(19)国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 116669760 A
(43)申请公布日 2023.08.29

(21)申请号 202180081878.0
(22)申请日 2021.12.03
(30)优先权数据
63/121,164 2020.12.03 US
(85)PCT国际申请进入国家阶段日
2023.06.05
(86)PCT国际申请的申请数据
PCT/US2021/061876 2021.12.03
(87)PCT国际申请的公布数据
WO2022/120217 EN 2022.06.09
(71)申请人 磨石生物公司
地址 美国加利福尼亚州
(72)发明人 K·朱斯 C·D·斯卡伦 L·吉特林 A·弗格森 R·卢梭 R·耶冷斯凯 J·X·孙 M·J·戴维斯 A·R·拉帕波特 C·D·帕尔默
(74)专利代理机构 北京市金杜律师事务所 11256
专利代理师 陈文平
(51)Int.Cl.
A61K 39/39(2006.01)

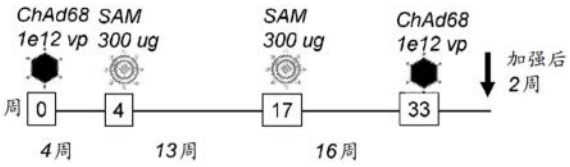
权利要求书18页 说明书128页
序列表（电子公布） 附图14页

(54)发明名称

同源腺病毒疫苗接种

(57)摘要

公开了与组合物相关的核苷酸、细胞和方法，包括它们作为疫苗的用途，包括用于同源初免/加强疫苗接种策略的载体和方法，诸如用于同源初免/加强疫苗接种策略的腺病毒载体。



CN 116669760 A

1.一种用于向受试者递送包含黑猩猩腺病毒(ChAdV)载体的组合物的方法,所述方法包括向所述受试者施用多个剂量的所述组合物,其中所述多个剂量至少包含第一剂量和第二剂量,并且其中所述第一剂量和所述第二剂量之间的时间段为至少27周。

2.一种用于向受试者递送包含黑猩猩腺病毒(ChAdV)载体的组合物的方法,所述方法包括向所述受试者施用多个剂量的所述组合物,其中所述多个剂量至少包含第一剂量和第二剂量,并且其中在施用第二剂量之前,确定ChAdV特异性中和抗体滴度低于中和阈值。

3.一种用于递送包含黑猩猩腺病毒(ChAdV)载体的组合物至受试者的方法,所述方法包括:

(a)向所述受试者施用第一剂量的所述组合物;

(b)确定或已经确定ChAdV特异性中和抗体滴度;和

(c)当确定ChAdV特异性中和抗体滴度低于中和阈值时,向所述受试者施用第二剂量的所述组合物。

4.如权利要求1-3中任一项所述的方法,其中所述第一剂量和所述第二剂量之间的时间段为至少27、至少28、至少29、至少30、至少31或至少32周。

5.如权利要求1-4中任一项所述的方法,其中所述第一剂量是初免剂量。

6.如权利要求1-5中任一项所述的方法,其中所述多个剂量包括三个或更多个剂量。

7.如权利要求6所述的方法,其中所述多个剂量中的一个或多个在所述第一剂量之前施用。

8.如权利要求1-7中任一项所述的方法,其中在所述第一剂量和所述第二剂量之间不施用额外剂量的所述组合物。

9.如权利要求2-8中任一项所述的方法,其中所述中和抗体滴度是NT50值,其计算为来自中和ChAdV病毒50%的所述免疫受试者血清的最小稀释度。

10.如权利要求9所述的方法,其中所述中和阈值是900或更小的NT50值。

11.如权利要求2-10中任一项所述的方法,其中确定所述中和抗体滴度包括以下步骤:

(1)在足以中和所述ChAdV病毒的条件下,使一种或多种稀释度的所述免疫受试者的血清与ChAdV病毒接触;和

(2)评估所述ChAdV病毒相对于未中和病毒的中和。

12.如权利要求11所述的方法,其中所述评估中和步骤包括测定由所述ChAdV病毒表达的报告构建体的表达。

13.如权利要求2-12中任一项所述的方法,其中所述中和阈值是在所述第二剂量中完全中和所述ChAdV病毒的最小中和抗体滴度。

14.如权利要求2-13中任一项所述的方法,其中所述中和阈值是所述第二剂量在所述受试者中诱导免疫响应的最小中和抗体滴度。

15.如以上方法权利要求中任一项所述的方法,其中所述ChAdV载体编码至少一种抗原。

16.如权利要求15所述的方法,其中所述至少一种抗原是非自身来源的肽,其中所述非自身来源的肽不由所述受试者的野生型基因编码。

17.如权利要求15或16所述的方法,其中所述至少一种抗原是肿瘤相关抗原。

18.如权利要求15-17中任一项所述的方法,其中所述肿瘤相关抗原是新抗原。

19.如权利要求15所述的方法,其中所述至少一种抗原是外来抗原。

20.如权利要求19所述的方法,其中所述外来抗原来自病原体、病毒、细菌、真菌或寄生虫。

21.如权利要求15-20中任一项所述的方法,其中所述多个剂量中的每一个包含相同的一种或多种抗原。

22.如以上方法权利要求中任一项所述的方法,其中肌内(IM)、皮内(ID)、皮下(SC)或静脉内(IV)施用所述组合物。

23.如以上方法权利要求中任一项所述的方法,其中肌内施用所述组合物。

24.如权利要求23所述的方法,其中在分开的注射部位施用所述肌内施用。

25.如权利要求24所述的方法,其中所述分开的注射部位在相对的三角肌中。

26.如权利要求24所述的方法,其中所述分开的注射部位在每一侧的臀肌或股直肌部位中。

27.如以上方法权利要求中任一项所述的方法,其中所述方法不包括施用免疫调节剂,或者所述方法在不存在免疫调节剂的情况下执行,任选地其中所述免疫调节剂是检查点抑制剂。

28.如以上方法权利要求中任一项所述的方法,其中所述方法还包括施用免疫调节剂。

29.如权利要求27或28所述的方法,其中所述免疫调节剂是抗CTLA4抗体或其抗原结合片段、抗PD-1抗体或其抗原结合片段、抗PD-L1抗体或其抗原结合片段、抗4-1BB抗体或其抗原结合片段、或抗OX-40抗体或其抗原结合片段。

30.如以上方法权利要求中任一项所述的方法,其还包括确定或已经确定所述受试者的HLA单倍型。

31.如以上方法权利要求1-30中任一项所述的方法,其中所述ChAdV载体包含:

(a)ChAdV骨架,其中所述ChAdV骨架包含:

(i)至少一个启动子核苷酸序列,和

(ii)至少一个聚腺苷酸化(聚(A))序列;和

(b)盒,其中所述盒包含:

(i)至少一个抗原编码核酸序列,其任选地包含:

a.表位编码核酸序列,其任选地包含至少一个使编码的表位序列不同于由野生型核酸序列编码的对应肽序列的改变,

b.任选的5'接头序列,和

c.任选的3'接头序列;并且

其中所述盒可操作地连接至所述至少一个启动子核苷酸序列和所述至少一个聚(A)序列。

32.如以上方法权利要求1-30中任一项所述的方法,其中所述ChAdV载体包含:

(a)ChAdV骨架,其中所述ChAdV骨架包含:

(i)修饰的ChAdV68序列,其至少包含SEQ ID NO:1中列出的序列的核苷酸2至36,518,其中所述核苷酸2至36,518缺少:(1)SEQ ID NO:1中示出的序列的核苷酸577至3403,其对应于E1缺失;(2)SEQ ID NO:1中示出的序列的核苷酸27,125至31,825,其对应于E3缺失;和(3)SEQ ID NO:1中示出的序列的核苷酸34,916至35,642,其对应于部分E4缺失;

(ii)CMV启动子核苷酸序列;和

(iii)SV40聚腺苷酸化(聚(A))序列;和

(b)盒,其中所述盒包含:

(i)至少一个抗原编码核酸序列,其包含:

a.表位编码核酸序列,其任选地包含至少一个使编码的表位序列不同于由野生型核酸序列编码的对应肽序列的改变,

b.任选的5'接头序列,和

c.任选的3'接头序列;并且

其中所述盒插入到所述E1缺失中,并且所述盒可操作地连接至所述CMV启动子核苷酸序列和所述SV40聚(A)序列。

33.如以上方法权利要求31中任一项所述的方法,其中所述表位编码核酸序列编码已知或疑似由细胞表面上的MHC I类和/或MHC II类呈递的表位,任选地其中所述细胞表面是肿瘤细胞表面或感染细胞表面,并且任选地其中所述细胞是所述受试者的细胞。

34.如以上方法权利要求31中任一项所述的方法,其中所述至少一个抗原编码酸序列编码能够被抗原加工成表位的多肽序列,任选地其中所述表位已知或疑似由细胞表面上的MHC I类呈递,任选地其中所述细胞的表面是肿瘤细胞表面或感染细胞表面。

35.如权利要求33或34所述的方法,其中所述细胞是选自由以下各项组成的组的肿瘤细胞:肺癌、黑色素瘤、乳腺癌、卵巢癌、前列腺癌、肾癌、胃癌、结肠癌、睾丸癌、头颈癌、胰腺癌、脑癌、B细胞淋巴瘤、急性骨髓性白血病、慢性骨髓性白血病、慢性淋巴细胞白血病、T细胞淋巴细胞白血病、非小细胞肺癌和小细胞肺癌,或者

其中所述细胞是选自由以下各项组成的组的感染细胞:病原体感染的细胞、病毒感染的细胞、细菌感染的细胞、真菌感染的细胞和寄生虫感染的细胞。

36.如权利要求35的方法,其中所述病毒感染细胞是HIV感染细胞。

37.如权利要求36所述的方法,其中所述表位编码核酸序列编码HIVGAG蛋白或表位。

38.如以上方法权利要求31或33-37中任一项所述的方法,其中所述ChAdV载体中的所述盒的每个元件的有序序列在式中描述,其从5'至3'包含

P_a-($L5_b$-N_c-$L3_d$)$_X$-($G5_e$-U_f)$_Y$-$G3_g$

其中P包含所述至少一个启动子序列,其可操作地连接至所述至少一个抗原编码核酸序列中的至少一个,其中a=1,

N包含所述表位编码核酸序列之一,其中c=1,

L5包含所述5'接头序列,其中b=0或1,

L3包含所述3'接头序列,其中d=0或1,

G5包含所述至少一个编码GPGPG氨基酸接头的核酸序列之一,其中e=0或1,

G3包含所述至少一个编码GPGPG氨基酸接头的核酸序列之一,其中g=0或1,

U包含所述至少一个MHC II类表位编码核酸序列之一,其中f=1,

X=1至400,其中对于每个X,对应的N_c是表位编码核酸序列,并且

Y=0、1或2,其中对于每个Y,对应的U_f是MHC II类表位编码核酸序列。

39.如权利要求38所述的方法,其中对于每个X,对应的N_c是不同的表位编码核酸序列。

40.如权利要求38或39所述的方法,其中对于每个Y,对应的U_f是不同的MHC II类表位编

码核酸序列。

41.如以上方法权利要求38-40中任一项所述的方法,其中

b=1,d=1,e=1,g=1,h=1,X=10,Y=2,

P是CMV启动子序列,

每个N编码长度为7-15个氨基酸的MHC I类表位、MHC II类表位、能够刺激B细胞响应的表位或其组合,

L5是编码所述表位的天然N末端氨基酸序列的天然5'接头序列,并且其中所述5'接头序列编码长度为至少3个氨基酸的肽,

L3是编码所述表位的天然C末端氨基酸序列的天然3'接头序列,并且其中所述3'接头序列编码长度为至少3个氨基酸的肽,

U是PADRE II类序列和破伤风类毒素MHC II类序列中的每一个,

所述ChAdV载体包含经修饰的ChAdV68序列,所述经修饰的ChAdV68序列至少包含SEQ ID NO:1中列出的序列的核苷酸2至36,518,其中所述核苷酸2至36,518缺少:(1)SEQ ID NO:1中示出的序列的核苷酸577至3403,其对应于E1缺失;(2)SEQ ID NO:1中示出的序列的核苷酸27,125至31,825,其对应于E3缺失;和(3)SEQ ID NO:1中示出的序列的核苷酸34,916至35,642,其对应于部分E4缺失,并且所述新抗原盒插入到所述E1缺失中,并且

所述I抗原编码核酸序列中的每一个编码长度为25个氨基酸的多肽。

42.如以上方法权利要求31或33-40中任一项所述的方法,其中所述盒整合在所述至少一个启动子核苷酸序列和所述至少一个聚(A)序列之间。

43.如以上方法权利要求31、33-40或42中任一项所述的方法,其中所述至少一个启动子核苷酸序列可操作地连接至所述盒。

44.如以上方法权利要求31、33-40或42-43中任一项所述的方法,其中所述ChAdV骨架包含ChAdV68载体骨架。

45.如权利要求44所述的方法,其中所述ChAdV68载体骨架包含SEQ ID NO:1中列出的序列。

46.如权利要求44或45所述的方法,其中所述ChAdV68载体骨架包含选自由以下各项组成的组的至少一个基因中的功能性缺失:参考ChAdV68基因组或参考SEQ ID NO:1中示出的序列的腺病毒E1A、E1B、E2A、E2B、E3、L1、L2、L3、L4和L5基因,任选地其中所述腺病毒骨架或修饰的ChAdV68序列在以下中完全缺失或功能性缺失:参考所述腺病毒基因组或参考SEQ ID NO:1中示出的序列的(1)E1A和E1B;或(2)E1A、E1B和E3,任选地其中所述E1基因通过参考SEQ ID NO:1中示出的序列的至少核苷酸577至3403的E1缺失而功能性缺失,并且任选地其中所述E3基因通过参考SEQ ID NO:1中示出的序列的至少核苷酸27,125至31,825的E3缺失而功能性缺失。

47.如权利要求44或45所述的方法,其中所述ChAdV68载体骨架包含参考ChAdV68基因组或参考SEQ ID NO:1中示出的序列的一个或多个基因或调控序列,任选地其中所述一个或多个基因或调控序列选自由以下各项组成的组:黑猩猩腺病毒反向末端重复序列(ITR)、E1A、E1B、E2A、E2B、E3、E4、L1、L2、L3、L4和L5基因。

48.如权利要求44-47中任一项所述的方法,其中所述ChAdV68载体骨架包含部分缺失的E4基因。

49.如权利要求48所述的方法,其中所述部分缺失的E4基因包含:

A.所述E4基因序列,其在SEQ ID NO:1中示出并至少缺少SEQ ID NO:1中示出的序列的核苷酸34,916至35,642,

B.所述E4基因序列,其在SEQ ID NO:1中示出并且至少缺少SEQ ID NO:1中示出的序列的核苷酸34,916至34,942、核苷酸34,952至35,305、SEQ ID NO:1中示出的序列的核苷酸35,302至35,642,并且其中所述载体至少包含SEQ ID NO:1中示出的序列的核苷酸2至36,518,

C.所述E4基因序列,其在SEQ ID NO:1中示出并且至少缺少SEQ ID NO:1中示出的序列的核苷酸34,980至36,516,并且其中所述载体至少包含SEQ ID NO:1中示出的序列的核苷酸2至36,518,

D.所述E4基因序列,其在SEQ ID NO:1中示出并且至少缺少SEQ ID NO:1中示出的序列的核苷酸34,979至35,642,并且其中所述载体至少包含SEQ ID NO:1中示出的序列的核苷酸2至36,518,

E.E4Orf2的至少部分缺失、完全缺失的E4Orf3、和E4Orf4的至少部分缺失的E4缺失,

F.E4Orf2的至少部分缺失、E4Orf3的至少部分缺失、和E4Orf4的至少部分缺失的E4缺失,

G.E4Orf1的至少部分缺失、完全缺失的E4Orf2、和E4Orf3的至少部分缺失的E4缺失,或

H.E4Orf2的至少部分缺失和E4Orf3的至少部分缺失的E4缺失。

50.如权利要求44所述的方法,其中所述ChAdV68载体骨架至少包含SEQ ID NO:1中列出的序列的核苷酸2至36,518,其中所述核苷酸2至36,518缺少:(1)SEQ ID NO:1中示出的序列的核苷酸577至3403,其对应于E1缺失;(2)SEQ ID NO:1中示出的序列的核苷酸27,125至31,825,其对应于E3缺失;和(3)SEQ ID NO:1中示出的序列的核苷酸34,916至35,642,其对应于部分E4缺失;任选地其中所述抗原盒插入到所述E1缺失中。

51.如权利要求44所述的方法,其中所述ChAdV68载体骨架包含SEQ ID NO:29369中列出的序列,任选地其中所述抗原盒插入到所述E1缺失中。

52.如权利要求44所述的方法,其中所述ChAdV68载体骨架至少包含SEQ ID NO:1中列出的序列的核苷酸2至36,518,其中所述核苷酸2至36,518缺少:

A.SEQ ID NO:1中示出的序列的核苷酸577至3403,其对应于E1缺失;

B.SEQ ID NO:1中示出的序列的核苷酸27,125至31,825,其对应于E3缺失;

C.SEQ ID NO:1中示出的序列的核苷酸34,916至35,642,其对应于部分E4缺失;

D.SEQ ID NO:1中示出的序列的核苷酸456至3014;

E.SEQ ID NO:1中示出的序列的核苷酸27,816至31,333;

F.SEQ ID NO:1中示出的序列的核苷酸3957至10346;

G.SEQ ID NO:1中示出的序列的核苷酸21787至23370;

H.SEQ ID NO:1中示出的序列的核苷酸33486至36193;或

其组合。

53.如权利要求44所述的方法,其中所述ChAdV68载体骨架至少包含SEQ ID NO:1中列出的序列的核苷酸2至36,518,其中所述核苷酸2至36,518缺少:(1)SEQ ID NO:1中示出的序列的核苷酸577至3403,其对应于E1缺失;和(2)SEQ ID NO:1中示出的序列的核苷酸27,

125至31,825,其对应于E3缺失。

54.如以上方法权利要求31、33-40或42-53中任一项所述的方法,其中所述其中所述盒插入所述ChAdV骨架的所述E1区、E3区和/或允许掺入所述盒的任何缺失的AdV区中。

55.如以上方法权利要求31、33-40或42-54中任一项所述的方法,其中所述ChAdV骨架由第一代、第二代或辅助依赖性腺病毒载体之一生成。

56.如以上方法权利要求31、33-40或42-55中任一项所述的方法,其中所述至少一个启动子核苷酸序列选自由以下各项组成的组:CMV、SV40、EF-1、RSV、PGK、HSA、MCK和EBV启动子序列。

57.如以上方法权利要求31、33-40或42-55中任一项所述的方法,其中所述至少一个启动子核苷酸序列是CMV启动子序列。

58.如以上方法权利要求中任一项所述的方法,其中所述表位编码核酸序列中的至少一个编码表位,所述表位在表达和翻译时能够由所述受试者的细胞上的MHC I类呈递。

59.如以上方法权利要求中任一项所述的方法,其中所述表位编码核酸序列中的至少一个编码表位,所述表位在表达和翻译时能够由所述受试者的细胞上的MHC II类呈递。

60.如以上方法权利要求31、33-40或42-59中任一项所述的方法,其中所述至少一个抗原编码核酸序列包含两个或更多个抗原编码核酸序列。

61.如权利要求60所述的方法,其中每个抗原编码核酸序列彼此直接连接。

62.如以上方法权利要求31-40或42-61中任一项所述的方法,其中每个抗原编码核酸序列通过编码接头的核酸序列连接至不同的抗原编码核酸序列。

63.如权利要求62所述的方法,其中所述接头连接两个表位编码核酸序列,或者将表位编码核酸序列连接至MHC II类表位编码核酸序列。

64.如权利要求63所述的方法,其中所述接头选自由以下各项组成的组:(1)长度为至少2、3、4、5、6、7、8、9或10个残基的连续甘氨酸残基;(2)长度为至少2、3、4、5、6、7、8、9或10个残基的连续丙氨酸残基;(3)两个精氨酸残基(RR);(4)丙氨酸、丙氨酸、酪氨酸(AAY);(5)被哺乳动物蛋白酶体高效加工的长度为至少2、3、4、5、6、7、8、9或10个氨基酸残基的共有序列;和(6)侧接来源于同源蛋白的抗原并且长度为至少2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20或2-20个氨基酸的一个或多个天然序列。

65.如权利要求62所述的方法,其中所述接头连接两个MHC II类表位编码核酸序列,或者将MHC II类序列连接至表位编码核酸序列。

66.如权利要求65所述的方法,其中所述接头包含序列GPGPG。

67.如以上方法权利要求31-40或42-66中任一项所述的方法,其中所述抗原编码核酸序列可操作地或直接连接至分开或连续的序列,所述分开或连续的序列增强所述抗原编码核酸序列的表达、稳定性、细胞运输、加工和呈递和/或免疫原性。

68.如权利要求67所述的方法,其中所述分开或连续的序列包含以下中的至少一种:泛素序列、经修饰以增加蛋白酶体靶向的泛素序列(例如,在位置76处含有Gly至Ala取代的所述泛素序列)、免疫球蛋白信号序列(例如,IgK)、主要组织相容性I类序列、溶酶体相关膜蛋白(LAMP)-1、人树突状细胞溶酶体相关膜蛋白和主要组织相容性II类序列;任选地,其中所述经修饰以增加蛋白酶体靶向的泛素序列是A76。

69.如以上方法权利要求中任一项所述的方法,其中所述表位编码核酸序列包含至少

一个改变,所述改变使得相对于翻译的对应野生型核酸序列,所述编码的表位对其对应MHC等位基因的结合亲和力增加。

70.如以上方法权利要求中任一项所述的方法,其中所述表位编码核酸序列包含至少一个改变,所述改变使得相对于翻译的对应野生型核酸序列,所述编码的表位对其对应MHC等位基因的结合稳定性增加。

71.如以上方法权利要求中任一项所述的方法,其中所述表位编码核酸序列包含至少一个改变,所述改变使得相对于翻译的对应野生型核酸序列,所述编码的表位在其对应MHC等位基因上呈递的可能性增加。

72.如以上方法权利要求中任一项所述的方法,其中所述至少一个改变包括点突变、移码突变、非移码突变、缺失突变、插入突变、剪接变异、基因组重排或蛋白酶体生成的剪接抗原。

73.如以上方法权利要求中任一项所述的方法,其中所述表位编码核酸序列编码已知或疑似在已知或疑似患有癌症的所述受试者中表达的表位。

74.如权利要求73所述的方法,其中所述癌症包括实体瘤。

75.如权利要求73或74所述的方法,其中所述癌症选自由以下各项组成的组:肺癌、黑色素瘤、乳腺癌、卵巢癌、前列腺癌、肾癌、胃癌、结肠癌、睾丸癌、头颈癌、胰腺癌、膀胱癌、脑癌、B细胞淋巴瘤、急性骨髓性白血病、成人急性成淋巴细胞性白血病、慢性骨髓性白血病、慢性淋巴细胞白血病、T细胞淋巴细胞白血病、非小细胞肺癌和小细胞肺癌。

76.如以上方法权利要求31-40或42-75中任一项所述的方法,其中所述至少一个抗原编码核酸序列包含至少2-10、2、3、4、5、6、7、8、9或10个抗原编码核酸序列,任选地其中每个抗原编码核酸序列编码不同的抗原编码核酸序列。

77.如以上方法权利要求31-40或42-75中任一项所述的方法,其中所述至少一个抗原编码核酸序列包含至少11-20、15-20、11-100、11-200、11-300、11-400、11、12、13、14、15、16、17、18、19、20或至多400个抗原编码核酸序列,任选地其中每个抗原编码核酸序列编码不同的抗原编码核酸序列。

78.如以上方法权利要求31-40或42-75中任一项所述的方法,其中所述至少一个抗原编码核酸序列包含至少11-20、15-20、11-100、11-200、11-300、11-400、11、12、13、14、15、16、17、18、19、20或至多400个抗原编码核酸序列。

79.如以上方法权利要求31-40或42-75中任一项所述的方法,其中所述至少一个抗原编码核酸序列包含至少2-400个抗原编码核酸序列,并且其中所述抗原编码核酸序列中的至少两个编码由细胞表面上的MHC I类呈递的表位序列或其部分。

80.如权利要求79所述的方法,其中所述MHC I类表位中的至少两个由肿瘤细胞表面上的MHC I类呈递。

81.如以上方法权利要求31-40或42-80中任一项所述的方法,其中所述表位编码核酸序列包含至少一个MHC I类表位编码核酸序列,并且其中每个抗原编码核酸序列编码长度在8至35个氨基酸之间的多肽序列,任选地长度为9-17、9-25、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34或35个氨基酸。

82.如以上方法权利要求31-40或42-81中任一项所述的方法,其中存在所述至少一个MHC II类表位编码核酸序列。

83.如以上方法权利要求31-40或42-81中任一项所述的方法,其中存在所述至少一个MHC II类表位编码核酸序列,并且其包含含至少一个改变的至少一个MHC II类表位编码核酸序列,所述改变使所述编码的表位序列不同于由野生型核酸序列编码的对应肽序列。

84.如以上方法权利要求31-40或42-83中任一项所述的方法,其中所述表位编码核酸序列包含MHC II类表位编码核酸序列,并且其中每个抗原编码核酸序列编码长度为12-20、12、13、14、15、16、17、18、19、20或20-40个氨基酸的多肽序列。

85.如以上方法权利要求31-40或42-84中任一项所述的方法,其中所述表位编码核酸序列包含MHC II类表位编码核酸序列,其中存在所述至少一个MHC II类表位编码核酸序列,并且其中所述至少一个MHC II类表位编码核酸序列包含至少一个通用MHC II类表位编码核酸序列,任选地其中所述至少一个通用序列包含破伤风类毒素和PADRE中的至少一种。

86.如以上方法权利要求31、33-40或42-85中任一项所述的方法,其中所述至少一个启动子核苷酸序列是诱导型的。

87.如以上方法权利要求31、33-40或42-85中任一项所述的方法,其中所述至少一个启动子核苷酸序列是非诱导型的。

88.如以上方法权利要求31、33-40或42-87中任一项所述的方法,其中所述至少一个聚(A)序列包含牛生长激素(BGH)SV40聚A序列。

89.如权利要求31、33-40或42-88中任一项所述的方法,其中所述至少一个聚(A)序列是至少20个、至少30个、至少40个、至少50个、至少60个、至少70个、至少80个或至少90个连续的A核苷酸。

90.如以上方法权利要求31、33-40或42-88中任一项所述的方法,其中所述至少一个聚(A)序列是至少100个连续A核苷酸。

91.如以上方法权利要求中任一项所述的方法,其中所述盒还包含以下中的至少一种:内含子序列、土拨鼠肝炎病毒转录后调控元件(WPRE)序列、内部核糖体进入序列(IRES)序列、编码2A自切割肽序列的核苷酸序列、编码弗林蛋白酶切割位点的核苷酸序列、或5'或3'非编码区中的已知增强可操作地连接至所述至少一个抗原编码核酸序列中的至少一个的mRNA的核输出、稳定性或翻译效率的序列。

92.如以上方法权利要求中任一项所述的方法,其中所述盒还包含报告基因,包括但不限于绿色荧光蛋白(GFP)、GFP变体、分泌型碱性磷酸酶、荧光素酶、荧光素酶变体或可检测肽或表位。

93.如权利要求92所述的方法,其中所述可检测肽或表位选自由HA标签、Flag标签、His标签或V5标签组成的组。

94.如以上方法权利要求中任一项所述的方法,其中所述一种或多种载体还包含一种或多种编码至少一种免疫调节剂的核酸序列。

95.如权利要求94所述的方法,其中所述免疫调节剂是抗CTLA4抗体或其抗原结合片段、抗PD-1抗体或其抗原结合片段、抗PD-L1抗体或其抗原结合片段、抗4-1BB抗体或其抗原结合片段、或抗OX-40抗体或其抗原结合片段。

96.如权利要求95所述的方法,其中所述抗体或其抗原结合片段是Fab片段、Fab'片段、单链Fv(scFv)、单特异性的或多重特异性的连接在一起的单结构域抗体(sdAb)(例如,骆驼科抗体结构域)或全长单链抗体(例如,具有通过柔性接头连接的重链和轻链的全长IgG)。

97.如权利要求95或96所述的方法，其中所述抗体的重链和轻链序列是被诸如2A或IRES的自切割序列分开的连续序列；或者所述抗体的重链和轻链序列通过诸如连续的甘氨酸残基的柔性接头连接。

98.如权利要求94所述的方法，其中所述免疫调节剂是细胞因子。

99.如权利要求98所述的方法，其中所述细胞因子是IL-2、IL-7、IL-12、IL-15或IL-21中的至少一种或其各自的变体。

100.如以上方法权利要求31-40或42-99中任一项所述的方法，其中至少一个表位编码核酸序列通过执行以下步骤来选择：

(a)从肿瘤、感染细胞或感染疾病生物体中获得外显子组、转录组或全基因组核苷酸测序数据中的至少一种，其中所述核苷酸测序数据用于获得代表一组抗原中的每一个的肽序列的数据；

(b)将所述每个抗原的肽序列输入呈递模型以生成每个所述抗原由细胞表面、任选肿瘤细胞表面或感染细胞表面上的所述MHC等位基因中的一个或多个呈递的一组数值可能性，已经至少基于接收到的质谱数据来鉴定的所述数值可能性组；和

(c)基于所述数值可能性组来选择所述抗原组的子集以生成一组选定的抗原，所述选定的抗原用于生成所述至少一个表位编码核酸序列。

101.如权利要求41所述的方法，其中所述表位编码核酸序列中的每一个通过执行以下步骤来选择：

(a)从肿瘤、感染细胞或感染疾病生物体中获得外显子组、转录组或全基因组核苷酸测序数据中的至少一种，其中所述核苷酸测序数据用于获得代表一组抗原中的每一个的肽序列的数据；

(b)将所述每个抗原的肽序列输入呈递模型以生成每个所述抗原由细胞表面、任选肿瘤细胞表面或感染细胞表面上的所述MHC等位基因中的一个或多个呈递的一组数值可能性，已经至少基于接收到的质谱数据来鉴定的所述数值可能性组；和

(c)基于所述数值可能性组来选择所述抗原组的子集以生成一组选定的抗原，所述选定的抗原用于生成所述至少20个表位编码核酸序列。

102.如权利要求100所述的方法，其中所述选定的表位组的数量为2-20。

103.如权利要求100-102中任一项所述的方法，其中所述呈递模型代表以下之间的依赖性：

(a)在肽序列的特定位置存在一个特定MHC等位基因和特定氨基酸的对；和

(b)此类在所述特定位置处包含所述特定氨基酸的肽序列由所述对的特定一个MHC等位基因呈递在细胞表面，任选的肿瘤细胞表面或感染细胞表面上的可能性。

104.如权利要求100-103中任一项所述的方法，其中选择所述选定的抗原组包括基于所述呈递模型来选择相对于未选定的抗原在所述细胞表面上呈递的可能性增加的抗原，任选地其中所述选定的抗原已被验证为由一个或多个特定HLA等位基因呈递。

105.如权利要求100-104中任一项所述的方法，其中选择所述选定的表位组包括基于所述呈递模型来选择相对于未选定的表位能够在所述受试者中诱导肿瘤特异性免疫响应的可能性增加的表位。

106.如权利要求100-105中任一项所述的方法，其中选择所述选定的抗原组包括基于

所述呈递模型来选择相对于未选定的抗原能够在所述受试者中诱导肿瘤特异性或感染性疾病特异性免疫响应的可能性增加的抗原。

107.如权利要求100-106中任一项所述的方法，其中选择所述选定的抗原组包括基于所述呈递模型来选择相对于未选定的抗原能够被专职性抗原呈递细胞(APC)呈递至初始T细胞的可能性增加的抗原，任选地其中所述APC是树突状细胞(DC)。

108.如权利要求100-107中任一项所述的方法，其中选择所述选定的抗原组包括基于所述呈递模型来选择相对于未选定的抗原，经由中枢或外周耐受而受到抑制的可能性降低的抗原。

109.如权利要求100-108中任一项所述的方法，其中选择所述选定的抗原组包括基于所述呈递模型来选择相对于未选定的抗原能够在所述受试者中诱导对正常组织的自身免疫响应的可能性降低的抗原。

110.如权利要求100-109中任一项所述的方法，其中外显子组或转录组核苷酸测序数据通过对肿瘤细胞或组织、感染细胞或感染性疾病生物体执行测序而获得。

111.如权利要求110所述的方法，其中所述测序是下一代测序(NGS)或任何大规模并行测序方法。

112.如以上权利要求中任一项所述的方法，其中所述盒包含由所述盒中的相邻序列形成的连接表位序列。

113.如权利要求112所述的方法，其中至少一个或每个连接表位序列对MHC具有大于500nM的亲和力。

114.如权利要求112或113所述的方法，其中每个连接表位序列是非自体的。

115.如以上权利要求中任一项所述的方法，其中所述MHC I类表位中的每一个被预测或验证为能够由存在于至少5％的群体中的至少一个HLA等位基因呈递。

116.如以上权利要求中任一项所述的方法，其中所述MHC I类表位中的每一个被预测或验证为能够由至少一个HLA等位基因呈递，其中每个抗原/HLA对在群体中具有至少0.01％的抗原/HLA患病率。

117.如以上权利要求中任一项所述的方法，其中所述MHC I类表位中的每一个被预测或验证为能够由至少一个HLA等位基因呈递，其中每个抗原/HLA对在群体中具有至少0.1％的抗原/HLA患病率。

118.如以上权利要求中任一项所述的方法，其中所述盒不编码包含翻译的野生型核酸序列的非治疗性MHC I类或II类表位核酸序列，其中所述非治疗性表位被预测为在所述受试者的MHC等位基因上展示。

119.如权利要求118所述的方法，其中所述非治疗性预测的MHC I类或II类表位序列是由所述盒中的相邻序列形成的连接表位序列。

120.如权利要求112-119所述的方法，其中所述预测基于通过将所述非治疗性表位的序列输入呈递模型而生成的呈递可能性。

121.如以上方法权利要求112-120中任一项所述的方法，其中所述盒中的所述抗原编码码核酸序列的顺序通过包括以下各项的一系列步骤来确定：

(a)生成对应于所述抗原编码核酸序列的不同顺序的一组候选盒序列；

(b)对于每个候选盒序列，基于所述候选盒序列中非治疗性表位的呈递来确定呈递评

分;和

(c)选择与低于预定阈值的呈递评分相关的候选盒序列作为用于抗原疫苗的所述盒序列。

122.如以上权利要求中任一项所述的方法,其中所述组合物被配制为包含药学上可接受的载剂的药物组合物。

123.如以上权利要求中任一项所述的方法,其中所述组合物包含含所述ChAdV载体的病毒颗粒。

124.如以上权利要求中任一项所述的方法,其中所述表位编码核酸序列中的一个或多个来源于所述受试者的肿瘤。

125.如以上权利要求中任一项所述的方法,其中所述表位编码核酸序列中的每一个来源于所述受试者的肿瘤。

126.如以上权利要求中任一项所述的方法,其中所述表位编码核酸序列中的一个或多个不来源于所述受试者的肿瘤。

127.如以上权利要求中任一项所述的方法,其中所述表位编码核酸序列中的每一个不来源于所述受试者的肿瘤。

128.如以上权利要求中任一项所述的方法,其中所述表位编码核酸序列包含选自由SEQ ID NO:57-29,364组成的组的表位。

129.如以上权利要求中任一项所述的方法,其中所述至少一个抗原编码核酸序列包含以下中的至少每一个:

(A)KRAS_G12C MHC I类表位编码核酸序列,其中所述KRAS_G12C MHC I类表位编码核酸序列编码选自由SEQ ID NO:14,954;19,848;和19,850组成的组的MHC I类表位,

(B)KRAS_G12D MHC I类表位编码核酸序列,其中所述KRAS_G12D MHC I类表位编码核酸序列编码选自由SEQ ID NO:19,749;19,865;和19,863组成的组的MHC I类表位,和

(C)KRAS_G12V MHC I类表位编码核酸序列,其中所述KRAS_G12V MHC I类表位编码核酸序列编码选自由SEQ ID NO:19,976;19,779;11,495;和19,974组成的组的MHC I类表位。

130.如以上权利要求中任一项所述的方法,其中所述至少一个抗原编码核酸序列包含:

(A)KRAS_G12C MHC I类表位编码核酸序列,

(B)KRAS_G12D MHC I类表位编码核酸序列,

(C)KRAS_G12V MHC I类表位编码核酸序列,

(D)KRAS Q61H MHC I类表位编码核酸序列,

(E)TP53_R213L MHC I类表位编码核酸序列,

(F)TP53_S127Y MHC I类表位编码核酸序列,

(G)TP53_R249M MHC I类表位编码核酸序列,

或其组合。

131.如以上方法权利要求中任一项所述的方法,其中所述方法还包括施用基于甲病毒的自扩增表达系统。

132.如权利要求131所述的方法,其中肌内(IM)、皮内(ID)、皮下(SC)或静脉内(IV)施用用于递送所述基于甲病毒的自扩增表达系统的所述组合物。

133.如权利要求131或132所述的方法,其中肌内施用用于递送所述基于甲病毒的自扩增表达系统的所述组合物。

134.如权利要求133所述的方法,其中在分开的注射部位施用所述肌内施用。

135.如权利要求134所述的方法,其中所述分开的注射部位在相对的三角肌中。

136.如权利要求134所述的方法,其中所述分开的注射部位在每一侧的臀肌或股直肌部位中。

137.如权利要求133-136中任一项所述的方法,其中所述一个或多个加强剂量的注射部位尽可能靠近所述初免剂量的注射部位。

138.如以上方法权利要求中任一项所述的方法,其中用于递送所述基于甲病毒的自扩增表达系统的所述组合物包含:

(A)所述基于甲病毒的自扩增表达系统,其中所述基于甲病毒的自扩增表达系统包含一种或多种载体,其中所述一种或多种载体包含:

(a)RNA甲病毒骨架,其中所述RNA甲病毒骨架包含:

(i)至少一个启动子核苷酸序列,和

(ii)至少一个聚腺苷酸化(聚(A))序列;和

(b)盒,其中所述盒包含:

(i)至少一个抗原编码核酸序列,其包含:

a.表位编码核酸序列,其任选地包含至少一个使编码的表位序列不同于由野生型核酸序列编码的对应肽序列的改变,

b.任选的5'接头序列,和

c.任选的3'接头序列;

(ii)任选地,可操作地连接至所述至少一个抗原编码核酸序列的第二启动子核苷酸序列;和

(iii)任选地,至少一个第二聚(A)序列,其中所述第二聚(A)序列是所述甲病毒的天然聚(A)序列或外源聚(A)序列,和

(B)脂质纳米颗粒(LNP),其中所述LNP封装所述基于甲病毒的自扩增表达系统。

139.如以上方法权利要求中任一项所述的方法,其中用于递送所述基于甲病毒的自扩增表达系统的所述组合物包含,

(A)所述基于甲病毒的自扩增表达系统,其中所述基于甲病毒的自扩增表达系统包含一种或多种载体,其中所述一种或多种载体包含:

(a)RNA甲病毒骨架,其中所述RNA甲病毒骨架包含SEQ ID NO:6中列出的核酸序列,其中所述RNA甲病毒骨架序列包含26S启动子核苷酸序列和聚(A)序列,其中所述26S启动子序列对于所述RNA甲病毒骨架是内源的,并且其中所述聚(A)序列对于所述RNA甲病毒骨架是内源的;和

(b)整合在所述26S启动子核苷酸序列和所述聚(A)序列之间的盒,其中所述盒可操作地连接至所述26S启动子核苷酸序列,并且其中所述盒包含至少一个抗原编码核酸序列,所述抗原编码核酸序列包含:

a.表位编码核酸序列,其任选地包含至少一个使编码的表位序列不同于由野生型核酸序列编码的对应肽序列的改变,

b.任选的5'接头序列,和

c.任选的3'接头序列;和

(B)脂质纳米颗粒(LNP),其中所述LNP封装所述基于甲病毒的自扩增表达系统。

140.如以上权利要求中任一项所述的方法,其中所述ChAdV载体的盒与用于递送所述基于甲病毒的自扩增表达系统的所述组合物的盒相同。

141.如以上方法权利要求中任一项所述的方法,其中用于递送所述基于甲病毒的自扩增表达系统的所述组合物中的所述盒的每个元件的有序序列在式中描述,其从5'至3'包含

P_a-($L5_b$-N_c-$L3_d$)$_X$-($G5_e$-U_f)$_Y$-$G3_g$

其中P包含所述第二启动子核苷酸序列,其中a=0或1,

N包含所述表位编码核酸序列之一,其中c=1,

L5包含所述5'接头序列,其中b=0或1,

L3包含所述3'接头序列,其中d=0或1,

G5包含所述至少一个编码GPGPG氨基酸接头的核酸序列之一,其中e=0或1,

G3包含所述至少一个编码GPGPG氨基酸接头的核酸序列之一,其中g=0或1,

U包含所述至少一个MHC II类表位编码核酸序列之一,其中f=1,

X=1至400,其中对于每个X,对应的N_c是表位编码核酸序列,并且

Y=0、1或2,其中对于每个Y,对应的U_f是MHC II类表位编码核酸序列。

142.如权利要求141所述的方法,其中对于每个X,对应的N_c是不同的表位编码核酸序列。

143.如权利要求141或142所述的方法,其中对于每个Y,对应的U_f是不同的MHC II类表位编码核酸序列。

144.如以上方法权利要求141-143中任一项所述的方法,其中

a=0,b=1,d=1,e=1,g=1,h=1,X=20,Y=2,

所述至少一个启动子核苷酸序列是由所述RNA甲病毒骨架提供的单个26S启动子核苷酸序列,

所述至少一个聚腺苷酸化聚(A)序列是由所述RNA甲病毒骨架提供的至少100个连续A核苷酸的聚(A)序列,

所述盒整合在所述26S启动子核苷酸序列和所述聚(A)序列之间,其中所述盒可操作地连接至所述26S启动子核苷酸序列和所述聚(A)序列,

每个N编码长度为7-15个氨基酸的MHC I类表位、MHC II类表位、能够刺激B细胞响应的表位或其组合,

L5是编码所述表位的天然N末端氨基酸序列的天然5'接头序列,并且其中所述5'接头序列编码长度为至少3个氨基酸的肽,

L3是编码所述表位的天然C末端氨基酸序列的天然3'接头序列,并且其中所述3'接头序列编码长度为至少3个氨基酸的肽,

U是PADRE II类序列和破伤风类毒素MHC II类序列中的每一个,

所述RNA甲病毒骨架是SEQ ID NO:6中列出的序列,并且

所述MHC I类表位编码核酸序列中的每一个编码长度在13和25个氨基酸之间的多肽。

145.如以上权利要求中任一项所述的方法,其中所述LNP包含选自由以下各项组成的

组的脂质:可电离的氨基脂质、磷脂酰胆碱、胆固醇、基于PEG的外壳脂质或其组合。

146.如以上权利要求中任一项所述的方法,其中所述LNP包含可电离的氨基脂质、磷脂酰胆碱、胆固醇和基于PEG的外壳脂质。

147.如权利要求145或146所述的方法,其中所述可电离的氨基脂质包含MC3样(二亚油醇基甲基-4-二甲基氨基丁酸酯)分子。

148.如以上权利要求中任一项所述的方法,其中所述LNP封装的表达系统具有约100nm的直径。

149.如以上方法权利要求138、140-143或145-148中任一项所述的方法,其中所述盒整合在所述至少一个启动子核苷酸序列和所述至少一个聚(A)序列之间。

150.如以上方法权利要求138、140-143或145-149中任一项所述的方法,其中所述至少一个启动子核苷酸序列可操作地连接至所述盒。

151.如以上方法权利要求138、140-143或145-150中任一项所述的方法,其中所述一种或多种载体包括一种或多种+链RNA载体。

152.如权利要求151所述的方法,其中所述一个或多个+链RNA载体包含5'7-甲基鸟苷(m7g)帽。

153.如权利要求151或152所述的方法,其中所述一个或多个+链RNA载体通过体外转录产生。

154.如以上方法权利要求138、140-143或145-153中任一项所述的方法,其中所述一个或多个载体在哺乳动物细胞内自复制。

155.如以上方法权利要求138、140-143或145-154中任一项所述的方法,其中所述RNA甲病毒骨架包含Aura病毒、摩根堡病毒、委内瑞拉马脑炎病毒、罗斯河病毒、森林脑炎病毒、辛德比斯病毒或马亚罗病毒的至少一个核苷酸序列。

156.如以上方法权利要求138、140-143或145-154中任一项所述的方法,其中所述RNA甲病毒骨架包含委内瑞拉马脑炎病毒的至少一个核苷酸序列。

157.如权利要求155或156所述的方法,其中所述RNA甲病毒骨架至少包含由所述Aura病毒、所述摩根堡病毒、所述委内瑞拉马脑炎病毒、所述罗斯河病毒、所述森林脑炎病毒、所述辛德比斯病毒或所述马亚罗病毒的核苷酸序列编码的用于非结构蛋白介导的扩增的序列、26S启动子序列、聚(A)序列、非结构蛋白1(nsP1)基因、nsP2基因、nsP3基因和nsP4基因。

158.如权利要求155或156所述的方法,其中所述RNA甲病毒骨架至少包含由所述Aura病毒、所述摩根堡病毒、所述委内瑞拉马脑炎病毒、所述罗斯河病毒、所述森林脑炎病毒、所述辛德比斯病毒或所述马亚罗病毒的核苷酸序列编码的用于非结构蛋白介导的扩增的序列、26S启动子序列和聚(A)序列。

159.如权利要求157或158所述的方法,其中用于非结构蛋白介导的扩增的序列选自由以下各项组成的组:假病毒5'UTR、51-nt CSE、24-nt CSE、26S亚基因组启动子序列、19-nt CSE、假病毒3'UTR或其组合。

160.如以上方法权利要求157-159中任一项所述的方法,其中所述RNA甲病毒骨架不编码结构病毒粒子蛋白衣壳E2和E1。

161.如权利要求160所述的方法,其中插入所述盒以替代所述Aura病毒、所述摩根堡病毒、所述委内瑞拉马脑炎病毒、所述罗斯河病毒、所述森林脑炎病毒、所述辛德比斯病毒或

所述马亚罗病毒的核苷酸序列内的结构病毒粒子蛋白。

162.如权利要求155或156所述的方法,其中所述委内瑞拉马脑炎病毒包含SEQ ID NO:3或SEQ ID NO:5的序列。

163.如权利要求155或156所述的方法,其中所述委内瑞拉马脑炎病毒包含有还包含碱基对7544和11175之间的缺失的SEQ ID NO:3或SEQ ID NO:5的序列。

164.如权利要求163所述的方法,其中所述RNA甲病毒骨架包含SEQ ID NO:6或SEQ ID NO:7中列出的序列。

165.如权利要求163或164所述的方法,其中所述盒插入位置7544以替代如SEQ ID NO:3或SEQ ID NO:5的序列中列出的碱基对7544和11175之间的所述缺失。

166.如权利要求161-165的方法,其中所述盒的插入提供包含所述nsP1-4基因和所述至少一个核酸序列的多顺反子RNA的转录,其中所述nsP1-4基因和所述至少一个核酸序列在单独的开放阅读框中。

167.如以上方法权利要求138、140-143或145-166中任一项所述的方法,其中所述至少一个启动子核苷酸序列是由所述RNA甲病毒骨架编码的所述天然26S启动子核苷酸序列。

168.如以上方法权利要求138、140-143或145-166中任一项所述的方法,其中所述至少一个启动子核苷酸序列是外源RNA启动子。

169.如以上方法权利要求138、140-143或145-168中任一项所述的方法,其中所述第二启动子核苷酸序列是26S启动子核苷酸序列。

170.如以上方法权利要求138、140-143或145-168中任一项所述的方法,其中所述第二启动子核苷酸序列包含多个26S启动子核苷酸序列,其中每个26S启动子核苷酸序列提供所述单独的开放阅读框中的一个或多个的转录。

171.如以上方法权利要求中任一项所述的方法,其中所述一个或多个载体各自的大小为至少300nt。

172.如以上方法权利要求中任一项所述的方法,其中所述一个或多个载体各自的大小为至少1kb。

173.如以上方法权利要求中任一项所述的方法,其中所述一个或多个载体各自的大小为2kb。

174.如以上方法权利要求中任一项所述的方法,其中所述一个或多个载体各自的大小小于5kb。

175.如以上方法权利要求138-143或145-174中任一项所述的方法,其中所述至少一个抗原编码核酸序列包含两个或更多个抗原编码核酸序列。

176.如权利要求175所述的方法,其中每个抗原编码核酸序列彼此直接连接。

177.如以上方法权利要求138-143或145-176中任一项所述的方法,其中每个抗原编码核酸序列通过编码接头的核酸序列连接至不同的抗原编码核酸序列。

178.如权利要求177所述的方法,其中所述接头连接两个MHC I类表位编码核酸序列,或者将MHC I类表位编码核酸序列连接至MHC II类表位编码核酸序列。

179.如权利要求178所述的方法,其中所述接头选自由以下各项组成的组:(1)长度为至少2、3、4、5、6、7、8、9或10个残基的连续甘氨酸残基;(2)长度为至少2、3、4、5、6、7、8、9或10个残基的连续丙氨酸残基;(3)两个精氨酸残基(RR);(4)丙氨酸、丙氨酸、酪氨酸(AAY);

(5)被哺乳动物蛋白酶体高效加工的长度为至少2、3、4、5、6、7、8、9或10个氨基酸残基的共有序列;和(6)侧接来源于同源蛋白的抗原并且长度为至少2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20或2-20个氨基酸的一个或多个天然序列。

180.如权利要求177所述的方法,其中所述接头连接两个MHC II类表位编码核酸序列,或者将MHC II类序列连接至MHC I类表位编码核酸序列。

181.如权利要求180所述的方法,其中所述接头包含序列GPGPG。

182.如以上方法权利要求138-143或145-181中任一项所述的方法,其中所述抗原编码核酸序列可操作地或直接连接至分开或连续的序列,所述分开或连续的序列增强所述抗原编码核酸序列的表达、稳定性、细胞运输、加工和呈递和/或免疫原性。

183.如权利要求182所述的方法,其中所述分开或连续的序列包含以下中的至少一种:泛素序列、经修饰以增加蛋白酶体靶向的泛素序列(例如,在位置76处含有Gly至Ala取代的所述泛素序列)、免疫球蛋白信号序列(例如,IgK)、主要组织相容性I类序列、溶酶体相关膜蛋白(LAMP)-1、人树突状细胞溶酶体相关膜蛋白和主要组织相容性II类序列;任选地,其中所述经修饰以增加蛋白酶体靶向的泛素序列是A76。

184.如以上方法权利要求138-143或145-183中任一项所述的方法,其中所述至少一个抗原编码核酸序列包含至少2-10、2、3、4、5、6、7、8、9或10个抗原编码核酸序列,任选地其中每个抗原编码核酸序列编码不同的抗原编码核酸序列。

185.如以上方法权利要求138-143或145-183中任一项所述的方法,其中所述至少一个抗原编码核酸序列包含至少11-20、15-20、11-100、11-200、11-300、11-400、11、12、13、14、15、16、17、18、19、20或至多400个抗原编码核酸序列,任选地其中每个抗原编码核酸序列编码不同的抗原编码核酸序列。

186.如权利要求138-143或145-183中任一项所述的组合物,其中所述至少一个抗原编码核酸序列包含至少11-20、15-20、11-100、11-200、11-300、11-400、11、12、13、14、15、16、17、18、19、20或至多400个抗原编码核酸序列。

187.如以上方法权利要求138-143或145-183中任一项所述的方法,其中所述至少一个抗原编码核酸序列包含至少2-400个抗原编码核酸序列,并且其中所述抗原编码核酸序列中的至少两个编码由细胞表面上的MHC I类呈递的表位序列或其部分。

188.如权利要求144所述的组合物,其中所述MHC I类表位中的至少两个由所述肿瘤细胞表面上的MHC I类呈递。

189.如以上方法权利要求138-143或145-188中任一项所述的方法,其中所述表位编码核酸序列包含至少一个MHC I类表位编码核酸序列,并且其中每个抗原编码核酸序列编码长度在8至35个氨基酸之间的多肽序列,任选地长度为9-17、9-25、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34或35个氨基酸。

190.如以上方法权利要求138-143或145-189中任一项所述的方法,其中存在所述至少一个MHC II类表位编码核酸序列。

191.如以上方法权利要求141-143或145-189中任一项所述的方法,其中存在所述至少一个MHC II类表位编码核酸序列,并且其包含含至少一个改变的至少一个MHC II类表位编码核酸序列,所述改变使所述编码的表位序列不同于由野生型核酸序列编码的对应肽序列。

192.如以上方法权利要求138-143或145-191中任一项所述的方法，其中所述表位编码核酸序列包含MHC II类表位编码核酸序列，并且其中每个抗原编码核酸序列编码长度为12-20、12、13、14、15、16、17、18、19、20或20-40个氨基酸的多肽序列。

193.如以上方法权利要求138-143或145-192中任一项所述的方法，其中所述表位编码核酸序列包含MHC II类表位编码核酸序列，其中存在所述至少一个MHC II类表位编码核酸序列，并且其中所述至少一个MHC II类表位编码核酸序列包含至少一个通用MHC II类表位编码核酸序列，任选地其中所述至少一个通用序列包含破伤风类毒素和PADRE中的至少一种。

194.如以上方法权利要求138、140-143或145-193中任一项所述的方法，其中所述至少一个启动子核苷酸序列或所述第二启动子核苷酸序列是诱导型的。

195.如以上方法权利要求138、140-143或145-193中任一项所述的方法，其中所述至少一个启动子核苷酸序列或所述第二启动子核苷酸序列是非诱导型的。

196.如以上方法权利要求138、140-143或145-195中任一项所述的方法，其中所述至少一个聚(A)序列包含所述甲病毒天然的聚(A)序列。

197.如以上方法权利要求138、140-143或145-195中任一项所述的方法，其中所述至少一个聚(A)序列包含所述甲病毒外源的聚(A)序列。

198.如以上方法权利要求138、140-143或145-197中任一项所述的方法，其中所述至少一个聚(A)序列可操作地连接至所述至少一个核酸序列中的至少一个。

199.如权利要求138、140-143或145-198中任一项所述的方法，其中所述至少一个聚(A)序列是至少20个、至少30个、至少40个、至少50个、至少60个、至少70个、至少80个或至少90个连续的A核苷酸。

200.如以上方法权利要求138、140-143或145-198中任一项所述的方法，其中所述至少一个聚(A)序列是至少100个连续A核苷酸。

201.如以上方法权利要求138-143或145-200中任一项所述的方法，其中所述表位编码核酸序列包含MHC I类表位编码核酸序列，并且其中所述MHC I类表位编码核酸序列通过执行以下步骤来选择：

(a)从所述肿瘤中获得外显子组、转录组或全基因组肿瘤核苷酸测序数据中的至少一种，其中所述肿瘤核苷酸测序数据用于获得代表一组表位中的每一个的肽序列的数据；

(b)将所述每个表位的肽序列输入呈递模型以生成所述表位中的每一个由所述肿瘤的肿瘤细胞表面上的所述MHC等位基因中的一个或多个呈递的一组数值可能性，已经至少基于接收到的质谱数据来鉴定的所述数值可能性组；和

(c)基于所述数值可能性组来选择所述表位组的子集以生成一组选定的表位，所述选定的表位用于生成所述MHC I类表位编码核酸序列。

202.如权利要求144所述的方法，其中所述MHC I类表位中的每一个编码核酸序列通过执行以下步骤来选择：

(a)从所述肿瘤中获得外显子组、转录组或全基因组肿瘤核苷酸测序数据中的至少一种，其中所述肿瘤核苷酸测序数据用于获得代表一组表位中的每一个的肽序列的数据；

(b)将所述每个表位的肽序列输入呈递模型以生成所述表位中的每一个由所述肿瘤的肿瘤细胞表面上的所述MHC等位基因中的一个或多个呈递的一组数值可能性，已经至少基

于接收到的质谱数据来鉴定的所述数值可能性组;和

(c)基于所述数值可能性组来选择所述表位组的子集以生成一组选定的表位,所述选定的表位用于生成所述至少20个MHC I类表位编码核酸序列。

203.如权利要求201所述的方法,其中所述选定的表位组的数量为2-20。

204.如权利要求201-203所述的方法,其中所述呈递模型代表以下之间的依赖性:

(a)在肽序列的特定位置存在一个特定MHC等位基因和特定氨基酸的对;和

(b)此类在所述特定位置处包含所述特定氨基酸的肽序列由所述对的特定一个MHC等位基因呈递在肿瘤细胞表面上的可能性。

205.如权利要求201-204所述的方法,其中选择所述选定的表位组包括基于所述呈递模型来选择相对于未选定的表位在所述肿瘤表面上呈递的可能性增加的表位。

206.如权利要求201-205所述的方法,其中选择所述选定的表位组包括基于所述呈递模型来选择相对于未选定的表位能够在所述受试者中诱导肿瘤特异性免疫响应的可能性增加的表位。

207.如权利要求201-206所述的方法,其中选择所述选定的表位组包括基于所述呈递模型来选择相对于未选定的表位能够被专职性抗原呈递细胞(APC)呈递至初始T细胞的可能性增加的表位,任选的其中所述APC是树突状细胞(DC)。

208.如权利要求201-207所述的方法,其中选择所述选定的表位组包括基于所述呈递模型来选择相对于未选定的表位,经由中枢或外周耐受而受到抑制的可能性降低的表位。

209.如权利要求201-208所述的方法,其中选择所述选定的表位组包括基于所述呈递模型来选择相对于未选定的表位能够在所述受试者中诱导对正常组织的自身免疫响应的可能性降低的表位。

210.如权利要求201-209所述的方法,其中外显子组或转录组核苷酸测序数据通过对所述肿瘤组织执行测序而获得。

211.如权利要求210所述的方法,其中所述测序是下一代测序(NGS)或任何大规模并行测序方法。

同源腺病毒疫苗接种

[0001] 相关申请的交叉引用

[0002] 本申请要求保护2020年12月3日提交的美国临时申请号63/121,164的权益,所述美国临时申请出于所有目的特此通过引用整体并入。

[0003] 序列表

[0004] 本申请含有已经通过电子方式以ASCII格式提交并且特此通过引用整体并入的序列表。所述ASCII副本创建于2020年12月3日,名为GSO-096_Sequence_Listing.txt,并且大小为6,925,547字节。

背景技术

[0005] 基于肿瘤特异性抗原的治疗性疫苗作为下一代个性化癌症免疫疗法具有极大的前景。[1-3]例如,鉴于新抗原生成的可能性相对较大,具有高突变负荷的癌症(诸如非小细胞肺癌(NSCLC)和黑色素瘤)是这种疗法的特别有吸引力的靶标。[4,5]早期证据显示,基于新抗原的疫苗接种可以引发T细胞响应[6],并且新抗原靶向细胞疗法可以在特定情况下导致选定的患者的肿瘤消退。[7]

[0006] 癌症和感染性疾病两种环境中抗原疫苗设计的一个问题是存在的许多编码突变中的哪一个生成“最佳”治疗性抗原,例如可以引发免疫力的抗原。

[0007] 除了当前的抗原预测方法的挑战之外,可以用于人类抗原递送的可用载体系统也存在许多挑战,其中许多来源于人类。例如,许多人由于先前的自然接触而对人类病毒具有预先存在的免疫力,而这种免疫力可能是将重组人病毒在疫苗用于接种策略中(诸如在癌症治疗或针对感染性疾病的疫苗接种中)进行抗原递送的主要障碍。虽然在解决以上问题的疫苗接种策略方面取得了一些进展,但仍然需要改进,特别是对于临床应用,诸如改进疫苗效力和功效。

发明内容

[0008] 本文公开了一种用于向受试者递送包含黑猩猩腺病毒(ChAdV)载体的组合物的方法,所述方法包括向受试者施用多个剂量的组合物,其中多个剂量至少包含第一剂量和第二剂量,并且其中第一剂量和第二剂量之间的时间段为至少27周。

[0009] 本文还公开了一种用于向受试者递送包含黑猩猩腺病毒(ChAdV)载体的组合物的方法,所述方法包括向受试者施用多个剂量的组合物,其中多个剂量至少包含第一剂量和第二剂量,并且其中在施用第二剂量之前,确定ChAdV特异性中和抗体滴度低于中和阈值。

[0010] 本文还公开了一种用于将包含黑猩猩腺病毒(ChAdV)载体的组合物递送至受试者的方法,所述方法包括:(a)向受试者施用第一剂量的组合物;(b)确定或已经确定ChAdV特异性中和抗体滴度;和(c)当确定ChAdV特异性中和抗体滴度低于中和阈值时,向受试者施用第二剂量的组合物。

[0011] 在一些方面,第一剂量和第二剂量之间的时间段为至少27周、至少28周、至少29周、至少30周、至少31周或至少32周。

[0012] 在一些方面，第一剂量是初免剂量。

[0013] 在一些方面，多个剂量包括三个或更多个剂量。在一些方面，多个剂量中的一个或多个在第一剂量之前施用。

[0014] 在一些方面，在第一剂量和第二剂量之间不施用额外剂量的组合物。

[0015] 在一些方面，中和抗体滴度是NT50值，其计算为中和ChAdV病毒50％的免疫受试者血清的最小稀释度。在一些方面，中和阈值是900或更小的NT50值。在一些方面，确定中和抗体滴度包括以下步骤：(1)在足以中和ChAdV病毒的条件下，将一种或多种稀释度的免疫受试者血清与ChAdV病毒接触；和(2)评估ChAdV病毒相对于未中和病毒的中和。在一些方面，评估中和步骤包括测定由ChAdV病毒表达的报告构建体的表达。

[0016] 在一些方面，中和阈值是在第二剂量中完全中和ChAdV病毒的最小中和抗体滴度。在一些方面，中和阈值是第二剂量在受试者中诱导免疫响应的最小中和抗体滴度。

[0017] 在一些方面，ChAdV载体编码至少一种抗原。在一些方面，至少一种抗原是非自身来源的肽，其中非自身来源的肽不由受试者的野生型基因编码。在一些方面，至少一种抗原是肿瘤相关抗原。在一些方面，肿瘤相关抗原是新抗原。

[0018] 在一些方面，至少一种抗原是外来抗原。在一些方面，外来抗原来自病原体、病毒、细菌、真菌或寄生虫。

[0019] 在一些方面，多个剂量中的每一个包含相同的一种或多种抗原。

[0020] 在一些方面，肌内(IM)、皮内(ID)、皮下(SC)或静脉内(IV)施用组合物。在一些方面，肌内施用组合物。在一些方面，在分开的注射部位施用肌内施用。在一些方面，分开的注射部位在相对的三角肌中。在一些方面，分开的注射部位在每一侧的臀肌或股直肌部位中。

[0021] 在一些方面，方法不包括施用免疫调节剂，或者方法在不存在免疫调节剂的情况下执行，任选地其中免疫调节剂是检查点抑制剂。

[0022] 在一些方面，方法还包括施用免疫调节剂。在一些方面，免疫调节剂是抗CTLA4抗体或其抗原结合片段、抗PD-1抗体或其抗原结合片段、抗PD-L1抗体或其抗原结合片段，抗4-1BB抗体或其抗原结合片段、或抗OX-40抗体或其抗原结合片段。

[0023] 在一些方面，方法还包括确定或已经确定受试者的HLA-单倍型。

[0024] 在一些方面，ChAdV载体包含：(a)ChAdV骨架，其中ChAdV骨架包含：(i)至少一个启动子核苷酸序列，和(ii)至少一个聚腺苷酸化(聚(A))序列；和(b)盒，其中盒包含：(i)至少一个抗原编码核酸序列，其任选地包含：a.表位编码核酸序列，其任选地包含至少一个改变，所述改变使编码的表位序列不同于由野生型核酸序列编码的对应肽序列，b.任选的5’接头序列，和c.任选的3’接头序列；并且其中盒可操作地连接至至少一个启动子核苷酸序列和至少一个聚(A)序列。

[0025] 在一些方面，ChAdV载体包含：(a)ChAdV骨架，其中ChAdV骨架包含：(i)修饰的ChAdV68序列，其至少包含SEQ ID NO:1中列出的序列的核苷酸2至36,518，其中核苷酸2至36,518缺少：(1)SEQ ID NO:1中示出的序列的核苷酸577至3403，其对应于E1缺失；(2)SEQ ID NO:1中示出的序列的核苷酸27,125至31,825，其对应于E3缺失；和(3)SEQ ID NO:1中示出的序列的核苷酸34,916至35,642，其对应于部分E4缺失；(ii)CMV启动子核苷酸序列；和(iii)SV40聚腺苷酸化(聚(A))序列；(b)盒，其中盒包含：(i)至少一个抗原编码核酸序列，其任选地包含：a.表位编码核酸序列，其任选地包含至少一个改变，所述改变使编码的表位

序列不同于由野生型核酸序列编码的对应肽序列,b.任选的5'接头序列,和c.任选的3'接头序列;并且其中盒插入到E1缺失中,并且盒可操作地连接至CMV启动子核苷酸序列和SV40聚(A)序列。

[0026] 在一些方面,表位编码核酸序列编码已知或疑似由细胞表面上的MHC I类和/或MHC II类呈递的表位,任选地其中细胞表面是肿瘤细胞表面或感染细胞表面,并且任选地其中细胞是受试者的细胞。在一些方面,至少一个抗原编码核酸序列编码能够被抗原加工成表位的多肽序列,任选地其中表位已知或疑似由细胞表面上的MHC I类呈递,任选地其中细胞表面是肿瘤细胞表面或感染细胞表面。在一些方面,细胞是选自由以下各项组成的组的肿瘤细胞:肺癌、黑色素瘤、乳腺癌、卵巢癌、前列腺癌、肾癌、胃癌、结肠癌、睾丸癌、头颈癌、胰腺癌、脑癌、B细胞淋巴瘤、急性骨髓性白血病、慢性骨髓性白血病、慢性淋巴细胞白血病、T细胞淋巴细胞白血病、非小细胞肺癌和小细胞肺癌,或其中细胞是选自由以下各项组成的组的感染细胞:病原体感染的细胞、病毒感染的细胞、细菌感染的细胞、真菌感染的细胞和寄生虫感染的细胞。

[0027] 在一些方面,病毒感染的细胞是HIV感染的细胞。在一些方面,表位编码核酸序列编码HIV GAG蛋白或表位。

[0028] 在一些方面,ChAdV载体中的盒的每个元件的有序序列在式中描述,其从5'至3'包含

[0029] P_a-($L5_b$-N_c-$L3_d$)$_X$-($G5_e$-U_f)$_Y$-$G3_g$

[0030] 其中P包含至少一个启动子序列,其可操作地连接至至少一个抗原编码核酸序列中的至少一个,其中a=1,N包含表位编码核酸序列之一,其中c=1,L5包含5'接头序列,其中b=0或1,L3包含3'接头序列,其中d=0或1,G5包含至少一个编码GPGPG氨基酸接头的核酸序列之一,其中e=0或1,G3包含至少一个编码GPGPG氨基酸接头的核酸序列之一,其中g=0或1,U包含至少一个MHC II类表位编码核酸序列之一,其中f=1,X=1至400,其中每个X对应的N_c是表位编码核酸序列,并且Y=0、1或2,其中每个Y对应的U_f是MHC II类表位编码核酸序列。

[0031] 在一些方面,每个X对应的N_c是不同的表位编码核酸序列。在一些方面,每个Y对应的U_f是不同的MHC II类表位编码核酸序列。

[0032] 在一些方面,b=1,d=1,e=1,g=1,h=1,X=10,Y=2,P是CMV启动子序列,每个N编码长度为7-15个氨基酸的MHC I类表位、MHC II类表位、能够刺激B细胞响应的表位或其组合,L5是编码表位的天然N末端氨基酸序列的天然5'接头序列,并且其中5'接头序列编码长度为至少3个氨基酸的肽,L3是编码表位的天然C末端氨基酸序列的天然3'接头序列,并且其中3'接头序列编码长度为至少3个氨基酸的肽,U分别是II类PADRE序列和破伤风类毒素MHC II类序列,ChAdV载体包含修饰的ChAdV68序列,所述ChAdV68序列至少包含SEQ ID NO:1中列出的序列的核苷酸2至36,518,其中核苷酸2至36,518缺少:(1)SEQ ID NO:1中示出的序列的核苷酸577至3403,其对应于E1缺失;(2)SEQ ID NO:1中示出的序列的核苷酸27,125至31,825,其对应于E3缺失;和(3)SEQ ID NO:1中示出的序列的核苷酸34,916至35,642,其对应于部分E4缺失,并且新抗原盒插入到E1缺失中,并且I抗原编码核酸序列中的每一个编码长度为25个氨基酸的多肽。

[0033] 在一些方面,盒整合在至少一个启动子核苷酸序列和至少一个聚(A)序列之间。在

一些方面,至少一个启动子核苷酸序列可操作地连接至盒。

[0034] 在一些方面,ChAdV骨架包含ChAdV68载体骨架。在一些方面,ChAdV68载体骨架包含SEQ ID NO:1中列出的序列。在一些方面,ChAdV68载体骨架包含选自由以下各项组成的组的至少一个基因中的功能性突变:参考ChAdV68基因组或参考SEQ ID NO:1中示出的序列的腺病毒E1A、E1B、E2A、E2B、E3、L1、L2、L3、L4和L5基因,任选地其中腺病毒骨架或修饰的ChAdV68序列在以下中完全缺失或功能性缺失:参考腺病毒基因组或参考SEQ ID NO:1中示出的序列的(1)E1A和E1B;或(2)E1A、E1B和E3,任选地其中E1基因通过参考SEQ ID NO:1中示出的序列的至少核苷酸577至3403的E1缺失而功能性缺失,并且任选地其中E3基因通过参考SEQ ID NO:1中示出的序列的至少核苷酸27,125至31,825的E3缺失而功能性缺失。在一些方面,ChAdV68载体骨架包含参考ChAdV68基因组或参考SEQ ID NO:1中示出的序列的一个或多个基因或调控序列,任选地其中一个或多个基因或调控序列选自由以下组成的组:黑猩猩腺病毒反向末端重复序列(ITR)、E1A、E1B、E2A、E2B、E3、E4、L1、L2、L3、L4和L5基因。

[0035] 在一些方面,ChAdV68载体骨架包含部分缺失的E4基因。在一些方面,部分缺失的E4基团包含:A.SEQ ID NO:1中示出的E4基因序列,其至少缺少SEQ ID NO:1中示出的序列的核苷酸34,916至35,642,B.SEQ ID NO:1中示出的E4基因序列,其至少缺少SEQ ID NO:1中示出的序列的核苷酸34,916至34,942、核苷酸34,952至35,305、SEQ ID NO:1中示出的序列的核苷酸35,302至35,642,并且其中载体至少包含SEQ ID NO:1中示出的序列的核苷酸2至36,518,C.SEQ ID NO:1中示出的E4基因序列,其至少缺少SEQ ID NO:1中示出的序列的核苷酸34,980至36,516,并且其中载体至少包含SEQ ID NO:1中示出的序列的核苷酸2至36,518,D.SEQ ID NO:1中示出的E4基因序列,其至少缺少SEQ ID NO:1中示出的序列的核苷酸34,979至35,642,并且其中载体至少包含SEQ ID NO:1中示出的序列的核苷酸2至36,518,E.E4Orf2的至少部分缺失、完全缺失的E4Orf3和E4Orf4的至少部分缺失的E4缺失,F.E4Orf2的至少部分缺失、E4Orf3的至少部分缺失和E4Orf4的至少部分缺失的E4缺失,G.E4Orf1的至少部分缺失、完全缺失的E4Orf2和E4Orf3的至少部分缺失的E4缺失,或H.E4Orf2的至少部分缺失和E4Orf3的至少部分缺失的E4缺失。

[0036] 在一些方面,ChAdV68载体骨架至少包含SEQ ID NO:1中列出的序列的核苷酸2至36,518,其中核苷酸2至36,518缺少:(1)SEQ ID NO:1中示出的序列的核苷酸577至3403,其对应于E1缺失;(2)SEQ ID NO:1中示出的序列的核苷酸27,125至31,825,其对应于E3缺失;和(3)SEQ ID NO:1中示出的序列的核苷酸34,916至35,642,其对应于部分E4缺失;任选地其中抗原盒插入到E1缺失中。在一些方面,ChAdV68载体骨架包含SEQ ID NO:29369中列出的序列,任选地其中抗原盒插入到E1缺失中。在一些方面,ChAdV68载体骨架至少包含SEQ ID NO:1中列出的序列的核苷酸2至36,518,其中核苷酸2至36,518缺少:A.SEQ ID NO:1中示出的序列的核苷酸577至3403,其对应于E1缺失;B.SEQ ID NO:1中示出的序列的核苷酸27,125至31,825,其对应于E3缺失;C.SEQ ID NO:1中示出的序列的核苷酸34,916至35,642,其对应于部分E4缺失;D.SEQ ID NO:1中示出的序列的核苷酸456至3014;E.SEQ ID NO:1中示出的序列的核苷酸27,816至31,333;F.SEQ ID NO:1中示出的序列的核苷酸3957至10346;G.SEQ ID NO:1中示出的序列的核苷酸21787至23370;H.SEQ ID NO:1中示出的序列的核苷酸33486至36193;或其组合。

[0037] 在一些方面,ChAdV68载体骨架至少包含SEQ ID NO:1中列出的序列的核苷酸2至36,518,其中核苷酸2至36,518缺少:(1)SEQ ID NO:1中示出的序列的核苷酸577至3403,其对应于E1缺失;和(2)SEQ ID NO:1中示出的序列的核苷酸27,125至31,825,其对应于E3缺失。

[0038] 在一些方面,其中盒插入ChAdV骨架的E1区、E3区和/或允许掺入盒的任何缺失的AdV区。

[0039] 在一些方面,ChAdV骨架由第一代、第二代或辅助依赖性腺病毒载体之一生成。在一些方面,至少一个启动子核苷酸序列选自由以下各项组成的组:CMV、SV40、EF-1、RSV、PGK、HSA、MCK和EBV启动子序列。在一些方面,至少一个启动子核苷酸序列是CMV启动子序列。

