



(12)发明专利申请

(10)申请公布号 CN 108779447 A

(43)申请公布日 2018.11.09

(21)申请号 201680068869.7

(22)申请日 2016.09.23

(30)优先权数据

62/233,618 2015.09.28 US

(85)PCT国际申请进入国家阶段日

2018.05.24

(86)PCT国际申请的申请数据

PCT/US2016/053413 2016.09.23

(87)PCT国际申请的公布数据

W02017/058658 EN 2017.04.06

(71)申请人 天普大学-联邦高等教育系统

地址 美国宾夕法尼亚州

(72)发明人 K·哈利利 W·胡

(74)专利代理机构 北京戈程知识产权代理有限公司 11314

代理人 程伟

(51)Int.Cl.

G12N 9/22(2006.01)

C12N 15/113(2010.01)

C12Q 1/686(2018.01)

C07K 14/16(2006.01)

权利要求书6页 说明书42页 附图28页

(54)发明名称

用于HIV感染的RNA引导治疗的方法和组合物

(57)摘要

本发明涉及用于逆转录病毒如人免疫缺陷病毒(HIV-1)中靶标序列的特异性裂解的组合物和方法。所述组合物可包括编码成簇规律间隔短回文重复序列(CRISPR)相关的核酸内切酶的核酸、以及与人免疫缺陷病毒中靶标序列互补的向导RNA序列,据信所述组合物可被输送至具有HIV感染或处于接触HIV感染风险下的受试者的细胞。

1. 一种组合物,用于令整合入被人免疫缺陷病毒(HIV)潜伏性感染的宿主细胞基因组中的前病毒DNA失活,所述组合物包含:

至少一种编码成簇规律间隔短回文重复序列(CRISPR)相关的核酸内切酶的单离的核酸序列,以及

至少一种向导RNA(gRNA),所述至少一种gRNA具有与前病毒HIV DNA的长末端重复序列(LTR)中的靶标序列互补的间隔序列序列。

2. 根据权利要求1所述的组合物,其中,所述至少一种gRNA包含与靶标核酸序列互补的核酸序列,所述核酸序列与SEQ ID NOS:1至SEQ ID NOS:66中的一个或多个序列及其片段、突变体、变体或组合的序列一致性为至少75%。

3. 根据权利要求1所述的组合物,其中,所述至少一种gRNA包含一核酸序列,所述核酸序列与SEQ ID NOS:1至SEQ ID NOS:66中的一个或多个序列及其片段、突变体、变体或组合的序列一致性为至少75%。

4. 根据权利要求1所述的组合物,其中,所述至少一种gRNA包含至少一个与靶标核酸序列互补的核酸序列,所述核酸序列包含SEQ ID NOS:1至SEQ ID NOS:66,及其片段、突变体、变体或组合。

5. 根据权利要求1所述的组合物,其中,所述至少一种gRNA包含至少一种核酸序列,所述至少一种核酸序列包含SEQ ID NOS:1至SEQ ID NOS:66,及其片段、突变体、变体或组合。

6. 根据权利要求1所述的组合物,其中,所述至少一种gRNA选自:gRNA A,其具有与前病毒DNA中的靶标序列SEQ ID NO:1或靶标序列SEQ ID NO:2互补的间隔序列序列;gRNA B,其具有与前病毒DNA中的靶标序列SEQ ID NO:3或靶标序列SEQ ID NO:4互补的间隔序列序列;或gRNA A与gRNA B的组合。

7. 一种用于令整合入被HIV潜伏性感染的宿主细胞基因组内的前病毒人免疫缺陷病毒(HIV)DNA失活的方法,所述方法包括下列步骤:

以组合物处理所述宿主细胞,所述组合物包含成簇规律间隔短回文重复序列(CRISPR)相关的核酸内切酶、以及至少一种向导RNA(gRNA),所述至少一种gRNA具有与前病毒HIV DNA的长末端重复序列(LTR)中的靶标序列互补的间隔序列序列;以及

令所述前病毒DNA失活。

8. 根据权利要求7所述的方法,其中,所述至少一种gRNA包含与靶标核酸序列互补的核酸序列,所述核酸序列与SEQ ID NOS:1至SEQ ID NOS:66中的一个或多个序列及其片段、突变体、变体或组合的序列一致性为至少75%。

9. 根据权利要求7所述的方法,其中,所述至少一种gRNA包含核酸序列,所述核酸序列与SEQ ID NOS:1至SEQ ID NOS:66中的一个或多个序列及其片段、突变体、变体或组合的序列一致性为至少75%。

10. 根据权利要求7所述的方法,其中,所述至少一种gRNA包含至少一个与靶标核酸序列互补的核酸序列,所述核酸序列包含SEQ ID NOS:1至SEQ ID NOS:66,及其片段、突变体、变体或组合。

11. 根据权利要求7所述的方法,其中,所述至少一种gRNA包含至少一种核酸序列,所述至少一种核酸序列包含SEQ ID NOS:1至SEQ ID NOS:66,及其片段、突变体、变体或组合。

12. 根据权利要求7所述的方法,其中,所述至少一种gRNA选自:gRNA A,其具有与前病

毒DNA中的靶标序列SEQ ID NO:1或靶标序列SEQ ID NO:2互补的间隔序列序列;gRNA B,其具有与前病毒DNA中的靶标序列SEQ ID NO:3或靶标序列SEQ ID NO:4互补的间隔序列序列;或gRNA A与gRNA B的组合。

13. 一种用于令整合入潜伏性感染HIV的宿主细胞基因组内的前病毒人免疫缺陷病毒(HIV)DNA失活的慢病毒表达载体组合物,所述慢病毒表达载体组合物包括:

编码成簇规律间隔短回文重复序列(CRISPR)相关的核酸内切酶的单离的核酸、以及至少一种向导RNA(gRNA),所述gRNA具有与前病毒HIV DNA的长末端重复序列(LTR)中的靶标序列互补的间隔序列序列,

所述CRISPR相关的核酸内切酶和所述至少一种gRNA被包括于至少一种慢病毒表达载体中,

其中,所述至少一种慢病毒表达载体诱导所述CRISPR相关的核酸内切酶和所述至少一种gRNA在宿主细胞内的表达。

14. 根据权利要求13所述的慢病毒表达载体组合物,其中,所述至少一种gRNA包含与靶标核酸序列互补的核酸序列,所述核酸序列与SEQ ID NOS:1至SEQ ID NOS:66中的一个或多个序列及其片段、突变体、变体或组合的序列一致性为至少75%。

15. 根据权利要求13所述的慢病毒表达载体组合物,其中,所述至少一种gRNA包含一核酸序列,所述核酸序列与SEQ ID NOS:1至SEQ ID NOS:66中的一个或多个序列及其片段、突变体、变体或组合的序列一致性为至少75%。

16. 根据权利要求13所述的慢病毒表达载体组合物,其中,所述至少一种gRNA包含至少一个与靶标核酸序列互补的核酸序列,所述核酸序列包含SEQ ID NOS:1至SEQ ID NOS:66,及其片段、突变体、变体或组合。

17. 根据权利要求13所述的慢病毒表达载体组合物,其中,所述至少一种gRNA包含至少一种核酸序列,所述至少一种核酸序列包含SEQ ID NOS:1至SEQ ID NOS:66,及其片段、突变体、变体或组合。

18. 根据权利要求13所述的慢病毒表达载体组合物,其中,所述至少一种gRNA选自:gRNA A,其具有与前病毒DNA中的靶标序列SEQ ID NO:1或靶标序列SEQ ID NO:2互补的间隔序列序列;gRNA B,其具有与前病毒DNA中的靶标序列SEQ ID NO:3或靶标序列SEQ ID NO:4互补的间隔序列序列;或gRNA A与gRNA B的组合。

19. 根据权利要求13所述的慢病毒表达载体组合物,其中,所述CRISPR相关的核酸内切酶和所述至少一种gRNA被整合入单一慢病毒表达载体中。

20. 根据权利要求13所述的慢病毒表达载体组合物,其中,所述CRISPR相关的核酸内切酶和所述至少一种gRNA被整合入独立的慢病毒表达载体中。

21. 一种消除被整合入潜伏性感染人免疫缺陷病毒(HIV)的经离体培养的宿主细胞基因组内的前病毒DNA的方法,所述方法包括下列步骤:

获得潜伏性感染HIV的宿主细胞群,其中,前病毒HIV DNA被整合入所述宿主细胞基因组内;

离体培养所述宿主细胞;

以组合物处理所述宿主细胞,所述组合物包含成簇规律间隔短回文重复序列(CRISPR)相关的核酸内切酶、以及至少一种向导RNA(gRNA),所述至少一种gRNA具有与前病毒HIV

DNA的长末端重复序列(LTR)中的靶标序列互补的间隔序列序列;以及

从所述宿主细胞基因组消除所述前病毒DNA。

22. 根据权利要求21的方法,其中,所述获得宿主细胞群的步骤进一步定义为获得人类宿主细胞群。

23. 根据权利要求21的方法,其中,所述获得宿主细胞群的步骤进一步定义为获得人类外周血单核细胞、或获得CD4⁺T细胞群。

24. 根据权利要求21的方法,其中,所述至少一种gRNA包含与靶标核酸序列互补的核酸序列,所述核酸序列与SEQ ID NOS:1至SEQ ID NOS:66中的一个或多个序列及其片段、突变体、变体或组合的序列一致性为至少75%。

25. 根据权利要求21的方法,其中,所述至少一种gRNA包含一核酸序列,所述核酸序列与SEQ ID NOS:1至SEQ ID NOS:66中的一个或多个序列及其片段、突变体、变体或组合的序列一致性为至少75%。

26. 根据权利要求21的方法,其中,所述至少一种gRNA包含至少一个与靶标核酸序列互补的核酸序列,所述核酸序列包含SEQ ID NOS:1至SEQ ID NOS:66,及其片段、突变体、变体或组合。

27. 根据权利要求21的方法,其中,所述至少一种gRNA包含至少一种核酸序列,所述至少一种核酸序列包含SEQ ID NOS:1至SEQ ID NOS:66,及其片段、突变体、变体或组合。

28. 根据权利要求21的方法,其中,所述至少一种gRNA选自:gRNA A,其具有与前病毒DNA中的靶标序列SEQ ID NO:1或靶标序列SEQ ID NO:2互补的间隔序列序列;gRNA B,其具有与前病毒DNA中的靶标序列SEQ ID NO:3或靶标序列SEQ ID NO:4互补的间隔序列序列;或gRNA A与gRNA B的组合。

29. 根据权利要求21的方法,其中,所述处理步骤进一步包括在潜伏性感染的T细胞内表达所述CRISPR相关的核酸内切酶和所述至少一种向导RNA(gRNA)的步骤。

30. 一种治疗具有潜伏性人免疫缺陷病毒(HIV)感染的T细胞的患者的方法,所述方法包括下列步骤:

从所述患者获得包括潜伏性感染的T细胞的细胞群,其中,前病毒HIV DNA被整合入所述T细胞基因组内;

离体培养所述潜伏性感染的T细胞;

以组合物处理所述潜伏性感染的T细胞,所述组合物包含成簇规律间隔短回文重复序列(CRISPR)相关的核酸内切酶、以及至少一种向导RNA(gRNA),所述至少一种gRNA具有与前病毒HIV DNA的长末端重复序列(LTR)中的靶标序列互补的间隔序列序列;

从所述T细胞基因组消除被整合的前病毒HIV DNA;

产生HIV被消除的T细胞群;

将所述HIV被消除的T细胞群输注入所述患者体内;以及

处理所述患者。

31. 根据权利要求30的方法,其中,所述获得潜伏性感染的T细胞群的步骤进一步定义为获得人类外周血单核细胞群、或获得CD4⁺T细胞群。

32. 根据权利要求30的方法,其中,所述至少一种gRNA包含与靶标核酸序列互补的核酸序列,所述核酸序列与SEQ ID NOS:1至SEQ ID NOS:66中的一个或多个序列及其片段、突变

体、变体或组合的序列一致性为至少75%。

33. 根据权利要求30的方法,其中,所述至少一种gRNA包含一核酸序列,所述核酸序列与SEQ ID NOS:1至SEQ ID NOS:66中的一个或多个序列及其片段、突变体、变体或组合的序列一致性为至少75%。

34. 根据权利要求30的方法,其中,所述至少一种gRNA包含至少一个与靶标核酸序列互补的核酸序列,所述核酸序列包含SEQ ID NOS:1至SEQ ID NOS:66,及其片段、突变体、变体或组合。

35. 根据权利要求30的方法,其中,所述至少一种gRNA包含至少一种核酸序列,所述至少一种核酸序列包含SEQ ID NOS:1至SEQ ID NOS:66,及其片段、突变体、变体或组合。

36. 根据权利要求30的方法,其中,所述至少一种gRNA选自:gRNA A,其具有与前病毒DNA中的靶标序列SEQ ID NO:1或靶标序列SEQ ID NO:2互补的间隔序列序列;gRNA B,其具有与前病毒DNA中的靶标序列SEQ ID NO:3或靶标序列SEQ ID NO:4互补的间隔序列序列;或gRNA A与gRNA B的组合。

37. 根据权利要求30的方法,其中,所述处理步骤进一步包括在所述潜伏性感染的T细胞内表达所述CRISPR相关的核酸内切酶和所述至少一种向导RNA (gRNA) 的步骤。

38. 一种防止处于人免疫缺陷病毒 (HIV) 感染风险下的患者的T细胞的HIV感染的方法,所述方法包括下述步骤:

确定患者处于HIV感染的风险下;

令处于HIV 1感染风险下的患者的T细胞暴露于有效量的表达载体组合物,所述表达载体组合物包括编码成簇规律间隔短回文重复序列 (CRISPR) 相关的核酸内切酶的单离的核酸、以及至少一种编码至少一种向导RNA (gRNA) 的单离的核酸,所述gRNA包括与HIV 1 DNA的LTR中的靶标序列互补的间隔序列序列;

在所述T细胞中稳定地表达所述CRISPR相关的核酸内切酶和所述至少一种gRNA;以及防止所述T细胞的HIV感染。

39. 根据权利要求38的方法,其中,所述暴露所述T细胞的步骤进一步定义为将所述T细胞暴露于体内。

40. 根据权利要求38的方法,其中,所述暴露所述T细胞的步骤进一步定义为将所述T细胞离体暴露,且在所述稳定地表达的步骤后,进行将所述T细胞输注如所述患者体内的步骤。

41. 根据权利要求38的方法,其中,所述表达载体组合物是慢病毒载体组合物。

42. 根据权利要求38的方法,其中,所述至少一种gRNA包含与靶标核酸序列互补的核酸序列,所述核酸序列与SEQ ID NOS:1至SEQ ID NOS:66中的一个或多个序列或其组合的序列一致性为至少75%。

43. 根据权利要求38的方法,其中,所述至少一种gRNA包含一核酸序列,所述核酸序列与SEQ ID NOS:1至SEQ ID NOS:66中的一个或多个序列及其片段、突变体、变体或组合的序列一致性为至少75%。

44. 根据权利要求38所述的方法,其中,所述至少一种gRNA包含至少一个与靶标核酸序列互补的核酸序列,所述核酸序列包含SEQ ID NOS:1至SEQ ID NOS:66,及其片段、突变体、变体或组合。

45. 根据权利要求38的方法,其中,所述至少一种gRNA包至少一种核酸序列,所述至少一种核酸序列包含SEQ ID NOS:1至SEQ ID NOS:66,及其片段、突变体、变体或组合。

46. 根据权利要求38所述的方法,其中,所述至少一种gRNA选自:gRNA A,其具有与前病毒DNA中的靶标序列SEQ ID NO:1或靶标序列SEQ ID NO:2互补的间隔序列序列;gRNA B,其具有与前病毒DNA中的靶标序列SEQ ID NO:3或靶标序列SEQ ID NO:4互补的间隔序列序列;或gRNA A与gRNA B的组合。

47. 一种用于根除被整合入哺乳动物受试者细胞中的HIV-1 DNA的药物组合物,所述药物组合物包括编码成簇规律间隔短回文重复序列(CRISPR)相关的核酸内切酶的单离的核酸序列;至少一种编码至少一种向导RNA(gRNA)的单离的核酸序列,所述gRNA与前病毒HIV-1 DNA的长末端重复序列(LTR)中的靶标序列互补;所述单离的核酸序列被包括在至少一种表达载体内。

48. 根据权利要求47所述的药物组合物,其中,所述至少一种gRNA包含与靶标核酸序列互补的核酸序列,所述核酸序列与SEQ ID NOS:1至SEQ ID NOS:66中的一个或多个序列或其片段、变体、突变体或组合的序列一致性为至少75%。

49. 根据权利要求47所述的药物组合物,其中,所述至少一种gRNA包含一核酸序列,所述核酸序列与SEQ ID NOS:1至SEQ ID NOS:66中的一个或多个序列及其片段、突变体、变体或组合的序列一致性为至少75%。

50. 根据权利要求47所述的药物组合物,其中,所述至少一种gRNA包含至少一个与靶标核酸序列互补的核酸序列,所述核酸序列包含SEQ ID NOS:1至SEQ ID NOS:66,及其片段、突变体、变体或组合。

51. 根据权利要求47所述的药物组合物,其中,所述至少一种gRNA包含至少一种核酸序列,所述至少一种核酸序列包含SEQ ID NOS:1至SEQ ID NOS:66,及其片段、突变体、变体或组合。

52. 根据权利要求47所述的药物组合物,其中,所述至少一种gRNA选自:gRNA A,其具有与前病毒DNA中的靶标序列SEQ ID NO:1或靶标序列SEQ ID NO:2互补的间隔序列序列;gRNA B,其具有与前病毒DNA中的靶标序列SEQ ID NO:3或靶标序列SEQ ID NO:4互补的间隔序列序列;或gRNA A与gRNA B的组合。

53. 根据权利要求47所述的药物组合物,其中,所述表达载体是慢病毒载体。

54. 一种治疗感染HIV-1的哺乳动物受试者的方法,所述方法包括下列步骤:确定哺乳动物受试者感染HIV-1,向所述受试者给药有效量的根据权利要求47所述的药物组合物;以及治疗所述受试者的HIV-1感染。

55. 一种编码成簇规律间隔短回文重复序列(CRISPR)相关的核酸内切酶的单离的核酸,及/或至少一种编码至少一种向导RNA(gRNA)的单离的核酸,所述gRNA包括与前病毒HIV DNA的长末端重复序列(LTR)中的靶标序列互补的间隔序列序列,

所述CRISPR相关的核酸内切酶和所述至少一种gRNA被包括在至少一种表达载体内,

其中,所述至少一种表达载体诱导所述CRISPR相关的核酸内切酶和所述至少一种gRNA在宿主细胞内表达。

56. 根据权利要求55所述的单离的核酸序列,其中,所述至少一种gRNA包含与靶标核酸序列互补的核酸序列,所述核酸序列与SEQ ID NOS:1至SEQ ID NOS:66中的一个或多个序

列或其片段、变体、突变体或组合的序列一致性为至少75%。

57. 根据权利要求55所述的单离的核酸序列,其中,所述至少一种gRNA包含一核酸序列,所述核酸序列与SEQ ID NOS:1至SEQ ID NOS:66中的一个或多个序列及其片段、突变体、变体或组合的序列一致性为至少75%。

58. 根据权利要求55所述的单离的核酸序列,其中,所述至少一种gRNA包含至少一个与靶标核酸序列互补的核酸序列,所述核酸序列包含SEQ ID NOS:1至SEQ ID NOS:66,及其片段、突变体、变体或组合。

59. 根据权利要求55所述的单离的核酸序列,其中,所述至少一种gRNA包含至少一种核酸序列,所述至少一种核酸序列包含SEQ ID NOS:1至SEQ ID NOS:66,及其片段、突变体、变体或组合。

60. 根据权利要求55所述的单离的核酸序列,其中,所述至少一种gRNA选自:gRNA A,其具有与前病毒DNA中的靶标序列SEQ ID NO:1或靶标序列SEQ ID NO:2互补的间隔序列序列;gRNA B,其具有与前病毒DNA中的靶标序列SEQ ID NO:3或靶标序列SEQ ID NO:4互补的间隔序列序列;或gRNA A与gRNA B的组合。

61. 根据权利要求55所述的单离的核酸序列,其中,所述CRISPR相关的核酸内切酶和所述至少一种gRNA被整合入单一慢病毒表达载体内。

62. 根据权利要求55所述的单离的核酸序列,其中,所述CRISPR相关的核酸内切酶和所述至少一种gRNA被整合入独立的表达载体内。

63. 一种用于治疗或预防HIV-1感染的试剂盒,所述试剂盒包括测得量的组合物,所述组合物包含至少一种编码成簇规律间隔短回文重复序列(CRISPR)相关的核酸内切酶的单离的核酸序列、以及至少一种编码一种或多种向导RNA(gRNA)的核酸序列,其中,所述一种或多种gRNA各自包括与HIV-1前病毒的长末端重复序列(LTR)中的靶标序列互补的间隔序列序列;以及,一种或多种选自封装材料、包含使用说明书的封装插页、无菌液体、注射器、和无菌容器所组成群组的物品。

64. 根据权利要求63所述的试剂盒,其中,所述一种或多种gRNA选自:gRNA A,其具有与前病毒DNA中的靶标序列SEQ ID NO:1或靶标序列SEQ ID NO:2互补的间隔序列序列;gRNA B,其具有与前病毒DNA中的靶标序列SEQ ID NO:3或靶标序列SEQ ID NO:4互补的间隔序列序列;或gRNA A与gRNA B的组合。

65. 根据权利要求63所述的试剂盒,其中,所述单离的核酸序列中的至少一个序列被包括于表达载体内。

66. 根据权利要求65所述的试剂盒,其中,所述表达载体是慢病毒表达载体。

67. 一种包含一个或多个核酸序列的单离的核酸序列,所述一个或多个核酸序列与SEQ ID NOS:1至SEQ ID NOS:66的一个或多个序列或其片段、变体、突变体或组合的序列一致性为至少75%。

68. 一种核酸序列,其包含SEQ ID NOS:1至SEQ ID NOS:66的任意一个或多个序列,或其片段、变体、突变体或组合。

用于HIV感染的RNA引导治疗的方法和组合物

[0001] 关于联邦资助研究的声明

[0002] 本发明在美国政府支持下,由国立卫生研究院(NIH)资助Kamel Khalili (P30MH092177)、Wenhui Hu (R01NS087971)、以及Wenhui Hu 和Kamel Khalili (R01NS087971)完成。美国政府可以拥有本发明的某些权利。

技术领域

[0003] 本发明涉及用于逆转录病毒如人免疫缺陷病毒(HIV-1)中靶标序列的特异性裂解的组合物和方法。所述组合物可包括编码成簇规律间隔短回文重复序列(CRISPR)相关的核酸内切酶的核酸、以及与人免疫缺陷病毒中靶标序列互补的向导RNA序列,据信所述组合物可被输送至具有HIV感染或处于接触HIV感染风险下的受试者的细胞。

背景技术

[0004] 由于全世界内超过3,500万人感染HIV-1,且新感染以大于200 万人每年的速率继续,因此,AIDS仍然是主要的公众卫生问题。抗逆转录病毒治疗(ART)有效地控制几乎所有HIV-1患者的病毒血症,并部分地修复原代宿主细胞(CD4⁺T细胞),但无法成功地从潜伏性感染的T 细胞消除HIV-1(Gandhi,et al,PLoS Med 7,e1000321(2010); Palella et al,N Engl J Med 338,853-860(1998))。在潜伏性感染的CD4⁺T细胞中,被整合的前病毒DNA拷贝保持在休眠状态,但当T 细胞被激活时,所述拷贝可被再次激活以产生能够复制的病毒,一旦抗逆转录病毒治疗中断,即导致病毒的快速回弹(Chun,et al.,Nature 387,183-188(1997);Chun,et al,Proc Natl Acad Sci USA 100, 1908-1913(2003);Finzi,et al., Science 278,1295-1300(1997); Hermankova,et al.,J Virol 77,7388-7392(2003); Siliciano,et al.,Nat Med 9,727-728(2003);Wong,et al,Science 278, 1291-1295(1997))。因此,大多数感染HIV-1的个体,甚至是那些对 ART应答良好的患者,由于感染HIV-1的储存细胞的持续存在,而必须维持终身的ART。在潜伏期,HIV感染的细胞产生很少或不产生病毒蛋白,因此避免了病毒的细胞病理效应并规避了被宿主免疫系统清除的命运。因为休眠的CD4⁺记忆T细胞隔室(Bruner,et al.,Trends Microbiol.23,192-203(2015))被认为是最卓越的潜伏性感染的细胞池,这一细胞隔室是以根除潜伏性HIV-1感染为目标的研究的关键焦点。

[0005] 最近,用来从这一细胞群中根除HIV-1的尝试已经主要使用“休克并杀死”途径,基本原理为,诱导CD4⁺记忆T细胞中的HIV再次激活,可通过细胞溶解或宿主免疫应答触发病毒生产细胞的消除。例如,对于建立病毒再激活而言,核染色质结构的后天性修饰是关键。因此,通过曲古抑菌素A(Trichostatin A(TSA))和伏立诺他(vorinostat) (SAHA)抑制组蛋白脱乙酰化酶(HDAC),导致细胞系中潜伏病毒的再激活(Quivy,et al.,J Virol 76, 11091-11093(2002);Pearson,et al,J Virol 82,12291-12303(2008);Friedman,et al,J Virol 85, 9078-9089(2011))。据此,已经通过离体方式测试了其它HDACi,包括伏立诺他、丙戊酸、帕比司他(panobinostat)和罗米地辛(romidepsin),在最佳案例中,所述HDACi导

致病毒血症的短暂增长 (Archin, et al., Nature 487, 482-485 (2012); Blazkova, et al., J. Infect. Dis 206, 765-769 (2012))。同样, 蛋白激酶C激动剂单独使用或与HDACi合用, 可潜在地再激活HIV (Laird, et al., J Clin Invest, 125, 1901-1912 (2015); Bullen, et al., Nature Med 20:425-429 (2014))。但是, 这一途径存在多种限制: i) 由于这一蓄积池中的大部分HIV基因组是非功能性的, 并非所有被整合的前病毒均能产生可复制的病毒 (Ho, et al., Cell 155, 540-551 (2013)); ii) 通过病毒产物检验已经发现, 从休眠的CD4⁺T细胞HIV-1蓄积池再次激活的CD4⁺T细胞的总数比通过基于PCR的检验检测到的被感染的细胞数小得多, 表明并非这一蓄积池中的所有细胞均被激活 (Eriksson, et al., PLoS Pathog 9, e1003174 (2013)); iii) 细胞毒素T淋巴细胞(CTL)免疫应答的强健度不足以消除被再次激活的被感染细胞 (Shan, et al., Immunity 36, 491-501 (2012)); 以及iv) 不能保护未感染的T细胞不被HIVgarnet, 因此未感染的T细胞支持病毒重建。

[0006] 已经显示, 成簇规律间隔短回文重复序列 (CRISPR) 相关的9 (Cas9) 核酸酶系统可用于宽范围的有机体内的基因编辑中, 所述有机体包括酵母、果蝇属 (*Drosophila*)、斑马鱼、线虫 (*C. elegans*) 和鼠, 且所述系统已经被多个实验室广泛用于指向人类疾病的体内外研究中 (Di Carlo et al., Nucl Acids Res 41:4336-4346 (2013); Gratz et al., Genetics 194, 1029-1035 (2013); Hwang et al., Nature Biotech 31, 227-229, (2013); Wang et al., 2013; Hu, et al., Proc Natl Acad Sci USA 111, 11461-11466 (2014))。在CRISPR/Cas9系统中, 组装了基因编辑复合物。每一复合物包括Cas9核酸酶和与前病毒DNA中的靶标序列互补的向导RNA (gRNA)。所述gRNA引导所述Cas9核酸酶啮合并裂解含有所述靶标序列的前病毒DNA链。所述Cas9/gRNA基因编辑复合物将一个或多个突变引入所述病毒DNA中。

[0007] 最近, 所述CRISPR/Cas9系统已经被修饰以令识别位于HIV-1长末端重复 (LTR) 序列内的具体DNA序列称为可能 (Hu, et al., Proc Natl Acad Sci USA 111, 11461-11466 (2014); Khalili et al., J Neurovirol 21, 310-321 (2015))。需要扩展现有CRISPR/Cas9介导疗法的治疗能力, 以包括从潜伏性感染患者T细胞根除被整合的HIV-1 DNA; 以及需要扩展诱导对处于感染风险下的患者T细胞内的HIV-1感染的抗性的能力。

发明内容

[0008] 用于人免疫缺陷病毒 (HIV) 感染的治愈性策略包括, 直接消除包括 CD4⁺T细胞在内的HIV阳性细胞中前病毒基因组的方法, 所述方法即便对宿主有害, 但益处有限。一种具体实施例中, 本发明提供用于治疗和预防逆转录病毒, 尤其是人免疫缺陷病毒HIV-1感染的组合物和方法。所述组合物和方法采用形成复合物的Cas9和至少一种gRNA, 在大多数案例中, 所述复合物消除宿主T细胞基因组中的前病毒HIV。优选的具体实施例中, 包括至少两种gRNA, 且各gRNA将CRISPR相关的核酸内切酶引导至HIV基因组的LTR中的不同靶标位点。

[0009] 具体地, 本发明提供Cas9/gRNA组合物, 所述组合物用于令被整合入被HIV潜伏性感染的宿主细胞基因组内的前病毒DNA失活。本发明还提供一种方法, 所述方法采用Cas9/gRNA组合物令宿主细胞中的前病毒HIV DNA失活。

[0010] 本发明进一步提供编码Cas9和至少一种gRNA的慢病毒载体, 用于令整合入被HIV潜伏性感染的宿主细胞基因组内的前病毒DNA失活。

[0011] 本发明还提供一种离体方法, 用于消除被HIV潜伏性感染的T细胞基因组内的前病

毒DNA。所述方法包括下述步骤：获得被HIV潜伏性感染的宿主细胞群，如AIDS患者的原代T细胞；离体培养所述宿主细胞；使用Cas9核酸内切酶和至少一种gRNA处理所述宿主细胞；以及，从所述宿主细胞基因组消除前病毒DNA。

[0012] 本发明再进一步提供治疗具有HIV感染的T细胞的患者的方法。所述方法包括实施下述步骤：如先前阐明的离体治疗方法；产生HIV 被消除的T细胞群；以及，将HIV被消除的T细胞群送回所述患者体内。

[0013] 本发明还提供用于令哺乳动物受试者细胞内被整合的HIV DNA失活的Cas9/gRNA药物组合物。

[0014] 本发明进一步提供一种方法，所述方法通过给药有效量的先前阐明的药物组合物，以治疗被HIV感染的哺乳动物受试者。

[0015] 本发明再进一步提供一种方法，用于预防处于HIV感染风险下患者的T细胞HIV感染。所述方法包括或离体或体内建立Cas9和gRNA 在患者T细胞内的稳定表达。

[0016] 本发明还提供试剂盒，便于先前阐明的治疗或预防HIV感染的方法的应用。

附图说明

[0017] 图1A至1D显示，CRISPR/Cas9消除了经PMA/TSA处理的潜伏性感染的人类T细胞系中的HIV-1表达。图1A：上半部分是被整合的HIV-1 前病毒DNA的结构组织的示意图，高亮所述长末端重复序列(LTR)、被所述LTR跨越的多种病毒基因、以及报告基因d2EGFP的定位。图1A 的下半部分是用于编辑的5' -LTR及靶标区域A(gRNA A)和B(gRNA B) 的核苷酸序列、以及用于结合多种转录因子的基序的例示性说明。位于+1的箭头描述转录起始位点。图1B是PMA/TSA治疗前和治疗后的 CD4⁺T细胞的EGFP流式细胞术和荧光显微成像的门控图，在仅表达Cas9 的对照细胞中显示PMA/TSA诱导的潜伏病毒再激活，但在表达Cas9和 gRNA两者的细胞中未显示再激活。图1C：对以表达Cas9±gRNA的质粒转染的细胞中的gRNA A、gRNA B和β-肌动蛋白RNA的基于RT-PCR 的检测。β-肌动蛋白是RNA负载对照。图1D：通过西方墨点分析检测对照细胞和HIV-1/EGFP表达被切除的细胞中的Cas9蛋白，β-微管蛋白用作蛋白质负载对照。

[0018] 图2A至2D显示，通过以病毒LTR为靶向的Cas9/gRNA，从宿主T 细胞基因组消除所整合的HIV-1 DNA。图2A：DNA分析显示，检测到 497核苷酸扩增子和504核苷酸扩增子，分别对应于对照细胞和共表达 Cas9及gRNA的细胞中的HIV-1 LTR。显示了对应于RRE和β-肌动蛋白的扩增子的位置。图2B：从以CRISPR/Cas9处理的细胞扩增的LTR DNA 的核苷酸组合物，以及用于所述病毒基因组多个区域的PCR扩增的引物的位置。显示了在移除所述病毒DNA片段后，7个核苷酸的插入缺失标记(InDel) 突变整合于5' -LTR的B基序与3' -LTR的B基序之间。用于gRNA B的种子序列以黑色强调。图2C和2D显示HIV-1整合入染色体1(图2C)和染色体16(图2D)中的位点。各图中显示了对PCR产物的DNA分析结果，所述PCR产物是通过源自被病毒DNA插入打断的细胞基因的特异性引物(P1和P2)而扩增的。基于琼脂糖凝胶上所见的DNA片段的桑格定序(Sanger sequencing)，所述图描述了在以 CRISPR/Cas9处理前含有全长度整合的HIV-1 DNA的各染色体，和以 Cas9/gRNA处理后的残留LTR DNA序列。图2C和2D中的星号指出次要DNA谱带，所述谱带表明，当使用5' -LTR或3' -LTR内的A靶标或B靶标时，病毒DNA被完全移除。

