655	QYRPTWVHV
40656	RFRWVHFHW
40657	SGRCNALNT
40658	TIRDSVFGC
40659	VAREDEWQN
40660	VCRTQESSA
40661	VHRHRAHKF
40662	VPRWVERAH
40663	VVRSYNFGE
40664	NDRHVPCWD
40665	AFRNKEFYV
40666	AHRDYACTN
40667	CMRGDKRNR
40668	CMRMDGNDR
40669	CNRPKCPPM
40670	DARFGLNFG
40671	DARFYITEG
40672	DHRETDAEE
40673	DYRNYFIWP
40674	EARCTDMCY
40675	EFRPVLECC
40676	EIRYYYPES
40677	ELRWDARPH
40678	EMRITNARC
40679	ERRMEQAQT
40680	ESRAGYMHY
40681	EYRCVYRKC
40682	EYRLCPQTS
40683	FHREQAAHQ
40684	HSRQSEETI
40685	IAREPLMAN
40686	IDRKDRRES
40687	IIRNRCAEP
40688	IYRLMHWDN
40689	LMRSTVKNG
40690	LPRCVIAAC
40691	LTREDFTQM
40692	MLRAMDIDA
40693	MVRDKNNDH
40694	NRRMVVAQS
40695	NTREYPEYI
40696	PMRMAIEFD
40697	PSRPCTYHS
40698	QGRWMYHWW
40699	QLREIQMIQ
40700	SNRCLTYEF
40701	SQRDEMFAL
40702	SQRHIECIG
40703	STRDWMSCM
40704	TDRMLAGQR
40705	TFRMSVTGV
40706	TIRSFNLEA
40707	TMRTHSSWW
40708	TYRVFMVHD
40709	VFRMVMEQY
40710	VYRLREKQE
40711	IGRHQMTDK
40712	ANRADWQHL
40713	QCRHTDQVK
40714	ETRGFVKCH
40715	ETRVREDAV
40716	QDRRPENRF
40717	CERSKCPKS
40718	DIRSYKGFQ
40719	QVRDIANMC
40720	VKRLENEWH
40721	YGRLREAMH
40722	DDRVKMFHY
40723	ANRMIWYPY

[1964]

40724	GHRPKLCDQ
40725	VNRGDLSLF
40726	QSRCDPVSP
40727	EHRDNTTRP
40728	GVRNALNLS
40729	GIRCISKAG
40730	LTRSGPIGQ
40731	MTRVSMTWR
40732	CCRQIWVQY
40733	STRPIACHS
40734	NNRALMSVE
40735	TVRHQMDAY
40736	NNRRQFGYV
40737	SSRMQMQGP
40738	MFEAIVTGL
40739	MHGAFVDTQ
40740	MNVAYHGFS
40741	MTQADPCIG
40742	NAYAAGNIG
40743	NHSANPHGY
40744	NIMAIHHDS
40745	NMVAMNEPR
40746	NVWAWKEVE
40747	QLIAIHDIN
40748	QTVAHENCD
40749	QTVAPGGHC
40750	RADATGVPN
40751	REDAYLAGH
40752	RSEAGIRQI
40753	SGLAFDEDD
40754	SHVAYKLLD
40755	TMVAFSSCL
40756	TSYAMWSDV
40757	VASAMRHHT
40758	VTEAFPTSV
40759	VVDAIAQAD
40760	WGLANNCHE
40761	GIWAYGCTS
40762	AACAIWKGE
40763	CGDAYDEHC
40764	DDSAQYWQV
40765	DHTACKKLC
40766	DSKADHCVN
40767	DSMADDQHL
40768	EASAFPHDQ
40769	ECHAIDQSD
40770	ECKARNSCG
40771	ECTAHGVHC
40772	EFDARVAHE
40773	EFFALDSQS
40774	EHGAVHNAN
40775	EVAASSAIH
40776	EVIAGHVGA
40777	FPSAPTEPG
40778	GKEAAYAVW
40779	GMWAQWCSN
40780	GVEAYKFPH
40781	HEQAVQKHI
40782	HNCAGWDDG
40783	HQVATQLIK
40784	HTKAYNHCV
40785	HVHAVPNFA
40786	IADAQTMQP
40787	IFVAPMPGM
40788	IHVALDQSL
40789	IRVAKKGAR
40790	ITGAMMYMG
40791	ITLACVYHD
40792	IVMADVYGN
40793	IVNACMVEK
40794	LEIALHACW
40795	LFTAFRQFT
40796	LTSAVGAGM
40797	MISAFTGSH
40798	MISAPPDTW
40799	MTFAHDTTN
40800	NEFAIQNAE
40801	NKEARGALD
40802	NTTASTDVP
40803	PCKAKEKAV
40804	QQDAMEQYF
40805	RAAAQGPPQ
40806	RFFAAELQW
40807	RLEAMYNWG
40808	SAIATHQYI
40809	SEYAAMRDS
40810	SGCAMLNMG
40811	SKCAHSFEH
40812	SNWARSSAE
40813	SRFASWEGM
40814	SRMAIDVIG
40815	STSADPHCG
40816	TQMAHMIAH
40817	TTCAHSKMG
40818	TTEADDPTT
40819	VGEAPGHWQ
40820	VVGADRCQA
40821	WEEACKLDT
40822	WHSAVQHLL
40823	WVDAFCCQC
40824	YEQAMEYET
40825	YNNADCHYH
40826	YTQAMCITA
40827	TQHALSWWS
40828	NVLARCDDI
40829	VRAAQRSEP
40830	TQNACKVRT
40831	AAKALMSCH
40832	DFHACYKCS
40833	ICAAPNSAK
40834	LSMAMQVGC
40835	MNKAMTTWA
40836	QCNALNSFV
40837	RTCAYESGT
40838	TKNAMEIQE
40839	TPNAFVFAS
40840	TQEACHHDP
40841	CMEAQGECY
40842	QMQAWYCQE
40843	GTNAYMAHK
40844	AVKAPSQMS
40845	QSDAVHEPA
40846	QGVACLDQL
40847	NVYATPVVC
40848	TNAAYEMWF
40849	DQLALTFED
40850	RTEAHLPAP
40851	LDGYMNHDC
40852	LILPMGVNN
40853	LMSVMFAAH
40854	LQDHMCHNG
40855	LQEDMSPYA
40856	LVIKMHQLW
40857	LYHEMTHAF
40858	MCAVMDIFG
40859	MMQRMVYQC
40860	MNSLMQVAW
40861	MVQRMVYQC
40862	MVTTMGSTE
40863	NCCNMQLTF
40864	NCHVMSINL
40865	NGCGMVWQR
40866	NLAGMLPKN
40867	NMTGMPNGT
40868	QALKMENDQ
40869	QFYMMPFSL
40870	QGFFMMALM
40871	QKNHMYHQM
40872	QSDMMCHCM
40873	QTCSMLKDR
40874	RATEMAWSW
40875	SAGFMQEAA
40876	SHTRMGAFS
40877	SIAWMHLTQ
40878	SNMGMKHYE
40879	STDFMIHHF
40880	THDSMFFAC
40881	VAGNMDDHA
40882	VPYWMNTQQ
40883	VSVDMNCAS
40884	VTSLMQMGC
40885	YDTRMQTLG
40886	YRGWMNMED
40887	MSLWMATMQ
40888	AFHHMMNTS
40889	AHTIMMRDF
40890	AMDLMGDCE
40891	AMMQMMKAN
40892	ATAPMCFYQ
40893	CCCPMILFN
40894	CGEDMKICQ
40895	CMFGMVNYT
40896	CVATMHCHQ
40897	DELKMGHNI
40898	DGTPMHEYE
40899	DKMVMVFKR
40900	DKWTMQHWE
40901	DMVKMADCY
40902	DYEDMDQYY
40903	EGNIMMTQH
40904	ETSPMMTKP
40905	FQSQMPKMA
40906	FSQTMIFGA
40907	GMIGMARLA

[1965]

40908	GVDPMVTHQ
40909	HKDTMHSFS
40910	HKFWMPNQH
40911	HMCWMATDK
40912	HMYGMQKHE
40913	INPKMHYAS
40914	LATTMQAFN
40915	LCEVMGCDA
40916	LENTMWKCA
40917	LGAVMTCHY
40918	LHQTMACQL
40919	LNGTMLAMS
40920	NANDMSSDR
40921	NCEPMNEWP
40922	NEPPMGLKY
40923	NESPMNDTC
40924	NITEMGAYP
40925	NTGEMAEYV
40926	PNILMMFEP
40927	PVGWMVQWP
40928	QCNHMDHVA
40929	QDNMMVRDY
40930	QISTMLARC
40931	QMATMFTCR
40932	QPISMQGQM
40933	QQVRMDRLD
40934	RDEDMHIAE
40935	RHQHMKYLD
40936	RTVDMNLKG
40937	SGLGMPNWT
40938	SGSCMWGYA
40939	SMIPMDSMP
40940	TPNNMWWME
40941	TQVLMHATW
40942	VAQTMHYHH
40943	VKDEMHYMR
40944	VLDQMHAPP
40945	VVNSMWLMH
40946	WFHNMLATN
40947	WGMCMQRPI
40948	YRTEMNEAH
40949	YVSQMWMQC
40950	YHAPMLMCN
40951	DPVEMCSYA
40952	DRTTMVGWY
40953	LPWHMMCMP
40954	AFCRMKFGG
40955	GTEMMGDHA
40956	PCWPMMIPG
40957	GTIRMPEHR
40958	FCSPMFKQE
40959	AKAGMSYSG
40960	ACAIMKMWY
40961	AHFQMNMSP
40962	CESKMCQEP
40963	CLYNMWQCE
40964	DTLPMAMDP
40965	DYGSMAASW
40966	EDFSMPFNH
40967	ESSSMNEEK
40968	IEVGMTFPE
40969	IHQYMTAHG
40970	LDKQMWQFT
40971	MEAMMKHME
40972	PQGIMWWTL
40973	TMWMMHSSP
40974	VGLSMHSMD
40975	VKGRMCNNF
40976	GSTHMNQEG
40977	YDYWMQAPT
40978	HNIHMWCNI
40979	MCLNMAKQG
40980	QTNTMQMLS
40981	NNCNMNVMK
40982	VLTTMWKNN
40983	FHDRMAWAR
40984	AQWYMWCWD
40985	YTELMKKAC
40986	CISPMRVHQ
40987	YASTMHSIP
40988	CEFPMGPSD
40989	LLPIMHTHW
40990	FQHRMEEDA
40991	STALMDANM
40992	EFGEMLGVD
40993	VTSHMATAC
40994	PAYRMECCE
40995	GAQSMQFSC
40996	YVIPMVPSS
40997	AAFVMNGDG
40998	FTEPMCMQG
40999	LHDSRNKSP
41000	LRDFENLIL
41001	LTEDGNWWS
41002	LVITFNQGC
41003	MGLSFNHTR
41004	MSTERNRAC
41005	PQQHTNFSK
41006	QCTVQNQFG
41007	QMELSNQWN
41008	QTVDHNITL
41009	SAKYDNVQI
41010	SAYSTNGDA
41011	SCQDVNNYH
41012	SHGYGNRFA
41013	SHYQDNIGF
41014	SSKPSNEMW
41015	TKKEINDCC
41016	TMVMNNWVA
41017	TQHSNNQAT
41018	WHKLHNRHV
41019	AHAFNNVVD
41020	AKHWFNESG
41021	AVVQANQVL
41022	AYPHRNVGL
41023	CKQQFNGQN
41024	CMNEWNKLG
41025	CMVDANHLW
41026	CNFNNNDHA
41027	CTLFSNFQY
41028	DSCKANTQS
41029	EAWMHNFNW
41030	EIYITNCTS
41031	ERHGFNHCN
41032	ESDPRNLGP
41033	ETEERNGFG
41034	GFGHSNGFL
41035	GISTFNLWK
41036	HMTEKNMGE
41037	HYMMNNYLV
41038	ICAIVNTCN
41039	ICSGVNADN
41040	IQVRWNAEL
41041	IRYILNPEF
41042	ISDCFNCAD
41043	LCAHYNPHN
41044	LHGNYNYSV
41045	LNSHINGHM
41046	LSCCTNVFC
41047	LSCMRNMSL
41048	LSDYVNDEG
41049	LVMPYNSAQ
41050	LYEERNAQM
41051	MCNGYNVSF
41052	MHVFSNVCG
41053	MQYTINYQD
41054	NETSLNQHM
41055	NFVYWNYLP
41056	NGVNINQCF
41057	NHVNANLRA
41058	NKQYFNQKD
41059	NSHDYNCFH
41060	NSLHNNEDP
41061	NTGMDNGQG
41062	NVLPFNCNF
41063	QHEYLNQKS
41064	QSQRLNGQI
41065	RTEYANHMH
41066	SFYSTNFQL
41067	SSGWLNSEN
41068	STCIDNGRT
41069	SVWWGNSAA
41070	TPMWVNMVC
41071	TVGVINLVR
41072	TYELLNPGT
41073	TYHCQNHDH
41074	VHQWTNMQF
41075	VPQTYNVEA
41076	VQDLFNQFC
41077	VVHCNNTSW
41078	WDCLLNKWR
41079	YGMYENFHC
41080	YRCHNNEKT
41081	IEEGKNALY
41082	GSATPNETP
41083	YVSLCNDGN
41084	ADHLHNVWI
41085	CHCHYNPFC
41086	CVMPSNCQD
41087	CYDMWNMTT
41088	DIQYKNHMN
41089	DKLLSNSWR

[1966]

41090	DTESHNETP	41137	QSKWQDHDY	41184	MYGFYGHDM	41231	RCQLFDHDE
41091	FKMDTNFAP	41138	QSVWLEHTL	41185	NAYQNSHYP	41232	LGKVQTHIR
41092	ITYEGNISP	41139	QSVYDMHTP	41186	QLYGRRHWH	41233	DVIHTTHSG
41093	LHLWRNSGE	41140	QVDGQCHTT	41187	QMMDFQHPA	41234	TKGRNDHME
41094	MTTHSNCGP	41141	RDSRPGHQY	41188	QNYWGRHPA	41235	NSTRWYHTG
41095	PTKYHNDMG	41142	RQFKQTHFQ	41189	QSHYRMHSA	41236	DTDNKPHFA
41096	QLSKVNVYE	41143	RSPHVHHQL	41190	QVNHVRHWS	41237	QGYSHLHAL
41097	RHSDHNYIG	41144	SAHTDRHNS	41191	RHICFQHCK	41238	CQAGTDHFP
41098	SNQHNNQSM	41145	TSPCDMHFS	41192	RIEIGQHYV	41239	CASSKSHYL
41099	SPGTENRGG	41146	TYNPWQHFE	41193	RSEVSSHLD	41240	ALGNRHHQC
41100	SQMQVNPLH	41147	VHCQTDHDG	41194	SANWFTHDK	41241	VQYYYGHMQ
41101	TLTEVNRLA	41148	WAHYKAHYP	41195	SVSLKVHHE	41242	GVFYVGHWQ
41102	TQEHINIWN	41149	WQDFYIHYA	41196	SYVTTQHST	41243	LVQQDWRFW
41103	TSHFKNCIM	41150	WRYMWPHRQ	41197	TASGIDHPK	41244	MCAYLHMFH
41104	TTWEQNVMD	41151	YAVMYQHHA	41198	TQAEYSHKS	41245	MEKVKCSFA
41105	TVCFENLGG	41152	YCLQYAHKG	41199	TTWGTGHMA	41246	PKNWAPMFT
41106	YALLHNAMW	41153	AFWNPTHFD	41200	VFNLYSHSH	41247	QQDIYIDFF
41107	MSARINTPD	41154	AMIPNLHME	41201	VIKPFAHHG	41248	QTEHLWMFH
41108	ERNMFNICH	41155	AMLGYQHQP	41202	YASYKCHGD	41249	SGHWNDYFH
41109	GRQSINEIA	41156	AQDIYEHMQ	41203	YVQSAQHQQ	41250	STIPIWNFL
41110	STGCYNWPQ	41157	ASYHWTHAG	41204	YYASLDHYL	41251	TEQFSRLFA
41111	WGQSYNYWP	41158	AVVIEKHDG	41205	NFWNQVHCV	41252	TQDSKMGFG
41112	AVKSCNTCK	41159	CMKSYDHGS	41206	MATNKYHIQ	41253	TVAWATAFE
41113	QESMCNYWQ	41160	CNVWPCHQE	41207	QYYNYCHMN	41254	VKKLAKYFH
41114	DKCLNNKFY	41161	DADLAAHWP	41208	GCMCIRHAY	41255	VKVDKSQFD
41115	HMAYCNKVC	41162	DAQLGMHFI	41209	YIYFALHTH	41256	VVNIRGSFQ
41116	GAMLNNNAS	41163	DKANCYHWL	41210	WFDRSWHIL	41257	AFEVSESFE
41117	GKTQYNHHN	41164	DKIPIDHFA	41211	NIAMPAHIP	41258	AGEMDEAFG
41118	TYDTRNHCL	41165	DKIPIDHFV	41212	HDPLAGHWP	41259	AMELADDFA
41119	NGNHYNAMR	41166	DVHHLTHSP	41213	QRQGQLHYW	41260	AVDHCFKFE
41120	DYSQVNCEK	41167	EANKAFHLA	41214	IACISMHNH	41261	DFKKPSLFR
41121	GTMKHNGRA	41168	EHGDEMHPP	41215	AQEHLDHCL	41262	DMQTICRFP
41122	EKIRQNAAP	41169	EHMNCMHCE	41216	ECTMRMHSH	41263	GMGWQVSFN
41123	NLERCNTMM	41170	GAGMRMHDT	41217	EDSFSFHTI	41264	GPIPSEVFP
41124	SRYVNNGTV	41171	HKASWCHDQ	41218	EYCETMHTN	41265	HCYIKEGFQ
41125	LRHEGYHSE	41172	HMMLRDHNN	41219	FMPYCVHSG	41266	HNMPSGQFT
41126	MEYWPIHCQ	41173	ICSIYGHHQ	41220	HSKNTVHKR	41267	HYESRLDFY
41127	MFLPTDHQF	41174	IDSQIYHML	41221	IMGIHLHPA	41268	ICKSDPVFD
41128	MQHEETHDY	41175	LCAEFPHGR	41222	IRNIAKHAL	41269	LEPCTDEFG
41129	MREQKSHGY	41176	LITQPVHQQ	41223	LDKRQDHNF	41270	LIMRISPFD
41130	MTHGHSHSF	41177	LQAGGEHNW	41224	MQAEACHRW	41271	LSGSIMEFN
41131	NLDMVTHAE	41178	LTLESPHPA	41225	NTHLFHHYL	41272	MVSENSTFW
41132	NTQGTVHIQ	41179	LYKWIVHNS	41226	QHTSAFHTM	41273	MYSRTIPFH
41133	NVDNHDHFE	41180	LYYITEHST	41227	WHKCLSHQD	41274	NKHEVGCFG
41134	PQPSEYHWN	41181	MAVSQSHQF	41228	WRKCLSHQD	41275	NKYEVGCFG
41135	QEHQYEHQS	41182	MAVWVKHQH	41229	NAYYDSHCE	41276	NMHWRYQFE
41136	QGPECTHNP	41183	MSSMLEHSW	41230	HDPPPMHWL	41277	PYPTHSMFS

[1967]

41278	QDSHVHSFF	41318	LQELSSQYA	41358	EQAMWMYPA	41398	TKSSIEYSA
41279	QGQYWHYFK	41319	LTVNLMKIA	41359	EQGHCEVAA	41399	TNLRQHTNA
41280	SCGHQTRFG	41320	NAIDCMSGA	41360	ESKWYYVMA	41400	TVSDGDWIA
41281	SCGHQTWFG	41321	NAWIDWSLA	41361	EVERDKVGA	41401	TVVLTLKEA
41282	SDHYCGTFY	41322	NQGTSAWGA	41362	FDEKRSESA	41402	VTSWFHEIA
41283	SHMMSHAFW	41323	NSLPFGMVA	41363	FPANYHSNA	41403	VTTVSGGCA
41284	SNSLYRAFL	41324	PFNKSKGLA	41364	FVSTQMPIA	41404	VVMCKSRDA
41285	SSVCTQIFT	41325	PNQVGTQGA	41365	GPNDWMFDA	41405	WVMHYYQYA
41286	TGEMDEAFG	41326	QDQQVHEGA	41366	GYWLWMYVA	41406	YSSPDLRPA
41287	TNDIWWQFS	41327	QLMIQGSYA	41367	HCDDVLWCA	41407	TVLVPVQAA
41288	TRTELANFF	41328	QSAYLIMVA	41368	HSAYFGVGA	41408	EVGIKYRAA
41289	VCWLWFYFG	41329	QTTFAERHA	41369	IEELGKPRA	41409	EVNWQCQYA
41290	VGICHPEFW	41330	SDHGRIMTA	41370	INGMKSAMA	41410	GSMFNPGAA
41291	WVTNQSDFA	41331	SLAHWDGRA	41371	INVPDLLGA	41411	MTNWFPACA
41292	YPQCAQWFW	41332	SSGWTQQMA	41372	ITSIRMYQA	41412	MVKTSCKPA
41293	HQKMTYAFT	41333	TFIMLGADA	41373	IVQPYHVEA	41413	MVKTVYPDA
41294	VRWGHTCFY	41334	VDYYSPYLA	41374	LHQPRMDCA	41414	PPEEQCGMA
41295	NSAECREFQ	41335	VGCIAQAAA	41375	LIYHTPVMA	41415	QFTETWPCA
41296	WKNNQKWFK	41336	VNDCNWCEA	41376	LSSCRGLIA	41416	RIIWPDVVA
41297	DIGSYKGFQ	41337	VRTPWQTVA	41377	LTIRDYGEA	41417	SSDTPANQA
41298	DSAHYCTFT	41338	WFNVKKTAA	41378	LVNVRYMHA	41418	TFVECIQYA
41299	EKSTILPFA	41339	YAPQQHMWA	41379	MHEPILKYA	41419	TGHFFTTAA
41300	EKSTILPFT	41340	YSQKQCPTA	41380	MICWWKQPA	41420	VGYGCKGYA
41301	EVEYGCWFE	41341	ACYQQCVSA	41381	MMNHSQLYA	41421	VSWSRMEQA
41302	GNMRQSRFP	41342	AKSSIEYSA	41382	MQHPYMQMA	41422	YRLHCCNRA
41303	HGLHEDSFY	41343	ALWDKFEAA	41383	MSTVSGSYA	41423	MRVQVFSDA
41304	TNNYPMPFD	41344	CDYHTLSPA	41384	NGFMAAVMA	41424	RTESWSSLA
41305	DMPPLHSFC	41345	CFENLQVMA	41385	NHADYMRNA	41425	QKWMQMDGA
41306	ASTMKSIFE	41346	CRDWDGFNA	41386	NMWPTFRAA	41426	TAAGYIVDA
41307	SLASEVQFQ	41347	CTTDTPQNA	41387	NTDEESGMA	41427	TQQTADRWA
41308	DLTTDTIFQ	41348	CYHECQGCA	41388	NTMMLQSAA	41428	FQIRQDNLA
41309	SADRTAMFH	41349	DIFFKCEWA	41389	QAPLKRAAA	41429	NTGMEGQMA
41310	PTYPYEWFQ	41350	DVGGRAVQA	41390	QCNRCDQYA	41430	IMTMQMDTA
41311	NMKPKDVFR	41351	EATDLCLNA	41391	QFIQYGTCA	41431	EYSSQHPYA
41312	SATIRFYFN	41352	EATDLCPNA	41392	QIGLKMTPA	41432	QPFCFSIGA
41313	FVPWCHSFG	41353	EDNHYQMIA	41393	QLNPWFYYA	41433	MGVPPLMQA
41314	ITEQYDAFA	41354	EFYNYEIMA	41394	REISVEAVA	41434	VFGSTRSAA
41315	NTMCQLNFC	41355	EGGSNMMNA	41395	SAKITWASA	41435	ATNYTMEPA
41316	LYQKDQRFS	41356	EKAMLLNGA	41396	SIVEFESAA	41436	HIHMSDRGA
41317	LFGTHPGQA	41357	EKTCVTREA	41397	STHMATQLA	41437	GSHDYYQSA

[1968] 实施例16

[1969] 在淋巴结中具有组织嗜性的AAV5变体

[1970] 此实施例描述了使用实施例1-实施例2中描述的方法和系统发现的在淋巴结中具有组织嗜性的工程化AAV5变体。

[1971] 分析CNS、肝脏、骨骼肌、心脏、肺、脾脏、淋巴结、骨髓、乳腺、皮肤、肾上腺、甲状腺、结肠、坐骨神经和脊髓组织以鉴定AAV5 VP1 581-589区域中驱动淋巴结嗜性的氨基酸残基。结果显示于图18G中，该图显示了相较于所有其他分析组织在淋巴结组织中AAV5 VP1的581至589区域中每个氨基酸的位置频率，其中衣壳多肽具有SEQ ID NO:2的序列。位置

("Xaan",n＝1-9)在x轴上表示并且氨基酸残基在y轴上表示。特定位置的特定氨基酸残基的频率在热图中作为强度表示。较深的方块表示特定氨基酸在特定位置的频率较高,较浅的方块表示特定氨基酸在特定位置的频率较低。

[1972] 分析相对于给定氨基酸残基出现在所有其他所收获组织中鉴定的变体中的指定位置的频率而言该给定氨基酸残基出现在淋巴结中鉴定的变体中对应于SEQ ID NO:1(推及于SEQ ID NO:2)的位置581至位置589的所述指定位置的频率,以确定优先靶向淋巴结组织的衣壳的一组序列规则。参考下文表31,和SEQ ID NO:2(AAV5 VP1),以下氨基酸由此可以在Xaa1-Xaa9任何一个或多个位置处以任何组合独立地突变,以提供与野生型AAV5 VP1衣壳(SEQ ID NO:1)相比具有增强的淋巴结嗜性的VP1衣壳,其中淋巴结嗜性在此是指优选淋巴结转导的特性,而不是优选转导所有其他所收获组织的特性。

[1973]

表31 淋巴结嗜性规则
Xaa1选自A、D、E、Q、S或T
Xaa1选自D、E或T
Xaa1为E
Xaa2选自A、H、I、S、T或V
Xaa2选自I、T或V
Xaa2为V
Xaa3选自A、E、H、I、T或V
Xaa3选自A、I、T或V
Xaa3为T
Xaa4选自A、D、E或P
Xaa4选自D或E

[1974]

Xaa4 为 E
Xaa5 选自 I、L、M、V 或 Y
Xaa5 选自 I、L、V 或 Y
Xaa5 为 L
Xaa6 选自 D、E、I、N 或 Q
Xaa6 选自 D、E 或 I
Xaa6 为 D
Xaa7 选自 A、E、G、Q 或 V
Xaa7 为 A、Q 或 V
Xaa7 为 V
Xaa8 选自 F、G、M 或 W
Xaa8 选自 F 或 W
Xaa8 为 W
Xaa9 选自 I、P、T 或 Y
Xaa9 为 I 或 P

[1975] 下文表32提供了在淋巴结组织中发现并且符合表31中提供的一种或多种规则的变体AAV VP1衣壳多肽的581至589区域的序列。因此，本公开提供了由具有SEQ ID NO:2的工程化AAV5 VP1衣壳多肽组成的rAAV，其中Xaa1至Xaa9区域是SEQ ID NO:19118-SEQ ID NO:20中的任一者，如表32中所公开。本文还涵盖由工程化AAV5 VP2衣壳多肽和工程化AAV5 VP3衣壳多肽组成的rAAV，所述多肽在对应于AAV5 VP1 581至589区域中的氨基酸的AAV5 VP2(氨基酸残基445至453)和AAV5 VP3(氨基酸残基389-397)中的区域处具有下文表中公开的序列。

[1976] 表32

[1977] 驱动淋巴结组织嗜性的AAV5 VP1衣壳多肽中581至589区域的序列3

[1978]

SEQ ID NO	581-589 序列
19118	ETTMWNSWA
19119	ENTMWQSSQ
19120	ERQAHNHYA
19121	EAKSRPQYA
19122	ECQQDDTDI
19123	EDICHWHHT
19124	EECIMQRSK
19125	EEERTCKHC
19126	EEFARIPGH
19127	EEPRCVNPC
19128	EFFCQADKA
19129	EGAGLYEFA
19130	EGPELTCDK
19131	EGTLRYTMS
19132	EHDTQDISV
19133	EHHPLGVMG
19134	EHHYCVPWS
19135	EHILTAERI
19136	EIAAKIFCR
19137	EIILIWKQC
19138	EIMSYPHTN
19139	EIQHVSHEH
19140	EKHGCDKAE
19141	ELNHKWHSV
19142	EMALVVSNQ
19143	EMHQFMECQ
19144	EMRFTNFCY
19145	ENIHNNDCT
19146	ENKEKEGQW
19147	ENTMAIIEH
19148	EPILYHEHK
19149	EPIQIFNDA
19150	EQHRDTMRP
19151	EQNACDHYG
19152	EQVFGFWNY
19153	ERMETFDCP
19154	ESCEDEGIL
19155	ESKASGANG
19156	ESMLPIMGR
19157	ESVPLEYAN
19158	ETCCFTNMG
19159	ETCGQGGYT
19160	ETDHCESWY
19161	ETPIEWDEQ
19162	EVEVHSTAL
19163	EVHIAPSAC
19164	EVKRITEQS
19165	EVKRWMMVD
19166	EVMAMVDVA
19167	EVPQNELNV
19168	EVRSIEYWN
19169	EVWHTKILI
19170	EWAFNYNQN
19171	EYKMFEKYE
19172	EDHWHGPGD
19173	FVALKAVID
19174	FVAPVDSFY
19175	FVFEGPMHQ
19176	FVKVRIWTQ
19177	FVQPTFIRH
19178	FVWYGPRNH
19179	GVACVQETD
19180	GVANFMLNN
19181	GVCEENEWM
19182	GVCQMCCWP
19183	GVDARHAQR
19184	GVEFCLTNG
19185	GVEWPPFMQ
19186	GVFMTWQWL
19187	GVHGINVMM
19188	GVHIMQCQF
19189	GVHIYACHY
19190	GVHQVPTWN
19191	GVLQMDWQE
19192	GVMQHWTCD
19193	GVMWDGVAP
19194	GVQMTVRQA
19195	GVTYAQRIC
19196	GVVFCDQKD
19197	GVVLFQGGA
19198	GVVNSMQTK
19199	GVVQQGAMI
19200	HVFVMDGPI
19201	HVIWDSMKL
19202	HVMDHVATS
19203	IVDFENRHP
19204	IVDPIFRQT
19205	IVFEGGIAS
19206	IVGAPCLVT
19207	IVGDCQFHV
19208	IVGSNTYGY
19209	IVHNVEYSN
19210	IVHSQPNSV
19211	IVLQLQQAK
19212	IVNALEFCK
19213	IVNNAQMHE
19214	IVQMHYINQ
19215	IVQVESCQN

[1979]

19216	IVSARDTHN	19257	MVQPMPQHH	19298	QVMNGSPKT	19339	TVATHCFIH
19217	KVDGMDIEG	19258	MVSFNYEIG	19299	QVQIGFEQG	19340	TVAVWWRNA
19218	KVDPPGIFW	19259	MVTKEYMCC	19300	QVVTWGYEY	19341	TVDAIQGWQ
19219	KVDQGLCQP	19260	MVTMGYWSS	19301	QVYFWGTHS	19342	TVDMRHFFT
19220	KVDWDGNID	19261	MVTPYWNQY	19302	RVCDKNLDL	19343	TVDSEHRLC
19221	KVEPHHPAD	19262	MVYPKRGEN	19303	RVDEKLDHD	19344	TVEAMHLGP
19222	KVNPKEFQD	19263	MVYSGQQEF	19304	RVDTNMRDR	19345	TVEDGQWSW
19223	KVPICAHKQ	19264	NVAFAHCGG	19305	RVEHWVNMS	19346	TVEFTYGWC
19224	KVQICDFWC	19265	NVDDHPTMC	19306	RVNEWGAYG	19347	TVEPTNYDG
19225	LVCLCHSCA	19266	NVDMRNVES	19307	RVPLYPYWT	19348	TVFHSTGVG
19226	LVDQYMWDN	19267	NVDTPSGNQ	19308	RVSPWDQWF	19349	TVGICFCPV
19227	LVDRVKHVD	19268	NVDWDSECS	19309	RVVMPLGSL	19350	TVGLLEMTR
19228	LVDYWAWNG	19269	NVECCLNRP	19310	SVAYKWDAF	19351	TVGTGCDTC
19229	LVEQERMGV	19270	NVEEKWEWW	19311	SVDACDYDD	19352	TVGWYQHDW
19230	LVFAKLDQN	19271	NVEFRYFYN	19312	SVDDCMHSA	19353	TVHNCLTMG
19231	LVGICDLCP	19272	NVESFPHEQ	19313	SVDNDRAWP	19354	TVLSTTQKD
19232	LVGSTSHVT	19273	NVEWWGHCA	19314	SVEAKSATC	19355	TVMHHDGRL
19233	LVMHQQRNY	19274	NVFATNCWV	19315	SVEFHEDFS	19356	TVMSWQQHE
19234	LVYQHDTAP	19275	NVGAHFAPF	19316	SVEGANAEV	19357	TVMWQTQIG
19235	LVYTELSGA	19276	NVGGPCYNL	19317	SVEGVIIDR	19358	TVNIILQGS
19236	MVATESHDK	19277	NVGGQNMEH	19318	SVEYYSMPQ	19359	TVQMYQDLE
19237	MVDDPGQQF	19278	NVGIGHGGN	19319	SVFQTHCDW	19360	TVSESRKHN
19238	MVDHWGDWN	19279	NVGIHFAQQ	19320	SVGPLQYWK	19361	TVTDEDDCN
19239	MVDHYICMY	19280	NVGNEHIAY	19321	SVGWGWMMG	19362	TVTIAVQNM
19240	MVDSFGKYL	19281	NVHIALEQH	19322	SVHEWDPCT	19363	TVVQWHAMI
19241	MVECEGEQY	19282	NVIQHAIGG	19323	SVHGKVSWF	19364	TVWTEVSPP
19242	MVEDSTCHW	19283	NVKFSIEHT	19324	SVHTIWPMA	19365	VVDPRPDWF
19243	MVEEANVPC	19284	NVLGTLVHV	19325	SVICWTHMA	19366	VVDRWDHWM
19244	MVEEGSWNI	19285	NVMDCESCC	19326	SVIYNEHVT	19367	VVDSNFCSM
19245	MVEGRQDLT	19286	NVSYEQNFM	19327	SVMKQELSV	19368	VVGCDKSDA
19246	MVENCDLPP	19287	PVATALRMP	19328	SVNMYMLPD	19369	VVHEAWCDT
19247	MVEWGGSNQ	19288	PVNLNCTVS	19329	SVSMASAHG	19370	VVHNFLDQS
19248	MVGCFYTWT	19289	PVSLCRFGG	19330	SVTACPNWN	19371	VVKEFTGHY
19249	MVGYQYPAA	19290	PVTGENHMN	19331	SVTASYGFS	19372	VVKICGCCR
19250	MVHEAELYA	19291	QVALRGATC	19332	SVVEDAMRG	19373	VVMLQSGMT
19251	MVKEDITEH	19292	QVDDCGHLF	19333	SVVPTNNDM	19374	VVMREWRCF
19252	MVKEQWQCT	19293	QVDMELPCW	19334	SVYAPTKWG	19375	VVNQGVACT
19253	MVLPANGFQ	19294	QVETTMIGH	19335	SVYWPTTEF	19376	VVQHIWWGI
19254	MVNALTDYG	19295	QVIQNVWER	19336	TVADMFFDG	19377	VVRWPGSEW
19255	MVNPWVVFY	19296	QVIQRRNVM	19337	TVAHVADCY	19378	VVSRNEVGA
19256	MVNQAWQTD	19297	QVMIECRQS	19338	TVAPCNQGL	19379	VVTFTNILQ

[1980]

19380	VVVCRPVAP	19421	PVKDCILWS	19462	LHTRTQCEP	19503	QYTHGTCFS
19381	VVYEKQDHM	19422	QVCSDWCEC	19463	LNTGYMWNQ	19504	RYTGTWQDC
19382	VVYPLAPYR	19423	QVDTRCYGG	19464	LPTSVNMMH	19505	SATYTEMQS
19383	VVYSNICVE	19424	QVMMDDSNV	19465	LQTAIPFWP	19506	SITHNDHGW
19384	WVASQGAQC	19425	QVSCNVEPK	19466	LSTTIDKDP	19507	SKTCEIGKT
19385	YVANVPVLW	19426	FPTQMKMCN	19467	LTTAVAIRE	19508	SNTAYFNGS
19386	YVLGRHCVF	19427	FPTWDKPPT	19468	LTTQLFLQE	19509	SSTAEPNGR
19387	YVQIKGLQL	19428	FTTCFSIMV	19469	LTTQYICGP	19510	STTAIDCMR
19388	YVTALATYD	19429	FTTFWVPQQ	19470	LTTWNTYVN	19511	SWTACTPCH
19389	SVTNGSTKH	19430	FTTLHNMEN	19471	LWTQISEFA	19512	SYTYRLTDQ
19390	LVCIRAGVE	19431	FYTCELACV	19472	MATEGACNP	19513	THTEFDVFN
19391	KVNCRMIIM	19432	GITTIWPDH	19473	MATMFTNWC	19514	TKTAFGDMG
19392	GVWMQGFHP	19433	GLTADLWHY	19474	MITCIDGIF	19515	TNTAFGCYG
19393	AVHTADVCC	19434	GMTYVSGHH	19475	MPTLCECGL	19516	TSTEYKMGL
19394	DVKSWAVCD	19435	GNTLIVTGP	19476	MSTMSMDSY	19517	TTTMLHRIN
19395	SVLHVQALA	19436	GSTTQHDGT	19477	MSTTELVAE	19518	TTTTQDSQR
19396	AVSPHWGNY	19437	GWTPHTLLR	19478	MWTGQLRQC	19519	TTTYQALNA
19397	CVVKRMAEF	19438	GYTANFYTN	19479	MYTQWSKAP	19520	TYTCNKWFR
19398	CVVNMEYCL	19439	HGTTIASGP	19480	NATNSQNMY	19521	VATKFETNT
19399	CVWSRHQIC	19440	HKTIDDIFT	19481	NATPYKAEY	19522	VATMAEDNP
19400	DVRLIGHEA	19441	HNTEPPGMP	19482	NFTVTNLQM	19523	VATTDGPVY
19401	DVVRQWQAP	19442	HPTNMSACE	19483	NITLGAACE	19524	VCTNLAHGA
19402	FVEEDLYEE	19443	HSTGSGTFH	19484	NNTNQCQNC	19525	VDTIFTDLH
19403	FVHMWSRLC	19444	HSTPFVFNS	19485	NQTEMSSNG	19526	VGTPLDTSV
19404	FVRPLEEFY	19445	IATNQSAIL	19486	NRTCNAGEW	19527	VITVFMEQN
19405	HVFQGDYIA	19446	IGTAYPQWR	19487	NRTGPAAWA	19528	VKTPGATDF
19406	HVGNLTAFN	19447	IGTIHPQAQ	19488	NSTAFFRDQ	19529	VPTNTFDTH
19407	IVANAQIPG	19448	IGTVTVGAY	19489	NSTIQGGAW	19530	VSTAVMPEG
19408	IVGYQNCQS	19449	IITMMCRSD	19490	NTTEISRLP	19531	VSTGLQDAC
19409	IVLGQMTCT	19450	IMTSVDSQR	19491	NYTIRVQDL	19532	VSTTGPIDG
19410	IVSEWEDWQ	19451	IRTRELDQY	19492	NYTNCEFGN	19533	VTTEMPMME
19411	IVSNSVCRV	19452	ISTEKNESY	19493	PATTADMHY	19534	VTTGYQCAS
19412	KVCVVAGPM	19453	ISTLFSSGD	19494	PPTHQMWYC	19535	VTTTTKVSN
19413	KVEVKRDYW	19454	ITTKYDCMG	19495	QATSAMIPS	19536	VWTQEMWDA
19414	KVVDRETYS	19455	ITTPFYKVG	19496	QCTCEGEAQ	19537	WDTVPKIRD
19415	KVYEIGEAP	19456	IWTRADLGR	19497	QDTELYCCM	19538	WFTGPWAMH
19416	MVEWEEEWD	19457	KMTVRQSDS	19498	QITQMGRQE	19539	WRTQCQSEK
19417	MVILRGMSF	19458	KTTDDTVGE	19499	QLTWQAHSW	19540	WYTYNMCER
19418	MVLPTKLWD	19459	KWTQFGWMD	19500	QMTPGWWKT	19541	YATFQDCNA
19419	MVSVVQEPN	19460	KWTRIGDSA	19501	QNTCLPVFY	19542	YATNQCYNG
19420	NVLSEVYGH	19461	LGTQEQGMP	19502	QTTSAMING	19543	YCTDQSETS

[1981]

19544	YITKGNLAR	19585	KAMESHDCC	19626	NHGEYFDWF	19667	TKMEYQGAT
19545	YMTFKGYHG	19586	KFEEVHASG	19627	NHYEMQDWE	19668	TMQEFVVKG
19546	YPTKDPFKN	19587	KHPEQYRKY	19628	NKLENCIGM	19669	TMYETLLMG
19547	YSTQCNTMK	19588	KHVEALKAY	19629	NNHEQEWRN	19670	TQNEHGKYD
19548	YTTDDGMSP	19589	KKYEECPLA	19630	NSREAEMLR	19671	TREEDDKVN
19549	YTTHRDTVT	19590	KMVEDEPSA	19631	NTHEFKFQF	19672	TRMEQPSGR
19550	YTTVDNSLQ	19591	KNEEDYALY	19632	NTHEFMMQD	19673	TSNEMKKHW
19551	KQTVFVDQE	19592	KPVEIDMEG	19633	NYDEKHVCH	19674	TTEEGQTFR
19552	RITQGESGM	19593	KQYELNNWC	19634	PFFERNIRE	19675	TWGENGAWC
19553	ISTKWDYIA	19594	KSSECGWAC	19635	PKFEFYEMI	19676	VCEEYYQMR
19554	QQTKNNSTD	19595	KTIEQTIPH	19636	PMEEMMHFC	19677	VFEEDAYGL
19555	WCTQFDEYT	19596	KTYERYQCA	19637	PTNELYDNE	19678	VFLEGMHFP
19556	KITVHDRIP	19597	LAKELCCIQ	19638	PYPEASVKM	19679	VKNEQTVFF
19557	DMTLQSVSS	19598	LANEKDSMC	19639	PYREGFRSP	19680	VLGELGDMG
19558	PLTIEVNCT	19599	LFREKMFNP	19640	QALECTCQR	19681	VSPECPSFT
19559	QCTAHNTCM	19600	LIVEHDCWF	19641	QARECEAMD	19682	WGAEKQWAF
19560	FNLEFSPTL	19601	LKCEYPPMS	19642	QASEHDTVV	19683	WGHEIMTEW
19561	FPQESNHDG	19602	LLDEVRQAG	19643	QHNEKPMDA	19684	WHDEKSRYL
19562	GAVEIDCLY	19603	LLWECEKGV	19644	QHYEFQKHC	19685	WIMEENAGQ
19563	GHAEQWFES	19604	LTLEQSVPM	19645	QIDEGMYSE	19686	WMIEMWNCM
19564	GHGEFMRDA	19605	LWYEVMIWY	19646	QINEFAEHF	19687	YADECAYGR
19565	GHVEVCTLS	19606	LYKENWQFP	19647	QMREIQAEQ	19688	YAGEHDMGQ
19566	GIGEEMFWD	19607	LYQENQHEW	19648	QNPEVHFSK	19689	YANEPQHHF
19567	GSPEVSVAS	19608	MFAEYSHIG	19649	QSEECMECR	19690	YMVEYAKYG
19568	GWSEAHKSD	19609	MFEEKNLHT	19650	QSNEDLADI	19691	YNEEEWFFA
19569	GYYEVVKMY	19610	MHKECEACR	19651	RAMEMQDQG	19692	YQVERGASA
19570	HGYEIQFFQ	19611	MHMEMNYSE	19652	RCNEYTGEK	19693	MGKEYSQCF
19571	HHDEKLKML	19612	MIAEFVSNY	19653	RIIEQDEIT	19694	TYEEENGFL
19572	HMDELIYCG	19613	MNEEQEHGA	19654	RINEQDECM	19695	NTCEGWSHF
19573	HNKEKWPTE	19614	MPDEQQQMW	19655	RWMELENAT	19696	GLFEGSPGA
19574	IACESDESW	19615	MQAELTMSA	19656	SAGEWGQMW	19697	MWHELIADQ
19575	IHYEHGMQC	19616	MRHEEQRLK	19657	SIDEFQENW	19698	NQREWHGLA
19576	ILLEFQAPM	19617	MRHEIVAMY	19658	SPDEFGTFR	19699	GDYQLLNSP
19577	IQQERNRYH	19618	MSAEVLLMT	19659	SPHEIFVWN	19700	GFIDLNSRR
19578	ISAEHIFAG	19619	MSCELTGMH	19660	SSEEWCQCE	19701	GIDWLDQYG
19579	ISFEWLELN	19620	MYVEYTVGM	19661	SSVEGIRER	19702	GIMLLPNGH
19580	ISSEAIFYP	19621	NAVEILLCG	19662	SWAENKQSF	19703	GSLTLLEQT
19581	ITMEILGQN	19622	NCNEPKYAA	19663	TAEEIWYER	19704	GYVSLWSMY
19582	KACESIQDL	19623	NEYECMCWY	19664	TCCEPSWND	19705	HCVALMMDR
19583	KAEEIENMC	19624	NGDEFTVHG	19665	TGNELYCWT	19706	HEEVLIAGS
19584	KAIESMLFM	19625	NHAEFLSGC	19666	THAEFEHMH	19707	HGLTLKWDC

[1982]

19708	HHESLGMHM	19749	RCLDLVNEP	19790	GTECIDWTA	19831	MSHTTDGDE
19709	HMEALMTNV	19750	SADQLHQFA	19791	GTNAKDGTP	19832	MTGYWDVEY
19710	HTGFLTKGM	19751	SAEDLYYHY	19792	GWDRHDGFV	19833	MTIPIDEMA
19711	HTICLAVTP	19752	SCVCLEAHQ	19793	HDKDFDWAN	19834	MTVHRDAWY
19712	HTWQLHNWN	19753	SDWPLMRYQ	19794	HHDTSDMSC	19835	MTYKWDCHV
19713	IDLNLNIQI	19754	SEVGLPCYQ	19795	HHGPMDWGA	19836	MWNIIDRHN
19714	IHDTLGAVA	19755	SFNQLICGL	19796	HIMCYDEWY	19837	NAPHADTPS
19715	KCEVLEVQD	19756	SHDDLRMQC	19797	HLYIKDMTD	19838	NCRHCDPNC
19716	KEVYLSTAA	19757	SPGILQMIG	19798	HSMVHDCFN	19839	NCSYHDASP
19717	KHVDLSNIP	19758	SREILDVVA	19799	HSWTRDRLE	19840	NDGKDDTYQ
19718	LGNQLAYGG	19759	SSQHLWHRD	19800	HWVMGDVKL	19841	NEDMWDYNH
19719	LHYPLVNDG	19760	STVTLHTHQ	19801	HYDQRDEHQ	19842	NFRTFDHSE
19720	LMDHLCTPN	19761	SYDGLDQVA	19802	HYHQSDWMH	19843	NHLTRDQWG
19721	LRGCLHLEH	19762	SYEILGDWC	19803	IAIPEDKQY	19844	NMLNADAMY
19722	LTNFLQNDN	19763	TCLNLSWYG	19804	IASLEDMNG	19845	NNSVCDKMS
19723	MFLPLHYKD	19764	TFGFLSDPQ	19805	IEMVRDCGY	19846	NQVGRDVMD
19724	MHELLQSVH	19765	THEHLANHN	19806	IHLKSDYVW	19847	NSDTGDMNP
19725	MIAQLHTWA	19766	TMEPLGWSF	19807	IKDWVDGIQ	19848	NTMHTDTME
19726	MNRWLLTDG	19767	TSEVLDHFC	19808	IQQTIDMVA	19849	NTMLTDTGS
19727	MPQSLRVCS	19768	TYMQLSKYC	19809	ITHMIDQDN	19850	QCMHDDWSS
19728	MYEWLAHCW	19769	VFHTLHTKG	19810	ITWPEDGSI	19851	QCMWPDIAG
19729	NANLLTEQY	19770	VGMPLRKEE	19811	KDYASDIHK	19852	QEKLQDMYG
19730	NEMRLECSY	19771	VGNCLHNYW	19812	KEFRMDSGS	19853	QFMDFDFHY
19731	NHKWLKCSL	19772	VIYQLGEMV	19813	KHHTSDTAV	19854	QLMPRDGLC
19732	NHVSLNCDP	19773	VREFLNGLN	19814	KIYDHDAWT	19855	QQEHADWWR
19733	NISWLHSFH	19774	WCWQLHDQF	19815	KSLCYDDWM	19856	QSENMDYWD
19734	NTCYLYAFF	19775	WDDLLSFWC	19816	KWVNVDDMR	19857	QTFDADGMH
19735	PDPVLSIKD	19776	WTLMLHNFM	19817	LADSDDPMD	19858	QTQPHDRHH
19736	PHCTLSFIG	19777	NMPSLRTRI	19818	LIASRDEAH	19859	QYHHIDYGK
19737	PRKSLWMEP	19778	ASFFLNPLE	19819	LINMWDFWP	19860	QYQNDDRAS
19738	PWAVLNQEP	19779	ATMNLSWGL	19820	LQHFIDEGA	19861	QYYRADGAK
19739	PYEQLGTTI	19780	FQDPKDTVG	19821	LQKVCDVYT	19862	RQVMTDRIN
19740	QFGSLVNSW	19781	FTNSDDDVS	19822	LTAKPDCCN	19863	RSEFADHWW
19741	QGYKLDNQM	19782	GCEACDFMG	19823	LTVVNDTVF	19864	SAEYWDMQD
19742	QKEYLMNSG	19783	GDNKIDMHN	19824	LWLARDLMC	19865	SAIWGDNSM
19743	QMWTLPSHN	19784	GHMWTDSMA	19825	MASFTDHPQ	19866	SDETRDKHD
19744	QPMVLNQDC	19785	GIHVTDLQE	19826	MEVAQDVIN	19867	SFSFCDQRT
19745	QQHRLTFMP	19786	GMQWPDSME	19827	MHCAVDAQC	19868	SIEDFDEDK
19746	QSQSLTEEA	19787	GNDIWDFMA	19828	MHRHMDHEG	19869	SMEPYDTFF
19747	QTNALLNGE	19788	GQNVMDYHA	19829	MNCQQDCVR	19870	SNRWQDTYY
19748	QWNGLNSCL	19789	GSSVDDGVS	19830	MQHVRDQFN	19871	SPEIWDMGA

[1983]

19872	SSSQCDVYE	19913	YSEMGDFYG	19954	NPQYGCVME	19995	MFSHGHVMF
19873	STLQSDPFE	19914	YSSFTDQNV	19955	NYHDYQVCY	19996	DCVLTMVDK
19874	STMFKDHFA	19915	YTFYKDCCC	19956	PGIFGCVGW	19997	QWFTKGVGE
19875	TARNHDHTA	19916	DNILRDVMA	19957	PSNNKCVWC	19998	FPGVFKNWV
19876	TCHDVDFTM	19917	QISVYDTKY	19958	QKHCQCVSV	19999	GASWYCSWM
19877	TGGLWDWFE	19918	STALMDANM	19959	QTGTGWVGN	20000	GFQWDAHWN
19878	THVGIDNIP	19919	MSGPNDTEF	19960	QWGTTNVIT	20001	GFVSYHSWD
19879	TIRYFDGCG	19920	GIWVRDTYN	19961	RADNTNVSW	20002	GIAVRSAWF
19880	TLHDNDSAC	19921	DHLSNDAEM	19962	RINHYHVWV	20003	GISDDAYWW
19881	TMEVSDQHN	19922	ATGLTDHQW	19963	RSCGEVVFT	20004	GMMLMPSWC
19882	TMGIYDMNE	19923	QSRCVDNTV	19964	RWDGQEVWK	20005	GNAACNAWP
19883	TMRTSDVYE	19924	GCAVIEVSA	19965	SGYYPWVLT	20006	GTPRDEQWY
19884	TNYQVDWDS	19925	GKWYNGVEL	19966	SHNFKQVMR	20007	GYNTANSWQ
19885	TQMMYDCMQ	19926	GRNNAWVHV	19967	SLHPVFVNS	20008	HLDWFNMWQ
19886	TSKSYDGQC	19927	GSMMFGVAQ	19968	SQFWAHVCP	20009	ICEHRIFWD
19887	TSSVVDLNA	19928	GTSWVLVQY	19969	SRVNCFVIV	20010	IDQNFCMWG
19888	VCALYDEQS	19929	GTYLHVVAD	19970	STILMPVHI	20011	IHHGCKMWI
19889	VCCYIDCVQ	19930	GWSDMTVVF	19971	TAVAWCVSS	20012	IPMLCRIWM
19890	VCSMNDSCF	19931	HHSIKNVCH	19972	TAVLREVFM	20013	IREPMPAWA
19891	VDCIIDPQA	19932	IDFITTVHY	19973	TCYTCCVWY	20014	ISAIAGMWG
19892	VEKGQDMDN	19933	IDHHFVVIE	19974	THCTKVVDT	20015	ITAATTTWK
19893	VGDQPDAGP	19934	IHEHHPVIF	19975	TIEVMSVWN	20016	IWYVRCWWA
19894	VGVQDDFMY	19935	ILGVHNVEA	19976	TMRTQCVAL	20017	KKKNCILWR
19895	VHYNVDQCA	19936	IPHHIFVQD	19977	TPLTCMVCA	20018	LADSQVFWT
19896	VMVKRDVAR	19937	IYVMPVVER	19978	TQFHCPVVK	20019	LPDKAAHWS
19897	VNGVTDMDD	19938	KDENNPVAP	19979	TQSSQTVAY	20020	LQPRTGPWK
19898	VPPSADDPH	19939	KEDMTSVGQ	19980	TSQQGIVSL	20021	LRPHIHAWE
19899	VPWVSDWRC	19940	KEVTSQVSY	19981	TYADMWVGC	20022	LSENITGWW
19900	VQLIKDWDQ	19941	KFVHTQVEC	19982	TYHTVSVCD	20023	LSPWNFGWD
19901	VSMRCDADY	19942	KNGCITVQP	19983	VGSWVEVQN	20024	LTVPHCSWQ
19902	VSQPFDPYP	19943	KQFNHMVTH	19984	VKQYWHVEE	20025	MASCKVAWS
19903	VSRIDDAAC	19944	KSVDKYVVD	19985	VMETHWVWA	20026	MCDWREKWT
19904	VTHWCDHSY	19945	LCHDKIVHV	19986	VNPTNKVHG	20027	MFVCRQFWK
19905	WDSCSDSWM	19946	LEVVQVVGG	19987	VTELRYVSH	20028	MGAANYSWY
19906	WFRMRDCCG	19947	MLPTCMVVN	19988	YAEVNPVIG	20029	MGSIWLEWH
19907	WLVIKDVNL	19948	MMASKFVGR	19989	YHQWPFVDM	20030	MIESVSHWD
19908	WMMQVDGYP	19949	MRMPDGVNT	19990	YKKAWMVMQ	20031	MIFPIYSWA
19909	WRKMEDNCM	19950	MWEATCVRS	19991	YNPSEEVDG	20032	MPMCHCGWN
19910	YEQGVDRTF	19951	NAELYPVEW	19992	YQGTMIVGS	20033	MREVHCTWM
19911	YGCKSDEGF	19952	NGEQAFVMV	19993	YSCKERVGV	20034	MTAADVCWG
19912	YQIQFDQII	19953	NHCYVEVGG	19994	SFSYDVVMN	20035	MTLDFSTWG

[1984]

20036	NGMQICCWY	20057	SKFVRYDWG	20078	WQCAMVMWI	20099	NCPSDQPSI
20037	NHVIYHDWE	20058	SLCADPDWP	20079	YPWCYMDWV	20100	NEVHYEMAI
20038	NIDHQFHWW	20059	SMWGPRAWT	20080	YTHDTGLWM	20101	NRSKASGDI
20039	NLKWHKRWT	20060	SPYNECHWL	20081	LFRTIPIWQ	20102	NSWFKPEFI
20040	NMIQDPYWN	20061	SQGITGRWV	20082	YWCPSQWWR	20103	NWYKDVIRI
20041	NMVNAGGWS	20062	STEDCVSWY	20083	KMAGIGCWY	20104	PAESCNQKI
20042	NQMFWNIWG	20063	TADVEPPWN	20084	MYDTMGAWC	20105	RYPHEVNGI
20043	NSGPFYCWT	20064	TALRQMAWH	20085	FTPMNRDRI	20106	SCQAVWCGI
20044	PLRWDESWG	20065	TFEPYSWWL	20086	GHYWAHEAI	20107	SRETFTKHI
20045	QANWTMMWD	20066	THGCMLTWT	20087	GIFTTNFEI	20108	TCADGMDHI
20046	QIKFVEYWG	20067	THHQMFEWT	20088	GMDVNHCGI	20109	TKECSHEVI
20047	QNRYSELWA	20068	TIASFMQWD	20089	GSCITMGSI	20110	TSRFEMHTI
20048	QSYWFQQWA	20069	TKAMIWHWG	20090	HSDLMVCQI	20111	VTAMFGQEI
20049	QTSSCEEWG	20070	TMRHWCRWP	20091	HYCYGEFGI	20112	VTNPKQEAI
20050	QYKQVLDWY	20071	TRMMMQEWP	20092	IACTSLPII	20113	WMGAMVHDI
20051	RIDAQQQWF	20072	VAVKEAQWH	20093	IAENFQGCI	20114	YGPKFWIAI
20052	RSADSWEWT	20073	VCHWSLCWL	20094	ICHAKATNI	20115	YMDMRNLVI
20053	RTIAWEKWQ	20074	VFHSTHWWY	20095	ILKLPWLNI	20116	WPNIKHQPI
20054	RWDAGRTWF	20075	VSEYYWEWN	20096	MNMTYYTEI	20117	YSDPKNMSI
20055	SEDQFHHWC	20076	VTSTILHWQ	20097	MSEAIGHKI		
20056	SIDNKWPWR	20077	WILNAPEWQ	20098	MTDLARTYI		

[1985] 实施例17

[1986] 在皮肤中具有组织嗜性的AAV5变体

[1987] 此实施例描述了使用实施例1-实施例2中描述的方法和系统发现的在皮肤中具有组织嗜性的工程化AAV5变体。

[1988] 分析CNS、肝脏、骨骼肌、心脏、肺、脾脏、淋巴结、骨髓、乳腺、皮肤、肾上腺、甲状腺、结肠、坐骨神经和脊髓组织以鉴定AAV5VP1 581-589区域中驱动皮肤嗜性的氨基酸残基。结果显示于图18J中，该图显示了相较于所有其他分析组织在皮肤组织中AAV5 VP1的581至589区域中每个氨基酸的位置频率，其中衣壳多肽具有SEQ ID NO:2的序列。位置("Xaan"，n＝1-9)在x轴上表示并且氨基酸残基在y轴上表示。特定位置的特定氨基酸残基的频率在热图中作为强度表示。较深的方块表示特定氨基酸在特定位置的频率较高，较浅的方块表示特定氨基酸在特定位置的频率较低。

[1989] 分析相对于给定氨基酸残基出现在所有其他所收获组织中鉴定的变体中的指定位置的频率而言该给定氨基酸残基出现在皮肤中鉴定的变体中对应于SEQ ID NO:1(推及于SEQ ID NO:2)的位置581至位置589的所述指定位置的频率，以确定优先靶向皮肤组织的衣壳的一组序列规则。参考下文表33，和SEQ ID NO:2(AAV5 VP1)，以下氨基酸由此可以在Xaa1-Xaa9任何一个或多个位置处以任何组合独立地突变，以提供与野生型AAV5 VP1衣壳(SEQ ID NO:1)相比具有增强的皮肤嗜性的VP1衣壳，其中皮肤嗜性在此是指优选皮肤转导的特性，而不是优选转导所有其他所收获组织的特性。

[1990]

表33 皮肤嗜性规则
Xaa1选自A、C、K、Q、R或T
Xaa1选自C、K或R
Xaa1为C
Xaa2选自A、C、I、S、T或V
Xaa2选自A、S、T或V
Xaa2为V
Xaa3选自A、C、F、G、M、Q、S或V
Xaa3选自A、C、F、M或Q
Xaa3为C
Xaa4选自C、K、L、P、R或W
Xaa4选自L、P或R
Xaa4为R
Xaa5选自F、H、I、M、V或Y
Xaa5选自M、V或Y
Xaa5为Y
Xaa6选自F、H、I、M、N、Q或S
Xaa6选自M、N或Q
Xaa6为N
Xaa7选自A、H、K、M、N、R或V
Xaa7为A、H、K或R
Xaa7为K
Xaa8选自A、F、G、H、S或Y
Xaa8选自A、F或S

[1991]

Xaa8为S
Xaa9选自A、E、G、P、Q、R或S
Xaa9选自A、Q或S
Xaa9为A

[1992] 下文表34提供了在皮肤组织中发现并且符合表33中提供的一种或多种规则的变体AAV VP1衣壳多肽的581至589区域的序列。因此，本公开提供了由具有SEQ ID NO:2的工程化AAV5 VP1衣壳多肽组成的rAAV，其中Xaa1至Xaa9区域是SEQ ID NO:31991-SEQ ID NO:32990中的任一者，如表34中所公开。本文还涵盖由工程化AAV5 VP2衣壳多肽和工程化AAV5 VP3衣壳多肽组成的rAAV，所述多肽在对应于AAV5 VP1 581至589区域中的氨基酸的AAV5 VP2(氨基酸残基445至453)和AAV5 VP3(氨基酸残基389-397)中的区域处具有下文表中公开的序列。

[1993] 表34

[1994] 驱动皮肤组织嗜性的AAV5 VP1衣壳多肽中581至589区域的序列

[1995]

SEQ ID NO	581-589 序列	SEQ ID NO	581-589 序列	SEQ ID NO	581-589 序列	SEQ ID NO	581-589 序列
31991	CAFAYTIQS	32027	CSSPQYFSM	32065	CAAQHYAQV	32103	CKYDITKMA
31992	CAFMTHCYA	32028	CSTYYCNGL	32066	CAGIQATGD	32104	CLELVKVYE
31993	CAGPWCHHE	32029	CSVKCAFGC	32067	CAIHTRSHS	32105	CLNTYHYQG
31994	CAIVCFDSF	32030	CSYPWDNLG	32068	CAIMFSAGP	32106	CLSSSYNSY
31995	CAMWHQKGM	32031	CTATYTIRD	32069	CAIMLQHMT	32107	CMHLVNMNY
31996	CANCTWVAE	32032	CTCQYGWTV	32070	CALPVLVAM	32108	CMQFIRIGS
31997	CASWYATDR	32033	CTFYYFKHH	32071	CALPYAWSR	32109	CMQFIRVGS
31998	CAVWANWLC	32034	CTKGWMVIG	32072	CALSMYSQE	32110	CMVHNHKGF
31999	CAWLATDFK	32035	CTLRTWSGM	32073	CALTYAWSR	32111	CMYWLTNQE
32000	CCDHVQSSW	32036	CTMFYEQSK	32074	CAVYFGKCA	32112	CNCHIRHTG
32001	CCMPMFRNA	32037	CTMKYQEHK	32075	CCRYYVNVE	32113	CNFQTNVHL
32002	CFAAYRMQG	32038	CTMQALRQA	32076	CEKLTQNVD	32114	CNTQVYGNY
32003	CFAKWWGYE	32039	CTTKFEKCW	32077	CEYFKRHPD	32115	CNVDRYNRP
32004	CFALKINYS	32040	CTTTFACPT	32078	CFHQYHQAH	32116	CPKPFLIQE
32005	CFCIMVNGY	32041	CTWMAFTAP	32079	CFHRYHQAH	32117	CQELRTQYA
32006	CFCQRPLWP	32042	CTWPRCDSF	32080	CFKIWDLGN	32118	CQEVYAHMS
32007	CFLRNNAEY	32043	CTWVTHGQQ	32081	CFMGLSFHW	32119	CQSVHYRAP
32008	CFRYAEIGP	32044	CVCGKMGHG	32082	CFVEPKTSC	32120	CRAAHSVCA
32009	CGAIHSKRD	32045	CVDPDDYGY	32083	CFVWEFNYP	32121	CRATHSVCA
32010	CHARGYIMP	32046	CVEIMYMNE	32084	CGDCYGWCQ	32122	CSGCYMKMP
32011	CHCMPLVIF	32047	CVEYWGQCN	32085	CGGFCTALE	32123	CSGTMLSQP
32012	CITHTLKDG	32048	CVGSRPEAA	32086	CHAQIQHGG	32124	CSHHVDRVM
32013	CKVMKMEGE	32049	CVIWDQAQN	32087	CHAQSIMGP	32125	CSHQKLMES
32014	CLVWERRFS	32050	CVLTQQFNA	32088	CHFMINAGR	32126	CSMIQNSFQ
32015	CMSMNHWMR	32051	CVMHSTDHH	32089	CHHFATESE	32127	CSTDHALVE
32016	CNAKMTDLA	32052	CVQRKDRGT	32090	CHHTTNTSQ	32128	CSTMVRSQT
32017	CPNIAPIPG	32053	CVQRKMQAV	32091	CHVPLNNCD	32129	CSTWLNQHM
32018	CQLTVHSFG	32054	CVQTKINCG	32092	CIKACYDHE	32130	CTCNLPYAI
32019	CQSSYSMYA	32055	CVQYNKCRS	32093	CIKSYNVMM	32131	CTEMLQRKG
32020	CQTRFYVMP	32056	CVRYYPRPA	32094	CILHKHSFP	32132	CTGQLMHKE
32021	CRGTIAVGI	32057	CVTMALMDQ	32095	CIMLNARVG	32133	CTHGFLTGP
32022	CRMFKIDAI	32058	CVTTSNYTM	32096	CIQLNKMGP	32134	CTHKVNSWI
32023	CRSRYEAQC	32059	CVTVTYDGR	32097	CIVTVPQQA	32135	CTIFNTVDQ
32024	CSIKCAFGC	32060	CVVHKQKEY	32098	CIWMQHFSL	32136	CTIRANWHF
32025	CSITHQRGC	32061	CVVHTMRWC	32099	CKFDMPSFF	32137	CTMLANHYQ
32026	CSQKIAYAK	32062	CVVKHLFGH	32100	CKNFHWMND	32138	CTPGFWSER
		32063	CWSYGQTAA	32101	CKQQINDAY	32139	CTQFTKHQE
		32064	CAAQHCAQV	32102	CKSHFDQVE	32140	CTTAWQGAW

[1996]

32141	CTTVKMWHT	32182	CSMPQQCME	32223	DVCKVMTRV	32264	YVDKRSTME
32142	CTTYRNLLS	32183	CTMCHDCSP	32224	DVGSWLNKC	32265	AVETGESAM
32143	CTYFSMMDR	32184	CTEISWGGV	32225	DVNHTKEFN	32266	AVATRAMYS
32144	CTYMQMAYG	32185	CVGYKSNTP	32226	DVRVDMNHA	32267	AVSCGQQRE
32145	CVFQCYRSP	32186	CSESYRGPD	32227	DVYNNVKLG	32268	DVAIMQGKG
32146	CVFVQQNFQ	32187	CWECCTDSG	32228	EVCRIERYM	32269	DVLMDKRHC
32147	CVHAASLMP	32188	EVMPVNRKH	32229	EVDIVTNNY	32270	DVNEPIQCQ
32148	CVHLGHWQW	32189	EVTRFSEPG	32230	EVIFMNCPL	32271	DVNLDPDVC
32149	CVLPVLVAM	32190	GVAYEEDYY	32231	EVNEEEVTY	32272	DVYKSPCTT
32150	CVLWQVHSE	32191	HVFSHWNCN	32232	EVWWEESKS	32273	EVGINSEGL
32151	CVNQHSLAM	32192	IVADLNMFY	32233	EVYPCMSYG	32274	EVGLCQSGR
32152	CVPISAAPV	32193	IVGMQMTGE	32234	FVHCTNALF	32275	EVLKSFMCQ
32153	CVRLIDWCC	32194	IVNSFQKMS	32235	FVTCNQYAP	32276	EVMFMKKHL
32154	CVVMKFSEH	32195	IVTTVSRTA	32236	GVFYTMASY	32277	EVQVTKAQE
32155	CVYWQQNHE	32196	KVEQYMQCH	32237	GVKCINSPV	32278	EVSHRVSAA
32156	CWGPKFVDF	32197	KVTMESFDL	32238	GVMNHIAWT	32279	EVTQLSCQV
32157	CWMPQQVAT	32198	LVDHMSTGC	32239	HVCMLEFGT	32280	EVVCSMYYG
32158	CYPVYKDWS	32199	LVGPTDQLP	32240	HVMQGAPTD	32281	FVASVPDMS
32159	CYTIKQMDI	32200	MVGQLMKES	32241	HVVMKHGES	32282	FVDFNHKQW
32160	CANHLNQTI	32201	NVFYVGEWP	32242	IVAKISWSE	32283	FVDTYEAPG
32161	CWAVTSMSS	32202	NVMLGAGTI	32243	IVCPAGQYH	32284	FVEPYTDHQ
32162	CTCWYNRGD	32203	NVQKQQTDA	32244	IVEPSGHFL	32285	FVTMLHWQS
32163	CTHPVHDNE	32204	NVQRLCCLQ	32245	IVVSMDQWL	32286	GVAVIRKMQ
32164	CYQVTIQHN	32205	PVSLFGLIH	32246	LVGPEGAWN	32287	GVEEFNAQD
32165	CTLMHSHTR	32206	QVCAMVQCR	32247	LVMQLYCAG	32288	GVFPNMFAP
32166	CIIGRQEWL	32207	SVKHKEMVG	32248	LVMRMCADG	32289	GVGMKCYSH
32167	CSRNRCWSG	32208	SVNYCPRGP	32249	LVTQDMKRY	32290	GVLPVQQGF
32168	CSIQINNCR	32209	SVTCLYTSF	32250	MVTQPCAWQ	32291	GVQPWESCT
32169	CGHRCEMNM	32210	TVDRCFQYA	32251	NVERIQSLG	32292	GVSAYCTSA
32170	CNSYNKPWM	32211	TVFYCTKMH	32252	NVFIATAQT	32293	GVWIMRSKR
32171	CMHLSCDWT	32212	TVGIEWLNW	32253	NVGIFWYCG	32294	GVYYYTKQP
32172	CYAEEPVWP	32213	TVGSWAPGS	32254	PVAWCDKLG	32295	HVAFQWSGQ
32173	CNSWSTGSL	32214	TVMRALVYN	32255	QVAYAFEAN	32296	HVALQNDWG
32174	CMNHRWSTT	32215	VVVDLWWHT	32256	QVHYMMCYE	32297	IVAHQCQYC
32175	CWHFFGADR	32216	WVGCTNHGQ	32257	QVLSQWFDN	32298	IVAKYGMTS
32176	CARLEGRGN	32217	YVVDLKNFG	32258	QVVQKAMAE	32299	IVDQRSVWV
32177	CMEAQGECY	32218	AVETEFRER	32259	SVMTTFCHN	32300	IVQPCQSNW
32178	CTKPIPCWM	32219	AVNCTIYGP	32260	TVGITMLPM	32301	IVRSALTNT
32179	CDCLHHCCM	32220	AVNTQDDLF	32261	TVHEIMENN	32302	KVCFKEDAF
32180	CIWSSPDQP	32221	AVRNVEVSR	32262	VVMPKESNE	32303	LVNEMWYHS
32181	CADPFSIRG	32222	AVSTWAMQS	32263	YVCGSGMEQ	32304	LVQHSVVWM

[1997]

32305	LVTGMMSMF	32346	PVLHIYWGL	32387	HTCHFGGFV	32428	RTCWFMGVG
32306	MVAQHGISL	32347	YVHPWFGVS	32388	ISCQSMCCE	32429	SGCPQRVAT
32307	MVAQYGISL	32348	KVNACGNLT	32389	ITCQKMAMN	32430	SHCWYWYWN
32308	MVEHNRLAN	32349	PVNLAQRDK	32390	KDCFNGECE	32431	SNCPYHTFY
32309	MVGMLYKTK	32350	EVAGCATAV	32391	KPCMFWMSR	32432	SNCVTDATA
32310	MVVYRENMG	32351	SVKLAIDQK	32392	KTCAVNVYL	32433	TNCAMQMYY
32311	MVWEVKFND	32352	LVGKYIDSM	32393	LACQIQWMT	32434	TTCRTWEAF
32312	MVWEVKFNG	32353	DVMQNQCKS	32394	MTCEYRTSM	32435	VACRENMLT
32313	NVCAMPMTH	32354	HVVPCCPFH	32395	MWCDYGTGW	32436	VLCHEADGP
32314	NVCTMPMTH	32355	DVISRNGRC	32396	QACFIAGQQ	32437	VYCPCDGPD
32315	NVMPSPMDR	32356	LVDCSNTCQ	32397	QDCCKQSVA	32438	YACHMGHYD
32316	NVQAYERAN	32357	AVEDYWRYF	32398	QGCKQMGSM	32439	YPCLMEKEY
32317	NVQRHHIFQ	32358	GVAQYLHHC	32399	SSCYHSTGA	32440	YSCMVNTHQ
32318	NVVQSFTAK	32359	QVEDGYWSY	32400	STCTIRGNP	32441	YSCSQAMHG
32319	PVQHIYWGL	32360	MVFLEQHGR	32401	TACTMFECQ	32442	AACENKVCF
32320	QVEVDHCWA	32361	LVQKWMRTC	32402	VACWYTKHS	32443	AMCLFETMQ
32321	QVGVCDQCQ	32362	VVISTKHTH	32403	VICQVQGRA	32444	AQCKLTTVE
32322	QVMENYELK	32363	DVDVVALPM	32404	VYCLTHAGP	32445	ATCWQPAAK
32323	QVMQNQCYT	32364	RVTIMFTGT	32405	YICRLWTQC	32446	DNCKIMFMC
32324	QVTMQPYSK	32365	AVNACSNSH	32406	YTCVMEHMH	32447	DPCPNMIWT
32325	QVTRDGKGG	32366	TVWIQSYSN	32407	AMCPTHWHQ	32448	DWCTYRAQD
32326	QVVQYLTNH	32367	PVNESVHIC	32408	DRCNFNCSE	32449	EACYMYHNT
32327	RVDVTNHRY	32368	SVRTCHQDE	32409	DSCPQIFRD	32450	EGCYFWKQV
32328	RVEPTTVPD	32369	DVTQHTSCT	32410	EWCLADGMT	32451	EICTASPQA
32329	RVQSQWDGC	32370	PVEMKFAVQ	32411	GACKFNGHK	32452	ESCLTGGNT
32330	SVCPTRYAD	32371	GVWMQGFHP	32412	GRCQYFQMA	32453	ETCQKYVAK
32331	SVVCTTEFS	32372	QVHMQRYYG	32413	GWCWYMRMG	32454	FICQIKSAS
32332	SVWASTVFH	32373	KVAHEMMSA	32414	IGCKVQSTN	32455	HACPKHYCR
32333	TVARGPVPG	32374	DVCKATTTE	32415	ISCKAHWQG	32456	HCCFKNGDM
32334	TVGNSKMQN	32375	AVHNVKDVH	32416	ITCVDNKMC	32457	HMCPFMDQA
32335	TVHWDITGC	32376	NVYATPVVC	32417	KSCEWMLQQ	32458	IACVSWFTC
32336	TVMSAFDWG	32377	NVMGTQAKE	32418	KTCEWGQDK	32459	ILCWHHYSV
32337	TVTTKVRDS	32378	DVKAFMGGV	32419	LCCPFDTDA	32460	ISCRTQSAC
32338	VVCMHHCDM	32379	RVNGSTCRH	32420	LKCMNHSAW	32461	ITCITWVNG
32339	VVGQKMTQT	32380	SVYMMTEGW	32421	LKCVFMRAE	32462	KACDYFMHG
32340	VVQQIRDFH	32381	HVTKLMTSH	32422	LSCLATYQS	32463	KHCHSMQFE
32341	WVEWPQMGS	32382	ELCTNWTAN	32423	MLCEPEGAR	32464	KSCLYFDAY
32342	WVVCRCAWE	32383	EPCQSHEFY	32424	NYCTYHSTH	32465	MACDYFMHG
32343	YVPDRDGCD	32384	ETCPYFDFD	32425	QSCATIIRE	32466	MACPKNISG
32344	YVQLKQEWY	32385	GGCVYWQMD	32426	QSCLQQIQA	32467	MACSMNVAL
32345	YVVGMGKFC	32386	GTCTDMGIC	32427	RSCTKEDYW	32468	MMCFNYFDR

[1998]

32469	NACSHQHLG	32510	EACMPWAHH	32551	KGDRFMVDQ	32592	SYARKVYDC
32470	NPCPNMIWT	32511	IMCHARMMA	32552	LAARIGEVE	32593	ICIRGARAD
32471	NYCRRTFHH	32512	AQCHDMWGH	32553	LIARYGKQW	32594	ERMCYEDSY
32472	QACRLTYWY	32513	ESCHLQHGE	32554	LMARIQNIT	32595	ESGFYATWN
32473	QACWLTYWY	32514	WACIIYMWN	32555	MADREMAAR	32596	FGGHYWPCP
32474	QMCTLTKAD	32515	GRPRTPLGN	32556	MCIRSGFIV	32597	GHGSYLAED
32475	RACHKQLWV	32516	GTIRKEKYY	32557	MEDRLAHYH	32598	GNQPYGDNQ
32476	RPCWPWDHL	32517	IPRRLQLPI	32558	NCIRNWQHA	32599	GQVGYQMHG
32477	RSCFEMKHI	32518	ITQRDNIAM	32559	NGIRIFLFG	32600	GYHMYQKTS
32478	RTCNLPYAI	32519	IYERHHCQA	32560	PMSRQLYSD	32601	ITMSYAHMC
32479	RYCSTIYYS	32520	KDNRRIYKQ	32561	QGVRILTAA	32602	IWRDYDQCY
32480	SACPWNFYC	32521	MALRSPFNL	32562	QWMRWLIVT	32603	QKRDYQFMP
32481	SGCFWRNTS	32522	MCDRAIRSD	32563	TADRWCDAS	32604	QMNSYPMYE
32482	SHCLYWMKH	32523	QMYRLAYSR	32564	TCTRKDDDM	32605	RADAYTQIA
32483	SICQTNFAV	32524	RLSRAHLPQ	32565	TMARFPCIS	32606	SCELYMARD
32484	SKCGFDVAA	32525	SIERHCRME	32566	VSERAHCGE	32607	ACSCYMHTG
32485	SLCTIHGAQ	32526	YDWRTMAEV	32567	WEHRFSASM	32608	DISMYMFQD
32486	TACHISHYD	32527	YSERYINAY	32568	WGIRNMELR	32609	DMNMYLGSQ
32487	TACYKNTCE	32528	AASRSHQSE	32569	YADRFACYQ	32610	ECIWYDNYF
32488	TACYPHSFG	32529	DGVRMHKDW	32570	YANRNFSPE	32611	GFYWYEMAL
32489	TICSMGQKD	32530	ETKRSDHEV	32571	YENRGREHN	32612	HYAAYEQAM
32490	TICYMSTSS	32531	LCARGHKSG	32572	AFDRHSDWA	32613	IKMLYHWKD
32491	TYCRSGDVQ	32532	MCLREIQMQ	32573	MEVRECGWD	32614	ISYDYAQAP
32492	VMCTWQACR	32533	MPQRYNRET	32574	MGQRYDIQG	32615	ITMPYAHMC
32493	WHCNFHWHL	32534	NCARGWMII	32575	QAVRRWWFF	32616	KTPWYTWHD
32494	WYCLTQRSD	32535	TGERTHMCL	32576	HHLRNVHHC	32617	MASLYPFGQ
32495	WYCMCIIRE	32536	VSRRSDWSE	32577	KATRDPKFP	32618	MFTPYRMQA
32496	YACATQQHL	32537	YFPRPFKMT	32578	NSTRNTSYY	32619	NGASYNQIE
32497	YACPKHYCR	32538	YMDRMSECY	32579	MTERYYWCQ	32620	QEMDYRMST
32498	YHCKFNWLG	32539	AFDRHSDLA	32580	MTQRVVHWG	32621	QHQHYELHQ
32499	YHCQNYAAS	32540	AGGRIGDSQ	32581	NNDRHQHLK	32622	QWSMYNMYD
32500	YKCAHDPTG	32541	DFVRTNIGV	32582	NILRYLICD	32623	RLTCYRMHK
32501	IYCWMPMSP	32542	EFVRNPCFE	32583	MPTRMQKGA	32624	SAAFYMDVD
32502	EHCNHDQGV	32543	ESARYIDFT	32584	DIGRFTAYY	32625	SGLEYMGQQ
32503	KACMSHQGR	32544	FAQRTTDHR	32585	EAWRNSSSG	32626	SIQMYRYQA
32504	MSCPAKIVV	32545	FHDRPFTFT	32586	AQARTIGLG	32627	TPLWYHRTE
32505	QNCIMLASN	32546	FQDREQFWY	32587	TKGRNDHME	32628	WCGCYHIFN
32506	GTCLFKTHK	32547	GWMRGEYNV	32588	GQFRSGAYR	32629	WWFMYWADI
32507	KPCNPVGQE	32548	ICVRWHCGP	32589	LNARTMRSA	32630	YPHKYWQMK
32508	RLCDTHKCL	32549	IRVRWHCGP	32590	MMKRNRFIC	32631	YRHEYAHYS
32509	EMCGGGLSC	32550	KCQRMDTWK	32591	EYSRMNKAS	32632	TAVLYETEF

[1999]

32633	ADHVYGTEQ	32674	EYSPYQVME	32715	AASGLNAFK	32756	TTDQFNKMQ
32634	AINEYWDVP	32675	QHFWYEMES	32716	AKSCFNGFE	32757	WSHTANKYP
32635	ASDGYFMEL	32676	QIQLYQNDE	32717	ASNLMNISF	32758	YILSMNGGE
32636	ASGCYATMW	32677	MCFPYSCAA	32718	ASQWLNDSF	32759	YQVKFNVMC
32637	DIKTYMGLQ	32678	QSEHYVLPC	32719	AWVMNNWWA	32760	YRPPCNRKD
32638	DKGGYGCHE	32679	RNLNYNTAV	32720	DKSEENRQE	32761	YSTAHNEMP
32639	DLVNYGNHG	32680	ATNKYANAI	32721	ETSCVNLFN	32762	DHWQFNTHM
32640	DYQWYQVAR	32681	LLAKYMVHT	32722	FKESRNQGA	32763	EAMVINTES
32641	EMRFYESGN	32682	ITEQYDAFA	32723	GAFYINNTF	32764	QHEQSNHIW
32642	EQLQYEPSH	32683	FRLDYDYKN	32724	GLTLANIGA	32765	FTSNRNWQA
32643	GILPYVKFQ	32684	LPTQYQQQN	32725	HEPLTNQVS	32766	RIVDGNACG
32644	GSTHYTDAN	32685	GQNGYGQCM	32726	HPVDTNAGA	32767	QPQSRNGAR
32645	ICISYEGGT	32686	HSDWYMKYP	32727	ITSLCNEFH	32768	KISGLNDAS
32646	ICQVYSKTG	32687	NMPAYQIQC	32728	ITSLWNEFH	32769	QESMCNYWQ
32647	IPMQYAHFK	32688	YISEYVMMV	32729	MHAERNANM	32770	TRETMNEWH
32648	IPVQYAHFK	32689	NGNHYNAMR	32730	MTDQFNKMQ	32771	GAMLNNNAS
32649	IQDDYQLMT	32690	DPGWYGLAP	32731	MYAERNANM	32772	RMTPNNCQR
32650	LADLYDLSH	32691	VQMQYERLF	32732	MYVWRNDCA	32773	KIAAWNDFL
32651	LEQFYEGNI	32692	AYHCYMSWM	32733	NFRTINDNM	32774	KTGATNLAW
32652	LGGDYFRSD	32693	GDVSYDAEA	32734	NIDIHNHEE	32775	SQAHQNRTC
32653	LKKWYETHG	32694	MMLGYMEQD	32735	PLKGTNQIF	32776	VHHKTNRSY
32654	MAEPYGLAT	32695	ESNKWNPYG	32736	PMVIFNTDY	32777	STVGQNHCG
32655	MCDDYETSQ	32696	GIHAHNSTY	32737	QCDYRNMHF	32778	KMTNSNENS
32656	MTKPYMTQC	32697	KFASANGPL	32738	QGTHHNHVQ	32779	LPSHCNLGS
32657	MYLGYDMDC	32698	LCKTFNQSD	32739	QKVPVNDHQ	32780	TPQPMNVAT
32658	NQSAYRVAE	32699	LGNMENPPP	32740	QMLNHNHHA	32781	KWTPREKDY
32659	QHEGYVRNG	32700	LRVLGNMTC	32741	QRMITNWGC	32782	MSQKGFKAG
32660	QQPSYMVSC	32701	MFQLPNHFC	32742	QTDIRNVSG	32783	RKWMKAKAT
32661	QTPAYCHHQ	32702	NATQKNGMY	32743	QWQTTNGGA	32784	SGNPTMKNW
32662	QYQWYAKIE	32703	NPKPNNYQY	32744	RADQSNPCQ	32785	SSVCFPKAG
32663	SAHWYCHRT	32704	SSDAENRCG	32745	RHFMINAGR	32786	TWTMNEKDS
32664	THYTYAVAQ	32705	DAGCSNDNG	32746	RTECCNVAA	32787	AHKPAKKWL
32665	TNGHYHCKY	32706	GCLYANIAQ	32747	RTSQENKWC	32788	AIWQPSKMD
32666	TTTIYRRNG	32707	GTATTNFDQ	32748	SAVEGNRHN	32789	ALYYPEKYA
32667	TTTIYSRNG	32708	LLFPRNVGG	32749	SCFEVNQAS	32790	ETEPLPKQD
32668	TYAKYALRG	32709	MHRCRNTDQ	32750	SLRNRNADY	32791	IMTPHWKAC
32669	VLNTYPCNH	32710	QAMTNNVRC	32751	SMVIFNTDY	32792	ITTEFMKAC
32670	YIRMYDIAS	32711	TLDVQNVQN	32752	SWICCNAST	32793	MIGISMKGC
32671	YTMCYSHQS	32712	TSGYCNIGA	32753	SYVHWNWVS	32794	MSSALEKCG
32672	YTQNYVFPP	32713	VAAQVNQSQ	32754	TAYIHNSEQ	32795	RQIEVTKWG
32673	SIDTYDHRN	32714	VSNWENVAG	32755	THNEVNYAQ	32796	TERKMHKAA

[2000]

32797	TRASVMKYQ	32838	KQHYAYLSP	32879	LSKVQQGSY	32920	LSAFMGRNA
32798	AEALHIKAY	32839	KSTIEPFSI	32880	LYQDLLYSE	32921	MTNIAWQKA
32799	ANLKHAKFG	32840	MSAECLHSD	32881	NIVESTESP	32922	QSAFMGRNA
32800	DFFYMDKLS	32841	NCIDMWWSS	32882	NNPVLVASC	32923	QSSLEAGAA
32801	DIIWAQKAA	32842	PKSEEYTSD	32883	NNTYVMSSG	32924	SEKFIIRTA
32802	GFNPVIKYS	32843	QHNHMWSSW	32884	QAEYFDASA	32925	SIGFWMRQA
32803	GLEWHTKDW	32844	QTMKFTSSW	32885	RSMWPPSSG	32926	TGQWTFATA
32804	HQVTIFKDR	32845	RHAPTDISP	32886	TAHCFTVSP	32927	WEFSLADVA
32805	ICWLITKTE	32846	THAVLIRSK	32887	TASLLWYSN	32928	AIVKMGNPA
32806	KMYSSAKSY	32847	TQESSRVSS	32888	TEEKCYISP	32929	ATNDKVHCA
32807	LKHPTWKEE	32848	TSLKIYHSP	32889	TQLPTMQSL	32930	DASESHFEA
32808	LPGYVEKHQ	32849	VEPISSQSY	32890	TSSPHGASG	32931	DQHEVQNVA
32809	MQNACFKDP	32850	WLMKWLESD	32891	VDRFICASW	32932	EWLPCGDLA
32810	NIYETLKRI	32851	DCIDMWWSS	32892	VSSLHDDSE	32933	FGRLHQWLA
32811	QKLNIQKQS	32852	DFYTVSMSC	32893	DHDFAQFSG	32934	GCHSFCEKA
32812	QLMMEEKVS	32853	FMKFIEGSQ	32894	NNGPIWISS	32935	HCSQMDTYA
32813	RASFNDKAG	32854	KTGLPDISA	32895	PDAMWMASS	32936	KAESFTYNA
32814	RCVQPQKEG	32855	LGTFTCSSP	32896	SDRMVSISA	32937	LITTGEFLA
32815	SAEHSSKWK	32856	LSNTCCMSN	32897	AAMSSSHSY	32938	MSNSIRWMA
32816	SLGCFHKWE	32857	QWQLSWCSC	32898	VHSFHVESY	32939	NAELSCYGA
32817	SQWMIWKCR	32858	QWTQWAMSC	32899	LRVSQSFSN	32940	NFAYTMLNA
32818	TFLCTLKEE	32859	RCWHIPNSS	32900	ATMEKCRSI	32941	QNSHVIHEA
32819	TFLYTLKEE	32860	SHGWFHQSQ	32901	LLFPLMSSS	32942	SGTYLYDNA
32820	WSAKSTKEQ	32861	TGAQAANSC	32902	NTECRQNSD	32943	TQTKMMDMA
32821	YYSPMVKAH	32862	TWSYNVTSF	32903	YGENISPSP	32944	YMTEWCSWA
32822	ETHQRCKMS	32863	YTVIRCVSN	32904	TWQSMVNSE	32945	EIVQPQCYA
32823	RTDVMMKVM	32864	ALRVGWTSE	32905	TSQYSSVST	32946	AGGQKPAFA
32824	SCSFCPKDC	32865	DKQAMSHSQ	32906	GAQSMQFSC	32947	AISCGKWYA
32825	ATVHMPKHE	32866	DTVYNVLSW	32907	KTYSAMYSP	32948	ALAWMEFCA
32826	DWAWTMKIA	32867	EATDNLTSN	32908	ENTMWQSSQ	32949	FFGTPYHMA
32827	PDLCQDKLN	32868	EDESLHMSW	32909	KLKESMESA	32950	GFELVKAYA
32828	LFWQFGKQG	32869	EDESRHMSW	32910	QWMHCQLSR	32951	HTMNGVLYA
32829	PRWQSEKCQ	32870	EHVNVTYSP	32911	MQQHMKYSL	32952	IDSWQPCDA
32830	TFTYEAKFW	32871	ESKSEQPSA	32912	ECRMAVQSC	32953	KCALIQDIA
32831	NHDPSIKDQ	32872	FLEAWANSM	32913	GAANRIQSW	32954	KHAQMSSDA
32832	RKWMNAKAT	32873	FRDGMIVSC	32914	QRNEMSASR	32955	KKGDSTREA
32833	IKTTMVKED	32874	GIGQPEDSP	32915	DMTLQSVSS	32956	LMMVILDTA
32834	FEEFTTISR	32875	ICATILRSA	32916	ENGAISEPA	32957	MAANCEVFA
32835	GGIFEQSSS	32876	KSGISEMSE	32917	HGNLMYTGA	32958	MAEYTDAAA
32836	GSFWFTDSK	32877	LGRDCSRSH	32918	ISVCKQMIA	32959	MWNYGHNYA
32837	IAGPAWFSM	32878	LNMATSLSR	32919	IWQYKDSDA	32960	MYGQNWNVA

[2001]

32961	NCIHNWQHA	32969	SSFPMWWIA	32977	DCDWQSAMA	32985	KRTIEHPCA
32962	NHFLASVLA	32970	SSGLNWRDA	32978	ALQFVSSTA	32986	RQMCMAHHA
32963	NSKIVDYCA	32971	STGSHGVCA	32979	AWDANLAFA	32987	TSEPHCSYA
32964	QCWFRGDYA	32972	WTKWECCVA	32980	QKGMMLDCA	32988	ICNMDFRCA
32965	QCWLRGDYA	32973	KMLPTCSDA	32981	VFGSTRSAA	32989	AHAWLQHWA
32966	RATCGISMA	32974	KPNPFVAWA	32982	RGGWCMRKA	32990	HIHMSDRGA
32967	RKHGLLFMA	32975	FGQFCCQVA	32983	INENMQWLA		
32968	RMAHTFWDA	32976	MPEKDGCPA	32984	PPMPVSHFA		

[2002] 实施例18

[2003] 在骨髓中具有组织嗜性的AAV5变体

[2004] 此实施例描述了使用实施例1-实施例2中描述的方法和系统发现的在骨髓中具有组织嗜性的工程化AAV5变体。

[2005] 分析CNS、肝脏、骨骼肌、心脏、肺、脾脏、淋巴结、骨髓、乳腺、皮肤、肾上腺、甲状腺、结肠、坐骨神经和脊髓组织以鉴定AAV5VP1 581-589区域中驱动骨髓嗜性的氨基酸残基。结果显示于图18H中，该图显示了相较于所有其他分析组织在骨髓组织中AAV5 VP1的581至589区域中每个氨基酸的位置频率，其中衣壳多肽具有SEQ ID NO:2的序列。位置(“Xaan”，n＝1-9)在x轴上表示并且氨基酸残基在y轴上表示。特定位置的特定氨基酸残基的频率在热图中作为强度表示。较深的方块表示特定氨基酸在特定位置的频率较高，较浅的方块表示特定氨基酸在特定位置的频率较低。

[2006] 分析相对于给定氨基酸残基出现在所有其他所收获组织中鉴定的变体中的指定位置的频率而言该给定氨基酸残基出现在骨髓中鉴定的变体中对应于SEQ ID NO:1(推及于SEQ ID NO:2)的位置581至位置589的所述指定位置的频率，以确定优先靶向骨髓组织的衣壳的一组序列规则。参考下文表35，和SEQ ID NO:2(AAV5 VP1)，以下氨基酸由此可以在Xaa1-Xaa9任何一个或多个位置处以任何组合独立地突变，以提供与野生型AAV5 VP1衣壳(SEQ ID NO:1)相比具有增强的骨髓嗜性的VP1衣壳，其中骨髓嗜性在此是指优选骨髓转导的特性，而不是优选转导所有其他所收获组织的特性。

表35 骨髓嗜性规则
Xaa1选自A、E、G、Q、S或T
Xaa1选自A、E或T
Xaa1为E
Xaa2选自A、I、Q、S、T、V或Y
Xaa2选自A、S、T
Xaa2为A
Xaa3选自A、G、I、M、Q、S或T
Xaa3选自A、Q或T
Xaa3为Q
Xaa4选自A、E、P、Q、T或V
Xaa4选自A、P或Q
Xaa4为Q
Xaa5选自F、I、L、M、Q、V或Y
Xaa5选自F、V或Y
Xaa5为V
Xaa6选自F、I、N、Q、S或V
Xaa6选自I、N、Q或S
Xaa6为S
Xaa7选自A、C、M、S或V
Xaa7为A、C或V
Xaa7为C
Xaa8选自A、C、D、G、M、S或Y
Xaa8选自A、M、S或Y
Xaa8为M
Xaa9选自D、E、G、L、P、S或Y
Xaa9选自D、E或P
Xaa9为P

[2007]

[2008] 下文表36提供了在骨髓组织中发现并且符合表35中提供的一种或多种规则的变体AAV VP1衣壳多肽的581至589区域的序列。因此，本公开提供了由具有SEQ ID NO:2的工程化AAV5 VP1衣壳多肽组成的rAAV，其中Xaa1至Xaa9区域是SEQ ID NO:4118-SEQ ID NO:5117中的任一者，如表36中所公开。本文还涵盖由工程化AAV5 VP2衣壳多肽和工程化AAV5 VP3衣壳多肽组成的rAAV，所述多肽在对应于AAV5 VP1 581至589区域中的氨基酸的AAV5 VP2(氨基酸残基445至453)和AAV5 VP3(氨基酸残基389-397)中的区域处具有下文表中公开的序列。

[2009] 表36

[2010] 驱动骨髓组织嗜性的AAV5 VP1衣壳多肽中581至589区域的序列

[2011]

SEQ ID NO	581-589 序列
4118	EAEDEIQKA
4119	EASWPMDWD
4120	ECIHDQCSS
4121	EDIDLKGAQ
4122	EEGPQRMNH
4123	EGCCYWSHA
4124	EHWPMAMFN
4125	EHWQKDYNH
4126	EIYWMMSNV
4127	EKKSCQCWH
4128	ELDSWWRIA
4129	ELKESMESA
4130	EMAQTSETP
4131	ENGHKFNEE
4132	ENQTRMTSW
4133	EPTSEVHHD
4134	EQFNHTWHG
4135	EQHDWYQKY
4136	EQIALHCAW
4137	ERFAFVADH
4138	ESSRYRMMM
4139	ETHEFEKLC
4140	ETRTTAWSA
4141	EVITWGHMQ
4142	EVKDKPSIF
4143	EVWNAHEEG
4144	EWESSSIIS
4145	EYTHESHWA
4146	EIEGMYWPL
4147	ETAWKWEWS
4148	EQLLSLCPS
4149	EKNQSTCET
4150	EYARGELEN
4151	EHPEWKCMY
4152	EMEHYPCYN
4153	EYSMSEQTN
4154	EKGGQMCAN
4155	ECDRLKNER
4156	EWCPGEPDG
4157	EEPRILYAG
4158	EHKLFMKPE
4159	EKIMVKRAF
4160	EAKEDWGVR
4161	EDKYMTKRH
4162	EGDLSGFQC
4163	EGKPRCFDW
4164	EHFQYQNHI
4165	EHKYFNSGH
4166	EHWYNRKGL
4167	EIWMWTQGC
4168	EKTAGMQTM
4169	ELPWQRWEN
4170	ELVPMKLRS

[2012]

4171	ENHEEKMAE	4212	MAGPCSSDT	4253	SAATFKGSP	4294	VARDHRAGT
4172	ENLSVNNLF	4213	MAIIRAEVN	4254	SAAYCSYLR	4295	VASLGLYCM
4173	ENMDDHYDT	4214	MAIRWHEVY	4255	SAAYFPARG	4296	WAPVHWMKD
4174	ENWCIRYAK	4215	MAIYKCAVN	4256	SACGAQYPP	4297	WAYNSVASE
4175	EQHLKPFEA	4216	MAKYPVNWG	4257	SAEAVVRLW	4298	YAASHMEMP
4176	EQKHCWSRD	4217	MANDKHDPQ	4258	SAEVQATTY	4299	YAHIVYRHG
4177	ERMLVHMRS	4218	MARQEGWCK	4259	SAGLRVCDR	4300	YAMDCMGGV
4178	EVQWNNPKP	4219	MASGSVEKM	4260	SAIFCLKKH	4301	YAQTDGCLE
4179	HAAGNEFCY	4220	MASHIWNQW	4261	SAIIVDCHY	4302	YATQNTFSY
4180	HACDTQIHP	4221	MAVGNLSEA	4262	SALDVQRPS	4303	YAVEYGWNQ
4181	HAENGAVWR	4222	NAALHSADW	4263	SALQISMLY	4304	AAHNYEGVA
4182	HAISRQQVE	4223	NAANQDHTW	4264	SANWVQRDM	4305	CATHMIRTR
4183	HALTTDYYN	4224	NAEAAWMTF	4265	SAPSMWWAS	4306	DADDQAVSG
4184	HANQIGWCL	4225	NAELMDAHG	4266	SAQNHDCWL	4307	DADFCPRMP
4185	HATSSTGDD	4226	NAFCGNHYL	4267	SARPFEFME	4308	DAGLDSAGY
4186	HAVMHSAVS	4227	NAFNYRLVI	4268	SASNEFDGL	4309	DAHALQNVI
4187	HAYNTFQFG	4228	NAGHYMEYK	4269	SAVAHMSET	4310	DATGKFNFN
4188	IADHTIADP	4229	NAGPCWSEE	4270	TACIAYVDN	4311	FADKWANQF
4189	IADNIMQAC	4230	NAMPRIQFP	4271	TADFNVRVE	4312	FANLFANTW
4190	IAGMTADDG	4231	NAMSNAAML	4272	TAFWMKGPQ	4313	FANLFTNTW
4191	IAHDGQADS	4232	NANPHHAVP	4273	TAGLSHFKS	4314	GANNQNSCE
4192	IAHIKKVSL	4233	NANQSCANN	4274	TAGLVTGLG	4315	GANQHDVGH
4193	IAILLHNFD	4234	NASHDMHDL	4275	TAHCHITYL	4316	GAQCLFFQN
4194	IANSKMWEA	4235	NAVQTLQLY	4276	TAHGCPTQH	4317	HAIHDCYPP
4195	IAYCTDAYE	4236	NAWFPMTNT	4277	TAKTHLFPW	4318	HAILYFYFP
4196	KACQNSNFW	4237	PACFRVMFV	4278	TALHIDWIP	4319	LAIQPCDER
4197	KACTDESGF	4238	PAMNFITFF	4279	TALMSSLQA	4320	MAGDYACAM
4198	KARDLEYIQ	4239	PAPKNQSTD	4280	TANLHTIEG	4321	MASSVVHSC
4199	KAYDTMKCA	4240	QAACTHPWS	4281	TANQAYSTI	4322	MATNREWWG
4200	KAYSEILGY	4241	QADQMPTQE	4282	TAPTIDYGG	4323	MAVWAKAAE
4201	LAANSVLYF	4242	QAEGGPNMP	4283	TATGMKHVY	4324	NARDPACDN
4202	LADWYSARI	4243	QAKTSYGGE	4284	TAVHLLDPD	4325	PADRNNLWQ
4203	LAEIMKCPP	4244	QALVHGAPL	4285	TAVRYEVCG	4326	PAECISDSQ
4204	LAFHLNGTS	4245	QAMMAQRAG	4286	TAWKQNVSD	4327	PATVKPMYW
4205	LAGCMGLQE	4246	QASQYTVKY	4287	TAYYEEFWR	4328	QALNRVQFN
4206	LAMAAHEEA	4247	QATHWPMWR	4288	VAAHEGRKS	4329	QATGMLYEE
4207	LAMNQSKAI	4248	RACHPPHLK	4289	VADRQNEHK	4330	QATVTKEYW
4208	LANLHEWEG	4249	RAEPMIYNF	4290	VAFSLMIMD	4331	RAIHDCYPP
4209	LASITTPPS	4250	RAEQAPVCN	4291	VAGFVSPTG	4332	SARDGISMH
4210	MAANNECLV	4251	RAQDTEMML	4292	VAGNGIPMP	4333	SAVYIQPDY
4211	MAGMMEEGS	4252	RAQLYGMDS	4293	VAGWATPQL	4334	TADVVPCHH

[2013]

4335	VADYLNEFT	4376	ISQWCANNK	4417	SIQMMQDHV	4458	ISQHTCSKD
4336	VANNGMAVK	4377	KDQCIKQAP	4418	SIQTTFNDD	4459	KEQPSQMHG
4337	VAWSQEESQ	4378	KIQPNDRTQ	4419	STQKYWNYE	4460	MQQANSDYA
4338	YAQQFTMPT	4379	KVQENEVLG	4420	TGQLNTGCV	4461	MTQCRDCWL
4339	QAAGETFYG	4380	LKQQFAEGM	4421	THQTARQNV	4462	MTQDFNCDY
4340	AACWTNMQW	4381	LVQLNMAGA	4422	TLQWEPKGV	4463	QRQQMSHDC
4341	HAPMDNAPF	4382	LVQSIWGEE	4423	TNQCICAAP	4464	QRQRMSHDC
4342	NACDWKNDY	4383	LWQLFTMQQ	4424	TNQNLGNDY	4465	RHQLLEWDS
4343	IAMLESWWG	4384	LYQQLESSG	4425	TSQSMVCCE	4466	RVQHCYGDM
4344	DAGRFTGTS	4385	MLQSHTCLG	4426	TTQWSQLQS	4467	SKQPNADSD
4345	TAPRIDQQE	4386	MNQFGNMKQ	4427	TVQSMTSFA	4468	VCQKYDEWG
4346	VATDSVHES	4387	MQQPFYRSM	4428	TVQTPPQYV	4469	VTQQYWACD
4347	MAEQMGCQA	4388	MYQVWCTDT	4429	TWQSLCDFE	4470	VTQQYWACN
4348	DARFDQWHH	4389	NCQHNEESP	4430	VCQLMKYQP	4471	WDQPNMHTF
4349	PACFDDGEF	4390	NDQFRVVDA	4431	VDQVFVEDR	4472	WGQHKHTVA
4350	LAEPWATSH	4391	NFQPMHWST	4432	VPQTDLDLC	4473	YCQEFEQYP
4351	PAVGFKDEQ	4392	NIQGVLDYR	4433	VQQIIGTIG	4474	YDQWECDIH
4352	HANHLMSHE	4393	NPQKWNPCI	4434	VSQPCLIVD	4475	YDQWERDIH
4353	LAHCDMMMC	4394	NQQCFDYYS	4435	VVQWFAQQV	4476	HDQMMGLER
4354	GADDHHLMT	4395	NVQIATTDS	4436	VWQQQFVWL	4477	TNQCIFKTM
4355	YAPKMWAEA	4396	NYQPFYVWA	4437	WDQFCRLMM	4478	CHQCIGQDN
4356	WANTYSENV	4397	PHQLHENWG	4438	WNQIMEYDY	4479	GDQFQDSTC
4357	RAIYNDMTE	4398	PLQLDHECT	4439	WNQPLMWCF	4480	THQYFASLA
4358	AAMGYAKMS	4399	PSQIDHAEQ	4440	YCQQTYTEY	4481	IGQHNVQTT
4359	AAPIRCYFE	4400	PWQYCELGA	4441	YEQKENQSY	4482	LKQCQSTMF
4360	CAHPEEYHQ	4401	QCQSEHRDI	4442	YGQWFDMSY	4483	ALQQNPSMT
4361	DAHFPGLAQ	4402	QCQSFGVTQ	4443	YIQHMHIQR	4484	MMQCRDCWL
4362	DAQCFSFGH	4403	QDQLYGRDE	4444	YIQTKQIMQ	4485	SCQEWGIQW
4363	HAVGPVMLE	4404	QDQQKQYDW	4445	YNQAVCQEP	4486	ANQFWKIHC
4364	GIQDEAWVA	4405	QHQHLCRTS	4446	YPQPFMLMF	4487	GHIQYMRST
4365	GQQSFQQVD	4406	QKQFNAQMG	4447	YTQLCHHGV	4488	GIIQTGCYK
4366	GSQMMQDQA	4407	QLQWFVKDS	4448	AIQSRRSVG	4489	GIVQGYEGS
4367	GVQLTIQQA	4408	QNQIEHWCA	4449	AKQCLQDYP	4490	GKTQFRKCS
4368	GVQPALYED	4409	QPQFMVGQI	4450	CFQQKHNLF	4491	GMYQKCYFA
4369	GYQPDMGMS	4410	QVQFDRCYL	4451	CGQSVSHAG	4492	GNCQQVQMG
4370	HCQIDVLRV	4411	RFQCEGHLD	4452	CQQYTLVQA	4493	GQDQCCRIW
4371	HTQFEPGYH	4412	RHQLLEWYD	4453	DDQFQDSTC	4494	GSSQGRARA
4372	HVQVDQSKS	4413	RTQPDGQTR	4454	FDQPTPDHP	4495	GSVQEGYDA
4373	HWQHETMAQ	4414	SDQQWKPTH	4455	GEQHQMPMV	4496	GTRQFEGNC
4374	INQLKHSAY	4415	SGQPFSYSK	4456	IGQHDVQTT	4497	GTTQDWRVP
4375	ISQCDTAII	4416	SHQQCVMDP	4457	IMQQWPAMF	4498	GVIQDQGMF

[2014]

4499	GVVQRATIE	4540	NVLQLDAFL	4581	VVIQDIRKD	4622	LKSQGCEAY
4500	GVYQKQSMQ	4541	NWAQAHWYG	4582	VVLQQVDPA	4623	SCCQHDNNA
4501	HDLQYPMSF	4542	NWDQTMSQD	4583	WERQYVPFV	4624	YHEQCQWHS
4502	HFHQMIFGG	4543	PEYQVERSE	4584	WIHQQFAEI	4625	NMMQYQWAI
4503	ICAQANIAR	4544	PMIQTDCRW	4585	WPCQQQSTW	4626	MIPQHARGP
4504	ICDQPGTTG	4545	PPPQIYLMA	4586	WSGQNYNGA	4627	TGRQYYSCC
4505	IGVQTSFQD	4546	PWVQTNIGL	4587	YFKQFMNCY	4628	GNLYVPSWG
4506	IHLQWQSMM	4547	QCFQEGWQD	4588	YMNQQNKDE	4629	GVDLVMMLK
4507	IIDQKYSWC	4548	QDIQLSHHT	4589	YSLQTEVND	4630	GVMLVGHKY
4508	INLQGHHWP	4549	QEVQFFATD	4590	YVSQSSYNY	4631	HDFHVKRRC
4509	ISCQEEISM	4550	QHVQRNYAS	4591	YYGQNYMGC	4632	HNDWVIMAP
4510	ISDQNVAHR	4551	QQCQTGTPG	4592	ARAQIQDEW	4633	HNVHVTTKW
4511	ITLQTANAP	4552	QSCQAIVED	4593	CVMQERQAN	4634	ICKLVMKED
4512	IVLQYKFME	4553	QTTQQWCMQ	4594	DLEQAMFMG	4635	IEIYVNFCT
4513	IVMQQWFPF	4554	QYSQIHMQK	4595	DPCQEWDSE	4636	IIDYVAHLL
4514	KCEQIENNP	4555	RGYQHEMMG	4596	FEAQIVLPG	4637	INELVEWCP
4515	KEIQFMTMH	4556	RQHQGHYDY	4597	IHLQWIWAN	4638	ISCSVESLS
4516	KLVQNMKRD	4557	SNEQLDDDR	4598	KHDQSGIAN	4639	KDKEVIVNC
4517	KQDQCMCWY	4558	SNIQDDPKG	4599	LSGQIDEMS	4640	KDMCVHKYT
4518	KYEQLLGCY	4559	SQAQCEMMD	4600	LYMQAPSEG	4641	KEWLVDESQ
4519	KYYQAFSKY	4560	SSDQQGWVT	4601	NCLQKMQIW	4642	KMKMVKEVC
4520	LNWQFVECQ	4561	SSLQCFDME	4602	NDLQDKSHQ	4643	KVEEVNVQQ
4521	LNYQIALCR	4562	SSMQMPSGN	4603	NVSQITQDC	4644	KWTEVDGSW
4522	LPHQYNVVY	4563	SYVQTEWVY	4604	PNNQARGQC	4645	LCEHVQNCM
4523	LQTQRHWIP	4564	TIAQTHWFW	4605	PVIQPCDER	4646	LDDKVADHW
4524	LSVQGLCSY	4565	TIKQCSFCL	4606	RWKQWCWEF	4647	LDKIVGMWS
4525	LTSQSNHWG	4566	TISQTQGYM	4607	VSYQTQDML	4648	LGDNVDCWQ
4526	LWMQHQHGS	4567	TQTQINQGR	4608	YDLQTIKQL	4649	LMGDVTHMF
4527	LWSQYGLWV	4568	TSMQSTQGK	4609	NITQLQCYC	4650	LRVWVMYLS
4528	LYLQNMHCE	4569	TTMQRLHFE	4610	FHPQTSDND	4651	LSCHVQMNS
4529	MDVQTMNDG	4570	TYAQRHVMG	4611	MEWQSEYIS	4652	LTSDVHIHM
4530	MSFQETTSI	4571	VDKQMQIYL	4612	DHPQMHFYW	4653	LYEAVAQME
4531	MTVQWFDIN	4572	VHCQADTAN	4613	LVAQQDSCE	4654	MFPLVGMIV
4532	MWDQTEMFN	4573	VINQNIQCQ	4614	LWKQWCWEF	4655	MMTMVCNAL
4533	MYSQRNCLY	4574	VMKQGNQEE	4615	LVIQPCDER	4656	MPASVGSYT
4534	NDRQHHHKC	4575	VPAQQSTFV	4616	KPVQCACQN	4657	MWSPVFNWL
4535	NNYQKLSFD	4576	VPHQLCHIG	4617	LIIQHPNMA	4658	NCEWVTKLA
4536	NSIQFSSWR	4577	VQVQGLAIM	4618	SFLQYHERT	4659	NETGVFRQA
4537	NTAQNEGWM	4578	VTAQAESWF	4619	GLPQSADVI	4660	NFCAVEQGV
4538	NTEQMPPLS	4579	VTAQQMNLG	4620	DQVQGAHLF	4661	NFGYVRAVG
4539	NTVQDMGVC	4580	VTVQKQFDH	4621	SGDQHLSEG	4662	NIWGVDATE

[2015]

4663	NPAPVDQHY	4704	VHNMVIAFD	4745	IVNTCSQSW	4786	WFMEPSPQR
4664	NQNAVECIQ	4705	VIGGVTSYW	4746	KRKEASTRW	4787	WMHEDSSQY
4665	NRDVVRAWH	4706	VRELVMWGY	4747	LFEHSSVAN	4788	WTFMDSTTD
4666	NTEHVLVTP	4707	VSGFVFVWR	4748	LGGLISCPG	4789	YCSPFSGDA
4667	NVFLVPDMD	4708	VTDYVPMNQ	4749	LHISLSKSC	4790	YLNKRSVNA
4668	NVVSVPVKS	4709	VTEIVCESG	4750	LISEFSHDK	4791	CISMYSEGD
4669	PESHVVSVW	4710	VVGVVIYTR	4751	LLTHISYME	4792	ISGLMSDYQ
4670	PNYVVQWNF	4711	VVVAVNADG	4752	LPGIQSTMR	4793	IWMLPSVTT
4671	PTTIVGWRD	4712	VWELVSHVQ	4753	LVSAYSEKF	4794	KQHNYSGRA
4672	QCDDVMISE	4713	YDSGVFTIS	4754	MDIDLSSHA	4795	KSGLQSITR
4673	QCTFVPFET	4714	YYHGVVDSE	4755	MGAFTSINI	4796	LSEVFSAQR
4674	QGKYVGMIM	4715	CHAAVTNLY	4756	MNVAFSMED	4797	LSMVFSKGE
4675	QGYYVNSVD	4716	DLIIVPWSN	4757	MTPAMSFSP	4798	MHFPWSGEA
4676	QMKMVNTSK	4717	LEILVAINE	4758	MTTRESICE	4799	QVPWESKLC
4677	QWCPVQMQC	4718	LMTWVEKGA	4759	NFDHISAST	4800	SKEEDSTML
4678	QYNVVKCVA	4719	LMTWVEKGD	4760	NLRYASNWP	4801	VEMCYSWNY
4679	RCDNVEACF	4720	NTYGVAFDT	4761	NMMSKSLDS	4802	YYTIASANV
4680	RDGPVVPPP	4721	QYHWVTRII	4762	NMSPFSSHS	4803	IPSADSWPM
4681	RIASVQWEE	4722	RPPHVGPAL	4763	NMSPMSESY	4804	LPRTSSQIG
4682	RPTLVNAHL	4723	RQEFVFDCY	4764	NQYTASNIC	4805	QWGFGSDET
4683	RQIPVEDSN	4724	SGKLVMCGY	4765	NTYSWSREW	4806	AVLLASDQR
4684	RTDLVAVCG	4725	SSGPVIILR	4766	QTGAASANR	4807	DWNWKSCEW
4685	RWSHVRVLW	4726	TCFHVAEMF	4767	QVFCPSVDA	4808	TSSNDSTAN
4686	SFAHVGLPI	4727	PQCHVMADR	4768	RGVGMSTDN	4809	IHGYHSCYF
4687	SGHMVVCCR	4728	CSDLVGTNC	4769	RVAEMSNTP	4810	HVMLLSPNQ
4688	SNAPVHTCD	4729	IRMNVVLTP	4770	SKNCLSFSW	4811	LSELFSAQR
4689	SPATVAAGC	4730	ISMEVNRGE	4771	SMVTDSRWT	4812	DIKMFSFHT
4690	SSIGVMGLY	4731	ALDRVSMAC	4772	SSGKDSNAA	4813	GHTGSWCDR
4691	SSLCVNDQW	4732	FFEGVNINY	4773	STDDESMPG	4814	GKYRIACDT
4692	STFRVGDCI	4733	PMYVVPEMC	4774	TECPKSLTS	4815	GMADLKCPS
4693	STTIVGWRD	4734	GKLDISNLC	4775	TGKDESHFY	4816	GMVEAECMD
4694	TCDYVCAQV	4735	GLCNPSIVT	4776	THFHTSREP	4817	GQICILCSE
4695	THTSVGGQC	4736	GQLYCSRYD	4777	TMEFTSPCS	4818	GQYFKECYD
4696	TRCDVEGYA	4737	GSNFGSIDK	4778	TNHGFSNWV	4819	GVNILTCEC
4697	TTYWVREMS	4738	GSNPFSRHG	4779	TQDWNSTAA	4820	GWKYEQCTD
4698	TVMTVAAMQ	4739	GYAKESAKR	4780	TTHWQSWIQ	4821	GWTPQLCSS
4699	TYRLVQQCM	4740	HHMRMSTLP	4781	TYGMCSTFQ	4822	HCSSMLCQS
4700	VFESVVITE	4741	HMDDSSICA	4782	VCLTHSFSN	4823	HNIPFYCMH
4701	VFTPVEKSG	4742	ICREESETP	4783	VIEMESGDF	4824	HRLMYDCHK
4702	VGDRVQKNS	4743	IGEKISAMP	4784	VNLVDSIRD	4825	HSCPLACMN
4703	VGTTVNVRS	4744	INEGESTYD	4785	VVNLFSFRS	4826	HTIGYQCHD

4827	HVAEYCCKH	4868	STMGTKCSM	4909	TSVMWACPG	4950	NTFLCQKMA
4828	HYDTHQCWF	4869	TDREIHCMN	4910	GICGTRSMK	4951	NTRFNVDME
4829	IEWDHLCSH	4870	TFFPGHCYQ	4911	GIMLSQDML	4952	PLTLADKMQ
4830	IHHGSICCT	4871	TLIEKLCED	4912	GIVADRSME	4953	PWYWPFEMY
4831	ITETRACAA	4872	TPMRCYCFG	4913	GLGYGHHMM	4954	QGNWTMLMQ
4832	KEEMMQCEE	4873	TTDTATCKS	4914	GMVFTNYMR	4955	QKECYPNMV
4833	KGEPENCST	4874	TTHKMTCCP	4915	GSELGGVMG	4956	QSGVDFAMA
4834	KQDMARCWW	4875	TTYKQECTT	4916	GVLPSYMMM	4957	QSLREENMQ
4835	KTDSINCQQ	4876	VDEWYVCMW	4917	HSPPCQIMP	4958	QVLKGEEMP
4836	LFEPINCYH	4877	VDRHGGCQF	4918	HWTTEEQMM	4959	RHDHMMNMA
4837	LSEVCLCSE	4878	VFSKFVCQL	4919	IMCTKARMD	4960	RSVVNEKMF
4838	MCDANHCNS	4879	VGSLYNCSG	4920	IMIFEQMMH	4961	RVCNNESMF
4839	MKGHNDCIS	4880	VNTLRHCQS	4921	IQGCQMGME	4962	SCVVTDNMR
4840	MMGVRNCHM	4881	VPTNLNCCQ	4922	ISHNNAHMQ	4963	SDHVLHEMC
4841	MSTRCHCRY	4882	VREHYKCGW	4923	ITTPYGHMY	4964	SEREMRHMP
4842	MTYKADCPK	4883	YENAMICEM	4924	KDLRSMEMC	4965	SFASKIAME
4843	NDDTMHCAY	4884	YTENKECCP	4925	KFMGYMVMP	4966	SLDHMPMMG
4844	NHTHITCTE	4885	AKNYFWCKP	4926	KFMGYTVMP	4967	SLTFFQYMD
4845	NMMIMHCHT	4886	CISILGCAH	4927	KHWFDPGMW	4968	SPSSHCVMI
4846	NNTTTRCLM	4887	FTEEGACGM	4928	KIMEPTFMC	4969	STSPYEHME
4847	NPLMQCCYL	4888	GGMDLICHG	4929	LCGLNTKMY	4970	SVGTRINMA
4848	NRADMICAK	4889	HWVMQMCIG	4930	LFHPDMMMP	4971	THSSLASMV
4849	NSAPSMCST	4890	IMEPLVCDD	4931	LGDDKYIMS	4972	TLCSQNWMM
4850	NTEPNNCWC	4891	KGCMDWCDS	4932	LNANMKEMF	4973	TNHIQNEMD
4851	NTVYDNCGE	4892	MKYDRPCCT	4933	LQTSTINMQ	4974	TNNEDTQMH
4852	NVKHCHCEI	4893	NNDHFVCYL	4934	LSFDFPTMC	4975	TRHYNCTML
4853	NWTYTFCQQ	4894	NTNMRHCGI	4935	MCALNPGMP	4976	TVVTKGTMP
4854	PFPMEQCWQ	4895	RSIPGGCVA	4936	MCFMQNDMR	4977	TWIEMDYME
4855	QCKWGDCAC	4896	SFYSHICRI	4937	MDWSFGKMG	4978	TYPKRMNMM
4856	QVHRPECRH	4897	TCHWIQCFN	4938	MHDGGMEMP	4979	VCDPWFAMP
4857	RFWLQCCTE	4898	TGEAQDCNP	4939	MILGEMSMY	4980	VDAPQPNMP
4858	SFENSICTK	4899	TVPMDHCWA	4940	MNLNKIMMY	4981	VMYMRDQMW
4859	SFYERACFS	4900	VMKIKRCEG	4941	MPYCEHWMI	4982	VPKWGTGMF
4860	SGFLTVCDR	4901	VNALFDCCY	4942	MSIPFCWMQ	4983	VSAYEQNMK
4861	SHCMLNCWA	4902	VTMNFHCGC	4943	MVKPCGSMF	4984	VTWHPLSMH
4862	SKILFDCTS	4903	VNDYDQCCM	4944	NFHRWDRMT	4985	VVEPQMRMC
4863	SPLITNCEA	4904	DYAENWCIE	4945	NIEWFKTMG	4986	WITLNQKMV
4864	SQVRTVCGF	4905	AWRPPNCEL	4946	NKEDNHMMM	4987	WWCHPMHMM
4865	SRNPMMCHP	4906	NPIKFQCAH	4947	NKNEMAMMG	4988	YDKTYMAMG
4866	STKNDICEG	4907	VYANENCAP	4948	NNCHYNIMA	4989	YQFNWGNMT
4867	STLPSKCGM	4908	PHEDCECLR	4949	NQFPATEMQ	4990	YTKETGGMV

[2016]

[2017]

4991	DKIPINGMS	5023	HQTPSELAP	5055	NSTNNGWQP	5087	TWSSIDMWP
4992	DVEKTMQMN	5024	HSAAPTEGP	5056	NSVKQNAAP	5088	VFLVSGMLP
4993	FHFIRHEMV	5025	HVFSIFHVP	5057	NTGLSDFPP	5089	VGLTMTGVP
4994	FQMAMCSMG	5026	IMRLDARCP	5058	NVTTAWNCP	5090	VRNIGGEWP
4995	GRFEEGAMY	5027	INTGFMNKP	5059	NWHWPPVFP	5091	WDCTMYHFP
4996	IGVMLNAMQ	5028	ISHPSVGEP	5060	NYDFYGMIP	5092	YPHTPHNCP
4997	KDRCQYEMA	5029	ITMPRMDNP	5061	NYGIAYNFP	5093	YPHTPYNCP
4998	MCGFETTMP	5030	KDVEFFHGP	5062	PLIIQQFWP	5094	AQPNGNEHP
4999	MHDDYRGMQ	5031	KSEENYNHP	5063	PSACMAGKP	5095	CISIIAITP
5000	MSRRKEAMP	5032	KVDDHAFEP	5064	QFEDMYKSP	5096	CISITAITP
5001	QRIWTTYMH	5033	KVDNEFYVP	5065	QHANFAVQP	5097	CVLSWCVRP
5002	RRFEEGAMY	5034	KVWPCNPDP	5066	QNSPCTAFP	5098	DFNDSNWLP
5003	SELFRANMD	5035	LCAIWTLCP	5067	QSCLFMQIP	5099	DMMVSGHVP
5004	VMCNMMQMT	5036	LFRFDMSPP	5068	QSTWAERWP	5100	DQLTWHTGP
5005	WNAPKVMMA	5037	LHCSPEIGP	5069	QTMLYWYLP	5101	DQMDEFLYP
5006	DKIPINGMP	5038	LHESLDDFP	5070	QTWKHMLGP	5102	GIDRPHFVP
5007	SFHPSWWMY	5039	LPCHAIVSP	5071	RCIWKCEGP	5103	HNNNHAIPP
5008	NCFDSEDMG	5040	LVTLNHNKP	5072	REDPSALEP	5104	LGNWHHNSP
5009	VTTNELGMH	5041	LYGDAQGGP	5073	RHDLYVWSP	5105	LHYNNERDP
5010	TTFPQVIMN	5042	MGGHMHEYP	5074	RYDYAKYHP	5106	MSEVMDFLP
5011	DENWNHNMF	5043	MMVCGDRSP	5075	SFCTHDHFP	5107	STAHRPWTP
5012	TSGEMLRML	5044	MPNMWQEWP	5076	SGVPYWAHP	5108	TELVRQVAP
5013	LIEVLNHMN	5045	MQEELVHCP	5077	SSIYEPIYP	5109	TITALMVDP
5014	LFDRCNNMW	5046	MRGIEVEQP	5078	STSCSPKWP	5110	VKLTHRVRP
5015	KFHSKERMN	5047	MYIRREGNP	5079	STTSQHHPP	5111	VQPVQLFCP
5016	YMNSLWHMR	5048	NCDERPHWP	5080	SVTSQWMPP	5112	YTCCENHWP
5017	GIVTCAGMC	5049	NEGPIMYSP	5081	SVVCQQDEP	5113	YPNDMTFEP
5018	GIHICKACP	5050	NEWNPLPKP	5082	TESDFAIGP	5114	AQPSMDYVP
5019	GNWSKVSKP	5051	NILCKHKVP	5083	TMTRIVHEP	5115	FKVRKFWDP
5020	GWENLFMTP	5052	NISGQLHTP	5084	TNSSGIFPP	5116	QTEEPDDYP
5021	HHPTLFHFP	5053	NMSGEYVFP	5085	TSALELQKP	5117	INEGLLYPP
5022	HPDNSQMSP	5054	NQCEIIDFP	5086	TWPRGAMHP		

[2018] 实施例19

[2019] 如通过机器学习鉴定的在CNS中具有组织嗜性的AAV5变体

[2020] 此实施例描述了如通过机器学习方法鉴定的展现CNS组织嗜性的工程化AAV5变体的优选特性。

[2021] 首先，对来自CNS组织的测序数据进行生物信息学预处理。生物信息学预处理步骤的概述如下。首先，将原始NGS配对末端读数合并，进行质量过滤(>Q30avg)和基数校正(fastp)。接下来，去除<150nt的读段(读取长度)。使用变体区域5'和3'上恒定序列的模糊匹配来确保读段的准确对齐(对齐、模糊对齐)。去除带有终止密码子的读段(终止)。使用另外的相关过滤器选择变体用于后续分析(例如，“去重”-显示在给定样本中检测到的独特UMI-var对的数量)。最后，将经过滤的变体移动到本公开的机器学习管道中。图19A显示了NGS数据的生物信息学预处理的各个步骤的结果。显示了通过过滤器的48个单独样本的逐步读取计数。图19B显示了CNS/非CNS变体序列的位置比较。位置1-9是指AAV5 VP1(SEQ ID

NO:2)的581-589位置。

[2022] 考虑到实验设计,开发了另外的过滤器。首先,由于文库的设计使得单个氨基酸残基由单个密码子编码,因此每个变体氨基酸序列应由大致一个核苷酸序列表示。因此,给定核苷酸序列的测序读段的数量应与编码氨基酸序列的读段的数量大致匹配(给定核苷酸序列的测序读段与其编码的氨基酸序列的比率应约等于1)。因此,在一些情况下,消除了此比率＜0.9的DNA变体。此外,由于在文库制备和测序过程中可能出现许多误差,因此在对测序读段进行去重后仅存在一次的DNA变体被消除。

[2023] 最后,由于从组织中分离出的变体序列必须在文库制备过程中进行扩增,因此起始材料中变体核苷酸序列的丰度应与其扩增后的丰度成正比例。因此,与由于PCR或测序误差产生的变体序列相比,起始材料中的变体序列应显示更大的扩增。因此,比较去重前后核苷酸序列的测序读段比率产生了双峰分布,较高比率的组可能代表起始群体中的变体,而较低比率的变体组可能是由PCR/测序误差造成。因此,作为任选的严格过滤器,在一些实施方案中,去除了低比率组变体。

[2024] 图20A显示了通过本文所述的方法分析的周围和CNS组织样本的实例。图20B显示了相较于非CNS变体从CNS回收的前1000个预测的/过滤的变体中的氨基酸位置丰度。机器学习过程的概述如下。经过滤的变体被二进制分类为对于靶组织存在(1)/不存在(0):在靶组织/非靶组织两者中都发现的变体(“共有”变体)被指定为不存在(0)。每个9aa变体序列表示为144个特征向量,每个位置由16个生物物理特性组成(下文表37)。将特征化变体输入到2个独立ML模型中:随机森林(RF)和基于直方图的梯度提升树(HGB)。训练和测试是通过交叉验证完成的,并且对组织存在/不存在的所有预测是在模型训练中未使用的保留数据上进行的(图22)。为每个训练样本分配权重以校正类别不平衡。通过两个模型之间的一致性来评估机器学习性能(图21A-图21B),并以两个模型的平均预测概率作为预测。输出排名靠前的预测变体。提取促成组织靶向性/去靶向性的重要特征(图23和图24A至图24C)作为沙普利加和解释(Shapley Additive Explanations,SHAP)值,表示个别特征对模型输出的贡献。

[2025] 为了适应从组织回收的变体的潜在稀疏采样,我们通过变体序列的特征化将模型集中在生物物理特性上。在该实施方案中,表37中列出的生物物理特性用于为每个变体构建144个特征向量作为ML的输入。在表37中,氨基酸生物物理特征源自以下氨基酸特性。“电荷”是指氨基酸侧链作为酸或碱的静电性能,在中性pH的水性溶剂中具有正电荷或负电荷。“磷酸化”是指氨基酸残基的官能团是否可以具有作为翻译后修饰添加的磷酸基团。“离子键”是指氨基酸残基侧链参与静电相互作用的能力。“氢键(Hydrogen_bond)”是指氨基酸残基侧链参与氢键的能力,“氢键供体(hbond_donor)”是指在中性pH条件下可以将氢原子提供给氢键的氨基酸残基侧链原子的数目,而“氢键受体(h-bond_acceptor)”是指在中性pH条件下可以接受氢键中的供体氢原子的氨基酸残基侧链原子的数目,并且“总潜在氢键(total_potential_h_bond)”是指可以参与氢键的氨基酸残基侧链原子的数目。“摩尔质量(Mol_mass)”是氨基酸残基的预测分子量,单位为道尔顿。“体积”是指给定氨基酸残基在水溶液中的预测体积,改编自Zamyatnin,A.A.,Protein volume in solution,Prog.Biophys.Mol.Biol.,24:107-123(1972)。“亲疏水性”表示给定氨基酸残基侧链的疏水(排斥水)或亲水(吸引水)特性,并且这些值改编自Kyte,J.和Doolittle,R.F.,

J.Mol.Biol.,157:105-132(1982)。“Goldman_engelman_steitz”是疏水性的一种特殊度量,因为它是指α螺旋中的氨基酸残基从非水性条件到水的自由能转移,并且个别氨基酸的值(kcal/残基)改编自Engelman,D.M.、Steitz,T.A.和Goldman A.,Annu.Rev.Biophys.Chem.,15:321-353(1986)。“柔性”是指多肽中氨基酸残基的对称/非对称分布,改编自Bhaskaran,R.&Ponnuswamy,P.R.,Int.J.Peptide and Protein Res.,32:4:241-255(1988)。“可突变性”是指给定氨基酸残基在进化间隔内发生变化的概率,并且是根据一个残基被另一个残基置换的相对频率来计算,在这种情况下,丙氨酸任意设置为“100”,如改编自Dayhoff,M.O.,Schwartz,R.M.,&Orcutt,B.C.Atlas of Protein Sequence and Structure,Vol.5,增刊3(1978)。另见万维网上的参考资料,网址为:imgt.org/IMGTeducation/Aide-memoire/_UK/aminoacids/charge/and imgt.org/IMGTeducation/Aide-memoire/_UK/aminoacids/abbreviation.html#refs,其中每一者以引用方式整体并入本文。

[2026] 表37-用于表征变体氨基酸的生物物理特性

[2027]

氨基酸	类别	电荷	亲疏水性	溶解性	磷酸化	平均柔性指数	离子键	氢键	亲水性	表面可接近性	可突变性	氢键供体	氢键受体	总潜在氢键	摩尔质量	体积	goldman_engelman_steitz
Y	两亲的	0	-1.3	63	1	0.42	0	1	-2.3	1.089	41	1	1	2	181	193.6	-0.7
W	两亲的	0	-0.9	97	0	0.31	0	1	-3.4	0.808	18	1	0	1	204	227.8	1.2
D	带电的	-1	-3.5	-55	0	0.51	1	1	3	1.283	106	0	4	4	133	111.1	-9.2
E	带电的	-1	-3.5	-31	0	0.5	1	1	3	1.445	102	0	4	4	147	138.4	-8.2
R	带电的	1	-4.5	-14	0	0.53	1	1	3	1.475	65	5	0	5	174	173.4	-12.3

K	带电的	1	-3.9	-23	0	0.47	1	1	3	1.545	56	3	0	3	146	168.6	-8.8
V	疏水的	0	4.2	76	0	0.39	0	0	-1.5	0.606	74	0	0	0	117	140	2.6
P	疏水的	0	-1.6	-46	0	0.51	0	0	0	1.236	56	0	0	0	115	112.7	-0.2
M	疏水的	0	1.9	74	0	0.3	0	0	-1.3	0.714	94	0	0	0	149	162.9	3.4
A	疏水的	0	1.8	41	0	0.36	0	0	-0.5	0.815	100	0	0	0	89	88.6	1.6
G	疏水的	0	-0.4	0	0	0.54	0	0	0	0.714	49	0	0	0	75	60.1	1
F	疏水的	0	2.8	100	0	0.31	0	0	-2.5	0.695	41	0	0	0	165	189.9	3.7
I	疏水的	0	4.5	99	0	0.46	0	0	-1.8	0.603	96	0	0	0	131	166.7	3.1
L	疏水的	0	3.8	97	0	0.37	0	0	-1.8	0.603	40	0	0	0	131	166.7	2.8
H	极性的	1	-3.2	8	0	0.32	1	1	-0.5	1.18	66	2	2	4	155	153.2	-3
N	极性的	0	-3.5	-28	0	0.46	0	1	3	1.296	134	2	2	4	132	114.1	-4.8
Q	极性的	0	-3.5	-10	0	0.49	0	1	0.2	1.348	93	2	2	4	146	143.8	-4.1
S	极性的	0	-0.8	-5	1	0.51	0	1	0.3	1.115	120	1	2	3	105	89	0.6
T	极性的	0	-0.7	13	1	0.44	0	1	-0.4	1.184	97	1	2	3	119	116.1	1.2
C	极性的	0	2.5	49	0	0.35	0	0	-1	0.394	20	0	0	0	121	108.5	2

[2028]

[2029] 下文描述了AAV5 VP1衣壳多肽的581至589区域中各位置处的有利生物物理特性和有利氨基酸残基，它们与如通过机器学习预测的较高的CNS组织嗜性概率相关。表38中描述的任何以下机器学习衍生的位置氨基酸偏好可以单独存在或彼此组合存在或与实施例5，表9中观察和描述的任何氨基酸偏好组合存在。

[2030]

表 38 机器学习衍生的 CNS 组织嗜性规则
位置 Xaa1 处低氨基酸溶解性 Xaa1 选自 K、R 或 Q
位置 Xaa1 处低氨基酸亲疏水性 Xaa1 选自 K 或 R
位置 Xaa1 处高平均氨基酸柔性指数 Xaa1 选自 D、E、R、K、G、I、N、Q 或 S
位置 Xaa1 处高氢键供体 Xaa1 选自 K 或 R
位置 Xaa1 处氨基酸可突变性 Xaa1 选自 K、R、P 或 H
位置 Xaa2 处低氨基酸溶解性 Xaa2 选自 R、K、Q 或 S
位置 Xaa2 处低氨基酸亲疏水性 Xaa2 选自 R、K、D、E、N、Q、H、P、Y、W、S 或 T
位置 Xaa2 处高氨基酸电荷 Xaa2 选自 R、K 或 H
位置 Xaa3 处高氨基酸溶解性 Xaa3 选自 A、M、V、W、L 或 I
位置 Xaa5 处高氨基酸溶解性 Xaa5 选自 C、M、V、W、L 或 I
位置 Xaa5 处高亲疏水性 Xaa5 选自 M、V 或 I
位置 Xaa5 处低平均氨基酸柔性指数 Xaa5 选自 M、W、F 或 C
位置 Xaa8 处高氨基酸溶解性 Xaa8 选自 H、V 或 I

[2031] 下文表39提供了表现出CNS组织嗜性并且符合表38中提供的一种或多种规则的变体AAV VP1衣壳多肽的581至589区域的序列。因此，本公开提供了由具有SEQ ID NO:2的工程化AAV5 VP1衣壳多肽组成的rAAV，其中Xaa1至Xaa9区域是SEQ ID NO:8118—SEQ ID NO:9117中的任一者，如表39中所公开。本文还涵盖由工程化AAV5 VP2衣壳多肽和工程化AAV5 VP3衣壳多肽组成的rAAV，所述多肽在对应于AAV5 VP1 581至589区域中的氨基酸的AAV5 VP2(氨基酸残基445至453)和AAV5 VP3(氨基酸残基389-397)中的区域处具有下文表中公开的序列。

[2032] 表39

[2033] 驱动CNS组织嗜性的AAV5 VP1衣壳多肽中581至589区域的序列

[2034]

SEQ ID NO	581-589 序列
8118	HWELVRCHD
8119	SQSRTIIQA
8120	KMAGIGCWD
8121	KSIMMSKLD
8122	ICIRGARTD
8123	CCKSKEACA
8124	RVAGEGIQS
8125	DNYWIIPFA
8126	TERDHKTIY
8127	VKHTAVDVL
8128	RGLSQDCCR
8129	KMEGNMSEL
8130	KNPPWICVW
8131	GVFRPDSFR
8132	AKCTIQNFQ
8133	KTVTKEVTP
8134	RTFYVMGTG
8135	HWEMIRCHD
8136	GKVQWSAQA
8137	VAANSRMMI
8138	KSPCYMSVQ
8139	KSGKVWYVL
8140	PWRARVGAR
8141	RGLTQDCCR
8142	RQYFEDSAG
8143	KSVQVYWHP
8144	IPINDREWG
8145	QIRSLITHA
8146	RMTPHNCQR

[2035]

8147	EACYWGHCS	8188	EMHAFWQRG	8229	QYWWPTTHD	8270	IANKSNCHK
8148	MMAVIRQKS	8189	FCKCFVKHC	8230	KTNFGDAIE	8271	VGRFNAYGE
8149	LSAKQRFFC	8190	RKKVEVVVA	8231	YIYMMLNLH	8272	GLWLNVTPR
8150	KFNTKYQLM	8191	NTRTGNDRL	8232	QWFAKGVGE	8273	NTHTGNARL
8151	QNKHNWWIT	8192	LADCMINHH	8233	QLMSCANVE	8274	MVKPCLDEW
8152	SRLPVHKCP	8193	LDTKTLASS	8234	DIEFCMKQY	8275	QIKAHSNVF
8153	IRNRYAQAP	8194	IVDSKVICV	8235	DYEKFVKWM	8276	LVDRSGRAA
8154	KVNCRMTIM	8195	RTFMKDMCE	8236	FCNYAAMSP	8277	TSKGKPTCC
8155	KSYLIANRH	8196	DAHSPWQVI	8237	CYWSMTNCT	8278	NCYCAYGWP
8156	KTTFDVTCY	8197	TRHPTITMH	8238	KWPVTANLH	8279	IWQSLKKAT
8157	MEDRVFACC	8198	SSRHRGGCQ	8239	DTRHKWCLE	8280	ARSGINGHA
8158	NLNNLVSCR	8199	RCTDHGESI	8240	FCVKSPRVD	8281	NWTRNMQWA
8159	DAIWEHIRL	8200	SSMCSGSLL	8241	DAIQWGSWQ	8282	NCLSRKKFA
8160	KWAIVGYAL	8201	GKLTFDTER	8242	SALYRRGHV	8283	CIMLNGCFP
8161	KFGFMYSEF	8202	RKDQSDLWS	8243	FMTVRNWAG	8284	KNGATNLAW
8162	KSTMLEHHA	8203	RSMWSAFKH	8244	KISTTWADT	8285	NNCKLAMVH
8163	GRWMCNEAA	8204	FRLDYNYKN	8245	MHNAHVIIE	8286	KATHCDKYA
8164	KDHQLWGVW	8205	RKKVEVVGE	8246	MTGCRETVD	8287	VFNPMLQDT
8165	ASSFCKFTS	8206	GYNYYWQHV	8247	QYKWHDVGV	8288	ARPEIEYMG
8166	PKRIRYRTT	8207	TKLMYNTVD	8248	VACDMCLVQ	8289	NSVHVHSSN
8167	TKITMPGYQ	8208	KTLHMPESK	8249	KSEWTCCPH	8290	SCPWKCMIV
8168	TFHQSDYWA	8209	GRPVWCVCM	8250	ARKPCNGWD	8291	CSNDINGRT
8169	GKRVGDTSC	8210	MVHRVQGDD	8251	TRYPTIAMH	8292	KMGMHDMAT
8170	FAQRDKQDR	8211	ATRYQICGE	8252	TFMFNSKWS	8293	KKLVNEGTW
8171	KTYSAMHSP	8212	DIRWLPYDL	8253	RAMAYHIYW	8294	QLNQWTFAG
8172	TRAYYDSYS	8213	HMHNLSVKP	8254	LYNMRMKDR	8295	FERRIGWDH
8173	RMTPNNFQR	8214	THRTMDDYF	8255	DIWWTLTRD	8296	VWNVMIKMC
8174	KWGNEGRMS	8215	NYFTFWVFI	8256	YHVHMMPPL	8297	TGGCYRMME
8175	TVRHRMDAY	8216	HGCKCQYSF	8257	KSGSCAMRV	8298	KPDPRQADQ
8176	RQQRNHASS	8217	VSANEDRAG	8258	TRWTEKRQA	8299	LSCKFGWDG
8177	TISQIPLVR	8218	YPVNMMPPL	8259	QLMPCTNVE	8300	KYPFKYYVT
8178	FNWLVRAYI	8219	GFRRTHVDQ	8260	MGAMKSDDF	8301	TMLQKKRQD
8179	MYTGRRCWR	8220	CWMLHHRCS	8261	AMKSHGCLT	8302	GYAQRSFFM
8180	QCHRGYIYE	8221	CVNLMPVNN	8262	QNGPILTSR	8303	ATDGYVAQD
8181	NYFAFWVFI	8222	TMRAGLENS	8263	YCCNQEPFT	8304	KMSTYHTYD
8182	AIDCCYSCG	8223	QGFSTDAKP	8264	NWCTMYRYC	8305	HVADNQVVQ
8183	KIYMQCDRE	8224	QIVSCDTET	8265	KDPDFYDWP	8306	AKTFVMSFD
8184	DRKQAGQIY	8225	QKGMMLDYA	8266	LWIKQHVKC	8307	TCNVSWEHE
8185	RDVGEENSA	8226	DSHRCLKAA	8267	TRLEIEYMG	8308	RRTDHCEPI
8186	IAKKNFMKT	8227	NTSEMLMAQ	8268	DMTHTGIMS	8309	RTDYSRLSD
8187	RKAYLNYVA	8228	GKIARDQLN	8269	SMHMRVCYT	8310	TDLRKRSNL

[2036]

8311	IKAWWHDHQ	8352	AHLHNQVIF	8393	QSVNEHHFP	8434	EWLRQPRMT
8312	KFKWEACCP	8353	KATVEMQSL	8394	RAAKYQNSH	8435	NQQRTNNNG
8313	VNVDCDVQQ	8354	CVRIGMRGP	8395	DYERADCWM	8436	QPRTVGVAG
8314	QYQSTGVVR	8355	HLSFMMSSF	8396	IRNYCRMTA	8437	ASWPCGRNV
8315	LIHEERKEQ	8356	TTMWIWFPF	8397	LAVWQQKME	8438	LNSMIMWDN
8316	ETNVVEDWY	8357	TTNMKWCAG	8398	INMSVSKDQ	8439	ESRSPNNPA
8317	LACTPRYIF	8358	FIQPTWDSY	8399	ACDKWDCCC	8440	PTSWLNDLT
8318	HGGCYRMME	8359	RCIWCHMQG	8400	SQPVCCVCM	8441	MIRKRFEGQ
8319	QSVNERHFH	8360	LVHRVQGGD	8401	GRALANCAE	8442	AMWAGLENS
8320	LWIKQHIKG	8361	NVRSPAVRD	8402	ITTHMNCAD	8443	KLTDIWAHM
8321	QLMPCANVA	8362	LKENWMEWG	8403	MKTTMQDYF	8444	TSYNCGRDP
8322	TQREGETSY	8363	YSASFGCQS	8404	HCIKVESVG	8445	HCVKVESVG
8323	HNDWQQVFS	8364	GLKYWMHGI	8405	HKGTHAFAM	8446	TACGSFHGC
8324	YEDNDVMQI	8365	QTGTGHGFE	8406	HYTEMSCDD	8447	MSVPKRQGP
8325	ARRAGLENS	8366	QMNYQHEHR	8407	DFWWTLTRD	8448	LKQAKQMEP
8326	GATWLFRSP	8367	AHLHNKFIF	8408	SMQYFFKTA	8449	DCDVWHANV
8327	RTLWALYAI	8368	STKQMTKSF	8409	LNCPSKDGT	8450	MGDWLGVPT
8328	DCIICSSKG	8369	QEENGKSVT	8410	CTHMKNVRT	8451	FCKHNAQTQ
8329	HMALVMQTV	8370	RVADEGIQP	8411	ANADNSFCQ	8452	QTKPMIGNE
8330	HSNMLSVHQ	8371	CGLSTHEYL	8412	EERPWEWET	8453	LSAQFNNQS
8331	TNGAQIMHV	8372	RANTESDQP	8413	CSHGTGAMC	8454	PVLLMRRMF
8332	STRPIACQS	8373	QSAEWYLNG	8414	FAIWDYESK	8455	ISTICPKIC
8333	CMYVPADQE	8374	CWQLFNEQN	8415	HLRFIMHEP	8456	TGSFMGCVD
8334	SPHYFADQW	8375	AADGYVTQD	8416	HLMPRSQCF	8457	MTGCREAVA
8335	GFDLMRQIN	8376	SANQRHHDW	8417	QLMVAFLCP	8458	ERGLDNKPR
8336	GHHQCYKAN	8377	KCVEWNAPA	8418	AHSPFLDHM	8459	NYCINKRAV
8337	EMNSDYSGI	8378	MHDTMGAWC	8419	KQAGYDLFH	8460	SGQPVDRMF
8338	SYWHVLLLV	8379	GDNMVQWGS	8420	DTMQWGSWQ	8461	QIMSAKVFG
8339	SADRTVMFH	8380	DVWCRKLMP	8421	SGFFVHTSS	8462	MFVVNRNWA
8340	RFTVKDGDA	8381	SYRTLFNGW	8422	AKIIAQLCV	8463	WPDPESEEV
8341	QCKPGAEGF	8382	MRQHMKYSL	8423	NWTGNMQWE	8464	MNTMLMSNS
8342	KAIQAFEHD	8383	PDQYLVVGD	8424	QREGNNYLT	8465	HFVDNQVVQ
8343	ATRYQICDE	8384	RELDKENCV	8425	NAFWTQNME	8466	NTRQRYHTW
8344	MWWREGAMR	8385	MTGQCRNGQ	8426	IKQVFNQIV	8467	TSLVMPSLP
8345	LFMTYSTRN	8386	TSLWQQMAS	8427	HAVSGPTTG	8468	DYRQCIDWE
8346	QPPNREAWC	8387	HSMFRHPCP	8428	NTSGLLMAQ	8469	VLRYKCYPS
8347	TPKSAGICC	8388	GNTKCPSWH	8429	MYLSFPVHA	8470	HWEIEWVYW
8348	KPTYEHRGF	8389	HECDTVAVN	8430	STLSRHGWG	8471	QVMMGQQTP
8349	CNYPIIWDR	8390	ERTGEKGIR	8431	QWMWQRLID	8472	RGCQVKCHT
8350	FTGTPPVGM	8391	GSGPVQSEE	8432	TQGTMDSWY	8473	EQWKTHQQN
8351	QGQTGKYVM	8392	DGWCQKLMP	8433	RPIHTRQNY	8474	YSKWARALS

[2037]

8475	QCMITISCP	8516	HAETDHAGL	8557	HAVNSDCKK	8598	SPSQKGYGP
8476	DVWWQKLMP	8517	MKMHQLNGC	8558	NIRPCVNQE	8599	QSHGGVVVQ
8477	VANTWQVHC	8518	SSERFRAEG	8559	QMECRGGAM	8600	WIQLAILVC
8478	RMMWQSHCG	8519	QGGCAPMDS	8560	VGVQCPNGG	8601	NTHTGNYRL
8479	VANITACAN	8520	REVWITGCG	8561	MQGMQTFWT	8602	EACYLGHCH
8480	PWAVEYSLV	8521	SDNGWCAHS	8562	QTMGIGGGP	8603	DEYQQESFP
8481	SYMWPHFPH	8522	YCCNQELFA	8563	RSQQDRVMA	8604	YRQQMFERQ
8482	FFEQEAHNR	8523	TFVYGEWDR	8564	QQGMCVTAI	8605	AQQSTMISL
8483	LCTGTERQG	8524	IFPSKIPCQ	8565	AIGPRLIDD	8606	AHQTPWAWI
8484	MAVGTVSAK	8525	YVTQMQHLE	8566	HGINVELGS	8607	SISRADLQG
8485	QSSSPVMWV	8526	NSRLLIKHP	8567	NVDGRHPMI	8608	LNVMSIHRC
8486	SSVVMVKQP	8527	KTFPAADMK	8568	DAARERHWE	8609	AGRFHKHYA
8487	EPRLYCAQA	8528	VGHVMIKMC	8569	QTGTGHWVE	8610	GEMTRCFRN
8488	YANIHNRHG	8529	GTDHVIVGY	8570	TCGQCVNGV	8611	QVKWTANPD
8489	EAYIYLGEA	8530	GAQSMQSSC	8571	EGSMFAERQ	8612	KTLSPHVFA
8490	PVETARCSG	8531	YAVQTNNIM	8572	HSCKYLMRV	8613	MGRPGACQA
8491	NAIPYFIQA	8532	WGGMSNNQG	8573	MCDWLGAPT	8614	FMQPYGHGE
8492	FCKCFMKHF	8533	QMVTQLNWG	8574	EVSMFAERP	8615	DCWCYTTKC
8493	EKLVMYHKC	8534	VPKQGAKQE	8575	ESAITSCAW	8616	LIRRHCHYP
8494	YEDCMDMYF	8535	DEYQQESFR	8576	RVCEMRTVD	8617	QMGVTVVQR
8495	GIWVWDTYN	8536	GGTRSGKVV	8577	AHQGCRPGQ	8618	VCYNQEKAA
8496	AEWDHKTIY	8537	CQNTAIRAI	8578	ANMISPYGL	8619	EMNSDRSGI
8497	TQIQIRQAD	8538	TIHPPWVTL	8579	EKHRCAKDR	8620	HLWMCPPRY
8498	RCNWQFAQH	8539	CMVSTPKLA	8580	TQQRSQATA	8621	TASLTDVIW
8499	ESRHCKYPY	8540	RIWWALYAI	8581	IPMSREFPE	8622	LETTMRHGE
8500	SSNCMRRAS	8541	SNTTDVFTW	8582	GNQLITMSG	8623	FIGFNNAEA
8501	GVGQFRLLS	8542	WTSSKAVYD	8583	YEMTSKCAY	8624	QQQRTVFGM
8502	KIDQTQWRR	8543	ESSFQDNVI	8584	TIGQIPWVW	8625	CYNQFHTYN
8503	HLSFLVSSF	8544	YAHAFHCQP	8585	HHGASLGKY	8626	LTISHPKSQ
8504	NAGMSFRNR	8545	RRPPKRGTT	8586	QMMPQMDCS	8627	AVNRIQQVH
8505	VMNRMEWSV	8546	QMYATMDTY	8587	DSWQYYNGL	8628	GWFMIQSMK
8506	LFMGMSDST	8547	EGGQCEKCK	8588	VTYYGHYSV	8629	WCNWKFAQH
8507	ANVVMRKCQ	8548	MGDWLGTPT	8589	TTVHMPKHE	8630	MSYLVKGVG
8508	GIYYQWRGQ	8549	GVKYATLHD	8590	QSHHFLNSA	8631	NTSRELIMP
8509	DVHAEHSPQ	8550	PWEIEWGYW	8591	TNTMIWNDC	8632	TVNPCVNGE
8510	QHPMTFSHE	8551	MYLLFPVHE	8592	KAIQASERD	8633	QMGHQLCVQ
8511	EPRLYCTQE	8552	HVLNGTGGP	8593	CRSTGMMRQ	8634	DERHKWCLE
8512	NAKEQVVNT	8553	NIGQSPTQL	8594	TSSLTDVIW	8635	MWHELITDQ
8513	EFCHDDILM	8554	VSCMWCQEN	8595	TQAKIECWS	8636	LTRSGPVGQ
8514	PLFRSEWLA	8555	TMSVHYGQV	8596	KYAHSSCED	8637	AQFHIHRWQ
8515	GNMLRNFGS	8556	NAAMITTVC	8597	NYENNFRAQ	8638	VCTPPLNMQ

[2038]

8639	SWSQLQHTA	8680	QYYWFAQFL	8721	EAYATVHMC	8762	VADWIGRTP
8640	VATYIHAMH	8681	LKVDHKIGD	8722	LMHTYVVQQ	8763	DLYIVMNQA
8641	STPQGKQRQ	8682	DSPAFHQSG	8723	DACAAHDHW	8764	NPIPGYFEQ
8642	TTTITNRLR	8683	WSCRSICPP	8724	DIMMKHCRR	8765	SCKWIQSNR
8643	SSLCRTYDN	8684	QHATTCTHP	8725	LTQYKALVQ	8766	QWGVISGHA
8644	MLEKRNHWG	8685	QYQHAVWAP	8726	MDKHKIGYA	8767	ESEFGHWWG
8645	NSELLRYTF	8686	LVCSNVDDR	8727	ESDWLLGTI	8768	KKAGAKDAL
8646	CGVGLTQCK	8687	GKEHMPHHM	8728	NVGYYKWQS	8769	CTCEPRSAW
8647	TGQQIHYKG	8688	SMSDCEVSF	8729	IECFPRESI	8770	YLYHSMMKQ
8648	DLEPIMTQA	8689	DHIWLHWME	8730	LSMPYAPTT	8771	ILRTIPNRT
8649	PIAQDKVWS	8690	WSTLYLMTS	8731	AVKILMHDH	8772	IEESIQNAD
8650	YSTIQGSGP	8691	MANDQRVDY	8732	THELHGHCS	8773	AMMPKTETH
8651	ATTTIQGKA	8692	SYWDVLLLV	8733	NSQTIMLHQ	8774	VVTLHSKAR
8652	ISICAQDIG	8693	KMGTLHNME	8734	HIHMSDRGT	8775	VCCSHADSA
8653	STNPNLTGD	8694	DHHRTSAVR	8735	TGRAWMWHS	8776	DGSSHFTMQ
8654	THELHGHCV	8695	MNTPCQQAP	8736	NSHAHGHMK	8777	VAYPIQTDA
8655	QEMMIVNCD	8696	NSTHNTSYY	8737	DYTIMPYIA	8778	YAQQSPHCE
8656	EADTFTLIV	8697	AVSMFAERQ	8738	SVTYHERHP	8779	YMLHAIMES
8657	GMNAYPAHS	8698	STDLSNHFQ	8739	CTAMGWVHN	8780	AAVITINGL
8658	NCRQLGDNM	8699	DSIIKDQSA	8740	SVTYHKWHP	8781	LNMVWGQKF
8659	ARCSMLDHG	8700	FCGKGANQI	8741	VSCEFGQEY	8782	NALPAQQHH
8660	ASTWMQSYS	8701	TQRSTQWSK	8742	KDLEYQRSN	8783	QKPTIWDHF
8661	HYCSCHNKY	8702	CHYQINKFA	8743	AKGPHWGQG	8784	YQDIRQCIF
8662	VSNVTIGCC	8703	QATPACPAV	8744	RYHHSCDSH	8785	TCTHFEFVD
8663	DNAFMHHHV	8704	DVFWPDSFR	8745	ACTWLHGTN	8786	TLEPAYWHE
8664	SKVHWDLPC	8705	FVIKPNQAT	8746	SIHNVLVQN	8787	PVHFWTQSK
8665	NPQHQIDIG	8706	CRKSKEVCA	8747	KVDQNTNWC	8788	RCKRAVQVV
8666	MWWCESAMR	8707	ASGFGHWWG	8748	FGEFTTDEN	8789	WCESWGPQG
8667	FIQCCQDTP	8708	RKRTAGVHF	8749	VWMRFDYVD	8790	AVVYKMNYH
8668	RMVCDGRID	8709	TYMYRSDCS	8750	MAKKREVGV	8791	KNHMMLYVN
8669	ELPRYYGQV	8710	MSVKKECGG	8751	MDEHFSVGC	8792	ANMCGCDSF
8670	VTAFTTCHY	8711	WHFLMTFYI	8752	EFRCRLIDS	8793	LFSDHGVSY
8671	TCSSTEGSG	8712	IADDYFMYH	8753	NHCHCIAHP	8794	TWSCCLHQY
8672	EWRTISYGN	8713	RIHDTMSDY	8754	IIMLGCTGR	8795	WTDSDSDCT
8673	QSTPTLGPK	8714	CNMPQQCME	8755	VTLTSKMCQ	8796	VIDCCYSCD
8674	GWALIVNDL	8715	TVGMRQAMM	8756	RAECRYQTT	8797	PIAQDQVLS
8675	VTILGLNPT	8716	KKIFCSRDY	8757	FDVHKMCIY	8798	LCRWYRYAN
8676	VTCHWHQGG	8717	NHDEKMWAP	8758	VSADTPYMN	8799	QNIRPGVPT
8677	QGSVVHSPS	8718	LCAVCGGCD	8759	THRTRHTDC	8800	RTTDHRADE
8678	NQARENWFH	8719	VHPQTHVTD	8760	DYYLNRASP	8801	QIHSEHCTT
8679	HIKVCTLGD	8720	QHEQSNRIW	8761	SVSTKPHAE	8802	QSHDGVMVQ

[2039]

8803	DLLANTICT	8844	HAETHHVGL	8885	EQAGAASMW	8926	LNQCAQAGC
8804	SANIVFSVC	8845	ATVDVKGSH	8886	VQKLTLNAK	8927	TATCGIEAE
8805	LSMTRELWH	8846	SSDGESHYI	8887	GHNHMYTQP	8928	TNLHNNHYW
8806	MLHKYLHMW	8847	RTLLPNANS	8888	VSKTPWERD	8929	AYSYYSVLP
8807	QYAITYQKG	8848	ITCNWCVMP	8889	QSCHVMNMP	8930	STMLLLARD
8808	VQMYWSSKD	8849	GSLLQGVSD	8890	IVLWFVNTT	8931	VAYYSGHMM
8809	GFNLMHMAG	8850	FCIMASDMA	8891	WCQHIPHNG	8932	IQQLSMVGS
8810	LFYSPMHTH	8851	QGAFQMMLP	8892	CIAREYTGY	8933	ARVKETCEW
8811	VKAKKMANP	8852	FMWTTHRWH	8893	YFVVNHFWS	8934	TFMFNSEGS
8812	NCMPTIAGS	8853	MFAQQMGVD	8894	STLEDEKPY	8935	CVPKTIACC
8813	SHWQMQVMA	8854	ALVMIILQG	8895	ARCHDMWGH	8936	ACYEFVNAC
8814	HIMQEGRSS	8855	ATISTESEV	8896	MAAIKWYMS	8937	TRMQFNKCH
8815	HFAMKCHNN	8856	LTCNWTHHG	8897	WMLHGWATL	8938	SRDIAKSST
8816	EIIQAHLVN	8857	TFKEYQGYH	8898	MVDCMKWHE	8939	YESCNLTRD
8817	VKNALCQHM	8858	IRQPIHMTP	8899	AGMYSFSGE	8940	STQPACAWP
8818	KTWDAHRSS	8859	QTAMWCWQE	8900	VALQTKIEQ	8941	MVGRNMWAV
8819	IRGFMQNIY	8860	HIEVCTLGA	8901	GTQNLLHGY	8942	CQMQVLNNI
8820	VSYAPMDAQ	8861	QTHLKLECE	8902	MTQYKALAQ	8943	YLVHEGHSR
8821	SGSMTQHHT	8862	GRNEQGYCH	8903	GCATYSCYV	8944	TGTQFMRWA
8822	AICTTASFD	8863	TFMQNTYQG	8904	QGDMGTWVG	8945	MCMQEDRLA
8823	YPTECNHLT	8864	LTNYHIAQE	8905	SAVVSMRND	8946	LATNHRHWL
8824	QSQHLAYKG	8865	ATTEKCRSI	8906	YPYGKCSNY	8947	GRWPYFLAC
8825	LKAQAVTVT	8866	ENAQMKCLM	8907	RVINTEWGR	8948	IPSFYWGDY
8826	CTVQFELGC	8867	HKLWYWAFF	8908	TTYHHGNTL	8949	GDDWIIWQA
8827	GTSQLQHMA	8868	AHAFRHKHF	8909	QQTTYDCLD	8950	MECIHMRND
8828	PIYCNALCS	8869	KFICTTITC	8910	LTHLQHSAT	8951	SIWDCNHRS
8829	VSKPDMHVS	8870	QPGIVICKK	8911	YMPLAMQYC	8952	NTVFQLYHT
8830	GPLVCISAG	8871	IMGYKCKWL	8912	TREPCQQCD	8953	GYGHTFWDA
8831	HKVKFGALH	8872	QPGYMNKLW	8913	IQDLQIAIK	8954	VHEEYSNSH
8832	TQGWLAAGH	8873	QCDAALNWS	8914	ATTNHDMQQ	8955	GTLACDQKE
8833	GPRLIVHKL	8874	ISTVCPKIF	8915	FCNKLTSVW	8956	FCCMYCQHP
8834	MFGFSYSDG	8875	NDQHAPCTE	8916	EAVCMDTFP	8957	ESSMWCTDC
8835	TMGMANCGM	8876	ASSHRHWVY	8917	SAALVADGP	8958	QTADWCHPV
8836	QSAIMSART	8877	AMMLFQTSY	8918	QPEIFPVMA	8959	MPHGRCNTI
8837	VWGMAYTAC	8878	SKEDLCHQQ	8919	HTESGQDET	8960	GITQHFVKH
8838	DATFMKNAF	8879	AKIHGAWNA	8920	TRNTWQQDF	8961	AMDCGQSFG
8839	TEGPMHRVW	8880	ASAPRVNEQ	8921	GVNTNMNGS	8962	LAYSQYVGC
8840	FFRICDSYI	8881	SCPWKCPIV	8922	CQAELIDVR	8963	NRNCTDAWH
8841	CTALVNEDL	8882	PNRYRPFCC	8923	TSTMCSWSM	8964	RSHFQEGVQ
8842	SQAHQNRAC	8883	DHLRDHPLY	8924	NAEPQQADI	8965	MFMAKQNAM
8843	AMIGHGAYA	8884	GNGTVGKAN	8925	IRQMAGGHC	8966	PIWSVYQDE

[2040]

8967	DKRIICQDN	9005	FCHAFPDAY	9043	GLWHGFYAV	9081	FWMKWESYG
8968	MFYAMVGPE	9006	VIGWHGTNA	9044	GGCCGGELN	9082	AMRDNSVAH
8969	ETCCSEWHV	9007	TQGVSTIGV	9045	LQYSPGQQT	9083	CWSYTNKNC
8970	MCKNGDSHG	9008	NDEQRNTLY	9046	TNNCYWCAP	9084	QANPYNHGI
8971	AIHAHTAAN	9009	TSQHKYAEE	9047	AVLGSNYLK	9085	IHVKFPPFD
8972	NVSSHQHCW	9010	HDCIWMECR	9048	MDTKVVYGT	9086	EYIQHSSAR
8973	CRCCWYQDQ	9011	DFAYRSTGV	9049	WADRQGALA	9087	TTTTPCNMP
8974	DMSDCEVSF	9012	DSATSKSAD	9050	VITLKNNCH	9088	CCTGLGSMG
8975	SMCWFNQVA	9013	IPASCGSTR	9051	AIAEIKIQS	9089	TVTHMSTWG
8976	AAPMVGEFY	9014	SMGWDHMMS	9052	VTMPMATWH	9090	NNMLHGAYR
8977	SPMSFLNVL	9015	MASTESDMC	9053	VHFIAWRGM	9091	WGQHLPMID
8978	TKGPKELMA	9016	QMEVHWHMR	9054	VIGVEMKCM	9092	MMKMCEYDG
8979	EGWTQFRPE	9017	AQWISPWPT	9055	TDAGMFAHA	9093	GANNMTAMN
8980	QAGDSMWAK	9018	QENDANRPQ	9056	ITKFKMQDH	9094	THAAYGTNL
8981	SADTNVVVE	9019	LVNQLFHSW	9057	AMCIVCIAP	9095	LSSLPMHYC
8982	LKSGQVRFT	9020	ACDQVNTFQ	9058	KSGYQCDSM	9096	ITEQYDAFV
8983	YIHDYSSPN	9021	MKSVFMHEW	9059	LALMIQQAH	9097	YTGIKVMGR
8984	QTEIAMGFT	9022	WGIKCLGTS	9060	LSAQDTMFH	9098	DVEHVSCPK
8985	QLCIPLTWE	9023	VSRSDWTMP	9061	YQAQLNNCW	9099	WHCEFTQHH
8986	EANCNHVGV	9024	TSTIPGVGT	9062	SCEATMHAI	9100	LSTQSAKEY
8987	SGMKGRTQI	9025	YHTFVECAQ	9063	TILGIPSSK	9101	ISDPNVGEP
8988	AEAIINDIW	9026	ASCYVWGPA	9064	DMTHTDIMP	9102	SVCSKWDGT
8989	LTEFTPYIG	9027	QHQSVMGFA	9065	LIGPNTKWG	9103	NNMWDTIFM
8990	FCAHNFTSD	9028	HMQAQIHYW	9066	EARMPWAHH	9104	NHKMVRDAL
8991	SFLMMMNCT	9029	KQHELIVRV	9067	HVHMCTPSV	9105	HVGIYEHCH
8992	LCMPADRQS	9030	AYCCFNRFQ	9068	LNTYYRLED	9106	SPKMMVSMT
8993	MTNSQHQHW	9031	TYMQCMVSG	9069	INCQTKKEM	9107	WMVRQKYMA
8994	PKHHVTWED	9032	CAVNTYQKH	9070	VPEPKGTLN	9108	LQWWIRMGK
8995	VSQPAWQQK	9033	RNINEYEHY	9071	STHMAGHYA	9109	CCGIRIEYG
8996	QAQSHIHYT	9034	SARVNHGRV	9072	APGMAFGGQ	9110	FAGFTPGWM
8997	EVLMAMSLF	9035	HSHMYDCQQ	9073	SVSAQQTMQ	9111	PHKDSMQMT
8998	GQALDKQYQ	9036	NTECWQNSD	9074	RTVYDVQDD	9112	DKTVQSLSL
8999	SNEPLDTCY	9037	LNCQKANQN	9075	GCYGNQPPT	9113	PWACEHVWH
9000	SQAVRKCCD	9038	TVMMLMDNG	9076	TATWGIEVE	9114	EYCEFMSEA
9001	QGSPLDSFQ	9039	MTTGLMAYE	9077	ECATNVGDR	9115	YGVMCPSHY
9002	LLFPLMGSS	9040	LLYRNMILS	9078	WDQKNLMDT	9116	PIIRGRSDS
9003	PNSPGVCSC	9041	QSHLSDKMY	9079	ERTLTKTGS	9117	EPMEANQCR
9004	MSTHVLKWV	9042	TMAMLGWMN	9080	IVNMNHSAM		

[2041] 实施例20

[2042] 如通过机器学习鉴定的在肝脏中具有组织嗜性的工程化AAV5变体

[2043] 此实施例描述了如通过机器学习方法鉴定的展现肝脏组织嗜性的工程化AAV5变体的优选特性。进行如实施例19和图19至图22中描述的相同生物信息学预处理、机器学习方法和分析以确定展现肝脏组织嗜性的工程化AAV5变体的优选特性。图25显示了一组对模型输出概率有贡献的前20个位置特征。沙普利加和解释(SHAP)值可用于询问特征对模型预测的相对贡献。如图26A至图26C所示，在组织靶向性和非靶向性变体之间进一步比较了这

些特征。图26A至图26C显示了排名靠前的预测位置特征的比较。如果发现特征对HGB和RF模型都重要，则选择这些特征。显示了相较于随机2％肝脏变体，前1000个ML预测的肝脏变体中的特征的概述。

[2044] 下文描述了AAV5 VP衣壳多肽的581至589区域中各位置处的有利生物物理特性和有利氨基酸残基，它们与较高的肝脏组织嗜性概率相关。表40中描述的任何以下机器学习衍生的位置氨基酸偏好可以单独存在或彼此组合存在或与实施例4，表6中观察和描述的任何氨基酸偏好组合存在。下文表40中列出了从两个ML模型(HGB和RF)中提取的前30个重要的肝脏嗜性特征之间共有的位置特征的概述。

[2045]

表 40
机器学习衍生的肝脏组织嗜性规则

位置 Xaa1 处高表面可接近性 Xaa1 选自 K、R 或 E
位置 Xaa1 处低亲疏水性(<-3.5) Xaa1 选自 K 或 R
位置 Xaa1 处低氨基酸可突变性 Xaa1 选自 H、P、K 或 R
位置 Xaa1 处低氨基酸溶解性 Xaa1 选自 Q、K 或 R
位置 Xaa2 处高表面可接近性 Xaa2 选自 E、R 或 K
位置 Xaa2 处低亲疏水性 Xaa2 选自 K 或 R
位置 Xaa2 处高氨基酸体积 Xaa2 选自 S、L、I、A、R 或 K

[2046]

表 40 机器学习衍生的肝脏组织嗜性规则
位置 Xaa3 处高可突变性 Xaa3 选自 N、I、A、M、E 或 D
位置 Xaa3 处低溶解性 Xaa3 选自 N、K、R 或 E
位置 Xaa4 处低亲疏水性 Xaa4 选自 K 或 R
位置 Xaa4 处高氨基酸体积 Xaa4 选自 K、R、I 或 L
位置 Xaa5 处中等氨基酸溶解性 Xaa5 选自 H 或 T
位置 Xaa8 处低表面可接近性 Xaa8 选自 V 或 C
位置 Xaa8 处低平均柔性指数 Xaa8 选自 W、V、M、A、F、L、H 或 C

[2047] 下文表41提供了表现出肝脏组织嗜性并且符合表40中提供的一种或多种规则的变体AAV VP1衣壳多肽的581至589区域的序列。因此，本公开提供了由具有SEQ ID NO:2的工程化AAV5 VP1衣壳多肽组成的rAAV，其中Xaa1至Xaa9区域是SEQ ID NO:44438-SEQ ID NO:45437中的任一者，如表41中所公开。本文还涵盖由工程化AAV5 VP2衣壳多肽和工程化AAV5 VP3衣壳多肽组成的rAAV，所述多肽在对应于AAV5 VP1 581至589区域中的氨基酸的AAV5VP2(氨基酸残基445至453)和AAV5 VP3(氨基酸残基389-397)中的区域处具有下文表中公开的序列。

