(19)中华人民共和国国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 113498439 A
(43)申请公布日 2021.10.12

(21)申请号 201980092591.0
(22)申请日 2019.12.20
(30)优先权数据
62/783133 2018.12.20 US
62/815335 2019.03.07 US
62/815845 2019.03.08 US
(85)PCT国际申请进入国家阶段日
2021.08.19
(86)PCT国际申请的申请数据
PCT/US2019/067758 2019.12.20
(87)PCT国际申请的公布数据
WO2020/132396 EN 2020.06.25
(71)申请人 波赛达治疗公司
地址 美国加利福尼亚州
(72)发明人 E·M·奥斯特塔 D·舍德洛克
(74)专利代理机构 中国专利代理(香港)有限公司 72001
代理人 黄登高 黄希贵
(51)Int.Cl.
C12N 15/85(2006.01)

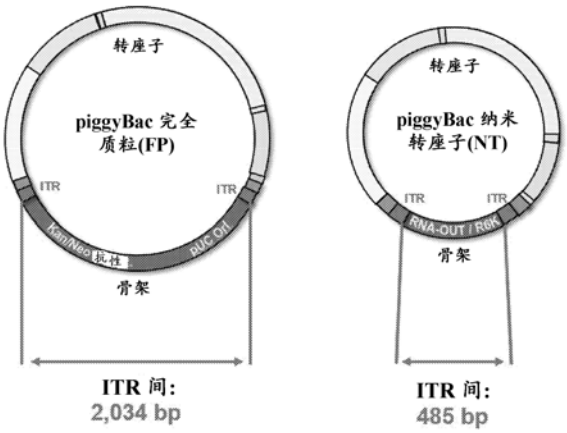
权利要求书3页 说明书61页
序列表112页 附图15页

(54)发明名称

纳米转座子组合物及使用方法

(57)摘要

公开了包含第一核酸序列和第二核酸序列的组合物,所述第一核酸序列包含:(a)第一反向末端重复(ITR),(b)第二ITR和(c)ITR内序列,其中所述ITR内序列包含转座子序列;和所述第二核酸序列包含ITR间序列,其中所述ITR间序列的长度为1至600个核苷酸,包括端点在内。优选地,组合物是纳米转座子。



CN 113498439 A

1. 一种组合物,其包含:

第一核酸序列,其包含:(a)第一反向末端重复(ITR),(b)第二ITR和(c) ITR内序列,其中所述ITR内序列包含转座子序列;和

第二核酸序列,其包含ITR间序列,其中所述ITR间序列的长度为1至600个核苷酸,包括端点在内。

2.权利要求1的组合物,其中所述ITR间序列的长度为1至100个核苷酸,包括端点在内。

3.权利要求1的组合物,其中所述第一核酸序列进一步包含复制起点序列。

4.权利要求1的组合物,其中所述第二核酸序列进一步包含复制起点序列。

5.权利要求3或4的组合物,其中所述复制起点序列的长度为1至450个核苷酸。

6.权利要求5的组合物,其中所述复制起点序列包含R6K复制起点。

7.权利要求1的组合物,其中所述第一核酸进一步包含编码第一可选择标记物的序列。

8.权利要求1的组合物,其中所述第二核酸序列进一步包含编码第一可选择标记物的序列。

9.权利要求7或8的组合物,其中所述第一可选择标记物的长度为1至200个核苷酸。

10.权利要求7或8的组合物,其中所述第一可选择标记物是蔗糖可选择标记物。

11.权利要求7或8的组合物,其中所述蔗糖可选择标记物是RNA-OUT可选择标记物。

12.权利要求1的组合物,其中所述第一核酸序列不包含重组位点、切除位点、连接位点或其组合。

13.权利要求1的组合物,其中所述第二核酸序列不包含重组位点、切除位点、连接位点或其组合。

14.权利要求1的组合物,其中所述第一核酸序列不包含编码外来DNA的序列。

15.权利要求1的组合物,其中所述第二核酸序列不包含编码外来DNA的序列。

16.权利要求1的组合物,其中所述第一核酸序列进一步包含至少一种外源序列和编码能够在哺乳动物细胞中表达外源序列的启动子的序列。

17.权利要求16的组合物,其中所述第一核酸序列进一步包含编码绝缘子的至少一种序列。

18.权利要求16的组合物,其中所述第一核酸序列进一步包含聚腺苷(聚A)序列。

19.权利要求16的组合物,其中编码能够在哺乳动物细胞中表达外源序列的启动子的序列能够在人细胞中表达外源序列。

20.权利要求19的组合物,其中所述启动子是组成型启动子。

21.权利要求19的组合物,其中所述启动子是诱导型启动子。

22.权利要求16的组合物,其中所述至少一种外源序列包含编码非天然存在的抗原受体的序列、编码治疗性多肽的序列或其组合。

23.权利要求22的组合物,其中所述非天然存在的抗原受体包括嵌合抗原受体(CAR)。

24.权利要求23的组合物,其中所述CAR包括:

(a)包含抗原识别区的胞外结构域,

(b)跨膜结构域,和

(c)包含至少一个共刺激结构域的胞内结构域。

25.权利要求24的组合物,其中所述抗原识别区包含至少一种单链可变片段(scFv)、单

结构域抗体、Centyrin或其组合。

26.权利要求25的组合物，其中所述单结构域抗体是VHH或VH。

27. 权利要求24的组合物，其中所述抗原识别区包含至少一种抗BCMA Centyrin。

28. 权利要求27的组合物，其中所述抗BCMA Centyrin包含SEQ ID NO: 29的氨基酸序列。

29. 权利要求24的组合物，其中所述抗原识别区包含至少一种抗BCMA VH。

30. 权利要求29的组合物，其中所述抗BCMA VH包含SEQ ID NO: 97的氨基酸序列。

31. 权利要求24的组合物，其中所述抗原识别区包含至少一种抗PSMA Centyrin。

32. 权利要求31的组合物，其中所述抗PSMA Centyrin包含SEQ ID NO: 94的氨基酸序列。

33.权利要求24的组合物，其中所述胞外结构域进一步包含信号肽。

34.权利要求24的组合物，其中所述CAR进一步包含在抗原识别区和跨膜结构域之间的铰链区。

35.权利要求24的组合物，其中所述跨膜结构域包含编码CD8跨膜结构域的序列。

36.权利要求24的组合物，其中所述至少一个共刺激结构域包括CD3ζ共刺激结构域、4-1BB共刺激结构域或其组合。

37.权利要求24的组合物，其中所述至少一个共刺激结构域包含CD3ζ共刺激结构域和4-1BB共刺激结构域，并且其中所述4-1BB共刺激结构域位于跨膜结构域和CD3ζ共刺激结构域之间。

38.权利要求XX的组合物，其中所述至少一种外源序列包含编码诱导型促凋亡多肽的序列、编码第二可选择标记物的序列、编码嵌合刺激受体(CSR)的序列、编码转座酶的序列、编码自切割肽的序列或其组合。

39.权利要求38的组合物，其中所述第二可选择标记物包含编码二氢叶酸还原酶(DHFR)突变蛋白酶的序列。

40.权利要求1的组合物，其中所述组合物是转座子。

41.权利要求40的组合物，其中所述转座子是piggyBac转座子。

42.一种多核苷酸，其包含编码权利要求1的组合物的核酸序列。

43.一种细胞，其包含权利要求1的组合物。

44.一种细胞群体，其中多个细胞群体被修饰为表达权利要求24的CAR。

45.权利要求44的细胞群体，其中多个修饰的细胞是多个修饰的免疫细胞。

46.权利要求44的细胞群体，其中多个修饰的细胞是多个修饰的T细胞。

47.权利要求44的细胞群体，其中至少50%的多个修饰的T细胞表达包含CD45RA和CD62L的一种或多种细胞表面标记物，并且不表达包含CD45RO的一种或多种细胞表面标记物。

48.一种组合物，其包含权利要求43的细胞。

49.一种组合物，其包含权利要求44的细胞群体。

50.一种药物组合物，其包含权利要求1、48或49中任一项的组合物和药学上可接受的载体。

51.权利要求1、48或49中任一项的组合物或权利要求50的药物组合物，其用于治疗有需要的受试者中的癌症。

52.权利要求51的用途,其中所述癌症是BCMA阳性癌症或PSMA阳性癌症。

53.权利要求51的用途,其中所述癌症是原发性肿瘤、转移癌、多重抗性癌症、进行性肿瘤或复发癌。

54.权利要求51的用途,其中所述癌症是肺癌、脑癌、头颈癌、乳腺癌、皮肤癌、肝癌、胰腺癌、胃癌、结肠癌、直肠癌、子宫癌、宫颈癌、卵巢癌、前列腺癌、睾丸癌、皮肤癌、食管癌、淋巴瘤、白血病、急性白血病、急性成淋巴细胞性白血病(ALL)、急性淋巴细胞性白血病、急性髓样白血病(AML)、急性粒细胞性白血病、慢性髓细胞白血病(CML)、慢性淋巴细胞性白血病(CLL)、多毛细胞白血病、骨髓增生异常综合征(MDS)、何杰金氏病、非何杰金氏淋巴瘤或多发性骨髓瘤。

纳米转座子组合物及使用方法

[0001] 相关申请的交叉引用

本申请要求于2018年12月20日提交的U.S.S.N. 62/783,133;于2019年3月7日提交的U.S.S.N. 62/815,335、以及于2019年3月8日提交的U.S.S.N. 62/815,845的优先权和权益。这些申请各自的内容整体引入本文作为参考。

[0002] 引入作为参考的序列表

于2019年12月19日创建且大小为295 MB,命名为"POTH-047_001WO Seq Listing_ST25.txt"的文件的内容在此整体引入作为参考。

技术领域

[0003] 本公开内容涉及分子生物学,且更具体而言,涉及纳米转座子、包含纳米转座子的细胞组合物、其制备方法和使用方法。

背景技术

[0004] 本领域一直存在关于用于基因疗法中的具有改善转座的组合物和方法的长期但未满足的需要。本公开内容提供了纳米转座子组合物、制备这些组合物的方法和使用这些组合物的方法,所述组合物包括对携带转座子序列的载体的非天然存在的结构改善,以改善转座功效,特别是作为修饰基因疗法细胞的方法用于人细胞中。

发明内容

[0005] 本公开内容提供了组合物,其包含:第一核酸序列和第二核酸序列,所述第一核酸序列包含:(a)第一反向末端重复(ITR),(b)第二ITR和(c) ITR内序列,其中所述ITR内序列包含转座子序列;和所述第二核酸序列包含ITR间序列,其中所述ITR间序列的长度为1至600个核苷酸,包括端点在内。在一个优选的方面,ITR间序列的长度为1至100个核苷酸,包括端点在内。组合物可以为转座子或可以为纳米转座子。在一个优选的方面,转座子是piggyBac转座子。

[0006] 第一核酸序列和/或第二核酸序列可以进一步包含复制起点序列。复制起点序列的长度可以为1至450个核苷酸。复制起点序列可以包含R6K复制起点。

[0007] 第一核酸序列和/或第二核酸序列可以进一步包含编码第一可选择标记物的序列。第一可选择标记物的长度可以为1至200个核苷酸。第一可选择标记物可以为蔗糖可选择标记物。在一个优选的方面,蔗糖可选择标记物是RNA-OUT可选择标记物。

[0008] 第一核酸序列和/或第二核酸序列可以不包含重组位点、切除位点、连接位点或其组合。第一核酸序列和/或第二核酸序列可以不包含编码外来DNA的序列。

[0009] 第一核酸序列可以进一步包含至少一种外源序列和编码能够在哺乳动物细胞中表达外源序列的启动子的序列。第一核酸序列可以进一步包含编码绝缘子的至少一种序列。第一核酸序列可以进一步包含聚腺苷(聚A)序列。编码能够在哺乳动物细胞中表达外源序列的启动子的序列能够在人细胞中表达外源序列。启动子可以为组成型启动子或诱导型

启动子。

[0010] 至少一种外源序列可以包含编码非天然存在的抗原受体的序列、编码治疗性多肽的序列或其组合。在一个优选的方面，非天然存在的抗原受体包括嵌合抗原受体(CAR)。CAR可以包括：(a)包含抗原识别区的胞外结构域，(b)跨膜结构域，和(c)包含至少一个共刺激结构域的胞内结构域。抗原识别区可以包含至少一种单链可变片段(scFv)、单结构域抗体、Centyrin或其组合。单结构域抗体可以为VHH或VH。

[0011] 抗原识别区可以包含至少一种抗BCMA Centyrin。优选地，抗BCMA Centyrin包含SEQ ID NO: 29的氨基酸序列。抗原识别区可以包含至少一种抗BCMA VH。优选地，抗BCMA VH包含SEQ ID NO: 97的氨基酸序列。抗原识别区可以包含至少一种抗PSMA Centyrin。优选地，抗PSMA Centyrin包含SEQ ID NO: 94的氨基酸序列。

[0012] 胞外结构域可以进一步包含信号肽。CAR进一步包含在抗原识别区和跨膜结构域之间的铰链区。跨膜结构域可以包含编码CD8跨膜结构域的序列。至少一个共刺激结构域可以包含CD3ζ共刺激结构域、4-1BB共刺激结构域或其组合。至少一个共刺激结构域可以包括CD3ζ共刺激结构域和4-1BB共刺激结构域，并且其中所述4-1BB共刺激结构域位于跨膜结构域和CD3ζ共刺激结构域之间。至少一种外源序列可以包含编码诱导型促凋亡多肽的序列、编码第二可选择标记物的序列、编码嵌合刺激受体(CSR)的序列、编码转座酶的序列、编码自切割肽的序列或其组合。第二可选择标记物可以包含编码二氢叶酸还原酶(DHFR)突变蛋白酶的序列。

[0013] 本公开内容还提供了包含编码如本文公开的组合物(例如，转座子或纳米转座子)的核酸序列的多核苷酸，和/或包含编码如本文公开的CAR的核酸序列的多核苷酸。

[0014] 本公开内容还提供了包含如本文公开的组合物(例如，转座子或纳米转座子)的细胞。本公开内容还提供了细胞群体，其中多个群体被修饰为表达如本文公开的CAR或组合物(例如，转座子或纳米转座子)。在一个方面，多个修饰的细胞是多个修饰的免疫细胞。在一个方面，多个修饰的细胞是多个修饰的T细胞。在一个方面，至少50%的多个修饰的T细胞表达包含CD45RA和CD62L的一种或多种细胞表面标记物，并且不表达包含CD45RO的一种或多种细胞表面标记物。

[0015] 本公开内容还提供了药物组合物，其包含如本文公开的CAR或组合物(例如，转座子或纳米转座子)，并且进一步包含药学上可接受的载体。

[0016] 本公开内容还提供了通过施用治疗有效量的如本文公开的组合物(例如转座子或纳米转座子)、CAR、细胞、细胞群体或药物组合物中的任一种，来治疗有需要的受试者的增殖病症的方法。在一个方面，增殖病症是癌症。癌症可以为BCMA阳性癌症或PSMA阳性癌症。癌症可以为原发性肿瘤、转移癌、多重抗性癌症、进行性肿瘤或复发癌。癌症可以为实体瘤或血液癌症。癌症可以为肺癌、脑癌、头颈癌、乳腺癌、皮肤癌、肝癌、胰腺癌、胃癌、结肠癌、直肠癌、子宫癌、宫颈癌、卵巢癌、前列腺癌、睾丸癌、皮肤癌、食管癌、淋巴瘤、白血病、急性白血病、急性成淋巴细胞性白血病(ALL)、急性淋巴细胞性白血病、急性髓样白血病(AML)、急性粒细胞性白血病、慢性髓细胞白血病(CML)、慢性淋巴细胞性白血病(CLL)、多毛细胞白血病、骨髓增生异常综合征(MDS)、何杰金氏病、非何杰金氏淋巴瘤或多发性骨髓瘤。

附图说明

[0017] 专利或申请文件含有至少一幅彩色绘图。具有彩色附图的本专利或专利申请公开的副本将在请求和支付必要费用后由专利局提供。

[0018] 图1是比较piggyBac完全质粒和piggyBac纳米转座子(NT)的图谱的一对示意图。

[0019] 图2是描绘在人泛T细胞中,用piggyBac NT改善的转座的图。

[0020] 图3是比较piggyBac NT和piggyBac短NT的图谱的一对示意图。

[0021] 图4是显示通过减少ITR间序列(例如,降低侧接ITR的距离),增强人泛T细胞中的piggyBac转座的图。

[0022] 图5是显示在人泛T细胞中,用抗BCMA嵌合抗原受体(CAR) NT和抗PSMA CAR NT增加的转座的一对图。

[0023] 图6是一系列图,其显示使用抗BCMA CAR NT或抗PSMA CAR NT产生的人CAR-T细胞能够杀死靶肿瘤细胞。

[0024] 图7是一系列图,其显示使用抗BCMA CAR NT或抗PSMA CAR NT产生的人CAR-T细胞在表型组成上是可比较的。

[0025] 图8是一系列图,其显示使用抗BCMA CAR NT或抗PSMA CAR NT产生的人CAR-T细胞具有相似的整合拷贝数。

[0026] 图9是凝胶电泳分析的照片,其证实了单体NT纯度与人泛T细胞中的转座效率相关。

[0027] 图10是显示单体NT纯度与人泛T细胞中的转座效率相关的一对图。

[0028] 图11是显示使用鼠异种移植模型,通过以应激剂量的全长质粒(FLP)相对于NT递送时,P-PSMA-101转座子的临床前评估的示意图。

[0029] 图12是一系列图,其显示当通过FLP相对于NT递送时,用P-PSMA-101转座子治疗的小鼠的肿瘤体积评价。

[0030] 图13是描绘编码BCMA Centyrin CAR (CARTyrin)的P-BCMA-101 piggyBac NT的示意图。纳米转座子编码ITR #1、绝缘子#1、EF1α启动子、BMCA CARTyrin、SV40 PA、绝缘子#2和ITR #2。该序列还编码纳米转座子元件RNA-OUT和R6K起点。

[0031] 图14是描绘编码PSMA CARTyrin的P-PSMA-101 piggyBac NT的示意图。纳米转座子编码ITR #1、绝缘子#1、EF1α启动子、PSMA CARTyrin、SV40 PA、绝缘子#2和ITR #2。该序列还编码纳米转座子元件RNA-OUT和R6K起点。

[0032] 图15是描绘编码BCMA VH CAR (VCAR)的P-BCMA-ALLO1 piggyBac纳米转座子的示意图。纳米转座子编码ITR #1、绝缘子#1、EF1α启动子、BMCA VCAR、SV40 PA、绝缘子#2和ITR #2。该序列还编码纳米转座子元件RNA-OUT和R6K起点。

[0033] 本文引用的所有文件,包括任何交叉引用或有关的专利或申请,在此整体引入本文作为参考用于所有目的,除非明确排除或以其它方式限制。任何文件的引用并非承认它是关于本文公开或请求保护的任何发明的现有技术,或者它单独或与任何其它一个或多个参考文献的任何组合教导、暗示或公开任何此类发明。进一步地,就本文件中术语的任何含义或定义与引入作为参考的文件中相同术语的任何含义或定义冲突而言,应以本文件中赋予该术语的含义或定义为准。

具体实施方式

[0034] 本公开内容提供了纳米转座子，包含纳米转座子的组合物和细胞，制备纳米转座子的方法，以及使用本文所述的纳米转座子、组合物和细胞的方法。

[0035] 本公开内容的纳米转座子被设计为使纳米转座子的ITR内序列降到最低，以使第一ITR序列和第二ITR序列尽可能接近，从而增加转座功效和效率。本公开内容的纳米转座子和包含纳米转座子的组合物在每一种细胞类型中都是有效的；然而，它们对于用于人细胞中特别有效。如本文所述，本公开内容的纳米转座子可以用于增加转座，并且因而增加对于人细胞的基因转移至多个细胞中足够高百分比的细胞。

[0036] 不希望受理论的束缚，通过使ITR间序列或距离降到最低，相应的转座酶可能更能够使两个ITR序列连接在一起，导致ITR内序列从纳米转座子中的增加切除和/或ITR内序列增加整合到靶位点内。此外，在本公开内容的优选方面，纳米转座子、其骨架和/或ITR间序列不包含外来DNA序列。特别是当与非纳米转座子相比时，外来DNA的缺乏进一步改善了转座功效和效率。

[0037] 本公开内容的组合物

本公开内容提供了组合物，其包含：第一核酸序列和第二核酸序列，所述第一核酸序列包含：(a)第一反向末端重复(ITR)或编码第一ITR的序列，(b)第二ITR或编码第二ITR的序列，以及(c) ITR内序列或编码ITR内的序列，其中所述ITR内序列包含转座子序列或编码转座子的序列；和所述第二核酸序列包含ITR间序列或编码ITR间的序列，其中所述ITR间序列的长度等于或小于700个核苷酸。第二核酸序列在本文中也称为骨架区或非整合区。在一个方面，组合物是环状DNA或线性DNA。在一个方面，组合物是质粒或载体。在一个方面，组合物是转座子。在一个优选的方面，组合物是纳米转座子。

[0038] 在一些方面，ITR间序列的长度等于或小于650个核苷酸、等于或小于600个核苷酸、等于或小于550个核苷酸、等于或小于500个核苷酸、等于或小于450个核苷酸、等于或小于400个核苷酸、等于或小于350个核苷酸、等于或小于300个核苷酸、等于或小于250个核苷酸、等于或小于200个核苷酸、等于或小于150个核苷酸、等于或小于100个核苷酸、等于或小于50个核苷酸、等于或小于25个核苷酸、或者等于或小于10个核苷酸。在一些方面，第二核酸序列的长度等于或小于700个核苷酸、等于或小于650个核苷酸、等于或小于600个核苷酸、等于或小于550个核苷酸、等于或小于500个核苷酸、等于或小于450个核苷酸、等于或小于400个核苷酸、等于或小于350个核苷酸、等于或小于300个核苷酸、等于或小于250个核苷酸、等于或小于200个核苷酸、等于或小于150个核苷酸、等于或小于100个核苷酸、等于或小于50个核苷酸、等于或小于25个核苷酸、或者等于或小于10个核苷酸。

[0039] 本公开内容提供了组合物，其包含：第一核酸序列和第二核酸序列，所述第一核酸序列包含：(a)第一反向末端重复(ITR)或编码第一ITR的序列，(b)第二ITR或编码第二ITR的序列，以及(c) ITR内序列或编码ITR内的序列，其中所述ITR内序列包含转座子序列或编码转座子的序列；和所述第二核酸序列包含ITR间序列或编码ITR间的序列，其中所述ITR间序列的长度为1至700个核苷酸，包括端点在内。第二核酸序列在本文中也称为骨架区或非整合区。在一个方面，组合物是环状DNA或线性DNA。在一个方面，组合物是质粒或载体。在一个方面，组合物是转座子。在一个优选的方面，组合物是纳米转座子。

[0040] 在一些方面，ITR间序列的长度为1至650个核苷酸、1至600个核苷酸、1至550个核

苷酸、1至500个核苷酸、1至450个核苷酸、1至400个核苷酸、1至350个核苷酸、1至300个核苷酸、1至250个核苷酸、1至200个核苷酸、1至150个核苷酸、1至100个核苷酸、1至50个核苷酸、1至25个核苷酸或1至10个核苷酸，每个范围都包括端点在内。在一些方面，第二核酸序列的长度为1至650个核苷酸、1至600个核苷酸、1至550个核苷酸、1至500个核苷酸、1至450个核苷酸、1至400个核苷酸、1至350个核苷酸、1至300个核苷酸、1至250个核苷酸、1至200个核苷酸、1至150个核苷酸、1至100个核苷酸、1至50个核苷酸、1至25个核苷酸或1至10个核苷酸，每个范围都包括端点在内。

[0041] 在一些方面，ITR间序列的长度为1至25个核苷酸、1至50个核苷酸、25至50个核苷酸、50至100个核苷酸、100至150个核苷酸、150至200个核苷酸、200至250个核苷酸、250至300个核苷酸、300至350个核苷酸、350至400个核苷酸、400至450个核苷酸、450至500个核苷酸、500至550个核苷酸、550至600个核苷酸、600至650个核苷酸、650至700个核苷酸，每个范围都包括端点在内。在一些方面，第二核酸序列的长度为1至25个核苷酸、1至50个核苷酸、25至50个核苷酸、50至100个核苷酸、100至150个核苷酸、150至200个核苷酸、200至250个核苷酸、250至300个核苷酸、300至350个核苷酸、350至400个核苷酸、400至450个核苷酸、450至500个核苷酸、500至550个核苷酸、550至600个核苷酸、600至650个核苷酸、650至700个核苷酸，每个范围都包括端点在内。

[0042] 在包括本公开内容的短纳米转座子(NTS)的一些方面，ITR间序列的长度为1至10个核苷酸、10至20个核苷酸、20至30个核苷酸、30至40个核苷酸、40至50个核苷酸、50至60个核苷酸、60至70个核苷酸、70至80个核苷酸、80至90个核苷酸、或90至100个核苷酸，每个范围都包括端点在内。在一些方面，包括本公开内容的短纳米转座子(NTS)，核酸序列的长度为1至10个核苷酸、10至20个核苷酸、20至30个核苷酸、30至40个核苷酸、40至50个核苷酸、50至60个核苷酸、60至70个核苷酸、70至80个核苷酸、80至90个核苷酸、或90至100个核苷酸，每个范围都包括端点在内。

[0043] 在一些方面，ITR内序列的长度大于或等于100个核苷酸、大于或等于500个核苷酸、大于或等于1000个核苷酸、大于或等于1500个核苷酸、大于或等于2000个核苷酸、大于或等于2500个核苷酸、大于或等于3000个核苷酸、大于或等于3500个核苷酸、大于或等于4000个核苷酸、大于或等于4500个核苷酸、大于或等于5000个核苷酸、大于或等于5500个核苷酸、大于或等于6000个核苷酸、大于或等于6500个核苷酸、大于或等于7000个核苷酸、大于或等于7500个核苷酸、大于或等于8000个核苷酸、大于或等于8500个核苷酸、大于或等于9000个核苷酸、大于或等于9500个核苷酸、大于或等于10000个核苷酸(10千碱基(kb))、大于或等于50000个核苷酸(50 kb)、大于或等于100000个核苷酸(100 kb)、大于或等于150000个核苷酸(150 kb)、大于或等于200000个核苷酸(200 kb)、大于或等于250000个核苷酸(250 kb)、大于或等于300000个核苷酸(300 kb)、大于或等于350000个核苷酸(350 kb)、大于或等于400000个核苷酸(400 kb)、大于或等于450000个核苷酸(450 kb)、大于或等于500000个核苷酸(50 kb)、或两者之间的任何数目的核苷酸。在一些方面，第二核酸序列的长度大于或等于100个核苷酸、大于或等于500个核苷酸、大于或等于1000个核苷酸、大于或等于1500个核苷酸、大于或等于2000个核苷酸、大于或等于2500个核苷酸、大于或等于3000个核苷酸、大于或等于3500个核苷酸、大于或等于4000个核苷酸、大于或等于4500个核苷酸、大于或等于5000个核苷酸、大于或等于5500个核苷酸、大于或等于6000个核苷酸、大

于或等于6500个核苷酸、大于或等于7000个核苷酸、大于或等于7500个核苷酸、大于或等于8000个核苷酸、大于或等于8500个核苷酸、大于或等于9000个核苷酸、大于或等于9500个核苷酸、大于或等于10000个核苷酸(10千碱基(kb))、大于或等于50000个核苷酸(50 kb)、大于或等于100000个核苷酸(100 kb)、大于或等于150000个核苷酸(150 kb)、大于或等于200000个核苷酸(200 kb)、大于或等于250000个核苷酸(250 kb)、大于或等于300000个核苷酸(300 kb)、大于或等于350000个核苷酸(350 kb)、大于或等于400000个核苷酸(400 kb)、大于或等于450000个核苷酸(450 kb)、大于或等于500000个核苷酸(50 kb) 、或两者之间的任何数目的核苷酸。

[0044] 组合物可以进一步包含复制起点序列或编码复制序列的序列。第一核酸序列或第二核酸序列可以进一步包含复制起点序列或编码复制序列的序列。优选地,第一核酸序列包含复制起点序列或编码复制序列的序列。

[0045] 在一些方面,复制起点序列的长度等于或小于450个核苷酸、等于或小于400个核苷酸、等于或小于350个核苷酸、等于或小于300个核苷酸、等于或小于250个核苷酸、等于或小于200个核苷酸、等于或小于150个核苷酸、等于或小于100个核苷酸、等于或小于50个核苷酸、等于或小于25个核苷酸、或者等于或小于10个核苷酸。在一些方面,复制起点序列的长度为1至450个核苷酸、1至400个核苷酸、1至350个核苷酸、1至300个核苷酸、1至250个核苷酸、1至200个核苷酸、1至150个核苷酸、1至100个核苷酸、1至50个核苷酸、1至25个核苷酸、或1至10个核苷酸,每个范围都包括端点在内。

[0046] 复制起点序列可以包含R6K复制起点。R6K复制起点可以包含R6K γ复制起点。复制起点序列可以包含微型复制起点。微型复制起点可以包含R6K微型复制起点。R6K微型复制起点可以包含R6K γ微型复制起点。R6Kγ微型复制起点的长度为281个核苷酸(281个碱基对),并且包含SEQ ID NO: 15的核酸序列、基本上由其组成或由其组成。

[0047] 组合物可以进一步包含第一可选择标记物或编码第一可选择标记物的序列。第一核酸序列或第二核酸序列可以进一步包含第一可选择标记物或编码第一可选择标记物的序列。优选地,第一核酸序列包含第一可选择标记物或编码第一可选择标记物的序列。

[0048] 在一些方面,第一可选择标记物的长度等于或小于450个核苷酸、等于或小于200个核苷酸、等于或小于150个核苷酸、等于或小于100个核苷酸、等于或小于50个核苷酸、等于或小于25个核苷酸、或者等于或小于10个核苷酸。在一些方面,第一可选择标记物的长度为1至200个核苷酸、1至150个核苷酸、1至100个核苷酸、1至50个核苷酸、1至25个核苷酸、或1至10个核苷酸,每个范围都包括端点在内。

[0049] 第一可选择标记物可以包含蔗糖可选择标记物、荧光标记物、细胞表面标记物或其组合。在一个优选的方面,第一可选择标记物包含蔗糖可选择标记物、基本上由其组成或由其组成。在一个优选的方面,蔗糖可选择标记物包含RNA-OUT选择标记物。RNA-OUT可选择标记物的长度为139个核苷酸(139个碱基对),并且包含SEQ ID NO: 16的核酸序列、基本上由其组成或由其组成。

[0050] 编码第一ITR的序列或编码第二ITR的序列可以包含TTAA、TTAT或TTAX识别序列。编码第一ITR的序列或编码第二ITR的序列可以包含至少2个、至少3个、至少4个、至少5个、至少6个、至少7个、至少8个、至少9个、至少10个、至少11个、至少12个、至少13个、至少14个、至少15个、至少16个、至少17个、至少18个、至少19个或至少20个核苷酸。

[0051] 编码第一ITR的序列或编码第二ITR的序列可以包含核酸序列、可以基本上由核酸序列组成或可以由核酸序列组成，所述核酸序列与SEQ ID NO: 24至少70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100% (或两者之间的任何百分比)等同。编码第一ITR的序列或编码第二ITR的序列可以包含核酸序列、可以基本上由核酸序列组成或可以由核酸序列组成，所述核酸序列与SEQ ID NO: 25至少70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100% (或两者之间的任何百分比)等同。编码第一ITR的序列或编码第二ITR的序列可以包含核酸序列、可以基本上由核酸序列组成或可以由核酸序列组成，所述核酸序列与SEQ ID NO: 26至少70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100% (或两者之间的任何百分比)等同。编码第一ITR的序列或编码第二ITR的序列可以包含核酸序列、可以基本上由核酸序列组成或可以由核酸序列组成，所述核酸序列与SEQ ID NO: 27至少70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100% (或两者之间的任何百分比)等同。编码第一ITR的序列或编码第二ITR的序列可以包含核酸序列、可以基本上由核酸序列组成或可以由核酸序列组成，所述核酸序列与SEQ ID NO: 2至少70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100% (或两者之间的任何百分比)等同。编码第一ITR的序列或编码第二ITR的序列可以包含核酸序列、可以基本上由核酸序列组成或可以由核酸序列组成，所述核酸序列与SEQ ID NO: 14至少70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100% (或两者之间的任何百分比)等同。

[0052] 在一个方面，编码第一ITR的序列包含SEQ ID NO: 24的核酸序列，并且编码第二ITR的第二序列包含SEQ ID NO: 25的核酸序列。在一个方面，编码第一ITR的序列包含SEQ ID NO: 24的核酸序列，并且编码第二ITR的第二序列包含SEQ ID NO: 26的核酸序列。在一个方面，编码第一ITR的序列包含SEQ ID NO: 24的核酸序列，并且编码第二ITR的第二序列包含SEQ ID NO: 27的核酸序列。

[0053] 第一核酸序列可以进一步包含至少一种外源序列，以及能够在哺乳动物细胞中表达外源序列的至少一种启动子。在一个优选的方面，启动子能够在人细胞中表达外源序列。在一个优选的方面，组合物的转座子序列包含至少一种外源序列，以及能够在哺乳动物细胞中表达外源序列的至少一种启动子。

[0054] 启动子可以为组成型启动子。启动子可以为诱导型启动子。启动子可以为细胞类型或组织类型特异性启动子。启动子可以为EF1a启动子(SEQ ID NO: 4)、CMV启动子、MND启动子、SV40启动子、PGK1启动子、Ubc启动子、CAG启动子、H1启动子或U6启动子。在一个优选的方面，启动子是EF1a启动子。在一个方面，第一核酸序列包含编码能够在哺乳动物细胞中表达第一外源序列的第一启动子的第一序列、以及编码能够在哺乳动物细胞中表达第二外源序列的第二启动子的第二序列，其中所述第一启动子是组成型启动子，并且其中所述第二启动子是诱导型启动子。在一个方面，编码第一启动子的第一序列和编码第二启动子的第二序列在相反方向上定向。

[0055] 至少一种外源序列包含以下、基本上由以下组成或由以下组成：编码非天然存在的抗原受体的序列、编码治疗性多肽的序列或其组合。非天然存在的抗原受体可以包含嵌合抗原受体(CAR)、T细胞受体(TCR)、嵌合刺激受体(CSR)、HLA I类组织相容性抗原、α链E重组多肽(HLA-E)、β-2-微球蛋白(B2M)重组多肽或其组合。TCR、CSR、HLA-E和B2M在本文中得到详细描述。在一个优选的方面，非天然存在的抗原受体包含CAR。

[0056] 至少一种外源序列可以进一步包含编码诱导型促凋亡多肽的序列、基本上由其组成或由其组成。诱导型促凋亡多肽在本文中得到详细描述。

[0057] 至少一种外源序列可以进一步包含编码第二可选择标记物的序列、基本上由其组成或由其组成。第二可选择标记物可以编码对于细胞活力和存活所必需的基因产物。当受到选择性细胞培养条件的挑战时，第二可选择标记物可以编码对于细胞活力和存活所必需的基因产物。选择性细胞培养条件可以包含对细胞活力或存活有害的化合物，并且其中基因产物赋予对化合物的抗性。选择基因的非限制性实例包括*neo* (赋予对新霉素的抗性)、DHFR (编码二氢叶酸还原酶并赋予对氨甲蝶呤的抗性)、TYMS (编码胸苷酸合成酶)、MGMT (编码O(6)-甲基鸟嘌呤-DNA甲基转移酶)、多药抗性基因(MDR1)、ALDH1 (编码醛脱氢酶1家族，成员A1)、FRANCF、RAD51C (编码RAD51旁系同源物C)、GCS (编码葡萄糖神经酰胺合酶)、NKX2.2 (编码NK2同源框2)或其任何组合。

[0058] 第二可选择标记物可以为可检测标记物。可检测标记物可以为荧光标记物、细胞表面标记物或代谢标记物。在一个优选的方面，第二可选择标记物包含编码二氢叶酸还原酶(DHFR)突变蛋白酶的序列。DHFR突变蛋白酶包含SEQ ID NO: 52的氨基酸序列、基本上由其组成或由其组成。DHFR突变蛋白酶由多核苷酸编码，所述多核苷酸包含SEQ ID NO: 53或SEQ ID NO: 11的核酸序列、基本由其组成或由其组成。DHFR突变蛋白酶的氨基酸序列可以进一步包含在位置80、113或153中的一个或多个位置处的突变。DHFR突变蛋白酶的氨基酸序列可以包含以下中的一个或多个：在位置80处的苯丙氨酸(F)或亮氨酸(L)取代、在位置113处的亮氨酸(L)或缬氨酸(V)取代、以及在位置153处的缬氨酸(V)或天冬氨酸(D)取代。

[0059] 至少一种外源序列可以进一步包含编码至少一种自切割肽的序列、基本上由其组成或由其组成。例如，自切割肽可以位于CAR和诱导型促凋亡多肽之间；或者，自切割肽可以位于CAR和第二可选择标记物之间。

[0060] 至少一种外源序列可以进一步包含编码至少两种自切割肽的序列、基本上由其组成或由其组成。例如，第一自切割肽位于CAR的上游或紧上游，而第二自切割肽位于CAR的下游或紧下游；或者，第一自切割肽和第二自切割肽侧接CAR。例如，第一自切割肽位于诱导型促凋亡多肽的上游或紧上游，而第二自切割肽位于诱导型促凋亡多肽的下游或紧下游；或者，第一自切割肽和第二自切割肽侧接诱导型促凋亡多肽。例如，第一自切割肽位于第二可选择标记物的上游或紧上游，而第二自切割肽位于第二可选择标记物的下游或紧下游；或者，第一自切割肽和第二自切割肽侧接第二可选择标记物。

[0061] 自切割肽的非限制性实例包括T2A肽、GSG-T2A肽、E2A肽、GSG-E2A肽、F2A肽、GSG-F2A肽、P2A肽或GSG-P2A肽。T2A肽包含氨基酸序列、基本上由氨基酸序列组成或由氨基酸序列组成，所述氨基酸序列与SEQ ID NO: 54至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100% (或两者之间的任何百分比)等同。GSG-T2A多肽包含氨基酸序列、基本上由氨基酸序列组成或由氨基酸序列组成，所述氨基酸序列与SEQ ID NO: 55至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100% (或两者之间的任何百分比)等同。GSG-T2A多肽由多核苷酸编码，所述多核苷酸包含核酸序列或由核酸序列组成，所述核酸序列与SEQ ID NO: 7、SEQ ID NO: 8、SEQ ID NO: 10、SEQ ID NO: 56至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100% (或两者之间的任何百分比)等同。E2A肽包含氨基酸序列、基本上由氨基酸序列组成或由氨基酸序列组成，所述氨基酸序列与SEQ ID NO: 57至少75%、80%、85%、90%、95%、96%、

97%、98%、99%或100% (或两者之间的任何百分比)等同。GSG-E2A肽包含氨基酸序列、基本上由氨基酸序列组成或由氨基酸序列组成,所述氨基酸序列与SEQ ID NO: 58至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100% (或两者之间的任何百分比)等同。F2A肽包含氨基酸序列、基本上由氨基酸序列组成或由氨基酸序列组成,所述氨基酸序列与SEQ ID NO: 59至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100% (或两者之间的任何百分比)等同。GSG-F2A肽包含氨基酸序列、基本上由氨基酸序列组成或由氨基酸序列组成,所述氨基酸序列与SEQ ID NO: 60至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100% (或两者之间的任何百分比)等同。P2A肽包含氨基酸序列、基本上由氨基酸序列组成或由氨基酸序列组成,所述氨基酸序列与SEQ ID NO: 61至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100%(或两者之间的任何百分比)等同。GSG-P2A肽包含氨基酸序列、基本上由氨基酸序列组成或由氨基酸序列组成,所述氨基酸序列与SEQ ID NO: 62至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100% (或两者之间的任何百分比)等同。

[0062] 包含至少一种外源序列和能够在哺乳动物细胞中表达外源序列的至少一种启动子的第一核酸序列,可以进一步包含编码绝缘子的至少一种序列。在一个方面,第一核酸序列可以包含编码第一绝缘子的第一序列和编码第二绝缘子的第二序列。在一些实施方案中,编码第一绝缘子或第二绝缘子的序列包含核酸序列、基本由核酸序列组成或由核酸序列组成,所述核酸序列与SEQ ID NO: 3或SEQ ID NO: 13至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100% (或两者之间的任何百分比)等同。

[0063] 包含至少一种外源序列和能够在哺乳动物细胞中表达外源序列的至少一种启动子的第一核酸序列,可以进一步包含聚腺苷(聚A)序列。包含至少一种外源序列、能够在哺乳动物细胞中表达外源序列的至少一种启动子和编码绝缘子的至少一种序列的第一核酸序列,可以进一步包含聚腺苷(聚A)序列。聚A序列可以分离自或衍生自病毒聚A序列。聚A序列可以分离自或衍生自(SV40)聚A序列。在一些实施方案中,编码第一绝缘子或第二绝缘子的序列包含核酸序列、基本由核酸序列组成或由核酸序列组成,所述核酸序列与SEQ ID NO: 12至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100% (或两者之间的任何百分比)等同。

[0064] 在一个方面,组合物不包含编码外来DNA的序列。在一个方面,第一核酸序列不包含编码外来DNA的序列。在一个方面,第二核酸序列不包含编码外来DNA的序列。在一个方面,组合物包含编码外来DNA的序列。在一个方面,第一核酸序列包含编码外来DNA的序列。在一个方面,第二核酸序列包含编码外来DNA的序列。外来DNA是这样的DNA序列,其并非衍生自或得自与外源序列将在其中表达的哺乳动物细胞相同的生物体。例如,外来DNA可以是来自病毒的DNA,而不是来自哺乳动物的DNA;或者外来DNA可以是来自爬行动物的DNA,而不是来自哺乳动物的DNA。在另一个方面,外来DNA可以来自一种哺乳动物,但该哺乳动物不同于外源序列将在其中表达的哺乳动物。例如,外来DNA来自大鼠而不是人。

[0065] 在一个方面,组合物不包含重组位点、切除位点、连接位点或其组合。在一个方面,组合物不包含重组事件、切除事件、连接事件或其组合的产物。在一个方面,组合物并非衍生自重组事件、切除事件、连接事件或其组合。

[0066] 在一个方面,第一核酸序列不包含重组位点、切除位点、连接位点或其组合。在一个方面,第一核酸序列不包含重组事件、切除事件、连接事件或其组合的产物。在一个方面,

第一核酸序列并非衍生自重组事件、切除事件、连接事件或其组合。

[0067] 在一个方面，第二核酸序列不包含重组位点、切除位点、连接位点或其组合。在一个方面，第二核酸序列不包含重组事件、切除事件、连接事件或其组合的产物。在一个方面，第二核酸序列并非衍生自重组事件、切除事件、连接事件或其组合。

[0068] 重组位点可以包含起因于重组事件的序列，可以包含其为重组事件的产物的序列，或可以包含重组酶的活性(例如，重组酶位点)。

[0069] 嵌合抗原受体(CAR)

本公开内容还提供了包含CAR的组合物(例如，纳米转座子)，其中所述CAR包括包含抗原识别区的胞外结构域；跨膜结构域和包含至少一个共刺激结构域的胞内结构域。CAR可以进一步包含在抗原识别结构域和跨膜结构域之间的铰链区。

[0070] 抗原识别区可以包含至少一种单链可变片段(scFv)、Centyrin、单结构域抗体或其组合。在一个方面，至少一种单结构域抗体是VHH。在一个方面，至少一种单结构域抗体是VH。

[0071] scFv

本公开内容的组合物(例如，转座子或纳米转座子)可以包含CAR；并且在一些方面，CAR的抗原识别区可以包含一种或多种scFv组合物，以识别并结合特异性靶蛋白/抗原。抗原识别区可以包含至少两种scFv。抗原识别区可以包含至少三种scFv。在一个方面，本公开内容的CAR是包含特异性结合两种不同抗原的至少两种scFv的双特异性CAR。

[0072] scFv组合物包含抗体的重链可变区和轻链可变区。scFv是免疫球蛋白的重(VH)链和轻(VL)链的可变区的融合蛋白，并且VH和VL结构域由短肽接头连接。尽管去除恒定区并引入接头，scFv仍可以保留原始免疫球蛋白的特异性。在一些方面，接头多肽包含SEQ ID NO: 33的氨基酸序列、基本上由其组成或由其组成。接头多肽可以由多核苷酸编码，所述多核苷酸包含SEQ ID NO: 34的核酸序列、基本上由其组成或由其组成。

[0073] Centyrin

本公开内容的组合物(例如，转座子或纳米转座子)可以包含CAR；并且在一些方面，CAR的抗原识别区可以包含一种或多种Centyrin组合物，以识别并结合特异性靶蛋白/抗原。特异性结合抗原的Centyrin可以用于将细胞(例如，细胞毒性免疫细胞)的特异性导向特异性抗原。包含Centyrin的CAR在本文中被称为CARTyrin。

[0074] 本公开内容的Centyrin可以包含蛋白质支架，其中所述支架能够特异性结合抗原。本公开内容的Centyrin可以包括包含至少一个纤连蛋白III型(FN3)结构域的共有序列的蛋白质支架，其中所述支架能够特异性结合抗原。至少一个纤连蛋白III型(FN3)结构域可以衍生自人蛋白质。人蛋白质可以为腱生蛋白C。共有序列包含氨基酸序列、基本上由氨基酸序列组成或由氨基酸序列组成，所述氨基酸序列与SEQ ID NO: 84至少74%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100% (或两者之间的任何百分比)等同，或者共有序列包含氨基酸序列、基本上由氨基酸序列组成或由氨基酸序列组成，所述氨基酸序列与SEQ ID NO: 85至少74%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100% (或两者之间的任何百分比)等同。共有序列由多核苷酸编码，所述多核苷酸包含核酸序列、基本上由核酸序列组成或由核酸序列组成，所述核酸序列与SEQ ID NO: 86至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100% (或两者之间的任何百分比)等同。

[0075] 共有序列可以在以下内的一个或多个位置处进行修饰:(a)包含在共有序列的位置13-16处的氨基酸残基TEDS (SEQ ID NO: 87)或由其组成的A-B环;(b)包含在共有序列的位置22-28处的氨基酸残基TAPDAAF (SEQ ID NO: 88)或由其组成的B-C环;(c)包含在共有序列的位置38-43处的氨基酸残基SEKVGE (SEQ ID NO: 89)或由其组成的C-D环;(d)包含在共有序列的位置51-54处的氨基酸残基GSER (SEQ ID NO: 90)或由其组成的D-E环;(e)包含在共有序列的位置60-64处的氨基酸残基GLKPG (SEQ ID NO: 91)或由其组成的E-F环;(f)包含在共有序列的位置75-81处的氨基酸残基KGGHRSN (SEQ ID NO: 92)或由其组成的F-G环;或(g) (a)-(f)的任何组合。本公开内容的Centyrin可以包含至少5个纤连蛋白III型(FN3)结构域、至少10个纤连蛋白III型(FN3)结构域、或至少15个纤连蛋白III型(FN3)结构域的共有序列。

[0076] 术语“抗体模拟物”预期描述这样的有机化合物,其特异性结合靶序列并具有不同于天然存在的抗体的结构。抗体模拟物可以包含蛋白质、核酸或小分子。本公开内容的抗体模拟物与之特异性结合的靶序列可以为抗原。抗体模拟物可以提供超过抗体的优良性质,包括但不限于优良的溶解性、组织渗透性、针对热和酶的稳定性(例如,对酶促降解的抗性)以及较低的生产成本。示例性抗体模拟物包括但不限于亲和体、afflilin、affimer、affitin、alphabody、anticalin和avimer (也称为亲合力多聚体)、DARPin (预设计锚蛋白重复蛋白)、Fynomer、Kunitz结构域肽和单体(monobody)。

[0077] 本公开内容的亲和体分子包含蛋白质支架,所述蛋白质支架包含不含任何二硫键的一个或多个α螺旋或者由其组成。优选地,本公开内容的亲和体分子包含三个α螺旋或由三个α螺旋组成。例如,本公开内容的亲和体分子可以包含免疫球蛋白结合结构域。本公开内容的亲和体分子可以包含蛋白A的Z结构域。

[0078] 本公开内容的affilin分子包含通过修饰例如γ-B晶体蛋白或泛素的暴露氨基酸而产生的蛋白质支架。affilin分子在功能上模拟抗体对抗原的亲和力,但在结构上不模拟抗体。在用于制备affililin的任何蛋白质支架中,在正确折叠的蛋白质分子中,对溶剂或可能的结合配偶体可接近的那些氨基酸被视为暴露的氨基酸。这些暴露的氨基酸中的任何一个或多个可以进行修饰,以特异性结合靶序列或抗原。

[0079] 本公开内容的affimer分子包括包含高度稳定的蛋白质的蛋白质支架,所述蛋白质被改造为展示肽环,所述肽环提供对于特异性靶序列的高亲和力结合位点。本公开内容的示例性affimer分子包含基于胱抑素蛋白或其三级结构的蛋白质支架。本公开内容的示例性affimer分子可以共享包含位于反平行的β折叠之上的α螺旋的共同三级结构。

[0080] 本公开内容的affitin分子包含人工蛋白质支架,其结构可以衍生自例如DNA结合蛋白(例如,DNA结合蛋白Sac7d)。本公开内容的affitin选择性地结合靶序列,其可以为抗原的全部或部分。本公开内容的示例性affitin通过以下进行制造:使DNA结合蛋白的结合表面上的一个或多个氨基酸序列随机化,并且使所得的蛋白经受核糖体展示和选择。本公开内容的affitin的靶序列可以在例如基因组中或者在肽、蛋白质、病毒或细菌的表面上找到。在一些方面,affitin分子可以用作酶的特异性抑制剂。本公开内容的affitin分子可以包括耐热蛋白质或其衍生物。

[0081] 本公开内容的alphabody分子也可以被称为细胞穿透alphabody (CPAB)。本公开内容的alphabody分子包含结合各种靶序列(包括抗原)的小蛋白(通常小于10 kDa)。

alphabody分子能够到达并结合细胞内靶序列。在结构上，本公开内容的alphabody分子包含形成单链α螺旋(类似于天然存在的卷曲螺旋结构)的人工序列。本公开内容的alphabody分子可以包括包含一种或多种氨基酸的蛋白质支架，所述一种或多种氨基酸被修饰，以特异性结合靶蛋白。不管分子的结合特异性如何，本公开内容的alphabody分子维持正确的折叠和热稳定性。

[0082] 本公开内容的anticalin分子包含与蛋白质或小分子中的靶序列或位点结合的人工蛋白质。本公开内容的anticalin分子可以包含衍生自人脂质运载蛋白的人工蛋白质。本公开内容的anticalin分子可以用于代替例如单克隆抗体或其片段。anticalin分子可以证实比单克隆抗体或其片段更优良的组织渗透性和热稳定性。本公开内容的示例性anticalin分子可以包含约180个氨基酸，具有大约20 kDa的质量。在结构上，本公开内容的anticalin分子包含桶状结构和附着的α螺旋，所述桶状结构包含通过环成对连接的反平行的β链。在一些方面，本公开内容的anticalin分子包含桶状结构和附着的α螺旋，所述桶状结构包含通过环成对连接的八条反平行的β链。

[0083] 本公开内容的avimer分子包含与靶序列(其也可以为抗原)特异性结合的人工蛋白质。本公开内容的avimer可以识别在同一靶或不同靶内的多重结合位点。当本公开内容的avimer识别多于一种靶时，avimer模拟双特异性抗体的功能。人工蛋白质avimer可以包含各自大约30-35个氨基酸的两种或更多种肽序列。这些肽可以经由一种或多种接头肽连接。avimer的一种或多种肽的氨基酸序列可以衍生自膜受体的A结构域。avimer具有刚性结构，其可以任选地包含二硫键和/或钙。与抗体相比，本公开内容的avimer可以证实更大的热稳定性。

[0084] 本公开内容的DARPin (预设计锚蛋白重复蛋白)包含对于靶序列具有高特异性和高亲和力的遗传改造的蛋白质、重组蛋白质或嵌合蛋白质。在一些方面，本公开内容的DARPin衍生自锚蛋白，并且任选地包含锚蛋白的至少三个重复基序(也称为重复结构单元)。锚蛋白介导高亲和力的蛋白质-蛋白质相互作用。本公开内容的DARPin包含大的靶相互作用表面。

[0085] 本公开内容的Fynomer包含小结合蛋白(约7 kDa)，其衍生自人Fyn SH3结构域，并且改造为以与抗体相等的亲和力和相等的特异性结合靶序列和分子。

[0086] 本公开内容的Kunitz结构域肽包括包含Kunitz结构域的蛋白质支架。Kunitz结构域包含用于抑制蛋白酶活性的活性位点。在结构上，本公开内容的Kunitz结构域包含富含二硫化物的α+β折叠。这种结构通过牛胰胰蛋白酶抑制剂来例证。Kunitz结构域肽识别特异性蛋白质结构，并且充当竞争性蛋白酶抑制剂。本公开内容的Kunitz结构域可以包含艾卡拉肽(衍生自人脂蛋白相关凝血抑制剂(LACI))。

[0087] 本公开内容的单体是大小与单链抗体可比较的小蛋白质(包含约94个氨基酸并且具有约10 kDa的质量)。这些遗传改造的蛋白特异性结合靶序列包括抗原。本公开内容的单体可以特异性靶向一种或多种不同的蛋白质或靶序列。在一些方面，本公开内容的单体包含模拟人纤连蛋白的结构的蛋白质支架，且更优选地，模拟纤连蛋白的第十细胞外III型结构域的结构。纤连蛋白的第十细胞外III型结构域以及其单体模拟物，含有形成桶的七个β折叠以及在每侧上的三个暴露的环，其对应于抗体的三个互补决定区(CDR)。与抗体的可变结构域的结构形成对比，单体缺乏金属离子的任何结合位点以及中心二硫键。可以通过修

饰环BC和FG来优化多特异性单体。本公开内容的单体可以包含adnectin。

[0088] VHH

本公开内容的组合物(例如,转座子或纳米转座子)可以包含CAR;并且在一些方面,CAR的抗原识别区可以包含至少一种单结构域抗体(SdAb),以识别并结合特异性靶蛋白/抗原。在一个方面,单结构域抗体是VHH。VHH是在骆驼科动物中发现的重链抗体。特异性结合抗原的VHH可以用于将细胞(例如,细胞毒性免疫细胞)的特异性导向特异性抗原。抗原识别区可以包含至少两种VHH。抗原识别区可以包含至少三种VHH。在一个方面,本公开内容的CAR是双特异性CAR,其包含特异性结合两种不同抗原的至少两种VHH。包含VHH的CAR在本文中被称为VCAR。

[0089] 本公开内容的至少一种VHH蛋白或VCAR可以任选地由细胞系、混合细胞系、永生化细胞或永生化细胞的克隆群体产生,如本领域众所周知的。参见例如,Ausubel等人,编辑,Current Protocols in Molecular Biology,John Wiley & Sons,Inc.,NY,N.Y. (1987-2001);Sambrook等人,Molecular Cloning: A Laboratory Manual,第2版,Cold Spring Harbor,N.Y. (1989);Harlow和Lane,Antibodies,a Laboratory Manual,Cold Spring Harbor,N.Y. (1989);Colligan等人,编辑,Current Protocols in Immunology,John Wiley & Sons,Inc.,NY (1994-2001);Colligan等人,Current Protocols in Protein Science,John Wiley & Sons,NY,N.Y.,(1997-2001)。

[0090] 可以改变、添加和/或缺失来自VHH蛋白的氨基酸,以减少免疫原性,或者减少、增强或修饰结合、亲和力、结合速率、解离速率、亲合力、特异性、半衰期、稳定性、溶解性或任何其它合适的特性,如本领域已知的。

[0091] 任选地,VHH蛋白可以进行改造,伴随对于抗原的高亲和力和其它有利的生物学性质的保留。为了实现这一目标,可以通过使用亲本序列和改造序列的三维模型,分析亲本序列和各种概念性改造产物的过程,来任选地制备VHH蛋白。三维模型是普遍可用的并且是本领域技术人员熟悉的。计算机程序是可获得的,其示出且展示所选候选序列的可能的三维构象结构,并且可以测量可能的免疫原性(例如,Monrovia,Calif.的Xencor,Inc.的Immunofilter程序)。这些展示的检查允许分析残基在候选序列的功能中的可能作用,即,分析影响候选VHH蛋白结合其抗原的能力的残基。以这种方式,可以从亲本序列和参考序列中选择且组合残基,使得实现所需的特性,例如对于靶抗原的亲和力。可替代地或除上述程序之外,可以使用其它合适的改造方法。可以使用核苷酸(DNA或RNA展示)或肽展示文库,例如在体外展示中,方便地实现筛选与相似蛋白质或片段特异性结合的VHH。可以用本公开内容的VHH或VCAR执行竞争性测定法,以便确定哪些蛋白质、抗体和其它拮抗剂与本公开内容的VHH或VCAR竞争结合靶蛋白和/或共享表位区域。如本领域普通技术人员容易知道的这些测定法,评估了拮抗剂或配体之间对于蛋白质上有限数目的结合位点的竞争。

[0092] VH

本公开内容的组合物(例如,转座子或纳米转座子)可以包含CAR;并且在一些方面,CAR的抗原识别区可以包含至少一种单结构域抗体(SdAb),以识别并结合特异性靶蛋白/抗原。在一个方面,单结构域抗体是VH。VH是衍生自常见IgG的单结构域结合剂。特异性结合抗原的VH可以用于将细胞(例如,细胞毒性免疫细胞)的特异性导向特异性抗原。抗原识别区可以包含至少两种VH。抗原识别区可以包含至少三种VH。在一个方面,本公开内容的

CAR是双特异性CAR,其包含特异性结合两种不同抗原的至少两种VH。

[0093] VH可以分离自或衍生自人序列。VH可以包含人CDR序列和/或人构架序列和非人或人源化序列(例如,大鼠Fc结构域)。在一些方面,VH是完全人源化的VH。在一些方面,VH既不是天然存在的抗体,也不是天然存在的抗体的片段。在一些方面,VH不是单克隆抗体的片段。在一些方面,VH是UniDab抗体(TeneoBio)。在一些方面,对VH进行修饰,以去除Fc结构域或其一部分。在一些方面,对VH的构架序列进行修饰,以例如改善表达、降低免疫原性或改善功能。

[0094] VH可以使用UniRat (TeneoBio)系统和"基于NGS的发现(NGS-based Discovery)"进行完全改造,以产生VH。使用这种方法,特异性VH并非天然存在的,而是使用完全改造的系统生成的。VH并非衍生自天然存在的单克隆抗体(mAb),所述单克隆抗体或直接从宿主(例如,小鼠、大鼠或人)中分离、或者直接从细胞或细胞系(杂交瘤)的单个克隆中分离。随后,并未从所述细胞系中克隆这些VH。相反,VH序列使用UniRat系统作为转基因进行完全改造,所述转基因包含具有大鼠Fc结构域的人可变区(VH结构域),并且因此是不含轻链的人/大鼠嵌合体,并且与标准mAb形式不同。天然大鼠基因被敲除,并且在大鼠中表达的唯一抗体来自具有与大鼠Fc连接的VH结构域的转基因(UniAb)。这些是UniRat中表达的唯一Ab。下一代测序(NGS)和生物信息学用于鉴定在免疫后,通过UniRat生成的重链抗体的完全抗原特异性储库。然后,使用独特的基因组装方法,以将抗体储库序列信息转换成全人重链抗体的大型集合,所述全人重链抗体可以在体外筛选各种功能。在一些方面,完全人源化的VH通过在体外将人VH结构域与人Fc融合而生成(以生成非天然存在的重组VH抗体)。在一些方面,VH是完全人源化的,但它们在体内表达为不含轻链的人/大鼠嵌合体(人VH、大鼠Fc)。在体内表达为不含轻链的人/大鼠嵌合体(人VH、大鼠Fc)的完全人源化的VH为约80kDa (相对于150 kDa)。

[0095] 本公开内容的CAR可以以选自以下的至少一种亲和力K_D结合人抗原:小于或等于10^{-9}M、小于或等于10^{-10}M、小于或等于10^{-11}M、小于或等于10^{-12}M、小于或等于10^{-13}M、小于或等于10^{-14}M、以及小于或等于10^{-15}M。K_D可以通过任何手段进行确定,所述手段包括但不限于表面等离振子共振。

[0096] 在一个方面,所公开的CAR的抗原识别区包含至少一种抗BCMA Centyrin。抗BCMA Centyrin包含氨基酸序列、基本上由氨基酸序列组成或由氨基酸序列组成,所述氨基酸序列与SEQ ID NO: 29至少95%、96%、97%、98%、99%或100% (或两者之间的任何百分比)等同。抗BCMA Centyrin由多核苷酸编码,所述多核苷酸包含核酸序列、基本上由核酸序列组成或由核酸序列组成,所述核酸序列与SEQ ID NO: 28至少95%、96%、97%、98%、99%或100% (或两者之间的任何百分比)等同。

[0097] 包含抗BCMA Centyrin的CAR在本文中被称为BCMA CARTyrin。在一个优选的方面,BCMA CARTyrin包含氨基酸序列、基本上由氨基酸序列组成或由氨基酸序列组成,所述氨基酸序列与SEQ ID NO: 30至少95%、96%、97%、98%、99%或100% (或两者之间的任何百分比)等同。BCMA CARTyrin由多核苷酸编码,所述多核苷酸包含核酸序列、基本上由核酸序列组成或由核酸序列组成,所述核酸序列与SEQ ID NO: 9至少95%、96%、97%、98%、99%或100% (或两者之间的任何百分比)等同。

[0098] 包含BCMA CARTyrin的本公开内容的组合物(例如,纳米转座子)包含氨基酸序列、

基本上由氨基酸序列组成或由氨基酸序列组成，所述氨基酸序列与SEQ ID NO: 17至少95%、96%、97%、98%、99%或100% (或两者之间的任何百分比)等同。包含BCMA CARTyrin的本公开内容的组合物(例如，纳米转座子)由多核苷酸编码，所述多核苷酸包含核酸序列、基本上由核酸序列组成或由核酸序列组成，所述核酸序列与SEQ ID NO: 1至少95%、96%、97%、98%、99%或100% (或两者之间的任何百分比)等同。包含BCMA CARTyrin的本公开内容的组合物(例如，纳米转座子)在本文中也被称为P-BCMA-101-转座子(如图13中所示)。

[0099] 在一个方面，所公开的CAR的抗原识别区包含至少一种抗PSMA Centyrin。抗PSMA Centyrin包含氨基酸序列、基本上由氨基酸序列组成或由氨基酸序列组成，所述氨基酸序列与SEQ ID NO: 94至少95%、96%、97%、98%、99%或100% (或两者之间的任何百分比)等同。抗PSMA Centyrin由多核苷酸编码，所述多核苷酸包含核酸序列、基本上由核酸序列组成或由核酸序列组成，所述核酸序列与SEQ ID NO: 93至少95%、96%、97%、98%、99%或100% (或两者之间的任何百分比)等同。

[0100] 包含抗PSMA Centyrin的CAR在本文中被称为PSMA CARTyrin。在一个优选的方面，PSMA CARTyrin包含氨基酸序列、基本上由氨基酸序列组成或由氨基酸序列组成，所述氨基酸序列与SEQ ID NO: 95至少95%、96%、97%、98%、99%或100% (或两者之间的任何百分比)等同。PSMA CARTyrin由多核苷酸编码，所述多核苷酸包含核酸序列、基本上由核酸序列组成或由核酸序列组成，所述核酸序列与SEQ ID NO: 19至少95%、96%、97%、98%、99%或100% (或两者之间的任何百分比)等同。

[0101] 包含PSMA CARTyrin的本公开内容的组合物(例如，纳米转座子)包含氨基酸序列、基本上由氨基酸序列组成或由氨基酸序列组成，所述氨基酸序列与SEQ ID NO: 20至少95%、96%、97%、98%、99%或100% (或两者之间的任何百分比)等同。包含PSMA CARTyrin的本公开内容的组合物(例如，纳米转座子)由多核苷酸编码，所述多核苷酸包含核酸序列、基本上由核酸序列组成或由核酸序列组成，所述核酸序列与SEQ ID NO: 18至少95%、96%、97%、98%、99%或100% (或两者之间的任何百分比)等同。包含PSMA CARTyrin的本公开内容的组合物(例如，纳米转座子)在本文中也被称为P-PSMA-101-转座子(如图14中所示)。

[0102] 在一个方面，所公开的CAR的抗原识别区包含至少一种抗BCMA VH。抗BCMA VH包含氨基酸序列、基本上由氨基酸序列组成或由氨基酸序列组成，所述氨基酸序列与SEQ ID NO: 97至少95%、96%、97%、98%、99%或100% (或两者之间的任何百分比)等同。抗BCMA VH由多核苷酸编码，所述多核苷酸包含核酸序列、基本上由核酸序列组成或由核酸序列组成，所述核酸序列与SEQ ID NO: 96至少95%、96%、97%、98%、99%或100% (或两者之间的任何百分比)等同。

[0103] 包含抗BCMA VH的CAR在本文中被称为BCMA VCAR。在一个优选的方面，BCMA VCAR包含氨基酸序列、基本上由氨基酸序列组成或由氨基酸序列组成，所述氨基酸序列与SEQ ID NO: 98至少95%、96%、97%、98%、99%或100% (或两者之间的任何百分比)等同。BCMA VCAR由多核苷酸编码，所述多核苷酸包含核酸序列、基本上由核酸序列组成或由核酸序列组成，所述核酸序列与SEQ ID NO: 22至少95%、96%、97%、98%、99%或100% (或两者之间的任何百分比)等同。

[0104] 包含BCMA VCAR的本公开内容的组合物(例如，纳米转座子)包含氨基酸序列、基本上由氨基酸序列组成或由氨基酸序列组成，所述氨基酸序列与SEQ ID NO: 23至少95%、

96%、97%、98%、99%或100% (或两者之间的任何百分比)等同。包含BCMA VCAR的本公开内容的组合物(例如,纳米转座子)由多核苷酸编码,所述多核苷酸包含核酸序列、基本上由核酸序列组成或由核酸序列组成,所述核酸序列与SEQ ID NO: 21至少95%、96%、97%、98%、99%或100% (或两者之间的任何百分比)等同。包含BCMA VCAR的本公开内容的组合物(例如,纳米转座子)在本文中也被称为P-BCMA-ALLO1-转座子(如图15中所示)。

[0105] 胞外结构域可以包含信号肽。信号肽可以包含编码人CD2、CD3δ、CD3ε、CD3γ、CD3ζ、CD4、CD8α、CD19、CD28、4-1BB或GM-CSFR信号肽的序列。在一个优选的方面,信号肽包含以下、基本上由以下组成或由以下组成:人CD8 α (CD8α)信号肽(SP)或其一部分。人CD8α SP包含氨基酸序列、基本上由氨基酸序列组成或由氨基酸序列组成,所述氨基酸序列与SEQ ID NO: 31至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100% (或两者之间的任何百分比)等同。优选地,人CD8α SP包含SEQ ID NO: 31的氨基酸序列、基本上由其组成或由其组成。

[0106] 人CD8α SP由多核苷酸编码,所述多核苷酸包含核酸序列、基本上由核酸序列组成或由核酸序列组成,所述核酸序列与SEQ ID NO: 32至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100% (或两者之间的任何百分比)等同。优选地,人CD8α SP由多核苷酸编码,所述多核苷酸包含SEQ ID NO: 32的氨基酸序列、基本上由其组成或由其组成。

[0107] 铰链结构域或铰链区可以包含人CD8α、IgG4、CD4序列或其组合。在一个优选的方面,铰链可以包含以下、基本上由以下组成或由以下组成:人CD8 α (CD8α)铰链或其一部分。人CD8a铰链包含氨基酸序列、基本上由氨基酸序列组成或由氨基酸序列组成,所述氨基酸序列与SEQ ID NO: 35至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100% (或两者之间的任何百分比)等同。优选地,人CD8α铰链结构域包含SEQ ID NO: 35的氨基酸序列、基本上由其组成或由其组成。

[0108] 人CD8α铰链由多核苷酸编码,所述多核苷酸包含核酸序列、基本上由核酸序列组成或由核酸序列组成,所述核酸序列与SEQ ID NO: 36至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100% (或两者之间的任何百分比)等同。优选地,人CD8α铰链结构域由多核苷酸编码,所述多核苷酸包含SEQ ID NO: 36的核酸序列、基本上由其组成或由其组成。

[0109] 跨膜结构域可以包含以下、基本上由以下组成或由以下组成:编码人CD2、CD3δ、CD3ε、CD3γ、CD3ζ、CD4、CD8α、CD19、CD28、4-1BB或GM-CSFR跨膜结构域的序列。优选地,跨膜结构域可以包含人CD8 α (CD8α)跨膜结构域或其一部分、基本上由其组成或由其组成。CD8a跨膜结构域包含氨基酸序列、基本上由氨基酸序列组成或由氨基酸序列组成,所述氨基酸序列与SEQ ID NO: 37至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100% (或两者之间的任何百分比)等同。优选地,人CD8α跨膜结构域包含SEQ ID NO: 37的氨基酸序列、基本上由其组成或由其组成。

[0110] CD8α跨膜结构域由多核苷酸编码,所述多核苷酸包含核酸序列、基本上由核酸序列组成或由核酸序列组成,所述核酸序列与SEQ ID NO: 38至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100% (或两者之间的任何百分比)等同。优选地,CD8α跨膜结构域由多核苷酸编码,所述多核苷酸包含SEQ ID NO: 38的核酸序列、基本上由其组成或由其组成。

[0111] 至少一个共刺激结构域可以包含以下、基本上由以下组成或由以下组成:人4-1BB、CD28、CD3 ζ (CD3ζ)、CD40、ICOS、MyD88、OX-40细胞内结构域或其任何组合。优选地,至

少一个共刺激结构域包含CD3ζ、4-1BB共刺激结构域或其组合。

[0112] 4-1BB细胞内结构域包含氨基酸序列、基本上由氨基酸序列组成或由氨基酸序列组成，所述氨基酸序列与SEQ ID NO: 39至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100% (或两者之间的任何百分比)等同。优选地，4-1BB细胞内结构域包含SEQ ID NO: 39的氨基酸序列、基本上由其组成或由其组成。