[0019] 图3A至3F描述全基因组测序的记过,其显示通过Cas9/gRNA和 gRNA A及gRNA B在人类T细胞中切除两拷贝HIV-1的前病毒。图3A 和3B是映射至名为BWA的HIV-1基因组(KM390026.1)的解读的整合基因组图,显示所述HIV-1前病毒DNA序列存在于表达Cas9但不表达 gRNA的对照细胞中(图3 A),但它们在表达Cas9和gRNA A及gRNA B 后的T细胞中完全不存在(图3B)。图3C和3D是结构变体CREST分析,鉴别位于由所表明的Cas9/gRNA A/B裂解后的解读支持的两个LTR的 5'端和3'端的两个断裂点。例示性说明了映射至HIV-1基因组(KM390026.1)的解读的整合基因组图(IGV)。图3E:通过名为CREST 的结构变体鉴别位于9389位点(红色箭头)的gRNA(图3A、3B)特异性断裂点。数值的紫色线指出,5'和3'-LTR在断裂后剩余的部分结合的位置。图3F:在移除精确切割位点之间的核苷酸,即通过gRNA A 从靶标A处5'-LTR的PAM(红色箭头)移除和通过gRNA B从靶标B处 3'-LTR的PAM(红色箭头)移除所述核苷酸后,位于连接位点(红色箭头)处的DNA序列的例示性说明。

[0020] 图4A至4E显示HIV-1导向的基因编辑策略对宿主基因的影响。图4A是染色体1的示意图,强调HIV-1前病毒DNA整合入细胞基因 RSNB1的位点,以及邻近基因的位置。图4B是在通过Cas9/gRNA切除病毒DNA之前和之后,位于接近前病毒整合位点的多个位置的基因的表达。通过逆转录和qPCR鉴别基因的表达,并将所述值归一化至 β -肌动蛋白转录本。图4C是染色体16节段的线性结构组织,例示性说明作为HIV-1 DNA整合位点的MSRB1的位置,以及作为病毒DNA插入处的MSRB1外显子2的核苷酸结构。显示了集中细胞基因接近MSRB1 的位置。图4D是SyberGreen qPCR的结果,例示性说明在HIV-1 DNA 根除之前和DNA根除之后,细胞中的MSRB1表达及其邻近基因表达。所述表显示用于从5个独立的对照细胞系和5个独立的HIV-1被根除的单细胞克隆体获得的各细胞基因转录本的靶标/参考。图4E是通过全基因组测序和生物信息学解释进行的脱靶评估。图解演示预测的脱靶位点的位置,在HIV-1 DNA被切除的T-细胞中,在所扩张的位于所筛选InDel位点侧翼的30、300和600bp内具有3至7个核苷酸错配。脱靶序列旁边的数字表明所述1200bp扩张序列的核苷酸数。在gRNA A 脱靶位点(蓝色)以绿色强调错配的核苷酸,而在gRNA B脱靶位点(紫色)以橙色强调错配的核苷酸。所述PAM序列加红色下划线。注意,脱靶定位原理所述InDel的位置,且在所预测的从PAM起始的第三核苷酸处未展现突变。

[0021] 图5A至5E:慢病毒(LV)介导的Cas9/gRNA递送阻抑了人类T细胞内的HIV-1感染。图5A:对以表达gRNA A/B、Cas9、或Cas9和gRNA A/B两者的LV处理的2D10 T细胞的PCR片段分析。显示了全长度扩增子(417bp)和切除位于gRNA A与gRNA B之间的190bp后的较小DNA 片段(227bp)的位置。显示270bp β -肌动蛋白DNA片段的扩增作为对照。图5B:表达GFP(HIV-1)和RFP(Cas9)的细胞的代表性散点图,表明在LV感染后,72.9%的2D10细胞表达Cas9,在以PMA/TSA诱导后,超过45%的这些细胞(31.8%)显示无GFP表达的证据,所述GFP表达是HIV-1 DNA消除的指征。图5C是原代CD4⁺T细胞中体外感染实验的实验过程安排。通过在磁性柱(Miltenyi Biotec)上进行阴性选择,将 CD4⁺T细胞从新鲜制备的、抗体标记的PBMC分离出来,随后以48小时的抗CD2/CD3/CD28处理进行激活,之后进行6天的人rIL-2介导扩张。之后,通过接种以HIV-1感染细胞,2天后,以含有lenti-Cas9且具有或不具有lenti-gRNA LTR A/B的慢病毒鸡尾酒混合物转导。4天后,收获细胞上清液和细胞,并分析HIV-1的存在。图5D:如实验过程中所述,以HIV-1_{JRFL}或HIV-1_{NL4-3}感染从血沉棕黄层新鲜分离的PBMC制备的CD4⁺T细胞,通过TaqMan qPCR测定HIV-1拷贝数,并归一化至 β -球蛋白基因拷贝数。感染

6天后,HIV-1_{JRFL}的拷贝数显著下降(48%),当将LV-Cas9/gRNA表达与那些接受LV-Cas9的相比时,甚至观察到了 HIV-1L4-3的更急剧减少。图5E:来自在LV-gRNA A/B的存在或不存在下以LV-Cas9治疗的HIV-1感染CD4⁺T细胞的LTR和β-肌动蛋白 DNA(对照)的PCR分析。显示398bp HIV-1 LTR和270bpβ-肌动蛋白扩增子的位置。

[0022] 图6A至6F显示,对HIV-1感染的患者的外周血单核细胞(PBMC)和CD4⁺T细胞中HIV-1复制的阻抑。图6A:以LV-Cas9或LV-Cas9加上LV-gRNA A/B(在材料与方法中揭示)处理来自两位HIV-1感染的志愿者的PBMC,并通过qPCR测定DNA拷贝。显然,在归一化至β-球蛋白DNA后,检测到病毒拷贝数的实质性减少。图6B:在含有人类IL-2(20U/ml)的介质中扩展从PBMC单离的CD4⁺T细胞,以LV-Cas9或LV-Cas9加上LV-gRNA A/B感染,4天后,通过qPCR测定病毒DNA拷贝数。与PBMC类似,在接受LV-Cas9/gRNA的细胞中观察到了比接受LV-Cas9的细胞剧烈下降的HIV-1 DNA拷贝数。图6C:收获使用表达Cas9或表达Cas9和gRNA A/B的慢病毒载体处理的CD4⁺T细胞,通过p24 Gag ELISA测定病毒复制。图6D:从患者样品单离的DNA经过慢病毒处理后,通过扩展-374/+43的引物进行PCR分析。显示417预期扩增子的位置。对照表示在感染第6天,来自HIV-1_{JRFL}感染的LTR DNA的扩增。图6E:使用跨越LTR的-416至-19的不同成套引物进行病毒LTR的PCR扩增(如图6D中所示)。显示了398bp扩增子的位置。图6F:来自患者2的LTR片段(显示于图6E中)的TA克隆和测序,在一些扩增的DNA中显示插入、删除和单核苷酸变异(SNV)。注意,所述检验方法不能检测大DNA消除,大DNA消除需要源自所述病毒基因组外部即位于整合位点侧翼的引物。

[0023] 图7A显示对以表达Cas9的质粒或表达Cas9加上gRNA的质粒转染的几种2D10克隆体的流式细胞术评估。以PMA/TSA处理所述+Cas9/-gRNA细胞,在71%至89%的细胞中诱导了HIV-1表达(GFP⁺)。相反,以Cas9和gRNA转染的细胞显示对于所述处理无明显应答(1%至3%)。图7B显示RT-PCR检验的结果,用于检测几种克隆2D10细胞在根除其潜伏的HIV-1基因组后的gRNA A和gRNA B,β-肌动蛋白mRNA水平用作RNA制备完整性和负载的对照。C11表示对照(+Cas9/-gRNA)细胞的RT-PCR。

[0024] 图8A和8B显示怀有整合的HIV-1前病毒DNA拷贝的人类T细胞的全基因组测序和生物学分析。图8A:HIV-1整合在染色体1和染色体16上的核苷酸水平上的细节显示于右侧。宿主染色体DNA序列以红色显示,整合的DNA序列以黑色显示。四个被删除的核苷酸(TAAG)加绿色下划线。基于称为结构变体的CREST,鉴别四个与HIV-1相关的染色体间易位(CTX)。图8B:通过NCBI、BLASTIN分析的染色体1和染色体16的示意图,强调HIV-1基因组与宿主染色体之间的相关性(LTR,长末端重复序列)。

[0025] 图9A和9B显示染色体1的描述HIV-1 DNA被整合处的RSNB1内的区域的部分的DNA测序结果。图9A是PAM的位置,以及与gRNA A和gRNA B向对应的LTR的核苷酸序列。强调LTR A和B。图9B:PCR片段的DNA测序,显示断裂点的精确位置和位于PAM下游3核苷酸处的7核苷酸插入。

[0026] 图10A:染色体16中宿主DNA的DNA测序,例示性说明HIV-1 DNA整合入MRSB1基因中的精确位点,并强调InDel突变的区域。图10B:在通过gRNA A裂解后,5'-LTR中的8核苷酸插入的位置(位于LTR A靶标),以及通过gRNA B裂解时,3核苷酸插入的位置(位于LTR B靶标)。

[0027] 图11A和11B显示细胞凋亡检验的结果,所述检验用来评估用于根除HIV-1的Cas9/

gRNA对于细胞凋亡的影响。图11B是柱状图,显示对14种仅以Cas9慢病毒感染且不经gRNA感染的T细胞克隆体实施的细胞凋亡检验的平均结果。对于每一样品,所述实验平行实施三次,数据表示为平均值和标准偏差。不同的颜色表示在不同的细胞凋亡阶段检测到的平均细胞百分比,如图下方表中所示。图11A的左图显示代表性样品的结果。图11B是柱状图,其显示对先前已经通过以Cas9慢病毒和gRNA慢病毒感染T细胞而根除细胞内HIV-1的T细胞克隆体实施相同的细胞凋亡检验的结果。图11B的左图显示代表性样品的结果。所述结果显示,以Cas9感染的克隆体与被根除的克隆体之间没有显著差异,表明gRNA并无影响细胞凋亡的细胞机制。

[0028] 图12A和12B显示细胞活力检验的结果,所述检验被用来探究所研发的用于根除HIV-1的Cas9/gRNA对细胞活力的影响。图12A是柱状图,显示对仅使用Cas9慢病毒感染而不使用gRNA感染的14种T细胞克隆体实施的细胞活力检验的平均结果。对于每一样品,所述实验平行进行三次,数据表示为平均值和标准偏差。生存细胞和死亡细胞的平均百分比分别表示为蓝色和红色。图12A的左图显示代表性样品的结果。图12B是柱状图,显示对先前已经通过以Cas9慢病毒和gRNA慢病毒感染T细胞而根除细胞内HIV-1的T细胞克隆体实施相同的细胞活力检验的结果。图12B的左图显示代表性样品的结果。结果显示,以Cas9感染的克隆体与被根除的克隆体之间没有显著差异,表明gRNA慢病毒并不诱导细胞死亡。

[0029] 图13A和13B显示细胞周期检验的结果,所述检验被用来探究所研发的用于根除HIV-1的Cas9/gRNA对细胞周期的影响。图13B是柱状图,显示对仅使用Cas9慢病毒感染而不使用gRNA感染的14种T细胞克隆体实施的细胞周期检验的平均结果。对于每一样品,所述实验平行进行三次,数据表示为平均值和标准偏差。在不同细胞周期阶段中检测到的细胞的百分比显示为不同颜色,如柱状图下方表格中所示。图13A的左图显示代表性样品的结果。图13B是柱状图,显示对先前已经通过以Cas9慢病毒和gRNA慢病毒感染T细胞而根除细胞内HIV-1的T细胞克隆体实施相同的细胞周期检验的结果。图13B的左图显示代表性样品的结果。结果显示,以Cas9感染的克隆体与被根除的克隆体之间没有显著差异,表明gRNA慢病毒并不影响细胞周期机制。

[0030] 图14是显示染色体的平均深度(左侧坐标)和覆盖率(右侧坐标)的图。X轴是染色体编号;左侧Y轴是每一染色体的平均深度,右侧Y轴是覆盖在每一染色体上的分数。

[0031] 图15A至15C显示保护HIV-1切除的T细胞系不再感染。图15A:通过西方墨点法检查集中潜伏性感染的T细胞在清除其HIV-1基因组后的Cas9表达(上图),并通过RT-PCR检查gRNA B的存在(中图)。α-微管蛋白和β-肌动蛋白的表达分别用作蛋白质和RNA的负载对照。图15B:以HIV-1感染具有Cas9及/或gRNA表达的T细胞,在感染后多个时间点,通过流式细胞术测定每一案例中的病毒感染。图15C显示图15B中所示实验的定量值。

[0032] 图16A和16B显示源自患者的原代PBMC和CD4⁺T细胞实验的结果。图16A:通过CNAC基础科学核心1(Temple University, Philadelphia)获得来自进行ART的四位HIV-1阳性患者的血液样品。AA:非裔美国人,His:西班牙裔。图16B是用于患者血液样品的实验流程的示意图。通过在磁性柱(Miltenyi Biotec)上的阴性选择从新鲜制备的、抗体标记的PBMC单离CD4⁺T细胞,随后通过48小时的抗CD2/CD3/CD28处理激活所述细胞,之后,进行6天的人rIL-2介导的扩张。同时,将来自相同血液样品的PBMC进行PHA激活,并使用人rIL-2进行类似的扩张。此后,以含有lenti-Cas9且具有或不具有lenti-gRNA LTR A/B的慢病毒鸡尾酒

混合物转导所述细胞。4天后,收获上清液和细胞,并分析其中是否存在HIV-1。图16C:通过FITC接合抗CD4抗体标记细胞的流式细胞术核对单离后的CD4⁺T细胞的纯度。显示单离后在CD4⁻去除的细胞群和CD4富集的细胞群中CD4阳性(GFP通道)细胞的代表性直方图。

[0033] 图17是显示源自患者的PBMC中HIV-1水平的图。显示对来自案例3和4的PBMC在以慢病毒Cas9感染或以Cas9加上慢病毒gRNA A和gRNA B感染后的p24ELISA检验结果。以细胞:微珠比为2:21的包覆抗CD2、CD3和CD28的微珠(Miltenyi Biotec)或PMA/TSA鸡尾酒混合物(PMA 25nM/TSA 250nM)处理细胞48小时,随后计数并测量上清液中的Gag p24。

[0034] 图18A至18B是用于每一样品中人 β -球蛋白(图18A、18B)和 HIV-1Gag(图18C、18D)基因拷贝数的绝对定量的扩增图谱和标准曲线。从对应于10 μ l/反应中10⁵基因组拷贝的3.3 μ g/ml开始,到对应于10 μ l/反应中10基因组拷贝的0.33ng/ml为止,制备从U1获得的单核细胞系的基因组DNA的连续稀释液。U1细胞含有整合入染色体2和染色体X的HIV-1前病毒的每基因组2个单一的、全长度拷贝,等于 β -球蛋白基因拷贝(每二倍体基因组2个)。

具体实施方式

[0035] 根据本发明的CRISPR-Cas9系统包括至少一种组合的基因编辑复合物,所述复合物包含CRISPR相关的核酸酶如Cas9,以及与位于被整合入哺乳动物基因组中的HIV前病毒DNA链上的靶标序列互补的向导RNA。每一基因标记复合物可裂解所述靶标序列内的DNA,造成令前病毒基因组失活的删除和其它突变。优选的具体实施例中,所述向导RNA与出现在所述HIV前病毒的两个LTR区域中每一个内的靶标序列互补。某些具体实施例中,所述gRNA与LTR的U3区域内的位点互补。其它具体实施例中,所述gRNA包括gRNA A,其与图1A中指定为“gRNA A”的区域中的靶标序列互补;以及gRNA B,其与图1A中指定为“gRNA B”的区域中的靶标序列互补。一种优选的具体实施例中,gRNA A与gRNA B两者以成对(“双链”)方式组合。

[0036] 定义

[0037] 除非另做定义,本为中使用的科技术语具有与本发明所属领域技术人员一般理解者相同的意义。尽管可使用与本文中揭示的那些相似或等效的任何方法和材料来实践对本发明的测试,优选本文中揭示的材料和方法。为了解释和主张本发明,将使用下述术语。亦应理解,本文中使用的术语仅用于揭示特定具体实施例的目的,而非意图限制。

[0038] 本文中使用的冠词“一(a)”和“一(an)”指的是一个或超过一个(即,至少一个)的所述词语法宾语。举例而言,“一元件”意为一个元件或超过一个元件。因此,举例而言,“一细胞”的说法包括相同类型的多个细胞。此外,就详细说明书及/或权利要求书中使用的术语“包括(including)”、“包括(includes)”、“具有(having)”、“具有(has)”、“具有(with)”或其变体而言,这些术语意图为以类似于术语“包含”的方式包容在内。

[0039] 如本文中所使用,术语“包含”或其变体关于定义或揭示项目、组合物、仪器、方法、过程、系统等的元件而使用时,意为开放性的,允许存在其它的元件,从而表明所定义或揭示的项目、组合物、仪器、方法、过程、系统包括那些具体指明的元件,或视情况而定,其等效物,且可包括其它元件,而仍落入所定义的项目、组合物、仪器、方法、过程、系统等的范畴/定义内。

[0040] 如本文中所使用,当“约”指代可测量的值如量、时距等时,意为涵盖从所指定数值

计的 $\pm 20\%$ 、 $\pm 10\%$ 、 $\pm 5\%$ 、 $\pm 1\%$ 、或 $\pm 0.1\%$ 的变动,而这些变动对于实施所揭露的方法是适宜的。或者,特别是关于生物学系统和过程,所述术语可意为处于5倍数量级内,亦可意为处于2倍数量级内。若特定的值揭示于本申请和权利要求书中,除非另做说明,应假定术语“约”意为处于所述特定值的可接受的误差范围内。

[0041] 如本文中所以使用,术语逆转录病毒如人免疫缺陷病毒(HIV)的“根除”,意为所述病毒不能复制,所述基因组被删除、片段化、降解、遗传性灭活、或任何其它物理学、生物学、化学或结构性表现形式,其防止所述病毒成为可传递的或感染任何其它细胞或受试者,导致所述病毒的体内清除。一些情况下,所述病毒基因组的片段可能是可检测的,但所述病毒不能复制、或感染等。

[0042] 如本文中所以使用,“有效量”意为提供治疗性或预防性益处的量。

[0043] “编码”指的是多核苷酸中特定核苷酸序列如基因、cDNA或mRNA的用作在生物学过程中合成其它聚合物和大分子的模板的固有性质,所述其它聚合物和大分子具有所定义的核苷酸序列(即,rRNA、tRNA和mRNA)或所定义的氨基酸序列以及自其所得的生物学性质。因此,若对应于一基因的mRNA的转录和转译在细胞或其它生物学系统中生产蛋白质,则所述基因编码蛋白质。作为与mRNA序列一致且通常提供在序列列表中的核苷酸序列的编码链,以及用作基因或cDNA转录模板的非编码链,两者均可指代为编码所述蛋白质或所述基因或cDNA的其它产物。

[0044] 如本文中所以使用,术语“表达”定义为通过其启动子驱动的特定核苷酸序列的转录及/或转译。

[0045] “表达载体”指代是包含重组多核苷酸的载体,所述重组多核苷酸包含可操作地链接至待表达核苷酸序列的表达控制序列。表达载体包含足够的用于表达的顺式作用元件;用于表达的其它元件可由宿主细胞供给或在体外表达系统中供给。表达载体包括所有所述领域已知的表达载体,如合并所述重组多核苷酸的粘粒、质粒(如,裸质粒或包含在脂质体中的质粒)和病毒(如,慢病毒、逆转录病毒、腺病毒、及腺相关病毒)。

[0046] “单离的”意为从天然状态改变或移除。例如,天然存在于活体动物体内的核酸或肽不是“单离的”,但部分地或完全地与其天然状态下共存材料单离的同一核酸或肽是“单离的”。单离的核酸或蛋白质可以大体上纯的形式存在,或可存在于非原生环境如宿主细胞中。

[0047] “单离的核酸”指的是已经与天然出现状态中位于其侧翼的序列单离的核酸片段或片段,即已经从其正常与所述片段相邻的序列移除的DNA片段,即在其天然出现的基因组中与所述片段相邻的序列。所述术语也用于已经大体上从天然伴生所述核酸的其它组分纯化的核酸,所述其它组分为在细胞中与所述核酸天然伴生的RNA或DNA或蛋白质。所述术语因此包括,例如,重组DNA,其被并入载体中、并入自主复制质粒或病毒中、或并入原核生物或真核生物的基因组DNA中,或作为不依赖其它序列的独立分子(即,作为通过PCR或限制酶消化产生的cDNA或基因组或cDNA片段)而存在。它还包括:作为编码额外多肽序列的杂交基因的一部分的重组DNA、互补DNA(cDNA)、天然及/或修饰单体或链接的线性或环状寡聚物或聚合物,包括脱氧核糖核苷酸、核糖核苷酸、其经取代的形式和 α -端基异构形式、肽核酸(PNA)、锁定核酸(LNA)、硫代磷酸酯、甲基磷酸酯等。

[0048] 所述核酸序列可以是“嵌合的”,换言之,由不同的区域组成。在本发明的语境中,

“嵌合”化合物是寡核苷酸,其含有两个或更多个化学区域,例如,DNA区域、RNA区域、PNA区域等。每一化学区域由至少一种单体单元即核苷酸构成。这些序列典型包含至少一个区域,在所述区域中,所述序列被修饰以展现一种或多种所希望的性质。

[0049] 术语“靶标核酸”序列指一种核酸(一般源自生物学样品),所述寡核苷酸被设计为特异性杂交至所述核酸。靶标核酸具有与相应的导向所述靶标的核酸序列互补的序列。术语“靶标核酸”可指代所述寡核苷酸被导向的较大核酸的特定子序列或指代整个序列(如,基因或 mRNA)。用法的差异将自语境中明显可见。

[0050] 在本发明的语境中,使用下述用于常见核酸碱基的缩写,“A”指腺嘌呤,“C”指胞嘧啶,“G”指鸟嘌呤,“T”指胸腺嘧啶,而“U”指尿嘧啶。

[0051] 除非另有说明,“编码氨基酸序列的核苷酸序列”包括作为彼此退化版本并编码相同氨基酸序列的所有核苷酸序列。短语“编码蛋白质或RNA的核苷酸序列”也可包括内含子,包括程度为所述编码蛋白质的核苷酸序列可在某些版本中含有内含子。

[0052] 免疫原性组合物的“肠胃外”给药包括皮下注射(s.c)、静脉注射(i.v.)、肌肉注射(i.m.)、胸骨注射、或输注技术。

[0053] 术语“患者”或“个体”或“受试者”在本文中可互换使用,且指代待治疗的哺乳动物受试者,以人力患者为优选。一些情况下,本发明的方法可用于实验动物、兽医应用、以及用于疾病的动物模型的研发中,所述动物包括但不限于,啮齿动物,包括小鼠、大鼠和仓鼠,以及灵长类。

[0054] 术语“多核苷酸”是核苷酸的链,也称为“核酸”。本文中,多核苷酸包括,但不限于,通过所述领域中可获得的任何手段获得的所有核酸序列,且包括天然出现的核酸和合成的核酸。

[0055] 术语“肽”、“多肽”、和“蛋白质”可互换使用,且指代由通过肽键共价链接的氨基酸残基构成的化合物。蛋白质或肽必须含有至少两个氨基酸,且蛋白质序列或肽序列可包含的氨基酸的最大数目并无限制。多肽包括任何包含通过肽键彼此连接的两个或更多氨基酸的肽或蛋白质。本文中,所述术语指代多种类型的短链和长链两者,短链在所述领域中一般也称为肽、寡肽和寡核苷酸,而长链在所述领域中通常称为蛋白质。“多肽”包括,例如,生物学活性片段、基本上同源的多肽、寡肽、同源二聚体、异源二聚体、多肽的变体、修饰多肽、衍生物、类似物、融合蛋白等。多肽包括天然肽、重组肽、合成肽、或其组合。

[0056] 术语“经转染的”或“经转化的”或“经转导的”意为将外源性核酸转移至或引入宿主细胞内的过程。“经转染的”或“经转化的”或“经转导的”细胞是已经使用外源性核酸转染、转化或转化的细胞。经转染/转化/转化的细胞包括原代受试者细胞及其后代。

[0057] “治疗”是以防止病变的病理学或症状发展或改变为意图而实施的干预措施。据此,“治疗”指治疗性处理和预防性或防御措施两者。“治疗”亦可具化为姑息疗法。需要处理者包括已经患有病变的和体内病变待防止的。据此,状态、病变或病症的“处理”包括:(1)防止或延缓正在可能苦于或易感所述状态、病变或症状但尚未经历或显现所述状态、病变或病症的临床或亚临床症状的人或其它哺乳动物体内发展的所述状态、病变或病症出现临床症状;(2)抑制所述状态、病变或病症,即,缓解、降低或延缓所述疾病的发展或其复发(在维持性治疗情况下)或其至少一种临床或亚临床症状;或(3)减轻所述疾病,即造成所述状态、病变或病症或其临床或亚临床症状的至少一者的退行。对被处理的个体的益处是统计学显

着的,或至少可为患者或医师察觉。

[0058] “载体”是物质的组合物,其包含单离的核酸且可用来将所述单离的核酸输送至细胞内部。载体的实例包括但不限于,线性多核苷酸、与离子化合物或两亲化合物联合的多核苷酸、质粒、和病毒。因此,术语“载体”包括自主复制的质粒或病毒。所述术语也视为包括促进将核酸转移至细胞内的非质粒化合物和非病毒化合物如,举例而言,聚赖氨酸化合物、脂质体等。病毒载体的实例包括,但不限于,腺病毒载体、腺病毒群载体、逆转录病毒载体等。

[0059] 术语“序列一致性百分比”指代任何给定的存疑序列与目标序列之间的一致性的程度。

[0060] 术语“外源性”表明,所述核酸或多肽是重组核酸构造的一部分或由所述重组核酸编码,或并非出于其天然环境中。例如,外源性核酸可以是来自被引入一种物种的另一物种的序列,即异源性核酸。典型地,这一外源性核酸经由重组核酸构造而被引入其它物种内。外源性核酸可以是一有机体原生的且已经被再次引入所述有机体细胞内的序列。包括原生序列的外源性核酸通常可通过链接至所述外源性核酸的非天然序列的存在而与天然出现的序列区分开来,所述非天然序列例如重组核酸构造中位于原生序列侧翼的非原生调节性序列。此外,稳定转化的外源性核酸典型被整合在除发现原生序列处之外的位置。

[0061] 术语“药学可接受的”(或“药理学可接受的”)是指,当酌情给药至动物或人时,不产生副作用、过敏反应或其它不良反应的分子整体和组合物。本文中使用的术语“药学可接受的载剂”包括任何和所有溶剂、分散截止、涂层、抗菌剂、等渗剂和吸收延迟剂、缓冲剂、赋形剂、粘合剂、润滑剂、凝胶、表面活性剂等,它们可用作药学可接受的物质的介质。

[0062] 若通过Swiss Prot.或GENBANK登录号指代氨基酸序列,则所述序列通过引用而并入本文。与所述登录号相关联的信息,如信号肽、胞外域、跨膜域、启动子序列和转译开始的鉴定,也通过引用而以其整体并入本文。

[0063] 基因:本文中揭露的所有基因、基因名称、及基因产物均试图与来自可应用本文揭露的组合物和方法的任何物种的同源染色体对应。应理解,当来自特定物种的基因或基因产物被揭露时,本揭示内容试图仅做例示性说明,而不应理解为限制,除非其所出现的语境中明确指出。因此,例如,对于本文中揭露的基因或基因产物,试图涵盖来自其它物种的同源及/或直系同源基因和基因产物。

[0064] 范围:贯穿本揭示内容,本发明的多个方面可以范围格式呈现。应理解,范围格式的说明仅仅为了表述的方便和简洁,而不应视为对本发明范畴的刚性限制。据此,对范围的说明应视为具有具体揭露的所有可能的子范围以及所述范围内的个体数值。例如,对范围如从1 至6的说明应视为具有具体揭露的子范围,如从1至3、从1至4、从 1至5、从2至4、从2至6、从3至6等,以及所述范围内的个体数值,例如,1、2、2.7、3、4、5、5.3、和6。这一规则可无视所述范围的广度而使用。

[0065] 组合物

[0066] 本文揭露的组合物可包括编码CRISPR相关的核酸内切酶如Cas9 的核酸。一些具体实施例中,亦可编码一种或多种与HIV的靶标序列互补的向导RNA。据此,一些具体实施例中,提供用于令整合为被人免疫缺陷病毒(HIV)潜伏性感染的宿主细胞基因组中的前病毒DNA失活的组合物,所述组合物包含至少一种编码成簇规律间隔短回文重复序列(CRISPR)相关的核酸内切酶的单离的核酸序列、以及至少一种向导RNA(gRNA),所述至少一种gRNA

具有与前病毒HIV DNA的长末端重复序列 (LTR)中的靶标序列互补的间隔序列序列。某些具体实施例中,所述至少一种gRNA包含与靶标核酸序列互补的核酸序列,所述靶标核酸序列与SEQ ID NO:1至SEQ ID NO:66及其片段、突变体、变体或组合的序列一致性为至少75%。其它具体实施例中,所述至少一种gRNA包含至少一种与靶标核酸序列互补的核酸序列,所述靶标核酸序列包含SEQ ID NO:1至SEQ ID NO:66、及其片段、突变体、变体或组合。某些具体实施例中,所述至少一种gRNA包含一核酸序列,所述核酸序列与 SEQ ID NO:1至SEQ ID NO:66及其片段、突变体、变体或组合的序列一致性为至少75%。其它具体实施例中,所述至少一种gRNA包含至少一个包含SEQ ID NO:1至SEQ ID NO:66及其片段、突变体、变体或组合的核酸序列。

[0067] 再其它具体实施例中,所述至少一种gRNA选自gRNA A,其具有与前病毒DNA中的靶标序列SEQ ID NO:1或靶标序列SEQ ID NO:2互补的间隔序列序列;gRNA B,其具有与前病毒DNA中的靶标序列SEQ ID NO:3或靶标序列SEQ ID NO:4互补的间隔序列序列;或gRNA A与 gRNA B的组合。

[0068] 所述单离的核酸可由载体编码,或涵盖于一种或多种递送载具和制剂中,如下文详述。

[0069] CRISPR相关的核酸内切酶:CRISPR/Cas9诱导突变借以令前病毒失活的机制可变。例如,所述突变可影响前病毒复制和病毒基因表达。所述突变可包含一个或多个删除。所述删除的尺寸可从一个核苷酸碱基对至约10,000个碱基对的范围内变化。一些具体实施例中,所述删除可包括全部或大体上全部的前病毒序列。一些具体实施例中,所述删除可根除前病毒。所述突变亦可包含一个或多个插入,换言之,一个或多个核苷酸碱基对加入前病毒序列中。所插入序列的尺寸可在例如约1个碱基对至约300个核苷酸碱基对的范围内变化。所述突变可包含一个或多个点突变,换言之,单个核苷酸被替换为另一核苷酸。可用的点突变是那些具有功能性后果的突变,例如导致氨基酸密码子转化为终止密码子的突变,或导致产生非功能性蛋白质的突变。

[0070] 已经鉴别了三种类型(I至III)的CRISPR。CRISPR簇含有间隔序列,且所述序列与先行可动因子互补。CRISPR簇被转录并加工入天然 CRISPR RNA (crRNA)中。CRISPR相关的核酸内切酶Cas9属于II型 CRISPR/Cas系统,且具有强烈的切割靶标DNA的核酸内切酶活性。Cas9 由含有约20个碱基对(bp)的独特靶标序列(称为间隔序列)的成熟 crRNA和用作以核酸内切酶III为目标的前crRNA的加工的向导的反式激活的小RNA(tracrRNA)所引导。经由所述crRNA上的间隔序列与所述靶标DNA上的互补序列(称为前间隔序列)之间的互补性碱基配对,所述crRNA:tracrRNA双链将Cas9导向靶标DNA。Cas9识别三核苷酸(NGG)前间隔序列邻近基序(PAM),以指定切割位点(自PAM计数的第3个核苷酸)。所述crRNA和tracrRNA可独立表达,或经由合成茎环(AGAAAU)工程化入人工融合小向导RNA(sgRNA),以模拟天然 crRNA/tracrRNA双链。此sgRNA,如shRNA,可经合成或体外转录,用于引导RNA的转染或从U6或H1启动的RNA表达载体表达,但所述人工sgRNA的裂解效率比那些具有单独表达的crRNA和tracrRNA的系统低。

[0071] 所述CRISPR相关的核酸内切酶可以是Cas9核酸酶。所述Cas9核酸酶可具有与野生型酿脓链球菌(*Streptococcus pyogenes*)序列一致的核苷酸序列。所述CRISPR相关的核酸内切酶可以是来自其它物种如其它链球菌种如嗜热链球菌(*Streptococcus*

thermophiles)的序列。Cas9核酸酶序列可源自其它物种,包括但不限于:达松维尔拟诺卡氏菌(*Nocardiopsis dassonvillei*)、原始链霉菌(*Streptomyces pristinaespiralis*)、绿产色链霉菌(*Streptomyces viridochromogenes*)、粉红链霉菌(*Streptomyces roseum*)、酸热脂环酸杆菌(*Alicyclobacillus acidocaldarius*)、假蕈状芽孢杆菌(*Bacillus pseudomycooides*)、还原硒盐芽孢杆菌(*Bacillus selenitireducens*)、兴安微小杆菌(*Exiguobacterium sibiricum*)、德氏乳杆菌(*Lactobacillus delbrueckii*)、唾液乳杆菌(*Lactobacillus salivarius*)、海洋微颤菌(*Microscilla marina*)、伯克氏菌目(*Burkholderiales bacterium*)、萘降解极单胞菌(*Polaromonas naphthalenivorans*)、极单胞菌(*Polaromonas sp.*)、海洋固氮蓝藻(*Crocospaera watsonii*)、蓝丝菌属(*Cyanothece sp.*)、铜绿微囊藻(*Microcystis aeruginosa*)、聚球藻属(*Synechococcus sp.*)、阿拉伯糖醋盐杆菌(*Acetohalobium arabaticum*)、丹氏制氮菌(*Ammonifex degensii*)、热角军纤维素菌(*Caldicelulosiruptor becsicii*)、矿菌(*Candidatus desulforudis*)、肉毒梭状芽孢杆菌(*Clostridium botulinum*)、艰难梭菌(*Clostridium difficile*)、大芬戈尔德菌(*Finegoldia magna*)、嗜热盐碱厌氧菌(*Natranaerobius thermophilus*)、嗜热丙酸降解发酵菌(*Pelotomaculum thermopropionicum*)、嗜酸性喜温硫杆菌(*Acidithiobacillus caldus*)、嗜酸性氧化亚铁硫杆菌(*Acidithiobacillus ferrooxidans*)、酒色别样着色菌(*Allochromatium vinosum*)、海杆菌属(*Marinobacter sp.*)、嗜盐亚消化球菌(*Nitrosococcus halophilus*)、亚硝化球菌(*Nitrosococcus watsoni*)、假交替单胞菌(*Pseudoalteromonas haloplanktis*)、消旋纤线杆菌(*Ktedonobacter racemifer*)、甲烷盐菌(*Methanohalobium evestigatum*)、鱼腥藻(*Anabaena variabilis*)、泡沫节球藻(*Nodularia spumigena*)、念珠藻属(*Nostoc sp.*)、极大螺旋藻(*Arthrospira maxima*)、钝顶螺旋藻(*Arthrospira platensis*)、螺旋藻属(*Arthrospira sp.*)、林氏藻属(*Lyngbya sp.*)、原型微鞘藻(*Microcoleus chthonoplastes*)、颤藻属(*Oscillatoria sp.*)、石袍藻(*Petrotoga mobilis*)、非洲栖热腔菌(*Thermosiphon africanus*)、或藻青菌(*Acaryochloris marina*)。绿脓假单胞菌(*Pseudomonas aeruginosa*)、大肠杆菌(*Escherichia coli*)、或其它经测序的细菌基因组和古生菌,或其它原核微生物,亦可为在本文中揭露的具体实施例中使用的Cas9序列的来源。