[2048] 表41

[2049] 驱动肝脏组织嗜性的AAV5 VP1衣壳多肽中581至589区域的序列

[2050]

SEQ ID NO	581-589 序列
44438	KAVAGCDFW
44439	KVQSAVEVT
44440	TTKPKQSCM
44441	QATKLQDVA
44442	RARHTEECA
44443	TSHKQSDCT
44444	KSGHTWDMS
44445	RVVTTSNCT
44446	VAFPMKWVC
44447	KDTFTKQWQ
44448	KAQSDCDFY
44449	KAVSSLDGV
44450	LVGRCYDLH
44451	KDGKCVSCG
44452	KCEVKQCDH
44453	SQVIRLACY
44454	ACNVCAHIR
44455	TTYKEMVSG
44456	AAMKLGNLA
44457	GSNRHMIYA
44458	AISKFIEGM
44459	ECGRLYEHE
44460	KADQFSRDR
44461	AAYKAQVLR
44462	ASWLDMCMR
44463	KYEHTTSPG
44464	QGLKCEQFE
44465	TKNTIVKAE
44466	RSNLLLQHM
44467	EKRMRAECE
44468	AIKMTNRCR
44469	TMAKYNEME
44470	RAERWGGCV
44471	NAKLEHAVN
44472	IMTKVHHVT
44473	QKRHCQTCT
44474	TKRDCNTQI
44475	RGEQKVYCS
44476	SRRLTTAHY
44477	SKKLTQQHM
44478	MGFRTSQYV
44479	AASRGNDCN
44480	KYETENGIN
44481	KAKLETWWD
44482	GMTKNTMGH
44483	ARSFVEKCR
44484	FGKYFWNSF
44485	VMVKWMRCN
44486	IVHKCIDIG
44487	VRQMNLLCR
44488	KAQYGFAVP
44489	KSEFVQSNF
44490	FTKIHWWFT
44491	TKMQVTSST
44492	MRATSHIVS
44493	KSTFAVLWA
44494	GATFCWNCG
44495	VKTQYCDMD
44496	KVGPAADFA
44497	AASFHAMDR
44498	EKVMRKPAD
44499	KVTIGTLAN
44500	AKMNYANVD
44501	SKQVCLAHE
44502	VAKMSFDCC
44503	RPSYKVKCQ
44504	ATHKIIGVE
44505	KSMHTVNEP
44506	VCARTNNLC
44507	TKVIYEKAV
44508	RSGVTYAWA
44509	RTQTFTMIP
44510	KCEMNISCH
44511	KPLQAFAFN
44512	AELRYQFYM
44513	DKRVIVKCE
44514	SNCLLGQCT
44515	QKQGFAINQ
44516	SKLLYSTSG
44517	KITFCWPCH
44518	RMQYTQSCS
44519	GTRTSFAMD
44520	VVKQTTGMD
44521	SATRAFEVQ
44522	SNGLQNIAF
44523	KMQSRLYCY
44524	TIKPSWLCA
44525	VSVRMHEMH
44526	KAMMDHAYD
44527	LASRAENHS
44528	KMQVDNANE
44529	MCMKAQRSG
44530	TAMTDVRDR
44531	RINRNNCRI
44532	MAQLLYSCQ
44533	QCRRMEGVE
44534	AMQRALINA
44535	DRKTQMIST
44536	NGFIQYDSK
44537	SKNAANSWD
44538	QKKFMECAQ
44539	SLFLEKSCG
44540	KVLECHGAA
44541	KIESTSSHC
44542	KHQFTGLHS
44543	QRKMNPDCM
44544	KMSTCMEHN
44545	AIGRNDQTT
44546	NAVVMMSNK
44547	RNIQGISWA
44548	AKSPITNGM
44549	SSNRYGDVT
44550	MRSGFQMST
44551	IAAEFKACR
44552	KLNSTMYVG
44553	LAVRTKQFH
44554	KQTSPLHIG
44555	KVANFFDEP
44556	TKQHFPLGE
44557	MIVKIERGV
44558	ACSKLTCCY
44559	LMYKMHQIL
44560	TAGKTWEGK
44561	KFHNSIVVT
44562	AACRISEVP
44563	ISTKAHQME
44564	SQKSHTVWA
44565	VKFLVQCHD
44566	SVTIWENFV
44567	KANPSNGFQ
44568	TTGWQVQCM
44569	SYVFAQIGC
44570	ATRQLFDCR
44571	GTHRTEMGR
44572	AQNRIGNQN
44573	MSNTAVQFV
44574	QKASYVEFD
44575	FGMKSKQMS

[2051]

44576	IQREVTWCT	44617	RAEAHWHNF	44658	TIFRLRKSC	44699	KKDTVNENA
44577	NVNREHRFA	44618	NKLIQEGCP	44659	MMSSQMKCG	44700	LSGRMAQQA
44578	MDRLLIKMA	44619	EKAGSGAWN	44660	HTARNVRMR	44701	ATLVKAFCA
44579	TRFMTDCMF	44620	SAGMKPMCC	44661	TGKSWTMLG	44702	ACTHLMCRP
44580	VKGQMYSCE	44621	RKSQKRMIN	44662	EAAKESRLA	44703	QVTKLPPFS
44581	KCEEQKPCP	44622	SAAIDRLQR	44663	GVKLQSYMT	44704	SCHMIAEFA
44582	VCLQFYQGR	44623	AQAYYGVCD	44664	DRRYFTYAG	44705	SGYFRIQLS
44583	MTRASHVMC	44624	MAMRVGSHQ	44665	LCYMLFEMT	44706	STKIGAKFL
44584	RADQQVRWA	44625	QKTPTMNEW	44666	TIKCWNWSN	44707	TSNFNCKMW
44585	RADRCSDLY	44626	NCTMIQRNY	44667	STGMWQCAQ	44708	GSQRIGIIP
44586	TSVTIVECV	44627	KCNSQPSSD	44668	KLTECPLCG	44709	QRAGSFGWD
44587	SNRTCAVSV	44628	IKQPYALTA	44669	KTFFYHAPC	44710	NMRMWEWTT
44588	VRTYVQACE	44629	GQQLIARCG	44670	RISDAYQCP	44711	IAVKAWFDH
44589	ALMKLWCAM	44630	EERPWYCKE	44671	SSKYAHYCA	44712	KSNKRCGHL
44590	QKNPTLDSA	44631	NRELRMHHE	44672	IATVWARCP	44713	MGLQRMEFC
44591	SMRRNKYWC	44632	TEQIFMKGC	44673	TIGRTQFCH	44714	KPSHYSCAF
44592	TKAEHVVMH	44633	QKRLGEEVE	44674	SSHMEFSAK	44715	KGATDVPSG
44593	VAVKHHTTC	44634	TTYKAMSGD	44675	MIQRVPTVA	44716	LAIKCMDFW
44594	FCKTSTSFN	44635	SHMFQYSHT	44676	SKTDKMEFY	44717	ADNFLVKHG
44595	SANLLSDFH	44636	TLKVHGIGG	44677	TTAIRMYCV	44718	SRVIRSDSC
44596	KSVMDDVVD	44637	KIQQNMETL	44678	KCQYCSDMY	44719	MKQFCTCMS
44597	KGETTQNYG	44638	KAMQPECGA	44679	TCKPIYHIE	44720	NCLLQWGQC
44598	KAMDGAALT	44639	ASYRMGDHP	44680	RVNRKQINI	44721	GAAVLSALH
44599	KMEVRMDHG	44640	AFAKLIHTE	44681	SGGKSQIVP	44722	KKNTAFAST
44600	TKTTTNISI	44641	KAEPNMALE	44682	MFRINTECC	44723	QRTATLEAW
44601	TIAREKGYP	44642	FRKSYWYTG	44683	NAATIYREA	44724	SKLTVAAVE
44602	AIKPSTWQQ	44643	MGGGQASWY	44684	ASTKQEYWR	44725	SMYRVMPCV
44603	MKFLLELWS	44644	VLHKAEICL	44685	TSALKPACN	44726	HGILQNCFI
44604	KGMMKDTEQ	44645	YCTQCRGGV	44686	TVNLNGESR	44727	LTVQLLECA
44605	IMKPFMEMW	44646	MCSQCAVLD	44687	SGTFAPFGW	44728	MMMRQHLYG
44606	ACVTCTTFK	44647	RVNTSSRAT	44688	ACAKDRYAS	44729	TAHFSFDTA
44607	AMKSRCRCE	44648	IRYAYDGNE	44689	KCMSALEQQ	44730	QKKLWYGPV
44608	MCKFSVWAL	44649	VATQLIHSY	44690	ALTMLIKGC	44731	VGHKLSGNC
44609	YNKAYGYMA	44650	IWKQLVEMQ	44691	SKTLFHMIQ	44732	TKLQFTSPH
44610	QRAHYIDCD	44651	KGGTENGWN	44692	VSVIKFKGE	44733	VGNFLYRAY
44611	LAAVWVVDG	44652	MSQRIWTVG	44693	KPAEAHSVW	44734	TSGTSYQCQ
44612	MAHGFQRWC	44653	NKQMFCKGP	44694	TRQLYGLHD	44735	RWQNAYEVS
44613	STKIQYFCE	44654	MTYRSEFMR	44695	KKVMGMLYH	44736	ARVMFDMSR
44614	EYVAWMKQW	44655	APGKGNRAC	44696	MKTPLAEAH	44737	SGQFFNVSL
44615	KSVQKECYF	44656	MFTRFMRCG	44697	KVEGCKSHY	44738	VSMMLCASL
44616	MCAQWCQCN	44657	RCMQSFTFH	44698	HVAMCSNLM	44739	KGGLQENFG

[2052]

44740	TALAKLHVG	44781	VYNFSARER	44822	QCRKDWGHY	44863	YKNHFMHMW
44741	ACQPSFAGR	44782	VIAKMCEQN	44823	FCNTLVISA	44864	VALVVVMSH
44742	ADKPLYQNE	44783	NCNLYYDHL	44824	AWYRTEQMW	44865	ACVTNLWGD
44743	DWVLWTRIC	44784	TIFTTYSMK	44825	TMKSWVHGC	44866	MMMRQEFCE
44744	VANAHWCVG	44785	LAGKAQMWQ	44826	SAAGMNDCL	44867	IRKYYHNGV
44745	SCAKTFAYS	44786	KMNGCQSTE	44827	KTTTCWADP	44868	KFAIAQPVS
44746	KIQQDIDMG	44787	MKQRTWSGK	44828	ARRIEEGYG	44869	TAMKQPTLP
44747	KLKTQELWS	44788	MMVVWQNMC	44829	LCMPLLTFD	44870	SVIIHPSYE
44748	IVMREAQAS	44789	FCVYFVDGK	44830	NGNVVMKEY	44871	VGAQFAVQF
44749	TKNFTKGAW	44790	SAMIAGLGK	44831	SSLTCSHSV	44872	SKNELFVMH
44750	LCAKEGSHE	44791	DRKLFQFSG	44832	QGLKIYCHP	44873	TACFNALSC
44751	TCLHVWQCG	44792	TSWTCPRFG	44833	AGVGSWQNE	44874	DCYKSIVGV
44752	TSHVMYQYC	44793	MMKILQHSA	44834	KGAIQQEWW	44875	VKFLVQQGQ
44753	AAVMQFMGV	44794	QIVVWFKEP	44835	AKNPYLWYQ	44876	GTQTRFDNS
44754	KMDWPNVLN	44795	VSRITTIMA	44836	ACSILYRHL	44877	KSGPVVCKI
44755	TNQRIPNCE	44796	SAGSCAHMQ	44837	NTWMNLESC	44878	VAHISSGAL
44756	RSKDGGNCW	44797	LSLRLERMC	44838	TGGQLFQCW	44879	ARNLAQGGY
44757	SHTIFLHSF	44798	KGTPNLVSE	44839	MNVKKSGSL	44880	VFGINGKCY
44758	TGRLQEAYA	44799	TMFSCKYIG	44840	MGLPVADWG	44881	FFTQCARVP
44759	SFCFCSANF	44800	VTNMRLEGV	44841	FSTSHFHLW	44882	TSWFKDCVG
44760	TASFFMSVM	44801	AVKLPMCVD	44842	GSVKGMSVS	44883	STQIWSKMH
44761	ACQRVEKYT	44802	KINLTQSAT	44843	YSGMFHEGR	44884	ISGTNRMVP
44762	RAIAGGECV	44803	SAQTHIKFM	44844	GWKPYAYAE	44885	GAHTCAKAG
44763	VAVKGWWGT	44804	TAVIRSNTA	44845	ACCLTMCSY	44886	TIVFSLHRD
44764	KRCVTHECR	44805	VFNRLIQSI	44846	STRTTSSSM	44887	MCAQMMGVE
44765	TAYMTFDQA	44806	QKAHIFDVT	44847	AAVPCSMQQ	44888	TSASCCQLT
44766	SSYRGVACE	44807	NSSKLIKYT	44848	STQMQTVGL	44889	KGRWVDPMA
44767	FCGSCVAFE	44808	GAGKSNCNP	44849	NFFLTAVEH	44890	MWKEITGQS
44768	AMGINMSGF	44809	RGDTQHCMY	44850	YAQSKVRMG	44891	SALITNADQ
44769	KVVPNFPAQ	44810	KKEYNNGWW	44851	AGGKAQAVE	44892	TCTTLTNQA
44770	TCSKDGSWW	44811	AAYQMTGIY	44852	TATLCANFC	44893	SSVHQMLTC
44771	ERSPTFESQ	44812	TAMHTHKLG	44853	TGTVKMGME	44894	AGKITWQND
44772	LKASYHYQR	44813	NHNMCAVVV	44854	TTYIRLGGP	44895	SGKTHSGEQ
44773	FANTLANNA	44814	DNRFMFVDT	44855	SVFLSQSCR	44896	MGHMCPGCF
44774	ATRQDGMAY	44815	KWAENTVSG	44856	AKYEKVGMA	44897	TSTLHFSAD
44775	SPASCQAAA	44816	HALLKMCEG	44857	ITQIWARVI	44898	KYETENGIK
44776	AMGFSPHER	44817	ISSTCKFGQ	44858	QMNRLGNWN	44899	RASVTMGSA
44777	SRYVVNSAH	44818	TRLKLDVWS	44859	QSSWHSKVW	44900	VTQKCREYW
44778	MANIWHKGA	44819	SAYQLTGTA	44860	KSDQSVCDQ	44901	FTTVCAGCV
44779	TKTPCWVWS	44820	ACVASIGMK	44861	LAASCGHVG	44902	MRVTYPWAD
44780	AWGSFIHMR	44821	FCRIHMMCW	44862	GATTCKQID	44903	AGAFMMPNW

[2053]

44904	SKSDKLECR	44945	ISQLVWKCV	44986	TYQFRNRSE	45027	THTKAQIGH
44905	GGTYKLNGC	44946	TAYRAIGPY	44987	SAAWQKRCP	45028	MGSRHFHFP
44906	KTIPVCAHS	44947	QRGTYNQMM	44988	RKEEFHACN	45029	GHYVCSMAY
44907	KADQGVKDW	44948	ATVGKVDIQ	44989	GSMFTCYLA	45030	DKHGNANWN
44908	KSTPLAEGV	44949	KTATNMNSF	44990	GESRLNEHW	45031	GMSRQSIAD
44909	VGKGMNRFN	44950	GCGLWVYSS	44991	IKTPYCSSG	45032	KAEHFNVDA
44910	KINTECMHD	44951	VTTSLMDSA	44992	FANSTQGCA	45033	IKMNHHIQG
44911	MAIIMSELG	44952	KSMMNESAV	44993	AACLNTQYR	45034	IFGKQHRLH
44912	TPKHQQGYY	44953	TIIACPRGS	44994	FISFHEGYR	45035	ATGTKISVE
44913	FREEYRRGG	44954	AAFGMNTFP	44995	HQMKAMNLG	45036	MAKINPLHG
44914	ACNYFSREC	44955	VLYKIWQEK	44996	ACMGIKHVH	45037	TTRHCAYAQ
44915	ATLKVMGHH	44956	TCRRYECSV	44997	IAQQTWQFE	45038	MIQFDMMHG
44916	KCTTEGTNA	44957	EKAMKSCTF	44998	ERYMFLNME	45039	KTKSGAVFT
44917	AMLMNTACG	44958	AQNLHFYAP	44999	HKTPYPHLR	45040	AQKLMMGQG
44918	TINRWMVAF	44959	AKSQCSKGL	45000	RAEMLHEAV	45041	VSNLTWPGP
44919	NGNMIWRGK	44960	VGTPTCQGV	45001	SGQHMPGCM	45042	ESKIVRYNN
44920	IVNTFWNYC	44961	GSHSHMLDR	45002	GYMLKSDVQ	45043	SAGRWCQTY
44921	SVYLKPTVS	44962	SAAYRYCEP	45003	YCVQTAFMH	45044	NGALQMVQW
44922	KHALAVSQP	44963	KVNPLNGHW	45004	FRAYFQSAQ	45045	GMFIQAVEA
44923	VGNQIPRGE	44964	GRIQFIMCG	45005	VSFTWNRCI	45046	AGNTTIKFE
44924	STHILEGWD	44965	KATQEVWHK	45006	ACASLINQM	45047	AYFSCMHVN
44925	KSVGGTEHP	44966	KKEAVGIMN	45007	TNRHAEAYI	45048	LGMTCVAHG
44926	QKNATAVNN	44967	GASSSYQNS	45008	SCQPAYGGG	45049	VAILALRCH
44927	FCSQHLTGP	44968	EREAHCNQA	45009	YTNLVIGVW	45050	TKQQVGDCW
44928	QCRVCQFHA	44969	VPHISFKCC	45010	MIFVISVGN	45051	TGLMLQQIC
44929	VAAQFSRFE	44970	LAHMAYPSV	45011	TFYVQVNDK	45052	NVHFLFNCA
44930	RSSSETLWS	44971	TALVQYNYP	45012	SANKVVHTD	45053	NYRVAEWVA
44931	HKKEFVMNS	44972	ESMRGQQWW	45013	KCAMALIHK	45054	AARIACFDS
44932	TKNPTQLER	44973	TVAFSMCRP	45014	MLFWFRCLL	45055	VCVTKQMAC
44933	TGMVLNGAM	44974	RCAMDIPHG	45015	DIEKLHHVF	45056	DRRPVCCYT
44934	MKSHQMWVM	44975	LAMPCSGLC	45016	RYVCPLDHY	45057	KNDKGQCYF
44935	VPKYWQQAC	44976	GKQFNMGDP	45017	SCCQEANGA	45058	LGIKSQDAA
44936	LVWRLHQAP	44977	NKTLVGIQS	45018	THYFRNMHG	45059	QSALMSKYW
44937	KSTAPEMMN	44978	TANIASVPG	45019	AFVPFGNDL	45060	TVGIKFQEM
44938	SFVIHSRTA	44979	MIVKTVMDV	45020	LGKVVYQAY	45061	YTKFMSGGM
44939	SMFIFLHIC	44980	VTWSCSEIA	45021	ASKLDTALT	45062	AFRSPNVMY
44940	AMASFWECG	44981	IGHGVQTAF	45022	MCKLLATSQ	45063	YSSLLALRA
44941	QSFIMYGDC	44982	KCNNHVNSS	45023	SLKAEYSGN	45064	TQAKLQCAE
44942	ASLKSMEYR	44983	TRVPTQSIF	45024	NTQTVNRMF	45065	TANHATMCD
44943	VGMANMTCQ	44984	TCHVSQDCT	45025	QVHKGWACV	45066	SPQPILDLQ
44944	MCGHHAQLG	44985	NTFMSTFLE	45026	SQRVWEQVC	45067	TTVLIQICE

[2054]

45068	SAASLIPTL	45109	AGMGIHAMP	45150	KAHDVLSNP	45191	TAALWKIEC
45069	MKVDSTFIH	45110	AGQLVQACH	45151	VTNLLMRAS	45192	TGQPYYNTA
45070	LGAANQMCH	45111	TRDTFFKMG	45152	VGWACQAFY	45193	AILPCSTPK
45071	MSTADMIRI	45112	STYVQMDGG	45153	MTNSSHKWS	45194	EKQVYIVRN
45072	KAALMGWEK	45113	TSTTKMLSC	45154	KAAWSEDGR	45195	MAKPAHAED
45073	NCVVLQKLH	45114	TANATVFHS	45155	ASGLYYSHG	45196	KVHEASYFP
45074	DRAPKLSWE	45115	TCEQEWKAR	45156	MTGLQRVAV	45197	VMIKTWASL
45075	VCNQFQHWH	45116	SAMPGASMN	45157	KPSTKMTLR	45198	HSATAVREG
45076	KISFEVTEY	45117	GASYANNVV	45158	IATKIRGFY	45199	GQIFMCQHI
45077	TFYIWNHCD	45118	AAFTHMGTF	45159	ARGVLNLED	45200	AVSLHEHLQ
45078	MCKQIIQGG	45119	AVKNWEVHE	45160	FMNMMGSCS	45201	AAMPVMKCW
45079	SCSVWRCMS	45120	VASFWHCGA	45161	FAHQHPGCP	45202	LGAMMALHC
45080	VSILMYQNQ	45121	AIYPTNFVK	45162	NCVSLAAHW	45203	RCGLSSDAT
45081	TSMFYEMCN	45122	NPALHNQCQ	45163	DSMRCTPVP	45204	LALQIQENR
45082	SYKHCSQVG	45123	QKMAVCDMP	45164	FDRLITQGE	45205	TARTPNDFE
45083	FAAQWFWTG	45124	KLHQPTWAA	45165	TGLADQDWW	45206	MGTFKLDFN
45084	TWATWLEGQ	45125	SAKYHMDVE	45166	NSIASKKSM	45207	TSTRANMLC
45085	SAQVCQFIQ	45126	QCVANYIAR	45167	MFYYAFTCG	45208	AFKLVSCLE
45086	NAMSQVSLA	45127	GWHVQSCAM	45168	NSMVGGMSE	45209	ISLPCHEVK
45087	SNQFFPQSG	45128	VPITCYNHG	45169	LVIPWARVE	45210	EKGYIFVID
45088	TGTLMGSQG	45129	DKQPHNTGF	45170	FAAMACDMT	45211	TTSVLTGTG
45089	MCNYSNSGE	45130	AQNGMAICE	45171	NSTSIFDAC	45212	FARQFMEYW
45090	TCSQPPCCG	45131	VLFQTMVED	45172	IAAHVVMCD	45213	FGHFVGLTG
45091	GHYIFNCSQ	45132	TRKAVCQPV	45173	SAVPVCSCI	45214	STMSIVRTP
45092	IGHRYYWSM	45133	VALKMFCTY	45174	NKNSTNQWE	45215	ASVTNVKGM
45093	GCHTIQFYR	45134	GANTEVLAG	45175	TRFEYWSNS	45216	MSKLHAGIH
45094	NCNMYHYFA	45135	MGAVTMVGS	45176	QMTMWANFY	45217	GIAFAVHGI
45095	GAQPRVVLS	45136	MSYLKTTCH	45177	SPNPLPWCD	45218	MFQTTILES
45096	KATCYMDAN	45137	TMGLHQDHP	45178	KAINPLKAT	45219	MAHSYQSFC
45097	SAKQWTTVE	45138	LMHPLGTCA	45179	SVSIFKGSN	45220	VTQMIANVQ
45098	NANTNGHMF	45139	TMGLAGKSA	45180	GNYISSMAC	45221	DRFTQQSAK
45099	TAGQPGILG	45140	YGQHWNPLT	45181	YATTTMSML	45222	VQKYVTQAF
45100	WGTYWFNMP	45141	ACYQLPSAN	45182	KVHTCGMNY	45223	SASTTRFIG
45101	VSSKSVHDV	45142	ATKEKHPFY	45183	NWNMWLKSL	45224	VAATAYMDE
45102	SPQSFPQFE	45143	MMKNTQFCE	45184	FAAITRFQS	45225	RGEHFKCSY
45103	SKAANNQCC	45144	KTNFRCEYY	45185	SVTAKNLCL	45226	HSNMWLDQY
45104	MGYQPAVGN	45145	MVAKEIDGP	45186	VCTICPVQM	45227	NKTSNIMYH
45105	AEMRTMIQA	45146	MGSLQNMEF	45187	NIKLTHFEM	45228	AAYSPMFWR
45106	GGTFIQGQG	45147	RAQPQERVV	45188	VTQISCVDT	45229	SVMPCVDAY
45107	KGTDDGQWQ	45148	TKAECMPYQ	45189	GAGMEASMD	45230	TAVPILDWT
45108	GCNIQNMAA	45149	VAMFNRLDR	45190	AAVSVHKCL	45231	SGMLFIYSG

[2055]

45232	NRLTIHTWY	45273	TARNLMGSS	45314	HKGFTALFQ	45355	KSAQMYGES
45233	VAVITQQYP	45274	KGNHKMKHE	45315	GTRTRIQQE	45356	RSATDYGNH
45234	NSRLIRPHG	45275	RIEDTRHNN	45316	ACTNMFAYD	45357	ASGTKWAHK
45235	RTIHPMRFI	45276	MCHQWMSCT	45317	QSMHCANLK	45358	SPVTLHNGE
45236	SKFEFQNAA	45277	STHYKIWAG	45318	LGSHYCNYA	45359	RKEFWIKFK
45237	MGRPVPSHS	45278	TMKAMMEVP	45319	VAMLEFQSA	45360	VAARYTAPN
45238	MAHNLVSLS	45279	ALAHEAGKR	45320	SSHMLELHI	45361	AVKQTIYMP
45239	KMMCNPTLL	45280	SSNIIQYHE	45321	HGVIPMNWI	45362	FKILKPGLG
45240	NIKLTHCNF	45281	TQKSGSFNN	45322	SGAKMAWVA	45363	AAIPFACMG
45241	AASSWVLME	45282	ALKLAQSHY	45323	SQWMLPCGE	45364	MCRLWHDAW
45242	MCLQPYNCY	45283	VCSHFCNEQ	45324	MPTHMWDAP	45365	AQKYILVGQ
45243	VTHSTPQCV	45284	MSAWKMRDR	45325	NVMMQTYSR	45366	IQIFCFWCG
45244	TSKVVSDAP	45285	TVTVWPVCC	45326	TCQGIGSVT	45367	NRKGQHVQL
45245	ASSIVFTAA	45286	NSKHLNTYV	45327	TCRVAEFVP	45368	TKKTCVTWG
45246	KLHSGGYLC	45287	VSHWITKVR	45328	MGVMLTSEW	45369	TVNMWGADD
45247	TAGIIYFYH	45288	TSMPYKAYG	45329	AQNLYSAFW	45370	GFKPEMRMA
45248	SGKQLQNTY	45289	VYLQHNCCW	45330	NHQLTYCSG	45371	HRKHKFVPI
45249	RFTMYSEWL	45290	AANVGMLGS	45331	MIKLRHWDV	45372	KLSWEEYVR
45250	TSRPISHEI	45291	KLQMNYEEK	45332	HCGPCLKGD	45373	VIAQQGNAR
45251	SGMLKPMYS	45292	LGGFHVSAS	45333	AAQYQAIGV	45374	TTAHVYDGM
45252	RTTMGVKME	45293	GATFLQFDG	45334	TNKMVGHGH	45375	VKQMKLEQH
45253	TCLTRHSQG	45294	SQMLSEHLE	45335	RCGVQQGSN	45376	MCVMSQEHL
45254	MADKKWRYC	45295	TPGWVNQIA	45336	TFMRNEWTY	45377	GRMQTSRCF
45255	SFYIHQEAN	45296	TGTFVQMDK	45337	AKYICQNYC	45378	TMRGTGAQH
45256	ASYMKSELA	45297	ATKQTPWAI	45338	GCARNNGAI	45379	HYRPFWSAC
45257	SYTRAVECN	45298	TMCHEMCSL	45339	GCGKLEIHF	45380	ELFMDHRSW
45258	SVQMEGLGW	45299	VGAACKMVA	45340	TAVQTHTGA	45381	FSQPCVNGQ
45259	GTKHTMQQC	45300	GQVTMQRCC	45341	ARMPRESAD	45382	RPFTNEGLC
45260	TSQHLQVLC	45301	TWTIKQIGE	45342	VMIPRMGCG	45383	VMSPFGKFC
45261	MHNRIQKGG	45302	SARQVLDPD	45343	YGYMANKMK	45384	TVKIYGEAS
45262	VIKSYVSTA	45303	VSKSWLNSF	45344	NSYTNFKGG	45385	AYRYEYCAI
45263	RLQYAGQFL	45304	GGMFLAEGS	45345	VLYSWWHFD	45386	IGTTWYGDT
45264	DKQFIAEWP	45305	NYVTLGNHM	45346	KCFEDLCST	45387	ERARSHECT
45265	SHEHCMKMC	45306	ISKHCHDQG	45347	VSNPMSCKK	45388	KIESIVGVT
45266	KTEDYQHHQ	45307	SAQSALDAL	45348	SCQHFMDDP	45389	MGASNMRFV
45267	AVYMAGRYM	45308	FAHGHSKAR	45349	VAAHWKMSN	45390	VANYCNRPE
45268	GATPLCDAN	45309	AYFYLGAEF	45350	AAVAMQMCE	45391	AGQTFIYVH
45269	GAYHVTNQT	45310	QKTQHPVWN	45351	ISQHCHCKL	45392	LAMMCHADQ
45270	SHQHHPWGA	45311	NLCVSLESY	45352	QWKQAQCCT	45393	MHVMSGHLF
45271	GATIVMTGH	45312	MSQRVSVPS	45353	GHRSRLSGA	45394	AMRQANAGV
45272	TAVAQVDPV	45313	KEKLQMNHM	45354	TCASLQSNQ	45395	QYSIWNFDH

[2056]

45396	TTHLGMNHF	45407	NATGAPRAW	45418	TEMSLPGME	45429	YGAAKQAMA
45397	MSLHKCDVR	45408	MKHDVKEWN	45419	NAKLGDACG	45430	AVTSSLRDL
45398	NFAMSMRDH	45409	FSTQQCCVE	45420	AMKKEGEWG	45431	TGQVQGFHY
45399	MKYEKNSCW	45410	EKKFHLKYQ	45421	SAHSTAVYR	45432	DKLMFGEWC
45400	GSSYLNQPT	45411	ALGQFATCD	45422	VTYSCVCTH	45433	AASLKPWMY
45401	MSRHCSMHD	45412	GGFRFTAQA	45423	GSNWIYEEA	45434	TRGLYMWSY
45402	MANATSREN	45413	IALSCQSHL	45424	VSVIFPQQG	45435	LSSRSEHPN
45403	EAYRNKNVC	45414	MSQFGMLAG	45425	MVIREWCQC	45436	NASMYGAQC
45404	KNNTKKQWP	45415	AMKFVNYFY	45426	MVMRQLFYN	45437	SIFFFAASP
45405	RVAFTLESF	45416	SMHLHMKDL	45427	VSYYMYEAP		
45406	AYKQIHDST	45417	ISTLLSYCG	45428	CVIRAHQLC		

[2057] 实施例21

[2058] 如通过机器学习鉴定的具有肝脏去靶向组织嗜性的AAV5变体

[2059] 此实施例描述了如通过机器学习方法鉴定的展现肝脏去靶向组织嗜性的工程化AAV5变体的优选特性。进行如实施例19和图19至图22中描述的相同生物信息学预处理、机器学习方法和分析以确定展现肝脏去靶向组织嗜性的工程化AAV5变体的优选特性。图27显示了一组对模型输出概率有贡献的前20个位置特征。沙普利加和解释(SHAP)值可用于询问特征对模型预测的相对贡献。如图28A至图28B所示，在组织靶向性和非靶向性变体之间进一步比较了这些特征。图28A至图28B显示了排名靠前的预测位置特征的比较。如果发现特征对HGB和RF模型都重要，则选择这些特征。显示了相较于随机2％肝脏去靶向性变体，前1000个ML预测的肝脏去靶向性变体中的特征的概述。

[2060] 下文描述了AAV5 VP1衣壳多肽的581至589区域中各位置处的有利生物物理特性和有利氨基酸残基，它们与较高的肝脏去靶向组织嗜性概率相关。任何以下机器学习衍生的位置氨基酸偏好描述于表42中。下文表42中列出了从两个ML模型(HGB和RF)中提取的前30个重要的肝脏嗜性特征之间共有的位置特征的概述。

[2061]

表 42 机器学习衍生的肝脏去靶向组织嗜性规则
位置 Xaa1 处低溶解性 Xaa1 选自 D 或 P
位置 Xaa1 处低可突变性 Xaa1 选自 C、K 或 L
位置 Xaa2 处低溶解性 Xaa2 选自 N、K、P、E 或 D
位置 Xaa2 处低亲疏水性 Xaa2 选自 D、E、R、K、H、N 或 Q
位置 Xaa2 处低电荷 Xaa2 选自 D 或 E
位置 Xaa2 处高总潜在氢键数 Xaa2 选自 H、N、Q、D、E 或 R
位置 Xaa2 处中等体积 Xaa2 选自 D、E、V、P、N 或 T
位置 Xaa3 处低溶解性 Xaa3 选自 P 或 D
位置 Xaa4 处中等体积 Xaa4 选自 D、E、V、P、N 或 T
位置 Xaa5 处低溶解性 Xaa5 选自 N、P、E 或 D
位置 Xaa8 处低溶解性 Xaa8 选自 K 或 Q
位置 Xaa8 处低亲疏水性 Xaa8 选自 K 或 R
位置 Xaa8 处高表面可接近性 Xaa8 选自 E、R 或 K

[2062] 下文表43提供了表现出肝脏去靶向组织嗜性并且符合表42中提供的一种或多种规则的变体AAV VP1衣壳多肽的581至589区域的序列。因此，本公开提供了由具有SEQ ID NO:2的工程化AAV5VP1衣壳多肽组成的rAAV，其中Xaa1至Xaa9区域是SEQ ID NO:46438-SEQ ID NO:47437中的任一者，如表43中所公开。本文还涵盖由工程化AAV5 VP2衣壳多肽和工程化AAV5 VP3衣壳多肽组成的rAAV，所述多肽在对应于AAV5 VP1 581至589区域中的氨基酸的AAV5 VP2(氨基酸残基445至453)和AAV5 VP3(氨基酸残基389-397)中的区域处具有下文表中公开的序列。

[2063] 表43

[2064] 驱动肝脏去靶向组织嗜性的AAV5 VP1衣壳多肽中581至589区域的序列

[2065]

SEQ ID NO	581-589 序列
46438	IDGDDEPRA
46439	PRSDDDQRF
46440	WDCEDERRH
46441	DDPDDLECC
46442	PESDPERTN
46443	REWWWCLRI
46444	DDHEEVFKA
46445	DEDECEGRM
46446	AEIEKDPRW
46447	PEPLPEPSK
46448	YECVERDRF
46449	PHPQATGPS
46450	QEHHPRDRW
46451	CDTEAEFRR
46452	MEPGEKSRP
46453	LEMGGKTRT
46454	CHPHDEHKT
46455	MDLFPPHRH
46456	YEDNDVMQI
46457	PDMKSVVKA
46458	KWPIEWGHY
46459	VDRLPPNKS
46460	MDWDDAIRG
46461	WDRPPVMRP
46462	YDWGRVTPN
46463	REPWQTTEH
46464	PWKHEPSEF
46465	FDSHPKIER
46466	LRDRDGNRM
46467	RRGMEPTKH
46468	GDERPGQRN
46469	PLDRDQATQ
46470	PKNCRWYRT
46471	PECGDRGDR
46472	WDHEPETKH
46473	EGPDPCNSA
46474	PECFPKGRN
46475	CDHDTKHPT
46476	PADFRDTRV
46477	CEYDDESQF
46478	DDEDVSEEI
46479	HDKPGKPPY
46480	KNPPWICVW
46481	PIEKDDNVD
46482	PKMDKVEPC
46483	PDPCTEGTC
46484	ADNCPKNDN
46485	PDHITKCDE
46486	DDHDDKLVA
46487	WEDEPPWKI
46488	PQPDENYPW
46489	FDREKRDGN
46490	GNPHEGQKF
46491	PEGSMDRRF
46492	GDWSERSIT
46493	PERWERPVW
46494	NDLCQDQRV
46495	TRPDESPRQ
46496	WECADRRPT
46497	YESCNLTRD
46498	PPPVHNDRS
46499	VHPEACMRN
46500	PDQVDRNTP
46501	REPNSGLVY
46502	EEPEFTGRG
46503	TDHNGEIRP
46504	CDRDPSWRD
46505	QDRNRVHRQ
46506	IDPYDCDPT
46507	PDQDDMMHG
46508	YEYRPELRA
46509	CVPEPWDYR
46510	MDPKHHCES
46511	QDQCDDKIE
46512	MHYCDRERK
46513	PPDRCTAEK
46514	RNPSQDQKG
46515	CDTHDVSKD
46516	DVADVPDTT
46517	DEKPMDGRG
46518	PDTSPYCKS
46519	QLPDKWPKS
46520	YRCCDDSKT
46521	PAECISDSQ
46522	PDPLYWIMN
46523	LEEGDHMHM
46524	GDHEPKHEH
46525	WEVEQRPRH
46526	CECWPEYKR
46527	PFPLNPERH
46528	DVISRNGRC
46529	PNEDKIKPD
46530	YPCDDTKPE
46531	DTDEENSPH
46532	IDPGMMFRG
46533	PEQRDRLGF
46534	DEPSDDDEE
46535	HHPPKQPPN
46536	CRWEDYDSR
46537	CPPQADRIQ
46538	PQPQFDMMQ
46539	EFPDGGHKM
46540	CNCPDQHPD
46541	WPKCDECRM
46542	ADNEDARPK
46543	PICRPDTDV
46544	APCHDKDRE
46545	LDPPDSACG
46546	VETEEKMIL
46547	MEDTDERSN
46548	CDRDKRHDH
46549	GEEADEYHM
46550	RDIWECIRR
46551	AECQEDRRG
46552	TEPDTVDES
46553	PDPGIRWPM
46554	PRDVDDSAY
46555	VEEDQSCKT

[2066]

46556	RQYFEDNAG	46597	PECIEGPNH	46638	LETAEKSKS	46679	PETTSRWKW
46557	YDQWERDIH	46598	HDNKDVNRN	46639	FDSCSRHRW	46680	RDYQDFICY
46558	RKPEWSCMV	46599	GEMTRCFRN	46640	DNTDPDGSW	46681	CDFQSSWDQ
46559	LNPQSPEPK	46600	PDKEKVMPA	46641	ADPQSVCCT	46682	IECFPRESI
46560	PPCKMKIDT	46601	WNSPDDERV	46642	PHCQTVSDI	46683	PERASEMFW
46561	PNDWITDPH	46602	VDRDDGKDY	46643	IEEPDTCLV	46684	NYPRDHRQH
46562	VERMENPKH	46603	DFEDELSKV	46644	PEMQNDMDS	46685	FEAHDRPSN
46563	DAPEPHAQK	46604	YEDGDHIVE	46645	DHLPEDKRY	46686	WEMNDMQDY
46564	PQEPEEGKT	46605	VEGEEILRI	46646	QENDANRPQ	46687	LDTCEMLKQ
46565	HDMDDVLNH	46606	CEWLDLMRT	46647	KQGCEKNPM	46688	QFPEFDRET
46566	VDCTPDRRN	46607	HEAEDWRPF	46648	HRCDPDMGR	46689	PVPHPEKEF
46567	PHQARKNKP	46608	KMPCEGEFT	46649	DVPRDRKEN	46690	EFDEFNTRG
46568	CEDEGKTWL	46609	GDPWEEWRK	46650	EDNNNDVKH	46691	GDQAGATRF
46569	PFPQMNSNA	46610	HERWDANPY	46651	IDFRPPNNT	46692	EPPTNDART
46570	ADLRPSVDY	46611	PGCCDLTNI	46652	DFPDKGELG	46693	AEVGTEHKV
46571	GEPWDTRIV	46612	PIIKGEPAD	46653	CIPDVAIMM	46694	DDPSGQLRL
46572	YKPKFDAMG	46613	PDTDDIIHW	46654	MEDHTAPRW	46695	EDPHSKKEH
46573	PIPQQGGTA	46614	EDLNEKSKS	46655	PENQQNHRD	46696	CRKEQGQRT
46574	MDGRDKVDP	46615	EVDDGNVGE	46656	CRKSETPWD	46697	PEAGGKSWD
46575	CEDQRQPME	46616	DDWNEAKRN	46657	RIPPMMIML	46698	PWDPEEDGD
46576	PDFHHKRDF	46617	PPDMPMSAD	46658	PQPWNQQAF	46699	ENPEMMGFE
46577	NEGCRDEKS	46618	QNPWDTDNH	46659	CEIQGGQRY	46700	MNPNEKGGL
46578	LQPNAVGKT	46619	PDTWAYAPH	46660	VDFHDGFND	46701	PDAAERKGV
46579	PKLCSHQTE	46620	TEPEIRNWF	46661	WDDHKLWRW	46702	EPPDGQVDA
46580	QVPDEGYSW	46621	FPSDPDGEH	46662	RDPFSPDPM	46703	FDANDCRPS
46581	NEPDKGYTM	46622	PPPVVSPKP	46663	HLPRDDPVM	46704	FRSHDREEN
46582	DLDDVCNKC	46623	FDPHNPCFR	46664	SDIPPNCRK	46705	IKCEEPMRL
46583	QDTDEGDQC	46624	ANAEPVHRC	46665	DPPRSKDDS	46706	FIPDTDNWG
46584	FDNPDEYVT	46625	LVPESQMHG	46666	NDPDTSRTA	46707	IDCRSRLEE
46585	YPDCEGGKM	46626	WERCPELNQ	46667	IDWPDNQKV	46708	CKEQPTERF
46586	DVNKPDEKW	46627	MDEHREQRS	46668	YDCENILRT	46709	DDMEDQQQY
46587	EVNDDHAKP	46628	SDPEKYPWD	46669	VEKPSEEPQ	46710	DKPEQLKEH
46588	PHQRNRMDP	46629	ERDDPRFSN	46670	LKPLPHFRE	46711	PDSHKQPDC
46589	QEPDSNRGV	46630	FDSSQEWKL	46671	HDEEEKNKA	46712	CDPSVEDRV
46590	SLPQPDTRY	46631	CRKNPTIHM	46672	PHGWEENDP	46713	PEGCHHQEE
46591	FEPCPRSCW	46632	EAEEMGLKH	46673	PLVGMDHRE	46714	FDPCQGGEQ
46592	KEPKMEKRG	46633	CKYLSHERL	46674	HRPEQLHEG	46715	NDEDNMSRH
46593	PGDHRFYWP	46634	IDPCDELFW	46675	PGPHDQGPY	46716	PECLPKNSC
46594	ADTANILRH	46635	CKPTSDMKE	46676	EVEDRSGWC	46717	AEDASRDLF
46595	KPCRDNQDV	46636	RPWEWWIRK	46677	PEPDWVQEC	46718	CNPRPQMVG
46596	LPPCYCQQY	46637	FEPQVEANT	46678	CRSEGWIHT	46719	CDDHAKRNA

[2067]

46720	MRNYPDARF	46761	LGDEPWAKF	46802	DDPEYTLCL	46843	WPPEPWCCC
46721	DNPPGRRKP	46762	LNHCDKEDH	46803	KFPYLPHRN	46844	IDPFCAVRI
46722	PDNSDATHQ	46763	WEDHDLHRN	46804	PVGNKDSRY	46845	DEYCQWIRS
46723	MDCKEWNKQ	46764	PALYRRKTD	46805	WDVDHKLRM	46846	PQDEWQNRV
46724	GNQCELNPH	46765	NEDIPDPPH	46806	PQHGAETPH	46847	LNPHVSQIW
46725	PNYKKRSTV	46766	WEIQDVIRV	46807	PHREEDDHC	46848	MDPYEVTSC
46726	LRCCGMETE	46767	PENWNTGKH	46808	CVPNPLGHI	46849	WDAFEPMRY
46727	PPDMQSAKW	46768	PDRYPAFEM	46809	NDTGPHGRH	46850	YDFSERKCF
46728	DHPSNDAEM	46769	HNPWFSCRW	46810	DDTENLREI	46851	MDPNHVHAG
46729	RIPGSRPPQ	46770	INQKETWKF	46811	VDPWDPAQP	46852	FDCFDRNLC
46730	WDPNQGGAT	46771	PVPYSGPKQ	46812	CDPWDWEIE	46853	YDTDDAQPN
46731	PSCAPCLEK	46772	YEREPRMFH	46813	EQPHDPSST	46854	WDPHKENTM
46732	IDSHCKMRE	46773	LDEEPVFVG	46814	PVPCNTVKF	46855	RDPNAQGWT
46733	CEYDGTMDT	46774	PVDADHQLW	46815	KMPNELGYY	46856	EWPHIDAGK
46734	EVADENGNY	46775	RPPHVGPAL	46816	KEREERRFW	46857	EMDEKLDGK
46735	PQPNIPVRW	46776	ADPTTETPH	46817	DFPEPYYGF	46858	CQPPEEGMV
46736	WDYNSIFPG	46777	PDWWPRCKP	46818	QDQRRQKRM	46859	PPANDVPHQ
46737	IHSNDQPKG	46778	YQPRSTPPG	46819	PRHTMLGPS	46860	IECSPRETI
46738	PHDLKHQRS	46779	NNGDDRRQH	46820	DDREKCNWM	46861	VDIHKEMRV
46739	PIPLMMHDE	46780	REWKDNCKQ	46821	FEQWKDQKV	46862	PAEHDPNQV
46740	FDCLPDREH	46781	PDDWSGTLY	46822	LRPCEGEQD	46863	PKPPPHYEQ
46741	PNDWHIAYG	46782	PEVYPGGTT	46823	CPTKPRYPS	46864	PGNPENDPN
46742	LDTKRTFKK	46783	PKCCCKCDW	46824	PSESKLMNY	46865	DWNLEDDRA
46743	CDRASDEGK	46784	PDCRWGEDN	46825	CDKNWTLQQ	46866	CEYWHQVRP
46744	DDSSQEKRW	46785	CDLPMSWRV	46826	PSPPPGHNA	46867	HDPCTQIKP
46745	PEHQDEQSC	46786	PEDCDDRIS	46827	IKPEPNWSS	46868	CECKPICAT
46746	TEPETYPAT	46787	QDCEPEYDF	46828	PEETRHNDC	46869	PRVDNGNQL
46747	FKDWDHHTV	46788	DVSEDHFND	46829	DDVQVEPKW	46870	PHESCECSG
46748	WKTCPQEAY	46789	PKQLDQPTN	46830	PQNGMLHYS	46871	ADTKRKRRT
46749	EEKDEQERA	46790	EFHDMDMVT	46831	FDSRHIEKP	46872	DEPRDWGAQ
46750	WPNTPWEPH	46791	CEHPESKDV	46832	HEKHSLDRY	46873	INEDAQEKM
46751	CPTGDQQKW	46792	TEHKDHFET	46833	ETPDDWNCY	46874	LNSEPKCRA
46752	RPCMQQDPH	46793	PSNEPSVKT	46834	CDAEAGFRR	46875	PMPNTGDEE
46753	CDAGLISKM	46794	GEVTGHERN	46835	EEFGPTKPM	46876	TVDEQGIQH
46754	EASDMSIEK	46795	PNPQSHLHI	46836	DVPQKAHND	46877	FDGCSKMPM
46755	AEPRAIWRH	46796	CPRCDSRPA	46837	IHPCEGSQT	46878	AEPDTIDWD
46756	FHISPKEKW	46797	MEDNTVCSV	46838	VDLQQEDIY	46879	DIWDTEVQE
46757	PIPPIKTHS	46798	QPPRRAMKM	46839	AFPDNSKTE	46880	PQDLKPGNS
46758	PDPQNHQFV	46799	SLPEEHPFE	46840	EELASSDKY	46881	IDRCDVLQH
46759	TEVEQLPRA	46800	ANPLFKCIS	46841	EEYTGECRG	46882	HDPWNVLDW
46760	IDPNNLTSM	46801	DQSEDLDQG	46842	NIDDKPSMQ	46883	EDPCSPAAQ

[2068]

46884	MNTNENCRK	46925	QDHENNLKT	46966	NDFSDEPCM	47007	IKPKDQAPL
46885	CKHVDWDSK	46926	ADSQDPLQQ	46967	IDPPGTESM	47008	ERCHDTEKY
46886	SEGMDRDRC	46927	PLESHESEF	46968	PYPQQSHLE	47009	CKKESTCAL
46887	YKELDRDPQ	46928	HDDTHRIVN	46969	VRPFDDGPM	47010	PKPVQPHND
46888	PRLQSRKYL	46929	CNEFKHERR	46970	TEPSKMPIW	47011	DPELDIPEE
46889	PGLWIAPDH	46930	NEARDGEPF	46971	PTMKFKDRD	47012	HEDTHVLRQ
46890	PYLHSKQDK	46931	PQNPSEWPP	46972	CHSGDILDT	47013	ESEEHRVMS
46891	PQCHDHRQD	46932	CRCDTKIRE	46973	PPKMMDNEF	47014	PMEEMHSHE
46892	EQPTQTPFS	46933	QEPLSKHGV	46974	PSWGGREEE	47015	DEEVEGMSV
46893	HPPKRHIPC	46934	INPEIRQIE	46975	LDQNNGPET	47016	DEPFFAIRP
46894	VDTQKTERA	46935	YDLSPKMYY	46976	NETPDAWNV	47017	YEMMPNIKY
46895	DVDDLVEIM	46936	EVSDYQQWV	46977	WEVPRWPPI	47018	EYPENSSHF
46896	CELNGKIDL	46937	RPPHKMSKE	46978	FDTDQRQSE	47019	WNPGPVLTH
46897	EEENCEQGN	46938	EDRCPGNPG	46979	WIPCQFSKI	47020	GERNGGWRL
46898	DEDERTLDI	46939	LMPDPVQGI	46980	PLDHRPHTF	47021	PADIDDDGN
46899	YPPSQEIKT	46940	SIDRDHPKE	46981	IWEDMKGKF	47022	QRWSDGDKC
46900	PHSEEGREI	46941	ADPQVEWAM	46982	WRIRPQQEI	47023	AKICDADQL
46901	EQPECETNV	46942	CQRQEPDRE	46983	PGHFEPTNY	47024	VEPRYDDPC
46902	PHCNDGSEP	46943	DAMEDSKTC	46984	MEPPDMQCP	47025	PSEHWGCDG
46903	CECWGCTRG	46944	PDPKAWSNY	46985	DVPFVNASD	47026	RNCCGPNPY
46904	PNDFTTHGA	46945	PQPKQCMCW	46986	LRTCPVWRQ	47027	PEQMDRAHG
46905	CDTNPQDQV	46946	DMSDEGQKT	46987	PIECMGQHW	47028	MVPDQYQMS
46906	PRWQSEECQ	46947	QEPFTESKS	46988	VDPSNAAFK	47029	WVPDLEDVQ
46907	VEATPKSPN	46948	YECDSEHEV	46989	DPPIDMSKD	47030	IDKITWQKH
46908	PEEGSPAEY	46949	PHGTEDWDY	46990	WKPGYGAKP	47031	DDWTPYLNE
46909	DTFQADKRY	46950	VRPHDVFYM	46991	CPPSETRVP	47032	PDQSEMTKA
46910	PYEKEKKKD	46951	FHPDKLMPM	46992	MQDRDTPTI	47033	FHPDRRNPT
46911	DFYHHDDRG	46952	QDRRDDHGG	46993	RHPEGRRVD	47034	QDTNETNDV
46912	DTEEPVVNP	46953	LKLRDEFNE	46994	PFKQGRENP	47035	WNEIDEGRW
46913	PNEPNEEHK	46954	CEVQDEWTG	46995	LDVTPQPEG	47036	DDNQSRGCP
46914	VRIYDRCEM	46955	EVPNNHHYT	46996	EVGDMAHQS	47037	EIQQHDNVI
46915	LRCDPWLPQ	46956	PEPSWKYRP	46997	PGHHPLEDH	47038	SPPDWWPFD
46916	LEWSEKINP	46957	CPWCDTQTY	46998	HLPVYAPKD	47039	MEPDMRQYS
46917	VTPEETVVI	46958	KDKCDYMTV	46999	DVEAMDFGT	47040	DIEFLDCET
46918	PDPPKQYED	46959	CATDPIHRV	47000	CTPGFWSER	47041	LEMQQLTKV
46919	HPGHDRPDF	46960	CESFPGPVM	47001	IDHIGDDRF	47042	IEELEEESH
46920	EHGHDDQHR	46961	PHVCGEGKI	47002	PQSHKLIES	47043	EDWRKDNNF
46921	DVENQDWAM	46962	LREEELMTR	47003	CNPPCTPFN	47044	PDMKDFKQE
46922	PHSEKYAKH	46963	WNPANMFNI	47004	DVCDDHKPT	47045	WPWKPPYSD
46923	PDCLDEGAA	46964	CDPCMECGS	47005	PPPLAGWMW	47046	LQPDRECIG
46924	QDCSDWTKF	46965	LHPPPHQCP	47006	YDPNMFPYT	47047	WDMMFPCKW

[2069]

47048	PHDPGECYP	47089	GPDEEWNIQ	47130	PDTSNKTFH	47171	DADDRVEAT
47049	PQDHMVNMQ	47090	PMDEFGHYV	47131	HERCSELEY	47172	CKDRPIIDC
47050	DNDEQLGGC	47091	PGEENEPPA	47132	TDANRLCRK	47173	NGDDHVFRE
47051	CKPHQQRIH	47092	ENDEASICM	47133	PNGCDRSGG	47174	DYIRIFDRH
47052	GDMNVWVRY	47093	PSRTEGDKM	47134	LESDSFDRD	47175	QDHEREEYK
47053	CEGAAWPQI	47094	PDRQQPMQH	47135	PDLQKRQGA	47176	AEAKERFEM
47054	SVDDFSSHV	47095	PNQLDKPAM	47136	YEYWPDVSW	47177	PDDEKIEHP
47055	VIPESMNHA	47096	EHPNDKIAI	47137	CPWWDFAEC	47178	MNPAVEKGM
47056	DDPCMRCNV	47097	CNMATSNRS	47138	DQKDVKGKM	47179	CDPQYDQED
47057	KKPWDDHTR	47098	PWPDHQRAC	47139	YELCPLPQH	47180	YEAHSDDRK
47058	SWPQGDYRF	47099	WPDSRNSPN	47140	DVEDCLICC	47181	MDPCIEPGR
47059	PVCPMKQRS	47100	FDDRKRDAR	47141	IMPEKMHTD	47182	CQDEDEAMQ
47060	DGEEDWMDP	47101	LYCEADRKI	47142	PHAEDRDNN	47183	PIAQDKVWS
47061	MESEESVKP	47102	PNENKGQGF	47143	TVPDVRDCQ	47184	PRLTDIREV
47062	QDTESKFRD	47103	TYPDVCIGW	47144	EQHRDTMRP	47185	VPPFLDIRD
47063	DDMDLIIKH	47104	PQQQEGGTE	47145	CVMPHNHRL	47186	PEHPDPFSQ
47064	ENPSNWHDW	47105	PPNFKIYKE	47146	PQEDEGVRQ	47187	YWCLDDPRD
47065	HHVSMDNKH	47106	LPAGDEKGE	47147	SQPNPEINW	47188	ERVEDGMDH
47066	NRDWFEWKM	47107	REPLVTEGG	47148	PIDGNSHKD	47189	RDYAPLLID
47067	CQPESVWHP	47108	LCDEGDMPS	47149	PGEPADIKT	47190	QVPEIQSMP
47068	PPCGKGTLC	47109	TEDFENHDG	47150	PDHWECIHD	47191	PRNGPYCDI
47069	PYMLCNSRS	47110	NEPQSPVGY	47151	QQPPLAPDV	47192	CDNGEDDHH
47070	GEPEAKGNG	47111	WDGIDRQNT	47152	YNPQNHVEG	47193	PFDTMDYRS
47071	VNLDEMPKA	47112	MKVTDKLDD	47153	GDNAKDFRG	47194	PNTHPEYSK
47072	ADRCEQIVL	47113	QDETDEWQR	47154	FNVSWENKP	47195	WNREDKRAQ
47073	VNPGQESGE	47114	QEFIGEERT	47155	PERLNGPTP	47196	PDWCQDPGM
47074	CMPDIMTWS	47115	QDDIDLASW	47156	FDPHRCSCH	47197	PQQPHWCDG
47075	ENHENCDQD	47116	PTDAALHGN	47157	DRLRDDYPC	47198	PPDPPGDNV
47076	MDDMDRRNS	47117	PTINDWKTK	47158	VSEDDESHQ	47199	SPPNLRAFP
47077	MDWCEYHRS	47118	PMDHEKRYH	47159	FESTQEPIA	47200	GEEDEQSNH
47078	PDHERVEYC	47119	YKCDDNFIP	47160	CKNFPMKNN	47201	CHPTHQIRC
47079	IRPIICRKG	47120	RDSNCHFDK	47161	PVPMSHIMQ	47202	PFPYHLIVC
47080	PPEHCEHTQ	47121	YNDFNRGEW	47162	HNGIEKTRE	47203	PNNQARGQC
47081	LEPPGLDSM	47122	PCDQTLIWQ	47163	KDPCADRWA	47204	KNPTWVMSD
47082	LKTCPEGNT	47123	DFGCKDEST	47164	EVVDNMRSQ	47205	YRHEEYWHN
47083	WDKCAWAWD	47124	CNLCTHAKC	47165	ADYKDWAEH	47206	LEHENGTPQ
47084	MRVCDELEP	47125	ANPDSSVMH	47166	PRNLPAYHS	47207	QDPHAHRDD
47085	MRHDPFHRR	47126	FRDGDAAEE	47167	PMSQICCDM	47208	VDLCRKQVK
47086	LVPDDRLWD	47127	PFPCQQDIF	47168	YDCWAKEAT	47209	PYWCDWPNG
47087	VWDDKSEAE	47128	VKNAERPRE	47169	ENPEVWIEQ	47210	YPCQNHWRT
47088	VNPCSIMKA	47129	CDFCRQFIK	47170	IERWGVIDV	47211	MIPCHDRTY

[2070]

47212	QPADPPNDS	47253	GNQDRTNRK	47294	HNVPDVRRN	47335	PQEDKQCKP
47213	PKHADALRM	47254	CQTRDTLEL	47295	EADERSKEQ	47336	DVVHLDWKE
47214	LEPCTDEFG	47255	EADEQKTRW	47296	PYIWQAMHR	47337	IVPCLWPIP
47215	PDCDPAHHD	47256	TEADEEWDT	47297	KDHIAEHRL	47338	WEPVNAAMA
47216	CRPQAIWLH	47257	GDKWIGCEL	47298	AKPKCPPEG	47339	QEPLRCIPG
47217	HEEIPQQPK	47258	ECDCRDTPC	47299	DDDIGADSV	47340	GNCEPSEPE
47218	CEDLNKPMM	47259	VQPQQMHKT	47300	DDPFENIGV	47341	DNVNNKQKG
47219	HNPSGPEFS	47260	PKCCSGMTT	47301	WDGRTNDKA	47342	PDWFKADDT
47220	GDFIPKQEV	47261	YNPFYGADM	47302	EGDDCMWCR	47343	KHPAELDAA
47221	CKTQERANS	47262	CPWCDMKPW	47303	VNRRRRDEM	47344	PTGHQRGTF
47222	SDDDREREA	47263	KDRDTCRAD	47304	TEVDDKICY	47345	WENEEDKVT
47223	FEPVSGETP	47264	SELTNWVRW	47305	EHGDEMHPP	47346	QNWYDKLPT
47224	PHSDSWCDR	47265	YNPAACVEL	47306	LEKFPYQPL	47347	SDARECDIM
47225	WHKRDDKTQ	47266	CVMPTIYRG	47307	FDEMHTCYD	47348	QDSQDQQQW
47226	PYPPMWRGM	47267	EIKPYDARN	47308	KCPDEMNNI	47349	DNCDAGKWD
47227	IEDVWECPY	47268	EFSGHDEPD	47309	IDTGDWLVG	47350	VKDLADPPG
47228	DYDWSDMKC	47269	VDIKPKGAI	47310	PVPCCHYDP	47351	PSAPNMWKQ
47229	IEMHEKKEE	47270	MEQWDPDEW	47311	PGTEDYSRW	47352	ENEETWIFY
47230	CEPRDCNFY	47271	VEVTPGFKF	47312	LNYQPCCRI	47353	WDNMDQTQG
47231	PTMRESMES	47272	PPNLCWNQH	47313	PAEEYHSRE	47354	EYPCSKMRY
47232	EVPFKTSTT	47273	SRVKDCHKA	47314	PEELYLYRK	47355	QCDNESDSM
47233	VSPHKTSRN	47274	ENEKPRTRC	47315	MDVSDQWPE	47356	SDPDIACVH
47234	ADHCNWSRT	47275	PHDWSPEFA	47316	WFPDHMSGI	47357	HDDQMHCSS
47235	SDHDNVFED	47276	PHQGDREGF	47317	CPPPSHHID	47358	CEEETIEQS
47236	SRPGDRGVQ	47277	IESDDIRTQ	47318	LPPMKQDEE	47359	HRAQDPQQF
47237	QEPDSYACM	47278	HDDPPGGGF	47319	WRPCAYCQT	47360	HRTDEPTQN
47238	DEECQRVPM	47279	EIQDLLCKA	47320	CDDNGCTRD	47361	DVGTGDDRQ
47239	VDNMSEGRQ	47280	CVEDHWLHI	47321	RYPEPDAQV	47362	VDACSHNRS
47240	EQEHDEEKP	47281	HIAEPKNKP	47322	QVPIMDDFF	47363	FDHQNWFEQ
47241	PEFGWWIAM	47282	HEWKCEKRV	47323	PKLPEYTDS	47364	VNCPDQHDV
47242	PDELMGWQQ	47283	YEATPHGRQ	47324	MNSEPVTET	47365	CDMAHMLPV
47243	CECAGWSYE	47284	YPPNEHWQC	47325	HVPENFVQC	47366	KDDCSMMPQ
47244	MEPFHVNKV	47285	LNPAGKIWD	47326	PDRCHSGDW	47367	YENKTDMKF
47245	HKDWHDADL	47286	PEVGSPYPS	47327	YPPPPNQHH	47368	EYPDNQWSQ
47246	PCNCDSWQN	47287	IERHEHNPY	47328	IFPYTDKFI	47369	CDWPIENRA
47247	NEDWEAPEG	47288	HDPPKLGCM	47329	PGDCYWLER	47370	VDHVKRCPV
47248	PIEAMHDGW	47289	NDDCEESQK	47330	HEHRPYFRV	47371	WRTSDKYER
47249	MIPQSEERW	47290	PVVERFNDP	47331	DLFNHDSPH	47372	PKIQRVTET
47250	YKTGDEYGL	47291	PFEEYEQTD	47332	YDDNNWGSA	47373	LDDWMMPVS
47251	PLDKEPMER	47292	VRPEKQRVG	47333	HDSGPTFDM	47374	HYPRNLVEE
47252	HEDHDPAQE	47293	YSDEQGICQ	47334	EIDLEDAQK	47375	NEFQEDKSV

[2071]

47376	DRGQDNHRI	47392	PHDNHRHGL	47408	IEQEDQGYH	47424	GVPIIEPRV
47377	PFNDGGIIE	47393	PIVDWKRQD	47409	EVLDDMKWN	47425	WPKGTKDEA
47378	CYPCLQDKG	47394	YRDMPGQHN	47410	FNPMVYAHK	47426	DRVHPDHLY
47379	CTPNHCGGK	47395	YYPMADPSC	47411	QVDEWGGHE	47427	NEQHNRDQQ
47380	SDTEKMHQQ	47396	PVNHKDQMT	47412	PMLQSRGWN	47428	QEPKETLMI
47381	GPCCPKYNC	47397	EYEWVDEPQ	47413	PYACHHPDF	47429	KRWMEVDTY
47382	TLPADKPKG	47398	PYPGGAVRA	47414	PHTNRECDC	47430	LNNKDPPPV
47383	FTEEDEIRE	47399	RKPWVASRV	47415	WYDDDLLAN	47431	FPHPDSSKH
47384	PDFQKVVSH	47400	NDPWTPISY	47416	PPGCGWYDN	47432	PDHSRLEYP
47385	CNPDHMYCQ	47401	YRFGPFEQE	47417	WQPTHPARF	47433	LEPIQLDMF
47386	HETCGENVC	47402	RRLPDFDDV	47418	EWDDPYGWE	47434	WPPCHITQL
47387	WELAQKHKP	47403	QPPESMVQL	47419	DQPWFGWQT	47435	CPLRDDDPA
47388	YYNEPKAKG	47404	DEWTNKRNS	47420	KHPDQWTMT	47436	EDENMKSTN
47389	CKTDEFDNH	47405	CRFTDDSSV	47421	YIDRDDSEI	47437	PTEPDEQYA
47390	KRPEICEWG	47406	PWNLPWRPG	47422	VDYRTTVKS		
47391	LDPPNSHVN	47407	YESNTKSDV	47423	CDDERMDQP		

[2072] 实施例22

[2073] 如通过机器学习鉴定的具有骨骼肌或心肌组织嗜性的AAV5变体

[2074] 此实施例描述了如通过机器学习方法鉴定的展现骨骼肌组织嗜性或心肌组织嗜性的工程化AAV5变体的优选特性。进行如实施例19和图19至图22中描述的相同生物信息学预处理、机器学习方法和分析以确定展现骨骼肌组织嗜性或心肌组织嗜性的工程化AAV5变体的优选特性。图29显示了一组对模型输出概率有贡献的前20个位置特征。沙普利加和解释(SHAP)值可用于询问特征对模型预测的相对贡献。如图30A至图30B所示，在组织靶向性和非靶向性变体之间进一步比较了这些特征。图30A至图30B显示了排名靠前的预测位置特征的比较。如果发现特征对HGB和RF模型都重要，则选择这些特征。显示了相较于随机2％肌肉变体，前1000个ML预测的骨骼肌或心肌变体中的特征的概述。

[2075] 下文描述了AAV5 VP1衣壳多肽的581至589区域中各位置处的有利生物物理特性和有利氨基酸残基，它们与较高的骨骼肌组织嗜性或心肌组织嗜性概率相关。任何以下机器学习衍生的位置氨基酸偏好描述于表44中。下文表44中列出了从两个ML模型(HGB和RF)中提取的前30个重要的骨骼肌组织嗜性或心肌组织嗜性特征之间共有的位置特征的概述。

[2076]

表44 机器学习衍生的骨骼肌组织嗜性或心肌组织嗜性规则
位置 Xaa1 处低溶解性 Xaa1 选自 D、E、R、K、P、N 或 Q
位置 Xaa1 处低亲疏水性 Xaa1 选自 D、E、R、K、Q、N、Y 或 P
位置 Xaa1 处高表面可接近性 Xaa1 选自 E、R 或 K
位置 Xaa2 处高亲疏水性 Xaa2 选自 V、I、F、L 或 C
位置 Xaa2 处低可突变性 Xaa2 选自 R、V、I、H 或 C
位置 Xaa2 处中等体积 Xaa2 选自 E、V 或 Q
位置 Xaa3 处低溶解性 Xaa3 选自 D、R 或 Q
位置 Xaa4 处低溶解性 Xaa4 选自 D、E、P 或 N
位置 Xaa4 处低电荷 Xaa4 选自 D 或 E
位置 Xaa5 处低氨基酸溶解性 Xaa5 选自 D、E、R、K、N 或 Q
位置 Xaa8 处低溶解性 Xaa8 选自 D、E、K、P 或 N
位置 Xaa8 处高柔性指数 Xaa8 选自 Q、S、P、E 或 D
位置 Xaa8 处高表面可接近性 Xaa8 选自 S、D、P、N、E、R 或 K

[2077] 下文表45提供了表现出骨骼组织嗜性或肌肉组织嗜性并且符合表44中提供的一种或多种规则的变体AAV VP1衣壳多肽的581至589区域的序列。因此，本公开提供了由具有SEQ ID NO:2的工程化AAV5 VP1衣壳多肽组成的rAAV，其中Xaa1至Xaa9区域是SEQ ID NO:25118-SEQ ID NO:26117中的任一者，如表45中所公开。本文还涵盖由工程化AAV5 VP2衣壳多肽和工程化AAV5 VP3衣壳多肽组成的rAAV，所述多肽在对应于AAV5 VP1 581至589区域中的氨基酸的AAV5 VP2(氨基酸残基445至453)和AAV5 VP3(氨基酸残基389-397)中的区域处具有下文表中公开的序列。

[2078] 表45

[2079] 驱动骨骼组织嗜性或肌肉组织嗜性的AAV5 VP1衣壳多肽中581至589区域的序列

[2080]

SEQ ID NO	581-589 序列
25118	RDQCRMEAG
25119	EIREFRASN
25120	DVCDDHKPT
25121	EVRIQCEID
25122	WERCLELNP
25123	PVMCVWHDN
25124	CWECCPDSG
25125	TGCNHNGRF
25126	PVICVWHDS
25127	RDQCRMEYG
25128	KGNPLDGDK
25129	EINEFRASN
25130	YSDAKNMSI
25131	DLLVATEWF
25132	PYGTIEYDE
25133	FPKVSECNW
25134	TQHEDHWLA
25135	NGVGRPGDN
25136	AADIFALND
25137	EDNIQIMMI
25138	AVCDDHKPN
25139	TQPEDHWLS
25140	FPKGSEGNW
25141	WEQCLELNQ
25142	TYEALQNSW
25143	WERCMELNQ
25144	RYQCRMESG
25145	SGDWDFGAW
25146	DLIVATAWF
25147	VFKPLDNRY
25148	DSYTDSWST
25149	PFACDRYEV
25150	SGDLDFGEW
25151	KDGRNPDCL
25152	DLSVATEWF
25153	SADIFAIND
25154	CWACCTDSG
25155	CWECYTDSG
25156	YKGMEKDPV
25157	TYESLQTSW
25158	QTNEEEYGN
25159	YYNHKIVEQ
25160	CWEGCTDSG
25161	AADIFAINA
25162	PYGPIEYDA
25163	EAGCKVWPT
25164	EFGCMYWGY
25165	DVAPMHQVA
25166	EVLIQCKID
25167	NKMIIFDSW
25168	DVADVPDTT
25169	HTIMEWSPT
25170	CGECCTDSG
25171	DVTTTSWIE
25172	ECGCMYWCY
25173	GNEDCNLAF
25174	HHFNVNLDI
25175	TQTEDHWLA
25176	ECGCMYWGC
25177	YSDPKNMYI
25178	EIHMFGPSE
25179	EAMPFRENQ
25180	YSDPMNMSI
25181	DLIVATELF
25182	KEFVQWAAA
25183	ASDMQFWSM
25184	DTSDLQVMK
25185	EVYKNDDGN
25186	NIDEMNCNA
25187	YKGMEKTPV
25188	NTVEKMQTL
25189	ESYPKLTQF
25190	RDAWKWKRY
25191	TIDDKQSWA
25192	ATMDGHMSM
25193	WGRCLELNQ
25194	CQMLCRTPV
25195	VFKTLDKRY
25196	EMAELWDTG
25197	EISSSNLDP
25198	DLIVATEGF
25199	FCDPAHHYE
25200	VMDEYEANP
25201	ETSEHNRMP
25202	ECGCMYWGS
25203	PYGLIEYDE
25204	EIAGNDTYM
25205	VDRLFRVEY
25206	PTEPDEQYA
25207	SVDQDQPGN
25208	GWECCTDSG
25209	SGDWDFGEG
25210	DMENPNAGP

[2081]

25211	VFKPLDKRS	25252	DALEWPMPP	25293	GFVDKMVMD	25334	TADDFQGDE
25212	DVIPRDGAE	25253	DIAPDWKGE	25294	IVHEWEVNH	25335	SQERMAYDG
25213	PVICVWHGN	25254	TKMNIFDSW	25295	EVMQPEASV	25336	EAMQDHTDC
25214	ITAEVTWQE	25255	MVRPTDQPA	25296	QTLDKWDGP	25337	CDWMANVAV
25215	IRTWGCKKQ	25256	ETEDQARSC	25297	DTFMMQSAH	25338	DVKPVQKDW
25216	GSHWYDCNP	25257	IEYWEHETK	25298	DAIPWWWSY	25339	ANEYQYDTM
25217	RLTEDMSDM	25258	TIVEDAPLA	25299	WERCLVLNQ	25340	MVVDRMLGW
25218	FDWMANVSV	25259	YKEMPTILC	25300	NVEEGMAAD	25341	MVVEYLTAG
25219	THESLQNSW	25260	TVLEDSMWA	25301	EAGPWKRDV	25342	MPAYSKLKW
25220	FQMLCRPPV	25261	AVCDWLIAE	25302	TIHDVAMMD	25343	AHEEDEAMQ
25221	EIHMVGQSE	25262	PSSCDWIPV	25303	MNLQIGGKG	25344	ETRPTVAIQ
25222	ETADRIHWI	25263	LSTELHHFG	25304	NNVRQVGNT	25345	VDNMSEGRQ
25223	PYGPIEYYE	25264	CTEGENDEH	25305	QFPGGDDEG	25346	SQESMDYDG
25224	CDHVTKNPT	25265	PMPDGMFGI	25306	VYVEKMGWY	25347	CVDDLASSF
25225	EVHTKDEHM	25266	CTVDGLHFH	25307	TVEHTNHDA	25348	EVKPSMRDD
25226	TVCNHHGRF	25267	DNYMQLNGV	25308	HQYVHDWRQ	25349	LCSQPIPEP
25227	QGDYSWWRY	25268	DVDGEHEVY	25309	NLIVATEWF	25350	ETRPGMSDS
25228	GTLDVPTMF	25269	EVQHELASG	25310	MLMHADMTP	25351	ETSFKIHEH
25229	SVICVWHDN	25270	FTGDFATFA	25311	DSSESLNQN	25352	DTVEGGYIQ
25230	HVDITDNDM	25271	TTDQQCTGP	25312	KCEFYQFCQ	25353	YYEHPINKK
25231	CDHATKHPT	25272	ESVEKSISN	25313	EVCSQWTQF	25354	SIDWINQQG
25232	TYEYLQNSW	25273	AMGEEINYQ	25314	MIDEVWHHW	25355	EITPYAKWP
25233	ATSEIQPFV	25274	VAGDIEWYG	25315	PNMPHAEDQ	25356	EMCERNAGI
25234	EVTEGCPMG	25275	DTQPDTLNP	25316	TFDDFDMCA	25357	DVCTNPWYT
25235	DNVDLDIQH	25276	AITDDNRMQ	25317	PWQQHRKMS	25358	NDYFRVYRK
25236	NVLGFDDGG	25277	TNDQQCTGH	25318	GDPWEEWRK	25359	SQEHMDYDG
25237	SNVMATQNW	25278	VCEELDREW	25319	EQLEDNDMQ	25360	LGPWPQTDQ
25238	QVTDNAAGM	25279	GNGYWMCMS	25320	MIHSQDMFE	25361	AGEFVQANW
25239	AKYLPGLYY	25280	HQYVHTWRQ	25321	QVFHQDDVN	25362	TVKPHDNTE
25240	HNCPWDQQY	25281	SYHMFDYGE	25322	HYKYDANRK	25363	QDEMMMACI
25241	CWECCTYSG	25282	CMHDEMQTS	25323	DMMNIEKMG	25364	VTWEDVYIN
25242	WNTIMTYES	25283	DFNSIDHSG	25324	WEHCYRVSW	25365	NVHESDTMF
25243	PQYVHNWRQ	25284	DASERNRWQ	25325	DILFWEDDD	25366	NIAGYEWWM
25244	HTGMEWSPT	25285	RHNPHDACF	25326	MTDEHLQYT	25367	WLLGMWNTG
25245	WDYSMAKKD	25286	QYVDQENCN	25327	LSVDRDYFL	25368	VINPHDRSV
25246	WKHCEREIN	25287	DASPQCDRD	25328	DVQQIHNKE	25369	TVIPYEMFD
25247	SIFDYDTFQ	25288	EIKPAAEIT	25329	MITNNCNIP	25370	PSTNLTDSA
25248	QTNGEEDGN	25289	DALTRVDGD	25330	NISWNESHD	25371	SVYDDFTST
25249	EAALENVNM	25290	DYQWPDVEE	25331	ETVHPSASN	25372	NIHGVNLQG
25250	WQKPINCEC	25291	DIVENDSLQ	25332	TKMIIFDAW	25373	DWEKVPCMK
25251	EIRMFGQSE	25292	VIVESEEYM	25333	IIAQWDQVN	25374	TQQVADYTK

[2082]

25375	FIGSSDKQS	25416	AHSEIMCDQ	25457	SQISVEFNC	25498	MVYEATELD
25376	CLDELVEED	25417	VTKLAEDVY	25458	EFWSMWVSD	25499	MTVDLYEVL
25377	ATLDLDSHA	25418	PEFMWCVPW	25459	DADELRSAK	25500	EVDKGLAQA
25378	DTQQEQSQC	25419	DTTTDMWVA	25460	EMRTYNKDY	25501	DSELLPEAH
25379	TAEIFAVND	25420	DVFIENQHW	25461	SSTELFQAG	25502	AVYDHAPQN
25380	DVVFNLWAP	25421	NVSLQSIHK	25462	TEPETYPAT	25503	WHHTHSHNC
25381	GSMSSTVPH	25422	ESGMEQISQ	25463	VMDDESHRG	25504	DIMPQWAFM
25382	NVVMLDQSY	25423	PNGCYTDCE	25464	DACEEGSSM	25505	DIRHYNISR
25383	QQDMIVTKQ	25424	AVKEFHDCF	25465	WYTQDIEKE	25506	VTLDGNRMG
25384	DFMFLISTH	25425	WVADHGDGS	25466	IGDHDCQEQ	25507	IIETQCQAI
25385	FQVDTFMFE	25426	EIGMTQPCT	25467	DVGNTYNHW	25508	EASHIFTGE
25386	IFEWGDATY	25427	WLYTSNFKD	25468	QSYTTMYQC	25509	QIAEACAQH
25387	LPPGIEWEH	25428	DDMEDQQQY	25469	ISGEIDYME	25510	ETTCNMGAK
25388	MDRLAEFKD	25429	EATTTSEKD	25470	ACCEFELTQ	25511	VFDMTNLRV
25389	DTVVKNMYH	25430	QTQDTMNLH	25471	TANEWETME	25512	LSPWHQTDQ
25390	EQIPVTWTT	25431	QVHHNNVDM	25472	AEKMDEIWK	25513	ECSHQDRSN
25391	WHMMSRDGN	25432	SVVLCNLCQ	25473	VTKPKWSNF	25514	TMMDAEVGL
25392	STEEGKGWF	25433	DAPDFMTFC	25474	VVVLSDFGE	25515	YLTELIQKT
25393	DVMMALATQ	25434	FPRIEGQGE	25475	GSHWYDYNQ	25516	EIVELRQTR
25394	SAEPQDANV	25435	PQLFKGPVE	25476	GMYESNQGS	25517	FPMVWFVLF
25395	SVTMDEEAH	25436	TTEWLDFDA	25477	DVDCVWHEE	25518	SNEWPTMIH
25396	AVPTSQAEY	25437	FEWVDWGYV	25478	LVYDPHFTS	25519	QISEHAGRS
25397	PCAMHNWCC	25438	MRNYPDARF	25479	NDVDLWHVD	25520	VVQEKSINW
25398	EVYVESQNE	25439	EIGSCPEWC	25480	AVEYAGWEK	25521	DAQPQRNGS
25399	CIDTYDIDS	25440	DNPPGRRKP	25481	FYQADIGWA	25522	EVWTSCMTP
25400	SNMELYCRV	25441	TTTDSPMHF	25482	NTEAEDHQD	25523	EATNFTIGR
25401	GTHEHEYGA	25442	SCDESEAYY	25483	RCDCKVHEF	25524	VTEPHSFTV
25402	DTLPVQVYQ	25443	WEHVLDYAK	25484	HLPCGWPAV	25525	LHEVQTILH
25403	NIFDTNVLT	25444	DASFVDQDE	25485	IPVHFDWFF	25526	EVVTQMEHL
25404	VSEDKEAIS	25445	GVNEAEFTI	25486	NVEGIVTAC	25527	WRVWQNLPQ
25405	VMHDIPNFF	25446	MIEVTSHST	25487	DVGNMAAIE	25528	TSEYMATGE
25406	ETMSRHEGM	25447	AVQHTEFNL	25488	ETLQAQHCD	25529	KDWRFEDEK
25407	ATEIDHWQW	25448	DIQGMLEAA	25489	NTADIHMKE	25530	IQAEQMSGC
25408	LSMDVTMAG	25449	DVVMTAMDY	25490	SPLEVWTPA	25531	DAKNWLDKN
25409	AIHWMDWCQ	25450	FKDWDHHTV	25491	KFVWDTWAW	25532	MPNFAWEAQ
25410	TVHFFSEET	25451	NVEHNNMGA	25492	KCFLCGSTT	25533	DLLNYGMFH
25411	EMAGPFVTT	25452	HVHEIMEEE	25493	WECADRRPT	25534	EQPFDEFKA
25412	DIHVNNFWN	25453	SVMTNDDID	25494	VIDDYFQMC	25535	HATGLVNVM
25413	EKQMDGINH	25454	GPIEFGWQI	25495	GATCTIVRY	25536	KEEFNQEPA
25414	QEDGDFSTF	25455	EFANTEHTS	25496	TIDCYAAMD	25537	EAYTRVGYE
25415	EITQTYDHL	25456	VTEDMIMNT	25497	PYCQQAANW	25538	MQEEMGPRK

[2083]

25539	DFTNQDHDN	25580	DTGSFTSTP	25621	DCSHEAHRY	25662	DHEMPESTF
25540	EVRANCESY	25581	LTEATPKWT	25622	INPELSEGC	25663	VKCFPCQLE
25541	DNTDPDGSW	25582	GDWSERSIT	25623	DTSSWGVIP	25664	NEPELKLQP
25542	TIGPSNFST	25583	TTCPMEKLN	25624	HYRDNKSLP	25665	DTCQAMHGD
25543	DNANFSGAD	25584	LCVEDNIQQ	25625	MNSEPVTET	25666	EWHFVNIGT
25544	EYMERDMSI	25585	DVAQYGLGN	25626	QMEWLAAGP	25667	MVQQYNGSE
25545	FDKLGTMWG	25586	TADMMPMNP	25627	IFTTADGEF	25668	DIERFSTVC
25546	CVDQLDDVT	25587	TDESEFESS	25628	MFPSEAKEI	25669	NIAAHSPGP
25547	ESGGYMRMC	25588	HDMSYHDPI	25629	DVKPGLFYT	25670	HNPCWPFMI
25548	VQTMPKPIY	25589	TVVERHEHG	25630	ETDQPQLAA	25671	IVDMPVRHP
25549	TYVDNNHMV	25590	CVETQWRAD	25631	DVLLYMSAT	25672	SADEDKCGW
25550	EIVVEFEDQ	25591	SFDIESARQ	25632	DIIFKLEAG	25673	SFPEAWPIA
25551	IPDCMNLSQ	25592	EVEVVKEFQ	25633	CDAGLISKM	25674	MIVSQQWVW
25552	HTVMEGSPT	25593	NTQTTDIHG	25634	MFTGPDVRH	25675	HIEYWDQRD
25553	PIIPIAEHV	25594	PQKGDDWLY	25635	GVMESQYIS	25676	SLMYEDDYP
25554	DVQGSGQWV	25595	PCEQMPYIM	25636	DCTLQPAMY	25677	DAAIYVMLD
25555	HTNQQANSP	25596	DEYDNVWQA	25637	PEAQRVAGI	25678	MQKNPPKRV
25556	STDDEYSLA	25597	DMRDSHAWP	25638	DVEMCQKMS	25679	TMVDWELGK
25557	TAVEWHYFD	25598	YTCHEEWPL	25639	WHGCYKIMF	25680	QTGLNEYAP
25558	MAPQSLWNM	25599	ESGLAPMNL	25640	VHMPYTEET	25681	DDMWCGGLY
25559	SSGEMDHVY	25600	PMDHEKRYH	25641	ENGTHQYND	25682	VIYFVGLIA
25560	TIDDYGPIS	25601	VPEPDDKPY	25642	QCAYSWWRY	25683	HHACGTREH
25561	DQPCKMWAR	25602	AIIWTTDQP	25643	CEWLGTCNH	25684	DISSNIYDI
25562	TTEAGTALE	25603	QSMPICEGH	25644	DTCYNNWKD	25685	AIECYKMGS
25563	ETDLKCHSY	25604	DAGPLSMRS	25645	LSAQRLVGD	25686	EESNRGYQH
25564	ACVEQEKAD	25605	VTEVQHMAK	25646	NLDMYSISQ	25687	MTEVVVSSM
25565	SMNEMYQYP	25606	WEGFYMLTL	25647	SILFYSETH	25688	HEGHPHLSC
25566	GIDTTTQAF	25607	MIMCMDCFP	25648	MDYDGMAVY	25689	ESCTGEWWA
25567	FPQMNEWEC	25608	DIAGMPHIS	25649	LCSDEVCWT	25690	MTDTVQEYA
25568	DVHTTCYHY	25609	GVDAGLVSE	25650	TIEGKSLYC	25691	PRNGPYCDI
25569	SQPNPEINW	25610	TIDVPMQYA	25651	ETSMPMEHT	25692	NKMTEARKH
25570	EVMLSKMEY	25611	MELEGHNLP	25652	VEINAKWGN	25693	DNGQVEETA
25571	IVDHCNMYP	25612	DEWYRSKED	25653	IEPWDCQIP	25694	DAYVQQPSF
25572	EFHQFGNHM	25613	VWDCPKWRH	25654	EEPEFTGRG	25695	FVYEWSNLV
25573	TVQMGEWGY	25614	SYDDTVICW	25655	ENDCRIRMA	25696	EVRGQGSHY
25574	NMSDGHEVC	25615	IVDNKCYYN	25656	CVMHVQCDH	25697	WFDCPPCCA
25575	DDNEYHVWS	25616	DSTPIMQRC	25657	AIATGWQDM	25698	WREFSEWTF
25576	DNWMGQDDK	25617	YWDCQHSYM	25658	EYHGYLEHT	25699	QLFAPDMDW
25577	DEMDAQMGA	25618	TKMIIFGSW	25659	IQWVHGCLF	25700	TYNDFYTAF
25578	TNDWMRLRV	25619	SVVIRQLVD	25660	NDCSDCQTL	25701	EVGSMPQDW
25579	IHYDYNTIC	25620	EVCIHEQCM	25661	AGMEIIVQC	25702	SKELGSVET

[2084]

25703	DVTISMQAG	25744	LAPCSKIYC	25785	STIPTPIWL	25826	ITSVFDNKA
25704	KYNFNLNKY	25745	FMWPPQIVV	25786	VQVQDWLPE	25827	DVITHWIGQ
25705	VSDWPQIKA	25746	HMQEPSIAS	25787	DWNCAALKP	25828	QQASSSGGF
25706	DTKSRQWDC	25747	MEPVMHYLN	25788	EVGPSWEQR	25829	EGDDLDRHY
25707	AYEGECIHY	25748	YCKCPTCNT	25789	MHDFLTSAQ	25830	MPELKFSYC
25708	MAEFEQWLE	25749	ADGPVLQMF	25790	NLPCSQQAM	25831	HQTHQHNIW
25709	CELHELFDG	25750	EVSSCDDFT	25791	DQRCNKAPG	25832	AFGTLEHEP
25710	TAEYIDISY	25751	CNEEMEFSC	25792	PLAFSPVQR	25833	TKPTHDVGS
25711	EIAQFSGVC	25752	MLTESQQSM	25793	ETDLAMKAE	25834	QVMQNESER
25712	DWNEEMAHL	25753	WPGFEKLPS	25794	NVEWRQACN	25835	SDESNWFGQ
25713	WQEYRMTDQ	25754	SPYCDDRGN	25795	ETTTCWQYF	25836	FSADVNNYD
25714	SNENICPTC	25755	QYCLTEASH	25796	AMPYEEVPA	25837	LWCCTDTEC
25715	LPVDVLVKT	25756	TSEWLWVDS	25797	ESAQFVAYS	25838	SMEEWGCES
25716	SADWRHKSP	25757	DNMVWRGAH	25798	VLDCAMCNP	25839	MHFCIWMRM
25717	EDGQSMCVS	25758	CDNDYLTAT	25799	AEPHFQNKS	25840	AHNESDDCG
25718	MHHEYAMAG	25759	VIEYNSNIY	25800	MNESFPTWN	25841	SHQFVQPNN
25719	DDHDINGGP	25760	HVSAIEKAQ	25801	EVQCDQIHP	25842	AIHMFGQSE
25720	MTEDQTWWQ	25761	QVATFECKQ	25802	MDQCFSYNQ	25843	PRKCSNEEV
25721	WNEPWYVRP	25762	LHIHRDVTI	25803	EVGSQRTAY	25844	HFPLGWLWL
25722	ECQQDSWHY	25763	NHEETEFGY	25804	DHSPVKSAY	25845	QQLELDGSL
25723	EVQRQMSNK	25764	MDSCYDYQS	25805	YICQQDLIM	25846	INHGWEKVA
25724	DMHSFKWWS	25765	EVLGCNYWP	25806	KIDHEQDCY	25847	HQHCAEFCH
25725	DHGEIPAFT	25766	STEDWATLH	25807	LPVEPIDCE	25848	DCSHNGMRE
25726	EYDKMMYAE	25767	CQTDKNATE	25808	GLMVGQWPH	25849	WINQQTLQC
25727	DADREDFWS	25768	ALEQEMVKF	25809	AHIWDTLIP	25850	QERMFEMTA
25728	QQEAEHKGV	25769	QVTFASASK	25810	ATIVGDSGN	25851	QINMSMHES
25729	AHEVTFVRP	25770	DDEVYVWGQ	25811	TANDYGPTE	25852	MVTARCWLS
25730	QPIDVSAAI	25771	QVMPRSIER	25812	NVELPKIHE	25853	PLWFEPALN
25731	DIYLKCFIK	25772	AVHTFQMWD	25813	MLPDHWVIT	25854	EIGYRKVYN
25732	EIFLENRFP	25773	LVGNADHGQ	25814	HPFNEAVKH	25855	ESVMEKSMF
25733	FTKTNDHNT	25774	NVGLPSYFA	25815	IISSTTSTG	25856	NVNHFGMGD
25734	WQALWCLVH	25775	EDYSGTKWG	25816	IKEHECRRI	25857	SVNVDNAEM
25735	AFKALDVQC	25776	TVPLKTYGG	25817	AVHQINVAA	25858	NEQFQCDVH
25736	DIHSTNWLA	25777	TYMPMDCKW	25818	LVMEALPMA	25859	NPNCTKQQY
25737	NLTDKCTAD	25778	FIVQQSIWK	25819	GQIWAAMRR	25860	QMGPHERHA
25738	NSEWWSSVL	25779	NEQHYEYWK	25820	MWVCKLFVC	25861	ETRMHMDAL
25739	TNEIMWCSP	25780	LDVTPQPEG	25821	PIHHGEHVY	25862	LNMYEASLS
25740	MVAPSQALQ	25781	ETSLPHWQG	25822	NHQCNGLAQ	25863	MKVTDKLDD
25741	APGEYMDVE	25782	DSKPVMMAN	25823	TCEEPYNFA	25864	ADETCAKKF
25742	ECAKTEWNF	25783	VYRHFDMTP	25824	LVHEKWEQF	25865	INQWLTCRW
25743	KDLHDAIHN	25784	DSLAAWSDK	25825	ALAKFRCWS	25866	HNVGDMFCN

[2085]

25867	SNQPQGQYC	25908	ICLHDHFRN	25949	PAFLGACGF	25990	ELEGMSRRE
25868	LEKVINVLH	25909	EMKGTLYAY	25950	ESEEHRVMS	25991	VDATTGDYQ
25869	SFVHMFQGP	25910	APDFRMITR	25951	CQGCVPFKG	25992	TVGFGLKAS
25870	DASYYSPVP	25911	PQWMPVNQI	25952	FIEFDAVTW	25993	WMEASITGP
25871	VNDDIYQQG	25912	EFKTQQFDP	25953	TTCSSNREW	25994	WTDGWHTMC
25872	LEMMGWPIG	25913	VNEWLDSIY	25954	VCNELQKYF	25995	SVTLGMKPN
25873	MEHALCTRT	25914	GSIQSVMWE	25955	VPCYHKGRW	25996	GHVQESWNC
25874	ECSCVKRGY	25915	DFMCDQFVC	25956	DVKCEPDRG	25997	STVAADWDH
25875	ESGFPNPYH	25916	PCSDSTELD	25957	LICWYQQGN	25998	ESEIAISWN
25876	STDWSPSLG	25917	DQECWQSSQ	25958	EMQTHYAAR	25999	NETLDSKHF
25877	ETDYIQMGA	25918	EEPLPSDQK	25959	AETDHYIPT	26000	TFMWLDRVG
25878	DIVTVIKRT	25919	QTMKLEHLK	25960	PQSDDWAMP	26001	TTDSWMDVE
25879	LAWDNQQFN	25920	PGTHWWWDC	25961	NHQWTQRRY	26002	DEAEEHGCT
25880	DTDHRNKFP	25921	MTVSWDSWK	25962	FEGNYTATK	26003	FTPPKLHYK
25881	MSNDRCSMG	25922	DAWKPHQFY	25963	CNCPDQHPD	26004	QYLTAEFTS
25882	FRQSVPHWA	25923	THEEYQLGS	25964	IFQCLYTGT	26005	MSTNNCNIS
25883	QVGQEAKNC	25924	TVIQTECNT	25965	VETYQINDT	26006	EAWSCGIYD
25884	YLEWDSWRC	25925	SNTFIDWQA	25966	FFEWLLDTG	26007	NYDTLLETH
25885	MSDFPISTP	25926	LYDWSCAQS	25967	SDALGPEWW	26008	NVEYVNWHW
25886	MQNLSSVTE	25927	WREPNCHGL	25968	MTNCKLHEQ	26009	FVMPISAQG
25887	STEYRNALN	25928	TMAEMCGSP	25969	IYEACFWWL	26010	SQESHMTWH
25888	VIVGNNVGI	25929	SETDWQRKC	25970	WCMQEGTMS	26011	GIVALDNAD
25889	DGPCDQMTN	25930	QVIIGYCES	25971	EIRLEHDGG	26012	TVMWKDYID
25890	MDEYFCGFT	25931	GFDECMFWY	25972	HVELSLWNQ	26013	PWCEDPGCS
25891	EPAQSQHGF	25932	PCTQIGEIY	25973	CWMEFMAGN	26014	TVIYVTDDA
25892	SFDYVMWGC	25933	FMDFFGMAS	25974	AWDYWNSQW	26015	TQDCSTMPS
25893	DVFKEHYHQ	25934	DQMLDNTAY	25975	NNITTPDAY	26016	SMCWLENGN
25894	HFVQPACEH	25935	SQDCQEVGG	25976	SVTMHSMEL	26017	MVDLETTYV
25895	DVANLKHHV	25936	SDHMSERVD	25977	TYDHIKFAP	26018	DNTQKQNDP
25896	VSEQIWNAD	25937	DVLQREEIC	25978	LQQMENASP	26019	NVDRCDSWQ
25897	IVPCLWPIP	25938	VIALYSLAY	25979	CTDDKMCPV	26020	TYEPCSTSH
25898	GHQEFKGSE	25939	EWNYGGLKC	25980	IDSCNTSLN	26021	DFHCQTERQ
25899	ANCSSEIYV	25940	EVQYGKYWA	25981	CKEMRMDQW	26022	TTAPMEMSL
25900	DSTHDIAHS	25941	DMTMYSNAV	25982	EPEHGAMDK	26023	NAVEGQSHK
25901	TTDRQCTGH	25942	VTNCHTFNN	25983	MVLPYQITN	26024	DNYSIKVGQ
25902	VISAMNVYP	25943	SIPCNHCFE	25984	ETHLINPGI	26025	FMMFHMGLD
25903	AHQFIQWLH	25944	MTQGYPSLM	25985	CDHDTKHPT	26026	ETRFDDRRA
25904	ETYRYGSQR	25945	DIASVRQTE	25986	CSAEIIDTC	26027	NTGWQEAER
25905	KGNPLAGDT	25946	IVDSASVCQ	25987	EVGFYMVPR	26028	WVFMSWCDV
25906	NTMQMPMIG	25947	AIEAKMTQV	25988	AIWSIEHES	26029	PCLITVAPH
25907	MYEVLNNDQ	25948	ALAKFQGWS	25989	MVNLGYAEC	26030	DIAVIGAGA

[2086]

26031	EATLITYCS	26053	FVCELQSAN	26075	DTAKVVELC	26097	IIDNVLCNG
26032	TTEDMQCDM	26054	ETKSHPVMV	26076	ETSCGTLMC	26098	MKMCNCPYQ
26033	ITMPRVRPT	26055	ENDEASICM	26077	EQQWIAPKH	26099	MILVQQLQN
26034	LFVMPAAIE	26056	VVHMFNDQM	26078	EVKQATIWC	26100	PSTGNIQFM
26035	ALEHTLFGT	26057	NTFILETFD	26079	TYDPFWMQH	26101	QHPGFGLVI
26036	MGCDVMKAS	26058	HDFNWVTKE	26080	THDTTTFAF	26102	IWWEHMGPA
26037	DSWPTLMHD	26059	WFVTQGAYM	26081	WVFQADWNT	26103	QWRWHKSAC
26038	CYHEECDGN	26060	FWATSHWGY	26082	WVTPQCIHL	26104	DSESRLCED
26039	ITTPQTQSH	26061	IPYTEKADI	26083	LGHDVNGHH	26105	ETSAEFEQV
26040	IHDWYEVGS	26062	HDEEEKNKA	26084	CHELIKKQW	26106	FLGSMWPRD
26041	SMQMTSHYH	26063	WLHYVEDTL	26085	SNWSKCAPT	26107	CVTLEQSSF
26042	TVAVFSIAM	26064	NWAFNARHD	26086	LNDYTHHQA	26108	HMTAMCLAG
26043	QGDEINEQA	26065	CSGYWMCMS	26087	ACEQENIQH	26109	NVELRHHGY
26044	PCSQYYPSC	26066	DFVLTGTIN	26088	LYDYMHMGS	26110	MALDQRSKP
26045	PFAAVSMVF	26067	DVNQNAKYD	26089	EVPLHTVCQ	26111	DFIYDGEDM
26046	MQLTNMEIE	26068	ITDIFHWSC	26090	VHGCGKQYH	26112	FDRHDKSCC
26047	VTESIDYVH	26069	HMRTMDAFE	26091	MHMDQPTMC	26113	TIGQHCADF
26048	TIAPYTYMD	26070	ESIQKHSVC	26092	VKFSNEMKW	26114	HVANGTDAC
26049	DIQLTHIHG	26071	MSEMEKQQF	26093	MELHKCNSP	26115	DSHTRRYDT
26050	MYMTVQQNV	26072	LMDEQPMWC	26094	ELHGTYNSC	26116	MIPQPVEPF
26051	PFHERGLMQ	26073	HWDVQVNNA	26095	SEAEMDWIG	26117	TVAHHFVET
26052	TYVNQGMGP	26074	EQNHAQLMP	26096	AFRTMDEVH		

[2087] 实施例23

[2088] 用衣壳包裹治疗性有效载荷的AAV5源性病毒体治疗神经系统疾病或疾患

[2089] 此实施例描述了用具有本文公开的任何一种工程化CNS嗜性变体AAV5 VP衣壳多肽的变体AAV5源性病毒体治疗神经系统疾病或疾患，其中变体AAV5病毒体衣壳包裹治疗性有效载荷。将编码AAV Rep、AAV5源性变体Cap和辅助蛋白的多核苷酸序列以及治疗性有效载荷在细胞中转染以产生变体AAV5病毒体，其中编码变体AAV5Cap的多核苷酸序列在从对应于VP1衣壳多肽中的581的位置至对应于589的位置的区域中包含至少一个突变，并且其中编码变体AAV5Cap的多核苷酸序列编码SEQ ID NO:7118-SEQ ID NO:10117中任一者的氨基酸序列。治疗性有效载荷是靶向由与神经系统疾病或疾患有关的基因编码的mRNA的指导RNA，或转基因。将衣壳包裹有效载荷的变体AAV5病毒体施用于受试者。受试者是人类或非人类动物。施用途径是全身施用途径。全身施用途径是静脉内施用。与衣壳包裹治疗性有效载荷的野生型AAV5病毒体相比，衣壳包裹治疗性有效载荷的变体AAV5病毒体在施用于受试者后表现出增强的CNS组织嗜性。在施用于受试者时，与衣壳包裹治疗性有效载荷的野生型AAV5病毒体相比，施用较低剂量的衣壳包裹治疗性有效载荷的变体AAV5病毒体。在施用于受试者后，神经系统疾病或疾患的至少一种症状得到缓解或受试者得到治愈。

[2090] 实施例24

[2091] 用衣壳包裹治疗性有效载荷的AAV5源性病毒体治疗阿尔茨海默氏病

[2092] 此实施例描述了用具有本文公开的任何一种工程化CNS嗜性变体AAV5 VP衣壳多肽的变体AAV5源性病毒体治疗阿尔茨海默氏病，其中变体AAV5病毒体衣壳包裹治疗性有效载荷。将编码AAV Rep、AAV5源性变体Cap和辅助蛋白的多核苷酸序列以及治疗性有效载荷

在细胞中转染以产生变体AAV5病毒体，其中编码变体AAV5Cap的多核苷酸序列在从对应于VP1衣壳多肽中的581的位置至对应于589的位置的区域中包含至少一个突变，并且其中编码变体AAV5Cap的多核苷酸序列编码SEQ ID NO:7118-SEQ ID NO:10117中任一者的氨基酸序列。每个这种变体都对所有非CNS组织去靶向，包括对心脏组织去靶向。治疗性有效载荷是靶向由与阿尔茨海默氏病有关的基因编码的mRNA的指导RNA，或转基因。指导RNA所靶向的mRNA是由编码淀粉样前体蛋白(APP)、α-突触核蛋白(SNCA)、Tau蛋白(MAPT)或载脂蛋白E(ApoE)的基因编码的mRNA。转基因是编码淀粉样前体蛋白(APP)、α-突触核蛋白(SNCA)、Tau蛋白(MAPT)或载脂蛋白E(ApoE)的基因。将衣壳包裹有效载荷的变体AAV5病毒体施用于受试者。受试者是人类或非人类动物。施用途径是全身施用途径。全身施用途径是静脉内施用。与衣壳包裹治疗性有效载荷的野生型AAV5病毒体相比，衣壳包裹治疗性有效载荷的变体AAV5病毒体在施用于受试者后表现出增强的CNS组织嗜性。在施用于受试者时，与衣壳包裹治疗性有效载荷的野生型AAV5病毒体相比，施用较低剂量的衣壳包裹治疗性有效载荷的变体AAV5病毒体。在施用于受试者后，阿尔茨海默氏病的至少一种症状得到缓解或受试者得到治愈。

[2093] 实施例25

[2094] 用衣壳包裹治疗性有效载荷的AAV5源性病毒体治疗帕金森氏病

[2095] 此实施例描述了用具有本文公开的任何一种工程化CNS嗜性变体AAV5 VP衣壳多肽的变体AAV5源性病毒体治疗帕金森氏病，其中变体AAV5病毒体衣壳包裹治疗性有效载荷。将编码AAV Rep、AAV5源性变体Cap和辅助蛋白的多核苷酸序列以及治疗性有效载荷在细胞中转染以产生变体AAV5病毒体，其中编码变体AAV5Cap的多核苷酸序列在从对应于VP1衣壳多肽中的581的位置至对应于589的位置的区域中包含至少一个突变，并且其中编码变体AAV5Cap的多核苷酸序列编码SEQ ID NO:7118-SEQ ID NO:10117中任一者的氨基酸序列。治疗性有效载荷是靶向由与帕金森氏病有关的基因编码的mRNA的指导RNA，或转基因。指导RNA所靶向的mRNA是由编码富亮氨酸重复激酶2(LRRK2)的基因编码的mRNA。转基因是LRRK2;GBA1;GBA1+α-突触核蛋白;AADC;GDNF;神经秩蛋白;GAD;NTN;hFOXG1;hKCNQ2;hFMR1;抗Tau/miRNA;EPM2A或EPM2B。将衣壳包裹有效载荷的变体AAV5病毒体施用于受试者。受试者是人类或非人类动物。施用途径是全身施用途径。全身施用途径是静脉内施用。与衣壳包裹治疗性有效载荷的野生型AAV5病毒体相比，衣壳包裹治疗性有效载荷的变体AAV5病毒体在施用于受试者后表现出增强的CNS组织嗜性。在施用于受试者时，与衣壳包裹治疗性有效载荷的野生型AAV5病毒体相比，施用较低剂量的衣壳包裹治疗性有效载荷的变体AAV5病毒体。在施用于受试者后，帕金森氏病的至少一种症状得到缓解或受试者得到治愈。

[2096] 实施例26

[2097] 用衣壳包裹治疗性有效载荷的AAV5源性病毒体治疗Tau蛋白病

[2098] 此实施例描述了用具有本文公开的任何一种工程化CNS嗜性变体AAV5 VP衣壳多肽的变体AAV5源性病毒体治疗Tau蛋白病，其中变体AAV5病毒体衣壳包裹治疗性有效载荷。将编码AAV Rep、AAV5源性变体Cap和辅助蛋白的多核苷酸序列以及治疗性有效载荷在细胞中转染以产生变体AAV5病毒体，其中编码变体AAV5Cap的多核苷酸序列在从对应于VP1衣壳多肽中的581的位置至对应于589的位置的区域中包含至少一个突变，并且其中编码变体

AAV5Cap的多核苷酸序列编码SEQ ID NO:7118-SEQ ID NO:10117中任一者的氨基酸序列。治疗性有效载荷是靶向由与Tau蛋白病有关的基因编码的mRNA的指导RNA,或转基因。指导RNA所靶向的mRNA是由编码Tau蛋白的基因(MAPT)编码的mRNA。转基因是MAPT。将衣壳包裹有效载荷的变体AAV5病毒体施用于受试者。受试者是人类或非人类动物。施用途径是全身施用途径。全身施用途径是静脉内施用。与衣壳包裹治疗性有效载荷的野生型AAV5病毒体相比,衣壳包裹治疗性有效载荷的变体AAV5病毒体在施用于受试者后表现出增强的CNS组织嗜性。在施用于受试者时,与衣壳包裹治疗性有效载荷的野生型AAV5病毒体相比,施用较低剂量的衣壳包裹治疗性有效载荷的变体AAV5病毒体。在施用于受试者后,Tau蛋白病的至少一种症状得到缓解或受试者得到治愈。

[2099] 实施例27

[2100] 用衣壳包裹治疗性有效载荷的AAV5源性病毒体治疗α-1抗胰蛋白酶缺乏症(AATD)

[2101] 此实施例描述了用具有本文公开的任何一种工程化CNS嗜性变体AAV5 VP衣壳多肽的变体AAV5源性病毒体治疗AATD,其中变体AAV5病毒体衣壳包裹治疗性有效载荷。将编码AAV Rep、AAV5源性变体Cap和辅助蛋白的多核苷酸序列以及治疗性有效载荷在细胞中转染以产生变体AAV5病毒体,其中编码变体AAV5Cap的多核苷酸序列在从对应于VP1衣壳多肽中的581的位置至对应于589的位置的区域中包含至少一个突变,并且其中编码变体AAV5Cap的多核苷酸序列编码SEQ ID NO:43438-SEQ ID NO:46437中任一者的氨基酸序列。治疗性有效载荷是靶向由与AATD有关的基因编码的mRNA的指导RNA,或转基因。指导RNA所靶向的mRNA是由编码SERPINA1的基因编码的mRNA。转基因是SERPINA1。将衣壳包裹有效载荷的变体AAV5病毒体施用于受试者。受试者是人类或非人类动物。施用途径是全身施用途径。全身施用途径是静脉内施用。与衣壳包裹治疗性有效载荷的野生型AAV5病毒体相比,衣壳包裹治疗性有效载荷的变体AAV5病毒体在施用于受试者后表现出增强的CNS组织嗜性。在施用于受试者时,与衣壳包裹治疗性有效载荷的野生型AAV5病毒体相比,施用较低剂量的衣壳包裹治疗性有效载荷的变体AAV5病毒体。在施用于受试者后,AATD的至少一种症状得到缓解或受试者得到治愈。

[2102] 实施例28

[2103] 用衣壳包裹治疗性有效载荷的AAV5源性病毒体治疗囊性纤维化

[2104] 此实施例描述了用具有本文公开的任何一种工程化CNS嗜性变体AAV5 VP衣壳多肽的变体AAV5源性病毒体治疗囊性纤维化,其中变体AAV5病毒体衣壳包裹治疗性有效载荷。将编码AAV Rep、AAV5源性变体Cap和辅助蛋白的多核苷酸序列以及治疗性有效载荷在细胞中转染以产生变体AAV5病毒体,其中编码变体AAV5 Cap的多核苷酸序列在从对应于VP1衣壳多肽中的581的位置至对应于589的位置的区域中包含至少一个突变,并且其中编码变体AAV5 Cap的多核苷酸序列编码SEQ ID NO:16118-SEQ ID NO:19117中任一者的氨基酸序列。治疗性有效载荷是靶向由与囊性纤维化有关的基因编码的mRNA的指导RNA,或转基因。指导RNA所靶向的mRNA是由编码CFTR的基因编码的mRNA。转基因是CFTR。将衣壳包裹有效载荷的变体AAV5病毒体施用于受试者。受试者是人类或非人类动物。施用途径是全身施用途径。全身施用途径是静脉内施用。与衣壳包裹治疗性有效载荷的野生型AAV5病毒体相比,衣壳包裹治疗性有效载荷的变体AAV5病毒体在施用于受试者后表现出增强的CNS组织嗜性。在施用于受试者时,与衣壳包裹治疗性有效载荷的野生型AAV5病毒体相比,施用较低

剂量的衣壳包裹治疗性有效载荷的变体AAV5病毒体。在施用于受试者后，囊性纤维化的至少一种症状得到缓解或受试者得到治愈。

[2105] 实施例29

[2106] 用衣壳包裹治疗性有效载荷的AAV5源性病毒体治疗杜兴氏肌营养不良

[2107] 此实施例描述了用具有本文公开的任何一种工程化CNS嗜性变体AAV5 VP衣壳多肽的变体AAV5源性病毒体治疗杜兴氏肌营养不良，其中变体AAV5病毒体衣壳包裹治疗性有效载荷。将编码AAV Rep、AAV5源性变体Cap和辅助蛋白的多核苷酸序列以及治疗性有效载荷在细胞中转染以产生变体AAV5病毒体，其中编码变体AAV5Cap的多核苷酸序列在从对应于VP1衣壳多肽中的581的位置至对应于589的位置的区域中包含至少一个突变，并且其中编码变体AAV5Cap的多核苷酸序列编码SEQ ID NO:25118-SEQ ID NO:26117、SEQ ID NO:28991-SEQ ID NO:31990或SEQ ID NO:13118-SEQ ID NO:16117中任一者的氨基酸序列。治疗性有效载荷是靶向由与杜兴氏肌营养不良有关的基因编码的mRNA的指导RNA，或转基因。指导RNA所靶向的mRNA是由编码DMD的基因编码的mRNA。转基因是DMD。将衣壳包裹有效载荷的变体AAV5病毒体施用于受试者。受试者是人类或非人类动物。施用途径是全身施用途径。全身施用途径是静脉内施用。与衣壳包裹治疗性有效载荷的野生型AAV5病毒体相比，衣壳包裹治疗性有效载荷的变体AAV5病毒体在施用于受试者后表现出增强的CNS组织嗜性。在施用于受试者时，与衣壳包裹治疗性有效载荷的野生型AAV5病毒体相比，施用较低剂量的衣壳包裹治疗性有效载荷的变体AAV5病毒体。在施用于受试者后，杜兴氏肌营养不良的至少一种症状得到缓解或受试者得到治愈。

[2108] 实施例30

[2109] 用衣壳包裹治疗性有效载荷的AAV5源性病毒体治疗额颞叶痴呆

[2110] 此实施例描述了用具有本文公开的任何一种工程化CNS嗜性变体AAV5 VP衣壳多肽的变体AAV5源性病毒体治疗额颞叶痴呆(FTD)，其中变体AAV5病毒体衣壳包裹治疗性有效载荷。将编码AAV Rep、AAV5源性变体Cap和辅助蛋白的多核苷酸序列以及治疗性有效载荷在细胞中转染以产生变体AAV5病毒体，其中编码变体AAV5 Cap的多核苷酸序列在从对应于VP1衣壳多肽中的581的位置至对应于589的位置的区域中包含至少一个突变，并且其中编码变体AAV5Cap的多核苷酸序列编码SEQ ID NO:7118-SEQ ID NO:10117中任一者的氨基酸序列。治疗性有效载荷是靶向由与FTD有关的基因编码的mRNA的指导RNA，或转基因。指导RNA所靶向的mRNA是由编码颗粒蛋白前体的基因编码的mRNA。转基因是颗粒蛋白前体。将衣壳包裹有效载荷的变体AAV5病毒体施用于受试者。受试者是人类或非人类动物。施用途径是全身施用途径。全身施用途径是静脉内施用。与衣壳包裹治疗性有效载荷的野生型AAV5病毒体相比，衣壳包裹治疗性有效载荷的变体AAV5病毒体在施用于受试者后表现出增强的CNS组织嗜性。在施用于受试者时，与衣壳包裹治疗性有效载荷的野生型AAV5病毒体相比，施用较低剂量的衣壳包裹治疗性有效载荷的变体AAV5病毒体。在施用于受试者后，FTD的至少一种症状得到缓解或受试者得到治愈。

[2111] 实施例31

[2112] 如通过机器学习鉴定的具有肾上腺组织嗜性的AAV5变体

[2113] 此实施例描述了如通过机器学习方法鉴定的展现肾上腺组织嗜性的工程化AAV5变体的优选特性。进行如实施例19和图19至图22中描述的相同生物信息学预处理、机器学

习方法和分析以确定展现肾上腺组织嗜性的工程化AAV5变体的优选特性。

[2114] 下文描述了AAV5 VP衣壳多肽的581至589区域中各位置处的有利生物物理特性和有利氨基酸残基，它们与较高的肾上腺组织嗜性概率相关。表46中描述的任何以下机器学习衍生的位置氨基酸偏好可以单独存在或彼此组合存在或与实施例7，表13中观察和描述的任何氨基酸偏好组合存在。下文表46中列出了从ML模型中提取的排名靠前的重要肾上腺嗜性特征之间共有的位置特征的概述。

[2115]

表 46 机器学习衍生的肾上腺组织嗜性规则
Xaa1 处低摩尔质量 Xaa1 选自 V、P、S 或 C
Xaa1 处低亲疏水性 Xaa1 选自 T、S、W 或 Y
Xaa2 处低亲疏水性 Xaa2 选自 R
Xaa2 处低可突变性 Xaa2 选自 C
Xaa2 处低溶解性 Xaa2 选自 K
Xaa3 处低平均柔性 Xaa3 选自 W、M 或 F
Xaa3 处高溶解性

[2116]

表 46 机器学习衍生的肾上腺组织嗜性规则
Xaa3 选自 M
Xaa4 处高表面可接近性 Xaa4 选自 K 或 R
Xaa4 处高平均柔性 Xaa4 选自 K、I 或 N
Xaa5 处中等可突变性 Xaa5 选自 R 或 H
Xaa5 处高 goldman engelman steitz Xaa5 选自 V 或 L
Xaa5 处低亲疏水性 Xaa5 选自 R
Xaa5 处高体积 Xaa5 选自 Y、R 或 F
Xaa6 处高溶解性 Xaa6 选自 Y、V、M、A 或 C
Xaa7 处中等可突变性 Xaa7 选自 V、H 或 R
Xaa7 处低溶解性 Xaa7 选自 R
Xaa8 处高平均柔性 Xaa8 选自 K、I 或 N
Xaa8 处高摩尔质量 Xaa8 选自 R 或 Y
Xaa9 处高可突变性 Xaa9 选自 N

[2117] 下文表47提供了表现出肾上腺组织嗜性并且符合表46中提供的一种或多种规则的变体AAV VP1衣壳多肽的581至589区域的序列。因此，本公开提供了由具有SEQ ID NO:2的工程化AAV5 VP1衣壳多肽组成的rAAV，其中Xaa1至Xaa9区域是SEQ ID NO:2118-SEQ ID

NO:3117中的任一者,如表47中所公开。本文还涵盖由工程化AAV5 VP2衣壳多肽和工程化AAV5 VP3衣壳多肽组成的rAAV,所述多肽在对应于AAV5 VP1 581至589区域中的氨基酸的AAV5VP2(氨基酸残基445至453)和AAV5 VP3(氨基酸残基389-397)中的区域处具有下文表中公开的序列。

[2118] 表47

[2119] 驱动肾上腺组织嗜性的AAV5 VP1衣壳多肽中581至589区域的序列

[2120]

SEQ ID NO	581-589 序列
2118	AAGGFLSGC
2119	AAVQKERYN
2120	ACYFMMRQN
2121	AFGHAMKQF
2122	AIGCIQNGG
2123	AIGCVQNGG
2124	AIGGFPNGG
2125	AIGYFQNGG
2126	AILKLTKVL
2127	AILKMTKVM
2128	AILKTTKVL
2129	AIMKMTKVL
2130	ALGHRHHQC

[2121]

2131	ALNKTLAEG	2172	CARFVERWH	2213	CLMALNRSG	2254	CVRHTENWR
2132	ANMFRNGML	2173	CARFVERWR	2214	CLMAMKRSG	2255	CVSLFARAM
2133	ANMFRNGMR	2174	CAVIHLNEA	2215	CLMGLKRSG	2256	CVSPFARAK
2134	ANYIQYVWE	2175	CAVPPSRYP	2216	CLSPFARAM	2257	CVSPFARAT
2135	AQMQYERLF	2176	CAVPTARYP	2217	CLVALKRSG	2258	CVSPFARGM
2136	ARAPNEGVI	2177	CAVPTSRYL	2218	CMKRYMMLP	2259	CVSPFARSM
2137	ARQWMCVHC	2178	CAVPTSRYT	2219	CPDEPKWFP	2260	CVSPFARVM
2138	ARTLYMINR	2179	CAVQKERYN	2220	CRKQGHGEK	2261	CVSPFGRAM
2139	ARVIEHKAG	2180	CAVVMMRQP	2221	CSIVRWMSP	2262	CVSPFVRAM
2140	ASMFWHYQV	2181	CCAKFKMCT	2222	CSKFIVAWE	2263	CVSQFARAM
2141	ATCKHTDHL	2182	CCEDMKIWQ	2223	CSKFIVGWE	2264	CVSSFARAM
2142	ATCYFVSSQ	2183	CCVFRHHYC	2224	CSKFIVVWE	2265	CVTTRHKQG
2143	ATFKHTDHF	2184	CDIDYDLGF	2225	CSSFRWMSP	2266	CVWPFKFDF
2144	ATFKHTDHW	2185	CECIVQMAT	2226	CSSVRWMGP	2267	CVWPMKFDF
2145	ATLKHTDHL	2186	CECLVQMGT	2227	CSSVRWMST	2268	CWDVTCQEN
2146	ATNKYANAT	2187	CFDIIQRYD	2228	CTFVMMRQP	2269	CYGRVAHLE
2147	ATVKHTDHL	2188	CFKFKFRLV	2229	CTGQRGFTN	2270	CYRFIQQSY
2148	AVGLNERNG	2189	CFKFKVRLA	2230	CTGVMMRQP	2271	CYVFRHHDC
2149	AVGNRHHQC	2190	CFKSKFSWQ	2231	CTIRGNSSC	2272	DASSNDRNT
2150	AWGMQDQRY	2191	CFNIIERYD	2232	CTLRGNASC	2273	DCESPEKQN
2151	AWMCRGVYC	2192	CGEDMKMWQ	2233	CTVGMMRQP	2274	DCESSEKPN
2152	AWMCRHVYC	2193	CGKDIKIWQ	2234	CTVLMMRQP	2275	DCQSPEKPN
2153	AWMCRNFYC	2194	CGKDMQIWQ	2235	CTVLMTRQP	2276	DEDVCKFVF
2154	AWMCRNVDC	2195	CHGRVAHHE	2236	CTVVLMRQP	2277	DEVACKFVF
2155	AWMCRNVHC	2196	CHLRDCSPI	2237	CTVVMMREP	2278	DEVDCKFVF
2156	AWMCRNVNC	2197	CHMDYDLGF	2238	CTVVMVRQP	2279	DEVVCKFMF
2157	AWMCRNVYY	2198	CHRFIQPSY	2239	CVDTYRMWQ	2280	DEVVCKIVF
2158	AWMCRSVYC	2199	CHRFIQRSY	2240	CVHQRMHVT	2281	DFKLFHNRE
2159	AWMCRTVYC	2200	CHRYIQQSY	2241	CVIQTSRGT	2282	DFKLFHNRP
2160	AWMFRNVYC	2201	CHSRDCSPL	2242	CVKFKVRLV	2283	DFKLFNNRA
2161	AWMRRNVYC	2202	CIKSMFSWQ	2243	CVLEFSRQW	2284	DFKVFHNRA
2162	AWRCRNVYC	2203	CILQTRKGN	2244	CVLKFSREW	2285	DFQLFHNRA
2163	AWVCRNVYC	2204	CKIGTQKGQ	2245	CVLKFSRHW	2286	DFRLFHNRA
2164	AWVMKDQQY	2205	CKIVAQKGQ	2246	CVLKFSRQG	2287	DFRWFIRMN
2165	AWVMKNQRY	2206	CKIVIQKGQ	2247	CVLKFSRQS	2288	DMKRPTRYN
2166	AYCHRAHNG	2207	CKIVTQKSQ	2248	CVLKFSRRW	2289	DMQRPPRYN
2167	CAGIHLNEG	2208	CKLVTQKGQ	2249	CVLKVSRQW	2290	DMQRPSRYN
2168	CAGIHLNES	2209	CKMVTQKGQ	2250	CVLKYSRQW	2291	DNNMIVRCV
2169	CAGIHMNEA	2210	CKMYMHEAR	2251	CVLQTSRGI	2292	DNQWMNSLE
2170	CARFVARWP	2211	CKMYMQESR	2252	CVMQTSRGI	2293	DNSMIVRCA
2171	CARFVEGWP	2212	CKMYMRESR	2253	CVNKLSYEN	2294	DQPHLHRIG

[2122]

2295	DRCLVEWVS	2336	FRKQAHGEK	2377	GTFKHTDHL	2418	ICTTLIQGM
2296	DRCLVKWVA	2337	FRKQGHGAK	2378	GVAYRESQA	2419	ICTTQIKGM
2297	DRMDFHFFF	2338	FRKQGHGEE	2379	GVDKLSYEN	2420	ICYKWWVMF
2298	DSKQIAKAM	2339	FRKQGHGQK	2380	GVHHRWRMT	2421	IDAALDMSW
2299	DVCNVLGWK	2340	FRKQGNGEK	2381	GVKPRWGLF	2422	IEVKYWNRH
2300	DVDVVARPM	2341	FRKQGPGEK	2382	GVKTRCDYD	2423	IEYVRSCPW
2301	DVKLFHNRA	2342	FRKQVHGEK	2383	GVWMQGIHP	2424	IFMAFDHLE
2302	DVRGLNREY	2343	FRNQVHGEK	2384	GWEKKINFQ	2425	IGMDKPVDA
2303	DVRNFFDHC	2344	FSSCCRQKG	2385	GWEKKINGQ	2426	IHMKPQESP
2304	DWHQYDGCE	2345	FTSNRNWEA	2386	GWEKKLNVQ	2427	IHMKQQDSP
2305	DWNQHDGCE	2346	FTSNRNWQD	2387	GWEKKVNVQ	2428	IHMKQQESA
2306	DWNQYAGCE	2347	FTSNRNWQP	2388	GWEKTINVQ	2429	IHMKQQEST
2307	DWNQYDGGE	2348	GAHHRCRMT	2389	GWFFKNYYC	2430	IHMKQQQSP
2308	DWNQYNGCE	2349	GAWMQGFHP	2390	GWKKKINVQ	2431	IHMKQRESP
2309	DWSQYDGCE	2350	GCAKFQMCT	2391	GWMCRNVYC	2432	IHMKTIVGE
2310	DYKQTAKAM	2351	GCAKTRFEN	2392	GWQKKINVQ	2433	IHMRQQESP
2311	ECAHVIKFN	2352	GCDKNRFEN	2393	HACCLLQQY	2434	IILHNMKSW
2312	ECHKMRCLE	2353	GCDKTRFAN	2394	HACSLLKQY	2435	IMLALMHED
2313	ECHKMRCVE	2354	GCDKTRFES	2395	HACSLLQQF	2436	IMLHRVMSD
2314	ECNKMRCIE	2355	GCDKTRFEY	2396	HIEWVSMQR	2437	IMMAHWFAH
2315	EETKKVKHW	2356	GCHKMRCIE	2397	HIQWVSMHR	2438	IMMAQWFAQ
2316	EHKKNINMH	2357	GEGIRKMNN	2398	HMEWVSMHR	2439	IMMAQWFGH
2317	EKRHISRAP	2358	GEGIRQMNT	2399	HRADIMCWN	2440	IMMAQWFSH
2318	EQRHISRAR	2359	GEGVRQMNN	2400	HREAIMCWN	2441	IMMAQWIAH
2319	EQRHISRTP	2360	GEVVCKFVF	2401	HREDIMCGN	2442	IMMHRMMSD
2320	ERVIEHRAG	2361	GFDRQDAHR	2402	HREDIMCWD	2443	IMMHRVMSE
2321	ERVLEHKAG	2362	GFKLFHNRA	2403	HREDIMCWI	2444	IMMHRVMSV
2322	EYDVNMRQM	2363	GMKTRCDDD	2404	HREDIRCWN	2445	IMMHRVMSY
2323	FGKDAFVSS	2364	GMKTRCDYA	2405	HRGDIMCWN	2446	IMMHRVMYD
2324	FGKDTFASS	2365	GMKTRCDYV	2406	HRKDIMCWN	2447	IMMLRVMSD
2325	FGKDTFLSS	2366	GMKTRCHYD	2407	HRQDIMCWN	2448	IMMNRVMSD
2326	FGKDTFVSA	2367	GMKTRCNYD	2408	HRTETECWC	2449	IMMQRVMSD
2327	FGKDTVVSS	2368	GMKTRCVYD	2409	HSCSLLQQY	2450	IMMVQWFAH
2328	FGMDTFVSS	2369	GMQTRCDYD	2410	HVGALTRSD	2451	IMMYRVMSD
2329	FMMHRVMSD	2370	GMRTRCDYD	2411	IARHIDHDL	2452	IMVALMHEG
2330	FMMHRVMSG	2371	GRAFVKIDT	2412	IAWHVDHDL	2453	IVCRMWHQP
2331	FMMMAKTQA	2372	GRICMQGMK	2413	ICHLRHIAA	2454	IWHKWWVMF
2332	FQKLRDVCF	2373	GRIFMLGMK	2414	ICHLRHIPV	2455	IWYKWWAMF
2333	FREQGHGEK	2374	GRIFVKIDT	2415	ICHLRHISG	2456	IWYKWWLMF
2334	FRKEGHGEK	2375	GRLFVKIDT	2416	ICHLRHNAG	2457	IWYKWWVMC
2335	FRKPGHGEK	2376	GRWFQIGTE	2417	ICHLRHVAG	2458	IWYKWWVMV

[2123]

2459	IWYKWWVMY	2500	KHKINAVER	2541	KTMEVHFQK	2582	KYKRPTINP
2460	KAIPENSNR	2501	KHKPMCVQL	2542	KTQWDEAIV	2583	KYQPPTINP
2461	KALEVHFQK	2502	KHKYEDPMV	2543	KTYFADTIC	2584	LAVDKSRLL
2462	KAMEFHFQK	2503	KHKYEDTMG	2544	KTYFADTIS	2585	LCGLMPRPA
2463	KAMPEPVTE	2504	KHQINTVER	2545	KTYFADTMG	2586	LCGVMPRPG
2464	KAQDFATEN	2505	KHQPMCVQL	2546	KTYFADTVG	2587	LFKPEDEDR
2465	KAQEFATED	2506	KHSSHMEAC	2547	KTYFAETIG	2588	LFKPEDPDR
2466	KAQEFATEH	2507	KIHYHFEDK	2548	KTYFANTIG	2589	LFKPEEQDR
2467	KAQKFATEN	2508	KINYHFEDE	2549	KTYFSDTIG	2590	LHRDPMPIM
2468	KAQQFATEN	2509	KINYHFEDQ	2550	KTYFWMSYC	2591	LHRDQMPFM
2469	KAREFATEN	2510	KISGLNDAT	2551	KTYFWMSYS	2592	LHRDQMPLM
2470	KASTEAGDK	2511	KKDEANGQT	2552	KTYFWTSYY	2593	LIMAHWLHM
2471	KATTEASDK	2512	KKDPMQTYP	2553	KTYFWVSYY	2594	LIMAYWLDM
2472	KCDKPMLEP	2513	KMGTEVSQQ	2554	KVDVFTVKD	2595	LIMSHWLDM
2473	KCDKPMLQP	2514	KMRMIFRKD	2555	KVDVITFKD	2596	LKRPQGEAE
2474	KCDTFRGME	2515	KNEPMCVQL	2556	KVDVITVQD	2597	LLMAHWLDM
2475	KCEPFRGME	2516	KPVEYASAP	2557	KVDVVTVKD	2598	LQKYYMNQP
2476	KCETFLGME	2517	KQDFHHGLA	2558	KVILDSFRR	2599	LQNYRDEDG
2477	KCGILYYQG	2518	KQLYHIPTC	2559	KVMPEPVPE	2600	LQPWWLGMI
2478	KCGNLYYQG	2519	KQQYVTMNA	2560	KVMPEPVTA	2601	LQTYHHVDG
2479	KCGSLNYQG	2520	KQRDISRAP	2561	KVMPEPVTV	2602	LQTYQPVDG
2480	KCGSLYCQG	2521	KQRHISRAP	2562	KVMPVPVTE	2603	LRFLSDADP
2481	KCGSLYHQG	2522	KQSAHMEAC	2563	KVMSEPVTE	2604	LRKEGHGEK
2482	KCGSLYSQG	2523	KQSPHMEAC	2564	KVTWMDLED	2605	LRYLSDTDP
2483	KCGSLYYQV	2524	KQSSHMEAF	2565	KVVLDSFRQ	2606	LSTNRLVNA
2484	KCGSMYYQG	2525	KRDLANGQF	2566	KVYMEPANW	2607	LSTNRMVSA
2485	KCGTLYYQG	2526	KRIIEHPCA	2567	KYEFLIKYI	2608	LSTNRMVTA
2486	KCYFMMRQN	2527	KRMDFHFFF	2568	KYEFLLKHI	2609	LYRDQMPIM
2487	KDCQYIGAC	2528	KRSSHMEAC	2569	KYEFLLKYL	2610	MCGHASKAV
2488	KDCQYIGSW	2529	KRVIEHKAG	2570	KYEFLLRYI	2611	MCGHASRAE
2489	KDCQYLGAW	2530	KRYFWMSYY	2571	KYEILNCNP	2612	MCGPASKAE
2490	KFELNQTAV	2531	KSMLEDRIC	2572	KYEVLLKYI	2613	MCHLRHIAG
2491	KFELNRKAV	2532	KSMWEARIC	2573	KYEYLLKYI	2614	MCRANGNMP
2492	KFETFRGME	2533	KSMWEDRIG	2574	KYIAVWDAR	2615	MCSLRHWQP
2493	KFGLNQKAV	2534	KSMWEDRTC	2575	KYIDFWDAR	2616	MIGMWNRSN
2494	KFIDVWDAR	2535	KSMWEERIC	2576	KYIDIWDAR	2617	MISMLNRSN
2495	KGEYQADAP	2536	KSMWENRIC	2577	KYIDVWEAR	2618	MRSIYMNIS
2496	KHDPMCGQL	2537	KSMWEYRIC	2578	KYKAPTINP	2619	MRWILTIME
2497	KHEQMCVQL	2538	KSVWEDRIC	2579	KYKHPTINP	2620	MSAQRHVQT
2498	KHERMCVQL	2539	KTFFWMSYY	2580	KYKPLTINP	2621	MSAQRYAQT
2499	KHETMCVQL	2540	KTFSESTCR	2581	KYKPPTMNP	2622	MSAQRYVQA

[2124]

2623	MSMLRMHEA	2664	NNQWTNSLE	2705	QAYCCMHQD	2746	RCEAFNWRE
2624	NAMCFGRAR	2665	NPCKINVGT	2706	QCDEFKMRM	2747	RCELFNWRE
2625	NAMCFGREK	2666	NQCKINFGT	2707	QCHKMRCIE	2748	RCEPFHWRE
2626	NAMCFGRTK	2667	NQCKINVGN	2708	QESNRCWCD	2749	RCEPFNLRE
2627	NAMCIGRAK	2668	NQCKINVGP	2709	QHKERMLFD	2750	RCEPFNWHE
2628	NAMGFGRAK	2669	NQCKLNVGT	2710	QHKERMVFD	2751	RCGHASKAE
2629	NATCFGRAK	2670	NQLKDVAAS	2711	QHQKNINMH	2752	RCGPFNWRE
2630	NCNKCINRT	2671	NQWKINVGT	2712	QLDEFKMRM	2753	RCGQVRNEM
2631	NCNKCKHRT	2672	NRAETECWC	2713	QMSWVQNWE	2754	RCKLFDHDE
2632	NCNKCKYRT	2673	NRCLVEWVA	2714	QMYHNHHTA	2755	RCVEDHFHK
2633	NCRHTILYV	2674	NRMDFHFFV	2715	QMYWAQNWE	2756	RCVLDHFHK
2634	NCSKCAKYD	2675	NRMDFNFFF	2716	QMYWGQNWE	2757	RCVQDHFHE
2635	NCVKLGCQS	2676	NRNKCKNRT	2717	QMYWIQNWE	2758	RCVQDHFHR
2636	NELQDQCGM	2677	NRPFNYFQR	2718	QMYWVQNLE	2759	RCVQDHFHT
2637	NEVVCKFVF	2678	NRTLYMINR	2719	QMYWVQSWE	2760	RCVQDHFNK
2638	NFSKCAKYV	2679	NSSYFEFSN	2720	QPVTRMTNE	2761	RCVQDHIHK
2639	NFSKCTKYD	2680	NSVKLGCQN	2721	QQLHSRCGM	2762	RCVQDHSHK
2640	NFSKGAKYD	2681	NTIKRCSQG	2722	QQRHISRAP	2763	RCVQDNFHK
2641	NGMDKAVDA	2682	NVDMSRRDF	2723	QRDLAHGQF	2764	RCVQHHFHK
2642	NGMDKPFDA	2683	NVMCFGRAK	2724	QRDLANGKS	2765	RCVRDHFHK
2643	NGMDKPIDA	2684	NVRNFFAHC	2725	QRDLANGLF	2766	RCYFMMRQN
2644	NGMDKPVDG	2685	NVRNFFEHC	2726	QRDLANGQI	2767	RDLSLHLMR
2645	NGMDKPVDS	2686	NVSKCAKYD	2727	QRDLANGQS	2768	RECLIERCG
2646	NGMDKPVNA	2687	NVWPNIAEM	2728	QRDLANGQV	2769	RECLIERCI
2647	NGMDQPVDA	2688	NVWTNISEM	2729	QRDLANGQY	2770	RECLIKRCS
2648	NGMDRPVDA	2689	NWNGRWNQC	2730	QRDMANGQF	2771	RECLVERCS
2649	NGMDTPVDA	2690	NWNKCKNRT	2731	QRDPANGQF	2772	RIERVDVFN
2650	NGVDKPVDA	2691	NWPCHDNRH	2732	QSAFRQSDS	2773	RKDEANGQA
2651	NGWESEQHM	2692	NWSYHDNRH	2733	QSTFRQCDS	2774	RMDEANGQT
2652	NHMKQQESP	2693	NWTGRWHQC	2734	QSTMRQYYA	2775	RQCLIERCS
2653	NLDMSRRAF	2694	NYCDRAHNG	2735	QSYCCMHQY	2776	RQMWMAHHA
2654	NLDMSRRDC	2695	PAVQKERYN	2736	QTYWVQNWE	2777	RRDAKVANG
2655	NMIKRCCQG	2696	PCFSIARYY	2737	QVSGMFHPA	2778	RRDEANGQT
2656	NMIKRCSKG	2697	PCYFMMRQN	2738	QVYCCMHQY	2779	RREDIMCWN
2657	NMIKRCSQA	2698	PKVTCRYRM	2739	QVYWVQNWE	2780	RRKDIMKER
2658	NMIKRCSQC	2699	PPLHSRCGM	2740	QWDEFKKRM	2781	RRYRELTEE
2659	NMIKRCSQS	2700	PQLHSRCRM	2741	QWDEFKMRR	2782	RSEPFNWRE
2660	NMIKRCSQV	2701	PQLLSRCGM	2742	QWDEFKVRM	2783	RTDAIHHQV
2661	NMIKWCSQG	2702	PQLNSRCGM	2743	QWDVFKMRM	2784	RTDAIHPQI
2662	NMVKRCSQG	2703	PQLRSRCGM	2744	RACSLLQQY	2785	RWVQDHFHK
2663	NNQWKNSLE	2704	PWVMKDQRY	2745	RCDKTRFEN	2786	RYCQVWGCA

[2125]

2787	RYDLFSYHQ	2828	SNCVQHNAW	2869	SWRHIYHEA	2910	TCFSIVRYY
2788	RYDLVSHHQ	2829	SNGAQHNAW	2870	SWRTRGNFE	2911	TCFSNARYY
2789	RYDLVSYHR	2830	SNGFEWACE	2871	SWSKCNMHV	2912	TCFSTARYY
2790	RYDLVSYYQ	2831	SNGFEWVCA	2872	SWVLIHCWN	2913	TCFSVARYY
2791	RYNLVSYHQ	2832	SNGFEWVGE	2873	SYDRQDAHR	2914	TCFTIARYY
2792	RYRQEWGCA	2833	SNGFEWVSE	2874	SYVFRHHYC	2915	TCFTVARYY
2793	SAGQKERYN	2834	SNGFEWVYE	2875	TACATYRSF	2916	TCGSLYYQG
2794	SAIQKERYN	2835	SNMFRNGMQ	2876	TACMTYRSF	2917	TCHFALHGC
2795	SALYGSKAI	2836	SNVFEWVCE	2877	TACVAYRSF	2918	TCHFMMRQN
2796	SAVQQERYN	2837	SRDEAKSWV	2878	TACVNYRSF	2919	TCHYALHAC
2797	SAVQRERYN	2838	SRICMLGMK	2879	TACVTYRGL	2920	TCNKCKNRT
2798	SAVRNERYN	2839	SRIILDYYS	2880	TACVTYRIF	2921	TCRYALHGC
2799	SAWYGSKSI	2840	SRPFNYIQR	2881	TAFVTYRSF	2922	TCSFMMRQN
2800	SCKGFHLQN	2841	SRSFNYFQR	2882	TAFVTYRSL	2923	TCSRFDNQA
2801	SCKSHQMHE	2842	SRWCQIGTE	2883	TAGHEGHSC	2924	TCVSIARYY
2802	SCRHIYREA	2843	SRWFKIGTE	2884	TAGHQGHIC	2925	TCWALFRGE
2803	SCRHLYHEA	2844	SRWFQIGTK	2885	TAGVTYRSF	2926	TCWDLFRGG
2804	SCRHNYHEA	2845	SSGDRRSDS	2886	TAITMWQYM	2927	TCYFMMGQN
2805	SCRRIYHEA	2846	SSKGFHLQN	2887	TAMCFGRAK	2928	TCYFMMREN
2806	SFDRQDVHR	2847	SSKQDFYGL	2888	TAMSMCMYG	2929	TCYFMMRKN
2807	SFDRQGAHR	2848	SSKSHRMHE	2889	TAMTICMYG	2930	TCYFMMRPN
2808	SFGHAMKQC	2849	SSMLEGVQS	2890	TANDHRVDA	2931	TCYFMMRQD
2809	SFKGFHLEN	2850	SSVDRRSDG	2891	TATDRRVDA	2932	TCYFMMRQK
2810	SFKGFHVQN	2851	SSVHALYFP	2892	TAVQKERYN	2933	TCYFMMRQT
2811	SFMGFHLQN	2852	SSVHTLYYP	2893	TAVSICMYG	2934	TCYFMMRRN
2812	SFVRQDAHR	2853	SSVQKERYN	2894	TAVTMWQSM	2935	TCYFMVRQN
2813	SGQKFIFCP	2854	STCYFVASL	2895	TAWVTYRSF	2936	TCYSIARYY
2814	SIGCFQNGG	2855	STDPYTSHH	2896	TAYVTYRSF	2937	TCYSMMRQN
2815	SIIFSNTCM	2856	STFKHTDHL	2897	TCCFMMRQN	2938	TCYVMMRQN
2816	SILLPGIQG	2857	STGFEWVCE	2898	TCCSIARYY	2939	TCYYMMRQN
2817	SIMHEGVQS	2858	STSYFVASQ	2899	TCDFMMRQN	2940	TDCVTYRSF
2818	SIMLPGIQA	2859	SVGPTVHQM	2900	TCFFMMRQN	2941	TFDLSNKVT
2819	SLNQAIHLE	2860	SVSKFQADK	2901	TCFGIARYY	2942	TFGWFMRIP
2820	SMGPTVHEM	2861	SVVQKERYN	2902	TCFNIARYY	2943	TFKGFHLQN
2821	SMNQAVHLE	2862	SWEKKINVQ	2903	TCFSIARCY	2944	TFSKCAKYD
2822	SMNQPIHLE	2863	SWGLIHGWN	2904	TCFSIARYD	2945	TFYFMMRQN
2823	SMRSQHADY	2864	SWGLIHWWN	2905	TCFSIARYF	2946	TGCVTYRSF
2824	SNCALHNAW	2865	SWGLIHYWN	2906	TCFSIARYS	2947	TGGLNERNG
2825	SNCAQHNSW	2866	SWGLVHCWN	2907	TCFSIDRYY	2948	TGGTFVRTM
2826	SNCAQYNAW	2867	SWKWDKYCN	2908	TCFSIGRYY	2949	TGMDKPVDA
2827	SNCSQHNAW	2868	SWMCRNVYC	2909	TCFSITRYY	2950	TGQATERAL

[2126]

2951	TGTDHRVDA	2992	TTCYIVASE	3033	VAWSNEDQK	3074	WATKWDVHR
2952	TGWSTCMIP	2993	TTCYLVASP	3034	VCHTFLRLG	3075	WCNTRLAMA
2953	TITKPINLF	2994	TTFKHTDHL	3035	VFCRMWHQP	3076	WCTSRLAMA
2954	TITKPSHLF	2995	TTGYFVASQ	3036	VGEKRMDRS	3077	WCTTRLTMA
2955	TLNKTLAES	2996	TTRYFVASQ	3037	VHEQLEGCK	3078	WCTTRLVMA
2956	TNTCECVDR	2997	TTWYFVASQ	3038	VILKMTKVL	3079	WWMHGRGGR
2957	TNYIQYVWD	2998	TTYYFVASQ	3039	VLGNRHHQC	3080	YACSLLQQY
2958	TPLMMCRAS	2999	TVGLNERNA	3040	VLKKFGHSF	3081	YAVQKERYN
2959	TPMMMCQAS	3000	TVGLNERNS	3041	VLMQYERLF	3082	YCHKTAVAA
2960	TRCYFVASQ	3001	TVGLNKRNG	3042	VLNKTLAES	3083	YCHQQDRIT
2961	TRDAKVSNG	3002	TVGVNERNG	3043	VLPKFGHSF	3084	YCNKTAVAA
2962	TREWMCVHC	3003	TVKMHFAAP	3044	VLQKFGHSV	3085	YCYKNAVAA
2963	TRHMGVSAG	3004	TVKTNFTSK	3045	VLRKFGHSF	3086	YCYKTAVAV
2964	TRKWMCVHC	3005	TVQSNFTSK	3046	VMKTRCDYD	3087	YCYRTAVAA
2965	TRLWMCVHC	3006	TWALAAKYE	3047	VMMHRVMSD	3088	YGEAQITSP
2966	TRMDFHFFF	3007	TWFSIARYY	3048	VPIEYMTEK	3089	YIKQICADA
2967	TRPWMCVHC	3008	TWKWDEYCN	3049	VPIPYMTEK	3090	YIKQICSDS
2968	TRQRMCVHC	3009	TWKWDKYCH	3050	VPIQYMAEK	3091	YINKNPHQM
2969	TRQWMCAHC	3010	TWKWDKYYN	3051	VPMQYMTEK	3092	YLHKNPHQM
2970	TRQWMCIHC	3011	TWKWDRYCN	3052	VQEQLESCK	3093	YLNKNHHQM
2971	TRQWMCLHC	3012	TWNKCNMHV	3053	VQKLRDLCC	3094	YLNKNLHQM
2972	TRQWMCVHG	3013	TWPYHDNRH	3054	VQMQYARLF	3095	YLNKNPHQR
2973	TRQWMCVHW	3014	TWQTKDFYN	3055	VQMQYERLY	3096	YLNKNPHQV
2974	TRQWMCVYC	3015	TWSKCNLHV	3056	VQMQYQRLF	3097	YLNQNPHQM
2975	TRQWMGVHC	3016	TWSKCNMRV	3057	VQMQYVRLF	3098	YLTKNPHQM
2976	TRQWMWVHC	3017	TWSKCNVHV	3058	VQTYQHVDG	3099	YMIKRCSQG
2977	TRTLYMIDR	3018	TWSKCSMHV	3059	VQVQYERLF	3100	YPTQRLHHL
2978	TRTLYRINR	3019	TWSKCTMHV	3060	VRIQYMTEK	3101	YRDEAKCWV
2979	TRTPYMINR	3020	TWSKFNMHV	3061	VRMQYERLF	3102	YRDEAKSWA
2980	TRWFQIGTE	3021	TWSRCNMHV	3062	VSAECYFDM	3103	YRDEAQSWV
2981	TRYFMMRQN	3022	TWTYCEQRH	3063	VSDEFYFDM	3104	YRDEDKSWV
2982	TSCDRAHNA	3023	TYCDRADNG	3064	VSDVCYFDM	3105	YRGEAKSWV
2983	TSCVTYRSF	3024	TYCDRAHDV	3065	VSGECYFDM	3106	YRHEAKSWV
2984	TTCFFVASQ	3025	TYCDRAHNA	3066	VTWMVRTSN	3107	YRKQGHGEK
2985	TTCFIVASQ	3026	TYCDRTHNG	3067	VVCRMWYQP	3108	YRMEYGTCK
2986	TTCYFVASE	3027	TYFDRAHTG	3068	VVEKRMERS	3109	YRMEYGTWE
2987	TTCYFVASP	3028	TYWDRAHNG	3069	VVFRMWHQP	3110	YSAQLRQQM
2988	TTCYFVASW	3029	TYYDRAHNG	3070	VVWMQGFHP	3111	YSVDRRSDS
2989	TTCYFVGSQ	3030	VACRMWHQP	3071	VVYRMWHQP	3112	YWCPNQWWR
2990	TTCYFVSSQ	3031	VAEKRMDRS	3072	VWRYFCVSL	3113	YWCPSKWWR
2991	TTCYFVVSQ	3032	VARHVDHDL	3073	VWYKWWVMF	3114	YWCPSQLWR

3115	YWCPSQWLR	3116	YWGLIHCWN	3117	YWGPSQWWR

[2127] 实施例32

[2128] 如通过机器学习鉴定的具有骨髓组织嗜性的AAV5变体

[2129] 此实施例描述了如通过机器学习方法鉴定的展现骨髓组织嗜性的工程化AAV5变体的优选特性。进行如实施例19和图19至图22中描述的相同生物信息学预处理、机器学习方法和分析以确定展现骨髓组织嗜性的工程化AAV5变体的优选特性。

[2130] 下文描述了AAV5 VP衣壳多肽的581至589区域中各位置处的有利生物物理特性和有利氨基酸残基，它们与较高的骨髓组织嗜性概率相关。表48中描述的任何以下机器学习衍生的位置氨基酸偏好可以单独存在或彼此组合存在或与实施例18，表35中观察和描述的任何氨基酸偏好组合存在。下文表48中列出了从ML模型中提取的排名靠前的重要骨髓嗜性特征之间共有的位置特征的概述。

[2131]

表 48
机器学习衍生的骨髓组织嗜性规则

Xaa1 处高亲疏水性 Xaa1 选自 V、I 或 L
Xaa1 处低可突变性 Xaa1 选自 Y、L、F 或 C
Xaa2 处低亲疏水性 Xaa2 选自 Y 或 W
Xaa2 处高摩尔质量 Xaa2 选自 W
Xaa2 处低表面可接近性 Xaa2 选自 W 或 A
Xaa2 处低亲水性 Xaa2 选自 W

表 48
机器学习衍生的骨髓组织嗜性规则

Xaa2 处低可突变性 Xaa2 选自 C
Xaa2 处低平均柔性 Xaa2 选自 W、M 或 F
Xaa5 处低平均柔性 Xaa5 选自 W、M 或 F
Xaa6 处低平均柔性 Xaa6 选自 W、M 或 F
Xaa6 处低可突变性 Xaa6 选自 Y、F、L 或 C
Xaa6 处高溶解性 Xaa6 选自 W、F、I 或 L
Xaa7 处低表面可接近性 Xaa7 选自 C 或者 Xaa7 处高表面可接近性 Xaa7 选自 D 或 N
Xaa7 处低可突变性 Xaa7 选自 C
Xaa7 处高溶解性 Xaa7 选自 C 或者 Xaa7 处低溶解性 Xaa7 选自 D
Xaa8 处低电荷 Xaa8 选自 D 或 E
Xaa8 处高可突变性 Xaa8 选自 D、E、A 或 T

[2132]

[2133]

表 48 机器学习衍生的骨髓组织嗜性规则
Xaa9 处高摩尔质量 Xaa9 选自 H 或 F
Xaa9 处低可突变性 Xaa9 选自 Y、F 或 L

[2134] 下文表49提供了表现出骨髓组织嗜性并且符合表48中提供的一种或多种规则的变体AAV VP1衣壳多肽的581至589区域的序列。因此,本公开提供了由具有SEQ ID NO:2的工程化AAV5 VP1衣壳多肽组成的rAAV,其中Xaa1至Xaa9区域是SEQ ID NO:5118-SEQ ID NO:6117中的任一者,如表49中所公开。本文还涵盖由工程化AAV5 VP2衣壳多肽和工程化AAV5 VP3衣壳多肽组成的rAAV,所述多肽在对应于AAV5 VP1 581至589区域中的氨基酸的AAV5VP2(氨基酸残基445至453)和AAV5 VP3(氨基酸残基389-397)中的区域处具有下文表中公开的序列。

[2135] 表49

[2136] 驱动骨髓组织嗜性的AAV5 VP1衣壳多肽中581至589区域的序列

[2137]

SEQ ID NO	581-589 序列
5118	AADVVPCHH
5119	ACFHVAEMF
5120	AEMHMVADF
5121	AESLCNELC
5122	AKKSCQCWH
5123	ANDYDQCCM
5124	ANGHKFNEE
5125	ANNEFIQPF
5126	ANQCIFKTM
5127	AQCLCNELC
5128	AQGLCNELC
5129	AQHDWYQKY
5130	AQPNMDYVP
5131	AQPSMDCVP
5132	AQPSMDYAP
5133	AQPSMDYVA
5134	AQPSMEYVP
5135	AQSLCNELF
5136	AQSLCNELG
5137	AQSLCNQLC
5138	AQSLFNELC
5139	AQSVCNELC
5140	AQTLCNELC
5141	ARAKIQDEW
5142	ARALIQDEW
5143	ARARIQDEW
5144	ARGQIQDEW
5145	ARPSMDYVP
5146	ARVQIQDEW
5147	AWESSSIIS
5148	AYAENWCIE
5149	CANHMIRTR
5150	CATHMIRTK
5151	CCEHKEYLN
5152	CCEPKGYLN
5153	CCRKDNKKS
5154	CCRMANKKS
5155	CCRMDDKKS
5156	CCRMDHKKS
5157	CCRMDIKKS
5158	CCRMDNKES
5159	CCRMDNKIS
5160	CCRMDNKKA
5161	CCRMDNKRS
5162	CCRMDNQKS
5163	CCRRDNKKS
5164	CDAEAEFQR
5165	CDAEDEFRR
5166	CDAKAEFRR
5167	CDAVAEFRR
5168	CDDEAEFRR
5169	CDGEAEFLR
5170	CDGEAEFRR
5171	CDGVAEFRR
5172	CFKQKHNLF
5173	CFLQKHNLF
5174	CFQEKHNLF
5175	CFQPKHNLF
5176	CFQQKHNIF
5177	CFQQKHNLI
5178	CFQQNHNLF
5179	CFQRKHNLF
5180	CGYHEFMVG
5181	CHCDAVECF
5182	CHFDAMECF
5183	CHKCIGQDN
5184	CHQCIGPDN
5185	CHQCIGQDS
5186	CHQCIGQDT
5187	CHQCVGQDN
5188	CHWDAMECF
5189	CIKIQLIDQ
5190	CIKIQMIDK
5191	CIKIRMIDQ
5192	CIKVQMIDQ
5193	CINITAITP
5194	CISITVITP
5195	CMAHIWEAH
5196	CMAHIWEGD
5197	CMDHIWEGH
5198	CMSHIWEGH
5199	CPQCIGQDN
5200	CRCDAMECF
5201	CWYDEFMVG
5202	CWYHEFMVA
5203	CWYHEFMVD
5204	CWYHEFMVR
5205	CWYHEFMVS
5206	CWYHEFMVV
5207	CWYHEFVVG
5208	CWYHEHMVG
5209	CWYHEVMVG
5210	CWYHKFMVG
5211	CWYLEFMVG
5212	CWYNEFMVG
5213	CWYPEFMVG
5214	CWYREFMVG
5215	CWYYEFMVG
5216	DACDWKNDY
5217	DACQEWDSE
5218	DADFCARMP
5219	DADFCPRMA
5220	DADFCPRMS
5221	DADFCPRMT
5222	DARFDQWHY
5223	DCAENWCIE
5224	DCIHDQCSS
5225	DDDCWVFYH
5226	DDDFWVFYH
5227	DDDHWVFYH
5228	DDDYWAFYH
5229	DDDYWGFYH
5230	DDGSWVFYH
5231	DFAENWCIE
5232	DGDYWVFYH
5233	DGGCWPREC
5234	DHAENWCIE
5235	DHELTGSEA
5236	DITQLQCYC
5237	DKNEMFMKW
5238	DLDCDWLHS
5239	DMDCEWLHS
5240	DPCQDWDSE
5241	DPCQEWNSE
5242	DPCQKWDSE
5243	DPCQVWDSE
5244	DPCREWDSE
5245	DPGQEWDSE
5246	DQCQEWDSE
5247	DRCQEWDSE
5248	DSCQEWDSE
5249	DSRFDQWHH
5250	DTCQEWDSE
5251	DYAEHWCIE
5252	DYAEIWCIE
5253	DYAEKWCIA
5254	DYAENGCIE
5255	DYAENWCIA
5256	DYAENWCIG
5257	DYAENWCME
5258	DYAENWCNE
5259	DYAENWCVE
5260	DYAENWWIE
5261	DYAETWCIE
5262	DYAQNWCIE
5263	DYAVNWCIE
5264	DYDENWCIE
5265	DYGENWCIE
5266	DYPENWCIE

[2138]

5267	DYTENWCIE	5308	EKKSCLCWH	5349	EQHHWYQKY	5390	FQMSMCSMG
5268	DYVENWCIE	5309	EKKSCPCWH	5350	EQHNWYQKY	5391	FQVRKFWDP
5269	EAEDEIQKS	5310	EKKSCQCWD	5351	EQKSCQCWH	5392	FRIIWGADM
5270	EAEDEIQRA	5311	EKKSCQCWN	5352	EQLLSLCHS	5393	FRILWQCEG
5271	EAGDEIQKA	5312	EKKSCRCWH	5353	EQLLSLCLS	5394	FRMAMCSMG
5272	EATGMLYEE	5313	EKKSFQCWH	5354	EQLLSLCPN	5395	FRPIWGADM
5273	EAWNAHEEG	5314	EKKSWQCWH	5355	EQLLSLCSS	5396	FRPIWGGDM
5274	ECESSSIIS	5315	EKKSYQCWH	5356	EQLLTLCPS	5397	FRTIWGADL
5275	ECIHAQCSS	5316	EKKTCQCWH	5357	EQLSSLCPS	5398	FRTIWGADV
5276	ECIHDKCSS	5317	EKLLSLCPS	5358	EQNDWYQKY	5399	FRTIWGPDM
5277	ECIHDQCCS	5318	EKNESTCET	5359	EQPLSLCPS	5400	FRTIWGTDM
5278	ECIHDQCSA	5319	EKNQCTCET	5360	EQVLSLCPS	5401	FRVRKFWDP
5279	ECIHDQCSP	5320	EKNQGTCET	5361	ERIHDQCSS	5402	FVDSMLHKW
5280	ECIHDQCST	5321	EKNQSTCEA	5362	ERKSCQCWH	5403	FWDRQWDDF
5281	ECIHDQCTS	5322	EKNQSTCEP	5363	ERNQSTCET	5404	FWLGLFPTK
5282	ECIHHQCSS	5323	EKNQSTCES	5364	ESEDEIQKA	5405	FWLGVFPAK
5283	ECIHVQCSS	5324	EKNQSTCGT	5365	ESGHKFNEE	5406	FWMGLFPAK
5284	EDFHEFAQT	5325	EKNVWNNKW	5366	ESIGMVCSD	5407	GATHMIRTR
5285	EDGHKFNEE	5326	EKQSCQCWH	5367	ETALKWEWS	5408	GCEAKEYLN
5286	EEKSCQCWH	5327	EKRSCQCWH	5368	ETAWKWDWS	5409	GCLTWAFRG
5287	EGIHDQCSS	5328	EKTQSTCET	5369	ETGHKFNEE	5410	GCVTWAFRC
5288	EHAEWKCMY	5329	EKTSCQCWH	5370	ETSWKWEWS	5411	GCVTWDFRG
5289	EHHEWKCMY	5330	ELPEWKCMY	5371	EVPWESKLC	5412	GDHDMLNSL
5290	EHPEGKCMY	5331	ENAHKFNEE	5372	EVWDAHEEG	5413	GEQEARNCS
5291	EHPELKCMY	5332	ENGDKFNEE	5373	EVWNSHEEG	5414	GEREGRNCS
5292	EHPEWKCLC	5333	ENGHEFNEE	5374	EWEGSSIIS	5415	GEYLRHTVT
5293	EHPEWKGMY	5334	ENGHKFDEE	5375	EWESNSIIS	5416	GEYLRNTIT
5294	EHPEWTCMY	5335	ENGHKFNAE	5376	EWESSSIIA	5417	GEYLRSTVT
5295	EHREWKCMY	5336	ENGHKFNEA	5377	EWESSSTIS	5418	GEYQRNTVT
5296	EHSEWKCMY	5337	ENGHKFNEK	5378	EWETSSIIS	5419	GGAAMWGRY
5297	EHTEWKCMY	5338	ENGHKFNKE	5379	EWGFGSDET	5420	GGAAMWRKY
5298	EINQSTCET	5339	ENGHKFNQE	5380	EWIHDQCSS	5421	GGAGMWRRY
5299	EKDQSTCET	5340	ENGHKFSEE	5381	EWLPWINER	5422	GGASMWRRY
5300	EKESCQCWH	5341	ENGHKFTEE	5382	FEMAMCSMG	5423	GGATMWRRY
5301	EKHQSTCET	5342	ENGHQFNEE	5383	FKIRKFWDP	5424	GGDAMWRRY
5302	EKIMVKRPF	5343	ENGHRFNEE	5384	FKMAMCSMG	5425	GGEERVGEA
5303	EKIMVKRTF	5344	ENPEWKCMY	5385	FKVGKFWDP	5426	GGGAMWRRY
5304	EKIMVKWAF	5345	ENVHKFNEE	5386	FKVIKFWDP	5427	GIEGMYWPL
5305	EKKACQCWH	5346	EQHDWCQKY	5387	FKVRKFLDP	5428	GILIMAYSW
5306	EKKPCQCWH	5347	EQHDWYKKY	5388	FKVRKFWDS	5429	GLDCEWLHS
5307	EKKSCKCWH	5348	EQHDWYLKY	5389	FQMAMGSMG	5430	GLLYDFNPE

[2139]

5431	GLLYDFNQG	5472	IGMIRDARE	5513	IRQQWPAMF	5554	IWDRQWDAF
5432	GLPHSADVI	5473	IGMIRDTHE	5514	IRTIWGADM	5555	IWDRQWDDY
5433	GLPQSADGI	5474	IGMIWDTRE	5515	ISALMSDYQ	5556	IWDRQWGDF
5434	GLPRSADVI	5475	IGMLRDTRE	5516	ISDLMHDVH	5557	IWDRQWNDF
5435	GLSVMTYTS	5476	IGVMLNAMK	5517	ISDLMHNAH	5558	IWDRQWYDF
5436	GMHHAVWAS	5477	IGVMLNDMQ	5518	ISDLMHNGH	5559	IWDSQWDDF
5437	GMNHAVGAS	5478	ILHMMHRGA	5519	ISDLMHNVN	5560	IWHRQWDDF
5438	GMNHAVWAN	5479	IMAPLVCDD	5520	ISDLMHSVH	5561	IWKLPSVTT
5439	GMNHAVWTS	5480	IMAPLVCDG	5521	ISDLMHYVH	5562	IWKMPPIDM
5440	GMNHSVWAS	5481	IMEALVCDD	5522	ISDLMNNVH	5563	IWMLPSITT
5441	GMNNAVWAS	5482	IMEHLVCDD	5523	ISDLVHNVH	5564	IWMLPSVAT
5442	GMNQAVWAS	5483	IMEHMVCDD	5524	ISDMMHNVH	5565	IWMLPSVIT
5443	GMNYAVWAS	5484	IMEPLACDD	5525	ISELMSDYQ	5566	IWVRQWDDF
5444	GMPQSADVI	5485	IMEPLVCAD	5526	ISGLMSDDQ	5567	IWYRQWDDF
5445	GMSHAVWAS	5486	IMEPLVCDA	5527	ISGLMSDHQ	5568	KCCLGKNSR
5446	GMTHAVWAS	5487	IMEPLVCDN	5528	ISGLMSDYE	5569	KFHSKERMS
5447	GNGHKFNEE	5488	IMEPLVCDV	5529	ISGLMSDYK	5570	KINIFAYTY
5448	GQSLCNELC	5489	IMEPLVCGD	5530	ISGLTSDYQ	5571	KINIFEYTC
5449	GVPQSADVI	5490	IMEPLVCND	5531	ISGLVSDYQ	5572	KINVFEYTY
5450	GWMCEVPWT	5491	IMEPPVCDD	5532	ISNLMHNVH	5573	KITIFEYTY
5451	GWMWEVHWT	5492	IMEPRVCDD	5533	ISQHTCNKD	5574	KKDVWNNKW
5452	GWMWEVPWR	5493	IMEPVVCDD	5534	ISQHTCSED	5575	KKHVWNNKW
5453	GWMWEVTWT	5494	IMERLVCDD	5535	ISQQTCSKD	5576	KKNAWNNKW
5454	HDEMMGLER	5495	IMERMVCDD	5536	ISRHTCSKD	5577	KKNEMFMKW
5455	HDIMFAHQM	5496	IMKPLVCDD	5537	ISSADSWPM	5578	KKNEWNNKW
5456	HDIMFSHQM	5497	IMKQWPAMF	5538	ISVLMSDYQ	5579	KKNGWNNKW
5457	HDIMFTHKM	5498	IMPQWPAMF	5539	ITCIEAQEN	5580	KKNVWDNKW
5458	HDIMFTHKR	5499	IMQEWPAMF	5540	ITCIEGLEN	5581	KKNVWHNKW
5459	HDIMFTHQR	5500	IMQPLVCDD	5541	ITCIEGQED	5582	KKNVWNNEW
5460	HDIMFTYQM	5501	IMQPWPAMF	5542	ITCIEGREN	5583	KKNVWNSKW
5461	HDKMMGLER	5502	IMQQWLAMF	5543	ITCLEGQEN	5584	KKNVWNTKW
5462	HDLMFTHQM	5503	IMQQWPAKF	5544	ITEPLVCDD	5585	KKNVWNYKW
5463	HDQMMDLER	5504	IMQQWPAMY	5545	ITQHTCSKD	5586	KKNVWSNKW
5464	HDQMMGMER	5505	IMQQWPDMF	5546	IWARQWDDF	5587	KKNVWYNKW
5465	HNIMFTHQM	5506	IMQQWPSMF	5547	IWDHQWDDF	5588	KKSVWNNKW
5466	IADVVPCHH	5507	IMQQWPTMF	5548	IWDMMHRGA	5589	KKYVWNNKW
5467	IAMLESWWS	5508	IMRQWPAMF	5549	IWDRKWDDF	5590	KLLPWINER
5468	ICFHVAEMF	5509	IMVPLVCDD	5550	IWDRLWDDF	5591	KLNIFEYTY
5469	IDMLESWWG	5510	INNEFIQPF	5551	IWDRPWDDF	5592	KNDHFVCYL
5470	IGEAQDCNP	5511	IPDLMHNVH	5552	IWDRQLDDF	5593	KNGHKFNEE
5471	IGMIRATRE	5512	IPGLMSDYQ	5553	IWDRQWADF	5594	KNNEFIQPF

[2140]

5595	KPVQCACEN	5636	LVDSMLHEW	5677	LWGSMEYEE	5718	MAGDCACAM
5596	KQHDWYQKY	5637	LVDSMLHRW	5678	LWKEWCWEF	5719	MDHPNDHFF
5597	KRLPWINER	5638	LVDSMLRKW	5679	LWKKWCWEF	5720	MDHTNDLFF
5598	KTNVWNNKW	5639	LVDSMPHKW	5680	LWKMPPIDM	5721	MEWQSEYVS
5599	KWGFGSDET	5640	LVDTMLHKW	5681	LWKPRHYDF	5722	MGEQMGCQA
5600	KWLPCINER	5641	LVHSMLHKW	5682	LWKPWCWEF	5723	MGEVMDFLP
5601	KWLPGINER	5642	LVIEPCDER	5683	LWKQGCWEF	5724	MKNVWNNKW
5602	KWLPWIHER	5643	LVIKPCDER	5684	LWKQLCWEF	5725	MLGHANDGL
5603	KWLPWIKER	5644	LVIQACDER	5685	LWKQSCWEF	5726	MLGHASDAL
5604	KWLPWISER	5645	LVIQHCDER	5686	LWKQWCGEF	5727	MMEPLVCDD
5605	KWLPWVNER	5646	LVIQPCAER	5687	LWKQWCLEF	5728	MSAVMDFLP
5606	KWLQWINER	5647	LVIQPCDEG	5688	LWKQWCSEF	5729	MSEIMDFLP
5607	KWLRWINER	5648	LVIQPCDEH	5689	LWKQWCWEC	5730	MSEVMDFMP
5608	KWMPWINER	5649	LVIQPCDEP	5690	LWKQWCWEI	5731	MSEVMDFPP
5609	KWPPWINER	5650	LVIQPCDKR	5691	LWKQWCWES	5732	MSEVMDFQP
5610	KWQPWINER	5651	LVIQPCNER	5692	LWKQWCWEV	5733	MSEVMDFRP
5611	KWRPWINER	5652	LVIQPYDER	5693	LWKQWCWEY	5734	MSGVMDFLP
5612	LADSMLHKW	5653	LVIQSCDER	5694	LWKQWCWGF	5735	MSQVMDFLP
5613	LAGQMDEMS	5654	LVIRPCDER	5695	LWKQWCWKF	5736	MSWNWWALY
5614	LAIQPCDER	5655	LVLQPCDER	5696	LWKQWFWEF	5737	MSWNWWNLY
5615	LATGMLYEE	5656	LVMQPCDER	5697	LWKQWGWEC	5738	MSWNWWTVY
5616	LDDSMLHKW	5657	LVNSMLHKW	5698	LWKRWCWEF	5739	MSWTWWTLY
5617	LDQMMGLER	5658	LVTQPCDER	5699	LWMQWCWEF	5740	MTEQMGCQA
5618	LGIQPCDER	5659	LVVQPCDER	5700	LWPSMEYEE	5741	MWASMEYEE
5619	LIDSMLHKW	5660	LWAAMEYEE	5701	LWQQWCWEF	5742	MWDRQWDDF
5620	LILIMAYSW	5661	LWAFLEYEE	5702	LWRQWCWEF	5743	MWKQWCWEF
5621	LPAWQQHGR	5662	LWAFMEYEE	5703	LWTQWCWEF	5744	NACDLKNDY
5622	LQGGFDHLK	5663	LWASMAYEE	5704	LWTSMEYEV	5745	NACDWKSDY
5623	LQGGFDHLP	5664	LWASMEFEE	5705	LWVRQWDDF	5746	NACDWKTDY
5624	LQGGFDHLS	5665	LWASMEHEE	5706	LWVSMEYEE	5747	NACDWKYDY
5625	LSASMEYEE	5666	LWASMESEE	5707	LYEIRFSDM	5748	NACDWRNDY
5626	LSGLIDEMS	5667	LWASMEYDK	5708	LYEIRFVDM	5749	NADVVPCHH
5627	LSGQIAEMS	5668	LWASMEYEA	5709	LYKIRFADM	5750	NAYDWKNDY
5628	LSGQIDEMA	5669	LWASMEYEQ	5710	LYKQAPSEG	5751	NCFHVAEMF
5629	LSGQIDEMP	5670	LWASMEYEV	5711	LYMPAPSEG	5752	NFFDSEDMG
5630	LSGQIDEMT	5671	LWASMGYEE	5712	LYMQATSEG	5753	NGGCWAREC
5631	LSGQIDEMY	5672	LWASMKYEE	5713	LYVIRFADM	5754	NGGCWLREC
5632	LSMVFSRGE	5673	LWASVEYEE	5714	LYVQAPSEG	5755	NGGCWPWEC
5633	LSVQIDEMS	5674	LWENYHSEF	5715	MAEPMGCQA	5756	NGGCWSREC
5634	LSVVFSKGE	5675	LWEQWCWEF	5716	MAEQMGCRA	5757	NGGCWTREC
5635	LVDIMLHKW	5676	LWESKEMCS	5717	MAEQMVCQA	5758	NGGFWPREC

[2141]