[0040] 在一些方面,表位编码核酸序列中的至少一个编码表位,所述表位在表达和翻译时能够由受试者细胞上的MHC I类呈递。在一些方面,表位编码核酸序列中的至少一个编码在表达和翻译时能够由受试者细胞上的MHC II类呈递的表位。

[0041] 在一些方面,至少一个抗原编码核酸序列包含两个或更多个抗原编码核酸序列。在一些方面,每个抗原编码核酸序列彼此直接连接。

[0042] 在一些方面,每个抗原编码核酸序列通过编码接头的核酸序列连接至不同的抗原编码核酸序列。在一些方面,接头连接两个表位编码核酸序列,或者将表位编码核酸序列连接至MHC II类表位编码核酸序列。在一些方面,接头选自由以下各项组成的组:(1)长度为至少2、3、4、5、6、7、8、9或10个残基的连续甘氨酸残基;(2)长度为至少2、3、4、5、6、7、8、9或10个残基的连续丙氨酸残基;(3)两个精氨酸残基(RR);(4)丙氨酸、丙氨酸、酪氨酸(AAY);(5)被哺乳动物蛋白酶体高效加工的长度为至少2、3、4、5、6、7、8、9或10个氨基酸残基的共有序列;和(6)侧接来源于同源蛋白的抗原并且长度为至少2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20或2-20个氨基酸的一个或多个天然序列。在一些方面,接头连接两个MHC II类表位编码核酸序列,或者将MHC II类序列连接至表位编码核酸序列。在一些方面,接头包含序列GPGPG。

[0043] 在一些方面,抗原编码核酸序列可操作地或直接连接至分开或连续的序列,所述分开或连续的序列增强抗原编码核酸序列的表达、稳定性、细胞运输、加工和呈递和/或免疫原性。在一些方面,分开或连续的序列包含以下中的至少一种:泛素序列、经修饰以增加蛋白酶体靶向的泛素序列(例如,在位置76处含有Gly至Ala取代的泛素序列)、免疫球蛋白信号序列(例如,IgK)、主要组织相容性I类序列、溶酶体相关膜蛋白(LAMP)-1、人树突状细胞溶酶体相关膜蛋白和主要组织相容性II类序列;任选地,其中经修饰以增加蛋白酶体靶向的泛素序列是A76。

[0044] 在一些方面,表位编码核酸序列包含至少一个改变,所述改变使得相对于翻译的对应野生型核酸序列,编码的表位对其对应MHC等位基因的结合亲和力增加。在一些方面,表位编码核酸序列包含至少一个改变,所述改变使得相对于翻译的对应野生型核酸序列,编码的表位对其对应MHC等位基因的结合稳定性增加。在一些方面,表位编码核酸序列包含至少一个改变,所述改变使得相对于翻译的对应野生型核酸序列,编码的表位在其对应MHC等位基因上的呈递可能性增加。在一些方面,至少一个改变包括点突变、移码突变、非移码突变、缺失突变、插入突变、剪接变异、基因组重排或蛋白酶体生成的剪接抗原。

[0045] 在一些方面，表位编码核酸序列编码已知或疑似在已知或疑似患有癌症的受试者中表达的表位。在一些方面，癌症包括实体瘤。在一些方面，癌症选自由以下各项组成的组：肺癌、黑色素瘤、乳腺癌、卵巢癌、前列腺癌、肾癌、胃癌、结肠癌、睾丸癌、头颈癌、胰腺癌、膀胱癌、脑癌、B细胞淋巴瘤、急性骨髓性白血病、成人急性成淋巴细胞性白血病、慢性骨髓性白血病、慢性淋巴细胞白血病、T细胞淋巴细胞白血病、非小细胞肺癌和小细胞肺癌。

[0046] 在一些方面，至少一个抗原编码核酸序列包含至少2-10、2、3、4、5、6、7、8、9或10个抗原编码核酸序列，任选地其中每个抗原编码核酸序列编码不同的抗原编码核酸序列。在一些方面，至少一个抗原编码核酸序列包含至少11-20、15-20、11-100、11-200、11-300、11-400、11、12、13、14、15、16、17、18、19、20或至多400个抗原编码核酸序列，任选地其中每个抗原编码核酸序列编码不同的抗原编码核酸序列。在一些方面，至少一个抗原编码核酸序列包含至少11-20、15-20、11-100、11-200、11-300、11-400、11、12、13、14、15、16、17、18、19、20或至多400个抗原编码核酸序列。

[0047] 在一些方面，至少一个抗原编码核酸序列包含至少2-400个抗原编码核酸序列，并且其中抗原编码核酸序列的至少两个编码由细胞表面上的MHC I类呈递的表位序列或其部分。在一些方面，MHC I类表位中的至少两个由肿瘤细胞表面上的MHC I类呈递。

[0048] 在一些方面，表位编码核酸序列包含至少一个MHC I类表位编码核酸序列，并且其中每个抗原编码核酸序列编码长度在8至35个氨基酸之间的多肽序列，任选地长度为9-17、9-25、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34或35个氨基酸。

[0049] 在一些方面，存在至少一个MHC II类表位编码核酸序列。

[0050] 在一些方面，存在至少一个MHC II类表位编码核酸序列，并且其包含至少一个MHC II类表位编码核酸序列，所述MHC II类表位编码核酸序列包含至少一个改变，所述改变使编码的表位序列不同于由野生型核酸序列编码的对应肽序列。

[0051] 在一些方面，表位编码核酸序列包含MHC II类表位编码核酸序列，并且其中每个抗原编码核酸序列编码长度为12-20、12、13、14、15、16、17、18、19、20或20-40个氨基酸的多肽序列。在一些方面，表位编码核酸序列包含MHC II类表位编码核酸序列，其中存在至少一个MHC II类表位编码核酸序列，并且其中至少一个MHC II类表位编码核酸序列包含至少一个通用MHC II类表位编码核酸序列，任选地其中至少一个通用序列包含破伤风类毒素和PADRE中的至少一种。

[0052] 在一些方面，至少一个启动子核苷酸序列是诱导型的。在一些方面，至少一个启动子核苷酸序列是非诱导型的。在一些方面，至少一个聚(A)序列包含牛生长激素(BGH)SV40聚A序列。在一些方面，至少一个聚(A)序列是至少20个、至少30个、至少40个、至少50个、至少60个、至少70个、至少80个或至少90个连续的A核苷酸。在一些方面，至少一个聚(A)序列是至少100个连续的A核苷酸。

[0053] 在一些方面，盒还包含以下中的至少一种：内含子序列、土拨鼠肝炎病毒转录后调控元件(WPRE)序列、内部核糖体进入序列(IRES)序列、编码2A自切割肽序列的核苷酸序列、编码弗林蛋白酶(Furin)切割位点的核苷酸序列、或5'或3'非编码区中的已知增强可操作地连接至至少一个抗原编码核酸序列中的至少一个的mRNA的核输出、稳定性或翻译效率的序列。在一些方面，盒还包含报告基因，包括但不限于绿色荧光蛋白(GFP)、GFP变体、分泌型

碱性磷酸酶、荧光素酶、荧光素酶变体或可检测肽或表位。在一些方面，可检测肽或表位选自由HA标签、Flag标签、His标签或V5标签组成的组。

[0054] 在一些方面，一种或多种载体还包含一种或多种编码至少一种免疫调节剂的核酸序列。在一些方面，免疫调节剂是抗CTLA4抗体或其抗原结合片段、抗PD-1抗体或其抗原结合片段、抗PD-L1抗体或其抗原结合片段、抗4-1BB抗体或其抗原结合片段、或抗OX-40抗体或其抗原结合片段。在一些方面，抗体或其抗原结合片段是Fab片段、Fab'片段、单链Fv(scFv)、单特异性的或多重特异性连接在一起的单结构域抗体(sdAb)(例如，骆驼科抗体结构域)或全长单链抗体(例如，具有通过柔性接头连接的重链和轻链的全长IgG)。在一些方面，抗体的重链和轻链序列是被自切割序列(诸如2A或IRES)分开的连续序列；或者抗体的重链和轻链序列通过柔性接头(诸如连续的甘氨酸残基)连接。在一些方面，免疫调节剂是细胞因子。在一些方面，细胞因子是IL-2、IL-7、IL-12、IL-15或IL-21中的至少一种或其各自的变体。

[0055] 在一些方面，至少一个表位编码核酸序列通过执行以下步骤来选择：(a)从肿瘤、感染细胞或感染疾病生物体中获得外显子组、转录组或全基因组核苷酸测序数据中的至少一种，其中核苷酸测序数据用于获得代表一组抗原中的每一个的肽序列的数据；(b)将每个抗原的肽序列输入呈递模型以生成每个抗原由细胞表面(任选的肿瘤细胞表面或感染细胞表面)上的一个或多个MHC等位基因呈递的一组数值可能性，已经至少基于接收到的质谱数据来鉴定的数值可能性组；和(c)基于数值可能性组来选择抗原组的子集以生成一组选定的抗原，所述选定的抗原用于生成至少一个表位编码核酸序列。在一些方面，表位编码核酸序列中的每一个通过执行以下步骤来选择：(a)从肿瘤、感染细胞或感染疾病生物体中获得外显子组、转录组或全基因组核苷酸测序数据中的至少一种，其中核苷酸测序数据用于获得代表一组抗原中的每一个的肽序列的数据；(b)将每个抗原的肽序列输入呈递模型以生成每个抗原由细胞表面(任选的肿瘤细胞表面或感染细胞表面)上的一个或多个MHC等位基因呈递的一组数值可能性，已经至少基于接收到的质谱数据来鉴定的数值可能性组；和(c)基于数值可能性组来选择抗原组的子集以生成一组选定的抗原，所述选定的抗原用于生成至少20个表位编码核酸序列。在一些方面，选定的表位组的数量是2-20。在一些方面，呈递模型代表以下之间的依赖性：(a)在肽序列的特定位置存在一个特定MHC等位基因和特定氨基酸的对；和(b)此类在特定位置处包含特定氨基酸的肽序列由所述对的一个特定MHC等位基因呈递在细胞表面(任选的肿瘤细胞表面或感染细胞表面)上的可能性。在一些方面，选择选定的抗原组包括基于呈递模型来选择相对于未选定的抗原在细胞表面上呈递的可能性增加的抗原，任选地其中选定的抗原已被验证为由一个或多个指定HLA等位基因呈递。在一些方面，选择选定的表位组包括基于呈递模型来选择相对于未选定的表位能够在受试者中诱导肿瘤特异性免疫响应的可能性增加的表位。在一些方面，选择选定的抗原组包括基于呈递模型来选择相对于未选定的抗原能够在受试者中诱导肿瘤特异性或感染性疾病特异性免疫响应的可能性增加的抗原。在一些方面，选择选定的抗原组包括基于呈递模型来选择相对于未选定的抗原能够被专职性抗原呈递细胞(APC)呈递至初始T细胞的可能性增加的抗原，任选的其中APC是树突状细胞(DC)。在一些方面，选择选定的抗原组包括基于呈递模型来选择相对于未选定的抗原，经由中枢或外周耐受而受到抑制的可能性降低的抗原。在一些方面，选择选定的抗原组包括基于呈递模型来选择相对于未选定的抗原能够在

受试者中诱导对正常组织的自身免疫响应的可能性降低的抗原。在一些方面，外显子组或转录组核苷酸测序数据通过对肿瘤细胞或组织、感染细胞或感染性疾病生物体执行测序而获得。在一些方面，测序是下一代测序(NGS)或任何大规模并行测序方法。

[0056] 在一些方面，盒包含由盒中的相邻序列形成的连接表位序列。在一些方面，至少一个或每个连接表位序列对MHC具有大于500nM的亲和力。在一些方面，每个连接表位序列是非自体的。

[0057] 在一些方面，每个MHC I类表位被预测或验证为能够由存在于至少5％的群体中的至少一个HLA等位基因呈递。在一些方面，每个MHC I类表位被预测或验证为能够由至少一个HLA等位基因呈递，其中每个抗原/HLA对在群体中具有至少0.01％的抗原/HLA患病率。在一些方面，每个MHC I类表位被预测或验证为能够由至少一个HLA等位基因呈递，其中每个抗原/HLA对在群体中具有至少0.1％的抗原/HLA患病率。

[0058] 在一些方面，盒不编码包含翻译的野生型核酸序列的非治疗性MHC I类或II类表位核酸序列，其中预测非治疗性表位在受试者的MHC等位基因上呈现。在一些方面，非治疗性预测的MHC I类或II类表位序列是由盒中的相邻序列形成的连接表位序列。在一些方面，预测基于通过将非治疗性表位的序列输入呈递模型而生成的呈递可能性。在一些方面，盒中的抗原编码核酸序列的顺序通过包含以下的一系列步骤来确定：(a)生成对应于抗原编码核酸序列的不同顺序的一组候选盒序列；(b)对于每个候选盒序列，基于候选盒序列中非治疗性表位的呈递来确定呈递评分；和(c)选择与低于预定阈值的呈递评分相关的候选盒序列作为抗原疫苗的盒序列。

[0059] 在一些方面，组合物被配制为包含药学上可接受的载剂的药物组合物。在一些方面，组合物包含有包含ChAdV载体的病毒颗粒。

[0060] 在一些方面，一个或多个表位编码核酸序列来源于受试者的肿瘤。在一些方面，每个表位编码核酸序列来源于受试者的肿瘤。在一些方面，一个或多个表位编码核酸序列不来源于受试者的肿瘤。在一些方面，每个表位编码核酸序列不来源于受试者的肿瘤。

[0061] 在一些方面，表位编码核酸序列包含选自由SEQ ID NO:57-29,364组成的组的表位。在一些方面，至少一个抗原编码核酸序列包含以下中的至少每一个：(A)KRAS_G12C MHC I类表位编码核酸序列，其中KRAS_G12C MHC I类表位编码核酸序列编码选自由SEQ ID NO:14,954;19,848;和19,850组成的组的MHC I类表位，(B)KRAS_G12D MHC I类表位编码核酸序列，其中KRAS_G12DMHC I类表位编码核酸序列编码选自由SEQ ID NO:19,749;19,865;和19,863组成的组的MHC I类表位，和(C)KRAS_G12V MHC I类表位编码核酸序列，其中KRAS_G12V MHC I类表位编码核酸序列编码选自由SEQ ID NO:19,976;19,779;11,495;和19,974组成的组的MHC I类表位。

[0062] 在一些方面，至少一个抗原编码核酸序列包含：(A)KRAS_G12CMHC I类表位编码核酸序列、(B)KRAS_G12D MHC I类表位编码核酸序列、(C)KRAS_G12V MHC I类表位编码核酸序列、(D)KRAS Q61H MHC I类表位编码核酸序列、(E)TP53_R213L MHC I类表位编码核酸序列、(F)TP53_S127Y MHC I类表位编码核酸序列、(G)TP53_R249M MHC I类表位编码核酸序列或其组合。

[0063] 在一些方面，方法还包括施用基于甲病毒的自扩增表达系统。在一些方面，肌内(IM)、皮内(ID)、皮下(SC)或静脉内(IV)施用用于递送基于甲病毒的自扩增表达系统的组

合物。在一些方面，肌内施用用于递送基于甲病毒的自扩增表达系统的组合物。在一些方面，在分开的注射部位施用肌内施用。在一些方面，分开的注射部位在相对的三角肌中。在一些方面，分开的注射部位在每一侧的臀肌或股直肌部位中。在一些方面，一个或多个加强剂量的注射部位尽可能靠近初免剂量的注射部位。

[0064] 在一些方面，用于递送基于甲病毒的自扩增表达系统的组合物包含：(A)基于甲病毒的自扩增表达系统，其中基于甲病毒的自扩增表达系统包含一种或多种载体，其中一种或多种载体包含：(a)RNA甲病毒骨架，其中RNA甲病毒骨架包含：(i)至少一个启动子核苷酸序列，和(ii)至少一个聚腺苷酸化(聚(A))序列；和(b)盒，其中盒包含：(i)至少一个抗原编码核酸序列，其包含：a.表位编码核酸序列，其任选地包含至少一个改变，所述改变使编码的表位序列不同于由野生型核酸序列编码的对应肽序列，b.任选的5'接头序列，和c.任选的3'接头序列；(ii)任选地，可操作地连接至至少一个抗原编码核酸序列的第二启动子核苷酸序列；和(iii)任选地，至少一个第二聚(A)序列，其中第二聚(A)序列是甲病毒的天然聚(A)序列或外源聚(A)序列，和(B)脂质纳米颗粒(LNP)，其中LNP封装基于甲病毒的自扩增表达系统。

[0065] 在一些方面，用于递送基于甲病毒的自扩增表达系统的组合物包含：(A)基于甲病毒的自扩增表达系统，其中基于甲病毒的自扩增表达系统包含一种或多种载体，其中一种或多种载体包含：(a)RNA甲病毒骨架，其中RNA甲病毒骨架包含SEQ ID NO:6中列出的核酸序列，其中RNA甲病毒骨架序列包含26S启动子核苷酸序列和聚(A)序列，其中26S启动子序列对于RNA甲病毒骨架是内源的，并且其中聚(A)序列对于RNA甲病毒骨架是内源的；和(b)整合在26S启动子核苷酸序列和聚(A)序列之间的盒，其中盒可操作地连接至26S启动子核苷酸序列，并且其中盒包含至少一个抗原编码核酸序列，所述抗原编码核酸序列包含：a.表位编码核酸序列，其任选地包含至少一个改变，所述改变使编码的表位序列不同于由野生型核酸序列编码的对应肽序列，b.任选的5'接头序列，和c.任选的3'接头序列；和(B)脂质纳米颗粒(LNP)，其中LNP封装基于甲病毒的自扩增表达系统。

[0066] 在一些方面，ChAdV载体的盒与用于递送基于甲病毒的自扩增表达系统的组合物的盒相同。

[0067] 在一些方面，用于递送基于甲病毒的自扩增表达系统的组合物中的盒的每个元件的有序序列在下式中描述，从5'至3'包含：P_a-($L5_b$-N_c-$L3_d$)$_X$-($G5_e$-U_f)$_Y$-$G3_g$，其中P包含第二启动子核苷酸序列，其中a＝0或1，N包含表位编码核酸序列之一，其中c＝1，L5包含5'接头序列，其中b＝0或1，L3包含3'接头序列，其中d＝0或1，G5包含至少一个编码GPPGG氨基酸接头的核酸序列之一，其中e＝0或1，G3包含至少一个编码GPGPG氨基酸接头的核酸序列之一，其中g＝0或1，U包含至少一个MHC II类表位编码核酸序列之一，其中f＝1，X＝1至400，其中对于每个X，对应的N_c是表位编码核酸序列，并且Y＝0、1或2，其中对于每个Y，对应的U_f是MHC II类表位编码核酸序列。

[0068] 在一些方面，每个X对应的N_c是不同的表位编码核酸序列。在一些方面，每个Y对应的U_f是不同的MHC II类表位编码核酸序列。在一些方面，a＝0、b＝1、d＝1、e＝1、g＝1、h＝1、X＝20、Y＝2，至少一个启动子核苷酸序列是由RNA甲病毒骨架提供的单个26S启动子核苷酸序列，至少一个聚腺苷酸化聚(A)序列是由RNA甲病毒骨架提供的至少100个连续A核苷酸的聚(A)序列，盒整合在26S启动子核苷酸序列和聚(A)序列之间，其中盒可操作地连接至

26S启动子核苷酸序列和聚(A)序列,每个N编码长度为7-15个氨基酸的MHC I类表位、MHC II类表位、能够刺激B细胞响应的表位或其组合,L5是编码表位的天然N末端氨基酸序列的天然5'接头序列,并且其中5'接头序列编码长度为至少3个氨基酸的肽,L3是编码表位的天然C末端氨基酸序列的天然3'接头序列,并且其中3'接头序列编码长度为至少3个氨基酸的肽,U分别是PADRE II类序列和破伤风类毒素MHC II类序列,RNA甲病毒骨架是SEQ ID NO:6中列出的序列,并且每个MHC I类表位编码核酸序列编码长度在13和25个氨基酸之间的多肽。

[0069] 在一些方面,LNP包含选自由以下各项组成的组的脂质:可电离的氨基脂质、磷脂酰胆碱、胆固醇、基于PEG的外壳脂质或其组合。在一些方面,LNP包含可电离的氨基脂质、磷脂酰胆碱、胆固醇和基于PEG的外壳脂质。在一些方面,可电离的氨基脂质包含MC3样(二亚油醇基甲基-4-二甲基氨基丁酸酯)分子。在一些方面,LNP封装的表达系统具有约100nm的直径。

[0070] 在一些方面,盒整合在至少一个启动子核苷酸序列和至少一个聚(A)序列之间。在一些方面,至少一个启动子核苷酸序列可操作地连接至盒。

[0071] 在一些方面,一个或多个载体包含一个或多个+链RNA载体。在一些方面,一个或多个+链RNA载体包含5'7-甲基鸟苷(m7g)帽。在一些方面,一个或多个+链RNA载体通过体外转录产生。

[0072] 在一些方面,一个或多个载体在哺乳动物细胞内自复制。在一些方面,RNA甲病毒骨架包含Aura病毒、摩根堡病毒(Fort Morgan virus)、委内瑞拉马脑炎病毒(Venezuelan equine encephalitis virus)、罗斯河病毒(Ross River virus)、森林脑炎病毒(Semliki Forest virus)、辛德比斯病毒(Sindbis virus)或马亚罗病毒(Mayaro virus)的至少一个核苷酸序列。在一些方面,RNA甲病毒骨架包含委内瑞拉马脑炎病毒的至少一个核苷酸序列。在一些方面,RNA甲病毒骨架至少包含由Aura病毒、摩根堡病毒、委内瑞拉马脑炎病毒、罗斯河病毒、森林脑炎病毒、辛德比斯病毒或马亚罗病毒的核苷酸序列编码的用于非结构蛋白介导的扩增的序列、26S启动子序列、聚(A)序列、非结构蛋白1(nsP1)基因、nsP2基因、nsP3基因和nsP4基因在一些方面,RNA甲病毒骨架至少包含由Aura病毒、摩根堡病毒、委内瑞拉马脑炎病毒、罗斯河病毒、森林脑炎病毒、辛德比斯病毒或马亚罗病毒的核苷酸序列编码的用于非结构蛋白介导的扩增的序列、26S启动子序列和聚(A)序列。在一些方面,用于非结构蛋白介导的扩增的序列选自由以下各项组成的组:假病毒5'UTR、51-nt CSE、24-nt CSE、26S亚基因组启动子序列、19-nt CSE、假病毒3'UTR或其组合。在一些方面,RNA甲病毒骨架不编码结构病毒粒子蛋白衣壳E2和E1。在一些方面,插入盒以替代Aura病毒、摩根堡病毒、委内瑞拉马脑炎病毒、罗斯河病毒、森林脑炎病毒、辛德比斯病毒或马亚罗病毒的核苷酸序列内的结构病毒粒子蛋白。在一些方面,委内瑞拉马脑炎病毒包含SEQ ID NO:3或SEQ ID NO:5的序列。在一些方面,委内瑞拉马脑炎病毒包含有还包含碱基对7544和11175之间的缺失的SEQ ID NO:3或SEQ ID NO:5的序列。在一些方面,RNA甲病毒骨架包含SEQ ID NO:6或SEQ ID NO:7中列出的序列。在一些方面,盒插入位置7544以替代如SEQ ID NO:3或SEQ ID NO:5的序列中列出的碱基对7544和11175之间的缺失。在一些方面,盒的插入提供包含nsP1-4基因和至少一个核酸序列的多顺反子RNA的转录,其中nsP1-4基因和至少一个核酸序列在单独的开放阅读框中。

[0073] 在一些方面，至少一个启动子核苷酸序列是由RNA甲病毒骨架编码的天然26S启动子核苷酸序列。在一些方面，至少一个启动子核苷酸序列是外源RNA启动子。在一些方面，第二启动子核苷酸序列是26S启动子核苷酸序列。在一些方面，第二启动子核苷酸序列包含多个26S启动子核苷酸序列，其中每个26S启动子核苷酸序列提供单独的开放阅读框中的一个或多个的转录。

[0074] 在一些方面，一个或多个载体各自的大小为至少300nt。在一些方面，一个或多个载体各自的大小为至少1kb。在一些方面，一个或多个载体各自的大小为2kb。在一些方面，一个或多个载体各自的大小小于5kb。

[0075] 在一些方面，至少一个抗原编码核酸序列包含两个或更多个抗原编码核酸序列。在一些方面，每个抗原编码核酸序列彼此直接连接。在一些方面，每个抗原编码核酸序列通过编码接头的核酸序列连接至不同的抗原编码核酸序列。在一些方面，接头连接两个MHC I类表位编码核酸序列，或者将MHC I类表位编码核酸序列连接至MHC II类表位编码核酸序列。在一些方面，接头选自由以下各项组成的组：(1)长度为至少2、3、4、5、6、7、8、9或10个残基的连续甘氨酸残基；(2)长度为至少2、3、4、5、6、7、8、9或10个残基的连续丙氨酸残基；(3)两个精氨酸残基(RR)；(4)丙氨酸、丙氨酸、酪氨酸(AAY)；(5)被哺乳动物蛋白酶体高效加工的长度为至少2、3、4、5、6、7、8、9或10个氨基酸残基的共有序列；和(6)侧接来源于同源蛋白的抗原并且长度为至少2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20或2-20个氨基酸的一个或多个天然序列。在一些方面，接头连接两个MHC II类表位编码核酸序列，或者将MHC II类序列连接至MHC I类表位编码核酸序列。在一些方面，接头包含序列GPGPG。

[0076] 在一些方面，抗原编码核酸序列可操作地或直接连接至分开或连续的序列，所述分开或连续的序列增强抗原编码核酸序列的表达、稳定性、细胞运输、加工和呈递和/或免疫原性。在一些方面，分开或连续的序列包含以下中的至少一种：泛素序列、经修饰以增加蛋白酶体靶向的泛素序列(例如，在位置76处含有Gly至Ala取代的泛素序列)、免疫球蛋白信号序列(例如，IgK)、主要组织相容性I类序列、溶酶体相关膜蛋白(LAMP)-1、人树突状细胞溶酶体相关膜蛋白和主要组织相容性II类序列；任选地，其中经修饰以增加蛋白酶体靶向的泛素序列是A76。

[0077] 在一些方面，至少一个抗原编码核酸序列包含至少2-10、2、3、4、5、6、7、8、9或10个抗原编码核酸序列，任选地其中每个抗原编码核酸序列编码不同的抗原编码核酸序列。在一些方面，至少一个抗原编码核酸序列包含至少11-20、15-20、11-100、11-200、11-300、11-400、11、12、13、14、15、16、17、18、19、20或至多400个抗原编码核酸序列，任选地其中每个抗原编码核酸序列编码不同的抗原编码核酸序列。在一些方面，至少一个抗原编码核酸序列包含至少11-20、15-20、11-100、11-200、11-300、11-400、11、12、13、14、15、16、17、18、19、20或至多400个抗原编码核酸序列。在一些方面，至少一个抗原编码核酸序列包含至少2-400个抗原编码核酸序列，并且其中抗原编码核酸序列的至少两个编码由细胞表面上的MHC I类呈递的表位序列或其部分。

[0078] 在一些方面，MHC I类表位中的至少两个由肿瘤细胞表面上的MHC I类呈递。在一些方面，表位编码核酸序列包含至少一个MHC I类表位编码核酸序列，并且其中每个抗原编码核酸序列编码长度在8至35个氨基酸之间的多肽序列，任选地长度为9-17、9-25、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34或35个

氨基酸。

[0079] 在一些方面,存在至少一个MHC II类表位编码核酸序列。在一些方面,存在至少一个MHC II类表位编码核酸序列,并且其包含至少一个MHC II类表位编码核酸序列,所述MHC II类表位编码核酸序列包含至少一个改变,所述改变使编码的表位序列不同于由野生型核酸序列编码的对应肽序列。在一些方面,表位编码核酸序列包含MHC II类表位编码核酸序列,并且其中每个抗原编码核酸序列编码长度为12-20、12、13、14、15、16、17、18、19、20或20-40个氨基酸的多肽序列。在一些方面,表位编码核酸序列包含MHC II类表位编码核酸序列,其中存在至少一个MHC II类表位编码核酸序列,并且其中至少一个MHC II类表位编码核酸序列包含至少一个通用MHC II类表位编码核酸序列,任选地其中至少一个通用序列包含破伤风类毒素和PADRE中的至少一种。

[0080] 在一些方面,至少一个启动子核苷酸序列或第二启动子核苷酸序列是诱导型的。在一些方面,至少一个启动子核苷酸序列或第二启动子核苷酸序列是非诱导型的。

[0081] 在一些方面,至少一个聚(A)序列包含甲病毒天然的聚(A)序列。在一些方面,至少一个聚(A)序列包含甲病毒外源的聚(A)序列。在一些方面,至少一个聚(A)序列可操作地连接至至少一个核酸序列中的至少一个。在一些方面,至少一个聚(A)序列是至少20个、至少30个、至少40个、至少50个、至少60个、至少70个、至少80个或至少90个连续的A核苷酸。在一些方面,至少一个聚(A)序列是至少100个连续的A核苷酸。

[0082] 在一些方面,表位编码核酸序列包含MHC I类表位编码核酸序列,并且其中MHC I类表位编码核酸序列通过执行以下步骤来选择:(a)从肿瘤中获得外显子组、转录组或全基因组肿瘤核苷酸测序数据中的至少一种,其中肿瘤核苷酸测序数据用于获得代表一组表位中的每一个的肽序列的数据;(b)将每个表位的肽序列输入呈递模型以生成每个表位由肿瘤的肿瘤细胞表面上的一个或多个MHC等位基因呈递的一组数值可能性,已经至少基于接收到的质谱数据来鉴定的数值可能性组;(c)基于数值可能性组来选择表位组的子集以生成一组选定的表位,所述选定的表位用于生成MHC I类表位编码核酸序列。在一些方面,MHC I类表位编码核酸序列中的每一个通过执行以下步骤来选择:(a)从肿瘤中获得外显子组、转录组或全基因组肿瘤核苷酸测序数据中的至少一种,其中肿瘤核苷酸测序数据用于获得代表一组表位中的每一个的肽序列的数据;(b)将每个表位的肽序列输入呈递模型以生成每个表位由肿瘤的肿瘤细胞表面上的一个或多个MHC等位基因呈递的一组数值可能性,已经至少基于接收到的质谱数据来鉴定的数值可能性组;和(c)基于数值可能性组来选择表位组的子集以生成一组选定的表位,所述选定的表位用于生成至少20个MHC I类表位编码核酸序列。在一些方面,选定的表位组的数量是2-20。在一些方面,呈递模型代表以下之间的依赖性:(a)在肽序列的特定位置存在一个特定MHC等位基因和特定氨基酸的对;和(b)此类在特定位置处包含特定氨基酸的肽序列由所述对的一个特定MHC等位基因呈递在肿瘤细胞表面上的可能性。在一些方面,选择选定的表位组包括基于呈递模型来选择相对于未选定的表位能够在肿瘤细胞表面上呈递的可能性增加的表位。在一些方面,选择选定的表位组包括基于呈递模型来选择相对于未选定的表位能够在受试者中诱导肿瘤特异性免疫响应的可能性增加的表位。在一些方面,选择选定的表位组包括基于呈递模型来选择相对于未选定的表位能够被专职性抗原呈递细胞(APC)呈递至初始T细胞的可能性增加的表位,任选的其中APC是树突状细胞(DC)。在一些方面,选择选定的表位组包括基于呈递模型来选择

相对于未选定的表位，经由中枢或外周耐受而受到抑制的可能性降低的表位。在一些方面，选择选定的表位组包括基于呈递模型来选择相对于未选定的表位能够在受试者中诱导对正常组织的自身免疫响应的可能性降低的表位。

[0083] 在一些方面，外显子组或转录组核苷酸测序数据通过对肿瘤细胞执行测序而获得。在一些方面，测序是下一代测序(NGS)或任何大规模并行测序方法。

附图说明

[0084] 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，其中：

[0085] 图(FIG.)1呈现了疫苗接种方案的示意图。

[0086] 图2A呈现了使用ELISpot测量的抗原特异性细胞免疫响应。在用六个重叠肽池(每个池有20个肽，长度为15个氨基酸)刺激的PBMC中测量抗原特异性IFN-γ的产生。示出了动物(n＝6)中每个池的平均值。每个样品的本底校正为DMSO对照，第0-4周应用计数转换因子。堆叠条形图从上到下分别示出了池1-6的响应。

[0087] 图2B呈现了使用ELISpot测量的抗原特异性细胞免疫响应。在用六个重叠肽池(每个池有20个肽，长度为15个氨基酸)刺激的PBMC中测量抗原特异性IFN-γ的产生。示出了个体灵长类动物#1-3的响应(分别为从上到下的图)。每个样品的本底校正为DMSO对照，第0-4周应用计数转换因子。堆叠条形图从上到下分别示出了池1-6的响应。

[0088] 图2C呈现了使用ELISpot测量的抗原特异性细胞免疫响应。在用六个重叠肽池(每个池有20个肽，长度为15个氨基酸)刺激的PBMC中测量抗原特异性IFN-γ的产生。示出了个体灵长类动物#4-6的响应(分别为从上到下的图)。每个样品的本底校正为DMSO对照，第0-4周应用计数转换因子。堆叠条形图从上到下分别示出了池1-6的响应。

[0089] 图3A呈现了使用ELISpot测量的GRANITE受试者G1的抗原特异性细胞免疫响应。在用CD8表位池刺激的PBMC中测量了抗原特异性IFN-γ的产生。图表示出了一式三份ELISpot孔的平均斑点形成单位(SFU)/10^6个PBMC+/-标准偏差(SD)。LOD＝测定检测限。ULOQ＝定量上限(5,000SFU/10^6个PBMC)。

[0090] 图3B呈现了使用ELISpot测量的GRANITE受试者G1的抗原特异性细胞免疫响应。在用迷你CD8表位池刺激的PBMC中测量了抗原特异性IFN-γ的产生。图表示出了一式三份ELISpot孔的平均斑点形成单位(SFU)/10^6个PBMC+/-标准偏差(SD)。LOD＝测定检测限。示出了完整时程。

[0091] 图3C呈现了使用ELISpot测量的GRANITE受试者G1的抗原特异性细胞免疫响应。在用迷你CD8表位池刺激的PBMC中测量了抗原特异性IFN-γ的产生。图表示出了一式三份ELISpot孔的平均斑点形成单位(SFU)/10^6个PBMC+/-标准偏差(SD)。LOD＝测定检测限。示出了图3B的第68周和第70周。

[0092] 图4呈现了使用ELISpot测量的GRANITE受试者G3的抗原特异性细胞免疫响应。在用CD8表位池刺激的PBMC中测量了抗原特异性IFN-γ的产生。图表示出了一式三份ELISpot孔的平均斑点形成单位(SFU)/10^6个PBMC+/-标准偏差(SD)。LOD＝测定检测限。

[0093] 图5呈现了使用ELISpot测量的GRANITE受试者G8的抗原特异性细胞免疫响应。在用CD8表位池刺激的PBMC中测量了抗原特异性IFN-γ的产生。图表示出了一式三份ELISpot孔的平均斑点形成单位(SFU)/10^6个PBMC+/-标准偏差(SD)。LOD＝测定检测限。

[0094] 图6呈现了使用ELISpot测量的GRANITE受试者G11的抗原特异性细胞免疫响应。在用CD8表位池刺激的PBMC中测量了抗原特异性IFN-γ的产生。图表示出了一式三份ELISpot孔的平均斑点形成单位(SFU)/10^6个PBMC+/-标准偏差(SD)。LOD＝测定检测限。

[0095] 图7呈现了同源ChAdV初免-加强疫苗策略的免疫细胞表征数据。左图:对于患者G1、G3、G8和G11,示出了在离体IFNg ELISpot中用患者特异性肽池刺激过夜的PBMC的ChAdV加强前后的ELISpot数据。图表示出了一式三份ELISpot孔的平均斑点形成单位(SFU)/10^6个PBMC。中图:使用患者特异性CD8池刺激后,通过MSD U-plex测定来分析ChAdV加强离体IFNg ELISpot测定(G1、G3、G8和G11)前后的上清液中的颗粒酶B(GRZB)的水平。图表示出了重复份孔中以pg/ml为单位的平均GRZB水平(减去本底)。右图:用DMSO或患者特异性CD8池刺激过夜后,通过离体荧光斑点测定来分析患者G11的ChAdV加强时间点前后的PBMC,以测量IFNg、GRZB和IL-2表达。图表示出了针对IFNγ、GRZB和IL-2堆叠的重复份荧光斑点孔的平均SFU/10^6PBMC+/-标准偏差(SD)。

[0096] 图8A示出了四聚体染色和记忆表型分析的流式细胞术数据以评估效应记忆群体。密度图示出了患者G1的的肽-HLA(HLA-B*44:05-PE)ChAd68加强前的识别。流程图示出了CD8四聚体$^+$群中的初始(CD45RA$^+$CCR7$^+$)、中央记忆(CD45RA$^-$CCR7$^+$)、效应记忆(CD45RA$^-$CCR7$^-$)和效应T记忆RA$^+$(CD45RA$^+$CCR7$^-$)细胞的频率。

[0097] 图8B示出了四聚体染色和记忆表型分析的流式细胞术数据以评估效应记忆群体。密度图示出了患者G1的的肽-HLA(HLA-B*44:05-APC)后ChAd68加强的识别。流程图示出了CD8四聚体$^+$群中的初始(CD45RA$^+$CCR7$^+$)、中央记忆(CD45RA$^-$CCR7$^+$)、效应记忆(CD45RA$^-$CCR7$^-$)和效应T记忆RA$^+$(CD45RA$^+$CCR7$^-$)细胞的频率。

[0098] 图8C示出了四聚体染色和记忆表型分析的流式细胞术数据以评估效应记忆群体。密度图示出了患者G1的的肽-HLA(HLA-A*02:01-BV421)前和后ChAd68加强的识别。饼状图示出了来自对应点状图的CD8四聚体$^+$群中的初始(CD45RA$^+$CCR7$^+$)、中央记忆(CD45RA$^-$CCR7$^+$)、效应记忆(CD45RA$^-$CCR7$^-$)和效应T记忆RA$^+$(CD45RA$^+$CCR7$^-$)细胞的频率。

具体实施方式

[0099] I.定义

[0100] 一般来讲,权利要求和说明书中使用的术语旨在被解释为具有本领域普通技术人员所理解的普通含义。下文定义了某些术语以提供额外的清晰性。如果普通含义与所提供的定义产生冲突,则使用所提供的定义。

[0101] 如本文所用,术语“抗原”是刺激免疫响应的物质。抗原可以是新抗原。抗原可以是“共有抗原”,其是在特定人群(例如,癌症患者的特定人群)中发现的抗原。

[0102] 如本文所用,术语“新抗原”是具有至少一种使其与对应的野生型抗原不同(例如经由肿瘤细胞中的突变或对肿瘤细胞具有特异性的翻译后修饰)的改变的抗原。新抗原可以包括多肽序列或核苷酸序列。突变可以包括移码或非移码插入缺失、错义或无义取代、剪接位点改变、基因组重排或基因融合、或产生neoORF的任何基因组或表达改变。突变还可以包括剪接变体。对肿瘤细胞具有特异性的翻译后修饰可以包括异常磷酸化。对肿瘤细胞具有特异性的翻译后修饰还可以包括蛋白酶体生成的剪接抗原。参见Liepe等人,A large fraction of HLA class I ligands are proteasome-generated spliced peptides;

Science.2016 Oct 21;354(6310):354-358。可以通过使用各种诊断方法(例如下文进一步描述的患者选择方法)来鉴定施用的受试者。

[0103] 如本文所用,术语“肿瘤抗原”是这样的抗原,其存在于受试者的肿瘤细胞或组织中但不存在于受试者对应的正常细胞或组织中,或者来源于已知或已发现与正常细胞或组织相比在肿瘤细胞或癌组织中具有表达改变的多肽。

[0104] 如本文所用,术语“基于抗原的疫苗”是基于一种或多种抗原(例如多种抗原)的疫苗组合物。疫苗可以是基于核苷酸的(例如,基于病毒、基于RNA或基于DNA)、基于蛋白质的(例如,基于肽)或其组合。

[0105] 如本文所用,术语“候选抗原”是产生可代表抗原的序列的突变或其他畸变。

[0106] 如本文所用,术语“编码区”是编码蛋白质的基因部分。

[0107] 如本文所用,术语“编码突变”是发生在编码区中的突变。

[0108] 如本文所用,术语“ORF”意指开放阅读框。

[0109] 如本文所用,术语“NEO-ORF”是由突变或其他畸变(诸如剪接)引起的肿瘤特异性ORF。

[0110] 如本文所用,术语“错义突变”是导致从一个氨基酸取代为另一个氨基酸的突变。

[0111] 如本文所用,术语“无义突变”是导致从氨基酸取代为终止密码子或导致去除规范起始密码子的突变。

[0112] 如本文所用,术语“移码突变”是导致蛋白质框架改变的突变。

[0113] 如本文所用,术语“插入缺失”是一个或多个核酸的插入或缺失。

[0114] 如本文所用,在两个或更多个核酸或多肽序列的上下文中,术语“同一性”百分比是指具有指定百分比的当针对最大对应而进行比较和比对时为相同的核苷酸或氨基酸残基的两个或更多个序列或子序列,如使用下文描述的序列比较算法之一(例如,BLASTP和BLASTN或其他技术人员可用的算法)或通过视觉检查来测量。根据应用,“同一性”百分比可以存在于所比较的序列的区域上,例如,存在于功能结构域上,或者可替代地存在于待比较的两个序列的全长上。

[0115] 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,必要时指定子序列坐标,并且指定序列算法程序参数。序列比较算法然后基于指定的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。或者,序列相似性或不相似性可以通过特定核苷酸的组合存在或不存在来确定,或者对于翻译的序列,通过选定序列位置处的氨基酸(例如,序列基序)的组合存在或不存在来确定。

[0116] 例如,用于比较的最佳序列比对可以通过Smith&Waterman,Adv.Appl.Math.2:482(1981)的局部同源性算法;通过Needleman&Wunsch,J.Mol.Biol.48:443(1970)的同源性比对算法;通过Pearson&Lipman,Proc.Nat'l.Acad.Sci.USA85:2444(1988)的相似性搜索方法;通过这些算法的计算机化实施(Wisconsin Genetics Software Package,Genetics Computer Group,575Science Dr.,Madison,Wis.中的GAP、BESTFIT、FASTA和TFASTA);或通过视觉检查(一般参见Ausubel等人,下文)来进行。

[0117] 适用于确定序列同一性和序列相似性百分比的算法的一个实例是Altschul等人,J.Mol.Biol.215:403-410(1990)中描述的BLAST算法。用于执行BLAST分析的软件可通过国

家生物技术信息中心公开获得。

[0118] 如本文所用,术语“非终止或通读”是导致天然终止密码子去除的突变。

[0119] 如本文所用,术语“表位”是通常由抗体或T细胞受体结合的抗原的指定部分。

[0120] 如本文所用,术语“免疫原性”是刺激免疫响应的能力,例如,经由T细胞、B细胞或两者刺激免疫响应的能力。

[0121] 如本文所用,术语“HLA结合亲和力”、“MHC结合亲和力”意指指定抗原与指定MHC等位基因之间结合的亲和力。

[0122] 如本文所用,术语“诱饵”是用于从样品中富集指定DNA或RNA序列的核酸探针。

[0123] 如本文所用,术语“变体”是受试者的核酸与用作对照的参考人类基因组之间的差异。

[0124] 如本文所用,术语“变体调用”是通常来自测序的变体存在的算法确定。

[0125] 如本文所用,术语“多态性”是种系变体,即在个体的所有携带DNA的细胞中发现的变体。

[0126] 如本文所用,术语“体细胞变体”是在个体的非种系细胞中产生的变体。

[0127] 如本文所用,术语“等位基因”是基因的一种形式或遗传序列的一种形式或蛋白质的一种形式。

[0128] 如本文所用,术语“HLA类型”是HLA基因等位基因的补体。

[0129] 如本文所用,术语“无义介导的降解”或“NMD”是由于提前终止密码子引起的细胞对mRNA的降解。

[0130] 如本文所用,术语“躯干突变(truncal mutation)”是起源于肿瘤发展早期并存在于肿瘤细胞的大部分中的突变。

[0131] 如本文所用,术语“亚克隆突变”是起源于肿瘤发展后期并仅存在于肿瘤细胞的子集中的突变。

[0132] 如本文所用,术语“外显子组”是编码蛋白质的基因组的子集。外显子组可以是基因组的集体外显子。

[0133] 如本文所用,术语“逻辑回归”是来自统计的二进制数据的回归模型,其中将因变量等于一的概率的对数建模为因变量的线性函数。

[0134] 如本文所用,术语“神经网络”是一种用于分类或回归的机器学习模型,由多层线性变换组成,随后是通常经由随机梯度下降和反向传播而训练的逐元素非线性。

[0135] 如本文所用,术语“蛋白质组”是由细胞、细胞群或个体表达和/或翻译的所有蛋白质的集合。

[0136] 如本文所用,术语“肽组”是细胞表面上由MHC-I或MHC-II呈递的所有肽的集合。肽组可指代细胞或细胞集合的性质(例如,肿瘤肽组,其意指包含肿瘤的所有细胞的肽组的联合,或感染性疾病肽组,其意指被感染性疾病感染的所有细胞的肽组的联合)。

[0137] 如本文所用,术语“ELISPOT”意指酶联免疫吸附斑点测定——这是监测人类和动物免疫响应的常用方法。

[0138] 如本文所用,术语“dextramer”是基于葡聚糖的肽-MHC多聚体,其用于流式细胞术中的抗原特异性T细胞染色。

[0139] 如本文所用,术语“耐受性或免疫耐受性”是对一种或多种抗原(例如自身抗原)的

免疫无响应状态。

[0140] 如本文所用,术语“中枢耐受”是通过缺失自身反应性T细胞克隆或通过促进自身反应性T细胞克隆分化为免疫抑制性调控性T细胞(Treg),从而在胸腺中受到影响的耐受。

[0141] 如本文所用,术语“外周耐受”是通过下调或使在中枢耐受中存活的自身反应性T细胞无应答或促进这些T细胞分化成Treg,从而在外周受到影响的耐受。

[0142] 术语“样品”可以包括通过包括静脉穿刺、排泄、射精、按摩、活检、针吸、灌洗样品、刮片、外科切口、或介入或本领域已知的其他方式的方式从受试者获取的单个细胞或多个细胞或细胞碎片或体液等分试样。

[0143] 术语“受试者”涵盖细胞、组织或生物体、人或非人,无论是体内、离体还是体外,雄性还是雌性。术语受试者包括哺乳动物,包括人类。

[0144] 术语“哺乳动物”涵盖人和非人两种,并且包括但不限于人、非人灵长类动物、犬科动物、猫科动物、鼠科动物、牛科动物、马科动物和猪科动物。

[0145] 术语“临床因素”是指受试者的状况的量度,例如疾病活性或严重程度。“临床因素”涵盖受试者健康状况的所有标志物,包括非样品标志物和/或受试者的其他特征,诸如但不限于年龄和性别。临床因素可以是可在确定的条件下从受试者样品(或样品群)或受试者的评价中获得的评分、值或一组值。临床因素也可以通过标志物和/或其他参数(诸如基因表达替代物)来预测。临床因素可以包括肿瘤类型、肿瘤亚型、感染类型、感染亚型和吸烟史。

[0146] 术语“来源于肿瘤的抗原编码核酸序列”是指从肿瘤获得(例如经由RT-PCR)的核酸序列;或通过对肿瘤进行测序然后使用测序数据合成核酸序列(例如经由本领域已知的各种合成或基于PCR的方法)获得的序列数据。衍生的序列可以包括核酸序列变体,诸如序列优化的核酸序列变体(例如,针对表达而密码子优化的和/或以其他方式优化的),其编码与获自肿瘤的对应天然核酸序列相同的多肽序列。

[0147] 术语“来源于感染的抗原编码核酸序列”是指从感染细胞或感染疾病生物体获得(例如经由RT-PCR)的核酸序列;或通过对感染细胞或感染疾病生物体进行测序然后使用测序数据合成核酸序列(例如经由本领域已知的各种合成或基于PCR的方法)获得的序列数据。衍生的序列可以包括核酸序列变体,诸如序列优化的核酸序列变体(例如,针对表达而密码子优化的和/或以其他方式优化的),其编码与对应的天然感染性疾病生物体核酸序列相同的多肽序列。衍生的序列可以包括核酸序列变体,其编码相对于天然感染性疾病生物体核酸序列具有一个或多个(例如,1、2、3、4或5个)突变的修饰的感染性疾病生物体多肽序列。例如,修饰的多肽序列相对于感染性疾病生物体蛋白质的天然多肽序列可以具有一个或多个错义突变。

[0148] 术语“甲病毒”是指披膜病毒(Togaviridae)科的成员,并且是正义单链RNA病毒。甲病毒通常分类为旧世界病毒(Old World),诸如辛德比斯病毒、罗斯河病毒、马亚罗病毒、基孔肯雅热病毒(Chikungunya)和森林脑炎病毒,或新世界病毒(New World),诸如东方马脑炎病毒(eastern equine encephalitis)、Aura病毒、摩根堡病毒或委内瑞拉马脑炎病毒及其衍生毒株TC-83。甲病毒通常是自复制的RNA病毒。

[0149] 术语“甲病毒骨架”是指允许病毒基因组自复制的甲病毒的最小序列。最小序列可以包括用于非结构蛋白介导的扩增的保守序列、非结构蛋白1(nsP1)基因、nsP2基因、nsP3

基因、nsP4基因和聚A序列，以及用于表达亚基因组病毒RNA(包括亚基因组(例如26S)启动子元件)的序列。

[0150] 术语“用于非结构蛋白介导的扩增的序列”包括本领域技术人员众所周知的甲病毒属保守序列元件(CSE)。CSE包括但不限于甲病毒5'UTR、51-nt CSE、24-nt CSE、亚基因组启动子序列(例如26S亚基因组启动子序列)、19-nt CSE和甲病毒3'非编码区。

[0151] 术语“RNA聚合酶”包括催化从DNA模板产生RNA多核苷酸的聚合酶。RNA聚合酶包括但不限于噬菌体来源的聚合酶，包括T3、T7和SP6。

[0152] 术语“脂质”包括疏水性和/或两亲性分子。脂质可以是阳离子的、阴离子的或中性的。脂质可以是合成的或天然来源的，并且在一些情况下是可生物降解的。脂质可以包括胆固醇、磷脂、脂质缀合物，包括但不限于聚乙二醇(PEG)缀合物(PEG化脂质)、蜡、油、甘油酯、脂肪和脂溶性维生素。脂质还可以包括二亚油醇基甲基-4-二甲基氨基丁酸酯(MC3)和MC3样分子。

[0153] 术语“脂质纳米颗粒”或“LNP”包括使用围绕水性内部的含脂质膜形成的囊泡样结构，也称为脂质体。脂质纳米颗粒包括基于脂质的组合物，其具有由表面活性剂稳定的固体脂质核心。核心脂质可以是脂肪酸、酰基甘油、蜡和这些表面活性剂的混合物。诸如磷脂、鞘磷脂、胆汁盐(牛磺胆酸钠)和固醇类(胆固醇)的生物膜脂可用作稳定剂。脂质纳米颗粒可以使用限定比例的不同脂质分子形成，包括但不限于限定比例的一种或多种阳离子、阴离子或中性脂质。脂质纳米颗粒可以将分子封装在外膜壳内，并且随后可以与靶细胞接触以将封装的分子递送至宿主细胞胞质溶胶。脂质纳米颗粒可以用非脂质分子进行修饰或功能化，包括在它们的表面上修饰或功能化。脂质纳米颗粒可以是单层的(单层)或多层的(多层)。脂质纳米颗粒可以与核酸复合。单层脂质纳米颗粒可以与核酸复合，其中核酸在水性内部。多层脂质纳米颗粒可与核酸复合，其中核酸在水性内部，或形成或夹在其中。

[0154] 缩写：MHC：主要组织相容性复合物；HLA：人类白细胞抗原，或人类MHC基因座；NGS：下一代测序；PPV：阳性预测值；TSNA：肿瘤特异性新抗原；FFPE：福尔马林固定的、石蜡包埋的；NMD：无义介导的降解；NSCLC：非小细胞肺癌；DC：树突状细胞。

[0155] 应当注意，除非上下文另外明确规定，否则如说明书和随附权利要求中所用，单数形式“一个/种(a/an)”和“所述(the)”包括复数个指示物。

[0156] 除非特别说明或从上下文中明显看出，否则如本文所用，术语“约”理解为在本领域的正常公差范围内，例如在平均值的2个标准偏差内。约可以理解为在所述值的10％、9％、8％、7％、6％、5％、4％、3％、2％、1％、0.5％、0.1％、0.05％或0.01％范围内。除非上下文另有明确规定，否则本文提供的所有数值均由术语约修饰。

[0157] 本文未直接定义的任何术语应理解为具有本发明领域内所理解的通常与其关联的含义。某些术语在本文中讨论以在描述本发明方面的组合物、装置、方法等以及如何制造或使用它们方面向从业者提供额外的指导。应当理解，可以用多于一种方式来表述同一事物。因此，替代语言和同义词可用于本文讨论的术语中的任何一个或多个。术语是否在本文中阐述或讨论并不重要。提供了一些同义词或可替代的方法、材料等。除非明确说明，否则列举一个或一些同义词或等同物不排除使用其他同义词或等同物。实例(包括术语实例)的使用仅出于说明性目的，并且不限制本文中本发明各方面的范围和含义。

[0158] 在本说明书的主体内引用的所有参考文献、授权专利和专利申请都特此出于所有

目的通过引用整体并入。

[0159] II.抗原鉴定

[0160] 肿瘤和正常外显子组和转录组的NGS分析的研究方法已在抗原鉴定领域描述和应用。[6,14,15]可以考虑进行某些优化，以提高临床环境中抗原鉴定的敏感度和特异性。这些优化可以分为两个领域，与实验室流程相关的领域和与NGS数据分析相关的领域。所描述的研究方法也可以用于在其他环境中鉴定抗原，例如鉴定来自感染性疾病生物体、受试者感染或受试者感染细胞的抗原。优化的实例是本领域技术人员已知的，例如在美国专利号10,055,540、美国申请公开号US20200010849A1、美国申请号16/606,577和国际专利申请公开WO2020181240A1、WO/2018/195357和WO/2018/208856中更详细描述的方法，所述专利各自出于所有目的通过引用整体并入本文。

[0161] 用于鉴定抗原(例如，来源于肿瘤或感染性疾病生物体的抗原)的方法包括鉴定可能在细胞表面上呈递(例如，在肿瘤细胞、感染细胞或免疫细胞(包括专职性抗原呈递细胞，诸如树突状细胞)上被MHC呈递)和/或可能具有免疫原性的抗原。例如，一种这样的方法包括以下步骤：从肿瘤、感染细胞或感染疾病生物体中获得外显子组、转录组或全基因组核苷酸测序和/或表达数据中的至少一种，其中核苷酸测序和/或表达数据用于获得代表一组抗原(例如，来源于肿瘤或感染性疾病生物体的抗原)中的每一个的肽序列的数据；将每个抗原的肽序列输入一个或多个呈递模型以生成每个抗原由细胞表面(诸如的受试者的肿瘤细胞或感染细胞)上的一个或多个MHC等位基因呈递的一组数值可能性，已经至少基于接收到的质谱数据来鉴定的数值可能性组；和基于数值可能性组来选择抗原组的子集以生成一组选定的抗原。

[0162] III.新抗原中肿瘤特异性突变的鉴定

[0163] 本文还公开了用于鉴定某些突变(例如，存在于癌细胞中的变体或等位基因)的方法。特别地，这些突变可以存在于患有癌症的受试者的癌细胞的基因组、转录组、蛋白质组或外显子组中，但不存在于受试者的正常组织中。用于鉴定对肿瘤具有特异性的新抗原(包括共有新抗原)的具体方法是本领域技术人员已知的，例如在美国专利号10,055,540、美国申请公开号US20200010849A1和国际专利申请公开号WO/2018/195357和WO/2018/208856中更详细描述的方法，所述专利各自出于所有目的通过引用整体并入本文。国际专利申请公开WO2019226941A1中更详细地描述了对肿瘤具有特异性的共有新抗原的实例，所述专利出于所有目的通过引用整体并入本文。共享新抗原包括但不限于KRAS相关突变(例如KRAS G12C、KRAS G12V、KRAS G12D和/或KRAS Q61H突变)。例如，KRAS相关MHC I类新表位可以包括参考野生型(WT)人类KRAS的那些突变，诸如参考一下示例性氨基酸序列：MTEYKLVVVGAGGVGKSALTIQLIQNHF VDEYDPTIEDSYRKQVVIDGETCLLDILDTAGQEEYSAMRDQYMRTGEGFLCVFAINNTKSFEDIHHYREQIKRVKDSEDVPMVLVGNKCDLPSRTVDTKQAQDLARSYGIPFIETSAKTRQRVEDAFYTLVREIRQYRLKKISKEEKTPGCVKIKKCIIM。

[0164] 如果肿瘤中的基因突变导致仅在肿瘤中的蛋白质的氨基酸序列发生变化，则它们可以被认为可用于肿瘤的免疫靶向。可用的突变包括：(1)导致蛋白质中的不同氨基酸的非同义突变；(2)通读突变，其中终止密码子被修饰或缺失，从而导致在C末端具有新的肿瘤特异性序列的较长蛋白质的翻译；(3)剪接位点突变，其导致成熟mRNA中包括内含子，从而导致独特的肿瘤特异性蛋白质序列；(4)染色体重排，其在2种蛋白质的连接处产生具有肿瘤

特异性序列的嵌合蛋白(即,基因融合);(5)移码突变或缺失,其导致具有新的肿瘤特异性蛋白质序列的新开放阅读框。突变还可以包括非移码插入缺失、错义或无义取代、剪接位点改变、基因组重排或基因融合、或产生neoORF的任何基因组或表达改变中的一种或多种。

[0165] 具有突变的肽或由例如肿瘤细胞中的剪接位点、移码、通读或基因融合突变引起的突变多肽可以通过对肿瘤细胞与正常细胞中的DNA、RNA或蛋白质进行测序来鉴定。

[0166] 突变还可以包括先前鉴定的肿瘤特异性突变。已知的肿瘤突变可见于癌症体细胞突变目录(COSMIC)数据库。

[0167] 多种方法可用于检测个体DNA或RNA中特定突变或等位基因的存在。此领域的进步提供了准确、简单且廉价的大规模SNP基因分型。例如,已经描述了若干种技术,包括动态等位基因特异性杂交(DASH)、微孔板阵列对角线凝胶电泳(MADGE)、焦磷酸测序、寡核苷酸特异性连接、TaqMan系统以及各种DNA“芯片”技术,诸如Affymetrix SNP芯片。这些方法利用通常通过PCR的目标遗传区域的扩增。还有其他方法,基于通过侵入性切割生成小信号分子,然后进行质谱法或固定挂锁探针和滚环扩增。下文总结了本领域已知的用于检测指定突变的若干种方法。

[0168] 基于PCR的检测方式可以包括同时多重扩增多个标志物。例如,本领域众所周知选择PCR引物以产生大小不重叠并且可以同时分析的PCR产物。或者,可以用差异标记并因此可以各自差异检测的引物来扩增不同的标志物。当然,基于杂交的检测方式允许对样品中的多个PCR产物进行差异检测。本领域已知的其他技术允许对多个标志物进行多重分析。

[0169] 已经开发了若干种方法来促进基因组DNA或细胞RNA中的单核苷酸多态性的分析。例如,单碱基多态性可以通过使用专门的抗核酸外切酶核苷酸来检测,如Mundy,C.R.(美国专利号4,656,127)中所公开的。根据所述方法,允许与紧邻多态性位点3’的等位基因序列互补的引物与获自特定动物或人的靶分子杂交。如果靶分子上的多态性位点含有与存在的特定抗外切核酸酶核苷酸衍生物互补的核苷酸,则该衍生物将掺入到杂交引物的末端上。这种掺入使引物对核酸外切酶具有抗性,从而允许对其进行检测。由于样品的抗核酸外切酶衍生物的身份是已知的,因此引物对核酸外切酶具有抗性的发现揭示了靶分子的多态性位点中存在的核苷酸与反应中使用的核苷酸衍生物互补。这种方法的优点是不需要测定大量无关的序列数据。

[0170] 基于溶液的方法可以用于确定多态性位点的核苷酸的身份。Cohen,D.等人(法国专利2,650,840;PCT申请号WO91/02087)。如美国专利号4,656,127的Mundy方法中那样,采用与紧邻多态性位点3’的等位基因序列互补的引物。所述方法使用标记的双脱氧核苷酸衍生物来确定该位点的核苷酸的身份,如果与多态性位点的核苷酸互补,则将所述双脱氧核苷酸衍生物掺入到引物的末端上。

[0171] Goelet,P.等人(PCT申请号92/15712)描述了被称为遗传位分析(Genetic Bit Analysis)或GBA的替代方法。Goelet,P.等人的方法使用标记的终止子和与多态性位点的3’序列互补的引物的混合物。因此,掺入的标记的终止子由存在于被评价的靶分子的多态性位点中的核苷酸决定并且与其互补。与Cohen等人(法国专利2,650,840;PCT申请号WO91/02087)的方法相反,Goelet,P.等人的方法可以是异相测定,其中将引物或靶分子固定至固相。

[0172] 已经描述了若干种用于测定DNA中的多态性位点的引物指导的核苷酸掺入程序

(Komher,J.S.等人,Nucl.Acids.Res.17:7779-7784(1989);Sokolov,B.P.,Nucl.Acids Res.18:3671(1990);Syvanen,A.-C.等人,Genomics 8:684-692(1990);Kuppuswamy,M.N.等人,Proc.Natl.Acad.Sci.(U.S.A.)88:1143-1147(1991);Prezant,T.R.等人,Hum.Mutat.1:159-164(1992);Ugozzoli,L.等人,GATA 9:107-112(1992);Nyren,P.等人,Anal.Biochem.208:171-175(1993))。这些方法与GBA的差异在于它们利用标记的脱氧核苷酸的掺入来区分多态性位点处的碱基。在这种形式中,由于信号与掺入的脱氧核苷酸数量成正比,同一核苷酸运行中出现的多态性可导致与运行长度成正比的信号(Syvanen,A.-C.等人,Amer.J.Hum.Genet.52:46-59(1993))。

[0173] 许多计划直接从数百万个平行的DNA或RNA分子中获取序列信息。实时单分子合成测序技术依赖于荧光核苷酸的检测,因为它们掺入到与正在测序的模板互补的新生DNA链中。在一个方法中,长度为30-50个碱基的寡核苷酸共价锚定在玻璃盖玻片的5'端。这些锚定链执行两个功能。第一,如果模板配置有与表面结合寡核苷酸互补的捕获尾,那么锚定链充当靶模板链的捕获位点。它们还充当形成序列阅读基础的模板定向引物延伸的引物。捕获引物用作序列测定的固定位置位点,所述序列测定使用染料接头的合成、检测和化学切割的多个循环以去除染料。每个循环包括添加聚合酶/标记的核苷酸混合物、漂洗、成像和染料切割。在替代方法中,聚合酶用荧光供体分子修饰并固定在载玻片上,而每个核苷酸都用附接至γ-磷酸盐的受体荧光部分进行颜色编码。系统检测荧光标记的聚合酶和荧光修饰的核苷酸之间的相互作用,因为核苷酸掺入到从头链中。还存在其他合成测序技术。

[0174] 可以使用任何合适的合成测序平台来鉴定突变。如上所述,目前有四种主要的合成测序平台可用:Roche/454Life Sciences的基因组测序仪、Illumina/Solexa的1G分析仪、Applied BioSystems的SOLiD系统和Helicos Biosciences的Heliscope系统。Pacific BioSciences和VisiGen Biotechnologies也描述了合成测序平台。在一些实施方案中,多个被测序的核酸分子与支持物(例如,固体支持物)结合。为了将核酸固定在支持物上,可以在模板的3'和/或5'端添加捕获序列/通用引发位点。可以通过将捕获序列与共价附接至支持物的互补序列杂交来将核酸结合至支持物。捕获序列(也称为通用捕获序列)是与附接至支持物的序列互补的可双重用作通用引物的核酸序列。

[0175] 作为捕获序列的替代方案,偶联对(诸如,例如抗体/抗原、受体/配体或如例如美国专利申请号2006/0252077中所述的亲和素-生物素对)的成员可以连接至每个片段以被捕获在涂有该偶联对的相应第二成员的表面上。

[0176] 在捕获之后,可以分析序列,例如通过单分子检测/测序,例如,如实施例和美国专利号7,283,337所述,包括模板依赖性合成测序。在合成测序中,表面结合分子在聚合酶的存在下暴露于多个标记的三磷酸核苷酸。模板的序列由掺入到生长链3'端中的标记的核苷酸的顺序决定。这可以实时完成,或可以以分步重复模式完成。对于实时分析,每个核苷酸可以掺入不同的光学标记,并且可以利用多个激光来刺激掺入的核苷酸。

[0177] 测序还可以包括其他大规模并行测序或下一代测序(NGS)技术和平台。大规模并行测序技术和平台的其他实例是Illumina HiSeq或MiSeq、Thermo PGM或Proton、Pac Bio RS II或Sequel、Qiagen's Gene Reader和Oxford Nanopore MinION。可以使用其他类似的当前大规模并行测序技术,以及这些技术的未来几代。

[0178] 可以利用任何细胞类型或组织来获得用于本文所述方法的核酸样品。例如,DNA或

RNA样品可以通过已知技术(例如静脉穿刺)从肿瘤或体液(例如血液)获得,或者从唾液获得。或者,可以对干燥样品(例如头发或皮肤)执行核酸测试。此外,可以从肿瘤中获得样品用于测序,并且可以从正常组织中获得另一个样品用于测序,其中正常组织与肿瘤的组织类型相同。可以从肿瘤中获得样品用于测序,并且可以从正常组织中获得另一个样品用于测序,其中正常组织相对于肿瘤具有不同的组织类型。

[0179] 肿瘤可以包括以下中的一种或多种:肺癌、黑色素瘤、乳腺癌、卵巢癌、前列腺癌、肾癌、胃癌、结肠癌、睾丸癌、头颈癌、胰腺癌、脑癌、B细胞淋巴瘤、急性骨髓性白血病、慢性骨髓性白血病、慢性淋巴细胞白血病、T细胞淋巴细胞白血病、非小细胞肺癌和小细胞肺癌。

[0180] 或者,可以使用蛋白质质谱法来鉴定或验证与肿瘤细胞上的MHC蛋白结合的突变肽的存在。肽可以从肿瘤细胞或从肿瘤中免疫沉淀的HLA分子中酸洗脱,然后使用质谱法进行鉴定。

[0181] IV.抗原

[0182] 抗原可以包括核苷酸或多肽。例如,抗原可以是编码多肽序列的RNA序列。因此,可用于疫苗的抗原可以包括核苷酸序列或多肽序列。

[0183] 本文公开了包含通过本文公开的方法鉴定的肿瘤特异性突变的分离的肽、包含已知的肿瘤特异性突变的肽、以及通过本文公开的方法鉴定的突变多肽或其片段。新抗原肽可以在其编码序列的上下文中描述,其中新抗原包括编码相关多肽序列的核苷酸序列(例如,DNA或RNA)。

[0184] 本文还公开了来源于已知或已发现与正常细胞或组织相比在肿瘤细胞或癌组织中具有改变表达的任何多肽(例如已知或已发现与正常细胞或组织相比在肿瘤细胞或癌组织中异常表达的任何多肽)的肽。可以在例如COSMIC数据库中找到可以来源于抗原肽的合适的多肽。COSMIC整理了关于人类癌症中的体细胞突变的综合信息。所述肽含有肿瘤特异性突变。肿瘤抗原(例如共有肿瘤抗原和肿瘤新抗原)可以包括但不限于美国申请号17/058,128中描述的那些,所述申请出于所有目的通过引用并入本文。抗原肽可以在其编码序列的上下文中描述,其中抗原包括编码相关多肽序列的核苷酸序列(例如,DNA或RNA)。

[0185] 本文还公开了来源于与感染性疾病生物体、受试者的感染或受试者的感染细胞相关的任何多肽的肽。抗原可以来源于感染性疾病生物体的核苷酸序列或多肽序列。感染性疾病生物体的多肽序列包括但不限于病原体来源肽、病毒来源肽、细菌来源肽、真菌来源肽和/或寄生虫来源肽。感染性疾病生物体包括但不限于严重急性呼吸系统综合征相关冠状病毒(SARS)、严重急性呼吸系统综合征冠状病毒2(SARS-CoV-2)、埃博拉病毒(Ebola)、HIV、乙型肝炎病毒(HBV)、流感、丙型肝炎病毒(HCV)、人乳头瘤病毒(HPV)、巨细胞病毒(CMV)、基孔肯雅热病毒、呼吸道合胞病毒(RSV)、登革热病毒(Dengue virus)、正粘病毒科(orthymyxoviridae)病毒和结核病。

[0186] 本文公开了包含通过本文公开的方法鉴定的感染性疾病生物体特异性抗原或表位的分离的肽、包含已知的感染性疾病生物体特异性抗原或表位的肽、以及通过本文公开的方法鉴定的突变多肽或其片段。抗原肽可以在其编码序列的上下文中描述,其中抗原包括编码相关多肽序列的核苷酸序列(例如,DNA或RNA)。

[0187] 本文所述的载体和相关组合物可以用于递送来自任何生物体的抗原,包括它们的毒素或其他副产物,以预防和/或治疗感染或与所述生物体或其副产物相关的其他不良反

应。

[0188] 可以掺入到疫苗中(例如,在盒中编码)的抗原包括免疫原,其可用于使人或非人动物免疫以抵抗病毒,诸如感染人和非人脊椎动物的致病病毒。抗原可选自多种病毒科。期望针对其产生免疫响应的期望病毒科的实例包括,小核糖核酸病毒科,其包括鼻病毒属(rhinovirus),约50%的普通感冒病例由其引起;肠道病毒属(enterovirus),包括脊髓灰质炎病毒(poliovirus)、柯萨奇病毒(coxsackievirus)、艾柯病毒(ec hovirus)和人类肠道病毒,诸如甲型肝炎病毒;以及主要在非人类动物中引起口蹄疫的口蹄疫病毒属(apthovirus)。在小核糖核酸病毒科中,靶抗原包括VP1、VP2、VP3、VP4和VPG。另一个病毒科包括杯状病毒科(calcivirus),它包括诺瓦克病毒群(Norwalk group of viru s),是流行性胃肠炎的重要病原体。又另一个期望用于靶向抗原以刺激人类和非人类动物免疫响应的病毒科是披膜病毒(togavirus)科,其包括甲病毒属,甲病毒属包括辛德比斯病毒、罗斯河病毒和委内瑞拉、东部和西部马脑炎,以及风疹病毒(rubivirus),包括风疹病毒(Rubella virus)。黄病毒科(Flaviviridae)包括登革热、黄热病(yellow fever)、日本脑炎(apanese encephalitis)、圣路易斯脑炎(St.Louis encephalitis)和蜱传脑炎病毒(tick borne encephalitis virus)。其他靶抗原可能由丙型肝炎或冠状病毒科生成,其包括许多非人类病毒,诸如传染性支气管炎病毒(家禽)、猪传染性胃肠病毒(猪)、猪血凝性脑脊髓炎病毒(猪)、猫科动物传染性腹膜炎病毒(猫)、猫科动物肠道冠状病毒(猫)、犬科动物冠状病毒(狗)和人类呼吸道冠状病毒,它们可能导致普通感冒和/或非甲、乙或丙型肝炎。在冠状病毒科中,靶抗原包括E1(也称为M或基质蛋白)、E2(也称为S或刺突蛋白)、E3(也称为HE或血凝素-elterose)糖蛋白(不是所有的冠状病毒都存在)或N(核衣壳)。仍有其他抗原可靶向针对弹状病毒科(rhabdovirus),其包括水泡病毒属(vesi culovirus)(例如,水泡性口炎病毒(Vesicular Stomatitis Virus))和一般的狂犬病病毒属(lyssavirus)(例如,狂犬病病毒(rabies))。在弹状病毒科内,合适的抗原可来源于G蛋白或N蛋白。丝状病毒科(filovirida e)包括出血热病毒(hemorrhagic fever virus),诸如马尔堡病毒(Marbur g virus)和埃博拉病毒(Ebola virus),可能是合适的抗原来源。副粘病毒科(paramyxovirus)包括副流感病毒1型、副流感病毒3型、牛副流感病毒3型、风疹病毒(rubulavirus)(腮腺炎病毒(mumps virus))、副流感病毒2型、副流感病毒4型、新城疫病毒(Newcastle disease virus)(鸡)、牛瘟(rinderpest)、麻疹病毒(morbillivirus)(包括麻疹(measles))和犬瘟热(canine distemper)和肺病毒(pneumovirus),其包括呼吸道合胞病毒(例如,糖(G)蛋白和融合(F)蛋白,其序列可从GenBank获得)。流感病毒分类在正粘病毒科中,并且可以是抗原(例如HA蛋白、N1蛋白)的合适来源。布尼亚病毒科(bunyavirus)包括布尼亚病毒属(加利福尼亚脑炎(California encephalitis),La Crosse)、静脉病毒属(phlebov irus)(裂谷热(Rift Valley Fever))、汉坦病毒属(hantavirus)(puremala是一种血红蛋白热病毒)、内罗病毒属(nairovirus)(内罗毕绵羊病(Nair obi sheep disease))和各种未指定的真菌病毒属。沙粒病毒科(arenavir us)提供了针对LCM和拉沙热病毒(Lassa fever virus)的抗原来源。呼肠孤病毒科(reovirus)包括呼肠孤病毒属、轮状病毒属(rotavirus)(引起儿童急性胃肠炎)、环状病毒属(orbiviruses)和科罗病毒属(cultivirus)(科罗拉多壁虱热、Lebombo(人类)、马脑病、蓝舌病)。逆转录病毒科(retrovirus)包括致癌病毒亚科(sub-family oncorivirinal),它涵盖诸如猫白血病病毒