[0072] 野生型酿脓链球菌(*Streptococcus pyogenes*) Cas9序列可经修饰。一种例示性且优选的CRISPR相关的核酸内切酶是Cas9核酸酶。所述Cas9核酸酶可具有与野生型酿脓链球菌序列一致的核苷酸序列。一些具体实施例中,所述CRISPR相关的核酸内切酶可以是来自另一物种如其它链球菌种如嗜热链球菌、绿脓假单胞菌、大肠杆菌、或其它经测序的细菌基因组和古生菌、或其它原核微生物的序列。或者,所述野生型酿脓链球菌Cas9序列可经修饰。所述核酸序列可以是优化用于在哺乳动物细胞内有效表达的密码子,即,“人源化”。人源化的Cas9核酸酶序列可以是,例如,由以Genbank保藏号KM099231.1 GI:669193757; KM099232.1 GI:669193761;或KM099233.1 GI:669193765列举的表达载体中的任一种编码的Cas9核酸酶序列。或者,所述Cas9核酸酶序列可以是,例如,可商购的载体如来自Addgene (Cambridge, MA)的PX330或PX260中含有的序列。一些具体实施例中,所述Cas9核酸内切酶可具有作为下述任何Cas9核酸内切酶序列的变体或片段的氨基酸序列:Genbank保藏号KM099231.1 GI:669193757、KM099232.1 GI:669193761、或KM099233.1 GI:669193765,或

PX330 或PX260 (Addgene, Cambridge, MA) 的Cas9氨基酸序列。

[0073] 所述Cas9核酸酶序列可以是突变的序列。例如, Cas9核酸酶可以在保守的HNH域和RuvC域中突变, 两个域均牵涉于链特异性裂解中。另一实例中, RuvC催化域中的天冬氨酸盐至丙氨酸(D10A)的突变令所述Cas9切口酶突变体(Cas9n)将DNA切口而非裂解, 以得到单链断裂, 且后来通过HDR进行的优先修复可潜在性地降低来自脱靶双链断裂的不希望的indel突变出现的频率。所述Cas9核苷酸序列可经修饰, 以编码Cas9的生物学活性变体, 且这些变体可具有或可包括, 例如, 由于含有一个或多个突变(如, 加入、删除、或替换突变或这些突变的组合)而不同于野生型Cas9的氨基酸序列。所述替换突变中的一个或多个可以是替换(如, 保守氨基酸替换)。例如, Cas9多肽的生物学活性变体可具有一氨基酸序列, 所述氨基酸序列与野生型Cas9多肽的序列一致性为至少或约50% (如, 至少或约50%、55%、60%、65%、70%、75%、80%、85%、90%、95%、97%、98%、或99%的序列一致性)。保守氨基酸替换典型包括下述各组的组内替换: 甘氨酸和丙氨酸; 缬氨酸、异亮氨酸和亮氨酸; 天冬氨酸和谷氨酸; 天冬酰胺、谷氨酰胺、丝氨酸和苏氨酸; 赖氨酸、组氨酸和精氨酸; 以及苯丙氨酸和酪氨酸。Cas9氨基酸序列中的氨基酸残基可以是非天然出现的氨基酸残基。天然出现的氨基酸残基包括那些天然地由遗传密码编码的氨基酸残基以及非标准氨基酸(如, 具有D-构型而非L-构型的氨基酸)。本发明的多肽亦可包括作为标准残基的修饰版本的氨基酸残基(如, 吡咯赖氨酸可用来替代赖氨酸, 而硒代半胱氨酸可用来替代半胱氨酸)。非天然出现的氨基酸残基是自然界中未发现但符合氨基酸的基本公式且可被并入肽中的那些氨基酸残基。这些包括D-别异亮氨酸((2R, 3S)-2-氨基-3-甲基戊酸)和L-环戊基甘氨酸((S)-2-氨基-2-环戊基乙酸)。再例如, 可查阅教科书或世界性网站(当前由加州理工学院(California Institute of Technology)维护的一个网站显示已经成功地并入功能性蛋白质中的非天然氨基酸的结构)。

[0074] 向导RNA序列: 本发明的组合物和方法可包括编码与HIV中靶标序列互补的向导RNA的序列。HIV的遗传变异性体现在已经揭示的多个组和亚型中。一批HIV序列提供于Los Alamos HIV数据库和概略(即, 序列数据库网站为<http://www.hiv.lanl.gov>)中。本发明的方法和组合物可应用于来自哪些各种组、亚型和循环重组形式的任何HIV。这些包括, 例如, HIV-1主要组(一旦称为M组)和次要组N组、O组或P组, 以及但不限于, 下述亚型: A、B、C、D、F、G、H、J和K或HIV组(例如, 但不限于, 下述N组、O组和P组中任一组)。

[0075] 所述向导RNA可以是与编码序列或非编码序列(即, 靶标序列)互补的序列。例如, 所述向导RNA可以是与HIV长末端重复序列(LTR)区域互补的序列。

[0076] 实施例部分中揭露的实验显示, 使用本发明的Cas9与gRNA组合物处理T淋巴细胞和原代人类T细胞, 最普遍是通过根除前病毒基因组而造成被整合的HIV-1前病毒失活。全基因组测序和总和生物信息学分析的结果排除了对正常宿主DNA的任何基因毒性。

[0077] 据此, 本发明涵盖用于令整合入被HIV潜伏性感染的宿主细胞基因组内的前病毒DNA失活的组合物。所述组合物包括至少一种编码CRISPR相关的核酸内切酶的单离的核酸序列、以及至少一种与前病毒 HIV DNA的长末端重复序列(LTR)中的靶标序列互补的gRNA。本发明还涵盖令整合入被HIV潜伏性感染的宿主细胞基因组内的前病毒DNA失活的方法。所述方法包括以组合物处理所述宿主细胞的步骤, 其中, 所述组合包括CRISPR相关的核酸内切酶、以及至少一种与前病毒HIV DNA的长末端重复序列(LTR)中的靶标序列互补的gRNA。

对于所述组合物和所述方法两者,优选的gRNA包括gRNA A、gRNA B、或最优选gRNA A与gRNA B的组合。

[0078] gRNA可包括含有约20个碱基对(bp)的称为“间隔序列”的独特靶标序列的成熟crRNA;反式激活的小RNA(tracrRNA),其作为向导用于pre-crRNA的以核酸酶III为目标的加工。经由所述crRNA上的间隔序列与所述靶标DNA上的互补性序列(也称为“前间隔序列”)之间的碱基配对,所述crRNA:tracrRNA双链将Cas9导向靶标DNA。本发明中,所述crRNA与tracrRNA可单独表达,或经由合成茎环(AGAAAU)而工程化入人工融合gRNA中,以模拟天然的crRNA/tracrRNA双链。此gRNA可被合成或体外转录,用于引导RNA的转染或从例如U6或H1启动的RNA表达载体表达。当将gRNA揭示为与靶标DNA序列互补时,应理解,事实上与所述靶标DNA序列互补的是所述gRNA的间隔序列。

[0079] 一旦通过gRNA引导靶标序列,Cas9识别三核苷酸(NGG)前间隔序列邻近基序(PAM)以指定切割位点(自PAM计数的第三个核苷酸)。

[0080] HIV-1的长末端重复序列(LTR)区域被再分成U3区域、R区域和U5区域。LTR含有基因表达所需的所有信号,且牵涉入前病毒至宿主细胞基因组内的整合中。例如,在U3内发现了基础或核心启动子、核心增强子和调节区域,而在R内发现了反式激活应答元件。HIV-1中,所述U5区域包括若干子区域,例如,TAR或反式作用应答元件,其牵涉入转录激活中;Poly A,其牵涉入二聚化作用和基因组封装中;Psi或引物结合位点;Psi或封装信号;DIS或二聚物起始位点。

[0081] 本发明的优选gRNA各自与HIV-1 LTR的U3区域中的靶标序列互补。gRNA A可以是与下述两个靶标序列互补的任何gRNA:

[0082] AGGGCCAGGGATCAGATATCCACTGACCTT (SEQ ID NO:1);或

[0083] ATCAGATATCCACTGACCTT (SEQ ID NO:2)。

[0084] gRNA B可以是与下述两个靶标序列互补的任何gRNA:

[0085] AGCTCGATGTCAGCAGTTCTTGAAGTACTC (SEQ ID NO:3);或

[0086] CAGCAGTTCTTGAAGTACTC (SEQ ID NO:4)。

[0087] SEQ ID NOS:1和SEQ ID NOS:3是30bp gRNA,它们用于在实施例部分中详述的实验中,其中,实现了gRNA在淋巴细胞宿主细胞中的稳定表达。SEQ ID NOS:2和SEQ ID NOS:4是截短的20bp gRNA,它们用于慢病毒载体的构建中。本发明的gRNA亦可包括来自位于一端的HIV-1 LTR的PAM序列,但PAM序列并不包括在实施例中所报导的gRNA内。例示性的包括PAM序列的gRNA A是

[0088] AGGGCCAGGGATCAGATATCCACTGACCTTTGG (SEQ ID NO:5)。例示性的包括PAM序列的gRNA B是AGCTCGATGTCAGCAGTTCTTGAAGTACTCCGG (SEQ ID NO:6)。

[0089] 根据本发明的gRNA序列可与所述靶标序列的正义链或反义链互补。它们可包括额外的可以与靶标序列互补或不互补的5'及/或3'序列。它们与靶标序列的互补性可能低于100%,例如,75%的互补性。所述gRNA序列可作为一种或多种不同序列的组合而使用,如复合构型。复合构型可包括2、3、4、5、6、7、8、9、10或更多不同向导RNA的组合。在实施例1和2中揭露的实验中发现,采用gRNA A和gRNA B的双链“二切割”策略在令病毒失活和根除HIV-1的两个LTR中每一个中由Cas9诱导的裂解之间的序列上,尤其有效。

[0090] 修饰的或突变的核酸序列:一些具体实施例中,任何核酸序列均可经修饰或自原

生核酸序列衍生,例如,通过引入核酸碱基、骨架等的突变、删除、替换、修饰而修饰或衍生。所述核酸序列包括载体、基因编辑剂、gRNA等。预期用于本发明的一些修饰核酸序列的实例包括那些包含经修饰的骨架的核酸序列,所述经修饰的骨架为,例如,硫代磷酸酯、磷酸三酯、甲基磷酸酯、短链烷基或环烷基类糖间链接、或短链杂原子或杂环类糖间链接。一些具体实施例中,经修饰的寡核苷酸包含那些具有硫代磷酸酯骨架的寡核苷酸和那些具有杂原子骨架、 $\text{CH}_2\text{-NH-O-CH}_2$ 、 $\text{CH}_2\text{-N(CH}_3\text{)-O-CH}_2$ [称为亚甲基(甲基亚胺基)或 MMI骨架]、 $\text{CH}_2\text{-O-N(CH}_3\text{)-CH}_2$ 、 $\text{CH}_2\text{-N(CH}_3\text{)-N(CH}_3\text{)-CH}_2$ 、和 $\text{O-N(CH}_3\text{)-CH}_2\text{-CH}_2$ 骨架的寡核苷酸,其中,原生磷酸二酯骨架表示为 O-P-O-CH_2)。De Mesmaeker et al. *Acc. Chem. Res.* 1995, 28:366-374 中揭露的酰胺骨架也是本发明的具体实施例。一些具体实施例中,核酸序列具有吗啉代骨架结构(Summerton and Weller, U.S. Pat. No. 5,034,506)、肽核酸(PNA)骨架,其中,所述寡核苷酸的磷酸二酯骨架被替换为聚酰胺骨架,且核酸碱基直接或键结至聚酰胺骨架的氮杂氮原子(Nielsen et al. *Science* 1991, 254, 1497)。所述核酸序列亦可包含一个或多个经取代的糖部分。所述核酸序列亦可具有替代呋喃戊糖基的糖模拟物如环丁基。

[0091] 所述核酸序列亦可额外地或者替代地包括核酸碱基(所述领域中一般简称为“碱基”)修饰或替换。本文中,“未修饰的”或“天然”核酸碱基包括腺嘌呤(A)、鸟嘌呤(G)、胸腺嘧啶(T)、胞嘧啶(C)和尿嘧啶(U)。经修饰的核酸碱基包括仅很少或短暂地发现于天然核酸中的核酸碱基,如,次黄嘌呤、6-甲基腺嘌呤、5-甲基嘧啶特别是5-甲基胞嘧啶(也称为5-甲基-2'-去氧胞嘧啶且所述领域中一般表示为 5-Me-C)、5-羟甲基胞嘧啶(HMC)、糖基HMC和龙胆二糖基HMC,以及合成性核酸碱基如2-氨基腺嘌呤、2-(甲基胺基)腺嘌呤、2-(咪唑基烷基)腺嘌呤、2-(氨基烷基胺基)腺嘌呤或其它杂取代的烷基腺嘌呤、2-硫尿嘧啶、2-硫胸腺嘧啶、5-溴尿嘧啶、5-羟甲基尿嘧啶、8-氮杂鸟嘌呤、7-去氮鸟嘌呤、 N^6 -(6-氨基己基)腺嘌呤和2,6-二氨基嘌呤(Kornberg, A., *DNA Replication*; W.H. Freeman & Co., San Francisco, 1980, pp75-77; Gebeyehu, G., et al. *Nucl. Acids Res.* 1987, 15:4513)。可包括所述领域中已知的“万能”碱基,如肌苷。已经显示,5-Me-C替换将核酸双链的稳定性增加0.6至1.2°C(Sanghvi, Y. S., Crooke, S.T. and Lebleu, B., eds., *Antisense Research and Applications*, CRC Press, Boca Raton, 1993, pp.276-278)。

[0092] 本发明的核酸序列的另一修饰牵涉将提升寡核苷酸的活性或细胞摄取的一个或多个部分或结合体化学链接至所述核酸序列。这些部分包括但不限于,脂质部分,如胆固醇部分、胆甾醇基部分(Letsinger et al., *Proc. Natl. Acad. Sci. USA* 1989, 86, 6553)、胆酸(Manoharan et al. *Bioorg. Med. Chem. Lett.* 1994, 4, 1053)、硫醚如己基-S-三苯甲硫醇(Manoharan et al. *Ann. N. Y. Acad. Sci.* 1992, 660, 306; Manoharan et al. *Bioorg. Med. Chem. Lett.* 1993, 3, 2765)、巯基胆固醇(Oberhauser et al., *Nucl. Acids Res.* 1992, 20, 533)、脂肪族链如十二碳二醇或十一烷基残基(Saison-Behmoaras et al. *EMBO J.* 1991, 10, 111; Kabanov et al. *FEBS Lett.* 1990, 259, 327; Svinarchuk et al. *Biochimie* 1993, 75, 49)、磷脂质如二-十六烷基-甘油或1,2-二-O-十六烷基-甘油-3-H-磷酸三乙铵(Manoharan et al. *Tetrahedron Lett.* 1995, 36, 3651; Shea et al. *Nucl. Acids Res.* 1990, 18, 3777)、聚胺或聚乙二醇链(Manoharan et al. *Nucleosides & Nucleotides* 1995, 14, 969)、或金刚烷乙酸(Manoharan et al. *Tetrahedron Lett.* 1995, 36, 3651)。不需要给定核酸序列的所有位置均被均匀地修饰,事实上,超过一种

前述修饰可并入单个核酸序列中,甚至位于核酸序列中的单个核苷酸上。

[0093] 一些具体实施例中,将所述RNA分子如crRNA、tracrRNA、gRNA 工程化为包含一个或多个经修饰的核酸碱基。例如,已知的对RNA分子的修饰可在例如《基因VI》(Genes VI)第9章(“遗传密码的诠释”(Interpreting the Genetic Code)),Lewis,ed.(1997,Oxford University Press,New York)和Grosjean与Benne编撰的《RNA的修饰和编辑》(Modification and Editing of RNA)(1998,ASM Press, Washington DC)中找到。经修饰的RNA组分包括下列:2'-O-甲基胞苷;N⁴-甲基胞苷;N⁴,2'-O-二甲基胞苷;N⁴-乙酰基胞苷;5-甲基胞苷;5,2'-O-二甲基胞苷;5-羟甲基胞苷;5-甲酰基胞苷;2'-O-甲基-5-甲酰基胞苷;3-甲基胞苷;2-硫胞苷;赖胞苷(lysidine);2'-O-甲基尿苷;2-硫代尿苷;2-硫代-2'-O-甲基尿苷;3,2'-O-二甲基尿苷;3-(3-氨基-3-羧基丙基)尿苷;4-硫代尿苷;核糖基胸腺嘧啶;5,2'-O-二甲基尿苷;5-甲基-2-硫代尿苷;5-羟基尿苷;5-甲氧基尿苷;尿苷-5-羟乙酸;尿苷-5-羟乙酸甲酯;5-羧基甲基尿苷;5-甲氧羰基甲基尿苷;5-甲氧羰基甲基-2'-O-甲基尿苷;5-甲氧羰基甲基-2'-硫代尿苷;5-氨基甲酰基甲基尿苷;5-氨基甲酰基甲基-2'-O-甲基尿苷;5-(羧基羟甲基)尿苷;5-(羧基羟甲基)尿苷甲酯;5-胺甲基-2-硫代尿苷;5-甲基胺甲基尿苷;5-甲基胺甲基-2-硫代尿苷;5-甲基胺甲基-2- 硫尿苷;5-羧基甲基胺甲基尿苷;5-羧基甲基胺甲基-2'-O-甲基-尿苷;5-羧基甲基胺甲基-2-硫代尿苷;二氢尿苷;二氢核糖基胸腺嘧啶;2'-甲基腺苷;2-甲基腺苷;N⁶-甲基腺苷;N⁶,N⁶-二甲基腺苷;N⁶,2'-O-三甲基腺苷;2-甲基硫代-N⁶-N-异戊烯基腺苷;N⁶-(顺-羟基异戊烯基)-腺苷;2-甲基硫代-N⁶-(顺-羟基异戊烯基)-腺苷;N⁶-缩水甘油基氨基甲酰基腺苷;N⁶-苏氨酰基氨基甲酰基腺苷;N⁶-甲基-N⁶-苏氨酰基氨基甲酰基腺苷;2-甲基硫代-N⁶-甲基-N⁶-苏氨酰基氨基甲酰基腺苷;N⁶-羟基正缬氨酰基氨基甲酰基腺苷;2-甲基硫代-N⁶-羟基正缬氨酰基氨基甲酰基腺苷;2'-O-核糖基腺苷(磷酸酯);肌苷;2'-O-甲基肌苷;1-甲基肌苷;1,2'-O-二甲基肌苷;2'-O-甲基鸟苷;1-甲基鸟苷;N²-甲基鸟苷;N²,N²-二甲基鸟苷;N²,2'-O-二甲基鸟苷;N²,N²,2'-O-三甲基鸟苷;2'-O-核糖基鸟苷(磷酸酯);7-甲基鸟苷;N²,7-二甲基鸟苷;N²,N²,7-三甲基鸟苷;怀俄苷(wyosine);甲基怀俄苷;修饰不足的羟基怀丁苷;怀丁苷(wybutosine);羟基怀丁苷;过氧怀丁苷;辫苷(queuosine);环氧辫苷;半乳糖基-辫苷;甘露糖基-辫苷;7-氰基-7-去氮鸟苷;古菌苷(arachaeosine)[亦称7-甲酰胺基-7-去氮鸟苷];和7-胺甲基-7-去氮鸟苷。

[0094] 本发明的单离的核酸分子可通过标准技术制备。例如,可使用聚合酶链反应(PCR)技术来获得含有本文中揭示的核苷酸序列的单离的核酸。多种PCR方法揭示在例如,Dieffenbach和Dveksler编撰的《PCR 引物:实验室手册》(PCR Primer:A Laboratory Manual,Cold Spring Harbor Laboratory Press,1995)中。通常,来自感兴趣的区域末端或超出所述范围的序列信息被用来设计寡核苷酸引物,所述引物与待扩增模板的相对链的序列一致或相似。多种PCR策略亦可获得,通过所述策略,可将位点特异性核苷酸序列修饰引入模板核酸中。

[0095] 单离的核酸亦可经化学合成为单个核酸分子(如,使用亚磷酰胺技术在3'至5'方向上的自动DNA合成),或合成为一系列寡核苷酸。例如,可合成含有所希望序列的一对或多对长寡核苷酸(如,>50至100个核苷酸),且每一对含有短节段的互补性(如,约15个核苷酸),因此,当所述寡核苷酸对被退火时,形成双链。使用DNA聚合酶来延伸所述寡核苷酸,每

寡核苷酸对得到单个的双链核酸,所述核酸随后可连接入载体中。

[0096] 两个核酸或其编码的多肽可揭示为彼此具有某种程度的一致性。例如,Cas9蛋白及其生物学活性变体可揭示为展现某种程度的一致性。可通过在蛋白质信心检索(Protein Information Research(PIR))网站(pir.georgetown.edu)中找出短Cas9序列,之后使用“短接近一致序列”局部序列排比检索基本工具(Basic Local Alignment Search Tool(BLAST))算法在NCBI网站(ncbi.nlm.nih.gov/blast)分析,完成比对。

[0097] 可测得与Cas9序列一致性的百分比,且所鉴别的变体可用作 CRISPR相关的核酸内切酶及/或检验其作为药物组合物的效能。天然出现的Cas9可以是查询序列,而Cas9蛋白的片段可以是目标序列。同样,Cas9蛋白的片段可以是查询序列,而其生物学活性变体可以是目标序列。为了测定序列一致性,可使用计算机程序ClustalW(version 1.83,缺省参数)将查询核酸或氨基酸序列与一个或多个目标核酸或氨基酸序列分别比对,所述程序运行跨越其整体长度而进行核酸或蛋白序列的比对(总体比对)。见Chenna et al,Nucleic acids Res. 31:3497-3500,2003。

[0098] 重组构建物和递送载体:用于包含在所述药物组合物中的例示性表达载体包括质粒载体和慢病毒载体,但本发明并不限于这些载体。大量的宿主/表达载体组合可用来表达本文中揭示的核酸序列。适当的表达载体包括而不限,源自例如噬菌体、杆状病毒和逆转录病毒的质粒和病毒载体。大量载体和表达系统可从诸如Novagen(Madison, WI)、Clontech(Palo Alto,CA)、Stratagene(La Jolla,CA)和 Invitrogen/Life Technologies(Carlsbad,CA)等公司商购。标记基因可赋予宿主细胞以可选择的表型。例如,标记物可赋予耐杀虫剂性,如对抗生素(如,卡那霉素、G418、博来霉素、或潮霉素)的抗性。表达载体可包括设计为促进对所表达的多肽的操纵或检测(如,纯化或定位)的标签序列。标签序列如绿色荧光蛋白(GFP)、谷胱甘肽S-转移酶(GST)、多组氨酸、c-myc、血球凝集素、或FLAG™标签(Kodak,New Haven, CT)序列典型被表达为与所编码的多肽融合。此类标签可插入多肽中的任意处,包括羧基端或氨基端。所述载体亦可包括复制的起源、支架附件区域(SAR)、调节区域等。术语“调节区域”指的是,影响转录或转译起始和速率、以及转录或转译产物的稳定性及/或移动性的核苷酸序列。调节区域包括而不限,启动子序列、增强子序列、应答元件、蛋白识别位点、可诱导元件、蛋白结合序列、5'和3'未转译区域(UTR)、转录开始位点、终止序列、聚腺苷酸化序列、核定位信号、和内含子。术语“可操作地链接”指的是,将调节区域和待转录的序列放置在核酸中,以影响这一序列的转录或转译。例如,为了令编码序列处于启动子的控制之下,所述多肽的转译阅读框的转译起始位点典型位于所述启动子下游的1至40个核苷酸之间。但是,启动子可位于转译起始位点上游多达约5,000个核苷酸处,或转录开始位点上游约2,000个核苷酸处。启动子典型包含至少一个核心(基础)启动子。启动子亦可包括至少一个控制元件,如增强子序列、上游元件或上游激活区域(UAR)。对待包括的启动子的选择基于若干因素,包括但不限于,效率、可选择性、可诱导性、所希望的表达水平、和细胞或组织优先的表达。对于所述领域技术人员,通过适宜地选择并相对于编码序列定位启动子和其它调节区域是例行公事。

[0099] 若需要,本发明的多核苷酸亦可与微递送载体如阳离子脂质体和腺病毒载体合用。对于脂质体制备、靶向和内容为的递送的综述见 Mannino and Gould-Fogerite, BioTechniques,6:682(1988)。也见 Feigner and Holm,Bethesda Res.Lab.Focus,1 1

(2):21 (1989) 和Maurer,R.A.,Bethesda Res.Lab.Focus,11 (2):25 (1989)。

[0100] 在实施例部分中揭露的实验中发现,慢病毒载体在实现本发明的 Cas9和gRNA在人类T淋巴细胞系中以及首次在人类T细胞原代培养物中的表达中有效,所述人类T细胞包括源自HIV-1⁺患者的T细胞。在来自HIV⁺患者的原代T细胞中,慢病毒递送的Cas9与gRNA A和gRNA B的联合表达显著降低了病毒拷贝数和病毒蛋白表达。这代表了与先前基因编辑领域相比在HIV⁺患者治疗中的关键优势。

[0101] 因此,本发明涵盖用于令整合入被HIV潜伏性感染的宿主细胞基因组内的前病毒DNA失活的慢病毒载体组合物。所述组合物包括编码 CRISPR相关的核酸内切酶的单离的核酸,以及至少一种编码至少一种向导gRNA的单离的核酸,其中,所述gRNA包括与前病毒HIV DNA的 LTR中的靶标序列互补的间隔序列,且所述单离的核酸被包括在至少一种慢病毒表达载体内。所述慢病毒表达载体诱导所述CRISPR相关的核酸内切酶和所述至少一种gRNA在宿主细胞内的表达。

[0102] 所有单离的核酸可被包括在单一慢病毒表达载体中,或所述核酸可被再分成慢病毒载体的任何适当组合。例如,所述CRISPR相关的核酸内切酶可并入第一慢病毒表达载体中,第一gRNA可并入第二慢病毒表达载体中,且第二gRNA可并入第三慢病毒表达载体中。当使用多个表达载体时,不需要所有的表达载体全是慢病毒载体。

[0103] 实施例2的结果也表明了将离体培养物中的潜伏性感染的T细胞与本发明的Cas9与gRNA组合物接触的可用性。发现,gRNA A与gRNA B的组合获得对被整合的HIV前病毒DNA的最优根除。关于这一能力的一种用途是过继性疗法,需要使用本发明的组合物离体培养患者的被 HIV感染的细胞,再将HIV被消除的细胞返回所述患者体内。

[0104] 本文中还提供重组构建体,其可用来改造细胞。如本文中所述,重组核酸构建体包含编码Cas9的核酸及/或与HIV中靶标序列互补的向导RNA,所述构建体可操作地链接至一调节区域,所述调节区域适用于表达Cas9及/或与所述细胞中HIV的靶标序列互补的向导RNA。应知晓,大量核酸可编码具有特定氨基酸序列的多肽。遗传密码的退火是所述领域中所周知的。对于很多氨基酸,存在不止一个用作所述氨基酸密码子的核苷酸三连体。例如,可使用适宜的用于特定有机体的密码子偏倚来修饰Cas9的编码序列中的密码子,因此可获得在所述有机体内的最优表达。

[0105] 几种递送方法可与本文中的分子包埋体在体外系统(细胞培养)和体内系统(动物和患者)中合用。一种具体实施例中,可采用慢病毒基因递送系统。这一系统提供基因在具有广泛趋向和用于大DNA插入的能力的分裂细胞和不分裂细胞中稳定的长期存在(Du11 et al,J Virol, 72:8463-8471 1998)。一种具体实施例中,可使用腺相关病毒(AAV) 作为递送方法。AAV是一种非致病性单链DNA病毒,近年来已经活跃地用于在体外系统和体内系统中递送治疗性基因(Choi et al,Curr Gene Ther,5:299-310,2005)。

[0106] 用于在体内外表达本文中揭示的任何多核苷酸包埋体的载体包括,例如,病毒载体(如,腺病毒Ad、AAV、慢病毒、水疱性口炎病毒(VSV)及逆转录病毒)、脂质体和其它含有脂质的复合物、以及其它能介导多核苷酸至宿主细胞的递送的大分子复合物。载体也可包含进一步调整基因递送及/或基因表达或在其它方面向靶标细胞提供有益性质的其它组分或功能。如下文中更详细揭示和例示性说明,此类其它组分包括,例如,影响结合细胞或以细胞为靶向的组分(包括介导细胞类型或组织特异性结合的组分);影响所述载体核酸被细胞

摄取的组分;影响所述多核苷酸被摄取后在所述细胞内定位的组分(如介导核定位的剂);以及影响所述多核苷酸表达的组分。此类组分也可包括标记物,如可检测及/或可选择的标记物,所述标记物可用来检测或选择已经摄取并正在表达被所述载体递送的核酸的细胞。此类组分可提供为所述载体的天然形式(如使用某些病毒载体,所述病毒载体具有介导结合和摄取的组分或功能性),或载体可经修饰以提供此类功能性。其它载体包括Chen等人在*BioTechniques*,34:167-171(2003)中揭示的那些载体。大量此类载体是所述领域中已知且常用的。“重组病毒载体”指的是包含一种或多种异源性基因产物或序列的病毒载体。由于很多病毒载体展现与封装相关的尺寸约束,典型通过替换所述病毒基因组的一个或多个部分而引入所述异源性基因产物或序列。此类病毒可变为复制缺陷型病毒,而复制缺陷型病毒需要在病毒复制和衣壳化过程中提供被删除的功能(使用诸如携带复制及/或衣壳化所必需的基因产物的辅助病毒或封装细胞系)。也已经揭示了其体内存在装载在病毒颗粒外部的待递送多核苷酸待的修饰病毒载体(见,如,Curiel,D T, et al.*PNAS* 88:8850-8854, 1991)。一些具体实施例中,所述载体是复制缺陷型载体。复制缺陷型重组腺病毒载体可根据已知技术生产。见,Quantin,et al,*Proc.Natl.Acad.Sci.USA*,89:2581-2584(1992);Stratford-Perricadet,et al.,*J.Clin.Invest.*, 90:626-630(1992);和Rosenfeld,et al,*Cell*,68:143-155(1992)。

[0107] 表达载体也可包括,例如,染色体DNA序列、非染色体DNA序列和合成DNA序列的片段。适当的载体包括SV40的衍生物和已知的细菌质粒,如大肠杆菌质粒col E1、pCR1、pBR322、pMal-C2、pET、pGEX、pMB9及其衍生物、质粒如RP4;噬菌体DNA,如噬菌体1的众多衍生物,如NM989,及其它噬菌体DNA,如M13和细丝状单链噬菌体DNA;酵母质粒,如2 μ 质粒及其衍生物,可用于真核生物细胞的载体,如可用于昆虫和哺乳动物细胞的载体;源自质粒与噬菌体DNA的组的载体,如已经修饰以采用噬菌体DNA或其它表达控制序列的质粒。

[0108] 其它载体包括病毒载体、融合蛋白和化学接合物。逆转录病毒载体包括莫洛尼鼠白血病病毒和HIV系病毒。一种HIV系病毒载体包含至少两个载体,其中,gag基因和pol基因来自HIV基因组,而env基因来自另一种病毒。DNA病毒载体包括痘病毒载体如天花病毒载体或禽痘病毒载体、疱疹病毒载体如I型单纯疱疹病毒(HSV)载体[Geller, A.I. et al, *J.Neurochem*,64:487(1995);Lim,F.,et al,in *DNA Cloning:Mammalian Systems*, D.Glover,Ed.(Oxford Univ.Press, Oxford England)(1995);Geller,A.I. et al,*Proc Natl.Acad.Sci.: U.S.A.:*90 7603(1993);Geller,A.L,et al,*Proc Natl.Acad.Sci USA:* 87:1 149(1990)]、腺病毒载体[LeGal LaSalle et al,*Science*, 259:988(1993);Davidson,et al,*Nat.Genet.*3:219(1993); Yang,et al,*J.Virol.*69:2004(1995)]和腺相关病毒载体[Kaplitt, M.G.,et al,*Nat.Genet.*8:148(1994)]。

[0109] 一些具体实施例中,所述载体是产生单链DNA的载体,其可产生在细胞内被表达的产物。见,例如,Chen et al,*BioTechniques*,34: 167-171(2003),所述文献通过引用而以其整体并入本文。

[0110] 本文中揭露的多核苷酸可与微递送载具如阳离子脂质体和腺病毒载体合用。对于脂质体制备过程、靶向和内容递送的综述,见Mannino and Gould-Fogerite, *BioTechniques*,6:682(1988)。也见,Feigner and Holm,*Bethesda Res.Lab.Focus*,1 1(2):21(1989)和Maurer, R.A.,*Bethesda Res.Lab.Focus*,1 1(2):25(1989)。