[0113] 4-1BB细胞内结构域由多核苷酸编码，所述多核苷酸包含核酸序列、基本上由核酸序列组成或由核酸序列组成，所述核酸序列与SEQ ID NO: 40至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100% (或两者之间的任何百分比)等同。优选地，4-1BB细胞内结构域由多核苷酸编码，所述多核苷酸包含SEQ ID NO: 40的核酸序列、基本上由其组成或由其组成。

[0114] CD3ζ细胞内结构域包含氨基酸序列、基本上由氨基酸序列组成或由氨基酸序列组成，所述氨基酸序列与SEQ ID NO: 41至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100%(或两者之间的任何百分比)等同。优选地，CD3ζ细胞内结构域包含SEQ ID NO: 41的氨基酸序列、基本上由其组成或由其组成。

[0115] CD3ζ细胞内结构域由多核苷酸编码，所述多核苷酸包含核酸序列、基本上由核酸序列组成或由核酸序列组成，所述核酸序列与SEQ ID NO: 42至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100% (或两者之间的任何百分比)等同。优选地，CD3ζ细胞内结构域由多核苷酸编码，所述多核苷酸包含SEQ ID NO: 42的核酸序列、基本上由其组成或由其组成。

[0116] 转座子和载体组合物

转座系统

本公开内容提供了转座子或纳米转座子，其包含：第一核酸序列和第二核酸序列，所述第一核酸序列包含：(a)第一反向末端重复(ITR)或编码第一ITR的序列，(b)第二ITR或编码第二ITR的序列，以及(c) ITR内序列或编码ITR内的序列，其中所述ITR内序列包含转座子序列或编码转座子的序列；和所述第二核酸序列包含ITR间序列或编码ITR间的序列，其中所述ITR间序列的长度等于或小于700个核苷酸。

[0117] 本公开内容的转座子或纳米转座子包含蛋白质支架(例如，包含至少一种scFv、单结构域抗体或Centyrin的CAR)。转座子或纳米转座子可以为包含编码蛋白质支架的序列的质粒DNA转座子(例如，包含至少一种scFv、单结构域抗体或Centyrin的CAR)，侧翼为两个顺式调控绝缘子元件。转座子或纳米转座子可以进一步包括包含编码转座酶的序列的质粒。编码转座酶的序列可以为DNA序列或RNA序列。优选地，编码转座酶的序列是mRNA序列。

[0118] 本公开内容的转座子或纳米转座子可以为piggyBac™ (PB)转座子。在一些方面，当转座子是PB转座子时，转座酶是piggyBac™ (PB)转座酶、piggyBac样(PBL)转座酶或Super piggyBac™ (SPB)转座酶。优选地，编码SPB转座酶的序列是mRNA序列。

[0119] PB转座子以及PB、PBL和SPB转座酶的非限制性实例在以下中详细描述：美国专利号6,218,182；美国专利号6,962,810；美国专利号8,399,643和PCT公开号WO 2010/099296。

[0120] PB、PBL和SPB转座酶识别在转座子的端部上的转座子特异性反向末端重复序列(ITR)，并且在染色体位点(TTAT靶序列)内的序列5'-TTAT-3'处或在染色体位点(TTAA靶序列)内的序列5'-TTAA-3'处在ITR之间插入内容物。PB或PBL转座子的靶序列可以包含以下

或由以下组成:5'-CTAA-3'、5'-TTAG-3'、5'-ATAA-3'、5'-TCAA-3'、5'AGTT-3'、5'-ATTA-3'、5'-GTTA-3'、5'-TTGA-3'、5'-TTTA-3'、5'-TTAC-3'、5'-ACTA-3'、5'-AGGG-3'、5'-CTAG-3'、5'-TGAA-3'、5'-AGGT-3'、5'-ATCA-3'、5'-CTCC-3'、5'-TAAA-3'、5'-TCTC-3'、5'TGAA-3'、5'-AAAT-3'、5'-AATC-3'、5'-ACAA-3'、5'-ACAT-3'、5'-ACTC-3'、5'-AGTG-3'、5'-ATAG-3'、5'-CAAA-3'、5'-CACA-3'、5'-CATA-3'、5'-CCAG-3'、5'-CCCA-3'、5'-CGTA-3'、5'-GTCC-3'、5'-TAAG-3'、5'-TCTA-3'、5'-TGAG-3'、5'-TGTT-3'、5'-TTCA-3'5'-TTCT-3'和5'-TTTT-3'。PB或PBL转座子系统对于可以包括在ITR之间的目的基因没有有效载荷限制。

[0121] 一种或多种PB、PBL和SPB转座酶的示例性氨基酸序列公开于美国专利号6,218,185;美国专利号6,962,810和美国专利号8,399,643中。在一个优选的方面,PB转座酶包含氨基酸序列或由氨基酸序列组成,所述氨基酸序列与SEQ ID NO: 63至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100% (或两者之间的任何百分比)等同。

[0122] PB或PBL转座酶可以包含氨基酸序列或由氨基酸序列组成,所述氨基酸序列具有在SEQ ID NO: 63的序列的位置30、165、282或538中的2个或更多个、3个或更多个或者每一个处的氨基酸取代。转座酶可以为SPB转座酶,其包含SEQ ID NO: 63的序列的氨基酸序列或由其组成,其中在位置30处的氨基酸取代可以为缬氨酸(V)取代异亮氨酸(I),在位置165处的氨基酸取代可以为丝氨酸(S)取代甘氨酸(G),在位置282处的氨基酸取代可以为缬氨酸(V)取代甲硫氨酸(M)),并且在位置538处的氨基酸取代可以为赖氨酸(K)取代天冬酰胺(N)。在一个优选的方面,SPB转座酶包含氨基酸序列或由氨基酸序列组成,所述氨基酸序列与SEQ ID NO: 64至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100% (或两者之间的任何百分比)等同。

[0123] 在其中转座酶包含在位置30、165、282和/或538处的上述突变的某些方面,PB、PBL和SPB转座酶可以进一步包含在SEQ ID NO: 63或SEQ ID NO: 64的序列的位置3、46、82、103、119、125、177、180、185、187、200、207、209、226、235、240、241、243、258、296、298、311、315、319、327、328、340、421、436、456、470、486、503、552、570和591中的一个或多个位置处的氨基酸取代,在PCT公开号WO 2019/173636和PCT/US2019/049816中更详细地描述。

[0124] 如PCT公开号WO 2019/173636和PCT/US2019/049816中更详细地描述的,PB、PBL或SPB转座酶可以分离自或衍生自昆虫、脊椎动物、甲壳类动物或尾索动物。在优选的方面,PB、PBL或SPB转座酶分离自或衍生自昆虫粉纹夜蛾(*Trichoplusia ni*) (GenBank登录号AAA87375)或家蚕(*Bombyx mori*) (GenBank登录号BAD11135)。

[0125] 高活性PB或PBL转座酶是比它衍生自其的天然存在的变体更具活性的转座酶。在一个优选的方面,高活性PB或PBL转座酶分离自或衍生自家蚕或热带爪蟾(*Xenopus tropicalis*)。高活性PB或PBL转座酶的实例公开于美国专利号6,218,185;美国专利号6,962,810、美国专利号8,399,643和WO 2019/173636中。高活性氨基酸取代的列表公开于美国专利号10,041,077中。

[0126] 在一些方面,PB或PBL转座酶是整合缺陷型的。整合缺陷型PB或PBL转座酶是这样的转座酶,其可以切除其相应的转座子,但以低于相应的野生型转座酶的频率整合切除的转座子。整合缺陷型PB或PBL转座酶的实例公开于美国专利号6,218,185;美国专利号6,962,810、美国专利号8,399,643和WO 2019/173636中。整合缺陷型氨基酸取代的列表公开于美国专利号10,041,077中。

[0127] 在一些方面,PB或PBL转座酶与核定位信号融合。与核定位信号融合的PB或PBL转座酶的实例公开于美国专利号6,218,185;美国专利号6,962,810、美国专利号8,399,643和WO 2019/173636中。

[0128] 本公开内容的转座子或纳米转座子可以为睡美人(Sleeping Beauty)转座子。在一些方面,当转座子是睡美人转座子时,转座酶是睡美人转座酶(例如,如美国专利号9,228,180中公开的)或高活性睡美人(SB100X)转座酶。在一个优选的方面,睡美人转座酶包含氨基酸序列或由氨基酸序列组成,所述氨基酸序列与SEQ ID NO: 65至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100% (或两者之间的任何百分比)等同。在一个优选的方面,高活性睡美人(SB100X)转座酶包含氨基酸序列或由氨基酸序列组成,所述氨基酸序列与SEQ ID NO: 66至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100% (或两者之间的任何百分比)等同。

[0129] 本公开内容的转座子或纳米转座子可以为Helraiser转座子。示例性的Helraiser转座子包括Helibat1,其包含核酸序列或由核酸序列组成,所述核酸序列与SEQ ID NO: 67至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100% (或两者之间的任何百分比)等同。在一些方面,当转座子是Helraiser转座子时,转座酶是Helitron转座酶(例如,如WO 2019/173636中公开的)。在一个优选的方面,Helitron转座酶包含氨基酸序列或由氨基酸序列组成,所述氨基酸序列与SEQ ID NO: 68至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100%(或两者之间的任何百分比)等同。

[0130] 本公开内容的转座子或纳米转座子可以为Tol2转座子。示例性Tol2转座子,包括反向重复、亚末端序列和Tol2转座酶,包含核酸序列或由核酸序列组成,所述核酸序列与SEQ ID NO: 69至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100% (或两者之间的任何百分比)等同。在一些方面,当转座子是Tol2转座子时,转座酶是Tol2转座酶(例如,如WO 2019/173636中公开的)。在一个优选的方面,Tol2转座酶包含氨基酸序列或由氨基酸序列组成,所述氨基酸序列与SEQ ID NO: 70至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100% (或两者之间的任何百分比)等同。

[0131] 本公开内容的转座子或纳米转座子可以为TcBuster转座子。在一些方面,当转座子是TcBuster转座子时,转座酶是TcBuster转座酶或高活性TcBuster转座酶(例如,如WO 2019/173636中公开的)。TcBuster转座酶可以包含天然存在的氨基酸序列或非天然存在的氨基酸序列,或者由其组成。在一个优选的方面,TcBuster转座酶包含氨基酸序列或由氨基酸序列组成,所述氨基酸序列与SEQ ID NO: 71至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100% (或两者之间的任何百分比)等同。编码TcBuster转座酶的多核苷酸可以包含天然存在的核酸序列或非天然存在的核酸序列,或者由其组成。在一个优选的方面,TcBuster转座酶由多核苷酸编码,所述多核苷酸包含核酸序列或由核酸序列组成,所述核酸序列与SEQ ID NO: 72至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100% (或两者之间的任何百分比)等同。

[0132] 在一些方面,当与如PCT公开号WO 2019/173636和PCT/US2019/049816中更详细地描述的野生型TcBuster转座酶相比时,突变型TcBuster转座酶包含一个或多个序列变异。

[0133] 本文公开的细胞递送组合物(例如,转座子)可以包含编码治疗性蛋白质或治疗剂的核酸。治疗性蛋白质的实例包括在PCT公开号WO 2019/173636和PCT/US2019/049816中公

开的那些治疗性蛋白质。

[0134] 载体系统

在一些方面，本公开内容的组合物(例如，纳米转座子)可以与另一种转座子或纳米转座子或者载体组合利用。本公开内容的载体可以为病毒载体或重组载体。病毒载体可以包含分离自或衍生自逆转录病毒、慢病毒、腺病毒、腺伴随病毒或其任何组合的序列。病毒载体可以包含分离自或衍生自腺伴随病毒(AAV)的序列。病毒载体可以包含重组AAV(rAAV)。示例性的腺伴随病毒和重组腺伴随病毒包含紧靠编码本公开内容的scFv或CAR的序列顺式定位的、两个或更多个反向末端重复(ITR)序列。示例性腺伴随病毒和重组腺伴随病毒包括但不限于所有血清型(例如AAV1、AAV2、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8和AAV9)。示例性腺伴随病毒和重组腺伴随病毒包括但不限于自互补的AAV (scAAV)和AAV杂合体，其含有一种血清型的基因组和另一种血清型的衣壳(例如，AAV2/5、AAV-DJ和AAV-DJ8)。示例性腺伴随病毒和重组腺伴随病毒包括但不限于rAAV-LK03。

[0135] 本公开内容的载体可以为纳米颗粒。纳米颗粒载体的非限制性实例包括核酸(例如，RNA、DNA、合成核苷酸、修饰的核苷酸或其任何组合)、氨基酸(L-氨基酸、D-氨基酸、合成氨基酸、修饰的氨基酸或其任何组合)、聚合物(例如，聚合物囊泡(polymersome))、胶束、脂质(例如，脂质体)、有机分子(例如，碳原子、片层、纤维、管)、无机分子(例如，磷酸钙或金)或其任何组合。纳米颗粒载体可以被动或主动转运跨越细胞膜。

[0136] 本文公开的细胞递送组合物(例如，转座子、载体)可以包含编码治疗性蛋白质或治疗剂的核酸。治疗性蛋白质的实例包括在PCT公开号WO 2019/173636和PCT/US2019/049816中公开的那些治疗性蛋白质。

[0137] 本公开内容的细胞和修饰的细胞

本公开内容的细胞和修饰的细胞可以为哺乳动物细胞。优选地，细胞和修饰的细胞是人细胞。本公开内容的细胞和修饰的细胞可以为免疫细胞。本公开内容的免疫细胞可以包含淋巴样祖细胞、天然杀伤(NK)细胞、T淋巴细胞(T细胞)、干性记忆T细胞(T_{SCM}细胞)、中枢记忆T细胞(T_{CM})、干细胞样T细胞、B淋巴细胞(B细胞)、抗原呈递细胞(APC)、细胞因子诱导的杀伤(CIK)细胞、髓样祖细胞、嗜中性粒细胞、嗜碱性粒细胞、嗜酸性粒细胞、单核细胞、巨噬细胞、血小板、红细胞、红血细胞(RBC)、巨核细胞或破骨细胞。

[0138] 免疫前体细胞可以包含可以分化成一种或多种类型的免疫细胞的任何细胞。免疫前体细胞可以包含可以自我更新并发育成免疫细胞的多能干细胞。免疫前体细胞可以包含造血干细胞(HSC)或其后代。免疫前体细胞可以包含可以发育成免疫细胞的前体细胞。免疫前体细胞可以包含造血祖细胞(HPC)。

[0139] 造血干细胞(HSC)是多潜能、自我更新的细胞。来自淋巴样和髓样谱系的所有分化的血细胞都起于HSC。HSC可以在成人骨髓、外周血、动员的外周血、腹膜透析流出液和脐带血中找到。

[0140] HSC可以分离自或衍生自原代或培养的干细胞。HSC可以分离自或衍生自胚胎干细胞、多潜能干细胞、多能干细胞、成体干细胞或诱导多能干细胞(iPSC)。

[0141] 免疫前体细胞可以包含HSC或HSC后代细胞。HSC后代细胞的非限制性实例包括多潜能干细胞、淋巴样祖细胞、天然杀伤(NK)细胞、T淋巴细胞(T细胞)、B淋巴细胞(B细胞)、髓样祖细胞、嗜中性粒细胞、嗜碱性粒细胞、嗜酸性粒细胞、单核细胞和巨噬细胞。

[0142] 通过所公开的方法产生的HSC可以保留“原始”干细胞的特征，所述干细胞在分离自或衍生自成体干细胞并且定型至单一谱系时，共享胚胎干细胞的特性。例如，通过所公开的方法产生的“原始”HSC在分裂之后保留其“干性”，并且不分化。因而，作为过继细胞疗法，通过所公开的方法产生的“原始”HSC不仅补充其数目，而且在体内扩增。当作为单一剂量施用时，通过所公开的方法产生的“原始”HSC可以是治疗上有效的。

[0143] 原始HSC可以为CD34+。原始HSC可以为CD34+和CD38-。原始HSC可以为CD34+、CD38-和CD90+。原始HSC可以为CD34+、CD38-、CD90+和CD45RA-。原始HSC可以为CD34+、CD38-、CD90+、CD45RA-和CD49f+。原始HSC可以为CD34+、CD38-、CD90+、CD45RA-和CD49f+。

[0144] 原始HSC、HSC和/或HSC后代细胞可以根据所公开的方法进行修饰，以表达外源序列(例如，嵌合抗原受体或治疗性蛋白质)。修饰的原始HSC、修饰的HSC和/或修饰的HSC后代细胞可以向前分化，以产生修饰的免疫细胞，包括但不限于修饰的T细胞、修饰的天然杀伤细胞和/或修饰的B细胞。

[0145] 修饰的免疫或免疫前体细胞可以为NK细胞。NK细胞可以为从淋巴样祖细胞分化的细胞毒性淋巴细胞。修饰的NK细胞可以衍生自修饰的造血干细胞和祖细胞(HSPC)或修饰的HSC。在一些方面，未活化的NK细胞衍生自CD3耗竭的白细胞去除术(leukapheresis)（含有CD14/CD19/CD56+细胞)。

[0146] 修饰的免疫或免疫前体细胞可以为B细胞。B细胞是在细胞表面上表达B细胞受体的一类淋巴细胞。B细胞受体与特异性抗原结合。修饰的B细胞可以衍生自修饰的造血干细胞和祖细胞(HSPC)或修饰的HSC。

[0147] 本公开内容的修饰的T细胞可以衍生自修饰的造血干细胞和祖细胞(HSPC)或修饰的HSC。与传统的生物制品和化学治疗剂不同，所公开的修饰的T细胞能够在抗原识别后快速繁殖，从而潜在地避免了重复治疗的需要。为了实现这一点，在一些方面，修饰的T细胞不仅驱动初始应答，而且还作为稳定的活记忆T细胞群体持续存在于患者中，以预防潜在的复发。可替代地，在一些方面，当不需要时，修饰的T细胞并不持续存在于患者中。

[0148] 大量努力已集中于并不通过抗原非依赖性(强直)信号传导导致T细胞耗尽的抗原受体分子的开发，以及含有早期记忆T细胞，尤其是干细胞记忆(T_{SCM})或干细胞样T细胞的修饰T细胞产物的开发。本公开内容的干细胞样修饰的T细胞显示出最大的自我更新能力和多潜能能力，以衍生中枢记忆(T_{CM}) T细胞或T_{CM}样细胞、效应记忆(T_{EM})和效应T细胞(T_E)，从而产生更好的肿瘤根除和长期修饰的T细胞移植物植入。分化的线性途径可能负责生成这些细胞：幼稚T细胞(T_N) > T_{SCM} > T_{CM} > T_{EM} > T_E > T_{TE}，由此T_N是直接产生T_{SCM}的亲代前体细胞，然后，其依次又直接产生T_{CM}等。本公开内容的T细胞的组合物可以包含每个亲代T细胞子集中的一个或多个，其中T_{SCM}细胞是最丰富的(例如，T_{SCM} > T_{CM} > T_{EM} > T_E > T_{TE})。

[0149] 免疫细胞前体可以分化成或能够分化成早期记忆T细胞、干细胞如T细胞、幼稚T细胞(T_N)、T_{SCM}、T_{CM}、T_{EM}、T_E或T_{TE}。免疫细胞前体可以为本公开内容的原始HSC、HSC或HSC后代细胞。免疫细胞可以为早期记忆T细胞、干细胞如T细胞、幼稚T细胞(T_N)、T_{SCM}、T_{CM}、T_{EM}、T_E或T_{TE}。

[0150] 本公开内容的方法可以修饰和/或产生修饰的T细胞群体，其中群体中至少2%、5%、10%、15%、20%、25%、30%、35%、40%、45%、50%、60%、65%、70%、75%、80%、85%、90%、95%、99%或两者之间的任何百分比的多个修饰的T细胞，表达早期记忆T细胞的一种或多种细胞表面标记物。修饰的早期记忆T细胞群体包含多个修饰的干细胞样T细胞。修饰的早期记忆T细胞群体

包含多个修饰的T_{SCM}细胞。修饰的早期记忆T细胞群体包含多个修饰的T_{CM}细胞。

[0151] 本公开内容的方法可以修饰和/或产生修饰的T细胞群体，其中群体中至少2%、5%、10%、15%、20%、25%、30%、35%、40%、45%、50%、60%、65%、70%、75%、80%、85%、90%、95%、99%或两者之间的任何百分比的多个修饰的T细胞，表达干细胞样T细胞的一种或多种细胞表面标记物。修饰的干细胞样T细胞群体包含多个修饰的T_{SCM}细胞。修饰的干细胞样T细胞群体包含多个修饰的T_{CM}细胞。

[0152] 在一些方面，群体中至少5%、至少10%、至少15%、至少20%、至少25%、至少30%、至少35%、至少40%、至少45%、至少50%、至少55%、至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%、至少99%或100%或两者之间的任何百分比的多个修饰的T细胞，表达干性记忆T细胞(T_{SCM})或T_{SCM}样细胞的一种或多种细胞表面标记物；并且其中一种或多种细胞表面标记物包含CD45RA和CD62L。细胞表面标记物可以包含CD62L、CD45RA、CD28、CCR7、CD127、CD45RO、CD95、CD95和IL-2Rβ中的一种或多种。细胞表面标记物可以包含CD45RA、CD95、IL-2Rβ、CCR7和CD62L中的一种或多种。

[0153] 在一些方面，群体中至少5%、至少10%、至少15%、至少20%、至少25%、至少30%、至少35%、至少40%、至少45%、至少50%、至少55%、至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%、至少99%或100%的多个修饰的T细胞，表达中枢记忆T细胞(T_{CM})或T_{CM}样细胞的一种或多种细胞表面标记物；并且其中一种或多种细胞表面标记物包含CD45RO和CD62L。细胞表面标记物可以包含CD45RO、CD95、IL-2Rβ、CCR7和CD62L中的一种或多种。

[0154] 本公开内容的方法可以修饰和/或产生修饰的T细胞群体，其中群体中至少2%、5%、10%、15%、20%、25%、30%、35%、40%、45%、50%、60%、65%、70%、75%、80%、85%、90%、95%、99%或两者之间的任何百分比的多个修饰的T细胞，表达幼稚T细胞(T_N)的一种或多种细胞表面标记物。细胞表面标记物可以包含CD45RA、CCR7和CD62L中的一种或多种。

[0155] 本公开内容的方法可以修饰和/或产生修饰的T细胞群体，其中群体中至少2%、5%、10%、15%、20%、25%、30%、35%、40%、45%、50%、60%、65%、70%、75%、80%、85%、90%、95%、99%或两者之间的任何百分比的多个修饰的T细胞，表达效应T细胞(修饰的T_{EFF})的一种或多种细胞表面标记物。细胞表面标记物可以包含CD45RA、CD95和IL-2Rβ中的一种或多种。

[0156] 本公开内容的方法可以修饰和/或产生修饰的T细胞群体，其中群体中至少2%、5%、10%、15%、20%、25%、30%、35%、40%、45%、50%、60%、65%、70%、75%、80%、85%、90%、95%、99%或两者之间的任何百分比的多个修饰的T细胞，表达干细胞样T细胞、干性记忆T细胞(T_{SCM})或中枢记忆T细胞(T_{CM})的一种或多种细胞表面标记物。

[0157] 群体中多个修饰的细胞包含转基因或编码转基因(例如CAR)的序列，其中群体中至少75%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%、至少99.5%、至少99.9%或100%的多个细胞包含转基因或编码转基因的序列，其中至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%、至少99.5%、至少99.9%或100%的修饰的细胞群体，表达包含CD34的一种或多种细胞表面标记物，或者其中至少约70%至约99%、约75%至约95%、或约85%至约95%的修饰的细胞群体，表达包含CD34的一种或多种细胞表面标记物(例如，包含细胞表面标记物表型CD34+)。

[0158] 群体中多个修饰的细胞包含转基因或编码转基因(例如CAR)的序列,其中群体中至少75%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%、至少99.5%、至少99.9%或100%的多个细胞包含转基因或编码转基因的序列,其中至少40%、至少45%、至少50%、至少55%、至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%、至少99.5%、至少99.9%或100%的修饰的细胞群体,表达包含CD34的一种或多种细胞表面标记物,并且不表达包含CD38的一种或多种细胞表面标记物,或者其中至少约45%至约90%、约50%至约80%、或约65%至约75%的修饰的细胞群体,表达包含CD34的一种或多种细胞表面标记物,并且不表达包含CD38的一种或多种细胞表面标记物(例如,包含细胞表面标记物表型CD34+和CD38-)。

[0159] 群体中多个修饰的细胞包含转基因或编码转基因(例如CAR)的序列,其中群体中至少75%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%、至少99.5%、至少99.9%或100%的多个细胞包含转基因或编码转基因的序列,其中至少0.1%、至少0.2%、至少0.3%、至少0.4%、至少0.5%、至少0.6%、至少0.7%、至少0.8%、至少0.9%、至少1%、至少1.5%、至少2%、至少3%、至少4%、至少5%、至少10%、至少15%、至少20%、至少25%、至少30%、至少35%、至少40%、至少45%、至少50%、至少55%、至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%、至少99.5%、至少99.9%或100%的修饰的细胞群体,表达包含CD34和CD90的一种或多种细胞表面标记物,并且不表达包含CD38的一种或多种细胞表面标记物,或者其中至少约0.2%至约40%、约0.2%至约30%、约0.2%至约2%、或0.5%至约1.5%的修饰的细胞群体,表达包含CD34和CD90的一种或多种细胞表面标记物,并且不表达包含CD38的一种或多种细胞表面标记物(例如,包含细胞表面标记物表型CD34+、CD38-和CD90+)。

[0160] 群体中多个修饰的细胞包含转基因或编码转基因(例如CAR)的序列,其中群体中至少75%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%、至少99.5%、至少99.9%或100%的多个细胞包含转基因或编码转基因的序列,其中至少0.1%、至少0.2%、至少0.3%、至少0.4%、至少0.5%、至少0.6%、至少0.7%、至少0.8%、至少0.9%、至少1%、至少1.5%、至少2%、至少3%、至少4%、至少5%、至少10%、至少15%、至少20%、至少25%、至少30%、至少35%、至少40%、至少45%、至少50%、至少55%、至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%、至少99.5%、至少99.9%或100%的修饰的细胞群体,表达包含CD34和CD90的一种或多种细胞表面标记物,并且不表达包含CD38和CD45RA的一种或多种细胞表面标记物,或者其中至少约0.2%至约40%、约0.2%至约30%、约0.2%至约2%、或0.5%至约1.5%的修饰的细胞群体,表达包含CD34和CD90的一种或多种细胞表面标记物,并且不表达包含CD38和CD45RA的一种或多种细胞表面标记物(例如,包含细胞表面标记物表型CD34+、CD38-、CD90+、CD45RA)。

[0161] 群体中多个修饰的细胞包含转基因或编码转基因(例如CAR)的序列,其中群体中至少75%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%、至少99.5%、至少99.9%或100%的多个细胞包含转基因或编码转基因

的序列，其中至少0.01%、至少0.02%、至少0.03%、至少0.04%、至少0.05%、至少0.06%、至少0.07%、至少0.08%、至少0.09%、至少0.1%、至少0.2%、至少0.3%、至少0.4%、至少0.5%、至少0.6%、至少0.7%、至少0.8%、至少0.9%、至少1%、至少1.5%、至少2%、至少3%、至少4%、至少5%、至少10%、至少15%、至少20%、至少25%、至少30%、至少35%、至少40%、至少45%、至少50%、至少55%、至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%、至少99.5%、至少99.9%或100%的修饰的细胞群体，表达包含CD34、CD90和CD49f的一种或多种细胞表面标记物，并且不表达包含CD38和CD45RA的一种或多种细胞表面标记物，或者其中至少约0.02%至约30%、约0.02%至约2%、约0.04%至约2%、或约0.04%至约1%的修饰的细胞群体，表达包含CD34、CD90和CD49f的一种或多种细胞表面标记物，并且不表达包含CD38和CD45RA的一种或多种细胞表面标记物(例如，包含细胞表面标记物表型CD34+、CD38-、CD90+、CD45RA和CD49f+)。

[0162] 群体中多个修饰的细胞包含转基因或编码转基因(例如CAR)的序列，其中群体中至少75%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%、至少99.5%、至少99.9%或100%的多个细胞包含转基因或编码转基因的序列，其中至少0.01%、至少0.02%、至少0.03%、至少0.04%、至少0.05%、至少0.06%、至少0.07%、至少0.08%、至少0.09%、至少0.1%、至少0.2%、至少0.3%、至少0.4%、至少0.5%、至少0.6%、至少0.7%、至少0.8%、至少0.9%、至少1%、至少1.5%、至少2%、至少3%、至少4%、至少5%、至少10%、至少15%、至少20%、至少25%、至少30%、至少35%、至少40%、至少45%、至少50%、至少55%、至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%、至少99.5%、至少99.9%或100%的修饰的细胞群体，表达包含CD34和CD90的一种或多种细胞表面标记物，并且不表达包含CD45RA的一种或多种细胞表面标记物，或者其中至少约0.2%至约5%、约0.2%至约3%、或约0.4%至约3%的修饰的细胞群体，表达包含CD34和CD90的一种或多种细胞表面标记物，并且不表达包含CD45RA的一种或多种细胞表面标记物(例如，包含细胞表面标记物表型CD34+、CD90+和CD45RA-)。

[0163] 产生和/或扩增免疫细胞或免疫前体细胞(例如，所公开的修饰的T细胞)的组合物和方法、以及用于维持或增强免疫细胞或免疫前体细胞(例如，所公开的修饰的T细胞)的细胞活力水平和/或干样表型的缓冲液，在本文其它地方公开，并且在美国专利号10,329,543和PCT公开号WO 2019/173636中更详细地公开。

[0164] 本公开内容的细胞和修饰的细胞可以为体细胞。本公开内容的细胞和修饰的细胞可以为分化的细胞。本公开内容的细胞和修饰的细胞可以为自体细胞或同种异体细胞。对同种异体细胞进行改造，以预防在施用于受试者之后对移植物植入的不良反应。同种异体细胞可以为任何类型的细胞。同种异体细胞可以为干细胞或可以衍生自干细胞。同种异体细胞可以为分化的体细胞。

[0165] 表达嵌合抗原受体的方法

本公开内容提供了在细胞的表面上表达CAR的方法。该方法包括(a)获得细胞群体；(b)在足以将CAR转移跨越细胞群体中的至少一种细胞的细胞膜的条件下，使细胞群体与包含CAR或编码CAR的序列的组合物接触，从而生成修饰的细胞群体；(c)在适合于整合编码CAR的序列的条件下，培养修饰的细胞群体；(d)从在细胞表面上表达CAR的修饰的细胞群

体扩增和/或选择至少一种细胞。

[0166] 在一些方面，细胞群体可以包含白细胞和/或CD4+和CD8+白细胞。细胞群体可以包含以优化比率的CD4+和CD8+白细胞。CD4+与CD8+白细胞的优化比率并不在体内天然发生。细胞群体可以包含肿瘤细胞。

[0167] 在一些方面，足以将CAR或编码CAR的序列、转座子或载体转移跨越细胞群体中的至少一种细胞的细胞膜的条件，包括在特定电压下施加一种或多种电脉冲、缓冲液和一种或多种补充因子中的至少一种。在一些方面，适合于整合编码CAR的序列的条件包括缓冲液和一种或多种补充因子中的至少一种。

[0168] 缓冲液可以包含PBS、HBSS、OptiMEM、BTXpress、Amaxa Nucleofector、人T细胞核转染缓冲液或其任何组合。一种或多种补充因子可以包含(a)重组人细胞因子、趋化因子、白介素或其任何组合；(b)盐、矿物质、代谢物或其任何组合；(c)细胞培养基；(d)细胞DNA传感、代谢、分化、信号转导、一种或多种细胞凋亡途径或其组合的抑制剂；以及(e)修饰或稳定一种或多种核酸的试剂。重组人细胞因子、趋化因子、白介素或其任何组合可以包含IL2、IL7、IL12、IL15、IL21、IL1、IL3、IL4、IL5、IL6、IL8、CXCL8、IL9、IL10、IL11、IL13、IL14、IL16、IL17、IL18、IL19、IL20、IL22、IL23、IL25、IL26、IL27、IL28、IL29、IL30、IL31、IL32、IL33、IL35、IL36、GM-CSF、IFN-γ、IL-1 α/IL-1F1、IL-1 β/IL-1F2、IL-12 p70、IL-12/IL-35 p35、IL-13、IL-17/IL-17A、IL-17A/F异源二聚体、IL-17F、IL-18/IL-1F4、IL-23、IL-24、IL-32、IL-32 β、IL-32 γ、IL-33、LAP (TGF-β 1)、淋巴毒素-α/TNF-β、TGF-β、TNF-α、TRANCE/TNFSF11/RANK L或其任何组合。盐、矿物质、代谢物或其任何组合可以包含HEPES、烟酰胺、肝素、丙酮酸钠、L-谷氨酰胺、MEM非必需氨基酸溶液、抗坏血酸、核苷、FBS/FCS、人血清、血清替代物、抗生素、pH调节剂、Earle's盐、2-巯基乙醇、人转铁蛋白、重组人胰岛素、人血清白蛋白、Nucleofector PLUS Supplement、KCL、$MgCl_2$、Na_2HPO_4、NAH_2PO_4、乳糖酸钠、甘露糖醇、琥珀酸钠、氯化钠、CINa、葡萄糖、$Ca(NO_3)_2$、Tris/HCl、K_2HPO_4、KH_2PO_4、聚乙烯亚胺、聚乙二醇、泊洛沙姆188、泊洛沙姆181、泊洛沙姆407、聚乙烯吡咯烷酮、Pop313、Crown-5或其任何组合。细胞培养基可以包含PBS、HBSS、OptiMEM、DMEM、RPMI 1640、AIM-V、X-VIVO 15、CellGro DC培养基、CTS OpTimizer T细胞扩增SFM、TexMACS培养基、PRIME-XV T细胞扩增培养基、ImmunoCult-XF T细胞扩增培养基或其任何组合。细胞DNA传感、代谢、分化、信号转导、一种或多种细胞凋亡途径或其组合的抑制剂包含以下的抑制剂：TLR9、MyD88、IRAK、TRAF6、TRAF3、IRF-7、NF-KB、1型干扰素、促炎细胞因子、cGAS、STING、Sec5、TBK1、IRF-3、RNA pol III、RIG-1、IPS-1、FADD、RIP1、TRAF3、AIM2、ASC、半胱天冬酶1、Pro-IL1B、PI3K、Akt、Wnt3A、糖原合酶激酶-3β (GSK-3 β)抑制剂(例如TWS119)或其任何组合。此类抑制剂的实例可以包含巴弗洛霉素、氯喹、奎纳克林、AC-YVAD-CMK、Z-VAD-FMK、Z-IETD-FMK或其任何组合。修饰或稳定一种或多种核酸的试剂包含pH调节剂、DNA结合蛋白、脂质、磷脂、CaPO4、具有或不具有NLS序列的净中性电荷DNA结合肽、TREX1酶或其任何组合。

[0169] 扩增和选择步骤可以并发或序贯地发生。扩增可以在选择之前发生。扩增可以在选择之后发生，并且任选地，进一步(即第二)选择可以在扩增之后发生。并发扩增和选择可以是同时的。扩增和/或选择步骤可以进行10至14天的时期，包括端点在内。

[0170] 扩增可以包含使修饰的细胞群体的至少一种细胞与抗原接触，以通过CAR刺激至少一种细胞，从而生成扩增的细胞群体。抗原可以在基底的表面上呈递。基底可以具有任何

形式，包括但不限于表面、孔、珠或多个珠以及基质。基底可以进一步包含顺磁性或磁性组分。抗原可以在基底的表面上呈递，其中所述基底是磁珠，并且其中磁体可以用于从修饰且扩增的细胞群体中去除或分离磁珠。抗原可以在细胞或人工抗原呈递细胞的表面上呈递。人工抗原呈递细胞可以包括但不限于肿瘤细胞和干细胞。

[0171] 在其中转座子或载体包含选择基因的一些方面，选择步骤包括使修饰的细胞群体中的至少一种细胞与选择基因对其赋予抗性的化合物接触，从而将表达选择基因的细胞鉴定为从选择中存活，并且将未能表达选择基因的细胞鉴定为未能从选择步骤中存活。

[0172] 本公开内容提供了组合物，其包含本文所述的方法的修饰、扩增且选择的细胞群体。

[0173] 用于在细胞的表面上表达CAR的方法的更详细描述公开于PCT公开号WO 2019/049816和PCT/US2019/049816中。

[0174] 本公开内容提供了细胞或细胞群体，其中所述细胞构成组合物，所述组合物包含(a)诱导型转基因构建体，其包含编码诱导型启动子的序列和编码转基因的序列，以及(b)受体构建体，其包含编码组成型启动子的序列和编码外源受体例如CAR的序列，其中在(a)的构建体和(b)的构建体整合到细胞的基因组序列内后，表达外源受体，并且其中外源受体在结合配体或抗原后转导细胞内信号，所述细胞内信号直接或间接靶向调控诱导型转基因(a)的表达的诱导型启动子，以修饰基因表达。

[0175] 组合物可以通过降低基因表达来修饰基因表达。组合物可以通过瞬时修饰基因表达(例如，对于配体与外源受体结合的持续时间)来修饰基因表达。组合物可以剧烈地修饰基因表达(例如，配体与外源受体可逆地结合)。组合物可以长期修饰基因表达(例如，配体与外源受体不可逆地结合)。

[0176] 外源受体可以包含相对于细胞的基因组序列的内源受体。示例性受体包括但不限于细胞内受体、细胞表面受体、跨膜受体、配体门控离子通道和G蛋白偶联受体。

[0177] 外源受体可以包含非天然存在的受体。非天然存在的受体可以为合成、修饰、重组、突变型或嵌合受体。非天然存在的受体可以包含分离自或衍生自T细胞受体(TCR)的一种或多种序列。非天然存在的受体可以包含分离自或衍生自支架蛋白的一种或多种序列。在一些方面，包括其中非天然存在的受体不包含跨膜结构域的那些方面，非天然存在的受体与第二跨膜、膜结合和/或细胞内受体相互作用，所述第二跨膜、膜结合和/或细胞内受体在与非天然存在的受体接触之后，转导细胞内信号。非天然存在的受体可以包含跨膜结构域。非天然存在的受体可以与转导细胞内信号的细胞内受体相互作用。非天然存在的受体可以包含细胞内信号传导结构域。非天然存在的受体可以为嵌合配体受体(CLR)。CLR可以为嵌合抗原受体(CAR)。

[0178] 编码诱导型启动子的序列包含编码NFκB启动子的序列、编码干扰素(IFN)启动子的序列或编码白介素-2启动子的序列。在一些方面，IFN启动子是IFNγ启动子。诱导型启动子可以分离自或衍生自细胞因子或趋化因子的启动子。细胞因子或趋化因子可以包含IL2、IL3、IL4、IL5、IL6、IL10、IL12、IL13、IL17A/F、IL21、IL22、IL23、转化生长因子β (TGFβ)、集落刺激因子2 (GM-CSF)、干扰素γ (IFNγ)、肿瘤坏死因子α (TNFα)、LTα、穿孔素、颗粒酶C(Gzmc)、颗粒酶B (Gzmb)、C-C基序趋化因子配体5 (CCL5)、C-C基序趋化因子配体4(Cc14)、C-C基序趋化因子配体3 (Cc13)、X-C基序趋化因子配体1 (Xc11)或LIF白介素6家

族细胞因子(Lif)。

[0179] 诱导型启动子可以分离自或衍生自包含表面蛋白质的基因的启动子,所述表面蛋白质涉及细胞分化、活化、耗尽和功能。在一些方面,基因包含CD69、CD71、CTLA4、PD-1、TIGIT、LAG3、TIM-3、GITR、MHCII、COX-2、FASL或4-1BB。

[0180] 诱导型启动子可以分离自或衍生自涉及CD代谢和分化的基因的启动子。诱导型启动子可以分离自或衍生自Nr4a1、Nr4a3、Tnfrsf9 (4-1BB)、Sema7a、Zfp36l2、Gadd45b、Dusp5、Dusp6和Neto2的启动子。

[0181] 在一些方面,诱导型转基因构建体包含或驱动抑制性检查点信号下游的信号传导组分、转录因子、细胞因子或细胞因子受体、趋化因子或趋化因子受体、细胞死亡或细胞凋亡受体/配体、代谢传感分子、赋予对癌症疗法的敏感性的蛋白质、以及癌基因或肿瘤抑制基因的表达。其非限制性实例公开于PCT公开号WO 2019/173636和PCT申请号PCT/US2019/049816中。

[0182] 武装的细胞(Armored Cells)

本公开内容的修饰的细胞(例如,CAR T细胞)可以进一步进行修饰,以增强其治疗潜力。可替代地或另外,修饰的细胞可以进一步进行修饰,以致使它们对免疫学和/或代谢检查点较不敏感。这种类型的修饰“武装”细胞,在修饰之后,所述细胞在此处可以被称为“武装的”细胞(例如,武装的T细胞)。可以通过例如在肿瘤免疫抑制微环境内,阻断和/或稀释天然递送至细胞的特异性检查点信号(例如检查点抑制),来产生武装的细胞。

[0183] 本公开内容的武装的细胞可以衍生自任何细胞,例如,T细胞、NK细胞、造血祖细胞、外周血(PB)衍生的T细胞(包括分离自或衍生自G-CSF动员的外周血的T细胞)、或脐带血(UCB)衍生的T细胞。武装的细胞(例如,武装的T细胞)可以包含以下中的一种或多种:嵌合配体受体(包含蛋白质支架、抗体、ScFv或抗体模拟物的CLR)/嵌合抗原受体(包含蛋白质支架、抗体、ScFv或抗体模拟物的CAR)、CARTyrin (包含Centyrin的CAR)和/或VCAR (包含骆驼科动物VHH或单结构域VH的CAR)。武装的细胞(例如,武装的T细胞)可以包含如本文公开的诱导型促凋亡多肽。武装的细胞(例如,武装的T细胞)可以包含外源序列。外源序列可以包含编码治疗性蛋白质的序列。示例性治疗性蛋白质可以为核、细胞质、细胞内、跨膜、细胞表面结合或分泌的蛋白质。由武装的细胞(例如,武装的T细胞)表达的示例性治疗性蛋白质,可以修饰武装的细胞的活性或可以修饰第二细胞的活性。武装的细胞(例如,武装的T细胞)可以包含选择基因或选择标记物。武装的细胞(例如,武装的T细胞)可以包含合成基因表达盒(在本文中也被称为诱导型转基因构建体)。

[0184] 本公开内容的修饰的细胞(例如,CAR T细胞)可以进一步进行修饰,以沉默或减少编码抑制性检查点信号的受体的一种或多种基因的表达,以产生武装的细胞(例如,武装的CAR T细胞)。抑制性检查点信号的受体在细胞表面上或在细胞的细胞质内表达。沉默或减少编码抑制性检查点信号的受体的基因的表达,导致在武装的细胞的表面上或在细胞质内抑制性检查点受体的蛋白质表达的丧失。因此,具有编码抑制性检查点受体的一种或多种基因的沉默或减少表达的武装的细胞对检查点信号是抗性的、不接受的或不敏感的。在这些抑制性检查点信号的存在下,武装的细胞对抑制性检查点信号的抗性或敏感性降低增强了武装的细胞的治疗潜力。抑制性检查点信号(和诱导免疫压制的蛋白质)的非限制性实例公开于PCT公开号WO 2019/173636中。可以被沉默的抑制性检查点信号的优选实例包括但

不限于PD-1和TGFβRII。

[0185] 本公开内容的修饰的细胞(例如,CAR T细胞)可以进一步进行修饰,以沉默或减少编码涉及检查点信号传导的细胞内蛋白质的一种或多种基因的表达,以产生武装的细胞(例如,武装的CAR T细胞)。可以通过靶向涉及检查点信号传导途径的任何细胞内信号传导蛋白,来增强修饰的细胞的活性,从而实现对一个或多个检查点途径的检查点抑制或干扰。涉及检查点信号传导的细胞内信号传导蛋白的非限制性实例公开于PCT公开号WO 2019/173636中。

[0186] 本公开内容的修饰的细胞(例如,CAR T细胞)可以进一步进行修饰,以沉默或减少编码阻碍疗法功效的转录因子的一种或多种基因的表达,以产生武装的细胞(例如,武装的CAR T细胞)。可以通过沉默或减少阻碍疗法功效的转录因子的表达(或阻遏功能),来增强或调节修饰的细胞的活性。可以进行修饰以沉默或减少表达或阻遏其功能的转录因子的非限制性实例,包括但不限于PCT公开号WO 2019/173636中公开的示例性转录因子。

[0187] 本公开内容的修饰的细胞(例如,CAR T细胞)可以进一步进行修饰,以沉默或减少编码细胞死亡或细胞凋亡受体的一种或多种基因的表达,以产生武装的细胞(例如,武装的CAR T细胞)。死亡受体与其内源性配体的相互作用导致细胞凋亡的开始。细胞死亡和/或细胞凋亡受体和/或配体的表达、活性或相互作用的破坏致使修饰的细胞对死亡信号较不接受,因而,使得武装的细胞在肿瘤环境中更有效。细胞死亡和/或细胞凋亡受体和配体的非限制性实例公开于PCT公开号WO 2019/173636中。可以进行修饰的细胞死亡受体的优选实例是Fas (CD95)。

[0188] 本公开内容的修饰的细胞(例如,CAR T细胞)可以进一步进行修饰,以沉默或减少编码代谢传感蛋白的一种或多种基因的表达,以产生武装的细胞(例如,武装的CAR T细胞)。通过修饰的细胞破坏免疫压制性肿瘤微环境(特征在于低水平的氧、pH、葡萄糖和其它分子)的代谢传感,导致T细胞功能的延长保留,并且因而,每个细胞杀死更多的肿瘤细胞。代谢传感基因和蛋白质的非限制性实例公开于PCT公开号WO 2019/173636中。优选的实例,HIF1a和VHL在缺氧环境中发挥T细胞功能的作用。武装的T细胞可能具有编码HIF1a或VHL的一种或多种基因的沉默或减少表达。

[0189] 本公开内容的修饰的细胞(例如,CAR T细胞)可以进一步进行修饰,以沉默或减少编码蛋白质的一种或多种基因的表达,所述蛋白质赋予对癌症疗法包括单克隆抗体的敏感性,以产生武装的细胞(例如,武装的CAR T细胞)。因此,在癌症疗法(例如,化学疗法、单克隆抗体疗法或另一种抗肿瘤治疗)的存在下,武装的细胞可以起作用,并且可以证实优良的功能或功效。涉及赋予对癌症疗法的敏感性的蛋白质的非限制性实例公开于PCT公开号WO 2019/173636中。

[0190] 本公开内容的修饰的细胞(例如,CAR T细胞)可以进一步进行修饰,以沉默或减少编码生长优势因子的一种或多种基因的表达,以产生武装的细胞(例如,武装的CAR T细胞)。沉默或减少癌基因的表达可以对于细胞赋予生长优势。例如,在CAR T细胞制造过程期间沉默或减少TET2基因的表达(例如,破坏表达),导致当与缺乏这种扩增能力的非武装的CAR T细胞相比较时,具有扩增和随后根除肿瘤的显著能力的武装的CAR T细胞的生成。该策略可以与安全开关(例如,本文所述的iC9安全开关)偶联,所述安全开关允许在来自受试者的不良反应或武装的CAR T细胞不受控制的生长的情况下,武装的CAR T细胞的靶向破

坏。生长优势因子的非限制性实例公开于PCT公开号WO 2019/173636中。

[0191] 本公开内容的修饰的细胞(例如,CAR T细胞)可以进一步进行修饰,以表达修饰/嵌合检查点受体,以产生本公开内容的武装的T细胞。

[0192] 修饰/嵌合检查点受体可以包含无效受体、诱饵受体或显性失活受体。无效受体、诱饵受体或显性失活受体可以为修饰/嵌合受体/蛋白质。无效受体、诱饵受体或显性失活受体可以被截断,用于细胞内信号传导结构域的表达。可替代地或另外,无效受体、诱饵受体或显性失活受体可以在细胞内信号传导结构域内的一个或多个氨基酸位置处突变,所述氨基酸位置对于有效信号传导是决定性或必需的。无效受体、诱饵受体或显性失活受体的截断或突变,可以导致受体丧失向细胞或细胞内传递或转导检查点信号的能力。

[0193] 例如,可以通过在武装的细胞(例如,武装的CAR T细胞)的表面上表达修饰的/嵌合的PD-1无效受体,来实现来自在肿瘤细胞的表面上表达的PD-L1受体的免疫压制性检查点信号的稀释或阻断,所述无效受体与也在武装的细胞的表面上表达的内源性(未修饰的)PD-1受体有效地竞争,以减少或抑制免疫压制性检查点信号通过武装的细胞的内源性PD-1受体的转导。在该非限制性实例中,两种不同受体之间与肿瘤细胞上表达的PD-L1结合的竞争,减少或减弱了有效检查点信号传导的水平,从而增强了表达PD-1无效受体的武装的细胞的治疗潜力。

[0194] 修饰/嵌合检查点受体可以包含无效受体、诱饵受体或显性失活受体,其为跨膜受体、膜相关或膜连接受体/蛋白质或细胞内受体/蛋白质。示例性无效、诱饵或显性失活的细胞内受体/蛋白质包括但不限于抑制性检查点信号下游的信号传导组分、转录因子、细胞因子或细胞因子受体、趋化因子或趋化因子受体、细胞死亡或细胞凋亡受体/配体、代谢传感分子、赋予对癌症疗法的敏感性的蛋白质、以及癌基因或肿瘤压制基因。细胞因子、细胞因子受体、趋化因子和趋化因子受体的非限制性实例公开于PCT公开号WO 2019/173636中。

[0195] 修饰/嵌合检查点受体可以包含开关受体(switch receptor)。示例性开关受体包含修饰/嵌合受体/蛋白质,其中天然或野生型细胞内信号传导结构域被不同的细胞内信号传导结构域转换或替换,所述不同的细胞内信号传导结构域或对蛋白质是非天然的和/或不是野生型结构域。例如,用刺激性信号传导结构域替换抑制性信号传导结构域将免疫压制信号转换为免疫刺激信号。可替代地,用不同的抑制性结构域替换抑制性信号传导结构域可以减少或增强抑制性信号传导的水平。经由与内源性野生型检查点受体(而不是开关受体)竞争结合免疫压制性肿瘤微环境内表达的同源检查点受体,开关受体的表达或过表达可以导致同源检查点信号的稀释和/或阻断。武装的细胞(例如,武装的CAR T细胞)可以包含编码开关受体的序列,导致一种或多种开关受体的表达,并且因而改变武装的细胞的活性。武装的细胞(例如,武装的CAR T细胞)可以表达开关受体,其靶向检查点受体下游的细胞内表达的蛋白质、转录因子、细胞因子受体、死亡受体、代谢传感分子、癌症疗法、癌基因和/或肿瘤抑制蛋白或基因。

[0196] 示例性开关受体可以包含蛋白质或可以衍生自蛋白质,所述蛋白质包括但不限于抑制性检查点信号下游的信号传导组分、转录因子、细胞因子或细胞因子受体、趋化因子或趋化因子受体、细胞死亡或细胞凋亡受体/配体、代谢传感分子、赋予对癌症疗法的敏感性的蛋白质、以及癌基因或肿瘤抑制基因。

[0197] 本公开内容的修饰的细胞(例如,CAR T细胞)可以进一步进行修饰,以表达介导条

件基因表达的CLR/CAR，以产生武装的T细胞。CLR/CAR与武装的T细胞的核中的条件基因表达系统的组合构成合成基因表达系统，其在同源配体与CLR或同源抗原与CAR结合时被条件激活。例如，该系统可以通过减少或限制在配体或抗原结合的位点处、在肿瘤环境处或在肿瘤环境内的合成基因表达，来帮助'武装'或增强修饰的T细胞的治疗潜力。

[0198] 基因编辑组合物和方法

通过将转基因引入细胞内来产生修饰的细胞。引入步骤可以包含经由非转座递送系统来递送核酸序列、转基因和/或基因组编辑构建体。

[0199] 将核酸序列、转基因和/或基因组编辑构建体离体、在体内、在体外或原位引入细胞内，可以包含局部递送、吸附、吸收、电穿孔、旋转感染(spin-fection)、共培养、转染、机械递送、声波递送、振动递送、磁转染或通过纳米颗粒介导的递送中的一种或多种。将核酸序列、转基因和/或基因组编辑构建体离体、在体内、在体外或原位引入细胞内，可以包含脂质体转染、磷酸钙转染、fugene转染和树枝状聚合物介导的转染。通过机械转染将核酸序列、转基因和/或基因组编辑构建体离体、在体内、在体外或原位引入细胞内，可以包含细胞挤压、细胞轰击或基因枪技术。通过纳米颗粒介导的转染将核酸序列、转基因和/或基因组编辑构建体离体、在体内、在体外或原位引入细胞内，可以包含脂质体递送、通过胶束递送和通过聚合物囊泡(polymerosome)递送。

[0200] 将核酸序列、转基因和/或基因组编辑构建体离体、在体内、在体外或原位引入细胞内可以包含非病毒载体。非病毒载体可以包含核酸。非病毒载体可以包含质粒DNA、线性双链DNA (dsDNA)、线性单链DNA (ssDNA)、DoggyBone™ DNA、纳米质粒、小环DNA、单链寡脱氧核苷酸(ssODN)、DDNA寡核苷酸、单链mRNA (ssRNA)和双链mRNA (dsRNA)。非病毒载体可以包含如本文所述的转座子。

[0201] 将核酸序列、转基因和/或基因组编辑构建体离体、在体内、在体外或原位引入细胞内可以包含病毒载体。病毒载体可以为非整合的非染色体载体。非整合的非染色体载体的非限制性实例包括腺伴随病毒(AAV)、腺病毒和疱疹病毒。病毒载体可以为整合染色体载体。整合染色体载体的非限制性实例包括腺相关载体(AAV)、慢病毒和γ逆转录病毒。

[0202] 将核酸序列、转基因和/或基因组编辑构建体离体、在体内、在体外或原位引入细胞内可以包含载体的组合。载体组合的非限制性实例包括病毒和非病毒载体、多个非病毒载体或多个病毒载体。载体组合的非限制性实例包括DNA衍生的载体和RNA衍生的载体的组合、RNA和逆转录酶的组合、转座子和转座酶的组合、非病毒载体和内切核酸酶的组合、以及病毒载体和内切核酸酶的组合。

[0203] 基因组修饰可以包含将核酸序列、转基因和/或基因组编辑构建体离体、在体内、在体外或原位引入细胞内，以稳定整合核酸序列、瞬时整合核酸序列、产生核酸序列的位点特异性整合、或产生核酸序列的偏向整合。核酸序列可以为转基因。

[0204] 基因组修饰可以包含将核酸序列、转基因和/或基因组编辑构建体离体、在体内、在体外或原位引入细胞内，以稳定整合核酸序列。稳定的染色体整合可以为随机整合、位点特异性整合或偏向整合。位点特异性整合可以为非辅助的或辅助的。辅助位点特异性整合与定点核酸酶共递送。定点核酸酶包含具有5'和3'核苷酸序列延伸的转基因，其含有与基因组整合位点的上游和下游区域的同源性百分比。具有同源核苷酸延伸的转基因使得能够通过同源重组、微同源性介导的末端连接或非同源末端连接的基因组整合。位点特异性整

合可以在安全港位点处发生。基因组安全港位点能够容纳新遗传材料的整合，其方式确保新近插入的遗传元件可靠地起作用(例如，以治疗有效的表达水平表达)，并且不对宿主基因组造成对宿主生物体造成风险的有害改变。潜在基因组安全港的非限制性实例包括人白蛋白基因的内含子序列、腺伴随病毒位点1 (AAVS1)、AAV病毒在染色体19上的天然存在的整合位点、趋化因子(C-C基序))受体5 (CCR5)基因的位点和小鼠Rosa26基因座的人直向同源物的位点。

[0205] 位点特异性转基因整合可以在破坏靶基因的表达的位点处发生。通过在内含子、外显子、启动子、遗传元件、增强子、抑制子、起始密码子、终止密码子和应答元件处的位点特异性整合，可以发生靶基因表达的破坏。通过位点特异性整合靶向的靶基因的非限制性实例包括TRAC、TRAB、PDI、任何免疫压制性基因和涉及同种异体排斥的基因。

[0206] 位点特异性转基因整合可以在导致靶基因的表达增强的位点处发生。通过在内含子、外显子、启动子、遗传元件、增强子、抑制子、起始密码子、终止密码子和应答元件处的位点特异性整合，可以发生靶基因表达的增强。

[0207] 酶可以用于在宿主基因组中产生链断裂，以促进转基因的递送或整合。酶可以产生单链断裂或双链断裂。断裂诱导酶的非限制性实例包括转座酶、整合酶、内切核酸酶、CRISPR-Cas9、转录激活因子样效应物核酸酶(TALEN)、锌指核酸酶(ZFN)、Cas-CLOVER™和CPF1。断裂诱导酶可以以DNA编码、以mRNA编码、作为蛋白质或作为具有指导RNA (gRNA)的核蛋白复合物递送至细胞。

[0208] 位点特异性转基因整合可以由载体介导的整合位点偏向进行控制。载体介导的整合位点偏向可以通过所选的慢病毒载体或所选的γ逆转录病毒载体进行控制。

[0209] 位点特异性转基因整合位点可以为非稳定的染色体插入。整合的转基因可以变得沉默、去除、切除或进一步修饰。基因组修饰可以为转基因的非稳定整合。非稳定整合可以为瞬时的非染色体整合、半稳定的非染色体整合、半持久的非染色体插入或非稳定的染色体插入。瞬时的非染色体插入可以为染色体外的或细胞质的。在一个方面，转基因的瞬时的非染色体插入并不整合到染色体内，并且修饰的遗传材料在细胞分裂期间不复制。

[0210] 基因组修饰可以为转基因的半稳定或持久的非染色体整合。DNA载体编码支架/基质附着区(S-MAR)模块，其与核基质蛋白结合，用于非病毒载体的附加型保留，允许在分裂细胞的核中的自主复制。

[0211] 基因组修饰可以为转基因的非稳定的染色体整合。整合的转基因可以变得沉默、去除、切除或进一步修饰。

[0212] 通过转基因插入对基因组的修饰可以经由以下发生：通过同源重组(HR)的宿主细胞指导的双链断裂修复(同源性介导的修复)、微同源性介导的末端连接(MMEJ)、非同源末端连接(NHEJ)、转座酶介导的修饰、整合酶介导的修饰、内切核酸酶介导的修饰或重组酶介导的修饰。通过转基因插入对基因组的修饰可以经由CRISPR-Cas9、TALEN、ZFN、Cas-CLOVER™和cpf1而发生。

[0213] 在涉及插入新的或现有的核苷酸/核酸的基因编辑系统中，除切割酶(例如，核酸酶、重组酶、整合酶或转座酶)之外，还必须将插入工具(例如，DNA模板载体、可转座元件(转座子或逆转录转座子)递送到细胞。用于重组酶的此类插入工具的实例可能包括DNA载体。其它基因编辑系统需要整合酶连同插入载体一起、转座酶连同转座子/逆转录转座子一起

等的递送。可以用作切割酶的重组酶的实例是CRE重组酶。可以用于插入工具中的整合酶的非限制性实例包括取自多种病毒(包括AAV、γ逆转录病毒和慢病毒)中的任一种的基于病毒的酶。可以用于插入工具中的转座子/逆转录转座子的实例在本文中更详细地描述。

[0214] 具有离体、在体内、在体外或原位基因组修饰的细胞可以为种系细胞或体细胞。修饰的细胞可以为人、非人、哺乳动物、大鼠、小鼠或犬细胞。修饰的细胞可以为分化的、未分化的或永生化的。修饰的未分化细胞可以为干细胞。修饰的未分化细胞可以为诱导多能干细胞。修饰的细胞可以为免疫细胞。修饰的细胞可以为T细胞、造血干细胞、天然杀伤细胞、巨噬细胞、树突状细胞、单核细胞、巨核细胞或破骨细胞。修饰的细胞可以在细胞休眠、处于激活状态、静止、处于间期、处于前期、处于中期、处于后期或处于末期时进行修饰。修饰的细胞可以为新鲜的、冷冻保存的、散装的、分选成亚群、来自全血、来自白细胞去除术或来自永生化的细胞系。用于从白细胞去除术产物或血液中分离细胞的详细描述公开于PCT公开号WO 2019/173636和PCT/US2019/049816中。

[0215] 本公开内容提供了基因编辑组合物和/或包含基因编辑组合物的细胞。基因编辑组合物可以包含编码DNA结合结构域的序列,以及编码核酸酶蛋白或其核酸酶结构域的序列。编码核酸酶蛋白的序列或编码其核酸酶结构域的序列可以包含DNA序列、RNA序列或其组合。核酸酶或其核酸酶结构域可以包含CRISPR/Cas蛋白、转录激活因子样效应物核酸酶(TALEN)、锌指核酸酶(ZFN)和内切核酸酶中的一种或多种。

[0216] 核酸酶或其核酸酶结构域可以包含核酸酶失活的Cas (dCas)蛋白和内切核酸酶。内切核酸酶可以包含Clo051核酸酶或其核酸酶结构域。基因编辑组合物可以包含融合蛋白。融合蛋白可以包含核酸酶失活的Cas9 (dCas9)蛋白和Clo051核酸酶或Clo051核酸酶结构域。基因编辑组合物可以进一步包含引导序列。引导序列包含RNA序列。

[0217] 本公开内容提供了包含与效应物可操作连接的小Cas9 (Cas9)的组合物。本公开内容提供了融合蛋白,其包含DNA定位组分和效应分子、基本上由其组成或由其组成,其中所述效应物包含小Cas9 (Cas9)。本公开内容的小Cas9构建体可以包括包含IIS型内切核酸酶的效应物。具有活性催化位点的金黄色葡萄球菌(Staphylococcus aureus) Cas9包含SEQ ID NO: 43的氨基酸序列。

[0218] 本公开内容提供了包含与效应物可操作地连接的失活的小Cas9 (dSaCas9)的组合物。本公开内容提供了融合蛋白,其包含DNA定位组分和效应分子、基本上由其组成或由其组成,其中所述效应物包含小的失活的Cas9 (dSaCas9)。本公开内容的小的失活的Cas9 (dSaCas9)构建体可以包括包含IIS型内切核酸酶的效应物。dSaCas9包含SEQ ID NO: 44的氨基酸序列,其包括D10A和N580A突变,以使催化位点失活。

[0219] 本公开内容提供了包含与效应物可操作地连接的失活的Cas9 (dCas9)的组合物。本公开内容提供了融合蛋白,其包含DNA定位组分和效应分子、基本上由其组成或由其组成,其中所述效应物包含失活的Cas9 (dCas9)。本公开内容的失活的Cas9 (dCas9)构建体可以包括包含IIS型内切核酸酶的效应物。

[0220] dCas9可以分离或衍生自化脓性链球菌(*Streptoccocus pyogenes*)。dCas9可以包含在氨基酸位置10和840处具有取代的dCas9,所述取代使催化位点失活。在一些方面,这些取代是D10A和H840A。dCas9可以包含SEQ ID NO: 45或SEQ ID NO: 46的氨基酸序列。

[0221] 示例性Clo051核酸酶结构域包含SEQ ID NO: 47的氨基酸序列、基本上由其组成

或由其组成。

[0222] 示例性dCas9-Clo051 (Cas-CLOVER)融合蛋白可以包含SEQ ID NO: 48的氨基酸序列、基本上由其组成或由其组成。示例性dCas9-Clo051融合蛋白可以由多核苷酸编码,所述多核苷酸包含SEQ ID NO: 49的核酸序列、基本上由其组成或由其组成。编码dCas9-Clo051融合蛋白的核酸可以为DNA或RNA。

[0223] 示例性dCas9-Clo051 (Cas-CLOVER)融合蛋白可以包含SEQ ID NO: 50的氨基酸序列、基本上由其组成或由其组成。示例性dCas9-Clo051融合蛋白可以由多核苷酸编码,所述多核苷酸包含SEQ ID NO: 51的核酸序列、基本上由其组成或由其组成。编码dCas9-Clo051融合蛋白的核酸可以为DNA或RNA。

[0224] 包含基因编辑组合物的细胞可以稳定或瞬时表达基因编辑组合物。优选地,基因编辑组合物是瞬时表达的。引导RNA可以包含与基因组DNA序列内的靶序列互补的序列。基因组DNA序列内的靶序列可以为基因组DNA序列的安全港位点内的靶序列。

[0225] 基因编辑组合物,包括Cas-CLOVER,以及使用这些组合物用于基因编辑的方法在美国专利公开号2017/0107541、2017/0114149、2018/0187185和美国专利号10,415,024中详细描述。

[0226] 还可以使用一种或多种基于聚(组氨酸)的胶束,将基因编辑工具递送至细胞。聚(组氨酸) (例如,聚(L组氨酸))是pH敏感性聚合物,由于咪唑环在不饱和的氮上提供孤电子对。即,聚(组氨酸)通过质子化-去质子化具有两性性质。特别地,在某些pH下,含聚(组氨酸)的三嵌段共聚物可以组装成在表面上具有带正电荷的聚(组氨酸)单元的胶束,从而使得能够与带负电荷的基因编辑分子复合。使用这些纳米颗粒以pH依赖性方式结合且释放蛋白质和/或核酸,可以提供有效和选择性的机制来执行所需的基因修饰。特别地,这种基于胶束的递送系统提供了关于荷电材料的大量灵活性、以及大的有效载荷容量和纳米颗粒有效载荷的靶向释放。在一个实例中,通过使用基于聚(组氨酸)的胶束来递送核酸酶,使得能够位点特异性切割双链DNA。不希望受特定理论的束缚,认为在由各种三嵌段共聚物形成的胶束中,疏水性嵌段聚集以形成核,在端部上留下亲水性嵌段和聚(组氨酸)嵌段,以形成一个或多个包围层。

[0227] 在一个方面,本公开内容提供了由亲水性嵌段、疏水性嵌段和荷电嵌段制成的三嵌段共聚物。在一些方面,亲水性嵌段可以为聚(环氧乙烷) (PEO),并且荷电嵌段可以为聚(L组氨酸)。可以使用的示例三嵌段共聚物是PEO-b-PLA-b-PHIS,其中在每个嵌段中可变数目的重复单元因设计而变。

[0228] 可以用作用于制备三嵌段共聚物的中间体的二嵌段共聚物可以具有与以下偶联的亲水性生物相容性聚(环氧乙烷) (PEO) (其在化学上与PEG同义):各种疏水性脂肪族聚(酸酐)、聚(核酸)、聚(酯)、聚(原酸酯)、聚(肽)、聚(磷腈)和聚(糖),包括但不限于聚(丙交酯) (PLA)、聚(乙交酯) (PLGA)、聚(乳酸共乙醇酸) (PLGA)、聚(ε-己内酯) (PCL)和聚(三亚甲基碳酸酯) (PTMC)。由100%聚乙二醇化表面构成的聚合物胶束具有改善的体外化学稳定性、加强的体内生物利用度和延长的血液循环半衰期。

[0229] 聚合囊泡、聚合物囊泡和基于聚(组氨酸)的胶束,包括包含三嵌段共聚物的那些,及其制备方法,在美国专利号7,217,427;7,868,512;6,835,394;8,808,748;10,456,452;美国公开号2014/0363496;2017/0000743;和2019/0255191;以及PCT公开号WO 2019/

126589中进一步详细描述。

[0230] 诱导型促凋亡多肽

本文公开的诱导型促凋亡多肽优于现有的诱导型多肽，因为本公开内容的诱导型促凋亡多肽的免疫原性低得多。诱导型促凋亡多肽是重组多肽，且因此是非天然存在的。进一步地，序列经重组以产生诱导型促凋亡多肽，其不包含宿主人免疫系统可以识别为“非自身的”非人序列，并且因而在受试者中诱导免疫应答，所述受试者接受诱导型促凋亡多肽、包含诱导型促凋亡多肽的细胞、或者包含诱导型促凋亡多肽或含有诱导型促凋亡多肽的细胞的组合物。

[0231] 本公开内容提供了包含配体结合区、接头和促凋亡肽的诱导型促凋亡多肽，其中所述诱导型促凋亡多肽不包含非人序列。在某些方面，非人序列包含限制位点。在某些方面，配体结合区可以为多聚体配体结合区。在某些方面，促凋亡肽是半胱天冬酶多肽。半胱天冬酶多肽的非限制性实例包括半胱天冬酶1、半胱天冬酶2、半胱天冬酶3、半胱天冬酶4、半胱天冬酶5、半胱天冬酶6、半胱天冬酶7、半胱天冬酶8、半胱天冬酶9、半胱天冬酶10、半胱天冬酶11、半胱天冬酶12和半胱天冬酶14。优选地，半胱天冬酶多肽是半胱天冬酶9多肽。半胱天冬酶9多肽可以为截短的半胱天冬酶9多肽。诱导型促凋亡多肽可以为非天然存在的。当半胱天冬酶是半胱天冬酶9或截短的半胱天冬酶9时，诱导型促凋亡多肽也可以被称为“iC9安全开关”。

[0232] 诱导型半胱天冬酶多肽可以包含(a)配体结合区、(b)接头和(c)半胱天冬酶多肽，其中所述诱导型促凋亡多肽不包含非人序列。在某些方面，诱导型半胱天冬酶多肽包含(a)配体结合区、(b)接头和(c)截短的半胱天冬酶9多肽，其中所述诱导型促凋亡多肽不包含非人序列。

[0233] 配体结合区可以包含FK506结合蛋白12 (FKBP12)多肽。包含FK506结合蛋白12(FKBP12)多肽的配体结合区的氨基酸序列可以包含在序列的位置36处的修饰。修饰可以为在位置36 (F36V)处的缬氨酸(V)取代苯丙氨酸(F)。FKBP12多肽可以包含氨基酸序列、基本上由氨基酸序列组成或由氨基酸序列组成，所述氨基酸序列与SEQ ID NO: 73至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100% (或两者之间的任何百分比)等同。FKBP12多肽可以由多核苷酸编码，所述多核苷酸包含核酸序列或由核酸序列组成，所述核酸序列与SEQ ID NO: 74至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100% (或两者之间的任何百分比)等同。