的人类和兽医疾病、HTLVI和HTLVII、慢病毒亚科(包括人类免疫缺陷病毒(HIV)、猿猴免疫缺陷病毒(SIV)、猫免疫缺陷病毒(FIV)、马传染性贫血病毒和泡沫病毒亚科(spumavirinal))。在慢病毒中,已经描述了许多合适的抗原并且它们可以很容易地选择。合适的HIV和SIV抗原的实例包括但不限于gag、pol、Vif、Vpx、VPR、Env、Tat、Nef和Rev蛋白、以及其各种片段。例如,Env蛋白的合适片段可包括其任何亚基,诸如gp120、gp160、gp41或其更小的片段,例如长度至少约8个氨基酸。类似地,可选择tat蛋白的片段。[参见,美国专利号5,891,994和美国专利号6,193,981]。另见D.H.Bar ouch等人,J.Virol.,75(5):2462-2467(March 2001)和R.R.Amara等人,Science,292:69-74(6Apr.2001)中描述的HIV和SIV蛋白。在另一实例中,HIV和/或SIV免疫原性蛋白或肽可用于形成融合蛋白或其他免疫原性分子。参见,例如,2001年8月2日公布的WO 01/54719和1999年4月8日公布的99/16884中描述的HIV-1Tat和/或Nef融合蛋白和免疫方案。本发明不限于本文所述的HIV和/或SI V免疫原性蛋白或肽。此外,对这些蛋白质的多种修饰已经被描述或者可以由本领域技术人员容易地进行。参见,例如,描述于美国专利号5,972,596中的修饰的gag蛋白。此外,可以单独或组合递送任何期望的HIV和/或SIV免疫原。此类组合可包括来自单个载体或来自多个载体的表达。乳多空病毒病毒科(papovavirus)包括多瘤病毒亚科(polyomavirus)(BKU和JCU病毒)和乳头瘤病毒亚科(papillomavirus)(与癌症或乳头状瘤的恶性进展相关)。腺病毒科包括引起呼吸道疾病和/或肠炎的病毒(EX、AD7、ARD、O.B.)。细小病毒科(parvovirus)猫细小病毒(猫肠炎)、猫泛白细胞减少症病毒(panleucopeniavirus)、犬细小病毒和猪细小病毒。疱疹病毒科(herpesvirus)包括甲型疱疹病毒亚科(alphaherpesvirinae),其涵盖单纯疱疹病毒属(simplexvirus)(HSV I、HSVII)、水痘病毒(varicellovirus)(伪狂犬病(pseudorabies)、水痘带状疱疹病毒(varicella zoster)),和乙型疱疹病毒亚科(betaherpesvirina e),其涵盖巨细胞病毒属(cytomegalovirus)(人CMV、鼠巨细胞病毒属(muromegalovirus)),和丙型疱疹病毒亚科(gammaherpesvirinae),其包括淋巴潜隐病毒属(lymphocryptovirus)、EBV(伯基特氏淋巴瘤(Burkitt s lymphoma))、传染性鼻气管炎、马立克氏病(Marek's disease)病毒和鼻病毒属(rhadinovirus)。痘病毒科包括脊索痘病毒亚科(chordopoxyiri nae),其涵盖正痘病毒属(orthopoxvirus)(天花(Variola)(天花(Smallpox))和牛痘(Vaccinia)(牛痘(Cowpox)))、副痘病毒(parapoxvirus)、禽痘病毒(avipoxvirus)、羊痘病毒(capripoxvirus)、兔痘病毒(leporipoxvirus)、猪痘病毒(suipoxvirus)和昆虫痘病毒亚科(entomopoxyirinae)。肝炎病毒科包括乙型肝炎病毒。一种可能是合适的抗原来源的未分类病毒是丁型肝炎病毒(Hepatitis delta virus)。仍其他的病毒来源可包括禽感染性法氏囊病病毒(avian infectious bursal disease virus)和猪呼吸和繁殖综合征病毒。甲病毒科包括马动脉炎病毒(equine arteritis virus)和各种脑炎病毒。

[0189] 可以掺入到疫苗中(例如,在盒中编码)的抗原还包括免疫原,其可用于使人或非人动物免疫以抵抗病原体,包括感染人和非人脊椎动物的细菌、真菌、寄生微生物或多细胞寄生虫。细菌病原体的实例包括致病性革兰氏阳性球菌,包括肺炎球菌(pneumococci);葡萄球菌(staphylococci);和链球菌(streptococci)。致病性革兰氏阴性球菌包括脑膜炎球菌(meningococcus);淋球菌(gonococcus)。致病性肠道革兰氏阴性杆菌包括肠杆菌科(enterobacteriaceae);假单胞菌属(pseudomona s)、不动杆菌属(acinetobacteria)和艾

肯氏菌属(eikenella);类鼻疽(mel ioidosis);沙门氏菌(salmonella);志贺氏菌(shigella);嗜血杆菌(haem ophilus)(流感嗜血杆菌(Haemophilus influenzae)、睡眠嗜血杆菌(Hae mophilus somnus));莫拉氏菌(moraxella);杜氏嗜血杆菌(H.ducreyi)(其导致软下疳);布鲁氏菌(brucella);土拉弗朗西斯菌(Franisella tul arensis)(其导致兔热病);耶尔森氏菌(yersinia)(巴斯德菌(pasteurella));念珠状链杆菌(streptobacillus moniliformis)和螺菌属(spirillum)。革兰氏阳性杆菌包括单核细胞增生李斯特菌(listeria monocytogenes);红斑丹毒丝菌(erysipelothrix rhusiopathiae);白喉棒状杆菌(Corynebacteri um diphtheria)(白喉);霍乱;炭疽杆菌(B.anthracis)(炭疽);腹股沟肉芽肿(donovanosis)(腹股沟肉芽肿(granuloma inguinale));和巴尔通体病(bartonellosis)。由致病性厌氧菌引起的疾病包括破伤风(tetanus);肉毒中毒(botulism);其他梭菌病(clostridia);结核病(tuberculosis);麻风病(leprosy);和其他分枝杆菌(mycobacteria)。指定细菌物种的实例是但不限于肺炎链球菌(Streptococcus pneumoniae)、化脓性链球菌(St reptococcus pyogenes)、无乳链球菌(Streptococcus agalactiae)、粪链球菌(Streptococcus faecalis)、卡他莫拉氏菌(Moraxella catarrhalis)、幽门螺杆菌(Helicobacter pylori)、脑膜炎奈瑟氏菌(Neisseria meningi tidis)、淋病奈瑟氏菌(Neisseria gonorrhoeae)、沙眼衣原体(Chlamydi a trachomatis)、肺炎衣原体(Chlamydia pneumoniae)、鹦鹉热衣原体(Chlamydia psittaci)、百日咳杆菌(Bordetella pertussis)、伤寒沙门氏菌(Salmonella typhi)、鼠伤寒沙门氏菌(Salmonella typhimurium)、猪霍乱沙门氏菌(Salmonella choleraesuis)、大肠杆菌(Escherichia coli)、志贺氏菌(Shigella)、霍乱弧菌(Vibrio cholerae)、白喉棒状杆菌(Corynebac terium diphtheriae)、结核分枝杆菌(Mycobacterium tuberculosis)、鸟分枝杆菌(Mycobacterium avium)、胞内分枝杆菌(Mycobacterium intr acellulare complex)复合物、奇异变形杆菌(Proteus mirabilis)、普通变形杆菌(Proteus vulgaris)、金黄色葡萄球菌(Staphylococcus aureus)、破伤风梭菌(Clostridium tetani)、问号钩端螺旋体(Leptospira interrogans)、伯氏疏螺旋体(Borrelia burgdorferi)、溶血巴斯德菌(Pasteurella haemolytica)、胸膜肺炎放线杆菌(Actinobacillus pleuropneumoniae)和鸡败血支原体(Mycoplasma gallisepticum)。致病性螺旋体疾病包括梅毒(syphilis);密螺旋体病(treponematoses):雅司病(yaws)、品他病(pinta)和地方性梅毒;和钩端螺旋体病(leptospirosis)。由高等致病菌和致病真菌引起的其他感染包括放线菌病(放射菌病);诺卡氏菌病(noca rdiosis);隐球菌病(cryptococcosis)(隐球菌(Cryptococcus))、芽生菌病(blastomycosis)(芽生菌(Blastomyces))、组织胞浆菌病(histoplasmosis)(组织胞浆菌(Histoplasma))和球孢子菌病(coccidioidomycosis)(球孢子菌(Coccidiodes));念珠菌病(candidiasis)(念珠菌(Candida))、曲霉菌病(aspergillosis)(曲霉菌(Aspergillis))和毛霉菌病(mucormycosis);孢子丝菌病(sporotrichosis);副球孢子菌病(paracoccidiodomycosis)、petriellidiosis、拟球孢子菌病(torulopsosis)、足分枝菌病(mycetoma)和着色真菌病(chromomycosis);和皮肤癣菌病(dermatophytosis)。立克次体(Rickettsial)感染包括斑疹伤寒(Typhus fever)、落基山斑疹热(Ro cky Mountain spotted fever)、Q热和立克次体痘(Rickettsialpox)。支原体(mycoplasma)和衣原体

(chlamydial)感染的实例包括:肺炎支原体(mycoplasma pneumoniae);性病性淋巴肉芽肿(lymphogranuloma ven ereum);鹦鹉热(psittacosis);和围产期衣原体感染。致病性真核生物涵盖致病性原生动物和蠕虫,由此产生的感染包括:阿米巴病(amebi asis);疟疾(malaria);利什曼病(例如,由利什曼原虫(Leishmania maj or)引起);锥虫病(trypanosomiasis);弓形虫病(toxoplasmosis)(例如,由弓形虫(Toxoplasma gondii)引起);卡氏肺孢子虫(Pneumocystis cari nii);Trichans;弓形虫;巴贝斯虫病(babesiosis);贾第鞭毛虫病(giar diasis)(例如,由贾第鞭毛虫(Giardia)引起);旋毛虫病(trichinosis)(例如,由毛滴虫(Trichomonas)引起);丝虫病(filariasis);血吸虫病(schis tosomiasis)(例如,由血吸虫(Schistosoma)引起);线虫(nematodes);吸虫(trematodes)或吸虫(flukes);和绦虫(cestode)(绦虫(tapeworm))感染。其他寄生虫感染可能由蛔虫(Ascaris)、鞭虫(Trichuris)、隐孢子虫(Cry ptosporidium)和卡氏肺孢子虫等引起。

[0190] 本文还公开了来源于与感染性疾病生物体、受试者的感染或受试者的感染细胞相关的任何多肽的肽。抗原可以来源于感染性疾病生物体的核酸序列或多肽序列。感染性疾病生物体的多肽序列包括但不限于病原体来源肽、病毒来源肽、细菌来源肽、真菌来源肽和/或寄生虫来源肽。感染性疾病生物体包括但不限于严重急性呼吸系统综合征相关冠状病毒(SARS)、严重急性呼吸系统综合征冠状病毒2(SARS-CoV-2)、埃博拉病毒(Ebola)、HIV、乙型肝炎病毒(HBV)、流感、丙型肝炎病毒(HCV)、人乳头瘤病毒(HPV)、巨细胞病毒(CMV)、基孔肯雅热病毒、呼吸道合胞病毒(RSV)、登革热病毒、正粘病毒科病毒和结核病。

[0191] 可以选择预测在细胞表面上呈递的抗原,所述细胞诸如肿瘤细胞、感染细胞或免疫细胞,包括专职性抗原呈递细胞,诸如树突状细胞。可以选择预期具有免疫原性的抗原。

[0192] 由抗原核苷酸序列编码的一种或多种多肽可以包含以下中的至少一种:与MHC的结合亲和力的IC50值小于1000nM、对于MHC I类肽,长度为8-15、8、9、10、11、12、13、14或15个氨基酸、在肽内或附近存在促进蛋白酶体切割的序列基序、以及存在促进TAP转运的序列基序。对于MHC II类肽,长度为6-30、6、7、8、9、10、11、12、13、14、15、16、17、18,19、20、21、22、23、24、25、26、27、28、29或30个氨基酸、在肽内或附近存在促进细胞外或溶酶体蛋白酶(例如,组织蛋白酶)切割或HLA-DM催化的HLA结合的序列基序。

[0193] 一种或多种抗原可以在肿瘤表面上呈递。一种或多种抗原可以在感染细胞表面上呈递。

[0194] 一种或多种抗原可以在患有肿瘤的受试者中具有免疫原性,例如能够在受试者中刺激T细胞响应和/或B细胞响应。一种或多种抗原可以在患有或疑似患有感染的受试者中具有免疫原性,例如能够在受试者中刺激T细胞响应和/或B细胞响应。一种或多种抗原可以在有感染风险的受试者中具有免疫原性,例如能够在受试者中刺激提供针对感染的免疫保护(即,免疫力)的T细胞响应和/或B细胞响应,例如,诸如刺激记忆T细胞、记忆B细胞或感染特异性抗体的产生。

[0195] 一种或多种抗原可以能够刺激B细胞响应,诸如产生识别所述一种或多种抗原的抗体(例如,识别肿瘤或感染性疾病抗原的抗体)。抗体可以识别线性多肽序列或识别二级和三级结构。因此,B细胞抗原可以包括线性多肽序列或具有二级和三级结构的多肽,包括但不限于全长蛋白质、蛋白质亚基、蛋白质结构域或已知或预测具有二级和三级结构的任

何多肽序列。能够刺激B细胞对肿瘤或感染性疾病抗原的响应的抗原可以分别是在肿瘤细胞表面或感染性疾病生物体上发现的抗原。能够引发B细胞对肿瘤或感染性疾病抗原的响应的抗原可以分别是在肿瘤或感染性疾病生物体内表达的细胞内新抗原。

[0196] 一种或多种抗原可以包括能够刺激T细胞响应的抗原(例如,包括预测的T细胞表位序列的肽)和能够刺激B细胞响应的不同抗原(例如,全长蛋白质、蛋白质亚基、蛋白质结构域)的组合。

[0197] 刺激受试者的自身免疫响应的一种或多种抗原可以排除在为受试者生成疫苗的上下文中的考虑因素之外。

[0198] 至少一个抗原肽分子(例如,表位序列)的大小可以包含但不限于约5、约6、约7、约8、约9、约10、约11、约12、约13、约14、约15、约16、约17、约18、约19、约20、约21、约22、约23、约24、约25、约26、约27、约28、约29、约30、约31、约32、约33、约34、约35、约36、约37、约38、约39、约40、约41、约42、约43、约44、约45、约46、约47、约48、约49、约50、约60、约70、约80、约90、约100、约110、约120个或更多个氨基分子残基以及其中可衍生的任何范围。在具体的实施方案中,抗原肽分子等于或少于50个氨基酸。

[0199] 抗原肽和多肽可以是:对于MHC I类,长度为15个或更少的残基,并且通常由约8至约11个残基组成,特别是9或10个残基;对于MHC II类,6-30个残基,包括在内。

[0200] 如果期望,可以通过若干种方式设计更长的肽。在一种情况下,当预测或已知肽在HLA等位基因上的呈递可能性时,较长的肽可能由以下任一种组成:(1)单个呈递的肽,向每个对应基因产物的N和C末端延伸2-5个氨基酸;(2)一些或所有呈递的肽与每个肽的延伸序列的串联。在另一种情况下,当测序揭示了肿瘤中存在长(大于10个残基)新表位序列时(例如,由于导致新的肽序列的移码、通读或内含子包含),较长的肽将由以下组成:(3)整条新型肿瘤特异性或感染性疾病特异性氨基酸—从而绕过了最强HLA呈递的较短肽的计算或基于体外测试的选择的需要。在这两种情况下,使用较长的肽允许患者细胞进行内源性加工,并且可能导致更有效的抗原呈递和T细胞响应刺激。较长的肽还可以包括肽的全长蛋白质、蛋白质亚基、蛋白质结构域及其组合,诸如分别在肿瘤或感染性疾病生物体中表达的那些。可以包括较长的肽(例如,全长蛋白质、蛋白质亚基或蛋白质结构域)及其组合以刺激B细胞响应。

[0201] 抗原肽和多肽可以在HLA蛋白上呈递。在一些方面,抗原肽和多肽以大于野生型肽的亲和力在HLA蛋白上呈递。在一些方面,抗原肽或多肽可以具有至少少于5000nM、至少少于1000nM、至少少于500nM、至少少于250nM、至少少于200nM、至少少于150nM、至少少于100nM、至少少于50nM或更少的IC50。

[0202] 在一些方面,抗原肽和多肽在施用于受试者时不诱导自身免疫响应和/或引起免疫耐受。

[0203] 还提供了包含至少两种或更多种抗原肽的组合物。在一些实施方案中,组合物含有至少两种不同的肽。至少两种不同的肽可以来源于相同的多肽。不同的多肽意指肽的长度、氨基酸序列或两者不同。肽可以包括肿瘤特异性突变。肿瘤特异性肽可以来源于已知或已发现含有肿瘤特异性突变的任何多肽,或者肽来源于已知或已发现与正常细胞或组织相比在肿瘤细胞或癌组织中具有改变表达的任何多肽(例如已知或已发现与正常细胞或组织相比在肿瘤细胞或癌组织中异常表达的任何多肽)。肽可以来源于已知或疑似与感染性疾

病生物体相关的任何多肽，或者肽来源于已知或已发现与正常细胞或组织相比在感染细胞中具有改变表达的任何多肽(例如，感染性疾病多核苷酸或多肽，包括表达受限于宿主细胞的感染性疾病多核苷酸或多肽)。可以在例如COSMIC数据库或AACR基因组学证据肿瘤信息交换(GENIE)数据库中找到可衍生出抗原肽的合适的多肽。COSMIC整理了关于人类癌症中的体细胞突变的综合信息。AACR GENIE汇总了临床级癌症基因组数据并将其与数万名癌症患者的临床结果联系起来。在一些方面，肿瘤特异性突变是特定癌症类型的驱动突变。肽可以包括KRAS突变(例如，KRAS G12C、KRAS G12V、KRAS G12D和/或KRAS Q61H突变)。

[0204] 可以修饰具有期望活性或性质的抗原肽和多肽以提供某些期望属性，例如改进的药理学特性，同时增加或至少保留未修饰的肽结合期望MHC分子并激活适当T细胞的基本上所有的生物活性。例如，可以对抗原肽和多肽进行各种变化，诸如保守或非保守的取代，其中此类变化可能在它们的用途中提供某些优势，诸如改善的MHC结合、稳定性或呈递。保守取代意指用生物学上和/或化学上相似的另一个氨基酸残基替换氨基酸残基，例如，一个疏水性残基替换另一个，或一个极性残基替换另一个。取代包括组合，诸如Gly、Ala;Val、Ile、Leu、Met;Asp、Glu;Asn、Gln;Ser、Thr;Lys、Arg;和Phe、Tyr。还可以使用D-氨基酸探测单个氨基酸取代的影响。此类修饰可以使用众所周知的肽合成程序进行，如例如Merrifield,Science232:341-347(1986),Barany&Merrifield,The Peptides,Gross&Meienhofer编辑(N.Y.,Academic Press),第1-284(1979)页;和Stewart&Young,Solid Phase Peptide Synthesis,(Rockford,Ill.,Pierce),第2版(1984)中所述。

[0205] 用各种氨基酸模拟物或非天然氨基酸对肽和多肽进行的修饰特别可用于增加肽和多肽的体内稳定性。可以通过多种方式测定稳定性。例如，肽酶和各种生物介质，诸如人血浆和血清，已经用于测试稳定性。参见例如，Verhoef等人,Eur.J.Drug Metab Pharmacokin.11:291-302(1986)。使用25％人血清(体积/体积)测定可以方便地确定肽的半衰期。方案通常如下。在使用前，通过离心将合并的人血清(AB型，非热灭活)脱脂。血清然后用RPMI组织培养基稀释至25％，并且用于测试肽稳定性。在预定的时间间隔，取出少量反应溶液并将其添加至6％三氯乙酸水溶液或乙醇中。将浑浊的反应样品冷却(4℃)15分钟，然后离心以沉淀沉淀的血清蛋白。然后使用稳定性特异性色谱条件通过反相HPLC确定肽的存在。

[0206] 可以修饰肽和多肽以提供除改善的血清半衰期之外的期望属性。例如，肽刺激CTL活性的能力可以通过与序列的键来增强，所述序列含有至少一个能够刺激辅助T细胞响应的表位。免疫原性肽/辅助T缀合物可以通过间隔分子连接。间隔物通常由相对较小的中性分子组成，诸如氨基酸或氨基酸模拟物，它们在生理条件下基本上不带电。间隔物通常选自例如Ala、Gly或非极性氨基酸或中性极性氨基酸的其他中性间隔物。应当理解，任选存在的间隔物不需要由相同的残基组成，因此可以是异源寡聚体或同源寡聚体。当存在时，间隔物通常是至少一个或两个残基，更通常是三至六个残基。或者，肽可以在没有间隔物的情况下连接至辅助T肽。

[0207] 抗原肽可以直接连接至辅助T肽，或者通过位于肽的氨基或羧基末端的间隔物而连接至辅助T肽。抗原肽或辅助T肽的氨基末端可以被酰化。示例性辅助T肽包括破伤风类毒素830-843、流感307-319、疟疾环子孢子382-398和378-389。

[0208] 蛋白质或肽可以通过本领域技术人员已知的任何技术来制备，包括通过标准分子

生物学技术表达蛋白质、多肽或肽、从天然来源分离蛋白质或肽、或化学合成蛋白质或肽。先前已经公开了对应于各种基因的核苷酸和蛋白质、多肽和肽序列，并且它们可以在本领域普通技术人员已知的计算机化数据库中找到。一个这样的数据库是位于美国国立卫生研究院网站上的国家生物技术信息中心的Genbank和GenPept数据库。可以使用本文公开的技术或本领域普通技术人员已知的技术来扩增和/或表达已知基因的编码区。或者，蛋白质、多肽和肽的各种商业制剂是本领域技术人员已知的。

[0209] 在另一方面，抗原包括编码抗原肽或其部分的核酸(例如，多核苷酸)。多核苷酸可以是例如单链和/或双链的DNA、cDNA、PNA、CNA、RNA(例如，mRNA)、或多核苷酸的天然或稳定形式(诸如，例如具有硫代磷酸酯骨架的多核苷酸)或其组合，并且它可含有或可不含有内含子。可以对编码抗原的多核苷酸序列进行序列优化以改善表达，诸如通过改善转录、翻译、转录后加工和/或RNA稳定性。例如，编码抗原的多核苷酸序列可以是密码子优化的。“密码子优化”在本文中是指根据给定生物体的密码子偏性，使用常用的同义密码子替换不常用的密码子。可以优化多核苷酸序列以改善转录后加工，例如优化以减少非预期的剪接，诸如通过去除剪接基序(例如，规范和/或隐蔽非规范剪接供体、分支和/或受体序列)和/或引入外源剪接基序(例如，剪接供体、分支和/或受体序列)以偏向有利的剪接事件。外源内含子序列包括但不限于来源于SV40(例如，SV40微型内含子)和来源于免疫球蛋白(例如，人β-珠蛋白基因)的那些。外源内含子序列可以掺入到启动子/增强子序列和抗原序列之间。Callendret等人(Virology.2007Jul 5;363(2):288-302)更详细地描述了用于表达载体的外源内含子序列，所述文献出于所有目的通过引用并入本文。可以优化多核苷酸序列以改善转录物稳定性，例如通过去除RNA不稳定基序(例如，富含AU的元件和3'UTR基序)和/或重复核苷酸序列。可以优化多核苷酸序列以改善准确转录，例如通过去除隐蔽的转录起始子和/或终止子。可以优化多核苷酸序列以改善翻译和翻译准确性，例如通过去除隐蔽的AUG起始密码子、提前聚A序列和/或二级结构基序。可以优化多核苷酸序列以改善转录物的核输出，诸如通过添加组成型转运元件(CTE)、RNA转运元件(RTE)或土拨鼠转录后调控元件(WPRE)。Callendret等人(Virology.2007Jul 5;363(2):288-302)更详细地描述了用于表达载体的核输出信号，所述文献出于所有目的通过引用并入本文。多核苷酸序列可以针对GC含量进行优化，例如以反映给定生物体的平均GC含量。序列优化可以平衡一种或多种序列性质，诸如转录、翻译、转录后加工和/或RNA稳定性。序列优化可以生成平衡转录、翻译、转录后加工和RNA稳定性中的每一个的最佳序列。序列优化算法是本领域技术人员已知的，诸如GeneArt(Thermo Fisher)、Codon Optimization Tool(IDT)、Cool Tool(University of Singapore)、SGI-DNA(La Jolla California)。可以单独地对抗原编码蛋白的一个或多个区域进行序列优化。

[0210] 又另一个方面提供能够表达多肽或其部分的表达载体。用于不同细胞类型的表达载体是本领域众所周知的，并且可以在不进行过度实验的情况下进行选择。一般来讲，DNA以正确的方向和正确的表达阅读框插入到表达载体(诸如质粒)中。如果必要，可以将DNA连接至由期望宿主识别的适当转录和翻译调控控制核苷酸序列，但此类控制通常在表达载体中可用。然后通过标准技术将载体引入到宿主中。指导可见于例如Sambrook等人(1989)Molecular Cloning,A Laboratory Manual,Cold Spring Harbor Laboratory,Cold Spring Harbor,N.Y。

[0211] V.疫苗组合物

[0212] 本文还公开了能够产生特异性免疫响应(例如肿瘤特异性免疫响应或感染性疾病生物体特异性免疫响应)的免疫原性组合物(例如疫苗组合物)。疫苗组合物通常包含例如使用本文所述的方法选择,或选自病原体来源肽、病毒来源肽、细菌来源肽、真菌来源肽和/或寄生虫来源肽的一种或多种抗原。疫苗组合物也可以称为疫苗。

[0213] 疫苗可以含有1至30个肽、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29或30个不同的肽、6、7、8、9、10 11、12、13或14个不同的肽、12、13或14个不同的肽。肽可以包括翻译后修饰。疫苗可以含有1至100或更多个核苷酸序列、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、50、51、52、53、54、55、56、57、58、59、60、61、62、63、64、65、66、67、68、69、70、71、72、73、74、75、76、77、78、79、80、81、82、83、84、85、86、87、88、89、90、91、92、93、94、95、96、97、98、99、100或更多个不同的核苷酸序列、6、7、8、9、10 11、12、13或14个不同的核苷酸序列或12、13或14个不同的核苷酸序列。疫苗可以含有1至30个抗原序列、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、50、51、52、53、54、55、56、57、58、59、60、61、62、63、64、65、66、67、68、69、70、71、72、73、74、75、76、77、78、79、80、81、82、83、84、85、86、87、88、89、90、91、92、93、94、95、96、97、98、99、100或更多个不同的抗原序列、6、7、8、9、10 11、12、13或14个不同的抗原序列或12、13或14个不同的抗原序列。

[0214] 疫苗可以含有1至30个抗原编码核酸序列、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、50、51、52、53、54、55、56、57、58、59、60、61、62、63、64、65、66、67、68、69、70、71、72、73、74、75、76、77、78、79、80、81、82、83、84、85、86、87、88、89、90、91、92、93、94、95、96、97、98、99、100或更多个不同的抗原编码核酸序列、6、7、8、9、10 11、12、13或14个不同的抗原编码核酸序列或12、13或14个不同的抗原编码核酸序列。抗原编码核酸序列可以指“抗原盒”的抗原编码部分。本文更详细地描述了抗原盒的特征。抗原编码核酸序列可以含有一个或多个表位编码核酸序列(例如,编码串联的T细胞表位的抗原编码核酸序列)。

[0215] 疫苗可以含有1至30个不同的表位编码核酸序列、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、50、51、52、53、54、55、56、57、58、59、60、61、62、63、64、65、66、67、68、69、70、71、72、73、74、75、76、77、78、79、80、81、82、83、84、85、86、87、88、89、90、91、92、93、94、95、96、97、98、99、100或更多个不同的表位编码核酸序列、6、7、8、9、10 11、12、13或14个不同的表位编码核酸序列或12、13或14个不同的表位编码核酸序列。表位编码核酸序列可以指单个表位序列的序列,诸如编码串联的T细胞表位的抗原编码核酸序列中的每个T细胞表位。

[0216] 疫苗可以含有表位编码核酸序列的至少两个重复。如本文所用,“迭代”(或可互换的“重复”)是指抗原编码核酸序列中的相同核酸表位编码核酸序列(包括本文所述的任选的5’接头序列和/或任选的3’接头序列)的两个或更多个迭代。在一个实例中,盒的抗原编

码核酸序列部分编码表位编码核酸序列的至少两个迭代。在另一个非限制性实例中，盒的抗原编码核酸序列部分编码多于一个不同表位，并且不同表位中的至少一个由编码不同表位的核酸序列的至少两个迭代(即，至少两个不同的表位编码核酸序列)编码。在说明性非限制性实例中，抗原编码核酸序列编码由表位编码核酸序列表位编码序列A(E_A)、表位编码序列B(E_B)、表位编码序列C(E_C)编码的表位A、B和C，并且具有不同表位中的至少一个的迭代的示例性抗原编码核酸序列如下式所说明，但不限于下式：

[0217] -一个不同表位的迭代(表位A的迭代)：

[0218] E_A-E_B-E_C-E_A；或

[0219] E_A-E_A-E_B-E_C

[0220] -多个不同表位的迭代(表位A、B和C的迭代)：

[0221] E_A-E_B-E_C-E_A-E_B-E_C；或

[0222] E_A-E_A-E_B-E_B-E_C-E_C

[0223] -多个不同表位的多个迭代(表位A、B和C的迭代)：

[0224] E_A-E_B-E_C-E_A-E_B-E_C-E_A-E_B-E_C；或

[0225] E_A-E_A-E_A-E_B-E_B-E_B-E_C-E_C-E_C

[0226] 以上实例不是限制性的，并且具有不同表位中的至少一个的迭代的抗原编码核酸序列可以以任何顺序或频率编码每个不同表位。例如，顺序和频率可以是不同表位的随机排列，例如，在具有表位A、B和C的示例中为式E_A-E_B-E_C-E_C-E_A-E_B-E_A-E_C-E_A-E_C-E_C-E_B。

[0227] 本文还提供了一种抗原编码盒，所述抗原编码盒具有至少一个从5'至3'由下式描述的抗原编码核酸序列：

[0228] $(E_x-(E^N_n)_y)_z$

[0229] 其中E表示包括不同的表位编码核酸序列的核苷酸序列，

[0230] n表示单独的不同表位编码核酸序列的数量并且是包括0的任何整数，

[0231] E^N表示核苷酸序列，其包含每个对应n的单独的不同表位编码核酸序列，

[0232] 对于z的每个迭代：对于每个n，x＝0或1，y＝0或1，并且x或y中的至少一者等于1，并且

[0233] z＝2或更大，其中抗原编码核酸序列包含E的至少两个迭代、给定的E^N或其组合。

[0234] 每个E或E^N可以独立地包含本文所述的任何表位编码核酸序列(例如，编码感染性疾病T细胞表位和/或新抗原表位的肽)。例如，每个E或E^N可以独立地包含从5'至3'由式($L5_b$-N_c-$L3_d$)描述的核苷酸序列，其中N包含与每个E或E^N相关的不同表位编码核酸序列，其中c＝1，L5包含5'接头序列，其中b＝0或1，并且L3包含3'接头序列，其中d＝0或1。可以使用的表位的接头在本文中进一步描述，例如，参见V.A.抗原盒。

[0235] 表位编码核酸序列的迭代(包括任选的5'接头序列和/或任选的3'接头序列)可以彼此直接线性连接(例如，E_A-E_A-……，如上所示)。表位编码核酸序列的迭代可以由一个或多个额外的核苷酸序列分开。一般来讲，表位编码核酸序列的迭代可以由适用于本文所述的组合物的任何大小的核苷酸序列分开。在一个实例中，表位编码核酸序列的迭代可以由单独的不同表位编码核酸序列分开(例如，E_A-E_B-E_C-E_A……，如上所示)。在迭代由单个单独的不同表位编码核酸序列分开，并且每个表位编码核酸序列(包括任选的5'接头序列和/或任选的3'接头序列)编码长度为25个氨基酸的肽的实例中，迭代可以相隔75个核苷酸，诸如

在由E_A-E_B-E_A……表示的抗原编码核酸中，E_A相隔75个核苷酸。在说明性实例中，具有序列VTNTEMFVTAPDNLGYMYEVQ WPGQTQPQIANCSVYDFFVWLHYYSVRDTVTNTEMFVTAPDNLG YMYEVQWPGQTQPQIANCSVYDFFVWLHYYSVRDT的抗原编码核酸序列编码25mer抗原Trp1(VTNTEMFVTAPDNLGYMYEVQWP GQ)和Trp2(TQPQIANCSVYDFFVWLHYYSVRDT)的迭代，Trp1的迭代由25mer Trp2分开，因此Trp1表位编码核酸序列的迭代与75个核苷酸的Trp2表位编码核酸序列分开。在迭代由2、3、4、5、6、7、8或9个单独的不同表位编码核酸序列分开，并且每个表位编码核酸序列(包括任选的5'接头序列和/或任选的3'接头序列)编码长度为25个氨基酸的肽的实例中，迭代可以分别与150、225、300、375、450、525、600或675个核苷酸分开。

[0236] 在一个实施方案中，选择不同的肽和/或多肽或编码它们的核苷酸序列，使得肽和/或多肽能够与不同的MHC分子缔合，诸如与不同的MHC I类分子和/或不同的MHC II类分子缔合。在一些方面，一种疫苗组合物包含能够与最常出现的MHC I类分子和/或不同的MHC II类分子缔合的肽和/或多肽的编码序列。因此，疫苗组合物可以包含能够与至少2种优选的、至少3种优选的或至少4种优选的MHC I类分子和/或不同的MHC II类分子缔合的不同片段。

[0237] 疫苗组合物可以能够刺激特异性细胞毒性T细胞响应和/或特异性辅助T细胞响应。疫苗组合物可以能够刺激特异性细胞毒性T细胞响应和特异性辅助T细胞响应。

[0238] 疫苗组合物可以能够刺激特异性B细胞响应(例如，抗体响应)。

[0239] 疫苗组合物可以能够刺激特异性细胞毒性T细胞响应、特异性辅助T细胞响应和/或特异性B细胞响应。疫苗组合物可以能够刺激特异性细胞毒性T细胞响应和特异性B细胞响应。疫苗组合物可以能够刺激特异性辅助T细胞响应和特异性B细胞响应。疫苗组合物可以能够刺激特异性细胞毒性T细胞响应、特异性辅助T细胞响应和特异性B细胞响应。

[0240] 疫苗组合物还可以包含佐剂和/或载剂。可用的佐剂和载剂的实例在下文中给出。组合物可以与载剂缔合，所述载剂诸如例如蛋白质或抗原呈递细胞，诸如例如能够将肽呈递至T细胞的树突状细胞(DC)。

[0241] 佐剂是混合到疫苗组合物中增加或以其他方式改变对抗原的免疫响应的任何物质。载剂可以是抗原能够与其缔合的支架结构，例如多肽或多糖。任选地，佐剂共价或非共价缀合。

[0242] 佐剂增加对抗原的免疫响应的能力通常表现为免疫介导反应的显著或实质性增加或疾病症状的减轻。例如，体液免疫的增加通常表现为针对抗原的抗体滴度的显著增加，而T细胞活性的增加通常表现为细胞增殖或细胞毒性或细胞因子分泌的增加。佐剂也可改变免疫响应，例如通过将主要体液响应或Th响应改变为主要细胞响应或Th响应。

[0243] 合适的佐剂包括但不限于1018ISS、明矾、铝盐、Amplivax、AS15、BCG、CP-870,893、CpG7909、CyaA、dSLIM、GM-CSF、IC30、IC31、咪喹莫特(Imiquimod)、ImuFact IMP321、IS Patch、ISS、ISCOMATRIX、JuvImmune、LipoVac、MF59、单磷酰脂A、Montanide IMS 1312、Montanide ISA 206、Montanide ISA 50V、Montanide ISA-51、OK-432、OM-174、OM-197-MP-EC、ONTAK、PepTel载体系统、PLG微颗粒、雷西莫特(resiquimod)、SRL172、病毒体和其他病毒样颗粒、YF-17D、VEGF trap、R848、β-葡聚糖、Pam3Cys、Aquila's QS21刺激子(Aquila Biotech,Worcester,Mass.,USA)，其来源于皂苷、分枝杆菌提取物和合成细菌细胞壁模拟物和其他专有佐剂，诸如Ribi's Detox.Quil或Superfos。诸如不完全弗氏佐剂(Freund's)

或GM-CSF的佐剂是可用的。先前已经描述了对树突状细胞具有特异性的若干种免疫佐剂(例如,MF59)及其制备(Dupuis M等人,Cell Immunol.1998;186(1):18-27;Allison A C;Dev Biol Stand.1998;92:3-11)。也可以使用细胞因子。若干种细胞因子直接影响树突状细胞向淋巴组织的迁移(例如,TNF-α),加速树突状细胞成熟为T淋巴细胞的高效抗原呈递细胞(例如,GM-CSF、IL-1和IL-4)(美国专利号5,849,589,通过引用整体具体并入本文)以及用作免疫佐剂(例如,IL-12)(Gabrilovich D I等人,JImmunother Emphasis Tumor Immunol.1996(6):414-418)。

[0244] 还报道了CpG免疫刺激性寡核苷酸增强佐剂在疫苗环境中的作用。也可使用其他TLR结合分子,诸如结合TLR 7、TLR 8和/或TLR 9的RNA。

[0245] 可用的佐剂的其他实例包括但不限于化学修饰的CpG(例如,CpR、Idera)、Poly(I:C)(例如,polyi:CI2U)、非CpG细菌DNA或RNA以及免疫活性小分子和抗体,诸如环磷酰胺、舒尼替尼(sunitinib)、贝伐珠单抗(bevacizumab)、西乐葆(celebrex)、NCX-4016、西地那非(sildenafil)、他达拉非(tadalafil)、伐地那非(vardenafil)、索拉非尼(sorafinib)、XL-999、CP-547632、帕唑帕尼(pazopanib)、ZD2171、AZD2171、伊匹单抗(ipilimumab)、曲美木单抗(tremelimumab)和SC58175,其可发挥治疗作用和/或作为佐剂。佐剂和添加剂的量和浓度可以由熟练的技术人员容易地确定而无需过度的实验。额外的佐剂包括集落刺激因子,诸如粒细胞巨噬细胞集落刺激因子(GM-CSF、沙格司亭(sargramostim))。

[0246] 疫苗组合物可以包含多于一种不同的佐剂。此外,治疗组合物可以包含任何佐剂物质,包括任何以上物质或其组合。还考虑疫苗和佐剂可以一起施用或以任何适当的顺序分开施用。

[0247] 载剂(或赋形剂)可以独立于佐剂存在。载剂的功能可以例如是增加特别是突变体的分子量以增加活性或免疫原性、赋予稳定性、增加生物活性或增加血清半衰期。此外,载剂可以帮助将肽呈递至T细胞。载剂可以是本领域技术人员已知的任何合适的载剂,例如蛋白质或抗原呈递细胞。载剂蛋白可以是但不限于钥孔血蓝蛋白、血清蛋白,诸如转铁蛋白、牛血清白蛋白、人血清白蛋白、甲状腺球蛋白或卵清蛋白、免疫球蛋白或激素,诸如胰岛素或棕榈酸。对于人类的免疫,载剂通常是人体可接受的并且安全的生理学上可接受的载剂。然而,破伤风类毒素和/或白喉类毒素是合适的载剂。或者,载剂可以是葡聚糖,例如琼脂糖。

[0248] 细胞毒性T细胞(CTL)识别与MHC分子结合的肽形式的抗原,而不是完整的外来抗原本身。MHC分子本身位于抗原呈递细胞的细胞表面。因此,如果存在肽抗原、MHC分子和APC的三聚体复合物,则可能激活CTL。对应地,如果不仅将肽用于激活CTL,而且如果额外添加具有相应MHC分子的APC,则可能增强免疫响应。因此,在一些实施方案中,疫苗组合物额外含有至少一种抗原呈递细胞。

[0249] 抗原还可以包括在基于病毒载体的疫苗平台中,诸如牛痘、鸡痘、自复制甲病毒、马拉巴病毒(marabavirus)、腺病毒(参见,例如,Tatsis等人,Adenoviruses,Molecular Therapy(2004)10,616—629)、或慢病毒,包括但不限于第二代、第三代、或杂交第二代/第三代慢病毒和经设计用于靶向指定细胞类型或受体的任一代重组慢病毒(参见,例如,Hu等人,Immunization Delivered by Lentiviral Vectors for Cancer and Infectious Diseases,Immunol Rev.(2011)239(1):45-61,Sakuma等人,Lentiviral vectors:basic

to translational,Biochem J.(2012)443(3):603-18,Cooper等人,Rescue of splicing-mediated intron loss maximizes expression in lentiviral vectors containing the human ubiquitin C promoter,Nucl.Acids Res.(2015)43(1):682-690,Zufferey等人,Self-Inactivating Lentivirus Vector for Safe and Efficient In Vivo Gene Delivery,J.Virol.(1998)72(12):9873-9880)。根据上文提及的基于病毒载体的疫苗平台的包装容量,这种方法可以递送编码一种或多种抗原肽的一种或多种核苷酸序列。序列可侧接非突变序列,可被接头分开,或者可在一个或多个靶向亚细胞区室的序列之前(参见,例如,Gros等人,Prospective identification of neoantigen-specific lymphocytes in the peripheral blood of melanoma patients,Nat Med.(2016)22(4):433-8、Stronen等人,Targeting of cancer neoantigens with donor-derived T cell receptor repertoires,Science.(2016)352(6291):1337-41、Lu等人,Efficient identification of mutated cancer antigens recognized by T cells associated with durable tumor regressions,Clin Cancer Res.(2014)20(13):3401-10)。引入到宿主中之后,感染细胞表达抗原,从而刺激宿主针对一种或多种肽的免疫(例如,CTL)响应。可用于免疫方案的牛痘载体和方法描述于例如美国专利号4,722,848中。另一种载体是BCG(卡介苗)。Stover等人(Nature351:456-460(1991))描述了BCG载体。根据本文的描述,可用于抗原的治疗性施用或免疫的多种其他疫苗载体(例如,伤寒沙门氏菌载体(Salmonella typhi)等)对本领域技术人员将是显而易见的。

[0250] V.A.抗原盒

[0251] 考虑到本文提供的教导,用于选择一种或多种抗原、克隆和构建“抗原盒”以及将其插入到病毒载体中的方法在本领域技术范围内。“抗原盒”或“盒”意指选定的一种或多种抗原(例如,抗原编码核酸序列)与转录抗原和表达转录产物所必需的其他调控元件的组合。选定的一种或多种抗原可以指不同的表位序列,例如盒中的抗原编码核酸序列可以编码一种表位编码核酸序列(或多种表位编码核酸序列),使得转录并表达表位。一种或多种抗原可以以允许转录的方式可操作地连接至调控组分。此类组分包括可以驱动抗原在病毒载体转染的细胞中表达的常规调控元件。因此,抗原盒也可以含有选定的启动子,其连接至抗原并且与其他任选的调控元件一起位于重组载体的选定病毒序列内。盒可以包括一种或多种抗原,诸如一种或多种病原体来源肽、病毒来源肽、细菌来源肽、真菌来源肽、寄生虫来源肽和/或肿瘤来源肽。盒可以具有一个或多个抗原编码核酸序列,诸如含有多个抗原编码核酸序列的盒,每个抗原编码核酸序列独立地可操作地连接至单独的启动子和/或使用其他多顺反子系统连接在一起,诸如2A核糖体跳跃序列元件(例如,E2A、P2A、F2A或T2A序列)或内部核糖体进入位点(IRES)序列元件。接头还可以具有切割位点,诸如TEV或弗林蛋白酶切割位点。具有切割位点的接头可以与其他元件(诸如多顺反子系统中的那些)结合使用。在非限制性说明性实例中,弗林蛋白酶切割位点可以与2A核糖体跳跃序列元件结合使用,使得弗林蛋白酶切割位点被配置为促进翻译后2A序列的去除。在含有多于一个抗原编码核酸序列的盒中,每个抗原编码核酸序列可以含有一个或多个表位编码核酸序列(例如,编码串联的T细胞表位的抗原编码核酸序列)。

[0252] 可用的启动子可以是组成型启动子或受调控的(诱导型)启动子,其将能够控制待表达的抗原量。例如,期望的启动子是巨细胞病毒立即早期启动子/增强子的启动子[参见,

例如,Boshart等人,Cell,41:521-530(1985)]。另一种期望的启动子包括劳斯肉瘤病毒(Rous sarcoma virus)LTR启动子/增强子。又另一种启动子/增强子序列是鸡细胞质β-肌动蛋白启动子[T.A.Kost等人,Nucl.Acids Res.,11(23):8287(1983)]。本领域技术人员可以选择其他合适的或期望的启动子。

[0253] 抗原盒还可以包括与病毒载体序列异源的核酸序列,包括为转录物的高效聚腺苷酸化提供信号的序列(聚(A)、poly-A或pA)和具有功能剪接供体和受体位点的内含子。用于本发明的示例性载体中的常见poly-A序列是源自乳多空病毒SV-40的序列。poly-A序列通常可以插入到基于抗原的序列之后和病毒载体序列之前的盒中。常见的内含子序列也可以来源于SV-40中,并且称为SV-40T内含子序列。抗原盒也可以含有位于启动子/增强子序列和抗原之间的这样的内含子。这些和其他常见载体元件的选择是常规的[参见,例如,Sambrook等人,"Molecular Cloning.A Laboratory Manual.",第2版,Cold Spring Harbor Laboratory,New York(1989)和其中引用的参考文献]并且许多此类序列可从商业和工业来源以及从Genbank获得。

[0254] 抗原盒可具有一种或多种抗原。例如,给定的盒可以包括1-10、1-20、1-30、10-20、15-25、15-20、1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20或更多个抗原。抗原可以直接彼此连接。抗原也可以通过接头彼此连接。抗原可以相对于彼此处于任何方向,包括N至C或C至N。

[0255] 如本文别处所述,抗原盒可以位于病毒载体中任何选定缺失的位点,诸如VEE骨架的缺失结构蛋白,或位于基于ChAd的载体的E1基因区缺失或E3基因区缺失的位点,以及其他可以选择的位点。

[0256] 抗原盒可以使用以下式描述,以描述每个元素的有序序列,从5'至3':

[0257] $(P_a\text{-}(L5_b\text{-}N_c\text{-}L3_d)_X)_Z\text{-}(P2_h\text{-}(G5_e\text{-}U_f)_Y)_W\text{-}G3_g$

[0258] 其中P和P2包含启动子核苷酸序列,N包含MHC I类表位编码核酸序列,L5包含5'接头序列,L3包含3'接头序列,G5包含编码氨基酸接头的核酸序列,G3包含至少一个编码氨基酸接头的核酸序列之一,U包含MHC II类抗原编码核酸序列,其中对于每个X,对应的Nc是表位编码核酸序列,其中对于每个Y,对应的Uf是MHC II类表位编码核酸序列(例如,通用MHC II类表位编码核酸序列)。通用序列可以包含破伤风类毒素和PADRE中的至少一种。通用序列可以包含破伤风类毒素肽。通用序列可以包含PADRE肽。通用序列可以包含破伤风类毒素和PADRE。可以通过选择存在的元素数量来进一步定义组成和有序序列,例如其中a=0或1,其中b=0或1,其中c=1,其中d=0或1,其中e=0或1,其中f=1,其中g=0或1,其中h=0或1,X=1至400,Y=0、1、2、3、4或5,Z=1至400,并且W=0、1、2、3、4或5。

[0259] 在一个实例中,存在的元素包括其中a=0、b=1、d=1、e=1、g=1、h=0、X=10、Y=2、Z=1和W=1,描述了其中不存在额外的启动子(例如,仅存在由载体骨架(诸如RNA甲病毒或ChAdV骨架)提供的启动子核苷酸序列),存在10个MHC I类表位,每个N存在5'接头,每个N存在3'接头,存在2个MHC II类表位,存在连接两个MHC II类表位的接头,存在将两个MHC II类表位的5'端连接至最终的MHC I类表位的3'接头的接头,并且存在将两个MHC II类表位的3'端连接至载体骨架(例如,ChAdV或RNA甲病毒骨架)的接头。

[0260] 将抗原盒的3'端连接至载体骨架(例如,RNA甲病毒骨架)的实例包括直接连接至由载体骨架提供的3'UTR元件,诸如3'19-nt CSE。将抗原盒的5'端连接至载体骨架(例如,

RNA甲病毒骨架)的实例包括直接连接至载体骨架的启动子或5'UTR元件,诸如亚基因组启动子序列(例如,26S亚基因组启动子序列)、甲病毒5'UTR、51-nt CSE或24-nt CSE。

[0261] 其他实例包括:在a=1的情况下,描述了除载体骨架(例如,ChAdV或RNA甲病毒骨架)提供的启动子核苷酸序列之外的启动子存在的位置;在a=1并且Z大于1的情况下,除载体骨架提供的启动子核苷酸序列之外的多个启动子存在的位置,每个驱动表达1个或多个不同的MHC I类表位编码核酸序列;在h=1的情况下,描述了驱动表达MHC II类表位编码核酸序列的单独的启动子存在的位置;以及在g=0的情况下,描述了MHC II类表位编码核酸序列(如果存在)直接连接至载体骨架(例如,ChAdV或RNA甲病毒骨架)。例如,ChAdV载体骨架可以具有置于CMV启动子/增强子控制下的抗原盒。

[0262] 其他实例包括其中存在的每个MHC I类表位可以具有5'接头、3'接头、两者都没有、或两者都有。在同一抗原盒中存在多于一个MHC I类表位的实例中,一些MHC I类表位可具有5'接头和3'接头两者,而其他MHC I类表位可具有5'接头、3'接头中的任一者,或两者都没有。在同一抗原盒中存在多于一个MHC I类表位的其他实例中,一些MHC I类表位可具有5'接头或3'接头,而其他MHC I类表位可具有5'接头、3'接头中的任一者,或两者都没有。

[0263] 在同一抗原盒中存在多于一个MHC II类表位的实例中,一些MHC II类表位可具有5'接头和3'接头两者,而其他MHC II类表位可具有5'接头、3'接头中的任一者,或两者都没有。在同一抗原盒中存在多于一个MHC II类表位的其他实例中,一些MHC II类表位可具有5'接头或3'接头,而其他MHC II类表位可具有5'接头、3'接头中的任一者,或两者都没有。

[0264] 其他实例包括其中存在的每个抗原可以具有5'接头、3'接头、两者都没有、或两者都有。在同一抗原盒中存在多于一个抗原的实例中,一些抗原可具有5'接头和3'接头两者,而其他抗原可具有5'接头、3'接头中的任一者,或两者都没有。在同一抗原盒中存在多于一个抗原的其他实例中,一些抗原可具有5'接头或3'接头,而其他抗原可具有5'接头、3'接头中的任一者,或两者都没有。

[0265] 启动子核苷酸序列P和/或P2可以与载体骨架(诸如RNA甲病毒骨架)提供的启动子核苷酸序列相同。例如,由RNA甲病毒骨架Pn和P2提供的启动子序列可以各自包含亚基因组启动子序列(例如,26S亚基因组启动子序列)或CMV启动子。启动子核苷酸序列P和/或P2可以与载体骨架(例如,ChAdV或RNA甲病毒骨架)提供的启动子核苷酸序列不同,并且可以彼此不同。

[0266] 5'接头L5可以是天然序列或非天然序列。非天然序列包括但不限于AAY、RR和DPP。3'接头L3也可以是天然序列或非天然序列。此外,L5和L3可以都是天然序列,都是非天然系列,或一个可以是天然的而另一个可以是非天然的。对于每个X,氨基酸接头的长度可以是2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、50、51、52、53、54、55、56、57、58、59、60、61、62、63、64、65、66、67、68、69、70、71、72、73、74、75、76、77、78、79、80、81、82、83、84、85、86、87、88、89、90、91、92、93、94、95、96、97、98、99、100或更多个氨基酸。对于每个X,氨基酸接头的长度还可以是至少3、至少4、至少5、至少6、至少7、至少8、至少9、至少10、至少11、至少12、至少13、至少14、至少15、至少16、至少17、至少18、至少19、至少20、至少21、至少22、至少23、至少24、至少25、至少26、至少27、至少28、至少29或至少30个氨基酸。

[0267] 对于每个Y,氨基酸接头G5的长度可以是2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、50、51、52、53、54、55、56、57、58、59、60、61、62、63、64、65、66、67、68、69、70、71、72、73、74、75、76、77、78、79、80、81、82、83、84、85、86、87、88、89、90、91、92、93、94、95、96、97、98、99、100或更多个氨基酸。对于每个Y,氨基酸接头的长度还可以是至少3、至少4、至少5、至少6、至少7、至少8、至少9、至少10、至少11、至少12、至少13、至少14、至少15、至少16、至少17、至少18、至少19、至少20、至少21、至少22、至少23、至少24、至少25、至少26、至少27、至少28、至少29或至少30个氨基酸。

[0268] 氨基酸接头G3的长度可以是2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、50、51、52、53、54、55、56、57、58、59、60、61、62、63、64、65、66、67、68、69、70、71、72、73、74、75、76、77、78、79、80、81、82、83、84、85、86、87、88、89、90、91、92、93、94、95、96、97、98、99、100或更多个氨基酸。G3的长度还可以是至少3、至少4、至少5、至少6、至少7、至少8、至少9、至少10、至少11、至少12、至少13、至少14、至少15、至少16、至少17、至少18、至少19、至少20、至少21、至少22、至少23、至少24、至少25、至少26、至少27、至少28、至少29或至少30个氨基酸。

[0269] 对于每个X,每个N可以编码MHC I类表位、MHC II类表位、能够刺激B细胞响应的表位/抗原或其组合。对于每个X,每个N可以编码MHC I类表位、MHC II类表位和能够刺激B细胞响应的表位/抗原的组合。对于每个X,每个N可以编码MHC I类表位和MHC II类表位的组合。对于每个X,每个N可以编码MHC I类表位和能够刺激B细胞响应的表位/抗原的组合。对于每个X,每个N可以编码MHC II类表位和能够刺激B细胞响应的表位/抗原的组合。对于每个X,每个N可以编码MHC II类表位。对于每个X,每个N可以编码能够刺激B细胞响应的表位/抗原。对于每个X,每个N可以编码长度为7-15个氨基酸的MHC I类表位。对于每个X,每个N还可以编码长度为5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29或30个氨基酸的MHC I类表位。对于每个X,每个N还编码长度为至少5、至少6、至少7、至少8、至少9、至少10、至少11、至少12、至少13、至少14、至少15、至少16、至少17、至少18、至少19、至少20、至少21、至少22、至少23、至少24、至少25、至少26、至少27、至少28、至少29或至少30个氨基酸的MHC I类表位。

[0270] 编码一种或多种抗原的盒可以是700个核苷酸或更少。编码一种或多种抗原的盒可以是700个核苷酸或更少,并且编码2种不同的表位编码核酸序列(例如,编码两种不同的感染性疾病或肿瘤衍生的编码免疫原性多肽的核酸序列)。编码一种或多种抗原的盒可以是700个核苷酸或更少,并且编码至少2种不同的表位编码核酸序列。编码一种或多种抗原的盒可以是700个核苷酸或更少,并且编码3种不同的表位编码核酸序列。编码一种或多种抗原的盒可以是700个核苷酸或更少,并且编码至少3种不同的表位编码核酸序列。编码一种或多种抗原的盒可以是700个核苷酸或更少,并且包括1-10、1-5、1、2、3、4、5、6、7、8、9、10或更多种抗原。

[0271] 编码一种或多种抗原的盒的长度可以在375-700个核苷酸之间。编码一种或多种抗原的盒的长度可以在375-700个核苷酸之间,并且编码2种不同的表位编码核酸序列(例如,编码两种不同的感染性疾病或肿瘤衍生的编码免疫原性多肽的核酸序列)。编码一种或

多种抗原的盒的长度可以在375-700个核苷酸之间，并且编码至少2种不同的表位编码核酸序列。编码一种或多种抗原的盒的长度可以在375-700个核苷酸之间，并且编码至少3种不同的表位编码核酸序列。编码一种或多种抗原的盒的长度在375-700个核苷酸之间，并且编码至少3种不同的表位编码核酸序列。编码一种或多种抗原的盒的长度可以在375-700个核苷酸之间，并且包括1-10、1-5、1、2、3、4、5、6、7、8、9、10或更多种抗原。

[0272] 编码一种或多种抗原的盒的长度可以是600、500、400、300、200或100个核苷酸或更少。编码一种或多种抗原的盒的长度可以是600、500、400、300、200或100个核苷酸或更少，并且编码2种不同的表位编码核酸序列。编码一种或多种抗原的盒的长度可以是600、500、400、300、200或100个核苷酸或更少，并且编码至少2种不同的表位编码核酸序列。编码一种或多种抗原的盒的长度可以是600、500、400、300、200或100个核苷酸或更少，并且编码3种不同的表位编码核酸序列。编码一种或多种抗原的盒的长度可以是600、500、400、300、200或100个核苷酸或更少，并且编码至少3种不同的表位编码核酸序列。编码一种或多种抗原的盒的长度可以是600、500、400、300、200或100个核苷酸或更少，并且包括1-10、1-5、1、2、3、4、5、6、7、8、9、10或更多种抗原。

[0273] 编码一种或多种抗原的盒的长度可以在375-600、375-500或375-400个核苷酸之间。编码一种或多种抗原的盒的长度可以在375-600、375-500或375-400个核苷酸之间，并且编码2种不同的表位编码核酸序列。编码一种或多种抗原的盒的长度可以在375-600、375-500或375-400个核苷酸之间，并且编码至少2种不同的表位编码核酸序列。编码一种或多种抗原的盒的长度可以在375-600、375-500或375-400个核苷酸之间，并且编码3种不同的表位编码核酸序列。编码一种或多种抗原的盒的长度可以在375-600、375-500或375-400个核苷酸之间，并且编码至少3种不同的表位编码核酸序列。编码一种或多种抗原的盒的长度可以在375-600、375-500或375-400个核苷酸之间，并且包括1-10、1-5、1、2、3、4、5、6、7、8、9、10或更多种抗原。

[0274] 在一些情况下，编码额外抗原和/或表位的盒中的抗原或表位相对于其他编码的抗原或表位可以是免疫显性表位。一般来讲，免疫显性是免疫响应仅针对一种或几种指定免疫原性肽的倾斜。免疫显性可以作为免疫监测方案的一部分进行评估。例如，免疫显性可以通过评价T细胞和/或B细胞对编码的抗原的响应来评估。

[0275] 免疫显性可以评估为免疫显性抗原的存在对一种或多种其他抗原的免疫响应的影响。例如，相对于不存在免疫显性抗原的情况下的免疫响应，免疫显性抗原及其各自的免疫响应(例如，免疫显性MHC I类表位)可以减少另一种抗原的免疫响应。这种减少可以使得免疫显性抗原存在下的免疫响应不被认为是治疗有效的响应。例如，如果与在不存在免疫显性MHC I类表位的情况下刺激的响应相比，T细胞对其他抗原的响应不再被认为是治疗有效响应，则MHC I类表位通常被认为是免疫显性的。相对于不存在免疫显性抗原的情况下的响应，免疫响应也可以降低至低于检测限或接近检测限。例如，如果与在不存在免疫显性MHC I类表位的情况下刺激的响应相比，T细胞对其他抗原的响应处于或低于检测限，则MHC I类表位通常被认为是免疫显性的。一般来讲，免疫显性的评估在两种都能够刺激免疫应答的抗原之间进行，例如，在施用于受试者的疫苗组合物中的两个T细胞表位之间进行，所述T细胞表位分别具有已知或预测呈递每个表位的同源MHC等位基因。免疫显性可以通过在存在和不存在可疑免疫显性抗原的情况下评价对其他抗原的相对免疫响应来评估。

[0276] 免疫显性可以评估为两种或更多种抗原之间的免疫响应的相对差异。免疫显性可以指代指定抗原相对于同一盒中编码的另一种抗原的5倍、10倍、20倍、30倍、40倍或50倍的免疫响应。免疫显性可以指代指定抗原相对于同一盒中编码的另一种抗原的100倍、200倍、300倍、400倍或500倍的免疫响应。免疫显性可以指代指定抗原相对于同一盒中编码的另一种抗原的1000倍、2000倍、3000倍、4000倍或5000倍的免疫响应。免疫显性可以指代指定抗原相对于同一盒中编码的另一种抗原的10,000倍的免疫响应。

[0277] 在一些情况下,可能期望避免含有免疫显性表位的疫苗组合物。例如,可能期望避免设计编码免疫显性表位的疫苗盒。不希望受理论的束缚,将免疫显性表位与额外表位一起施用和/或编码可降低对额外表位的免疫响应,包括潜在地最终降低针对额外表位的疫苗效力。作为说明性非限制性实例,包括TP53相关新表位的疫苗组合物可具有偏向TP53相关新表位的免疫响应,例如T细胞响应,所述TP53相关新表位负面影响(例如,将免疫响应降低至免疫响应不是治疗有效响应和/或低于检测限的程度)对疫苗组合物中的其他抗原或表位(例如,疫苗组合物中的一种或多种KRAS相关新表位)的免疫响应。因此,疫苗组合物可以设计成不含有免疫显性表位,诸如将疫苗盒(例如,(新)抗原编码盒)设计成不编码免疫显性表位。例如,相对于在免疫显性MHC I类表位不存在的情况下施用另一种表位时的免疫响应,当在疫苗组合物中施用于受试者时,盒不编码降低对盒中编码的另一种表位的免疫响应的表位。在另一实例中,相对于在免疫显性MHC I类表位不存在的情况下施用另一种表位时的免疫响应,当在疫苗组合物中施用于受试者时,盒不编码将对盒中编码的另一种表位的免疫响应降低至低于检测限的表位。在另一实例中,盒不编码降低对盒中编码的另一种表位的免疫响应的表位,其中相对于在免疫显性MHC I类表位不存在的情况下施用另一种表位时的免疫响应,当在疫苗组合物中施用于受试者时,免疫响应不是治疗有效的响应。在另一实例中,盒不编码相对于施用于受试者的疫苗组合物中的同一盒中编码的另一种表位刺激5倍、10倍、20倍、30倍、40倍或50倍或更高免疫响应的表位,其中每种抗原能够刺激受试者的免疫响应。在另一实例中,盒不编码相对于施用于受试者的疫苗组合物中的同一盒中编码的另一种表位刺激100倍、200倍、300倍、400倍、500倍或更高免疫响应的表位,其中每种抗原能够刺激受试者的免疫响应。在另一实例中,盒不编码相对于施用于受试者的疫苗组合物中的同一盒中编码的另一种表位刺激1000倍、2000倍、3000倍、4000倍、5000倍或更高免疫响应的表位,其中每种抗原能够刺激受试者的免疫响应。在另一实例中,盒不编码相对于施用于受试者的疫苗组合物中的同一盒中编码的另一种表位导致10,000倍或更高免疫响应的表位,其中每种抗原能够刺激受试者的免疫响应。

[0278] V.B.免疫调节剂

[0279] 本文所述的载体,诸如本文所述的ChAdV载体或本文所述的甲病毒载体,可以包含编码至少一种抗原的核酸,并且相同或不同的载体可以包含编码至少一种免疫调节剂的核酸。免疫调节剂可以包括结合分子(例如抗体,诸如scFv),其结合并阻断免疫检查点分子的活性。免疫调节剂可以包括细胞因子,诸如IL-2、IL-7、IL-12(包括IL-12p35、p40、p70和/或p70融合构建体)、IL-15或IL-21。免疫调节剂可以包括修饰的细胞因子(例如,pegIL-2)。载体可以包含抗原盒和一种或多种编码免疫调节剂的核酸分子。

[0280] 可以靶向用于阻断或抑制的说明性免疫检查点分子包括但不限于CTLA-4、4-1BB(CD137)、4-1BBL(CD137L)、PDL1、PDL2、PD1、B7-H3、B7-H4、BTLA、HVEM、TIM3、GAL9、LAG3、

TIM3、B7H3、B7H4、VISTA、KIR、2B4(属于CD2分子家族,并且在所有NK、γδ和记忆CD8+(αβ)T细胞上表达)、CD160(也称为BY55)和CGEN-15049。免疫检查点抑制剂包括抗体或其抗原结合片段或其他结合蛋白,它们结合并阻断或抑制以下中的一种或多种的活性:CTLA-4、PDL1、PDL2、PD1、B7-H3、B7-H4、BTLA、HVEM、TIM3、GAL9、LAG3、TIM3、B7H3、B7H4、VISTA、KIR、2B4、CD160和CGEN-15049。说明性免疫检查点抑制剂包括曲美木单抗(CTLA-4阻断抗体)、抗OX40、PD-L1单克隆抗体(抗B7-H1;MEDI4736)、伊匹单抗、MK-3475(PD-1阻断剂)、纳武单抗(Nivolumab)(抗PD1抗体)、CT-011(抗PD1抗体)、BY55单克隆抗体、AMP224(抗PDL1抗体)、BMS-936559(抗PDL1抗体)、MPLDL3280A(抗PDL1抗体)、MSB0010718C(抗PDL1抗体)和Yervoy/伊匹单抗(抗CTLA-4检查点抑制剂)。可以使用本领域的普通技术将抗体编码序列工程化成载体,诸如C68。示例性方法在Fang等人,Stable antibody expression at therapeutic levels using the 2A peptide.Nat Biotechnol.2005年5月;23(5):584-90.Epub 2005年4月17日中描述;所述文献出于所有目的通过引用并入本文。

[0281] V.C.疫苗设计和制造的其他考虑因素

[0282] V.C.1.确定一组涵盖所有肿瘤亚克隆的肽

[0283] 躯干肽意指由所有或大多数肿瘤亚克隆呈递的肽,可以优先包括在疫苗中。任选地,如果不存在预测呈递并且高概率具有免疫原性的躯干肽,或者如果预测呈递并且高概率具有免疫原性的躯干肽的数量足够小以致于疫苗中可以包括额外的非躯干肽,然后可以通过估计肿瘤亚克隆的数量和身份并选择肽来确定其他肽的优先级,以最大限度地增加疫苗涵盖的肿瘤亚克隆的数量。

[0284] V.C.2.抗原优先排序

[0285] 在应用了所有以上抗原过滤器之后,仍然有比疫苗技术可以支持的更多的候选抗原可用于疫苗包括。此外,可能保留关于抗原分析的各个方面的不确定性,并且候选疫苗抗原的不同性质之间可能存在权衡。因此,代替选择过程的每个步骤的预定过滤器,可以考虑集成的多维模型,所述模型将候选抗原放置在具有至少以下轴的空间中并使用集成方法优化选择。

[0286] 1.自身免疫或耐受的风险(种系风险)(通常优选较低的自身免疫风险)

[0287] 2.测序伪影的概率(通常优选较低的伪影概率)

[0288] 3.免疫原性的概率(通常优选较高的免疫原性概率)

[0289] 4.呈递概率(通常优选较高的呈递概率)

[0290] 5.基因表达(通常优选较高的表达)

[0291] 6.HLA基因的覆盖(参与呈递一组抗原的大量HLA分子可降低肿瘤、感染性疾病和/或感染细胞经由HLA分子的下调或突变而逃避免疫攻击的概率)

[0292] 7.HLA类别的覆盖(覆盖HLA-I和HLA-II可增加治疗响应的概率并降低肿瘤或感染性疾病逃避的可能性)

[0293] 此外,任选地,如果预测抗原由在患者肿瘤或感染细胞的全部或部分中丢失或失活的HLA等位基因呈递,则它们可以从疫苗接种中取消优先级(例如,排除)。HLA等位基因丢失可以通过基因座的体细胞突变、杂合性丢失或纯合缺失发生。用于检测HLA等位基因体细胞突变的方法是本领域众所周知的,例如(Shukla等人,2015)。同样充分描述了用于检测体细胞LOH和纯合缺失(包括HLA基因座)的方法。(Carter等人,2012;McGranahan等人,2017;

Van Loo等人,2010)。如果质谱数据表明预测的抗原不是由预测的HLA等位基因呈递,也可以取消抗原的优先级。

[0294] V.D.自扩增RNA载体

[0295] 一般来讲,所有自扩增RNA(SAM)载体含有来源于自复制病毒的自扩增骨架。术语“自扩增骨架”是指允许病毒基因组自复制的自复制病毒的最小序列。例如,允许甲病毒自复制的最小序列可以包括用于非结构蛋白介导的扩增的保守序列(例如,非结构蛋白1(nsP1)基因、nsP2基因、nsP3基因、nsP4基因和/或聚A序列)。自扩增骨架还可以包括用于表达亚基因组病毒RNA的序列(例如,甲病毒的26S启动子元件)。SAM载体可以是正义RNA多核苷酸或负义RNA多核苷酸,诸如具有来源于正义或负义自复制病毒的骨架的载体。自复制病毒包括但不限于甲病毒、黄病毒(例如,昆津病毒)、麻疹病毒和弹状病毒(例如,狂犬病病毒和水泡性口炎病毒)。来源于自复制病毒的SAM载体系统的实例在Lundstrom(Molecules.2018Dec 13;23(12).pii:E3310.doi:10.3390/molecules23123310)中更详细地描述,所述文献出于所有目的通过引用并入本文。

[0296] V.D.1.甲病毒生物学

[0297] 甲病毒是指披膜病毒科的成员,并且是正义单链RNA病毒。成员通常分类为旧世界病毒(Old World),诸如辛德比斯病毒、罗斯河病毒、马亚罗病毒、基孔肯雅热病毒(Chikungunya)和森林脑炎病毒,或新世界病毒(New World),诸如东方马脑炎病毒(eastern equine encephalitis)、Aura病毒、摩根堡病毒或委内瑞拉马脑炎病毒及其衍生毒株TC-83(Strauss Microbial Review 1994)。天然甲病毒基因组的长度通常为约12kb,其前三分之二含有编码非结构蛋白(nsP)的基因,所述非结构蛋白形成用于病毒基因组的自复制的RNA复制复合物,后三分之一含有编码用于病毒粒子生产的结构蛋白的亚基因组表达盒(Frolov RNA 2001)。

[0298] 甲病毒的模型生命周期涉及若干个不同的步骤(Strauss Microbial Review 1994,Jose Future Microbiol 2009)。病毒附接至宿主细胞后,病毒粒子与内吞区室内的膜融合,从而导致基因组RNA最终释放到胞质溶胶中。呈正链方向并且包含5’甲基鸟苷酸帽和3’聚A尾的基因组RNA被翻译以产生形成复制复合物的非结构蛋白nsP1-4。在感染早期,正链随后被复合物复制到负链模板中。在目前的模型中,随着感染的进展,复制复合物被进一步处理,所得的加工复合物转换为将负链转录为全长正链基因组RNA,以及含有结构基因的26S亚基因组正链RNA。已经鉴定出甲病毒的若干个保守序列元件(CSE)可能在各种RNA复制步骤中发挥作用,所述元件包括:从负链模板复制正链RNA时的5’UTR的互补链、从基因组模板复制负链合成时的51-nt CSE、从负链转录亚基因组RNA时nsPs和26S RNA之间的连接区中的24-nt CSE、以及从正链模板合成负链时的3’19-nt CSE。

[0299] 在各种RNA种类复制之后,病毒颗粒通常在病毒的自然生命周期中组装。26S RNA被翻译,并且所得蛋白质进一步加工产生结构蛋白,包括衣壳蛋白、糖蛋白E1和E2,以及两个小多肽E3和6K(Strauss1994)。病毒RNA发生衣壳化,通常仅对基因组RNA具有特异性的衣壳蛋白被包装,随后是病毒体组装和膜表面的出芽。

[0300] V.D.2.作为递送载体的甲病毒

[0301] 甲病毒(包括甲病毒序列、特征和其他元素)可以用于生成基于甲病毒的递送载体(也称为甲病毒载体、甲病毒病毒载体、甲病毒疫苗载体、自复制RNA(srRNA)载体或自扩增

mRNA(SAM)载体)。甲病毒先前被工程化用作表达载体系统(Pushko 1997,Rheme 2004)。甲病毒提供若干种优势,特别是在期望异源抗原表达的疫苗环境中。由于其在宿主胞质溶胶中自复制的能力,甲病毒载体通常能够在细胞内产生高拷贝数的表达盒,从而导致高水平的异源抗原产生。此外,载体通常是瞬时的,从而导致改善了生物安全性并减少了对载体的免疫耐受的诱导。一般来讲,与其他标准病毒载体(诸如人腺病毒)相比,公众对甲病毒载体也缺乏预先存在的免疫力。基于甲病毒的载体通常也导致对感染细胞的细胞毒性响应。在某种程度上,细胞毒性在疫苗环境中可能很重要,以正确刺激对表达的异源抗原的免疫响应。然而,期望的细胞毒性程度可能是一种平衡行为,因此已经开发了若干种减毒甲病毒,包括VEE的TC-83毒株。因此,本文所述的抗原表达载体的实例可以利用允许高水平抗原表达的甲病毒骨架,刺激对抗原的强烈免疫响应,不刺激对载体本身的免疫响应,并且可以以安全的方式使用。此外,抗原表达盒可以设计为通过优化载体使用的甲病毒序列来刺激不同水平的免疫响应,所述序列包括但不限于来源于VEE或其减毒衍生物TC-83的序列。