5759	NGGGWPREC	5800	PQCHVTADR	5841	QWGFASDET	5882	SFHPSWWVY
5760	NINQLQCYC	5801	PQCLVMADR	5842	QWGFESDET	5883	SFHPTWWMY
5761	NITQLQCCC	5802	PQCRVMADR	5843	QWGFGADET	5884	SFHRSWWMY
5762	NKHEMFMKW	5803	PQSLCNELC	5844	QWGFGSDEA	5885	SFHSSWWMY
5763	NKNAMFMKW	5804	PQYHVMADR	5845	QWGFGSDEN	5886	SFHTSWWMY
5764	NKNGMFMKW	5805	PVIQPCDER	5846	QWGFGSDEP	5887	SFILRLICS
5765	NKNVWNNKW	5806	PWGFGSDET	5847	QWGFGSGET	5888	SFILRLISF
5766	NLDCEWLHS	5807	PWKIDFWYH	5848	QWGFGSHET	5889	SFILRPISS
5767	NMEPLVCDD	5808	PWKVDFWDH	5849	QWGFGSNET	5890	SFNPSWWMY
5768	NMTQLQCYC	5809	PWLSEFMHW	5850	QWGFGSVET	5891	SFNSHICRI
5769	NNDDFVCYL	5810	QAAGMLYEE	5851	QWGFVSDET	5892	SFSSHICRI
5770	NNDHCVCYL	5811	QANGMLYEE	5852	QWGIGSDET	5893	SFYNHICRI
5771	NNDHFICYL	5812	QATRMLYEE	5853	QWGSGSDET	5894	SFYPSWWMY
5772	NNDHFLCYL	5813	QATVMLYEE	5854	QWGVGSDET	5895	SFYSHICHI
5773	NNDHFVCFL	5814	QDFHEFAET	5855	QWGYGSDET	5896	SFYSHICPI
5774	NNDHFVCHL	5815	QDFHEFAPT	5856	QWKQWCWEF	5897	SFYSHICRM
5775	NNDHFVCNL	5816	QDFHEFAQA	5857	QWVFGSDET	5898	SFYSLICRI
5776	NNDHFVCYM	5817	QDFHEFDQT	5858	RDFHEFAQT	5899	SFYSNICRI
5777	NNDHIVCYL	5818	QDFHEFPQT	5859	RDIMFTHQM	5900	SFYSPICRI
5778	NNDHSVCYL	5819	QDFPEFAQT	5860	RHELLEWDS	5901	SFYSYICRI
5779	NNDHVVCYL	5820	QDIHEFAQT	5861	RHQLLEWAS	5902	SGFPMWDYW
5780	NNDHYVCYL	5821	QFMCEIECR	5862	RHQLMEWDS	5903	SGIPMWDHW
5781	NNDRFVCYL	5822	QFMFEIECK	5863	RHQLVEWDS	5904	SGIPMWDYR
5782	NNHHFVCYL	5823	QHPEWKCMY	5864	RHRLLEWDS	5905	SGIPMWEYW
5783	NPCQEWDSE	5824	QKGISEFQR	5865	RKNVWNNKW	5906	SGIPMWNYW
5784	NRGMNAWTA	5825	QKNVWNNKW	5866	RMDWVAECR	5907	SGIPVWDYW
5785	NSCDWKNDY	5826	QMDWVAGCR	5867	RNNEFIQPF	5908	SGIQMWDYW
5786	NTCIEGQEN	5827	QMDWVGECR	5868	RRQLLEWDS	5909	SGIRMWDYW
5787	NVGQITQDC	5828	QNCFFMHWW	5869	RSSTFEHHK	5910	SGNPMWDYW
5788	NWFDSEDMG	5829	QNGHKFNEE	5870	RSTTFERHK	5911	SILIMAYSW
5789	NYAENWCIE	5830	QNWFFLHWW	5871	RVIQPCDER	5912	SIYSHICRI
5790	NYRVHGGAG	5831	QPTGMLYEE	5872	RVTRDERWG	5913	SMNSLWHMR
5791	NYRVHGSAA	5832	QQCHVMADR	5873	RWGFGSDEA	5914	SQSLCNELC
5792	NYRVRGSAG	5833	QQHDWYQKY	5874	RWKQWCWEF	5915	SRAQIQDEW
5793	PATGMLYEE	5834	QVPWESRLC	5875	SCFHVAEMF	5916	SRTIWGADM
5794	PKELFPTCM	5835	QVTGMLYEE	5876	SCILRLISS	5917	SSILRLISS
5795	PLCHVMADR	5836	QVWNAHEEG	5877	SCYSHICRI	5918	SSIPMWDYW
5796	PMDWVAECR	5837	QWAFGSDET	5878	SFDPSWWMY	5919	SVPMDHCWA
5797	PNNEFIQPF	5838	QWESKEMGS	5879	SFFLRLISS	5920	SWDRQWDDF
5798	PQCHVMDDR	5839	QWESKEMYS	5880	SFHPIWWMY	5921	SWKIDFWDH
5799	PQCHVMSDR	5840	QWGCGSDET	5881	SFHPSWWMS	5922	SYAVLQRTM

[2142]

5923	SYILRLISS
5924	SYYSHICRI
5925	TADAVPCHH
5926	TADVVLCHH
5927	TADVVPCDH
5928	TAEAQDCNP
5929	TAEVVPCHH
5930	TANVVPCHH
5931	TCEAQDCNP
5932	TCEGREVDS
5933	TCEGRKLDS
5934	TCFDSEDMG
5935	TCFHAAEMF
5936	TCFHFAEMF
5937	TCFHGAEMF
5938	TCFHIAEMF
5939	TCFHVAAMF
5940	TCFHVDEMF
5941	TCFHVTEMF
5942	TCFNVAEMF
5943	TCFPVAEMF
5944	TCFRVAEMF
5945	TCGRDEFAF
5946	TCGRDLFAF
5947	TCGRDQCAF
5948	TCGRDQFAY
5949	TCGRDQFSF
5950	TCIHVAEMF
5951	TFFHVAEMF
5952	TFHPSWWMY
5953	TGDVVPCHH
5954	TGEAEDCNP
5955	TGEAQDCNA
5956	TGEAQDCSP
5957	TGFHVAEMF
5958	TGGCWPREC
5959	TGIPMWDYW
5960	TGKAQDCNP
5961	TGMSELADE
5962	TGPMDHCWA
5963	THFPWSGEA
5964	THQLLEWDS
5965	TMEPLVCDD
5966	TMSVVQCER
5967	TNDEFIQPF
5968	TNDHFVCYL
5969	TNHCIFKTM
5970	TNHEFIQPF
5971	TNIEFIQPF
5972	TNMSELADE
5973	TNNAVIQPF
5974	TNNECIQPF
5975	TNNEFFQPF
5976	TNNEFIEPF
5977	TNNEFIQLF
5978	TNNEFIQQF
5979	TNNEFIQRF
5980	TNNEFIQSF
5981	TNNEFIRPF
5982	TNNEFLQPF
5983	TNNEFVQPF
5984	TNNEIIQPF
5985	TNNEVIQPF
5986	TNNQFIQPF
5987	TNQCLFKTM
5988	TNSEFIQPF
5989	TNTEFIQPF
5990	TNYEFIQPF
5991	TRAQIQDEW
5992	TSDVVPCHH
5993	TSMNELADE
5994	TSNEFIQPF
5995	TVPMDHCWP
5996	TVPMDHCWS
5997	TWFHVAEMF
5998	TWLPWINER
5999	TYAVLKRTM
6000	VATDSVNES
6001	VGEERVGEP
6002	VGGERVGEA
6003	VGVMLNAMQ
6004	VKKIKRCEG
6005	VMEPLVCDD
6006	VMYPWISCL
6007	VNDDDQCCM
6008	VNDHDQCCM
6009	VNDNDQCCM
6010	VNDYAQCCM
6011	VNDYDKCCM
6012	VNDYDQCCR
6013	VNDYDQCCV
6014	VNDYDQCFM
6015	VNDYDQCWM
6016	VNDYDQCYM
6017	VNDYEQCCM
6018	VNDYVQCCM
6019	VNGHKFNEE
6020	VNGYDQCCM
6021	VNHYDQCCM
6022	VNNYDQCCM
6023	VQKLFNSTV
6024	VQLLSLCPS
6025	VSDLMHNVH
6026	VSGQIDEMS
6027	VVIQPCDER
6028	VWAAMEYEE
6029	VWDCQWDDF
6030	VWEMPPIDM
6031	VWESSSIIS
6032	VWKKPPIDM
6033	VWKMLPIDM
6034	VWKMPPIDT
6035	VWKMPPIDV
6036	VWKMPPLDM
6037	VWKMPPMDM
6038	VWKMPPVDM
6039	VWKMPQIDM
6040	VWKMPSIDM
6041	VWKMSPIDM
6042	VWKQWCWEF
6043	VWKRPPIDM
6044	VWKTPPIDM
6045	VWKVPPIDM
6046	VWQMPPIDM
6047	VWTMPPIDM
6048	WAKCQFGAT
6049	WCRMDNKKS
6050	WCRWPFYSE
6051	WHCDAMECF
6052	WILIMASSW
6053	WILIMAYAW
6054	WILIMAYSL
6055	WILIMAYTW
6056	WILIMEYSW
6057	WILIMSYSW
6058	WILIMTYSW
6059	WILIMVYSW
6060	WILSMAYSW
6061	WIMIMAYSW
6062	WIPIMAYSW
6063	WIQIMAYSW
6064	WIRIMAYSW
6065	WIVIMAYSW
6066	WLIHLWTQF
6067	WLIPLWAQF
6068	WLIPLWPQF
6069	WLIPLWTEF
6070	WLIPMWTQF
6071	WLIPPWTQF
6072	WLIPVWTQF
6073	WLIRLWTQF
6074	WLLIMAYSW
6075	WLLPLWTQF
6076	WPSIWKNEQ
6077	WPSIWNNER
6078	WQMRMQNKF
6079	WSKCQFGAA
6080	WSKCQFGAM
6081	WSKCRFGAT
6082	WSKFQFGAT
6083	WSKWQFGAT
6084	WSLIMAYSW
6085	WSQCQFGAT
6086	WSRCQFGAT

[2143]

6087	WTIACIQRG
6088	WTNDCIQRG
6089	WVDSMLHKW
6090	WVLIMAYSW
6091	YASYKAMED
6092	YCFDSEDMG
6093	YDIMFTHQM
6094	YDQMMGLER

6095	YFAADERWG
6096	YFQQKHNLF
6097	YFTADEQWG
6098	YGGCWPREC
6099	YISDSVSHG
6100	YKVRKFWDP
6101	YPNGMTFEP
6102	YQEHFHNLI

6103	YQENFHNWI
6104	YQNDMTFEP
6105	YRSHEKFQF
6106	YRSPEKFEF
6107	YRSPEQFQF
6108	YRTPEKFQF
6109	YSDEMDWSK
6110	YSNEMDWNK

6111	YSSSKAMED
6112	YVGDSVSHG
6113	YVSDCVSHG
6114	YVSDGVSHG
6115	YVSDSVNHG
6116	YVSDSVSHS
6117	YVSDSVSRG

[2144] 实施例33

[2145] 如通过机器学习鉴定的具有结肠组织嗜性的AAV5变体

[2146] 此实施例描述了如通过机器学习方法鉴定的展现结肠组织嗜性的工程化AAV5变体的优选特性。进行如实施例19和图19至图22中描述的相同生物信息学预处理、机器学习方法和分析以确定展现结肠组织嗜性的工程化AAV5变体的优选特性。

[2147] 下文描述了AAV5 VP衣壳多肽的581至589区域中各位置处的有利生物物理特性和有利氨基酸残基，它们与较高的结肠组织嗜性概率相关。表50中描述的任何以下机器学习衍生的位置氨基酸偏好可以单独存在或彼此组合存在或与实施例14，表27中观察和描述的任何氨基酸偏好组合存在。下文表50中列出了从ML模型中提取的排名靠前的重要结肠嗜性特征之间共有的位置特征的概述。

[2148]

表 50
机器学习衍生的结肠组织嗜性规则

Xaa1 处高摩尔质量 Xaa1 选自 Y 或 W
Xaa1 处高溶解性 Xaa1 选自 W、F、I 或 L
Xaa2 处低溶解性 Xaa2 选自 D

[2149]

表 50 机器学习衍生的结肠组织嗜性规则
Xaa2 处低可突变性 Xaa2 选自 P 或 K
Xaa2 处中等摩尔质量 Xaa2 选自 D、E、N、K、M、Q、I 或 L
Xaa2 处低亲疏水性 Xaa2 选自 D、E、R、K、H、N 或 Q
Xaa3 处低可突变性 Xaa3 选自 K、V、P 或 C
Xaa3 处高溶解性 Xaa3 选自 W、F、I 或 L
Xaa5 处高平均柔性 Xaa5 选自 S、P、G、R、E 或 D
Xaa5 处高表面可接近性 Xaa5 选自 D 或 N
Xaa6 处低亲疏水性 Xaa6 选自 R
Xaa6 处低可突变性 Xaa6 选自 Y、R、F 或 L
Xaa6 处低溶解性 Xaa6 选自 R 或 Q
Xaa6 处高表面可接近性 Xaa6 选自 E、R 或 K
Xaa6 处高平均柔性 Xaa6 选自 G 或 R
Xaa8 处低溶解性 Xaa8 选自 D

[2150] 下文表51提供了表现出结肠组织嗜性并且符合表50中提供的一种或多种规则的

变体AAV VP1衣壳多肽的581至589区域的序列。因此，本公开提供了由具有SEQ ID NO:2的工程化AAV5 VP1衣壳多肽组成的rAAV，其中Xaa1至Xaa9区域是SEQ ID NO:11118-SEQ ID NO:12117中的任一者，如表51中所公开。本文还涵盖由工程化AAV5 VP2衣壳多肽和工程化AAV5 VP3衣壳多肽组成的rAAV，所述多肽在对应于AAV5 VP1 581至589区域中的氨基酸的AAV5VP2(氨基酸残基445至453)和AAV5 VP3(氨基酸残基389-397)中的区域处具有下文表中公开的序列。

[2151] 表51

[2152] 驱动结肠组织嗜性的AAV5 VP1衣壳多肽中581至589区域的序列

[2153]

SEQ ID NO	581-589 序列
11118	ACCKVRRDC
11119	ACPVEFVPW
11120	ACYIDAAPN
11121	ACYLRNEDR
11122	ADASRTELD
11123	ADASRTEPD
11124	ADASRTKRD
11125	ADERPNATD
11126	ADIMKEIIF
11127	ADIMKESIV
11128	ADIMKVSIF
11129	ADIMMESIF
11130	ADIMQESIF
11131	ADIMRESIF
11132	ADINRWSYV
11133	ADITKESIF
11134	ADLHNNHEV
11135	ADLNNNHED
11136	ADRDGPHWY
11137	ADRGGPHWD
11138	AECCLEPKH
11139	AEDGADMIT
11140	AEGCKFYNG
11141	AELMMGSII
11142	AEQQHRLIE
11143	AFCRGHYKA
11144	AFCRGHYPA
11145	AFCRGHYRA
11146	AFCRGLYQA
11147	AFCRGRYQA
11148	AFDTFYAFS
11149	AFHHDRKYY
11150	AFKEFSLCC
11151	AFLYAAGWC
11152	AFNLDRKYY
11153	AFNQDRKYY
11154	AFPQNEHQR
11155	AFVTFYAFR
11156	AGPVEFFPW
11157	AGRQEKPDG
11158	AGRQEKPYV
11159	AGWQKCYDS
11160	AGWQMCDDS
11161	AGWQTCYDS
11162	AHDLATEWE
11163	AIDYLIDIL
11164	AINHEKTDW
11165	AKDPKPIRK
11166	ALNHDRKYY
11167	ALPDTNHAD
11168	ALSNDNAIL
11169	ALSNHNAVL
11170	AMHMTRGKC

[2154]

11171	AMIWPQVLD	11212	CMKIAFWGC	11253	DKHHSRIDV	11294	EGDEQEFVG
11172	AMQKTRGKC	11213	CMKMAFWEC	11254	DKHHSRIHG	11295	EGFPFEGIS
11173	ANCDMKDPW	11214	CMPTARICM	11255	DKHHSRITG	11296	EGFRFEGII
11174	ANCDMTDRW	11215	CQIWAKNEY	11256	DKHHSRIYG	11297	EGHNYEAML
11175	ANCNKRHNW	11216	CQIWGKNEF	11257	DKHHSRMDG	11298	EGIYFVKHD
11176	ANNHEKTDL	11217	CQLWAKNEF	11258	DKHHSRNDG	11299	EGQNYEALL
11177	ANNHEKTEW	11218	CQPHWEKGD	11259	DKHHSRVDG	11300	EGQNYEAVL
11178	ANNHVKTDW	11219	CRKMMQRIS	11260	DKHLSRIDG	11301	EGQNYESML
11179	APPPHEVKI	11220	CSGSGHLIY	11261	DKHNSRIDG	11302	EGQNYETML
11180	APPPHEVKS	11221	CWDERNVFC	11262	DKHPSRIDG	11303	EGRNYEAML
11181	APPPREVKF	11222	CWFMGEGCH	11263	DKHRSRIDG	11304	EIAYDYFPM
11182	APPQMCDFI	11223	CWLIWKLDR	11264	DKLHSRIDG	11305	EIILKDMCH
11183	AQFKDRSSF	11224	CWVLWKLDR	11265	DKNHSRIDG	11306	EIILKDMWN
11184	ARPKNLPLS	11225	DCKCDMFNC	11266	DKPHSRIDG	11307	EIIVKDMWH
11185	AVTNHCIDR	11226	DCPLSNSYE	11267	DKYHSRIDG	11308	EKAPGFICF
11186	AVWQMCYDS	11227	DDKIAVGWP	11268	DMHHSRIDG	11309	EKCGPCARY
11187	AWDGSIEFL	11228	DDKIAVRWH	11269	DNGFKLQKS	11310	EKCMCHDIS
11188	AYDAPRWFM	11229	DDVDCCDRK	11270	DNQCSDKGL	11311	EKHDGLMRR
11189	AYFQARLLL	11230	DFIMKYHAM	11271	DPPNDMGAV	11312	EKPDGLMRQ
11190	AYIMKESIF	11231	DFNHDRKYY	11272	DQHHSRIDG	11313	EKPEPRYKE
11191	AYNHDRKYY	11232	DFPQNEHQR	11273	DRDCNLTSR	11314	EKPQPWYKE
11192	CDKMCAPQA	11233	DFQEFSLCC	11274	DRHHSRIDG	11315	EKPSITCRE
11193	CDPHGMVGC	11234	DGFMDFKHH	11275	DSRCPAAYF	11316	EMCMDVPSV
11194	CDRKGAGNC	11235	DGMRDFKPG	11276	DTFDRIFCS	11317	EQCLAAAGE
11195	CDRMYAPQA	11236	DGVVPVTNL	11277	DWDHMFEME	11318	EQCLAAAVE
11196	CERKGAGHC	11237	DGYMDFKHL	11278	DWSHIFGRG	11319	EQCLAAGAE
11197	CERKGAGNG	11238	DGYMDFKRH	11279	DYNWHACSI	11320	EQCLAAPAE
11198	CERKGSGNC	11239	DHALATEWE	11280	EAIHRQACT	11321	EQCLAASAE
11199	CERRGAGNC	11240	DHDLATGWE	11281	EDIMKESIF	11322	EQCVAAAAE
11200	CFCKCGMPS	11241	DHDLTTEWE	11282	EDNKWHGWH	11323	EQDIYGAIV
11201	CFDCSRAHP	11242	DHDLVTEWE	11283	EEAPGFFCF	11324	EQPVGQSHR
11202	CFDRGKAFW	11243	DHVLATEWE	11284	EEAPGFICY	11325	EQTEGQSHR
11203	CGFPGFSPV	11244	DKDCDLTSR	11285	EEATGFICF	11326	ESINRQACT
11204	CGRWNFQYL	11245	DKDGDRFYI	11286	EEDPGFICF	11327	EVAPGFICF
11205	CGSSGHLIY	11246	DKDHSRIDG	11287	EEGPGFICF	11328	EVEYDYFPM
11206	CHPLTAPSV	11247	DKHDSRIDG	11288	EEGPGFIFF	11329	EVIYFVQHD
11207	CIHTCESII	11248	DKHHGRIDG	11289	EETPGFICF	11330	EVLYFVKHD
11208	CINPCESII	11249	DKHHSRFDG	11290	EFKEFSLCC	11331	EVWCQVEPN
11209	CLLDGKSDL	11250	DKHHSRIDC	11291	EFPIDRRFP	11332	EWACAHAMR
11210	CMCPPRLEV	11251	DKHHSRIDD	11292	EGDEHEFAG	11333	EWDQPNYIE
11211	CMKFAFWEC	11252	DKHHSRIDR	11293	EGDEPEFAG	11334	EWVCAHAMR

[2155]

11335	EYFQARLWL	11376	FLGCYRPYA	11417	FWTGRNPHN	11458	GNCNKRHNW
11336	EYRQPQSIA	11377	FLGFYCPYA	11418	FYPGERHNG	11459	GPKKGGQAM
11337	FACPVLWGQ	11378	FLGMGYAIF	11419	GCDGYALAI	11460	GPKQGGQSM
11338	FANPVLWGQ	11379	FLGNTWLGP	11420	GCDPYNMVD	11461	GPVLPDVGG
11339	FASPILWGQ	11380	FLGWYCPYA	11421	GDDDCCDRK	11462	GPVMPDVGA
11340	FASPVLWGR	11381	FLGYYCPYA	11422	GDHHGMVGC	11463	GRSSHPPIY
11341	FCWRDNEPE	11382	FLRNAWLGP	11423	GDKIKFSDR	11464	GSKSGGADD
11342	FDGPSVRGE	11383	FLRNTWLDP	11424	GDKMGICHC	11465	GSKYGGDDD
11343	FDKMYAPQA	11384	FLRNTWLGA	11425	GDKMGIHHC	11466	GSNSGGDDD
11344	FDMCIEYSN	11385	FLRNTWLGS	11426	GDKMYAPQA	11467	GVRRVKIEW
11345	FDMWGVGIP	11386	FLRNTWLSP	11427	GDPHGMGGC	11468	GWIEDKREV
11346	FDNHANTWA	11387	FLRNTWLVP	11428	GDPHGMLGC	11469	GWSEEKREV
11347	FDSPVLWGQ	11388	FMCNIGCFT	11429	GDPPGMVGC	11470	HDDWKKQPN
11348	FEQTRFPMQ	11389	FMGCYCPYA	11430	GDPQGMVGC	11471	HGAERFDRF
11349	FEREFHYYD	11390	FNVCTMWAQ	11431	GDVDCCDRT	11472	HGAERVDRV
11350	FEREFKYYD	11391	FPDCKNKAF	11432	GDYHYRVNW	11473	HGLVDFDPH
11351	FERPPKLGI	11392	FPLRENDDC	11433	GEAPGFICF	11474	HKHHSRIDG
11352	FEWEFQYYD	11393	FPTCWRYAK	11434	GENKVNDPA	11475	HNHTRKDPV
11353	FHDNARLFL	11394	FPTIGLMAC	11435	GEYHCRVNW	11476	HVANEEVGG
11354	FHLREKDDC	11395	FPYCKNKPF	11436	GEYKLNDPA	11477	HVGNEEVSG
11355	FHYDCQDKV	11396	FQRNTWLGP	11437	GEYKVNDPP	11478	HWRVQLPWN
11356	FHYDCRDNV	11397	FRIRCRVNK	11438	GEYRVNDPA	11479	ICHANNVRV
11357	FIAHKGQQW	11398	FRIRRRVNK	11439	GFCRGHYQA	11480	ICKARIEKY
11358	FIEFHVILV	11399	FRRNTWLGP	11440	GFMCNLCGP	11481	ICKPRLEKY
11359	FIGKGYAIF	11400	FSEQHAVND	11441	GFQTKFHLC	11482	ICPANNVRG
11360	FIGMGYAIC	11401	FSEQHAVNT	11442	GGDGYAIAI	11483	ICPANNVRL
11361	FIVHKGRQW	11402	FSEQHVVNN	11443	GGRHGYLHD	11484	ICPVNNVRV
11362	FIVLKGQQW	11403	FSERHAVNN	11444	GGWQMCYDS	11485	ICRPRIEKY
11363	FIVMGYAIF	11404	FSGWFRVCD	11445	GGYKVNDPA	11486	ICSANNVRV
11364	FIVNKGQQW	11405	FSPVERHNG	11446	GIIHHCVSE	11487	ICSIDNNNY
11365	FLERYQHMD	11406	FTSPVLWGQ	11447	GIIRHCFSE	11488	ICSMDNKNY
11366	FLFMEPRNG	11407	FVCMEPRNG	11448	GIIRHCLSE	11489	IDHDNWDMR
11367	FLGCYCHYA	11408	FVFMDPRNG	11449	GIIRHCVIE	11490	IEKDEMPVC
11368	FLGCYCLYA	11409	FVFVEPRNG	11450	GIIRNCVSE	11491	IERPPKLGT
11369	FLGCYCPDA	11410	FVKGEFHFH	11451	GKHHSRIDG	11492	IFLNIFELR
11370	FLGCYCPSA	11411	FVQPFLHDG	11452	GLRDGNSIF	11493	IFLNIFEML
11371	FLGCYCPYG	11412	FWCRPRLHY	11453	GMIPYNWSW	11494	IGANGWQCC
11372	FLGCYCPYS	11413	FWPGHNPHN	11454	GMPATELWL	11495	IGLVDFDPH
11373	FLGCYCRYA	11414	FWSRMIPLV	11455	GMRRDGLDV	11496	IHAPKLLWL
11374	FLGCYFPYA	11415	FWTGHNPHI	11456	GMRWDRDQL	11497	IHCPDSWSQ
11375	FLGCYGPYA	11416	FWTGHNPPN	11457	GMSSHTPIY	11498	IHCPEAWSQ

[2156]

11499	IHCPESWSL	11540	KDIRGIWDI	11581	KQCLAAAAE	11622	LSICKPVMP
11500	IHGPCTPWD	11541	KDIRVIWDI	11582	KSEWQLMYG	11623	LWDHGASYK
11501	IHHEGRSVC	11542	KDMRDIWDI	11583	KTADEFDCD	11624	LWHHGASYM
11502	IHPPKLLWL	11543	KDSWAGYQM	11584	KWERYRDFI	11625	LYDQSRSQL
11503	IHSPKLLWL	11544	KDWMPSYAL	11585	KYKREEYEK	11626	LYPCKMPMN
11504	IHTNMRVIS	11545	KDYWAEYQM	11586	LCKLPEIRK	11627	LYYPSRSQL
11505	IHTNMRVIV	11546	KDYWAGCQM	11587	LDAKMVQPS	11628	LYYQSRSKL
11506	IHTPRLLWL	11547	KDYWAGYKM	11588	LDCDSVWSP	11629	MCKPRIEKY
11507	IIKFEWQEV	11548	KDYWAGYRM	11589	LDLKMVQPS	11630	MCLIDGKDW
11508	IIKFEWQKF	11549	KDYWAVYQM	11590	LEGEGMWAF	11631	MCSIDNKNY
11509	IIKFEWQQV	11550	KEGCHGRMG	11591	LERKPRNDC	11632	MDKHGCCCL
11510	IIKFEWQRV	11551	KEGCHVRVG	11592	LERKPRNYF	11633	MDKLGCCCW
11511	IKDQNKEWN	11552	KELSCDQNW	11593	LERKPRSYC	11634	MDMQGGKVC
11512	ILRNTWLGP	11553	KFLLAFEVT	11594	LERQPRNYC	11635	MDQHGCCCW
11513	IMNEAHYRF	11554	KFLYAAGWC	11595	LESEGMWSF	11636	MDWCVQQSS
11514	IMNKDHYRF	11555	KGIVTDRIL	11596	LHHEGRSVC	11637	MDWCVQQST
11515	IMNRAHYRF	11556	KHENKKDVL	11597	LKHPPQPKV	11638	MEADHKSDP
11516	INGSNTLTN	11557	KIMQFCWDF	11598	LKQINRHPC	11639	MKGCFVHHS
11517	IPECWWRWH	11558	KKFLCWEWS	11599	LMDAIDSIW	11640	MKGYFIHHS
11518	IPEFWLRWH	11559	KLCNPVVFV	11600	LMVAIDSIW	11641	MKGYFVHHY
11519	IPEIWWRWH	11560	KMCAYQPDT	11601	LNYCDQYGD	11642	MMPCMVNFG
11520	IPKDEMPVC	11561	KMCHDERQV	11602	LNYCDQYVG	11643	MNKCIENRG
11521	IPPLKTEDN	11562	KMCPDERKV	11603	LNYERRDNY	11644	MNKWIENRW
11522	IPRNVYMCD	11563	KMCPDERQL	11604	LPAYTTPIV	11645	MQFVDRLAM
11523	IPRTVYMCD	11564	KMCTDERQV	11605	LPFGNKFPD	11646	MQFVDRWAL
11524	IQKDEMPFC	11565	KMFPDERQV	11606	LPGPVTVSW	11647	MRCCNVPWS
11525	IQKDEMPVW	11566	KMGPDERQV	11607	LPKPHFDKQ	11648	MRDRMKDEE
11526	IQKYEMPVC	11567	KMLIDEPTL	11608	LPKPYFDKK	11649	MVIHPAISR
11527	IQRDEMPVC	11568	KMLIDQATL	11609	LPLLKTEDN	11650	MWDHRNETM
11528	IRPNDSSFH	11569	KMLMDEATL	11610	LPLTDAPMV	11651	MWDNRNESM
11529	IRPNEGSFH	11570	KMLNDEATL	11611	LPPLETEDN	11652	NAMKRSPDV
11530	IRPNESSFP	11571	KMRIDEATL	11612	LPPLKTEAN	11653	NAMQRSPDF
11531	ISGKEMNST	11572	KMWPDERQV	11613	LPPLKTEDD	11654	NCINATDGD
11532	ISNEMSKIL	11573	KNFEELDND	11614	LPPLKTEDT	11655	NCLVDFDPH
11533	IVADNTVCA	11574	KNFEKLDNE	11615	LPPLKTKDN	11656	NCMNATDWD
11534	IVDDHTVCA	11575	KNFEKLDYD	11616	LPPLKTVDN	11657	NCNNELTDR
11535	IVDTNTATK	11576	KNFEKLNND	11617	LPPLRTEDN	11658	NDMAQSYRE
11536	IVPRFLVEA	11577	KNKCIENRW	11618	LQRLGYSPD	11659	NDMEQSDRE
11537	IWPANNVRV	11578	KPCPVRPPH	11619	LQTYTTPIV	11660	NDMEQSYHE
11538	KDDWAGYQM	11579	KPDDEFDCD	11620	LRIDPKDSF	11661	NELAYQDWS
11539	KDIREIWDI	11580	KPVYLYCSD	11621	LSICHPVMP	11662	NELEYQHWS

[2157]

11663	NFCPENAWQ	11704	NMPWNMPWN	11745	PFCRGHYQA	11786	RCDLKEMWK
11664	NFEDQESHV	11705	NMPWNMQLN	11746	PFISPKQVS	11787	RCYLKEMRK
11665	NFGDLESHV	11706	NMPWNMQWK	11747	PFISPREVS	11788	RDCHMRITD
11666	NFITQEMYS	11707	NMPWNMQWS	11748	PFIYTDLGP	11789	RDCWMVFII
11667	NFNTQEMYS	11708	NMPWNVQWN	11749	PFLYTDWGP	11790	RDFVAGQQP
11668	NFTTKEMYS	11709	NMQTKSAHK	11750	PFNNWFAAG	11791	RDFVGEQQP
11669	NFTTMNMTC	11710	NNHMVAKIA	11751	PGWQMCYDS	11792	RDGKWALEN
11670	NFTTQEMYY	11711	NPCHGYLKM	11752	PINFPYEWL	11793	RDGKWVREN
11671	NFWMDMHRI	11712	NPPCARECR	11753	PKFAFNQCY	11794	RDYHMLITD
11672	NGGTWANVE	11713	NPPCARGGR	11754	PKGMINWPS	11795	RGFHHADNV
11673	NGHNELTDR	11714	NPPCARKWR	11755	PKVRINWPS	11796	RGFNHADNG
11674	NGLDDFDPH	11715	NPPCARQWR	11756	PKVVINWPS	11797	RHIYRKGWG
11675	NGLVDCDPH	11716	NPRCGIFTL	11757	PMCMEPFPM	11798	RLRDGNSIF
11676	NGLVDFAPH	11717	NQGLPCCQI	11758	PMNRDTLIS	11799	RMCPYEAWK
11677	NGLVDFDPL	11718	NQGMHCCQI	11759	PPPLNRKWD	11800	RMCTYEPWK
11678	NGLVDFDRH	11719	NQGMPCCRI	11760	PQAYTTPIV	11801	RPEFNAFQA
11679	NGLVDFDTH	11720	NQNNLEYPI	11761	PQLYFTTDK	11802	RSFIDIVGT
11680	NGLVDFVPH	11721	NRHPPAAIS	11762	PQTKKDCPA	11803	RTQPIGVDP
11681	NGLVHFDPH	11722	NRNPPAAII	11763	PRKMMQRIS	11804	RYCWMVVII
11682	NGNNELADR	11723	NRNPPEAIS	11764	PSDLKLVNN	11805	SADYNFGFS
11683	NGNNELTDK	11724	NRNSFHQFE	11765	PSTIKHVDP	11806	SCGWDHGNL
11684	NGNNGLTDR	11725	NRPWQMRWA	11766	PSTIKRIDP	11807	SCPANNVRV
11685	NGRVDFDPH	11726	NRTSFHKFE	11767	PWDGSIEFL	11808	SCPVEFFPW
11686	NGWEGSEDD	11727	NSITANCSK	11768	QCLLWNDCK	11809	SDASRTERD
11687	NGWSDVDRQ	11728	NSITENCAK	11769	QELFHGKIA	11810	SDKMGIYHC
11688	NHNVPACII	11729	NSITENCSN	11770	QFPRHEPIQ	11811	SDLHNNHED
11689	NIEQMLAIG	11730	NTKVKDEDN	11771	QFPRQEPLQ	11812	SDNWGAQCC
11690	NKHHSRIDG	11731	NVLGPPSFN	11772	QFYNRSDDL	11813	SDPHGMVGC
11691	NMHSEQTPL	11732	NVNNELTDR	11773	QIDQFCPIV	11814	SFGPSQVCC
11692	NMKFSVGMN	11733	NVTANYCCF	11774	QINPFYMWF	11815	SFIGNAQWC
11693	NMKISVDMN	11734	NYDHFIAAS	11775	QINQFCQIV	11816	SFIGNARWC
11694	NMKTESAHK	11735	NYGKDMHKL	11776	QMCPDERQV	11817	SFLAHNNFH
11695	NMKTKSADK	11736	NYNNFIAAS	11777	QPCLVRPPH	11818	SFLANNHFH
11696	NMKTKSAHE	11737	NYRESRSDL	11778	QPCTVRPPH	11819	SFLANNTFH
11697	NMKTKSAHT	11738	PAIHNHMCT	11779	QSMWETDRN	11820	SFNHDRKYY
11698	NMKTKSAPK	11739	PAIRNHMCP	11780	QWCWMAVSI	11821	SFYEWPHCE
11699	NMKTKSARK	11740	PAYKNSGGD	11781	QWNMFQVIF	11822	SGICCCDNA
11700	NMKTKSDHK	11741	PDFVAEQQP	11782	RAAKVIPPT	11823	SGIGCCDNS
11701	NMKTKTAHK	11742	PETLWANIW	11783	RADKFIPPT	11824	SGWKMCYDS
11702	NMKTRSAHK	11743	PETLWTDIW	11784	RADKVITPT	11825	SGWQMCYDS
11703	NMPWNMEWN	11744	PETPWTNIW	11785	RAYKVIPPT	11826	SHDQALMFE

[2158]

11827	SHFYEDARS	11868	TRNPPAAIS	11909	VNCDKRHNW	11950	WDHNCQATI
11828	SIIRHCVSE	11869	TRPSAASIR	11910	VNCNKRHNL	11951	WDHNYQDTI
11829	SIQPFLHDG	11870	TRPSAASVG	11911	VNCNKRHNM	11952	WDNNCQDTI
11830	SLNGYTLVC	11871	TRPSASSIG	11912	VNCNKRRNW	11953	WDQNCQDTI
11831	SLPEYSNVK	11872	TSICYDQIQ	11913	VNCNKRYNW	11954	WERMGQIEC
11832	SNPWPMGKV	11873	TSPPDQVPV	11914	VNCNNRHNC	11955	WERMGQIED
11833	SPHHDKDPF	11874	TWDGSIEFL	11915	VNCNRRHNW	11956	WERMGQIEH
11834	SPMINLWPS	11875	VADHRPMYS	11916	VNCSKRHNW	11957	WERMGQIES
11835	SRKMMKRIS	11876	VAIVDGWIE	11917	VNCYKRHNW	11958	WERMGQIQY
11836	SRKMMPRIS	11877	VAPSESSWE	11918	VNGNKRHNW	11959	WERMGQMEY
11837	SRQCPLEIG	11878	VCLVDGKDW	11919	VNLQDLSDR	11960	WEWDRPFMK
11838	SSDLKLVNS	11879	VCPANNVRV	11920	VNPSDQTNH	11961	WFFHYKPWM
11839	SSGGDHGNL	11880	VCTIYYMCG	11921	VNSSDQNNH	11962	WFMKCNWWT
11840	SSIHHQCCT	11881	VDCNKRHNW	11922	VPDFDPFWS	11963	WGCCEEMDS
11841	SSIQHHCCT	11882	VDHCTAIGA	11923	VPIFDGWIE	11964	WINTGPLRH
11842	SSVWDHGNL	11883	VDIFDGWIE	11924	VPNFAAFCM	11965	WKIKPDHRV
11843	SVQLFLHDG	11884	VDPFCAVRI	11925	VPPACRHKK	11966	WKLKNCKSD
11844	SVQPFLHDA	11885	VDPMGEAGT	11926	VPPGCRHQK	11967	WLFEGKLDT
11845	SVQPFLYDG	11886	VDVDCCDRK	11927	VPTFDPFWS	11968	WLFYGCEPN
11846	SVQPVLHDG	11887	VEFADNPWN	11928	VQFVDRWAM	11969	WLPHGPRME
11847	SYDNWDFCC	11888	VERGNKIAS	11929	VQTKKDCPA	11970	WLPNRRIDD
11848	TCSIDNKNY	11889	VERSNKISS	11930	VRFDMCKWT	11971	WLQDGNSIF
11849	TCTWIGAGS	11890	VFCRGHYQA	11931	VRFEMMSWT	11972	WLRDGDSIF
11850	TDASRTERD	11891	VGTKEGPDP	11932	VTDFHTDML	11973	WLRDGNSIS
11851	TDIMKESIF	11892	VHHEGRSFC	11933	VTDFNTDMQ	11974	WLRDGNSIY
11852	TDYWAGYQM	11893	VHHEGRSVF	11934	VVFINYNDQ	11975	WLREGNSIF
11853	TFPDGEMYQ	11894	VHHKGRSVC	11935	VVFTNYNDK	11976	WLRGGNSIF
11854	TGLVDFDPH	11895	VHLRENDDC	11936	VVPPGNVIL	11977	WLRVGNSIF
11855	TGTWIGTGS	11896	VHPEGRSVC	11937	VVPPGNVLP	11978	WMGWDRDQL
11856	TGWQMCYDS	11897	VHQEGRSVC	11938	VVYRDTCNG	11979	WMPNRRIDA
11857	TIDDHSFDQ	11898	VHREGRSVC	11939	VWDHRNESM	11980	WMPNRRIDG
11858	TINCKWCFC	11899	VHYEGRSVC	11940	VYNWHACAI	11981	WMRCDRDQL
11859	TINPHKSDW	11900	VISRECYEE	11941	WAEDAYRRF	11982	WMRLDRDQL
11860	TISCKWCVC	11901	VKCKDKYQS	11942	WCACVFEYW	11983	WMRWARDQL
11861	TNICYDQIK	11902	VKCQDKYQY	11943	WCERVFEYW	11984	WMRWDQDQL
11862	TNICYDQIR	11903	VKHHSRIDG	11944	WCEWVFEYW	11985	WMRWDRAQL
11863	TNNHEKTDW	11904	VLGLPPYPE	11945	WCFKKELCN	11986	WMRWDREQL
11864	TPCCLHHPA	11905	VLLYMRNNQ	11946	WCKLPEIRR	11987	WMSDMLKAV
11865	TPPPHEVKF	11906	VMAAEQLNQ	11947	WCQCVFEYW	11988	WMWWDRDQL
11866	TQMFAMPPS	11907	VMIWPQILD	11948	WCQLPEIRK	11989	WNIPNRSQA
11867	TQPPLMNIL	11908	VMPTDCILV	11949	WCYLMNLES	11990	WNNEHLKVW

[2159]

11991	WNNFYCGCN	12023	YDINYYPKD	12055	YHTNNVEPW	12087	YNMVHGDGP
11992	WPFPVQFCD	12024	YDINYYPNE	12056	YHTSNMEPW	12088	YNMVHSDEP
11993	WPHHAKDPF	12025	YDIPKSTCS	12057	YHTTNMEPW	12089	YNPRIALMN
11994	WPHHDKDPV	12026	YDNSKYCKN	12058	YHTYNMEPW	12090	YNPRSALMN
11995	WPHHNKDPF	12027	YDTNYYPKE	12059	YIVHKGQQW	12091	YNTVHGDEP
11996	WPRDGNSIF	12028	YDYGGRVQK	12060	YIWLVMFFS	12092	YNVCTLWAQ
11997	WPSEGLIGA	12029	YEAKSLQPN	12061	YKAGDRFYI	12093	YNVCTMLAQ
11998	WPVPVQVCD	12030	YEENVLFMG	12062	YKDCDRFYI	12094	YNVVHGDEP
11999	WQIWAKNEF	12031	YEINYYPKE	12063	YKDDDRFYI	12095	YPEDRLNMR
12000	WRFDGCEPN	12032	YETGMWLNR	12064	YKDGDRIYI	12096	YPLGSAGSD
12001	WRFYGCEPT	12033	YFDCSRAHS	12065	YKDSDRFYI	12097	YPQHNACMQ
12002	WRFYGCERN	12034	YFDCSRANP	12066	YKHHSRIDG	12098	YPTNNMEPW
12003	WRFYGCESN	12035	YFDCSRGHP	12067	YKYGDRFYI	12099	YQAAARQNA
12004	WRGCMFMSK	12036	YFDCSRTHP	12068	YLPSAELIK	12100	YQADPRQNA
12005	WRNCMFMSK	12037	YFDRSRAHP	12069	YMWICPPDQ	12101	YQATPRQNA
12006	WRSYGCEPN	12038	YFDWSRAHP	12070	YMWIHPPDQ	12102	YQAYGRGHV
12007	WSSEGLIGA	12039	YGCNRSYQD	12071	YMWISPPDQ	12103	YQDAPRQNA
12008	WVWNGCRDM	12040	YHANNMEPW	12072	YMWIYAPDQ	12104	YQIFQYPLH
12009	WYAQCWKAF	12041	YHDLATEWE	12073	YMWIYPPDK	12105	YQLFQYPWH
12010	WYCWMVFII	12042	YHDNALLFL	12074	YMWIYPPDL	12106	YQLLHTIEN
12011	WYDAGHKCK	12043	YHENVLFMG	12075	YMWIYPPDP	12107	YQLLHTIEQ
12012	WYDAGNNCK	12044	YHINNMEPW	12076	YMWIYPPGQ	12108	YRTLQCKWE
12013	WYDAVNKCK	12045	YHPHHMYWR	12077	YMWIYPTDQ	12109	YRWIYPPDQ
12014	WYDDGNKCK	12046	YHPRTALMN	12078	YMWIYQPDQ	12110	YRYLWPIVR
12015	WYECVFEYW	12047	YHQHNACMR	12079	YMWNYPPDQ	12111	YTGLNELQN
12016	WYYAGNKCK	12048	YHTIDMEPW	12080	YMWSYPPDQ	12112	YTICDVPDP
12017	YAHYCWDPI	12049	YHTINMEPW	12081	YMWTYPPDQ	12113	YTPNFRIGD
12018	YAVAVGISS	12050	YHTNNMEPL	12082	YNMAHGDEP	12114	YVIPFEMDD
12019	YDFGGRVQT	12051	YHTNNMERW	12083	YNMVDGDEP	12115	YVIPVEMDN
12020	YDHHKIIWA	12052	YHTNNMETW	12084	YNMVHGDEA	12116	YYDCSRAHP
12021	YDHHKIIWV	12053	YHTNNMVPW	12085	YNMVHGDES	12117	YYPCVQYRQ
12022	YDINYYPEE	12054	YHTNNREPW	12086	YNMVHGDET		

[2160] 实施例34

[2161] 如通过机器学习鉴定的具有心脏组织嗜性的AAV5变体

[2162] 此实施例描述了如通过机器学习方法鉴定的展现心脏组织嗜性的工程化AAV5变体的优选特性。进行如实施例19和图19至图22中描述的相同生物信息学预处理、机器学习方法和分析以确定展现心脏组织嗜性的工程化AAV5变体的优选特性。

[2163] 下文描述了AAV5 VP衣壳多肽的581至589区域中各位置处的有利生物物理特性和有利氨基酸残基，它们与较高的心脏组织嗜性概率相关。表52中描述的任何以下机器学习衍生的位置氨基酸偏好可以单独存在或彼此组合存在或与实施例13，表25中观察和描述的任何氨基酸偏好组合存在。下文表52中列出了从ML模型中提取的排名靠前的重要心脏嗜性特征之间共有的位置特征的概述。

[2164]

表 52 机器学习衍生的心脏组织嗜性规则
Xaa1 处低溶解性 Xaa1 选自 N 或 E
Xaa1 处低亲疏水性 Xaa1 选自 H、N、Q、P、Y、D 或 E
Xaa1 处高可突变性 Xaa1 选自 A 或 E
Xaa2 处高亲疏水性 Xaa2 选自 V 或 I
Xaa2 处中等可突变性 Xaa2 选自 V
Xaa2 处中等体积 Xaa2 选自 V、E 或 Q
Xaa2 处高溶解性 Xaa2 选自 V 或 M
Xaa3 处低溶解性 Xaa3 选自 R 或 Q
Xaa4 处低表面可接近性 Xaa4 选自 C
Xaa4 处高溶解性 Xaa4 选自 C

[2165]

表 52 机器学习衍生的心脏组织嗜性规则
Xaa4 处低电荷 Xaa4 选自 D、E、Y、W、V、P、M、A、G、F、I、L、N、Q、S、T 或 C
Xaa4 处高亲疏水性 Xaa4 选自 C
Xaa5 处高表面可接近性 Xaa5 选自 D、E、R、K、N 或 Q
Xaa5 处低溶解性 Xaa5 选自 D
Xaa6 处低可突变性 Xaa6 选自 C
Xaa6 处低溶解性 Xaa6 选自 D
Xaa8 处高表面可接近性 Xaa8 选自 D 或 N
Xaa8 处高平均柔性 Xaa8 选自 D、R、P、G 或 S
Xaa9 处中等摩尔质量 Xaa9 选自 N、D、L 或 I

[2166] 下文表53提供了表现出心脏组织嗜性并且符合表52中提供的一种或多种规则的变体AAV VP1衣壳多肽的581至589区域的序列。因此，本公开提供了由具有SEQ ID NO:2的工程化AAV5 VP1衣壳多肽组成的rAAV，其中Xaa1至Xaa9区域是SEQ ID NO:14118-SEQ ID NO:15117中的任一者，如表53中所公开。本文还涵盖由工程化AAV5 VP2衣壳多肽和工程化AAV5 VP3衣壳多肽组成的rAAV，所述多肽在对应于AAV5 VP1 581至589区域中的氨基酸的AAV5VP2(氨基酸残基445至453)和AAV5 VP3(氨基酸残基389-397)中的区域处具有下文表中公开的序列。

[2167] 表53

[2168] 驱动心脏组织嗜性的AAV5 VP1衣壳多肽中581至589区域的序列

[2169]

SEQ ID NO	581-589 序列
14118	AADFFAIND
14119	AADIFAISD
14120	AADIFGIND
14121	AADIFVIND
14122	AADIYAIND
14123	AADMFAIND
14124	AAYIFAIND
14125	ACDCHHWDQ
14126	ACIEWNWGE
14127	AGDCCLAER
14128	AHECCLPNR
14129	AIQPRDMYS
14130	AKECTGLKN
14131	AKGMEKLRT
14132	AKNGDVLEM
14133	ALPCDNWRG
14134	APRCESCNM
14135	AQIDCITGL
14136	ASDYCTTER
14137	ATMNLSWAL
14138	AVCDDHKPN
14139	AVDWHCEVL
14140	AVESKDIPQ
14141	AVGDEQPTD
14142	CCECGHWPN
14143	CDDCYLINR
14144	CDICKHEEY
14145	CDNCYAEPH
14146	CDWMASVSV
14147	CDWMPNVSV
14148	CDWMSNVSV
14149	CEYGWILKR
14150	CGDWDFGEW
14151	CKCCRMEKC
14152	CKFRDFAES
14153	CKMLCRPPV
14154	CKPGDFVRK
14155	CLECCTDSA
14156	CPDCGPTGP
14157	CPECSVDPW
14158	CQMLCRSPM
14159	CRDCCLRGW
14160	CRECSNERV
14161	CRECVHNRC
14162	CRICRATIW
14163	CSDPKNMSI
14164	CWECCNDSG
14165	CWECCPDSG
14166	CWECCTDGG
14167	CWECCTDNG
14168	CWECCTDRV
14169	CWECCTDSV
14170	CWECCTDSW
14171	CWECFTDSG
14172	CWECGTDSG
14173	CWECYTDSG
14174	CWEWCTDSG
14175	CWGCCTDSG
14176	CWVCCTDSG
14177	CYWCGWDNR
14178	DAMDQREQY
14179	DAVDIVKDE
14180	DCNQFNWSE
14181	DDDDFAEKG
14182	DDEILKEDG
14183	DDIVGKPKW
14184	DDVEKDETN
14185	DERERKTKD
14186	DFKALDVPC
14187	DFKVLDVQC
14188	DGCDDHKPN
14189	DGKIEPIRF
14190	DGMRCKEKK
14191	DGVGRPGDN
14192	DHRIDNIDA
14193	DIDDLNFAT
14194	DIDFMECAF
14195	DIEEWEMAE
14196	DIEGQNNGM
14197	DIEHIRDPP
14198	DIEMFNENG
14199	DIEMNITAL
14200	DIGDIWMSQ
14201	DIIEYQKVE
14202	DIIPKDAAP
14203	DIKVLDSQP
14204	DILCICETS
14205	DINNNNCGY
14206	DIQPFDYYL
14207	DIQPNDGCS
14208	DIRDEDNIG
14209	DIRDSWTES
14210	DITEEDVTL
14211	DITHHNDGK
14212	DIVDHDHAG
14213	DIVECIQYN
14214	DIYNDSWST
14215	DLCRPNPKN

[2170]

14216	DLIVATELF	14257	DVCDDHQPN	14298	DVVEQRECP	14339	EIEYIDAGG
14217	DLIVVTEWF	14258	DVCDDHTPN	14299	DVYDDHKPN	14340	EIGCVDSAK
14218	DLPNSQPPH	14259	DVCDDPKPN	14300	DVYEKMWAN	14341	EIGMQIAQE
14219	DLSEPTEWF	14260	DVCPFPDGK	14301	DYFDMNLTD	14342	EIGTFDNSE
14220	DLTVATEWF	14261	DVCPVDVQN	14302	DYGPEDFSM	14343	EIHDQQIVF
14221	DMDCPEDQQ	14262	DVDDIHVSF	14303	EAAPRDSDD	14344	EIHDVDFGM
14222	DMFQLKPQN	14263	DVDEISDGV	14304	EAECKMWPT	14345	EIHEFFHNR
14223	DMVDLDVNP	14264	DVDIHEWPP	14305	EAELGNVNM	14346	EIHEFGASN
14224	DNVNMNDAI	14265	DVDMIDTCP	14306	EAGCRMWPT	14347	EIHEFPASN
14225	DQRCNKAPG	14266	DVDPWDVTE	14307	EARIQCKID	14348	EIHEFRAAN
14226	DRCDSIILA	14267	DVDQMHHND	14308	EARMDIWPI	14349	EIHEFRAPN
14227	DSDAKNMSI	14268	DVDTVEMES	14309	EAVNRPNMK	14350	EIHEFRASI
14228	DSDPKNMSI	14269	DVEEKDNKA	14310	ECGCMYWCY	14351	EIHEFRAST
14229	DSDSKNMSI	14270	DVEETEAAY	14311	ECGCMYWGC	14352	EIHEFRAYN
14230	DSDTKNMSI	14271	DVEHCNTNA	14312	ECGCMYWVY	14353	EIHEFRDSN
14231	DSNDQDTPE	14272	DVEHQDWGG	14313	EDDWQRMLY	14354	EIHEFRGSN
14232	DSSEVDMNF	14273	DVEIHKWGG	14314	EDRIQCKID	14355	EIHEFRSSN
14233	DSYDDSWST	14274	DVETTDNLK	14315	EDRMDKIFT	14356	EIHEFSASN
14234	DSYNDGWST	14275	DVEWDNMKQ	14316	EDTGEREAW	14357	EIHEISTHD
14235	DSYNDSWSP	14276	DVGDDHKPN	14317	EDYAACFEF	14358	EIHGMDTCP
14236	DTAPKQEAE	14277	DVGEDEFQP	14318	EEGMQMISQ	14359	EIHMFEQSE
14237	DTIQKKMDH	14278	DVGEWTMNN	14319	EERCYDGSW	14360	EIHPWHRTN
14238	DTIWADINN	14279	DVGILDFNG	14320	EFHEFRASN	14361	EIIDTSCNC
14239	DVADWNGYD	14280	DVGPMDCQP	14321	EFRIQCKID	14362	EIISMNDFM
14240	DVAERVEQI	14281	DVHEYDHPE	14322	EIAIYNFAN	14363	EILDYCSYN
14241	DVAGLDKNA	14282	DVHFFENSR	14323	EIAPNMCPP	14364	EILVRCEVD
14242	DVAPMDQVE	14283	DVIPRDGAE	14324	EIAPREVVN	14365	EIMPDSIMC
14243	DVAPMHEVE	14284	DVKDQSIML	14325	EICPVDRTQ	14366	EIMPNHNMI
14244	DVAPMHQDE	14285	DVKEIDELC	14326	EICSFEEGH	14367	EINDQRPWP
14245	DVAPMHQIE	14286	DVMPMPEAQ	14327	EIDDDMLEE	14368	EINEFRASN
14246	DVAPMHRVE	14287	DVNEDGFNM	14328	EIDDTDQRW	14369	EINPKSEDE
14247	DVAPMNQVE	14288	DVNENEWTD	14329	EIDDVGEPG	14370	EIPEFRASN
14248	DVCDDHKAN	14289	DVRDSMKRD	14330	EIDEFSASN	14371	EIQDIVKID
14249	DVCDDHKPH	14290	DVSDFIQHF	14331	EIDQKDISQ	14372	EIQPLNAVN
14250	DVCDDHKPI	14291	DVSEGDDST	14332	EIDRGDSYK	14373	EIRCHHKFA
14251	DVCDDHKPS	14292	DVSQIPITF	14333	EIEALMKER	14374	EIRDMAHHG
14252	DVCDDHKPT	14293	DVTEYQWST	14334	EIEEDDEWY	14375	EIRDTHEMQ
14253	DVCDDHKQN	14294	DVTPNSWIE	14335	EIEEFMKMR	14376	EIREFRASN
14254	DVCDDHKRN	14295	DVTPTSWSE	14336	EIEQDNEPG	14377	EIREQCIGP
14255	DVCDDHKSN	14296	DVTTLNESQ	14337	EIEQSFLDS	14378	EIRFEGNME
14256	DVCDDHKTN	14297	DVVDFSIVK	14338	EIEVTAWAC	14379	EIRIQCKID

[2171]

14380	EIRPWDCKW	14421	ETAGKSIHH	14462	EVIPTCHFG	14503	EVRMQCKID
14381	EIRTIDWQW	14422	ETEDYRTSN	14463	EVKPKDEAH	14504	EVRNKCKID
14382	EISDIHDLR	14423	ETEEIQKGH	14464	EVKQWNEVM	14505	EVRNQCKID
14383	EISEYCQPS	14424	ETHEFRASN	14465	EVLIQCKID	14506	EVRSECKID
14384	EISPYNDVE	14425	ETMDKLWAN	14466	EVMHWNIEW	14507	EVRSKCKID
14385	EISQIPPGW	14426	ETMNQSWGL	14467	EVMMFSCYE	14508	EVRSQCKID
14386	EITCLCVHS	14427	ETNEEEDGN	14468	EVMSLESFE	14509	EVRTKCKID
14387	EITDFQYEN	14428	ETQMFEIGE	14469	EVMYRCLMP	14510	EVRTKNEAW
14388	EITDKDVCT	14429	EVACFDDKN	14470	EVPEIEHDK	14511	EVRTQCKID
14389	EITEFWGNY	14430	EVACLGVIN	14471	EVPIQCKID	14512	EVRTQIIMD
14390	EITEITMSK	14431	EVAETDNFF	14472	EVPTRDCCD	14513	EVRVQCKID
14391	EITPKDSHT	14432	EVAHENFRM	14473	EVQCDQIHP	14514	EVRYNPFHW
14392	EIVDFDNYF	14433	EVAHLDATK	14474	EVQDDSKAN	14515	EVSELDGAW
14393	EIVEQCGNR	14434	EVAPECKFE	14475	EVQDDYESN	14516	EVSEVKENG
14394	EIVHQDDVN	14435	EVCDLIHDN	14476	EVQDLDMMD	14517	EVSGTDELI
14395	EIYEFQMWD	14436	EVCEEEFDP	14477	EVQDTDCSP	14518	EVSLWEIGC
14396	EIYEFRASN	14437	EVDDKYHED	14478	EVQEDCMFE	14519	EVTDPNMLD
14397	EKQMDGINH	14438	EVDEGHNNE	14479	EVQIQCKID	14520	EVTDYSGMF
14398	EKYFKMFDN	14439	EVDGRNRGQ	14480	EVRCFHELE	14521	EVTEGCPMG
14399	ELRCTKLQN	14440	EVDPLEIAL	14481	EVREKEPND	14522	EVTPMQDID
14400	ELRIQCKID	14441	EVECRHETC	14482	EVRESDFRE	14523	EVTQIWLFP
14401	ELSDSMDSN	14442	EVEELWNPF	14483	EVRFEMRYE	14524	EVVDCNWLM
14402	EMHECPMAW	14443	EVEEMNWGK	14484	EVRIECKID	14525	EVVEFHWQQ
14403	EMKCIDVED	14444	EVEEQIELI	14485	EVRIHCKLD	14526	EVVVRNMPT
14404	EMRTYNKDY	14445	EVESLPCSL	14486	EVRIPCKID	14527	EVWIQCKID
14405	ENYPWEQLE	14446	EVEWSCRFE	14487	EVRIQCEID	14528	EVYDHDTYI
14406	EPRCENYME	14447	EVGGMGEVK	14488	EVRIQCKFD	14529	EVYHICGAN
14407	EQCRGEPDE	14448	EVGIQCKID	14489	EVRIQCKFG	14530	EWIWCYDRG
14408	EQKDNMMMN	14449	EVGIQCKVD	14490	EVRIQCKIN	14531	EWTMFIEAH
14409	ERRCAWFNA	14450	EVGPLNEWG	14491	EVRIQCKIV	14532	FCDCGHWPN
14410	ESDEREEAD	14451	EVGTHDCLD	14492	EVRIQCKIY	14533	FCECGHWPN
14411	ESDPINEAN	14452	EVHDFHYIP	14493	EVRIQCKLD	14534	FDCHTFGYR
14412	ESGCKMWPT	14453	EVHDKNTQF	14494	EVRIQCKLV	14535	FDECFPEYW
14413	ESHEIRDHN	14454	EVHEARPYF	14495	EVRIQCKMD	14536	FDKMFREIL
14414	ESQEAEWGW	14455	EVHEFRASN	14496	EVRIQCQID	14537	FDRCRNANW
14415	ESRCGDKGC	14456	EVHELENIN	14497	EVRIQCTID	14538	FDRPGRNIL
14416	ESRCVDNTV	14457	EVHMIPEMQ	14498	EVRIQFKID	14539	FEDCGWDKT
14417	ESTDFDELN	14458	EVHTEDYGK	14499	EVRIQWKID	14540	FEIGGKTRT
14418	ESTELRTGN	14459	EVHVSEIMK	14500	EVRIRCKID	14541	FENCRDAGR
14419	ESTNIPMAF	14460	EVIEWCNYE	14501	EVRIRCKMD	14542	FERCEGDNH
14420	ESVPKHEYM	14461	EVIMQCMMK	14502	EVRIYCKID	14543	FHCICAMGH

[2172]

14544	FHSLCAMGH	14585	GRYCECVEH	14626	INYISFIDR	14667	LWQECDEAW
14545	FKFEPWEYW	14586	GSDEFMAYG	14627	INYISFVAR	14668	MDKEIETWP
14546	FKHSLRQWT	14587	GTMNLSWGL	14628	INYISFVDP	14669	MDPQDEMKF
14547	FKHSQREWT	14588	GVCDDHKPN	14629	INYISFVDS	14670	MDRCPPERT
14548	FKHSRREWT	14589	GWECCTDSG	14630	INYISFVNR	14671	MDSCYDRVR
14549	FKKEDAGQF	14590	HARCVDNTV	14631	INYSSFVDR	14672	MDWFKGERS
14550	FKQDDRSEV	14591	HCECGHWPN	14632	ISYISFVDR	14673	MEGSCKQER
14551	FKVLPWWYH	14592	HDCKLKPIM	14633	ITYISFVDR	14674	MEILGTMKQ
14552	FLRCKKWHI	14593	HDRCFASYC	14634	IVDQWNMGI	14675	MEQQPGPKR
14553	FMRPCDMGM	14594	HDRCFDSTA	14635	IWDCMVEKF	14676	MGDCQTLSY
14554	FNRCGRPKN	14595	HGVGRSGDN	14636	KDDCGTQER	14677	MHPRPRPEV
14555	FNVDCGTWF	14596	HHFHVSLDI	14637	KDDFDPDQG	14678	MIDFREDVQ
14556	FPAFRGGSF	14597	HHFQVNLDI	14638	KDEIEDFKS	14679	MIEPYEMHY
14557	FPIDFVPFR	14598	HHFYVNLDI	14639	KDKFEWLTH	14680	MINDMKMFF
14558	FPKASECNW	14599	HHYHVNLDI	14640	KDWRGWEER	14681	MKWINIIDW
14559	FPKECECNW	14600	HKGMEKNPV	14641	KDWWEDDKQ	14682	MMFENHQRN
14560	FPKGSECTW	14601	HKYIDWIDG	14642	KDWYAWEEW	14683	MNRQRDHIR
14561	FPKHDWPTS	14602	HPFHVNLDI	14643	KEGCETDKS	14684	MPKFEDVDH
14562	FPPFCWSGM	14603	HPKCSEQYR	14644	KEMCQWDAQ	14685	MQIECTWSH
14563	FPPWPLDLV	14604	HPRCATDHA	14645	KRDCPWHNL	14686	MQLDCTWHN
14564	FPRFGDDRM	14605	HQRCPDRAD	14646	KSRCADNTV	14687	MRGCMDKTR
14565	FPRGSECNW	14606	HRQCNLLDI	14647	KVIEFDMDI	14688	MSDEWMCNN
14566	FPSCHQCNI	14607	HRSCDRMIM	14648	LAECMEEMC	14689	MTRCVDADY
14567	FQPWPKYWC	14608	HSDPKNMSI	14649	LCDYSWWRY	14690	MVEEQEEND
14568	FQWRSEMKN	14609	HTPDRTEFE	14650	LCPCRLWDW	14691	MVEICCSNK
14569	FRHSLREWT	14610	HTRCVDHDT	14651	LCTDKMDGD	14692	MVMIWDHDQ
14570	FRMENRNAN	14611	HVDDLNDWA	14652	LDECCYDTR	14693	MVNPIDWQC
14571	FRQSPWPVV	14612	HVKCLDYQI	14653	LDEYARPQQ	14694	MVRCVEGDD
14572	FRVTPKPEA	14613	IDICGIEHR	14654	LDRCLVIVQ	14695	MYECGNIGE
14573	FSDPKNMSI	14614	IDRMGMILN	14655	LDWAEHFNC	14696	NCECGHWPN
14574	FWECCTDSG	14615	IDYISFVDR	14656	LEFGGKTRT	14697	NERWDHDQL
14575	FWGCNWDPK	14616	IGIGRSGDN	14657	LEIGAKTRT	14698	NESIDDIEQ
14576	FWQCPWQED	14617	IIAQRDQLN	14658	LEIGGKTKT	14699	NHWNQELDF
14577	GGDWDFGEW	14618	IIGDFSQLG	14659	LERCLGLNQ	14700	NIALRDSMN
14578	GKSADLLKQ	14619	IIKEMDIHW	14660	LESAGKTRT	14701	NIDAWDNHN
14579	GPDCECQRG	14620	IIPDREDIS	14661	LHECRKEIH	14702	NIVDDQYGN
14580	GQPTPWLDH	14621	IITQRDQVN	14662	LKDCHGDGR	14703	NKPFDAFKW
14581	GRDCKYINW	14622	IMDCDTNKY	14663	LKGCPCDTM	14704	NKWHPMIVI
14582	GRGCQIEKY	14623	INDISFVDR	14664	LNCQCWPER	14705	NLDCRPVNG
14583	GRPCCGYNY	14624	INFISFVDR	14665	LNGCDSRMR	14706	NMFPLKPQN
14584	GRPCEPILQ	14625	INYISFDDR	14666	LTDGWDPER	14707	NMFQLKPQH

[2173]

14708	NMFQLKPQT	14749	PIICVWHDN	14790	PVICVWHGN	14831	QCDCSWWRY
14709	NMIQLKPQN	14750	PKGSNMFDW	14791	PVICVWHHN	14832	QCDYSWWRD
14710	NNMLIRPNF	14751	PKKVFCPEE	14792	PVICVWHNN	14833	QCDYSWWRH
14711	NSDPKNMPI	14752	PKQCKVHEQ	14793	PVICVWPDN	14834	QCDYSWWRN
14712	NSDPKNMSI	14753	PKSCKFIDT	14794	PVICVWYDN	14835	QDGTGEPKG
14713	NVNEWEECE	14754	PKYFECIYY	14795	PVIFVWHDN	14836	QFRCVDNKV
14714	NVREYDRYE	14755	PMDCGKHKK	14796	PVIYIWHDN	14837	QIEFIWDAH
14715	NVVEQHEYN	14756	PMFHPRQYP	14797	PVLCVWHDK	14838	QIEMTEISQ
14716	NWEFNARND	14757	PMMISYIQR	14798	PVLCVWHDN	14839	QISEGEVLN
14717	PAEGGNVSD	14758	PMRCGDDAK	14799	PVMCVWHDN	14840	QKNEEEDGN
14718	PAFCQRPEQ	14759	PNCKQWVDN	14800	PVNCVWHDN	14841	QKPCKVEQV
14719	PCAMHNWGF	14760	PNDLKFETF	14801	PVQCNWERE	14842	QKRCCWHRC
14720	PCDMHNWGC	14761	PNICPLDEQ	14802	PVRCTKPRD	14843	QLDSWVEKR
14721	PCGMHNWGC	14762	PNKLSGGEE	14803	PVSCVWHDN	14844	QMEEIQIGH
14722	PCGPIEYDE	14763	PNRCHDDRG	14804	PVTCVWHDN	14845	QNLEIIGAG
14723	PCPCDHGGK	14764	PPVRPWEEM	14805	PVVCVWHDN	14846	QPICRRDKR
14724	PCSQPIHEP	14765	PQFFSWSEY	14806	PWLDDSEME	14847	QPMEWKMVR
14725	PCTMPYVEN	14766	PQHAKCGEF	14807	PWLWHKSAC	14848	QPRCLDNTV
14726	PCTMTYVED	14767	PQRVLPMEY	14808	PWQCKGPIN	14849	QPRCYRNIS
14727	PDCLRPVMK	14768	PREFEWYPH	14809	PWQFCYPWT	14850	QSDEYNIAF
14728	PDHCDPMPE	14769	PRGCDCLQS	14810	PWRVQKHGV	14851	QSRCFDNTV
14729	PDHCVKNER	14770	PRRCDNWKT	14811	PWRWHKSAW	14852	QSRCLDNTL
14730	PDICVWHDN	14771	PSGPIQYDE	14812	PWRWHKSDC	14853	QSRCVDITV
14731	PDIDMDMDY	14772	PSKCEWTDN	14813	PWRWHKSPC	14854	QSRCVDNPG
14732	PDLRPRDAC	14773	PSRCFDNTV	14814	PWRWHKSSC	14855	QSRCVDNRV
14733	PDRFPQIKS	14774	PSRCVDNTV	14815	PWRWHKSTC	14856	QSRCVDNSV
14734	PDTEKEMRF	14775	PTMNLSWGL	14816	PWRWHKSVC	14857	QSRCVDTTV
14735	PDTYVRERY	14776	PTNEEEDGN	14817	PWRWHRSAC	14858	QTHEEDDGN
14736	PEAGDQFKT	14777	PVACVWHDN	14818	PWRWRKSAC	14859	QTHEEEDGN
14737	PEEMKYPLK	14778	PVFCVWHDN	14819	PWRWYKSAC	14860	QTNDKEDGN
14738	PEVFTDIRP	14779	PVGFDPPPF	14820	PWSCQPEQA	14861	QTNEEEAGN
14739	PEWGTSCNW	14780	PVHDFDAAY	14821	PWTYDKEFY	14862	QTNEEEDAN
14740	PFACDGYEM	14781	PVHSHNIDN	14822	PWVFTYPSN	14863	QTNEEEDGH
14741	PGGCRGIRY	14782	PVICDWHDN	14823	PYGPFEYDE	14864	QTNEEEDGT
14742	PGICGEPGK	14783	PVICGWHDN	14824	PYGWRNIDQ	14865	QTNEEEDSN
14743	PGICVWHDN	14784	PVICGWHEN	14825	PYIMEFHAF	14866	QTNEEEDVN
14744	PGKCANERM	14785	PVICLWHDN	14826	PYIMEFHPC	14867	QTNEGEDGN
14745	PGPCVMEIW	14786	PVICVWHAN	14827	PYIMEFHPV	14868	QTNEQEDGN
14746	PHFHVNLDI	14787	PVICVWHDH	14828	PYIMEFHTC	14869	QTSVEEDGN
14747	PHPCMFPTM	14788	PVICVWHDS	14829	PYSMEFHPF	14870	QTTEEEDGN
14748	PICQGMMEN	14789	PVICVWHDT	14830	QANEEEDGN	14871	QTYEEEDGN

[2174]

14872	QVGDDRSFA	14913	SGRCPFSSK	14954	TRPCGTLSH	14995	WDYSMTKKP
14873	QVRIQCKID	14914	SIEECCEEM	14955	TTHEGHEDN	14996	WEFTMAPEV
14874	QWDDDQWST	14915	SMFQLKPQN	14956	TVICVWHDN	14997	WEKCGWPPI
14875	QWPDDQWST	14916	SNDACIQSF	14957	TVPCMWQTY	14998	WELCMELNQ
14876	QYRCSDSAF	14917	SNECWNEMA	14958	TVWEQDLLE	14999	WEPCLELNQ
14877	RDCCDAWSR	14918	SNGDCIQAF	14959	TWAPWCCDC	15000	WERCLELDQ
14878	RDECRMESG	14919	SNRCHDDRG	14960	TWAPWCCYY	15001	WERCLELHQ
14879	RDKCRMESG	14920	SNVDCIQSF	14961	TWGPWCCYF	15002	WERCLELNL
14880	RDPCRMESG	14921	SPERMDYDG	14962	TWPPWCCYF	15003	WERCLELNP
14881	RDPLNDSRY	14922	SPKGSECNW	14963	TWRCGDIGH	15004	WERCLEQNH
14882	RDQCRMEAG	14923	SQQWCEQPD	14964	TYRCHDDRG	15005	WERCLEVNQ
14883	RDQCRMECG	14924	SRPCCKDIR	14965	VADCKQPRF	15006	WERCLGLNQ
14884	RDQCRMEDG	14925	SSDPKNMSI	14966	VDCSKILRR	15007	WERCLGLTQ
14885	RDQCRMEFG	14926	STMDLSWGL	14967	VDDRPKPEH	15008	WERCLQLNQ
14886	RDQCRMEHG	14927	STMNLSWGF	14968	VDRLFRMEF	15009	WERCLVLNQ
14887	RDQCRMESA	14928	SVDYEDEVW	14969	VDRQVRMEY	15010	WERCMELNQ
14888	RDQCRMESD	14929	SVEPKSMTG	14970	VDSHKIIRR	15011	WERCQELNH
14889	RDQCRMESR	14930	SVETLDQSL	14971	VEIRGKTRT	15012	WGAFRGVCE
14890	RDQCRMESS	14931	SVICVWHDN	14972	VHEGKTFDI	15013	WHGFHKIMF
14891	RDQCRMESV	14932	SVLRRDEVN	14973	VHQDKIIGE	15014	WIAPRHMGL
14892	RDQCRMETC	14933	SVMDIDCHQ	14974	VICEFEEGN	15015	WKAFRGGIG
14893	RDQCRMEYG	14934	SWLDNILPK	14975	VITEDDQAE	15016	WKDCKQMKM
14894	RDQCRMKSG	14935	SYFWCFMDC	14976	VNEWEADAE	15017	WKFGDKLIM
14895	RDQCRVESG	14936	TADIFAIND	14977	VQLCDLLRR	15018	WKGAPALER
14896	RDSHDEFEW	14937	TAEIFAVND	14978	VRYSDMFWW	15019	WKHCEREIN
14897	REDCWPPEN	14938	TDKYCDQKR	14979	WAEHDFPKG	15020	WKIDPCVLQ
14898	REHIGFMKF	14939	TFSEIRDQC	14980	WCDDDMRRQ	15021	WKLGDGIQD
14899	REIGGKTRT	14940	THPWEMTKE	14981	WCTEEWEEK	15022	WKRPPGGEH
14900	REKCHQHQC	14941	THRCHDDRG	14982	WDCSMTKKH	15023	WKSHEWDPH
14901	RSRCVDNTV	14942	TIDSWFQER	14983	WDFSGRIDN	15024	WLECNPPEQ
14902	RWGCDLPQI	14943	TIGEINDAN	14984	WDHREERTH	15025	WLHPYSYCA
14903	RYQCRMESG	14944	TKFHDGDWC	14985	WDICSGLRQ	15026	WNLQDKPGE
14904	SADIFAIND	14945	TKMLDRDAE	14986	WDKCIPMHW	15027	WNPARKDDI
14905	SCECGHWPI	14946	TNRCHDDRC	14987	WDLVDPPKY	15028	WNSCPWQLP
14906	SCECGHWPN	14947	TNRCHDDRD	14988	WDMNGKDGS	15029	WPCHDAPIW
14907	SDPGEILKR	14948	TNRCHDDRV	14989	WDPRRPVKK	15030	WPIGEKMVF
14908	SGDGDFGEW	14949	TNRCHDHRG	14990	WDRCLELNK	15031	WPKERVPKG
14909	SGDLDFGEW	14950	TNRCHDYRG	14991	WDSLEWDRA	15032	WPNDDMTRM
14910	SGDWAFGEW	14951	TPPWESIRK	14992	WDYSMPKKH	15033	WQFTGKPTL
14911	SGDWDFGKW	14952	TQPFSKDED	14993	WDYSMSKKH	15034	WQHIMKPQI
14912	SGDWNFGEW	14953	TREHQIISP	14994	WDYSMTKKN	15035	WQMLCRPPV

[2175]

15036	WQRCLELNQ	15057	YCECGHWPD	15078	YIRPEQPVC	15099	YSDPKNMPI
15037	WRQFPCAEQ	15058	YCECGHWPH	15079	YKDRDICRL	15100	YSDPKNMSV
15038	WRSCPKGLV	15059	YCECGHWPI	15080	YKEHEAVQF	15101	YSDPKNMTL
15039	WSRCVDNTV	15060	YCECGHWPY	15081	YKFYPGLPT	15102	YSDPKNMTT
15040	WTIEPEWYF	15061	YCECGHWQN	15082	YKGMEKNAG	15103	YSDPKNMYI
15041	WTKDDEQLE	15062	YCECGHWSN	15083	YKGMEKNRV	15104	YSDPKTMSI
15042	WVHPFSYCA	15063	YCECGHWTS	15084	YKGMEKTPV	15105	YSDTKKMSI
15043	WVHPYTYCA	15064	YCECGNWPN	15085	YKMCDHDDC	15106	YSGPKNMSI
15044	WVHQYSYCA	15065	YCECGRWPI	15086	YKPCMTIDV	15107	YSHPKHMSI
15045	WVHTYSYCA	15066	YCECGRWPN	15087	YKQHEGEQF	15108	YSYPKNMTI
15046	WVLDISHCG	15067	YCECGYWPN	15088	YPHCRLLGP	15109	YVDPYEMYG
15047	WVRCLELKQ	15068	YCEGGHWPN	15089	YPHMDCSEE	15110	YVFGQKDEM
15048	WVYPYSYCA	15069	YCGCGIHQQ	15090	YRDCEQMEV	15111	YVIEQKDEM
15049	WWRVPCESH	15070	YDESVRISR	15091	YRDCRQFNR	15112	YVIGQKDEV
15050	WYDWIHERT	15071	YEEFPEIIM	15092	YRVLPRIKD	15113	YVLDSWRTS
15051	WYICQKDSM	15072	YEETRVEAR	15093	YSDAKNMSI	15114	YWECGHWPN
15052	WYKCQDVDR	15073	YELYQDIDN	15094	YSDGKNMSI	15115	YWGCPDQVR
15053	YADPKNMSI	15074	YERCHHQGD	15095	YSDHQNMSI	15116	YYNHKIFEP
15054	YCECEHWPN	15075	YGPCCHFQF	15096	YSDPKDMSI	15117	YYNPKIVEP
15055	YCECGHCPI	15076	YHKFGDMDI	15097	YSDPKNISM		
15056	YCECGHSPN	15077	YHRCDEFFL	15098	YSDPKNMFI		

[2176] 实施例35

[2177] 如通过机器学习鉴定的具有淋巴结组织嗜性的AAV5变体

[2178] 此实施例描述了如通过机器学习方法鉴定的展现淋巴结组织嗜性的工程化AAV5变体的优选特性。进行如实施例19和图19至图22中描述的相同生物信息学预处理、机器学习方法和分析以确定展现淋巴结组织嗜性的工程化AAV5变体的优选特性。

[2179] 下文描述了AAV5 VP衣壳多肽的581至589区域中各位置处的有利生物物理特性和有利氨基酸残基，它们与较高的淋巴结组织嗜性概率相关。表54中描述的任何以下机器学习衍生的位置氨基酸偏好可以单独存在或彼此组合存在或与实施例16，表31中观察和描述的任何氨基酸偏好组合存在。下文表54中列出了从ML模型中提取的排名靠前的重要淋巴结嗜性特征之间共有的位置特征的概述。

[2180]

表 54 机器学习衍生的淋巴结组织嗜性规则
Xaa1 处高平均柔性 Xaa1 选自 D、E、P、G、Q、S 或 R
Xaa1 处高氢键供体 Xaa1 选自 R
Xaa1 处高摩尔质量 Xaa1 选自 Y、W、R 或 F
Xaa2 处低溶解性 Xaa2 选自 N 或 E
Xaa3 处低平均柔性 Xaa3 选自 W、M 或 F
Xaa3 处低可突变性 Xaa3 选自 R、H、K、P、Y、F、L 或 C
Xaa4 处低可突变性 Xaa4 选自 C
Xaa5 处高可突变性 Xaa5 选自 N
Xaa5 处中等摩尔质量 Xaa5 选自 D、I、L 或 N
Xaa6 处高摩尔质量 Xaa6 选自 Y、W、R 或 F
Xaa6 处高平均柔性 Xaa6 选自 G 或 R
Xaa7 处高平均柔性 Xaa7 选自 D、E、K、P、I、N、Q 或 S
Xaa7 处低溶解性 Xaa7 选自 N 或 E

[2181]

表 54 机器学习衍生的淋巴结组织嗜性规则
Xaa8 处低溶解性 Xaa8 选自 N、E 或 D
Xaa8 处中等可突变性 Xaa8 选自 R 或 H
Xaa9 处低可突变性 Xaa9 选自 P 或 K
Xaa9 处高平均柔性 Xaa9 选自 D、E、P 或 S
Xaa9 处高溶解性 Xaa9 选自 M 或 V

[2182] 下文表55提供了表现出淋巴结组织嗜性并且符合表54中提供的一种或多种规则的变体AAV VP1衣壳多肽的581至589区域的序列。因此,本公开提供了由具有SEQ ID NO:2的工程化AAV5 VP1衣壳多肽组成的rAAV,其中Xaa1至Xaa9区域是SEQ ID NO:20118-SEQ ID NO:21117中的任一者,如表55中所公开。本文还涵盖由工程化AAV5 VP2衣壳多肽和工程化AAV5 VP3衣壳多肽组成的rAAV,所述多肽在对应于AAV5 VP1 581至589区域中的氨基酸的AAV5VP2(氨基酸残基445至453)和AAV5 VP3(氨基酸残基389-397)中的区域处具有下文表中公开的序列。

[2183] 表55

[2184] 驱动淋巴结组织嗜性的AAV5 VP1衣壳多肽中581至589区域的序列

[2185]

SEQ ID NO	581-589 序列
20118	ACKLDHMPP
20119	ADCCSCKAN
20120	ADDCTHKNH
20121	ADSCTHKNH
20122	ADYCAHKNH
20123	ADYCPHKNH
20124	ADYCTHKTH
20125	ADYCTHRNH
20126	ADYGTHKNH
20127	ADYWTHKNH
20128	AECCSIFSS
20129	AERLCIEWD
20130	AESKCDIGT
20131	AESRCDVGT
20132	AEWSWAEHS
20133	AFWGTDVWG
20134	AFWSPDVWG
20135	AGFRIDFRC
20136	AGGWLFFWH
20137	AGNNEDTQL
20138	AGQWGSLNP
20139	AGSNEDKQL
20140	AGTDQTMHM
20141	AHRWSWSES
20142	AISQDMFMQ
20143	AISQDMLMQ
20144	AMLCNRVCE
20145	ANCCSCKAD
20146	ANCCSCKDN
20147	ANCCSCKPN
20148	ANCCSCKTN
20149	ANCCSCKTS
20150	ANCCSGKAN
20151	ANCYSCKAN
20152	ASASHDFMG
20153	ASCEDEGIL
20154	ASGSNDFMG
20155	ATDRVVVNF
20156	ATSRCDVQT
20157	AVPWEWQEG
20158	CCAALIQSG
20159	CCDALIQSS
20160	CCDDLIQSG
20161	CCDPLIQSG
20162	CDMWWYQDP
20163	CEKYNLPKI
20164	CKDRPIIDC
20165	CLREAVVGI
20166	CVLNMEYCL
20167	CVQRLLFKC
20168	CVVNMEYYL
20169	DAAIKMDES
20170	DCIFFHCPK
20171	DDFGPNHRH
20172	DDHQPMQSR
20173	DDRGFTRSM
20174	DDWSWAEHS
20175	DECIMQRSK
20176	DEFDPNHRH
20177	DEFGPNHGH
20178	DEFGPNHIH
20179	DEFGPNHKH
20180	DEFGPNHRR
20181	DEFGPNHRY
20182	DEFGPNYRH
20183	DEFGPSHRH
20184	DEFGTNHRH
20185	DEIGPNHRH
20186	DEMGYDQLD
20187	DERCYDQLD
20188	DEWAWAEHS
20189	DEWPWAEHS
20190	DEWSGAEHS
20191	DEWSLAEHS
20192	DEWSWAAHS
20193	DEWSWAEHA
20194	DEWSWAERS
20195	DEWSWAGHS
20196	DEWSWAQHS
20197	DEWSWSEHS
20198	DEWSWTEHS
20199	DEWSWVEHS
20200	DEWTWAEHS
20201	DEWWSGHAH
20202	DEWWSGHAP
20203	DEYGPNHRH
20204	DFHSRTNIN
20205	DFLDKSTFF
20206	DFLHGTSIM
20207	DFLHVTSTM
20208	DFLPVTSIM
20209	DFLYVTSIM
20210	DGEWGSLNP
20211	DGFRIEFRC
20212	DGLKNDDRK
20213	DGNDQTMHM
20214	DGPDQTMHM
20215	DGQWGSLNL
20216	DGQWGSLNQ
20217	DGQWGSLNS
20218	DGTDPTMHM
20219	DGTDQTMHV
20220	DGTGQTMHM
20221	DGTNQTMHM
20222	DHQGPKLSK
20223	DMYMQQIWR
20224	DNNSPQEAS
20225	DNRCYKYWN
20226	DPICKVEDW
20227	DPIWLIKEC
20228	DPIWRVEDW
20229	DPIWTVEDW
20230	DPKWANCDC
20231	DPMWEREMF
20232	DPSWKVEDW
20233	DPVWKVEDW
20234	DQDAGGYDS
20235	DQWSWAEHS
20236	DRQWGSLNP
20237	DSKNQNCFK
20238	DSLHVTSIM
20239	DSNYQGTKH
20240	DSPNQNCFK
20241	DSQNENCFK
20242	DSQNQKCFK
20243	DSQNQNRFK
20244	DSRNQNCFK
20245	DTFGLRFDD
20246	DTFGLRIDD
20247	DTFGQRVDD
20248	DTFWLRVDD
20249	DTQNQNCFK
20250	DTRCDKYWN
20251	DTRCSKYWN
20252	DTRCYKYWT
20253	DVQWGSLNP
20254	DWTDQTMHM
20255	EAHTNSIGK
20256	EANWVYIEA
20257	EARTNSICK
20258	ECEQDDTDI
20259	ECGLWNFSG
20260	ECKITESLT
20261	EDHLHGPGD
20262	EDHWLGPGD
20263	EDHWRGPGD
20264	EDRWHGPGD
20265	EDYWHGPGD
20266	EEAHSVKFV

[2186]

20267	EEANSVKFA	20308	EHHPVGVMG	20349	ESKYVFPNQ	20390	FFIFPEDED
20268	EEARCVNPC	20309	EHHTLGVMG	20350	ESKYVIPNL	20391	FFIFPEHED
20269	EEAYSVKFA	20310	EHLNLGYHT	20351	ESKYVIPYQ	20392	FFIFPEIED
20270	EEDHSVKFA	20311	EHRNLGYHT	20352	ESKYVISNQ	20393	FFIFPENAD
20271	EEFARIPGN	20312	EHVRGMEER	20353	ESKYVVPNQ	20394	FFIFPENEA
20272	EEFDRIPGH	20313	EHVRGMVDR	20354	ESMERWGYD	20395	FFIFPENEV
20273	EEFSRIPGH	20314	EHYTQDISV	20355	ESTNDNITI	20396	FFIFPENGD
20274	EEPETYTQM	20315	EIAHKLDDI	20356	ESTTDNIAI	20397	FFIFPENQD
20275	EEPRCGNPC	20316	EIALKLDDL	20357	ETCGPGGYT	20398	FFIFPENVD
20276	EEPRCVDPC	20317	EINEDYAKI	20358	ETHYNEFCK	20399	FFIFPETED
20277	EEPRCVNPG	20318	EKDQRGIGS	20359	ETRCYKYWN	20400	FFIFPEYED
20278	EEPRCVNPW	20319	EKRCHAIAL	20360	ETVETHRDR	20401	FFIFRENED
20279	EEPRCVNPY	20320	ELAGKAFDH	20361	EVGGVRIGL	20402	FFIFSENED
20280	EEPRCVNTC	20321	ELYVILIWW	20362	EVKRWMMAD	20403	FFIFTENED
20281	EEPRCVSPC	20322	EMPIEWDEQ	20363	EVKRWMMDD	20404	FFIIPENED
20282	EEPRCVTPC	20323	EMRFNNFCY	20364	EVKRWMMGD	20405	FFIVPENED
20283	EFCENLTAI	20324	ENDTQDISV	20365	EVKRWMMVG	20406	FFIYPENED
20284	EFCENMTSI	20325	ENTMSIIEH	20366	EVKRWMTVD	20407	FFLFPENED
20285	EFFCEADKA	20326	ENTMTIIEH	20367	EVPKNELNV	20408	FFNINGSVE
20286	EFFCPADKA	20327	EPIKIFNDA	20368	EVQRWMMVD	20409	FFVFPENED
20287	EFFCQADKT	20328	EPIQIFDDA	20369	EVRRWMMVD	20410	FGLKNDDRK
20288	EFFCQADKV	20329	EPIQIFNDS	20370	EWAQPMIMD	20411	FHALWAECF
20289	EFFCQANKA	20330	EPIQIFSDA	20371	EWIQPMIMD	20412	FIDAKIWDV
20290	EFFCQVDKA	20331	EPIQIVNDA	20372	EWTQPLIMD	20413	FIIFPENED
20291	EFFFQADKA	20332	EPIQMFNDA	20373	EWTQPMIMG	20414	FIYKCLFGF
20292	EFQQDDTDI	20333	EPIQSFNDA	20374	FATPKEGFD	20415	FLVFMIIEG
20293	EGIKFQTEL	20334	EPIWKVEDW	20375	FCDGCAPED	20416	FLVFMLIEV
20294	EGKYMDYFM	20335	EPLQIFNDA	20376	FCDRCAAED	20417	FNGNDMLMV
20295	EGNSNWVFQ	20336	EPNIHAKCV	20377	FCDRCAPEG	20418	FPEHKDQGI
20296	EGPQNELNV	20337	EPPLSPSAS	20378	FCDRCAPEV	20419	FQEHKDKGI
20297	EGVGLYEFA	20338	EPVFGFWNY	20379	FCDRCAPEY	20420	FQEHKDQGV
20298	EHDNQDISV	20339	EQHNLGYHT	20380	FCDRCAPKD	20421	FTNAAPVQD
20299	EHDTQAISV	20340	EQIASYDRV	20381	FCDRCATED	20422	FWESDRVMM
20300	EHDTQDIIV	20341	ERDTQDISV	20382	FCDRCDPED	20423	FWKSDRVMV
20301	EHDTQDSSV	20342	ESCEDEGLL	20383	FCDRCTPED	20424	FWKTDRVMM
20302	EHHALGVMG	20343	ESCEDEGML	20384	FCGRCAPED	20425	FYGPNWEFH
20303	EHHDCVPWS	20344	ESCEDEGTL	20385	FFFFPENED	20426	FYIFPENED
20304	EHHHSFGCP	20345	ESCEDEGVL	20386	FFICPENED	20427	FYLPNWEFH
20305	EHHNLGYHK	20346	ESCEDGGIL	20387	FFIFAENED	20428	FYQQFFEGI
20306	EHHNVGYHT	20347	ESCEDQGIL	20388	FFIFHENED	20429	FYREFFEGI
20307	EHHPLGAMG	20348	ESCQDEGIL	20389	FFIFLENED	20430	FYRKFFEGI

[2187]

20431	FYRPFFEGI	20472	GEFWTWSRT	20513	HFCANRHIA	20554	IPYKNQNNF
20432	FYRPNWECH	20473	GEIPDRVAC	20514	HFCDNRHIA	20555	ISEWNFQDW
20433	FYRPNWEFD	20474	GEIPHRVSC	20515	HFCFNRHIA	20556	IVCQNNANS
20434	FYRPNWEFL	20475	GEIPNRVAC	20516	HFCLNRHIA	20557	IYRPNWEFH
20435	FYRPNWEFY	20476	GEIPPRVAC	20517	HFCVNRHIE	20558	KDHNREHII
20436	FYRPNWEIH	20477	GEPVDRNTC	20518	HFCVNRHNA	20559	KDNECVSHT
20437	FYRPNWEVH	20478	GEWSWAEHS	20519	HFCVNRNIA	20560	KDPETYTQM
20438	FYRQFFAGI	20479	GFLHVTSIM	20520	HFCVNRPIA	20561	KEPEAYTQM
20439	FYRQFFEGF	20480	GFPDEGSLQ	20521	HFFVNRHIA	20562	KEPENYTQM
20440	FYRQFFEGL	20481	GFPDEKSLQ	20522	HFGVNRHIA	20563	KEPETYIQM
20441	FYRQFFEGV	20482	GFPDERNLQ	20523	HGLDLGCAV	20564	KEPETYTKM
20442	FYRQFFQGI	20483	GFPDERSLE	20524	HGLKNDDRK	20565	KEPGTYTQM
20443	FYRQYFEGI	20484	GFPDERSLK	20525	HGTDQTMHM	20566	KEPKTYTQM
20444	GARKNGFMQ	20485	GFTDERSLQ	20526	HIRLWPFEG	20567	KERETYTQM
20445	GARMNGFMR	20486	GLELGHVWG	20527	HPMDDACKQ	20568	KESETYTQM
20446	GARMNGFTQ	20487	GPFHQDARA	20528	HSANRDHFV	20569	KETETYTQM
20447	GARTNGFMQ	20488	GPFKQDARA	20529	HSEDRDHFV	20570	KGIDMAMEL
20448	GCQQDDTDI	20489	GPFLQDARA	20530	HSEHRDHFV	20571	KILTDIFDS
20449	GDCACKFSA	20490	GPFQEDARA	20531	HSENRDHFG	20572	KIWNFTKMM
20450	GDCECKFAA	20491	GPFQQDARS	20532	HSENRDHIV	20573	KIWNVTKVM
20451	GDCECKFSS	20492	GPFQQDDRA	20533	HSENRDPFV	20574	KIWSVTKMM
20452	GDCECTFSA	20493	GPFQQDGRA	20534	HSENRDYFV	20575	KKPETYTQM
20453	GDCGCKFSA	20494	GPVQQDARA	20535	HSGNRDHFV	20576	KMQCLTFCD
20454	GDHLGNVWY	20495	GSISNVKAN	20536	HSKNRDHFV	20577	KNIDYLKWN
20455	GDHWHGPGD	20496	GSQCGVPND	20537	HTHYNEFCK	20578	KNNEYLKWN
20456	GDIPHRVAC	20497	GTFSNVKAN	20538	HTKYECYRF	20579	KPHHWNKKT
20457	GDMWWYQDP	20498	GTISNVKDN	20539	HTSLLHFDH	20580	KPPHWNKET
20458	GDPADRNTC	20499	GTISNVKPN	20540	HVCVNRHIA	20581	KPPHWNNKT
20459	GDPGDRNTC	20500	GTISNVKSN	20541	HWPTHWNIP	20582	KQPETYTQM
20460	GDPLDRNTC	20501	GTISNVKTN	20542	HWWFRKRWV	20583	KRNEDVLEQ
20461	GDPVDHNTC	20502	GTISNVNAN	20543	HWWYRKRWE	20584	KSMERWAYD
20462	GDPVDRNAC	20503	GTISNVRAN	20544	HYTKSISHS	20585	KSMGRWGYD
20463	GDRLDREVF	20504	GTKPDTFHG	20545	IDMWWYQDP	20586	KTHYNEFCK
20464	GDYCTHKNH	20505	GTVSNVKAN	20546	IFDTRWSNV	20587	KTIWGWSDC
20465	GDYECKFSA	20506	GVDMCGLAE	20547	IFNYSNHND	20588	KTWNVTKMM
20466	GDYLGNIWY	20507	GVISCPDCI	20548	IFPYSNHND	20589	KYISEWMEL
20467	GEAQQRVVK	20508	GVISCPECL	20549	IINCNGMNW	20590	LAMDCRPEH
20468	GECECKFSA	20509	GYAGVESLG	20550	IINRTGMNW	20591	LDDPDTSGQ
20469	GEFGPNHRH	20510	GYAGVGILG	20551	IIRLWPFEG	20592	LFCVNRHIA
20470	GEFMEQGCG	20511	HANQDPWDG	20552	ILGQWISYV	20593	LGICCYRDP
20471	GEFWNWSRP	20512	HATLDPWDG	20553	IPDWLWKHV	20594	LGMDWGMTD

[2188]

20595	LGSLNGAMA	20636	LSEWNFHDW	20677	MHSPLIQYG	20718	NRILDVEHD
20596	LGSSNGAMD	20637	LSICCYRDP	20678	MIEDRDYFI	20719	NSENRDHFV
20597	LGSSNGAVA	20638	LSMDCRPEH	20679	MIETKKRVI	20720	NTAEGLGWV
20598	LHCALGQHS	20639	LTMDCGPEH	20680	MISPLIQYG	20721	NTIEGLGWV
20599	LHCELGEHS	20640	LTMDCRAEH	20681	MISWWKVDF	20722	NTTEGLGWG
20600	LHCELGQHP	20641	LTMDCRPED	20682	MKDAGEWLV	20723	NTTQGLGWV
20601	LHCGLGQHS	20642	LTMDCRPER	20683	MKFIQMIWV	20724	NVMEKYECR
20602	LHCKLGQHS	20643	LTMDCRPEY	20684	MLSALIQYG	20725	NVMEKYECS
20603	LHCVLGQHS	20644	LTMDCRPGH	20685	MLSHLIQYG	20726	NWSAPQFSD
20604	LHGKVSMHR	20645	LTMDCRPKH	20686	MLSPLIQYC	20727	NWSGPQFTD
20605	LHWELGQHS	20646	LTMDCRPKR	20687	MLSTLIQYG	20728	NWSTPQFTD
20606	LIFPNGDFQ	20647	LTMDGRPEH	20688	MNFIQMILV	20729	PECCGIFSS
20607	LIPLWPFEG	20648	LTMGCRPEH	20689	MNFIQMLWV	20730	PECCSICSS
20608	LIQLWPFEG	20649	LTMPNSEGK	20690	MNFISVHMC	20731	PECCSIFAS
20609	LIRLWQFEG	20650	LTMPNSEQK	20691	MNICVLDGP	20732	PECCSIFCS
20610	LIRLWRFEG	20651	LTPDWIGHY	20692	MNICVLDGT	20733	PECCSIFSA
20611	LKMDCRPEH	20652	LTPHWIGHS	20693	MNSAQTMVY	20734	PECCSIFST
20612	LLICCYRDP	20653	LTPHWIVHY	20694	MNSTQTMAY	20735	PECCSIFYS
20613	LLINDWISP	20654	LWFCCYRDP	20695	MNVDDLPIN	20736	PECCSIVSS
20614	LLVYFHVPN	20655	LWICCYQDP	20696	MSHCLYVFG	20737	PECCSIYSS
20615	LNVADLPIN	20656	LWICCYRDA	20697	MTCNDMAFY	20738	PECCSLFSS
20616	LNVDDMPIN	20657	LWICFYRDP	20698	MTIDSMAPC	20739	PECCSSFSS
20617	LNVHDLPIN	20658	LWMCCYRDP	20699	MTKWDRLHD	20740	PECCSTFSS
20618	LNVVDLPIN	20659	LWNCCYRDP	20700	MTKWDRWRD	20741	PECCSVFSS
20619	LPADLVDDW	20660	LWVCCYRDP	20701	MWICCYRDP	20742	PECCTIFSS
20620	LPADLVQDW	20661	MAGINGMAY	20702	MWYLWWNCL	20743	PECWSIFSS
20621	LPADLVYDW	20662	MAGLNGMAN	20703	MYLDCCDFK	20744	PECYSIFSS
20622	LPDDLVHDW	20663	MAGLNGMAS	20704	MYLDCRNFK	20745	PEELGYCYV
20623	LPPHWIGHY	20664	MAGLNGMEY	20705	MYMDCCNFK	20746	PEELLYCYV
20624	LPSDLVHDW	20665	MAGLNGVAY	20706	MYVDCCNFK	20747	PEELWYCHV
20625	LPVDLVHDW	20666	MAGVNGMAY	20707	NATEGLGWV	20748	PEEQWYCYV
20626	LQMANRDPF	20667	MAHNGNWNR	20708	NCVLFECCG	20749	PEFCSIFSS
20627	LQMANRDSC	20668	MDMSKAWMI	20709	NCVLYEFCG	20750	PEKGMVEHQ
20628	LQMANRDTF	20669	MELGKDWIQ	20710	NCVVYECCG	20751	PERENLERC
20629	LQMDNRDSF	20670	MFENWNLGV	20711	NEWSWAEHS	20752	PEWCSIFSS
20630	LQMPNRDSF	20671	MFETKKRVV	20712	NGARADAKK	20753	PFAHMESDA
20631	LQMSNRDSF	20672	MGGLNGMAY	20713	NGLDLGCGV	20754	PFEHMASDA
20632	LQTHSICRD	20673	MGIDKAMEL	20714	NGTDQTMHM	20755	PFEHMDSDA
20633	LQTKSICRD	20674	MGIDMAMEF	20715	NGTRVDAKK	20756	PHQWSWSES
20634	LRCELGQHS	20675	MHCELGQHS	20716	NKVNWNHSS	20757	PHRWSWIES
20635	LRMDCRPEH	20676	MHDDNLNWN	20717	NNGFDRYID	20758	PHRWSWSEP

[2189]

20759	PILWTNHSN	20800	QSDKNWYGM	20841	REFWTWSRP	20882	RPHVCHQCD
20760	PIPWTNHSN	20801	QSDQNWYSM	20842	REPETYTQM	20883	RPWYHENDG
20761	PIQWTNHIN	20802	QSKRCLFDH	20843	RERPQQGQQ	20884	RPWYHEYGG
20762	PIQWTNHNN	20803	QTDYNEFCK	20844	RFCVNRHIA	20885	RRPNAHVDC
20763	PIQWTNPSN	20804	QTHNNEFCK	20845	RFPDERSLQ	20886	RSDVCDNYD
20764	PISWCTQDL	20805	QTHYNEFCE	20846	RGGPQQGQQ	20887	RSHVCHQCD
20765	PISWCTQHV	20806	QTHYNEFCR	20847	RGHVCHQCD	20888	RSMERWGYD
20766	PLINQCREL	20807	QTIGGWSDC	20848	RGRPEQGQQ	20889	RTDAQIGWQ
20767	PMLCNRGCE	20808	QTIWGWGDC	20849	RGRPQEGQQ	20890	RTDELIGWQ
20768	PMLCNRLCE	20809	QTIWGWIDC	20850	RGRPQQGEQ	20891	RTDEQIGWR
20769	PMLCNRVGE	20810	QTIWGWNDC	20851	RGRPQQGKQ	20892	RTDQQIGWQ
20770	PMLWNRVCE	20811	QTIWGWSDF	20852	RGRPQQGQR	20893	RVGCLNVWG
20771	PMMCNRVCE	20812	QTIWVWSDC	20853	RGRSQQGQQ	20894	RVHVCHQCD
20772	PPLGAWWYG	20813	QTNYNEFCK	20854	RHFLAWSEI	20895	RVMMDWQPD
20773	PPPNTFHDS	20814	QVCSDWCEG	20855	RHIDFFFGG	20896	RWICCYRDP
20774	PQCCSIFSS	20815	RAHECHQCD	20856	RHNDYETHN	20897	RYADMMSEY
20775	PRNNDDWRT	20816	RAHVCDQCD	20857	RHRWSWSES	20898	RYCLAWSEI
20776	PRQLWEVWQ	20817	RAHVCHHCD	20858	RHSDYETHN	20899	RYFFAWSEI
20777	PSRDCDSWE	20818	RAHVCHPCD	20859	RHTDYETHT	20900	RYFLAWIEI
20778	PTIWGWSDC	20819	RAHVCHQCA	20860	RHWWGAACV	20901	RYFLAWSAI
20779	PTMDCRPEH	20820	RAHVCNQCD	20861	RIDHNGGWF	20902	RYFLAWSES
20780	PTRCLPLWT	20821	RAHVCQQCD	20862	RIGHNGGGF	20903	RYFLAWTEI
20781	PVEWAYWHM	20822	RARPHQGQQ	20863	RIRLWPFEG	20904	RYFLDWSEI
20782	QAMRCLFDH	20823	RDQLDREVF	20864	RIWNVTKMM	20905	RYFSAWSEI
20783	QARENLERC	20824	RDRLDRAVF	20865	RKPVEWHWE	20906	RYFWAWSEI
20784	QECCSIFSS	20825	RDRLDREAF	20866	RLDLFIGRT	20907	RYILAWSEI
20785	QEGENLERC	20826	RDRLDREGF	20867	RLGLFMGRT	20908	RYVLAWSEI
20786	QEPETYTQM	20827	RDRLDRELF	20868	RLINDWIAP	20909	SAPRLIEDR
20787	QEPHPFSEQ	20828	RDRLDREVC	20869	RLINDWIFP	20910	SCDRCAPED
20788	QERENFERC	20829	RDRLDREVS	20870	RLINDWIST	20911	SCHTDDAKP
20789	QERENLDRC	20830	RDRLDREVV	20871	RLNCLENHS	20912	SCKFDHMPP
20790	QERENVERC	20831	RDRLDREVY	20872	RLNNDWISP	20913	SCKLAHMPP
20791	QIGCCMLMW	20832	RDRLDRQVF	20873	RLNSLDNHS	20914	SCKLDHMAP
20792	QIIEFEERV	20833	RDRLDWEVF	20874	RLSNDWISP	20915	SCKLDHMSP
20793	QITCCMLMW	20834	RDRLGREGF	20875	RLVNDWISP	20916	SCKWDHMPP
20794	QIWNVTKMM	20835	RDRLGREVF	20876	RMGSKEQED	20917	SCNPEDSIV
20795	QLTCPLLDD	20836	RDRMDREVF	20877	RMQCLTYCD	20918	SCNPEGSIG
20796	QMLCNRVCE	20837	RDRRDREVF	20878	RNPISVCSV	20919	SDDGDHCTI
20797	QMPMPPVAD	20838	RDRSYEQMD	20879	RNPVDWHWE	20920	SDDGYDIFG
20798	QNHYNEFCK	20839	RDRVDREVF	20880	RNPVEWHWK	20921	SDHGMHRDG
20799	QNIEYLKWN	20840	RECIPMNSA	20881	RNPVEWNWE	20922	SDMCWYQDP

[2190]

20923	SDMGWYQDP	20964	TCVELPERI	21005	VLGQWIGYV	21046	YCLKNDDRK
20924	SDMSWYQDP	20965	TDHNREHIS	21006	VLGRWISYV	21047	YCNEFDEIV
20925	SDMWCYQDP	20966	TDHNREHMI	21007	VMEITNLWH	21048	YCNEYDEII
20926	SDMWGYQDP	20967	TDIDKQRDK	21008	VMGITHLWH	21049	YCPAANWNW
20927	SDMWLYQDP	20968	TDVGDYWHA	21009	VNCCSCKAN	21050	YCPEANWNL
20928	SDMWWHQDP	20969	TDVMLAEWT	21010	VNVHNWRCS	21051	YCPEANWSW
20929	SDMWWYEDP	20970	TENDIGCAT	21011	VPHNFEFYT	21052	YCPEASWNW
20930	SDMWWYKDP	20971	TGGWMFFWH	21012	VSADKCNLN	21053	YCPETNWNW
20931	SDMWWYPDP	20972	TGIDMAMEL	21013	VSAGKCNVN	21054	YCPGCGLMW
20932	SDMWWYQAP	20973	TGLDFYRWF	21014	VSCEDEGIL	21055	YCPVANWNW
20933	SDMWWYQDA	20974	THNDIGCAT	21015	VSQNQNCFK	21056	YCREANWNW
20934	SDMWWYQDH	20975	TKLNIRDRW	21016	VTKWDRWHD	21057	YCTEYDEIV
20935	SDMWWYQDR	20976	TKTGNNVMH	21017	VWKSDRVMM	21058	YDHGMHRAG
20936	SDMWWYQDT	20977	TKWNNLHIS	21018	VYRPNWEFH	21059	YDHGMHRDA
20937	SDMWWYQHP	20978	TMLCNRVCE	21019	VYRQFFEGI	21060	YDHGMHRYG
20938	SDMWWYQVP	20979	TMPMPPVAD	21020	VYTVNDGCV	21061	YDHGWKMAC
20939	SDMWWYQYP	20980	TNCCSCKAN	21021	WEDSMFCQY	21062	YDIHPRFFH
20940	SDMWWYRDP	20981	TNGFDKYID	21022	WEHSMFCKY	21063	YDLCEHRWD
20941	SDTWWYQDP	20982	TNIDGLWPI	21023	WEHSMFCRY	21064	YDPGMHRDG
20942	SECCSIFSS	20983	TNNNDRFQV	21024	WELSMFCQY	21065	YEWSWAEHS
20943	SEGEHQCKA	20984	TNNVSEFMP	21025	WLYIWEKWP	21066	YFCVNRHIA
20944	SEMWWYQDP	20985	TRLNDEDRG	21026	WLYTWEKWT	21067	YFIFPENED
20945	SFWGPDVWG	20986	TSGSHDFMG	21027	WLYTWGKWP	21068	YGIIMKIDL
20946	SGLKNDDRK	20987	TSMERWGYD	21028	WLYTWVKWP	21069	YGIRLKIDL
20947	SGMWWYQDP	20988	TSPADHKWR	21029	WNPSNMFNI	21070	YGIRNGKIG
20948	SMLCNRVCE	20989	TSSADHKWR	21030	WRKAIPYVM	21071	YGKISTIIN
20949	SNCCSCKAN	20990	TSTNDRFQV	21031	WRKGITYVM	21072	YGLENDDRK
20950	SNDCNLFDL	20991	TTLNLRDRW	21032	WRKGTPYVM	21073	YGLKDDDRK
20951	SNDCSLCDL	20992	TTMNIRDRW	21033	WRKGVPYVM	21074	YGLKNDARK
20952	SNMWWYQDP	20993	TTMRGFYHC	21034	WRKVIPYVM	21075	YGLKNDDCK
20953	SNNCNLCDL	20994	TTTGNNVLH	21035	WRQGIPYVM	21076	YGLKNDDRE
20954	SPFQQDARA	20995	TTWNNLHII	21036	WRRGIPYVM	21077	YGLKNDDRR
20955	SSGSHDFMG	20996	TYITKMDLV	21037	WVVNFIRQV	21078	YGLKNDDSK
20956	SVFDNAFNN	20997	VCDRCAPED	21038	WWIIEEHFK	21079	YGLKNDGRK
20957	SVFDNGFDN	20998	VEQLCIEWD	21039	WWIIEEHGK	21080	YGLKNDNRK
20958	SVLPDWGDS	20999	VERLCIEWV	21040	WWVIEEHCK	21081	YGLKNDYRK
20959	SVPRLIEDG	21000	VESNCIDDV	21041	WYFLAWSEI	21082	YGLMNDDRK
20960	SWKLDHMPP	21001	VFQWINYYM	21042	YALKNDDRK	21083	YGLQNDDRK
20961	TCKLDHMPP	21002	VGTDQTMHM	21043	YCHEYDEIV	21084	YGMKNDDRK
20962	TCNPEDSIG	21003	VHCELGQHS	21044	YCIFEYNGR	21085	YGMKNDDRR
20963	TCSELPERI	21004	VHISNMEGD	21045	YCIGCVDAI	21086	YGRKNDDRK

[2191]

21087	YGVGYLMPA	21095	YMKHFKNPS	21103	YMKRFRNPS	21111	YTFGLRVDD
21088	YGVKNDDRK	21096	YMKHFRDPS	21104	YMMHFRNPS	21112	YTNVDGFPI
21089	YHLIAIGPS	21097	YMKHFRNAS	21105	YMRHFRNPS	21113	YTPGNNVTR
21090	YHLIAIWPP	21098	YMKHFRNSS	21106	YNISIMEAG	21114	YVFDNAFDN
21091	YHLIAKWPS	21099	YMKHIRNPS	21107	YQPGCDCSC	21115	YVKISTVIN
21092	YHLIDIWPS	21100	YMKHVRNPS	21108	YQPGCDCTC	21116	YVLKNDDRK
21093	YHPGCDCAC	21101	YMKNFRNPS	21109	YSENRDHFV	21117	YWNEYDEIV
21094	YLVGYLMPA	21102	YMKQFRNPS	21110	YSERKGEYG		

[2192] 实施例36

[2193] 如通过机器学习鉴定的具有乳腺组织嗜性的AAV5变体

[2194] 此实施例描述了如通过机器学习方法鉴定的展现乳腺组织嗜性的工程化AAV5变体的优选特性。进行如实施例19和图19至图22中描述的相同生物信息学预处理、机器学习方法和分析以确定展现乳腺组织嗜性的工程化AAV5变体的优选特性。

[2195] 下文描述了AAV5 VP衣壳多肽的581至589区域中各位置处的有利生物物理特性和有利氨基酸残基，它们与较高的乳腺组织嗜性概率相关。表56中描述的任何以下机器学习衍生的位置氨基酸偏好可以单独存在或彼此组合存在或与实施例11，表21中观察和描述的任何氨基酸偏好组合存在。下文表56中列出了从ML模型中提取的排名靠前的重要乳腺嗜性特征之间共有的位置特征的概述。

[2196]

表 56
机器学习衍生的乳腺组织嗜性规则

Xaa1 处低表面可接近性 Xaa1 选自 C
Xaa1 处中等摩尔质量 Xaa1 选自 C
Xaa2 处高表面可接近性 Xaa2 选自 D、N 或 Q

[2197]

表 56 机器学习衍生的乳腺组织嗜性规则
Xaa2 处低亲疏水性 Xaa2 选自 D、E、R、K、H、N 或 Q
Xaa3 处高平均柔性 Xaa3 选自 D、E、R、P、G 或 S
Xaa3 处中等可突变性 Xaa3 选自 R 或 H
Xaa4 处高可突变性 Xaa4 选自 M、I、Q 或 T
Xaa4 处高溶解性 Xaa4 选自 W、F、I 或 L
Xaa4 处高表面可接近性 Xaa4 选自 E、R 或 K
Xaa5 处高溶解性 Xaa5 选自 W、F、I 或 L
Xaa5 处低可突变性 Xaa5 选自 Y、F 或 L
Xaa6 处高亲疏水性 Xaa6 选自 V、I 或 L
Xaa6 处中等摩尔质量 Xaa6 选自 D、I、L 或 N
Xaa8 处低表面可接近性 Xaa8 选自 C
Xaa8 处低可突变性 Xaa8 选自 C、R 或 H
Xaa9 处中等可突变性 Xaa9 选自 R 或 H

[2198] 下文表57提供了表现出乳腺组织嗜性并且符合表56中提供的一种或多种规则的

变体AAV VP1衣壳多肽的581至589区域的序列。因此，本公开提供了由具有SEQ ID NO:2的工程化AAV5 VP1衣壳多肽组成的rAAV，其中Xaa1至Xaa9区域是SEQ ID NO:23118-SEQ ID NO:24117中的任一者，如表57中所公开。本文还涵盖由工程化AAV5 VP2衣壳多肽和工程化AAV5 VP3衣壳多肽组成的rAAV，所述多肽在对应于AAV5 VP1 581至589区域中的氨基酸的AAV5 VP2(氨基酸残基445至453)和AAV5 VP3(氨基酸残基389-397)中的区域处具有下文表中公开的序列。

[2199] 表57

[2200] 驱动乳腺组织嗜性的AAV5 VP1衣壳多肽中581至589区域的序列

[2201]

SEQ ID NO	581-589 序列
23118	AAAQFGKHT
23119	AASKVKLHH
23120	ACDRFLAMY
23121	ACDRFMTMY
23122	ACDRFMVMY
23123	ACDRIMAMY
23124	ACDRVMAMY
23125	ACDRYIAEY
23126	ACDRYMAMY
23127	ADQCKWWCM
23128	ADQCQWWCV
23129	AELMALAPC
23130	AELMVLAQC
23131	AFDRFMAMY
23132	AFTIKDDWW
23133	AGAKYHDWL
23134	AGGKLKTIG
23135	AGGKLQTLV
23136	AIGQFMMHL
23137	AIQQFMMHL
23138	AIRLFMMHL
23139	AIRQFMMHR
23140	AIRQFMMHV
23141	AIRQFTMHL
23142	AIRRFMMHL
23143	AIWPAAWMD
23144	AKLMALAQC
23145	AKVNCVYYG
23146	AKVSTLNGM
23147	ALREFWTWW
23148	ALRQFMMHL
23149	ANCKHDHCH
23150	ANCKPDHCH

[2202]

23151	ANCKQHHCH	23192	CFCTYEIEH	23233	CIGLRLAWD	23274	CINSDGKYS
23152	ANCRQDHCH	23193	CFCTYKIQH	23234	CIGLRLDWD	23275	CIPSDRHFG
23153	ANRKQDHCH	23194	CFCTYKNEH	23235	CIGRFDHKQ	23276	CIPTYRHLV
23154	ANRQFMMHL	23195	CFGFVNSVT	23236	CIGWWGMMG	23277	CIQWWGMMG
23155	ANWKQDHCH	23196	CFGTYKIEH	23237	CIGYKWFPG	23278	CIRSHCRQH
23156	ANYKQDHCH	23197	CFGVVNSVN	23238	CIHADGKYS	23279	CIRSSTDQP
23157	AQISTLNGM	23198	CFRTYDLLS	23239	CIIPKLKER	23280	CIRTDWQRI
23158	AQNGLGTVQ	23199	CFWMYDLLS	23240	CIIPKMEER	23281	CISRFDHKL
23159	ARGKLQTLG	23200	CFYTYKIEH	23241	CIIPKMKAR	23282	CISRYDHKQ
23160	ARGPMLAVK	23201	CGHWTYDCY	23242	CIIPKMKEK	23283	CITADGKYS
23161	ARVSTLNGM	23202	CGLIMEMCC	23243	CIIPKMKKR	23284	CITTMNVKT
23162	ASARLYYCW	23203	CGLIMEMCL	23244	CIIPKMMER	23285	CITTYSHFG
23163	ASDIFDGHQ	23204	CGLIMEMWW	23245	CIIPKMRER	23286	CIVLRLAWD
23164	ASDKMDGGS	23205	CGLIMEVCW	23246	CIIPKVKER	23287	CIVPKMKER
23165	ASDKMDIGS	23206	CGLIMQMCW	23247	CIIPRMKER	23288	CIVWWGMMG
23166	ASERLHYCW	23207	CGLIMVMCW	23248	CIIRKMKER	23289	CKHPSWYMD
23167	ASERLYYCL	23208	CGLLMEMCW	23249	CIISKMKER	23290	CKKWTNQAH
23168	ASKQWDHKH	23209	CGMIMEMCW	23250	CIKNDWQRI	23291	CKPWTNQAH
23169	ASNKMDVGS	23210	CGPIMEMCW	23251	CIKPDWQRI	23292	CKQWTNQAQ
23170	ASQRLYYCW	23211	CGVIMEMCW	23252	CIKTDWERI	23293	CKVHKHWDR
23171	ASRQFMMHL	23212	CHFQYKNQA	23253	CIKTDWQGI	23294	CKVHRNWDR
23172	ASVRLYYCW	23213	CHFRYQNQA	23254	CIKTDWQRN	23295	CKVRKNWDR
23173	ASYKDQRYY	23214	CHMSFNQCY	23255	CIKTDWRRI	23296	CLAIFRLQS
23174	ATQKITGTS	23215	CHRKWEKQE	23256	CILPKMKER	23297	CLGATWVAA
23175	ATRQFMMHL	23216	CIALRLAWD	23257	CILSSPEQP	23298	CLGSTWAAA
23176	ATTQWGMCN	23217	CICSSPDHP	23258	CIMGHCRQH	23299	CLGSTWFAA
23177	AVHPMAYCF	23218	CICSSPEQP	23259	CIMPKMKER	23300	CLGSTWGAA
23178	AWDRFMAMY	23219	CIDLQLAWD	23260	CIMSDCRQH	23301	CLGSTWVAT
23179	CAHKEHVCH	23220	CIDLRLAWG	23261	CIMSHCREH	23302	CMCLKKFYY
23180	CAHWSYDCY	23221	CIDLRLAWH	23262	CIMSHCRKH	23303	CMCLKKYYY
23181	CAHWTYECY	23222	CIDLRLDWD	23263	CIMSHCRQD	23304	CMCLRKCYY
23182	CAHWTYNCY	23223	CIDLRLVWD	23264	CIMSHCRQY	23305	CMCVKKCYY
23183	CAHWTYYCY	23224	CIDLWLAWD	23265	CIMSHFRQH	23306	CMENYHQQY
23184	CAHWYYDQR	23225	CIEWGGMMG	23266	CIMSHGRQH	23307	CMMSHCRQH
23185	CANKENVCH	23226	CIEWWDMMG	23267	CIMSHNRQH	23308	CQDFRHDCA
23186	CANKGHVCH	23227	CIEWWGKMG	23268	CIMSHWRQH	23309	CQDFRHECS
23187	CANWTYDCY	23228	CIEWWGMMA	23269	CIMSHYRQH	23310	CQGKWEKQE
23188	CATKEHVCH	23229	CIEWWGMVG	23270	CIMSRCRQH	23311	CQKRIIMQA
23189	CEGWGTYCL	23230	CIEWWSMMG	23271	CIMTDWQRI	23312	CQPKWEKQE
23190	CEGWVTYCI	23231	CIGAHWHIP	23272	CINADGKYP	23313	CQQRLIMQA
23191	CFCILEKMT	23232	CIGINLDHH	23273	CINRFDHKQ	23314	CQREWEKQE

[2203]

23315	CQRKWEKQA	23356	CVYREIQTA	23397	DQKWEKYGL	23438	EEYDTVAAP
23316	CQRRWEKQE	23357	CVYREITTA	23398	DQKWEKYGM	23439	EFDIDQCCT
23317	CRGKSHEHT	23358	CVYRELKTA	23399	DQKWSVHMG	23440	EGDIFDGHQ
23318	CRGKSHNHT	23359	CVYREVKTA	23400	DQVWSVHMG	23441	EGIWQGMDA
23319	CRGKSRDHT	23360	CWAIFRLRS	23401	DREWSVHMG	23442	EHVIIEKNY
23320	CRGQSHDHT	23361	CWAIFRMQS	23402	DSKWLDRPV	23443	EHVILKKNY
23321	CRLIMEMCW	23362	CWAIVRLQS	23403	DSSKVKLHH	23444	EHVIMKKNY
23322	CRRKSHDHT	23363	CWAQLNTDP	23404	DTDKKGQGI	23445	EIKVVNSKD
23323	CSIWEDRSG	23364	CWAQLNTDS	23405	DTDKQGQGF	23446	EIKYRRCRA
23324	CSMERRNEL	23365	CWDIFRLQS	23406	DTDKQGQGL	23447	EIKYRRFMA
23325	CSMSHCRQH	23366	CWDIYNACD	23407	DTDKQGQGM	23448	EIKYRRSMA
23326	CSPNFQDQK	23367	CWGIFRLQS	23408	DTDKQGQGV	23449	EIREGEECC
23327	CSRATWGAE	23368	CWGQLNTDT	23409	DTDKQSQGI	23450	EIRVVNSRD
23328	CTDRMVKAA	23369	CWPQLNTDT	23410	DTNKQGQGI	23451	EISPAAWMD
23329	CTDRMVKAT	23370	CWSIFRLQS	23411	DVDYRSKAW	23452	EKMESPECL
23330	CTGNYHQQY	23371	CWVMMKTQE	23412	DVDYRYEAW	23453	EKTWLVVWC
23331	CTHKDIEMD	23372	CYCTYKIEH	23413	DVFIDWSGR	23454	EMDAQGAMW
23332	CTHKQSIAV	23373	DAAKFGKHT	23414	DVKTDTSRD	23455	EMRMNIQCE
23333	CTHRMVKAS	23374	DASKVKLHR	23415	DVRQILLCH	23456	EPEGALDKT
23334	CTLFSGRTA	23375	DDLGAGRSE	23416	DVRQILVCH	23457	EPEGELDRT
23335	CTQSFVMYQ	23376	DDNWHAWFM	23417	DVRQIMMCH	23458	EPEGELYKT
23336	CTVFSGRAA	23377	DDVGAGRSD	23418	DVRQLLMCH	23459	EPEGEVDKT
23337	CTVSSGRTA	23378	DDVMWQLIH	23419	DYFQYNEAR	23460	EQEAYILRI
23338	CVAIEMSVA	23379	DDWMPQWVD	23420	DYYQYNEAP	23461	ESDIFDAHQ
23339	CVCLKKCYY	23380	DEFIWLVMH	23421	EAISYVENV	23462	ESDIFDGHE
23340	CVCTYKIEH	23381	DEISNQNSY	23422	ECDIFDGHQ	23463	ESDIFDGHK
23341	CVFREIKTA	23382	DGGVTTYGG	23423	ECDILDGHQ	23464	ESDIFDGHP
23342	CVGINLEHH	23383	DHVIIKKNY	23424	ECDRYIAEH	23465	ESDIFDGHR
23343	CVGINLYHH	23384	DIRPSVSDS	23425	ECDRYIAQY	23466	ESDIFDGRQ
23344	CVGLNLDHH	23385	DKEWSVHMG	23426	ECDRYISEY	23467	ESDIFDVHQ
23345	CVHGQYYQE	23386	DLDAQGAMW	23427	ECGFIRLNE	23468	ESDIFHGHQ
23346	CVLIMEMCW	23387	DLGMINNMA	23428	ECKPIAGDR	23469	ESDIFVGHQ
23347	CVLYYNQQC	23388	DMDAKGAMW	23429	ECYHVLANT	23470	ESDISDGHQ
23348	CVMCYNDCK	23389	DMDAQGSMW	23430	EDIHMIAGH	23471	ESDIVDGHP
23349	CVMSHCRQH	23390	DMSIATWNP	23431	EDTHMIAGP	23472	ESDIVDGHQ
23350	CVQGQYNQE	23391	DPDKQGQGI	23432	EDTHMIVGH	23473	ESDPMQLFT
23351	CVQGQYYQA	23392	DPKWEKYAM	23433	EEFDTIAAP	23474	ESDPMQRFT
23352	CVQGRYYQE	23393	DPKWEKYGT	23434	EEFDTLAAP	23475	ESDVFEGHQ
23353	CVQTMNGRS	23394	DQAWSVHMG	23435	EEFDTVAAR	23476	ESKSCGLKG
23354	CVVINLDHH	23395	DQEWSVHMC	23436	EEFDTVATP	23477	ESKSCGLRG
23355	CVYREIKTS	23396	DQEWSVNMG	23437	EEFDTVGAP	23478	ESNIFDGHQ

[2204]

23479	ESTSCGLTG	23520	GHFTVLNWG	23561	HYCKIIPCH	23602	ITRTYQERQ
23480	ESVDWIKCE	23521	GNCKQDHCH	23562	HYCKIIPSH	23603	IVRDKTQDD
23481	ESWPAAWMD	23522	GRHTAVATG	23563	HYCKIIPYN	23604	KANMWLACV
23482	ETDIFDGHQ	23523	GSERLYYCW	23564	HYCKIITYH	23605	KCGFIRLNE
23483	EVEAIRGIR	23524	GSVMVAITG	23565	HYCKILPYH	23606	KCVIGKGRR
23484	EVEIFRLEG	23525	GYGRMNKAS	23566	HYCKLIPYH	23607	KDFWEHECF
23485	EVIAYVENV	23526	HASICQGGD	23567	HYCKMIPYH	23608	KDFWEHLCF
23486	FDWMPQWVD	23527	HAYIYETWP	23568	HYCKNIPYH	23609	KDFWEHQCV
23487	FGGQLRVMY	23528	HDCDMQASE	23569	HYCKTIPYH	23610	KDFWERQCF
23488	FGRERVWDS	23529	HDKWGWWRE	23570	HYCMIIPYH	23611	KDFWEYQCF
23489	FGSLYRRRG	23530	HDRKACMIA	23571	HYCQIIPYH	23612	KDFWGHQCF
23490	FIGLRYKDQ	23531	HDVMWQLIH	23572	HYCRIIPYH	23613	KDFWKHQCF
23491	FKLWKLFQA	23532	HDWMPQWVD	23573	HYFKIIPYH	23614	KDRMGINQC
23492	FKMDMTKRS	23533	HEAKLVPKA	23574	HYGKIIPYH	23615	KDRMGLNQR
23493	FKMWKLFKA	23534	HEFIWLEMH	23575	HYYKIIPYH	23616	KDRMGLNQW
23494	FKRWKLFQA	23535	HEYIYETLP	23576	ICKIRSRLD	23617	KDRMVLNQC
23495	FNIPCLFSR	23536	HEYIYVTWP	23577	ICTIGSRLD	23618	KDTHMIAGH
23496	FNIQCLCSR	23537	HFDFNHQCQ	23578	IDRKACMIA	23619	KDVTRLRCP
23497	FNIQCLFTR	23538	HFLEVAWGS	23579	IGVQFLKDH	23620	KDVWEHQCF
23498	FNIQCLISR	23539	HHRHIANNG	23580	IIGDKTQDD	23621	KEFDTVAAP
23499	FNIQCLYSR	23540	HKHMILYHF	23581	IIRDKTEDD	23622	KFDIDQCCT
23500	FNIRCLFSR	23541	HMDIQMVMP	23582	IIRDKTQDG	23623	KKFWDTVIY
23501	FNLQCLFSR	23542	HMDIQMYMP	23583	IIRDKTQND	23624	KKHMILCHF
23502	FQTSMYCDW	23543	HMDIQVFMP	23584	IIRDKTRDD	23625	KQTQIRRHM
23503	FRGHAMWDA	23544	HMSIATWYP	23585	IIRDRTQDD	23626	KSDIFDGHQ
23504	FRGHGMWYA	23545	HMTFYLSRD	23586	IIRDTTQDD	23627	KSVDWIKCE
23505	FRGHGTWDA	23546	HPSICEGGD	23587	IIREKTQDD	23628	KYAYILDCA
23506	FRGHGVWDA	23547	HSEMAMLMQ	23588	IIRGKTQDD	23629	KYAYSLYCA
23507	FRGLGMWDA	23548	HSEMVMWMQ	23589	IIRYKTQDD	23630	KYGYSLDCA
23508	FSEQNVSEA	23549	HSGELYDAF	23590	IMRDKTQDD	23631	KYTYSLDCA
23509	FSIQCLFSR	23550	HSGELYDCF	23591	IPRTYQEMQ	23632	KYWNYAEAK
23510	FTEQIGRQM	23551	HSGELYDSF	23592	IQHQSRRHM	23633	LADIEYSHD
23511	FTQQIDRQM	23552	HSGIHCCCG	23593	IQLKAVPSC	23634	LAGKEGAHF
23512	FVCLRYKDQ	23553	HSNFFHLND	23594	IQLKAVTCC	23635	LCASRGLDA
23513	FVGWYLNSP	23554	HSNFFHMTD	23595	IQLKTVTSC	23636	LCGAIIKAS
23514	FVGWYWDSP	23555	HVDGNIACN	23596	IRGHGMWDA	23637	LCGLIIKAS
23515	FVGWYWNPP	23556	HVDSNIACT	23597	IRLKAVTSC	23638	LCGPIIKAA
23516	FVGWYWNSH	23557	HVDSNISCN	23598	IRRHQMTDK	23639	LCGPIIKAP
23517	FVGWYWNSR	23558	HVDSNLACN	23599	ISRDKTQDD	23640	LCGPIIKAY
23518	FVMGYNDCK	23559	HVDSNVACN	23600	ITRTYQEME	23641	LCGPIIKPS
23519	GCGLVQRFM	23560	HVNSNIACN	23601	ITRTYQEMK	23642	LCGPIIKSS

[2205]

23643	LCGPIIRAS	23684	MANKWLACV	23725	MQPPSRRRM	23766	NDVMWHLIH
23644	LCGPILKAS	23685	MANRWLACV	23726	MQPQFRRHM	23767	NDVMWQLIR
23645	LCGPIVKAS	23686	MASMWLACV	23727	MQPQSRRDM	23768	NDVMWQLIY
23646	LCGPLIKAS	23687	MATMWLACV	23728	MQPQSRRHK	23769	NDVMWQLVH
23647	LCGQIIKAS	23688	MCGLRYYWA	23729	MQPQSRRNM	23770	NDWMPQWVD
23648	LCTSRGLDP	23689	MCGPIIKAS	23730	MQPRGRRHM	23771	NEISNQNTY
23649	LDRKACLIA	23690	MCLLIRACE	23731	MQPRSRRHM	23772	NFDFNHQCH
23650	LDRKACMIG	23691	MCLWILACE	23732	MQRRSRRHM	23773	NFGIYHLQG
23651	LDRKACMIS	23692	MCLWIRPCE	23733	MQTESRRHM	23774	NKHCIAMPC
23652	LDRKACMMA	23693	MCRHIIPDN	23734	MQTQSLRHM	23775	NKYCIAMAC
23653	LDRKACMVA	23694	MEHQSRRHM	23735	MQTQSRRNM	23776	NKYCIAMPF
23654	LDRKACVIA	23695	MEPESRRHM	23736	MRIIQGDNY	23777	NKYCITMPC
23655	LDRKAFMIA	23696	MEPQSRRHI	23737	MRPQCRRHM	23778	NKYCLAMPC
23656	LDRKAWMIA	23697	MEPQSRRHM	23738	MRPQSRRHK	23779	NMDAQGAMW
23657	LDRKDCMIA	23698	MFEKYFQSM	23739	MRPQSRRHM	23780	NNGEWDVCY
23658	LDRKPCMIA	23699	MFMKDELIA	23740	MRPQSRRHR	23781	NPKWEKYGM
23659	LDRKSCMIA	23700	MFMKNELMA	23741	MRVIQGDNC	23782	NQDWLLSVP
23660	LDRKVCMIA	23701	MHKHFEDHK	23742	MSNMWLACV	23783	NQEWSVHMG
23661	LDRQACMIA	23702	MHKHFKGHK	23743	MSPKPAYWG	23784	NVDVRESEW
23662	LDRRACMIA	23703	MHKHYKDHK	23744	MSTKPAYWA	23785	NWHEKRHRG
23663	LFTDYWEQG	23704	MHKRFKDHK	23745	MSTKPEYWG	23786	NYKPSFLTR
23664	LGILYRRRG	23705	MHLLKDFCT	23746	MTAKMAGRM	23787	PCAIEESHD
23665	LGSPYRRRG	23706	MHQHFKDHK	23747	MTNMWLACV	23788	PCDRFMAMY
23666	LIRDKTQDD	23707	MHRHFKDHK	23748	MVAIKQYHT	23789	PEDFENHDG
23667	LKAQFHWGA	23708	MIDWGNKCT	23749	MVGVTRDAS	23790	PLRAEMSFH
23668	LKDQFHWCA	23709	MKPQGRRHM	23750	MVVIKQYQT	23791	QCDRVWLNQ
23669	LKDRFHWGA	23710	MKPQSRRHL	23751	MVVIKRYHT	23792	QCEFIRLNE
23670	LKMDMTKHS	23711	MLAGWWQDT	23752	MWFFMKFAK	23793	QCGFIRLDG
23671	LKVDMTKRS	23712	MLGEWWPGC	23753	MWGEWWPGG	23794	QCGFIRLHE
23672	LLFFMKFAK	23713	MLNWGNKCT	23754	MWVIQGDNY	23795	QCGFIRLNG
23673	LQMDMTKRS	23714	MNKHFKDHK	23755	MYKHFKDHK	23796	QCGFIRLNQ
23674	LQPQMRRHM	23715	MPKHFKDHK	23756	NDGMVHNCW	23797	QCGFIRVNE
23675	LQPQNRRHM	23716	MQHLSRRHM	23757	NDHWHAWFM	23798	QCGFLRLNE
23676	LQPQSRRHV	23717	MQHQCRRHM	23758	NDKWGWWRQ	23799	QCHIIVAPE
23677	LRCQMGDHW	23718	MQHRSRRHM	23759	NDLMWQLIH	23800	QCHIMVARE
23678	LRMDMTKRS	23719	MQLQCRRHM	23760	NDNWDAWFM	23801	QCHINVARE
23679	LRVIQGDNY	23720	MQLRSRRHM	23761	NDNWHAWFR	23802	QCPIIVARE
23680	LTHKRESEW	23721	MQPDSRRHM	23762	NDNWHAWFV	23803	QDFWEHQCF
23681	LTPKRESEC	23722	MQPHIRRHM	23763	NDNWHAWIM	23804	QEFDTVAAP
23682	LTRKRESEC	23723	MQPHSRRDM	23764	NDNWHEWFM	23805	QEHMILCHF
23683	LWFYMKFAK	23724	MQPLSRRHM	23765	NDVMWELIH	23806	QETWLVVWC

[2206]

23807	QFGFIRLNE	23848	QQEAFILRI	23889	RETKVFELK	23930	RYCKIIPYH
23808	QFHIDQCCT	23849	QQEAYILKI	23890	RFKLLMTGR	23931	RYWNYAEAK
23809	QGGFIRLNA	23850	QQEAYILRL	23891	RKTWLVVWC	23932	SAGIFYAHK
23810	QGGFIRLNG	23851	QQEAYILRV	23892	RLKPSAVQR	23933	SAYIFYAHK
23811	QIGCNVMIS	23852	QQEEYILRI	23893	RLRNCGSNG	23934	SCGLVERFM
23812	QIGFNVMLS	23853	QQEGYILRI	23894	RLVRDSKNT	23935	SDQCQWWCM
23813	QIKYRRCMA	23854	QQESYILRI	23895	RLVRDSKTN	23936	SGGLVQRFM
23814	QIWPAAWMD	23855	QQEVYILRI	23896	RMIQPEHVV	23937	SGNIYESGQ
23815	QKDMIICHF	23856	QQHMILCHF	23897	RMKWYLLCV	23938	SIRDKTQDD
23816	QKDMILCHF	23857	QSDDWIKCE	23898	RMKWYLTCV	23939	SIRQFMMHL
23817	QKDWGRRVT	23858	QSDIFDGHQ	23899	RMPIELMHG	23940	SIRQYMMHL
23818	QKEAYILRI	23859	QSDRVWLNP	23900	RNNWKVDSQ	23941	SKGCIIAMQ
23819	QKHMIICHF	23860	QSEMAMWMQ	23901	RNNWKVNSR	23942	SQVSTLNGM
23820	QKHMILCGF	23861	QSGWLLEAK	23902	RNYWCMCGA	23943	SSHAANFCG
23821	QKHMILCRF	23862	QSIDWIKCE	23903	RPATVHASF	23944	SSHATTFCG
23822	QKHMILCYF	23863	QSLDWIKCE	23904	RPEMVHASF	23945	SSKYEMKFL
23823	QKHMLLCHF	23864	QSVDWIKGE	23905	RPERVHASF	23946	STDKDIEMD
23824	QKHMMLCHF	23865	QSVDWIKYE	23906	RQEAYILRI	23947	STHKDIEMG
23825	QKHMSLCHF	23866	QSVDWIQCE	23907	RQGCLLAWQ	23948	STHKDIEMN
23826	QKHMTLCHF	23867	QSVDWIRCE	23908	RQPQSRRHM	23949	STHKDIVMD
23827	QKHRILCHF	23868	QSVDWLKCE	23909	RRDELLIPD	23950	STTQWGVCN
23828	QKHRLLCHF	23869	QSVDWMKCE	23910	RRIELLIPD	23951	SWHEKRHRG
23829	QKHTILCHF	23870	QSVDWVKCE	23911	RRKWYLPCV	23952	SWSMQSKST
23830	QKHVILCHF	23871	QSVVWIKCE	23912	RRNELLLPD	23953	SYGQIEGYQ
23831	QKIWLVVWC	23872	QSVYWIKCE	23913	RRSELLIPD	23954	TCAVFVIHC
23832	QKKMILCHF	23873	QVKYRTCDW	23914	RRTELLIPD	23955	TCDRFMAMY
23833	QKNMILCHF	23874	QVTSDGKIV	23915	RSVDWIKCE	23956	TCFIGKGRR
23834	QKPMILCHF	23875	QYSRMNKAS	23916	RVADLMYIG	23957	TCGIGKGRR
23835	QKPWLVVWC	23876	RAKTCIYKY	23917	RVGDVMYIG	23958	TCIIGKGRR
23836	QKRMILCHF	23877	RAKTCLYKC	23918	RVSDVMYIG	23959	TCIIGQGRR
23837	QKTWLIVWC	23878	RCDGLWCHF	23919	RVTDVMYIG	23960	TCLIGKGRR
23838	QKTWLVGWC	23879	RCGFIRLNE	23920	RVVRDSKNN	23961	TCVIAKGRR
23839	QKTWLVVWF	23880	RCGFIRLNG	23921	RWGFIRLNE	23962	TCVIEKGRR
23840	QKTWLVVWV	23881	RCMKLVGTA	23922	RWRHCGSNG	23963	TCVIGKGQR
23841	QKTWLVVWY	23882	RCRNCGSNG	23923	RWRICGSNG	23964	TCVIGQGRR
23842	QNYWCVCGA	23883	RDFWEHQCF	23924	RWRNCGGNG	23965	TCVIGRGRR
23843	QPDPMKQWS	23884	RDQCVASCA	23925	RWRNCGNNG	23966	TCVIVKGRR
23844	QPKEWVIKD	23885	REAKVFEIK	23926	RWRNCGSNA	23967	TEDFENHDA
23845	QPREWVIKE	23886	REAKVFEPK	23927	RWRNCSSNG	23968	TIRQFMMHL
23846	QQDPMKQWT	23887	REPKVFELK	23928	RWRNGGSNG	23969	TKGCIIDMQ
23847	QQEACILRI	23888	RESKVFELK	23929	RWRSCGSNG	23970	TKLNCVYYG

[2207]

23971	TKVNCIYYG	24008	VSDPMQQFT	24045	YDWMAQWVD	24082	YGWMPQWVD
23972	TLRAAMSFH	24009	VVHQMAYCF	24046	YDWMHQWID	24083	YHGHIANNG
23973	TLRAEMSFR	24010	VWDWEHKCT	24047	YDWMPEWVD	24084	YHMGFNQCY
23974	TLRAEMSIH	24011	VWFFMKFAK	24048	YDWMPHWGD	24085	YHMNFNQCY
23975	TLRAVMSFH	24012	VWNGEHKCT	24049	YDWMPHWVE	24086	YHMSFKQCY
23976	TNCKQDHCH	24013	WAAIEYSHD	24050	YDWMPKWVD	24087	YHMTFNQCY
23977	TRVNCVYYG	24014	WADIDYSHD	24051	YDWMPPWVD	24088	YHRHIADNG
23978	TSERLYYCW	24015	WADIEYGHD	24052	YDWMPQLVD	24089	YHRHTANNG
23979	TSHATNFCG	24016	WADIEYSHA	24053	YDWMPQWAD	24090	YHRNIANNG
23980	TSNKDQRYY	24017	WAGIEYSHD	24054	YDWMPQWDA	24091	YIRDGGSHD
23981	TSSKDQRYY	24018	WAYIEYSHD	24055	YDWMPQWDD	24092	YIRDGVSHG
23982	TSYKDERYY	24019	WCGPIIKAS	24056	YDWMPQWFD	24093	YIRDGVSHH
23983	TSYKNQRYY	24020	WDFHSIFPG	24057	YDWMPQWGA	24094	YIRVGGNNQ
23984	TTTQWGMCN	24021	WDYHIIFPG	24058	YDWMPQWGD	24095	YIRVGVSHD
23985	TVLAWGLCN	24022	WDYHSICPG	24059	YDWMPQWID	24096	YLHEKRHRG
23986	TVRADMSFH	24023	WDYHSIFAG	24060	YDWMPQWIG	24097	YLRSLKCGD
23987	TYKPSFWTR	24024	WDYHSIFPC	24061	YDWMPQWIV	24098	YNWMPQWVD
23988	TYWNYAESK	24025	WDYHSIYPG	24062	YDWMPQWLG	24099	YQASIYCDW
23989	VASKTLWET	24026	WDYHSVFPG	24063	YDWMPQWVA	24100	YQTSIYCDW
23990	VCFKLFHVG	24027	WDYHTIFPG	24064	YDWMPQWVH	24101	YSDVLIYAN
23991	VDRKACMIA	24028	WQHGREADH	24065	YDWMPQWVV	24102	YSDVLVCAN
23992	VFTDYWEKG	24029	WSNFALWCQ	24066	YDWMPQWVY	24103	YSDVMICAN
23993	VFTDYWERG	24030	WSSFAVWCQ	24067	YDWMRQWVD	24104	YVDSNIACN
23994	VFVGKNAAW	24031	WSSFTLWCQ	24068	YDWMSQWVD	24105	YVLESNYCE
23995	VGSLYRRRG	24032	YAQSFSQGI	24069	YDWMTQWVD	24106	YVRDGVSHD
23996	VIRDKTQDD	24033	YAWMPQWVD	24070	YDWRPQWVD	24107	YVRQILMCH
23997	VLDAIHWHW	24034	YCFINSAYP	24071	YDWVPQWVD	24108	YVVESNYCQ
23998	VQPESRRHM	24035	YCPRLVEAS	24072	YDYDMQASE	24109	YVWMPQWVD
23999	VRFTVLNWG	24036	YDCDMKASE	24073	YEAPTILNP	24110	YWHEQRHRG
24000	VRIQPQHNK	24037	YDCDMQASG	24074	YEWMPQCVD	24111	YWHERRHRG
24001	VRLKPQHNK	24038	YDCDVQASE	24075	YFCILDKMT	24112	YWREKRHRG
24002	VRLQPKHNK	24039	YDCVMQASE	24076	YFCIMEKMT	24113	YYCKIIPYH
24003	VRLQPQHDK	24040	YDFDMQASE	24077	YFCIQEKMT	24114	YYFQYNEAP
24004	VRLQPQHDR	24041	YDGMPQWVD	24078	YFCSWNTQE	24115	YYKPSFWTR
24005	VRLQPQHSR	24042	YDLMPKWVD	24079	YFDFNHQCQ	24116	YYWHVSAFG
24006	VRLQPQLNK	24043	YDLMPQWVD	24080	YGHEKRHRG	24117	YYWHVSTFW
24007	VRLQPQRNK	24044	YDWDMQASE	24081	YGLIMEMCW		

[2208] 实施例37

[2209] 如通过机器学习鉴定的具有肺组织嗜性的AAV5变体

[2210] 此实施例描述了如通过机器学习方法鉴定的展现肺组织嗜性的工程化AAV5变体的优选特性。进行如实施例19和图19至图22中描述的相同生物信息学预处理、机器学习方法和分析以确定展现肺组织嗜性的工程化AAV5变体的优选特性。

[2211] 下文描述了AAV5 VP衣壳多肽的581至589区域中各位置处的有利生物物理特性和有利氨基酸残基，它们与较高的肺组织嗜性概率相关。表58中描述的任何以下机器学习衍

生的位置氨基酸偏好可以单独存在或彼此组合存在或与实施例12，表23中观察和描述的任何氨基酸偏好组合存在。下文表58中列出了从ML模型中提取的排名靠前的重要肺嗜性特征之间共有的位置特征的概述。

[2212]

表 58
机器学习衍生的肺组织嗜性规则

Xaa1 处高可突变性 Xaa1 选自 D、E、M、A、I、Q 或 T
Xaa2 处高摩尔质量 Xaa2 选自 F
Xaa2 处低可突变性 Xaa2 选自 Y、F 或 L
Xaa3 处低可突变性 Xaa3 选自 K、V、P 或 H
Xaa3 处低亲疏水性 Xaa3 选自 K 或 R
Xaa4 处低可突变性 Xaa4 选自 K 或 P
Xaa4 处高平均柔性 Xaa4 选自 D、E、P 或 S
Xaa5 处低平均柔性 Xaa5 选自 W、M 或 F

[2213]

表58 机器学习衍生的肺组织嗜性规则
Xaa5处高溶解性 Xaa5选自W、F、I或L
Xaa6处中等可突变性 Xaa6选自R或H
Xaa6处高表面可接近性 Xaa6选自T
Xaa7处低可突变性 Xaa7选自C
Xaa7处高溶解性 Xaa7选自W、V、M、F、I或L
Xaa8处高可突变性 Xaa8选自D、E、M、A、I、Q或T
Xaa8处低亲疏水性 Xaa8选自R或K
Xaa9处高平均柔性 Xaa9选自R或G

[2214] 下文表59提供了表现出肺组织嗜性并且符合表58中提供的一种或多种规则的变体AAV VP1衣壳多肽的581至589区域的序列。因此,本公开提供了由具有SEQ ID NO:2的工程化AAV5 VP1衣壳多肽组成的rAAV,其中Xaa1至Xaa9区域是SEQ ID NO:17118-SEQ ID NO:1 8117中的任一者,如表59中所公开。本文还涵盖由工程化AAV5 VP2衣壳多肽和工程化AAV5 VP3衣壳多肽组成的rAAV,所述多肽在对应于AAV5 VP1 581至589区域中的氨基酸的AAV5 VP2(氨基酸残基445至453)和AAV5 VP3(氨基酸残基389-397)中的区域处具有下文表中公开的序列。

[2215] 表59

[2216] 驱动肺组织嗜性的AAV5 VP1衣壳多肽中581至589区域的序列

SEQ ID NO	581-589 序列
17118	AAKAIHYWM
17119	AAMPFYWFE
17120	AANKMNAER
17121	ACSTVMDFS
17122	ACSTVTDFN
17123	ADADMMWMA
17124	ADCSSSDWS
17125	AEKAIHYWM
17126	AFCCDHNAA
17127	AFEDYGSWE
17128	AFKAPRRGV
17129	AFKATHRGV
17130	AFKATRGGV
17131	AFKATRLGV
17132	AFKATRQGV
17133	AFKATRRAV
17134	AFKATRRGA
17135	AFKATRRGM
17136	AFKATRRSV
17137	AFKSTRRGV
17138	AFKTTRRGV
17139	AFMPKESKE
17140	AFNPQEMWA
17141	AFNQQEVWA
17142	AFQSTDWGW
17143	AFVHWPIHG
17144	AFVNWPLHG
17145	AFVNWTIHG
17146	AFVTWPIHG
17147	AGDFSGYYG
17148	AGETNNGGY
17149	AGSASNWAD
17150	AGVFRKITM
17151	AGYASHWAD
17152	AGYASNWGD
17153	AHCHFSNMQ
17154	AHMVENDTP
17155	AHWDFSNMQ
17156	AIHDQLEAM
17157	AKAFIEHTY
17158	AKEQGAVGR
17159	AKGAYNQGN
17160	AKGFLEHTY
17161	AKGSYNQGN
17162	AKGTYNHGN
17163	AKHFYQKFH
17164	AKISAAWDR
17165	AKNAMHNFF
17166	AKNYMHCYW
17167	AKRDHHYFA
17168	AKTDNHVLH
17169	ALTYCHWTD
17170	ALTYCNWAD
17171	AMKNLKRTG
17172	AMVCQAIVF
17173	AMVCQIIVF
17174	AMVCRVIVF
17175	ANSSNHSQE
17176	AQHNSESNI
17177	ARHVNFEMS
17178	ASKMSHKEA
17179	ASKRSGKQP
17180	ASMEHMKKA
17181	ASRIYCRHE
17182	ASRSLERQA
17183	ASYFMEITR
17184	ATEQWDFCR
17185	ATSVFCVFR
17186	ATTSNNLTR
17187	ATTSSNLTR
17188	ATTSYNLAR
17189	AVDPTFKPH
17190	AVEFLAHQY
17191	AVKCTEYSA
17192	AVNPAFKPH
17193	AVYPLFAPE
17194	AWKMMHCDI
17195	AWKMMHCHN
17196	AWKMMHFDN
17197	AWKMMHGDN
17198	AWKMMNCDN
17199	AWQMMHCDN
17200	AWRMMHCDN
17201	AWYASNWAD
17202	AYKNCHRHG
17203	AYRMFNNVQ
17204	CAKACDTSS
17205	CAKDCDASS
17206	CAKNCDTSS
17207	CARDCDTSS
17208	CDYVQREWD
17209	CDYVQRKWD
17210	CDYVQRRWD
17211	CERRALKMI
17212	CESIMHPDQ
17213	CFKWTIVNG
17214	CFRQDRIIF
17215	CIESFIENH
17216	CIWQHNVTQ
17217	CKAHFGFME
17218	CKDHFGFMV
17219	CKDHFGVME
17220	CKEAFDVNE
17221	CLRQDRIIS
17222	CPEAFDVNE

[2217]

17223	CQDHFGFME	17264	DKTYMHCYW	17305	ECKQQGVRR	17346	EKMGYVHKG
17224	CSKDCDTSS	17265	DMSHTFVAA	17306	EFCEMQSLC	17347	EMKDLKRTG
17225	CSPPIVMDR	17266	DNAFVHCNQ	17307	EFCHALVCQ	17348	EMKNLKRAG
17226	CSYFMEITR	17267	DNAIVHCDQ	17308	EFEDYGGWE	17349	EMKNLKRNG
17227	CTAPIVMDR	17268	DNAIVHCNE	17309	EFEMSVFSN	17350	EMKNLKRTC
17228	CTKDCDTSS	17269	DNAIVHCNL	17310	EFFIFGAID	17351	EMKNLKRTS
17229	CTPPIAMDR	17270	DNAMVHCNQ	17311	EFFIIGAIG	17352	EMKNLRRTG
17230	CTPPIFMDR	17271	DNAVVHCNQ	17312	EFFIIGAIV	17353	EMRPQKMSR
17231	CTPPIIMDR	17272	DNGIVHCNQ	17313	EFFIIGAVD	17354	ENAQFLVWP
17232	CTPPILMDR	17273	DRCMQAKHG	17314	EFFIIGTID	17355	ENAQILVWA
17233	CTPPIVTDR	17274	DRCMQTKHA	17315	EFFIIVAID	17356	ENAQVLVWA
17234	CTPPTVMDR	17275	DRNYMHCYW	17316	EFFIVGAID	17357	ENAQYLVWA
17235	CTSFACKSF	17276	DSAIVHCNQ	17317	EFGHAMVCQ	17358	ENDSVSAIA
17236	CVKDCDTSS	17277	DSKEEGLAF	17318	EFITSVQCS	17359	ENEFIYCHP
17237	DADNYMIYQ	17278	DSMAMSARA	17319	EFKKEFEAF	17360	ENEWTYHQG
17238	DADYYMIYR	17279	DSMGMTARA	17320	EFKMSGFSN	17361	ENGQFLVWA
17239	DARVSNIRQ	17280	DSREEGLTF	17321	EFKMSVFSS	17362	ENHRITGKS
17240	DCRWYGNNS	17281	DSRQLHHLG	17322	EFKMSVFTN	17363	ENHSVSAIG
17241	DCTNNELKR	17282	DSRQLHHTG	17323	EFKMSVISN	17364	ENSQASAFG
17242	DCTNNKVKR	17283	DSRQLHRIG	17324	EFKMSVVSN	17365	ENSQFLVWA
17243	DCTNTKLKR	17284	DSRQLHYIG	17325	EFKMTVFSN	17366	ENSQNSAFG
17244	DEPCEELAF	17285	DSRRLHHIG	17326	EFKQEFVAF	17367	ENSQTSPFG
17245	DFIQSDTEF	17286	DTAIVHCNQ	17327	EFKQEYEAF	17368	ENVMHSEFR
17246	DFVQSDTGF	17287	DTNQLNDLS	17328	EFRMSVFSN	17369	ENVQFLVWA
17247	DGCSLYAFS	17288	DTTPNESDR	17329	EFTRKSGKF	17370	ENYRITGQS
17248	DGDYYMIYQ	17289	DVCTVFVTA	17330	EFTRRSCKF	17371	EPRPCNFWC
17249	DGTNNKLKR	17290	DVDYYMIYQ	17331	EHGCTHMGN	17372	EPRSCNFWG
17250	DHCGPIAWP	17291	DVVRVQKMA	17332	EHKQQGVRR	17373	EPRSCNFWW
17251	DHPYQHIEG	17292	DWASNNWWP	17333	EHQPFSMFS	17374	EPRSCNYWC
17252	DICCNYDAQ	17293	DWMTRHKVI	17334	EHRQMHDAY	17375	EQMGYVHKA
17253	DIGMRPSGW	17294	DWMTRNKVI	17335	EHRQMHDSF	17376	ERAEFEMKQ
17254	DKHFYEKFH	17295	DWMTRYKLI	17336	EHRQMHDSS	17377	ERNDFHWYM
17255	DKNYMHCHW	17296	DWMTRYKVL	17337	EHSCNHMGN	17378	ERSCTHMGN
17256	DKNYMHWYW	17297	DWSANNWWP	17338	EHSYTHMGN	17379	ERSDFHWCM
17257	DKNYMHYYW	17298	EAEYWSCGR	17339	EHTRQDVCD	17380	ERSEFEMKE
17258	DKNYTHCYW	17299	EAEYWTCGR	17340	EKDHFGVMQ	17381	ERSGFHWYM
17259	DKRDNHYFA	17300	EAHYTNECR	17341	EKHHFDVMQ	17382	ESKMNHKEA
17260	DKRDYHYFA	17301	EARDSCWWG	17342	EKLMYDKDL	17383	ESKMSHKQA
17261	DKRNHHYFA	17302	EARYCCWWG	17343	EKLNHETGY	17384	ESMFRDSDK
17262	DKRYHHYFA	17303	EARYSCWSG	17344	EKLSHEAGY	17385	ESMGICHNS
17263	DKSYMHCYW	17304	EARYTCWWG	17345	EKLTHEAGY	17386	ETIRDMSGF

[2218]

17387	ETMAQSCYR	17428	FGKLRKLKA	17469	GTRLKGNIM	17510	HVASNYDTS
17388	ETRSCNFWC	17429	FGRYRLLKG	17470	GVEFLAHRY	17511	HVGTHGWHR
17389	EVASQFFEH	17430	FGRYRLWEG	17471	GVKDNNRRL	17512	HVITHGWHR
17390	EVFKIDMIP	17431	FGRYRLWKA	17472	GWDKLMMLA	17513	HVVHLDDYW
17391	EVGEMGILR	17432	FGRYRLWQV	17473	GWHCFSFVV	17514	HVVTHGWQR
17392	EVGRIGWMS	17433	FGRYRLWRG	17474	GWKMMHCDN	17515	HVVTLGWHR
17393	EVQKWFRKS	17434	FHKKFHSKF	17475	GWNCFSFVV	17516	HVVTNGWHR
17394	EVRYSCWWG	17435	FHKQFHSEF	17476	GYEPTHSPR	17517	IAIKQRDGA
17395	EVTDEHCCW	17436	FHKQFHSRF	17477	GYKDCHRHG	17518	IANKQQLAR
17396	EVTNEHCCL	17437	FHKQYHSKF	17478	GYKPTHSPR	17519	IAVKQPDGA
17397	EVTNEHGCW	17438	FLKQFHSKF	17479	GYQMFNNVQ	17520	IAVPCASFY
17398	EVVKIDMIS	17439	FNIDCFNAA	17480	GYQPTHSHR	17521	IAVPSCWAR
17399	EVVKIDMTP	17440	FNKQFHSKF	17481	GYQPTNSPR	17522	ICHDCMNYR
17400	EWKMMHCDN	17441	FNVDCFSAA	17482	GYRPTHSPR	17523	ICVPSCWAR
17401	EWWRYMAYY	17442	FRKQFHSKF	17483	HATGVGKDR	17524	IEHDFNSAF
17402	EYICFNCYH	17443	GDYVQRQWD	17484	HAVTHGWHR	17525	IFHDCMNSR
17403	EYKEQGVRR	17444	GERWAHNML	17485	HDAQFLPDE	17526	IFHDFMNYR
17404	EYKKQGVRR	17445	GERWPHHML	17486	HGDHFDRNS	17527	IFSPQQLGA
17405	EYKMSHKEA	17446	GERWPHNMW	17487	HIEMRNSHM	17528	IGEKPVTDR
17406	EYKPQGVRR	17447	GERWPNNML	17488	HIEMRNTHM	17529	IGGMFSRNA
17407	EYKQEGVRR	17448	GERWTHNML	17489	HIEMRNVHM	17530	IGGMYSRHA
17408	EYKQKGVRR	17449	GFDAEHGYQ	17490	HIKCTHEDN	17531	IGRDIDTAF
17409	EYKQPGVRR	17450	GGNMRSSMI	17491	HIQCTHEDD	17532	IGRDVDTSF
17410	EYKQQCVRR	17451	GHIHARDGA	17492	HIQCTHKDN	17533	IHEWIANVE
17411	EYKQQGLRR	17452	GIKQTLVRR	17493	HIQCTHQDN	17534	IHEWLANVQ
17412	EYKQQSVRR	17453	GIRLKGDIM	17494	HMQCTHEDS	17535	IHGLIPMDS
17413	EYQQQGVRR	17454	GIRLRGNIM	17495	HRIYIMKKM	17536	IHKQFHSKF
17414	EYRQQGVRR	17455	GKRWPHNML	17496	HRIYIMKQM	17537	IHQWIANVQ
17415	EYTKQGVRR	17456	GLRQDRIIF	17497	HRNYIMKEM	17538	IIANHSDMA
17416	FARKKCDTW	17457	GNQPTHSPR	17498	HSAGVGKDR	17539	IIANHSHMS
17417	FCKLQKLKA	17458	GSETNNGGY	17499	HSEVKCVER	17540	IIANHSHMV
17418	FCKLRKLKG	17459	GSIIESSAR	17500	HSNGVGKDR	17541	IIKPTRAPA
17419	FCKLRKLRA	17460	GSIIKSSAR	17501	HSPTYKWKR	17542	IIVNHSHMA
17420	FCKMRKLKA	17461	GSIIQGSAR	17502	HSTGIGKDR	17543	IKHDFHSAF
17421	FCKSRKLKA	17462	GSISQSSAR	17503	HSTGLGKDR	17544	IKHDFNNAF
17422	FEAPFNLSF	17463	GSIVQSSAR	17504	HSTVVGKDR	17545	IKHDFNSAC
17423	FEAPINLGF	17464	GSKMSHKEA	17505	HTDGLFTDD	17546	IKHDFNSDF
17424	FEAPINMSF	17465	GSMIQSSAR	17506	HTKMTHGTA	17547	IKHDFNSGF
17425	FEVPINLSF	17466	GSNIQSSAR	17507	HTNCLFTDD	17548	IKHDFNSPF
17426	FFKLRKLKA	17467	GSQPTHSPR	17508	HTRSKDFWN	17549	IKHDFNSSF
17427	FGAPINLSF	17468	GTPPIVMDR	17509	HVAANYDTF	17550	IKHDINSAF

[2219]

17551	IKHDYNSAF	17592	ITNKQQLGR	17633	KNVMHSQFR	17674	LQKKRGRAC
17552	IKHNFNSAF	17593	ITNKQQLPR	17634	KNVMHSVFR	17675	LQSFRSMMD
17553	IKNAMHNFF	17594	ITNKQQLSR	17635	KPLIMWPDF	17676	LQSPLLWER
17554	IKNDFNSAF	17595	ITNKQQQAR	17636	KPLLMLPDF	17677	LRAPYNFWW
17555	IKRDFNSAF	17596	ITNKRQLAR	17637	KPMLMWPDF	17678	LRIYIMKEM
17556	IKYDFNSAF	17597	ITSKQQLAR	17638	KPRSCNFWC	17679	LRKKRGRAC
17557	ILLYSCVKD	17598	ITTKQQLSR	17639	KPVLMWPDF	17680	LRTPFNFWW
17558	IMLYSCVQD	17599	ITTQFMSKH	17640	KQAPCNSGH	17681	LRTPYDFWW
17559	INATPNLRT	17600	IVAISCYDN	17641	KQFCCCCRA	17682	LRTSYNFWW
17560	INRENRRGC	17601	IVGISCYDD	17642	KQFCCCRKA	17683	LRTTYNFWW
17561	INVDCFNAA	17602	IVGNSCYDN	17643	KQFCCCWKA	17684	LRYSFDQCM
17562	IQHDFNSAF	17603	IVGTSCYDN	17644	KQFCFCCKA	17685	LSTGVGKDR
17563	IRADLDVDH	17604	IVKDNNRRL	17645	KQFCWCCKA	17686	LSYVQGDWR
17564	IRHDFNSAF	17605	IWEKPATDR	17646	KRFCCCCKA	17687	LTAVSGFRR
17565	ISAIADLDR	17606	IWEKPVADR	17647	KSESAASDN	17688	LTFVQGDWR
17566	ISAIADMDQ	17607	IWEKPVTAR	17648	KSEVKCVER	17689	LTPVSVFRR
17567	ISAPPNLRT	17608	IWEKSVTDR	17649	KSKMSHKEA	17690	LTRCNEKGN
17568	ISAPSCWAR	17609	IWEKTVTDR	17650	KSRIMAPLW	17691	LTTGIWKNW
17569	ISDPSCWAR	17610	IYSPNRQDR	17651	KTDYCYQCL	17692	LTYAQGDWR
17570	ISEDIKDNG	17611	IYVPSCWAR	17652	KTIRDMSGF	17693	LTYVEGDWR
17571	ISEDIKENA	17612	KAEYWNCGR	17653	KVGEMGIIR	17694	LTYVKGDWR
17572	ISEDIKENS	17613	KCEMISLSR	17654	KVGEMGLLR	17695	LTYVQGNWR
17573	ISGPSCWAR	17614	KEFCCCCKA	17655	KVGEMSILR	17696	LTYVQRDWR
17574	ISIPSCWAR	17615	KFGSFWPTG	17656	KVGNEKSCI	17697	LVEFYSENL
17575	ISLPSCWAR	17616	KGDHFDRNS	17657	KVHTFFANE	17698	LVHELNQIK
17576	ISNKQQLAR	17617	KHFCCCCKA	17658	KVQPVRRFT	17699	LVHNSPFDF
17577	ISVHSCWAR	17618	KHVMHSEFR	17659	KWEKPVTDR	17700	LVKDEDNNC
17578	ISVPGCWAR	17619	KHWRHTMFL	17660	KYKQQGVRR	17701	LVKMDWKEA
17579	ISVPICWAR	17620	KKDHFDVMQ	17661	LAAVSVFRR	17702	LVNELNQIK
17580	ISVPRCWAR	17621	KKEPCNSGH	17662	LACACQDAA	17703	LVRVINERQ
17581	ISVPSCWER	17622	KLREVGRLC	17663	LAETEKLWN	17704	LYCRMHNWA
17582	ISVPSCWGR	17623	KMEWFPHNS	17664	LASANHGGN	17705	LYCRMNDWA
17583	ISVPSCWSR	17624	KMKNLKRTG	17665	LCMHTMWSR	17706	MASIFQYPR
17584	ISVPTCWAR	17625	KMRQQKMSR	17666	LFDDVMVGR	17707	MCDDLMVGR
17585	ISVSSCWAR	17626	KNAPTNQPC	17667	LFGESMVGT	17708	MCKSYERGA
17586	ISVTSCWAR	17627	KNEWIYHQG	17668	LGKVINERQ	17709	MCMHKMWSR
17587	ITNKEQLAR	17628	KNEWNYHQG	17669	LGNFIDMPA	17710	MCYGFWCEI
17588	ITNKQELAR	17629	KNEWTYHQA	17670	LHYAFDQCM	17711	MFDDLMVAR
17589	ITNKQLLAR	17630	KNEWTYHQC	17671	LKIPFNMTF	17712	MFDDLMVCR
17590	ITNKQPLAR	17631	KNVKHSEFR	17672	LKKERGRAC	17713	MFDDLMVGH
17591	ITNKQQLER	17632	KNVMHSAFR	17673	LKKKRGRSC	17714	MFDDLMVSR

[2220]

17715	MFDDVMVGR	17756	MQKAHDRAA	17797	NGSNCFWCC	17838	PRYMGTKDA
17716	MFDEVMVGR	17757	MQKEHDRSA	17798	NHAPAKNKA	17839	PSRLEHYGL
17717	MFDNLMVGR	17758	MQKPHDRSA	17799	NICCNYDAK	17840	PTKMTHGTS
17718	MFDYLMVGR	17759	MQKTHDRSA	17800	NICCNYDAP	17841	PTSFAETDY
17719	MFGDLMVGR	17760	MQKWTEIGC	17801	NKNAMHNFF	17842	PVHTFFANE
17720	MFHDCMNYR	17761	MQLKGSWNR	17802	NKNYMHCYW	17843	QAHTFFANE
17721	MFNDLMVGR	17762	MQNEKNDHP	17803	NLEMLAMFP	17844	QAISDMPFP
17722	MGIRDWKRV	17763	MQNEKNDTP	17804	NLEMLAMFR	17845	QFFWSLHAV
17723	MGIREWRRV	17764	MQNEKNENP	17805	NMRSTEMGV	17846	QFGTFWPTG
17724	MGIWKSQRH	17765	MQTKGSWNR	17806	NNAPIWLHG	17847	QGEPLMRYS
17725	MGTREWKRV	17766	MRDEQDDMP	17807	NNGVFHWKD	17848	QHQAFSMFS
17726	MHNDKLMDE	17767	MRNDKWMDE	17808	NNPYQHIEG	17849	QHQPFTMFS
17727	MHNDKWMDA	17768	MRNEKNDNP	17809	NNSVFHWKD	17850	QHSCTHMGN
17728	MHSPNRQDR	17769	MRTPYNFWW	17810	NNVKIQYMG	17851	QICPSFDAD
17729	MHYSFDQCM	17770	MSRTCADWT	17811	NNWAFHWKD	17852	QINPSFDAD
17730	MIHWDYHMK	17771	MTAVSVFRR	17812	NNWGFHWKD	17853	QIYPSFEAD
17731	MIHWDYNME	17772	MTKQNKIHR	17813	NNWVCHWKD	17854	QIYTSFDAD
17732	MIHWDYQME	17773	MTSIFQHPR	17814	NNWVFHWKA	17855	QKATTDMYW
17733	MILWDYHME	17774	MTYVQGDWR	17815	NNWVFHWKH	17856	QKSPLLWER
17734	MINWDYHME	17775	MVDGLMKEY	17816	NNWVFPWKD	17857	QKTPTDMYW
17735	MIRSETQWA	17776	MVIREWKRV	17817	NNWVIHWKD	17858	QMKNLKRTG
17736	MIWSNRKFM	17777	MVKCTEYSA	17818	NPRSKDFWN	17859	QMVYAQKNM
17737	MIWYNRKCM	17778	MVKVINERQ	17819	NQCGLKMYP	17860	QNAWIECQL
17738	MIWYNRKVM	17779	MVWYNRKFM	17820	NSAIKEEMA	17861	QNTWIQCQL
17739	MIWYNRQFM	17780	MWKCKSTYA	17821	NSAIRQEMA	17862	QNVMHSEFR
17740	MIYWDYHME	17781	MYNPNRQDR	17822	NSDMKETGD	17863	QPRSCNFWC
17741	MKDEHDDMP	17782	MYSANRQDR	17823	NSRQLHHIG	17864	QQIPLLWER
17742	MKDEQDDTP	17783	MYSPNHQDR	17824	NSRSKDFWN	17865	QQQPFSMFS
17743	MKETMDCSY	17784	MYSPNREDR	17825	NSSNCFWCC	17866	QQSPPLWER
17744	MKKAHDRSA	17785	MYSPNRHDR	17826	NTEPMWSCS	17867	QQSPQLWDR
17745	MKKKRGRAC	17786	MYSPNRKDR	17827	NTRPRRLHM	17868	QQSPQLWER
17746	MKLDYDVGC	17787	MYSPNRQAR	17828	NTRRMRLHM	17869	QQSTLLWER
17747	MKNEKNDNP	17788	MYSQNRQDR	17829	NTRSKDFWY	17870	QRHDLEVDL
17748	MKWDYDVGG	17789	MYSTNRQDR	17830	NTSNCFWCC	17871	QSIRDMSGF
17749	MMKPWACLE	17790	MYTPNRQDR	17831	NVMPLIEYR	17872	QSQVKCVER
17750	MMKPWVFLE	17791	NASNCFWCF	17832	NVMTLIQYR	17873	QSRIMAPLW
17751	MMKQWVCLE	17792	NASTCFWCC	17833	PAEINKMHC	17874	QSSFAETDY
17752	MMMMRATMD	17793	NCRCYGNNS	17834	PQAKIDMEF	17875	QTIRDMGGF
17753	MMMMRVAMD	17794	NESPYHMCL	17835	PQSPLLWER	17876	QTIRDMNGF
17754	MNTALFVNA	17795	NFKTYRFCE	17836	PQVKVDMEF	17877	QTTFAETDY
17755	MQIKVSWNR	17796	NFVQYVHNY	17837	PRRLLHFSQ	17878	QTWAAHCTR

[2221]

17879	QTWSAHCSR	17920	SFKWPIVNG	17961	SSYFMEVTR	18002	TFFVTASWS
17880	QTWTAHCSR	17921	SFKWTILNG	17962	SSYFMKITR	18003	TFHQCERSV
17881	QYVCFNCYH	17922	SFKWTIVDG	17963	SSYFTEITR	18004	TFKNYRFCE
17882	RACIPKFKE	17923	SFKWTIVNA	17964	SSYIMEITR	18005	TFKTYRYCE
17883	RACVPKFKQ	17924	SFKWTIVND	17965	SSYVMEITR	18006	TFPNPVHRG
17884	RAEILQLQQ	17925	SFKWTIVNV	17966	SSYYMEITR	18007	TFRTYRFCE
17885	RARIPKFKQ	17926	SFKWTIVTG	17967	STDDKFWLG	18008	TFSNPVHRV
17886	RASIFQHPR	17927	SGKDNNRRL	17968	STDNRFWLG	18009	TFVNWPIHG
17887	RCDEWAEDQ	17928	SHIHKDSCW	17969	STDTKFWLG	18010	TKIFTENDQ
17888	RCDEWAKDK	17929	SIMKYDEMG	17970	STEAAHENP	18011	TKNAMHHFF
17889	RCDEWAKDR	17930	SIMKYDQMV	17971	STEASYENP	18012	TKNDMHNFF
17890	RCDEWAKNQ	17931	SIMLNECNT	17972	STESAYENP	18013	TKNGMHNFF
17891	RCDEWARDQ	17932	SIRLKGNIM	17973	STQQVFWNL	18014	TKNPMHNFF
17892	RCDEWVKDQ	17933	SKISATWDR	17974	STRPMRLHM	18015	TKTAMHNFF
17893	RCVPWMNQW	17934	SKLHKQYCC	17975	STSILDKLI	18016	TLTTEMIWA
17894	RDAQFLPDQ	17935	SKYDTDSAC	17976	STSSLDKLV	18017	TNWVFHWKD
17895	RFDDLMVGR	17936	SLEMRPRQI	17977	STTAYNLTR	18018	TPEAEHHKI
17896	RGDEWAKDQ	17937	SLKDNNRRL	17978	SVHAKMEGC	18019	TRGMRDFCY
17897	RGEILELQQ	17938	SLKMRPRKI	17979	SVHSKMQGC	18020	TRMVLEWTY
17898	RGEILQLQR	17939	SLKMRPRQS	17980	SVHSTMEGC	18021	TSEVFTVNC
17899	RHQPCECYE	17940	SLKMRPRQT	17981	SVKDINRRL	18022	TSQYVNREI
17900	RHQPFSMFS	17941	SLQMRPRQI	17982	SVKDNNGRL	18023	TSVPSCWAR
17901	RHWRHAMFL	17942	SLTASNWRD	17983	SVKDNNQRL	18024	TTEQWDVCR
17902	RHWRLTMFL	17943	SMVCQVIVF	17984	SVKDNNRIL	18025	TTEVEHHKI
17903	RIMVHEWGD	17944	SNDKNIHNG	17985	SVKDNNRKL	18026	TVKCTGYSA
17904	RLDHAETRA	17945	SNWVFHWKD	17986	SVKDNNRTL	18027	TVKDENAAC
17905	RMLAYKQRH	17946	SPREKNDYW	17987	SVKDTNRRL	18028	TVKEENAAF
17906	RVDHAESRA	17947	SPREKNRYW	17988	SVKNNNRRL	18029	TVKEENATC
17907	RVVDLVDVY	17948	SRIHKDICW	17989	SVMQSIDSS	18030	TVKEENSAC
17908	RWEKPVTDR	17949	SRINKDSCW	17990	SVNSKMEGC	18031	TVKEENTAC
17909	SAEAAYENP	17950	SRLHKDSCW	17991	SVRDNNRRL	18032	TVRCTEYSA
17910	SDCSYSDWS	17951	SRTMKDWWG	17992	SWAASNWRD	18033	TWKMMHCDN
17911	SEVHLEHNY	17952	SSAIKQEMA	17993	SWKMMHCDN	18034	VARPWAMNS
17912	SFCCDHNAG	17953	SSEQTGPHM	17994	SWTSSNWRD	18035	VATQSCKWM
17913	SFCCDHNAS	17954	SSHFMEITR	17995	SWTTSNWRD	18036	VAVTITDFD
17914	SFCCDHNGA	17955	SSIIQSSAR	17996	SWTVSNWRD	18037	VAVTVRDFD
17915	SFDERGVWC	17956	SSKQAGPHM	17997	SYQMFNNVQ	18038	VCKWMGRQQ
17916	SFGSCNTCG	17957	SSRPRMKQP	17998	TAQYLNREI	18039	VCSTVMDFN
17917	SFHDEEGLF	17958	SSTASNWRD	17999	TAREDYQHA	18040	VEADNNDAD
17918	SFKATRRGV	17959	SSYFMAITR	18000	TARLLEKYN	18041	VFNQQEMWA
17919	SFKLTIVNG	17960	SSYFMEIAR	18001	TCMGFTSFP	18042	VGKWMGKQQ