[0234] 接头区可以包含氨基酸序列、基本上由氨基酸序列组成或由氨基酸序列组成，所述氨基酸序列与SEQ ID NO: 75至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100% (或两者之间的任何百分比)等同，或者接头区可以由多核苷酸编码，所述多核苷酸包含核酸序列或由核酸序列组成，所述核酸序列与SEQ ID NO: 76至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100% (或两者之间的任何百分比)等同。在一些方面，编码接头的核酸序列不包含限制位点。

[0235] 截短的半胱天冬酶9多肽可以包含在序列的位置87处不包含精氨酸(R)的氨基酸序列。可替代地或另外，截短的半胱天冬酶9多肽可以包含在序列的位置282处不包含丙氨酸(A)的氨基酸序列。截短的半胱天冬酶9多肽可以包含氨基酸序列、基本上由氨基酸序列组成或由氨基酸序列组成，所述氨基酸序列与SEQ ID NO: 77至少75%、80%、85%、90%、95%、

96%、97%、98%、99%或100% (或两者之间的任何百分比)等同,或者截短的半胱天冬酶9多肽可以由多核苷酸编码,所述多核苷酸包含核酸序列或由核酸序列组成,所述核酸序列与SEQ ID NO: 78至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100% (或两者之间的任何百分比)等同。

[0236] 在某些方面,当多肽包含截短的半胱天冬酶9多肽时,诱导型促凋亡多肽包含氨基酸序列、基本上由氨基酸序列组成或由氨基酸序列组成,所述氨基酸序列与SEQ ID NO: 79至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100% (或两者之间的任何百分比)等同,或者诱导型促凋亡多肽由多核苷酸编码,所述多核苷酸包含核酸序列或由核酸序列组成,所述核酸序列与SEQ ID NO: 5、SEQ ID NO: 6或SEQ ID NO: 80至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100% (或两者之间的任何百分比)等同。

[0237] 诱导型促凋亡多肽可以在本领域已知的任何启动子的转录调控下在细胞中表达,所述启动子能够启动和/或调控诱导型促凋亡多肽在该细胞中的表达。

[0238] 诱导型促凋亡多肽的活化可以通过例如由诱导剂介导的化学诱导的二聚化(CID)来完成,以产生条件控制的蛋白质或多肽。由于不稳定二聚化试剂的降解或单体竞争性抑制剂的施用,促凋亡多肽不仅是诱导型,而且这些多肽的诱导也是可逆的。

[0239] 在某些方面,当配体结合区包含在位置36处具有缬氨酸(V)取代苯丙氨酸(F)(F36V)的FKBP12多肽时,诱导剂可以包含AP1903,一种合成药物(CAS索引名称:2-哌啶羧酸、1-[(2S)-1-氧代-2-(3,4,5-三甲氧基苯基)丁基]-、1,2-乙二基双[亚氨基(2-氧代-2,1-乙二基)氧基-3,1-亚苯基[(1R)-3-(3,4-二甲氧基苯基)亚丙基]]酯、[2S-[1 (R*),2R*[S*[S*[1 (R*),2R*]]]]]-(9Cl) CAS登记号:195514-63-7;分子式:C78H98N4O20;分子量:1411.65));AP20187 (CAS登记号:195514-80-8和分子式:C82H107N5O20)或AP20187类似物,例如AP1510。如本文使用的,诱导剂AP20187、AP1903和AP1510可以互换使用。

[0240] 诱导型促凋亡肽和诱导这些肽的方法在美国专利公开号WO 2019/0225667和PCT公开号WO2018/068022中详细描述。

[0241] 嵌合刺激受体和重组HLA-E多肽

对于施用于任何患者“普遍”安全的过继细胞组合物需要同种异体反应性的显著减少或消除。为此,可以修饰本公开内容的细胞(例如,同种异体细胞),以中断T细胞受体(TCR)和/或一类主要组织相容性复合物(MHC)的表达或功能。TCR介导移植物抗宿主(GvH)反应,而MHC介导宿主抗移植物(HvG)反应。在优选的方面,消除TCR的任何表达和/或功能,以预防可以引起受试者死亡的T细胞介导的GvH。因此,在一个优选的方面,本公开内容提供了纯的TCR阴性同种异体T细胞组合物(例如,组合物中的每种细胞以如此低的水平表达,以致于或无法检测或不存在)。

[0242] MHC I类(MHC-I,具体而言,HLA-A、HLA-B和HLA-C)的表达和/或功能被减少或消除,以预防HvG并因而改善细胞在受试者中的移植物植入。改善的移植物植入导致细胞的更长持久性,并且因此对于受试者更大的治疗窗口。具体而言,MHC-1的结构元件β-2-微球蛋白(B2M)的表达和/或功能被减少或消除。

[0243] 上述策略引起进一步的挑战。T细胞中的T细胞受体(TCR)敲除(KO)导致CD3-ζ(CD3z或CD3ζ)的表达丧失,所述CD3-ζ是TCR复合物的部分。TCR-KO T细胞中的CD3ζ丧失急剧减少了使用标准刺激/激活试剂(包括但不限于激动剂抗CD3 mAb),最佳地激活且扩增这

些细胞的能力。当TCR复合物的任何一种组分的表达或功能中断时，该复合物的所有组分都丧失，包括TCR-α (TCRα)、TCR-β (TCRβ)、CD3-γ (CD3γ)、CD3-ε (CD3ε)、CD3-δ (CD3δ)和CD3-ζ (CD3ζ)。CD3ε和CD3ζ两者均为T细胞激活和扩增所需要的。激动剂抗CD3 mAb通常识别CD3ε和复合物内可能的另一种蛋白质，其依次又向CD3ζ发信号。CD3ζ提供了用于T细胞激活的主要刺激(连同次要共刺激信号一起)，用于最佳激活和扩增。在正常条件下，完全T细胞激活取决于与第二信号结合的TCR的接合，所述第二信号由一种或多种共刺激受体(例如，CD28、CD2、4-1BBL)介导，其增强免疫应答。然而，当TCR不存在时，当使用标准激活/刺激试剂(包括激动剂抗CD3 mAb)刺激时，T细胞扩增严重减少。事实上，当使用标准激活/刺激试剂(包括激动剂抗CD3 mAb)刺激时，T细胞扩增减少到正常扩增水平的仅20-40%。

[0244] 因此，本公开内容提供了非天然存在的嵌合刺激受体(CSR)，其包括：(a)包含激活组分的胞外结构域，其中所述激活组分分离自或衍生自第一蛋白质；(b)跨膜结构域；以及(c)包含至少一个信号转导结构域的胞内结构域，其中所述至少一个信号转导结构域分离自或衍生自第二蛋白质；其中所述第一蛋白质和第二蛋白质并非等同的。

[0245] 激活组分可以包含以下中的一种或多种的一部分：T细胞受体(TCR)的组分、TCR复合物的组分、TCR共受体的组分、TCR共刺激蛋白的组分、TCR抑制蛋白的组分、细胞因子受体和激活组分的激动剂与之结合的趋化因子受体。激活组分可以包含激动剂与之结合的CD2细胞外结构域或其一部分。

[0246] 信号转导结构域可以包含以下中的一种或多种：人信号转导结构域的组分、T细胞受体(TCR)、TCR复合物的组分、TCR共受体的组分、TCR共刺激蛋白的组分、TCR抑制蛋白的组分、细胞因子受体和趋化因子受体。信号转导结构域可以包含CD3蛋白或其一部分。CD3蛋白可以包含CD3ζ蛋白或其一部分。

[0247] 胞内结构域可以进一步包含细胞质结构域。细胞质结构域可以分离自或衍生自第三蛋白质。第一蛋白质和第三蛋白质可以是等同的。胞外结构域可以进一步包含信号肽。信号肽可以衍生自第四蛋白质。第一蛋白质和第四蛋白质可以是等同的。跨膜结构域可以分离自或衍生自第五蛋白质。第一蛋白质和第五蛋白质可以是等同的。

[0248] 在一些方面，激活组分并不结合天然存在的分子。在一些方面，激活组分结合天然存在的分子，但CSR在激活组分与天然存在的分子结合后并不转导信号。在一些方面，激活组分结合非天然存在的分子。在一些方面，激活组分并不结合天然存在的分子，而是结合非天然存在的分子。CSR可以在激活组分与非天然存在的分子结合后选择性地转导信号。

[0249] 在一个优选的方面，本公开内容提供了非天然存在的嵌合刺激受体(CSR)，其包括：(a)包含信号肽和激活组分的胞外结构域，其中所述信号肽包含CD2信号肽或其一部分，并且其中所述激活组分包含激动剂与之结合的CD2细胞外结构域或其一部分；(b)跨膜结构域，其中所述跨膜结构域包含CD2跨膜结构域或其一部分；以及(c)包含细胞质结构域和至少一个信号转导结构域的胞内结构域，其中所述细胞质结构域包含CD2细胞质结构域或其一部分，并且其中所述至少一个信号转导结构域包含CD3ζ蛋白或其一部分。在一些方面，非天然CSR包含与SEQ ID NO: 81至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100% (或两者之间的任何百分比)等同的氨基酸序列。在一个优选的方面，非天然存在的CSR包含SEQ ID NO: 81的氨基酸序列。

[0250] 本公开内容还提供了非天然存在的嵌合刺激受体(CSR)，其中胞外结构域包含修

饰。当与激活组分或第一蛋白质的野生型序列相比较时，修饰可以包含激活组分或第一蛋白质的氨基酸序列的突变或截短。激活组分的氨基酸序列的突变或截短可以包含激动剂与之结合的CD2细胞外结构域或其一部分的突变或截短。CD2细胞外结构域的突变或截短可以减少或消除与天然存在的CD58的结合。在一些方面，包含突变或截短的CD2细胞外结构域包含与SEQ ID NO: 82至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100% (或两者之间的任何百分比)等同的氨基酸序列。在一个优选的方面，包含突变或截短的CD2细胞外结构域包含SEQ ID NO: 82的氨基酸序列。

[0251] 在一个优选的方面，本公开内容提供了非天然存在的嵌合刺激受体(CSR)，其包括：(a)包含信号肽和激活组分的胞外结构域，其中所述信号肽包含CD2信号肽或其一部分，并且其中所述激活组分包含激动剂与之结合的CD2细胞外结构域或其一部分，并且其中激动剂与之结合的CD2细胞外结构域或其一部分包含突变或截短；(b)跨膜结构域，其中所述跨膜结构域包含CD2跨膜结构域或其一部分；以及(c)包含细胞质结构域和至少一个信号转导结构域的胞内结构域，其中所述细胞质结构域包含CD2细胞质结构域或其一部分，并且其中所述至少一个信号转导结构域包含CD3ζ蛋白或其一部分。在一些方面，非天然CSR包含与SEQ ID NO: 83至少75%、80%、85%、90%、95%、96%、97%、98%、99%或100% (或两者之间的任何百分比)等同的氨基酸序列。在一个优选的方面，非天然存在的CSR包含SEQ ID NO: 83的氨基酸序列。

[0252] 本公开内容提供了编码本文公开的任何CSR的核酸序列。本公开内容提供了包含编码本文公开的任何CSR的核酸序列的转座子或载体。

[0253] 本公开内容提供了包含本文公开的任何CSR的细胞。本公开内容提供了包含编码本文公开的任何CSR的核酸序列的细胞。本公开内容提供了包含载体的细胞，所述载体包含编码本文公开的任何CSR的核酸序列。本公开内容提供了包含转座子的细胞，所述转座子包含编码本文公开的任何CSR的核酸序列。

[0254] 本文公开的修饰的细胞可以为同种异体细胞或自体细胞。在一些优选的方面，修饰的细胞是同种异体细胞。在一些方面，修饰的细胞是自体T细胞或修饰的自体CAR T细胞。在一些优选的方面，修饰的细胞是同种异体T细胞或修饰的同种异体CAR T细胞。

[0255] 本公开内容提供了包含本文公开的任何CSR的组合物。本公开内容提供了包含编码本文公开的任何CSR的核酸序列的组合物。本公开内容提供了包含载体的组合物，所述载体包含编码本文公开的任何CSR的核酸序列。本公开内容提供了包含转座子的组合物，所述转座子包含编码本文公开的任何CSR的核酸序列。本公开内容提供了包含本文公开的修饰的细胞的组合物、或包含本文公开的多种修饰的细胞的组合物。

[0256] 本公开内容提供了修饰的T淋巴细胞(T细胞)，其包含：(a)编码T细胞受体(TCR)的内源序列的修饰，其中所述修饰减少或消除了TCR的表达或活性水平；以及(b)嵌合刺激受体(CSR)，其包括：(i)包含激活组分的胞外结构域，其中所述激活组分分离自或衍生自第一蛋白质；(ii)跨膜结构域；以及(iii)包含至少一个信号转导结构域的胞内结构域，其中所述至少一个信号转导结构域分离自或衍生自第二蛋白质；其中所述第一蛋白质和第二蛋白质并非等同的。

[0257] 修饰的T细胞可以进一步包含诱导型促凋亡多肽。修饰的T细胞可以进一步包含编码β-2-微球蛋白(B2M)的内源序列的修饰，其中所述修饰减少或消除了主要组织相容性复

合物(MHC) I类(MHC-I)的表达或活性水平。

[0258] 修饰的T细胞可以进一步包含非天然存在的多肽,其包含HLA I类组织相容性抗原,α链E (HLA-E)多肽。包含HLA-E多肽的非天然存在的多肽可以进一步包含B2M信号肽。包含HLA-E多肽的非天然存在的多肽可以进一步包含B2M多肽。包含HLA-E多肽的非天然存在的多肽可以进一步包含接头,其中所述接头置于B2M多肽和HLA-E多肽之间。包含HLA-E多肽的非天然存在的多肽可以进一步包含肽和B2M多肽。包含HLA-E的非天然存在的多肽可以进一步包含置于B2M信号肽和肽之间的第一接头,以及置于B2M多肽和编码HLA-E的肽之间的第二接头。

[0259] 修饰的T细胞可以进一步包含非天然存在的抗原受体、编码治疗性多肽的序列或其组合。非天然存在的抗原受体可以包含嵌合抗原受体(CAR)。

[0260] CSR可以在修饰的T细胞中瞬时表达。CSR可以在修饰的T细胞中稳定表达。包含HLA-E多肽的多肽可以在修饰的T细胞中瞬时表达。包含HLA-E多肽的多肽可以在修饰的T细胞中稳定表达。诱导型促凋亡多肽可以在修饰的T细胞中瞬时表达。诱导型促凋亡多肽可以在修饰的T细胞中稳定表达。非天然存在的抗原受体或编码治疗性蛋白质的序列可以在修饰的T细胞中瞬时表达。非天然存在的抗原受体或编码治疗性蛋白质的序列可以在修饰的T细胞中稳定表达。

[0261] 如本文详细描述的,基因编辑组合物,包括但不限于包含dCas9-Clo051的RNA引导的融合蛋白,可以用于靶向且降低或消除内源性T细胞受体的表达。在优选的方面,基因编辑组合物靶向并缺失编码内源性T细胞受体的基因、基因的一部分或基因的调控元件(例如启动子)。用于生成引导RNA (gRNA)模板的引物(包括T7启动子、基因组靶序列和gRNA支架)的非限制性实例公开于PCT申请号PCT/US2019/049816中,所述引导RNA模板用于靶向且缺失TCR-α (TCR-α)、靶向且缺失TCR-β (TCR-β)、以及靶向且缺失β-2-微球蛋白(β2M)。

[0262] 基因编辑组合物,包括但不限于包含dCas9-Clo051的RNA引导的融合蛋白,可以用于靶向且降低或消除内源性MHCI、MHCII或MHC激活物的表达。在优选的方面,基因编辑组合物靶向且缺失编码内源性MHCI、MHCII或MHC激活物的一种或多种组分的基因、基因的一部分或基因的调控元件(例如启动子)。用于靶向且缺失MHC激活物的引导RNA (gRNA)的非限制性实例公开于PCT申请号PCT/US2019/049816中。

[0263] 非天然存在的嵌合刺激受体,编码TCR-α (TCR-α)、TCR-β (TCR-β)和/或β-2-微球蛋白(β2M)的内源序列的遗传修饰,以及包含HLA I类组织相容性抗原α链E (HLA-E)多肽的非天然存在的多肽的详细描述,公开于PCT申请号PCT/US2019/049816中。

[0264] 制剂、剂量和施用模式

本公开内容提供了用于施用本文所述的组合物的制剂、剂量和方法。

[0265] 所公开的组合物和药物组合物可以进一步包含任何合适的辅助剂中的至少一种,所述辅助剂例如但不限于稀释剂、粘合剂、稳定剂、缓冲液、盐、亲脂性溶剂、防腐剂、佐剂等等。药学上可接受的辅助剂是优选的。此类无菌溶液的非限制性实例和制备方法是本领域众所周知的,例如但不限于Gennaro,编辑,Remington's Pharmaceutical Sciences,第18版,Mack Publishing Co. (Easton,Pa.) 1990,以及"Physician's Desk Reference",第52版,Medical Economics (Montvale,N.J.) 1998。如本领域众所周知的或如本文所述的,可以常规选择药学上可接受的载体,其适合于蛋白质支架、片段或变体组合物的施用模式、

溶解性和/或稳定性。

[0266] 适合于使用的药物赋形剂和添加剂的非限制性实例包括蛋白质、肽、氨基酸、脂质和碳水化合物(例如,糖,包括单糖、二糖、三糖、四糖和寡糖;衍生糖,例如糖醇、醛糖酸、酯化糖等等;以及多糖或糖聚合物),其可以单独或组合存在,单独或组合构成按重量或体积计的1-99.99%。蛋白质赋形剂的非限制性实例包括血清白蛋白,例如人血清白蛋白(HSA)、重组人白蛋白(rHA)、明胶、酪蛋白等等。也可以在缓冲能力方面起作用的代表性的氨基酸/蛋白质组分,包括丙氨酸、甘氨酸、精氨酸、甜菜碱、组氨酸、谷氨酸、天冬氨酸、半胱氨酸、赖氨酸、亮氨酸、异亮氨酸、缬氨酸、甲硫氨酸、苯丙氨酸、阿斯巴甜等等。一种优选的氨基酸是甘氨酸。

[0267] 适合于使用的碳水化合物赋形剂的非限制性实例包括单糖,例如果糖、麦芽糖、半乳糖、葡萄糖、D-甘露糖、山梨糖等等;二糖,例如乳糖、蔗糖、海藻糖、纤维二糖等等;多糖,例如棉子糖、松三糖、麦芽糖糊精、葡聚糖、淀粉等等;以及糖醇,例如甘露糖醇、木糖醇、麦芽糖醇、乳糖醇、木糖醇、山梨糖醇(葡萄糖醇)、肌醇等等。优选地,碳水化合物赋形剂是甘露糖醇、海藻糖和/或棉子糖。

[0268] 组合物还可以包括缓冲液或pH调节剂;通常,缓冲液是由有机酸或碱制备的盐。代表性缓冲液包括有机酸盐,例如柠檬酸、抗坏血酸、葡糖酸、碳酸、酒石酸、琥珀酸、乙酸或苯二甲酸的盐;Tris、盐酸氨丁三醇或磷酸盐缓冲液。优选的缓冲液是有机酸盐,例如柠檬酸盐。

[0269] 另外,所公开的组合物可以包括聚合赋形剂/添加剂,例如聚乙烯吡咯烷酮、蔗聚糖(聚合糖)、葡萄糖结合剂(dextrate) (例如环糊精,如2-羟丙基-β-环糊精)、聚乙二醇、调味剂、抗微生物剂、甜味剂、抗氧化剂、抗静电剂、表面活性剂(例如聚山梨醇酯,例如“TWEEN 20”和“TWEEN 80”)、脂质(例如磷脂、脂肪酸)、类固醇(例如胆固醇)和螯合剂(例如EDTA)。

[0270] 许多已知和开发的模式可以用于施用治疗有效量的本文公开的组合物或药物组合物。施用模式的非限制性实例包括推注、经颊、输注、关节内、支气管内、腹内、囊内、软骨内、腔内、体腔内、小脑内、脑室内、结肠内、宫颈内、胃内、肝内、病灶内、肌内、心肌内、鼻内、眼内、骨内(intraosseous)、骨内(intraosteal)、骨盆内、心包内、腹膜内、胸膜内、前列腺内、肺内、直肠内、肾内、视网膜内、脊柱内、滑膜内、胸腔内、子宫内、瘤内、静脉内、膀胱内、经口、肠胃外、直肠、舌下、皮下、经皮或阴道手段。

[0271] 本公开内容的组合物可以制备用于特别是以液体溶液或悬浮液的形式的肠胃外(皮下、肌内或静脉内)或任何其它施用;用于特别是以半固体形式例如但不限于乳膏和栓剂的阴道或直肠施用;用于例如但不限于以片剂或胶囊的形式的经颊或舌下施用;或鼻内施用,例如但不限于粉末、滴鼻剂或气溶胶或某些药剂的形式;或经皮施用,例如但不限于凝胶、软膏、洗剂、悬浮液或贴剂递送系统,其具有化学增强剂例如二甲基亚砜,以或修饰皮肤结构或增加经皮贴剂中的药物浓度(Junginger等人,于“Drug Permeation Enhancement;” Hsieh,D. S.,编辑,第59-90页(Marcel Dekker,Inc. New York 1994),或具有氧化剂,其使得能够将含有蛋白质和肽的制剂施加到皮肤上(WO 98/53847),或施加电场以产生瞬时转运途径,例如电穿孔,或增加荷电药物通过皮肤的流动性,例如离子电渗疗法,或施加超声波,例如超声导入(美国专利号4,309,989和4,767,402) (上述公开和专利

整体引入本文作为参考)。

[0272] 对于肠胃外施用,本文公开的任何组合物可以配制为与药学上可接受的肠胃外媒介物结合或分开提供的溶液、悬浮液、乳状液、颗粒、粉末或冻干粉末。用于肠胃外施用的制剂可以含有作为常见赋形剂的无菌水或盐水、聚亚烷基二醇例如聚乙二醇、植物起源的油、氢化萘等等。用于注射的水性或油性悬浮液可以根据已知方法,通过使用适当的乳化剂或增湿剂和助悬剂来制备。用于注射的药剂可以为无毒、非经口施用的稀释剂,例如在溶剂中的水溶液、无菌可注射溶液或悬浮液。作为可用的媒介物或溶剂,水、林格氏溶液、等渗盐水等是允许的;作为普通溶剂或助悬溶剂,可以使用无菌不挥发性油。为了这些目的,可以使用任何种类的不挥发性油和脂肪酸,包括天然或合成或半合成的脂肪油或脂肪酸;天然或合成或半合成的甘油单酯或甘油二酯或甘油三酯。肠胃外施用是本领域已知的,并且包括但不限于常规注射手段,如美国专利号5,851,198中所述的气压无针注射装置、以及如美国专利号5,839,446中所述的激光穿孔装置。

[0273] 用于经口施用的制剂依赖佐剂(例如,间苯二酚和非离子表面活性剂,例如聚氧乙烯油基醚和正十六烷基聚乙烯醚)的共施用,以人为地增加肠壁的渗透性,以及酶促抑制剂(例如,胰腺胰蛋白酶抑制剂、二异丙基氟磷酸酯(DFF)和抑肽酶(trasylol))的共施用,以抑制酶促降解。在美国专利号6,309,663中描述了用于递送亲水性药剂包括蛋白质和蛋白质支架以及至少两种表面活性剂的组合的制剂,所述制剂预期用于经口、经颊、粘膜、鼻、肺、阴道、跨膜或直肠施用。用于经口施用的固体型剂型的活性组成成分化合物可以与至少一种添加剂混合,所述添加剂包括蔗糖、乳糖、纤维素、甘露糖醇、海藻糖、棉子糖、麦芽糖醇、葡聚糖、淀粉、琼脂、精氨酸、几丁质、脱乙酰壳多糖、果胶、黄蓍树胶、阿拉伯树胶、明胶、胶原、酪蛋白、白蛋白、合成或半合成聚合物和甘油酯。这些剂型还可以含有其它类型的添加剂,例如惰性稀释剂、润滑剂(例如硬脂酸镁、对羟基苯甲酸酯)、防腐剂(例如山梨酸、抗坏血酸、α-生育酚)、抗氧化剂(例如半胱氨酸)、崩解剂、粘合剂、增稠剂、缓冲剂、甜味剂、调味剂、加香剂等。

[0274] 片剂和丸剂可以进一步加工成肠包衣制剂。用于经口施用的液体制剂包括可允许医学用途的乳状液、糖浆剂、酏剂、悬浮液和溶液制剂。这些制剂可以含有通常用于所述领域中的惰性稀释剂,例如水。脂质体也已被描述为用于胰岛素和肝素的药物递送系统(美国专利号4,239,754)。最近,混合氨基酸(类蛋白)的人工聚合物的微球已用于递送药物(美国专利号4,925,673)。此外,美国专利号5,879,681和美国专利号5,871,753中描述且用于经口递送生物活性剂的载体化合物是本领域已知的。

[0275] 对于肺施用,优选地,本文所述的组合物或药物组合物以有效到达肺的下气道或鼻窦的粒径递送。组合物或药物组合物可以通过本领域已知的各种吸入或鼻装置中的任一种进行递送,所述装置用于通过吸入的治疗剂施用。能够将气溶胶化制剂沉积在患者的窦腔或肺泡中的这些装置包括计量吸入器、雾化器(例如喷射雾化器、超声雾化器)、干粉发生器、喷雾器等等。所有此类装置都可以使用适合于施用的制剂,用于将本文所述的组合物或药物组合物以气溶胶分配。这种气溶胶可以由溶液(水性和非水性)或固体颗粒组成。另外,可以通过在压力下迫使至少一种蛋白质支架的悬浮液或溶液通过喷嘴,来生产包括本文所述的组合物或药物组合物的喷雾剂。在计量吸入器(MDI)中,推进剂、本文所述的组合物或药物组合物、以及任何赋形剂或其它添加剂作为包括液化压缩气体的混合物包含在罐中。

计量阀的致动将混合物作为气溶胶释放，所述气溶胶优选含有在小于约10 μm、优选约1 μm至约5 μm、并且最优选约2 μm至约3 μm的尺寸范围中的颗粒。肺施用、制剂和有关装置的更详细描述公开于PCT公开号WO 2019/049816中。

[0276] 对于通过粘膜表面的吸收，组合物包括包含多个亚微米颗粒、粘膜粘附大分子、生物活性肽和水性连续相的乳状液，其通过实现乳状液颗粒的粘膜粘附来促进通过粘膜表面的吸收(美国专利号5,514,670)。适合于施加本公开内容的乳状液的粘液表面可以包括角膜、结膜、颊、舌下、鼻、阴道、肺、胃、肠和直肠施用途径。用于阴道或直肠施用的制剂，例如栓剂，可以含有作为赋形剂的例如聚亚烷基二醇、凡士林、可可脂等等。用于鼻内施用的制剂可以为固体，并且含有作为赋形剂的例如乳糖，或者可以为滴鼻剂的水性或油性溶液。对于经颊施用，赋形剂包括糖、硬脂酸钙、硬脂酸镁、预胶化淀粉等等(美国专利号5,849,695)。粘膜施用和制剂的更详细描述公开于PCT公开号WO 2019/049816中。

[0277] 对于经皮施用，本文公开的组合物或药物组合物被封装在递送装置中，所述递送装置例如脂质体或聚合物纳米颗粒、微粒、微胶囊或微球体(除非另有说明，否则统称为微粒)。许多合适的装置是已知的，包括由合成聚合物和天然聚合物制成的微粒，所述合成聚合物例如聚羟基酸，例如聚乳酸、聚乙醇酸及其共聚物，聚原酸酯、聚酐和聚磷腈，所述天然聚合物例如胶原、聚氨基酸、白蛋白和其它蛋白质、海藻酸盐和其它多糖及其组合(美国专利号5,814,599)。经皮施用、制剂和合适装置的更详细描述公开于PCT公开号WO 2019/049816中。

[0278] 可以期望在延长的时间段内，例如从单次施用开始一周到一年的时期，将所公开的化合物递送至受试者。可以利用各种缓慢释放、贮库或植入剂型。例如，剂型可以含有在体液中具有低溶解度的化合物的药学上可接受的无毒盐，例如，(a)与多元酸的酸加成盐，所述多元酸例如磷酸、硫酸、柠檬酸、酒石酸、鞣酸、扑酸、海藻酸、聚谷氨酸、萘单磺酸或萘二磺酸、聚半乳糖醛酸等等；(b)与多价金属阳离子的盐，所述多价金属阳离子例如锌、钙、铋、钡、镁、铝、铜、钴、镍、镉等等，或与有机阳离子的盐，所述有机阳离子由例如N,N'-二苄基乙二胺或乙二胺形成；或(c) (a)和(b)的组合，例如鞣酸锌盐。另外，所公开的化合物或优选地相对不溶性的盐，例如刚刚描述的那些盐，可以在适合于注射的凝胶中进行配制，所述凝胶例如具有例如芝麻油的单硬脂酸铝凝胶。特别优选的盐是锌盐、鞣酸锌盐、双羟萘酸盐等等。另一种类型的用于注射的缓慢释放贮库制剂，将含有分散用于封装在缓慢降解、无毒、无抗原性聚合物中的化合物或盐，所述聚合物例如如美国专利号3,773,919中描述的聚乳酸/聚乙醇酸聚合物。化合物或优选相对不溶性的盐，例如上文描述的那些盐，也可以在胆固醇基质硅橡胶丸剂中进行配制，特别是用于动物中。另外的缓慢释放、贮库或植入制剂，例如气体或液体脂质体，是文献中已知的(美国专利号5,770,222和“Sustained and Controlled Release Drug Delivery Systems”，J. R. Robinson编辑，Marcel Dekker，Inc.，N.Y.，1978)。

[0279] 合适的剂量是本领域众所周知的。参见例如，Wells等人，编辑，Pharmacotherapy Handbook，第2版，Appleton and Lange，Stamford，Conn. (2000)；PDR Pharmacopoeia，Tarascon Pocket Pharmacopoeia 2000，豪华版，Tarascon Publishing，Loma Linda，Calif. (2000)；Nursing 2001 Handbook of Drugs，第21版，Springhouse Corp.，Springhouse，Pa.，2001；Health Professional's Drug Guide 2001，编辑，Shannon，

Wilson,Stang,Prentice-Hall,Inc,Upper Saddle River,N.J。优选剂量可以任选地包括约0.1-99和/或100-500 mg/kg/施用,或者其任何范围、值或分数,或者达到约0.1-5000 μg/ml血清浓度/单次施用或多重施用的血清浓度,或者其任何范围、值或分数。关于本文公开的组合物或药物组合物的优选剂量范围是约1 mg/kg,直到约3、约6或约12 mg/kg受试者的体重。

[0280] 可替代地,施用的剂量可以取决于已知因素而变,所述因素例如特定药剂的药效学特性、以及其施用模式和途径;接受者的年龄、健康和重量;症状的性质和程度、并发治疗的种类、治疗频率和所需效应。通常,活性成分的剂量可以为约0.1至100毫克/千克体重。通常,0.1至50,且优选0.1至10毫克/千克/施用或以持续释放形式有效获得所需结果。

[0281] 作为非限制性实例,人或动物的治疗可以使用单一、输注或重复剂量,在第1-40天中的至少一天,或者可替代地或另外,在第1-52周中的至少一周,或者可替代地或另外,在1-20年中的至少一年,或其任何组合,作为约0.1至100 mg/kg或者其任何范围、值或分数/天的本文公开的组合物或药物组合物的一次性或周期性剂量提供。

[0282] 适合于内部施用的剂型一般含有约0.001毫克至约500毫克的活性成分/单位或容器。在这些药物组合物中,基于组合物的总重量,活性成分通常以按重量计约0.5-99.999%的量存在。

[0283] 有效量可以包含约0.001至约500 mg/kg/单次(例如,推注)、多重或连续施用的量,以达到0.01-5000 μg/ml血清浓度/单次、多重或连续施用的血清浓度,或者其中的任何有效范围或值,如使用已知方法来完成且确定的,如本文所述或有关领域已知的。

[0284] 在其中待施用于有需要的受试者的组合物是如本文公开的修饰的细胞的方面,可以施用约$1x10^3$至$1x10^{15}$个细胞;约$1x10^4$至$1x10^{12}$个细胞;约$1x10^5$至$1x10^{10}$个细胞;约$1x10^6$至$1x10^9$个细胞;约$1x10^6$至$1x10^8$个细胞;约$1x10^6$至$1x10^7$个细胞;或约$1x10^6$至$25x10^6$个细胞的细胞。在一个方面,施用约$5x10^6$至$25x10^6$个细胞的细胞。

[0285] 所公开的组合物和药物组合物的药学上可接受的赋形剂、制剂、剂量和施用方法的更详细描述公开于PCT公开号WO 2019/049816中。

[0286] 使用本公开内容的组合物的方法

本公开内容提供了所公开的组合物用于改善转座效率的用途。具体而言,该方法包括使一种细胞或多种细胞与组合物接触,所述组合物包含:第一核酸序列和第二核酸序列,所述第一核酸序列包含:(a)第一反向末端重复(ITR)或编码第一ITR的序列,(b)第二ITR或编码第二ITR的序列,以及(c) ITR内序列或编码ITR内的序列,其中所述ITR内序列包含转座子序列或编码转座子的序列;和所述第二核酸序列包含ITR间序列或编码ITR间的序列,其中所述ITR间序列的长度等于或小于700个核苷酸,其中当与包含大于700个核苷酸的第二核酸序列或ITR间序列的等同组合物相比较时,细胞或多种细胞内的转座效率得到改善。在一个方面,转座效率改善了至少5%、至少10%、至少15%、至少20%、至少25%、至少30%、至少35%、至少40%、至少45%、至少50%、至少55%、至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少90%。

[0287] 本公开内容提供了所公开的组合物或药物组合物用于治疗细胞、组织、器官、动物或受试者中的疾病或病症的用途,如本领域已知的或如本文所述的,使用所公开的组合物和药物组合物,例如用治疗有效量的组合物或药物组合物施用或接触细胞、组织、器官、动

物或受试者。在一个方面，受试者是哺乳动物。优选地，受试者是人。术语“受试者”和“患者”在本文中可互换使用。

[0288] 本公开内容提供了用于调节或治疗细胞、组织、器官、动物或受试者中的至少一种恶性疾病或病症的方法。优选地，恶性疾病是癌症。恶性疾病或病症的非限制性实例包括白血病、急性白血病、急性成淋巴细胞性白血病(ALL)、急性淋巴细胞性白血病、B细胞、T细胞或FAB ALL、急性髓样白血病(AML)、急性粒细胞性白血病、慢性髓细胞白血病(CML)、慢性淋巴细胞性白血病(CLL)、多毛细胞白血病、骨髓增生异常综合征(MDS)、淋巴瘤、何杰金氏病、恶性淋巴瘤、非何杰金氏淋巴瘤、伯基特氏淋巴瘤、多发性骨髓瘤、卡波西氏肉瘤、结肠直肠癌、胰腺癌、鼻咽癌、恶性组织细胞增生症、副肿瘤综合征/恶性肿瘤的高钙血症、实体瘤、膀胱癌、乳腺癌、结肠直肠癌、子宫内膜癌、头癌、颈癌、遗传性非息肉病性癌、何杰金氏淋巴瘤、肝癌、肺癌、非小细胞肺癌、卵巢癌、胰腺癌、前列腺癌、肾细胞癌、睾丸癌、腺癌、肉瘤、恶性黑色素瘤、血管瘤、转移性疾病、癌症有关的骨吸收、癌症有关的骨痛等等。

[0289] 在优选的方面，恶性疾病或病症的治疗包括过继细胞疗法。例如，在一个方面，本公开内容提供了表达至少一种公开的蛋白质支架和/或包含蛋白质支架的CAR (例如，用本公开内容的组合物递送至细胞的scFv、单结构域抗体、Centyrin)的修饰的细胞，其已进行选择和/或扩增用于施用于有需要的受试者。修饰的细胞可以配制用于在任何温度包括室温和体温下贮存。修饰的细胞可以配制用于冷冻贮存和随后的解冻。修饰的细胞可以在药学上可接受的载体中进行配制，用于从无菌包装中直接施用于受试者。修饰的细胞可以在具有细胞活力和/或CAR表达水平的指示剂的药学上可接受的载体中进行配制，以确保细胞功能和CAR表达的最低水平。修饰的细胞可以用一种或多种试剂以规定的密度在药学上可接受的载体中进行配制，以抑制进一步扩增和/或预防细胞死亡。

[0290] 任何可以包括将有效量的本文公开的任何组合物或药物组合物施用于需要此类调节、治疗或疗法的细胞、组织、器官、动物或受试者。此类方法可以任选地进一步包括用于治疗此类疾病或病症的共施用或组合疗法，其中本文公开的任何组合物或药物组合物的施用进一步包括之前、同时和/或之后施用至少一种化学治疗剂(例如，烷化剂、有丝分裂抑制剂、放射性药物)。

[0291] 在一些方面，受试者在施用之后并未发展移植物抗宿主(GvH)和/或宿主抗移植物(HvG)。在一个方面，施用是全身的。全身施用可以为本领域已知并在本文中详细描述的任何手段。优选地，全身施用是通过静脉内注射或静脉内输注。在一个方面，施用是局部的。局部施用可以为本领域已知并在本文中详细描述的任何手段。优选地，局部施用是通过瘤内注射或输注、脊柱内注射或输注、脑室内注射或输注、眼内注射或输注、或者骨内注射或输注。

[0292] 在一些方面，治疗有效剂量是单次剂量。在一些方面，单次剂量是同时制造的至少2、5、10、15、20、25、30、35、40、45、50、55、60、65、70、75、80、85、90、95、100、或两者之间的任何数目的剂量之一。在一些方面，当组合物是自体细胞或同种异体细胞时，剂量是足以使细胞植入和/或持续足够的时间以治疗疾病或病症的量。

[0293] 在一个实例中，本公开内容提供了治疗有需要的受试者中的癌症的方法，其包括向受试者施用组合物，所述组合物包含蛋白质支架或包含蛋白质支架的CAR (例如，scFv、单结构域抗体、Centyrin)，所述抗体或CAR与肿瘤细胞上的抗原特异性结合。在其中组合物

包含修饰的细胞或细胞群体的方面，细胞或细胞群体可以为自体的或同种异体的。

[0294] 在本文描述的治疗方法的一些方面，可以修改或终止治疗。具体而言，在其中用于治疗的组合物包含诱导型促凋亡多肽的方面，可以通过使细胞与诱导剂接触，在细胞中选择性地诱导细胞凋亡。可以响应例如恢复的征兆或疾病严重性/进展降低的征兆、疾病缓解/停止的征兆、和/或不良事件的发生来修改或终止治疗。在一些方面，该方法包括施用诱导剂的抑制剂以抑制细胞疗法的修改的步骤，从而恢复细胞疗法的功能和/或功效(例如，当疾病的征兆或症状再次出现、或者严重性增加和/或不良事件得到解决时)。

[0295] 蛋白质支架生产、筛选和纯化

本公开内容的至少一种蛋白质支架(例如，单克隆抗体、嵌合抗体、单结构域抗体、VHH、VH、单链可变片段(scFv)、Centyrin、抗原结合片段(Fab)或Fab片段)，可以任选地由细胞系、混合细胞系、永生化细胞或永生化细胞的克隆群体产生，如本领域众所周知的。参见例如，Ausubel等人，编辑，Current Protocols in Molecular Biology，John Wiley & Sons，Inc.，NY，N.Y. (1987-2001)；Sambrook等人，Molecular Cloning: A Laboratory Manual，第2版，Cold Spring Harbor，N.Y. (1989)；Harlow和Lane，Antibodies，a Laboratory Manual，Cold Spring Harbor，N.Y. (1989)；Colligan等人，编辑，Current Protocols in Immunology，John Wiley & Sons，Inc.，NY (1994-2001)；Colligan等人，Current Protocols in Protein Science，John Wiley & Sons，NY，N.Y.，(1997-2001)。

[0296] 可以改变、添加和/或缺失来自蛋白质支架的氨基酸，以减少免疫原性，或者减少、增强或修饰结合、亲和力、结合速率、解离速率、亲合力、特异性、半衰期、稳定性、溶解性或任何其它合适的特性，如本领域已知的。

[0297] 任选地，蛋白质支架可以进行改造，伴随对于抗原的高亲和力和其它有利的生物学性质的保留。为了实现这一目标，可以通过使用亲本序列和改造序列的三维模型，分析亲本序列和各种概念性改造产物的过程，来任选地制备支架蛋白质。三维模型是普遍可用的并且是本领域技术人员熟悉的。计算机程序是可获得的，其示出且展示所选候选序列的可能的三维构象结构，并且可以测量可能的免疫原性(例如，Monrovia，Calif.的Xencor，Inc.的Immunofilter程序)。这些展示的检查允许分析残基在候选序列的功能中的可能作用，即，分析影响候选蛋白质支架结合其抗原的能力的残基。以这种方式，可以从亲本序列和参考序列中选择且组合残基，使得实现所需的特性，例如对于靶抗原的亲和力。可替代地或除上述程序之外，可以使用其它合适的改造方法。

[0298] 可以使用核苷酸(DNA或RNA展示)或肽展示文库，例如在体外展示中，方便地实现与相似蛋白质或片段特异性结合的蛋白质支架的筛选。该方法涉及对于具有所需功能或结构的各个成员筛选大型肽集合。展示的核苷酸或肽序列的长度可以为3至5000个或更多个核苷酸或氨基酸，频繁为5-100个氨基酸长，且经常为约8至25个氨基酸长。除用于生成肽文库的直接化学合成方法之外，还已描述了几种重组DNA方法。一种类型涉及在细菌噬菌体或细胞的表面上的肽序列展示。每种细菌噬菌体或细胞含有编码特定展示肽序列的核苷酸序列。此类方法在PCT专利公开号WO 91/17271、WO 91/18980、WO 91/19818和WO 93/08278中进行描述。

[0299] 用于生成肽文库的其它系统具有体外化学合成和重组方法两者的方面。参见PCT专利公开号WO 92/05258、WO 92/14843和WO 96/19256。还参见美国专利号5,658,754；和5,

643,768。肽展示文库、载体和筛选试剂盒从此类供应商如Invitrogen (Carlsbad、Calif.)和Cambridge Antibody Technologies (Cambridgeshire、UK)商购可得。参见例如,转让给Enzon的美国专利号4,704,692、4,939,666、4,946,778、5,260,203、5,455,030、5,518,889、5,534,621、5,656,730、5,763,733、5,767,260、5856456;转让给Dyax的5,223,409、5,403,484、5,571,698、5,837,500;转让给Affymax的5,427,908、5,580,717;转让给Cambridge Antibody Technologies的5,885,793;转让给Genentech的5,750,373;转让给Xoma、Colligan,同上;Ausubel,同上;或Sambrook,同上的5,618,920、5,595,898、5,576,195、5,698,435、5,693,493、5,698,417。

[0300] 本公开内容的蛋白质支架可以以宽范围的亲和力(KD)结合人或其它哺乳动物蛋白质。在一个优选的方面,本公开内容的至少一种蛋白质支架可以任选地以高亲和力结合靶蛋白,例如,其KD等于或小于约10^{-7} M,例如但不限于,0.1-9.9 (或其中的任何范围或值)X 10^{-8}、10^{-9}、10^{-10}、10^{-11}、10^{-12}、10^{-13}、10^{-14}、10^{-15}或者其中的任何范围或值,如通过表面等离振子共振或Kinexa方法确定的,如通过本领域技术人员实践的。

[0301] 可以使用任何合适的方法在实验上确定蛋白质支架对于抗原的亲和力或亲合力。(参见例如,Berzofsky等人,“Antibody-Antigen Interactions,” In Fundamental Immunology,Paul,W. E.,编辑,Raven Press: New York,N.Y. (1984);Kuby,Janis Immunology,W.H. Freeman and Company: New York,N.Y. (1992);以及本文所述的方法)。如果在不同条件(例如盐浓度、pH)下测量,则所测量的特定蛋白质支架-抗原相互作用的亲和力可以不同。因此,亲和力和其它抗原结合参数(例如,KD、Kon、Koff)的测量优选用蛋白质支架和抗原的标准化溶液,以及标准化缓冲液例如本文所述的缓冲液来进行。

[0302] 可以用蛋白质支架执行竞争性测定法,以便确定哪些蛋白质、抗体和其它拮抗剂与蛋白质支架竞争结合靶蛋白和/或共享表位区域。如本领域普通技术人员容易知道的这些测定法,评估了拮抗剂或配体之间对于蛋白质上有限数目的结合位点的竞争。蛋白质和/或抗体在竞争之前或之后是固定或不溶解的,并且例如通过倾析(其中蛋白质/抗体是预先不溶解的)、或通过离心(其中蛋白质/抗体在竞争反应后是沉淀的),将与靶蛋白结合的样品与未结合的样品分开。另外,竞争性结合可以通过以下进行确定:功能是通过蛋白质支架与靶蛋白的结合还是结合的缺乏而改变,例如,蛋白质支架是抑制还是加强例如标记的酶促活性。可以使用ELISA和其它功能测定性,如本领域众所周知的。

[0303] 核酸分子

本公开内容的编码蛋白质支架的核酸分子可以为RNA的形式,例如mRNA、hnRNA、tRNA或任何其它形式,或者是DNA的形式,包括但不限于通过克隆获得或合成产生的cDNA和基因组DNA,或其任何组合。DNA可以为三链、双链或单链的,或其任何组合。DNA或RNA的至少一条链的任何部分可以为编码链,也称为有义链,或者它可以为非编码链,也被称为反义链。

[0304] 本公开内容的分离的核酸分子可以包括包含开放读码框(ORF)的核酸分子,任选地,具有一个或多个内含子,例如但不限于,至少一种蛋白质支架的至少一个指定部分;包含与靶蛋白结合的蛋白质支架或环区的编码序列的核酸分子;以及包含核苷酸序列的核酸分子,所述核苷酸序列与上文所述的那些基本上不同,但由于遗传密码的简并性,仍编码如本文所述和/或如本领域已知的蛋白质支架。当然,遗传密码是本领域众所周知的。因此,对

于本领域技术人员,生成编码本公开内容的特定蛋白质支架的此类简并核酸变体将是常规的。参见例如,Ausubel等人,同上,并且此类核酸变体包括在本公开内容中。

[0305] 如本文指示的,包含编码蛋白质支架的核酸的本公开内容的核酸分子,可以包括但不限于编码蛋白质支架片段的氨基酸序列本身的那些核酸分子;整个蛋白质支架或其一部分的编码序列;蛋白质支架、片段或一部分的编码序列,以及另外的序列,例如至少一种信号前导区或融合肽的编码序列,伴随或不伴随上述另外的编码序列,例如至少一个内含子,连同另外的非编码序列,包括但不限于非编码5'和3'序列,例如在转录、mRNA加工包括剪接和多聚腺苷酸化信号(例如,核糖体结合和mRNA的稳定性)中起作用的转录的非翻译序列;编码另外的氨基酸例如提供另外功能性的那些氨基酸的另外的编码序列。因此,可以将编码蛋白质支架的序列与标记物序列例如编码肽的序列融合,所述肽促进包含蛋白质支架片段或一部分的融合蛋白质支架的纯化。

[0306] 与如本文所述的多核苷酸选择性杂交的多核苷酸

本公开内容提供了在选择性杂交条件下与本文公开的多核苷酸杂交的分离的核酸。因此,多核苷酸可以用于分离、检测和/或定量包含此类多核苷酸的核酸。例如,本公开内容的多核苷酸可以用于鉴定、分离或扩增保藏文库中的部分或全长克隆。多核苷酸可以为从人或哺乳动物核酸文库中分离的基因组或cDNA序列,或者以其它方式与来自人或哺乳动物核酸文库的cDNA互补的基因组或cDNA序列。

[0307] 优选地,cDNA文库包含至少80%的全长序列,优选地,至少85%或90%的全长序列,并且更优选地,至少95%的全长序列。cDNA文库可以进行标准化,以增加稀有序列的表示。通常但非排他地对于相对于互补序列具有减少的序列同一性的序列采用低或中等严格性杂交条件。可以任选地对于具有更高同一性的序列采用中等和高严格性条件。低严格性条件允许具有约70%序列同一性的序列的选择性杂交,并且可以用于鉴定直向同源或旁系同源序列。

[0308] 任选地,多核苷酸将编码由本文所述的多核苷酸编码的蛋白质支架的至少一部分。多核苷酸包括可以用于与编码本公开内容的蛋白质支架的多核苷酸选择性杂交的核酸序列。参见例如,Ausubel,同上;Colligan,同上,各自整体引入本文作为参考。

[0309] 核酸的构建

本公开内容的分离的核酸可以使用(a)重组方法、(b)合成技术、(c)纯化技术和/或(d)其组合进行制备,如本领域众所周知的。

[0310] 核酸可以方便地包含除本公开内容的多核苷酸之外的序列。例如,可以将包含一个或多个内切核酸酶限制位点的多克隆位点插入核酸内,以帮助多核苷酸的分离。另外,可以插入可翻译序列,以帮助本公开内容的翻译的多核苷酸的分离。例如,六组氨酸标记物序列提供了方便的手段来纯化本公开内容的蛋白质。本公开内容的核酸,排除编码序列,任选地是用于克隆和/或表达本公开内容的多核苷酸的载体、衔接子或接头。

[0311] 可以将另外的序列加入这种克隆和/或表达序列中,以优化其在克隆和/或表达中的功能,以帮助多核苷酸的分离,或改善多核苷酸引入细胞内。克隆载体、表达载体、衔接子和接头的使用是本领域众所周知的。(参见例如,Ausubel,同上;或Sambrook,同上)。

[0312] 用于构建核酸的重组方法

本公开内容的分离的核酸组合物,例如RNA、cDNA、基因组DNA或其任何组合,可以

使用本领域技术人员已知的任何数目的克隆方法从生物来源中获得。在一些方面，在严格条件下，与本公开内容的多核苷酸选择性杂交的寡核苷酸探针用于鉴定cDNA或基因组DNA文库中的所需序列。RNA的分离以及cDNA和基因组文库的构建是本领域普通技术人员众所周知的。(参见例如，Ausubel，同上；或Sambrook，同上)。

[0313] 核酸筛选和分离方法

可以使用基于本公开内容的多核苷酸序列的探针来筛选cDNA或基因组文库。探针可以用于与基因组DNA或cDNA序列杂交，以分离相同或不同生物体中的同源基因。本领域技术人员将了解，可以在测定法中采用各种严格性程度的杂交；并且杂交或洗涤介质都可以为严格的。随着用于杂交的条件变得更严格，探针和靶之间必须存在更大程度的互补性用于双链体形成发生。严格性程度可以通过温度、离子强度、pH和部分变性溶剂如甲酰胺的存在中的一种或多种进行控制。例如，通过例如在0%至50%的范围内操纵甲酰胺浓度来改变反应物溶液的极性，方便地改变杂交的严格性。可检测结合所需的互补性(序列同一性)程度将依照杂交介质和/或洗涤介质的严格性而变。互补性程度最佳为100%，或70-100%，或者其中的任何范围或值。然而，应当理解，探针和引物中的微小序列变化可以通过减少杂交和/或洗涤介质的严格性进行补偿。

[0314] RNA或DNA的扩增方法是本领域众所周知的，并且可以基于本文呈现的教导和指导，无需过度实验，根据本公开内容而使用。

[0315] 已知的DNA或RNA扩增方法包括但不限于聚合酶链反应(PCR)和有关的扩增过程(参见例如，给予Mullis等人的美国专利号4,683,195、4,683,202、4,800,159、4,965,188；给予Tabor等人的4,795,699和4,921,794；给予Innis的5,142,033；给予Wilson等人的5,122,464；给予Innis 的5,091,310；给予Gyllensten等人的5,066,584；给予Gellensten等人的4,889,818；给予Silver等人的4,994,370，给予Biswas的4,766,067；给予Ringold的4,656,134)，以及RNA介导的扩增，其使用针对靶序列的反义RNA作为用于双链DNA合成的模板(给予Malek等人的美国专利号5,130,238，具有商品名NASBA)，所述参考文献的全部内容引入本文作为参考。(参见例如，Ausubel，同上；或Sambrook，同上。)

例如，聚合酶链反应(PCR)技术可以用于直接从基因组DNA或cDNA文库中扩增本公开内容的多核苷酸和有关基因的序列。PCR和其它体外扩增方法也可以是有用的，例如，以克隆编码待表达的蛋白质的核酸序列，制备核酸以用作探针用于检测样品中所需mRNA的存在、用于核酸测序或用于其它目的。足以指导技术人员完成体外扩增方法的技术实例在以下中找到：Berger，同上，Sambrook，同上和Ausubel，同上，以及Mullis等人，美国专利号4,683,202 (1987)；以及Innis等人，PCR Protocols A Guide to Methods and Applications，编辑，Academic Press Inc.，San Diego，Calif. (1990)。用于基因组PCR扩增的商购可得的试剂盒是本领域已知的。参见例如，Advantage-GC Genomic PCR Kit (Clontech)。另外，例如，T4基因32蛋白(Boehringer Mannheim)可以用于改善长PCR产物的产率。

[0316] 用于构建核酸的合成方法

本公开内容的分离的核酸也可以通过经由已知方法的直接化学合成来制备(参见例如，Ausubel等人，同上)。化学合成一般产生单链寡核苷酸，其可以通过与互补序列杂交，或通过使用单链作为模板用DNA聚合酶的聚合，转换成双链DNA。本领域技术人员将认识到，

虽然DNA的化学合成可以限于约100个或更多个碱基的序列,但可以通过较短序列的连接来获得较长序列。

[0317] 重组表达盒

本公开内容进一步提供了包含本公开内容的核酸的重组表达盒。本公开内容的核酸序列,例如编码本公开内容的蛋白质支架的cDNA或基因组序列,可以用于构建可以引入至少一种所需宿主细胞内的重组表达盒。重组表达盒通常包含与转录起始调控序列可操作地连接的本公开内容的多核苷酸,所述转录起始调控序列将指导多核苷酸在预期的宿主细胞中的转录。异源和非异源(即内源)启动子两者均可以用于指导本公开内容的核酸的表达。

[0318] 在一些方面,充当启动子、增强子或其它元件的分离的核酸可以被引入本公开内容的多核苷酸的非异源形式的适当位置中(上游、下游或内含子中),以便上调或下调本公开内容的多核苷酸的表达。例如,内源性启动子可以通过突变、缺失和/或取代在体内或体外加以改变。

[0319] 表达载体和宿主细胞

本公开内容还涉及包括本公开内容的分离的核酸分子的载体、用重组载体遗传改造的宿主细胞、以及通过重组技术生产至少一种蛋白质支架,如本领域众所周知的。参见例如,Sambrook等人,同上;Ausubel等人,同上,各自整体引入本文作为参考。

[0320] 可以任选地将多核苷酸连接到含有可选择标记物的载体,用于在宿主中繁殖。一般地,质粒载体在沉淀物例如磷酸钙沉淀物、或具有荷电脂质的复合物中引入。如果载体是病毒,则它可以使用适当的包装细胞系在体外进行包装,然后转导到宿主细胞内。

[0321] DNA插入物应该与适当的启动子可操作地连接。表达构建体将进一步含有用于转录起始、终止的位点,以及在转录区域中,用于翻译的核糖体结合位点。由构建体表达的成熟转录物的编码部分将优选包括在开始处的翻译起始、以及在待翻译的mRNA的端部处适当地放置的终止密码子(例如UAA、UGA或UAG),其中UAA和UAG优选用于哺乳动物或真核细胞表达。

[0322] 表达载体将优选地但任选地包括至少一种可选择标记物。此类标记物包括例如但不限于用于真核细胞培养的氨苄青霉素、zeocin (*Sh bla*基因)、嘌呤霉素(*pac*基因)、潮霉素B (*hygB*基因)、G418/遗传霉素(*neo*基因)、DHFR (编码二氢叶酸还原酶并赋予对氨甲蝶呤的抗性)、霉酚酸或谷氨酰胺合成酶(GS,美国专利号5,122,464;5,770,359;5,827,739)、杀稻瘟菌素(*bsd*基因)抗性基因,以及用于在大肠杆菌(E. coli)和其它细菌或原核生物中培养的氨苄青霉素、zeocin (*Sh bla*基因)、嘌呤霉素(*pac*基因)、潮霉素B (*hygB*基因)、G418/遗传霉素(*neo*基因)、卡那霉素、壮观霉素、链霉素、羧苄青霉素、博莱霉素、红霉素、多粘菌素B或四环素抗性基因(上述专利在此整体引入作为参考)。用于上述宿主细胞的适当培养基和条件是本领域已知的。合适的载体对于技术人员将是显而易见的。将载体构建体引入宿主细胞内可以通过磷酸钙转染、DEAE-葡聚糖介导的转染、阳离子脂质介导的转染、电穿孔、转导、感染或其它已知方法来实现。此类方法在本领域中进行描述,例如Sambrook,同上,第1-4和16-18章;Ausubel,同上,第1、9、13、15、16章。

[0323] 表达载体将优选地但任选地包括至少一种可选择的细胞表面标记物,用于分离通过本公开内容的组合物和方法修饰的细胞。本公开内容的可选择的细胞表面标记物包含将

细胞或细胞子集与另一个限定的细胞子集区分开的表面蛋白质、糖蛋白或蛋白质组。优选地，可选择的细胞表面标记物将通过本公开内容的组合物或方法修饰的那些细胞与未通过本公开内容的组合物或方法修饰的那些细胞区分开。此类细胞表面标记物包括例如但不限于“指定簇”或“分类决定簇”蛋白(经常缩写为“CD”)，例如截短或全长形式的CD19、CD271、CD34、CD22、CD20、CD33、CD52或其任何组合。细胞表面标记物进一步包括自杀基因标记物RQR8 (Philip B等人Blood. 2014 Aug 21;124(8):1277-87)。

[0324] 表达载体将优选地但任选地包括至少一种可选择的药物抗性标记物，用于分离通过本公开内容的组合物和方法修饰的细胞。本公开内容的可选择的药物抗性标记物可以包含野生型或突变型Neo、DHFR、TYMS、FRANCF、RAD51C、GCS、MDR1、ALDH1、NKX2.2或其任何组合。

[0325] 本公开内容的至少一种蛋白质支架可以以修饰的形式例如融合蛋白进行表达，并且不仅可以包括分泌信号，而且还可以包括另外的异源功能区域。例如，另外的氨基酸特别是荷电氨基酸的区域，可以加入蛋白质支架的N末端，以改善在宿主细胞中、在纯化过程中或在随后的处理和贮存过程中的稳定性和持久性。另外，可以将肽部分加入本公开内容的蛋白质支架中，以促进纯化。可以在蛋白质支架或其至少一个片段的最终制备之前去除此类区域。此类方法在许多标准实验室手册中进行描述，所述实验室手册例如Sambrook，同上，第17.29-17.42和18.1-18.74章；Ausubel，同上，第16、17和18章。

[0326] 本领域普通技术人员知道可用于表达编码本公开内容的蛋白质的核酸的众多表达系统。可替代地，本公开内容的核酸可以通过在宿主细胞中开启(通过操纵)而在宿主细胞中表达，所述宿主细胞含有编码本公开内容的蛋白质支架的内源性DNA。此类方法是本领域众所周知的，例如，如整体引入本文作为参考的美国专利号5,580,734、5,641,670、5,733,746和5,733,761中所述的。

[0327] 可用于产生蛋白质支架、其指定部分或变体的细胞培养物的示例是如本领域已知的细菌、酵母和哺乳动物细胞。哺乳动物细胞系统经常是单层细胞的形式，尽管也可以使用哺乳动物细胞悬浮液或生物反应器。本领域已开发了能够表达完整的糖基化蛋白的许多合适的宿主细胞系，并且包括COS-1 (例如ATCC CRL 1650)、COS-7 (例如ATCC CRL-1651)、HEK293、BHK21 (例如ATCC CRL-10)、CHO (例如ATCC CRL 1610)和BSC-1 (例如ATCC CRL-26)细胞系，Cos-7细胞、CHO细胞、hep G2细胞、P3X63Ag8.653、SP2/0-Ag14、293细胞、HeLa细胞等等，其可从例如美国典型培养物保藏中心(American Type Culture Collection)，Manassas，Va. (www.atcc.org)容易获得。优选的宿主细胞包括淋巴样起源的细胞，例如骨髓瘤和淋巴瘤细胞。特别优选的宿主细胞是P3X63Ag8.653细胞(ATCC登录号CRL-1580)和SP2/0-Ag14细胞(ATCC登录号CRL-1851)。在一个优选的方面，重组细胞是P3X63Ab8.653或SP2/0-Ag14细胞。

[0328] 用于这些细胞的表达载体可以包括下述表达控制序列中的一种或多种，例如但不限于复制起点；启动子(例如，晚期或早期SV40启动子、CMV启动子(美国专利号5,168,062；5,385,839)、HSV tk启动子、pgk (磷酸甘油酸激酶)启动子、EF-1 α启动子(美国专利号5,266,491)、至少一种人启动子；增强子和/或加工信息位点，例如核糖体结合位点、RNA剪接位点、多聚腺苷酸化位点(例如，SV40大T Ag 多聚A添加位点)和转录终止子序列。参见例如，Ausubel等人，同上；Sambrook等人，同上。可用于生产本公开内容的核酸或蛋白质的其

它细胞是已知的，和/或例如可得自美国典型培养物保藏中心的细胞系和杂交瘤目录(www.atcc.org)或者其它已知或商业来源。

[0329] 当采用真核宿主细胞时，通常将多聚腺苷酸花或转录终止子序列引入载体内。终止子序列的实例是来自牛生长激素基因的多聚腺苷酸化序列。还可以包括用于转录物的准确剪接的序列。剪接序列的实例是来自SV40的VP1内含子(Sprague等人，J. Virol. 45:773-781 (1983))。另外，如本领域已知的，可以将控制宿主细胞中的复制的基因序列引入载体内。

[0330] 蛋白质支架纯化

可以通过众所周知的方法，从重组细胞培养物中回收且纯化蛋白质支架，所述方法包括但不限于蛋白A纯化、硫酸铵或乙醇沉淀、酸提取、阴离子或阳离子交换层析、磷酸纤维素层析、疏水性相互作用层析、亲和层析、羟磷灰石层析和凝集素层析。高效液相层析(“HPLC”)也可以用于纯化。参见例如，Colligan，Current Protocols in Immunology，或Current Protocols in Protein Science，John Wiley & Sons，NY，N.Y.，(1997-2001)，例如，第1、4、6、8、9、10章，各自整体引入本文作为参考。

[0331] 本公开内容的蛋白质支架包括纯化的产物、化学合成程序的产物和通过重组技术从原核或真核宿主产生的产物，所述宿主包括例如大肠杆菌、酵母、高等植物、昆虫和哺乳动物细胞。取决于重组生产程序中采用的宿主，本公开内容的蛋白质支架可以为糖基化的或可以为非糖基化的。此类方法在许多标准实验室手册中进行描述，所述实验室手册例如Sambrook，同上，节段17.37-17.42；Ausubel，同上，第10、12、13、16、18和20章，Colligan，Protein Science，同上，第12-14章，全部整体引入本文作为参考。

[0332] 氨基酸密码

构成本公开内容的蛋白质支架的氨基酸经常是缩写的。氨基酸命名可以通过经由其单字母代码、其三字母代码、名称或三核苷酸密码子指定氨基酸得到指示，如本领域众所周知的(参见Alberts，B.等人，Molecular Biology of The Cell，第三版，Garland Publishing，Inc.，New York，1994)。本公开内容的蛋白质支架可以包括来自自发或突变和/或人为操纵的一个或多个氨基酸取代、缺失或添加，如本文指定的。可以通过本领域已知的方法，例如定点诱变或丙氨酸扫描诱变(例如，Ausubel，同上，第8、15章；Cunningham和Wells，Science 244:1081-1085 (1989))，来鉴定本公开内容的蛋白质支架中对于功能必需的氨基酸。后一程序在分子中的每一个残基处引入单个丙氨酸突变。然后测试所得到的突变型分子的生物活性，例如但不限于至少一种中和活性。对于蛋白质支架结合关键的位点也可以通过结构分析进行鉴定，所述结构分析例如结晶、核磁共振或光亲和标记(Smith等人，J. Mol. Biol. 224:899-904 (1992)和de Vos等人，Science 255:306-312 (1992))。

[0333] 如本领域技术人员将了解的，本公开内容包括本公开内容的至少一种生物活性蛋白质支架。生物活性蛋白质支架具有的比活性是天然(非合成)、内源性或有关和已知的蛋白质支架的比活性的至少20%、30%或40%，且优选至少50%、60%或70%，且最优选至少80%、90%或95%-99%或更多。测定且定量酶促活性和底物特异性的方法是本领域技术人员众所周知的。

[0334] 在另一个方面，本公开内容涉及如本文所述的蛋白质支架和片段，其通过有机部

分的共价附着进行修饰。此类修饰可以产生具有改善的药代动力学性质(例如,增加的体内血清半衰期)的蛋白质支架片段。有机部分可以为线性或支链的亲水聚合物基团、脂肪酸基团或脂肪酸酯基团。在特定方面,亲水聚合物基团可以具有约800至约120,000道尔顿的分子量,并且可以为聚烷烃二醇(例如,聚乙二醇(PEG)、聚丙二醇(PPG))、碳水化合物聚合物、氨基酸聚合物或聚乙烯吡咯烷酮,并且脂肪酸或脂肪酸酯基团可以包含约八至约四十个碳原子。

[0335] 本公开内容的修饰的蛋白质支架和片段可以包含一个或多个有机部分,其直接或间接地与抗体共价键合。与本公开内容的蛋白质支架或片段键合的每个有机部分可以独立地是亲水聚合物基团、脂肪酸基团或脂肪酸酯基团。如本文使用的,术语"脂肪酸"涵盖一元羧酸和二元羧酸。如该术语在本文中使用的,"亲水聚合物基团"指在水中比在辛烷中更可溶的有机聚合物。例如,聚赖氨酸在水中比在辛烷中更可溶。因此,通过聚赖氨酸的共价附着修饰的蛋白质支架由本公开内容涵盖。适合于修饰本公开内容的蛋白质支架的亲水聚合物可以为线性或支链的,并且包括例如聚烷烃二醇(例如,PEG、单甲氧基-聚乙二醇(mPEG)、PPG等等)、碳水化合物(例如,葡聚糖、纤维素、寡糖、多糖等等)、亲水性氨基酸的聚合物(例如聚赖氨酸、聚精氨酸、聚天冬氨酸等等)、聚烷烃氧化物(例如聚环氧乙烷、聚环氧丙烷等等)和聚乙烯吡咯烷酮。优选地,修饰本公开内容的蛋白质支架的亲水聚合物作为分开的分子实体具有约800至约150,000道尔顿的分子量。例如,可以使用PEG5000和PEG20,000,其中下标是以道尔顿计的聚合物的平均分子量。亲水聚合物基团可以由1至约6个烷基、脂肪酸或脂肪酸酯基团取代。由脂肪酸或脂肪酸酯基团取代的亲水聚合物可以通过采用合适的方法进行制备。例如,包含胺基团的聚合物可以偶联到脂肪酸或脂肪酸酯的羧酸酯,并且脂肪酸或脂肪酸酯上的活化羧酸酯(例如,用N,N-羰基二咪唑激活)可以偶联到聚合物上的羟基。

[0336] 适合于修饰本公开内容的蛋白质支架的脂肪酸和脂肪酸酯可以为饱和的,或者可以包含一个或多个不饱和单元。适合于修饰本公开内容的蛋白质支架的脂肪酸包括例如正十二酸(C12,月桂酸)、正十四酸(C14,肉豆蔻酸)、正十八酸(C18,硬脂酸)、正二十酸(C20,花生酸)、正二十二酸(C22,山萮酸)、正三十酸(C30)、正四十酸(C40)、顺式-Δ9-十八酸(C18,油酸)、所有顺式Δ5,8,11,14-二十碳四烯酸C20,花生四烯酸)、辛二酸、十四烷二酸、十八烷二酸、二十二烷二酸等等。合适的脂肪酸酯包括包含直链或支链低级烷基的二羧酸单酯。低级烷基可以包含1至约12个,优选1至约6个碳原子。

[0337] 可以使用合适的方法,例如通过与一种或多种改性剂反应,来制备修饰的蛋白质支架和片段。如该术语在本文中使用的,"改性剂"指包含活化基团的合适的有机基团(例如,亲水聚合物、脂肪酸、脂肪酸酯)。"活化基团"是化学部分或官能团,其可以在适当条件下与第二化学基团反应,从而在改性剂和第二化学基团之间形成共价键。例如,胺反应性活化基团包括亲电子基团,例如甲苯磺酸酯、甲磺酸酯、卤素(氯、溴、氟、碘)、N-羟基琥珀酰亚胺酯(NHS)等等。可以与硫醇反应的活化基团包括例如马来酰亚胺、碘乙酰基、丙烯酰基、吡啶基二硫化物、5-硫醇-2-硝基苯甲酸硫醇(TNB-硫醇)等等。醛官能团可以与含胺或酰肼的分子偶联,并且叠氮基团可以与三价磷基团反应,以形成氨基磷酸酯或磷酰亚胺(phosphorimide)键合。将活化基团引入分子内的合适方法是本领域已知的(参见例如,Hermanson,G. T.,Bioconjugate Techniques,Academic Press: San Diego,Calif.

