



(12)发明专利申请

(10)申请公布号 CN 110446781 A

(43)申请公布日 2019. 11. 12

(21)申请号 201880020639.2

(22)申请日 2018.02.15

(30)优先权数据

62/459,203 2017.02.15 US

(85)PCT国际申请进入国家阶段日

2019.09.24

(86)PCT国际申请的申请数据

PCT/US2018/018370 2018.02.15

(87)PCT国际申请的公布数据

WO2018/152325 EN 2018.08.23

(71)申请人 蓝鸟生物公司

地址 美国马萨诸塞州剑桥宾尼街60号

(72)发明人 李百胜 亚历山大·阿斯特拉罕

(74)专利代理机构 北京英赛嘉华知识产权代理有限公司 11204

代理人 王达佐 洪欣

(51)Int.Cl.

C12N 5/078(2006.01)

C12N 5/10(2006.01)

C12N 9/22(2006.01)

C12N 15/09(2006.01)

C12N 15/66(2006.01)

C12N 15/90(2006.01)

A61K 31/713(2006.01)

A61K 35/761(2006.01)

A61K 38/46(2006.01)

权利要求书7页 说明书78页

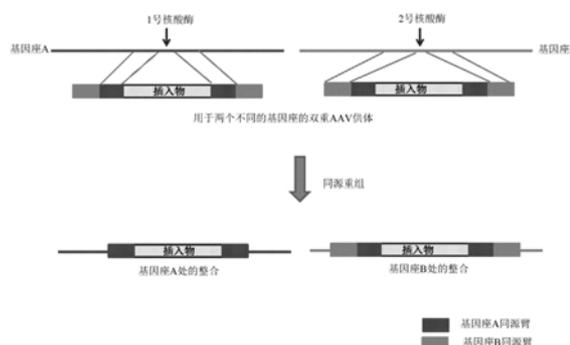
序列表17页 附图9页

(54)发明名称

供体修复模板多重基因组编辑

(57)摘要

本公开提供改良的多重基因组编辑组合物和方法。本公开还提供用于预防、治疗或改善血红蛋白病、癌症、感染性疾病、自身免疫疾病、炎性疾病或免疫缺陷的至少一种症状的基因组编辑细胞。



1. 一种DNA供体修复模板,其包含:
 - (a) 针对第一目标位点的第一对同源臂;
 - (b) 针对第二目标位点的第二对同源臂;和
 - (c) 一个或多个转基因。
2. 根据权利要求1所述的DNA供体修复模板,其中所述第一对同源臂包含与所述第一目标位点的DNA序列5'同源的5'同源臂;和与所述第一目标位点的DNA序列3'同源的3'同源臂。
3. 根据权利要求1或权利要求2所述的DNA供体修复模板,其中所述第二对同源臂包含与所述第二目标位点的DNA序列5'同源的5'同源臂;和与所述第二目标位点的DNA序列3'同源的3'同源臂。
4. 根据权利要求1至3中任一项所述的DNA供体修复模板,其中针对所述第一目标位点和/或所述第二目标位点的所述5'和3'同源臂的长度是独立地选自约100bp到约2500bp。
5. 根据权利要求1至4中任一项所述的DNA供体修复模板,其中针对所述第一目标位点和/或所述第二目标位点的所述5'和3'同源臂的长度是独立地选自约100bp到约1500bp。
6. 根据权利要求1至5中任一项所述的DNA供体修复模板,其中针对所述第一目标位点和/或所述第二目标位点的所述5'和3'同源臂的长度是独立地选自约100bp到约1000bp。
7. 根据权利要求1至6中任一项所述的DNA供体修复模板,其中针对所述第一目标位点和/或所述第二目标位点的所述5'和3'同源臂的长度是独立地选自约100bp到约500bp。
8. 根据权利要求1至7中任一项所述的DNA供体修复模板,其中针对所述第一目标位点和/或所述第二目标位点的所述5'和3'同源臂的长度是约500bp。
9. 根据权利要求1至8中任一项所述的DNA供体修复模板,其中针对所述第一目标位点和/或所述第二目标位点的所述5'和3'同源臂的长度是约300bp。
10. 根据权利要求1至9中任一项所述的DNA供体修复模板,其中针对所述第一目标位点和/或所述第二目标位点的所述5'和3'同源臂的长度是约100bp。
11. 根据权利要求1至10中任一项所述的DNA供体修复模板,其中所述第一目标位点的5'同源臂位于所述第二目标位点的5'同源臂的5'。
12. 根据权利要求1至11中任一项所述的DNA供体修复模板,其中所述第一目标位点的3'同源臂位于所述第二目标位点的3'同源臂的5'。
13. 根据权利要求1至11中任一项所述的DNA供体修复模板,其中所述第一目标位点的3'同源臂位于所述第二目标位点的3'同源臂的3'。
14. 根据权利要求1至11中任一项所述的DNA供体修复模板,其中所述第一目标位点的5'同源臂和3'同源臂侧接第一转基因并且所述第二目标位点的5'同源臂和3'同源臂侧接第二转基因。
15. 根据权利要求1至14中任一项所述的DNA供体修复模板,其中所述第一目标位点和所述第二目标位点在不同基因中。
16. 根据权利要求1至15中任一项所述的DNA供体修复模板,其中所述第一目标位点或所述第二目标位点包含TCR α 基因中的被工程改造的核酸酶裂解位点。
17. 根据权利要求1至15中任一项所述的DNA供体修复模板,其中所述第一目标位点或所述第二目标位点包含免疫系统检查点基因或编码免疫抑制性信号传导组件的基因中的

被工程改造的核酸酶裂解位点。

18. 根据权利要求1至15中任一项所述的DNA供体修复模板,其中所述第一目标位点是TCR α 基因中的被工程改造的核酸酶裂解位点并且所述第二目标位点是免疫系统检查点基因或编码免疫抑制性信号传导组件的基因中的被工程改造的核酸酶裂解位点。

19. 根据权利要求1至15中任一项所述的DNA供体修复模板,其中所述第一目标位点是第一免疫系统检查点基因或编码免疫抑制性信号传导组件的基因中的被工程改造的核酸酶裂解位点,并且所述第二目标位点是第二免疫系统检查点基因或编码免疫抑制性信号传导组件的基因中的被工程改造的核酸酶裂解位点。

20. 根据权利要求17至19中任一项所述的DNA供体修复模板,其中所述免疫系统检查点基因是独立地选自由以下组成的群组:程序性细胞死亡蛋白质1 (PD-1)、淋巴细胞活化基因3蛋白质 (LAG-3)、T细胞免疫球蛋白结构域和粘蛋白结构域蛋白质3 (TIM-3)、细胞毒性T淋巴细胞抗原-4 (CTLA-4)、带状T淋巴细胞弱化子 (BTLA)、T细胞免疫球蛋白和基于免疫受体酪氨酸的抑制性基元结构域 (TIGIT)、T细胞活化的V-结构域Ig抑制因子 (VISTA) 和杀手机体免疫球蛋白样受体 (KIR) 基因。

21. 根据权利要求17至19中任一项所述的DNA供体修复模板,其中所述编码免疫抑制性信号传导组件的基因是独立地选自由以下组成的群组:白细胞介素受体10 α 、转化生长因子 β 受体I (TGF β RI)、转化生长因子 β 受体II (TGF β RII)、芳基炔受体 (AHR)、血清和糖皮质激素调节的激酶1 (SGK1)、结节性硬化症复合物2 (TSC2)、冯希派尔-林道肿瘤抑制因子 (von Hippel-Lindau tumor suppressor; VHL)、腺苷A2a受体 (A2AR) 和Cbl原癌基因B (CBLB)。

22. 根据权利要求16至21中任一项所述的DNA供体修复模板,其中所述被工程改造的核酸酶是选自由以下组成的群组:兆核酸酶、megaTAL、TALEN、ZFN和CRISPR/Cas核酸酶。

23. 根据权利要求16至22中任一项所述的DNA供体修复模板,其中所述被工程改造的核酸酶是由选自由以下组成的群组的LAGLIDADG导向核酸内切酶 (LHE) 工程改造的兆核酸酶: I-AabMI、I-AaeMI、I-AniI、I-ApaMI、I-CapIII、I-CapIV、I-CkaMI、I-CpaMI、I-CpaMII、I-CpaMIII、I-CpaMIV、I-CpaMV、I-CpaV、I-CraMI、I-EjeMI、I-GpeMI、I-GpiI、I-GzeMI、I-GzeMII、I-GzeMIII、I-HjeMI、I-LtrII、I-LtrI、I-LtrWI、I-MpeMI、I-MveMI、I-NcrII、I-NcrI、I-NcrMI、I-OheMI、I-OnuI、I-OsoMI、I-OsoMII、I-OsoMIII、I-OsoMIV、I-PanMI、I-PanMII、I-PanMIII、I-PnoMI、I-ScuMI、I-SmaMI、I-SscMI和I-Vdi141I。

24. 根据权利要求16至23中任一项所述的DNA供体修复模板,其中所述被工程改造的核酸酶是由选自由以下组成的群组的LHE工程改造的兆核酸酶: I-CpaMI、I-HjeMI、I-OnuI、I-PanMI和SmaMI。

25. 根据权利要求16至24中任一项所述的DNA供体修复模板,其中所述被工程改造的核酸酶是I-OnuI LHE。

26. 根据权利要求16至24中任一项所述的DNA供体修复模板,其中所述被工程改造的核酸酶是I-SmaMI LHE。

27. 根据权利要求16至22中任一项所述的DNA供体修复模板,其中所述被工程改造的核酸酶是包含TALE DNA结合结构域和根据权利要求22至26中任一项所述的被工程改造的兆核酸酶的megaTAL。

28. 根据权利要求27所述的DNA供体修复模板,其中所述TALE结合结构域包含约9.5个

TALE重复单元到约15.5个TALE重复单元。

29. 根据权利要求27所述的DNA供体修复模板,其中所述TALE结合结构域包含约9.5个TALE重复单元、约10.5个TALE重复单元、约11.5个TALE重复单元、约12.5个TALE重复单元、约13.5个TALE重复单元、约14.5个TALE重复单元或约15.5个TALE重复单元。

30. 根据权利要求1至29中任一项所述的DNA供体修复模板,其中所述一个或多个转基因包含聚核苷酸,所述聚核苷酸编码免疫效能增强子、免疫抑制信号减弱子或被工程改造的抗原受体。

31. 根据权利要求1至30中任一项所述的DNA供体修复模板,其中所述一个或多个转基因进一步包含RNA聚合酶II启动子,所述RNA聚合酶II启动子可操作地连接到所述编码免疫效能增强子、免疫抑制信号减弱子或被工程改造的抗原受体的聚核苷酸。

32. 根据权利要求30或权利要求31所述的DNA供体修复模板,其中所述一个或多个转基因编码一个或多个CAR。

33. 根据权利要求32所述的DNA供体修复模板,其中所述一个或多个CAR是选自由以下组成的群组:抗BCMACAR和抗CD19 CAR。

34. 根据权利要求31所述的DNA供体修复模板,其中所述RNA聚合酶II启动子是选自由以下组成的群组:短EF1 α 启动子、长EF1 α 启动子、人类ROSA 26基因座、泛素C (UBC) 启动子、磷酸甘油酸激酶-1 (PGK) 启动子、巨细胞病毒增强子/鸡 β -肌动蛋白 (CAG) 启动子、 β -肌动蛋白启动子和骨髓增生肉瘤病毒增强子、负调控区缺失、d1587rev引物结合位点取代 (MND) 的启动子。

35. 根据权利要求30所述的DNA供体修复模板,其中所述聚核苷酸进一步编码一个或多个自裂解病毒肽,所述一个或多个自裂解病毒肽可操作地连接到一个或多个免疫效能增强子、免疫抑制信号减弱子或被工程改造的抗原受体;穿插在一个或多个免疫效能增强子、免疫抑制信号减弱子或被工程改造的抗原受体之间;和/或侧接一个或多个免疫效能增强子、免疫抑制信号减弱子或被工程改造的抗原受体。

36. 根据权利要求35所述的DNA供体修复模板,其中所述自裂解病毒肽是2A肽。

37. 根据权利要求30至36中任一项所述的DNA供体修复模板,其中所述聚核苷酸进一步包含异源聚腺苷酸化信号。

38. 根据权利要求30至37中任一项所述的DNA供体修复模板,其中所述一个或多个免疫抑制信号减弱子包含对抗免疫抑制因子的酶功能;结合免疫抑制因子的胞外结构域,任选地其中所述胞外结构域是抗体或其抗原结合片段;结合免疫抑制因子的胞外结构域和跨膜结构域;或结合免疫抑制因子的胞外结构域、跨膜结构域和不能转换免疫抑制信号的被修饰的胞内结构域。

39. 根据权利要求30至37中任一项所述的DNA供体修复模板,其中所述一个或多个免疫效能增强子是独立地选自由以下组成的群组:双特异性T细胞接合分子 (BiTE)、免疫增强因子和翻转受体。

40. 根据权利要求30至37中任一项所述的DNA供体修复模板,其中所述一个或多个免疫效能增强子是独立地选自由以下组成的群组:细胞因子、趋化因子、细胞毒素、细胞因子受体和其变异体。

41. 根据权利要求30至35中任一项所述的DNA供体修复模板,其中所述一个或多个被工

程改造的抗原受体是独立地选自由以下组成的群组：被工程改造的TCR、CAR、DARIC或 ζ 因子。

42. 根据权利要求41所述的DNA供体修复模板，其中所述一个或多个被工程改造的抗原受体是CAR。

43. 根据权利要求42所述的DNA供体修复模板，其中所述一个或多个CAR是选自由以下组成的群组：抗BCMACAR和抗CD19 CAR。

44. 一种病毒载体，其包含根据权利要求1至43中任一项所述的DNA供体修复模板。

45. 根据权利要求44所述的病毒载体，其中所述病毒载体是重组型腺相关病毒载体(rAAV)或逆转录病毒。

46. 根据权利要求45所述的病毒载体，其中所述rAAV具有一个或多个来自AAV2的ITR。

47. 根据权利要求45或权利要求46所述的病毒载体，其中所述rAAV具有选自由以下组成的群组的血清型：AAV1、AAV2、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAV9和AAV10。

48. 根据权利要求45所述的病毒载体，其中所述rAAV具有AAV6血清型。

49. 根据权利要求45所述的病毒载体，其中所述逆转录病毒是慢病毒。

50. 根据权利要求49所述的病毒载体，其中所述慢病毒是整合酶缺陷型慢病毒。

51. 一种细胞，其包含根据权利要求1至43中任一项所述的DNA供体修复模板或根据权利要求44至50中任一项所述的病毒载体。

52. 根据权利要求51所述的细胞，其中已经通过同源定向修复在所述第一目标位点和所述第二目标位点处插入所述一个或多个转基因。

53. 根据权利要求51或权利要求52所述的细胞，其中所述细胞是造血细胞。

54. 根据权利要求51至53中任一项所述的细胞，其中所述细胞是免疫效应细胞。

55. 根据权利要求51至54中任一项所述的细胞，其中所述细胞是CD3⁺、CD4⁺、CD8⁺或其组合。

56. 根据权利要求51至55中任一项所述的细胞，其中所述细胞是T细胞。

57. 根据权利要求51至56中任一项所述的细胞，其中所述细胞是细胞毒性T淋巴细胞(CTL)、肿瘤浸润淋巴细胞(TIL)或辅助T细胞。

58. 根据权利要求51至57中任一项所述的细胞，其中所述细胞的来源是外周血液单核细胞、骨髓、淋巴结组织、脐带血、胸腺组织、来自感染位点的组织、腹水、胸腔积液、脾组织或肿瘤。

59. 一种组合物，其包含根据权利要求1至43中任一项所述的DNA供体修复模板、根据权利要求44至50中任一项所述的病毒载体或根据权利要求51至58中任一项所述的细胞。

60. 一种组合物，其包含生理学上可接受的赋形剂和根据权利要求1至43中任一项所述的DNA供体修复模板、根据权利要求44至50中任一项所述的病毒载体或根据权利要求51至58中任一项所述的细胞。

61. 一种用于治疗、预防或改善癌症、感染性疾病、自身免疫疾病、炎性疾病和免疫缺陷或与其相关的病状的至少一种症状的方法，其包含向个体给予有效量的根据权利要求59或权利要求60所述的组合物，和任选地向个体给予有效量的一种或多种编码一种或多种被工程改造的核酸酶的聚核苷酸，所述一种或多种被工程改造的核酸酶被设计成在第一目标位点处引入DSB并且在第二目标位点处引入DSB。

62. 一种用于治疗实体癌的方法,其包含向个体给予有效量的根据权利要求59或权利要求60所述的组合物,和任选地向个体给予有效量的一种或多种编码一种或多种被工程改造的核酸酶的聚核苷酸,所述一种或多种被工程改造的核酸酶被设计成在第一目标位点和第二目标位点处引入DSB。

63. 根据权利要求62所述的方法,其中所述实体癌包含肝癌、胰腺癌、肺癌、乳癌、卵巢癌、前列腺癌、睾丸癌、膀胱癌、脑癌、肉瘤、头颈癌、骨癌、甲状腺癌、肾癌或皮肤癌。

64. 一种用于治疗恶性血液肿瘤的方法,其包含向个体给予有效量的根据权利要求59或权利要求60所述的组合物,和任选地向个体给予有效量的一种或多种编码一种或多种被工程改造的核酸酶的聚核苷酸,所述一种或多种被工程改造的核酸酶被设计成在第一目标位点和第二目标位点处引入DSB。

65. 根据权利要求64所述的方法,其中所述恶性血液肿瘤是白血病、淋巴瘤或多发性骨髓瘤。

66. 根据权利要求1至15中任一项所述的DNA供体修复模板,其中所述第一目标位点或所述第二目标位点包含TCR α 基因中的被工程改造的核酸酶裂解位点。

67. 根据权利要求1至15中任一项所述的DNA供体修复模板,其中所述第一目标位点或所述第二目标位点在以下中包含被工程改造的核酸酶裂解位点:有助于抑制 γ -球蛋白基因表达和HbF的基因,包括(但不限于)BCL11A基因座、KLF1基因座、SOX6基因座、GATA1基因座和LSD1基因座; β -球蛋白基因座的地中海型贫血等位基因;或 β -球蛋白基因座的镰状化等位基因。

68. 根据权利要求67所述的DNA供体修复模板,其中所述被工程改造的核酸酶是选自由以下组成的群组:兆核酸酶、megaTAL、TALEN、ZFN和CRISPR/Cas核酸酶。

69. 根据权利要求67至68中任一项所述的DNA供体修复模板,其中所述被工程改造的核酸酶是由选自由以下组成的群组的LAGLIDADG导向核酸内切酶(LHE)工程改造的兆核酸酶:I-AabMI、I-AaeMI、I-AniI、I-ApaMI、I-CapIII、I-CapIV、I-CkaMI、I-CpaMI、I-CpaMII、I-CpaMIII、I-CpaMIV、I-CpaMV、I-CpaV、I-CraMI、I-EjeMI、I-GpeMI、I-GpiI、I-GzeMI、I-GzeMII、I-GzeMIII、I-HjeMI、I-LtrII、I-LtrI、I-LtrWI、I-MpeMI、I-MveMI、I-NcrII、I-NcrI、I-NcrMI、I-OheMI、I-OnuI、I-OsoMI、I-OsoMII、I-OsoMIII、I-OsoMIV、I-PanMI、I-PanMII、I-PanMIII、I-PnoMI、I-ScuMI、I-SmaMI、I-SscMI和I-Vdi141I。

70. 根据权利要求67至69中任一项所述的DNA供体修复模板,其中所述被工程改造的核酸酶是由选自由以下组成的群组的LHE工程改造的兆核酸酶:I-CpaMI、I-HjeMI、I-OnuI、I-PanMI和SmaMI。

71. 根据权利要求67至70中任一项所述的DNA供体修复模板,其中所述被工程改造的核酸酶是I-OnuI LHE。

72. 根据权利要求67至70中任一项所述的DNA供体修复模板,其中所述被工程改造的核酸酶是I-SmaMI LHE。

73. 根据权利要求67至72中任一项所述的DNA供体修复模板,其中所述被工程改造的核酸酶是包含TALE DNA结合结构域和根据权利要求68至72中任一项所述的被工程改造的兆核酸酶的megaTAL。

74. 根据权利要求73所述的DNA供体修复模板,其中所述TALE结合结构域包含约9.5个

TALE重复单元到约15.5个TALE重复单元。

75. 根据权利要求73所述的DNA供体修复模板,其中所述TALE结合结构域包含约9.5个TALE重复单元、约10.5个TALE重复单元、约11.5个TALE重复单元、约12.5个TALE重复单元、约13.5个TALE重复单元、约14.5个TALE重复单元或约15.5个TALE重复单元。

76. 根据权利要求67至75中任一项所述的DNA供体修复模板,其中所述一个或多个转基因包含编码以下的聚核苷酸: β -球蛋白、 δ -球蛋白、 γ -球蛋白、 β -球蛋白^{A-T87Q}、 β -球蛋白^{A-T87Q/K120E/K95E}或 β -球蛋白^{A-T87Q/G16D/E22A}。

77. 根据权利要求67至76中任一项所述的DNA供体修复模板,其中所述一个或多个转基因进一步包含RNA聚合酶II启动子,所述RNA聚合酶II启动子可操作地连接到所述编码球蛋白或抗镰状化球蛋白的聚核苷酸。

78. 根据权利要求77所述的DNA供体修复模板,其中所述RNA聚合酶II启动子是选自由以下组成的群组:人类 β -球蛋白LCR和启动子、短EF1 α 启动子、长EF1 α 启动子、人类ROSA 26基因座、泛素C (UBC) 启动子、磷酸甘油酸激酶-1 (PGK) 启动子、巨细胞病毒增强子/鸡 β -肌动蛋白 (CAG) 启动子、 β -肌动蛋白启动子和骨髓增生肉瘤病毒增强子、负调控区缺失、dl587rev引物结合位点取代 (MND) 的启动子。

79. 根据权利要求78所述的DNA供体修复模板,其中所述聚核苷酸进一步编码一个或多个自裂解病毒肽,所述一个或多个自裂解病毒肽可操作地连接到一个或多个球蛋白或抗镰状化球蛋白;穿插在一个或多个球蛋白或抗镰状化球蛋白之间;和/或侧接一个或多个球蛋白或抗镰状化球蛋白。

80. 一种病毒载体,其包含根据权利要求67至79中任一项所述的DNA供体修复模板。

81. 根据权利要求80所述的病毒载体,其中所述病毒载体是重组型腺相关病毒载体 (rAAV) 或逆转录病毒。

82. 根据权利要求81所述的病毒载体,其中所述rAAV具有一个或多个来自AAV2的ITR。

83. 根据权利要求81或权利要求82所述的病毒载体,其中所述rAAV具有选自由以下组成的群组的血清型:AAV1、AAV2、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAV9和AAV10。

84. 根据权利要求81所述的病毒载体,其中所述逆转录病毒是慢病毒。

85. 根据权利要求81所述的病毒载体,其中所述慢病毒是整合酶缺陷型慢病毒。

86. 一种细胞,其包含根据权利要求67至79中任一项所述的DNA供体修复模板或根据权利要求80至85中任一项所述的病毒载体。

87. 根据权利要求86所述的细胞,其中已经通过同源定向修复在所述第一目标位点和所述第二目标位点处插入所述一个或多个转基因。

88. 根据权利要求86或权利要求87所述的细胞,其中所述细胞是造血细胞。

89. 根据权利要求86至88中任一项所述的细胞,其中所述细胞是CD34⁺细胞。

90. 根据权利要求86至89中任一项所述的细胞,其中所述细胞是CD133⁺细胞。

91. 一种组合物,其包含根据权利要求67至79中任一项所述的DNA供体修复模板、根据权利要求80至85中任一项所述的病毒载体或根据权利要求86至90中任一项所述的细胞。

92. 一种组合物,其包含生理学上可接受的赋形剂和根据权利要求67至79中任一项所述的DNA供体修复模板、根据权利要求80至85中任一项所述的病毒载体或根据权利要求86至90中任一项所述的细胞。

93. 一种用于治疗个体中的血红蛋白病的方法,其包含向所述个体给予根据权利要求86至90中任一项所述的细胞或根据权利要求91或权利要求92所述的组合物。

94. 一种用于改善个体中的血红蛋白病的至少一种症状的方法,其包含向所述个体给予根据权利要求86至90中任一项所述的细胞或根据权利要求91或权利要求92所述的组合物。

95. 一种用于治疗个体中的地中海型贫血的方法,其包含向所述个体给予有效量的根据权利要求86至90中任一项所述的细胞或根据权利要求91或权利要求92所述的组合物。

96. 根据权利要求95所述的方法,其中所述地中海型贫血是 β -地中海型贫血。

97. 一种用于治疗个体中的镰状细胞病的方法,其包含向所述个体给予有效量的根据权利要求86至90中任一项所述的细胞或根据权利要求91或权利要求92所述的组合物。

98. 一种用于治疗个体中的 β -地中海型贫血的方法,其包含向所述个体给予有效量的根据权利要求86至90中任一项所述的细胞或根据权利要求91或权利要求92所述的组合物。

供体修复模板多重基因组编辑

[0001] 相关申请案的交叉参考

[0002] 本申请案要求2017年2月15日提交的美国临时申请案第62/459,203号在35 U.S.C. §119(e) 下的权利,其以全文引用的方式并入本文中。

[0003] 关于序列表的陈述

[0004] 以文本格式代替纸本拷贝提供与本申请案相关的序列表,并且在此以引用的方式并入本说明书中。含有序列表的文本文件的名称是BLBD_084_01WO_ST25.txt。文本文件是30KB,于2018年2月15日创建,并且通过EFS网络以电子方式与说明书一起提交。

技术领域

[0005] 本公开涉及改良的用于基因组编辑的供体修复模板组合物。更具体地说,本公开涉及被设计成同时靶向基因组中的多个位点的单一供体修复模板。

背景技术

[0006] 仍难以实现有效和同时编辑多个遗传基因座的能力。此外,已证实尤其难以实现使用DNA修复模板和编辑核酸酶(如锌指核酸酶(ZFN)、TALE核酸酶(TALEN)、CRISPR/Cas9核酸酶系统和megaTAL核酸酶)的基因组驱动多个目标基因座处的同源性定向修复的能力。

[0007] 尽管腺相关病毒(AAV) DNA修复模板通常因为其具有低免疫原性并且可以有效驱动同源性定向重组而被选择用于编辑单一基因组基因座,但现有设计不适用于同时靶向多个遗传基因座。用于解决靶向多个遗传基因座的问题的普遍解决方案仍然是使用多个核酸酶和多个DNA供体修复模板。这种解决方案效率低下并且不适用于制备供临床使用的复杂的基于基因组编辑细胞的治疗法。