[0111] 本发明的某些具体实施例中,非病毒载体可用来实行转染。核酸的非病毒递送方法包括脂质转染、核转染、显微注射、基因枪、病毒体、脂质体、免疫脂质体、聚阳离子或脂质:核酸接合物、裸DNA、人工病毒体、及试剂增强的DNA摄取。脂质转染揭示于例如专利US 5,049,386、US 4,946,787和US 4,897,355中,而脂质转染试剂也是商售的(如,Transfectam和Lipofectin)。适用于核苷酸的有效受体识别脂质转染的阳离子脂质和中性脂质包括在Jessee拥有的专利US 7,166,298和Jesse拥有的专利US 6,890,554中揭示的那些脂质,所述专利各自的内容通过引用并入本文。递送可以至细胞(如,体外给药或离体给药)或靶标组织(如,体内给药)。

[0112] 合成载体典型是基于阳离子脂质或聚合物的,所述阳离子脂质或聚合物可与带负电的核酸复合以形成直径为100nm量级的颗粒。所述复合物保护核酸不被核酸酶降解。此外,细胞和局部递送策略必须解决对于内化、释放、以及分布于适宜的亚细胞隔室中的需要。系统性递送策略遭遇额外障碍,例如,阳离子递送载具与血液组分的强烈相互作用、被网状内皮系统摄取、肾过滤、毒性、以及载体以所述细胞为靶向的能力。修饰所述阳离子非病毒体的表面,可最小化其与血液组分的相互作用、降低网状内皮系统的摄取、减轻其毒性、并增加其与靶标细胞的结合亲和性。血浆蛋白的结合(也称为调理素作用)是RES 借以识别循环纳米颗粒的主要机制。例如,巨噬细胞,如肝脏中的 Kupffer细胞,经由清道夫受体识别经调理素作用的纳米颗粒。

[0113] 本发明的核酸序列可递送至受试者的适宜细胞。这可通过,例如,使用聚合性的、生物可降解的微粒或微胶囊递送载具,令其尺寸对于通过噬菌细胞如巨噬细胞进行的吞噬作用为最优。例如,可使用直径为1至10 μm 的PLGA(聚乙丙交酯)微粒。所述多核苷酸被密封在这些微粒内,这些微粒被巨噬细胞摄取并在所述细胞内被逐步降解,从而释放所述多核苷酸。一旦被释放,DNA即在所述细胞内被表达。第二类型的为了试图不被细胞直接摄取,而是主要用作核酸的缓释蓄积池,其仅在当通过生物降解而从所述微粒中释放时被细胞摄取。这里聚合性颗粒应因此足够大,以至于妨碍吞噬作用(即,大于5 μm ,且优选大于20 μm)。另一种实现所述核酸的摄取的途径是使用脂质体,优选通过标准方法使用。核酸可单独并入这些递送载具内,或与组织特异性抗体如以一般作为HIV感染的潜伏性感染蓄积池的细胞类型为靶标的抗体一起并入。或者,可制备由质粒或其它通过静电力或共价力粘附至聚-L-赖氨酸的载体组成的分子复合物。聚-L-赖氨酸结合至配体,而所述配体可结合至靶标细胞上的受体。“裸DNA”(即,无递送载具)至肌肉内、皮肤内、或皮下位点的递送,是另一种实现体内表达的手段。在相关多核苷酸(如,表达载体)中,所述核酸序列编码单离的核酸序列,所述单离的核酸序列包含编码CRISPR/Cas的序列及/或与 HIV的靶标序列互补的向导RNA,如上文所述。

[0114] 一些具体实施例中,载体的递送也可由外泌体所介导。外泌体是由多种细胞类型释放的脂质纳米载具。它们通过在细胞之间运输核酸和蛋白质而介导细胞内通讯。外泌体含有源自内吞途径的RNA、miRNA、和蛋白质。它们可通过内吞作用、融合或两者而被靶标细胞摄取。外泌体可被驾驭以将核酸递送至具体的靶标细胞。

[0115] 本发明的表达构建体也可通过纳米线团的手段递送。纳米线团是茧样的DNA纳米复合体(Sun, et al, J. Am. Chem. Soc. 2014, 136: 14722-14725)。它们可负载用于被靶标细胞摄取的核酸,并将其释放在靶标细胞的细胞质中。构建纳米线团、令其负载、以及设计释

放分子的方法可在Sun等人的文献(Sun W, et al, J. Am. Chem. Soc. 2014, 136:14722-14725; Sun W, et al, Angew. Chem. Int. Ed. 2015: 12029-12033)中找到。

[0116] 所述核酸和载体也可应用于装置(如,导管)的表面或包含在泵、贴剂、或任何其它药物递送装置内。本文中揭露的所述核酸和载体可单独给药,或在药学可接受的赋形剂或载剂(如,生理盐水)的存在下以混合物方式给药。基于给药模式和路径而选择所述赋形剂或载剂。适用于药物制剂的药物载剂和药学必需品可在《雷明顿药物科学》(Remington's Pharmaceutical Sciences (E.W.Martin))中找到,所述书是本领域内周知的参考书,也可在USP/NF(美国药典和国家处方集)中找到。

[0117] 本发明的一些具体实施例中,使用脂质体实现进入细胞或组织内的转染。核酸的脂质体制剂的药理学很大程度上由所述核酸被密封在脂质体双层内部的程度决定。密封的核酸被保护而不被核酸酶降解,而那些仅仅与脂质体表面关联的核酸不受保护。密封的核酸共享延长的循环寿命和未受损脂质体的生物分布,而表明关联的核酸则在它们与所述脂质体解离时进行裸核酸的吞噬作用。可使用传统无源负载技术,如乙醇点滴法(如在SALP中)、逆向蒸发法、和乙醇稀释法(如在SNALP中),将核酸截留在脂质体内部。

[0118] 脂质体递送系统提供稳定的制剂,提供改善的药代动力学,以及某种程度的“被动”或“生理学”组织靶向。亲水性和疏水性材料如潜在的化疗剂的密封是已知的。见,例如, Schneider拥有的专利US 5,466,468,其揭露可肠道外给药的包含合成性脂质的脂质体制剂; Hostetler等人拥有的专利US 5,580,571,其揭露接合至磷脂质的核苷酸类似物; Nyqvist拥有的专利US 5,626,869,其揭露药物组合物,其中,药学活性的化合物是包含在所定义的脂质系统内的发夹或其片段,所述脂质系统包含至少一种两亲性的极性脂质组分以及至少一种非极性脂质组分。

[0119] 脂质体和聚合物泡囊可含有多种溶液和化合物。某些具体实施例中,本发明的复合物与聚合物泡囊偶合或密封在聚合物泡囊中。作为一类人工载体,聚合物泡囊是将容器密封在其内部的微小的中空球,使用两亲性合成嵌段共聚物形成所述载体膜而制作。常见的聚合物泡囊在其核心中含有水性溶液,且可用于密封并保护敏感性分子如药物、酶、其它蛋白和肽、以及DNA和RNA片段。所述聚合物泡囊膜提供物理屏障,所述屏障将所密封的材料与外部材料如见于生物学系统中的外部材料分离。聚合物泡囊可通过已知技术从双重乳液生成,见 Lorenceau et al., 2005, Generation of Polymerosomes from Double-Emulsions, Langmuir 21 (20):9183-6,所述文献通过引用并入本文。

[0120] 本发明的一些具体实施例中,非病毒载体可经修饰以实行靶向递送和转染。PEG化(即,使用聚乙二醇修饰表面)是一种有力的方法,用来降低非病毒载体的调理素作用和聚集性并且最小化由网状内皮系统造成清除,导致静脉(i.v.)给药后的循环寿命延长。PEG化的纳米颗粒因此一般被称为“潜行”纳米颗粒。未从所述循环中快速清除的纳米颗粒将有机会遭遇被感染的细胞。

[0121] 本发明的一些具体实施例中,使用对组织的独特环境和外部刺激产生应答的靶向控释系统。金纳米棒在近红外区域具有强吸收谱带,所吸收的光能随后通过金纳米棒被转化为热能,即所谓的“光热效应”。因为近红外光可深深地渗透入组织,金纳米棒的表面可使用核酸修饰而用于控释。当所述经修饰的金纳米棒被红外光辐射时,由于光热效应诱导的热变性而释放核酸。所释放的核酸的量取决于光辐射的功率和暴露时间。

[0122] 无论组合物是否作为核酸或多肽而给药,它们均以促进被哺乳动物摄取的途径配制。可用的载体系统和制剂揭示于上文中。一些具体实施例中,所述载体可将所述组合物递送至具体的细胞类型。但本发明并不限于此,其它DNA递送方法如使用诸如磷酸钙、DEAE葡聚糖、脂质体、脂质复合物、表面活性剂、和全氟液体化学品的化学转染也符合预期,还预期物理递送方法如电穿孔、显微注射、弹道颗粒、和“基因枪”系统。

[0123] 其它具体实施例中,所述组合物包含已经使用一种或多种 CRISPR/Cas载体和gRNA改造或转染的细胞。一些具体实施例中,本发明的方法可离体应用。换言之,可从受试者身体移除细胞,并在培养物中使用所述组合物处理所述细胞,以切除例如HIV序列,再将被处理的细胞返回受试者身体。所述细胞可以是所述受试者的细胞,或他们可以是单体型匹配的细胞或细胞系。可辐射所述细胞以防止复制。一些具体实施例中,所述细胞是人类白细胞抗原(HLA)匹配的细胞、自体同源的细胞、细胞系、或其组合。其它啥方式中,所述细胞可以是干细胞。例如,胚胎干细胞或人工多能干细胞(诱导多能干细胞(iPS 细胞))。已经从包括人类在内的多个动物物种构建了胚胎干细胞(ES 细胞)和人工多能干细胞(诱导多能干细胞,iPS细胞)。这些类型的多能干细胞将是用于再生医学的最有用的细胞来源,引物这些细胞能通过适当诱导其分化而分化为几乎所有器官,并保持其积极分裂的能力,同时维持其多能性。特别地,可从自源性体细胞构建iPS细胞,因此,与通过摧毁陪陪产生的ES细胞相比,前者不大可能造成伦理和社会问题。再者,作为自源性细胞的iPS细胞,令避免排异反应成为可能,而排异反应是再生医学或移植疗法的最大障碍。

[0124] 根据所建立的方法制备用于再输注的转导细胞。培养约2至4周的一段时间后,所述细胞的数目可以是 1×10^6 至 1×10^{10} 之间。就此而言,不同患者的细胞和不同类型的细胞的生长特征不尽相同。将所转导的细胞再输注之前约72小时,取等量小样,分析表型、和表达所述治疗剂的细胞的百分比。对于给药,可以通过所述细胞类型的LD₅₀、所述细胞类型在多种浓度下的副作用、以及患者的体重和整体健康状况,确定本发明的细胞的给药速率。给药可经由单剂量或分剂量实施。亦可使用外源性给药的因子而动员成人干细胞,所述因子刺激其成人干细胞的产生且出自可能包括但不限于骨髓或脂肪组织的组装或空间。

[0125] 因此,本发明涵盖一种方法,其消除被整合入离体培养的被HIV 潜伏性感染的宿主细胞基因组中的前病毒DNA,其中,前病毒HIV DNA 被整合入所述宿主细胞基因组中。所述方法包括下述步骤:获得被HIV 潜伏性感染的宿主细胞群;离体培养所述宿主细胞;使用一组合物出来所述宿主细胞,所述组合物包括CRISPR相关的核酸内切酶和至少一种与前病毒HIV DNA的LTR中的靶标序列互补的gRNA;以及,从所述宿主细胞基因组消除所述前病毒DNA。当加入下述额外的步骤时,相同的方法步骤也可用于处理所述被潜伏性感染的宿主细胞的供体:产生 HIV被消除的T细胞群;将所述HIV被消除的T细胞群输注如所述患者体内;以及,治疗所述患者。

[0126] 先前阐述的在实施例部分中揭示的慢病毒递送系统,是优选的用于CRISPR相关的核酸内切酶和gRNA在患者T细胞或其它被潜伏性感染的宿主细胞内进行离体转导的系统。或者,可采用任何适当的表达载体系统,包括但不限于,先前枚举的那些系统。

[0127] 已经被证实在离体处理被潜伏性感染的T细胞方面有效的组合物和方法,若通过一种或多种适当表达载体的手段递送时,非常可能在体内仍有效。因此,本发明涵盖一种药物组合物,用于令被整合在哺乳动物受试者细胞中的HIV DNA失活,所述组合物包括编码

CRISPR相关的核酸内切酶的单离的核酸序列、以及至少一种编码至少一种与前病毒HIV DNA的LTR中的靶标序列互补的gRNA的单离的核酸序列。优选包括gRNA A与gRNA B的组合。亦优选所述药物组合物还包括至少一种表达载体,在所述表达载体内,所述单离的核酸序列被编码。

[0128] 本发明还涵盖治疗被HIV感染的哺乳动物受试者的方法,所述方法包括下述步骤:确定哺乳动物受试者被HIV感染,给药有效量的先前阐述的药物组合物至所述受试者,以及,治疗所述受试者的HIV感染。

[0129] 可通过所述领域技术人员已知的途径来制备根据本发明的药物组合物。例如,上文揭示的核酸和载体可配制于组合物中,用于应用至组织培养中的细胞或用于给药至患者或受试者。这些组合物可以药学领域中周知的方式制备,且可通过多种途径给药,给药途径取决于是否为局部治疗或全身性治疗且取决于待治疗的面积。给药可以是外用给药(包括眼用和给药至黏膜,包括鼻内、阴道和直肠递送)、肺部给药(如,通过粉末或气溶胶的吸入或吹入,包括通过喷雾器吸入或吹入;气管内、鼻内、表皮和透皮给药)、眼部给药、口服给药或肠胃外给药。用于眼部递送的方法可包括外用给药(滴眼液);结膜下、眼周或玻璃体内注射;或通过经外科手术置于结膜囊内的气泡式导管或眼科插入体引入。肠胃外给药包括静脉、动脉、皮下、腹膜内或肌肉注射或输注;或颅内,如鞘内或心室内给药。肠胃外给药可以是快速灌注剂量形式,或可以是,例如,通过连续输液泵给药。用于外用给药的药物组合物和制剂可包括透皮贴剂、油膏、洗液、乳清剂、凝胶剂、滴剂、栓剂、喷雾剂、液体、粉末等。传统的药学载体、水性基、粉末基质、油性基、增稠剂等可能是必需的或所希望的。

[0130] 本发明还包括药物组合物,其含有,作为活性成分,本文中揭示的核酸和载体,与一种或多种药学可接受的载体组合。术语“药学可接受的”(或“药理学可接受的”)指的是,当酌情给药至动物或人类时,不产生副作用、过敏反应或其它不良反应的分子整体和组合物。本文中,术语“药学可接受的载体”包括可用作药学可接受的物质的介质的任何和全部溶剂、分散接枝、包衣、抗菌剂、等渗和吸收延迟剂、缓冲剂、赋形剂、粘合剂、润滑剂、凝胶、表面活性剂等。在本发明的组合物的制作中,所述活性成分典型与赋形剂混合,由赋形剂吸收或密封在例如胶囊、片、袋、纸、或其它容器的载体内。当所述赋形剂用作稀释剂时,其可以是固体、半固体、或液体材料(如,生理盐水),其作为所述活性成分的载具、载体或介质而发挥作用。因此,所述组合物可以是片剂、丸剂、粉末剂、菱锭剂、袋剂、扁囊剂、酞剂、混悬剂、乳剂、溶液剂、糖浆剂、气溶胶剂(作为固体或处于液体介质中)、洗剂、乳清剂、油膏剂、凝胶剂、软胶囊剂、硬胶囊剂、栓剂、无菌注射溶液剂、和无菌封装的粉末剂。如所述领域中已知的,稀释剂的类型可依据所预计的给药途径而变。所得组合物可包括其它剂,如防腐剂。一些具体实施例中,所述载体可以是,或可包括,基于脂质的胶体或基于聚合物的胶体。一些具体实施例中,所述载体材料可以是配制为脂质体、水凝胶、微粒、纳米颗粒、或嵌段共聚物胶束的胶体。注意,所述载体材料可形成胶囊,且所述材料可以是基于聚合物的胶体。

[0131] 一些具体实施例中,本发明的组合物可配制为纳米颗粒,例如,由与DNA复合的高分子量线性聚乙烯亚胺(LPEI)的芯和环绕所述芯的聚乙二醇改性(PEG化)低分子量LPEI的壳组成的纳米颗粒。一些具体实施例中,所述组合物被配制为其内部密封所述组合物的纳米颗粒。L-PEI已经被用来有效地在体内将基因递送至多种器官如肺、脑、胰腺、视网膜、膀胱中及肿瘤中。L-PEI能有效地在体外和体内浓缩、稳定化并递送核酸。所述核酸和载体亦

可应用于装置(如,导管)的表面或包含在泵、贴、或其它药物递送装置中。本发明的核酸和载体可单独给药,或在药学可接受的赋形剂或载剂(如,生理盐水)的存在下以混合物形式给药。基于给药模式和途径选择所述赋形剂或载剂。适用于药物制剂中的药物载剂及药学必需品可在《雷明顿药物科学》(Remington's Pharmaceutical Sciences (E.W.Martin))中找到,所述书是本领域内周知的参考书,也可在USP/NF(美国药典和国家处方集)中找到。

[0132] 一些具体实施例中,所述组合物可配制为纳米颗粒,所述纳米颗粒密封表达Cas9或变体Cas9的核酸和至少一种与靶标HIV的gRNA序列;或其可包括编码这些组分的载体。或者,所述组合物可配制为密封CRISPR相关的核酸内切酶的纳米颗粒,所述多肽由一种或多种本发明的核酸组合物编码。

[0133] 在治疗HIV感染的方法中,可使用标准临床测试如用来检测所述受试者血清中HIV抗体或HIV多肽p24的存在的免疫检验来鉴别受试者,或通过HIV核酸扩增检验鉴别。将提供至所述受试者并导致所述感染症状完全解决、所述感染严重性下降、或所述感染进展变慢的此组合物的量,视为治疗有效量。本发明的方法也可包括监控步骤以帮助优化给药剂量、时机以及预期后果。一些本发明的方法中,可首先确定患者是否具有潜伏性HIV感染,随后决定是否使用本文中揭示的一种或多种组合物治疗所述患者。

[0134] 当在潜在的宿主细胞中稳定表达时,本发明的组合物减少或防止新的HIV感染。示例性方法和结果揭露在实施例部分中。据此,本发明涵盖防止处于HIV感染风险下的患者的T细胞HIV感染的方法。所述方法包括下述步骤:确定患者处于HIV感染的风险下;令所述患者的T细胞接触有效量的表达载体组合物,所述组合物包括编码CRISPR 相关的核酸内切酶的单离的核酸、以及至少一种编码至少一种与HIV DNA的LTR中的靶标序列互补的gRNA的核酸;在所述T细胞内稳定地表达CRISPR相关的核酸内切酶和所述至少一种gRNA;以及,防止所述 T细胞的HIV感染。

[0135] 处于具有HIV感染风险下的受试者可以是,例如,从事不安全性行为的任何性活跃个体,即从事性活动而不使用避孕套的个体;具有另一性传播感染的性活跃个体;静脉注射吸毒者;或未受割礼的男性。处于具有HIV感染风险下的受试者也可以是,例如,其职业可能将他或她带至与HIV感染群体接触的个体,如医护人员或现场急救员。处于具有HIV感染风险下的受试者可以是,例如,监禁场所中的同住者或性工作者,换言之,以性活动换取或非金融品如食物、药物或庇护所的个体。

[0136] 本发明也包括便于先前阐述的治疗和预防HIV感染的方法的应用的试剂盒。所述试剂盒包括测得量的组合物,所述组合物包括至少一种编码CRISPR相关的核酸内切酶的单离的核酸序列、以及至少一种编码一种或多种gRNA的核酸序列,其中,每一gRNA包括与HIV前病毒的长末端重复序列(LTR)中的靶标序列互补的间隔序列。所述试剂盒还包括选自封装材料、包含使用说明书的封装插页、无菌液体、注射器和无菌容器所组成群组的一项或多项。gRNA A和gRNA B是优选的gRNA。一种优选的具体实施例中,所述核酸序列被包括在表达载体如实施例1中详述的慢病毒表达载体中。所述试剂盒还可包括适当的稳定剂、载剂分子、芳香剂等,视预期用途而定。

[0137] 实施例

[0138] 实施例1:材料和方法

[0139] 细胞培养

[0140] 1. 稳定的细胞系。先前已经揭示了Jurkat 2D10报告细胞系 (Pearson, et al., J Virol 82, 12291-12303 (2008)), 所述细胞系在含有10% FBS和庆大霉素(10 μ g/ml)的RPMI培养基中培养。2x 10⁶个细胞使用10 μ g对照pX260质粒或各5 μ g的pX260 LTR-A和pX260 LTR-B质粒电穿孔 (Neon System, Invitrogen, 3x 10ms/1350V脉冲)。48小时后, 以含有0.5 μ g/ml嘌呤霉素的培养基替换先前使用的培养基。进行一周的选择后, 移除嘌呤霉素, 令细胞再生长一周。之后, 将细胞稀释至10细胞/ml的浓度, 并以50 μ l/孔的量置于96孔板中。2周后, 使用Guava EASYCYTE迷你型流式细胞分析仪对单细胞克隆体进行GFP标签化的HIV-1报告基因再激活 (以PMA 25nM/TSA 250nM处理12小时) 筛选。使用非反应性的克隆体进行进一步分析。

[0141] 2. 原代CD4⁺细胞单离和扩展。通过CNAC基础科学核心I获得血沉棕黄层和患者血液样品 (天普大学药学院 (Temple University School of Medicine, Philadelphia))。使用Ficoll-Paque试剂通过密度梯度离心从人类外周血单离PBMC。使用HBSS缓冲液将血液的血沉棕黄层样品的体积调节为30ml, 轻柔地铺展在15ml的Ficoll-Paque液垫上, 以1500RPM离心30分钟。收集含有PBMC的层, 在HBSS缓冲液中洗涤3次并计数。使用人CD4⁺T细胞单离试剂盒 (Miltenyi Biotec) 进行CD4⁺T细胞的进一步单离。以生物素接合的抗体鸡尾酒混合物 (抗CD8、CD 14、CD 15、CD 16、CD 19、CD36、CD56、CD 123、CD235 α 、TCR γ / δ) 标记细胞 (10⁷个), 随后与接合有抗生物素抗体和抗CD61抗体的磁珠混合, 并使用MACS LS柱分离。收集溢流的未标记的呈递所述CD4⁺富集馏分的细胞, 并通过CD4-FITC FACS证实纯度 (94 至97% CD4⁺阳性, 见图17)。之后, 使用T细胞激活/扩展试剂盒 (Miltenyi Biotec), 根据制造商的方案扩展细胞。简单地说, 将2.5x 10⁶细胞/ml与涂覆抗CD2、CD3、CD28抗体的磁珠以细胞:磁珠为2:1 的比混合。2天后, 用移液器轻柔地吸取细胞以打破团块, 加入一体积适当新鲜制备的含有人rIL-2的生长培养基。每3天替换一次培养基。所有原代细胞均在以浓度为20U/ml的人rIL-2补充的具有10% FBS 和庆大霉素 (10 μ g/ml) 的RPMI中生长 (NIH AIDS试剂程序, AIDS、NIAID、NIH: 人rIL-2的部分来自Maurice Gately博士 (Hoffmann-La Roche Inc.))。所有牵涉AIDS患者和体内感染细胞的过程均在BL2⁺实验室内进行。

[0142] 慢病毒递送

[0143] 1. 克隆慢病毒方法。先前揭示了“一体化”含有LTR靶标A和B 的pX260-U6-DR-BB-DR-Cbh-NLS-hSpCas9-NLSH1-短tracr-PGK-puro (Addgene 42229) 载体 (Hu, et al, Proc Natl Acad Sci USA 111, 11461 -11466 (2014))。对于至原代细胞内的慢病毒递送, 将表达用于LTR 靶标A和B的gRNA的DNA节段缩短至20个核苷酸 (表1中第5部分), 并首先亚克隆入pX330-U6-Chimeric_BB-CBh-hSpCas9 (Addgene 42230) 的U6-嵌合-gRNA表达盒中。随后, 使用MluI/BamHI延伸的引物对所述全gRNA表达盒进行PCR扩增 (T560/T561, 见表1中第5部分)、消化、并插入pKLV-U6gRNA (Bbs1) -PGKpuro2ABFP (Addgene 50946) 的 MluI/BamHI位点内。

[0144] 2. 慢病毒封装和纯化。通过使用pMDLg/pRRE (Addgene 12251)、pRSV-Rev (Addgene 12253) 和pCMV-VSV-G (Addgene 8454) 共转染 HEK293T细胞, 将所得pKLV-U6-LTR A/B-PGKpuro2ABFP封装入慢病毒颗粒内。对于将Cas9封装入慢病毒颗粒内, 使用下述载体: pCW-Cas9 (Addgene 50661)、psPAX2 (Addgene 12260)、和pCMV-VSV-G (Addgene 8454)。对

于一些实验,使用pLV-EF1 a-Cas9v1-T2A-RFP慢病毒 (Biosettia Inc.)。在氯喹(50 μ M)的存在下,使用CaPO4沉淀法和封装慢病毒载体混合物以30 μ g总DNA/2.5x 10⁶细胞/100mm培养皿共转染HEK 293T细胞。次日,替换所述培养基,在24小时和48小时后收集上清液,以3000RPM澄清10分钟,经0.45 μ m过滤,并通过超速离心法浓缩(2小时,25000RPM,使用20%蔗糖液垫)。通过轻柔搅动过夜将慢病毒微丸再次悬浮在HBSS中,取等量小样,在HEK 293T细胞中测定滴度。通过FLAG免疫细胞化学测定pCW-Cas9慢病毒的滴度,通过BFP荧光显微镜检查术测定pKLV-U6-LTR A/B-PGKpuro2ABFP 慢病毒的滴度。

[0145] 3. 原代细胞的慢病毒转导。在转导之前24小时,替换生长培养基,通过使用涂覆有抗CD2/CD3/CD28抗体的磁珠(Miltenyi Biotec)以细胞/磁珠为2:1的比激活细胞。次日,使用12.5x 10⁵IU的 pCW-Cas9慢病毒,与25x 10⁵IU的pKLV空慢病毒或12.5x 10⁵IU的每一个pKLV-LTR靶标A及pKLV-LTR靶标B慢病毒一起(总MOI 15),使用慢病毒感染2.5x 10⁵细胞。在32 $^{\circ}$ C在150 μ l含有8 μ g/ml聚凝胺的培养液中,将细胞以2700RPM物种培养2小时,随后再次悬浮,放置4小时,接着加入150 μ l的生长培养基。次日,细胞在1ml的 PBS中洗涤3次,并在含有人rIL-2(20U/ml)的生长培养基中培养。

[0146] 病毒检验和检测

[0147] 1. 病毒原液。如下述过程制备HIV-1JRFL粗原液:制备使用HIV-1 感染6天的PBMC的上清液,以3000RPM澄清10分钟,并经0.45 μ m 过滤。通过以pNL4-3-EGFP-P2A-Nef质粒转染HEK 293T细胞,制备 HIV-1 L4-3-EGFP-P2A-Nef报告病毒,并对其处理以用作慢病毒原液(见上文)。使用Gag p24 ELISA测定HIV-1JRFL滴度,通过被感染的HEK 293T细胞的GFP-FACS测定HIV-1NL4-3-EGFP-P2A-Nef滴度。

[0148] 2. 体外HIV-1感染。将从原代PBMC制备的CD4⁺T细胞激活并扩展一周,之后进行HIV-1感染。通过下述完成感染:使用粗HIV-1原液,在300ng的Gag p24/10⁶细胞/1ml下,在32 $^{\circ}$ C的含有8 μ g/ml聚凝胺的无血清培养基中,以2700RPM物种培养2小时,随后再次悬浮并放置4小时,之后在PBS中洗涤3次,最终在含有人rIL-2(20U/ml)的生长培养基中培养。在CD4⁺T细胞感染的情形中,激活细胞并扩展一周,再进行HIV-1感染。通过使用经稀释的病毒原液在8 μ g/ml的聚凝胺的存在下进行所述细胞的简单过夜培养而不经物种培养,感染Jurkat 2D10细胞。

[0149] 3. HIV-1 DNA检测和定量。根据制造商的方案,使用NUCLEOSPIN Tissue试剂盒(Macherey-Nagel)从细胞分类基因组DNA。对于LTR特异性PCR(见表1第1部分),使用FAIL SAFE PCR试剂盒和缓冲液D(Epicentre),令100ng所提取的DNA进行PCR,PCR条件如下:98 $^{\circ}$ C, 5分钟,30次循环(98 $^{\circ}$ C, 30秒;55 $^{\circ}$ C, 30秒;72 $^{\circ}$ C, 30秒),72 $^{\circ}$ C反应7分钟,并在2%琼脂糖凝胶中解析。使用Long Range PCR试剂盒(Qiagen)对250ng基因组DNA实施整合位点特异性PCR(见表1第2部分),PCR条件如下:93 $^{\circ}$ C, 3分钟,35次循环(93 $^{\circ}$ C, 15秒;55 $^{\circ}$ C, 30秒;62 $^{\circ}$ C, 7.5分钟)。令PCR产物进行琼脂糖凝胶电泳,凝胶纯化,克隆入TA载体(Invitrogen)中,并送至Sanger测序(Genewiz)。使用对于HIV-1 Gag基因为特异性的TAQMAN qPCR并使用细胞 β -球蛋白基因作为参考,将HIV-1 DNA定量(见表1第6部分)。在qPCR之前,将来自被感染细胞的基因组DNA稀释至10ng/ μ l,随后每反应/孔取 5 μ l(=50ng)。根据来自M.K.Liszewski et al,Methods,47(4): 254-260(2009)的简化步骤,使用铂Taq DNA聚合酶(Invitrogen)制备反应混合物。从U1细胞(NIH AIDS试剂程序,分化的AIDS、NIAID、NIH:

HIV-1感染的细胞(U1),来自Dr.Thomas Folks,Folks,et al, Science 238,800-802 (1987))基因组DNA的连续稀释制备标准品,由于每二倍体基因组含有两种单拷贝的HIV-1前病毒,等于 β -球蛋白基因拷贝数。qPCR条件为,对于Gag基因:98℃,5分钟,45次循环(98℃,15秒;62℃,30秒,同时撷取);对于 β -球蛋白基因:98℃,5分钟,45次循环(98℃,15秒;62℃,30秒,同时撷取;72℃,1分钟)。进行反应,并使用LightCycler480(Roche)分析数据。

[0150] 4.p24 ELISA。通过令来自感染的细胞的上清液进行p24 Gag抗原捕获ELISA(ABL Inc.),定量感染水平。对于归一化,记录总细胞数和上清液体积。

[0151] 宿主基因组分析

[0152] 1.基因组DNA制备,全基因组测序和生物信息学分析。验证来自亲本2D10T细胞的单一亚克隆对照C11和实验性AB5的靶标切割效率和对HIV-1EGFP报告基因再激活的功能性阻抑。根据制造商的方案,使用NUCLEOSPIN Tissue试剂盒(Macherey-Nagel)单离所述基因组DNA。所述基因组DNA提交至Novogene生物信息学研究所(novogene.com/en/)用于WGS和生物信息学分析。简单地说,进一步在1%的琼脂糖凝胶上鉴别DNA品质,使用NANOPHOTOMETER®分光光度计(IMPLEN,CA,USA)检查DNA纯度,并使用QUBIT® DNA检验试剂盒在QUBIT®2.0荧光计(Life Technologies,CA,USA)中测量DNA浓度。采用每样品总量为1.5 μ g的DNA,使用Truseq Nano DNA HT样品制备试剂盒(Illumina USA)并遵循制造商建议生成测序文库,且加入索引编码加以确定每一样品的序列归属。通过超声波处理将DNA样品片段化为350bp大小,随后将DNA片段末端打磨、A-截短尾、并连接至全长度转接物用于Illumina测序,并进行进一步PCR扩增。最后,纯化PCR产物(AMPure XP系统),通过Agilent2100生物分析仪分析所述文库的尺寸分布,并使用实时PCR定量。在cBot聚类生成系统上,使用Hiseq X HD PE聚类试剂盒(Illumina),根据制造商的说明书,实时被索引编码的样品的聚类分析。在聚类生产后,在Illumina Hiseq X Ten平台上对所述文库中的制剂测序,并生成配对端解读。通过碱基读出将最初的原始数据转化为测序解读并记录在FASTQ文件中,所述文件含有序列信息(解读)和相应的测序品质信息。在使用转接物滤除任何解读后(≥ 10 核苷酸与所述转接物比对,允许 $\leq 10\%$ 的错配), $\geq 10\%$ 的未鉴别核苷酸、 $> 50\%$ 的具有 < 5 的phred品质的碱基、或推定的PCR副本,对照样品的总计342.67Gb清洁解读(约109.25x覆盖率)和AB5样品样品的总计369.55Gb(112.72x)被保留用于进一步组装。采用Burrows-Wheeler Aligner(BWA)软件(Li and Durbin,Bioinformatics 25,1754-1760(2009))将配对端清洁解读映射至参考的人基因组(UCSC hg19)和HIV-1基因组(KM390026.1)。随后,使用Picard Samtools(Li H,et al.Bioinformatics 25,2078-2079(2009); broadinstitute.github.io/picard/),GATK(DePristo,Banks,et al. Nat Genetics 43,491-498(2011))和Samtools(Li,Handsaker et al.Bioinformatics 25,2078-2079(2009))来进行副本移除、局部再比对、和碱基品质再校准,以生成最终的用于计算序列覆盖率和深度的BAM文件。使用Python和PyVCF(0.6.0版本)以及PyFasta软件包(0.5.0版本)以若干准则过滤备选的indel。通过比较对照物(C11)与实验组(AB5)之间的差异,关注Cas9/LTR-gRNA(AB5组)对宿主基因组的潜在脱靶效应。通过muTect检测所述SNP(Cibulskis,Lawrence, et al.Nat Biotechnol 31,213-219(2013)),通过Strelka检测所述indel(Saunders,Wong,et al.,Bioinformatics 28,1811-1817(2012)),并通过CREST检测结构变体(SV)(Wang,Mullighan,et al. Nat Methods 8,652-654(2011))。独特的存