(1996))。活化基团可以直接与有机基团(例如,亲水聚合物、脂肪酸、脂肪酸酯)键合,或通过接头部分键合,所述接头部分例如其中一个或多个碳原子可以替换为杂原子例如氧、氮或硫的二价C1-C12基团。合适的接头部分包括例如四甘醇、—(CH2)3—、—NH—(CH2)6—NH—、—(CH2)2—NH—和—CH2—O—CH2—CH2—O—CH2—CH2—O—CH—NH—。包含接头部分的改性剂可以例如通过以下产生:在1-乙基-3-(3-二甲氨基丙基)碳二亚胺(EDC)的存在下,使单-Boc-烷基二胺(例如,单-Boc-乙二胺、单-Boc-二氨基己烷)与脂肪酸反应,以在游离胺和脂肪酸羧酸酯之间形成酰胺键。Boc保护基团可以通过用三氟乙酸(TFA)处理从产物中去除,以暴露可以偶联到另一种羧酸酯的伯胺,如所述的,或者可以与马来酸酐反应,并且使所得到的产物环化,以产生脂肪酸的活化的马来酰亚胺衍生物。(参见例如,Thompson等人,WO 92/16221,其全部教导引入本文作为参考。)

本公开内容的修饰的蛋白质支架可以通过使蛋白质支架或片段与改性剂反应来产生。例如,有机部分可以通过采用胺反应性改性剂,例如PEG的NHS酯,以非位点特异性方式与蛋白质支架键合。包含有机部分(其与本公开内容的蛋白质支架的特异性位点键合)的修饰的蛋白质支架和片段可以使用合适的方法进行制备,所述方法例如反向蛋白酶解(Fisch等人,Bioconjugate Chem.,3:147-153 (1992);Werlen等人,Bioconjugate Chem.,5:411-417 (1994);Kumaran等人,Protein Sci. 6 (10):2233-2241 (1997);Itoh等人,Bioorg. Chem.,24(1): 59-68 (1996);Capellas等人,Biotechnol. Bioeng.,56 (4):456-463 (1997)),以及Hermanson,G. T.,Bioconjugate Techniques,Academic Press:San Diego,Calif. (1996)中描述的方法。

[0338] 定义

如本公开内容自始至终使用的,单数形式“一个”、“一种”和“该/所述”包括复数所指物,除非上下文另有明确说明。因此,例如,提及“方法”包括多种此类方法,并且提及“剂量”包括提及本领域技术人员已知的一个或多个剂量及其等价物等等。

[0339] 术语“约”或“大约”意指在如通过本领域普通技术人员确定的特定值的可接受误差范围内,这将部分取决于如何测量或确定该值,例如,测量系统的局限性。例如,“约”可以意指在1个或多个标准差内。可替代地,“约”可以意指至多20%、或至多10%、或至多5%、或至多1%的给定值的范围。可替代地,特别是关于生物系统或过程,该术语可以意指在值的一个数量级内,优选在5倍内,且更优选在2倍内。当本申请和权利要求中描述特定值时,除非另有说明,否则术语“约”应该假定在特定值的可接受误差范围内的含义。

[0340] 本公开内容提供了分离的或基本上纯化的多核苷酸或蛋白质组合物。“分离的”或“纯化的”多核苷酸或蛋白质、或其生物活性部分,实质上或基本上不含如在其天然存在的环境中发现的,通常与多核苷酸或蛋白质伴随或相互作用的组分。因此,分离的或纯化的多核苷酸或蛋白质当通过重组技术生产时基本上不含其它细胞材料或培养基,或者当化学合成时基本上不含化学前体或其它化学制品。最佳地,“分离的”多核苷酸不含在多核苷酸衍生自其的生物体的基因组DNA中,天然侧接多核苷酸的序列(最佳地蛋白质编码序列) (即,位于多核苷酸的5’和3’端处的序列)。例如,在各个方面,分离的多核苷酸可以含有少于约5 kb、4 kb、3 kb、2 kb、1 kb、0.5 kb或0.1 kb的核苷酸序列,其天然侧接多核苷酸衍生自其的细胞的基因组DNA中的多核苷酸。基本上不含细胞材料的蛋白质包括具有小于约30%、20%、10%、5%或1% (按干重计)的污染蛋白质的蛋白质制剂。当重组生产本公开内容的蛋白

质或其生物活性部分时，最佳地，培养基代表少于约30%、20%、10%、5%或1% (按干重计)的化学前体或非目的蛋白质化学制品。

[0341] 本公开内容提供了所公开的DNA序列以及由这些DNA序列编码的蛋白质的片段和变体。在本公开内容自始至终使用的，术语“片段”指DNA序列的一部分或氨基酸序列的一部分以及因此由此编码的蛋白质。包含编码序列的DNA序列的片段可以编码蛋白质片段，其保留天然蛋白质的生物活性，并且因此保留如本文所述的对靶DNA序列的DNA识别或结合活性。可替代地，可用作杂交探针的DNA序列的片段一般不编码保留生物活性或不保留启动子活性的蛋白质。因此，DNA序列的片段的范围可以为至少约20个核苷酸、约50个核苷酸、约100个核苷酸，且直到本公开内容的全长多核苷酸。

[0342] 本公开内容的核酸或蛋白质可以通过模块化方法进行构建，所述模块化方法包括在靶载体中预组装单体单元和/或重复单元，所述靶载体随后可以组装成最终的目的载体。本公开内容的多肽可以包含本公开内容的重复单体，并且可以通过在靶载体中预组装重复单元通过模块化方法进行构建，所述靶载体随后可以组装成最终的目的载体。本公开内容提供了通过该方法产生的多肽以及编码这些多肽的核酸序列。本公开内容提供了宿主生物体和细胞，其包含编码这种模块化方法产生的多肽的核酸序列。

[0343] 术语“抗体”以最广泛的含义使用，并且具体地涵盖单克隆抗体(包括激动剂和拮抗剂抗体)和具有多表位特异性的抗体组合物。使用如本文定义的其抗体的天然或合成类似物、突变体、变体、等位基因、同源物和直向同源物(在本文中统称为“类似物”)也在其范围内。因此，根据其一个方面，术语“其抗体”在其最广泛的含义上也涵盖此类类似物。一般地，在此类类似物中，与如本文定义的其抗体相比，一个或多个氨基酸残基可能已被替换、缺失和/或添加。

[0344] 如本文使用的，“抗体片段”及其所有语法变体被定义为完整抗体的一部分，其包含完整抗体的抗原结合位点或可变区，其中所述部分不含完整抗体的Fc区的恒定重链结构域 (即CH2、CH3和CH4，取决于抗体同种型)。抗体片段的实例包括Fab、Fab'、Fab'-SH、F(ab')$_2$和Fv片段；双抗体；其为具有一级结构的多肽的任何抗体片段，所述多肽由邻接氨基酸残基的一个不间断序列组成(在本文中被称为“单链抗体片段”或“单链多肽”)，包括但不限于(1)单链Fv (scFv)分子，(2)仅含有一个轻链可变结构域的单链多肽、或其含有轻链可变结构域的三个CDR的片段，而不含相关的重链部分，以及(3)仅含有一个重链可变区的单链多肽、或其含有重链可变区的三个CDR的片段，而不含相关的轻链部分；以及由抗体片段形成的多特异性或多价结构。在包含一条或多条重链的抗体片段中，重链可以含有在完整抗体的非Fc区中发现的任何恒定结构域序列(例如，IgG同种型中的CHI)，和/或可以含有在完整抗体中发现的任何铰链区序列，和/或可以含有与重链的铰链区序列或恒定结构域序列融合或位于其中的亮氨酸拉链序列。该术语进一步包括单结构域抗体(“sdAB”)，其一般指具有单个单体可变抗体结构域的抗体片段(例如，来自骆驼科动物)。此类抗体片段类型通过本领域普通技术人员将容易理解。

[0345] “结合”指大分子之间(例如，蛋白质和核酸之间)的序列特异性、非共价相互作用。结合相互作用的并非所有组分都需要是序列特异性的(例如，与DNA骨架中的磷酸残基接触)，只要相互作用整体上是序列特异性的。

[0346] 术语“包含”预期意指组合物和方法包括所叙述的要素，但不排除其它要素。当用

于定义组合物和方法时，“基本上由……组成”应该意指排除当用于预期目的时，对组合具有任何重要意义的其它要素。因此，基本上由如本文定义的要素组成的组合物将不排除痕量杂质或惰性载体。“由……组成”应该意指排除多于痕量要素的其它成分和基本的方法步骤。由这些过渡术语各自定义的方面在本公开内容的范围内。

[0347] 术语“表位”指多肽的抗原决定簇。表位可以包含空间构象中的三个氨基酸，其为表位所独有的。一般地，表位由至少4、5、6或7个此类氨基酸组成，并且更通常地，由至少8、9或10个此类氨基酸组成。确定氨基酸的空间构象的方法是本领域已知的，并且包括例如x射线晶体学和二维核磁共振。

[0348] 如本文使用的，“表达”指多核苷酸通过其转录成mRNA的过程，和/或转录的mRNA通过其随后被翻译成肽、多肽或蛋白质的过程。如果多核苷酸衍生自基因组DNA，则表达可以包括在真核细胞中的mRNA剪接。

[0349] “基因表达”指将包含在基因中的信息转换成基因产物。基因产物可以为基因的直接转录产物(例如，mRNA、tRNA、rRNA、反义RNA、核酶、shRNA、微小RNA、结构RNA或任何其它类型的RNA)，或通过mRNA的翻译产生的蛋白质。基因产物还包括通过过程如加帽、多聚腺苷酸化、甲基化和编辑进行修饰的RNA，以及通过例如甲基化、乙酰化、磷酸化、泛素化、ADP-核糖基化、肉豆蔻化和糖基化进行修饰的蛋白质。

[0350] 基因表达的“调节”或“调控”指基因活性的变化。表达的调节可以包括但不限于基因激活和基因阻遏。

[0351] 术语“可操作地连接的(operatively linked)”或其等价物(例如，“可操作地连接的(linked operatively)”)意指两个或更多个分子相对于彼此定位，使得它们能够相互作用以实现可归于一个或两个分子或其组合的功能。

[0352] 公开了非共价连接的组分以及制备且使用非共价连接的组分的方法。各种组分可以采取如本文所述的各种不同形式。例如，可以使用非共价连接的(即，可操作地连接的)蛋白质，以允许暂时的相互作用，其避免了本领域中的一个或多个问题。非共价连接的组分例如蛋白质结合和解离的能力，仅或主要在其中此类结合对于所需活性需要的情况下实现功能性结合。键合可以具有足以允许所需效应的持续时间。

[0353] 公开了用于将蛋白质导向生物体的基因组中的特异性基因座的方法。该方法可以包括提供DNA定位组分和提供效应分子的步骤，其中所述DNA定位组分和效应分子能够经由非共价键合可操作地连接。

[0354] 术语“scFv”指单链可变片段。scFv是用接头肽相连接的免疫球蛋白的重链(VH)和轻链(VL)可变区的融合蛋白。接头肽的长度可以为约5至40个氨基酸，或约10至30个氨基酸，或约5、10、15、20、25、30、35或40个氨基酸。单链可变片段缺少完整抗体分子中发现的恒定Fc区，并且因此缺少用于纯化抗体的常见结合位点(例如蛋白G)。该术语进一步包括其为细胞内抗体(其为在细胞的细胞质中稳定的抗体)，并且可以与细胞内蛋白质结合的scFv。

[0355] 术语“单结构域抗体”意指具有能够选择性结合特异性抗原的单个单体可变抗体结构域的抗体片段。单结构域抗体一般是约110个氨基酸长的肽链，包含重链抗体或常见IgG的一个可变结构域(VH)，其一般具有与完整抗体相似的对抗原的亲和力，但更耐热且针对洗涤剂和高浓度尿素更稳定。实例是衍生自骆驼科动物的抗体或鱼抗体。可替代地，单结构域抗体可以由具有四条链的常见鼠或人IgG制成。

[0356] 如本文使用的，术语“特异性结合(specifically bind)”和“特异性结合(specific binding)”指抗体、抗体片段或纳米抗体优先结合存在于不同抗原的均质混合物中的特定抗原的能力。在一些方面，特异性结合相互作用将区别样品中的期望抗原和不期望抗原。在一些方面，多于约10倍至100倍或更多(例如，多于约1000倍或10,000倍)。“特异性”指免疫球蛋白或免疫球蛋白片段例如纳米抗体相对于不同的抗原靶优先结合一种抗原靶的能力，并不一定暗示高亲和力。

[0357] “靶位点”或“靶序列”是定义结合分子与之结合的核酸的一部分的核酸序列，条件是存在用于结合的充分条件。

[0358] 术语“核酸”或“寡核苷酸”或“多核苷酸”指共价连接在一起的至少两个核苷酸。单链的描述也定义了互补链的序列。因此，核酸也可以涵盖所描绘的单链的互补链。本公开内容的核酸还涵盖保留相同结构或编码相同蛋白质的基本上等同的核酸及其互补体。

[0359] 本公开内容的探针可以包含可以在严格杂交条件下与靶序列杂交的单链核酸。因此，本公开内容的核酸可以指在严格杂交条件下杂交的探针。

[0360] 本公开内容的核酸可以为单链或双链。即使当分子的大部分是单链时，本公开内容的核酸也可以含有双链序列。即使当分子的大部分是双链时，本公开内容的核酸也可以含有单链序列。本公开内容的核酸可以包括基因组DNA、cDNA、RNA或其杂交体。本公开内容的核酸可以含有脱氧核糖核苷酸和核糖核苷酸的组合。本公开内容的核酸可以含有碱基的组合，所述碱基包括尿嘧啶、腺嘌呤、胸腺嘧啶、胞嘧啶、鸟嘌呤、肌苷、黄嘌呤、次黄嘌呤、异胞嘧啶和异鸟嘌呤。可以合成本公开内容的核酸，以包含非天然氨基酸修饰。本公开内容的核酸可以通过化学合成方法或通过重组方法来获得。

[0361] 本公开内容的核酸，或其整个序列、或其任何部分，可以为非天然存在的。本公开内容的核酸可以含有并非天然存在的一个或多个突变、取代、缺失或插入，致使整个核酸序列非天然存在。本公开内容的核酸可以含有一个或多个一式两份的、反向的或重复的序列，其所得的序列并非天然存在的，致使整个核酸序列非天然存在。本公开内容的核酸可以含有并非天然存在的修饰的、人工的或合成的核苷酸，致使整个核酸序列非天然存在。

[0362] 考虑到遗传密码的冗余，多个核苷酸序列可以编码任何特定的蛋白质。本文考虑了所有此类核苷酸序列。

[0363] 如本公开内容自始至终使用的，术语“可操作地连接的”指基因的表达，所述基因处于它与之在空间上相连接的启动子的控制下。启动子可以置于处于其控制下的基因的5’(上游)或3’(下游)。启动子和基因之间的距离可以与衍生启动子的基因中该启动子和它控制的基因之间的距离大致相同。可以适应启动子和基因之间的距离变化，而无启动子功能的丧失。

[0364] 如本公开内容自始至终使用的，术语“启动子”指能够赋予、激活或增强细胞中的核酸表达的合成或天然衍生的分子。启动子可以包含一个或多个特异性转录调控序列，以进一步增强其表达和/或改变空间表达和/或时间表达。启动子还可以包含远端增强子或阻遏元件，其可以位于距转录起始位点多达数千个碱基对处。启动子可以衍生自包括病毒、细菌、真菌、植物、昆虫和动物的来源。启动子可以关于表达在其中发生的细胞、组织或器官，或者关于表达在其下发生的发育阶段，或者响应外部刺激如生理应激、病原体、金属离子或诱导剂，组成性或差异性地调控基因组分的表达。启动子的代表性实例包括细菌噬菌体T7

启动子、细菌噬菌体T3启动子、SP6启动子、lac操纵子-启动子、tac启动子、SV40晚期启动子、SV40早期启动子、RSV-LTR启动子、CMV IE启动子、EF-1 α启动子、CAG启动子、SV40早期启动子或SV40晚期启动子和CMV IE启动子。

[0365] 如本公开内容自始至终使用的，术语“基本上互补的”指在8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、30、35、40、45、50、55、60、65、70、75、80、85、90、95、100、180、270、360、450、540个或更多个核苷酸或氨基酸的区域上，第一序列与第二序列的互补体至少60%、65%、70%、75%、80%、85%、90%、95%、97%、98%或99%等同，或者两个序列在严格杂交条件下杂交。

[0366] 如本公开内容自始至终使用的，术语“基本上等同的”指在8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、30、35、40、45、50、55、60、65、70、75、80、85、90、95、100、180、270、360、450、540个或更多个核苷酸或氨基酸的区域上，第一序列和第二序列至少60%、65%、70%、75%、80%、85%、90%、95%、97%、98%或99%等同，或就核酸而言，如果第一序列与第二序列的互补体基本上互补。

[0367] 如本公开内容自始至终使用的，当用于描述核酸时，术语“变体”指(i)参考核苷酸序列的一部分或片段；(ii)参考核苷酸序列或其一部分的互补体；(iii)与参考核酸或其互补体基本上等同的核酸；或(iv)在严格条件下与参考核酸、其互补体或基本上等同于其的序列杂交的核酸。

[0368] 如本公开内容自始至终使用的，术语“载体”指含有复制起点的核酸序列。载体可以为病毒载体、细菌噬菌体、细菌人工染色体或酵母人工染色体。载体可以为DNA或RNA载体。载体可以为自复制的染色体外载体，且优选为DNA质粒。载体可以包含氨基酸与DNA序列、RNA序列、或DNA和RNA序列两者的组合。

[0369] 如本公开内容自始至终使用的，当用于描述肽或多肽时，术语“变体”指这样的肽或多肽，其通过氨基酸的插入、缺失或保守取代而在氨基酸序列中不同，但保留至少一种生物活性。变体还可以意指具有氨基酸序列的蛋白质，其与保留至少一种生物活性的具有氨基酸序列的参考蛋白质基本上等同。

[0370] 氨基酸的保守取代，即用具有相似性质(例如，亲水性、荷电区域的程度和分布)的不同氨基酸替换氨基酸，在本领域中被认为通常涉及微小变化。如本领域理解的，通过考虑氨基酸的亲水指数，可以部分地鉴定这些微小变化。Kyte等人，J. Mol. Biol. 157: 105-132 (1982)。氨基酸的亲水指数基于其疏水性和电荷的考虑。具有相似亲水指数的氨基酸可以被取代，且仍保留蛋白质功能。在一个方面，亲水指数为±2的氨基酸被取代。氨基酸的亲水性也可以用于揭示将导致蛋白质保留生物学功能的取代。在肽的背景下氨基酸的亲水性的考虑允许计算该肽的最大局部平均亲水性，已报告与抗原性和免疫原性良好关联的一种有用度量。美国专利号4,554,101，完全引入本文作为参考。

[0371] 具有相似亲水性值的氨基酸的取代可以导致保留生物活性，例如免疫原性的肽。可以用亲水性值在彼此±2内的氨基酸来执行取代。氨基酸的疏水性指数和亲水性值两者均受该氨基酸的特定侧链的影响。与该观察一致，与生物学功能相容的氨基酸取代被理解为取决于氨基酸且特别是这些氨基酸的侧链的相对相似性，如通过疏水性、亲水性、电荷、大小和其它性质所揭示的。

[0372] 如本文使用的，“保守”氨基酸取代可以如下表A、B或C中阐述的进行定义。在一些

方面，融合多肽和/或编码此类融合多肽的核酸包括已通过修饰编码本公开内容的多肽的多核苷酸引入的保守取代。氨基酸可以根据物理性质以及对二级蛋白质结构和三级蛋白质结构的贡献进行分类。保守取代是用一种氨基酸取代具有相似性质的另一种氨基酸。示例性的保守取代在表A中阐述。

[0373] 表A -- 保守取代I

侧链特性		氨基酸
脂肪族的	非极性的	GAPILVF
	极性的 - 非荷电的	CSTMNQ
	极性的 – 荷电的	DEKR
芳香族的		HFWY
其它		NQDE

可替代地，如表B中阐述的，保守氨基酸可以如Lehninger，(Biochemistry，第二版;Worth Publishers，Inc. NY，N.Y. (1975)，第71-77页)中所述进行分组。

[0374] 表B -- 保守取代II

侧链特性		氨基酸
非极性(疏水的)	脂肪族的:	ALIVP
	芳香族的:	FWY
	含硫的:	M
	边界的:	GY
非荷电的-极性的	羟基:	STY
	酰胺:	NQ
	巯基:	C
	边界的:	GY
带正电的(碱性的):		KRH
带负电的(酸性的):		DE

可替代地，示例性保守取代在表C中阐述。

[0375] 表C -- 保守取代III

原始残基	示例性取代
Ala (A)	Val Leu Ile Met
Arg (R)	Lys His
Asn (N)	Gln

Asp (D)	Glu
Cys (C)	Ser Thr
Gln (Q)	Asn
Glu (E)	Asp
Gly (G)	Ala Val Leu Pro
His (H)	Lys Arg
Ile (I)	Leu Val Met Ala Phe
Leu (L)	Ile Val Met Ala Phe
Lys (K)	Arg His
Met (M)	Leu Ile Val Ala
Phe (F)	Trp Tyr Ile
Pro (P)	Gly Ala Val Leu Ile
Ser (S)	Thr
Thr (T)	Ser
Trp (W)	Tyr Phe Ile
Tyr (Y)	Trp Phe Thr Ser
Val (V)	Ile Leu Met Ala

应当理解,本公开内容的多肽预期包括荷有氨基酸残基的一个或多个插入、缺失或取代或其任何组合,以及除氨基酸残基的插入、缺失或取代外的修饰的多肽。本公开内容的多肽或核酸可以含有一个或多个保守取代。

[0376] 如本公开内容自始至终使用的,术语“多于一个”上述氨基酸取代指2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19或20个或更多个所述氨基酸取代。术语“多于一个”可以指2、3、4或5个所述氨基酸取代。

[0377] 本公开内容的多肽和蛋白质,或其整个序列、或其任何部分,可以为非天然存在的。本公开内容的多肽和蛋白质可以含有并非天然存在的一个或多个突变、取代、缺失或插入,致使整个氨基酸序列非天然存在。本公开内容的多肽和蛋白质可以含有一个或多个一式两份的、反向的或重复的序列,其所得的序列并非天然存在的,致使整个氨基酸序列非天然存在。本公开内容的多肽和蛋白质可以含有并非天然存在的修饰的、人工的或合成的氨基酸,致使整个氨基酸序列非天然存在。

[0378] 如本公开内容自始至终使用的,“序列同一性”可以通过使用独立的可执行BLAST引擎程序用于比对(blast)两个序列(bl2seq)进行确定,所述程序可以从美国国家生物技术信息中心(National Center for Biotechnology Information) (NCBI) ftp网站检索,使用缺省参数(Tatusova和Madden,FEMS Microbiol Lett.,1999,174,247-250;其整体引入本文作为参考)。当在两个或更多个核酸或多肽序列的背景下使用时,术语“等同的”或“同一性”指在每个序列的指定区域上相同的残基的指定百分比。百分比可以通过以下进行计算:最佳地比对两个序列,在指定区域上比较两个序列,确定等同残基在两个序列中出现的位置数目,以得到匹配位置的数目,将匹配位置的数目除以指定区域中的位置总数目,并且将结果乘以100,以得到序列同一性的百分比。在其中两个序列具有不同长度,或比对产生一个或多个交错端部且指定的比较区域仅包括单个序列的情况下,单个序列的残基包括

在计算的分母而不是分子中。当比较DNA和RNA时，胸腺嘧啶(T)和尿嘧啶(U)可以被视为等价的。可以手动地或通过使用计算机序列算法如BLAST或BLAST 2.0来执行同一性。

[0379] 如本公开内容自始至终使用的，术语“内源的”指与靶基因或它引入其内的宿主细胞天然相关的核酸或蛋白质序列。

[0380] 如本公开内容自始至终使用的，术语“外源的”指与靶基因或它引入其内的宿主细胞并非天然相关的核酸或蛋白质序列，包括天然存在的核酸例如DNA序列的非天然存在的多重拷贝、或位于非天然存在的基因组位置中的天然存在的核酸序列。

[0381] 本公开内容提供了将包含DNA序列的多核苷酸构建体引入宿主细胞内的方法。“引入”意指以这样的方式将多核苷酸构建体呈递给细胞，使得构建体获得对宿主细胞的内部的接近。本公开内容的方法并不依赖用于将多核苷酸构建体引入宿主细胞内的特定方法，仅依赖于多核苷酸构建体获得对宿主的一种细胞的内部的接近。用于将多核苷酸构建体引入细菌、植物、真菌和动物内的方法是本领域已知的，包括但不限于稳定转化方法、瞬时转化方法和病毒介导的方法。

[0382] 实施例1:用PiggyBac纳米转座子改善的转座

评价了缩短piggyBac转座子质粒骨架，从而降低ITR之间的距离对进入人泛T细胞的基因组内的转座效率的作用。构建了完全尺寸的piggyBac质粒(FP)（图1）、piggyBac纳米转座子(NT)（图1)和piggyBac短纳米转座子(NTS)（图3)，各自包括编码GFP的转座子。FP骨架编码细菌pUC复制起点以及Kan/Neo抗性基因。NT骨架编码不含抗生素的蔗糖可选择的纳米质粒骨架，其包含RNA-OUT元件以及R6K微型复制起点。NTS与NT的不同之处在于:骨架RNA-OUT元件和R6K微型复制起点置于转座子元件的内部，从而进一步减少了ITR之间的距离。图1示出了完整质粒和NT纳米转座子之间的差异。虽然转座子的大小保持恒定(3,606 bp)，但NT骨架更短，有效地将ITR侧翼的距离从2,034 bp减少到493 bp（减少到小于1/4)，如表1中详述的。图3示出了piggyBac NT和piggyBac NTS之间的差异。虽然转座子的大小从3,614 bp增加到4,069（以掺入RNA-OUT和R6K序列)，但NTS骨架更短，有效地将ITR侧翼的距离从485 bp减少到48 bp（减少到小于1/10)，如表2中详述的。

[0383] 表1

	质粒大小	转座子大小	ITR侧翼的距离
PB-FP	5,640	3,606	2,034
PB-NT	4,099	3,606	493
改变倍数	1.38	1.00	4.13

表2

	质粒大小	转座子大小	ITR侧翼的距离
PB-NT	4,099	3,614	485
PB-NTS	4,117	4,069	48
改变倍数	1.00	0.89	10.10

经由电穿孔(EP)，将piggyBac FP或piggyBac NT伴随或不伴随编码Super piggyBac转座酶(SPB)的mRNA递送至人泛T细胞(图2)。另外，FP和NT以等摩尔或等质量的量递送至细胞。在EP之后，使用标准TCR激活试剂刺激细胞，并且15天后通过FACS评价GFP表达。这些数据显示了，当与FP GFP转座子质粒相比较时，将ITR侧翼的距离减少为小于1/4

(从2,034 bp到493 bp),导致在等摩尔和等质量的量两者下更高水平的GFP转座。另外,GFP表达是转座子稳定整合的结果,因为在不存在SPB的情况下电穿孔的T细胞中未检测到GFP表达。

[0384] 图1和图2证实了缩短piggyBac转座子质粒骨架增加了转座子进入人泛T细胞内的转座效率。然而,由于总质粒大小在FP和NT之间并不相同,因此尚不清楚的是通过NT增强的转座效率是更少的质粒(等摩尔;递送到细胞的DNA更少)、递送更多的总质粒 (等质量)、还是piggyBac ITR侧翼的距离更短的结果。由于递送至人泛T细胞的DNA可以引发免疫调节效应并且可以是毒性的,因此通过NT增强的转座效率可能是递送更少的总DNA (等摩尔)或递送更多的质粒(等质量)的结果。为了测试这一点,构建了NTS,其中纳米质粒骨架被重新定位到转座子内,放置在绝缘子和ITR之间(图3)。虽然NT和NTS的大小保持恒定(4,099和4,117 bp),但NTS中ITR侧翼的距离是1/10 (从485 bp到48 bp) (表2)。经由电穿孔(EP),将等摩尔/等质量的NT或NTS量连同编码SPB的mRNA一起递送至人泛T细胞。在EP之后,使用标准TCR激活试剂刺激细胞,并且15天后通过FACS评价GFP表达。这些数据显示了,当与NT GFP纳米转座子相比较时,将ITR侧翼的距离减少为小于1/10 (从485 bp到48 bp),同时保持总质粒大小不变,导致通过NTS更高水平的GFP转座(图4)。

[0385] 实施例2:BCMA CAR和PSMA CAR纳米转座子的转座

编码抗BCMA CAR和抗PSMA CAR的完全尺寸的piggyBac质粒(FP)或piggyBac纳米转座子(NT),经由电穿孔(EP)连同编码Super piggyBac转座酶的mRNA一起以等质量的量递送至人泛T细胞(图5)。在EP之后,在不存在选择试剂的情况下,使用标准TCR激活试剂刺激细胞,并且5天后通过FACS评价CAR表达。这些数据显示了,当与FP转座子质粒相比较时,两种NT均导致了在等质量的量下更高水平的转座。当CAR-T细胞由来自两个不同正常供体的人泛T细胞产生时,情况确实如此。如本文所述产生使用完全尺寸的piggyBac质粒(FP)或piggyBac纳米转座子(NT)产生的抗BCMA CAR和抗PSMA CAR T细胞。通过CAR-T细胞以指示的效应子/靶比率杀死被改造为表达BCMA (K562.BCMA)或PSMA (K562.PSMA)的K562细胞(图6)。这些数据显示了,无论是使用FP还是NT产生的所有CAR-T细胞,都能够以抗原依赖性方式杀死靶肿瘤细胞。对于由来自两个不同正常供体的人泛T细胞产生的CAR-T细胞,情况确实如此。图7是一系列图,其显示了使用抗BCMA CAR或抗PSMA CAR纳米转座子(NT)产生的人CAR-T细胞在表型组成方面是可比较的。如本文所述的,制造了使用完全尺寸的piggyBac质粒(FP)或piggyBac纳米转座子(NT)产生的抗BCMA CAR和抗PSMA CAR T细胞。执行了记忆T细胞标记物和激活/耗尽标记物(数据未显示)的表型分析。这些数据显示了,无论是使用FP还是NT产生的所有CAR-T细胞,都显示出以下的相似表型组成:CD45RA+CD62L+ (Tscm)、CD45RA-CD62L+ (Tcm)、CD45RA-CD62L- (Tem)和CD45RA+CD62L- (Teff)细胞。另外,观察到CCR7 (CD197)、CD127、CD27、LAG3、TIM3、CXCR3、PD-1和CD25的可比较的表达水平(数据未显示)。对于由来自两个不同正常供体的人泛T细胞产生的CAR-T细胞,情况确实如此。通过定量PCR测量整合转座子的平均拷贝数。这些数据显示了,在两个不同的供体中,无论是使用FP还是NT产生的所有CAR-T细胞,都显示出相似的转座子整合拷贝数(图8)。

[0386] 通过以不同比率混合高度多聚体批次(7%单体)与高度单体批次(87%单体),来产生以不同单体纯度的抗BCMA CAR piggyBac纳米转座子。两个批次均通过基因测序确认为在单体或多聚体结构中的一级水平下是等同的,并且仅在三级水平下不同。在不存在限制

性消化的情况下,将混合抗BCMA CAR NT的每个新批次在琼脂糖凝胶上运行,以揭示所得的单体/多聚体纳米转座子比率;生产了各种单体纯度的批次(7%、32%、45%、59%、65%、72%和87%)。在凝胶上,多聚体NT比单体NT迁移得更慢(图9)。图9中描绘的凝胶条带通过矩形加框,以示出多聚体(顶部)和单体(底部)纳米转座子;编号从上到下,从左到右进行:1 (多聚体)和2 (单体) [7%单体纯度]、3和4 [32%单体纯度]、5和6 [45%单体纯度]、7和8 [59%单体纯度]、9和10 [65%单体纯度]、11和12 [72%单体纯度]、13和14 [空白]、15和16 [87%单体纯度]。

[0387] 以不同单体纯度的抗BCMA CAR piggyBac NT,经由电穿孔(EP)连同编码Super piggyBac转座酶的mRNA一起递送至人泛T细胞(图10)。作为对照,还以等摩尔量递送以94%单体纯度的完全尺寸的抗BCMA CAR质粒(FP)。在EP之后,在不存在选择试剂的情况下,使用标准TCR激活试剂刺激细胞,并且5天后通过FACS评价CAR表达。这些数据显示了,在两个分开的供体(供体#3和供体#2)中,单体纯度积极影响转座效率。另外,这些数据显示了,当与等摩尔量的FP转座子质粒相比较时,NT导致更高水平的转座。

[0388] 实施例3:P-PSMA-101纳米转座子的临床前评估

当使用鼠异种移植模型,通过全长质粒(FLP)相对于纳米转座子(NT)以‘应激’剂量递送时,P-PSMA-101转座子的功效在临床前场景中进行评估。使用皮下(SC)注射到NSG小鼠内的表达萤光素酶的LNCaP细胞系(LNCaP.luc)的鼠异种移植模型,用于评价在来自两个不同的正常供体的总CAR-T细胞的两种不同‘应激’剂量(2.5x10^6或4x10^6)下,如通过全长质粒(FLP)或纳米转座子(NT)递送的P-PSMA-101转座子的体内抗肿瘤功效(图11)。使用或FLP或NT递送,使用P-PSMA-101转座子的piggyBac (PB)递送来产生所有CAR-T细胞。小鼠在腋窝中用LNCaP进行注射,并且当肿瘤确立时(通过卡尺测量100-200 mm^3)进行治疗。通过IV注射用两种不同‘应激’剂量($2.5x10^6$或$4x10^6$)的P-PSMA-101CAR-T治疗小鼠,用于在检测通过FLP和NT的转座子递送之间的功效方面可能的功能差异中的更大分辨率。对于对照小鼠(黑色)、供体#1 FLP小鼠(红色)、供体#1 NT小鼠(蓝色)、供体#2 FLP小鼠(橙色)和供体#2 NT小鼠(绿色),通过卡尺测量的肿瘤体积评价,如展示为具有误差条(顶部)和各个小鼠(底部)的组平均值(图12)。y轴显示了通过卡尺测量评价的肿瘤体积(mm^3)。x轴显示了在T细胞治疗后的天数。针对确立的SCLNCaP.luc实体瘤与FLP和对照小鼠相比,如通过卡尺测量的,通过NT递送的以‘应激’剂量的P-PSMA-101转座子证实了增强的抗肿瘤功效。

序列表

<110> 波赛达治疗公司

<120> 纳米转座子组合物及使用方法

<130> POTH-047/001WO (325002-2401)

<150> 62/783,133

<151> 2018-12-20

<150> 62/815,335

<151> 2019-03-07

<150> 62/815,845

<151> 2019-03-08

<160> 98

<170> PatentIn version 3.5

<210> 1

<211> 6198

<212> DNA

<213> 人工序列

<220>

<223> 表达 BCMA CARTyrin 的 P-BCMA-101 纳米转座子

<400> 1

[0001]

tgtacataga ttaaccctag aaagataatc atattgtgac gtacgttaaa gataatcatg 60
cgtaaaattg acgcatgtgt tttatcggtc tgtatatcga ggtttattta ttaatttgaa 120
tagatattaa gttttattat atttacactt acatactaat aataaattca acaaacaatt 180
tatttatgtt tatttattta ttaaaaaaaa acaaaaactc aaaatttctt ctataaagta 240
acaaaacttt tatcgaatac ctgcagcccg ggggatgcag agggacagcc ccccccaaaa 300
gcccccaggg atgtaattac gtccctcccc cgctaggggg cagcagcgag ccgcccgggg 360
ctccgctccg gtccggcgct cccccccgcat ccccgagccg gcagcgtgcg gggacagccc 420
gggcacgggg aaggtggcac gggatcgctt tcctctgaac gcttctcgct gctctttgag 480
cctgcagaca cctgggggga tacggggaaa agttgactgt gcctttcgat cgaaccatgg 540
acagttagct ttgcaaagat ggataaagtt ttaaacagag aggaatcttt gcagctaatg 600
gaccttctag gtcttgaaag gagtgggaat tggctccggt gcccgtcagt gggcagagcg 660
cacatcgccc acagtccccg agaagttggg gggaggggtc ggcaattgaa ccggtgccta 720
gagaaggtgg cgcggggtaa actgggaaag tgatgtcgtg tactggctcc gccttttcc 780
cgagggtggg ggagaaccgt atataagtgc agtagtcgcc gtgaacgttc tttttcgcaa 840
cgggtttgcc gccagaacac aggtaagtgc cgtgtgtggt tcccgcgggc ctggcctctt 900
tacgggttat ggcccttgcg tgccttgaat tacttccacc tggctgcagt acgtgattct 960
tgatcccgag cttcgggttg gaagtgggtg ggagagttcg aggccttgcg cttaaggagc 1020
cccttcgcct cgtgcttgag ttgaggcctg gcctgggcgc tggggccgcc gcgtgcgaat 1080
ctggtggcac cttcgcgcct gtctcgctgc tttcgataag tctctagcca tttaaaattt 1140
ttgatgacct gctgcgacgc tttttttctg gcaagatagt cttgtaaatg cgggccaaga 1200
tctgcacact ggtatttcgg tttttggggc cgcgggcggc gacggggccc gtgcgtccca 1260
gcgcacatgt tcggcgaggc ggggcctgcg agcgcggcca ccgagaatcg gacggggta 1320
gtctcaagct ggccggcctg ctctggtgcc tggcctcgcg ccgccgtgta tcgccccgcc 1380
ctgggcggca aggctggccc ggtcggcacc agttgcgtga gcggaaagat ggccgcttcc 1440
cggccctgct gcagggagct caaaatggag gacgcggcgc tcgggagagc gggcgggtga 1500

[0002]

gtcacccaca caaaggaaaa gggcctttcc gtcctcagcc gtcgcttcat gtgactccac 1560
ggagtaccgg gcgccgtcca ggcacctcga ttagttctcg agcttttgga gtacgtcgtc 1620
tttaggttgg ggggaggggt tttatgcgat ggagtttccc cacactgagt gggtggagac 1680
tgaagttagg ccagcttggc acttgatgta attctccttg gaatttgccc tttttgagtt 1740
tggatcttgg ttcattctca agcctcagac agtggttcaa agttttttc ttccatttca 1800
ggtgtcgtga gaattctaat acgactcact atagggtgtg ctgtctcatc attttggcaa 1860
agattggcca ccaagcttgc caccatgggg gtccaggtcg agactatttc accaggggat 1920
gggcgaacat ttccaaaaag gggccagact tgcgtcgtgc attacaccgg gatgctggag 1980
gacgggaaga aagtggacag ctccagggat cgcaacaagc ccttcaagtt catgctggga 2040
aagcaggaag tgatccgagg atgggaggaa ggcgtggcac agatgtcagt cggccagcgg 2100
gccaaactga ccattagccc tgactacgct tatggagcaa caggccaccc agggatcatt 2160
ccccctcatg ccaccctggt cttcgatgtg gaactgctga agctggaggg aggaggagga 2220
tccggatttg gggacgtggg ggccctggag tctctgcgag gaaatgccga tctggcttac 2280
atcctgagca tggaaccctg cggccactgt ctgatcatta acaatgtgaa cttctgcaga 2340
gaaagcggac tgcgaacacg gactggctcc aatattgact gtgagaagct gcggagaagg 2400
ttctctagtc tgcactttat ggtcgaagtg aaaggggatc tgaccgccaa gaaaatggtg 2460
ctggccctgc tggagctggc tcagcaggac catggagctc tggattgctg cgtggtcgtg 2520
atcctgtccc acgggtgcca ggcttctcat ctgcagttcc ccggagcagt gtacggaaca 2580
gacggctgtc ctgtcagcgt ggagaagatc gtcaacatct tcaacggcac ttcttgccct 2640
agtctggggg gaaagccaaa actgttcttt atccaggcct gtggcgggga acagaaagat 2700
cacggcttcg aggtggccag caccagccct gaggacgaat caccagggag caaccctgaa 2760
ccagatgcaa ctccattcca ggagggactg aggacctttg accagctgga tgctatctca 2820
agcctgccca ctcctagtga cattttcgtg tcttacagta ccttcccagg ctttgtctca 2880
tggcgcgatc ccaagtcagg gagctggtac gtggagacac tggacgacat ctttgaacag 2940
tgggcccatt cagaggacct gcagagcctg ctgctgcgag tggcaaacgc tgtctctgtg 3000
aagggcatct acaaacagat gcccgggtgc ttcaattttc tgagaaagaa actgttcttt 3060
aagacttccg gatctggaga gggaagggga agcctgctga cctgtggaga cgtggaggaa 3120
aacccaggac caatggcact gccagtcacc gccctgctgc tgcctctggc tctgctgctg 3180
cacgcagcta gaccaatgct gcctgcacca aagaacctgg tggtgagccg gatcacagag 3240
gactccgcca gactgtcttg gaccgcccct gacgccgcct tcgattcctt tccaatccgg 3300
tacatcgaga cactgatctg gggcgaggcc atctggctgg acgtgcccgg ctctgagagg 3360
agctacgatc tgacaggcct gaagcctggc accgagtatg cagtggtcat cacaggagtg 3420
aagggcggca ggttcagctc ccctctggtg gcctctttta ccacaaccac aacccctgcc 3480
cccagacctc ccacacccgc ccctaccatc gcgagtcagc ccctgagtct gagacctgag 3540
gcctgcaggc cagctgcagg aggagctgtg cacaccaggg gcctggactt cgcctgcgac 3600
atctacattt gggcaccact ggccgggacc tgtggagtgc tgctgctgag cctggtcatc 3660
acactgtact gcaagagagg caggaagaaa ctgctgtata ttttcaaaca gcccttcatg 3720
cgccccgtgc agactaccca ggaggaagac gggtgctcct gtcgattccc tgaggaagag 3780
gaaggcgggt gtgagctgcg cgtgaagttt agtcgatcag cagatgcccc agcttacaaa 3840
cagggacaga accagctgta taacgagctg aatctgggcc gccgagagga atatgacgtg 3900
ctggataagc ggagaggacg cgaccccgaa atgggaggca agcccaggcg caaaaaccct 3960
caggaaggcc tgtataacga gctgcagaag gacaaaatgg cagaagccta ttctgagatc 4020
ggcatgaagg gggagcgacg gagaggcaaa gggcacgatg ggctgtacca gggactgagc 4080
accgccacaa aggacaccta tgatgctctg catatgcagg cactgcctcc aaggggaagt 4140

[0003]

ggagaaggac gaggatcact gctgacatgc ggcgacgtgg aggaaaaccc tggcccaatg 4200
gtcgggtctc tgaattgtat cgtcgccgtg agtcagaaca tgggcattgg gaagaatggc 4260
gatttcccat ggccacctct gcgcaacgag tcccgatact ttcagcggat gacaactacc 4320
tcctctgtgg aagggaaaca gaatctggtc atcatgggaa agaaaacttg gttcagcatt 4380
ccagagaaga accggcccct gaaaggcaga atcaatctgg tgctgtcccg agaactgaag 4440
gagccaccac agggagctca ctttctgagc cggtccctgg acgatgcact gaagctgaca 4500
gaacagcctg agctggccaa caaagtcgat atggtgtgga tcgtcggggg aagttcagtg 4560
tataaggagg ccatgaatca ccccggccat ctgaaactgt tcgtcacacg gatcatgcag 4620
gactttgaga gcgatacttt ctttcctgaa attgacctgg agaagtacaa actgctgccc 4680
gaatatcctg gcgtgctgtc cgatgtccag gaagagaaag gcatcaaata caagttcgag 4740
gtctatgaga agaatgacta ataaggtacc gatcacatat gcctttaatt aaacactagt 4800
tctatagtgt cacctaaatt ccctttagtg agggttaatg gccgtaggcc gccagaattg 4860
ggtccagaca tgataagata cattgatgag tttggacaaa ccacaactag aatgcagtga 4920
aaaaaatgct ttatttgtga aatttgtgat gctattgctt tatttgtaac cattataagc 4980
tgcaataaac aagttaacaa caacaattgc attcatttta tgtttcaggt tcagggggag 5040
gtgtgggagg ttttttcgga ctctaggacc tgcgcatgcg cttggcgtaa tcatggtcat 5100
agctgtttcc tgttttcccc gtatcccccc aggtgtctgc aggctcaaag agcagcgaga 5160
agcgttcaga ggaaagcgat cccgtgccac cttccccgtg cccgggctgt ccccgcacgc 5220
tgccggctcg gggatgcggg gggagcgccg gaccggagcg gagccccggg cggctcgctg 5280
ctgcccccta gcgggggagg gacgtaatta catccctggg ggctttgggg gggggctgtc 5340
cctctcaccg cggtggagct ccagcttttg ttcgaattgg ggccccccct cgagggtatc 5400
gatgatatct ataacaagaa aatatatata taataagtta tcacgtaagt agaacatgaa 5460
ataacaatat aattatcgta tgagttaaat cttaaaagtc acgtaaaaga taatcatgcg 5520
tcattttgac tcacgcggtc gttatagttc aaaatcagtg acacttaccg cattgacaag 5580
cacgcctcac gggagctcca agcggcgact gagatgtcct aaatgcacag cgacggattc 5640
gcgctattta gaaagagaga gcaatatttc aagaatgcat gcgtcaattt tacgcagact 5700
atctttctag ggttaatcta gctagcctta agggcgcagc ccgcctaatg agcgggcttt 5760
tttttggctt gttgtccaca accgttaaac cttaaaagct ttaaaagcct tatatattct 5820
ttttttttctt ataaaactta aaaccttaga ggctatttaa gttgctgatt tatattaatt 5880
ttattgttca aacatgagag cttagtacgt gaaacatgag agcttagtac gttagccatg 5940
agagcttagt acgttagcca tgagggttta gttcgttaaa catgagagct tagtacgtta 6000
aacatgagag cttagtacgt actatcaaca ggttgaactg ctgatccacg ttgtggtaga 6060
attggtaaag agagtcgtgt aaaatatcga gttcgcacat cttgttgtct gattattgat 6120
ttttggcgaa accatttgat catatgacaa gatgtgtatc taccttaact taatgatttt 6180
gataaaaatc attaggta 6198

<210> 2
<211> 238
<212> DNA
<213> 人工序列
<220>
<223> ITR
<400> 2

ccctagaaag ataatcatat tgtgacgtac gttaaagata atcatgcgta aaattgacgc 60
atgtgtttta tcggtctgta tatcgaggtt tatttattaa tttgaataga tattaagttt 120

tattatattt acacttacat actaataata aattcaacaa acaatttatt tatgtttatt 180
tatttattaa aaaaaaacaa aaactcaaaa tttcttctat aaagtaacaa aactttta 238

<210> 3
<211> 232
<212> DNA
<213> 人工序列
<220>
<223> 绝缘子
<400> 3

gagggacagc ccccccccaa agcccccagg gatgtaatta cgtccctccc ccgctagggg 60
gcagcagcga gccgcccggg gctccgctcc ggtccggcgc tccccccgca tccccgagcc 120
ggcagcgtgc ggggacagcc cgggcacggg gaaggtggca cgggatcgct ttcctctgaa 180
cgcttctcgc tgctctttga gcctgcagac acctgggggg atacggggaa aa 232

<210> 4
<211> 1264
<212> DNA
<213> 人工序列
<220>
<223> EF1α 启动子
<400> 4

[0004]

agctttgcaa agatggataa agttttaaac agagaggaat ctttgcagct aatggacctt 60
ctaggtcttg aaaggagtgg gaattggctc cggtgcccgt cagtgggcag agcgcacatc 120
gcccacagtc cccgagaagt tgggggggagg ggtcggcaat tgaaccggtg cctagagaag 180
gtggcgcggg gtaaactggg aaagtgatgt cgtgtactgg ctccgccttt ttcccgaggg 240
tgggggagaa ccgtatataa gtgcagtagt cgccgtgaac gttcttttтc gcaacgggtt 300
tgccgccaga acacaggtaa gtgccgtgtg tggttcccgc gggcctggcc tctttacggg 360
ttatggccct tgcgtgcctt gaattacttc cacctggctg cagtacgtga ttcttgatcc 420
cgagcttcgg gttggaagtg ggtgggagag ttcgaggcct tgcgcttaag gagccccttc 480
gcctcgtgct tgagttgagg cctggcctgg gcgctggggc cgccgcgtgc gaatctggtg 540
gcaccttcgc gcctgtctcg ctgctttcga taagtctcta gccatttaaa atttttgatg 600
acctgctgcg acgctttttt tctggcaaga tagtcttgta aatgcgggcc aagatctgca 660
cactggtatt tcggtttttg gggccgcggg cggcgacggg gcccgtgcgt cccagcgcac 720
atgttcggcg aggcggggcc tgcgagcgcg gccaccgaga atcggacggg ggtagtctca 780
agctggccgg cctgctctgg tgcctggcct cgcgccgccg tgtatcgccc cgccctgggc 840
ggcaaggctg gcccggtcgg caccagttgc gtgagcggaa agatggccgc ttcccggccc 900
tgctgcaggg agctcaaaat ggaggacgcg gcgctcggga gagcgggcgg gtgagtcacc 960
cacacaaagg aaaagggcct ttccgtcctc agccgtcgct tcatgtgact ccacggagta 1020
ccgggcgccg tccaggcacc tcgattagtt ctcgagcttt tggagtacgt cgtctttagg 1080
ttgggggggag gggttttatg cgatggagtt tccccacact gagtgggtgg agactgaagt 1140
taggccagct tggcacttga tgtaattctc cttggaattt gccctttttg agtttggatc 1200
ttggttcatt ctcaagcctc agacagtggt tcaaagtttt tttcttccat ttcaggtgtc 1260
gtga 1264

<210> 5
<211> 1185

<212> DNA
<213> 人工序列
<220>
<223> 诱导型促凋亡多肽
<400> 5

[0005]

atgggggtcc aggtcgagac tatttcacca ggggatgggc gaacatttcc aaaaaggggc 60
cagacttgcg tcgtgcatta caccgggatg ctggaggacg ggaagaaagt ggacagctcc 120
agggatcgca acaagccctt caagttcatg ctgggaaagc aggaagtgat ccgaggatgg 180
gaggaaggcg tggcacagat gtcagtcggc cagcgggcca aactgaccat tagccctgac 240
tacgcttatg gagcaacagg ccacccaggg atcattcccc ctcatgccac cctggtcttc 300
gatgtggaac tgctgaagct ggagggagga ggaggatccg gatttgggga cgtgggggcc 360
ctggagtctc tgcgaggaaa tgccgatctg gcttacatcc tgagcatgga accctgcggc 420
cactgtctga tcattaacaa tgtgaacttc tgcagagaaa gcggactgcg aacacggact 480
ggctccaata ttgactgtga gaagctgcgg agaaggttct ctagtctgca ctttatggtc 540
gaagtgaaag gggatctgac cgccaagaaa atggtgctgg ccctgctgga gctggctcag 600
caggaccatg gagctctgga ttgctgcgtg gtcgtgatcc tgtcccacgg gtgccaggct 660
tctcatctgc agttccccgg agcagtgtac ggaacagacg gctgtcctgt cagcgtggag 720
aagatcgtca acatcttcaa cggcacttct tgccctagtc tgggggggaaa gccaaaactg 780
ttctttatcc aggcctgtgg cggggaacag aaagatcacg gcttcgaggt ggccagcacc 840
agccctgagg acgaatcacc agggagcaac cctgaaccag atgcaactcc attccaggag 900
ggactgagga cctttgacca gctggatgct atctcaagcc tgcccactcc tagtgacatt 960
ttcgtgtctt acagtacctt cccaggcttt gtctcatggc gcgatcccaa gtcagggagc 1020
tggtacgtgg agacactgga cgacatcttt gaacagtggg cccattcaga ggacctgcag 1080
agcctgctgc tgcgagtggc aaacgctgtc tctgtgaagg gcatctacaa acagatgccc 1140
gggtgcttca attttctgag aaagaaactg ttctttaaga cttcc 1185

<210> 6
<211> 1185
<212> DNA
<213> 人工序列
<220>
<223> 诱导型促凋亡多肽
<400> 6

atgggggtcc aggtcgagac tatttcacca ggggatgggc gaacatttcc aaaaaggggc 60
cagacttgcg tcgtgcatta caccgggatg ctggaggacg ggaagaaagt ggacagctcc 120
agggatcgca acaagccctt caagttcatg ctgggaaagc aggaagtgat ccgaggatgg 180
gaggaaggcg tggcacagat gtcagtcggc cagcgggcca aactgaccat tagccctgac 240
tacgcttatg gagcaacagg ccacccaggg atcattcccc ctcatgccac cctggtcttc 300
gatgtggaac tgctgaagct ggagggagga ggaggatccg gatttgggga cgtgggggcc 360
ctggagtctc tgcgaggaaa tgccgatctg gcttacatcc tgagcatgga accctgcggc 420
cactgtctga tcattaacaa tgtgaacttc tgcagagaaa gcggactgcg aacacggact 480
ggctccaata ttgactgtga gaagctgcgg agaaggttct ctagtctgca ctttatggtc 540
gaagtgaaag gggatctgac cgccaagaaa atggtgctgg ccctgctgga gctggctcag 600
caggaccatg gagctctgga ttgctgcgtg gtcgtgatcc tgtcccacgg gtgccaggct 660
tctcatctgc agttccccgg agcagtgtac ggaacagacg gctgtcctgt cagcgtggag 720

aagatcgtca acatcttcaa cggcacttct tgccctagtc tgggggaaa gccaaaactg 780
ttctttatcc aggcctgtgg cggggaacag aaagatcacg gcttcgaggt ggccagcacc 840
agccctgagg acgaatcacc agggagcaac cctgaaccag atgcaactcc attccaggag 900
ggactgagga cctttgacca gctggatgct atctcaagcc tgcccactcc tagtgacatt 960
ttcgtgtctt acagtacctt cccaggcttt gtctcatggc gcgatcccaa gtcagggagc 1020
tggtacgtgg agacactgga cgacatcttt gaacagtggg cccattcaga ggacctgcag 1080
agcctgctgc tgcgagtggc aaacgctgtc tctgtgaagg gcatctacaa acagatgccc 1140
gggtgcttca attttctgag aaagaaactg ttctttaaga cttcc 1185

<210> 7
<211> 63
<212> DNA
<213> 人工序列
<220>
<223> T2A 序列
<400> 7
ggatctggag agggaagggg aagcctgctg acctgtggag acgtggagga aaacccagga 60
cca 63

<210> 8
<211> 63
<212> DNA
<213> 人工序列
<220>
<223> T2A 序列
<400> 8
ggatctggag agggaagggg aagcctgctg acctgtggag acgtggagga aaacccagga 60
cca 63

[0006]

<210> 9
<211> 1002
<212> DNA
<213> 人工序列
<220>
<223> BCMA CARTyrin
<400> 9
atggcactgc cagtcaccgc cctgctgctg cctctggctc tgctgctgca cgcagctaga 60
ccaatgctgc ctgcaccaaa gaacctggtg gtgagccgga tcacagagga ctccgccaga 120
ctgtcttgga ccgcccctga cgccgccttc gattcctttc caatccggta catcgagaca 180
ctgatctggg gcgaggccat ctggctggac gtgcccggct ctgagaggag ctacgatctg 240
acaggcctga agcctggcac cgagtatgca gtggtcatca caggagtgaa gggcggcagg 300
ttcagctccc ctctggtggc ctcttttacc acaaccacaa cccctgcccc cagacctccc 360
acacccgccc ctaccatcgc gagtcagccc ctgagtctga gacctgaggc ctgcaggcca 420
gctgcaggag gagctgtgca caccaggggc ctggacttcg cctgcgacat ctacatttgg 480
gcaccactgg ccgggacctg tggagtgctg ctgctgagcc tggtcatcac actgtactgc 540
aagagaggca ggaagaaact gctgtatatt ttcaaacagc ccttcatgcg ccccgtgcag 600
actacccagg aggaagacgg gtgctcctgt cgattccctg aggaagagga aggcgggtgt 660

gagctgcgcg tgaagtttag tcgatcagca gatgccccag cttacaaaca gggacagaac 720
cagctgtata acgagctgaa tctgggccgc cgagaggaat atgacgtgct ggataagcgg 780
agaggacgcg accccgaaat gggaggcaag cccaggcgca aaaaccctca ggaaggcctg 840
tataacgagc tgcagaagga caaaatggca gaagcctatt ctgagatcgg catgaagggg 900
gagcgacgga gaggcaaagg gcacgatggg ctgtaccagg gactgagcac cgccacaaag 960
gacacctatg atgctctgca tatgcaggca ctgcctccaa gg 1002

<210> 10
<211> 63
<212> DNA
<213> 人工序列
<220>
<223> T2A 序列
<400> 10

ggaagtggag aaggacgagg atcactgctg acatgcggcg acgtggagga aaaccctggc 60
cca 63

<210> 11
<211> 561
<212> DNA
<213> 人工序列
<220>
<223> DHFR 选择
<400> 11

[0007]

atggtcgggt ctctgaattg tatcgtcgcc gtgagtcaga acatgggcat tgggaagaat 60
ggcgatttcc catggccacc tctgcgcaac gagtcccgat actttcagcg gatgacaact 120
acctcctctg tggaagggaa acagaatctg gtcatcatgg gaaagaaaac ttggttcagc 180
attccagaga agaaccggcc cctgaaaggc agaatcaatc tggtgctgtc ccgagaactg 240
aaggagccac cacagggagc tcactttctg agccggtccc tggacgatgc actgaagctg 300
acagaacagc ctgagctggc caacaaagtc gatatggtgt ggatcgtcgg gggaagttca 360
gtgtataagg aggccatgaa tcaccccggc catctgaaac tgttcgtcac acggatcatg 420
caggactttg agagcgatac tttctttcct gaaattgacc tggagaagta caaactgctg 480
cccgaatatc ctggcgtgct gtccgatgtc caggaagaga aaggcatcaa atacaagttc 540
gaggtctatg agaagaatga c 561

<210> 12
<211> 127
<212> DNA
<213> 人工序列
<220>
<223> 多聚 A sv40
<400> 12

cagacatgat aagatacatt gatgagtttg gacaaaccac aactagaatg cagtgaaaaa 60
aatgctttat ttgtgaaatt tgtgatgcta ttgctttatt tgtaaccatt ataagctgca 120
ataaaca 127

<210> 13
<211> 231

<212> DNA
<213> 人工序列
<220>
<223> 绝缘子
<400> 13
ttttccccgt atccccccag gtgtctgcag gctcaaagag cagcgagaag cgttcagagg 60
aaagcgatcc cgtgccacct tccccgtgcc cgggctgtcc ccgcacgctg ccggctcggg 120
gatgcggggg gagcgccgga ccggagcgga gccccgggcg gctcgctgct gccccctagc 180
gggggaggga cgtaattaca tccctggggg ctttgggggg gggctgtccc t 231
<210> 14
<211> 309
<212> DNA
<213> 人工序列
<220>
<223> ITR
<400> 14
gatatctata acaagaaaat atatatataa taagttatca cgtaagtaga acatgaaata 60
acaatataat tatcgtatga gttaaatctt aaaagtcacg taaaagataa tcatgcgtca 120
ttttgactca cgcggtcgtt atagttcaaa atcagtgaca cttaccgcat tgacaagcac 180

gcctcacggg agctccaagc ggcgactgag atgtcctaaa tgcacagcga cggattcgcg 240
ctatttagaa agagagagca atatttcaag aatgcatgcg tcaattttac gcagactatc 300
[0008] tttctaggg 309
<210> 15
<211> 281
<212> DNA
<213> 人工序列
<220>
<223> R6K 复制起点
<400> 15
ggcttgttgt ccacaaccgt taaaccttaa aagctttaaa agccttatat attcttttt 60
ttcttataaa acttaaaacc ttagaggcta tttaagttgc tgatttatat taattttatt 120
gttcaaacat gagagcttag tacgtgaaac atgagagctt agtacgttag ccatgagagc 180
ttagtacgtt agccatgagg gtttagttcg ttaaacatga gagcttagta cgttaaacat 240
gagagcttag tacgtactat caacaggttg aactgctgat c 281
<210> 16
<211> 139
<212> DNA
<213> 人工序列
<220>
<223> RNA-OUT
<400> 16
gtagaattgg taaagagagt cgtgtaaaat atcgagttcg cacatcttgt tgtctgatta 60
ttgatttttg gcgaaaccat ttgatcatat gacaagatgt gtatctacct taacttaatg 120

attttgataa aaatcatta 139

<210> 17

<211> 958

<212> PRT

<213> 人工序列

<220>

<223> P-BCMA-101 氨基酸序列

<400> 17

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

[0009]

Ser Gly Phe Gly Asp Val Gly Ala Leu Glu Ser Leu Arg Gly Asn Ala
115 120 125
Asp Leu Ala Tyr Ile Leu Ser Met Glu Pro Cys Gly His Cys Leu Ile
130 135 140
Ile Asn Asn Val Asn Phe Cys Arg Glu Ser Gly Leu Arg Thr Arg Thr
145 150 155 160
Gly Ser Asn Ile Asp Cys Glu Lys Leu Arg Arg Arg Phe Ser Ser Leu
165 170 175
His Phe Met Val Glu Val Lys Gly Asp Leu Thr Ala Lys Lys Met Val
180 185 190
Leu Ala Leu Leu Glu Leu Ala Gln Gln Asp His Gly Ala Leu Asp Cys
195 200 205
Cys Val Val Val Ile Leu Ser His Gly Cys Gln Ala Ser His Leu Gln
210 215 220
Phe Pro Gly Ala Val Tyr Gly Thr Asp Gly Cys Pro Val Ser Val Glu
225 230 235 240
Lys Ile Val Asn Ile Phe Asn Gly Thr Ser Cys Pro Ser Leu Gly Gly
245 250 255
Lys Pro Lys Leu Phe Phe Ile Gln Ala Cys Gly Gly Glu Gln Lys Asp
260 265 270
His Gly Phe Glu Val Ala Ser Thr Ser Pro Glu Asp Glu Ser Pro Gly
275 280 285

[0010]

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

[0011]

Lys Phe Ser Arg Ser Ala Asp Ala Pro Ala Tyr Lys Gln Gly Gln Asn
645 650 655
Gln Leu Tyr Asn Glu Leu Asn Leu Gly Arg Arg Glu Glu Tyr Asp Val
660 665 670
Leu Asp Lys Arg Arg Gly Arg Asp Pro Glu Met Gly Gly Lys Pro Arg
675 680 685
Arg Lys Asn Pro Gln Glu Gly Leu Tyr Asn Glu Leu Gln Lys Asp Lys
690 695 700
Met Ala Glu Ala Tyr Ser Glu Ile Gly Met Lys Gly Glu Arg Arg Arg
705 710 715 720
Gly Lys Gly His Asp Gly Leu Tyr Gln Gly Leu Ser Thr Ala Thr Lys
725 730 735
Asp Thr Tyr Asp Ala Leu His Met Gln Ala Leu Pro Pro Arg Gly Ser
740 745 750
Gly Glu Gly Arg Gly Ser Leu Leu Thr Cys Gly Asp Val Glu Glu Asn
755 760 765
Pro Gly Pro Met Val Gly Ser Leu Asn Cys Ile Val Ala Val Ser Gln
770 775 780
Asn Met Gly Ile Gly Lys Asn Gly Asp Phe Pro Trp Pro Pro Leu Arg
785 790 795 800
Asn Glu Ser Arg Tyr Phe Gln Arg Met Thr Thr Thr Ser Ser Val Glu
805 810 815
Gly Lys Gln Asn Leu Val Ile Met Gly Lys Lys Thr Trp Phe Ser Ile
820 825 830
Pro Glu Lys Asn Arg Pro Leu Lys Gly Arg Ile Asn Leu Val Leu Ser
835 840 845
Arg Glu Leu Lys Glu Pro Pro Gln Gly Ala His Phe Leu Ser Arg Ser
850 855 860
Leu Asp Asp Ala Leu Lys Leu Thr Glu Gln Pro Glu Leu Ala Asn Lys
865 870 875 880
Val Asp Met Val Trp Ile Val Gly Gly Ser Ser Val Tyr Lys Glu Ala
885 890 895
Met Asn His Pro Gly His Leu Lys Leu Phe Val Thr Arg Ile Met Gln
900 905 910
Asp Phe Glu Ser Asp Thr Phe Phe Pro Glu Ile Asp Leu Glu Lys Tyr
915 920 925
Lys Leu Leu Pro Glu Tyr Pro Gly Val Leu Ser Asp Val Gln Glu Glu
930 935 940
Lys Gly Ile Lys Tyr Lys Phe Glu Val Tyr Glu Lys Asn Asp
945 950 955

<210> 18
<211> 6204
<212> DNA
<213> 人工序列

<220>

<223> 表达 PSMA CARTyrin 的 P-PSMA-101 纳米转座子

<400> 18

[0012]

tgtacataga ttaaccctag aaagataatc atattgtgac gtacgttaaa gataatcatg 60
cgtaaaattg acgcatgtgt tttatcggtc tgtatatcga ggtttattta ttaatttgaa 120
tagatattaa gttttattat atttacactt acatactaat aataaattca acaaacaatt 180
tatttatgtt tatttattta ttaaaaaaaa acaaaaactc aaaatttctt ctataaagta 240
acaaaacttt tatcgaatac ctgcagcccg ggggatgcag agggacagcc ccccccaaaa 300
gcccccaggg atgtaattac gtccctcccc cgctaggggg cagcagcgag ccgcccgggg 360
ctccgctccg gtccggcgct cccccccgcat ccccgagccg gcagcgtgcg gggacagccc 420
gggcacgggg aaggtggcac gggatcgctt tcctctgaac gcttctcgct gctctttgag 480
cctgcagaca cctgggggga tacggggaaa agttgactgt gcctttcgat cgaaccatgg 540
acagttagct ttgcaaagat ggataaagtt ttaaacagag aggaatcttt gcagctaatg 600
gaccttctag gtcttgaaag gagtgggaat tggctccggt gcccgtcagt gggcagagcg 660
cacatcgccc acagtccccg agaagttggg gggaggggtc ggcaattgaa ccggtgccta 720
gagaaggtgg cgcggggtaa actgggaaag tgatgtcgtg tactggctcc gccttttcc 780
cgagggtggg ggagaaccgt atataagtgc agtagtcgcc gtgaacgttc tttttcgcaa 840
cgggtttgcc gccagaacac aggtaagtgc cgtgtgtggt tcccgcgggc ctggcctctt 900
tacgggttat ggcccttgcg tgccttgaat tacttccacc tggctgcagt acgtgattct 960
tgatcccgag cttcgggttg gaagtgggtg ggagagttcg aggccttgcg cttaaggagc 1020
cccttcgcct cgtgcttgag ttgaggcctg gcctgggcgc tggggccgcc gcgtgcgaat 1080
ctggtggcac cttcgcgcct gtctcgctgc tttcgataag tctctagcca tttaaaattt 1140
ttgatgacct gctgcgacgc tttttttctg gcaagatagt cttgtaaatg cgggccaaga 1200
tctgcacact ggtatttcgg tttttggggc cgcgggcggc gacggggccc gtgcgtccca 1260
gcgcacatgt tcggcgaggc ggggcctgcg agcgcggcca ccgagaatcg gacgggggta 1320
gtctcaagct ggccggcctg ctctggtgcc tggcctcgcg ccgccgtgta tcgccccgcc 1380
ctgggcggca aggctggccc ggtcggcacc agttgcgtga gcggaaagat ggccgcttcc 1440
cggccctgct gcagggagct caaaatggag gacgcggcgc tcgggagagc gggcgggtga 1500
gtcacccaca caaaggaaaa gggcctttcc gtcctcagcc gtcgcttcat gtgactccac 1560
ggagtaccgg gcgccgtcca ggcacctcga ttagttctcg agcttttgga gtacgtcgtc 1620
tttaggttgg ggggaggggt tttatgcgat ggagtttccc cacactgagt gggtggagac 1680
tgaagttagg ccagcttggc acttgatgta attctccttg gaatttgccc tttttgagtt 1740
tggatcttgg ttcattctca agcctcagac agtggttcaa agtttttttc ttccatttca 1800
ggtgtcgtga gaattctaat acgactcact atagggtgtg ctgtctcatc attttggcaa 1860
agattggcca ccaagcttgc caccatgggg gtccaggtcg agactatttc accaggggat 1920
gggcgaacat ttccaaaaag gggccagact tgcgtcgtgc attacaccgg gatgctggag 1980
gacgggaaga aagtggacag ctccagggat cgcaacaagc ccttcaagtt catgctggga 2040
aagcaggaag tgatccgagg atgggaggaa ggcgtggcac agatgtcagt cggccagcgg 2100
gccaaactga ccattagccc tgactacgct tatggagcaa caggccaccc agggatcatt 2160
ccccctcatg ccaccctggt cttcgatgtg gaactgctga agctggaggg aggaggagga 2220
tccggatttg gggacgtggg ggccctggag tctctgcgag gaaatgccga tctggcttac 2280
atcctgagca tggaaccctg cggccactgt ctgatcatta acaatgtgaa cttctgcaga 2340
gaaagcggac tgcgaacacg gactggctcc aatattgact gtgagaagct gcggagaagg 2400
ttctctagtc tgcactttat ggtcgaagtg aaaggggatc tgaccgccaa gaaaatggtg 2460