发明内容

[0008] 本文中涵盖改良的免疫效应细胞组合物和使用多重基因组编辑组合物制备其的方法。

[0009] 在各种实施例中,提供一种DNA供体修复模板,其包含:用于第一目标位点的第一对同源臂;用于第二目标位点的第二对同源臂;和一个或多个转基因。

[0010] 在具体实施例中,第一对同源臂包含与第一目标位点的DNA序列5'同源的5'同源臂;和与第一目标位点的DNA序列3'同源的3'同源臂。

[0011] 在具体实施例中,第二对同源臂包含与第二目标位点的DNA序列5'同源的5'同源臂;和与第二目标位点的DNA序列3'同源的3'同源臂。

[0012] 在某些实施例中,用于第一目标位点和/或第二目标位点的5'和3'同源臂的长度是独立地选自约100bp到约2500bp。

[0013] 在一些实施例中,用于第一目标位点和/或第二目标位点的5'和3'同源臂的长度是独立地选自约100bp到约1500bp。

[0014] 在其它实施例中,用于第一目标位点和/或第二目标位点的5'和3'同源臂的长度

是独立地选自约100bp到约1000bp。

[0015] 在具体实施例中,用于第一目标位点和/或第二目标位点的5'和3'同源臂的长度是独立地选自约100bp到约500bp。

[0016] 在其它实施例中,用于第一目标位点和/或第二目标位点的5'和3'同源臂的长度是约500bp。

[0017] 在一些实施例中,用于第一目标位点和/或第二目标位点的5'和3'同源臂的长度是约300bp。

[0018] 在其它实施例中,用于第一目标位点和/或第二目标位点的5'和3'同源臂的长度是约100bp。

[0019] 在具体实施例中,第一目标位点的5'同源臂位于第二目标位点的5'同源臂的5'。

[0020] 在某些实施例中,第一目标位点的3'同源臂位于第二目标位点的3'同源臂的5'。

[0021] 在一些实施例中,第一目标位点的3'同源臂位于第二目标位点的3'同源臂的3'。

[0022] 在其它实施例中,第一目标位点的5'同源臂和3'同源臂侧接第一转基因并且第二目标位点的5'同源臂和3'同源臂侧接第二转基因。

[0023] 在其它实施例中,第一目标位点和第二目标位点在不同的基因中。

[0024] 在具体实施例中,第一目标位点或第二目标位点包含TCR α 基因中的被工程改造的核酸酶裂解位点。

[0025] 在一些实施例中,第一目标位点或第二目标位点包含免疫系统检查点基因或编码免疫抑制性信号传导组件的基因中的被工程改造的核酸酶裂解位点。

[0026] 在具体实施例中,第一目标位点是TCR α 基因中的被工程改造的核酸酶裂解位点并且第二目标位点是免疫系统检查点基因或编码免疫抑制性信号传导组件的基因中的被工程改造的核酸酶裂解位点。

[0027] 在某些实施例中,第一目标位点是第一免疫系统检查点基因或编码免疫抑制性信号传导组件的基因中的被工程改造的核酸酶裂解位点,并且第二目标位点是第二免疫系统检查点基因或编码免疫抑制性信号传导组件的基因中的被工程改造的核酸酶裂解位点。

[0028] 在其它实施例中,免疫系统检查点基因是独立地选自由以下组成的群组:程序性细胞死亡蛋白质1 (PD-1)、淋巴细胞活化基因3蛋白质 (LAG-3)、T细胞免疫球蛋白结构域和粘蛋白结构域蛋白质3 (TIM-3)、细胞毒性T淋巴细胞抗原-4 (CTLA-4)、带状T淋巴细胞弱化子 (BTLA)、T细胞免疫球蛋白和基于免疫受体酪氨酸的抑制性基元结构域 (TIGIT)、T细胞活化的V-结构域Ig抑制因子 (VISTA) 和杀手细胞免疫球蛋白样受体 (KIR) 基因。

[0029] 在某些实施例中,编码免疫抑制性信号传导组件的基因是独立地选自由以下组成的群组:白细胞介素受体10 α 、转化生长因子 β 受体I (TGF β RI)、转化生长因子 β 受体II (TGF β RII)、芳基烃受体 (AHR)、血清和糖皮质激素调节的激酶1 (SGK1)、结节性硬化症复合物2 (TSC2)、冯希派尔-林道肿瘤抑制因子 (von Hippel-Lindau tumor suppressor; VHL)、腺苷A2a受体 (A2AR) 和Cb1原癌基因B (CBLB)。

[0030] 在具体实施例中,被工程改造的核酸酶是选自由以下组成的群组:兆核酸酶、megaTAL、TALEN、ZFN和CRISPR/Cas核酸酶。

[0031] 在其它实施例中,被工程改造的核酸酶是由选自由以下组成的群组的LAGLIDADG导向核酸内切酶 (LHE) 工程改造的兆核酸酶: I-AabMI、I-AaeMI、I-AniI、I-ApaMI、I-

CapIII、I-CapIV、I-CkaMI、I-CpaMI、I-CpaMII、I-CpaMIII、I-CpaMIV、I-CpaMV、I-CpaV、I-CraMI、I-EjeMI、I-GpeMI、I-GpiI、I-GzeMI、I-GzeMII、I-GzeMIII、I-HjeMI、I-LtrII、I-LtrI、I-LtrWI、I-MpeMI、I-MveMI、I-NcrII、I-NcrI、I-NcrMI、I-OheMI、I-OnuI、I-OsoMI、I-OsoMII、I-OsoMIII、I-OsoMIV、I-PanMI、I-PanMII、I-PanMIII、I-PnoMI、I-ScuMI、I-SmaMI、I-SscMI和I-Vdi141I。

[0032] 在一些实施例中,被工程改造的核酸酶是由选自由以下组成的群组的LHE工程改造的兆核酸酶:I-CpaMI、I-HjeMI、I-OnuI、I-PanMI和SmaMI。

[0033] 在具体实施例中,被工程改造的核酸酶是I-OnuI LHE。

[0034] 在其它实施例中,被工程改造的核酸酶是I-SmaMI LHE。

[0035] 在某些实施例中,被工程改造的核酸酶是megaTAL,其包含TALE DNA结合结构域和被工程改造的兆核酸酶。

[0036] 在其它实施例中,TALE结合结构域包含约9.5个TALE重复单元到约15.5个TALE重复单元。

[0037] 在具体实施例中,TALE结合结构域包含约9.5个TALE重复单元、约10.5个TALE重复单元、约11.5个TALE重复单元、约12.5个TALE重复单元、约13.5个TALE重复单元、约14.5个TALE重复单元或约15.5个TALE重复单元。

[0038] 在其它实施例中,一个或多个转基因包含编码以下的聚核苷酸:免疫效能增强子、免疫抑制信号减弱子或被工程改造的抗原受体。

[0039] 在某些实施例中,一个或多个转基因进一步包含可操作地连接到编码以下的聚核苷酸的RNA聚合酶II启动子:免疫效能增强子、免疫抑制信号减弱子或被工程改造的抗原受体。

[0040] 在一些实施例中,一个或多个转基因编码一个或多个CAR。

[0041] 在具体实施例中,一个或多个CAR是选自由以下组成的群组:抗BCMA CAR和抗CD19 CAR。

[0042] 在其它实施例中,RNA聚合酶II启动子是选自由以下组成的群组:短EF1 α 启动子、长EF1 α 启动子、人类ROSA 26基因座、泛素C (UBC) 启动子、磷酸甘油酸激酶-1 (PGK) 启动子、巨细胞病毒增强子/鸡 β -肌动蛋白 (CAG) 启动子、 β -肌动蛋白启动子和骨髓增生肉瘤病毒增强子、负调控区缺失、d1587rev引物结合位点取代 (MND) 的启动子。

[0043] 在其它实施例中,聚核苷酸进一步编码一种或多种与以下可操作地连接、穿插在以下之间和/或与以下侧接的自裂解型病毒肽:免疫效能增强子、免疫抑制信号减弱子或被工程改造的抗原受体。

[0044] 在其它实施例中,自裂解型病毒肽是2A肽。

[0045] 在具体实施例中,聚核苷酸进一步包含异源聚腺苷酸化信号。

[0046] 在某些实施例中,一个或多个免疫抑制信号减弱子包含对抗免疫抑制因子的酶功能;结合免疫抑制因子的胞外结构域,任选地其中胞外结构域是抗体或其抗原结合片段;结合免疫抑制因子的胞外结构域和跨膜结构域;或结合免疫抑制因子的胞外结构域、跨膜结构域和不能转换免疫抑制信号的被修饰的胞内结构域。

[0047] 在具体实施例中,一个或多个免疫效能增强子是独立地选自由以下组成的群组:双特异性T细胞接合分子 (BiTE)、免疫增强因子和翻转受体。

- [0048] 在一些实施例中,一个或多个免疫效能增强子是独立地选自由以下组成的群组:细胞因子、趋化因子、细胞毒素、细胞因子受体和其变异体。
- [0049] 在具体实施例中,一个或多个被工程改造的抗原受体是独立地选自由以下组成的群组:被工程改造的TCR、CAR、DARIC或 ζ 因子。
- [0050] 在其它实施例中,一个或多个被工程改造的抗原受体是CAR。
- [0051] 在具体实施例中,一个或多个CAR是选自由以下组成的群组:抗BCMA CAR和抗CD19 CAR。
- [0052] 在各种实施例中,提供包含本文中所涵盖的DNA供体修复模板的病毒载体。
- [0053] 在其它实施例中,病毒载体是重组型腺相关病毒载体(rAAV)或逆转录病毒。
- [0054] 在某些实施例中,rAAV具有来自AAV2的一个或多个ITR。
- [0055] 在具体实施例中,rAAV具有选自由以下组成的群组的血清型:AAV1、AAV2、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAV9和AAV10。
- [0056] 在一些实施例中,rAAV具有AAV6血清型。
- [0057] 在一些实施例中,逆转录病毒是慢病毒。
- [0058] 在其它实施例中,慢病毒是整合酶缺陷型慢病毒。
- [0059] 在各种实施例中,提供包含本文中所涵盖的DNA供体修复模板和/或病毒载体的细胞。
- [0060] 在某些实施例中,已经通过同源性定向修复在第一目标位点和第二目标位点处插入一个或多个转基因。
- [0061] 在一些实施例中,细胞是造血细胞。
- [0062] 在具体实施例中,细胞是免疫效应细胞。
- [0063] 在某些实施例中,细胞是CD3⁺、CD4⁺、CD8⁺或其组合。
- [0064] 在其它实施例中,细胞是T细胞。
- [0065] 在其它实施例中,细胞是细胞毒性T淋巴细胞(CTL)、肿瘤浸润淋巴细胞(TIL)或辅助T细胞。
- [0066] 在具体实施例中,细胞来源是外周血液单核细胞、骨髓、淋巴结组织、脐带血、胸腺组织、来自感染位点的组织、腹水、胸腔积液、脾组织或肿瘤。
- [0067] 在各种实施例中,提供包含本文中所涵盖的DNA供体修复模板、病毒载体或细胞的组合物。
- [0068] 在各种实施例中,提供一种组合物,其包含生理学上可接受的赋形剂和本文中所涵盖的DNA供体修复模板、病毒载体或细胞。
- [0069] 在各种实施例中,提供用于治疗、预防或改善癌症、感染性疾病、自身免疫疾病、发炎性疾病和免疫缺陷或与其相关的病状的至少一种症状的方法,其包含向个体给予有效量的本文中所涵盖的组合物,和任选地向个体给予有效量的一种或多种编码一种或多种被工程改造的核酸酶的聚核苷酸,所述一种或多种被工程改造的核酸酶被设计成在第一目标位点处引入DSB并且在第二目标位点处引入DSB。
- [0070] 在各种实施例中,提供用于治疗实体癌的方法,其包含向个体给予有效量的本文中所涵盖的组合物,和任选地向个体给予有效量的一种或多种编码一种或多种被工程改造的核酸酶的聚核苷酸,所述一种或多种被工程改造的核酸酶被设计成在第一目标位点和第

二目标位点处引入DSB。

[0071] 在一些实施例中,实体癌包含肝癌、胰腺癌、肺癌、乳癌、卵巢癌、前列腺癌、睾丸癌、膀胱癌、脑癌、肉瘤、头颈癌、骨癌、甲状腺癌、肾癌或皮肤癌。

[0072] 在各种实施例中,提供用于治疗恶性血液肿瘤的方法,其包含向个体给予有效量的本文中所涵盖的组合物,和任选地向个体给予有效量的一种或多种编码一种或多种被工程改造的核酸酶的聚核苷酸,所述一种或多种被工程改造的核酸酶被设计成在第一目标位点和第二目标位点处引入DSB。

[0073] 在具体实施例中,恶性血液肿瘤是白血病、淋巴瘤或多发性骨髓瘤。

附图说明

[0074] 图1显示用于使用多个AAV供体修复模板在多个目标基因座处驱动同源定向修复的常用和效率低下的策略。

[0075] 图2显示用于使用单一AAV供体修复模板在多个目标基因座处有效驱动同源定向修复的代表性本发明的策略。

[0076] 图3A显示同时靶向PD-1基因和TCR α 恒定区的单一AAV供体修复模板的代表性设计。

[0077] 图3B显示同源定向修复实验的结果,所述实验使用PD-1和TCR α megaTAL以及单一AAV供体修复模板,所述修复模板同时靶向PD-1基因和TCR α 恒定区以引入编码GFP的聚核苷酸序列。

[0078] 图4显示TCR α 恒定区的基因敲除效率,如通过不含CD3表达所测量。单独的TCR α megaTAL或其与PD-1megaTAL一起显示超过80%的基因敲除效率(泳道3和4)。

[0079] 图5显示与未用PD-1megaTAL处理的细胞(泳道1、3和5)相比,在24小时PMA/离子霉素刺激之后,用PD-1megaTAL处理的细胞(泳道2和4)中的PD-1表达减少。

[0080] 图6A显示同时靶向PD-1基因和TCR α 恒定区的单一AAV供体修复模板的代表性设计。

[0081] 图6B显示同源定向修复实验的结果,所述实验使用PD-1和TCR α megaTAL以及单一AAV供体修复模板,所述修复模板同时靶向PD-1基因和TCR α 恒定区以引入编码BCMA CAR的聚核苷酸序列。

[0082] 图7显示双重AAV供体的代表性实施例。单一插入物或两个不同的插入物可用于双重AAV供体。视用途而定,插入物和同源臂的位置和定向可以具有多种组合。可以增加不同同源臂的数目以用于靶向具有单一AAV供体的多个目标基因座。

[0083] 序列标识符的简要说明

[0084] SEQ ID NO:1阐述AAV供体修复模板的聚核苷酸序列。

[0085] SEQ ID NO:2阐述AAV供体修复模板的聚核苷酸序列。

[0086] SEQ ID NO:3-13阐述各种连接子的氨基酸序列。

[0087] SEQ ID NO:14-38阐述蛋白酶裂解位点和自裂解型多肽裂解位点的氨基酸序列。

具体实施方式

[0088] A. 概述

[0089] 本文中所涵盖的各种实施例通常部分涉及改良的基因组编辑组合物和使用其制备改良的免疫效应细胞组合物的方法。改良的基因组编辑组合物和方法使得能够使用单一DNA供体修复模板编辑基因组中的多个目标位点,由此极大地简化和提高治疗性细胞群体中的多重基因组编辑效率。预期具有多个基因组编辑物的治疗性细胞用于治疗或预防多种病状,包括(但不限于)血红蛋白病、癌症、感染性疾病、自身免疫疾病、炎症性疾病和免疫缺陷。

[0090] 在各种实施例中,提供一种DNA供体修复模板,其实现一个或多个目标位点的基因组编辑。

[0091] 在各种实施例中,提供一种载体,其包含实现一个或多个目标位点的基因组编辑的DNA供体修复模板。

[0092] 在各种实施例中,提供一种细胞,其包含一种或多种被工程改造的核酸酶和实现一个或多个目标位点的基因组编辑的DNA供体修复模板。在具体实施例中,细胞是造血细胞,包括(但不限于)造血干细胞、造血祖细胞、CD34⁺细胞、免疫效应细胞和T细胞。

[0093] 在具体实施例中,适用于本文中所涵盖的多重基因组编辑的目标位点包括(但不限于)一个或多个有助于抑制 γ -球蛋白基因表达、血红蛋白病、地中海型贫血和镰状细胞病的基因座。

[0094] 基因组编辑的造血干细胞、造血祖细胞或具有多个基因组编辑物的CD34⁺细胞可以在以下中包含一种或多种编辑物:有助于抑制 γ -球蛋白基因表达和HbF的基因,包括(但不限于)BCL11A基因座、KLF1基因座、SOX6基因座、GATA1基因座和LSD1基因座; β -球蛋白基因座的地中海型贫血等位基因;或 β -球蛋白基因座的镰状化等位基因。不希望受任何具体理论约束,预期被设计成通过基因组编辑来精确破坏或修饰球蛋白基因表达的被工程改造的核酸酶可以增强治疗性血色素表达的产生并且治疗、预防或改善血红蛋白病的至少一种症状。

[0095] 在具体实施例中,本文中所涵盖的适用于多重基因组编辑的目标位点包括(但不限于)一个或多个有助于T细胞受体(TCR)信号传导的基因座、免疫系统检查点基因和编码免疫抑制性信号传导组件的基因。

[0096] 具有多个基因组编辑物的被基因组编辑的T细胞可以在基因中包含一种或多种编辑物,所述编辑物有助于TCR信号传导,包括(但不限于)TCR α (TCR α)基因座和TCR β (TCR β)基因座。

[0097] 具有多个基因组编辑物的被基因组编辑的T细胞可以在编码免疫系统检查点基因或编码免疫抑制性信号传导组件的基因中包含或进一步包含一种或多种编辑物。包含降低免疫系统检查点基因和免疫抑制性信号传导组件的功能的编辑物的T细胞在肿瘤微环境中对T细胞消耗的抗性更高,并且具有增加的T细胞耐久性和持久性。

[0098] 免疫系统检查点基因的说明性实例包括(但不限于):程序性细胞死亡蛋白质1(PD-1)、淋巴细胞活化基因3蛋白质(LAG-3)、T细胞免疫球蛋白结构域和粘蛋白结构域蛋白质3(TIM-3)、细胞毒性T淋巴细胞抗原-4(CTLA-4)、带状T淋巴细胞弱化子(BTLA)、T细胞免疫球蛋白和基于免疫受体酪氨酸的抑制性基元结构域(TIGIT)、T细胞活化的V-结构域Ig抑制因子(VISTA)和杀手细胞免疫球蛋白样受体(KIR)基因。

[0099] 编码免疫抑制性信号传导组件的基因的说明性实例包括(但不限于):白细胞介素

10受体 α (IL-10R α)、转化生长因子 β 受体1 (TGFBR1)、转化生长因子 β 受体2 (TGFBR2)、芳香基炔受体 (AHR)、血清/糖皮质激素调节的激酶1 (SGK1)、结节性硬化症复合物2 (TSC2)、腺苷A2A受体 (A2AR)、冯希派尔-林道肿瘤抑制因子 (VHL) 和Cb1原癌基因B (CBLB)。

[0100] 不希望受任何具体的理论约束,预期包含一个或多个破坏一个或多个TCR信号传导组件的功能的基因组编辑物的T细胞将提供更安全和更有效的过继性细胞疗法,因为其基本上不含功能性内源性TCR表达,由此减少潜在的移植排斥并且因为其释放辅助信号传导组件以增强被工程改造的抗原受体T细胞的功效。此外,在免疫系统检查点基因或免疫抑制性信号传导组件中包含一种或多种基因组编辑物的T细胞将提供过继性细胞疗法,所述过继性细胞疗法在肿瘤微环境中具有降低的对T细胞消耗的敏感性、增加的T细胞耐久性和增加的T细胞持久性。

[0101] 在各种实施例中,细胞包含多个被工程改造的核酸酶,其在多个目标位点处诱导DSB,和DNA供体修复模板,其通过同源性定向修复 (HDR) 来编码插入多个目标位点中的一个或多个转基因。

[0102] 具体实施例中涵盖的适用于DNA供体修复模板中的转基因的说明性实例包括(但不限于):球蛋白基因和抗镰状化球蛋白基因。

[0103] 具体实施例中涵盖的适用于DNA供体修复模板中的转基因的说明性实例包括(但不限于):免疫效能增强子、免疫抑制信号减弱子或被工程改造的抗原受体。不希望受任何具体的理论约束,预期与仅包含基因分裂编辑物或仅包含免疫效能增强子、免疫抑制信号减弱子或被工程改造的抗原受体的表达的过继性细胞疗法相比,破坏编码有助于TCR信号传导、免疫系统检查点和/或免疫抑制性信号传导的基因的多个目标位点并且将一个或多个编码免疫效能增强子、免疫抑制信号减弱子或被工程改造的抗原受体的转基因引入基因座中可以使所得过继性细胞疗法具有更好的安全性、功效、耐久性和/或持久性概况。

[0104] 在各种实施例中,提供一种编辑多个目标位点处的细胞的基因组的方法,其是通过引入一种或多种编码一种或多种被工程改造的核酸酶的聚核苷酸和实现一个或多个目标位点的基因组编辑的DNA供体模板来进行。

[0105] 在各种实施例中,提供用于使用本文中所涵盖的基因组编辑的细胞疗法治疗、预防或改善血红蛋白病、癌症、感染性疾病、自身免疫疾病、炎症性疾病和免疫缺陷或与其相关的病状的至少一种症状的方法。

[0106] 除非明确相反地指出,否则具体实施例的实践将使用在所属领域的技术范围内的化学、生物化学、有机化学、分子生物学、微生物学、重组型DNA技术、遗传学、免疫学以及细胞生物学的常规方法,其中许多方法出于说明的目的而在下文中描述。在文献中全面说明这类技术。参见例如Sambrook等人,《分子克隆:实验指南 (Molecular Cloning: A Laboratory Manual)》(第3版,2001);Sambrook等人,《分子克隆:实验指南》(第2版,1989);Maniatis等人,《分子克隆:实验指南》(1982);Ausubel等人,《现代分子生物学实验技术 (Current Protocols in Molecular Biology)》(John Wiley and Sons,2008年7月更新);《精编分子生物学实验指南:现代分子生物学实验技术方法概要 (Short Protocols in Molecular Biology: A of Methods from Current Protocols in Molecular Biology)》,Greene Pub.Associates and Wiley-Interscience;Glover,《DNA克隆:实用方法 (DNA Cloning: A Practical Approach)》,第I卷和第II卷 (IRL Press,牛津 (Oxford),1985);

Anand,《复杂基因组分析技术(Techniques for the Analysis of Complex Genomes)》, (Academic Press, 纽约(New York), 1992);《转录与翻译(Transcription and Translation)》(B.Hames和S.Higgins编, 1984);Perbal,《分子克隆实践指南(A Practical Guide to Molecular Cloning)》(1984);Harlow和Lane,《抗体(Antibodies)》, (Cold Spring Harbor Laboratory Press, 纽约冷泉港(Cold Spring Harbor, N.Y.), 1998);《现代免疫学方法(Current Protocols in Immunology)》, Q.E.Coligan, A.M.Kruisbeek, D.H.Margulies, E.M.Shevach和W.Strober编, 1991);《免疫学年评(Annual Review of Immunology)》;以及如《免疫学进展(Advances in Immunology)》等杂志中的专论。

[0107] B. 定义

[0108] 除非另有定义, 否则本文中使用的所有技术和科学术语都具有与本发明所属领域的普通技术人员通常所了解相同的含义。尽管可以使用与本文中所描述的方法和材料类似或等效的任何方法和材料来实践或测试具体实施例, 但本文中描述组合物、方法和材料的优选实施例。出于本发明的目的, 下文定义以下术语。

[0109] 本文中使用的冠词“一(a/an)”和“所述(the)”是指冠词的语法对象中的一个或超过一个(即至少一个或一个或多个)。作为实例, “一个元件”意指一个元件或一个或多个元件。

[0110] 替代方案(例如“或”)的使用应理解为意指替代方案中的一个、两个或其任何组合。

[0111] 术语“和/或”应理解成意指替代方案中的一个或两个。

[0112] 如本文中所使用, 术语“约”或“大约”是指数量、水平、数值、数目、频率、百分比、尺度、尺寸、量、重量或长度与参考数量、水平、数值、数目、频率、百分比、尺度、尺寸、量、重量或长度相比, 变化多达15%、10%、9%、8%、7%、6%、5%、4%、3%、2%或1%。在一个实施例中, 术语“约”或“大约”是指关于参考数量、水平、值、数目、频率、百分比、维度、尺寸、量、重量或长度, 数量、水平、值、数目、频率、百分比、维度、尺寸、量、重量或长度范围±15%、±10%、±9%、±8%、±7%、±6%、±5%、±4%、±3%、±2%或±1%。

[0113] 在一个实施例中, 范围, 例如1到5、约1到5或约1到约5是指所述范围所涵盖的每个数值。举例来说, 在一个非限制性和仅说明性实施例中, 范围“1到5”与表述1、2、3、4、5; 或1.0、1.5、2.0、2.5、3.0、3.5、4.0、4.5或5.0; 或1.0、1.1、1.2、1.3、1.4、1.5、1.6、1.7、1.8、1.9、2.0、2.1、2.2、2.3、2.4、2.5、2.6、2.7、2.8、2.9、3.0、3.1、3.2、3.3、3.4、3.5、3.6、3.7、3.8、3.9、4.0、4.1、4.2、4.3、4.4、4.5、4.6、4.7、4.8、4.9或5.0等效。

[0114] 如本文中所使用, 术语“基本上”是指数量、水平、值、数目、频率、百分比、尺寸、大小、量、重量或长度是参考数量、水平、值、数目、频率、百分比、尺寸、大小、量、重量或长度的80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更高。在一个实施例中, “基本上相同”是指某个数量、水平、值、数目、频率、百分比、尺寸、大小、量、重量或长度产生的作用, 例如生理作用, 与参考数量、水平、值、数目、频率、百分比、尺寸、大小、量、重量或长度产生的作用大致相同。

[0115] 在本说明书通篇, 除非上下文另外要求, 否则“包含(comprise/comprises/comprising)”一词应理解为暗指包括所述步骤或要素或一组步骤或要素, 但不排除任何其它步骤或要素或任何其它组步骤或要素。“由……组成”意图包括并且限于短语“由……组

成”之后的任何事物。因此,短语“由……组成”指示所列要素是需要或必需的,并且不能存在其它要素。“基本上由……组成”意图包括短语后所列的任何要素,并且限于不干扰或影响本发明中规定的所列要素的活性或作用的其它要素。因此,短语“基本上由……组成”指示所列要素是必不可少的或必选的,并且不能存在显著影响所列要素的活性或作用的其它要素。