[2222]

18043	VGVTVTDFD	18062	VSKPIEVGY	18081	VVGGDKAAY	18100	WVDKQATDR
18044	VHALIPMDS	18063	VSMSSIYDK	18082	VVKVINERQ	18101	YCRCYGNNS
18045	VHGLIPMDT	18064	VSMSSNSDK	18083	VVSQTYRNR	18102	YIMVHEWGD
18046	VHRQMHDSY	18065	VSNQCCKAM	18084	VWKQYILVI	18103	YMILQHDAL
18047	VICATQDAA	18066	VSRALERQA	18085	VWMQHMCLD	18104	YMTLEHDAL
18048	VICATQDPS	18067	VSRSLERKA	18086	VWQQYILII	18105	YQSGGILRM
18049	VICPTQDPA	18068	VSRSLERQG	18087	WCRQTMMDL	18106	YQSGGIRRM
18050	VICTTQDPA	18069	VSRSLERQT	18088	WDAMDDKKA	18107	YQTGGIHRM
18051	VIHDQLETM	18070	VSRSLKRQA	18089	WETDMDMDA	18108	YSHTYKWKR
18052	VINMNRRWS	18071	VSRSLQRQA	18090	WFRRHDYFY	18109	YSLTYKWKR
18053	VKKKRGRAC	18072	VSRSQERQA	18091	WFRRHSYFY	18110	YSPTYMWKR
18054	VKNYMHCYW	18073	VSVPSCWAR	18092	WGDKKATDR	18111	YSQLFTFFE
18055	VKSFRSMMD	18074	VSVTVTDFD	18093	WGDKMAADA	18112	YSSTYKWKR
18056	VMRWPWDDT	18075	VTEHQQRGV	18094	WGDKPATDR	18113	YSTTYGGRH
18057	VPRSCNFWC	18076	VTKCNEKGN	18095	WGDKRATDR	18114	YVTEVNSMI
18058	VRTPYNFWW	18077	VTNKQQLAR	18096	WKNDMDMDA	18115	YWVDFVMPF
18059	VSETNNGGY	18078	VTNQCCKTM	18097	WKTEMDMDA	18116	YWVNFVVPF
18060	VSINTNQPG	18079	VTSQCCKAM	18098	WNCVMPDNA	18117	YWVTFVMPF
18061	VSKMSHKEA	18080	VTTGIWKTW	18099	WNTDMDMDA		

[2223] 实施例38

[2224] 如通过机器学习鉴定的具有坐骨神经组织嗜性的AAV5变体

[2225] 此实施例描述了如通过机器学习方法鉴定的展现坐骨神经组织嗜性的工程化AAV5变体的优选特性。进行如实施例19和图19至图22中描述的相同生物信息学预处理、机器学习方法和分析以确定展现坐骨神经组织嗜性的工程化AAV5变体的优选特性。

[2226] 下文描述了AAV5 VP衣壳多肽的581至589区域中各位置处的有利生物物理特性和有利氨基酸残基，它们与较高的坐骨神经组织嗜性概率相关。表60中描述的任何以下机器学习衍生的位置氨基酸偏好可以单独存在或彼此组合存在或与实施例8，表15中观察和描述的任何氨基酸偏好组合存在。下文表60中列出了从ML模型中提取的排名靠前的重要坐骨神经嗜性特征之间共有的位置特征的概述。

[2227]

表 60 机器学习衍生的坐骨神经组织嗜性规则
Xaa1 处高平均柔性 Xaa1 选自 G 或 R
Xaa1 处低溶解性 Xaa1 选自 R 或 Q
Xaa1 处低可突变性 Xaa1 选自 C、L、F、Y、R、K、P 或 H
Xaa1 处高体积 Xaa1 选自 Y 或 F
Xaa2 处高表面可接近性 Xaa2 选自 E、R 或 K
Xaa3 处中等可突变性 Xaa3 选自 H 或 R
Xaa3 处中等平均柔性 Xaa3 选自 V 或 Y
Xaa4 处高可突变性 Xaa4 选自 N
Xaa4 处高平均柔性 Xaa4 选自 I、N、G 或 R
Xaa4 处低溶解性 Xaa4 选自 N
Xaa6 处低可突变性 Xaa6 选自 C、L、F 或 Y
Xaa6 处高体积 Xaa6 选自 K、M、I 或 L
Xaa7 处低可突变性 Xaa7 选自 L、F 或 Y

[2228]

表60 机器学习衍生的坐骨神经组织嗜性规则
Xaa7处中等摩尔质量 Xaa7选自D、I、L或N
Xaa8处高表面可接近性 Xaa8选自S、Y、T、D、P、H或N
Xaa9处低可突变性 Xaa9选自C、H或R
Xaa9处中等溶解性 Xaa9选自Q、T或C
Xaa9处低表面可接近性 Xaa9选自C

[2229] 下文表61提供了表现出坐骨神经组织嗜性并且符合表60中提供的一种或多种规则的变体AAV VP1衣壳多肽的581至589区域的序列。因此，本公开提供了由具有SEQ ID NO:2的工程化AAV5 VP1衣壳多肽组成的rAAV，其中Xaa1至Xaa9区域是SEQ ID NO:26991-SEQ ID NO:27990中的任一者，如表61中所公开。本文还涵盖由工程化AAV5 VP2衣壳多肽和工程化AAV5 VP3衣壳多肽组成的rAAV，所述多肽在对应于AAV5 VP1 581至589区域中的氨基酸的AAV5VP2(氨基酸残基445至453)和AAV5 VP3(氨基酸残基389-397)中的区域处具有下文表中公开的序列。

[2230] 表61

[2231] 驱动坐骨神经组织嗜性的AAV5 VP1衣壳多肽中581至589区域的序列

[2232]

SEQ ID NO	581-589 序列
26991	AAYLYNVPM
26992	AIRIAMCDW
26993	AIRTAMCDW
26994	AIRTTMCHW
26995	ANGCAIEYH
26996	ANGCTIENH
26997	ANGGAIENH
26998	ANGGAIENN
26999	ARYRMRYHG
27000	ASGCAIENH
27001	ATFGWMNWF
27002	ATSRQAVWA
27003	ATYLYNIPM
27004	AVLLWCEVQ
27005	AYPRNAKYI
27006	AYSRNAKYT
27007	CADWPCMSL
27008	CAGWACMSL
27009	CAKFNNAIH
27010	CAKFTNAIY
27011	CAKFTNTIH
27012	CAKFTTAIH
27013	CANWACMSL
27014	CCCEALMPK
27015	CCCEALMSQ
27016	CCCEAMMPQ
27017	CCCEAVMPQ
27018	CCCEVLMPQ
27019	CCCPDATNN
27020	CCCSDATNN
27021	CCGQVEFGW
27022	CCSGALMPQ
27023	CDGQVEFGW
27024	CEDPLCQSS
27025	CEDWACMSL
27026	CEHPLCQSS
27027	CEIPLCQSS
27028	CENPLCKSS
27029	CENPLCPSS
27030	CENPLCQSN
27031	CENPLGQSS
27032	CENPLRQSS
27033	CENPLWQSS
27034	CENPQCQSS
27035	CENSLCQSS
27036	CETPLCQSS
27037	CEVNEEYDW
27038	CFAEPITDC
27039	CFGAEGYAG
27040	CFNEPITDC
27041	CFTAPITDC
27042	CFTEPIADC
27043	CFTEPINDC
27044	CFTEPITDR
27045	CFTEPITNC
27046	CFTESITDC
27047	CFTETITDC
27048	CGGKVEFGW
27049	CGGQAEFGW
27050	CGGQFEFGW
27051	CGGQIEFGW
27052	CGGQIEVGW
27053	CGGQVEAGW
27054	CGGQVECGW
27055	CGGQVEFDW
27056	CGGQVEFSL
27057	CGGQVEFSW
27058	CGGQVEFVW
27059	CGGQVEIGW
27060	CGGQVKFGW
27061	CGGQVQFGW
27062	CGKFTNAIH
27063	CGVQVEFGW
27064	CHPQHHYSH
27065	CHPQHRHSH
27066	CHPQLHHSH
27067	CKDTRKMRE
27068	CKEHRKMRE
27069	CKEPGKMRE
27070	CKEPRKMRA
27071	CKEPRKMRK
27072	CKEPRQMRE
27073	CKEPRRMRE
27074	CKETRKMRE
27075	CKKPRKMRE
27076	CKVNEEYDW
27077	CQIRFDRHH
27078	CQNPLCQSS
27079	CQVNAEYDW
27080	CQVNEAYDW
27081	CQVNEEDDW
27082	CQVNEEFDW
27083	CQVNEEYNW
27084	CRCKDAWNM
27085	CRCNAAWNM
27086	CRCNDAGNM
27087	CRCNDAWHR
27088	CRCNDAWNK
27089	CRCNDAWNR
27090	CRCNDAWNV
27091	CRCNDAWTM
27092	CRCNDEWNM
27093	CRCNEAWNM
27094	CRCNGAWNM
27095	CRCSDAWNM
27096	CRFNDAWNM
27097	CRGQVEFGW
27098	CRVNEEYDW
27099	CSSCNHRVF
27100	CVFGWNNYQ
27101	CVGQVEFDW
27102	CVGQVEFGW
27103	CVKFTNAIH
27104	CVWSVGTEW
27105	CWAADSKPC
27106	CYPQHHHSH
27107	DDLDAWQTS
27108	DDLDLWQTS
27109	DDLDVGQTS
27110	DDLDVWETS
27111	DDLDVWKTS
27112	DDLDVWRTS
27113	DDMDVWQTS
27114	DDQDVWQTS
27115	DDRDVWQTS
27116	DFEMWDWVG
27117	DFEVWDWFG
27118	DFNHIGWWQ
27119	DFSHIAWWQ
27120	DFTHIAWWQ
27121	DFYFSGFLI
27122	DFYISGFLI
27123	DFYLSGFLI
27124	DPADKPSNS
27125	DPQDKPSNS
27126	DRIMISIQP
27127	DRVMISIQQ
27128	DRVMVSIQP
27129	DVMMPALKV
27130	DVMMRGLKV
27131	DVMMSGLKV
27132	DWRMVCVHC
27133	EACLYNIPM
27134	EAYFYNIPM
27135	EAYLCNIPM
27136	EAYLYHIPM
27137	EAYLYMIPM
27138	EAYLYNFPM
27139	EAYLYNIRM

[2233]

27140	EAYLYNISM	27181	FHVRFDRHH	27222	FPPIVKRHY	27263	FSIRFDRHH
27141	EAYLYNVPM	27182	FIAANQNHM	27223	FPPTVKRHH	27264	FSPGPLAYN
27142	EAYLYSIPM	27183	FKEPRKMRE	27224	FQIGFDRHH	27265	FSTGPLSYN
27143	EAYMYNIPM	27184	FKIRFDRHH	27225	FQIKFDRHH	27266	FSVGPLAYN
27144	EEYLYNIPM	27185	FKMDCGSCG	27226	FQILFDRHH	27267	FTPIVKRHH
27145	EFKRMTCNR	27186	FKMDCHGCG	27227	FQIPFDRHH	27268	FTPQKDNSC
27146	EGSRKWFAC	27187	FKMDCHSCC	27228	FQIQFDRHH	27269	FTPQKDTSC
27147	EGYLYNIPM	27188	FKMDCHSFG	27229	FQIQIDRHH	27270	FTQEKDTSC
27148	EIGIMRHGH	27189	FKMDCHSGG	27230	FQIRFARHH	27271	FTQQKDASC
27149	EIILDDEQR	27190	FKMDCHTCG	27231	FQIRFDGHH	27272	FVAANQNNM
27150	EIILDDQKR	27191	FKMDGHSCG	27232	FQIRFDRDH	27273	GAHNELVGY
27151	EIIVDDQQR	27192	FKMDWHSCG	27233	FQIRFDRHC	27274	GASRQAVWA
27152	EIIVDEQQR	27193	FKMGCHSCG	27234	FQIRFDRHD	27275	GAYLYNIPM
27153	EIMLDDQQR	27194	FKMVCHSCG	27235	FQIRFDRHR	27276	GAYLYNLPM
27154	EIVLDDQQR	27195	FKRDCHSCG	27236	FQIRFDRNR	27277	GCCSDATNN
27155	EKMIDTCLQ	27196	FKRICLEAG	27237	FQIRFDRPH	27278	GEVVVCILR
27156	EKMRHTCLQ	27197	FKRICWEAA	27238	FQIRFDRRQ	27279	GFAADSKPC
27157	EKMTHPCLQ	27198	FKRICWEGG	27239	FQIRFDRYH	27280	GFNHIAWWQ
27158	EKMTHTCLE	27199	FKRICWEPG	27240	FQIRFGLHH	27281	GFTGMSDQD
27159	EKMTHTCLK	27200	FKRICWEVG	27241	FQIRFHRHH	27282	GFYLTREKK
27160	EKMTHTCLR	27201	FKRICWGAG	27242	FQIRFNRHH	27283	GGGQVEIGW
27161	EKMTHTCVQ	27202	FKRICWQAG	27243	FQIRFYRHH	27284	GIANDDDRE
27162	EKMTHTCVR	27203	FKRMCWEAG	27244	FQLRFDRHH	27285	GIANDDDWA
27163	EKMTHTWLQ	27204	FKRTCWEAG	27245	FQMRFDRHQ	27286	GIANDDDWE
27164	EPYLYNIPM	27205	FKRVCWEAG	27246	FQSRFDRHH	27287	GIANLDDWE
27165	EQMTHTCLP	27206	FKVDCHSCG	27247	FQSRFDRHR	27288	GIANVADWA
27166	EQMTHTCLQ	27207	FKVNRDDSW	27248	FQTRFDRHH	27289	GIANVADWE
27167	ERMTHTCLR	27208	FKVRFDRHH	27249	FQVRFDRHH	27290	GIANVDAWV
27168	ESHVRPLNM	27209	FLIRFDRHH	27250	FRIRFDRHH	27291	GIANVDDLE
27169	ETMKHKKRE	27210	FNAGPLAYN	27251	FRIRFDRHY	27292	GIANVDDWK
27170	ETMQYKKRE	27211	FNLPFCYHQ	27252	FSAGALAYN	27293	GIANVDDWR
27171	EWTRRMAAP	27212	FNLSFCYHL	27253	FSAGGLAYN	27294	GIANVDDWT
27172	EWTRRMEAS	27213	FNLSICYHQ	27254	FSAGPLAYC	27295	GIANVDEWK
27173	EWTRRMEAT	27214	FPGVMEAEI	27255	FSAGPLAYY	27296	GIANVDHWE
27174	EWTRRMEDP	27215	FPGVMQGEI	27256	FSAGPLDYN	27297	GIANVDNWE
27175	FAAANQNHM	27216	FPHIVKRHH	27257	FSAGPLSYI	27298	GIANVDYWE
27176	FAPIVKRHH	27217	FPIRFDRHH	27258	FSAGPLVYN	27299	GIANVEAWE
27177	FDLSFCYHQ	27218	FPIRFDRHL	27259	FSAGPMDYN	27300	GIANVGDWG
27178	FEKNSSQMR	27219	FPIRFDRHR	27260	FSAGSLAYN	27301	GIDNFDDWE
27179	FGGQVELGW	27220	FPIRFDRPH	27261	FSAGTLAYS	27302	GIDNVDDLE
27180	FHPYFPTYK	27221	FPPIVKQHH	27262	FSDGPLAYN	27303	GIDNVDDWG

[2234]

27304	GIDNVDEWE	27345	GWAADSKSC	27386	IWGIYVHYC	27427	KSRRGKLKW
27305	GIDNVDGWE	27346	GWAADSKTC	27387	IWGIYVLYF	27428	KSRRGKLPL
27306	GIDNVDYWE	27347	GWAADSRPC	27388	IWGIYVYYF	27429	KSRRGKLPW
27307	GIGNVDDWE	27348	GWAADSTPC	27389	KAGCFVDVW	27430	KSRRGKLRW
27308	GIPNVDGWE	27349	GWAADTKPC	27390	KARFHRERS	27431	KSRRGKLSW
27309	GISNVDDWE	27350	GWAAEAKPC	27391	KDLDVWQTS	27432	KSRRGKLTL
27310	GISNVDDWG	27351	GWAAGSTPC	27392	KDRQIDQYH	27433	KSRRGKMTW
27311	GISRQAVWA	27352	GWAANSKPC	27393	KENNTNRID	27434	KSRRGKRTW
27312	GITNVDDWE	27353	GWADDSKPC	27394	KGAENIHWG	27435	KSRRGKVTW
27313	GITNVDGWE	27354	GWADDSKTC	27395	KGAPWAYNC	27436	KSRRVKLTW
27314	GIVNVDDWE	27355	GWAGDSKPC	27396	KGDAYIHWG	27437	KSWGWVDTA
27315	GIWRSREHY	27356	GWASDSKPC	27397	KGDENIHWC	27438	KTGHFNQRG
27316	GIWRTREHH	27357	GWATDSKPC	27398	KGDEYIHWA	27439	KTRFHREPS
27317	GKANDDDWE	27358	GWAVDSKPC	27399	KGDEYIHWC	27440	KTRFHRERN
27318	GKANVDDRE	27359	GWAVNSKPC	27400	KGDEYIQGG	27441	KTRFHRERT
27319	GLPNVDDWE	27360	GWDDDSKPC	27401	KGDEYIRWG	27442	KTRRGRLTW
27320	GNWRTREHY	27361	GWGVDSKPC	27402	KGDEYIYWG	27443	KTWFHRERS
27321	GRANVDDCE	27362	GWHQKCMGD	27403	KGDEYVHWG	27444	KVGCSVDVW
27322	GRCNDAWNM	27363	GWSADSKPC	27404	KGDGYIHWG	27445	KVGCVVDVW
27323	GSHARPLNM	27364	GWTADSKPC	27405	KGDVHIHWG	27446	LEKNGSQMR
27324	GSHVGPLNM	27365	GWYKKCMGD	27406	KGDVYIHWG	27447	LEKNSAQMR
27325	GSHVGTLNM	27366	GWYQKCMGN	27407	KGGAYIHWG	27448	LEKNSSEMR
27326	GSHVQPLNM	27367	GWYQKCVGD	27408	KGGCFVDVW	27449	LEKNSSKMR
27327	GSHVRPLHM	27368	GWYQKYMGD	27409	KGGEYIHWC	27450	LEKNSSLMR
27328	GSHVRPLSM	27369	HFNINIYNC	27410	KGGLWAYNC	27451	LEKNSSPMR
27329	GSHVRSLNM	27370	HFYVSGFLI	27411	KGGPLAYNC	27452	LEKNSSQMK
27330	GTANVDDLE	27371	HGSRNWFAC	27412	KGGPWAHNC	27453	LEKNSSQTR
27331	GTHNELVGY	27372	HLACLKMEM	27413	KGGPWGYNC	27454	LEKNSSQVR
27332	GTSRQALWA	27373	HLDCLKMQM	27414	KGGQWAYNC	27455	LEKNTSQMR
27333	GTSRQAVWS	27374	HLDCLQMEM	27415	KGGRWAYNC	27456	LEQNSSQMR
27334	GTSRRAVWA	27375	HLDCLRMEM	27416	KGGVYIHWG	27457	LERNSSQMR
27335	GVTDLQNIC	27376	HLNCLKMEM	27417	KGHEYIHWG	27458	LLDCLKMEM
27336	GVVNVDDWE	27377	HMDCLKMEM	27418	KGNEYIHWG	27459	LMENHEFQD
27337	GVYRSMFQE	27378	HQDCLKMEM	27419	KGPEYIHWG	27460	LQDNINYSY
27338	GWAADAKPC	27379	HWRMACVHC	27420	KGTEYIHWG	27461	LYLPLCHRI
27339	GWAADSEHC	27380	HWRMGCVHC	27421	KGVAYIHWG	27462	MASIDKQNV
27340	GWAADSEPC	27381	HWRMVCVHG	27422	KPRFHRERS	27463	MCTGELDGT
27341	GWAADSKAC	27382	HWRMVCVYC	27423	KSGRGKLTW	27464	MCTGELSGT
27342	GWAADSKHC	27383	IRDWQAYWL	27424	KSRLGKLTW	27465	MCTGEWNGT
27343	GWAADSKLC	27384	IRDWQCYWL	27425	KSRRAKLTW	27466	MCTGKLNGT
27344	GWAADSKRC	27385	IRDWQGHWL	27426	KSRRGKLAW	27467	MCTRELNGT

[2235]

27468	MEKNSSQMR	27509	NFYISGFLI	27550	QDSRNWFAC	27591	REVKVCILR
27469	MEMTQEWMA	27510	NFYLGGFLI	27551	QGGRNWFAC	27592	REVLVCILK
27470	MEQNLQLTM	27511	NFYLIGFLI	27552	QGSGNWFAC	27593	REVMACILR
27471	MFTGELNGT	27512	NFYLSGFLM	27553	QGSRNWFGC	27594	REVMCCILR
27472	MICGTHDNV	27513	NFYLSGFVI	27554	QGSRNWFSC	27595	REVMFCILR
27473	MIFGTHDHV	27514	NFYLTGFLI	27555	QGSRNWYAC	27596	REVMGCILR
27474	MIFGTHDNF	27515	NFYLTGYLI	27556	QGTRNWFAC	27597	REVMGCMLR
27475	MIFGTHDNI	27516	NFYVSGFLI	27557	QIGRFAMAH	27598	REVMICILR
27476	MIIGTHDNV	27517	NKIDVQWLQ	27558	QIGRFAMPH	27599	REVMLCILR
27477	MMVTKEFNI	27518	NMECEKFLT	27559	QIGRFAMYH	27600	REVMVCIII
27478	MMVTKEFNL	27519	NMECERFLT	27560	QIGRFTMSH	27601	REVMVCIIK
27479	MNLCYSSMM	27520	NMECQQFLT	27561	QNSRHRYAW	27602	REVMVCIIR
27480	MQIRFDRHH	27521	NQIDVQWIQ	27562	QSRRGKLTW	27603	REVMVCIIT
27481	MQMAQEWMA	27522	NQIDVQWLK	27563	QSSRNWFAC	27604	REVMVCILG
27482	MQMMQMCVP	27523	NQLDVQWLQ	27564	QTLGWAMGT	27605	REVMVCILK
27483	MQMTQEWMG	27524	NQPCINTYW	27565	RAEGWFKYD	27606	REVMVCIPR
27484	MQMTQKWMA	27525	NQPCVNTCW	27566	RAHNGLVGY	27607	REVMVCIRR
27485	MTAYFKALC	27526	NSRRGRLTW	27567	RANMCDCFE	27608	REVMVCIVI
27486	MWTGELNGT	27527	NVMMPGLKV	27568	RANNWTRRH	27609	REVMVCIVR
27487	NDLDVWQTS	27528	PELLWCEVQ	27569	RASICDYFE	27610	REVMVCLLR
27488	NELLWCEVQ	27529	PEVMVCILR	27570	RASMCDFFE	27611	REVMVCTLR
27489	NEMLWCEVQ	27530	PIFGWMNWF	27571	RASMCDGFE	27612	REVMVCVLK
27490	NFCHSGFLI	27531	PKYAWSTSP	27572	RASMCDYFE	27613	REVMVCVLR
27491	NFDLSGFLI	27532	PKYHWHMYH	27573	RASMWDCFE	27614	RFIGMSDQD
27492	NFEMWDWFG	27533	PRLQIRKYL	27574	RCHNELVGY	27615	RFNGMSDQD
27493	NFHINIYNC	27534	PSCVTQESI	27575	RDSMCDCFE	27616	RFPGMSDQD
27494	NFHLSGFLI	27535	PSTTVMNVM	27576	RDVIVCILR	27617	RFTGMGDQD
27495	NFNIHIYNC	27536	PTFGWMDWF	27577	REAMVCILR	27618	RFTGMNDQD
27496	NFNINIFNC	27537	PTFGWMNLF	27578	REDMACILR	27619	RFTGMSAQD
27497	NFNINISNC	27538	PTFGWMNWV	27579	REDMGCILR	27620	RFTGMSDED
27498	NFNINTYNC	27539	PTFGWMNWY	27580	REDMLCILR	27621	RFTGMSDKD
27499	NFNINVYNC	27540	PTFVWMNWF	27581	REDMVCMLR	27622	RFTGMSDKV
27500	NFNITIYNC	27541	PTYGWMNWF	27582	REDMVGILR	27623	RFTGMSDQY
27501	NFNLSGFLI	27542	PWGDMKDPQ	27583	REDTVCILR	27624	RFTGMSDRD
27502	NFNMNIYNC	27543	PWGDMKDSH	27584	REFMVCILR	27625	RFTGMSGQD
27503	NFSINIYNC	27544	PWGVMKDPH	27585	REGMLCILR	27626	RFTGMSHQD
27504	NFSTNIYNC	27545	PWGYMKDPH	27586	REGMVCILR	27627	RFTGMSYQD
27505	NFTINIYNC	27546	QAYLYNIPM	27587	REGVVCILR	27628	RFTGMTDQD
27506	NFYCSGFLI	27547	QCRLHMQCP	27588	REIMVCILR	27629	RFTGVSDQD
27507	NFYFSGFLL	27548	QCRQSHEQE	27589	RELMVCILR	27630	RGHNGKGPM
27508	NFYHSGFLI	27549	QCSRNWFAC	27590	RETMVCILR	27631	RGSMCDCFE

[2236]

27632	RGVMQQMNH	27673	RTHNEIVGS	27714	RTHNWLVGY	27755	RYGNVAHHA
27633	RIANVDEWE	27674	RTHNEIVGY	27715	RTHSELVGY	27756	RYSNVAHYA
27634	RIHNELVGY	27675	RTHNELFGH	27716	RTHSQLVGY	27757	RYTGMSDQD
27635	RIPNVDDWE	27676	RTHNELFGS	27717	RTHTKLVGY	27758	SATNWECSK
27636	RLDCLKMEM	27677	RTHNELFGY	27718	RTMGQFYYE	27759	SCCADATNN
27637	RMSGMLCTV	27678	RTHNELGGY	27719	RTNNEIVGY	27760	SCCSDAANK
27638	RMSGMLFTL	27679	RTHNELHGY	27720	RTNNELAGY	27761	SCCSDAANN
27639	RMSGMVCTL	27680	RTHNELIGY	27721	RTNNELIGY	27762	SCCSDANNN
27640	RNHNEIVGY	27681	RTHNELLCY	27722	RTNNELVGS	27763	SCCSDATDN
27641	RNHNELVGC	27682	RTHNELLGF	27723	RTPNALVGY	27764	SCCSDATNH
27642	RNHNELVGY	27683	RTHNELVAH	27724	RTQGWFKYD	27765	SCCSDATQN
27643	RNSRHKYAW	27684	RTHNELVCD	27725	RTQNELVGS	27766	SCCSDATSN
27644	RPDLNEIVC	27685	RTHNELVCH	27726	RTQNELVVY	27767	SCCYDATNN
27645	RPDLNEVEC	27686	RTHNELVCS	27727	RTQNGLVGY	27768	SCTTRVWIC
27646	RPGLNEIEC	27687	RTHNELVCY	27728	RTRNELFGY	27769	SCWSDATNN
27647	RPHNELVGY	27688	RTHNELVFY	27729	RTRNELLGY	27770	SCYSDATNN
27648	RPHNGLVGY	27689	RTHNELVGE	27730	RTRNELVGY	27771	SEVMVCVLR
27649	RSHNGLVGY	27690	RTHNELVGF	27731	RTRNELVVY	27772	SEYHWHMYH
27650	RSHVQPLNM	27691	RTHNELVGG	27732	RTRNKLVGY	27773	SFCSDATNN
27651	RSRRGKLTW	27692	RTHNELVGH	27733	RTRSELVGY	27774	SFCSDATNS
27652	RSRRGKMTW	27693	RTHNELVGQ	27734	RTYNELVGY	27775	SFYLSGFLI
27653	RSSMCDCFE	27694	RTHNELVGV	27735	RVGMQQMNH	27776	SKCHWHMYH
27654	RSWRGKLTW	27695	RTHNELVIY	27736	RVMAEMVWE	27777	SKFAWSTSP
27655	RTDNEIVGY	27696	RTHNELVRH	27737	RVSGMLCTL	27778	SKFHWHMYH
27656	RTDNELVGD	27697	RTHNELVRY	27738	RVVMQPMNH	27779	SKYAWSTAP
27657	RTDNGLVGY	27698	RTHNELVSF	27739	RVVMVCILR	27780	SKYAWSTSQ
27658	RTEGWFEYD	27699	RTHNELVSY	27740	RWAADSKPC	27781	SKYAWSTSS
27659	RTEGWFKSD	27700	RTHNELVVF	27741	RWDNGKGPM	27782	SKYAWSTTP
27660	RTEGWFKYA	27701	RTHNELVVY	27742	RWGDMKDPH	27783	SKYDWSTSP
27661	RTEGWFKYE	27702	RTHNEPVGY	27743	RWHNGKAPM	27784	SKYGWSTSP
27662	RTEGWVKYD	27703	RTHNEVVGC	27744	RWHNGKEPM	27785	SKYHWDMYH
27663	RTHIELVGY	27704	RTHNEVVGD	27745	RWHNGKGHM	27786	SKYHWHIYP
27664	RTHNALVCY	27705	RTHNEVVGS	27746	RWHNGKGPK	27787	SKYHWHMYL
27665	RTHNALVSY	27706	RTHNGLVGF	27747	RWHNGKGRM	27788	SKYHWHMYR
27666	RTHNALVVY	27707	RTHNGLVGS	27748	RWHNGKGSM	27789	SKYHWNMYH
27667	RTHNDLVGH	27708	RTHNKLGGY	27749	RWHNGKVPM	27790	SKYHWPMYH
27668	RTHNDLVVY	27709	RTHNKLVGN	27750	RWHNGQGPM	27791	SKYHWQIYH
27669	RTHNEFVGY	27710	RTHNKLVGY	27751	RWHNGRGPM	27792	SKYHWRMYH
27670	RTHNEIFGY	27711	RTHNQLVGC	27752	RWHTGKGPM	27793	SKYLWHMYH
27671	RTHNEIIGY	27712	RTHNQLVGY	27753	RWNNGKGPM	27794	SKYNWHMYH
27672	RTHNEIVGF	27713	RTHNVLVGS	27754	RWRMVCVHC	27795	SKYNWHMYQ

[2237]

27796	SKYPWSTSP	27837	TELLWCKVQ	27878	TSSHDEAHR	27919	VQDTINYSY
27797	SKYQWHMYH	27838	TELLWCQVQ	27879	TTFGWMNWF	27920	VQDTVNYSY
27798	SKYRWHMYH	27839	TELLWCVVQ	27880	TTGINRHCC	27921	VQENINYSY
27799	SKYSWSTSP	27840	TELLWFEVQ	27881	TTMPDAPCR	27922	VQGRINYSY
27800	SKYTWSTSP	27841	TELLWGEVQ	27882	TTSHDEAHK	27923	VQGTINYSY
27801	SKYVWSTSP	27842	TELLWREVQ	27883	TVLLWCDVQ	27924	VQNKINYSY
27802	SKYYWHMYH	27843	TELLWSEIQ	27884	TVLLWCEVH	27925	VQVKINYSY
27803	SPTNWECSN	27844	TELSWCEVQ	27885	TVLLWCEVQ	27926	VQVRINYSY
27804	SRYAWSTSQ	27845	TELVWCEVK	27886	TWEIQIDCI	27927	VQVTINYSY
27805	SRYHWHMYH	27846	TEMLWCEVK	27887	TWETQIDCI	27928	VRDKINYSY
27806	STFGWMNWF	27847	TEMLWCEVL	27888	TWGSQIDCI	27929	VRDWQGYWL
27807	STMTDAPCG	27848	TEMLWCEVQ	27889	TWQSQIDCI	27930	VSRRGKLTW
27808	STTNWECSN	27849	TERLWCEVL	27890	TYPRNAKYT	27931	VTMPDAPCG
27809	SVCSDATNN	27850	TFNINIYNC	27891	VEDKINYSY	27932	VWEANPPFR
27810	SVMMSPHLM	27851	TFWMPDDNC	27892	VEKNSSQMR	27933	VWEDNAPFR
27811	SWAADSKPC	27852	TFWMPGNNC	27893	VGYRSMFQE	27934	VWEDNPPVR
27812	SYARKAYDC	27853	TFYLSGFLI	27894	VIGIIRHGH	27935	VWENNPPFR
27813	SYPRRVYDC	27854	TGLLWCEVQ	27895	VISNVDDWE	27936	WEKNSSQMR
27814	TALLWCEVL	27855	TISHDEAHK	27896	VLYRSMFQE	27937	WEVMVCILR
27815	TAYMTAMYC	27856	TKLLWCEVK	27897	VQAQINYSY	27938	WHAMVNSGP
27816	TCCSDATNN	27857	TKYAWSTSP	27898	VQDEINYSY	27939	WHDMGNSGP
27817	TDLLWCELQ	27858	TKYHWHMYH	27899	VQDKFNYSY	27940	WMCDFSCIQ
27818	TDLLWCEVH	27859	TMGEVVINL	27900	VQDKINFSY	27941	WMCEFSCIK
27819	TDRQIAQYH	27860	TMSEVVINL	27901	VQDKINYGY	27942	WRDMVNSGP
27820	TDRQIDQYL	27861	TNGCAIENH	27902	VQDKINYNY	27943	YAEVAVQKE
27821	TDRQIDQYR	27862	TPGIHRHCC	27903	VQDKINYSH	27944	YANSIQMSF
27822	TDRQIGQYH	27863	TPGISRHCC	27904	VQDKISYSY	27945	YAQVAVQQE
27823	TDRQIHQYH	27864	TPGITRHCC	27905	VQDKITYSY	27946	YGGQVEFGW
27824	TDRQVDQYH	27865	TPGIYRHCC	27906	VQDKLNYSH	27947	YHGVICKEL
27825	TELFWCDVQ	27866	TPGLNRHCC	27907	VQDKLNYSY	27948	YHPHFPTYK
27826	TELFWCEVL	27867	TQLLWCEVQ	27908	VQDKRNYSY	27949	YHPYFPAYK
27827	TELLLCEVH	27868	TRRSWECSP	27909	VQDKSNYSY	27950	YHPYFPTHK
27828	TELLWCAVQ	27869	TRYRMPYHG	27910	VQDKVNYSY	27951	YHPYFPTNK
27829	TELLWCEIK	27870	TRYRMRYHC	27911	VQDMLNYSY	27952	YHPYFPTYQ
27830	TELLWCEIQ	27871	TRYRMRYHS	27912	VQDMVNYSY	27953	YHPYFPTYR
27831	TELLWCELQ	27872	TRYRMRYNG	27913	VQDNINYSH	27954	YHPYFSTYK
27832	TELLWCELR	27873	TSCGYHDNN	27914	VQDQINYSY	27955	YHPYVPTYR
27833	TELLWCEVA	27874	TSCGYNDNS	27915	VQDRINYSY	27956	YIANPCANC
27834	TELLWCEVE	27875	TSCRYNDNN	27916	VQDRVNYSY	27957	YIDNPCGNC
27835	TELLWCEVG	27876	TSRSAAADL	27917	VQDTINFSY	27958	YIDNPCTNC
27836	TELLWCGVQ	27877	TSRTEAADL	27918	VQDTINYSF	27959	YIDTPCANC

[2238]

27960	YIHNPCANC	27968	YKVNRDVSW	27976	YNGVIGKEL	27984	YSNSIKMSL
27961	YKLNRDDSW	27969	YKVNRDYSW	27977	YNGVLCKEL	27985	YSNSVQMSF
27962	YKMDCHSCG	27970	YKVNRNDSW	27978	YNGVMCKEL	27986	YSNTIQMSL
27963	YKVNGDDSW	27971	YNGGICKEL	27979	YNGVTCKEL	27987	YTGVICQEL
27964	YKVNLDDSW	27972	YNGVFCKEL	27980	YNGVVCKEL	27988	YWRMVCVHC
27965	YKVNQDDSW	27973	YNGVICKVL	27981	YQIRFDRHH	27989	YYMHWINSG
27966	YKVNRDASW	27974	YNGVICQEL	27982	YRVNRDDSW	27990	YYMHWVNAG
27967	YKVNRDGSW	27975	YNGVICREL	27983	YSAGPLAYN		

[2239] 实施例39

[2240] 如通过机器学习鉴定的具有骨骼肌组织嗜性的AAV5变体

[2241] 此实施例描述了如通过机器学习方法鉴定的展现骨骼肌组织嗜性的工程化AAV5变体的优选特性。进行如实施例19和图19至图22中描述的相同生物信息学预处理、机器学习方法和分析以确定展现骨骼肌组织嗜性的工程化AAV5变体的优选特性。

[2242] 下文描述了AAV5 VP衣壳多肽的581至589区域中各位置处的有利生物物理特性和有利氨基酸残基，它们与较高的骨骼肌组织嗜性概率相关。表62中描述的任何以下机器学习衍生的位置氨基酸偏好可以单独存在或彼此组合存在或与实施例9，表17中观察和描述的任何氨基酸偏好组合存在。下文表62中列出了从ML模型中提取的排名靠前的重要骨骼肌嗜性特征之间共有的位置特征的概述。

[2243]

表 62
机器学习衍生的骨骼肌组织嗜性规则

Xaa1 处高平均柔性 Xaa1 选自 G 或 R 或者 Xaa1 处低平均柔性 Xaa1 选自 W、M、F 或 H
Xaa1 处高摩尔质量 Xaa1 选自 R、F 或 W

[2244]

表 62 机器学习衍生的骨骼肌组织嗜性规则
Xaa2 处低亲疏水性 Xaa2 选自 K 或 R
Xaa2 处低可突变性 Xaa2 选自 C、R 或 H
Xaa2 处高平均柔性 Xaa2 选自 G 或 R
Xaa3 处高平均柔性 Xaa3 选自 G 或 R
Xaa4 处高亲水性 Xaa4 选自 D、E、R、K 或 N
Xaa4 处低可突变性 Xaa4 选自 C、R 或 H
Xaa5 处低摩尔质量 Xaa5 选自 A
Xaa5 处低平均柔性 Xaa5 选自 A 或 L
Xaa5 处高可突变性 Xaa5 选自 D、A 或 E
Xaa6 处低平均柔性 Xaa6 选自 W、M 或 F
Xaa6 处低可突变性 Xaa6 选自 C
Xaa6 处高摩尔质量 Xaa6 选自 W
Xaa7 处低 goldman engelman steitz Xaa7 选自 R

[2245]

表62 机器学习衍生的骨骼肌组织嗜性规则
Xaa7处高平均柔性 Xaa7选自D、R、P、G或S
Xaa7处高可突变性 Xaa7选自R、H或N
Xaa7处低溶解性 Xaa7选自R或Q
Xaa8处高亲水性 Xaa8选自D、E、R、K或N
Xaa9处低可突变性 Xaa9选自Y、F或L

[2246] 下文表63提供了表现出骨骼肌组织嗜性并且符合表62中提供的一种或多种规则的变体AAV VP1衣壳多肽的581至589区域的序列。因此,本公开提供了由具有SEQ ID NO:2的工程化AAV5 VP1衣壳多肽组成的rAAV,其中Xaa1至Xaa9区域是SEQ ID NO:29991-SEQ ID NO:30990中的任一者,如表63中所公开。本文还涵盖由工程化AAV5 VP2衣壳多肽和工程化AAV5 VP3衣壳多肽组成的rAAV,所述多肽在对应于AAV5 VP1 581至589区域中的氨基酸的AAV5VP2(氨基酸残基445至453)和AAV5 VP3(氨基酸残基389-397)中的区域处具有下文表中公开的序列。

[2247] 表63

[2248] 驱动骨骼肌组织嗜性的AAV5 VP1衣壳多肽中581至589区域的序列

[2249]

SEQ ID NO	581-589 序列
29991	AAGDIGSVM
29992	AAGNIGSVM
29993	AANLLWRLW
29994	AANLLWRLY
29995	ACAWIESFT
29996	ACPDHHRWL
29997	ACPNVWWHM
29998	ACQCKVIGQ
29999	ACQCRVIWQ
30000	ACTWPGVGQ
30001	ACWDTHQSI
30002	ACWDTKQSI
30003	ACWDTLQSI
30004	ACWDTQKSI
30005	ACWDTQQGI
30006	ACWRHVQWN
30007	AFGEFDNHS
30008	AFIHIDREC
30009	AGGHCTYVA
30010	AGWIGACMA
30011	AHWNAIMRW
30012	AIFEPHRWH
30013	AIFEPRRWP
30014	AIIEPRRWH
30015	AIKWRSQDK
30016	AIRVDDSTI
30017	AIVSMDDWH
30018	AKREIENCH
30019	AMHKRWTND
30020	ANNVYCVKK
30021	APANTPWRK
30022	APRPPRSAR
30023	ARDNESNLI
30024	ARSGAQSMM
30025	ARSGFASNT
30026	ARSSFASNN
30027	ARSSFASNS
30028	ARSSFPSNT
30029	ARSSFSSNT
30030	ARSSFTSNT
30031	ARSSVASNT
30032	ASTEVFYMI
30033	ATGIDNNAT
30034	AWSCVHADV
30035	AWWARCKQQ
30036	AWWTPCKQQ
30037	AWWTRCKQL
30038	CADVRENAN
30039	CCEIWLASA
30040	CCELWLASS
30041	CCRHWRPTT
30042	CCRYWRPTP
30043	CGGKIDIGL
30044	CHQENFIFR
30045	CHQENFVFH
30046	CHQENFVFP
30047	CIETGLLNG
30048	CIRQYDAIH
30049	CITYEGWAD
30050	CKVSESAKN
30051	CPPICEGGT
30052	CTETTVRWP
30053	CTKTTVRWP
30054	CTPTTVRWP
30055	CTQSSVRWP
30056	CTQTAVRWP
30057	CTQTTIRWP
30058	CTQTTVRWR
30059	CTQTTVRYP
30060	CTRSTVRWP
30061	CTRTTVRWP
30062	CVGMDSYNG
30063	CYGHLTAMC
30064	DARASPDCL
30065	DARPQWDIP
30066	DARSSPDCI
30067	DCEEFDRTT
30068	DCLVQKAAL
30069	DCNVNPACF
30070	DCWAIHSSW
30071	DCWDTQQSI
30072	DEVHGFLSG
30073	DEVYSFLSG
30074	DFHSYCMSF
30075	DFIHVGDWE
30076	DFTHVGDWD
30077	DGDPNFRIQ
30078	DHGYRNNSM
30079	DILIDHDFF
30080	DKQHDCRAS
30081	DKQQDCRAY
30082	DQQHDCRAY
30083	DRSSLCADE
30084	DRYSLCADA
30085	DSGRIYAWI
30086	DSNVQFCID
30087	DSRSSPDCL
30088	DSTLQFCID
30089	DTRKWMKWH
30090	DVPIDHDFF
30091	DVVGDAACP
30092	DWWTRCKQQ
30093	EARLQWDIP
30094	EARPHWDIP
30095	EARPQWDIL
30096	EARPQWDIS
30097	EARPQWDIT
30098	EARPQWDLP
30099	EARPQWDSP
30100	EARPQWDVP
30101	ECWQHVQWN
30102	ECWRHVKWN
30103	ECWRHVQLN
30104	ECWRHVQWI
30105	ECWRHVQWK
30106	ECWRHVQWS
30107	ECWRQVQWN
30108	ECYRKVVFR
30109	ECYSIYCEQ
30110	EDNIRMMMI
30111	EFAMDWVSD
30112	EFWGSMRNG
30113	EFWSSMRTG
30114	EGRPQWDIP
30115	EILAAWTPP
30116	EKCFQWTSP
30117	EKCFQWTSS
30118	EKCYQWTSA
30119	EKDLSHATF
30120	EKDLSQDTF
30121	EKTSRHWKT
30122	EKVHCMWLH
30123	EQEGDYVDM

[2250]

30124	EQGAVWTPP	30165	FDMFASPDA	30206	FMKCESVNS	30247	FYGQVTTSC
30125	EQGTVWTPP	30166	FDMFASPGD	30207	FMKGENVNS	30248	FYKWWINLG
30126	EQVRAFRQI	30167	FDVFASPDD	30208	FMTCENVNS	30249	GARPQWDIP
30127	ERAGFKMMD	30168	FFELSKTVF	30209	FMWHANSCG	30250	GARSSPDCL
30128	ERAGFQMMH	30169	FFELSQTEF	30210	FMWHANSGA	30251	GASWHDREH
30129	ERAGVKMMH	30170	FFRDCPYES	30211	FMWYANSCA	30252	GATEVFYMI
30130	ERAGYKMMH	30171	FFRDYPYEY	30212	FNINAQDWD	30253	GCDMDAYSN
30131	ERASFKMMH	30172	FGGKFDIGL	30213	FNVTAQDWD	30254	GCDVDAYSN
30132	ERDGFKMMH	30173	FGGKIAIGL	30214	FPHHLWVGG	30255	GDCGASNCY
30133	ERGGFKMMH	30174	FGGKIDIGV	30215	FPYHLWIGG	30256	GFHNTELPI
30134	ERVGFKMMH	30175	FGGKIDIVL	30216	FPYHLWVAG	30257	GGAKALRFN
30135	ESRPQWDIP	30176	FGGKIDVGL	30217	FPYHVWVGG	30258	GGARALRFI
30136	ETCEMFTPG	30177	FGGKIHIGL	30218	FPYRLWVGG	30259	GGARALRFS
30137	ETGNDNNAT	30178	FGGKINIGL	30219	FRECLQIVN	30260	GGGCEDWPM
30138	ETPINSEKG	30179	FGGKLDIGL	30220	FRECLQLVT	30261	GGPEIKHYI
30139	EVGTDPSKN	30180	FGGTIDIGL	30221	FRECMQIVT	30262	GHWNAIMRW
30140	EVRPQWDIP	30181	FGLVHQKNA	30222	FREGLQIVT	30263	GIDWEDVWE
30141	FAGKIDIGL	30182	FGRKIDIGL	30223	FRGCLQIVT	30264	GIKWRSQDK
30142	FAGPNSYWP	30183	FGVVHQKDA	30224	FRGIETEQG	30265	GIRDDDSTI
30143	FAMFASPDD	30184	FHHLTEYHG	30225	FRGIETGEG	30266	GIRGDDSTI
30144	FAVMIHYHN	30185	FHHQPEYHG	30226	FRGIETGQV	30267	GIRVDDNTI
30145	FAVMVHYPN	30186	FHHQTEYRG	30227	FRGIETVQG	30268	GIRVDDSTK
30146	FAVMVRYHN	30187	FHHQTGYHG	30228	FRGTETGQG	30269	GIRVDDSTV
30147	FCAAEWGCL	30188	FHHQTQYHG	30229	FRGVETGQG	30270	GMDEVEREG
30148	FCPAEWACL	30189	FHKWGINLG	30230	FRKCLQIVT	30271	GMEEGEREG
30149	FCPAEWGCV	30190	FHKWWIHLG	30231	FRRWQSNTT	30272	GMEELEREG
30150	FCPAEWGGL	30191	FHKWWINLA	30232	FTETTVRWP	30273	GNIYHFNNV
30151	FCPAEWGYL	30192	FHKWWINLS	30233	FTQPTVRWP	30274	GPDGAKLNK
30152	FCPAKWGCL	30193	FHKWWINQG	30234	FVLVHQKDA	30275	GPDGAKRHK
30153	FCPEHWFDI	30194	FHKWWISLG	30235	FWKFALLSF	30276	GPDGARRNK
30154	FCPEHWFNM	30195	FHKWWITLG	30236	FWKLALHSF	30277	GPDKCTVNL
30155	FCPEHWFNN	30196	FHKWWLNLG	30237	FWPQIQYRM	30278	GPGLKAWPC
30156	FCPEHWFNS	30197	FHRQTEYHG	30238	FYGKSAPVG	30279	GPLATTDPV
30157	FCPEHWLNI	30198	FLKWWINLG	30239	FYGKTALVG	30280	GPTEVFYMI
30158	FCPEHWVNI	30199	FMGATIPED	30240	FYGKTAPVA	30281	GRDWAFCDY
30159	FCPQHWFNI	30200	FMGVTIPAD	30241	FYGPVTKSC	30282	GRHWACCDY
30160	FCSAEWGCL	30201	FMGVTIPEG	30242	FYGQLTKSC	30283	GRHWAFCDD
30161	FCTAEWGCL	30202	FMGVTIPKD	30243	FYGQTAPVG	30284	GRHWAFCDF
30162	FCTEHWFNI	30203	FMGVTIPVD	30244	FYGQVTESC	30285	GRHWAFCDH
30163	FDLFASPDD	30204	FMKCENFNS	30245	FYGQVTMSC	30286	GRHWAFCDN
30164	FDMFASPAD	30205	FMKCENVNT	30246	FYGQVTQSC	30287	GRHWAFCGY

[2251]

30288	GRHWAFCHY	30329	HDMLQWSIK	30370	IINKRHDKF	30411	KHEADYVDM
30289	GRHWAFFDY	30330	HDMWQWTIK	30371	INQMARHPK	30412	KHEPRQSPR
30290	GRHWAFGDY	30331	HDQVKCLRK	30372	INVEYFDTW	30413	KHKARQSPR
30291	GRHWAVCDY	30332	HDTWQWSIK	30373	INVEYFEAW	30414	KHKLRQSPR
30292	GRHWDFCDY	30333	HEIDGWQTA	30374	INVEYFVAW	30415	KHMPRQSPR
30293	GRHWSFCDY	30334	HEIDVWQMA	30375	IRTSFHVNA	30416	KHSELWREN
30294	GRHWTFCDY	30335	HEVDGWQMA	30376	ISNNDLRFV	30417	KKCFQWTSA
30295	GRPWAFCDY	30336	HEVHSFLSG	30377	ISNNDLRVG	30418	KKVRAFRQI
30296	GRQWAVCDY	30337	HKRNLYALQ	30378	ISNNDVRFG	30419	KMKDVEENG
30297	GRRWAFCDY	30338	HLHFHWGLW	30379	ISTNDLRFG	30420	KMKPVQRVG
30298	GRVSGHVHW	30339	HLWRTFEAQ	30380	ITCKFTQNG	30421	KMMSFSEVD
30299	GRYSLCADE	30340	HLYSDYWFG	30381	ITCKMDFWQ	30422	KQEGDDVDM
30300	GRYWAFCDY	30341	HLYSEYWFD	30382	ITTGSRPTK	30423	KQEGDSVDM
30301	GSAEVFYMI	30342	HRDAHLLME	30383	IVFPSLARG	30424	KQVRAFREI
30302	GSNEVFYMI	30343	HRYSLCADE	30384	IWPQSQYRM	30425	KQVRAFRPI
30303	GSPEVFYMI	30344	HSHFHWGLW	30385	KAHKTVQDK	30426	KQVRAFRQL
30304	GSTELFYMI	30345	HTCSEIVDG	30386	KAMSVPWGK	30427	KQVRAFRQM
30305	GSTEVFCMI	30346	HTCSELVEG	30387	KANGSMDHH	30428	KQVRAFRQS
30306	GSTEVFYMV	30347	HTILSHDPW	30388	KANYACDWD	30429	KQVRAFRQV
30307	GSTEVFYVI	30348	HTIWSHDQW	30389	KCFSMDFSG	30430	KQVRAFRRI
30308	GSTKVFYMI	30349	HTLLLSREP	30390	KCMPGCNDW	30431	KQVRDFRQI
30309	GVEPMSRAA	30350	IACMKCPNI	30391	KCMPNVGWR	30432	KQVRPFRQI
30310	GVRDAVHWC	30351	IANNDLRFG	30392	KCNVNPACF	30433	KQVRSFRQI
30311	GWSARSVNM	30352	ICEMGPYHG	30393	KCQWVGREN	30434	KQVRVFRQI
30312	HAISVGEWD	30353	ICMSQEKRC	30394	KCVSMAFSG	30435	KRVHFEGGD
30313	HARGIWDFH	30354	IDVGAYVWS	30395	KCVSMDFAG	30436	KSMTELDQP
30314	HCAHYWMQD	30355	IFEQCQRMR	30396	KCVSMDFTG	30437	KTCELFTPG
30315	HCIRHFCEV	30356	IFRDCPYEY	30397	KCVSMGFSG	30438	KTCEMFTPA
30316	HCLLQKAAL	30357	IGIVFQRDF	30398	KCVSMVFSG	30439	KTCEMFTPS
30317	HCSCHFCEV	30358	IGLVFQRDF	30399	KCVSMYFSG	30440	KTCERFTPG
30318	HCSHHFCEV	30359	IGVVFPRDF	30400	KCVSVDFSG	30441	KTCGMFTPG
30319	HCSRHFCEA	30360	IGVVFQRDC	30401	KCWRHVQWN	30442	KTELAMVID
30320	HCSRHFCKV	30361	IGVVFQRDI	30402	KDHNRWNRV	30443	KTELDRDSM
30321	HCSRHFCQV	30362	IGVVFQRDV	30403	KDPNRWNRL	30444	KTEWAMAID
30322	HCSRHFGEV	30363	IGVVFQRDY	30404	KDVGDDAHG	30445	KTEWAMVLD
30323	HCSRHYCEV	30364	IGVVFQRVF	30405	KDYNRWNRL	30446	KTEWAMVVD
30324	HCTRHFCEV	30365	IHVEYFDAW	30406	KEHNRWNRL	30447	KTISGFQDK
30325	HCTRYWMQD	30366	IHWNAIMRW	30407	KFAKDWVSD	30448	KTKWAMVID
30326	HCYRKVVFR	30367	IIKWRSQDK	30408	KFAMDWDSD	30449	KTMSGPWGK
30327	HDIDGWQMA	30368	IINKDHDKF	30409	KFAMDWVSV	30450	KVNYACDWH
30328	HDIWQWSIK	30369	IINKHHDRF	30410	KFVSMDFSG	30451	KVTYACDWD

[2252]

30452	KWDSIFVEH	30493	MNLVKIRNS	30534	NMGIASPWH	30575	PGDPQEDTV
30453	KWDSIFVQR	30494	MPVKKEYGG	30535	NMVDENRFW	30576	PGIDLWHMA
30454	KYAANWLKD	30495	MRHWAFCDY	30536	NMVHEHRFW	30577	PGIEWWHMA
30455	KYADNWLKA	30496	MRIHFEGGD	30537	NMVHESRFW	30578	PGWIGACMA
30456	KYADNWVKD	30497	MRVHFEGGG	30538	NMVHQNRFW	30579	PHFDKCQHF
30457	KYADYWLKD	30498	MRVHFQGGD	30539	NMVRENRFW	30580	PIDWEDVWE
30458	KYSDNWLKD	30499	MRVRFEGGD	30540	NSISVGEWD	30581	PKAMAKICE
30459	KYSEVWREN	30500	MSSLKGSDE	30541	NTSCWKLEN	30582	PKGMAKICA
30460	KYSKLWREN	30501	MSWLGAAGI	30542	NTVIRYRHS	30583	PKGMAKIFE
30461	KYWMRHVVM	30502	MTLLMGHES	30543	NVNEDVQWQ	30584	PKGMAKIWE
30462	LATSMCYQS	30503	MVCIHADTD	30544	NWVNGFVDH	30585	PKGMARICE
30463	LCHFLRTCK	30504	MWLAKARQG	30545	NWWAIHSSW	30586	PKGMSKICE
30464	LDKIEHDWN	30505	MWVGKARQG	30546	NYVLFVDHE	30587	PKGMVKICE
30465	LFCIVQSRE	30506	MYKYGQDWY	30547	PAWDLVEML	30588	PKGNRTNQI
30466	LIDRIWFNW	30507	NAINVGEWD	30548	PAWDLVGMF	30589	PKVMAKICE
30467	LSGPIHYEP	30508	NAISVGKWD	30549	PAWDLVGMH	30590	PMHQRWTND
30468	LSGSIHYES	30509	NALSVGEWD	30550	PAWDLVGMI	30591	PMMKKLCCG
30469	LSVKMDRAK	30510	NARSSPDCL	30551	PAWDPVGML	30592	PNGEKCERF
30470	LSVQMDRSK	30511	NCSRHFCEV	30552	PAWELVGMV	30593	PQHISADPT
30471	LSVRMDRAK	30512	NCWAIHGSW	30553	PAWHLVGML	30594	PQHISSDPT
30472	LTGQEAGKG	30513	NCWAILSSW	30554	PCDFDNQPH	30595	PQLISPDPT
30473	LVCIRADTD	30514	NDLGDDAHG	30555	PCHFDNQPH	30596	PQRISPDPT
30474	LVGKYWPRD	30515	NDVADDAHG	30556	PCLDHHRWL	30597	PRDDESNLI
30475	LVGPYWPRD	30516	NDVGDDDHG	30557	PCNFDSQPH	30598	PRDKIWFMN
30476	MADHCCANS	30517	NDVGDEAHG	30558	PCNVDNQPH	30599	PRDNVSNLI
30477	MADHCCTDS	30518	NDVGGDAHG	30559	PCPDHHRLL	30600	PRFCISLKC
30478	MAFSSMHQG	30519	NDVGNDAHG	30560	PCPDHHRRL	30601	PRGNESNLI
30479	MASLRGSDE	30520	NEIDGWQMA	30561	PCPDHHRWV	30602	PRMSYEIKG
30480	MCYQNIAFQ	30521	NEIMMWRAN	30562	PCPDHNRWL	30603	PRTSYEFKG
30481	MFGECYNNG	30522	NETMMWRTN	30563	PCPDNHRWL	30604	PSKPFSKEE
30482	MFKYYPYCD	30523	NFLDYHDMV	30564	PCPDPHRWL	30605	PSQPFSEEE
30483	MFLECYNNG	30524	NFMDYNDMV	30565	PCPEHHRWL	30606	PTWDLVGML
30484	MFQSYPYCD	30525	NFMEYHDMV	30566	PCPNVWWHM	30607	PWCECHRLF
30485	MFQYYQYCD	30526	NHQADEHDA	30567	PCQCKVIWQ	30608	PWSWCEFNK
30486	MFRECYDNG	30527	NIRPSDASG	30568	PCTDHHRWL	30609	PWSWCEYNQ
30487	MFRECYKNG	30528	NIRPSDASS	30569	PCTFDNQPH	30610	PWWTRCKQQ
30488	MFRECYNSG	30529	NIRPSDTSA	30570	PCYRKVVFR	30611	PYGEKCERF
30489	MFREWYNNG	30530	NIRQSDASA	30571	PDGERCERF	30612	PYTHWKLNN
30490	MFRKCYNNG	30531	NIRRSDASA	30572	PDWRKIKKF	30613	QARIEDMAK
30491	MGDGLWVKV	30532	NKRNLYAMQ	30573	PFMKDGECE	30614	QCEWQKMEH
30492	MHFKYPQRG	30533	NKRNLYAVQ	30574	PFMKDVKCE	30615	QCFLCGSAT

[2253]

30616	QCHFLRNCK	30657	REQIACPWY	30698	RLELDNNEA	30739	RVCPFCCCF
30617	QCMPGCNDL	30658	REQIARPGY	30699	RLEQYNNEA	30740	RVCTFCGCF
30618	QCPVECDHV	30659	REQISRPWY	30700	RLERANNEA	30741	RVDINVDCY
30619	QCVSMDFSG	30660	REQLARPWY	30701	RLERDDNEA	30742	RVDNDMADG
30620	QCYREVVFR	30661	RETLEQVAY	30702	RLERDHNEA	30743	RVERDNNEA
30621	QFAMDWVSD	30662	RFNLPNFYD	30703	RLERDINEA	30744	RVKRDNNEA
30622	QIRIEDMAK	30663	RFNWPIFYD	30704	RLERDNIEA	30745	RWEKMAYSW
30623	QKGNRTDQI	30664	RFNWPNVYD	30705	RLERDNNAA	30746	RYCDAWPDV
30624	QKGNRTTQI	30665	RFPFRAQVW	30706	RLERDNNAV	30747	SAWDLVGML
30625	QKRGIEAVD	30666	RFRFLAQVW	30707	RLERDNNED	30748	SCENWGRQH
30626	QKRIEDMAK	30667	RFRFRAKVW	30708	RLERDNNEP	30749	SCETGGRQH
30627	QLNICQATF	30668	RFRFRAQGW	30709	RLERDNNES	30750	SCETWGRHH
30628	QMHNTFCGW	30669	RFRFRAQIW	30710	RLERDNNET	30751	SCETWGRPH
30629	QMHSTFCSW	30670	RFRFRTQVW	30711	RLERDNTEA	30752	SCETWSRQH
30630	QRDKVWFMN	30671	RFRFRVQVW	30712	RLERDTNEA	30753	SCETWVRQH
30631	QRDRIWFMN	30672	RFRIRAQVW	30713	RLERHNNEA	30754	SCEWSKQCK
30632	QSVDVGLMH	30673	RFRVRAQVW	30714	RLERNNNEA	30755	SCMDGTTGV
30633	QTGDWWQYG	30674	RFTWPNFYD	30715	RLGRDNNEA	30756	SCMHGTTGV
30634	QTGNWWEYG	30675	RHDWQHRTV	30716	RLKRDNNEA	30757	SCMNGATGV
30635	QTGNWWKYG	30676	RHELTMSHH	30717	RLKRVNNEA	30758	SCMNGPTGV
30636	QTGNWWQYA	30677	RHELTMSQY	30718	RLQRDNNEA	30759	SCMNGTAGV
30637	QTGTWWQYG	30678	RHQPLHYFH	30719	RLYSDYWFD	30760	SCMNGTTVV
30638	QTISGFKDK	30679	RHVPLHYFH	30720	RMGYEVDWQ	30761	SCMNVTTGV
30639	QTISGFQYK	30680	RIDFNVDCY	30721	RMMYFDKFT	30762	SCPWPTGEE
30640	QTIVAWHHS	30681	RIDINVDCH	30722	RNLDYNTAV	30763	SCQTWGRQH
30641	QTLSGFQDK	30682	RIDINVDWY	30723	RNLNYNTGV	30764	SCRHWRPTP
30642	QTQIEDMAK	30683	RIDINVNCY	30724	RNPNYNTAV	30765	SCVTWGRQH
30643	QTRIDDMAK	30684	RIDINVYCY	30725	RNVNYNTAV	30766	SCWDTQQSI
30644	QTRIEDMTK	30685	RIDISVDCY	30726	RPERDNNEA	30767	SDKWYRVSF
30645	QTRIQDMAK	30686	RIDLNVDCY	30727	RQQIARPWY	30768	SFQHAQWMG
30646	QTRRLEPYH	30687	RIDVNVDCY	30728	RRCCISLKC	30769	SGWIAACMA
30647	QTRVEDMAK	30688	RIGINVDCY	30729	RRDSDSVCK	30770	SGWIRACMA
30648	QVCIHADTD	30689	RIMFFDKFT	30730	RRFCVSLKC	30771	SGWNGACMA
30649	RADVDMTSI	30690	RIMYFDEFT	30731	RTECVPDTK	30772	SHFESHQWC
30650	RCEIWLASS	30691	RIMYFDKCT	30732	RTLDSEYFF	30773	SHGESHQWC
30651	RCISNEGMQ	30692	RIMYFDRFT	30733	RTLDSEYWF	30774	SIMMPDYSC
30652	RCISNEWME	30693	RINWPNFYD	30734	RTQCAPDTK	30775	SITDANLGF
30653	RCISNEWMP	30694	RITYFDKFT	30735	RTQCVPATK	30776	SKIYAVDQS
30654	RDHNRWNRL	30695	RKDTPKAGH	30736	RTQCVPDAK	30777	SKRDIENCH
30655	REEIARPWY	30696	RLARDNNES	30737	RTQWVPDTK	30778	SKRDVQSMM
30656	REKIARPWY	30697	RLARENNEA	30738	RTVDSEYLF	30779	SKREIENCQ

[2254]

30780	SKREIETCH	30821	SWCECHRSF	30862	VAEMHWMVN	30903	VMQNVWCSW
30781	SKRELENCH	30822	SWCECHRVF	30863	VAIGFHWHA	30904	VMSCHSRDH
30782	SKRKIENCH	30823	SWCGCHRLF	30864	VAWAVSYVC	30905	VPKNVNRDM
30783	SKWACHPKM	30824	SWCVCHRLF	30865	VAWGVSYVG	30906	VPKTANRDM
30784	SKWNCHPKM	30825	SWNELKRNQ	30866	VCDIDAYSN	30907	VPKTVDRDM
30785	SLNELKRDQ	30826	SWSWCEFNQ	30867	VCEMGPYDG	30908	VPTTVNRDM
30786	SMAKHFTSS	30827	SYTHLKLNN	30868	VCEMGPYNG	30909	VQDKCTVNL
30787	SMHKIQYRA	30828	SYTPWKLNN	30869	VCEMGPYRG	30910	VRHWAFCDY
30788	SQEFVCAQE	30829	TAWDLVGML	30870	VCEMGQYHG	30911	VSVQMDRAK
30789	SRDNESNLI	30830	TAWRKIKKF	30871	VCERGPYHG	30912	VTGKMDFWQ
30790	SREQVNCDA	30831	TCKWVGREN	30872	VCQMGPYHG	30913	VTKNVWCSW
30791	SRISFVCSD	30832	TCMFRVYDN	30873	VCVAIMKDQ	30914	VTKWRSQDQ
30792	SRKQVNCDA	30833	TCPNVWWHK	30874	VCVAIMKGK	30915	VTMIMMDCH
30793	SRLQVNCDA	30834	TCPNVWWHL	30875	VCVAIMQDK	30916	VTRDLKKEQ
30794	SRLSFVCSD	30835	TCPNVWWRM	30876	VCVMGPYHG	30917	VVDMCRKGK
30795	SRQPVNCDA	30836	TCQWFGREN	30877	VCWDTQQSI	30918	VVRDLPNQY
30796	SRQQANCDA	30837	TDWPKIKKF	30878	VCWMAGAQM	30919	VWFGKCVNC
30797	SRQQVHCDA	30838	TDWRKFKKF	30879	VDKAWWNLI	30920	VWLKSLDRM
30798	SRQQVNCDV	30839	TDWRKLKKF	30880	VDKVWWNLV	30921	VWPKGMREM
30799	SRQQVNCEA	30840	TDWRKMKKF	30881	VDRSYGPSG	30922	VWSCVHTDV
30800	SRQQVNGDA	30841	TDWRKVKKF	30882	VFGDFDNHS	30923	WDVVAADGP
30801	SRQRVNCDA	30842	TDWRRIKKF	30883	VFVACHVNG	30924	WDVVATDVP
30802	SRSMSSSME	30843	TEVKGWESD	30884	VGDLNFRIQ	30925	WERAWLFGG
30803	SRSMYSGME	30844	TFMDYHDMV	30885	VHLNADTKF	30926	WERFWHEYW
30804	SRVAFVCSD	30845	TGFHNTFKD	30886	VHWDAIMRW	30927	WHGRLMITF
30805	SSAIQCDAA	30846	THQGDEHDA	30887	VHWNAIMRG	30928	WHQENFVFR
30806	SSGIECDAA	30847	TISDIPGEP	30888	VHWNAIMRL	30929	WKDPFNRKI
30807	SSGIQCDAT	30848	TISDIPGKP	30889	VHWNALMRW	30930	WLANGWFRE
30808	SSGIQCEAA	30849	TLDECCRRG	30890	VHWNASMRW	30931	WLANSWFRG
30809	SSGLQCDAA	30850	TLPCYQRYT	30891	VHWQQCGHQ	30932	WLDNSWFRE
30810	SSILTWTWI	30851	TMPCYQRYT	30892	VIDRIWFNW	30933	WLPNSWFRE
30811	SSRIQCDAA	30852	TNVEYFDAW	30893	VIKGRSQDK	30934	WLSNSWFRE
30812	STDNHWGHG	30853	TNWREMNAR	30894	VIKLRSQDK	30935	WMCRNHHFD
30813	STQATVRWP	30854	TNWRKIKKF	30895	VIKWGSQDK	30936	WMCRNHPFN
30814	STQEDFDMF	30855	TPRPPRSGR	30896	VIKWRGQDK	30937	WMCRNHPVD
30815	STQTSVRWP	30856	TPRPSRSAR	30897	VIKWRSLDK	30938	WMCRSHPFD
30816	SVQWYRVSF	30857	TPRTPRSAR	30898	VIKWRSRDK	30939	WRCLCARID
30817	SWCDCHRLF	30858	TRSGAQAMM	30899	VIQWRSQDK	30940	WRCMCARIA
30818	SWCECHRLL	30859	TSGTCMPIK	30900	VIRWRSQDK	30941	WRCMCARLD
30819	SWCECHRLV	30860	TWMVVHRNN	30901	VLPCQRKRV	30942	WRCMCARMD
30820	SWCECHRMF	30861	TYQWVGREN	30902	VLWNAIMRW	30943	WRCMCARTD

[2255]

30944	WRCMCARVD	30956	WRRFASGWE	30968	WYGHLNAMC	30980	YMGIASPWH
30945	WRCMCSRID	30957	WRRFASWLE	30969	WYSWRVRNR	30981	YMWWEEGTL
30946	WRCMCTRID	30958	WRRFSSWWE	30970	YARGIWDIH	30982	YMWWQAGTL
30947	WRCWYAPLT	30959	WRRFVSWWE	30971	YARGIWDVH	30983	YRQQVNCDA
30948	WRELSNTGY	30960	WRTADEGDA	30972	YARGTWDFH	30984	YRRLQSNTT
30949	WREWSHTGY	30961	WRTADESDT	30973	YCSRHFCEV	30985	YRRWESNTT
30950	WREWSNAGY	30962	WRTAHESDA	30974	YDRGIWDFH	30986	YRRWQSNTI
30951	WREWSNTAY	30963	WRTGDESDA	30975	YEVHSFLSG	30987	YRRWQSNTS
30952	WREWSNTCY	30964	WTETTVRWP	30976	YHDLCFAWY	30988	YVKFNVNIN
30953	WREWSNTVY	30965	WTQTTVRWP	30977	YHDLFFDWY	30989	YVKFSVDIN
30954	WRGWSNTGY	30966	WVANSWFRE	30978	YHDLFFGWY	30990	YVKISVNIN
30955	WRQWSNTGY	30967	WVCRNHPFD	30979	YKREIENCH		

[2256] 实施例40

[2257] 如通过机器学习鉴定的具有皮肤组织嗜性的AAV5变体

[2258] 此实施例描述了如通过机器学习方法鉴定的展现皮肤组织嗜性的工程化AAV5变体的优选特性。进行如实施例19和图19至图22中描述的相同生物信息学预处理、机器学习方法和分析以确定展现皮肤组织嗜性的工程化AAV5变体的优选特性。

[2259] 下文描述了AAV5 VP衣壳多肽的581至589区域中各位置处的有利生物物理特性和有利氨基酸残基，它们与较高的皮肤组织嗜性概率相关。表64中描述的任何以下机器学习衍生的位置氨基酸偏好可以单独存在或彼此组合存在或与实施例17，表33中观察和描述的任何氨基酸偏好组合存在。下文表64中列出了从ML模型中提取的排名靠前的重要皮肤嗜性特征之间共有的位置特征的概述。

[2260]

表 64
机器学习衍生的皮肤组织嗜性规则

Xaa1 处低表面可接近性 Xaa1 选自 C
Xaa1 处低体积 Xaa1 选自 C

[2261]

表 64 机器学习衍生的皮肤组织嗜性规则
Xaa1 处低可突变性 Xaa1 选自 C
Xaa2 处高表面可接近性 Xaa2 选自 R 或 K
Xaa2 处高平均柔性 Xaa2 选自 K、I 或 N
Xaa2 处低可突变性 Xaa2 选自 P 或 K
Xaa3 处高亲疏水性 Xaa3 选自 I 或 V
Xaa4 处低可突变性 Xaa4 选自 L、F 或 Y
Xaa4 处低平均柔性 Xaa4 选自 W、H、F 或 M
Xaa5 处高平均柔性 Xaa5 选自 G、R、K、I 或 N
Xaa6 处高平均柔性 Xaa6 选自 G、R、K、I 或 N
Xaa8 处高表面可接近性 Xaa8 选自 M、G 或 F
Xaa8 处低平均柔性 Xaa8 选自 H、F、M 或 W
Xaa8 处低可突变性 Xaa8 选自 L、F 或 Y

[2262]

表 64 机器学习衍生的皮肤组织嗜性规则
Xaa9 处高平均柔性 Xaa9 选自 D、E、R、K、P 或 G
Xaa9 处高可突变性 Xaa9 选自 D、E、R、V、A 或 H

[2263] 下文表65提供了表现出皮肤组织嗜性并且符合表64中提供的一种或多种规则的变体AAV VP1衣壳多肽的581至589区域的序列。因此，本公开提供了由具有SEQ ID NO:2的工程化AAV5 VP1衣壳多肽组成的rAAV，其中Xaa1至Xaa9区域是SEQ ID NO:32991-SEQ ID NO:33990中的任一者，如表65中所公开。本文还涵盖由工程化AAV5 VP2衣壳多肽和工程化AAV5 VP3衣壳多肽组成的rAAV，所述多肽在对应于AAV5 VP1 581至589区域中的氨基酸的AAV5VP2(氨基酸残基445至453)和AAV5 VP3(氨基酸残基389-397)中的区域处具有下文表中公开的序列。

[2264] 表65

[2265] 驱动皮肤组织嗜性的AAV5 VP1衣壳多肽中581至589区域的序列

[2266]

SEQ ID NO	581-589 序列
32991	AADAETPWF
32992	AADPRADCP
32993	AADPRGDCP
32994	AADPRIDCP
32995	AADPRVDCA
32996	AADPRVDCL
32997	AADPRVDCR
32998	AADPRVECP
32999	AADPTKSWC
33000	AADSETPWY
33001	AASPYITYR
33002	AEHCYRHKH
33003	AEPFPCDWS
33004	AFFYMDKLS

[2267]

33005	AIVCRVNYP	33046	CAVMKFSEH	33087	CHAPSIMGP	33128	CILHKHSVP
33006	AKRDMDPAQ	33047	CAVYFAKCA	33088	CHCHIRHTG	33129	CILHKNSFP
33007	AKRDMDPAS	33048	CAVYFGKCV	33089	CHFMFNAGR	33130	CILQKHSFP
33008	AKRNMDPAP	33049	CAVYFGRCA	33090	CHFMIHAGR	33131	CILYKHSFP
33009	ANETNDNFH	33050	CAVYYGKCA	33091	CHFMINAAR	33132	CIMHKHSFP
33010	ANVGCIMIL	33051	CAYFKRHPD	33092	CHFMINACR	33133	CIQACYDHE
33011	ASARYIDFT	33052	CAYFSMMDR	33093	CHFMINAGH	33134	CIQFTKHQE
33012	ASWCSDERC	33053	CAYHLNQTI	33094	CHFMINDGR	33135	CIQLSKMGP
33013	ATAGCIMIL	33054	CCLYYVNVE	33095	CHFMINSGR	33136	CIQLTKMGP
33014	ATDPRVDCP	33055	CCTIKQMDI	33096	CHFMINTGR	33137	CIRSYNVMM
33015	ATVGFIMIL	33056	CCVWEFNYP	33097	CHFMINVGR	33138	CITACYDHE
33016	AVATRAMYY	33057	CDHQFIDQA	33098	CHFMLNAGR	33139	CITMVRSQT
33017	AVDPRVDCP	33058	CEHFKRHPD	33099	CHFMSNAGR	33140	CKAHCPGQY
33018	CAFQCYRSP	33059	CERLTQNVD	33100	CHFMTNAGR	33141	CKCHFDQVE
33019	CAFVQQNFQ	33060	CEYFKRDPD	33101	CHFMVNAGR	33142	CKDFHWMND
33020	CAGMQATGD	33061	CEYFKRHPA	33102	CHFVINAGR	33143	CKFDITKMA
33021	CAHHLNQTI	33062	CEYFKRHPG	33103	CHHATNTSQ	33144	CKFDMASFF
33022	CAIFNTVDQ	33063	CEYFKRHTD	33104	CHHFATASE	33145	CKFDMPCFF
33023	CAIHTQSHS	33064	CEYFKRRPD	33105	CHHFATETE	33146	CKFDMPGFF
33024	CAIMCSAGP	33065	CEYFKRYPD	33106	CHHFATQSE	33147	CKFDMPSCF
33025	CAIMFSAGQ	33066	CEYIKRHPD	33107	CHIMINAGR	33148	CKFDMPSFI
33026	CAIMFSAGT	33067	CFHPYHQAH	33108	CHPVYKDWS	33149	CKFDMPSFV
33027	CAIMLKHMT	33068	CFHQYNQAH	33109	CHSMINAGR	33150	CKFDMPSFY
33028	CAIMLQHMA	33069	CFLWEFNYP	33110	CHTIKQMDI	33151	CKFDMPSVF
33029	CAIMLQHMI	33070	CFMALSFHW	33111	CHVMINAGR	33152	CKFDMPTFF
33030	CALHTRSHS	33071	CFMGLSFHL	33112	CHYMINAGR	33153	CKFDMQGFF
33031	CALMFSAGP	33072	CFMVLSFHW	33113	CIKACNDHE	33154	CKFDMQSFF
33032	CALPVRVAM	33073	CFPVYKDWS	33114	CIKACYDRE	33155	CKFDMSSFF
33033	CALPYAGSR	33074	CFVGLSFHW	33115	CIKACYDYE	33156	CKFDMTSFF
33034	CALSMYSHE	33075	CFVLEFNYP	33116	CIKACYHHE	33157	CKFHMPSFF
33035	CANHLNKTI	33076	CFVWAFNYP	33117	CIKACYNHE	33158	CKFNMPSFF
33036	CANHLNQTM	33077	CFVWEFDYP	33118	CIKAYNVMM	33159	CKFYIPSFF
33037	CANHLNRTI	33078	CFVWEFNDP	33119	CIKAYYDHE	33160	CKGHFDQVE
33038	CANHLSQTI	33079	CFVWEFNHP	33120	CIKGCYDHE	33161	CKHFHWMND
33039	CANHVNQTI	33080	CFVWEFNSP	33121	CIKPCYDHE	33162	CKIDMPSFF
33040	CAPISAAPV	33081	CFVWEFSYP	33122	CIKSCYDHE	33163	CKIHFDQVE
33041	CARLIDWCC	33082	CFVWGFNYP	33123	CIKSYNVMV	33164	CKNFHLMND
33042	CASMFSAGP	33083	CFVWGINYP	33124	CIKTYNVMM	33165	CKNFHRMND
33043	CATHLNQTI	33084	CFVWQFNYP	33125	CIKYYNVMM	33166	CKNFHWMDD
33044	CATVKMWHT	33085	CFVWVFNYP	33126	CILHKHNFP	33167	CKNFHWMHD
33045	CAVMFSAGP	33086	CFYQYHQAH	33127	CILHKHSSP	33168	CKNFHWMNE

[2268]

33169	CKNFHWMNH	33210	CKYDMPSFF	33251	CNHKFIDQA	33292	CTCNLPSAI
33170	CKNFHWMSD	33211	CKYDNTKMA	33252	CNHKVNSWI	33293	CTCNVPYAI
33171	CKNFHWRND	33212	CKYDTTKMA	33253	CNHPFIDQA	33294	CTCWYNRGE
33172	CKNFLWMND	33213	CKYDVTKMA	33254	CNHQFIDQG	33295	CTEVLQRKG
33173	CKNFPWMND	33214	CKYFHWMND	33255	CNHQFIDQS	33296	CTGCYMKMP
33174	CKNFYWMND	33215	CLELRTQYA	33256	CNHQFIDQT	33297	CTGIQATGD
33175	CKNHFDQVE	33216	CLELVKAYE	33257	CNHQFIDQV	33298	CTGLLMHKE
33176	CKNIHWMND	33217	CLELVKDYE	33258	CNHQFLDQA	33299	CTGQLMHEE
33177	CKPQINDAY	33218	CMFHNHKGF	33259	CNPQFIDQA	33300	CTGQLMHQE
33178	CKQQINDAC	33219	CMFQCYRSP	33260	CNYQTNVHL	33301	CTGQLMRKE
33179	CKQQINDAD	33220	CMFVQQNFQ	33261	CPKPFLIKE	33302	CTGWYNRGD
33180	CKQQINDAF	33221	CMHLSCDWA	33262	CPKPVLIQE	33303	CTHAVHDNE
33181	CKQQINDAH	33222	CMHLVNMNS	33263	CQELRTQHA	33304	CTHGFLTAP
33182	CKQQINDAN	33223	CMKACYDHE	33264	CQELRTQYT	33305	CTHGFLTSP
33183	CKQQINDAS	33224	CMKFIRVGS	33265	CQEVHAHMS	33306	CTHGFLTVP
33184	CKQQINDDY	33225	CMPFIRVGS	33266	CQEVYAHMT	33307	CTHGVLTGP
33185	CKQQINDGY	33226	CMQFIRAGS	33267	CQKWRSIWK	33308	CTHKVNTWI
33186	CKQQINDSY	33227	CMQFIRIGS	33268	CQQVYAHMS	33309	CTHPVHDHE
33187	CKQQINHAY	33228	CMQFIRLGS	33269	CQSHFDQVE	33310	CTHPVNDNE
33188	CKQQLNDAY	33229	CMQFIRVAS	33270	CRAAHPVCA	33311	CTHQFIDQA
33189	CKRQINDAY	33230	CMQFIRVGG	33271	CRAAHSICA	33312	CTHQVHDNE
33190	CKSFHWMND	33231	CMQFIRVGN	33272	CRAAYSVCA	33313	CTHTVHDNE
33191	CKSHFAQVE	33232	CMQFLRVGS	33273	CRATHSVCA	33314	CTHVFLTGP
33192	CKSHFDKVE	33233	CMQFNRVGS	33274	CRAVHSVCA	33315	CTIFHTVDQ
33193	CKSHFNQVE	33234	CMQFSRVGS	33275	CRFMINAGR	33316	CTIFNNVDQ
33194	CKSHIDQVE	33235	CMQFTRVGS	33276	CRTAHSVCA	33317	CTIFNTVDK
33195	CKSHVDQVE	33236	CMQFVRVGS	33277	CRVAHSVCA	33318	CTIFTTVDQ
33196	CKTFHWMND	33237	CMQVIRVGS	33278	CSAMVRSQT	33319	CTIMFSAGP
33197	CKTHFDQVE	33238	CMQYIRVGS	33279	CSGCHMKMP	33320	CTIRANLHF
33198	CKTQCPGQY	33239	CMRFIRVGS	33280	CSGCYMKMS	33321	CTIYRNLLS
33199	CKVDMPSFF	33240	CMSWLTNQE	33281	CSHPVHDNE	33322	CTKACYDHE
33200	CKYDFTKMA	33241	CMVHNHKGS	33282	CSLHKHSFP	33323	CTKMLQRKG
33201	CKYDIPKMA	33242	CMVHNHQGF	33283	CSLPYAWSR	33324	CTKSYNVMM
33202	CKYDISKMA	33243	CMYWLNNQE	33284	CSMIQNSFP	33325	CTLKVNSWI
33203	CKYDITKMG	33244	CMYWLTNKE	33285	CSTMVRSQA	33326	CTLMHIHTR
33204	CKYDITKMS	33245	CNCNLPYAI	33286	CSVCYMKMP	33327	CTLMHSHAR
33205	CKYDITKMT	33246	CNCPIRHTG	33287	CSVWEFNYP	33328	CTLMHSHNR
33206	CKYDITKMV	33247	CNCQIRHTG	33288	CTAMLQRKG	33329	CTLMHSHTI
33207	CKYDITMMA	33248	CNFMINAGR	33289	CTAVKMWHT	33330	CTLMLSHTR
33208	CKYDITQMA	33249	CNFQTNVHM	33290	CTCGYNRGD	33331	CTMLANHDQ
33209	CKYDLTKMA	33250	CNFQTSVHL	33291	CTCHIRHTG	33332	CTMLANHHQ

[2269]

33333	CTMLANHYK	33374	CVFQCYRTP	33415	CYPVYQDWS	33456	DYCRRTFHH
33334	CTMLANHYP	33375	CVFQYYRSP	33416	CYPVYTDWS	33457	DYQPCHVHM
33335	CTMLANHYR	33376	CVFVPQNFQ	33417	CYSVYKDWS	33458	DYSNQRVDR
33336	CTMLANNYQ	33377	CVFVQENFQ	33418	CYTIKHMDI	33459	DYTNQRVDR
33337	CTMLANYYQ	33378	CVFVQQHFQ	33419	CYTIKKMDI	33460	EDESRHISW
33338	CTMMANHYQ	33379	CVFVQQSFQ	33420	CYTIKQMAI	33461	EDEYRHMSW
33339	CTMVANHYQ	33380	CVFVQQTFQ	33421	CYTIKQMDL	33462	EDKSRHMSW
33340	CTMWANHYQ	33381	CVFVRQNFQ	33422	CYTIKQMDM	33463	EFVTRKAAD
33341	CTNFHWMND	33382	CVGMKFSEH	33423	CYTIKQMDS	33464	EFVTRKAAT
33342	CTNVKMWHT	33383	CVHLVHWQW	33424	CYTIKQMGI	33465	EGEDRQAFW
33343	CTNYHLKAG	33384	CVHQYHQAH	33425	CYTIKQMHI	33466	EHCNHDQGG
33344	CTNYHVKAV	33385	CVLMKFSEH	33426	CYTIKQMNI	33467	EIFERSMMY
33345	CTPVKMWHT	33386	CVLVQQNFP	33427	CYTIKQVDI	33468	EKAPWGANH
33346	CTQFIKHQE	33387	CVMGLSFHW	33428	CYTMKQMDI	33469	ESARYIDFA
33347	CTQFPKHQE	33388	CVRMIDWCC	33429	CYTVYKDWS	33470	ESARYLDFT
33348	CTQFTQHQE	33389	CVSISAAPV	33430	DEPCPCDWS	33471	FAEPYTDHQ
33349	CTQFTTHQE	33390	CVVMKFNEH	33431	DEPFPCDWA	33472	FAERTTDHR
33350	CTQLNKMGP	33391	CVVMKFSED	33432	DEPFPCDWT	33473	FAQRATDHR
33351	CTTAWQGTW	33392	CVVMKFSER	33433	DESFPCDWS	33474	FAQRNTDHR
33352	CTTFKMWHT	33393	CVVMKFSKH	33434	DFVRTNLGV	33475	FAQRTTHHR
33353	CTTGWQGAW	33394	CVVMKFSVH	33435	DHANQRVDR	33476	FAQRTTNHR
33354	CTTVKMLHT	33395	CVVMKFTEH	33436	DIFERSMMF	33477	FFGTPYHMS
33355	CTTVKMWHA	33396	CVVMKYSEH	33437	DKAGYGCHE	33478	FGHWAQDAS
33356	CTTVKMWHI	33397	CVVMTFSEH	33438	DKEDPQQWL	33479	FGQLIIKHR
33357	CTTVKMWNT	33398	CVVVQQNFQ	33439	DKEEPEQWL	33480	FGQRTTDHR
33358	CTTVQMWHT	33399	CVVWEFNYP	33440	DKGGNGCHE	33481	FHDRHFTFT
33359	CTTYRNLMS	33400	CVYQCYRSP	33441	DKGGYGCNE	33482	FHDRPFNFT
33360	CTTYRNLPS	33401	CVYWLTNQE	33442	DKGGYGGHE	33483	FHDRPFTFA
33361	CTTYRNVLS	33402	CWAVTSMGS	33443	DKGGYSCHE	33484	FICLIKSAS
33362	CTVFNTVDQ	33403	CWMPQQVAA	33444	DKSEESRQE	33485	FICQIKSAA
33363	CTVLANHYQ	33404	CWVVTSMSS	33445	DKSPWGANH	33486	FKCAHDPTG
33364	CTYFSMMDP	33405	CYAIKQMDI	33446	DKVGYGCHE	33487	FKMQMQYDS
33365	CTYFSMTDR	33406	CYHTTNTSQ	33447	DQDGRCYKG	33488	FKVQMQYDN
33366	CTYKVNSWI	33407	CYHVYKDWS	33448	DQDGRCYKQ	33489	FRAGMIVSC
33367	CTYVSMMDR	33408	CYMGLSFHW	33449	DSEPSRGDH	33490	FRDAMIVSC
33368	CVAMKFSEH	33409	CYPFYKDWS	33450	DSQPSRGDY	33491	FRDGMIFSC
33369	CVEPYTDHQ	33410	CYPLYKDWS	33451	DYAHQRVDR	33492	FRDGMTVSC
33370	CVFQCNRSP	33411	CYPVSKDWS	33452	DYANERVDR	33493	FRDRPFTFT
33371	CVFQCYGSP	33412	CYPVYKAWS	33453	DYANKRVDR	33494	FRHGMIVSC
33372	CVFQCYRGP	33413	CYPVYKDWT	33454	DYANQRGDR	33495	FRNGMIVSC
33373	CVFQCYRNP	33414	CYPVYKNWS	33455	DYATQRVDR	33496	FTQQRCAQC

[2270]

33497	FTQQRCGQC	33538	HCCFKTGDM	33579	KGDRFMVDP	33620	LKHPAWKEE
33498	FTQRTTDHR	33539	HCCFKYGDM	33580	KGKPWGHHN	33621	LKHPIWKEE
33499	FTRQRCVQC	33540	HCCFRNGDM	33581	KHCHSMQFG	33622	LKHPNWKEE
33500	FVEPYTDHR	33541	HKVNLKGDE	33582	KHCRSMQFE	33623	LKHPPWKEE
33501	FVEPYTHHQ	33542	HKVTIFKDR	33583	KHCYSMQFE	33624	LKHPSWKEE
33502	FVQPYTDHQ	33543	HQLEFDFVG	33584	KIVCRVNYP	33625	LKHPTWREE
33503	FYWCCNREC	33544	HQMEFDFVG	33585	KKLNIQKQS	33626	LKHTTWKEE
33504	FYWCYTREC	33545	HRVTIFKDR	33586	KLCFKEDAF	33627	LKLPTWKEE
33505	FYWFCTREC	33546	HYCRSGDVQ	33587	KMCSSAKSY	33628	LKPPTWKEE
33506	GAAVIRKMQ	33547	IAAQIDDRE	33588	KMCTLTKAD	33629	MAANCEVFG
33507	GFAVIRKMQ	33548	IAAQSDDRQ	33589	KMFSSAKSY	33630	MANENRGIR
33508	GGWIMRSKR	33549	IAARSDDRE	33590	KMYPSAKSY	33631	MANENRGVH
33509	GIAVIRKMQ	33550	IADQSDDRE	33591	KMYSSVKSY	33632	MANENRVIH
33510	GILPSVKFQ	33551	IAVCVIRDN	33592	KQCHSMQFE	33633	MASENRGIH
33511	GILPYGKFQ	33552	ICARWHCGP	33593	KSAMPMNDD	33634	MATPERQMP
33512	GILPYVKVQ	33553	ICLLITKTE	33594	KSARCTLHM	33635	MAVNCEVFA
33513	GKCGFDVAA	33554	ICLRWHCGP	33595	KSCLYFDAC	33636	MCIRSGFVV
33514	GKFDMPSFF	33555	ICVRWHCGS	33596	KSCLYFDAF	33637	MCIRSGIIV
33515	GLAVIRKMQ	33556	ICVRWHCGT	33597	KSCLYFDSY	33638	MCLRSGFIV
33516	GLEHCRPGT	33557	ICWLFTKTE	33598	KSSWTDSFG	33639	MCVRSGFIV
33517	GLEWHTKNW	33558	ICWLIAKTE	33599	KSVRCKLHM	33640	MCWLITKTE
33518	GLQWHTKDW	33559	ICWLIKKTE	33600	KVCFKEDAI	33641	MEDRLAHYY
33519	GNETSDNFH	33560	ICWLIPKTE	33601	KVCFKEDGF	33642	MEDRVAHYH
33520	GSQWLNDSF	33561	ICWLISKTE	33602	KVCFKEDPF	33643	MKHPTWKEE
33521	GVAAIRKMQ	33562	ICWLITKNE	33603	KVCFKEDSF	33644	MKIPNIQTI
33522	GVAGIRKMQ	33563	ICWLITKTG	33604	KVCYKEDAF	33645	MKMPRWGMP
33523	GVAIIRKMQ	33564	ICWLITKTQ	33605	KWCLSIDYR	33646	MKMTRWGML
33524	GVALIRKMQ	33565	ICWSITKTE	33606	KWFLSIDYP	33647	MKMTRWGMQ
33525	GVAVFRKMQ	33566	ICWWITKTE	33607	KWFLSIDYS	33648	MKMTRWGMS
33526	GVAVIRKME	33567	ISACVIRDN	33608	KWFLSMDYR	33649	MKTPTIQTI
33527	GVAVIRKMR	33568	ISCRTQSDC	33609	KWFPSIDYR	33650	MKVTRWGMP
33528	GVEWHTKDW	33569	ISVCVIRDT	33610	KWFVSIDYR	33651	MQNACFKNP
33529	GVFPNMFAH	33570	IVQQIRDFH	33611	KWVLSIDYR	33652	MSVCVIRDN
33530	GVFPNMFSP	33571	IWECPRAGS	33612	KYARCKLHM	33653	MVAHNRLAN
33531	GVTVIRKMQ	33572	IWECPRLGS	33613	LAAQSDDRE	33654	MVANCEVFA
33532	GVVVIRKMQ	33573	KADRFMVDQ	33614	LAARIGEVK	33655	MVEHNRLAD
33533	GVWIVRSKR	33574	KAEHEEMWN	33615	LAARIGKVE	33656	MVEHNRLAH
33534	GYANQRVDR	33575	KAEYFDASA	33616	LAARSGEVE	33657	MVEHNRLDN
33535	HCCFKNGDL	33576	KCAFNGIDE	33617	LCVRWHCGP	33658	MVEHNRLGN
33536	HCCFKNVDM	33577	KCDYRNMHF	33618	LCWLITKTE	33659	MVEHNRVAN
33537	HCCFKSGDM	33578	KELLCSDDA	33619	LGIRNMELR	33660	MVVYREHMG

[2271]

33661	MVVYRENMA	33702	QCDYRDMHF	33743	RINPDIIAP	33784	RSCFEMKHL
33662	MVVYRETMG	33703	QCDYRNMDF	33744	RISHDIIAP	33785	RSCFEMRHI
33663	MYAERNADM	33704	QCDYRNMNF	33745	RISPDIIAL	33786	RSCFVMKHI
33664	MYVWRNDFA	33705	QCHYRNMHF	33746	RISPDIIAS	33787	RSCYEMKHI
33665	MYVWRNDGA	33706	QCNYRNMHF	33747	RISPDIIGP	33788	RSMWPPSSV
33666	NAIRIFLFG	33707	QCVYRNMHF	33748	RISPDILAP	33789	RSVTPCHSN
33667	NCCFKNGDM	33708	QFDYRNMHF	33749	RISPDTIAP	33790	RSVTPCHYS
33668	NCIHNWKHA	33709	QGGRILTSA	33750	RISPEIIAP	33791	RSVTPCHYT
33669	NCIHNWQHT	33710	QGKSRDVLD	33751	RISPNIIAP	33792	RSVTPCHYY
33670	NFFYMDKLS	33711	QGNDAIVWQ	33752	RISRDIIAP	33793	RTHCHAKAD
33671	NFQHVRMLE	33712	QHCNHDQGV	33753	RIVCRVNYP	33794	RTHGNAKAD
33672	NFYETLKRI	33713	QKIPVNDHQ	33754	RKHGLLFMG	33795	RTHGPAKAD
33673	NGIRIFLFA	33714	QKLDIQKQS	33755	RKHGLLFMT	33796	RTSWKWEYY
33674	NGIRIFLFV	33715	QKLKIQKQS	33756	RKHGLLFMV	33797	RVAVTNHRY
33675	NGIRNFLFG	33716	QKLNIKKQS	33757	RKHGMLFMA	33798	RVCFKEDAF
33676	NGIRSFLFG	33717	QKLNIQKQP	33758	RKHPTWKEE	33799	RVDATNHRY
33677	NGLRIFLFG	33718	QKLNMQKQS	33759	RKLNIQKQS	33800	RVDVTNHRF
33678	NGVRIFLFG	33719	QKLNVQKQS	33760	RLHGMGDVP	33801	RVEPTTVAD
33679	NKCAHDPTG	33720	QKLPVNDHQ	33761	RMAHNFWDA	33802	RVESQWDGC
33680	NKCGFDVAA	33721	QKVHVNDHQ	33762	RMAHTFWNA	33803	RVQSQWHGC
33681	NSCRRTFHH	33722	QKVNIQKQS	33763	RMAPTFWDA	33804	RWFLSIDYR
33682	NYANQRVDR	33723	QKVPLNDHQ	33764	RMARTFWDA	33805	SADPRVDCP
33683	NYCRRTFHL	33724	QKVPVSDHQ	33765	RMHVLVIYT	33806	SADSETPWF
33684	NYCRRTFHP	33725	QKVPVTDHQ	33766	RMSPDIIAP	33807	SCWLITKTE
33685	NYCRRTFHQ	33726	QKVQVNDHQ	33767	RMYSSAKSY	33808	SEHCYRHQH
33686	NYCRRTFHR	33727	QMCTLTKTD	33768	RPCFEMKHI	33809	SEWMIWKCR
33687	NYCRRTFHY	33728	QQVTIFKDR	33769	RQEWRSIWK	33810	SGCFWRHTS
33688	NYCRRTFYH	33729	QSARYIDFT	33770	RQKLRSIWK	33811	SGCFWRNAS
33689	PANHARNPE	33730	QTKGRDGAC	33771	RQKWRIIWK	33812	SGCFWRNTP
33690	PANHARNPQ	33731	QVKSRDVLD	33772	RQKWRNIWK	33813	SGCFWRSTS
33691	PANHARNPR	33732	QWTRWLIVT	33773	RQKWRSFWK	33814	SGCFWRTTS
33692	PANHARTPK	33733	RACFEMKHI	33774	RQKWRSIWE	33815	SGIRIFLFG
33693	PVCPTRYAD	33734	RADVTNHRY	33775	RQKWRSIWQ	33816	SHCLYWMRH
33694	PVNHARNPK	33735	RAQSQWDGC	33776	RQKWRSIWR	33817	SHCMYWMKH
33695	QAESRDVLD	33736	RASCGISMA	33777	RQKWRSIWT	33818	SHCVYWMKH
33696	QAKSRDGLD	33737	RAVTPCHYN	33778	RQKWRSMWK	33819	SHDVRMHML
33697	QAKSRDVLA	33738	RCCFKNGDM	33779	RQKWRSVWK	33820	SHFLYWMKH
33698	QAKSRDVLN	33739	RCVQPQKEA	33780	RQQWRSIWK	33821	SIVCRVNYP
33699	QAKSRDVMD	33740	RCVQPQKKG	33781	RQRWRSIWK	33822	SKCGFDVAD
33700	QCAYRNMHF	33741	RDLMYTDKA	33782	RQTWRSIWK	33823	SKCGFDVAG
33701	QCCFKNGDM	33742	RHKWRSIWK	33783	RSATPCHYN	33824	SKCGFDVAP

[2272]

33825	SKCGFDVAV	33866	TIVCRVNYS	33907	WAEYRKTHA	33948	YFNSWIRNH
33826	SKCGFNVAA	33867	TKKKLQCES	33908	WAEYRKTHH	33949	YFQSMGHMW
33827	SKCSFDVAA	33868	TKKKMQCEA	33909	WAEYRQTHD	33950	YFRAMGHMW
33828	SKRDMDPAP	33869	TKKKMQCEY	33910	WAIRNMELR	33951	YFTGWIRNH
33829	SLGNRNADY	33870	TKKKMQCQS	33911	WAKYRKTHD	33952	YGCFWRNTS
33830	SLQNRNADY	33871	TKQKMQCES	33912	WEHRFPASM	33953	YHCQNYVAS
33831	SLRDRNADY	33872	TLVCRVNYP	33913	WFAARQHFP	33954	YIALRATFQ
33832	SMDTSDWWE	33873	TMKLGDMDG	33914	WFIARQHFP	33955	YIDKIAGWL
33833	SQWMIWKCQ	33874	TSEWRTDDR	33915	WFVARKHFP	33956	YIGLRATFE
33834	SVCPNRYAD	33875	TVAMVMDFC	33916	WFVARQHFA	33957	YIGLRATFK
33835	SVCPTRHAD	33876	TVEHNRLAN	33917	WFVARQHFT	33958	YIGLRATFP
33836	SVCPTRYAE	33877	TVVCRVNYP	33918	WFVARQNFP	33959	YIGVRATFQ
33837	SVCPTRYAH	33878	VAAQSDDRE	33919	WFVGRQHFP	33960	YINKIAGWW
33838	SVCPTRYAN	33879	VAQQIRDFH	33920	WGIRNMELG	33961	YIPDRDGCD
33839	SVRNRNADY	33880	VAQRTTDHR	33921	WGIRNMELQ	33962	YIVLRATFQ
33840	TAAWRTDDR	33881	VCMHLRQGG	33922	WGIRNMEPR	33963	YKCAHDPNG
33841	TADPRVDCP	33882	VCMHVRQGA	33923	WGIRNMGLR	33964	YKCAHDPTE
33842	TADRWFDAS	33883	VCMPVRQGG	33924	WGIRNMKLR	33965	YKCAYDPTG
33843	TADRWGDAS	33884	VCWLITKTE	33925	WGIRNMVLR	33966	YKCGHDPTG
33844	TAELRTDDR	33885	VCWMITKTE	33926	WGIRNTELR	33967	YKCPHDPTG
33845	TAEPKMDHE	33886	VCWVITKTE	33927	WGNRNMELR	33968	YKCSHDPTG
33846	TAEWRNDDR	33887	VDDKHITGY	33928	WGSRNMELR	33969	YKCTHDPTG
33847	TAEWRTDAR	33888	VDDKNITCY	33929	WGVRNMELR	33970	YKCVHDPTG
33848	TAEWRTDDW	33889	VDDKYITCY	33930	WLVCRCAWE	33971	YKFDMPSFF
33849	TAEWRTDYR	33890	VEDKHITCY	33931	WSAKSTKKQ	33972	YKGAHDPTG
33850	TAGWRTDDR	33891	VEPFPCDWS	33932	WSAKTTKEQ	33973	YKGGYGCHE
33851	TAKWRTDDR	33892	VGGDYFRSD	33933	WSIRNMELR	33974	YLGLRATFQ
33852	TCILRQSPT	33893	VKILSEMMT	33934	WVVCRCDWE	33975	YQDGRCYKE
33853	TCLLRQSQT	33894	VKLLSEMMN	33935	WVVCRCSWE	33976	YSGQHFRTL
33854	TCPRKDDDM	33895	VKMLSEMMN	33936	WVVCRCTWE	33977	YTMCKANWN
33855	TCTRKDDDK	33896	VPVDNLWND	33937	YAARFACYQ	33978	YVHDRDGCD
33856	TCTRKDDDT	33897	VVAVIRKMQ	33938	YACAIQQHL	33979	YVPDRAGCD
33857	TCWLITKTE	33898	VVCMNHCDM	33939	YADRFAGYQ	33980	YVPDRDGCA
33858	TFVCRVNYP	33899	VVEHNRLAN	33940	YANLGREHN	33981	YVPDRDGCC
33859	TGDRFMVDQ	33900	VVEPYTDHQ	33941	YAPDRDGCD	33982	YVPDRDGCG
33860	TGGRIGDSQ	33901	VVQQIRDFN	33942	YDEQWNDNS	33983	YVPDRDGYD
33861	TIVCRINYP	33902	VVQQIRDFP	33943	YDIMNISPF	33984	YVPDRDVCD
33862	TIVCRVDYP	33903	VVQQIRDFR	33944	YDIMTISPC	33985	YVPDRNGCD
33863	TIVCRVHYP	33904	VVQQLRDFH	33945	YDMMTISPF	33986	YVPDRVGCD
33864	TIVCRVNYH	33905	VVQQSRDFH	33946	YENLAREHN	33987	YVPYRDGCD
33865	TIVCRVNYR	33906	VWECPRVGS	33947	YFASWIRNH	33988	YVRDRDGCD
33989	YVTDRDGCD	33990	YYCRRTFHH				

[2273] 实施例41

[2274] 如通过机器学习鉴定的具有脊髓组织嗜性的AAV5变体

[2275] 此实施例描述了如通过机器学习方法鉴定的展现脊髓组织嗜性的工程化AAV5变

体的优选特性。进行如实施例19和图19至图22中描述的相同生物信息学预处理、机器学习方法和分析以确定展现脊髓组织嗜性的工程化AAV5变体的优选特性。

[2276] 下文描述了AAV5 VP衣壳多肽的581至589区域中各位置处的有利生物物理特性和有利氨基酸残基，它们与较高的脊髓组织嗜性概率相关。表66中描述的任何以下机器学习衍生的位置氨基酸偏好可以单独存在或彼此组合存在或与实施例10，表19中观察和描述的任何氨基酸偏好组合存在。下文表66中列出了从ML模型中提取的排名靠前的重要脊髓嗜性特征之间共有的位置特征的概述。

[2277]

表 66
机器学习衍生的脊髓组织嗜性规则

Xaa1 处高体积 Xaa1 选自 F、W 或 Y
Xaa1 处低可突变性 Xaa1 选自 Y、F、L 或 C
Xaa1 处高溶解性 Xaa1 选自 W、F、I 或 L
Xaa1 处低平均柔性 Xaa1 选自 F、M 或 W
Xaa2 处低亲疏水性 Xaa2 选自 P 或 Y
Xaa3 处低亲水性 Xaa3 选自 Y、W、V、M、F、I 或 L
Xaa3 处高溶解性 Xaa3 选自 W、F、I 或 L

[2278]

表 66 机器学习衍生的脊髓组织嗜性规则
Xaa6 处高体积 Xaa6 选自 W、R、K、M、I 或 L
Xaa6 处高摩尔质量 Xaa6 选自 W
Xaa8 处高摩尔质量 Xaa8 选自 W、E、K、M、H 或 Q
Xaa8 处高体积 Xaa8 选自 W、K、M、I 或 L
Xaa8 处高 goldman engelman steitz Xaa8 选自 V 或 L
Xaa9 处高亲疏水性 Xaa9 选自 V 或 I
Xaa9 处高溶解性 Xaa9 选自 W、F、I 或 L

[2279] 下文表67提供了表现出脊髓组织嗜性并且符合表66中提供的一种或多种规则的变体AAV VP1衣壳多肽的581至589区域的序列。因此，本公开提供了由具有SEQ ID NO:2的工程化AAV5 VP1衣壳多肽组成的rAAV，其中Xaa1至Xaa9区域是SEQ ID NO:35438-SEQ ID NO:36437中的任一者，如表67中所公开。本文还涵盖由工程化AAV5 VP2衣壳多肽和工程化AAV5 VP3衣壳多肽组成的rAAV，所述多肽在对应于AAV5 VP1 581至589区域中的氨基酸的AAV5VP2(氨基酸残基445至453)和AAV5 VP3(氨基酸残基389-397)中的区域处具有下文表中公开的序列。

[2280] 表67

[2281] 驱动脊髓组织嗜性的AAV5 VP1衣壳多肽中581至589区域的序列

[2282]

SEQ ID NO	581-589 序列
35438	ACLLFLVGI
35439	ACYWLLNWF
35440	ACYWWLNWF
35441	ACYWWMNWF
35442	ADHPENFSD
35443	ADWLRRNIW
35444	AEFRPKISR
35445	AFIQQNNQD
35446	AFIQQSDQD
35447	AFIQQSNED
35448	AFIQQSTQD
35449	AFIQRSNQD
35450	AFVQQSNQD
35451	AFWYSMAFP
35452	AGDRTISMW
35453	AHFRAYTQY
35454	AHKWFCEGS
35455	AHKWVCEDS
35456	AKNVFEEHV
35457	ALFAPNGEV
35458	ALFYMEKMP
35459	AMVLFLVDI
35460	AMWIRGIWG
35461	ANDMCNYYG
35462	AQGNKTYVI
35463	AQKKCLSFG
35464	AQKWFCEDS
35465	AQQKCLSFG
35466	AQVNKTYLI
35467	AQVNKTYVI
35468	AQVTKTYVI
35469	ATMFSVIHQ
35470	ATWVHYDMY
35471	AVIQQSNQD
35472	AWFDMEKMP
35473	AWFHMEKMP
35474	AWFNMEKMP
35475	AWFYKEKMP
35476	AWFYMEEMP
35477	AWFYMEKMS
35478	AWFYMENMP
35479	AWFYMGKMP
35480	AWFYREKMP
35481	AWLQGGIYN
35482	AWSYMEKMP
35483	CAVKQECSR
35484	CCCSEIGVN
35485	CHVMSVPCA
35486	CKNWYINMH
35487	CQWTPEIFQ
35488	CTQCYGDMP
35489	CWAEFRMRD
35490	CWGEFRMRV
35491	CWTEFRMRV
35492	CWVEFRMRV
35493	CWWTPEIFQ
35494	DAKSPCDSV
35495	DAKSPCESF
35496	DAVNDLTWP
35497	DCIPDWSST
35498	DCYWWLNWF
35499	DEANKTYVI
35500	DEDNKTYVI
35501	DEVNKTYVI
35502	DFILDWSST
35503	DFIPDWSSI
35504	DFIRDWSST
35505	DFITDWSST
35506	DFVPDWSST
35507	DFWYGMAFP
35508	DFWYSMACP
35509	DFWYSMAVP
35510	DFWYSMTFP
35511	DFWYSRAFP
35512	DFWYSVAFP
35513	DFWYTMAFP
35514	DHVNKPYVI
35515	DIIPDWSST
35516	DIRIVTDGR
35517	DIWYSMAFP
35518	DKGNKTYVI
35519	DKNVFEEDV
35520	DKNVFEQHV
35521	DKNVYEEHV
35522	DNVNKTYVI
35523	DPGNKTYVI
35524	DPVNKTYVL
35525	DPVNKTYVV
35526	DPVTKTYVI
35527	DQAEVLSDH
35528	DQANKTYVI
35529	DQDNKTYVI
35530	DQDTKTYVI
35531	DQGNKTYVI
35532	DQGSKTYVI
35533	DQINKPYVI
35534	DQINKTYVI
35535	DQINKTYVS
35536	DQLNKTYVI
35537	DQVDKTYVI
35538	DQVNKNYVI
35539	DQVNKPYEI
35540	DQVNKTHVI
35541	DQVNKTNVI
35542	DQVNKTSVI
35543	DQVNKTYAI
35544	DQVNKTYEI
35545	DQVNKTYML
35546	DQVNKTYVF
35547	DQVNKTYVN
35548	DQVNKTYVV
35549	DQVNMTYVI
35550	DQVNNTYMI

[2283]

35551	DQVNQTYVI	35592	EMKCMGAGI	35633	FIIRCMVIW	35674	FRNWYINMH
35552	DQVNRTYVI	35593	EMKCTGAGI	35634	FIIRCVLIW	35675	FRNWYMNMH
35553	DQVSKTYVI	35594	EMKWLGAGI	35635	FIIRCVLVW	35676	FTFGGLAFA
35554	DQVTKTYVI	35595	EMKYIGAGI	35636	FIIRGMLIW	35677	FTIGGLAFA
35555	DQVTQTYVI	35596	EMWIGGIWG	35637	FIIRWMLIW	35678	FTNTKKQAF
35556	DRNVFEEHV	35597	EMWIRGVWG	35638	FILRCLLIW	35679	FTNWYINMD
35557	DRVNKTYLI	35598	EMWVRGVWG	35639	FILRCMLIC	35680	FTNWYINMH
35558	DRVNKTYVS	35599	EQEIKNSTQ	35640	FILRCMLIW	35681	FTNWYINMN
35559	DVNLKITTH	35600	EQELKDSTQ	35641	FIVGGLAFA	35682	FTVAGLAFA
35560	DVNLQIATH	35601	EQELKNSTL	35642	FIVRCMLIW	35683	FTVGGLAFG
35561	DVYYCPEAI	35602	EQELKTSTQ	35643	FKFISYIGD	35684	FTVGGLAFP
35562	DWMQGGIYH	35603	EQEVKNSTQ	35644	FKHWYINMH	35685	FTVGGLAFT
35563	DWMQGGIYY	35604	EQQLKNSTQ	35645	FKICFSPTV	35686	FTVGGLAFV
35564	DYAQYYRTW	35605	ERFRAYTQY	35646	FKNGYINMH	35687	FTVGGLGFA
35565	DYDPDIKDP	35606	ETFHNTATI	35647	FKNHYEGGP	35688	FTVGGLSFA
35566	DYDPDLKDQ	35607	ETWVHYDMY	35648	FKNLYINMH	35689	FTVGGLTFA
35567	DYGPDIKDQ	35608	EYGTKVKMD	35649	FKNWFINMH	35690	FTVGGMAFA
35568	ECLLFLIGI	35609	EYSTKVKMD	35650	FKNWHINMH	35691	FWCHIIGIY
35569	ECMLFLVGI	35610	EYVTKVKMD	35651	FKNWSINMH	35692	FWCNIIGIY
35570	ECVSKWIVF	35611	FAEASITRW	35652	FKNWYIHMH	35693	FWGNIIGIY
35571	EDWIRRNIW	35612	FASAFANTD	35653	FKNWYINMD	35694	FWSNFIGIY
35572	EDWLRCNIW	35613	FASAFSNSD	35654	FKNWYINMF	35695	FWSNIIGFY
35573	EDWLRRDIW	35614	FASAFVNSD	35655	FKNWYINML	35696	FWSNIIGIC
35574	EFIQQSNQD	35615	FASVFANSD	35656	FKNWYINMN	35697	FWSNIIGIH
35575	EHFRAYAQY	35616	FCMEWEVMD	35657	FKNWYINMP	35698	FWSNIIGIN
35576	EHFRAYIQY	35617	FFMAWEVMD	35658	FKNWYINMR	35699	FWSNIIGLY
35577	EHFRDYTQY	35618	FFMDWEVMD	35659	FKNWYINMY	35700	FWSNIIGMY
35578	EHFRSYTQY	35619	FFMEWEGMD	35660	FKNWYINRH	35701	FWSNIIGVY
35579	EHWLRRNIW	35620	FFMEWELMD	35661	FKNWYINTH	35702	FWSNIMGIY
35580	EIIPQFAYF	35621	FFMEWEVKD	35662	FKNWYINVH	35703	FWSNLIGIY
35581	EINREVREN	35622	FFMEWEVRD	35663	FKNWYISMH	35704	FWSNMIGIY
35582	EINRVVRQN	35623	FFMEWEVTD	35664	FKNWYLHMH	35705	FWSNVIGIY
35583	EISREVRQN	35624	FFMEWEVVD	35665	FKNWYNNMH	35706	FWSSIIGIY
35584	EKELKNSTQ	35625	FFMEWKVMD	35666	FKNWYVNMH	35707	FWSTIIGIY
35585	ELIYKFKTD	35626	FFMEWQVMD	35667	FKSWYINMH	35708	FWTNIIGIY
35586	ELIYKIETD	35627	FFYECYSNK	35668	FKVGGLAFA	35709	FYMEWEVMD
35587	ELIYKMKTD	35628	FIIGCMLIW	35669	FKYWYINMH	35710	GCYWWLNWF
35588	ELVNKTYVI	35629	FIIHCMLIW	35670	FLIRCMLIW	35711	GDEYYDIMG
35589	ELVYKIKTD	35630	FIIRAMLIW	35671	FMIRCMLIW	35712	GDHTENFSD
35590	EMINACFQT	35631	FIIRCMLIL	35672	FPEASLTRW	35713	GEFRPKISR
35591	EMKCIGAAI	35632	FIIRCMLVW	35673	FQNWYINMH	35714	GFIQQSNQD

[2284]

35715	GHVMTVPCA	35756	IRNMFEQYT	35797	KWHGGPGVM	35838	MKYQDTHTI
35716	GIKYKDILS	35757	IRSHTLTYC	35798	KWHGGQGVV	35839	MKYWDTHTF
35717	GINREVRQN	35758	IRSQTITDC	35799	KWHGGRGVV	35840	MYFFVWRFM
35718	GKNVFEEHV	35759	IRSQTLTDW	35800	KWLGGPGVV	35841	MYIFGWRFM
35719	GQGNKTYVI	35760	IRSQTLTDY	35801	KWLYGWPIK	35842	MYIFVWGFM
35720	GSASRNQRD	35761	IRSQTLTNC	35802	KWYGGPGVV	35843	MYIFVWRFT
35721	GSSPGKGWM	35762	IRSQTVTDC	35803	KYNMMYHMD	35844	MYIFVWRIM
35722	GSSSRNQRD	35763	IRSRTLTDC	35804	KYNMRYHMD	35845	MYIFVWRVM
35723	GTWVHYDMF	35764	IRWLRLHKK	35805	KYNMVYHMD	35846	MYIFYWRFM
35724	GWAEFRMRV	35765	ISWIIRADI	35806	KYNVLYHMD	35847	MYIIVWRFM
35725	GWAEFRVRV	35766	ISWINRADI	35807	KYSMLYHMD	35848	MYISVWRFM
35726	GWSEFRMRV	35767	ISWITRADV	35808	LATMSLTWW	35849	MYIVVWRFM
35727	HPVNKTYVI	35768	ISWITRAGI	35809	LCYWWLNWF	35850	MYIYFWRFM
35728	HQGNKTYVI	35769	ISWITRTDI	35810	LEKWGHMHT	35851	MYIYVWRFM
35729	HRFQEKGSR	35770	ITSQTLTDC	35811	LKICFGPTV	35852	MYLFVWRFM
35730	HRFQEKGTH	35771	ITWLHYDMY	35812	LKIGFSPTV	35853	NANDVQWYD
35731	HRFREKGSH	35772	IWFYMEKMP	35813	LPEIYLSFC	35854	NAPDVQWYD
35732	HTLLWYGIW	35773	IWSNIIGIY	35814	LPKRYLSFC	35855	NATDFQWYD
35733	HTLLWYGIY	35774	IWYWWLNWF	35815	LPQIYLSFC	35856	NATDGQWHD
35734	HVYCCPEAI	35775	KAVWDPSSM	35816	LPTIYLSFC	35857	NATDVEWYD
35735	ICHCCMAWG	35776	KCWDCHDAG	35817	LQICFSPTV	35858	NATDVKWYD
35736	ICHCCMVWG	35777	KCWDCHEGG	35818	LRICFSPTV	35859	NATDVLWYD
35737	ICHCCMYWG	35778	KEDFGWKHV	35819	LRWVRLHKK	35860	NATDVRWYD
35738	ICHWCMDWG	35779	KESRSQTGD	35820	LSKIYLSFC	35861	NCDRSLAEY
35739	ICYCCMDWG	35780	KEVFGWKHM	35821	LSSCYYTTY	35862	NCHIDRVQD
35740	ICYWWLNCF	35781	KGVWDPGSM	35822	LSWITRADI	35863	NCHVDHVQD
35741	ICYWWLNWF	35782	KGVWDPSST	35823	LVFLRGQHY	35864	NCKNLTYYI
35742	IDRWWVIDV	35783	KHIGSYTQF	35824	LWFYMEEMP	35865	NFIPDWSST
35743	IHIGTNSYA	35784	KHNMLYHMD	35825	LWFYMEKMP	35866	NFWYSMAFP
35744	IHTATNSYA	35785	KHYKWGHGQ	35826	LWSNMIGIY	35867	NIDAHERYR
35745	IHTGTNSYV	35786	KINREVRQN	35827	MFRIEAVCP	35868	NKNVFEEHV
35746	IIYECYSNK	35787	KKWSCHRGH	35828	MKFQDTHTF	35869	NLCRHLFQC
35747	IKHIAAESR	35788	KKYQDTHTF	35829	MKRDAAFCA	35870	NLCRHLFQF
35748	IKHIATETR	35789	KMWIRGIWG	35830	MKREAAVCA	35871	NNYKFQQDG
35749	IKHISTESR	35790	KRIKMPPPR	35831	MKRETAFCP	35872	NNYKGQQDG
35750	IKNWYINMH	35791	KRIKMPPTK	35832	MKRETAFCS	35873	NNYKLQQDG
35751	IKNWYINMP	35792	KWHAGPGVV	35833	MKRETAFCT	35874	NNYKVQEDG
35752	IKSQTLTDC	35793	KWHCGPGVV	35834	MKRETAVCA	35875	NPAYGRGHI
35753	IMFHCREMG	35794	KWHGGPGAV	35835	MKRETGFCA	35876	NPELGKCEI
35754	IMVHCRQMG	35795	KWHGGPGDV	35836	MKYKDTHTF	35877	NPGWRYMWF
35755	IMVHCRVMG	35796	KWHGGPGFV	35837	MKYQDTHTC	35878	NPGWRYMWW

[2285]

35879	NQAYGRGHI	35920	PGNRTISMW	35961	RKWSCHRAH	36002	SLMRFLASI
35880	NQVNKTYVI	35921	PHKWFCEDS	35962	RKWSCHRCH	36003	SLMRFLHSI
35881	NRAYGRGHV	35922	PPDIGTQFI	35963	RKWSCHRGL	36004	SLMRFLPAI
35882	NSWITRADI	35923	PPHIGTEFI	35964	RKWSCHRGY	36005	SLMRFLPST
35883	NSWRHICDH	35924	PPHIGTKFI	35965	RKWSCHRVH	36006	SLMRFLPSV
35884	NVAAHERYR	35925	PPHLGTQFI	35966	RKWSCNRGH	36007	SLMRFLPYI
35885	NVCRHLFQF	35926	PPHMGTQFI	35967	RKWYCHRGH	36008	SLMRFLRSI
35886	NVDAHERYK	35927	PPHTGTQFI	35968	RPHIGTQFI	36009	SLMRFLSSI
35887	NVDAPERYR	35928	PPLIGTQFI	35969	RPVQEDGCG	36010	SLMRFLTSI
35888	NVDVHERYR	35929	PRFQEKGSH	35970	RQNDTHRAN	36011	SLMRFPPSI
35889	NVGAHERYR	35930	PRHIGTQFI	35971	RQVQRYWEC	36012	SLMRFVPSI
35890	NVYYCPEAI	35931	PWFYLEKMP	35972	RQWSCHRGH	36013	SLMRVLPSI
35891	NWCKHLFQF	35932	PWFYMEEMP	35973	RRFQEKGSH	36014	SLVRFLLSI
35892	NWCRDLFQF	35933	QCLLFLVGI	35974	RRIESYCIT	36015	SLVRFLPSI
35893	NWCRHLFEC	35934	QEKWGHKHT	35975	RRIKMPPPK	36016	SMHDSNFNV
35894	NWCRHLFEF	35935	QEKWGHMHA	35976	RTVKEDGCG	36017	SMHDYNFNI
35895	NWCRHLFKF	35936	QEKWGHVHT	35977	RVPADWACA	36018	SMHDYTFNV
35896	NWCRHLFKY	35937	QTIIGNVWV	35978	RWSYMKAGK	36019	SNDKCNYYG
35897	NWCRHLFQC	35938	QWFYMEKMP	35979	RWSYMQAGI	36020	SNDMCDYYG
35898	NWCRHLFQI	35939	RAHWSDIAD	35980	RWSYMQTGK	36021	SNDMCNHYG
35899	NWCRHLFQV	35940	RAPADWACA	35981	RWSYMQVGK	36022	SNDMCNSYG
35900	NWCRHLFQY	35941	RDGSNWLCI	35982	RWSYMRAGK	36023	SNNMCNYYG
35901	NWCRHLFRF	35942	RDPADWACG	35983	RYIFVWRLM	36024	SPYWLMRMG
35902	NWCRHLIQF	35943	RDPADWACS	35984	RYNMLYHMD	36025	SQDCGRPEI
35903	NWCRHLVQF	35944	RDPADWTCA	35985	RYPADWACA	36026	SRMRFLPSI
35904	NWCRHRFQF	35945	RDPADWVCA	35986	SAAWQFGES	36027	SRMRLLPSI
35905	NWCRILFQF	35946	RDPSDWACA	35987	SAAWQFGVS	36028	SSASVNHTS
35906	NWCRLLFQF	35947	RDPTDWACA	35988	SDDMCNYYG	36029	SSDMCNYYG
35907	NWCRNLFQF	35948	RDTADWACA	35989	SDMIKLAEI	36030	SSFYMEKMP
35908	NWCRQLFQV	35949	REDFGWKHM	35990	SGADHYHLK	36031	SSSPGKGWK
35909	NWCRVLFHF	35950	RHIGSYTQF	35991	SGFYMEMMP	36032	SSWITRADI
35910	NWFRHLFQF	35951	RHIHRYWEC	35992	SGFYMEQMP	36033	STDMCNYYG
35911	NWGRHLFQF	35952	RHKWFCEAS	35993	SGFYMERMP	36034	SVADHYHRK
35912	NWWRHLFQF	35953	RHPADWACA	35994	SHKWFCEDS	36035	SVADHYNLK
35913	NYAQHYRTW	35954	RHVHPYWEC	35995	SHYWPMCWS	36036	SVADYYHLK
35914	NYSQYYRTW	35955	RHVHQYWEC	35996	SLFYMDKMP	36037	SVAHHYHLK
35915	PADRTISMW	35956	RHVHRFWEC	35997	SLFYMEKML	36038	SVARQMHSW
35916	PDDRTISMW	35957	RHVHRYWVC	35998	SLFYMEKMP	36039	SVMRFLPSI
35917	PGDRTISKW	35958	RIMHSWYND	35999	SLLRFQPSI	36040	SVSDHYHLK
35918	PGDRTITMW	35959	RKEQDHWDK	36000	SLMLFLPSI	36041	SWCFMEKMP
35919	PGDRTLSMW	35960	RKVESYCIT	36001	SLMRCLPSI	36042	SWCRHLFEF

[2286]

36043	SWCRHLFQF	36084	SWFYVEKIP	36125	THKWFCEDT	36166	TYAQYYRTW
36044	SWCYIEKMP	36085	SWFYVEKMP	36126	THKWFCEHS	36167	VAFRPKISR
36045	SWCYMEEMP	36086	SWIYMDKMP	36127	THKWFCENI	36168	VCCWWLNWF
36046	SWCYMEQMP	36087	SWIYMEKMT	36128	THKWFCENS	36169	VCSWWLNWF
36047	SWCYMVKMP	36088	SWIYMEQMP	36129	THKWFCEVS	36170	VCYWWLDWF
36048	SWFCVEKMP	36089	SWTYMEKMP	36130	THKWFCEYS	36171	VCYWWLHWI
36049	SWFDMEKMP	36090	SWVCMEKMP	36131	THKWFCGDI	36172	VCYWWLIWF
36050	SWFFMEKMP	36091	SWVYMEKMQ	36132	THKWFCQDS	36173	VCYWWLNLF
36051	SWFFMVKMP	36092	SWVYMEKMS	36133	THKWFCVDS	36174	VCYWWLNWC
36052	SWFFVEKMP	36093	SWVYMEMMP	36134	THKWICEDI	36175	VCYWWLNWI
36053	SWFHMEKMP	36094	SWVYMEQMP	36135	THKWVCEDS	36176	VCYWWLNWY
36054	SWFHMEKMQ	36095	SWVYMKKMP	36136	THRWFCEDS	36177	VCYWWLSWL
36055	SWFNMEQMP	36096	SWYCMEKMP	36137	THRWFCEES	36178	VCYWWLTWF
36056	SWFSKEKMP	36097	SWYYDERSN	36138	THRWLCEDS	36179	VCYWWMNWF
36057	SWFSVEKMP	36098	SWYYMEEMP	36139	TLMRFLPSI	36180	VEFGPKISR
36058	SWFYIERMP	36099	SWYYMEKMP	36140	TPKWFCEDN	36181	VEFKPKISR
36059	SWFYKETMP	36100	TAEKFICFP	36141	TPKWFCEDS	36182	VEFRHKISR
36060	SWFYMAQMP	36101	TCARSLAEY	36142	TPKWVCEDS	36183	VEFRPEISR
36061	SWFYMDKMT	36102	TCDRSLAQY	36143	TRKWFCEDS	36184	VEFRPKFSR
36062	SWFYMDMMP	36103	TCDRSLGEY	36144	TSERFICFP	36185	VEFRPKLSR
36063	SWFYMEAMP	36104	TCDRSLPEY	36145	TTDHFICFP	36186	VEFRPKMSR
36064	SWFYMEDMP	36105	TCDRSLTEY	36146	TTELYICFP	36187	VEFRPKTSR
36065	SWFYMEEVP	36106	TCDRSLVEY	36147	TWCRHLFQF	36188	VEFRPKVSR
36066	SWFYMEGMP	36107	TCDRTLAEY	36148	TWFHMEKMP	36189	VEFRPRISR
36067	SWFYMEKIR	36108	TCKNLTYSI	36149	TWFYMDKMP	36190	VEFRPTISR
36068	SWFYMEKMA	36109	TCKNMTYYI	36150	TWFYMEKMP	36191	VEFRRKISR
36069	SWFYMEKMI	36110	TCKSLTYYI	36151	TWFYMEKMQ	36192	VEFRSKISR
36070	SWFYMEKMR	36111	TCNRSLAEY	36152	TWFYMEQMP	36193	VEFRTKISR
36071	SWFYMEKMT	36112	TCRNLTYYI	36153	TWLCGWPIK	36194	VEFTPKISR
36072	SWFYMEKRP	36113	TCTNLTYYI	36154	TWLFGWPIK	36195	VEYRPKISR
36073	SWFYMEKVP	36114	TDKWFCVDS	36155	TWLGGCKSV	36196	VFIQQSNQD
36074	SWFYMEMMP	36115	THIASYTQF	36156	TWLYGWPFK	36197	VFYWWLNWF
36075	SWFYMERMP	36116	THICSYTQF	36157	TWLYGWPIQ	36198	VGYWWLNWF
36076	SWFYMETMP	36117	THIGSYTEF	36158	TWLYGWPIR	36199	VIWVHYDMY
36077	SWFYMEWMP	36118	THIGTYTQF	36159	TWLYGWPIT	36200	VKFRPKISR
36078	SWFYMKKMP	36119	THIVSYTQF	36160	TWLYGWPVK	36201	VKKCHTCGA
36079	SWFYMQKMP	36120	THKWCCEDS	36161	TWLYGWSIK	36202	VKNWYINMH
36080	SWFYMVKMP	36121	THKWFCDYS	36162	TWLYGWTIK	36203	VKWVRLHKK
36081	SWFYRDKMP	36122	THKWFCEDC	36163	TWMYGWPIK	36204	VNWVHYDMY
36082	SWFYREKLP	36123	THKWFCEDI	36164	TWPVGCKSV	36205	VQVNKTYVI
36083	SWFYVEEMP	36124	THKWFCEDN	36165	TWRGGCKSV	36206	VRWLRLHKK

[2287]

36207	VRWVRIHKK	36248	WDNWSNTPI	36289	WICKMWSWW	36330	WKSVCPSRQ
36208	VRWVRLHEK	36249	WDNWSNTQM	36290	WIHAMTYWC	36331	WKTVCPSRH
36209	VRWVRLHKQ	36250	WDNWSNTQV	36291	WIHMPHDRH	36332	WLCQMWSWW
36210	VRWVRLHKR	36251	WDNWSNTRI	36292	WINMAHDRH	36333	WLPAMTYWC
36211	VRWVRLHKT	36252	WDNWSTTQI	36293	WINMHHDRH	36334	WMLLGLTWF
36212	VRWVRLHRK	36253	WDNWTNTQI	36294	WINMPDDRH	36335	WMLLGMAWC
36213	VRWVRLHTK	36254	WDNWTNTQM	36295	WINMPHARH	36336	WMLLGMHWC
36214	VRYWWLNCF	36255	WDSWSNTQI	36296	WINMPHDGH	36337	WMLLGMNWC
36215	VSWITRADI	36256	WDVKFWESM	36297	WINMPHDKH	36338	WMLLGMPWC
36216	VTVGGLAFA	36257	WDVKYWETM	36298	WINMPHDRD	36339	WMLLGMTLC
36217	VTWMHYDMY	36258	WDVMFWETM	36299	WINMPHDRL	36340	WMLLGMTWF
36218	VTWVHYAMY	36259	WENWSNSQI	36300	WINMPHDRN	36341	WMLLGMTWG
36219	VTWVHYDMF	36260	WFCKKWSWW	36301	WINMPHDRP	36342	WMLLGMTWW
36220	VTWVHYDMH	36261	WFCKMLSWW	36302	WINMPHDRR	36343	WMLLGMTWY
36221	VTWVHYNMY	36262	WFCKMWGWG	36303	WINMPHDRY	36344	WMLLGTTWC
36222	VTWVHYVMY	36263	WFCKMWGWW	36304	WINMPHGRH	36345	WMLLGVTWC
36223	VTWVHYYMY	36264	WFCKMWIWW	36305	WINMPLDRH	36346	WMLLVMTWC
36224	VVFLRDQHY	36265	WFCKMWSGW	36306	WINMPNDRH	36347	WMLNQSRGW
36225	VWSNIIGIY	36266	WFCKMWSLW	36307	WINMPPDRH	36348	WMLQGMTWC
36226	VWYWWLNWF	36267	WFCKMWSWG	36308	WINMPRDRH	36349	WMMLGMTWC
36227	VWYWWLNWL	36268	WFCKMWSWL	36309	WINMTHDRH	36350	WMMMGLTWC
36228	VYYWWLNWF	36269	WFCKMWTWW	36310	WINRPHDRH	36351	WMNMPHDRH
36229	WANFKTLNR	36270	WFCKVWSWW	36311	WINVPHDRH	36352	WMPNQSRVW
36230	WANFQTLNK	36271	WFCNLWSWW	36312	WIPAKTYWC	36353	WMVLGMNWC
36231	WAPTPWRTH	36272	WFCQMWSWW	36313	WIPAMAYWC	36354	WMVLGMTWC
36232	WCCKMWSWW	36273	WFCRMWSWW	36314	WIPAMPYWC	36355	WNCSVYTCI
36233	WCMSQQYED	36274	WFCTLWSWW	36315	WIPAMTYLC	36356	WNNWSNTQI
36234	WDGKFWETM	36275	WFCTMWSWW	36316	WIPAMTYWF	36357	WRLLGMTWC
36235	WDIKFWETM	36276	WFFKMWSWW	36317	WIPAMTYWW	36358	WSCKMWSWW
36236	WDLKFWETM	36277	WFGKMWSWW	36318	WIPAMTYWY	36359	WTLLGMTWC
36237	WDMSQQYED	36278	WFNMPHDRH	36319	WIPARTYWC	36360	WVCKMWGWW
36238	WDNLRNTQI	36279	WFWKMWSWW	36320	WIPAVTYWC	36361	WVCKMWSWW
36239	WDNSSNTQI	36280	WFYKMWSWW	36321	WIPGMTYWC	36362	WVNLCADLI
36240	WDNWINTQI	36281	WGNWSNTQI	36322	WIPSMTYWC	36363	WVPAMTYWC
36241	WDNWSDTQI	36282	WGVKFWETM	36323	WIPTMTYWC	36364	WVPAPWRTH
36242	WDNWSHTQI	36283	WHCSVYTCS	36324	WIPVVTYWC	36365	WVPPPWRTH
36243	WDNWSITQI	36284	WHRDEMLTP	36325	WISAMTYWC	36366	WVPTHWRTH
36244	WDNWSNTEI	36285	WHREEMLTT	36326	WISMPHDRH	36367	WVPTPGRTH
36245	WDNWSNTKI	36286	WHVKFWETM	36327	WITAMTYWC	36368	WVPTPWQTH
36246	WDNWSNTKK	36287	WHWDEMLTT	36328	WITMPHDRH	36369	WVPTPWRSH
36247	WDNWSNTLI	36288	WIAAMTYWC	36329	WKNVCPSRQ	36370	WVPTPWRTN

[2288]

36371	WVPTPWRTQ	36388	YAQPDWFMT	36405	YLESKTIHP	36422	YTALMWVWA
36372	WVPTPWRTY	36389	YAQRDWFMA	36406	YPELGKGEI	36423	YTELGKCEI
36373	WVPTTWRTH	36390	YAWFYWADY	36407	YPELGQCEI	36424	YTLLWYGIC
36374	WVTLCAVLI	36391	YDQPDWFMA	36408	YPQPDWFMA	36425	YTNFEAVLI
36375	WVTTPWRTH	36392	YDYWPMCWS	36409	YPYWPMCWS	36426	YTNFEDALI
36376	WYMSKQYED	36393	YFMEWEVMD	36410	YSWFNWADY	36427	YTNFEDVII
36377	WYMSQEYED	36394	YFWYSMAFP	36411	YSWFYWADC	36428	YTNFEGVLI
36378	WYMSQQNED	36395	YHYWHMCWS	36412	YSWFYWADF	36429	YTNFENVLI
36379	WYMSQQSED	36396	YHYWPMCWA	36413	YSWFYWADS	36430	YTNIEDVLI
36380	WYMSQQYAD	36397	YHYWPMCWC	36414	YSWFYWAVY	36431	YTNLDDVLI
36381	WYMSQQYKD	36398	YHYWPMCWP	36415	YSWFYWPDY	36432	YTWFYWADY
36382	WYMSQQYQD	36399	YHYWPMCWT	36416	YSWFYWSDY	36433	YWSHIIGIY
36383	WYMSQRYED	36400	YHYWPMCWY	36417	YSWFYWTDY	36434	YWSNIIGIY
36384	WYMTQQYED	36401	YHYWPMGWS	36418	YSWFYWVDY	36435	YYAQYYRTW
36385	WYTSQQYED	36402	YIIRCMLIW	36419	YSWIYWADY	36436	YYWFYWADY
36386	WYVKFWETM	36403	YKNWYINMH	36420	YSWVYWADY	36437	YYWFYWSDY
36387	WYVSQQYED	36404	YLEAKTIHA	36421	YSWYYWADY		

[2289] 实施例42

[2290] 如通过机器学习鉴定的具有脾脏组织嗜性的AAV5变体

[2291] 此实施例描述了如通过机器学习方法鉴定的展现脾脏组织嗜性的工程化AAV5变体的优选特性。进行如实施例19和图19至图22中描述的相同生物信息学预处理、机器学习方法和分析以确定展现脾脏组织嗜性的工程化AAV5变体的优选特性。

[2292] 下文描述了AAV5 VP衣壳多肽的581至589区域中各位置处的有利生物物理特性和有利氨基酸残基，它们与较高的脾脏组织嗜性概率相关。表68中描述的任何以下机器学习衍生的位置氨基酸偏好可以单独存在或彼此组合存在或与实施例6，表11中观察和描述的任何氨基酸偏好组合存在。下文表68中列出了从ML模型中提取的排名靠前的重要脾脏嗜性特征之间共有的位置特征的概述。

[2293]

表 68 机器学习衍生的脾脏组织嗜性规则
Xaa1 处低溶解性 Xaa1 选自 D 或 P 或者 高溶解性 Xaa1 选自 F、I 或 L
Xaa1 处低亲疏水性 Xaa1 选自 Y 或 P
Xaa1 处低可突变性 Xaa1 选自 C、K 或 P
Xaa2 处低溶解性 Xaa2 选自 D、Q 或 R
Xaa2 处低亲疏水性 Xaa2 选自 D、E、R、K、H、N 或 Q
Xaa2 处低电荷 Xaa2 选自 D 或 E
Xaa2 处低体积 Xaa2 选自 T、N、P 或 D
Xaa2 处高平均柔性 Xaa2 选自 D、E、R、P、G、Q 或 S
Xaa3 处低溶解性 Xaa3 选自 D、E、P 或 N
Xaa3 处低亲疏水性 Xaa3 选自 D、E、H、N、Q 或 P
Xaa4 处低亲疏水性 Xaa4 选自 K 或 R
Xaa5 处低溶解性 Xaa5 选自 D、E、P 或 N

[2294]

表 68 机器学习衍生的脾脏组织嗜性规则
Xaa5 处高平均柔性 Xaa5 选自 D、E、R、P、G、Q 或 S
Xaa6 处低可突变性 Xaa6 选自 C
Xaa8 处高表面可接近性 Xaa8 选自 E、R 或 K
Xaa8 处低溶解性 Xaa8 选自 E、P、R、K、N 或 Q
Xaa8 处中等体积 Xaa8 选自 E、D、R、K、V、P、M、I、L、H、N、Q 或 T
Xaa9 处中等摩尔质量 Xaa9 选自 E、D、K、M、I、L、H 或 N

[2295] 下文表69提供了表现出脾脏组织嗜性并且符合表68中提供的一种或多种规则的变体AAV VP1衣壳多肽的581至589区域的序列。因此，本公开提供了由具有SEQ ID NO:2的工程化AAV5 VP1衣壳多肽组成的rAAV，其中Xaa1至Xaa9区域是SEQ ID NO:38438-SEQ ID NO:39437中的任一者，如表69中所公开。本文还涵盖由工程化AAV5 VP2衣壳多肽和工程化AAV5 VP3衣壳多肽组成的rAAV，所述多肽在对应于AAV5 VP1 581至589区域中的氨基酸的AAV5 VP2(氨基酸残基445至453)和AAV5 VP3(氨基酸残基389-397)中的区域处具有下文表中公开的序列。

[2296] 表69

[2297] 驱动脾脏组织嗜性的AAV5 VP1衣壳多肽中581至589区域的序列

[2298]

SEQ ID NO	581-589 序列
38438	ADDMCCSPN
38439	ADEDDCDDP
38440	ADEMEEEHG
38441	ADEMWPIRF
38442	ADHCNWSRT
38443	ADHCSWSRT
38444	ADMCECVKG
38445	ADNRTKPLC
38446	ADPDQKAMN
38447	ADQRDCWTT
38448	ADREPWPRN
38449	ADSRVRPKW
38450	AECKCCVKM
38451	AENQDDWQF
38452	AGDMWRHKM
38453	AGDTDMPED
38454	AGWEEAWCM
38455	AHDCDEPAL
38456	AIENNYCPM
38457	AIPHWDDKH
38458	APDFNREKN
38459	APDHWMEKM
38460	ARKCDFPAL
38461	CDAKRTPGM
38462	CDAWGELRG
38463	CDCKEQEED
38464	CDCKGRIPH
38465	CDDFCMCHC
38466	CDDHNHPRA
38467	CDDKQSHKE
38468	CDDKRFSKK
38469	CDDRERQQN
38470	CDDWGWDQM
38471	CDEIEMQGT
38472	CDELPMAGH
38473	CDEPHFWRC
38474	CDERRWDRP
38475	CDIRCENMG
38476	CDMCREWKW
38477	CDNKCWPHD
38478	CDNKPKNLS
38479	CDNRRNWWE
38480	CDPCNEFNK
38481	CDPMDACQE
38482	CDPNDMREH
38483	CDPRFVEAM
38484	CDPRTLPFY
38485	CDPWDWDIE
38486	CDPWDWEIE
38487	CDQFCEWRQ
38488	CDQFHWDKK
38489	CDRRCFDKH
38490	CDSCCGMHV
38491	CDTWPWSKK
38492	CDWCDQFRH
38493	CDYCDPRKF
38494	CDYRCLWYC
38495	CECYEPFEG
38496	CEDMDQWPW
38497	CELPDRCPQ
38498	CERVPVFSF
38499	CFDHRPLPC
38500	CFPCCWAKF
38501	CGAWWWNHM
38502	CGDEWAPRL
38503	CGDIDHPKS
38504	CGDWAWDRC
38505	CGEKCWHDI
38506	CGHDSKPRC
38507	CGHWCCSDY
38508	CGIEDKPEN
38509	CGLCRYYDM
38510	CGPCCLWGM
38511	CGPEPCVRQ
38512	CGPWNEPEY
38513	CGTHCPLRQ
38514	CGYDCQPRF
38515	CHKFDRELW
38516	CHNHPMPPL
38517	CKDEFWPLM
38518	CKDFRMSPL
38519	CKDYCRDEL
38520	CKPYEWWVM
38521	CKYTDDEEQ
38522	CMTDWSGMM
38523	CNCKDYVEV
38524	CNDQHMEKD
38525	CNPPEEPVR
38526	CNTQPSPKD
38527	CPEYRMARF
38528	CPPPCSYLL
38529	CPQWCEYYE
38530	CQCDDNWEI
38531	CQERERGKW
38532	CQSECAKRG
38533	CRDCRDGEN
38534	CRDEGLEYE
38535	CRDGDEPRG
38536	CRDWGWICA
38537	CRNDNIERK
38538	CRPLDWKPD
38539	CRYWCSYKD
38540	CWDGCHRRL
38541	CWDHGAPEL
38542	CWEDCQYHE
38543	CYDDWQDRI
38544	DCDMFCMFN
38545	DDARRCHRK
38546	DDCFQRCRN
38547	DDDCNYTRN
38548	DDDEDLNYD
38549	DDDHPRPRS
38550	DDDNWHHKN
38551	DDDPGDAYW
38552	DDECEMYYN
38553	DDEDGGHYM
38554	DDEPKGIRM
38555	DDEVDGDHP
38556	DDGCNYTRN
38557	DDGPDTNRK
38558	DDGPDTYRR
38559	DDGPETYRK
38560	DDGVMDSVD
38561	DDHEEGFKA
38562	DDHQENTKT
38563	DDKGPRNSM
38564	DDKHDRIHH
38565	DDLKCWNNY
38566	DDMHDNMVH
38567	DDNCNYTRN
38568	DDNHPTPTL
38569	DDNREPQDT
38570	DDNWDVFRD
38571	DDPCVEFET
38572	DDPFDVTLQ
38573	DDPFWCQRL
38574	DDPQEVHGF
38575	DDPQRKGYC
38576	DDPWGDRHT
38577	DDQLEGWID
38578	DDQPDNNIY
38579	DDQRLLDDI
38580	DDQTDYPPT
38581	DDQYRIEQH
38582	DDRANVPKH
38583	DDSGPKSQI
38584	DDSIEHFKG
38585	DDSWPIPLG
38586	DDTEPMMNK
38587	DDTEPMPVR
38588	DDTNEIEEY
38589	DDWKGWLRF
38590	DDWKPCEFA
38591	DDWMRPKRC
38592	DDYGHDDKN
38593	DECSESMRN
38594	DEGEEEPKL
38595	DEHKDPDEA
38596	DENKDEFWM
38597	DEQCRTHRK
38598	DEQHGGEMD

[2299]

38599	DESCEWMNQ	38646	DQWNDECFF	38693	EMHPCFWIF	38740	FKGCSGELV
38600	DESEERIQD	38647	DRERWWPET	38694	EPSRCRPHV	38741	FKMCPCVKG
38601	DESSDDDEE	38648	DRKFERQTM	38695	EQKQEPEFW	38742	FKMSDILPH
38602	DGCFDAQIN	38649	DRKKPVPEW	38696	ERCEGIQFC	38743	FKPCMCDPM
38603	DGDHDARIH	38650	DRNQWLPRH	38697	ERDAGDTIN	38744	FKPDPCFDD
38604	DGETDELRM	38651	DRPRPYCRN	38698	ERHGEEIRD	38745	FKPTNIEKG
38605	DGFLDGEWW	38652	DRQYPPPNH	38699	ERHYDIMLY	38746	FKWWDTMNQ
38606	DGFMDGEWC	38653	DRSPWCPQF	38700	ERKFYENRM	38747	FKYTDHGIQ
38607	DGKTSREVC	38654	DRSQRCNRG	38701	ERQKDRLPD	38748	FLNDWSNWF
38608	DGPKDRIFN	38655	DRTGRCMPP	38702	ESRYCMWGM	38749	FNPFDCDIG
38609	DGPVPWMMC	38656	DRYLGWQFC	38703	ETRELITMM	38750	FNTFDCEIG
38610	DGSDGMSTN	38657	DVDCCCVYN	38704	EVDEDPAFW	38751	FRCNERRWN
38611	DGVFPRCGN	38658	DWDCKTSRN	38705	EVDFCSAIM	38752	FRCYPNQRD
38612	DHCKWWLDV	38659	DWKPADNKC	38706	EVEHDVAVW	38753	FRDRDGNRM
38613	DHDCKCEEA	38660	DWNSDWPEY	38707	EVTMGCMIN	38754	FRGGEMAKN
38614	DHDGDVPVF	38661	DWQRHPPKC	38708	FDCDDVIER	38755	FRNENEFEP
38615	DHHCDRPQG	38662	DWRPPCFRF	38709	FDDCDAGRG	38756	FRPKCGPIH
38616	DHQCEFPMV	38663	DWSKEIMGM	38710	FDDDDKAKC	38757	FRPKECALQ
38617	DHTCNWPPN	38664	DWTCNGHRH	38711	FDDYDLTVE	38758	FRQCGEEYH
38618	DICFRVSMM	38665	DYDWSDMKC	38712	FDETDFEAM	38759	FRQLDRERQ
38619	DIEDTQMSW	38666	DYENDEDGV	38713	FDFPDCHGW	38760	FRQWQTMET
38620	DIEQWIAAM	38667	DYKCCRPEY	38714	FDHDDEHKC	38761	FRSEEKAEN
38621	DKDFWSAPN	38668	EDDFKPARD	38715	FDHHEKWNN	38762	FRSRNPPKH
38622	DKHLEWPNH	38669	EDEGDLWFY	38716	FDHRPNCHD	38763	FRSYGIRLI
38623	DKKNECNWH	38670	EDEPDGKMG	38717	FDNPDQRNA	38764	FRTWPDDET
38624	DKTCDCEMT	38671	EDEQSGERE	38718	FDNRDSDKG	38765	FWEHADEEG
38625	DLSDQYPEL	38672	EDGHKVGFM	38719	FDNRRHWWE	38766	FYDGPREMW
38626	DLWCMDAEW	38673	EDHDIPRRN	38720	FDPFQFCVG	38767	FYVGPAMPM
38627	DLWTHWNPC	38674	EDNRWIPPL	38721	FDPHHAKQQ	38768	GADWDEPKE
38628	DMDEGSSSN	38675	EDPGDPDMM	38722	FDPIYEMVE	38769	GDCIDGKPH
38629	DMDKDGPRE	38676	EDPHEEFDR	38723	FDSFDEHHN	38770	GDDHDIEFM
38630	DMWDWTYRM	38677	EDQCCFELP	38724	FDWISKCSW	38771	GDDHKSHRK
38631	DNDTWQPRI	38678	EDTQPLFRH	38725	FDWQWEDKC	38772	GDDLDCWKN
38632	DNEWDCAGN	38679	EEAHRGERC	38726	FDYKWMNAY	38773	GDEDHGSKM
38633	DNHREAWQD	38680	EEDHEQMCN	38727	FEDWGWPKG	38774	GDELHFLRM
38634	DNKRRVPEC	38681	EFDHWTMRM	38728	FEHRDNWKM	38775	GDERPGRRN
38635	DNQCEDPKP	38682	EFEHELDIW	38729	FERWDRKEH	38776	GDLCRYWRD
38636	DPEMEYMGM	38683	EFPVPLDFW	38730	FERYDKPYM	38777	GDPDDDEIW
38637	DPKPRHDND	38684	EGAEGWTQD	38731	FGEMDGGCY	38778	GEAIDNEKW
38638	DQDICGWVW	38685	EGDPPMPVF	38732	FGGCDGTPW	38779	GEEEDRMIH
38639	DQDRRCDQP	38686	EHDCQDDVN	38733	FGWLDCCNQ	38780	GESCDEKSD
38640	DQECDDHGI	38687	EHEWDAPSW	38734	FGWPDEQQW	38781	GIPCWQDEW
38641	DQHGWRRPG	38688	EHKYRDMKH	38735	FHETQTPKT	38782	GKPEDGFDD
38642	DQKCERGRI	38689	EHPMELEFC	38736	FKACDSECD	38783	GPMRRRINI
38643	DQNHRVDRH	38690	EIDFGTPET	38737	FKCGNWGII	38784	GRCQPWYFH
38644	DQRHDHRQM	38691	EIDYWWCPH	38738	FKDYRCGWC	38785	GRDCDEAPD
38645	DQSSGWPRD	38692	EMDLCHYSN	38739	FKEGDAKKD	38786	GRDFYWKIE

[2300]

38787	GRECSAQRM	38834	IDDKKWIHA	38881	IRHCDGDCE	38928	LDPCAWVKV
38788	GRHGDDLEV	38835	IDDRIWNCD	38882	IRHEDGCTY	38929	LDPFAQPFS
38789	GRPQCWYKH	38836	IDDWLMDKP	38883	IRPCAPARR	38930	LDPTDPGRW
38790	GRTCDYHDM	38837	IDEWDCIDS	38884	IRPDGEEIQ	38931	LDRLEAPTG
38791	HDATPTPRY	38838	IDFHDCFTK	38885	IRPNSSLEL	38932	LDRNRVHRQ
38792	HDCCEWWRA	38839	IDFNDWIPW	38886	IRPTGTTPN	38933	LDSCDPRLK
38793	HDCWAEWIA	38840	IDGDDEHRA	38887	IRSCPMKCP	38934	LDTCEMLKQ
38794	HDDGDGFQF	38841	IDHKEAFFE	38888	ITCDWYYQM	38935	LDTGEMLKQ
38795	HDDIEAIRQ	38842	IDIHEREHF	38889	IVSPADIVP	38936	LDTKRWPKY
38796	HDEGPEFKD	38843	IDKWEWNRD	38890	IWERDPDGE	38937	LDWKDKERG
38797	HDELWVDRM	38844	IDLQDCNPW	38891	IYCMPYIRA	38938	LDYRDIARD
38798	HDESGEWLE	38845	IDMTDMVLW	38892	KDCNHRAEN	38939	LEAGEKWNF
38799	HDHKWKPPE	38846	IDPEGESIS	38893	KDDCAAWNC	38940	LEAHEGEKF
38800	HDKWENPGQ	38847	IDPFDWEGM	38894	KDEDDIWID	38941	LEDLRLPRM
38801	HDLQPGMNN	38848	IDPFWTFKF	38895	KDHGHQLPP	38942	LENKWCYKM
38802	HDPCTQIKP	38849	IDPPEQDIM	38896	KDHLGGKRN	38943	LENRPELND
38803	HDPWNVLDW	38850	IDPYDCDPT	38897	KDHMGCPPH	38944	LESCSPIVS
38804	HDYEPKCMG	38851	IECNDDCPE	38898	KDHWHPPKY	38945	LEWIHDWYH
38805	HEFKDDMPH	38852	IEKFDDPPA	38899	KDKMKWLKE	38946	LEWYECVVT
38806	HENKEEQKD	38853	IEPWEHMSN	38900	KDMCYLSKK	38947	LGDCPGEPH
38807	HEQWDNIPC	38854	IEQKGWRRD	38901	KDNFWMYKT	38948	LGLCDREMQ
38808	HFECRTMPN	38855	IESDPCVAG	38902	KDNGNKMDE	38949	LGSFDDSIS
38809	HFEFWNGGM	38856	IEWDPWIHL	38903	KDRTQIWKQ	38950	LHDWWMREM
38810	HGDHRGNRE	38857	IEWFRWMMW	38904	KDSPWFERT	38951	LHQRDRIEF
38811	HGSMDSVYW	38858	IFDPRVKAW	38905	KDSQRHDKE	38952	LKCGSLIDV
38812	HHQHNEIEL	38859	IGENMWIWH	38906	KDVFWENKF	38953	LKLADWWTT
38813	HKECPRQRH	38860	IHPMRIGPW	38907	KDWSELYEY	38954	LKLRDEVNE
38814	HKGFEPERQ	38861	IHTVPDWVM	38908	KDYIPQLRG	38955	LKMERCMIA
38815	HKPCHGTEF	38862	IIVEVEICH	38909	KDYNDQEEN	38956	LKPFEDLLN
38816	HMEDWMLEM	38863	IKDGPEIDP	38910	KEKIQDYPE	38957	LKQWDPWLC
38817	HNDTWQPRL	38864	IKFWDGLRD	38911	KEWFDWCIT	38958	LLDCRGPEV
38818	HNEGERWPN	38865	IKPCDEPVD	38912	KPLKCWEPQ	38959	LNDHEMPGM
38819	HQCMWDWLT	38866	IKPIGFWVN	38913	KRDYWCWPM	38960	LNRCDAPKL
38820	HRCCETTIN	38867	IKVKPDNIM	38914	LCSFDDSII	38961	LNWPDYHCM
38821	HRCPIDSRI	38868	IKWPDEHIW	38915	LCSFGDSIS	38962	LQCREGLPM
38822	HRECGVWDF	38869	INPYGTPLN	38916	LDCMRWGFN	38963	LQDEPEPPH
38823	HRSDDDQRF	38870	IPWCDYFFA	38917	LDCPPAPKM	38964	LQDFRFCKQ
38824	HRSDPPQKM	38871	IQDCNSIRY	38918	LDCRPEHPD	38965	LQDPACDPS
38825	HTCRLDVNM	38872	IQERMSIED	38919	LDDEESLVF	38966	LQHRDSVES
38826	HVEEWAEKM	38873	IQLGEFWKE	38920	LDDWDFKPW	38967	LRCKWWRTV
38827	HWSIDAPRK	38874	IQWMDMDPW	38921	LDECEIECN	38968	LRDFWECRL
38828	IDAKPYGRE	38875	IRDFGERRI	38922	LDEFKMEPQ	38969	LRDWDGNRM
38829	IDAWEEHYE	38876	IRDIDYLPW	38923	LDGFDTNLE	38970	LREVDYPSM
38830	IDCCSFSPT	38877	IRDMSCLPS	38924	LDGTDAVRL	38971	LREWKWCPD
38831	IDCQRWCQQ	38878	IRGDEEDRN	38925	LDHKWPDYI	38972	LRPCEGGQD
38832	IDCTPWIQT	38879	IRGDEMLED	38926	LDIRWYEHV	38973	LRPDIAWEC
38833	IDDDPEPTR	38880	IRGREDMLQ	38927	LDLWRNWKI	38974	LRPGGGLKK

[2301]

38975	LRTIWSPRM	39022	NDPGNRPRE	39069	PDNEDARPK	39116	PGAIDWQIM
38976	LRTNDCNFH	39023	NDPRRPDSN	39070	PDPADKCTY	39117	PGCCDRLVE
38977	LRVMDPLEG	39024	NDRCEKPDI	39071	PDPEGEDEP	39118	PGDFGTAED
38978	LRYNDLKIH	39025	NDRPEADEH	39072	PDPGIRWPM	39119	PGGMCWWER
38979	LTGCSDEVQ	39026	NDSMGEPKS	39073	PDPHEDGNP	39120	PGGNDCLLF
38980	LWADGWELK	39027	NDSYGQPKV	39074	PDPLGWLAH	39121	PGPKPWKEI
38981	LYQCTDPEK	39028	NDTGPHGRH	39075	PDPPKKYED	39122	PGPNDVNHC
38982	LYQPGEPKS	39029	NDTNRALKM	39076	PDPPKQYED	39123	PGSHGPPVC
38983	MDACEGQNN	39030	NDYIDEFPY	39077	PDPQQMSRR	39124	PGTMDPPVN
38984	MDARDECYQ	39031	NDYVEEHPM	39078	PDPRYWIMN	39125	PGYNRLEPS
38985	MDCRREHRH	39032	NECCDRHPM	39079	PDPSDKGTY	39126	PHCEEGPED
38986	MDDFEAIRW	39033	NEDHENPSN	39080	PDPSPAKGV	39127	PHDFWVIED
38987	MDDSEVLKY	39034	NEDQEEDPC	39081	PDPSPYCKS	39128	PHDNDAYFW
38988	MDEEFWPFN	39035	NEYRPPPQV	39082	PDPWPWMDD	39129	PHDRRKNDN
38989	MDERDAHEE	39036	NGYFDAMRH	39083	PDQDDMMHG	39130	PHEIEYYMN
38990	MDERWWIKI	39037	NHCDENWYW	39084	PDQFENSFK	39131	PHEWDMCAE
38991	MDFEEMWVW	39038	NHCKDWEEH	39085	PDQFIRWPM	39132	PHNNHWIEP
38992	MDHKDGYRQ	39039	NKPIWYVEM	39086	PDQHDPSNL	39133	PHNQWMPKS
38993	MDHLHKPKT	39040	NPENHKCRF	39087	PDQSHIDLM	39134	PHRHGYRKD
38994	MDNHKKPRA	39041	NQEKTEPKC	39088	PDRCDLTDF	39135	PHRQDRMIW
38995	MDPDPKCGM	39042	NQEMEYHPS	39089	PDRCNYVRH	39136	PHTCCGPEE
38996	MDPSHGKIN	39043	NQTCDQEEM	39090	PDSGEQPVA	39137	PHTDRLDIN
38997	MEGFDGWRM	39044	NRDWFEWKM	39091	PDSPGVCRC	39138	PHWFNWPMC
38998	MEMYRWCRH	39045	NRDWRFFRI	39092	PDTMWREHG	39139	PHYNTVPEN
38999	MFDCRFPRY	39046	NREQDELKW	39093	PDTYHYCEW	39140	PKCSNWLNM
39000	MFDFWNLTM	39047	NREWPDYRD	39094	PDVNQGPWL	39141	PKCTDWNGV
39001	MGECDWPMH	39048	NRGVDRLKW	39095	PDVSERMMD	39142	PKDNPIEPK
39002	MGMEDDWKH	39049	NRPEGRGEC	39096	PECGDRGDR	39143	PKDTKGPET
39003	MHPWDALDM	39050	NRVLEIERK	39097	PEDCPDGKA	39144	PKEQPCAEY
39004	MHYFRYCKN	39051	NRYFDAMRH	39098	PEDEWGFEY	39145	PKEWDKPPH
39005	MKDFWTEKW	39052	NTDFECVGN	39099	PEEDEHDVC	39146	PKHEDCCPE
39006	MMNHHDTRP	39053	NVCEQDIYE	39100	PEEDSPWIM	39147	PKKYDRDVM
39007	MNPMAYTRN	39054	PDALDPCDP	39101	PEFGWWIAM	39148	PKPWETSRC
39008	MQPFQMHRG	39055	PDAYDDRNH	39102	PEHHFRWPH	39149	PKQGPAIFN
39009	MRCDGWIQF	39056	PDCEPGNVE	39103	PEHKDHFET	39150	PKSCGDGND
39010	MRCDLKKKE	39057	PDDCPWPNN	39104	PEKLDLINY	39151	PKTCNCKVT
39011	MRDWEASKM	39058	PDDFDKFMW	39105	PEKYDIWPV	39152	PKYHESAKS
39012	MRQFGCPTE	39059	PDDFEGGPL	39106	PEMEDASLG	39153	PLQDGCEMH
39013	NDCMHNDRW	39060	PDDHESPRD	39107	PENDHGMEA	39154	PMHYWKSEM
39014	NDDEPGWKQ	39061	PDDHETPTN	39108	PESFPAERC	39155	PMPLCCWSL
39015	NDDEPPSRL	39062	PDDIHRFPK	39109	PESNDKRES	39156	PNALDMCVE
39016	NDDFDCQEA	39063	PDEQSLPFD	39110	PETGKAPHI	39157	PNEPWCPVV
39017	NDDLDCCVK	39064	PDEVPRCRA	39111	PETIRGGKC	39158	PNESNEIEH
39018	NDDRCFIRA	39065	PDKLDRFVG	39112	PEVNEGIHL	39159	PNNAWIDKQ
39019	NDEIDAQYH	39066	PDLFEKCRI	39113	PEYPKKGKC	39160	PNNEENDAM
39020	NDEIECIGH	39067	PDLRDHWQM	39114	PEYYDDDHS	39161	PNNNDQDKT
39021	NDMFPSFDC	39068	PDNAQGPQC	39115	PFPRDWHDA	39162	PNPHDRPRW

[2302]

39163	PNQLDKPAM	39210	QDKMPCWQF	39257	RPWEWWIRK	39304	VDEFNVWKH
39164	PNQRNEIEH	39211	QDMGEKLTH	39258	RRPEWLKNY	39305	VDFFDWLDH
39165	PNWGIEPGF	39212	QDMHDKEPA	39259	SDCHENDPH	39306	VDHITKWKF
39166	PQCKPLVPM	39213	QDPCSWLKN	39260	SDDWRDFDN	39307	VDPCDDGNA
39167	PQGCDKYLW	39214	QDPHHNHRH	39261	SDENDVDDC	39308	VDQHERARN
39168	PQHYDKWEI	39215	QDPQWWMRC	39262	SDENEMEPH	39309	VDQRDMATD
39169	PQLPELEKW	39216	QDRYGVPEN	39263	SDEWPYYRQ	39310	VDWFARWDS
39170	PQNGHICKG	39217	QDTDEGDQC	39264	SDNREWMRY	39311	VEFLDEIRH
39171	PQNYDQDKV	39218	QENTPGYEG	39265	SDPCKVLRH	39312	VEKMREWPF
39172	PQPNRTSPH	39219	QEPTEITPM	39266	SDSADDPKF	39313	VEPYEGVRM
39173	PQSEPAPYD	39220	QKPSDEKRC	39267	SDSKGRERV	39314	VLDCDNPVD
39174	PQSFWGSEM	39221	QKQCDRTAW	39268	SDWDPSWDT	39315	VNPMAYTRN
39175	PRCDPCMGV	39222	QMFMPYCPF	39269	SDYSELPQH	39316	VRDCRWCIV
39176	PREGDWEGM	39223	QQPWCWWEK	39270	SEHEPDITN	39317	VREFEEDKN
39177	PREMWECQC	39224	QRCDDWIHS	39271	SEKFGIPRD	39318	VRPYADSRN
39178	PREQRPSKS	39225	QRCKPRCPM	39272	SEPPDWNTK	39319	VRRDDPPRP
39179	PRGLPTPMN	39226	QRDPEDFET	39273	SGPRTWIKK	39320	VRSCGGHDI
39180	PRNCPEDWS	39227	QRMHDCCDL	39274	SHDIPVPIM	39321	VRSWDTDFE
39181	PRNYEAAEC	39228	QVECSWDKY	39275	SHHNPGPRT	39322	VYWREEIRI
39182	PRRMPLVPH	39229	RDAKDGGRP	39276	SHYPNRPKG	39323	WDCEDGWQQ
39183	PRSDDDQRF	39230	RDAYWAEDN	39277	SKDWPPFKA	39324	WDDHERGQV
39184	PRSGDDQRF	39231	RDCFWCIGE	39278	SKQYDPPKP	39325	WDDRPALKG
39185	PRTDGCPWA	39232	RDCHKGIKN	39279	SMPLWQDSN	39326	WDDRQCEQH
39186	PRYQDRCMW	39233	RDDCWPFNP	39280	SMPMWWPKH	39327	WDEKPAGQV
39187	PTPDGSDKN	39234	RDDLRLWPF	39281	SNERACEEV	39328	WDFCKCADI
39188	PTVPQCPRH	39235	RDDLWLWPF	39282	SQPTGWEEN	39329	WDFNDGITP
39189	PVPHSCERR	39236	RDHNEKINY	39283	SRDGDEPNH	39330	WDNFNPPKQ
39190	PWFHKRYKN	39237	RDHNEQINC	39284	SREFWWDGT	39331	WDNWDDREP
39191	PWNISCPKE	39238	RDIKPWFEM	39285	SREWDCGSH	39332	WDPGEGNNF
39192	PWPKWCWPA	39239	RDLEKGDEG	39286	SRPPDWPRT	39333	WDPLEHRLH
39193	PWQFGHAKQ	39240	RDPAWIWDY	39287	TDDCYGPKG	39334	WDPSKVNIM
39194	PWTYRGEDF	39241	RDTMFEWKF	39288	TDDHERGDC	39335	WDQCGWHRQ
39195	PYDQRTDDA	39242	REEFPLCNQ	39289	TDDNGEIRP	39336	WDRASWLEQ
39196	PYGPMCTMG	39243	REEKDWFKH	39290	TDDWYGPKC	39337	WDRCDCEIP
39197	PYKMNHPRE	39244	RELDCMWEN	39291	TDEIPEMRY	39338	WDRCPMPRN
39198	PYNCRRCPD	39245	REPHDAKNN	39292	TDFSDRLKW	39339	WDSCEGREP
39199	PYNPNVDPH	39246	REPYWCLGY	39293	TDHCNWSRT	39340	WDSHDFEIH
39200	QDAMDPWRI	39247	REQLWVHPG	39294	TDHKDGFRI	39341	WDSKRLNPP
39201	QDCVERIKS	39248	REYPKAHRE	39295	TDHNGEIRT	39342	WDWIEVYRC
39202	QDEIDPSEH	39249	RHDTWHDEG	39296	TDNRWMDKW	39343	WDYRNLNRQ
39203	QDEIRMPFN	39250	RKDYWSDDG	39297	TDPPDGFDM	39344	WECQSRDED
39204	QDEMDGYDQ	39251	RLEWWIHPM	39298	TDSCEDCKI	39345	WEDEPPWKI
39205	QDFLPVLKW	39252	RNDFWPAGN	39299	TEFECFWNT	39346	WEEQRIPSE
39206	QDHCEPFEP	39253	RNQDWGDRQ	39300	TESLEKGPM	39347	WEEWDMMRC
39207	QDHKDECNE	39254	RPEMRCFRN	39301	TREGPQLKC	39348	WELPCEWTY
39208	QDHLADWRN	39255	RPKCNLIEM	39302	TRLKELLIW	39349	WEPHPYQTN
39209	QDKHEPGKW	39256	RPNCYKAEH	39303	TWDRECEYD	39350	WFDSRKEEW

[2303]

39351	WFECWTYPM	39373	WKTTEAPRD	39395	WRDQGFSTN	39417	YDNKPLPKH
39352	WGCKENWNM	39374	WLDDRLEFF	39396	WREMPMSEL	39418	YDSCDCSRF
39353	WGDCDPRVG	39375	WNDYDEGEC	39397	WRHWDGLNL	39419	YDTRWQPRC
39354	WGDCDYCEK	39376	WNPADGKQN	39398	WRNQGDLVP	39420	YDWKAADGW
39355	WGHKPCKGY	39377	WNPWDQLEF	39399	WRNRPCLNE	39421	YDYRSYPVA
39356	WGPCAFGQE	39378	WPCPWFDHM	39400	WRQCEFGVE	39422	YEAKETERP
39357	WGPKPCRGY	39379	WPPEPWCCC	39401	WRSDGPIER	39423	YEVDEEIKI
39358	WGSCECHKC	39380	WPPWEWYED	39402	WRSEEVEHC	39424	YEVKEEPVS
39359	WGTKPCKGY	39381	WPSCGCTVE	39403	WRSRDHDHP	39425	YGCFDRCMR
39360	WHANEVLKK	39382	WPVDSWIEN	39404	WRTCDMPYF	39426	YGDFEWCKR
39361	WHDTNLEPH	39383	WQANHWDRT	39405	WRYAEWWNQ	39427	YGPEDLHKW
39362	WHSWWTNIM	39384	WQEDQVGRQ	39406	WRYTERIVH	39428	YHDCRLHRP
39363	WKACPLRET	39385	WQPWRHERH	39407	WWDCEPDTT	39429	YKDARWPLW
39364	WKAYPWQPT	39386	WQQFWWNEC	39408	WWDTDGNFG	39430	YKECHWHKE
39365	WKCAEFRTC	39387	WQYYPCIDK	39409	WYDNAGNKN	39431	YKTVDVGPH
39366	WKCPDKLDN	39388	WRAPGPLEN	39410	WYPDEADHA	39432	YPNKDGERN
39367	WKDAEKKDD	39389	WRCATIPIA	39411	WYQDWHPRM	39433	YPWFDGWQW
39368	WKDHTVPDH	39390	WRCCQFWII	39412	YDCCCGWKD	39434	YQYFELPKW
39369	WKDNESSEF	39391	WRCDDRDNC	39413	YDCIPRCKP	39435	YRHEEYWHN
39370	WKDRDQCLH	39392	WRCPRLIDW	39414	YDDAEMRGH	39436	YRHEPEDEG
39371	WKLVDLGIG	39393	WRDHPEDIQ	39415	YDEMYYCVG	39437	YRPGPTIRT
39372	WKSCGMDSL	39394	WRDMDFYIN	39416	YDHRCYVRG		