在于AB5组中的indel的总数为32,399,且通过公用数据库(dbSNP)(Sherry,Ward,et al. Nucl Acids Res 29,308-311(2001))和杂合indel过滤。随后,如先前所揭示,提取所述indel位点的上游300bp(600bp)至下游300bp(600bp)的序列(Hu,Kaminski,et al. Proc Natl Acad Sci USA 111, 11461-11466(2014);Veres,Gosis,et al. Cell Stem Cell 15, 27-30(2014))。提取从所述indel位点的上游300bp(600bp)至下游300bp(600bp)的序列,随后与预计的潜在脱靶序列LTR-A/B+NRG。同样,SV分析检测到AB5组中的42处删除和10处插入,并将±300bp(600bp)处的提取序列与预计的脱靶序列LTR-A/B+NRG比较。为了确定HIV-1的整合位点,使用CREST(Wang,Mullighan,et al. Nat Methods 8,652-654(2011))检测与HIV-1基因组相关的对照样品的SV。

[0153] 2. Surveyor检验。使用SURVEYOR突变检测试剂盒(Transgenomic)并根据制造商的方案,测试来自6个预计脱靶位点的PCR产物中突变的存在(表1第1部分)。简单地说,将异种PCR产物在95℃变性10分钟,通过使用热循环仪的逐步冷却进行杂交。之后,在0.25μl的SURVEYOR增强子S和15mM的MgCl₂S的存在下,在42℃,使用0.25μl的SURVEYOR核酸酶将300ng的杂交DNA(9μl)消化4小时。随后,加入终止液,将样品与未消化的对照物一起在2%琼脂糖凝胶中解析。

[0154] 3. 逆转录和PCR。使用RNeasy试剂盒(Qiagen)从Jurkat细胞提取总RNA,并进行柱DNase I消化。之后,使用0.5μg的RNA进行M-MLV逆转录反应(Invitrogen)。对于gRNA表达筛选,将特异性反向引物(pX260-crRNA-37R,表1,第3部分)用于RT反应中,随后进行使用靶标A或B正义寡核苷酸作为正向引物(表1,第5部分)的标准PCR,再进行琼脂糖凝胶电泳。对于检查邻近基因,将表达寡dT引物混合物用于RT中,使用mRNA特异性引物对并使用和b-肌动蛋白作为参考,令cDNA进行SYBERGREEN实时PCR(Roche)(表1,第4部分)。

[0155] 流式细胞分析

[0156] 使用Guava EASYCYTE迷你型流式细胞仪(Guava Technologies),在活体细胞中定量Jurkat 2D10细胞中的GFP和RFP表达。对于HIV-1报告基因病毒滴度,将HEK 293T细胞在感染48小时后胰蛋白酶化,在4%多聚甲醛中洗涤并固定10分钟,随后在PBS中洗涤3次,进行GFP FACS分析。通过使用CD4V5FITC抗体(BDBiosciences)直接标记,再通过FACS,检查原代T细胞中的CD4表达。

[0157] 膜联蛋白检验

[0158] 将Jurkat细胞洗涤、计数并在PBS中稀释至 1×10^5 细胞/ml的密度。对于每一样品,将100μl的细胞悬浮液与100μl的室温膜联蛋白V-PE染色试剂(Guava Nexin Reagent)混合,并在暗处于室温培养20分钟。培养后,使用Guava EasyCyte迷你型流式细胞仪撷取样品。

[0159] 使用碘化丙啶染色评估细胞活力。将PI溶液加入200μl的活体细胞悬浮液中,直至PI最终浓度为10μg/ml。样品在暗处于室温培养5分钟。培养后,使用Guava EASYCYTE迷你型流式细胞仪撷取样品。

[0160] 细胞周期分析

[0161] 细胞以1x PBS洗涤,随后再次悬浮在250μl的室温1x PBS中。将这一悬浮液逐滴加入1ml的-20℃88%乙醇中,直至乙醇的最终浓度为70%。细胞在-20℃固定过夜,随后洗涤,使用10μg/ml的碘化丙啶和100μg/ml的RNase A溶液在1x PBS中于37℃培养30分钟。所述

样品随后于4℃冷却,并使用Guava EASYCYTE迷你型流式细胞仪摄取。

[0162] 西方墨点法,免疫细胞化学

[0163] 通过下述制备全细胞裂解液:在冰上将Jurkat细胞在TNN缓冲液(用于哺乳动物细胞,50mM Tris pH 7.4、150mM NaCl、1%Nonidet P-40、5mM EDTA pH 8、1x蛋白酶抑制剂鸡尾酒混合物(Sigma))中培养30分钟,随后通过在4℃全速旋转10分钟而预澄清。将50μg的裂解液在1x Laemli缓冲液中变性,通过SDS-聚丙烯酰胺凝胶电泳在Tris-甘氨酸缓冲液中分离,之后转移至硝基纤维素膜(BioRad)上。所述膜在5%牛奶/PBST中封闭1小时,随后以鼠抗-flag M2单克隆抗体(1:1000,Sigma)或鼠抗-α-微管蛋白单克隆抗体(1:2000)培养。以PBST洗涤后,使用接合羊抗鼠抗体(1:10,000)将所述膜在室温培养1小时。使用Odyssey红外成像系统(LI-COR Biosciences)扫描所述膜并分析。将细胞在4孔腔载片中培养,次日以4%多聚甲醛/PBS固定10分钟。洗涤3次后,使用鼠抗-flag M2单克隆抗体(1:1000,Sigma)将细胞在0.1%Triton X-100、2%BSA/PBS中于室温培养2小时。洗涤3次后,使用羊抗鼠FITC二次抗体(1:200)培养细胞,随后使用Hoechst 33258培养5分钟。以PBS冲洗3次后,抗衰落水性封片剂(Biomed)封盖细胞,并在Leica DMI6000B荧光显微镜下分析。

[0164] 统计分析。所呈现的±SD来自三次实验,并通过学生t测试或ANOVA和Newman-Keuls多项比较试验评估。通常,将p值<0.05或0.01视为统计学显著。

[0165] 表1显示本研究中使用的DNA寡核苷酸的序列。

[0166] 表1

[0167]

引物	序列
1. PCR	
LTR -453/S	5'-TGGAAGGGCTAATCACTCCCAAC-3' (SEQ ID NO: 7)
LTR -374/S	5'-TTAGCAGAACTACACACCAGGGCC-3' (SEQ ID NO: 8)
LTR +43/AS	5'-CCG AGAGCTCCCAGGCTCAGATCT-3' (SEQ ID NO: 8)
LTR -417/S	5'-GATCTGTGGATCTACCACACACA-3' (SEQ ID NO: 10)
LTR -19/AS	5'-GCTGCTTATATGTAGCATCTGAG-3' (SEQ ID NO: 11)
RRE/S	5'-CGCCAAGCTTGAATAGGAGCTTTGTTCC-3' (SEQ ID NO: 12)
RRE/AS	5'-CTAGGATCCAGGAGCTGTTGATCCTTAGG-3' (SEQ ID NO: 13)
LTR-A-0T-1/S	5'-GTGGACTTTGGATGGTGAGATAG-3' (SEQ ID NO: 14)
LTR-A-0T-1/AS	5'-GCCTGGCAAGAGTGAAGTGAAGTC-3' (SEQ ID NO: 15)
LTR-A-0T-2/S	5'-AAG ATAATGAGTTGTGGCAGAGC-3' (SEQ ID NO: 16)
LTR-A-0T-2/AS	5'-TCTACCTGGTAATCCAGCATCTGG-3' (SEQ ID NO: 17)
LTR-A-0T-3/S	5'-ATAGGAGGAAGGCACCAAGAGGG-3' (SEQ ID NO: 18)
LTR-A-0T-3/AS	5'-AATGATGCTTTGGTCTACTCCT-3' (SEQ ID NO: 19)
LTR-A-0T-4/S	5'-TGCTCTTGCTACTCTGGCATGTAC-3' (SEQ ID NO: 20)
LTR-A-0T-4/ AS	5'-AATCTACCTCTGAGAGCTGCAGG-3' (SEQ ID NO: 21)
LTR-A-0T-5/S	5'-TCAGACACAGCTGAAGCAGAGGC-3' (SEQ ID NO: 22)
LTR-A-0T-5/AS	5'-ATGCCAGTGTCAAGTAGATGTCAG-3' (SEQ ID NO: 23)
LTR-A-0T-6/S	5'-TCAAGATCAGCCAGAGTGCACATG-3' (SEQ ID NO: 24)
LTR-A-0T-6/AS	5'-TGCTCTTCCGAGCCTCTCTGGAG-3' (SEQ ID NO: 25)
b-肌动蛋白 S	5'-CTACAATGAGCTGCGTGTGGC-3' (SEQ ID NO: 26)
b-肌动蛋白 AS	5'-CAGGTCCAGACGCAGGATGGC-3' (SEQ ID NO: 27)
2. 长范围 PCR	
D10-Chr1-5' Arm/F (6-29)	5'-GAGCACAGGACTCATTCAACAGT-3' (SEQ ID NO: 28)
D10-Chr1 -3' Arm/R (276-250)	5'-TTTGTATGTCAACAGACAGTATCCAG-3' (SEQ ID NO: 29)
D10-Ch16 MSRB1-S	5'-TGTG ATACTTCGAGCGGCT-3' (SEQ ID NO: 30)
D10-Ch16 MSRB1-AS	5'-GG AAAGGCGGGAGCTGATGA-3' (SEQ ID NO: 31)
3. gRNA RT 和 PCR	
pX260-crRNA-3' /R	5'-TGGGACCATTCAAAAACAGCAT-3' (SEQ ID NO: 32)

[0168]

4. 邻近基因 qPCR

RSBN1/F	5'-GTAAGGCCAGGAGAACAGATG-3' (SEQ ID NO: 33)
RSBN1/R	5'-TCAAAGAGAACTTCGCGGG-3' (SEQ ID NO: 34)
PHTF1/F	5'-CCCAAGTTGTGTCCATCCTATC-3' (SEQ ID NO: 35)
PHTF1/R	5'-AGACACCCCATACCCAAAC-3' (SEQ ID NO: 36)
MAGI3/F	5'-G ACACCGCAGTAATTCAGTTG-3' (SEQ ID NO: 37)
MAGI3/R	5'-AGCAAGACGAAGGATGAACAG-3' (SEQ ID NO: 38)
PTPN22/F	5'-TTTGCCCTATGATTATAGCCGG-3' (SEQ ID NO: 39)
PTPN22/R	5'-GTTGTAGATAAAGGACCCTGGG-3' (SEQ ID NO: 40)
AP4B1-AS1/F	5'-AGAAGGAAAAGGAGCAGACAC-3' (SEQ ID NO: 41)
AP4B1-AS1/R	5'-AGAAAGTGGAGGTGCTGTG-3' (SEQ ID NO: 42)
HS3ST6/F	5'-CTTCTACTCAACGCCACCA-3' (SEQ ID NO: 43)
HS3ST6/R	5'-AAGGGCCGGTAGAACTCC-3' (SEQ ID NO: 44)
RPL3L/F	5'-AACAATGCATCCACCAGCTA-3' (SEQ ID NO: 45)
RPL3L/R	5'-GTAATGACCCGCTTCTTGGT-3' (SEQ ID NO: 46)
MSRB1/F	5'-GAAGCTTAGGCCACATCTC-3' (SEQ ID NO: 47)
MSRB1/R	5'-CTGGAAGGGTTGACCAGAG-3' (SEQ ID NO: 48)
NDUFB 10/F	5'-GC ATGTATGAAGCCGAAATG-3' (SEQ ID NO: 49)
NDUFB 10/R	5'-TGAAGTCTCCACTTCCTTG-3' (SEQ ID NO: 50)
RPS2/F	5'-GCCTCTCTCAAGGATGAGGT-3' (SEQ ID NO: 51)
RPS2/R	5'-CAACAAATGCCTTGAACCTG-3' (SEQ ID NO: 52)
b-肌动蛋白 S	5'-CTAC AATGAGCTGCGTGTGGC-3' (SEQ ID NO: 53)
b-肌动蛋白 AS	5'-C AGGTCCAGACGCAGGATGGC-3' (SEQ ID NO: 54)

5. 靶标 A 和 B 寡核苷酸和克隆

LTR-A S/F/5'	5'-CACCGATCAGATATCCACTGACCTT-3' (SEQ ID NO: 55)
LTR-A S R/3'	5'-AAACAAGGTCAGTGGATATCIGATC-3' (SEQ ID NO: 56)
LTR-B AS/F/5'	5'-CACCGCAGCAGTTCTTGAAGTACTC-3' (SEQ ID NO: 57)
LTR-B AS/R/3'	5'-AAACGAGTACTTCAAGAACTGCTGC-3' (SEQ ID NO: 58)
T560	5'-TATGGGCCACCGGTGAGGGCTATTTCCCATGATTC-3' (SEQ ID NO: 59)
T561	5'-TGTGGATCCTCGAGGCGGGCCATTTACCGTAAGTTATG-3' (SEQ ID NO: 60)

6. Taqman qPCR

HIV-Gag-RTfw	5'-CATGTTTTCAGCATTATCAGAAGGA-3' (SEQ ID NO: 61)
HIV-Gag-RTrev	5'-TGCTTGATGTCCCCCACT-3' (SEQ ID NO: 62)
HIV-RT 探针	5' /56-FAM/-CCACCCACAAGATTTAAACACC-BHQ-3' (SEQ ID NO: 63)
b-球蛋白 RTfw	5'-CCCTTGGACCCAGAGGTTCT-3' (SEQ ID NO: 64)
b-球蛋白 RTrev	5'-CGAGCACTTCTTGCCATGA-3' (SEQ ID NO: 65)
b-球蛋白 RT 探针	5'-FAM-GCGAGCATCTGTCCACTCCTGATGCTGTTATGGGCGCTCGC-TAM A-3' (SEQ ID NO: 66)

[0169] 实施例2:用于抑制人类T淋巴细胞中潜伏HIV-1的再激活的 CRISPR/Cas9系统

[0170] Cas9/gRNA抑制人类T细胞中潜伏性HIV-1的HIV-1再激活。以确定根据本发明的CRISPR/Cas9系统能否消除人类T淋巴细胞系2D10 中的HIV-1基因组为目标,实施最初的实

验。这些细胞中存在单轮 HIV-1PN-3 的被整合拷贝,所述HIV-1PN-3的基因组缺乏编码大多数Gag-Pol多蛋白的序列但涵盖全长度5'LTR和3'LTR,且包括编码替换潜伏状态的Nef的标记蛋白绿色荧光蛋白(GFP)(图1A)。因为被整合的前病毒的均匀特性,2D10是初次构建HIV-1根除的原理论证的适当细胞系。使用促炎剂如肉豆蔻酸乙酸佛波酯(PMA)及/或HDAC抑制剂曲古抑菌素A(TSA)处理稳定表达Cas9但不表达gRNA的克隆2D10细胞,深深地刺激HIV-1启动子活性,导致在超过90%的被处理细胞内产生所述病毒蛋白和GFP(图1B,左图),提供便利的用于研究病毒潜伏性和再激活的细胞培养模型。分别设计为以跨越nt-287/-254(gRNA A)和nt-146/-113(gRNA B)处的LTR U3区域的所有病毒隔离群中的高保守序列为靶标的Cas9与gRNA A/B的共表达(图1A),完全消除了PMA/TSA诱导的GFP产生,表明对表达Cas9表达质粒和gRNA表达质粒两者的预先选择的T细胞混合克隆群体中HIV-1基因表达的抑制(图1B,右图)。分别通过RT-PCR和西方墨点法鉴别gRNA和Cas9的表达(图1C、1D)。通过对随机选择的具有或不具有gRNA表达的Cas9阳性克隆细胞的GFP产生的流式细胞分析检测显示,HIV-1表达被完全从表达Cas9表达质粒和gRNA表达质粒两者的细胞中消除(图7A)。另外发现,在仅表达单一gRNA(A或B)的很多克隆体中,GFP产生被有效阻断至与通过共表达A与B两者所得者类似的水平(图7B;也见图7A),证明任一单一构型的gRNA的表达可启动两个LTR处的裂解,从而实现前病毒DNA的根除。

[0171] HIV-1前病毒DNA在人类T细胞中的整合位点以及将病毒DNA从宿主细胞染色体中切除。通过2D10细胞的全基因组测序(WGS),核实HIV-1前病毒DNA整合的位点。采用结构变异(SV)的CREST(“剪裁揭示结构”(clipping reveals structure))检测(Wang, Mulligham, et al, Nat Methods 8, 652-654 (2011))来评估由前病毒DNA整合入宿主基因组造成的断点,并使用hg19基因组和HIV-1基因组KM390026.1作为参考基因组来读出所述DNA序列。通过CTX鉴别、设计了4种与HIV-1 DNA相关的染色体间易位(图8A、8B)。检测到HIV-1 5'LTR与P163.3:1991382之间及HIV-1 3'LTR与P613.3:1991378之间的断点,映射至蛋氨酸亚砷还原酶B1 MSRBI基因(NM_01332)的外显子2,并对应于用于2D10细胞中前病毒的先前映射位置(Pearson, R. et al. J Virol 76, 11091-11093 (2002); Jadowsky, J.K. et al. Mol Cell Biol 34, 1911-1928 (2014))。此外,两个CTX被映射至具有位于P13.2: 114338315与HIV-1 5'LTR之间的断点以及位于HIV-1 3'LTR与P13.2: 114338320之间的另一断点的染色体1。另外应注意,染色体1PI 3.2中两个断点之间的4个核苷酸TAAG被删除。染色体1中先前未被交联剂加成映射检出的HIV-1前病毒被整合入圆形精子细胞碱性蛋白1(RSBN1)基因(NM_018364)的第二内含子(114339984-114320431)中。对用于HIV-1 DNA整合入染色体1和染色体16中的位点的被鉴别的一致序列的图解显示在图8A、8B中。

[0172] LTR DNA的短范围扩增检验显示,对照细胞中的预期的497-bp DNA片段,以及在使用Cas9/gRNA A和B处理后相似大小(504bp)的第二DNA片段(图2A)。所述PCR扩增的直接DNA测序结果提供证据,证实经Cas9/gRNA处理的细胞中观察到的504-bp DNA片段是通过将残留5'LTR接合至经由Cas9/gRNA B裂解后剩余的3'LTR而创建的(图2B)。在所述克隆细胞的5'及3'融合位点的接合中,也检测到了亦具有7个核苷酸的插入的Indel突变(图2B)。对应于位于所述病毒基因组中心的Rev应答元件(RRE)的257-bp PCR扩增缺失,证实Cas9/gRNA B移除了跨越所述两个长末端重复序列之间的DNA序列(图2A)。使用源自RSBN1的第二内含子的成对引物对表达Cas9但不表达gRNA的2D10对照细胞进行长范围PCR分析,结果

证实了存在一6130-bp DNA片段,所述片段对应于被整合的HIV-1基因组加上源自其染色体1的侧翼DNA序列(图2C)。还存在呈递来自染色体1其它拷贝的宿主细胞DNA序列的264核苷酸DNA片段(在凝胶谱图底部显示)。在使用Cas9/gRNA A和B处理的细胞中,对应于所述被整合的HIV-1基因组的具有6130个核苷酸的DNA片段完全不存在。相反,PCR扩增产生了一具有909个核苷酸的较小DNA片段。对所述扩增子的测序证实,跨越5'LTR的B域与3'LTR的B域之间的被整合的病毒DNA被切除(图9A、9B)。再一次,检测到来自另一染色体的宿主基因组扩增的一具有264个核苷酸的DNA片段(图2C)。

[0173] 通过使用对应于MSRB1基因的第二外显子的引物对的长范围PCR检查染色体16中是否存在HIV-1前病毒DNA,并比较其在经Cas9/gRNA A/B处理的细胞中的状态。结果显示,染色体16中不存在预期的HIV-1基因组的5467-bp DNA片段及其侧翼宿主DNA。相反,检测到一较小的759-bp DNA片段,所述片段反映了经gRNA A裂解后的5'LTR中残留U3区域与经gRNA B裂解后的3'LTR中剩余U3区域的接合(图2D)。对所述759-bp DNA片段的直接测序鉴别了病毒DNA切除的位点(图10A、10B)。发现一由扩增来自染色体16另一拷贝的宿主DNA而导致的较小的110-bp DNA片段。这些观察结果提供强有力的证据,证实所使用的基因编辑分子有效地消除了HIV-1基因组的被整合的前病毒DNA的多个拷贝,所述多个拷贝分散于多条染色体中。

[0174] 从宿主细胞消除跨越5'LTR与3'LTR之间以及断点位置之间的HIV-1 DNA序列。为进一步验证所述基于Cas9/gRNA处理的基因编辑策略在从潜伏性感染的T细胞中消除HIV-1前病毒DNA的效率,使用GATK检测(Depristo, et al., Nat Genetics 43, 491-498 (2011))相对于参考HIV-1 DNA (GenBank保藏号KM3900261)分析对照细胞和HIV-1被根除的细胞中,所述HIV-1基因组中插入/删除(InDel)和单核苷酸多态性(SNP)的出现。与图2A至2D中显示的结果一致,来自全基因组测序的解读映射至5'-LTR和3'-LTR以及所述对照2D10细胞中的前病毒基因组,通过基因组测序,支持HIV-1 DNA的深度覆盖的精确性和可靠性所述(图3A)。在经Cas9/gRNA处理的2D10细胞中,被整合的基因组观点(IGV)显示,具有映射至3'LTR的解读且对应于HIV-1前病毒DNA的大DNA片段被完全移除(图3B)。映射至两个LTR之间的整体前病毒基因组的解读完全不存在,证实宿主细胞中被整合的HIV-1基因组的两个拷贝被完全消除,且证实单一克隆细胞中的Cas9/gRNA表达可达到100%基因编辑/消除,可能归因于通过稳定表达的Cas9/gRNA进行的重复基因组编辑。再者,这些结果证实,在病毒5'LTR和3'LTR裂解后且切除基因组后,病毒DNA很可能被降解,因此,宿主基因组中不出现转录或再整合。

[0175] 为了确定两个LTR的Cas9/gRNA A/B诱导裂解后后的修复事件,使用对来自HIV-1被切除的细胞的DNA中的结构变体(SV)的BWA检测(Wang, et al, Nat Methods 8, 652-654 (2011)),且所述检测鉴别了大的插入及/或删除的断点。结果证实,没有来自一条染色体的被切除的HIV-1 DNA被插入宿主基因组中及/或另一条染色体上的被整合前病毒DNA拷贝中,进一步排除所述被切除的病毒DNA再次整合入宿主细胞基因组中的见解。但是,3个断点被鉴别为由于删除对应于整合入宿主基因组中的病毒DNA位点的DNA片段而造成的。一个左侧断点位于5'LTR末端的核苷酸636(=HIV:9710)处,由10个解读支持。一个右侧断点展现两种模式,一种位于HIV:9073(=HIV:-3),由具有2个C→G转化和4个C→T转化的6个解读支持;另一种位于HIV:9075(=HIV:-1),由63个解读支持(如3C、3D)。注意,这两个断点可

实际反映,在通过Cas9/gRNA A或B在5'LTR和3'LTR切除全部前病毒DNA并随后将所述DNA在所述裂解位点再次接合后,LTR整体 634个核苷酸的存在。第三断点位于3'LTR中部的核苷酸9389(=HIV: 313)处,具有由87/161解读支持的C插入和由69/161解读支持的 CTAAGTT插入(图3E)。这一断点表示裂解后的DNA在5'LTR和3'LTR 的位点A和位点B处的接合(图3F)。

[0176] 切除HIV-1前病毒DNA对于邻近基因表达和脱靶效应的效果。调查CRISPR/Cas9介导的从RSBN1基因切除HIV-1前病毒DNA的影响。测定来自RSBN1和几种很靠近前病毒插入位点的其它细胞基因的RNA 生产的水平,如图4A中所示。5种对照物和5种HIV-1被根除的单细胞克隆体的RT-PCR的结果表明,对RSBN1的表达水平没有显著效果,但在邻近RNA的水平中检测到了小于0.4倍的较小变动(图4B),这可能无法归因于基因编辑策略且可能不影响其蛋白质的整体表达。同样,从染色体16消除HIV-1基因组显示,对整合位点即MSRB1基因及其周围基因的表达没有显著影响(图4C、4D)。基于若干与细胞健康相关的参数,包括细胞活力、细胞周期进展和细胞凋亡,使用几种通过 Cas9/gRNA A和B根除HIV-1后的克隆细胞,评估Cas9/gRNA A和B 的效果。未发现在通过Cas9/gRNA A和B消除HIV-1前病毒DNA后对宿主细胞生命迹象产生持久且显著的有害效应(图11A、11B、12A、12B、13A、13B)。

[0177] 为扩展潜在脱靶的分析的范畴,比较使用Cas9/gRNAs A/B系统处理的2D10细胞进行且结论为完全根除前病毒HIV-1 DNA的全基因组测序的InDel结果。为了提高InDel检测信心,寻求100x覆盖的全基因组测序,但统计学分析显示,实际达成的总覆盖率,对于对照细胞为 109.3x,而对于HIV-1被根除的细胞为112.7x(表2)。对于每一染色体的覆盖水平不同,范围是,对于染色体1为>96x,且对于染色体16 为>110x(图14)。使用人类(hg19)基因组作为参考序列,对于InDel,在对照物(+Cas9/-gRNA)中鉴别出1,361,311个(<50bp插入/删除)且在HIV-1被根除的细胞中鉴别出1,358,399个(表3);以及,对于单核苷酸多态性(SNP),在对照物中鉴别出3,973,098个且在HIV-1被根除的细胞中鉴别出3,961,395个(表4)。对照物与HIV-1被根除的细胞之间的比较生物信息学分析,鉴别出32,399体InDel(通过Strelka 检出的小插入/删除)、46,614体SNV(单核苷酸变异,通过MuTect检出)和52SV(包括通过CREST检出的大InDel在内的结构变异),后两组分布于不同的基因组区域内(表5)。

[0178] 在丢弃公用数据库dbSNP中找到的小InDel后,在HIV-1被根除的细胞中鉴别出30,156 InDel和43,858 SNV。滤除杂合突变,这一数字降为989 InDel。为了确定这些被过滤的InDel是否是由Cas9/gRNA A/B编辑系统造成的新生突变,检查各自过滤InDel的±30bp、±300 bp或±600bp侧翼序列,并使用Blastn(e-值截留:1000)将它们与由处于0至7个错配的序列相似性预计的潜在gRNA脱靶宿主基因组位点比较,且与HIV-1上靶序列比较。若无与gRNA A和B的靶标的任何错配,则在所提取的经过滤的InDel60,600和1200bp序列附近无脱靶位点。在所提取的60-bp序列内,无脱靶位点,但在来自PAM NRG(其必须是100%匹配)的>12核苷酸比对长度中存在7处错配。在所提取的600-bp序列中,无脱靶位点,但发现3处对于gRNA A或B的靶标的错配。由于4至7处错配,在来自PAM的20bp比对长度中仅发现一处具有6处错配的潜在脱靶位点,另一处位于来自PAM的12bp比对长度中的具有3处错配的靶标A潜在脱靶位点,以及一个额外的位于来自PAM的16bp长度的具有4处错配的靶标B潜在脱靶位点。在3 处错配的所提取的1200bp序列中,没有发现靶标A脱靶位点,但发现一处位于来自PAM的13bp的具有2处错配的靶标B潜在脱靶位点。遵循对于所述1200-bp序列的3至7处错配

的标准,仅发现6处靶标A潜在脱靶位点和2处靶标B潜在脱靶位点(图4E)。这些数据一起提供了强有力的证据,证实在HIV-1基因组被切除的细胞中检测到的indel无一处于任何潜在脱靶位点的靶标A或B的60bp内,与通过允许最高7处错配的检索标准所预测者相同。通过将检索序列扩展至600或1200bp,鉴别出相对罕见的脱靶位点,包括各种树木的错配和比对长度。由于与最后12bp种子序列加上PAM NRG的完美匹配,所述indel无一落入60至1200DNA序列的检索区域内。这些数据的整体诠释证实了这些细胞中以及其它细胞类型中的在前Surveyor检验结果(W Hu, et al. Proc Natl Acad Sci USA 111,11461-11466 (2014)),并通过非常严格的分析建立,通过Cas9/gRNA HIV-1 DNA切除系统无法产生对宿主T细胞基因组的脱靶效应。

[0179] HIV-1对HIV-1被根除的细胞的感染性。选择几种T细胞克隆体,其前病毒DNA通过Cas9/gRNA根除,并将Cas9和gRNA的表达维持在多种水平,以评估被HIV-1新感染的程度。如图15A中所示,克隆体C7表达Cas9但不表达gRNA B,而克隆体AB8未显示可检测水平的Cas9但含有gRNA B。选择具有等量的gRNA B和不同水平的Cas9表达的两种其它克隆体AB9和AB5进行再感染研究。将这些细胞以HIV-1 NL4-gFP感染,再通过流式细胞术进行病毒复制的纵向评估显示,单独表达Cas9或gRNA B的细胞可被HIV-1感染,且在這些研究进程的整个过程中支持病毒复制(感染后18天)(图15B)。相比之下,表达Cas9和gRNA B两者的细胞对HIV-1感染具有抗性,且不支持病毒复制。表达较高水平的Cas9的AB5,其病毒复制抗性似乎比显示减少的Cas9表达的AB9更高(图15A、15B)。图15C总结了图15B中显示结果的定量值。结果表明,Cas9和LTR引导的gRNA在细胞内的存在,有效防止了人类T细胞培养物不被HIV-1造成新感染。

[0180] 慢病毒介导的Cas9/gRNA递送阻抑CD4⁺T细胞的HIV-1感染。测试Cas9/gRNA阻抑从健康个体制备的CD4⁺T细胞的HIV-1感染的能力。因为慢病毒载体的高转导效率和低毒性,选择慢病毒载体递送Cas9表达DNA和gRNA表达DNA。LV转导的结果显示,HIV-1 LTR DNA被表达Cas9和gRNA两者的LV有效裂解,但使用仅表达Cas9的LV转导的对照细胞中无裂解(图5A)。注意,所述gRNA并不裂解缺少U3调节区域的LV LTR,因此,它们在LV基因组的表达上无效。据此,流式细胞分析显示,当使用LV-Cas9/gRNA转导时,潜伏性感染的T细胞中被整合的HIV-1基因组得以功能性失活(图5B)。再一次,未发现细胞死亡可能与原代细胞中Cas9/gRNA相关的证据,坚定了图11A、11B中显示的观察结果。一旦在T细胞系中通过LV进行编辑分子的基因递送的功效得以证实,即使用HIV-1_{JRFL}或HIV-1_{PNL4-3}感染CD4⁺T细胞的原代培养物,随后使用对照LV Cas9或者使用LV Cas9加上LV gRNA转导所述原代培养物(图5C)。与对照物相比,在经LV Cas9/gRNA处理的CD4⁺T细胞中见到HIV-1拷贝数的实质性减少(图5D)。在CD4⁺T细胞中,病毒DNA的扩增在所述对照细胞中显示预期的398-bp扩增,以及在经LV Cas9/gRNA转导的细胞中显示具有较低强度的类似大小的DNA片段(图5E)。

[0181] 在含有HIV-1基因组、从在天普大学医院AIDS临床中心常规随访的HIV-1⁺患者获得的PBMC和CD4⁺T细胞中,评估慢病毒递送的Cas9/gRNA的HIV-1基因组编辑能力。对于这一概念验证研究,最初,力图从4位患者(TUR0001至TUR0004;案例1至4)制备PBMC和CD4⁺T细胞,所述患者正在进行抗逆转录病毒治疗并展现对于治疗的不同应答,如通过病毒负载检验和CD4⁺细胞百分比所测定(图11A)。用来制备PBMC和CD4⁺T细胞、检查CD4⁺T细胞、以及慢病毒治疗时间轴和细胞收获的过程显示在图16B中。通过FITC接合抗CD4抗体的流式细胞分

析来验证CD4⁺T细胞的纯度(图16C)。

[0182] 使用慢病毒-Cas9和慢病毒-Cas9/gRNA转导PBMC的结果显示,表达Cas9和gRNA的细胞群的病毒拷贝数的实质性将是,在案例1中减少81%,在案例2中减少91%(图6A)。在CD4⁺T细胞的慢病毒转导后获得类似的结果,其显示,当表达Cas9和gRNA两者时,与仅表达Cas9的对照细胞相比,案例1中的病毒拷贝降低>92%,而案例2中降低56%(图6B)。用于β-球蛋白和Gag基因拷贝数的绝对定量的标准曲线和扩增图谱显示在图18A至18D中。对CD4⁺T细胞中Gag p24基因产生的检查证实,当使用慢病毒-Cas9/gRNA对所述细胞进行单转导时,与使用慢病毒-Cas9所见结果相比,案例1中的病毒复制减少71%,而案例2中减少62%(图6C)。另外,在通过慢病毒递送Cas9/gRNA后,检查从案例3和案例4获得的PBMC中的Gag p24的水平。这一研究的结果显示,在使用治疗性慢病毒转导所述细胞后,来自案例3和案例4的HIV-1p24产生分别减少39%和54%(图17)。

[0183] 之后,通过扩增所述病毒DNA并测序,评估所述患者样品中通过Cas9/gRNA引入的突变的特性。在案例1和案例2中,使用跨越-374/+43的引物进行的CD4⁺T细胞的初始基因扩增未能检测到任何谱带,在缺乏gRNA表达的对照样品中观察到DNA谱带(图6D)。这一观察结果证明,案例1中的HIV-1基因组序列可能不同于用于基因扩增的引物的基因组序列(图6D)。在案例2中,在未处理的细胞中检测出预期的DNA片段,通过Cas9/gRNA引入的突变可能已经消除了PCR引物对DNA序列的识别,因此干扰DNA扩增。使用另一套识别所述LTR的引物,导致在所有样品中产生预期的398核苷酸扩增子(图6E)。可能与来自使用Cas9/gRNA治疗的2D10细胞的结果(显示于图2A中)类似,见于gRNA表达的存在中的一些397核苷酸DNA片段,是在切除整个HIV-1编码序列后病毒LTR的剩余5'和3'序列接合的结果。对扩增子的测序证实了Cas9/gRNA在编辑预期位置处病毒基因组上的效果,且显示了LTR中PAM序列内及/或所述序列后的InDel和单核苷酸变异(SNV)突变的存在(图6F)。