[0013]

ctggccctgc tggagctggc tcagcaggac catggagctc tggattgctg cgtggtcgtg 2520
atcctgtccc acgggtgcca ggcttctcat ctgcagttcc ccggagcagt gtacggaaca 2580
gacggctgtc ctgtcagcgt ggagaagatc gtcaacatct tcaacggcac ttcttgccct 2640
agtctggggg gaaagccaaa actgttcttt atccaggcct gtggcgggga acagaaagat 2700
cacggcttcg aggtggccag caccagccct gaggacgaat caccagggag caaccctgaa 2760
ccagatgcaa ctccattcca ggagggactg aggacctttg accagctgga tgctatctca 2820
agcctgccca ctcctagtga cattttcgtg tcttacagta ccttcccagg ctttgtctca 2880
tggcgcgatc ccaagtcagg gagctggtac gtggagacac tggacgacat ctttgaacag 2940
tgggcccatt cagaggacct gcagagcctg ctgctgcgag tggcaaacgc tgtctctgtg 3000
aagggcatct acaaacagat gcccgggtgc ttcaattttc tgagaaagaa actgttcttt 3060
aagacttccg gatctggaga gggaagggga agcctgctga cctgtggaga cgtggaggaa 3120
aacccaggac caatggcact gccagtcacc gccctgctgc tgcctctggc tctgctgctg 3180
cacgcagcta gaccaatgct gcctgcacca aagaacctgg tggtgtctcg ggtgaccgag 3240
gactctgcca gactgagctg ggccatcgac gagcagaggg attggttcga gagctttctg 3300
atccagtatc aggagtccga gaaagtgggc gaggccatcg tgctgacagt gcctggcagc 3360
gagcggtcct atgatctgac cggcctgaag ccaggcacag agtacaccgt gtccatctac 3420
ggcgtgtatc acgtgtacag gtccaatcct ctgtctgcca tcttcaccac aaccacaacc 3480
cctgccccca gacctcccac acccgccccc accatcgcga gtcagccccc gagtctgaga 3540
cctgaggcct gcaggccagc tgcaggagga gctgtgcaca ccaggggcct ggacttcgcc 3600
tgcgacatct acatttgggc accactggcc gggacctgtg gagtgctgct gctgagcctg 3660
gtcatcacac tgtactgcaa gagaggcagg aagaaactgc tgtatatttt caaacagccc 3720
ttcatgcgcc ccgtgcagac tacccaggag gaagacgggt gctcctgtcg attccctgag 3780
gaagaggaag gcgggtgtga gctgcgcgtg aagtttagtc gatcagcaga tgccccagct 3840
tacaaacagg gacagaacca gctgtataac gagctgaatc tgggccgccg agaggaatat 3900
gacgtgctgg ataagcggag aggacgcgac cccgaaatgg gaggcaagcc caggcgcaaa 3960
aaccctcagg aaggcctgta taacgagctg cagaaggaca aaatggcaga agcctattct 4020
gagatcggca tgaaggggga gcgacggaga ggcaaagggc acgatgggct gtaccaggga 4080
ctgagcaccg ccacaaagga cacctatgat gctctgcata tgcaggcact gcctccaagg 4140
ggaagtggag aaggacgagg atcactgctg acatgcggcg acgtggagga aaaccctggc 4200
ccaatggtcg ggtctctgaa ttgtatcgtc gccgtgagtc agaacatggg cattgggaag 4260
aatggcgatt tcccatggcc acctctgcgc aacgagtccc gatactttca gcggatgaca 4320
actacctcct ctgtggaagg gaaacagaat ctggtcatca tgggaaagaa aacttggttc 4380
agcattccag agaagaaccg gcccctgaaa ggcagaatca atctggtgct gtcccgagaa 4440
ctgaaggagc caccacaggg agctcacttt ctgagccggt ccctggacga tgcactgaag 4500
ctgacagaac agcctgagct ggccaacaaa gtcgatatgg tgtggatcgt cgggggaagt 4560
tcagtgtata aggaggccat gaatcacccc ggccatctga aactgttcgt cacacggatc 4620
atgcaggact ttgagagcga tactttcttt cctgaaattg acctggagaa gtacaaactg 4680
ctgcccgaat atcctggcgt gctgtccgat gtccaggaag agaaaggcat caaatacaag 4740
ttcgaggtct atgagaagaa tgactaataa ggtaccgatc acatatgcct ttaattaaac 4800
actagttcta tagtgtcacc taaattccct ttagtgaggg ttaatggccg taggccgcca 4860
gaattgggtc cagacatgat aagatacatt gatgagtttg gacaaaccac aactagaatg 4920
cagtgaaaaa aatgctttat ttgtgaaatt tgtgatgcta ttgctttatt tgtaaccatt 4980
ataagctgca ataaacaagt taacaacaac aattgcattc attttatgtt tcaggttcag 5040
ggggaggtgt gggaggtttt ttcggactct aggacctgcg catgcgcttg gcgtaatcat 5100

ggtcatagct gtttcctgtt ttccccgtat ccccccaggt gtctgcaggc tcaaagagca 5160
gcgagaagcg ttcagaggaa agcgatcccg tgccaccttc cccgtgcccg ggctgtcccc 5220
gcacgctgcc ggctcgggga tgcgggggga gcgccggacc ggagcggagc cccgggcggc 5280
tcgctgctgc cccctagcgg gggagggacg taattacatc cctgggggct ttgggggggg 5340
gctgtccctc tcaccgcggt ggagctccag cttttgttcg aattggggcc ccccctcgag 5400
ggtatcgatg atatctataa caagaaaata tatatataat aagttatcac gtaagtagaa 5460
catgaaataa caatataatt atcgtatgag ttaaatctta aaagtcacgt aaaagataat 5520
catgcgtcat tttgactcac gcggtcgtta tagttcaaaa tcagtgacac ttaccgcatt 5580
gacaagcacg cctcacggga gctccaagcg gcgactgaga tgtcctaaat gcacagcgac 5640
ggattcgcgc tatttagaaa gagagagcaa tatttcaaga atgcatgcgt caatttttacg 5700
cagactatct ttctagggtt aatctagcta gccttaaggg cgcagcccgc ctaatgagcg 5760
ggctttttttt tggcttgttg tccacaaccg ttaaacctta aaagctttaa aagccttata 5820
tattcttttt tttcttataa aacttaaaac cttagaggct atttaagttg ctgatttata 5880
ttaattttat tgttcaaaca tgagagctta gtacgtgaaa catgagagct tagtacgtta 5940
gccatgagag cttagtacgt tagccatgag ggtttagttc gttaaacatg agagcttagt 6000
acgttaaaca tgagagctta gtacgtacta tcaacaggtt gaactgctga tccacgttgt 6060
ggtagaattg gtaaagagag tcgtgtaaaa tatcgagttc gcacatcttg ttgtctgatt 6120
attgattttt ggcgaaacca tttgatcata tgacaagatg tgtatctacc ttaacttaat 6180
gattttgata aaaatcatta ggta 6204

<210> 19
<211> 1008
<212> DNA
<213> 人工序列
<220>
<223> PSMA CARTyrin
<400> 19

[0014]

atggcactgc cagtcaccgc cctgctgctg cctctggctc tgctgctgca cgcagctaga 60
ccaatgctgc ctgcaccaaa gaacctggtg gtgtctcggg tgaccgagga ctctgccaga 120
ctgagctggg ccatcgacga gcagagggat tggttcgaga gctttctgat ccagtatcag 180
gagtccgaga aagtgggcga ggccatcgtg ctgacagtgc ctggcagcga gcggtcctat 240
gatctgaccg gcctgaagcc aggcacagag tacaccgtgt ccatctacgg cgtgtatcac 300
gtgtacaggt ccaatcctct gtctgccatc ttcaccacaa ccacaacccc tgcccccaga 360
cctcccacac ccgcccctac catcgcgagt cagcccctga gtctgagacc tgaggcctgc 420
aggccagctg caggaggagc tgtgcacacc aggggcctgg acttcgcctg cgacatctac 480
atttgggcac cactggccgg gacctgtgga gtgctgctgc tgagcctggt catcacactg 540
tactgcaaga gaggcaggaa gaaactgctg tatattttca aacagccctt catgcgcccc 600
gtgcagacta cccaggagga agacgggtgc tcctgtcgat tccctgagga agaggaaggc 660
gggtgtgagc tgcgcgtgaa gtttagtcga tcagcagatg ccccagctta caaacaggga 720
cagaaccagc tgtataacga gctgaatctg ggccgccgag aggaatatga cgtgctggat 780
aagcggagag gacgcgaccc cgaaatggga ggcaagccca ggcgcaaaaa ccctcaggaa 840
ggcctgtata acgagctgca gaaggacaaa atggcagaag cctattctga gatcggcatg 900
aaggggggagc gacggagagg caaagggcac gatgggctgt accagggact gagcaccgcc 960
acaaaggaca cctatgatgc tctgcatatg caggcactgc ctccaagg 1008

<210> 20

<211> 960
<212> PRT
<213> 人工序列
<220>
<223> P-PSMA-101 氨基酸序列
<400> 20

[0015]

Met Gly Val Gln Val Glu Thr Ile Ser Pro Gly Asp Gly Arg Thr Phe
1 5 10 15
Pro Lys Arg Gly Gln Thr Cys Val Val His Tyr Thr Gly Met Leu Glu
20 25 30
Asp Gly Lys Lys Val Asp Ser Ser Arg Asp Arg Asn Lys Pro Phe Lys
35 40 45
Phe Met Leu Gly Lys Gln Glu Val Ile Arg Gly Trp Glu Glu Gly Val
50 55 60
Ala Gln Met Ser Val Gly Gln Arg Ala Lys Leu Thr Ile Ser Pro Asp
65 70 75 80
Tyr Ala Tyr Gly Ala Thr Gly His Pro Gly Ile Ile Pro Pro His Ala
85 90 95
Thr Leu Val Phe Asp Val Glu Leu Leu Lys Leu Glu Gly Gly Gly Gly
100 105 110
Ser Gly Phe Gly Asp Val Gly Ala Leu Glu Ser Leu Arg Gly Asn Ala
115 120 125
Asp Leu Ala Tyr Ile Leu Ser Met Glu Pro Cys Gly His Cys Leu Ile
130 135 140
Ile Asn Asn Val Asn Phe Cys Arg Glu Ser Gly Leu Arg Thr Arg Thr
145 150 155 160
Gly Ser Asn Ile Asp Cys Glu Lys Leu Arg Arg Arg Phe Ser Ser Leu
165 170 175
His Phe Met Val Glu Val Lys Gly Asp Leu Thr Ala Lys Lys Met Val
180 185 190
Leu Ala Leu Leu Glu Leu Ala Gln Gln Asp His Gly Ala Leu Asp Cys
195 200 205
Cys Val Val Val Ile Leu Ser His Gly Cys Gln Ala Ser His Leu Gln
210 215 220
Phe Pro Gly Ala Val Tyr Gly Thr Asp Gly Cys Pro Val Ser Val Glu
225 230 235 240
Lys Ile Val Asn Ile Phe Asn Gly Thr Ser Cys Pro Ser Leu Gly Gly
245 250 255
Lys Pro Lys Leu Phe Phe Ile Gln Ala Cys Gly Gly Glu Gln Lys Asp
260 265 270
His Gly Phe Glu Val Ala Ser Thr Ser Pro Glu Asp Glu Ser Pro Gly
275 280 285
Ser Asn Pro Glu Pro Asp Ala Thr Pro Phe Gln Glu Gly Leu Arg Thr
290 295 300

[0016]

```
Phe Asp Gln Leu Asp Ala Ile Ser Ser Leu Pro Thr Pro Ser Asp Ile
305                 310                 315                 320
Phe Val Ser Tyr Ser Thr Phe Pro Gly Phe Val Ser Trp Arg Asp Pro
                325                 330                 335
Lys Ser Gly Ser Trp Tyr Val Glu Thr Leu Asp Asp Ile Phe Glu Gln
            340                 345                 350
Trp Ala His Ser Glu Asp Leu Gln Ser Leu Leu Leu Arg Val Ala Asn
        355                 360                 365
Ala Val Ser Val Lys Gly Ile Tyr Lys Gln Met Pro Gly Cys Phe Asn
    370                 375                 380
Phe Leu Arg Lys Lys Leu Phe Phe Lys Thr Ser Gly Ser Gly Glu Gly
385                 390                 395                 400
Arg Gly Ser Leu Leu Thr Cys Gly Asp Val Glu Glu Asn Pro Gly Pro
                405                 410                 415
Met Ala Leu Pro Val Thr Ala Leu Leu Leu Pro Leu Ala Leu Leu Leu
            420                 425                 430
His Ala Ala Arg Pro Met Leu Pro Ala Pro Lys Asn Leu Val Val Ser
        435                 440                 445
Arg Val Thr Glu Asp Ser Ala Arg Leu Ser Trp Ala Ile Asp Glu Gln
    450                 455                 460
Arg Asp Trp Phe Glu Ser Phe Leu Ile Gln Tyr Gln Glu Ser Glu Lys
465                 470                 475                 480
Val Gly Glu Ala Ile Val Leu Thr Val Pro Gly Ser Glu Arg Ser Tyr
                485                 490                 495
Asp Leu Thr Gly Leu Lys Pro Gly Thr Glu Tyr Thr Val Ser Ile Tyr
            500                 505                 510
Gly Val Tyr His Val Tyr Arg Ser Asn Pro Leu Ser Ala Ile Phe Thr
        515                 520                 525
Thr Thr Thr Thr Pro Ala Pro Arg Pro Pro Thr Pro Ala Pro Thr Ile
    530                 535                 540
Ala Ser Gln Pro Leu Ser Leu Arg Pro Glu Ala Cys Arg Pro Ala Ala
545                 550                 555                 560
Gly Gly Ala Val His Thr Arg Gly Leu Asp Phe Ala Cys Asp Ile Tyr
                565                 570                 575
Ile Trp Ala Pro Leu Ala Gly Thr Cys Gly Val Leu Leu Leu Ser Leu
            580                 585                 590
Val Ile Thr Leu Tyr Cys Lys Arg Gly Arg Lys Lys Leu Leu Tyr Ile
        595                 600                 605
Phe Lys Gln Pro Phe Met Arg Pro Val Gln Thr Thr Gln Glu Glu Asp
    610                 615                 620
Gly Cys Ser Cys Arg Phe Pro Glu Glu Glu Glu Gly Gly Cys Glu Leu
625                 630                 635                 640
Arg Val Lys Phe Ser Arg Ser Ala Asp Ala Pro Ala Tyr Lys Gln Gly
                645                 650                 655
```

[0017]

Gln Asn Gln Leu Tyr Asn Glu Leu Asn Leu Gly Arg Arg Glu Glu Tyr
660 665 670
Asp Val Leu Asp Lys Arg Arg Gly Arg Asp Pro Glu Met Gly Gly Lys
675 680 685
Pro Arg Arg Lys Asn Pro Gln Glu Gly Leu Tyr Asn Glu Leu Gln Lys
690 695 700
Asp Lys Met Ala Glu Ala Tyr Ser Glu Ile Gly Met Lys Gly Glu Arg
705 710 715 720
Arg Arg Gly Lys Gly His Asp Gly Leu Tyr Gln Gly Leu Ser Thr Ala
725 730 735
Thr Lys Asp Thr Tyr Asp Ala Leu His Met Gln Ala Leu Pro Pro Arg
740 745 750
Gly Ser Gly Glu Gly Arg Gly Ser Leu Leu Thr Cys Gly Asp Val Glu
755 760 765
Glu Asn Pro Gly Pro Met Val Gly Ser Leu Asn Cys Ile Val Ala Val
770 775 780
Ser Gln Asn Met Gly Ile Gly Lys Asn Gly Asp Phe Pro Trp Pro Pro
785 790 795 800
Leu Arg Asn Glu Ser Arg Tyr Phe Gln Arg Met Thr Thr Thr Ser Ser
805 810 815
Val Glu Gly Lys Gln Asn Leu Val Ile Met Gly Lys Lys Thr Trp Phe
820 825 830
Ser Ile Pro Glu Lys Asn Arg Pro Leu Lys Gly Arg Ile Asn Leu Val
835 840 845
Leu Ser Arg Glu Leu Lys Glu Pro Pro Gln Gly Ala His Phe Leu Ser
850 855 860
Arg Ser Leu Asp Asp Ala Leu Lys Leu Thr Glu Gln Pro Glu Leu Ala
865 870 875 880
Asn Lys Val Asp Met Val Trp Ile Val Gly Gly Ser Ser Val Tyr Lys
885 890 895
Glu Ala Met Asn His Pro Gly His Leu Lys Leu Phe Val Thr Arg Ile
900 905 910
Met Gln Asp Phe Glu Ser Asp Thr Phe Phe Pro Glu Ile Asp Leu Glu
915 920 925
Lys Tyr Lys Leu Leu Pro Glu Tyr Pro Gly Val Leu Ser Asp Val Gln
930 935 940
Glu Glu Lys Gly Ile Lys Tyr Lys Phe Glu Val Tyr Glu Lys Asn Asp
945 950 955 960

<210> 21
<211> 6248
<212> DNA
<213> 人工序列
<220>
<223> 表达 BCMA VCAR 的 P-BCMA-ALLO1 纳米转座子

<400> 21

[0018]

tgtacataga ttaaccctag aaagataatc atattgtgac gtacgttaaa gataatcatg 60
cgtaaaattg acgcatgtgt tttatcggtc tgtatatcga ggtttattta ttaatttgaa 120
tagatattaa gttttattat atttacactt acatactaat aataaattca acaaacaatt 180
tatttatgtt tatttattta ttaaaaaaaa acaaaaactc aaaatttctt ctataaagta 240
acaaaacttt tatcgaatac ctgcagcccg ggggatgcag agggacagcc ccccccaaaa 300
gcccccaggg atgtaattac gtccctcccc cgctaggggg cagcagcgag ccgcccgggg 360
ctccgctccg gtccggcgct ccccccgcat ccccgagccg gcagcgtgcg gggacagccc 420
gggcacgggg aaggtggcac gggatcgctt tcctctgaac gcttctcgct gctctttgag 480
cctgcagaca cctgggggga tacggggaaa agttgactgt gcctttcgat cgaaccatgg 540
acagttagct ttgcaaagat ggataaagtt ttaaacagag aggaatcttt gcagctaatg 600
gaccttctag gtcttgaaag gagtgggaat tggctccggt gcccgtcagt gggcagagcg 660
cacatcgccc acagtccccg agaagttggg gggaggggtc ggcaattgaa ccggtgccta 720
gagaaggtgg cgcggggtaa actgggaaag tgatgtcgtg tactggctcc gccttttcc 780
cgagggtggg ggagaaccgt atataagtgc agtagtcgcc gtgaacgttc tttttcgcaa 840
cgggtttgcc gccagaacac aggtaagtgc cgtgtgtggt tcccgcgggc ctggcctctt 900
tacgggttat ggcccttgcg tgccttgaat tacttccacc tggctgcagt acgtgattct 960
tgatcccgag cttcgggttg gaagtgggtg ggagagttcg aggccttgcg cttaaggagc 1020
cccttcgcct cgtgcttgag ttgaggcctg gcctgggcgc tggggccgcc gcgtgcgaat 1080
ctggtggcac cttcgcgcct gtctcgctgc tttcgataag tctctagcca tttaaaattt 1140
ttgatgacct gctgcgacgc tttttttctg gcaagatagt cttgtaaatg cgggccaaga 1200
tctgcacact ggtatttcgg tttttggggc cgcgggcggc gacggggccc gtgcgtccca 1260
gcgcacatgt tcggcgaggc ggggcctgcg agcgcggcca ccgagaatcg gacgggggta 1320
gtctcaagct ggccggcctg ctctggtgcc tggcctcgcg ccgccgtgta tcgccccgcc 1380
ctgggcggca aggctggccc ggtcggcacc agttgcgtga gcggaaagat ggccgcttcc 1440
cggccctgct gcagggagct caaaatggag gacgcggcgc tcgggagagc gggcgggtga 1500
gtcacccaca caaaggaaaa gggcctttcc gtcctcagcc gtcgcttcat gtgactccac 1560
ggagtaccgg gcgccgtcca ggcacctcga ttagttctcg agcttttgga gtacgtcgtc 1620
tttaggttgg ggggaggggt tttatgcgat ggagtttccc cacactgagt gggtggagac 1680
tgaagttagg ccagcttggc acttgatgta attctccttg gaatttgccc tttttgagtt 1740
tggatcttgg ttcattctca agcctcagac agtggttcaa agttttttc ttccatttca 1800
ggtgtcgtga gaattctaat acgactcact atagggtgaa gcttgccacc atgggggtcc 1860
aggtggaaac aatctctccg ggggatgggc ggacattccc taaaaggggc cagacctgcg 1920
tggtgcatta caccggcatg ctggaagatg gcaagaaggt ggacagcagc cgggacagaa 1980
acaagccctt caagttcatg ctgggcaagc aagaagtgat cagaggctgg gaagagggcg 2040
tcgcccagat gtctgttgga cagagagcca agctgacaat cagccccgat tacgcctatg 2100
gcgccacagg acaccctggc atcattcctc cacatgccac actggtgttc gacgtggaac 2160
tgctgaagct ggaaggcggc ggaggatctg gctttggaga tgtgggagcc ctggaaagcc 2220
tgagaggcaa tgccgatctg gcctacatcc tgagcatgga accttgcggc cactgcctga 2280
ttatcaacaa cgtgaacttc tgtagagaga gcggcctgcg gaccagaacc ggcagcaata 2340
tcgattgcga gaagctgcgg cggagattca gcagcctgca cttcatggtg gaagtgaagg 2400
gcgacctgac cgccaagaaa atggtgctgg ctctgctgga actggcccag caagatcatg 2460
gcgccctgga ttgctgtgtg gtcgtgatcc tgtctcacgg ctgtcaggcc agccaccttc 2520
aattccctgg cgccgtgtat ggcacagatg gctgtcctgt gtccgtggaa aagatcgtga 2580

[0019]

acatcttcaa cggcaccagc tgtcctagcc tcggcggaaa gcccaagctg ttcttcatcc 2640
aagcctgtgg cggcgagcag aaggatcacg gatttgaggt ggccagcaca agccccgagg 2700
atgagtctcc tggaagcaac cctgagcctg acgccacacc tttccaagag ggcctgagaa 2760
ccttcgacca gctggacgct atcagctccc tgcctacacc tagcgacatc ttcgtgtcct 2820
acagcacatt ccccggcttt gtgtcttggc gggaccctaa gtctggctct tggtacgtgg 2880
aaaccctgga cgacatcttt gagcagtggg ctcacagcga ggacctccag tctctgctgc 2940
tgagagtggc caatgccgtg tccgtgaagg gcatctacaa gcagatgcct ggctgcttca 3000
acttcctgcg gaagaagctg tttttcaaga ccagcggcag cggcgaaggc agaggatccc 3060
ttttgacatg cggcgatgtg gaagagaacc ccggacctat ggctctgcct gtgacagctc 3120
tgcttctgcc tctggcactg cttcttcatg cggcgcgccc tgaagttcag ctgcttgaat 3180
ctggcggagg cctggttcaa cctggcggat ctctgagact gagctgtgcc gccagcggct 3240
tcaccttcag caattacgcc atgacctgga tcagacaggc ccctggcaaa ggcctggaat 3300
gggtgtccgg aattacaggc gacggcggca gcacctttta cgccgattct gtgaagggca 3360
gattcaccat cagccgggac aacagcaaga acaccctgta cctgcagatg aacagcctga 3420
gagccgagga caccgccgtg tactactgcg tgaaggactg gaacaccacc atgatcaccg 3480
agagaggcca gggcacactg gtcaccgtgt cctctacaac aacaccggcg cctagacctc 3540
caacaccagc tcctacaatc gcgagtcagc ccctgtctct cagacccgaa gcctgcaggc 3600
cagctgcagg aggagctgtg cacaccaggg gcctggactt cgcctgcgac atctacattt 3660
gggcaccact ggccgggacc tgtggagtgc tgctgctgag cctggtcatc acactgtact 3720
gcaagagagg caggaagaaa ctgctgtata ttttcaaaca gcccttcatg cgccccgtgc 3780
agactaccca ggaggaagac gggtgctcct gtcgattccc tgaggaagag gaaggcgggt 3840
gtgagctgcg cgtgaagttt agtcgatcag cagatgcccc agcttacaaa cagggacaga 3900
accagctgta taacgagctg aatctgggcc gccgagagga atatgacgtg ctggataagc 3960
ggagaggacg cgaccccgaa atgggaggca agcccaggcg caaaaaccct caggaaggcc 4020
tgtataacga gctgcagaag gacaaaatgg cagaagccta ttctgagatc ggcatgaagg 4080
gggagcgacg gagaggcaaa gggcacgatg ggctgtacca gggactgagc accgccacaa 4140
aggacaccta tgatgctctg catatgcagg cactgcctcc aaggggaagt ggagaaggac 4200
gaggatcact gctgacatgc ggcgacgtgg aggaaaaccc tggcccaatg gtcgggtctc 4260
tgaattgtat cgtcgccgtg agtcagaaca tgggcattgg gaagaatggc gatttcccat 4320
ggccacctct gcgcaacgag tcccgatact ttcagcggat gacaactacc tcctctgtgg 4380
aagggaaaca gaatctggtc atcatgggaa agaaaacttg gttcagcatt ccagagaaga 4440
accggcccct gaaaggcaga atcaatctgg tgctgtcccg agaactgaag gagccaccac 4500
agggagctca ctttctgagc cggtccctgg acgatgcact gaagctgaca gaacagcctg 4560
agctggccaa caaagtcgat atggtgtgga tcgtcggggg aagttcagtg tataaggagg 4620
ccatgaatca ccccggccat ctgaaactgt tcgtcacacg gatcatgcag gactttgaga 4680
gcgatacttt ctttcctgaa attgacctgg agaagtacaa actgctgccc gaatatcctg 4740
gcgtgctgtc cgatgtccag gaagagaaag gcatcaaata caagttcgag gtctatgaga 4800
agaatgacta ataaggtacc gatcacatat gcctttaatt aaacactagt tctatagtgt 4860
cacctaaatt ccctttagtg agggttaatg gccgtaggcc gccagaattg ggtccagaca 4920
tgataagata cattgatgag tttggacaaa ccacaactag aatgcagtga aaaaaatgct 4980
ttatttgtga aatttgtgat gctattgctt tatttgtaac cattataagc tgcaataaac 5040
aagttaacaa caacaattgc attcatttta tgtttcaggt tcagggggag gtgtgggagg 5100
ttttttcgga ctctaggacc tgcgcatgcg cttggcgtaa tcatggtcat agctgtttcc 5160
tgttttcccc gtatcccccc aggtgtctgc aggctcaaag agcagcgaga agcgttcaga 5220

ggaaagcgat cccgtgccac cttccccgtg cccgggctgt ccccgcacgc tgccggctcg 5280
gggatgcggg gggagcgccg gaccggagcg gagccccggg cggctcgctg ctgccccCta 5340
gcgggggagg gacgtaatta catccctggg ggctttgggg gggggctgtc cctctcaccg 5400
cggtggagct ccagcttttg ttcgaattgg ggccccccct cgagggtatc gatgatatct 5460
ataacaagaa aatatatata taataagtta tcacgtaagt agaacatgaa ataacaatat 5520
aattatcgta tgagttaaat cttaaaagtc acgtaaaaga taatcatgcg tcattttgac 5580
tcacgcggtc gttatagttc aaaatcagtg acacttaccg cattgacaag cacgcctcac 5640
gggagctcca agcggcgact gagatgtcct aaatgcacag cgacggattc gcgctattta 5700
gaaagagaga gcaatatttc aagaatgcat gcgtcaattt tacgcagact atctttctag 5760
ggttaatcta gctagcctta agggcgcagc ccgcctaatg agcgggcttt tttttggctt 5820
gttgtccaca accgttaaac cttaaaagct ttaaaagcct tatatattct tttttttctt 5880
ataaaactta aaaccttaga ggctatttaa gttgctgatt tatattaatt ttattgttca 5940
aacatgagag cttagtacgt gaaacatgag agcttagtac gttagccatg agagcttagt 6000
acgttagcca tgagggttta gttcgttaaa catgagagct tagtacgtta aacatgagag 6060
cttagtacgt actatcaaca ggttgaactg ctgatccacg ttgtggtaga attggtaaag 6120
agagtcgtgt aaaatatcga gttcgcacat cttgttgtct gattattgat ttttggcgaa 6180
accatttgat catatgacaa gatgtgtatc taccttaact taatgatttt gataaaaatc 6240
attaggta 6248

<210> 22
<211> 1086
<212> DNA
<213> 人工序列

[0020]

<220>
<223> BCMA VCAR
<400> 22

atggctctgc ctgtgacagc tctgcttctg cctctggcac tgcttcttca tgcggcgcgc 60
cctgaagttc agctgcttga atctggcgga ggcctggttc aacctggcgg atctctgaga 120
ctgagctgtg ccgccagcgg cttcaccttc agcaattacg ccatgacctg gatcagacag 180
gcccctggca aaggcctgga atgggtgtcc ggaattacag gcgacggcgg cagcaccttt 240
tacgccgatt ctgtgaaggg cagattcacc atcagccggg acaacagcaa gaacaccctg 300
tacctgcaga tgaacagcct gagagccgag gacaccgccg tgtactactg cgtgaaggac 360
tggaacacca ccatgatcac cgagagaggc cagggcacac tggtcaccgt gtcctctaca 420
acaacaccgg cgcctagacc tccaacacca gctcctacaa tcgcgagtca gcccctgtct 480
ctcagacccg aagcctgcag gccagctgca ggaggagctg tgcacaccag gggcctggac 540
ttcgcctgcg acatctacat ttgggcacca ctggccggga cctgtggagt gctgctgctg 600
agcctggtca tcacactgta ctgcaagaga ggcaggaaga aactgctgta tattttcaaa 660
cagcccttca tgcgccccgt gcagactacc caggaggaag acgggtgctc ctgtcgattc 720
cctgaggaag aggaaggcgg gtgtgagctg cgcgtgaagt ttagtcgatc agcagatgcc 780
ccagcttaca aacagggaca gaaccagctg tataacgagc tgaatctggg ccgccgagag 840
gaatatgacg tgctggataa gcggagagga cgcgaccccg aaatgggagg caagcccagg 900
cgcaaaaacc ctcaggaagg cctgtataac gagctgcaga aggacaaaat ggcagaagcc 960
tattctgaga tcggcatgaa gggggagcga cggagaggca aagggcacga tgggctgtac 1020
cagggactga gcaccgccac aaaggacacc tatgatgctc tgcatatgca ggcactgcct 1080
ccaagg 1086

<210> 23
<211> 570
<212> PRT
<213> 人工序列
<220>
<223> P-BCMA-ALL01 氨基酸序列
<400> 23

[0021]

Met Ala Leu Pro Val Thr Ala Leu Leu Leu Pro Leu Ala Leu Leu Leu
1 5 10 15
His Ala Ala Arg Pro Glu Val Gln Leu Leu Glu Ser Gly Gly Gly Leu
20 25 30
Val Gln Pro Gly Gly Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe
35 40 45
Thr Phe Ser Asn Tyr Ala Met Thr Trp Ile Arg Gln Ala Pro Gly Lys
50 55 60
Gly Leu Glu Trp Val Ser Gly Ile Thr Gly Asp Gly Gly Ser Thr Phe
65 70 75 80
Tyr Ala Asp Ser Val Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn Ser
85 90 95
Lys Asn Thr Leu Tyr Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr
100 105 110
Ala Val Tyr Tyr Cys Val Lys Asp Trp Asn Thr Thr Met Ile Thr Glu
115 120 125
Arg Gly Gln Gly Thr Leu Val Thr Val Ser Ser Thr Thr Thr Pro Ala
130 135 140
Pro Arg Pro Pro Thr Pro Ala Pro Thr Ile Ala Ser Gln Pro Leu Ser
145 150 155 160
Leu Arg Pro Glu Ala Cys Arg Pro Ala Ala Gly Gly Ala Val His Thr
165 170 175
Arg Gly Leu Asp Phe Ala Cys Asp Ile Tyr Ile Trp Ala Pro Leu Ala
180 185 190
Gly Thr Cys Gly Val Leu Leu Leu Ser Leu Val Ile Thr Leu Tyr Cys
195 200 205
Lys Arg Gly Arg Lys Lys Leu Leu Tyr Ile Phe Lys Gln Pro Phe Met
210 215 220
Arg Pro Val Gln Thr Thr Gln Glu Glu Asp Gly Cys Ser Cys Arg Phe
225 230 235 240
Pro Glu Glu Glu Glu Gly Gly Cys Glu Leu Arg Val Lys Phe Ser Arg
245 250 255
Ser Ala Asp Ala Pro Ala Tyr Lys Gln Gly Gln Asn Gln Leu Tyr Asn
260 265 270
Glu Leu Asn Leu Gly Arg Arg Glu Glu Tyr Asp Val Leu Asp Lys Arg
275 280 285
Arg Gly Arg Asp Pro Glu Met Gly Gly Lys Pro Arg Arg Lys Asn Pro

290 295 300

Gln Glu Gly Leu Tyr Asn Glu Leu Gln Lys Asp Lys Met Ala Glu Ala
305 310 315 320
Tyr Ser Glu Ile Gly Met Lys Gly Glu Arg Arg Arg Gly Lys Gly His
325 330 335
Asp Gly Leu Tyr Gln Gly Leu Ser Thr Ala Thr Lys Asp Thr Tyr Asp
340 345 350
Ala Leu His Met Gln Ala Leu Pro Pro Arg Gly Ser Gly Glu Gly Arg
355 360 365
Gly Ser Leu Leu Thr Cys Gly Asp Val Glu Glu Asn Pro Gly Pro Met
370 375 380
Val Gly Ser Leu Asn Cys Ile Val Ala Val Ser Gln Asn Met Gly Ile
385 390 395 400
Gly Lys Asn Gly Asp Phe Pro Trp Pro Pro Leu Arg Asn Glu Ser Arg
405 410 415
Tyr Phe Gln Arg Met Thr Thr Thr Ser Ser Val Glu Gly Lys Gln Asn
420 425 430
Leu Val Ile Met Gly Lys Lys Thr Trp Phe Ser Ile Pro Glu Lys Asn
435 440 445
Arg Pro Leu Lys Gly Arg Ile Asn Leu Val Leu Ser Arg Glu Leu Lys
450 455 460
[0022] Glu Pro Pro Gln Gly Ala His Phe Leu Ser Arg Ser Leu Asp Asp Ala
465 470 475 480
Leu Lys Leu Thr Glu Gln Pro Glu Leu Ala Asn Lys Val Asp Met Val
485 490 495
Trp Ile Val Gly Gly Ser Ser Val Tyr Lys Glu Ala Met Asn His Pro
500 505 510
Gly His Leu Lys Leu Phe Val Thr Arg Ile Met Gln Asp Phe Glu Ser
515 520 525
Asp Thr Phe Phe Pro Glu Ile Asp Leu Glu Lys Tyr Lys Leu Leu Pro
530 535 540
Glu Tyr Pro Gly Val Leu Ser Asp Val Gln Glu Glu Lys Gly Ile Lys
545 550 555 560
Tyr Lys Phe Glu Val Tyr Glu Lys Asn Asp
565 570

<210> 24
<211> 35
<212> DNA
<213> 人工序列
<220>
<223> ITR
<400> 24
ccctagaaag atagtctgcg taaaattgac gcatg 35
<210> 25

[0023]

<210> 63
<211> DNA
<212> 人工序列
<213>
<220>
<223> ITR
<400> 25
ccctagaaag ataatcatat tgtgacgtac gttaaagata atcatgcgta aaattgacgc 60
atg 63
<210> 26
<211> 63
<212> DNA
<213> 人工序列
<220>
<223> ITR
<400> 26
ccctagaaag ataatcatat tgtgacgtac gttaaagata atcatgcgta aaattgacgc 60
atg 63
<210> 27
<211> 241
<212> DNA
<213> 人工序列
<220>
<223> ITR
<400> 27
ttaaccctag aaagataatc atattgtgac gtacgttaaa gataatcatg tgtaaaattg 60
acgcatgtgt tttatcggtc tgtatatcga ggtttattta ttaatttgaa tagatattaa 120
gttttattat atttacactt acatactaat aataaattca acaaacaatt tatttatgtt 180
tatttattta ttaaaaaaaa caaaaactca aaatttcttc tataaagtaa caaaactttt 240
a 241
<210> 28
<211> 270
<212> PRT
<213> 人工序列
<220>
<223> BCMA centyrin
<400> 28
Ala Thr Gly Cys Thr Gly Cys Cys Thr Gly Cys Ala Cys Cys Ala Ala
1 5 10 15
Ala Gly Ala Ala Cys Cys Thr Gly Gly Thr Gly Gly Thr Gly Ala Gly
20 25 30
Cys Cys Gly Gly Ala Thr Cys Ala Cys Ala Gly Ala Gly Gly Ala Cys
35 40 45
Thr Cys Cys Gly Cys Cys Ala Gly Ala Cys Thr Gly Thr Cys Thr Thr
50 55 60

[0024]

Gly Gly Ala Cys Cys Gly Cys Cys Cys Cys Thr Gly Ala Cys Gly Cys
65 70 75 80
Cys Gly Cys Cys Thr Thr Cys Gly Ala Thr Thr Cys Cys Thr Thr Thr
85 90 95
Cys Cys Ala Ala Thr Cys Cys Gly Gly Thr Ala Cys Ala Thr Cys Gly
100 105 110
Ala Gly Ala Cys Ala Cys Thr Gly Ala Thr Cys Thr Gly Gly Gly Gly
115 120 125
Cys Gly Ala Gly Gly Cys Cys Ala Thr Cys Thr Gly Gly Cys Thr Gly
130 135 140
Gly Ala Cys Gly Thr Gly Cys Cys Cys Gly Gly Cys Thr Cys Thr Gly
145 150 155 160
Ala Gly Ala Gly Gly Ala Gly Cys Thr Ala Cys Gly Ala Thr Cys Thr
165 170 175
Gly Ala Cys Ala Gly Gly Cys Cys Thr Gly Ala Ala Gly Cys Cys Thr
180 185 190
Gly Gly Cys Ala Cys Cys Gly Ala Gly Thr Ala Thr Gly Cys Ala Gly
195 200 205
Thr Gly Gly Thr Cys Ala Thr Cys Ala Cys Ala Gly Gly Ala Gly Thr
210 215 220
Gly Ala Ala Gly Gly Gly Cys Gly Gly Cys Ala Gly Gly Thr Thr Cys
225 230 235 240
Ala Gly Cys Thr Cys Cys Cys Cys Thr Cys Thr Gly Gly Thr Gly Gly
245 250 255
Cys Cys Thr Cys Thr Thr Thr Thr Ala Cys Cys Ala Cys Ala
260 265 270

<210> 29
<211> 90
<212> PRT
<213> 人工序列
<220>
<223> BCMA centyrin
<400> 29

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

85 90

<210> 30
<211> 334
<212> PRT
<213> 人工序列
<220>
<223> BCMA CARTyrin
<400> 30

Met Ala Leu Pro Val Thr Ala Leu Leu Leu Pro Leu Ala Leu Leu Leu
1 5 10 15

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

[0025]

275 280 285

Met Ala Glu Ala Tyr Ser Glu Ile Gly Met Lys Gly Glu Arg Arg Arg

290 295 300

Gly Lys Gly His Asp Gly Leu Tyr Gln Gly Leu Ser Thr Ala Thr Lys

305 310 315 320

Asp Thr Tyr Asp Ala Leu His Met Gln Ala Leu Pro Pro Arg

325 330

<210> 31

<211> 21

<212> PRT

<213> 人工序列

<220>

<223> CD8 信号肽

<400> 31

Met Ala Leu Pro Val Thr Ala Leu Leu Leu Pro Leu Ala Leu Leu Leu

1 5 10 15

His Ala Ala Arg Pro

20

<210> 32

<211> 63

<212> DNA

<213> 人工序列

[0026]

<220>

<223> CD8 信号肽

<400> 32

atggctctgc ctgtgacagc tctgcttctg cctctggcac tgcttcttca tgcggcgcgc 60

cct 63

<210> 33

<211> 15

<212> PRT

<213> 人工序列

<220>

<223> 接头

<400> 33

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser

1 5 10 15

<210> 34

<211> 45

<212> DNA

<213> 人工序列

<220>

<223> 接头

<400> 34

ggcggaggcg gtagcggtgg cggaggtagc ggaggtggtg gatct 45

<210> 35
<211> 45
<212> PRT
<213> 人工序列
<220>
<223> CD8a 铰链
<400> 35
Thr Thr Thr Pro Ala Pro Arg Pro Pro Thr Pro Ala Pro Thr Ile Ala
1 5 10 15
Ser Gln Pro Leu Ser Leu Arg Pro Glu Ala Cys Arg Pro Ala Ala Gly
20 25 30
Gly Ala Val His Thr Arg Gly Leu Asp Phe Ala Cys Asp
35 40 45
<210> 36
<211> 135
<212> DNA
<213> 人工序列
<220>
<223> CD8a 铰链
<400> 36
accacaacac cggcgcctag acctccaaca ccagctccta caatcgcgag tcagcccctg 60
tctctcagac ccgaagcctg caggccagct gcaggaggag ctgtgcacac caggggcctg 120
gacttcgcct gcgac 135

[0027]

<210> 37
<211> 24
<212> PRT
<213> 人工序列
<220>
<223> CD8a TM
<400> 37
Ile Tyr Ile Trp Ala Pro Leu Ala Gly Thr Cys Gly Val Leu Leu Leu
1 5 10 15
Ser Leu Val Ile Thr Leu Tyr Cys
20
<210> 38
<211> 72
<212> DNA
<213> 人工序列
<220>
<223> CD8a TM
<400> 38
atctacattt gggcaccact ggccgggacc tgtggagtgc tgctgctgag cctggtcatc 60
acactgtact gc 72
<210> 39

<211> 42
<212> PRT
<213> 人工序列
<220>
<223> 41BB ICS
<400> 39
Lys Arg Gly Arg Lys Lys Leu Leu Tyr Ile Phe Lys Gln Pro Phe Met
1 5 10 15
Arg Pro Val Gln Thr Thr Gln Glu Glu Asp Gly Cys Ser Cys Arg Phe
20 25 30
Pro Glu Glu Glu Glu Gly Gly Cys Glu Leu
35 40
<210> 40
<211> 126
<212> DNA
<213> 人工序列
<220>
<223> 41BB ICS
<400> 40
aagagaggca ggaagaaact gctgtatatt ttcaaacagc ccttcatgcg ccccgtgcag 60
actacccagg aggaagacgg gtgctcctgt cgattccctg aggaagagga aggcgggtgt 120
gagctg 126

[0028]

<210> 41
<211> 112
<212> PRT
<213> 人工序列
<220>
<223> CD3z ICS
<400> 41
Arg Val Lys Phe Ser Arg Ser Ala Asp Ala Pro Ala Tyr Lys Gln Gly
1 5 10 15
Gln Asn Gln Leu Tyr Asn Glu Leu Asn Leu Gly Arg Arg Glu Glu Tyr
20 25 30
Asp Val Leu Asp Lys Arg Arg Gly Arg Asp Pro Glu Met Gly Gly Lys
35 40 45
Pro Arg Arg Lys Asn Pro Gln Glu Gly Leu Tyr Asn Glu Leu Gln Lys
50 55 60
Asp Lys Met Ala Glu Ala Tyr Ser Glu Ile Gly Met Lys Gly Glu Arg
65 70 75 80
Arg Arg Gly Lys Gly His Asp Gly Leu Tyr Gln Gly Leu Ser Thr Ala
85 90 95
Thr Lys Asp Thr Tyr Asp Ala Leu His Met Gln Ala Leu Pro Pro Arg
100 105 110
<210> 42

<211> 336
<212> DNA
<213> 人工序列
<220>
<223> 41BB ICS
<400> 42
cgcgtgaagt ttagtcgatc agcagatgcc ccagcttaca aacagggaca gaaccagctg 60
tataacgagc tgaatctggg ccgccgagag gaatatgacg tgctggataa gcggagagga 120
cgcgaccccg aaatgggagg caagcccagg cgcaaaaacc ctcaggaagg cctgtataac 180
gagctgcaga aggacaaaat ggcagaagcc tattctgaga tcggcatgaa gggggagcga 240
cggagaggca aagggcacga tgggctgtac cagggactga gcaccgccac aaaggacacc 300
tatgatgctc tgcatatgca ggcactgcct ccaagg 336
<210> 43
<211> 1053
<212> PRT
<213> 人工序列
<220>
<223> SCas9 - 40-18163
<400> 43

[0029]

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

[0030]

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

[0031]

```
545                 550                 555                 560
Arg Ser Val Ser Phe Asp Asn Ser Phe Asn Asn Lys Val Leu Val Lys
                565                 570                 575
Gln Glu Glu Asn Ser Lys Lys Gly Asn Arg Thr Pro Phe Gln Tyr Leu
            580                 585                 590
Ser Ser Ser Asp Ser Lys Ile Ser Tyr Glu Thr Phe Lys Lys His Ile
        595                 600                 605
Leu Asn Leu Ala Lys Gly Lys Gly Arg Ile Ser Lys Thr Lys Lys Glu
    610                 615                 620
Tyr Leu Leu Glu Glu Arg Asp Ile Asn Arg Phe Ser Val Gln Lys Asp
625                 630                 635                 640
Phe Ile Asn Arg Asn Leu Val Asp Thr Arg Tyr Ala Thr Arg Gly Leu
                645                 650                 655
Met Asn Leu Leu Arg Ser Tyr Phe Arg Val Asn Asn Leu Asp Val Lys
            660                 665                 670
Val Lys Ser Ile Asn Gly Gly Phe Thr Ser Phe Leu Arg Arg Lys Trp
        675                 680                 685
Lys Phe Lys Lys Glu Arg Asn Lys Gly Tyr Lys His His Ala Glu Asp
    690                 695                 700
Ala Leu Ile Ile Ala Asn Ala Asp Phe Ile Phe Lys Glu Trp Lys Lys
705                 710                 715                 720
Leu Asp Lys Ala Lys Lys Val Met Glu Asn Gln Met Phe Glu Glu Lys
                725                 730                 735
Gln Ala Glu Ser Met Pro Glu Ile Glu Thr Glu Gln Glu Tyr Lys Glu
            740                 745                 750
Ile Phe Ile Thr Pro His Gln Ile Lys His Ile Lys Asp Phe Lys Asp
        755                 760                 765
Tyr Lys Tyr Ser His Arg Val Asp Lys Lys Pro Asn Arg Glu Leu Ile
    770                 775                 780
Asn Asp Thr Leu Tyr Ser Thr Arg Lys Asp Asp Lys Gly Asn Thr Leu
785                 790                 795                 800
Ile Val Asn Asn Leu Asn Gly Leu Tyr Asp Lys Asp Asn Asp Lys Leu
                805                 810                 815
Lys Lys Leu Ile Asn Lys Ser Pro Glu Lys Leu Leu Met Tyr His His
            820                 825                 830
Asp Pro Gln Thr Tyr Gln Lys Leu Lys Leu Ile Met Glu Gln Tyr Gly
        835                 840                 845
Asp Glu Lys Asn Pro Leu Tyr Lys Tyr Tyr Glu Glu Thr Gly Asn Tyr
    850                 855                 860
Leu Thr Lys Tyr Ser Lys Lys Asp Asn Gly Pro Val Ile Lys Lys Ile
865                 870                 875                 880
Lys Tyr Tyr Gly Asn Lys Leu Asn Ala His Leu Asp Ile Thr Asp Asp
                885                 890                 895
Tyr Pro Asn Ser Arg Asn Lys Val Val Lys Leu Ser Leu Lys Pro Tyr
```

900 905 910
Arg Phe Asp Val Tyr Leu Asp Asn Gly Val Tyr Lys Phe Val Thr Val
915 920 925
Lys Asn Leu Asp Val Ile Lys Lys Glu Asn Tyr Tyr Glu Val Asn Ser
930 935 940
Lys Cys Tyr Glu Glu Ala Lys Lys Leu Lys Lys Ile Ser Asn Gln Ala
945 950 955 960
Glu Phe Ile Ala Ser Phe Tyr Asn Asn Asp Leu Ile Lys Ile Asn Gly
965 970 975
Glu Leu Tyr Arg Val Ile Gly Val Asn Asn Asp Leu Leu Asn Arg Ile
980 985 990
Glu Val Asn Met Ile Asp Ile Thr Tyr Arg Glu Tyr Leu Glu Asn Met
995 1000 1005
Asn Asp Lys Arg Pro Pro Arg Ile Ile Lys Thr Ile Ala Ser Lys
1010 1015 1020
Thr Gln Ser Ile Lys Lys Tyr Ser Thr Asp Ile Leu Gly Asn Leu
1025 1030 1035
Tyr Glu Val Lys Ser Lys Lys His Pro Gln Ile Ile Lys Lys Gly
1040 1045 1050

[0032]

<210> 44
<211> 1053
<212> PRT
<213> 人工序列
<220>
<223> dSaCas9 - 40-18164
<400> 44
Met Lys Arg Asn Tyr Ile Leu Gly Leu Ala Ile Gly Ile Thr Ser Val
1 5 10 15
Gly Tyr Gly Ile Ile Asp Tyr Glu Thr Arg Asp Val Ile Asp Ala Gly
20 25 30
Val Arg Leu Phe Lys Glu Ala Asn Val Glu Asn Asn Glu Gly Arg Arg
35 40 45
Ser Lys Arg Gly Ala Arg Arg Leu Lys Arg Arg Arg Arg His Arg Ile
50 55 60
Gln Arg Val Lys Lys Leu Leu Phe Asp Tyr Asn Leu Leu Thr Asp His
65 70 75 80
Ser Glu Leu Ser Gly Ile Asn Pro Tyr Glu Ala Arg Val Lys Gly Leu
85 90 95
Ser Gln Lys Leu Ser Glu Glu Glu Phe Ser Ala Ala Leu Leu His Leu
100 105 110

Ala Lys Arg Arg Gly Val His Asn Val Asn Glu Val Glu Glu Asp Thr
115 120 125
Gly Asn Glu Leu Ser Thr Lys Glu Gln Ile Ser Arg Asn Ser Lys Ala

[0033]

130 135 140

Leu Glu Glu Lys Tyr Val Ala Glu Leu Gln Leu Glu Arg Leu Lys Lys
145 150 155 160
Asp Gly Glu Val Arg Gly Ser Ile Asn Arg Phe Lys Thr Ser Asp Tyr
165 170 175
Val Lys Glu Ala Lys Gln Leu Leu Lys Val Gln Lys Ala Tyr His Gln
180 185 190
Leu Asp Gln Ser Phe Ile Asp Thr Tyr Ile Asp Leu Leu Glu Thr Arg
195 200 205
Arg Thr Tyr Tyr Glu Gly Pro Gly Glu Gly Ser Pro Phe Gly Trp Lys
210 215 220
Asp Ile Lys Glu Trp Tyr Glu Met Leu Met Gly His Cys Thr Tyr Phe
225 230 235 240
Pro Glu Glu Leu Arg Ser Val Lys Tyr Ala Tyr Asn Ala Asp Leu Tyr
245 250 255
Asn Ala Leu Asn Asp Leu Asn Asn Leu Val Ile Thr Arg Asp Glu Asn
260 265 270
Glu Lys Leu Glu Tyr Tyr Glu Lys Phe Gln Ile Ile Glu Asn Val Phe
275 280 285
Lys Gln Lys Lys Lys Pro Thr Leu Lys Gln Ile Ala Lys Glu Ile Leu
290 295 300
Val Asn Glu Glu Asp Ile Lys Gly Tyr Arg Val Thr Ser Thr Gly Lys
305 310 315 320
Pro Glu Phe Thr Asn Leu Lys Val Tyr His Asp Ile Lys Asp Ile Thr
325 330 335
Ala Arg Lys Glu Ile Ile Glu Asn Ala Glu Leu Leu Asp Gln Ile Ala
340 345 350
Lys Ile Leu Thr Ile Tyr Gln Ser Ser Glu Asp Ile Gln Glu Glu Leu
355 360 365
Thr Asn Leu Asn Ser Glu Leu Thr Gln Glu Glu Ile Glu Gln Ile Ser
370 375 380
Asn Leu Lys Gly Tyr Thr Gly Thr His Asn Leu Ser Leu Lys Ala Ile
385 390 395 400
Asn Leu Ile Leu Asp Glu Leu Trp His Thr Asn Asp Asn Gln Ile Ala
405 410 415
Ile Phe Asn Arg Leu Lys Leu Val Pro Lys Lys Val Asp Leu Ser Gln
420 425 430
Gln Lys Glu Ile Pro Thr Thr Leu Val Asp Asp Phe Ile Leu Ser Pro
435 440 445
Val Val Lys Arg Ser Phe Ile Gln Ser Ile Lys Val Ile Asn Ala Ile
450 455 460
Ile Lys Lys Tyr Gly Leu Pro Asn Asp Ile Ile Ile Glu Leu Ala Arg
465 470 475 480
Glu Lys Asn Ser Lys Asp Ala Gln Lys Met Ile Asn Glu Met Gln Lys

485 490 495
Arg Asn Arg Gln Thr Asn Glu Arg Ile Glu Glu Ile Ile Arg Thr Thr
500 505 510
Gly Lys Glu Asn Ala Lys Tyr Leu Ile Glu Lys Ile Lys Leu His Asp
515 520 525
Met Gln Glu Gly Lys Cys Leu Tyr Ser Leu Glu Ala Ile Pro Leu Glu
530 535 540
Asp Leu Leu Asn Asn Pro Phe Asn Tyr Glu Val Asp His Ile Ile Pro
545 550 555 560
Arg Ser Val Ser Phe Asp Asn Ser Phe Asn Asn Lys Val Leu Val Lys
565 570 575
Gln Glu Glu Ala Ser Lys Lys Gly Asn Arg Thr Pro Phe Gln Tyr Leu
580 585 590
Ser Ser Ser Asp Ser Lys Ile Ser Tyr Glu Thr Phe Lys Lys His Ile
595 600 605
Leu Asn Leu Ala Lys Gly Lys Gly Arg Ile Ser Lys Thr Lys Lys Glu
610 615 620

Tyr Leu Leu Glu Glu Arg Asp Ile Asn Arg Phe Ser Val Gln Lys Asp
625 630 635 640
Phe Ile Asn Arg Asn Leu Val Asp Thr Arg Tyr Ala Thr Arg Gly Leu
645 650 655
[0034]
Met Asn Leu Leu Arg Ser Tyr Phe Arg Val Asn Asn Leu Asp Val Lys
660 665 670
Val Lys Ser Ile Asn Gly Gly Phe Thr Ser Phe Leu Arg Arg Lys Trp
675 680 685
Lys Phe Lys Lys Glu Arg Asn Lys Gly Tyr Lys His His Ala Glu Asp
690 695 700
Ala Leu Ile Ile Ala Asn Ala Asp Phe Ile Phe Lys Glu Trp Lys Lys
705 710 715 720
Leu Asp Lys Ala Lys Lys Val Met Glu Asn Gln Met Phe Glu Glu Lys
725 730 735
Gln Ala Glu Ser Met Pro Glu Ile Glu Thr Glu Gln Glu Tyr Lys Glu
740 745 750
Ile Phe Ile Thr Pro His Gln Ile Lys His Ile Lys Asp Phe Lys Asp
755 760 765
Tyr Lys Tyr Ser His Arg Val Asp Lys Lys Pro Asn Arg Glu Leu Ile
770 775 780
Asn Asp Thr Leu Tyr Ser Thr Arg Lys Asp Asp Lys Gly Asn Thr Leu
785 790 795 800
Ile Val Asn Asn Leu Asn Gly Leu Tyr Asp Lys Asp Asn Asp Lys Leu
805 810 815
Lys Lys Leu Ile Asn Lys Ser Pro Glu Lys Leu Leu Met Tyr His His
820 825 830

Asp Pro Gln Thr Tyr Gln Lys Leu Lys Leu Ile Met Glu Gln Tyr Gly
835 840 845
Asp Glu Lys Asn Pro Leu Tyr Lys Tyr Tyr Glu Glu Thr Gly Asn Tyr
850 855 860
Leu Thr Lys Tyr Ser Lys Lys Asp Asn Gly Pro Val Ile Lys Lys Ile
865 870 875 880
Lys Tyr Tyr Gly Asn Lys Leu Asn Ala His Leu Asp Ile Thr Asp Asp
885 890 895
Tyr Pro Asn Ser Arg Asn Lys Val Val Lys Leu Ser Leu Lys Pro Tyr
900 905 910
Arg Phe Asp Val Tyr Leu Asp Asn Gly Val Tyr Lys Phe Val Thr Val
915 920 925
Lys Asn Leu Asp Val Ile Lys Lys Glu Asn Tyr Tyr Glu Val Asn Ser
930 935 940
Lys Cys Tyr Glu Glu Ala Lys Lys Leu Lys Lys Ile Ser Asn Gln Ala
945 950 955 960
Glu Phe Ile Ala Ser Phe Tyr Asn Asn Asp Leu Ile Lys Ile Asn Gly
965 970 975
Glu Leu Tyr Arg Val Ile Gly Val Asn Asn Asp Leu Leu Asn Arg Ile
980 985 990
Glu Val Asn Met Ile Asp Ile Thr Tyr Arg Glu Tyr Leu Glu Asn Met
995 1000 1005

[0035]

Asn Asp Lys Arg Pro Pro Arg Ile Ile Lys Thr Ile Ala Ser Lys
1010 1015 1020
Thr Gln Ser Ile Lys Lys Tyr Ser Thr Asp Ile Leu Gly Asn Leu
1025 1030 1035
Tyr Glu Val Lys Ser Lys Lys His Pro Gln Ile Ile Lys Lys Gly
1040 1045 1050

<210> 45
<211> 1368
<212> PRT
<213> 人工序列
<220>
<223> dCas9 - 40-18165
<220>
<221> misc_feature
<222> (1)..(1)
<223> Xaa 可以为任何天然存在的氨基酸
<400> 45
Xaa Asp Lys Lys Tyr Ser Ile Gly Leu Ala Ile Gly Thr Asn Ser Val
1 5 10 15
Gly Trp Ala Val Ile Thr Asp Glu Tyr Lys Val Pro Ser Lys Lys Phe
20 25 30
Lys Val Leu Gly Asn Thr Asp Arg His Ser Ile Lys Lys Asn Leu Ile

[0036]

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

385 390 395 400
Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile Pro His Gln Ile His Leu
405 410 415
Gly Glu Leu His Ala Ile Leu Arg Arg Gln Glu Asp Phe Tyr Pro Phe
420 425 430
Leu Lys Asp Asn Arg Glu Lys Ile Glu Lys Ile Leu Thr Phe Arg Ile
435 440 445
Pro Tyr Tyr Val Gly Pro Leu Ala Arg Gly Asn Ser Arg Phe Ala Trp
450 455 460
Met Thr Arg Lys Ser Glu Glu Thr Ile Thr Pro Trp Asn Phe Glu Glu
465 470 475 480
Val Val Asp Lys Gly Ala Ser Ala Gln Ser Phe Ile Glu Arg Met Thr
485 490 495
Asn Phe Asp Lys Asn Leu Pro Asn Glu Lys Val Leu Pro Lys His Ser
500 505 510
Leu Leu Tyr Glu Tyr Phe Thr Val Tyr Asn Glu Leu Thr Lys Val Lys
515 520 525
Tyr Val Thr Glu Gly Met Arg Lys Pro Ala Phe Leu Ser Gly Glu Gln
530 535 540
Lys Lys Ala Ile Val Asp Leu Leu Phe Lys Thr Asn Arg Lys Val Thr
545 550 555 560
[0037] Val Lys Gln Leu Lys Glu Asp Tyr Phe Lys Lys Ile Glu Cys Phe Asp
565 570 575
Ser Val Glu Ile Ser Gly Val Glu Asp Arg Phe Asn Ala Ser Leu Gly
580 585 590
Thr Tyr His Asp Leu Leu Lys Ile Ile Lys Asp Lys Asp Phe Leu Asp
595 600 605
Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp Ile Val Leu Thr Leu Thr
610 615 620
Leu Phe Glu Asp Arg Glu Met Ile Glu Glu Arg Leu Lys Thr Tyr Ala
625 630 635 640
His Leu Phe Asp Asp Lys Val Met Lys Gln Leu Lys Arg Arg Arg Tyr
645 650 655
Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu Ile Asn Gly Ile Arg Asp
660 665 670
Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe Leu Lys Ser Asp Gly Phe
675 680 685
Ala Asn Arg Asn Phe Met Gln Leu Ile His Asp Asp Ser Leu Thr Phe
690 695 700
Lys Glu Asp Ile Gln Lys Ala Gln Val Ser Gly Gln Gly Asp Ser Leu
705 710 715 720
His Glu His Ile Ala Asn Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly
725 730 735
Ile Leu Gln Thr Val Lys Val Val Asp Glu Leu Val Lys Val Met Gly

[0038]

```
                740                 745                 750
Arg His Lys Pro Glu Asn Ile Val Ile Glu Met Ala Arg Glu Asn Gln
        755                 760                 765
Thr Thr Gln Lys Gly Gln Lys Asn Ser Arg Glu Arg Met Lys Arg Ile
    770                 775                 780
Glu Glu Gly Ile Lys Glu Leu Gly Ser Gln Ile Leu Lys Glu His Pro
785                 790                 795                 800
Val Glu Asn Thr Gln Leu Gln Asn Glu Lys Leu Tyr Leu Tyr Tyr Leu
                805                 810                 815
Gln Asn Gly Arg Asp Met Tyr Val Asp Gln Glu Leu Asp Ile Asn Arg
            820                 825                 830
Leu Ser Asp Tyr Asp Val Asp Ala Ile Val Pro Gln Ser Phe Leu Lys
        835                 840                 845
Asp Asp Ser Ile Asp Asn Lys Val Leu Thr Arg Ser Asp Lys Asn Arg
    850                 855                 860
Gly Lys Ser Asp Asn Val Pro Ser Glu Glu Val Val Lys Lys Met Lys
865                 870                 875                 880
Asn Tyr Trp Arg Gln Leu Leu Asn Ala Lys Leu Ile Thr Gln Arg Lys
                885                 890                 895
Phe Asp Asn Leu Thr Lys Ala Glu Arg Gly Gly Leu Ser Glu Leu Asp
            900                 905                 910
Lys Ala Gly Phe Ile Lys Arg Gln Leu Val Glu Thr Arg Gln Ile Thr
        915                 920                 925
Lys His Val Ala Gln Ile Leu Asp Ser Arg Met Asn Thr Lys Tyr Asp
    930                 935                 940
Glu Asn Asp Lys Leu Ile Arg Glu Val Lys Val Ile Thr Leu Lys Ser
945                 950                 955                 960
Lys Leu Val Ser Asp Phe Arg Lys Asp Phe Gln Phe Tyr Lys Val Arg
                965                 970                 975
Glu Ile Asn Asn Tyr His His Ala His Asp Ala Tyr Leu Asn Ala Val
            980                 985                 990
Val Gly Thr Ala Leu Ile Lys Lys  Tyr Pro Lys Leu Glu  Ser Glu Phe
        995                 1000                1005
Val Tyr  Gly Asp Tyr Lys Val  Tyr Asp Val Arg Lys  Met Ile Ala
    1010                1015                1020
Lys Ser  Glu Gln Glu Ile Gly  Lys Ala Thr Ala Lys  Tyr Phe Phe
    1025                1030                1035
Tyr Ser  Asn Ile Met Asn Phe  Phe Lys Thr Glu Ile  Thr Leu Ala
    1040                1045                1050
Asn Gly  Glu Ile Arg Lys Arg  Pro Leu Ile Glu Thr  Asn Gly Glu
    1055                1060                1065
Thr Gly  Glu Ile Val Trp Asp  Lys Gly Arg Asp Phe  Ala Thr Val
    1070                1075                1080
Arg Lys  Val Leu Ser Met Pro  Gln Val Asn Ile Val  Lys Lys Thr
```

1085 1090 1095
Glu Val Gln Thr Gly Gly Phe Ser Lys Glu Ser Ile Leu Pro Lys
1100 1105 1110
Arg Asn Ser Asp Lys Leu Ile Ala Arg Lys Lys Asp Trp Asp Pro
1115 1120 1125
Lys Lys Tyr Gly Gly Phe Asp Ser Pro Thr Val Ala Tyr Ser Val
1130 1135 1140
Leu Val Val Ala Lys Val Glu Lys Gly Lys Ser Lys Lys Leu Lys
1145 1150 1155
Ser Val Lys Glu Leu Leu Gly Ile Thr Ile Met Glu Arg Ser Ser
1160 1165 1170
Phe Glu Lys Asn Pro Ile Asp Phe Leu Glu Ala Lys Gly Tyr Lys
1175 1180 1185
Glu Val Lys Lys Asp Leu Ile Ile Lys Leu Pro Lys Tyr Ser Leu
1190 1195 1200
Phe Glu Leu Glu Asn Gly Arg Lys Arg Met Leu Ala Ser Ala Gly
1205 1210 1215
Glu Leu Gln Lys Gly Asn Glu Leu Ala Leu Pro Ser Lys Tyr Val
1220 1225 1230
Asn Phe Leu Tyr Leu Ala Ser His Tyr Glu Lys Leu Lys Gly Ser
1235 1240 1245
[0039] Pro Glu Asp Asn Glu Gln Lys Gln Leu Phe Val Glu Gln His Lys
1250 1255 1260
His Tyr Leu Asp Glu Ile Ile Glu Gln Ile Ser Glu Phe Ser Lys
1265 1270 1275
Arg Val Ile Leu Ala Asp Ala Asn Leu Asp Lys Val Leu Ser Ala
1280 1285 1290
Tyr Asn Lys His Arg Asp Lys Pro Ile Arg Glu Gln Ala Glu Asn
1295 1300 1305
Ile Ile His Leu Phe Thr Leu Thr Asn Leu Gly Ala Pro Ala Ala
1310 1315 1320
Phe Lys Tyr Phe Asp Thr Thr Ile Asp Arg Lys Arg Tyr Thr Ser
1325 1330 1335
Thr Lys Glu Val Leu Asp Ala Thr Leu Ile His Gln Ser Ile Thr
1340 1345 1350
Gly Leu Tyr Glu Thr Arg Ile Asp Leu Ser Gln Leu Gly Gly Asp
1355 1360 1365

<210> 46
<211> 1368
<212> PRT
<213> 人工序列
<220>
<223> dCas9 - 40-18166
<400> 46

[0040]

```
Met Asp Lys Lys Tyr Ser Ile Gly Leu Ala Ile Gly Thr Asn Ser Val
1               5                  10                  15
Gly Trp Ala Val Ile Thr Asp Glu Tyr Lys Val Pro Ser Lys Lys Phe
            20                  25                  30
Lys Val Leu Gly Asn Thr Asp Arg His Ser Ile Lys Lys Asn Leu Ile
        35                  40                  45
Gly Ala Leu Leu Phe Asp Ser Gly Glu Thr Ala Glu Ala Thr Arg Leu
    50                  55                  60
Lys Arg Thr Ala Arg Arg Arg Tyr Thr Arg Arg Lys Asn Arg Ile Cys
65                  70                  75                  80
Tyr Leu Gln Glu Ile Phe Ser Asn Glu Met Ala Lys Val Asp Asp Ser
                85                  90                  95
Phe Phe His Arg Leu Glu Glu Ser Phe Leu Val Glu Glu Asp Lys Lys
            100                 105                 110
His Glu Arg His Pro Ile Phe Gly Asn Ile Val Asp Glu Val Ala Tyr
        115                 120                 125
His Glu Lys Tyr Pro Thr Ile Tyr His Leu Arg Lys Lys Leu Val Asp
    130                 135                 140
Ser Thr Asp Lys Ala Asp Leu Arg Leu Ile Tyr Leu Ala Leu Ala His
145                 150                 155                 160
Met Ile Lys Phe Arg Gly His Phe Leu Ile Glu Gly Asp Leu Asn Pro
                165                 170                 175
Asp Asn Ser Asp Val Asp Lys Leu Phe Ile Gln Leu Val Gln Thr Tyr
            180                 185                 190
Asn Gln Leu Phe Glu Glu Asn Pro Ile Asn Ala Ser Gly Val Asp Ala
        195                 200                 205
Lys Ala Ile Leu Ser Ala Arg Leu Ser Lys Ser Arg Arg Leu Glu Asn
    210                 215                 220
Leu Ile Ala Gln Leu Pro Gly Glu Lys Lys Asn Gly Leu Phe Gly Asn
225                 230                 235                 240
Leu Ile Ala Leu Ser Leu Gly Leu Thr Pro Asn Phe Lys Ser Asn Phe
                245                 250                 255
Asp Leu Ala Glu Asp Ala Lys Leu Gln Leu Ser Lys Asp Thr Tyr Asp
            260                 265                 270
Asp Asp Leu Asp Asn Leu Leu Ala Gln Ile Gly Asp Gln Tyr Ala Asp
        275                 280                 285

Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp Ala Ile Leu Leu Ser Asp
    290                 295                 300
Ile Leu Arg Val Asn Thr Glu Ile Thr Lys Ala Pro Leu Ser Ala Ser
305                 310                 315                 320
Met Ile Lys Arg Tyr Asp Glu His His Gln Asp Leu Thr Leu Leu Lys
                325                 330                 335
Ala Leu Val Arg Gln Gln Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe
```

[0041]

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

690 695 700

Lys Glu Asp Ile Gln Lys Ala Gln Val Ser Gly Gln Gly Asp Ser Leu
705 710 715 720
His Glu His Ile Ala Asn Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly
725 730 735
Ile Leu Gln Thr Val Lys Val Val Asp Glu Leu Val Lys Val Met Gly
740 745 750
Arg His Lys Pro Glu Asn Ile Val Ile Glu Met Ala Arg Glu Asn Gln
755 760 765
Thr Thr Gln Lys Gly Gln Lys Asn Ser Arg Glu Arg Met Lys Arg Ile
770 775 780
Glu Glu Gly Ile Lys Glu Leu Gly Ser Gln Ile Leu Lys Glu His Pro
785 790 795 800

Val Glu Asn Thr Gln Leu Gln Asn Glu Lys Leu Tyr Leu Tyr Tyr Leu
805 810 815
Gln Asn Gly Arg Asp Met Tyr Val Asp Gln Glu Leu Asp Ile Asn Arg
820 825 830
Leu Ser Asp Tyr Asp Val Asp Ala Ile Val Pro Gln Ser Phe Leu Lys
835 840 845
Asp Asp Ser Ile Asp Asn Lys Val Leu Thr Arg Ser Asp Lys Asn Arg
850 855 860
[0042]
Gly Lys Ser Asp Asn Val Pro Ser Glu Glu Val Val Lys Lys Met Lys
865 870 875 880
Asn Tyr Trp Arg Gln Leu Leu Asn Ala Lys Leu Ile Thr Gln Arg Lys
885 890 895
Phe Asp Asn Leu Thr Lys Ala Glu Arg Gly Gly Leu Ser Glu Leu Asp
900 905 910
Lys Ala Gly Phe Ile Lys Arg Gln Leu Val Glu Thr Arg Gln Ile Thr
915 920 925
Lys His Val Ala Gln Ile Leu Asp Ser Arg Met Asn Thr Lys Tyr Asp
930 935 940
Glu Asn Asp Lys Leu Ile Arg Glu Val Lys Val Ile Thr Leu Lys Ser
945 950 955 960
Lys Leu Val Ser Asp Phe Arg Lys Asp Phe Gln Phe Tyr Lys Val Arg
965 970 975
Glu Ile Asn Asn Tyr His His Ala His Asp Ala Tyr Leu Asn Ala Val
980 985 990
Val Gly Thr Ala Leu Ile Lys Lys Tyr Pro Lys Leu Glu Ser Glu Phe
995 1000 1005
Val Tyr Gly Asp Tyr Lys Val Tyr Asp Val Arg Lys Met Ile Ala
1010 1015 1020
Lys Ser Glu Gln Glu Ile Gly Lys Ala Thr Ala Lys Tyr Phe Phe
1025 1030 1035