[2304] 实施例43

[2305] 如通过机器学习鉴定的具有甲状腺组织嗜性的AAV5变体

[2306] 此实施例描述了如通过机器学习方法鉴定的展现甲状腺组织嗜性的工程化AAV5变体的优选特性。进行如实施例19和图19至图22中描述的相同生物信息学预处理、机器学习方法和分析以确定展现甲状腺组织嗜性的工程化AAV5变体的优选特性。

[2307] 下文描述了AAV5 VP衣壳多肽的581至589区域中各位置处的有利生物物理特性和有利氨基酸残基，它们与较高的甲状腺组织嗜性概率相关。表70中描述的任何以下机器学习衍生的位置氨基酸偏好可以单独存在或彼此组合存在或与实施例15，表29中观察和描述的任何氨基酸偏好组合存在。下文表70中列出了从ML模型中提取的排名靠前的重要甲状腺嗜性特征之间共有的位置特征的概述。

表 70
机器学习衍生的甲状腺组织嗜性规则

Xaa1 处高可突变性 Xaa1 选自 N
Xaa2 处低表面可接近性 Xaa2 选自 F、G 或 M
Xaa3 处高溶解性 Xaa3 选自 F
Xaa3 处低可突变性 Xaa3 选自 Y、F、L 或 C
Xaa3 处中等摩尔质量 Xaa3 选自 D、E、R、K、V、P、M、I、L、N、Q、T 或 C
Xaa3 处低表面可接近性 Xaa3 选自 V、I、L 或 C
Xaa4 处高 goldman engelman steitz Xaa4 选自 L 或 V
Xaa4 处低表面可接近性 Xaa4 选自 V、M、A、G、F、I 或 L
Xaa4 处低摩尔质量 Xaa4 选自 D、A、G、I、L 或 N
Xaa5 处高溶解性 Xaa5 选自 C、L、F、M、V 或 Y 或者 Xaa5 处低溶解性 Xaa5 选自 D
Xaa5 处低平均柔性 Xaa5 选自 F、M 或 W
Xaa6 处低平均柔性 Xaa6 选自 F、M 或 W

[2308]

[2309]

表 70 机器学习衍生的甲状腺组织嗜性规则
Xaa7 处高可突变性 Xaa7 选自 N
Xaa7 处低体积 Xaa7 选自 P、N 或 T
Xaa8 处低平均柔性 Xaa8 选自 F、M 或 W
Xaa8 处低表面可接近性 Xaa8 选自 M、G 或 F
Xaa9 处低可突变性 Xaa9 选自 R、K、P、H 或 C
Xaa9 处低亲疏水性 Xaa9 选自 R

[2310] 下文表71提供了表现出甲状腺组织嗜性并且符合表70中提供的一种或多种规则的变体AAV VP1衣壳多肽的581至589区域的序列。因此,本公开提供了由具有SEQ ID NO:2的工程化AAV5 VP1衣壳多肽组成的rAAV,其中Xaa1至Xaa9区域是SEQ ID NO:41438-SEQ ID NO:42437中的任一者,如表71中所公开。本文还涵盖由工程化AAV5 VP2衣壳多肽和工程化AAV5 VP3衣壳多肽组成的rAAV,所述多肽在对应于AAV5 VP1 581至589区域中的氨基酸的AAV5VP2(氨基酸残基445至453)和AAV5 VP3(氨基酸残基389-397)中的区域处具有下文表中公开的序列。

[2311] 表71

[2312] 驱动甲状腺组织嗜性的AAV5 VP1衣壳多肽中581至589区域的序列

[2313]

SEQ ID NO	581-589 序列
41438	AACAVWKGE
41439	AANWFTHDK
41440	ACWLWFYFG
41441	ADSADRVYN
41442	AFHHMMNTP
41443	AFKKPSLFR
41444	AFYNYEIMA
41445	AHAIMMRDF
41446	AHHALSFWR
41447	AHHEVSFWR
41448	AHNIMMRDF
41449	AHTIMMRDV
41450	AHTIMMRHF
41451	AHTIMMRYF
41452	AIHEPAENR
41453	AIHEPAGDR
41454	AIHKPAEDR
41455	AIHVPAEDR
41456	AKIPIDHFA
41457	AKSSIEFSA
41458	AKSSIEHSA
41459	AKSSIEYSG
41460	AMHEPAEDR
41461	AMMKMMKAN
41462	AMMLMMKAN
41463	AMMQMMKTN
41464	AMMQMMRAN
41465	ANERHWTSF
41466	ANVRHWTYF
41467	ARAMYVESF
41468	ARSSIEYSA
41469	ARTMYGESF
41470	ASDTMNCGS
41471	ASYHWTHAS
41472	ASYHWTHVG
41473	ATERHWTYF
41474	ATRDWMSCM
41475	AVDEVGSSN
41476	AVHEPAEDR
41477	AVILSSMYS
41478	AVTVDGNNR
41479	AWMDCPIMC
41480	AYDEKKNQM
41481	AYDKKKNQM
41482	AYDQRKNQM
41483	AYLINMLWM
41484	AYNQKKNQM
41485	CFENLQIMA
41486	CGGQFWQGI
41487	CKDEFIVGN
41488	CKIPIDHFA
41489	CKQKFNGQN
41490	CKQQFIGQN
41491	CKQQFTGQN
41492	CKQRFNGQN
41493	CMESYDHGS
41494	CMFGMVNYP
41495	CMFGTVNYT
41496	CMKSFDHGS
41497	CMQKDGNDR
41498	CMRGDKRKR
41499	CMRGDQRNR
41500	CMRGDTRNR
41501	CMRKDGNDR
41502	CMRMDCNDR
41503	CMRMDGNNR
41504	CMRMVGNDR
41505	CMRSDKRNR
41506	CMRVDGNDR
41507	CMVDANHVW
41508	CNFNNNDHV
41509	CTLFTNFQY
41510	CTLISNFQY
41511	CVGQFWEGI
41512	CVGQFWKGI
41513	CVGQFWRGI
41514	CVVQFWQGI
41515	DANDMSSDR
41516	DARFGINFG
41517	DARFGLNFA
41518	DARFYIAEG
41519	DARFYISEG
41520	DCVIDWCQM
41521	DDSARYWQV
41522	DDSSWDNLF
41523	DELKMGHNS
41524	DFEYTCNQH
41525	DFEYTCNQS
41526	DFQKPSLFR
41527	DGAHFRSCF
41528	DGIIDWCQM
41529	DGSQAGPDK
41530	DGVIDWCKM
41531	DGVIDWRQM
41532	DGVVDWCQM
41533	DHIACKKLC
41534	DHTACKKVC
41535	DHTACKRLC
41536	DHTGCKKLC
41537	DIFFKCEWV
41538	DIFFKCVWA
41539	DIHHLCNIF
41540	DIHHWCNIC
41541	DIHPWCNIF
41542	DINAKKNIC
41543	DINGKKNIR
41544	DINGKKNIW
41545	DINGKKNVC
41546	DINGRKNIC
41547	DIQHWCNIF
41548	DIYHFNHKS
41549	DIYHWCNIF
41550	DKEDDHRRA
41551	DKIAIDHFA
41552	DKIPIAHFA
41553	DKIPIDHFT
41554	DKIPMDHFA
41555	DKIPVDHFA
41556	DKISIDHFA
41557	DKLPIDHFA
41558	DKMLMVFKR
41559	DKMPIDHFA
41560	DKVPIDHFA
41561	DLHHWCNIF
41562	DLNGKKNIC
41563	DQLKMGHNI
41564	DRCFRDDFE
41565	DRCFRDDMV
41566	DRWTMQHWE
41567	DSGKANTQS
41568	DSKADHFVN
41569	DSKADHGVN
41570	DSKADNCVN
41571	DSKADPCVN
41572	DSKADRCVN
41573	DSKADYCVN
41574	DSMADDPHL
41575	DSMGDDQHL
41576	DSRFNESLG
41577	DTVPSHLGC
41578	DVLKMGHNI
41579	DVVIDWCQM
41580	DWIVDPPKP
41581	DWIVDTAKP
41582	DWIVDTPKA
41583	DWIVDTPKL
41584	DWIVDTPQP
41585	DWMDCSIMC
41586	ECERAKGEW
41587	ECERPKGEW
41588	ECERTKGGW
41589	ECERTKGKW
41590	ECERTKGQW
41591	ECERTRGEW
41592	ECKRTKGEW
41593	EDDGDDNIP
41594	EEIAGHVGA
41595	EFHQDMFEC
41596	EFSGYTNQE
41597	EFYDYEIMA
41598	EFYNYEIMT
41599	EFYNYGIMA
41600	EFYNYQIMA
41601	EHGAVHNAS
41602	EHGAVLNAN
41603	EHHGFNHCN
41604	EHMTCMHCE
41605	EIHQDMFEF
41606	EKAMLLNGT

[2314]

41607	EKAMLMNGA	41654	FPKFCKCIS	41701	GYRINMLWM	41748	IPWEYAMND
41608	EKAMVLNGA	41655	FSDTMNCGS	41702	HAGSGSNVL	41749	IPWPYAMND
41609	EKEDDHRRP	41656	FSQTMIFCA	41703	HAGTGSNVL	41750	ISDCFTCAD
41610	EKFMVMNSC	41657	FTRDWMSCM	41704	HHICFQHCK	41751	ISGMKSAMA
41611	ELMWWQWDR	41658	FVEPCMAIL	41705	HIHHWCNIF	41752	ITETPGNTH
41612	EMSSEMEKL	41659	FVEPCMAIS	41706	HKISLFKVN	41753	ITGAKMYMG
41613	ERFMVMNSF	41660	FVEPCMTIW	41707	HNLMGWYGP	41754	ITGAMMHMG
41614	ERFMVMNSW	41661	FVERCMAIW	41708	HQTLDMERF	41755	ITGATMYMG
41615	ERFMVMSSC	41662	FVQPCMAIW	41709	HQTLDMKKF	41756	ITGIRMYQA
41616	ERGAVHNAN	41663	GAYFFDKYK	41710	HQTLDMKRV	41757	ITSIGMYQA
41617	ERHAFNHCN	41664	GAYFFDKYS	41711	HRCFRDDME	41758	ITSIRMDQA
41618	ERHGFNHCK	41665	GAYFFHKYN	41712	HSAYFRVGA	41759	ITSIRMFQA
41619	ERHGFNHCS	41666	GCQGFFNMK	41713	HTGPGSNVL	41760	ITSIRMYKA
41620	ERHGFNHWN	41667	GCQGFFNMP	41714	HVEYVIWPG	41761	ITSIRMYRA
41621	ERHGINHCN	41668	GDAALYNWF	41715	HVEYVIWPV	41762	ITSIWMYQA
41622	ERHGVNHCN	41669	GDGSWDNLF	41716	HVKYVIWPD	41763	ITSLRMYQA
41623	ERQTLAMDR	41670	GDSSLDNLF	41717	HWISFERHP	41764	ITSNRMYQA
41624	ESMWYYVMA	41671	GDSSWDNVF	41718	HYEICHISK	41765	IVGADRCQA
41625	ESSWPLQWN	41672	GDSVLYNWF	41719	IADIVWQGQ	41766	IVNACMAEK
41626	ESTISPNHK	41673	GDTSWDNLF	41720	IAHPFKCGG	41767	IVQSFWNIP
41627	ESVDHWGDW	41674	GELKMGHNI	41721	ICAIANTCN	41768	IVVAVNQSL
41628	ETAMGTNLY	41675	GFTVDGNNR	41722	IDDDNWEQP	41769	IWDFKCNQC
41629	ETAMLLNGA	41676	GGPKAANDW	41723	IDDGDDNIP	41770	IWEDYLHVG
41630	ETAYRENPC	41677	GIHHWCNIF	41724	IEIALHACW	41771	IWKDYLHMG
41631	ETGPSMFSF	41678	GKMVMVFKR	41725	IFAAPMPGM	41772	IWKDYLHVA
41632	ETGYGENPC	41679	GMRGDKRNR	41726	IFHSNPVVR	41773	IWKDYLYVG
41633	ETGYQENPC	41680	GRFMVMNSC	41727	IFHSTPIVR	41774	IYEERNAQM
41634	ETGYRENSC	41681	GSSWPLLWN	41728	IFVAPMPGR	41775	IYICGPMIN
41635	EVSKFWRHT	41682	GSYHWTHAG	41729	IFVASMPGM	41776	IYIREPMIN
41636	EVTVDGNNR	41683	GTADCRNSS	41730	IFWMKCGWR	41777	IYIRGPMLN
41637	EVVAGHVGA	41684	GTGEWMHMA	41731	IGNRNVNRP	41778	IYIRVPMIN
41638	EYCHWWYKL	41685	GVAKAANDW	41732	IINWLYNEP	41779	IYLRGPMIN
41639	EYERLGDSW	41686	GVNVDGNNR	41733	IISIRMYQA	41780	IYRLREKQE
41640	EYMNCMHCE	41687	GVPKAANDL	41734	IKFWTDNTD	41781	IYTRGPMIN
41641	FAEPCMAIW	41688	GVPRAANDW	41735	IMIEFAYMD	41782	IYVRGPMIN
41642	FCERNMMAQ	41689	GVPVDGNNR	41736	IMLEFAYMD	41783	KAFGWFHFD
41643	FCNRNVNRP	41690	GVSVDGNNR	41737	IMVEFADMD	41784	KAIGWFHFH
41644	FKFWTDNAD	41691	GVTGDGNNR	41738	IMVEFSYMD	41785	KAIGWFHFN
41645	FKFWTDNTE	41692	GVTIDGNNR	41739	IMVEFTYMD	41786	KAIGWFYFD
41646	FKFWTDNTN	41693	GVTLDGNNR	41740	INDCCTMPD	41787	KALGWFHFD
41647	FKSIQCSWG	41694	GVTVDANNR	41741	INDCCTMSD	41788	KAMGWFHFD
41648	FKSVQCSWA	41695	GVTVDGNDR	41742	INVPDILGA	41789	KAVAVNQSL
41649	FMVEFAYMD	41696	GWMDCSIMC	41743	INVPDLLGG	41790	KDAGDDNIP
41650	FNQTMIFGA	41697	GYLINMLLM	41744	INVPDLLGS	41791	KDDADDNIP
41651	FNVPDLLGA	41698	GYLISMLWM	41745	IPKFCKCSS	41792	KDDGDDNIL
41652	FPANYYSNA	41699	GYLIYMLWM	41746	IPKVCKCIS	41793	KDDGDDNIQ
41653	FPGNYHSNA	41700	GYLTNMLWM	41747	IPQFCKCIS	41794	KDDGDNNIP

[2315]

41795	KDDGGDNIP	41842	KVVAVNRSL	41889	LMEFTEDGS	41936	LWCNENAME
41796	KDGGDDNIP	41843	KWEQEPNRN	41890	LMEFTEEGN	41937	LWCNENAMG
41797	KDNGDDNIP	41844	KWRQAEYDE	41891	LMEITEDGN	41938	LWCNENAMN
41798	KDRMLAGQR	41845	KWTRTGSSF	41892	LMEPYGNCD	41939	LWCNENPMD
41799	KDSHVHSFF	41846	KYCHWWYRL	41893	LMRGTVKNG	41940	LWCTENAMD
41800	KDYGDDNIP	41847	KYCLWWYKL	41894	LMRSTVRNG	41941	LWDFKCTQC
41801	KECDPQHWW	41848	KYCNWWYKL	41895	LMSCIKFSC	41942	LWGNENAMD
41802	KEMRKQGMW	41849	KYRHWWYKL	41896	LMSCIPFSC	41943	LYAERNAQM
41803	KEYAPQHWW	41850	KYWHWWYKL	41897	LMSCIQISC	41944	LYEERHAQM
41804	KEYDLQHWW	41851	LAETPGNTH	41898	LMSCLQFSC	41945	LYEERNSQM
41805	KEYDPEHWW	41852	LAIALHACW	41899	LNGMKSAMA	41946	LYEQRNAQM
41806	KEYDPQNWW	41853	LAIPQPVNG	41900	LNVPDLLGA	41947	LYHHDILQT
41807	KEYDPQYWW	41854	LATTMQGFN	41901	LPKFCKCIS	41948	LYKERNAQM
41808	KEYDPRHWW	41855	LATTMQPFN	41902	LPRCFIAAC	41949	LYQERNAQM
41809	KFSGYTNQE	41856	LATTMRAFN	41903	LPRCVIAAY	41950	LYRWIVHNS
41810	KFYNYEIMA	41857	LCAEFPHCR	41904	LPRYVIAAC	41951	LYTEYITKH
41811	KGRWMYHWW	41858	LCAEYPHGR	41905	LSCCANVFC	41952	LYYIAEHST
41812	KHCHWWYKL	41859	LCTEFPHGR	41906	LSCCNNVFC	41953	LYYIPEHST
41813	KIEMLCCPG	41860	LDALPMTQA	41907	LSCCTNGFC	41954	MAGYLEKPF
41814	KIEMLCCQR	41861	LEIAMHACW	41908	LSCCTNIFC	41955	MAVTQSHQF
41815	KIEMLCCQV	41862	LEINWWYPG	41909	LSDFVNDEG	41956	MFVAPMPGM
41816	KIHQDMFEC	41863	LELALHACW	41910	LSDHSQAAV	41957	MGAYIMWDC
41817	KIKMLCCQG	41864	LELNWWYPG	41911	LSDYVNEEG	41958	MHVFSNVFG
41818	KIMWWQWDR	41865	LEVNWWFPG	41912	LSDYVNNEG	41959	MIANYEGEM
41819	KKRRDTHIY	41866	LEVNWWYAG	41913	LSDYVTDEG	41960	MIGNYEGEM
41820	KLEMLCCQG	41867	LEVNWWYPA	41914	LSLSGITQL	41961	MIHLPFGKR
41821	KLYGRRHWH	41868	LEVNWWYPV	41915	LSSCRGMIA	41962	MIHLPFGQR
41822	KMQNENRQM	41869	LEVNWWYQG	41916	LSYHTPVMA	41963	MIHMPFGER
41823	KNDGDDNIP	41870	LFAEFPHGR	41917	LTAREKKCG	41964	MIVIYEGEM
41824	KPEWTMMYD	41871	LFTAFRKFT	41918	LTDYVNDEG	41965	MIVNYEGEV
41825	KPQWAMMYD	41872	LFTAFRQFA	41919	LTEIPGNTH	41966	MIVNYEGKM
41826	KPYEDIARM	41873	LFTAFRQFN	41920	LTEKPGNTH	41967	MIVRCSNTW
41827	KQEMQGVCR	41874	LFWMKCGWR	41921	LTEPPGNTH	41968	MIVSYEGEM
41828	KQYDPQHWW	41875	LGDYVNDEG	41922	LTERPGNTH	41969	MLEQEPNRN
41829	KRETLAMDR	41876	LGERTMDRI	41923	LTETPGNTQ	41970	MLFAENNGL
41830	KRFMVMNSC	41877	LHTIDANRR	41924	LTQEDFTQM	41971	MSSMLEHAW
41831	KRHGFNHCN	41878	LIMRIPPFD	41925	LTQTPGNTH	41972	MSSMLERSW
41832	KSFYMCQCG	41879	LIMRISSFD	41926	LTRADFTQM	41973	MTARRESAY
41833	KSFYMCRCA	41880	LIMRLSPFD	41927	LTREDFAQM	41974	MTARRESSY
41834	KSFYMFRCG	41881	LIVRISPFD	41928	LTREDFNQM	41975	MTARRESTS
41835	KSSMLEHSW	41882	LKNTMWKCA	41929	LTREDFSQM	41976	MTARRGSTY
41836	KSTMCDNDA	41883	LKSCIQFSC	41930	LTREDITQM	41977	MWEKEPNRN
41837	KTEMSDMNW	41884	LKVNWWYPG	41931	LTRQDFTQM	41978	MWEQEANRN
41838	KTEMSDMYW	41885	LLKFIHLNE	41932	LTRVDFTQM	41979	MWEQEPNQN
41839	KTYVWEHES	41886	LLKIYESNQ	41933	LTSIRMYQA	41980	MWFAANNGL
41840	KVLICQNGQ	41887	LMEFTEDCN	41934	LVGADRCQA	41981	MWGWPGDSF
41841	KVVAVDQSL	41888	LMEFTEDGD	41935	LWCNENAMA	41982	NAALYGGWI

[2316]

41983	NACTGTWWR	42030	NGVINDCEF	42077	NSLWCQKCP	42124	NWENSPWCY
41984	NAELFGGWI	42031	NGVNMNQCF	42078	NSMWCEKCP	42125	NWEPCWWAY
41985	NAELSGGWI	42032	NGVSINQCF	42079	NSMWCQECP	42126	NWEPCWWDH
41986	NAELYGGWV	42033	NGWLWWQRY	42080	NSMWCQKCL	42127	NWEPCWWSH
41987	NAEQYGGWI	42034	NGYLVDIPR	42081	NSMWCQKCR	42128	NWEPCWWTH
41988	NAEVYGGWI	42035	NGYLVNIPR	42082	NSMWCQKCT	42129	NWERCWWAH
41989	NAGLYGGWI	42036	NGYLVYIPR	42083	NSMWCQQCP	42130	NWVDSPWCY
41990	NALVLQNVR	42037	NHADYMRDA	42084	NSMWCRKCP	42131	PDRMLAGQR
41991	NANDMGSDR	42038	NHADYMRHA	42085	NSRHNNEDP	42132	PHICFQHCK
41992	NANDMSGDR	42039	NHADYMRKA	42086	NSSCIQECW	42133	PHTIMMRDF
41993	NANDMSSDH	42040	NHADYMRNV	42087	NSSLIQECW	42134	PIAKFANMS
41994	NANDMSSNR	42041	NHTACKKLC	42088	NSSWIEECW	42135	PLDWKLNNR
41995	NANDMSSYR	42042	NHVDANLRA	42089	NSSWIQECL	42136	PLDWPLNNR
41996	NANVLSSDR	42043	NHVNANRRA	42090	NSSWIQQCW	42137	PLYGRRHWH
41997	NANVMSSDR	42044	NHVNDNLRA	42091	NSSWLQECW	42138	PNILMMFEA
41998	NAQHAWENG	42045	NHVNTNLRA	42092	NSVHNNEDP	42139	PNIVMMFEP
41999	NAVLYGGWI	42046	NIMSYSSWN	42093	NTADCRNSS	42140	PNMLMMFEP
42000	NAYENSHYP	42047	NKMVMVFKR	42094	NTDEESGMP	42141	PSPHEWMAF
42001	NAYLVHIPR	42048	NLKWDWQNH	42095	NTKYAERSG	42142	PSRPCTYHT
42002	NAYQCSHYP	42049	NLKWREHHP	42096	NTQFAERSG	42143	PWAQTEKPN
42003	NAYQNSNYP	42050	NLKWRETHP	42097	NTQYADRSG	42144	PWGETEKPN
42004	NAYQTSHYP	42051	NMCEYANVQ	42098	NTQYAERCG	42145	PWIPNQIKR
42005	NCALERMPS	42052	NMCKCANVQ	42099	NTQYAERPG	42146	PWTETEKPN
42006	NCAPARMPS	42053	NMCKFANVQ	42100	NTQYAERSA	42147	QAWWKIFWH
42007	NCAQERMPS	42054	NMCKYANVP	42101	NTQYSERSG	42148	QEMRKKGMW
42008	NCLVNLCTY	42055	NMCKYANVR	42102	NTRFADQTE	42149	QEMRKRGMW
42009	NCPPERMPS	42056	NMCKYPNVQ	42103	NTRFPDQAE	42150	QEMRQQGMW
42010	NCVVHLCTY	42057	NMCKYSNVQ	42104	NTRFPDQTA	42151	QFSGYTNQV
42011	NCVVNLCKY	42058	NMCRYANVQ	42105	NTRFSDQTE	42152	QFYNYEIMA
42012	NCVVNPCTY	42059	NMHWQYQFE	42106	NTSWIQECW	42153	QGQWMYHWW
42013	NEAPMGLKY	42060	NMHWRYKFE	42107	NTYQNSHYP	42154	QGRWMYNWW
42014	NEHPMGLKY	42061	NMIPMDSMP	42108	NVFVAYCKY	42155	QGRWMYYWW
42015	NERPMGLKY	42062	NMLPFNCNF	42109	NVIVAYCKS	42156	QGRWVYHWW
42016	NESPMGLKY	42063	NMPWRYQFE	42110	NVLAFNCNF	42157	QIHQDMFEC
42017	NFWDQVHCV	42064	NMYWRYQFE	42111	NVLPFNCHF	42158	QIMRISPFD
42018	NFWNQVHCG	42065	NNCNMNVMR	42112	NVLPFNGNF	42159	QIYGRRHWH
42019	NFWNQVHFV	42066	NNMHSDWCV	42113	NVLPVNCNF	42160	QKAMLLNGA
42020	NFWSQVHCV	42067	NNMWCQKCP	42114	NVLSHTMME	42161	QLYGRRHLH
42021	NGAPERMPS	42068	NNVNANLRA	42115	NVLVHTMME	42162	QMSSEMEQL
42022	NGCFVYWWL	42069	NPSQMDMDT	42116	NVMGHTMME	42163	QNSNVMGCM
42023	NGFMPAVMA	42070	NQAHCILNC	42117	NVNDMSSDR	42164	QRHGFNHCN
42024	NGRLLWQRY	42071	NQLQMWYRV	42118	NVQCNRANG	42165	QSKWYYVMA
42025	NGRLWWERY	42072	NQTHCIINC	42119	NVQCTRADG	42166	QSRAGYMHY
42026	NGRLWWKRY	42073	NRFASWEGM	42120	NVQCTRSNG	42167	QSVDHWGDL
42027	NGRLWWQRC	42074	NSDEESGME	42121	NVYVAYCKS	42168	QVHEWVFPR
42028	NGRLWWQRF	42075	NSLHNNEDA	42122	NWADSPWCY	42169	QVLICKNGQ
42029	NGVIINQCF	42076	NSLHNNEHP	42123	NWEHCWWAH	42170	QVLICQNAQ

[2317]

42171	QVLICQNGL	42218	SEYAGMRDS	42265	STAACRNSS	42312	TNDIWLQFS
42172	QVLICRNGQ	42219	SFTLLVGDY	42266	STADCRNNS	42313	TNDIWWQFA
42173	QVLVCQNGQ	42220	SGCAMHNMG	42267	STADCRSSS	42314	TNDIWWQFC
42174	QVVICQNGQ	42221	SGCAMLDMG	42268	STADGRNSS	42315	TNDIWWQFF
42175	QVVIDHPKY	42222	SGCAMLHMG	42269	STADYRNSS	42316	TNDIWWQVS
42176	QWIRTGSSF	42223	SGCAMLNFG	42270	STAGCRNSS	42317	TNDIWWQYS
42177	QWTRAGSSF	42224	SGCAMLNMA	42271	STGDCRNSS	42318	TNKMLACSS
42178	QYCHWWYKL	42225	SGCAMLNME	42272	STPDCRNSS	42319	TQDIYEHMQ
42179	RAIGWFHFD	42226	SGCAMLNMV	42273	STSADPNCG	42320	TRFASWEGM
42180	REISVKAVA	42227	SGCAMLNVG	42274	STSADPYCG	42321	TSEFVAISM
42181	REISVQAVA	42228	SGCAMLYMG	42275	STSGDPHCG	42322	TTEMSDMHW
42182	REMRKQGMW	42229	SGCAMPNMG	42276	STTDCRNSS	42323	TTSADPHCG
42183	RFFAAELQL	42230	SGCAMVNMG	42277	STTFASSWE	42324	TTWVYWNEF
42184	RFFAAEPQW	42231	SGCGMLNMG	42278	STWDWMSCM	42325	TVSGIDHPK
42185	RFFAAQLQW	42232	SGCVMLNMG	42279	SVIQNEMRI	42326	TVVINDCEF
42186	RFFADELQW	42233	SGDYYDGMP	42280	SVLGHTMME	42327	TVWVYWNEF
42187	RFFAVELQW	42234	SGFASWEGM	42281	SVLQNEMRI	42328	TYDQKKNQM
42188	RHICFKHCK	42235	SGLGMPNWP	42282	SVNMRFAKH	42329	VCAEFPHGR
42189	RHMCFQHCK	42236	SGWAMLNMG	42283	SVNMRFTEH	42330	VCAIVNTCN
42190	RMATMFTCR	42237	SHMMNHAFW	42284	SVNWFTHDK	42331	VCNRNVNRP
42191	RMHDACFDN	42238	SHMMSNAFW	42285	SVVQNEMPI	42332	VCWLWFYCG
42192	RMHDACYDT	42239	SIKMCQSTV	42286	SVVQTEMRI	42333	VCWLWFYFA
42193	RMHDGCYDN	42240	SIKVCESTV	42287	SWIDMPMMQ	42334	VCWLWFYFV
42194	RMHDTCYDN	42241	SIKVCQSNV	42288	SWTALGSVT	42335	VCWVWFYFG
42195	RMNDACYDN	42242	SIKVCRSTV	42289	SWTALSSVP	42336	VCWWWFYFG
42196	RMRDACYDN	42243	SIPSIGSEE	42290	SWTAVGSVP	42337	VFHSNPIVR
42197	RMYDACYDN	42244	SIVDFESAA	42291	SWTIVEQVH	42338	VFVAPMPGM
42198	RPEYANHMH	42245	SKSSIEYSA	42292	SYLINMLWM	42339	VFWMKCGWK
42199	RQTLDMKRF	42246	SLHSIGSEE	42293	SYMMSHAFW	42340	VFWMQCGWR
42200	RSFYMCRCG	42247	SLPSIGNEE	42294	TACAGTWWR	42341	VGIPERAMP
42201	RTEYADHMH	42248	SLPSLGSEE	42295	TASAIDHPK	42342	VGWLWFYFG
42202	RTKYANHMH	42249	SLPSVGSEE	42296	TASGIDHPQ	42343	VHKWDTLRD
42203	RTYLWEHES	42250	SMCKYANVQ	42297	TASGIDHPT	42344	VHMGEKMWS
42204	RWFAENNGL	42251	SMIPMDSMT	42298	TASGMDHPK	42345	VHMGQKMWN
42205	RWRCRMRGN	42252	SMIPMDSVP	42299	TAVINDCEF	42346	VIARCSNTW
42206	RYFAAELQW	42253	SMSPMDSMP	42300	TAWIYWNEF	42347	VIIRCSNTW
42207	SANDMSSDR	42254	SMVPMDSMP	42301	TDDGDDNIP	42348	VILRCSNTW
42208	SANWCTHDK	42255	SNERHWTYF	42302	TDRMMAGQR	42349	VIVNYEGEM
42209	SANWFAHDK	42256	SNMMSHAFW	42303	TFVAPMPGM	42350	VIVRCANTW
42210	SASWFTHDK	42257	SNWWFMFEA	42304	TGVINNCEF	42351	VIVRCPNTW
42211	SATWFTHDK	42258	SPQHCIGMP	42305	TGVLNDCEF	42352	VIVRCSNTG
42212	SCNIVEQVH	42259	SRFASWDGM	42306	TGVMNDCEF	42353	VIVRGSNTW
42213	SDAVLYNWF	42260	SRFASWEGV	42307	TGVTNDCEF	42354	VKFWTDNTD
42214	SEDIDYVHL	42261	SSAYYDGMP	42308	TGWVYWNEF	42355	VLVRCSNTW
42215	SEYAAMRDP	42262	SSDFYDGMP	42309	THHELSFWR	42356	VMEFTEDGN
42216	SEYAAMRDT	42263	SSDYYAGMP	42310	TIHEPAEDR	42357	VNDCCTMRD
42217	SEYAAMRDY	42264	SSDYYDGMT	42311	TKVNHIMRP	42358	VPDLFNQFC

[2318]

42359	VPHPFKCGG	42379	WDCLLHKWR	42399	WMSKQMFKG	42419	YVSQMWMRC
42360	VSDFKIYMG	42380	WDCLLNRWR	42400	WMSQQMFKG	42420	YVSQVWMQC
42361	VSDTLNCGS	42381	WDCLVNKWR	42401	WNCLLNKWR	42421	YWLLDHANF
42362	VSFLPDEYC	42382	WECLLNKWR	42402	WSSVAHPGN	42422	YYAICHISK
42363	VSFLPNQYC	42383	WGEACKLDT	42403	WVGQFWQGI	42423	YYEFCHISK
42364	VSYHWTHAG	42384	WHCLLNKWR	42404	WWAPPWHAL	42424	YYEICDISK
42365	VTAFKIYMG	42385	WHGMRMIRA	42405	WWTDMSSWK	42425	YYEICHICK
42366	VTAREEKCG	42386	WHGMRMIRT	42406	WWTDTSSWT	42426	YYEICHIST
42367	VTDFEIYMG	42387	WHGMRMLRP	42407	WWTNMSSWT	42427	YYEICHLSK
42368	VTGAMMYMG	42388	WHGMRMVRP	42408	WYGMRMIRP	42428	YYEICNISK
42369	VTIHERRCN	42389	WKVIYHQPG	42409	YGYLVHIPR	42429	YYEICPISK
42370	VTSEWMHMA	42390	WKVIYHQRG	42410	YHEICHISK	42430	YYEICQISK
42371	VTVHEHRCN	42391	WMEHYGNCD	42411	YHTACKKLC	42431	YYEICRISK
42372	VVICKRFVG	42392	WMEPYDNCD	42412	YKFWTDNTD	42432	YYEICYISK
42373	VVTVDGNNR	42393	WMEPYGNCE	42413	YLKDVDFDS	42433	YYELCHISK
42374	VWFAENNGL	42394	WMEPYGNCN	42414	YPQCVQWFW	42434	YYGICHISK
42375	VWMDCSIMC	42395	WMEPYGNYD	42415	YPRCAQWFW	42435	YYKICHISK
42376	VYEERNAQM	42396	WMESYGNCD	42416	YQHWAMVQR	42436	YYQICHISK
42377	VYRLRERQE	42397	WMGPYGNCD	42417	YQHWTMDQR	42437	YYVICHISK
42378	VYRVREKQE	42398	WMKSYDHGS	42418	YVEYVIWPD		

[2319] 实施例44

[2320] 用衣壳包裹治疗性有效载荷的AAV5源性病毒体治疗雷特综合征

[2321] 此实施例描述了用具有本文公开的任何一种工程化CNS嗜性变体AAV5 VP衣壳多肽的变体AAV5源性病毒体治疗雷特综合征，其中变体AAV5病毒体衣壳包裹治疗性有效载荷。将编码AAV Rep、AAV5源性变体Cap和辅助蛋白的多核苷酸序列以及治疗性有效载荷在细胞中转染以产生变体AAV5病毒体，其中编码变体AAV5Cap的多核苷酸序列在从对应于VP1衣壳多肽中的581的位置至对应于589的位置的区域中包含至少一个突变，并且其中编码变体AAV5Cap的多核苷酸序列编码SEQ ID NO:7118-SEQ ID NO:10117中任一者的氨基酸序列。治疗性有效载荷是靶向由与雷特综合征有关的基因编码的mRNA的指导RNA、靶向与雷特综合征有关的基因中的提前终止密码子(PTC)的阻抑tRNA，或转基因。指导RNA所靶向的mRNA是由编码MECP2的基因编码的mRNA。阻抑tRNA靶向MECP2中的PTC。转基因是MECP2。将衣壳包裹有效载荷的变体AAV5病毒体施用于受试者。受试者是人类或非人类动物。施用途径是全身施用途径。全身施用途径是静脉内施用。与衣壳包裹治疗性有效载荷的野生型AAV5病毒体相比，衣壳包裹治疗性有效载荷的变体AAV5病毒体在施用于受试者后表现出增强的CNS组织嗜性。在施用于受试者时，与衣壳包裹治疗性有效载荷的野生型AAV5病毒体相比，施用较低剂量的衣壳包裹治疗性有效载荷的变体AAV5病毒体。在施用于受试者后，雷特综合征的至少一种症状得到缓解或受试者得到治愈。

[2322] 实施例45

[2323] 富集变体的测定

[2324] 此实施例示出了在以下各组织中富集变体的测定:肾上腺、骨髓、CNS组织、结肠、心脏、肺、淋巴结、乳腺、骨骼肌或心肌、坐骨神经、骨骼肌、皮肤、脊髓、脾脏和甲状腺。

[2325] 衣壳变体可以根据观察到的其丰度进行分析。在如实施例19中所述过滤之后，基于动物之间、或组织切片、或CNS区域或测序样本之间的重复观察数量对变体进行评分。这

导致变体频率的分布，并且可以通过统计方法相对于所有其他变体的观察频率分布根据变体的观察数量对变体进行排名。

[2326] 此外，组织内的变体“富集”可以使用以下等式或该等式的变型来计算：

$$\text{富集评分} = \frac{\text{组织1中的变体1 UM1计数/组织1中变体UM1计数的总和}}{\text{所有组织中的变体1 UM1计数/所有组织中变体UM1计数的总和}}$$

[2327]

$$\frac{VARjk \,/\, \sum_{j=1}^{n} VARjk}{\sum_{k=1}^{m} VARjk \,/\, \sum_{k=1}^{m}\sum_{j=1}^{n} VARjk}$$

[2328] 其中VAR是对于所有m个组织中所有n个变体，组织根据k中变体j的UMI计数。

[2329] 如同观察数量，这导致各组织内变体富集频率的分布，并且可以通过统计方法相对于所有其他变体的富集频率分布根据变体的富集对变体进行排名。

[2330] 然后可以将得到的观察和富集相对评分求和，或取平均值，或以其他方式作积分运算，并且可以根据此以积分法运算的观察/富集评分对变体进行排名以鉴定可能感染给定组织的候选物。

[2331] 为了基于以积分法运算的观察和富集评分生成各组织中存在的衣壳变体的列表，优先考虑观察到的变体。根据以积分法运算的评分对变体进行排名，并且包括排名靠前的最多1000个变体。

[2332] 各组织的富集序列描述于下文表中。本文还涵盖由工程化AAV5 VP2衣壳多肽和工程化AAV5 VP3衣壳多肽组成的rAAV，所述多肽在对应于AAV5 VP1 581至589区域中的氨基酸的AAV5 VP2(氨基酸残基445至453)和AAV5 VP3(氨基酸残基389-397)中的区域处具有下文表中公开的序列。

[2333] 表72

[2334] 驱动肾上腺组织嗜性的AAV5 VP1衣壳多肽中581至589区域的序列

[2335]

SEQ ID NO	581-589 序列
3118	KIEYGHWIH
3119	TIDVKRWGQ
3120	CAIPRHVFP
3121	CSYKNQQHQ
3122	AAQKGQHYS
3123	AEMPTVNNI
3124	AERTKELHS
3125	AGWHNTFWG
3126	AHDAYTKNS
3127	ANVHSGQCQ
3128	AQVFSVGWT
3129	ASHRYILAE
3130	ATMTFHFSP
3131	ATVLEASPD
3132	AVHLKAACP
3133	AVMHKNADY
3134	AVMSIPFQS
3135	AVQSCWTGT
3136	CFHRCLFIL
3137	CHLSTCRWP
3138	CIHMLKLAN
3139	CIHQYGRVH
3140	CILRHIKEN
3141	CKMRQRPFR
3142	CLIRNVAGG
3143	CPHKFLMER
3144	CPRCYHRAQ
3145	CSGSKMRMG
3146	CTANRAMML
3147	CTQRRMTFP
3148	CTSYRLSTP
3149	CVMSYINWH
3150	CVNYNLGND
3151	CYKWIDMCN
3152	DHEWYCSLA
3153	DILRYHTQL
3154	DSFQTENSM
3155	DTWIMSQVH
3156	DVKAPHGCV
3157	EDGRADWHV
3158	EHDCYMAVV
3159	ETALLEFRR
3160	EVCQDYQQL
3161	EWKRRLKRR
3162	FCHLIMRYY
3163	FCNYVATDD
3164	FHLRSHWMP
3165	GECRESCVT
3166	GKSDCEHMY
3167	GMHETSVND
3168	GPPLFTMIG
3169	GTTQARIGS
3170	GTVCNNECR
3171	GYIQVHVFN
3172	GYQRAYVPF
3173	HALTWMRAK
3174	HFPYNFWYY
3175	HKIYGSTPY
3176	HSQPDKTGA
3177	HSYNWDEGG
3178	HVMTRNNVY
3179	IIQKTQHNT
3180	IKVQMECHM
3181	ILYNHWVIG
3182	IMEEARTYC
3183	IMSNYCLTA
3184	INKVQTFGL
3185	IRCCTKPLT
3186	KCATYPSHD
3187	KIGMEEVMS
3188	KILIRDCGV
3189	KKKYWKMED
3190	KMLYSSYTQ
3191	KNEWPCIGF
3192	KNKTSEYAC
3193	KTFFHLSFA
3194	KWADQRNSF
3195	LCMLAQIME
3196	LLQKMWAPT
3197	LMNRRADER
3198	LVHCFLKSA
3199	LYCLCKNWC
3200	LYGLKMMYN
3201	MANMVWQNW
3202	MGAMALELA
3203	MSKIKQRLN
3204	MVNRPATCV
3205	MVTPHDHWH
3206	NCDSKVVGH

[2336]

3207	NCLKEIVVD	3248	VSAYRCSFE	3289	CSPSMHPSP	3330	IMLSCRNMK
3208	NHHCCTPSP	3249	VVMHCAKEF	3290	CSPSVHLSP	3331	IMLSYHNMK
3209	NHRWNFGKS	3250	VYHRLCRDT	3291	CSVTRRYWG	3332	KGEAMCQSL
3210	NLVDCECRV	3251	VYHRNATWV	3292	CTMYRNVSW	3333	KGEAMYQSL
3211	NTCEGWSHF	3252	WPEQCNVNL	3293	CVGWQSTAP	3334	KGEAVCQSL
3212	PHMPYSDQQ	3253	YAQSVSESN	3294	CVGWRSAAP	3335	KGESMCQSL
3213	PQYSVCGGP	3254	YCERCQDFH	3295	CVGWRSTAP	3336	KGGAMCQSL
3214	PSARSEQQD	3255	YDEYPRNAT	3296	CVGWWSTAP	3337	KSDPNRWAG
3215	QAGGSGFNE	3256	YDNYWAYNP	3297	CVHWYSKSS	3338	KSFAGTTHM
3216	QISAYGRPL	3257	YGVCLWNNG	3298	CVIPPRGCV	3339	KSVAGTTHM
3217	QQNGSNNFM	3258	YITQQYASH	3299	CVIPPRGWV	3340	KVEAMCQSL
3218	QSTPTLGPK	3259	YNPQQGKSK	3300	CVQWILNSA	3341	LEWYPNHHN
3219	QVGTHNHMI	3260	YPYAILVDI	3301	CVQWIRNSA	3342	LEWYPNHLN
3220	QWVHKQNLD	3261	YWQWNTSGN	3302	CVQWIRNSV	3343	LEWYPNHRN
3221	QYADFKAQQ	3262	YWVPDCYNF	3303	CYPSMHLSP	3344	LEWYPNHSN
3222	QYYMTIHHF	3263	GVHATCMHA	3304	DHDEIFVAR	3345	LEWYPYHRN
3223	RAETTKQLM	3264	ADRQCLQGP	3305	DHDEMFVAR	3346	LICFALNAP
3224	RCWITEVSY	3265	ASDPNRWAG	3306	DHNEIFVAR	3347	MVYCNRKSN
3225	RELSVQQTI	3266	ATGTYNLQK	3307	DIQCANEGY	3348	QFHVPQSDS
3226	RGKGHGTDL	3267	CIHRYSKSS	3308	EGIHCWKEM	3349	QFHVPQSGS
3227	RGLKCYMRN	3268	CIHWYGKSS	3309	EGIHCWKGM	3350	QFRVPQSDS
3228	RVHTIDLTH	3269	CIHWYSKSA	3310	EIDPNRWAG	3351	QNNNLYTQV
3229	RYQPVENIE	3270	CIHWYSKSS	3311	ESAPNRWAG	3352	SAGMLAQMC
3230	SCCWRGSME	3271	CIHWYSKSY	3312	ESDPNRGAG	3353	SAGMRAQMC
3231	SHVCQEYMN	3272	CLVSYSQGQ	3313	ESDPNRLAG	3354	SAGMRTQMC
3232	SLPQTDQHV	3273	CLVSYSRGQ	3314	ESDPNRWAC	3355	SAGMWAQMC
3233	SNAEMSQLF	3274	CMHWYSKSS	3315	ESDPNRWAD	3356	SAGVRAQMC
3234	SNKLMMDVM	3275	CMQTDKLMN	3316	ESDPNRWAG	3357	SAVMRAQMC
3235	SPRCGGPFP	3276	CMQTGKLMN	3317	ESDPNRWSG	3358	SEQYVAYSG
3236	SVCMHIGSQ	3277	CMQTVELMN	3318	FSLFCGQYS	3359	SKQYVAYSE
3237	SVGWSFEPK	3278	CMQTVKLMK	3319	FSLFCRQYS	3360	SKQYVAYSG
3238	TDLEQGMDQ	3279	CMQTVKLMN	3320	GICFALNAP	3361	SKQYVTYSG
3239	TFDLWQWGH	3280	CMQTVKLMS	3321	GIQCANEGY	3362	TASKYIGRT
3240	TTEWKPSIL	3281	CMQTVKLMT	3322	GLTIDHQLV	3363	TATEYIGRT
3241	TYLHQPEDS	3282	CMQTVRLMN	3323	GLTTDHQLV	3364	TATKYIGRT
3242	TYSNDGLFN	3283	CSMARRYWG	3324	GLTTDPQLV	3365	TKPSMEVNC
3243	VAKITMNTW	3284	CSMTRGYWG	3325	GLTTNHQLV	3366	VEANSACQN
3244	VCQVGFLND	3285	CSMTRRYWG	3326	GVQCANEGY	3367	VEANSECQN
3245	VGQPTEKSF	3286	CSMTRRYWS	3327	IICFALNAP	3368	VEANSEFQN
3246	VHCKVWGQC	3287	CSPGMHLSP	3328	IMLSCHNMK	3369	VEANSGCQN
3247	VMEGQRGWT	3288	CSPSMHLSP	3329	IMLSCHNMR	3370	VICFALNAP

[2337]

3371	VICFALNAS	3412	PCQQLMCFM	3453	DARFDQWHH	3494	AAGHHANLG
3372	VICFALNTP	3413	TVNRAMREE	3454	ATGTYNLQE	3495	AAGYARTYD
3373	VSLFCRQYS	3414	EVLKFCKEW	3455	TTTNQFMFQ	3496	AANLNEGAK
3374	WCVHAYWQV	3415	SMSHNGQCL	3456	VCRLDNGPQ	3497	AANPHMHMR
3375	WRVHAYWQV	3416	TNMLAMRGD	3457	EYACMLMTQ	3498	AASHLIVHE
3376	WRVHSYWQV	3417	MLAFDPMGR	3458	YDYWMQAPT	3499	AATITYAPS
3377	YIHWYSKSS	3418	ASFFLNPLE	3459	SHAQPIASH	3500	AATMAPWMA
3378	WGQSYNYWP	3419	QKAIVAQND	3460	DFNGSNWLP	3501	AAVKKLSDD
3379	SIHSIVKGE	3420	ISMRSECED	3461	RVNGSTCRH	3502	AAYLPSLVH
3380	GMLEFMNMT	3421	ISWISHASP	3462	HIHMSDRGA	3503	ACCRMTTQQ
3381	QMYWVQNWE	3422	CVEQYVEQG	3463	NAYYDSHCE	3504	ACDFIGQNE
3382	STDPHTSHH	3423	STTAYHQDY	3464	THVKRFDYE	3505	ACELQEMAM
3383	TLTEVNRLA	3424	GWDMLSVAK	3465	KKHKGWCRA	3506	ACGQAIHHS
3384	SCASCRFPL	3425	MITNNCNIS	3466	MPWLDKPPC	3507	ACIFRHQFY
3385	ETHWDIDSN	3426	TCHPNMSVP	3467	RRTDHGEPI	3508	ACLEVYCQD
3386	EAAHQTNIW	3427	MEHTMDFGV	3468	GNECFHNVE	3509	ACLSSPIFE
3387	MGTEWTQSY	3428	KEFQLQMEP	3469	YPTMTFPEE	3510	ACMTVYRSC
3388	QQSFFFGEQ	3429	TIVVIPIKS	3470	KVNCRMIIM	3511	ACQPSPSYT
3389	CVDTNRMWQ	3430	SSIYKEIQE	3471	WPNIKHQPI	3512	ACTQRQEEM
3390	DVCKATTTE	3431	GTRIYGKVS	3472	QCTAHNTCM	3513	ACYCYFMFI
3391	TTYPHNTHG	3432	EFVLGCQGI	3473	DMTLQSVSS	3514	ACYRQHATK
3392	KGLITENNY	3433	QSGMLTTFA	3474	YFPSIQCYE	3515	ADIWIQFDA
3393	LHKAYVTDS	3434	YVDHWTLGD	3475	RYPCCQSPF	3516	ADRLINSSK
3394	KQATNNNGW	3435	SRYVNNGTV	3476	WKATPGGQC	3517	ADYQTQTYN
3395	CFARTLSYD	3436	AAQYMGIMK	3477	GHLFNNNEW	3518	AECLSTKWV
3396	MPYPLNVKE	3437	AEDTPYTIY	3478	MNSFYRAEW	3519	AEQWKLGSI
3397	EQTHTWLCR	3438	DCVLTMVDK	3479	QWFTKGVGE	3520	AERDCEKGP
3398	ELKYVMMTA	3439	MTIKMEQQY	3480	INYISFVDR	3521	AERMFHQHF
3399	LARTTTNSE	3440	NSQLSMTFG	3481	KMAGIGCWY	3522	AFCSMNATS
3400	TLTTDKTYS	3441	NMPSLRTRI	3482	LEIGGKTRT	3523	AFIEQVMNC
3401	TRSYFDQSC	3442	EIHMFGQSE	3483	NGVGRSGDN	3524	AFMAIMWTL
3402	PYIMEFHPF	3443	MVMMREPSN	3484	AGMYKCQED	3525	AFQQPHKAN
3403	KIRDCDGNV	3444	MNANETHGA	3485	YSDPKNMSI	3526	AGDRYTSED
3404	EIRQFVKGA	3445	WTTSIAVDT	3486	DVKSWAVCD	3527	AGISLWCHN
3405	APFFTFTNH	3446	NDVWQPHPN	3487	TVNWYSGFG	3528	AGLTMQKCE
3406	TQCKQLTVM	3447	GLSMVLHFV	3488	AAAGFSAPS	3529	AGSSCNTFH
3407	CQNTAVRAI	3448	RCADNPYFR	3489	AAARDRRQT	3530	AHAEFAVDP
3408	KAVLFLKDC	3449	AFLHGQMWQ	3490	AACAQVYHS	3531	AHAQYQSVA
3409	DVTQHTSCT	3450	NANYWFYYD	3491	AACGVLWSE	3532	AHMPREWMA
3410	VCQYCELGG	3451	AIYTTASFD	3492	AAEREKAHG	3533	AHTAIGSCP
3411	AVVCKMNYH	3452	EAGCKMWPT	3493	AAGFVTCLD	3534	AHVCLSGQW

[2338]

3535	AHVDYNTCD	3576	AQGTQCANL	3617	ATVMTVHQG	3658	CAYAYANLC
3536	AHVGVELAE	3577	AQKQADINA	3618	ATYNPGVNW	3659	CCEFKMWRG
3537	AHYHVHSHG	3578	AQRMAKIWA	3619	AVCSMHGTI	3660	CCMWKNSWE
3538	AIAHGSDYE	3579	ARELLHWYG	3620	AVDIIAWDT	3661	CCNSRNVFP
3539	AIDHRTMCY	3580	ARMPVSCLP	3621	AVEEMVAKG	3662	CCTKLQVLA
3540	AIEKQPNFE	3581	ARSRTLSDG	3622	AVEKCMTLH	3663	CCTLKRCYD
3541	AIESSMHCC	3582	ARTECHHCG	3623	AVEPHCFGH	3664	CDGQLDFHW
3542	AIGKSNQGL	3583	ARYKLMPHH	3624	AVGCYTMSG	3665	CFFRVRLCI
3543	AIIHQSTAW	3584	ASCRLAQCC	3625	AVKMFIDRN	3666	CFHIRNIDR
3544	AILDQNSLQ	3585	ASEGAGSWE	3626	AVMAQKIAD	3667	CFHPTDQGW
3545	AINTIPMAQ	3586	ASEPITFVS	3627	AVQKEYQYL	3668	CFKYTFMQH
3546	AISQIMSNE	3587	ASFGYMIDD	3628	AVSQFQCSQ	3669	CGAKMSRAA
3547	AITAHLAHD	3588	ASGIIKSNG	3629	AVVEHAVNR	3670	CGTIVHWST
3548	AITQMEWIN	3589	ASGPIGTAA	3630	AVVRSVTRD	3671	CGTTIAYCP
3549	AIVPESAAQ	3590	ASHGDGVGT	3631	AWDSKSNWD	3672	CHFTQHKSD
3550	AKANRCEAE	3591	ASKPVELGH	3632	AWEMYHVID	3673	CHHWFNHAT
3551	AKMPQCAWA	3592	ASLMKPEAV	3633	AWGTMNHWL	3674	CHLTTNVNE
3552	AKMQNMACA	3593	ASMAYCAQE	3634	AWHMYEMSE	3675	CIAIITRGP
3553	AKNLTMLCR	3594	ASMKRQENL	3635	AWTEYSLSC	3676	CIARKMVDD
3554	AKSYANDWA	3595	ASMTDFVGV	3636	AWYHHCHYD	3677	CIFPRNMAT
3555	ALHHFVKAQ	3596	ASMWQHFVL	3637	AYAFVGRDN	3678	CIGSIPMAA
3556	ALLMQTFAH	3597	ASQGVWQSH	3638	AYDRMGGTY	3679	CIHAIPRER
3557	ALMVFKSMW	3598	ASSTAWQWP	3639	AYEQMPESC	3680	CIHARMVGM
3558	ALMVNQGNA	3599	ASTLCNRGQ	3640	AYHNVKEWN	3681	CIHDHKECV
3559	ALNLSWWPM	3600	ASTWQGHTD	3641	AYSSTLHAE	3682	CINWTRKAP
3560	ALTAITTSS	3601	ASVESRIRI	3642	AYYEPRGWN	3683	CIRLLCRDK
3561	ALTNHMICP	3602	ATAAENIYE	3643	CAAQEQRSR	3684	CIRYEMGAA
3562	ALTYRDSMY	3603	ATAVTANAN	3644	CAAWRQTQE	3685	CITQRQAMW
3563	AMLKYAYPG	3604	ATCKPDSCD	3645	CAAWYYGRT	3686	CIVLRSMGQ
3564	AMSWTQKGG	3605	ATDQFGHHD	3646	CAFRTAQMA	3687	CIVVYSRSQ
3565	AMVGYHCRE	3606	ATGKMMFPS	3647	CALIKHQQP	3688	CIWKHSHFV
3566	ANFPVNTLE	3607	ATHCINRAH	3648	CAMFRNVAQ	3689	CKAHYHDNF
3567	ANGSVNCEE	3608	ATIGRSCWD	3649	CAMKGNVAQ	3690	CKDAMNTGG
3568	ANGTLAECG	3609	ATMPWDGNE	3650	CAMLKMKYD	3691	CKFLLNQAR
3569	ANMYSNVTS	3610	ATMQESDRI	3651	CAMPFRQSV	3692	CKIASNEAF
3570	ANTSGWTSG	3611	ATSPSEHTN	3652	CAMSRMVFS	3693	CKMEFMRVE
3571	APDQANVYE	3612	ATSSQVATI	3653	CAMWTSASP	3694	CLCLTWAFG
3572	APMSYSDMD	3613	ATTFTRHNV	3654	CAMYYHRAC	3695	CLGYEKQCP
3573	APRVHSGNC	3614	ATTIWHNHM	3655	CAQWIKEQQ	3696	CLHYTIVNN
3574	APSPKAVYF	3615	ATVFMQETT	3656	CASAFVSYQ	3697	CLMEVQWGT
3575	AQEFIPSHP	3616	ATVHMERDC	3657	CAVGRHMAL	3698	CMDEVHMCT

[2339]

3699	CMMPYSRTN	3740	CVDQHHLTP	3781	DAIRWNKMV	3822	DPPKEPKEG
3700	CMYQSRAAY	3741	CVFAFGMKH	3782	DALMREAGH	3823	DPRDYDHHS
3701	CNTPYWQQA	3742	CVFAHNRGF	3783	DAMYYNFIT	3824	DPTGEGDII
3702	CPFLMNRAF	3743	CVFMDRMAY	3784	DAQSSPTAF	3825	DQGVMRQNE
3703	CQQRYNVGF	3744	CVFPRHHMF	3785	DARFKPMMC	3826	DQMGHTWQG
3704	CSAYYMKCS	3745	CVGLVMRPQ	3786	DASMYYECS	3827	DQMIQVEFT
3705	CSCWYQVQF	3746	CVGPFGSMM	3787	DATHMMCMH	3828	DQQPYVSGL
3706	CSDNLYCAM	3747	CVHCYNRGY	3788	DCEPREGYW	3829	DQRYGTSFS
3707	CSECWCKYW	3748	CVHIRNIGW	3789	DFGAEKASE	3830	DRCYFGMGH
3708	CSHDFTQYQ	3749	CVHTVIRGW	3790	DFHYNSDWW	3831	DRNWYRHPG
3709	CSIQWWQWG	3750	CVIMIGELS	3791	DFLIHGYSG	3832	DRRPVDCMN
3710	CSKLHDHIN	3751	CVIQRMHVY	3792	DFMKMQQHG	3833	DSASIPNYE
3711	CSLQQMSIR	3752	CVKWCNRCS	3793	DFYTMAEHF	3834	DSFRKGTFA
3712	CSNMYVASE	3753	CVLASRADR	3794	DGFKMQFDR	3835	DSFTYTHGW
3713	CSQPMSMPN	3754	CVLMRWAWP	3795	DGTDWDQFS	3836	DSIFPPHTI
3714	CSQSFQWSK	3755	CVMPRHRFP	3796	DHSFHWAYD	3837	DSMLNGMGW
3715	CSSMWFLWG	3756	CVMSHLSGK	3797	DHSHNHNFF	3838	DSQAAPTCP
3716	CSSPMSLQD	3757	CVMTYWEAH	3798	DHVPISTYF	3839	DSRTVFVDN
3717	CSTSRMASY	3758	CVMVWAEKG	3799	DIERSSRSM	3840	DSVDWFTSG
3718	CSVNRHHYW	3759	CVNHARQMS	3800	DIKCTMTES	3841	DTFGVACYL
3719	CSYELITNG	3760	CVNPTTTFG	3801	DIMCGQEHK	3842	DTGMDAKYP
3720	CTFHTMRGH	3761	CVPSYYNNR	3802	DITINSFGS	3843	DTKSYQAPA
3721	CTFIRHAMQ	3762	CVSYMFSQC	3803	DKCRKDWHF	3844	DTKWLHWMH
3722	CTFLRSTFA	3763	CVTSLYGQY	3804	DKKFLECVI	3845	DTTNRNMYS
3723	CTFNQIRGP	3764	CVVEYPWTW	3805	DKLLMPFDR	3846	DVAEAWKHA
3724	CTFQPNRVC	3765	CVVGKAKLA	3806	DKLYVDVMA	3847	DVAWHHWTF
3725	CTGFMGWCM	3766	CWCCWNNPG	3807	DKMIAMVAD	3848	DVDRGRSMS
3726	CTHPNALSP	3767	CWEWTMINC	3808	DKNDKWHAM	3849	DVDRIMPQP
3727	CTIAIRDSS	3768	CWNDEPEKC	3809	DKYLCYTTP	3850	DVDWTYMTL
3728	CTIMAMQSH	3769	CWNYVHRAQ	3810	DLYVSQKAH	3851	DVHFKMTMA
3729	CTIRKQQMT	3770	CWRVDRNHS	3811	DMAQPQAAE	3852	DVLCLHNHF
3730	CTKFEHRSL	3771	CWSNVHRGS	3812	DMEWEHCHL	3853	DVLFTDVGF
3731	CTKHMLQGE	3772	CWSVRHAWQ	3813	DMFMIHAVQ	3854	DVTFRWGIC
3732	CTMHMDAKY	3773	CWWLWNARE	3814	DMHMWVHGM	3855	DVTLMENAR
3733	CTMWCYGCP	3774	CYCCLRCHW	3815	DMMYVRKDA	3856	DVWFNHTVS
3734	CTNVRMKYE	3775	CYCTNEHPS	3816	DMYQFPMHG	3857	DVYPPQGLM
3735	CTRQTTFHG	3776	CYPACICTQ	3817	DNQMKNEMR	3858	DVYQLHAMQ
3736	CTVSKNYGQ	3777	CYVYRHVIC	3818	DNRCIIGHN	3859	DWAKSDRIT
3737	CTYQYQAGM	3778	CYYLSIVVC	3819	DNSTQIIGE	3860	DWFIQYHRG
3738	CTYTPLRAP	3779	DAASDMFDP	3820	DPCWIWKTI	3861	DWSWYFQGS
3739	CVCYHNSSQ	3780	DAGYFTLHA	3821	DPGRFQIGN	3862	DYAAISSMA

[2340]

3863	DYAKQDSSC	3904	EGKWIGLQT	3945	EQARIDTIG	3986	EVYITGHEI
3864	DYNLAGLLP	3905	EGMKSWNST	3946	EQDQTKTMH	3987	EWKFCWDAI
3865	DYRQWDQQY	3906	EGRSSENND	3947	EQELPVTLK	3988	EWQHVGATC
3866	DYVQVNQQP	3907	EGSPYFMAK	3948	EQPQKPSAW	3989	EYCRSNFQH
3867	EAASADCQD	3908	EHALEHTQL	3949	EQTKVMHCC	3990	EYESCICAS
3868	EAAYGCKIG	3909	EHELLKQYE	3950	EQVSKGVYA	3991	EYLRFIAHD
3869	EADKYPYCP	3910	EHGMILRLG	3951	ERGPTDNLH	3992	FACMRQMAY
3870	EADWNIGDT	3911	EHLFLIHEA	3952	ERHESHQEF	3993	FATGKHHYL
3871	EAERIVRGP	3912	EHRQKMAST	3953	ERINKFNAD	3994	FATHTYDYW
3872	EAFCLHGWA	3913	EHRTAFFYS	3954	ERRSAESLQ	3995	FAVPSKVTM
3873	EAGFWTQQA	3914	EHVFEEVHS	3955	ESAQSEGEF	3996	FCYRMTDDR
3874	EAGSFQHCA	3915	EIDNVCRQA	3956	ESDIGHPYK	3997	FDMEGFREL
3875	EAHNVFDNP	3916	EIGCAHAEE	3957	ESDIMRSPM	3998	FERTVHTDK
3876	EAHPHWWVT	3917	EIGRNCALF	3958	ESEQSQDNH	3999	FGDWKHGDG
3877	EANDSLKAN	3918	EILLSMMEQ	3959	ESGPKWRFP	4000	FGPGSHGNN
3878	EANLSNVRA	3919	EILQPQVMI	3960	ESNGFVNGV	4001	FKGPYHYQS
3879	ECFPSIQNL	3920	EIMCKGTMG	3961	ESSRTGSHA	4002	FLFYLNMHE
3880	ECGLSHDCG	3921	EIQQVGRDF	3962	ESTESHFCE	4003	FLMRRNGAD
3881	ECHVRDSHL	3922	EIRHWGWLV	3963	ESTQMCTCC	4004	FLTERCQTE
3882	ECIKQSECA	3923	EISNVRCAP	3964	ESVKFGKMP	4005	FMNHFMMWT
3883	ECIQVVRGS	3924	EIVPTMHCK	3965	ESYTLYQHL	4006	FMQKECVEE
3884	ECKDVEDIF	3925	EIVYQHISQ	3966	ETACNHTAW	4007	FNADIAHCD
3885	ECMLCPEAS	3926	EKHNFTQMV	3967	ETASQQHDA	4008	FNMLVLETW
3886	ECQTSPVAY	3927	EKKFNHQIN	3968	ETCNRGRLF	4009	FNSSFMGQA
3887	ECRAGDPVN	3928	EKLYSVSVA	3969	ETFAMNAAW	4010	FPHKWNAMS
3888	ECSLKDHHS	3929	EKRLESAGW	3970	ETHFYEVKF	4011	FPVDMQWMY
3889	ECTANTHYY	3930	EKSNLFTVR	3971	ETLAKGYDW	4012	FQGPIAPME
3890	ECWMQNEHA	3931	EKYMYHDVQ	3972	ETLRMFGVY	4013	FSAMFRRHC
3891	EDCALHQSE	3932	ELDFIIMPG	3973	ETMPNKLQY	4014	FSGNLSTYA
3892	EDHWLFWKQ	3933	ELNGAWWKA	3974	ETNIDNRAM	4015	FSKLTGQDF
3893	EETILIDCH	3934	EMGAVGSFK	3975	ETRINIEAH	4016	FSLKTESRI
3894	EEVMLGCLQ	3935	EMHRNELAM	3976	ETTNAAWFR	4017	FTDDDVNRF
3895	EFAADAHGM	3936	EMMVKPSYC	3977	ETVSGMHER	4018	FTGLCDIWQ
3896	EFAGITDMG	3937	EMRIQTAQQ	3978	EVCFDVCGE	4019	FTRALLDMG
3897	EFCEYAQAT	3938	EMVLCVTHE	3979	EVCMAVVDN	4020	FVRTKEDLV
3898	EFHQQVNTW	3939	ENDQIAHEQ	3980	EVDHQFHFC	4021	GADDKGTHC
3899	EFMHAYEAT	3940	ENGYMHQYL	3981	EVGPGSHCE	4022	GADIYIETS
3900	EFSECHAFD	3941	ENIKTPFGT	3982	EVNHMMCDM	4023	GADKLFMYG
3901	EFSNPCSFG	3942	ENTYQDTIP	3983	EVQKCMFVE	4024	GADNVQKIC
3902	EFVKAFMQI	3943	EPKDVSHMT	3984	EVRWIPWHY	4025	GADPFEGQN
3903	EGHGAVLNP	3944	EPQSFNHWS	3985	EVSLFSCEF	4026	GAEIEQLAF

[2341]

4027	GAEYDQRVW	4050	GCRFVILDH	4073	GMAVINEVE	4096	GVCFKLTDN
4028	GAHFIMSSL	4051	GCTKYGRHN	4074	GMGKYNQWW	4097	GVFQCPQSV
4029	GAQFCTQAR	4052	GCVVKNTHS	4075	GMGMVEIVG	4098	GVHQKNCPE
4030	GARHNGIFT	4053	GDKWISIQD	4076	GMWVNPMAL	4099	GVKAREIEI
4031	GARQFNCLE	4054	GEMCNVSGN	4077	GNQHWDVCG	4100	GVNDLRHVA
4032	GASKLSHER	4055	GFDQRGMEG	4078	GQCYCYVQP	4101	GVNHRMDHF
4033	GASRYQWGN	4056	GGTEIAMCF	4079	GQSGLHVCY	4102	GVRVFMSDV
4034	GATDCGSCP	4057	GHAMFNSRD	4080	GRIQRMESD	4103	GVSMCYNWN
4035	GAVECSEYK	4058	GHFQTMMAQ	4081	GSAQEPIWA	4104	GVSSTHQHR
4036	GAVPHPVCA	4059	GHITCATMS	4082	GSILSIGFH	4105	GVVANLNST
4037	GAVQKLIQC	4060	GHKQCFHQT	4083	GSSWIDQQG	4106	GVVKMENMA
4038	GAVSANWQS	4061	GHNCCHMIE	4084	GSTRCVKWV	4107	GVYTTEQCV
4039	GCASIGIHE	4062	GHVQESWMW	4085	GTEHHPFFD	4108	GWEYHPAFG
4040	GCDWLMCEL	4063	GIAQKNRNP	4086	GTERLPAHQ	4109	GWTNYALMT
4041	GCERMERVW	4064	GIEHLQNTS	4087	GTEYERHTF	4110	GYHWTNTAS
4042	GCFFSQRGY	4065	GIMCSGMCP	4088	GTFWRDSQL	4111	GYLPLEIIA
4043	GCFYTQRAY	4066	GIMSVMGIG	4089	GTGKHAWCY	4112	GYMPCMCQD
4044	GCGKLGVDR	4067	GINHRHVSA	4090	GTHYDASAG	4113	GYVQDFHGP
4045	GCHQLMNFG	4068	GKKDQVKGG	4091	GTIRLDTQY	4114	HAEQALSVA
4046	GCIRSISWP	4069	GKTELHVEV	4092	GTLEKRAIF	4115	HAGSCTLND
4047	GCIRVDSGF	4070	GLGFKHSMG	4093	GTLSYMMWW	4116	HAGSWQYGP
4048	GCMHYVLTE	4071	GLVEIHYTE	4094	GTSTIRAEH	4117	HAMPFCGRS
4049	GCMQSIHMQ	4072	GLYPHQAMF	4095	GTYIELACF		

[2342] 表73

[2343] 驱动骨髓组织嗜性的AAV5 VP1衣壳多肽中581至589区域的序列

SEQ ID NO	581-589 序列						
6118	DVDVVFRWV	6126	AAQTLECHK	6137	ASKFEIACF	6148	CYNDNWVDV
6119	GNMHNFSAW	6127	ADHLACALQ	6138	ATCLIVCEA	6149	DAHKSMKSQ
6120	QVVKFTGYH	6128	ADSDVHSEP	6139	ATNNADCGV	6150	DCLHCWDFC
6121	VQMNLDAAP	6129	AHENYRYHQ	6140	CARANCVLS	6151	DENWRDTCG
6122	YSFATGNAA	6130	AHLNCDKSM	6141	CFDQPLYID	6152	DFLEEHSHM
6123	MVDPYGCEG	6131	AHRCGCFKC	6142	CMRHGDRCA	6153	DFNFLRQQA
6124	CEMQVQLCV	6132	AILKMTKVL	6143	CRIDHMETY	6154	DFTQLRMSG
6125	AVDHYSNGA	6133	AKTGGEPDS	6144	CVMEDNEAP	6155	DGSKNQHSY
		6134	AKVNMMKQF	6145	CWEEGMPFY	6156	DIAQLILTT
		6135	ANNTYPVCA	6146	CWILHHRCG	6157	DIGRNVGEC
		6136	AQWISPWPT	6147	CYEHDHGQA	6158	DILGKGEPC

[2344]

6159	DIMSYATGI	6200	HWCIMKGMR	6241	MGHYCHARR	6282	RSMYANNDV
6160	DKNYPTHKQ	6201	ICEYFRVQA	6242	MHQFRCHYD	6283	RYNNCPNLW
6161	DLGCVQENF	6202	ICEYRYCIF	6243	MMAPLKVFD	6284	SAMSRAVYY
6162	DQECNEIGF	6203	IDAWGHFSE	6244	MNMEDRCTF	6285	SEKTQEVCL
6163	DTDCHSYMD	6204	IGYSTWPRE	6245	MNNRWSDCD	6286	SGEGITKMY
6164	DTNDYTSCD	6205	ILNTCLVDQ	6246	MTEPVLLDT	6287	SILITPYAN
6165	DVIRSESRR	6206	IMDVRVKRS	6247	MVQHIEQWE	6288	SLCWEVDLM
6166	DVSYHKKMH	6207	IMVMINTPY	6248	NCWLPDNPI	6289	SLYITPYHT
6167	DVVSMKWGV	6208	INTNYPCLN	6249	NFKQHDTFC	6290	SMYTNCTFF
6168	ELASSWLVI	6209	IPCILELVV	6250	NGFYWVFVM	6291	SQKQICGDT
6169	ELLEFNAMQ	6210	IPVMPCNCE	6251	NLCPNTVDV	6292	SWLPENPVH
6170	ENAYKTYPL	6211	IWILRARGD	6252	NMGCPMMTE	6293	SYAYSHSAT
6171	ERNRLHHRY	6212	KDDTQSHMF	6253	NNSHLYHDN	6294	TNVYQNVEA
6172	ESHALHMCN	6213	KEQRPPNAG	6254	NWDSPSNKS	6295	TPPCCVIQF
6173	ESSYWDKPK	6214	KFVQYQHYV	6255	NYFPTHAAW	6296	TQEACHHDP
6174	EVVRQKVIC	6215	KGRWIFEYA	6256	PDQVDRNTP	6297	TRSYFDQSC
6175	EYSWSAQYN	6216	KHDGEIRVL	6257	PEKAKMVWH	6298	TSNTFQRHG
6176	FAAMCNKWD	6217	KHIVGQDGG	6258	PPADPATES	6299	TVKYHQDQD
6177	FCSTYNAGT	6218	KIYLNQILR	6259	PQTMMCKRW	6300	TVNNDPNGI
6178	FEQSEHKPT	6219	KLQMGQDNS	6260	PSMLWSVVV	6301	VCHAFTCNL
6179	FSHAHNWMT	6220	KLSDHEMMY	6261	QEGNGFHNT	6302	VCTIRAECP
6180	FSHRFDAIF	6221	KQNDTVKIP	6262	QEKLWEQGS	6303	VDTIRFPRM
6181	FSTPARINN	6222	KTDRMAMMM	6263	QEVRSEHYS	6304	VEKTHNKWM
6182	FTAQLAVWM	6223	LAFTMMWPW	6264	QFGPGCFKP	6305	VESMNHASM
6183	GDMCEDMCT	6224	LATFWGHQD	6265	QQIWLCQEI	6306	VGICRLKAV
6184	GHDFQKHAE	6225	LCAPTQWYM	6266	QQLCVDESH	6307	VHYYACHRF
6185	GLHQFSICM	6226	LCGKAGLYH	6267	QRPTCKNTA	6308	VKEPATYKY
6186	GPAVPEQYG	6227	LDLCMLFLF	6268	QSMVSCYEG	6309	VNAGEVYRS
6187	GRMLVDRHA	6228	LEMQRGMET	6269	QTRKNEIGI	6310	VPSHMPHHP
6188	GTGPMHHKW	6229	LFDKDIARK	6270	QWMLRKEVQ	6311	VTDAAMFMG
6189	GYDPPEGMV	6230	LFKFHSGLQ	6271	QWTACQATN	6312	VTNPLEGRI
6190	GYYSMHHEI	6231	LGLTEGQSM	6272	QYMLAEDWR	6313	VVEMNGDDD
6191	HATTVEFCI	6232	LNGWWDCTR	6273	RADGLATVV	6314	VVGVPARAQ
6192	HESKTAHRP	6233	LPCTSIMWY	6274	RCNMMGKGA	6315	VVQYRLDPP
6193	HFHQQRKTY	6234	LPEHIQQMH	6275	RDRYCDPHQ	6316	VYTSTQPQE
6194	HGEPFCMAP	6235	LPENRNKRH	6276	RDRYCDPLQ	6317	WTHTVTSDW
6195	HHQQCKLDW	6236	LPKSKYMFQ	6277	RKFWLAGKF	6318	WWDHLPELM
6196	HKMECAQSM	6237	LQANEWNMC	6278	RLQIIMYRV	6319	YAFNTPWMT
6197	HPHQIPIDI	6238	LQQPIVHIE	6279	RMKGAKRWY	6320	YVEFEDQNG
6198	HRPEQLHEG	6239	LVMCPPFFG	6280	RNEYQSACW	6321	YVYDAQRQD
6199	HSWMRKSQC	6240	LYDVVHRIW	6281	RRNMPIPSW	6322	YWRYHYYYM

[2345]

6323	QNKAFLDAV	6364	AIIQWACWA	6405	ASMHEETCL	6446	AAWDIELNE
6324	FDHNAKVLD	6365	WPTMLSAHM	6406	DMTLQSVSS	6447	AAWVSGKLY
6325	YNNIMQIMV	6366	DFNGSNWLP	6407	EVAFTTSSD	6448	ACALDGASN
6326	FWLGLFPAK	6367	FTEPMCMQG	6408	SPQQPSVFH	6449	ACEFGSLNQ
6327	NNDHYDSES	6368	KTDKAGLME	6409	KSTYYNLKW	6450	ACFHANNWH
6328	MCQNVEKEY	6369	KWEQRSKMS	6410	SQERMDYDG	6451	ACHVQLSDA
6329	QMDWVAECR	6370	TLHYFHGYQ	6411	MNLQIGSKG	6452	ACKHEISHA
6330	MSSRKEFLT	6371	MNHQCMLWD	6412	EIHMFGQSE	6453	ACLNMWEER
6331	SCRMVYGRA	6372	IYCLCQDSD	6413	MPWLDKPPC	6454	ACPSAELCY
6332	AVHKQTAAW	6373	TLNSRHPEC	6414	QCTAHNTCM	6455	ACTSSQVKG
6333	DCAWMQKEW	6374	ACTKQKTVQ	6415	DQRHGNVSP	6456	ACVFKCWWP
6334	QKNVSTKYD	6375	KTCLCIFTK	6416	EAGCKMWPT	6457	ADCQPQSFN
6335	DYVDQCQVS	6376	SSNFNHHID	6417	ATMNLSWGL	6458	ADWLYMQSC
6336	TAVLYETEF	6377	QFMMGQQTP	6418	QAKMILDGT	6459	ADYDMLTAY
6337	VPDYVPMNW	6378	IAVASHAHG	6419	INYISFVDR	6460	AEIMCSQDH
6338	TSRDGGWAY	6379	KTGMGDLFL	6420	WKATPGGQC	6461	AEIQNWFAK
6339	LPLPMQSMC	6380	RITQGESGM	6421	YCECGHWPN	6462	AEQHTLGMG
6340	DGGRANTLH	6381	STVGQNHCG	6422	SNVACIQAF	6463	AFDTMVCGT
6341	QADNGTMRV	6382	KVNCRMIIM	6423	QSRCVDNTV	6464	AFFTNMSPS
6342	MFSHGHVMF	6383	YVDHWTLGD	6424	YSDPKNMSI	6465	AFGPYWMEK
6343	PYHNPWECR	6384	TMPDRDPTY	6425	ASKQGTHDY	6466	AFHLWKEDQ
6344	TTYPHNTHG	6385	QKAIVAQND	6426	NHSLCWDSK	6467	AFKSFDSHI
6345	EQTYYKITW	6386	CMEVTCFVV	6427	HSMMPGWPS	6468	AFRAICAHG
6346	TSLCAKPHF	6387	LESVKMDCN	6428	EWMQNWISM	6469	AFWIEQMDY
6347	FCHEVFSHT	6388	VLKSNLTGM	6429	AAADVEGTT	6470	AGDLWEIGS
6348	CKVDIHDYM	6389	AWMCCARHD	6430	AAAPWCEFR	6471	AGEHSRSER
6349	ETSTHCCKV	6390	RNDNTCHAY	6431	AAAWPEWTM	6472	AGFPFEKFA
6350	PAYRMECCE	6391	MYDTMGAWC	6432	AACNEDHHW	6473	AGGHIHADH
6351	GMVELGSCW	6392	ILEAIVNFW	6433	AACSRHYES	6474	AGQFRKLPY
6352	EAMTCTKNH	6393	AHAWLQHWA	6434	AAELYKCEN	6475	AGWPTSREM
6353	DNVFSNSPN	6394	ATGLTDHQW	6435	AAFEGEACF	6476	AHMGQVWWF
6354	HILIMQGVS	6395	DVKSWAVCD	6436	AAFEWMRAP	6477	AHSQLVSDR
6355	MHGIARMQQ	6396	PDLWYEKSS	6437	AAGWRDGRR	6478	AHVLNTFGN
6356	HMDHCQCHE	6397	SECEVLCSL	6438	AAHCEYSTI	6479	AIEFWDDAL
6357	EKFWFKTMN	6398	RYPCCQSPF	6439	AAKTFYWAR	6480	AIFHACNSG
6358	QLDTYMEGR	6399	MNSFYRAEW	6440	AALASDHDF	6481	AIFRSEITY
6359	NMKPKDVFR	6400	RTRISQMFG	6441	AAQFLHIVK	6482	AIGNLTGAQ
6360	QPGVRTDDY	6401	RDVGEDNSA	6442	AASCGYHGW	6483	AIQASMDAF
6361	WCNWQFAQH	6402	ATGTYNLQE	6443	AASQAQSLT	6484	AIRCESDAA
6362	GILVQGGSM	6403	DQEAQVVTG	6444	AAVCTQFSL	6485	AISVGEVTT
6363	ILKQASCKG	6404	PMELWRTWY	6445	AAVGSAQIE	6486	AISVRHQHH

[2346]

6487	AITDALLYR	6528	AQFSHLHSY	6569	ATTYWHRGG	6610	CHNTANHIQ
6488	AITSASAKN	6529	AQHSHHECY	6570	ATYYWQFVW	6611	CHSYFWMPD
6489	AIVGWKTPP	6530	AQSMPYSRH	6571	AVELAPCTN	6612	CIATQAFMY
6490	AIWIEEFFH	6531	AQWNQLYGS	6572	AVFWTHQEE	6613	CIHQDAPPH
6491	AKAHVCTIP	6532	ARATWPRTA	6573	AVMPNHWQW	6614	CISQAVMPP
6492	AKDQLNITR	6533	ARAVDGLNN	6574	AVMYPMRLH	6615	CKERCGYAN
6493	AKEVTTCNW	6534	ARFIREHEW	6575	AVNLKQFKT	6616	CKLQKCYYM
6494	AKHTGVTFG	6535	ARLWIPGEP	6576	AVNPAWKHD	6617	CLCYFCCDA
6495	AKKMWHCDQ	6536	ARRMDDDWI	6577	AVNWKQDGY	6618	CLHCPIRRR
6496	AKYGPWENN	6537	ASCAFAEQY	6578	AVRDHELNN	6619	CLSACDLLV
6497	AKYQWQSYP	6538	ASCTYYHKW	6579	AVSMACLNT	6620	CMMMYNTCR
6498	ALAWKNYDK	6539	ASEHYMEVF	6580	AVVDWQALE	6621	CMMQCMDMM
6499	ALCTENAER	6540	ASGIEGEQG	6581	AVVPLMESA	6622	CMNPRSYYF
6500	ALGPLHQGD	6541	ASGWIQAKT	6582	AVYEALRCL	6623	CNAKQQAAW
6501	ALMPSMFLS	6542	ASIEKAEMH	6583	AVYFVPCRP	6624	CNCASRTEE
6502	ALQQNSLGE	6543	ASMDHNLDD	6584	AWEMEMEYE	6625	CNDLRNQWE
6503	ALVYAHWWG	6544	ASMTKHNNF	6585	AWNPSGFMG	6626	CNNPTTRRW
6504	ALYEADQFD	6545	ASNHRLEKY	6586	AWTRCIVCA	6627	CNWSKSSIA
6505	ALYQIKWGF	6546	ASQPFNPWV	6587	AWVMSVMLA	6628	CPGVLQSCD
6506	AMDVIIEGA	6547	ASSARSIFE	6588	AWVNCISGS	6629	CPQQCIWYG
6507	AMEYLRLQH	6548	ASTAPVADQ	6589	AYAETSTHP	6630	CQVEKVENE
6508	AMGTTEVQV	6549	ASTNSVWHV	6590	AYDNHYLAP	6631	CQVPAWYDP
6509	AMHLQERMP	6550	ASVMQLLCP	6591	AYQHHEGVV	6632	CRVNDVMQN
6510	AMHNREGAI	6551	ASWLYDNNC	6592	AYTESQHNL	6633	CSCTTDTDP
6511	AMIQCAMPP	6552	ATAMPQVRY	6593	CAADCGKGW	6634	CSDWYVECK
6512	AMWWGQAMG	6553	ATCCYTHFD	6594	CADPQERCD	6635	CSEVLEDNA
6513	ANAAIREHL	6554	ATDDQMIKG	6595	CAQNLNLWL	6636	CSLPHVVHY
6514	ANAERCNLA	6555	ATDHSHPWP	6596	CASAVQEDC	6637	CSQRIHSHC
6515	ANCRKDPMP	6556	ATDPCACNI	6597	CASEHSCWH	6638	CSSDYSTLS
6516	ANENCSKGS	6557	ATDVVDECH	6598	CATFQNDVL	6639	CTCQAKIFS
6517	ANGPHTDCI	6558	ATELYMNGH	6599	CCELLIENT	6640	CTDRFAIGA
6518	ANILWYHAS	6559	ATERWLEGN	6600	CDGRRLSMH	6641	CTFPNVHEA
6519	ANKCLDMRD	6560	ATHDAMVTN	6601	CDVQQPTMM	6642	CTHSAFHYV
6520	ANLMSAKME	6561	ATKCYEYTW	6602	CDYHFKNWG	6643	CTIWKQCGH
6521	ANPLCELQQ	6562	ATMPQECGA	6603	CEPLVEIHG	6644	CTMCYYQNN
6522	APRLTEACL	6563	ATNHSHPWP	6604	CFDYMELSP	6645	CTQMVIHQE
6523	APSWRNHMC	6564	ATQFGMTED	6605	CFEDCFYCH	6646	CTTAIGPSG
6524	APVLITNQQ	6565	ATQSVWMLN	6606	CFEPQCVIN	6647	CTVILPKVG
6525	AQDHRHSDW	6566	ATQTDYNIP	6607	CFVPDTEWG	6648	CTWFDCSQE
6526	AQDWSATQG	6567	ATSLPAHYD	6608	CGQQNFRIP	6649	CVDTHHANV
6527	AQEKRLYGQ	6568	ATTLHERTH	6609	CHADLFWEQ	6650	CVHATFDYT

[2347]

6651	CVTTDHCAW	6692	DEIQIGGSF	6733	DLQTSYHDQ	6774	DSWTGWNFA
6652	CWENTNVTC	6693	DEMTFEQLF	6734	DMARALKFF	6775	DTCKYLWTP
6653	CWEVVNTFR	6694	DEQHICMQT	6735	DMCTMAYTG	6776	DTCSRKHWM
6654	CWFIKRACA	6695	DESEMTSGG	6736	DMEYSVCHS	6777	DTDEYQKAQ
6655	CWYVWDLWP	6696	DEWWWSREE	6737	DMKMYEFGP	6778	DTEWKYKHN
6656	DACFVNQPN	6697	DFASRCNYV	6738	DMLMNLMEP	6779	DTKMQHTCT
6657	DACSHTWAD	6698	DFCFNESYW	6739	DMNFTEQYA	6780	DTKWYPSTG
6658	DAEDTPATV	6699	DFHCYSTER	6740	DMTWTMGRQ	6781	DTMLGPQMQ
6659	DAGKFRKYV	6700	DFMTLVTHQ	6741	DNACHTPFW	6782	DTMTCNRHQ
6660	DAHQENLNW	6701	DFQAYSWNH	6742	DNAVLDGVP	6783	DTNHATMAC
6661	DAKDMMMTC	6702	DFRHVDKHA	6743	DNMPQPVCH	6784	DTNITDQWR
6662	DALFNHAAV	6703	DFTQGMCTP	6744	DNRFIQDAR	6785	DTQYTRDFV
6663	DALSYMQME	6704	DGDAQSCAY	6745	DNSSEGSWF	6786	DTSWDRIFV
6664	DANEFILMH	6705	DGDCLQRHL	6746	DNVYKGPWA	6787	DTYMEADSI
6665	DANNTLQHS	6706	DGFNYELQP	6747	DPCPYIDHQ	6788	DVAEFTGVV
6666	DANYYFVWQ	6707	DGGHMMNQD	6748	DPYPQGRQK	6789	DVETTKAIN
6667	DAWPNQQMW	6708	DGHNFDGHN	6749	DQADCNWTW	6790	DVFKHLQCT
6668	DCAQNWSGS	6709	DGIWTHKAD	6750	DQASYDIHG	6791	DVGACFIQD
6669	DCEWMSTGA	6710	DGQFIFSQN	6751	DQCPLRSGT	6792	DVGPMNGPR
6670	DCGFQTAHT	6711	DGYNMLWDP	6752	DQKSQMQTR	6793	DVLPAVECP
6671	DCHKWDTNQ	6712	DHCEMHDCW	6753	DQLEFICWP	6794	DVLTSQELY
6672	DCKMHQAAM	6713	DHFHMELTK	6754	DQNSYGCIP	6795	DVMHDSRCS
6673	DCKWLALYP	6714	DHQSYYTQK	6755	DQYVPATGY	6796	DVMPIECQT
6674	DCMQDPAYR	6715	DIEFYCAYP	6756	DRAVKFSIL	6797	DVNIVCMDD
6675	DCMVMSQGH	6716	DIEMFGSEM	6757	DRGDQVMQT	6798	DVRSMYVMG
6676	DCNFRHQDP	6717	DIEYQSMGS	6758	DRGFAHLCQ	6799	DVTHGDPAK
6677	DCPEWGVGC	6718	DIGWLAPTA	6759	DRSTQDKFD	6800	DVTKHALAP
6678	DCQMILGDY	6719	DIICDCDMF	6760	DRTVQVRVD	6801	DVVDGEASQ
6679	DCQPMEWWE	6720	DILLDSRHS	6761	DRTYTYCWS	6802	DVVHGGTQL
6680	DCRHFQSQE	6721	DITLIMSVH	6762	DRVMMQRND	6803	DVVKNQGIG
6681	DCTEIVSGM	6722	DKGQLCSLQ	6763	DRWGMEEHQ	6804	DVVQNDCEF
6682	DCTMYCKQS	6723	DKIWIWQAS	6764	DRWQIPSKL	6805	DVYWDECVF
6683	DCTPYQGGH	6724	DKLARSVSC	6765	DSATFKSAD	6806	DWACQEDAG
6684	DCYSKRFTP	6725	DKMELNAHW	6766	DSECKNCMC	6807	DWCWYTCSW
6685	DCYYLKYSA	6726	DKQQAKDQL	6767	DSGGNMQVY	6808	DWNYRHWFN
6686	DDEYNLTGL	6727	DKQRPMCVC	6768	DSGKQPADQ	6809	DWVETAQES
6687	DDPNYAWEA	6728	DKQWFWFGH	6769	DSHCEVRFE	6810	DWYMHDQPM
6688	DDQLYAWGM	6729	DKTSCMSFR	6770	DSIQRLSAG	6811	DYEQSQKWW
6689	DDSLLFHMD	6730	DKTYLTRME	6771	DSNTGQKAY	6812	DYGTAGYPR
6690	DEASDLQYA	6731	DLAVMGKCT	6772	DSVEWTVHE	6813	DYIMHHHCG
6691	DEHPHPEHC	6732	DLMWYEGAS	6773	DSVRMLGKP	6814	DYLCINGFS

[2348]

6815	DYNSKMDFH	6856	EFCKANGPL	6897	ELCKTVFLA	6938	ESCILEEHM
6816	DYRPENTGC	6857	EFDGCQYFA	6898	ELGAAVMVC	6939	ESCIMQCMI
6817	DYTTAIYTD	6858	EFEQFTRYA	6899	ELGFWHGKG	6940	ESCLFHMRQ
6818	DYVRRAWMP	6859	EFHQITGSS	6900	ELLNPDNPI	6941	ESENNKFGH
6819	EAAWVVRKR	6860	EFMMPMIYD	6901	ELMLGNMFW	6942	ESEPNKHNP
6820	EADNTQRSS	6861	EFMPTVDTK	6902	ELMMNLIGM	6943	ESFPNKIQG
6821	EADSKFVDQ	6862	EFQWWNMSK	6903	ELRCRDSDQ	6944	ESGVEDMQN
6822	EAFAYDHKD	6863	EFTHWMQMQ	6904	ELRPITWVD	6945	ESHHKGDHM
6823	EAGFDAPGR	6864	EFTQLAYTY	6905	ELTKKVKAS	6946	ESKSHWTIG
6824	EAGIDKCCH	6865	EGCWHGHTF	6906	EMIHRHQVD	6947	ESKVGYELG
6825	EAHMIGRHT	6866	EGHAFNTFN	6907	EMKKLCDAC	6948	ESLTHEDGY
6826	EAICVEREY	6867	EGKPRLHDA	6908	EMKPWCQAL	6949	ESRAVLSMP
6827	EAKLQQEIE	6868	EGKTDVMSY	6909	EMMLVHQPS	6950	ESRLTMTCP
6828	EAMGCILRG	6869	EGWTVMVDQ	6910	EMNAVFSTD	6951	ESVNRYFKQ
6829	EAMHGDAIW	6870	EHCNHDMNT	6911	EMTKNDSFM	6952	ETAYTTRTE
6830	EARAHQNEI	6871	EHCQEMRWH	6912	EMYFMSEGT	6953	ETDHRKHYC
6831	EASANWLKC	6872	EHEAMLVND	6913	ENIQSIADV	6954	ETECWHKSH
6832	EASGLHKIF	6873	EHEFLGAKS	6914	ENKEYPDSN	6955	ETEMICEFQ
6833	EASNRMIDN	6874	EHFRACVVF	6915	ENMCNQGIG	6956	ETFACPSCR
6834	EATLGGWCD	6875	EHFSTYMRE	6916	ENRLCVGTV	6957	ETGMFKFSG
6835	EATPFVRFR	6876	EHGYSHTLA	6917	ENVVAIQTL	6958	ETGWKQNIE
6836	EATVWTWNP	6877	EHQEQQLVH	6918	ENYKEIPYG	6959	ETHFCIRNS
6837	EAWHFYEWP	6878	EHSYIMANP	6919	EPTASKHDR	6960	ETHQMGMDR
6838	EAYVNKDQW	6879	EHTNLEKYA	6920	EPTNIPDGF	6961	ETILDQHGG
6839	ECDFPMGTE	6880	EHTTWHQYT	6921	EPVRRTVMG	6962	ETISRVGMQ
6840	ECEVSHFID	6881	EIATHWQKG	6922	EPWMPSIAL	6963	ETLFNSYGR
6841	ECHHLWCDY	6882	EICHYNEQY	6923	EPWSIGKAA	6964	ETLPLYWQC
6842	ECHQMNEDQ	6883	EIDAMIPMT	6924	EPYCEGVAW	6965	ETNKQLNLW
6843	ECMQRYIGV	6884	EIDLVTRWE	6925	EQETHEALG	6966	ETQQNIWSW
6844	ECNQNWDEF	6885	EIEEQGWSG	6926	EQHVIAMDH	6967	ETQSPMCCL
6845	ECQNQFGWM	6886	EIFLNPQYY	6927	EQIKQLNWA	6968	ETSCTCEVC
6846	ECRCIPTVW	6887	EIGYRHRDC	6928	EQMKYYWFE	6969	ETSGQGVLW
6847	ECRSEMGLP	6888	EIGYWCGCA	6929	EQNEQKDMW	6970	ETTWNVICP
6848	ECTRGELTH	6889	EIHQVTHYE	6930	EQVPFCFTE	6971	EVADPMTWV
6849	ECVHKHSGS	6890	EKHLLNCDS	6931	ERACSMADR	6972	EVCPGSIRS
6850	ECYHMELMT	6891	EKKFGMSGD	6932	ERCHYPYIK	6973	EVERKVNQE
6851	EDTRFHQWT	6892	EKMVAEFQC	6933	ERETLGDQR	6974	EVGNPTQYD
6852	EEAEQEEWN	6893	EKQFGVKNR	6934	ERGMQVWPP	6975	EVHYCGNCG
6853	EEMHDDGDA	6894	EKTWHQRNM	6935	ERKYIMQFD	6976	EVIMSHLIT
6854	EESQRDMIT	6895	ELATLVDWN	6936	ERRSHTETA	6977	EVKYFWHCA
6855	EFAAEMTPA	6896	ELAWFMNTE	6937	ERYRGVIAC	6978	EVLMDWSWE

[2349]

6979	EVLQLLKAD	7014	FAAYDMVSE	7049	FMNVYFEHA	7084	FWRGITCNF
6980	EVLYGGQQW	7015	FACPEALGT	7050	FMQVGGINT	7085	FYRLDRCEM
6981	EVNSGRDWW	7016	FADPHTWFS	7051	FMTSLRDRF	7086	GACDVLSGE
6982	EVQGFIDDR	7017	FAEDIHYGA	7052	FNCEGQLYE	7087	GADEPQATL
6983	EVQHIHFPY	7018	FAIEPGLCT	7053	FNDEYPCSQ	7088	GADYFTECH
6984	EVQKILFLQ	7019	FATIVVHEE	7054	FNDSEPEVI	7089	GAGHYITAW
6985	EVQRNGIGM	7020	FATVSPGWI	7055	FNFQKTMDC	7090	GANAIWTEE
6986	EVSHFKVDE	7021	FAYENRLCD	7056	FNLEEWWSP	7091	GAYEYASKM
6987	EVSIQPQMN	7022	FCFIEDLAM	7057	FNLQYAGLQ	7092	GCASEHGRS
6988	EVSYTIGDI	7023	FCFMYPFHW	7058	FNVPFAKDD	7093	GCCQGSPFA
6989	EVVHLMMRG	7024	FCFYNRCWQ	7059	FPRESSNLL	7094	GCELIPIRP
6990	EVWSPMHNG	7025	FCHLEQCMI	7060	FQAGLANCA	7095	GCEMKYQSF
6991	EWGWAWLVG	7026	FCSCCHCRT	7061	FQFTAMWYM	7096	GCLELEHYE
6992	EWKMANVMQ	7027	FCVINELGT	7062	FQGHSRDGM	7097	GCTIAEGAC
6993	EWRHAWNCW	7028	FCYQKLNYN	7063	FQLYQNHGA	7098	GCTQGEQAT
6994	EWRMFINTA	7029	FDMQLTAYA	7064	FQMTKQTQS	7099	GCVGAGCDM
6995	EWTPFMNEG	7030	FDQPIQQCP	7065	FQMVEQQPQ	7100	GDCFKQTFH
6996	EWYLRHHMI	7031	FEHVIRTMW	7066	FQNLLDGCW	7101	GDGNRPTDI
6997	EWYQIRHID	7032	FEKWRHPRH	7067	FQTRPSCSF	7102	GDHCSLSEE
6998	EYAIKICNH	7033	FGAKYHDGQ	7068	FRVRWWHYR	7103	GDHKTIVMF
6999	EYASTACYD	7034	FGEPANVNA	7069	FRWNIDVTY	7104	GDKYCGVHT
7000	EYDHIDVWL	7035	FGGLCWLVV	7070	FSANMNADI	7105	GDLFWKKIH
7001	EYEGYKSAA	7036	FGGLVIDAS	7071	FSDGLNHSD	7106	GDYTKMVNQ
7002	EYEPIASGP	7037	FGKQRNEDY	7072	FSQHFNRYP	7107	GECTYVGGD
7003	EYEWLMAAA	7038	FGLWQQAGY	7073	FSQLMNTPP	7108	GFAQYGMAL
7004	EYFDNRCAQ	7039	FHAEWKGPP	7074	FTFDQPSGY	7109	GFEQYDCEM
7005	EYFHVVGVD	7040	FHTRVNACD	7075	FTHPSTDQQ	7110	GFFCFVTQW
7006	EYFPTARTF	7041	FICQQGDSL	7076	FTILLFVHN	7111	GFMEHPNSA
7007	EYKSRGREV	7042	FIRGSPMNM	7077	FTLLNSEGF	7112	GFMEKCYDG
7008	EYMERWFND	7043	FIRMIIVDD	7078	FTMRYQKLH	7113	GFNALQQAP
7009	EYSDSMVWP	7044	FKAEYQSGH	7079	FVETTQVMN	7114	GFNCEKPDA
7010	EYVRWNKQC	7045	FLINPAWQG	7080	FVKSFNNDV	7115	GGGHCWQNA
7011	EYYLLGCAD	7046	FLWGIRVLA	7081	FVQSSCQQS	7116	GGHMKMQME
7012	EYYRKQCAI	7047	FMGKTDVNM	7082	FVVDKYIYA	7117	GGRAFAGYK
7013	FAASSEIYG	7048	FMLTMDFWA	7083	FWNNNLFQS		

[2350] 表74

[2351] 驱动CNS组织嗜性的AAV5 VP1衣壳多肽中581至589区域的序列

[2352]

SEQ ID NO	581-589 序列
9118	AAMESCAEI
9119	EKYQIHWDR
9120	VTGTYNLQE
9121	ATGTYSLQE
9122	ATGTYNLRE
9123	ATDTYNLQE
9124	MWTKFNEYG
9125	QTLGGHMWT
9126	TLCHTDNMV
9127	GIAYEVLWD
9128	SRLPVHECP
9129	TTGTYNLQE
9130	MWTEFNEDG
9131	TFVYGERDR
9132	GQQTIEMSY
9133	NYAMRRDSY
9134	KPLASIGQY
9135	AAFVMNGDG
9136	RVAGEGIQP
9137	TCGHTRFFI
9138	ESAEPRMIP
9139	TTSQNSEAM
9140	NTQVISRWT
9141	AGCNTQNGH
9142	VGPLYAYYS
9143	QHSRAYMWY
9144	DIRIVSDGK
9145	TMAQKGLWF
9146	QCEYSNCCY
9147	NGTQCLNMD
9148	MTDNNGPLF
9149	CPEHNQVGV
9150	EMYVQRHSL
9151	LINSIVCGD
9152	ACKPCNGWD
9153	PLTIEVNCT
9154	SKGWADLCP
9155	DMTLQSVSS
9156	RVTIMFTGT
9157	RMFWDTSDR
9158	SMQYFFKTE
9159	SFKGFHLQN
9160	TTYPHNTHG
9161	KSTMLEHHD
9162	TCKWYEESI
9163	DICQKQSPN
9164	ILEAIVNFW
9165	DYEVCGNWS
9166	EWNQVWWHD
9167	SPNGRGLCG
9168	TSMEFTQHQ
9169	ASFFLNPLE
9170	MITNNCNIS
9171	MAEQMERDF
9172	VLTTMWKNN
9173	IATHNHILN
9174	MYDTMGAWC
9175	CKQLFCELG
9176	MWTEFNENG
9177	ECVSDYACC
9178	AGFSQFDCF
9179	RGKPILNQQ
9180	KSAQVYWHP
9181	AIGPALCLT
9182	QNKHNWWIM
9183	MASQGGQVR
9184	CGHLLRRAE
9185	WGDNKNFMG
9186	FVFAMIPGF
9187	TFMFNSEWS
9188	MAQPYGRTA
9189	CKNHYMMAA
9190	NSGVLIGME
9191	LAESKKRNK
9192	AFNGANSWT
9193	YTELMKKAC
9194	ETNDNMILH
9195	QFMMGQQTP
9196	VKAWWHDHQ
9197	GKIAHDQLN
9198	QVHMQRYYG
9199	MFVVNQNWA
9200	ILINDREWG
9201	DVTQTRYCY
9202	QTGVMAAFG
9203	ESRQVQDCY
9204	IMNPTAQNN
9205	CCCVPRSIY
9206	NRIQMTDFQ
9207	QPYWVAVDP
9208	KMLHKVWID
9209	QWMHCQLSR
9210	GHMFETFWG
9211	NCPMYDMIS
9212	CCCQFSFWN
9213	TLEPAYRHE
9214	NLSNLVSCR
9215	DRKQADQIY
9216	QPKMWMMAS
9217	KEFQLQMEP
9218	QMKGATEFC
9219	LETMMRHGE
9220	LCKFANGAA
9221	FAQQIINAS
9222	IMYLKRCCT
9223	EKRLEGWVY
9224	NVWSPAVRD
9225	DEYQQESFQ
9226	EICMRSGFI
9227	AVKAPSQMS
9228	QTDILRDVQ
9229	DNYWIVPFA
9230	INMPVSKDQ
9231	ENQSTMSHT
9232	FNTPYAHIG
9233	LESVKMDCN
9234	WCEAVGDYA
9235	SMSHNGQCL
9236	HAETDHVGL
9237	QCTLEISYV
9238	SGQPVDCMF
9239	AAAWSCMQW
9240	RTVPTQEMS
9241	LTPRSEGLP
9242	GDVSYDAEA
9243	QHKWHDVGV
9244	TAFQSLQQM
9245	REFAISGGN
9246	PHAETHWYT
9247	NLAPMIQVC
9248	STGSSKDQD
9249	HFEVRSVQR
9250	TMRLFWDMG
9251	NFVNHQPLP
9252	GKGQWTAQE
9253	VHGQIDLSL
9254	YISEYVMMV
9255	GVRNALNLS
9256	HMEQCWCNE
9257	MFFPGFVAH
9258	QCFLWNDCQ
9259	LIHWLKHDL
9260	KCAWIEVTP
9261	YPVHMMPPL
9262	PCCWEQSTN
9263	CSMPQQCME
9264	RRTDHGEPI
9265	GSKIIMHTE
9266	SQMKYAHNQ

[2353]

9267	EFNWLLWSG	9308	SDSKVQWGM	9349	KTMGIGGGP	9390	GITMHPMHH
9268	AEDTPYTIY	9309	WDGWVTTHF	9350	RVWSLEWHL	9391	LPPPEEKWL
9269	RTAEDCPFQ	9310	QTNGHPYFP	9351	EQPNMGDYG	9392	TTTNQFMFQ
9270	AKCTILNFQ	9311	QTSHQDQDC	9352	FGEFTMDEN	9393	RCGEIMGLD
9271	WFSIIMSKL	9312	HMGIAYNDH	9353	KFGFMYNEF	9394	MQLRIHTNS
9272	QTAEWYCGA	9313	TFVRKPSLM	9354	ERFHSYPLG	9395	ICHIRTRGN
9273	MHGIARMQQ	9314	QQLCYMSDE	9355	FLYTFMFQH	9396	MSGPNDTEF
9274	VTKYCDHEQ	9315	KVAEGQNML	9356	GSVMLMVIR	9397	NTQYSAKND
9275	DFNGSNWLP	9316	LHYRPTEQC	9357	KMAGIGCWY	9398	KQKTQNRNA
9276	IHFTSQSCL	9317	KTGMGDLFL	9358	RTRISQMFG	9399	TRETMNEWH
9277	QVKWTADPD	9318	AQCHDMWGH	9359	AQARTIGLG	9400	ERQAHNHYA
9278	CIQPTWDSY	9319	ATGLTDHQW	9360	MIRPSINGS	9401	SKGWADPCP
9279	DVKAFMGGV	9320	QWAPLQTWA	9361	TKGRNDHME	9402	RTEAHLPAP
9280	DPGWYGLAP	9321	SPHNCQHMG	9362	ATIAQWCVH	9403	GNECFHNVE
9281	DHLSNDAEM	9322	ISMRSECED	9363	RYNNCPNLW	9404	ECIFAPRMS
9282	NSIGWDSAS	9323	KIAAWNDFL	9364	FKIEWSQDI	9405	YFPSIQCYE
9283	TNFPTSMHG	9324	RFDQFWDTY	9365	KKHKGWCRA	9406	SSIYKEIQE
9284	RQTTYDCLD	9325	HQCNCLLSW	9366	GAANRIQSW	9407	THVKRFDYE
9285	SAKCCINAV	9326	MAMWIADTH	9367	NWTGNMQWA	9408	LYMMALLPT
9286	VHHKTNRSY	9327	AVHTADVCC	9368	CMEVTCFVV	9409	SGALKNREE
9287	FGSMQANDR	9328	SCAPDTQFG	9369	RVEMGCEPR	9410	EASIQASWN
9288	AVETGESAM	9329	CWILHHRCS	9370	SQIQIRQAD	9411	ACIPQMHSA
9289	MVENGLRLQ	9330	QVGIEFVVC	9371	KIGCMYSSQ	9412	LHSYQWRLG
9290	AIEHGAWKC	9331	TIVVIPIKS	9372	MWHELIADQ	9413	ACVERLHHK
9291	SMMEQKEQQ	9332	MPWLDKPPC	9373	PQEAQIGAP	9414	TVNWYSGFG
9292	AFLHGQMWQ	9333	YMSWYQIRN	9374	KTQIHMGDE	9415	IAKQNFMKT
9293	IEFTHSRKV	9334	AVYLDHVFS	9375	LPSHCNLGS	9416	FTQAIHHPD
9294	NWKLPTCMT	9335	KACMSHQGR	9376	KVESTTQLC	9417	YHCCFAMQS
9295	EHKGCMQAI	9336	TSIEGLKTM	9377	AYAMRNYTS	9418	NLERCNTMM
9296	ATGTYNLQE	9337	AWSYTGACY	9378	IIRQSEQYH	9419	HMHNLPVKP
9297	NGNHYNAMR	9338	QCVDCPPYC	9379	AHAWLQHWA	9420	HCWVNEKCE
9298	QHEQSNHIW	9339	VTCNWCVMP	9380	HEVMMCPAP	9421	FADKWTNQF
9299	IGRLNASTP	9340	KCADLIIRS	9381	EVAFTTSSD	9422	RQAWPKDGM
9300	MSNLEYYEA	9341	NVCDNGHEI	9382	PLNQEQIPQ	9423	CSESYRGPD
9301	TSCCYESSS	9342	MMKRNRFIC	9383	TSVIRLFWI	9424	VLKSNLTGM
9302	WINEANCLM	9343	SRVMIHVYV	9384	QFAQLWCAK	9425	NAYYDSHCE
9303	VESAMINQA	9344	WKATPGGQC	9385	GIWVRDTYN	9426	YPAQYTAFS
9304	EPRLYCAQE	9345	WTTSIAVDT	9386	LCTSTERQG	9427	GLFEGSPGA
9305	WPNIKHQPI	9346	DTHHVHGWQ	9387	DRVNLMWKD	9428	VTMIHNAFH
9306	FTHSCCCLS	9347	MGYAKHDSH	9388	CYMNFIVDG	9429	IYMSTGNTP
9307	SFSDTHLGW	9348	GCEEVGRCQ	9389	QISKPSVQC	9430	YIGMKSCSG

[2354]

9431	MFMTNVNSF	9472	TQCKQLTVM	9513	EFVLGCQGI	9554	DTAFMHHHV
9432	IAQSEADYW	9473	QSSHCKQMV	9514	QRSVFHQRS	9555	FQWRTEMNN
9433	AVHNVKDVH	9474	WGGMPNNQG	9515	PIWQGYWPW	9556	SGDWDFGEW
9434	HIHMSDRGA	9475	FLTCDQFEW	9516	HSMMPGWPS	9557	MNLQIGSKG
9435	GLFHPQFGQ	9476	RVNGSTCRH	9517	GVHHRCRMT	9558	CNGYWMCMS
9436	GPRHIEEDG	9477	ETHAAMNSM	9518	IQAVMASGW	9559	SQERMDYDG
9437	AYHCYMSWM	9478	DAKLTMHSA	9519	IAKIGNCVW	9560	SVGYIYDQS
9438	ITASIKDQW	9479	RYPCCQSPF	9520	CVDTNRMWQ	9561	GTWSDPWPN
9439	PAHINDQYM	9480	QGENAQHHN	9521	ARWFSTVEQ	9562	YPTMTFPEE
9440	AAMSSSHSY	9481	TSNHNMALE	9522	VFKPLDKRY	9563	AAQYMGIMK
9441	MNSFYRAEW	9482	NPGWQTIGQ	9523	EVRIQCKID	9564	HTVMEWSPT
9442	QATGMHRWQ	9483	CAWHMKQGN	9524	NGVGRSGDN	9565	GQHTTSHAG
9443	SSRMQMQGP	9484	NVMGTQAKE	9525	CVEQYVEQG	9566	RYQSHCWLH
9444	MEHTMDFGV	9485	TYIQYPTNA	9526	LSAQRLVCD	9567	LSTMFPEQT
9445	YLEMYSKDW	9486	DGMHRELYS	9527	KGNPLDGDT	9568	NIGVIPHIQ
9446	QWFTKGVGE	9487	QSAHPEPMC	9528	PVICVWHDN	9569	PWTTPENYT
9447	KNCIMLASN	9488	QAKMILDGT	9529	EIHMFGQSE	9570	YNCPWDQQY
9448	AYAPSKNAM	9489	VHVHCMIGS	9530	YCECGHWPN	9571	QWHDDQWST
9449	GLSMVLHFV	9490	TAALGHSFC	9531	TYESLQNSW	9572	DYAQPTGPD
9450	SPQQPSVFH	9491	DTRRNLCCD	9532	INYISFVDR	9573	FKHSLREWT
9451	PGGCYRMME	9492	ENTMWQSSQ	9533	MWGCKLFVC	9574	AARMVGEFY
9452	KTCLCIFTK	9493	NQREWHGLA	9534	CQMLCRPPV	9575	GSMPSTVPH
9453	ACHTGQSGM	9494	PDLWYEKSS	9535	RDQCRMESG	9576	ALAKFQCWS
9454	DIEFLDCEN	9495	ESSQEIKTC	9536	WVHPYSYCA	9577	AADIFAIND
9455	WTLQVMKIN	9496	YVDHWTLGD	9537	QSRCVDNTV	9578	QHHWKWSTE
9456	DVKSWAVCD	9497	NHSLCWDSK	9538	GRALGKKQG	9579	AAGAFLSGC
9457	LCAVCGDCD	9498	SECEVLCSL	9539	SNVACIQAF	9580	FDVQFMRVD
9458	DQEAQVVTG	9499	DQRHGNVSP	9540	HQYVHNWRQ	9581	ALQYCQGKD
9459	HFQPVFMQP	9500	MCFPYSCAA	9541	SVLHVQALA	9582	NERDQCQWP
9460	GHLFNNNEW	9501	ASKQGTHDY	9542	ATMNLSWGL	9583	CRMITMMNE
9461	YDYWMQAPT	9502	GNKHCMLGT	9543	WERCLELNQ	9584	EWHARHHYW
9462	ASMHEETCL	9503	GTKLIWPYQ	9544	TKMIIFDSW	9585	STSWLNDLT
9463	DCVLTMVDK	9504	QCTAHNTCM	9545	CWECCTDSG	9586	DEGYIMFAP
9464	MLPLEMKNL	9505	ACDQVNTVQ	9546	YSDPKNMSI	9587	QMVTQLNWE
9465	IMTIIPYGP	9506	IECSPRESI	9547	FPKGSECNW	9588	VVTQYVDAP
9466	LIEKQRWML	9507	RCADNPYFR	9548	EAGCKMWPT	9589	TVIRDSRSY
9467	VSVQSRAQL	9508	QESMCNYWQ	9549	KDGRNPHCL	9590	DEAKKSCWQ
9468	THSSCWYQD	9509	AGMYKCQED	9550	LEIGGKTRT	9591	AVNEAQDNR
9469	VFNLMLQDK	9510	SQAHQNRTC	9551	YKGMEKNPV	9592	NERSGSVLE
9470	HIKVEHEEV	9511	MGRQKDNVY	9552	DLIVATEWF	9593	TYASTRNFH
9471	MMLGYMEQD	9512	NANYWFYYD	9553	WHTIMTYES	9594	IRRSMQQQN

[2355]

9595	GFQRTHVDQ	9636	SFKSVSQYN	9677	NDMQMNWVH	9718	QFMGVSSHS
9596	HKHRMCKAR	9637	QAAVTKEYW	9678	MHDGYRGMQ	9719	VHPGFEKLA
9597	TAGQQPTFM	9638	EIIQAHLIN	9679	WVEPEQCFG	9720	NTFEQCTEH
9598	VQMNLDAAP	9639	KTGATNLAW	9680	TRCCEPYFQ	9721	NNRRQFGYV
9599	KDHQLWEVW	9640	GQFRSGAYR	9681	LSAKQQFFC	9722	QQSFFFGEQ
9600	GDCIIVISW	9641	PMMIIPYVD	9682	ECTKSMQSG	9723	QSMVACYTI
9601	RDVGEDNSA	9642	EPLVCISAG	9683	NCHQLGDNM	9724	DSMVYNRPQ
9602	PWMHTMLNP	9643	KGGLSTTNA	9684	PCQQLMCFM	9725	TRDAKVANG
9603	ETSTHCCKV	9644	KSTYYNLKW	9685	YTKYNWKAA	9726	SVKDPMDWG
9604	QVNMFMMPC	9645	DVGIGVDQG	9686	GTIRMPEHR	9727	TMTVRMEQP
9605	NADMSFRNR	9646	MCDNFPSAH	9687	VCHNPDPFT	9728	ACNTLEHMA
9606	FRLDYDYKN	9647	IKHITRNTY	9688	SLTQYQWRG	9729	KIRDCDGNV
9607	TNAAYEMWF	9648	GNMHNFSAW	9689	TLNSRHPEC	9730	IAYHYDQQN
9608	TTTITNRLC	9649	CSIQINNCR	9690	IMTMQMDTA	9731	QTNEEEDGN
9609	KIQPMMVKA	9650	MREKLVHDC	9691	YMHFRQYPI	9732	NHILTPVVF
9610	LFRAYAWFQ	9651	EQPHISIKG	9692	QTMCITIHP	9733	QSYPTMYQC
9611	DVIRDMIAE	9652	EMSISHGCH	9693	QEAGYTRHN	9734	ECGCMYWGY
9612	ALTHDDGRT	9653	RSDEMYVTP	9694	MNHQCMLWD	9735	EGVNRPNMK
9613	KTYSAMYSP	9654	SFSYDVVMN	9695	PWYNWDRRC	9736	HATGLVNFM
9614	NCMWISSTQ	9655	GDMCEDMCT	9696	DASQIPITF	9737	NMFQLKPQN
9615	TILPKKTLC	9656	TVTQGSIGH	9697	IHRVVCRDH	9738	NDVWQPHPN
9616	VTTHMNCAD	9657	RKDQSDLWP	9698	HHMPIWLGL	9739	TQPEDHWLA
9617	PPRFLQTTG	9658	MAAITQNVA	9699	AEVPYVMQQ	9740	IWITCPYAA
9618	VSVPKRQGP	9659	NSTPMACME	9700	CVLSRSTMR	9741	TCHPNMSVP
9619	VFGSTRSAA	9660	PLSSEWRWA	9701	AKRECTHIP	9742	SNVMATKNW
9620	TSDQCGNED	9661	CHLKICMHS	9702	TISQIPWVR	9743	VDRLFRMEY
9621	GAKYIQNTD	9662	DVWWTLTRD	9703	STRPIACHS	9744	FNIQCLFSR
9622	MRYHWDEAG	9663	DATTSTECH	9704	AWKMAHNSF	9745	PYGPIEYDE
9623	SSEYSQFAI	9664	AMLQKKRQD	9705	RNDNTCHAY	9746	DYFQYNEAP
9624	GVFWPDSFR	9665	TIAMPSWTY	9706	IEHCSNRYA	9747	TTCYFVASQ
9625	GAYSLHDWR	9666	NGYVVDNHS	9707	AVEDYWRYF	9748	AVVCKMNYH
9626	MTALPPSKN	9667	GHHQCYKAD	9708	KHIVGQDGG	9749	QVSFVGCNH
9627	HGFKCQYSF	9668	SANQRYHDW	9709	ERFKTFCEE	9750	ATGTYNLQA
9628	YGGSEHWGP	9669	DCAWMQKEW	9710	LVQKWMRTC	9751	RKWMNAKAT
9629	TTAYHRSDH	9670	KVNCRMIIM	9711	DARDQQMGT	9752	ENHPTWEKS
9630	NTHTGNDRL	9671	KCRYKDCQC	9712	NMPSLRTRI	9753	KAVLFLKDC
9631	YPQPCSCHS	9672	NTSELLMAQ	9713	GSQQTTLSC	9754	ARYMYEKDA
9632	VSGFDRVAF	9673	QGKQLYHMC	9714	GYFYALFFM	9755	SHAQPIASH
9633	SRLYKKRLF	9674	NWYTTWIDE	9715	GFRILDACL	9756	ATGTHNLQE
9634	CCKSKEVCA	9675	QAGNSMWAK	9716	ETDCMEGGH	9757	IPRPFFTMH
9635	TYDTRNHCL	9676	RMYATEHCS	9717	YPDTADIFW	9758	SCQLLNDGG

[2356]

9759	DGEFLHEVL	9800	LAAIQHCAM	9841	KNAVLNASD	9882	EAYITYNKA
9760	LFMTYSARN	9801	RQIGTGNME	9842	DVYMAVVGS	9883	CTHHNLKAA
9761	QVSFNARYA	9802	AQQQILNSN	9843	CIMLNGCFQ	9884	AWSDLMNCW
9762	DILAFDGSC	9803	CASSKSHYL	9844	RHWYIALYA	9885	MTIKMEQQY
9763	NTGVKFCAC	9804	CLYHSMMKQ	9845	CWQLFNKQN	9886	KATHCDKYV
9764	IGQYFNFHF	9805	FKPFAQQWN	9846	KPCNPVGQE	9887	AHEDPNTWC
9765	AEHMVVDHE	9806	KAVSYHHDG	9847	SSVVMAKQP	9888	AWMCCARHD
9766	MQHLPPAMT	9807	ICHTEVMFN	9848	NAKYSMLYG	9889	QWEEVWWEA
9767	RAEQKFVLL	9808	QCQSQEWFH	9849	GMWAQDSSG	9890	NAIPYFVQA
9768	QGKEFICGP	9809	WCTQFDEYT	9850	YPKRCAEGD	9891	ITSHWLSTN
9769	KCNDIVLYH	9810	PACKEWCFC	9851	VKHTAVDVP	9892	KSPCCMSVQ
9770	FQCKMFRIW	9811	IHNRYAQAP	9852	GEGVEQMDG	9893	AGYQCGYNE
9771	STWSANRMT	9812	ASSRFGSYG	9853	MCRPVMMIE	9894	ETNGVEDWY
9772	QCGCAPMDS	9813	EFGEMLGVD	9854	GTNAYMAHK	9895	SMKQMTKSF
9773	CCLWSYWAF	9814	CEMQVQLCV	9855	NIQLLHSYC	9896	YACWLQDAL
9774	QFCICCDGM	9815	DFMWYHICS	9856	MWAETADPA	9897	KWEQRSKMS
9775	CVFSRNNYQ	9816	EFDWLLGTI	9857	QIKTHSNVF	9898	KTQTHTSEP
9776	LTCMHASTP	9817	IATKSNCHK	9858	RLEITRYIW	9899	GNVLLGCQP
9777	GYAQRGFFM	9818	QPFCFSIGA	9859	VTMTTKSVW	9900	HAKQANGMV
9778	HLMPRNQCF	9819	QFHVWFEQK	9860	TMPDRDPTY	9901	KCEGEEVGR
9779	DIATMRNTW	9820	EMSPIDWHY	9861	TVWIQSYSN	9902	HSTNLFEAK
9780	GKVQWSAQE	9821	ARAYYDSYS	9862	CVKPFLTGW	9903	GWADFRELQ
9781	AHQTAWAWI	9822	TVRHQMDAY	9863	LIWSVYQDE	9904	ETEIRQMDG
9782	GCQCTGHPI	9823	VKFRITTEC	9864	MPHHSPDCW	9905	KITVHDRIP
9783	EPLYSWVGI	9824	SDHQYRGCL	9865	HGYQMAYAQ	9906	VCGWWHQYW
9784	EIYSQLTDL	9825	TSSDHYALP	9866	QVYMMEDNM	9907	IADDYFMCH
9785	KMEVNMSEL	9826	GTDLSNHFQ	9867	VTNWLSLQS	9908	GVWMPHGWR
9786	NVTTMFGQK	9827	RAERYFKDE	9868	MVFLEQHGR	9909	MSLQHHFCT
9787	CKVDIHDYM	9828	SKGDLCHQQ	9869	KRTDLHAAD	9910	YNNDMKQRR
9788	ENCMSRMKP	9829	EAQPSEING	9870	QTGTGHWFE	9911	QGVACLDQL
9789	HPLPNTEIC	9830	TVCGSFHGC	9871	TMGLGMNAI	9912	RITQGESGM
9790	KIGVNSHYD	9831	QTPKYFVHQ	9872	VLKSCDRHG	9913	ASGHMTMCM
9791	GGDGRVHNK	9832	KTDKAGLME	9873	TYWCPLMMN	9914	NTNRAHYWH
9792	SIRYDFYVP	9833	VPTGRRVKL	9874	ANGTRECQS	9915	TIKQQMEMM
9793	SNTPMPYHG	9834	GAQSMQFSC	9875	DSQYHVKCG	9916	CQNTAVRAI
9794	HVTMNQLYS	9835	TQEACHHDP	9876	MALPLHVLS	9917	AVAEIRPEP
9795	QHLMTFSHE	9836	MKTIMQDYF	9877	TIWVHEYGM	9918	YSFATGNAA
9796	GREITQVMS	9837	MFSHGHVMF	9878	QTQICHSWM	9919	SYLWRSVYC
9797	GRWMCNEAE	9838	SVSTKPHVE	9879	TEEVQFVVH	9920	AVDHYSNGA
9798	TNMLAMRGD	9839	RAIYNDMTE	9880	DSATFKSAD	9921	ASIATMDRI
9799	MYTGMRCWR	9840	MFTPCQEGL	9881	CGCQVKCHT	9922	NHKMVQDAL

[2357]

9923	QRALPHPIK	9964	MVMMREPSN	10005	YYNHKIVEP	10046	CAEGQRALT
9924	LTHNTPFPK	9965	MDGRDEIIH	10006	AKPTHDVGS	10047	QAIDMEFPL
9925	SIAQDQVWS	9966	VCRLDNGPQ	10007	HTNQQANAP	10048	FNESGWPTW
9926	CCRQIWVQY	9967	MTGIGKCWN	10008	VTKLAENVY	10049	NARVMAKRE
9927	CWGPIMHPF	9968	EIHTGRPGS	10009	DVCDDHKPN	10050	LNDPHCSQN
9928	GNKSPDECF	9969	QIIYRCRST	10010	IAHWKCHYP	10051	YFADAWLKT
9929	GHRPKLCDQ	9970	KCLPYMEMF	10011	YWEKVPCMK	10052	NIHYHMRRN
9930	QVACHMTNP	9971	PNGYRPFCC	10012	NWEFNARHD	10053	QMEVHWHMW
9931	QYVQCAAKD	9972	QGQSGTQYA	10013	VRLQPQHNK	10054	REEWKNASN
9932	QDEPMNLSC	9973	LVSQWDYDM	10014	KCTEYVHTC	10055	ISTKFGKEI
9933	GMRTHSVMN	9974	ESCHLQHGE	10015	VPTPWMGIS	10056	AIHQESMIQ
9934	GTNHRDCLV	9975	TSEPHCSYA	10016	HHFHVNLDI	10057	QEYHIYFPF
9935	LMKYQWIGN	9976	LSVQAQWDS	10017	DSRTQGGLC	10058	TLTHCYSKD
9936	KIDQTQWQR	9977	FVPWCHSFG	10018	VRADNMPVM	10059	GCLTSECAY
9937	TLHYFHGYQ	9978	MKEDNKFAV	10019	LLPIMHTHW	10060	AWFNTYSVH
9938	QSQWWDTCN	9979	KTGWDSHWS	10020	ISLVWMHYP	10061	GYGHTFRDA
9939	WKPEQRHTP	9980	TGMTFNSFR	10021	MWTEFNEHG	10062	FIDEKQCHE
9940	IYCLCQDSD	9981	NGYLVHACQ	10022	MWAEFNEYG	10063	HMSFCNFKV
9941	DMKQKAEYT	9982	QVEWFMAQY	10023	QVHIGSWPP	10064	SVETARCSG
9942	CIPTYRHFG	9983	YTWPYPETH	10024	YEMTSKCTY	10065	SSEHYSASF
9943	VRQFGTWAT	9984	IQCKYIHFT	10025	HMSQAGVVS	10066	LMKTSPSYW
9944	ICIRGARAD	9985	SNVNLMNSR	10026	DQSMMTIGQ	10067	YGNSVTCTS
9945	TSSKHLVFD	9986	ATKFFCSSE	10027	MPSWYIAQA	10068	KHITWNHCL
9946	MIKLINCVN	9987	MREMPMVGA	10028	GAGPVQSEE	10069	DTSKTGKHV
9947	IESNTNKFP	9988	SMQEYQKHI	10029	FGGMTPYGV	10070	LTQYKALAQ
9948	GIKKPNEDN	9989	KSMPPGQGD	10030	VCSNVRSYC	10071	TRIVEPDAN
9949	AMTRCKAGH	9990	ECRMAVQSC	10031	HGNSFLVHW	10072	LADCMLNHH
9950	HGNMVLVIG	9991	VRWGHTCFY	10032	MWTEFNEYW	10073	VPMWAHWQN
9951	LIQSYPHQC	9992	VKRLENEWH	10033	DITSLGCIM	10074	SVTPDSITT
9952	GGALARSSI	9993	TNRCHDDRG	10034	NRPLMSCCF	10075	QRNNLFVHD
9953	QKAIVAQND	9994	TTKNAISFE	10035	EFWWMNYMG	10076	IHGIDCYEI
9954	KTLPFAIYH	9995	EAELENVNM	10036	TSQYSSVST	10077	GIDLMRQIN
9955	KIQPTNHHE	9996	IQNPAMWGT	10037	MGTEWTQSY	10078	MKNQPTGAS
9956	KMGMHDMAM	9997	CDWMANVSV	10038	MWTEFDEYG	10079	QKSMLNSIQ
9957	QENDERIWY	9998	EQTHTWLCR	10039	KQLFMKFWN	10080	ICNMDFRCA
9958	APFFTFTNH	9999	AKYLPELYY	10040	EFKCVSDCL	10081	KTVLKMQDM
9959	YSDMFHKDH	10000	EIHEFRASN	10041	KTIAQTYYG	10082	RLGMNWKQI
9960	RMVLAYGNF	10001	WLIGMWNTG	10042	VQMQYERLF	10083	TPKSEGICC
9961	GQDCWGLAM	10002	WHVQSKRWT	10043	MDVQQIISG	10084	GVITVDAKG
9962	PRWQSEKCQ	10003	TTDQQCTGH	10044	NPIKFQCAH	10085	AIGPDRSNN
9963	NTAKLGWGL	10004	NISPYEMAV	10045	NSVHVHASN	10086	EQEISDQCR

[2358]

10087	MQYRGMSAS	10095	RWGVMDLCN	10103	KSELNRFTY	10111	MMTPMAVEP
10088	MISQMGLKN	10096	NVEDVGMIE	10104	SALYRWGHV	10112	KISTTWADK
10089	AFARQEYAN	10097	QSWWHSDTC	10105	HRAGMSNIP	10113	EVDQYYSLA
10090	TYLHQPEDS	10098	AADGYVAQD	10106	MARQCHSFH	10114	GTMSFPTGC
10091	TKLSMMGCE	10099	ITTPLHHHN	10107	ICKHLNVWP	10115	LMQWYNRCW
10092	EELRYTCHT	10100	NAQRLAVGM	10108	TNGKSSSYS	10116	GTGHIGNAT
10093	GSCCLNNET	10101	HVIGMGAGE	10109	HAVNFDCKK	10117	GGSYVSAPD
10094	KKTLYEMAC	10102	RQYFEDNAG	10110	QISAWNEFF		

[2359] 表75

[2360] 驱动结肠组织嗜性的AAV5 VP1衣壳多肽中581至589区域的序列

[2361]

SEQ ID NO	581-589 序列						
12118	AQWYMWCWD	12141	GKIAHDQLN	12167	QSQWWDTCN	12193	MFMTNVNSF
12119	ASHYHEKVW	12142	MCFPYSCAA	12168	LMKYQWIGN	12194	SSRMQMQGP
12120	IANWRWGPD	12143	LSVQAQWDS	12169	QISKPSVQC	12195	RYPCCQSPF
12121	RVNGSTCRH	12144	MMNQCQVWN	12170	AIEHGAWKC	12196	NARVMAKRE
12122	YNRLMHQLW	12145	TQEACHHDP	12171	SLTQYQWRG	12197	GLSMVLHFV
12123	ETTMMIVSP	12146	DFNGSNWLP	12172	TIVVIPIKS	12198	VSVQSRAQL
12124	STAKTGTTT	12147	HFEVRSVQR	12173	CSMPQQCME	12199	DMTLQSVSS
12125	LPPLKTEDN	12148	MAEQMERDF	12174	QTAEWYCGA	12200	MTDNNGPLF
12126	QTQICHSWM	12149	TGMAMSEQS	12175	GNECFHNVE	12201	PLTIEVNCT
12127	QSLWYWQQE	12150	PCCWEQSTN	12176	VGPLYAYYS	12202	YVDHWTLGD
12128	KMGMHDMAM	12151	NWYTTWIDE	12177	WTLQVMKIN	12203	NQREWHGLA
12129	IRPQYAKMD	12152	CCKSKEVCA	12178	TSEPHCSYA	12204	SECEVLCSL
12130	MNANETHGA	12153	LESVKMDCN	12179	AVHTADVCC	12205	CAWHMKQGN
12131	MSVAAGRGM	12154	PHVSGEGKI	12180	QHSRAYMWY	12206	ATGTYNLQE
12132	QTMCITIHP	12155	TYDTRNHCL	12181	HMEQCWCNE	12207	KCADLIIRS
12133	CMEAQGECY	12156	MAAITQNVA	12182	AYAPSKNAM	12208	QSRCVDNTV
12134	FPQTMHAQA	12157	PDLWYEKSS	12183	DNYWIVPFA	12209	ATMNLSWGL
12135	KPLASIGQY	12158	AHAWLQHWA	12184	QRSVFHQRS	12210	SVLHVQALA
12136	CCCQFSFWN	12159	MAMWIADTH	12185	HEVMMCPAP	12211	KMAGIGCWY
12137	DAELMWQIE	12160	LCAVCGDCD	12186	KTMGIGGGP	12212	YDYWMQAPT
12138	ENTMWQSSQ	12161	RITQGESGM	12187	TLNSRHPEC	12213	HSMMPGWPS
12139	KIDQTQWQR	12162	GNKSPDECF	12188	PLNQEQIPQ	12214	MFWTTPSHW
12140	EQEISDQCR	12163	KNCIMLASN	12189	AQCHDMWGH	12215	TKHCFYGGC
		12164	EKRLEGWVY	12190	KACMSHQGR	12216	AMNWWAAPQ
		12165	LETMMRHGE	12191	TTTNQFMFQ	12217	AMNWWDAPH
		12166	EQPNMGDYG	12192	VLKSNLTGM	12218	AMNWWDAPP

[2362]

12219	AMNWWGAPQ	12260	IGDPKVNEY	12301	PEFPGEVDA	12342	VQCIHPDYR
12220	CDRPDVNWF	12261	IYTYQWDWI	12302	PEFPGEVDS	12343	VQCIHRDYR
12221	CDRPDVNWV	12262	IYTYQWDWT	12303	PEFPGEVDV	12344	VQCIRHDYR
12222	CHSLHQVWA	12263	IYTYQWGWT	12304	PEFPGEVYA	12345	WKPPLSDAP
12223	CHSLHQVWD	12264	KCHYRMFQI	12305	PGFPGEVDA	12346	WPSMQCAAH
12224	CHSLHQVWG	12265	KDTEGLEAP	12306	PKFPGEVDA	12347	WPSMQCAAN
12225	CHSLRQVWD	12266	KDTEGLGAP	12307	PQCLHHHWY	12348	WPSMQCAAR
12226	CKTMKDIER	12267	LASQKRDCD	12308	PQCLHNHWY	12349	LSSQKRDCD
12227	DMQQNDPCC	12268	LEDPKVNEY	12309	PQCLHSHWY	12350	IEDPKVNEY
12228	DVDFRAWAT	12269	LINHWMHDN	12310	PQCLRNHWY	12351	SCSPDDWET
12229	DVHLFPTEQ	12270	LINKWMHDN	12311	PQCLYNHWY	12352	AMNWWDAPQ
12230	ECHYRMFQI	12271	LINQWMHNN	12312	QPVGLFAFA	12353	LINQWMHDN
12231	ECPYRMFQI	12272	LLNQWMHDN	12313	QSKSEGIHI	12354	AAEQRFNPE
12232	EHHWGDIID	12273	LSSQKRDCA	12314	QTDILDYAW	12355	AAILMSNAG
12233	ESAHECEHD	12274	LSSQKRDCG	12315	QTDILDYDW	12356	ACMMMWPYV
12234	ESAHECGHD	12275	LSSQKRDCN	12316	QTDILDYYW	12357	ADTGSTAVI
12235	ESAHEWEHD	12276	LSSQKRDGD	12317	RDRPDVNWV	12358	AESVEDYHT
12236	EYHYRMFQI	12277	LSSQRRDCD	12318	RLEINMLHC	12359	AFELHYHHQ
12237	FAVECVDCT	12278	MCATQHCEA	12319	RLEINMLNC	12360	AFKFKMLGQ
12238	FEVECVDCT	12279	MCEQYWTED	12320	RLEINMRNC	12361	AFKWSHIHL
12239	FEVGCVDCT	12280	MCEQYWTGD	12321	RLGINMLNC	12362	AGVQYTWCG
12240	FLGFKRYGG	12281	MGEQYWTED	12322	SCHQCDLCD	12363	AIDQFSGCK
12241	FLSFERYGG	12282	MGRMELDGV	12323	SCHQCNLCD	12364	AIEYRKYDR
12242	FLSFKRYGC	12283	MHMQQNQMM	12324	SCHQCSLCD	12365	AIFMMSGPG
12243	FLSFKRYGG	12284	MHMRQNQMM	12325	SCIFRHWLT	12366	AKWMITIGC
12244	FLSFKRYGS	12285	MHMWQNQMM	12326	SCSPDAWET	12367	ALDWCTRQA
12245	FLSFQRYGG	12286	MRATQHCEA	12327	SCSPDDWGT	12368	ALICVNVFG
12246	FLSFRRYGG	12287	MSRMELDGV	12328	SCSPDGWET	12369	AMHTQGSLG
12247	FNLTAHEDN	12288	MSRMEQDGV	12329	SCSPGDWET	12370	AMTQEQMAD
12248	GLEINMLNC	12289	MSRMGLDGV	12330	SLQRWFCRT	12371	ANNSCYRPY
12249	HCAKFECMF	12290	MSWMELDGV	12331	SLQRWVCRT	12372	APLPWWVVP
12250	IANTNPRGR	12291	NDDPGQAVG	12332	STIHAIEQG	12373	ARRYNPELQ
12251	IEAPKVNEY	12292	NVHLFPNEQ	12333	SYHQCNLCD	12374	ARSPWMHMS
12252	IEDPEVNEY	12293	NVHLFPTAQ	12334	TMNWWDAPQ	12375	ASDAKPAQL
12253	IEDPKVNAY	12294	NVHLFPTEQ	12335	TSVQSVDRG	12376	ASDDQRDSK
12254	IEDPKVNEC	12295	NVHLFPTGQ	12336	TSVRSVDRD	12377	ASHPCYSMH
12255	IEDPKVNEH	12296	NVHLFPTKQ	12337	TSVRSVDRG	12378	ATDIWCAFC
12256	IEDPKVSEY	12297	PACYSDVLN	12338	VQCIHHDYH	12379	ATNDSPHLA
12257	IEDPRVNEY	12298	PEFLGEVDA	12339	VQCIHHDYR	12380	ATSWNTAIG
12258	IEDPTVNEY	12299	PEFPGAVDA	12340	VQCIHHGYR	12381	ATYVYERFE
12259	IEDRKVNEY	12300	PEFPGEGDA	12341	VQCIHHNYR	12382	AWCITYGLG

[2363]

12383	AYERYDCYD	12424	DGSPKNTSS	12465	EICSHLCWA	12506	GELVRRLSQ
12384	AYTKMAAQG	12425	DHIQMHRFN	12466	EIILGGYHN	12507	GFCYPNKMI
12385	AYVAVASQW	12426	DINANMDKR	12467	EKGWSEPGN	12508	GFTQDPVGH
12386	AYVPEFRDT	12427	DIQSNCSEA	12468	EKQQIEIAS	12509	GGDTGEYLF
12387	CATRKDEMF	12428	DISEQNRMC	12469	EMMSQSACA	12510	GHEPNKHFQ
12388	CEECQEMLT	12429	DISQHPIGS	12470	ENESSHTNP	12511	GHQVTEYMF
12389	CFRICVSEY	12430	DIYLSNGTY	12471	ENKVCVYMM	12512	GHYVIHKTN
12390	CGEDMKIWQ	12431	DKCMLMDAS	12472	EPYQKDCEL	12513	GIITPSMGS
12391	CGGCLNLGT	12432	DNENCVQKT	12473	EQQAPDHVK	12514	GMNPQLECP
12392	CGGTHGMMG	12433	DQNEEGTKP	12474	ERGVMCHFW	12515	GQPINGKQE
12393	CHVDRAMHH	12434	DRIGYCNNW	12475	ESENPMGPY	12516	GSARQDGKP
12394	CHVSVPVDC	12435	DRQGCNAPR	12476	ESHMMRISE	12517	GSCLGISDM
12395	CIANVDNFD	12436	DSMWRTKEL	12477	ETQTHYTSC	12518	GSMQIAQEG
12396	CIEYTHRLA	12437	DSMYASSMP	12478	ETWLMHHHV	12519	GSMQSPTRR
12397	CILNTCSFC	12438	DSRQVKKFN	12479	ETYIGGCCM	12520	GSNSHLPLQ
12398	CIPPLSDKF	12439	DSTWLIAQQ	12480	EVGWIPPYM	12521	GSNSHLPLR
12399	CKAPAQDFR	12440	DTGVYSSWE	12481	EVNYVKCND	12522	GSTSKDSSM
12400	CKAPTRSYR	12441	DTSADKSGW	12482	EVTNRTIPR	12523	GTHPCGCPI
12401	CKSIGCMDE	12442	DTWISHWWC	12483	EWKCSQNLT	12524	GTSFNVHGA
12402	CNDMTAQYV	12443	DVKSENHDW	12484	EWKYSQNLT	12525	GVDHFFWVS
12403	CPDQYRAYP	12444	DVNVAASWV	12485	EYNQCTFGM	12526	GVHSAFYQT
12404	CQGFVKGQS	12445	DVVQEFWAK	12486	EYYQSVKMS	12527	GVIWIHDEN
12405	CRINMERVS	12446	DWHHEHNPN	12487	FAQLHAVGN	12528	GYNHVDSWS
12406	CTKKDNMDN	12447	DWQRVKTMW	12488	FASAMDVTL	12529	HCHLCSLRG
12407	CTNRHTPFI	12448	DWSTIEWTA	12489	FDRHRWQPL	12530	HDERMWAMD
12408	CWATEWQMP	12449	DYWSYPNYF	12490	FEHFDLSAS	12531	HEPRTICTF
12409	CWEQEWPFW	12450	DYYESWWMH	12491	FLGNIKSYE	12532	HEQIGPEYR
12410	CWQNKINHI	12451	EAHQSTLST	12492	FLQNVGFFN	12533	HFGYMHCSQ
12411	CYVMRRIYC	12452	EAKFLGGCW	12493	FMMAFKLDC	12534	HNEHKVGYG
12412	DADQAAKAC	12453	EANHYIEIQ	12494	FMQPIKQMD	12535	HPLKIPIGK
12413	DAGMCEKED	12454	EAPKVTVWG	12495	FNPWDKYGW	12536	HPPSMIRVV
12414	DANIASSGR	12455	EASMKWWSN	12496	FQPHIRDVG	12537	HQNHPPAFD
12415	DCMVFINYL	12456	ECDHHEEHY	12497	FQYRPAEHV	12538	HSFEIHDEG
12416	DDCTNGSME	12457	ECQCCECQD	12498	FRVLQGNSV	12539	HSNPRCVFF
12417	DDVNNVAIA	12458	ECRMRLDNS	12499	FTETTHCGH	12540	HTDQDRHAE
12418	DDWQKLGVP	12459	EDRNCSECL	12500	FTSGHGYME	12541	HTSPCWGSY
12419	DELYTGPCA	12460	EEKPNYHTA	12501	FVHKTEGED	12542	HVEMPAPGC
12420	DFNTKTSSL	12461	EFDYYCLKT	12502	FVHTVTVNP	12543	HWIHMVNQQ
12421	DGEFQAYWE	12462	EFITKNWEH	12503	FYFPIGMRP	12544	ICFWTIAQW
12422	DGGMIRGVS	12463	EHECPNSGM	12504	GDGRAETWQ	12545	ICKPEGGFK
12423	DGNKREWLS	12464	EHETTMAEA	12505	GEDCPSAEE	12546	IDNMPKCRP

[2364]

12547	IDVHNGFDQ	12588	KTFPAADVK	12629	MKIMCENVR	12670	NVTETGPDA
12548	IDVRPHFAD	12589	KVEFTATVM	12630	MNMEKLQID	12671	NYASEEFLT
12549	IEMSEQPNC	12590	KVNIYSDTL	12631	MNRIMMVAD	12672	NYMGIDEQK
12550	IETRMESMG	12591	LAERWEQNN	12632	MNSPRRECI	12673	NYSMPVDWS
12551	IFSHWFAGW	12592	LAHEPYWSP	12633	MQIILEFPR	12674	PCVEPNAAM
12552	IGCLYTSNW	12593	LALCDAVND	12634	MQTAHPVYL	12675	PDARGMHFL
12553	IGDQAMECS	12594	LALHSQFCE	12635	MRNHGDFDF	12676	PEHPRDWLM
12554	IKCEPWVSG	12595	LASPKQMVA	12636	MSILASVRV	12677	PHAGQGNKG
12555	ILAQENRWA	12596	LCAPYNREV	12637	MSIRGNWGY	12678	PHGPITNPE
12556	IMEWKNTMV	12597	LCGCCMMHY	12638	MSKPCRVCE	12679	PIAIKDQGW
12557	IMMHRVMSD	12598	LCGNDKKCD	12639	MTDPGMKRR	12680	PIASMFNGP
12558	IMRDMVLPV	12599	LCHFTKGVY	12640	MWVQRETGG	12681	PKILFMFSF
12559	IMTNYHYLD	12600	LDQGMYFHK	12641	NAAIWEHVH	12682	PLMIARYDP
12560	IMVKRTVNA	12601	LECWNKIAR	12642	NACTKVKMT	12683	PMAPDWRNS
12561	INCKPSVQF	12602	LGYHVKDFM	12643	NAYNFTKDC	12684	PMTLPLECW
12562	INVDKDQFL	12603	LIAFLVDRI	12644	NCENKQNVI	12685	PNEWYLWAC
12563	IQCSTNSWE	12604	LIMPLTTQM	12645	NCPCHTDAY	12686	PPNMTSCDW
12564	IQFVTWRFE	12605	LMKMWTSMG	12646	NCREHPEEG	12687	PQDQSCAGQ
12565	IQGPSWHKN	12606	LPDFWRTRN	12647	NDIECEMAS	12688	PQHWKSVGA
12566	IQQVDMMCP	12607	LPIFVQGDF	12648	NEYEKLFPR	12689	PTAMWLTHM
12567	IRWVEPMCQ	12608	LRAILDCGW	12649	NFMSMYGFE	12690	PVATYNVSR
12568	ISPCPTVVA	12609	LSATMVWAQ	12650	NHFPILIDH	12691	PVEPSCQMK
12569	ISPNMQWYP	12610	LSIQDNRGQ	12651	NHKLHTGPG	12692	PVEVMQWKP
12570	IVFQGDTGT	12611	LSTSSQLTP	12652	NHWMGEKEC	12693	PYEPQTINC
12571	IVTVNCAGE	12612	LSVAVQESN	12653	NIGFAIVDP	12694	PYVDKWANH
12572	IWCNGQMDD	12613	LSWMACAQA	12654	NIGKERMEN	12695	QAAPVHAKG
12573	IYTYFIEWF	12614	LSYVMDTAA	12655	NIGTHMGHT	12696	QAEATHAEF
12574	KAFPIDLIA	12615	LTQFSCVSH	12656	NIYVTRAHC	12697	QATCTRYGH
12575	KDDHCAFWN	12616	LWVWWMKDH	12657	NKSPTWTES	12698	QCGCCKMKW
12576	KEGVDQEFF	12617	LYEEHHACW	12658	NLHTPAGCM	12699	QCIQIDMER
12577	KEKQNAKCL	12618	MAASGMLCY	12659	NLVCRESSC	12700	QDDEYHMTP
12578	KHHPTDADN	12619	MADGRPRMP	12660	NPCIKASYF	12701	QEIAKKWSN
12579	KICMYHLID	12620	MATCYHDHW	12661	NPNGPLHYD	12702	QETGPKSST
12580	KIFWTTEFH	12621	MCAQPCNLP	12662	NQMPKEKWD	12703	QFKDSRGAG
12581	KKTEEHHCQ	12622	MCVHKFLTG	12663	NRTGVTPST	12704	QGDPVKFMS
12582	KMMHLPQEE	12623	MDGPWMGKL	12664	NRVDAMIHY	12705	QGKAKQIQH
12583	KMVQDEMWC	12624	MEHVPTRDQ	12665	NRWMMGMKP	12706	QGTHSQNDD
12584	KQEPMHCSA	12625	MEMPLMDVI	12666	NSTEQEIIQ	12707	QIAQQPCMW
12585	KSDQVWQGE	12626	MGTQYQFTG	12667	NTMTSAPVC	12708	QIDNVNQDP
12586	KSTIFKMRH	12627	MHAETDIQM	12668	NTNMAWKQF	12709	QITKKDNLM
12587	KTECFFSMH	12628	MHFTSWPVD	12669	NVTELHNPK	12710	QKLVHKESF

[2365]

12711	QNEKYGKTL	12752	SHMQDKNTT	12793	TGQNGWSKT	12834	VRAMPFDPD
12712	QQGPLQRIM	12753	SHVEPVMGF	12794	TGVVDTAYS	12835	VRNTEMDWF
12713	QQKRVCDQM	12754	SIRWFYTDC	12795	THSKSDCAY	12836	VSAPFCHIY
12714	QRPWHDLTY	12755	SIVNHNEAR	12796	THTVRDSHI	12837	VSMWYDRDD
12715	QSAQSHSHV	12756	SIWGCNVGG	12797	TIDRCNWHY	12838	VTFMITMFH
12716	QSNHEFRWF	12757	SKHEWDMHV	12798	TILHGMLLE	12839	VVDHFFWVS
12717	QVAVDQSPF	12758	SLDRFNQSG	12799	TITHEQPGP	12840	VVITDGGQV
12718	QVYALSKHW	12759	SLILMHEQT	12800	TIVPYEYNQ	12841	VYMIFSGWQ
12719	QWDRCLVQV	12760	SLNWWFWGG	12801	TMDPCRKSC	12842	WALAPELWY
12720	QWFVQGITL	12761	SMGPLWDPL	12802	TMGVYSFDW	12843	WAMDNMNRC
12721	QYKVKERGM	12762	SMHTVINKG	12803	TMMLVVLRG	12844	WDHRYQTGH
12722	QYNSHFLYG	12763	SNDCRNEFT	12804	TNGIGGKKE	12845	WDQCTMINQ
12723	RCDQHDDQW	12764	SNSCEYQIM	12805	TQLHWNDGR	12846	WDSPQTNFT
12724	RCTDCSYFY	12765	SPCFYPDCS	12806	TQQALPLAC	12847	WFHTRHPQE
12725	RDQQIHHED	12766	SPICLQHDR	12807	TSAPDRLSS	12848	WGPCKHPWQ
12726	REVKPMHVH	12767	SQKNSRHLS	12808	TTEQIRDID	12849	WGVVKMVLC
12727	RNDVCCEMQ	12768	SSALITRKE	12809	TTVSCVCKG	12850	WHDKLRNNK
12728	RNLYENQGC	12769	SSSFWENKP	12810	TVLRRCNYE	12851	WIDGIVDIC
12729	RPCSKSHEF	12770	SSSSHLAAD	12811	TVYAWLLSM	12852	WINGPASRY
12730	RPESEHPEG	12771	SSVCKTACW	12812	TWFVGQSTR	12853	WIQLNHQEV
12731	RPREGEWPG	12772	STARDLYLM	12813	TWQRTENNK	12854	WITAFDSIG
12732	RSLHYGNSV	12773	STSHPDNSI	12814	TYLMNQFHM	12855	WITVKEFST
12733	RSPPCDSFP	12774	SVFDVWCDQ	12815	VCKMPMYQH	12856	WLISNSPQW
12734	RTNDHPDKH	12775	SVMQGSSIQ	12816	VCRCTKYTA	12857	WRCPPEQWW
12735	RVDCETPYH	12776	SVWSTRMQK	12817	VDPNLQYPW	12858	WVSADTYEH
12736	RVGNERLHH	12777	SWEILHDCH	12818	VDWFTAPAP	12859	YASLWESGW
12737	RWYHYLLRN	12778	SWHWMTTQN	12819	VFCNYQKKA	12860	YCMTVICSE
12738	RYCLLDEIE	12779	SWYGNEEDQ	12820	VGSKWHVLG	12861	YDECQKIFV
12739	SCKHDDMGK	12780	SYHDHEQGI	12821	VGYSTVGEW	12862	YDGCMWMPF
12740	SCKWIQSNR	12781	SYQLPILHS	12822	VHHDLGESK	12863	YDQKGGQCW
12741	SCLDHQDID	12782	SYSYVNDEA	12823	VHHHGYPTS	12864	YEDIIWNMH
12742	SCNCKKNKI	12783	TCKGMDNAI	12824	VHKEWHWQC	12865	YENVGSKAW
12743	SDKGRQHYN	12784	TCTEPIVEC	12825	VHMYPMTQV	12866	YGHRRDEAR
12744	SDNNAPEFQ	12785	TCVQMEHVK	12826	VKPNETDLH	12867	YHSHLDAGQ
12745	SEGVYNQVY	12786	TCYWDNCWY	12827	VLGFDHDPQ	12868	YKLNCFARP
12746	SFHVAQGDY	12787	TDRWTRVMS	12828	VLLMDGRTF	12869	YLDCPHTQF
12747	SFIAPCDLS	12788	TEDQYNNIH	12829	VLRYVNSAG	12870	YPMAKNRGD
12748	SGCTWDMSP	12789	TERWGWLPH	12830	VMELMMHNH	12871	YPRHWTLQH
12749	SGPDTVFFT	12790	TFWCAQLGV	12831	VPLVNDGRC	12872	YQHVWSQER
12750	SHASGIECT	12791	TFWIRMAAG	12832	VPTICMATY	12873	YQNPDAPAL
12751	SHESTNAGA	12792	TGLYYEMQE	12833	VQNEADMNP	12874	YQQQLMMNE

[2366]

12875	YSDAVYCSM	12916	DVETNNWRC	12957	GQNPKMSAE	12998	LVVEKPSES
12876	YTDALGMEP	12917	DVHDSQMCY	12958	GSCCLHIHH	12999	MACKKAQDV
12877	YTPLPAVNG	12918	DYNDFNWYF	12959	GTQHEKMYQ	13000	MAKRVELVT
12878	YWLCYPKLC	12919	ELGTYDGMG	12960	GTTQCTVHT	13001	MCHSFPTRP
12879	EKLTADAAI	12920	EMQMEERLP	12961	GYDSSEAFE	13002	MEVEAEKPN
12880	INSESEWNP	12921	EPGRMHHSW	12962	GYVTEGFDE	13003	MMKYYDAIG
12881	QTTEFALVQ	12922	EQCLAAAAE	12963	HCPCKCTWL	13004	MMMPQMWSG
12882	SGLYSQCGG	12923	ESRPMIQPF	12964	HERHATLMR	13005	MNRNVEMDH
12883	SDVHMFCML	12924	EVCGHQKIQ	12965	HLLGNLEYM	13006	MTDKLYQED
12884	QSNLMEFHS	12925	EVHVTACGN	12966	HNVCTMWAQ	13007	MTNWLMRGR
12885	EAQNIHDPH	12926	EVVRIENPM	12967	HQIQFCPCP	13008	MVNNWDWFY
12886	HTTLNYMSN	12927	EWFCYCGTC	12968	HTQEELITP	13009	NCDFMVASI
12887	AATPQKNQE	12928	EYHKTQWGH	12969	HTYCYWQSC	13010	NCQPYGEPT
12888	AFCRGHYQA	12929	FALIYHKDE	12970	IAAKVQNHD	13011	NCSHYTQHY
12889	AGAPPLAVH	12930	FASPVLWGQ	12971	ICAAPNSAK	13012	NGLLDFDPH
12890	AGWQMCFDS	12931	FEGFIERQH	12972	IDHDNWDML	13013	NGLVDFDPH
12891	AGWQMCYGS	12932	FGHRYDSAY	12973	IIGQTSIQD	13014	NIAPGGTIF
12892	AHPILQGWD	12933	FGNPMQTCP	12974	IKEHIPYWD	13015	NLEAMDCNY
12893	AHYQQSTDR	12934	FGVNDMDSA	12975	IMNKAHYRF	13016	NMCTTFRLL
12894	AIGHGAWKC	12935	FHCICWTHM	12976	IQAGHQATM	13017	NMKTKSAHK
12895	AMCMLSHIH	12936	FKYDCYQRN	12977	IVDMNTATK	13018	NNHRLNDAN
12896	ANELSWHPR	12937	FLGCYCPYA	12978	KCFTHAIVG	13019	NTMQSPAWV
12897	AQGCKFYNG	12938	FLRNTWLGP	12979	KDWIPSYAL	13020	NVTPTDREG
12898	ATVTRARSM	12939	FMPYCVHSG	12980	KEGCHVRMG	13021	PCGRTCFNG
12899	CACHMKQGN	12940	FPSWMMKAP	12981	KMQGGNGYG	13022	PMMSFIFHH
12900	CDSLSLLQN	12941	FSTAHMFGW	12982	KNAYNMTIQ	13023	PQTAAQKSD
12901	CFCQCGMPS	12942	FWCRPRLHD	12983	KPEILHFHD	13024	PTEVTVSDM
12902	CHHSGESHQ	12943	FYHHLPEEP	12984	KTAENVWAM	13025	PTQWAMGDY
12903	CKGRANWYM	12944	GAGRMNCNV	12985	KTESIVKLQ	13026	QAAYCRCTF
12904	CKWLAIPCR	12945	GANMMHMMN	12986	KTIQHGTAD	13027	QATHDLMHQ
12905	CQVGLQNCW	12946	GCLISECAY	12987	KYSDFCCPV	13028	QATLHHCTN
12906	CTKMHTMIE	12947	GDKMGIYHC	12988	KYTEALSCG	13029	QCLLWNDCQ
12907	CWALNGQLG	12948	GEMPYCNRQ	12989	LAMCALDNF	13030	QGQEKEWYY
12908	DALPERRYF	12949	GFKVEQPYE	12990	LAMRALDNF	13031	QHMELQEHD
12909	DAMQKISCV	12950	GHTNCFAMS	12991	LCSTEGMHT	13032	QPCIPVQIA
12910	DFIYHHQAP	12951	GIDEAKADY	12992	LCSTKGMHT	13033	QQLYFTTDK
12911	DGVVTVTNL	12952	GKYWLQRSV	12993	LDTKRTFKK	13034	QVIPVTYRK
12912	DHDLATEWE	12953	GLFHPQFVQ	12994	LGEATWQAV	13035	QVKMGVIDQ
12913	DKHHSRIDG	12954	GNISTFCWQ	12995	LMMKCCAMH	13036	QVLATMNNK
12914	DMNMYLGSQ	12955	GQHHEETIY	12996	LPDRDWHKW	13037	QWRQMDRMQ
12915	DSKITQFNH	12956	GQILFRSHT	12997	LQTKKDCPA	13038	RDDQPPMVW

[2367]

13039	RDFVAEQQP	13059	STHPLNLYH	13079	YHDLFFAWY	13099	IQKDEMPVC
13040	REERYFKDE	13060	SVQPFLHDG	13080	YHTNNMEPW	13100	GTQFWMSVE
13041	RIEMHQSVT	13061	TADRQLVHS	13081	YMHVPPSIH	13101	QMSWGKYHQ
13042	RKAEQEDVY	13062	TCGKMHGPN	13082	YMWIYPPDQ	13102	KDYWAGYQM
13043	RLVTYPHTA	13063	TLWDNTNTP	13083	YNVCTMWAQ	13103	ICDLYSQER
13044	RNATKIKLT	13064	TPSWFNVGA	13084	YTQLYCSHL	13104	EEAPGFICF
13045	RNCLTDAWD	13065	TRGHHCAED	13085	YVPCERCEW	13105	DLPNDMGAV
13046	RSDQECSKA	13066	TVTCPEVRF	13086	YVQSYWNES	13106	GMHATELWL
13047	RTCKTIGYY	13067	VNQLHIHVQ	13087	YWERHCMTI	13107	LMHPQMTMC
13048	RTEPTFVDP	13068	VPWYEGPYE	13088	YYGWLLVMV	13108	AAKKGRDLH
13049	RTYQYEYEV	13069	VRFEMMSCT	13089	AGWQMCYDS	13109	RHIYRKGWR
13050	SCWRETMKE	13070	VTKCGWMLQ	13090	VNCNKRHNW	13110	GEFADNPWN
13051	SLENEARRC	13071	WDNSKGYNV	13091	PYCSIKQAS	13111	ERQMKLRCA
13052	SLHIHGNGP	13072	WNNEDLKVW	13092	CQTHWEKGD	13112	TDSSDQLQI
13053	SMGMCELPI	13073	YATQMRHLE	13093	EMISNNCHP	13113	YNMVHGDEP
13054	SNAHMQVGC	13074	YCLMDKEDT	13094	RQRYMGMQF	13114	NIKMLVHCG
13055	SNAYHAWKN	13075	YDGKTFLHW	13095	SYLFSKHSE	13115	MKGYFVHHS
13056	SQDLFCNAA	13076	YFDCSRAHP	13096	KAISLHAHG	13116	MLRTQIHCA
13057	SSGWDHGNL	13077	YGTKCERCY	13097	GIGEMQCTE	13117	CKDNTFVSY
13058	SSTWQNVDK	13078	YGTKSERCY	13098	ILERYQHMD		

[2368] 表76

[2369] 驱动心脏组织嗜性的AAV5 VP1衣壳多肽中581至589区域的序列

[2370]

SEQ ID NO	581-589 序列	15132	EQTMFNSPP	15148	IDPRRNNDL	15164	PFTQHVAYI
15118	ADIDQIMCQ	15133	FAMVVTTEY	15149	KINPSMPMD	15165	PFTQHVAYS
15119	ADIDQIVCQ	15134	FFPHDGTIV	15150	KPDTSVGIG	15166	PQQFAWNND
15120	ADIDQVMCQ	15135	FYQMFHASG	15151	KPDTSVGIV	15167	PTAIQAYMR
15121	AHKYDWEYL	15136	GFPPRWKIP	15152	KWPCHWHHE	15168	QHVSTEPPY
15122	ARSDWRCMC	15137	GHCCSNGPT	15153	KWPCHWHHG	15169	QIADGWGWE
15123	AWCEGNCLE	15138	GKQQCPAQT	15154	LCINKYFNY	15170	QMCATMNHE
15124	CCCIMSVMC	15139	GKQQCPTQT	15155	LEPKMKLAD	15171	QMQHCMDYL
15125	CCCIMSVMG	15140	GNRTEGYFT	15156	LEPKMKLAE	15172	QVGWKECVM
15126	CSRIMALSN	15141	HDCALYGRF	15157	LVCEVACKD	15173	RCQQFHGQI
15127	DNRTEGYFT	15142	HLNPMTGSN	15158	MFTQNAGKI	15174	RDINIDPCC
15128	ECHLLSDYD	15143	HNVMCDACC	15159	MSCPKFHIV	15175	SAMVVTTEY
15129	EDWHGWNEH	15144	ICKYPMIPH	15160	NEDSLWNMH	15176	SFTQHVAYS
15130	ELYMQMHWP	15145	ICKYPMISH	15161	NEDSPWNMH	15177	SKQQCPAQT
15131	ENRLRLAHE	15146	ICMYPMIPH	15162	NEDSPWSMH	15178	SNRTEGYFT
		15147	IDEFWYGPF	15163	NNRCKDQTE	15179	TDIDQIMCQ

[2371]

15180	TDRHDGVFA	15221	AKVEMEEWN	15262	AWMPFIDSG	15303	CNPAKSKDG
15181	TEYQNARTM	15222	ALAWQMRAQ	15263	AWPDIMWAS	15304	CPCKGRRLG
15182	TITCEADNW	15223	AMNDELKMQ	15264	AWRGSKMSN	15305	CPCLYCNLA
15183	VFRNRWLEW	15224	AMVIIYTDF	15265	AWSVQGPYG	15306	CPEFGFICE
15184	VIEVEQACT	15225	AMVLMFLYQ	15266	AWVDQEIPA	15307	CQIPKKIGM
15185	VIKVEQACT	15226	ANNQHRARA	15267	AWVGAEHAG	15308	CQTAFYVGH
15186	VPNPKNFGW	15227	APFMSEDCS	15268	AYEHITDGR	15309	CRDRDMAHQ
15187	VSVVQNEAL	15228	APLMTAFDP	15269	AYTDTNAGY	15310	CRLKKYYLK
15188	YCSIIGRSC	15229	APMSQWHTV	15270	AYVEKMRFD	15311	CSPRKSHFC
15189	YFALCPPYT	15230	APNGYTMFY	15271	AYYHEMAVP	15312	CSRFMDSQW
15190	YHAFYWQPY	15231	APSAHHTHL	15272	CAAMQQGRD	15313	CSVDKEDMD
15191	AADHWVCQF	15232	APYWDCFDR	15273	CAEAIIKNF	15314	CTKLSGLNA
15192	AAFQCAKQG	15233	AQGPMSVQL	15274	CAHSNKGGA	15315	CTKWQRTLG
15193	AAGMHWNAA	15234	ARPIFPWIV	15275	CAMQPCLAV	15316	CTLEFPSFL
15194	AAHMGHGWP	15235	ASTVVYAEM	15276	CAVDKCMPH	15317	CVGIMTNGK
15195	AANQEQECL	15236	ASWLADNLC	15277	CCSSFRDMV	15318	CVGWAMGLG
15196	AAPCPDRFV	15237	ATAFRHRIN	15278	CCWLANGGG	15319	CVHKQGENE
15197	AASDSNRIH	15238	ATARLNWDA	15279	CDMLTRMQY	15320	CVIQSNTHV
15198	ACCDKTRPE	15239	ATDQYWFPT	15280	CDPWSWTQI	15321	CVQPTYADH
15199	ACMFWHRCG	15240	ATEIKGSHS	15281	CDSTRFCCV	15322	CVTLEMQQG
15200	ADGRGMDCP	15241	ATESKWDQD	15282	CEVARWLDY	15323	CWAMDWGDT
15201	AEENWRQCA	15242	ATHIHYHGL	15283	CEWKHHAKF	15324	CWQFTRCAS
15202	AEHDLKIVS	15243	ATIEGDGVL	15284	CFHDPYEKQ	15325	CYYHTCLGG
15203	AEVEQYAEA	15244	ATIRVITIG	15285	CGCYSPFLE	15326	DACFYGNFE
15204	AFDPTFIDT	15245	ATNYVVRHT	15286	CGSDCNPAY	15327	DAENCMLGW
15205	AFEAFDYDH	15246	ATSVMAWWD	15287	CGYQLFECF	15328	DAGYCQSGV
15206	AFFLTPAEP	15247	ATWRICSSG	15288	CHNEDLCVW	15329	DALSFQSSA
15207	AFGNLEVHM	15248	AVDFCQTYD	15289	CHTMQDIKD	15330	DANRKDLRG
15208	AGAPLIHQQ	15249	AVHAAYCYW	15290	CIGCPNEWC	15331	DAQAGWRFN
15209	AGAQQYWAD	15250	AVHAGPQAI	15291	CIQHIAESE	15332	DASSGFTWT
15210	AGECQLLMT	15251	AVKPREAWM	15292	CISEMDDQR	15333	DATDWKCAW
15211	AHNCNGHVP	15252	AVMYNDTYD	15293	CISFSNKWV	15334	DAVKAEVMW
15212	AHNWPAMYT	15253	AVPCQLSGA	15294	CKAAGPWTD	15335	DAWAKGYEW
15213	AHNYCGYYV	15254	AVQERFLQN	15295	CKDSFASVW	15336	DAWSANEAW
15214	AHQPDQYCK	15255	AVQSQTKNM	15296	CKPRLCEVP	15337	DAYWMQWWL
15215	AIAHCVQIE	15256	AVQWVEHEK	15297	CKYMIHPDC	15338	DCASFPAVD
15216	AIEGLTGRD	15257	AVRCDPYRF	15298	CLYPLQLEW	15339	DCQGWTAWD
15217	AIEYMNMGF	15258	AVSTTVDWY	15299	CMDQYRIQF	15340	DCVKAIQGW
15218	AIKDMRLAL	15259	AVTTPDVYV	15300	CNACQAIEC	15341	DCVQMNVNV
15219	AISSHWQWP	15260	AVVAMNHHT	15301	CNKAVQKTS	15342	DDCWQDRHV
15220	AKNEECDCN	15261	AWGEMNFSV	15302	CNMPQITGA	15343	DDVWKELDV

[2372]

15344	DEAYRMPLQ	15385	DPHQGIAGE	15426	EAEEYEPWK	15467	EKVMQMHMQ
15345	DEGAAAKAI	15386	DPPIVQQWH	15427	EAFQEDAHP	15468	ELWKAMPPE
15346	DEGTSQTKT	15387	DQKWDTLKP	15428	EANSEYNVY	15469	EMEINPTVE
15347	DEHDGCKKS	15388	DQRRKTNIR	15429	EATKSVETR	15470	EMHMTNTGW
15348	DEMYVSSQF	15389	DQTTLRLMG	15430	EAVTARVQH	15471	EMTKYLNHA
15349	DFAQRNSAC	15390	DQYWAWADS	15431	ECDDYDCHY	15472	ENHCWSAKE
15350	DFGMKNYLA	15391	DRLEDKESQ	15432	ECTRSGKWG	15473	ENNAERWMF
15351	DFKFEMTDR	15392	DRMQFFVRN	15433	EDHVMTKKD	15474	ENVQCELMQ
15352	DFLYSPIGI	15393	DRMTCDTDR	15434	EDIEKNTFE	15475	ENWTWKPEE
15353	DFSTPYQFG	15394	DSEWTIDEH	15435	EDKQAQTCQ	15476	EPMYFIVCS
15354	DGCTSKDNQ	15395	DSGEVRTAS	15436	EDRSSCVWM	15477	EPVNFNIQV
15355	DGIHSLTGW	15396	DSKVFEWKC	15437	EEEPLSEMA	15478	EQDGHNINE
15356	DGNQLCRLT	15397	DSQLTNACG	15438	EEKMSVDRK	15479	EQEQFHICG
15357	DGNYIRVDV	15398	DTELVHNGW	15439	EETVSKRKK	15480	EQMLWNQIS
15358	DGPCAQYRG	15399	DTLSSDAFL	15440	EFIFHIHDQ	15481	EQQWNSVNE
15359	DGPMFATDI	15400	DTMCYAWCK	15441	EFMQSFFIE	15482	ERIMHLHES
15360	DHHGAGYNR	15401	DTQPWVPCG	15442	EGEDYMDIE	15483	ERMPLTCMS
15361	DHISDYQHY	15402	DTRESGFWV	15443	EGNQFWMFS	15484	ERSEHPCMF
15362	DIILRIRLK	15403	DTRIKMMLP	15444	EGQTSNIWE	15485	ESCWPQVTG
15363	DIKAILSKD	15404	DVAKKNMDG	15445	EGVTTLQSS	15486	ESEPIVYGW
15364	DILGVHLGS	15405	DVASFGHEI	15446	EHDTMEGFG	15487	ESHNTTEGH
15365	DIQWQHEWM	15406	DVCMLVDLP	15447	EHGQEEMIM	15488	ESLENPNDR
15366	DISCNLQFL	15407	DVCMSNELT	15448	EHTAMWMHG	15489	ETDNFSDCA
15367	DKMLKTDLM	15408	DVDCCEFCT	15449	EICNKGQPD	15490	ETFVTCTGA
15368	DKNSCCLNP	15409	DVDFHWQGM	15450	EIEGVKANV	15491	ETHHGQMFC
15369	DKPVGRYGH	15410	DVDHKWQYD	15451	EIGGNPIYD	15492	ETHMYERRE
15370	DKQFVEDNQ	15411	DVDPMPCFY	15452	EIGYCMTCD	15493	ETHVVQNDN
15371	DKRFQKSYS	15412	DVGIMDYWS	15453	EIKLPMEQD	15494	ETILAPSGT
15372	DLMMHYEHK	15413	DVLQHQGAL	15454	EIKPMVASH	15495	ETLMTVKKC
15373	DLQMSVNTG	15414	DVMMCFHWR	15455	EILNHTRLW	15496	ETVRQGNPE
15374	DLSQMNKNY	15415	DVPDKPWYH	15456	EIMDTKQCS	15497	EVARKVQNE
15375	DMGEVQVGI	15416	DVQLDQEWD	15457	EIMTMEFQD	15498	EVAWMGRQN
15376	DMNPDACVR	15417	DVQRKTSGE	15458	EINTPFFMS	15499	EVDPWGHME
15377	DMQVLNTDD	15418	DVRFISEQG	15459	EINVTQTHQ	15500	EVHVNWKDE
15378	DMSEVDKVF	15419	DVRPPAVNE	15460	EIQQMDDAP	15501	EVIDHGHME
15379	DMTIAAWAC	15420	DWPCDDFRK	15461	EIVATYGQN	15502	EVLEWRHYQ
15380	DMTKVTPFL	15421	DYLSFCWKD	15462	EIVELATQY	15503	EVNISGEYV
15381	DNACVHCAE	15422	DYNKDPKLV	15463	EKHAFSYSF	15504	EVTEPVKAI
15382	DNCWLNDSA	15423	EACMHRTVL	15464	EKHMTMAGI	15505	EVTKHLVAC
15383	DNEPNNRKS	15424	EADFVGPNS	15465	EKMAREWQT	15506	EVWTPCFPI
15384	DNISWSAWD	15425	EAEEHCTFW	15466	EKMWPLNWG	15507	EWCNEWCGP