[0184] 表2显示映射率和覆盖率。

[0185] 表2

[0186]

样品	+Cas9/+gRNA	+Cas9/-gRNA
总计	2304621804 (100%)	2153253838 (100%)
拷贝	42615862 (19.60%)	326664344 (15.50%)
映射	2175107441 (94.38%)	2108094482 (97.90%)
适当地映射	2133746358 (92.59%)	2057204364 (95.54%)
PE映射	2173896582 (94.33%)	2107021448 (97.85%)

[0187]

SE 映射	2421718 (0.11 %)	2146068 (01.10%)
具有映射至不同 chr 的配对物	930716 (0.41%)	9569904 (0.33%)
具有映射至不同 chr 的配对物 (mapQ \geq 5)	6944857 (0.30%)	7044381 (0.33%)
平均测序深度	112.72	109.25
覆盖率	99.67%	99.69%
至少 4X 覆盖率	99.48%	99.51%
至少 10X 覆盖率	99.00%	99.08%
至少 20X 覆盖率	97.29%	97.57%

[0188] 总计:总清洁解读的数目

[0189] 拷贝:复制解读的数目

[0190] 映射:映射至参考基因组的总解读的数目(百分比)

[0191] 适当地映射:映射至参考基因组且方向正确的解读的数目

[0192] PE映射:映射至参考基因组的成对端解读的数目(百分比)

[0193] SE映射:映射至参考基因组的单端解读的数目

[0194] 具有映射至不同chr的配对物:映射至不同染色体的配对解读的数目(百分比)

[0195] 具有映射至不同chr的配对物 (mapQ \geq 5):映射至不同染色体且 MAQ $>$ 5的配对解读的数目(百分比)

[0196] 平均测序深度:映射至参考基因组的平均测序深度

[0197] 覆盖率:所述基因组的序列覆盖率

[0198] 至少4X覆盖率:全部基因组碱基中,深度为 $>$ 4X的碱基的百分比[0199] 至少10X覆盖率:全部基因组碱基中,深度为 $>$ 10X的碱基的百分比[0200] 至少20X覆盖率:全部基因组碱基中,深度为 $>$ 20X的碱基的百分比

[0201] 表3显示不同基因组区域内插入/删除(InDel)的分布。

[0202] 表3

[0203]

样品	+Cas9/+gRNA	+Cas9/-gRNA	+Cas9/+gRNA 过量 +Cas9/-gRNA
CDS	1701	1746	164
移码_删除	866	910	124
移码_插入	279	275	33
非移码_删除	235	232	2
非移码_插入	187	196	0
终止增益	16	11	1
终止损失	1	1	0
未知	117	121	4
内含子	537492	538344	12229
UTR3	12154	12144	354
UTR5	1450	1446	58
剪接	498	500	20
ncRNA_外显子	2638	2629	78
ncRNA_内含子	71426	71581	1758
ncRNA_UTR3	389	388	16
ncRNA_UTR5	45	50	2
ncRNA_剪接	80	74	1
上游	9229	9256	209
下游	10296	10204	199
基因间	711001	712949	17310
总计	1358399	1361311	32399

[0204] 标注样品:样品名

[0205] CDS:外显子区域内InDel的数目

[0206] 移码_删除:造成蛋白质编码序列中移码变化的一个或多个核苷酸的删除。删除长度不是3的倍数。

[0207] 移码_插入:造成蛋白质编码序列中移码变化的一个或多个核苷酸的插入。插入长度不是3的倍数。

[0208] 非移码_删除:非移码删除,不改变编码蛋白框的删除,删除长度不是3的倍数。

[0209] 非移码_插入:非移码插入,不改变编码蛋白框的插入,插入长度不是3的倍数。

- [0210] 终止增益:导致变异位点处的终止密码子的即刻创建的移码插入/ 删除、非移码插入/删除或批量替换。
- [0211] 终止损失:导致变异位点处的终止密码子的即刻消除的移码插入/ 删除、非移码插入/删除或批量替换。
- [0212] 未知:未知功能(缘于数据库文献中基因结构定义中的多个错误)
- [0213] 内含子:内含子区域内InDel的数目
- [0214] UTR3:3'UTR区域中InDel的数目
- [0215] UTR5:5'UTR区域中InDel的数目
- [0216] 剪接:4bp剪接连接区域中InDel的数目
- [0217] ncRNA_外显子:非编码RNA外显子区域中InDel的数目
- [0218] ncRNA_内含子:非编码RNA内含子区域中InDel的数目
- [0219] ncRNA_UTR3:非编码RNA的3'UTR区域中InDel的数目
- [0220] ncRNA_UTR5:非编码RNA的5'UTR区域中InDel的数目
- [0221] ncRNA剪接:非编码RNA的4bp剪接连接区域中InDel的数目
- [0222] 上游:转录起始位点的1kb上游区域中InDel的数目
- [0223] 下游:转录结束位点的1kb下游区域中InDel的数目
- [0224] 基因间:基因间区域中InDel的数目
- [0225] 总计:InDel的总数目
- [0226] 表4列出不同基因组区域内的单核苷酸多态性(SNP)分布
- [0227]

样品	+Cas9/+gRNA	+Cas9/-gRNA	+Cas9/+gRNA 过量 +Cas9/-gRNA
CDS	29643	29982	1085
同义_SNP	13747	13894	347
错义_SNP	14924	15091	679
终止增益	378	394	43
终止损失	15	15	0
未知	579	588	16

[0228]

内含子	1369993	1374507	17921
UTR3	28450	28628	564
UTR5	6626	6659	193
剪接	845	882	32
ncRNA_外显子	12827	12918	175
ncRNA_内含子	206923	207516	2352
ncRNA_UTR3	809	824	11
ncRNA_UTR5	164	171	5
ncRNA_剪接	139	141	2
上游	26386	26577	493
下游	25707	25782	403
基因间	2252883	2258511	23378
总计	3961395	3973098	46614

[0229] 标注样品:样品名

[0230] CDS:外显子区域内InDel的数目

[0231] 同义_SNP:不造成氨基酸改变的单核苷酸改变

[0232] 错义_SNP:造成氨基酸改变的单核苷酸改变

[0233] 终止增益:导致变异位点处的终止密码子的即刻创建的错义_SNP

[0234] 终止损失:导致变异位点处的终止密码子的即刻消除的错义_SNP

[0235] 未知:未知功能(缘于数据库文献中基因结构定义中的多个错误)

[0236] 内含子:内含子区域内体SNP的数目

[0237] UTR3:3'UTR区域中体SNP的数目

[0238] UTR5:5'UTR区域中体SNP的数目

[0239] 基因间:基因间区域中体SNP的数目

[0240] ncRNA_外显子:非编码RNA外显子区域中体SNP的数目

[0241] ncRNA_内含子:非编码RNA内含子区域中体SNP的数目

[0242] 上游:转录起始位点的1kb上游区域中体SNP的数目

[0243] 下游:转录结束位点的1kb下游区域中体SNP的数目

[0244] 剪接:10bp剪接连接区域中体SNP的数目

[0245] ncRNA_UTR3:非编码RNA的3'UTR区域中体SNP的数目

[0246] ncRNA_UTR5:非编码RNA的5'UTR区域中体SNP的数目

[0247] ncRNA剪接:非编码RNA的10bp剪接连接区域中体SNP的数目

[0248] 总计:体SNP的总数目

[0249] 表5是SVN、InDel和SV的汇总

[0250] 表5

[0251]

	总 InDel	体 InDel ^a	dbSNP 筛选后	均聚物筛选后
+Cas9/-gRNA	1361311			
+Cas9/+gRNA	1358399	32399	30156	989

	总 SNV	体 SNV ^b	dbSNP 筛选后
+Cas9/-gRNA	3973098		
+Cas9/+gRNA	3961395	46614	43848

	总 SV ^c	体 SV ^d
+Cas9/-gRNA	3433	
+Cas9/+gRNA	3487	52

[0252] ^a体InDel意为:与通过Strelka检出的对照细胞系相比较,+Cas9/+gRNA中的具体InDel

[0253] ^b体SNV意为:与通过MuTect检出的对照细胞系相比较,+Cas9/+gRNA中的具体SNV

[0254] ^c总SV:仅包括通过Crest检出的SV类型的删除和插入

[0255] ^d体SV意为:与通过Crest检出的对照细胞系相比较,+Cas9/+gRNA 中的具体SV(删除和插入)

[0256] 讨论

[0257] 总之,结果显示,慢病毒递送的Cas9/gRNA A/B显着减少了来自 HIV-1感染患者的PBMC和CD4⁺T-细胞中的病毒拷贝数和蛋白质水平。使用被引导至所述LTR内的引物的PCR扩增并检测了未被完全从这些细胞删除的残留病毒DNA片段,所述片段仍受Cas9/gRNA的影响并含有邻近PAM序列的InDel突变。这些发现证实,CRISPR/Cas9在HIV-1 患者的PBMC中发挥有效的抗病毒活性。

[0258] ART治疗不能从被感染患者体内根除HIV-1,所述患者因此必须进行终身治疗。本文中揭示的新治疗策略将实现永久性缓解,其令患者可以停止ART,并降低随之而来的成本和潜在的长期副作用。所研发的从人类CD4⁺T细胞根除被整合的HIV-1拷贝的CRISPR/Cas9技术,抑制原代培养的人类CD4⁺T细胞中的HIV-1感染,且阻抑HIV-1⁺患者的外周血单克隆细胞(PBMC)和CD4⁺T细胞中的离体病毒复制。它们还解决了进一步的关键问题,提供这一基因编辑有效阻碍病毒复制而不对宿主DNA 造成基因毒性或诱发经由宿主细胞途径的破坏性效应的证据。这一研究中,作为第一步,使用克隆2D10细胞系作为人类T细胞潜伏性模型来建立:(i) the ability of Cas9/gRNA在使用超深全基因组测序而移除被整合的HIV-1 DNA拷贝的整个编码序列中的能力,以及(ii) 调查其与脱靶效应相关和细胞活力的安全性。一旦实现这些目的,所述研究的关注点即转移至原代细胞培养物以及患者样品上,以检查

CRISPR/Cas9在实验室设定中影响病毒DNA负载的效率。

[0259] 发现,CRISPR/Cas9编辑分散在染色体中的病毒DNA的多个拷贝。使用识别LTR U3区域内特异性DNA基序的Cas9加上gRNA A和B的对潜伏性感染的T细胞的联合治疗,有效消除了跨越所述两个LTR之间的整个病毒DNA片段。通过位于PAM上游的恰恰三个核苷酸处的位点的宿主DNA修复,将剩余染色体1中通过Cas9和gRNA B裂解5'LTR 和3'LTR的位点与染色体16中通过Cas9和gRNA A及B裂解的位点连接在一起。对经CRISPR/Cas9处理的HIV-1感染的2D10细胞的全基因组评估清楚地证实,被整合的病毒DNA被从RSBN1基因的第二内含子和MSRB1基因的外显子2消除。为了解决特异性和潜在的脱靶效应及副作用,通过全基因组测序和生物信息学分析实施空前详细水平的综合分析。这些透露了对照细胞的基因组中天然出现的突变和gRNA A 及B介导的HIV-1 DNA根除。所发现的突变包括天然出现的InDel、碱基切除和碱基替换,它们全部被预期或多或少地出现在培养物中快速生长的细胞内,所述细胞包括Jurkat 2D10细胞。关键问题是本文中的发现:由于没有序列被视为与任何此突变位点的1200个核苷酸内的gRNA A或B等同,这些突变无一缘于基因编辑系统。再者,这一用于 HIV-1 DNA切除的方法对邻近和远端细胞基因不具副作用,通过在所培养细胞中评估的所有体外测量,显示对细胞活力、细胞周期进展或增殖无影响,且并不诱发细胞凋亡,因此强烈支持其在转译期的安全性。发现,在若干通路后,Cas9和gRNA的表达水平削弱并最终消失,但只要Cas9和单或多态性gRNA存在,则对细胞的保护,以令其不产生新的HIV-1感染。

[0260] 另一待解决的关键转译可行性问题是,CRISPR/Cas9介导的HIV-1 根除能否防止或阻抑最相关人类和患者靶标细胞区中的HIV-1感染。发现,在来自HIV-1感染患者的PBMC和CD4⁺T细胞中,慢病毒递送的 Cas9/gRNA A/B显著降低了病毒拷贝数和蛋白水平。使用在LTR中导向的成套引物,这些细胞中,被扩增和检测的残留病毒DNA片段并未完全删除,仍受到Cas9/gRNA的影响并含有邻近PAM序列的InDel突变体。这些发现证实,CRISPR/Cas9在HIV-1患者的PBMC中发挥有效的抗病毒活性。还发现,经由慢病毒递送将Cas9/gRNAs A/B引入原代培养的人类CD4⁺HIV-1_{JRFL}或HIV-1_{NL4-3}感染的T细胞中,显著减少了病毒拷贝数,证实稳定整合的HIV-1 引导的Cas9和gRNA(与目前使用的gRNA A和B截然不同)赋予细胞系以对HIV-1感染的抗性。鉴于CRISPR/Cas9 可以被整合的DNA序列以及游离基因DNA序列为靶向的见解,很可能被整合的以及预整合的、可自由活动的细胞内HIV-1 DNA被Cas9/gRNA 编辑,此见解由Cas9/gRNA对多种人类病毒和任一构型的质粒DNA的编辑活性得以验证。

[0261] 注意,在这些研究的过程中,由于本研究的目标为确定病毒感染的生产阶段过程中对病毒阻抑的程度,在使用CRISPR/Cas9治疗之前不包括。观察到显著水平的阻抑,提供CRISPR/Cas9有效地令宿主染色体HIV-1 DNA的功能活性整合拷贝失能的证据。这一见解得到使用2D10 CD4⁺T细胞的观察记过的支持,在所述细胞中,被整合入染色体1和染色体16中的HIV-1的潜伏性拷贝被CRISPR/Cas9有效地消除。总之,本文中的发现综合且总结性地显示,人类T细胞中被宿主整合的HIV-1的整体编码序列得以根除,对于此系统对患者体内T细胞导向的HIV-1治疗的可译性提供强有力的支持。

[0262] 已经以例示性说明的方式解释本发明,且应理解,所使用的术语意图为所述词汇的说明特性而非限制。显然,鉴于上述教导,可对本发明做出多种修饰和改变。因此应理解,在所附权利要求书的范畴内,本发明可在具体揭示内容以外予以实践。

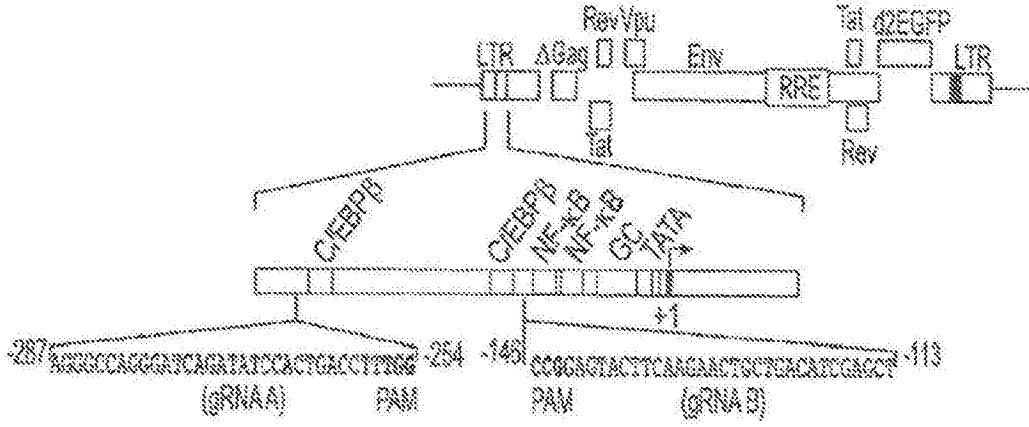


图1A

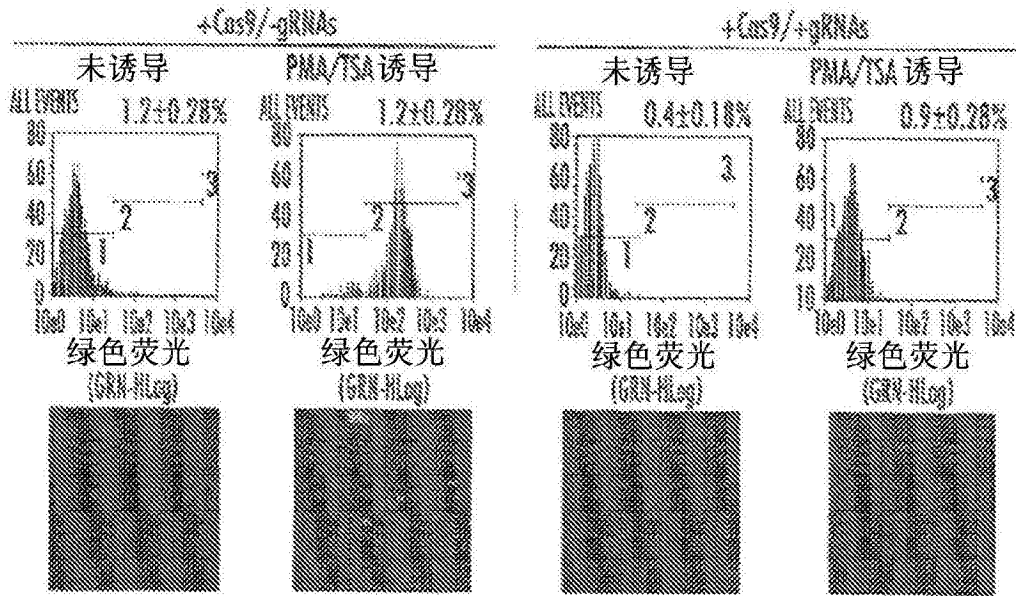


图1B

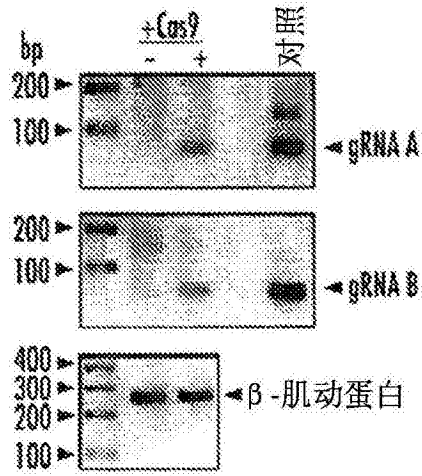


图1C

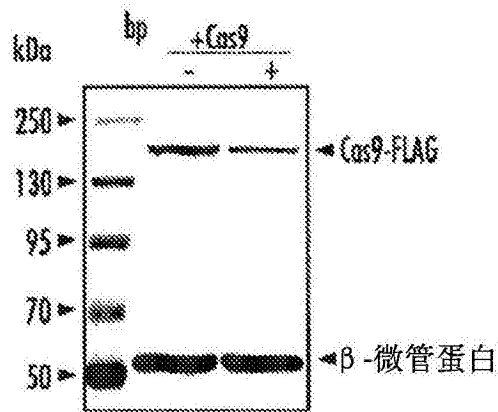


图1D

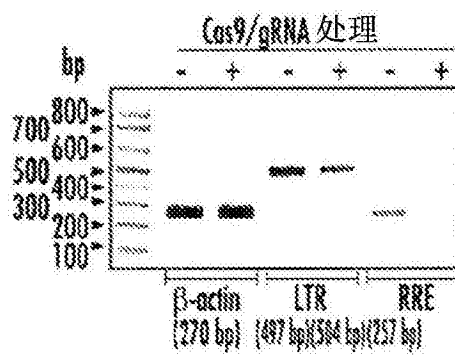


图2A

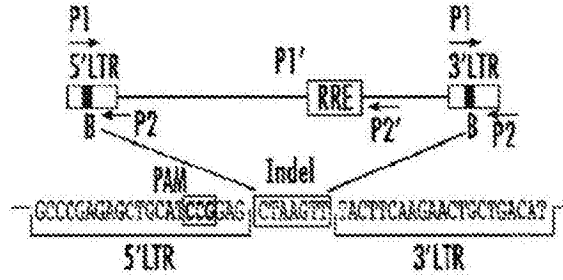


图2B

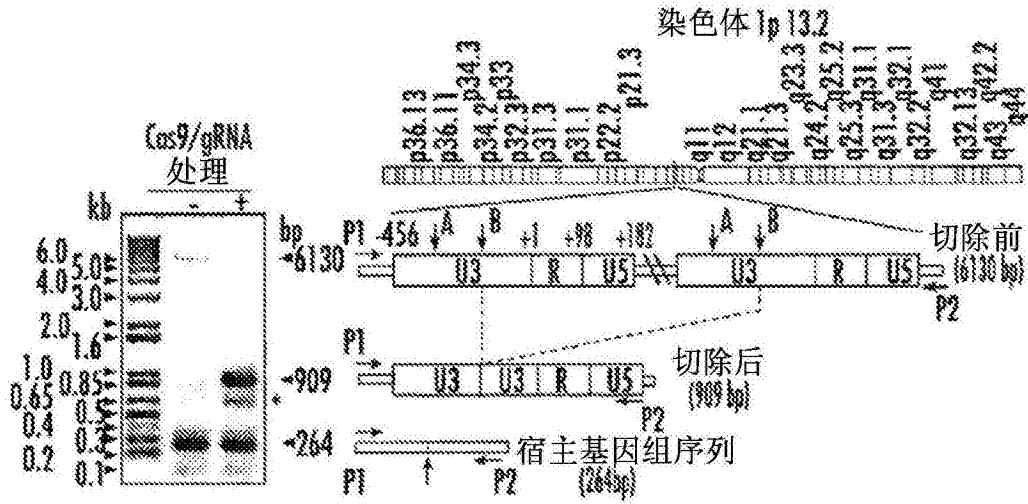


图2C

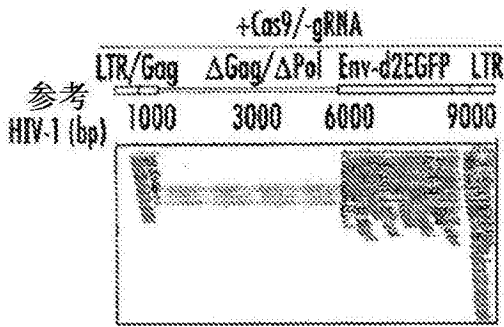
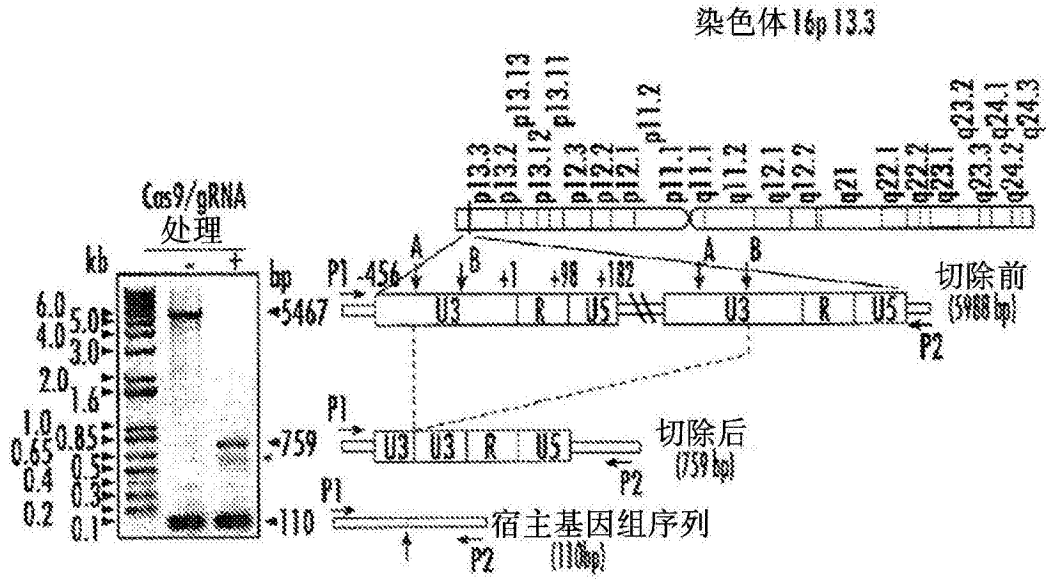