[0043]

```
Tyr Ser  Asn Ile Met Asn Phe  Phe Lys Thr Glu Ile  Thr Leu Ala
    1040                1045                1050
Asn Gly  Glu Ile Arg Lys Arg  Pro Leu Ile Glu Thr  Asn Gly Glu
    1055                1060                1065
Thr Gly  Glu Ile Val Trp Asp  Lys Gly Arg Asp Phe  Ala Thr Val
    1070                1075                1080
Arg Lys  Val Leu Ser Met Pro  Gln Val Asn Ile Val  Lys Lys Thr
    1085                1090                1095
Glu Val  Gln Thr Gly Gly Phe  Ser Lys Glu Ser Ile  Leu Pro Lys
    1100                1105                1110
Arg Asn  Ser Asp Lys Leu Ile  Ala Arg Lys Lys Asp  Trp Asp Pro
    1115                1120                1125
Lys Lys  Tyr Gly Gly Phe Asp  Ser Pro Thr Val Ala  Tyr Ser Val
    1130                1135                1140
Leu Val  Val Ala Lys Val Glu  Lys Gly Lys Ser Lys  Lys Leu Lys
    1145                1150                1155
Ser Val  Lys Glu Leu Leu Gly  Ile Thr Ile Met Glu  Arg Ser Ser
    1160                1165                1170
Phe Glu  Lys Asn Pro Ile Asp  Phe Leu Glu Ala Lys  Gly Tyr Lys
    1175                1180                1185
Glu Val  Lys Lys Asp Leu Ile  Ile Lys Leu Pro Lys  Tyr Ser Leu
    1190                1195                1200
Phe Glu  Leu Glu Asn Gly Arg  Lys Arg Met Leu Ala  Ser Ala Gly
    1205                1210                1215
Glu Leu  Gln Lys Gly Asn Glu  Leu Ala Leu Pro Ser  Lys Tyr Val
    1220                1225                1230
Asn Phe  Leu Tyr Leu Ala Ser  His Tyr Glu Lys Leu  Lys Gly Ser
    1235                1240                1245
Pro Glu  Asp Asn Glu Gln Lys  Gln Leu Phe Val Glu  Gln His Lys
    1250                1255                1260
His Tyr  Leu Asp Glu Ile Ile  Glu Gln Ile Ser Glu  Phe Ser Lys
    1265                1270                1275
Arg Val  Ile Leu Ala Asp Ala  Asn Leu Asp Lys Val  Leu Ser Ala
    1280                1285                1290
Tyr Asn  Lys His Arg Asp Lys  Pro Ile Arg Glu Gln  Ala Glu Asn
    1295                1300                1305
Ile Ile  His Leu Phe Thr Leu  Thr Asn Leu Gly Ala  Pro Ala Ala
    1310                1315                1320
Phe Lys  Tyr Phe Asp Thr Thr  Ile Asp Arg Lys Arg  Tyr Thr Ser
    1325                1330                1335
Thr Lys  Glu Val Leu Asp Ala  Thr Leu Ile His Gln  Ser Ile Thr
    1340                1345                1350
Gly Leu  Tyr Glu Thr Arg Ile  Asp Leu Ser Gln Leu  Gly Gly Asp
    1355                1360                1365
```

<210> 47
<211> 199
<212> PRT
<213> 人工序列
<220>
<223> Clo051 - 40-18030
<400> 47

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

[0044]

<210> 48
<211> 1591
<212> PRT
<213> 人工序列
<220>
<223> dCas9-Clo051 - 40-18168
<400> 48

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

[0045]

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

[0046]

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

[0047]

```
Thr Lys Val Lys Tyr Val Thr Glu Gly Met Arg Lys Pro Ala Phe Leu
            740                 745                 750
Ser Gly Glu Gln Lys Lys Ala Ile Val Asp Leu Leu Phe Lys Thr Asn
        755                 760                 765
Arg Lys Val Thr Val Lys Gln Leu Lys Glu Asp Tyr Phe Lys Lys Ile
    770                 775                 780
Glu Cys Phe Asp Ser Val Glu Ile Ser Gly Val Glu Asp Arg Phe Asn
785                 790                 795                 800
Ala Ser Leu Gly Thr Tyr His Asp Leu Leu Lys Ile Ile Lys Asp Lys
                805                 810                 815
Asp Phe Leu Asp Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp Ile Val
            820                 825                 830
Leu Thr Leu Thr Leu Phe Glu Asp Arg Glu Met Ile Glu Glu Arg Leu
        835                 840                 845
Lys Thr Tyr Ala His Leu Phe Asp Asp Lys Val Met Lys Gln Leu Lys
    850                 855                 860
Arg Arg Arg Tyr Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu Ile Asn
865                 870                 875                 880
Gly Ile Arg Asp Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe Leu Lys
                885                 890                 895
Ser Asp Gly Phe Ala Asn Arg Asn Phe Met Gln Leu Ile His Asp Asp
            900                 905                 910
Ser Leu Thr Phe Lys Glu Asp Ile Gln Lys Ala Gln Val Ser Gly Gln
        915                 920                 925
Gly Asp Ser Leu His Glu His Ile Ala Asn Leu Ala Gly Ser Pro Ala
    930                 935                 940
Ile Lys Lys Gly Ile Leu Gln Thr Val Lys Val Val Asp Glu Leu Val
945                 950                 955                 960
Lys Val Met Gly Arg His Lys Pro Glu Asn Ile Val Ile Glu Met Ala
                965                 970                 975
Arg Glu Asn Gln Thr Thr Gln Lys Gly Gln Lys Asn Ser Arg Glu Arg
            980                 985                 990
Met Lys Arg Ile Glu Glu Gly Ile  Lys Glu Leu Gly Ser  Gln Ile Leu
        995                 1000                1005
Lys Glu  His Pro Val Glu Asn  Thr Gln Leu Gln Asn  Glu Lys Leu
    1010                1015                1020
Tyr Leu  Tyr Tyr Leu Gln Asn  Gly Arg Asp Met Tyr  Val Asp Gln
    1025                1030                1035
Glu Leu  Asp Ile Asn Arg Leu  Ser Asp Tyr Asp Val  Asp Ala Ile
    1040                1045                1050
Val Pro  Gln Ser Phe Leu Lys  Asp Asp Ser Ile Asp  Asn Lys Val
    1055                1060                1065
Leu Thr  Arg Ser Asp Lys Asn  Arg Gly Lys Ser Asp  Asn Val Pro
    1070                1075                1080
```

[0048]

```
Ser Glu  Glu Val Val Lys Lys  Met Lys Asn Tyr Trp  Arg Gln Leu
    1085                1090                1095
Leu Asn  Ala Lys Leu Ile Thr  Gln Arg Lys Phe Asp  Asn Leu Thr
    1100                1105                1110
Lys Ala  Glu Arg Gly Gly Leu  Ser Glu Leu Asp Lys  Ala Gly Phe
    1115                1120                1125
Ile Lys  Arg Gln Leu Val Glu  Thr Arg Gln Ile Thr  Lys His Val
    1130                1135                1140
Ala Gln  Ile Leu Asp Ser Arg  Met Asn Thr Lys Tyr  Asp Glu Asn
    1145                1150                1155
Asp Lys  Leu Ile Arg Glu Val  Lys Val Ile Thr Leu  Lys Ser Lys
    1160                1165                1170
Leu Val  Ser Asp Phe Arg Lys  Asp Phe Gln Phe Tyr  Lys Val Arg
    1175                1180                1185
Glu Ile  Asn Asn Tyr His His  Ala His Asp Ala Tyr  Leu Asn Ala
    1190                1195                1200
Val Val  Gly Thr Ala Leu Ile  Lys Lys Tyr Pro Lys  Leu Glu Ser
    1205                1210                1215
Glu Phe  Val Tyr Gly Asp Tyr  Lys Val Tyr Asp Val  Arg Lys Met
    1220                1225                1230
Ile Ala  Lys Ser Glu Gln Glu  Ile Gly Lys Ala Thr  Ala Lys Tyr
    1235                1240                1245
Phe Phe  Tyr Ser Asn Ile Met  Asn Phe Phe Lys Thr  Glu Ile Thr
    1250                1255                1260
Leu Ala  Asn Gly Glu Ile Arg  Lys Arg Pro Leu Ile  Glu Thr Asn
    1265                1270                1275
Gly Glu  Thr Gly Glu Ile Val  Trp Asp Lys Gly Arg  Asp Phe Ala
    1280                1285                1290
Thr Val  Arg Lys Val Leu Ser  Met Pro Gln Val Asn  Ile Val Lys
    1295                1300                1305
Lys Thr  Glu Val Gln Thr Gly  Gly Phe Ser Lys Glu  Ser Ile Leu
    1310                1315                1320
Pro Lys  Arg Asn Ser Asp Lys  Leu Ile Ala Arg Lys  Lys Asp Trp
    1325                1330                1335
Asp Pro  Lys Lys Tyr Gly Gly  Phe Asp Ser Pro Thr  Val Ala Tyr
    1340                1345                1350
Ser Val  Leu Val Val Ala Lys  Val Glu Lys Gly Lys  Ser Lys Lys
    1355                1360                1365
Leu Lys  Ser Val Lys Glu Leu  Leu Gly Ile Thr Ile  Met Glu Arg
    1370                1375                1380
Ser Ser  Phe Glu Lys Asn Pro  Ile Asp Phe Leu Glu  Ala Lys Gly
    1385                1390                1395
Tyr Lys  Glu Val Lys Lys Asp  Leu Ile Ile Lys Leu  Pro Lys Tyr
    1400                1405                1410
```

Ser Leu Phe Glu Leu Glu Asn Gly Arg Lys Arg Met Leu Ala Ser
1415 1420 1425
Ala Gly Glu Leu Gln Lys Gly Asn Glu Leu Ala Leu Pro Ser Lys
1430 1435 1440
Tyr Val Asn Phe Leu Tyr Leu Ala Ser His Tyr Glu Lys Leu Lys
1445 1450 1455
Gly Ser Pro Glu Asp Asn Glu Gln Lys Gln Leu Phe Val Glu Gln
1460 1465 1470
His Lys His Tyr Leu Asp Glu Ile Ile Glu Gln Ile Ser Glu Phe
1475 1480 1485
Ser Lys Arg Val Ile Leu Ala Asp Ala Asn Leu Asp Lys Val Leu
1490 1495 1500
Ser Ala Tyr Asn Lys His Arg Asp Lys Pro Ile Arg Glu Gln Ala
1505 1510 1515
Glu Asn Ile Ile His Leu Phe Thr Leu Thr Asn Leu Gly Ala Pro
1520 1525 1530
Ala Ala Phe Lys Tyr Phe Asp Thr Thr Ile Asp Arg Lys Arg Tyr
1535 1540 1545
Thr Ser Thr Lys Glu Val Leu Asp Ala Thr Leu Ile His Gln Ser
1550 1555 1560
Ile Thr Gly Leu Tyr Glu Thr Arg Ile Asp Leu Ser Gln Leu Gly
1565 1570 1575

[0049]

Gly Asp Gly Ser Pro Lys Lys Lys Arg Lys Val Ser Ser
1580 1585 1590

<210> 49
<211> 4776
<212> DNA
<213> 人工序列
<220>
<223> dCas9-Clo051 - 40-18169
<400> 49

atggcaccaa agaagaaaag aaaagtggag ggcatcaagt caaacatcag cctgctgaaa 60
gacgaactgc ggggacagat tagtcacatc agtcacgagt acctgtcact gattgatctg 120
gccttcgaca gcaagcagaa tagactgttt gagatgaaag tgctggaact gctggtcaac 180
gagtatggct tcaagggcag acatctgggc gggtctagga aacctgacgg catcgtgtac 240
agtaccacac tggaagacaa cttcggaatc attgtcgata ccaaggctta ttccgagggc 300
tactctctgc caattagtca ggcagatgag atggaaaggt acgtgcgcga aaactcaaat 360
agggacgagg aagtcaaccc caataagtgg tgggagaatt tcagcgagga agtgaagaaa 420
tactacttcg tctttatctc aggcagcttc aaagggaagt ttgaggaaca gctgcggaga 480
ctgtccatga ctaccggggt gaacggatct gctgtcaacg tggtcaatct gctgctgggc 540
gcagaaaaga tcaggtccgg ggagatgaca attgaggaac tggaacgcgc catgttcaac 600
aattctgagt ttatcctgaa gtatggaggc gggggaagcg ataagaaata ctccatcgga 660
ctggccattg gcaccaattc cgtgggctgg gctgtcatca cagacgagta caaggtgcca 720
agcaagaagt tcaaggtcct ggggaacacc gatcgccaca gtatcaagaa aaatctgatt 780

[0050]

ggagccctgc tgttcgactc aggcgagact gctgaagcaa cccgactgaa gcggactgct 840
aggcgccgat atacccggag aaaaaatcgg atctgctacc tgcaggaaat tttcagcaac 900
gagatggcca aggtggacga tagtttcttt caccgcctgg aggaatcatt cctggtggag 960
gaagataaga aacacgagcg gcatcccatc tttggcaaca ttgtggacga agtcgcttat 1020
cacgagaagt accctactat ctatcatctg aggaagaaac tggtggactc caccgataag 1080
gcagacctgc gcctgatcta tctggccctg gctcacatga tcaagttccg ggggcatttt 1140
ctgatcgagg gagatctgaa ccctgacaat tctgatgtgg acaagctgtt catccagctg 1200
gtccagacat acaatcagct gtttgaggaa aacccaatta atgcctcagg cgtggacgca 1260
aaggccatcc tgagcgccag actgtccaaa tctaggcgcc tggaaaacct gatcgctcag 1320
ctgccaggag agaagaaaaa cggcctgttt gggaatctga ttgcactgtc cctgggcctg 1380
acacccaact tcaagtctaa ttttgatctg gccgaggacg ctaagctgca gctgtccaaa 1440
gacacttatg acgatgacct ggataacctg ctggctcaga tcggcgatca gtacgcagac 1500
ctgttcctgg ccgctaagaa tctgagtgac gccatcctgc tgtcagatat tctgcgcgtg 1560
aacacagaga ttactaaggc cccactgagt gcttcaatga tcaaaagata tgacgagcac 1620
catcaggatc tgaccctgct gaaggctctg gtgaggcagc agctgcccga gaaatacaag 1680
gaaatcttct ttgatcagag caagaatgga tacgccggct atattgacgg cggggcttcc 1740
caggaggagt tctacaagtt catcaagccc attctggaaa agatggacgg caccgaggaa 1800
ctgctggtga agctgaatcg ggaggacctg ctgagaaaac agaggacatt tgataacgga 1860
agcatccctc accagattca tctgggcgaa ctgcacgcca tcctgcgacg gcaggaggac 1920
ttctacccat ttctgaagga taaccgcgag aaaatcgaaa agatcctgac cttcagaatc 1980
ccctactatg tggggcctct ggcacgggga aatagtagat ttgcctggat gacaagaaag 2040
tcagaggaaa ctatcacccc ctggaacttc gaggaagtgg tcgataaagg cgctagcgca 2100
cagtccttca ttgaaaggat gacaaatttt gacaagaacc tgccaaatga gaaggtgctg 2160
cccaaacaca gcctgctgta cgaatatttc acagtgtata acgagctgac taaagtgaag 2220
tacgtcaccg aagggatgcg caagcccgca ttcctgtccg gagagcagaa gaaagccatc 2280
gtggacctgc tgtttaagac aaatcggaaa gtgactgtca aacagctgaa ggaagactat 2340
ttcaagaaaa ttgagtgttt cgattcagtg gaaatcagcg gcgtcgagga caggtttaac 2400
gcctccctgg ggacctacca cgatctgctg aagatcatca aggataagga cttcctggac 2460
aacgaggaaa atgaggacat cctggaggac attgtgctga cactgactct gtttgaggat 2520
cgcgaaatga tcgaggaacg actgaagact tatgcccatc tgttcgatga caaagtgatg 2580
aagcagctga aaagaaggcg ctacaccgga tggggacgcc tgagccgaaa actgatcaat 2640
gggattagag acaagcagag cggaaaaact atcctggact ttctgaagtc cgatggcttc 2700
gccaacagga acttcatgca gctgattcac gatgactctc tgaccttcaa ggaggacatc 2760
cagaaagcac aggtgtctgg ccagggggac agtctgcacg agcatatcgc aaacctggcc 2820
ggcagccccg ccatcaagaa agggattctg cagaccgtga aggtggtgga cgaactggtc 2880
aaggtcatgg gacgacacaa acctgagaac atcgtgattg agatggcccg cgaaaatcag 2940
acaactcaga agggccagaa aaacagtcga gaacggatga agagaatcga ggaaggcatc 3000
aaggagctgg ggtcacagat cctgaaggag catcctgtgg aaaacactca gctgcagaat 3060
gagaaactgt atctgtacta tctgcagaat ggacgggata tgtacgtgga ccaggagctg 3120
gatattaaca gactgagtga ttatgacgtg gatgccatcg tccctcagag cttcctgaag 3180
gatgactcca ttgacaacaa ggtgctgacc aggtccgaca agaaccgcgg caaatcagat 3240
aatgtgccaa gcgaggaagt ggtcaagaaa atgaagaact actggaggca gctgctgaat 3300
gccaagctga tcacacagcg gaaatttgat aacctgacta aggcagaaag aggaggcctg 3360
tctgagctgg acaaggccgg cttcatcaag cggcagctgg tggagacaag acagatcact 3420

[0051]

```
aagcacgtcg ctcagattct ggatagcaga atgaacacaa agtacgatga aaacgacaag     3480
ctgatcaggg aggtgaaagt cattactctg aaatccaagc tggtgtctga ctttagaaag     3540
gatttccagt tttataaagt cagggagatc aacaactacc accatgctca tgacgcatac     3600
ctgaacgcag tggtcgggac cgccctgatt aagaaatacc ccaagctgga gtccgagttc     3660
gtgtacggag actataaagt gtacgatgtc cggaagatga tcgccaaatc tgagcaggaa     3720
attggcaagg ccaccgctaa gtatttcttt tacagtaaca tcatgaattt ctttaagacc     3780
gaaatcacac tggcaaatgg ggagatcaga aaaaggcctc tgattgagac caacggggag     3840
acaggagaaa tcgtgtggga caagggaagg gattttgcta ccgtgcgcaa agtcctgtcc     3900
atgccccaag tgaatattgt caagaaaact gaagtgcaga ccgggggatt ctctaaggag     3960
agtattctgc ctaagcgaaa ctctgataaa ctgatcgccc ggaagaaaga ctgggacccc     4020
aagaagtatg gcgggttcga ctctccaaca gtggcttaca gtgtcctggt ggtcgcaaag     4080
gtggaaaagg ggaagtccaa gaaactgaag tctgtcaaag agctgctggg aatcactatt     4140
atggaacgca gctccttcga gaagaatcct atcgattttc tggaagccaa gggctataaa     4200
gaggtgaaga aagacctgat cattaagctg ccaaaatact cactgtttga gctggaaaac     4260
ggacgaaagc gaatgctggc aagcgccgga gaactgcaga agggcaatga gctggccctg     4320
ccctccaaat acgtgaactt cctgtatctg gctagccact acgagaaact gaaggggtcc     4380
cctgaggata acgaacagaa gcagctgttt gtggagcagc acaaacatta tctggacgag     4440
atcattgaac agatttcaga gttcagcaag agagtgatcc tggctgacgc aaatctggat     4500
aaagtcctga gcgcatacaa caagcaccga gacaaaccaa tccgggagca ggccgaaaat     4560
atcattcatc tgttcaccct gacaaacctg ggcgcccctg cagccttcaa gtattttgac     4620
accacaatcg atcggaagag atacacttct accaaagagg tgctggatgc taccctgatc     4680
caccagagta ttaccggcct gtatgagaca cgcatcgacc tgtcacagct gggaggcgat     4740
gggagcccca agaaaaagcg gaaggtgtct agttaa                               4776
```

<210> 50
<211> 1588
<212> PRT
<213> 人工序列
<220>
<223> dCas9-Clo051 - 40-18170
<400> 50

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

[0052]

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

[0053]

Ser Asn Phe Asp Leu Ala Glu Asp Ala Lys Leu Gln Leu Ser Lys Asp
465 470 475 480
Thr Tyr Asp Asp Asp Leu Asp Asn Leu Leu Ala Gln Ile Gly Asp Gln
485 490 495
Tyr Ala Asp Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp Ala Ile Leu
500 505 510
Leu Ser Asp Ile Leu Arg Val Asn Thr Glu Ile Thr Lys Ala Pro Leu
515 520 525
Ser Ala Ser Met Ile Lys Arg Tyr Asp Glu His His Gln Asp Leu Thr
530 535 540
Leu Leu Lys Ala Leu Val Arg Gln Gln Leu Pro Glu Lys Tyr Lys Glu
545 550 555 560
Ile Phe Phe Asp Gln Ser Lys Asn Gly Tyr Ala Gly Tyr Ile Asp Gly
565 570 575
Gly Ala Ser Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu
580 585 590
Lys Met Asp Gly Thr Glu Glu Leu Leu Val Lys Leu Asn Arg Glu Asp
595 600 605
Leu Leu Arg Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile Pro His Gln
610 615 620
Ile His Leu Gly Glu Leu His Ala Ile Leu Arg Arg Gln Glu Asp Phe
625 630 635 640
Tyr Pro Phe Leu Lys Asp Asn Arg Glu Lys Ile Glu Lys Ile Leu Thr
645 650 655
Phe Arg Ile Pro Tyr Tyr Val Gly Pro Leu Ala Arg Gly Asn Ser Arg
660 665 670
Phe Ala Trp Met Thr Arg Lys Ser Glu Glu Thr Ile Thr Pro Trp Asn
675 680 685
Phe Glu Glu Val Val Asp Lys Gly Ala Ser Ala Gln Ser Phe Ile Glu
690 695 700
Arg Met Thr Asn Phe Asp Lys Asn Leu Pro Asn Glu Lys Val Leu Pro
705 710 715 720
Lys His Ser Leu Leu Tyr Glu Tyr Phe Thr Val Tyr Asn Glu Leu Thr
725 730 735
Lys Val Lys Tyr Val Thr Glu Gly Met Arg Lys Pro Ala Phe Leu Ser
740 745 750
Gly Glu Gln Lys Lys Ala Ile Val Asp Leu Leu Phe Lys Thr Asn Arg
755 760 765
Lys Val Thr Val Lys Gln Leu Lys Glu Asp Tyr Phe Lys Lys Ile Glu
770 775 780
Cys Phe Asp Ser Val Glu Ile Ser Gly Val Glu Asp Arg Phe Asn Ala
785 790 795 800
Ser Leu Gly Thr Tyr His Asp Leu Leu Lys Ile Ile Lys Asp Lys Asp
805 810 815

Phe Leu Asp Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp Ile Val Leu
820 825 830
Thr Leu Thr Leu Phe Glu Asp Arg Glu Met Ile Glu Glu Arg Leu Lys
835 840 845
Thr Tyr Ala His Leu Phe Asp Asp Lys Val Met Lys Gln Leu Lys Arg
850 855 860
Arg Arg Tyr Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu Ile Asn Gly
865 870 875 880
Ile Arg Asp Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe Leu Lys Ser
885 890 895
Asp Gly Phe Ala Asn Arg Asn Phe Met Gln Leu Ile His Asp Asp Ser
900 905 910
Leu Thr Phe Lys Glu Asp Ile Gln Lys Ala Gln Val Ser Gly Gln Gly
915 920 925
Asp Ser Leu His Glu His Ile Ala Asn Leu Ala Gly Ser Pro Ala Ile
930 935 940
Lys Lys Gly Ile Leu Gln Thr Val Lys Val Val Asp Glu Leu Val Lys
945 950 955 960
Val Met Gly Arg His Lys Pro Glu Asn Ile Val Ile Glu Met Ala Arg
965 970 975
Glu Asn Gln Thr Thr Gln Lys Gly Gln Lys Asn Ser Arg Glu Arg Met
980 985 990

[0054]

Lys Arg Ile Glu Glu Gly Ile Lys Glu Leu Gly Ser Gln Ile Leu Lys
995 1000 1005
Glu His Pro Val Glu Asn Thr Gln Leu Gln Asn Glu Lys Leu Tyr
1010 1015 1020
Leu Tyr Tyr Leu Gln Asn Gly Arg Asp Met Tyr Val Asp Gln Glu
1025 1030 1035
Leu Asp Ile Asn Arg Leu Ser Asp Tyr Asp Val Asp Ala Ile Val
1040 1045 1050
Pro Gln Ser Phe Leu Lys Asp Asp Ser Ile Asp Asn Lys Val Leu
1055 1060 1065
Thr Arg Ser Asp Lys Asn Arg Gly Lys Ser Asp Asn Val Pro Ser
1070 1075 1080
Glu Glu Val Val Lys Lys Met Lys Asn Tyr Trp Arg Gln Leu Leu
1085 1090 1095
Asn Ala Lys Leu Ile Thr Gln Arg Lys Phe Asp Asn Leu Thr Lys
1100 1105 1110
Ala Glu Arg Gly Gly Leu Ser Glu Leu Asp Lys Ala Gly Phe Ile
1115 1120 1125
Lys Arg Gln Leu Val Glu Thr Arg Gln Ile Thr Lys His Val Ala
1130 1135 1140
Gln Ile Leu Asp Ser Arg Met Asn Thr Lys Tyr Asp Glu Asn Asp
1145 1150 1155

[0055]

```
Lys Leu  Ile Arg Glu Val Lys  Val Ile Thr Leu Lys  Ser Lys Leu
    1160                1165                1170
Val Ser  Asp Phe Arg Lys Asp  Phe Gln Phe Tyr Lys  Val Arg Glu
    1175                1180                1185
Ile Asn  Asn Tyr His His Ala  His Asp Ala Tyr Leu  Asn Ala Val
    1190                1195                1200
Val Gly  Thr Ala Leu Ile Lys  Lys Tyr Pro Lys Leu  Glu Ser Glu
    1205                1210                1215
Phe Val  Tyr Gly Asp Tyr Lys  Val Tyr Asp Val Arg  Lys Met Ile
    1220                1225                1230
Ala Lys  Ser Glu Gln Glu Ile  Gly Lys Ala Thr Ala  Lys Tyr Phe
    1235                1240                1245
Phe Tyr  Ser Asn Ile Met Asn  Phe Phe Lys Thr Glu  Ile Thr Leu
    1250                1255                1260
Ala Asn  Gly Glu Ile Arg Lys  Arg Pro Leu Ile Glu  Thr Asn Gly
    1265                1270                1275
Glu Thr  Gly Glu Ile Val Trp  Asp Lys Gly Arg Asp  Phe Ala Thr
    1280                1285                1290
Val Arg  Lys Val Leu Ser Met  Pro Gln Val Asn Ile  Val Lys Lys
    1295                1300                1305
Thr Glu  Val Gln Thr Gly Gly  Phe Ser Lys Glu Ser  Ile Leu Pro
    1310                1315                1320
Lys Arg  Asn Ser Asp Lys Leu  Ile Ala Arg Lys Lys  Asp Trp Asp
    1325                1330                1335
Pro Lys  Lys Tyr Gly Gly Phe  Asp Ser Pro Thr Val  Ala Tyr Ser
    1340                1345                1350
Val Leu  Val Val Ala Lys Val  Glu Lys Gly Lys Ser  Lys Lys Leu
    1355                1360                1365
Lys Ser  Val Lys Glu Leu Leu  Gly Ile Thr Ile Met  Glu Arg Ser
    1370                1375                1380
Ser Phe  Glu Lys Asn Pro Ile  Asp Phe Leu Glu Ala  Lys Gly Tyr
    1385                1390                1395
Lys Glu  Val Lys Lys Asp Leu  Ile Ile Lys Leu Pro  Lys Tyr Ser
    1400                1405                1410
Leu Phe  Glu Leu Glu Asn Gly  Arg Lys Arg Met Leu  Ala Ser Ala
    1415                1420                1425
Gly Glu  Leu Gln Lys Gly Asn  Glu Leu Ala Leu Pro  Ser Lys Tyr
    1430                1435                1440
Val Asn  Phe Leu Tyr Leu Ala  Ser His Tyr Glu Lys  Leu Lys Gly
    1445                1450                1455
Ser Pro  Glu Asp Asn Glu Gln  Lys Gln Leu Phe Val  Glu Gln His
    1460                1465                1470
Lys His  Tyr Leu Asp Glu Ile  Ile Glu Gln Ile Ser  Glu Phe Ser
    1475                1480                1485
```

Lys Arg Val Ile Leu Ala Asp Ala Asn Leu Asp Lys Val Leu Ser
1490 1495 1500
Ala Tyr Asn Lys His Arg Asp Lys Pro Ile Arg Glu Gln Ala Glu
1505 1510 1515
Asn Ile Ile His Leu Phe Thr Leu Thr Asn Leu Gly Ala Pro Ala
1520 1525 1530
Ala Phe Lys Tyr Phe Asp Thr Thr Ile Asp Arg Lys Arg Tyr Thr
1535 1540 1545
Ser Thr Lys Glu Val Leu Asp Ala Thr Leu Ile His Gln Ser Ile
1550 1555 1560
Thr Gly Leu Tyr Glu Thr Arg Ile Asp Leu Ser Gln Leu Gly Gly
1565 1570 1575
Asp Gly Ser Pro Lys Lys Lys Arg Lys Val
1580 1585

<210> 51
<211> 4767
<212> DNA
<213> 人工序列
<220>
<223> dCas9-Clo051 - 40-18171
<400> 51

[0056]

atgcctaaga agaagcggaa ggtggaaggc atcaaaagca acatctccct cctgaaagac 60
gaactccggg ggcagattag ccacattagt cacgaatacc tctccctcat cgacctggct 120
ttcgatagca agcagaacag gctctttgag atgaaagtgc tggaactgct cgtcaatgag 180
tacgggttca agggtcgaca cctcggcgga tctaggaaac cagacggcat cgtgtatagt 240
accacactgg aagacaactt tgggatcatt gtggatacca aggcatactc tgagggttat 300
agtctgccca tttcacaggc cgacgagatg gaacggtacg tgcgcgagaa ctcaaataga 360
gatgaggaag tcaaccctaa caagtggtgg gagaacttct ctgaggaagt gaagaaatac 420
tacttcgtct ttatcagcgg gtccttcaag ggtaaatttg aggaacagct caggagactg 480
agcatgacta ccggcgtgaa tggcagcgcc gtcaacgtgg tcaatctgct cctgggcgct 540
gaaaagattc ggagcggaga gatgaccatc gaagagctgg agagggcaat gtttaataat 600
agcgagttta tcctgaaata cggtggcggt ggatccgata aaaagtattc tattggttta 660
gccatcggca ctaattccgt tggatgggct gtcataaccg atgaatacaa agtaccttca 720
aagaaattta aggtgttggg gaacacagac cgtcattcga ttaaaaagaa tcttatcggt 780
gccctcctat tcgatagtgg cgaaacggca gaggcgactc gcctgaaacg aaccgctcgg 840
agaaggtata cacgtcgcaa gaaccgaata tgttacttac aagaaatttt tagcaatgag 900
atggccaaag ttgacgattc tttctttcac cgtttggaag agtccttcct tgtcgaagag 960
gacaagaaac atgaacggca ccccatcttt ggaaacatag tagatgaggt ggcatatcat 1020
gaaaagtacc caacgattta tcacctcaga aaaaagctag ttgactcaac tgataaagcg 1080
gacctgaggt taatctactt ggctcttgcc catatgataa agttccgtgg gcactttctc 1140
attgagggtg atctaaatcc ggacaactcg gatgtcgaca aactgttcat ccagttagta 1200
caaacctata atcagttgtt tgaagagaac cctataaatg caagtggcgt ggatgcgaag 1260
gctattctta gcgcccgcct ctctaaatcc cgacggctag aaaacctgat cgcacaatta 1320
cccggagaga agaaaaatgg gttgttcggt aaccttatag cgctctcact aggcctgaca 1380

[0057]

```
ccaaatttta agtcgaactt cgacttagct gaagatgcca aattgcagct tagtaaggac 1440
acgtacgatg acgatctcga caatctactg gcacaaattg gagatcagta tgcggactta 1500
tttttggctg ccaaaaacct tagcgatgca atcctcctat ctgacatact gagagttaat 1560
actgagatta ccaaggcgcc gttatccgct tcaatgatca aaaggtacga tgaacatcac 1620
caagacttga cacttctcaa ggccctagtc cgtcagcaac tgcctgagaa atataaggaa 1680
atattctttg atcagtcgaa aaacgggtac gcaggttata ttgacggcgg agcgagtcaa 1740
gaggaattct acaagtttat caaacccata ttagagaaga tggatgggac ggaagagttg 1800
cttgtaaaac tcaatcgcga agatctactg cgaaagcagc ggactttcga caacggtagc 1860
attccacatc aaatccactt aggcgaattg catgctatac ttagaaggca ggaggatttt 1920
tatccgttcc tcaaagacaa tcgtgaaaag attgagaaaa tcctaacctt tcgcatacct 1980
tactatgtgg gacccctggc ccgagggaac tctcggttcg catggatgac aagaaagtcc 2040
gaagaaacga ttactccatg gaattttgag gaagttgtcg ataaaggtgc gtcagctcaa 2100
tcgttcatcg agaggatgac caactttgac aagaatttac cgaacgaaaa agtattgcct 2160
aagcacagtt tactttacga gtatttcaca gtgtacaatg aactcacgaa agttaagtat 2220
gtcactgagg gcatgcgtaa acccgccttt ctaagcggag aacagaagaa agcaatagta 2280
gatctgttat tcaagaccaa ccgcaaagtg acagttaagc aattgaaaga ggactacttt 2340
aagaaaattg aatgcttcga ttctgtcgag atctccgggg tagaagatcg atttaatgcg 2400
tcacttggta cgtatcatga cctcctaaag ataattaaag ataaggactt cctggataac 2460
gaagagaatg aagatatctt agaagatata gtgttgactc ttaccctctt tgaagatcgg 2520
gaaatgattg aggaaagact aaaaacatac gctcacctgt tcgacgataa ggttatgaaa 2580
cagttaaaga ggcgtcgcta tacgggctgg ggacgattgt cgcggaaact tatcaacggg 2640
ataagagaca agcaaagtgg taaaactatt ctcgattttc taaagagcga cggcttcgcc 2700
aataggaact ttatgcagct gatccatgat gactctttaa ccttcaaaga ggatatacaa 2760
aaggcacagg tttccggaca aggggactca ttgcacgaac atattgcgaa tcttgctggt 2820
tcgccagcca tcaaaaaggg catactccag acagtcaaag tagtggatga gctagttaag 2880
gtcatgggac gtcacaaacc ggaaaacatt gtaatcgaga tggcacgcga aaatcaaacg 2940

actcagaagg ggcaaaaaaa cagtcgagag cggatgaaga gaatagaaga gggtattaaa 3000
gaactgggca gccagatctt aaaggagcat cctgtggaaa atacccaatt gcagaacgag 3060
aaactttacc tctattacct acaaaatgga agggacatgt atgttgatca ggaactggac 3120
ataaaccgtt tatctgatta cgacgtcgat gccattgtac cccaatcctt tttgaaggac 3180
gattcaatcg acaataaagt gcttacacgc tcggataaga accgagggaa aagtgacaat 3240
gttccaagcg aggaagtcgt aaagaaaatg aagaactatt ggcggcagct cctaaatgcg 3300
aaactgataa cgcaaagaaa gttcgataac ttaactaaag ctgagagggg tggcttgtct 3360
gaacttgaca aggccggatt tattaaacgt cagctcgtgg aaacccgcca aatcacaaag 3420
catgttgcac agatactaga ttcccgaatg aatacgaaat acgacgagaa cgataagctg 3480
attcgggaag tcaaagtaat cactttaaag tcaaaattgg tgtcggactt cagaaaggat 3540
tttcaattct ataaagttag ggagataaat aactaccacc atgcgcacga cgcttatctt 3600
aatgccgtcg tagggaccgc actcattaag aaatacccga agctagaaag tgagtttgtg 3660
tatggtgatt acaaagttta tgacgtccgt aagatgatcg cgaaaagcga acaggagata 3720
ggcaaggcta cagccaaata cttcttttat tctaacatta tgaatttctt taagacggaa 3780
atcactctgg caaacggaga gatacgcaaa cgacctttaa ttgaaaccaa tggggagaca 3840
ggtgaaatcg tatgggataa gggccgggac ttcgcgacgg tgagaaaagt tttgtccatg 3900
ccccaagtca acatagtaaa gaaaactgag gtgcagaccg gagggttttc aaaggaatcg 3960
```

attcttccaa aaaggaatag tgataagctc atcgctcgta aaaaggactg ggacccgaaa 4020
aagtacggtg gcttcgatag ccctacagtt gcctattctg tcctagtagt ggcaaaagtt 4080
gagaagggaa aatccaagaa actgaagtca gtcaaagaat tattggggat aacgattatg 4140
gagcgctcgt cttttgaaaa gaaccccatc gacttccttg aggcgaaagg ttacaaggaa 4200
gtaaaaaagg atctcataat taaactacca aagtatagtc tgtttgagtt agaaaatggc 4260
cgaaaacgga tgttggctag cgccggagag cttcaaaagg ggaacgaact cgcactaccg 4320
tctaaatacg tgaatttcct gtatttagcg tcccattacg agaagttgaa aggttcacct 4380
gaagataacg aacagaagca acttttgtt gagcagcaca aacattatct cgacgaaatc 4440
atagagcaaa tttcggaatt cagtaagaga gtcatcctag ctgatgccaa tctggacaaa 4500
gtattaagcg catacaacaa gcacagggat aaacccatac gtgagcaggc ggaaaatatt 4560
atccatttgt ttactcttac caacctcggc gctccagccg cattcaagta ttttgacaca 4620
acgatagatc gcaaacgata cacttctacc aaggaggtgc tagacgcgac actgattcac 4680
caatccatca cgggattata tgaaactcgg atagatttgt cacagcttgg gggtgacgga 4740
tcccccaaga agaagaggaa agtctga 4767

<210> 52
<211> 187
<212> PRT
<213> 人工序列
<220>
<223> DHFR 突变蛋白 - 40-17012
<400> 52

[0058]

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

180 185

<210> 53
<211> 561
<212> DNA
<213> 人工序列
<220>
<223> DHFR 突变蛋白 - 40-17095
<400> 53

atggtcgggt ctctgaattg tatcgtcgcc gtgagtcaga acatgggcat tgggaagaat 60
ggcgatttcc catggccacc tctgcgcaac gagtcccgat actttcagcg gatgacaact 120
acctcctctg tggaagggaa acagaatctg gtcatcatgg gaaagaaaac ttggttcagc 180
attccagaga agaaccggcc cctgaaaggc agaatcaatc tggtgctgtc ccgagaactg 240
aaggagccac cacagggagc tcactttctg agccggtccc tggacgatgc actgaagctg 300
acagaacagc ctgagctggc caacaaagtc gatatggtgt ggatcgtcgg gggaagttca 360
gtgtataagg aggccatgaa tcaccccggc catctgaaac tgttcgtcac acggatcatg 420
caggactttg agagcgatac tttctttcct gaaattgacc tggagaagta caaactgctg 480
cccgaatatc ctggcgtgct gtccgatgtc caggaagaga aaggcatcaa atacaagttc 540
gaggtctatg agaagaatga c 561

<210> 54
<211> 18
<212> PRT
<213> Homo Sapiens

[0059]

<400> 54

Glu Gly Arg Gly Ser Leu Leu Thr Cys Gly Asp Val Glu Glu Asn Pro
1 5 10 15
Gly Pro

<210> 55
<211> 21
<212> PRT
<213> Homo Sapiens
<400> 55

Gly Ser Gly Glu Gly Arg Gly Ser Leu Leu Thr Cys Gly Asp Val Glu
1 5 10 15
Glu Asn Pro Gly Pro
20

<210> 56
<211> 63
<212> DNA
<213> Homo Sapiens
<400> 56

ggatctggag agggaagggg aagcctgctg acctgtggag acgtggagga aaacccagga 60
cca 63

<210> 57

<211> 20
<212> PRT
<213> Homo Sapiens
<400> 57
Gln Cys Thr Asn Tyr Ala Leu Leu Lys Leu Ala Gly Asp Val Glu Ser
1 5 10 15
Asn Pro Gly Pro
20
<210> 58
<211> 23
<212> PRT
<213> Homo Sapiens
<400> 58
Gly Ser Gly Gln Cys Thr Asn Tyr Ala Leu Leu Lys Leu Ala Gly Asp
1 5 10 15
Val Glu Ser Asn Pro Gly Pro
20
<210> 59
<211> 22
<212> PRT
<213> Homo Sapiens
[0060] <400> 59
Val Lys Gln Thr Leu Asn Phe Asp Leu Leu Lys Leu Ala Gly Asp Val
1 5 10 15
Glu Ser Asn Pro Gly Pro
20
<210> 60
<211> 25
<212> PRT
<213> Homo Sapiens
<400> 60
Gly Ser Gly Val Lys Gln Thr Leu Asn Phe Asp Leu Leu Lys Leu Ala
1 5 10 15
Gly Asp Val Glu Ser Asn Pro Gly Pro
20 25
<210> 61
<211> 19
<212> PRT
<213> Homo Sapiens
<400> 61
Ala Thr Asn Phe Ser Leu Leu Lys Gln Ala Gly Asp Val Glu Glu Asn
1 5 10 15
Pro Gly Pro

<210> 62
<211> 22
<212> PRT
<213> Homo Sapiens
<400> 62
Gly Ser Gly Ala Thr Asn Phe Ser Leu Leu Lys Gln Ala Gly Asp Val
1 5 10 15
Glu Glu Asn Pro Gly Pro
20
<210> 63
<211> 594
<212> PRT
<213> 人工序列
<220>
<223> T2A - 40-18099
<400> 63
Met Gly Ser Ser Leu Asp Asp Glu His Ile Leu Ser Ala Leu Leu Gln
1 5 10 15
Ser Asp Asp Glu Leu Val Gly Glu Asp Ser Asp Ser Glu Ile Ser Asp
20 25 30
His Val Ser Glu Asp Asp Val Gln Ser Asp Thr Glu Glu Ala Phe Ile
35 40 45
Asp Glu Val His Glu Val Gln Pro Thr Ser Ser Gly Ser Glu Ile Leu
50 55 60
Asp Glu Gln Asn Val Ile Glu Gln Pro Gly Ser Ser Leu Ala Ser Asn
65 70 75 80
Arg Ile Leu Thr Leu Pro Gln Arg Thr Ile Arg Gly Lys Asn Lys His
85 90 95
Cys Trp Ser Thr Ser Lys Ser Thr Arg Arg Ser Arg Val Ser Ala Leu
100 105 110
Asn Ile Val Arg Ser Gln Arg Gly Pro Thr Arg Met Cys Arg Asn Ile
115 120 125
Tyr Asp Pro Leu Leu Cys Phe Lys Leu Phe Phe Thr Asp Glu Ile Ile
130 135 140
Ser Glu Ile Val Lys Trp Thr Asn Ala Glu Ile Ser Leu Lys Arg Arg
145 150 155 160
Glu Ser Met Thr Gly Ala Thr Phe Arg Asp Thr Asn Glu Asp Glu Ile
165 170 175
Tyr Ala Phe Phe Gly Ile Leu Val Met Thr Ala Val Arg Lys Asp Asn
180 185 190
His Met Ser Thr Asp Asp Leu Phe Asp Arg Ser Leu Ser Met Val Tyr
195 200 205
Val Ser Val Met Ser Arg Asp Arg Phe Asp Phe Leu Ile Arg Cys Leu
210 215 220

[0061]

[0062]

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

Cys Lys Lys Val Ile Cys Arg Glu His Asn Ile Asp Met Cys Gln Ser
580 585 590
Cys Phe

<210> 64
<211> 594
<212> PRT
<213> 人工序列
<220>
<223> SPB 转座酶-40-14484
<400> 64

[0063]

Met Gly Ser Ser Leu Asp Asp Glu His Ile Leu Ser Ala Leu Leu Gln
1 5 10 15
Ser Asp Asp Glu Leu Val Gly Glu Asp Ser Asp Ser Glu Val Ser Asp
20 25 30
His Val Ser Glu Asp Asp Val Gln Ser Asp Thr Glu Glu Ala Phe Ile
35 40 45
Asp Glu Val His Glu Val Gln Pro Thr Ser Ser Gly Ser Glu Ile Leu
50 55 60
Asp Glu Gln Asn Val Ile Glu Gln Pro Gly Ser Ser Leu Ala Ser Asn
65 70 75 80
Arg Ile Leu Thr Leu Pro Gln Arg Thr Ile Arg Gly Lys Asn Lys His
85 90 95
Cys Trp Ser Thr Ser Lys Ser Thr Arg Arg Ser Arg Val Ser Ala Leu
100 105 110
Asn Ile Val Arg Ser Gln Arg Gly Pro Thr Arg Met Cys Arg Asn Ile
115 120 125
Tyr Asp Pro Leu Leu Cys Phe Lys Leu Phe Phe Thr Asp Glu Ile Ile
130 135 140
Ser Glu Ile Val Lys Trp Thr Asn Ala Glu Ile Ser Leu Lys Arg Arg
145 150 155 160
Glu Ser Met Thr Ser Ala Thr Phe Arg Asp Thr Asn Glu Asp Glu Ile
165 170 175
Tyr Ala Phe Phe Gly Ile Leu Val Met Thr Ala Val Arg Lys Asp Asn
180 185 190
His Met Ser Thr Asp Asp Leu Phe Asp Arg Ser Leu Ser Met Val Tyr
195 200 205
Val Ser Val Met Ser Arg Asp Arg Phe Asp Phe Leu Ile Arg Cys Leu
210 215 220
Arg Met Asp Asp Lys Ser Ile Arg Pro Thr Leu Arg Glu Asn Asp Val
225 230 235 240
Phe Thr Pro Val Arg Lys Ile Trp Asp Leu Phe Ile His Gln Cys Ile
245 250 255
Gln Asn Tyr Thr Pro Gly Ala His Leu Thr Ile Asp Glu Gln Leu Leu

[0064]

260 265 270

Gly Phe Arg Gly Arg Cys Pro Phe Arg Val Tyr Ile Pro Asn Lys Pro
275 280 285

Ser Lys Tyr Gly Ile Lys Ile Leu Met Met Cys Asp Ser Gly Thr Lys
290 295 300

Tyr Met Ile Asn Gly Met Pro Tyr Leu Gly Arg Gly Thr Gln Thr Asn
305 310 315 320

Gly Val Pro Leu Gly Glu Tyr Tyr Val Lys Glu Leu Ser Lys Pro Val
325 330 335

His Gly Ser Cys Arg Asn Ile Thr Cys Asp Asn Trp Phe Thr Ser Ile
340 345 350

Pro Leu Ala Lys Asn Leu Leu Gln Glu Pro Tyr Lys Leu Thr Ile Val
355 360 365

Gly Thr Val Arg Ser Asn Lys Arg Glu Ile Pro Glu Val Leu Lys Asn
370 375 380

Ser Arg Ser Arg Pro Val Gly Thr Ser Met Phe Cys Phe Asp Gly Pro
385 390 395 400

Leu Thr Leu Val Ser Tyr Lys Pro Lys Pro Ala Lys Met Val Tyr Leu
405 410 415

Leu Ser Ser Cys Asp Glu Asp Ala Ser Ile Asn Glu Ser Thr Gly Lys
420 425 430

Pro Gln Met Val Met Tyr Tyr Asn Gln Thr Lys Gly Gly Val Asp Thr
435 440 445

Leu Asp Gln Met Cys Ser Val Met Thr Cys Ser Arg Lys Thr Asn Arg
450 455 460

Trp Pro Met Ala Leu Leu Tyr Gly Met Ile Asn Ile Ala Cys Ile Asn
465 470 475 480

Ser Phe Ile Ile Tyr Ser His Asn Val Ser Ser Lys Gly Glu Lys Val
485 490 495

Gln Ser Arg Lys Lys Phe Met Arg Asn Leu Tyr Met Ser Leu Thr Ser
500 505 510

Ser Phe Met Arg Lys Arg Leu Glu Ala Pro Thr Leu Lys Arg Tyr Leu
515 520 525

Arg Asp Asn Ile Ser Asn Ile Leu Pro Lys Glu Val Pro Gly Thr Ser
530 535 540

Asp Asp Ser Thr Glu Glu Pro Val Met Lys Lys Arg Thr Tyr Cys Thr
545 550 555 560

Tyr Cys Pro Ser Lys Ile Arg Arg Lys Ala Asn Ala Ser Cys Lys Lys
565 570 575

Cys Lys Lys Val Ile Cys Arg Glu His Asn Ile Asp Met Cys Gln Ser
580 585 590

Cys Phe

<210> 65

<211> 340
<212> PRT
<213> 人工序列
<220>
<223> SB100X 转座酶- 40-14485
<400> 65

[0065]

Met Gly Lys Ser Lys Glu Ile Ser Gln Asp Leu Arg Lys Lys Ile Val
1 5 10 15
Asp Leu His Lys Ser Gly Ser Ser Leu Gly Ala Ile Ser Lys Arg Leu
20 25 30
Lys Val Pro Arg Ser Ser Val Gln Thr Ile Val Arg Lys Tyr Lys His
35 40 45
His Gly Thr Thr Gln Pro Ser Tyr Arg Ser Gly Arg Arg Arg Val Leu
50 55 60
Ser Pro Arg Asp Glu Arg Thr Leu Val Arg Lys Val Gln Ile Asn Pro
65 70 75 80
Arg Thr Thr Ala Lys Asp Leu Val Lys Met Leu Glu Glu Thr Gly Thr
85 90 95
Lys Val Ser Ile Ser Thr Val Lys Arg Val Leu Tyr Arg His Asn Leu
100 105 110
Lys Gly Arg Ser Ala Arg Lys Lys Pro Leu Leu Gln Asn Arg His Lys
115 120 125
Lys Ala Arg Leu Arg Phe Ala Thr Ala His Gly Asp Lys Asp Arg Thr
130 135 140
Phe Trp Arg Asn Val Leu Trp Ser Asp Glu Thr Lys Ile Glu Leu Phe
145 150 155 160
Gly His Asn Asp His Arg Tyr Val Trp Arg Lys Lys Gly Glu Ala Cys
165 170 175
Lys Pro Lys Asn Thr Ile Pro Thr Val Lys His Gly Gly Gly Ser Ile
180 185 190
Met Leu Trp Gly Cys Phe Ala Ala Gly Gly Thr Gly Ala Leu His Lys
195 200 205
Ile Asp Gly Ile Met Arg Lys Glu Asn Tyr Val Asp Ile Leu Lys Gln
210 215 220
His Leu Lys Thr Ser Val Arg Lys Leu Lys Leu Gly Arg Lys Trp Val
225 230 235 240
Phe Gln Met Asp Asn Asp Pro Lys His Thr Ser Lys Val Val Ala Lys
245 250 255
Trp Leu Lys Asp Asn Lys Val Lys Val Leu Glu Trp Pro Ser Gln Ser
260 265 270
Pro Asp Leu Asn Pro Ile Glu Asn Leu Trp Ala Glu Leu Lys Lys Arg
275 280 285
Val Arg Ala Arg Arg Pro Thr Asn Leu Thr Gln Leu His Gln Leu Cys
290 295 300

Gln Glu Glu Trp Ala Lys Ile His Pro Thr Tyr Cys Gly Lys Leu Val
305 310 315 320
Glu Gly Tyr Pro Lys Arg Leu Thr Gln Val Lys Gln Phe Lys Gly Asn
325 330 335
Ala Thr Lys Tyr
340

<210> 66
<211> 340
<212> PRT
<213> 人工序列
<220>
<223> hyperSB100X 转座酶-40-14486

[0066]

<400> 66
Met Gly Lys Ser Lys Glu Ile Ser Gln Asp Leu Arg Lys Arg Ile Val
1 5 10 15
Asp Leu His Lys Ser Gly Ser Ser Leu Gly Ala Ile Ser Lys Arg Leu
20 25 30
Ala Val Pro Arg Ser Ser Val Gln Thr Ile Val Arg Lys Tyr Lys His
35 40 45
His Gly Thr Thr Gln Pro Ser Tyr Arg Ser Gly Arg Arg Arg Val Leu
50 55 60
Ser Pro Arg Asp Glu Arg Thr Leu Val Arg Lys Val Gln Ile Asn Pro
65 70 75 80
Arg Thr Thr Ala Lys Asp Leu Val Lys Met Leu Glu Glu Thr Gly Thr
85 90 95
Lys Val Ser Ile Ser Thr Val Lys Arg Val Leu Tyr Arg His Asn Leu
100 105 110
Lys Gly His Ser Ala Arg Lys Lys Pro Leu Leu Gln Asn Arg His Lys
115 120 125
Lys Ala Arg Leu Arg Phe Ala Thr Ala His Gly Asp Lys Asp Arg Thr
130 135 140
Phe Trp Arg Asn Val Leu Trp Ser Asp Glu Thr Lys Ile Glu Leu Phe
145 150 155 160
Gly His Asn Asp His Arg Tyr Val Trp Arg Lys Lys Gly Glu Ala Cys
165 170 175
Lys Pro Lys Asn Thr Ile Pro Thr Val Lys His Gly Gly Gly Ser Ile
180 185 190
Met Leu Trp Gly Cys Phe Ala Ala Gly Gly Thr Gly Ala Leu His Lys
195 200 205
Ile Asp Gly Ile Met Asp Ala Val Gln Tyr Val Asp Ile Leu Lys Gln
210 215 220
His Leu Lys Thr Ser Val Arg Lys Leu Lys Leu Gly Arg Lys Trp Val
225 230 235 240
Phe Gln His Asp Asn Asp Pro Lys His Thr Ser Lys Val Val Ala Lys

245 250 255
Trp Leu Lys Asp Asn Lys Val Lys Val Leu Glu Trp Pro Ser Gln Ser
260 265 270
Pro Asp Leu Asn Pro Ile Glu Asn Leu Trp Ala Glu Leu Lys Lys Arg
275 280 285
Val Arg Ala Arg Arg Pro Thr Asn Leu Thr Gln Leu His Gln Leu Cys
290 295 300
Gln Glu Glu Trp Ala Lys Ile His Pro Asn Tyr Cys Gly Lys Leu Val
305 310 315 320
Glu Gly Tyr Pro Lys Arg Leu Thr Gln Val Lys Gln Phe Lys Gly Asn
325 330 335
Ala Thr Lys Tyr
340

<210> 67
<211> 5296
<212> DNA
<213> 人工序列
<220>
<223> Helraiser转座子 - 40-18174
<400> 67

[0067]

tcctatataa taaaagagaa acatgcaaat tgaccatccc tccgctacgc tcaagccacg 60
cccaccagcc aatcagaagt gactatgcaa attaacccaa caaagatggc agttaaattt 120
gcatacgcag gtgtcaagcg ccccaggagg caacggcggc cgcgggctcc caggaccttc 180
gctggccccg ggaggcgagg ccggccgcgc ctagccacac ccgcgggctc ccgggacctt 240
cgccagcaga gagcagagcg ggagagcggg cggagagcgg gaggtttgga ggacttggca 300
gagcaggagg ccgctggaca tagagcagag cgagagagag ggtggcttgg agggcgtggc 360
tccctctgtc accccagctt cctcatcaca gctgtggaaa ctgacagcag ggaggaggaa 420
gtcccacccc cacagaatca gccagaatca gccgttggtc agacagctct cagcggcctg 480
acagccagga ctctcattca cctgcatctc agaccgtgac agtagagagg tgggactatg 540
tctaaagaac aactgttgat acaacgtagc tctgcagccg aaagatgccg gcgttatcga 600
cagaaaatgt ctgcagagca acgtgcgtct gatcttgaaa gaaggcggcg cctgcaacag 660
aatgtatctg aagagcagct actggaaaaa cgtcgctctg aagccgaaaa acagcggcgt 720
catcgacaga aaatgtctaa agaccaacgt gcctttgaag ttgaaagaag gcggtggcga 780
cgacagaata tgtctagaga acagtcatca acaagtacta ccaataccgg taggaactgc 840
cttctcagca aaaatggagt acatgaggat gcaattctcg aacatagttg tggtggaatg 900
actgttcgat gtgaattttg cctatcacta aatttctctg atgaaaaacc atccgatggg 960
aaatttactc gatgttgtag caaagggaaa gtctgtccaa atgatataca ttttccagat 1020
tacccggcat atttaaaaag attaatgaca aacgaagatt ctgacagtaa aaatttcatg 1080
gaaaatattc gttccataaa tagttctttt gcttttgctt ccatgggtgc aaatattgca 1140
tcgccatcag gatatgggcc atactgtttt agaatacacg gacaagttta tcaccgtact 1200
ggaactttac atccttcgga tggtgtttct cggaagtttg ctcaactcta tattttggat 1260
acagccgaag ctacaagtaa aagattagca atgccagaaa accagggctg ctcagaaaga 1320
ctcatgatca acatcaacaa cctcatgcat gaaataaatg aattaacaaa atcgtacaag 1380
atgctacatg aggtagaaaa ggaagcccaa tctgaagcag cagcaaaagg tattgctccc 1440

[0068]

acagaagtaa caatggcgat taaatacgat cgtaacagtg acccaggtag atataattct 1500
ccccgtgtaa ccgaggttgc tgtcatattc agaaacgaag atggagaacc tccttttgaa 1560
agggacttgc tcattcattg taaaccagat cccaataatc caaatgccac taaaatgaaa 1620
caaatcagta tcctgtttcc tacattagat gcaatgacat atcctattct ttttccacat 1680
ggtgaaaaag gctggggaac agatattgca ttaagactca gagacaacag tgtaatcgac 1740
aataatacta gacaaaatgt aaggacacga gtcacacaaa tgcagtatta tggatttcat 1800
ctctctgtgc gggacacgtt caatcctatt ttaaatgcag gaaaattaac tcaacagttt 1860
attgtggatt catattcaaa aatggaggcc aatcggataa atttcatcaa agcaaaccaa 1920
tctaagttga gagttgaaaa atatagtggt ttgatggatt atctcaaatc tagatctgaa 1980
aatgacaatg tgccgattgg taaaatgata atacttccat catcttttga gggtagtccc 2040
agaaatatgc agcagcgata tcaggatgct atggcaattg taacgaagta tggcaagccc 2100
gatttattca taaccatgac atgcaacccc aaatgggcag atattacaaa caatttacaa 2160
cgctggcaaa aagttgaaaa cagacctgac ttggtagcca gagtttttaa tattaagctg 2220
aatgctcttt taaatgatat atgtaaattc catttatttg gcaaagtaat agctaaaatt 2280
catgtcattg aatttcagaa acgcggactg cctcacgctc acatattatt gatattagat 2340
agtgagtcca aattacgttc agaagatgac attgaccgta tagttaaggc agaaattcca 2400
gatgaagacc agtgtcctcg acttttttcaa attgtaaaat caaatatggt acatggacca 2460
tgtggaatac aaaatccaaa tagtccatgt atggaaaatg gaaaatgttc aaagggatat 2520
ccaaaagaat ttcaaaatgc gaccattgga aatattgatg gatatcccaa atacaaacga 2580
agatctggta gcaccatgtc tattggaaat aaagttgtcg ataacacttg gattgtccct 2640
tataacccgt atttgtgcct taaatataac tgtcatataa atgttgaagt ctgtgcatca 2700
attaaaagtg tcaaatattt atttaaatac atctataaag ggcacgattg tgcaaatatt 2760
caaatttctg aaaaaaatat tatcaatcat gacgaagtac aggacttcat tgactccagg 2820
tatgtgagcg ctcctgaggc tgtttggaga ctttttgcaa tgcgaatgca tgaccaatct 2880
catgcaatca caagattagc tattcatttg ccaaatgatc agaatttgta ttttcatacc 2940
gatgattttg ctgaagtttt agatagggct aaaaggcata actcgacttt gatggcttgg 3000
ttcttattga atagagaaga ttctgatgca cgtaattatt attattggga gattccacag 3060
cattatgtgt ttaataattc tttgtggaca aaacgccgaa agggtgggaa taaagtatta 3120
ggtagactgt tcactgtgag ctttagagaa ccagaacgat attaccttag acttttgctt 3180
ctgcatgtaa aaggtgcgat aagttttgag gatctgcgaa ctgtaggagg tgtaacttat 3240
gatacatttc atgaagctgc taaacaccga ggattattac ttgatgacac tatctggaaa 3300
gatacgattg acgatgcaat catccttaat atgcccaaac aactacggca acttttgca 3360
tatatatgtg tgtttggatg tccttctgct gcagacaaat tatgggatga gaataaatct 3420
cattttattg aagatttctg ttggaaatta caccgaagag aaggtgcctg tgtgaactgt 3480
gaaatgcatg cccttaacga aattcaggag gtattcacat tgcatggaat gaaatgttca 3540
catttcaaac ttccggacta tcctttatta atgaatgcaa atacatgtga tcaattgtac 3600
gagcaacaac aggcagaggt tttgataaat tctctgaatg atgaacagtt ggcagccttt 3660
cagactataa cttcagccat cgaagatcaa actgtacacc ccaaatgctt tttcttggat 3720
ggtccaggtg gtagtggaaa aacatatctg tataaagttt taacacatta tattagaggt 3780
cgtggtggta ctgttttacc cacagcatct acaggaattg ctgcaaattt acttcttggt 3840
ggaagaacct ttcattccca atataaatta ccaattccat taaatgaaac ttcaatttct 3900
agactcgata taaagagtga agttgctaaa accattaaaa aggcccaact tctcattatt 3960
gatgaatgca ccatggcatc cagtcatgct ataaacgcca tagatagatt actaagagaa 4020
attatgaatt tgaatgttgc atttggtggg aaagttctcc ttctcggagg ggattttcga 4080

caatgtctca gtattgtacc acatgctatg cgatcggcca tagtacaaac gagtttaaag 4140
tactgtaatg tttggggatg tttcagaaag ttgtctctta aaacaaatat gagatcagag 4200
gattctgctt atagtgaatg gttagtaaaa cttggagatg gcaaacttga tagcagtttt 4260
catttaggaa tggatattat tgaaatcccc catgaaatga tttgtaacgg atctattatt 4320
gaagctacct ttggaaatag tatatctata gataatatta aaaatatatc taaacgtgca 4380
attctttgtc caaaaaatga gcatgttcaa aaattaaatg aagaaatttt ggatatactt 4440

gatggagatt ttcacacata tttgagtgat gattccattg attcaacaga tgatgctgaa 4500
aaggaaaatt ttcccatcga atttcttaat agtattactc cttcgggaat gccgtgtcat 4560
aaattaaaat tgaaagtggg tgcaatcatc atgctattga gaaatcttaa tagtaaatgg 4620
ggtctttgta atggtactag atttattatc aaaagattac gacctaacat tatcgaagct 4680
gaagtattaa caggatctgc agagggagag gttgttctga ttccaagaat tgatttgtcc 4740
ccatctgaca ctggcctccc atttaaatta attcgaagac agtttcccgt gatgccagca 4800
tttgcgatga ctattaataa atcacaagga caaactctag acagagtagg aatattccta 4860
cctgaacccg ttttcgcaca tggtcagtta tatgttgctt tctctcgagt tcgaagagca 4920
tgtgacgtta aagttaaagt tgtaaatact tcatcacaag ggaaattagt caagcactct 4980
gaaagtgttt ttactcttaa tgtggtatac agggagatat tagaataagt ttaatcactt 5040
tatcagtcat tgtttgcatc aatgttgttt ttatatcatg tttttgttgt ttttatatca 5100
tgtctttgtt gttgttatat catgttgtta ttgtttattt attaataaat ttatgtatta 5160
ttttcatata cattttactc atttcctttc atctctcaca cttctattat agagaaaggg 5220
caaatagcaa tattaaaata tttcctctaa ttaattccct ttcaatgtgc acgaatttcg 5280
tgcaccgggc cactag 5296

[0069]

<210> 68
<211> 1496
<212> PRT
<213> 人工序列
<220>
<223> Helitron转座酶- 40-14488
<400> 68

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

[0070]

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

[0071]

```
465                 470                 475                 480
Glu Asn Asp Asn Val Pro Ile Gly Lys Met Ile Ile Leu Pro Ser Ser
                485                 490                 495
Phe Glu Gly Ser Pro Arg Asn Met Gln Gln Arg Tyr Gln Asp Ala Met
            500                 505                 510
Ala Ile Val Thr Lys Tyr Gly Lys Pro Asp Leu Phe Ile Thr Met Thr
        515                 520                 525
Cys Asn Pro Lys Trp Ala Asp Ile Thr Asn Asn Leu Gln Arg Trp Gln
    530                 535                 540
Lys Val Glu Asn Arg Pro Asp Leu Val Ala Arg Val Phe Asn Ile Lys
545                 550                 555                 560
Leu Asn Ala Leu Leu Asn Asp Ile Cys Lys Phe His Leu Phe Gly Lys
                565                 570                 575
Val Ile Ala Lys Ile His Val Ile Glu Phe Gln Lys Arg Gly Leu Pro
            580                 585                 590
His Ala His Ile Leu Leu Ile Leu Asp Ser Glu Ser Lys Leu Arg Ser
        595                 600                 605
Glu Asp Asp Ile Asp Arg Ile Val Lys Ala Glu Ile Pro Asp Glu Asp
    610                 615                 620
Gln Cys Pro Arg Leu Phe Gln Ile Val Lys Ser Asn Met Val His Gly
625                 630                 635                 640
Pro Cys Gly Ile Gln Asn Pro Asn Ser Pro Cys Met Glu Asn Gly Lys
                645                 650                 655
Cys Ser Lys Gly Tyr Pro Lys Glu Phe Gln Asn Ala Thr Ile Gly Asn
            660                 665                 670
Ile Asp Gly Tyr Pro Lys Tyr Lys Arg Arg Ser Gly Ser Thr Met Ser
        675                 680                 685
Ile Gly Asn Lys Val Val Asp Asn Thr Trp Ile Val Pro Tyr Asn Pro
    690                 695                 700
Tyr Leu Cys Leu Lys Tyr Asn Cys His Ile Asn Val Glu Val Cys Ala
705                 710                 715                 720
Ser Ile Lys Ser Val Lys Tyr Leu Phe Lys Tyr Ile Tyr Lys Gly His
                725                 730                 735
Asp Cys Ala Asn Ile Gln Ile Ser Glu Lys Asn Ile Ile Asn His Asp
            740                 745                 750
Glu Val Gln Asp Phe Ile Asp Ser Arg Tyr Val Ser Ala Pro Glu Ala
        755                 760                 765
Val Trp Arg Leu Phe Ala Met Arg Met His Asp Gln Ser His Ala Ile
    770                 775                 780
Thr Arg Leu Ala Ile His Leu Pro Asn Asp Gln Asn Leu Tyr Phe His
785                 790                 795                 800
Thr Asp Asp Phe Ala Glu Val Leu Asp Arg Ala Lys Arg His Asn Ser
                805                 810                 815
Thr Leu Met Ala Trp Phe Leu Leu Asn Arg Glu Asp Ser Asp Ala Arg
```

820 825 830
Asn Tyr Tyr Tyr Trp Glu Ile Pro Gln His Tyr Val Phe Asn Asn Ser
835 840 845
Leu Trp Thr Lys Arg Arg Lys Gly Gly Asn Lys Val Leu Gly Arg Leu
850 855 860
Phe Thr Val Ser Phe Arg Glu Pro Glu Arg Tyr Tyr Leu Arg Leu Leu
865 870 875 880
Leu Leu His Val Lys Gly Ala Ile Ser Phe Glu Asp Leu Arg Thr Val
885 890 895
Gly Gly Val Thr Tyr Asp Thr Phe His Glu Ala Ala Lys His Arg Gly
900 905 910
Leu Leu Leu Asp Asp Thr Ile Trp Lys Asp Thr Ile Asp Asp Ala Ile
915 920 925
Ile Leu Asn Met Pro Lys Gln Leu Arg Gln Leu Phe Ala Tyr Ile Cys
930 935 940
Val Phe Gly Cys Pro Ser Ala Ala Asp Lys Leu Trp Asp Glu Asn Lys
945 950 955 960
Ser His Phe Ile Glu Asp Phe Cys Trp Lys Leu His Arg Arg Glu Gly
965 970 975
Ala Cys Val Asn Cys Glu Met His Ala Leu Asn Glu Ile Gln Glu Val
980 985 990
[0072] Phe Thr Leu His Gly Met Lys Cys Ser His Phe Lys Leu Pro Asp Tyr
995 1000 1005
Pro Leu Leu Met Asn Ala Asn Thr Cys Asp Gln Leu Tyr Glu Gln
1010 1015 1020
Gln Gln Ala Glu Val Leu Ile Asn Ser Leu Asn Asp Glu Gln Leu
1025 1030 1035
Ala Ala Phe Gln Thr Ile Thr Ser Ala Ile Glu Asp Gln Thr Val
1040 1045 1050
His Pro Lys Cys Phe Phe Leu Asp Gly Pro Gly Gly Ser Gly Lys
1055 1060 1065
Thr Tyr Leu Tyr Lys Val Leu Thr His Tyr Ile Arg Gly Arg Gly
1070 1075 1080
Gly Thr Val Leu Pro Thr Ala Ser Thr Gly Ile Ala Ala Asn Leu
1085 1090 1095
Leu Leu Gly Gly Arg Thr Phe His Ser Gln Tyr Lys Leu Pro Ile
1100 1105 1110
Pro Leu Asn Glu Thr Ser Ile Ser Arg Leu Asp Ile Lys Ser Glu
1115 1120 1125
Val Ala Lys Thr Ile Lys Lys Ala Gln Leu Leu Ile Ile Asp Glu
1130 1135 1140
Cys Thr Met Ala Ser Ser His Ala Ile Asn Ala Ile Asp Arg Leu
1145 1150 1155
Leu Arg Glu Ile Met Asn Leu Asn Val Ala Phe Gly Gly Lys Val

[0073]