[0116] 在本说明书通篇,对“一个实施例”、“实施例”、“特定实施例”、“相关实施例”、“某一实施例”、“其它实施例”、“各种实施例”或“另一实施例”或其复数形式或组合的参考意谓在至少一个实施例中包括关于所述实施例描述的具体特性、结构或特征。因此,贯穿本说明书中的各处出现的前述短语不一定全部指同一个实施例。此外,在一个或多个实施例中,具体特性、结构或特征可以任何合适方式组合。还应理解,在一个实施例中对一个特征的正面叙述用作在一个具体实施例中排除所述特征的基础。

[0117] “分离的细胞”是指从体内组织或器官获得并且基本上不含细胞外基质的非天然存在的细胞,例如自然界中不存在的细胞、经过修饰的细胞、工程改造的细胞等。在一个实施例中,被分离的细胞包含一个或多个不天然存在于所述细胞中的聚核苷酸序列。

[0118] 术语“干细胞”是指能够满足以下条件的未分化细胞:(1)长期自我更新,或能够产生初始细胞的至少一个相同拷贝,(2)在单细胞水平下分化成多个,并且在一些情况下,仅一个特殊的细胞类型;以及(3)实现组织的体内功能性再生。根据干细胞的发育潜力将干细胞细分为全能、亚全能、多能以及寡能/单能。“自我更新”是指细胞具有产生未改变的子细胞并且产生特殊的细胞类型(效能)的独特能力。自我更新可以通过两种方式实现。不对称细胞分裂产生一个与亲本细胞相同的子细胞和一个与亲本细胞不同的子细胞,所述与亲本细胞不同的子细胞是祖细胞或分化细胞。对称细胞分裂产生两个相同的子细胞。

[0119] 如本文中所使用,术语“祖代”或“祖细胞”是指具有自我更新并且分化成多个成熟细胞的能力的细胞。许多祖细胞沿着单一谱系分化,但可能具有相当广泛的增殖能力。

[0120] 如本文中所使用的术语“原代细胞”在所属领域中是已知的,是指已从组织分离并且已建立用于体外或离体生长的细胞。相应的细胞经历了非常少的(如果有的话)群体倍增,并且因此与传代细胞系相比,它们更能代表它们所衍生的组织的主要功能组分,因此代表了体内状态的更具代表性的模型。从各种组织获得样品的方法和建立原代细胞系的方法在所属领域中是众所周知的(参见例如Jones和Wise,《分子生物学方法(Methods Mol Biol.)》1997)。在具体实施例中,细胞是原代细胞。用于本文中所涵盖的方法中的原生细胞是来源于脐带血、胎盘血、动员后外周血和骨髓。

[0121] 如本文中所使用,术语“细胞群体”是指可以由任何数目和/或任何组合的如本文中其它地方所述的同质或异质细胞类型构成的多个细胞。细胞群体可以包含约10%、约20%、约30%、约40%、约50%、约60%、约70%、约80%、约90%或约100%的待编辑的目标细胞类型。

[0122] 术语“造血干细胞”或“HSC”是指可以产生生物体的所有血细胞类型的多能干细胞,包括骨髓(例如单核细胞和巨噬细胞、嗜中性粒细胞、嗜碱粒细胞、嗜酸性粒细胞、红细胞、巨核细胞/血小板、树突状细胞)和淋巴谱系(例如T细胞、B细胞、NK细胞)以及所属领域中已知的其它细胞(参看Fei,R.等人,美国专利案第5,635,387号;McGlave等人,美国专利案第5,460,964号;Simmons,P.等人,美国专利案第5,677,136号;Tsukamoto等人,美国专利

案第5,750,397号;Schwartz等人,美国专利案第5,759,793号;DiGuisto等人,美国专利案第5,681,599号;Tsukamoto等人,美国专利案第5,716,827号)。当移植到被致死性照射过的动物或人类中时,造血干细胞和祖细胞可以重建(repopulate)红细胞、嗜中性粒细胞-巨噬细胞、巨核细胞以及淋巴造血细胞池。

[0123] 如本文中所使用,术语“CD34⁺细胞”是指在其细胞表面上表达CD34蛋白质的细胞。如本文中所使用,“CD34”是指通常充当细胞-细胞粘附因子的细胞表面糖蛋白(例如唾液粘蛋白(sialomucin)蛋白质)。CD34⁺是造血干细胞和祖细胞的细胞表面标记。CD34⁺是造血干细胞和祖细胞的细胞表面标记。适用于本文中所涵盖的具体实施例中的造血细胞包括(但不限于):CD34⁺细胞。适用于本文中所涵盖的具体实施例中的造血细胞包括(但不限于):CD34⁺CD38^{Lo}CD90⁺CD45RA⁻细胞、CD34⁺、CD59⁺、Thy1/CD90⁺、CD38^{Lo/-}、C-kit/CD117⁺和Lin⁽⁻⁾细胞以及CD34⁺、CD133⁺细胞。

[0124] “免疫效应细胞”是免疫系统具有一种或多种效应功能(例如细胞毒性细胞杀伤活性、细胞因子的分泌、ADCC和/或CDC的诱导)的任何细胞。具体实施例中所涵盖的说明性免疫效应细胞是T淋巴细胞,特别是细胞毒性T细胞(CTL;CD8⁺T细胞)、TIL和辅助T细胞(HTL;CD4⁺T细胞)。在一个实施例中,免疫效应细胞包括自然杀手(NK)细胞。在一个实施例中,免疫效应细胞包括自然杀手T(NKT)细胞。

[0125] 术语“T细胞”或“T淋巴细胞”是所属领域中公认的并且意图包括胸腺细胞、原生T淋巴细胞、未成熟的T淋巴细胞、成熟T淋巴细胞、静息T淋巴细胞或活化的T淋巴细胞。T细胞可以是T辅助(Th)细胞,例如T辅助1(Th1)或T辅助2(Th2)细胞。T细胞可以是辅助T细胞(HTL;CD4⁺T细胞)CD4⁺T细胞、细胞毒性T细胞(CTL;CD8⁺T细胞)、肿瘤浸润细胞毒性T细胞(TIL;CD8⁺T细胞)、CD4⁺CD8⁺T细胞、CD4⁻CD8⁻T细胞或任何其它T细胞子集。在一个实施例中,T细胞是NKT细胞。适用于具体实施例中的其它说明性T细胞群体包括原生T细胞和记忆T细胞。

[0126] 在具体实施例中,“强效T细胞”和“幼T细胞”可以互换使用并且是指其中T细胞能够增殖并且伴随分化减少的T细胞表型。在具体实施例中,幼T细胞具有“原生T细胞”的表型。在具体实施例中,幼T细胞包含以下生物标记物中的一种或多种或全部:CD62L、CCR7、CD28、CD27、CD122、CD127、CD197和CD38。在一个实施例中,幼T细胞包含以下生物标记物中的一种或多种或全部:CD62L、CD127、CD197和CD38。在一个实施例中,幼T细胞不表达CD57、CD244、CD160、PD-1、CTLA4、TIM3和LAG3。

[0127] 如本文中所使用,术语“增殖”是指细胞分裂增加,无论细胞的对称或不对称分裂。在具体实施例中,“增殖”是指造血细胞的对称或不对称分裂。当被处理的样品中的细胞与未被处理的样品中的细胞相比数目增加时,发生“增殖增加”。

[0128] 如本文中所使用,术语“分化”是指一种降低细胞的效能或增殖或使细胞转换成发育受限程度更高的状态的方法。在具体实施例中,分化T细胞获取免疫效应细胞功能。

[0129] 如本文中所使用,术语“T细胞制备”或“制备T细胞的方法”或类似术语是指产生T细胞的治疗组合物的方法,所述制备方法可以包含以下步骤中的一个或多个或全部:收集、刺激、活化、基因组编辑和扩增。

[0130] 术语“离体”一般是指在生物体外部发生的活动,如在生物体外部的人工环境中,在活组织中或活组织上进行的实验或测量,优选其中天然条件的变化最小。在具体实施例

中，“离体”程序涉及从生物体获得活细胞或组织并且在实验室装置中，通常在无菌条件下培养或调节，且取决于情况，典型地持续数小时或长达约24小时，但包括长达48或72小时。在某些实施例中，可以收集和冷冻这类组织或细胞，并且日后解冻以便进行离体处理。持续时间长于几天的使用活细胞或组织的组织培养实验或程序典型地被视为“试管内”，但在某些实施例中，这个术语可以与离体互换使用。

[0131] 术语“体内”通常是指在生物体内进行的活动，如细胞自我更新和细胞增殖或扩增。在一个实施例中，术语“体内扩增”是指细胞群体体内数目增加的能力。在一个实施例中，细胞是被体内工程改造或修饰的。

[0132] 术语“刺激”是指由刺激分子(例如TCR/CD3复合物)与其同源配位体结合，由此介导信号转导事件(包括但不限于)通过TCR/CD3复合物进行的信号转导)而诱导的原初反应。

[0133] “刺激分子”是指与同源刺激配位体特异性结合的T细胞上的分子。

[0134] 如本文中所示使用，“刺激配位体”意指满足以下条件的配位体：当存在于抗原呈现细胞(例如aAPC、树突状细胞、B细胞等)上时，可以与T细胞上的同源结合搭配物(本文中称为“刺激分子”)特异性结合，由此通过T细胞介导原初反应，包括(但不限于)活化、免疫反应起始、增殖等。刺激配位体包括(但不限于)CD3配位体，例如抗CD3抗体和CD2配位体，例如抗CD2抗体，和肽，例如CMV、HPV、EBV肽。

[0135] 术语“活化”是指T细胞已被充分刺激以诱导可检测的细胞增殖的状态。在具体实施例中，活化也可以与所诱导的细胞因子产生和可检测的效应功能相关。术语“活化的T细胞”尤其是指正在增殖的T细胞。由单独的TCR产生的信号不足以实现T细胞的完全活化，并且还需要一个或多个二级或共刺激信号。因此，T细胞活化包含通过TCR/CD3复合物产生的初级刺激信号和一个或多个二级共刺激信号。共刺激可以由接受初级活化信号(如由CD3/TCR复合物或CD2进行的刺激)的T细胞的增殖和/或细胞因子产量证明。

[0136] “共刺激信号”是指一种信号，其与初级信号(如TCR/CD3接合)组合，引起T细胞增殖、细胞因子产生和/或具体分子(例如CD28)的上调或下调。

[0137] “共刺激配位体”是指结合共刺激分子的分子。共刺激配位体可以是可溶的或提供于表面上。“共刺激分子”是指与共刺激配位体(例如抗CD28抗体)特异性结合的T细胞上的同源结合搭配物。

[0138] “抗原(Ag)”是指可以刺激动物中的抗体产生或T细胞应答的化合物、组合物或物质，例如脂质、碳水化合物、多糖、糖蛋白、肽或核酸，包括注射或吸收到动物中的组合物(如包括肿瘤特异性蛋白质的组合物)。抗原与特异性体液或细胞免疫产物反应，所述产物包括由异源抗原(如所公开的抗原)诱导的产物。“目标抗原”或“相关目标抗原”是本文中所涵盖的被工程改造的抗原受体的结合结构域经设计以结合的抗原。在一个实施例中，抗原是MHC-肽复合物，如I类MHC-肽复合物或II类MHC-肽复合物。

[0139] “表位”或“抗原决定簇”是指与结合剂结合的抗原的区域。

[0140] 如本文中所示使用，“分离的聚核苷酸”是指已经从在天然存在的状态下与其侧接的序列纯化的聚核苷酸，例如从通常与片段相邻的序列移出的DNA片段。“分离的聚核苷酸”还指互补DNA(cDNA)、重组型DNA或在自然界中不存在并且通过人力制得的其它聚核苷酸。

[0141] 如本文中所示使用，“分离的蛋白质”、“分离的肽”或“分离的多肽”等是指体外合成、

从细胞环境以及从与细胞中其它组分的缔合分离和/或纯化的肽或多肽分子,即,其不与体内物质明显缔合。

[0142] “提高”或“促进”或“增加”或“扩增”或“增强”通常是指本文中所涵盖的组合物产生、引发或引起比由媒剂或对照物所引起的反应更大的反应(即,生理反应)的能力。从所属领域中的理解和本文中的说明显而易见,可测量的反应可以包括催化活性、结合亲和力、结合位点特异性、结合位点选择性、持久性、细胞溶解活性增加和/或促炎性细胞因子增加等。“增加”或“增强”的量典型地是“统计显著”的量,并且可以包括使媒剂或对照物所产生的反应增加1.1、1.2、1.5、2.3、4、5、6、7、8、9、10、15、20、30倍或更多倍(例如500、1000倍)(包括在其间并且高于1的所有整数和小数点,例如1.5、1.6、1.7、1.8等)。

[0143] “减少”或“减小”或“减低”或“降低”或“减轻”或“去除”或“抑制”或“减弱”通常是指本文中所涵盖的组合物产生、引发或引起比由媒剂或对照物所引起的反应更小的反应(即,生理反应)的能力。可测量的反应可以包括脱靶结合亲和力、脱靶裂解特异性、T细胞耗竭等的减小。“减少”或“降低”的量典型地是“统计显著”的量,并且可以包括使媒剂或对照物所产生的反应(参考反应)减小1.1、1.2、1.5、2.3、4、5、6、7、8、9、10、15、20、30倍或更多倍(例如500、1000倍)(包括在其间并且高于1的所有整数和小数点,例如1.5、1.6、1.7、1.8)。

[0144] “维持(maintain)”或“保持(preserve)”或“维持(maintenance)”或“无变化”或“无显著变化”或“无显著减少”通常是指本文中所涵盖的组合物产生、引发或引起与由媒剂或对照物所引起的反应基本上类似或相当的生理反应(即,下游作用)的能力。相当的反应是与参考反应无显著差异或可测量的差异的反应。

[0145] 如本文中所使用,术语“特异性结合亲和力”或“特异性结合(specifically binds)”或“特异性结合(specifically bound)”或“特异性结合(specific binding)”或“特异性靶向”描述一个分子与另一个分子结合,例如与背景结合相比,多肽的DNA结合结构域以更大的结合亲和力结合于DNA。如果结合结构域以例如大于或等于约 10^5M^{-1} 的亲和力或 K_a (即,具体结合相互作用的平衡缔合常数,其以 $1/\text{M}$ 为单位)与目标位点结合或缔合,则所述结合结构域“特异性结合”于目标位点。在某些实施例中,结合结构域以大于或等于约 10^6M^{-1} 、 10^7M^{-1} 、 10^8M^{-1} 、 10^9M^{-1} 、 10^{10}M^{-1} 、 10^{11}M^{-1} 、 10^{12}M^{-1} 或 10^{13}M^{-1} 的 K_a 结合于目标位点。“高亲和力”结合结构域是指具有至少 10^7M^{-1} 、至少 10^8M^{-1} 、至少 10^9M^{-1} 、至少 10^{10}M^{-1} 、至少 10^{11}M^{-1} 、至少 10^{12}M^{-1} 、至少 10^{13}M^{-1} 或更大的 K_a 的结合结构域。

[0146] 或者,亲和力可以定义为具体结合相互作用的平衡解离常数(K_d),其以 M 为单位(例如 10^{-5}M 到 10^{-13}M ,或更小)。具体实施例中涵盖的包含一个或多个DNA结合结构域的核酸酶变异体针对DNA目标位点的亲和力可以使用常规技术,例如酵母细胞表面呈现,或通过使用被标记的配位体的结合缔合或置换分析法容易地测定。

[0147] 在一个实施例中,特异性结合的亲和力比背景结合大约2倍、比背景结合大约5倍、比背景结合大约10倍、比背景结合大约20倍、比背景结合大约50倍、比背景结合大约100倍、或比背景结合大约1000倍或更多。

[0148] 术语“选择性结合(selectively binds)”或“选择性结合(selectively bound)”或“选择性结合(selectively binding)”或“选择性靶向”描述在多个脱靶分子存在下,一个分子优先结合于目标分子(在靶结合)。在具体实施例中,HE或megaTAL选择性结合在靶DNA结合位点的频率比HE或megaTAL结合脱靶DNA目标结合位点的频率高约5、10、15、20、25、

50、100或1000倍。

[0149] “在靶”是指目标位点序列。

[0150] “脱靶”是指与目标位点序列类似但不一致的序列。

[0151] “目标位点”或“目标序列”是在存在足以实现结合和/或裂解的条件下,界定核酸中结合分子将结合和/或裂解的部分的染色体或染色体外核酸序列。当提到仅关于目标位点或目标序列的一条链的聚核苷酸序列或SEQ ID NO.时,应理解,核酸酶变体所结合和/或裂解的目标位点或目标序列是双链的并且包含参考序列和其补体。

[0152] “重组”是指两个聚核苷酸之间遗传信息交换的过程,包括(但不限于)通过非同源末端接合(NHEJ)和同源重组进行的供体捕获。出于本公开的目的,“同源重组(HR)”是指在例如通过同源性定向修复(HDR)机制修复双链断裂期间发生的这类交换的特殊形式。这一程序需要核苷酸序列同源性,使用“供体”分子作为模板来修复“目标”分子(即,经历双链断裂的分子),并且不同地称为“不产生交换的基因转换(non-crossover gene conversion)”或“短链基因转化(short tract gene conversion)”,因为其引起遗传信息从供体转移到目标。不希望受任何具体理论约束,这类转移可以涉及在断裂的目标与供体之间形成的异双螺旋DNA的错配校正,和/或“合成依赖性链粘接(synthesis-dependent strand annealing)”,其中供体被用于再合成遗传信息,所述遗传信息将变成目标的一部分;和/或相关过程。这类特殊HR通常引起目标分子序列的改变,使得供体聚核苷酸序列的一部分或全部被并入目标聚核苷酸中。

[0153] “NHEJ”或“非同源末端接合”是指在无供体修复模板或同源序列存在下双链断裂的消除。NHEJ可以在断裂位点处引起插入和缺失。NHEJ是由若干子路径介导的,这些子路径各自具有不同的突变结果。经典的NHEJ路径(cNHEJ)需要KU/DNA-PKcs/Lig4/XRCC4复合物,通过最低限度的处理,将各末端重新接合在一起,并且通常引起断裂的精确修复。替代性NHEJ路径(altNHEJ)在消除dsDNA断裂方面也具有活性,但这些路径明显更易于诱变并且通常无法精确修复以插入和缺失为标志的断裂。尽管不希望受任何具体理论约束,但预期通过末端处理酶(例如外切核酸酶,例如Trex2)修饰dsDNA断裂可以使修复向altNHEJ路径偏移。

[0154] “裂解”是指DNA分子的共价主链的断裂。裂解可以通过多种方法引发,包括(但不限于)磷酸二酯键的酶促或化学水解。单链裂解和双链裂解都是可能的。双链裂解可以作为两个不同的单链裂解事件的结果发生。DNA裂解可以引起产生钝端或交错端。在某些实施例中,本文中所涵盖的多肽和核酸酶变体,例如导向核酸内切酶变体、megaTAL等被用于靶向双链DNA裂解。核酸内切酶裂解识别位点可以在任一个DNA链上。

[0155] “外源”分子是通常不存在于细胞中,而是通过一种或多种遗传、生物化学或其它方法引入细胞中的分子。例示性外源分子包括(但不限于)小型有机分子、蛋白质、核酸、碳水化合物、脂质、糖蛋白、脂蛋白、多糖、以上分子的任何被修饰的衍生物或包含一个或多个以上分子的任何复合物。用于将外源分子引入细胞中的方法是所属领域的技术人员已知的并且包括(但不限于)脂质介导的转移(即,脂质体,包括中性和阳离子性脂质)、电致孔、直接注射、细胞融合、粒子轰击、生物聚合物纳米粒子、磷酸钙共沉淀、DEAE-葡聚糖介导的转移以及病毒载体介导的转移。

[0156] “内源”分子是在具体环境条件下通常存在于处于具体发育阶段的具体细胞中的

分子。其它内源分子可以包括蛋白质。

[0157] “基因”是指编码基因产物的DNA区域,以及调节基因产物的产生的所有DNA区域,无论这类调节序列是否邻近于编码和/或转录序列。基因包括(但不限于)启动子序列、增强子、沉默子、隔离子、边界元件、终止子、聚腺苷酸化序列、转录后反应元件、翻译调节序列,如核糖体结合位点和内部核糖体进入位点、复制起点、基质连接位点以及基因座控制区。

[0158] “基因表达”是指基因中所含信息转化成基因产物。基因产物可以是基因(例如 mRNA、tRNA、rRNA、反义RNA、核糖核酸酶、结构RNA或任何其它类型的RNA)的直接转录产物或由mRNA翻译产生的蛋白质。基因产物还包括通过如封端、聚腺苷酸化、甲基化和编辑等方法修饰的RNA,以及通过例如甲基化、乙酰化、磷酸化、泛素化、ADP-核糖基化、肉豆蔻酸化和糖基化修饰的蛋白质。

[0159] 如本文中所使用,术语“基因工程改造”或“基因修饰”是指将呈DNA或RNA形式的额外遗传物质染色体或染色体外添加到细胞中的总遗传物质中。基因修饰可以靶向或不靶向细胞基因组中的具体位点。在一个实施例中,基因修饰具有位点特异性。在一个实施例中,基因修饰不具有位点特异性。

[0160] 如本文中所使用,术语“基因组编辑”是指在细胞基因组中的目标位点处遗传物质的取代、缺失和/或引入,其恢复、校正、破坏和/或改变基因或基因产物的表达。具体实施例中涵盖的基因组编辑包含在细胞基因组中的目标位点处或附近将一个或多个核酸酶变异体引入细胞中以产生DNA病灶,任选地在供体修复模板存在下。

[0161] 如本文中所使用,术语“基因疗法”是指将额外遗传物质引入细胞中的总遗传物质中以恢复、校正或改变基因或基因产物的表达或达到表达治疗性多肽的目的。在具体实施例中,通过基因组编辑将遗传物质引入细胞基因组中以恢复、校正、破坏或改变基因或基因产物的表达或达到表达治疗性多肽的目的被认为是基因疗法。

[0162] 如本文中所使用,术语“受试者”和“个体”通常可以互换使用并且是指展现可以用本文中其它地方涵盖的核酸酶变异体、基因组编辑组合物、基因疗法载体、基因组编辑载体、基因组编辑的细胞以及方法治疗的免疫病症的症状的任何动物。适合的个体(例如患者)包括实验动物(如小鼠、大鼠、兔或天竺鼠)、农畜和家畜或宠物(如猫或狗)。包括非人类灵长类动物并且优选地,人类个体。典型个体包括患有免疫病症、已诊断患有免疫病症或具有罹患免疫病症的风险的人类患者。

[0163] 如本文中所使用,术语“患者”是指已诊断患有可以用本文中其它地方所涵盖的核酸酶变异体、基因组编辑组合物、基因疗法载体、基因组编辑载体、基因组编辑的细胞以及方法治疗的免疫病症的个体。

[0164] 如本文中所使用,“自体”是指其中供体和受体是同一名个体之细胞。

[0165] 如本文中所使用,“同种异体”是指其中供体和受体物种相同,但细胞在遗传学上不同的细胞。

[0166] 如本文中所使用,“同基因”是指其中供体和受体物种相同、供体和受体是不同个体并且供体细胞和受体细胞在遗传学上一致的细胞。

[0167] 如本文中所使用,“异基因”是指其中供体和受体物种不相同的细胞。

[0168] 如本文中所使用,“治疗(treatment/treating)”包括针对疾病或病理学病状的症状或病理的任何有益或所需的作用,并且可以包括所治疗的疾病或病状,例如血红蛋白病、

癌症、GVHD、感染性疾病、自身免疫疾病、发炎性疾病和免疫缺陷的一种或多种可测量的标记物的甚至最小降低。治疗可以任选地涉及疾病或病状的进程的延缓。“治疗”并非必须指示完全根除或治愈疾病或病状或其相关症状。

[0169] 如本文中所使用,“预防(prevent)”和类似词语,如“预防(prevention/prevented/preventing)”等指示用于预防、抑制疾病或病状,例如血红蛋白病、癌症、GVHD、感染性疾病、自身免疫疾病、发炎性疾病和免疫缺陷的发生或复发,或降低其发生或复发的可能性的方法。其也指延缓疾病或病状的发作或复发或延缓疾病或病状的症状出现或复发。如本文中所使用,“预防”和类似词语还包括在疾病或病状发作或复发之前降低疾病或病状的强度、影响、症状和/或负荷。

[0170] 如本文中所使用,短语“改善……的至少一种症状”是指减少所治疗的个体的疾病或病状(例如血红蛋白病、癌症、GVHD、感染性疾病、自身免疫疾病、发炎性疾病和免疫缺陷)的一种或多种症状。在具体实施例中,所治疗的血红蛋白病或血红蛋白病病状是 β -地中海型贫血,其中所改善的一种或多种症状包括(但不限于)虚弱、疲乏、外观苍白、黄疸、面部骨骼畸形、生长缓慢、腹部肿胀、深色尿、缺铁(在不存在输液的情况下)、需要频繁输液。在具体实施例中,所治疗的血红蛋白病或血红蛋白病病状是镰状细胞病(SCD),其中所改善的一种或多种症状包括(但不限于)贫血;不可解释的疼痛发作,如腹部、胸腔、骨骼或关节疼痛;手部或脚部肿胀;腹部肿胀;发热;频繁感染;皮肤或甲床苍白;黄疸;生长延缓;视力问题;中风病征或症状;缺铁(在不存在输液的情况下)、需要频繁输液。在具体实施例中,所治疗的疾病或病状是癌症,其中所改善的一种或多种症状包括(但不限于)虚弱、疲乏、呼吸急促、容易瘀伤和出血、频繁感染、淋巴结肿大、腹胀或腹痛(由于腹部器官增大)、骨骼或关节疼痛、骨折、非计划性体重减轻、食欲不振、盗汗、持续性轻度发热和排尿减少(由于肾功能损伤)。

[0171] 如本文中所使用,术语“量”是指足以实现有益或所需的预防性或治疗性结果(包括临床结果)的DNA供体修复模板、核酸酶变异体、基因组编辑组合物或基因组编辑的细胞的“有效的量”或“有效量”。

[0172] “预防有效量”是指足以实现所需预防结果的DNA供体修复模板、核酸酶变异体、基因组编辑组合物或基因组编辑的细胞的量。典型地但未必,因为预防性剂量是在疾病之前或在疾病的较早阶段用于个体,所以预防有效量小于治疗有效量。

[0173] DNA供体修复模板、核酸酶变异体、基因组编辑组合物或基因组编辑的细胞的“治疗有效量”可以根据如疾病状态、个体的年龄、性别和体重以及在个体中引起所需反应的能力等因素而变化。治疗有效量还是治疗有益作用超过任何毒性或有害作用的量。术语“治疗有效量”包括可有效地“治疗”个体(例如患者)的量。当指示治疗量时,待给予的具体实施例中所涵盖的组合物确切量可以由医师根据本说明书并且考虑个体在年龄、体重、肿瘤尺寸、感染或转移程度以及患者(个体)的状况来确定。