图 3A

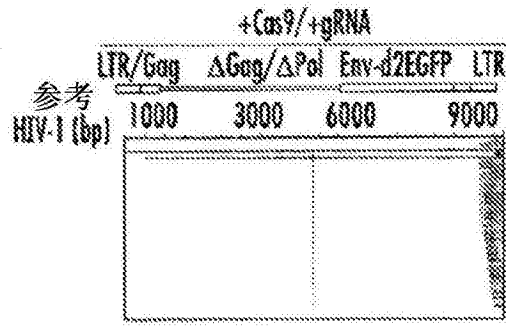


图 3B

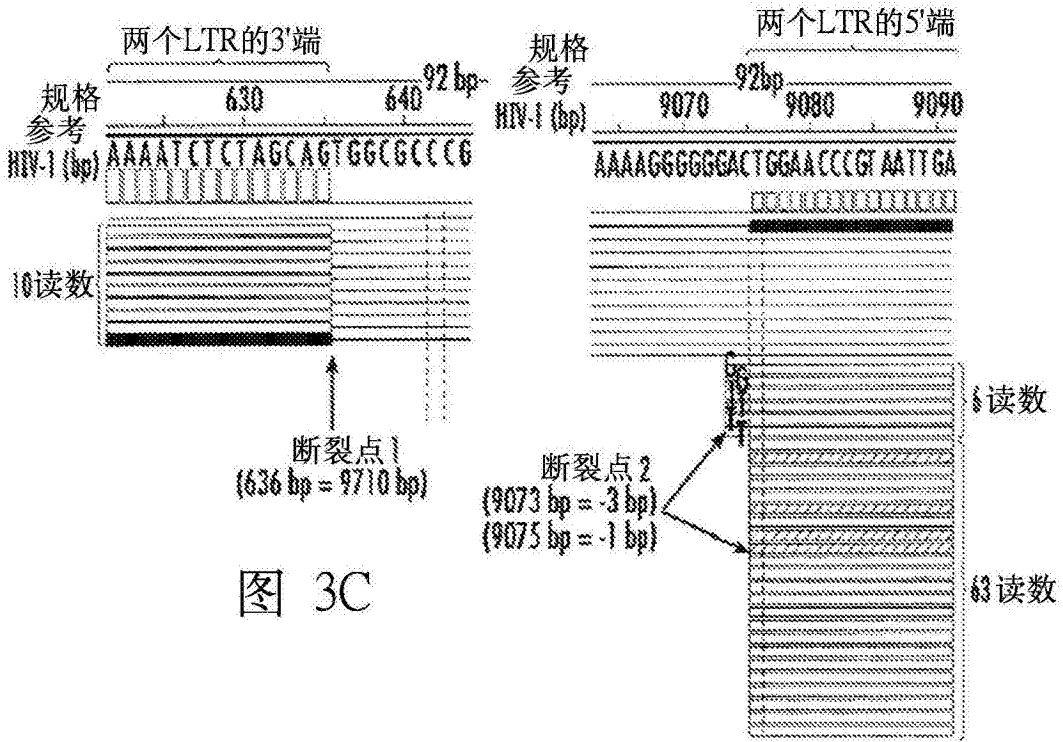


图 3C

图 3D

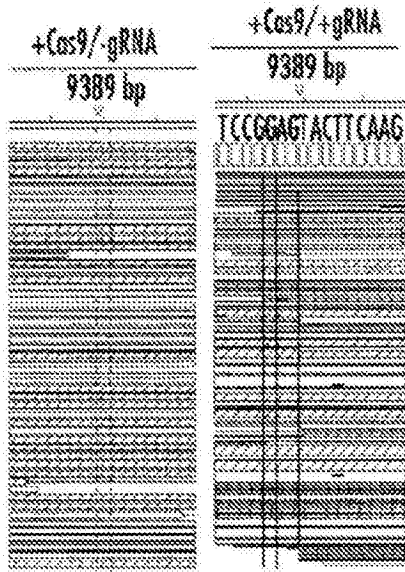


图 3E

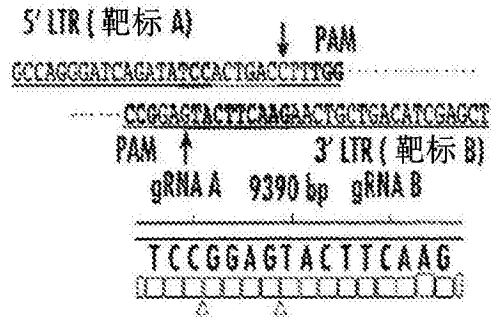


图 3F

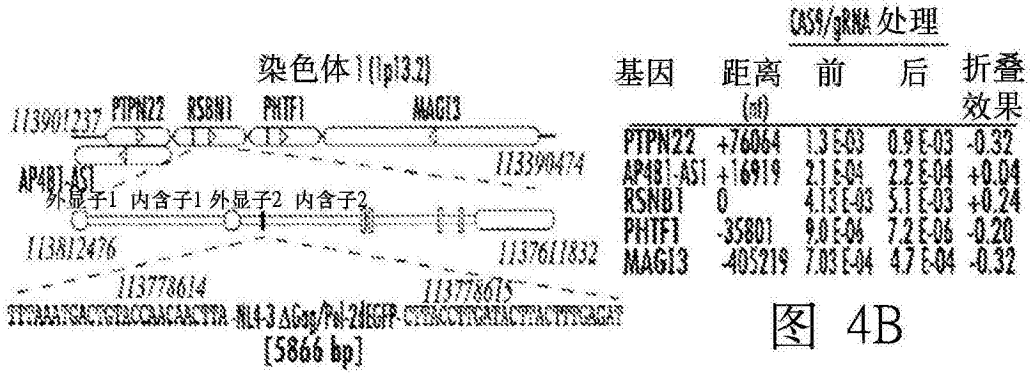


图 4A

图 4B

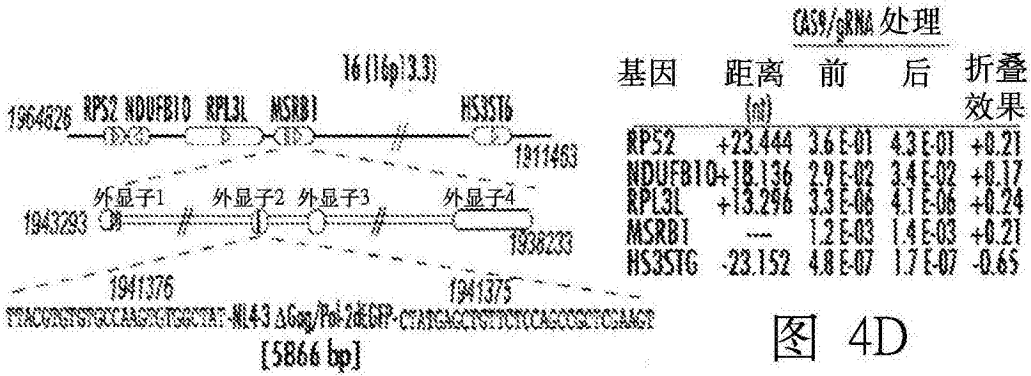


图 4C

图 4D

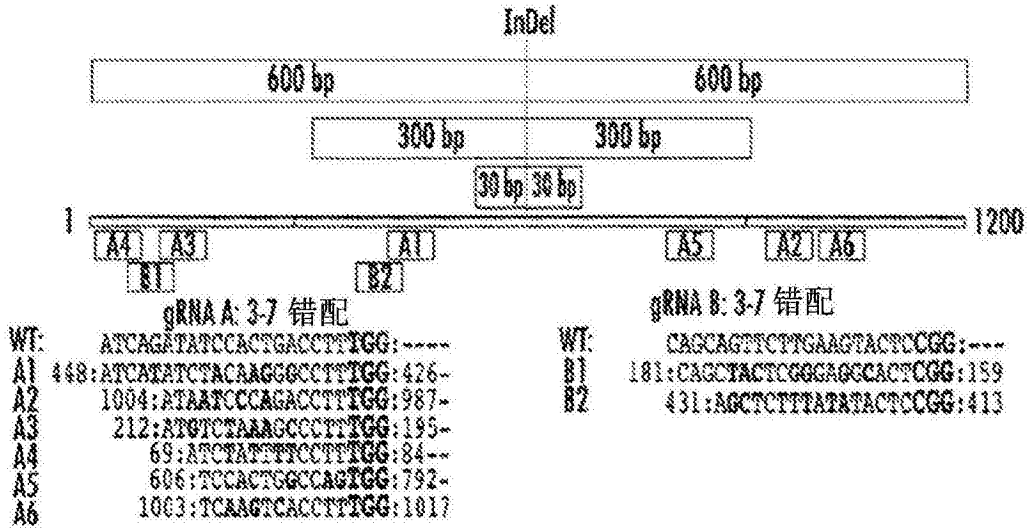


图4E

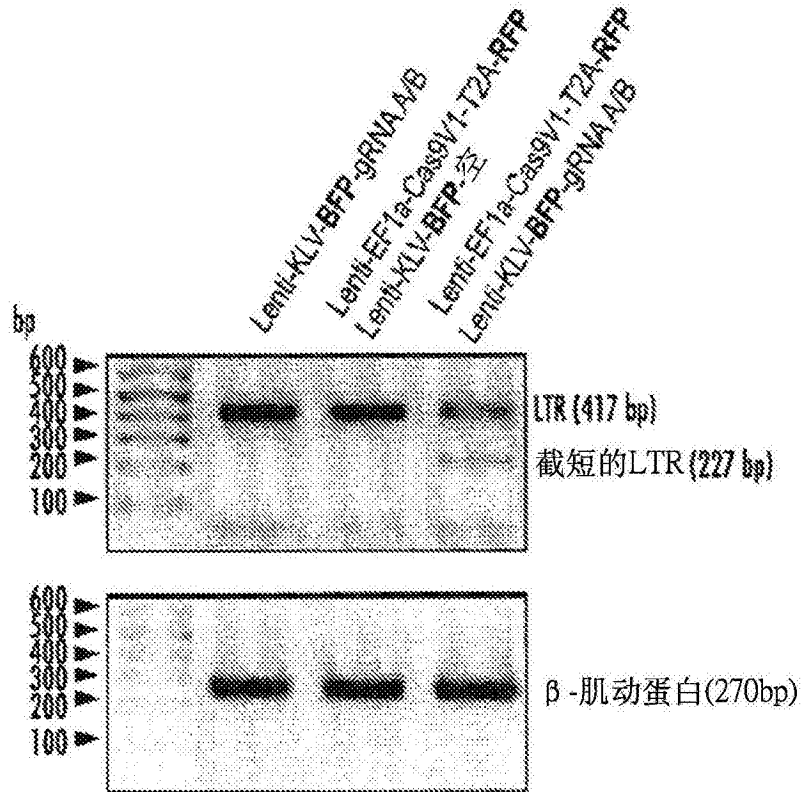


图5A

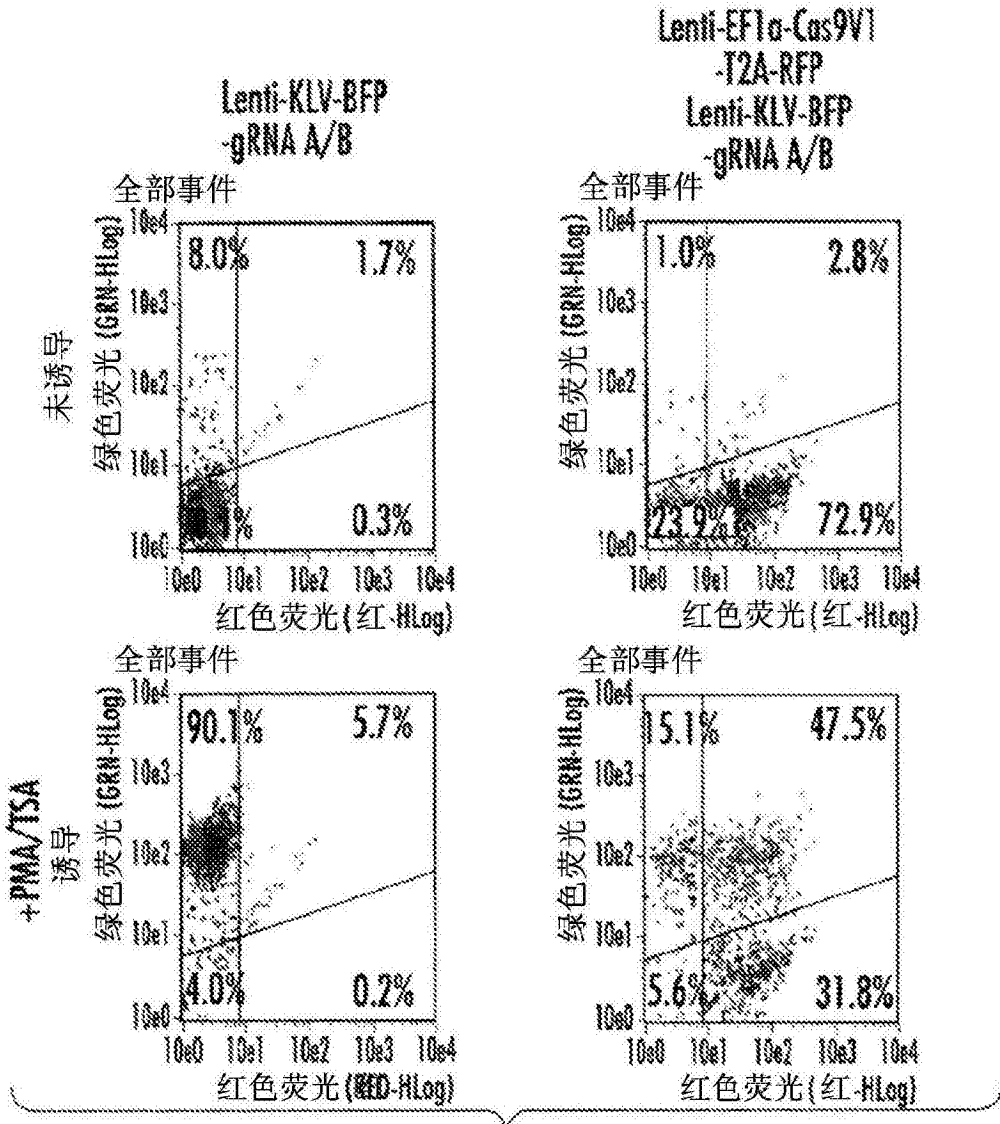


图 5B

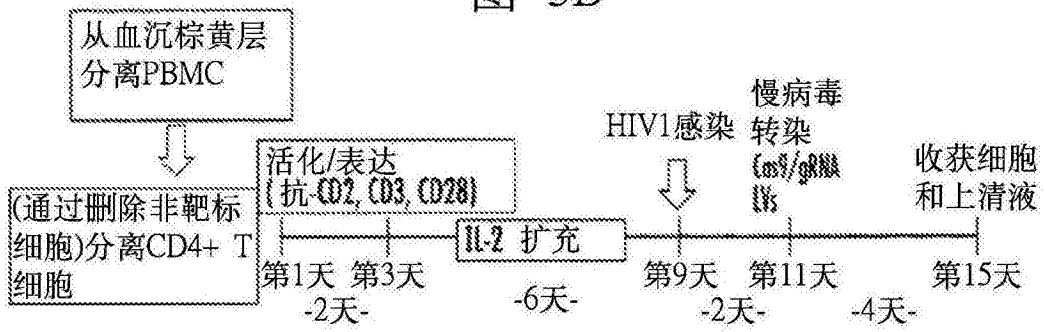


图 5C

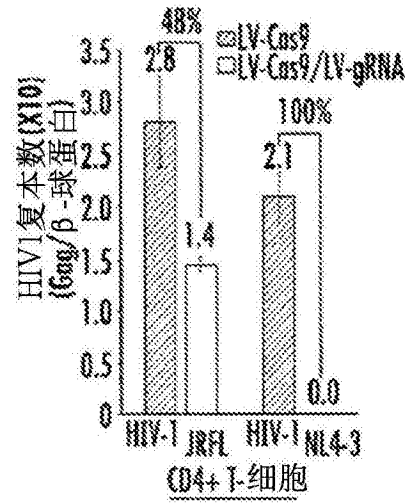


图5D

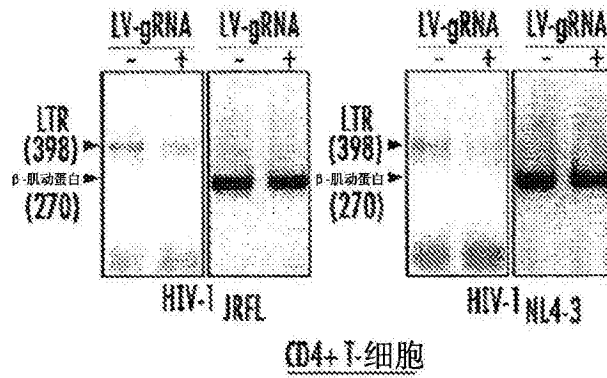


图5E

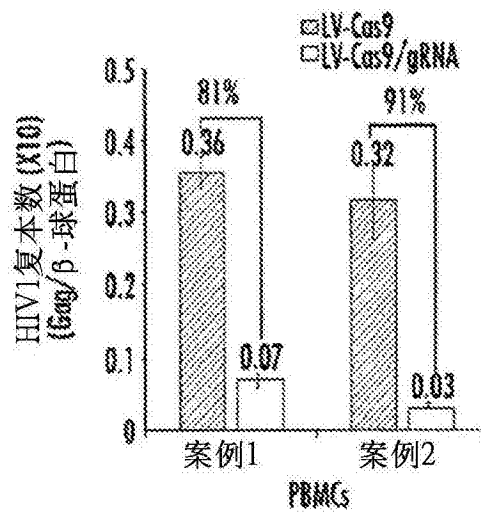


图6A

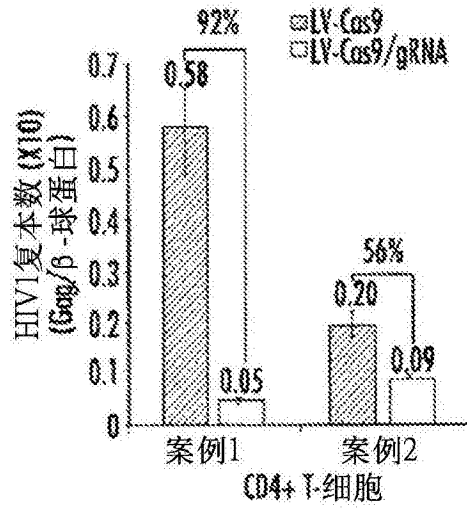


图6B

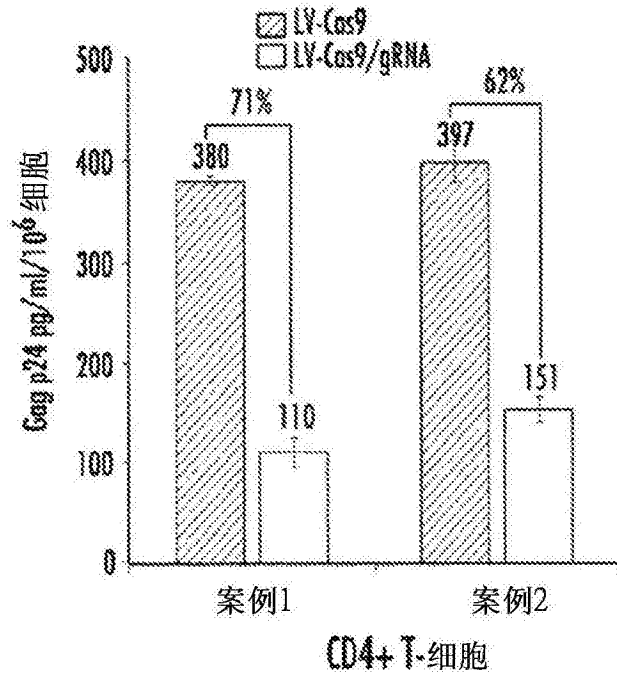


图6C

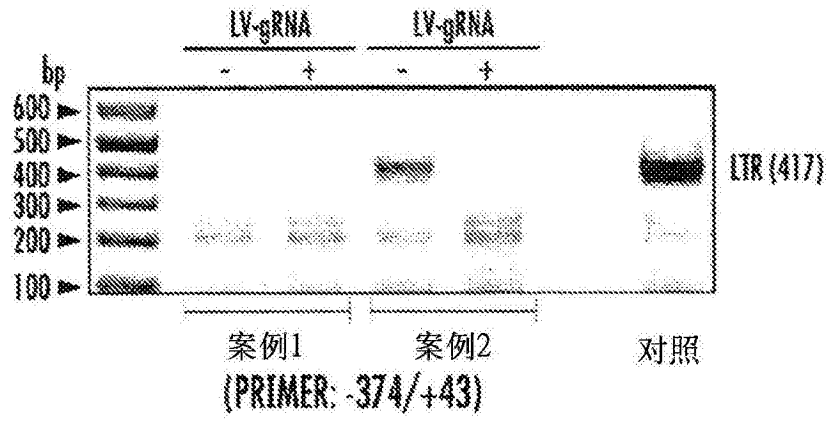


图6D

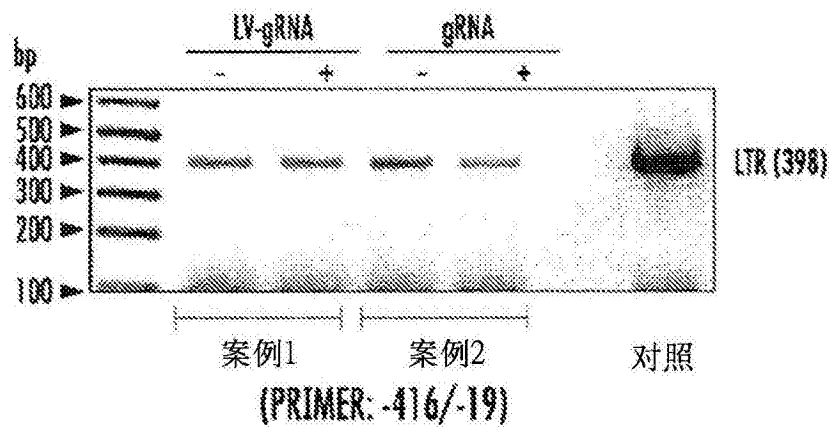


图6E

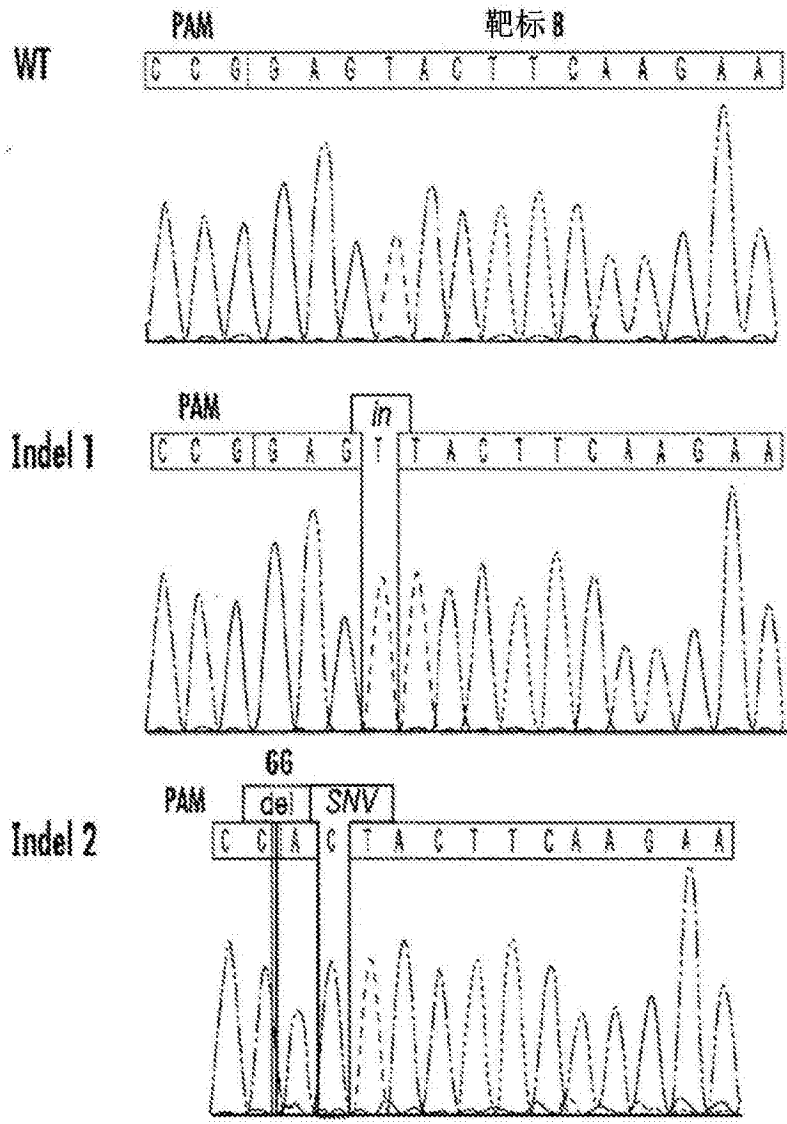


图6F

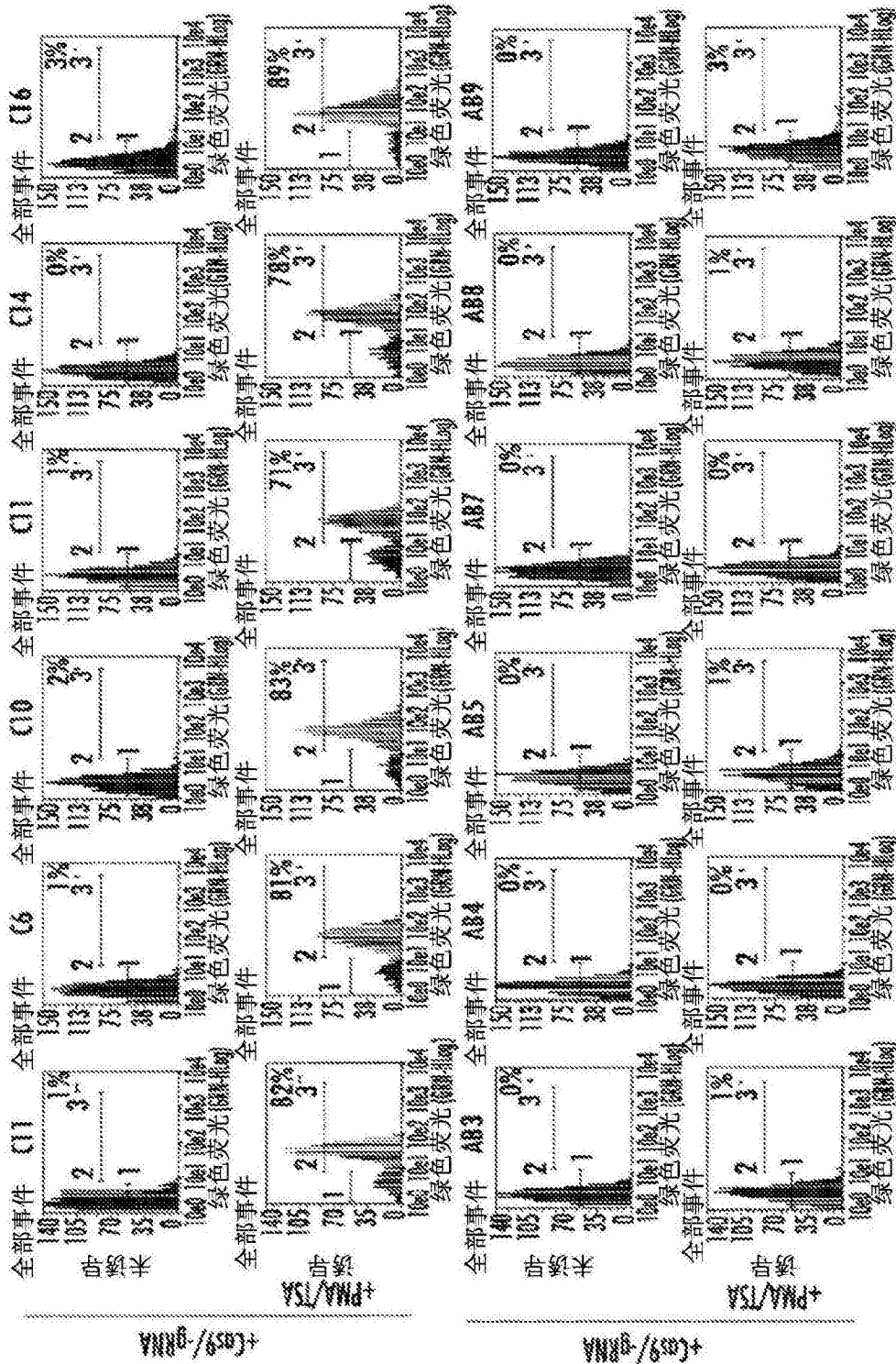


图7A

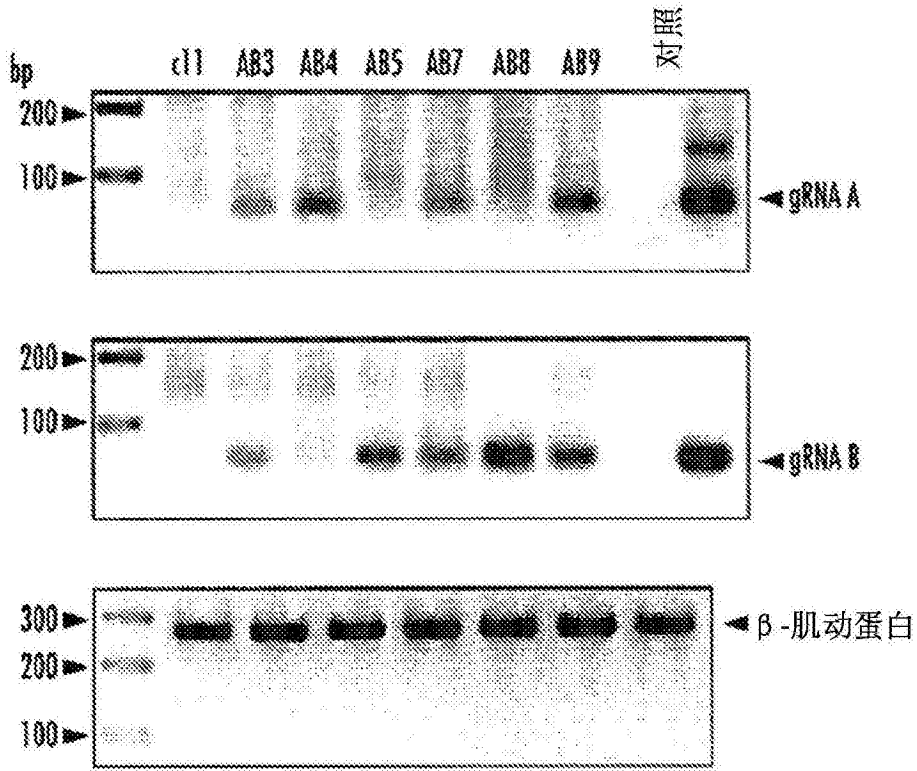


图7B

染色体	起始	终结	基因名	功能	区域	类型
I	114333320	114333420	K58M	固有的		CG
ChrXus	I	I	~	基因间的		CG
HIV						
插入I	ChrXus	9789	9709	~	基因间的	CG
I	114338815	114338815	K58M	内含子的		CG
HIV						
16	1991378	1991378	MSR31	外显子的		CG
ChrXus	I	I	~	基因间的		CG
插入II						
ChrXus	9789	9709	~	基因间的		CG
16	1991382	1991382	~	外显子的		CG

图8A

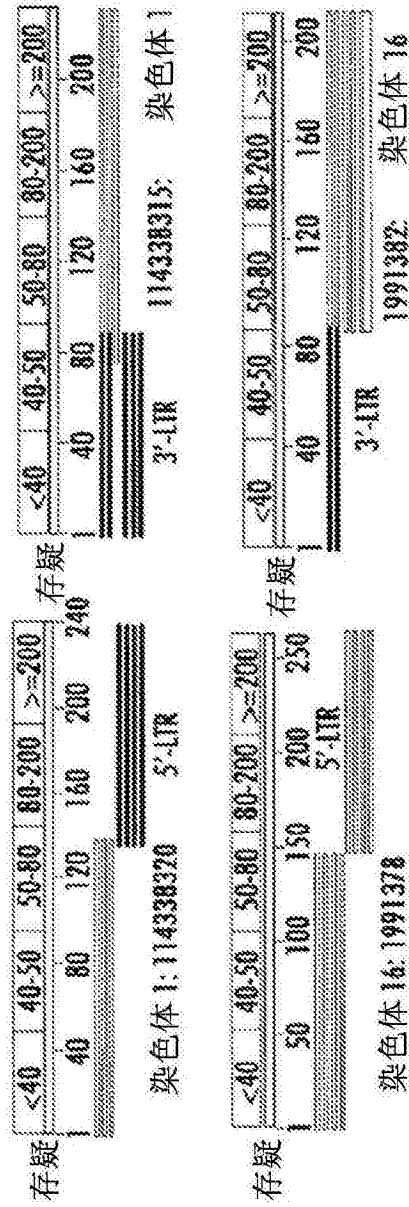


图8B

CR1 靶标/CR43118 序列比对

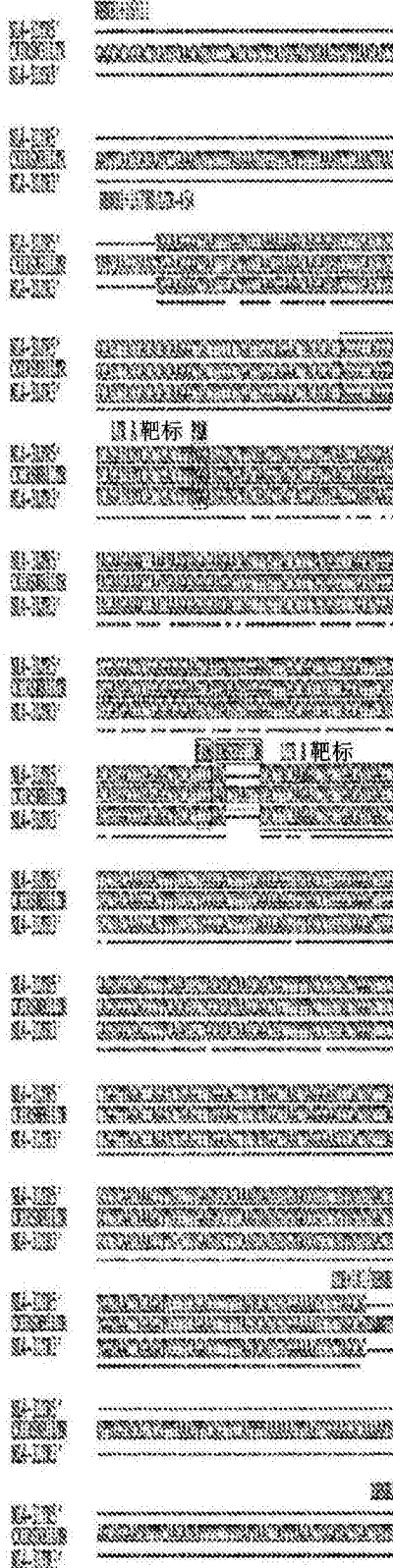


图 9A

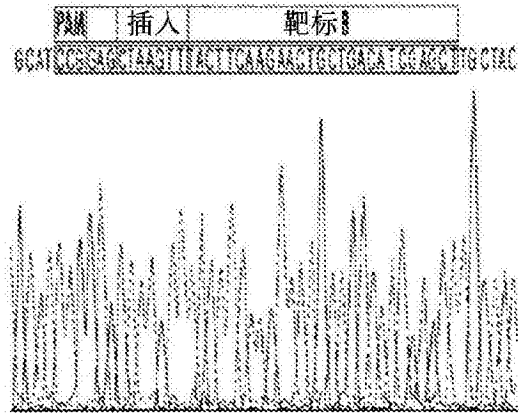


图 9B

CH16 MSRBI LTR/NL4-3 LTRs 序列比对

```

MSRBI+1648
NL4-3LTR5'
CH16MSRBI LTR
NL4-3LTR3'
-----
MSRBI+1926/LTR -454
NL4-3LTR5'
CH16MSRBI LTR
NL4-3LTR3'
-----
ATCTGTGGATCTAGCCACACACAGGCTACTTCCCTGATTTGGUNGA ACTACACACCAGGGC
ATCTGTGGATCTAGCCACACACAGGCTACTTCCCTGATTTGGUNGA ACTACACACCAGGGC
ATCTGTGGATCTAGCCACACACAGGCTACTTCCCTGATTTGGUNGA ACTACACACCAGGGC
NL4-3LTR5'
CH16MSRBI LTR
NL4-3LTR3'
-----
          LTR A 靶标          插入          PAN
CAGGGATCAGATATCCACTGAC-----CTTTGGATGGTCTTCAAGTTAGTACAGT
CAGGGATCAGATATCCACTGAC-----CTTTGGATGGTCTTCAAGTTAGTACAGT
CAGGGATCAGATATCCACTGAC-----CTTTGGATGGTCTTCAAGTTAGTACAGT
NL4-3LTR5'
CH16MSRBI LTR
NL4-3LTR3'
-----
TGAACCAGAGCAAGTACTTCTGCCAANTANGGAGAGAACACACAGCTTGTACACCCCTAT
TGAGCCAGAGCAAGTACTTCTGCCAANTANGGAGAGAACACACAGCTTGTACACCCCTAT
TGAGCCAGAGCAAGTACTTCTGCCAANTANGGAGAGAACACACAGCTTGTACACCCCTAT
NL4-3LTR5'
CH16MSRBI LTR
NL4-3LTR3'
-----
GAGCCAGCATGGGATGGAGACCCGGAGGGAGAGTATTAGTGTGGAGTTTGACAGCCT
GAGCCAGCATGGGATGGAGACCCGGAGGGAGAGTATTAGTGTGGAGTTTGACAGCCT
GAGCCAGCATGGGATGGAGACCCGGAGGGAGAGTATTAGTGTGGAGTTTGACAGCCT
NL4-3LTR5'
CH16MSRBI LTR
NL4-3LTR3'
-----
          PAN          插入          LTR B 靶标
CCTAGCATTTTCATCAGTGGCCCGAGAGCTGCATCCGAGT-----ACTTCAAGACTGCTG
CCTAGCATTTTCATCAGTGGCCCGAGAGCTGCATCCGAGT-----ACTTCAAGACTGCTG
CCTAGCATTTTCATCAGTGGCCCGAGAGCTGCATCCGAGT-----ACTTCAAGACTGCTG
NL4-3LTR5'
CH16MSRBI LTR
NL4-3LTR3'
-----
ACATCGAGCTTTCTACAAGGACTTTCCGCTGGGACTTTCCAGGGAGGTTGTGGCTGGG
ACATCGAGCTTTCTACAAGGACTTTCCGCTGGGACTTTCCAGGGAGGTTGTGGCTGGG
ACATCGAGCTTTCTACAAGGACTTTCCGCTGGGACTTTCCAGGGAGGTTGTGGCTGGG
NL4-3LTR5'
CH16MSRBI LTR
NL4-3LTR3'
-----
CGGGACTGGGGAGTGGCGAGCCCTCAGATGCTACTATATTTGGAGCTGCTTTTGGCTGTA
CGGGACTGGGGAGTGGCGAGCCCTCAGATGCTACTATATTTGGAGCTGCTTTTGGCTGTA
CGGGACTGGGGAGTGGCGAGCCCTCAGATGCTACTATATTTGGAGCTGCTTTTGGCTGTA
NL4-3LTR5'
CH16MSRBI LTR
NL4-3LTR3'
-----
CTGGGTCTCTCTGGTTAGACACAGATCTGAGCCTGGGAGCTCTCTGGCTAACTAGGGAACC
CTGGGTCTCTCTGGTTAGACACAGATCTGAGCCTGGGAGCTCTCTGGCTAACTAGGGAACC
CTGGGTCTCTCTGGTTAGACACAGATCTGAGCCTGGGAGCTCTCTGGCTAACTAGGGAACC
NL4-3LTR5'
CH16MSRBI LTR
NL4-3LTR3'
-----
CACTGCTTAAGCCTCAATAAAGCTTGCCTTGAGTCTCAAAGTAGTGTGGCCGCTCTGT
CACTGCTTAAGCCTCAATAAAGCTTGCCTTGAGTCTCAAAGTAGTGTGGCCGCTCTGT
CACTGCTTAAGCCTCAATAAAGCTTGCCTTGAGTCTCAAAGTAGTGTGGCCGCTCTGT
NL4-3LTR5'
CH16MSRBI LTR
NL4-3LTR3'
-----
TGTGTGACTCTGGTAACTAGAGATCCCTCAGACCCTTTTACTCAGTGTGGAAAATCTCTA
TGTGTGACTCTGGTAACTAGAGATCCCTCAGACCCTTTTACTCAGTGTGGAAAATCTCTA
TGTGTGACTCTGGTAACTAGAGATCCCTCAGACCCTTTTACTCAGTGTGGAAAATCTCTA
NL4-3LTR5'
CH16MSRBI LTR
NL4-3LTR3'
-----
LTR +181/MSRBI +1927 { 复制 } MSRBI+1958
GCAC-----
GCACCTATGAGCTGTTCTCCAGCCGCTCGAAGTATCCACA
GCAC-----
***