1160 1165 1170
Leu Leu Leu Gly Gly Asp Phe Arg Gln Cys Leu Ser Ile Val Pro
1175 1180 1185
His Ala Met Arg Ser Ala Ile Val Gln Thr Ser Leu Lys Tyr Cys
1190 1195 1200
Asn Val Trp Gly Cys Phe Arg Lys Leu Ser Leu Lys Thr Asn Met
1205 1210 1215
Arg Ser Glu Asp Ser Ala Tyr Ser Glu Trp Leu Val Lys Leu Gly
1220 1225 1230
Asp Gly Lys Leu Asp Ser Ser Phe His Leu Gly Met Asp Ile Ile
1235 1240 1245
Glu Ile Pro His Glu Met Ile Cys Asn Gly Ser Ile Ile Glu Ala
1250 1255 1260
Thr Phe Gly Asn Ser Ile Ser Ile Asp Asn Ile Lys Asn Ile Ser
1265 1270 1275
Lys Arg Ala Ile Leu Cys Pro Lys Asn Glu His Val Gln Lys Leu
1280 1285 1290
Asn Glu Glu Ile Leu Asp Ile Leu Asp Gly Asp Phe His Thr Tyr
1295 1300 1305
Leu Ser Asp Asp Ser Ile Asp Ser Thr Asp Asp Ala Glu Lys Glu
1310 1315 1320
Asn Phe Pro Ile Glu Phe Leu Asn Ser Ile Thr Pro Ser Gly Met
1325 1330 1335
Pro Cys His Lys Leu Lys Leu Lys Val Gly Ala Ile Ile Met Leu
1340 1345 1350
Leu Arg Asn Leu Asn Ser Lys Trp Gly Leu Cys Asn Gly Thr Arg
1355 1360 1365
Phe Ile Ile Lys Arg Leu Arg Pro Asn Ile Ile Glu Ala Glu Val
1370 1375 1380
Leu Thr Gly Ser Ala Glu Gly Glu Val Val Leu Ile Pro Arg Ile
1385 1390 1395
Asp Leu Ser Pro Ser Asp Thr Gly Leu Pro Phe Lys Leu Ile Arg
1400 1405 1410
Arg Gln Phe Pro Val Met Pro Ala Phe Ala Met Thr Ile Asn Lys
1415 1420 1425
Ser Gln Gly Gln Thr Leu Asp Arg Val Gly Ile Phe Leu Pro Glu
1430 1435 1440
Pro Val Phe Ala His Gly Gln Leu Tyr Val Ala Phe Ser Arg Val
1445 1450 1455
Arg Arg Ala Cys Asp Val Lys Val Lys Val Val Asn Thr Ser Ser
1460 1465 1470
Gln Gly Lys Leu Val Lys His Ser Glu Ser Val Phe Thr Leu Asn
1475 1480 1485
Val Val Tyr Arg Glu Ile Leu Glu

1490 1495

<210> 69
<211> 4682
<212> PRT
<213> 人工序列
<220>
<223> Tol2转座子 - 40-14491
<400> 69

Cys Ala Gly Ala Gly Gly Thr Gly Thr Ala Ala Ala Gly Thr Ala Cys
1 5 10 15
Thr Thr Gly Ala Gly Thr Ala Ala Thr Thr Thr Thr Ala Cys Thr Thr
20 25 30
Gly Ala Thr Thr Ala Cys Thr Gly Thr Ala Cys Thr Thr Ala Ala Gly
35 40 45
Thr Ala Thr Thr Ala Thr Thr Thr Thr Thr Gly Gly Gly Gly Ala Thr
50 55 60
Thr Thr Thr Thr Ala Cys Thr Thr Thr Ala Cys Thr Thr Gly Ala Gly
65 70 75 80
Thr Ala Cys Ala Ala Thr Thr Ala Ala Ala Ala Ala Thr Cys Ala Ala
85 90 95
Thr Ala Cys Thr Thr Thr Thr Ala Cys Thr Thr Thr Thr Ala Cys Thr
100 105 110
Thr Ala Ala Thr Thr Ala Cys Ala Thr Thr Thr Thr Thr Thr Thr Ala
115 120 125
Gly Ala Ala Ala Ala Ala Ala Ala Ala Gly Thr Ala Cys Thr Thr Thr
130 135 140
Thr Thr Ala Cys Thr Cys Cys Thr Thr Ala Cys Ala Ala Thr Thr Thr
145 150 155 160
Thr Ala Thr Thr Thr Ala Cys Ala Gly Thr Cys Ala Ala Ala Ala Ala
165 170 175
Gly Thr Ala Cys Thr Thr Ala Thr Thr Thr Thr Thr Thr Gly Gly Ala
180 185 190
Gly Ala Thr Cys Ala Cys Thr Thr Cys Ala Thr Thr Cys Thr Ala Thr
195 200 205
Thr Thr Thr Cys Cys Cys Thr Thr Gly Cys Thr Ala Thr Thr Ala Cys
210 215 220
Cys Ala Ala Ala Cys Cys Ala Ala Thr Thr Gly Ala Ala Thr Thr Gly
225 230 235 240
Cys Gly Cys Thr Gly Ala Thr Gly Cys Cys Cys Ala Gly Thr Thr Thr
245 250 255
Ala Ala Thr Thr Thr Ala Ala Ala Thr Gly Thr Thr Ala Thr Thr Thr
260 265 270
Ala Thr Thr Cys Thr Gly Cys Cys Thr Ala Thr Gly Ala Ala Ala Ala
275 280 285

[0074]

[0075]

Thr Cys Gly Thr Thr Thr Thr Cys Ala Cys Ala Thr Thr Ala Thr Ala
290 295 300
Thr Gly Ala Ala Ala Thr Thr Gly Gly Thr Cys Ala Gly Ala Cys Ala
305 310 315 320
Thr Gly Thr Thr Cys Ala Thr Thr Gly Gly Thr Cys Cys Thr Thr Thr
325 330 335
Gly Gly Ala Ala Gly Thr Gly Ala Cys Gly Thr Cys Ala Thr Gly Thr
340 345 350
Cys Ala Cys Ala Thr Cys Thr Ala Thr Thr Ala Cys Cys Ala Cys Ala
355 360 365
Ala Thr Gly Cys Ala Cys Ala Gly Cys Ala Cys Cys Thr Thr Gly Ala
370 375 380
Cys Cys Thr Gly Gly Ala Ala Ala Thr Thr Ala Gly Gly Gly Ala Ala
385 390 395 400
Ala Thr Thr Ala Thr Ala Ala Cys Ala Gly Thr Cys Ala Ala Thr Cys
405 410 415
Ala Gly Thr Gly Gly Ala Ala Gly Ala Ala Ala Ala Thr Gly Gly Ala
420 425 430
Gly Gly Ala Ala Gly Thr Ala Thr Gly Thr Gly Ala Thr Thr Cys Ala
435 440 445
Thr Cys Ala Gly Cys Ala Gly Cys Thr Gly Cys Gly Ala Gly Cys Ala
450 455 460
Gly Cys Ala Cys Ala Gly Thr Cys Cys Ala Ala Ala Ala Thr Cys Ala
465 470 475 480
Gly Cys Cys Ala Cys Ala Gly Gly Ala Thr Cys Ala Ala Gly Ala Gly
485 490 495
Cys Ala Cys Cys Cys Gly Thr Gly Gly Cys Cys Gly Thr Ala Thr Cys
500 505 510
Thr Thr Cys Gly Cys Gly Ala Ala Thr Thr Cys Thr Thr Thr Thr Cys
515 520 525
Thr Thr Thr Ala Ala Gly Thr Gly Gly Thr Gly Thr Ala Ala Ala Thr
530 535 540
Ala Ala Ala Gly Ala Thr Thr Cys Ala Thr Thr Cys Ala Ala Gly Ala
545 550 555 560
Thr Gly Ala Ala Ala Thr Gly Thr Gly Thr Cys Cys Thr Cys Thr Gly
565 570 575
Thr Cys Thr Cys Cys Cys Gly Cys Thr Thr Ala Ala Thr Ala Ala Ala
580 585 590
Gly Ala Ala Ala Thr Ala Thr Cys Gly Gly Cys Cys Thr Thr Cys Ala
595 600 605
Ala Ala Ala Gly Thr Thr Cys Gly Cys Cys Ala Thr Cys Ala Ala Ala
610 615 620
Cys Cys Thr Ala Ala Gly Gly Ala Ala Gly Cys Ala Thr Ala Thr Thr
625 630 635 640

[0076]

```
Gly Ala Gly Gly Thr Ala Ala Gly Thr Ala Cys Ala Thr Thr Ala Ala
                645                 650                 655
Gly Thr Ala Thr Thr Thr Thr Gly Thr Thr Thr Thr Ala Cys Thr Gly
            660                 665                 670
Ala Thr Ala Gly Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr
        675                 680                 685
Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr
    690                 695                 700
Thr Gly Gly Gly Thr Gly Thr Gly Cys Ala Thr Gly Thr Thr Thr Thr
705                 710                 715                 720
Gly Ala Cys Gly Thr Thr Gly Ala Thr Gly Gly Cys Gly Cys Gly Cys
                725                 730                 735
Cys Thr Thr Thr Thr Ala Thr Ala Thr Gly Thr Gly Thr Ala Gly Thr
            740                 745                 750
Ala Gly Gly Cys Cys Thr Ala Thr Thr Thr Thr Cys Ala Cys Thr Ala
        755                 760                 765
Ala Thr Gly Cys Ala Thr Gly Cys Gly Ala Thr Thr Gly Ala Cys Ala
    770                 775                 780
Ala Thr Ala Thr Ala Ala Gly Gly Cys Thr Cys Ala Cys Gly Thr Ala
785                 790                 795                 800
Ala Thr Ala Ala Ala Ala Thr Gly Cys Thr Ala Ala Ala Ala Thr Gly
                805                 810                 815
Cys Ala Thr Thr Thr Gly Thr Ala Ala Thr Thr Gly Gly Thr Ala Ala
            820                 825                 830
Cys Gly Thr Thr Ala Gly Gly Thr Cys Cys Ala Cys Gly Gly Gly Ala
        835                 840                 845
Ala Ala Thr Thr Thr Gly Gly Cys Gly Cys Cys Thr Ala Thr Thr Gly
    850                 855                 860
Cys Ala Gly Cys Thr Thr Thr Gly Ala Ala Thr Ala Ala Thr Cys Ala
865                 870                 875                 880
Thr Thr Ala Thr Cys Ala Thr Thr Cys Cys Gly Thr Gly Cys Thr Cys
                885                 890                 895
Thr Cys Ala Thr Thr Gly Thr Gly Thr Thr Thr Gly Ala Ala Thr Thr
            900                 905                 910
Cys Ala Thr Gly Cys Ala Ala Ala Ala Cys Ala Cys Ala Ala Gly Ala
        915                 920                 925
Ala Ala Ala Cys Cys Ala Ala Gly Cys Gly Ala Gly Ala Ala Ala Thr
    930                 935                 940
Thr Thr Thr Thr Thr Thr Cys Cys Ala Ala Ala Cys Ala Thr Gly Thr
945                 950                 955                 960
Thr Gly Thr Ala Thr Thr Gly Thr Cys Ala Ala Ala Ala Cys Gly Gly
                965                 970                 975
Thr Ala Ala Cys Ala Cys Thr Thr Thr Ala Cys Ala Ala Thr Gly Ala
            980                 985                 990
```

[0077]

```
Gly Gly Thr Thr Gly Ala Thr Thr  Ala Gly Thr Thr Cys  Ala Thr Gly
        995                 1000                1005
Thr Ala  Thr Thr Ala Ala Cys  Thr Ala Ala Cys Ala  Thr Thr Ala
    1010                1015                1020
Ala Ala  Thr Ala Ala Cys Cys  Ala Thr Gly Ala Gly  Cys Ala Ala
    1025                1030                1035
Thr Ala  Cys Ala Thr Thr Thr  Gly Thr Thr Ala Cys  Thr Gly Thr
    1040                1045                1050
Ala Thr  Cys Thr Gly Thr Thr  Ala Ala Thr Cys Thr  Thr Thr Gly
    1055                1060                1065
Thr Thr  Ala Ala Cys Gly Thr  Thr Ala Gly Thr Thr  Ala Ala Thr
    1070                1075                1080
Ala Gly  Ala Ala Ala Thr Ala  Cys Ala Gly Ala Thr  Gly Thr Thr
    1085                1090                1095
Cys Ala  Thr Thr Gly Thr Thr  Thr Gly Thr Thr Cys  Ala Thr Gly
    1100                1105                1110
Thr Thr  Ala Gly Thr Thr Cys  Ala Cys Ala Gly Thr  Gly Cys Ala
    1115                1120                1125
Thr Thr  Ala Ala Cys Thr Ala  Ala Thr Gly Thr Thr  Ala Ala Cys
    1130                1135                1140
Ala Ala  Gly Ala Thr Ala Thr  Ala Ala Ala Gly Thr  Ala Thr Thr
    1145                1150                1155
Ala Gly  Thr Ala Ala Ala Thr  Gly Thr Thr Gly Ala  Ala Ala Thr
    1160                1165                1170
Thr Ala  Ala Cys Ala Thr Gly  Thr Ala Thr Ala Cys  Gly Thr Gly
    1175                1180                1185
Cys Ala  Gly Thr Thr Cys Ala  Thr Thr Ala Thr Thr  Ala Gly Thr
    1190                1195                1200
Thr Cys  Ala Thr Gly Thr Thr  Ala Ala Cys Thr Ala  Ala Thr Gly
    1205                1210                1215
Thr Ala  Gly Thr Thr Ala Ala  Cys Thr Ala Ala Cys  Gly Ala Ala
    1220                1225                1230
Cys Cys  Thr Thr Ala Thr Thr  Gly Thr Ala Ala Ala  Ala Gly Thr
    1235                1240                1245
Gly Thr  Thr Ala Cys Cys Ala  Thr Cys Ala Ala Ala  Ala Cys Thr
    1250                1255                1260
Ala Ala  Thr Gly Thr Ala Ala  Thr Gly Ala Ala Ala  Thr Cys Ala
    1265                1270                1275
Ala Thr  Thr Cys Ala Cys Cys  Cys Thr Gly Thr Cys  Ala Thr Gly
    1280                1285                1290
Thr Cys  Ala Gly Cys Cys Thr  Thr Ala Cys Ala Gly  Thr Cys Cys
    1295                1300                1305
Thr Gly  Thr Gly Thr Thr Thr  Thr Thr Gly Thr Cys  Ala Ala Thr
    1310                1315                1320
```

[0078]

```
Ala Thr  Ala Ala Thr Cys Ala  Gly Ala Ala Ala Thr  Ala Ala Ala
    1325                1330                1335
Ala Thr  Thr Ala Ala Thr Gly  Thr Thr Thr Gly Ala  Thr Thr Gly
    1340                1345                1350
Thr Cys  Ala Cys Thr Ala Ala  Ala Thr Gly Cys Thr  Ala Cys Thr
    1355                1360                1365
Gly Thr  Ala Thr Thr Thr Cys  Thr Ala Ala Ala Ala  Thr Cys Ala
    1370                1375                1380
Ala Cys  Ala Ala Gly Thr Ala  Thr Thr Thr Ala Ala  Cys Ala Thr
    1385                1390                1395
Thr Ala  Thr Ala Ala Ala Gly  Thr Gly Thr Gly Cys  Ala Ala Thr
    1400                1405                1410
Thr Gly  Gly Cys Thr Gly Cys  Ala Ala Ala Thr Gly  Thr Cys Ala
    1415                1420                1425
Gly Thr  Thr Thr Thr Ala Thr  Thr Ala Ala Ala Gly  Gly Gly Thr
    1430                1435                1440
Thr Ala  Gly Thr Thr Cys Ala  Cys Cys Cys Ala Ala  Ala Ala Ala
    1445                1450                1455
Thr Gly  Ala Ala Ala Ala Thr  Ala Ala Thr Gly Thr  Cys Ala Thr
    1460                1465                1470
Thr Ala  Ala Thr Gly Ala Cys  Thr Cys Gly Cys Cys  Cys Thr Cys
    1475                1480                1485
Ala Thr  Gly Thr Cys Gly Thr  Thr Cys Cys Ala Ala  Gly Cys Cys
    1490                1495                1500
Cys Gly  Thr Ala Ala Gly Ala  Cys Cys Thr Cys Cys  Gly Thr Thr
    1505                1510                1515
Cys Ala  Thr Cys Thr Thr Cys  Ala Gly Ala Ala Cys  Ala Cys Ala
    1520                1525                1530
Gly Thr  Thr Thr Ala Ala Gly  Ala Thr Ala Thr Thr  Thr Thr Ala
    1535                1540                1545
Gly Ala  Thr Thr Thr Ala Gly  Thr Cys Cys Gly Ala  Gly Ala Gly
    1550                1555                1560
Cys Thr  Thr Thr Cys Thr Gly  Thr Gly Cys Cys Thr  Cys Cys Ala
    1565                1570                1575
Thr Thr  Gly Ala Gly Ala Ala  Thr Gly Thr Ala Thr  Gly Thr Ala
    1580                1585                1590
Cys Gly  Gly Thr Ala Thr Ala  Cys Thr Gly Thr Cys  Cys Ala Thr
    1595                1600                1605
Gly Thr  Cys Cys Ala Gly Ala  Ala Ala Gly Gly Thr  Ala Ala Thr
    1610                1615                1620
Ala Ala  Ala Ala Ala Cys Ala  Thr Cys Ala Ala Ala  Gly Thr Ala
    1625                1630                1635
Gly Thr  Cys Cys Ala Thr Gly  Thr Gly Ala Cys Ala  Thr Cys Ala
    1640                1645                1650
```

[0079]

```
Gly Thr  Gly Gly Gly Thr Thr  Ala Gly Thr Thr Ala  Gly Ala Ala
    1655                 1660                 1665
Thr Thr  Thr Thr Thr Thr Gly  Ala Ala Gly Cys Ala  Thr Cys Gly
    1670                 1675                 1680
Ala Ala  Thr Ala Cys Ala Thr  Thr Thr Thr Gly Gly  Thr Cys Cys
    1685                 1690                 1695
Ala Ala  Ala Ala Ala Thr Ala  Ala Cys Ala Ala Ala  Ala Cys Cys
    1700                 1705                 1710
Thr Ala  Cys Gly Ala Cys Thr  Thr Thr Ala Thr Thr  Cys Gly Gly
    1715                 1720                 1725
Cys Ala  Thr Thr Gly Thr Ala  Thr Thr Cys Thr Cys  Thr Thr Cys
    1730                 1735                 1740
Cys Gly  Gly Gly Thr Cys Thr  Gly Thr Thr Gly Thr  Cys Ala Ala
    1745                 1750                 1755
Thr Cys  Cys Gly Cys Gly Thr  Thr Cys Ala Cys Gly  Ala Cys Thr
    1760                 1765                 1770
Thr Cys  Gly Cys Ala Gly Thr  Gly Ala Cys Gly Cys  Thr Ala Cys
    1775                 1780                 1785
Ala Ala  Thr Gly Cys Thr Gly  Ala Ala Thr Ala Ala  Ala Gly Thr
    1790                 1795                 1800
Cys Gly  Thr Ala Gly Gly Thr  Thr Thr Thr Gly Thr  Thr Ala Thr
    1805                 1810                 1815
Thr Thr  Thr Thr Gly Gly Ala  Cys Cys Ala Ala Ala  Ala Thr Gly
    1820                 1825                 1830
Thr Ala  Thr Thr Thr Thr Cys  Gly Ala Thr Gly Cys  Thr Thr Cys
    1835                 1840                 1845
Ala Ala  Ala Thr Ala Ala Thr  Thr Cys Thr Ala Cys  Cys Thr Ala
    1850                 1855                 1860
Ala Cys  Cys Cys Ala Cys Thr  Gly Ala Thr Gly Thr  Cys Ala Cys
    1865                 1870                 1875
Ala Thr  Gly Gly Ala Cys Thr  Ala Cys Thr Thr Thr  Gly Ala Thr
    1880                 1885                 1890
Gly Thr  Thr Thr Thr Thr Ala  Thr Thr Ala Cys Cys  Thr Thr Thr
    1895                 1900                 1905
Cys Thr  Gly Gly Ala Cys Ala  Thr Gly Gly Ala Cys  Ala Gly Thr
    1910                 1915                 1920
Ala Thr  Ala Cys Cys Gly Thr  Ala Cys Ala Thr Ala  Cys Ala Thr
    1925                 1930                 1935
Thr Thr  Thr Cys Ala Gly Thr  Gly Gly Ala Gly Gly  Gly Ala Cys
    1940                 1945                 1950
Ala Gly  Ala Ala Ala Gly Cys  Thr Cys Thr Cys Gly  Gly Ala Cys
    1955                 1960                 1965
Thr Ala  Ala Ala Thr Cys Thr  Ala Ala Ala Ala Thr  Ala Thr Cys
    1970                 1975                 1980
```

[0080]

Thr Thr Ala Ala Ala Cys Thr Gly Thr Gly Thr Thr Cys Cys Gly
1985 1990 1995
Ala Ala Gly Ala Thr Gly Ala Ala Cys Gly Gly Ala Gly Gly Thr
2000 2005 2010
Gly Thr Thr Ala Cys Gly Gly Gly Cys Thr Thr Gly Gly Ala Ala
2015 2020 2025
Cys Gly Ala Cys Ala Thr Gly Ala Gly Gly Gly Thr Gly Ala Gly
2030 2035 2040
Thr Cys Ala Thr Thr Ala Ala Thr Gly Ala Cys Ala Thr Cys Thr
2045 2050 2055
Thr Thr Thr Cys Ala Thr Thr Thr Thr Thr Gly Gly Gly Thr Gly
2060 2065 2070
Ala Ala Cys Thr Ala Ala Cys Cys Cys Thr Thr Thr Ala Ala Thr
2075 2080 2085
Gly Cys Thr Gly Thr Ala Ala Thr Cys Ala Gly Ala Gly Ala Gly
2090 2095 2100
Thr Gly Thr Ala Thr Gly Thr Gly Thr Ala Ala Thr Thr Gly Thr
2105 2110 2115
Thr Ala Cys Ala Thr Thr Thr Ala Thr Thr Gly Cys Ala Thr Ala
2120 2125 2130
Cys Ala Ala Thr Ala Thr Ala Ala Ala Thr Ala Thr Thr Thr Ala
2135 2140 2145
Thr Thr Thr Gly Thr Thr Gly Thr Thr Thr Thr Thr Ala Cys Ala
2150 2155 2160
Gly Ala Gly Ala Ala Thr Gly Cys Ala Cys Cys Cys Ala Ala Ala
2165 2170 2175
Thr Thr Ala Cys Cys Thr Cys Ala Ala Ala Ala Ala Cys Thr Ala
2180 2185 2190
Cys Thr Cys Thr Ala Ala Ala Thr Thr Gly Ala Cys Ala Gly Cys
2195 2200 2205
Ala Cys Ala Gly Ala Ala Gly Ala Gly Ala Ala Ala Gly Ala Thr
2210 2215 2220
Cys Gly Gly Gly Ala Cys Cys Thr Cys Cys Ala Cys Cys Cys Ala
2225 2230 2235
Thr Gly Cys Thr Thr Cys Cys Ala Gly Cys Ala Gly Thr Ala Ala
2240 2245 2250
Gly Cys Ala Ala Cys Thr Gly Ala Ala Ala Gly Thr Thr Gly Ala
2255 2260 2265
Cys Thr Cys Ala Gly Thr Thr Thr Thr Cys Cys Cys Ala Gly Thr
2270 2275 2280
Cys Ala Ala Ala Cys Ala Thr Gly Thr Gly Thr Cys Thr Cys Cys
2285 2290 2295
Ala Gly Thr Cys Ala Cys Thr Gly Thr Gly Ala Ala Cys Ala Ala
2300 2305 2310

[0081]

Ala Gly Cys Thr Ala Thr Ala Thr Thr Ala Ala Gly Gly Thr Ala
2315 2320 2325
Cys Ala Thr Cys Ala Thr Thr Cys Ala Ala Gly Gly Ala Cys Thr
2330 2335 2340
Thr Cys Ala Thr Cys Cys Thr Thr Thr Cys Ala Gly Cys Ala Cys
2345 2350 2355
Thr Gly Thr Thr Gly Ala Thr Cys Thr Gly Cys Cys Ala Thr Cys
2360 2365 2370
Ala Thr Thr Thr Ala Ala Ala Gly Ala Gly Cys Thr Gly Ala Thr
2375 2380 2385
Thr Ala Gly Thr Ala Cys Ala Cys Thr Gly Cys Ala Gly Cys Cys
2390 2395 2400
Thr Gly Gly Cys Ala Thr Thr Thr Cys Thr Gly Thr Cys Ala Thr
2405 2410 2415
Thr Ala Cys Ala Ala Gly Gly Cys Cys Thr Ala Cys Thr Thr Thr
2420 2425 2430
Ala Cys Gly Cys Thr Cys Cys Ala Ala Gly Ala Thr Ala Gly Cys
2435 2440 2445
Thr Gly Ala Ala Gly Cys Thr Gly Cys Thr Cys Thr Gly Ala Thr
2450 2455 2460
Cys Ala Thr Gly Ala Ala Ala Cys Ala Gly Ala Ala Ala Gly Thr
2465 2470 2475
Gly Ala Cys Thr Gly Cys Thr Gly Cys Cys Ala Thr Gly Ala Gly
2480 2485 2490
Thr Gly Ala Ala Gly Thr Thr Gly Ala Ala Thr Gly Gly Ala Thr
2495 2500 2505
Thr Gly Cys Ala Ala Cys Cys Ala Cys Ala Ala Cys Gly Gly Ala
2510 2515 2520
Thr Thr Gly Thr Thr Gly Gly Ala Cys Thr Gly Cys Ala Cys Gly
2525 2530 2535
Thr Ala Gly Ala Ala Ala Gly Thr Cys Ala Thr Thr Cys Ala Thr
2540 2545 2550
Thr Gly Gly Thr Gly Thr Ala Ala Cys Thr Gly Cys Thr Cys Ala
2555 2560 2565
Cys Thr Gly Gly Ala Thr Cys Ala Ala Cys Cys Cys Thr Gly Gly
2570 2575 2580
Ala Ala Gly Thr Cys Thr Thr Gly Ala Ala Ala Gly Ala Cys Ala
2585 2590 2595
Thr Thr Cys Cys Gly Cys Thr Gly Cys Ala Cys Thr Thr Gly Cys
2600 2605 2610
Cys Thr Gly Cys Ala Ala Ala Ala Gly Ala Thr Thr Ala Ala Thr
2615 2620 2625
Gly Gly Gly Cys Thr Cys Thr Cys Ala Thr Ala Cys Thr Thr Thr
2630 2635 2640

[0082]

```
Thr Gly  Ala Gly Gly Thr Ala  Cys Thr Gly Gly Cys  Cys Ala Gly
    2645                 2650                 2655
Thr Gly  Cys Cys Ala Thr Gly  Ala Ala Thr Gly Ala  Thr Ala Thr
    2660                 2665                 2670
Cys Cys  Ala Cys Thr Cys Ala  Gly Ala Gly Thr Ala  Thr Gly Ala
    2675                 2680                 2685
Ala Ala  Thr Ala Cys Gly Thr  Gly Ala Cys Ala Ala  Gly Gly Thr
    2690                 2695                 2700
Thr Gly  Thr Thr Thr Gly Cys  Ala Cys Ala Ala Cys  Cys Ala Cys
    2705                 2710                 2715
Ala Gly  Ala Cys Ala Gly Thr  Gly Gly Thr Thr Cys  Cys Ala Ala
    2720                 2725                 2730
Cys Thr  Thr Thr Ala Thr Gly  Ala Ala Gly Gly Cys  Thr Thr Thr
    2735                 2740                 2745
Cys Ala  Gly Ala Gly Thr Thr  Thr Thr Thr Gly Gly  Thr Gly Thr
    2750                 2755                 2760
Gly Gly  Ala Ala Ala Ala Cys  Ala Ala Thr Gly Ala  Thr Ala Thr
    2765                 2770                 2775
Cys Gly  Ala Gly Ala Cys Thr  Gly Ala Gly Gly Cys  Ala Ala Gly
    2780                 2785                 2790
Ala Ala  Gly Gly Thr Gly Thr  Gly Ala Ala Ala Gly  Thr Gly Ala
    2795                 2800                 2805
Thr Gly  Ala Cys Ala Cys Thr  Gly Ala Thr Thr Cys  Thr Gly Ala
    2810                 2815                 2820
Ala Gly  Gly Cys Thr Gly Thr  Gly Gly Thr Gly Ala  Gly Gly Gly
    2825                 2830                 2835
Ala Ala  Gly Thr Gly Ala Thr  Gly Gly Thr Gly Thr  Gly Gly Ala
    2840                 2845                 2850
Ala Thr  Thr Cys Cys Ala Ala  Gly Ala Thr Gly Cys  Cys Thr Cys
    2855                 2860                 2865
Ala Cys  Gly Ala Gly Thr Cys  Cys Thr Gly Gly Ala  Cys Cys Ala
    2870                 2875                 2880
Ala Gly  Ala Cys Gly Ala Thr  Gly Gly Cys Thr Thr  Cys Gly Ala
    2885                 2890                 2895
Ala Thr  Thr Cys Cys Ala Gly  Cys Thr Ala Cys Cys  Ala Ala Ala
    2900                 2905                 2910
Ala Cys  Ala Thr Cys Ala Ala  Ala Ala Gly Thr Gly  Thr Gly Cys
    2915                 2920                 2925
Cys Thr  Gly Thr Cys Ala Cys  Thr Thr Ala Cys Thr  Thr Ala Ala
    2930                 2935                 2940
Cys Cys  Thr Ala Gly Thr Cys  Thr Cys Ala Ala Gly  Cys Gly Thr
    2945                 2950                 2955
Thr Gly  Ala Thr Gly Cys Cys  Cys Ala Ala Ala Ala  Ala Gly Cys
    2960                 2965                 2970
```

[0083]

Thr Cys Thr Cys Thr Cys Ala Ala Ala Thr Gly Ala Ala Cys Ala
2975 2980 2985
Cys Thr Ala Cys Ala Ala Gly Ala Ala Ala Cys Thr Cys Thr Ala
2990 2995 3000
Cys Ala Gly Ala Thr Cys Thr Gly Thr Cys Thr Thr Thr Gly Gly
3005 3010 3015
Cys Ala Ala Ala Thr Gly Cys Cys Ala Ala Gly Cys Thr Thr Thr
3020 3025 3030
Ala Thr Gly Gly Ala Ala Thr Ala Ala Ala Ala Gly Cys Ala Gly
3035 3040 3045
Cys Cys Gly Ala Thr Cys Gly Gly Cys Thr Cys Thr Ala Gly Cys
3050 3055 3060
Ala Gly Cys Thr Gly Ala Ala Gly Cys Thr Gly Thr Thr Gly Ala
3065 3070 3075
Ala Thr Cys Ala Gly Ala Ala Ala Gly Cys Cys Gly Gly Cys Thr
3080 3085 3090
Thr Cys Ala Gly Cys Thr Thr Thr Thr Ala Ala Gly Gly Cys Cys
3095 3100 3105
Ala Ala Ala Cys Cys Ala Ala Ala Cys Gly Cys Gly Gly Thr Gly
3110 3115 3120
Gly Ala Ala Thr Thr Cys Ala Ala Cys Thr Thr Thr Thr Ala Thr
3125 3130 3135
Gly Gly Cys Thr Gly Thr Thr Gly Ala Cys Ala Gly Ala Ala Thr
3140 3145 3150
Thr Cys Thr Thr Cys Ala Ala Ala Thr Thr Thr Gly Cys Ala Ala
3155 3160 3165
Ala Gly Ala Ala Gly Cys Ala Gly Gly Ala Gly Ala Ala Gly Gly
3170 3175 3180
Cys Gly Cys Ala Cys Thr Thr Cys Gly Gly Ala Ala Thr Ala Thr
3185 3190 3195
Ala Thr Gly Cys Ala Cys Cys Thr Cys Thr Cys Thr Thr Gly Ala
3200 3205 3210
Gly Gly Thr Thr Cys Cys Ala Ala Thr Gly Thr Ala Ala Gly Thr
3215 3220 3225
Gly Thr Thr Thr Thr Thr Cys Cys Cys Cys Thr Cys Thr Ala Thr
3230 3235 3240
Cys Gly Ala Thr Gly Thr Ala Ala Ala Cys Ala Ala Ala Thr Gly
3245 3250 3255
Thr Gly Gly Gly Thr Thr Gly Thr Thr Thr Thr Thr Gly Thr Thr
3260 3265 3270
Thr Ala Ala Thr Ala Cys Thr Cys Thr Thr Thr Gly Ala Thr Thr
3275 3280 3285
Ala Thr Gly Cys Thr Gly Ala Thr Thr Thr Cys Thr Cys Cys Thr
3290 3295 3300

[0084]

```
Gly Thr  Ala Gly Gly Thr Thr  Thr Ala Ala Thr Cys  Cys Ala Gly
    3305                 3310                 3315
Cys Ala  Gly Ala Ala Ala Thr  Gly Cys Thr Gly Thr  Thr Cys Thr
    3320                 3325                 3330
Thr Gly  Ala Cys Ala Gly Ala  Gly Thr Gly Gly Gly  Cys Cys Ala
    3335                 3340                 3345
Ala Cys  Ala Cys Ala Ala Thr  Gly Cys Gly Thr Cys  Cys Ala Gly
    3350                 3355                 3360
Thr Thr  Gly Cys Ala Ala Ala  Ala Gly Thr Ala Cys  Thr Cys Gly
    3365                 3370                 3375
Ala Cys  Ala Thr Cys Thr Thr  Gly Cys Ala Ala Gly  Cys Gly Gly
    3380                 3385                 3390
Ala Ala  Ala Cys Gly Ala Ala  Thr Ala Cys Ala Cys  Ala Gly Cys
    3395                 3400                 3405
Thr Gly  Gly Gly Gly Thr Gly  Gly Cys Thr Gly Cys  Thr Gly Cys
    3410                 3415                 3420
Cys Thr  Ala Gly Thr Gly Thr  Cys Cys Ala Thr Cys  Ala Gly Thr
    3425                 3430                 3435
Thr Ala  Ala Gly Cys Thr Thr  Gly Ala Ala Ala Cys  Thr Thr Cys
    3440                 3445                 3450
Ala Gly  Cys Gly Ala Cys Thr  Cys Cys Ala Cys Cys  Ala Thr Thr
    3455                 3460                 3465
Cys Thr  Cys Thr Cys Ala Gly  Gly Thr Ala Cys Thr  Gly Thr Gly
    3470                 3475                 3480
Ala Cys  Cys Cys Ala Cys Thr  Thr Gly Thr Gly Gly  Ala Thr Gly
    3485                 3490                 3495
Cys Cys  Cys Thr Ala Cys Ala  Ala Cys Ala Ala Gly  Gly Ala Ala
    3500                 3505                 3510
Thr Cys  Cys Ala Ala Ala Cys  Ala Cys Gly Ala Thr  Thr Cys Ala
    3515                 3520                 3525
Ala Gly  Cys Ala Thr Ala Thr  Gly Thr Thr Thr Gly  Ala Ala Gly
    3530                 3535                 3540
Ala Thr  Cys Cys Thr Gly Ala  Gly Ala Thr Cys Ala  Thr Ala Gly
    3545                 3550                 3555
Cys Ala  Gly Cys Thr Gly Cys  Cys Ala Thr Cys Cys  Thr Thr Cys
    3560                 3565                 3570
Thr Cys  Cys Cys Thr Ala Ala  Ala Thr Thr Thr Cys  Gly Gly Ala
    3575                 3580                 3585
Cys Cys  Thr Cys Thr Thr Gly  Gly Ala Cys Ala Ala  Ala Thr Gly
    3590                 3595                 3600
Ala Thr  Gly Ala Ala Ala Cys  Cys Ala Thr Cys Ala  Thr Ala Ala
    3605                 3610                 3615
Ala Ala  Cys Gly Ala Gly Gly  Thr Ala Ala Ala Thr  Gly Ala Ala
    3620                 3625                 3630
```

[0085]

Thr Gly Cys Ala Ala Gly Cys Ala Ala Cys Ala Thr Ala Cys Ala
3635 3640 3645
Cys Thr Thr Gly Ala Cys Gly Ala Ala Thr Thr Cys Thr Ala Ala
3650 3655 3660
Thr Cys Thr Gly Gly Gly Cys Ala Ala Cys Cys Thr Thr Thr Gly
3665 3670 3675
Ala Gly Cys Cys Ala Thr Ala Cys Cys Ala Ala Ala Ala Thr Thr
3680 3685 3690
Ala Thr Thr Cys Thr Thr Thr Thr Ala Thr Thr Thr Ala Thr Thr
3695 3700 3705
Thr Ala Thr Thr Thr Thr Thr Gly Cys Ala Cys Thr Thr Thr Thr
3710 3715 3720
Thr Ala Gly Gly Ala Ala Thr Gly Thr Thr Ala Thr Ala Thr Cys
3725 3730 3735
Cys Cys Ala Thr Cys Thr Thr Thr Gly Gly Cys Thr Gly Thr Gly
3740 3745 3750
Ala Thr Cys Thr Cys Ala Ala Thr Ala Thr Gly Ala Ala Thr Ala
3755 3760 3765
Thr Thr Gly Ala Thr Gly Thr Ala Ala Ala Gly Thr Ala Thr Thr
3770 3775 3780
Cys Thr Thr Gly Cys Ala Gly Cys Ala Gly Gly Thr Thr Gly Thr
3785 3790 3795
Ala Gly Thr Thr Ala Thr Cys Cys Cys Thr Cys Ala Gly Thr Gly
3800 3805 3810
Thr Thr Thr Cys Thr Thr Gly Ala Ala Ala Cys Cys Ala Ala Ala
3815 3820 3825
Cys Thr Cys Ala Thr Ala Thr Gly Thr Ala Thr Cys Ala Thr Ala
3830 3835 3840
Thr Gly Thr Gly Gly Thr Thr Thr Gly Gly Ala Ala Ala Thr Gly
3845 3850 3855
Cys Ala Gly Thr Thr Ala Gly Ala Thr Thr Thr Thr Ala Thr Gly
3860 3865 3870
Cys Thr Ala Ala Ala Ala Thr Ala Ala Gly Gly Gly Ala Thr Thr
3875 3880 3885
Thr Gly Cys Ala Thr Gly Ala Thr Thr Thr Thr Ala Gly Ala Thr
3890 3895 3900
Gly Thr Ala Gly Ala Thr Gly Ala Cys Thr Gly Cys Ala Cys Gly
3905 3910 3915
Thr Ala Ala Ala Thr Gly Thr Ala Gly Thr Thr Ala Ala Thr Gly
3920 3925 3930
Ala Cys Ala Ala Ala Ala Thr Cys Cys Ala Thr Ala Ala Ala Ala
3935 3940 3945
Thr Thr Thr Gly Thr Thr Cys Cys Cys Ala Gly Thr Cys Ala Gly
3950 3955 3960

[0086]

```
Ala Ala  Gly Cys Cys Cys Cys  Thr Cys Ala Ala Cys  Cys Ala Ala
    3965                 3970                 3975
Ala Cys  Thr Thr Thr Thr Cys  Thr Thr Thr Gly Thr  Gly Thr Cys
    3980                 3985                 3990
Thr Gly  Cys Thr Cys Ala Cys  Thr Gly Thr Gly Cys  Thr Thr Gly
    3995                 4000                 4005
Thr Ala  Gly Gly Cys Ala Thr  Gly Gly Ala Cys Thr  Ala Cys Ala
    4010                 4015                 4020
Thr Cys  Ala Gly Ala Gly Thr  Gly Cys Ala Thr Cys  Thr Gly Gly
    4025                 4030                 4035
Ala Gly  Cys Cys Thr Thr Thr  Gly Gly Ala Cys Cys  Ala Cys Ala
    4040                 4045                 4050
Ala Gly  Ala Ala Gly Gly Ala  Ala Thr Thr Gly Gly  Cys Cys Ala
    4055                 4060                 4065
Ala Cys  Ala Gly Thr Thr Cys  Ala Thr Cys Thr Gly  Ala Thr Gly
    4070                 4075                 4080
Ala Thr  Gly Ala Ala Gly Ala  Thr Thr Thr Thr Thr  Thr Cys Gly
    4085                 4090                 4095
Cys Thr  Thr Cys Thr Thr Thr  Gly Ala Ala Ala Cys  Cys Gly Ala
    4100                 4105                 4110
Cys Ala  Ala Cys Ala Cys Ala  Thr Gly Ala Ala Gly  Cys Cys Ala
    4115                 4120                 4125
Gly Cys  Ala Ala Ala Gly Ala  Gly Thr Thr Gly Gly  Ala Thr Gly
    4130                 4135                 4140
Gly Ala  Thr Ala Thr Cys Thr  Gly Gly Cys Cys Thr  Gly Thr Gly
    4145                 4150                 4155
Thr Thr  Thr Cys Ala Gly Ala  Cys Ala Cys Cys Ala  Gly Gly Gly
    4160                 4165                 4170
Ala Gly  Thr Cys Thr Cys Thr  Gly Cys Thr Cys Ala  Cys Gly Thr
    4175                 4180                 4185
Thr Thr  Cys Cys Thr Gly Cys  Thr Ala Thr Thr Thr  Gly Cys Ala
    4190                 4195                 4200
Gly Cys  Cys Thr Cys Thr Cys  Thr Ala Thr Cys Ala  Ala Gly Ala
    4205                 4210                 4215
Cys Thr  Ala Ala Thr Ala Cys  Ala Cys Cys Thr Cys  Thr Thr Cys
    4220                 4225                 4230
Cys Cys  Gly Cys Ala Thr Cys  Gly Gly Cys Thr Gly  Cys Cys Thr
    4235                 4240                 4245
Gly Thr  Gly Ala Gly Ala Gly  Gly Cys Thr Thr Thr  Thr Cys Ala
    4250                 4255                 4260
Gly Cys  Ala Cys Thr Gly Cys  Ala Gly Gly Ala Thr  Thr Gly Cys
    4265                 4270                 4275
Thr Thr  Thr Thr Cys Ala Gly  Cys Cys Cys Cys Ala  Ala Ala Ala
    4280                 4285                 4290
```

[0087]

Gly Ala Gly Cys Thr Ala Gly Gly Cys Thr Thr Gly Ala Cys Ala
4295 4300 4305
Cys Thr Ala Ala Cys Ala Ala Thr Thr Thr Thr Gly Ala Gly Ala
4310 4315 4320
Ala Thr Cys Ala Gly Cys Thr Thr Cys Thr Ala Cys Thr Gly Ala
4325 4330 4335
Ala Gly Thr Thr Ala Ala Ala Thr Cys Thr Gly Ala Gly Gly Thr
4340 4345 4350
Thr Thr Thr Ala Cys Ala Ala Cys Thr Thr Thr Gly Ala Gly Thr
4355 4360 4365
Ala Gly Cys Gly Thr Gly Thr Ala Cys Thr Gly Gly Cys Ala Thr
4370 4375 4380
Thr Ala Gly Ala Thr Thr Gly Thr Cys Thr Gly Thr Cys Thr Thr
4385 4390 4395
Ala Thr Ala Gly Thr Thr Thr Gly Ala Thr Ala Ala Thr Thr Ala
4400 4405 4410
Ala Ala Thr Ala Cys Ala Ala Ala Cys Ala Gly Thr Thr Cys Thr
4415 4420 4425
Ala Ala Ala Gly Cys Ala Gly Gly Ala Thr Ala Ala Ala Ala Cys
4430 4435 4440
Cys Thr Thr Gly Thr Ala Thr Gly Cys Ala Thr Thr Thr Cys Ala
4445 4450 4455
Thr Thr Thr Ala Ala Thr Gly Thr Thr Thr Thr Thr Thr Gly Ala
4460 4465 4470
Gly Ala Thr Thr Ala Ala Ala Ala Gly Cys Thr Thr Ala Ala Ala
4475 4480 4485
Cys Ala Ala Gly Ala Ala Thr Cys Thr Cys Thr Ala Gly Thr Thr
4490 4495 4500
Thr Thr Cys Thr Thr Thr Cys Thr Thr Gly Cys Thr Thr Thr Thr
4505 4510 4515
Ala Cys Thr Thr Thr Thr Ala Cys Thr Thr Cys Cys Thr Thr Ala
4520 4525 4530
Ala Thr Ala Cys Thr Cys Ala Ala Gly Thr Ala Cys Ala Ala Thr
4535 4540 4545
Thr Thr Thr Ala Ala Thr Gly Gly Ala Gly Thr Ala Cys Thr Thr
4550 4555 4560
Thr Thr Thr Thr Ala Cys Thr Thr Thr Thr Ala Cys Thr Cys Ala
4565 4570 4575
Ala Gly Thr Ala Ala Gly Ala Thr Thr Cys Thr Ala Gly Cys Cys
4580 4585 4590
Ala Gly Ala Thr Ala Cys Thr Thr Thr Thr Ala Cys Thr Thr Thr
4595 4600 4605
Thr Ala Ala Thr Thr Gly Ala Gly Thr Ala Ala Ala Ala Thr Thr
4610 4615 4620

Thr Thr Cys Cys Cys Thr Ala Ala Gly Thr Ala Cys Thr Thr Gly
4625 4630 4635
Thr Ala Cys Thr Thr Thr Cys Ala Cys Thr Thr Gly Ala Gly Thr
4640 4645 4650
Ala Ala Ala Ala Thr Thr Thr Thr Thr Gly Ala Gly Thr Ala Cys
4655 4660 4665
Thr Thr Thr Thr Thr Ala Cys Ala Cys Cys Thr Cys Thr Gly
4670 4675 4680

<210> 70
<211> 649
<212> PRT
<213> 人工序列
<220>
<223> Tol2转座酶- 40-14490

[0088]

<400> 70
Met Glu Glu Val Cys Asp Ser Ser Ala Ala Ala Ser Ser Thr Val Gln
1 5 10 15
Asn Gln Pro Gln Asp Gln Glu His Pro Trp Pro Tyr Leu Arg Glu Phe
20 25 30
Phe Ser Leu Ser Gly Val Asn Lys Asp Ser Phe Lys Met Lys Cys Val
35 40 45
Leu Cys Leu Pro Leu Asn Lys Glu Ile Ser Ala Phe Lys Ser Ser Pro
50 55 60
Ser Asn Leu Arg Lys His Ile Glu Arg Met His Pro Asn Tyr Leu Lys
65 70 75 80
Asn Tyr Ser Lys Leu Thr Ala Gln Lys Arg Lys Ile Gly Thr Ser Thr
85 90 95
His Ala Ser Ser Ser Lys Gln Leu Lys Val Asp Ser Val Phe Pro Val
100 105 110
Lys His Val Ser Pro Val Thr Val Asn Lys Ala Ile Leu Arg Tyr Ile
115 120 125
Ile Gln Gly Leu His Pro Phe Ser Thr Val Asp Leu Pro Ser Phe Lys
130 135 140
Glu Leu Ile Ser Thr Leu Gln Pro Gly Ile Ser Val Ile Thr Arg Pro
145 150 155 160
Thr Leu Arg Ser Lys Ile Ala Glu Ala Ala Leu Ile Met Lys Gln Lys
165 170 175
Val Thr Ala Ala Met Ser Glu Val Glu Trp Ile Ala Thr Thr Thr Asp
180 185 190
Cys Trp Thr Ala Arg Arg Lys Ser Phe Ile Gly Val Thr Ala His Trp
195 200 205
Ile Asn Pro Gly Ser Leu Glu Arg His Ser Ala Ala Leu Ala Cys Lys
210 215 220
Arg Leu Met Gly Ser His Thr Phe Glu Val Leu Ala Ser Ala Met Asn

225 230 235 240
Asp Ile His Ser Glu Tyr Glu Ile Arg Asp Lys Val Val Cys Thr Thr
245 250 255
Thr Asp Ser Gly Ser Asn Phe Met Lys Ala Phe Arg Val Phe Gly Val
260 265 270
Glu Asn Asn Asp Ile Glu Thr Glu Ala Arg Arg Cys Glu Ser Asp Asp
275 280 285
Thr Asp Ser Glu Gly Cys Gly Glu Gly Ser Asp Gly Val Glu Phe Gln
290 295 300
Asp Ala Ser Arg Val Leu Asp Gln Asp Asp Gly Phe Glu Phe Gln Leu
305 310 315 320
Pro Lys His Gln Lys Cys Ala Cys His Leu Leu Asn Leu Val Ser Ser
325 330 335
Val Asp Ala Gln Lys Ala Leu Ser Asn Glu His Tyr Lys Lys Leu Tyr
340 345 350
Arg Ser Val Phe Gly Lys Cys Gln Ala Leu Trp Asn Lys Ser Ser Arg
355 360 365
Ser Ala Leu Ala Ala Glu Ala Val Glu Ser Glu Ser Arg Leu Gln Leu
370 375 380
Leu Arg Pro Asn Gln Thr Arg Trp Asn Ser Thr Phe Met Ala Val Asp
385 390 395 400
[0089] Arg Ile Leu Gln Ile Cys Lys Glu Ala Gly Glu Gly Ala Leu Arg Asn
405 410 415
Ile Cys Thr Ser Leu Glu Val Pro Met Phe Asn Pro Ala Glu Met Leu
420 425 430
Phe Leu Thr Glu Trp Ala Asn Thr Met Arg Pro Val Ala Lys Val Leu
435 440 445
Asp Ile Leu Gln Ala Glu Thr Asn Thr Gln Leu Gly Trp Leu Leu Pro
450 455 460
Ser Val His Gln Leu Ser Leu Lys Leu Gln Arg Leu His His Ser Leu
465 470 475 480
Arg Tyr Cys Asp Pro Leu Val Asp Ala Leu Gln Gln Gly Ile Gln Thr
485 490 495
Arg Phe Lys His Met Phe Glu Asp Pro Glu Ile Ile Ala Ala Ala Ile
500 505 510
Leu Leu Pro Lys Phe Arg Thr Ser Trp Thr Asn Asp Glu Thr Ile Ile
515 520 525
Lys Arg Gly Met Asp Tyr Ile Arg Val His Leu Glu Pro Leu Asp His
530 535 540
Lys Lys Glu Leu Ala Asn Ser Ser Ser Asp Asp Glu Asp Phe Phe Ala
545 550 555 560
Ser Leu Lys Pro Thr Thr His Glu Ala Ser Lys Glu Leu Asp Gly Tyr
565 570 575
Leu Ala Cys Val Ser Asp Thr Arg Glu Ser Leu Leu Thr Phe Pro Ala

580 585 590

Ile Cys Ser Leu Ser Ile Lys Thr Asn Thr Pro Leu Pro Ala Ser Ala
595 600 605

Ala Cys Glu Arg Leu Phe Ser Thr Ala Gly Leu Leu Phe Ser Pro Lys
610 615 620

Arg Ala Arg Leu Asp Thr Asn Asn Phe Glu Asn Gln Leu Leu Leu Lys
625 630 635 640

Leu Asn Leu Arg Phe Tyr Asn Phe Glu
645

<210> 71
<211> 636
<212> PRT
<213> 人工序列
<220>
<223> TcBuster转座酶- 40-18214
<400> 71

[0090]

Met Met Leu Asn Trp Leu Lys Ser Gly Lys Leu Glu Ser Gln Ser Gln
1 5 10 15

Glu Gln Ser Ser Cys Tyr Leu Glu Asn Ser Asn Cys Leu Pro Pro Thr
20 25 30

Leu Asp Ser Thr Asp Ile Ile Gly Glu Glu Asn Lys Ala Gly Thr Thr
35 40 45

Ser Arg Lys Lys Arg Lys Tyr Asp Glu Asp Tyr Leu Asn Phe Gly Phe
50 55 60

Thr Trp Thr Gly Asp Lys Asp Glu Pro Asn Gly Leu Cys Val Ile Cys
65 70 75 80

Glu Gln Val Val Asn Asn Ser Ser Leu Asn Pro Ala Lys Leu Lys Arg
85 90 95

His Leu Asp Thr Lys His Pro Thr Leu Lys Gly Lys Ser Glu Tyr Phe
100 105 110

Lys Arg Lys Cys Asn Glu Leu Asn Gln Lys Lys His Thr Phe Glu Arg
115 120 125

Tyr Val Arg Asp Asp Asn Lys Asn Leu Leu Lys Ala Ser Tyr Leu Val
130 135 140

Ser Leu Arg Ile Ala Lys Gln Gly Glu Ala Tyr Thr Ile Ala Glu Lys
145 150 155 160

Leu Ile Lys Pro Cys Thr Lys Asp Leu Thr Thr Cys Val Phe Gly Glu
165 170 175

Lys Phe Ala Ser Lys Val Asp Leu Val Pro Leu Ser Asp Thr Thr Ile
180 185 190

Ser Arg Arg Ile Glu Asp Met Ser Tyr Phe Cys Glu Ala Val Leu Val
195 200 205

Asn Arg Leu Glu Asn Ala Lys Cys Gly Phe Thr Leu Gln Met Asp Glu
210 215 220

[0091]

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

Lys Glu Leu Met Pro Phe Val Thr Thr Tyr Leu Cys Glu Lys Ser Phe
580 585 590
Ser Val Tyr Val Ala Thr Lys Thr Lys Tyr Arg Asn Arg Leu Asp Ala
595 600 605
Glu Asp Asp Met Arg Leu Gln Leu Thr Thr Ile His Pro Asp Ile Asp
610 615 620
Asn Leu Cys Asn Asn Lys Gln Ala Gln Lys Ser His
625 630 635

<210> 72
<211> 1911
<212> PRT
<213> 人工序列
<220>
<223> TcBuster转座酶- 40-18215
<400> 72

[0092]

Ala Thr Gly Ala Thr Gly Thr Thr Gly Ala Ala Thr Thr Gly Gly Cys
1 5 10 15

Thr Gly Ala Ala Ala Ala Gly Thr Gly Gly Ala Ala Ala Gly Cys Thr
20 25 30
Thr Gly Ala Ala Ala Gly Thr Cys Ala Ala Thr Cys Ala Cys Ala Gly
35 40 45
Gly Ala Ala Cys Ala Gly Ala Gly Thr Thr Cys Cys Thr Gly Cys Thr
50 55 60
Ala Cys Cys Thr Thr Gly Ala Gly Ala Ala Cys Thr Cys Thr Ala Ala
65 70 75 80
Cys Thr Gly Cys Cys Thr Gly Cys Cys Ala Cys Cys Ala Ala Cys Gly
85 90 95
Cys Thr Cys Gly Ala Thr Thr Cys Thr Ala Cys Ala Gly Ala Thr Ala
100 105 110
Thr Thr Ala Thr Cys Gly Gly Thr Gly Ala Ala Gly Ala Gly Ala Ala
115 120 125
Cys Ala Ala Ala Gly Cys Thr Gly Gly Thr Ala Cys Cys Ala Cys Cys
130 135 140
Thr Cys Thr Cys Gly Cys Ala Ala Gly Ala Ala Gly Cys Gly Gly Ala
145 150 155 160
Ala Ala Thr Ala Thr Gly Ala Cys Gly Ala Gly Gly Ala Cys Thr Ala
165 170 175
Thr Cys Thr Gly Ala Ala Cys Thr Thr Cys Gly Gly Thr Thr Thr Thr
180 185 190
Ala Cys Ala Thr Gly Gly Ala Cys Thr Gly Gly Cys Gly Ala Cys Ala
195 200 205
Ala Gly Gly Ala Thr Gly Ala Gly Cys Cys Cys Ala Ala Cys Gly Gly
210 215 220

[0093]

Ala Cys Thr Thr Thr Gly Thr Gly Thr Gly Ala Thr Thr Thr Gly Cys
225 230 235 240
Gly Ala Gly Cys Ala Gly Gly Thr Ala Gly Thr Cys Ala Ala Cys Ala
245 250 255
Ala Thr Thr Cys Cys Thr Cys Ala Cys Thr Thr Ala Ala Cys Cys Cys
260 265 270
Gly Gly Cys Cys Ala Ala Ala Cys Thr Gly Ala Ala Ala Cys Gly Cys
275 280 285
Cys Ala Thr Thr Thr Gly Gly Ala Cys Ala Cys Ala Ala Ala Gly Cys
290 295 300
Ala Thr Cys Cys Gly Ala Cys Gly Cys Thr Thr Ala Ala Ala Gly Gly
305 310 315 320
Cys Ala Ala Gly Ala Gly Cys Gly Ala Ala Thr Ala Cys Thr Thr Cys
325 330 335
Ala Ala Ala Ala Gly Ala Ala Ala Ala Thr Gly Thr Ala Ala Cys Gly
340 345 350
Ala Gly Cys Thr Cys Ala Ala Thr Cys Ala Ala Ala Ala Gly Ala Ala
355 360 365
Gly Cys Ala Thr Ala Cys Thr Thr Thr Thr Gly Ala Gly Cys Gly Ala
370 375 380
Thr Ala Cys Gly Thr Ala Ala Gly Gly Gly Ala Cys Gly Ala Thr Ala
385 390 395 400
Ala Cys Ala Ala Gly Ala Ala Cys Cys Thr Cys Cys Thr Gly Ala Ala
405 410 415
Ala Gly Cys Thr Thr Cys Thr Thr Ala Thr Cys Thr Cys Gly Thr Cys
420 425 430
Ala Gly Thr Thr Thr Gly Ala Gly Ala Ala Thr Ala Gly Cys Thr Ala
435 440 445
Ala Ala Cys Ala Gly Gly Gly Cys Gly Ala Gly Gly Cys Ala Thr Ala
450 455 460
Thr Ala Cys Cys Ala Thr Ala Gly Cys Gly Gly Ala Gly Ala Ala Gly
465 470 475 480
Thr Thr Gly Ala Thr Cys Ala Ala Gly Cys Cys Thr Thr Gly Cys Ala
485 490 495
Cys Cys Ala Ala Gly Gly Ala Thr Cys Thr Gly Ala Cys Ala Ala Cys
500 505 510
Thr Thr Gly Cys Gly Thr Ala Thr Thr Thr Gly Gly Ala Gly Ala Ala
515 520 525

Ala Ala Ala Thr Thr Cys Gly Cys Gly Ala Gly Cys Ala Ala Ala Gly
530 535 540
Thr Thr Gly Ala Thr Cys Thr Cys Gly Thr Cys Cys Cys Cys Cys Thr
545 550 555 560
Gly Thr Cys Cys Gly Ala Cys Ala Cys Gly Ala Cys Thr Ala Thr Thr

[0094]

```
                565                 570                 575
Thr Cys Ala Ala Gly Gly Cys Gly Ala Ala Thr Cys Gly Ala Ala Gly
            580                 585                 590
Ala Cys Ala Thr Gly Ala Gly Thr Thr Ala Cys Thr Thr Cys Thr Gly
        595                 600                 605
Thr Gly Ala Ala Gly Cys Cys Gly Thr Gly Cys Thr Gly Gly Thr Gly
    610                 615                 620
Ala Ala Cys Ala Gly Gly Thr Thr Gly Ala Ala Ala Ala Ala Thr Gly
625                 630                 635                 640
Cys Thr Ala Ala Ala Thr Gly Thr Gly Gly Gly Thr Thr Thr Ala Cys
                645                 650                 655
Gly Cys Thr Gly Cys Ala Gly Ala Thr Gly Gly Ala Cys Gly Ala Gly
            660                 665                 670
Thr Cys Ala Ala Cys Ala Gly Ala Thr Gly Thr Thr Gly Cys Cys Gly
        675                 680                 685
Gly Thr Cys Thr Thr Gly Cys Ala Ala Thr Cys Cys Thr Gly Cys Thr
    690                 695                 700
Thr Gly Thr Gly Thr Thr Thr Gly Thr Thr Ala Gly Gly Thr Ala Cys
705                 710                 715                 720
Ala Thr Ala Cys Ala Thr Gly Ala Ala Ala Gly Cys Thr Cys Thr Thr
                725                 730                 735
Thr Thr Gly Ala Gly Gly Ala Gly Gly Ala Thr Ala Thr Gly Thr Thr
            740                 745                 750
Gly Thr Thr Cys Thr Gly Cys Ala Ala Ala Gly Cys Ala Cys Thr Thr
        755                 760                 765
Cys Cys Cys Ala Cys Thr Cys Ala Gly Ala Cys Gly Ala Cys Ala Gly
    770                 775                 780
Gly Gly Gly Ala Gly Gly Ala Gly Ala Thr Thr Thr Thr Cys Ala Ala
785                 790                 795                 800
Thr Cys Thr Thr Cys Thr Cys Ala Ala Thr Gly Cys Cys Thr Ala Thr
                805                 810                 815
Thr Thr Cys Gly Ala Ala Ala Ala Gly Cys Ala Cys Thr Cys Cys Ala
            820                 825                 830
Thr Cys Cys Cys Ala Thr Gly Gly Ala Ala Thr Cys Thr Gly Thr Gly
        835                 840                 845
Thr Thr Ala Cys Cys Ala Cys Ala Thr Thr Thr Gly Cys Ala Cys Ala
    850                 855                 860
Gly Ala Cys Gly Gly Thr Gly Cys Cys Ala Ala Gly Gly Cys Ala Ala
865                 870                 875                 880
Thr Gly Gly Thr Ala Gly Gly Ala Gly Thr Thr Ala Thr Thr Ala Ala
                885                 890                 895
Ala Gly Gly Ala Gly Thr Cys Ala Thr Ala Gly Cys Gly Ala Gly Ala
            900                 905                 910
Ala Thr Ala Ala Ala Ala Ala Ala Ala Cys Thr Cys Gly Thr Cys Cys
```

915 920 925
Cys Thr Gly Ala Thr Ala Thr Ala Ala Ala Ala Gly Cys Thr Ala Gly
930 935 940
Cys Cys Ala Cys Thr Gly Thr Thr Gly Cys Cys Thr Gly Cys Ala Thr
945 950 955 960
Cys Gly Cys Cys Ala Cys Gly Cys Thr Thr Thr Gly Gly Cys Thr Gly
965 970 975
Thr Ala Ala Ala Gly Cys Gly Ala Ala Thr Ala Cys Cys Gly Ala Ala
980 985 990
Thr Gly Cys Ala Thr Thr Gly Cys Ala Cys Gly Ala Gly Gly Thr Gly
995 1000 1005
Cys Thr Cys Ala Ala Thr Gly Ala Cys Gly Cys Thr Gly Thr Thr
1010 1015 1020
Ala Ala Ala Ala Thr Gly Ala Thr Cys Ala Ala Cys Thr Thr Cys
1025 1030 1035

Ala Thr Cys Ala Ala Gly Thr Cys Thr Cys Gly Gly Cys Cys Gly
1040 1045 1050
Thr Thr Gly Ala Ala Thr Gly Cys Gly Cys Gly Cys Gly Thr Cys
1055 1060 1065
Thr Thr Cys Gly Cys Thr Thr Thr Gly Cys Thr Gly Thr Gly Thr
1070 1075 1080
Gly Ala Cys Gly Ala Thr Thr Thr Gly Gly Gly Gly Ala Gly Cys
1085 1090 1095
Cys Thr Gly Cys Ala Thr Ala Ala Ala Ala Ala Thr Cys Thr Thr
1100 1105 1110
Cys Thr Thr Cys Thr Thr Cys Ala Thr Ala Cys Cys Gly Ala Ala
1115 1120 1125
Gly Thr Gly Ala Gly Gly Thr Gly Gly Cys Thr Gly Thr Cys Thr
1130 1135 1140
Ala Gly Ala Gly Gly Ala Ala Ala Gly Gly Thr Gly Cys Thr Gly
1145 1150 1155
Ala Cys Cys Cys Gly Ala Thr Thr Thr Thr Gly Gly Gly Ala Ala
1160 1165 1170
Cys Thr Gly Ala Gly Ala Gly Ala Thr Gly Ala Ala Ala Thr Thr
1175 1180 1185
Ala Gly Ala Ala Thr Thr Thr Thr Cys Thr Thr Cys Ala Ala Cys
1190 1195 1200
Gly Ala Ala Ala Gly Gly Gly Ala Ala Thr Thr Thr Gly Cys Cys
1205 1210 1215
Gly Gly Gly Ala Ala Ala Thr Thr Gly Ala Ala Cys Gly Ala Cys
1220 1225 1230
Ala Cys Cys Ala Gly Thr Thr Gly Gly Thr Thr Gly Cys Ala Ala
1235 1240 1245

[0095]

[0096]

```
Ala Ala  Thr Thr Thr Gly Gly  Cys Ala Thr Ala Thr  Ala Thr Ala
    1250                1255                1260
Gly Cys  Thr Gly Ala Cys Ala  Thr Ala Thr Thr Cys  Ala Gly Thr
    1265                1270                1275
Thr Ala  Thr Cys Thr Gly Ala  Ala Thr Gly Ala Ala  Gly Thr Thr
    1280                1285                1290
Ala Ala  Thr Cys Thr Thr Thr  Cys Cys Cys Thr Gly  Cys Ala Ala
    1295                1300                1305
Gly Gly  Gly Cys Cys Gly Ala  Ala Thr Ala Gly Cys  Ala Cys Ala
    1310                1315                1320
Ala Thr  Cys Thr Thr Cys Ala  Ala Gly Gly Thr Ala  Ala Ala Thr
    1325                1330                1335
Ala Gly  Cys Cys Gly Cys Ala  Thr Thr Ala Ala Cys  Ala Gly Thr
    1340                1345                1350
Ala Thr  Thr Ala Ala Ala Thr  Cys Ala Ala Ala Gly  Thr Thr Gly
    1355                1360                1365
Ala Ala  Gly Thr Thr Gly Thr  Gly Gly Gly Ala Ala  Gly Ala Gly
    1370                1375                1380
Thr Gly  Thr Ala Thr Ala Ala  Cys Gly Ala Ala Ala  Ala Ala Thr
    1385                1390                1395
Ala Ala  Cys Ala Cys Thr Gly  Ala Gly Thr Gly Thr  Thr Thr Thr
    1400                1405                1410
Gly Cys  Gly Ala Ala Cys Cys  Thr Cys Ala Ala Cys  Gly Ala Thr
    1415                1420                1425
Thr Thr  Thr Thr Thr Gly Gly  Ala Ala Ala Cys Thr  Thr Cys Ala
    1430                1435                1440
Ala Ala  Cys Ala Cys Thr Gly  Cys Gly Thr Thr Gly  Gly Ala Thr
    1445                1450                1455
Cys Cys  Ala Ala Ala Cys Cys  Thr Gly Ala Ala Gly  Thr Cys Thr
    1460                1465                1470
Ala Ala  Thr Ala Thr Thr Thr  Thr Gly Gly Ala Ala  Cys Ala Thr
    1475                1480                1485
Cys Thr  Cys Ala Ala Cys Gly  Gly Thr Cys Thr Thr  Ala Ala Gly
    1490                1495                1500
Ala Ala  Cys Ala Cys Cys Thr  Thr Thr Cys Thr Gly  Gly Ala Gly
    1505                1510                1515
Thr Ala  Thr Thr Thr Thr Cys  Cys Ala Cys Cys Thr  Ala Cys Gly
    1520                1525                1530
Thr Gly  Thr Ala Ala Thr Ala  Ala Thr Ala Thr Cys  Thr Cys Cys
    1535                1540                1545
Thr Gly  Gly Gly Thr Gly Gly  Ala Gly Ala Ala Thr  Cys Cys Thr
    1550                1555                1560
Thr Thr  Cys Ala Ala Thr Gly  Ala Ala Thr Gly Cys  Gly Gly Thr
    1565                1570                1575
```

[0097]

Ala Ala Cys Gly Thr Cys Gly Ala Thr Ala Cys Ala Cys Thr Cys
1580 1585 1590
Cys Cys Ala Ala Thr Ala Ala Ala Ala Gly Ala Gly Ala Gly Gly
1595 1600 1605
Gly Ala Ala Cys Ala Ala Thr Thr Gly Ala Thr Thr Gly Ala Cys
1610 1615 1620
Ala Thr Ala Cys Gly Gly Ala Cys Thr Gly Ala Thr Ala Cys Gly
1625 1630 1635
Ala Cys Ala Thr Thr Gly Ala Ala Ala Thr Cys Thr Thr Cys Ala
1640 1645 1650
Thr Thr Cys Gly Thr Gly Cys Cys Thr Gly Ala Thr Gly Gly Thr
1655 1660 1665
Ala Thr Ala Gly Gly Ala Cys Cys Ala Thr Thr Cys Thr Gly Gly
1670 1675 1680
Ala Thr Cys Ala Ala Ala Cys Thr Gly Ala Thr Gly Gly Ala Cys
1685 1690 1695
Gly Ala Ala Thr Thr Thr Cys Cys Ala Gly Ala Ala Ala Thr Thr
1700 1705 1710
Ala Gly Cys Ala Ala Ala Cys Gly Ala Gly Cys Thr Gly Thr Cys
1715 1720 1725
Ala Ala Ala Gly Ala Gly Cys Thr Cys Ala Thr Gly Cys Cys Ala
1730 1735 1740
Thr Thr Thr Gly Thr Ala Ala Cys Cys Ala Cys Thr Thr Ala Cys
1745 1750 1755
Cys Thr Cys Thr Gly Thr Gly Ala Gly Ala Ala Ala Thr Cys Ala
1760 1765 1770
Thr Thr Thr Thr Cys Cys Gly Thr Cys Thr Ala Thr Gly Thr Ala
1775 1780 1785
Gly Cys Cys Ala Cys Ala Ala Ala Ala Ala Cys Ala Ala Ala Ala
1790 1795 1800
Thr Ala Thr Cys Gly Ala Ala Ala Thr Ala Gly Ala Cys Thr Thr
1805 1810 1815
Gly Ala Thr Gly Cys Thr Gly Ala Ala Gly Ala Cys Gly Ala Thr
1820 1825 1830
Ala Thr Gly Cys Gly Ala Cys Thr Cys Cys Ala Ala Cys Thr Thr
1835 1840 1845
Ala Cys Thr Ala Cys Thr Ala Thr Cys Cys Ala Thr Cys Cys Ala
1850 1855 1860
Gly Ala Cys Ala Thr Thr Gly Ala Cys Ala Ala Cys Cys Thr Thr
1865 1870 1875
Thr Gly Thr Ala Ala Cys Ala Ala Cys Ala Ala Gly Cys Ala Gly
1880 1885 1890
Gly Cys Thr Cys Ala Gly Ala Ala Ala Thr Cys Cys Cys Ala Cys
1895 1900 1905

Thr Gly Ala
1910

<210> 73
<211> 107
<212> PRT
<213> 人工序列
<220>
<223> FKBP12 -40-18091
<400> 73

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

[0098]

<210> 74
<211> 321
<212> DNA
<213> 人工序列
<220>
<223> FKBP12 -40-18092
<400> 74

ggggtccagg tcgagactat ttcaccaggg gatgggcgaa catttccaaa aaggggccag 60
acttgcgtcg tgcattacac cgggatgctg gaggacggga agaaagtgga cagctccagg 120
gatcgcaaca agcccttcaa gttcatgctg ggaaagcagg aagtgatccg aggatgggag 180
gaaggcgtgg cacagatgtc agtcggccag cgggccaaac tgaccattag ccctgactac 240
gcttatggag caacaggcca cccagggatc attccccctc atgccaccct ggtcttcgat 300
gtggaactgc tgaagctgga g 321