[0174] 如本文中所使用,术语“血红蛋白病”或“血红蛋白病病状”是指一组多种遗传性血液病症,其涉及由血色素的结构和/或合成的改变而引起的异常血色素分子的存在。

[0175] “免疫病症”是指引起免疫系统反应的疾病。在具体实施例中,术语“免疫病症”是指癌症、移植物抗宿主疾病、自身免疫疾病或免疫缺陷。在一个实施例中,免疫病症涵盖感染性疾病。

[0176] 如本文中所使用,术语“癌症”通常涉及其中异常细胞不受控制地分裂且可侵袭邻近组织的一类疾病或病状。

[0177] 如本文中所使用,术语“恶性”是指一种癌症,其中肿瘤细胞群组呈现以下中的一种或多种:不受控生长(即,超过正常限度的分裂)、侵袭(即,浸入和破坏相邻组织)和转移(即,通过淋巴或血液扩散到身体中的其它位置)。

[0178] 如本文中所使用,术语“转移”是指癌症从身体的一个部分扩散到另一部分。由扩散的细胞形成的肿瘤称为“转移瘤”或“转移”。转移瘤含有与原始(原发性)肿瘤中的细胞类似的细胞。

[0179] 如本文中所使用,术语“良性”或“非恶性”是指可以长大但不会扩散到身体的其它部分的肿瘤。良性肿瘤是自限性的并且通常不会侵袭或转移。

[0180] “癌细胞”或“肿瘤细胞”是指癌变生长或组织的单独的细胞。肿瘤通常是指由细胞异常生长形成的肿胀或病灶,其可以是良性、癌前或恶性的。大部分癌症形成肿瘤,但有一些,例如白血病,未必形成肿瘤。对于形成肿瘤的癌症,术语癌症(细胞)和肿瘤(细胞)可以互换地使用。个体中肿瘤的量是“肿瘤负荷”,其可以按肿瘤的数目、体积或重量形式测量。

[0181] “移植物抗宿主疾病”或“GVHD”是指在细胞、组织或实体器官移植之后发生的并发症。GVHD可以在干细胞或骨髓移植之后发生,其中移植的供体细胞攻击移植物接受者的身体。人类中的急性GVHD在移植后约60天内发生并且在细胞溶解性淋巴细胞的作用下引起皮肤、肝和肠道的损伤。慢性GVHD是在稍后发生并且是主要影响皮肤,引起B细胞多克隆活化以及Ig和自身抗体过量产生的一种全身性自身免疫疾病。实体器官移植的移植物抗宿主疾病(SOT-GVHD)是以两种形式发生。较常见类型是抗体介导的,其中在具有A型、B型或AB型血液的接受者中,来自具有O型血液的供体的抗体攻击接受者的红血球,引起轻度短暂性溶血性贫血。SOT-GVHD的第二种形式是与高死亡率相关的一种细胞类型,其中供体衍生的T细胞针对免疫不同的宿主组织,最通常在皮肤、肝、胃肠道和骨髓中产生免疫攻击,在这些器官中引起并发症。

[0182] “移植物抗白血病”或“GVL”是指供体的移植组织(如骨髓或外周血液)中的免疫细胞针对个人的白血病细胞的免疫反应。

[0183] “自身免疫疾病”是指身体针对其自身组织的一些成分产生免疫原性(即,免疫系统)反应的疾病。换句话说,免疫系统丧失其将身体内的某一组织或系统识别为“自身的”能力,并且靶向且攻击所述组织或系统,就好象它是外来的。自身免疫性疾病的说明性实例包括(但不限于):关节炎、发炎性肠病、桥本氏甲状腺炎(Hashimoto's thyroiditis)、格雷夫斯氏病(Grave's disease)、狼疮、多发性硬化、风湿性关节炎、溶血性贫血、抗免疫性甲状腺炎、全身性红斑狼疮、乳糜泻、克罗恩氏病(Crohn's disease)、结肠炎、糖尿病、硬皮病、牛皮癣等。

[0184] “免疫缺陷”意指患者免疫系统因疾病或因施用化学物质而受损的状态。这种病状使系统缺乏防御外来物质所需的血细胞数目和类型。免疫缺陷病状或疾病是所属领域中已知的并且包括例如获得性免疫缺陷综合症(AIDS)、重度联合免疫缺陷疾病(SCID)、选择性IgA缺乏症、常见变异型免疫缺陷病、X连锁无丙种球蛋白血症、慢性肉芽肿病、高IgM综合症、韦斯考特-奥德里奇综合症(Wiskott-Aldrich Syndrome, WAS)以及糖尿病。

[0185] “感染性疾病”是指可以在人与人之间或在生物体与生物体之间传播的一种疾病,

并且是由微生物或病毒试剂引起(例如普通感冒)。感染性疾病是所属领域中已知的并且包括例如肝炎、性传播疾病(例如衣原体病、淋病)、结核病、HIV/AIDS、白喉、乙型肝炎、丙型肝炎、霍乱以及流感。

[0186] C. 供体修复模板

[0187] 本文中所涵盖的供体修复模板适用于使用单一供体分子和一种或多种被工程改造的核酸酶的多重基因组编辑。本文中所涵盖的供体修复模板与现有方法相比提供显著改良,所述现有方法是繁琐的并且通常需要多轮编辑和单独的供体模板以及用于每个待编辑的基因座的相应的被工程改造的核酸酶。在具体实施例中,使用多个被工程改造的核酸酶在多个目标位点中引入DSB,并且在靶向多个目标位点的单一DNA供体修复模板存在下,通过同源性定向修复(HDR)机制修复每个目标位点处的DSB。

[0188] 在具体实施例中,使用DNA供体修复模板将一个或多个序列插入基因组中的多个位点中。在具体优选实施例中,使用供体修复模板修复或修饰基因组中的多个目标位点序列。

[0189] 在各种实施例中,通过用包含供体修复模板的腺相关病毒(AAV)、逆转录病毒(例如慢病毒、IDLV等)、单纯疱疹病毒、腺病毒或牛痘病毒载体转导细胞,将供体修复模板引入造血细胞(例如HSC、CD34⁺细胞、T细胞等)中。

[0190] 在某些实施例中,通过将包含供体修复模板的单链DNA、双链DNA、质粒或微型环DNA引入细胞中,将供体修复模板引入造血细胞(例如HSC、CD34⁺细胞、T细胞等)中。

[0191] 在各种实施例中,DNA供体修复模板包含对应于基因组中的多个目标位点的多个同源臂。

[0192] 在各种实施例中,DNA供体修复模板包含对应于基因组中的多个目标位点的多对同源臂,其中每对同源臂侧接一个或多个转基因。

[0193] 在一些实施例中,DNA供体修复模板包含对应于基因组中的多个目标位点的多对同源臂,其中每对同源臂侧接不同转基因。

[0194] 在各种实施例中,DNA供体修复模板包含对应于基因组中的多个目标位点的多对同源臂,其中每对同源臂侧接一个或多个转基因,其中至少两对同源臂侧接相同转基因。

[0195] “同源臂对”是指一组两个同源臂。在具体实施例中,一对同源臂包含5'同源臂和3'同源臂。“5'同源臂”是指与目标位点(例如双链断裂位点)的DNA序列5'一致或几乎一致或同源的聚核苷酸序列。“3'同源臂”是指与目标位点的DNA序列3'一致或几乎一致或同源的聚核苷酸序列。在具体实施例中,一对同源臂包含一个包含聚核苷酸序列的同源臂,所述聚核苷酸序列包括双链断裂的目标位点,其中目标位点中具有突变以最小化目标位点的再裂解。在具体实施例中,一对同源臂中的任一个或两个同源臂独立地位于距目标位点约100bp、约200bp、约300bp、约400bp、约500bp、约600bp、约700bp、约800bp、约900bp、约1000bp、约1100bp、约1200bp、约1300bp、约1400bp、约1500bp、约1600bp、约1700bp、约1800bp、约1900bp、约2000bp、约2100bp、约2200bp、约2300bp、约2400bp、约2500bp、约2600bp、约2700bp、约2800bp、约2900bp或约3000bp处,包括距目标位点的所有中间距离。

[0196] 具体实施例中所涵盖的合适长度的同源臂的说明性实例可以独立地选择并且包括(但不限于):约100bp、约200bp、约300bp、约400bp、约500bp、约600bp、约700bp、约800bp、约900bp、约1000bp、约1100bp、约1200bp、约1300bp、约1400bp、约1500bp、约1600bp、约

1700bp、约1800bp、约1900bp、约2000bp、约2100bp、约2200bp、约2300bp、约2400bp、约2500bp、约2600bp、约2700bp、约2800bp、约2900bp或约3000bp,或更长的同源臂,包括所有中间长度的同源臂。

[0197] 合适的同源臂长度的其它说明性实例包括(但不限于):约100bp到约600bp、约100bp到约500bp、约100bp到约400bp、约100bp到约300bp、约100bp到约200bp、约200bp到约600bp、约200bp到约500bp、约200bp到约400bp、约200bp到约300bp、约300bp到约600bp、约300bp到约500bp、约100bp到约3000bp、约200bp到约3000bp、约300bp到约3000bp、约400bp到约3000bp、约500bp到约3000bp、约500bp到约2500bp、约500bp到约2000bp、约750bp到约2000bp、约750bp到约1500bp或约1000bp到约1500bp,包括所有中间长度的同源臂。

[0198] 在具体实施例中,DNA供体修复模板中的任何5'和3'同源臂的长度是独立地选自约100bp、约200bp、约300bp、约400bp、约500bp或约600bp。在一个实施例中,5'同源臂是约300bp并且3'同源臂是约300bp。在一个实施例中,DNA供体修复模板中的多个5'同源臂中的每一个是约300bp并且DNA供体修复模板中的多个3'同源臂中的每一个是约300bp。

[0199] 在具体实施例中,DNA供体修复模板包含多对同源臂和一个或多个转基因。

[0200] 在具体实施例中,DNA供体修复模板包含多对同源臂和转基因。在某些实施例中,每一对同源臂包含与目标位点的DNA序列5'同源的5'同源臂和与目标位点的DNA序列3'同源的3'同源臂,其中每一对同源臂与不同目标位点相关。在具体实施例中,DNA供体修复模板包含第一对同源臂,其包含与第一目标位点的DNA序列5'同源的5'同源臂和与第一目标位点的DNA序列3'同源的3'同源臂;以及与第二目标位点的DNA序列5'同源的5'同源臂和与第二目标位点的DNA序列3'同源的3'同源臂,其中第一目标位点与第二目标位点不相同。

[0201] 在具体实施例中,每一对同源臂侧接相同转基因序列,每个5'同源臂位于转基因的5'并且每个3'同源臂位于转基因的3'。在具体实施例中,DNA供体修复模板包含5'同源臂,其与第一目标位点的DNA序列5'同源并且位于与第二目标位点的DNA序列5'同源的5'同源臂的5',两个5'同源臂都位于转基因的5';和3'同源臂,其与第一目标位点的DNA序列3'同源并且位于与第二目标位点的DNA序列3'同源的3'同源臂的3',两个3'同源臂都位于转基因的3'。

[0202] 在具体实施例中,每一对同源臂侧接相同转基因序列,每个5'同源臂位于转基因的5'并且每个3'同源臂位于转基因的3'。在具体实施例中,DNA供体修复模板包含5'同源臂,其与第一目标位点的DNA序列5'同源并且位于与第二目标位点的DNA序列5'同源的5'同源臂的5',两个5'同源臂都位于转基因的5';和3'同源臂,其与第一目标位点的DNA序列3'同源并且位于与第二目标位点到DNA序列3'同源的3'同源臂的5',两个3'同源臂都位于转基因的3'。

[0203] 在具体实施例中,每一对同源臂侧接不同转基因序列,第一对同源臂侧接第一转基因并且第二对同源臂侧接第二转基因。在具体实施例中,DNA供体修复模板包含第一对同源臂,其包含与第一目标位点的DNA序列5'同源并且位于第一转基因的5'的5'同源臂,和与第一目标位点的DNA序列3'同源并且位于第一转基因的3'的3'同源臂;和第二对同源臂,其包含与第二目标位点的DNA序列5'同源并且位于第二转基因的5'的5'同源臂,和与第二目标位点的DNA序列3'同源并且位于第二转基因的3'的3'同源臂。

[0204] 在优选实施例中,DNA供体修复模板编码一个或多个转基因,所述转基因提高基于

细胞的基因疗法在治疗或预防血红蛋白病或改善血红蛋白病的至少一种症状方面的功效。在具体实施例中，DNA供体修复模板包含一个或多个编码球蛋白多肽的转基因，包括(但不限于) β -球蛋白多肽、抗镰状化 β -球蛋白多肽(例如 $\beta^{\text{A-T87Q}}$ 、 $\beta^{\text{A-K120E}}$ 、 $\beta^{\text{A-K95E}}$ 或其组合)或 γ -球蛋白多肽。

[0205] 在优选实施例中，DNA供体修复模板编码一个或多个转基因，所述转基因提高过继性细胞疗法的安全性、功效、耐久性和/或持久性。在具体实施例中，DNA供体修复模板包含一个或多个编码免疫效能增强子、免疫抑制信号减弱子或被工程改造的抗原受体的转基因。在具体实施例中，DNA供体模板被设计成将一个或多个转基因插入(通过HDR)多个编码基因的基因座，所述转基因有助于TCR信号传导、免疫系统检查点和/或免疫抑制信号传导。

[0206] 供体修复模板可以进一步包含本文中其它地方所涵盖的一种或多种聚核苷酸，如启动子和/或增强子、非翻译区(UTR)、Kozak序列、聚腺苷酸化信号、其它限制酶位点、多克隆位点、内部核糖体进入位点(IRES)、重组酶识别位点(例如LoxP、FRT和Att位点)、终止密码子、转录终止信号以及编码自裂解多肽的聚核苷酸、表位标签。

[0207] 1. 免疫效能增强子

[0208] 在具体实施例中，本文中所涵盖的被基因组编辑的免疫效应细胞通过在编码免疫效能增强子的供体修复模板存在下，将DSB引入一个或多个编码TCR信号传导组件(例如TCR α 基因的恒定区)、免疫系统检查点或免疫抑制信号传导组件的基因中而变得更强效。如本文中所使用，术语“免疫效能增强子”是指刺激和/或增强T细胞活化和/或功能的非天然存在的分子、免疫增强因子，以及在T细胞中将来自肿瘤微环境的免疫抑制信号转化成免疫刺激信号的非天然存在的多肽。

[0209] 在具体实施例中，DNA供体模板被设计成将免疫效能增强子插入一个基因座中并且将一个或多个编码另一免疫效能增强子、免疫抑制信号减弱子或被工程改造的抗原受体的转基因插入一个或多个不同基因座中。

[0210] 在具体实施例中，免疫效能增强子是选自由以下组成的群组：双特异性T细胞接合子(BiTE)分子；免疫增强因子，包括(但不限于)细胞因子、趋化因子、细胞毒素和/或细胞因子受体；以及翻转受体。

[0211] 在一些实施例中，免疫效能增强子、免疫增强因子或翻转受体是包含蛋白质不稳定结构域的融合多肽。

[0212] a. 双特异性T细胞接合子(BiTE)分子

[0213] 在具体实施例中，本文中所涵盖的被基因组编辑的免疫效应细胞通过在编码双特异性T细胞接合子(BiTE)分子的供体修复模板存在下，将DSB引入一个或多个编码TCR信号传导组件(例如TCR α 基因的恒定区)、免疫系统检查点或免疫抑制信号传导组件的基因中而变得更强效。BiTE分子是包含结合目标抗原的第一结合结构域、本文中其它地方所涵盖的连接子或间隔子以及结合免疫效应细胞上的刺激或共刺激分子的第二结合结构域的由两部分组成的分子。第一结合结构域和第二结合结构域可以独立地选自配位体、受体、抗体或其抗原结合片段、凝集素以及碳水化合物。

[0214] 在具体实施例中，第一结合结构域和第二结合结构域是抗原结合结构域。

[0215] 在具体实施例中，第一结合结构域和第二结合结构域是抗体或其抗原结合片段。在一个实施例中，第一结合结构域和第二结合结构域是单链可变片段(scFv)。

[0216] 在具体实施例中,可以被第一结合结构域识别和结合的目标抗原的说明性实例包括(但不限于): α 叶酸受体、滋养层糖蛋白(TPBG); α v β 6整合素;B细胞成熟抗原(BCMA);B7-H3;B7-H6;CD276;CD16;CD19;CD20;CD22;CD30;CD33;CD44;CD44v6;CD44v7/8;CD70;CD79a;CD79b;CD123;CD138;CD171;癌胚抗原(CEA)多肽;硫酸软骨素蛋白聚糖4(CSPG4);表皮生长因子受体(EGFR);erb-b2受体酪氨酸激酶2(ERBB2);EGFR基因扩增和变异体III(EGFRvIII);上皮细胞粘附分子(EPCAM);艾普瑞林A2(ephA2);成纤维细胞活化蛋白质 α (FAP);胚胎乙酰胆碱受体(胚胎AChR);bDGalpNAc(1-4)[aNeu5Ac(2-8)aNeu5Ac(2-3)]bDGalp(1-4)bDG1cp(1-1)Cer(GD2);aNeu5Ac(2-8)aNeu5Ac(2-3)bDGalp(1-4)bDG1cp(1-1)Cer(GD3);磷脂酰肌醇聚糖-3(GPC3);人类白细胞抗原A1 MAGE家族成员A1(HLA-A1⁺MAGEA1);人类白细胞抗原A2 MAGE家族成员A1(HLA-A2⁺MAGEA1);人类白细胞抗原A3 MAGE家族成员A1(HLA-A3⁺MAGEA1);MAGEA1;人类白细胞抗原A1纽约食管鳞状细胞癌1(HLA-A1⁺NY-ESO-1);人类白细胞抗原A2纽约食管鳞状细胞癌1(HLA-A2⁺NY-ESO-1);人类白细胞抗原A3纽约食管鳞状细胞癌1(HLA-A3⁺NY-ESO-1);白细胞介素受体11 α (IL-11R α);白细胞介素-13受体子单元 α -2(IL-13R α 2); λ 轻链;路易斯-Y(lewis-Y); κ 轻链;间皮素;细胞表面相关粘蛋白1(MUC1);细胞表面相关粘蛋白16(MUC16)、神经细胞粘附分子(NCAM)、自然杀手群组2D(NKG2D)配位体、NY-ESO-1、黑素瘤中的优先表达抗原(PRAME)、前列腺干细胞抗原(PSCA)、前列腺特异性膜抗原(PSMA)、受体酪氨酸激酶样孤儿受体1(ROR1)、滑膜肉瘤X(SSX)、存活素、肿瘤相关糖蛋白72(TAG72)、肿瘤内皮标记物1(TEM1)、肿瘤内皮标记物5(TEM5)、肿瘤内皮标记物7(TEM7)、肿瘤内皮标记物8(TEM8)、血管内皮生长因子受体2(VEGFR2)和威尔姆氏肿瘤1(Wilm' tumor 1;WT-1)。

[0217] 目标抗原的其它说明性实施例包括MHC-肽复合物,任选地其中肽由以下处理:FR α 、TPBG、 α v β 6整合素、BCMA、B7-H3、B7-H6、CD276、CD16、CD19、CD20、CD22、CD30、CD33、CD44、CD44v6、CD44v7/8、CD70、CD79a、CD79b、CD123、CD138、CD171、CEA、CSPG4、EGFR、ERBB2、EGFRvIII、EPCAM、EphA2、FAP、胚胎AChR、GD2、GD3、GPC3、HLA-A1⁺MAGEA1、HLA-A2⁺MAGEA1、HLA-A3⁺MAGEA1、MAGEA1、HLA-A1⁺NY-ESO-1、HLA-A2⁺NY-ESO-1、HLA-A3⁺NY-ESO-1、IL-11R α 、IL-13R α 2、 λ 轻链、路易斯-Y、 κ 轻链、间皮素、MUC1、MUC16、NCAM、NKG2D配位体、NY-ESO-1、PRAME、PSCA、PSMA、ROR1、SSX、存活素、TAG72、TEM1、TEM5、TEM7、TEM8、VEGFR2和WT-1。

[0218] 在具体实施例中,免疫效应细胞上被第二结合结构域识别和结合的刺激或共刺激分子的说明性实例包括(但不限于):CD3 γ 、CD3 δ 、CD3 ϵ 、CD3 ζ 、CD28、CD40、CD80、CD86、CD134、CD137以及CD278。

[0219] 在具体实施例中,使用多个被工程改造的核酸酶诱导基因组中的多个目标位点中的DSB,并且将包含转基因的DNA供体修复模板引入细胞且通过同源重组插入至少一个目标位点中,所述转基因编码BiTE并且由对应于基因组中的多个目标位点的多个同源臂侧接。

[0220] 在具体实施例中,使用多个被工程改造的核酸酶诱导基因组中的多个目标位点中的DSB,并且将包含转基因的DNA供体修复模板引入细胞且通过同源重组插入两个或更多个目标位点中,所述转基因编码BiTE并且由对应于基因组中的多个目标位点的多个同源臂侧接。

[0221] 在具体实施例中,使用多个被工程改造的核酸酶诱导基因组中的多个目标位点中的DSB,并且将包含转基因的DNA供体修复模板引入细胞且通过同源重组插入每个目标位点

中,所述转基因编码BiTE并且由对应于基因组中的多个目标位点的多个同源臂侧接。

[0222] b. 免疫增强因子

[0223] 在具体实施例中,本文中所涵盖的被基因组编辑的免疫效应细胞通过在编码免疫增强因子的供体修复模板存在下,将DSB引入一个或多个编码TCR信号传导组件(例如TCR α 基因的恒定区)、免疫系统检查点或免疫抑制信号传导组件的基因中而变得更强效。免疫增强因子是指增强免疫效应细胞中的免疫反应的具体细胞因子、趋化因子、细胞毒素和细胞因子受体。

[0224] 在具体实施例中,供体修复模板编码选自以下组成的群组的细胞因子:IL-2、胰岛素、IFN- γ 、IL-7、IL-21、IL-10、IL-12、IL-15以及TNF- α 。

[0225] 在具体实施例中,供体修复模板编码选自以下组成的群组的趋化因子:MIP-1 α 、MIP-1 β 、MCP-1、MCP-3以及RANTES。

[0226] 在具体实施例中,供体修复模板编码选自以下组成的群组的细胞毒素:穿孔素(Perforin)、颗粒酶A和颗粒酶B。

[0227] 在具体实施例中,供体修复模板编码选自以下组成的群组的细胞因子受体:IL-2受体、IL-7受体、IL-12受体、IL-15受体以及IL-21受体。

[0228] 在具体实施例中,使用多个被工程改造的核酸酶诱导基因组中的多个目标位点中的DSB,并且将包含转基因的DNA供体修复模板引入细胞且通过同源重组插入至少一个目标位点中,所述转基因编码免疫增强因子并且由对应于基因组中的多个目标位点的多个同源臂侧接。

[0229] 在具体实施例中,使用多个被工程改造的核酸酶诱导基因组中的多个目标位点中的DSB,并且将包含转基因的DNA供体修复模板引入细胞且通过同源重组插入两个或更多个目标位点中,所述转基因编码免疫增强因子并且由对应于基因组中的多个目标位点的多个同源臂侧接。

[0230] 在具体实施例中,使用多个被工程改造的核酸酶诱导基因组中的多个目标位点中的DSB,并且将包含转基因的DNA供体修复模板引入细胞且通过同源重组插入每个目标位点中,所述转基因编码免疫增强因子并且由对应于基因组中的多个目标位点的多个同源臂侧接。

[0231] c. 翻转受体

[0232] 在具体实施例中,本文中所涵盖的被基因组编辑的免疫效应细胞通过“翻转”或“逆转”免疫抑制因子的免疫抑制信号而变得对耗竭具有更高的抗性,所述免疫抑制信号是由肿瘤微环境针对阳性免疫刺激信号引发。在一个实施例中,通过在编码翻转受体的供体修复模板存在下,将DSB引入一个或多个编码TCR信号传导组件(例如TCR α 基因的恒定区)、免疫系统检查点或免疫抑制信号传导组件的基因中来工程改造T细胞。如本文中所使用,术语“翻转受体”是指将来自肿瘤微环境的免疫抑制信号转化成T细胞中的免疫刺激信号的非天然存在的多肽。在优选实施例中,翻转受体是指一种多肽,其包含结合免疫抑制因子的胞外结构域、跨膜结构域和将免疫刺激信号转导到T细胞中的胞内结构域。

[0233] 在一个实施例中,供体修复模板包含翻转受体,其包含结合免疫抑制细胞因子的胞外结构域或胞外结合结构域、跨膜结构域和免疫增强细胞因子受体的胞内结构域。

[0234] 在具体实施例中,翻转受体包含结合免疫抑制细胞因子的胞外结构域,其是以下

的细胞外细胞因子结合结构域:IL-4受体、IL-6受体、IL-8受体、IL-10受体、IL-13受体或TGF β 受体;从CD4、CD8 α 、CD27、CD28、CD134、CD137、CD3多肽、IL-2受体、IL-7受体、IL-12受体、IL-15受体或IL-21受体分离的跨膜;和从IL-2受体、IL-7受体、IL-12受体、IL-15受体或IL-21受体分离的胞内结构域。

[0235] 在具体实施例中,翻转受体包含结合免疫抑制细胞因子的胞外结构域,其是结合IL-4、IL-6、IL-8、IL-10、IL-13或TGF β 的抗体或其抗原结合片段;从CD4、CD8 α 、CD27、CD28、CD134、CD137、CD3多肽、IL-2受体、IL-7受体、IL-12受体、IL-15受体或IL-21受体分离的跨膜;和从IL-2受体、IL-7受体、IL-12受体、IL-15受体或IL-21受体分离的胞内结构域。

[0236] 在一个实施例中,供体修复模板包含翻转受体,其包含结合免疫抑制因子的胞外结构域、跨膜结构域和一个或多个胞内共刺激信号传导结构域和/或初级信号传导结构域。

[0237] 适用于具体实施例中所涵盖的翻转受体的具体实施例中的胞外结构域的说明性实例包括(但不限于):受体的细胞外配位体结合结构域,其包含ITIM和/或ITSM。

[0238] 适用于具体实施例中所涵盖的翻转受体的具体实施例中的胞外结构域的其它说明性实例包括(但不限于)以下的细胞外配位体结合结构域:PD-1、LAG-3、TIM-3、CTLA-4、BTLA、CEACAM1、TIGIT、TGF β RII、IL4R、IL6R、CXCR1、CXCR2、IL10R、IL13R α 2、TRAILR1、RCAS1R和FAS。

[0239] 在一个实施例中,胞外结构域包含选自由以下组成的群组的受体的细胞外配位体结合结构域:PD-1、LAG-3、TIM3、CTLA-4、IL10R、TIGIT以及TGF β RII。