[2373]

15508	EWDEWLISF	15549	FVSSYYAIH	15590	GSSHFNDLG	15631	HVMPWLCQW
15509	EWIHLYCCF	15550	FVSTESVWN	15591	GTCHGPVLM	15632	HVMTLVDYD
15510	EWIPYAVGN	15551	FYIMMEKML	15592	GTCQVQLHS	15633	HVNTEQKYY
15511	EYDYIFNHH	15552	GACCKQIGE	15593	GTEIKTLHG	15634	HWNCNCRNA
15512	EYEADTITC	15553	GAKWQAQPD	15594	GTEVIGLNT	15635	HWPQRLNEH
15513	EYMGFRCDR	15554	GATQHECWC	15595	GVFLGYAQR	15636	HWSWFASFV
15514	EYMSIMMNH	15555	GAYDTVWGW	15596	GVISDGKYA	15637	HWTNQDSNK
15515	EYNPTNHIL	15556	GCATWHNLY	15597	GVKDFDVCT	15638	HYHQYTRLP
15516	EYPAEFAQC	15557	GCMLRSVTG	15598	GVMVNPCNM	15639	IADVIHSKW
15517	EYQQLARHW	15558	GCNLLSVSN	15599	GVSAEGKAQ	15640	IAETLIGVL
15518	FACMPNEFL	15559	GDAFSCVFP	15600	GVSGMQITG	15641	IAHGASAYL
15519	FAFDTQMVF	15560	GDDPMFFEK	15601	GVTMNTECE	15642	IAYSLLANY
15520	FAPPTHTLE	15561	GEHTGMMLY	15602	GVVATDCEI	15643	ICMFEEEQW
15521	FCHNIFTEC	15562	GFMAVVQSW	15603	GVYMESAVR	15644	ICPPFQRQV
15522	FDHSERQYH	15563	GGRCYFWED	15604	GWETPLVWF	15645	ICPYAEVSD
15523	FEKTRIAAW	15564	GGRLYKPEW	15605	GWWNHYLFW	15646	ICSNPWNKE
15524	FEMGQKSLS	15565	GHASWAMWY	15606	GYIYHKMYT	15647	IDCQGKCDW
15525	FFEFRPSTK	15566	GHGHKWIYS	15607	HAAETEIHQ	15648	IEDDPLHAE
15526	FGMVMPMWA	15567	GHMYLLDME	15608	HCDTHVNLE	15649	IEHRFMIAG
15527	FGPDSTHLC	15568	GHTCPCKWD	15609	HCKSGQQML	15650	IEPIHNWKK
15528	FGQLLMTDG	15569	GIDTDDVGH	15610	HCTCYDAIE	15651	IESVTHLCD
15529	FGWYYFWDP	15570	GIGMFCVTN	15611	HDKWKWSDR	15652	IGELDEACD
15530	FHSCSRMGL	15571	GKARRCATW	15612	HENYYREGA	15653	IGFMDNTSC
15531	FIDTACSHG	15572	GKDVKQCCD	15613	HFLQMEHGH	15654	IGIPIHMWD
15532	FIGDTYAGG	15573	GKLGLDCQD	15614	HFMYDMNFQ	15655	IGQSIVTMQ
15533	FIGMQSALE	15574	GKVVDAHMG	15615	HILKHVHTP	15656	IHMFRMWPE
15534	FITMRQEGP	15575	GLQNKGYPT	15616	HISFHTIPG	15657	IHPQIAEGM
15535	FIVILNNSD	15576	GLSIEISTP	15617	HITGEYIAP	15658	IIDVFPRHT
15536	FNPDTDQDH	15577	GLSLHKSDH	15618	HKESSEGAT	15659	IIFFLCRDP
15537	FPFGTMHGS	15578	GMADMVHGP	15619	HLNNCHWRR	15660	IKQPANDTD
15538	FPFLTKSPK	15579	GMASHSHYI	15620	HMGQMYWHM	15661	IKYGMRHTN
15539	FPQLPQNEA	15580	GMQTIEQFN	15621	HMVDINILA	15662	ILVMVWLGQ
15540	FSMDKQQFQ	15581	GNEFIECRQ	15622	HNGVMMHSH	15663	IMRWHYGMV
15541	FTCEYTYAY	15582	GNHWGWSIS	15623	HNKACITAL	15664	INASSCDYW
15542	FTELSMVQD	15583	GPEMTVAQT	15624	HPNNFNRAH	15665	INRNYWDFI
15543	FTKQADELP	15584	GQTPCSSWG	15625	HPVHYNMCS	15666	INSELPVDV
15544	FTNFEYSME	15585	GQVDNMMAT	15626	HSCTMMGGD	15667	INYSSVATV
15545	FTQGGEVYA	15586	GRELEIEPQ	15627	HSMGIFSNL	15668	IPQPTMMQD
15546	FTSWMDVQD	15587	GSDRATGKE	15628	HSMVQWNST	15669	IQEYIQDAD
15547	FTTQPSVAW	15588	GSEFCILNE	15629	HVEMANAGK	15670	IRESWIGQF
15548	FVPQVWQIP	15589	GSKPSSKQS	15630	HVIPHNRSE	15671	IRYVTAYYK

[2374]

15672	ISEIMPSNP	15713	KNDMLKKFG	15754	LNTDVAEAF	15795	MGRWENIWC
15673	ISMYQSTQV	15714	KPIPQTVWL	15755	LNVGKKGYG	15796	MHDMLMGQM
15674	ISNWRCSQK	15715	KPQAHYYTI	15756	LQQCEPMKS	15797	MHFQAWSGA
15675	ITCSNEKDF	15716	KQLLTRMYM	15757	LQSEHNEDE	15798	MHIGVFTQA
15676	ITENRGYYH	15717	KQPEPTPLA	15758	LSICDQYDV	15799	MIAFWSTWR
15677	ITSNELPME	15718	KRMQEMTCR	15759	LSMRCMDQW	15800	MIAWWGSPG
15678	IVEFMHCDD	15719	KSASPTQPY	15760	LSNFSWQYG	15801	MIKPLTCNK
15679	IVEMYAKNN	15720	KSPAIHTEE	15761	LSVKHVHYS	15802	MIKQTGAEW
15680	IVHDEPRHN	15721	KSSGVWMCC	15762	LTCMWGVYP	15803	MINNSKTPY
15681	IVHDIWGQA	15722	KTEWYDMKN	15763	LTTLVEAIP	15804	MIRTQCCWI
15682	IVKYRAYPG	15723	KVEENLRSM	15764	LTTQWNAKG	15805	MISQMMSNA
15683	IVNKNCSQP	15724	KVIPEDQNC	15765	LTVFSNGGV	15806	MIYWTVDQQ
15684	IVPFGCGGF	15725	KWDSTCYRW	15766	LTWVFGPFS	15807	MKASKDGGK
15685	IWKCLNTQN	15726	KWTEFDRCD	15767	LVDTWNNHR	15808	MKEEWGHIP
15686	IWTCCKGGG	15727	KYEEQYHFH	15768	LWCKTWHNY	15809	MLSWVYPCM
15687	IWVDYEITR	15728	KYQPDGWCY	15769	LWDGEAWCW	15810	MMAMKTDSH
15688	IYEARTIVR	15729	LAAGVWDQE	15770	LWNMIKPLY	15811	MMCQIMHSC
15689	KALPGMSAG	15730	LACKREAMW	15771	LYELQQFGA	15812	MMNGMMGGN
15690	KAPGVCTRS	15731	LAVCKTHEP	15772	LYHWWGYIE	15813	MMQEIGAEA
15691	KCDMKRGQF	15732	LCHPGEHWQ	15773	LYQSTFDWR	15814	MMRGPESDR
15692	KCELMSNQT	15733	LCVTGFDTH	15774	MAADRWPLG	15815	MMVQQMFFL
15693	KCWYKFITN	15734	LDESWSVAG	15775	MAGQSAAYW	15816	MNESTHIHP
15694	KDALFDFPC	15735	LDPNRCMDQ	15776	MAHIEGVAN	15817	MPCCWWVVT
15695	KDDGMCYGR	15736	LGFWETQIM	15777	MASDTDYSF	15818	MPCPLPKYR
15696	KDEDWNWCD	15737	LGHMQASTP	15778	MAVMIANAA	15819	MPGYQIHNA
15697	KDFWAMSMV	15738	LGNQYWWHF	15779	MCGLTKFRP	15820	MPKDWVPDA
15698	KDIGKAAGY	15739	LHEIHLETN	15780	MCPPYQQAE	15821	MPLTHSMHN
15699	KDVMFREMH	15740	LHLPQLLHD	15781	MCSSQHSYS	15822	MPNAFYQAC
15700	KECEQLMFP	15741	LHSAWMHLA	15782	MDDDNDPVF	15823	MPSQNGTKH
15701	KFDHDSEFY	15742	LIAHWIHSF	15783	MDKDSEIMG	15824	MQLGEHSWY
15702	KFEPRYGDC	15743	LIAPYFGQL	15784	MEEYNTTTC	15825	MQLLGDSEY
15703	KGIHEPISC	15744	LICTMSDMN	15785	MEGGHPYRT	15826	MQSQEDHSP
15704	KHDSDWMWY	15745	LIDHIQNQH	15786	MESKYWHSD	15827	MRVWRYCDR
15705	KHIQDARWF	15746	LKEMTWINW	15787	MFHDTLGMF	15828	MSSCKKMDA
15706	KHWREMITV	15747	LKQFRHMDG	15788	MFMRLGTLP	15829	MTGEKTGNA
15707	KIDDDSRVY	15748	LLYMAHCPP	15789	MFNMYEGSQ	15830	MTLSWWGIG
15708	KIQYDDFDG	15749	LMDGCTPFM	15790	MFNWAEMED	15831	MTMLKMAGR
15709	KIVDIDWWA	15750	LMEVYRSKA	15791	MFTDNTFKC	15832	MTNGNTVGA
15710	KLMWFASTD	15751	LMHYGMVWP	15792	MGIQTNYEA	15833	MTNIMQQGK
15711	KLSNDRFIY	15752	LMSKENNCN	15793	MGNMPASMA	15834	MVDGDLPMK
15712	KMQTVMRAL	15753	LNGECVVGY	15794	MGRTFFEEF	15835	MVDKGFAGQ

[2375]

15836	MVESELMMC	15877	NMVEYKQTA	15918	PTDFIWPMM	15959	QSEFLVKHI
15837	MVGCCCHFG	15878	NNDCENPPL	15919	PTDQDQMLE	15960	QSMDCNWDG
15838	MVGPAQIEL	15879	NNEFDVQHS	15920	PTKMHNHPW	15961	QSPSDIQCT
15839	MVHFTSHPI	15880	NNMEIMKTC	15921	PVKLETHES	15962	QTADHRKGM
15840	MVLTDAHTA	15881	NPKQTHWDY	15922	PWGRSLVHC	15963	QTATMAVWN
15841	MVQCMGQLA	15882	NPRGEFMAH	15923	PWPSMACQK	15964	QTDHFDSCR
15842	MVQPGHFES	15883	NQLQPICKP	15924	PYAPKNCCV	15965	QTRAIVLLD
15843	MVVASSASD	15884	NSAHELHSP	15925	QAIGRDQSF	15966	QTYAQTHRT
15844	MWDSVYYPF	15885	NSGTCPAPH	15926	QAMWGNGQG	15967	QVDWLKSSF
15845	MYAPRCQTE	15886	NSSTVCNNR	15927	QASVHVMAD	15968	QVEDELANY
15846	MYGYTMNSG	15887	NSVMQMCVD	15928	QDHALVYRN	15969	QVEMKDCDQ
15847	MYSCWTVNT	15888	NTFMPNGAN	15929	QDYYSKPEW	15970	QVFPCDYEN
15848	MYSNLVHWS	15889	NTHMAHKSE	15930	QEGIRTDVS	15971	QVMRCDFDD
15849	MYSYPHRRS	15890	NTQTLMTNG	15931	QEMPLPMKS	15972	QVNLENHSC
15850	MYTTHCHGM	15891	NTSLVRESE	15932	QFAEWWTMA	15973	QVSVDCSQC
15851	NAGLVNLGK	15892	NVECVQQQW	15933	QFILSHRRD	15974	QVVVIVWQA
15852	NAQEEPESW	15893	NVHDDHGAR	15934	QFLPCNFRC	15975	QVWINNVNI
15853	NARYDKDNL	15894	NVMAGHDHE	15935	QFMTCQNYP	15976	QWGTMMKNQ
15854	NAVEQYVHH	15895	NVMFEQDDR	15936	QFVSWYTYR	15977	QWMTGMCWE
15855	NCKEKEDCQ	15896	NVQMQTLSW	15937	QGEMPDMFK	15978	QYAHISAHS
15856	NCMASGWRG	15897	NVWTCDVCS	15938	QGNSPYHEA	15979	QYQSVSNFS
15857	NCWINTRLP	15898	NWFKMQAGM	15939	QGQPEVVVW	15980	RCAQYEVAG
15858	NDAKNLEGW	15899	NYAETPCYS	15940	QGQPTDASG	15981	RCLKRKTCF
15859	NEKEYSWAA	15900	NYNFYATHV	15941	QGYHFAWND	15982	RCTGFDDFE
15860	NEKHMSYCF	15901	NYQNNMTFN	15942	QHALNQAVA	15983	RDETSFRHC
15861	NERHTFCVD	15902	PDVLHPNND	15943	QHIHEIGKK	15984	RDSDKERCK
15862	NEWLMYLLT	15903	PFRTDVRGS	15944	QIAMATMFR	15985	RFHTDYRRH
15863	NFALATTTE	15904	PGGTPRPEY	15945	QIDNVRHHP	15986	RFVMRPKES
15864	NFTHREKEC	15905	PGKTWLWEL	15946	QIMANACQN	15987	RGEQGVLTL
15865	NGKCCHPGP	15906	PGLDSTMHY	15947	QITDMQDMQ	15988	RGLFDNWIF
15866	NHALMHGFQ	15907	PHPKSHKQH	15948	QLVEAQLPP	15989	RGTYSYSAR
15867	NHCINSTEL	15908	PIKYKHRQP	15949	QLVEKEDMN	15990	RHLDQIPWP
15868	NHESKDASW	15909	PIRREWHLW	15950	QMSPAGCTE	15991	RHSWHPHFE
15869	NIGFQMTMG	15910	PKEQYEWKD	15951	QNKVTQREC	15992	RHYFYQDDD
15870	NIGYANLWC	15911	PKVDKRHKL	15952	QPSPGHAYP	15993	RIHEGSISG
15871	NIQFQVGNY	15912	PLASMWQHD	15953	QQDAAFHWT	15994	RKKGELCSR
15872	NIQWITDVG	15913	PLMLLSKAM	15954	QQGFIVYWH	15995	RKWCEFDAI
15873	NIVHYCNPS	15914	PLTFRGHNH	15955	QQIHMGHQA	15996	RMCCSDSGE
15874	NKSDEQLVQ	15915	PMEKFYCGC	15956	QQTLFRTGS	15997	RMEQQGVMN
15875	NLTFCLFMG	15916	PNTERTYEC	15957	QQYRASNFA	15998	RMESHIAQT
15876	NMGVYQFLG	15917	PQHAGRHHE	15958	QSCVTLQTN	15999	RMHMPWCHG

[2376]

16000	RNPHMDQSM	16030	SIQPLAATT	16060	STGQDQEHT	16090	TEKMSHCSE
16001	RPIHDFGVY	16031	SKMEEALRM	16061	STHGQASAI	16091	TESQSGGQP
16002	RTDQMELAT	16032	SLRLELRPQ	16062	STKQVEALG	16092	TFARRADYA
16003	RTETKSTYF	16033	SMKHRMQAF	16063	STVNTSSPE	16093	TFPVGYNYT
16004	RTYWLDDGK	16034	SMLQFLSDD	16064	SVCLDIPCI	16094	TGEWQQGKP
16005	RVDQEGGRD	16035	SMYFKPTQD	16065	SVDYTNALT	16095	THMTIDDAS
16006	RYECYRSPF	16036	SNEMCGFAM	16066	SVENYERCQ	16096	TIGVDPSCQ
16007	SAESRQGES	16037	SNLQIQDAK	16067	SVRFDGNDW	16097	TIKHILNHK
16008	SATPVGKEY	16038	SNSPTCYAY	16068	SVTVNNVHC	16098	TIKNYCQGR
16009	SAVHRNDVA	16039	SNTMRLNAE	16069	SWAVLHQAQ	16099	TIKYDLHFA
16010	SAVHYTYMD	16040	SNYAPFYCD	16070	SWDMEWDGP	16100	TIMPRHRHA
16011	SCAFVDEWS	16041	SPHCYESTP	16071	SWDVPLYNP	16101	TITTFWDFE
16012	SCDAVDGLW	16042	SPMVKKSGT	16072	SYCNLQFMR	16102	TKRPFKGQW
16013	SCDLKHDSG	16043	SPMYEMSAF	16073	SYEKSVNMG	16103	TLHHNSDAK
16014	SCTDFNSMG	16044	SPNHDAGQN	16074	SYQAVFGSI	16104	TLIYNAQVT
16015	SDHYNPEPY	16045	SQDTEHCWA	16075	SYVQRHCVF	16105	TMDQKNDMD
16016	SEEESKEGG	16046	SQGCQVKQC	16076	TAFSGWKEQ	16106	TMGTAAKSH
16017	SEFCAWKCD	16047	SQMPHSKAA	16077	TAGLGIESV	16107	TMPVDVMIP
16018	SELVRWKHS	16048	SQSNFIIME	16078	TAMDCMCGT	16108	TMPWQGNHC
16019	SEQNRGSHT	16049	SRHWRPQSC	16079	TAMKFPGGI	16109	TMQNTEPGI
16020	SERKDQRMW	16050	SSGLYWKKN	16080	TASPAFDES	16110	TMREFGEGD
16021	SEWQIFFFN	16051	SSLNYSPGA	16081	TASSWLMHQ	16111	TNADSWDQE
16022	SFRIANQTN	16052	STAAYQFEE	16082	TATGMNRCE	16112	TNGVPPWWS
16023	SFYSTGIWG	16053	STAEDQADH	16083	TATSWWAFK	16113	TNHINFGWI
16024	SGPDMTAGN	16054	STAPDNTAF	16084	TAVWMDKTS	16114	TNHSFWIQG
16025	SHRWVMKEH	16055	STDALWQEN	16085	TCFMGWRCL	16115	TPDCYSMGA
16026	SHSHQPSAS	16056	STDCIFFQH	16086	TCRMGMEIG	16116	TPNCGINHE
16027	SICGAVQNF	16057	STDSMDAEA	16087	TDQTCSKCI	16117	TQCNTDQRT
16028	SIKGRDWGD	16058	STEHLGRND	16088	TEEICQLRE		
16029	SIMVSTHNN	16059	STEPSGIHQ	16089	TEKDNEEHQ		

[2377] 表77

[2378] 驱动肺组织嗜性的AAV5 VP1衣壳多肽中581至589区域的序列

[2379]

SEQ ID NO	581-589 序列						
18118	QTNTMQMLS	18122	DILMFEWFD	18129	DIRMFEWVD	18136	EGKQAYFFP
18119	GPRHIEEDG	18123	DIRMFAWFD	18130	DIRMFKWFD	18137	ELKQAYFFP
18120	RKADVEYQG	18124	DIRMFELFD	18131	DIRMFVWFD	18138	EPKRLLRTV
18121	SCDMQQAHQ	18125	DIRMFEWFD	18132	DIRMIEWFD	18139	EPKRLLWTV
		18126	DIRMFEWFG	18133	DIRMVEWFD	18140	EWEQAYFFP
		18127	DIRMFEWFN	18134	DIRVFEWFD	18141	EWKPAYFFP
		18128	DIRMFEWFY	18135	DMRMFEWFD	18142	EWKQACFFP

[2380]

18143	EWKQASFFP	18184	SKNLYTVAG	18225	MCSHNLPKS	18266	ADRQSQMHL
18144	EWKQAYFFA	18185	TMTSTLQST	18226	MHTMTNAMK	18267	AFLELAMYP
18145	EWKQAYFFP	18186	VWKQAYFFP	18227	MNTHEQWGG	18268	AFTEGQLNE
18146	EWKRAYFFP	18187	GTKLIWPYQ	18228	MSLMPHGWH	18269	AGNRWWWSI
18147	EWRQAYFFP	18188	CSTLAKHDE	18229	MTDLMYIQP	18270	AHASWKCWF
18148	GIRMFEWFD	18189	DVIHTTHSG	18230	MVECLYYHG	18271	AIGTLAPNL
18149	GMAHRFRGR	18190	DFMWYHICS	18231	MVFLQCQQN	18272	AILMQQGAG
18150	GMAHWFRGR	18191	DVEFCMKQY	18232	MWNTSNLFH	18273	AIMHPPYCE
18151	GMTHRFRGR	18192	MMLGYMEQD	18233	NVNLQMMMT	18274	AIQSGNWSS
18152	HIESFTYCN	18193	DFNGSNWLP	18234	NYNGYSDAQ	18275	AKLEQMDSA
18153	HTESFAYCN	18194	AGRFKFPQP	18235	PSTKPPKCK	18276	AKTHCHCWW
18154	HTESFTYCN	18195	AMYQCPSMD	18236	QDWSDFNDC	18277	ALHHCIKKA
18155	HTESFTYCS	18196	ANWKHWQWR	18237	QNVQLENSQ	18278	ALIQFMMDQ
18156	HTESFTYCT	18197	ATVNSDTLT	18238	SKTKKRFAW	18279	AMAYSWTMA
18157	HTKSFTYCN	18198	AVGHFEYVL	18239	SRYDTNCWW	18280	AMCFNKCWQ
18158	HTQFINGRG	18199	AVWTMLEHC	18240	STHNRLDWQ	18281	AMSQQHDHQ
18159	HTQFTDGRG	18200	CQVWWENRD	18241	TGDEIVGTY	18282	AMSTLGECP
18160	HTQFTNGRG	18201	CQYAVTKMG	18242	TSIPPRHNP	18283	APQFYVSYD
18161	MSHSRIQEM	18202	CSGVDCTGA	18243	VAAYNFPTC	18284	APTKVHCQG
18162	NIRMFEWFD	18203	DCSSCMKDL	18244	VMEGVRNME	18285	AQNPFDIDA
18163	NISFPNQSE	18204	DRKLGDAHG	18245	WTHMYENID	18286	AQNREGAKL
18164	NTESFTYCN	18205	DTARWDVMT	18246	WYNQGAEYH	18287	AQTFCSFDR
18165	QLDARWYRC	18206	DVVNCTVHT	18247	WYWRLDCLW	18288	ARSGINGHE
18166	QLNARWYRC	18207	EICSRGRTI	18248	YATIPGVFP	18289	ASGCLQQIS
18167	QLNARWYRW	18208	ELGHCLLYE	18249	YCHRSDCDR	18290	ASIDVASWG
18168	QLNAWWYRC	18209	ELQSFMQYG	18250	YMDEMITEL	18291	ASTPWCWFE
18169	QLSARWYRC	18210	ERFKTFCEE	18251	AAMPTHHTM	18292	ATDNSDTLT
18170	RLDYRHYSK	18211	FASSHWYQG	18252	AANIMISQA	18293	ATGIDPLGA
18171	RLNYRHCSK	18212	FYLWDTRET	18253	AANVTQDTS	18294	ATMIKFECT
18172	RLNYRHYSK	18213	GEANMAPSR	18254	AAPKYSAHQ	18295	ATYISAHDA
18173	RLNYRHYSR	18214	GMSEQIAIT	18255	AAQPHFVVA	18296	AVEHIERSF
18174	RLNYRHYST	18215	GWSTDWTET	18256	AASDWCRYS	18297	AVISPERYH
18175	RLNYRPYSK	18216	IASKFNICC	18257	AASWQEFFP	18298	AVNDWNNDW
18176	RLNYWHYSK	18217	IFVYCEKMP	18258	AAYFVGENC	18299	AVTRPSNKM
18177	SKNLCTMAG	18218	ITSSPNHIF	18259	ACEWKQHWA	18300	AWDPMTSTF
18178	SKNLYTMAC	18219	IVDTSARYQ	18260	ACGKMHGPN	18301	AWGGFNAEM
18179	SKNLYTMAD	18220	KANWVESME	18261	ACGVVLKAA	18302	AWILFWWFV
18180	SKNLYTMAG	18221	KTGWGGITG	18262	ACKQFKKHA	18303	AYCSEERSP
18181	SKNLYTMAV	18222	KWKVITTCY	18263	ACLQQDIAM	18304	AYENVQHPN
18182	SKNLYTMSG	18223	LLCTCGWCE	18264	ADAHQRTVS	18305	AYEPFASSF
18183	SKNLYTMVG	18224	LSYENDAFQ	18265	ADGQMLQAH	18306	AYNRMMLTC

[2381]

18307	CAFPVNITA	18348	DMLKFHFNY	18389	EIHDEASCY	18430	FQRLEMKQH
18308	CAGVTSHGE	18349	DMMDNQIAS	18390	EIHKEMGFG	18431	FQRWEMKQH
18309	CAICNAQCP	18350	DNIRYPCHS	18391	EIMFYMRKQ	18432	FSGKYDQRG
18310	CAQSVHFHK	18351	DNQFFNGDY	18392	EINIKKKQM	18433	FSQYPSAFA
18311	CAYEVYDGK	18352	DPVWLPFMF	18393	EINKMAALG	18434	FSRPFEQYH
18312	CFCKNGTYL	18353	DQFARHKKY	18394	EIRTVESAG	18435	FVAQMWDDH
18313	CGEWLNECS	18354	DRPKKNLCC	18395	EISWIPREN	18436	FYRQLDEAG
18314	CHHFNMPSQ	18355	DRPSTGWYL	18396	EKHSVENYG	18437	GASSKSLDA
18315	CILHVTDNG	18356	DSDIVNDYC	18397	EKIDWQSCV	18438	GDIYYKCWV
18316	CKSNWHIFK	18357	DTKVKKEFG	18398	EMAICAEHS	18439	GDLWNNHLE
18317	CLMSNITHI	18358	DTMLRLCYP	18399	EMGRLPNHN	18440	GFEDRGNQE
18318	CPFQHLGPK	18359	DTNWMQWNH	18400	EMVDQIVDF	18441	GFMLRMSTD
18319	CQQYWNAAW	18360	DTSLNDGWM	18401	EMYPEAAAK	18442	GFSLFMMNR
18320	CSAVYLANH	18361	DVCCQFSAY	18402	ENYPPSQCA	18443	GGMTVLIEY
18321	CVDEYYNSS	18362	DVLVATQVQ	18403	EPNHPESLA	18444	GHQLMFFKF
18322	CVIDWKAMS	18363	DYQRKTRYS	18404	EQKVGNQLS	18445	GIKEQRRFV
18323	CVTKTMWTG	18364	EACKCSVVP	18405	ERCMSEWSF	18446	GIRMDLCTF
18324	CWRSGSMNL	18365	EAFYLQKPG	18406	ERWSGTRRG	18447	GITHDFMAG
18325	CWSITNQDH	18366	EAIQDAHQM	18407	ERYHFMTHQ	18448	GIVIKPHGE
18326	CYASMSASY	18367	EANKGASQD	18408	ESKYNRKYN	18449	GIYPQFTWQ
18327	CYDWYEGAE	18368	ECCNMSPPG	18409	ESRDAVCPD	18450	GKEFALSAF
18328	CYSENDHQV	18369	ECGRFAPME	18410	ETCHTYAQH	18451	GLATLGLSG
18329	DAAHMGYEP	18370	ECHKCEVCV	18411	ETHGDRKSR	18452	GLGMVHRMH
18330	DAGHSFQIF	18371	EDRPMAMME	18412	ETICCDSAH	18453	GMMWRSKWI
18331	DAGPIWYSS	18372	EEDKYNVVQ	18413	ETMGVHYCE	18454	GMRCWRHRL
18332	DANQRATNS	18373	EEKKLTVIM	18414	ETRMQRRSQ	18455	GMRYSEWHW
18333	DANSLNSWP	18374	EFGTYDNIG	18415	ETWPLIDDA	18456	GMVDVCDCC
18334	DARSMAPVE	18375	EFRYGPQYG	18416	EVADCLKHF	18457	GQHTIHKNS
18335	DAYSVHHYV	18376	EFVIAEETS	18417	EVCNEMPEP	18458	GQTVQPCST
18336	DCNSLINVA	18377	EGANYHWDH	18418	EVELDERLS	18459	GSCCLNNET
18337	DCRHTILYV	18378	EGAQQESVG	18419	EVELYGCIC	18460	GSLTMPDAH
18338	DDMLMQFQH	18379	EGARYTKFH	18420	EVESQEHLH	18461	GSYYVNWDQ
18339	DFGCMDYGP	18380	EGGMGPQQS	18421	EVTRETDYM	18462	GTCLQDNYN
18340	DFHSYLKQQ	18381	EGKWMSVHH	18422	EYCMLEDYQ	18463	GTGWEASAD
18341	DGFQQGEGI	18382	EGRPQYTDP	18423	EYNRRYLLE	18464	GTGWGASAD
18342	DINRSGNQC	18383	EGSFVLSTE	18424	EYQNYQLVT	18465	GTLSSSIVE
18343	DKWCCEMTA	18384	EHAPMVHAL	18425	EYTRTQWAH	18466	GTNQVCLFT
18344	DLEYIWHWG	18385	EHQRMKFRN	18426	FHHLNTPEM	18467	GVDWRHQQN
18345	DLIGSGPWF	18386	EHVQKVNYP	18427	FIFEEQRWA	18468	GVISEGISA
18346	DMHECIGCS	18387	EHYCVTAWY	18428	FKETWLFPT	18469	GVMVIAGMG
18347	DMHRQHLIE	18388	EIDLWRTPT	18429	FLCCPKECP	18470	GVNQYDEMI

[2382]

18471	GVVQMVIGD	18512	IVRTMDNEQ	18553	LNKYNPHKT	18594	MNSCNNYCM
18472	GWMDLMSAR	18513	IYATNIAWP	18554	LQDLTQLCK	18595	MPVERNVTP
18473	GYDHHGTEH	18514	IYMHKIALF	18555	LRQCPLNKL	18596	MQEDCMSGE
18474	HCMEYIHHA	18515	IYSQIGRFD	18556	LSDQCLDLA	18597	MQNEDVPHT
18475	HDCEFLKDE	18516	KCAGWMPSS	18557	LSNQQWYAG	18598	MQQECVDMW
18476	HEVERRRAK	18517	KDVEECNSL	18558	LSRSGHDQE	18599	MQQGYGIFA
18477	HGMQLEAFF	18518	KEYPRIGRM	18559	LSTDNNIQF	18600	MRFILLADW
18478	HHNIKEAWM	18519	KGSNDGADK	18560	LSTMEWKTN	18601	MSQCIVMHR
18479	HIESYMCRD	18520	KHQSTCGQD	18561	LSTNPQHCG	18602	MTMQDINGQ
18480	HIRHRDENG	18521	KMDFWAVPI	18562	LTAMKKQNA	18603	MVWNTSMDP
18481	HKNRLLRSV	18522	KQSELGIGC	18563	LTEVKWCNC	18604	MYTMQQSGI
18482	HKSRATHTG	18523	KSSNYHSWP	18564	LTGYLNDMT	18605	NAAPISHIF
18483	HNHILIQTS	18524	KTEQGGDAE	18565	LVDKNMFIN	18606	NCENEWRLN
18484	HPMDGLIQV	18525	KTTDECWHV	18566	LVTQWVPAY	18607	NCERRKSAG
18485	HSEHQGRMV	18526	KVEDVGDYA	18567	LWAWSGEPK	18608	NCGSLPVTE
18486	HSFWIAQMQ	18527	KVTEAAKTM	18568	LWNDMHFHG	18609	NERPYEGMG
18487	HSGIGLTMQ	18528	KWCMFFDQH	18569	MAGRDDGRA	18610	NFFSLHMMS
18488	HSWLENSAC	18529	KYELSQCFG	18570	MAIIRLAEM	18611	NGCECHECW
18489	HVEYVQKYE	18530	KYHQAGQLC	18571	MAKSSRNDQ	18612	NGFQTRMRV
18490	HVFQVSDED	18531	LADSSMLWE	18572	MAPKFSSYG	18613	NGWMNQSAG
18491	HWHNVISKD	18532	LADVQFLYV	18573	MASFPMENK	18614	NIQWQWMTG
18492	IADTPDALQ	18533	LADWSYEAS	18574	MCQHEQATF	18615	NLNVMFSME
18493	IEDISKKWA	18534	LATDVPVSE	18575	MCTSVTMNG	18616	NQEGQGMFT
18494	IFQPCFRDW	18535	LCFQSVHEE	18576	MEHSTHDDD	18617	NSAHIMKDG
18495	IGQSTQYCA	18536	LCTAVFRDQ	18577	MENYQWLAP	18618	NSMHGDCQW
18496	IHQSVIQKG	18537	LFMHTMWGA	18578	MFGCIIGDF	18619	NTCPVESAH
18497	IISRYDNDD	18538	LFNCTVWLE	18579	MGSKRKPWT	18620	NTDNNWTSG
18498	IKYSIDHFN	18539	LFRYQRKTQ	18580	MGVKVDGTN	18621	NTNLLEITH
18499	IMELTHHNA	18540	LGDWAPASI	18581	MHDGYRGMQ	18622	NTTYRTEYF
18500	INAVPWQLL	18541	LGNSFKRAR	18582	MHHCYQTFA	18623	NVVFKQSAH
18501	IPAGMTPWC	18542	LGQQAMPLK	18583	MHIAQMLGP	18624	NVWVIKDIC
18502	IQAPTENNN	18543	LHFKIEQVD	18584	MHSNLHWYG	18625	NVYFSWMWT
18503	IRPMYETEE	18544	LHNNMNAFG	18585	MHSPCSGPW	18626	NWTCLMSIP
18504	ISAPGCVFC	18545	LIALDMCHP	18586	MIDVLSWLE	18627	NYPLQWGVE
18505	ISDRMNYMN	18546	LIQSYVNTG	18587	MIESQGDLQ	18628	PATDQIDEC
18506	ISERVHQTE	18547	LIRAKHRTM	18588	MILPSMAWC	18629	PATRVFADI
18507	ISQFGANVH	18548	LKKEVQRYQ	18589	MISCGGDDH	18630	PDWMDWNHY
18508	ISSDKKIMG	18549	LKSEVSIME	18590	MKDHWDFYG	18631	PFVKSNIAL
18509	ITAPRMVET	18550	LMYWKNGMS	18591	MMIGTYMGE	18632	PKQLFWYWW
18510	ITGCIQKAK	18551	LNAWVNRHS	18592	MMNQTCCHM	18633	PLFAITTIE
18511	ITKWSLGNP	18552	LNHKYERMT	18593	MMVVRGKVE	18634	PLTMGHGEC

[2383]

18635	PQHNCGYNN	18676	RADHTIDFS	18717	SREICQLEY	18758	TSQQFITAY
18636	PSENNKRKK	18677	RCDTSQWWQ	18718	SSPWKDFRQ	18759	TSSWLQCYI
18637	PVDWIIAAE	18678	REGWPGHKY	18719	SSQIKTSSF	18760	TSTWPGVIA
18638	QAIFVRQLN	18679	RENLCNVPP	18720	STACAESGW	18761	TTAISLIGE
18639	QCGHDHQYT	18680	RGGLIHFST	18721	STHPYMLSK	18762	TTAYSPNDR
18640	QCKYSERGT	18681	RHLDIQNHD	18722	STHQPDCFR	18763	TTNLEQMIP
18641	QFAEKPGSW	18682	RHMESLNSA	18723	STMCMVRDS	18764	TTQDGQMYS
18642	QFEGGDSMG	18683	RIHLWENNP	18724	STTHLQYYC	18765	TTQSMQSGS
18643	QGTGRYAWG	18684	RLCWMTPCC	18725	SVGPQLYLT	18766	TVCMFWVWC
18644	QGYRKLMNQ	18685	RLENFHEAH	18726	SVGTWQNMG	18767	TVDPFKRIW
18645	QHHTYMCIS	18686	RMELRMIFP	18727	SVHKAENWW	18768	TVEEAVWAN
18646	QIMFNHQNL	18687	RMEPLHCCE	18728	SVLMQMGSE	18769	TVMMPMRLD
18647	QIVGNQKRE	18688	RPYPAMFKG	18729	SVTTCHMPL	18770	TVNICLQHQ
18648	QKMEMFEGS	18689	RQHHVGNEV	18730	TADHNLFDY	18771	TWNCTFTGG
18649	QLETFTVHN	18690	RQPPLKWKR	18731	TADWEDQTP	18772	VAHQSPNMM
18650	QNDVMMGNH	18691	RSNEKGMNG	18732	TATEGVWGC	18773	VAQPCGQYF
18651	QQSFFFGEQ	18692	RTDFLACHC	18733	TCCYDQRGP	18774	VCIDHRAYS
18652	QQSSMATEY	18693	RTINFSLMI	18734	TCDSNATWY	18775	VCQQWLTQN
18653	QQWDVAMMC	18694	RTMYWPSVN	18735	TCEHRLMAT	18776	VDYPRKCDK
18654	QRVDQQSLG	18695	RVDFTPIVH	18736	TCGTTANFE	18777	VGEPLAMFM
18655	QSEWWRRHA	18696	RVGMDPRDG	18737	TFWYTGHWG	18778	VGQWFFKAG
18656	QSHSMMYFG	18697	RWAEMHFID	18738	TGEYTNSWD	18779	VIDCCYSCG
18657	QSHSTDYCS	18698	RWAYLEEFR	18739	THGCGQIMP	18780	VIEQHIQQF
18658	QSQPDNKGM	18699	SAGWDGKVC	18740	THLSCRMWG	18781	VKWDKKTKQ
18659	QSQYKAECK	18700	SAVNYFMFG	18741	TIGIPYSHD	18782	VLFTIGKDS
18660	QSSGEERGT	18701	SAYVVCSHD	18742	TKKDLDRQW	18783	VLHRTSYKH
18661	QSSWDGQDD	18702	SCDPRHVDE	18743	TLALDCKWP	18784	VMEFDKPSA
18662	QSTQTSYIS	18703	SEANMAPSR	18744	TLCIREIMM	18785	VMKGVDKFH
18663	QSTSDECNL	18704	SELDQYTMA	18745	TLTMMQGCS	18786	VNRCCYPNY
18664	QTDIWMMWA	18705	SFLTTAYEA	18746	TMFAVTVKN	18787	VPEHMDRDQ
18665	QTDSVMECT	18706	SGHPRHMCY	18747	TMGFLWYEQ	18788	VQGLLIFVS
18666	QTTHFRYFH	18707	SHMPFSDMC	18748	TNKMLGRAP	18789	VQHVRYWHS
18667	QTTQKSEER	18708	SHMQMNWVD	18749	TPGRNNQGM	18790	VSFPTNTMQ
18668	QVDTRYMNE	18709	SHTECHTWG	18750	TPIGVQKLA	18791	VSLPHSKHL
18669	QVGMVAISQ	18710	SIISDEYFT	18751	TPSQQAEQC	18792	VSNHWMYHL
18670	QVLWANNHD	18711	SITPRDEKP	18752	TQAPIWENL	18793	VTAWWDTAE
18671	QVSFEWYKC	18712	SKIKPADPN	18753	TQEHSHEVS	18794	VTEVYERST
18672	QWPLNPYGK	18713	SLSVKFKTN	18754	TSDVRLTEE	18795	VTGFSFTPN
18673	QWRILDVVK	18714	SMQQAHGWK	18755	TSGHMRNMF	18796	VTGTMLKDA
18674	QYAMSYAGE	18715	SPGSCQAFE	18756	TSIDETELF	18797	VTKMPDFAP
18675	QYYEHEFLD	18716	SQNQYIKAA	18757	TSIWINAVK	18798	VTQYNPKMG

[2384]

18799	VTRQVIESQ	18840	TIDWCEPHC	18881	LHKAYVTDS	18922	IAKIGNCVW
18800	VTWSTHQWL	18841	SFSHTDLEA	18882	AVDHYSNGA	18923	YFPSIQCYE
18801	VVWDVDYYF	18842	AIEHGAWKC	18883	DMKQKAEYT	18924	NMPSLRTRI
18802	VWDALQNPM	18843	TLTTDKTYS	18884	IMYLKRCCT	18925	CKNHYMMAA
18803	VYVQEGNCS	18844	KDGRNPHCL	18885	DQRHGNVSP	18926	SECEVLCSL
18804	WAHTNMVYT	18845	WHCVAPNSF	18886	RVTIMFTGT	18927	ERQAHNHYA
18805	WFCFEDSPS	18846	ACIPQMHSA	18887	TTYPHNTHG	18928	ERFHSYPLG
18806	WFMCDYNGD	18847	NYAMRRDSY	18888	PPRFLQTTG	18929	MAEQMERDF
18807	WGWSPWKKI	18848	IYMSTGNTP	18889	YPTMTFPEE	18930	DQEAQVVTG
18808	WKTKVENPE	18849	GYRNYFIWP	18890	GCEEVGRCQ	18931	ACVERLHHK
18809	WWCWMAGPV	18850	MFTPCQEGL	18891	MAAITQNVA	18932	NHSLCWDSK
18810	WWTHKYFYL	18851	GSTWGASWH	18892	WPTMLSAHM	18933	MYDTMGAWC
18811	YAHEEDQNP	18852	GQFRSGAYR	18893	KCLPYMEMF	18934	QRNEMSASR
18812	YAWISEYEA	18853	VYILEGGIH	18894	AHAWLQHWA	18935	GHLFNNNEW
18813	YDVTKLGEQ	18854	NTQVISRWT	18895	PDLWYEKSS	18936	MLPLEMKNL
18814	YEAWKVNID	18855	CYMNFIVDG	18896	DTRRNLCCD	18937	ATGTYNLQE
18815	YHCCWTSHP	18856	NHKMVQDAL	18897	LIEKQRWML	18938	YVDHWTLGD
18816	YHNIATPIP	18857	DVYMAVVGS	18898	ENQSTMSHT	18939	IAQSEADYW
18817	YPNTWCVVV	18858	GITMHPMHH	18899	QWMHCQLSR	18940	LINSIVCGD
18818	YQGPAVLCA	18859	QMTSSTNCA	18900	ICHIRTRGN	18941	LVQKWMRTC
18819	YRTHCFKGH	18860	MHGIARMQQ	18901	YISEYVMMV	18942	AAMESCAEI
18820	YSCHFNQKT	18861	RVWSLEWHL	18902	QATGMHRWQ	18943	DMTLQSVSS
18821	YWNGRLKWD	18862	EVYCYQLQS	18903	MFMTNVNSF	18944	AEDTPYTIY
18822	ITTSNVEND	18863	QENDERIWY	18904	KTVETYAAH	18945	HFQPVFMQP
18823	CCFPHYFVG	18864	AMTRCKAGH	18905	HIKVEHEEV	18946	LSPWPQTDQ
18824	MFELHDHQF	18865	KPLASIGQY	18906	KMAGIGCWY	18947	AGMYKCQED
18825	RLCDTHKCL	18866	GAANRIQSW	18907	QCTAHNTCM	18948	PHAETHWYT
18826	NNRALMSVE	18867	GHMFETFWG	18908	RTRISQMFG	18949	YPDTADIFW
18827	DSCIKGLAC	18868	EAQPSEING	18909	MGYAKHDSH	18950	VTMIHNAFH
18828	MREKLVHDC	18869	ECLEVKRTC	18910	PIWQGYWPW	18951	DLIVATEWF
18829	TTYHHRNTL	18870	FGDHLRVMY	18911	RCGEIMGLD	18952	HSMMPGWPS
18830	TWNTWQQDF	18871	WTHCTNNQI	18912	LPSHCNLGS	18953	IWITCPYAA
18831	DSASPIMGM	18872	NAYYDSHCE	18913	PAHINDQYM	18954	PYGPIEYDE
18832	MMAMKQGFY	18873	EKYQIHWDR	18914	DRVNLMWKD	18955	ATMNLSWGL
18833	KACMSHQGR	18874	QALVGNAAE	18915	QCVDCPPYC	18956	NGVGRSGDN
18834	TIGEMVNPS	18875	QSSSSYLWG	18916	AVHTADVCC	18957	RCADNPYFR
18835	NWNHCHQWR	18876	LESVKMDCN	18917	NQREWHGLA	18958	WERCLELNQ
18836	SEDKDQGAQ	18877	QRSVFHQRS	18918	HIHMSDRGA	18959	SNVACIQAF
18837	CAVAGGEWM	18878	NDVWQPHPN	18919	PCCWEQSTN	18960	INYISFVDR
18838	NTMLSWGSP	18879	EDQWISNPF	18920	AFLHGQMWQ	18961	LEIGGKTRT
18839	CSIRGFSAF	18880	DHLSNDAEM	18921	PLTIEVNCT	18962	YDYWMQAPT

[2385]

18963	QSAHPEPMC	19002	DNNFKDKEE	19041	GFKKFHFTF	19080	KVMTAEIAD
18964	ASMHEETCL	19003	DSFYPQLQN	19042	GKERFHRID	19081	LACACQDAP
18965	DMIDPVRHG	19004	DSQLISTLA	19043	GMSLKNYEE	19082	LDGSCYHWG
18966	ADRSNSSCF	19005	DTINWIASQ	19044	GNYHLDQDE	19083	LKATCTAMD
18967	AEYSTNLHS	19006	DTLKSDNYI	19045	GPNGDVARY	19084	LTTVYNWQH
18968	AFSYCMAHM	19007	DTRGKAFFP	19046	GQQTMEMSY	19085	LVEMGCEPR
18969	AIDYKTDWT	19008	DVITVYWDS	19047	GTMHALKDE	19086	LWTMDTVRP
18970	AKACHLWGK	19009	DVTDDQIHM	19048	GTQSERDIQ	19087	LYVGWMNCL
18971	ALQMAYEWS	19010	DVVKSWQMG	19049	GVEFVPIGN	19088	MCYHTGHFA
18972	ANDIAEHWG	19011	DYSLLWMYE	19050	HCIFKDPKI	19089	MGGMIVMCQ
18973	ANVCIITHC	19012	ECAFIHVDE	19051	HEVDRPAPC	19090	MHCRTFSHN
18974	APFMMQCFS	19013	ECMQYFPEW	19052	HITNYGYCG	19091	MHMRTNEGS
18975	APYPSWPTY	19014	EDRLIVRMD	19053	HNPATECRC	19092	MILTEGIIE
18976	AQWRIMFIA	19015	EHITEQDLW	19054	HPLTQMVLE	19093	MLHMMPWID
18977	ARDCWWSDG	19016	EIGMEYGHS	19055	HPLTQMVLG	19094	MMAMKQRFY
18978	ARDRPHRGS	19017	EINNAIASR	19056	HRRRAKVWV	19095	MMGGKFGTT
18979	ASCRCLNWH	19018	EIQQMCVKW	19057	HTRYWGFAL	19096	MNNEVWKAS
18980	ASQRSGRQP	19019	EIYMCKEGN	19058	IAGMHEFQQ	19097	MNQHWPPQI
18981	ATEHVRSWW	19020	EMSWKMKPT	19059	IFILIEWHH	19098	MPLSMKAKK
18982	ATIMFIEIS	19021	ENHPPWLHC	19060	IKTITYHST	19099	MSDTMDMGE
18983	ATKCVLGLG	19022	ERDSLWFGK	19061	ILWSQHEVL	19100	MSIMCCSDC
18984	ATMYMERAD	19023	ESPQIHIYS	19062	INISTPVNC	19101	MSNVCICTG
18985	ATYEHSCTK	19024	ETVQDVGLH	19063	ISHQCHQIH	19102	MTGHNESFT
18986	AVNHAQCGF	19025	EVSWVPTTK	19064	ISTQMSWDG	19103	MTVAGQSRD
18987	CARLGGRGN	19026	EVVNSQFNL	19065	ITFCCRAAK	19104	MTWTRNSGT
18988	CCTILNMSG	19027	EYEWVDEPQ	19066	ITQYTTYEG	19105	MVDREKKKE
18989	CFEFLWEHE	19028	EYFKLRQFG	19067	IVAEHDPGF	19106	MVSSETGSH
18990	CLVRIQKDH	19029	EYSNETHQE	19068	IVMPALMCK	19107	MWKRHCGMG
18991	CQAIINADT	19030	FEADNTMVH	19069	IYKLFQHGM	19108	MWMQMLSAM
18992	CSTMPKLPE	19031	FGHFITKKM	19070	IYRLFQHGM	19109	MYCLEFMND
18993	CTMSPGYDQ	19032	FIKLTTLPS	19071	KADNRNWSC	19110	MYTGMRCWH
18994	CVVMAVDNT	19033	FLEAMQRTP	19072	KAQPSEING	19111	NAPNDLKNY
18995	CWGWAPPDD	19034	FLQYSMSGN	19073	KCLPYMKMF	19112	NASCQYRGA
18996	CWPQSCKPE	19035	FMDECLRVD	19074	KCTAEKAQC	19113	NASCQYRGV
18997	DGENMDEMA	19036	FNSSEQQQM	19075	KCTVEKAQC	19114	NATHMYNAH
18998	DIGRFIAYY	19037	FSHKISVWC	19076	KEREERRFW	19115	NGCIVAEAA
18999	DIKLVEAYA	19038	FTHWVNEQL	19077	KKYFPTGIG	19116	NGYIVHACQ
19000	DKDEFQMAP	19039	FVKVVWGHK	19078	KLDYHGVRQ	19117	NGYQARTKS
19001	DNDMYHSMQ	19040	GDNMDQTQG	19079	KMVQVEHYT		

[2386] 表78

[2387] 驱动淋巴结组织嗜性的AAV5 VP1衣壳多肽中581至589区域的序列

[2388]

SEQ ID NO	581-589 序列
21118	AVDCHNSWI
21119	VQQYHITQE
21120	AAAKWDDVW
21121	AACKEKRNN
21122	AACTMAVAR
21123	AACYDVNKA
21124	AADLKGYIG
21125	AAELNSSQS
21126	AAENCQCPF
21127	AAEWKVGQA
21128	AAGEEWSAL
21129	AAMSCADAS
21130	AANESTHDC
21131	AAPYLTGRD
21132	AASPITVLK
21133	AASQMMNHG
21134	AATMNGQCM
21135	AAVGYFHTS
21136	AAVMESFWR
21137	AAVMSLCTV
21138	ACATTNFYD
21139	ACDPSHKHQ
21140	ACDPYSINF
21141	ACETWQWRE
21142	ACEWWFNKE
21143	ACIMQMTQD
21144	ACKLLDSHL
21145	ACMPICMSA
21146	ACTYRMTWH
21147	ACVDLVVDK
21148	ACVKHLYDW
21149	ADAWETPYP
21150	ADCHTQNVA
21151	ADCRYNLAC
21152	ADCSDQMSQ
21153	ADFDGILWY
21154	ADIQFARMP
21155	ADLDINQYW
21156	ADLHKRHNW
21157	ADNLTPHQW
21158	AEHKWPEEL
21159	AEQFVDVAG
21160	AESPVAYQQ
21161	AEWVEQAVH
21162	AFEHQYCHL
21163	AFEIVHGAN
21164	AFETLRNVK
21165	AFGEFVILA
21166	AFGHFPGIG
21167	AFHMMEKWE
21168	AFISKWVHP
21169	AFITQVPWY
21170	AFMNTYVAD
21171	AFMPWMHTY
21172	AFMQTICDM
21173	AFQYSNQAH
21174	AFRDISKMI
21175	AFRLLGYSY
21176	AFRNFTHCP
21177	AFSLNNCTH
21178	AFTCCGYEN
21179	AGDQKYVGC
21180	AGEEYTTDF
21181	AGEITKSMF
21182	AGHRDRAET
21183	AGLYLTNED
21184	AGNAMDFCF
21185	AGNEVQTMP
21186	AGQYKLIFW
21187	AGYCSQWRF
21188	AHAECKFWS
21189	AHAKWWDYQ
21190	AHALNEWYA
21191	AHAMVVNQG
21192	AHCYVLRHE
21193	AHELSYSLP
21194	AHFDYDIFW
21195	AHGLFMCMN
21196	AHHEVCWQH
21197	AHHLQICHT
21198	AHLWVQGSH
21199	AHMHTHAIM
21200	AHNQINDNL
21201	AHPHFEENG
21202	AHSWWHRAE
21203	AIAECIGFT
21204	AICQLLGDF
21205	AIEPKEDWM
21206	AIGHGENYR
21207	AIHETIAGT
21208	AIHYLWQDQ
21209	AIMPVRDGQ
21210	AIPCEMNAN
21211	AIQLSQDLM
21212	AISVMHYDM
21213	AITDTSCCN
21214	AITTQGDFY
21215	AITYQGQGV
21216	AIVPITFNM
21217	AIVYKNWAK
21218	AIWRANEST
21219	AIYSLDFEQ
21220	AKFFPHTNT
21221	AKGLNLDIM
21222	AKHLFDAAP
21223	AKHRCMSEW
21224	AKMWECNPK
21225	AKNSLYNYH
21226	AKYWNPELQ
21227	ALHWWQCLT
21228	ALINQPDSA
21229	ALMMYFENQ
21230	ALTEQCVYG
21231	ALVLWDMLH
21232	AMAGPQFSA
21233	AMAHKGMDA
21234	AMAQPVKCA
21235	AMHITQYKT
21236	AMHQNIFGI
21237	AMIHIAMMP
21238	AMQQWVSGC
21239	AMSLQYKNE
21240	AMTSRGSYP
21241	AMVFNMQLG
21242	AMVSSWLVG
21243	AMYVMYHGM
21244	ANDCRGMSE
21245	ANDLLTMQF
21246	ANFGSHAAM
21247	ANHFESSRS
21248	ANHPNMLIP
21249	ANMIRKGYG
21250	ANSFEVCGM
21251	ANSYYHQYC
21252	ANVEQAMCA
21253	ANVPMFRMD
21254	ANYFPTKFP
21255	APGCVNQLH
21256	APLTSYPTL
21257	APPVFYMPQ
21258	APTVINMLK
21259	APWQFNHQW
21260	AQCPMTWSL
21261	AQEVTITHE
21262	AQLLLIVMA
21263	AQRADWTYL
21264	AQWPLHRDT
21265	AQYGRNFAG
21266	ARMQEPIQV

[2389]

21267	ARNANCPRM	21308	ATYCEQYMY	21349	CAYQAMDFF	21390	CKFMYLMKE
21268	ARQINEHQI	21309	AVAGVVWDN	21350	CCCNRCHIY	21391	CKLTMKRQN
21269	ARVDDGQMT	21310	AVAMGGQDD	21351	CCFSTFECY	21392	CKSTHDMAW
21270	ASANNPVFF	21311	AVANWFQHM	21352	CCGDTRCET	21393	CLHCGTFST
21271	ASASYEGEQ	21312	AVCEGEITP	21353	CCGLPARLL	21394	CLHCHSLQC
21272	ASAVAKEWD	21313	AVCSKELDH	21354	CCIRSFVCA	21395	CLMLQSHHT
21273	ASDGIEKPC	21314	AVDHMAMES	21355	CCKTYQRAG	21396	CLRAVDNCF
21274	ASDHLNPFV	21315	AVDQINMEQ	21356	CCMTYNNQG	21397	CMQSESMYC
21275	ASDPWNWWP	21316	AVELSMRAV	21357	CCPGNGDIW	21398	CNATWGVFA
21276	ASGWIQLMC	21317	AVEPYCSML	21358	CCPMRHANF	21399	CNAVKNNDM
21277	ASMAYLKGI	21318	AVESTAIES	21359	CCSGNTGMY	21400	CNEVYTLIH
21278	ASNFHSGLY	21319	AVFHQGSGA	21360	CDHGCWCKG	21401	CNNNYTNWA
21279	ASNFVMEFP	21320	AVITWSQYH	21361	CDHVVCMAP	21402	CNNYSMVHR
21280	ASNKICNFK	21321	AVKPSVEPW	21362	CDWAAYHCW	21403	CNRLHDTWT
21281	ASREIVKHP	21322	AVMPQGHYG	21363	CEGIQVVDF	21404	CNRSEIQFR
21282	ASTHASESQ	21323	AVMTAQGWV	21364	CEVTINGAE	21405	CNVFRDPYL
21283	ASTSQCQFQ	21324	AVNPFPCQT	21365	CFAELSNVA	21406	CPCFYINYP
21284	ASTTIGQEM	21325	AVQSGIETC	21366	CGDHYRKGH	21407	CPETRHVVH
21285	ASVELGGYS	21326	AVSAAMQMD	21367	CGHLQQFIL	21408	CPKDCRIVI
21286	ASVKDFTVQ	21327	AVSDKPYQH	21368	CGIGVNMHH	21409	CPLSWIQQE
21287	ASYGDGPMW	21328	AVSFYQYGC	21369	CGIWNNNAF	21410	CPQPSNRTW
21288	ATAEGEDCT	21329	AVTHKWCEE	21370	CGRQMEDAA	21411	CPQQIQHYN
21289	ATECAWYHG	21330	AVTTLLYPI	21371	CGTQSFIWG	21412	CPTEVREKK
21290	ATFICLTSI	21331	AWAKDVMQC	21372	CGWCGLSKI	21413	CQGFGVPWD
21291	ATHEYASWD	21332	AWVISQNEN	21373	CHAIFCRCA	21414	CQHLHTSYW
21292	ATHPVMLDA	21333	AYASSQDSL	21374	CHDGSFNPH	21415	CQKALMQLG
21293	ATICSGWDR	21334	AYENTFPNP	21375	CHGQYVTWG	21416	CQTCLQFCD
21294	ATIGWERYN	21335	AYEPFWCWG	21376	CHHRMHRAQ	21417	CQVEFQMTT
21295	ATIMVCRFE	21336	AYHHVDYHW	21377	CHMGINRLN	21418	CRDHYDRSC
21296	ATKELAVQW	21337	AYNQNIGTG	21378	CHSPFNHWG	21419	CRFEHQYNC
21297	ATKLDFHPT	21338	AYTDLFNQC	21379	CHWYCSKVP	21420	CRKLQSIEA
21298	ATLTFQQHH	21339	CADFMVSWG	21380	CIDEGVSPD	21421	CRTNGHYYM
21299	ATMLRKEVW	21340	CADWCNTQP	21381	CIETSDQER	21422	CRVEPVDSD
21300	ATMSPECMR	21341	CADYQYQQD	21382	CIEYTPMGR	21423	CSACTPFSK
21301	ATNWEGIDS	21342	CAIGYWEHD	21383	CIGFYMCAG	21424	CSADMNSMP
21302	ATNYVTNHD	21343	CAMMDDRGD	21384	CIGVINGMN	21425	CSADMQQDI
21303	ATRVGGSTW	21344	CAQLFWLPQ	21385	CIHYVQHCQ	21426	CSDASHHFG
21304	ATSAHSCSS	21345	CASDLMENR	21386	CIILRSNDG	21427	CSEAEGDWY
21305	ATSINCNAW	21346	CASDSLVNG	21387	CILCQPHKW	21428	CSFNVMNTA
21306	ATVDYWQAW	21347	CAVNWNVKE	21388	CIWKNVPPW	21429	CSKPDCQWD
21307	ATWTDSWKV	21348	CAVQFNYCV	21389	CKDAYRSHA	21430	CSMQNNRMY

[2390]

21431	CSMYYCWTG	21472	DAAYSNWFF	21513	DDVNKMRKE	21554	DHQLTMEGM
21432	CSNSQDFGA	21473	DACKFSSNI	21514	DDVWMWQWV	21555	DHSMYATQE
21433	CSTEMQMNC	21474	DADEPWLDR	21515	DDYKKWRHP	21556	DICDKNHEG
21434	CSYQRTEYF	21475	DADRAIMHE	21516	DDYKPHCKF	21557	DIEFDITCD
21435	CTATPEIIH	21476	DAEDCHTHK	21517	DECYNLFTV	21558	DIGEFWTAH
21436	CTEININQG	21477	DAEGQVCYM	21518	DEDDYNMPY	21559	DIHRFAEQF
21437	CTFVRELQC	21478	DAETVDSQF	21519	DEHLITMQT	21560	DIHVITKGI
21438	CTGDVDQSE	21479	DAFLIQELC	21520	DEQWSPTTI	21561	DIKQHQLWG
21439	CTGLAQHWW	21480	DAGKPQYSI	21521	DEYVPEAYF	21562	DILDVVKQG
21440	CTGQRHASP	21481	DAHDYDSNF	21522	DFAACQDYL	21563	DIMINDLYS
21441	CTHHYMEQH	21482	DAHNYWGWQ	21523	DFAFENWMG	21564	DIQHLTESG
21442	CTKHMDGQA	21483	DAHWVFKTN	21524	DFAWQECWC	21565	DIQQHNIAA
21443	CTNDWCRSD	21484	DAIPYKQPT	21525	DFEKVSEQA	21566	DIQVPLAYR
21444	CTNLISSYC	21485	DAKFCGQGH	21526	DFGCQEQMT	21567	DIRRMVASL
21445	CTQVIEECP	21486	DALYEMGTT	21527	DFGPRCCFG	21568	DISMESWMP
21446	CTTYQDEMN	21487	DAMPSDRWW	21528	DFGVQKSNE	21569	DITETQCGS
21447	CTWWIPLVY	21488	DANFTMMTD	21529	DFHDQYVID	21570	DIYYSMRGR
21448	CTYTELQYM	21489	DANNKKYHV	21530	DFIWQQCSN	21571	DKACMLWDS
21449	CVDQLSLGH	21490	DARHYFGLQ	21531	DFRFGTRFG	21572	DKCAIEFAS
21450	CVHAISCVG	21491	DARPDYEMQ	21532	DFSRMAGGP	21573	DKGALEFWI
21451	CVIGRWLYM	21492	DARTIVLTH	21533	DFTAFKNDM	21574	DKGEWQCYK
21452	CVISMAYQW	21493	DASDCCYSY	21534	DGCPSGNWP	21575	DKGFIQHVA
21453	CVLKYNHHY	21494	DASLWSVPM	21535	DGELGKFDI	21576	DKMTLPFSP
21454	CVMMQAGGE	21495	DASMHGTVP	21536	DGFDFEKIG	21577	DKPCSICGC
21455	CVSYSGQHD	21496	DASPTRLVY	21537	DGGFQQAEW	21578	DKRWSANQW
21456	CVVNAHQMR	21497	DASVNRVIE	21538	DGHSDLLLS	21579	DKWPPLEPH
21457	CVVRAQSAM	21498	DATKIPMQQ	21539	DGKHLDTDP	21580	DLAWQYQPA
21458	CVVTMSMSD	21499	DATQNHFQY	21540	DGKKTMHET	21581	DLEDAHVGL
21459	CWMFPDYWN	21500	DAVDVAMAM	21541	DGQLHMFYA	21582	DLHGIGVLA
21460	CWMSFSAYE	21501	DAVYYLFQA	21542	DGSGSEWHR	21583	DLIAGNWPA
21461	CYESTEGEH	21502	DAWHSTCNS	21543	DGTADDKRY	21584	DLPSPRRID
21462	CYHEYCCNM	21503	DAYYDHRGC	21544	DGTETIYNM	21585	DLRVMHWHH
21463	CYKYGPPSQ	21504	DCGWLAYEW	21545	DGTYANHMG	21586	DLVCAQWRE
21464	CYMVFMFKA	21505	DCKCPQPDL	21546	DGTYKCPEP	21587	DMAAFHYAG
21465	CYQDIEAGY	21506	DCPWKAFHA	21547	DHCNFMAPY	21588	DMAANHTFP
21466	CYQFKKLTN	21507	DCRGLSSIF	21548	DHDWLQNMG	21589	DMALYHSNV
21467	CYQSEKNGY	21508	DCSVFWIGA	21549	DHEVVQNMP	21590	DMDCYTMME
21468	CYSGVAMAS	21509	DCTFTQQQW	21550	DHKLIQMHT	21591	DMDSPFPWD
21469	CYVTKSWDH	21510	DDKPWVIMA	21551	DHLICQNGP	21592	DMDTYYGHG
21470	CYWFMYKPK	21511	DDLQPMERT	21552	DHMWVNWSQ	21593	DMFTFAHNS
21471	DAAQTIQYP	21512	DDVNGAVIM	21553	DHPFNVKQE	21594	DMGFVNRHE

[2391]

21595	DMIHYPLYG	21636	DRLQCDECS	21677	DTILEQYGP	21718	DVKPGWMGD
21596	DMLMLVTHD	21637	DRNTAMTGG	21678	DTIQWDQGI	21719	DVKWMESWL
21597	DMMGFWDKP	21638	DRQINWQQE	21679	DTKMVPYMP	21720	DVKYLTHQW
21598	DMNELAWAY	21639	DRSDKGCLY	21680	DTKPENNWI	21721	DVLLLEHER
21599	DMRVGNGYF	21640	DRSVHVRAS	21681	DTMHCLSCD	21722	DVLTNLNGF
21600	DMSVIQDGA	21641	DRVEERAYG	21682	DTMPTMDLM	21723	DVNFVGEPE
21601	DMTRNPEWE	21642	DRYLFTHWH	21683	DTQWQEWCE	21724	DVNSYKLYG
21602	DNAACIEMS	21643	DRYQWAKRH	21684	DTRDNWLIR	21725	DVQTIRMEH
21603	DNAEMNSFD	21644	DSALQFIDR	21685	DTRTPAVWP	21726	DVRERDSDD
21604	DNCYSCEAT	21645	DSAMFGQYK	21686	DTSQTSNNT	21727	DVRTEPTIT
21605	DNEGWLMAE	21646	DSANVGWCT	21687	DTSRCLAYW	21728	DVRTVTQHG
21606	DNEVWLSAV	21647	DSAPGKMWP	21688	DTWTLDFEV	21729	DVSQSDWQP
21607	DNHALEFDC	21648	DSAPLQSQL	21689	DTYAEIMYG	21730	DVTAQRAWK
21608	DNMQRKWAG	21649	DSAPPPNVI	21690	DTYPLVIGS	21731	DVTAYSESI
21609	DNQEPSIRS	21650	DSDEGSEHI	21691	DVAMTSLIA	21732	DVTEDVVSP
21610	DNQSKESQA	21651	DSDLYEGWM	21692	DVCDSIGYC	21733	DVTENREVW
21611	DNSDFKWHW	21652	DSEWLHWHA	21693	DVCEQRNWL	21734	DVTIYWNDY
21612	DNVPYCYYQ	21653	DSGFAMWTP	21694	DVDDSQCEY	21735	DVTLNEGVV
21613	DNYINQEAA	21654	DSHFYDQWQ	21695	DVDDWFHQS	21736	DVTPEDACV
21614	DPAHTTKIC	21655	DSHPIAHGF	21696	DVDHICEGH	21737	DVVALQASK
21615	DPDRTPSQH	21656	DSMKGTMFQ	21697	DVDMDTNRG	21738	DVVANLTSS
21616	DPEEHNNYG	21657	DSMMHDCHF	21698	DVENHYMGS	21739	DVVEQFHWT
21617	DPEFQQWTH	21658	DSMNGVEAS	21699	DVEPMKNDF	21740	DVVIVDCAN
21618	DPIDEVNAQ	21659	DSNEEQRQG	21700	DVESRYDCG	21741	DVVVSQMAM
21619	DPNRIDNQY	21660	DSPEQFGLD	21701	DVFCMLASH	21742	DVWAQSQVL
21620	DPSEVRARV	21661	DSQHTAGWP	21702	DVFIQQAYS	21743	DVWLQIDGQ
21621	DQCFMTYDT	21662	DSQLIKQWM	21703	DVFMGARFE	21744	DVWWVEGEY
21622	DQDNPSCWL	21663	DSQPMFACV	21704	DVFPCQKTE	21745	DWASPKSYV
21623	DQEPNWNGV	21664	DSREAVAVD	21705	DVFQCEELS	21746	DWCMSMVWT
21624	DQFDSHSQG	21665	DSRQDHWFE	21706	DVGFNRISR	21747	DWDQQRQYF
21625	DQGCKLRRC	21666	DSVHSYEQK	21707	DVGIQHRPL	21748	DWGAMWVCG
21626	DQGMGSCKD	21667	DSVPNLKAY	21708	DVGLFHGGF	21749	DWGHRIFEH
21627	DQHGASTVF	21668	DSYPNQACT	21709	DVGPQKNIV	21750	DWHHPPVDH
21628	DQIPMKSMM	21669	DTCTTCVSL	21710	DVHMADGNN	21751	DWKPIYEAS
21629	DQKMYVDED	21670	DTCVTAANH	21711	DVILTYDRA	21752	DWPIRSAPI
21630	DQKSVQCYE	21671	DTDTRCIPC	21712	DVIRTHDGV	21753	DYCTGACDC
21631	DQLEGGESN	21672	DTEGDHREG	21713	DVKCMQNTF	21754	DYDVASFYH
21632	DQLGLRDNT	21673	DTELSMVEP	21714	DVKGNNGGG	21755	DYEGPIRPN
21633	DQQPALDAQ	21674	DTFFHGAFE	21715	DVKIVWSGE	21756	DYEHYWHET
21634	DQWSILTFS	21675	DTGIWAYVG	21716	DVKLAGMIH	21757	DYKLHPIGY
21635	DRAKECQCE	21676	DTHRIIMDD	21717	DVKPAMNQA	21758	DYKMDVMPC

[2392]

21759	DYNMMHWAA	21800	ECTVMQRGA	21841	EGWFMKKQD	21882	EIRSHDTGH
21760	DYQRCTFLP	21801	ECWFCDIWR	21842	EGYEYHTQR	21883	EISWQDRDL
21761	EACSNSHEK	21802	EDLHQWAQH	21843	EHATESWAG	21884	EITFNQYGH
21762	EADCVHRVW	21803	EDMFCPTKA	21844	EHATIENWP	21885	EITWESNIA
21763	EADVSLMHC	21804	EEDFRNEAM	21845	EHDGVWMYQ	21886	EIVDLTVME
21764	EAEAEVIIR	21805	EEDMQRTVQ	21846	EHEDTIAHD	21887	EIVEHDYRE
21765	EAECTWASL	21806	EEEAGKWMC	21847	EHFPQQSGD	21888	EIVYMMCPC
21766	EAELKTNFM	21807	EELEQIWHG	21848	EHGPCMSGC	21889	EKAGNLERK
21767	EAESGGAYK	21808	EESLDVVMK	21849	EHIGCGYWQ	21890	EKDQYKMGS
21768	EAEYRNWNP	21809	EFCYEPDWD	21850	EHIRFGHYW	21891	EKEQFSHKG
21769	EAGMFVHRA	21810	EFETFWEVG	21851	EHKPDGANY	21892	EKGMIDLVY
21770	EAGQWLGSP	21811	EFGTMHYAP	21852	EHLKLLIAC	21893	EKGMNMHAG
21771	EAHAWEKTE	21812	EFGTSQKDA	21853	EHMSICDAP	21894	EKHQLSVQQ
21772	EAHEEWSFD	21813	EFGYSPGVS	21854	EHNSEDCIC	21895	EKQWLSRQD
21773	EAHVNVYKY	21814	EFIHAENWY	21855	EHPMENTYQ	21896	EKSGYHFSC
21774	EAKCMHDFS	21815	EFKARMVTP	21856	EHQFMWQDH	21897	EKVICLYRG
21775	EAMIWNNTM	21816	EFLVNTSWL	21857	EHRDHDAWA	21898	EKVTLNMDM
21776	EANQILSAH	21817	EFMCSCMLN	21858	EHVKCNHFF	21899	EKYMNFMCP
21777	EANSSQSQF	21818	EFMTYETNV	21859	EIAGSIMHL	21900	ELADQNQIQ
21778	EANWEHSCI	21819	EFNSIMHWW	21860	EIAKWNCEL	21901	ELAWIWQIH
21779	EAQTTGSTW	21820	EFQWRACFQ	21861	EIAQLQFDQ	21902	ELCCNDMFK
21780	EASLIKEDK	21821	EFRMASSDA	21862	EICAFDVGW	21903	ELERTVKEF
21781	EASSLYSQQ	21822	EFSDVQRTA	21863	EICVIECEQ	21904	ELGYSEKVC
21782	EASWSLHGQ	21823	EFSITMAPW	21864	EIDYHICSG	21905	ELLTCIECG
21783	EATGESFWF	21824	EFSKMSGTA	21865	EIEMAVAYA	21906	ELNFCWQQR
21784	EAWHPGCCQ	21825	EFVEKWLWG	21866	EIETSAREQ	21907	ELYMNNHTE
21785	EAWMCVEAQ	21826	EGAELEWDH	21867	EIFQYMHHH	21908	EMAELNQHV
21786	EAWSMNAPY	21827	EGAQLDVMS	21868	EIGIGAPDY	21909	EMDQVPWGF
21787	EAYFDTAEC	21828	EGDMTTNQY	21869	EIHHGHAVP	21910	EMEFPDMCM
21788	ECAPWVVLS	21829	EGEELRVAY	21870	EIIAQKWTQ	21911	EMEWQDSWW
21789	ECCIGDNMM	21830	EGEEYTLPS	21871	EIIMYEQHM	21912	EMGFGKQGM
21790	ECEMMDYSK	21831	EGFLGFGQN	21872	EIIPTSMFG	21913	EMGMPSYMK
21791	ECFATGEFM	21832	EGFYLEGDP	21873	EIITVRDDE	21914	EMKEAWLSA
21792	ECHKCSKLN	21833	EGKEWIFYE	21874	EIIYRFVDA	21915	EMMVLCENE
21793	ECKLSVCGY	21834	EGMKTQQTF	21875	EILSDDDMP	21916	EMPTTCKNC
21794	ECMHAEWQA	21835	EGMWPHSAF	21876	EIMHRCHFW	21917	EMQCYQTYH
21795	ECMLIGTDG	21836	EGQRLTWAQ	21877	EIMIQVIRD	21918	EMSECSITA
21796	ECPDYMRKD	21837	EGRQHFVYA	21878	EIMMEHDNK	21919	EMTPQQFWH
21797	ECQMHRTFA	21838	EGSQWRTSI	21879	EIMWPLAGN	21920	EMWICHFGG
21798	ECRPWGMHE	21839	EGTMSNGWQ	21880	EINLGHQHN	21921	EMWSRGMVN
21799	ECSESAEAK	21840	EGVRLNLYG	21881	EIQMILTAV	21922	EMYPFSKQL

[2393]

21923	ENAEIKDHM	21964	ERIQWTHMG	22005	ETKIYDCNG	22046	EVRGAEANH
21924	ENAIKHHDS	21965	ERTAPSSRG	22006	ETKLAIDEF	22047	EVSKLATTY
21925	ENAQRHNYA	21966	ERTMVHQYW	22007	ETKSVLDEH	22048	EVSQFGIMN
21926	ENDAEVPDH	21967	ERVDQENCH	22008	ETMCPKGPL	22049	EVTFWTCEG
21927	ENEAMAKWF	21968	ERVTDEHEI	22009	ETMNQIGSC	22050	EVTIRSHYQ
21928	ENFMLRPCN	21969	ERYHMRWEP	22010	ETNLFEHST	22051	EVTLIPDSW
21929	ENHASAHVR	21970	ESANYMNLF	22011	ETPGTKQCS	22052	EVVHSMSDG
21930	ENHIEFWWP	21971	ESCYMMMNW	22012	ETQRDRHTN	22053	EVVNVPMRT
21931	ENLESRLQH	21972	ESGMLFYFM	22013	ETRDYLEAD	22054	EVYRNTMKC
21932	ENLHCDCLY	21973	ESKAHAQTP	22014	ETRQMDGNN	22055	EWAGKQTDM
21933	ENLRMGKWE	21974	ESKTTSCAF	22015	ETSMRFESL	22056	EWAYMGSYP
21934	ENVKLDQCD	21975	ESLRAENMP	22016	ETSRMNVDA	22057	EWDDTSNSG
21935	ENWHLVAHR	21976	ESMGFEECG	22017	ETTCYAYME	22058	EWHFPKAMN
21936	ENYCQCCLE	21977	ESMKLIQCD	22018	ETVERQCEQ	22059	EWNDYPSWY
21937	ENYRHHTGP	21978	ESMNIMTYK	22019	EVAETSHSY	22060	EWRHVELHY
21938	EPEWLGHEE	21979	ESNDNGNQW	22020	EVANGKGDH	22061	EYAGHHCFY
21939	EPGPMQCWD	21980	ESNPVCVLK	22021	EVCLGQRYA	22062	EYCMCMWGA
21940	EPGQFFHDC	21981	ESRCAWLSD	22022	EVDRHLGMQ	22063	EYCNVMIPC
21941	EPLDNVGCG	21982	ESRKEIGYA	22023	EVEACKYEY	22064	EYDDVPEPQ
21942	EPMNGKVME	21983	ESRRHHWAH	22024	EVECTHRGE	22065	EYEQDCRFQ
21943	EPNAQMTAA	21984	ESSDISNYL	22025	EVEVWSVYH	22066	EYHDKHWAG
21944	EPQEVENMT	21985	ESTLLLCHG	22026	EVFPYGCQN	22067	EYHTEWIGM
21945	EPTLVGEVP	21986	ESTNLQCVL	22027	EVGMIVHPD	22068	EYIGEELHK
21946	EPTNIWAVH	21987	ESVTSLDQQ	22028	EVGTCYFSK	22069	EYILWHWCT
21947	EPTSATWVY	21988	ESYETQLAC	22029	EVHKNDKMH	22070	EYISTQADA
21948	EPVHLPEWM	21989	ESYFSETQP	22030	EVHSMDTQW	22071	EYKYYTYIF
21949	EQFGCGQWL	21990	ESYIVFDNT	22031	EVIQTSDNR	22072	EYMATEHGN
21950	EQKMKQIAQ	21991	ESYKCIRDC	22032	EVKFEMLCQ	22073	EYMHYYGSI
21951	EQKSTQTFH	21992	ETAHYSMRE	22033	EVKHKQDWA	22074	EYPLESTYK
21952	EQLDSHWWK	21993	ETANQEWER	22034	EVKNKGGFC	22075	EYPLPVTLQ
21953	EQLDVKYWH	21994	ETASKNVQT	22035	EVKPCLVVQ	22076	EYSQLHPSP
21954	EQLIEWSSH	21995	ETCILAVDG	22036	EVKTRMCAG	22077	EYTCRLLYD
21955	EQNMLDADC	21996	ETDEIYIVG	22037	EVLKAQREV	22078	EYWIGQQSC
21956	EQQPMASFT	21997	ETDQQYDQY	22038	EVLLSTDEA	22079	EYWQISAWS
21957	EQQYLAQGK	21998	ETEEIVNMA	22039	EVLPYEWCE	22080	FACEIQYSD
21958	EQTATMRWD	21999	ETEGYLFSP	22040	EVMIRKLDM	22081	FAEDCQRFD
21959	EQVYYWNTP	22000	ETEPPDTQG	22041	EVMYSERTW	22082	FALNGIEKW
21960	EQWYDEIPH	22001	ETESMMFCN	22042	EVNMEMQLA	22083	FASRYDGHH
21961	EQYYFNQLR	22002	ETHTYDDCS	22043	EVQGEKCLN	22084	FAVFCWAER
21962	ERFYQHAAK	22003	ETIAMQYRC	22044	EVQQFHRDQ	22085	FAYCESEEQ
21963	ERGYNMAQH	22004	ETKCYAMLG	22045	EVQRWSGYH	22086	FCDACLGGY

[2394]

22087	FCELYFHCA	22095	FEQYIRNIK	22103	FICDVQMGD	22111	FKEELMKYQ
22088	FCLDPNFYY	22096	FERLFMGWT	22104	FIEFYMGEQ	22112	FKFEQENNS
22089	FCTGASHAH	22097	FEVMRAIRC	22105	FIEGEMHIR	22113	FKKPMVCHV
22090	FDDCTWQGP	22098	FFTKISWSW	22106	FIGLQWSYP	22114	FKLVSGDQW
22091	FDHATMQFG	22099	FGTAVMVWE	22107	FIPISHQHP	22115	FKQIDHYPR
22092	FDRCNESYC	22100	FGWDAYCRV	22108	FIRSYNVEC	22116	FMAELKTDD
22093	FEPHGCQIE	22101	FHQYHYQAM	22109	FIRVSWWLA	22117	FMKSKLETV
22094	FEQADMMQM	22102	FHWWGAVVK	22110	FIYTDFRMT		

[2395] 表79

[2396] 驱动乳腺组织嗜性的AAV5 VP1衣壳多肽中581至589区域的序列

[2397]

SEQ ID NO	581-589 序列						
24118	FGDHLRVMY	24144	FWMNCWVIV	24172	YATQMQHLE	24200	MYLRDHNGL
24119	LPTQYQQQN	24145	GQYIAVNYS	24173	YMGSTNIWM	24201	MYLRDHNSL
24120	TKKIIESSS	24146	HYEMPQAMQ	24174	GVGPSHGER	24202	MYLRDHNVL
24121	IMYLKRCCT	24147	KCGNEGRMS	24175	GVMHHGIFD	24203	MYLRDHSGL
24122	TSDQCGNED	24148	LQKGWANCE	24176	AISMYWSSW	24204	MYLRDHTGL
24123	GQFPHYGIF	24149	LSTLITMMD	24177	AVQCYIHPS	24205	MYLRDRNGL
24124	YTELMKKAC	24150	LTDFHPKML	24178	CSTIHAGHP	24206	MYLRGHNGL
24125	LGMHYKDNC	24151	MAMKLIEGN	24179	DIGMYWSSW	24207	MYLRNHNGL
24126	KMGTMDFNH	24152	NFCKQGNTC	24180	DISMCWSSW	24208	NISMYWSSW
24127	QWDDEWDMD	24153	NIRCGNDAS	24181	DISMDWSSW	24209	QCYAIAHLI
24128	THACMYHEY	24154	NQQPGQHSQ	24182	DISMYWGSW	24210	RCCAIAHLI
24129	LEQCHDHCA	24155	NVVCVRYKS	24183	DISMYWSGW	24211	RCYAIADLI
24130	TKVNCVYYG	24156	QATMDAWVH	24184	DISMYWSSL	24212	RCYAIAHLI
24131	QELGVNCAA	24157	QIPKAWNRY	24185	DISMYWSSW	24213	RCYAIAHLL
24132	QKDWGPRVT	24158	QTTTNDSSQ	24186	DMSMYWSSW	24214	RCYAIAHLV
24133	APMTCKEFL	24159	SADRTAMFH	24187	GISMYWSSW	24215	RCYAIAHVI
24134	ASWFPVLFP	24160	SARTDCQCK	24188	KPMINEIVY	24216	RCYAIAHWI
24135	CTYAAMMSY	24161	SDSKVQWGM	24189	LHALINQGS	24217	RCYAIALLI
24136	CVQCECNQN	24162	SEDHFGFWV	24190	LHALVNQGS	24218	RCYAIAPLI
24137	DMFRPDVRC	24163	SHWFFHFDL	24191	LHALVNQGY	24219	RCYAIARLI
24138	DTDNKPHFA	24164	STSWLNDLT	24192	LRALVNQGS	24220	RCYAIGHLI
24139	DTHIFIRHA	24165	TEFSQQYKC	24193	MCLRDHNGL	24221	RCYAISHLI
24140	EFHFKCRTD	24166	VAFEYGERA	24194	MPMINEIVY	24222	RCYAITHLI
24141	EHHAMMLHE	24167	VCGWWHQYW	24195	MYLCDHNGL	24223	RCYAIVHLI
24142	FSHMPTSGG	24168	VHTEMSNID	24196	MYLRDDNGL	24224	RCYAVAHLI
24143	FVCPRAFNM	24169	VTQRPLHHS	24197	MYLRDHDGL	24225	RCYDIAHLI
		24170	WLYLAFAQN	24198	MYLRDHNCL	24226	RCYSIAHLI
		24171	WWKPCNCQY	24199	MYLRDHNDL	24227	RCYTIAHLI

[2398]

24228	RGYAIAHLI	24269	TCTNSEPHP	24310	VHHKTNRSY	24351	HSMMPGWPS
24229	RWYAIAHLI	24270	ECQYEMNDD	24311	RVNGSTCRH	24352	DQEAQVVTG
24230	SIQCYIHPS	24271	THAQMSLRD	24312	NGNHYNAMR	24353	ISLVWMHYP
24231	SVQCCIHPS	24272	KYASMDIYS	24313	WTTSIAVDT	24354	AACNGLEYQ
24232	SVQCYIHPS	24273	AQSKSLMRC	24314	FTEPMCMQG	24355	AADCYHWFR
24233	SVQCYIRPS	24274	TVMQCGTMP	24315	DVKSWAVCD	24356	AAEWDRFAE
24234	SVQCYMHPS	24275	TAFQSLQQM	24316	AYHCYMSWM	24357	AAGLQVDSW
24235	SVQCYVHPS	24276	HMDEVCSKD	24317	RYPCCQSPF	24358	AAMASAFVY
24236	SVQFYIHPS	24277	YVSLCNDGN	24318	AHAWLQHWA	24359	AAMLGWESE
24237	TPMIDEIVY	24278	RVEYVTDSH	24319	HQCNCLLSW	24360	AASPKTMNE
24238	TPMINEIFY	24279	MCQNVEKEY	24320	TNFPTSMHG	24361	ACDRFMAMY
24239	TPMINEIVC	24280	NSTPMACME	24321	QHLMTFSHE	24362	ACSVSICTE
24240	TPMINEIVY	24281	TSEPHCSYA	24322	QRNEMSASR	24363	ACYDYYQQV
24241	TPMINEMVY	24282	DPGWYGLAP	24323	WDGWVTTHF	24364	ACYSYEFEE
24242	TPMINEVVY	24283	CWGPIMHPF	24324	PAHINDQYM	24365	ADCFIFAEY
24243	TPMINGIVY	24284	LSVQAQWDS	24325	VLKSNLTGM	24366	AELEIMWDD
24244	TPMISEIVY	24285	IAKIGNCVW	24326	ATGTYNLQE	24367	AFGVYPSVS
24245	TPMVNEIVY	24286	KAVSYHHDG	24327	LINSIVCGD	24368	AFNHVHEGT
24246	VYLRDHNGL	24287	SLTQYQWRG	24328	ACHTGQSGM	24369	AFTIEDDWW
24247	YSTIHAGHP	24288	MATCYMYAD	24329	DHLSNDAEM	24370	AGVQENVSV
24248	RVTIMFTGT	24289	APFFTFTNH	24330	SECEVLCSL	24371	AILTVCIGA
24249	GTRINTCNG	24290	TRMQAVDHE	24331	VFNLMLQDK	24372	AINMSGSGN
24250	AVGSLRTFQ	24291	FWMQTEHGT	24332	ILINDREWG	24373	AKNDNLCAD
24251	CNTPYEFWA	24292	GCEEVGRCQ	24333	QWFTKGVGE	24374	AKRQKVVLS
24252	DLRLMACCG	24293	IKHITRNTY	24334	NVMGTQAKE	24375	AKWRGTLIN
24253	RRRFGISGL	24294	WKWYGLPHH	24335	YSDPKNMSI	24376	ALCDYMSWH
24254	VTYYSMTQQ	24295	ESTFWQHTF	24336	LEIGGKTRT	24377	AMEYRSWHT
24255	MASRKSMID	24296	SQMKYAHNQ	24337	ASMHEETCL	24378	ANCIRPFTS
24256	GMVELGSCW	24297	AEVPYVMQQ	24338	WERCLELNQ	24379	ANCKQDHCH
24257	TVTTTKWRD	24298	TCGHTRFFI	24339	MPWLDKPPC	24380	ANICTIESD
24258	QHTQQLVQN	24299	NTSGELIMP	24340	ATMNLSWGL	24381	ANIRGRSMK
24259	SGNMYESGE	24300	DRVNLMWKD	24341	WVHPYSYCA	24382	ANQVTGSWK
24260	QDFRMNAGA	24301	AVHTADVCC	24342	QSRCVDNTV	24383	ANTQYPAHM
24261	GLSMVLHFV	24302	EFVLGCQGI	24343	NRIQMTDFQ	24384	ANWPMACDQ
24262	AMTRCKAGH	24303	DVKAFMGGV	24344	SNVACIQAF	24385	APFRKQMPG
24263	GSNRNPIMM	24304	TTSQNSEAM	24345	INYISFVDR	24386	APWHSNDMR
24264	ASAGRNWTA	24305	YASSCQNAM	24346	EVRIQCKID	24387	AQNVLDTVQ
24265	TYWTKQSEV	24306	QSMVACYTI	24347	SVLHVQALA	24388	AQNVLGTVQ
24266	CIWSSPDQP	24307	MGYAKHDSH	24348	CWILHHRCS	24389	AQVSTLNGM
24267	NYLQMCWFN	24308	SSRMQMQGP	24349	KMAGIGCWY	24390	AREDPCAPR
24268	TTYPHNTHG	24309	SQIQIRQAD	24350	LCAVCGDCD	24391	ARGWDYKCP

[2399]

24392	ASERLYYCW	24433	CGWWFNAAE	24474	CSVITQTTG	24515	DDLGAGRSD
24393	ASILQFTVD	24434	CHAEYSHEQ	24475	CTDKLLNAG	24516	DDLSMTNDC
24394	ASKPTRKVT	24435	CHFQYQNQA	24476	CTDRMVKAS	24517	DEEFYPPAW
24395	ASQYQQSLS	24436	CHLLLTDCY	24477	CTENYHQQY	24518	DEFIWLEMH
24396	ATAINQRQA	24437	CHMMDWQCG	24478	CTGQYMGYI	24519	DETVFRNPV
24397	ATEQCLNYL	24438	CHMQFMEAG	24479	CTHQQSIAV	24520	DFEYFSTKD
24398	ATIRQTLLN	24439	CHTEYSHEQ	24480	CTLKQNMPA	24521	DGFQYDAHC
24399	ATTCITSDD	24440	CIDLRLAWD	24481	CTQLSVNMQ	24522	DGVQVDWGR
24400	ATTKGMSGW	24441	CIEWWGMMG	24482	CTQSFVLYQ	24523	DHKEYMDKG
24401	ATTLINQEW	24442	CIIPKMKER	24483	CTVFSGRTA	24524	DHRQWKVLA
24402	AVGHFYTNS	24443	CILHRYVIE	24484	CTVRQWHKP	24525	DICHYIRNE
24403	AVHTADVFC	24444	CILLKKCMA	24485	CVAIDMSVA	24526	DIMRYINTA
24404	AVIFMSVCE	24445	CINADGKYS	24486	CVEYLKQLY	24527	DIRTSVSDS
24405	AVTQQVDEN	24446	CIQAVIGGT	24487	CVGDREPWG	24528	DITQMVVGC
24406	AWCNIDMST	24447	CIRLTEYLG	24488	CVGINLDHH	24529	DIVIKIGMG
24407	AWSCIMSYD	24448	CIRTMNGSW	24489	CVGKQIMAE	24530	DKGCKCQWG
24408	AYENFHWYD	24449	CISRFDHKQ	24490	CVIHPPNFD	24531	DKKPVEDDM
24409	AYHDSGHEW	24450	CKQWTNQAH	24491	CVLYYNQHC	24532	DKVVQQCKD
24410	AYYQLQMSS	24451	CKVHKNWDR	24492	CVQGQYYQE	24533	DLEHKHQFF
24411	CADVPITMA	24452	CLGSTWVAA	24493	CVQHNLVGY	24534	DLRCDPQMW
24412	CAEQWAAML	24453	CMCLKKCYY	24494	CVQTMNGKS	24535	DMDAQGAMW
24413	CAHWTYDCY	24454	CMQQYDNGQ	24495	CVYREIKTA	24536	DMYFTRHTW
24414	CAHWYYEQR	24455	CMREQMNAC	24496	CWAIFRLQS	24537	DMYSMVKHN
24415	CANKEHVCH	24456	CNKDINQSC	24497	CWAQLNTDT	24538	DNFPWCTGN
24416	CANWKDHGY	24457	CNWLFMTQW	24498	CWEMMKTQE	24539	DQCVLSRGH
24417	CAQHRDTAF	24458	CPGPMMQMA	24499	CWNIYNACD	24540	DQEWSVHMG
24418	CASLKVFDS	24459	CPKPIGCGG	24500	CWSSWQKRD	24541	DQFCTYDYT
24419	CASTYTGIN	24460	CQDFRHECA	24501	CYDGWWCPS	24542	DQGPSESDL
24420	CATAMFRYQ	24461	CQQRIIMQA	24502	CYKFVGAAD	24543	DQKVMFECY
24421	CATKSMMYG	24462	CQRKWEKQE	24503	CYRDVRARG	24544	DQQCAQMGV
24422	CAWLSQTAS	24463	CRTELGSMY	24504	CYVYWTTCP	24545	DRQCCKLTF
24423	CCMPLQSFG	24464	CSAKVPMPT	24505	DAAQFGKHT	24546	DSHRLPGNW
24424	CEGFEMRVE	24465	CSAQIKAAC	24506	DACWGHEYA	24547	DSMHYYMMH
24425	CEGWGTYCI	24466	CSCAMYNFH	24507	DAGTNSFCG	24548	DSNYEEKVI
24426	CEHGNENWP	24467	CSDQYQTQN	24508	DAGVTTYGG	24549	DSQYMCVHR
24427	CFGFVNSVN	24468	CSGRLGSLY	24509	DASKVKLHH	24550	DSRQFENQC
24428	CFMWNWHHE	24469	CSHHLMTFG	24510	DASSYLHLC	24551	DSTGRPMMG
24429	CFRMYDLLS	24470	CSHNFQDQK	24511	DCAKFPFMI	24552	DSVAHSSQK
24430	CFTNNVVWT	24471	CSIWEDKSG	24512	DCEDIKKWE	24553	DTDRGLVYE
24431	CGCSEVTPH	24472	CSMECRNEL	24513	DCGQFSIIE	24554	DTHAVEFAA
24432	CGLIMEMCW	24473	CSTTMNVKT	24514	DDHWDWLCY	24555	DTHWWNILC

[2400]

24556	DTYGIQDFQ	24597	EIKYRRCMA	24638	EVMCSDVPK	24679	GFHLEEKYD
24557	DVDFEWSQW	24598	EIRECEECC	24639	EVNCTHLYN	24680	GFKQYATYI
24558	DVEPEYCAG	24599	EITDWGVVP	24640	EYARKTKGW	24681	GGHFDGWQR
24559	DVFIDWSGQ	24600	EITYRRCMA	24641	EYPYRYEFS	24682	GGMWLQNVS
24560	DVIMGPIAN	24601	EIYMRHNFP	24642	EYRWMPNHL	24683	GGWFMGVMG
24561	DVINRYECG	24602	EKAHSWQDP	24643	FACQCKGCS	24684	GHREFQMPY
24562	DVKQQPLAI	24603	EKISICSQY	24644	FAKYMNHGG	24685	GHVAYPPDQ
24563	DVKTDTSPD	24604	EKMESPECI	24645	FDEPMIHQW	24686	GIARMFRWV
24564	DVLLADTEP	24605	EKPYHEEKF	24646	FEYIWVYWY	24687	GIGYKWFPG
24565	DVRDNEPER	24606	ELLFMNAMS	24647	FFCPKHKWP	24688	GLMPVMNYT
24566	DVRQILMCH	24607	EMARNIEQN	24648	FICLRYKDQ	24689	GNATKIKLN
24567	DVWQNSGAL	24608	EMRMNIQWE	24649	FIWVCHTYK	24690	GNRRCWVPF
24568	DWNQVDNYY	24609	ENCEPMHCE	24650	FKMWKLFQA	24691	GQGGNCKSH
24569	DWPAQKGWA	24610	ENCHITKMD	24651	FLELMQIGL	24692	GSHEYWEMA
24570	DWQPMLHMN	24611	ENSYTHDTI	24652	FLRYMVCAH	24693	GSLFQNYQW
24571	DYLAMNLHF	24612	ENTLFMRYH	24653	FMEQDIGHA	24694	GSYMACRMD
24572	DYMHMLTAQ	24613	EPKQEEWYV	24654	FNGSICDTY	24695	GTCNFQMDS
24573	DYMQNQYGT	24614	EPVTFVGQA	24655	FNNNDVHPI	24696	GTDCFWWHE
24574	EAEIFRLEG	24615	EPYFCCCVP	24656	FPSAYQHQH	24697	GTSFVGNPL
24575	EALMPMLGA	24616	ERDHVNISC	24657	FQSMGEIWC	24698	GTYVCQNHD
24576	EASTQWLMQ	24617	ERSWYYDQY	24658	FRGHGMWDA	24699	GVAHPEAVC
24577	EATCDPPYN	24618	ESCEGILTF	24659	FRLCSDVRC	24700	GVTEHNEVC
24578	EATIMDQVV	24619	ESDPMQQFT	24660	FSNLVEVSK	24701	GVVDESYDP
24579	EATKQSWTH	24620	ESEYNSSYN	24661	FSTCTSCCA	24702	GVVGLNTCV
24580	EAYLTIDKL	24621	ESKSCGLTG	24662	FTCPYKFGM	24703	GWCNRYFWD
24581	ECDRYIAEY	24622	ESNLTPWSQ	24663	FTEQIDRQM	24704	GWIRYKNSE
24582	ECGLWWGSE	24623	ESRWMDEHT	24664	FTMASAQAD	24705	GYFGKCKRI
24583	ECKPIAGDK	24624	ESTHKFMQD	24665	FVCLGRKGP	24706	HADMEQGNS
24584	ECPNVNTAQ	24625	ESYSTAEMA	24666	FVGWYWNSP	24707	HADSNIACN
24585	ECVANPFQP	24626	ETEWAWVTQ	24667	FWKTKSGGA	24708	HAESMVCHH
24586	ECYHVLANS	24627	ETLPIEIRS	24668	FWQMKSCLP	24709	HASLNLNPS
24587	EDTHMIAGH	24628	ETSKWHAVC	24669	FWVLDMCQA	24710	HCIWNWSDA
24588	EEFDTVAAP	24629	ETVLHQWAA	24670	FWYPKDRMV	24711	HEAKQVPKA
24589	EFMLSSKDI	24630	EVAMLEEDK	24671	GACERRDAM	24712	HERCSELEY
24590	EFYHFMYKD	24631	EVCLRHYWQ	24672	GADQCMLWE	24713	HEYIYETWP
24591	EGVESMLYA	24632	EVDCKVSHV	24673	GAIYFNKLG	24714	HFYVGSPED
24592	EHDWITVYE	24633	EVDVHDEWV	24674	GCAPLVQDD	24715	HHQYIDCFV
24593	EHVIIKKNY	24634	EVEAIRGIQ	24675	GCAYQAVNP	24716	HIDPDHWHG
24594	EICVVTGEQ	24635	EVEFMSYRW	24676	GCISTLNGM	24717	HKLYEDRIN
24595	EIEMMGRGE	24636	EVISYVENV	24677	GCMKLVGTA	24718	HKNDLAAVQ
24596	EIKVVNSRD	24637	EVLHRDLHM	24678	GEVVPHISW	24719	HKSDMWWTA

[2401]

24720	HLPAPELYL	24761	IQGPEAIMK	24802	KTNMHAKSV	24843	LRYYIEMDR
24721	HMDIQMFMP	24762	IQGVQWMNH	24803	KTNRTYMWR	24844	LSQMLAHTA
24722	HNFCMYWPQ	24763	IREEMHEQF	24804	KVEACQSVE	24845	LSVWAIRND
24723	HNHSADSFA	24764	ISGLAQMGM	24805	KVGAHVAQE	24846	LTHKRESEC
24724	HPELLDRFE	24765	ISICYHMFP	24806	KVGIHDSLV	24847	LTMGFEIGW
24725	HPSSTWSIL	24766	ISMQICAGW	24807	KVKWKDKDL	24848	LTNPWHYFA
24726	HQNKFEWSP	24767	ITCQSNEQH	24808	KVYYIISHQ	24849	LTVIIFDDA
24727	HRNEQLIRA	24768	ITGTRFAEV	24809	KWIMWIEYN	24850	LVAMVPQSY
24728	HSGELYDGF	24769	ITMPAEAFQ	24810	KYHNVPSKI	24851	LVDQGKCMQ
24729	HSMLETTGK	24770	ITTKYQIDH	24811	KYQYYQEGP	24852	LWHWLWLKT
24730	HSNFFHMND	24771	ITTPQLNSV	24812	LAGCLCYTN	24853	LWTRMVYFG
24731	HSYQQCAVG	24772	IVDHLPRHG	24813	LAGCLWYTN	24854	LYAHGNCYP
24732	HTPQVLKGE	24773	IVGLATTAM	24814	LAICHSNLA	24855	LYASVMQSE
24733	HTTAHQNWA	24774	IVHPTLNTP	24815	LANMPNMGF	24856	LYKMVTKLH
24734	HTTFALVMH	24775	IVMREHGAC	24816	LATQACSCN	24857	MAPTVVRKK
24735	HVDSNIACN	24776	IWADLECGR	24817	LCASRGLDP	24858	MATPKEVMQ
24736	HVKSPTFQT	24777	IWDLAGDAM	24818	LCGPIIKAS	24859	MCAGWCDDQ
24737	HVKVHEMMI	24778	IYNQRQMCP	24819	LCIRLWTSA	24860	MCGLRYYWP
24738	HVMPKQGFN	24779	KACDRQIYE	24820	LCPNWWRYQ	24861	MCLWIRACE
24739	HYCKIIPYH	24780	KACEMYASF	24821	LDRKACMIA	24862	MCPNHKPCW
24740	HYEYCFGIF	24781	KADAKVPCC	24822	LEAELMGHC	24863	MESLCVYFY
24741	IAGTYTNWY	24782	KADVRSAQH	24823	LEVDMHASY	24864	MFMKNELIA
24742	IAMDNQNTD	24783	KASNDYLCR	24824	LFEQKMVWS	24865	MGCQSQKPK
24743	IAQDTGTLH	24784	KCDGLWCHF	24825	LFSPYEACC	24866	MGMFNTMAE
24744	ICKIGSRLD	24785	KDRMGLNQC	24826	LFVVTTTNN	24867	MHETHTWGA
24745	IEGHWAQFP	24786	KDVTKLRCP	24827	LGSLYRRRG	24868	MHMSTLDAK
24746	IEIATVNVH	24787	KEPITAEHK	24828	LGYHVQEAG	24869	MHVQQQRYG
24747	IFMVHDSDY	24788	KFAGAGIGW	24829	LHCQMGDHW	24870	MIYKDYNHA
24748	IFVGKNAAW	24789	KGKLEYRIR	24830	LHEIQKPNH	24871	MIYPCMTSC
24749	IGIQFLKDH	24790	KGLDVPHWG	24831	LHSQADGLY	24872	MKHESDGGF
24750	IGTYGWIEA	24791	KHCEILCNW	24832	LHTPQHDCH	24873	MLAGWWQET
24751	IHDASTKDQ	24792	KIAEWEDWM	24833	LKDQFHWGA	24874	MMIVQCSNK
24752	IHMHEVGVY	24793	KICSQLDSP	24834	LKELIQFGN	24875	MMKWTEYMV
24753	IIKDVGDFG	24794	KKFWDTVIC	24835	LKMDMTKRS	24876	MMTLKNTMD
24754	IIPQVDQYY	24795	KKNCREPEQ	24836	LLPPMMMGQ	24877	MNFNSLAQW
24755	IIRDKTQDD	24796	KMQAMGVNW	24837	LNLYWFWSA	24878	MNLPKSHDC
24756	IKCDKCQDQ	24797	KPTKYEWLH	24838	LNPMTQNHE	24879	MNMYYTGEV
24757	IKGRFDWDC	24798	KREAQMVGN	24839	LPLQVSNMA	24880	MNNIYECIP
24758	IMEYVHKPG	24799	KSLIQQSWD	24840	LPRLNLVDR	24881	MNSHNHTST
24759	IQAATNLWQ	24800	KSMVIRHAQ	24841	LQDVIVDCE	24882	MNTPTFHQI
24760	IQGEVTVTA	24801	KTCHPEGMP	24842	LQTPQHHGA	24883	MPTMMGSYM

[2402]

24884	MQFRTNNAN	24925	NKDDTCFQF	24966	PVIWLPVDS	25007	QRVMSGKPL
24885	MQNISCPQC	24926	NKYCIAMPC	24967	PWDKEEPQG	25008	QSDPQPVIW
24886	MQPRYNCEG	24927	NLGETITNA	24968	PYAIEESHD	25009	QSGIHCCCG
24887	MQSFVGVMW	24928	NMIHLWTTE	24969	PYIRRVHVS	25010	QSGWFLEAK
24888	MRALSVSQG	24929	NMNQNPTMS	24970	QAACAMATT	25011	QSHFCFENA
24889	MRTMFYTSP	24930	NMYWCMQIH	24971	QADKQFYGN	25012	QSIPEWSCA
24890	MRVIQGDNY	24931	NNRETMIHG	24972	QAGYQSALY	25013	QSNGQKDSQ
24891	MSFTYVGGS	24932	NNREWDVCY	24973	QANGENLDH	25014	QSVDWIKCE
24892	MSTKPAYWG	24933	NPPDVGCPS	24974	QCDRVWLNP	25015	QTDQYWKMG
24893	MSTKPVYWG	24934	NPQSLCYSC	24975	QCIPKQKDY	25016	QTGWLHSYG
24894	MTDLQLIER	24935	NQHWLLSVP	24976	QCQTLDCCD	25017	QTKTGEIQK
24895	MTEILQITG	24936	NQMPCWVFF	24977	QCWYCELCP	25018	QTKTGKSAT
24896	MTEPFMQDL	24937	NSCFIAENC	24978	QESEVHVFG	25019	QVHTKWDGM
24897	MTIGYFVQA	24938	NSGIATADH	24979	QFQLVTIQG	25020	QVQYRTCDW
24898	MTNRMQALQ	24939	NSIQYGQAP	24980	QGAFHTAWT	25021	QWTEKTFEG
24899	MTRFKIYHE	24940	NSKSPSEGC	24981	QGEKCVPDF	25022	QYEPLGRAM
24900	MTTHCNTAI	24941	NSQPWYRHS	24982	QGIMTSCQT	25023	QYKSTQRRN
24901	MVNCAVTYQ	24942	NSTAHAFAY	24983	QGVRMVGNS	25024	QYYMLSESA
24902	MVSDTQHGK	24943	NSTAHPMMW	24984	QHCMHSMVE	25025	RAKTCIYKC
24903	MVSYDVLGK	24944	NTHVCNLHE	24985	QHQHDLPTG	25026	RAMCILHNQ
24904	MVSYNVLGK	24945	NTLFWDGNP	24986	QIAVVCSQR	25027	RAMPTAAQL
24905	MVTTQQIVC	24946	NVEERNRHF	24987	QIAVVCSQW	25028	RDLPIQFGP
24906	MVVIKQYHT	24947	NVEKQVAMG	24988	QIFKALVGE	25029	RDMDMQMWC
24907	MVVVTRDAS	24948	NVHHTTQIP	24989	QIGWDDAIT	25030	REAKVFELK
24908	MWGEWWPGC	24949	NVMSALDLS	24990	QIISETGWE	25031	REGCYNTQE
24909	MWQEFQWQY	24950	NVTQWDEDL	24991	QKFCNIHEY	25032	RFELLMTGR
24910	MYIPRNGEP	24951	NVVQCNGPH	24992	QKQTSISFS	25033	RGFWVEKSG
24911	NACMVTGCG	24952	NWFGINSWH	24993	QKTWLVVWC	25034	RGSAKDAWG
24912	NAEWQRDTF	24953	PANYSMPNC	24994	QMATPVDMT	25035	RISSEGMWC
24913	NASPCQIFG	24954	PDGYRRSHG	24995	QMGCTFQTP	25036	RLHIDDEDH
24914	NCFVMYCGR	24955	PDWCQDPGM	24996	QMGLLDTYP	25037	RLVRDSKNN
24915	NDKWGWWRE	24956	PEELYLYRK	24997	QMYMDFTQP	25038	RMEELINDK
24916	NDVMWQLIH	24957	PEYRQWMLI	24998	QNYWCMCGA	25039	RMVTWDQSE
24917	NFDFNHQCQ	24958	PFLFTVIHN	24999	QPKEWVIKE	25040	RNMCRHNDI
24918	NFEIYHLQC	24959	PFVHDDMAR	25000	QPRNDCNEH	25041	RPETVHASF
24919	NFGIYHLQC	24960	PGEENEPPA	25001	QQAMYTSSG	25042	RQGCLLAWP
24920	NFNWTRADC	24961	PIWTQHHTM	25002	QQDPMKQWS	25043	RQICNEPWN
24921	NFRTINAHP	24962	PLQQYDGCS	25003	QQEVEQTMG	25044	RRNELLIPD
24922	NGDPVFRMA	24963	PQHCNTMWG	25004	QQEYFDDMQ	25045	RSEMVNTQM
24923	NIEGIEFNA	24964	PQIDSLQDY	25005	QQSAVGAHY	25046	RSQWHNSID
24924	NIVGATEWD	24965	PRNDDNCLP	25006	QQVYEMKWH	25047	RSSEYLTQQ

[2403]

25048	RTLVKGAQC	25066	SFASMYNIK	25084	SQGNTRYSY	25102	SVFSFHSDT
25049	RTQGCGPSR	25067	SFMYKTQAL	25085	SRCLRHQEC	25103	SWDVRVTAQ
25050	RVADVMYIG	25068	SFYKNGMSN	25086	SRMDMNHGL	25104	SWSMQNKST
25051	RVECRPWGE	25069	SGGFCNYAT	25087	SRTMYKWMP	25105	SWVSADEVP
25052	RVMDNACDM	25070	SIGPASNAL	25088	SRTNIEWGG	25106	SYGQIECYQ
25053	RWDYYWTGM	25071	SIHMAEGVE	25089	SRYKCQHRN	25107	SYLRAHHKD
25054	RWRNCGSNG	25072	SIHRWNWMG	25090	SSDQLIMWG	25108	TAALMGNMA
25055	SACIFYAHK	25073	SINPESQGS	25091	SSHATNFCG	25109	TALDNSTGI
25056	SAEMWVKDE	25074	SISPSMGEK	25092	SSHTTNFCG	25110	TAQLMQTAC
25057	SALPCNTAQ	25075	SIVHPPTTS	25093	SSTGMQSST	25111	TAYLNNWVP
25058	SARTEERES	25076	SKGCIIDMQ	25094	SSTTHSRMQ	25112	TAYPEYCGN
25059	SATTMPRTD	25077	SKLLVTRNH	25095	STFPKWTSE	25113	TCAPQGKQE
25060	SCFINSAYP	25078	SKRLVTRNH	25096	STHFRTEDR	25114	TCAVFVFHC
25061	SCGLVQRFM	25079	SLKDVDLWH	25097	STLVYPCGP	25115	TCEEYYNPF
25062	SCGQVEHYC	25080	SLYKHEIWW	25098	STSAQPNFD	25116	TCELMQVMH
25063	SCQMYTYGI	25081	SMFPFNMGP	25099	STTQWGMCN	25117	TCQNIAMQL
25064	SCRLRFNYA	25082	SMTDCDFHQ	25100	STVSQMNYF		
25065	SCTTCHSWQ	25083	SNHPNMHTY	25101	SVCPSHMWM		

[2404] 表80

[2405] 驱动坐骨神经组织嗜性的AAV5 VP1衣壳多肽中581至589区域的序列

[2406]

SEQ ID NO	581-589 序列	28007	KVDAEGFHF	28026	EHGYMRYTK	28045	FPPFYAPCT
27991	EIILDDQQR	28008	VWEDNPPFR	28027	EIECLQTVA	28046	FPPFYTLCT
27992	LDDNAWRHR	28009	SVLMSPHLM	28028	EIEFLQTVA	28047	FTAYPQAVM
27993	LGGKPYALD	28010	LEKNSSQMR	28029	EIEFLQTVS	28048	GAIVFDLCD
27994	GSQQTTLSC	28011	CRCNDAWNM	28030	EIEFLQTVV	28049	GAWDYMYEC
27995	TWESQIDCI	28012	AAPFMWRDL	28031	EIEILQTVA	28050	GCFWAANSC
27996	VVGLKPMHM	28013	ACKYPPHVR	28032	EIEVLQTVA	28051	GCFWAANSG
27997	NFNINIYNC	28014	ANKMFMEQA	28033	EIFLDDQQR	28052	GCFWATNSG
27998	ANHATGGQS	28015	ANKMFMGQA	28034	EIILDAQQR	28053	GIILDDQQR
27999	EFKRMTCNK	28016	ANKMVMEQA	28035	EIILDDEQR	28054	GMPTNHQYI
28000	NKHQGMHRM	28017	AWSEFQWFD	28036	EIILDDQQC	28055	GMPTNRQYI
28001	KGGPWAYNC	28018	CYRCLPPEA	28037	EIILDDQQH	28056	GMSTNRQYI
28002	VLIPIGQHY	28019	DIDRHAHFD	28038	EIILDDQRR	28057	GNLIDWQHI
28003	DPFECTVWP	28020	DKGYSLYQP	28039	EIILDGQQR	28058	GNLIDWRHI
28004	TVQWNFAHM	28021	DRDWSDVRR	28040	EIILDVQQR	28059	HADFERHSQ
28005	TAYMTDMYC	28022	DRGWSDVRR	28041	EIILGDQQR	28060	HADFGRHSQ
28006	SATNWECSN	28023	EHGYMRCTK	28042	ELALGVLWK	28061	HLSFMDLGS
		28024	EHGYMRCTR	28043	FPPFCALCT	28062	HLSFMNLCS
		28025	EHGYMRCTT	28044	FPPFYALCT	28063	HLSFMNLDS

[2407]

28064	HLSFMNLGI	28105	MSFQKTKWS	28146	TFKVNWMDN	28187	CLSMFVGNT
28065	HLSFMNLGS	28106	NCLPWGQFC	28147	TFRFNWMDN	28188	CVKLYVNED
28066	HLSFMSLGS	28107	NCLPWSQFC	28148	TFTFNWMDN	28189	CVKLYVNGD
28067	HLSFVNLGS	28108	NCLSWSQFC	28149	THCPFQHDI	28190	CVRKCDLCG
28068	IALTVTRFP	28109	NLRECTIMN	28150	THCQFPHDI	28191	DMPWVDYMG
28069	IQHWMSHSH	28110	NLREGTIMN	28151	THCQFQHDI	28192	DMTFEAYPT
28070	IQHWMSHSR	28111	NLWECTIMN	28152	THCQFQHDM	28193	DMTFEPYPT
28071	ISHCHFDHK	28112	NRGCSDRIS	28153	TNKMFMEQA	28194	DMTFETYPT
28072	ISHSHFDHK	28113	NWLPWSQFC	28154	VAIVFDLCD	28195	DNDVEMAAK
28073	ISHYHFDHE	28114	NYIEACAMN	28155	VAIVFGLCD	28196	DSDNGPNIF
28074	ISHYHFDHK	28115	PREICWSHS	28156	VAWDCMYEC	28197	EAHPHMMSQ
28075	ISHYHFDHR	28116	PVCSTHFWQ	28157	VAWDDMYEC	28198	EAQPHMMSQ
28076	ISHYHFDQK	28117	PVSMGTKRE	28158	VAWDYMYEC	28199	ECGRYGVFS
28077	ISHYHFGHK	28118	PVSMGTRRE	28159	VAWDYMYEG	28200	EFMEVMDQV
28078	ITQCTMMGV	28119	QCRHEGSNG	28160	VAWGYMYEC	28201	EGGFQDFCE
28079	KLPGSHAER	28120	QCRHEWSNG	28161	WANLRQMDL	28202	EGGFRDFWE
28080	KLPGSHTER	28121	QCYTFNDTK	28162	WGNLRQMDL	28203	EMLRCCDCT
28081	KLPGSRAER	28122	QHNYVSVVE	28163	WSSREYLHK	28204	EQLDVQEVY
28082	KYASVLEQA	28123	QHNYVSVVG	28164	WVNLRQMDL	28205	EVDVSNDNC
28083	KYASVPEQA	28124	QHVGLQAEK	28165	WVNLRQMGL	28206	EVDVSNNNC
28084	LDLKWWCTF	28125	RENTFDQIR	28166	WVNLWQMDL	28207	EVSDGTPTV
28085	MFFQKPKWS	28126	RHGHGHHYD	28167	YYYSSKSSV	28208	FALWSNCAE
28086	MIALRMEGN	28127	RHSHGHHYD	28168	YYYSSTSSV	28209	FALWSNCGE
28087	MIDLRMEGN	28128	RHSHGHHYG	28169	WSSREYWHR	28210	FALWSNCVE
28088	MIDLRMEGT	28129	RHSHGHHYY	28170	WSSRGYWHK	28211	FDCHNASKP
28089	MIDLRMGGN	28130	RHSHGPHYD	28171	WSSREDWHK	28212	FLTEMQWHA
28090	MIDLRVEGN	28131	RLSFMNLGS	28172	MHNPLWFSS	28213	FLTGMQWHA
28091	MIDLWMEGN	28132	SIEVRPPLC	28173	ACVDHMFER	28214	FVADWIWPA
28092	MQEQQKAYE	28133	SIEVWPPLC	28174	AGYTIHTVA	28215	FVADWIWPT
28093	MQEQQKTYA	28134	SKAKYKTQM	28175	AGYTIHTVE	28216	FVADWIWSA
28094	MQEQQKTYE	28135	SMEFEDQCV	28176	AIHLPSACW	28217	GCFTFWDHQ
28095	MQEQQRTYE	28136	SMEFGDQCV	28177	AITMDAQPN	28218	GCLQVNQME
28096	MQQSEQCYP	28137	SMEVEDQCV	28178	ATIDHNSPG	28219	GCNMIPKHD
28097	MQQSEQCYQ	28138	SVEVRPPLC	28179	AVGSEDSSD	28220	GDMPERETG
28098	MQQSEQFYQ	28139	TFCFVQTLV	28180	CAHIMDELG	28221	GHVSTSNQG
28099	MQQSGQCYQ	28140	TFCFVQTRV	28181	CDINVLIMM	28222	GQHSLVTMY
28100	MREQQKTYE	28141	TFKFNWMDN	28182	CDINVLIMV	28223	GVQFRYSMP
28101	MSFQEPKWS	28142	TFKFNWMDS	28183	CDKLDKPAV	28224	GVRVQDSHH
28102	MSFQKPKWS	28143	TFKFNWMDT	28184	CDKLDKPTV	28225	GVTHKAKHE
28103	MSFQKPMWS	28144	TFKFNWMGN	28185	CFAHMQIAN	28226	GVWVQDSHH
28104	MSFQKPRWS	28145	TFKFNWMYN	28186	CISNHNSTY	28227	HSDYEMNDD

[2408]

28228	HYTEGYRQT	28269	QASFKQMGD	28310	MFRQFYIQG	28351	NAIPYFVQA
28229	IADHFLAHF	28270	QASFKQMGG	28311	STVGQNHCG	28352	CVEQYVEQG
28230	IADLCNCEN	28271	QATCYDTSP	28312	EFCHQHSAI	28353	AVAEIRPEP
28231	IGAPVDTIN	28272	QFTHNQAAN	28313	TGWNFPVSC	28354	ASKQGTHDY
28232	IGTAFRNSG	28273	QGIQIGKCQ	28314	GCLTSECAY	28355	SSEYSQFAI
28233	KCESDYLPE	28274	QHVAFNRWH	28315	TMFYCHGPN	28356	TMTVRMEQP
28234	KTNWWDGQC	28275	QILDQQEPP	28316	HITPMITMW	28357	MWGCKLFVC
28235	LAIRKIANH	28276	QINNVVDNR	28317	AITRHVADT	28358	AGMYKCQED
28236	LAIRKITNH	28277	QISMNVAFR	28318	QRSVFHQRS	28359	TYESLQNSW
28237	LARYCEKTP	28278	QMAEFTYYG	28319	VKRLENEWH	28360	TKMIIFDSW
28238	LCMESIMMI	28279	QMFDTWKCW	28320	MMAMKQGFY	28361	LCKFANGAA
28239	LCMESLMMI	28280	QTSSYDAHC	28321	EGRTISYGN	28362	INYISFVDR
28240	LCTESLMMI	28281	QTSSYDTHC	28322	STCYPFAMG	28363	RQAWPKDGM
28241	LDRRSSAYS	28282	QVTGIPWFY	28323	QIIYRCRST	28364	QSAHPEPMC
28242	LFSAWNDFA	28283	QWEDRTQMS	28324	GTAVSRQSK	28365	YCECGHWPN
28243	LFSAWNDFV	28284	RATCYDTSP	28325	GCEEVGRCQ	28366	SQERMDYDG
28244	LMCYMQKIT	28285	REMEYQGER	28326	QHCFRAMED	28367	WINEANCLM
28245	LMGYMQKIT	28286	RPPNRVATG	28327	MNSFYRAEW	28368	SNVACIQAF
28246	LMSTMFNTF	28287	SAMPETTIF	28328	VCHNPDPFT	28369	YTKYNWKAA
28247	LQVHWWLLL	28288	SAMPGTTIF	28329	NGYLVHACQ	28370	NYAMRRDSY
28248	LTIPLHAWS	28289	SQPMLKVNA	28330	NPGWQTIGQ	28371	PVICVWHDN
28249	LYGQVVAID	28290	SSSDLCQNR	28331	LCAVCGDCD	28372	SFKSVSQYN
28250	MAIFDSAAV	28291	TCMPFGTSQ	28332	ATLMSDGGG	28373	ENTMWQSSQ
28251	MFWSGNKQG	28292	TFKLTTGGE	28333	YFPSIQCYE	28374	NAYYDSHCE
28252	MQCDFLHDT	28293	THDDCQWWE	28334	GAANRIQSW	28375	KCADLIIRS
28253	MQKDVEQMW	28294	TIHLPSACW	28335	MFMTNVNSF	28376	EKRLEGWVY
28254	MSMCYGIKG	28295	TIKLTTGGE	28336	VSVQSRAQL	28377	KDGRNPHCL
28255	NAMTQGNHQ	28296	VIMWNYEHY	28337	PLTIEVNCT	28378	QSRCVDNTV
28256	NAMTQINHQ	28297	VSVSIYSGW	28338	KTMGIGGGP	28379	ATMNLSWGL
28257	NAMTQVNHQ	28298	VTMVKRKGT	28339	NRIQMTDFQ	28380	YSDPKNMSI
28258	NAMTQVNHR	28299	YDDWYSNVY	28340	DFNGSNWLP	28381	PDLWYEKSS
28259	NHNPTNRNC	28300	YPMPMCCQP	28341	FAIWKRNES	28382	TFVRKPSLM
28260	NKPCKYMYK	28301	YPMPMCCQS	28342	HFQPVFMQP	28383	PYGPIEYDE
28261	NPTTVCLKM	28302	WCRTPQREN	28343	DQEAQVVTG	28384	EPRLYCAQE
28262	NTHEGTDGN	28303	GMTNEVQAA	28344	SECEVLCSL	28385	DRKQADQIY
28263	NTTPNCNSW	28304	EGGFRDFCE	28345	AAMESCAEI	28386	KGNPLDGDT
28264	NTWPYPENS	28305	RYEISERRC	28346	NQREWHGLA	28387	GLFEGSPGA
28265	NYLQFGHKQ	28306	SAISCQCAK	28347	NYVNQGMGP	28388	DIEFLDCEN
28266	PGTQPYMGW	28307	EVQRNIYSP	28348	NISPYEMAV	28389	EIHMFGQSE
28267	PQEIKSTSC	28308	HCTEGYRQT	28349	ATGTYNLQE	28390	WGGMPNNQG
28268	PRECHQSHQ	28309	SAPEFSVIC	28350	HATGLVNFM	28391	SVLHVQALA

[2409]

28392	ASMHEETCL	28433	CIIVRSTAA	28474	GRAHGNVRI	28515	MESGTLLGV
28393	MHGIARMQQ	28434	CIKYNSSIG	28475	GRVLLMAQD	28516	MFCEFPKCG
28394	EAGCKMWPT	28435	CQSNKQQFC	28476	GSFPCYSLG	28517	MFGTKMSQY
28395	KMAGIGCWY	28436	CVFSWNNYQ	28477	GSFPCYSLS	28518	MFREFPKCG
28396	AVYLDHVFS	28437	CYEITMNQG	28478	GSTEVFYMI	28519	MIDLGNRGF
28397	GHLFNNNEW	28438	DCLTSECAY	28479	GTAVSKQSK	28520	MIDWGGEWG
28398	SNSHPFQSH	28439	DFETWDWFG	28480	GTNHSAEVY	28521	MMLHKQKDY
28399	LSAQRLVCD	28440	DFNHIAWWQ	28481	GTRATKHHV	28522	MMWKENAGW
28400	TVNWYSGFG	28441	DHHGIDTTL	28482	HAPPHGTWA	28523	MSECKNHAY
28401	YIGMKSCSG	28442	DIWLWDTFA	28483	HAVNDHSWM	28524	MSNYHITHE
28402	NHSLCWDSK	28443	DRPNHSTDG	28484	HCKVKECQY	28525	MTAHKENHP
28403	PIWQGYWPW	28444	DSMFIIMNF	28485	HDEWLGHGM	28526	MTAQSPLIP
28404	AELFYDAMY	28445	DSMFTIMNF	28486	HIDRPPCCM	28527	MTDYFKALC
28405	CQVNEEYDW	28446	DVASWCQCP	28487	HWMEQPPYG	28528	NCDHWFGTL
28406	DTEAIMVVD	28447	DYENCFESY	28488	IQCNLCEAW	28529	NDKSWLKHC
28407	ESWRTLCQI	28448	DYGYVSINF	28489	IQYNLCEAW	28530	NETDVGTDV
28408	GHHNTPDMQ	28449	DYTTHWQYA	28490	IWGICVHYF	28531	NQREWHGLS
28409	GWDKKQNCG	28450	EEKYQPCLK	28491	KENTTNRID	28532	NRGYANVAW
28410	IKADVDAME	28451	EFKRMTRNK	28492	KHHMFVDKH	28533	NTFVKNEFW
28411	LTTTVPTKQ	28452	EGTQTHSAT	28493	KKDEWHCYP	28534	NTLTMQSYL
28412	MAEPVKRCK	28453	EIKFPACGL	28494	KMCWWPEAM	28535	NTLTMRSYL
28413	MCHKVMDII	28454	EKLNKTWAG	28495	KQNASDQVH	28536	NTNYWFYYD
28414	MSECKNPAY	28455	ENDHRDYNI	28496	KQVDPLDFW	28537	NVEGTRMRP
28415	TEVEISASF	28456	ENYMCYHSL	28497	KSDFRVLHQ	28538	NVEHSHCVG
28416	TTVAGPHTT	28457	EQAPLDVMF	28498	KSEYVVGVP	28539	PKLCSHQTE
28417	VSFWYDCHH	28458	ETMEHKKRE	28499	KTRFHRERS	28540	PRLQSRKYL
28418	AACYDYKDH	28459	ETMQHKKRA	28500	KVDAEGSHF	28541	PSCLTQESI
28419	AAKLPVAKG	28460	ETMQHKKRE	28501	KVMVVDSVY	28542	PSTTIMNVM
28420	AEFLTYEQK	28461	EVADENGNY	28502	LATMWFCAD	28543	PTFGWMNWF
28421	AFWQDLYYV	28462	EVCMAVVDN	28503	LCYHFWKVN	28544	QCERYHAAD
28422	AGGGSVQCN	28463	FKRICWEAG	28504	LEMHYKDNC	28545	QCRLDMQCP
28423	AHLNCDKSM	28464	FNLSSCYHQ	28505	LFRSHWDAQ	28546	QFEDYWAIN
28424	AMATCQMQH	28465	FTEYEHRAS	28506	LKASLREQW	28547	QGQKQNTTM
28425	AMGRLHDSK	28466	FVAANQNHM	28507	LMENYEIQD	28548	QHFVQGTYP
28426	AMMLTDLML	28467	GAAVSRKGY	28508	LMQKWMRTC	28549	QICPQSMMP
28427	APHWMEISS	28468	GAEQHTNWP	28509	LYLPHCHRI	28550	QKDERHCYP
28428	AVMWTVLHF	28469	GDIRTYPAE	28510	LYLSHCHRI	28551	QMFMQMSGH
28429	CADWACMSL	28470	GFTQNYAQN	28511	LYMWQNQHV	28552	QNSRHKYAW
28430	CAVKQASAM	28471	GLIHSTHTV	28512	MATCYMYVD	28553	QTLGCEMGT
28431	CCWFVYEST	28472	GMTNEDQAA	28513	MCSYFNNRE	28554	QTLGWEMGT
28432	CFTESITDC	28473	GNPHLIEHC	28514	MEFYGRAGN	28555	QTVETYAAH

[2410]

28556	QVGQYQNTI	28597	VTYTNGMAE	28638	HAYQTMSEQ	28679	CVWSVGAEW
28557	RANDGYYIN	28598	VVTVHVRGH	28639	CFTEPITDC	28680	DTAATMEPK
28558	RDGFQEFTP	28599	WQIGGCPEE	28640	MTFVCNFGT	28681	EPRPETEGC
28559	RFTGMSDQD	28600	WTHTYATGR	28641	TSCGYNDNN	28682	MVTERKPWS
28560	RNMSRQEGM	28601	YAEVAVQQE	28642	HPMFPNTED	28683	SYDKEHAVW
28561	RRFVQGTYP	28602	YAMPLYFWE	28643	AFAAYCSRE	28684	RANTWTRRH
28562	RTMGMDWFT	28603	YAPQNEKPE	28644	CAKFTNAIH	28685	HLTILHCPE
28563	RVAMQQMNH	28604	DFKFHSTGS	28645	NHQDNSAYP	28686	YCGWKEFRR
28564	SAWQWAQMN	28605	MMVTKEVNL	28646	IWGIYVHYF	28687	MASFVEHCI
28565	SCAVCGDCD	28606	RTEGWFKYD	28647	YWWGFCWGL	28688	KAMGMLEWA
28566	SCTARAWIC	28607	TSMWGLETS	28648	KVEGYGWTR	28689	EEKDPHMDS
28567	SESQEGVYG	28608	FPGVMEGEI	28649	YDQYGDRTM	28690	MDNNYHQFG
28568	SGALKNREG	28609	TAKQHFMVW	28650	EIGIIRHGH	28691	WIQWVRKMG
28569	SKYHWHMYH	28610	FKMDCHSCG	28651	VDHTLDFYH	28692	NEWSRWLTG
28570	SRDAAAYSA	28611	ATWPERWQN	28652	TVVINACLE	28693	FQNVEIEKF
28571	STEWAALGV	28612	LMENYEFQD	28653	LRHWFFSNM	28694	QNMFGNQPY
28572	STFISSGIR	28613	ESYHYKLMT	28654	AMTRQNGWM	28695	CKQRNCDGN
28573	SYIPFDTQK	28614	QPIHFQQRD	28655	WMCEFSCIQ	28696	LCNNADMRC
28574	TAAGYIVDV	28615	VVNQQEFYY	28656	LTTKDDKEQ	28697	VKQSMPMKG
28575	TAYMIDMYC	28616	AQKITDEVC	28657	PECKGAVDP	28698	MANLKHQLA
28576	TCNVSWKHA	28617	DVRSSMDHC	28658	FQTPVANVC	28699	EDFVTTSEL
28577	TCSYFNNRE	28618	FNLSFCYHQ	28659	QVGQYQNMI	28700	LKRTVGVCQ
28578	TGKVDGMQY	28619	RWHNGKGPM	28660	QNADQAMSG	28701	ALILYARKH
28579	TLPKMYMWP	28620	RYSNVAHHA	28661	RFVLCDYME	28702	ASIKFEGQE
28580	TMRTTMCIK	28621	FAMEQVKCD	28662	HSEMHNHHA	28703	SQKQFSMQH
28581	TMRTTMCIT	28622	DRVMISIQP	28663	TAFANAHTF	28704	CDNAGEFGQ
28582	TPGINRHCC	28623	AGSGSVQCN	28664	KVGCFVDVW	28705	NRCQFGKDW
28583	TTFITETMD	28624	LPTLRVADK	28665	NWRMVHGRW	28706	VTNIFNGPR
28584	TTHQTAFEW	28625	DTDDWTCVF	28666	GVNDLQNIC	28707	GYKCGCVCC
28585	TVNEALTCH	28626	SWHNNWTMD	28667	CVVNRGSIT	28708	ACEQIFEGI
28586	TVRWCHWCL	28627	WHDMVNSGP	28668	DTHTCSGFS	28709	AHGWSQGAS
28587	TWAMHPSHE	28628	RQCTDEECW	28669	DFEMWDWFG	28710	AHGWSQGES
28588	VAWSCDKFA	28629	RVVMQQMNH	28670	TRRSWECAP	28711	AHRWSQGES
28589	VEDMGQVEA	28630	DRLNHSTDG	28671	SCTTRAWIC	28712	AIHEVQSSF
28590	VGQNFNTGR	28631	FLIQGDLTF	28672	MIEPNDHGK	28713	AILKITKVL
28591	VGTDCATSL	28632	FPPIVKRHH	28673	KKDERHCYP	28714	AMINPKNDH
28592	VPSLNTSAG	28633	LCEGQQWHA	28674	CHCQMEAEW	28715	AMINPNNDH
28593	VQKITDEVC	28634	SGSIYHHMC	28675	RVEMQYWEN	28716	AMTNPNNDH
28594	VSDDYECGI	28635	VVMWQFDSQ	28676	DWTINDGWS	28717	ANEEMMAAD
28595	VTQWKYGDT	28636	RPDLNEIEC	28677	TVREDCCIN	28718	APAPTLREW
28596	VTRNEIWFP	28637	SADNKIGGP	28678	SSQKYKDWE	28719	AQSPGIHQE

[2411]

28720	AQSPGMHQE	28761	DLYTYLMEN	28802	FADLILSDG	28843	HVVWRQSSH
28721	ATDQPANLC	28762	DLYTYMMEN	28803	FGDKMTVTS	28844	IAHAAMHDG
28722	ATDQPVNLC	28763	DLYTYVMEN	28804	FSPNTPRSD	28845	IAHEAMHDG
28723	ATGTQLNQD	28764	DMEEHTHSC	28805	FSPNTQRSD	28846	IAHEAMRDG
28724	AVHNVKDVH	28765	DNLCIGAFG	28806	FWSHVGLQE	28847	IAHEAMYDG
28725	CASSYPRDL	28766	DNLCIGASG	28807	FWSPVGLQE	28848	IAHGAMHDG
28726	CASYYPRAL	28767	DNPCIGAFG	28808	GASYFSQPV	28849	ICFQHQIRK
28727	CASYYPRDL	28768	DQRMLLKEL	28809	GCTWEQHAE	28850	ICFQYQIRK
28728	CEFEKMDYG	28769	DQRMLLKGL	28810	GCTWEQHGE	28851	ICMAKNGWW
28729	CEFKKMDYG	28770	DSCSIKCDE	28811	GCTWEQHVE	28852	ICMDKHGWW
28730	CEFKKVDYG	28771	DSCSITCDE	28812	GKSHVDQDK	28853	ICMDKNGWW
28731	CFCKYDMQM	28772	DSCTHHEGF	28813	GKSHVEQDK	28854	ICMDKSGWW
28732	CFCKYEMQM	28773	DSHTANFRG	28814	GKSHVNQDK	28855	IIGFMPPTG
28733	CFCTYDMQM	28774	DSHTVNFRG	28815	GKSPVDQDK	28856	ISNPLMTQM
28734	CGMCHECTL	28775	DSYTHHEGF	28816	GQDCWGLAM	28857	ITTAYPECD
28735	CGTAYWTAY	28776	DYASLAGQA	28817	GRRYPECTT	28858	IVAIAHGEH
28736	CGTTYWTAY	28777	DYASLEGQA	28818	GRWYPECTT	28859	IVASAHGEH
28737	CPKAWEKAA	28778	DYASLEGRA	28819	GSSYFGQPV	28860	KAIPQGSSC
28738	CPKDWAKAA	28779	DYASLGGQA	28820	GSSYFIQPV	28861	KARQDDVNT
28739	CPKDWEKAA	28780	DYASLVGQA	28821	GSSYFNQPV	28862	KHDIYHTYT
28740	CPKDWGKAA	28781	DYGPKDFFA	28822	GSSYFSQPV	28863	KHNPSRHGI
28741	CPKGWEKAA	28782	DYGPKEFFA	28823	GTSYFSQPV	28864	KQVNFNAKN
28742	CQMLLHAFK	28783	EAIFAMTST	28824	GTYGSSFYM	28865	KSDTQCRFF
28743	CQMLLRAFK	28784	EAYAMKQNT	28825	GVAHYHHHP	28866	KSDTQWRFF
28744	CQTETILCG	28785	EDQVASHWI	28826	GVAHYHHHQ	28867	KTVPECESS
28745	CQTKTILCG	28786	EFMCVDQLE	28827	GVAHYHHRQ	28868	KTVQECESS
28746	CQTKTMLCG	28787	EFMCVDQSE	28828	GVAHYQHHQ	28869	KTVQEGESS
28747	CTKECMLWP	28788	EGVIISLLE	28829	GVAHYRHHQ	28870	KTVQEWESS
28748	CTKEYMLWP	28789	EMMQTCVQH	28830	GVAPYHHHQ	28871	LAAMQQGPW
28749	CVECNQPVR	28790	EMMRTCVQH	28831	GVFECSTCQ	28872	LAMRRFDNF
28750	CVMCHACTL	28791	EQICSEIDQ	28832	GYTWEQHVE	28873	LCFQHQIRK
28751	CVMCHECTL	28792	EQVMDTWQL	28833	HAVWRQSSH	28874	LFNKNDCSM
28752	DAGCMGGWF	28793	ESLSGGEHP	28834	HCWGHAFFF	28875	LFNKNGCSM
28753	DCARQHYYA	28794	ESLSNGEHP	28835	HCWGHEFFF	28876	LFNKNGCSV
28754	DICLEPAQT	28795	ESLSSGEHP	28836	HCWGHGFFF	28877	LFRAYAWFQ
28755	DICLETAQT	28796	ESLYSGEHP	28837	HCYSTSEME	28878	LGRKPYALD
28756	DICLETTQT	28797	ESQSSGEHP	28838	HCYSTSETE	28879	LGSDFNARD
28757	DICSETAQT	28798	ETQYWYARQ	28839	HTDDLKKVI	28880	LLTGQNSSH
28758	DKSHVDQDK	28799	EVFLESQDE	28840	HTIITNQPY	28881	LMWIALEAW
28759	DLYTYIMEN	28800	EVFRESQDE	28841	HVFSANVDK	28882	LPWIHTNAN
28760	DLYTYIMGN	28801	EVKYKGNYN	28842	HVFSANVNK	28883	LSHKPADPT

[2412]

28884	LSMQWAVIH	28911	MYQRASDQN	28938	QRVQVGPPA	28965	SYCPVVQWD
28885	LSMQWAVLH	28912	NATSIYDID	28939	QSDTQCRFF	28966	SYCQVVQWD
28886	LSMQWVVLH	28913	NATSIYDVD	28940	RHVQVGPPA	28967	TDPLVMMEH
28887	LTCYHNELS	28914	NFEMQPRAD	28941	RIHYELYIC	28968	TECLCVYNF
28888	LVMRRFDNF	28915	NFMLVNGPL	28942	RPEAVICWN	28969	TLSSGARDM
28889	MADLIVEHD	28916	NNEPYQKFN	28943	RPEDVICWN	28970	TMMFEQNSK
28890	MAEHRKQEH	28917	NNEPYRKFN	28944	SAAAVPSWH	28971	TTDQPANLC
28891	MCHRASDQN	28918	NNFNDAAGL	28945	SAADVPSWH	28972	TVETDHWAK
28892	MCQRASDQN	28919	NNFNDEAGL	28946	SAAVKSEPP	28973	VAEHRKQEH
28893	MDGQAWKEA	28920	NPDYRNEMP	28947	SADVKSEPP	28974	VAEHRQQEH
28894	MGHKPVLSS	28921	NTRPFEKDI	28948	SAHMYHAGY	28975	VDGQAWKEA
28895	MIRADIRLI	28922	PERRIGQWY	28949	SATQNWNEH	28976	VKEQTYAAN
28896	MIRAVIRLI	28923	PGICKANWT	28950	SATVKSEPP	28977	VLAGLECQY
28897	MIREAIRLI	28924	PGIRKANWT	28951	SDTWAMYHH	28978	VNEEMMAAD
28898	MIREDIRLI	28925	PMFPDSMSS	28952	SGTWAMYHH	28979	VTDRNAWVG
28899	MIREGIRLI	28926	PMSPDSMSS	28953	SHATFMMGN	28980	VTEHRKQEH
28900	MIREVIRLI	28927	QCSYCQQVY	28954	SKSHVDQDK	28981	VTGTQLNQD
28901	MIREVMRLI	28928	QHVPVGPPA	28955	SRDWPASMN	28982	VTNRWKRKS
28902	MIRGVIRLI	28929	QHVQVGPPA	28956	SRDWPESMN	28983	VTNRWKRMS
28903	MIRVVIRLI	28930	QIMWFPPEE	28957	SRDWSESMN	28984	VVECASHSG
28904	MLTGQNSSH	28931	QLCSPPMLD	28958	SSPNTQRSD	28985	VVNQFEHWM
28905	MLYICMMAT	28932	QLCYPPMLD	28959	STHIYWICP	28986	WIHSELYIC
28906	MNMCFGQST	28933	QLERTGVHP	28960	STHIYWICS	28987	WIHYELYIC
28907	MSMQWAVLH	28934	QLHYYNVGH	28961	STHVYWICP	28988	WSFSERPGS
28908	MTTVIYDSC	28935	QMFQGLWSH	28962	STVPQMRWT	28989	WTGYSWSVA
28909	MVEMDHWAK	28936	QPERTGVHP	28963	SVSPHTPMH	28990	YAVMSHPAG
28910	MVETDHWAK	28937	QQERFPHSW	28964	SVSQHTPMH		

[2413] 表81

[2414] 驱动骨骼肌组织嗜性的AAV5 VP1衣壳多肽中581至589区域的序列

[2415]

SEQ ID NO	581-589 序列						
30991	ACIFVMLPL	30996	APSKPWLHF	31004	CDPHWDWWE	31012	DDCQFDNEN
30992	ACIFVVLPL	30997	ARLCKRKQA	31005	CDPHWGWCE	31013	DDFGTLPRM
30993	AGKEVSAGM	30998	AWVEHAAFM	31006	CFVHTSCQN	31014	DENIKIIYG
30994	AKKGMRHDQ	30999	AYIFVVLPL	31007	CPWFGLCNW	31015	DENIKVIYG
30995	AKKGMWHDQ	31000	CAWDKHHKM	31008	CPWFSLCNW	31016	DIDPLRNSG
		31001	CAWDKHHKV	31009	CPWFSLWNW	31017	DIDQLRNSC
		31002	CDPHWAWCE	31010	CRTCDEDGA	31018	DIDQLRNSG
		31003	CDPHWDWCE	31011	CRTCDEDVA	31019	DKLKFVKQW

[2416]

31020	DKPKFVKQW	31061	ICNKQEFVA	31102	NAPCMWMPY	31143	TWVEHAAFM
31021	DLRSPGFYA	31062	IDHCGCTQP	31103	NAPCMWMSY	31144	VDLHHPIYR
31022	DMTPKNMQM	31063	IFRCFGVEY	31104	NDFHIEPIH	31145	VVKFSWKPW
31023	DRAMYHKVV	31064	ILARWSEKE	31105	NEDPGAGCN	31146	VVPQPPLET
31024	DRAMYRKVG	31065	IPAFPLRMV	31106	NEDPGAGCY	31147	VVQFSWKPW
31025	DRAMYRKVV	31066	IPEFPLRMV	31107	NGFHIEPIH	31148	WNPHHDFEA
31026	DRTMYRKVV	31067	IPEFPLRRV	31108	NGFRIEPIH	31149	WNPHHDFED
31027	DVYQFWTSF	31068	IVMYPFCST	31109	NLQSMFCEP	31150	WNPHHDFEG
31028	EACYMCRNV	31069	KENLKTNWW	31110	NQPSSDDRQ	31151	WNPHRDFED
31029	EAYDVAAVN	31070	KQEGDYVDM	31111	NVAQPTIHR	31152	YGNCRGHRL
31030	EAYDVATVN	31071	KVKACEGSK	31112	NVTQPTIHR	31153	YGNCWGHRL
31031	EFECFHNGP	31072	KVKSCAGSK	31113	PFECEANEQ	31154	YIDQLRNSG
31032	EFGCFHNGP	31073	KVKSCEESK	31114	PFGCEANEQ	31155	YNPYRFINK
31033	ELNLPVRWY	31074	KVKSCEGAK	31115	PHEQKMLDR	31156	YNPYRFMNK
31034	ERRMEVLNA	31075	KVKSCEGSK	31116	PPCGLKCMN	31157	AENLPRWGM
31035	ETYDVAAVN	31076	KVKSWEGSK	31117	PPCGLKCMS	31158	CIETGILNG
31036	EVECFHNGP	31077	KVRSCEGSK	31118	PQAQEQFHI	31159	CTQTTVRWP
31037	FTDTHEMVL	31078	KWMQDITYF	31119	QFFHTLWKY	31160	EKVHCMWWH
31038	FYPRPELTG	31079	LALNAPNFC	31120	QFSHTLWKY	31161	ETHEMNVCM
31039	GCIQFESTG	31080	LELPDSIKV	31121	QNWHMKNKM	31162	FCPAEWGCL
31040	GDGYLEDSW	31081	LELPVSIKV	31122	QNWHMKNNM	31163	GRHWAFCDY
31041	GMAPKNMQM	31082	LLPLFTPYV	31123	QSTQICHNR	31164	HCSRHFCEV
31042	GMTPKNMQM	31083	LMWIGQVVE	31124	RFWPRKVQP	31165	HQQSEAFVS
31043	GNKDCYGSM	31084	LMWTGQVVE	31125	RFWQRKVQP	31166	IACMGEVQG
31044	GRNFQESLQ	31085	LQAQEQFHI	31126	RKDDQMYAL	31167	KAHQTVQDK
31045	GTMLTTEDL	31086	LRNCICNDH	31127	RKEGCTSHG	31168	KCVSMDFSG
31046	GVWQIHARE	31087	LRTMVQNDH	31128	RSGMARYAY	31169	MYEYGQDWY
31047	GVWQIPARE	31088	LRTMVRNDH	31129	RTMLTTEDL	31170	NAISVGEWD
31048	GVWRIHARE	31089	LTGWFLIWV	31130	SHYFCTMST	31171	NKRNLYALQ
31049	GYTTMQTLH	31090	LVLNAPNFC	31131	SPPCCKKIL	31172	NMVHENRFW
31050	HLDAIWPLW	31091	LWTMVQNDH	31132	SWKLKLCQT	31173	PAWDLVGML
31051	HLDEIWPLW	31092	MAGYNDSCM	31133	SWRLKLCQT	31174	PKGMAKICE
31052	HLDEIWPSW	31093	MAGYNDSWM	31134	TCCLPWQRI	31175	QKKGIEAVD
31053	HLDESWPLW	31094	MAGYNGSCM	31135	TGKEASAGM	31176	RFRFRAQVW
31054	HMQWMGFRG	31095	MAGYSDSCM	31136	TGKEDSAGM	31177	RLERDNNEA
31055	HSRCSHYDW	31096	MIEWQFCTS	31137	TGKEGSAGM	31178	VIKWRSQDK
31056	HTPHDDYRS	31097	MMIGNQMHC	31138	TGKEVSAGM	31179	WLANSWFRE
31057	HYMSMTTYE	31098	MPGYNDSCM	31139	TKKGMRHDQ	31180	YMWWQEGTL
31058	IAYTLQFQW	31099	MRGNCEIKR	31140	TSFSFPDCM	31181	YVVGDAACP
31059	ICNEQEFVA	31100	MTGYNDSCM	31141	TVSCGPADM	31182	KFAMDWVSD
31060	ICNEQEVVA	31101	NAPCMWMPD	31142	TVSCGPTDM	31183	LNELPRAEY

[2417]

31184	ISLQEMMGK	31225	ATGLHITTC	31266	CWTDNDDMD	31307	DSRRIYAWI
31185	VGECVASAP	31226	ATHERDFFV	31267	CYGHLNAMC	31308	DSYCVNHFH
31186	MREMPMVGA	31227	ATHEVITPA	31268	DAEHSPNIM	31309	DTAYVWCEM
31187	AAANYDQIE	31228	ATNWGPMSP	31269	DAKHHQMVH	31310	DTNRCMMQS
31188	AACCPFAIM	31229	ATQEFLGLN	31270	DAKPYQRNI	31311	DTQLDDIAG
31189	AAGMSGQFE	31230	ATVATYINT	31271	DATATGQWS	31312	DTYPMQSAD
31190	AANEKPTHQ	31231	AVGEVDAYS	31272	DAWTMASAM	31313	DVDDLVEIM
31191	AAVQFVLSC	31232	AVIYSSYNW	31273	DCEEFDRTP	31314	DVEDHDSQQ
31192	ACMSRAMFD	31233	AVMYYGYES	31274	DCKWEHHSD	31315	DVIANIHSR
31193	ACQCKVIWQ	31234	AVWIEPRQW	31275	DCPQDEQMN	31316	DVMHNHVKA
31194	ACTWPGVGP	31235	AVYLNEFNM	31276	DDEPKTDWC	31317	DYRQYLEVQ
31195	ADQSWYIKG	31236	AWFCPLLYN	31277	DDMWQRDWF	31318	EADKQYDDQ
31196	AEIYNCEFK	31237	AWSARSVNM	31278	DDPAGHCYV	31319	EAGMKQLCA
31197	AEVEADQRM	31238	AYGWENQFE	31279	DEHKRLDIC	31320	EAIQFPSRG
31198	AFANALACY	31239	AYHELTSCA	31280	DEVPNKCDC	31321	EAQDKWAGA
31199	AFEYFGNPI	31240	CADSKMTST	31281	DFKPKQHDH	31322	EAQSIACPT
31200	AFNRIGRYT	31241	CAHFAMNNH	31282	DFNWQDSGY	31323	EAYMMVCSH
31201	AFYYLGHDS	31242	CAVCPALSA	31283	DFSSINCQS	31324	ECLFLRNVE
31202	AGSTAMQSQ	31243	CCRHWRPTP	31284	DGMCNGGHP	31325	ECNRMAMAQ
31203	AGTSFDMMP	31244	CEHDMCWPT	31285	DGVSDHQRP	31326	ECNVNPACF
31204	AHMENELEM	31245	CEQVQIGDE	31286	DHTDMLCFG	31327	ECSDWTHMQ
31205	AHNYSLAEQ	31246	CFICCHCFQ	31287	DIFPNCKTD	31328	ECWRHVQWN
31206	AHRQLDAEW	31247	CGEHDKFKS	31288	DIFPSMNMS	31329	ECYSIYCEE
31207	AIAWWHEPM	31248	CGGCEDWPM	31289	DIFQLTDDS	31330	EDNIRIMMI
31208	AIDWEDVWE	31249	CHMYHRREG	31290	DIHEWQVMS	31331	EDYHYTHTE
31209	AIEMQRQTD	31250	CHPHDEHKT	31291	DIHGSQEFA	31332	EEAYVNCSV
31210	AIFEPRRWH	31251	CHQENFVFR	31292	DINENMSVE	31333	EFDEFNTRG
31211	AIQLHFEWG	31252	CIASFTDGK	31293	DIVCKFASP	31334	EFESGLLDH
31212	AKGNTNNNA	31253	CKVNESAKN	31294	DLWRTFEAQ	31335	EFHSYCMSF
31213	ALVHKDLYG	31254	CNPRPQMVG	31295	DNHDCGRSH	31336	EFIHIDREC
31214	AMMELMWTD	31255	CNTKFLSQS	31296	DNHGYDRHY	31337	EFKCVDATE
31215	ANHHCGLAC	31256	CQKTFELHG	31297	DNRCCWHMN	31338	EFMEAHCGC
31216	ANHLFVIEH	31257	CQMPERWTN	31298	DNRWWVAFS	31339	EFMPEISTG
31217	APDDENCDF	31258	CSGRGEIEF	31299	DNVIMTTVG	31340	EFRHLKHGG
31218	AQATESSQP	31259	CTCDYTAQY	31300	DPSNCMRIM	31341	EGLENNDMF
31219	AQVEYTYET	31260	CTCEGGSTA	31301	DQEVCTAMD	31342	EHAGCQLWS
31220	AQYPTLEYA	31261	CVLGHIHHG	31302	DQLNVNESG	31343	EHEQLPSCY
31221	ASFFSSVAQ	31262	CVMQRAADL	31303	DQNIFLSID	31344	EHFFDDIHG
31222	ATDALSTSL	31263	CVQPHVDPM	31304	DRSDGGFVG	31345	EHKLVMCFP
31223	ATDNHWGHG	31264	CVTCTNVYG	31305	DSGEGLHNK	31346	EHKWYTTLG
31224	ATENVMWTC	31265	CWSQEANRG	31306	DSLLMREAY	31347	EIHHFEYWS

[2418]

31348	EIIHVVGND	31389	EYGVVKTPP	31430	GLRHDPAQF	31471	HNFYPVTGG
31349	EILKNYTEA	31390	EYMNYEHES	31431	GMAYLQNEG	31472	HPEEFRGFT
31350	EIRPQDTEM	31391	EYSQNHSEV	31432	GMQWMDRQA	31473	HPGTYAIHA
31351	EKAKRISGW	31392	FAAHYGQND	31433	GMVQKQHGR	31474	HPHAGGINK
31352	EKPMMQNEC	31393	FAAPNSYWP	31434	GNVVWLARP	31475	HPVWVRSDH
31353	EKYSIIAHS	31394	FAKTDVWCG	31435	GPDGAKRNK	31476	HQETFFHGP
31354	ELAYLWPPH	31395	FANQGAIHP	31436	GPIGMDWED	31477	HRDPHLLME
31355	ELHSVNYML	31396	FCPEHWFNI	31437	GPISGRYSV	31478	HSFLDNNAD
31356	EMCVQHNTS	31397	FDMFASPDD	31438	GPVVSPCKP	31479	HSGMDHAQY
31357	EMEAKAFQD	31398	FDNSCPKRW	31439	GQFEDLLVL	31480	HSMQGAPLY
31358	EMFYCGAYH	31399	FFALNPNYP	31440	GQGLKAWPC	31481	HSRWFEDYE
31359	EMMASAQSN	31400	FFELSQTVF	31441	GQVAFDKDY	31482	HSTCPCVVR
31360	EMMPLHQQP	31401	FGGKIDIGL	31442	GRHDWMDKE	31483	HSYPTAFYY
31361	ENFSFYKCT	31402	FHCRGYLGQ	31443	GSLFSHFWV	31484	HTGMEHMLY
31362	ENIHYACSN	31403	FHHAMTEWQ	31444	GSYNRLHDN	31485	HTKQMDVDM
31363	ENLVINHIS	31404	FKPVNAPQL	31445	GTDRNYAQC	31486	HVIDHTGIL
31364	ENNKWYSAW	31405	FLESWEKDQ	31446	GTEQVSAEA	31487	HVKINGTGK
31365	ENRCFSDEQ	31406	FLPPNCIEW	31447	GVHHNEHCR	31488	HWYMMEDYV
31366	EPHWFFAES	31407	FLWLMQAIN	31448	GVKFFEYDQ	31489	HYFTKYKVD
31367	EPQATQKAI	31408	FMEQYMMVD	31449	GVMVWNQLR	31490	IAGLFLKTY
31368	EQRIVQQAT	31409	FMGQQNVAC	31450	GVYMEHNFE	31491	IAMISLYNG
31369	ERCESHRAS	31410	FPQHMNVCV	31451	GVYSNCHDG	31492	IANCNTKHL
31370	ERQTCWRGS	31411	FRAQLPTCS	31452	GWYCMRQKI	31493	IAVQQFYAA
31371	ESADMACSQ	31412	FRGIETGQG	31453	GYCGLMQYP	31494	IAYQLTEIT
31372	ESHSYVIAA	31413	FSGCYCMRI	31454	GYTDDGVGY	31495	ICFTPMHDS
31373	ESRCVLEQG	31414	FYGQVTKSC	31455	HADHAMAYF	31496	ICLPKMYFG
31374	ESTNSHWGA	31415	FYQTIGQQE	31456	HASLQCMKT	31497	IDIDMCYDY
31375	ESVDVGLMH	31416	GAFMDMLIS	31457	HCEGGQTIV	31498	IDPPHKMLD
31376	ETEPHTFFS	31417	GAIDYHTQM	31458	HDLSYIDMA	31499	IFGCTSAAE
31377	ETGIDNNAT	31418	GAMLAQLSC	31459	HEDRKHAGE	31500	IGCHRMLNP
31378	ETNLECGQK	31419	GASWNDREH	31460	HEDVGWCSR	31501	IGNCDTFTK
31379	ETQYCRGQQ	31420	GEPELPNNA	31461	HEEFEYWNL	31502	IGVVFQRDF
31380	EVAHKLKHA	31421	GFFIYQNRI	31462	HEHKRSSTE	31503	IHHALQCHK
31381	EVDEVSPSA	31422	GGARALRFN	31463	HEIDGWQMA	31504	IHVTSKTVK
31382	EVGTNPSKN	31423	GGIQLDREW	31464	HFDEITAGT	31505	IIQGPVKAN
31383	EVRVDPCCL	31424	GHRKAFLAT	31465	HFERSPYWI	31506	ILMSMYING
31384	EVSHDAQCT	31425	GIFQYMSKH	31466	HFQPLFEYN	31507	ILVNQQNGQ
31385	EVTEVQTKA	31426	GIGPIMSAM	31467	HGTEVCHWS	31508	IMEITNVHE
31386	EWAMWKPYA	31427	GIIQVRMEW	31468	HHEEHDIFA	31509	IMELTHHNA
31387	EWSDEHLKA	31428	GINQEYTKC	31469	HHSFTIPYQ	31510	INEMARHPK
31388	EWVREDGLP	31429	GITLTELQF	31470	HILHLYICW	31511	INYTYLWAE

[2419]

31512	IPGYAEFCE	31553	LATYGEWHN	31594	LTRPRVQKL	31635	MMEYTHSFE
31513	IPGYLEIHR	31554	LCDYMHMGS	31595	LTSRSDFDL	31636	MNLPDPSVG
31514	IPTYRNRDH	31555	LCEENSVDP	31596	LVDMCRKGK	31637	MNVPWYKDR
31515	IQGVLPAMA	31556	LCFHYMLAQ	31597	LVGQYWPRD	31638	MPCFSVAAA
31516	IQMDFMLWR	31557	LCWGPLHIS	31598	LVGSYQHFQ	31639	MPHPEIRPG
31517	IRPTQRPCA	31558	LCYFGSRWY	31599	LVTRGVSPL	31640	MQEPMRCAI
31518	ISGKSHGDR	31559	LCYRTEKWD	31600	LWDQTQIAI	31641	MSFLFFVFM
31519	ISQWEHSCW	31560	LDSQWMYYN	31601	LYAVENQHW	31642	MSLLMGHES
31520	ISTDCAEFK	31561	LEEAVAAGN	31602	LYCQQAANW	31643	MTEVMREIR
31521	ITETEAALY	31562	LESTSFCTY	31603	LYDHYAAAS	31644	MTIILPMTS
31522	ITKHHMRNC	31563	LEYDWNGWW	31604	LYNWMKLIH	31645	MTKPWSIIE
31523	ITSEGFSFT	31564	LFCIVQSRQ	31605	MADLKWEDS	31646	MTKQTLCHL
31524	IVANGHWLQ	31565	LGWSNYECA	31606	MAFVSKYIM	31647	MTQLCEPWL
31525	IVEQCQRMR	31566	LHIRYQSAP	31607	MAHDIPVPV	31648	MVIQAIWPV
31526	IVFQSLARG	31567	LHNSYWSLA	31608	MAHQLNHNS	31649	MVIQDGHEY
31527	IVTMGITEI	31568	LHWCSQHGH	31609	MATDISGWF	31650	MVMVMEAHN
31528	IYEIMDYLA	31569	LIGFRTMFG	31610	MAWTHCHKP	31651	MVQWEWCDN
31529	IYFMFYCSV	31570	LILGAKFMG	31611	MCLCNMVIQ	31652	MYLEKDLSG
31530	IYPQCALQW	31571	LITWDGPTF	31612	MCSARRQDT	31653	MYQLCHAFA
31531	KATYMEFDT	31572	LKFWASEKG	31613	MCTQPSVKP	31654	NAEFDNNRG
31532	KCEEFAHPQ	31573	LKYQFWTYD	31614	MCYQNIAFP	31655	NAIPWINMD
31533	KCFLCGSAT	31574	LMMPYMGTH	31615	MDEQRQHEE	31656	NAVAQIVSD
31534	KCFWAELWA	31575	LNITLVMGC	31616	MDFPSCEMC	31657	NCCPGFAIA
31535	KEAVLNHWT	31576	LNPHVSQIW	31617	MDHQLCIQY	31658	NCPWPTGEE
31536	KHELVTRMY	31577	LPCKEVCCR	31618	MDPILRCLH	31659	NCRMHNAWF
31537	KHKPRQSPR	31578	LPGTPPDNC	31619	MDPRAHGKV	31660	NDDRCSIEL
31538	KHVCLVSWY	31579	LPKWKEEIY	31620	MEQWDPDEW	31661	NEQWWVWYR
31539	KMEWLAAGP	31580	LPPMKQDEE	31621	MFELFQEGW	31662	NETMMWRAN
31540	KMMAFSEVD	31581	LQHHKYSSM	31622	MFGGFVKMK	31663	NFHYIGTLG
31541	KNNEPPMFF	31582	LRCIDPPCQ	31623	MFIKWRRYW	31664	NFIFMQDMT
31542	KSYCFMCPT	31583	LRNNNCSDI	31624	MFMWWNYFP	31665	NFMDYHDMV
31543	KTCEMFTPG	31584	LRRWDDFQT	31625	MFRISRFMA	31666	NGMLETDHV
31544	KTTILEDCC	31585	LRSIEEKLS	31626	MGILKGEFS	31667	NIYMPHTCL
31545	KTVEWWGRW	31586	LRSLVAEDF	31627	MGTSCCNYE	31668	NMATANTCV
31546	KVVDCGRYQ	31587	LSDRCDVSG	31628	MHLQKHIGC	31669	NMDGKNCRT
31547	KWADAEYGH	31588	LSEVMMQPG	31629	MHYLVMRDA	31670	NNEMPGWRA
31548	KWCYHLFCQ	31589	LSHIKVEQD	31630	MIDPKMGDR	31671	NNMPMDAEH
31549	LANVWVCME	31590	LSKNCEYWE	31631	MIHKKDGMY	31672	NNMSAHHSK
31550	LAPEIECRH	31591	LSTGNIQFM	31632	MIMQCNCSM	31673	NNYHYPHIG
31551	LATSKCSAM	31592	LTKPMPIIL	31633	MKQLSNDGE	31674	NNYLELHRY
31552	LATSMCYQA	31593	LTMISMFSD	31634	MMDLAYHWQ	31675	NPSSHNQLN

[2420]

31676	NQNIWKHNV	31717	PWLAIPRGG	31758	QRQNVQQTW	31799	RWEQMAYSW
31677	NQRPYDQWA	31718	PWQSGSDGD	31759	QSDAILKYY	31800	SAGVVVQSI
31678	NRGCYEMIN	31719	PYASQECIS	31760	QSDLQHWWY	31801	SATMVADKE
31679	NRTFKAECH	31720	PYPAIELGK	31761	QSKELSCMK	31802	SAYMALHPL
31680	NSIFCLENQ	31721	PYSPHKQSA	31762	QSKSGQHIA	31803	SCEWPKQCK
31681	NSNQDQWRI	31722	QANVNGHDF	31763	QTCTSQDYS	31804	SCTHWNTWR
31682	NTCCWKLEN	31723	QCAFIPNSV	31764	QTLESMIYH	31805	SDAPEHDGQ
31683	NTVIRYRQS	31724	QCEHDRQIM	31765	QTMFYEIGG	31806	SDLKIETAY
31684	NVCLEPMDP	31725	QCPVECEHV	31766	QTTQTDCTK	31807	SDQWYRVSF
31685	NVEWQVRGP	31726	QCYRKVVFR	31767	QTVEMSATV	31808	SHMVVTSQG
31686	NVQWADIGG	31727	QDIWPVCRE	31768	QVDEWGGHE	31809	SHVESHQWC
31687	NYKQPVQAP	31728	QEDDFRNKS	31769	QVDNREGDV	31810	SIDVTSCAF
31688	NYNITKFFG	31729	QEKIAMRTN	31770	QVKCMQMWP	31811	SIHPLNQEE
31689	NYWQAQEFY	31730	QFTSQNEGP	31771	QVVHNEFIH	31812	SIRVYPELS
31690	PAEVWTHSL	31731	QFVGRERQV	31772	QVVSTDAPI	31813	SISPIRDMD
31691	PCCQEHCQR	31732	QGFYTHTGY	31773	QVWITAKEE	31814	SIVSIVCPN
31692	PCNFDNQPH	31733	QHMPRHIWH	31774	QYDAFAVLA	31815	SKMAGFPHW
31693	PDAAYWFDG	31734	QHWQFCTNA	31775	QYMDQDVAI	31816	SKWTCHPKM
31694	PDCPYKCYC	31735	QIADQTGIV	31776	RADEDMTSI	31817	SLKMVDVAV
31695	PDGEKCERF	31736	QIGPTISEY	31777	RADHWGSWP	31818	SMFKYEARD
31696	PDNYIQNHH	31737	QIMNYCYFN	31778	RENIHCEGL	31819	SMHEKQHFE
31697	PDQWDFEQY	31738	QISCLEHGL	31779	RHELTMSHY	31820	SNLDFPNQE
31698	PEFMTEMCT	31739	QISYHIQHH	31780	RHEPLHYFH	31821	SNPIMAKGV
31699	PFHMALAIE	31740	QITQHPCEQ	31781	RHWLNGCSN	31822	SNPQSVCRG
31700	PFYHMCAAP	31741	QITSMKCAN	31782	RIDINVDCY	31823	SNPRTQTMV
31701	PGCMIVSQN	31742	QIVQSRSAV	31783	RIELMACTT	31824	SNSPHPAVF
31702	PGELGDENF	31743	QIYNWEKMT	31784	RIKYEHCQL	31825	SNVILDFEN
31703	PHNWRSVSW	31744	QKGNRTNQI	31785	RIMYFDKFT	31826	SNVVKQASD
31704	PMMQKLCCG	31745	QLIVCDTGH	31786	RITYFDDQC	31827	SNYRALYGH
31705	PNNIGGMKP	31746	QMGSRMEMC	31787	RNGFRWTVL	31828	SRDGWPNCI
31706	PPGSHKGIN	31747	QMHSTFCGW	31788	RPEAYWVPV	31829	SREPIIKGH
31707	PPPDRTFAA	31748	QMPIGGCSE	31789	RRDSDSVFK	31830	SSGGIWYFD
31708	PQEFREMCP	31749	QNSHKHWDH	31790	RRFCISLKC	31831	SSMEMWTCN
31709	PQLKFGQLY	31750	QPCVDWPYF	31791	RSAEGISHW	31832	SSRFAPPPE
31710	PQNGMLHYS	31751	QQAPGQMFY	31792	RSHHHQRMM	31833	SSRPPSGWV
31711	PRAENDESR	31752	QQDRLWTTG	31793	RTIEEINTS	31834	SSSQCCTPI
31712	PTITDRFFQ	31753	QQKDQAWRH	31794	RTVEIESWV	31835	STHHFSYGE
31713	PTTNWHYVQ	31754	QQMIRQLSE	31795	RVDDENGGN	31836	SVEAGIKEG
31714	PTYHWDSKT	31755	QQMWGCACC	31796	RVDHDMADG	31837	SVKEEMFSP
31715	PVEHVEKEM	31756	QQNVLEHHF	31797	RVESQACFQ	31838	SVLESDICK
31716	PWHEYAGFT	31757	QQSPVEQCY	31798	RVSVDESDK	31839	SVNDTVDYF

[2421]

31840	SVWDMVKMQ	31878	TSFMCQKFP	31916	VLVNVDQLG	31954	YADMDQAVQ
31841	SWCPEYWEI	31879	TSGPCMPIK	31917	VMPNWKEQT	31955	YAMSRKQTI
31842	SWHYLAMQC	31880	TSQHAIWNA	31918	VMSCYSRDH	31956	YATVEEDLQ
31843	SYHAHQPHC	31881	TSTLVSNNV	31919	VMTFRQKAC	31957	YEHWYEHDD
31844	SYKQVTAPR	31882	TTEFVRIKG	31920	VPGPMIYEF	31958	YERLFDEQT
31845	SYLSFGAPW	31883	TTNDDNCNN	31921	VPKTVNRDM	31959	YFRDQEEGQ
31846	TAAEQKADN	31884	TTSEKNFDD	31922	VQSVRRITG	31960	YFSLLMENC
31847	TAHCTDWQH	31885	TTTRDEAFM	31923	VRMICTSSC	31961	YHSDYHHNN
31848	TASMMECAD	31886	TVCTTKSMG	31924	VRVWDYKGT	31962	YIEPNHHSG
31849	TATTHQRCP	31887	TVRCMIQLG	31925	VSALQQTQQ	31963	YIQRGTDSM
31850	TCPNVWWHM	31888	TVSDDWCWI	31926	VSCHWGLAN	31964	YLVRQWAIM
31851	TCVIRNMMI	31889	TWITTGQFV	31927	VSRLRQNHW	31965	YMSKHEADI
31852	TEAPWWSGH	31890	TWMFVHRNN	31928	VTDHAWTMP	31966	YNNNHLMMC
31853	TEEVDSMAC	31891	TWSDNHAAQ	31929	VTMLMAGMQ	31967	YQQQMTFHN
31854	TEGNWDSCW	31892	TWTNKLCDM	31930	VTSYWEFQE	31968	YQYFPDNKK
31855	TEKAIVYYP	31893	TYEDHVDRF	31931	VTYHHHNFL	31969	YRHGSVDGP
31856	TFGYMATQV	31894	TYIWSMLDR	31932	VVNFVGAEG	31970	YSCNRLTRF
31857	TGFHNTFRD	31895	TYQNELACY	31933	VVWMNSNHD	31971	YTQLNASQW
31858	THCHRHHID	31896	VATMAHYGY	31934	VWFYPQWDS	31972	YTTFQGYWA
31859	THIHSGVYG	31897	VCIGYDHTP	31935	VWSCVHADV	31973	YVAEPYYHG
31860	THKWDADAK	31898	VCPLPNQLH	31936	VYWFPDQMM	31974	YVCNICPTP
31861	TIATVRTCT	31899	VCSQTTLWH	31937	WCNGHGSIF	31975	YVGQTSSSE
31862	TIDLVYLDK	31900	VDKVWWNLI	31938	WDLYLKTMG	31976	YVGTIPTFS
31863	TIIPIDWTE	31901	VETCRDCEG	31939	WEAGECDLM	31977	AAEENEAAD
31864	TILPTCKQA	31902	VFGEFDNHS	31940	WEKFWHEYW	31978	AAGPDALRP
31865	TIPCYQRYT	31903	VFHFQTKGM	31941	WEWPMTWTT	31979	AAGTVCLPM
31866	TISDIPGQP	31904	VFKFDFSYP	31942	WHIFVVIGI	31980	AAGTVEFGH
31867	TKFMPRGYI	31905	VGDPNFRIQ	31943	WIPHCKCPA	31981	AAGWGGATS
31868	TLFYHDWYM	31906	VHAQTIQRE	31944	WKAHTEHLG	31982	AAMEVMVKC
31869	TLSSHWKER	31907	VHHTEPERQ	31945	WKCWYAPLT	31983	AAMHESCNW
31870	TMTKRDDFW	31908	VHWPQCGHQ	31946	WLIWGIIEQ	31984	AAPCCLYYT
31871	TNTAFMNGT	31909	VIHAPTHKT	31947	WMCRNHPFD	31985	ACAWIENFT
31872	TNTSMPAAW	31910	VIPPSIFNN	31948	WQMTTRYYM	31986	ACHATTLRY
31873	TNWREMNPR	31911	VIYHTSYYW	31949	WSSTQDQTN	31987	ACKQDSELV
31874	TPWFWQETA	31912	VKECRRTDQ	31950	WTVGDEQPM	31988	ACMSRAMFN
31875	TQTELYCCP	31913	VKSEWYDEQ	31951	WVEPQTTIY	31989	ACTRPMVLA
31876	TRHCHECLG	31914	VLGDWSNIC	31952	WVEPSSKPP	31990	ACVEHYVAQ
31877	TSDEAEHYT	31915	VLIPILVAQ	31953	WWQTFCYGE		