<210> 75
<211> 5
<212> PRT
<213> 人工序列
<220>
<223> 接头-40-18093
<400> 75

Gly Gly Gly Gly Ser

1 5

<210> 76

<211> 15

<212> PRT

<213> 人工序列

<220>

<223> 接头-40-18094

<400> 76

Gly Gly Ala Gly Gly Ala Gly Gly Ala Gly Gly Ala Thr Cys Cys
1 5 10 15

<210> 77

<211> 282

<212> PRT

<213> 人工序列

<220>

<223> 截短的 Cas9-40-18095

<400> 77

[0099]

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

210 215 220

Ser Gly Ser Trp Tyr Val Glu Thr Leu Asp Asp Ile Phe Glu Gln Trp
225 230 235 240
Ala His Ser Glu Asp Leu Gln Ser Leu Leu Leu Arg Val Ala Asn Ala
245 250 255
Val Ser Val Lys Gly Ile Tyr Lys Gln Met Pro Gly Cys Phe Asn Phe
260 265 270
Leu Arg Lys Lys Leu Phe Phe Lys Thr Ser
275 280

<210> 78
<211> 843
<212> DNA
<213> 人工序列
<220>
<223> 截短的 Cas9-40-18096
<400> 78

[0100]

tttggggacg tgggggccct ggagtctctg cgaggaaatg ccgatctggc ttacatcctg 60
agcatggaac cctgcggcca ctgtctgatc attaacaatg tgaacttctg cagagaaagc 120
ggactgcgaa cacggactgg ctccaatatt gactgtgaga agctgcggag aaggttctct 180
agtctgcact ttatggtcga agtgaaaggg gatctgaccg ccaagaaaat ggtgctggcc 240
ctgctggagc tggctcagca ggaccatgga gctctggatt gctgcgtggt cgtgatcctg 300
tcccacgggt gccaggcttc tcatctgcag ttccccggag cagtgtacgg aacagacggc 360
tgtcctgtca gcgtggagaa gatcgtcaac atcttcaacg gcacttcttg ccctagtctg 420
gggggaaagc caaaactgtt ctttatccag gcctgtggcg gggaacagaa agatcacggc 480
ttcgaggtgg ccagcaccag ccctgaggac gaatcaccag ggagcaaccc tgaaccagat 540
gcaactccat tccaggaggg actgaggacc tttgaccagc tggatgctat ctcaagcctg 600
cccactccta gtgacatttt cgtgtcttac agtaccttcc caggctttgt ctcatggcgc 660
gatcccaagt cagggagctg gtacgtggag acactggacg acatctttga acagtgggcc 720
cattcagagg acctgcagag cctgctgctg cgagtggcaa acgctgtctc tgtgaagggc 780
atctacaaac agatgcccgg gtgcttcaat tttctgagaa agaaactgtt ctttaagact 840
tcc 843

<210> 79
<211> 394
<212> PRT
<213> 人工序列
<220>
<223> 诱导型促凋亡多肽
<400> 79

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

[0101]

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

<210> 80
<211> 1182
<212> DNA
<213> 人工序列
<220>
<223> 诱导型促凋亡多肽
<400> 80

[0102]

ggggtccagg tcgagactat ttcaccaggg gatgggcgaa catttccaaa aaggggccag 60
acttgcgtcg tgcattacac cgggatgctg gaggacggga agaaagtgga cagctccagg 120
gatcgcaaca agcccttcaa gttcatgctg ggaaagcagg aagtgatccg aggatgggag 180
gaaggcgtgg cacagatgtc agtcggccag cgggccaaac tgaccattag ccctgactac 240
gcttatggag caacaggcca cccagggatc attccccctc atgccaccct ggtcttcgat 300
gtggaactgc tgaagctgga gggaggagga ggatccggat ttggggacgt gggggccctg 360
gagtctctgc gaggaaatgc cgatctggct tacatcctga gcatggaacc ctgcggccac 420
tgtctgatca ttaacaatgt gaacttctgc agagaaagcg gactgcgaac acggactggc 480
tccaatattg actgtgagaa gctgcggaga aggttctcta gtctgcactt tatggtcgaa 540
gtgaaagggg atctgaccgc caagaaaatg gtgctggccc tgctggagct ggctcagcag 600
gaccatggag ctctggattg ctgcgtggtc gtgatcctgt cccacgggtg ccaggcttct 660
catctgcagt tccccggagc agtgtacgga acagacggct gtcctgtcag cgtggagaag 720
atcgtcaaca tcttcaacgg cacttcttgc cctagtctgg ggggaaagcc aaaactgttc 780
tttatccagg cctgtggcgg ggaacagaaa gatcacggct tcgaggtggc cagcaccagc 840
cctgaggacg aatcaccagg gagcaaccct gaaccagatg caactccatt ccaggaggga 900
ctgaggacct ttgaccagct ggatgctatc tcaagcctgc ccactcctag tgacattttc 960
gtgtcttaca gtaccttccc aggctttgtc tcatggcgcg atcccaagtc agggagctgg 1020
tacgtggaga cactggacga catctttgaa cagtgggccc attcagagga cctgcagagc 1080
ctgctgctgc gagtggcaaa cgctgtctct gtgaagggca tctacaaaca gatgcccggg 1140
tgcttcaatt ttctgagaaa gaaactgttc tttaagactt cc 1182

<210> 81
<211> 463
<212> PRT
<213> 人工序列
<220>
<223> CSR-CD2z-46-17062
<400> 81

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

[0103]

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

Arg Gly Lys Gly His Asp Gly Leu Tyr Gln Gly Leu Ser Thr Ala Thr
435 440 445
Lys Asp Thr Tyr Asp Ala Leu His Met Gln Ala Leu Pro Pro Arg
450 455 460

<210> 82
<211> 184
<212> PRT
<213> 人工序列
<220>
<223> 具有 D11H-46-17119 的 CD2 ECD
<400> 82

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

[0104]

<210> 83
<211> 463
<212> PRT
<213> 人工序列
<220>
<223> CSR CD2z-D111H-46-17118
<400> 83

Met Ser Phe Pro Cys Lys Phe Val Ala Ser Phe Leu Leu Ile Phe Asn
1 5 10 15

[0105]

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

Asn Gln Leu Tyr Asn Glu Leu Asn Leu Gly Arg Arg Glu Glu Tyr Asp
370 375 380
Val Leu Asp Lys Arg Arg Gly Arg Asp Pro Glu Met Gly Gly Lys Pro
385 390 395 400
Arg Arg Lys Asn Pro Gln Glu Gly Leu Tyr Asn Glu Leu Gln Lys Asp
405 410 415
Lys Met Ala Glu Ala Tyr Ser Glu Ile Gly Met Lys Gly Glu Arg Arg
420 425 430
Arg Gly Lys Gly His Asp Gly Leu Tyr Gln Gly Leu Ser Thr Ala Thr
435 440 445
Lys Asp Thr Tyr Asp Ala Leu His Met Gln Ala Leu Pro Pro Arg
450 455 460

<210> 84
<211> 89
<212> PRT
<213> Homo Sapiens
<400> 84

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

[0106]

<210> 85
<211> 90
<212> PRT
<213> Homo Sapiens
<400> 85

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

Ser Asn Pro Leu Ser Ala Glu Phe Thr Thr
85 90

<210> 86
<211> 270
<212> DNA
<213> Homo Sapiens
<400> 86
atgctgcctg caccaaagaa cctggtggtg tctcatgtga cagaggatag tgccagactg 60
tcatggactg ctcccgacgc agccttcgat agttttatca tcgtgtaccg ggagaacatc 120
gaaaccggcg aggccattgt cctgacagtg ccagggtccg aacgctctta tgacctgaca 180
gatctgaagc ccggaactga gtactatgtg cagatcgccg gcgtcaaagg aggcaatatc 240
agcttccctc tgtccgcaat cttcaccaca 270

<210> 87
<211> 4
<212> PRT
<213> 人工序列
<220>
<223> A-B 环
<400> 87
Thr Glu Asp Ser
1

[0107]

<210> 88
<211> 7
<212> PRT
<213> 人工序列
<220>
<223> A-B 环
<400> 88
Thr Ala Pro Asp Ala Ala Phe
1 5

<210> 89
<211> 6
<212> PRT
<213> 人工序列
<220>
<223> A-B 环
<400> 89
Ser Glu Lys Val Gly Glu
1 5

<210> 90
<211> 4
<212> PRT
<213> 人工序列
<220>

<223> D-E 环
<400> 90
Gly Ser Glu Arg
1
<210> 91
<211> 5
<212> PRT
<213> 人工序列
<220>
<223> E-F 环
<400> 91
Gly Leu Lys Pro Gly
1 5
<210> 92
<211> 7
<212> PRT
<213> 人工序列
<220>
<223> F-G 环
<400> 92
Lys Gly Gly His Arg Ser Asn
1 5

[0108]

<210> 93
<211> 411
<212> DNA
<213> 人工序列
<220>
<223> PSMA centyrin
<400> 93
atgctgcctg caccaaagaa cctggtggtg tctcgggtga ccgaggactc tgccagactg 60
agctgggcca tcgacgagca gagggattgg ttcgagagct ttctgatcca gtatcaggag 120
tccgagaaag tgggcgaggc catcgtgctg acagtgcctg gcagcgagcg gtcctatgat 180
ctgaccggcc tgaagccagg cacagagtac accgtgtcca tctacggcgt gtatcacgtg 240
tacaggtcca atcctctgtc tgccatcttc accacaacca caacccctgc ccccagacct 300
cccacacccg cccctaccat cgcgagtcag cccctgagtc tgagacctga ggcctgcagg 360
ccagctgcag gaggagctgt gcacaccagg ggcctggact tcgcctgcga c 411
<210> 94
<211> 92
<212> PRT
<213> 人工序列
<220>
<223> PSMA centyrin
<400> 94
Met Leu Pro Ala Pro Lys Asn Leu Val Val Ser Arg Val Thr Glu Asp

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

<210> 95
<211> 336
<212> PRT
<213> 人工序列
<220>
<223> PSMA CARTyrin
<400> 95
Met Ala Leu Pro Val Thr Ala Leu Leu Leu Pro Leu Ala Leu Leu Leu
1 5 10 15

[0109]
His Ala Ala Arg Pro Met Leu Pro Ala Pro Lys Asn Leu Val Val Ser
20 25 30
Arg Val Thr Glu Asp Ser Ala Arg Leu Ser Trp Ala Ile Asp Glu Gln
35 40 45
Arg Asp Trp Phe Glu Ser Phe Leu Ile Gln Tyr Gln Glu Ser Glu Lys
50 55 60
Val Gly Glu Ala Ile Val Leu Thr Val Pro Gly Ser Glu Arg Ser Tyr
65 70 75 80
Asp Leu Thr Gly Leu Lys Pro Gly Thr Glu Tyr Thr Val Ser Ile Tyr
85 90 95
Gly Val Tyr His Val Tyr Arg Ser Asn Pro Leu Ser Ala Ile Phe Thr
100 105 110
Thr Thr Thr Thr Pro Ala Pro Arg Pro Pro Thr Pro Ala Pro Thr Ile
115 120 125
Ala Ser Gln Pro Leu Ser Leu Arg Pro Glu Ala Cys Arg Pro Ala Ala
130 135 140
Gly Gly Ala Val His Thr Arg Gly Leu Asp Phe Ala Cys Asp Ile Tyr
145 150 155 160
Ile Trp Ala Pro Leu Ala Gly Thr Cys Gly Val Leu Leu Leu Ser Leu
165 170 175
Val Ile Thr Leu Tyr Cys Lys Arg Gly Arg Lys Lys Leu Leu Tyr Ile
180 185 190
Phe Lys Gln Pro Phe Met Arg Pro Val Gln Thr Thr Gln Glu Glu Asp

195 200 205

Gly Cys Ser Cys Arg Phe Pro Glu Glu Glu Glu Gly Gly Cys Glu Leu
210 215 220
Arg Val Lys Phe Ser Arg Ser Ala Asp Ala Pro Ala Tyr Lys Gln Gly
225 230 235 240
Gln Asn Gln Leu Tyr Asn Glu Leu Asn Leu Gly Arg Arg Glu Glu Tyr
245 250 255
Asp Val Leu Asp Lys Arg Arg Gly Arg Asp Pro Glu Met Gly Gly Lys
260 265 270
Pro Arg Arg Lys Asn Pro Gln Glu Gly Leu Tyr Asn Glu Leu Gln Lys
275 280 285
Asp Lys Met Ala Glu Ala Tyr Ser Glu Ile Gly Met Lys Gly Glu Arg
290 295 300
Arg Arg Gly Lys Gly His Asp Gly Leu Tyr Gln Gly Leu Ser Thr Ala
305 310 315 320
Thr Lys Asp Thr Tyr Asp Ala Leu His Met Gln Ala Leu Pro Pro Arg
325 330 335

<210> 96
<211> 489
<212> DNA
<213> 人工序列
<220>
[0110]
<223> BCMA VH
<400> 96

gaagttcagc tgcttgaatc tggcggaggc ctggttcaac ctggcggatc tctgagactg 60
agctgtgccg ccagcggctt caccttcagc aattacgcca tgacctggat cagacaggcc 120
cctggcaaag gcctggaatg ggtgtccgga attacaggcg acggcggcag caccttttac 180
gccgattctg tgaagggcag attcaccatc agccgggaca acagcaagaa caccctgtac 240
ctgcagatga acagcctgag agccgaggac accgccgtgt actactgcgt gaaggactgg 300
aacaccacca tgatcaccga gagaggccag ggcacactgg tcaccgtgtc ctctacaaca 360
acaccggcgc ctagacctcc aacaccagct cctacaatcg cgagtcagcc cctgtctctc 420
agacccgaag cctgcaggcc agctgcagga ggagctgtgc acaccagggg cctggacttc 480
gcctgcgac 489

<210> 97
<211> 118
<212> PRT
<213> 人工序列
<220>
<223> BCMA VH
<400> 97

Glu Val Gln Leu Leu Glu Ser Gly Gly Gly Leu Val Gln Pro Gly Gly
1 5 10 15
Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Thr Phe Ser Asn Tyr
20 25 30

Ala Met Thr Trp Ile Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Val
35 40 45
Ser Gly Ile Thr Gly Asp Gly Gly Ser Thr Phe Tyr Ala Asp Ser Val
50 55 60
Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn Ser Lys Asn Thr Leu Tyr
65 70 75 80
Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Tyr Cys
85 90 95
Val Lys Asp Trp Asn Thr Thr Met Ile Thr Glu Arg Gly Gln Gly Thr
100 105 110
Leu Val Thr Val Ser Ser
115

<210> 98
<211> 362
<212> PRT
<213> 人工序列
<220>
<223> BCMA VCAR
<400> 98

[0111]

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

Tyr Ala Asp Ser Val Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn Ser
85 90 95
Lys Asn Thr Leu Tyr Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr
100 105 110
Ala Val Tyr Tyr Cys Val Lys Asp Trp Asn Thr Thr Met Ile Thr Glu
115 120 125
Arg Gly Gln Gly Thr Leu Val Thr Val Ser Ser Thr Thr Thr Pro Ala
130 135 140
Pro Arg Pro Pro Thr Pro Ala Pro Thr Ile Ala Ser Gln Pro Leu Ser
145 150 155 160
Leu Arg Pro Glu Ala Cys Arg Pro Ala Ala Gly Gly Ala Val His Thr
165 170 175
Arg Gly Leu Asp Phe Ala Cys Asp Ile Tyr Ile Trp Ala Pro Leu Ala
180 185 190

[0112]

Gly Thr Cys Gly Val Leu Leu Leu Ser Leu Val Ile Thr Leu Tyr Cys
195 200 205
Lys Arg Gly Arg Lys Lys Leu Leu Tyr Ile Phe Lys Gln Pro Phe Met
210 215 220
Arg Pro Val Gln Thr Thr Gln Glu Glu Asp Gly Cys Ser Cys Arg Phe
225 230 235 240
Pro Glu Glu Glu Glu Gly Gly Cys Glu Leu Arg Val Lys Phe Ser Arg
245 250 255
Ser Ala Asp Ala Pro Ala Tyr Lys Gln Gly Gln Asn Gln Leu Tyr Asn
260 265 270
Glu Leu Asn Leu Gly Arg Arg Glu Glu Tyr Asp Val Leu Asp Lys Arg
275 280 285
Arg Gly Arg Asp Pro Glu Met Gly Gly Lys Pro Arg Arg Lys Asn Pro
290 295 300
Gln Glu Gly Leu Tyr Asn Glu Leu Gln Lys Asp Lys Met Ala Glu Ala
305 310 315 320
Tyr Ser Glu Ile Gly Met Lys Gly Glu Arg Arg Arg Gly Lys Gly His
325 330 335
Asp Gly Leu Tyr Gln Gly Leu Ser Thr Ala Thr Lys Asp Thr Tyr Asp
340 345 350
Ala Leu His Met Gln Ala Leu Pro Pro Arg
355 360

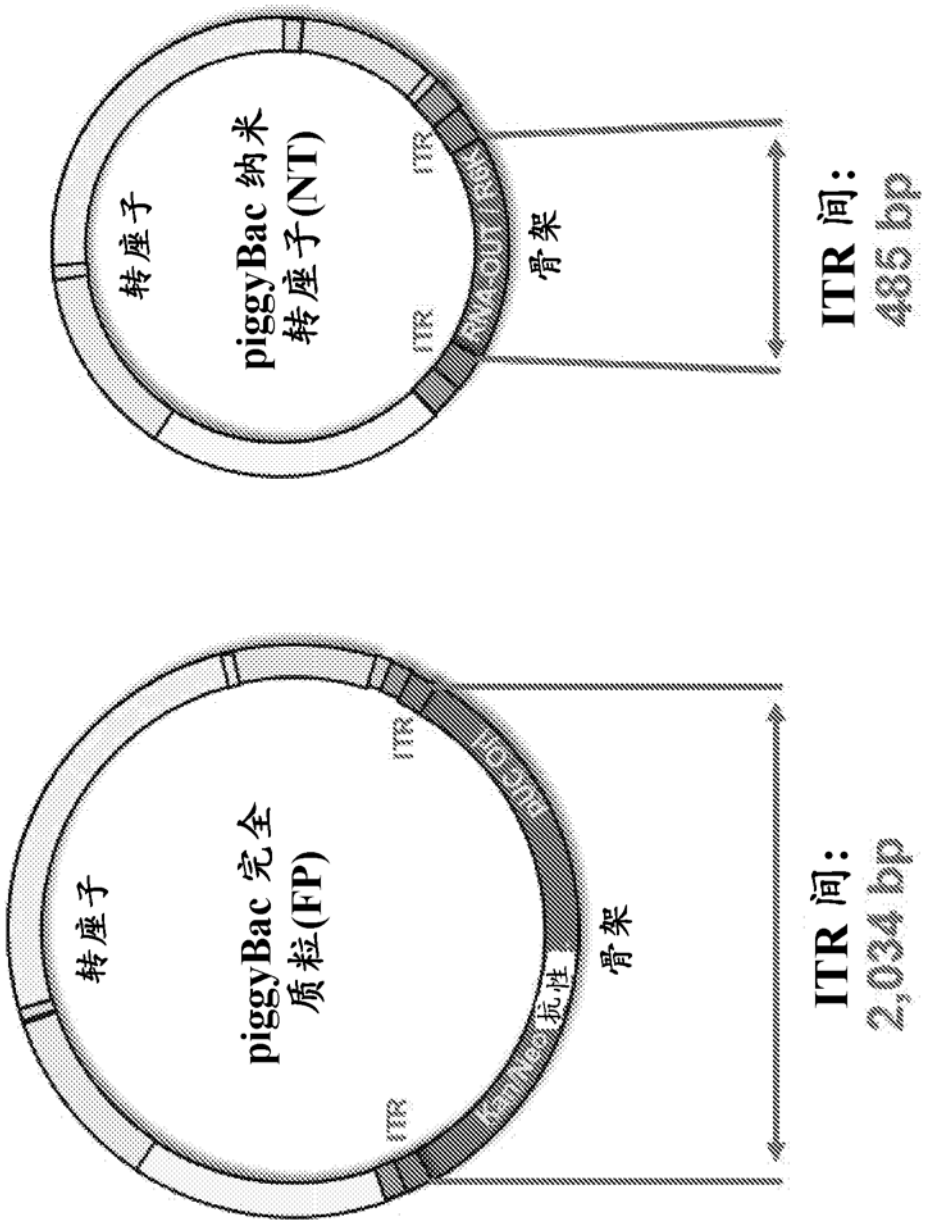

转座子
piggyBac 纳米
转座子(NT)
ITR
ITR
骨架
ITR 间:
485 bp
转座子
piggyBac 完全
质粒(FP)
ITR
ITR
骨架
ITR 间:
2,034 bp

图 1

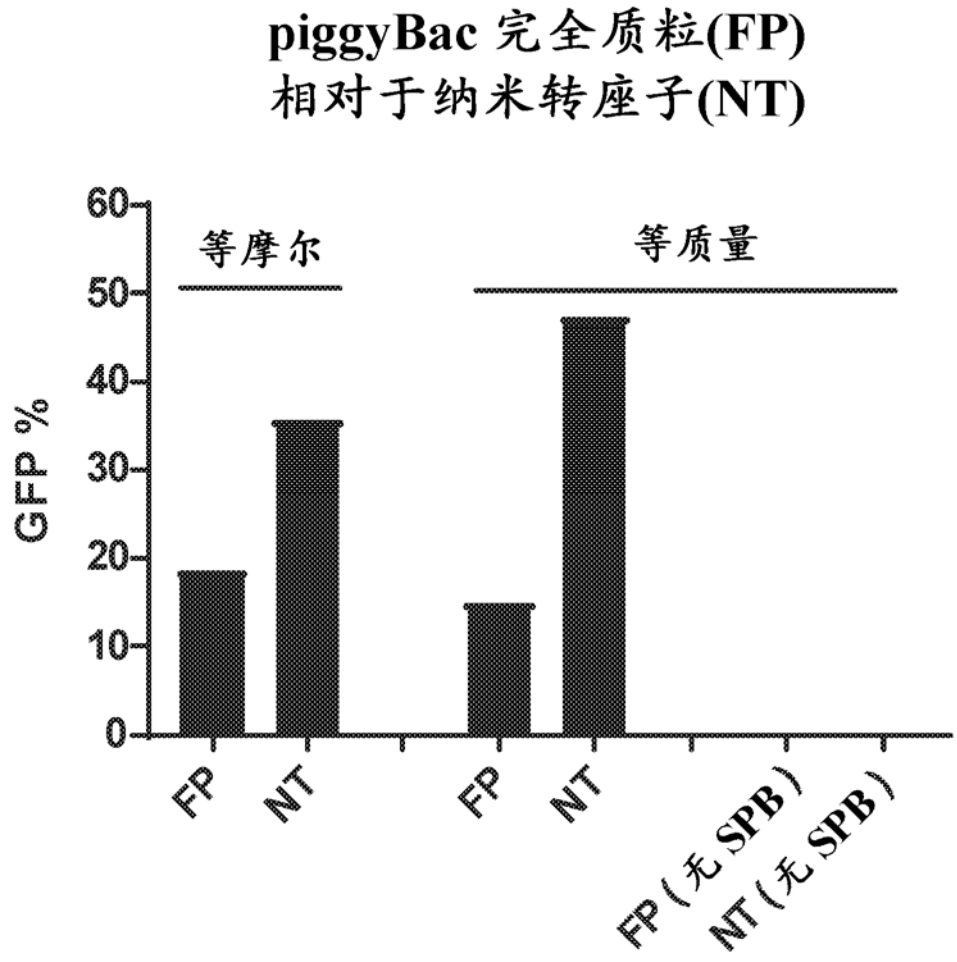

piggyBac 完全质粒(FP)
相对于纳米转座子(NT)
GFP %
60
50
40
30
20
10
0
等摩尔
等质量
FP
NT
FP
NT
FP(无 SPB)
NT(无 SPB)

图 2

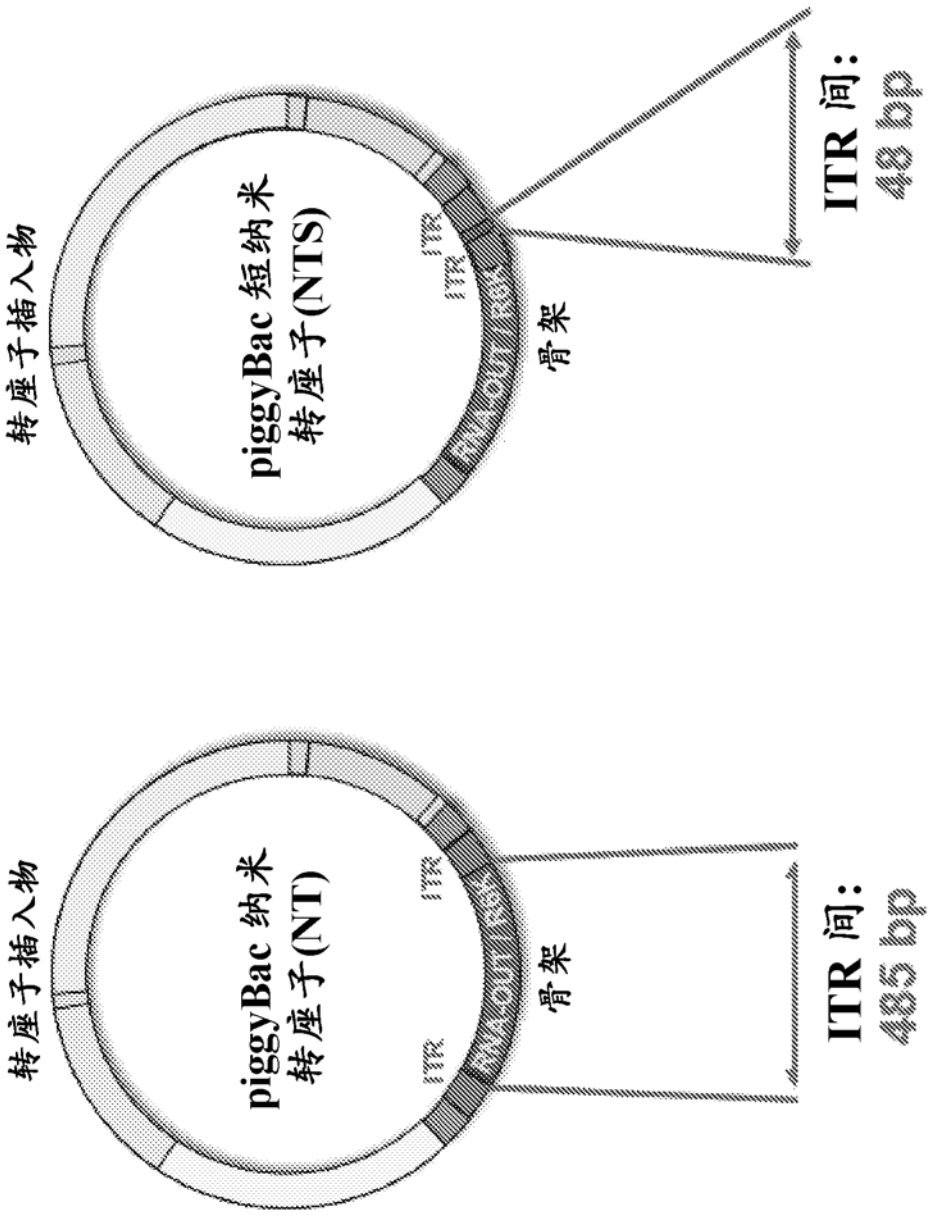

转座子插入物
piggyBac 短纳米
转座子(NTS)
ITR
ITR
骨架
ITR 间:
48 bp
转座子插入物
piggyBac 纳米
转座子(NT)
ITR
ITR
骨架
ITR 间:
485 bp

图 3

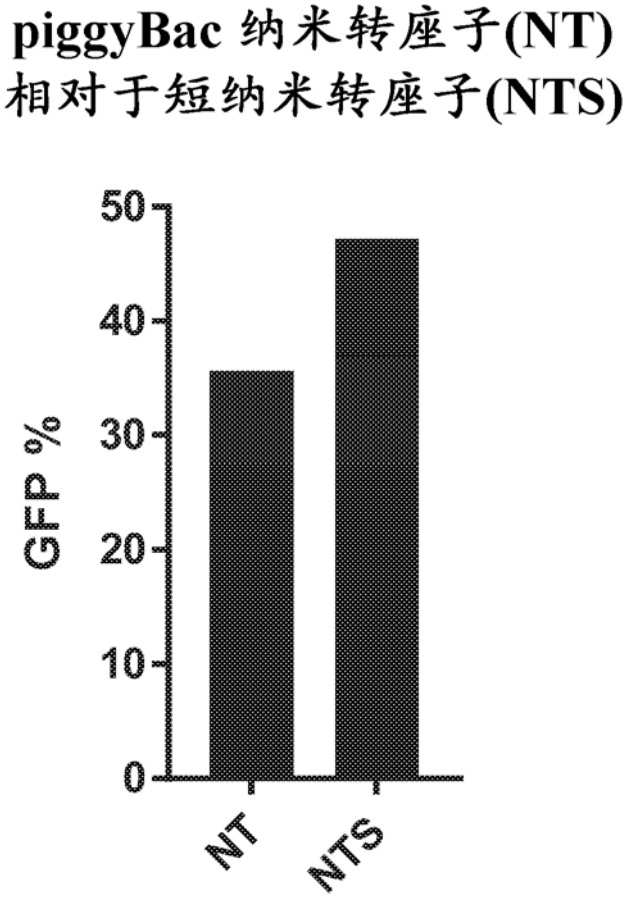


图 4

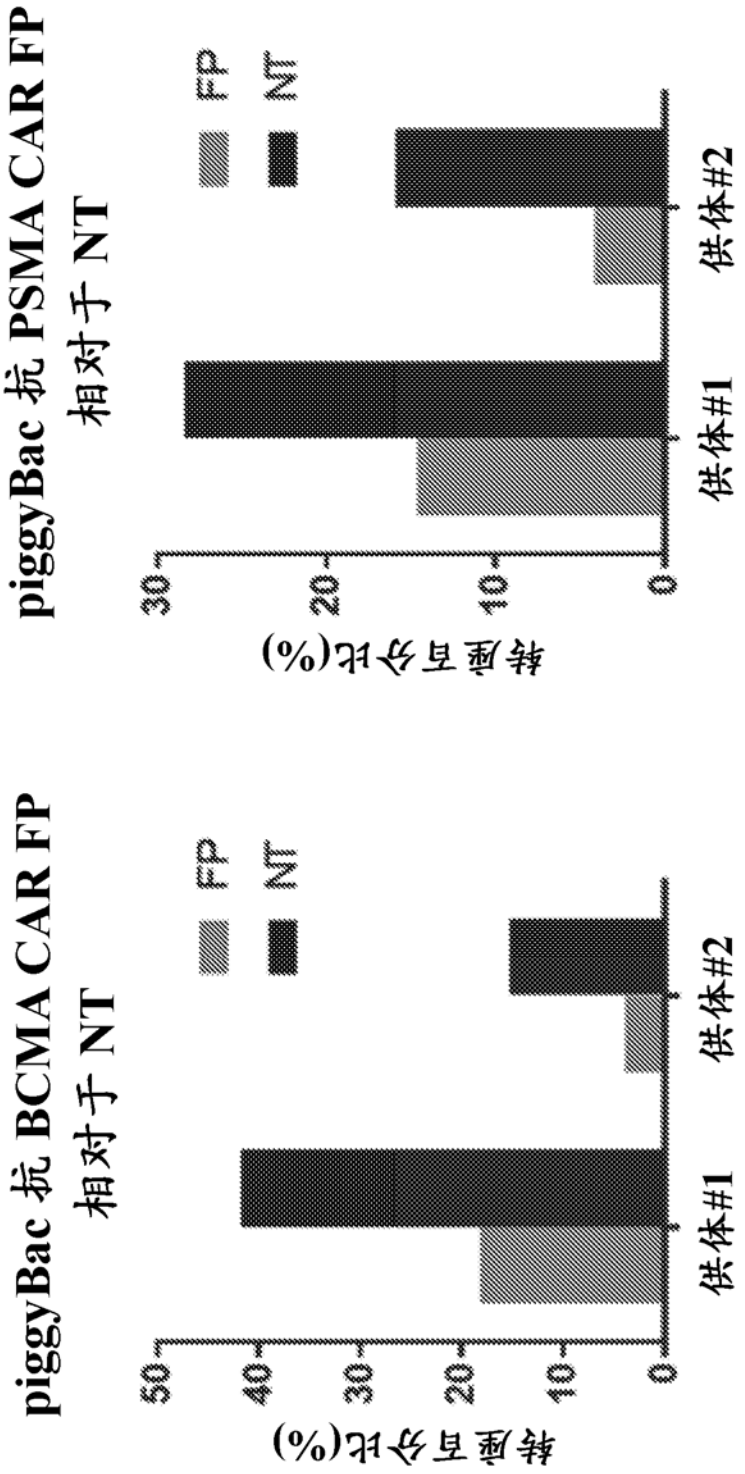

piggyBac 抗 PSMA CAR FP
相对于 NT
FP
NT
转座百分比(%)
0
10
20
30
供体#1
供体#2
piggyBac 抗 BCMA CAR FP
相对于 NT
FP
NT
转座百分比(%)
0
10
20
30
40
50
供体#1
供体#2

图 5

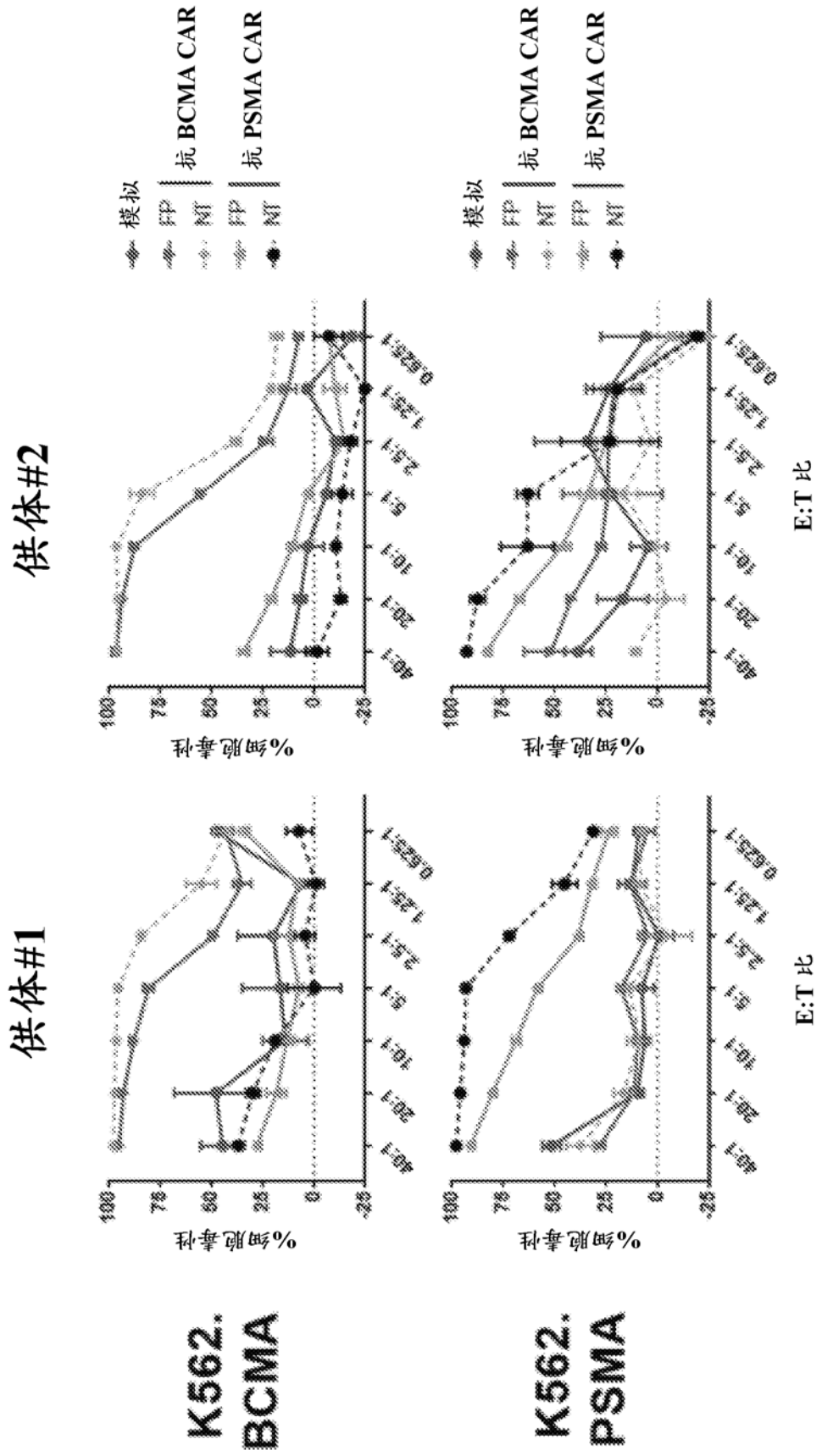

供体#1
供体#2
K562. BCMA
K562. PSMA
模拟
EP
MT
抗 BCMA CAR
抗 PSMA CAR
% 细胞毒性
E:T 比

图 6

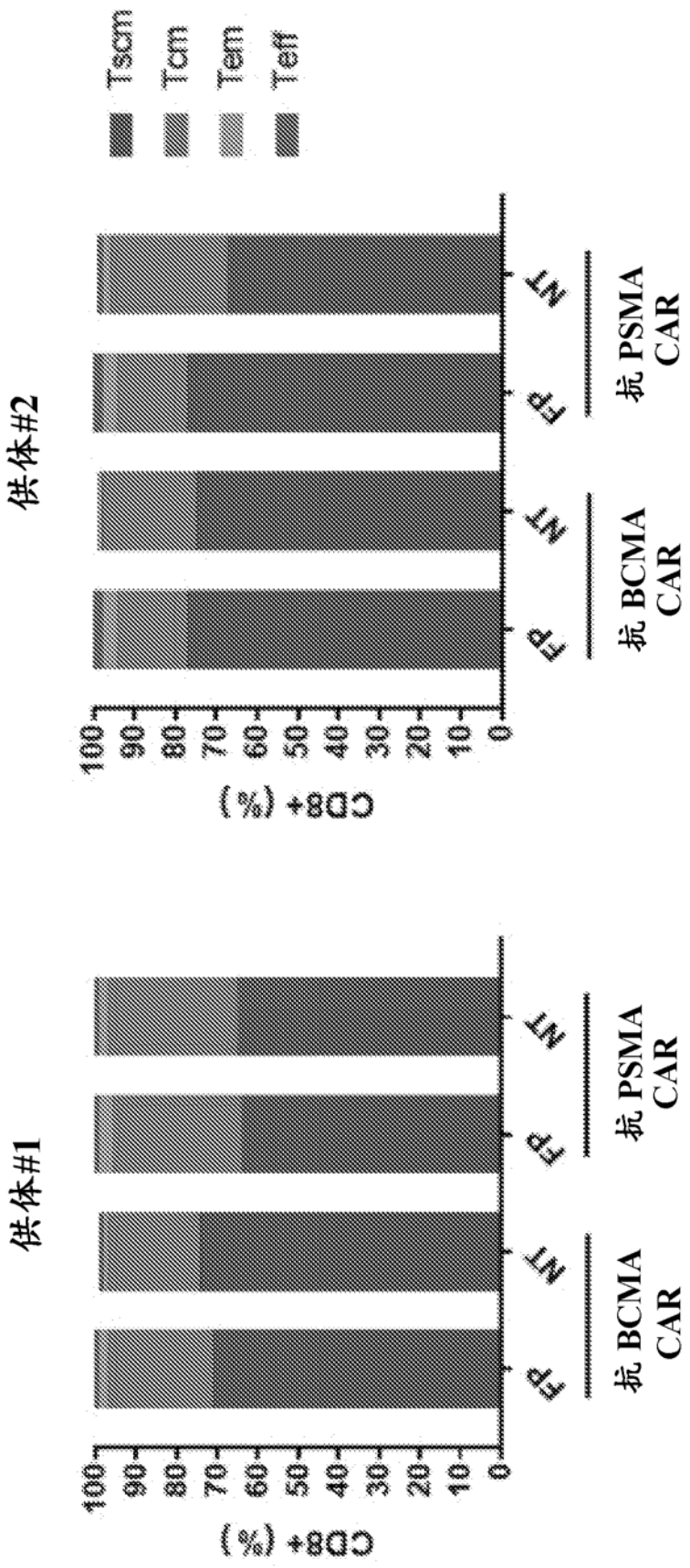

Tscm
Tcm
Tem
Teff
供体#2
CD8+ (%)
100
90
80
70
60
50
40
30
20
10
0
FP
NT
FP
NT
抗 BCMA CAR
抗 PSMA CAR
供体#1
CD8+ (%)
100
90
80
70
60
50
40
30
20
10
0
FP
NT
FP
NT
抗 BCMA CAR
抗 PSMA CAR

图 7

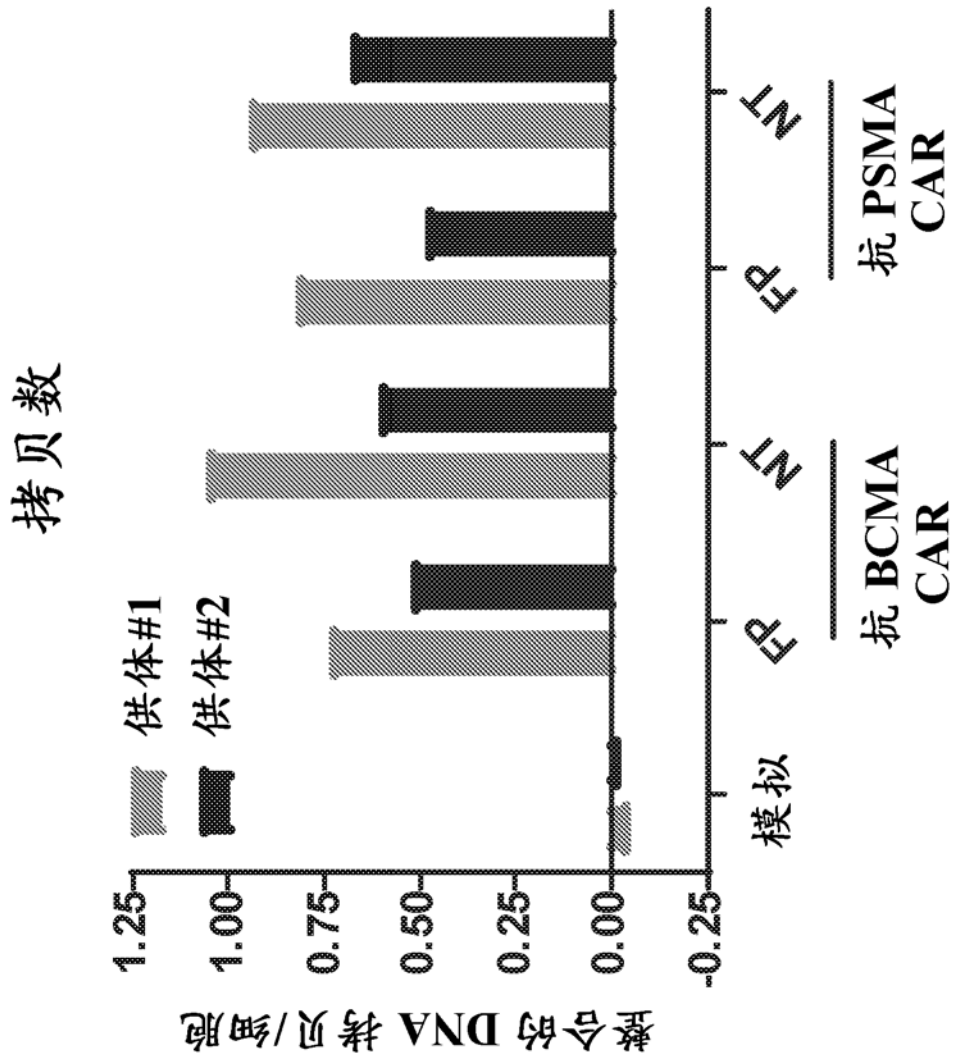

整合的 DNA 拷贝/细胞
拷贝数
1.25
1.00
0.75
0.50
0.25
0.00
-0.25
供体#1
供体#2
模拟
CP
NT
CP
NT
抗 BCMA CAR
抗 PSMA CAR

图 8

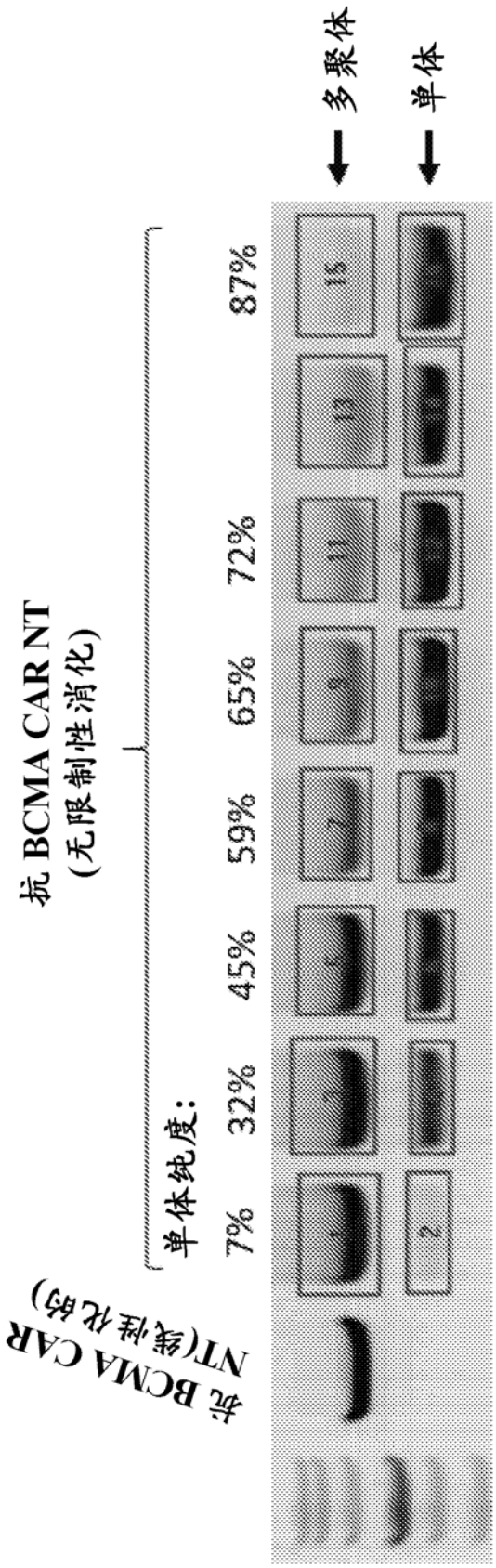

抗 BCMA CAR NT
(无限制性消化)
单体纯度:
7%
32%
45%
59%
65%
72%
87%
抗 BCMA CAR NT(限制性消化的)
单体
多聚体

图 9

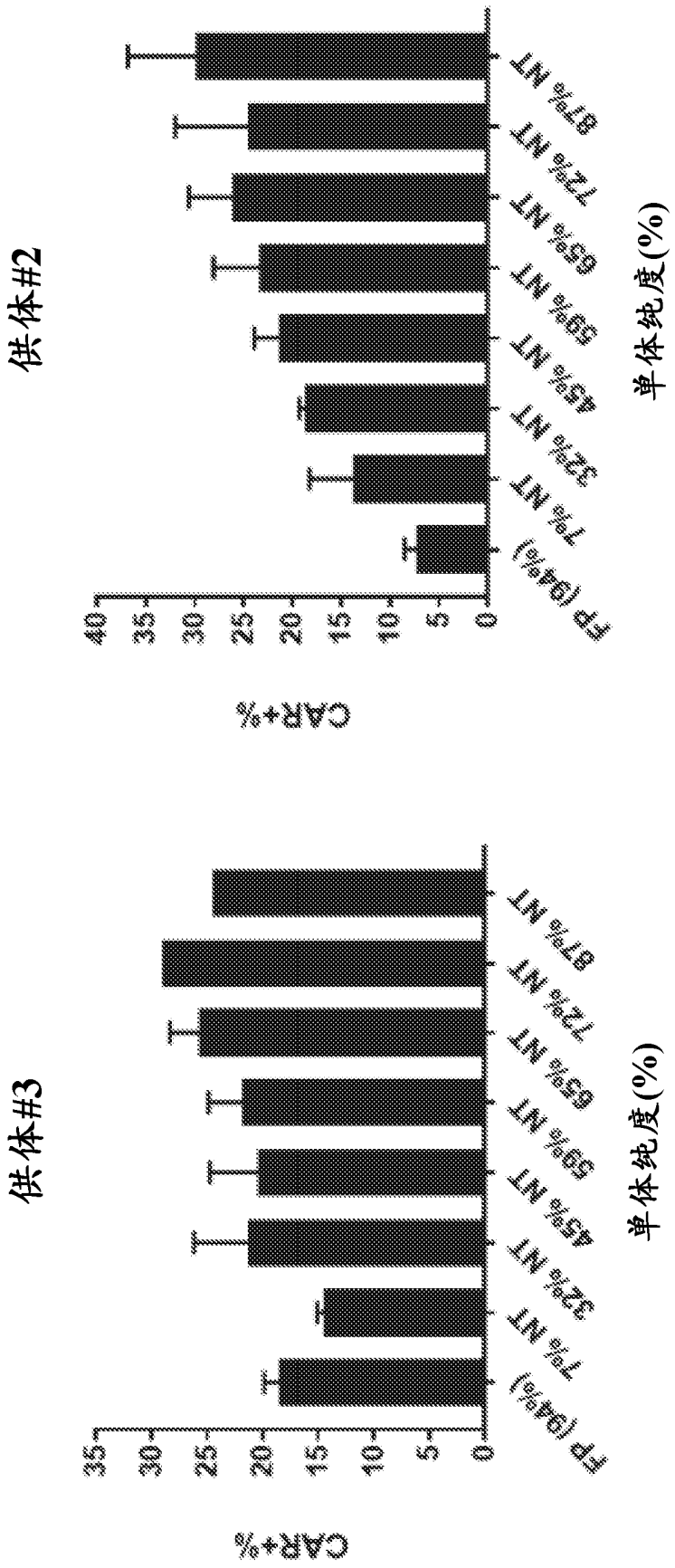

供体#2
CAR+%
单体纯度(%)
FP (94%)
70% NT
32% NT
45% NT
59% NT
65% NT
72% NT
87% NT
供体#3
CAR+%
单体纯度(%)
FP (94%)
70% NT
32% NT
45% NT
59% NT
65% NT
72% NT
87% NT

图 10

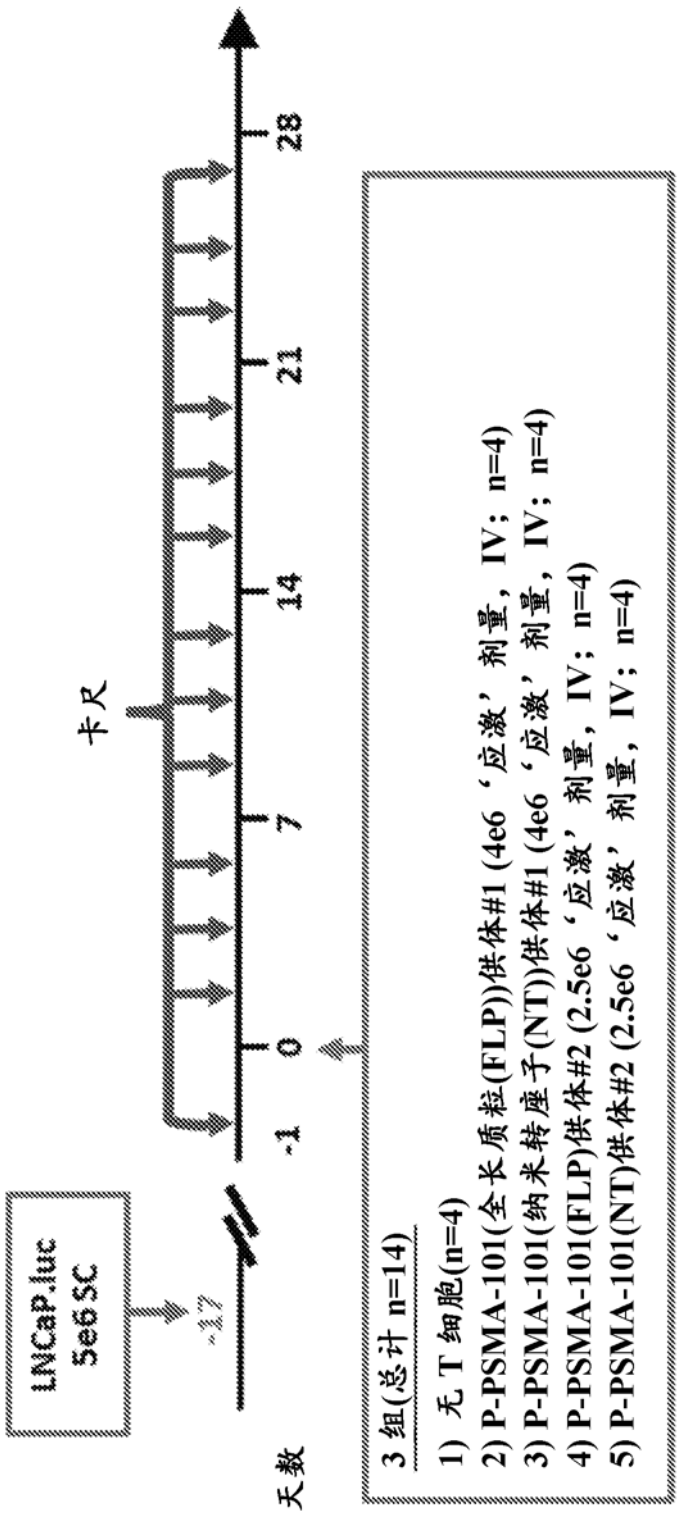
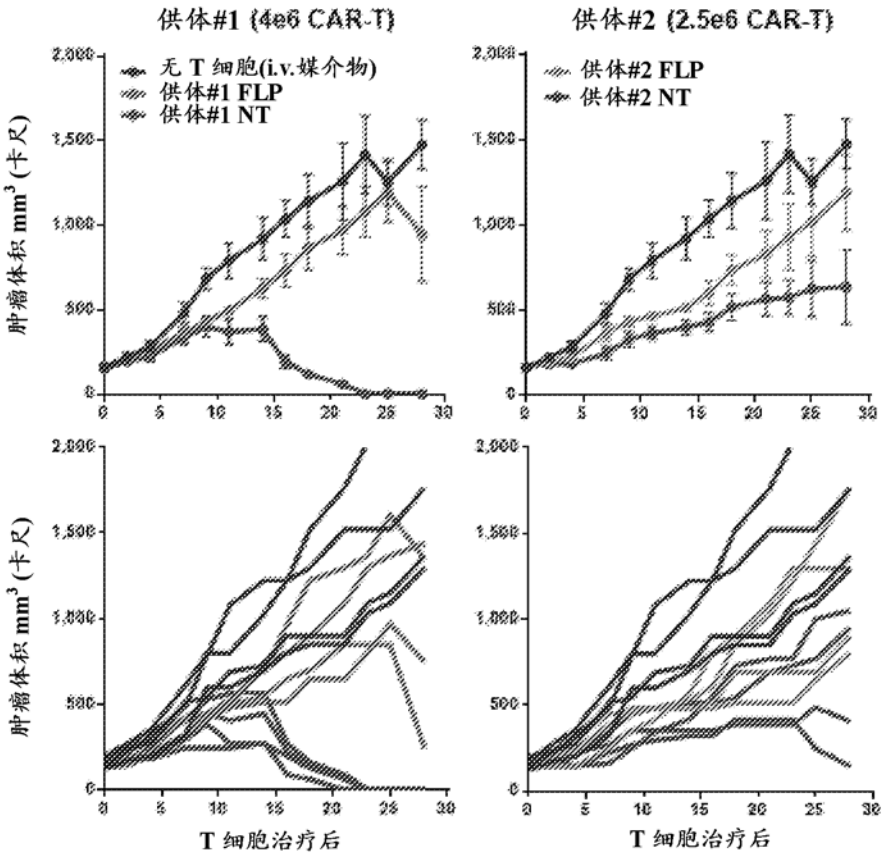

LNCaP.luc
5e6 SC
-17
-1
0
7
14
21
28
天数
卡尺
3 组(总计 n=14)
1) 无 T 细胞(n=4)
2) P-PSMA-101(全长质粒(FLP))供体#1 (4e6‘应激’剂量，IV；n=4)
3) P-PSMA-101(纳米转座子(NT))供体#1 (4e6‘应激’剂量，IV；n=4)
4) P-PSMA-101(FLP)供体#2 (2.5e6‘应激’剂量，IV；n=4)
5) P-PSMA-101(NT)供体#2 (2.5e6‘应激’剂量，IV；n=4)
研究设计

图 11


供体#1 (4e6 CAR-T)
供体#2 (2.5e6 CAR-T)
无 T 细胞(i.v.媒介物)
供体#1 FLP
供体#1 NT
供体#2 FLP
供体#2 NT
肿瘤体积 mm3 (卡尺)
肿瘤体积 mm3 (卡尺)
T 细胞治疗后
T 细胞治疗后

卡尺测量

图 12

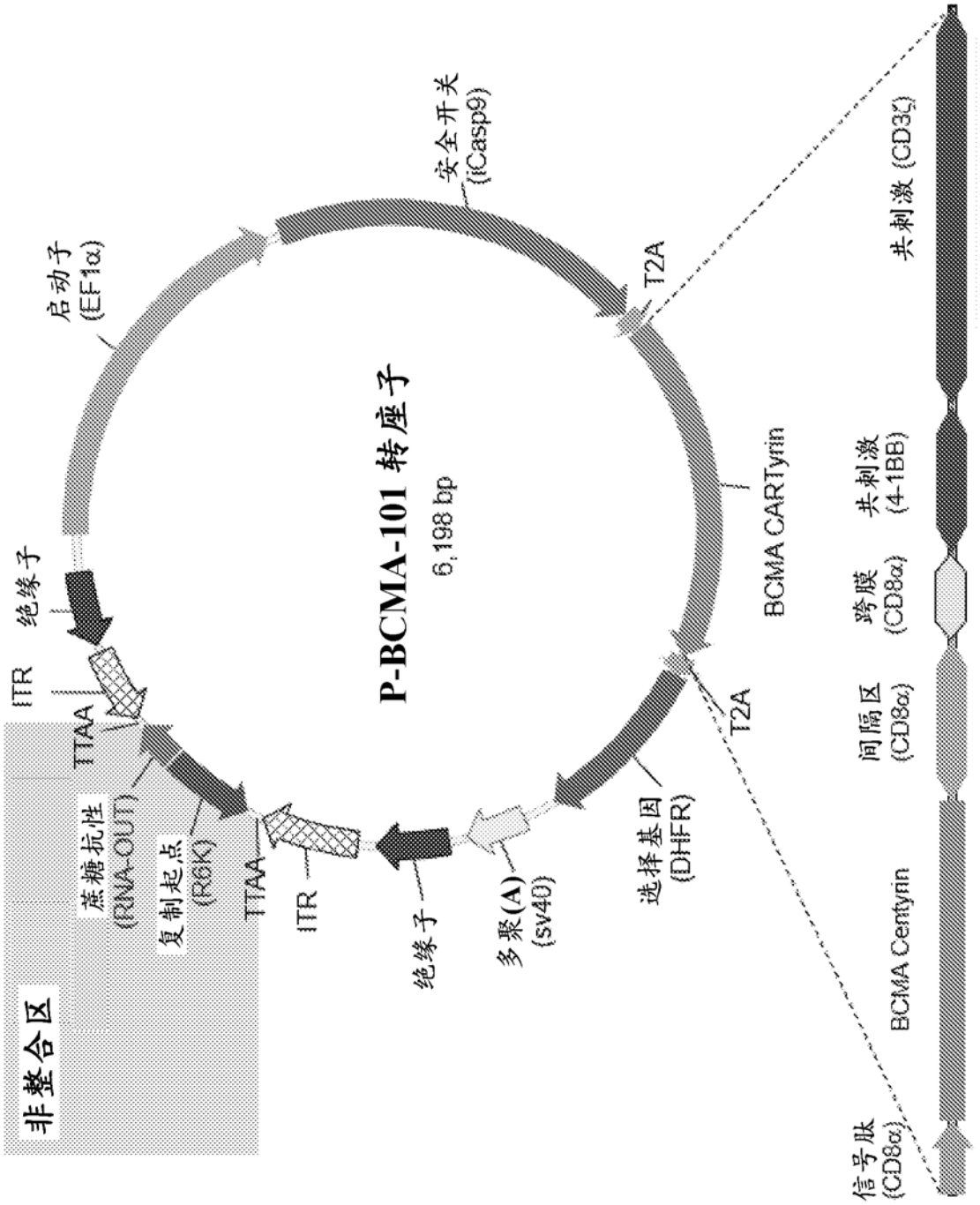

P-BCMA-101 转座子
6,198 bp
非整合区
ITR
TTAA
蒸腾抗性
(RNA-OUT)
复制起点
(R6K)
TTAA
ITR
绝缘子
多聚(A)
(sv40)
选择基因
(DHFR)
T2A
BCMA CARTyrin
T2A
安全开关
(iCasp9)
启动子
(EF1α)
绝缘子
信号肽
(CD8α)
BCMA Centyrin
间隔区
(CD8α)
跨膜
(CD8α)
共刺激
(4-1BB)
共刺激 (CD3ζ)

图 13

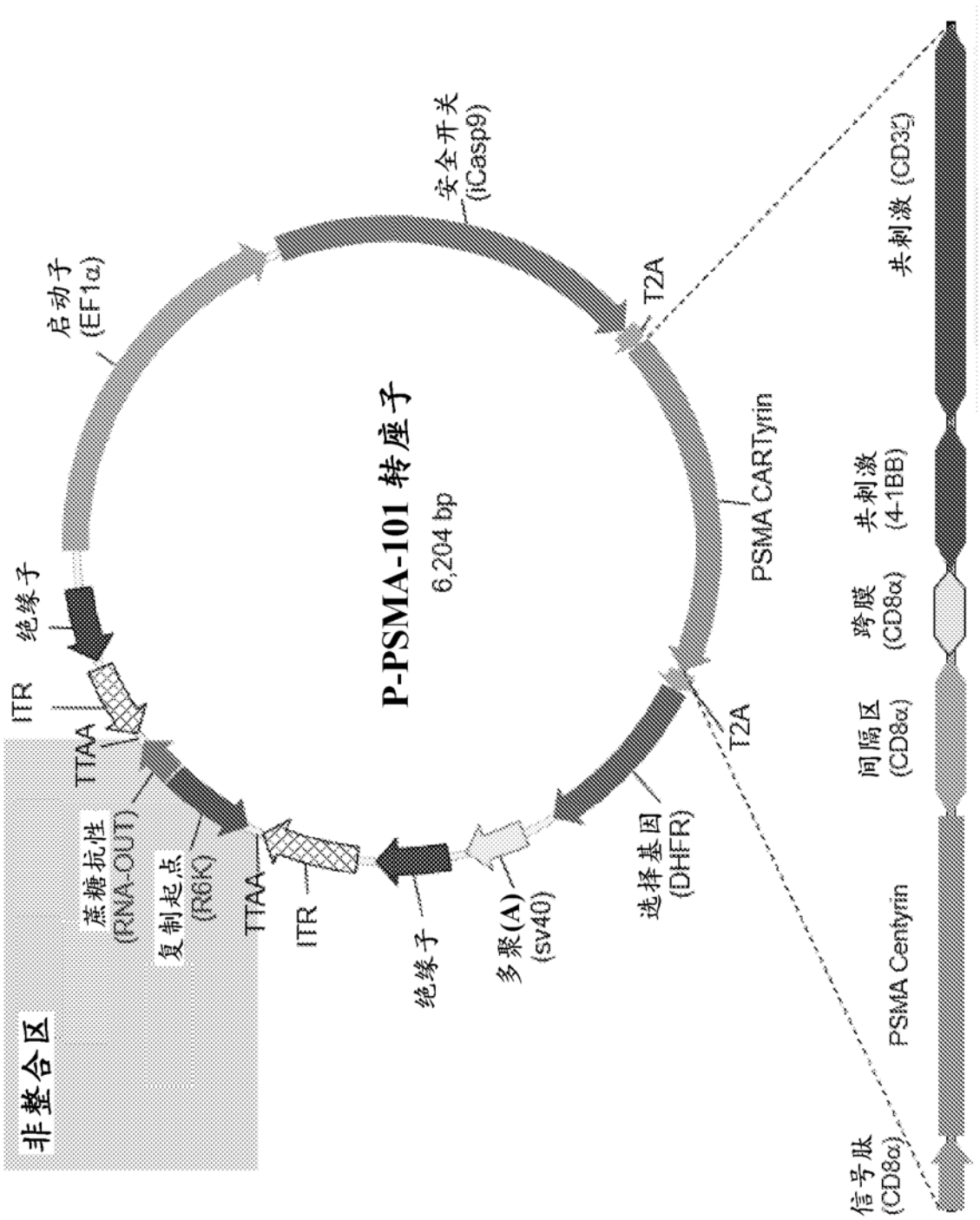

P-PSMA-101 转座子
6,204 bp
非整合区
ITR
绝缘子
启动子 (EF1α)
安全开关 (iCasp9)
T2A
PSMA CARTyrin
跨膜 (CD8α)
共刺激 (4-1BB)
共刺激 (CD3ζ)
间隔区 (CD8α)
信号肽 (CD8α)
PSMA Centyrin
T2A
选择基因 (DHFR)
多聚(A) (sv40)
绝缘子
ITR
TTAA
复制起点 (R6K)
蔗糖抗性 (RNA-OUT)
TTAA
ITR

图 14

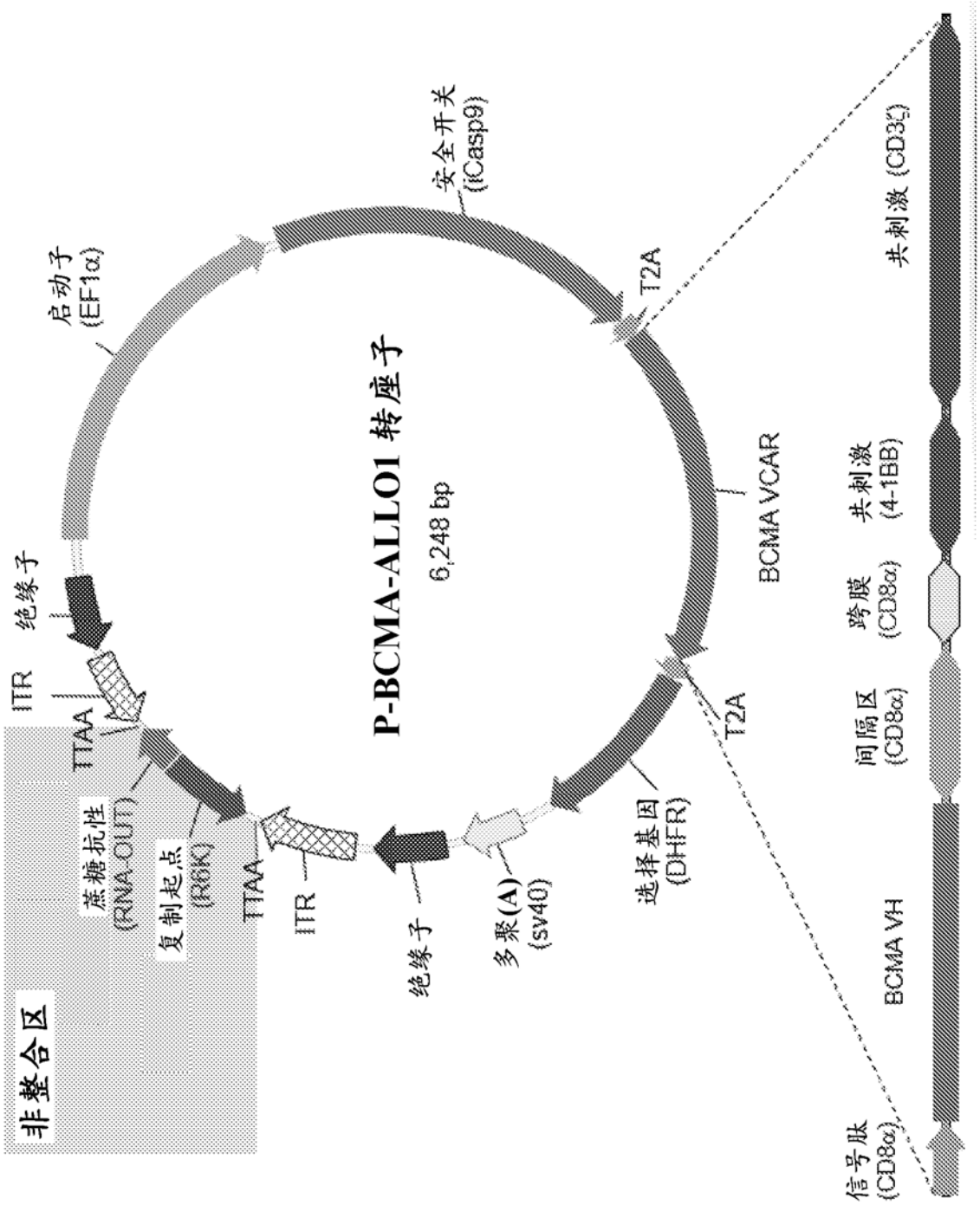

P-BCMA-ALLO1 转座子
6,248 bp
非整合区
ITR
绝缘子
启动子 (EF1α)
TTAA
蔗糖抗性 (RNA-OUT)
复制起点 (R6K)
TTAA
ITR
绝缘子
多聚(A) (sv40)
选择基因 (DHFR)
T2A
BCMA VCAR
T2A
安全开关 (iCasp9)
信号肽 (CD8α)
BCMA VH
间隔区 (CD8α)
跨膜 (CD8α)
共刺激 (4-1BB)
共刺激 (CD3ζ)

图 15