[0240] 在一个实施例中,供体修复模板包含翻转受体,其包含结合免疫抑制细胞因子的胞外结构域、跨膜结构域和一个或多个胞内共刺激信号传导结构域和/或初级信号传导结构域。

[0241] 适用于具体实施例中所涵盖的翻转受体的具体实施例中的跨膜结构域的说明性实例包括(但不限于)以下蛋白质的跨膜结构域:PD-1、LAG-3、TIM-3、CTLA-4、IL10R、TIGIT,和T细胞受体、CD δ 、CD3 ϵ 、CD γ 、CD3 ζ 、CD4、CD5、CD8 α 、CD9、CD16、CD22、CD27、CD28、CD33、CD37、CD45、CD64、CD80、CD86、CD134、CD137或CD154的TGF β RII α 或 β 链。在具体实施例中,优选的是选择与TCR信号传导复合物(例如CD3)相关的跨膜结构域,以增加免疫刺激信号。

[0242] 在各种实施例中,翻转受体包含引发免疫刺激信号的胞内结构域。如本文中所使用,术语「胞内结构域」是指免疫刺激基元或结构域,包括(但不限于)免疫受体酪氨酸活化基元(ITAM)、共刺激信号传导结构域、初级信号传导结构域或另一种与在T细胞中引发免疫刺激信号相关的细胞内结构域。

[0243] 适用于具体实施例中所涵盖的翻转受体的具体实施例中的胞内结构域的说明性实例包括(但不限于)包含ITAM基元的结构域。

[0244] 适用于具体实施例中所涵盖的翻转受体的具体实施例中的胞内结构域的其它说明性实例包括(但不限于)从以下分离的共刺激信号传导结构域:TLR1、TLR2、TLR3、TLR4、TLR5、TLR6、TLR7、TLR8、TLR9、TLR10、CARD11、CD2、CD7、CD27、CD28、CD30、CD40、CD54(ICAM)、CD83、CD134(OX40)、CD137(4-1BB)、CD278(ICOS)、DAP10、LAT、NKD2C、SLP76、TRIM或ZAP70。

[0245] 适用于具体实施例中所涵盖的翻转受体的具体实施例中的胞内结构域的其它说明性实例包括(但不限于):从IL-2受体、IL-7受体、IL-12受体、IL-15受体或IL-21受体分离的胞内结构域。

[0246] 适用于具体实施例中所涵盖的翻转受体的具体实施例中的胞内结构域的其它说明性实例包括(但不限于)从以下分离的初级信号传导结构域:FcR γ 、FcR β 、CD3 γ 、CD3 δ 、CD3 ϵ 、CD3 ζ 、CD22、CD79a、CD79b和CD66d。

[0247] 在具体实施例中,翻转受体包含胞外结构域,其包含来自PD-1、LAG-3、TIM-3、CTLA-4、IL10R、TIGIT或TGF β RII的细胞外结构域;跨膜结构域,其来自CD3多肽、CD4、CD8 α 、CD28、CD134、CD137、PD-1、LAG-3、TIM-3、CTLA-4、IL10R和TGF β RII;和胞内结构域,其来自CD28、CD134、CD137、CD278和/或CD3 ζ 。

[0248] 在具体实施例中,翻转受体包含胞外结构域,其包含来自PD-1、LAG-3、TIM-3、CTLA-4、IL10R、TIGIT或TGF β RII的细胞外结构域;跨膜结构域,其来自CD3多肽、CD4、CD8 α 、CD28、CD134或CD137;和胞内结构域,其来自CD28、CD134、CD137、CD278和/或CD3 ζ 。

[0249] 在具体实施例中,使用多个被工程改造的核酸酶诱导基因组中的多个目标位点中的DSB,并且将包含转基因的DNA供体修复模板引入细胞且通过同源重组插入至少一个目标位点中,所述转基因编码翻转受体并且由对应于基因组中的多个目标位点的多个同源臂侧接。

[0250] 在具体实施例中,使用多个被工程改造的核酸酶诱导基因组中的多个目标位点中的DSB,并且将包含转基因的DNA供体修复模板引入细胞且通过同源重组插入两个或更多个目标位点中,所述转基因编码翻转受体并且由对应于基因组中的多个目标位点的多个同源臂侧接。

[0251] 在具体实施例中,使用多个被工程改造的核酸酶诱导基因组中的多个目标位点中的DSB,并且将包含转基因的DNA供体修复模板引入细胞且通过同源重组插入每个目标位点中,所述转基因编码翻转受体并且由对应于基因组中的多个目标位点的多个同源臂侧接。

[0252] i. PD-1翻转受体

[0253] PD-1表达于T细胞上并且通过肿瘤微环境中的免疫抑制因子经历免疫抑制。PD-L1和PD-L2的表达与一些人类恶性病的预后相关。PD-L1/PD-1信号传导路径是T细胞耗竭的一个重要调节路径。PD-L1在癌细胞和基质细胞中大量表达,并且使用单克隆抗体阻断PD-L1/PD-1可以增强T细胞抗肿瘤功能。PD-L2还结合于PD-1并且不利地调节T细胞功能。

[0254] 在一个实施例中,在通过同源重组插入DSB处的一个或多个基因中的供体修复模板存在下,通过被工程改造的核酸酶在一个或多个编码TCR信号传导组件(例如TCR α 基因的恒定区)、免疫系统检查点或免疫抑制性信号传导组件的基因中引入DSB,所述供体修复模板包含一对或多对侧接转基因的同源臂,所述转基因编码PD-1翻转受体。

[0255] 具体实施例中所涵盖的PD-1翻转受体包含人类PD-1受体的细胞外配位体结合结构域;来自PD-1、CD3多肽、CD4、CD8 α 、CD28、CD134或CD137的跨膜结构域;和来自CD28、CD134、CD137、CD278和/或CD3 ζ 的胞内结构域。

[0256] ii. LAG-3翻转受体

[0257] 淋巴细胞活化基因-3 (LAG-3) 是对T细胞功能具有不同生物作用的细胞表面分子。LAG-3信号传导与自身免疫反应的CD4⁺调节性T细胞抑制相关。此外,LAG-3表达在CD8⁺T细胞的抗原刺激时增加并且与肿瘤微环境中的T细胞耗竭相关。LAG-3的体内抗体阻断与抗原特异性CD8⁺T细胞的积聚增加和效应功能相关。一个群组证实,给予抗LAG-3抗体与特异性抗肿瘤疫苗接种的组合引起肿瘤中活化的CD8⁺T细胞的显著增加和肿瘤实质的破坏。Grosso

等人(2007).《临床研究杂志(J Clin Invest.)》117(11):3383-3392。

[0258] 在一个实施例中,在通过同源重组插入DSB处的一个或多个基因中的供体修复模板存在下,通过被工程改造的核酸酶在一个或多个编码TCR信号传导组件(例如TCR α 基因的恒定区)、免疫系统检查点或免疫抑制性信号传导组件的基因中引入DSB,所述供体修复模板包含一对或多对侧接转基因的同源臂,所述转基因编码LAG-3翻转受体。

[0259] 具体实施例中涵盖的LAG-3翻转受体包含人类LAG-3受体的细胞外配位体结合结构域;来自LAG-3、CD3多肽、CD4、CD8 α 、CD28、CD134或CD137的跨膜结构域;和来自CD28、CD134、CD137、CD278和/或CD3 ζ 的胞内结构域。

[0260] iii.TIM-3翻转受体

[0261] 已确定T细胞免疫球蛋白-3(TIM-3)是阴性调节分子并且在免疫耐受性方面起作用。TIM-3表达鉴别在癌症中和在慢性感染期间耗竭的T细胞。表达TIM-3的CD4⁺和CD8⁺T细胞产生的细胞因子量降低并且对抗原起反应的增殖性更低。增加的TIM-3表达与降低的T细胞增殖和降低的IL-2、TNF和IFN- γ 产量相关。阻断TIM-3信号传导路径可以恢复增殖并且增加抗原特异性T细胞中的细胞因子产量。

[0262] TIM-3被共表达并且与癌胚抗原细胞粘附分子1(CEACAM1)形成杂二聚体,所述癌胚抗原细胞粘附分子1是另一种众所周知的表达于被活化的T细胞上并且涉及T-Cell1抑制的分子。存在CEACAM1可以赋予TIM-3抑制功能。通过每个分子的紧密相关的膜远端N端结构域,CEACAM1通过形成顺式杂二聚相互相用而有助于TIM-3的成熟和细胞表面表达。CEACAM1和TIM-3还通过其N端结构域反式结合。

[0263] 在一个实施例中,在通过同源重组插入DSB处的一个或多个基因中的供体修复模板存在下,通过被工程改造的核酸酶在一个或多个编码TCR信号传导组件(例如TCR α 基因的恒定区)、免疫系统检查点或免疫抑制性信号传导组件的基因中引入DSB,所述供体修复模板包含一对或多对侧接转基因的同源臂,所述转基因编码TIM-3翻转受体。

[0264] 具体实施例中涵盖的TIM-3翻转受体包含人类TIM-3受体的细胞外配位体结合结构域;来自TIM-3、CD3多肽、CD4、CD8 α 、CD28、CD134或CD137的跨膜结构域;和来自CD28、CD134、CD137、CD278和/或CD3 ζ 的胞内结构域。

[0265] iv.CTLA-4翻转受体

[0266] CTLA4主要表达于T细胞上,其在T细胞上调节早期T细胞活化的幅度。CTLA4对抗T细胞共刺激受体CD28的活性。CD28不影响T细胞活化,除非TCR首先由同源抗原接合。在进行抗原识别后,CD28信号传导强烈扩增TCR信号传导以活化T细胞。CD28和CTLA4共有一致配位体:CD80(亦称为B7.1)和CD86(亦称为B7.2)。CTLA4对于两种配位体具有显著更高的整体亲和力,并且通过在结合CD80和CD86方面胜过CD28以及积极地向T细胞递送抑制信号而减弱T细胞活化。CTLA4还通过来自CD28接合的CD80和CD86的螯合作用以及积极地从抗原呈现细胞(APC)表面去除CD80和CD86而实现信号传导独立性T细胞抑制。

[0267] 在一个实施例中,在通过同源重组插入DSB处的一个或多个基因中的供体修复模板存在下,通过被工程改造的核酸酶在一个或多个编码TCR信号传导组件(例如TCR α 基因的恒定区)、免疫系统检查点或免疫抑制性信号传导组件的基因中引入DSB,所述供体修复模板包含一对或多对侧接转基因的同源臂,所述转基因编码CTLA-4翻转受体。

[0268] 具体实施例中涵盖的CTLA-4翻转受体包含人类CTLA-4受体的细胞外配位体结

合结构域；来自CTLA-4、CD3多肽、CD4、CD8 α 、CD28、CD134或CD137的跨膜结构域；和来自CD28、CD134、CD137、CD278和/或CD3 ζ 的胞内结构域。

[0269] v. TIGIT翻转受体

[0270] T细胞免疫球蛋白和基于免疫受体酪氨酸的抑制性基元[ITIM]结构域(TIGIT)是鉴别为在多个实体肿瘤类型中持续大量表达的T细胞共抑制受体。TIGIT限制抗肿瘤和其它CD8 $^+$ T细胞依赖性慢性免疫反应。TIGIT大量表达于人类和鼠类肿瘤浸润性T细胞上。已证实，TIGIT的基因消融或抗体阻断可以体外和体内增强NK细胞杀伤和CD4 $^+$ T细胞激活，并且可以加重CD4 $^+$ T细胞依赖性自身免疫性疾病(如实验性自身免疫性脑炎)的严重度(Goding等人, 2013, Joller等人, 2011, Levin等人, 2011, Lozano等人, 2012, Stanietsky等人, 2009, Stanietsky等人, 2013, Stengel等人, 2012, Yu等人, 2009)。相反，给予TIGIT-Fc融合蛋白质或促效性抗TIGIT抗体可以抑制体外T细胞活化和体内CD4 $^+$ T细胞依赖性迟发型过敏(Yu等人, 2009)。TIGIT可能通过在结合于CD155方面胜过其相对共刺激受体CD226来发挥其免疫抑制作用。

[0271] 在癌症和慢性病毒感染的模型中，TIGIT和PD-L1的抗体共阻断协同并且特异性增强CD8 $^+$ T细胞效应功能，分别引起显著肿瘤和病毒清除作用。

[0272] 在一个实施例中，在通过同源重组插入DSB处的一个或多个基因中的供体修复模板存在下，通过被工程改造的核酸酶在一个或多个编码TCR信号传导组件(例如TCR α 基因的恒定区)、免疫系统检查点或免疫抑制性信号传导组件的基因中引入DSB，所述供体修复模板包含一对或多对侧接转基因的同源臂，所述转基因编码TIGIT翻转受体。

[0273] 具体实施例中所涵盖的TIGIT翻转受体包含人类TIGIT受体的细胞外配位体结合结构域；来自TIGIT、CD3多肽、CD4、CD8 α 、CD28、CD134或CD137的跨膜结构域；和来自CD28、CD134、CD137、CD278和/或CD3 ζ 的胞内结构域。

[0274] vi. TGF β RII翻转受体

[0275] 转化生长因子- β (TGF β)是由肿瘤细胞和免疫细胞产生的免疫抑制性细胞因子，其可极化免疫系统的多个臂。肿瘤细胞和肿瘤浸润性淋巴细胞过度产生免疫抑制性细胞因子(包括TGF β)将有助于免疫抑制性肿瘤微环境。TGF β 通常与肿瘤转移和侵袭、抑制免疫细胞功能和癌症患者中的不良预后相关。肿瘤特异性CTL中通过TGF β RII进行的TGF β 信号传导减弱其在肿瘤中的功能和出现率，并且用单克隆抗体阻断CD8 $^+$ T细胞上的TGF β 信号传导将实现更快速的肿瘤监测和在肿瘤位点处存在更多的CTL。

[0276] 在一个实施例中，在通过同源重组插入DSB处的一个或多个基因中的供体修复模板存在下，通过被工程改造的核酸酶在一个或多个编码TCR信号传导组件(例如TCR α 基因的恒定区)、免疫系统检查点或免疫抑制性信号传导组件的基因中引入DSB，所述供体修复模板包含一对或多对侧接转基因的同源臂，所述转基因编码TGF β RII翻转受体。

[0277] 具体实施例中所涵盖的TGF β RII翻转受体包含人类TGF β RII受体的细胞外配位体结合结构域；来自TGF β RII、CD3多肽、CD4、CD8 α 、CD28、CD134或CD137的跨膜结构域；和来自CD28、CD134、CD137、CD278和/或CD3 ζ 的胞内结构域。

[0278] 2. 免疫抑制信号减弱子

[0279] 困扰现有过继性细胞疗法的一个局限或问题是由肿瘤微环境介导的耗竭引起的免疫效应细胞反应性不足。耗竭的T细胞具有明显不同于原始T细胞、效应T细胞或记忆T细

胞的特有分子特征。其被定义为细胞因子表达和效应功能降低的T细胞。

[0280] 在具体实施例中,通过减少或阻断免疫抑制因子的信号传导使本文中所涵盖的基因组编辑的免疫效应细胞对耗竭具有更高的抗性。在一个实施例中,通过在编码免疫抑制信号减弱子的供体修复模板存在下,将DSB引入一个或多个编码TCR信号传导组件(例如TCR α 基因的恒定区)、免疫系统检查点或免疫抑制信号传导组件的基因中来工程改造T细胞。

[0281] 在具体实施例中,DNA供体模板被设计成将免疫抑制信号减弱子插入一个基因座中并且将一个或多个编码免疫效能增强子、另一免疫抑制信号减弱子或被工程改造的抗原受体的转基因插入一个或多个不同基因座中。

[0282] 如本文中所使用,术语“免疫抑制信号减弱子”是指减少免疫抑制信号从肿瘤微环境转导到T细胞的非天然存在的多肽。在一个实施例中,免疫抑制信号减弱子是结合免疫抑制因子的抗体或其抗原结合片段。在优选实施例中,因为免疫抑制信号减弱子包含结合免疫抑制因子的胞外结构域,和任选地,跨膜结构域,以及任选地,被修饰的胞内结构域(例如胞内信号传导结构域),所以所述减弱子是指引起针对具体免疫抑制因子或信号传导路径的抑制性、减弱性或显性负作用的多肽。

[0283] 在具体实施例中,胞外结构域是识别和结合免疫抑制因子的细胞外结合结构域。

[0284] 在具体实施例中,被修饰的胞内结构域突变以减少或抑制免疫抑制信号。合适的突变策略包括(但不限于)氨基酸取代、添加或缺失。合适的突变还包括(但不限于)截短胞内结构域以去除信号传导结构域;使胞内结构域突变以去除对于信号传导基元活性来说重要的残基;以及使胞内结构域突变以阻断受体循环。在具体实施例中,胞内结构域当存在时不会转导免疫抑制信号,或具有显著减少的信号传导。

[0285] 因此,在一些实施例中,免疫抑制信号减弱子充当来自肿瘤微环境的一个或多个免疫抑制因子的接收器(sink)并且抑制T细胞中的相应免疫抑制性信号传导路径。

[0286] 一种免疫抑制信号是由色氨酸分解代谢介导的。癌细胞中由吲哚胺2,3-双加氧酶(IDO)引起的色氨酸分解代谢引起产生犬尿氨酸,已证实其对肿瘤微环境中的T细胞具有免疫抑制作用。参见例如Platten等人(2012)《癌症研究(Cancer Res.)》72(21):5435-40。

[0287] 在一个实施例中,供体修复模板包含具有犬尿氨酸酶活性的酶。

[0288] 适用于具体实施例中的具有犬尿氨酸酶活性的酶的说明性实例包括(但不限于)L-犬尿氨酸水解酶。

[0289] 在一个实施例中,供体修复模板包含编码免疫抑制信号减弱子的一个或多个聚核苷酸,所述免疫抑制信号减弱子减少或阻断由免疫抑制因子介导的免疫抑制性信号传导。

[0290] 具体实施例中所涵盖的作为免疫抑制信号减弱子的目标的免疫抑制因子的说明性实例包括(但不限于):程序性死亡配位体1(PD-L1)、程序性死亡配位体2(PD-L2)、转化生长因子 β (TGF β)、巨噬细胞集落刺激因子1(M-CSF1)、肿瘤坏死因子相关细胞凋亡诱导配位体(TRAIL)、在SiSo细胞上表达的受体结合癌症抗原配位体(RCAS1)、Fas配位体(FasL)、CD47、白细胞介素-4(IL-4)、白细胞介素-6(IL-6)、白细胞介素-8(IL-8)、白细胞介素-10(IL-10)以及白细胞介素-13(IL-13)。

[0291] 在各种实施例中,免疫抑制信号减弱子包含结合免疫抑制因子的抗体或其抗原结合片段。

[0292] 在各种实施例中,免疫抑制信号减弱子包含结合免疫抑制因子的胞外结构域。

[0293] 在具体实施例中,免疫抑制信号减弱子包含结合免疫抑制因子的胞外结构域和跨膜结构域。

[0294] 在另一个实施例中,免疫抑制信号减弱子包含结合免疫抑制因子的胞外结构域、跨膜结构域以及被修饰的胞内结构域,所述被修饰的胞内结构域不转导免疫抑制信号或具有显著降低的转导免疫抑制信号的能力。

[0295] 如本文中所使用,术语“胞外结构域”是指抗原结合结构域。在一个实施例中,胞外结构域是将来自肿瘤微环境的免疫抑制信号转导至T细胞中的免疫抑制受体的细胞外配位体结合结构域。在具体实施例中,胞外结构域是指包含免疫受体酪氨酸抑制性基元(ITIM)和/或免疫受体酪氨酸转换基元(ITSM)的受体的细胞外配位体结合结构域。

[0296] 适用于免疫抑制信号减弱子的具体实施例中的胞外结构域的说明性实例包括(但不限于)抗体或其抗原结合片段,或从以下多肽分离的细胞外配位体结合结构域:程序性细胞死亡蛋白质1(PD-1)、淋巴细胞活化基因3蛋白质(LAG-3)、T细胞免疫球蛋白结构域和粘蛋白结构域蛋白质3(TIM3)、细胞毒性T淋巴细胞抗原-4(CTLA-4)、带状疱疹淋巴细胞弱化子(BTLA)、T细胞免疫球蛋白和基于免疫受体酪氨酸的抑制性基元结构域(TIGIT)、转化生长因子 β 受体II(TGF β RII)、巨噬细胞集落刺激因子1受体(CSF1R)、白细胞介素4受体(IL4R)、白细胞介素6受体(IL6R)、趋化因子(C-X-C基元)受体1(CXCR1)、趋化因子(C-X-C基元)受体2(CXCR2)、白细胞介素10受体子单元 α (IL10R)、白细胞介素13受体子单元 α 2(IL13R α 2)、肿瘤坏死因子相关细胞凋亡诱导配位体(TRAILR1)、在SiSo细胞上表达的受体结合癌症抗原(RCAS1R)以及Fas细胞表面死亡受体(FAS)。

[0297] 在一个实施例中,胞外结构域包含选自由以下组成的群组的受体的细胞外配位体结合结构域:PD-1、LAG-3、TIM3、CTLA-4、IL10R、TIGIT、CSF1R以及TGF β RII。

[0298] 在具体实施例中,可以使用多种跨膜结构域。适用于具体实施例中所涵盖的免疫抑制信号减弱子的具体实施例中的跨膜结构域的说明性实例包括(但不限于)以下蛋白质的跨膜结构域:T细胞受体的 α 或 β 链、CD δ 、CD3 ϵ 、CD γ 、CD3 ζ 、CD4、CD5、CD8 α 、CD9、CD 16、CD22、CD27、CD28、CD33、CD37、CD45、CD64、CD80、CD86、CD134、CD137、CD152、CD154和PD-1。

[0299] 在具体实施例中,本文中所涵盖的过继性细胞疗法包含抑制或阻断来自肿瘤微环境的免疫抑制TGF β 信号通过TGF β RII进行的转导的免疫抑制信号减弱子。在一个实施例中,免疫抑制信号减弱子包含有包含TGF β RII细胞外配体结合的胞外结构域、TGF β RII跨膜结构域以及截短的非功能性TGF β RII胞内结构域。在另一个实施例中,免疫抑制信号减弱子包含有包含TGF β RII细胞外配体结合的胞外结构域、TGF β RII跨膜结构域并且不含胞内结构域。

[0300] 在具体实施例中,使用多个被工程改造的核酸酶诱导基因组中的多个目标位点中的DSB,并且将包含转基因的DNA供体修复模板引入细胞且通过同源重组插入至少一个目标位点中,所述转基因编码免疫抑制信号减弱子并且由对应于基因组中的多个目标位点的多个同源臂侧接。

[0301] 在具体实施例中,使用多个被工程改造的核酸酶诱导基因组中的多个目标位点中的DSB,并且将包含转基因的DNA供体修复模板引入细胞且通过同源重组插入两个或更多个目标位点中,所述转基因编码免疫抑制信号减弱子并且由对应于基因组中的多个目标位点的多个同源臂侧接。

[0302] 在具体实施例中,使用多个被工程改造的核酸酶诱导基因组中的多个目标位点中

的DSB,并且将包含转基因的DNA供体修复模板引入细胞且通过同源重组插入每个目标位点中,所述转基因编码免疫抑制信号减弱子并且由对应于基因组中的多个目标位点的多个同源臂侧接。

[0303] 3. 被工程改造的抗原受体

[0304] 在具体实施例中,本文中所涵盖的被基因组编辑的免疫效应细胞包含被工程改造的抗原受体。在一个实施例中,通过在编码被工程改造的抗原受体的供体修复模板存在下,将DSB引入一个或多个编码TCR信号传导组件(例如TCR α 基因的恒定区)、免疫系统检查点或免疫抑制性信号传导组件的基因中来工程改造T细胞。

[0305] 在具体实施例中,DNA供体模板被设计成将被工程改造的抗原受体插入一个基因座中并且将一个或多个编码免疫效能增强子、免疫抑制信号减弱子或另一被工程改造的抗原受体的转基因插入一个或多个不同基因座中。

[0306] 在具体实施例中,被工程改造的抗原受体是被工程改造的T细胞受体(TCR)、嵌合抗原受体(CAR)、DARIC受体或其组件,或嵌合细胞因子受体。

[0307] 在具体实施例中,使用多个被工程改造的核酸酶诱导基因组中的多个目标位点中的DSB,并且将包含转基因的DNA供体修复模板引入细胞且通过同源重组插入至少一个目标位点中,所述转基因编码被工程改造的抗原受体并且由对应于基因组中的多个目标位点的多个同源臂侧接。

[0308] 在具体实施例中,使用多个被工程改造的核酸酶诱导基因组中的多个目标位点中的DSB,并且将包含转基因的DNA供体修复模板引入细胞且通过同源重组插入两个或更多个目标位点中,所述转基因编码被工程改造的抗原受体并且由对应于基因组中的多个目标位点的多个同源臂侧接。

[0309] 在具体实施例中,使用多个被工程改造的核酸酶诱导基因组中的多个目标位点中的DSB,并且将包含转基因的DNA供体修复模板引入细胞且通过同源重组插入每个目标位点中,所述转基因编码被工程改造的抗原受体并且由对应于基因组中的多个目标位点的多个同源臂侧接。

[0310] a. 被工程改造的TCR

[0311] 在具体实施例中,本文中所涵盖的被基因组编辑的免疫效应细胞包含被工程改造的TCR。在一个实施例中,通过在编码被工程改造的TCR的供体修复模板存在下,将DSB引入一个或多个编码TCR信号传导组件(例如TCR α 基因的恒定区)、免疫系统检查点或免疫抑制性信号传导组件的基因中来工程改造T细胞。

[0312] 在一个实施例中,本文中所涵盖的被工程改造的T细胞包含在TCR α 等位基因处插入的被工程改造的TCR和以下中的一个或多个:免疫抑制信号减弱子、翻转受体、嵌合抗原受体(CAR)、DARIC受体或其组件,或嵌合细胞因子受体被插入一个或多个编码另一TCR信号传导组件、免疫系统检查点或免疫抑制性信号传导组件的基因中的DSB中。

[0313] 天然存在的T细胞受体包含两个子单元,即 α 链和 β 链子单元,其各自是由每个T细胞的基因组中的重组事件产生的特有蛋白质。可以关于针对具体目标抗原的选择性来筛选TCR文库。通过这种方式,可以选择对目标抗原具有高亲合力和反应性的天然TCR,克隆且随后引入用于过继性免疫疗法的T细胞群体中。