[0302] 已经使用甲病毒序列工程化了若干种表达载体设计策略(Pushko1997)。在一种策略中,甲病毒载体设计包括在结构蛋白基因下游插入26S启动子序列元件的第二个拷贝,然后是异源基因(Frolov 1993)。因此,除了天然的非结构蛋白和结构蛋白之外,还产生了表达异源蛋白的额外亚基因组RNA。在这种系统中,存在用于产生感染性病毒粒子的所有元素,因此可以在未感染细胞中发生表达载体的重复感染。

[0303] 另一种表达载体设计利用辅助病毒系统(Pushko 1997)。在这种策略中,结构蛋白被异源基因替换。因此,在仍然完整的非结构基因介导的病毒RNA自复制之后,26S亚基因组RNA提供异源蛋白的表达。传统上,然后反式提供表达结构蛋白的额外载体以产生感染性病毒,诸如通过细胞系的共转染。在USPN 8,093,021中详细描述了一种系统,其出于所有目的通过引用整体并入本文。辅助载体系统提供了限制形成传染性颗粒的可能性的益处,因此提高了生物安全性。此外,辅助载体系统减少了总载体长度,潜在地提高了复制和表达效率。因此,本文所述的抗原表达载体的实例可以利用甲病毒骨架,其中结构蛋白被抗原盒替换,所得载体减少了生物安全性问题,同时由于减少了整体表达载体的大小而促进了高效表达。

[0304] V.D.3.自复制病毒的体外产生

[0305] 本领域众所周知的用于RNA产生的方便技术是体外转录(IVT)。在这种技术中,首先通过本领域技术人员众所周知的技术产生期望载体的DNA模板,所述技术包括标准分子生物学技术,诸如克隆、限制性消化、连接、基因合成(例如,化学和/或酶促合成)和聚合酶链反应(PCR)。

[0306] DNA模板在期望转录成RNA的序列(例如,SAM)的5'端含有RNA聚合酶启动子。启动子包括但不限于噬菌体聚合酶启动子,诸如T3、T7、K11或SP6。根据选择的指定RNA聚合酶启动子序列,除了期望的序列之外,还可以转录额外的5'核苷酸。例如,规范的T7启动子可以被称为序列TAATACGACTCACTATAGG,其中使用DN A模板TAATACGACTCACTATAGGN产生期望序列N的IVT反应将产生mRNA序列GG-N。一般来讲,并且不希望受理论的束缚,T7聚合酶更高效地转录以鸟苷开始的RNA转录物。在不期望额外的5'核苷酸的情况下(例如,无额外的GG),包含在DNA模板中的RNA聚合酶启动子可以是产生仅含有期望序列的5'核苷酸的转录物的序列,所述转录物例如具有衍生SAM载体的自复制病毒的天然5'序列的SAM。例如,最小的T7

启动子可以被称为序列TAATACGACTC ACTATA,其中使用DNA模板TAATACGACTCACTATAN产生期望序列N的IVT反应将产生mRNA序列N。同样,被称为序列ATTT AGGTGACACTATA的最小SP6启动子可以用于生成无额外5'核苷酸的转录物。在典型的IVT反应中,DNA模板与适当的RNA聚合酶、缓冲剂和核苷酸(NTP)一起孵育。

[0307] 所得的RNA多核苷酸可以任选地进一步修饰,包括但不限于添加5'帽结构,诸如7-甲基鸟苷或相关结构,以及任选地修饰3'端以包括聚腺苷酸(聚A)尾。在改良的IVT反应中,通过在IVT期间添加帽类似物,RNA以共转录方式被5'帽结构加帽。帽类似物可以包括二核苷酸(m^7G-ppp-N)帽类似物或三核苷酸(m^7G-ppp-N-N)帽类似物,其中N代表核苷酸或修饰的核苷酸(例如,核糖核苷,包括但不限于腺苷、鸟苷、胞苷和尿苷)。示例性帽类似物及其在IVT反应中的用途也在美国专利号10,519,189中更详细地描述,所述专利出于所有目的通过引用并入本文。如所讨论的,T7聚合酶更高效地转录以鸟苷开始的RNA转录物。为了改善不以鸟苷开始的模板的转录效率,可以使用三核苷酸帽类似物(m^7G-ppp-N-N)。三核苷酸帽类似物可以相对于使用二核苷酸帽类似物(m^7G-ppp-N)的IVT反应将转录效率提高2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20倍或更多倍。

[0308] 也可以在转录后添加5'帽结构,诸如使用含有mRNA 2'-O-甲基转移酶和S-腺苷甲硫氨酸的牛痘加帽系统(例如,NEB目录号M2080)。

[0309] 所得的RNA多核苷酸可以任选地与所述加帽技术分开或除所述加帽技术之外进一步修饰,包括但限于修饰3'端以包括聚腺苷酸(聚A)尾。

[0310] 然后可以使用本领域众所周知的技术纯化RNA,诸如苯酚-氯仿提取或柱纯化(例如,基于色谱法的纯化)。

[0311] V.D.4.经由脂质纳米颗粒的递送

[0312] 疫苗载体设计中需要考虑的一个重要方面是针对载体本身的免疫力(Riley 2017)。这可以是对载体本身预先存在的免疫力的形式,诸如某些人类腺病毒系统,或者是在施用疫苗后对载体发展免疫力的形式。如果多次施用相同疫苗,诸如单独的初免和加强剂量,或者如果要使用相同疫苗载体系统来递送不同的抗原盒,则后者是一个重要的考虑因素。

[0313] 在甲病毒载体的情况下,标准递送方法是先前讨论的辅助病毒系统,其反式提供衣壳、E1和E2蛋白以产生感染性病毒颗粒。然而,重要的是要注意,E1和E2蛋白通常是中和抗体的主要靶标(Strauss1994)。因此,如果感染性颗粒被中和抗体靶向,则使用甲病毒载体将感兴趣的抗原递送至靶细胞的功效可能会降低。

[0314] 病毒颗粒介导的基因递送的替代方案是使用纳米材料递送表达载体(Riley 2017)。重要的是,纳米材料媒介物可以由非免疫原性材料制成,并且通常避免刺激对递送载体本身的免疫。这些材料可以包括但不限于脂质、无机纳米材料和其他聚合材料。脂质可以是阳离子的、阴离子的或中性的。材料可以是合成的或天然来源的,并且在一些情况下是可生物降解的。脂质可以包括脂肪、胆固醇、磷脂、脂质缀合物,包括但不限于聚乙二醇(PEG)缀合物(PEG化脂质)、蜡、油、甘油酯和脂溶性维生素。

[0315] 脂质纳米颗粒(LNP)是一种有吸引力的递送系统,因为脂质的两亲性能够形成膜和囊泡样结构(Riley 2017)。一般来讲,这些囊泡通过吸收到靶细胞的膜中并且将核酸释放到胞质溶胶中来递送表达载体。此外,LNP可以进一步修饰或功能化以促进指定细胞类型

的靶向。LNP设计中的另一个考虑因素是靶向效率和细胞毒性之间的平衡。脂质组合物通常包括阳离子、中性、阴离子和两亲性脂质的定义混合物。在一些情况下，包括指定脂质以防止LNP聚集、防止脂质氧化或提供促进附接额外部分的功能性化学基团。脂质组成会影响整体LNP大小和稳定性。在一个实例中，脂质组合物包含二亚油醇基甲基-4-二甲基氨基丁酸酯(MC3)或MC3样分子。MC3和MC3样脂质组合物可以配制成包括一种或多种其他脂质，诸如PEG或PEG缀合脂质、固醇或中性脂质。

[0316] 直接暴露于血清的核酸载体(诸如表达载体)可能具有若干种不期望的后果，包括血清核酸酶对核酸的降解或游离核酸对免疫系统的脱靶刺激。因此，甲病毒载体的封装可以用于避免降解，同时也避免潜在的脱靶效应。在某些实例中，甲病毒载体完全封装在递送媒介物内，诸如在LNP的水性内部中。可以通过本领域技术人员众所周知的技术将甲病毒载体封装在LNP内，诸如在微流体液滴生成装置上进行微流体混合和液滴生成。此类装置包括但不限于标准T形接头装置或流量聚焦装置。在一个实例中，将期望的脂质制剂(诸如含有MC3或MC3样的组合物)与甲病毒递送载体和其他期望的剂并行地提供至液滴生成装置，使得递送载体和期望的剂完全封装在基于MC3或MC3样的LNP的内部中。在一个实例中，液滴生成装置可以控制所产生的LNP的大小范围和大小分布。例如，LNP的大小范围可以是1至1000纳米，例如1、10、50、100、500或1000纳米。生成液滴后，可以进一步处理或修饰封装表达载体的递送媒介物以制备它们用于施用。

[0317] V.E.黑猩猩腺病毒(ChAd)

[0318] V.E.1.使用黑猩猩腺病毒的病毒递送

[0319] 用于递送一种或多种抗原(例如，经由抗原盒)的疫苗组合物可以通过提供黑猩猩来源的腺病毒核苷酸序列、多种新型载体和表达黑猩猩腺病毒基因的细胞系来产生。黑猩猩C68腺病毒(本文中也称为ChAdV68)的核苷酸序列可以在用于抗原递送的疫苗组合物中使用(参见SEQ ID NO:1)。C68腺病毒衍生载体的使用在USPN 6,083,716中更详细地描述，所述专利出于所有目的通过引用整体并入本文。基于ChAdV68的载体和递送系统在美国申请公开号US20200197500A1和国际专利申请公开WO2020243719A1中详细描述，所述专利各自出于所有目的通过引用并入本文。

[0320] 在另一方面，本文提供了一种重组腺病毒，其包含黑猩猩腺病毒(诸如C68)的DNA序列和抗原盒，所述抗原盒可操作地连接至指导其表达的调控序列。重组病毒能够感染哺乳动物细胞，优选人细胞，并且能够在细胞中表达抗原盒产物。在这种载体中，可以缺失天然黑猩猩E1基因和/或E3基因和/或E4基因。可以将抗原盒插入到任何这些基因缺失位点中。抗原盒可以包括期望针对其引发的免疫响应的抗原。

[0321] 在另一方面，本文提供感染了黑猩猩腺病毒(诸如C68)的哺乳动物细胞。

[0322] 在又另一个方面，提供了一种新型哺乳动物细胞系，其表达黑猩猩腺病毒基因(例如，来自C68的基因)或其功能片段。

[0323] 在又另一个方面，本文提供了一种用于将抗原盒递送至哺乳动物细胞的方法，其包括将有效量的黑猩猩腺病毒(诸如C68)引入到细胞中的步骤，所述黑猩猩腺病毒已被工程化以表达抗原盒。

[0324] 又另一个方面提供了一种用于在哺乳动物宿主中刺激免疫响应以治疗癌症的方法。方法可以包括向宿主施用有效量的重组黑猩猩腺病毒(诸如C68)的步骤，所述重组黑猩

猩腺病毒包含编码来自肿瘤的一种或多种抗原的抗原盒，免疫响应针对所述肿瘤进行靶向。

[0325] 又另一个方面提供了一种用于在哺乳动物宿主中刺激免疫响应以治疗或预防受试者的疾病(诸如感染性疾病)的方法。方法可以包括向宿主施用有效量的重组黑猩猩腺病毒(诸如C68)的步骤，所述重组黑猩猩腺病毒包含编码诸如来自感染性疾病的一种或多种抗原的抗原盒，免疫响应针对所述感染性疾病进行靶向。

[0326] 还公开了一种非猿猴哺乳动物细胞，其表达从SEQ ID NO:1的序列获得的黑猩猩腺病毒基因。所述基因可以选自由SEQ ID NO:1的腺病毒E1A、E1B、E2A、E2B、E3、E4、L1、L2、L3、L4和L5组成的组。

[0327] 还公开了一种包含黑猩猩腺病毒DNA序列的核酸分子，所述DNA序列包含从SEQ ID NO:1的序列获得的基因。所述基因可以选自由SEQ ID NO:1的所述黑猩猩腺病毒E1A、E1B、E2A、E2B、E3、E4、L1、L2、L3、L4和L5基因组成的组。在一些方面，核酸分子包含SEQ ID NO:1。在一些方面，核酸分子包含SEQ ID NO:1的序列，其缺少选自由SEQ ID NO:1的E1A、E1B、E2A、E2B、E3、E4、L1、L2、L3、L4和L5基因组成的组的至少一个基因。

[0328] 还公开了一种载体，其包含获自SEQ ID NO:1的黑猩猩腺病毒DNA序列和抗原盒，所述抗原盒可操作地连接至一个或多个调控序列，所述调控序列指导所述盒在异源宿主细胞中的表达，任选地其中黑猩猩腺病毒DNA序列至少包含复制和病毒粒子衣壳化所必需的顺式元件、侧接抗原盒和调控序列的顺式元件。在一些方面，黑猩猩腺病毒DNA序列包含选自由SEQ ID NO:1的E1A、E1B、E2A、E2B、E3、E4、L1、L2、L3、L4和L5基因序列组成的组的基因。在一些方面，载体可以缺少E1A和/或E1B基因。

[0329] 本文还公开了一种腺病毒载体，其包含：部分缺失的E4基因，其包含缺失或部分缺失的E4orf2区和缺失或部分缺失的E4orf3区、以及任选的缺失或部分缺失的E4orf4区。部分缺失的E4可以包含SEQ ID NO:1中示出的序列的至少核苷酸34,916至35,642的E4缺失，并且其中载体至少包含SEQ ID NO:1中列出的序列的核苷酸2至36,518。部分缺失的E4可以包含SEQ ID NO:1中示出的序列的核苷酸34,916至34,942的至少部分缺失、SEQ ID NO:1中示出的序列的核苷酸34,952至35,305的至少部分缺失、SEQ ID NO:1中示出的序列的核苷酸35,302至35,642的至少部分缺失的E4缺失，并且其中载体至少包含SEQ ID NO:1中列出的序列的核苷酸2至36,518。部分缺失的E4可以包含SEQ ID NO:1中示出的序列的至少核苷酸34,980至36,516的E4缺失，并且其中载体至少包含SEQ ID NO:1中列出的序列的核苷酸2至36,518。部分缺失的E4可以包含SEQ ID NO:1中示出的序列的至少核苷酸34,979至35,642的E4缺失，并且其中载体至少包含SEQ ID NO:1中列出的序列的核苷酸2至36,518。部分缺失的E4可以包含E4Orf2的至少部分缺失、完全缺失的E4Orf3、和E4Orf4的至少部分缺失的E4缺失。部分缺失的E4可以包含E4Orf2的至少部分缺失、E4Orf3的至少部分缺失和E4Orf4的至少部分缺失的E4缺失。部分缺失的E4可以包含E4Orf1的至少部分缺失、完全缺失的E4Orf2和E4Orf3的至少部分缺失的E4缺失。部分缺失的E4可以包含E4Orf2的至少部分缺失和E4Orf3的至少部分缺失的E4缺失。部分缺失的E4可以包含E4Orf1的起始位点和E4Orf5的起始位点之间的E4缺失。部分缺失的E4可以是与E4Orf1的起始位点相邻的E4缺失。部分缺失的E4可以是与E4Orf2的起始位点相邻的E4缺失。部分缺失的E4可以是与E4Orf3的起始位点相邻的E4缺失。部分缺失的E4可以是与E4Orf4的起始位点相邻的E4缺

失。E4缺失可以是至少50、至少100、至少200、至少300、至少400、至少500、至少600、至少700、至少800、至少900、至少1000、至少1100、至少1200、至少1300、至少1400、至少1500、至少1600、至少1700、至少1800、至少1900或至少2000个核苷酸。E4缺失可以是至少700个核苷酸。E4缺失可以是至少1500个核苷酸。E4缺失可以是50或更少、100或更少、200或更少、300或更少、400或更少、500或更少、600或更少、700或更少、800或更少、900或更少、1000或更少、1100或更少、1200或更少、1300或更少、1400或更少、1500或更少、1600或更少、1700或更少、1800或更少、1900或更少或2000或更少的核苷酸。E4缺失可以是750个或更少的核苷酸。E4缺失可以是至少1550个核苷酸或更少。

[0330] 部分缺失的E4基因可以是SEQ ID NO:1中示出的E4基因,其至少缺少SEQ ID NO:1中示出的序列的核苷酸34,916至35,642。部分缺失的E4基因可以是SEQ ID NO:1中示出的E4基因序列,其缺少SEQ ID NO:1中示出的E4基因序列,并且其缺少SEQ ID NO:1中所示的序列的核苷酸34,916至34,942、核苷酸34,952至35,305,以及SEQ ID NO:1中示出的序列的核苷酸35,302至35,642。部分缺失的E4基因可以是SEQ ID NO:1中示出的E4基因序列,其至少缺少SEQ ID NO:1中示出的序列的核苷酸34,980至36,516。部分缺失的E4基因可以是SEQ ID NO:1中示出的E4基因序列,其至少缺少SEQ ID NO:1中示出的序列的核苷酸34,979至35,642。具有部分缺失的E4基因的腺病毒载体可以具有盒,其中盒包含至少一个有效负载核酸序列,并且其中盒包含至少一个启动子序列,所述启动子序列可操作地连接至至少一个有效负载核酸序列。具有部分缺失的E4基因的腺病毒载体可以具有SEQ ID NO:1中示出的ChAdV68序列的一个或多个基因或调控序列,任选地其中一个或多个基因或调控序列包含SEQ ID NO:1中示出的序列的黑猩猩腺病毒反向末端重复序列(ITR)、E1A、E1B、E2A、E2B、E3、E4、L1、L2、L3、L4和L5基因中的至少一种。具有部分缺失的E4基因的腺病毒载体可以具有SEQ ID NO:1中示出的序列的核苷酸2至34,916,其中部分缺失的E4基因在核苷酸2至34,916的3',并且任选地核苷酸2至34,916额外缺少SEQ ID NO:1中示出的序列的核苷酸577至3403,其对应于E1缺失,且/或缺少SEQ ID NO:1中示出的序列的核苷酸27,125至31,825,其对应于E3缺失。具有部分缺失的E4基因的腺病毒载体可以具有SEQ ID NO:1中示出的序列的核苷酸35,643至36,518,并且其中部分缺失的E4基因在核苷酸35,643至36,518的5'。具有部分缺失的E4基因的腺病毒载体可以具有SEQ ID NO:1中示出的序列的核苷酸2至34,916,其中部分缺失的E4基因在核苷酸2至34,916的3',核苷酸2至34,916额外缺少SEQ ID NO:1中示出的序列的核苷酸577至3403,其对应于E1缺失,并且缺少SEQ ID NO:1中示出的序列的核苷酸27,125至31,825,其对应于E3缺失。具有部分缺失的E4基因的腺病毒载体可以具有SEQ ID NO:1中示出的序列的核苷酸2至34,916,其中部分缺失的E4基因在核苷酸2至34,916的3',核苷酸2至34,916额外缺少SEQ ID NO:1中示出的序列的核苷酸577至3403,其对应于E1缺失,并且缺少SEQ ID NO:1中示出的序列的核苷酸27,125至31,825,其对应于E3缺失,并且具有SEQ ID NO:1中示出的序列的核苷酸35,643至36,518,并且其中部分缺失的E4基因在核苷酸35,643至36,518的5'。

[0331] 部分缺失的E4基因可以是SEQ ID NO:1中示出的E4基因序列,其至少缺少SEQ ID NO:1中示出的序列的核苷酸34,916至35,642、SEQ ID NO:1中示出的序列的核苷酸2至34,916,其中部分缺失的E4基因在核苷酸2至34,916的3'端,核苷酸2至34,916额外缺少SEQ ID NO:1中示出的序列的核苷酸577至3403,其对应于E1缺失,并且缺少SEQ ID NO:1中示出的

序列的核苷酸27,125至31,825,其对应于E3缺失,并且具有SEQ ID NO:1中示出的序列的核苷酸35,643至36,518,并且其中部分缺失的E4基因在核苷酸35,643至36,518的5'。

[0332] 本文还公开了一种用本文公开的载体(诸如经工程化以表达抗原盒的C68载体)转染的宿主细胞。本文还公开了一种人细胞,其表达通过将本文公开的载体引入到细胞中而引入其中的选定基因。

[0333] 本文还公开了一种用于将抗原盒递送至哺乳动物细胞的方法,其包括将有效量的本文公开的载体(诸如经工程化以表达抗原盒的C68载体)引入到所述细胞中。

[0334] 本文还公开了一种用于产生抗原的方法,其包括将本文公开的载体引入到哺乳动物细胞中、在合适的条件下培养细胞、以及产生抗原。

[0335] V.E.2.E1表达互补细胞系

[0336] 为了生成本文所述缺失任何基因的重组黑猩猩腺病毒(Ad),缺失的基因区域的功能(如果对病毒的复制和感染性至关重要)可以通过辅助病毒或细胞系(即互补或包装细胞系)提供给重组病毒。例如,为了生成复制缺陷型黑猩猩腺病毒载体,可以使用表达人或黑猩猩腺病毒的E1基因产物的细胞系;这种细胞系可以包括HEK293或其变体。可以遵循用于生成表达黑猩猩E1基因产物的细胞系的方案(USPN6,083,716的实施例3和4)来生成表达任何选定的黑猩猩腺病毒基因的细胞系。

[0337] 可以使用AAV增强测定来鉴定黑猩猩腺病毒E1表达细胞系。这种测定可用于鉴定通过使用其他未表征的腺病毒(例如来自其他物种的腺病毒)的E1基因制备的细胞系中的E1功能。USPN 6,083,716的实施例4B中描述了该测定。

[0338] 选定的黑猩猩腺病毒基因,例如E1,可以在启动子的转录控制下用于在选定的亲代细胞系中表达。为了这个目的,可以采用诱导型或组成型启动子。诱导型启动子包括可由锌诱导的绵羊金属硫蛋白启动子、或可由糖皮质激素(特别是地塞米松(dexamethasone))诱导的小鼠乳腺肿瘤病毒(MMTV)启动子。其他诱导型启动子,诸如国际专利申请WO95/13392中鉴定的那些,也可以用于产生包装细胞系,所述申请通过引用并入本文。也可以采用控制黑猩猩腺病毒基因表达的组成型启动子。

[0339] 可以选择亲代细胞以生成表达任何期望C68基因的新细胞系。不限于,这种亲代细胞系可以是HeLa[ATCC登录号CCL 2]、A549[ATCC登录号CCL 185]、KB[CCL 17]、Detroit[例如,Detroit 510,CCL 72]和WI-38[CCL 75]细胞。可以从其他来源获得其他合适的亲代细胞系。亲代细胞系可以包括CHO、HEK293或其变体、911、HeLa、A549、LP-293、PER.C6或AE1-2a。

[0340] E1表达细胞系可以用于生成重组黑猩猩腺病毒E1缺失载体。使用与表达一种或多种其他黑猩猩腺病毒基因产物的程序基本上相同的程序构建的细胞系可用于生成在编码那些产物的基因中缺失的重组黑猩猩腺病毒载体。此外,表达其他人Ad E1基因产物的细胞系也可用于生成黑猩猩重组Ad。

[0341] V.E.3.作为载体的重组病毒颗粒

[0342] 本文公开的组合物可以包含病毒载体,其将至少一种抗原递送至细胞。此类载体包含黑猩猩腺病毒DNA序列(诸如C68)和抗原盒,所述抗原盒可操作地连接至指导盒表达的调控序列。C68载体能够在感染的哺乳动物细胞中表达盒。C68载体可以在一个或多个病毒基因中功能性缺失。抗原盒包含受到一个或多个调控序列(诸如启动子)控制的至少一种抗

原。任选的辅助病毒和/或包装细胞系可以向黑猩猩病毒载体提供缺失的腺病毒基因的任何必需产物。

[0343] 术语“功能性缺失”意指足够量的基因区域被去除或以其他方式改变，例如通过突变或修饰，使得基因区域不再能够产生基因表达的一种或多种功能产物。可以导致功能性缺失的突变或修饰包括但不限于无义突变，诸如引入提前终止密码子和去除规范和非规范起始密码子、改变mRNA剪接或其他转录加工的突变、或其组合。如果期望，可以去除整个基因区。

[0344] 形成本文公开的载体的核酸序列的修饰，包括序列缺失、插入和其他突变，可以使用标准分子生物学技术生成并且在本发明的范围内。

[0345] V.E.4.病毒质粒载体的构建

[0346] 可用于本发明的黑猩猩腺病毒C68载体包括重组的、有缺陷的腺病毒，即在E1a或E1b基因中功能性缺失并且任选地携带其他突变(例如其他基因中的温度敏感突变或缺失)的黑猩猩腺病毒序列。预计这些黑猩猩序列也可用于由其他腺病毒和/或腺相关病毒序列形成杂交载体。由人腺病毒制备的同源腺病毒载体在已发表的文献中描述[参见，例如，上文引用的Kozarsky I和II，以及其中引用的参考文献，美国专利号5,240,846]。

[0347] 在可用于将抗原盒递送至人(或其他哺乳动物)细胞的黑猩猩腺病毒C68载体的构建中，可以在载体中采用许多腺病毒核酸序列。包含最小黑猩猩C68腺病毒序列的载体可以与辅助病毒结合使用以产生感染性重组病毒颗粒。辅助病毒提供最小黑猩猩腺病毒载体的病毒感染性和繁殖所需的基本基因产物。当在其他功能性病毒载体中仅产生黑猩猩腺病毒基因的一个或多个选定缺失时，可以通过在选定的包装细胞系中繁殖病毒以在病毒载体生产过程中提供缺失的基因产物，所述包装细胞系反式提供缺失的基因功能。

[0348] V.E.5.重组最小腺病毒

[0349] 最小的黑猩猩Ad C68病毒是一种病毒颗粒，仅包含复制和病毒包壳所需的腺病毒顺式元件。也就是说，载体含有腺病毒的顺式作用5’和3’反向末端重复序列(ITR)(其作为复制起点)和天然5’包装/增强子结构域(含有包装线性Ad基因组所必需的序列和E1启动子的增强子元件)。参见，例如，国际专利申请WO96/13597中描述的用于制备“最小”人Ad载体的技术，并且所述申请通过引用并入本文。

[0350] V.E.6.其他缺陷型腺病毒

[0351] 重组的复制缺陷型腺病毒也可以含有比最小黑猩猩腺病毒序列更多的序列。这些其他Ad载体的特征可以在于病毒基因区域的各个部分的缺失，以及通过任选使用辅助病毒和/或包装细胞系形成的感染性病毒颗粒。

[0352] 作为一个实例，可以通过缺失C68腺病毒立即早期基因E1a和延迟早期基因E1b的全部或足够部分来形成合适的载体，以消除其正常的生物学功能。当在黑猩猩腺病毒转化的互补细胞系上生长时，复制缺陷型E1缺失病毒能够复制和产生感染性病毒，所述细胞系含有功能性腺病毒E1a和E1b基因，其反式提供对应的基因产物。基于与已知腺病毒序列的同源性，预计与本领域的人重组E1缺失腺病毒一样，所得重组黑猩猩腺病毒能够感染多种细胞类型并能够表达一种或多种抗原，但是不能在大多数不携带黑猩猩E1区DNA的细胞中复制，除非细胞以非常高的感染复数被感染。

[0353] 作为另一个实例，可以从形成重组病毒一部分的黑猩猩腺病毒序列中去除全部或

部分的C68腺病毒延迟早期基因E3。

[0354] 也可以构建具有E4基因缺失的黑猩猩腺病毒C68载体。又另一种载体可以含有延迟早期基因E2a中缺失。

[0355] 也可以在黑猩猩C68腺病毒基因组的L1至L5的任何晚期基因中进行缺失。类似地，中间基因IX和IVa2中的缺失可以用于一些目的。可以在其他结构或非结构腺病毒基因中进行其他缺失。

[0356] 上文讨论的缺失可以单独使用，即腺病毒序列可以仅含有E1的缺失。或者，可以以任何组合使用有效破坏或降低其生物活性的整个基因或其部分的缺失。例如，在一个示例性载体中，腺病毒C68序列可以具有E1基因和E4基因的缺失、或E1、E2a和E3基因的缺失、或E1和E3基因的缺失、或E1、E2a和E4基因的缺失，具有或不具有E3的缺失等。如上文所讨论的，此类缺失可以与其他突变(诸如温度敏感突变)结合使用，以获得期望的结果。

[0357] 包含一种或多种抗原的盒任选地插入到黑猩猩C68 Ad病毒的任何缺失区域中。或者，如果期望，可以将盒插入到现有基因区域中以破坏该区域的功能。

[0358] V.E.7.辅助病毒

[0359] 根据用于携带抗原盒的病毒载体的黑猩猩腺病毒基因含量，可以使用辅助腺病毒或非复制病毒片段来提供足够的黑猩猩腺病毒基因序列以产生含有抗原盒的感染性重组病毒颗粒。

[0360] 可用的辅助病毒含有选定的腺病毒基因序列，所述序列不存在于腺病毒载体构建体中且/或不由转染载体的包装细胞系表达。辅助病毒可以是复制缺陷型的，并且除上述序列外还含有多种腺病毒基因。辅助病毒可以与本文所述的E1表达细胞系组合使用。

[0361] 对于C68，“辅助”病毒可以是通过用SspI剪切C68基因组的C末端而形成的片段，其从病毒的左端去除约1300bp。然后将剪切的病毒与质粒DNA共转染到E1表达细胞系中，从而通过与质粒中的C68序列的同源重组来形成重组病毒。

[0362] 辅助病毒也可以形成聚阳离子缀合物，如Wu等人，J.Biol.Chem.，264:16985-16987(1989)；K.J.Fisher和J.M.Wilson，Biochem.J.，299:49(1994年4月1日)中所述。辅助病毒可以任选地含有报告基因。许多此类报道基因是本领域已知的。辅助病毒上报告基因的存在不同于腺病毒载体上的抗原盒，所述报告基因的存在允许独立监测Ad载体和辅助病毒。这种第二报告基因用于在纯化后实现所得重组病毒和辅助病毒之间的分离。

[0363] V.E.8.病毒颗粒的组装和细胞系的感染

[0364] 可以使用常规技术来实现将选定的腺病毒、抗原盒和其他载体元件的DNA序列组装成各种中间质粒和穿梭载体，以及使用质粒和载体产生重组病毒颗粒。此类技术包括cDNA的常规克隆技术、体外重组技术(例如，Gibson组装)、使用腺病毒基因组的重叠寡核苷酸序列、聚合酶链反应和提供期望核苷酸序列的任何合适的方法。采用标准转染和共转染技术，例如CaPO4沉淀技术或脂质体介导的转染方法，诸如lipofectamine。采用的其他常规方法包括病毒基因组的同源重组、在琼脂覆盖层中形成病毒斑块、测量信号生成的方法等。

[0365] 例如，在含有期望抗原盒的病毒载体的构建和组装之后，可以在辅助病毒的存在下将载体体外转染到包装细胞系中。同源重组发生在辅助序列和载体序列之间，其允许载体中的腺病毒-抗原序列得以复制并且包装到病毒衣壳中，从而产生重组病毒载体颗粒。

[0366] 所得重组黑猩猩C68腺病毒可用于将抗原盒转移至选定细胞。在使用包装细胞系

中生长的重组病毒进行的体内实验中,E1缺失的重组黑猩猩腺病毒展示了将盒转移至非黑猩猩(优选人类)细胞中的效用。

[0367] V.E.9.重组病毒载体的使用

[0368] 所得的含有抗原盒的重组黑猩猩C68腺病毒(通过腺病毒载体和辅助病毒或腺病毒载体和包装细胞系的合作产生,如上所述)因此提供了可以将抗原体内或离体递送至受试者的高效基因转移媒介物。

[0369] 根据公开的基因治疗方法将上述重组载体施用于人。携带抗原盒的黑猩猩病毒载体可以施用于患者,优选悬浮在生物相容溶液或药学上可接受的递送媒介物中。合适的媒介物包括无菌盐水。已知为药学上可接受的载剂并且为本领域技术人员所熟知的其他水性和非水性等渗无菌注射液和水性和非水性无菌混悬液可用于此目的。

[0370] 黑猩猩腺病毒载体以以下量施用,所述量足以转导人类细胞并提供足够水平的抗原转移和表达以提供治疗益处,而没有过度的不良作用或具有医学上可接受的生理作用,所述作用可由医学领域的技术人员确定。常规和药学上可接受的施用途径包括但不限于直接递送至肝脏、鼻内、静脉内、肌内、皮下、皮内、口服和其他肠胃外施用途径。如果期望,可以将施用途径组合。

[0371] 病毒载体的剂量将主要取决于以下因素,诸如所治疗的病状、患者的年龄、体重和健康状况,因此可能因患者而异。将调整剂量以平衡治疗益处与任何副作用,并且此类剂量可根据重组载体所用于的治疗应用而变化。可以监测抗原的表达水平以确定剂量施用的频率。

[0372] 重组的复制缺陷型腺病毒可以以“药学有效量”施用,即,一定量的在施用途径中有效转染期望的细胞并提供足够水平的选定基因表达以提供疫苗接种益处(即,一些可测量水平的保护性免疫)的重组腺病毒。包含抗原盒的C68载体可以与佐剂共同施用。佐剂可以与载体分离(例如,明矾)或编码在载体内,特别是如果佐剂是蛋白质。佐剂是本领域众所周知的。

[0373] 常规和药学上可接受的施用途径包括但不限于鼻内、静脉内、肌内、气管内、皮下、皮内、经直肠、口服和其他肠胃外施用途径。如果期望,可以将施用途径组合,或根据免疫原或疾病进行调整。例如,在狂犬病的预防中,优选皮下、气管内和鼻内途径。施用途径将主要取决于所治疗疾病的性质。

[0374] 可以监测对抗原的免疫水平以确定是否需要加强剂。例如,在评估血清中的抗体滴度后,可能期望进行任选的加强免疫

[0375] VI.治疗和制造方法

[0376] 还提供了一种通过向受试者施用一种或多种抗原(例如使用本文公开的方法鉴定的多种抗原)来刺激受试者的肿瘤特异性免疫响应、针对肿瘤进行疫苗接种、治疗和/或减轻受试者的癌症症状的方法。

[0377] 还提供了一种通过向受试者施用一种或多种抗原(例如使用本文公开的方法鉴定的多种抗原)来刺激受试者的感染性疾病生物体特异性免疫响应、针对感染性疾病生物体进行疫苗接种、治疗和/或减轻受试者的与感染性疾病生物体相关的感染症状的方法。

[0378] 在一些方面,受试者已被诊断患有癌症或有发展癌症的风险。受试者可以是人、狗、猫、马或任何期望肿瘤特异性免疫响应的动物。肿瘤可以是任何实体瘤,诸如乳腺癌、卵

巢癌、前列腺癌、肺癌、肾癌、胃癌、结肠癌、睾丸癌、头颈癌、胰腺癌、脑癌、黑色素瘤，以及其他组织器官肿瘤和血液肿瘤，诸如淋巴瘤和白血病，包括急性骨髓性白血病、慢性骨髓性白血病、慢性淋巴细胞白血病、T细胞淋巴细胞白血病和B细胞淋巴瘤。

[0379] 在一些方面，受试者已被诊断患有感染或有感染的风险(例如，年龄、地理/旅行和/或工作相关的增加的感染风险或易感性，或有季节性和/或新型疾病感染的风险)。

[0380] 可以以足以刺激CTL响应的量施用抗原。可以以足以刺激T细胞响应的量施用抗原。可以以足以刺激B细胞响应的量施用抗原。可以以足以刺激T细胞响应和B细胞响应两者的量施用抗原。

[0381] 抗原可以单独施用或与其他治疗剂组合施用。治疗剂可以包括靶向感染性疾病生物体的那些，诸如抗病毒剂或抗生素剂。

[0382] 此外，还可以向受试者施用抗免疫抑制剂/免疫刺激剂，诸如检查点抑制剂。例如，还可以向受试者施用抗CTLA抗体或抗PD-1或抗PD-L1。通过抗体阻断CTLA-4或PD-L1可以增强患者对癌细胞的免疫响应。特别地，在遵循疫苗接种方案时，CTLA-4阻断已经显示是有效的。

[0383] 可以确定包括在疫苗组合物中的每种抗原的最佳量和最佳给药方案。例如，可以制备抗原或其变体用于静脉内(i.v.)注射、皮下(s.c.)注射、皮内(i.d.)注射、腹膜内(i.p.)注射、肌内(i.m.)注射。注射方法包括皮下、皮内、肌内和静脉内。DNA或RNA注射方法包括皮内、肌内、皮下和静脉内。疫苗组合物的其他施用方法是本领域技术人员已知的。

[0384] 可以编译疫苗，使得存在于组合物中的抗原的选择、数量和/或量是组织、癌症、感染性疾病和/或患者特异性的。例如，可以通过给定组织中亲代蛋白质的表达模式或通过患者的突变或疾病状态来指导肽的精确选择。选择可以取决于具体的癌症类型、具体的感染性疾病(例如，受试者感染或有感染风险的具体感染性疾病分离柱/毒株)、疾病的状态、疫苗接种的目的(例如，预防或靶向正在发生的疾病)、早期治疗方案、患者的免疫状态，当然还有患者的HLA单倍型。此外，根据特定患者的个人需要，疫苗可以含有个性化组分。实例包括根据抗原在特定患者中的表达来改变抗原的选择或在第一轮或治疗方案后调整第二次治疗。

[0385] 可以通过使用各种诊断方法(例如下文进一步描述的患者选择方法)来鉴定施用抗原疫苗的患者。患者选择可能涉及鉴定一个或多个基因的突变或表达模式。患者选择可能涉及鉴定正在感染的感染性疾病。患者选择可能涉及鉴定感染性疾病的感染风险。在一些情况下，患者选择涉及鉴定患者的单倍型。可以并行地执行各种患者选择方法，例如，测序诊断可以鉴定患者的突变和单倍型。可以按顺序执行各种患者选择方法，例如，一个诊断测试鉴定突变并且单独的诊断测试鉴定患者的单倍型，并且其中每个测试可以是相同(例如，都是高通量测序)或不同(例如，一个是高通量测序，并且另一个是Sanger测序)的诊断方法。

[0386] 对于用作癌症或感染性疾病疫苗的组合物，在正常组织中大量表达的具有类似正常自身肽的抗原可以在本文所述的组合物中避免或以低量存在。另一方面，如果已知患者的肿瘤或感染细胞表达大量的某种抗原，则用于治疗这种癌症或感染的相应药物组合物可以大量存在且/或可以包括对这种特定抗原或这种抗原的途径具有特异性的多于一种抗原。

[0387] 可以将包含抗原的组合物施用于已经罹患癌症或感染的个体。在治疗应用中，组合物以足以刺激对肿瘤抗原或感染性疾病生物体抗原的有效CTL响应并治愈或至少部分阻止症状和/或并发症的量施用于受试者。将足够实现此目标的量定义为“治疗有效剂量”。这种用途的有效量将取决于例如组合物、施用方式、所治疗疾病的阶段和严重程度、患者的体重和一般健康状况以及处方医师的判断。应当记住，组合物通常可以用于严重疾病状态，即危及生命或潜在地危及生命的情况，尤其是当癌症已经转移或感染性疾病生物体已引起器官损伤和/或其他免疫病理时。在这种情况下，考虑到外来物质的最小化和抗原的相对无毒性质，可以施用显著过量的这些组合物，并且治疗医师可能期望施用显著过量的这些组合物。

[0388] 对于治疗用途，施用可以在肿瘤的检测或手术切除时开始，或在感染的检测或治疗时开始。这之后可以是加强剂量，直到至少症状基本减轻并且此后持续一段时间，或者认为提供了免疫力(例如，产生记忆B细胞或T细胞群或抗原特异性B细胞或抗体)。

[0389] 用于治疗性治疗的药物组合物(例如，疫苗组合物)旨在用于肠胃外、局部(topical)、经鼻、口服或局部(local)施用。药物组合物可以肠胃外施用，例如静脉内、皮下、皮内或肌内施用。可以在手术切除部位施用组合物以刺激对肿瘤的局部免疫响应。可以施用组合物以靶向受试者的指定感染组织和/或细胞。本文公开了用于肠胃外施用的组合物，其包含抗原溶液，并且疫苗组合物溶解或悬浮在可接受的载剂(例如，水性载剂)中。可以使用多种水性载剂，例如，水、缓冲水、0.9％盐水、0.3％甘氨酸、透明质酸等。这些组合物可以通过常规的、众所周知的灭菌技术进行灭菌，或可以进行无菌过滤。所得水溶液可以按原样或以冻干形式包装使用，冻干制剂在施用前与无菌溶液组合。根据需要，组合物可含有药学上可接受的辅助物质以接近生理条件，诸如pH调节和缓冲剂、张力调节剂、润湿剂等，例如，醋酸钠、乳酸钠、氯化钠、氯化钾、氯化钙、脱水山梨糖醇单月桂酸酯、三乙醇胺油酸酯等。

[0390] 抗原也可以经由脂质体施用，脂质体将抗原靶向至特定的细胞组织，诸如淋巴组织。脂质体也可用于增加半衰期。脂质体包括乳液、泡沫、胶束、不溶性单层、液晶、磷脂分散体、层状层等。在这些制剂中，待递送的抗原作为脂质体的一部分单独掺入，或与分子或其他治疗或免疫原性组合物结合掺入，所述分子例如与淋巴样细胞中普遍存在的受体结合，诸如结合CD45抗原的单克隆抗体。因此，可以将充满期望抗原的脂质体导向淋巴样细胞部位，然后脂质体在那里递送选定的治疗/免疫原性组合物。脂质体可以由标准的囊泡形成脂质形成，所述脂质通常包括中性和带负电荷的磷脂和固醇，诸如胆固醇。脂质的选择通常通过考虑例如脂质体大小、酸不稳定性和脂质体在血流中的稳定性来指导。多种方法可用于制备脂质体，如Szoka等人，Ann.Rev.Biophys.Bioeng.9;467(1980)、美国专利号4,235,871、4,501,728、4,501,728、4,837,028和5,019,369中所述。

[0391] 为了靶向免疫细胞，待掺入到脂质体中的配体可以包括例如对期望免疫系统细胞的细胞表面决定簇具有特异性的抗体或其片段。脂质体悬浮液可以静脉内、局部、局部等方式施用，其剂量尤其根据施用方式、所递送的肽和所治疗疾病的阶段而变化。

[0392] 出于治疗或免疫目的，编码肽和任选的本文所述的肽中的一种或多种的核酸也可以施用于患者。许多方法被方便地用于将核酸递送至患者。例如，核酸可以作为“裸DNA”直接递送。这种方法在例如Wolff等人，Science 247:1465-1468(1990)以及美国专利号5,

580,859和5,589,466中描述。核酸也可以使用弹道递送来施用,例如,如美国专利号5,204,253中所述。可以施用仅由DNA组成的颗粒。或者,可以将DNA粘附至颗粒,诸如金颗粒。用于递送核酸序列的方法可以包括使用或不使用电穿孔的病毒载体、mRNA载体和DNA载体。

[0393] 核酸也可以与阳离子化合物(诸如阳离子脂质)复合来递送。脂质介导的基因递送方法描述于例如9618372WOAWO 96/18372;9324640WOAWO 93/24640;Mannino&Gould-Fogerite,BioTechniques6(7):682-691(1988);美国专利号5,279,833Rose美国专利号5,279,833;9106309WOAWO 91/06309;和Felgner等人,Proc.Natl.Acad.Sci.USA 84:7413-7414(1987)。

[0394] 抗原还可以包括在基于病毒载体的疫苗平台中,诸如牛痘、鸡痘、自复制甲病毒、马拉巴病毒、腺病毒(参见,例如,Tatsis等人,Adenoviruses,Molecular Therapy(2004)10,616—629)、或慢病毒,包括但不限于第二代、第三代、或杂交第二代/第三代慢病毒和经设计用于靶向指定细胞类型或受体的任一代重组慢病毒(参见,例如,Hu等人,Immunization Delivered by Lentiviral Vectors for Cancer and Infectious Diseases,Immunol Rev.(2011)239(1):45-61,Sakuma等人,Lentiviral vectors:basic to translational,Biochem J.(2012)443(3):603-18,Cooper等人,Rescue of splicing-mediated intron loss maximizes expression in lentiviral vectors containing the human ubiquitin C promoter,Nucl.Acids Res.(2015)43(1):682-690,Zufferey等人,Self-Inactivating Lentivirus Vector for Safe and Efficient In Vivo Gene Delivery,J.Virol.(1998)72(12):9873-9880)。根据上文提及的基于病毒载体的疫苗平台的包装容量,这种方法可以递送编码一种或多种抗原肽的一种或多种核苷酸序列。序列可侧接非突变序列,可被接头分开,或者可在一个或多个靶向亚细胞区室的序列之前(参见,例如,Gros等人,Prospective identification of neoantigen-specific lymphocytes in the peripheral blood of melanoma patients,Nat Med.(2016)22(4):433-8、Stronen等人,Targeting of cancer neoantigens with donor-derived T cell receptor repertoires,Science.(2016)352(6291):1337-41、Lu等人,Efficient identification of mutated cancer antigens recognized by T cells associated with durable tumor regressions,Clin Cancer Res.(2014)20(13):3401-10)。引入到宿主中之后,感染细胞表达抗原,并从而刺激宿主针对一种或多种肽的免疫(例如,CTL)响应。可用于免疫方案的牛痘载体和方法描述于例如美国专利号4,722,848中。另一种载体是BCG(卡介苗)。Stover等人(Nature351:456-460(1991))描述了BCG载体。根据本文的描述,可用于抗原的治疗性施用或免疫的多种其他疫苗载体(例如,伤寒沙门氏菌载体(Salmonella typhi)等)对本领域技术人员将是显而易见的。

[0395] 一种施用核酸的方式使用编码一个或多个表位的小基因构建体。为了产生编码选定CTL表位(小基因)以在人细胞中表达的DNA序列,表位的氨基酸序列被反向翻译。人类密码子使用表用于指导每种氨基酸的密码子选择。这些表位编码DNA序列直接相连,产生连续的多肽序列。为了优化表达和/或免疫原性,可以将额外的元件掺入到小基因设计中。可以反向翻译并包括在小基因序列中的氨基酸序列的实例包括:辅助T淋巴细胞、表位、前导(信号)序列和内质网保留信号。此外,CTL表位的MHC呈递可以通过包括合成的(例如,聚丙氨酸)或天然存在的与CTL表位相邻的侧翼序列来改善。通过组装编码小基因的正链和负链的

寡核苷酸,将小基因序列转化为DNA。使用众所周知的技术在适当的条件下合成、磷酸化、纯化和退火重叠寡核苷酸(30-100个碱基长)。使用T4 DNA连接酶来连接寡核苷酸的末端。然后可以将这种编码CTL表位多肽的合成小基因克隆到期望的表达载体中。

[0396] 纯化的质粒DNA可以使用多种制剂制备用于注射。这些中最简单的是在无菌磷酸盐缓冲盐水(PBS)中重建冻干DNA。已经描述了多种方法,并且新技术可以是可用的。如上所述,核酸方便地与阳离子脂质一起配制。此外,统称为保护性、相互作用、非冷凝(PINC)的糖脂、融合脂质体、肽和化合物也可以与纯化的质粒DNA复合,以影响变量,诸如稳定性、肌内分散或运输至指定器官或细胞类型。

[0397] 还公开了一种制造疫苗的方法,其包括执行本文公开的方法的步骤;和产生包含多种抗原或多种抗原的子集的疫苗。

[0398] 可以使用本领域已知的方法来制造本文公开的抗原。例如,本文公开的产生抗原或载体(例如,包括编码一种或多种抗原的至少一种序列的载体)的方法可以包括在适合表达抗原或载体的条件下培养宿主细胞,以及纯化抗原或载体,其中宿主细胞包含至少一种编码抗原或载体的多核苷酸。标准纯化方法包括色谱技术、电泳、免疫、沉淀、透析、过滤、浓缩和色谱聚焦技术。

[0399] 宿主细胞可以包括中国仓鼠卵巢(CHO)细胞、NSO细胞、酵母或HEK293细胞。可以用一种或多种多核苷酸转化宿主细胞,所述多核苷酸包含至少一种编码本文公开的抗原或载体的核酸序列,任选地其中分离的多核苷酸还包含启动子序列,所述启动子序列可操作地连接至至少一种编码抗原或载体的核酸序列。在某些实施方案中,分离的多核苷酸可以是cDNA。

[0400] VII.抗原使用和施用

[0401] 还可以使用的疫苗接种方法、方案和时间表包括但不限于国际申请公开WO2021092095中描述的那些,所述申请出于所有目的通过引用并入本文。

[0402] 初免/加强策略中的每个载体通常包括有包括抗原的盒。盒可以包括约1-50个抗原,它们由间隔物分开,诸如通常围绕每个抗原的天然序列或其他非天然间隔物序列,诸如AAY。盒还可以包括MHCII抗原,诸如破伤风类毒素抗原和PADRE抗原,它们可以被认为是通用的II类抗原。盒还可以包括靶向序列,诸如泛素靶向序列。此外,每个疫苗剂量可以与免疫调节剂结合(例如,同时、之前或之后)施用于受试者。每个疫苗剂量可以与检查点抑制剂(CPI)结合(例如,同时、之前或之后)施用于受试者。CPI可以包括抑制CTLA4、PD1和/或PDL1的那些,诸如抗体或其抗原结合部分。此类抗体可以包括曲美木单抗或德瓦鲁单抗(durvalumab)。每个疫苗剂量可以与细胞因子(诸如IL-2、IL-7、IL-12(包括IL-12p35、p40、p70和/或p70融合构建体)、IL-15或IL-21)结合(例如,同时、之前或之后)施用于受试者。每个疫苗剂量可以与修饰的细胞因子(例如,pegIL-2)结合(例如,同时、之前或之后)施用于受试者。

[0403] 疫苗接种方案可以用于向受试者给药一种或多种抗原。初免疫苗和加强疫苗可以用于向受试者进行给药。初免疫苗可以与本文所述的任何抗原编码载体一起使用,诸如基于ChAdV68的载体(例如,SEQ ID NO:1或2中示出的序列)。加强剂量可以与本文所述的任何抗原编码载体一起使用,诸如基于ChAdV68的载体(例如,SEQ ID NO:1或2中示出的序列)或基于SAM的载体(例如,SEQ ID NO:3或4中示出的序列)。可以施用一个或多个加强剂量并且

所述加强剂量可以是连续施用相同的加强疫苗(例如,连续施用相同的基于ChAdV68的载体或连续施用相同的基于SAM的载体)或者可以是连续施用不同的加强疫苗(例如,施用基于SAM的载体之后施用基于ChAdV68的载体)。连续施用不同的疫苗可以包括不同疫苗的任何组合。例如,疫苗策略可以使用基于ChAdV68的初免,然后使用一个或多个基于SAM的加强,以及在基于SAM的加强后使用基于ChAdV68的加强。说明性非限制性疫苗策略包括但不限于:ChAdV初免-SAM加强-SAM加强-ChAdV加强;或ChAdV初免-SAM加强-SAM加强-SAM加强-SAM加强-ChAdV加强。

[0404] 基于ChAdV68的疫苗可以以$1x10^{11}$个病毒颗粒至$1x10^{12}$个病毒颗粒的剂量施用。基于ChAdV68的疫苗可以以$1x10^{11}$个病毒颗粒的剂量施用。基于ChAdV68的疫苗可以以$5x10^{11}$个病毒颗粒的剂量施用。基于ChAdV68的疫苗可以以$1x10^{12}$个病毒颗粒的剂量施用。基于ChAdV68的疫苗的选定剂量将取决于例如组合物、施用方式、所治疗疾病的阶段和严重程度、患者的体重和一般健康状况以及处方医师的判断。

[0405] 基于SAM的疫苗可以以10-300μg RNA范围内的剂量施用。基于SAM的疫苗可以以100-300μg RNA范围内的剂量施用。基于SAM的疫苗可以以100μg RNA的剂量施用。基于SAM的疫苗可以以300μg RNA的剂量施用。基于SAM的疫苗的选定剂量将取决于例如组合物、施用方式、所治疗疾病的阶段和严重程度、患者的体重和一般健康状况以及处方医师的判断。

[0406] 可以将初免疫苗注射(例如,肌内注射)到受试者中。可以使用每剂量的双侧注射。例如,可以使用ChAdV68(C68)的一次或多次注射(例如,$1x10^{12}$个病毒颗粒的总剂量);可以使用选自0.001至1ug RNA范围内(特别是0.1或1ug)的低疫苗剂量的SAM载体的一次或多次注射;或者可以使用选自1至100ug RNA范围内(特别是10、100或300ug)的高疫苗剂量的SAM载体的一次或多次注射。

[0407] 可以在初免疫苗接种后注射(例如,肌内注射)疫苗加强(加强疫苗)。可以使用每剂量的双侧注射。例如,可以使用ChAdV68(C68)的一次或多次注射(例如,$1x10^{12}$个病毒颗粒的总剂量);可以使用选自0.001至1ug RNA范围内(特别是0.1或1ug)的低疫苗剂量的SAM载体的一次或多次注射;或者可以使用选自1至100ug RNA范围内(特别是10、100或300ug)的高疫苗剂量的SAM载体的一次或多次注射。

[0408] 可以在初免后约每1、2、3、4、5、6、7、8、9或10周,例如每4周和/或8周施用加强疫苗。可以在初免后每4周施用加强疫苗。可以在初免后每6周施用加强疫苗。可以在初免后每12周施用加强疫苗。在疫苗接种方案的过程期间,可以以不同的间隔施用加强剂量。例如,说明性非限制性实例包括初免-4w-加强-12w-加强-12w-加强;或初免-4w-加强-6w-加强-6w-加强-6w-加强-6w-加强,其中"w"代表周。

[0409] 一种或多种疫苗施用可以包括一种或多种检查点抑制剂的共同施用。说明性免疫检查点抑制剂包括曲美木单抗(CTLA-4阻断抗体)、抗OX40、PD-L1单克隆抗体(抗B7-H1;MEDI4736)、伊匹单抗、MK-3475(PD-1阻断剂)、纳武单抗(抗PD1抗体)、CT-011(抗PD1抗体)、BY55单克隆抗体、AMP224(抗PDL1抗体)、BMS-936559(抗PDL1抗体)、MPLDL3280A(抗PDL1抗体)、MSB0010718C(抗PDL1抗体)和Yervoy/伊匹单抗(抗CTLA-4检查点抑制剂)。在说明性非限制性实例中,纳武单抗、Yervoy/伊匹单抗或其组合。

[0410] 也可以将抗CTLA-4(例如,曲美木单抗)施用于受试者。例如,可以在肌内疫苗注射(ChAdV68初免或SAM低剂量)部位的附近皮下施用抗CTLA4,以确保引流至同一淋巴结。曲美

木单抗是选择性人IgG2单抗CTLA-4抑制剂。靶标抗CTLA-4(曲美木单抗)皮下剂量通常为70-75mg(特别是75mg),剂量范围为例如1-100mg或5-420mg。

[0411] 在某些情况下,可以使用抗PD-L1抗体,诸如德瓦鲁单抗(MEDI 4736)。德瓦鲁单抗是一种选择性、高亲和力的人IgG1单抗,阻断PD-L1与PD-1和CD80的结合。德瓦鲁单抗通常每4周静脉注射20mg/kg。

[0412] 可以在疫苗施用之前、期间和/或之后执行免疫监测。这种监测可以告知安全性和功效以及其他参数。

[0413] 为了执行免疫监测,通常使用PBMC。PBMC可以在初免疫苗接种之前和初免疫苗接种之后(例如,4周和8周)分离。PBMC可以在加强疫苗接种之前和每次加强疫苗接种之后(例如,4周和8周)收获。

[0414] 免疫响应,诸如T细胞响应和B细胞响应,可以作为免疫监测方案的一部分进行评估。例如,可以监测和/或评估本文所述的疫苗组合物刺激免疫响应的能力。如本文所用,“刺激免疫响应”是指免疫响应的任何增加,诸如引发免疫响应(例如,初免疫苗,其刺激未接受过治疗的受试者(naïve subject)中的免疫响应的引发)或增强免疫响应(例如,加强疫苗,其刺激受试者的免疫响应的增强,所述受试者具有对抗原的预先存在的免疫响应,诸如被初免细胞引发的预先存在的免疫响应)。可以使用本领域已知的一种或多种方法或通过细胞毒性测定来测量T细胞响应,所述方法诸如ELISpot、细胞内细胞因子染色、细胞因子分泌和细胞表面捕获、T细胞增殖、MHC多聚体染色。可以通过使用ELISpot测定来测量细胞因子(诸如IFN-γ)的诱导,从PBMC监测T细胞对疫苗中编码的表位的响应。可以通过使用流式细胞术来测量细胞内或细胞外捕获的细胞因子(诸如IFN-γ)的诱导,从PBMC监测指定CD4或CD8 T细胞对疫苗中编码的表位的响应。可以通过使用MHC多聚体染色测量表达对表位/MHC I类复合物具有特异性的T细胞受体的T细胞群,从PBMC监测指定CD4或CD8T细胞对疫苗中编码的表位的响应。可以通过在3H-胸苷、溴脱氧尿苷和羧基荧光素-二乙酸酯-琥珀酰亚胺酯(CFSE)掺入后测量T细胞群的离体扩增,从PBMC监测指定CD4或CD8 T细胞对疫苗中编码的表位的响应。对疫苗中编码的表位具有特异性的的PBMC衍生的T细胞的抗原识别能力和裂解活性可以通过铬释放测定或替代的比色细胞毒性测定在功能上进行评估。

[0415] 可以使用本领域已知的一种或多种方法来测量B细胞响应,所述方法诸如用于确定以下的测定:B细胞分化(例如,分化成浆细胞)、B细胞或浆细胞增殖、B细胞或浆细胞激活(例如,上调共刺激标志物,诸如CD80或CD86)、抗体类别转换和/或抗体生产(例如,ELISA)。

[0416] 疫苗接种方案可以包括针对感兴趣的疫苗组合物(诸如基于ChAdV68的疫苗)评估中和抗体滴度。例如,基于ChAdV68的疫苗可以作为初免剂量施用,并且在重新施用之前,在确定ChAdV特异性中和抗体滴度低于中和阈值后,将基于ChAdV68的疫苗作为加强剂量重新施用。中和抗体滴度可以是NT50值,其计算为中和ChAdV病毒50%的免疫受试者血清的最小稀释度。确定中和抗体滴度可以包括以下步骤:(1)在足以中和ChAdV病毒的条件下,将一种或多种稀释度的免疫受试者血清与ChAdV病毒接触;和(2)评估ChAdV病毒相对于未中和病毒的中和。

[0417] 可以在施用本文所述的任何疫苗组合物之后监测受试者的疾病状态。例如,可以使用从受试者分离的细胞游离DNA(cfDNA)来监测疾病状态。此外,可以使用从受试者分离的cfDNA来监测疫苗疗法的功效。cfDNA检测可以包括以下步骤:a.从受试者中分离或已经

分离出cfDNA;b.对分离的cfDNA进行测序或已经对其进行测序;c.确定或已经确定cfDNA中相对于受试者的野生型种系核酸序列的一个或多个突变的频率,以及d.从步骤(c)中评估或已经评估受试者的疾病状态。方法还可以包括,在以上步骤(c)之后,d.向给定受试者执行步骤(a)-(c)的多于一次迭代,并且比较在多于一次迭代中确定的一个或多个突变的频率;和f.从步骤(d)评估或已经评估受试者的疾病状态。可以在不同的时间点执行多于一次的迭代,诸如在施用疫苗组合物之前执行的步骤(a)-(c)的第一次迭代和在施用疫苗组合物之后执行的步骤(a)-(c)的第二次迭代。步骤(c)可以包括比较:在多于一次迭代中确定的一个或多个突变的频率,或比较在第一次迭代中确定的一个或多个突变的频率与在第二次迭代中确定的一个或多个突变的频率。在后续迭代或第二次迭代中确定的一个或多个突变频率的增加可以被评估为疾病进展。在后续迭代或第二次迭代中确定的一个或多个突变频率的减少可以被评估为响应。在一些方面,响应是完全响应(CR)或部分响应(PR)。可以在评估步骤之后向受试者施用疗法,诸如在cfDNA中的一个或多个突变的频率的评估表明受试者患有疾病的情况下。cfDNA分离步骤可以使用离心将cfDNA从细胞或细胞碎片中分离出来。cfDNA可以从全血中分离出来,诸如通过分离血浆层、血沉棕黄层和红血。cfDNA测序可以使用下一代测序(NGS)、Sanger测序、双重测序、全外显子组测序、全基因组测序、从头测序、定相测序、靶向扩增子测序、鸟枪法测序或其组合,并且可包括在测序前富集一个或多个感兴趣的多核苷酸区域(例如,已知或疑似编码一个或多个突变、编码区和/或肿瘤外显子组多核苷酸的多核苷酸)的cfDNA。富集cfDNA可包括将一种或多种多核苷酸探针与一种或多种感兴趣的多核苷酸区域杂交,所述多核苷酸探针可以被修饰(例如,生物素化)。一般来讲,可以同时或并行监测任何数量的突变。

[0418] 同源疫苗接种方案可以包括同源剂量之间的间隔以提高第二剂量的功效。例如,基于ChAdV68的疫苗可以作为初免剂量施用,并且包括在再次施用基于ChAdV68的疫苗作为加强剂量之前的间隔以提高功效,诸如降低ChAdV特异性中和抗体滴度对加强剂量的疗效的影响。例如,初免剂量可诱导中和抗体的产生,中和抗体然后随着时间的推移逐渐减少。在本文所述的基于ChAdV68的疫苗的说明性非限制性实例中,间隔为至少27周。间隔可以是27周。间隔可以是28周。间隔可以是29周。间隔可以是30周。间隔可以是31周。间隔可以是32周。间隔可以是33周。间隔可以是至少27周。间隔可以是至少28周。间隔可以是至少29周。间隔可以是至少30周。间隔可以是至少31周。间隔可以是至少32周。间隔可以是至少33周。

[0419] 在同源初免-加强策略中,基于ChAdV68的疫苗施用之间的间隔可以短至8周。间隔可以是8周。间隔可以是9周。间隔可以是10周。间隔可以是11周。间隔可以是12周。间隔可以是13周。间隔可以是14周。间隔可以是15周。间隔可以是16周。间隔可以是17周。间隔可以是18周。间隔可以是19周。间隔可以是20周。间隔可以是21周。间隔可以是23周。间隔可以是24周。间隔可以是25周。间隔可以是26周。

[0420] 在同源初免-加强策略中,基于ChAdV68的疫苗施用之间的间隔可以是至少8周。间隔可以是至少9周。间隔可以是至少10周。间隔可以是至少11周。间隔可以是至少12周。间隔可以是至少13周。间隔可以是至少14周。间隔可以是至少15周。间隔可以是至少16周。间隔可以是至少17周。间隔可以是至少18周。间隔可以是至少19周。间隔可以是至少20周。间隔可以是至少21周。间隔可以是至少23周。间隔可以是至少24周。间隔可以是至少25周。间隔可以是至少26周。

[0421] 在同源初免-加强策略中，基于ChAdV68的疫苗施用之间的间隔可以是2个月。间隔可以是2.5个月。间隔可以是3个月。间隔可以是3.5个月。间隔可以是4个月。间隔可以是4.5个月。间隔可以是5个月。间隔可以是5.5个月。间隔可以是6个月。间隔可以是6.5个月。间隔可以是7个月。间隔可以是7.5个月。间隔可以是8个月。间隔可以是8.5个月。间隔可以是至少2个月。间隔可以是至少2.5个月。间隔可以是至少3个月。间隔可以是至少3.5个月。间隔可以是至少4个月。间隔可以是至少4.5个月。间隔可以是至少5个月。间隔可以是至少5.5个月。间隔可以是至少6个月。间隔可以是至少6.5个月。间隔可以是至少7个月。间隔可以是至少7.5个月。间隔可以是至少8个月。间隔可以是至少8.5个月。

[0422] VIII.HLA肽的分离和检测

[0423] 在组织样品的裂解和溶解后，使用经典的免疫沉淀(IP)方法分离HLA肽分子(55-58)。澄清的裂解物用于HLA特异性IP。国际专利申请公开WO/2018/208856中更详细地描述了实例和方法，所述专利出于所有目的通过引用整体并入本文。

[0424] IX.呈递模型

[0425] 呈递模型可以用于鉴定患者体内肽呈递的可能性。各种呈递模型是本领域技术人员已知的，例如在美国专利号10,055,540、美国申请公开号US20200010849A1和US20110293637和国际专利申请公开WO/2018/195357、WO/2018/208856和WO2016187508中更详细描述的呈递模型，所述专利各自出于所有目的通过引用整体并入本文。

[0426] X.训练模块

[0427] 训练模块可以用于构建一个或多个基于训练数据集的呈递模型，所述呈递模型生成肽序列是否将被与肽序列相关的MHC等位基因呈递的可能性。各种训练模块是本领域技术人员已知的，例如在美国专利号10,055,540、美国申请公开号US20200010849A1和国际专利申请公开WO/2018/195357和WO/2018/208856中更详细描述的呈递模型，所述专利各自出于所有目的通过引用整体并入本文。训练模块可以构建呈递模型以基于每个等位基因预测肽的呈递可能性。训练模块还可以构建呈递模型以预测存在两个或更多个MHC等位基因的多等位基因环境中的肽的呈递可能性。

[0428] XI.预测模块

[0429] 预测模块可以用于接收序列数据并使用呈递模型在序列数据中选择候选抗原。具体地，序列数据可以是从患者的肿瘤组织细胞、患者的感染细胞或感染性疾病生物体本身中提取的DNA序列、RNA序列和/或蛋白质序列。预测模块可通过将从患者的正常组织细胞中提取的序列数据与从患者的肿瘤组织细胞中提取的序列数据进行比较以鉴定含有一个或多个突变的部分，从而鉴定作为突变肽序列的候选新抗原。预测模块可鉴定候选抗原，所述候选抗原是病原体来源肽、病毒来源肽、细菌来源肽、真菌来源肽和寄生虫来源肽，诸如通过将从患者的正常组织细胞中提取的序列数据与从患者的感染细胞中提取的序列数据进行比较以鉴定含有一个或多个感染性疾病生物体相关抗原的部分。预测模块可通过将从患者的正常组织细胞中提取的序列数据与从患者的肿瘤组织细胞中提取的序列数据进行比较以鉴定不适当表达的候选抗原，从而鉴定与正常细胞或组织相比在肿瘤细胞或癌组织中具有改变的表达的候选抗原。预测模块可通过将从患者的正常组织细胞中提取的序列数据与从患者的感染组织细胞中提取的序列数据进行比较以鉴定表达的候选抗原(例如，鉴定表达的对感染性疾病具有特异性的多核苷酸和/或多肽)，从而鉴定与正常细胞或组织相比

在感染细胞或感染组织中表达的候选抗原。

[0430] 呈递模块可以将一个或多个呈递模型应用于加工过的肽序列以估计肽序列的呈递可能性。具体地，预测模块可通过将呈递模型应用于候选抗原以选择可能在肿瘤HLA分子或感染细胞HLA分子上呈递的一种或多种候选抗原肽序列。在一种实施方式中，呈递模块选择估计的呈递可能性高于预定阈值的的候选抗原序列。在另一种实施方式中，呈递模型选择具有最高估计呈递可能性的N个候选抗原序列(其中N通常是可以在疫苗中递送的表位的最大数量)。可以将包括选择用于给定患者的候选抗原的疫苗注射到受试者体内以刺激免疫响应。

[0431] XI.B.盒设计模块

[0432] XI.B.1综述

[0433] 盒设计模块可以用于基于选定的候选肽生成疫苗盒序列以注射到患者体内。例如，盒设计模块可以用于生成编码串联表位序列(诸如串联T细胞表位)的序列。各种盒设计模块是本领域技术人员已知的，例如在美国专利号10,055,540、美国申请公开号US20200010849A1和国际专利申请公开WO/2018/195357和WO/2018/208856中更详细描述的盒设计模块，所述专利各自出于所有目的通过引用整体并入本文。

[0434] 可基于选定的肽来生成一组治疗性表位，所述选定的肽由与高于预定阈值的呈递可能性相关的预测模块确定，其中呈递可能性由呈递模型确定。然而，应当理解，在其他实施方案中，可基于多种方法中的任何一种或多种方法(单独或组合)生成治疗性表位组，例如，基于与患者的HLA I类或II类等位基因的结合亲和力或预测结合亲和力、与患者的HLA I类或II类等位基因的结合稳定性或预测结合稳定性、随机取样等。

[0435] 治疗性表位可能对应于选定的肽本身。除了选定的肽之外，治疗性表位还可包括C末端和/或N末端侧翼序列。N末端和C末端侧翼序列可以是治疗性疫苗表位在其来源蛋白背景下的天然N末端和C末端侧翼序列。治疗性表位可以代表固定长度的表位。治疗性表位可以代表可变长度的表位，其中表位的长度可以根据例如C或N侧翼序列的长度而变化。例如，C末端侧翼序列和N末端侧翼序列可以各自具有2-5个残基的不同长度，从而产生16种可能的表位选择。

[0436] 盒设计模块还可以通过考虑跨越盒中一对治疗性表位之间的连接处的连接表位的呈递来生成盒序列。连接表位是新的非自身但不相关的表位序列，其由于在盒中串联治疗性表位和接头序列的过程而出现在盒中。连接表位的新序列不同于盒本身的治疗性表位。

[0437] 盒设计模块可以生成降低连接表位在患者体内呈递的可能性的盒序列。具体地，当将盒注射到患者体内时，连接表位有可能被患者的HLA I类或HLA II类等位基因呈递，并分别刺激CD8或CD4 T细胞响应。此类反应是不期望的，因为对连接表位具有反应性的T细胞没有治疗益处，并且可能通过抗原竞争减少对盒中的选定治疗性表位的免疫响应。[76]

[0438] 盒设计模块可以迭代通过一个或多个候选盒，并且确定与该盒序列相关的连接表位的呈递评分低于数值阈值的盒序列。连接表位呈递评分是与盒中连接表位的呈递可能性相关的量，并且连接表位呈递评分的值越高表示盒的连接表位将被HLA I类或HLA II类或两者呈递的可能性越高。

[0439] 在一个实施方案中，盒设计模块可确定与候选盒序列中最低的连接表位呈递评分

相关的盒序列。

[0440] 盒设计模块可迭代一个或多个候选盒序列，确定候选盒的连接表位呈递评分，并且鉴定与低于阈值的连接表位呈递评分相关的最佳盒序列。

[0441] 盒设计模块还可以检查一个或多个候选盒序列，以鉴定候选盒序列中的任何连接表位是否是针对给定患者(针对给定患者设计疫苗)的自身表位。为了实现这一点，盒设计模块根据已知的数据库(诸如BLAST)检查连接表位。在一个实施方案中，盒设计模块可被配置为设计避免连接自身表位的盒。

[0442] 盒设计模块可以执行强力方法并且迭代通过所有或大多数可能的候选盒序列，以选择具有最小连接表位呈递评分的序列。然而，随着疫苗容量的增加，此类候选盒的数量可能过分的大。例如，对于20个表位的疫苗容量，盒设计模块必须迭代通过约10^{18}个可能的候选盒以确定具有最低连接表位呈递评分的盒。这种确定在计算上可能是繁重的(就所需的计算处理资源而言)，并且有时难以处理，因为盒设计模块要在合理的时间内完成以生成用于患者的疫苗。此外，考虑每个候选盒的可能连接表位可能甚至更加繁重。因此，盒设计模块可基于迭代通过多个候选盒序列的方式来选择盒序列，所述候选盒序列的数量显著小于用于强力方法的候选盒序列的数量。

[0443] 盒设计模块可以生成随机生成或至少伪随机生成的候选盒的子集，并选择与低于预定阈值的连接表位呈递评分相关的候选盒作为盒序列。此外，盒设计模块可从具有最低连接表位呈递评分的子集中选择候选盒作为盒序列。例如，盒设计模块可为一组20个选定表位生成约100万个候选盒的子集，并且选择具有最小连接表位呈递评分的候选盒。尽管生成随机盒序列的子集并从子集中选择具有低连接表位呈递评分的盒序列相对于强力方法可能不是最佳的，但它需要的计算资源显著更少，从而使它的实施在技术上是可行的。此外，与这种更高效的技术相反，执行强力方法可能只导致连接表位呈递评分的微小或甚至可忽略的改进，因此使其从资源分配的角度来看不值得。盒设计模块可以通过将盒的表位序列制定成非对称旅行商问题(TSP)来确定改进的盒配置。给定节点列表和每对节点之间的距离，TSP确定与最短总距离相关的节点序列，以精确地访问每个节点一次并返回到原始节点。例如，给定彼此之间距离已知的城市A、B和C，TSP的解决方案生成一个封闭的城市序列，其中精确地访问每个城市一次所经过的总距离是可能路径中最小的。当一对节点之间的距离不对称时，TSP的非对称版本确定节点的最佳序列。例如，从节点A到节点B的“距离”可能不同于从节点B到节点A的“距离”。通过使用非对称TSP解决改进的最优盒，盒设计模块可以找到导致在盒的表位之间的连接处的呈递评分降低的盒序列。非对称TSP的解决方案指示治疗性表位的序列，所述治疗性表位对应于表位应在盒中串联以最大限度地减少盒连接处的连接表位呈递评分的顺序。通过这种方法确定的盒序列可以产生这样的序列，其连接表位的呈递显著更少，同时可能需要比随机采样方法显著更少的计算资源，尤其是当生成的候选盒式序列的数量很大时。用于优化盒设计的不同计算方法和比较的说明性实例在美国专利号10,055,540、美国申请公开号US20200010849A1和国际专利申请公开WO/2018/195357和WO/2018/208856中更详细描述，所述专利各自出于所有目的通过引用整体并入本文。

[0444] 本领域技术人员可以选择用于包括在共享抗原疫苗中的共享(新)抗原序列和用这种疫苗治疗的合适患者，例如，如美国申请号17/058,128中所述，所述申请出于所有目的