[2422] 表82

[2423] 驱动皮肤组织嗜性的AAV5 VP1衣壳多肽中581至589区域的序列

[2424]

SEQ ID NO	581-589 序列
33991	NMWYSFVMI
33992	FWLGLFPAK
33993	DSQYHVKCG
33994	AIEWKSTVT
33995	AIEWRSTVT
33996	ANAHQAESF
33997	ANAHQAGSF
33998	ANLEQSLMA
33999	ANLEQSLMT
34000	ANLKQSLMA
34001	ANTHQAESF
34002	ATKNRWPIA
34003	ATKNRWPIV
34004	CCHGWHNGA
34005	CCHGWHNGS
34006	CCHGWYNGS
34007	CECEKLGTS
34008	CEGEKLGTS
34009	CEYEKLGTS
34010	CFHGWHNGS
34011	CIALLEQYM
34012	CNATYQQND
34013	CSCQWYISV
34014	CSCRWYISV
34015	CSYRWYISV
34016	CYLWCAYQH
34017	CYLWCDYQH
34018	CYLWCDYQR
34019	CYLWCDYRH
34020	CYLWCVYQH
34021	CYLWYDYQH
34022	DHPPVKEVS
34023	DNLNFNSPT
34024	DRPPVKEVS
34025	DTYQNYMFM
34026	DVHPSNNVQ
34027	DVHPSNTVQ
34028	EGPWMFAQG
34029	EGPWMFTQG
34030	EYQCWGRFT
34031	EYQCWGWFT
34032	GYELQLQMN
34033	HMYDNELNP
34034	IAEPQWHKQ
34035	ITEPQWHKQ
34036	MCFMAFRFD
34037	MDYCCEHFP
34038	MHYCCEHFP
34039	MLFLKRNIA
34040	MRFMAFRFD
34041	MRFMAFRFG
34042	MRFMAFWFD
34043	PTPPMQHTA
34044	QAAHEPQIC
34045	RHIVHKRTE
34046	RMYDNELNP
34047	RRIVHKRTA
34048	RRIVHKRTE
34049	RRIVHKRTG
34050	RSKCPVEIH
34051	RSKCPVEIR
34052	TIEWKSTVT
34053	TNAHQAESF
34054	TNLEQSLMA
34055	TTHPMQHTA
34056	TTPPMQHTA
34057	WCHGWHNGS
34058	WPAQGRITT
34059	YSCRWYISV
34060	AAVQMSSLS
34061	ACANTDQRG
34062	ACMSILEKE
34063	AEHQNNCAF
34064	AELPEGWMC
34065	AGFYEHRSN
34066	AHGCTNAMS
34067	AIRISSLVT
34068	AIVHHGKEP
34069	AKNYQFQVG
34070	ALDSCMVSE
34071	AMIHRNDHC
34072	ARGMCQDLQ
34073	ASEVHHAEF
34074	ASHCMILHT
34075	ASVDNSGSP
34076	ASWCLNHMA
34077	ATSTRQQHA
34078	AVDPAWHGC
34079	AVSAFQQDN
34080	AVSMKAFAI
34081	CAAHVVLTF
34082	CAKAKKMTA
34083	CAQQIYSQT
34084	CCARIAVAE
34085	CDTESLAML
34086	CEHCCITQF
34087	CEHCCITRF
34088	CFAFQNMDD
34089	CFQSNITAN
34090	CFRSADNWW
34091	CGAHIMSGH
34092	CIQWWEMWN
34093	CKQYIVTDV
34094	CLNRIQTAS
34095	CMMPWQSQS
34096	CMTLILAQE
34097	CMVWTNVSP
34098	CNRSIQCMV
34099	CPGRYWKVG
34100	CPKLIRQFA
34101	CPNVWARPW
34102	CQCRYFDWC
34103	CSAFTAHQS
34104	CTFDCNGYW
34105	CTFSTSEQP
34106	CTSWVWMQG
34107	CVAWHAHSA
34108	CVKQTFTDS
34109	CVLPLFQET
34110	CVMAGAIHR
34111	CVNPMAVAC
34112	CVQYTRIGR
34113	CVSPENGMM
34114	CVWLKQMYE
34115	CWMNTHIGE
34116	CWQYQMVTA
34117	DASPVSRSV
34118	DFVLRSDPA
34119	DGTHMDTWL
34120	DIVGPNTGF
34121	DKGFKTNSL
34122	DKHLFDMCR
34123	DLYTISTRD
34124	DMFMWNTQE
34125	DMQCYKRPG
34126	DMTNCQRCR
34127	DQRWFNQHA
34128	DTYLNVMFE
34129	DVAKKRTAG
34130	DVCRKGEQF
34131	DVPLNECDT
34132	DWRRNTEKQ
34133	DYAVMKIAM
34134	EAEEGMGSW
34135	EARMWHHQW
34136	ECLESGIQT
34137	EGWLMMKTG
34138	EHIHGKFFL
34139	EIERDTSHE

[2425]

34140	EIHTSEMDL	34181	HNMCPCVDY	34222	LQMNYKAQN	34263	NFQMEMDAF
34141	EIKSWHQVY	34182	HSGPFCRHG	34223	LSAYLTEAG	34264	NFSLTNNGQ
34142	EKNSPWQFA	34183	HSSPFCRHG	34224	LSIPKEQCW	34265	NILVLCQNN
34143	ELCWHHGVA	34184	HSSSKEACF	34225	LSMWRWESE	34266	NKLDQDNWK
34144	ENKHHEAFY	34185	HTNWKDGYG	34226	LSNFFGAMQ	34267	NKYCFPTSW
34145	ERMNMIRYA	34186	HTYPAHWLQ	34227	LSSTFPHWH	34268	NLMAHGVAP
34146	ESKIEEAME	34187	HWCMLINVE	34228	LTAKPCTAP	34269	NMPNTHCNG
34147	ESVHDNGGH	34188	IDIDRGHYG	34229	LTGLQEMAA	34270	NQVCWPRIV
34148	ETHHNACSP	34189	IFGCIQSAR	34230	LVQFIYPFE	34271	NRLSFMHYG
34149	ETHRMTWTT	34190	IGFQYGSVE	34231	LVQFTYPFE	34272	NTEFSHLFP
34150	ETHTVDADM	34191	IGVSLLNDP	34232	LWSLRSANF	34273	NVNVNKHNH
34151	EVNIECCNG	34192	IIKESVYCQ	34233	LYLMDIAMC	34274	NVNVNKHNR
34152	EVWTWQDAH	34193	IISMTNRGP	34234	LYLMDITMC	34275	NVRDTMRGF
34153	FITGRHSGG	34194	IMFYTNERV	34235	LYNLFMHTE	34276	NVRVEFIQC
34154	FPCMSWKQA	34195	IMSKTPSFL	34236	MAIVDPRSC	34277	NVVHVGDVK
34155	FVFPKHVTL	34196	INVEGGFEY	34237	MAQKKCKCP	34278	PACNCPKTH
34156	FVVPISAYQ	34197	IPMEWEHAG	34238	MAYCYLMHN	34279	PIWPGVLQH
34157	GACNYMAHM	34198	IRVHDNRQQ	34239	MDGSYHWDQ	34280	PTIKTQFSC
34158	GACVMFVGR	34199	ISVDCMQRI	34240	MFMCYINDN	34281	PTMSLFNEQ
34159	GAQCFIMTP	34200	ITAPKWQEK	34241	MGVEVTQAM	34282	QAKYVTESA
34160	GHCACWVAG	34201	ITVEAVHHC	34242	MIDRHWTAS	34283	QAMKHAFYQ
34161	GHFLTTCSF	34202	IWDVRQDMD	34243	MIRPSINGS	34284	QATIWNNRE
34162	GIGFVDMQL	34203	KFDTFIGYQ	34244	MITLLPNWP	34285	QCCQQPCHN
34163	GIVDCIAAM	34204	KFWCGHMEE	34245	MLFKAQAFF	34286	QCCQTLFQS
34164	GNDKFLVHN	34205	KMTSVSWSV	34246	MLFMAQAFF	34287	QCFDNQDEH
34165	GQVATISQF	34206	KNDLEAFSA	34247	MMCTFDFAG	34288	QEAGECKWC
34166	GSSQIVKAE	34207	KNGDTMAVG	34248	MMEIQESCE	34289	QHQCYYNQQ
34167	GTCPYRAIS	34208	KPKCTYPLA	34249	MPEHGMDNW	34290	QIMCNTTGH
34168	GTFMSSRQS	34209	KTDILDTIV	34250	MPPWYRHHL	34291	QKSWIVGEK
34169	GTTAKTEQI	34210	KTQEFEVQP	34251	MQLEANKYH	34292	QPIFYDRSQ
34170	GTVNFEKMS	34211	KVESTTQLC	34252	MRCIEVKGF	34293	QQDGDQQPL
34171	GVNDYRDAG	34212	KWSYNHFRV	34253	MSDHLSNGA	34294	QQKAISSLP
34172	GYAMKYLNA	34213	LAKCGRWSV	34254	MTHDRAMQC	34295	QSNNALMGN
34173	HAYWCQMFM	34214	LATLLAQIY	34255	MTQCNTEHL	34296	QTFSFAPGY
34174	HCDVIVGDW	34215	LCWFLLGVP	34256	MTTVHWHSS	34297	QVMDLVYPT
34175	HCVPHVGYQ	34216	LENCMWGGN	34257	MVLHEMFVY	34298	RAEPFPEEK
34176	HFMCCWFKR	34217	LFWPDCHFP	34258	MVNPCLLGN	34299	RCEHDKLCL
34177	HHAASATCF	34218	LGLQEWNTP	34259	MVTCAGPTV	34300	RCWFSELHE
34178	HIQGMNMWG	34219	LGMPCAWCP	34260	NAGSSGQGA	34301	REVYGLFMD
34179	HKMRAIWFA	34220	LIYPWKPGA	34261	NCKYRMELE	34302	RGLMCEMMN
34180	HMGMMQCQE	34221	LPEKWNIWK	34262	NCVRYKNMD	34303	RHMFHEAGW

[2426]

34304	RKPPFGPPY	34345	TEWEQNWTP	34386	QKNVSTKYD	34427	FGPCLMNER
34305	RLMSGAIVE	34346	TFGTTAIHN	34387	YCHQEDRIT	34428	FGQLIIKHP
34306	RNKHDMIGC	34347	TGNASNDCW	34388	SQETFQKMD	34429	FKMFNWPAY
34307	RRDCMDSAF	34348	THFSKMSDQ	34389	SRYVNNGTV	34430	FSGVYKQYE
34308	RVAAPWLPE	34349	TKCSWNHIS	34390	DGDFLQNHH	34431	GDSYYMWKA
34309	RVILAMGPY	34350	TLCIVPPLI	34391	ICHTEVMFN	34432	GERGHGERY
34310	SAHQFQDYQ	34351	TMSTCQQML	34392	QTANVKAMY	34433	GFHDCSMMI
34311	SAPQFQDYQ	34352	TNQLYRGFW	34393	SAICQNYDA	34434	GIHEIVVNQ
34312	SDMPHPPLS	34353	TSFMPPSEA	34394	MALTVYWKG	34435	GKCIMQMCM
34313	SGCRYNQEQ	34354	TSGFFHMYY	34395	DTSKTGKHV	34436	GLSMVMHFV
34314	SGLQLMWHH	34355	TSHCIPEVE	34396	TQCKQLTVM	34437	GSMQHRRDC
34315	SHKFMDLMM	34356	TTIGWHVCQ	34397	KTMGIGGGP	34438	GSMRHQRDC
34316	SHQLIRREN	34357	TYCKSYRTR	34398	RVMMGNNQA	34439	HGYEHMMSG
34317	SIRYDFYVP	34358	TYKCWIADN	34399	MNSFYRAEW	34440	HTERWHMLG
34318	SLCSMLCTR	34359	VAVDLSWCG	34400	AAGHIGSVM	34441	HTHDWYSAG
34319	SLCWMHAHH	34360	VFASEWCGS	34401	AAILTDYDE	34442	HVYEHMMSG
34320	SMASVSEIK	34361	VGTKHDVWY	34402	AATGFTFFE	34443	IAGISMHNH
34321	SMEQYWTSN	34362	VKKRDNQST	34403	ACDQQQRES	34444	IDDPTMFEV
34322	SMGVPQAAN	34363	VKKRVNQST	34404	AHAPQFGPC	34445	IHWHCWPKW
34323	SMMPKCKAS	34364	VMKHQHSGR	34405	APEIHDYCA	34446	IIQQYGCST
34324	SMQCEVVGP	34365	VQGYKDSNE	34406	AYMWEEQAT	34447	IMQPPWHTG
34325	SMQCEVVGS	34366	VSCYTHRQD	34407	CSMPSVNSD	34448	KEPMPRGFL
34326	SMQEYQKHI	34367	VTFCVHRYQ	34408	CSSCNCHCS	34449	KICHEDISV
34327	SMYQFADPM	34368	VTTSRWELI	34409	CTKAYRKPG	34450	KQALYDTFK
34328	SQDIESLWP	34369	VYILEGGIH	34410	DCMLVPVSG	34451	KQMPLTNVP
34329	SRMQFHNQE	34370	WFRITDTCG	34411	DEPFPCDWS	34452	LAYPHERDG
34330	SRWCPKVPF	34371	WIITKSTEF	34412	DGRQVWFFQ	34453	LHYDTMKNE
34331	SSGEGKAED	34372	WMTCFPMQN	34413	DHSQNTEPT	34454	LLANYMVHT
34332	SSRPKYFCV	34373	WRDFFNFSH	34414	DKFSLPSQD	34455	LPIEQVKWF
34333	STCHVGKMQ	34374	WTKMQQNDT	34415	DNVTTKKVQ	34456	LWSWYDQPS
34334	STCSIDAQM	34375	WVAVLNLNA	34416	DQCPGHISV	34457	MCWSWQDAY
34335	SVCWSNKVP	34376	WWQLMNLLP	34417	DYCNKKRPV	34458	MQTPQAVMS
34336	SVDRQFVHS	34377	YCGIYWNSF	34418	EAYANMFHG	34459	MSMKYLEQT
34337	SVQLWGYQP	34378	YEHIDWTFC	34419	EDEQHLWND	34460	MTFQDAATY
34338	SVRYECSHG	34379	YLHKRCWPS	34420	EEGTVPEVL	34461	MTIDIDQWC
34339	SYQSCESMQ	34380	YRLLTCASA	34421	ELPRNYPDP	34462	MTVIWIWSH
34340	SYSWKQYGV	34381	YSALNSYAC	34422	EQMPLTNVP	34463	NNVEQHCQV
34341	TALGIYAHK	34382	YSQLFQMTN	34423	EQPSVNYLR	34464	PEYQFECIF
34342	TCIRAMTSC	34383	YTFWINNHA	34424	EQTGDFVHH	34465	PGWSVANCS
34343	TCRPSDWAW	34384	DTSPVEYAH	34425	ERQARNHYA	34466	PKVNLKGDE
34344	TCYQNQRSH	34385	GMVDDKWLY	34426	EVNCDWHWQ	34467	PQVMWPGWF

[2427]

34468	PSTHHTQGM	34509	AQGHKCYCS	34550	IPKYKQFFD	34591	SCEHCNKPY
34469	PVEELYGTF	34510	AQTNKTFMK	34551	ISAQCPHVH	34592	SCHKYTDGN
34470	QFFACGVYD	34511	ASMPVWKHY	34552	ISHWQQRAE	34593	SDPDIACVH
34471	QPTSHMNMN	34512	ATIMMHQFG	34553	ISKAQQHNP	34594	SFCTMQHSV
34472	QTIWFIDDN	34513	AYVEHNASR	34554	KHDLTQQRQ	34595	SHQWSRQMM
34473	RDLMYTDQA	34514	CKTHCPGQY	34555	KHVITGGWW	34596	SNTYAGFQW
34474	RIEWEMISD	34515	CNHQFIDQA	34556	KKPWDDHTR	34597	SRMGLELVN
34475	SCATAHARE	34516	CSCSYQDMM	34557	KLESMAMSM	34598	SSMIMCGYV
34476	SHEKYWYFS	34517	CVNMEMMTA	34558	KLVCVDAYQ	34599	TGTNLHQCQ
34477	SKCIMQMCM	34518	CYEQVNSND	34559	KMNTGDNYH	34600	TMHIKRNSH
34478	SMFMGRRTE	34519	DCADMGYRR	34560	KQPTPVNGI	34601	TMQCISHLS
34479	SPYSCWMQE	34520	DCEWMIRLG	34561	KSARCKLHM	34602	TMRGSNHEP
34480	SRYLYYNYH	34521	DDEYPAHWD	34562	KSLSGEKNY	34603	TSMKTHFDR
34481	SSYYFMGTP	34522	DHLPMNQER	34563	LDMWGDFLL	34604	TSNHENRAD
34482	SVNRMDQDH	34523	DHQWTDAWV	34564	LDQSMDIAP	34605	TWNEPHQNQ
34483	TCELTMCYL	34524	DIGYRSTTF	34565	LQSHSQCFR	34606	VDDQHPTQK
34484	THSSCWYQA	34525	DRAYRPWND	34566	LTWGANAYT	34607	VIFMPIVHC
34485	TIRPVMWGH	34526	DSKSYWMDH	34567	LVCFADIMQ	34608	VLVDHPVNE
34486	TKPPLGCSS	34527	DVGCLQNMY	34568	LYRLVAQMN	34609	WYQSYLNWS
34487	TQTNKTFMK	34528	EAYTNMFHG	34569	MDQLVINSK	34610	YAQHEFRLD
34488	TVCEPKCAA	34529	ECRMYWFED	34570	MIKTVAKKE	34611	YERQMATPG
34489	TVICIPSGS	34530	EMMHKMDAG	34571	MILYFWKSH	34612	YHPLQSDFG
34490	VIWYCFCTL	34531	ESRQVHVGN	34572	MLAPIDCNT	34613	YPHHEQHAT
34491	VMGIVEPSK	34532	ESSYVNGIG	34573	MLGPLWKDQ	34614	ATQFYQACI
34492	VMWNQNKTE	34533	FASYDNTHL	34574	MNTRKDPPL	34615	GFHDRSMMI
34493	VVHNVKDVH	34534	FGEDGGFKG	34575	MTGWSSKQS	34616	NCWVNEKCE
34494	WTSILDHSS	34535	FGPRENNHN	34576	MTYQHVAGA	34617	QGSKQMYQA
34495	WVASSYKPG	34536	FHGNYQHDS	34577	MVCGNQRNN	34618	RTHGHAKAD
34496	WYQPRFGRQ	34537	FLCLPQGHI	34578	MVLADNEFI	34619	TSKSISHMD
34497	YDAVWRKYT	34538	FQKVVEWQA	34579	NNDRWAIPQ	34620	YTQFCAAMW
34498	YDEHWNDNS	34539	FQRPNFSLT	34580	NPGWYGLAP	34621	AQDINRLQA
34499	YDIMTISPF	34540	FQSLNKNQS	34581	NQKGHDYSQ	34622	IAQSEADYW
34500	YHKDYMCLG	34541	FVTQQHSHE	34582	NQVEVIYQS	34623	DAVIHTGFA
34501	YTGRFTCMS	34542	FWNHHDVAR	34583	NTSYVTMCT	34624	DHLNEEQTK
34502	AATTYNQHA	34543	GLCSLNNMP	34584	PAFVCESPQ	34625	LHQFEAYYK
34503	ACEQQQRES	34544	GSMQHQRDC	34585	PKPPLGCSS	34626	KQQMMIVGN
34504	AGDWYIALI	34545	HEWWGPVSH	34586	QCEMKHHLA	34627	YKTSWGQHV
34505	AGHLQFTCP	34546	HKELILQPT	34587	QDPHKCRRS	34628	VDNTDPWIM
34506	AGVRTQCDN	34547	HVQHWMNDR	34588	QPAAKNKWN	34629	ANIPIEMVC
34507	ALCSCNQLF	34548	IADIDLTSF	34589	QQNTEVKWE	34630	TGNHEDPDN
34508	AMRKIGVWH	34549	ICQPDMKQS	34590	QVVVRVQDH	34631	MMAMKQGFY

[2428]

34632	ALTHDDGRT	34673	NGVGRSGDN	34714	CPTYKNKDQ	34755	NSRLLIKHL
34633	AEHCYRHQH	34674	MNLQIGSKG	34715	HVGAGGMPG	34756	NCKPDQIAM
34634	IEFGVKISG	34675	NWTGNMQWA	34716	VTHMTCVEG	34757	TAMVSECKD
34635	KELLCSHDA	34676	EVRIQCKID	34717	AYMWEGQAT	34758	YDYYKMREG
34636	LGDQKLVDE	34677	SNVACIQAF	34718	KCGNVLCEV	34759	SAPEFSVIC
34637	HCTETHHTN	34678	ATMNLSWGL	34719	EHDWRMFHQ	34760	REAWGIRDS
34638	TVGFRFDRA	34679	QSAHPEPMC	34720	GPKEHNLCD	34761	KIETRPTCT
34639	GGAEIVSAD	34680	DHSQNPEPT	34721	ANSLEGGFE	34762	HCTEGYRQT
34640	YIQFEAYAN	34681	EECPNCEAP	34722	YLHERTDKD	34763	MCLNMAKQG
34641	KVHFAKVRI	34682	GWNYMFGSA	34723	VAEAGYCYK	34764	QNPNCKVMW
34642	LIAYKSHAH	34683	HHTMPIDWN	34724	KECCFPPCE	34765	YFHWHDCIG
34643	SPQQPSVFH	34684	GTMPMTGQC	34725	IGEVFDAFS	34766	KYINFDHHQ
34644	PAEAKQRFD	34685	AQAAAVICW	34726	AFDGEMSWE	34767	SFFHQTINP
34645	MKVSNWVFP	34686	DQLAHNAQE	34727	IEMPTESKH	34768	KAGYQEGSC
34646	TIAFNGAVC	34687	INMPCPNTK	34728	YRLHCCNRA	34769	IACISMHNH
34647	MNTNVMQKD	34688	RMWMEQYQA	34729	DAMAYGQFM	34770	QPGVRTDDY
34648	PANHARNPK	34689	IMQMHPASG	34730	ITELMGCLH	34771	QCQNWRHLS
34649	YYTKHEQDP	34690	IPSHMDAHQ	34731	RGMAHHHLC	34772	LIEMCCSSL
34650	RQAWPKDGM	34691	HTDSDCRHS	34732	TGGWQRGAF	34773	PTEFRVCPV
34651	QTAEWYCGA	34692	SYLPRHHKG	34733	FCWSMCEAC	34774	DVCNQSGMS
34652	THVKRFDYE	34693	ASDCDGKTW	34734	FCIKDQHVR	34775	RYEISERRC
34653	EISICTAKF	34694	QIIMLQHEN	34735	PVTHPQTSK	34776	NCERWFQIE
34654	CSECQAIHS	34695	EICHEDISV	34736	RSGVQCCAE	34777	DRREKQLMK
34655	KERMQVSKA	34696	KTTPAPDQG	34737	IWQQCNQKN	34778	ENQARCIFE
34656	FHNAPEHHF	34697	SWYLYYNYH	34738	KQSDKWGDD	34779	ESTHMLYEM
34657	EPRLYCAQE	34698	LSIRNGMCD	34739	MIVHQVSQI	34780	TWHQGWAEY
34658	DIEFLDCEN	34699	QFMYCAKGN	34740	TSNKKGEAA	34781	EGGFRDFCE
34659	DATKEMSWK	34700	HSSCKNCPW	34741	TMFYCHGPN	34782	AFARQEYAN
34660	QQAMWLHSS	34701	DGDYACFKN	34742	LCMNDDRLN	34783	ASSFCKFAS
34661	CADPQERCD	34702	QFLGGPKIQ	34743	DCLKKNHSL	34784	ASSRFGSYG
34662	TTYPHNTHG	34703	RAVHARQYQ	34744	RMDCKWEDL	34785	CHQPKNSDR
34663	AVNQWMHSL	34704	EHTQRMDMA	34745	TGWNFPVSC	34786	CIAKQIARC
34664	CTNYHVKAG	34705	IVTMQMDSW	34746	GSQQMQASQ	34787	CMYHEQPGR
34665	WPNIKHQPI	34706	QEINDLKMH	34747	NSKACRGCT	34788	CPQCDVMTS
34666	KNCIMLASN	34707	YTYLHQSSG	34748	QCLWVDSNV	34789	CSHETGAMC
34667	PLTIEVNCT	34708	DDRRNTKWE	34749	EFCHQHSAI	34790	CSNCRQIMG
34668	ATGTYNLQE	34709	VCCSIGPEH	34750	HQNLLICNE	34791	CSVCLKYEQ
34669	WATWPEPLQ	34710	ALKSFNWCS	34751	FAGFTPVWM	34792	CTLKTQAET
34670	YVDHWTLGD	34711	TNDWGEKNA	34752	NSPHSGVGD	34793	CVKFSHHHA
34671	GLFEGSPGA	34712	CWPEAMIDD	34753	DKTCYLTCP	34794	DANHCMWWQ
34672	SVLHVQALA	34713	GAETEPKDS	34754	CHGFSGCQT	34795	DCVKSSDGI

[2429]

34796	DGCPMTRDL	34837	MIHTPDAGC	34878	SWRWHHWHW	34919	AVVIAQKVV
34797	DGLFEGHIP	34838	MIKFPAKQQ	34879	SYVMNLLGQ	34920	AWELVVHQA
34798	DIKWQHGAH	34839	MMGSNHCSS	34880	TAHIMTDVI	34921	AWNFIRVGN
34799	DMYDHEWLM	34840	MNSFYRAEL	34881	TIDGMMGDH	34922	CAANVTKMG
34800	DPKCFVWSY	34841	MSPFFWEQG	34882	TLPKMHMWP	34923	CAAPTFADK
34801	DWIVMSQKA	34842	MTSTQIRAG	34883	TQQYTWRDA	34924	CAEHLTPSE
34802	EFNWLLWSG	34843	MYMEQACYD	34884	TSYLRKIHS	34925	CAGLLGFHW
34803	EIAHDTSMS	34844	NADEKYNNT	34885	TTAGKFAAQ	34926	CAGVQTTWW
34804	EIGHKQMAI	34845	NAPMMACQI	34886	TWARYKVGE	34927	CAHYFAHAG
34805	EIVQPQYYA	34846	NGLFEGHIP	34887	TYDTFGCME	34928	CAKWFDVQI
34806	EMHKVPSGM	34847	NWDRGDKCI	34888	TYRTRHTDC	34929	CAQSWCSFE
34807	EVTAVEHIW	34848	QCMVMTEAD	34889	VKHLFYWLP	34930	CCFLINWDR
34808	EWTKKTADH	34849	QDFRMNAGA	34890	VLKSCDRHG	34931	CCGWYHEFG
34809	FLHISSWYQ	34850	QDFWMNAGA	34891	VVKHGVTTN	34932	CCMEVFMYN
34810	FNMNTNDSG	34851	QILPHMYAY	34892	WGATSEFMG	34933	CFTMPVHWG
34811	GCKKSEWIT	34852	QNNVITVDW	34893	WHCVAPNSF	34934	CHIQLHAGI
34812	GGSMLGNGY	34853	QPFLHRPCA	34894	WNDLVRCEE	34935	CIHYKEQYA
34813	GNKLKNDQM	34854	QPFLHRPCV	34895	WWQPVEYAW	34936	CIHYKEQYE
34814	HCLAGESYV	34855	QQPRKINLC	34896	YHAPMLMCN	34937	CIIHQSKYG
34815	HFHEVAWGS	34856	QTMCITIHP	34897	YLAMMMSGD	34938	CIMSRQQDC
34816	HPETMIYGW	34857	QWDEFKMRM	34898	YLAMTMSGD	34939	CIRRDDHHH
34817	ICNAMLGSE	34858	QYQTHPMMH	34899	YQNHYQKKW	34940	CIRWFDWLG
34818	IFNHLLPTE	34859	QYWWPATHD	34900	YSVVMRTSY	34941	CIWYTQMGT
34819	IRVMFGKHN	34860	RCEPFNWRE	34901	YWDACYDQK	34942	CKETVETVY
34820	KAIWQGTEN	34861	REVHVNQNN	34902	DADCYVPAD	34943	CKMQLFSHG
34821	KFICTTIAC	34862	RGSQQMETH	34903	EGKVPGQFA	34944	CLGWCTACP
34822	KICPTEYAS	34863	RKKVEVVVE	34904	ACDLVKKDQ	34945	CLTALMKLS
34823	KIVNQENMK	34864	RLKLSAVQR	34905	ACEAYQDEK	34946	CLVRYFRDG
34824	KNTMCDNDA	34865	RSTWGHIQR	34906	AFSQQDYSF	34947	CMTKTSKMS
34825	KSDKVWYVL	34866	RVNWEQRMN	34907	AGQLNFCSI	34948	CQCSIMSIE
34826	KTQTHTSEP	34867	RYQPFSRQW	34908	AHNVRCATA	34949	CRGHIWSMD
34827	KWEQRSKMS	34868	SALAHHANF	34909	ALCWKQMDH	34950	CSACIASAT
34828	LMAQWEQTG	34869	SALAHRANF	34910	AMCWYYHMQ	34951	CSLKGPPTH
34829	LMAYCYANN	34870	SAMCSGSLL	34911	ANVISPQSN	34952	CSSSLFSAC
34830	LMHGVMDHV	34871	SANQAHGEF	34912	ASCNTHHEV	34953	CSTPITAAV
34831	LTQKNLSWQ	34872	SANQRYHDW	34913	ASGTKYRCN	34954	CSTPITTAV
34832	LWPQVMEQR	34873	SAVQEERYN	34914	ASTMFIDVE	34955	CTCAGVVTC
34833	MAGHYFNSR	34874	SAVQKERYN	34915	ATHNVHMFH	34956	CTNLYAMAC
34834	MAIDNLECK	34875	SDHQYRGCL	34916	ATMAYRHVN	34957	CTQHMYVKD
34835	MAIISQIEF	34876	SMTTWWMFP	34917	AVDCSDHSY	34958	CTWKMMWGG
34836	MATISQIEF	34877	SNSHPFQSH	34918	AVEQWVNHD	34959	CTWTTGEAS

[2430]

34960	CVCLAEYVV	34968	CWMGIKRQD	34976	DIHNYLMSE	34984	DTCRRLINE
34961	CVFYLPHSP	34969	CWMLQHKMM	34977	DISMNQDVN	34985	DVDWANSLL
34962	CVMSRWNYD	34970	CWSINYNYY	34978	DKRCNKWDP	34986	DVRLIGHEA
34963	CVQHFHWDK	34971	CWTIRHGGP	34979	DMRLQDHVR	34987	DYARCQSVW
34964	CVRFLNRVA	34972	CWYLWYNMP	34980	DNFFTQYLA	34988	EAEPNGFGL
34965	CVTRCMARD	34973	CYYMLQAAD	34981	DQESLNAHY	34989	EAGDVRRFG
34966	CVVYHNHHT	34974	DDDTENDFM	34982	DREQNEMLG	34990	EAGPHSDTW
34967	CWAQMNHVF	34975	DGGQTHVSY	34983	DSSSSIGHW		

[2431] 表83

[2432] 驱动脊髓组织嗜性的AAV5 VP1衣壳多肽中581至589区域的序列

[2433]

SEQ ID NO	581-589 序列						
36438	GCLQIERRH	36464	KHMPTLQWQ	36492	ADQHTSPTS	36520	MMLSDDRTP
36439	EVTRNHHFN	36465	KSTEQMHRA	36493	AIILLCHAE	36521	NIDQWKHQQ
36440	CCKPQCKCQ	36466	KVNCRMIIM	36494	ATRCWWREI	36522	QISMAHFLP
36441	MVFLEQHGR	36467	LNLARYEAP	36495	CIIARSTAA	36523	QMEQWHRGG
36442	YITCQYSHH	36468	MGMWKVWAQ	36496	CNIHWTMAE	36524	QNWAALACP
36443	ILVPYYTGS	36469	MTGDSGKFR	36497	CNNRLDVFN	36525	QVSVSVIAW
36444	AWSDLMNCW	36470	NIHAVFVQS	36498	CVHERANAN	36526	SITEQPLAR
36445	MIEPYTSYN	36471	NIHDVFVQS	36499	DIRPNMFWS	36527	SMAICMHDN
36446	NWYTTWIDE	36472	NITLVMGQH	36500	DNAVNSHVF	36528	SMQSAEPSC
36447	MMKRNRFIC	36473	QCCVTYVNS	36501	DNYYYWGSL	36529	STHQSEFTT
36448	ACVERLHHK	36474	QEFYCWIDG	36502	EDFVTTSEL	36530	STVYRWDMP
36449	ECRMAVQSC	36475	QNEGRGSIV	36503	ETANGMCST	36531	SVMLKANEV
36450	APFFTFTNH	36476	RTAMFESRL	36504	EVKQGFTQG	36532	SVYHCWRSQ
36451	ASLVKWDFE	36477	SHARKVYDC	36505	EWHYRNKDA	36533	TAEVRWPCP
36452	ATMWGDCDY	36478	SIYQKWVTN	36506	GEPDHSPAN	36534	TASFAQRRS
36453	AWKMAHNSF	36479	SKCPWFHNG	36507	GIKSCWQNS	36535	TDDPPLLTY
36454	CHHDVNVPI	36480	SKNIRNYDI	36508	GITEQPLAR	36536	TFCESMPWV
36455	CIMLNGCFQ	36481	SQCVMQFPL	36509	GSKGMAFCS	36537	TIRYIGMSG
36456	DLCAGNKSQ	36482	SYARKVYDC	36510	GVNEFLDTN	36538	TKKQMEWQM
36457	DLCEGNKSQ	36483	TAALGHSFC	36511	HMGYVNNQS	36539	TQQTGWVYN
36458	DLCGGNKSQ	36484	TQGMYVESY	36512	HQHEAPMDI	36540	TTYIPASAA
36459	DPMKIFVSH	36485	TTLLVQAEW	36513	HTTPCARSR	36541	VCHNPDPFT
36460	EKEWRPLQW	36486	VIIYPQFKC	36514	KSDGQMRQS	36542	VSSDKQVLE
36461	ESNGMYFDR	36487	YGGMEFEDC	36515	KTYSAMYSP	36543	WECQWDMSM
36462	GNCVYHCYE	36488	YGGVEFEDC	36516	LFYMKEDMR	36544	WFCYLSPMI
36463	HITCQYSHH	36489	YQTLCMHQD	36517	LPTLRVADK	36545	WFEPSNLIK
		36490	YSKVFYFFS	36518	MHMSYHWHE	36546	LGMHYKDNC
		36491	YTQAMCITA	36519	MHMSYHWRE	36547	ATGTYDLQE

[2434]

36548	APKAPKMYH	36589	CMEAQGECY	36630	NVCDNGHEI	36671	CSYIESWQI
36549	DMVTTRSPF	36590	ISLVWMHYP	36631	HFQPVFMQP	36672	CTGHFSCQQ
36550	DQVNKTYVI	36591	NNGPIWISS	36632	HMEQCWCNE	36673	CTYTMATNR
36551	DWSVWIMET	36592	AHLHNQFMP	36633	LPPPEEKWL	36674	DAHKFARYP
36552	FLTTVLRVC	36593	RTEAHLPAP	36634	ESAEPRMIP	36675	DAKSPCDSF
36553	HLTTIPGDR	36594	EGASPLACT	36635	ICIRGARAD	36676	DAVNDLTCP
36554	IFYECYSNK	36595	DANHCMWWQ	36636	KKHKGWCRA	36677	DDTEMSYCY
36555	MYIFVWRFM	36596	EKYQIHWDR	36637	AEVPYVMQQ	36678	DFWYSMAFP
36556	PEKVWSDLK	36597	KQEWHVHTE	36638	MAEQMERDF	36679	DGAWVKMMY
36557	SWFYMEKMP	36598	KLYGLNHSC	36639	ITASIKDQW	36680	DQAELLSDH
36558	TWLYGWPIK	36599	QSSSSYLWG	36640	ETDCMEGGH	36681	DQFRALDLR
36559	VMPIAWQQD	36600	DIRIVSDGK	36641	QFAQLWCAK	36682	DVASYCQLQ
36560	VQKCHTCGA	36601	MALPLHVLS	36642	KACMSHQGR	36683	EAGMWPSKA
36561	VQKCRTCGA	36602	QTMCITIHP	36643	GLFEGSPGA	36684	EAWMVMNDQ
36562	WMLLGMTWC	36603	RHWYIALYA	36644	AYHCYMSWM	36685	ECRSNDYDR
36563	FTNTKEQAF	36604	LSAKQQFFC	36645	GQQTIEMSY	36686	EDVFYQCNH
36564	YSWFYWADY	36605	GEGVEQMDG	36646	ATGTYNLQE	36687	EFRRNELGA
36565	VRWVRLHKK	36606	RQIGTGNME	36647	GHLFNNNEW	36688	EFWVCTRNH
36566	VCYWWLNWF	36607	IMTIIPYGP	36648	FAQQIINAS	36689	EGEARCHWH
36567	THKWFCEDS	36608	VNDDYECGI	36649	AGMYKCQED	36690	EHFRAYTQY
36568	DPYSSKYES	36609	NTQYSAKND	36650	LEIGGKTRT	36691	EIKCSGGYA
36569	EDWLRRNIW	36610	SRTICVHGN	36651	NWTGNMQWA	36692	EKRESFGCG
36570	KHYKWGHCQ	36611	KHIVGQDGG	36652	DIEFLDCEN	36693	ENEMYMKQD
36571	ESVPLRETT	36612	KCLPYMEMF	36653	KMAGIGCWY	36694	EQDRGPMQC
36572	SSSPGKGWM	36613	KAYMMMEGP	36654	QCTAHNTCM	36695	ESMKFRTFS
36573	GSTSRNQRD	36614	SVTYHEWHP	36655	LHGIINQDN	36696	FASAFANSD
36574	AHTQYQGQN	36615	YSDMFHKDH	36656	AAGREYWLY	36697	FMHSYEKFD
36575	KFDVIRQAM	36616	GNECFHNVE	36657	AAVPVNKLT	36698	FNIAESTPM
36576	GTIPKQHEF	36617	IAKIGNCVW	36658	ADQSYNEYY	36699	FNPNFVTKT
36577	QKAIVTQND	36618	SGNQIVQAP	36659	AELCNNIIK	36700	FNVFIDRSN
36578	MDYHCQVED	36619	ASFFLNPLE	36660	AEYTHDCAM	36701	FQRRTQLIT
36579	ERQSTHGQY	36620	DMKQKAEYT	36661	AGHVFWDGE	36702	FTVFYRLQW
36580	KCWDCHEAG	36621	MMEATHKLS	36662	AIGPSWMGV	36703	FTVGGLAFA
36581	NHDEKLWAP	36622	GCEEVGRCQ	36663	AIRPHHAKS	36704	GDEYYDTMG
36582	DGMHRELYS	36623	CIHPVLVTM	36664	AQKMEDQSL	36705	GDLSRDQED
36583	AKMWQLENP	36624	YDWGRVTLN	36665	AQKMKDQSL	36706	GFRTQDMIE
36584	LTNRTQESW	36625	QKAIVAQND	36666	ATEMFVQHD	36707	GPCMETQHH
36585	STALMDANM	36626	TFTYEAKFW	36667	ATIAQWCIH	36708	GSEQKIHIN
36586	KQAGYDFFH	36627	ISWISHASP	36668	AYQWFMSQP	36709	GTQSYLVKN
36587	SPQQPSVFH	36628	PCQQLMCFM	36669	CDSIYHHMP	36710	HATFTWRCL
36588	TTYPHNTHG	36629	VSVQSRAQL	36670	CHFKYPRIE	36711	HFWITEQMD

[2435]

36712	HGCVETNEH	36753	NAMLWRMHN	36794	RTEACTSLP	36835	TWHNPRNDK
36713	HMLKMEMFN	36754	NCCQEHNVG	36795	RTQWCIADP	36836	VCQYCELRG
36714	HNDQYEHDE	36755	NEGKWFWAS	36796	RTYYEENGR	36837	VRMVPESEW
36715	HVDNDNMHE	36756	NHDEKLWVP	36797	SATDMTMLT	36838	VTNWLSLQY
36716	HVYYCPEAI	36757	NHPVGWDND	36798	SDYEKFLAY	36839	VVLPWGLRE
36717	IACPMGIHD	36758	NHVTMCVLR	36799	SEHWHEDAW	36840	WKFMFEKWR
36718	IAESYRMMH	36759	NIVDCDDVR	36800	SGNSQNAFM	36841	WKTVCPSRQ
36719	IDSWMMIPS	36760	NLGWQEWDS	36801	SHQHMWHQT	36842	WLAPCAKCG
36720	IGEPQAYMT	36761	NNPFSRYMP	36802	SIETMDTNA	36843	WNCSVYTCS
36721	INHGVNHQW	36762	NSMWLCTNQ	36803	SILNAMSLQ	36844	WSPLYQTVR
36722	ISMRSECKD	36763	NSVMMTMIA	36804	SNYFNNFGF	36845	YIHTVENND
36723	ISWITRADI	36764	NSVVMTMIA	36805	SNYFNNFWF	36846	YQNHYQKKW
36724	IVNLQHYHG	36765	NTIKVEDLG	36806	SPPNLRAFP	36847	IKHIATESR
36725	IVSTSGNND	36766	NWGDQRRHV	36807	SPSIDRFWV	36848	VMGTGRMQH
36726	IWEDMKGKF	36767	NYAQYYRAW	36808	SPYWVMRMG	36849	NWGDQRQHV
36727	KAHLVLAGD	36768	NYCTYHSTH	36809	SQKVHIDNN	36850	LVFLRDQHY
36728	KAMTDAANW	36769	PADHKQTSV	36810	STHYASTYL	36851	CRWTPEIFQ
36729	KCAITETSD	36770	PFCESACHI	36811	SVADHYHLK	36852	KRIKMPPPK
36730	KFTTAMTEF	36771	PGDRTISMW	36812	SVAMQMHSW	36853	SWYYDEQSN
36731	KIVESHGTN	36772	PGMMYQFAW	36813	SVCWIWWTF	36854	NNYKVQQDG
36732	KKTEEMPVT	36773	PHVCGEGKI	36814	SVKPHYSCW	36855	HHQLHLKDS
36733	KPCTQAAEP	36774	PKGWADPCP	36815	SVVYEAIWA	36856	QDDVLDVTE
36734	KSYSCEQTS	36775	PNWGLDMAP	36816	SWHAEVVNQ	36857	VKNHYEGGP
36735	KTHVMHWEV	36776	PTCIIMKGN	36817	TAIHMNWRQ	36858	KPSQCLPHN
36736	KTYSVMYSP	36777	PTFHRIGSY	36818	TAPLVYPHC	36859	DYDPDIKDQ
36737	KVCNHELFG	36778	QAARMQKYQ	36819	TASDAENYQ	36860	CDSIYRHMP
36738	LAAMSLTWW	36779	QAVSAVDKC	36820	TCDRSLAEY	36861	GAMMGNSLP
36739	LDGVFFEFF	36780	QDLEWIAGC	36821	TCDWSLAEY	36862	CTQGYGDMP
36740	LFRSHTCHH	36781	QHHPINCAT	36822	TCGYTHIFS	36863	MIYTGWDMM
36741	LRTQLLLQD	36782	QHMVTSHNP	36823	TCQHNQHVA	36864	IGKEFIVMP
36742	LSDSYVRKF	36783	QHNTSYPSF	36824	TCSPDEQLD	36865	TVNTYEHCW
36743	LVKGREQFN	36784	QPTLPDEHE	36825	TEALAQFMA	36866	ECASVDISY
36744	LVQCHHVHC	36785	QTVPPMVQE	36826	TEKEFMYIC	36867	VNSCRHAWY
36745	MAMWARTEP	36786	QWIVFTCII	36827	THIGSYTQF	36868	AVFAPNGEV
36746	MDSQHNKWQ	36787	QWSDIMNCD	36828	THQYWYGQE	36869	KIESRNHYS
36747	MEQFDHCPI	36788	QYAHHYNIE	36829	TIKGQHGMT	36870	NSQYAHEPP
36748	MFRIEAACP	36789	QYVDMYTQQ	36830	TISDHGAEW	36871	AIGHMTQLP
36749	MGAHFMNPP	36790	RCDTHRSCT	36831	TLLFAHDYA	36872	QIIPQFAYF
36750	MKTFWDATF	36791	RDSSNWLCI	36832	TRATSLTPC	36873	YFMTWGKCD
36751	MVGPSWDTT	36792	RNYLPLQFD	36833	TTIAQWCVH	36874	IKHGGWGFH
36752	MYYLMQKTG	36793	RSMLGCHTN	36834	TTNQNSEAM	36875	ECVSKWVVF

[2436]

36876	SPPNWRAFP	36917	VFLPIARDQ	36958	CTQQEASMM	36999	ETHYPKCQE
36877	AFIQQSNQD	36918	FPEASITRW	36959	CVLQENANA	37000	ETHYPKYQE
36878	VSNQFKYND	36919	PKGVSCPNL	36960	CVTYLDAWA	37001	ETKFRLFPK
36879	QMHRLFNDE	36920	KHDGEIRVL	36961	CWGFACLLI	37002	EVMIAAALP
36880	SKMEEVMDP	36921	DSDPITVDL	36962	CWWFACLLI	37003	EVMIAAALS
36881	MEQPTCMME	36922	SDCEKFLAY	36963	DADSSIQLG	37004	EVVGYPIIC
36882	QPGTQEVLQ	36923	DNVVMWWNQ	36964	DADSTIQLG	37005	EWDDWYSQS
36883	GCQGIPRMQ	36924	KDFNFEAQG	36965	DAGIGMKFY	37006	FAQLKWMMS
36884	YSDNLERCG	36925	PYGQCMRTF	36966	DELWWIANG	37007	FICQIPETK
36885	YANQAMDYM	36926	TNKPKNVLM	36967	DFAYQCYLN	37008	FWHCINACP
36886	TWPGGCKSV	36927	TLGMLLENI	36968	DFRTFNQES	37009	FWPCINACP
36887	AVDDCESER	36928	DVNLQITTH	36969	DHEHYTWFG	37010	GAPRYHESR
36888	KGVWDPSSM	36929	ENDYVHGCM	36970	DHHTCLEQY	37011	GATRQEHAF
36889	MTGIVLQNN	36930	IDLSHCRCT	36971	DKRAWDWLR	37012	GATWQEHAF
36890	QVNYYTKGM	36931	NRTAGSCAC	36972	DLMKIDKWD	37013	GELWWIANG
36891	QIGSHSFSC	36932	LAKPKCKLT	36973	DMHMLNCCE	37014	GFVQKVNDT
36892	EINREVRQN	36933	ATGLPARFP	36974	DMHMLNCCG	37015	GICWPNWYH
36893	QTIIGNVWG	36934	NHAWTKPWQ	36975	DPLKHCYDN	37016	GQNIVEMSW
36894	EIGLNGAGT	36935	CCTQFKCVS	36976	DRHNHDVAI	37017	GVYKTNQPQ
36895	RTVQEDGCG	36936	AHVFEPQRV	36977	DRHNHDVTI	37018	GYEGFEIVD
36896	NCHIDHVQD	36937	ALFCYNDHV	36978	DSRQQMSLE	37019	GYEGSEIVD
36897	EIKCSVKLE	36938	AMMPCVGWQ	36979	DTEYQTNAN	37020	GYEYVMSIK
36898	WVPTPWRTH	36939	AQLCCHNET	36980	DVAHKDDML	37021	HCSPDDMKC
36899	CTSTMHAFQ	36940	ASCEWSWRL	36981	DVQQANGRY	37022	HDFPFWSKG
36900	MSANMMTKG	36941	AVHWTQKRE	36982	DVTMAEVIG	37023	HIHFDEWNC
36901	LFNYEGRTQ	36942	AVTKCMKTP	36983	DVWTLPECS	37024	HKGFLQEWI
36902	CPNLPGMIN	36943	AWAEWCEYS	36984	DYEVCGNWS	37025	HPAGYEKWW
36903	SAAWQFGGS	36944	AWKDIPSAA	36985	DYEYVMSIK	37026	HQKSEWPNT
36904	MMATAPELQ	36945	AWMHAQLSR	36986	DYNESDMQN	37027	HSDRDEIEE
36905	NYAQYYRTW	36946	AWMHTQLSR	36987	EEMLVTRAS	37028	HSDRDEIEG
36906	SQDCGRPES	36947	AWTEWCEYS	36988	EHETVWHNP	37029	HVLHQDFFI
36907	HEAQKEDCN	36948	AYYQTICAA	36989	EHNIYNNGL	37030	HVLHQDFFT
36908	EMDSITQFH	36949	CCDLWDDMS	36990	EIMQPMQGM	37031	HWNTGMLTN
36909	FHMNSWLSR	36950	CCGPSATAQ	36991	ELNCLSRYV	37032	HWYVVLRDW
36910	YWIPTHVLN	36951	CFCMKEWPN	36992	ENGPVRRIG	37033	IAKDMHEMS
36911	WYEMMVGYH	36952	CIPPPGYKD	36993	ENQRMRQIN	37034	IAVLHSSQR
36912	WDVKFWETM	36953	CKEAWADCT	36994	ERTDTNLWS	37035	IFALIPDAN
36913	SGTSIHTYL	36954	CKNVETQMA	36995	ERTDTNVWS	37036	ILDNDQVTY
36914	ECVEWNAPD	36955	CSAAFQLQP	36996	ESGSTRNDG	37037	ILQCHEAIG
36915	MIEGHAAPR	36956	CSTKAHRQY	36997	ESYEHQMHR	37038	IMLMCTPRK
36916	PILLCHAPE	36957	CTEQIHRHV	36998	ETHRTSNNC	37039	IPGFKHSDW

[2437]

37040	ISTTFQWYS	37081	NLRIHAHRT	37122	TCLHDYGYG	37163	LHWPITVQP
37041	ITGAVQPFL	37082	NPMRPSVEQ	37123	TLFCYNDHV	37164	VLNKNWTMC
37042	KGGACDACI	37083	NVPIRALQK	37124	TNAYQYHVG	37165	ISDCMDNYW
37043	KKGWEFMYG	37084	PIALYANCS	37125	TPEHGLMVG	37166	YCVNMQYDV
37044	KKGWEVMYG	37085	PIELYANCS	37126	TQCEIFEAH	37167	NSEREYAIR
37045	KVDPIRHGQ	37086	PIGLYANCS	37127	TRKVAEEMW	37168	RAQSHMEDY
37046	KVPTDGANS	37087	PLTTDCSPN	37128	TRPVEMQPG	37169	EKFSANKFQ
37047	KYVRWFQAW	37088	QGFADGRPT	37129	TVEPEYFQA	37170	RSGQGEGHM
37048	LAESHYPFE	37089	QHLEINCVS	37130	TVEPEYFQD	37171	GSHQLMWIG
37049	LDTKWDQHI	37090	QLVYNMWDC	37131	TVEPEYFRD	37172	YMWGTELGC
37050	LDTKWDQHV	37091	QMCMNISHY	37132	TVFMNAHWA	37173	VTNLYEGSG
37051	LGGMHPCNN	37092	QMCPSCCVT	37133	TVNEGEYGP	37174	EMWHHLSTP
37052	LGGMHPYNN	37093	QPDYEEPPS	37134	TWSYDPASS	37175	WWRATFGQA
37053	LINKFPQQE	37094	QPWVPSILH	37135	VAEPSWQNN	37176	NWQWIAVQD
37054	LISKFPQQE	37095	QVETVPYLS	37136	VCTLNNHTM	37177	HNLYGVYWG
37055	LMEMTLRMG	37096	QYGHKQFEY	37137	VGNETHLGD	37178	NQSANWLQN
37056	LNNHMKKYP	37097	QYRVDSYPF	37138	VGNKTHLGD	37179	TGGQNYAHE
37057	LRISKWMTG	37098	RCQDKVMPY	37139	VHITVCRYG	37180	CSEPDPAWN
37058	LVCDSGYAA	37099	RCSHDDMKC	37140	VIYCHPMVP	37181	AFSSERVYA
37059	LVRIWHETM	37100	RCSPDVMKC	37141	VMDWWLDEN	37182	CCILWESGT
37060	LWECYGTHG	37101	RFDMWWLFI	37142	VPHTTCHTH	37183	CPQCDVMTS
37061	MAFMQGLGW	37102	RHKDTEWFI	37143	VPHTTYHTH	37184	QSSFWPAPT
37062	MDGQACLNQ	37103	RIEMFWCNN	37144	VQKWGLMQQ	37185	SKWHFYLRQ
37063	MDNQEDRCE	37104	RIEMFWCTN	37145	VSGSTWNDG	37186	VRFTRTSLQ
37064	MFALIPDAN	37105	RIKMFWCNN	37146	VTEFIMHCG	37187	KAEMKMGCN
37065	MHAQRMKYP	37106	RLDCNQFWE	37147	VTEVIMHCG	37188	MANHSISAD
37066	MHTQRMKYP	37107	SALTEPPFP	37148	VTHTVHQET	37189	IDRCCWDYA
37067	MLMYLGHLT	37108	SAPTELPFP	37149	VVVVNGVTE	37190	EPGLLWHTQ
37068	MLMYLGHRT	37109	SCGQFEWDN	37150	VYQRFCGKR	37191	HTHDVRVHK
37069	MNEDCMQTS	37110	SGSWEAFGF	37151	WINPTEMGA	37192	WKPAAYSNC
37070	MNHEHGPIG	37111	SGSWEDFGF	37152	WIVLSDTIG	37193	CGHTACSQC
37071	MPRHKHSQT	37112	SGSWEGFGF	37153	WKHCWMLMQ	37194	DICLTHDHC
37072	MSDDVVLAL	37113	SHVFEPQRV	37154	WLRILRAPH	37195	REIRDDYIY
37073	MTVQNVPEM	37114	SIEMFWCNN	37155	WLRILRAPY	37196	VYTPKYQDL
37074	MVDGPAASG	37115	SMHTWQTAK	37156	WSQEQITFD	37197	KCESWKDFF
37075	MYYVTQMCN	37116	SMHTWQTTK	37157	YNGKWPRLT	37198	QNEPEKQVK
37076	NAWMYVNMS	37117	SNERRHMCA	37158	YNMCSAGSE	37199	LASYESVEY
37077	NAYPWHLFC	37118	SPMGVAMNG	37159	YNMYSAGSE	37200	HMHFDEWNC
37078	NAYSWHLFC	37119	SSIPWNYYT	37160	YSGCWWQHG	37201	KPFHVEMSL
37079	NIIKCVECC	37120	SVTYLDAWA	37161	YSNRIPIGF	37202	GRHYIAAGY
37080	NIIKCVEGC	37121	TCAGFCAPE	37162	YVPPNYTEQ	37203	TTMTHSTLD

[2438]

37204	AEICHKRED	37245	AWCWQVILM	37286	GNECFHNVA	37327	MKYQDTHTF
37205	DAAEELRNA	37246	CAPKQECSR	37287	GNHTFNQDQ	37328	MTIQVFMSD
37206	LIAHDHKCF	37247	CDVWWFPNL	37288	GQNLNRGVC	37329	MVASVAAEE
37207	YTKIESVHL	37248	CNRQKYTAK	37289	GTAGPCRSS	37330	MYMHHALSP
37208	ALMMDDITP	37249	CNSSESHYY	37290	GWCWQVILM	37331	NGKTNPCQK
37209	TQSHWNKMS	37250	CWTMLMFST	37291	GYNKWGRNE	37332	NGTQCLNMG
37210	KTFQYMGPT	37251	CYQHLNTMN	37292	HEGLIMRLA	37333	NGTQCPNMD
37211	SNHESMHMD	37252	DANHCMWWR	37293	HIGNKHIEC	37334	NISWIQDYM
37212	SCGQFEWGN	37253	DDHCCTTPT	37294	HMEQCRCNE	37335	NMMTLWSHL
37213	VEPYMFAPM	37254	DEPTLEEDM	37295	HMYATEHCS	37336	NPGWRYMWC
37214	TTETRNWVL	37255	DGTLAILCT	37296	IACRTHQTE	37337	NQDEKLWAP
37215	TAQQFAHSG	37256	DGTLAIMCA	37297	IARRTHQTE	37338	NSEKVYGMF
37216	EANPYTATK	37257	DGTLAIMCT	37298	ICHCCMDWG	37339	NSNHYNAMR
37217	FQGMCQLRE	37258	DGYQKYQEQ	37299	ICIRGARAA	37340	NTEPTLVTR
37218	IIIQRSRTY	37259	DIACLMNST	37300	ICIRGVRAD	37341	NTEYSAKND
37219	DFTEEGEAY	37260	DIEGYGHSW	37301	ICTFYQPSP	37342	NVDAHEQYR
37220	MVKLDAQDA	37261	DISQTTITK	37302	IHSPFMPYH	37343	NVDRQHAQG
37221	EIMQPLQGM	37262	DKLPMAMDR	37303	ITTARNQYN	37344	NVRNGTNTC
37222	VPNRNCQFS	37263	DKNVFEEHV	37304	IVSPFQTGT	37345	NVYVTPVVC
37223	CIHSVHFAA	37264	DNCQVENHL	37305	IYRLCQDSD	37346	NYRCFDPMH
37224	MFAKWMSGS	37265	DPPHRNCTI	37306	KAEVKSPAP	37347	PMMIIPYVA
37225	ESESWIHMP	37266	DQEKCLSFG	37307	KATMTNCCF	37348	PNGHHWMIP
37226	SALTELPFP	37267	EARMPDDWS	37308	KCEGEEAGR	37349	QAPVIPIMT
37227	FGLQMMYIV	37268	EARMPNDWP	37309	KEDFGWKHM	37350	QHNSCKDAW
37228	PAPMSEYNF	37269	EFGFKIHRY	37310	KKEQDHRDK	37351	QMNCLKLGP
37229	ECALAQYGS	37270	EIFHACNSG	37311	KQAGYGFFH	37352	QNIATMRGF
37230	SHMGKHHHM	37271	ELQYYKEHT	37312	KYEVPGKYG	37353	QNICLGVPT
37231	ITAKQEVTS	37272	ELYGLNHSC	37313	LGSPFYERG	37354	QSFPCMFKD
37232	AADCGSFDH	37273	EMGIRGIWG	37314	LHSYQWPLG	37355	QSPVPNRYA
37233	AADCGSFYH	37274	ESCHLQHEE	37315	LKVVQAYEC	37356	QTFSEGWRL
37234	ADHTENFSD	37275	ESCHPQHGE	37316	LNNRTQESW	37357	QVNKGAPQE
37235	AEDCGSFDH	37276	EVCTDDGNV	37317	LQLIVVSQT	37358	RADWSDIAD
37236	AEDTPYNIY	37277	EYATKVKMD	37318	LQYPTACGC	37359	RCNLYLEYS
37237	AESPQRWAL	37278	FHCAMEGRT	37319	LSAKRQFFC	37360	RDGDANCSL
37238	AEWSEGECA	37279	FSRHWECVQ	37320	LTNRTQESR	37361	REKWGHMHT
37239	AEWSEWECA	37280	FSRRWECVQ	37321	MAIQVFMYD	37362	REQHIAGAD
37240	AEWSEWEGA	37281	FWSNIIGIY	37322	MAIQVIMSD	37363	RGLVQDCCR
37241	AGQKHWHYS	37282	GFNMNNMVA	37323	MGTILQEVA	37364	RHVHRYWEC
37242	AIDHHTEVV	37283	GHLDPFDCG	37324	MHAEDVTSV	37365	RVKAYAPTT
37243	ATFTFNEDN	37284	GLCIYDHRH	37325	MISAPPDTY	37366	RWSYMQADK
37244	AVCMEEREP	37285	GMSMSQSHK	37326	MKRGTAFCA	37367	SAMYGSTWA

[2439]

37368	SAVYTAACN	37386	TISWLDTEV	37404	WRMFESAPG	37422	EMCHSMGDF
37369	SGMTWNFHA	37387	TISWPDTEV	37405	YAQPDWVMA	37423	EMCVMSFGY
37370	SHIQLNAWW	37388	TKDAWNGLC	37406	YCMWYATPP	37424	FFGWKKSNY
37371	SIFMEVAWC	37389	TMSNYGCPQ	37407	YHVPYQHLP	37425	FGREGCPTE
37372	SIQRNCAFW	37390	TQIMRVISV	37408	YHWHINSFD	37426	FVELHCHTE
37373	SLMRFLPSI	37391	TTQRDGQWY	37409	YPELGKCEI	37427	GDTVMLVPS
37374	SMGWAILEG	37392	TVGPAGISE	37410	YTGHTNNGR	37428	GGPFKVASS
37375	SMPVGTEID	37393	VDGLQCCCP	37411	YTGLMWVWA	37429	GTFHNTATI
37376	SSASVNHTN	37394	VDIMRNDTG	37412	AMVLFLVDN	37430	HTRVVESTG
37377	SSDMQVCWT	37395	VEFRPKISR	37413	CALKQECSR	37431	IAKIGNCVC
37378	SSDNSHQGR	37396	VGIMWNDTG	37414	CHNQMWQIC	37432	IHTGTNSYA
37379	SSKFHERQH	37397	VNWFVRAYI	37415	CISQTFRYY	37433	IMVHCREMG
37380	STDNTGTAC	37398	VSYWFQWGS	37416	CWAEFRMRV	37434	IPQVCQFGQ
37381	STGSCAPLY	37399	VTATSPGIE	37417	CWMLGRDCC	37435	IRSQTLTDC
37382	TAVKSEAST	37400	VVTKPAYFV	37418	DGAAPLTFC	37436	IVHKLESVN
37383	TDHTENFSD	37401	WADLWQYQW	37419	DIAMNPRMC	37437	KEARSQTGD
37384	TFANPPENA	37402	WATPPLQRC	37420	DITLWMATE		
37385	THGIYGTHD	37403	WRIRPQQEI	37421	DLMWHVNAI		

[2440] 表84

[2441] 驱动脾脏组织嗜性的AAV5 VP1衣壳多肽中581至589区域的序列

[2442]

SEQ ID NO	581-589 序列						
39438	CTTYYCWQE	39455	LVELSIKQQ	39475	CECHWAYPS	39495	EPMPMPVWN
39439	TMMYYESGQ	39456	LWEHTFDCD	39476	GFQRTHVDQ	39496	HYVFGMSST
39440	YNMVHGDEP	39457	VQSFENQAC	39477	MEPHYWEWP	39497	GHNVTAMAI
39441	WGSLEWDRA	39458	VFRVKCAAD	39478	RYMNAAKFF	39498	LWSRFISMD
39442	DSGTKKRMH	39459	QSDWISIPC	39479	MNEYIMSMH	39499	VPTPWMGIS
39443	QHELEGNRF	39460	RANTWTRRH	39480	IYMSTGNTP	39500	IAYHYDQQN
39444	TEESLVRTY	39461	KVPSQLVFE	39481	QQTKNNSTD	39501	IEHCSNRYA
39445	DDYYQLHES	39462	ENCMSRMKP	39482	EIYSQLTDL	39502	FAQQIINAS
39446	LLEHPEKDL	39463	RLTEDMPDM	39483	VTKLAENVY	39503	YVLYSWRTS
39447	LWQHTCIHG	39464	NNGVHIFCE	39484	WDYSMTKKH	39504	VTNWLSLQS
39448	SDFTLGYHH	39465	SHMSYGQVQ	39485	FKHSLREWT	39505	QMVTQLNWE
39449	MIEQWDACS	39466	DVCDDHKPN	39486	ALAKFQCWS	39506	IQNPAMWGT
39450	PQYDCWKGS	39467	GDNMVQWVS	39487	FCDLAHHYE	39507	EFAHEWYVN
39451	NHYPDQEAE	39468	GYASIVKRD	39488	PWTTPENYT	39508	PYGQCMRTF
39452	EQPHCKWTA	39469	DNYMQLNCV	39489	NLAPMIQVC	39509	INYISFVDR
39453	SHNNNPSGE	39470	DVAPMHQVE	39490	EAELENVNM	39510	GVNDLQNIC
39454	EVADIVSLA	39471	QLRCEYQHT	39491	YMALDLIPY	39511	CSKATWGAE
		39472	EASIQASWN	39492	TTKNAISFE	39512	DYVDQCQVS
		39473	PCAMHNWGC	39493	STLMPDEMA	39513	MCGCGWAKG
		39474	MNANETHGA	39494	LSAQRLVCD	39514	NAELQQADI

[2443]

39515	ESSQEIKTC	39562	NMEPLKLQA	39609	VTMIHNAFH	39656	RDQCRMESG
39516	AIRQFMMHL	39563	CWECCTDSG	39610	SFKSVSQYN	39657	QIKTHSNVF
39517	QTQICHSWM	39564	GNECFHNVE	39611	AFGLTTGHT	39658	DSAKQWHHQ
39518	WFDRSWHIL	39565	LSTMFPEQT	39612	CVLSRSTMR	39659	SVLHVQALA
39519	FRAVYYYHS	39566	TAFQSLQQM	39613	LCKFANGAA	39660	ETNDNMILH
39520	TFHQSDYWE	39567	NGVGRSGDN	39614	QSYPTMYQC	39661	FQWRTEMNN
39521	CGCQVKCHT	39568	QAWPKTEYI	39615	QPYWVAVDP	39662	FQTCDFWAT
39522	WNWLKVMHE	39569	YNNDMKQRR	39616	CKYKTKSIR	39663	TVNWYSGFG
39523	MASQGGQVR	39570	MLHTSECVQ	39617	QCTLEISYV	39664	YNFDYMKDG
39524	DVEFCMKQY	39571	GSHWYDCNQ	39618	EICMRSGFI	39665	KCADLIIRS
39525	INMPVSKDQ	39572	PTNYGWMWR	39619	WVDQNVAQD	39666	CCLWSYWAF
39526	QMCVWMTRQ	39573	AHEDPNTWC	39620	YYNHKIVEP	39667	FQTLILSGL
39527	SSEYSQFAI	39574	QSMVACYTI	39621	QHEQSNHIW	39668	ILINDREWG
39528	DSIIKDQSD	39575	PWRWHKSAC	39622	YVDHWTLGD	39669	EPEGELDKT
39529	EHPEWKCMY	39576	WINEANCLM	39623	NIQLLHSYC	39670	MYAGWLNHG
39530	WKATPGGQC	39577	LSPWPQTDQ	39624	IMNPTAQNN	39671	FLTCDQFEW
39531	STARHNMQL	39578	LETMMRHGE	39625	ATMNLSWGL	39672	NMFQLKPQN
39532	RYQSHCWLH	39579	LGSPFYECG	39626	HTNYTVGYE	39673	PYGPIEYDE
39533	TFAKSDWLA	39580	ACQFYISRP	39627	PWMHTMLNP	39674	GGDGRVHNK
39534	NSQLSMTFG	39581	VFKPLDKRY	39628	EAGPRKRDV	39675	VDRLFRMEY
39535	RFDQFWDTY	39582	TFVRKPSLM	39629	LEIGGKTRT	39676	QWFTKGVGE
39536	QTSHQDQDC	39583	NDMQMNWVH	39630	ARWFSTVEQ	39677	QSRCVDNTV
39537	DHELTRSEA	39584	HAKQANGMV	39631	KSDAQRNVA	39678	QIIYRCRST
39538	MITNNCNIS	39585	VKFRITTEC	39632	WERCLELNQ	39679	KDGRNPHCL
39539	FNESGWPTW	39586	CVEQYVEQG	39633	RCADNPYFR	39680	AIGPALCLT
39540	MWGCKLFVC	39587	MQPMESATW	39634	QTDILRDVQ	39681	TQPEDHWLA
39541	SKGWADLCP	39588	VEDRVFACC	39635	MDGRDEIIH	39682	TGMTFNSFR
39542	HIHDTMSDY	39589	EVRIQCKID	39636	ASDMQFWST	39683	NAIPYFVQA
39543	WHTIMTYES	39590	KQKTQNRNA	39637	CSIQINNCR	39684	DVKSWAVCD
39544	SFVHMFQGH	39591	KVAEGQNML	39638	PIWQGYWPW	39685	VLPMGPVDP
39545	WVHPYSYCA	39592	YPTMTFPEE	39639	GAANRIQSW	39686	CDHVTKHPT
39546	QVSFNARYA	39593	ASKQGTHDY	39640	SGDWDFGEW	39687	SNVACIQAF
39547	FSTSYCHED	39594	YVIGQKDEM	39641	TQCKQLTVM	39688	THVKRFDYE
39548	WLIGMWNTG	39595	PHASGYTCI	39642	MHGIARMQQ	39689	VFNLMLQDK
39549	NSTMEHSRW	39596	EKRLEGWVY	39643	LMKTSPSYW	39690	ECGCMYWGY
39550	KYEFLLKYI	39597	NLSNLVSCR	39644	KMHGNERCF	39691	MYTGMRCWR
39551	HCDSPAATP	39598	YTELMKKAC	39645	RYPCCQSPF	39692	MDNCPPMWN
39552	DIEFLDCEN	39599	QCTAHNTCM	39646	AVYLDHVFS	39693	ATMWMQWWD
39553	YSDPKNMSI	39600	TWAPWCCYF	39647	ITAARNQYN	39694	TNSCECVDR
39554	ACIPQMHSA	39601	ICHIRTRGN	39648	VGNVMIKMC	39695	GHLFNNNEW
39555	NCMHLQGCQ	39602	CIQPTWDSY	39649	HHFHVNLDI	39696	ASMHEETCL
39556	SRVMIHVYV	39603	EQELKNSTQ	39650	NRIQMTDFQ	39697	VCDGHTAFP
39557	SMFTHQLAQ	39604	NGYLVHACQ	39651	QISKPSVQC	39698	TISQIPWVR
39558	DVTPTSWIE	39605	NADMSFRNR	39652	DVMQYQNHM	39699	KNCIMLASN
39559	WCEAVGDYA	39606	DQRHGNVSP	39653	MQHTDWTSG	39700	QGKQLYHMC
39560	GDMYYWKLQ	39607	TWYMNWMWH	39654	CWILHHRCS	39701	YFPSIQCYE
39561	GWDMLSVAK	39608	KGNPLDGDT	39655	FSWPIHEPD	39702	KMAGIGCWY

[2444]

39703	GGALARSSI	39750	QIRSLITHD	39797	CQQYAAWAV	39844	GIRFFEKGI
39704	PDLWYEKSS	39751	QFAQLWCAK	39798	CRIASFKAI	39845	GMHMPCDKN
39705	NVWTNIAEM	39752	RYNNCPNLW	39799	CSAYCHACE	39846	GQAKYWMGT
39706	HIKVEHEEV	39753	EVAFTTSSD	39800	CSSDRDGLP	39847	GYKPDRLHD
39707	EHVQCSADM	39754	YVKMCSMRN	39801	DDGHHVPIM	39848	HMVINQMDM
39708	LYMMALLPT	39755	HTVMEWSPT	39802	DDMFAQYFS	39849	HPRQGWMHS
39709	ASIATMDRI	39756	TFMFNSEWS	39803	DHSMWTWAE	39850	IAGYRHNTD
39710	YKGMEKNPV	39757	AGMYKCQED	39804	DICNATCWV	39851	IFHNYDLTG
39711	QMKGATEFC	39758	FCILASDMA	39805	DITQTDMWG	39852	INDARQHFG
39712	AADIFAIND	39759	DLIVATEWF	39806	DIWTAKENV	39853	INVAMQKEG
39713	NSGVLIGME	39760	QGFSTDVKP	39807	DLAMTMVNP	39854	IRGDEMLED
39714	TSNHNMALE	39761	FMIMAKTQA	39808	DSGPQTQMV	39855	ISNYVGFHG
39715	LIEKQRWML	39762	QSAHPEPMC	39809	DTAINQCWI	39856	IVLHMSEGY
39716	SGFFVHTSN	39763	YDYWMQAPT	39810	DVEPSECSW	39857	KAMPEQLWP
39717	QSSHCKQMV	39764	VVCRMWHQP	39811	DVQNANQMF	39858	KAVSYQNPF
39718	RDYIERREP	39765	GSKIIMHTE	39812	DVRYELPYR	39859	KMHRRIWGA
39719	RQAWPKDGM	39766	NHSLCWDSK	39813	EAVGMPLCN	39860	KNTETMYSE
39720	QHHWKWSTE	39767	TSSKHLVFD	39814	EDWYWMNLC	39861	KNWPMDKHL
39721	YCECGHWPN	39768	TKMIIFDSW	39815	EFKVYQNCH	39862	KVNEMVAEV
39722	HMGIAYNDH	39769	ENTMWQSSQ	39816	EKYTFCIFL	39863	KVTEYCDCG
39723	LCAVCGDCD	39770	PVICVWHDN	39817	ELLYTLMFH	39864	LIICTWKAG
39724	TPLMMCQAS	39771	NWTGNMQWA	39818	EMEAQLLGE	39865	LIMEAMLFS
39725	TYESLQNSW	39772	MNLQIGSKG	39819	EPCLVNISH	39866	LLRRNDELK
39726	GRALGKKQG	39773	DRKQADQIY	39820	EQNASMTMH	39867	LMDCHSVIP
39727	YTKYNWKAA	39774	GGNAGRDTT	39821	ERDHIPCYA	39868	LMQMDKPNG
39728	TATCGIEVE	39775	WGGMPNNQG	39822	ESSNTQPWH	39869	LPHMCKCKG
39729	MPWLDKPPC	39776	HSMMPGWPS	39823	ETASDPQTI	39870	LTMADGLMR
39730	TRQWMCVHC	39777	DSDTWAMHE	39824	ETKYSYEVG	39871	LTNEVREEN
39731	DQEAQVVTG	39778	EGIMNHAPT	39825	ETTDGYGQV	39872	LTWISHGVN
39732	MVENGLRLQ	39779	FQVMDHANP	39826	ETVWNDEQS	39873	MAMEYIEMY
39733	EAGCKMWPT	39780	NSNSMNCVS	39827	EVALQMFSQ	39874	MARIYPGTC
39734	AIGCFQNGG	39781	QQVLGLAGN	39828	EVCQNINEA	39875	MAYMEHHVG
39735	GLFEGSPGA	39782	QSSPDQRAL	39829	EVMPITGYG	39876	MEMEYEEFM
39736	YLNKNPHQM	39783	AAPIAHSMR	39830	FHELRHVDT	39877	MKQDSTQGM
39737	CQMLCRPPV	39784	ADPTTETPH	39831	FKDWNTVGM	39878	MPLCGYNVR
39738	VNLPYGYAK	39785	AFNEMEQGD	39832	FKTRETTEI	39879	MTDQPPSKV
39739	LMTPQVKQA	39786	AHAPWCWAV	39833	FMHHTVRSF	39880	MVWNVAMMD
39740	SQERMDYDG	39787	AMTFSVSCS	39834	FMPAKSVAD	39881	NDELCVYCW
39741	NTHTGNDRL	39788	ANYALHYDR	39835	FMRSVQWNA	39882	NGVEQIKHE
39742	QGHKTMHVF	39789	AREFWHCRD	39836	FSGMETRPP	39883	NMVMFVMDN
39743	FGGMTPYGV	39790	ATNIKESVN	39837	FTLVATNIH	39884	NQICHQYDA
39744	SECEVLCSL	39791	AVHAFPWSF	39838	FVSSVTQVA	39885	NSTVHWQQN
39745	EIHMFGQSE	39792	AVKDMKKWD	39839	GAGDMPMLD	39886	NTQWYMMNA
39746	QAKMILDGT	39793	CDEDLYVTP	39840	GDQGPRQQL	39887	NVEDNRKVF
39747	VSDECYFDM	39794	CHIQVDHHF	39841	GFHPIPWFP	39888	NYRNQKNCH
39748	REWMFTLMA	39795	CMDGQLKWF	39842	GIEMDLCNG	39889	QRATTKRHA
39749	CNGYWMCMS	39796	CPWNSQVTL	39843	GIGPQLNWA	39890	QVHPVNHWG

[2445]

39891	QVQVVDEWS	39938	AAEEGHTWS	39985	AIAEYAQQG	40032	ARITWNVCT
39892	QYQHERGWL	39939	AAFQPDVNF	39986	AIAWYMGEA	40033	ARKQIVGMD
39893	RGGYIVAGI	39940	AAHEKQCAA	39987	AIECSWGWI	40034	ARQHMRRQP
39894	RGYWDVREG	39941	AAHTWHVAW	39988	AIKMDGDFQ	40035	ASADVHYAQ
39895	SAECMTVSS	39942	AAHYNDGCT	39989	AILPMLMQS	40036	ASANRKQMD
39896	SIWHYQHMQ	39943	AAISMLWVE	39990	AINIMDQEA	40037	ASAWQMMMP
39897	SIYKYTALN	39944	AAKEATSEG	39991	AIPDYQIYE	40038	ASCEMTQTN
39898	SMLLSNQNQ	39945	AAKWKAKCA	39992	AISHNMKTP	40039	ASDQFMADE
39899	SNQCLPYDC	39946	AANPIESVD	39993	AISSCARQI	40040	ASHKDGWAT
39900	SSAVIMKFP	39947	AAPLKRGLP	39994	AITMSMPTW	40041	ASHNFSDEG
39901	STHTVSWDM	39948	AASHTEGFL	39995	AITVGCGWA	40042	ASNMHANAH
39902	STQPSAKQV	39949	AASPCNDFW	39996	AIVEYWSSC	40043	ASRDYDLAP
39903	SVDYVEGYC	39950	AASQCFCMP	39997	AIYYCLSDW	40044	ASSSCNGGH
39904	TAMAYTSCE	39951	AATFKSRHA	39998	AKCTWRSYH	40045	ASTLGCDCC
39905	TCQHAPSPM	39952	AAWHANLSR	39999	AKEDQHQFQ	40046	ASVTRVAVL
39906	TFEQIPCWT	39953	ACDVKKHFG	40000	AKEYMMLSY	40047	ASVVSDECN
39907	TFHVCQTMG	39954	ACHAYPSAW	40001	AKLKQAWSG	40048	ATAFGGEYY
39908	TMQVMEVIS	39955	ACMEAIMTT	40002	AKMDWKLPM	40049	ATCYLMKLD
39909	TNLRPMTER	39956	ACTDFMSAC	40003	AKQQPEKQY	40050	ATDDECCVG
39910	TNNTMRHEF	39957	ACVEYIDSA	40004	AKSCWPSVW	40051	ATDDVRCLW
39911	TRFSSQGKM	39958	ADAVQLMCP	40005	AKTHNPNTD	40052	ATDGYSKWD
39912	TSHAQTQMD	39959	ADCDYGRKF	40006	ALCSRDSSN	40053	ATDPERSQW
39913	TSRTRNCTH	39960	ADCIFESDG	40007	ALLAMQLAT	40054	ATDTWDRAS
39914	TYALDMVST	39961	ADMYANENE	40008	AMLHTMGHW	40055	ATEYQPNTW
39915	TYGLYDDSY	39962	ADPKYVPTT	40009	AMMNWQMMN	40056	ATFSCDPWN
39916	VAMYYESEM	39963	ADQCNPLCA	40010	AMMTVEYFN	40057	ATGEEDVNW
39917	VDNKFMEHK	39964	ADQCTMLRF	40011	AMREWRVFP	40058	ATGFGTEDS
39918	VDPCANEGH	39965	AEMCKTLIT	40012	AMTSMIQSG	40059	ATHEMKGMN
39919	VIEEMTMVC	39966	AEPCKLKHF	40013	AMYGKAAGD	40060	ATIQHEAGM
39920	VIMPMQPRC	39967	AEPRHDKAC	40014	AMYNPYTQK	40061	ATKFFRMET
39921	VMLHLIHYP	39968	AFKGHYMAA	40015	ANCQPSRVC	40062	ATLTWNYSQ
39922	VPCIWVLKD	39969	AFNSGMAKI	40016	ANHTKLKMY	40063	ATNFDVGHE
39923	VSTCCQANG	39970	AFVENMWNF	40017	ANMENWVQS	40064	ATNSNYEHR
39924	VVECSGRAN	39971	AGALPQAWA	40018	ANMSATYYE	40065	ATNWMRDQT
39925	VWVVRNHAD	39972	AGCFFTGIK	40019	APAPNVRYS	40066	ATSEQRIVV
39926	WCAGATMQW	39973	AGDHFLLVG	40020	APNMMNWIA	40067	ATSLAHCVC
39927	WEKLSQRNR	39974	AGRMRFAKM	40021	APVIGWGAY	40068	ATTDNYHDR
39928	WEMNDMQDY	39975	AGSYYGVIN	40022	AQDNNNCAA	40069	ATTHETVLY
39929	WINIPCHCI	39976	AHCVPWSWT	40023	AQEPVQGNY	40070	ATTNSSGSH
39930	WKCIVPGQD	39977	AHEATSCMV	40024	AQHVAIFNL	40071	AVAMLVDVW
39931	YMVWSMMMI	39978	AHGCGFAWP	40025	AQMYRWVMT	40072	AVAWKPRMP
39932	AAAMCGYQV	39979	AHKQMDWHT	40026	AQPLEHHYN	40073	AVCNDDHSA
39933	AAAPSDDEP	39980	AHLPLMLGT	40027	AQQEVKCGY	40074	AVCRPSTSY
39934	AADREKSWM	39981	AHQGVAGTY	40028	AQSHGTEMF	40075	AVDCYGLFS
39935	AADTQSYIR	39982	AHVDNNLLF	40029	AQTWGGAGS	40076	AVELETQWW
39936	AADTYVNHM	39983	AIAAENDCT	40030	ARDHGKWNW	40077	AVFYPYITD
39937	AADYDDQQW	39984	AIAEHWEGR	40031	ARDLYHRSE	40078	AVGHNGSHR

[2446]

40079	AVLNSWPAC	40126	CKNGPAYVP	40173	CVCQYHQTN	40220	DGQCPNAWV
40080	AVMPYCELE	40127	CLFYDEAMG	40174	CVDPGVSWP	40221	DGTHNHLQG
40081	AVQDKVCTC	40128	CLHVPHNSL	40175	CVIHMKTKI	40222	DGTYLMFHA
40082	AVVLTDWNE	40129	CLLVAECLQ	40176	CVKQYMFAQ	40223	DGVDKEWYE
40083	AWAFHHAVP	40130	CLQFKCRLS	40177	CVMDAMAIP	40224	DHQANRTCC
40084	AWANYIRHE	40131	CLQHTSMKC	40178	CVMTEDFQF	40225	DHTWQHEGH
40085	AWEFVNYPP	40132	CLSDNYHNQ	40179	CVSESNIWH	40226	DICAQSPGA
40086	AWEVNGHYH	40133	CLVSIKDCH	40180	CVYRHQREY	40227	DIDSVQEVP
40087	AWWFEAPMV	40134	CMACTDIYM	40181	CWDSTLYGY	40228	DIDVGHSFT
40088	AYAPEPMCP	40135	CMGNEDMST	40182	CYEPYHRHV	40229	DIFLNKCTY
40089	AYLNFYDGP	40136	CMIEKNMDY	40183	CYFFARNEM	40230	DIGTRVCWH
40090	AYLQFQIQG	40137	CMPLSMGTR	40184	CYGHRQIKD	40231	DIHNDIQHG
40091	AYMTSNSFG	40138	CMQNSIKEC	40185	CYKLSNYIV	40232	DILPVQAPG
40092	AYSRATTHI	40139	CMRDCNHYG	40186	CYQWHHDGT	40233	DIMPYASQG
40093	AYVAKKNMN	40140	CNASKFWQD	40187	DACKTINQF	40234	DIQIQTYQS
40094	CANEHNHTA	40141	CNQQIQFFG	40188	DACLYWAPW	40235	DIQLYNNAF
40095	CAPVSTRKA	40142	CNRQGCMHE	40189	DADAWQMVT	40236	DISAMAGWY
40096	CAQSIPEAW	40143	CNVYRFHEA	40190	DAGMPWFMN	40237	DIWQASTRA
40097	CATCHTLHM	40144	CPCHEIHPQ	40191	DAGPSVVVQ	40238	DKESWCEVT
40098	CCCEENYRR	40145	CPFFHNKLA	40192	DAKCGRMVA	40239	DKHTYEHGM
40099	CCEQFELGS	40146	CPQCLPVTE	40193	DAMPSCEFS	40240	DKRCFQQSN
40100	CCLPGWGSL	40147	CPYCARQHE	40194	DAQLLVMRM	40241	DKSSTARGM
40101	CCNLSDEHH	40148	CPYQDVFQH	40195	DAQWPLQRD	40242	DKYDTDWAF
40102	CCRSGDCMP	40149	CQFEYQIFQ	40196	DARFTFVES	40243	DKYSNKCLR
40103	CCTVTQAPE	40150	CQHLVRQGL	40197	DCAEWLVRD	40244	DLEGTEQSL
40104	CCYHTQLLH	40151	CQLSDSFLF	40198	DCGRAQMMA	40245	DLEYIWRQD
40105	CDMALVGGD	40152	CQMGFCCHS	40199	DCHGAFDRE	40246	DLIMGRKKQ
40106	CDNNWPIKQ	40153	CQQFYERVD	40200	DCSLMGSDP	40247	DLMLNFHLD
40107	CDWSNIYDV	40154	CQQGSPNDW	40201	DCSRHHMAQ	40248	DLQHFMSYP
40108	CDYLLAQCC	40155	CQVESIGPH	40202	DCSTQRTIW	40249	DLTQRISMS
40109	CEDFCMVNR	40156	CRCWSQGHQ	40203	DCWSEQQPM	40250	DLVNCNNAN
40110	CEFRPVVDH	40157	CREDKLEAR	40204	DCYHIDMGC	40251	DMAFMHSIC
40111	CEHTGSFQS	40158	CRIPMPARC	40205	DDAPLKITA	40252	DMCKKMLSW
40112	CFDRSIGYE	40159	CRVWEPLRM	40206	DDDKKHACR	40253	DMIGSQAEL
40113	CFHYLMNTW	40160	CSEMMQIHC	40207	DDGQCDRKG	40254	DMVYLAGNA
40114	CFMDGWPFV	40161	CSESGLWCS	40208	DEHCDSPAI	40255	DNEDVLHMG
40115	CGAICDAVV	40162	CSKGDCQPH	40209	DEMKWLNCS	40256	DNGWFENMA
40116	CGEPPHIQD	40163	CSLLFVSCV	40210	DFCKQCNGL	40257	DNQCNEITL
40117	CGRHWDPWF	40164	CSMYWDWLF	40211	DFDWEVNPT	40258	DNRDQHWRT
40118	CGTGMNGAF	40165	CSTHIAQSE	40212	DFEPNEEHY	40259	DNRNERRQS
40119	CHTLKSGEH	40166	CSTSFSQRG	40213	DFLVFSVQG	40260	DNTVNTIIW
40120	CHYPYCVPL	40167	CSYNDMDRC	40214	DFSENGHTF	40261	DNVGRVNSV
40121	CIGFAMESG	40168	CTCMTKCVD	40215	DFSERRIVV	40262	DNWKPTRIG
40122	CIQRYTTEQ	40169	CTELVQFIG	40216	DFSQFNQGH	40263	DPEQAYTKH
40123	CIRVQEFGG	40170	CTFSIACDH	40217	DFTLSFDAH	40264	DPLLKSEWG
40124	CIVEEMAHW	40171	CTGIPKCVK	40218	DGPLTCEQA	40265	DPNNYKGYY
40125	CIYEQERKM	40172	CTNQHKLSS	40219	DGPMSKQEL	40266	DPPLHWVTM

[2447]

40267	DQAFEWFIY	40310	DVHVPCVIG	40353	ECKVYSGGK	40396	EHQLAHESA
40268	DQCKTLFYT	40311	DVIMCNCPK	40354	ECQRPHPIQ	40397	EHQLPQVDF
40269	DQGQVGCFV	40312	DVMSMFTDV	40355	ECRMIWNHK	40398	EHRFRVKLA
40270	DQGYMPTKT	40313	DVRSGNFYP	40356	ECRPWVWPG	40399	EHRQEQTRL
40271	DQLNQQAMS	40314	DVSICVTCT	40357	ECSAFHWQP	40400	EIASGSMYW
40272	DQPFMNRQH	40315	DVSWHYAQG	40358	ECSCEEIPC	40401	EIATNYYSQ
40273	DQRICGEKC	40316	DVTDNGTCC	40359	ECSQMISSE	40402	EICQINHSG
40274	DQWLMQCKN	40317	DVVNEDTGP	40360	ECTMTNASN	40403	EIEMDGAHW
40275	DRCNKIWPD	40318	DVWEHSHPS	40361	ECWNWESFI	40404	EIEVFWKMH
40276	DRGWQNWTY	40319	DVWTPWNAS	40362	ECWNYLMSW	40405	EIGIRLGQD
40277	DRKQAQRNG	40320	DWAKYWHFP	40363	EDCLFPACF	40406	EIGWHMNYV
40278	DRNTIRKMA	40321	DWESDYYQN	40364	EDDPVSVAD	40407	EIGYRSMMT
40279	DRPHGQMMT	40322	DWFQCHCAA	40365	EDFHCFPLS	40408	EIHEGQLAQ
40280	DSKLLCCSC	40323	DWILELMST	40366	EDIAPAHGT	40409	EILKPGAHM
40281	DSKMAGQAN	40324	DWKLFNAWK	40367	EDIKQQMVC	40410	EIMISIGHT
40282	DSKQAFVMT	40325	DWMEGSENG	40368	EDKRDAWAR	40411	EIMTRYGGH
40283	DSKTIQVYW	40326	DYELTPSWY	40369	EDMTVNWSS	40412	EISCKKWNQ
40284	DSLIQDFLH	40327	DYGESKLEP	40370	EDQLCHRTH	40413	EIVYSVFGL
40285	DSLLANQDG	40328	DYGMPECVD	40371	EDQWVQMCA	40414	EIWASQQVN
40286	DSQYQAIQF	40329	DYILLTTWE	40372	EDWHTFYWK	40415	EKEFVHSCD
40287	DSRHLACYT	40330	DYRWIKGLA	40373	EECVTGLHA	40416	EKEPLHCDG
40288	DSRILQAGD	40331	DYSQFGGGQ	40374	EEEPLYVDN	40417	EKFMLWFWG
40289	DSRMCPPYN	40332	EAAFIDAHC	40375	EFDQIGLFN	40418	EKGCVGPVS
40290	DSRQPNAGL	40333	EADWDSCWM	40376	EFLLHHCAP	40419	EKHTMTWHN
40291	DSRVDSDDY	40334	EAEWVLCCK	40377	EFLPKSDQH	40420	EKKFRTADA
40292	DSVNSCAFP	40335	EAGPLKDCC	40378	EFMSHSTGA	40421	EKMCEWLTD
40293	DTDWIDCMG	40336	EAHMRRYSK	40379	EFNPAMWIV	40422	EKMFTLSDS
40294	DTEFRVMDW	40337	EAHSVRYLM	40380	EFPGKHWVW	40423	EKYGTQREN
40295	DTISEQMRT	40338	EALACQDPG	40381	EFRQMGACG	40424	EKYHVDHHA
40296	DTMLKFVST	40339	EAMRRSNLH	40382	EFVVLPEAI	40425	EKYTYNMGT
40297	DTNIDLKPP	40340	EAMRRYRSP	40383	EFYQQAYQH	40426	ELCKGNCFL
40298	DTRIDSVYE	40341	EANDYNQNC	40384	EGEQPSMAI	40427	ELFSFGHRE
40299	DTRQWDDGL	40342	EANRIDAVS	40385	EGHGCDPWF	40428	ELTLRMNWG
40300	DTSQYGNHV	40343	EANRVLCMN	40386	EGIKKEGMP	40429	EMAIRWDYT
40301	DTVQLMGSQ	40344	EAPLPKWQC	40387	EGLNDDSRA	40430	EMAMNGRTG
40302	DTYELESQL	40345	EASCYWEGR	40388	EGLSMNEYW	40431	EMCGSSKVY
40303	DVAQWFEAS	40346	EASMGRNGH	40389	EGQTSSECF	40432	EMESWDYCQ
40304	DVARCMDCQ	40347	EAYPQKECA	40390	EGRLTMKHH	40433	EMGEHHSTL
40305	DVASKDDWP	40348	EAYTVSGYA	40391	EGYQHPNEF	40434	EMHMIKKHY
40306	DVCKYQHNP	40349	ECANHEKTY	40392	EHDPYMKVF	40435	EMIKTAWAS
40307	DVCYAGLTE	40350	ECDFRMIWD	40393	EHDWPADRI	40436	EMTNAETFT
40308	DVDDRMEMC	40351	ECEFPDVYA	40394	EHEQDCEQI	40437	ENESCMHIE
40309	DVFNAGPPF	40352	ECGEFTCML	40395	EHPIDWRKE		

[2448] 表85

[2449] 驱动甲状腺组织嗜性的AAV5 VP1衣壳多肽中581至589区域的序列

[2450]

SEQ ID NO	581-589 序列
42438	KGVWCKIDV
42439	QQFPHNRQC
42440	DLKERMRGF
42441	ISTKWDYIA
42442	ISWISHASP
42443	DKTVQTSAR
42444	RMYATEHCS
42445	GLWLNVAPR
42446	SMHYYDTFE
42447	FNIHGDLLC
42448	WMSEQMFKG
42449	ILNMKACDV
42450	VFKRVVGFA
42451	QMQNENRQM
42452	RTREPFWWW
42453	MQYRGMSAS
42454	VSDTMNCGS
42455	EAEPWDYWA
42456	IAWADGLEF
42457	IAWCDGLEF
42458	IAWSDGLEF
42459	IAWSDGMEF
42460	IAWYDGLEF
42461	IVWSDGLEF
42462	MAWSDGLEF
42463	MWIEFNEYG
42464	MWPEFNEYG
42465	MWTECNEYG
42466	MWTEFDEYG
42467	MWTEFNAYG
42468	MWTEFNECG
42469	MWTEFNESG
42470	MWTEFNEYE
42471	MWTEFNEYV
42472	MWTEFNGYG
42473	MWTEFNKYG
42474	MWTEFNVYG
42475	MWTEFSEYG
42476	MWTEFTEYG
42477	MWTESNEYG
42478	MWTEVNEYG
42479	MWTGFNEYG
42480	SAWSDGLEF
42481	VWTEFNEYG
42482	MWAEFNEYG
42483	MWTEFNEYW
42484	MWTEFNENG
42485	MWTEFNEHG
42486	LFLQLYRHC
42487	LWRSLQDGN
42488	QISVYDTKY
42489	AMHCQPAFA
42490	NAIPYFVQA
42491	GLNICKNQQ
42492	MMVQWGAMA
42493	MTSTSSRGE
42494	DGQNRDQMC
42495	DKIDTCCFL
42496	NTQVISRWT
42497	GTMCREQMI
42498	KTCLCIFTK
42499	EASIQASWN
42500	HACSLLQQY
42501	RVTIMFTGT
42502	LWGWPLHKL
42503	TKQFVHNED
42504	AQIQPAMFP
42505	KGNSVTHVQ
42506	LIHWLKHDL
42507	ICHTEVMFN
42508	RVMMGNNQA
42509	SPQQPSVFH
42510	LTLRQGYIT
42511	EFVLGCQGI
42512	YSSHHSMDY
42513	SVTYHEWHP
42514	GCLQMPCHE
42515	MVENGLRLQ
42516	SRCTVMDPG
42517	IVQPMCYQQ
42518	DVKAFMGGV
42519	AYAPSKNAM
42520	NVYMRTCNM
42521	NAAYQMAAY
42522	LVQKWMRTC
42523	EKHFMQVDR
42524	LEIVMHSLP
42525	NQREWHGLA
42526	LADGCCCVN
42527	AVETGESAM
42528	RCADNPYFR
42529	ETHTVSREC
42530	QAAWIPFAG
42531	AVEDYWRYF
42532	QYVQCAAKD
42533	ACIPQMHSA
42534	AVTGHVIGA
42535	WGDARMIGP
42536	TGMAMSEQS
42537	VFNLMLQDK
42538	VTCNWCVMP
42539	DANHCMWWQ
42540	QISKPSVQC
42541	WTTSIAVDT
42542	QTSHQDQDC
42543	QAKMILDGT
42544	IHNRYAQAP
42545	NWYTTWIDE
42546	CMEVTCFVV
42547	AYRVKEVAY
42548	APVHPTQSM
42549	QKAIVAQND
42550	AQGFKPSDG
42551	NANYWFYYD
42552	TRETMNEWH
42553	GVWMPHGWR
42554	AVHNVKDVH
42555	ETSTHCCKV
42556	FQDERTTHH
42557	KSNLCYHVQ
42558	DTRRNLCCD
42559	AEHMVVDHE
42560	DVYTMHELH
42561	MKEDNKFAV
42562	MFFPGFVAH
42563	RFDQFWDTY
42564	HFQPVFMQP
42565	DIEFLDCEN
42566	QCLQTNPYN
42567	RTMDCEYTG
42568	MSGPNDTEF
42569	IGDEYATHT
42570	SFSDTHLGW
42571	YFPSIQCYE
42572	IRLMTPDEN
42573	ICHIRTRGN
42574	MCDNFPSAH
42575	SSIYKEIQE
42576	PMMIIPYVD
42577	PLTIEVNCT
42578	NVCDNGHEI
42579	KIAAWNDFL
42580	HQCNCLLSW
42581	GAANRIQSW
42582	DVWWTLTRD
42583	CSESYRGPD
42584	TISQIPWVR
42585	MFMTNVNSF
42586	GKGQWTAQE
42587	SGALKNREE
42588	LHSYQWRLG
42589	DVKSWAVCD
42590	LPSHCNLGS
42591	GQQTIEMSY
42592	QWFTKGVGE
42593	QRNEMSASR
42594	YPAQYTAFS
42595	DHLSNDAEM
42596	DQEAQVVTG
42597	ASMHEETCL
42598	TTYPHNTHG
42599	QFAQLWCAK
42600	MTDNNGPLF
42601	RYPCCQSPF
42602	CAWHMKQGN
42603	TSVIRLFWI
42604	TTTNQFMFQ
42605	MMLGYMEQD
42606	GLFEGSPGA
42607	AEDTPYTIY
42608	QHKWHDVGV
42609	NRIQMTDFQ
42610	YVDHWTLGD

[2451]

42611	ACVERLHHK	42658	KSNDSDNEF	42705	AWDANLAFV	42752	MNKFREEYG
42612	DQRHGNVSP	42659	LMWTRNTAH	42706	CAHTDKVHH	42753	MPQGAVNSD
42613	ATGTYNLQE	42660	LNIQCTNQA	42707	CEIVVSHVE	42754	MSSFMASRE
42614	EAGCKMWPT	42661	LRHANRRTY	42708	CLRSDIGKQ	42755	MTNNKMMYE
42615	KQKTQNRNA	42662	LVKAHCAWS	42709	CNEKGYHWH	42756	NCVVNLCTY
42616	ASKQGTHDY	42663	MAELCMEGH	42710	CTLSFMEDK	42757	NEPLFMMGS
42617	TYESLQNSW	42664	MYMPATMAV	42711	CVRFTPTHN	42758	NGVWCKIDV
42618	MNLQIGSKG	42665	NAEDVDQGE	42712	DCAWQTNTS	42759	NLSPINWHL
42619	GRALGKKQG	42666	NCCFVYWWL	42713	DCMCNFAGM	42760	NQLPEMIET
42620	EVRIQCKID	42667	NDRHVPCRD	42714	DFFPMKIAD	42761	NRDGVPQSY
42621	PHAETHWYT	42668	NLMTDWWDI	42715	DHRYMMSSV	42762	NSSMIEWDC
42622	WERCLELNQ	42669	PPEQFQSGK	42716	DISLKLCHQ	42763	NTAHLQDGW
42623	AGMYKCQED	42670	PVQKWMRTC	42717	DITWPNHSY	42764	NTRLVFGIT
42624	TVNWYSGFG	42671	QACLQIMWN	42718	EEGATSGWP	42765	NTSIWERDQ
42625	ATMNLSWGL	42672	QGYRSSEMQ	42719	EIQPMMRKP	42766	NTWVPHEFY
42626	QSSHCKQMV	42673	QKALNVHHA	42720	EMTLQSVSS	42767	NVFVAYCKS
42627	YSDPKNMSI	42674	QMHMVHKCE	42721	EPKYWWVNF	42768	NVMRVPTGQ
42628	GHLFNNNEW	42675	QNMLMQAGV	42722	ESILPGKIL	42769	PRNQLGPGI
42629	DINTCRVCY	42676	QYIDTYHWT	42723	FASLCNNGL	42770	PSSQFKCDQ
42630	MSDQKEHMS	42677	QYQTHPMMH	42724	FASLSGPPN	42771	PTIACGNMV
42631	AGCSISWTP	42678	RNKSPDECF	42725	FTCNHLAFY	42772	PWKYCCMDQ
42632	AVIVTADMQ	42679	RTAHRDDFA	42726	GEEADEYHM	42773	PYVDKWANH
42633	AWQQNTFWY	42680	RVEYPPFTS	42727	GGMAHAIIN	42774	QIHRAGAKQ
42634	AWTQASRGN	42681	SFLVMNRMG	42728	GLENVVVGL	42775	QNRLVNRQH
42635	AWYKAWSHV	42682	TDLWYEKSS	42729	GNDFFVNIY	42776	QRDFAWQFC
42636	DFFLMKIAD	42683	TEAKRTIDW	42730	HASSEHVQD	42777	QRQGQPHYW
42637	DLDFWNGQS	42684	TGCYHVHMD	42731	HMGPYMESE	42778	QSFSYDAKE
42638	DVQEAMNLV	42685	TICKCKISN	42732	IAMRTLNQK	42779	RFKTCRYMY
42639	DVRLATAWH	42686	TIQQCMHLG	42733	IAMTAILCD	42780	RKVRIDGDT
42640	DVYATPVVC	42687	TPQNNAMSM	42734	IATWGEQQM	42781	RNEKGYHWH
42641	ECVEGQSCW	42688	TRRTCPIAM	42735	IECEWPQDE	42782	RPEDVPVAP
42642	ETKYWWVNF	42689	VRWGYTCFY	42736	IECKWPQDK	42783	RSFMCGLIG
42643	EVTQCDFFV	42690	YFEHMMITT	42737	ITPKFMFPS	42784	RYKSRGALT
42644	FTGASFHIA	42691	YFHMIDQIP	42738	KASEHLPGF	42785	SATILANEV
42645	GCMYIRHAY	42692	YNDFNRGEW	42739	KGLTIQKYA	42786	SIVDCQGAF
42646	GIDAHAIIY	42693	YSNCQPTHM	42740	KMCTHDLAC	42787	SLVPVSKTR
42647	GTMCRERMI	42694	ADGWWMDSL	42741	KPQPTGIFT	42788	SWTCCATQC
42648	GYMCIRHAY	42695	AHLEMMNQF	42742	KTDFVFLPQ	42789	SYDEICSTV
42649	HAYSLLQQY	42696	ALLMFSLSN	42743	KVVPAEHQY	42790	TAILNQEIH
42650	HLAPYCQQA	42697	AMLDSHSAF	42744	LAGPKLWEQ	42791	TFSPARHPK
42651	HLTMWLVMT	42698	AQSRMHTMD	42745	LAHMIEITA	42792	TGHPLSNCL
42652	IHDLVPAYD	42699	ARTMYVESF	42746	LCGHCGESC	42793	TLGYVWSVG
42653	IKQFWTKKM	42700	ASEYFPVFK	42747	LCKYCFTAD	42794	TNERFTIQC
42654	ITHRYGWNL	42701	ASIWPEGKD	42748	LRSLWEAWH	42795	TQDIWYNYV
42655	KGMHSMPGI	42702	AVICKRFVG	42749	LTKFLHKGD	42796	VCYCSSTAK
42656	KLTIDTACR	42703	AVIMTADMQ	42750	MGYKRAASR	42797	VDQPTTHWG
42657	KQAGMLLGW	42704	AVKANDTVH	42751	MMFRPNSAH	42798	VHGQYPCQE

[2452]

42799	VHNWFLTWW	42846	LFSKEKHHR	42893	HIQCQGDSW	42940	KNTMCDNDA
42800	VLTTIWKNN	42847	LMSPYNFLT	42894	NNDIRIFGN	42941	AVKMPVEQS
42801	VNAQQNWGV	42848	LMYQDDMFE	42895	WAPLPYNWC	42942	FTANEDQCH
42802	WANTQWQDD	42849	LQHQLHVYE	42896	LVDGYRMRW	42943	CRRGQLLYE
42803	WGYNYQCAF	42850	LSQCQSYVA	42897	DLCVHANVR	42944	DKIMMAEMD
42804	WLKYKKTWI	42851	LTKPCWPWA	42898	TLLPHWQAT	42945	PIAFGCERA
42805	WPDQNMFAV	42852	MAAGNTTCA	42899	SIPMGNNWG	42946	ETGYRENPC
42806	WYTFLQFIT	42853	MCSSDLEGC	42900	FNTLGLLFD	42947	RFADYCKEE
42807	YCNAYMIGM	42854	MDCCARQAP	42901	ENMHKMGML	42948	NWVFYTNFP
42808	MEHVVSVQP	42855	MPSCINGPY	42902	RNSHGWRLD	42949	WECQIQDMK
42809	NTRFPDQTE	42856	MTNRCAKYK	42903	QCARPWYRS	42950	DIYHFNHRS
42810	ACLSPDHDR	42857	MVTFRENPA	42904	MWGWPGDAF	42951	NSFANNFVM
42811	AMCMLSHIH	42858	NASVGAWAQ	42905	KVIRKDSCV	42952	DAQWSFEGA
42812	ANRNCVTPH	42859	NEQVATPMA	42906	MSPNTKDCT	42953	RINCNYGQM
42813	ARMINDNDP	42860	NIGYMNVLF	42907	NGRLWWQRY	42954	CHPPMTNMG
42814	ATWSADVTA	42861	NLRECTIMN	42908	TYSQPPVSG	42955	WPDSNMTGV
42815	CNFTVCNMS	42862	NSLLRHGEQ	42909	HFTHTESAH	42956	QTIMRHAQT
42816	CTAEMAPMT	42863	QLHQCPAVG	42910	SRFWNRWDF	42957	VERLCACQA
42817	CTDGKHASW	42864	QSQPDHVSW	42911	CLGPWNVKS	42958	VMMYNDSIS
42818	DEISDPEEF	42865	QYECESDNS	42912	CCMSKMPDE	42959	AWDANLAFA
42819	DGAQDKWAC	42866	RNDYSDQYD	42913	MTHPYWEDS	42960	DAGKFGVDI
42820	DLIYKGQNG	42867	RSDQHSPHM	42914	THNVKTDYP	42961	NQSVCYSSV
42821	DRPKRQVLY	42868	SCQRNQQTN	42915	LDYHFPFEA	42962	LRHDVFSQA
42822	DSHIDQFGF	42869	SFHNCMLDP	42916	WETLEHWRG	42963	DFAAQQRAQ
42823	DTEKGQAWW	42870	SSTNNQTVD	42917	DANWRTWQH	42964	SAGPFFDHY
42824	EDILEITKH	42871	STCVTEMFY	42918	CVVPYYRFT	42965	CMASVNYFD
42825	EMAGMREFQ	42872	TCFNQNNTW	42919	HHEPYGSNC	42966	CWARMQGYM
42826	ENNPSHTQL	42873	TFFQAVVIA	42920	THHHVLEVP	42967	CFVDSVFMY
42827	EPYKQMGAG	42874	TMASWECYQ	42921	AAMILESHK	42968	APRNVHCGT
42828	ERTLTKAGS	42875	TMDSIQWFE	42922	CMRGKWDRL	42969	GRCFMPPGR
42829	ETVMMEVRD	42876	TQKAHKNPG	42923	VTMIYDLHQ	42970	KNHAHEFIE
42830	EVTKCEPPS	42877	TTDSKLLGH	42924	LDHHAEVLQ	42971	IIAVSMRGR
42831	EVTKCGPPS	42878	TVEDRRKQG	42925	MFYPFQCGN	42972	LPRMPECSV
42832	GQKSYDGCH	42879	VADCPHHFD	42926	FASHQMNMM	42973	TTMVQMGAW
42833	GVDNQDFSW	42880	VHPYAESVD	42927	PPESQESST	42974	TLVSCDQGN
42834	HAKCNRNIK	42881	VVMVQHRTA	42928	PDPKAWSNY	42975	SADRMISHE
42835	HQHDDRPDM	42882	VVPHPFPAY	42929	IRDLVPAYD	42976	MPHHSPDCW
42836	HVMEWLQCP	42883	WWEACPAAS	42930	QCPCSRLIC	42977	TPSQMDMDT
42837	IAGPIEIGM	42884	YCTTLSMES	42931	TICKCTISN	42978	SCNFEFLWM
42838	IKQYPQNKP	42885	YFVMQWFHY	42932	DEEFCEMSA	42979	TTVDYAGTG
42839	IQAVMASGW	42886	YMVQDRDEW	42933	LPRPGILME	42980	WKCCQDQLL
42840	ITDEYFMNM	42887	YMYSNMQTM	42934	PYGKAAEKL	42981	SWFAGQWEW
42841	KACSMQQEW	42888	GTSEWMHMA	42935	EKMRYPPMI	42982	EITDKYWGE
42842	KGVALCDYQ	42889	TVGTAKSQI	42936	SFRCCWDWH	42983	CRGSLSNAN
42843	KIYMQCDQE	42890	DEKFGQVED	42937	KPPGSRWIT	42984	KAYMMMEGP
42844	KPEWNHTSA	42891	GTPDMRLTL	42938	LACVAHWPS	42985	VPLGFNFAG
42845	LANHAKMGL	42892	STAFASSWE	42939	MSHEMFWVD	42986	AVSNWLETS

[2453]

42987	MWTVARMFP	43034	AAWGYNQFF	43081	ASLDQAEQN	43128	CQIPEGHEG
42988	ITHKYGWNL	43035	AAWVDKWWV	43082	ATENEHAMP	43129	CQVKRNATG
42989	QGRWACWYS	43036	AAYMSSECL	43083	ATNQRKAYT	43130	CSLRLIAES
42990	NHASFKRFM	43037	ACCCHENVL	43084	ATTNMFVDT	43131	CTALCYSRQ
42991	RNMALHHPT	43038	ACEHPLNYD	43085	ATVKYNCHW	43132	CTVIDINKN
42992	MWTKFNEYG	43039	ACGVTQMNP	43086	AVAFGSQED	43133	CTWLTQHQF
42993	HLLKKDASD	43040	ACQYDVETL	43087	AVGFCQYSS	43134	CVAIENVNF
42994	GQAGVHNLS	43041	AEDVDEGCD	43088	AVNHTPIWA	43135	CVRQTVTLC
42995	SFFREDYSF	43042	AEIRNVACW	43089	AVNYGHEGG	43136	CVSHYASQM
42996	RCTDCSYFY	43043	AFFMDHMWH	43090	AVTDKCDNP	43137	CWPFAINHI
42997	IKRAYDNNG	43044	AGINCQWCP	43091	AVTSYDNMQ	43138	CWTMFCPPT
42998	SQQHCQHDQ	43045	AHFPFFVYW	43092	AWTLNNNHW	43139	CWTQMKFWG
42999	QGDFFNICP	43046	AIFSREENG	43093	AYTDIWNFW	43140	CYHQYETRC
43000	CMQKYIPDV	43047	AIGIDHQAF	43094	CAAMAYEKG	43141	CYKSTQGAE
43001	PSLWKEKVW	43048	AIGNFHMAY	43095	CAEVQGETS	43142	DAGDENAPE
43002	YGTAMDSLY	43049	AIIWGAIQE	43096	CASGFGRLA	43143	DAKMGAEFS
43003	IHQCIVCAP	43050	AIQMPDVTY	43097	CAWVMQWHH	43144	DAKNECASW
43004	HLAPYCQQE	43051	AKDMHVANE	43098	CCHFDEFYM	43145	DALLHRHYE
43005	LMQSGQKDF	43052	AKLSNCAIP	43099	CDNSENFLT	43146	DAREYWRYQ
43006	EVDQYYSLA	43053	AKMGFDRFT	43100	CDTRWYCLC	43147	DASPKGGAS
43007	KIQMAVSLD	43054	AKSHHDDSA	43101	CEAEKVEQP	43148	DAWLHEQGT
43008	LINQWMHDN	43055	ALEYIRQGT	43102	CEGVANPWE	43149	DCNAADPGR
43009	HTHDVRVHK	43056	ALHHYPVYH	43103	CFPLCSFST	43150	DCQYVVNDC
43010	MPASNWVAE	43057	ALHVVMDVG	43104	CFTPHESMQ	43151	DCRMCERHV
43011	RTVHRDDFA	43058	AMDYLMEHV	43105	CGESWAGFN	43152	DCRWMSIMG
43012	TSRDGGWAY	43059	AMGNQRNYY	43106	CGPNVSMMP	43153	DDGKTYCLY
43013	KNETFIAST	43060	AMHAQQNRE	43107	CGVVESGMN	43154	DDRNVWSAD
43014	EVAGCATAV	43061	AMIASHHCQ	43108	CHKPYHLGI	43155	DDVRQGDCW
43015	MWTEFNEDG	43062	AMKFAGECA	43109	CHYEYSKFG	43156	DEIDIGREG
43016	QPGHTYNAP	43063	AMKFAGEGA	43110	CIETVHKHF	43157	DEMLKQFNG
43017	MESHHHAHA	43064	AMMHRNNET	43111	CIWPYQAIQ	43158	DFNVTAFRD
43018	DCVCCSELV	43065	ANAYQINGQ	43112	CKPLECQLV	43159	DFVGCSHWG
43019	TKQPVHFLG	43066	ANCMDFHIT	43113	CLCLIVDCE	43160	DGTPISCVG
43020	AMNWWDAPQ	43067	ANFYNEEAT	43114	CLLGEVMFD	43161	DHDDTRLHR
43021	YFMTCSTYP	43068	ANHHQTYDT	43115	CLTDSQDLP	43162	DHDKREIVT
43022	GSFMSCEKN	43069	ANQISVRGV	43116	CMHWGWIPM	43163	DHFYIAKQY
43023	VDQHGWNDM	43070	APCWEGSMR	43117	CMKIFKFNS	43164	DHGFEGGWP
43024	HHSWQPQKY	43071	APESARPAN	43118	CMMTYMHLG	43165	DIWHTQDME
43025	QFCTSCASN	43072	APIWNDSLM	43119	CMNCRHQCL	43166	DKAQIKMYG
43026	KMLLVMDGD	43073	APQVSKRFA	43120	CMQKIDHSW	43167	DKDEHDTQD
43027	ERTCIATMW	43074	APSTFANWG	43121	CMSVSTRWY	43168	DKKQSCDTS
43028	GVDNRQWLF	43075	APVNMENPI	43122	CMVCPGREE	43169	DKQLMMLHT
43029	NEAPYSTWQ	43076	AQQCSRLMF	43123	CNDRYMFAH	43170	DKRWQKMVR
43030	SGNIDVHTM	43077	AQRPCDDWQ	43124	CPCTFADEQ	43171	DKSQMHKWG
43031	AADNYSWVV	43078	AQTMSHCFW	43125	CPKPSDGSF	43172	DLCFMEQLH
43032	AAHRGECMA	43079	ARAAQLSYW	43126	CPLTSWHFE	43173	DLEVTSASQ
43033	AAVPRMWAM	43080	ARQCNPNCT	43127	CQHRTHCCM	43174	DLHDVVFVV

[2454]

43175	DMHGIENAR	43222	EDCNTNERS	43269	ETQSAAYDE	43316	GATCNMFSY
43176	DMPEDSIHW	43223	EDIQFDGDS	43270	ETTWQCIDP	43317	GAVWQQGSD
43177	DNKSIVMSQ	43224	EDKTQEMWI	43271	EVAQNTPHQ	43318	GDSQLNHWM
43178	DNNVQKCGQ	43225	EDLLMNSNT	43272	EVGDCMPQE	43319	GEVEFVPRN
43179	DNQPTVCHT	43226	EEPAMKIFM	43273	EVGQHEATI	43320	GFTTMYKHG
43180	DNRPPLFWQ	43227	EFEHFNYQD	43274	EVSTQTKLS	43321	GHEVQNWPM
43181	DPEYTVVLT	43228	EFRWQQSTP	43275	EVTYWQEWT	43322	GHHQRTTIN
43182	DQHPSFLGC	43229	EFTDQADWE	43276	EVYQGMLYF	43323	GHKHGIFAG
43183	DQNESGFGC	43230	EFTQDNWGQ	43277	EVYQQQKDH	43324	GHKTPSGPG
43184	DQWQATIAE	43231	EFYPSTPCS	43278	EWKMWYLQG	43325	GHMCDCTSM
43185	DRCQTVIDW	43232	EGNWWHEYA	43279	EWTDWRDEG	43326	GIDENHAGL
43186	DRKCMIPNK	43233	EGTDGDMFS	43280	EYEFEGNAP	43327	GIGCPQQWE
43187	DRQQMDCYQ	43234	EHESMCCGW	43281	EYRQAQSFS	43328	GIIPWMHFM
43188	DRTPLHQQQ	43235	EHKLWEGFA	43282	FAEQEHMAS	43329	GIKANIGTK
43189	DSAELHWPH	43236	EHYQKHCYL	43283	FAIQEYHIT	43330	GKCENKLST
43190	DSFTFLKKA	43237	EIFQAFLLS	43284	FATLQAAQD	43331	GKHKTYEPS
43191	DSGILETCD	43238	EIQRNHAIQ	43285	FATTTYADC	43332	GLQAFCFQS
43192	DSHAYWMYP	43239	EIWMMNGEY	43286	FCHYLVYDD	43333	GLQHYGEID
43193	DSNFSYLGH	43240	EKAELVRDL	43287	FFRCEKITA	43334	GMCENVSSF
43194	DSQFKMAAW	43241	EKCLLMNAR	43288	FISTFERHI	43335	GMPEDSIHW
43195	DTAAQKWEA	43242	EKLYSTLQP	43289	FITEIQDYT	43336	GMQYIHVTG
43196	DTDTVLQSS	43243	ELAEYGNGW	43290	FITRHPVYD	43337	GPPMYKDCM
43197	DTEWNHHNT	43244	ELEVKSLWW	43291	FKYMWNFSN	43338	GPYKNITFQ
43198	DTFQPHWYS	43245	ELEWMMHCA	43292	FLKNKDCLL	43339	GQGFRRQHR
43199	DTKHSSYDS	43246	EMMYKKLDQ	43293	FLLQYMQPS	43340	GQVEMFADR
43200	DTMTKEKRH	43247	EMTMEGQDI	43294	FLVQLSDGV	43341	GQVSSTWWH
43201	DTMVDEYDE	43248	EMVQGDTHE	43295	FMGFCGPYQ	43342	GRNNNTTEQ
43202	DVCSTFGWM	43249	ENDFLCAWT	43296	FMNPIGVNG	43343	GSAPFQCIG
43203	DVDQEQDPI	43250	ENMWGYANH	43297	FMSSRVINF	43344	GSHQFSGPH
43204	DVGIHEWCT	43251	ENNEYWHVE	43298	FPAFPKEAC	43345	GSNPPMMQY
43205	DVGLSGFER	43252	EPEMQNQNQ	43299	FQHDRHTYL	43346	GTGTCMWAE
43206	DVHLFGEFF	43253	EPRVHTWRK	43300	FQIPFCRSF	43347	GTQPVNWLI
43207	DVLFEGQGS	43254	ERDYCAYTS	43301	FQMHHKWDL	43348	GTVMRGMAL
43208	DVNPASHFC	43255	ERKMCFCHP	43302	FSAMYPHMN	43349	GVAPLMDQA
43209	DVQSWCRQG	43256	ESAWKHFVP	43303	FTALLHYDI	43350	GVEFQHNQA
43210	DWKPEMDGS	43257	ESKWNNPDA	43304	FTAYMQEFA	43351	GVGCEILGA
43211	DWRMKKSSY	43258	ESMLAIMST	43305	FTCPVSHFE	43352	GVHPDNWRN
43212	DWSHTQYGA	43259	ESNRMMGCG	43306	FTPWPQCCQ	43353	GVQLAMSDC
43213	DWVCMDWCQ	43260	ESRANATLR	43307	FTTQFIDGK	43354	GVQSLHFNS
43214	EADDHDQNR	43261	ESVHEGTSS	43308	FTWTQCVVP	43355	GVTCKDMYH
43215	EAHISAEFK	43262	ESYPNNHDL	43309	FVEVRMMCH	43356	GWSQEGHFA
43216	EAIFHDVEM	43263	ESYYKPDFQ	43310	FVKMRLSDD	43357	GWWGLVHNQ
43217	EAKTSYRMW	43264	ETFFLPNCN	43311	FVTVWRTTG	43358	GWWYYGQCR
43218	EARETAIMC	43265	ETMCGNELG	43312	FVWNTMSTE	43359	GYASHQHWQ
43219	EATQHETRV	43266	ETMDREEDK	43313	FYLPMEHMD	43360	HAGCSTRTG
43220	ECMCVQFCQ	43267	ETMFYHIDH	43314	FYTETNHAP	43361	HAIHKWPRW
43221	ECTMIEYMG	43268	ETQAGAPCN	43315	GAGLYMVMG	43362	HALGRVANR

[2455]

43363	HAMLNEICD	43382	HPPPLPWQS	43401	IDWVGHMGA	43420	IMGNYIQAN
43364	HATQRSADA	43383	HQDQPEWAK	43402	IECNYCKFY	43421	IMMDWEGNW
43365	HAYLAGWCN	43384	HRIAEICRP	43403	IECYKGMFC	43422	IMWVLNELE
43366	HCQDLQIGM	43385	HSEMLDLQE	43404	IEPEGHLVT	43423	INEGIGRHF
43367	HCWAQYLRF	43386	HSQGFDVTQ	43405	IEYPNKWRW	43424	INHGTTATF
43368	HGARGEQFA	43387	HTADMRAYD	43406	IFAVMADMS	43425	INMPAVMKF
43369	HGCYVAHCG	43388	HTLFHRDCA	43407	IFHPQWKQG	43426	IPGTRIAWA
43370	HGEMCAFFQ	43389	HVATKEECT	43408	IFKHCKHDE	43427	IQCYYKGEQ
43371	HGRCIWPMW	43390	HVEPLEETM	43409	IGAGFYYKP	43428	IRFGSREAK
43372	HGYADCKAH	43391	HYDPYNRTL	43410	IGEAFASTG	43429	IRGQANSCV
43373	HHCNTSEFV	43392	HYYTPDRMT	43411	IGLLTSLHD	43430	ISEVENWYD
43374	HHEIKVIFE	43393	IACMENLSL	43412	IGRVKCART	43431	ISNKLTMPV
43375	HHILNVWTE	43394	IAFWIDGQC	43413	IGVMPGIVA	43432	ISTMIEYSN
43376	HHNFRMAMN	43395	IAHWCHKCW	43414	IHSAHHWAS	43433	ITEYAQQDC
43377	HIAIVCLMD	43396	IATGRESAI	43415	IHSQMIEAP	43434	ITHHRDNWF
43378	HLLHVCLEK	43397	IATGWESAI	43416	IIGRFESMS	43435	ITKIMRQAE
43379	HMIRGYCDY	43398	IAWEKCDAW	43417	IINPLRHGD	43436	IVMPVAAFH
43380	HNGMDTQWH	43399	ICHSGGWST	43418	ILDPQHWSM	43437	IVQCFVDSQ
43381	HNHHECRQM	43400	ICRREWNPV	43419	ILTHVSDAC		

[2456] 表86

[2457] 驱动肝脏组织嗜性的AAV5 VP1衣壳多肽中581至589区域的序列

[2458]

SEQ ID NO	581-589 序列	45455	QCGSMQGVY	45475	VFKLISNQA	45495	ECHVCTHKA
45438	AFEKVNCPD	45456	REAEYWRDQ	45476	VVVRSNECA	45496	ELPRYYGQM
45439	ILIFIQNCP	45457	SSFYAQREG	45477	YHNQDRHYL	45497	ETHRRCKMS
45440	ASQRTHHVD	45458	VIISRHMDS	45478	HIKVEHEEV	45498	FAKHCVMTA
45441	ATNLINKLF	45459	VTHLFGSQA	45479	ACASTSPNA	45499	FDEALWKNN
45442	EVLLGCQNY	45460	AAEAGIHQA	45480	AHSRMDWTP	45500	FTMRMASGF
45443	FYMLSNQGS	45461	ACQSYWNEA	45481	AIGRRRLTH	45501	GTHTQQWHH
45444	GSVGMLDGQ	45462	DATFCTYMR	45482	AKGHDKACC	45502	HVFFAPLEG
45445	GVNIMFRQQ	45463	DKTSRLADQ	45483	AMGSLMRAS	45503	HYMGLLCGC
45446	KGEPGMNYY	45464	KANMCEKIC	45484	ARNSMPFGM	45504	IFDVYARWT
45447	KIAVYADWA	45465	KPNNFDTQF	45485	CGTGEDCED	45505	IHIGKNDNC
45448	KNAWMDCVN	45466	NINKFTCHE	45486	DKCNVTRAD	45506	IHKFCFKGY
45449	KNGAGDAHY	45467	NSAIEYQNE	45487	DMSDCYEHA	45507	ISGFTKHEW
45450	LAKMYWDMG	45468	QATPSEWEN	45488	DNILRDVMA	45508	IVKQCYDCH
45451	LTQSFQTYH	45469	QSLMCWNMG	45489	DVANHITQM	45509	KAANDHHVA
45452	MIVQQSTGC	45470	SARYQIEGC	45490	DVDWNMGVE	45510	KHVILRQLP
45453	NRDYYRGNY	45471	SIVHSGLSF	45491	DVNQYTMRH	45511	KIVSNESMT
45454	QAVQLWSHV	45472	TFYPTQNQH	45492	DWRNKRNHA	45512	KVDGCQKLC
		45473	TSNWLNFFP	45493	DYVFWGFPC	45513	LMYGVGTCF
		45474	VAAQSDTVN	45494	EAAKLVEPT	45514	LNNHHGTGS

[2459]

45515	MFQWRWIDG	45556	ETLKMYWEK	45597	ADSQSRVMP	45638	AKGLTHLTC
45516	MGFPKMSAW	45557	GFSPISSPC	45598	ADYHEGMFP	45639	AKLWRYRWI
45517	MQVVLQHTS	45558	GIDTISTSI	45599	AEHQHLACV	45640	AKMLKDRAA
45518	MYYQIENDP	45559	GSEWKTFGD	45600	AEMRTMGSQ	45641	AKMPYMSTH
45519	NSICMGGWK	45560	IMMCVWRSF	45601	AFAGTLQHP	45642	AKQMTAAQM
45520	NVCTKRENE	45561	KAFDNELHN	45602	AFAWAHCPV	45643	AKSCEESRP
45521	NYCKYNKMR	45562	KGEMEMSMA	45603	AFMPQSWTQ	45644	AKSEFSVSI
45522	PESVKMDCN	45563	KKCDTMIGQ	45604	AFMRNFDYP	45645	AKSGWTLME
45523	QVYYVHTER	45564	KMEHMWKMP	45605	AFMSTGMLA	45646	ALAHNFNGM
45524	RASHDGCCD	45565	MNGFCFGEA	45606	AFMVTASNL	45647	ALDTFCRWP
45525	RIFWDFHAT	45566	MVKQAQTMC	45607	AFNHWHYRN	45648	ALGLNSNHG
45526	RVPRCVYEC	45567	NGKQFESGP	45608	AFRSVVIKW	45649	ALMTWVRDC
45527	TDALTDSMH	45568	NSASYCNPT	45609	AGCNRCWKQ	45650	ALWYSKKEC
45528	TEDDRWRHA	45569	NTLVNMESP	45610	AGDKWMASF	45651	AMASVRGGD
45529	TIWERQSFH	45570	TNSVMQNME	45611	AGGFWCEPF	45652	AMENWYSYQ
45530	TQGMYVESY	45571	TTGPKHCSV	45612	AGKWWWQYH	45653	AMLDGMTSG
45531	TSNIFQRHG	45572	YSTYSEKSC	45613	AGVTLPVTL	45654	AMLMSKEVA
45532	TVNVDSQYN	45573	YVETTKVWP	45614	AHIAVFDCD	45655	AMMQGVSNQ
45533	VASSHRESQ	45574	TSERHSCMV	45615	AHLDMWWSN	45656	AMNVIASHF
45534	VIKVNTVNP	45575	ATAFKMLNT	45616	AHMCTQQSG	45657	AMSCYTLWC
45535	VSGMLAACN	45576	SPNGRGLCG	45617	AHMTEIAFA	45658	AMSFTQFCS
45536	VVNVWGCAP	45577	AACWTVPQH	45618	AHNMDVWCI	45659	AMSPIVDAP
45537	WNIVLMFYC	45578	AAESQGHRH	45619	AHPPTQYNH	45660	AMSRLMMFS
45538	WWVVVPCFM	45579	AAHCIPNHD	45620	AHTLRQFWE	45661	AMVEPQCQG
45539	YCTGLGSMD	45580	AAISSGFPA	45621	AIACAHMGG	45662	AMWFCSVGD
45540	YFALNDITH	45581	AAKVFPDGM	45622	AIEFHEAFT	45663	AMWRLNNVF
45541	YNLQWFKHY	45582	AAMCCHLGV	45623	AILFLKGND	45664	ANCTQENQT
45542	YSDNGRRFD	45583	AAQCCTFGT	45624	AINGTGTFT	45665	ANECKQNVE
45543	YSIIWRVYR	45584	AATASGYAS	45625	AINWPMHIE	45666	ANHVTDTSC
45544	YSMDVMMSI	45585	AAVFNDVEK	45626	AIQHPNQWC	45667	ANKMIVLNH
45545	YVHKPPDQG	45586	AAVGINTRE	45627	AIQNEENPG	45668	ANKSVGEWR
45546	EVRSTRKNH	45587	AAVGSNEFA	45628	AIQTTETEE	45669	ANLFYSDRD
45547	LDGPIGMGT	45588	AAWMGNAVM	45629	AISDKQPWR	45670	ANLLGHYQA
45548	GAYSLHDWR	45589	ACAYSVASE	45630	AITPCLEFK	45671	ANMLTNCDT
45549	IMVERWQHA	45590	ACCKAGFLV	45631	AIVEGVKHY	45672	ANMNCDKFS
45550	MCDNFPSAH	45591	ACEYKVVNQ	45632	AIVPIWTTN	45673	ANNVHCVKK
45551	NGMDKPVDA	45592	ACHNNVYHG	45633	AIVQFDVEE	45674	ANNVTQAER
45552	AAWQSEIAN	45593	ACKANTAQM	45634	AIVTASCPP	45675	ANQCHMGYT
45553	ACILDGFTL	45594	ACKFTEISM	45635	AKCCSQLDY	45676	ANTRFPMKD
45554	CAYNAPDWP	45595	ACLTKNWTR	45636	AKCYKEWWP	45677	APDTWVPVY
45555	EFTPEHRCW	45596	ADNVMHWWT	45637	AKDQYGLLS	45678	APGNHNNWW

[2460]

45679	APHCEKENG	45720	ATYKIDTSC	45761	CCSNKSRFV	45802	CMFWANCSQ
45680	APISWQMAL	45721	AVAIRSWCG	45762	CCSPLCQQP	45803	CMGSGFAEL
45681	APSMRKIIP	45722	AVAMEFNVV	45763	CCVPMMNVC	45804	CMKMWDCGK
45682	APSNPRKVA	45723	AVDPMCFDE	45764	CCVTAWKNQ	45805	CMMPVLEEC
45683	AQAEKPCSY	45724	AVESICSTE	45765	CDAIYVSCG	45806	CMRWFLLGS
45684	AQAKEWKTK	45725	AVFIPGQFW	45766	CDCLFMLHQ	45807	CNEVPSVMP
45685	AQCYKFNGH	45726	AVGMCQMDL	45767	CDNGYETNE	45808	CNSQKSYAH
45686	AQHTVWRFA	45727	AVLIAGAGH	45768	CFCQGFRKR	45809	CNSWNCAYA
45687	AQLPFCQQD	45728	AVLKAEVVK	45769	CFEHWADFE	45810	CPDVGFQSG
45688	AQQKGGITF	45729	AVPFPWGCC	45770	CFEKKQTDG	45811	CPHLFVFDW
45689	AQSAISVMV	45730	AVQLWWWHN	45771	CGEQKDQIG	45812	CPPSNNPCE
45690	AQSAWHSWP	45731	AVRFYSTPI	45772	CGHLYNRDF	45813	CPSRTWAIE
45691	ARLACCSAD	45732	AVRWSQTPG	45773	CGLDYAMCQ	45814	CPTIIVKDW
45692	ARMYYDDQP	45733	AVTVVGTQI	45774	CGRHWKEWE	45815	CPVANTTHP
45693	ARRLWQNDQ	45734	AVVAQCQSW	45775	CGRQCKACF	45816	CQEFNSHTD
45694	ARVQENHVG	45735	AVVGWTAGQ	45776	CHCQWMCQN	45817	CQNFSSLPW
45695	ASGMQGSWH	45736	AVYIMTQLQ	45777	CHGAHQKLE	45818	CQSVSANHI
45696	ASIMQTKGI	45737	AWAEPDMRW	45778	CHMPWHIHE	45819	CRARRNPRL
45697	ASKPGIIFD	45738	AWDAESCYK	45779	CHMSTVMKV	45820	CRIRKRHIT
45698	ASMPFLSYT	45739	AWEKIHRSP	45780	CHQETFLDA	45821	CSCDEMEID
45699	ASNILDLPE	45740	AWITCLIAP	45781	CHTGLMDAM	45822	CSDACQAMV
45700	ASQWKIHGA	45741	AWMHDTCTK	45782	CHVQFDGFG	45823	CSDMLTKCN
45701	ASSPFSQVR	45742	AWRSVASNH	45783	CICREGKFY	45824	CSEHENLGI
45702	ASTCKKLAM	45743	AWTHHFTLD	45784	CIMACQDRT	45825	CSEKFNLHC
45703	ASYSILKQR	45744	AWWRVTHEQ	45785	CIMGMQWQF	45826	CSETQVSTG
45704	ATAHIAPPS	45745	AYAPLSHNN	45786	CIRHQNHYW	45827	CSQPHQMAS
45705	ATDTMQFIW	45746	AYCEIPTNP	45787	CITAEQCGE	45828	CSQRCMSFG
45706	ATIAQWCVH	45747	AYDHVLESR	45788	CITPQFTSS	45829	CSRSDLLCG
45707	ATKLGCSRI	45748	AYEYRNHEF	45789	CIVKRMECF	45830	CSVWCIYYR
45708	ATLGHVMMC	45749	AYFPNSWGQ	45790	CKAMTLAEM	45831	CSYPAHRFH
45709	ATLTNQHAE	45750	AYNRVWSCG	45791	CKCYEMEVS	45832	CTEVMAMGT
45710	ATMMKWDAN	45751	AYPQCFQNP	45792	CKEPHTMHE	45833	CTFSSDKSD
45711	ATMWLEDQC	45752	AYQAKYDLA	45793	CKKPLGTRI	45834	CTKLVNWYT
45712	ATRAMQVLE	45753	AYTETSYLH	45794	CKNEYRIHT	45835	CTLRCIMHT
45713	ATRLAFHQG	45754	AYTLCLCNE	45795	CKSVNPRMM	45836	CTMRSGVGT
45714	ATRLITYAV	45755	CANCPMFPT	45796	CKVEWWHGN	45837	CTPTILRGV
45715	ATSFTQFCS	45756	CAQFKMTYF	45797	CKYEISSGL	45838	CTQSQCHLC
45716	ATSGQTYSV	45757	CARQPMPPA	45798	CLVSCPALN	45839	CTTEVRGTP
45717	ATSMCMHPR	45758	CAYAIMPCE	45799	CLYVPLFAD	45840	CTTMKMNQA
45718	ATSPHLAWG	45759	CCFQFELEC	45800	CMEGGLMDW	45841	CTWFIQGPK
45719	ATVKFSDCV	45760	CCIFIWWLY	45801	CMEMMKWAN	45842	CTYLAKLKC

[2461]

45843	CVAAWGENG	45884	DHLKIRGAY	45925	DMWDEGDGD	45966	DSTRIITQY
45844	CVDNPFVLV	45885	DHQKRNVIQ	45926	DNADMFNIG	45967	DSYTMNWFY
45845	CVFMGMMWD	45886	DHQLVLVSG	45927	DNAKRPMTI	45968	DTAAEVNMQ
45846	CVGRWWNFY	45887	DHTVLFEGQ	45928	DNENAIPMG	45969	DTAEYRAGY
45847	CVHISDFGD	45888	DHWITGPCR	45929	DNGNMFQQH	45970	DTDCARAIG
45848	CVIQGHSIE	45889	DHYLGHASY	45930	DNKLTEPCN	45971	DTDRQVKQS
45849	CVMAIESRG	45890	DICWKGWCF	45931	DNKYYMSKC	45972	DTFEYTNAF
45850	CVNRQHRIN	45891	DIDRFTDVQ	45932	DNMEFESPG	45973	DTKFYKIAY
45851	CVPASGGWD	45892	DIECKEIAI	45933	DNNAMLVCT	45974	DTLQRACML
45852	CVTMKEFAM	45893	DIEITMQEN	45934	DNPVTRCTD	45975	DTLWYVNFA
45853	CWRQYVAGQ	45894	DIGPQDHWP	45935	DNSWFVQGG	45976	DTMMMPSSE
45854	CYCTCEQPD	45895	DIHWMDHDF	45936	DNVYTMHQW	45977	DTMTTCYNS
45855	CYELTMPLD	45896	DIIEKVNEM	45937	DPMYIPVNH	45978	DTQNMKSSK
45856	DADRNNMSV	45897	DILERMNRA	45938	DQCLCHTQA	45979	DTRWWGMYA
45857	DAEMLCYDY	45898	DINMNAQER	45939	DQGLWENIP	45980	DTSSHECCD
45858	DAFWSNLHL	45899	DIRKDQHKN	45940	DQYFEGWRK	45981	DTVHSGWQD
45859	DAHWGALSV	45900	DIVIKKCNA	45941	DRCVMTDVL	45982	DTVNQQKTW
45860	DAIPKKWYG	45901	DKASNQHNG	45942	DRFMEFTGA	45983	DVCKKGQCV
45861	DALANNESR	45902	DKFAAHALA	45943	DRFWSKYKD	45984	DVCSQGQVP
45862	DALSMLRAP	45903	DKGHHNSMH	45944	DRGMPTNWC	45985	DVEMMIQRA
45863	DALTNRASH	45904	DKIDFSSVM	45945	DRGQSFQMF	45986	DVEYYDHCF
45864	DASMMWSRG	45905	DKKFGQFMS	45946	DRGTVNRDC	45987	DVGYSAVWA
45865	DASYEYQSL	45906	DKKWEEICA	45947	DRSILHAHP	45988	DVKWGTSNL
45866	DCAQTGEYG	45907	DKLELIGLD	45948	DRTIIQNIP	45989	DVLPMEWCV
45867	DCKNHYQSS	45908	DKNFWVHWF	45949	DRTSYPLHL	45990	DVLWGNRDS
45868	DCMWYNRID	45909	DKTEMVIID	45950	DSAWYYWHP	45991	DVMAMAPSC
45869	DCPRSKMMT	45910	DLASVGRWN	45951	DSCHSRIIL	45992	DVNAGTRFP
45870	DCPTYYWRD	45911	DLCKMKEDK	45952	DSELVRDYI	45993	DVTDCKWGF
45871	DCSCHELAL	45912	DLFRQERLF	45953	DSIGLFRGG	45994	DVWYPEFIG
45872	DCYHICVTP	45913	DLINWWHLQ	45954	DSKIAHGRQ	45995	DWAHWAAYP
45873	DDELCPQCN	45914	DLKPWNPAW	45955	DSLRKYQCR	45996	DWCTDMTRK
45874	DDHQSTWMN	45915	DLRLKVISR	45956	DSLTVNMAH	45997	DWHNLVEYP
45875	DDRCVQAAK	45916	DLSRSFLAQ	45957	DSMAAHCAV	45998	DWLNLGRYM
45876	DFAVSTSFD	45917	DLVIIYMFH	45958	DSNGIAHVP	45999	DYFREHMCP
45877	DFCMKEKAV	45918	DMELQFMYA	45959	DSNKWAGLG	46000	DYKLRARNA
45878	DGCVHNQAI	45919	DMGNYPCFE	45960	DSNTHLPIG	46001	DYNGLSVPI
45879	DGPICKGIY	45920	DMMLRHCVF	45961	DSPQIVDVY	46002	DYYQWDGQI
45880	DGTHPCWNG	45921	DMRSNLPWH	45962	DSQGINKLH	46003	EADRWSSPN
45881	DHCKTALPN	45922	DMSHECVDA	45963	DSRHSTFCL	46004	EAGRQAWTD
45882	DHERQRWLH	45923	DMSHECVGA	45964	DSRHYANFM	46005	EAHLDWIYP
45883	DHIRQHTCS	45924	DMVDLEMYP	45965	DSTPDRSIW	46006	EAHWFTNYN

[2462]

46007	EAKPWSGYL	46048	EISPHKILP	46089	EQIHQLRFG	46130	EWMTCGVSG
46008	EAKRYHVDG	46049	EITGGHASW	46090	EQKWQWCQY	46131	EWSMLQYNH
46009	EANMKEGLI	46050	EIVWYFQDT	46091	EQQMLLQLS	46132	EWYIANISD
46010	EATIKALVF	46051	EKAMRNGQA	46092	EQYMVNTRD	46133	EYELTQPAP
46011	EATWTHADE	46052	EKCCCLQWE	46093	ERCPWSVGE	46134	EYIFENEMC
46012	ECDPRQMMC	46053	EKEIKMLFG	46094	ERENSLWMS	46135	EYLLCRGEA
46013	ECHQGSNME	46054	EKGPYWHQH	46095	ERFLLEAVE	46136	FADHFEAMF
46014	ECMDQQATW	46055	EKHAMECVS	46096	ERGKWHALS	46137	FAEHEIEHS
46015	ECMQQCFVR	46056	EKIAPNHQA	46097	ERQPFTMQW	46138	FAIIRNDSH
46016	ECQPFATVT	46057	EKKMWVYTG	46098	ERSWTHAHW	46139	FAIMSSDHP
46017	ECVYNPANN	46058	EKLMAYIML	46099	ERTTKAEGY	46140	FALNTEISA
46018	EEKYKDAQW	46059	EKLPIQMFN	46100	ERVARLQFT	46141	FAPFLKAGH
46019	EEPGSQMTV	46060	EKMPWLISN	46101	ESAKFVCEY	46142	FAQMNKDRC
46020	EERPWKRDL	46061	EKNWYQHIW	46102	ESCEDGIRT	46143	FARHPLKCA
46021	EETEGIEPH	46062	EKQTRNQDG	46103	ESGFRKLRV	46144	FASQHAVQA
46022	EFDAAVIFK	46063	ELHHPIRRS	46104	ESHMHDHWI	46145	FCFWMPTTN
46023	EFGYEHIVP	46064	ELMASNRCH	46105	ESILVCATP	46146	FCGKLTHVG
46024	EFRNSSCTN	46065	ELMFVMHGT	46106	ESIQLRSSQ	46147	FDWRCGCWS
46025	EGAHNKFDR	46066	ELPMMNMAQ	46107	ESSFVHEEQ	46148	FEQPPMHHQ
46026	EGCIPGGWS	46067	ELSFKTAQG	46108	ESTEPMSAT	46149	FEVTMTVPE
46027	EGGSCNTQM	46068	ELTMANCMN	46109	ESTLKTTYA	46150	FFQRFLSAW
46028	EGNFIMGCS	46069	EMERKNMFE	46110	ETCTEPKYW	46151	FGAMTSDTI
46029	EGNQEGLNM	46070	EMGCDNKSF	46111	ETEMTSWQN	46152	FGAYYNDRG
46030	EGNREGLNM	46071	EMHLYHICH	46112	ETKAVISAQ	46153	FGETYNHSH
46031	EGNYMAGLG	46072	EMTSTMKSH	46113	ETLERDWDY	46154	FGEWITFYF
46032	EGWHFPAME	46073	EMYRALGHG	46114	ETMKDVRFM	46155	FGHKIEAPH
46033	EHDGTPFNE	46074	ENGPMCDSN	46115	ETNQLSGYC	46156	FGQHVWQGF
46034	EHDHQPSYA	46075	ENHSPNVDP	46116	ETPKFNWSN	46157	FGRYFQFDL
46035	EHERFCQGR	46076	ENILAIPAL	46117	ETRTDVPML	46158	FGWWYKTFT
46036	EHHRRGAFK	46077	ENKWQPKQA	46118	ETTAAFKPG	46159	FHAPTECMA
46037	EHMFNKDCT	46078	ENNSRQLIN	46119	ETYDNNLQA	46160	FHFQFTMDK
46038	EHQRGHACN	46079	EPCMRNLFG	46120	ETYVCWFGS	46161	FHHLFDCNL
46039	EHSYLLHDT	46080	EPIRTEGAE	46121	EVATSYARV	46162	FHLACIHAQ
46040	EICSEINMK	46081	EPLCHQIMP	46122	EVDQAPMMI	46163	FHMTRYKMT
46041	EIEQEEGGC	46082	EPLVDEFGG	46123	EVHRSELQR	46164	FHRPSEGQE
46042	EIGLYWIKC	46083	EPRECFHIE	46124	EVIGFLYHA	46165	FIRLFMKLN
46043	EIGYKQMYF	46084	EPWVMKLNK	46125	EVMAMCDRY	46166	FKECTNGMH
46044	EIKASHNYD	46085	EQADYRAVQ	46126	EVMPMCDRY	46167	FKMATAECP
46045	EINQGDEDG	46086	EQERYEWDP	46127	EVQLTYLHE	46168	FKPCKDDDV
46046	EIQSVMPQP	46087	EQGCQLVTT	46128	EVVTKHDSR	46169	FKVTEYELP
46047	EISNLRQGF	46088	EQHPRMNFH	46129	EWCKQWMRF	46170	FMDEIAVTM

[2463]

46171	FMGFMGVMG	46212	FYPRSWPQA	46253	GIANTSDSY	46294	GRSLSQSEW
46172	FMQIWNKYS	46213	GAARYGYNI	46254	GICCILWDR	46295	GRYQSSQFC
46173	FMRMEMMGQ	46214	GADSSGGWN	46255	GIDSMVIQR	46296	GSASHNVWY
46174	FMRPFLEPD	46215	GAFIRDFDI	46256	GIHEIMVQR	46297	GSDYWLCDK
46175	FNFCFNHQF	46216	GAGSVQEIH	46257	GIHMVTAMS	46298	GSFYAQYEE
46176	FNGIDQMRY	46217	GAGTTMRPD	46258	GIKFTQQDL	46299	GSFYCQWNN
46177	FNMLLQVNR	46218	GAHWWGGCV	46259	GILRQIEKP	46300	GSGKMWQED
46178	FNTTCQMSI	46219	GAIPLLNWL	46260	GISICGQNW	46301	GSLLFNPEE
46179	FPICNQERI	46220	GAKMQIAFW	46261	GKDTPNSFP	46302	GSMKEMMMR
46180	FPQPPCTYA	46221	GAKQHMPCY	46262	GKGEVLAVA	46303	GSMVASSWC
46181	FPTNQNDPH	46222	GAVNSEVCE	46263	GKGTMEFGN	46304	GSNQRDCSR
46182	FPTQMHYNA	46223	GCCWSYYCG	46264	GKGYGQHDG	46305	GSRPTPADW
46183	FQAKMHTLA	46224	GCDTTKFDA	46265	GKILWMSAR	46306	GSSHYMEFY
46184	FQDWPVTWC	46225	GCFIYWKDS	46266	GKMQVLHSA	46307	GSYTKPTNV
46185	FQQQTGTRC	46226	GCVKDRCWM	46267	GKNMLIDGG	46308	GTAHNVFMG
46186	FRFSTQHEA	46227	GCWNERCQN	46268	GKPYTQIIR	46309	GTCLDQTGC
46187	FRLNLLKTN	46228	GCYWACWLT	46269	GKSTCKRDV	46310	GTETHCDFE
46188	FSEVTLQES	46229	GDNYTAWVG	46270	GLEMRNQAM	46311	GTGAFTQVP
46189	FSKTGHCKF	46230	GEFMMFDRG	46271	GLFQQAMQG	46312	GTIAFSDNN
46190	FSQCPITCC	46231	GEGWNDDLG	46272	GLGCLMQGN	46313	GTIPMNCFW
46191	FSQGPGMHF	46232	GEREYWNWE	46273	GLSVGQRIE	46314	GTKMFQESI
46192	FSQHQEHMR	46233	GFACLDHFD	46274	GLYERLACC	46315	GTMFWGWTG
46193	FSQRVEKTW	46234	GFDIVLKDF	46275	GMHYSQCQH	46316	GTNNIINVR
46194	FSSEAADYD	46235	GFELNMRIS	46276	GMKAVFIGD	46317	GTQHFWMDH
46195	FSTFTKPSP	46236	GFGNTFCIT	46277	GMWVNHGAK	46318	GTRWIHDMA
46196	FSVPLGATT	46237	GFGTVSMER	46278	GNEHCVKYA	46319	GTSQEKRGQ
46197	FSVSAQHSC	46238	GFGYMWCKG	46279	GNEKIERYD	46320	GVAAVGLYK
46198	FTFRVGEVN	46239	GFLLTYGED	46280	GNEMPNYGG	46321	GVIVPISRS
46199	FTHGSVLHD	46240	GFMIGNQYL	46281	GNHVNYNMH	46322	GVLKQANQQ
46200	FTMLGMEAI	46241	GFRVIISNW	46282	GPEKLTWGP	46323	GVQNTALPR
46201	FTNPHAFYP	46242	GFSCLICFD	46283	GPPHYRNPE	46324	GVTSPVGTC
46202	FTQHIVGET	46243	GFTQTCMTN	46284	GQFIFDEFV	46325	GVVDNQQCL
46203	FTTKYVYFD	46244	GGLLYIECY	46285	GQGFMLWNG	46326	GVWPHYENK
46204	FVDHYYGHQ	46245	GGTIDGHVG	46286	GQGKNWVSW	46327	GWCITKARN
46205	FVNHTRCNP	46246	GGVKFNNWW	46287	GQGMCDRKT	46328	GWDHFMQNT
46206	FVRHDVIGQ	46247	GHMFVNIDQ	46288	GQLNCMKGS	46329	GWMDYIAAH
46207	FVTHWEYYP	46248	GHMMMMECW	46289	GQQENPVFT	46330	GWYNWMIQY
46208	FVTMMQIAF	46249	GHSLTMRAH	46290	GQTSPLMYP	46331	GYESKNTWY
46209	FVTMRYEGH	46250	GHTARDIHA	46291	GRAPMCAAA	46332	GYEWVQEIE
46210	FWTSYTTNP	46251	GHTHSQPAA	46292	GRLVVHQFK	46333	GYFPNHWCK
46211	FYGQNYWWY	46252	GHTPGGSFY	46293	GRNMFLWWW	46334	GYHCTGAHS

[2464]

46335	GYICEGTCH	46361	HEWAGELQT	46387	HMEMPYKLG	46413	HSMQTKFDW
46336	GYNERFEHK	46362	HEYGFGEAR	46388	HMKCDVWQP	46414	HSSALFDME
46337	GYNPMMSAA	46363	HFANICVRG	46389	HMLKKEDMN	46415	HSTQAGGMY
46338	GYSQKPRTP	46364	HFEHLWCQS	46390	HMTTVNLQY	46416	HSTTACVHN
46339	GYTYNQQVC	46365	HFGHYLGMF	46391	HNGKSVLEW	46417	HSVMWGKEN
46340	HADAENQVM	46366	HFISHDYHN	46392	HNKMTYNYE	46418	HTDKWWRDI
46341	HAIPHTMWH	46367	HFMPNNFLA	46393	HNNMFFWDR	46419	HTEKGGEQA
46342	HAIQEGTVP	46368	HFVNSYCWW	46394	HNRAHENDK	46420	HTIHGAGLI
46343	HAIYWIPQV	46369	HGAQFARCS	46395	HNSVVNDHE	46421	HTIPVPMAA
46344	HAMATECGP	46370	HGNEVDHSM	46396	HNTQNGFAD	46422	HTNTCLFEE
46345	HAMLSHGCC	46371	HHAVMNKMD	46397	HNWYKRIVQ	46423	HTQIIGGSD
46346	HANESISMT	46372	HHEVTTKFQ	46398	HPGIDLVCG	46424	HTQWFNMGE
46347	HASYQSSFG	46373	HHFHSGCCW	46399	HQCFWIYGS	46425	HVDQYYRSA
46348	HATSQELCQ	46374	HHPPELDRM	46400	HQFQFYQWM	46426	HVDTEKGRE
46349	HATTFEWDS	46375	HHTDRNGGD	46401	HQPPIMNQG	46427	HVELSQLHS
46350	HAYALAVEH	46376	HHVLKGDMA	46402	HRCTMQPDI	46428	HVEWVNAEK
46351	HCCSANNSP	46377	HIDGATVSG	46403	HRCTPYNRD	46429	HVLPVCYRM
46352	HCELISRIA	46378	HIKWDCWCP	46404	HRDHDITDF	46430	HVMFPWGMW
46353	HCELLPVNE	46379	HINAFDNIR	46405	HRFSPPMET	46431	HVMYQHWMM
46354	HCFMAHCSY	46380	HISWWPACG	46406	HRHFEMSVI	46432	HVNIEHQGT
46355	HCMNWMING	46381	HITTLYMEA	46407	HRMVIRKSQ	46433	HVSDCCKSQ
46356	HCQILPGWA	46382	HKLVAHERH	46408	HRVQCDEAP	46434	HVVKYPLSG
46357	HCVNEHAWE	46383	HKPMCDRVF	46409	HSCTSGNPG	46435	HVYRGNGME
46358	HDGFPWADE	46384	HLMGLCAMS	46410	HSESIDVWS	46436	HWDECNYVG
46359	HERNWPFAC	46385	HLQSISHCH	46411	HSETQRHQM	46437	HWDQFNGQF
46360	HESCLKLAG	46386	HLVGMSLTQ	46412	HSGPWSNCE		

[2465] 实施例46

[2466] 用衣壳包裹治疗性有效载荷的AAV5源性病毒体治疗脊髓性肌萎缩症(SMA)

[2467] 此实施例描述了用具有本文公开的任何一种工程化CNS嗜性变体AAV5 VP衣壳多肽的变体AAV5源性病毒体治疗脊髓性肌萎缩症(SMA),其中变体AAV5病毒体衣壳包裹治疗性有效载荷。将编码AAVRep、AAV5源性变体Cap和辅助蛋白的多核苷酸序列以及治疗性有效载荷在细胞中转染以产生变体AAV5病毒体,其中编码变体AAV5Cap的多核苷酸序列在从对应于VP1衣壳多肽中的581的位置至对应于589的位置的区域中包含至少一个突变,并且其中编码变体AAV5Cap的多核苷酸序列编码SEQ ID NO:34991-SEQ ID NO:37437中任一者的氨基酸序列。治疗性有效载荷是靶向由与SMA有关的基因编码的mRNA的指导RNA,或转基因。指导RNA所靶向的mRNA是由编码运动神经元存活(SMN)蛋白的基因编码的mRNA。转基因是SMN。将衣壳包裹有效载荷的变体AAV5病毒体施用于受试者。受试者是人类或非人类动物。施用途径是全身施用途径。全身施用途径是静脉内施用。与衣壳包裹治疗性有效载荷的野生型AAV5病毒体相比,衣壳包裹治疗性有效载荷的变体AAV5病毒体在施用于受试者后表现出增强的脊髓嗜性。在施用于受试者时,与衣壳包裹治疗性有效载荷的野生型AAV5病毒体相比,施用较低剂量的衣壳包裹治疗性有效载荷的变体AAV5病毒体。在施用于受试者后,SMA的至少一种症状得到缓解或受试者得到治愈。

[2468] 2.某些序列

[2469] ＞AAV5

[2470] VP1,来自Chiorini等人,J.Virol.73(2):1309-1319(1999)

MSFVDHPPDWLEEVGEGLREFLGLEAGPPKPKPNQQHQDQA RGLVLPGYN

YLGPGNGLDRGEPVNRADEVAREHDISYNEQLEAGDNPYLK YNHADAEFQ

EKLADDTSFGGNLGKAVFQAKKRVLEPFGLVEEGAKTAPTGK RIDDHFPK

RKKARTEEDSKPSTSSDAEAGPSGSQQLQIPAQPASSLGADTM SAGGGGP

LGDNNQGADGVGNASGDWHCDSTWMGDRVVTKSTRTWVL PSYNNHQYREI

[2471]

KSGSVDGSNANAYFGYSTPWGYFDFNRFHSHWSPRDWQRLI NNYWGFRPR

SLRVKIFNIQVKEVTVQDSTTTIANNLTSTVQVFTDDDYQLPY VVGNGTE

GCLPAFPPQVFTLPQYGYATLNRDNTENPTERSSFFCLEYFPS KMLRTGN

NFEFTYNFEEVPFHSSFAPSQNLFKLANPLVDQYLYRFVSTNN TGGVQFN

KNLAGRYANTYKNWFPGPMGRTQGWNLGSGVNRASVSAFA TTNRMELEGA

SYQVPPQPNGMTNNLQGSNTYALENTMIFNSQPANPGTTATY LEGNMLIT

SESETQPVNRVAYNVGGQMATNNQSSTTAP**ATGTYNLQE**IVP GSVWMERD

[2472] VYLQGPIWAKIPETGAHFHPSPAMGGFGLKHPPPMMLIKNTP VPGNITSF

SDVPVSSFITQYSTGQVTVEMEWELKKENSKRWNPEIQYTNN YNDPQFVD

FAPDSTGEYRTTRPIGTRYLTRPL**[SEQ ID NO:1]**

[2473] ＞AAV5文库支架

MSFVDHPPDWLEEVGEGLREFLGLEAGPPKPKPNQQHQDQA RGLVLPGYN

YLGPGNGLDRGEPVNRADEVAREHDISYNEQLEAGDNPYLK YNHADAEFQ

EKLADDTSFGGNLGKAVFQAKKRVLEPFGLVEEGAKTAPTGK RIDDHFPK

[2474] RKKARTEEDSKPSTSSDAEAGPSGSQQLQIPAQPASSLGADTM SAGGGGP

LGDNNQGADGVGNASGDWHCDSTWMGDRVVTKSTRTWVL PSYNNHQYREI

KSGSVDGSNANAYFGYSTPWGYFDFNRFHSHWSPRDWQRLI NNYWGFRPR

SLRVKIFNIQVKEVTVQDSTTTIANNLTSTVQVFTDDDYQLPY

VVGNGTE

GCLPAFPPQVFTLPQYGYATLNRDNTENPTERSSFFCLEYFPS KMLRTGN

NFEFTYNFEEVPFHSSFAPSQNLFKLANPLVDQYLYRFVSTNN TGGVQFN

[2475]

KNLAGRYANTYKNWFPGPMGRTQGWNLGSGVNRASVSAFA TTNRMELEGA

SYQVPPQPNGMTNNLQGSNTYALENTMIFNSQPANPGTTATY LEGNMLIT

SESETQPVNRVAYNVGGQMATNNQSSTTAP

[2476] Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8、Xaa9

IVPGSVWMERDVYLQGPIWAKIPETGAHFHPSPAMGGFGLKH PPPMMLIK

[2477]

NTPVPGNITSFSDVPVSSFITQYSTGQVTVEMEWELKKENSKR WNPEIQY

TNNYNDPQFVDFAPDSTGEYRTTRPIGTRYLTRPL**[SEQ ID NO:2]**

[2478] Xaa1、Xaa2、Xaa3、Xaa4、Xaa5、Xaa6、Xaa7、Xaa8和Xaa9各自独立地选自A、R、N、D、C、E、Q、G、H、I、L、K、M、F、P、S、T、W、Y和V

[2479] SEQ ID NO:3

[2480]

MSFVDHPPDWLEEVGEGLREFLGLEAGPPKPKPNQQHQDQA RGLVLPGYN

YLGPGNGLDRGEPVNRADEVAREHDISYNEQLEAGDNPYLK YNHADAEFQ

EKLADDTSFGGNLGKAVFQAKKRVLEPFGLVEEGAKTAPTGK RIDDHFPK

RKKARTEEDSKPSTSSDAEAGPSGSQQLQIPAQPASSLGADTM SAGGGGP

LGDNNQGADGVGNASGDWHCDSTWMGDRVVTKSTRTWVL PSYNNHQYREI

KSGSVDGSNANAYFGYSTPWGYFDFNRFHSHWSPRDWQRLI NNYWGFRPR

[2481] SLRVKIFNIQVKEVTVQDSTTTIANNLTSTVQVFTDDDYQLPY VVGNGTE

GCLPAFPPQVFTLPQYGYATLNRDNTENPTERSSFFCLEYFPS KMLRTGN

NFEFTYNFEEVPFHSSFAPSQNLFKLANPLVDQYLYRFVSTNN TGGVQFN

KNLAGRYANTYKNWFPGPMGRTQGWNLGSGVNRASVSAFA TTNRMELEGA

SYQVPPQPNGMTNNLQGSNTYALENTMIFNSQPANPGTTATY LEGNMLIT

SESETQPVNRVAYNVGGQMATNNQSSTTAP**ATGT**

[2482] Xaa5

[2483] **NLQE**IVPGSVWMERDVYLQGPIWAKIPETGAHFHPSPAMGG

FGLKHPPPMM

[2484] LIKNTPVPGNITSFSDVPVSSFITQYSTGQVTVEMEWELKKEN SKRWNPEI

QYTNNYNDPQFVDFAPDSTGEYRTTRPIGTRYLTRPL

[2485] 其中Xaa5为V(Y585V)

[2486] SEQ ID NO:4

MSFVDHPPDWLEEVGEGLREFLGLEAGPPKPKPNQQHQDQA RGLVLPGYN

YLGPGNGLDRGEPVNRADEVAREHDISYNEQLEAGDNPYLK YNHADAEFQ

EKLADDTSFGGNLGKAVFQAKKRVLEPFGLVEEGAKTAPTGK RIDDHFPK

RKKARTEEDSKPSTSSDAEAGPSGSQQLQIPAQPASSLGADTM SAGGGGP

[2487]

LGDNNQGADGVGNASGDWHCDSTWMGDRVVTKSTRTWVL PSYNNHQYREI

KSGSVDGSNANAYFGYSTPWGYFDFNRFHSHWSPRDWQRLI NNYWGFRPR

SLRVKIFNIQVKEVTVQDSTTTIANNLTSTVQVFTDDDYQLPY VVGNGTE

GCLPAFPPQVFTLPQYGYATLNRDNTENPTERSSFFCLEYFPS KMLRTGN

NFEFTYNFEEVPFHSSFAPSQNLFKLANPLVDQYLYRFVSTNN TGGVQFN

KNLAGRYANTYKNWFPGPMGRTQGWNLGSGVNRASVSAFA TTNRMELEGA

[2488]

SYQVPPQPNGMTNNLQGSNTYALENTMIFNSQPANPGTTATY LEGNMLIT

SESETQPVNRVAYNVGGQMATNNQSSTTAP**ATGTYN**

[2489] Xaa7

QEIVPGSVWMERDVYLQGPIWAKIPETGAHFHPSPAMGGFGL KHPPPMMLI

[2490] KNTPVPGNITSFSDVPVSSFITQYSTGQVTVEMEWELKKENS KRWNPEIQY

TNNYNDPQFVDFAPDSTGEYRTTRPIGTRYLTRPL

[2491] 其中Xaa7为T(L587T)

[2492] SEQ ID NO:5

MSFVDHPPDWLEEVGEGLREFLGLEAGPPKPKPNQQHQDQA RGLVLPGYN

YLGPGNGLDRGEPVNRADEVAREHDISYNEQLEAGDNPYLK YNHADAEFQ

[2493]

EKLADDTSFGGNLGKAVFQAKKRVLEPFGLVEEGAKTAPTGK RIDDHFPK

RKKARTEEDSKPSTSSDAEAGPSGSQQLQIPAQPASSLGADTM

SAGGGGP

LGDNNQGADGVGNASGDWHCDSTWMGDRVVTKSTRTWVL PSYNNHQYREI

KSGSVDGSNANAYFGYSTPWGYFDFNRFHSHWSPRDWQRLI NNYWGFRPR

SLRVKIFNIQVKEVTVQDSTTTIANNLTSTVQVFTDDDYQLPY VVGNGTE

GCLPAFPPQVFTLPQYGYATLNRDNTENPTERSSFFCLEYFPS KMLRTGN

NFEFTYNFEEVPFHSSFAPSQNLFKLANPLVDQYLYRFVSTNN TGGVQFN

[2494]

KNLAGRYANTYKNWFPGPMGRTQGWNLGSGVNRASVSAFA TTNRMELEGA

SYQVPPQPNGMTNNLQGSNTYALENTMIFNSQPANPGTTATY LEGNMLIT

SESETQPVNRVAYNVGGQMATNNQSSTTAP**Xaa1TGT**

YNLQEIVPGSVWMERDVYLQGPIWAKIPETGAHFHPSPAMG GFGLKHPPPMM

LIKNTPVPGNITSFSDVPVSSFITQYSTGQVTVEMEWELKKEN SKRWNPEI

QYTNNYNDPQFVDFAPDSTGEYRTTRPIGTRYLTRPL

[2495] 其中Xaa1为T(A581T)

[2496] SEQ ID NO:6

MSFVDHPPDWLEEVGEGLREFLGLEAGPPKPKPNQQHQDQARGLVLPGYN

YLGPGNGLDRGEPVNRADEVAREHDISYNEQLEAGDNPYLKYNHADAEFQ

EKLADDTSFGGNLGKAVFQAKKRVLEPFGLVEEGAKTAPTGKRIDDHFPK

RKKARTEEDSKPSTSSDAEAGPSGSQQLQIPAQPASSLGADTMSAGGGGP

LGDNNQGADGVGNASGDWHCDSTWMGDRVVTKSTRTWVLPSYNNHQYREI

[2497] KSGSVDGSNANAYFGYSTPWGYFDFNRFHSHWSPRDWQRLINNYWGFRPR

SLRVKIFNIQVKEVTVQDSTTTIANNLTSTVQVFTDDDYQLPYVVGNGTE

GCLPAFPPQVFTLPQYGYATLNRDNTENPTERSSFFCLEYFPSKMLRTGN

NFEFTYNFEEVPFHSSFAPSQNLFKLANPLVDQYLYRFVSTNNTGGVQFN

KNLAGRYANTYKNWFPGPMGRTQGWNLGSGVNRASVSAFATTNRMELEGA

SYQVPPQPNGMTNNLQGSNTYALENTMIFNSQPANPGTTATYLEGNMLIT

[2498] SESETQPVNRVAYNVGGQMATNNQSSTTAP**A**

[2499] Xaa

GTYNLQEIVPGSVWMERDVYLQGPIWAKIPETGAHFHPSPA MGGFGLKHPP

[2500] PMMLIKNTPVPGNITSFSDVPVSSFITQYSTGQVTVEMEWELK KENSKRWN

PEIQYTNNYNDPQFVDFAPDSTGEYRTTRPIGTRYLTRPL

[2501] 其中Xaa2为A(T582A)

[2502] SEQ ID NO:7

MSFVDHPPDWLEEVGEGLREFLGLEAGPPKPKPNQQHQDQA RGLVLPGYN

YLGPGNGLDRGEPVNRADEVAREHDISYNEQLEAGDNPYLK YNHADAEFQ

EKLADDTSFGGNLGKAVFQAKKRVLEPFGLVEEGAKTAPTGK RIDDHFPK

[2503] RKKARTEEDSKPSTSSDAEAGPSGSQQLQIPAQPASSLGADTM SAGGGGP

LGDNNQGADGVGNASGDWHCDSTWMGDRVVTKSTRTWVL PSYNNHQYREI

KSGSVDGSNANAYFGYSTPWGYFDFNRFHSHWSPRDWQRLI NNYWGFRPR

SLRVKIFNIQVKEVTVQDSTTTIANNLTSTVQVFTDDDYQLPY

VVGNGTE

GCLPAFPPQVFTLPQYGYATLNRDNTENPTERSSFFCLEYFPS KMLRTGN

NFEFTYNFEEVPFHSSFAPSQNLFKLANPLVDQYLYRFVSTNN TGGVQFN

[2504]

KNLAGRYANTYKNWFPGPMGRTQGWNLGSGVNRASVSAFA TTNRMELEGA

SYQVPPQPNGMTNNLQGSNTYALENTMIFNSQPANPGTTATY LEGNMLIT

SESETQPVNRVAYNVGGQMATNNQSSTTAP**ATG**

[2505] Xaa4

YNLQEIVPGSVWMERDVYLQGPIWAKIPETGAHFHPSPAMG GFGLKHPPP

[2506] MMLIKNTPVPGNITSFSDVPVSSFITQYSTGQVTVEMEWELK KENSKRWN

PEIQYTNNYNDPQFVDFAPDSTGEYRTTRPIGTRYLTRPL

[2507] 其中Xaa4为A(T584A)

[2508] SEQ ID NO:8

MSFVDHPPDWLEEVGEGLREFLGLEAGPPKPKPNQQHQDQA RGLVLPGYN

[2509]

YLGPGNGLDRGEPVNRADEVAREHDISYNEQLEAGDNPYLK YNHADAEFQ

EKLADDTSFGGNLGKAVFQAKKRVLEPFGLVEEGAKTAPTGK RIDDHFPK

RKKARTEEDSKPSTSSDAEAGPSGSQQLQIPAQPASSLGADTM SAGGGGP

LGDNNQGADGVGNASGDWHCDSTWMGDRVVTKSTRTWVL PSYNNHQYREI

KSGSVDGSNANAYFGYSTPWGYFDFNRFHSHWSPRDWQRLI NNYWGFRPR

SLRVKIFNIQVKEVTVQDSTTTIANNLTSTVQVFTDDDYQLPY VVGNGTE

[2510]

GCLPAFPPQVFTLPQYGYATLNRDNTENPTERSSFFCLEYFPS KMLRTGN

NFEFTYNFEEVPFHSSFAPSQNLFKLANPLVDQYLYRFVSTNN TGGVQFN

KNLAGRYANTYKNWFPGPMGRTQGWNLGSGVNRASVSAFA TTNRMELEGA

SYQVPPQPNGMTNNLQGSNTYALENTMIFNSQPANPGTTATY LEGNMLIT

SESETQPVNRVAYNVGGQMATNNQSSTTAP**ATGT**

[2511] Xaa5

NXaa7QEIVPGSVWMERDVYLQGPIWAKIPETGAHFHPSPAM GGFGLKHPPPMM

[2512]

LIKNTPVPGNITSFSDVPVSSFITQYSTGQVTVEMEWELKKEN

SKRWNPEI

[2513]

QYTNNYNDPQFVDFAPDSTGEYRTTRPIGTRYLTRPL

[2514] 其中Xaa5为V(Y585V)并且Xaa7为T(L587T)

[2515] SEQ ID NO:9ATGTYNLQE(AAV5 VP1衣壳的581-589唾液酸结合区,和Xaa1-Xaa9的参考)

[2516] SEQ ID NO:111(图10,中间变体DNA)AATCCAGCTATGTTCAATTGCGATTAC

[2517] SEQ ID NO:112(图10,中间变体;氨基酸)NPAMFNCDY

[2518] SEQ ID NO:113(图10,右边变体,DNA)AAGTTGACCGCCCACATCTACGCCTTG

[2519] SEQ ID NO:114(图10,右边变体;氨基酸)KLTAHIYAL

[2520] SEQ ID NO:115-1114(表8,其中列出的AA代表SEQ ID NO:2的Xaa1-Xaa9)

[2521] AAV5 VP2序列(SEQ ID NO:1115)