[0314] 在一个实施例中,通过在一个或多个TCR α 对偶基因中的DSB处引入供体修复模板

来修饰T细胞,所述供体修复模板包含编码TCR的子单元的聚核苷酸,其中TCR子单元具有形成TCR的能力,所述TCR赋予T细胞针对表达目标抗原的肿瘤细胞的特异性。在具体实施例中,所述子单元与天然存在的子单元相比具有一个或多个氨基酸取代、缺失、插入或修饰,只要所述子单元保留形成TCR并且使被转染的T细胞能够将目标细胞导向且参与免疫相关细胞因子信号传导的能力即可。被工程改造的TCR优选还以高亲和力结合呈现相关肿瘤相关肽的目标细胞,并且任选地介导体内呈现相关肽的目标细胞的有效杀伤。

[0315] 编码被工程改造的TCR的核酸优选是从T细胞的(天然存在的)染色体中的其天然环境分离,并且可以并入合适的载体(如本文中其他地方所描述)中。在具体实施例中,核酸和包含其的载体都可以被转移到细胞,优选T细胞中。接着,被修饰的T细胞能够表达由一个或多个被转导的核酸编码的TCR的一条或多条链。在优选实施例中,由于被工程改造的TCR被引入通常不表达具体TCR的T细胞中,所以所述被工程改造的TCR是外源TCR。被工程改造的TCR的重要方面是其对由主要组织相容复合物(MHC)或类似免疫组件所呈现的肿瘤抗原具有高亲和力。相比于被工程改造的TCR,CAR被工程改造成以与MHC无关的方式结合目标抗原。

[0316] TCR可以被表达成具有其它多肽附接到本发明的TCR的 α 链或 β 链的氨基末端或羧基末端部分,只要所附接的其它多肽不干扰所述 α 链或 β 链形成功能性T细胞受体的能力以及MHC依赖性抗原识别即可。

[0317] 由具体实施例中涵盖的被工程改造的TCR识别的抗原包括(但不限于)癌症抗原,包括在血液癌和实体肿瘤上的抗原。说明性抗原包括(但不限于)FR α 、TPBG、 α v β 6整合素、BCMA、B7-H3、B7-H6、CD276、CD16、CD19、CD20、CD22、CD30、CD33、CD44、CD44v6、CD44v7/8、CD70、CD79a、CD79b、CD123、CD138、CD171、CEA、CSPG4、EGFR、ERBB2、EGFRvIII、EPCAM、EphA2、FAP、胚胎AchR、GD2、GD3、GPC3、HLA-A1+MAGEA1、HLA-A2+MAGEA1、HLA-A3+MAGEA1、MAGEA1、HLA-A1+NY-ESO-1、HLA-A2+NY-ESO-1、HLA-A3+NY-ESO-1、IL-11R α 、IL-13R α 2、 λ 轻链、路易斯-Y、 κ 轻链、间皮素、MUC1、MUC16、NCAM、NKG2D配位体、NY-ESO-1、PRAME、PSCA、PSMA、ROR1、SSX、存活素、TAG72、TEM1、TEM5、TEM7、TEM8、VEGFR2和WT-1。

[0318] 在一个实施例中,供体修复模板包含编码RNA聚合酶II启动子或第一自裂解病毒肽的聚核苷酸,和编码整合于一个被修饰的和/或非功能性TCR α 等位基因中的被工程改造的TCR的 α 链和/或 β 链的聚核苷酸。

[0319] 在一个实施例中,供体修复模板包含编码RNA聚合酶II启动子或第一自裂解病毒肽的聚核苷酸,和编码整合于一个被修饰的和/或非功能性TCR α 等位基因中的被工程改造的TCR的 α 链和/或 β 链的聚核苷酸。

[0320] 在具体实施例中,供体修复模板从5'到3'包含与TCR α 等位基因同源的5'同源臂、编码第一自裂解病毒肽的聚核苷酸、编码被工程改造的TCR的 α 链的聚核苷酸、编码第二自裂解病毒肽的聚核苷酸,和编码被工程改造的TCR的 β 链聚核苷酸,和与TCR α 等位基因同源的3'同源臂。在这种情况下,另一TCR α 等位基因可以是功能性的,或可以因DSB而具有降低的功能或呈现非功能性并且由NHEJ修复。在一个实施例中,另一TCR α 等位基因已由本文中所涵盖的被工程改造的核酸酶修饰并且可以具有降低的功能或呈现非功能性。

[0321] 在某一实施例中,两个TCR α 对偶基因都被修饰并且具有降低的功能或是非功能性的。

[0322] b. 嵌合抗原受体 (CAR)

[0323] 在具体实施例中,本文中所涵盖的被工程改造的免疫效应细胞包含一个或多个再引导针对肿瘤细胞的细胞毒性的嵌合抗原受体 (CAR)。CAR是将针对目标抗原(例如肿瘤抗原)的基于抗体的特异性与T细胞受体活化细胞内结构域组合以产生呈现特异性抗肿瘤细胞免疫活性的嵌合蛋白的分子。如本文中所使用,术语“嵌合”描述由来自不同来源的不同蛋白质或DNA部分构成。在一个实施例中,通过在编码CAR的供体修复模板存在下,将DSB引入一个或多个编码TCR信号传导组件(例如TCR α 基因的恒定区)、免疫系统检查点或免疫抑制信号传导组件的基因中来工程改造T细胞。

[0324] 在一个实施例中,本文中所涵盖的被工程改造的T细胞包含在TCR α 等位基因处插入的CAR和以下中的一个或多个:免疫抑制信号减弱子、翻转受体、被工程改造的T细胞受体(TCR)的 α 和/或 β 链、嵌合抗原受体(CAR)、DARIC受体或其组件,或嵌合细胞因子受体被插入一个或多个编码TCR信号传导组件、免疫系统检查点或免疫抑制性信号传导组件的基因中的DSB中。

[0325] 在各种实施例中,被基因组编辑的T细胞在两个或更多个遗传基因座中表达一个或多个CAR。CAR可以是相同或不同的。在一个实施例中,借助于DNA供体修复模板将相同的CAR插入各种基因组位置中,所述DNA供体修复模板包含与侧接编码CAR的转基因的各种基因组位置同源的同源臂对。

[0326] 在具体实施例中,DNA供体模板被设计成将CAR插入一个基因座中并且将一个或多个编码免疫效能增强子、免疫抑制信号减弱子或另一被工程改造的抗原受体的转基因插入一个或多个不同基因座中。

[0327] 在各种实施例中,CAR包含结合于特异性目标抗原(又称为结合结构域或抗原特异性结合结构域)的细胞外结构域、跨膜结构域和细胞内信号传导结构域。CAR的主要特征是其能够再引导免疫效应细胞特异性,由此触发增殖、细胞因子产生、吞噬作用或能够以与主要组织相容性(MHC)无关的方式介导目标抗原表达细胞的细胞死亡的分子的产生,利用单克隆抗体、可溶性配位体或细胞特异性共受体的细胞特异性靶向能力。

[0328] 在具体实施例中,CAR包含特异性结合于肿瘤细胞上表达的目标多肽,例如目标抗原的细胞外结合结构域。如本文中所使用,术语“结合结构域”、“细胞外结构域”、“细胞外结合结构域”、“抗原结合结构域”、“抗原特异性结合结构域”以及“细胞外抗原特异性结合结构域”可以互换使用并且提供能够特异性结合于相关目标抗原的嵌合受体,例如CAR或DARIC。结合结构域可以包含能够特异性识别并且结合于生物分子(例如细胞表面受体或肿瘤蛋白质、脂质、多糖或其它细胞表面目标分子,或其组件)的任何蛋白质、多肽、寡肽或肽。结合结构域包括相关生物分子的任何天然存在的、合成的、半合成的或以重组方式产生的结合搭配物。

[0329] 在具体实施例中,细胞外结合结构域包含抗体或其抗原结合片段。

[0330] “抗体”是指一种结合剂,其是特异性识别并且结合目标抗原(如肽、脂质多糖或含有抗原决定子(如免疫细胞所识别的那些)的核酸)的表位的多肽,其包含至少一个轻链或重链免疫球蛋白可变区。抗体包括抗原结合片段,例如骆驼Ig(骆驼抗体或其VHH片段)、IgNAR、DARPin、FN3片段、Fab片段、Fab'片段、F(ab)'₂片段、F(ab)'₃片段、Fv、单链Fv抗体(“scFv”)、双scFv、(scFv)₂、微型抗体、双功能抗体、三功能抗体、四功能抗体、二硫键稳定

的Fv蛋白质(“dsFv”)以及单结构域抗体(sdAb、纳米抗体(Nanobody))或它们的其它抗体片段。所述术语还包括被基因工程改造的形式,如嵌合抗体(例如人类化鼠抗体)、异结合抗体(如双特异性抗体)和其抗原结合片段。还参见《皮尔斯目录和手册(Pierce Catalog and Handbook)》,1994-1995(Pierce Chemical Co., Rockford, IL);Kuby, J.,《免疫学(Immunology)》,第3版,W.H.Freeman&Co., New York, 1997。

[0331] 在一个优选实施例中,结合结构域是scFv。

[0332] 在另一个优选实施例中,结合结构域是骆驼抗体。

[0333] 在具体实施例中,CAR包含结合选自由以下组成的群组的抗原的细胞外结构域:FR α 、TPBG、 α v β 6整合素、BCMA、B7-H3、B7-H6、CD276、CD16、CD19、CD20、CD22、CD30、CD33、CD44、CD44v6、CD44v7/8、CD70、CD79a、CD79b、CD123、CD138、CD171、CEA、CSPG4、EGFR、ERBB2、EGFRvIII、EPCAM、EphA2、FAP、胚胎AchR、GD2、GD3、GPC3、HLA-A1+MAGEA1、HLA-A2+MAGEA1、HLA-A3+MAGEA1、MAGEA1、HLA-A1+NY-ESO-1、HLA-A2+NY-ESO-1、HLA-A3+NY-ESO-1、IL-11R α 、IL-13R α 2、 λ 轻链、路易斯-Y、 κ 轻链、间皮素、MUC1、MUC16、NCAM、NKG2D配位体、NY-ESO-1、PRAME、PSCA、PSMA、ROR1、SSX、存活素、TAG72、TEM1、TEM5、TEM7、TEM8、VEGFR2和WT-1。

[0334] 在具体实施例中,CAR包含细胞外结合结构域,例如结合抗原的抗体或其抗原结合片段,其中所述抗原是MHC-肽复合物,如I类MHC-肽复合物或II类MHC-肽复合物。

[0335] 在某些实施例中,CAR在多个结构域之间包含连接子残基。“可变区连接序列”是满足以下条件的氨基酸序列:其将重链可变区连接到轻链可变区,并且提供与所述两个子结合结构域的相互作用相容的间隔子功能,使得所得多肽保持针对与包含相同轻链和重链可变区的抗体相同的目标分子的特异性结合亲和力。在具体实施例中,CAR包含一个、两个、三个、四个或五个或更多个连接子。在具体实施例中,连接子的长度是约1到约25个氨基酸、约5到约20个氨基酸,或约10到约20个氨基酸,或任何中间长度的氨基酸。在一些实施例中,连接子的长度是1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25或更多个氨基酸。

[0336] 在具体实施例中,CAR的结合结构域之后是一个或多个“间隔子结构域”,间隔子结构域是指使抗原结合结构域移动远离效应细胞表面以便能实现适当细胞/细胞接触、抗原结合和活化的区域(Patel等人,《基因疗法(Gene Therapy)》,1999;6:412-419)。间隔子结构域可以来源于天然、合成、半合成或重组型来源。在某些实施例中,间隔子结构域是免疫球蛋白的一部分,所述免疫球蛋白包括(但不限于)一个或多个重链恒定区,例如CH2和CH3。间隔子结构域可以包括天然存在的免疫球蛋白铰链区或改变的免疫球蛋白铰链区的氨基酸序列。

[0337] 在一个实施例中,间隔子结构域包含IgG1、IgG4或IgD的CH2和CH3。

[0338] 在一个实施例中,CAR的结合结构域连接到一个或多个“铰链结构域”,铰链结构域在定位远离效应细胞表面的抗原结合结构域以便能实现适当细胞/细胞接触、抗原结合和活化中起作用。CAR通常在结合结构域与跨膜结构域(TM)之间包含一个或多个铰链结构域。铰链结构域可以来源于天然、合成、半合成或重组型来源。铰链结构域可以包括天然存在的免疫球蛋白铰链区或改变的免疫球蛋白铰链区的氨基酸序列。

[0339] 适用于本文中所述的CAR中的说明性铰链结构域包括来源于1型膜蛋白(如CD8 α 和CD4)的细胞外区域的铰链区,其可以是来自这些分子的野生型铰链区或可以改变。在另

一个实施例中,铰链结构域包含CD8 α 铰链区。

[0340] 在一个实施例中,铰链是PD-1铰链或CD152铰链。

[0341] “跨膜结构域”是CAR中融合细胞外结合部分和细胞内信号传导结构域并且将CAR锚定到免疫效应细胞的质膜的部分。TM结构域可以来源于天然、合成、半合成或重组型来源。

[0342] 说明性TM结构域可以来源于(即,至少包含)T细胞受体的 α 或 β 链的跨膜区、CD3 δ 、CD3 ϵ 、CD3 γ 、CD3 ζ 、CD4、CD5、CD8 α 、CD9、CD16、CD22、CD27、CD28、CD33、CD37、CD45、CD64、CD80、CD86、CD134、CD137、CD152、CD154以及PD-1。

[0343] 在一个实施例中,CAR包含来源于CD8 α 的TM结构域。在另一个实施例中,本文中涵盖的CAR包含来源于CD8 α 的TM结构域以及长度优选在1、2、3、4、5、6、7、8、9或10个氨基酸之间的短寡肽或多肽连接子,所述连接子连接CAR的TM结构域和细胞内信号传导结构域。甘氨酸-丝氨酸连接子提供尤其合适的连接子。

[0344] 在具体实施例中,CAR包含细胞内信号传导结构域。“细胞内信号传导结构域”是指CAR中参与将有效CAR结合于目标抗原的信息转导到免疫效应细胞内部以引起效应细胞功能的部分,所述效应细胞功能例如是活化、细胞因子产生、增殖和细胞毒性活性,包括细胞毒性因子释放到结合CAR的目标细胞中,或由抗原结合于细胞外CAR结构域引起的其它细胞反应。

[0345] 术语“效应功能”是指细胞的特有功能。T细胞的效应功能例如可以是细胞溶解活性或帮助或活性,包括细胞因子的分泌。因此,术语“细胞内信号传导结构域”是指转导效应功能信号且引导细胞执行特有功能的蛋白质的部分。尽管通常可以使用整个细胞内信号传导结构域,但在许多情况下无需使用整个结构域。就使用细胞内信号传导结构域的截短部分来说,可以使用这类截短部分代替整个结构域,只要其转导效应功能信号即可。术语细胞内信号传导结构域意图包括足以转导效应功能信号的细胞内信号传导结构域的任何截短部分。

[0346] 众所周知,仅由TCR产生的信号不足以完全活化T细胞并且还需要二次或共刺激信号。因此,T细胞活化可以被认为是由二种不同类别的细胞内信号传导结构域介导:通过TCR(例如TCR/CD3复合物)起始抗原依赖性初级活化的初级信号传导结构域以及以与抗原无关的方式起作用以提供二级或共刺激信号的共刺激信号传导结构域。在优选实施例中,CAR包含细胞内信号传导结构域,所述细胞内信号传导结构域包含一个或多个“共刺激信号传导结构域”和一个“初级信号传导结构域”。

[0347] 初级信号传导结构域以刺激性方式或以抑制性方式调节TCR复合物的初级活化。以刺激性方式起作用的初级信号传导结构域可以含有信号传导基元,所述信号传导基元称为基于免疫受体酪氨酸的活化基元或ITAM。

[0348] 适用于具体实施例中所涵盖的CAR中的含有ITAM的初级信号传导结构域的说明性实例包括来源于FcR γ 、FcR β 、CD3 γ 、CD3 δ 、CD3 ϵ 、CD3 ζ 、CD22、CD79a、CD79b以及CD66d的初级信号传导结构域。在具体优选实施例中,CAR包含CD3 ζ 初级信号传导结构域和一个或多个共刺激信号传导结构域。细胞内初级信号传导结构域和共刺激信号传导结构域可以按任何次序串联连接到跨膜结构域的羧基端。

[0349] 在具体实施例中,CAR包含一个或多个共刺激信号传导结构域以增强表达CAR受体

的T细胞的功效和扩增。如本文中所使用,术语“共刺激信号传导结构域”或“共刺激结构域”是指共刺激分子的细胞内信号传导结构域。

[0350] 适用于具体实施例中所涵盖的CAR中的这类共刺激分子的说明性实例包括TLR1、TLR2、TLR3、TLR4、TLR5、TLR6、TLR7、TLR8、TLR9、TLR10、CARD11、CD2、CD7、CD27、CD28、CD30、CD40、CD54 (ICAM)、CD83、CD134 (OX40)、CD137 (4-1BB)、CD278 (ICOS)、DAP10、LAT、NKD2C、SLP76、TRIM和ZAP70。在一个实施例中,CAR包含一个或多个选自由CD28、CD137和CD134组成的群组的共刺激信号传导结构域,以及一个CD3 ζ 初级信号传导结构域。

[0351] 在各种实施例中,CAR包含:结合选自由以下组成的群组的抗原的细胞外结构域:BCMA、CD19、CSPG4、PSCA、ROR1和TAG72;从选自由以下组成的群组的多肽分离的跨膜结构域:CD4、CD8 α 、CD154和PD-1;从选自由以下组成的群组的多肽分离的一个或多个细胞内共刺激信号传导结构域:CD28、CD134和CD137;以及从选自由以下组成的群组的多肽分离的信号传导结构域:FcR γ 、FcR β 、CD3 γ 、CD3 δ 、CD3 ϵ 、CD3 ζ 、CD22、CD79a、CD79b和CD66d。

[0352] c. DARIC受体

[0353] 在具体实施例中,被工程改造的免疫效应细胞包含一个或多个DARIC受体。如本文中所使用,术语“DARIC”或“DARIC受体”是指多链被工程改造的抗原受体。在一个实施例中,通过在编码DARIC的一个或多个组件的供体修复模板存在下,将DSB引入一个或多个编码TCR信号传导组件(例如TCR α 基因的恒定区)、免疫系统检查点或免疫抑制性信号传导组件的基因中来工程改造T细胞。

[0354] 在一个实施例中,被工程改造的T细胞包含一个不在TCR α 等位基因处插入的DARIC,并且以下中的一个或多个被插入一个或多个TCR α 对偶基因中的DSB中:免疫抑制信号减弱子、翻转受体、被工程改造的T细胞受体(TCR)的 α 和/或 β 链、嵌合抗原受体(CAR)或DARIC受体或其组件。

[0355] 在具体实施例中,DNA供体模板被设计成将DARIC插入一个基因座中并且将一个或多个编码免疫效能增强子、免疫抑制信号减弱子或另一被工程改造的抗原受体的转基因插入一个或多个不同基因座中。

[0356] DARIC架构和组件的说明性实例公开于PCT公开案第W02015/017214号和美国专利公开案第20150266973号中,其各自以全文引用的方式并入本文中。

[0357] 在一个实施例中,供体修复模板包含以下DARIC组件:信号传导多肽,其包含第一多聚化结构域、第一跨膜结构域和一个或多个细胞内共刺激信号传导结构域和/或初级信号传导结构域;以及结合多肽,其包含结合结构域、第二多聚化结构域和任选地第二跨膜结构域。功能性DARIC包含促进在细胞表面上形成DARIC受体复合物的桥接因子,并且所述桥接因子与信号传导多肽和结合多肽的多聚化结构域缔合并安置在信号传导多肽和结合多肽的多聚化结构域之间。

[0358] 在具体实施例中,第一多聚化结构域和第二多聚化结构域与选自由以下组成的群组的桥接因子缔合:雷帕霉素(rapamycin)或其雷帕霉素类似物、库马霉素(coumermycin)或其衍生物、赤霉素(gibberellin)或其衍生物、脱落酸(ABA)或其衍生物、甲氨蝶呤(methotrexate)或其衍生物、环孢菌素A(cyclosporin A)或其衍生物、FKCsA或其衍生物、针对FKBP(SLF)的甲氧苄氨嘧啶(trimethoprim;Tmp)合成配位体或其衍生物,和其任何组合。

[0359] 雷帕霉素类似物的说明性实例包括美国专利案第6,649,595号中所公开的那些,所述雷帕霉素类似物结构以全文引用的方式并入本文中。在某些实施例中,桥接因子是与雷帕霉素相比免疫抑制作用显著降低的雷帕霉素类似物。“显著降低的免疫抑制作用”是指如在临床上或在具有人免疫抑制活性的适当体外(例如抑制T细胞增殖)或体内替代物中所测量,雷帕霉素类似物的免疫抑制作用是关于等摩尔量的雷帕霉素所观察或预期的免疫抑制作用的至少低于0.1至0.005倍。在一个实施例中,“显著降低的免疫抑制作用”是指雷帕霉素类似物在这类体外分析法中的EC₅₀值比在同一分析法中关于雷帕霉素所观察到的EC₅₀值大至少10至250倍。

[0360] 雷帕霉素类似物的其它说明性实例包括(但不限于)依维莫司(everolimus)、诺瓦莫司(novolimus)、吡美莫司(pimecrolimus)、地磷莫司(ridaforolimus)、他克莫司(tacrolimus)、坦罗莫司(temsirolimus)、优米莫司(umiroliimus)以及佐他莫司(zotarolimus)。

[0361] 在某些实施例中,多聚化结构域将与桥接因子缔合,所述桥接因子是雷帕霉素或其雷帕霉素类似物。举例来说,第一多聚化结构域和第二多聚化结构域是选自FKBP和FRB的一对。FRB结构域是能够与FKBP蛋白质和雷帕霉素或其雷帕霉素类似物形成三联复合物的多肽区域(蛋白质“结构域”)。FRB结构域存在于多种天然存在的蛋白质中,包括来自人类和其它物种的mTOR蛋白质(在文献中也称为FRAP、RAP1或RAFT);酵母蛋白质,包括Tor1和Tor2;以及念珠菌属FRAP同源物。关于这些蛋白质的核苷酸序列、克隆以及其它方面的信息是所属领域中已知的。举例来说,人类mTOR的蛋白质序列寄存编号是GenBank寄存编号L34075.1(Brown等人,《自然(Nature)》369:756,1994)。

[0362] 适用于本文中所涵盖的具体实施例中的FRB结构域通常含有至少约85到约100个氨基酸残基。在某些实施例中,基于GenBank寄存编号L34075.1的氨基酸序列,用于本公开的融合蛋白质中的FRB氨基酸序列将包含93个氨基酸的序列Ile-2021到Lys-2113和T2098L突变。用于具体实施例中所涵盖的DARIC中的FRB结构域将能够与结合于雷帕霉素或其雷帕霉素类似物的FKBP蛋白质的复合物结合。在某些实施例中,FRB结构域的肽序列包含(a)至少跨越人类mTOR中所指示的93个氨基酸的区域或同源蛋白质的相应区域的天然存在的肽序列;(b)天然存在的FRB的变体,其中所述天然存在的肽的最多约十个氨基酸,或约1到约5个氨基酸,或约1到约3个氨基酸,或在一些实施例中仅一个氨基酸已经缺失、插入或被取代;或(c)由能够与编码天然存在的FRB结构域的DNA分子选择性杂交的核酸分子或由因遗传密码的简并性而能够与编码天然存在的FRB结构域的DNA分子选择性杂交的DNA序列编码的肽。

[0363] FKBP(FK506结合蛋白质)是针对大环内酯(如FK506、FK520和雷帕霉素)的胞质受体,并且在各物种界间高度保守。FKBP是能够结合于雷帕霉素或其雷帕霉素类似物并且进一步与含有FRB的蛋白质或融合蛋白质形成三联复合物的蛋白质或蛋白质结构域。FKBP结构域也可以称为“雷帕霉素结合结构域”。关于各种FKBP物种的核苷酸序列、克隆和其它方面的信息是所属领域中已知的(参见例如Staendart等人,《自然》346:671,1990(人类FKBP12);Kay,《生物化学杂志(Biochem. J.)》314:361,1996)。其它哺乳动物物种、酵母以及在其它生物体中的同源FKBP蛋白质也是所属领域中已知的并且可以用于本文中所公开的融合蛋白质中。具体实施例中所涵盖的FKBP结构域将能够结合于雷帕霉素或其雷帕霉素类

似物并且涉及具有含有FRB蛋白质的三联复合物(可以通过用于检测这类结合的任何手段直接或间接地测定)。

[0364] 适用于具体实施例中所涵盖的DARIC中的FKBP结构域的说明性实例包括(但不限于):天然存在的FKBP肽序列,优选从人类FKBP12蛋白质(GenBank寄存编号AAA58476.1)分离的FKBP肽序列或从其、从另一个人类FKBP、从鼠类或其它哺乳动物FKBP或从某一其它动物、酵母或真菌FKBP分离的肽序列;天然存在的FKBP序列的变异体,其中所述天然存在的肽中最多约十个氨基酸,或约1到约5个氨基酸或约1到约3个氨基酸,或在一些实施例中仅一个氨基酸已经缺失、插入或被取代;或由能够与编码天然存在的FKBP的DNA分子选择性杂交的核酸分子或由因遗传密码的简并性而能够与编码天然存在的FKBP的DNA分子选择性杂交的DNA序列编码的肽序列。

[0365] 适用于具体实施例中所涵盖的DARIC中的多聚化结构域对的其它说明性实例包括(但不限于)包括来自FKBP与FRB、FKBP与钙调神经磷酸酶、FKBP与亲环蛋白、FKBP与细菌DHFR、钙调神经磷酸酶与亲环蛋白、PYL1与ABI1、或GIB1与GAI,或其变异体。

[0366] 在其它实施例中,抗桥接因子阻断信号传导多肽和结合多肽与桥接因子的缔合。举例来说,环孢菌素或FK506可以用作抗桥接因子以滴定去除雷帕霉素并且因此,由于仅结合一个多聚化结构域而停止信号传导。在某些实施例中,抗桥接因子(例如环孢菌素、FK506)是一种免疫抑制剂。举例来说,可以使用免疫抑制性抗桥接因子来阻断或最小化具体实施例中所涵盖的DARIC组件的功能,并且同时在临床环境中抑制或阻断不合需要的或病理性炎性反应。

[0367] 在一个实施例中,第一多聚化结构域包含FRB T2098L,第二多聚化结构域包含FKBP12,并且桥接因子是雷帕霉素类似物AP21967。

[0368] 在另一个实施例中,第一多聚化结构域包含FRB,第二多聚化结构域包含FKBP12,并且桥接因子是雷帕霉素、坦罗莫司或依维莫司。

[0369] 在具体实施例中,信号传导多肽第一跨膜结构域和结合多肽包含第二跨膜结构域或GPI锚。第一和第二跨膜结构域的说明性实例是从独立地选自由以下组成的群组的多肽分离:CD3 δ 、CD3 ϵ 、CD3 γ 、CD3 ζ 、CD4、CD5、CD8 α 、CD9、CD16、CD22、CD27、CD28、CD33、CD37、CD45、CD64、CD80、CD86、CD134、CD137、CD152、CD154和PD-1。