通过引用并入本文。候选共享(新)抗原的质谱(MS)验证可以作为选择过程的一部分执行。

[0445] 盒设计模块还可以通过考虑疫苗中编码的额外蛋白质序列来生成盒序列。例如，用于生成编码串联的T细胞表位的序列的盒设计模块可以考虑已经由疫苗中存在的额外蛋白质序列(例如，全长蛋白质序列)编码的T细胞表位，诸如通过去除已经由候选序列列表中的额外蛋白质序列编码的T细胞表位。

[0446] 盒设计模块还可以通过考虑序列的大小来生成盒序列。不希望受理论的束缚，一般来讲，增加的盒大小会对疫苗方面产生负面影响，诸如疫苗生产和/或疫苗功效。在一个实例中，盒设计模块可以考虑重叠序列，诸如重叠的T细胞表位序列。一般来讲，含有重叠T细胞表位序列的单个序列(也称为“框架”)比单独连接单个T细胞表位序列更高效，因为它减少了编码多个肽所需的序列大小。因此，在说明性实例中，用于生成编码串联的T细胞表位的序列的盒设计模块可以考虑扩展候选T细胞表位以编码一个或多个额外的T细胞表位的成本/益处，诸如确定呈递额外的T细胞表位的MHC在额外人群覆盖中获得的益处与增加序列大小的成本。

[0447] 盒设计模块还可以通过考虑验证的表位生成的免疫响应的刺激幅度来生成盒序列。

[0448] 盒设计模块还可以通过考虑编码的表位在群体中的呈递来生成盒序列，例如，至少一种免疫原性表位在一定比例的群体中由至少一种HLA呈递，例如被至少85％、90％或95％的群体呈递(例如，四个主要种族群体，即欧洲人(EUR)、非裔美国人(AFA)、亚洲和太平洋岛民(APA)和西班牙裔(HIS)中的HLA-A、HLA-B和HLA-C基因)。作为说明性非限制性实例，盒设计模块还可以生成盒序列，使得至少一种HLA存在于至少85％、90％或95％的群体中，所述群体呈递至少一种验证的表位或呈递至少4、5、6或7种预测表位。

[0449] 盒设计模块还可以通过考虑提高潜在安全性的其他方面来生成盒序列，诸如限制编码可能存在安全风险的功能蛋白、功能蛋白结构域、功能蛋白亚基或功能蛋白片段或限制编码所述蛋白质的可能性。在一些情况下，盒设计模块可以限制编码的肽的序列大小，使得小于翻译的对应全长蛋白质的50％、49％、48％、47％、46％、45％、45％、43％、42％、41％、40％、39％、38％、37％、36％、35％、34％或33％。在某些情况下，盒设计模块可以限制编码的肽的序列大小，使得单个连续序列小于翻译的对应全长蛋白质的50％，但多于一个序列可能来自相同的翻译的对应全长蛋白质，并且一起编码超过50％。在说明性实例中，如果含有重叠T细胞表位序列的单个序列(“框架”)大于翻译的对应全长蛋白质的50％，则所述框架可以分成多个框架(例如，f1、f2等)，使得每个框架都小于翻译的对应全长蛋白质的50％。盒设计模块还可以限制编码的肽的序列大小，使得单个连续序列小于翻译的对应全长蛋白质的49％、48％、47％、46％、45％、45％、43％、42％、41％、40％、39％、38％、37％、36％、35％、34％或33％。当从相同的基因编码多个框架时，多个框架可以具有彼此重叠的序列，换句话说，每个框架分别编码相同的序列。当从相同基因编码多个框架时，可以将来源于相同基因的两个或更多个核酸序列排序，使得如果第二核酸序列跟随(紧跟或不紧跟)对应基因中的第一核酸序列，则第一核酸序列不能被第二核酸序列紧跟或不能与第二核酸序列连接。例如，如果相同基因中有3个框架(f1、f2、f3，按氨基酸位置的递增顺序排列)：

[0450] -不允许以下盒排序：

[0451] ○f2紧跟f1

[0452] ○f3紧跟f2

[0453] ○f3紧跟f1

[0454] -允许以下盒排序：

[0455] ○f2紧跟f3

[0456] ○f1紧跟f2

[0457] XIII.示例性计算机

[0458] 计算机可以用于本文所述的任何计算方法。本领域的技术人员将认识到计算机可以具有不同的结构。计算机的实例是本领域技术人员已知的，例如在美国专利号10,055,540、美国申请公开号US20200010849A1和国际专利申请公开WO/2018/195357和WO/2018/208856中更详细描述的计算机，所述专利各自出于所有目的通过引用整体并入本文。

[0459] XIV.实施例

[0460] 以下是用于进行本发明的特定实施方案的实施例。实施例仅出于说明性目的而提供，并且不意图以任何方式限制本发明的范围。已努力确保关于所用数值(例如数量、温度等)的准确性，但当然应虑及一些实验误差和偏差。

[0461] 除非另外指示，否则本发明的实施将采用属于本领域的技能的常规蛋白质化学、生物化学、重组DNA技术和药理学方法。此类技术充分说明于文献中。参见例如T.E.Creighton,Proteins:Structures and Molecular Properties(W.H.Freeman and Company,1993);A.L.Lehninger,Biochemistry(Worth Publishers,Inc.,当前版本);Sambrook等,Molecular Cloning:A Laboratory Manual(第2版,1989);Methods In Enzymology(S.Colowick和N.Kaplan编,Academic Press,Inc.);Remington's Pharmaceutical Sciences,第18版(Easton,Pennsylvania:Mack Publishing Company,1990);Carey和Sundberg Advanced Organic Chemistry第3版(Plenum Press)第A和B卷(1992)。

[0462] XIV.A.ChAd抗原盒递送载体

[0463] 在一个实例中，黑猩猩腺病毒(ChAdV)被工程化成抗原盒的递送载体。在另一实例中，全长ChAdV68载体是基于缺失E1(核苷酸457至3014)和E3(核苷酸27,816-31,332)序列的AC_000011.1(来自美国专利6083716的序列2)合成的。插入受CMV启动子/增强子控制的报告基因以替代缺失的E1序列。将该克隆转染到HEK293细胞中并不产生感染性病毒。为了确认野生型C68病毒的序列，从ATCC获得分离株VR-594，传代，然后独立测序(SEQ ID NO:10)。将AC_000011.1序列与野生型ChAdV68病毒的ATCC VR-594序列(SEQ ID NO:10)进行比较时，鉴定出6个核苷酸差异。在一个实例中，基于AC_000011.1生成了修饰的ChAdV68载体，所述AC_000011.1的对应ATCC VR-594核苷酸在五个位置处被取代(ChAdV68.5WTnt SEQ ID NO:1)。

[0464] 在另一实例中，基于AC_000011.1生成了修饰的ChAdV68载体，所述AC_000011.1缺失E1(核苷酸577至3403)和E3(核苷酸27,816-31,332)序列并且对应的ATCC VR-594核苷酸在四个位置处被取代。插入受CMV启动子/增强子控制的GFP报告子(ChAdV68.4WTnt.GFP;SEQ ID NO:11)或模型新抗原盒(ChAdV68.4WTnt.MAG25mer;SEQ ID NO:12)以替代缺失的E1序列。

[0465] 在另一实例中，基于AC_000011.1生成了修饰的ChAdV68载体，所述AC_000011.1缺

失E1(核苷酸577至3403)和E3(核苷酸27,125-31,825)序列并且对应的ATCC VR-594核苷酸在五个位置处被取代。插入受CMV启动子/增强子控制的GFP报告子(ChAdV68.5WTnt.GFP;SEQ ID NO:13)或模型新抗原盒(ChAdV68.5WTnt.MAG25mer;SEQ ID NO:2)以替代缺失的E1序列。

[0466] 在另一实例中,基于AC_000011.1生成了抗原表达系统的经修饰的ChAdV68载体("chAd68-空-E4缺失"SEQ ID NO:29369),所述AC_000011.1缺失E1(核苷酸577至3403)、E3(核苷酸27,125-31,825)和E4区域(核苷酸34,916至35,642)序列并且对应的ATCC VR-594(独立测序的全长VR-594C68 SEQ ID NO:10)核苷酸在五个位置处被取代。对应的ATCC VR-594核苷酸在五个位置处被取代的全长ChAdV68 AC_000011.1序列被称为"ChAdV68.5WTnt"(SEQ IDNO:1)。插入受CMV启动子/增强子控制的抗原盒以替代缺失的E1序列。

[0467] 相关载体描述如下:

[0468] -全长ChAdVC68序列"ChAdV68.5WTnt"(SEQ ID NO:1);AC_000011.1序列,其对应的ATCC VR-594核苷酸在五个位置处被取代。

[0469] -ATCC VR-594C68(SEQ ID NO:10);独立测序的;全长C68

[0470] -ChAdV68.4WTnt.GFP(SEQ ID NO:11);缺失E1(核苷酸577至3403)和E3(核苷酸27,816-31,332)序列的AC_000011.1;对应的ATCC VR-594核苷酸在四个位置处被取代;插入受CMV启动子/增强子控制的GFP报告子以替代缺失的E1

[0471] -ChAdV68.4WTnt.MAG25mer(SEQ ID NO:12);缺失E1(核苷酸577至3403)和E3(核苷酸27,816-31,332)序列的AC_000011.1;对应的ATCC VR-594核苷酸在四个位置处被取代;插入受CMV启动子/增强子控制的模型新抗原盒以替代缺失的E1

[0472] -ChAdV68.5WTnt.GFP(SEQ ID NO:13);缺失E1(核苷酸577至3403)和E3(核苷酸27,125-31,825)序列的AC_000011.1;对应的ATCC VR-594核苷酸在五个位置处被取代;插入受CMV启动子/增强子控制的GFP报告基因以替代缺失的E1

[0473] -ChAd68-空-E4缺失(SEQ ID NO:29369):AC_000011.1,其缺失E1(核苷酸577至3403)、E3(核苷酸27,125-31,825)和E4区域(核苷酸34,916至35,642)序列并且对应的ATCC VR-594(独立测序的全长VR-594C68 SEQ ID NO:10)核苷酸在五个位置处被取代

[0474] -ChAdV68-GAG(全长ChAd68-CMV-SIVGag-SV40聚A;SEQ ID NO:29371);缺失E1(核苷酸577至3403)和E3(核苷酸27,125-31,825)序列的AC_000011.1;对应的ATCC VR-594核苷酸在五个位置处被取代;插入受CMV启动子/增强子控制的全长密码子优化的SIV_{SME543} GAG以替代缺失的E1

[0475] 在293F细胞中的腺病毒生产

[0476] ChAdV68病毒生产在293F细胞中进行,所述细胞在8%CO_2的孵育箱内的293 FreeStyle™(ThermoFisher)培养基中生长。在感染当天,将细胞稀释至10^6个细胞/mL,具有98%的存活力,并且每次生产运行在1L摇瓶(Corning)中使用400mL。每次感染使用目标MOI大于3.3的4mL三级病毒原液。将细胞孵育48-72小时,直到如通过台盼蓝测量的存活力小于70%。然后将感染的细胞通过离心、全速台式离心机收集,并且在1XPBS中洗涤,重新离心,然后重悬在20mL的10mM Tris pH7.4中。将细胞沉淀通过冻融3次来裂解,并通过在4,300Xg下离心5分钟来澄清。

[0477] 通过CsCl离心进行腺病毒纯化

[0478] 将病毒DNA通过CsCl离心纯化。进行了两次不连续的梯度运行。第一次是从细胞组分中纯化病毒，而第二次是进一步精进与细胞组分的分离并且将缺陷型颗粒与感染性颗粒分离。

[0479] 将10mL的1.2(26.8g CsCl溶解在92mL的10mM Tris pH 8.0中)CsCl添加到异质同晶聚合物管中。然后使用移液器小心地添加8mL的1.4CsCl(53g CsCl溶解在87mL的10mM Tris pH 8.0中)，递送至管底部。将澄清的病毒小心地层铺在1.2层的顶部。如果需要，则添加更多10mM Tris以平衡管。然后将管置于SW-32Ti转子中，并在10℃下离心2小时30分钟。然后将管移至层流柜并使用18号针头和10mL注射器拉动病毒条带。注意不要去除污染的宿主细胞DNA和蛋白质。然后将条带用10mM Tris pH 8.0稀释至少2倍，并且如前所述在不连续梯度上分层。如前所述进行运行，不同之处是这次运行进行过夜。第二天，小心地拉动条带以避免拉动任何缺陷型颗粒条带。然后使用Slide-a-Lyzer™盒(Pierce)针对ARM缓冲液(20mM Tris pH 8.0、25mM NaCl、2.5％甘油)透析病毒。这执行了3次，每次缓冲液交换1小时。然后将病毒等分以在-80℃下储存。

[0480] 腺病毒病毒测定

[0481] VP浓度通过使用基于1.1x 10^{12}个病毒颗粒(VP)的消光系数的OD 260测定执行，相当于OD260 nm处的吸光度值为1。在病毒裂解缓冲液(0.1％ SDS、10mM Tris pH 7.4、1mM EDTA)中制备腺病毒的两种稀释液(1:5和1:10)。在两种稀释度下一式两份测量OD，并且VP浓度/mL通过将OD260值X稀释因子X1.1x 10^{12}VP相乘来测量。

[0482] 感染单位(IU)滴度通过病毒原液的有限稀释测定来计算。病毒最初在DMEM/5％NS/1X PS中稀释100倍，随后接着使用10倍稀释液稀释至1x 10^{-7}。然后将100μL的这些稀释液添加到293A细胞中，所述细胞在至少一小时前以3e5个细胞/孔接种在24孔板中。这是一式两份进行的。将板在37℃下在CO2(5％)孵育箱中孵育48小时。然后将细胞用1XPBS洗涤，然后用100％冷甲醇(-20℃)固定。然后将板在-20℃下孵育最少20分钟。将孔用1XPBS洗涤，然后在室温下在1XPBS/0.1％ BSA中封闭1小时。将兔抗Ad抗体(Abcam,Cambridge,MA)以1:8,000稀释度添加到封闭缓冲液(0.25ml/孔)中，并且在室温下孵育1小时。每个孔用0.5mL PBS洗涤4次。每个孔添加稀释1000倍的HRP缀合的山羊抗兔抗体(Bethyl Labs,Montgomery Texas)，并且在最后一轮洗涤前孵育1小时。执行5次PBS洗涤，并且使用含有0.01％ H_2O_2的Tris缓冲盐水中的DAB(二氨基联苯胺四盐酸盐)底物(0.67mg/mL DAB于50mM Tris pH 7.5、150mM NaCl中)来显影板。在计数前显影孔5分钟。在10X物镜下使用每个视野产生4-40个染色细胞的稀释度对细胞进行计数。使用的视野是0.32mm^2网格，其每个视野相当于24孔板上的625个网格。感染性病毒数/mL可以通过将每个网格的染色细胞数乘以每个视野的网格数乘以稀释因子10来确定。类似地，在使用GFP表达细胞时，可以使用荧光而不是衣壳染色来确定每毫升表达GFP的病毒颗粒的数量。

[0483] XIV.B.基于甲病毒的抗原盒递送SAM载体

[0484] 通过缺失位于26S亚基因组启动子的3'的VEE的结构蛋白(缺失VEE序列7544至11,175;基于Kinney等人1986进行编号;SEQ ID NO:6)，由自复制委内瑞拉马脑炎(VEE)病毒(Kinney,1986,Virology152:400-413)生成用于抗原表达系统的RNA甲病毒骨架。为了生成自扩增mRNA("SAM")载体，缺失的序列被抗原序列替换。含有20种模型抗原的代表性SAM载体是"VEE-MAG25mer"(SEQ ID NO:4)。以具有MAG25mer盒的所述抗原盒为特征的载体可以

被本文所述的盒和/或全长蛋白替换，诸如全长SIV GAG，参见下文。

[0485] 全长密码子优化的SIV_{SME543} GAG核苷酸序列；

[0486] ATGGGAGCCAGGAACAGTGTGCTCTCAGGCAAGAAGGCAGATGAGCTGGAGAAGATCAGGCTGAGACCC
AATGGCAAGAAGAAGTACATGCTGAAGCATGTGGTCTGGGCAGCCAATGAGCTGGACAGGTTTGGCCTGGCAGAGTC
CCTGCTGGACAACAAGGAGGGCTGCCAGAAGATCCTGTCAGTGCTGGCCCCCCTGGTGCCCACTGGCTCAGAGAACC
TGAAGAGCCTCTACAACACAGTGTGTGTGATTTGGTGCATCCATGCAGAGGAGAAGGTGAAGCACACTGAGGAGGCC
AAGCAGATTGTGCAGAGGCACCTGGTGGTGGAGACTGGCACAGCTGACAAGATGCCAGCCACCTCCAGGCCCACAGC
ACCCCCCTCTGGCAGGGGGGGCAACTACCCAGTCCAGCAAGTGGGGGGCAACTATGTGCACCTGCCCCTGAGCCCCA
GAACCCTGAATGCCTGGGTCAAGCTGGTGGAGGAGAAGAAGTTTGGAGCAGAGGTGGTGCCTGGCTTCCAGGCCCTG
TCAGAGGGATGCACTCCCTATGACATCAACCAGATGCTGAACTGTGTGGGAGAGCACCAGGCAGCAATGCAGATCAT
CAGGGAGATCATCAATGAGGAGGCTGCTGACTGGGACCTCCAGCACCCCCAGCCTGGACCCCTCCCTGCAGGCCAGC
TGAGGGAGCCCAGAGGGAGTGACATAGCAGGCACCACCTCCACAGTGGAGGAGCAGATCCAGTGGATGTACAGGCAG
CAGAACCCCATCCCTGTGGGCAACATCTACAGGAGGTGGATCCAGCTGGGCCTCCAGAAGTGTGTCAGGATGTACAA
CCCAACCAACATCCTGGATGTGAAGCAGGGCCCCAAGGAGCCCTTCCAGTCTTATGTGGACAGGTTCTACAAGAGCC
TGAGAGCTGAGCAGACAGACCCTGCTGTGAAGAACTGGATGACCCAGACACTGCTGATCCAGAATGCCAATCCTGAC
TGCAAGCTGGTGCTGAAGGGGCTGGGGATGAATCCAACCCTGGAGGAGATGCTGACAGCCTGCCAGGGCATTGGGGG
ACCTGGACAGAAGGCCAGGCTCATGGCAGAGGCTCTCAAGGAGGCCCTCAGACCAGACCAGCTGCCATTTGCTGCTG
TGCAGCAGAAGGGCCAGAGGAGGACCATCAAGTGCTGGAACTGTGGCAAGGAGGGCCACTCTGCCAGGCAGTGCAGA
GCCCCCAGGAGGCAGGGCTGCTGGGGCTGTGGAAAGACAGGCCATGTGATGGCCAAGTGCCCAGAGAGGCAGGCAGG
CTTCCTGGGCTTTGGCCCCTGGGGCAAGAAGCCAAGAAACTTCCCCATGGCCCAGATGCCCCAGGGCCTGACCCCCA
CAGCCCCCCCAGAGGACCCAGCTGTGGACCTGCTGAAGAACTACATGAAGATGGGCAGGAAGCAGAGGGAGAACAGG
GAGAGACCCTACAAGGAGGTGACTGAGGACCTGCTGCACCTGAACTCCCTGTTTGGGGAGGACCAGTGA

[0487] 全长SIV_{SME543} GAG氨基酸序列；

[0488] MGARNSVLSGKKADELEKIRLRPNGKKKYMLKHVVWAANELDRFGLAESLLDNKEGCQKILSVLAPLV
PTGSENLKSLYNTVCVIWCIHAEEKVKHTEEAKQIVQRHLVVETGTADKMPATSRPTAPPSGRGGNYPVQQVGGNY
VHLPLSPRTLNAWVKLVEEKKFGAEVVPGFQALSEGCTPYDINQMLNCVGEHQAAMQIIREIINEEAADWDLQHPQ
PGPLPAGQLREPRGSDIAGTTSTVEEQIQWMYRQQNPIPVGNIYRRWIQLGLQKCVRMYNPTNILDVKQGPKEPFQ
SYVDRFYKSLRAEQTDPAVKNWMTQTLLIQNANPDCKLVLKGLGMNPTLEEMLTACQGIGGPGQKARLMAEALKEA
LRPDQLPFAAVQQKGQRRTIKCWNCGKEGHSARQCRAPRRQGCWGCGKTGHVMAKCPERQAGFLGFGPWGKKPRNF
PMAQMPQGLTPTAPPEDPAVDLLKNYMKMGRKQRENRERPYKEVTEDLLHLNSLFGEDQ*。

[0489] 体外转录生成RNA

[0490] 对于体内研究，根据以下方案生成SAM RNA，如下所示：

[0491] (1)通过TriLink Biotechnologies进行纯化并且用Enzymatic Cap1封端，或

[0492] (2)生成为“AU-SAM”载体。将缺少规范3’二核苷酸GG的修饰的T7 RNA聚合酶启动子(TAATACGACTCACTATA)添加至SAM载体的5’端以生成体外转录模板DNA(SEQ ID NO:29370;缺失7544至11,175,无插入的抗原盒)。反应条件如下所述：

[0493] -1x转录缓冲液(40mM Tris-HCL[pH7.9]、10mM二硫苏糖醇、2mM亚精胺、0.002％Triton X-100和27mM氯化镁)，使用的最终浓度为1x T7 RNA聚合酶混合物(E2040S)；0.025mg/mL DNA转录模板(通过限制性消化线性化)；8mM CleanCap试剂AU(目录号N-7114)

和各10mM的ATP、三磷酸胞苷(CTP)、GTP和三磷酸尿苷(UTP)

[0494] -将转录反应物在37℃下孵育2小时,并在37℃下用DNA酶I缓冲液中的最终2U DNA酶I(AM2239)/0.001mg DNA转录模板处理1小时。

[0495] -将SAM通过RNeasy Maxi(QIAGEN,75162)纯化

[0496] 除了共转录添加5'帽结构外,还可以在转录后使用含有mRNA 2'-O-甲基转移酶和S-腺苷甲硫氨酸的牛痘加帽系统(NEB目录号M2080)以酶促方式添加7-甲基鸟苷或相关的5'帽结构。

[0497] XV.ChAdV增强方案在NHP中生成T细胞响应

[0498] 疫苗研究在Mamu A01印度恒河猴(NHP)中进行,以使用表达全长SIV抗原GAG的ChAdV68载体和SAM载体证明免疫原性。图1说明了疫苗接种策略,其中在第0周和第33周进行ChAdV68疫苗接种(1e12 vp/动物),并且在第4周和第17周进行基于VEE甲病毒的SAM疫苗接种(AU-SAM帽)。

[0499] 免疫接种

[0500] 使用(1)ChAdV68-GAG(SEQ ID NO:29371)或(2)300ug的基于VEE的SAM载体(SEQ ID NO:6)的$1x10^{12}$个病毒颗粒(每次注射$5x10^{11}$个病毒颗粒)对Mamu-A*01印度恒河猴进行双侧肌内免疫,所述载体也经工程化以含有表达全长SIV抗原GAG的盒。ChAdV68和SAM的疫苗加强在初免疫苗接种后的指定时间施用。

[0501] 免疫监测

[0502] 将PBMC使用淋巴细胞分离培养基(LSM,MP Biomedicals)和LeucoSep分离管(Greiner Bio-One)在初免疫苗接种后的指定时间分离,并且重悬在含有10% FBS和青霉素/链霉素的RPMI中。使用碘化丙锭染色在Attune NxT流式细胞仪(Thermo Fisher)上对细胞进行计数以排除死亡和细胞凋亡的细胞。然后将细胞调整至适当的活细胞浓度用于后续分析。对于研究中的每只猴子,使用ELISpot或流式细胞术方法测量T细胞响应。通过使用离体酶联免疫斑点(ELISpot)分析测量细胞因子(诸如IFN-γ)的诱导,从PBMC监测T细胞对疫苗中编码的6个不同的SIV GAG表位池(6个重叠肽池,每个20个肽,15个氨基酸)的响应。ELISpot分析根据ELISpot协调指南{DOI:10.1038/nprot.2015.068}使用猴IFNg ELISpotPLUS试剂盒(MABTECH)进行。将200,000个PBMC与10uM的所示肽在涂覆IFNg抗体的96孔板中孵育16小时。使用碱性磷酸酶显影斑点。反应计时10分钟,并且通过在自来水下运行板来终止。使用AID vSpot Reader Spectrum对斑点进行计数。对于ELISpot分析,将饱和度大于50%的孔记录为"多得无法计数"。将重复孔偏差大于10%的样品从分析中排除。然后使用以下公式对斑点计数进行校正以获得孔汇合度:斑点计数+2x(斑点计数x%汇合度/[100%-%汇合度])。通过从抗原刺激孔中减去负肽刺激孔中的斑点计数来校正负本底。最后,将标记为多得无法计数的孔设置为观察到的最高校正值,四舍五入到最接近的百位。

[0503] 结果

[0504] 评估了疫苗接种后外周血单核细胞(PBMC)中的抗原特异性细胞免疫响应。如图2所示,在ChAdV初免后观察到免疫响应,以及在每个SAM加强剂量后观察到免疫响应增加。值得注意的是,在第33周施用ChAdV加强剂量后,也观察到免疫响应增强。具体地,如呈现了6只灵长类动物的每个池的平均值的图2A所示,观察到第34周的平均值相对于第33周增加4.0倍,并且第35周的平均值相对于第33周增加6.6倍(第33周平均SFC:2497;第34周平均

SFC:9930;第35周平均SFC:16445)。图2B呈现了非常高响应的个体灵长类动物#1-3(大于18,000SFC/1e6 PBMC)的结果,在第35周相对于第33周分别增加了6.6、4.0和16.6倍。图2C呈现了高响应的个体灵长类动物#4-6(小于18,000SFC/1e6 PBMC)的结果,灵长类动物#4和#6在第35周相对于第33周分别增加了5.8倍和4.7倍,并且灵长类动物#4在第34周相对于第33周增加了54.8倍。结果表明疫苗接种策略,特别是ChAdV加强剂量,导致了稳健的抗原特异性免疫响应。

[0505] XVI.ChAdV中和抗体在NHP中随时间减少

[0506] 疫苗研究在Mamu A01印度恒河猴(NHP)中进行,以使用表达SIV抗原的载体证明免疫原性。三组NHP用ChAdV68.5WTnt.MAG25mer(SEQ ID NO:2)和检查点抑制剂抗CTLA-4抗体伊匹单抗(第5组和第6组)或不用检查点抑制剂(第4组)进行免疫。抗体静脉内(第5组)或皮下(第6组)施用。在第0周和第32周施用ChAdV68疫苗接种(1e12vp/动物),在其之间额外施用基于VEE甲病毒的SAM疫苗接种。

[0507] 中和抗体滴度

[0508] 在治疗后指示的时间点收集血清。将NHP血清在56℃下热灭活30分钟。在DMEM中将血清稀释两倍,从1:20降至1:10,240。将体积为50uL的ChAdV-LacZ病毒(7e6 IU/mL)与等体积的稀释血清混合,并且在37℃下孵育1小时。然后将20uL的这种混合物添加到预先接种7e4 HEK293细胞/mL的96孔板的每个孔中。在每个稀释度下执行一式三份的孔,并且将细胞在37℃下孵育过夜。第二天,使用Galacto-Star™(Applied Biosystems)化学发光试剂盒测量β-半乳糖苷酶活性。相对于在细胞培养基中稀释的未中和病毒的表达,测量与稀释血清混合的病毒的中和百分比。中和滴度(NT50)计算为中和病毒50%的血清的最小稀释度,并且这是使用非线性回归曲线拟合计算的。

[0509] 结果

[0510] 在疫苗接种方案的过程期间评估了ChAdV中和抗体滴度,并且结果呈现在表2中。在疫苗接种前(第0天),NT50值表明预先存在的ChAdV中和抗体没有有意义的存在。疫苗接种后,在所有测试组中,ChAdV中和抗体滴度在第7周增加,并在第24周进一步增加。在第32周,ChAdV治疗组的中和滴度然后下降到未接受IPI的组的第24周值的大约三分之一(第4组;NT50为2527对比846),并且下降至共施用IPI的组的大约一半(第5/6组;NT50为1184对比601)。值得注意的是,在第32周施用ChAdV加强后五周,同一组的滴度分别再次增加至5664和5096。结果表明,在ChAdV初免剂量后预期的初始增加后,针对ChAdV的同源抗体逐渐减弱。

[0511] 表2-疫苗接种后的ChAdV中和抗体滴度

[0512]

治疗后时间	ChAd 疫苗接种	第 4 组: ChAd 无 IPI	第 5/6 组: ChAd 加 IPI
第 0 天	初免	<20	<20
第 7 周	n/a	1004	476
第 24 周	n/a	2527	1184
第 32 周	加强	846	601
第 37 周	n/a	5664	5096

[0513] XVII.ChAdV增强方案在人类受试者中生成T细胞响应

[0514] 方法

[0515] 个性化的新抗原癌症疫苗(“GRANITE”)与免疫检查点阻断在晚期癌症患者中组合施用。GRANITE异源初免/加强疫苗方案包括(1)用作初免疫苗接种[GRT-C901]的ChAdV和(2)在LNP中配制的SAM,其用于GRT-C901后的加强疫苗接种[GRT-R902]。ChAdV载体基于具有SEQ ID NO:1的序列的经修饰ChAdV68序列,所述SEQ ID NO:1的序列具有E1(核苷酸577至3403)缺失和E3(核苷酸27,125-31,825)缺失。SAM载体基于具有SEQ ID NO:6中列出的核酸序列的RNA甲病毒骨架。GRT-C901和GRT-R902均表达相同的20种个性化新抗原以及两种通用CD4 T细胞表位(PADRE和破伤风类毒素)。肿瘤用于全外显子组和转录组测序以检测体细胞突变,并且血液用于HLA分型和种系外显子组变异的检测/减去以使用EDGE算法为受试者生成个性化的新抗原盒。

[0516] GRANITE方案经由双侧肌内注射(例如,在每个三角肌中)与免疫检查点阻断(特别是SC IV纳武单抗)组合来施用疫苗。

[0517] GRT-C901是复制缺陷型、E1和E3缺失的腺病毒载体,基于黑猩猩腺病毒68。载体含有编码20个新抗原以及两个通用CD4 T细胞表位(PADRE和破伤风类毒素)的表达盒。GRT-C901以5×10^{11}vp/mL配制在溶液中,并在相对的三角肌中的2个双侧疫苗注射部位中的每一个处肌内注射1.0mL。受试者的GRT-C901盒(表3A)和确定的HLA(表3B)评估如下。

[0518] 表3A-GRANITE盒

[0519]

患者	盒序列
G1	LAEQVQKADEFIQANATNKLTVIAEQGDGCSGHKNQNQIDTLQEV RLRFLGQLSQDRDFIMTLNTSLHRSWWMENDTCVCSTVAEFSHQ CSHAGGRPGNWPMNTIAEAVIEMVNRGQIQITINGFVSQKQIADR AAHNESPGNNVVDISVSEEAEVDVREREIQRDREPKRARDLQLQA EQKGLTEVIQSLKMENEEFKKQGRVKDLESLFHQSEVELAAALSD KGKISRHEMLQVLHLMVGVQVTEEQLGLFEGDEYATLMMQCKE GAQKEGLMCLAFKNDATEILHSHVVKPVPAHPSGIGWLITFCSKG EDCLTSQTRLRLSHNMKLSISYLREYAKANDWLQFIIHLNTAQDEL VTFSDELAQLYHHVCLCAPFPAWSAFGKEEHDEALKNTWNLHTD AYLLYTPSQIVLTAILSSASRAGGPKVDIDVPDVNFEAPEGKLKGPK

[0520]

	FNFDVAGYIVGANTETYLLEKSRAIREEQKLPEADVAVLRTFLRQQ ASILS
G3	SKELVEKSEAVRKMSSSGKLWKPRQMASMGTLAFHEYGRPFLIIK DQDRKDTRIAKRMAMLIYTDFLCMAPISFFYLSLLRMYLSPPCIHC LGPIKLELLAVKLLIERATNQGWMLRVKWLYHLAVRASSDGEGT MSLPASVDGSPVSPSCSPPSPSALIQEMEERLWEKEQEVAERYHFL EERLARLGLTEVYQWYLDLFFGMHVQEYGSDWPPPNTRRVYISY LFIYQSLDAIDGRQARRTNSSSPLGNWVQQQAYGMLAGIPVTICT YPFVFMDARRVPPKDLRVKKNLKKFRYVKLVKKNRQNGMHTLY LLDIKVKEQSLEPNYADTLISQESFEKKGFLSAPQSLGDGGLVCPY MEFLKNENNELPKLQWSLSLAEKSILHEECIKSPVVETVPAPPDPL EPPSLPAQRPGPPTPAAAHSVKGGYQESSLLRVLVTLNDLEPKAAK QGKDLTVLDHTLPTALTSPELSSWHSSFVAVTEEGTKAAAATGIGF TVT
G8	APRLGSYSGTTILHLLNSTSNNLYLPPACSSSSSSLFLAVVARRLGR GSVSYQHKFQDDDQTHVKGSLKAGFFGTPPAVISQMLLLDSPHKE PIRLRYKLPKVRERHFSESTYIDNALSRLTLGNEHKVLLTGTPLQKT VEELFSLLHFLPYHRSQSSSSVLNNKSMDSINYPSDLSDQENAPPIL PRRPSEGLGPSPHLEVLLLSMRDMNITKLTSVDAPLFNAPGHDKM LSPNFDVHHTAMLTRGEQHLHKSLKLYQVIFKGEIGKGNLGGIAA QPHVNGGPLYSKNSISIAPPPPPMAPAPPARTSRGQVSERANEAGG QVGAKMKWTNLLYKAGGVKIRDDERLLQNANFDLRTSQSVRAM LAYNMRDQHLMTEYKLVVVGADGVGKSALTIQLIQGIALVMRVD GAVSSCFRQLFLAQQRGSRKERQVYSKALNRLFGVEASGRRKDL AKRLLVGKSDSVDAEKSMLSKLEQSAERCVSTLLNLIQTKVNYVV QE
G11	AAGPSCALKAGKTASGAGEGGALSVDSESHWVAPEHTKRSVPQD PDFAEVGTYWPVNSDQAVLELDLGWPLTFGFVQVLPDADTLLHLP RKVLQMDFLVHNSPRSTPGFILDTTSNKQTPQKKQLREHQLRLMS ARACYERYSGNQVLFCMQSALAKERRELRQAQREAEMDSIPIKSD LVNEEATGTVPCIPGAAQALHKNCPHIVVGTPGHILALARNKSLNL AAQQGEPHPEFSVDSPSPGLNMELEVLLIVLVMMSLVLGKFVPVN WEPPQADSALRMPSTGSIAPCVPPLSTSLLFRFNFLAPELPAFSEFS TSETMGHSTVTSAQKSSQIGNSQLLKRHVQRTECQNSQSRKEQVN TELKAVPAQEVKMNEPKLSRPTAAIVLKGEKGDRGPKGYEWQEV EKGSLPRVLLQTPAAPQRCTTIHYNYMCNSSYMGGMNRRPILTIC KECGKAFAVSSHLSRHERIHTGEKIASGSELSESAFSYFITACAVIILT

[0521] 表3B-GRANITE受试者HLA

[0522]

患者	HLA-A	HLA-A	HLA-B	HLA-B	HLA-C	HLA-C
G1	A*02:01	A*32:01	B*18:01	B*44:05	C*02:02	C*07:04
G3	A*02:01	A*11:01	B*27:02	B*35:01	C*04:01	C*02:02
G8	A*03:01	A*33:01	B*42:01	B*53:01	C*04:01	C*17:01
G11	A*01:01	A*02:01	B*07:02	B*15:01	C*03:03	C*07:02

[0523] GRT-R902是来源于甲病毒的SAM载体。GRT-R902载体编码病毒蛋白和RNA扩增所需的5'和3'RNA序列，但不编码结构蛋白。SAM载体在包括4种脂质的LNP中配制：可电离的氨基脂质、磷脂酰胆碱、胆固醇和基于PEG的外壳脂质，以封装SAM并形成LNP。GRT-R902载体含有

与分别用于每个患者的GRT-C901相同的新抗原表达盒。将GRT-R902配制成1mg/mL的溶液，并在相对的三角肌中的2个双侧疫苗注射部位中的每一个处进行肌内注射(优选三角肌，可以使用每一侧上的臀肌[背侧或腹侧]或股直肌)。加强疫苗接种部位尽可能靠近初免疫苗接种部位。注射体积基于待施用的剂量。受试者G1和G3中的SAM剂量水平为30μg，对于受试者G8为100μg，并且对于受试者G11为300μg。剂量水平量明确指代SAM载体的量，即，它不指代其他组分，诸如LNP。LNP:SAM的比率约为24:1。因此，LNP的剂量分别为720μg、2400μg或7200μg。

[0524] 纳武单抗是人单克隆IgG4抗体(参见SEQ ID NO:29522和29523)，其阻断PD-1与其配体PD-L1和PD-L2的相互作用。将纳武单抗配制成10mg/mL的溶液，并且通过0.2微米至1.2微米孔径的低蛋白结合在线过滤器以方案指定的剂量作为静脉内输注(480mg)施用。它不作为静脉内推注或团注注射施用。纳武单抗输注后立即用稀释剂冲洗以清洁管线。纳武单抗在每次疫苗接种(即，GRT-C901或GRT-R902中的每一种)后在同一天与或不与伊匹单抗一起施用。纳武单抗的剂量和途径基于美国食品和药物管理局批准的剂量和途径。

[0525] 评估了GRANITE受试者在疫苗接种后的免疫响应。对受试者执行抽血并收集PBMC。通过IFN-γELISpot评估T细胞响应。用于再刺激的肽如下所示。将外周血单核细胞(PBMC)以2x 10^5细胞/孔(离体)或10^5细胞/孔(IVS后)涂铺，并在最小表位肽(8-11聚体)或对照的存在下在涂覆有抗人干扰素-γ抗体的ELISpot板中刺激过夜。在加湿的5% CO_2、37℃孵育箱中孵育20小时后，取出细胞并根据标准方案显影ELISpot板。将数据报告为每10^6个脾细胞的斑点形成细胞(SFC)。

[0526] 离体IFNγELISpot

[0527] 通过离体ELISpot测定执行产生IFNg的T细胞的检测(Janetzki等人，2005)。简而言之，收获细胞，计数，并且以4x 10^6个细胞/ml(离体PBMC)或2x 10^6个细胞/ml(IVS扩增细胞)重悬在培养基中，并且在DMSO(VWR International)、植物凝集素-L(PHA-L;Sigma-Aldrich,Natick,MA,USA)、CEF肽池或同源肽的存在下在涂覆有抗人IFNg捕获抗体(Mabtech,Cincinatti,OH,USA)的ELISpot Multiscreen板(EMD Millipore)中培养。在5% CO_2、37℃的加湿孵育箱中孵育18-24小时后，收集上清液，从板中取出细胞，并且使用抗人IFNg检测抗体(Mabtech)、Vectastain亲和素过氧化物酶复合物(Vector Labs,Burlingame,CA,USA)和AEC底物(BD Biosciences,San Jose,CA,USA)检测膜结合的IFNg。使ELISpot板干燥、避光保存并寄到Zellnet Consulting,Inc.(Fort Lee,NJ,USA)用于标准化评价(Janetzki等人，2015)。数据以每百万个细胞的斑点形成单位(SFU)表示。

[0528] ELISpot上清液的MSD

[0529] 使用U-plex测定MSD U-PLEX生物标志物测定(Meso Scale Discovery Inc.,Rockville,MD,USA)对离体ELISpot上清液中分泌的IL-2、TNFa和颗粒酶B(GRZB)进行检测。根据制造商的说明书进行测定。使用每种细胞因子的已知标准品的系列稀释液来计算分析物浓度(pg/ml)。对于图形数据表示，低于标准曲线最小范围的值表示为零。

[0530] 三重荧光斑点

[0531] 根据制造商的说明书(MABtech)对患者样品的子集执行多重荧光斑点测定，以检测产生IFNg、IL-2和颗粒酶B的T细胞。简而言之，将患者PBMC解冻并且以2x 10^6个细胞/ml静置过夜。第二天，收集细胞，计数，并且以4x 10^6个细胞/ml重悬，然后进行涂铺。将预先涂

覆的荧光斑点板用PBS清洗并且用培养基封闭，然后添加患者特异性肽池和细胞(2x 10^5个PBMC/孔)。在5% CO_2、37℃的加湿孵育箱中孵育18-24小时后，使用抗人IFNg-BAM/抗BAM-490、生物素化抗人颗粒酶B/链霉亲和素550和抗人IL-2-WASP/抗WASP-650抗体对、然后用荧光增强剂来检测分泌的IFNg、IL-2和颗粒酶B。在AID iSpot读板器(Autoimmun Diagnostika)上对板进行成像和计数。数据以每百万个细胞的斑点形成单位(SFU)表示。

[0532] I类四聚体染色

[0533] HLAI类四聚体由Flex-T™生物素化HLA单体(BioLegend,San Diego,CA,USA)或MBL International(Woburn,MA,USA)生成，目标肽根据制造商的说明书通过UV交换负载到单体上。通过ELISA(BioLegend，根据制造商的说明书)评估通过UV交换的目标肽的成功负载。可替代地，生物素化单体在内部生成并与目标肽预先折叠。将肽-HLA(pHLA)单体使用荧光团缀合的链霉亲和素(来自BioLegend的PE；来自eBioscience,San Diego,CA,USA的APC；BV421)组装成四聚体。使用生成的荧光团标记的肽-MHC四聚体，对PBMC进行四聚体染色，以帮助识别抗原特异性T细胞。将患者PBMC样品解冻并重悬在含FAC缓冲液(PBS，10％FBS[体积/体积])的96孔V型底板中。染色前，将样品与50nM达沙替尼(Dasatinib)(Sigma Aldrich,St.Louis,MO,USA)一起在5％CO_2的37℃孵育箱中孵育30分钟。随后，将负载有感兴趣的目标肽的HLA I类荧光团标记的四聚体直接添加到每个对应的PBMC样品。在需要将多种四聚体添加到同一样品中的情况下，用等量的每种四聚体制成混合物，然后混合并同时添加。将样品在室温下在黑暗中孵育1小时。孵育后，将样品在FACS缓冲液中洗涤，并且用含有纯化的抗人HLA-A、B、C抗体、抗人CD8-PerCP-Cy5.5、抗人CCR7-PE-Cy7、抗人CD3-BV421(都来自BioLegend)、抗人CD45RA-SB702、抗人CD4-APC-eF780(都来自eBioscience)和Live/Dead Green(ThermoFisher)的抗体主要混合物染色。混合样品并且在4℃下在避光条件下在黑暗中孵育20分钟。将样品用FAC缓冲液洗涤一次，然后重悬在FAC缓冲液中，然后在BD LSRFortessa™流式细胞仪(BD Biosciences)上采集。

[0534] 流式细胞术分析

[0535] 使用FlowJo软件(FlowJo,LLC,Ashland,OR,USA)分析在流式细胞仪上采集的样品。人类样品的门控策略如下：淋巴细胞(SSC-A对比FSC-A)、单细胞(FSC-H对比FSC-A)、活细胞(FSC-A对比LD-AF488)、$CD3^+$细胞(FSC-A对比CD3-BV421)、$CD8^+Tet^+$(CD8-PerCP-Cy5.5对比四聚体-PE或四聚体-APC或对比四聚体-BV421)、记忆表型(CD45RA-SB702对比CCR7-PE-Cy7)。结果表示为阳性细胞群％(亲代频率)。数据显示为在指示处减去本底。

[0536] 表4-再刺激测定中的GRANITE患者G1肽

[0537]

肽序列(氨基酸)	肽池	
TVAEFSHQC	CD8 池	小型 1
VLTAILSSA	CD8 池	小型 1
FIMTLNTSL	CD8 池	小型 1
VDIDVPDVNF	CD8 池	小型 1
FEAPEGKL	CD8 池	小型 1
FEAPEGKLK	CD8 池	小型 1
LSISYLREY	CD8 池	小型 1
IEMVNRGQIQI	CD8 池	小型 1
IDVPDVNF	CD8 池	小型 1
IEMVNRGQI	CD8 池	小型 1
EHDEALKNTW	CD8 池	小型 1
EEHDEALKNTW	CD8 池	小型 1
NQIDTLQEV	CD8 池	小型 2
SLFHQSEVEL	CD8 池	小型 2
YIVGANTETYL	CD8 池	小型 2
IVGANTETY	CD8 池	小型 2
YIVGANTETY	CD8 池	小型 2
GEDCLTSQTRL	CD8 池	小型 2
NESPGNNVV	CD8 池	小型 2
TEILHSHVV	CD8 池	小型 2
AEVDVREREI	CD8 池	小型 2
SDELAQLY	CD8 池	小型 2
NESPGNNVVDI	CD8 池	小型 2

	DELVTFSDEL	CD8 池	小型 2
	DEFIQANA	CD8 池	小型 2
	DELAQLYHH	CD8 池	小型 2
	DEFIQANAT	CD8 池	小型 2
	KLPEADVAVL	CD8 池	小型 3
	PEADVAVL	CD8 池	小型 3
	ADVAVLRTF	CD8 池	小型 3
	PEADVAVLRTF	CD8 池	小型 3
[0538]	DVAVLRTF	CD8 池	小型 3
	EADVAVLRTF	CD8 池	小型 3
	DEYATLMM	CD8 池	小型 3
	GLTEVIQSL	CD8 池	小型 4
	QVLHLMVGV	CD8 池	小型 4
	VLHLMVGV	CD8 池	小型 4
	HEMLQVLHL	CD8 池	小型 4
	TEVIQSLKM	CD8 池	小型 4
	AEQKGLTEVI	CD8 池	小型 4

[0539] 表5—再刺激测定中的GRANITE患者G3肽

肽序列(氨基酸)	肽池	
LVEKSEAVRK	CD8 池	小型 1
ASMGTLAFHEY	CD8 池	小型 1
ASMGTLAFH	CD8 池	小型 1
MGTLAFHEY	CD8 池	小型 1
GTLAFHEY	CD8 池	小型 1
ATNQGWMLR	CD8 池	小型 1
GTMSLPASV	CD8 池	小型 1
RVPPKDLRVK	CD8 池	小型 1
DTLISQESFEK	CD8 池	小型 1
ISQESFEKK	CD8 池	小型 1
TLISQESFEK	CD8 池	小型 1
ISQESFEK	CD8 池	小型 1
FLKNENNEL	CD8 池	小型 1
SPVVETVPA	CD8 池	小型 1
LEPPSLPAQR	CD8 池	小型 1
MAMLIYTDF	CD8 池	小型 2

[0540]

[0541]

AMLIYTDFL	CD8 池	小型 2
SPSALIQEM	CD8 池	小型 2
SDWPPPNTR	CD8 池	小型 2
DAIDGRQAR	CD8 池	小型 2
MLAGIPVTI	CD8 池	小型 2
GMLAGIPVTI	CD8 池	小型 2
GMLAGIPVT	CD8 池	小型 2
TLYLLDIKV	CD8 池	小型 2
SFVAVTEEGTK	CD8 池	小型 2
FVAVTEEGTK	CD8 池	小型 2
YLSPPCIHC	CD8 池	小型 3
ARLGLTEVY	CD8 池	小型 3
FLEERLARL	CD8 池	小型 3
RLARLGLTEV	CD8 池	小型 3
LARLGLTEVY	CD8 池	小型 3
LARLGLTEV	CD8 池	小型 3
SLLRVLVTL	CD8 池	小型 3
VLVTLNDLEPK	CD8 池	小型 3
VLDHTLPTA	CD8 池	小型 4
TVLDHTLPTA	CD8 池	小型 4
LPTALTSPEL	CD8 池	小型 4
TVLDHTLPT	CD8 池	小型 4
TVLDHTLPTAL	CD8 池	小型 4
VLDHTLPTAL	CD8 池	小型 4

[0542] 表6-再刺激测定中的GRANITE患者G8肽

[0543]

肽序列(氨基酸)	肽池	
SSSSVLNNK	CD8 池	小型 1
SQSSSSVLNNK	CD8 池	小型 1
HVNGGPLYSK	CD8 池	小型 1
QSSSSVLNNK	CD8 池	小型 1
QTHVKGSLK	CD8 池	小型 1
ENAPPILPR	CD8 池	小型 1
VIFKGEIGK	CD8 池	小型 1
GSYSGTTILH	CD8 池	小型 1
VVVGADGVGK	CD8 池	小型 1

VNGGPLYSK	CD8 池	小型 1
SMRDMNITK	CD8 池	小型 2
TLLNLIQTK	CD8 池	小型 2
STLLNLIQTK	CD8 池	小型 2
YIDNALSRL	CD8 池	小型 2
TSRGQVSER	CD8 池	小型 2
SMRDMNITKL	CD8 池	小型 2
VGKSDSVDAEK	CD8 池	小型 2
TYIDNALSR	CD8 池	小型 2
KSDSVDAEK	CD8 池	小型 2
NLIQTKVNYV	CD8 池	小型 2
LLNLIQTK	CD8 池	小型 2
DSVDAEKSMLS	CD8 池	小型 2
DLRTSQSVR	CD8 池	小型 3
VLLTGTPLQK	CD8 池	小型 3
KVLLTGTPLQK	CD8 池	小型 3
LLTGTPLQK	CD8 池	小型 3
NFDLRTSQSVR	CD8 池	小型 3
QVYSKALNR	CD8 池	小型 4
LLYKAGGVK	CD8 池	小型 4
NFDVHHTAM	CD8 池	小型 4
DSPHKEPIRLR	CD8 池	小型 4
SLFLAVVAR	CD8 池	小型 4
AVSSCFRQL	CD8 池	小型 4
SSSSSLFL	CD8 池	小型 4
SSSSSSLFL	CD8 池	小型 4
DGAVSSCFR	CD8 池	小型 4
SSSSLFLAV	CD8 池	小型 4
LFLAVVAR	CD8 池	小型 4
QVYSKALNRL	CD8 池	小型 4
DVHHTAMLTR	CD8 池	小型 4

[0544]

[0545] 表7-再刺激测定中的GRANITE患者G11肽

肽序列(氨基酸)	肽池	
SGAGEGGALSV	CD8 池	小型 1
GAGEGGALSV	CD8 池	小型 1

[0546]

[0547]

KQTPQKKQL	CD8 池	小型 1
SGAGEGGAL	CD8 池	小型 1
GEGGALSV	CD8 池	小型 1
TASGAGEGGAL	CD8 池	小型 1
GAGEGGAL	CD8 池	小型 1
AGEGGALSV	CD8 池	小型 1
RQAQREAEM	CD8 池	小型 1
APEHTKRSV	CD8 池	小型 1
VLPDADTLLHL	CD8 池	小型 1
SIAPCVPPL	CD8 池	小型 1
SQIGNSQLL	CD8 池	小型 2
SQIGNSQL	CD8 池	小型 2
TPGHILAL	CD8 池	小型 3
IVVGTPGHIL	CD8 池	小型 3
GTPGHILAL	CD8 池	小型 3
VVGTPGHIL	CD8 池	小型 3
VGTPGHILAL	CD8 池	小型 3
VGTPGHIL	CD8 池	小型 3
IPGAAQAL	CD8 池	小型 3
VPCIPGAAQAL	CD8 池	小型 3
NSDQAVLEL	CD8 池	小型 3
NSDQAVLELDL	CD8 池	小型 3
LPRVLLQTP	CD8 池	小型 3
FLAPELPAF	CD8 池	小型 4
GSELSESAFSY	CD8 池	小型 4
FSVDSPSPGL	CD8 池	小型 4
LSESAFSY	CD8 池	小型 4
KAFAVSSHL	CD8 池	小型 4
KEQVNTEL	CD8 池	小型 4
ELSESAFSY	CD8 池	小型 4
SAFSYFITA	CD8 池	小型 4
VLGKFVPV	CD8 池	小型 4
SVDSPSPGL	CD8 池	小型 4
FLAPELPAFS	CD8 池	小型 4
LVLGKFVPV	CD8 池	小型 4
RLMSARACY	CD8 池	小型 4
SRPTAAIVL	CD8 池	小型 4

[0548]

SYMGGMNR	CD8 池	小型 4

[0549] 结果

[0550] 通过IFN-γELISpot评估GRANITE受试者G1和G3的CD8 T细胞免疫响应。如图3所示,受试者G1早在第68周使用最初用作初免疫苗[GRT-C901]的ChAdV载体施用加强剂量后2周就表现出增加的CD8 T细胞响应。图3A示出了受试者G1对整个CD8表位池的响应。图3B(完整的研究时间线)和图3C(仅第68周和第70周)示出了受试者G1对CD8表位小型池的响应。如图4所示,受试者G3也早在第36周使用最初用作初免疫苗[GRT-C901]的ChAdV载体施用加强剂量后1周(第37周)就表现出增加的CD8 T细胞响应。如图5所示,受试者G8也表现出在第36周使用最初用作初免疫苗[GRT-C901]的ChAdV载体施用加强剂量后至多29周检测到的增加的CD8 T细胞响应。如图6所示,受试者G11也早在第27周使用最初用作初免疫苗[GRT-C901]的ChAdV载体施用加强剂量后3周(第30周)就表现出增加的CD8 T细胞响应,并且增加的T细胞响应持续时间在使用ChAdV载体加强后持续至多24周。

[0551] 还确定了免疫细胞特征(例如,功能和免疫表型)以评估使用ChAdV载体的加强的功效。如图7所示,ChAdV加强增加了肿瘤患者的抗原特异性IFNγ和颗粒酶B CD8 T细胞响应。还观察到最小的IL-2产生,这与主要由CD8 T细胞驱动的响应一致。

[0552] 然后评估效应记忆T细胞群。效应记忆T细胞存在于外周循环和组织中,保留细胞毒性功能,并且具有高增殖能力以在再次暴露后对病原体/抗原呈递细胞提供立即响应。如图8A(加强前)和图8B(加强后)所示,ChAdV加强扩大了具有HLA-B*44:05受限细胞的效应记忆表型的抗原特异性T细胞的频率。如图8C所示,ChAdV加强(右图)扩大了具有HLA-A*02:01受限细胞的效应记忆表型的抗原特异性T细胞的频率。在所有情况下,加强还增加了总四聚体阳性HLA限制性抗原特异性T细胞的百分比。因此,数据证明,ChAdV加强刺激了总体抗原特异性T细胞的扩增,并且特别是具有效应记忆表型的T细胞的扩增。

[0553] 结果证明,当早在施用相同载体作为初免剂量后27周作为加强剂量施用时,GRANITE ChAdV载体GRT-C901刺激了稳健的CD8 T细胞免疫响应,从而增加总细胞群并且促进细胞功能和分化(例如,激活、细胞因子产生和免疫表型),并且相对于前ChAdV加强增加的T细胞响应可以持续至少29周。因此,数据表明ChAdV载体在同源初免/加强疫苗接种策略中是有效的。

[0554] XVIII.HIV治疗中的ChAdV加强方案

[0555] HIV受试者在初免/加强疫苗方案中进行治疗,所述方案包括用作初免疫苗接种的基于ChAdV的疫苗,以及配制在LNP中的基于SAM的疫苗和用于加强疫苗接种的基于ChAdV的疫苗的组合。受试者包括目前接受ART疗法的HIV+个体。ChAdV载体基于具有SEQ ID NO:1的序列的经修饰的ChAdV68序列,所述SEQ ID NO:1的序列具有E1(核苷酸577至3403)缺失和E3(核苷酸27,125-31,825)缺失,或基于具有SEQ ID NO:29369的序列的经修饰的ChAdV68序列,所述SEQ ID NO:29369的序列缺失E1(核苷酸577至3403)、E3(核苷酸27,125-31,825)和E4区域(核苷酸34,916至35,642)序列。SAM载体基于具有SEQ ID NO:6中列出的核酸序列的RNA甲病毒骨架。基于ChAdV的疫苗和基于SAM的疫苗均编码HIV抗原,诸如密码子优化的全长GAG,以及两种通用CD4 T细胞表位(PADRE和破伤风类毒素)。

[0556] 疫苗经由双侧肌内注射施用(例如,在每个三角肌中)。基于ChAdV的疫苗以$5x10^{11}$个病毒颗粒的总剂量施用,并根据需要调整为$1x10^{11}$或$1x10^{12}$个病毒颗粒。基于SAM的疫苗以300μg RNA的剂量施用,并且根据需要降级至100μg RNA。

[0557] 评估了两种疫苗方案。第一群组使用以下方案治疗:ChAdV-4w-SAM-12w-SAM-12w-ChAdV。第二群组使用以下方案治疗:ChAdV-4w-SAM-6w-SAM-6w-SAM-6w-SAM-6w-ChAdV。

[0558] 对受试者进行监测,并且结果证明初免/加强疫苗方案在治疗HIV中是有效的。

[0559] XIX.ChAdV加强治疗方案

[0560] 受试者在初免/加强疫苗方案中进行治疗,所述方案包括用作初免疫苗接种和加强疫苗接种的基于ChAdV的疫苗。方案还可以包括与基于SAM的疫苗的组合,所述基于SAM的疫苗配制在LNP中用于加强疫苗接种。ChAdV载体基于具有SEQ ID NO:1的序列的经修饰的ChAdV68序列,所述SEQ ID NO:1的序列具有E1(核苷酸577至3403)缺失和E3(核苷酸27,125-31,825)缺失,或基于具有SEQ ID NO:29369的序列的经修饰的ChAdV68序列,所述SEQ ID NO:29369的序列缺失E1(核苷酸577至3403)、E3(核苷酸27,125-31,825)和E4区域(核苷酸34,916至35,642)序列。SAM载体基于具有SEQ ID NO:6中列出的核酸序列的RNA甲病毒骨架。基于ChAdV的疫苗和基于SAM的疫苗(如果适用)编码抗原盒,所述抗原盒编码与受试者被诊断患有、预测患有或有患有疾病风险相关的串联T细胞表位和/或全长蛋白质,以及两种通用CD4 T细胞表位(PADRE和破伤风类毒素)。

[0561] 疫苗经由双侧肌内注射施用(例如,在每个三角肌中)。基于ChAdV的疫苗以$5x10^{11}$个病毒颗粒的总剂量施用,并根据需要调整为$1x10^{11}$或$1x10^{12}$个病毒颗粒。基于SAM的疫苗以300μg RNA的剂量施用,并且根据需要降级,例如根据需要降级至10μg和/或100μgRNA。

[0562] 评估的疫苗方案包括同源的基于ChAdV的初免/加强疫苗方案,其中在两个月或约2个月(例如,在8周或约8周)施用加强。

[0563] 对受试者进行监测,并且结果证明初免/加强疫苗方案在治疗受试者和/或使受试者免疫中是有效的。

[0564] 序列

[0565] 表A

[0566] 参考序列表,SEQ ID NO.10,755-21,015。为清楚起见,预测与给定HLA等位基因肽相关并且具有EDGE评分大于0.001的HLA等位基因的每种肽都分配有唯一的SEQ ID.NO。每个以上序列标识符都与以下各项相关:肽的氨基酸序列、HLA亚型、对应于肽的基因名称、与肽相关的突变,以及肽:HLA对的患病率是否大于0.1%(标注为“1”)或小于0.1%(标注为“0”)。

[0567] 表A在2018年5月23日提交的美国临时申请号62/675,559中完整公开,其特此通过引用整体并入。

[0568] AACR GENIE结果

[0569] 参考序列表,SEQ ID NO.21,016-29,357。为清楚起见,预测与给定HLA等位基因肽相关并且具有EDGE评分大于0.001并且患病率大于0.1%的HLA等位基因的每个肽都分配有唯一的SEQ ID.NO。每个以上序列标识符都与肽对应的基因名称和突变、HLA亚型和肽的氨基酸序列相关。

[0570] 额外的MS验证的新抗原(SEQ ID NO 29358-29364)

[0571]

SEQ ID NO:	基因	突变	HLA I 类亚型	限制性肽
29358	CTNNB1	S37Y 点突变	HLA-A*02:01	YLDSGIHYGA
29359	CHD4	CHD4_K73fs	HLA-B*08:01	TVRAATIL
29360	CTNNB1	CTNNB1_S45P	A*11:01	TTAPPLSGK
29361	CTNNB1	CTNNB1_T41A	A*11:01	ATAPSLSGK
29362	KRAS	KRAS_G12V	A*03:01	VVGAVGVGK
29363	KRAS	KRAS_Q61R	A*01:01	ILDTAGREEY
29364	TP53	TP53_R213L	A*02:01	YLDDRNTFL

[0572] 表1.2

[0573] 参考序列表,SEQ ID NO.57-10,754。在至少0.98%的癌症病例中,与基因相关的预测共享抗原以至少10TPM的水平表达。每个以上序列标识符都与肽的基因名称、氨基酸序列、Ensembl ID和一个或多个对应HLA等位基因相关。

[0574] 某些序列

[0575] 本文指代的载体、盒和抗体在下文描述并且由SEQ ID NO指代。

[0576]

曲美木单抗 VL (SEQ ID NO:16)
曲美木单抗 VH (SEQ ID NO:17)
曲美木单抗 VH CDR1 (SEQ ID NO:18)
曲美木单抗 VH CDR2 (SEQ ID NO:19)
曲美木单抗 VH CDR3 (SEQ ID NO:20)
曲美木单抗 VL CDR1 (SEQ ID NO:21)
曲美木单抗 VL CDR2 (SEQ ID NO:22)
曲美木单抗 VL CDR3 (SEQ ID NO:23)
德瓦鲁单抗(MEDI4736) VL (SEQ ID NO:24)
MEDI4736 VH (SEQ ID NO:25)
MEDI4736 VH CDR1 (SEQ ID NO:26)
MEDI4736 VH CDR2 (SEQ ID NO:27)
MEDI4736 VH CDR3 (SEQ ID NO:28)
MEDI4736 VL CDR1 (SEQ ID NO:29)
MEDI4736 VL CDR2 (SEQ ID NO:30)
MEDI4736 VL CDR3 (SEQ ID NO:31)
UbA76-25merPDTT 核苷酸(SEQ ID NO:32)
UbA76-25merPDTT 多肽(SEQ ID NO:33)
MAG-25merPDTT 核苷酸(SEQ ID NO:34)
MAG-25merPDTT 多肽(SEQ ID NO:35)
Ub7625merPDTT_NoSFL 核苷酸(SEQ ID NO:36)
Ub7625merPDTT_NoSFL 多肽(SEQ ID NO:37)
ChAdV68.5WTnt.MAG25mer (SEQ ID NO:2); 缺失 E1 (核苷酸 577 至 3403)和 E3 (核苷酸 27,125- 31,825)序列的 AC_000011.1; 对应的 ATCC VR-594 核苷酸在五个位置处被取代; 插入受 CMV 启动子/增强子控制的模型新抗原盒以替代缺失的 E1; 盒的 3'处 SV40

[0577]

聚 A
委内瑞拉马脑炎病毒[VEE] (SEQ ID NO:3) GenBank: L01442.2
VEE-MAG25mer (SEQ ID NO:4); 含有 MAG-25merPDTT 核苷酸(碱基 30-1755)
委内瑞拉马脑炎病毒株 TC-83 [TC-83](SEQ ID NO:5) GenBank: L01443.1
VEE 递送载体(SEQ ID NO:6); 缺失核苷酸 7544-11175 的 VEE 基因组[去除了甲病毒结构蛋白]
TC-83 递送载体(SEQ ID NO:7); 缺失核苷酸 7544-11175 的 TC-83 基因组[去除了甲病毒结构蛋白]
VEE 生产载体(SEQ ID NO:8); 缺失核苷酸 7544-11175 的 VEE 基因组，加上 5' T7 启动子，加上 3'限制性位点
TC-83 生产载体(SEQ ID NO:9); 缺失核苷酸 7544-11175 的 TC-83 基因组，加上 5' T7 启动子，加上 3'限制性位点
VEE-UbAAY (SEQ ID NO:14); 插入了 MHC I 类小鼠肿瘤表位 SIINFEKL 和 AH1-A5 的 VEE 递送载体
VEE-荧光素酶(SEQ ID NO:15); 在 7545 处插入荧光素酶基因的 VEE 递送载体
泛素(SEQ ID NO:38)>UbG76 0-228
泛素 A76 (SEQ ID NO:39)>UbA76 0-228
HLA-A2 (MHC I 类)信号肽(SEQ ID NO:40)>MHC 信号肽 0-78
HLA-A2 (MHC I 类)跨膜结构域(SEQ ID NO:41)>HLA A2 跨膜结构域 0-201
IgK 前导序列(SEQ ID NO:42)>IgK 前导序列 0-60
人 DC-Lamp (SEQ ID NO:43)>人 DCLAMP 0-3178
小鼠 LAMP1 (SEQ ID NO:44)>小鼠 Lamp1 0-1858
人 Lamp1 cDNA (SEQ ID NO:45)>人 Lamp1 0-2339
破伤风类毒素核酸序列(SEQ ID NO:46)
破伤风类毒素氨基酸序列(SEQ ID NO:47)
PADRE 核苷酸序列(SEQ ID NO:48)
PADRE 氨基酸序列(SEQ ID NO:49)
WPRE (SEQ ID NO:50)>WPRE 0-593
IRES (SEQ ID NO:51)>eGFP_IRES_SEAP_插入 1746-2335
GFP (SEQ ID NO:52)
SEAP (SEQ ID NO:53)
萤火虫荧光素酶(SEQ ID NO:54)
FMDV 2A (SEQ ID NO:55)
伊匹单抗重链(SEQ ID NO: 29365)

[0578]

QVQLVESGGGVVQPGRSLRLSCAASGFTFSSYTMHWVRQAPG KGLEWVTFISYDGNNKYYADSVKGRFTISRDNSKNTLYLQMN SLRAEDTAIYYCARTGWLGPFDYWGQGTLVTVSSASTKGPSVF PLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVHT FPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDK RVEPKSCDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPE VTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNST YRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKG QPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNG QPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVM HEALHNHYTQKSLSLSPGK
伊匹单抗轻链(SEQ ID NO: 29366)
EIVLTQSPGTLSLSPGERATLSCRASQSVGSSYLAWYQQKPGQA PRLLIYGAFSRATGIPDRFSGSGSGTDFTLTISRLEPEDFAVYYCQ QYGSSPWTFGQGTKVEIKRTVAAPSVFIFPPSDEQLKSGTASVV CLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSL SSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC
纳武单抗重链(SEQ ID NO: 29367)
QVQLVESGGGVVQPGRSLRLDCKASGITFSNSGMHWVRQAPG KGLEWVAVIWYDGSKRYYADSVKGRFTISRDNSKNTLFLQMN SLRAEDTAVYYCATNDDYWGQGTLVTVSSASTKGPSVFPLAPC SRSTSESTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQ SSGLYSLSSVVTVPSSSLGTKTYTCNVDHKPSNTKVDKRVESK YGPPCPPCPAPEFLGGPSVFLFPPKPKDTLMISRTPEVTCVVVD VSQEDPEVQFNWYVDGVEVHNAKTKPREEQFNSTYRVVSVLT VLHQDWLNGKEYKCKVSNKGLPSSIEKTISKAKGQPREPQVYT LPPSQEEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKT TPPVLDSDGSFFLYSRLTVDKSRWQEGNVFSCSVMHEALHNHY TQKSLSLSLGK
纳武单抗轻链(SEQ ID NO:29368)
EIVLTQSPATLSLSPGERATLSCRASQSVSSYLAWYQQKPGQAP RLLIYDASNRATGIPARFSGSGSGTDFTLTISSLEPEDFAVYYCQ QSSNWPRTFGQGTKVEIKRTVAAPSVFIFPPSDEQLKSGTASVV CLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSL SSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC
chAd68-空-E4 缺失(SEQ ID NO: 29369)
CATCTTCAATAATATACCTCAAACTTTTTGTGCGCGTTAATATG CAAATGAGGCGTTTGAATTTGGGGAGGAAGGGCGGTGATTG GTCGAGGGATGAGCGACCGTTAGGGGCGGGGCGAGTGACG TTTTGATGACGTGGTTGCGAGGAGGAGCCAGTTTGCAAGTT CTCGTGGGAAAAGTGACGTCAAACGAGGTGTGGTTTGAACA CGGAAATACTCAATTTTCCCGCGCTCTCTGACAGGAAATGAG

[0579]

GTGTTTCTGGGCGGATGCAAGTGAAAACGGGCCATTTTCGC
GCGAAAACTGAATGAGGAAGTGAAAATCTGAGTAATTTCGC
GTTTATGGCAGGGAGGAGTATTTGCCGAGGGCCGAGTAGAC
TTTGACCGATTACGTGGGGGTTTCGATTACCGTGTTTTTCACC
TAAATTTCCGCGTACGGTGTCAAAGTCCGGTGTTTTTACGTA
GGTGTCAGCTGATCGCCAGGGTATTTAAACCTGCGCTCTCCA
GTCAAGAGGCCACTCTTGAGTGCCAGCGAGAAGAGTTTTCT
CCTCCGCGCCGCGAGTCAGATCTACACTTTGAAAGTAGGGata
aGGTAGCGAGTGAGTAGTGTTCTGGGGCGGGGGAGGACCTG
CATGAGGGCCAGAATAACTGAAATCTGTGCTTTTCTGTGTGT
TGCAGCAGCATGAGCGGAAGCGGCTCCTTTGAGGGAGGGGT
ATTCAGCCCTTATCTGACGGGGCGTCTCCCCTCCTGGGCGGG
AGTGCGTCAGAATGTGATGGGATCCACGGTGGACGGCCGGC
CCGTGCAGCCCGCGAACTCTTCAACCCTGACCTATGCAACCC
TGAGCTCTTCGTCGTTGGACGCAGCTGCCGCCGCAGCTGCT
GCATCTGCCGCCAGCGCCGTGCGCGGAATGGCCATGGGCGC
CGGCTACTACGGCACTCTGGTGGCCAACTCGAGTTCCACCAA
TAATCCCGCCAGCCTGAACGAGGAGAAGCTGTTGCTGCTGA
TGGCCCAGCTCGAGGCCTTGACCCAGCGCCTGGGCGAGCTG
ACCCAGCAGGTGGCTCAGCTGCAGGAGCAGACGCGGGCCG
CGGTTGCCACGGTGAAATCCAAATAAAAAATGAATCAATAAA
TAAACGGAGACGGTTGTTGATTTTAACACAGAGTCTGAATCT
TTATTTGATTTTTCGCGCGCGGTAGGCCCTGGACCACCGGTC
TCGATCATTGAGCACCCGGTGGATCTTTTCCAGGACCCGGTA
GAGGTGGGCTTGGATGTTGAGGTACATGGGCATGAGCCCGT
CCCGGGGGGTGGAGGTAGCTCCATTGCAGGGCCTCGTGCTCG
GGGGTGGTGTTGTAAATCACCCAGTCATAGCAGGGGCGCAG
GGCATGGTGTTGCACAATATCTTTGAGGAGGAGACTGATGGC
CACGGGCAGCCCTTTGGTGTAGGTGTTTACAAATCTGTTGAG
CTGGGAGGGATGCATGCGGGGGGAGATGAGGTGCATCTTGG
CCTGGATCTTGAGATTGGCGATGTTACCGCCCAGATCCCGCC
TGGGGTTCATGTTGTGCAGGACCACCAGCACGGTGTATCCG
GTGCACTTGGGGAATTTATCATGCAACTTGGAAGGGAAGGC
GTGAAAGAATTTGGCGACGCCTTTGTGCCCGCCCAGGTTTTC
CATGCACTCATCCATGATGATGGCGATGGGCCCGTGGGCGGC
GGCCTGGGCAAAGACGTTTCGGGGGTCGGACACATCATAGT
TGTGGTCCTGGGTGAGGTCATCATAGGCCATTTTAATGAATTT
GGGGCGGAGGGTGCCGGACTGGGGGACAAAGGTACCCTCG
ATCCCGGGGGCGTAGTTCCCCTCACAGATCTGCATCTCCCAG
GCTTTGAGCTCGGAGGGGGGGATCATGTCCACCTGCGGGGC
GATAAAGAACACGGTTTCCGGGGCGGGGGAGATGAGCTGGG
CCGAAAGCAAGTTCCGGAGCAGCTGGGACTTGCCGCAGCCG
GTGGGGCCGTAGATGACCCCGATGACCGGCTGCAGGTGGTA

[0580]