```

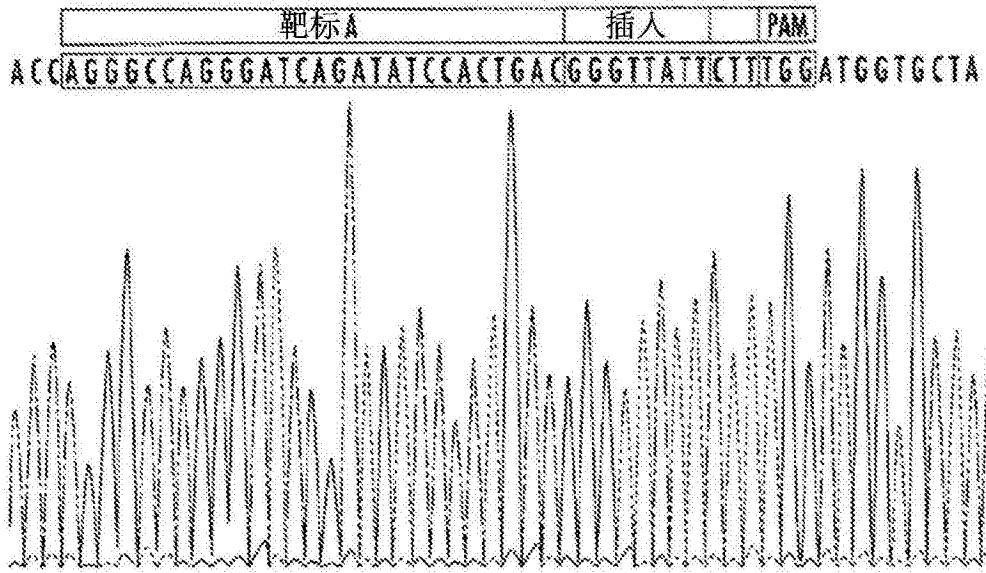


图10A

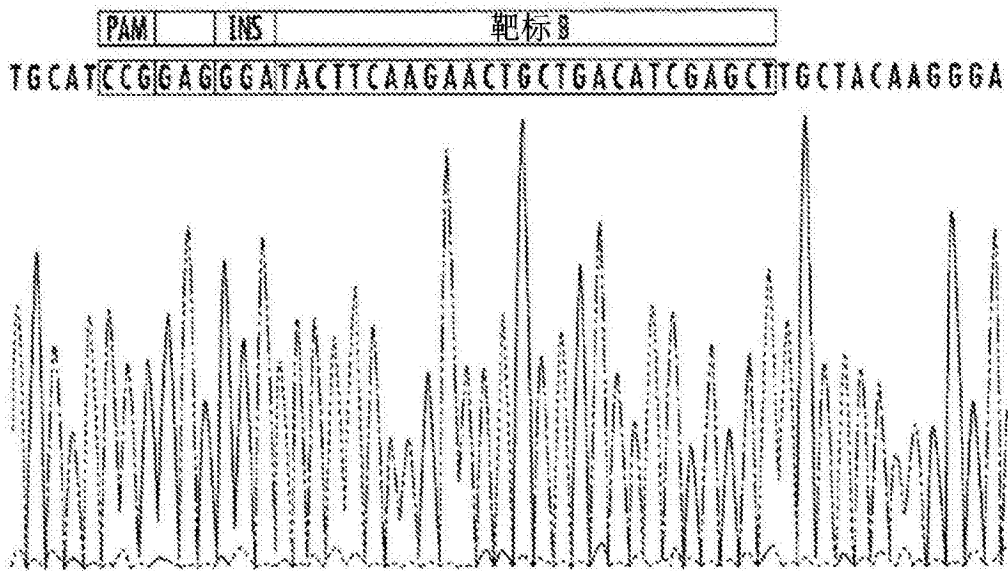


图10B

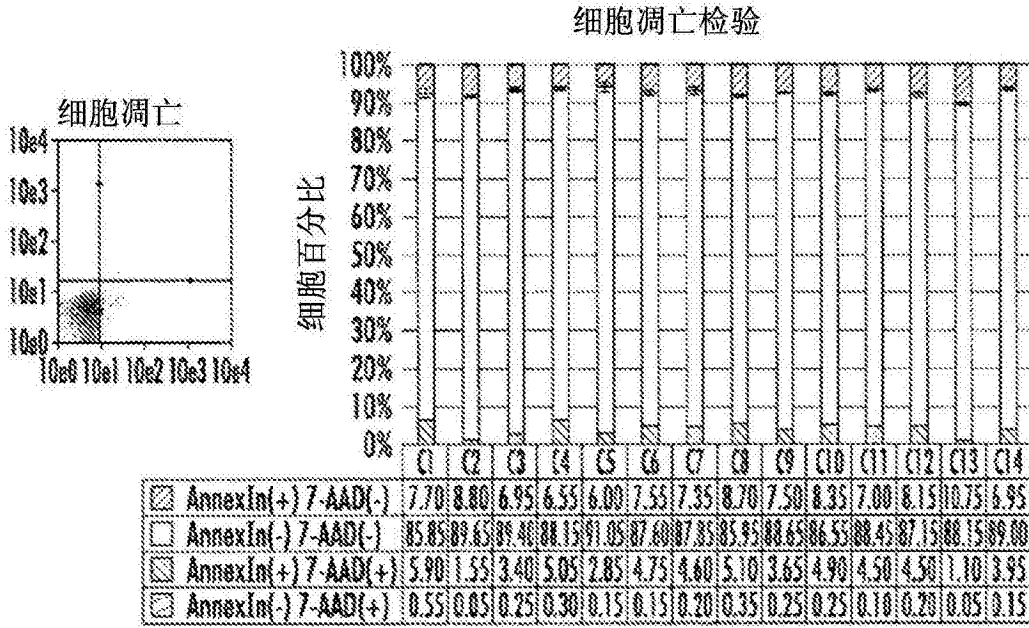


图11A

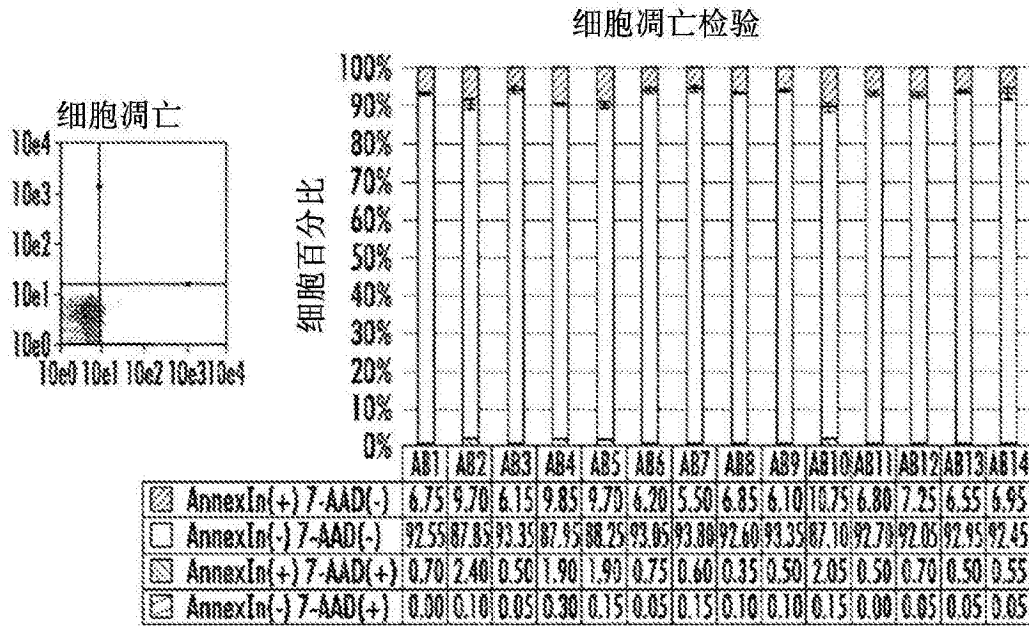


图11B

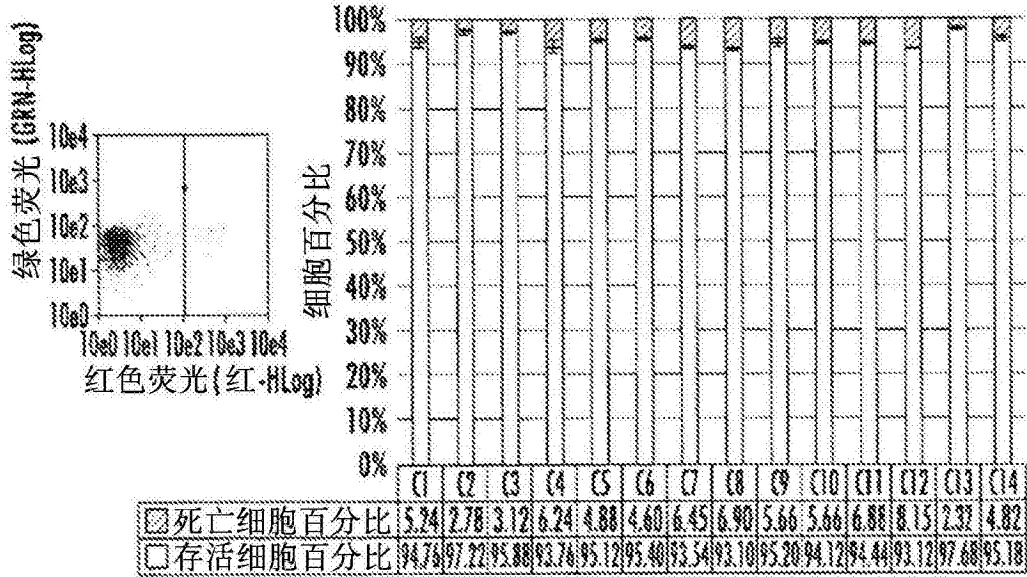


图12A

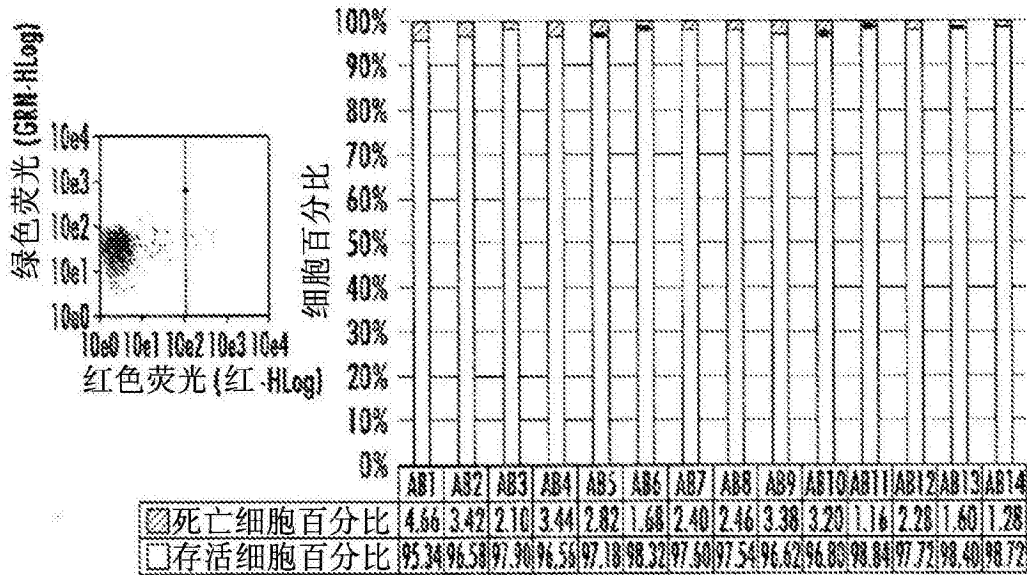


图12B

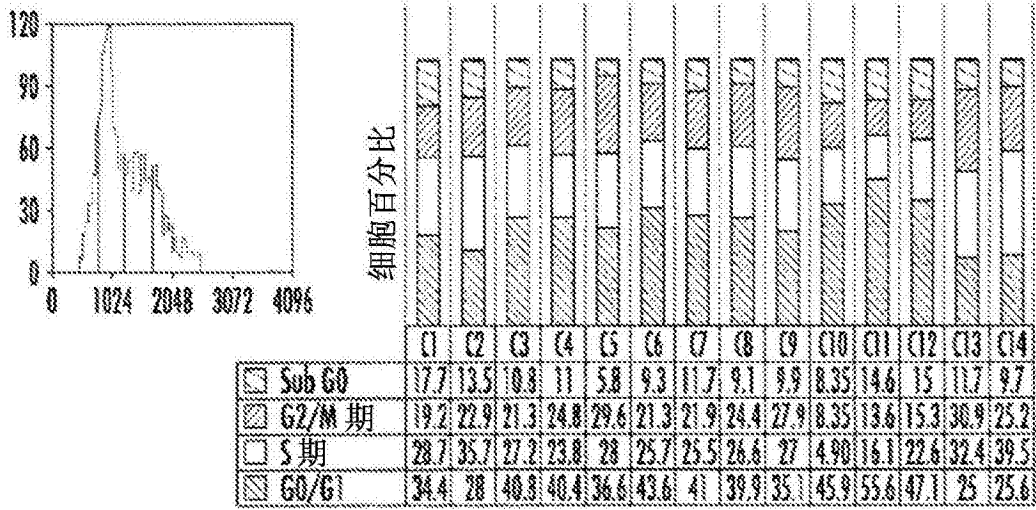


图13A

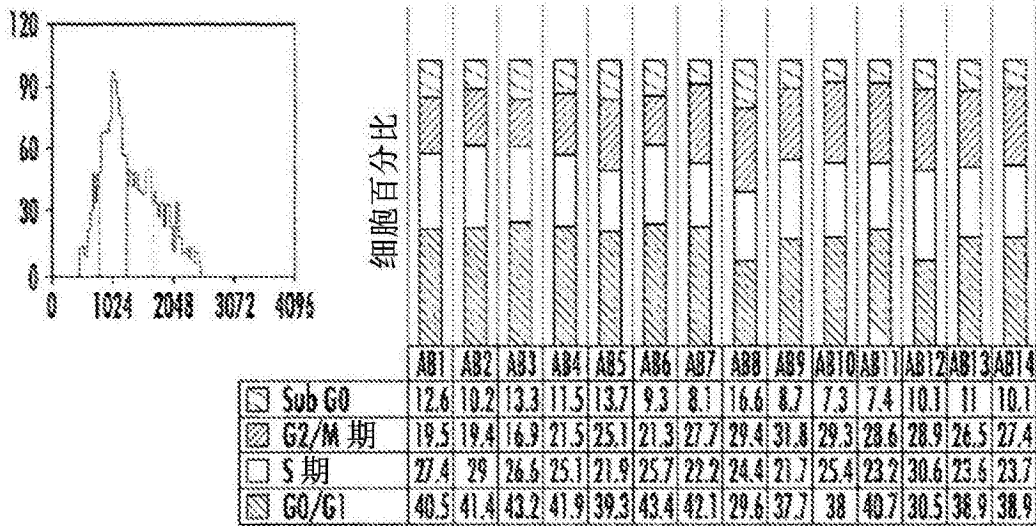


图13B

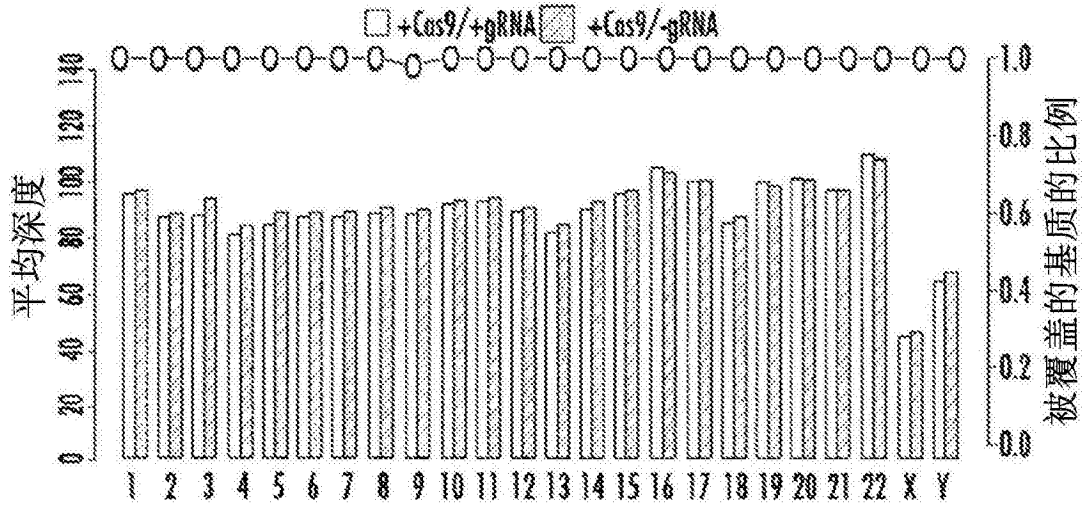


图14

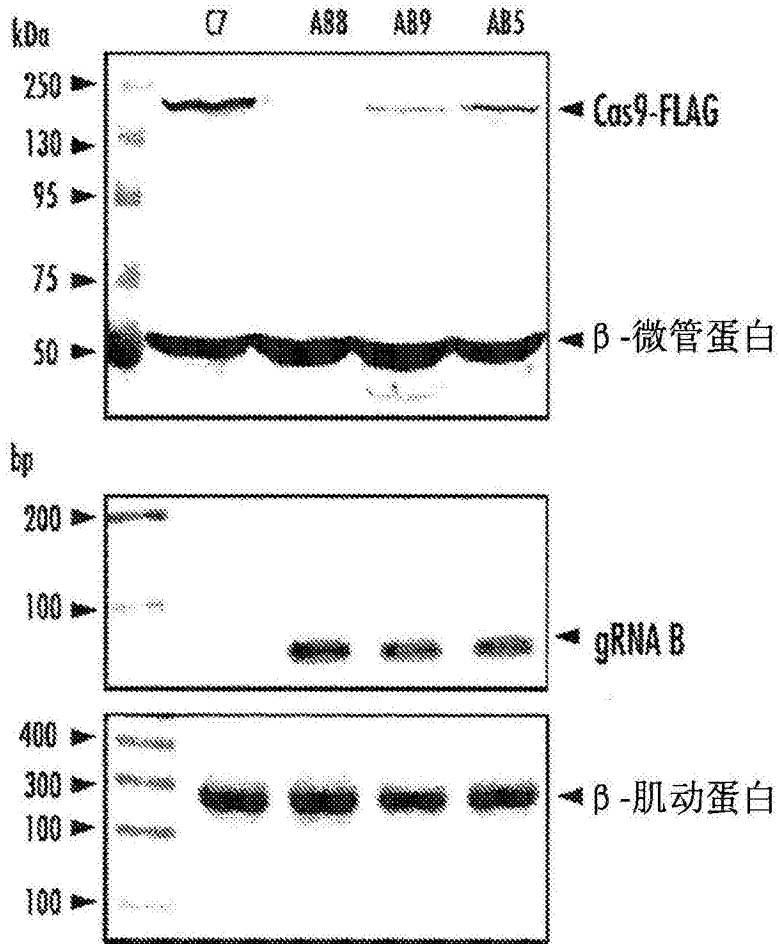
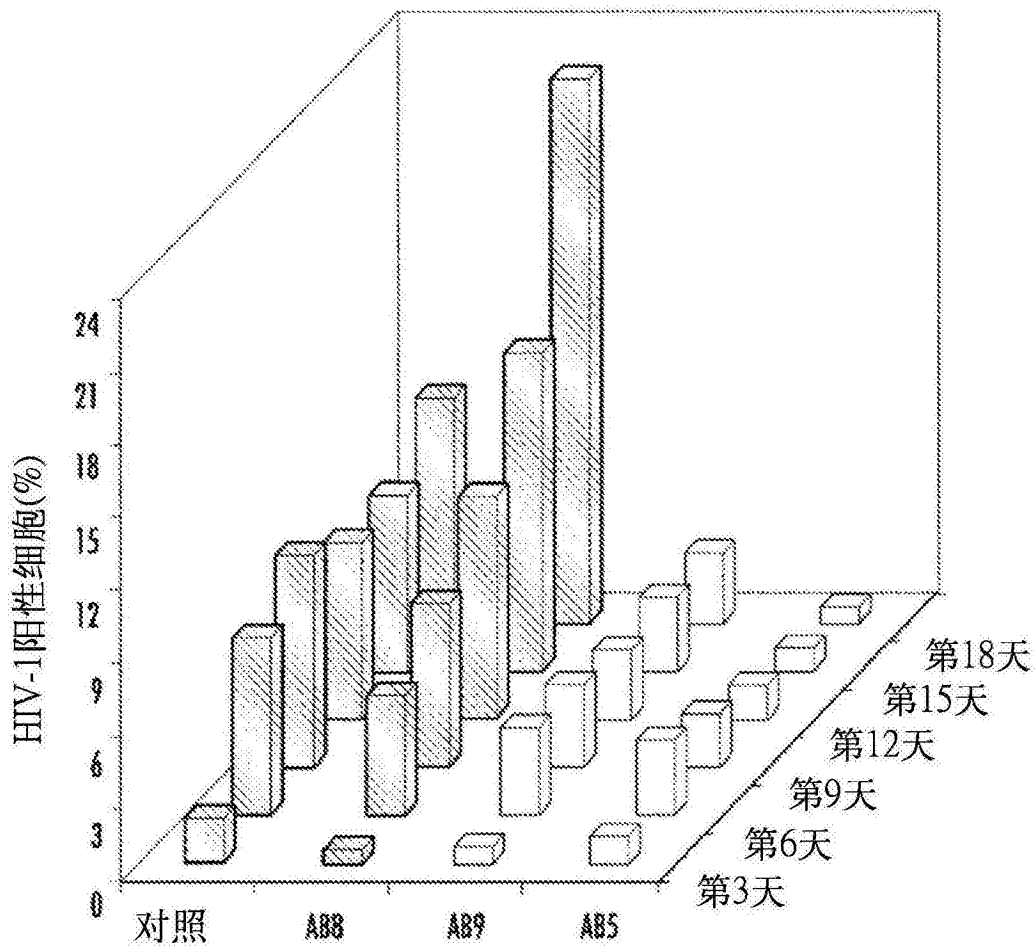


图15A



% HIV-1阳性细胞

天数	对照	ABB	AB9	AB5
3	1.9 (± 0.71)	0.6 (± 0.21)	0.8 (± 0.07)	1.2 (± 0.42)
6	7.4 (± 1.41)	5.1 (± 0.64)	3.7 (± 0.92)	3.2 (± 0.21)
9	8.7 (± 0.28)	6.9 (± 0.42)	3.5 (± 0.64)	2.3 (± 0.14)
12	7.3 (± 0.14)	9.2 (± 0.57)	2.9 (± 0.99)	1.4 (± 1.13)
15	7.2 (± 0.11)	13.2 (± 0.30)	3.2 (± 0.14)	1.0 (± 0.20)
18	9.4 (± 0.21)	22.6 (± 1.70)	2.9 (± 0.07)	0.8 (± 0.07)

图15C

	年龄/性别	种族	VL	CD4	CD4%	ART
案例1 (TUR0001)	43/M	AA	12.45	716	11	阿扎那韦,利托那韦,雷特格韦
案例2 (TUR0001)	50/M	AA	<20.0	630	30	利托那韦,达鲁拉韦,玉米细条病毒, 多替拉韦
案例3 (TUR0001)	44/F	HIS	<20.0	714	39	替诺福韦/恩曲他滨,阿扎那韦, 利托那韦
案例4 (TUR0001)	36/M	AA	523	16	8	拉米夫定,利托那韦,雷特格韦, 达鲁拉韦

图16A

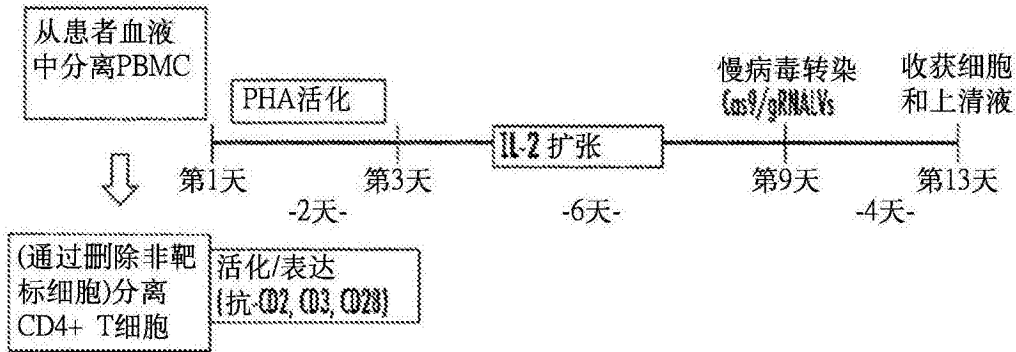


图16B

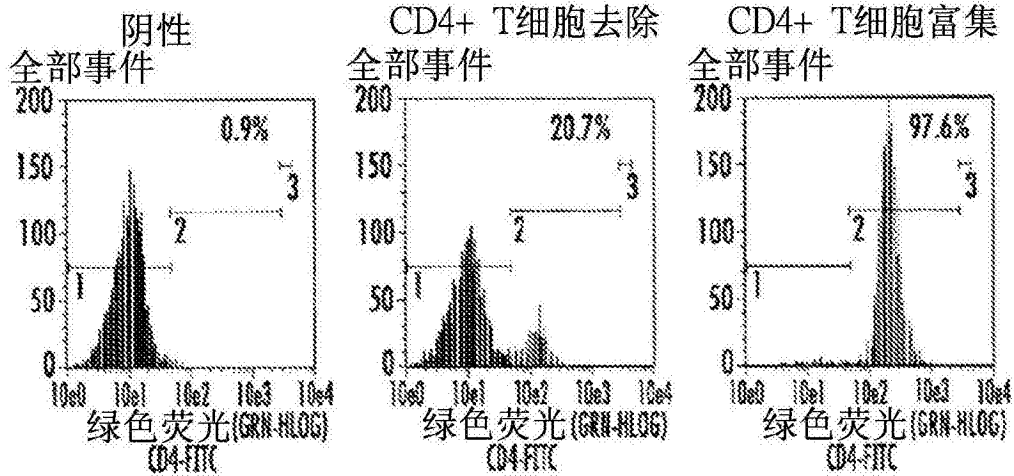


图16C

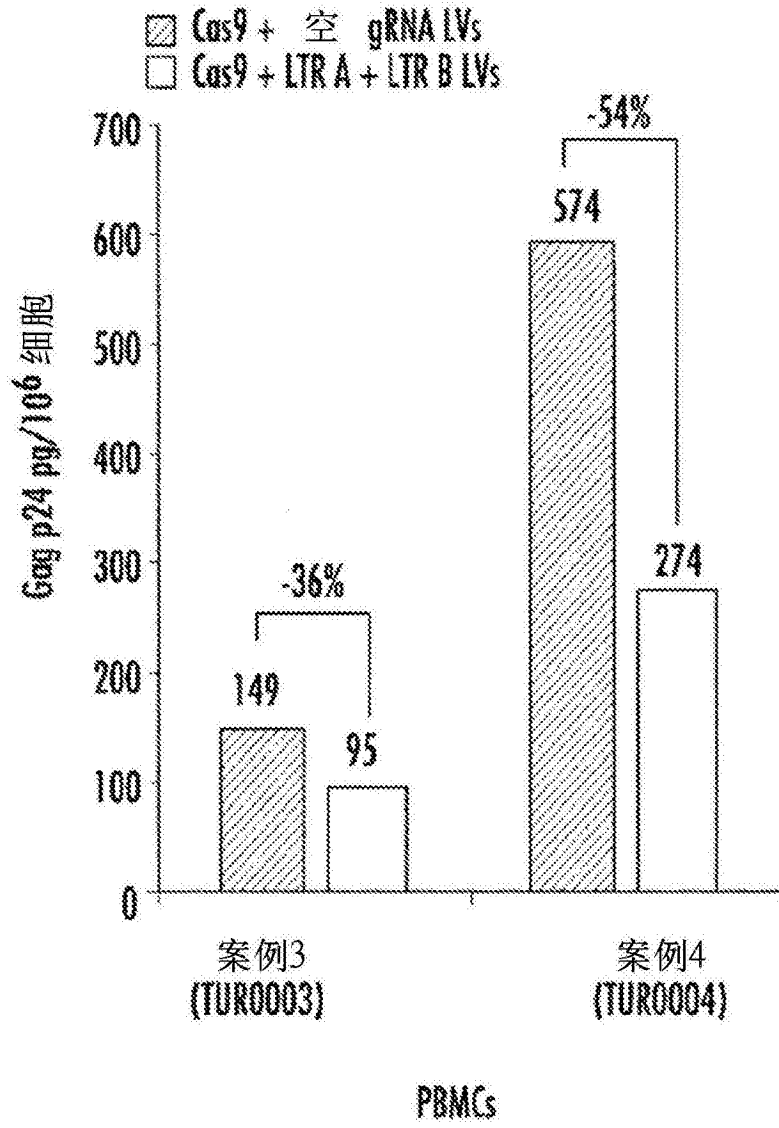


图17

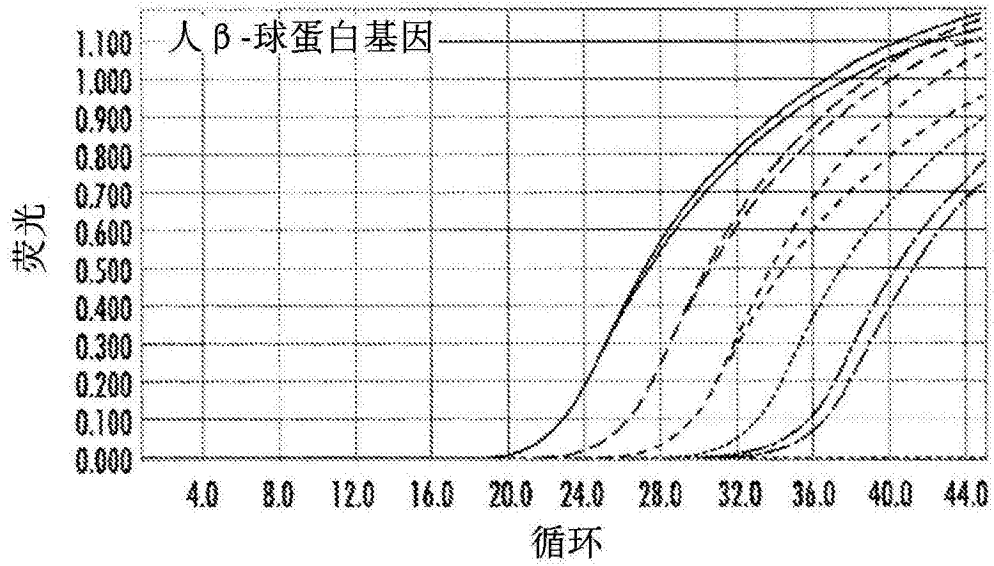


图18A

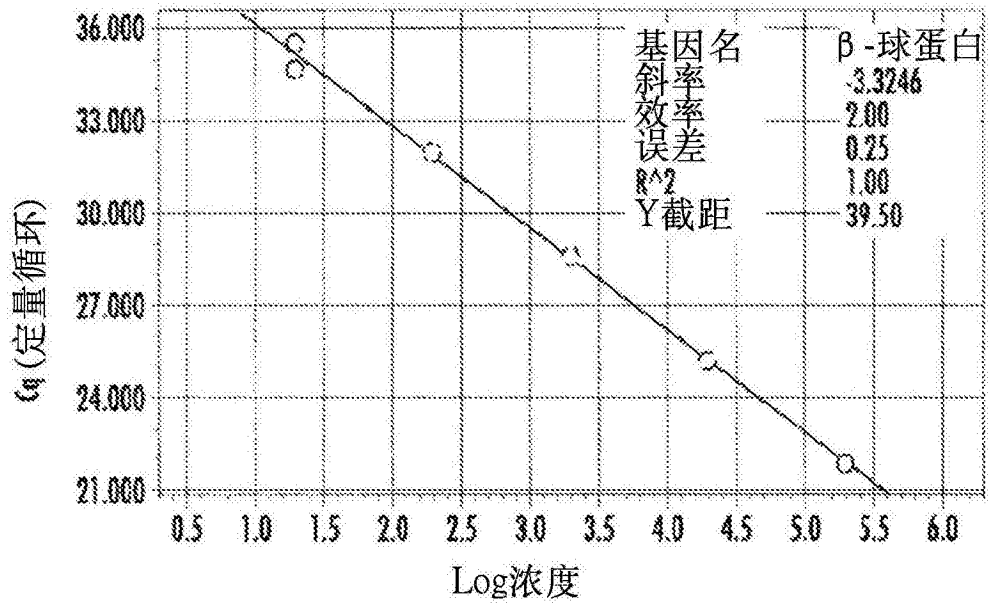


图18B

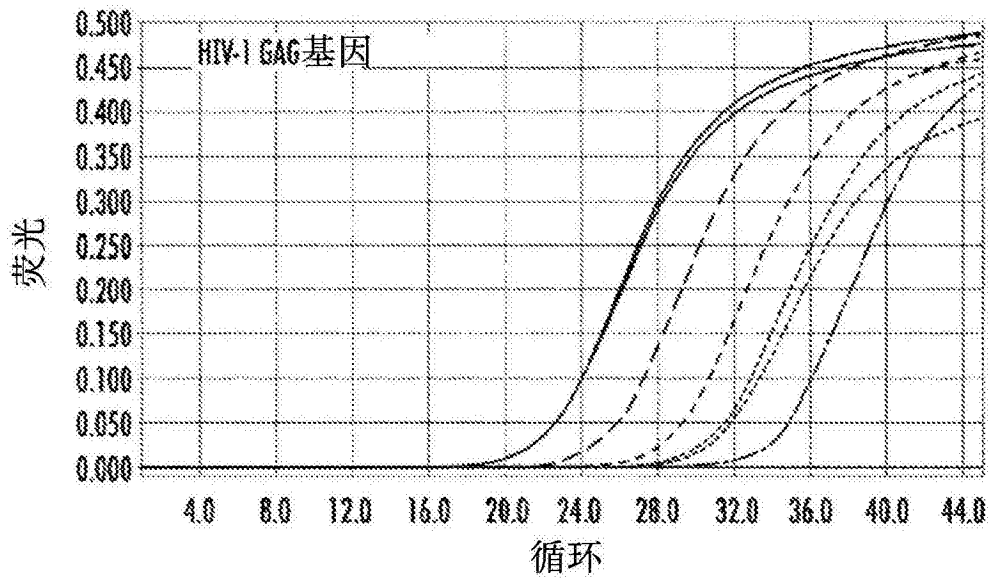


图18C

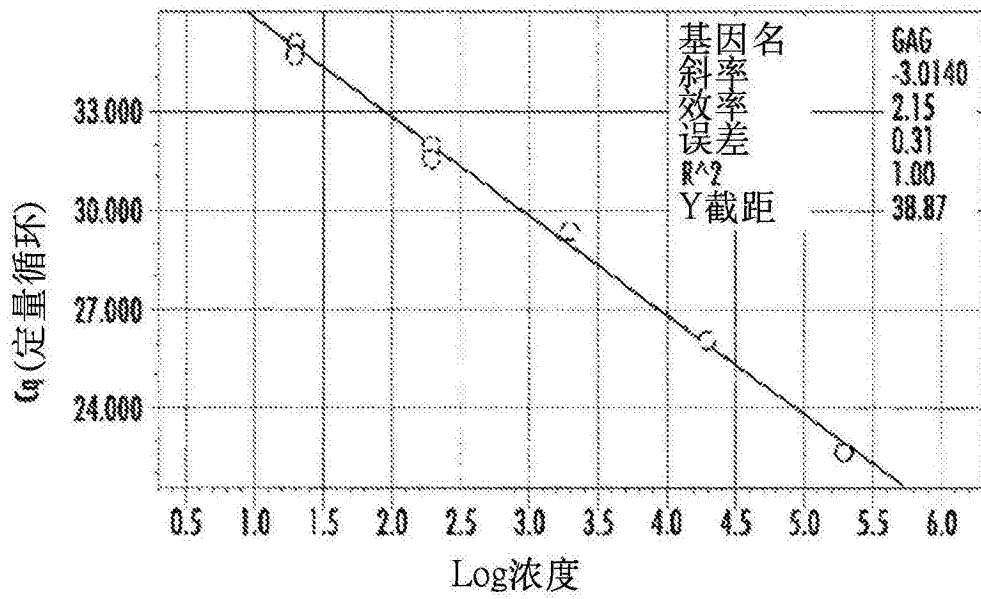


图18D