[0370] 在一个实施例中,信号传导多肽包含一个或多个细胞内共刺激信号传导结构域和/或初级信号传导结构域。

[0371] 适用于具体实施例中所涵盖的DARIC信号传导组件中的初级信号传导结构域的说明性实例包括来源于FcR γ 、FcR β 、CD3 γ 、CD3 δ 、CD3 ϵ 、CD3 ζ 、CD22、CD79a、CD79b以及CD66d的初级信号传导结构域。在具体优选实施例中,DARIC信号传导组件包含CD3 ζ 初级信号传导结构域和一个或多个共刺激信号传导结构域。细胞内初级信号传导结构域和共刺激信号传导结构域可以按任何次序串联连接到跨膜结构域的羧基端。

[0372] 适用于具体实施例中所涵盖的DARIC信号传导组件中的这类共刺激分子的说明性实例包括TLR1、TLR2、TLR3、TLR4、TLR5、TLR6、TLR7、TLR8、TLR9、TLR10、CARD11、CD2、CD7、CD27、CD28、CD30、CD40、CD54(ICAM)、CD83、CD134(OX40)、CD137(4-1BB)、CD278(ICOS)、DAP10、LAT、NKD2C、SLP76、TRIM和ZAP70。在一个实施例中,DARIC信号传导组件包含一个或多个选自由CD28、CD137和CD134组成的群组的共刺激信号传导结构域,以及一个CD3 ζ 初级

信号传导结构域。

[0373] 在具体实施例中,DARIC结合组件包含结合结构域。在一个实施例中,结合结构域是抗体或其抗原结合片段。

[0374] 抗体或其抗原结合片段包含特异性地识别并且结合目标抗原(如肽、脂质、多糖或含有抗原决定子(如由免疫细胞识别的那些)的核酸)的表位的至少一个轻链或重链免疫球蛋白可变区。抗体包括抗原结合片段,例如骆驼Ig(骆驼抗体或其VHH片段)、Ig NAR、DARpins、FN3片段、Fab片段、Fab'片段、F(ab)'2片段、F(ab)'3片段、Fv、单链Fv抗体("scFv")、双scFv、(scFv)₂、微型抗体、双功能抗体、三功能抗体、四功能抗体、二硫键稳定的Fv蛋白质("dsFv")以及单结构域抗体(sdAb、纳米抗体)或它们的其它抗体片段。所述术语还包括被基因工程改造的形式,如嵌合抗体(例如人类化鼠抗体)、异结合抗体(如双特异性抗体)和其抗原结合片段。还参见《皮尔斯目录与手册(Pierce Catalog and Handbook)》,1994-1995(Pierce Chemical Co.,Rockford,IL);Kuby,J.,《免疫学(Immunology)》,第3版,W.H.Freeman&Co.,New York,1997。

[0375] 在一个优选实施例中,结合结构域是scFv。

[0376] 在另一个优选实施例中,结合结构域是骆驼抗体。

[0377] 在具体实施例中,DARIC结合组件包含结合选自由以下组成的群组的抗原的细胞外结构域:FR α 、TPBG、 α v β 6整合素、BCMA、B7-H3、B7-H6、CD276、CD16、CD19、CD20、CD22、CD30、CD33、CD44、CD44v6、CD44v7/8、CD70、CD79a、CD79b、CD123、CD138、CD171、CEA、CSPG4、EGFR、ERBB2、EGFRvIII、EPCAM、EphA2、FAP、胚胎AChR、GD2、GD3、GPC3、HLA-A1⁺MAGEA1、HLA-A2⁺MAGEA1、HLA-A3⁺MAGEA1、MAGEA1、HLA-A1⁺NY-ESO-1、HLA-A2⁺NY-ESO-1、HLA-A3⁺NY-ESO-1、IL-11R α 、IL-13R α 2、 λ 轻链、路易斯-Y、 κ 轻链、间皮素、MUC1、MUC16、NCAM、NKG2D配位体、NY-ESO-1、PRAME、PSCA、PSMA、ROR1、SSX、存活素、TAG72、TEM1、TEM5、TEM7、TEM8、VEGFR2和WT-1。

[0378] 在一个实施例中,DARIC结合组件包含细胞外结构域,例如结合MHC-肽复合物(如I类MHC-肽复合物或II类MHC-肽复合物)的抗体或其抗原结合片段。

[0379] 在具体实施例中,本文中所涵盖的DARIC组件包含连接两个蛋白质、多肽、肽、结构域、区域或基元的连接子或间隔子。在某些实施例中,连接子包含约两个到约35个氨基酸,或约四个到约20个氨基酸,或约八个到约15个氨基酸,或约15个到约25个氨基酸。在其它实施例中,间隔子可以具有具体结构,如抗体CH₂CH₃结构域、铰链结构域等。在一个实施例中,间隔子包含IgG1、IgG4或IgD的CH₂和CH₃结构域。

[0380] 在具体实施例中,本文中所涵盖的DARIC组件包含一个或多个“铰链结构域”,所述铰链结构域在定位结构域以便能实现适当细胞/细胞接触、抗原结合和活化中起作用。DARIC可以在结合结构域与多聚化结构域和/或跨膜结构域(TM)之间或在多聚化结构域与跨膜结构域之间包含一个或多个铰链结构域。铰链结构域可以来源于天然、合成、半合成或重组型来源。铰链结构域可以包括天然存在的免疫球蛋白铰链区或改变的免疫球蛋白铰链区的氨基酸序列。在具体实施例中,铰链是CD8 α 铰链或CD4铰链。

[0381] 在一个实施例中,DARIC包含信号传导多肽,其包含FRB T2098L的第一多聚化结构域、CD8跨膜结构域、4-1BB共刺激结构域和CD3 ζ 初级信号传导结构域;结合多肽包含结合CD19的scFv、FKBP12的第二多聚化结构域和CD4跨膜结构域;并且桥接因子是雷帕霉素类似物AP21967。

[0382] 在一个实施例中,DARIC包含信号传导多肽,其包含FRB的第一多聚化结构域、CD8跨膜结构域、4-1BB共刺激结构域和CD3 ζ 初级信号传导结构域;结合多肽包含结合CD19的scFv、FKBP12的第二多聚化结构域和CD4跨膜结构域;并且桥接因子是雷帕霉素、坦罗莫司或依维莫司。

[0383] d. ζ 因子

[0384] 在具体实施例中,本文中所涵盖的被工程改造的免疫效应细胞包含一个或多个嵌合细胞因子受体。在一个实施例中,通过在编码 ζ 因子的供体修复模板存在下,将DSB引入一个或多个编码TCR信号传导组件(例如TCR α 基因的恒定区)、免疫系统检查点或免疫抑制信号传导组件的基因中来工程改造T细胞。

[0385] 在一个实施例中,本文中所涵盖的被工程改造的T细胞是一个不在TCR α 等位基因处插入的嵌合细胞因子受体,并且以下中的一个或多个被插入一个或多个TCR α 对偶基因中的DSB中:免疫抑制信号减弱子、翻转受体、被工程改造的T细胞受体(TCR)的 α 和/或 β 链、嵌合抗原受体(CAR)、DARIC受体或其组件或嵌合细胞因子受体。

[0386] 在具体实施例中,DNA供体模板被设计成将 ζ 因子插入一个基因座中并且将一个或多个编码免疫效能增强子、免疫抑制信号减弱子或另一被工程改造的抗原受体的转基因插入一个或多个不同基因座中。

[0387] 在各种实施例中,被基因组编辑的T细胞表达将细胞毒性再引导向肿瘤细胞的嵌合细胞因子受体。 ζ 因子是嵌合跨膜免疫受体,其包含细胞外结构域、跨膜区和细胞内信号传导结构域,所述细胞外结构域包含可溶性受体配位体,其连接到能够将细胞外结构域系栓到细胞表面的载体区域。 ζ 因子当在T淋巴细胞的表面上表达时,将T细胞活性引导到表达所述可溶性受体配位体特异性针对的受体的细胞。 ζ 因子嵌合免疫受体再引导T细胞的抗原特异性,并且用于治疗多种癌症,特别是通过人类恶性疾病利用的自分泌/旁分泌细胞因子系统进行治疗。

[0388] 在具体实施例中,嵌合细胞因子受体包含免疫抑制细胞因子或其细胞因子受体结合变体、连接子、跨膜结构域和细胞内信号传导结构域。

[0389] 在具体实施例中,细胞因子或其细胞因子受体结合变体是选自由以下组成的群组:白细胞介素-4(IL-4)、白细胞介素-6(IL-6)、白细胞介素-8(IL-8)、白细胞介素-10(IL-10)和白细胞介素-13(IL-13)。

[0390] 在某些实施例中,连接子包含CH₂CH₃结构域、铰链结构域等。在一个实施例中,连接子包含IgG1、IgG4或IgD的CH₂和CH₃结构域。在一个实施例中,连接子包含CD8 α 或CD4铰链结构域。

[0391] 在具体实施例中,跨膜结构域是选自由以下组成的群组:T细胞受体的 α 或 β 链、CD3 δ 、CD3 ϵ 、CD3 γ 、CD3 ζ 、CD4、CD5、CD8 α 、CD9、CD16、CD22、CD27、CD28、CD33、CD37、CD45、CD64、CD80、CD86、CD134、CD137、CD152、CD154和PD-1。

[0392] 在具体实施例中,细胞内信号传导结构域是选自由以下组成的群组:含有ITAM的初级信号传导结构域和/或共刺激结构域。

[0393] 在具体实施例中,细胞内信号传导结构域是选自由以下组成的群组:FcR γ 、FcR β 、CD3 γ 、CD3 δ 、CD3 ϵ 、CD3 ζ 、CD22、CD79a、CD79b和CD66d。

[0394] 在具体实施例中,细胞内信号传导结构域是选自由以下组成的群组:TLR1、TLR2、

TLR3、TLR4、TLR5、TLR6、TLR7、TLR8、TLR9、TLR10、CARD11、CD2、CD7、CD27、CD28、CD30、CD40、CD54 (ICAM)、CD83、CD134 (OX40)、CD137 (4-1BB)、CD278 (ICOS)、DAP10、LAT、NKD2C、SLP76、TRIM和ZAP70。

[0395] 在一个实施例中,嵌合细胞因子受体包含一个或多个选自由CD28、CD137和CD134组成的群组的共刺激信号传导结构域,以及一个CD3 ζ 初级信号传导结构域。

[0396] 4. 球蛋白基因

[0397] 在具体实施例中,本文中所涵盖的被基因组编辑的造血细胞包含球蛋白基因或其抗镰状化变异体。在一个实施例中,通过在编码球蛋白基因或其抗镰状化变异体的供体修复模板存在下,将DSB引入一个或多个编码以下的基因中来工程改造造血细胞:有助于抑制 γ -球蛋白基因表达和HbF的基因,包括(但不限于)BCL11A基因座、KLF1基因座、SOX6基因座、GATA1基因座和LSD1基因座; β -球蛋白基因座的地中海型贫血等位基因;或 β -球蛋白基因座的镰状化等位基因。

[0398] 在具体实施例中,球蛋白或其抗镰状化变异体包括(但不限于) β -球蛋白、 δ -球蛋白、 γ -球蛋白、 β -球蛋白^{A-T87Q}、 β -球蛋白^{A-T87Q/K120E/K95E}或 β -球蛋白^{A-T87Q/G16D/E22A}。

[0399] 在具体实施例中,使用多个被工程改造的核酸酶诱导基因组中的多个目标位点中的DSB,并且将包含转基因的DNA供体修复模板引入细胞且通过同源重组插入至少一个目标位点中,所述转基因编码球蛋白或其抗镰状化变异体并且由对应于基因组中的多个目标位点的多个同源臂侧接。

[0400] 在具体实施例中,使用多个被工程改造的核酸酶诱导基因组中的多个目标位点中的DSB,并且将包含转基因的DNA供体修复模板引入细胞且通过同源重组插入两个或更多个目标位点中,所述转基因编码球蛋白或其抗镰状化变异体并且由对应于基因组中的多个目标位点的多个同源臂侧接。

[0401] D. 核酸酶

[0402] 通过用靶向多个基因座的被工程改造的核酸酶实现的多重基因组编辑来产生具体实施例中所涵盖的免疫效应细胞组合物。在具体实施例中,将被工程改造的核酸酶引入免疫效应细胞以诱导多个基因组目标位点中的DSB,所述基因组目标位点包括(但不限于)编码TCR信号传导组件(例如TCR α 基因的恒定区)、免疫系统检查点或免疫抑制性信号传导组件的基因。

[0403] 免疫系统检查点基因的说明性实例包括(但不限于):PD-1、LAG-3、TIM-3、CTLA-4、BTLA、TIGIT、VISTA和KIR。

[0404] 编码免疫抑制性信号传导组件的基因的说明性实例包括(但不限于):IL-10 α 、TGF β R1、TGF β R2、AHR、SGK1、TSC2、VHL、A2AR和CBLB。

[0405] 通过用靶向多个基因座的被工程改造的核酸酶实现的多重基因组编辑来产生具体实施例中所涵盖的造血细胞组合物。在具体实施例中,将被工程改造的核酸酶引入造血细胞中以诱导多个基因组目标位点中的DSB,所述基因组目标位点包括(但不限于)编码有助于抑制 γ -球蛋白基因表达和HbF的多肽的基因。

[0406] 抑制 γ -球蛋白基因表达和HbF的多肽的说明性实例包括(但不限于):BCL11A、KLF1、SOX6、GATA1和LSD1。

[0407] 具体实施例中所涵盖的被工程改造的核酸酶在目标序列中产生单链DNA切口或双

链DNA断裂物(DSB)。此外,可以通过使用两个产生单链切口的核酸酶(切口酶)来在目标DNA中获得DSB。每个切口酶使DNA的一条链裂解并且使用两个或更多个切口酶可以在目标DNA序列中产生双链断裂物(例如交错的双链断裂物)。在优选实施例中,核酸酶与供体修复模板组合使用,其中通过DSB处的同源重组将所述供体修复模板引入DNA断裂位点处的目标序列中。

[0408] 在本文中的具体实施例中所涵盖的适用于基因组编辑的被工程改造的核酸酶包含一个或多个DNA结合结构域和一个或多个DNA裂解结构域(例如一个或多个核酸内切酶和/或核酸外切酶结构域),以及任选地一个或多个本文中所涵盖的连接子。“被工程改造的核酸酶”是指包含一个或多个DNA结合结构域和一个或多个DNA裂解结构域的核酸酶,其中核酸酶被设计和/或修饰成结合与DNA裂解目标序列相邻的DNA结合目标序列。被工程改造的核酸酶可以由天然存在的核酸酶或预先被工程改造的核酸酶设计和/或修饰。具体实施例中所涵盖的被工程改造的核酸酶可以进一步包含一个或多个其它功能性结构域,例如末端处理酶的末端处理酶结构域,所述末端处理酶呈现5-3'核酸外切酶、5-3'碱性核酸外切酶、3-5'核酸外切酶(例如Trex2)、5'翻转核酸内切酶、解螺旋酶或模板独立性DNA聚合酶活性。

[0409] 可以被工程改造成结合和裂解目标序列的核酸酶的说明性实例包括(但不限于)导向核酸内切酶(兆核酸酶)、megaTAL、转录活化因子样效应物核酸酶(TALEN)、锌指核酸酶(ZFN)和丛集规则间杂型短回文重复序列(CRISPR)/Cas核酸酶系统。

[0410] 在具体实施例中,本文中所涵盖的核酸酶包含一个或多个异源DNA结合和裂解结构域(例如ZFN、TALEN、megaTAL)(Boissel等人,2014;Christian等人,2010)。在其它实施例中,可以改变天然存在的核酸酶的DNA结合结构域以结合于所选择的目标位点(例如已经被工程改造成结合于与同源结合位点不同的位点的兆核酸酶)。举例来说,兆核酸酶已经被设计成结合与其同源结合位点不同的目标位点(Boissel等人,2014)。在具体实施例中,核酸酶需要核酸序列以使核酸酶靶向目标位点(例如CRISPR/Cas)。

[0411] 1. 导向核酸内切酶/兆核酸酶

[0412] 在各种实施例中,将多个导向核酸内切酶或兆核酸酶引入细胞中并且被工程改造成结合于和在多个基因组目标位点中引入单链切口或双链断裂物(DSB),所述基因组目标位点包括(但不限于)编码TCR信号传导组件(例如TCR α 基因的恒定区)、免疫系统检查点或免疫抑制性信号传导组件的基因。“导向核酸内切酶”与“兆核酸酶”可以互换使用并且是指满足以下条件的天然存在的核酸酶或被工程改造的兆核酸酶:其识别具有12-45个碱基对的裂解位点,并且通常基于序列和结构基元分成五个家族:LAGLIDADG、GIY-YIG、HNH、His-Cys盒以及PD-(D/E)XK。

[0413] 被工程改造的HE在自然界中不存在并且可以通过重组型DNA技术或随机突变诱发来获得。被工程改造的HE可以通过在天然存在的HE或预先被工程改造的HE中产生一个或多个氨基酸变化(例如突变、取代、添加或删除一个或多个氨基酸)来获得。在具体实施例中,被工程改造的HE包含DNA识别界面的一个或多个氨基酸变化。

[0414] 具体实施例中所涵盖的被工程改造的HE可以进一步包含一个或多个连接子和/或其它功能性结构域,例如末端处理酶的末端处理酶结构域,其呈现5-3'核酸外切酶、5-3'碱性核酸外切酶、3-5'核酸外切酶(例如Trex2)、5'翻转核酸内切酶、解螺旋酶或模板独立性

DNA聚合酶活性。在具体实施例中,将被工程改造的HE与末端处理酶一起引入T细胞中,所述末端处理酶呈现5-3'核酸外切酶、5-3'碱性核酸外切酶、3-5'核酸外切酶(例如Trex2)、5'翻转核酸内切酶、解螺旋酶或模板独立性DNA聚合酶活性。HE和3'处理酶可以分开地引入,例如在不同载体或独立mRNA中,或一起引入,例如以融合蛋白质形式,或在通过病毒自裂解肽或IRES元件隔开的多顺反子构筑体中。

[0415] “DNA识别界面”是指与核酸目标碱基以及相邻残基相互作用的HE氨基酸残基。对于每个HE,DNA识别界面包含侧链-侧链和侧链-DNA接触的广泛网状结构,其中大部分必然是识别具体核酸目标序列所特有的。因此,对应于具体核酸序列的DNA识别界面的氨基酸序列显著变化并且是任何天然或被工程改造的HE的特征。作为非限制性实例,具体实施例中涵盖的被工程改造的HE可以通过构筑HE变异体的文库来衍生,其中改变位于天然HE(或预先被工程改造的HE)的DNA识别界面中的一个或多个氨基酸残基。可以使用裂解分析法,针对每个预测的TCR α 基因座目标位点的目标裂解活性来筛选文库(参见例如Jarjour等人,2009.《核酸研究(Nuc.Acids Res.)》37(20):6871-6880)。

[0416] LAGLIDADG导向核酸内切酶(LHE)是研究最充分的兆核酸酶家族,主要在古细菌中以及绿藻和真菌中的细胞器DNA中编码,并且显示最高的总体DNA识别特异性。LHE包含每条蛋白质链一个或两个LAGLIDADG催化基元并且分别以同源二聚体或单链单体形式起作用。LAGLIDADG蛋白质的结构研究鉴别出高度保守的核心结构(Stoddard 2005),以 $\alpha\beta\alpha\beta\alpha$ 折叠为特征,并且LAGLIDADG基元属于这一折叠的第一个螺旋。LHE的高效和特异性裂解表示得到新颖、高度特异性核酸内切酶的蛋白质骨架。然而,将LHE工程改造成结合并且裂解非天然或非典型目标位点需要选择适当的LHE骨架、检查目标基因座、选择推定的目标位点以及在目标位点中多达三分之二的碱基对位置处广泛地改变LHE以改变其DNA接触点和裂解特异性。

[0417] 可以用于设计被工程改造的LHE的LHE的说明性实例包括(但不限于)I-AabMI、I-AaeMI、I-AniI、I-ApaMI、I-CapIII、I-CapIV、I-CkaMI、I-CpaMI、I-CpaMII、I-CpaMIII、I-CpaMIV、I-CpaMV、I-CpaV、I-CraMI、I-EjeMI、I-GpeMI、I-GpiI、I-GzeMI、I-GzeMII、I-GzeMIII、I-HjeMI、I-LtrII、I-LtrI、I-LtrWI、I-MpeMI、I-MveMI、I-NcrII、I-NcrI、I-NcrMI、I-OheMI、I-OnuI、I-OsoMI、I-OsoMII、I-OsoMIII、I-OsoMIV、I-PanMI、I-PanMII、I-PanMIII、I-PnoMI、I-ScuMI、I-SmaMI、I-SscMI和I-Vdi141I。

[0418] 可以用于设计被工程改造的LHE的LHE的其它说明性实例包括(但不限于)I-CreI和I-SceI。

[0419] 在一个实施例中,被工程改造的LHE是选自由以下组成的群组:I-CpaMI、I-HjeMI、I-OnuI、I-PanMI和SmaMI。

[0420] 在一个实施例中,被工程改造的LHE是I-OnuI。

[0421] 在一个实施例中,由天然I-OnuI产生靶向人类TCR α 基因的被工程改造的I-OnuI LHE。在优选实施例中,由预先被工程改造的I-OnuI产生靶向人类TCR α 基因的被工程改造的I-OnuI LHE。

[0422] 在具体实施例中,被工程改造的I-OnuI LHE包含DNA识别界面中的一个或多个氨基酸取代。在具体实施例中,I-OnuI LHE与I-OnuI或I-OnuI的被工程改造的变异体的DNA识别界面具有至少70%、至少71%、至少72%、至少73%、至少74%、至少75%、至少76%、至少

77%、至少78%、至少79%、至少80%、至少81%、至少82%、至少83%、至少84%、至少85%、至少86%、至少87%、至少88%、至少89%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%或至少99%序列一致性(Taekuchi等人2011.《美国国家科学院院刊(Proc Natl Acad Sci U.S.A.)》2011年8月9日;108(32):13077-13082)。

[0423] 在一个实施例中,I-OnuI LHE与I-OnuI或I-OnuI的被工程改造的变异体的DNA识别界面具有至少70%,更佳至少80%,更佳至少85%,更佳至少90%,更佳至少95%,更佳至少97%,更佳至少99%序列一致性(Taekuchi等人2011.《美国国家科学院院刊》2011年8月9日;108(32):13077-13082)。

[0424] 在具体实施例中,被工程改造的I-OnuI LHE在DNA识别界面中,具体地说在位于I-OnuI的位置24-50、68到82、180到203和223到240的子结构域中包含一个或多个氨基酸取代或修饰。

[0425] 在一个实施例中,被工程改造的I-OnuI LHE在位于整个I-OnuI序列内的任何地方的其它位置处包含一个或多个氨基酸取代或修饰。可以被取代和/或修饰的残基包括(但不限于)直接或通过水分子接触核酸目标或与核酸主链或核苷酸碱基相互作用的氨基酸。在一个非限制性实例中,本文中所涵盖的被工程改造的I-OnuI LHE包含一个或多个取代和/或修饰,优选在至少一个选自由以下位置组成的位置群组的位置中包含至少5个,优选至少10个,优选至少15个,更优选至少20个,甚至更优选至少25个:I-OnuI的19、24、26、28、30、32、34、35、36、37、38、40、42、44、46、48、68、70、72、75、76 77、78、80、82、168、180、182、184、186、188、189、190、191、192、193、195、197、199、201、203、223、225、227、229、231、232、234、236、238、240。

[0426] 2.MegaTAL

[0427] 在各种实施例中,将多个megaTAL引入细胞中并且被工程改造成结合和在多个基因组目标位点中引入DSB,所述基因组目标位点包括(但不限于)编码抑制 γ -球蛋白基因表达的多肽、TCR信号传导组件(例如TCR α 基因的恒定区)、免疫系统检查点或免疫抑制性信号传导组件的基因。“megaTAL”是指被工程改造的核酸酶,其包含被工程改造的TALE DNA结合结构域和被工程改造的兆核酸酶,并且任选地包含一个或多个连接子和/或其它功能性结构域,例如末端处理酶的末端处理酶结构域,所述末端处理酶呈现5-3'核酸外切酶、5-3'碱性核酸外切酶、3-5'核酸外切酶(例如Trex2)、5'翻转核酸内切酶、解螺旋酶或模板独立性DNA聚合酶活性。在具体实施例中,可以用末端处理酶将megaTAL引入T细胞中,所述末端处理酶呈现5-3'核酸外切酶、5-3'碱性核酸外切酶、3-5'核酸外切酶(例如Trex2)、5'翻转核酸内切酶、解螺旋酶或模板独立性DNA聚合酶活性。megaTAL和3'处理酶可以分开地引入,例如在不同载体或独立mRNA中,或一起引入,例如以融合蛋白质形式,或在通过病毒自裂解肽或IRES元件隔开的多顺反子构筑体中。

[0428] “TALE DNA结合结构域”是转录活化因子样效应子(TALE或TAL效应子)的DNA结合部分,其模拟植物转录活化因子以操纵植物转录物组(参见例如Kay等人,2007.《科学(Science)》318:648-651)。具体实施例中涵盖的TALE DNA结合结构域是从头或由天然存在的TALE(例如来自野油菜黄单胞菌野油菜致病变种(*Xanthomonas campestris* pv.vesicatoria)、大豆黄单胞菌(*Xanthomonas gardneri*)、半透明黄单胞菌(*Xanthomonas*

translucens)、地毯草黄单胞菌(*Xanthomonas axonopodis*)、番茄黄单胞菌(*Xanthomonas perforans*)、苜蓿黄单胞菌(*Xanthomonas alfalfa*)、柑橘黄单胞菌(*Xanthomonas citri*)、辣椒-番茄黄单胞菌(*Xanthomonas euvesicatoria*)和水稻黄单胞菌(*Xanthomonas oryzae*)的AvrBs3,以及来自茄科雷尔氏菌(*Ralstonia solanacearum*)的brg11和hpx17)工程改造的。用于衍生和设计DNA结合结构域的TALE蛋白质的说明性实例公开于美国专利案第9,017,967号和其中引用的参考文献中,其全部以全文引用的方式并入本文中。

[0429] 在具体实施例中, megaTAL包含TALE DNA结合结构域,所述TALE DNA结合结构域包含涉及TALE DNA结合结构域与其对应目标DNA序列的结合的一个或多个重复单元。单一“重复单元”(又称为“重复”)的长度典型地是33-35个氨基酸。每个TALE DNA结合结构域重复单元典型地在所述重复的12和/或13号位置处包括1或2个DNA结合残基,其构成重复可变双残基(Repeat Variable Di-Residue, RVD)。已测定用于这些TALE DNA结合结构域的DNA识别的天然(典型)码,使得在12和13号位置的HD序列引起与胞嘧啶(C)的结合、NG结合于T、NI结合于A、NN结合于G或A并且NG结合于T。在某些实施例中,涵盖非典型(非典型性)RVD。