GTTGAGGGAGAGACAGCTGCCGTCCTCCCGGAGGAGGGGG
GCCACCTCGTTCATCATCTCGCGCACGTGCATGTTCTCGCGC
ACCAGTTCCGCCAGGAGGCGCTCTCCCCCCAGGGATAGGAG
CTCCTGGAGCGAGGCGAAGTTTTTCAGCGGCTTGAGTCCGT
CGGCCATGGGCATTTTGGAGAGGGTTTGTTGCAAGAGTTCC
AGGCGGTCCCAGAGCTCGGTGATGTGCTCTACGGCATCTCGA
TCCAGCAGACCTCCTCGTTTCGCGGGTTGGGACGGCTGCGG
GAGTAGGGCACCAGACGATGGGCGTCCAGCGCAGCCAGGG
TCCGGTCCTTCCAGGGTCGCAGCGTCCGCGTCAGGGTGGTC
TCCGTCACGGTGAAGGGGTGCGCGCCGGGCTGGGCGCTTGC
GAGGGTGCGCTTCAGGCTCATCCGGCTGGTCGAAAACCGCT
CCCGATCGGCGCCCTGCGCGTCGGCCAGGTAGCAATTGACC
ATGAGTTCGTAGTTGAGCGCCTCGGCCGCGTGGCCTTTGGCG
CGGAGCTTACCTTTGGAAGTCTGCCCGCAGGCGGGACAGAG
GAGGGACTTGAGGGCGTAGAGCTTGGGGGCGAGGAAGACG
GACTCGGGGGCGTAGGCGTCCGCGCCGCAGTGGGCGCAGA
CGGTCTCGCACTCCACGAGCCAGGTGAGGTCGGGCTGGTCG
GGGTCAAAAACCAGTTTCCCGCCGTTCTTTTTGATGCGTTTC
TTACCTTTGGTCTCCATGAGCTCGTGTCCCCGCTGGGTGACA
AAGAGGCTGTCCGTGTCCCCGTAGACCGACTTTATGGGCCG
GTCCTCGAGCGGTGTGCCGCGGTCCTCCTCGTAGAGGAACC
CCGCCCACTCCGAGACGAAAGCCCGGGTCCAGGCCAGCAC
GAAGGAGGCCACGTGGGACGGGTAGCGGTCGTTGTCCACCA
GCGGGTCCACCTTTTCCAGGGTATGCAAACACATGTCCCCCT
CGTCCACATCCAGGAAGGTGATTGGCTTGTAAGTGTAGGCCA
CGTGACCGGGGGGTCCCGGCCGGGGGGGGTATAAAAGGGTGCG
GGTCCCTGCTCGTCCTCACTGTCTTCCGGATCGCTGTCCAGG
AGCGCCAGCTGTTGGGGTAGGTATTCCCTCTCGAAGGCGGG
CATGACCTCGGCACTCAGGTTGTCAGTTTCTAGAAACGAGG
AGGATTTGATATTGACGGTGCCGGCGGAGATGCCTTTCAAGA
GCCCCTCGTCCATCTGGTCAGAAAAGACGATCTTTTTGTTGT
CGAGCTTGGTGGCGAAGGAGCCGTAGAGGGCGTTGGAGAG
GAGCTTGGCGATGGAGCGCATGGTCTGGTTTTTTTCCTTGTC
GGCGCGCTCCTTGGCGGCGATGTTGAGCTGCACGTACTCGC
GCGCCACGCACTTCCATTCGGGGAAGACGGTGGTCAGCTCG
TCGGGCACGATTCTGACCTGCCAGCCCCGATTATGCAGGGTG
ATGAGGTCCACACTGGTGGCCACCTCGCCGCGCAGGGGCTC
ATTAGTCCAGCAGAGGCGTCCGCCCTTGCGCGAGCAGAAGG
GGGGCAGGGGGTCCAGCATGACCTCGTCGGGGGGGTCGGCA
TCGATGGTGAAGATGCCGGGCAGGAGGTCGGGGTCAAAGTA
GCTGATGGAAGTGGCCAGATCGTCCAGGGCAGCTTGCCATT
CGCGCACGGCCAGCGCGCGCTCGTAGGGACTGAGGGGCGT
GCCCCAGGGCATGGGATGGGTAAGCGCGGAGGCGTACATGC

[0581]

CGCAGATGTCGTAGACGTAGAGGGGCTCCTCGAGGATGCCG
ATGTAGGTGGGGTAGCAGCGCCCCCCGCGGATGCTGGCGCG
CACGTAGTCATACAGCTCGTGCGAGGGGGCGAGGAGCCCCG
GGCCCAGGTTGGTGCGACTGGGCTTTTCGGCGCGGTAGACG
ATCTGGCGGAAAATGGCATGCGAGTTGGAGGAGATGGTGGG
CCTTTGGAAGATGTTGAAGTGGGCGTGGGGCAGTCCGACCG
AGTCGCGGATGAAGTGGGCGTAGGAGTCTTGCAGCTTGGCG
ACGAGCTCGGCGGTGACTAGGACGTCCAGAGCGCAGTAGTC
GAGGGTCTCCTGGATGATGTCATACTTGAGCTGTCCCTTTTG
TTTCCACAGCTCGCGGTTGAGAAGGAACTCTTCGCGGTCCTT
CCAGTACTCTTCGAGGGGGAACCCGTCCTGATCTGCACGGTA
AGAGCCTAGCATGTAGAACTGGTTGACGGCCTTGTAGGCGC
AGCAGCCCTTCTCCACGGGGAGGGCGTAGGCCTGGGCGGCC
TTGCGCAGGGAGGTGTGCGTGAGGGCGAAAGTGTCCCTGAC
CATGACCTTGAGGAACTGGTGCTTGAAGTCGATATCGTCGCA
GCCCCCCTGCTCCCAGAGCTGGAAGTCCGTGCGCTTCTTGTA
GGCGGGGTTGGGCAAAGCGAAAGTAACATCGTTGAAGAGG
ATCTTGCCCGCGCGGGGCATAAAGTTGCGAGTGATGCGGAA
AGGTTGGGGCACCTCGGCCCGGTTGTTGATGACCTGGGCGG
CGAGCACGATCTCGTCGAAGCCGTTGATGTTGTGGCCCACGA
TGTAGAGTTCCACGAATCGCGGACGGCCCTTGACGTGGGGC
AGTTTCTTGAGCTCCTCGTAGGTGAGCTCGTCGGGGTCGCTG
AGCCCGTGCTGCTCGAGCGCCCAGTCGGCGAGATGGGGGTT
GGCGCGGAGGAAGGAAGTCCAGAGATCCACGGCCAGGGCG
GTTTGCAGACGGTCCCGGTACTGACGGAACTGCTGCCCGAC
GGCCATTTTTTCGGGGGTGACGCAGTAGAAGGTGCGGGGGT
CCCCGTGCCAGCGATCCCATTTGAGCTGGAGGGCGAGATCG
AGGGCGAGCTCGACGAGCCGGTCGTCCCCGGAGAGTTTCAT
GACCAGCATGAAGGGGACGAGCTGCTTGCCGAAGGACCCCA
TCCAGGTGTAGGTTTCCACATCGTAGGTGAGGAAGAGCCTTT
CGGTGCGAGGATGCGAGCCGATGGGGAAGAACTGGATCTCC
TGCCACCAATTGGAGGAATGGCTGTTGATGTGATGGAAGTAG
AAATGCCGACGGCGCGCCGAACACTCGTGCTTGTGTTTATAC
AAGCGGCCACAGTGCTCGCAACGCTGCACGGGATGCACGTG
CTGCACGAGCTGTACCTGAGTTCCTTTGACGAGGAATTTCAG
TGGGAAGTGGAGTCGTGGCGCCTGCATCTCGTGCTGTACTAC
GTCGTGGTGGTCGGCCTGGCCCTCTTCTGCCTCGATGGTGGT
CATGCTGACGAGCCCGCGCGGGAGGCAGGTCCAGACCTCGG
CGCGAGCGGGTCGGAGAGCGAGGACGAGGGCGCGCAGGCC
GGAGCTGTCCAGGGTCCTGAGACGCTGCGGAGTCAGGTCAG
TGGGCAGCGGCGGCGCGCGGTTGACTTGCAGGAGTTTTTCC
AGGGCGCGCGGGAGGTCCAGATGGTACTTGATCTCCACCGC
GCCATTGGTGGCGACGTCGATGGCTTGCAGGGTCCCGTGCC

[0582]

CCTGGGGTGTGACCACCGTCCCCCGTTTCTTCTTGGGCGGCT
GGGGCGACGGGGGCGGTGCCTCTTCCATGGTTAGAAGCGGC
GGCGAGGACGCGCGCCGGGCGGCAGGGGCGGCTCGGGGCC
CGGAGGCAGGGGCGGCAGGGGCACGTCGGCGCCGCGCGCG
GGTAGGTTCTGGTACTGCGCCCGGAGAAGACTGGCGTGAGC
GACGACGCGACGGTTGACGTCCTGGATCTGACGCCTCTGGG
TGAAGGCCACGGGACCCGTGAGTTTGAACCTGAAAGAGAG
TTCGACAGAATCAATCTCGGTATCGTTGACGGCGGCCTGCCG
CAGGATCTCTTGCACGTCGCCCGAGTTGTCCTGGTAGGCGAT
CTCGGTCATGAACTGCTCGATCTCCTCCTCTTGAAGGTCTCC
GCGGCCGGCGCGCTCCACGGTGGCCGCGAGGTCGTTGGAGA
TGCGGCCCATGAGCTGCGAGAAGGCGTTCATGCCCGCCTCG
TTCCAGACGCGGCTGTAGACCACGACGCCCTCGGGATCGCg
GGCGCGCATGACCACCTGGGCGAGGTTGAGCTCCACGTGGC
GCGTGAAGACCGCGTAGTTGCAGAGGCGCTGGTAGAGGTAG
TTGAGCGTGGTGGCGATGTGCTCGGTGACGAAGAAATACAT
GATCCAGCGGCGGAGCGGCATCTCGCTGACGTCGCCCAGCG
CCTCCAAACGTTCCATGGCCTCGTAAAAGTCCACGGCGAAG
TTGAAAAACTGGGAGTTGCGCGCCGAGACGGTCAACTCCTC
CTCCAGAAGACGGATGAGCTCGGCGATGGTGGCGCGCACCT
CGCGCTCGAAGGCCCCCGGGAGTTCCTCCACTTCCTCTTCTT
CCTCCTCCACTAACATCTCTTCTACTTCCTCCTCAGGCGGCA
GTGGTGGCGGGGGAGGGGGCCTGCGTCGCCGGCGGCGCAC
GGGCAGACGGTCGATGAAGCGCTCGATGGTCTCGCCGCGCC
GGCGTCGCATGGTCTCGGTGACGGCGCGCCCGTCCTCGCGG
GGCCGCAGCGTGAAGACGCCGCCGCGCATCTCCAGGTGGCC
GGGGGGGTCCCCGTTGGGCAGGGAGAGGGCGCTGACGATG
CATCTTATCAATTGCCCCGTAGGGACTCCGCGCAAGGACCTG
AGCGTCTCGAGATCCACGGGATCTGAAAACCGCTGAACGAA
GGCTTCGAGCCAGTCGCAGTCGCAAGGTAGGCTGAGCACGG
TTTCTTCTGGCGGGTCATGTTGGTTGGGAGCGGGGCGGGCG
ATGCTGCTGGTGATGAAGTTGAAATAGGCGGTTCTGAGACG
GCGGATGGTGGCGAGGAGCACCAGGTCTTTGGGCCCGGCTT
GCTGGATGCGCAGACGGTCGGCCATGCCCCAGGCGTGGTCC
TGACACCTGGCCAGGTCCTTGTAGTAGTCCTGCATGAGCCGC
TCCACGGGCACCTCCTCCTCGCCCGCGCGGCCGTGCATGCG
CGTGAGCCCGAAGCCGCGCTGGGGCTGGACGAGCGCCAGG
TCGGCGACGACGCGCTCGGCGAGGATGGCTTGCTGGATCTG
GGTGAGGGTGGTCTGGAAGTCATCAAAGTCGACGAAGCGGT
GGTAGGCTCCGGTGTTGATGGTGTAGGAGCAGTTGGCCATG
ACGGACCAGTTGACGGTCTGGTGGCCCGGACGCACGAGCTC
GTGGTACTTGAGGCGCGAGTAGGCGCGCGTGTCGAAGATGT
AGTCGTTGCAGGTGCGCACCAGGTACTGGTAGCCGATGAGG

[0583]

AAGTGCGGCGGCGGCTGGCGGTAGAGCGGCCATCGCTCGGT
GGCGGGGGCGCCGGGCGCGAGGTCCTCGAGCATGGTGCGGT
GGTAGCCGTAGATGTACCTGGACATCCAGGTGATGCCGGCGG
CGGTGGTGGAGGCGCGCGGGAACTCGCGGACGCGGTTCCA
GATGTTGCGCAGCGGCAGGAAGTAGTTCATGGTGGGCACGG
TCTGGCCCGTGAGGCGCGCGCAGTCGTGGATGCTCTATACGG
GCAAAAACGAAAGCGGTCAGCGGCTCGACTCCGTGGCCTG
GAGGCTAAGCGAACGGGTTGGGCTGCGCGTGTACCCCGGTT
CGAATCTCGAATCAGGCTGGAGCCGCAGCTAACGTGGTATTG
GCACTCCCGTCTCGACCCAAGCCTGCACCAACCCTCCAGGA
TACGGAGGCGGGTCGTTTTGCAACTTTTTTTTGGAGGCCGGA
TGAGACTAGTAAGCGCGGAAAGCGGCCGACCGCGATGGCTC
GCTGCCGTAGTCTGGAGAAGAATCGCCAGGGTTGCGTTGCG
GTGTGCCCCGGTTCGAGGCCGGCCGGATTCCGCGGCTAACG
AGGGCGTGGCTGCCCCGTCGTTTCCAAGACCCCATAGCCAG
CCGACTTCTCCAGTTACGGAGCGAGCCCCTCTTTTGTTTTGT
TTGTTTTTGCCAGATGCATCCCGTACTGCGGCAGATGCGCCC
CCACCACCCTCCACCGCAACAACAGCCCCCTCCACAGCCGG
CGCTTCTGCCCCCGCCCCAGCAGCAACTTCCAGCCACGACC
GCCGCGGCCGCCGTGAGCGGGGCTGGACAGAGTTATGATCA
CCAGCTGGCCTTGGAAGAGGGCGAGGGGCTGGCGCGCCTG
GGGGCGTCGTCGCCGGAGCGGCACCCGCGCGTGCAGATGAA
AAGGGACGCTCGCGAGGCCTACGTGCCCAAGCAGAACCTGT
TCAGAGACAGGAGCGGCGAGGAGCCCGAGGAGATGCGCGC
GGCCCGGTTCCACGCGGGGCGGGAGCTGCGGCGCGGCCTG
GACCGAAAGAGGGTGCTGAGGGACGAGGATTTCGAGGCGG
ACGAGCTGACGGGGATCAGCCCCGCGCGCGCGCACGTGGCC
GCGGCCAACCTGGTCACGGCGTACGAGCAGACCGTGAAGG
AGGAGAGCAACTTCCAAAAATCCTTCAACAACCACGTGCGC
ACCCTGATCGCGCGCGAGGAGGTGACCCTGGGCCTGATGCA
CCTGTGGGACCTGCTGGAGGCCATCGTGCAGAACCCCACCA
GCAAGCCGCTGACGGCGCAGCTGTTCCTGGTGGTGCAGCAT
AGTCGGGACAACGAAGCGTTCAGGGAGGCGCTGCTGAATAT
CACCGAGCCCGAGGGCCGCTGGCTCCTGGACCTGGTGAACA
TTCTGCAGAGCATCGTGGTGCAGGAGCGCGGGCTGCCGCTG
TCCGAGAAGCTGGCGGCCATCAACTTCTCGGTGCTGAGTTT
GGGCAAGTACTACGCTAGGAAGATCTACAAGACCCCGTACG
TGCCCATAGACAAGGAGGTGAAGATCGACGGGTTTTACATG
CGCATGACCCTGAAAGTGCTGACCCTGAGCGACGATCTGGG
GGTGTACCGCAACGACAGGATGCACCGTGCGGTGAGCGCCA
GCAGGCGGCGCGAGCTGAGCGACCAGGAGCTGATGCATAGT
CTGCAGCGGGCCCTGACCGGGGCCGGGACCGAGGGGGAGA
GCTACTTTGACATGGGCGCGGACCTGCACTGGCAGCCCAGC

[0584]

CGCCGGGCCTTGGAGGCGGCGGCAGGACCCTACGTAGAAGA
GGTGGACGATGAGGTGGACGAGGAGGGCGAGTACCTGGAA
GACTGATGGCGCGACCGTATTTTTGCTAGATGCAACAACAAC
AGCCACCTCCTGATCCCGCGATGCGGGCGGCGCTGCAGAGC
CAGCCGTCCGGCATTAACTCCTCGGACGATTGGACCCAGGC
CATGCAACGCATCATGGCGCTGACGACCCGCAACCCCGAAG
CCTTTAGACAGCAGCCCCAGGCCAACCGGCTCTCGGCCATC
CTGGAGGCCGTGGTGCCCTCGCGCTCCAACCCCACGCACGA
GAAGGTCCTGGCCATCGTGAACGCGCTGGTGGAGAACAAGG
CCATCCGCGGCGACGAGGCCGGCCTGGTGTACAACGCGCTG
CTGGAGCGCGTGGCCCGCTACAACAGCACCAACGTGCAGAC
CAACCTGGACCGCATGGTGACCGACGTGCGCGAGGCCGTGG
CCCAGCGCGAGCGGTTCCACCGCGAGTCCAACCTGGGATCC
ATGGTGGCGCTGAACGCCTTCCTCAGCACCCAGCCCGCCAA
CGTGCCCCGGGGCCAGGAGGACTACACCAACTTCATCAGCG
CCCTGCGCCTGATGGTGACCGAGGTGCCCCAGAGCGAGGTG
TACCAGTCCGGGCCGGACTACTTCTTCCAGACCAGTCGCCA
GGGCTTGCAGACCGTGAACCTGAGCCAGGCTTTCAAGAACT
TGCAGGGCCTGTGGGGCGTGCAGGCCCCGGTCGGGGACCGC
GCGACGGTGTCGAGCCTGCTGACGCCGAACTCGCGCCTGCT
GCTGCTGCTGGTGGCCCCCTTCACGGACAGCGGCAGCATCA
ACCGCAACTCGTACCTGGGCTACCTGATTAACCTGTACCGCG
AGGCCATCGGCCAGGCGCACGTGGACGAGCAGACCTACCAG
GAGATCACCCACGTGAGCCGCGCCCTGGGCCAGGACGACCC
GGGCAACCTGGAAGCCACCCTGAACTTTTTGCTGACCAACC
GGTCGCAGAAGATCCCGCCCCAGTACGCGCTCAGCACCGAG
GAGGAGCGCATCCTGCGTTACGTGCAGCAGAGCGTGGGCCT
GTTCCTGATGCAGGAGGGGGCCACCCCCAGCGCCGCGCTCG
ACATGACCGCGCGCAACATGGAGCCCAGCATGTACGCCAGC
AACCGCCCGTTCATCAATAAACTGATGGACTACTTGCATCGG
GCGGCCGCCATGAACTCTGACTATTTCACCAACGCCATCCTG
AATCCCCACTGGCTCCCGCCGCCGGGGTTCTACACGGGCGA
GTACGACATGCCCGACCCCAATGACGGGTTCCTGTGGGACG
ATGTGGACAGCAGCGTGTTCTCCCCCCGACCGGGTGCTAAC
GAGCGCCCCTTGTGGAAGAAGGAAGGCAGCGACCGACGCC
CGTCCTCGGCGCTGTCCGGCCGCGAGGGTGCTGCCGCGGCG
GTGCCCGAGGCCGCCAGTCCTTTCCCGAGCTTGCCCTTCTCG
CTGAACAGTATCCGCAGCAGCGAGCTGGGCAGGATCACGCG
CCCGCGCTTGCTGGGCGAAGAGGAGTACTTGAATGACTCGC
TGTTGAGACCCGAGCGGGAGAAGAACTTCCCCAATAACGGG
ATAGAAAGCCTGGTGGACAAGATGAGCCGCTGGAAGACGTA
TGCGCAGGAGCACAGGGACGATCCCCGGGCGTCGCAGGGG
GCCACGAGCCGGGGCAGCGCCGCCCGTAAACGCCGGTGGC

[0585]

ACGACAGGCAGCGGGGACAGATGTGGGACGATGAGGACTC
CGCCGACGACAGCAGCGTGTTGGACTTGGGTGGGAGTGGTA
ACCCGTTCGCTCACCTGCGCCCCCGTATCGGGCGCATGATGT
AAGAGAAACCGAAAATAAATGATACTCACCAAGGCCATGGC
GACCAGCGTGCGTTCGTTTCTTCTCTGTTGTTGTTGTATCTAG
TATGATGAGGCGTGCGTACCCGGAGGGTCCTCCTCCCTCGTA
CGAGAGCGTGATGCAGCAGGCGATGGCGGCGGCGGCGATGC
AGCCCCCGCTGGAGGCTCCTTACGTGCCCCCGCGGTACCTG
GCGCCTACGGAGGGGCGGAACAGCATTCGTTACTCGGAGCT
GGCACCCTTGTACGATACCACCCGGTTGTACCTGGTGGACAA
CAAGTCGGCGGACATCGCCTCGCTGAACTACCAGAACGACC
ACAGCAACTTCCTGACCACCGTGGTGCAGAACAATGACTTC
ACCCCCACGGAGGCCAGCACCCAGACCATCAACTTTGACGA
GCGCTCGCGGTGGGGCGGCCAGCTGAAAACCATCATGCACA
CCAACATGCCCAACGTGAACGAGTTCATGTACAGCAACAAG
TTCAAGGCGCGGGTGATGGTCTCCCGCAAGACCCCCAATGG
GGTGACAGTGACAGAGGATTATGATGGTAGTCAGGATGAGC
TGAAGTATGAATGGGTGGAATTTGAGCTGCCCGAAGGCAAC
TTCTCGGTGACCATGACCATCGACCTGATGAACAACGCCATC
ATCGACAATTACTTGGCGGTGGGGCGGCAGAACGGGGTGCT
GGAGAGCGACATCGGCGTGAAGTTCGACACTAGGAACTTCA
GGCTGGGCTGGGACCCCGTGACCGAGCTGGTCATGCCCGGG
GTGTACACCAACGAGGCTTTCCATCCCGATATTGTCTTGCTG
CCCGGCTGCGGGGTGGACTTCACCGAGAGCCGCCTCAGCAA
CCTGCTGGGCATTCGCAAGAGGCAGCCCTTCCAGGAAGGCT
TCCAGATCATGTACGAGGATCTGGAGGGGGGCAACATCCCC
GCGCTCCTGGATGTCGACGCCTATGAGAAAAGCAAGGAGGA
TGCAGCAGCTGAAGCAACTGCAGCCGTAGCTACCGCCTCTA
CCGAGGTCAGGGGCGATAATTTTGCAAGCGCCGCAGCAGTG
GCAGCGGCCGAGGCGGCTGAAACCGAAAGTAAGATAGTCAT
TCAGCCGGTGGAGAAGGATAGCAAGAACAGGAGCTACAAC
GTACTACCGGACAAGATAAACACCGCCTACCGCAGCTGGTA
CCTAGCCTACAACTATGGCGACCCCGAGAAGGGCGTGCGCT
CCTGGACGCTGCTCACCACCTCGGACGTCACCTGCGGCGTG
GAGCAAGTCTACTGGTCGCTGCCCGACATGATGCAAGACCC
GGTCACCTTCCGCTCCACGCGTCAAGTTAGCAACTACCCGGT
GGTGGGCGCCGAGCTCCTGCCCGTCTACTCCAAGAGCTTCTT
CAACGAGCAGGCCGTCTACTCGCAGCAGCTGCGCGCCTTCA
CCTCGCTTACGCACGTCTTCAACCGCTTCCCCGAGAACCAGA
TCCTCGTCCGCCCGCCCGCGCCCACCATTACCACCGTCAGTG
AAAACGTTCCTGCTCTCACAGATCACGGGACCCTGCCGCTG
CGCAGCAGTATCCGGGGAGTCCAGCGCGTGACCGTTACTGA
CGCCAGACGCCGCACCTGCCCCTACGTCTACAAGGCCCTGG

[0586]

GCATAGTCGCGCCGCGCGTCCTCTCGAGCCGCACCTTCTAAA
TGTCCATTCTCATCTCGCCCAGTAATAACACCGGTTGGGGCC
TGCGCGCGCCCAGCAAGATGTACGGAGGCGCTCGCCAACGC
TCCACGCAACACCCCGTGCGCGTGCGCGGGCACTTCCGCGC
TCCCTGGGGCGCCCTCAAGGGCCGCGTGCGGTCGCGCACCA
CCGTCGACGACGTGATCGACCAGGTGGTGGCCGACGCGCGC
AACTACACCCCCGCCGCCGCGCCCGTCTCCACCGTGGACGC
CGTCATCGACAGCGTGGTGGCcGACGCGCGCCGGTACGCCC
GCGCCAAGAGCCGGCGGCGGCGCATCGCCCGGCGGCACCG
GAGCACCCCCGCCATGCGCGCGGCGCGAGCCTTGCTGCGCA
GGGCCAGGCGCACGGGACGCAGGGCCATGCTCAGGGCGGC
CAGACGCGCGGCTTCAGGCGCCAGCGCCGGCAGGACCCGG
AGACGCGCGGCCACGGCGGCGGCAGCGGCCATCGCCAGCAT
GTCCCGCCCGCGGCGAGGGAACGTGTACTGGGTGCGCGACG
CCGCCACCGGTGTGCGCGTGCCCGTGCGCACCCGCCCCCCT
CGCACTTGAAGATGTTCACTTCGCGATGTTGATGTGTCCCAG
CGGCGAGGAGGATGTCCAAGCGCAAATTCAAGGAAGAGATG
CTCCAGGTCATCGCGCCTGAGATCTACGGCCCTGCGGTGGTG
AAGGAGGAAAGAAAGCCCCGCAAAATCAAGCGGGTCAAAA
AGGACAAAAAGGAAGAAGAAAGTGATGTGGACGGATTGGT
GGAGTTTGTGCGCGAGTTCGCCCCCCGGCGGCGCGTGCAGT
GGCGCGGGCGGAAGGTGCAACCGGTGCTGAGACCCGGCAC
CACCGTGGTCTTCACGCCCGGCGAGCGCTCCGGCACCGCTT
CCAAGCGCTCCTACGACGAGGTGTACGGGGATGATGATATTC
TGGAGCAGGCGGCCGAGCGCCTGGGCGAGTTTGCTTACGGC
AAGCGCAGCCGTTCCGCACCGAAGGAAGAGGCGGTGTCCAT
CCCGCTGGACCACGGCAACCCCACGCCGAGCCTCAAGCCCG
TGACCTTGCAGCAGGTGCTGCCGACCGCGGCGCCGCGCCGG
GGGTTCAAGCGCGAGGGCGAGGATCTGTACCCCACCATGCA
GCTGATGGTGCCCAAGCGCCAGAAGCTGGAAGACGTGCTGG
AGACCATGAAGGTGGACCCGGACGTGCAGCCCGAGGTCAA
GGTGCGGCCCATCAAGCAGGTGGCCCCGGGCCTGGGCGTGC
AGACCGTGGACATCAAGATTCCCACGGAGCCCATGGAAACG
CAGACCGAGCCCATGATCAAGCCCAGCACCAGCACCATGGA
GGTGCAGACGGATCCCTGGATGCCATCGGCTCCTAGTCGAA
GACCCCGGCGCAAGTACGGCGCGGCCAGCCTGCTGATGCCC
AACTACGCGCTGCATCCTTCCATCATCCCCACGCCGGGCTAC
CGCGGCACGCGCTTCTACCGCGGTCATACCAGCAGCCGCCG
CCGCAAGACCACCACTCGCCGCCGCCGTCGCCGCACCGCCG
CTGCAACCACCCCTGCCGCCCTGGTGCGGAGAGTGTACCGC
CGCGGCCGCGCACCTCTGACCCTGCCGCGCGCGCGCTACCA
CCCGAGCATCGCCATTTAAACTTTCGCCtGCTTTGCAGATCAA
TGGCCCTCACATGCCGCCTTCGCGTTCCCATTACGGGCTACC

[0587]

GAGGAAGAAAACCGCGCCGTAGAAGGCTGGCGGGGAACGG
GATGCGTCGCCACCACCACCGGCGGCGGCGCGCCATCAGCA
AGCGGTTGGGGGGAGGCTTCCTGCCCGCGCTGATCCCCATCA
TCGCCGCGGCGATCGGGGCGATCCCCGGCATTGCTTCCGTGG
CGGTGCAGGCCTCTCAGCGCCACTGAGACACACTTGGAAAC
ATCTTGTAATAAACCaATGGACTCTGACGCTCCTGGTCCTGTG
ATGTGTTTTCGTAGACAGATGGAAGACATCAATTTTTCGTCC
CTGGCTCCGCGACACGGCACGCGGCCGTTCATGGGCACCTG
GAGCGACATCGGCACCAGCCAACTGAACGGGGGCGCCTTCA
ATTGGAGCAGTCTCTGGAGCGGGCTTAAGAATTTCGGGTCC
ACGCTTAAAACCTATGGCAGCAAGGCGTGGAACAGCACCAC
AGGGCAGGCGCTGAGGGATAAGCTGAAAGAGCAGAACTTC
CAGCAGAAGGTGGTCGATGGGCTCGCCTCGGGCATCAACGG
GGTGGTGGACCTGGCCAACCAGGCCGTGCAGCGGCAGATCA
ACAGCCGCCTGGACCCGGTGCCGCCCGCCGGCTCCGTGGAG
ATGCCGCAGGTGGAGGAGGAGCTGCCTCCCCTGGACAAGCG
GGGCGAGAAGCGACCCCGCCCCGATGCGGAGGAGACGCTG
CTGACGCACACGGACGAGCCGCCCCCGTACGAGGAGGCGG
TGAAACTGGGTCTGCCCACCACGCGGCCCATCGCGCCCCTG
GCCACCGGGGTGCTGAAACCCGAAAAGCCCGCGACCCTGG
ACTTGCCTCCTCCCCAGCCTTCCCGCCCCTCTACAGTGGCTA
AGCCCCTGCCGCCGGTGGCCGTGGCCCGCGCGCGACCCGGG
GGCACCGCCCGCCCTCATGCGAACTGGCAGAGCACTCTGAA
CAGCATCGTGGGTCTGGGAGTGCAGAGTGTGAAGCGCCGCC
GCTGCTATTAAACCTACCGTAGCGCTTAACTTGCTTGTCTGTG
TGTGTATGTATTATGTCGCCGCCGCCGCTGTCCACCAGAAGG
AGGAGTGAAGAGGCGCGTCGCCGAGTTGCAAGATGGCCAC
CCCATCGATGCTGCCCCAGTGGGCGTACATGCACATCGCCGG
ACAGGACGCTTCGGAGTACCTGAGTCCGGGTCTGGTGCAGT
TTGCCCGCGCCACAGACACCTACTTCAGTCTGGGGAACAAG
TTTAGGAACCCCACGGTGGCGCCCACGCACGATGTGACCAC
CGACCGCAGCCAGCGGCTGACGCTGCGCTTCGTGCCCGTGG
ACCGCGAGGACAACACCTACTCGTACAAAGTGCGCTACACG
CTGGCCGTGGGCGACAACCGCGTGCTGGACATGGCCAGCAC
CTACTTTGACATCCGCGGCGTGCTGGATCGGGGCCCTAGCTT
CAAACCCTACTCCGGCACCGCCTACAACAGTCTGGCCCCCA
AGGGAGCACCCAACACTTGTCAGTGGACATATAAAGCCGAT
GGTGAAACTGCCACAGAAAAAACCTATACATATGGAAATGC
ACCCGTGCAGGGCATTAACATCACAAAAGATGGTATTCAACT
TGGAACTGACACCGATGATCAGCCAATCTACGCAGATAAAAC
CTATCAGCCTGAACCTCAAGTGGGTGATGCTGAATGGCATGA
CATCACTGGTACTGATGAAAAGTATGGAGGCAGAGCTCTTAA
GCCTGATACCAAAATGAAGCCTTGTTATGGTTCTTTTGCCAA

[0588]

GCCTACTAATAAAGAAGGAGGTCAGGCAAATGTGAAAACAG
GAACAGGCACTACTAAAGAATATGACATAGACATGGCTTTCT
TTGACAACAGAAGTGCGGCTGCTGCTGGCCTAGCTCCAGAA
ATTGTTTTGTATACTGAAAATGTGGATTTGGAAACTCCAGATA
CCCATATTGTATACAAAGCAGGCACAGATGACAGCAGCTCTT
CTATTAATTTGGGTCAGCAAGCCATGCCCAACAGACCTAACT
ACATTGGTTTCAGAGACAACTTTATCGGGCTCATGTACTACA
ACAGCACTGGCAATATGGGGGTGCTGGCCGGTCAGGCTTCT
CAGCTGAATGCTGTGGTTGACTTGCAAGACAGAAACACCGA
GCTGTCCTACCAGCTCTTGCTTGACTCTCTGGGTGACAGAAC
CCGGTATTTCAGTATGTGGAATCAGGCGGTGGACAGCTATGA
TCCTGATGTGCGCATTATTGAAAATCATGGTGTGGAGGATGA
ACTTCCCAACTATTGTTTCCCTCTGGATGCTGTTGGCAGAAC
AGATACTTATCAGGGAATTAAGGCTAATGGAACTGATCAAAC
CACATGGACCAAAGATGACAGTGTCAATGATGCTAATGAGAT
AGGCAAGGGTAATCCATTCGCCATGGAAATCAACATCCAAGC
CAACCTGTGGAGGAACTTCCTCTACGCCAACGTGGCCCTGT
ACCTGCCCGACTCTTACAAGTACACGCCGGCCAATGTTACCC
TGCCCACCAACACCAACACCTACGATTACATGAACGGCCGG
GTGGTGGCGCCCTCGCTGGTGGACTCCTACATCAACATCGGG
GCGCGCTGGTCGCTGGATCCCATGGACAACGTGAACCCCTT
CAACCACCACCGCAATGCGGGGCTGCGCTACCGCTCCATGC
TCCTGGGCAACGGGCGCTACGTGCCCTTCCACATCCAGGTG
CCCCAGAAATTTTTCGCCATCAAGAGCCTCCTGCTCCTGCCC
GGGTCCTACACCTACGAGTGGAACTTCCGCAAGGACGTCAA
CATGATCCTGCAGAGCTCCCTCGGCAACGACCTGCGCACGG
ACGGGGCCTCCATCTCCTTCACCAGCATCAACCTCTACGCCA
CCTTCTTCCCCATGGCGCACAACACGGCCTCCACGCTCGAG
GCCATGCTGCGCAACGACACCAACGACCAGTCCTTCAACGA
CTACCTCTCGGCGGCCAACATGCTCTACCCCATCCCGGCCAA
CGCCACCAACGTGCCCATCTCCATCCCCTCGCGCAACTGGGC
CGCCTTCCGCGGCTGGTCCTTCACGCGTCTCAAGACCAAGG
AGACGCCCTCGCTGGGCTCCGGGTTCGACCCCTACTTCGTCT
ACTCGGGCTCCATCCCCTACCTCGACGGCACCTTCTACCTCA
ACCACACCTTCAAGAAGGTCTCCATCACCTTCGACTCCTCCG
TCAGCTGGCCCGGCAACGACCGGCTCCTGACGCCCAACGAG
TTCGAAATCAAGCGCACCGTCGACGGCGAGGGCTACAACGT
GGCCCAGTGCAACATGACCAAGGACTGGTTCCTGGTCCAGA
TGCTGGCCCACTACAACATCGGCTACCAGGGCTTCTACGTGC
CCGAGGGCTACAAGGACCGCATGTACTCCTTCTTCCGCAACT
TCCAGCCCATGAGCCGCCAGGTGGTGGACGAGGTCAACTAC
AAGGACTACCAGGCCGTCACCCTGGCCTACCAGCACAACAA
CTCGGGCTTCGTCGGCTACCTCGCGCCCACCATGCGCCAGG

[0589]

GCCAGCCCTACCCCGCCAACTACCCCTACCCGCTCATCGGCA
AGAGCGCCGTCACCAGCGTCACCCAGAAAAAGTTCCTCTGC
GACAGGGTCATGTGGCGCATCCCCTTCTCCAGCAACTTCATG
TCCATGGGCGCGCTCACCGACCTCGGCCAGAACATGCTCTAT
GCCAACTCCGCCCACGCGCTAGACATGAATTTCGAAGTCGA
CCCCATGGATGAGTCCACCCTTCTCTATGTTGTCTTCGAAGTC
TTCGACGTCGTCCGAGTGCACCAGCCCCACCGCGGCGTCAT
CGAGGCCGTCTACCTGCGCACCCCCTTCTCGGCCGGTAACGC
CACCACCTAAGCTCTTGCTTCTTGCAAGCCATGGCCGCGGGC
TCCGGCGAGCAGGAGCTCAGGGCCATCATCCGCGACCTGGG
CTGCGGGCCCTACTTCCTGGGCACCTTCGATAAGCGCTTCCC
GGGATTCATGGCCCCGCACAAGCTGGCCTGCGCCATCGTCA
ACACGGCCGGCCGCGAGACCGGGGGCGAGCACTGGCTGGC
CTTCGCCTGGAACCCGCGCTCGAACACCTGCTACCTCTTCGA
CCCCTTCGGGTTCTCGGACGAGCGCCTCAAGCAGATCTACC
AGTTCGAGTACGAGGGCCTGCTGCGCCGCAGCGCCCTGGCC
ACCGAGGACCGCTGCGTCACCCTGGAAAAGTCCACCCAGAC
CGTGCAGGGTCCGCGCTCGGCCGCCTGCGGGCTCTTCTGCT
GCATGTTCCTGCACGCCTTCGTGCACTGGCCCGACCGCCCCA
TGGACAAGAACCCCACCATGAACTTGCTGACGGGGGTGCCC
AACGGCATGCTCCAGTCGCCCCAGGTGGAACCCACCCTGCG
CCGCAACCAGGAGGCGCTCTACCGCTTCCTCAACTCCCACT
CCGCCTACTTTCGCTCCCACCGCGCGCGCATCGAGAAGGCC
ACCGCCTTCGACCGCATGAATCAAGACATGTAAACCGTGTGT
GTATGTTAAATGTCTTTAATAAACAGCACTTTCATGTTACACA
TGCATCTGAGATGATTTATTTAGAAATCGAAAGGGTTCTGCC
GGGTCTCGGCATGGCCCGCGGGCAGGGACACGTTGCGGAAC
TGGTACTTGGCCAGCCACTTGAACTCGGGGATCAGCAGTTT
GGGCAGCGGGGTGTCGGGGAAGGAGTCGGTCCACAGCTTC
CGCGTCAGTTGCAGGGCGCCCAGCAGGTCGGGCGCGGAGAT
CTTGAAATCGCAGTTGGGACCCGCGTTCTGCGCGCGGGAGT
TGCGGTACACGGGGTTGCAGCACTGGAACACCATCAGGGCC
GGGTGCTTCACGCTCGCCAGCACCGTCGCGTCGGTGATGCT
CTCCACGTCGAGGTCCTCGGCGTTGGCCATCCCGAAGGGGG
TCATCTTGCAGGTCTGCCTTCCCATGGTGGGCACGCACCCGG
GCTTGTGGTTGCAATCGCAGTGCAGGGGGATCAGCATCATCT
GGGCCTGGTCGGCGTTCATCCCCGGGTACATGGCCTTCATGA
AAGCCTCCAATTGCCTGAACGCCTGCTGGGCCTTGGCTCCCT
CGGTGAAGAAGACCCCGCAGGACTTGCTAGAGAACTGGTTG
GTGGCGCACCCGGCGTCGTGCACGCAGCAGCGCGCGTCGTT
GTTGGCCAGCTGCACCACGCTGCGCCCCCAGCGGTTCTGGG
TGATCTTGGCCCGGTCGGGGTTCTCCTTCAGCGCGCGCTGCC
CGTTCTCGCTCGCCACATCCATCTCGATCATGTGCTCCTTCTG

[0590]

GATCATGGTGGTCCCGTGCAGGCACCGCAGCTTGCCCTCGG
CCTCGGTGCACCCGTGCAGCCACAGCGCGCACCCGGTGCAC
TCCCAGTTCTTGTGGGCGATCTGGGAATGCGCGTGCACGAA
GCCCTGCAGGAAGCGGCCCATCATGGTGGTCAGGGTCTTGT
TGCTAGTGAAGGTCAGCGGAATGCCGCGGTGCTCCTCGTTG
ATGTACAGGTGGCAGATGCGGCGGTACACCTCGCCCTGCTCG
GGCATCAGCTGGAAGTTGGCTTTCAGGTCGGTCTCCACGCG
GTAGCGGTCCATCAGCATAGTCATGATTTCCATACCCTTCTCC
CAGGCCGAGACGATGGGCAGGCTCATAGGGTTCTTCACCAT
CATCTTAGCGCTAGCAGCCGCGGCCAGGGGGTCGCTCTCGTC
CAGGGTCTCAAAGCTCCGCTTGCCGTCCTTCTCGGTGATCCG
CACCGGGGGGGTAGCTGAAGCCCACGGCCGCCAGCTCCTCCT
CGGCCTGTCTTTCGTCCTCGCTGTCCTGGCTGACGTCCTGCA
GGACCACATGCTTGGTCTTGCGGGGTTTCTTCTTGGGCGGCA
GCGGCGGCGGAGATGTTGGAGATGGCGAGGGGGAGCGCGA
GTTCTCGCTCACCACTACTATCTCTTCCTCTTCTTGGTCCGAG
GCCACGCGGCGGTAGGTATGTCTCTTCGGGGGCAGAGGCGG
AGGCGACGGGCTCTCGCCGCCGCGACTTGGCGGATGGCTGG
CAGAGCCCCTTCCGCGTTCGGGGGTGCGCTCCCGGCGGCGC
TCTGACTGACTTCCTCCGCGGCCGGCCATTGTGTTCTCCTAG
GGAGGAACAACAAGCATGGAGACTCAGCCATCGCCAACCTC
GCCATCTGCCCCCACCGCCGACGAGAAGCAGCAGCAGCAGA
ATGAAAGCTTAACCGCCCCGCCGCCCAGCCCCGCCACCTCC
GACGCGGCCGTCCCAGACATGCAAGAGATGGAGGAATCCAT
CGAGATTGACCTGGGCTATGTGACGCCCGCGGAGCACGAGG
AGGAGCTGGCAGTGCGCTTTTCACAAGAAGAGATACACCAA
GAACAGCCAGAGCAGGAAGCAGAGAATGAGCAGAGTCAGG
CTGGGCTCGAGCATGACGGCGACTACCTCCACCTGAGCGGG
GGGGAGGACGCGCTCATCAAGCATCTGGCCCGGCAGGCCAC
CATCGTCAAGGATGCGCTGCTCGACCGCACCGAGGTGCCCC
TCAGCGTGGAGGAGCTCAGCCGCGCCTACGAGTTGAACCTC
TTCTCGCCGCGCGTGCCCCCCAAGCGCCAGCCCAATGGCAC
CTGCGAGCCCAACCCGCGCCTCAACTTCTACCCGGTCTTCGC
GGTGCCCGAGGCCCTGGCCACCTACCACATCTTTTTCAAGAA
CCAAAAGATCCCCGTCTCCTGCCGCGCCAACCGCACCCGCG
CCGACGCCCTTTTCAACCTGGGTCCCGGCGCCCGCCTACCTG
ATATCGCCTCCTTGGAAGAGGTTCCCAAGATCTTCGAGGGTC
TGGGCAGCGACGAGACTCGGGCCGCGAACGCTCTGCAAGG
AGAAGGAGGAGAGCATGAGCACCACAGCGCCCTGGTCGAG
TTGGAAGGCGACAACGCGCGGCTGGCGGTGCTCAAACGCA
CGGTCGAGCTGACCCATTTCGCCTACCCGGCTCTGAACCTGC
CCCCCAAAGTCATGAGCGCGGTCATGGACCAGGTGCTCATC
AAGCGCGCGTCGCCCATCTCCGAGGACGAGGGCATGCAAGA

[0591]

CTCCGAGGAGGGCAAGCCCGTGGTCAGCGACGAGCAGCTG
GCCCGGTGGCTGGGTCCTAATGCTAGTCCCCAGAGTTTGGAA
GAGCGGCGCAAACTCATGATGGCCGTGGTCCTGGTGACCGT
GGAGCTGGAGTGCCTGCGCCGCTTCTTCGCCGACGCGGAGA
CCCTGCGCAAGGTCGAGGAGAACCTGCACTACCTCTTCAGG
CACGGGTTCGTGCGCCAGGCCTGCAAGATCTCCAACGTGGA
GCTGACCAACCTGGTCTCCTACATGGGCATCTTGCACGAGAA
CCGCCTGGGGCAGAACGTGCTGCACACCACCCTGCGCGGGG
AGGCCCGGCGCGACTACATCCGCGACTGCGTCTACCTCTACC
TCTGCCACACCTGGCAGACGGGCATGGGCGTGTGGCAGCAG
TGTCTGGAGGAGCAGAACCTGAAAGAGCTCTGCAAGCTCCT
GCAGAAGAACCTCAAGGGTCTGTGGACCGGGTTCGACGAG
CGCACCACCGCCTCGGACCTGGCCGACCTCATTTTCCCCGAG
CGCCTCAGGCTGACGCTGCGCAACGGCCTGCCCGACTTTAT
GAGCCAAAGCATGTTGCAAAACTTTCGCTCTTTCATCCTCGA
ACGCTCCGGAATCCTGCCCGCCACCTGCTCCGCGCTGCCCTC
GGACTTCGTGCCGCTGACCTTCCGCGAGTGCCCCCCGCCGC
TGTGGAGCCACTGCTACCTGCTGCGCCTGGCCAACTACCTGG
CCTACCACTCGGACGTGATCGAGGACGTCAGCGGCGAGGGC
CTGCTCGAGTGCCACTGCCGCTGCAACCTCTGCACGCCGCA
CCGCTCCCTGGCCTGCAACCCCCAGCTGCTGAGCGAGACCC
AGATCATCGGCACCTTCGAGTTGCAAGGGCCCAGCGAAGGC
GAGGGTTCAGCCGCCAAGGGGGGTCTGAAACTCACCCCGG
GGCTGTGGACCTCGGCCTACTTGCGCAAGTTCGTGCCCGAG
GACTACCATCCCTTCGAGATCAGGTTCTACGAGGACCAATCC
CATCCGCCCAAGGCCGAGCTGTCGGCCTGCGTCATCACCCA
GGGGGCGATCCTGGCCCAATTGCAAGCCATCCAGAAATCCC
GCCAAGAATTCTTGCTGAAAAAGGGCCGCGGGGTCTACCTC
GACCCCCAGACCGGTGAGGAGCTCAACCCCGGCTTCCCCCA
GGATGCCCCGAGGAAACAAGAAGCTGAAAGTGGAGCTGCC
GCCCGTGGAGGATTTGGAGGAAGACTGGGAGAACAGCAGT
CAGGCAGAGGAGGAGGAGATGGAGGAAGACTGGGACAGCA
CTCAGGCAGAGGAGGACAGCCTGCAAGACAGTCTGGAGGA
AGACGAGGAGGAGGCAGAGGAGGAGGTGGAAGAAGCAGC
CGCCGCCAGACCGTCGTCCTCGGCGGGGGAGAAAGCAAGC
AGCACGGATACCATCTCCGCTCCGGGTCGGGGTCCCGCTCG
ACCACACAGTAGATGGGACGAGACCGGACGATTCCCGAACC
CCACCACCCAGACCGGTAAGAAGGAGCGGCAGGGATACAA
GTCCTGGCGGGGGCACAAAAACGCCATCGTCTCCTGCTTGC
AGGCCTGCGGGGGCAACATCTCCTTCACCCGGCGCTACCTG
CTCTTCCACCGCGGGGTGAACTTTCCCCGCAACATCTTGCAT
TACTACCGTCACCTCCACAGCCCCTACTACTTCCAAGAAGAG
GCAGCAGCAGCAGAAAAAGACCAGCAGAAAACCAGCAGCT

[0592]

AGAAAATCCACAGCGGCGGCAGCAGGTGGACTGAGGATCG
CGGCGAACGAGCCGGCGCAAACCCGGGAGCTGAGGAACCG
GATCTTTCCCACCCTCTATGCCATCTTCCAGCAGAGTCGGGG
GCAGGAGCAGGAACTGAAAGTCAAGAACCGTTCTCTGCGCT
CGCTCACCCGCAGTTGTCTGTATCACAAGAGCGAAGACCAA
CTTCAGCGCACTCTCGAGGACGCCGAGGCTCTCTTCAACAA
GTACTGCGCGCTCACTCTTAAAGAGTAGCCCGCGCCCGCCC
AGTCGCAGAAAAAGGCGGGAATTACGTCACCTGTGCCCTTC
GCCCTAGCCGCCTCCACCCATCATCATGAGCAAAGAGATTCC
CACGCCTTACATGTGGAGCTACCAGCCCCAGATGGGCCTGGC
CGCCGGTGCCGCCCAGGACTACTCCACCCGCATGAATTGGCT
CAGCGCCGGGCCCGCGATGATCTCACGGGTGAATGACATCC
GCGCCCACCGAAACCAGATACTCCTAGAACAGTCAGCGCTC
ACCGCCACGCCCCGCAATCACCTCAATCCGCGTAATTGGCCC
GCCGCCCTGGTGTACCAGGAAATTCCCCAGCCCACGACCGT
ACTACTTCCGCGAGACGCCCAGGCCGAAGTCCAGCTGACTA
ACTCAGGTGTCCAGCTGGCGGGCGGCGCCACCCTGTGTCGT
CACCGCCCCGCTCAGGGTATAAAGCGGCTGGTGATCCGGGG
CAGAGGCACACAGCTCAACGACGAGGTGGTGAGCTCTTCGC
TGGGTCTGCGACCTGACGGAGTCTTCCAACTCGCCGGATCG
GGGAGATCTTCCTTCACGCCTCGTCAGGCCGTCCTGACTTTG
GAGAGTTCGTCCTCGCAGCCCCGCTCGGGTGGCATCGGCAC
TCTCCAGTTCGTGGAGGAGTTCACTCCCTCGGTCTACTTCAA
CCCCTTCTCCGGCTCCCCCGGCCACTACCCGGACGAGTTCAT
CCCGAACTTCGACGCCATCAGCGAGTCGGTGGACGGCTACG
ATTGAAACTAATCACCCCCTTATCCAGTGAAATAAAGATCATA
TTGATGATGATTTTACAGAAATAAAAAATAATCATTTGATTTG
AAATAAAGATACAATCATATTGATGATTTGAGTTTAACAAAAA
AATAAAGAATCACTTACTTGAAATCTGATACCAGGTCTCTGT
CCATGTTTTCTGCCAACACCACTTCACTCCCCTCTTCCCAGC
TCTGGTACTGCAGGCCCCGGCGGGCTGCAAACTTCCTCCAC
ACGCTGAAGGGGATGTCAAATTCCTCCTGTCCCTCAATCTTC
ATTTTATCTTCTATCAGATGTCCAAAAAGCGCGTCCGGGTGG
ATGATGACTTCGACCCCGTCTACCCCTACGATGCAGACAACG
CACCGACCGTGCCCTTCATCAACCCCCCCTTCGTCTCTTCAG
ATGGATTCCAAGAGAAGCCCCTGGGGGTGTTGTCCCTGCGA
CTGGCCGACCCCGTCACCACCAAGAACGGGGAAATCACCCT
CAAGCTGGGAGAGGGGGTGGACCTCGATTCCTCGGGAAAAC
TCATCTCCAACACGGCCACCAAGGCCGCCGCCCCTCTCAGTT
TTTCCAACAACACCATTTCCCTTAACATGGATCACCCCTTTTA
CACTAAAGATGGAAAATTATCCTTACAAGTTTCTCCACCATTA
AATATACTGAGAACAAGCATTCTAAACACACTAGCTTTAGGT
TTTGGATCAGGTTTAGGACTCCGTGGCTCTGCCTTGGCAGTA

[0593]

CAGTTAGTCTCTCCACTTACATTTGATACTGATGGAAACATAA
AGCTTACCTTAGACAGAGGTTTGCATGTTACAACAGGAGATG
CAATTGAAAGCAACATAAGCTGGGCTAAAGGTTTAAAATTTG
AAGATGGAGCCATAGCAACCAACATTGGAAATGGGTTAGAG
TTTGGAAGCAGTAGTACAGAAACAGGTGTTGATGATGCTTAC
CCAATCCAAGTTAAACTTGGATCTGGCCTTAGCTTTGACAGT
ACAGGAGCCATAATGGCTGGTAACAAAGAAGACGATAAACT
CACTTTGTGGACAACACCTGATCCATCACCAAACTGTCAAAT
ACTCGCAGAAAATGATGCAAAACTAACACTTTGCTTGACTAA
ATGTGGTAGTCAAATACTGGCCACTGTGTCAGTCTTAGTTGT
AGGAAGTGGAAACCTAAACCCCATTACTGGCACCGTAAGCA
GTGCTCAGGTGTTTCTACGTTTTGATGCAAACGGTGTTCTTTT
AACAGAACATTCTACACTAAAAAAATACTGGGGGTATAGGCA
GGGAGATAGCATAGATGGCACTCCATATACCAATGCTGTAGG
ATTCATGCCCAATTTAAAAGCTTATCCAAAGTCACAAAGTTC
TACTACTAAAAATAATATAGTAGGGCAAGTATACATGAATGGA
GATGTTTCAAAACCTATGCTTCTCACTATAACCCTCAATGGTA
CTGATGACAGCAACAGTACATATTCAATGTCATTTTCATACAC
CTGGACTAATGGAAGCTATGTTGGAGCAACATTTGGGGCTAA
CTCTTATACCTTCTCATACATCGCCCAAGAATGAACACTGTAT
CCCACCCTGCATGCCAACCCTTCCCACCCCACTCTGTGGAAC
AAACTCTGAAACACAAAATAAAATAAAGTTCAAGTGTTTTAT
TGATTCAACAGTTTTACAGGATTCGAGCAGTTATTTTTCCTCC
ACCCTCCCAGGACATGGAATACACCACCCTCTCCCCCCGCAC
AGCCTTGAACATCTGAATGCCATTGGTGATGGACATGCTTTT
GGTCTCCACGTTCCACACAGTTTCAGAGCGAGCCAGTCTCG
GGTCGGTCAGGGAGATGAAACCCTCCGGGCACTCCCGCATC
TGCACCTCACAGCTCAACAGCTGAGGATTGTCCTCGGTGGT
CGGGATCACGGTTATCTGGAAGAAGCAGAAGAGCGGCGGTG
GGAATCATAGTCCGCGAACGGGATCGGCCGGTGGTGTCGCA
TCAGGCCCCGCAGCAGTCGCTGCCGCCGCCGCTCCGTCAAG
CTGCTGCTCAGGGGGTCCGGGTCCAGGGACTCCCTCAGCAT
GATGCCCACGGCCCTCAGCATCAGTCGTCTGGTGCGGCGGG
CGCAGCAGCGCATGCGGATCTCGCTCAGGTCGCTGCAGTAC
GTGCAACACAGAACCACCAGGTTGTTCAACAGTCCATAGTT
CAACACGCTCCAGCCGAAACTCATCGCGGGAAGGATGCTAC
CCACGTGGCCGTCGTACCAGATCCTCAGGTAAATCAAGTGGT
GCCCCCTCCAGAACACGCTGCCCACGTACATGATCTCCTTGG
GCATGTGGCGGTTCACCACCTCCCGGTACCACATCACCCTCT
GGTTGAACATGCAGCCCCGGATGATCCTGCGGAACCACAGG
GCCAGCACCGCCCCGCCCGCCATGCAGCGAAGAGACCCCGG
GTCCCGGCAATGGCAATGGAGGACCCACCGCTCGTACCCGT
GGATCATCTGGGAGCTGAACAAGTCTATGTTGGCACAGCAC

[0594]

AGGCATATGCTCATGCATCTCTTCAGCACTCTCAACTCCTCG GGGGTCAAAACCATATCCCAGGGCACGGGGAACTCTTGCAG GACAGCGAACCCCGCAGAACAGGGCAATCCTCGCACAGAA CTTACATTGTGCATGGACAGGGTATCGCAATCAGGCAGCACC GGGTGATCCTCCACCAGAGAAGCGCGGGTCTCGGTCTCCTC ACAGCGTGGTAAGGGGGCCGGCCGATACGGGTGATGGCGGG ACGCGGCTGATCGTGTTCGCGACCGTGTCATGATGCAGTTGC TTTCGGACATTTTCGTACTTGCTGTAGCAGAACCTGGTCCGG GCGCTGCACACCGATCGCCGGCGGCGGTCTCGGCGCTTGGA ACGCTCGGTGTTGAAATTGTAAAACAGCCACTCTCTCAGAC CGTGCAGCAGATCTAGGGCCTCAGGAGTGATGAAGATCCCA TCATGCCTGATGGCTCTGATCACATCGACCACCGTGGAATGG GCCAGACCCAGCCAGATGATGCAATTTTGTTGGGTTTCGGTG ACGGCGAGCCTCGGGAACAACGATGAAGTAAATGCAAGCG GTGCGTTCCAGCATGGTTAGTTAGCTGATCTGTAGAAAAAAC AAAAATGAACATTAAACCATGCTAGCCTGGCGAACAGGTGG GTAAATCGTTCTCTCCAGCACCAGGCAGGCCACGGGGTCTC CGGCGCGACCCTCGTAAAAATTGTCGCTATGATTGAAAACCA TCACAGAGAGACGTTCCCGGTGGCCGGCGTGAATGATTCGA CAAGATGAATACACCCCCGGAACATTGGCGTCCGCGAGTGA AAAAAAGCGCCCGAGGAAGCAATAAGGCACTACAATGCTCA GTCTCAAGTCCAGCAAAGCGATGCCATGCGGATGAAGCACA AAATTCTCAGGTGCGTACAAAATGTAATTACTCCCCTCCTGC ACAGGCAGCAAAGCCCCCGATCCCTCCAGGTACACATACAA AGCCTCAGCGTCCATAGCTTACCGAGCAGCAGCACACAACA GGCGCAAGAGTCAGAGAAAGGCTGAGCTCTAACCTGTCCAC CCGCTCTCTGCTCAATATATAGCCCAGATCTACACTGACGTAA AGGCCAAAGTCTAAAAATACCCGCCAAATAATCACACACGC CCAGCACACGCCCAGAAACCGGTGACACACTCAAAAAAATA CGCGCACTTCCTCAAACGCCCAAAACTGCCGTCATTTCCGG GTTCCCACGCTACGTCATCAAAACACGACTTTCAAATTCCGT CGACCGTTAAAAACGTCACCCGCCCCGCCCCTAACGGTCGC CCGTCTCTCAGCCAATCAGCGCCCCGCATCCCCAAATTCAAA CACCTCATTTGCATATTAACGCGCACAAAAAGTTTGAGGTATA TTATTGATGATG
SAM 体外转录模板 DNA (SEQ ID NO: 29370)；缺失核苷酸 7544-11175 的 VEE 基因组，加上最小 5’T7 启动子(粗斜体)
TAATACGACTCACTATAATGggcggcgcatgagagaagcccagaccaattacctacccaaaATGGagaaagttcacgttgacatcgaggaagacagcccattcctcagagctttgcagcggagcttcccgcagtttgaggtagaagccaagcaggtcactgataatgaccatgctaatgccagagcgttttcgcatctggcttcaaaactgatcgaaacggaggtggacccatccgacacgatccttgacattggaagtgcgcccgcccgcagaatgtattctaagcacaagtatcattgtatctgtccgatgagatgtg

[0595]

cggaagatccggacagattgtataagtatgcaactaagctgaagaaaaactgtaaggaaataactg
ataaggaattggacaagaaaatgaaggagctcgccgccgtcatgagcgaccctgacctggaaact
gagactatgtgcctccacgacgacgagtcgtgtcgctacgaagggcaagtcgctgtttaccaggat
gtatacgcggttgacggaccgacaagtctctatcaccaagccaataagggagttagagtcgcctac
tggataggctttgacaccacccctttatgtttaagaacttggctggagcatatccatcatactctacca
actgggccgacgaaaccgtgttaacggctcgtaacataggcctatgcagctctgacgttatggagc
ggtcacgtagagggatgtccattcttagaaagaagtatttgaaaccatccaacaatgttctattctctgt
tggctcgaccatctaccacgagaagagggacttactgaggagctggcacctgccgtctgtatttcac
ttacgtggcaagcaaaattacacatgtcggtgtgagactatagttagttgcgacgggtacgtcgttaa
aagaatagctatcagtccaggcctgtatgggaagccttcaggctatgctgctacgatgcaccgcga
gggattcttgtgctgcaaagtgacagacacattgaacggggagagggtctcttttcccgtgtgcacg
tatgtgccagctacattgtgtgaccaaatgactggcatactggcaacagatgtcagtgcggacgacg
cgcaaaaactgctggttgggctcaaccagcgtatagtcgtcaacggtcgcacccagagaaacacc
aataccatgaaaaattacctttgcccgtagtggcccaggcatttgctaggtgggcaaaggaatataa
ggaagatcaagaagatgaaaggccactaggactacgagatagacagttagtcatggggtgttgttg
ggcttttagaaggcacaagataacatctatttataagcgcccggatacccaaaccatcatcaaagtg
aacagcgatttccactcattcgtgctgcccaggataggcagtaacacattggagatcgggctgaga
acaagaatcaggaaaatgttagaggagcacaaggagccgtcacctctcattaccgccgaggacgt
acaagaagctaagtgcgcagccgatgaggctaaggaggtgcgtgaagccgaggagttgcgcgc
agctctaccacctttggcagctgatgttgaggagcccactctggaagccgatgtcgacttgatgttac
aagaggctggggccggctcagtggagacacctcgtggcttgataaaggttaccagctacgctggc
gaggacaagatcggctcttacgctgtgctttctccgcaggctgtactcaagagtgaaaaattatcttg
catccaccctctcgctgaacaagtcatagtgataacacactctggccgaaaagggcgttatgccgtg
gaaccataccatggtaaagtagtggtgccagagggacatgcaatacccgtccaggactttcaagct
ctgagtgaaagtgccaccattgtgtacaacgaacgtgagttcgtaaacaggtacctgcaccatattg
ccacacatggaggagcgctgaacactgatgaagaatattacaaaactgtcaagcccagcgagcac
gacggcgaatacctgtacgacatcgacaggaaacagtgcgtcaagaaagaactagtcactgggct
agggctcacaggcgagctggtggatcctcccttccatgaattcgcctacgagagtctgagaacacg
accagccgctccttaccaagtaccaaccataggggtgtatggcgtgccaggatcaggcaagtctg
gcatcattaaaagcgcagtcaccaaaaaagatctagtggtgagcgccaagaaagaaaactgtgca
gaaattataagggacgtcaagaaaatgaaagggctggacgtcaatgccagaactgtggactcagt
gctcttgaatggatgcaaacaccccgtagagaccctgtatattgacgaagcttttgcttgtcatgcag
gtactctcagagcgctcatagccattataagacctaaaaaggcagtgctctgcggggatcccaaac
agtgcggtttttttaacatgatgtgcctgaaagtgcattttaaccacgagatttgcacacaagtcttcca
caaaagcatctctcgccgttgcactaaatctgtgacttcggtcgtctcaaccttgttttacgacaaaaa
aatgagaacgacgaatccgaaagagactaagattgtgattgacactaccggcagtaccaaacctaa
gcaggacgatctcattctcacttgtttcagagggtgggtgaagcagttgcaaatagattacaaaggc
aacgaaataatgacggcagctgcctctcaagggctgacccgtaaaggtgtgtatgccgttcggtac
aaggtgaatgaaaatcctctgtacgcacccacctcagaacatgtgaacgtcctactgacccgcacg
gaggaccgcatcgtgtggaaaacactagccggcgacccatggataaaaacactgactgccaagta
ccctgggaatttcactgccacgatagaggagtggcaagcagagcatgatgccatcatgaggcaca
tcttggagagaccggaccctaccgacgtcttccagaataaggcaaacgtgtgttgggccaaggcttt
agtgccggtgctgaagaccgctggcatagacatgaccactgaacaatggaacactgtggattatttt

[0596]

gaaacggacaaagctcactcagcagagatagtattgaaccaactatgcgtgaggttctttggactcg
atctggactccggtctatttttctgcacccactgttccgttatccattaggaataatcactgggataactc
cccgtcgcctaacatgtacgggctgaataaagaagtggtccgtcagctctctcgcaggtacccaca
actgcctcgggcagttgccactggaagagtctatgacatgaacactggtacactgcgcaattatgat
ccgcgcataaacctagtacctgtaaacagaagactgcctcatgctttagtcctccaccataatgaaca
cccacagagtgacttttcttcattcgtcagcaaattgaagggcagaactgtcctggtggtcggggaa
aagttgtccgtcccaggcaaaatggttgactggttgtcagaccggcctgaggctaccttcagagctc
ggctggatttaggcatcccaggtgatgtgcccaaatatgacataatatttgttaatgtgaggaccccat
ataaataccatcactatcagcagtgtgaagaccatgccattaagcttagcatgttgaccaagaaagct
tgtctgcatctgaatcccggcggaacctgtgtcagcataggttatggttacgctgacagggccagcg
aaagcatcattggtgctatagcgcggcagttcaagttttcccgggtatgcaaaccgaaatcctcactt
gaagagacggaagttctgtttgtattcattgggtacgatcgcaaggcccgtacgcacaatccttacaa
gctttcatcaaccttgaccaacatttatacaggttccagactccacgaagccggatgtgcaccctcat
atcatgtggtgcgaggggatattgccacggccaccgaaggagtgattataaatgctgctaacagca
aaggacaacctggcggaggggtgtgcggagcgctgtataagaaattcccggaaagcttcgattta
cagccgatcgaagtaggaaaagcgcgactggtcaaaggtgcagctaaacatatcattcatgccgta
ggaccaaacttcaacaaagtttcggaggttgaaggtgacaaacagttggcagaggcttatgagtcc
atcgctaagattgtcaacgataacaattacaagtcagtagcgattccactgttgtccaccggcatctttt
ccgggaacaaagatcgactaacccaatcattgaaccatttgctgacagctttagacaccactgatgc
agatgtagccatatactgcagggacaagaaatgggaaatgactctcaaggaagcagtggctagga
gagaagcagtggaggagatatgcatatccgacgactcttcagtgacagaacctgatgcagagctg
gtgagggtgcatccgaagagttctttggctggaaggaagggctacagcacaagcgatggcaaaac
tttctcatatttggaagggaccaagtttcaccaggcggccaaggatatagcagaaattaatgccatgt
ggcccgttgcaacggaggccaatgagcaggtatgcatgtatatcctcggagaaagcatgagcagt
attaggtcgaaatgccccgtcgaagagtcggaagcctccacaccacctagcacgctgccttgcttgt
gcatccatgccatgactccagaaagagtacagcgcctaaaagcctcacgtccagaacaaattactg
tgtgctcatcctttccattgccgaagtatagaatcactggtgtgcagaagatccaatgctcccagccta
tattgttctcaccgaaagtgcctgcgtatattcatccaaggaagtatctcgtggaaacaccaccggta
gacgagactccggagccatcggcagagaaccaatccacagaggggacacctgaacaaccacca
cttataaccgaggatgagaccaggactagaacgcctgagccgatcatcatcgaagaggaagaag
aggatagcataagtttgctgtcagatggcccgacccaccaggtgctgcaagtcgaggcagacattc
acgggccgccctctgtatctagctcatcctggtccattcctcatgcatccgactttgatgtggacagttt
atccatacttgacaccctggagggagctagcgtgaccagcggggcaacgtcagccgagactaact
cttacttcgcaaagagtatggagtttctggcgcgaccggtgcctgcgcctcgaacagtattcaggaa
ccctccacatcccgctccgcgcacaagaacaccgtcacttgcacccagcagggcctgctcgagaa
ccagcctagtttccaccccgccaggcgtgaatagggtgatcactagagaggagctcgaggcgctt
accccgtcacgcactcctagcaggtcggtctcgagaaccagcctggtctccaacccgccaggcgt
aaatagggtgattacaagagaggagtttgaggcgttcgtagcacaacaacaatgacggtttgatgc
gggtgcatacatcttttcctccgacaccggtcaagggcatttacaacaaaaatcagtaaggcaaacg
gtgctatccgaagtggtgttggagaggaccgaattggagatttcgtatgccccgcgcctcgaccaa
gaaaaagaagaattactacgcaagaaattacagttaaatcccacacctgctaacagaagcagatac
cagtccaggaaggtggagaacatgaaagccataacagctagacgtattctgcaaggcctagggca
ttatttgaaggcagaaggaaaagtggagtgctaccgaaccctgcatcctgttcctttgtattcatctagt

[0597]

gtgaaccgtgccttttcaagccccaaggtcgcagtggaagcctgtaacgccatgttgaaagagaac tttccgactgtggcttcttactgtattattccagagtacgatgcctatttggacatggttgacggagcttc atgctgcttagacactgccagtttttgccctgcaaagctgcgcagctttccaaagaaacactcctattt ggaacccacaatacgatcggcagtgccttcagcgatccagaacacgctccagaacgtcctggcag ctgccacaaaaagaaattgcaatgtcacgcaaatgagagaattgcccgtattggattcggcggcctt taatgtggaatgcttcaagaaatatgcgtgtaataatgaatattgggaaacgtttaaagaaaaccccat caggcttactgaagaaaacgtggtaaattacattaccaaattaaaaggaccaaaagctgctgctctttt tgcgaagacacataatttgaatatgttgcaggacataccaatggacaggtttgtaatggacttaaaga gagacgtgaaagtgactccaggaacaaaacatactgaagaacggcccaaggtacaggtgatcca ggctgccgatccgctagcaacagcgtatctgtgcggaatccaccgagagctggttaggagattaaa tgcggtcctgcttccgaacattcatacactgtttgatatgtcggctgaagactttgacgctattatagcc gagcacttccagcctggggattgtgttctggaaactgacatcgcgtcgtttgataaaagtgaggacg acgccatggctctgaccgcgttaatgattctggaagacttaggtgtggacgcagagctgttgacgct gattgaggcggctttcggcgaaatttcatcaatacatttgcccactaaaactaaatttaaattcggagc catgatgaaatctggaatgttcctcacactgtttgtgaacacagtcattaacattgtaatcgcaagcag agtgttgagagaacggctaaccggatcaccatgtgcagcattcattggagatgacaatatcgtgaaa ggagtcaaatcggacaaattaatggcagacaggtgcgccacctggttgaatatggaagtcaagatt atagatgctgtggtgggcgagaaagcgccttatttctgtggagggtttattttgtgtgactccgtgacc ggcacagcgtgccgtgtggcagacccccctaaaaaggctgtttaagcttggcaaacctctggcagc agacgatgaacatgatgatgacaggagaagggcattgcatgaagagtcaacacgctggaaccga gtgggtattctttcagagctgtgcaaggcagtagaatcaaggtatgaaaccgtaggaacttccatcat agttatggccatgactactctagctagcagtgttaaatcattcagctacctgagaggggcccctataa ctctctacggcTAAcctgaatggactacgactTatcacgcccaaacatttacagccgcggtgtca aaaaccgcgtggacgtggttaacatccctgctgggaggatcagccgtaattattataattggcttggt gctggctactattgtggccatgtacgtgctgaccaaccagaaacataattgaatacagcagcaattg gcaagctgcttacatagaactcgcggcgattggcatgccgccttaaaatttttattttatttttctttttctttt tccgaatcggattttgtttttaatatttcAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAA
ChAdV68-GAG (全长 ChAd68-CMV-SIVGag-SV40 聚 A; SEQ ID NO: 29371); 缺失 E1 (核苷酸 577 至 3403)和 E3 (核苷酸 27,125-31,825)序列的 AC_000011.1; 对应的 ATCC VR-594 核苷酸在五个位置处被取代; 插入受 CMV 启动子/增强子控制的全长密码子优化的 SIVSME543 GAG 以替代缺失的 E1
gGATCTGCGACGCGAGGCTGGATGGCCTTCCCCATTATGATTC TTCTCGCTTCCGGCGGCATCGGGATGCCCGCGTTGCAGGCCA TGCTGTCCAGGCAGGTAGATGACGACCATCAGGGACAGCTT CAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGT TGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATC ACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGAC AGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCT CGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCT

[0598]

GTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAG
CTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTC
CAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACC
GCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGG
TAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAAC
AGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTT
CTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAG
TATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAA
AAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCT
GGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGC
AGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACG
GGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGAT
TTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTT
TTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAG
TTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGT
CTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGA
TAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTG
CAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTAT
CAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAG
TGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAAcTGT
TGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTG
CGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGC
TCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGAT
CAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCG
GTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTG
GCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATT
CTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGG
TGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCG
ACCGAGTTGCTCTTGCCCGGCGTCAACACGGGATAATACCGC
GCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACG
TTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAG
ATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTC
AGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAAC
AGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACA
CGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTG
AAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTT
GAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACA
TTTCCCCGAAAAGTGCCACCTGACGTCTAAGAAACCATTATT
ATCATGACATTAACCTATAAAAATAGGCGTATCACGAGGCCCT
TTCGTCTTgGCTGTGTCCTGCTTATCCACAACATTTTGCGCAC
GGTTATGTGGACCGGTCCGTTAATTAACATCaTCAATAATATA
CCTCAAACTTTTTGTGCGCGTTAATATGCAAATGAGGCGTTT
GAATTTGGGGAGGAAGGGCGGTGATTGGTCGAGGGATGAGC

[0599]

GACCGTTAGGGGCGGGGCGAGTGACGTTTTGATGACGTGGT
TGCGAGGAGGAGCCAGTTTGCAAGTTCTCGTGGGAAAAGTG
ACGTCAAACGAGGTGTGGTTTGAACACGGAAATACTCAATT
TTCCCGCGCTCTCTGACAGGAAATGAGGTGTTTCTGGGCGGA
TGCAAGTGAAAACGGGCCATTTTCGCGCGAAAACTGAATGA
GGAAGTGAAAATCTGAGTAATTTCGCGTTTATGGCAGGGAG
GAGTATTTGCCGAGGGCCGAGTAGACTTTGACCGATTACGTG
GGGGTTTCGATTACCGTGTTTTTCACCTAAATTTCCGCGTACG
GTGTCAAAGTCCGGTGTTTTTACGTAGGTGTCAGCTGATCGC
CAGGGTATTTAAACCTGCGCTCTCCAGTCAAGAGGCCACTCT
TGAGTGCCAGCGAGAAGAGTTTTCTCCTCCGCGCCGCGAGT
CAGATCTACACTTTGAAAGTAGGGATAACAGGGTAATgacattga
ttattgactagttGttaaTAGTAATCAATTACGGGGTCATTAGTTCATAG
CCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGC
CCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTC
AATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTC
CATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCAC
TTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTA
TTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCC
AGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCT
ACGTATTAGTCATCGCTATTACCATGgTGATGCGGTTTTGGCA
GTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATT
TCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTG
GCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCC
CGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGG
AGGTCTATATAAGCAGAgcTCGTTTAGTGAACCGTCAGATCGC
CTGGAACGCCATCCACGCTGTTTTGACCTCCATAGAAGACAG
CGATCGCGccaccATGGGAGCCAGGAACAGTGTGCTCTCAGGC
AAGAAGGCAGATGAGCTGGAGAAGATCAGGCTGAGACCCA
ATGGCAAGAAGAAGTACATGCTGAAGCATGTGGTCTGGGCA
GCCAATGAGCTGGACAGGTTTGGCCTGGCAGAGTCCCTGCT
GGACAACAAGGAGGGCTGCCAGAAGATCCTGTCAGTGCTGG
CCCCCCTGGTGCCCACTGGCTCAGAGAACCTGAAGAGCCTC
TACAACACAGTGTGTGTGATTTGGTGCATCCATGCAGAGGAG
AAGGTGAAGCACACTGAGGAGGCCAAGCAGATTGTGCAGA
GGCACCTGGTGGTGGAGACTGGCACAGCTGACAAGATGCCA
GCCACCTCCAGGCCCACAGCACCCCCCTCTGGCAGGGGGGG
CAACTACCCAGTCCAGCAAGTGGGGGGCAACTATGTGCACC
TGCCCCTGAGCCCCAGAACCCTGAATGCCTGGGTCAAGCTG
GTGGAGGAGAAGAAGTTTGGAGCAGAGGTGGTGCCTGGCT
TCCAGGCCCTGTCAGAGGGATGCACTCCCTATGACATCAACC
AGATGCTGAACTGTGTGGGAGAGCACCAGGCAGCAATGCAG
ATCATCAGGGAGATCATCAATGAGGAGGCTGCTGACTGGGA