[2522]

TAPTGKRIDDHFPKRKKARTEEDSKPSTSSDAEAGPSGSQQLQ
IPAQPASSLGADTMSAGGGGPLGDNNQGADGVGNASGDWHCDS
TWMGDRVVTKSTRTWVLPSYNNHQYREIKSGSVDGSNANAYFG
YSTPWGYFDFNRFHSHWSPRDWQRLINNYWGFRPRSLRVKIFNIQ
VKEVTVQDSTTTIANNLTSTVQVFTDDDYQLPYVVGNGTEGCLPA
FPPQVFTLPQYGYATLNRDNTENPTERSSFFCLEYFPSKMLRTGNN
FEFTYNFEEVPFHSSFAPSQNLFKLANPLVDQYLYRFVSTNNTGGV
QFNKNLAGRYANTYKNWFPGPMGRTQGWNLGSGVNRASVSAFA
TTNRMELEGASYQVPPQPNGMTNNLQGSNTYALENTMIFNSQPA
NPGTTATYLEGNMLITSESETQPVNRVAYNVGGQMATNNQSSTTA
P**ATGTYNLQE**IVPGSVWMERDVYLQGPIWAKIPETGAHFHPSPA

[2523]

MGGFGLKHPPPMMLIKNTPVPGNITSFSDVPVSSFITQYSTGQVTV
EMEWELKKENSKRWNPEIQYTNNYNDPQFVDFAPDSTGEYRTTR
PIGTRYLTRPL

[2524] AAV5 VP3序列(SEQ ID NO:1116)

[2525]

MSAGGGGPLGDNNQGADGVGNASGDWHCDSTWMGDRVV
TKSTRTWVLPSYNNHQYREIKSGSVDGSNANAYFGYSTPWGYFD
FNRFHSHWSPRDWQRLINNYWGFRPRSLRVKIFNIQVKEVTVQDS
TTTIANNLTSTVQVFTDDDYQLPYVVGNGTEGCLPAFPPQVFTLPQ
YGYATLNRDNTENPTERSSFFCLEYFPSKMLRTGNNFEFTYNFEEV
PFHSSFAPSQNLFKLANPLVDQYLYRFVSTNNTGGVQFNKNLAGR
YANTYKNWFPGPMGRTQGWNLGSGVNRASVSAFATTNRMELEG
ASYQVPPQPNGMTNNLQGSNTYALENTMIFNSQPANPGTTATYLE
GNMLITSESETQPVNRVAYNVGGQMATNNQSSTTAP**ATGTYNLQ**
EIVPGSVWMERDVYLQGPIWAKIPETGAHFHPSPAMGGFGLKHPP
PMMLIKNTPVPGNITSFSDVPVSSFITQYSTGQVTVEMEWELKKE
NSKRWNPEIQYTNNYNDPQFVDFAPDSTGEYRTTRPIGTRYLTRPL

[2526] 3.等效方案及以引用方式的并入

[2527] 本文所引用的所有参考文献出于所有目的以引用方式并入,其程度就如同具体地且个别地指出各个别出版物、数据库条目(例如,Genbank序列或GeneID条目)、专利申请或专利以引用方式整体并入。申请人意在依据37 C.F.R.§1.57(b)(1)将此以引用方式并入的声明与每一个别出版物、数据库条目(例如,Genbank序列或GeneID条目)、专利申请或专利相关联,其中每一者都遵照37 C.F.R.§1.57(b)(2)明确标识,即使所述引用与专门的以引用方式并入的声明不直接相邻。在说明书中包含专门的以引用方式并入的声明(如果有的话)不会以任何方式削弱这个以引用方式并入的一般声明。本文引用参考文献并不意味着承认该参考文献是相关的现有技术,也不构成对这些出版物或文件的内容或日期的任何承认。

[2528] 虽然已参考优选实施方案和各种替代实施方案特别地显示和描述了本发明,但是相关领域技术人员应了解,可在不脱离本发明的精神和范围的情况下在其中进行形式和细节的各种改变。

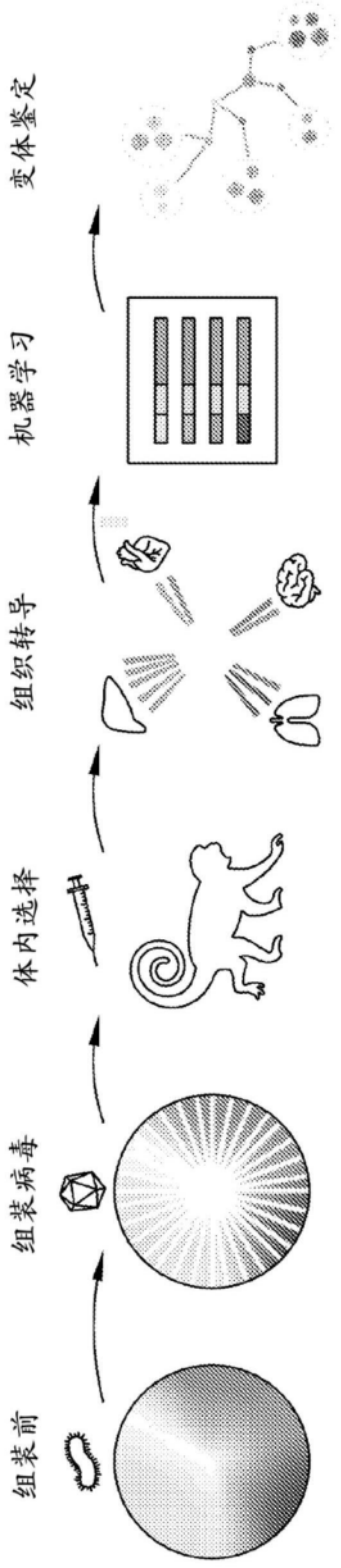


图1

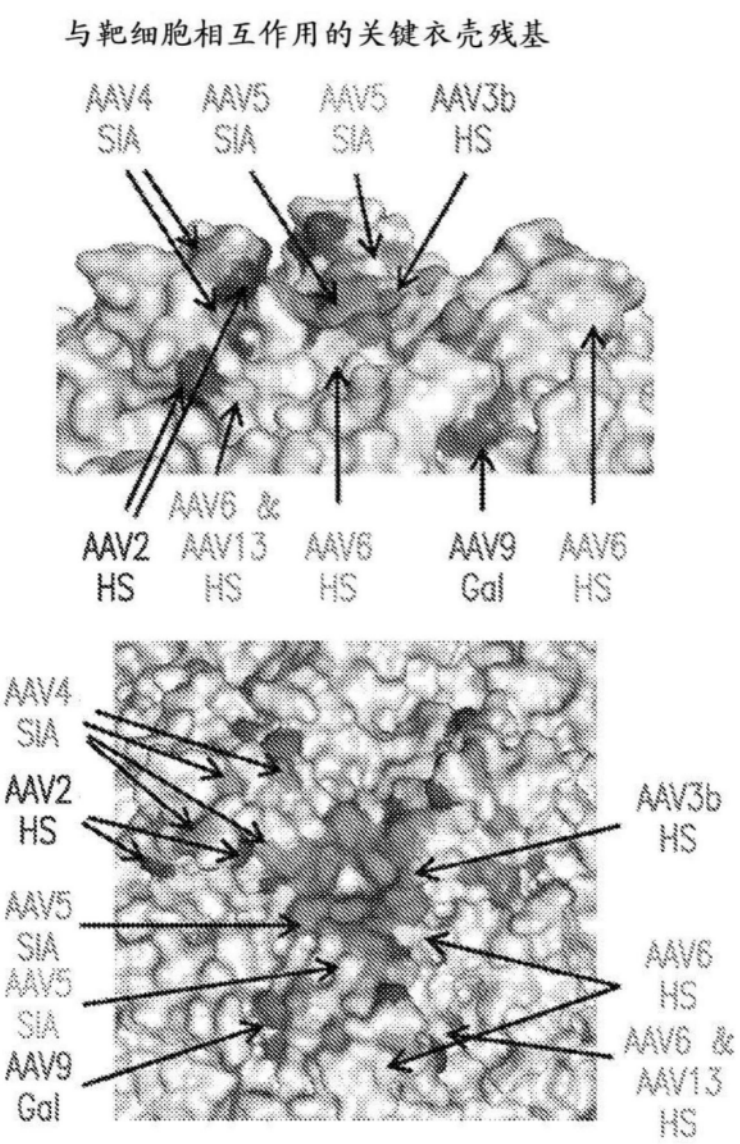

与靶细胞相互作用的关键衣壳残基
AAV4
SIA
AAV5
SIA
AAV5
SIA
AAV3b
HS
AAV6 &
AAV13
HS
AAV2
HS
AAV6
HS
AAV9
Gal
AAV6
HS
AAV4
SIA
AAV2
HS
AAV5
SIA
AAV5
SIA
AAV9
Gal
AAV3b
HS
AAV6
HS
AAV6 &
AAV13
HS

图2A

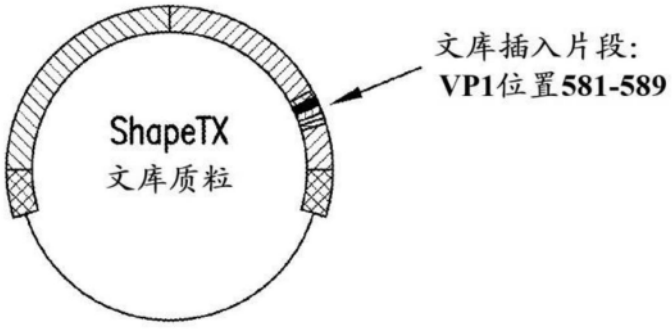

ShapeTX
文库质粒
文库插入片段:
VP1位置581-589

图2B

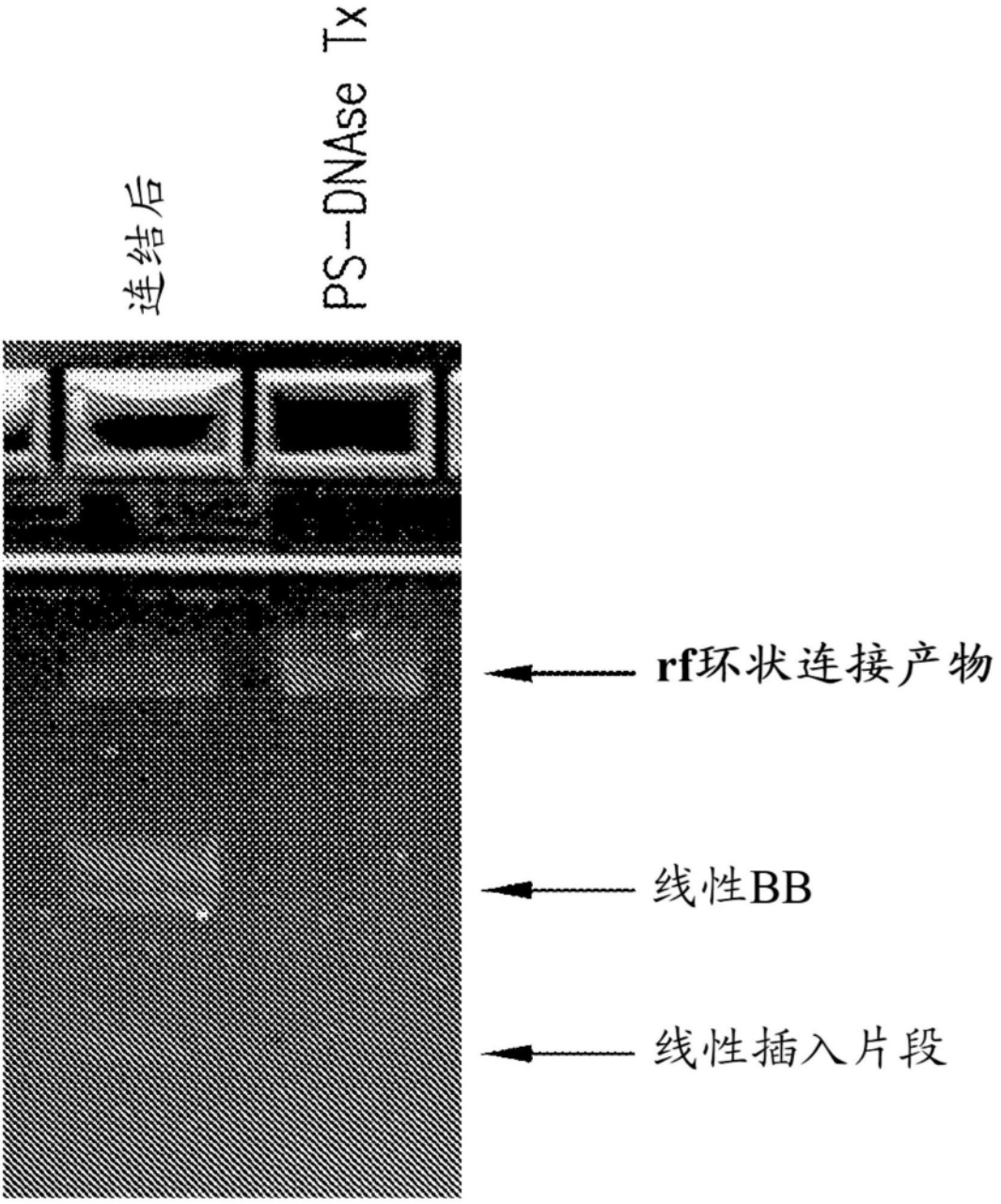


图3

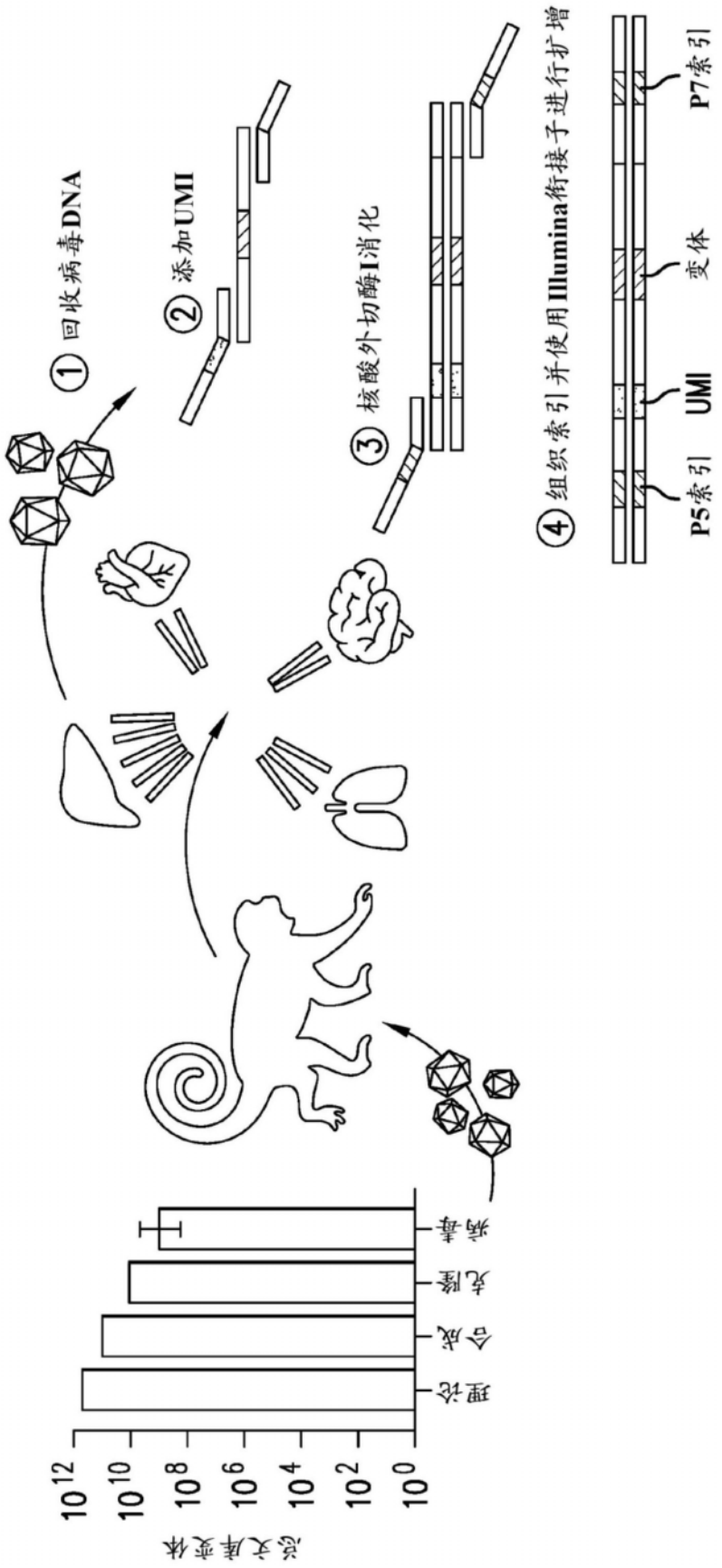

① 回收病毒DNA
② 添加UMI
③ 核酸外切酶I消化
④ 组织索引并使用Illumina衔接子进行扩增
P5索引
UMI
变体
P7索引
总文库实体
10^{12}
10^{10}
10^{8}
10^{6}
10^{4}
10^{2}
10^{0}
理论
心脏
皮肾
病毒

图4

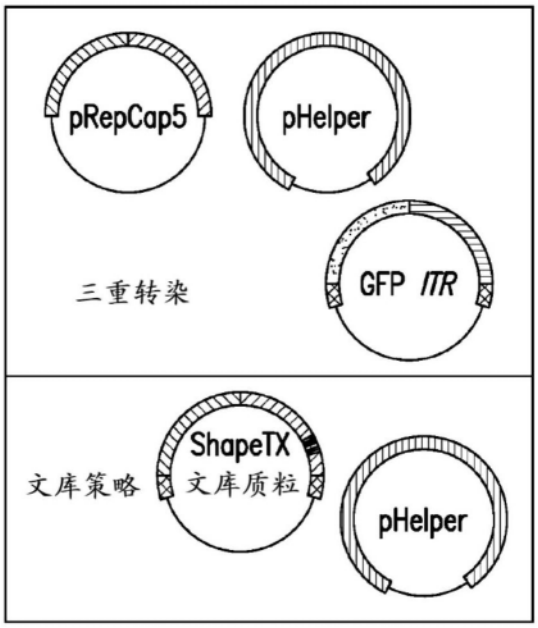

pRepCap5
pHelper
三重转染
GFP ITR
文库策略
ShapeTX
文库质粒
pHelper

图5A

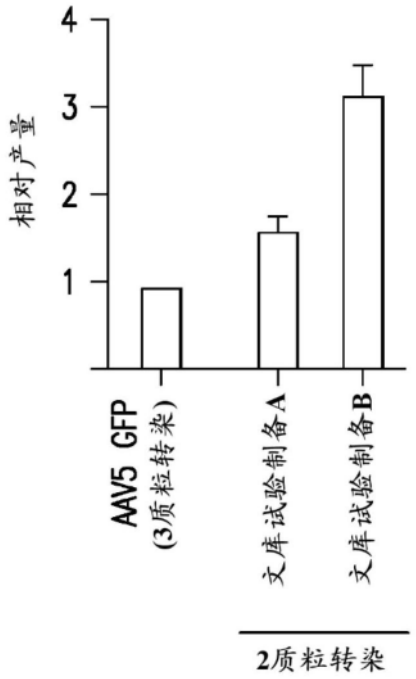

相对产量
4
3
2
1
AAV5 GFP
(3质粒转染)
文库试验制备A
文库试验制备B
2质粒转染

图5B

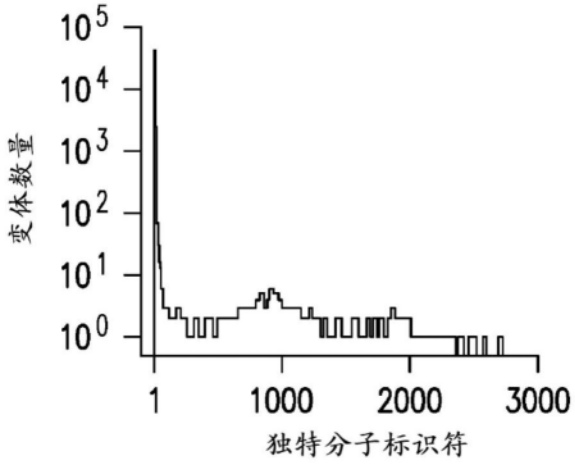


图5C

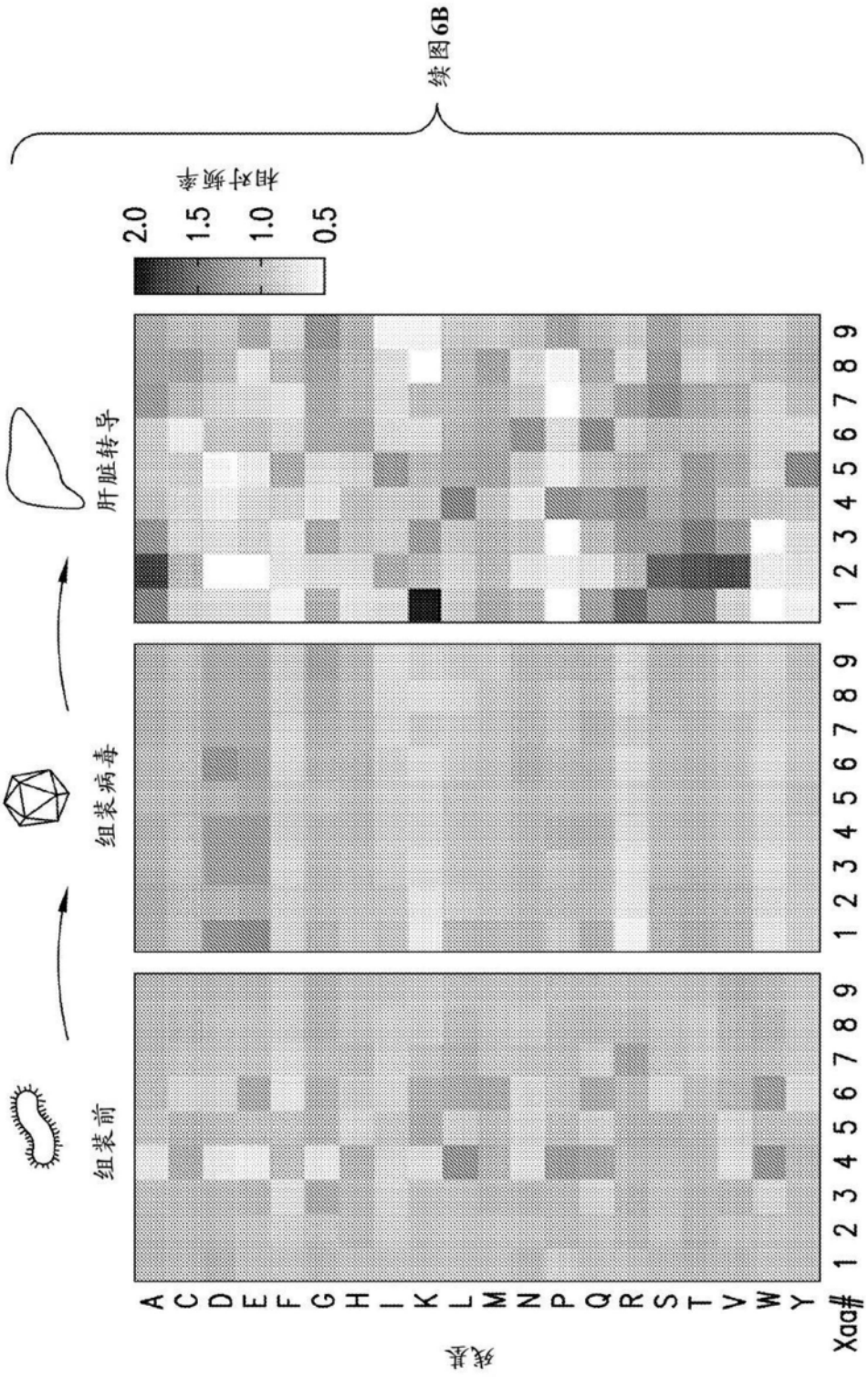
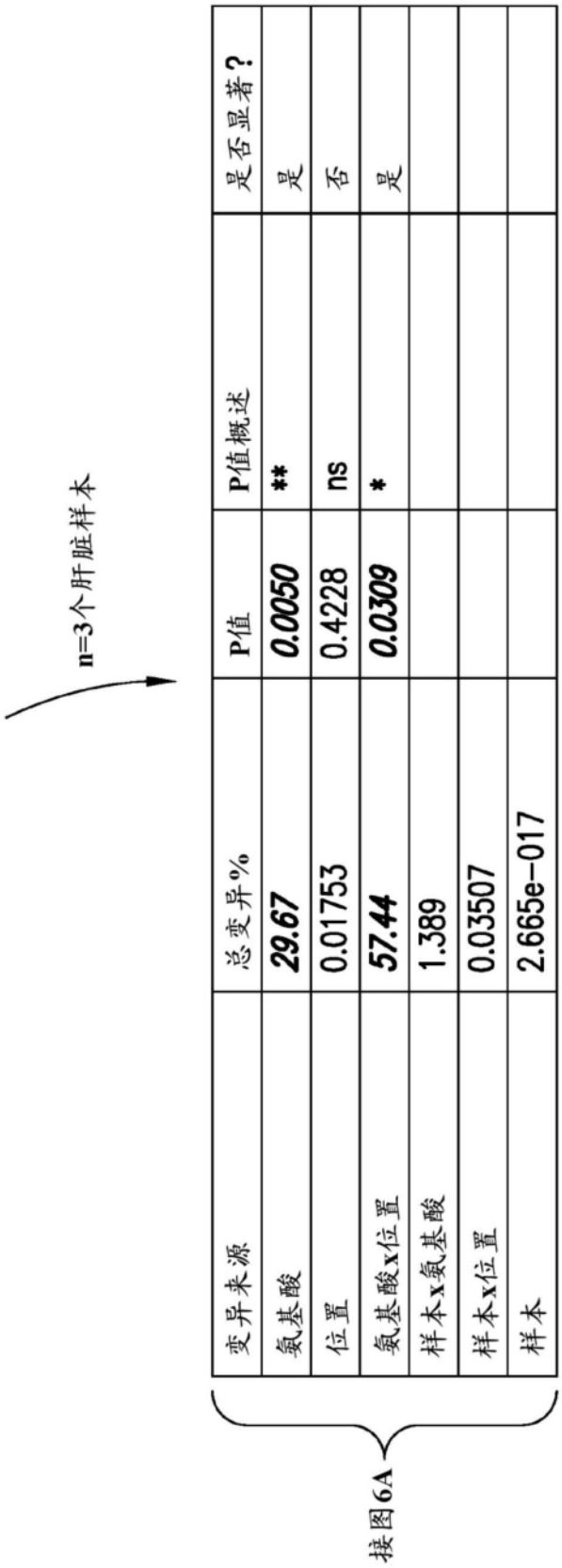

续图6B
相对频率
2.0
1.5
1.0
0.5
肝脏转导
组装病毒
组装前
残基
Xaa#
A C D E F G H I K L M N P Q R S T V W Y
1 2 3 4 5 6 7 8 9

图6A



变异来源	总变异%	P值	P值概述	是否显著?
氨基酸	***29.67***	***0.0050***	**	是
位置	0.01753	0.4228	ns	否
氨基酸x位置	***57.44***	***0.0309***	*	是
样本x氨基酸	1.389			
样本x位置	0.03507			
样本	2.665e-017			

图6B

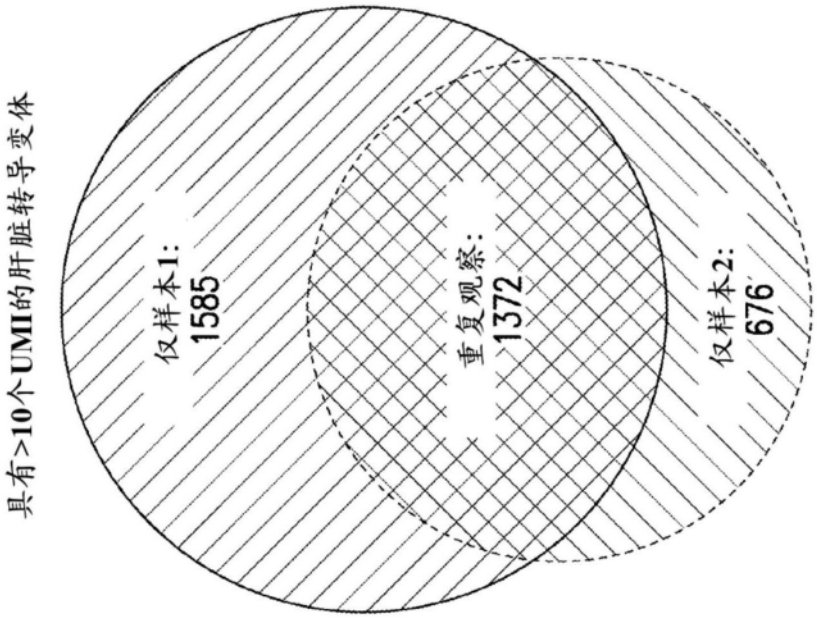

具有>10个UMI的肝脏转导变体
仅样本1:
1585
重复观察:
1372
仅样本2:
676

图7A

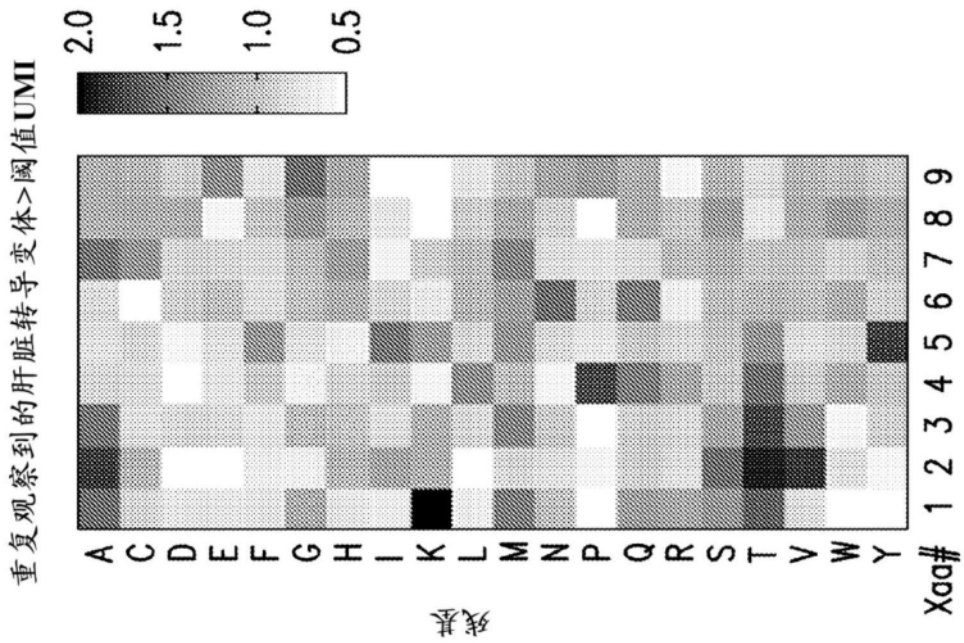

重复观察到的肝脏转导变体>阈值UMI
2.0
1.5
1.0
0.5
A C D E F G H I K L M N P Q R S T V W Y
残基
Xaa# 1 2 3 4 5 6 7 8 9

图7B

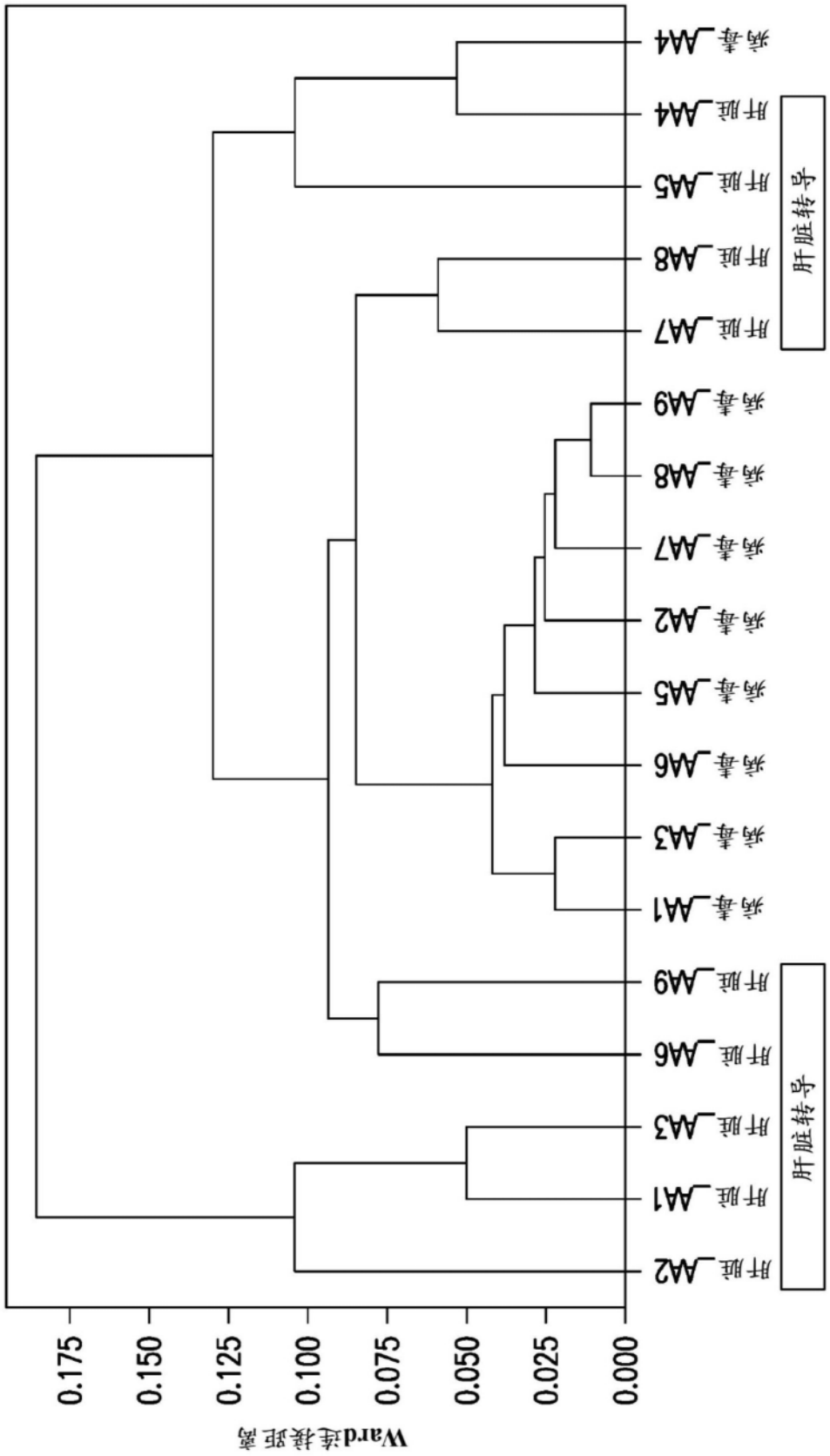

肝脏_AA2
肝脏_AA1
肝脏_AA3
肝脏_AA6
肝脏_AA9
病毒_AA1
病毒_AA3
病毒_AA6
病毒_AA5
病毒_AA2
病毒_AA7
病毒_AA8
病毒_AA9
肝脏_AA7
肝脏_AA8
肝脏_AA5
肝脏_AA4
病毒_AA4
含转脏肝
含转脏肝
0.000
0.025
0.050
0.075
0.100
0.125
0.150
0.175
Ward连接距离

图8

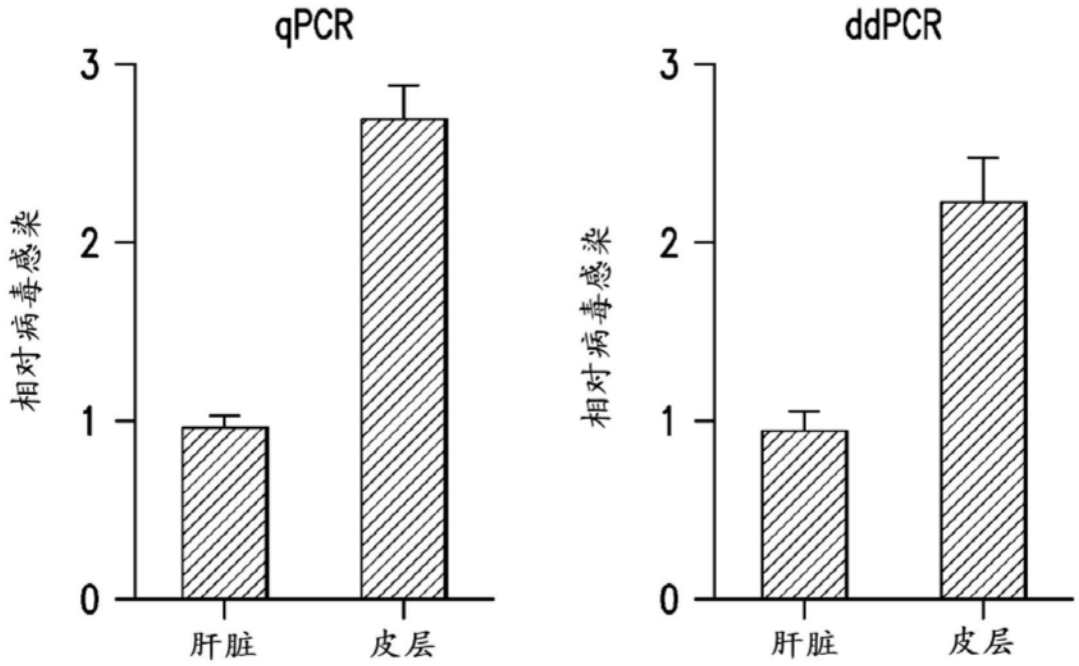


图9

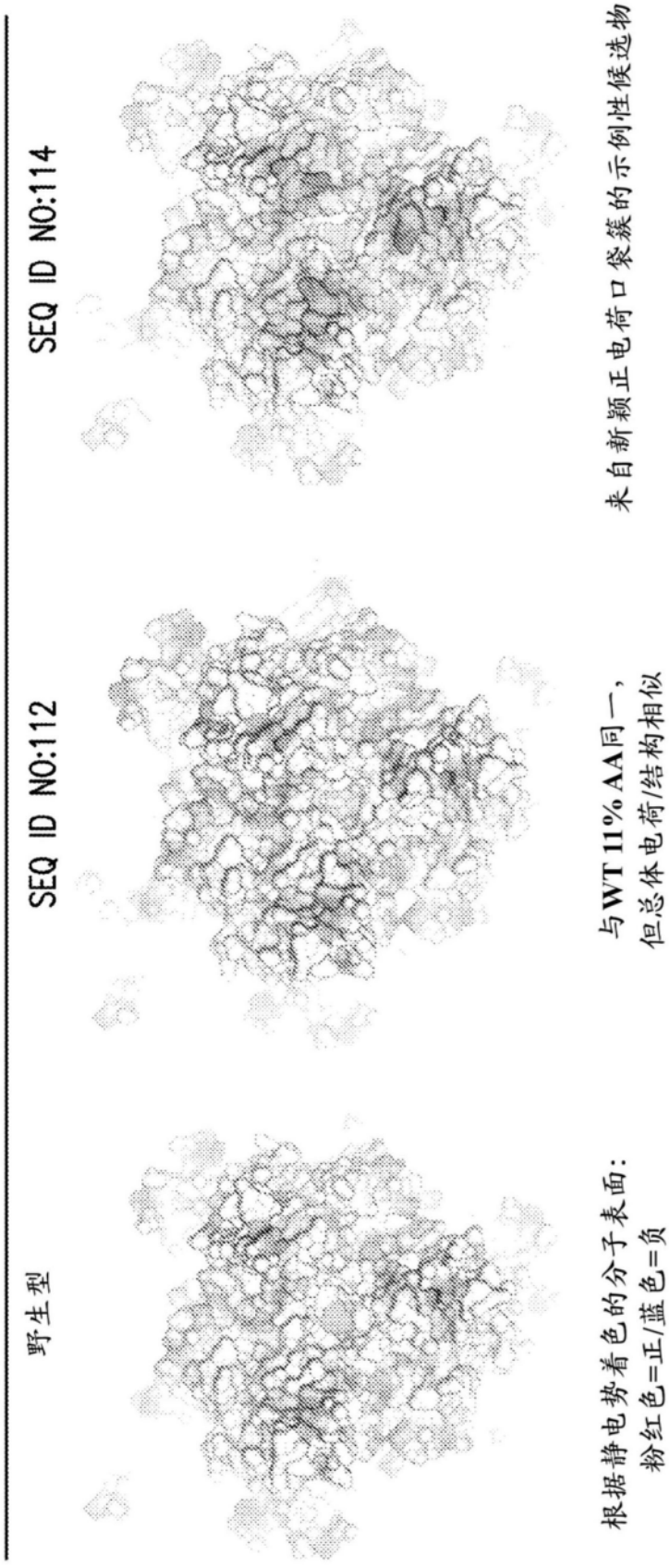

野生型
SEQ ID NO:112
SEQ ID NO:114
根据静电势着色的分子表面：
粉红色=正/蓝色=负
与WT 11% AA同一，
但总体电荷/结构相似
来自新颖正电荷口袋簇的示例性候选物
-10 0 +10

图10



图11A

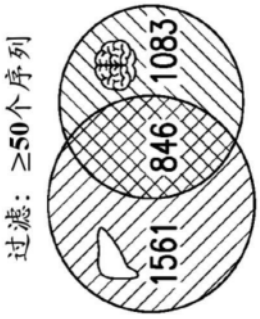


图11B

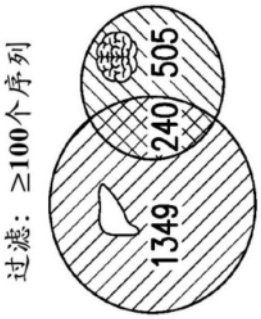


图11C

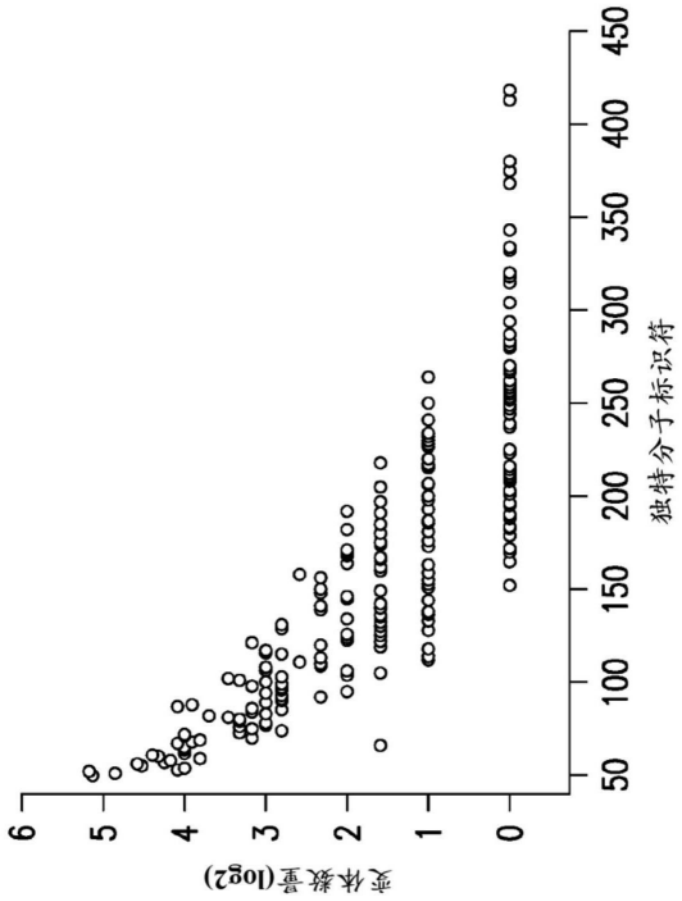


图11D

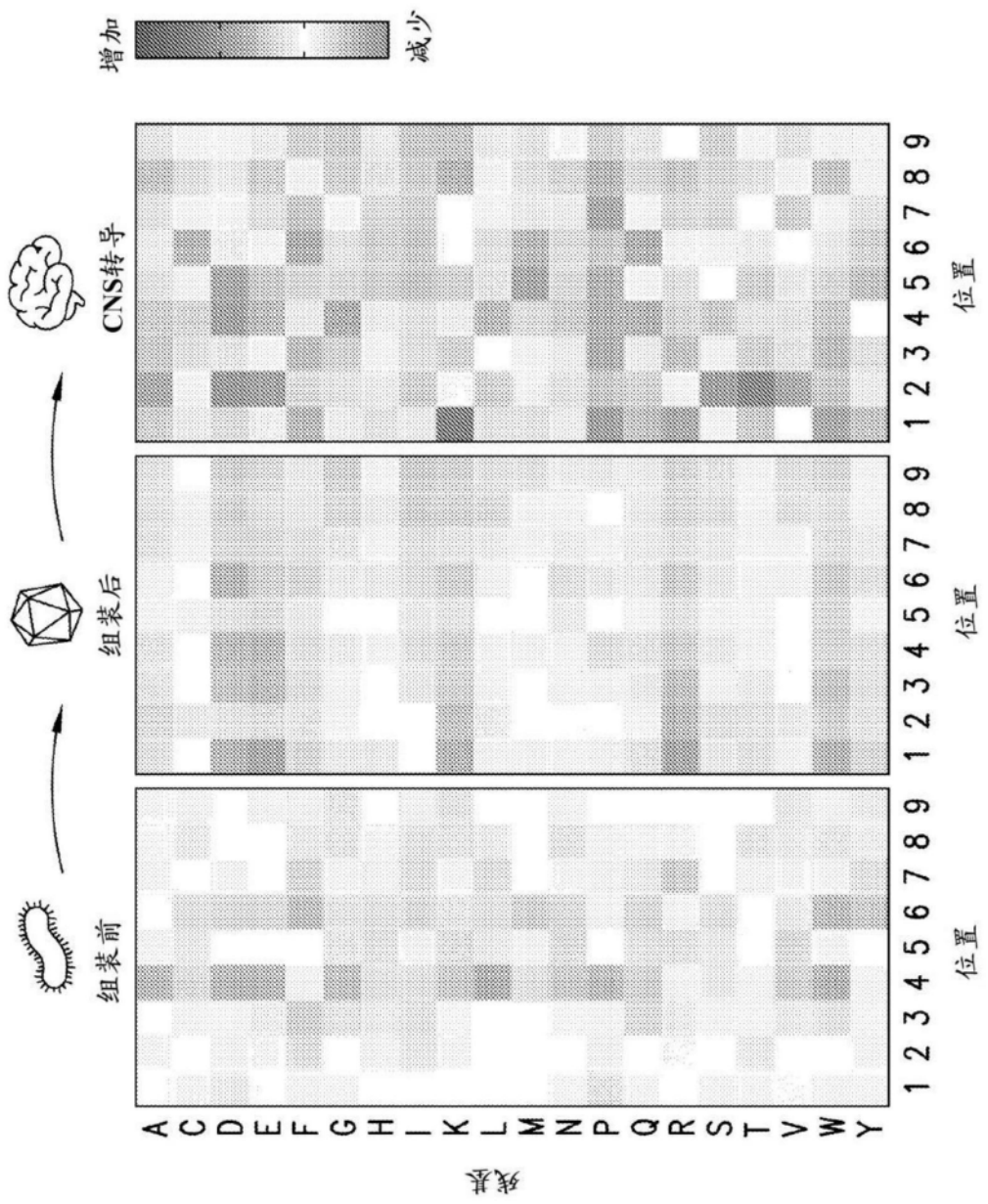


图12

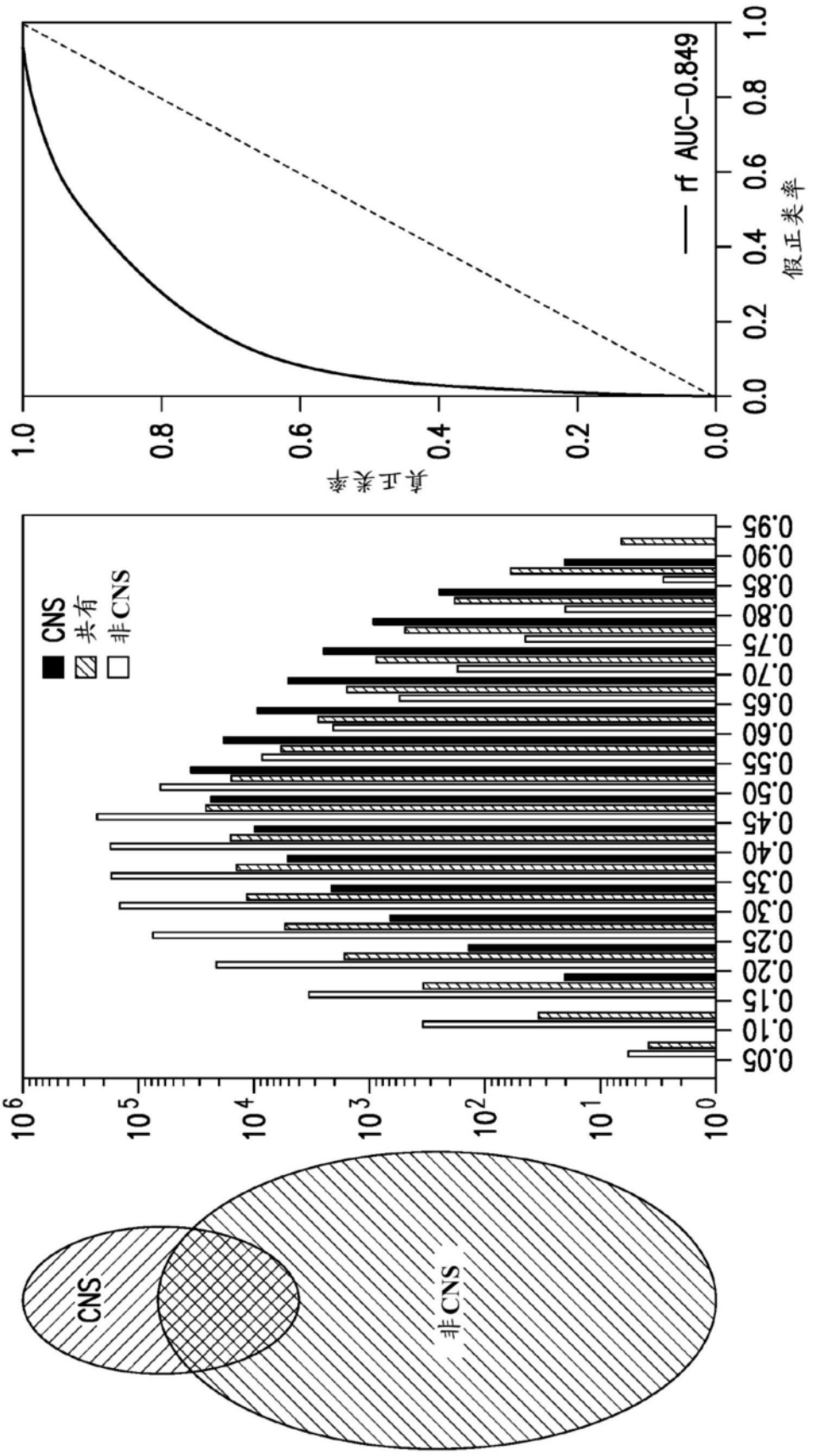


图13

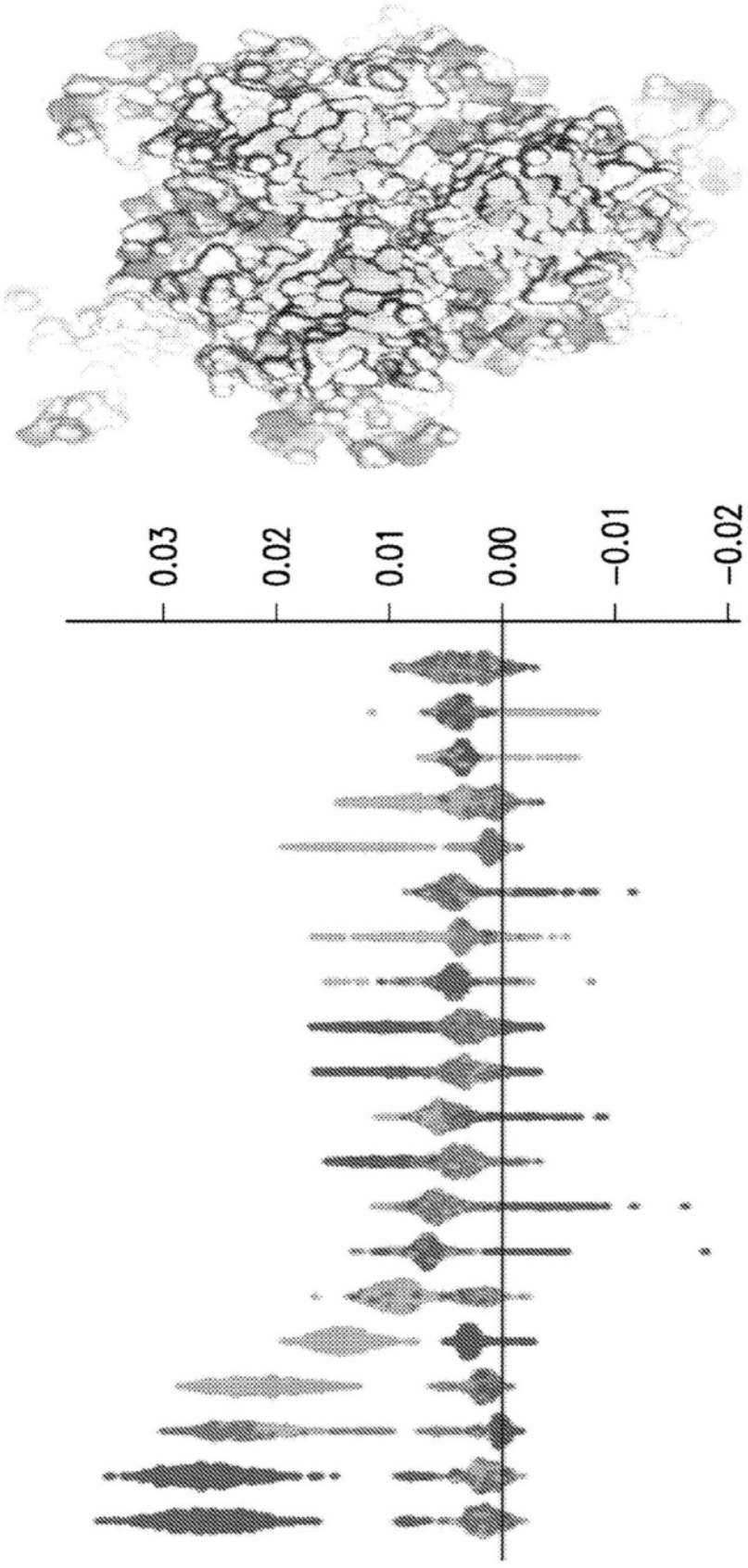


图14

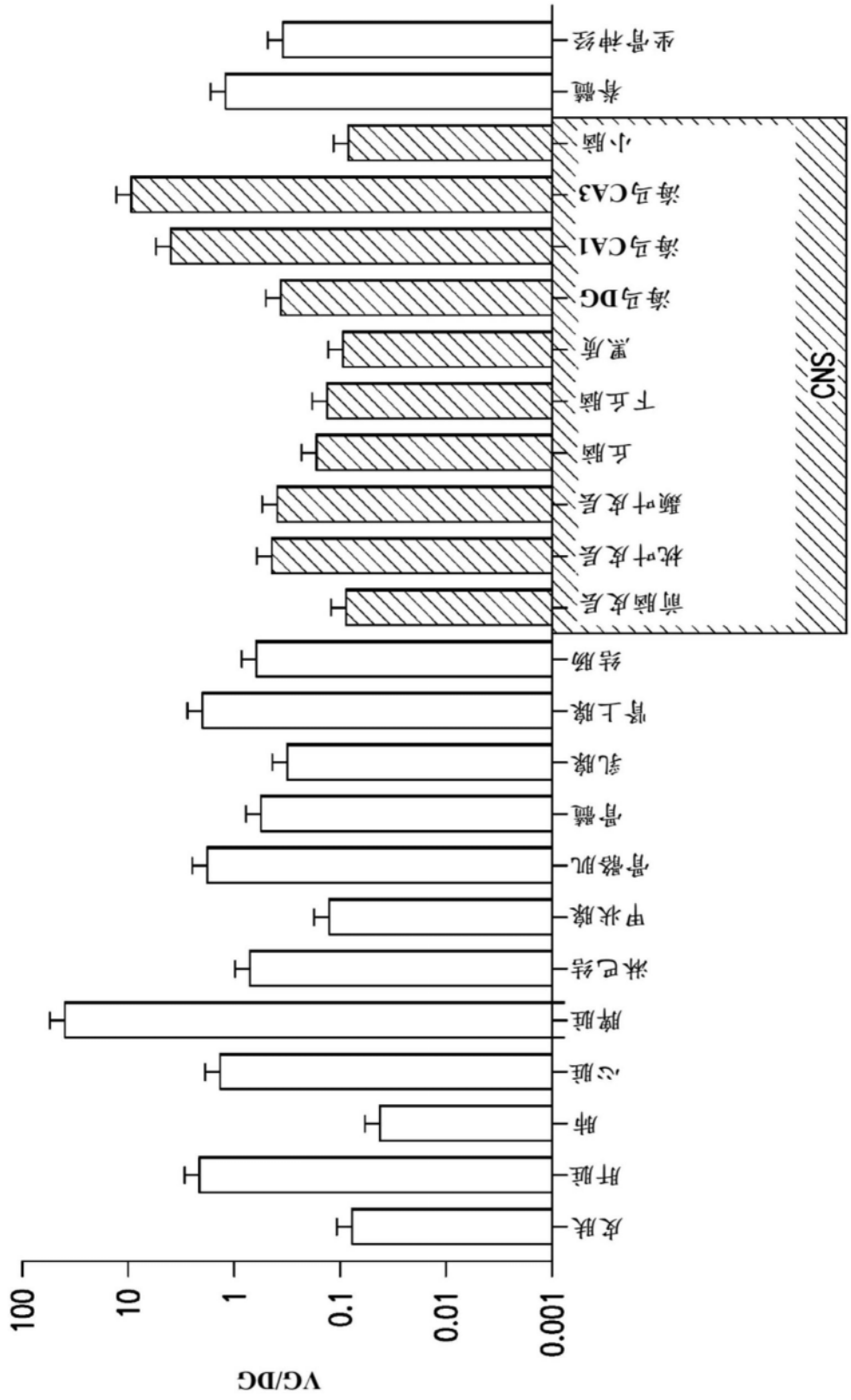

VG/DG
100
10
1
0.1
0.01
0.001
皮肤
肝脏
肺
心脏
脾脏
淋巴结
甲状腺
骨骼肌
骨髓
乳腺
肾上腺
结肠
CNS
前脑皮层
枕叶皮层
颞叶皮层
丘脑
下丘脑
黑质
海马DG
海马CA1
海马CA3
小脑
脊髓
坐骨神经

图15

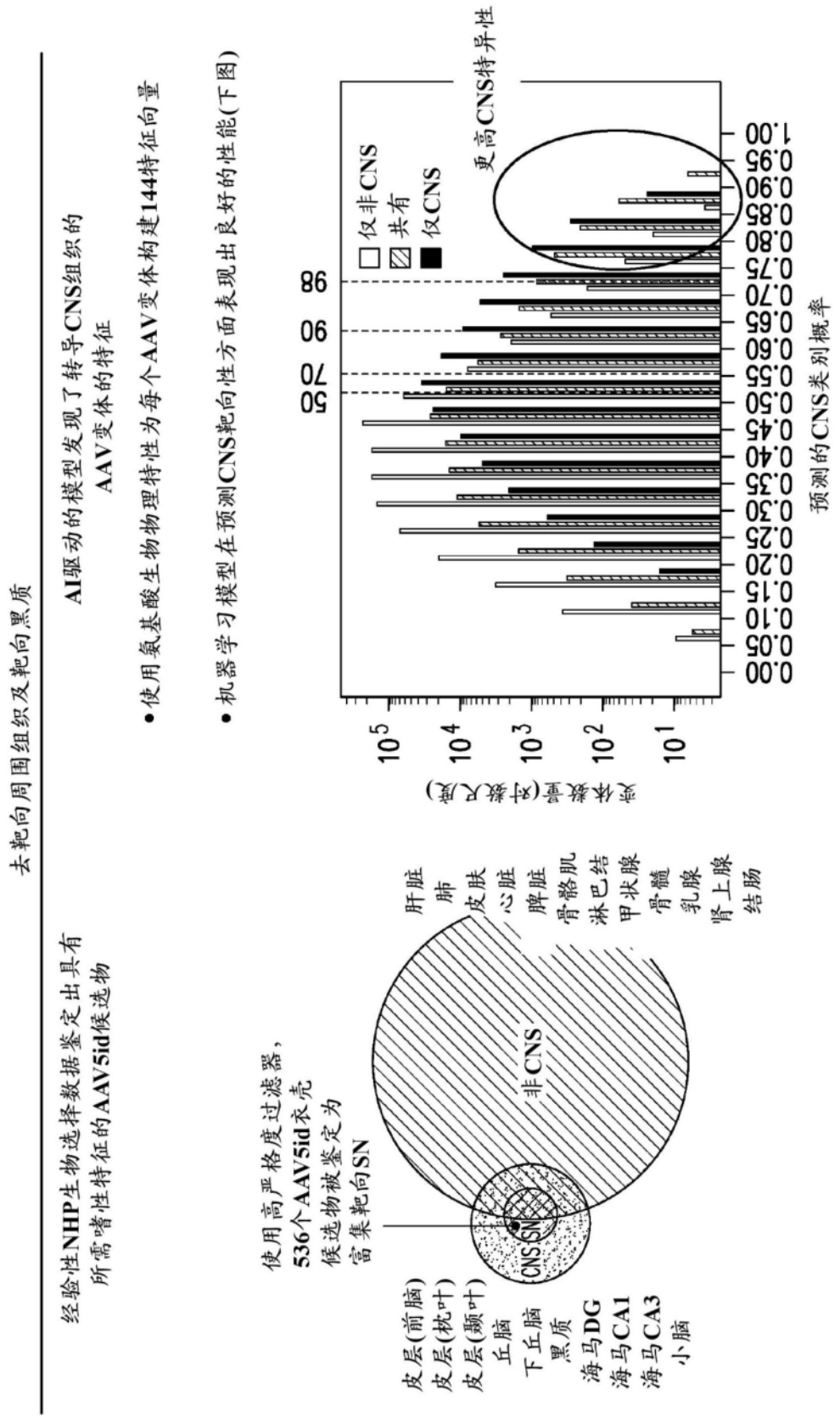

去靶向周围组织及靶向黑质
经验性NHP生物选择数据鉴定出具有
所需嗜性特征的AAV5id候选物
AI驱动的模型发现了转导CNS组织的
AAV变体的特征
• 使用氨基酸生物物理特性为每个AAV变体构建144特征向量
• 机器学习模型在预测CNS靶向性方面表现出良好的性能(下图)
使用高严格度过滤器，
536个AAV5id衣壳
候选物被鉴定为
富集靶向SN
皮层(前脑)
皮层(枕叶)
皮层(颞叶)
丘脑
下丘脑
黑质
海马DG
海马CA1
海马CA3
小脑
CNS
SN
非CNS
肝脏
肺
皮肤
心脏
脾脏
骨骼肌
淋巴结
甲状腺
骨髓
乳腺
肾上腺
结肠
仅非CNS
共有
仅CNS
更高CNS特异性
预测的CNS类别概率

图16

去靶向肝脏及靶向心肌/骨骼肌
经验性NHP生物选择数据鉴定出具有
所需嗜性特征的AAV5id候选物
AI驱动的模型发现了避免肝脏转导的AAV变体的特征
• 使用氨基酸生物物理特性为每个AAV变体构建144特征向量
• 在肝脏保留性变体中，心肌+骨骼肌指导变体选择中共有的特征
负类
共有
正类
更高肝脏去靶向性
预测的肝脏去靶向性类别概率
病体数量(对数尺度)
心脏
肝脏
骨骼肌
使用高严格度过滤器，
797个AAV5id衣壳
候选物符合所需特征

图17

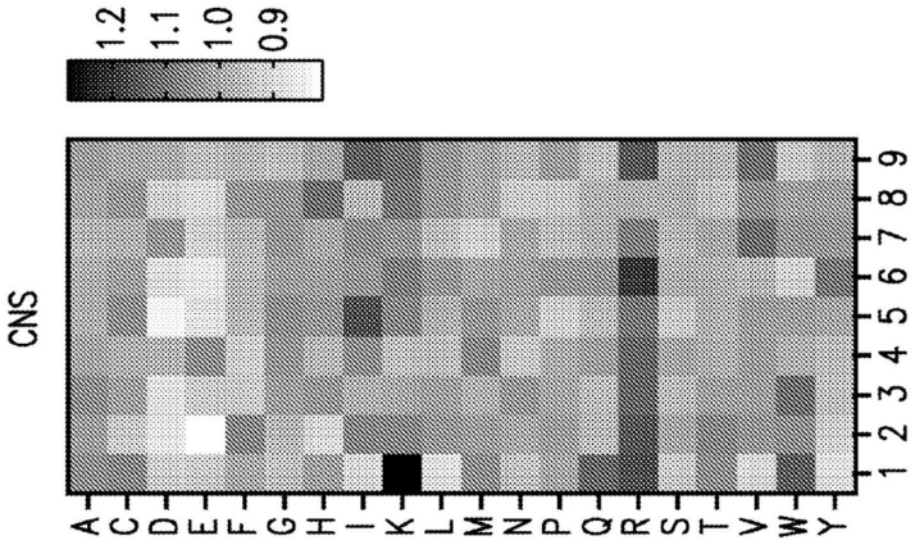

CNS
1.2
1.1
1.0
0.9
A C D E F G H I K L M N P Q R S T V W Y
1 2 3 4 5 6 7 8 9

图18A

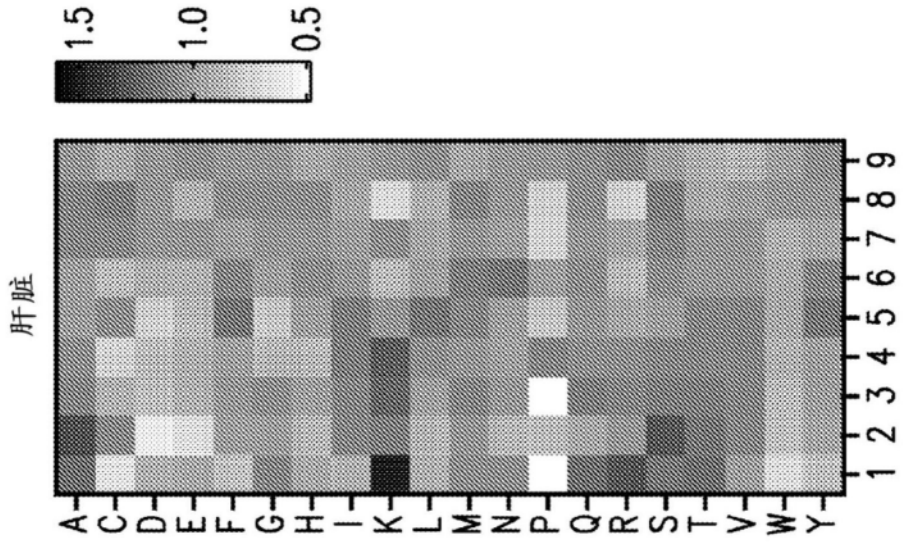

肝脏
1.5
1.0
0.5
A C D E F G H I K L M N P Q R S T V W Y
1 2 3 4 5 6 7 8 9

图18B

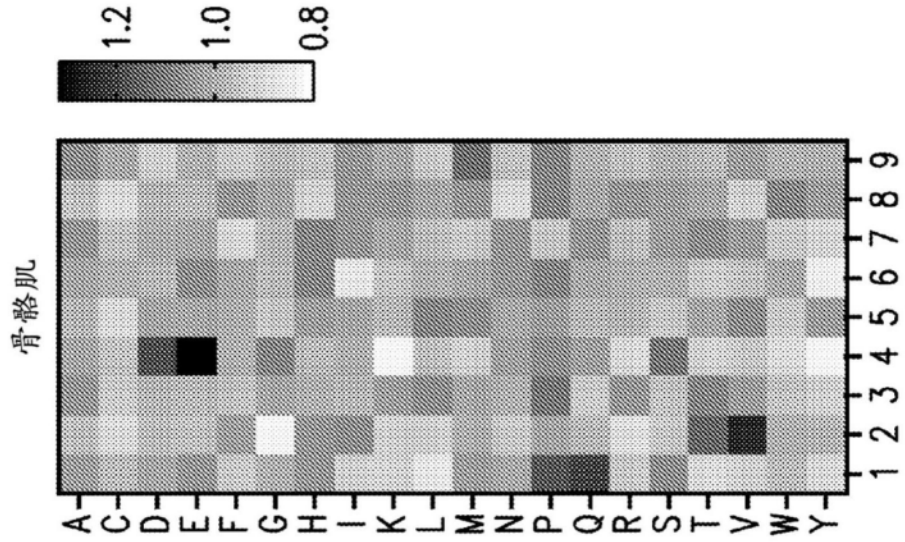

1.2
1.0
0.8
骨骼肌
1 2 3 4 5 6 7 8 9
A C D E F G H I K L M N P Q R S T V W Y

图18C

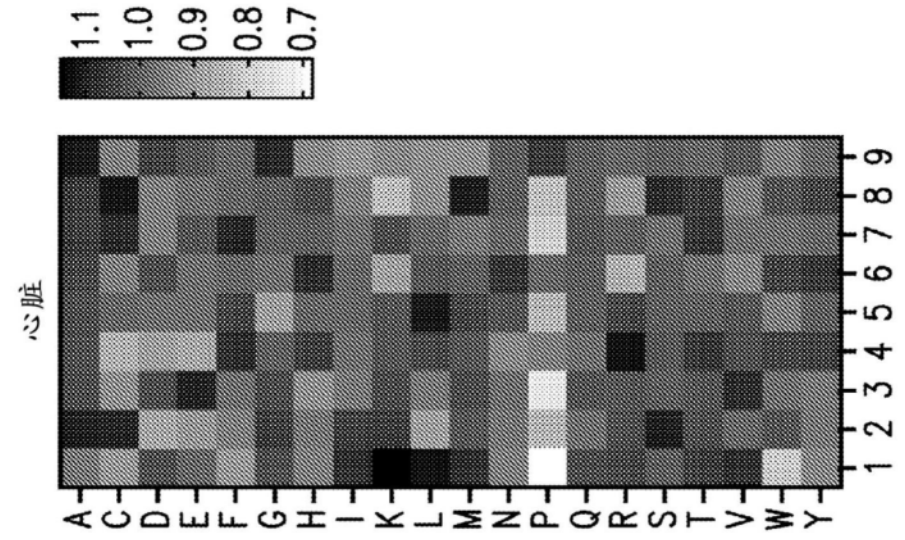

1.1
1.0
0.9
0.8
0.7
心脏
1 2 3 4 5 6 7 8 9
A C D E F G H I K L M N P Q R S T V W Y

图18D

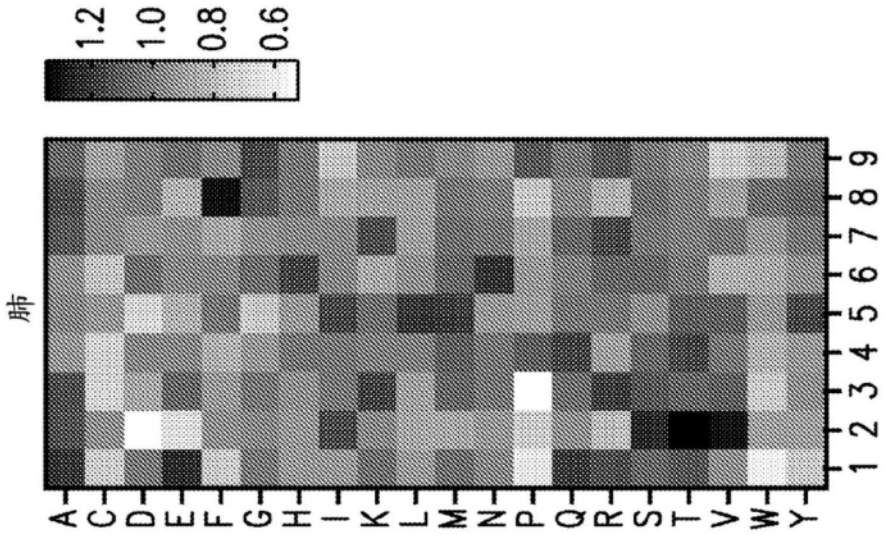

1.2
1.0
0.8
0.6
肺
1 2 3 4 5 6 7 8 9
A C D E F G H I K L M N P Q R S T V W Y

图18E

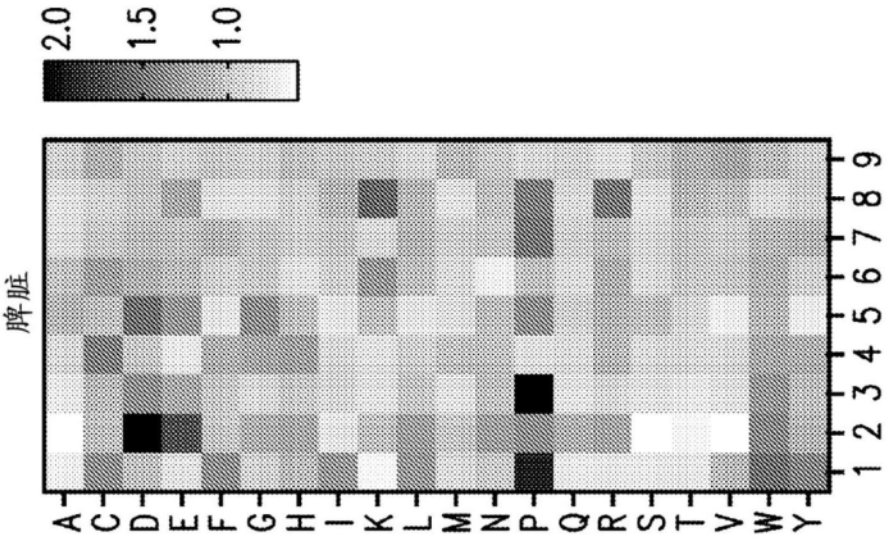

2.0
1.5
1.0
脾脏
1 2 3 4 5 6 7 8 9
A C D E F G H I K L M N P Q R S T V W Y

图18F

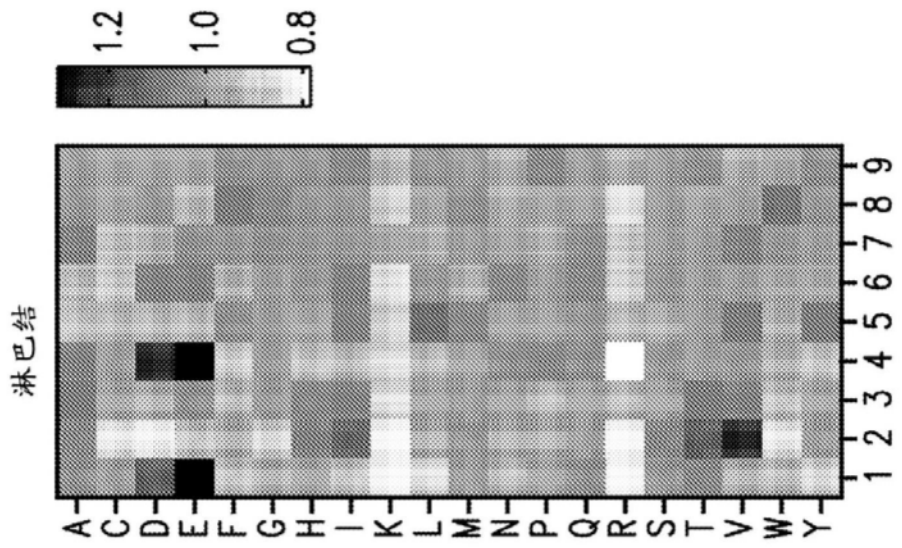

1.2
1.0
0.8
淋巴结
1 2 3 4 5 6 7 8 9
A C D E F G H I K L M N P Q R S T V W Y

图18G

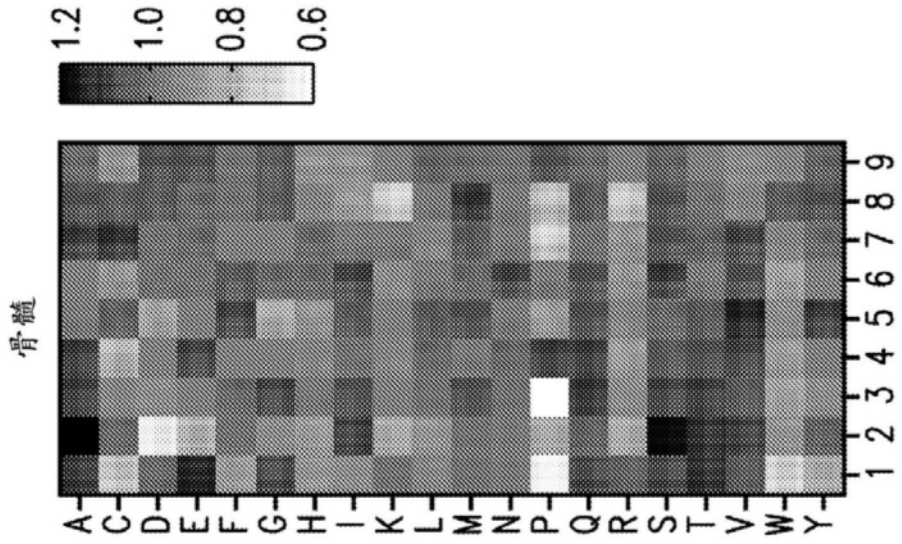

1.2
1.0
0.8
0.6
骨髓
1 2 3 4 5 6 7 8 9
A C D E F G H I K L M N P Q R S T V W Y

图18H

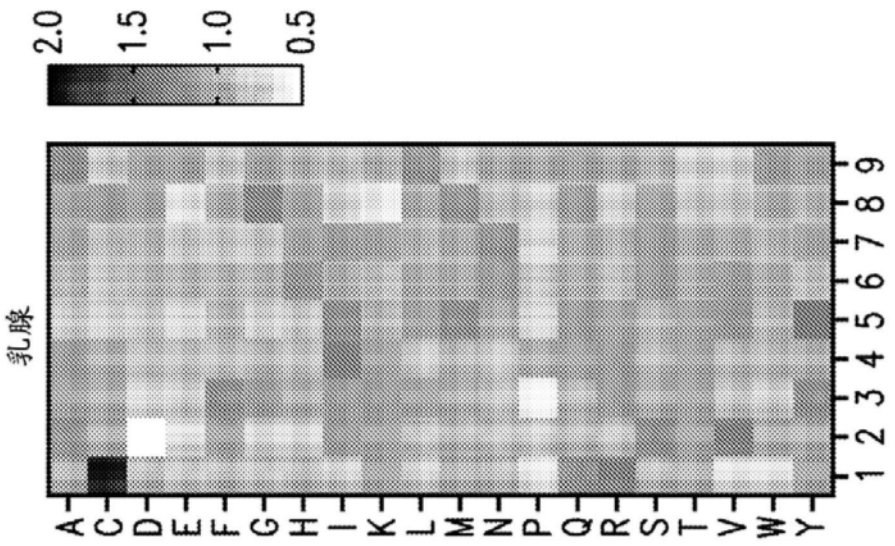

2.0
1.5
1.0
0.5
乳腺
1 2 3 4 5 6 7 8 9
A C D E F G H I K L M N P Q R S T V W Y

图18I

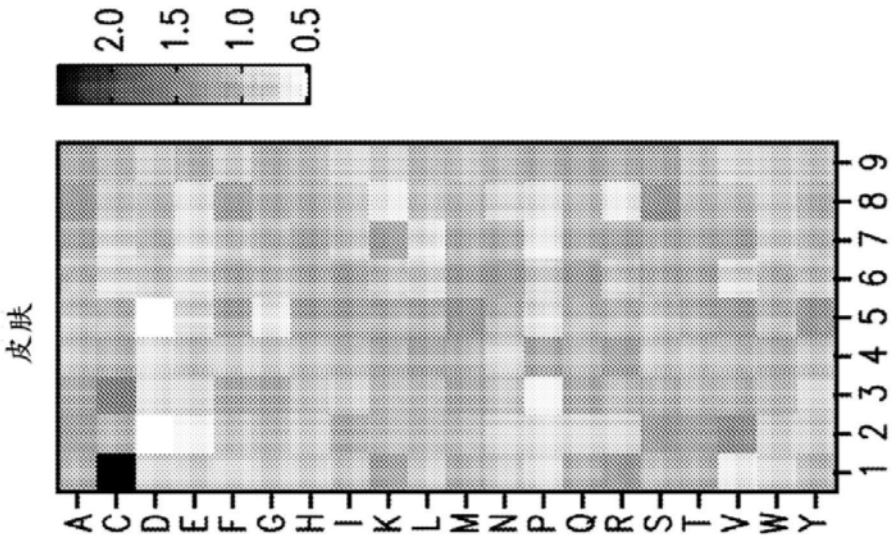

2.0
1.5
1.0
0.5
皮肤
1 2 3 4 5 6 7 8 9
A C D E F G H I K L M N P Q R S T V W Y

图18J

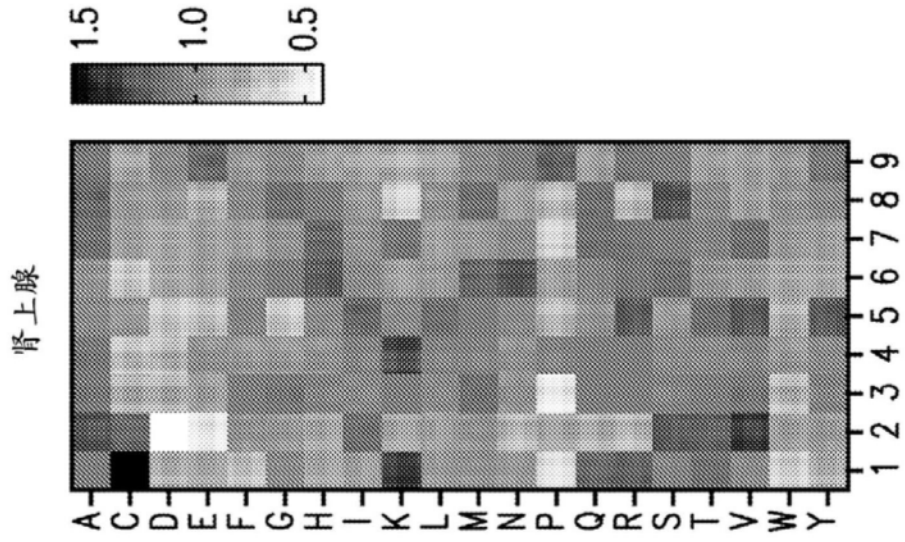

1.5
1.0
0.5
肾上腺
1 2 3 4 5 6 7 8 9
A C D E F G H I K L M N P Q R S T V W Y

图18K

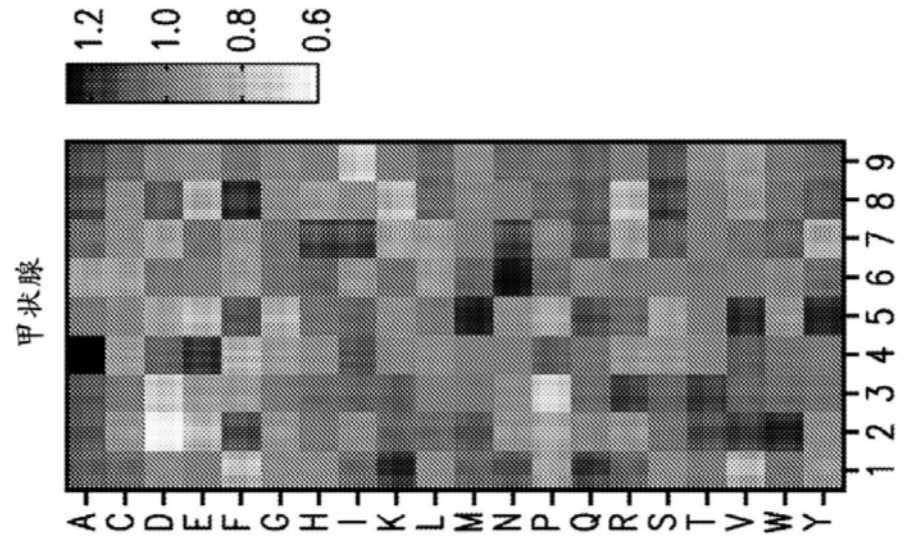

1.2
1.0
0.8
0.6
甲状腺
1 2 3 4 5 6 7 8 9
A C D E F G H I K L M N P Q R S T V W Y

图18L

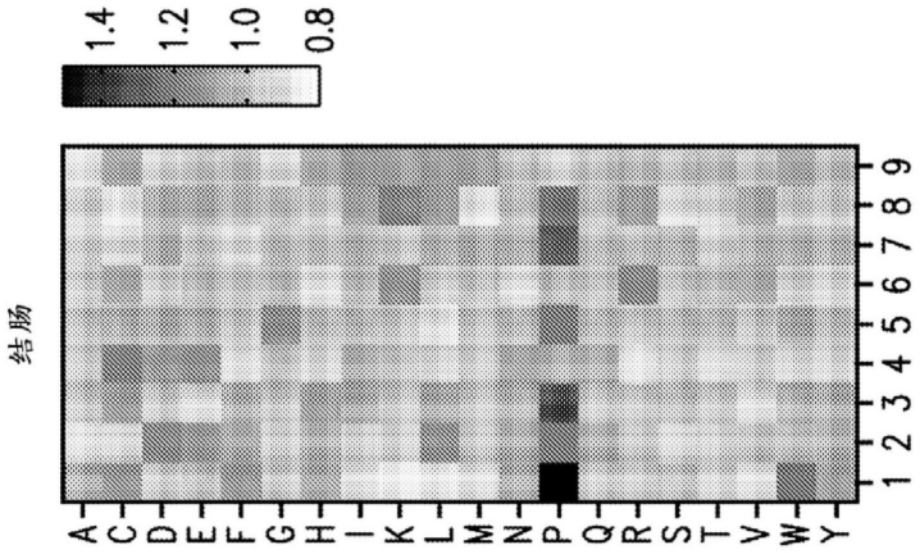

结肠
1.4
1.2
1.0
0.8
A C D E F G H I K L M N P Q R S T V W Y
1 2 3 4 5 6 7 8 9

图18M

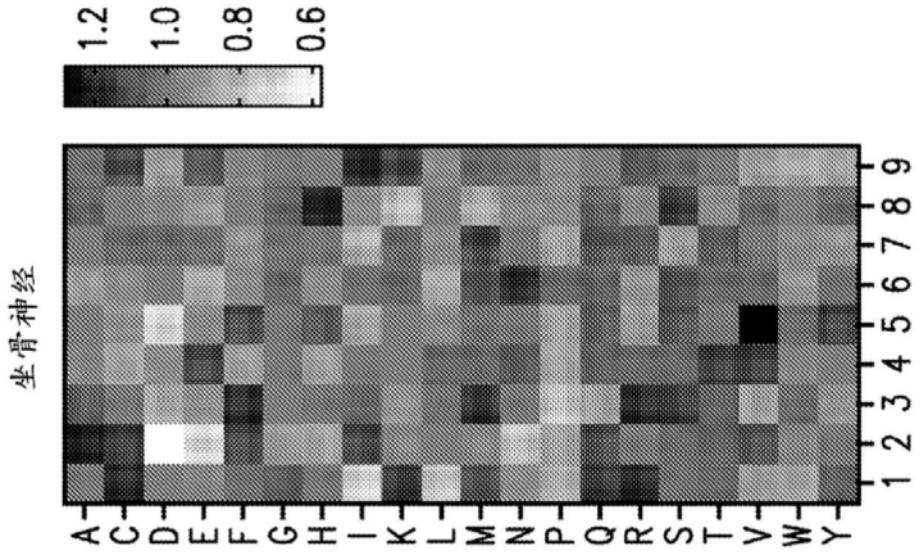

坐骨神经
1.2
1.0
0.8
0.6
A C D E F G H I K L M N P Q R S T V W Y
1 2 3 4 5 6 7 8 9

图18N

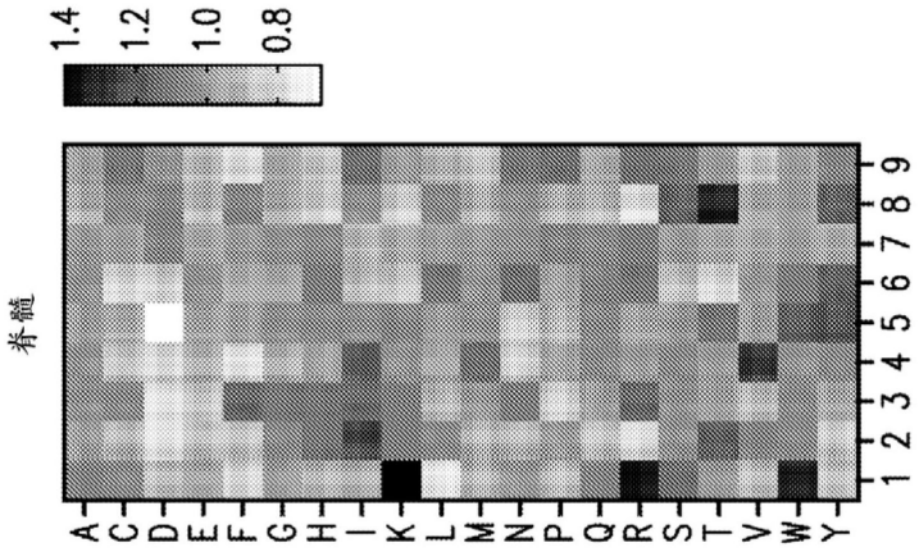


图180

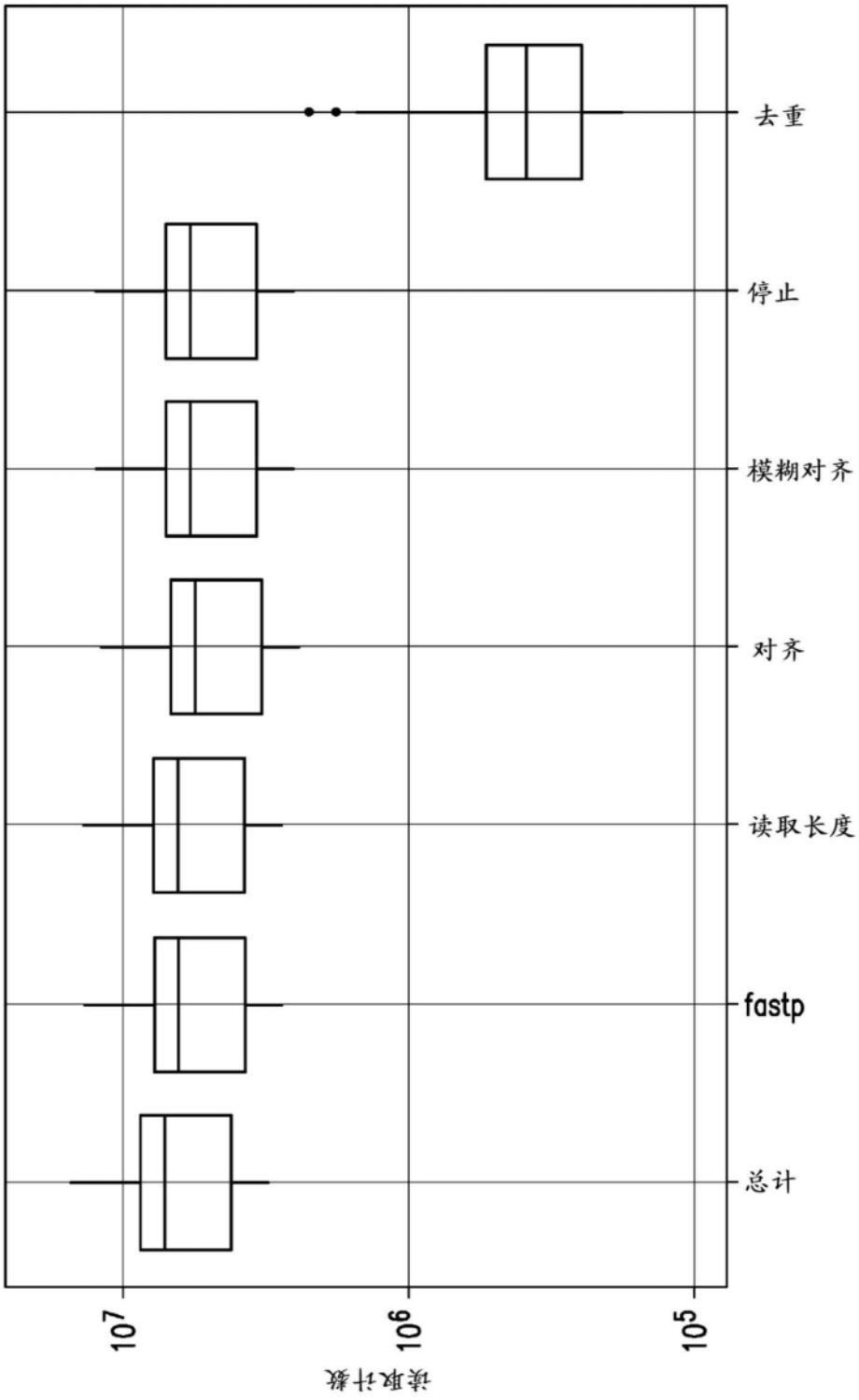


图19A

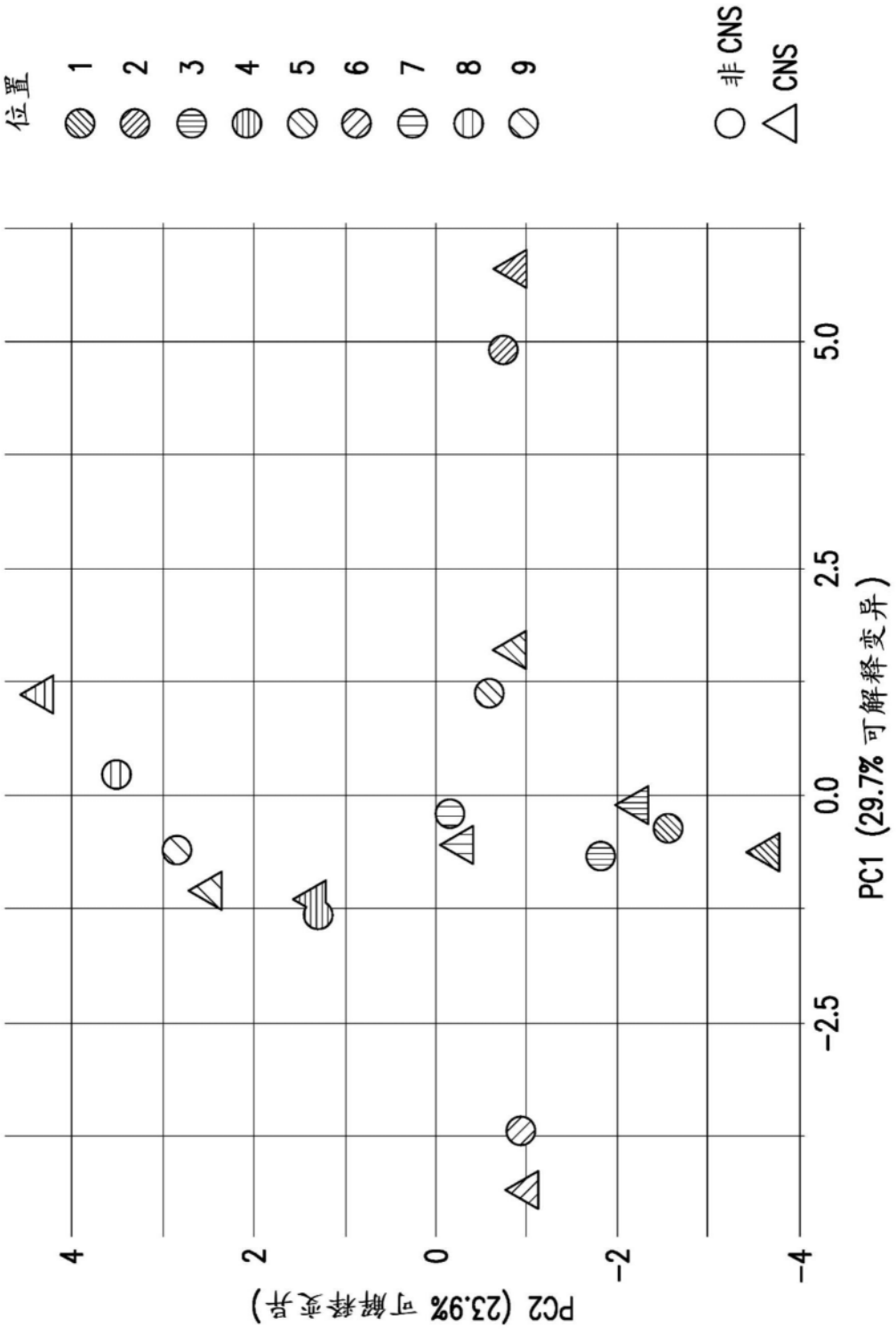
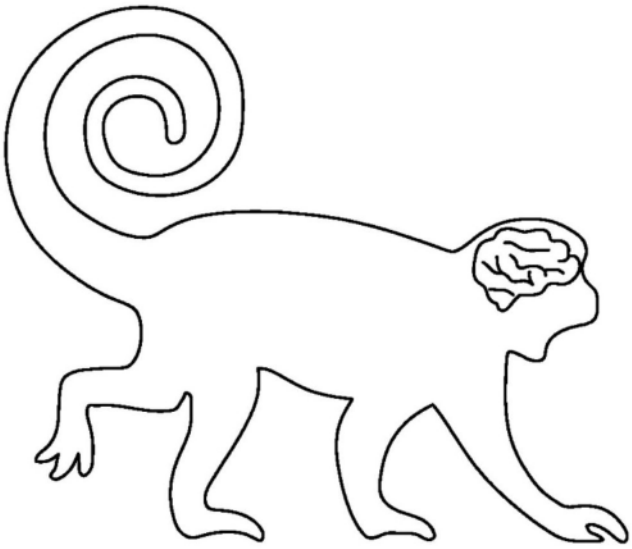

位置
1
2
3
4
5
6
7
8
9
非 CNS
CNS
PC2 (23.9% 可解释变异)
4
2
0
−2
−4
−2.5
0.0
2.5
5.0
PC1 (29.7% 可解释变异)

图19B



非CNS:	
皮肤	CNS:
肝脏	皮层(前脑)
肺	皮层(枕叶)
心脏	皮层(颞叶)
脾脏	丘脑
淋巴结	下丘脑
甲状腺	黑质
骨骼肌	海马DG
骨髓	海马CA1
乳腺	海马CA3
肾上腺	小脑
结肠	

图20A

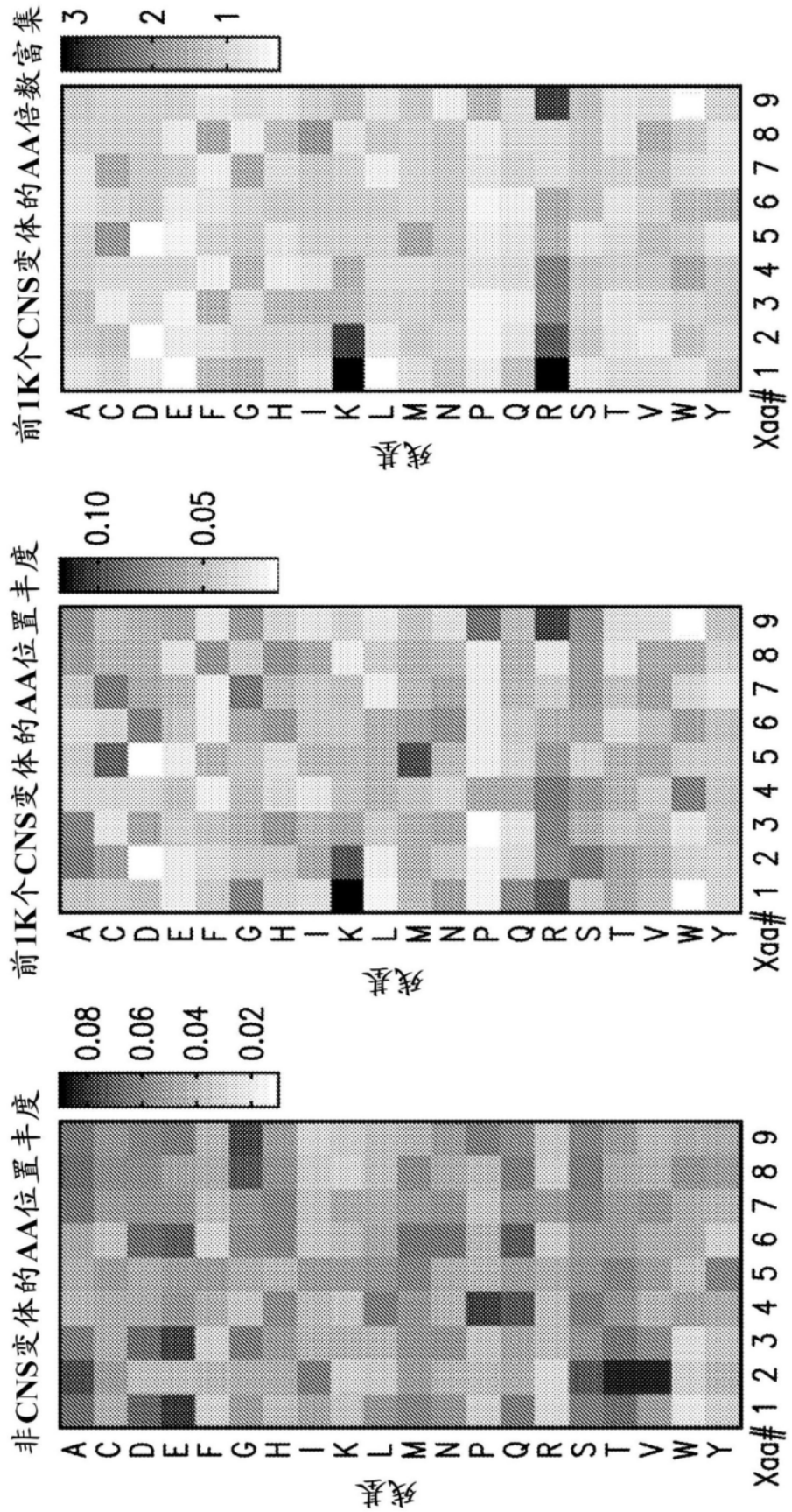

非CNS变体的AA位置丰度
0.08
0.06
0.04
0.02
残基
A C D E F G H I K L M N P Q R S T V W Y
Xaa# 1 2 3 4 5 6 7 8 9
前1K个CNS变体的AA位置丰度
0.10
0.05
残基
A C D E F G H I K L M N P Q R S T V W Y
Xaa# 1 2 3 4 5 6 7 8 9
前1K个CNS变体的AA倍数富集
3
2
1
残基
A C D E F G H I K L M N P Q R S T V W Y
Xaa# 1 2 3 4 5 6 7 8 9

图20B

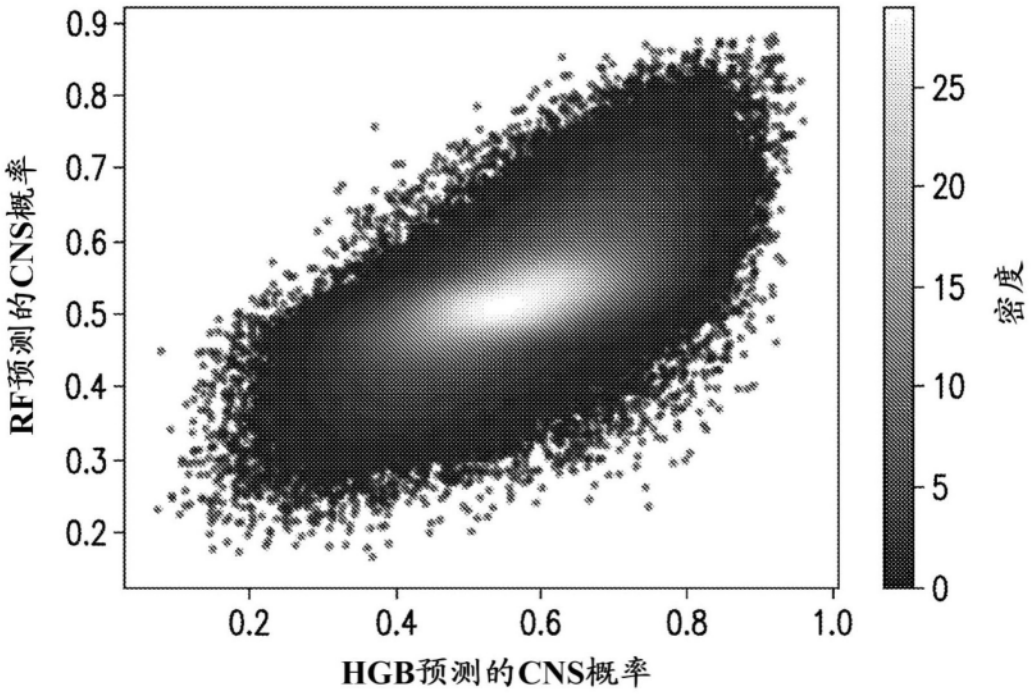

RF预测的CNS概率
HGB预测的CNS概率
密度
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.2
0.4
0.6
0.8
1.0
25
20
15
10
5
0

图21A

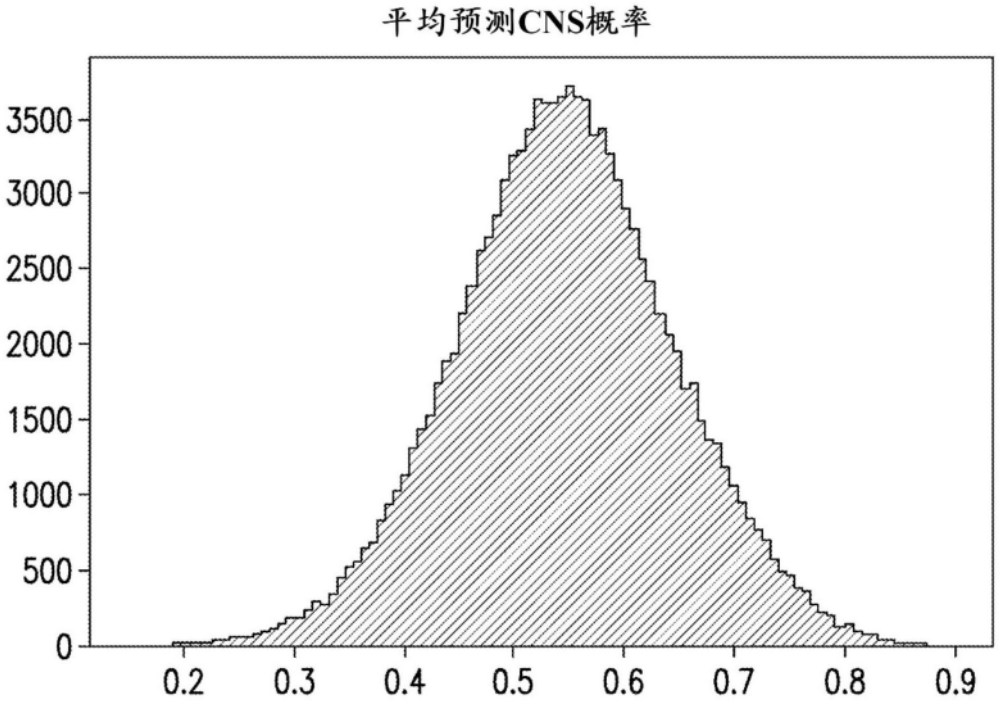

平均预测CNS概率
3500
3000
2500
2000
1500
1000
500
0
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9

图21B

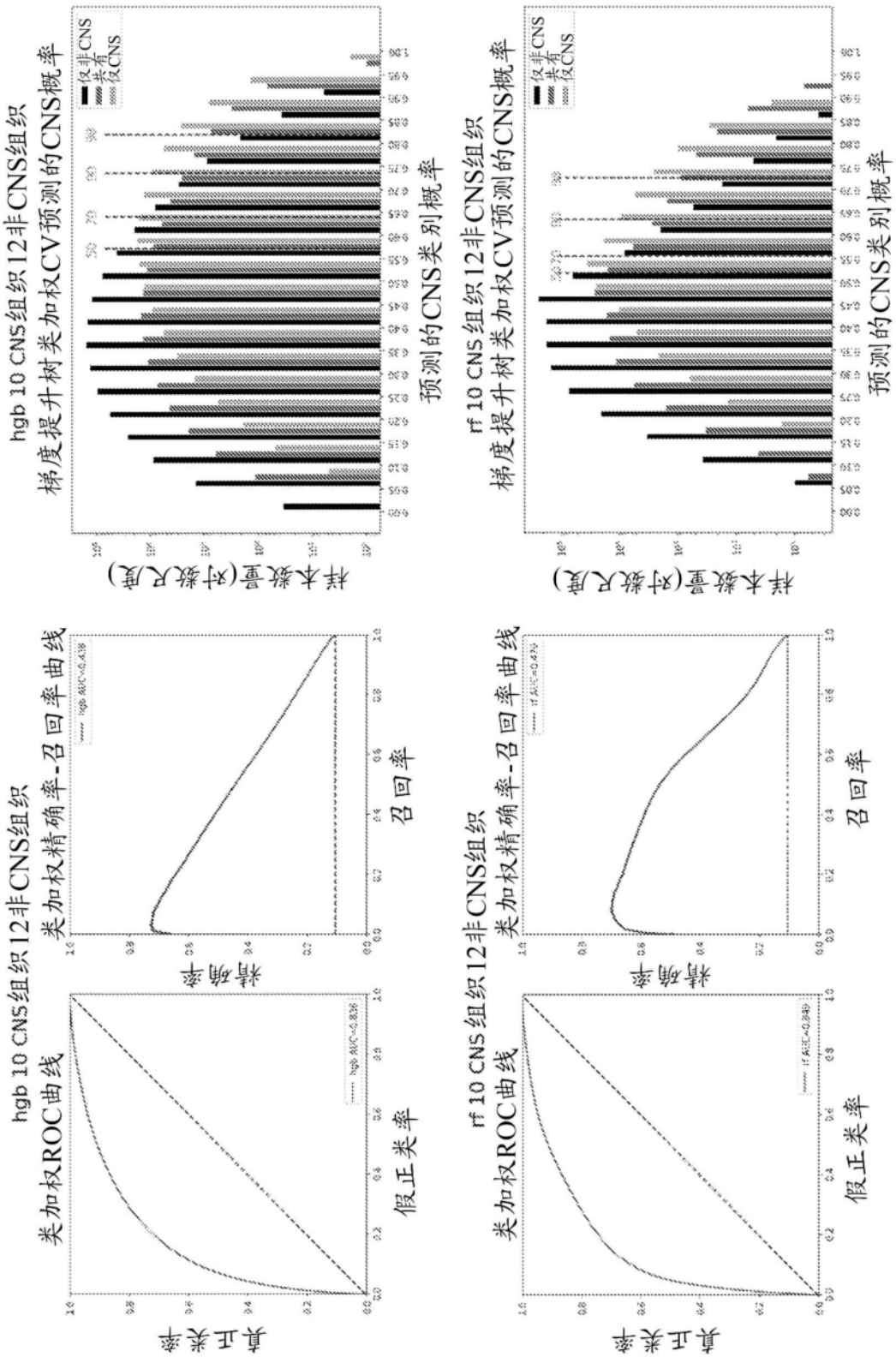

hgb 10 CNS组织12非CNS组织
类加权ROC曲线
真正类率
假正类率
hgb AUC=0.838
hgb 10 CNS组织12非CNS组织
类加权精确率-召回率曲线
精确率
召回率
hgb AUC=0.438
hgb 10 CNS组织12非CNS组织
梯度提升树类加权CV预测的CNS概率
样本数量(对数尺度)
预测的CNS类别概率
仅非CNS
共有
仅CNS
rf 10 CNS组织12非CNS组织
类加权ROC曲线
rf AUC=0.843
rf 10 CNS组织12非CNS组织
类加权精确率-召回率曲线
rf AUC=0.470
rf 10 CNS组织12非CNS组织
梯度提升树类加权CV预测的CNS概率

图22

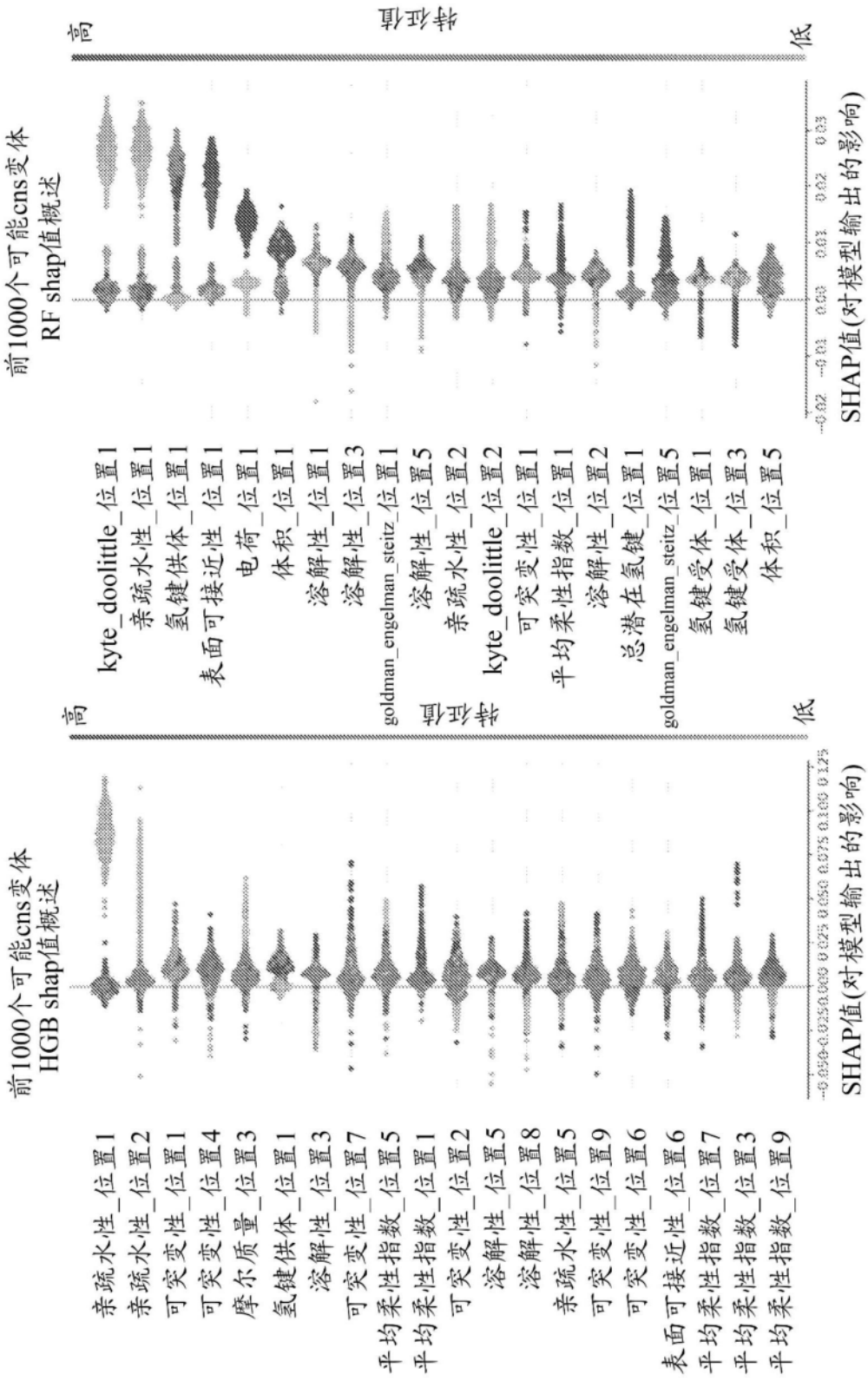

前1000个可能cns变体
HGB shap值概述
亲疏水性_位置1
亲疏水性_位置2
可突变性_位置1
可突变性_位置4
摩尔质量_位置3
氢键供体_位置1
溶解性_位置3
可突变性_位置7
平均柔性指数_位置5
平均柔性指数_位置1
可突变性_位置2
溶解性_位置5
溶解性_位置8
亲疏水性_位置5
可突变性_位置9
可突变性_位置6
表面可接近性_位置6
平均柔性指数_位置7
平均柔性指数_位置3
平均柔性指数_位置9
高
低
特征值
SHAP值(对模型输出的影响)
前1000个可能cns变体
RF shap值概述
kyte_doolittle_位置1
亲疏水性_位置1
氢键供体_位置1
表面可接近性_位置1
电荷_位置1
体积_位置1
溶解性_位置1
溶解性_位置3
goldman_engelman_steitz_位置1
溶解性_位置5
亲疏水性_位置2
kyte_doolittle_位置2
可突变性_位置1
平均柔性指数_位置1
溶解性_位置2
总潜在氢键_位置1
goldman_engelman_steitz_位置5
氢键受体_位置1
氢键受体_位置3
体积_位置5
高
低
特征值
SHAP值(对模型输出的影响)

图23

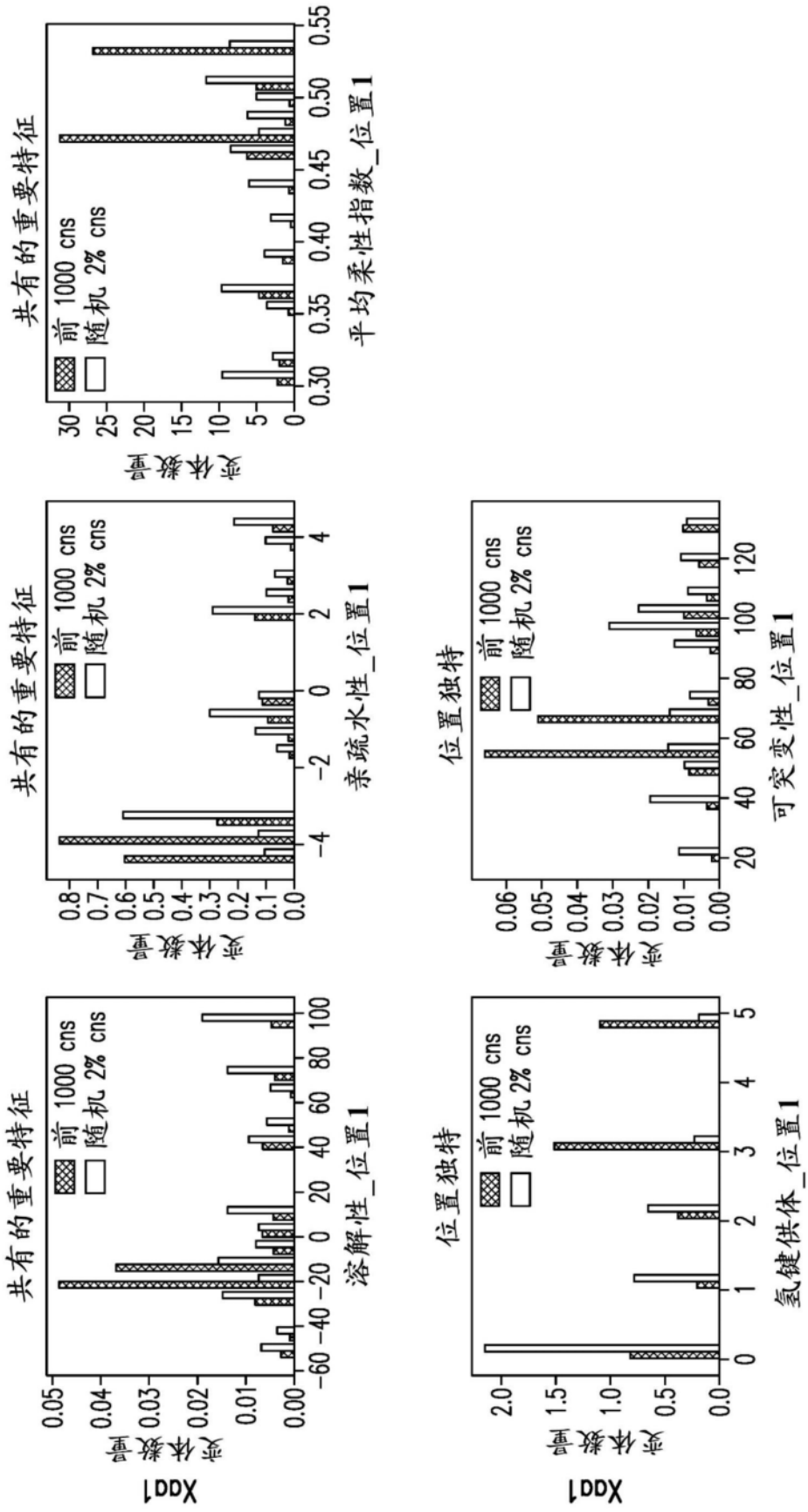

共有的重要特征
前 1000 cns
随机 2% cns
XGB1
特征重要性
溶解性_位置1
亲疏水性_位置1
平均柔性指数_位置1
位置独特
氢键供体_位置1
可突变性_位置1

图24A

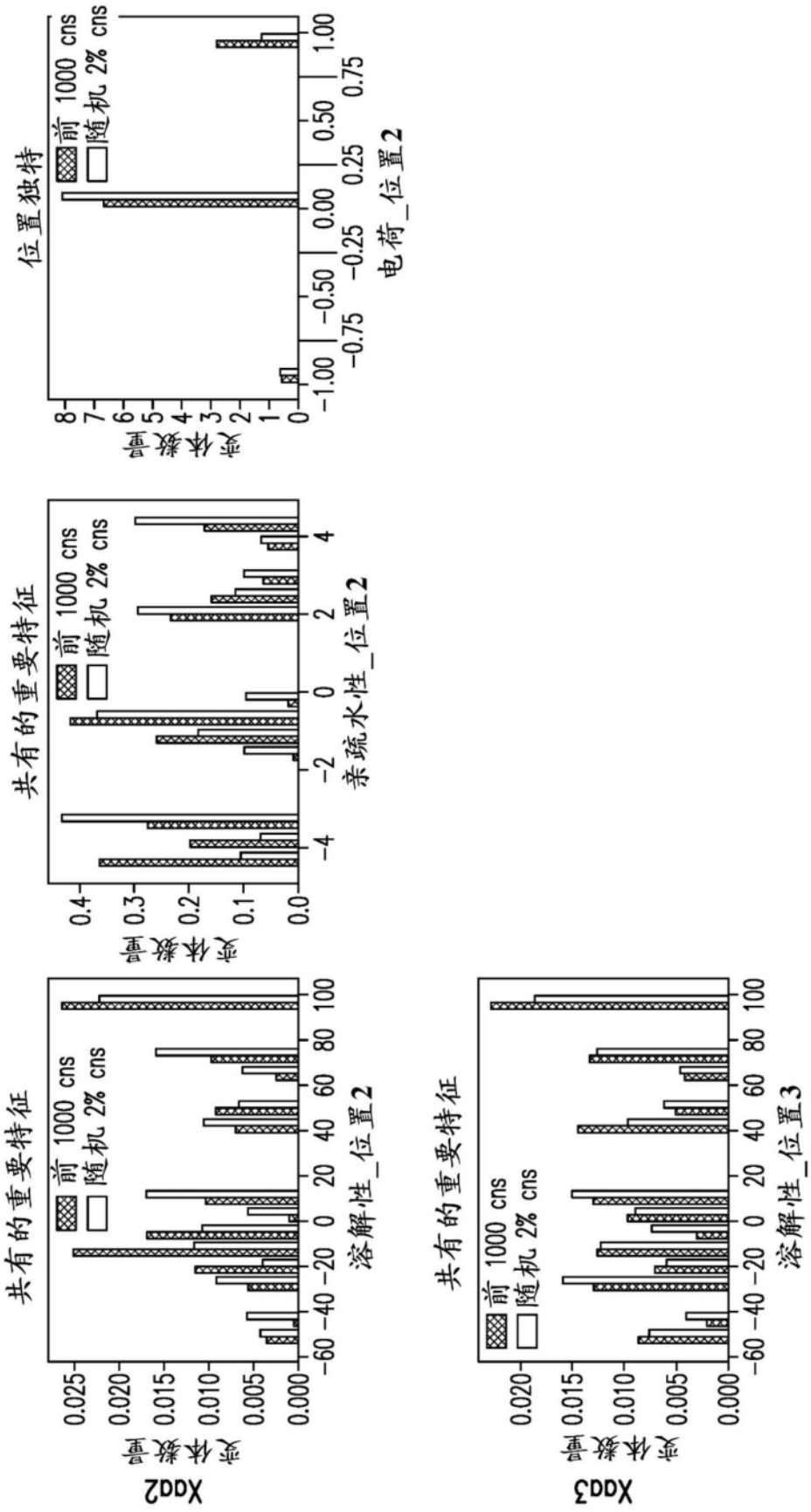

位置独特
前 1000 cns
随机 2% cns
特征贡献
电荷_位置2
共有的重要特征
亲疏水性_位置2
溶解性_位置2
XDD2
共有的重要特征
溶解性_位置3
XDD3

图24B

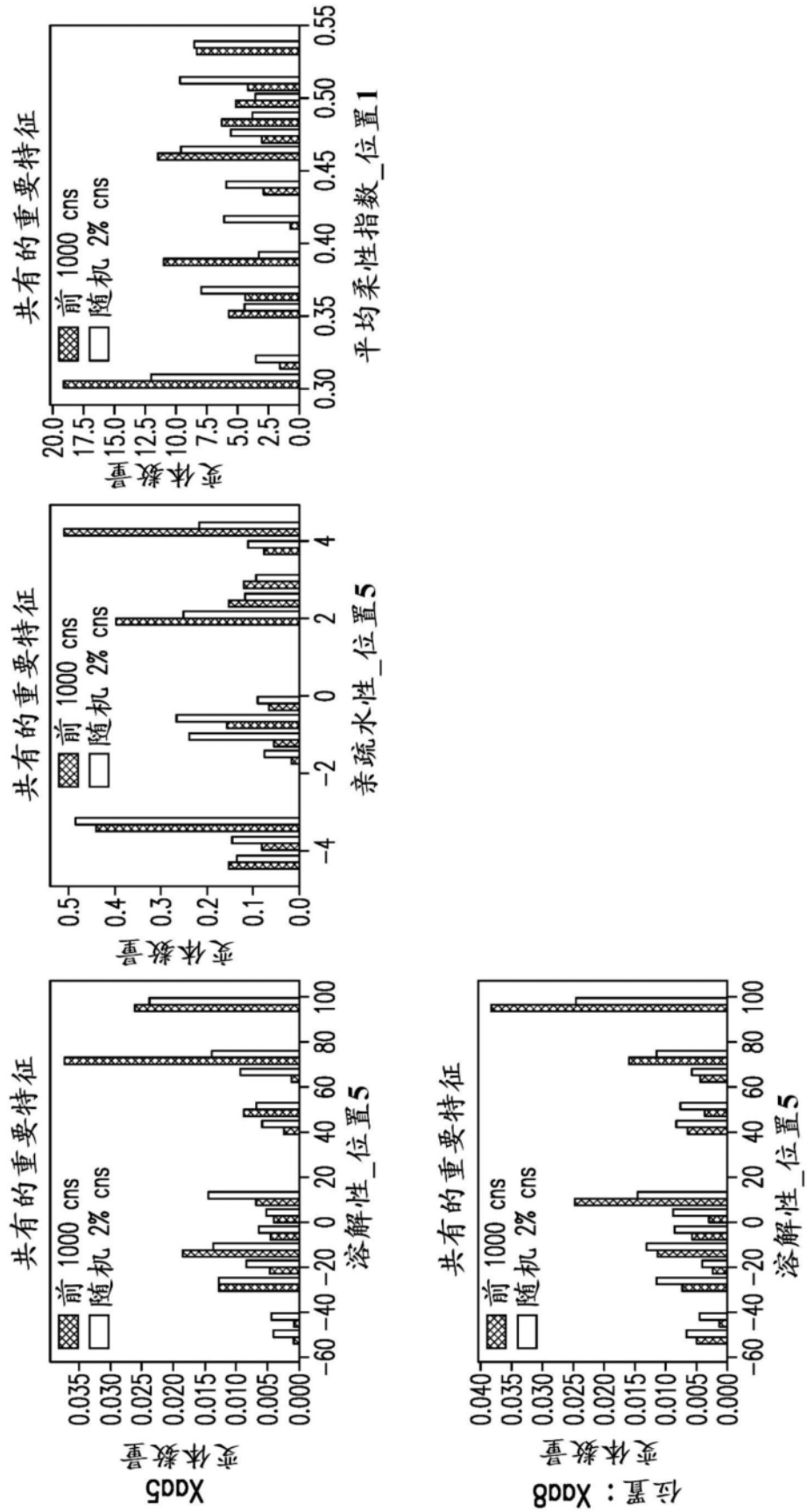

共有的重要特征
前 1000 cns
随机 2% cns
溶体数量
平均柔性指数_位置1
亲疏水性_位置5
溶解性_位置5
XDD5
XDD8

图24C

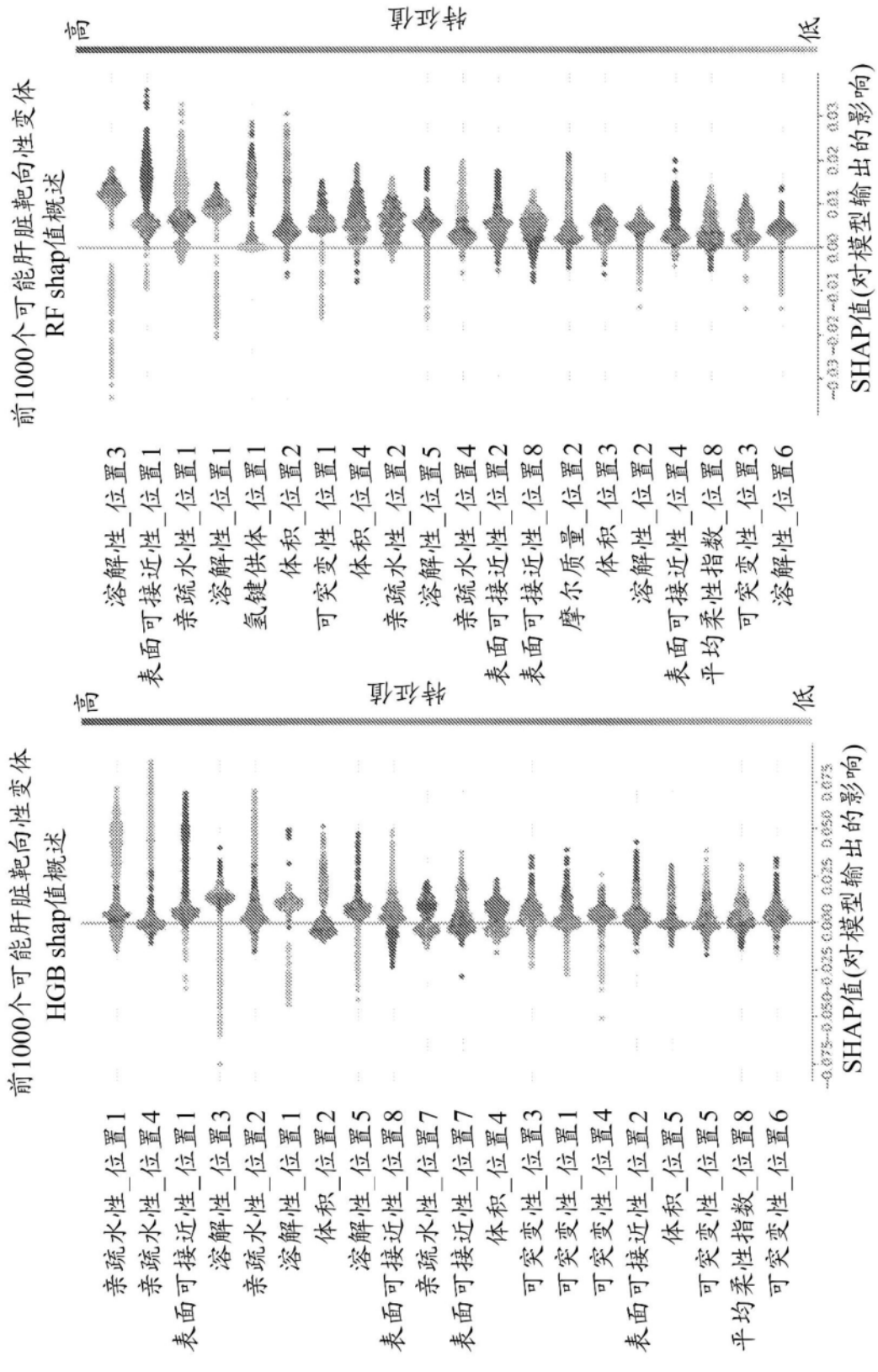

前1000个可能肝脏靶向性变体
HGB shap值概述
亲疏水性_位置1
亲疏水性_位置4
表面可接近性_位置1
溶解性_位置3
亲疏水性_位置2
溶解性_位置1
体积_位置2
溶解性_位置5
表面可接近性_位置8
亲疏水性_位置7
表面可接近性_位置7
体积_位置4
可突变性_位置3
可突变性_位置1
可突变性_位置4
表面可接近性_位置2
体积_位置5
可突变性_位置5
平均柔性指数_位置8
可突变性_位置6
高
特征值
低
-0.075 -0.050 -0.025 0.000 0.025 0.050 0.075
SHAP值(对模型输出的影响)
前1000个可能肝脏靶向性变体
RF shap值概述
溶解性_位置3
表面可接近性_位置1
亲疏水性_位置1
溶解性_位置1
氢键供体_位置1
体积_位置2
可突变性_位置1
体积_位置4
亲疏水性_位置2
溶解性_位置5
亲疏水性_位置4
表面可接近性_位置2
表面可接近性_位置8
摩尔质量_位置2
体积_位置3
溶解性_位置2
表面可接近性_位置4
平均柔性指数_位置8
可突变性_位置3
溶解性_位置6
高
特征值
低
-0.03 -0.02 -0.01 0.00 0.01 0.02 0.03
SHAP值(对模型输出的影响)

图25

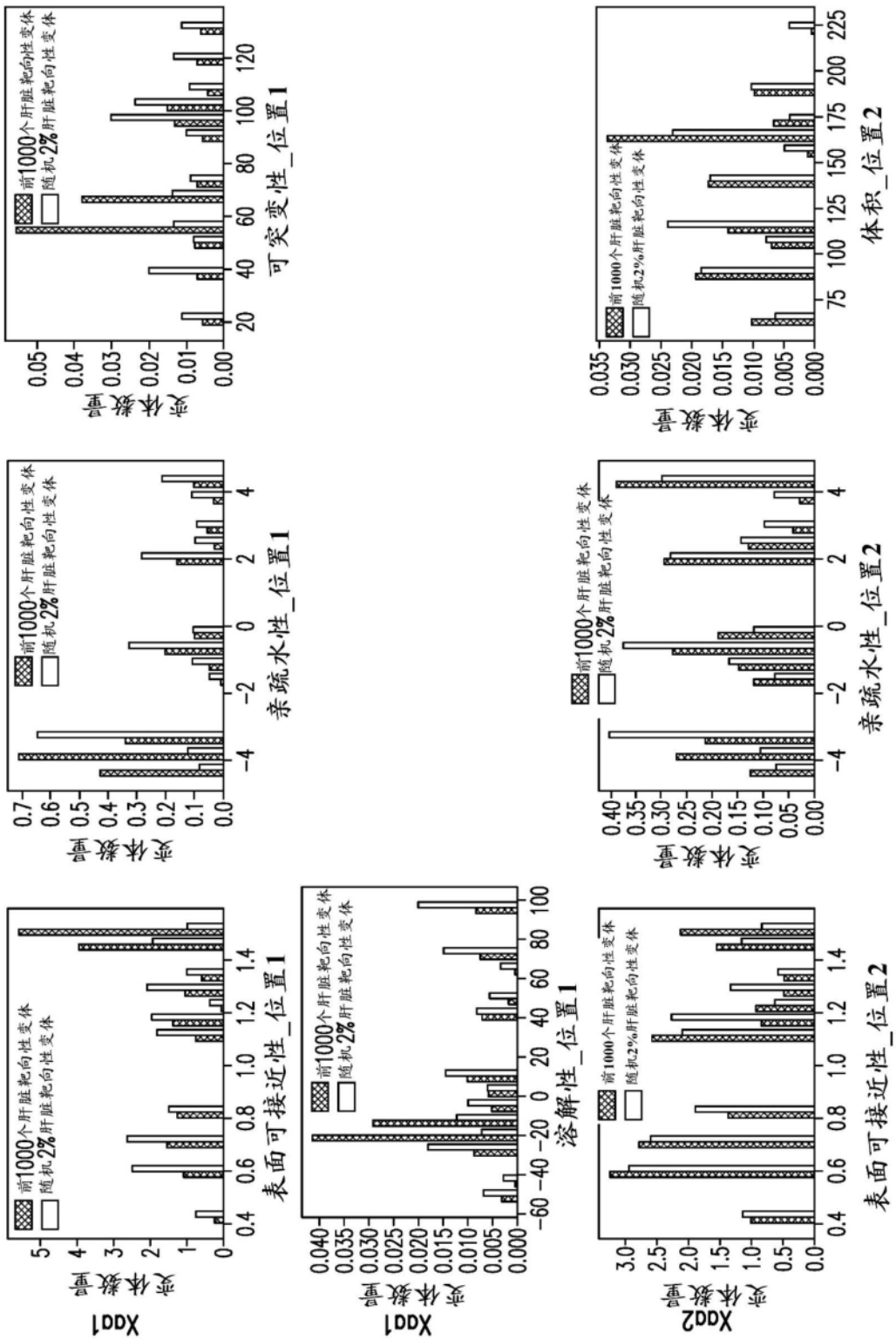

前1000个肝脏靶向性变体
随机2%肝脏靶向性变体
XGD1
XGD2
物体数量
表面可接近性_位置1
亲疏水性_位置1
溶解性_位置1
可突变性_位置1
表面可接近性_位置2
亲疏水性_位置2
体积_位置2

图26A

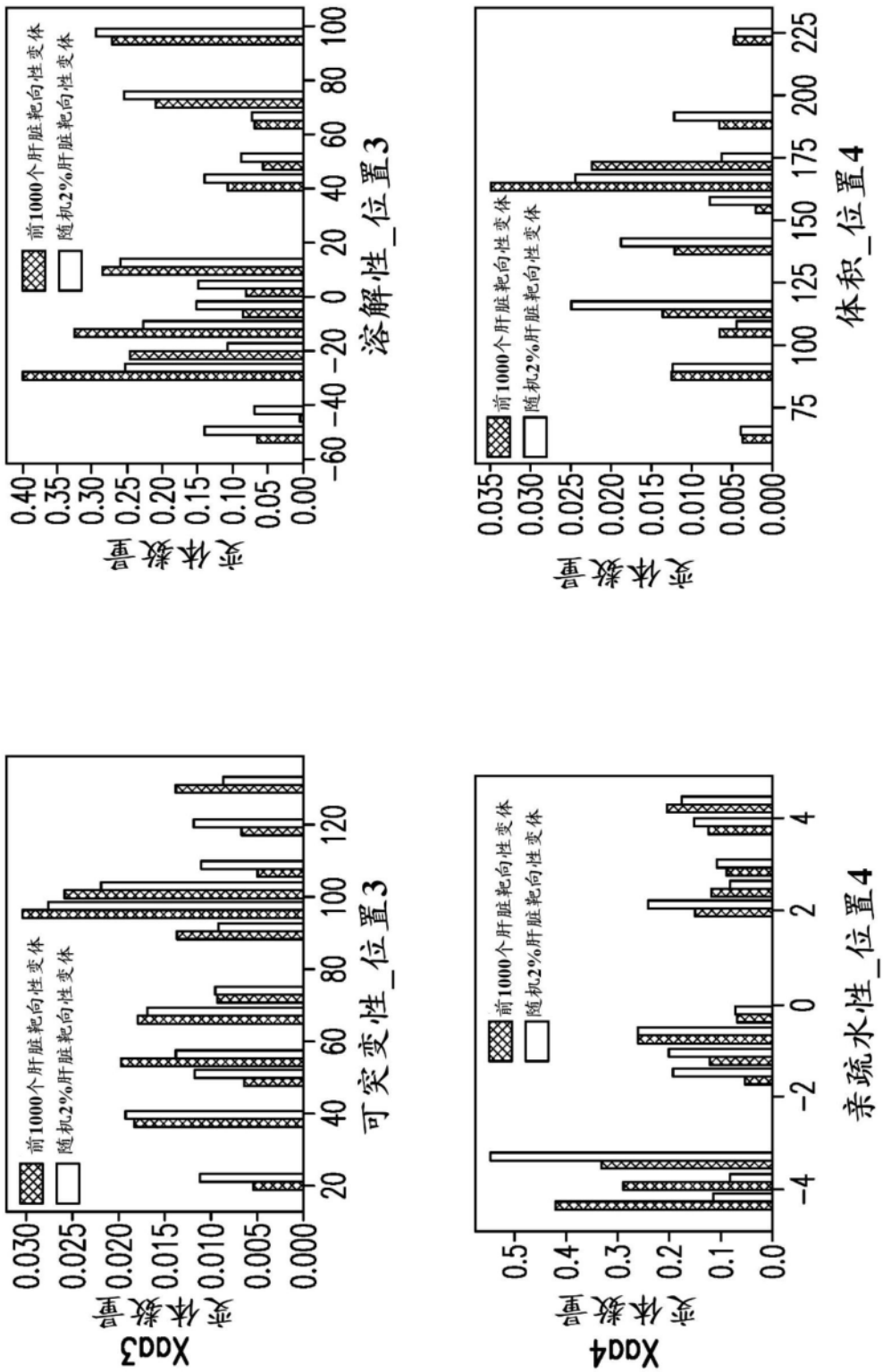

前1000个肝脏靶向性变体
随机2%肝脏靶向性变体
溶解性_位置3
体积_位置4
可突变性_位置3
亲疏水性_位置4
XDD3
XDD4
箱体数量

图26B

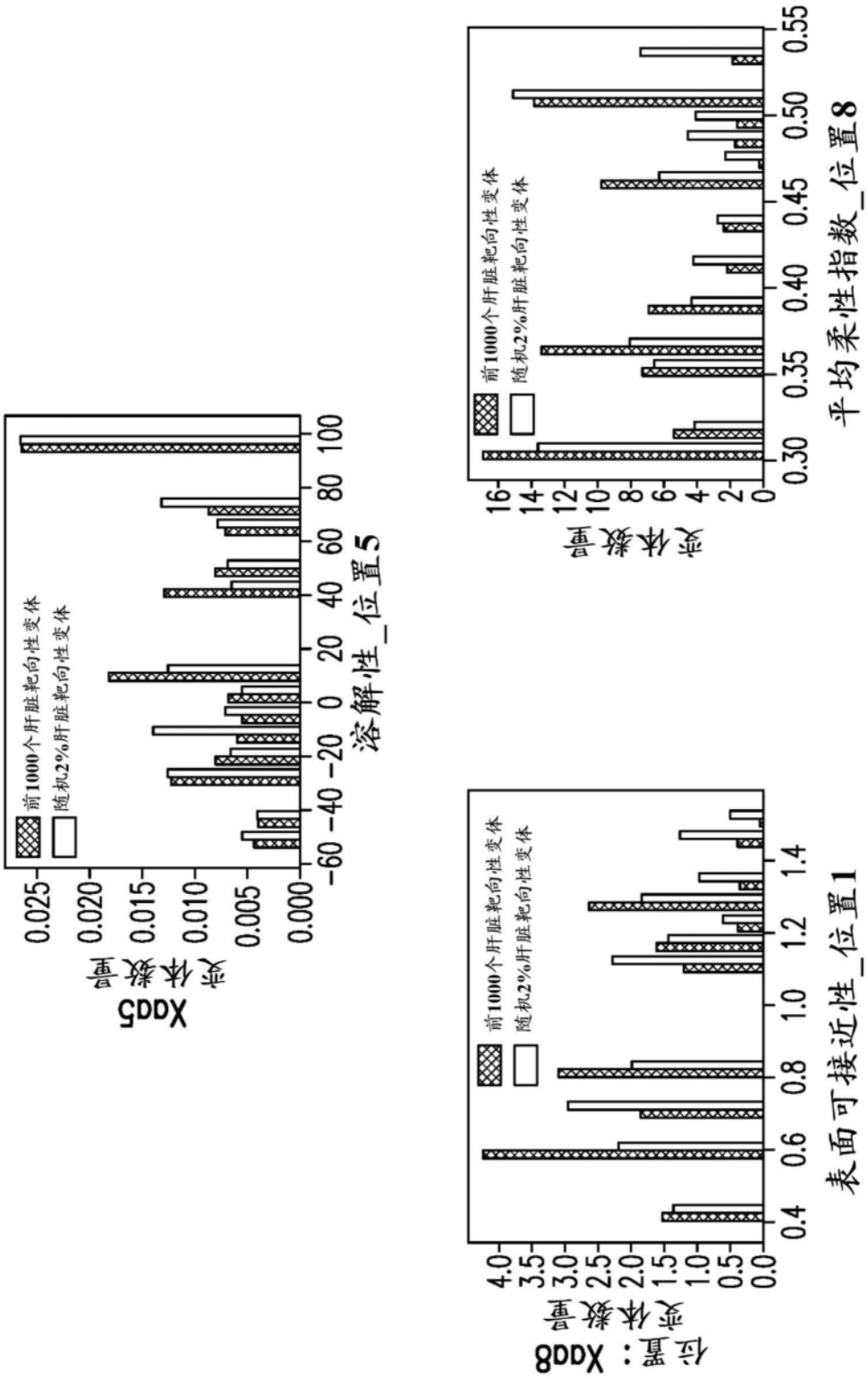

前1000个肝脏靶向性变体
随机2%肝脏靶向性变体
XDD5
溶体数量
溶解性_位置5
XDD8
溶体数量
表面可接近性_位置1
溶体数量
平均柔性指数_位置8

图26C

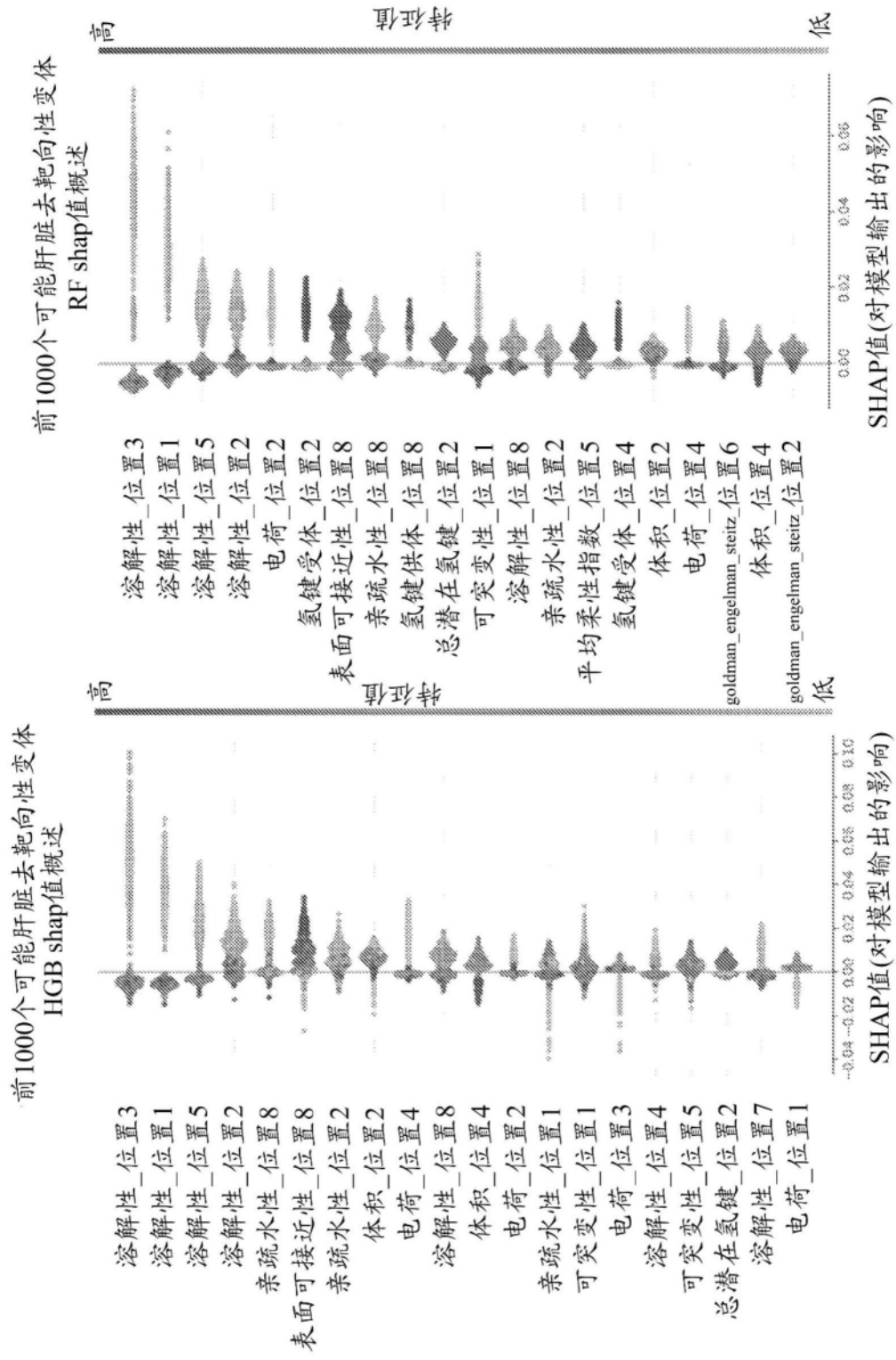

前1000个可能肝脏去靶向性变体
HGB shap值概述
溶解性_位置3
溶解性_位置1
溶解性_位置5
溶解性_位置2
亲疏水性_位置8
表面可接近性_位置8
亲疏水性_位置2
体积_位置2
电荷_位置4
溶解性_位置8
体积_位置4
电荷_位置2
亲疏水性_位置1
可突变性_位置1
电荷_位置3
溶解性_位置4
可突变性_位置5
总潜在氢键_位置2
溶解性_位置7
电荷_位置1
-0.04 -0.02 0.00 0.02 0.04 0.06 0.08 0.10
SHAP值(对模型输出的影响)
前1000个可能肝脏去靶向性变体
RF shap值概述
溶解性_位置3
溶解性_位置1
溶解性_位置5
溶解性_位置2
电荷_位置2
氢键受体_位置2
表面可接近性_位置8
亲疏水性_位置8
氢键供体_位置8
总潜在氢键_位置2
可突变性_位置1
溶解性_位置8
亲疏水性_位置2
平均柔性指数_位置5
氢键受体_位置4
体积_位置2
电荷_位置4
goldman_engelman_steitz_位置6
体积_位置4
goldman_engelman_steitz_位置2
高
特征值
低
0.00 0.02 0.04 0.06
SHAP值(对模型输出的影响)

图27

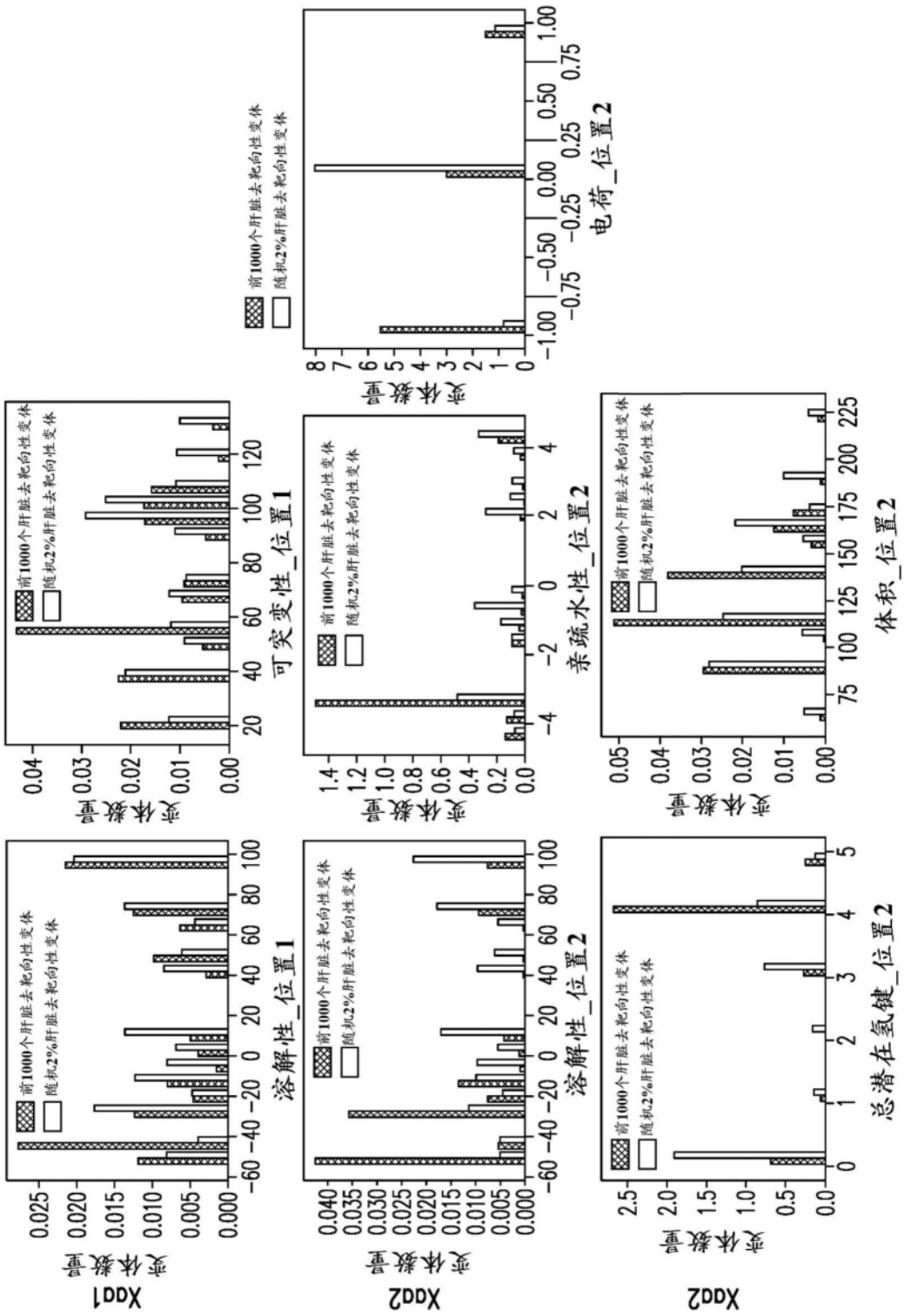

前1000个肝脏去靶向性变体
随机2%肝脏去靶向性变体
XGB1
XGB2
XGB2
穷体数量
溶解性_位置1
溶解性_位置2
可突变性_位置1
亲疏水性_位置2
电荷_位置2
体积_位置2
总潜在氢键_位置2

图28A

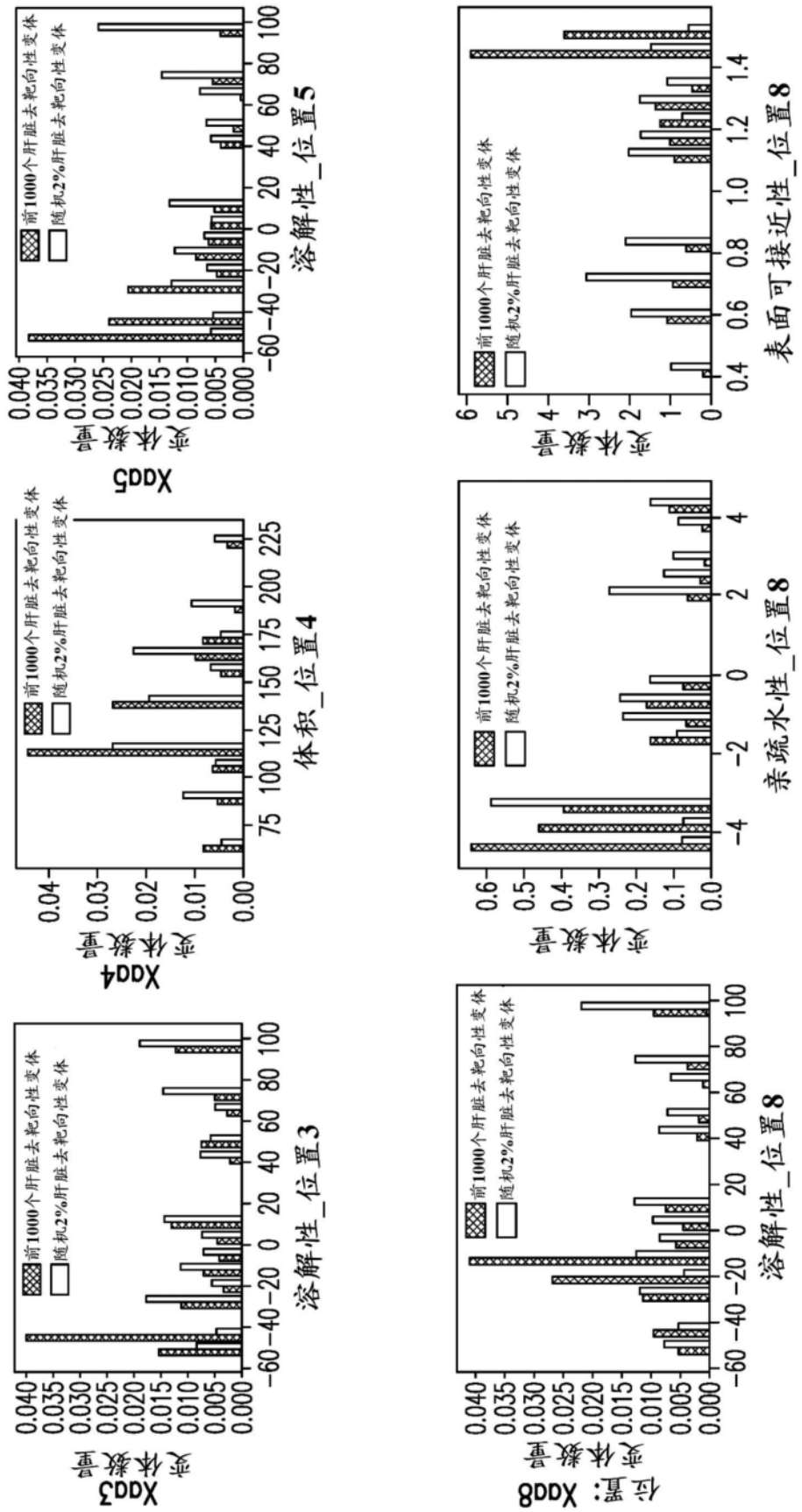

前1000个肝脏去靶向性变体
随机2%肝脏去靶向性变体
溶解性_位置3
体积_位置4
溶解性_位置5
亲疏水性_位置8
表面可接近性_位置8
溶解性_位置8
位置: XDD8
XDD3
XDD4
XDD5
穴体数量

图28B

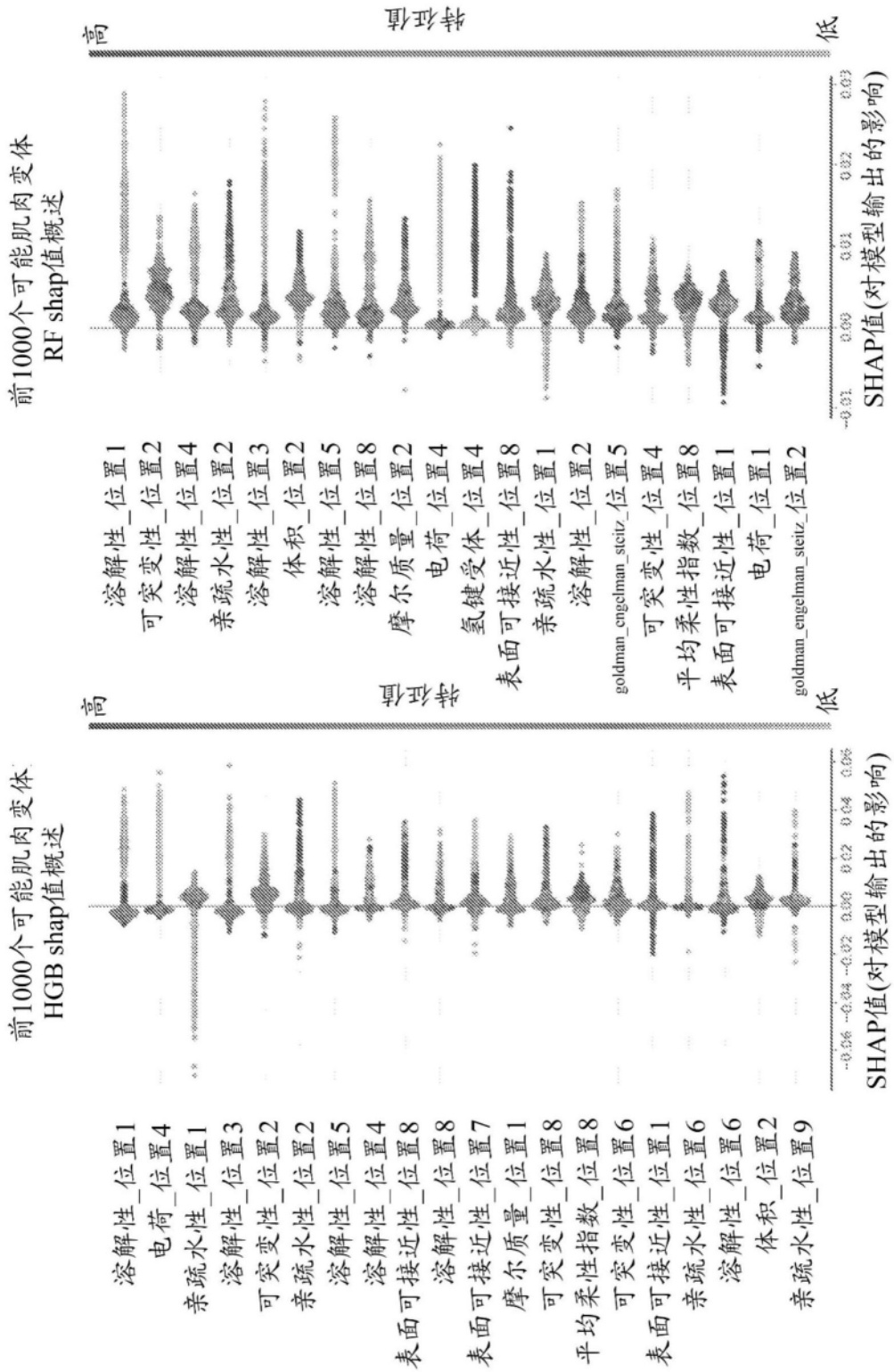

前1000个可能肌肉变体
HGB shap值概述
溶解性_位置1
电荷_位置4
亲疏水性_位置1
溶解性_位置3
可突变性_位置2
亲疏水性_位置2
溶解性_位置5
溶解性_位置4
表面可接近性_位置8
溶解性_位置8
表面可接近性_位置7
摩尔质量_位置1
可突变性_位置8
平均柔性指数_位置8
可突变性_位置6
表面可接近性_位置1
亲疏水性_位置6
溶解性_位置6
体积_位置2
亲疏水性_位置9
高
特征值
低
-0.06 -0.04 -0.02 0.00 0.02 0.04 0.06
SHAP值(对模型输出的影响)
前1000个可能肌肉变体
RF shap值概述
溶解性_位置1
可突变性_位置2
溶解性_位置4
亲疏水性_位置2
溶解性_位置3
体积_位置2
溶解性_位置5
溶解性_位置8
摩尔质量_位置2
电荷_位置4
氢键受体_位置4
表面可接近性_位置8
亲疏水性_位置1
溶解性_位置2
goldman_engelman_steiz_位置5
可突变性_位置4
平均柔性指数_位置8
表面可接近性_位置1
电荷_位置1
goldman_engelman_steiz_位置2
高
特征值
低
-0.01 0.00 0.01 0.02 0.03
SHAP值(对模型输出的影响)

图29

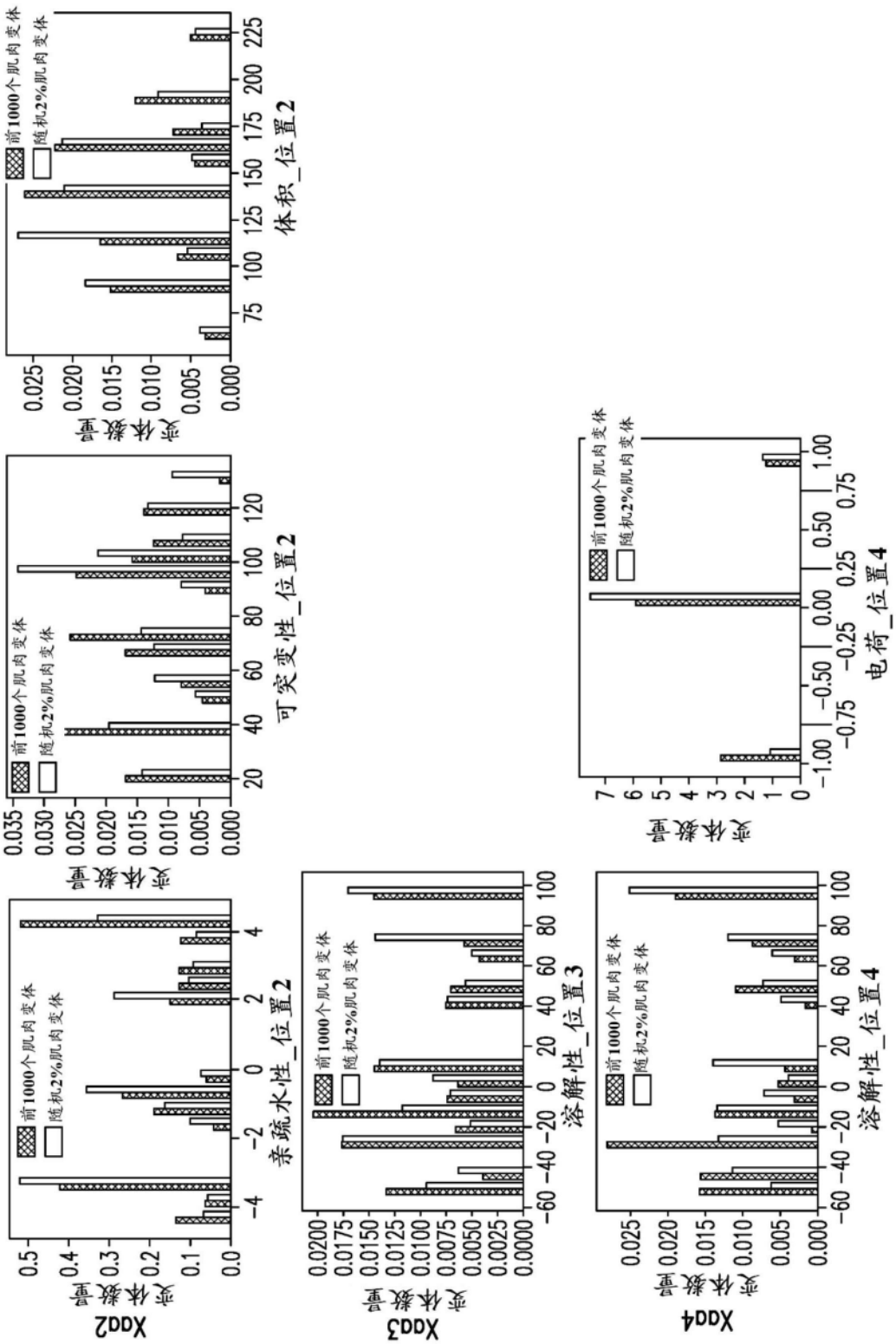

前1000个肌肉变体
随机2%肌肉变体
体积_位置2
可突变性_位置2
电荷_位置4
亲疏水性_位置2
溶解性_位置3
溶解性_位置4
特体数量
XQQ2
XQQ3
XQQ4

图30A

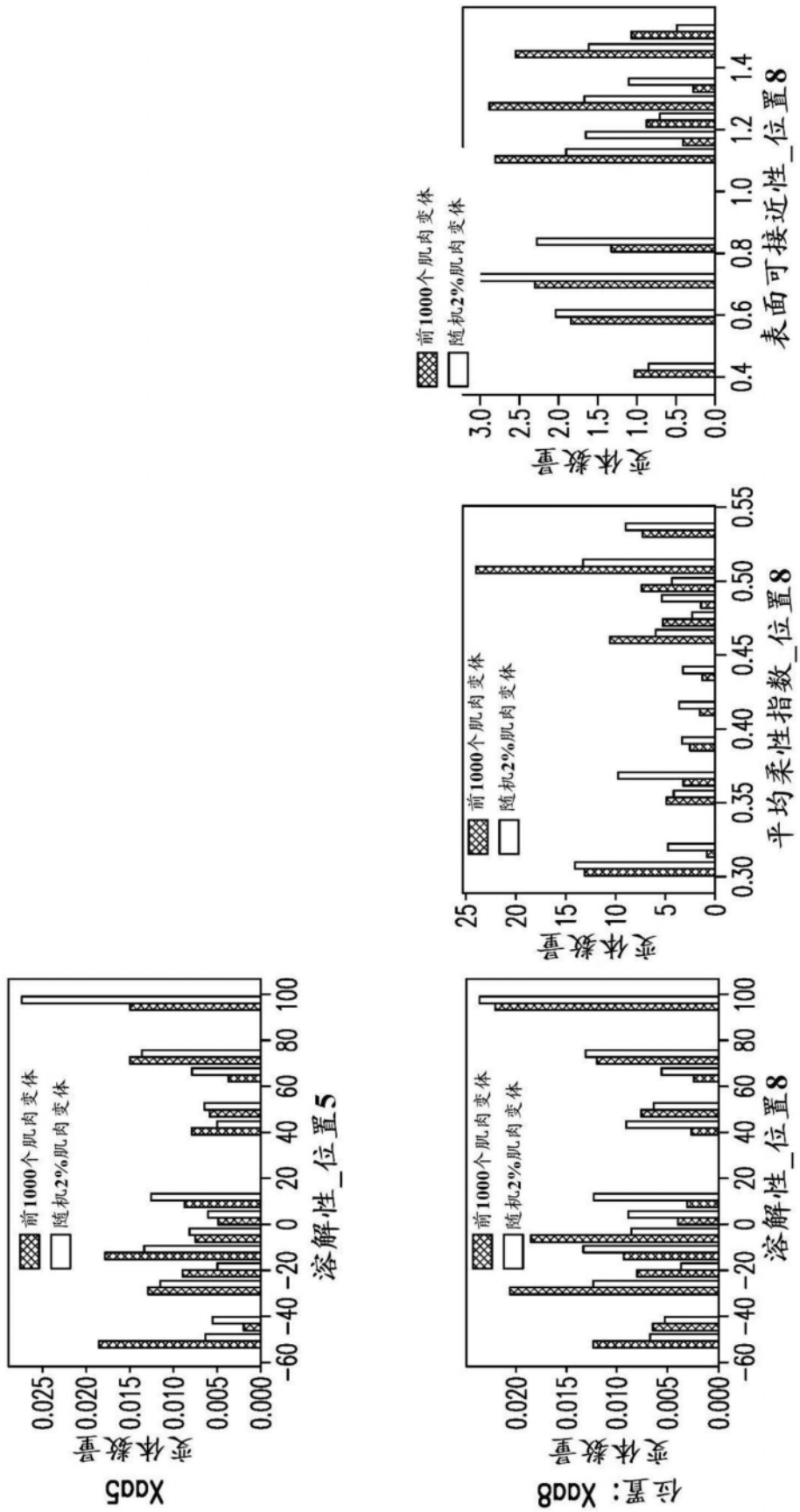

前1000个肌肉变体
随机2%肌肉变体
溶体数量
溶解性_位置5
溶解性_位置8
平均柔性指数_位置8
表面可接近性_位置8
位置: XQQ8
XQQ5

图30B