[0600]

CCTCCAGCACCCCCAGCCTGGACCCCTCCCTGCAGGCCAGC
TGAGGGAGCCCAGAGGGAGTGACATAGCAGGCACCACCTCC
ACAGTGGAGGAGCAGATCCAGTGGATGTACAGGCAGCAGA
ACCCCATCCCTGTGGGCAACATCTACAGGAGGTGGATCCAG
CTGGGCCTCCAGAAGTGTGTCAGGATGTACAACCCAACCAA
CATCCTGGATGTGAAGCAGGGCCCCAAGGAGCCCTTCCAGT
CTTATGTGGACAGGTTCTACAAGAGCCTGAGAGCTGAGCAG
ACAGACCCTGCTGTGAAGAACTGGATGACCCAGACACTGCT
GATCCAGAATGCCAATCCTGACTGCAAGCTGGTGCTGAAGG
GGCTGGGGATGAATCCAACCCTGGAGGAGATGCTGACAGCC
TGCCAGGGCATTGGGGGACCTGGACAGAAGGCCAGGCTCAT
GGCAGAGGCTCTCAAGGAGGCCCTCAGACCAGACCAGCTG
CCATTTGCTGCTGTGCAGCAGAAGGGCCAGAGGAGGACCAT
CAAGTGCTGGAACTGTGGCAAGGAGGGCCACTCTGCCAGGC
AGTGCAGAGCCCCCAGGAGGCAGGGCTGCTGGGGCTGTGG
AAAGACAGGCCATGTGATGGCCAAGTGCCCAGAGAGGCAG
GCAGGCTTCCTGGGCTTTGGCCCCTGGGGCAAGAAGCCAAG
AAACTTCCCCATGGCCCAGATGCCCCAGGGCCTGACCCCCA
CAGCCCCCCCAGAGGACCCAGCTGTGGACCTGCTGAAGAAC
TACATGAAGATGGGCAGGAAGCAGAGGGAGAACAGGGAGA
GACCCTACAAGGAGGTGACTGAGGACCTGCTGCACCTGAAC
TCCCTGTTTGGGGAGGACCAGTGATAGGTTTAAACTCCCATT
TAAATGTGAGGGTTAATGCTTCGAGCAGACATGATAAGATAC
ATTGATGAGTTTGGACAAACCACAACTAGAATGCAGTGAAA
AAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTG
TAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATT
GCATTCATTTTATGTTTCAGGTTCAGGGGGAGATGTGGGAGG
TTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTAAAATAA
CTATAACGGTCCTAAGGTAGCGAGTGAGTAGTGTTCTGGGGC
GGGGGAGGACCTGCATGAGGGCCAGAATAACTGAAATCTGT
GCTTTTCTGTGTGTTGCAGCAGCATGAGCGGAAGCGGCTCCT
TTGAGGGAGGGGTATTCAGCCCTTATCTGACGGGGCGTCTCC
CCTCCTGGGCGGGAGTGCGTCAGAATGTGATGGGATCCACG
GTGGACGGCCGGCCCGTGCAGCCCGCGAACTCTTCAACCCT
GACCTATGCAACCCTGAGCTCTTCGTCGTTGGACGCAGCTGC
CGCCGCAGCTGCTGCATCTGCCGCCAGCGCCGTGCGCGGAA
TGGCCATGGGCGCCGGCTACTACGGCACTCTGGTGGCCAAC
TCGAGTTCCACCAATAATCCCGCCAGCCTGAACGAGGAGAA
GCTGTTGCTGCTGATGGCCCAGCTCGAGGCCTTGACCCAGC
GCCTGGGCGAGCTGACCCAGCAGGTGGCTCAGCTGCAGGA
GCAGACGCGGGCCGCGGTTGCCACGGTGAAATCCAAATAAA
AAATGAATCAATAAATAAACGGAGACGGTTGTTGATTTTAAC
ACAGAGTCTGAATCTTTATTTGATTTTTCGCGCGCGGTAGGC

[0601]

CCTGGACCACCGGTCTCGATCATTGAGCACCCGGTGGATCTT
TTCCAGGACCCGGTAGAGGTGGGCTTGGATGTTGAGGTACAT
GGGCATGAGCCCGTCCCGGGGGTGGAGGTAGCTCCATTGCA
GGGCCTCGTGCTCGGGGGTGGTGTTGTAAATCACCCAGTCAT
AGCAGGGGCGCAGGGCATGGTGTTGCACAATATCTTTGAGG
AGGAGACTGATGGCCACGGGCAGCCCTTTGGTGTAGGTGTT
TACAAATCTGTTGAGCTGGGAGGGATGCATGCGGGGGGAGA
TGAGGTGCATCTTGGCCTGGATCTTGAGATTGGCGATGTTAC
CGCCCAGATCCCGCCTGGGGTTCATGTTGTGCAGGACCACC
AGCACGGTGTATCCGGTGCACTTGGGGAATTTATCATGCAAC
TTGGAAGGGAAGGCGTGAAAGAATTTGGCGACGCCTTTGTG
CCCGCCCAGGTTTTCCATGCACTCATCCATGATGATGGCGAT
GGGCCCGTGGGCGGCGGCCTGGGCAAAGACGTTTCGGGGG
TCGGACACATCATAGTTGTGGTCCTGGGTGAGGTCATCATAG
GCCATTTTAATGAATTTGGGGCGGAGGGTGCCGGACTGGGG
GACAAAGGTtCCCTCGATCCCGGGGGCGTAGTTCCCCTCACA
GATCTGCATCTCCCAGGCTTTGAGCTCGGAGGGGGGGATCAT
GTCCACCTGCGGGGCGATAAAGAACACGGTTTCCGGGGCGG
GGGAGATGAGCTGGGCCGAAAGCAAGTTCCGGAGCAGCTG
GGACTTGCCGCAGCCGGTGGGGCCGTAGATGACCCCGATGA
CCGGCTGCAGGTGGTAGTTGAGGGAGAGACAGCTGCCGTCC
TCCCGGAGGAGGGGGGCCACCTCGTTCATCATCTCGCGCAC
GTGCATGTTCTCGCGCACCAGTTCCGCCAGGAGGCGCTCTCC
CCCCAGGGATAGGAGCTCCTGGAGCGAGGCGAAGTTTTTCA
GCGGCTTGAGTCCGTCGGCCATGGGCATTTTGGAGAGGGTTT
GTTGCAAGAGTTCCAGGCGGTCCCAGAGCTCGGTGATGTGC
TCTACGGCATCTCGATCCAGCAGACCTCCTCGTTTCGCGGGT
TGGGACGGCTGCGGGAGTAGGGCACCAGACGATGGGCGTCC
AGCGCAGCCAGGGTCCGGTCCTTCCAGGGTCGCAGCGTCCG
CGTCAGGGTGGTCTCCGTCACGGTGAAGGGGTGCGCGCCGG
GCTGGGCGCTTGCGAGGGTGCGCTTCAGGCTCATCCGGCTG
GTCGAAAACCGCTCCCGATCGGCGCCCTGCGCGTCGGCCAG
GTAGCAATTGACCATGAGTTCGTAGTTGAGCGCCTCGGCCGC
GTGGCCTTTGGCGCGGAGCTTACCTTTGGAAGTCTGCCCGC
AGGCGGGACAGAGGAGGGACTTGAGGGCGTAGAGCTTGGG
GGCGAGGAAGACGGACTCGGGGGCGTAGGCGTCCGCGCCG
CAGTGGGCGCAGACGGTCTCGCACTCCACGAGCCAGGTGAG
GTCGGGCTGGTCGGGGTCAAAAACCAGTTTCCCGCCGTTCT
TTTTGATGCGTTTCTTACCTTTGGTCTCCATGAGCTCGTGTCC
CCGCTGGGTGACAAAGAGGCTGTCCGTGTCCCCGTAGACCG
ACTTTATGGGCCGGTCCTCGAGCGGTGTGCCGCGGTCCTCCT
CGTAGAGGAACCCCGCCCACTCCGAGACGAAAGCCCGGGTC
CAGGCCAGCACGAAGGAGGCCACGTGGGACGGGTAGCGGT

[0602]

CGTTGTCCACCAGCGGGTCCACCTTTTCCAGGGTATGCAAAC
ACATGTCCCCCTCGTCCACATCCAGGAAGGTGATTGGCTTGT
AAGTGTAGGCCACGTGACCGGGGGTCCCGGCCGGGGGGGTA
TAAAAGGGTGCGGGTCCCTGCTCGTCCTCACTGTCTTCCGGA
TCGCTGTCCAGGAGCGCCAGCTGTTGGGGTAGGTATTCCCTC
TCGAAGGCGGGCATGACCTCGGCACTCAGGTTGTCAGTTTCT
AGAAACGAGGAGGATTTGATATTGACGGTGCCGGCGGAGAT
GCCTTTCAAGAGCCCCTCGTCCATCTGGTCAGAAAAGACGA
TCTTTTTGTTGTCGAGCTTGGTGGCGAAGGAGCCGTAGAGG
GCGTTGGAGAGGAGCTTGGCGATGGAGCGCATGGTCTGGTT
TTTTTCCTTGTCGGCGCGCTCCTTGGCGGCGATGTTGAGCTG
CACGTACTCGCGCGCCACGCACTTCCATTCGGGGAAGACGG
TGGTCAGCTCGTCGGGCACGATTCTGACCTGCCAGCCCCGAT
TATGCAGGGTGATGAGGTCCACACTGGTGGCCACCTCGCCG
CGCAGGGGCTCATTAGTCCAGCAGAGGCGTCCGCCCTTGCG
CGAGCAGAAGGGGGGCAGGGGGTCCAGCATGACCTCGTCG
GGGGGGTCGGCATCGATGGTGAAGATGCCGGGCAGGAGGTC
GGGGTCAAAGTAGCTGATGGAAGTGGCCAGATCGTCCAGGG
CAGCTTGCCATTCGCGCACGGCCAGCGCGCGCTCGTAGGGA
CTGAGGGGCGTGCCCCAGGGCATGGGATGGGTAAGCGCGGA
GGCGTACATGCCGCAGATGTCGTAGACGTAGAGGGGCTCCT
CGAGGATGCCGATGTAGGTGGGGTAGCAGCGCCCCCCGCGG
ATGCTGGCGCGCACGTAGTCATACAGCTCGTGCGAGGGGGC
GAGGAGCCCCGGGCCCAGGTTGGTGCGACTGGGCTTTTCGG
CGCGGTAGACGATCTGGCGGAAAATGGCATGCGAGTTGGAG
GAGATGGTGGGCCTTTGGAAGATGTTGAAGTGGGCGTGGGG
CAGTCCGACCGAGTCGCGGATGAAGTGGGCGTAGGAGTCTT
GCAGCTTGGCGACGAGCTCGGCGGTGACTAGGACGTCCAGA
GCGCAGTAGTCGAGGGTCTCCTGGATGATGTCATACTTGAGC
TGTCCCTTTTGTTTCCACAGCTCGCGGTTGAGAAGGAACTCT
TCGCGGTCCTTCCAGTACTCTTCGAGGGGGAACCCGTCCTGA
TCTGCACGGTAAGAGCCTAGCATGTAGAACTGGTTGACGGC
CTTGTAGGCGCAGCAGCCCTTCTCCACGGGGAGGGCGTAGG
CCTGGGCGGCCTTGCGCAGGGAGGTGTGCGTGAGGGCGAA
AGTGTCCCTGACCATGACCTTGAGGAACTGGTGCTTGAAGT
CGATATCGTCGCAGCCCCCCTGCTCCCAGAGCTGGAAGTCCG
TGCGCTTCTTGTAGGCGGGGTTGGGCAAAGCGAAAGTAACA
TCGTTGAAGAGGATCTTGCCCGCGCGGGGCATAAAGTTGCG
AGTGATGCGGAAAGGTTGGGGCACCTCGGCCCGGTTGTTGA
TGACCTGGGCGGCGAGCACGATCTCGTCGAAGCCGTTGATG
TTGTGGCCCACGATGTAGAGTTCCACGAATCGCGGACGGCC
CTTGACGTGGGGCAGTTTCTTGAGCTCCTCGTAGGTGAGCTC
GTCGGGGTCGCTGAGCCCGTGCTGCTCGAGCGCCCAGTCGG

[0603]

CGAGATGGGGGTTGGCGCGGAGGAAGGAAGTCCAGAGATC
CACGGCCAGGGCGGTTTGCAGACGGTCCCGGTACTGACGGA
ACTGCTGCCCGACGGCCATTTTTTCGGGGGTGACGCAGTAG
AAGGTGCGGGGGTCCCCGTGCCAGCGATCCCATTTGAGCTG
GAGGGCGAGATCGAGGGCGAGCTCGACGAGCCGGTCGTCC
CCGGAGAGTTTCATGACCAGCATGAAGGGGACGAGCTGCTT
GCCGAAGGACCCCATCCAGGTGTAGGTTTCCACATCGTAGGT
GAGGAAGAGCCTTTCGGTGCGAGGATGCGAGCCGATGGGGA
AGAACTGGATCTCCTGCCACCAATTGGAGGAATGGCTGTTGA
TGTGATGGAAGTAGAAATGCCGACGGCGCGCCGAACACTCG
TGCTTGTGTTTATACAAGCGGCCACAGTGCTCGCAACGCTGC
ACGGGATGCACGTGCTGCACGAGCTGTACCTGAGTTCCTTTG
ACGAGGAATTTCAGTGGGAAGTGGAGTCGTGGCGCCTGCAT
CTCGTGCTGTACTACGTCGTGGTGGTCGGCCTGGCCCTCTTC
TGCCTCGATGGTGGTCATGCTGACGAGCCCGCGCGGGAGGC
AGGTCCAGACCTCGGCGCGAGCGGGTCGGAGAGCGAGGAC
GAGGGCGCGCAGGCCGGAGCTGTCCAGGGTCCTGAGACGC
TGCGGAGTCAGGTCAGTGGGCAGCGGCGGCGCGCGGTTGA
CTTGCAGGAGTTTTTCCAGGGCGCGCGGGAGGTCCAGATGG
TACTTGATCTCCACCGCGCCATTGGTGGCGACGTCGATGGCT
TGCAGGGTCCCGTGCCCCTGGGGTGTGACCACCGTCCCCCG
TTTCTTCTTGGGCGGCTGGGGCGACGGGGGCGGTGCCTCTT
CCATGGTTAGAAGCGGCGGCGAGGACGCGCGCCGGGCGGC
AGGGGCGGCTCGGGGCCCGGAGGCAGGGGCGGCAGGGGCA
CGTCGGCGCCGCGCGCGGGTAGGTTCTGGTACTGCGCCCGG
AGAAGACTGGCGTGAGCGACGACGCGACGGTTGACGTCCT
GGATCTGACGCCTCTGGGTGAAGGCCACGGGACCCGTGAGT
TTGAACCTGAAAGAGAGTTCGACAGAATCAATCTCGGTATC
GTTGACGGCGGCCTGCCGCAGGATCTCTTGCACGTCGCCCG
AGTTGTCCTGGTAGGCGATCTCGGTCATGAACTGCTCGATCT
CCTCCTCTTGAAGGTCTCCGCGGCCGGCGCGCTCCACGGTG
GCCGCGAGGTCGTTGGAGATGCGGCCCATGAGCTGCGAGAA
GGCGTTCATGCCCGCCTCGTTCCAGACGCGGCTGTAGACCA
CGACGCCCTCGGGATCGCgGGCGCGCATGACCACCTGGGCG
AGGTTGAGCTCCACGTGGCGCGTGAAGACCGCGTAGTTGCA
GAGGCGCTGGTAGAGGTAGTTGAGCGTGGTGGCGATGTGCT
CGGTGACGAAGAAATACATGATCCAGCGGCGGAGCGGCATC
TCGCTGACGTCGCCCAGCGCCTCCAAACGTTCCATGGCCTCG
TAAAAGTCCACGGCGAAGTTGAAAAACTGGGAGTTGCGCGC
CGAGACGGTCAACTCCTCCTCCAGAAGACGGATGAGCTCGG
CGATGGTGGCGCGCACCTCGCGCTCGAAGGCCCCCGGGAGT
TCCTCCACTTCCTCTTCTTCCTCCTCCACTAACATCTCTTCTA
CTTCCTCCTCAGGCGGCAGTGGTGGCGGGGGAGGGGGCCTG

[0604]

CGTCGCCGGCGGCGCACGGGCAGACGGTCGATGAAGCGCTC
GATGGTCTCGCCGCGCCGGCGTCGCATGGTCTCGGTGACGG
CGCGCCCGTCCTCGCGGGGCCGCAGCGTGAAGACGCCGCCG
CGCATCTCCAGGTGGCCGGGGGGGTCCCCGTTGGGCAGGGA
GAGGGCGCTGACGATGCATCTTATCAATTGCCCCGTAGGGAC
TCCGCGCAAGGACCTGAGCGTCTCGAGATCCACGGGATCTG
AAAACCGCTGAACGAAGGCTTCGAGCCAGTCGCAGTCGCA
AGGTAGGCTGAGCACGGTTTCTTCTGGCGGGTCATGTTGGTT
GGGAGCGGGGCGGGCGATGCTGCTGGTGATGAAGTTGAAAT
AGGCGGTTCTGAGACGGCGGATGGTGGCGAGGAGCACCAG
GTCTTTGGGCCCGGCTTGCTGGATGCGCAGACGGTCGGCCAT
GCCCCAGGCGTGGTCCTGACACCTGGCCAGGTCCTTGTAGT
AGTCCTGCATGAGCCGCTCCACGGGCACCTCCTCCTCGCCCG
CGCGGCCGTGCATGCGCGTGAGCCCGAAGCCGCGCTGGGGC
TGGACGAGCGCCAGGTCGGCGACGACGCGCTCGGCGAGGA
TGGCTTGCTGGATCTGGGTGAGGGTGGTCTGGAAGTCATCA
AAGTCGACGAAGCGGTGGTAGGCTCCGGTGTTGATGGTGTA
GGAGCAGTTGGCCATGACGGACCAGTTGACGGTCTGGTGGC
CCGGACGCACGAGCTCGTGGTACTTGAGGCGCGAGTAGGCG
CGCGTGTCGAAGATGTAGTCGTTGCAGGTGCGCACCAGGTA
CTGGTAGCCGATGAGGAAGTGCGGCGGCGGCTGGCGGTAGA
GCGGCCATCGCTCGGTGGCGGGGGCGCCGGGCGCGAGGTCC
TCGAGCATGGTGCGGTGGTAGCCGTAGATGTACCTGGACATC
CAGGTGATGCCGGCGGCGGTGGTGGAGGCGCGCGGGAACT
CGCGGACGCGGTTCCAGATGTTGCGCAGCGGCAGGAAGTAG
TTCATGGTGGGCACGGTCTGGCCCGTGAGGCGCGCGCAGTC
GTGGATGCTCTATACGGGCAAAAACGAAAGCGGTCAGCGGC
TCGACTCCGTGGCCTGGAGGCTAAGCGAACGGGTTGGGCTG
CGCGTGTACCCCGGTTCGAATCTCGAATCAGGCTGGAGCCG
CAGCTAACGTGGTATTGGCACTCCCGTCTCGACCCAAGCCTG
CACCAACCCTCCAGGATACGGAGGCGGGTCGTTTTGCAACT
TTTTTTTGGAGGCCGGATGAGACTAGTAAGCGCGGAAAGCG
GCCGACCGCGATGGCTCGCTGCCGTAGTCTGGAGAAGAATC
GCCAGGGTTGCGTTGCGGTGTGCCCCGGTTCGAGGCCGGCC
GGATTCCGCGGCTAACGAGGGCGTGGCTGCCCCGTCGTTTC
CAAGACCCCATAGCCAGCCGACTTCTCCAGTTACGGAGCGA
GCCCCTCTTTTGTTTTGTTTGTTTTTGCCAGATGCATCCCGTA
CTGCGGCAGATGCGCCCCCACCACCCTCCACCGCAACAACA
GCCCCCTCCACAGCCGGCGCTTCTGCCCCCGCCCCAGCAGC
AACTTCCAGCCACGACCGCCGCGGCCGCCGTGAGCGGGGCT
GGACAGAGTTATGATCACCAGCTGGCCTTGGAAGAGGGCGA
GGGGCTGGCGCGCCTGGGGGCGTCGTCGCCGGAGCGGCAC
CCGCGCGTGCAGATGAAAAGGGACGCTCGCGAGGCCTACGT

[0605]

GCCCAAGCAGAACCTGTTCAGAGACAGGAGCGGCGAGGAG
CCCGAGGAGATGCGCGCGGCCCGGTTCCACGCGGGGCGGG
AGCTGCGGCGCGGCCTGGACCGAAAGAGGGTGCTGAGGGA
CGAGGATTTCGAGGCGGACGAGCTGACGGGGATCAGCCCCG
CGCGCGCGCACGTGGCCGCGGCCAACCTGGTCACGGCGTAC
GAGCAGACCGTGAAGGAGGAGAGCAACTTCCAAAAATCCTT
CAACAACCACGTGCGCACCCTGATCGCGCGCGAGGAGGTGA
CCCTGGGCCTGATGCACCTGTGGGACCTGCTGGAGGCCATC
GTGCAGAACCCCACCAGCAAGCCGCTGACGGCGCAGCTGTT
CCTGGTGGTGCAGCATAGTCGGGACAACGAAGCGTTCAGGG
AGGCGCTGCTGAATATCACCGAGCCCGAGGGCCGCTGGCTC
CTGGACCTGGTGAACATTCTGCAGAGCATCGTGGTGCAGGA
GCGCGGGCTGCCGCTGTCCGAGAAGCTGGCGGCCATCAACT
TCTCGGTGCTGAGTTTGGGCAAGTACTACGCTAGGAAGATCT
ACAAGACCCCGTACGTGCCCATAGACAAGGAGGTGAAGATC
GACGGGTTTTACATGCGCATGACCCTGAAAGTGCTGACCCTG
AGCGACGATCTGGGGGTGTACCGCAACGACAGGATGCACCG
TGCGGTGAGCGCCAGCAGGCGGCGCGAGCTGAGCGACCAG
GAGCTGATGCATAGTCTGCAGCGGGCCCTGACCGGGGCCGG
GACCGAGGGGGAGAGCTACTTTGACATGGGCGCGGACCTGC
ACTGGCAGCCCAGCCGCCGGGCCTTGGAGGCGGCGGCAGG
ACCCTACGTAGAAGAGGTGGACGATGAGGTGGACGAGGAG
GGCGAGTACCTGGAAGACTGATGGCGCGACCGTATTTTTGCT
AGATGCAACAACAACAGCCACCTCCTGATCCCGCGATGCGG
GCGGCGCTGCAGAGCCAGCCGTCCGGCATTAACTCCTCGGA
CGATTGGACCCAGGCCATGCAACGCATCATGGCGCTGACGA
CCCGCAACCCCGAAGCCTTTAGACAGCAGCCCCAGGCCAAC
CGGCTCTCGGCCATCCTGGAGGCCGTGGTGCCCTCGCGCTCC
AACCCCACGCACGAGAAGGTCCTGGCCATCGTGAACGCGCT
GGTGGAGAACAAGGCCATCCGCGGCGACGAGGCCGGCCTG
GTGTACAACGCGCTGCTGGAGCGCGTGGCCCGCTACAACAG
CACCAACGTGCAGACCAACCTGGACCGCATGGTGACCGACG
TGCGCGAGGCCGTGGCCCAGCGCGAGCGGTTCCACCGCGAG
TCCAACCTGGGATCCATGGTGGCGCTGAACGCCTTCCTCAGC
ACCCAGCCCGCCAACGTGCCCCGGGGCCAGGAGGACTACAC
CAACTTCATCAGCGCCCTGCGCCTGATGGTGACCGAGGTGC
CCCAGAGCGAGGTGTACCAGTCCGGGCCGGACTACTTCTTC
CAGACCAGTCGCCAGGGCTTGCAGACCGTGAACCTGAGCCA
GGCTTTCAAGAACTTGCAGGGCCTGTGGGGCGTGCAGGCCC
CGGTCGGGGACCGCGCGACGGTGTCGAGCCTGCTGACGCCG
AACTCGCGCCTGCTGCTGCTGCTGGTGGCCCCCTTCACGGA
CAGCGGCAGCATCAACCGCAACTCGTACCTGGGCTACCTGAT
TAACCTGTACCGCGAGGCCATCGGCCAGGCGCACGTGGACG

[0606]

AGCAGACCTACCAGGAGATCACCCACGTGAGCCGCGCCCTG
GGCCAGGACGACCCGGGCAACCTGGAAGCCACCCTGAACTT
TTTGCTGACCAACCGGTCGCAGAAGATCCCGCCCCAGTACG
CGCTCAGCACCGAGGAGGAGCGCATCCTGCGTTACGTGCAG
CAGAGCGTGGGCCTGTTCCTGATGCAGGAGGGGGCCACCCC
CAGCGCCGCGCTCGACATGACCGCGCGCAACATGGAGCCCA
GCATGTACGCCAGCAACCGCCCGTTCATCAATAAACTGATGG
ACTACTTGCATCGGGCGGCCGCCATGAACTCTGACTATTTCA
CCAACGCCATCCTGAATCCCCACTGGCTCCCGCCGCCGGGGT
TCTACACGGGCGAGTACGACATGCCCGACCCCAATGACGGG
TTCCTGTGGGACGATGTGGACAGCAGCGTGTTCTCCCCCCG
ACCGGGTGCTAACGAGCGCCCCTTGTGGAAGAAGGAAGGC
AGCGACCGACGCCCGTCCTCGGCGCTGTCCGGCCGCGAGGG
TGCTGCCGCGGCGGTGCCCGAGGCCGCCAGTCCTTTCCCGA
GCTTGCCCTTCTCGCTGAACAGTATCCGCAGCAGCGAGCTG
GGCAGGATCACGCGCCCGCGCTTGCTGGGCGAAGAGGAGTA
CTTGAATGACTCGCTGTTGAGACCCGAGCGGGAGAAGAACT
TCCCCAATAACGGGATAGAAAGCCTGGTGGACAAGATGAGC
CGCTGGAAGACGTATGCGCAGGAGCACAGGGACGATCCCCG
GGCGTCGCAGGGGGCCACGAGCCGGGGCAGCGCCGCCCGT
AAACGCCGGTGGCACGACAGGCAGCGGGGACAGATGTGGG
ACGATGAGGACTCCGCCGACGACAGCAGCGTGTTGGACTTG
GGTGGGAGTGGTAACCCGTTCGCTCACCTGCGCCCCCGTATC
GGGCGCATGATGTAAGAGAAACCGAAAATAAATGATACTCA
CCAAGGCCATGGCGACCAGCGTGCGTTCGTTTCTTCTCTGTT
GTTGTTGTATCTAGTATGATGAGGCGTGCGTACCCGGAGGGT
CCTCCTCCCTCGTACGAGAGCGTGATGCAGCAGGCGATGGC
GGCGGCGGCGATGCAGCCCCCGCTGGAGGCTCCTTACGTGC
CCCCGCGGTACCTGGCGCCTACGGAGGGGCGGAACAGCATT
CGTTACTCGGAGCTGGCACCCTTGTACGATACCACCCGGTTG
TACCTGGTGGACAACAAGTCGGCGGACATCGCCTCGCTGAA
CTACCAGAACGACCACAGCAACTTCCTGACCACCGTGGTGC
AGAACAATGACTTCACCCCCACGGAGGCCAGCACCCAGACC
ATCAACTTTGACGAGCGCTCGCGGTGGGGCGGCCAGCTGAA
AACCATCATGCACACCAACATGCCCAACGTGAACGAGTTCAT
GTACAGCAACAAGTTCAAGGCGCGGGTGATGGTCTCCCGCA
AGACCCCCAATGGGGTGACAGTGACAGAGGATTATGATGGT
AGTCAGGATGAGCTGAAGTATGAATGGGTGGAATTTGAGCT
GCCCGAAGGCAACTTCTCGGTGACCATGACCATCGACCTGAT
GAACAACGCCATCATCGACAATTACTTGGCGGTGGGGCGGC
AGAACGGGGTGCTGGAGAGCGACATCGGCGTGAAGTTCGA
CACTAGGAACTTCAGGCTGGGCTGGGACCCCGTGACCGAGC
TGGTCATGCCCGGGGTGTACACCAACGAGGCTTTCCATCCCG

[0607]

ATATTGTCTTGCTGCCCGGCTGCGGGGTGGACTTCACCGAGA
GCCGCCTCAGCAACCTGCTGGGCATTCGCAAGAGGCAGCCC
TTCCAGGAAGGCTTCCAGATCATGTACGAGGATCTGGAGGG
GGGCAACATCCCCGCGCTCCTGGATGTCGACGCCTATGAGA
AAAGCAAGGAGGATGCAGCAGCTGAAGCAACTGCAGCCGT
AGCTACCGCCTCTACCGAGGTCAGGGGCGATAATTTTGCAAG
CGCCGCAGCAGTGGCAGCGGCCGAGGCGGCTGAAACCGAA
AGTAAGATAGTCATTCAGCCGGTGGAGAAGGATAGCAAGAA
CAGGAGCTACAACGTACTACCGGACAAGATAAACACCGCCT
ACCGCAGCTGGTACCTAGCCTACAACTATGGCGACCCCGAG
AAGGGCGTGCGCTCCTGGACGCTGCTCACCACCTCGGACGT
CACCTGCGGCGTGGAGCAAGTCTACTGGTCGCTGCCCGACA
TGATGCAAGACCCGGTCACCTTCCGCTCCACGCGTCAAGTTA
GCAACTACCCGGTGGTGGGCGCCGAGCTCCTGCCCGTCTAC
TCCAAGAGCTTCTTCAACGAGCAGGCCGTCTACTCGCAGCA
GCTGCGCGCCTTCACCTCGCTTACGCACGTCTTCAACCGCTT
CCCCGAGAACCAGATCCTCGTCCGCCCGCCCGCGCCCACCA
TTACCACCGTCAGTGAAAACGTTCCTGCTCTCACAGATCACG
GGACCCTGCCGCTGCGCAGCAGTATCCGGGGAGTCCAGCGC
GTGACCGTTACTGACGCCAGACGCCGCACCTGCCCCTACGT
CTACAAGGCCCTGGGCATAGTCGCGCCGCGCGTCCTCTCGA
GCCGCACCTTCTAAATGTCCATTCTCATCTCGCCCAGTAATAA
CACCGGTTGGGGCCTGCGCGCGCCCAGCAAGATGTACGGAG
GCGCTCGCCAACGCTCCACGCAACACCCCGTGCGCGTGCGC
GGGCACTTCCGCGCTCCCTGGGGCGCCCTCAAGGGCCGCGT
GCGGTCGCGCACCACCGTCGACGACGTGATCGACCAGGTGG
TGGCCGACGCGCGCAACTACACCCCCGCCGCCGCGCCCGTC
TCCACCGTGGACGCCGTCATCGACAGCGTGGTGGCcGACGC
GCGCCGGTACGCCCGCGCCAAGAGCCGGCGGCGGCGCATCG
CCCGGCGGCACCGGAGCACCCCCGCCATGCGCGCGGCGCGA
GCCTTGCTGCGCAGGGCCAGGCGCACGGGACGCAGGGCCAT
GCTCAGGGCGGCCAGACGCGCGGCTTCAGGCGCCAGCGCC
GGCAGGACCCGGAGACGCGCGGCCACGGCGGCGGCAGCGG
CCATCGCCAGCATGTCCCGCCCGCGGCGAGGGAACGTGTAC
TGGGTGCGCGACGCCGCCACCGGTGTGCGCGTGCCCGTGCG
CACCCGCCCCCCTCGCACTTGAAGATGTTCACTTCGCGATGT
TGATGTGTCCCAGCGGCGAGGAGGATGTCCAAGCGCAAATT
CAAGGAAGAGATGCTCCAGGTCATCGCGCCTGAGATCTACG
GCCCTGCGGTGGTGAAGGAGGAAAGAAAGCCCCGCAAAAT
CAAGCGGGTCAAAAAGGACAAAAAGGAAGAAGAAAGTGAT
GTGGACGGATTGGTGGAGTTTGTGCGCGAGTTCGCCCCCCG
GCGGCGCGTGCAGTGGCGCGGGCGGAAGGTGCAACCGGTG
CTGAGACCCGGCACCACCGTGGTCTTCACGCCCGGCGAGCG

[0608]

CTCCGGCACCGCTTCCAAGCGCTCCTACGACGAGGTGTACG
GGGATGATGATATTCTGGAGCAGGCGGCCGAGCGCCTGGGC
GAGTTTGCTTACGGCAAGCGCAGCCGTTCCGCACCGAAGGA
AGAGGCGGTGTCCATCCCGCTGGACCACGGCAACCCCACGC
CGAGCCTCAAGCCCGTGACCTTGCAGCAGGTGCTGCCGACC
GCGGCGCCGCGCCGGGGGTTCAAGCGCGAGGGCGAGGATC
TGTACCCCACCATGCAGCTGATGGTGCCCAAGCGCCAGAAG
CTGGAAGACGTGCTGGAGACCATGAAGGTGGACCCGGACGT
GCAGCCCGAGGTCAAGGTGCGGCCCATCAAGCAGGTGGCCC
CGGGCCTGGGCGTGCAGACCGTGGACATCAAGATTCCCACG
GAGCCCATGGAAACGCAGACCGAGCCCATGATCAAGCCCAG
CACCAGCACCATGGAGGTGCAGACGGATCCCTGGATGCCAT
CGGCTCCTAGTCGAAGACCCCGGCGCAAGTACGGCGCGGCC
AGCCTGCTGATGCCCAACTACGCGCTGCATCCTTCCATCATC
CCCACGCCGGGCTACCGCGGCACGCGCTTCTACCGCGGTCAT
ACCAGCAGCCGCCGCCGCAAGACCACCACTCGCCGCCGCCG
TCGCCGCACCGCCGCTGCAACCACCCCTGCCGCCCTGGTGC
GGAGAGTGTACCGCCGCGGCCGCGCACCTCTGACCCTGCCG
CGCGCGCGCTACCACCCGAGCATCGCCATTTAAACTTTCGCCt
GCTTTGCAGATCAATGGCCCTCACATGCCGCCTTCGCGTTCC
CATTACGGGCTACCGAGGAAGAAAACCGCGCCGTAGAAGGC
TGGCGGGGAACGGGATGCGTCGCCACCACCACCGGCGGCG
GCGCGCCATCAGCAAGCGGTTGGGGGGAGGCTTCCTGCCCG
CGCTGATCCCCATCATCGCCGCGGCGATCGGGGCGATCCCCG
GCATTGCTTCCGTGGCGGTGCAGGCCTCTCAGCGCCACTGA
GACACACTTGGAAACATCTTGTAATAAACCaATGGACTCTGA
CGCTCCTGGTCCTGTGATGTGTTTTCGTAGACAGATGGAAGA
CATCAATTTTTCGTCCCTGGCTCCGCGACACGGCACGCGGCC
GTTCATGGGCACCTGGAGCGACATCGGCACCAGCCAACTGA
ACGGGGGCGCCTTCAATTGGAGCAGTCTCTGGAGCGGGCTT
AAGAATTTCGGGTCCACGCTTAAAACCTATGGCAGCAAGGC
GTGGAACAGCACCACAGGGCAGGCGCTGAGGGATAAGCTG
AAAGAGCAGAACTTCCAGCAGAAGGTGGTCGATGGGCTCGC
CTCGGGCATCAACGGGGTGGTGGACCTGGCCAACCAGGCCG
TGCAGCGGCAGATCAACAGCCGCCTGGACCCGGTGCCGCCC
GCCGGCTCCGTGGAGATGCCGCAGGTGGAGGAGGAGCTGCC
TCCCCTGGACAAGCGGGGCGAGAAGCGACCCCGCCCCGATG
CGGAGGAGACGCTGCTGACGCACACGGACGAGCCGCCCCC
GTACGAGGAGGCGGTGAAACTGGGTCTGCCCACCACGCGGC
CCATCGCGCCCCTGGCCACCGGGGTGCTGAAACCCGAAAAG
CCCGCGACCCTGGACTTGCCTCCTCCCCAGCCTTCCCGCCCC
TCTACAGTGGCTAAGCCCCTGCCGCCGGTGGCCGTGGCCCG
CGCGCGACCCGGGGGCACCGCCCGCCCTCATGCGAACTGGC

[0609]

AGAGCACTCTGAACAGCATCGTGGGTCTGGGAGTGCAGAGT
GTGAAGCGCCGCCGCTGCTATTAAACCTACCGTAGCGCTTAA
CTTGCTTGTCTGTGTGTGTATGTATTATGTCGCCGCCGCCGCT
GTCCACCAGAAGGAGGAGTGAAGAGGCGCGTCGCCGAGTT
GCAAGATGGCCACCCCATCGATGCTGCCCCAGTGGGCGTACA
TGCACATCGCCGGACAGGACGCTTCGGAGTACCTGAGTCCG
GGTCTGGTGCAGTTTGCCCGCGCCACAGACACCTACTTCAG
TCTGGGGAACAAGTTTAGGAACCCCACGGTGGCGCCCACGC
ACGATGTGACCACCGACCGCAGCCAGCGGCTGACGCTGCGC
TTCGTGCCCGTGGACCGCGAGGACAACACCTACTCGTACAA
AGTGCGCTACACGCTGGCCGTGGGCGACAACCGCGTGCTGG
ACATGGCCAGCACCTACTTTGACATCCGCGGCGTGCTGGATC
GGGGCCCTAGCTTCAAACCCTACTCCGGCACCGCCTACAAC
AGTCTGGCCCCCAAGGGAGCACCCAACACTTGTCAGTGGAC
ATATAAAGCCGATGGTGAAACTGCCACAGAAAAAACCTATAC
ATATGGAAATGCACCCGTGCAGGGCATTAACATCACAAAAGA
TGGTATTCAACTTGGAACTGACACCGATGATCAGCCAATCTA
CGCAGATAAAACCTATCAGCCTGAACCTCAAGTGGGTGATGC
TGAATGGCATGACATCACTGGTACTGATGAAAAGTATGGAGG
CAGAGCTCTTAAGCCTGATACCAAAATGAAGCCTTGTTATGG
TTCTTTTGCCAAGCCTACTAATAAAGAAGGAGGTCAGGCAAA
TGTGAAAACAGGAACAGGCACTACTAAAGAATATGACATAG
ACATGGCTTTCTTTGACAACAGAAGTGCGGCTGCTGCTGGC
CTAGCTCCAGAAATTGTTTTGTATACTGAAAATGTGGATTTGG
AAACTCCAGATACCCATATTGTATACAAAGCAGGCACAGATG
ACAGCAGCTCTTCTATTAATTTGGGTCAGCAAGCCATGCCCA
ACAGACCTAACTACATTGGTTTCAGAGACAACTTTATCGGGC
TCATGTACTACAACAGCACTGGCAATATGGGGGTGCTGGCCG
GTCAGGCTTCTCAGCTGAATGCTGTGGTTGACTTGCAAGAC
AGAAACACCGAGCTGTCCTACCAGCTCTTGCTTGACTCTCTG
GGTGACAGAACCCGGTATTTCAGTATGTGGAATCAGGCGGTG
GACAGCTATGATCCTGATGTGCGCATTATTGAAAATCATGGTG
TGGAGGATGAACTTCCCAACTATTGTTTCCCTCTGGATGCTG
TTGGCAGAACAGATACTTATCAGGGAATTAAGGCTAATGGAA
CTGATCAAACCACATGGACCAAAGATGACAGTGTCAATGAT
GCTAATGAGATAGGCAAGGGTAATCCATTCGCCATGGAAATC
AACATCCAAGCCAACCTGTGGAGGAACTTCCTCTACGCCAA
CGTGGCCCTGTACCTGCCCGACTCTTACAAGTACACGCCGGC
CAATGTTACCCTGCCCACCAACACCAACACCTACGATTACAT
GAACGGCCGGGTGGTGGCGCCCTCGCTGGTGGACTCCTACA
TCAACATCGGGGCGCGCTGGTCGCTGGATCCCATGGACAAC
GTGAACCCCTTCAACCACCACCGCAATGCGGGGCTGCGCTA
CCGCTCCATGCTCCTGGGCAACGGGCGCTACGTGCCCTTCCA

[0610]

CATCCAGGTGCCCCAGAAATTTTTCGCCATCAAGAGCCTCCT
GCTCCTGCCCGGGTCCTACACCTACGAGTGGAACTTCCGCA
AGGACGTCAACATGATCCTGCAGAGCTCCCTCGGCAACGAC
CTGCGCACGGACGGGGCCTCCATCTCCTTCACCAGCATCAAC
CTCTACGCCACCTTCTTCCCCATGGCGCACAACACGGCCTCC
ACGCTCGAGGCCATGCTGCGCAACGACACCAACGACCAGTC
CTTCAACGACTACCTCTCGGCGGCCAACATGCTCTACCCCAT
CCCGGCCAACGCCACCAACGTGCCCATCTCCATCCCCTCGCG
CAACTGGGCCGCCTTCCGCGGCTGGTCCTTCACGCGTCTCA
AGACCAAGGAGACGCCCTCGCTGGGCTCCGGGTTCGACCCC
TACTTCGTCTACTCGGGCTCCATCCCCTACCTCGACGGCACC
TTCTACCTCAACCACACCTTCAAGAAGGTCTCCATCACCTTC
GACTCCTCCGTCAGCTGGCCCGGCAACGACCGGCTCCTGAC
GCCCAACGAGTTCGAAATCAAGCGCACCGTCGACGGCGAGG
GCTACAACGTGGCCCAGTGCAACATGACCAAGGACTGGTTC
CTGGTCCAGATGCTGGCCCACTACAACATCGGCTACCAGGGC
TTCTACGTGCCCGAGGGCTACAAGGACCGCATGTACTCCTTC
TTCCGCAACTTCCAGCCCATGAGCCGCCAGGTGGTGGACGA
GGTCAACTACAAGGACTACCAGGCCGTCACCCTGGCCTACC
AGCACAACAACTCGGGCTTCGTCGGCTACCTCGCGCCCACC
ATGCGCCAGGGCCAGCCCTACCCCGCCAACTACCCCTACCCG
CTCATCGGCAAGAGCGCCGTCACCAGCGTCACCCAGAAAAA
GTTCCTCTGCGACAGGGTCATGTGGCGCATCCCCTTCTCCAG
CAACTTCATGTCCATGGGCGCGCTCACCGACCTCGGCCAGA
ACATGCTCTATGCCAACTCCGCCCACGCGCTAGACATGAATT
TCGAAGTCGACCCCATGGATGAGTCCACCCTTCTCTATGTTG
TCTTCGAAGTCTTCGACGTCGTCCGAGTGCACCAGCCCCAC
CGCGGCGTCATCGAGGCCGTCTACCTGCGCACCCCCTTCTCG
GCCGGTAACGCCACCACCTAAGCTCTTGCTTCTTGCAAGCCA
TGGCCGCGGGCTCCGGCGAGCAGGAGCTCAGGGCCATCATC
CGCGACCTGGGCTGCGGGCCCTACTTCCTGGGCACCTTCGAT
AAGCGCTTCCCGGGATTCATGGCCCCGCACAAGCTGGCCTG
CGCCATCGTCAACACGGCCGGCCGCGAGACCGGGGGCGAG
CACTGGCTGGCCTTCGCCTGGAACCCGCGCTCGAACACCTG
CTACCTCTTCGACCCCTTCGGGTTCTCGGACGAGCGCCTCAA
GCAGATCTACCAGTTCGAGTACGAGGGCCTGCTGCGCCGCA
GCGCCCTGGCCACCGAGGACCGCTGCGTCACCCTGGAAAAG
TCCACCCAGACCGTGCAGGGTCCGCGCTCGGCCGCCTGCGG
GCTCTTCTGCTGCATGTTCCTGCACGCCTTCGTGCACTGGCC
CGACCGCCCCATGGACAAGAACCCCACCATGAACTTGCTGA
CGGGGGTGCCCAACGGCATGCTCCAGTCGCCCCAGGTGGAA
CCCACCCTGCGCCGCAACCAGGAGGCGCTCTACCGCTTCCT
CAACTCCCACTCCGCCTACTTTCGCTCCCACCGCGCGCGCAT

[0611]

CGAGAAGGCCACCGCCTTCGACCGCATGAATCAAGACATGT
AAACCGTGTGTGTATGTTAAATGTCTTTAATAAACAGCACTTT
CATGTTACACATGCATCTGAGATGATTTATTTAGAAATCGAAA
GGGTTCTGCCGGGTCTCGGCATGGCCCGCGGGCAGGGACAC
GTTGCGGAACTGGTACTTGGCCAGCCACTTGAACTCGGGGA
TCAGCAGTTTGGGCAGCGGGGTGTCGGGGAAGGAGTCGGTC
CACAGCTTCCGCGTCAGTTGCAGGGCGCCCAGCAGGTCGGG
CGCGGAGATCTTGAAATCGCAGTTGGGACCCGCGTTCTGCG
CGCGGGAGTTGCGGTACACGGGGTTGCAGCACTGGAACACC
ATCAGGGCCGGGTGCTTCACGCTCGCCAGCACCGTCGCGTC
GGTGATGCTCTCCACGTCGAGGTCCTCGGCGTTGGCCATCCC
GAAGGGGGTCATCTTGCAGGTCTGCCTTCCCATGGTGGGCA
CGCACCCGGGCTTGTGGTTGCAATCGCAGTGCAGGGGGATC
AGCATCATCTGGGCCTGGTCGGCGTTCATCCCCGGGTACATG
GCCTTCATGAAAGCCTCCAATTGCCTGAACGCCTGCTGGGCC
TTGGCTCCCTCGGTGAAGAAGACCCCGCAGGACTTGCTAGA
GAACTGGTTGGTGGCGCACCCGGCGTCGTGCACGCAGCAGC
GCGCGTCGTTGTTGGCCAGCTGCACCACGCTGCGCCCCCAG
CGGTTCTGGGTGATCTTGGCCCGGTCGGGGTTCTCCTTCAGC
GCGCGCTGCCCGTTCTCGCTCGCCACATCCATCTCGATCATG
TGCTCCTTCTGGATCATGGTGGTCCCGTGCAGGCACCGCAGC
TTGCCCTCGGCCTCGGTGCACCCGTGCAGCCACAGCGCGCA
CCCGGTGCACTCCCAGTTCTTGTGGGCGATCTGGGAATGCGC
GTGCACGAAGCCCTGCAGGAAGCGGCCCATCATGGTGGTCA
GGGTCTTGTTGCTAGTGAAGGTCAGCGGAATGCCGCGGTGC
TCCTCGTTGATGTACAGGTGGCAGATGCGGCGGTACACCTCG
CCCTGCTCGGGCATCAGCTGGAAGTTGGCTTTCAGGTCGGTC
TCCACGCGGTAGCGGTCCATCAGCATAGTCATGATTTCCATAC
CCTTCTCCCAGGCCGAGACGATGGGCAGGCTCATAGGGTTCT
TCACCATCATCTTAGCGCTAGCAGCCGCGGCCAGGGGGTCGC
TCTCGTCCAGGGTCTCAAAGCTCCGCTTGCCGTCCTTCTCGG
TGATCCGCACCGGGGGGTAGCTGAAGCCCACGGCCGCCAGC
TCCTCCTCGGCCTGTCTTTCGTCCTCGCTGTCCTGGCTGACG
TCCTGCAGGACCACATGCTTGGTCTTGCGGGGTTTCTTCTTG
GGCGGCAGCGGCGGCGGAGATGTTGGAGATGGCGAGGGGG
AGCGCGAGTTCTCGCTCACCACTACTATCTCTTCCTCTTCTTG
GTCCGAGGCCACGCGGCGGTAGGTATGTCTCTTCGGGGGCA
GAGGCGGAGGCGACGGGCTCTCGCCGCCGCGACTTGGCGGA
TGGCTGGCAGAGCCCCTTCCGCGTTCGGGGGTGCGCTCCCG
GCGGCGCTCTGACTGACTTCCTCCGCGGCCGGCCATTGTGTT
CTCCTAGGGAGGAACAACAAGCATGGAGACTCAGCCATCGC
CAACCTCGCCATCTGCCCCCACCGCCGACGAGAAGCAGCAG
CAGCAGAATGAAAGCTTAACCGCCCCGCCGCCCAGCCCCGC

[0612]

CACCTCCGACGCGGCCGTCCCAGACATGCAAGAGATGGAGG
AATCCATCGAGATTGACCTGGGCTATGTGACGCCCGCGGAGC
ACGAGGAGGAGCTGGCAGTGCGCTTTTCACAAGAAGAGATA
CACCAAGAACAGCCAGAGCAGGAAGCAGAGAATGAGCAGA
GTCAGGCTGGGCTCGAGCATGACGGCGACTACCTCCACCTG
AGCGGGGGGGAGGACGCGCTCATCAAGCATCTGGCCCGGCA
GGCCACCATCGTCAAGGATGCGCTGCTCGACCGCACCGAGG
TGCCCCTCAGCGTGGAGGAGCTCAGCCGCGCCTACGAGTTG
AACCTCTTCTCGCCGCGCGTGCCCCCCAAGCGCCAGCCCAA
TGGCACCTGCGAGCCCAACCCGCGCCTCAACTTCTACCCGG
TCTTCGCGGTGCCCGAGGCCCTGGCCACCTACCACATCTTTT
TCAAGAACCAAAAGATCCCCGTCTCCTGCCGCGCCAACCGC
ACCCGCGCCGACGCCCTTTTCAACCTGGGTCCCGGCGCCCG
CCTACCTGATATCGCCTCCTTGGAAGAGGTTCCCAAGATCTT
CGAGGGTCTGGGCAGCGACGAGACTCGGGCCGCGAACGCT
CTGCAAGGAGAAGGAGGAGAGCATGAGCACCACAGCGCCC
TGGTCGAGTTGGAAGGCGACAACGCGCGGCTGGCGGTGCTC
AAACGCACGGTCGAGCTGACCCATTTCGCCTACCCGGCTCT
GAACCTGCCCCCCAAAGTCATGAGCGCGGTCATGGACCAGG
TGCTCATCAAGCGCGCGTCGCCCATCTCCGAGGACGAGGGC
ATGCAAGACTCCGAGGAGGGCAAGCCCGTGGTCAGCGACG
AGCAGCTGGCCCGGTGGCTGGGTCCTAATGCTAGTCCCCAG
AGTTTGGAAGAGCGGCGCAAACTCATGATGGCCGTGGTCCT
GGTGACCGTGGAGCTGGAGTGCCTGCGCCGCTTCTTCGCCG
ACGCGGAGACCCTGCGCAAGGTCGAGGAGAACCTGCACTA
CCTCTTCAGGCACGGGTTCGTGCGCCAGGCCTGCAAGATCT
CCAACGTGGAGCTGACCAACCTGGTCTCCTACATGGGCATCT
TGCACGAGAACCGCCTGGGGCAGAACGTGCTGCACACCACC
CTGCGCGGGGAGGCCCGGCGCGACTACATCCGCGACTGCGT
CTACCTCTACCTCTGCCACACCTGGCAGACGGGCATGGGCGT
GTGGCAGCAGTGTCTGGAGGAGCAGAACCTGAAAGAGCTC
TGCAAGCTCCTGCAGAAGAACCTCAAGGGTCTGTGGACCGG
GTTCGACGAGCGCACCACCGCCTCGGACCTGGCCGACCTCA
TTTTCCCCGAGCGCCTCAGGCTGACGCTGCGCAACGGCCTG
CCCGACTTTATGAGCCAAAGCATGTTGCAAAACTTTCGCTCT
TTCATCCTCGAACGCTCCGGAATCCTGCCCGCCACCTGCTCC
GCGCTGCCCTCGGACTTCGTGCCGCTGACCTTCCGCGAGTG
CCCCCCGCCGCTGTGGAGCCACTGCTACCTGCTGCGCCTGG
CCAACTACCTGGCCTACCACTCGGACGTGATCGAGGACGTC
AGCGGCGAGGGCCTGCTCGAGTGCCACTGCCGCTGCAACCT
CTGCACGCCGCACCGCTCCCTGGCCTGCAACCCCCAGCTGC
TGAGCGAGACCCAGATCATCGGCACCTTCGAGTTGCAAGGG
CCCAGCGAAGGCGAGGGTTCAGCCGCCAAGGGGGGTCTGA

[0613]

AACTCACCCCGGGGCTGTGGACCTCGGCCTACTTGCGCAAG
TTCGTGCCCGAGGACTACCATCCCTTCGAGATCAGGTTCTAC
GAGGACCAATCCCATCCGCCCAAGGCCGAGCTGTCGGCCTG
CGTCATCACCCAGGGGGCGATCCTGGCCCAATTGCAAGCCAT
CCAGAAATCCCGCCAAGAATTCTTGCTGAAAAAGGGCCGCG
GGGTCTACCTCGACCCCCAGACCGGTGAGGAGCTCAACCCC
GGCTTCCCCCAGGATGCCCCGAGGAAACAAGAAGCTGAAA
GTGGAGCTGCCGCCCGTGGAGGATTTGGAGGAAGACTGGGA
GAACAGCAGTCAGGCAGAGGAGGAGGAGATGGAGGAAGAC
TGGGACAGCACTCAGGCAGAGGAGGACAGCCTGCAAGACA
GTCTGGAGGAAGACGAGGAGGAGGCAGAGGAGGAGGTGG
AAGAAGCAGCCGCCGCCAGACCGTCGTCCTCGGCGGGGGA
GAAAGCAAGCAGCACGGATACCATCTCCGCTCCGGGTCGGG
GTCCCGCTCGACCACACAGTAGATGGGACGAGACCGGACGA
TTCCCGAACCCCACCACCCAGACCGGTAAGAAGGAGCGGCA
GGGATACAAGTCCTGGCGGGGGCACAAAAACGCCATCGTCT
CCTGCTTGCAGGCCTGCGGGGGCAACATCTCCTTCACCCGG
CGCTACCTGCTCTTCCACCGCGGGGTGAACTTTCCCCGCAAC
ATCTTGCATTACTACCGTCACCTCCACAGCCCCTACTACTTCC
AAGAAGAGGCAGCAGCAGCAGAAAAAGACCAGCAGAAAA
CCAGCAGCTAGAAAATCCACAGCGGCGGCAGCAGGTGGACT
GAGGATCGCGGCGAACGAGCCGGCGCAAACCCGGGAGCTG
AGGAACCGGATCTTTCCCACCCTCTATGCCATCTTCCAGCAG
AGTCGGGGGCAGGAGCAGGAACTGAAAGTCAAGAACCGTT
CTCTGCGCTCGCTCACCCGCAGTTGTCTGTATCACAAGAGCG
AAGACCAACTTCAGCGCACTCTCGAGGACGCCGAGGCTCTC
TTCAACAAGTACTGCGCGCTCACTCTTAAAGAGTAGCCCGC
GCCCGCCCAGTCGCAGAAAAAGGCGGGAATTACGTCACCTG
TGCCCTTCGCCCTAGCCGCCTCCACCCATCATCATGAGCAAA
GAGATTCCCACGCCTTACATGTGGAGCTACCAGCCCCAGATG
GGCCTGGCCGCCGGTGCCGCCCAGGACTACTCCACCCGCAT
GAATTGGCTCAGCGCCGGGCCCGCGATGATCTCACGGGTGA
ATGACATCCGCGCCCACCGAAACCAGATACTCCTAGAACAGT
CAGCGCTCACCGCCACGCCCCGCAATCACCTCAATCCGCGTA
ATTGGCCCGCCGCCCTGGTGTACCAGGAAATTCCCCAGCCCA
CGACCGTACTACTTCCGCGAGACGCCCAGGCCGAAGTCCAG
CTGACTAACTCAGGTGTCCAGCTGGCGGGCGGCGCCACCCT
GTGTCGTCACCGCCCCGCTCAGGGTATAAAGCGGCTGGTGAT
CCGGGGCAGAGGCACACAGCTCAACGACGAGGTGGTGAGC
TCTTCGCTGGGTCTGCGACCTGACGGAGTCTTCCAACTCGCC
GGATCGGGGAGATCTTCCTTCACGCCTCGTCAGGCCGTCCTG
ACTTTGGAGAGTTCGTCCTCGCAGCCCCGCTCGGGTGGCATC
GGCACTCTCCAGTTCGTGGAGGAGTTCACTCCCTCGGTCTAC

[0614]

TTCAACCCCTTCTCCGGCTCCCCCGGCCACTACCCGGACGAG
TTCATCCCGAACTTCGACGCCATCAGCGAGTCGGTGGACGG
CTACGATTGAAACTAATCACCCCCTTATCCAGTGAAATAAAG
ATCATATTGATGATGATTTTACAGAAATAAAAAATAATCATTT
GATTTGAAATAAAGATACAATCATATTGATGATTTGAGTTTAA
CAAAAAAATAAAGAATCACTTACTTGAAATCTGATACCAGGT
CTCTGTCCATGTTTTCTGCCAACACCACTTCACTCCCCTCTTC
CCAGCTCTGGTACTGCAGGCCCCGGCGGGCTGCAAACTTCC
TCCACACGCTGAAGGGGATGTCAAATTCCTCCTGTCCCTCAA
TCTTCATTTTATCTTCTATCAGATGTCCAAAAAGCGCGTCCGG
GTGGATGATGACTTCGACCCCGTCTACCCCTACGATGCAGAC
AACGCACCGACCGTGCCCTTCATCAACCCCCCCTTCGTCTCT
TCAGATGGATTCCAAGAGAAGCCCCTGGGGGTGTTGTCCCT
GCGACTGGCCGACCCCGTCACCACCAAGAACGGGGAAATCA
CCCTCAAGCTGGGAGAGGGGGTGGACCTCGATTCCTCGGGA
AAACTCATCTCCAACACGGCCACCAAGGCCGCCGCCCCTCT
CAGTTTTTCCAACAACACCATTTCCCTTAACATGGATCACCC
CTTTTACACTAAAGATGGAAAATTATCCTTACAAGTTTCTCCA
CCATTAAATATACTGAGAACAAGCATTCTAAACACACTAGCT
TTAGGTTTTGGATCAGGTTTAGGACTCCGTGGCTCTGCCTTG
GCAGTACAGTTAGTCTCTCCACTTACATTTGATACTGATGGAA
ACATAAAGCTTACCTTAGACAGAGGTTTGCATGTTACAACAG
GAGATGCAATTGAAAGCAACATAAGCTGGGCTAAAGGTTTA
AAATTTGAAGATGGAGCCATAGCAACCAACATTGGAAATGG
GTTAGAGTTTGGAAGCAGTAGTACAGAAACAGGTGTTGATG
ATGCTTACCCAATCCAAGTTAAACTTGGATCTGGCCTTAGCTT
TGACAGTACAGGAGCCATAATGGCTGGTAACAAAGAAGACG
ATAAACTCACTTTGTGGACAACACCTGATCCATCACCAAACT
GTCAAATACTCGCAGAAAATGATGCAAAACTAACACTTTGCT
TGACTAAATGTGGTAGTCAAATACTGGCCACTGTGTCAGTCT
TAGTTGTAGGAAGTGGAAACCTAAACCCCATTACTGGCACCG
TAAGCAGTGCTCAGGTGTTTCTACGTTTTGATGCAAACGGTG
TTCTTTTAACAGAACATTCTACACTAAAAAAATACTGGGGGT
ATAGGCAGGGAGATAGCATAGATGGCACTCCATATACCAATG
CTGTAGGATTCATGCCCAATTTAAAAGCTTATCCAAAGTCAC
AAAGTTCTACTACTAAAAATAATATAGTAGGGCAAGTATACAT
GAATGGAGATGTTTCAAAACCTATGCTTCTCACTATAACCCTC
AATGGTACTGATGACAGCAACAGTACATATTCAATGTCATTTT
CATACACCTGGACTAATGGAAGCTATGTTGGAGCAACATTTG
GGGCTAACTCTTATACCTTCTCATACATCGCCCAAGAATGAA
CACTGTATCCCACCCTGCATGCCAACCCTTCCCACCCCACTC
TGTGGAACAAACTCTGAAACACAAAATAAAATAAAGTTCAA
GTGTTTTATTGATTCAACAGTTTTACAGGATTCGAGCAGTTAT

[0615]

TTTTCCTCCACCCTCCCAGGACATGGAATACACCACCCTCTC
CCCCCGCACAGCCTTGAACATCTGAATGCCATTGGTGATGGA
CATGCTTTTGGTCTCCACGTTCCACACAGTTTCAGAGCGAGC
CAGTCTCGGGTCGGTCAGGGAGATGAAACCCTCCGGGCACT
CCCGCATCTGCACCTCACAGCTCAACAGCTGAGGATTGTCCT
CGGTGGTCGGGATCACGGTTATCTGGAAGAAGCAGAAGAGC
GGCGGTGGGAATCATAGTCCGCGAACGGGATCGGCCGGTGG
TGTCGCATCAGGCCCCGCAGCAGTCGCTGCCGCCGCCGCTC
CGTCAAGCTGCTGCTCAGGGGGTCCGGGTCCAGGGACTCCC
TCAGCATGATGCCCACGGCCCTCAGCATCAGTCGTCTGGTGC
GGCGGGCGCAGCAGCGCATGCGGATCTCGCTCAGGTCGCTG
CAGTACGTGCAACACAGAACCACCAGGTTGTTCAACAGTCC
ATAGTTCAACACGCTCCAGCCGAAACTCATCGCGGGAAGGA
TGCTACCCACGTGGCCGTCGTACCAGATCCTCAGGTAAATCA
AGTGGTGCCCCCTCCAGAACACGCTGCCCACGTACATGATCT
CCTTGGGCATGTGGCGGTTCACCACCTCCCGGTACCACATCA
CCCTCTGGTTGAACATGCAGCCCCGGATGATCCTGCGGAACC
ACAGGGCCAGCACCGCCCCGCCCGCCATGCAGCGAAGAGA
CCCCGGGTCCCGGCAATGGCAATGGAGGACCCACCGCTCGT
ACCCGTGGATCATCTGGGAGCTGAACAAGTCTATGTTGGCAC
AGCACAGGCATATGCTCATGCATCTCTTCAGCACTCTCAACT
CCTCGGGGGTCAAAACCATATCCCAGGGCACGGGGAACTCT
TGCAGGACAGCGAACCCCGCAGAACAGGGCAATCCTCGCAC
AGAACTTACATTGTGCATGGACAGGGTATCGCAATCAGGCAG
CACCGGGTGATCCTCCACCAGAGAAGCGCGGGTCTCGGTCT
CCTCACAGCGTGGTAAGGGGGCCGGCCGATACGGGTGATGG
CGGGACGCGGCTGATCGTGTTCGCGACCGTGTCATGATGCA
GTTGCTTTCGGACATTTTCGTACTTGCTGTAGCAGAACCTGG
TCCGGGCGCTGCACACCGATCGCCGGCGGCGGTCTCGGCGC
TTGGAACGCTCGGTGTTGAAATTGTAAAACAGCCACTCTCTC
AGACCGTGCAGCAGATCTAGGGCCTCAGGAGTGATGAAGAT
CCCATCATGCCTGATGGCTCTGATCACATCGACCACCGTGGA
ATGGGCCAGACCCAGCCAGATGATGCAATTTTGTTGGGTTTC
GGTGACGGCGGGGGAGGGAAGAACAGGAAGAACCATGATT
AACTTTTAATCCAAACGGTCTCGGAGTACTTCAAAATGAAGA
TCGCGGAGATGGCACCTCTCGCCCCCGCTGTGTTGGTGGAA
AATAACAGCCAGGTCAAAGGTGATACGGTTCTCGAGATGTTC
CACGGTGGCTTCCAGCAAAGCCTCCACGCGCACATCCAGAA
ACAAGACAATAGCGAAAGCGGGAGGGTTCTCTAATTCCTCA
ATCATCATGTTACACTCCTGCACCATCCCCAGATAATTTTCAT
TTTTCCAGCCTTGAATGATTCGAACTAGTTCcTGAGGTAAATC
CAAGCCAGCCATGATAAAGAGCTCGCGCAGAGCGCCCTCCA
CCGGCATTCTTAAGCACACCCTCATAATTCCAAGATATTCTGC

[0616]

TCCTGGTTCACCTGCAGCAGATTGACAAGCGGAATATCAAAA
TCTCTGCCGCGATCCCTGAGCTCCTCCCTCAGCAATAACTGT
AAGTACTCTTTCATATCCTCTCCGAAATTTTTAGCCATAGGAC
CACCAGGAATAAGATTAGGGCAAGCCACAGTACAGATAAAC
CGAAGTCCTCCCCAGTGAGCATTGCCAAATGCAAGACTGCTA
TAAGCATGCTGGCTAGACCCGGTGATATCTTCCAGATAACTG
GACAGAAAATCGCCCAGGCAATTTTTAAGAAAATCAACAAA
AGAAAAATCCTCCAGGTGGACGTTTAGAGCCTCGGGAACAA
CGATGAAGTAAATGCAAGCGGTGCGTTCCAGCATGGTTAGTT
AGCTGATCTGTAGAAAAAACAAAAATGAACATTAAACCATG
CTAGCCTGGCGAACAGGTGGGTAAATCGTTCTCTCCAGCACC
AGGCAGGCCACGGGGTCTCCGGCGCGACCCTCGTAAAAATT
GTCGCTATGATTGAAAACCATCACAGAGAGACGTTCCCGGT
GGCCGGCGTGAATGATTCGACAAGATGAATACACCCCCGGA
ACATTGGCGTCCGCGAGTGAAAAAAAGCGCCCGAGGAAGC
AATAAGGCACTACAATGCTCAGTCTCAAGTCCAGCAAAGCG
ATGCCATGCGGATGAAGCACAAAATTCTCAGGTGCGTACAA
AATGTAATTACTCCCCTCCTGCACAGGCAGCAAAGCCCCCGA
TCCCTCCAGGTACACATACAAAGCCTCAGCGTCCATAGCTTA
CCGAGCAGCAGCACACAACAGGCGCAAGAGTCAGAGAAAG
GCTGAGCTCTAACCTGTCCACCCGCTCTCTGCTCAATATATAG
CCCAGATCTACACTGACGTAAAGGCCAAAGTCTAAAAATAC
CCGCCAAATAATCACACACGCCCAGCACACGCCCAGAAACC
GGTGACACACTCAAAAAAATACGCGCACTTCCTCAAACGCC
CAAAACTGCCGTCATTTCCGGGTTCCCACGCTACGTCATCAA
AACACGACTTTCAAATTCCGTCGACCGTTAAAAACGTCACC
CGCCCCGCCCCTAACGGTCGCCCGTCTCTCAGCCAATCAGCG
CCCCGCATCCCCAAATTCAAACACCTCATTTGCATATTAACGC
GCACAAAAAGTTTGAGGTATATTATTGATGATGgTTAATTAAcg
GCCCGGGC

[0617] 参考文献

[0618] 1.Desrichard,A,Snyder,A&Chan、T.A.Cancer Neoantigens and APPlications for Immunotherapy.Clin.Cancer Res.Off.J.Am.Assoc.Cancer Res.(2015)doi:10.1158/1078-0432.CCR-14-3175

[0619] 2.Schumacher,T.N.&Schreiber,R.D.Neoantigens in cancer immunotherapy.Science 348,69-74(2015).

[0620] 3.Gubin,M.M.,Artyomov,M.N.,Mardis,E.R.&Schreiber,R.D.Tumor neoantigens:building a framework for personalized cancer immunotherapy.J.Clin.Invest.125,3413-3421(2015).

[0621] 4.Rizvi,N.A.et al.Cancer immunology.Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer.Science 348,124-128(2015).

[0622] 5.Snyder,A.et al.Genetic basis for clinical response to CTLA-4

blockade in melanoma.N.Engl.J.Med.371,2189-2199(2014).

[0623] 6.Carreno,B.M.et al.Cancer immunotherapy.A dendritic cell vaccine increases the breadth and diversity of melanoma neoantigen-specific T cells.Science 348,803-808(2015).

[0624] 7.Tran,E.et al.Cancer immunotherapy based on mutation-specific CD4+T cells in a patient with epithelial cancer.Science 344,641-645(2014).

[0625] 8.Hacohen,N.&Wu,C.J.-Y.United States Patent ApPlication:20110293637-COMPOSITIONS AND METHODS OF IDENTIFYING TUMOR SPECIFIC NEOANTIGENS.(A1).at<http://appft1.uspto.gov/netacgi/nph-Parser?Sect1＝PTO1&Sect2＝HITOFF&d＝PG01&p＝1&u＝/netahtml/PTO/srchnum.html&r＝1&f＝G&1＝50&s1＝20110293637.PGNR.>

[0626] 9.Lundegaard,C.,Hoof,I.,Lund,O.&Nielsen,M.State of the art and challenges in sequence based T-cell epitope prediction.Immunome Res.6 Suppl 2,S3(2010).

[0627] 10.Yadav,M.et.al.Predicting immunogenic tumour mutations by combining mass spectrometry and exome sequencing.Nature 515,572-576(2014).

[0628] 11.Bassani-Sternberg,M.,Pletscher-Frankild,S.,Jensen,L.J.&Mann,M.Mass spectrometry of human leukocyte antigen class I peptidomes reveals strong effects of protein abundance and turnover on antigen presentation.Mol.CellProteomics MCP 14,658-673(2015).

[0629] 12.Van Allen,E.M.et al.Genomic correlates of response to CTLA-4 blockade in metastatic melanoma.Science 350,207-211(2015).

[0630] 13.Yoshida,K&Ogawa,S.Splicing factor mutations and cancer.Wiley Interdiscip.Rev.RNA 5,445-459(2014).

[0631] 14.Cancer Genome Atlas Research Network.Comprehensive molecular profiling of lung adenocarcinoma.Nature 511,543-550(2014).

[0632] 15.Rajasagi,M.et al.Systematic identification of personal tumor-specific neoantigens in chronic lymphocytic leukemia.Blood 124,453-462(2014).

[0633] 16.Downing,S.R.et al.United States Patent APPlication:0120208706-OPTIMIZATION OF MULTIGENE ANALYSIS OF TUMOR SAMPLES.(A1).at<http://appft1uspto.gov/netacgi/nph-Parser?Sect1＝PTO1&Sect2＝HITOFF&d＝PG01&p＝1&u＝/netahtml/PTO/srchnum.html&r＝1&f＝G&l＝50&s1＝20120208706PGNR>

[0634] 17.Target Capture for NextGen Sequencing-IDT.at<http://www.idtdna.com/pages/products/nextgen/target-capture>

[0635] 18.Shukla,S.A.et al.Comprehensive analysis of cancer-associated somatic mutations in class I HLA genes.Nat.Biotechnol.33,1152-1158(2015).

[0636] 19.Cieslik,M.et al.The use of exome capture RNA-seq for highly degraded RNA with application to clinical cancer sequencing.Genome Res.25,1372-1381(2015).

[0637] 20.Bodini,M.et al.The hidden genomic landscape of acute myeloid

leukemia:subclonal structure revealed by undetected mutations.Blood 125,600-605(2015).

[0638] 21.Saunders,C.T.et al.Strelka:accurate somatic small-variant calling from sequenced tumor-normal sample pairs.Bioinforma.Oxf.Engl.28,1811-1817(2012).

[0639] 22.Cibulskis,K.et al.Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples.Nat.Biotechnol.31,213-219(2013).

[0640] 23.Wilkerson,M.D.et al.Integrated RNA and DNA sequencing improves mutation detection in low purity tumors.Nucleic Acids Res.42,e107(2014).

[0641] 24.Mose,L.E.,Wilkerson,M.D.,Hayes,D.N.,Perou,C.M.&Parker,J.S.ABRA:improved coding indel detection via assembly-based realignment.Bioinforma.Oxf.Engl.30,2813-2815(2014).

[0642] 25.Ye,K.,Schulz,M.H.,Long,Q.,Apweiler,R.&Ning,Z.Pindel:a pattern growth approach to detect break points of large deletions and medium sized insertions from paired-end short reads.Bioinforma.Oxf.Engl.25,2865-2871(2009).

[0643] 26.Lam,H.Y.K.et al.Nucleotide-resolution analysis of structural variants using BreakSeq and a breakpointlibrary.Nat.Biotechnol.28,47-55(2010).

[0644] 27.Frampton,G.M.et al.Development and validation of a clinical cancer genomic profiling test based on massively parallel DNA sequencing.Nat.Biotechnol.31,1023-1031(2013).

[0645] 28.Boegel,S.et al.HLA tyPing from RNA-Seq sequence reads.Genome Med.4,102(2012).

[0646] 29.Liu,C.et al.ATHLATES:accurate typing of human leukocyte antigen through exome sequencing.Nucleic Acids Res.41,e142(2013).

[0647] 30.Mayor,N.P.et al.HLA Typing for the Next Generation.PloS One 10,e0127153(2015).

[0648] 31.Roy,C.K.,Olson,S.,Graveley,B.R.,Zamore,P.D.&Moore,M.J.Assessing long-distance RNA sequence connectivity via RNA-templated DNA-DNA ligation.eLife 4,(2015).

[0649] 32.Song,L.&Florea,L.CLASS:constrained transcript assembly of RNA-seq reads.BMC Bioinformatics 14 SuPPl 5,S14(2013).

[0650] 33.Maretty,L,Sibbesen,J.A.&Krogh,A.Bayesian transcriptome assembly.Genome Biol.15,501(2014).

[0651] 34.Pertea,M.et al.StringTie enables improved reconstruction of a transcriptome from RNA-seq reads.Nat.Biotechnol.33,290-295(2015).

[0652] 35.Roberts,A.,Pimentel,H.,Trapnell,C.&Pachter,L.Identification of novel transcripts in annotated genomes using RNA-Seq.Bioinforma.Oxf.Engl.

(2011).doi:10.1093/bioinformatics/btr355

[0653] 36.Vitting-Seerup,K.,Porse,B.T.,Sandelin,A.&Waage,J.spliceR:an R package for classification of alternative splicingand prediction of coding potential from RNA-seq data.BMC Bioinformatics 15,81(2014).

[0654] 37.Rivas,M.A.et al.Human genomics.Effect of predicted protein-truncating genetic variants on the human transcriptome.Science.348,666-669(2015).

[0655] 38.Skelly,D.A.,Johansson,M.,Madeoy,J.,Wakefield,J.&Akey,J.M.Apowerful and flexible statistical framework for testing hypotheses of allele-specific gene expression from RNA-seq data.Genome Res.21,1728-1737(2011).

[0656] 39.Anders,S.,Pyl,P.T.&Huber,W.HTSeq--a Python framework to work with high-throughput sequencing data.Bioinforma.Oxf.Engl.31,166-169(2015).

[0657] 40.Furney,S.J.et al.SF3B1 mutations are associated with alternative sDlicing in uveal melanoma.Cancer Discov.(2013).doi:10.1158/2159-8290.CD-13-0330

[0658] 41.Zhou,Q.et al.A chemical genetics approach for the functional assessment of novel cancer genes.Cancer Res.(2015).doi:10.11158/0008-5472.CAN-14-2930

[0659] 42.Maguire,S.L.et al.SF3B1 mutatiohs constitute a novel therapeutic target in breast cancer.J.Pathol.235,571-580(2015).

[0660] 43.Carithers,L.J.et.al.A Novel Approach to High-Quality Postmortem Tissue Procurement:The GTEx Project.Biopreservation Biobanking 13,311-319(2015).

[0661] 44.Xu,G.et.al.RNA CoMPASS:a dual approach for pathogen and host transcriptome analysis of RNA-seq datasets.PloS One 9,e89445(2014).

[0662] 45.Andreatta,M.&Nielsen,M.Gapped sequence alignment using artificial neural networks:application to the MHC class I system.Bioinforma.Oxf.Engl.(2015).doi:10.1093/bioinformatics/btv639

[0663] 46.Jørgensen,K.W.,Rasmussen,M.,Buus,S.&Nielsen,M.NetMHCstab-predicting stability of peptide-MHC-1 complexes;impacts for cytotoxic T lymphocyte epitope discovery.Immunology 141,18-26(2014).

[0664] 47.Larsen,M.V.et al.An integrative approach to CTL epitope prediction:a combined algorithm integrating MHC class I binding,TAP transport efficiency,and proteasomal cleavage predictions.Eur.J.Immunol.35,2295-2303(2005).

[0665] 48.Nielsen,M.,Lundegaard,C.,Lund,O.&Keşmir,C.The role of the proteasome in generating cytotoxic T-cell epitopes:insights obtained from improved predictions of proteasomal cleavage.Immunogenetics 57,33-41(2005).

[0666] 49.Boisvert,F.-M.et al.A Quantitative Spatial Proteomics Analysis of

Proteome Turnover in Human Cells.Mol.Cell.Proteomics 11,M111.011429-M111.011429(2012).

[0667] 50.Duan,F.et al.Genomic and bioinformatic profiling of mutational neoepitopes reveals new rules to predict anticancer immunogenicity.J.Exp.Med.211,2231-2248(2014).

[0668] 51.Janeway's Immunobiology:9780815345312:Medicine&Health Science Books@Amazon.com.at<http://www.amazon.com/Janeways-Immunobiology-Kenneth-Murphy/dp/0815345313>

[0669] 52.Calis,J.J.A.et al.Properties of MHC Class I Presented Peptides That Enhance Immunogenicity.PLoS Comput.Biol.9,e1003266(2013).

[0670] 53.Zhang,J.et al.Intratumor heterogeneity in localized lung adenocarcinomas delineated by multiregion sequencing.Science 346,256-259(2014)

[0671] 54.Walter,M.J.et al.Clonal architecture of secondary acute myeloid leukemia.N.Engl.J.Med.366,1090-1098(2012).

[0672] 55.Hunt DF,Henderson RA,Shabanowitz J,Sakaguchi K,Michel H,Sevilir N,Cox AL,Appella E,Engelhard VH.Charactehzation of peptides bound to the class I MHC molecule HLA-A2.1 by mass spectrometry.Science 1992.255:1261-1263.

[0673] 56.Zarling AL,Polefrone JM,Evans AM,Mikesh LM,Shabanowitz J,Lewis ST,Engelhard VH,Hunt DF.Identification of class I MHC-associated phosphopeptides as targets for cancer immunotherapy.Proc Natl Acad Sci U S A.2006 Oct 3;103(40):14889-94.

[0674] 57.Bassani-Sternberg M,P1etscher-Frankild S,Jensen LJ,Mann M.Mass spectrometry of human leukocyte antigen class I peptidomes reveals strong effects of protein abundance and tumover on antigen presentation.Mol Cell Proteomics.2015 Mar;14(3):658-73.doi:10.1074/mcp.M114.042812.

[0675] 58.Abelin JG,Trantham PD,Penny SA,Patterson AM,Ward ST,Hildebrand WH,Cobbold M,Bai DL,Shabanowitz J,Hunt DF.Complementary IMAC enrichment methods for HLA-associated phosphopeptide identification by mass spectrometry.Nat Protoc.2015 Sep;10(9):1308-18.doi:10.1038/nprot.2015.086.Epub 2015 Aug 6

[0676] 59.Barnstable CJ,Bodmer WF,Brown G,Galfre G,Milstein C,Williams AF,Ziegler A.Production of monoclonal antibodies to group A erythrocytes,HLA and other human cell surface antigens-new tools for genetic analysis.Cell.1978 May;14(1):9-20.

[0677] 60.Goldman JM,Hibbin J,Keamey L,Orchard K,Th'ng KH.HLA-DR monoclonal antibodies inhibit the proliferation of normal and chronic granulocytic leukaemia myeloid prohenitor cells.Br I Haematol.1982 Nov;52(3):411-20.

[0678] 61.Eng JK,Jahan TA,Hoopmann MR.Comet:an open-source MS/MS sequence database search tool.Proteomics.2013 Jan;13(1):22-4.doi:10.1002/

pmic.201200439.Epub 2012 Dec 4.

[0679] 62.Eng JK,Hoopmann MR,Jahan TA,Egertson JD,Noble WS,MacCoss MJ.A deeper look into Comet--implementation and features.J Am Soc Mass Spectrom.2015 Nov;26(11):1865-74.doi:10.1007/s13361-015-1179-x.Epub 2015 Jun 27.

[0680] 63.Lukas Käll,Jesse Canterbury,Jason Weston,William Stafford Noble and Michael J.MacCoss.Semi-supervised learning for peptide identification from shotgun proteomics datasets.Nature Methods 4:923-925,November 2007

[0681] 64.Lukas Käll,John D.Storey,Michael J.MacCoss and William Stafford Noble.Assigning confidence measures to peptides identified by tandem mass spectrometry.Journal of Proteome Research,7(1):29-34,January 2008

[0682] 65.Lukas Käll,John D.Storey and William Stafford Noble.Nonparametric estimation of posterior error probabilities associated with peptides identified by tandem mass spectrometry.Bioinformatics,24(16):i42-i48,August 2008

[0683] 66.Kinney RM,BJ Johnson,VL Brown,DW Trent.Nucleotide Sequence of the 26 S mRNA of the Virulent Trinidad Donkey Strain of Venezuelan Equine Encephalitis Virus and Deduced Sequence of the Encoded Stmctural Proteins.Virology 152(2),400-413.1986 Jul 30.

[0684] 67.Jill E Slansky,Frédéhque M Rattis,Lisa F Boyd,Tarek Fahmy,Elizabeth M Jaffee,Jonathan P Schneck,David H Margulies,Drew M Pardoll.Enhanced Antigen-Specific Antitumor Immunity with Altered Peptide Ligands that Stabilize the MHC-Peptide-TCR Complex.Immunity,Volume 13,Issue 4,1 October 2000,Pages 529-538.

[0685] 68.A Y Huang,P H Gulden,A S Woods,M C Thomas,C D Tong,W Wang,V H Engelhard,G Pastemack,R Cotter,D Hunt,D M Pardoll,and E M Jaffee.The immunodominant major histocompatibility complex class I-restricted antigen of a murine colon tumor derives from an endogenous retroviral gene product Proe Natl Acad Sci U S A.;93(18):9730-9735,1996 Sep 3.

[0686] 69.JOHNSON,BARBARA J.B.,RICHARD M.KINNEY,CRYSTLE L.KOST AND DENNIS W.TRENT.Molecular Determinants of Alphavirus Neurovimlence:Nucleotide and Deduced Protein Sequence Changes during Attenuation of Venezuelan Equine Encephalitis Virus.J Gen Virol 67:1951-1960,1986.

[0687] 70.Aarnoudse,C.A.,Krüse,M.,Konopitzky,R.,Brouwenstijn,N.,and Schrier,P.I.(2002).TCR reconstitution in Jurkat reporter cells facilitates the identification of novel tumor antigens by cDNA expression cloning.Int J Cancer 99,7-13.

[0688] 71.Alexander,J.,Sidney,J.,Southwood,S.,Ruppert,J.,Oseroff,C.,Maewal,A.,Snoke,K.,Serra,H.M.,Kubo,R.T.,and Sette,A.(1994).Development of high

potency universal DR-restricted helper epitopes by modification of high affinity DR-blocking peptides Immunity 1,751-761.

[0689] 72.Banu,N.,Chia,A.,Ho,Z.Z.,Garcia,A.T.,Paravasivam,K.,Grotenbreg,G.M.,Bertoletti,A.,and Gehring,A.J.(2014).Building and optimizing a virus-specific T cell receptor library for targeted immunotherapy in viral infections.Scientific Reports 4,4166.

[0690] 73.Cornet,S.,Miconnet,I.,Menez,J.,Lemonnier,F.,and Kosmatopoulos,K.(2006).Optimal organization of a polypeptide-based candidate cancer vaccine composed of cryptic tumor peptides with enhanced immunogenicity.Vaccine 24,2102-2109.

[0691] 74.Depla,E.,van der Aa,A.,Livingston,B.D.,Crimi,C.,Allosery,K.,de Brabandere,V.,Krakover,J.,Murthy,S.,Huang,M.,Power,S.,et al.(2008).Rational design of a multiepitope vaccine encoding T-lymphocyte epitopes for treatment of chronic hepatitis B virus infections.Journal of Virology 82,435-450.

[0692] 75.Ishioka,G.Y.,Fikes,J.,Hermanson,G.,Livingston,B.,Crimi,C.,Qin,M.,del Guercio,M.F.,Oseroff,C.,Dahlberg,C.,Alexander,J.,et al.(1999).Utilization of MHC class I transgenic mice for development of minigene DNA vaccines encoding multiple HLA-restricted CTL epitopes.J Immunol 162,3915-3925.

[0693] 76.Janetzki,S.,Price,L.,Schroeder,H.,Britten,C.M.,Welters,M.J.P.,and Hoos,A.(2015).Guidelines for the automated evaluation of.Elispot assays.Nat Protoc 10,1098-1115.

[0694] 77.Lyons,G.E.,Moore,T.,Brasic,N.,Li,M.,Roszkowski,J.J.,and Nishimura,M.I.(2006).Influence of human CD8 on antigen recognition by T-cell receptor-transduced cells.Cancer Res 66,11455-11461.

[0695] 78.Nagai,K.,Ochi,T.,Fujiwara,H.,An,J.,Shirakata,T.,Mineno,J.,Kuzushima,K.,Shiku,H.,Melenhorst,J.J.,Gostick,E.,et al(2012).Aurora kinase A-specific T-cell receptor gene transfer redirects T lymphocytes to display effective antileukemia reactivity.Blood 119,368-376.

[0696] 79.Panina-Bordignon,P.,Tan,A.,Termijtelen,A.,Demotz,S.,Corradin,G.,and Lanzavecchia,A.(1989).Universally immunogenic T cell epitopes:promiscuous binding to human MHC class II and promiscuous recognition by T cells.Eur J Immunol 19,2237-2242.

[0697] 80.Vitiello,A.,Marchesini,D.,Furze,J.,Sherman,L.A.,and Chesnut,R.W.(1991).Analysis of the HLA-restricted influenza-specific cytotoxic T lymphocyte response in transgenic mice carrying a chimeric human-mouse class I major histocompatibility complex.J Exp Med 173,1007-1015.

[0698] 81.Yachi,P.P.,Ampudia,J.,Zal,T.,and Gascoigne,N.R.J.(2006).Altered peptide ligands induce delayed CD8-T cell receptor interaction--a role for CD8 in distinguishing antigen quality.Immunity 25,203-211.

[0699] 82.Pushko P,Parker M,Ludwig GV,Davis NL,Johnston RE,Smith JF.Replicon-helper systems from attenuated Venezuelan equine encephalitis virus:expression of heterologous genes in vitro and immunization against heterologous pathogens in vivo.Virology.1997 Dec 22;239(2):389-401.

[0700] 83.Strauss,JH and E G Strauss.The alphaviruses:gene expression, replication,and evolution.Microbiol Rev.1994 Sep;58(3):491-562.

[0701] 84.Rhême C,Ehrengruber MU,Grandgirard D.Alphaviral cytotoxicity and its implication in vector development.Exp Physiol.2005 Jan;90(1):45-52.Epub 2004 Nov 12.

[0702] 85.Riley,Michael K.II,and Wilfred Vermarris.Recent Advances in Nanomaterials for Gene Delivery-A Review.Nanomaterials 2017,7(5),94.

[0703] 86.Frolov I,Hardy R,Rice CM.Cis-acting RNA elements at the 5′end of Sindbis virus eenome RNA regulate minus-and plus-strand RNA synthesis.RNA.2001 Nov;7(11):1638-51.

[0704] 87.Jose J,Snyder JE,Kuhn RJ.A structural and functional perspective of alphavirus replication and assembly.Future Microbiol.2009 Sep;4(7):837-56.

[0705] 88.Bo Li and C.olin N.Dewey.RSEM:accurate transcript quantification from RNA-Seq data with or without a referenfe genome.BMC Bioinformatics,12: 323,August 2011

[0706] 89.Hillary Pearson,Tariq Daouda,Diana Paol a Granados,Chantal Durette,Eric Bonneil,Mathieu Courcelles,Anja Rodenbrock,Jean-Philippe Laverdure,Caroline Côté,Sylvie Mader,Sébastien Lemieux,Pierre Thibault,and Claude Perreault.MHC class I-associated peptides derive from selective regions of the human genome.The Journal of Clinical Investigation,2016,

[0707] 90.Juliane Liepe,Fabio Marino,John Sidney,Anita Jeko,Daniel E.Bunting,Alessandro Sette,Peter M.Kloetzel,Michael P.H.Stumpf,Albert J.R.Heck,Michele Mishto.A large fraction of HLA class I ligands are proteasome-generated spliced peptides.Science,21,October 2016.

[0708] 91.Mommen GP.,Marino,F.,Meiring HD.,Poelen,MC.,van Gaans-van den Brink,JA.,Mohammed S.,Heck AJ.,and van Els CA.Sampling From the Proteome to the Human Leukocyte Antigen-DR(HLA-DR)Ligandome Proceeds Via High Specificity.Mol Cell Proteomics 15(4):1412-1423,April 2016.

[0709] 92.Sebastian Kreiter,Mathias Vormehr,Niels van de Roemer,Mustafa Diken,Martin Löwer,Jan Diekmann,Sebastian Boegel,Barbara Schrörs,Fulvia Vascotto,John C.Castle,Arbel D.Tadmor,Stephen P.Schoenberger,Christoph Huber, Özlem Türeci,and Ugur Sahin.Mutant MHC class II epitopes drive therapeutic immune responses to caner.Nature 520,692-696,April 2015.

[0710] 93.Tran E.,Turcotte S.,Gros A.,Robbins P.F.,Lu Y.C.,Dudley M.E., Wunderlich J.R.,Somerville R.P.,Hogan K.,Hinrichs C.S.,Parkhurst M.R.,Yang

J.C.,Rosenberg S.A.Cancer immunotherapy based on mutation-specific CD4+T cells in a patient with epithelial cancer.Science 344(6184)641-645,May 2014.

[0711] 94.Andreatta M.,Karosiene E.,Rasmussen M.,Stryhn A.,Buus S.,Nielsen M.Accurate pan-specific prediction of peptide-MHC class II binding affinity with improved binding core identification.Immunogenetics 67(11-12)641-650,November 2015.

[0712] 95.Nielsen,M.,Lund,O.NN-align.An artificial neural network-based alignment algorithm for MHC class II peptide binding prediction.BMC Bioinformatics 10:296,September 2009.

[0713] 96.Nielsen,M.,Lundegaard,C.,Lund,O.Prediction of MHC class II binding afrinity using SMM-align,a novel stabilization matrix alignment method.BMC Bioinformatics 8:238,July 2007.

[0714] 97.Zhang,J.,et al.PEAKS DB:de novo sequencing assisted database search for sensitive and accurate peptide identification.Molecular&Cellular Proteomics.11(4):1-8.1/2/2012.

[0715] 98.Jensen,Kamilla Kjaergaard,et al.“Improved Methods for Prediting Peptide Binding Affinity to MHC Class II Molecules.”Immunology,2018,doi:10.1111/imm.12889.

[0716] 99.Carter,S.L.,Cibulskis,K.,Helman,E.,McKenna,A.,Shen,H.,Zack,T.,Laird,P.W.,Onofrio,R.C.,Winckler,W.,Weir,B.A.,et al.(2012).Absolute quantification of somatic DNA alteratiohs in human cancer.Nat.Biotechnol.30,413-421

[0717] 100.McGranahan,N.,Rosenthal,R.,Hiley,C.T.,Rowan,A.J.,Watkins,T.B.K.,Wilson,G.A.,Birkbak,N.J.,Veeriah,S.,Van Loo,P.,Herrero,J.,et al.(2017).Allele-Specific HLA Loss and Immune Escape in Lung Cancer Evolution.Cell 171,1259-1271.e11.

[0718] 101.Shukla,S.A.,Rooney,M.S.,Rajasagi,M.,Tiao,G.,Dixon,P.M.,Lawrence,M.S.,Stevens,J.,Lane,W.J.,Dellagatta,J.L.,Steelman,S.,et al.(2015).Comprehensive ahalysis of cancer-associated somatic mutations in class I HLA genes.Nat.Biotechnol33,1152-1158.

[0719] 102.Van Loo,P.,Nordgard,S.H.,Lingjaerde,O.C.,Russnes,H.G.,Rye,I.H.,Sun,W.,Weigman,V.J.,Marynen,P.,Zetterberg,A.,Naume,B.,et al.(2010).Allele-specific copy number analysis of tumors.Proc.Natl.Acad.Sci.U.S.A.107,16910-16915.

[0720] 103.Van Loo,P.,Nordgard,S.H.,Lingjaerde,O.C.,Russnes,H.G.,Rye,I.H.,Sun,W.,Weigman,V.J.,Marynen,P.,Zetterberg,A.,Naume,B.,et al.(2010).Allele-specific copy number analysis of tumors.Proc.Natl,Acad.Sci.U.S.A.107,16910-16915.

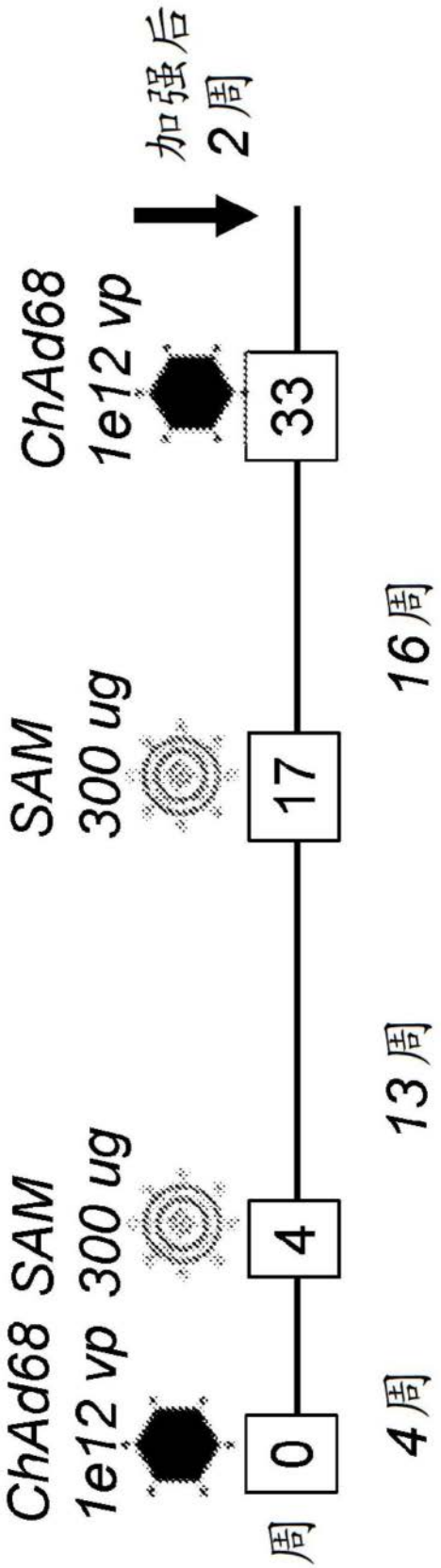

ChAd68
1e12 vp
SAM
300 ug
ChAd68
1e12 vp
加强后
2周
周
0
4
17
33
4周
13周
16周

图1

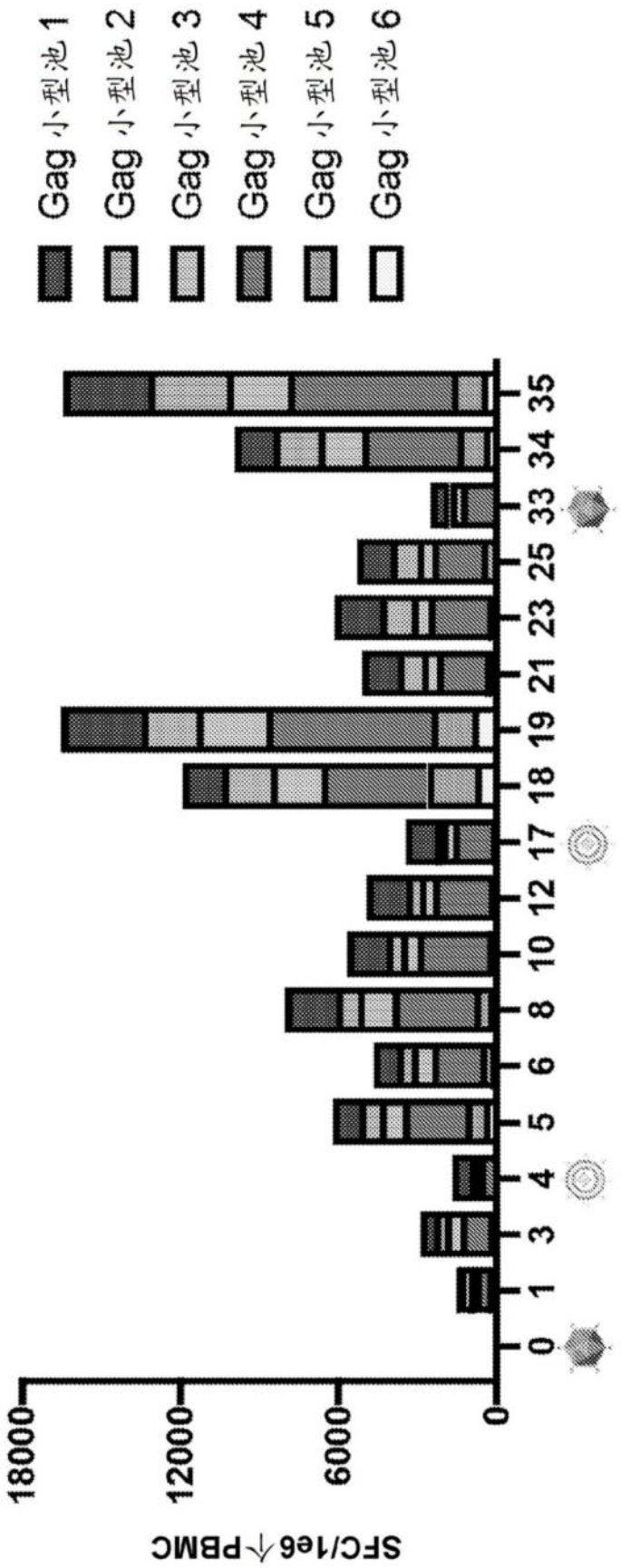

Gag小型池1
Gag小型池2
Gag小型池3
Gag小型池4
Gag小型池5
Gag小型池6
SFC/1e6↓PBMC
18000
12000
6000
0
0
1
3
4
5
6
8
10
12
17
18
19
21
23
25
33
34
35

图2A

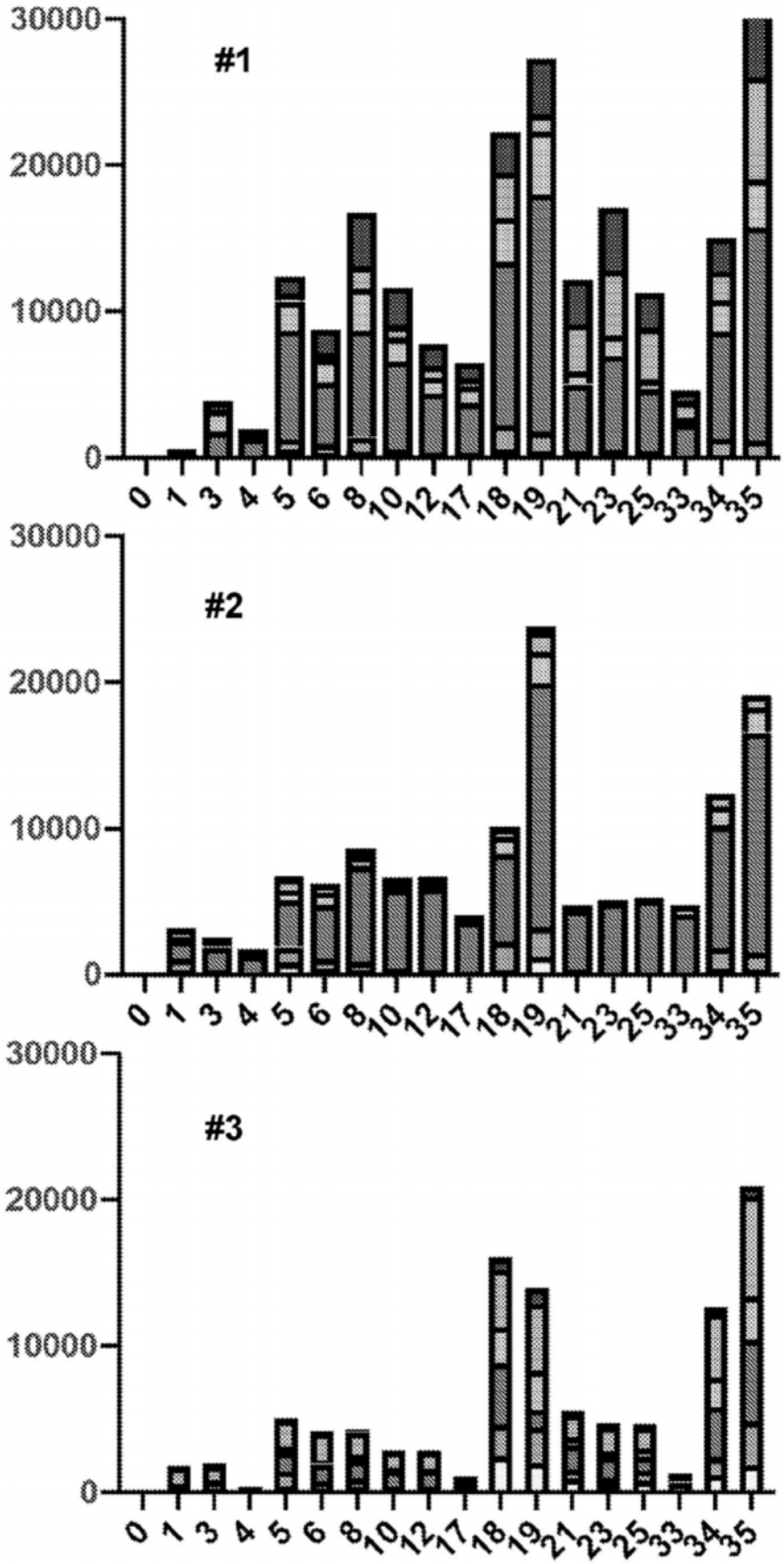

30000
20000
10000
0
#1
0 1 3 4 5 6 8 10 12 17 18 19 21 23 25 33 34 35
30000
20000
10000
0
#2
0 1 3 4 5 6 8 10 12 17 18 19 21 23 25 33 34 35
30000
20000
10000
0
#3
0 1 3 4 5 6 8 10 12 17 18 19 21 23 25 33 34 35

图2B

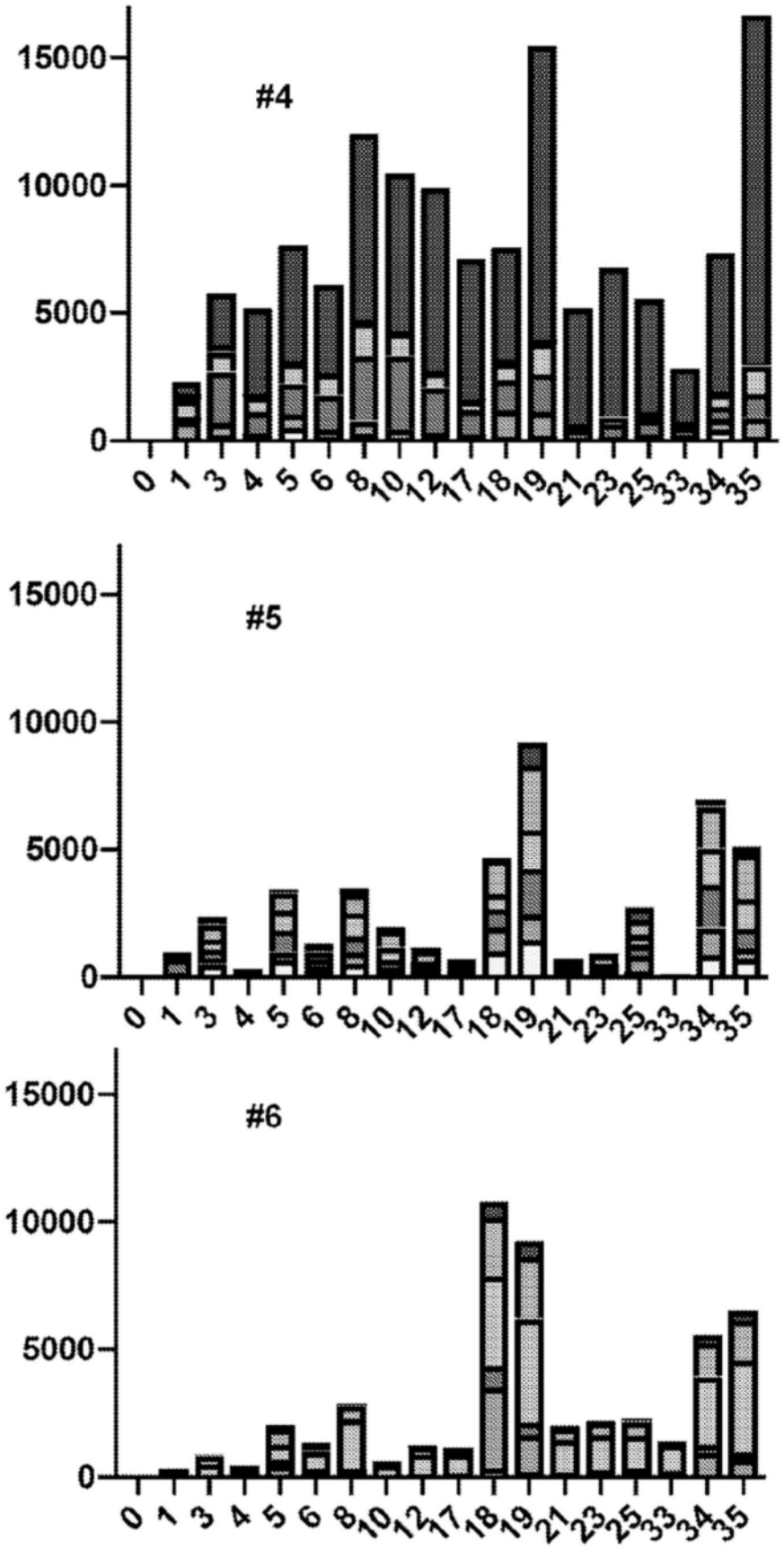

#4
15000
10000
5000
0
0 1 3 4 5 6 8 10 12 17 18 19 21 23 25 33 34 35
#5
15000
10000
5000
0
0 1 3 4 5 6 8 10 12 17 18 19 21 23 25 33 34 35
#6
15000
10000
5000
0
0 1 3 4 5 6 8 10 12 17 18 19 21 23 25 33 34 35

图2C

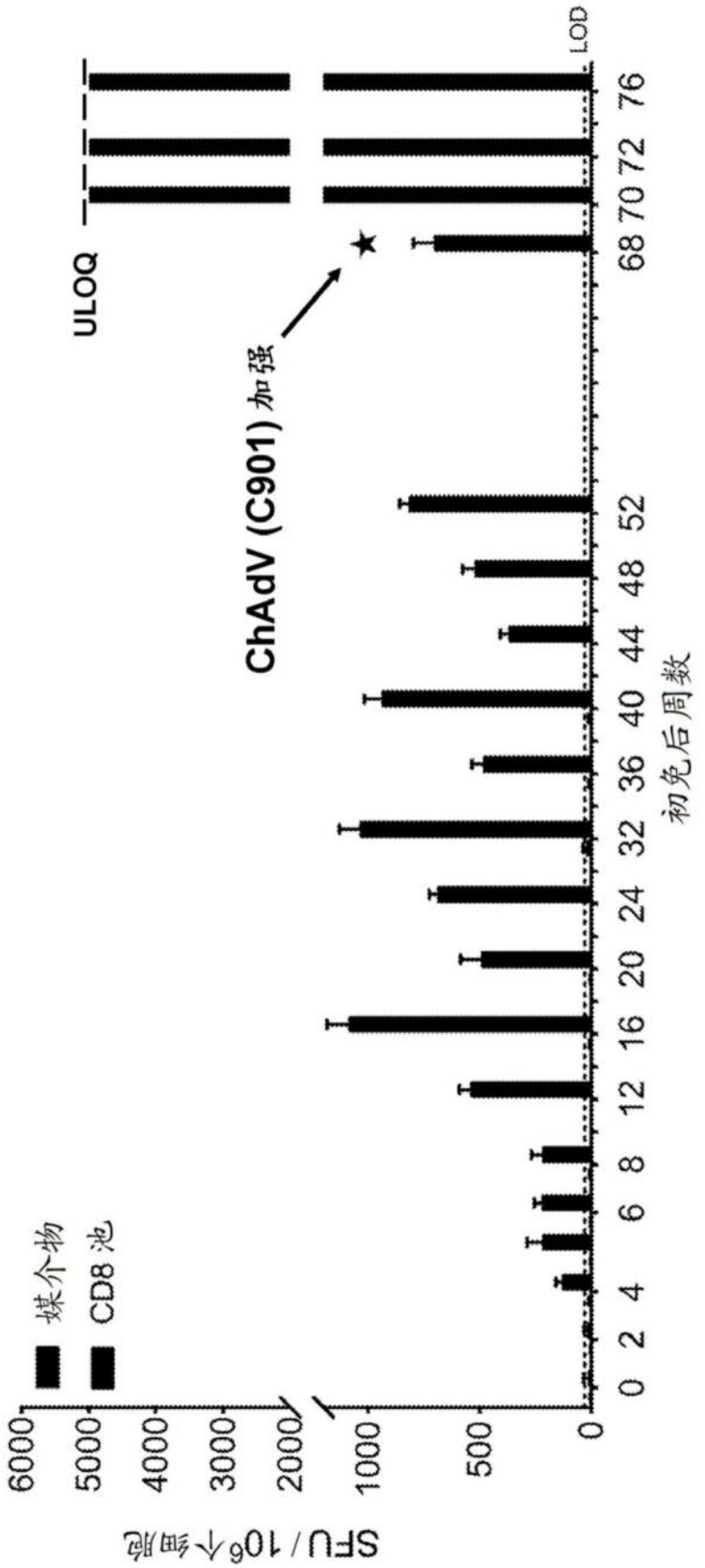

媒介物
CD8 池
SFU / 10⁶个细胞
6000
5000
4000
3000
2000
1000
500
0
0 2 4 6 8 12 16 20 24 32 36 40 44 48 52 68 70 72 76
初免后周数
ChAdV (C901) 加强
ULOQ
LOD

图3A

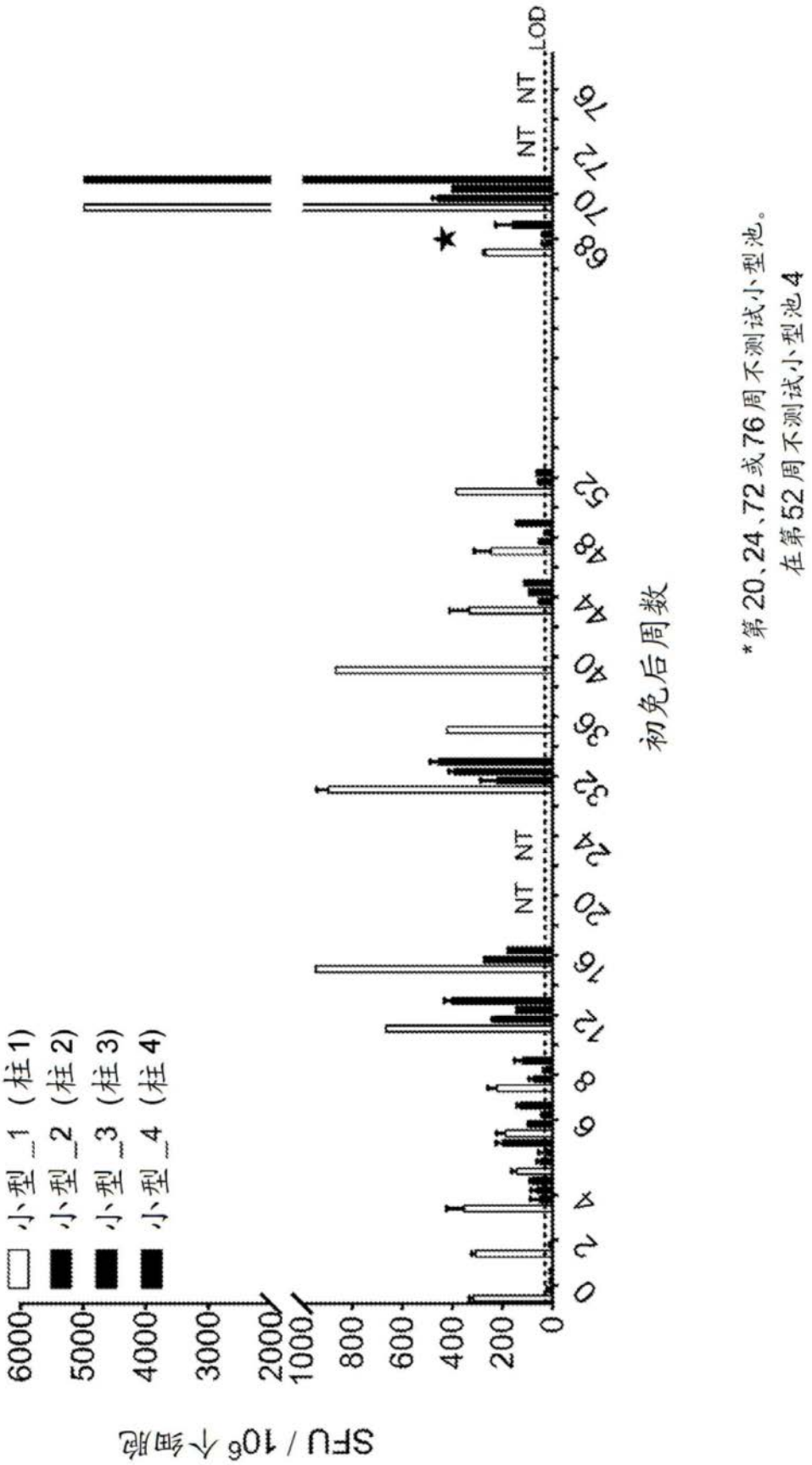

小型_1（柱1）
小型_2（柱2）
小型_3（柱3）
小型_4（柱4）
SFU / 10⁶个细胞
初免后周数
NT
NT
LOD
*第20、24、72或76周不测试小型池。
在第52周不测试小型池4

图3B

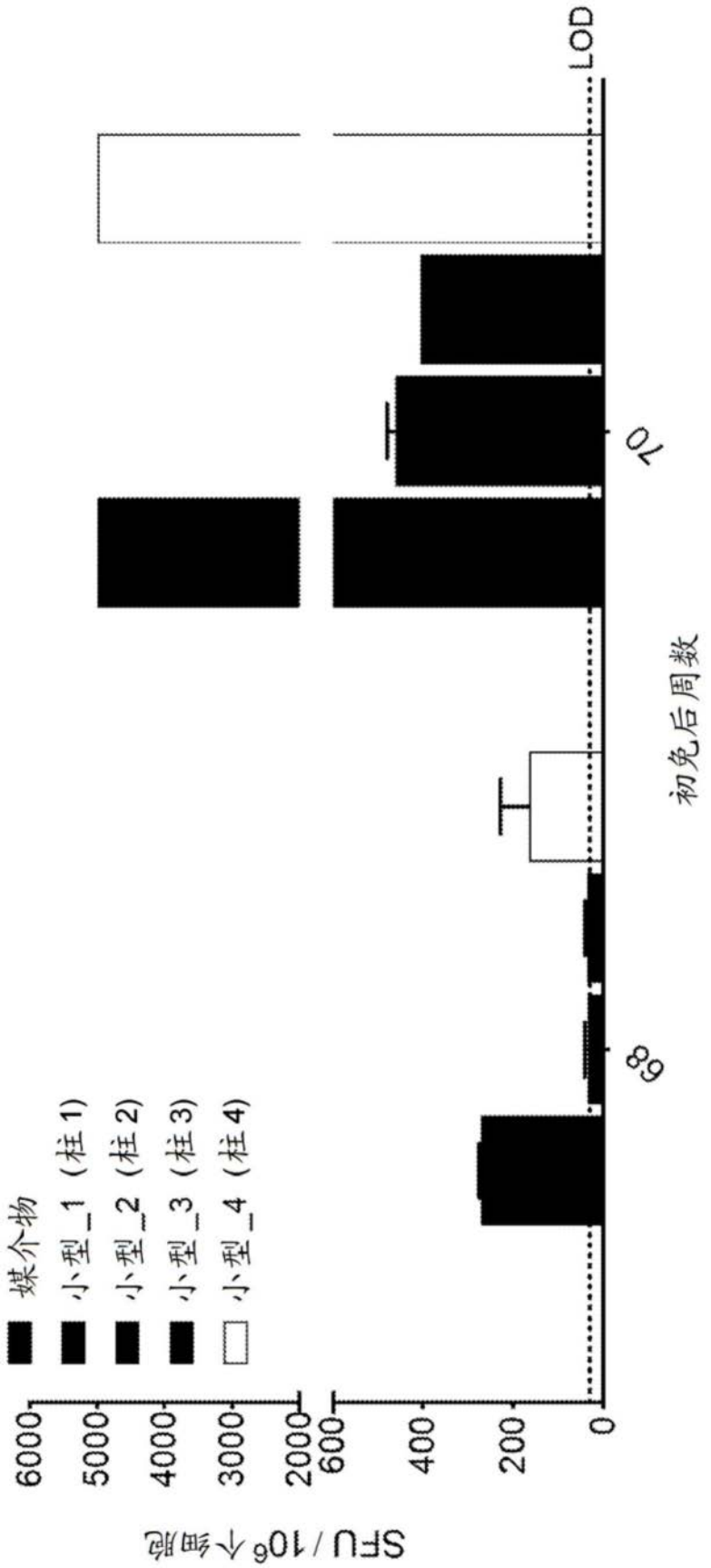

SFU / 10⁶ 个细胞
6000
5000
4000
3000
2000
600
400
200
0
媒介物
小型_1 (柱 1)
小型_2 (柱 2)
小型_3 (柱 3)
小型_4 (柱 4)
8
9
10
LOD
初免后周数

图3C

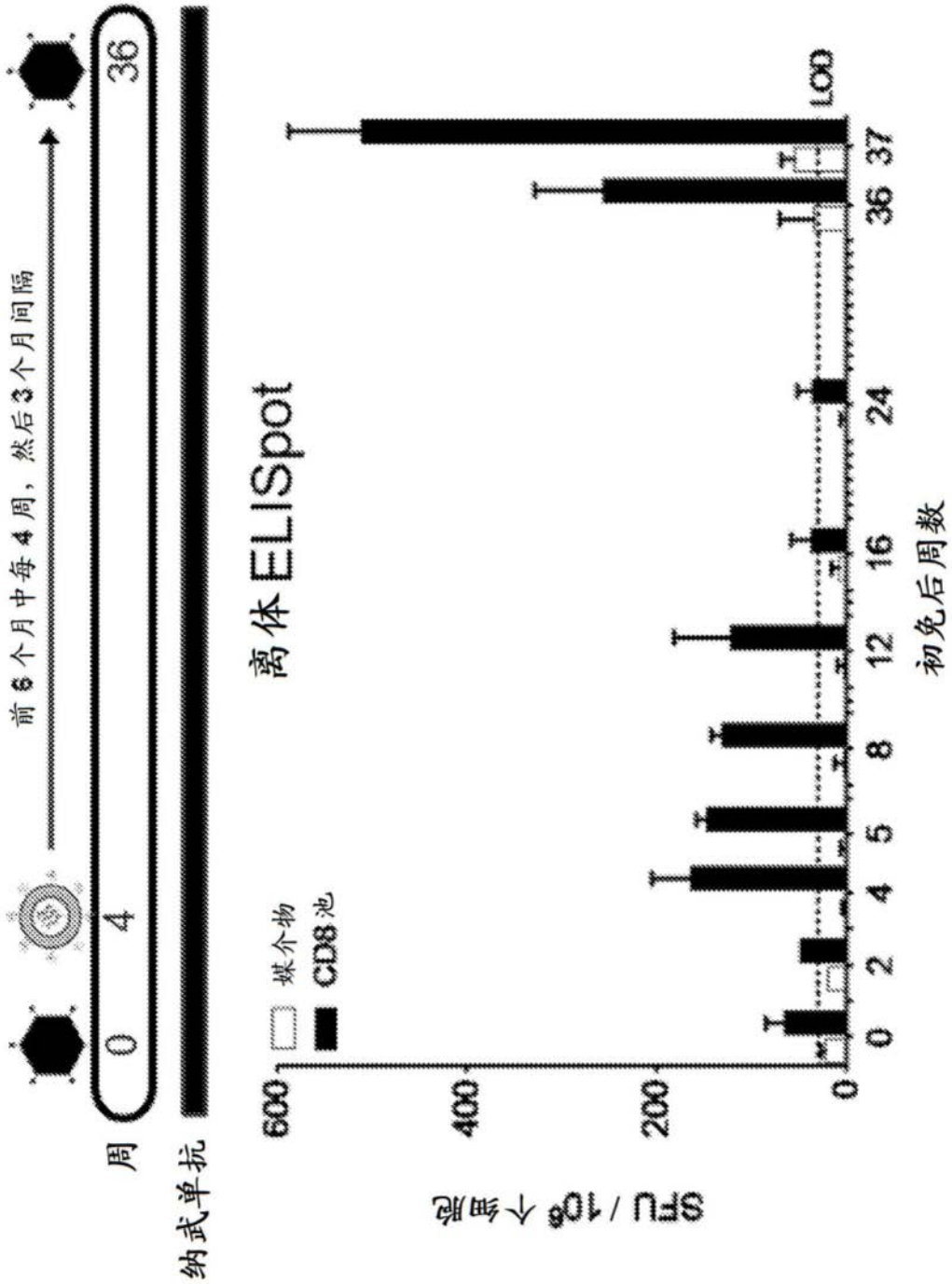

周
0
4
36
前6个月中每4周，然后3个月间隔
纳武单抗
离体ELISpot
媒介物
CD8池
SFU / 10⁶个细胞
600
400
200
0
0
2
4
5
8
12
16
24
36
37
LOD
初免后周数

图4

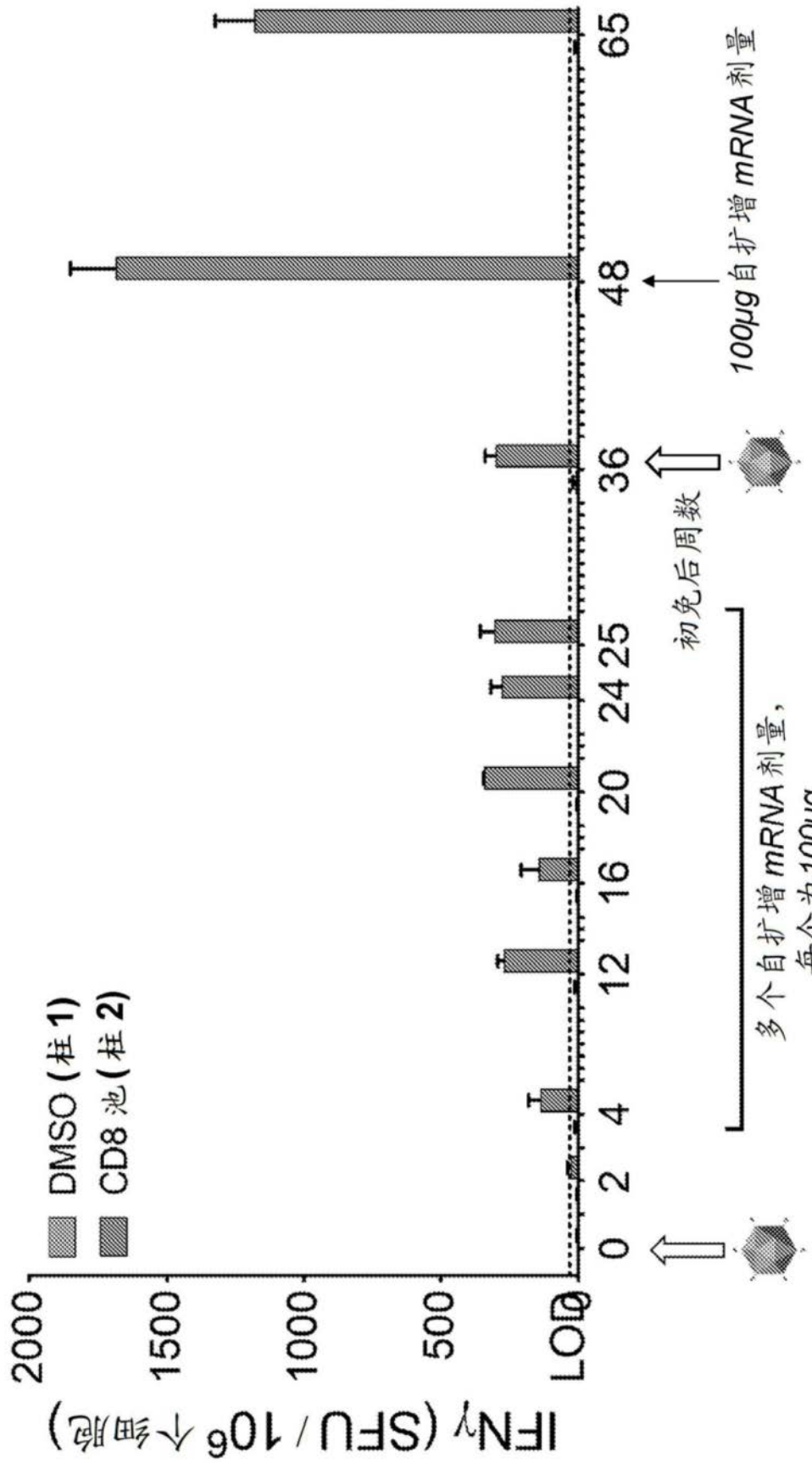

IFNγ (SFU / 10⁶ 个细胞)
2000
1500
1000
500
LOD
DMSO (柱 1)
CD8 池 (柱 2)
0
2
4
12
16
20
24
25
36
48
65
初免后周数
100μg 自扩增 mRNA 剂量
多个自扩增 mRNA 剂量，
每个为 100μg

图5

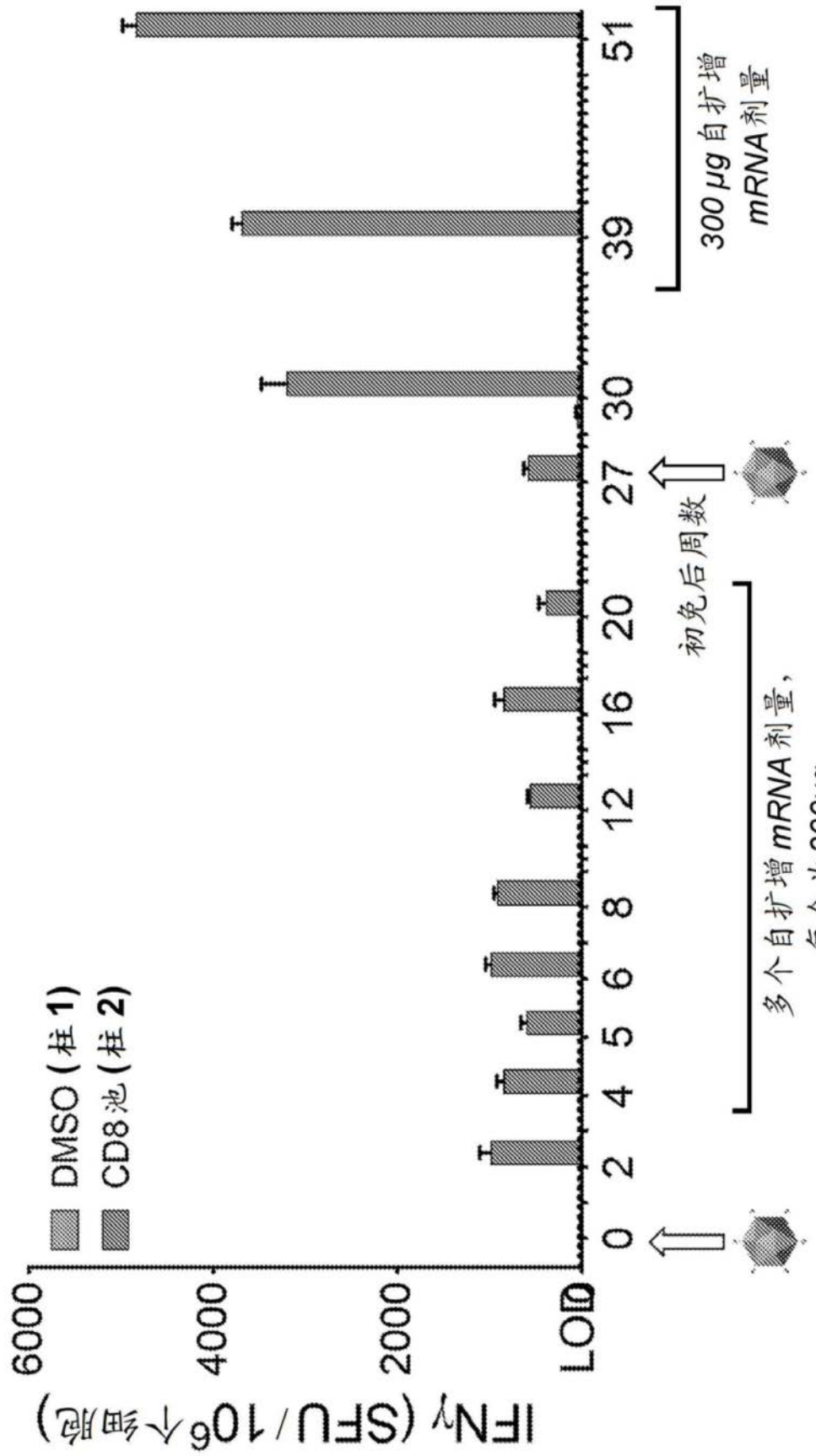

IFNγ (SFU/10⁶个细胞)
6000
4000
2000
LOD
DMSO (柱1)
CD8池 (柱2)
0
2
4
5
6
8
12
16
20
27
30
39
51
初免后周数
多个自扩增mRNA剂量，
每个为300μg
300 μg 自扩增
mRNA剂量

图6

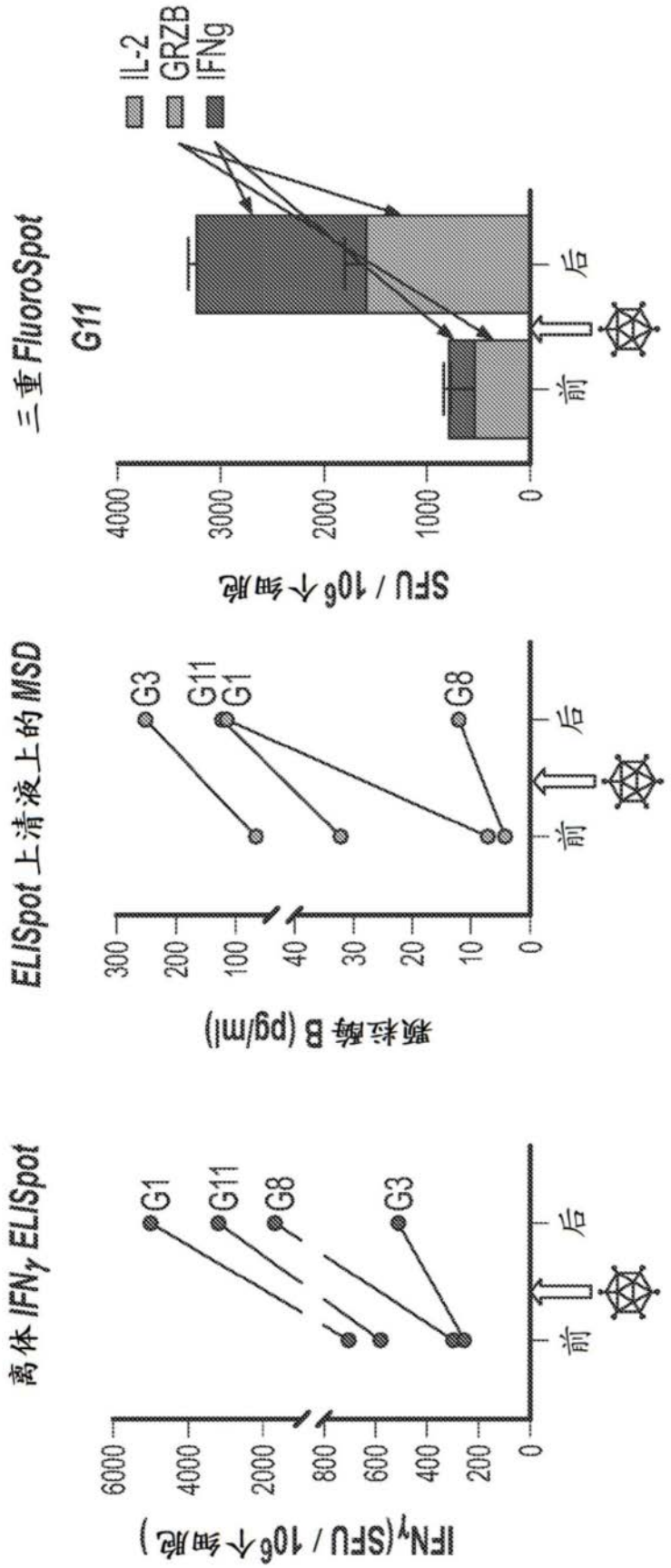

三重 FluoroSpot
G11
IL-2
GRZB
IFNg
SFU / 10⁶个细胞
4000
3000
2000
1000
0
前
后
ELISpot 上清液上的 MSD
颗粒酶 B (pg/ml)
G3
G11
G1
G8
300
200
100
40
30
20
10
0
离体 IFNγ ELISpot
IFNγ(SFU / 10⁶个细胞)
G1
G11
G8
G3
6000
4000
2000
800
600
400
200
0

图7

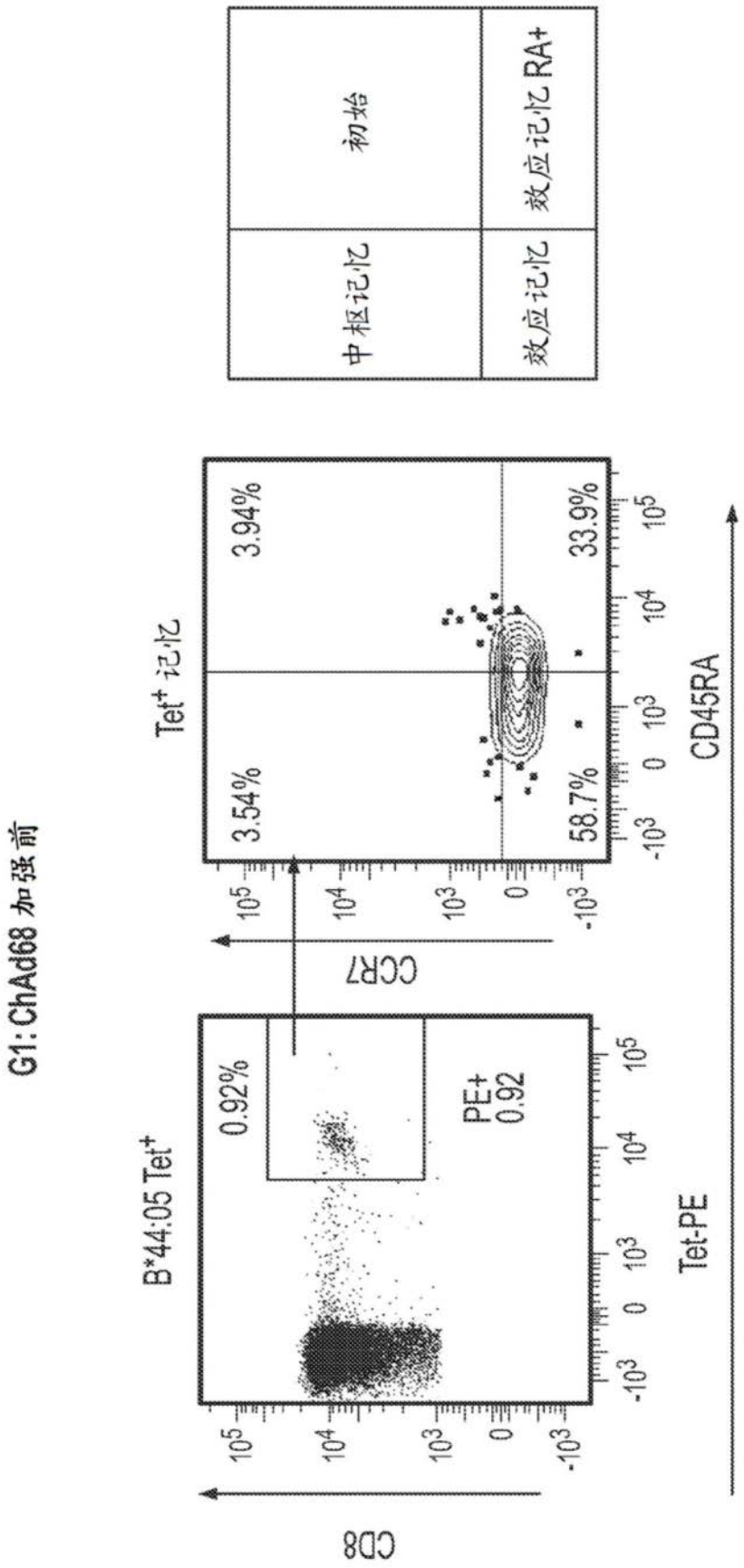

G1: ChAd68 加强前
B*44:05 Tet+
0.92%
PE+
0.92
CD8
Tet-PE
Tet+ 记忆
3.54%
3.94%
58.7%
33.9%
CCR7
CD45RA
中枢记忆
初始
效应记忆
效应记忆 RA+

图8A

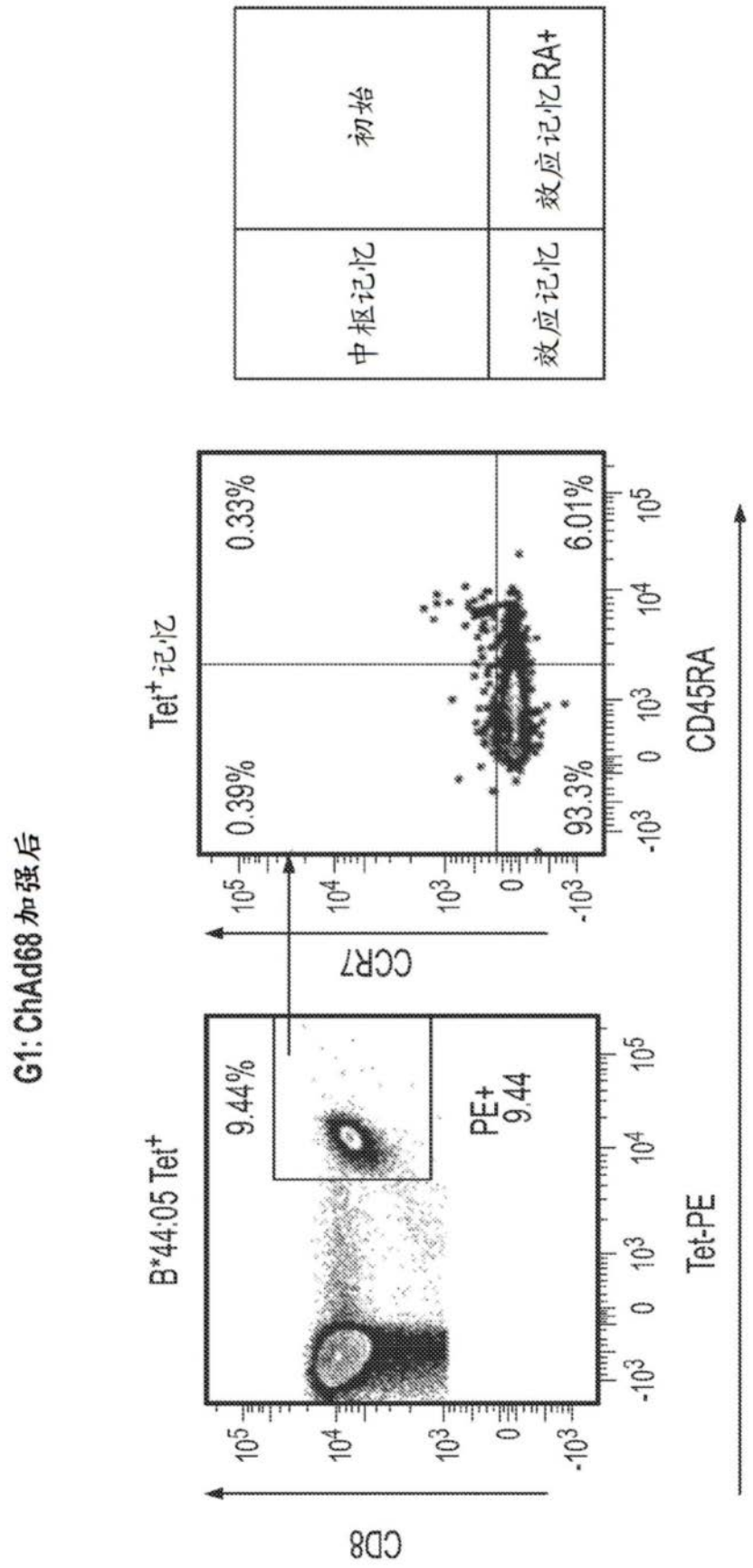
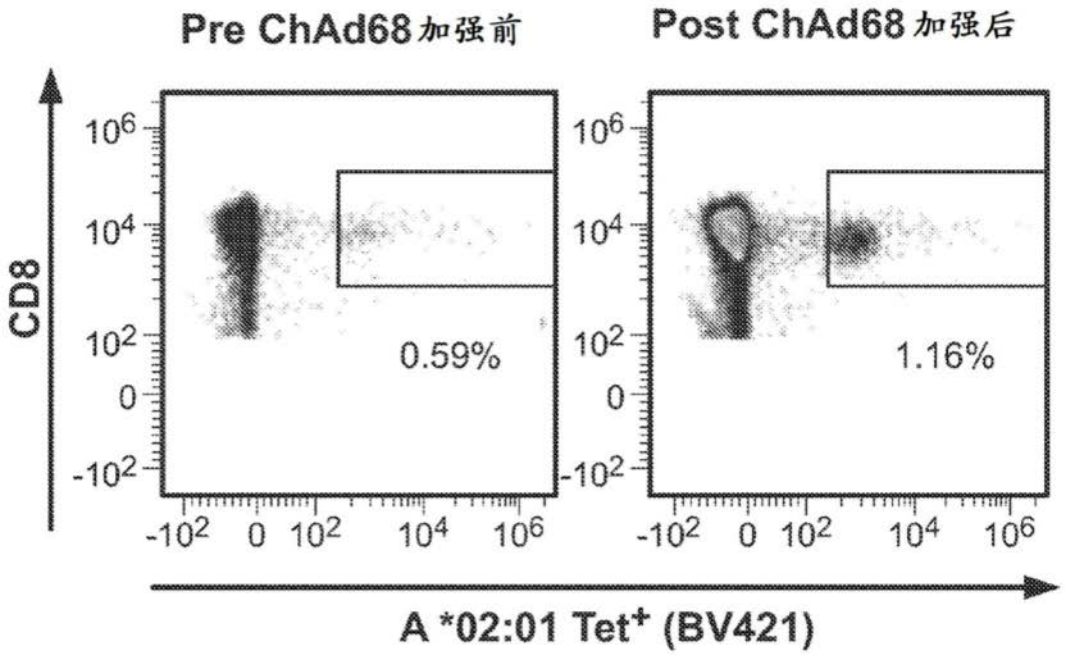

G1: ChAd68加强后
B*44:05 Tet+
9.44%
PE+
9.44
CD8
Tet-PE
Tet+记忆
0.39%
0.33%
93.3%
6.01%
CCR7
CD45RA
中枢记忆
初始
效应记忆
效应记忆RA+

图8B


Pre ChAd68加强前
Post ChAd68加强后
CD8
0.59%
1.16%
A *02:01 Tet+ (BV421)

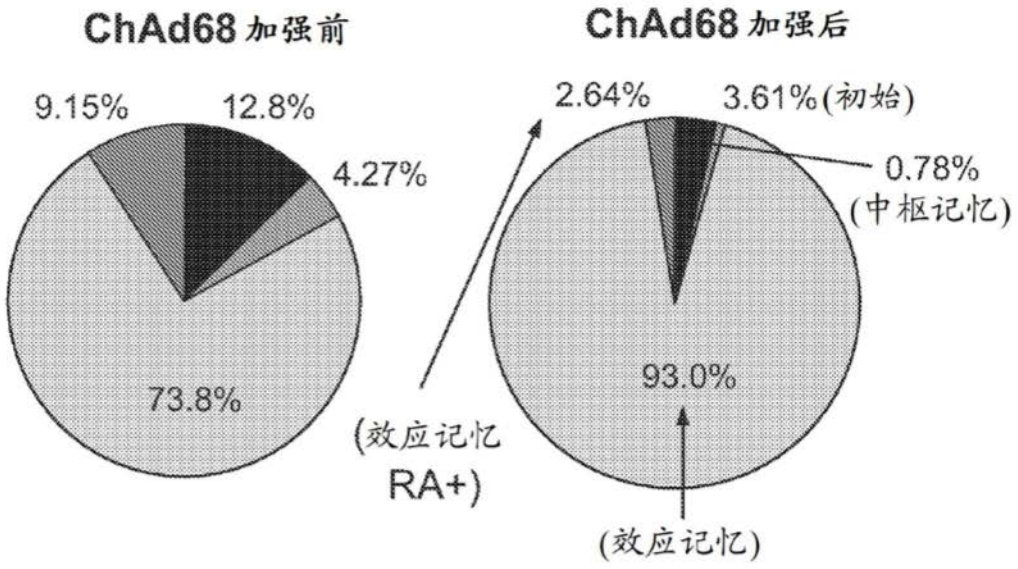

ChAd68加强前
ChAd68加强后
9.15%
12.8%
4.27%
73.8%
2.64%
3.61%(初始)
0.78%
(中枢记忆)
93.0%
(效应记忆
RA+)
(效应记忆)

图8C