[0430] 适用于在具体实施例中涵盖的具体megaTAL中的非标准RVD的说明性实例包括(但不限于)用于鸟嘌呤(G)的识别的HH、KH、NH、NK、NQ、RH、RN、SS、NN、SN、KN;用于腺嘌呤(A)的识别的NI、KI、RI、HI、SI;用于胸腺嘧啶(T)的识别的NG、HG、KG、RG;用于胞嘧啶(C)的识别的RD、SD、HD、ND、KD、YG;用于A或G的识别的NV、HN;和用于A或T或G或C的识别的H*、HA、KA、N*、NA、NC、NS、RA、S,其中(*)意指不存在13号位置处的氨基酸。适用于具体实施例中涵盖的具体megaTAL中的RVD的其它说明性实例还包括以全文引用的方式并入本文中的美国专利案第8,614,092号中所公开的那些。

[0431] 在具体实施例中,本文中所涵盖的megaTAL包含有包含3到30个重复单元的TALE DNA结合结构域。在某些实施例中, megaTAL包含3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29或30个TALE DNA结合结构域重复单元。在优选实施例中,本文中所涵盖的megaTAL包含有包含5-16个重复单元,更优选7-15个重复单元,更优选9-12个(专利案不明显)重复单元并且更优选9、10或11个重复单元的TALE DNA结合结构域。

[0432] 在具体实施例中,本文中所涵盖的megaTAL包含有包含3到30个重复单元和其它单个截短的TALE重复单元的TALE DNA结合结构域,所述截短的TALE重复单元包含位于一组TALE重复单元的C端的20个氨基酸,即,其它C端半TALE DNA结合结构域重复单元(本文中其它地方公开的C-帽的氨基酸-20到-1,见下文)。因此,在具体实施例中,本文中所涵盖的megaTAL包含有包含3.5到30.5个重复单元的TALE DNA结合结构域。在某些实施例中, megaTAL包含3.5、4.5、5.5、6.5、7.5、8.5、9.5、10.5、11.5、12.5、13.5、14.5、15.5、16.5、17.5、18.5、19.5、20.5、21.5、22.5、23.5、24.5、25.5、26.5、27.5、28.5、29.5或30.5个TALE DNA结合结构域重复单元。在优选实施例中,本文中所涵盖的megaTAL包含TALE DNA结合结构域,其包含5.5-13.5个重复单元,更优选7.5-12.5个重复单元,更优选9.5-11.5个重复单元并且更优选9.5、10.5或11.5个重复单元。

[0433] 在具体实施例中, megaTAL包含“N端结构域(NTD)”多肽、一个或多个TALE重复结构域/单元、“C端结构域(CTD)”多肽和被工程改造的兆核酸酶。

[0434] 如本文中所使用,术语“N端结构域(NTD)”多肽是指侧接天然存在的TALE DNA结合

结构域的N端部分或片段的序列。NTD序列(如果存在)可以具有任何长度,只要TALE DNA结合结构域重复单元保留结合DNA的能力即可。在具体实施例中,NTD多肽包含在TALE DNA结合结构域N端的至少120个到至少140个或更多个氨基酸(0是最N端重复单元的氨基酸1)。在具体实施例中,NTD多肽包含在TALE DNA结合结构域N端的至少约120、121、122、123、124、125、126、127、128、129、130、131、132、133、134、135、136、137、138、139或至少140个氨基酸。在一个实施例中,本文中所涵盖的megaTAL包含在黄单胞菌属TALE蛋白质的至少约氨基酸+1到+122到至少约+1到+137的NTD多肽(0是最N端重复单元的氨基酸1)。在具体实施例中,NTD多肽包含在黄单胞菌属TALE蛋白质的TALE DNA结合结构域N端的至少约122、123、124、125、126、127、128、129、130、131、132、133、134、135、136或137个氨基酸。在一个实施例中,本文中所涵盖的megaTAL包含在雷尔氏菌属TALE蛋白质的至少氨基酸+1到+121的NTD多肽(0是最N端重复单元的氨基酸1)。在具体实施例中,NTD多肽包含在雷尔氏菌属TALE蛋白质的TALE DNA结合结构域N端的至少约121、122、123、124、125、126、127、128、129、130、131、132、133、134、135、136或137个氨基酸。

[0435] 如本文中所使用,术语“C端结构域(CTD)”多肽是指侧接天然存在的TALE DNA结合结构域的C端部分或片段的序列。CTD序列(如果存在)可以具有任何长度,只要TALE DNA结合结构域重复单元保留结合DNA的能力即可。在具体实施例中,CTD多肽包含在TALE DNA结合结构域的最后一个完整重复的C端的至少20个到至少85个或更多个氨基酸(前20个氨基酸是在最后一个C端完整重复单元C端的半重复单元)。在具体实施例中,CTD多肽包含在TALE DNA结合结构域的最后一个完整重复C端的至少约20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、443、44、45、46、47、48、49、50、51、52、53、54、55、56、57、58、59、60、61、62、63、64、65、66、67、68、69、70、71、72、73、74、75、76、77、78、79、80、81、82、83、84或至少85个氨基酸。在一个实施例中,本文中所涵盖的megaTAL包含在黄单胞菌属TALE蛋白质的至少约氨基酸-20到-1的CTD多肽(-20是在最后一个C端完整重复单元的C端的半重复单元的氨基酸1)。在具体实施例中,CTD多肽包含在黄单胞菌属TALE蛋白质的TALE DNA结合结构域的最后一个完整重复的C端的至少约20、19、18、17、16、15、14、13、12、11、10、9、8、7、6、5、4、3、2或1个氨基酸。在一个实施例中,本文中所涵盖的megaTAL包含在雷尔氏菌属TALE蛋白质的至少约氨基酸-20到-1的CTD多肽(-20是在最后一个C端完整重复单元的C端的半重复单元的氨基酸1)。在具体实施例中,CTD多肽包含在雷尔氏菌属TALE蛋白质的TALE DNA结合结构域的最后一个完整重复的C端的至少约20、19、18、17、16、15、14、13、12、11、10、9、8、7、6、5、4、3、2或1个氨基酸。

[0436] 在具体实施例中,本文中所涵盖的megaTAL包含融合多肽,所述融合多肽包含被工程改造成结合目标序列的TALE DNA结合结构域、被工程改造成结合和裂解目标序列的兆核酸酶以及任选地NTD和/或CTD多肽,任选地通过本文中其它地方涵盖的一个或多个连接子多肽彼此接合。不希望受任何具体理论约束,预期包含TALE DNA结合结构域和任选地NTD和/或CTD多肽的megaTAL与连接子多肽融合,所述连接子多肽又与被工程改造的兆核酸酶融合。因此,TALE DNA结合结构域结合与由兆核酸酶的DNA结合结构域结合的目标序列相距约1、2、3、4、5、6、7、8、9、10、11、12、13、14或15个核苷酸内的DNA目标序列。通过这种方式,本文中所涵盖的megaTAL增加基因组编辑的特异性和效率。

[0437] 在具体实施例中,本文中所涵盖的megaTAL包含一个或多个TALE DNA结合重复单

元和选自由以下组成的群组的被工程改造的LHE: I-AabMI、I-AaeMI、I-AniI、I-ApaMI、I-CapIII、I-CapIV、I-CkaMI、I-CpaMI、I-CpaMII、I-CpaMIII、I-CpaMIV、I-CpaMV、I-CpaV、I-CraMI、I-CreI、I-SceI、I-EjeMI、I-GpeMI、I-GpiI、I-GzeMI、I-GzeMII、I-GzeMIII、I-HjeMI、I-LtrII、I-LtrI、I-LtrWI、I-MpeMI、I-MveMI、I-NcrII、I-NcrI、I-NcrMI、I-OheMI、I-OnuI、I-OsoMI、I-OsoMII、I-OsoMIII、I-OsoMIV、I-PanMI、I-PanMII、I-PanMIII、I-PnoMI、I-ScuMI、I-SmaMI、I-SscMI和-Vdi141I,或优选是I-CpaMI、I-HjeMI、I-OnuI、I-PanMI和SmaMI,或更优选是I-OnuI。

[0438] 在具体实施例中,本文中所涵盖的megaTAL包含NTD、一个或多个TALE DNA结合重复单元、CTD和选自由以下组成的群组的被工程改造的LHE: I-AabMI、I-AaeMI、I-AniI、I-ApaMI、I-CapIII、I-CapIV、I-CkaMI、I-CpaMI、I-CpaMII、I-CpaMIII、I-CpaMIV、I-CpaMV、I-CpaV、I-CraMI、I-CreI、I-SceI、I-EjeMI、I-GpeMI、I-GpiI、I-GzeMI、I-GzeMII、I-GzeMIII、I-HjeMI、I-LtrII、I-LtrI、I-LtrWI、I-MpeMI、I-MveMI、I-NcrII、I-NcrI、I-NcrMI、I-OheMI、I-OnuI、I-OsoMI、I-OsoMII、I-OsoMIII、I-OsoMIV、I-PanMI、I-PanMII、I-PanMIII、I-PnoMI、I-ScuMI、I-SmaMI、I-SscMI和I-Vdi141I,或优选是I-CpaMI、I-HjeMI、I-OnuI、I-PanMI和SmaMI,或更优选是I-OnuI。

[0439] 在具体实施例中,本文中所涵盖的megaTAL包含NTD、约9.5到约11.5个TALE DNA结合重复单元和选自由以下组成的群组的被工程改造的I-OnuI LHE: I-AabMI、I-AaeMI、I-AniI、I-ApaMI、I-CapIII、I-CapIV、I-CkaMI、I-CpaMI、I-CpaMII、I-CpaMIII、I-CpaMIV、I-CpaMV、I-CpaV、I-CraMI、I-CreI、I-SceI、I-EjeMI、I-GpeMI、I-GpiI、I-GzeMI、I-GzeMII、I-GzeMIII、I-HjeMI、I-LtrII、I-LtrI、I-LtrWI、I-MpeMI、I-MveMI、I-NcrII、I-NcrI、I-NcrMI、I-OheMI、I-OnuI、I-OsoMI、I-OsoMII、I-OsoMIII、I-OsoMIV、I-PanMI、I-PanMII、I-PanMIII、I-PnoMI、I-ScuMI、I-SmaMI、I-SscMI和I-Vdi141I,或优选是I-CpaMI、I-HjeMI、I-OnuI、I-PanMI和SmaMI,或更优选是I-OnuI。

[0440] 在具体实施例中,本文中所涵盖的megaTAL包含具有约122个氨基酸到137个氨基酸的NTD;约9.5、约10.5或约11.5个结合重复单元;具有约20个氨基酸到约85个氨基酸的CTD;和选自由以下组成的群组的被工程改造的I-OnuI LHE: I-AabMI、I-AaeMI、I-AniI、I-ApaMI、I-CapIII、I-CapIV、I-CkaMI、I-CpaMI、I-CpaMII、I-CpaMIII、I-CpaMIV、I-CpaMV、I-CpaV、I-CraMI、I-CreI、I-SceI、I-EjeMI、I-GpeMI、I-GpiI、I-GzeMI、I-GzeMII、I-GzeMIII、I-HjeMI、I-LtrII、I-LtrI、I-LtrWI、I-MpeMI、I-MveMI、I-NcrII、I-NcrI、I-NcrMI、I-OheMI、I-OnuI、I-OsoMI、I-OsoMII、I-OsoMIII、I-OsoMIV、I-PanMI、I-PanMII、I-PanMIII、I-PnoMI、I-ScuMI、I-SmaMI、I-SscMI和I-Vdi141I,或优选是I-CpaMI、I-HjeMI、I-OnuI、I-PanMI和SmaMI,或更优选是I-OnuI。

[0441] 3. Talen

[0442] 在各种实施例中,将多个转录活化因子样效应物核酸酶(TALEN)引入细胞中并且被工程改造成结合于和在多个基因组目标位点中引入单链切口或双链断裂物(DSB),所述基因组目标位点包括(但不限于)编码抑制 γ -球蛋白基因表达的多肽、TCR信号传导组件(例如TCR α 基因的恒定区)、免疫系统检查点或免疫抑制性信号传导组件的基因。“TALEN”是指被工程改造的核酸酶,其包含本文中其它地方涵盖的被工程改造的TALE DNA结合结构域和核酸内切酶结构域(或其核酸内切酶半结构域),并且任选地包含一个或多个连接子和/

或其它功能性结构域,例如末端处理酶的末端处理酶结构域,所述末端处理酶呈现5-3'核酸外切酶、5-3'碱性核酸外切酶、3-5'核酸外切酶(例如Trex2)、5'翻转核酸内切酶、解螺旋酶或模板独立性DNA聚合酶活性。在具体实施例中,可以将TALEN与末端处理酶一起引入T细胞中,所述末端处理酶呈现5-3'核酸外切酶、5-3'碱性核酸外切酶、3-5'核酸外切酶(例如Trex2)、5'翻转核酸内切酶、解螺旋酶或模板独立性DNA聚合酶活性。TALEN和3'处理酶可以分开地引入,例如在不同载体或独立mRNA中,或一起引入,例如以融合蛋白质形式,或在通过病毒自裂解肽或IRES元件隔开的多顺反子构筑体中。

[0443] 在一个实施例中,用两个TALEN实现靶向双链裂解,所述两个TALEN各自包含可以用于重建催化活性裂解结构域的核酸内切酶半结构域。在另一个实施例中,使用单一多肽实现靶向双链裂解,所述单一多肽包含TALE DNA结合结构域和两个核酸内切酶半结构域。

[0444] 具体实施例中所涵盖的TALEN包含NTD、包含约3到30个重复单元(例如约3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29或30个重复单元)的TALE DNA结合结构域和核酸内切酶结构域或半结构域。

[0445] 具体实施例中所涵盖的TALEN包含NTD、包含约3.5到30.5个重复单元(例如约3.5、4.5、5.5、6.5、7.5、8.5、9.5、10.5、11.5、12.5、13.5、14.5、15.5、16.5、17.5、18.5、19.5、20.5、21.5、22.5、23.5、24.5、25.5、26.5、27.5、28.5、29.5或30.5个重复单元)的TALE DNA结合结构域、CTD和核酸内切酶结构域或半结构域。

[0446] 具体实施例中所涵盖的TALEN包含如本文中其它地方所公开的具有约121个氨基酸到约137个氨基酸的NTD;包含约9.5到约11.5个重复单元(即,约9.5、约10.5或约11.5个重复单元)的TALE DNA结合结构域;具有约20个氨基酸到约85个氨基酸的CTD;和核酸内切酶结构域或半结构域。

[0447] 在具体实施例中,TALEN包含限制性核酸内切酶型核酸内切酶结构域。限制核酸内切酶(限制酶)存在于多种物种中,并且能够序列特异性结合于DNA(在识别位点处)和使结合位点处或附近的DNA裂解。某些限制酶(例如IIS型)使远离识别位点的位点处的DNA裂解并且具有可以分开的结合和核酸内切酶结构域。在一个实施例中,TALEN包含来自至少一个IIS型限制酶的核酸内切酶结构域(或核酸内切酶半结构域)和一个或多个本文中其它地方涵盖的TALE DNA结合结构域。

[0448] 适用于具体实施例中所涵盖的TALEN中的IIS型限制性核酸内切酶结构域的说明性实例包括在“rebase.neb.com/cgi-bin/sublist?S”公开的至少1633种IIS型限制核酸内切酶的核酸内切酶结构域。

[0449] 适用于具体实施例中所涵盖的TALEN中的IIS型限制性核酸内切酶结构域的其他说明性实例包括选自由以下组成的群组的核酸内切酶的结构域:Aar I、Ace III、Aci I、Alo I、Alw26 I、Bae I、Bbr7 I、Bbv I、Bbv II、BbvC I、Bcc I、Bce83 I、BceA I、Bcef I、Bcg I、BciV I、Bfi I、Bin I、Bmg I、Bpu10 I、BsaX I、Bsb I、BscA I、BscG I、BseR I、BseY I、Bsi I、Bsm I、BsmA I、BsmF I、Bsp24 I、BspG I、BspM I、BspNC I、Bsr I、BsrB I、BsrD I、BstF5 I、Btr I、Bts I、Cdi I、CjeP I、Drd II、EarI、Eci I、Eco31 I、Eco57I、Eco57M I、Esp3 I、Fau I、Fin I、Fok I、Gdi II、Gsu I、Hga I、Hin4 II、Hph I、Ksp632I、Mbo II、Mly I、Mme I、Mnl I、Pfl1108、I Ple I、Ppi I、Psr I、RleA I、Sap I、SfaN I、Sim I、SspD5 I、Sth132 I、Sts I、TspDT I、TspGW I、Tth111 II、UbaP I、Bsa I和BsmB I。

[0450] 在一个实施例中,本文中所涵盖的TALEN包含Fok I IIS型限制性核酸内切酶的核酸内切酶结构域。

[0451] 在一个实施例中,本文中所涵盖的TALEN包含TALE DNA结合结构域和来自至少一个IIS型限制性核酸内切酶的核酸内切酶半结构域以增强裂解特异性,任选地,其中所述核酸内切酶半结构域包含一个或多个可以最小化或防止同源二聚的氨基酸取代或修饰。

[0452] 具体实施例中涵盖的适用于具体实施例中的裂解半结构域的说明性实例包括美国专利公开案第20050064474号、第20060188987号、第20080131962号、第20090311787号、第20090305346号、第20110014616号和第20110201055号中所公开的裂解半结构域,其各自以全文引用的方式并入本文中。

[0453] 4. 锌指核酸酶

[0454] 在各种实施例中,将多个锌指核酸酶(ZFN)引入细胞中并且被工程改造成结合于和在多个基因组目标位点中引入单链切口或双链断裂物(DSB),所述基因组目标位点包括(但不限于)编码抑制 γ -球蛋白基因表达的多肽、TCR信号传导组件(例如TCR α 基因的恒定区)、免疫系统检查点或免疫抑制性信号传导组件的基因。“ZFN”是指被工程改造的核酸酶,其包含一个或多个锌指DNA结合结构域和核酸内切酶结构域(或其核酸内切酶半结构域),并且任选地包含一个或多个连接子和/或其它功能性结构域,例如末端处理酶的末端处理酶结构域,所述末端处理酶呈现5-3'核酸外切酶、5-3'碱性核酸外切酶、3-5'核酸外切酶(例如Trex2)、5'翻转核酸内切酶、解螺旋酶或模板独立性DNA聚合酶活性。在具体实施例中,可以将ZFN与末端处理酶一起引入T细胞中,所述末端处理酶呈现5-3'核酸外切酶、5-3'碱性核酸外切酶、3-5'核酸内切酶(例如Trex2)、5'翻转核酸内切酶、解螺旋酶或模板独立性DNA聚合酶活性。ZFN和3'处理酶可以分开地引入,例如在不同载体或独立mRNA中,或一起引入,例如以融合蛋白质形式,或在通过病毒自裂解肽或IRES元件隔开的多顺反子构筑体中。

[0455] 在一个实施例中,使用两个ZFN实现靶向双链裂解,所述两个ZFN各自包含可以用于重建催化活性裂解结构域的核酸内切酶半结构域。在另一个实施例中,用单一多肽实现靶向双链裂解,所述单一多肽包含一个或多个锌指DNA结合结构域和两个核酸内切酶半结构域。

[0456] 在一个实施例中,ZNF包含本文中其它地方涵盖的TALE DNA结合结构域、锌指DNA结合结构域和本文中其它地方涵盖的核酸内切酶结构域(或核酸内切酶半结构域)。

[0457] 在一个实施例中,ZNF包含锌指DNA结合结构域和本文中其它地方涵盖的兆核酸酶。

[0458] 在具体实施例中,ZFN包含锌指DNA结合结构域,其具有一、二、三、四、五、六、七或八个或更多个锌指基元和一个核酸内切酶结构域(或核酸内切酶半结构域)。典型地,单个锌指基元的长度是约30个氨基酸。锌指基元包括标准C₂H₂锌指和非标准锌指,例如C₃H锌指和C₄锌指。

[0459] 锌指结合结构域可以被工程改造成结合任何DNA序列。已经鉴别用于既定3bp DNA目标序列的候选锌指DNA结合结构域并且已经设计模块组装策略以用于将多个结构域连接成靶向相应复合DNA目标序列的多指肽。也可以使用所属领域中已知的其它合适的方法设计和构筑编码锌指DNA结合结构域的核酸,例如噬菌体呈现、随机突变诱发、组合文库、计算

机/合理设计、亲和力选择、PCR、从cDNA或基因组文库克隆、合成构筑等(参见例如美国专利案第5,786,538号;Wu等人,《美国科学院院报(PNAS)》92:344-348(1995);Jamieson等人,《生物化学(Biochemistry)》33:5689-5695(1994);Rebar和Pabo,《科学》263:671-673(1994);Choo和Klug,《美国科学院院报》91:11163-11167(1994);Choo和Klug,《美国科学院院报》91:11168-11172(1994);Desjarlais和Berg,《美国科学院院报》90:2256-2260(1993);Desjarlais和Berg,《美国科学院院报》89:7345-7349(1992);Pomerantz等人,《科学》267:93-96(1995);Pomerantz等人,《美国科学院院报》92:9752-9756(1995);Liu等人,《美国科学院院报》94:5525-5530(1997);Griesman和Pabo,《科学》275:657-661(1997);Desjarlais和Berg,《美国科学院院报》91:11-99-11103(1994))。

[0460] 单独的锌指基元结合于三个或四个核苷酸序列。锌指结合结构域被工程改造以结合的序列(例如目标序列)的长度将决定被工程改造的锌指结合结构域中锌指基元的数目。举例来说,对于其中锌指基元不结合于重叠子位点的ZFN,由两指结合结构域结合六核苷酸目标序列;由三指结合结构域结合九核苷酸目标序列等。在具体实施例中,目标位点中单独的锌指基元的DNA结合位点无需是相邻的,但可以由一个或若干个核苷酸隔开,取决于多指结合结构域中锌指基元之间的连接子序列的长度和性质。

[0461] 在具体实施例中,本文中所涵盖的ZNF包含锌指DNA结合结构域,其包含二、三、四、五、六、七或八个或更多个锌指基元,以及来自至少一个IIS型限制酶的核酸内切酶结构域或半结构域和一个或多个本文中其它地方涵盖的TALE DNA结合结构域。

[0462] 在具体实施例中,本文中所涵盖的ZNF包含锌指DNA结合结构域,其包含三、四、五、六、七或八个或更多个锌指基元,以及来自至少一个选自由以下组成的群组的IIS型限制酶的核酸内切酶结构域或半结构域:Aar I、Ace III、Aci I、Alo I、Alw26 I、Bae I、Bbr7 I、Bbv I、Bbv II、BbvC I、Bcc I、Bce83 I、BceA I、Bcef I、Bcg I、BciV I、Bfi I、Bin I、Bmg I、Bpu10 I、BsaX I、Bsb I、BscAI、BscG I、BseR I、BseY I、Bsi I、Bsm I、BsmA I、BsmF I、Bsp24 I、BspG I、BspM I、BspNC I、Bsr I、BsrB I、BsrD I、BstF5 I、Btr I、Bts I、Cdi I、CjeP I、Drd II、EarI、Eci I、Eco31 I、Eco57 I、Eco57M I、Esp3 I、Fau I、Fin I、Fok I、Gdi II、Gsu I、Hga I、Hin4 II、Hph I、Ksp632 I、Mbo II、Mly I、Mme I、Mnl I、Pfl11108、I Ple I、Ppi I、Psr I、RleA I、Sap I、SfaN I、Sim I、SspD5I、Sth132 I、Sts I、TspDT I、TspGW I、Tth111 II、UbaP I、Bsa I和BsmB I。

[0463] 在具体实施例中,本文中所涵盖的ZNF包含锌指DNA结合结构域,其包含三、四、五、六、七或八个或更多个锌指基元,和来自Fok I IIS型限制性核酸内切酶的核酸内切酶结构域或半结构域。

[0464] 在一个实施例中,本文中所涵盖的ZFN包含锌指DNA结合结构域和来自至少一个IIS型限制性核酸内切酶的核酸内切酶半结构域以增强裂解特异性,任选地,其中所述核酸内切酶半结构域包含一个或多个可以最小化或防止同源二聚的氨基酸取代或修饰。

[0465] 5. CRISPR/Cas核酸酶系统

[0466] 在各种实施例中,将CRISPR(丛集规则间杂短回文重复序列)/Cas(CRISPR相关)核酸酶系统引入细胞中并且被工程改造成结合于和在多个基因组目标位点中引入单链切口或双链断裂物(DSB),所述基因组目标位点包括(但不限于)编码抑制 γ -球蛋白基因表达的多肽、TCR信号传导组件(例如TCR α 基因的恒定区)、免疫系统检查点或免疫抑制性信号传导

组件的基因。CRISPR/Cas核酸酶系统是基于可以用于哺乳动物基因组工程改造的细菌系统的最近被工程改造的核酸酶系统。参见例如Jinek等人(2012)《科学》337:816-821; Cong等人(2013)《科学》339:819-823; Mali等人(2013)《科学》339:823-826; Qi等人(2013)《细胞(Cell)》152:1173-1183; Jinek等人(2013), 《eLife》2:e00471; David Segal(2013)《eLife》2:e00563; Ran等人(2013)《自然实验手册(Nature Protocols)》8(11):2281-2308; Zetsche等人(2015)《细胞》163(3):759-771, 其各自以全文引用的方式并入本文中。

[0467] 在一个实施例中, CRISPR/Cas核酸酶系统包含Cas核酸酶和一个或多个将Cas核酸酶募集到目标位点的RNA, 例如转录活化crRNA(tracrRNA)和CRISPR RNA(crRNA), 或单一向导RNA(sgRNA)。crRNA和tracrRNA可以被工程改造成一个本文中称为“单一向导RNA”或“sgRNA”的聚核苷酸序列。

[0468] 在一个实施例中, 对于位于编码抑制 γ -球蛋白基因表达的多肽、TCR信号传导组件(例如TCR α 基因的恒定区)、免疫系统检查点或免疫抑制性信号传导组件的基因内的原间隔子目标序列的位点特异性DNA识别和位点特异性裂解, Cas核酸酶被工程改造成双链DNA核酸内切酶或切口酶或催化死亡Cas, 并且与和tracrRNA或sgRNA形成目标复合物。原间隔子基元邻接短原间隔子相邻基元(PAM), 其在募集Cas/RNA复合物中起作用。Cas多肽识别对Cas多肽具有特异性的PAM基元。相应地, CRISPR/Cas系统可以用于靶向和裂解由对具体Cas多肽具有特异性的具体3'PAM序列侧接的双链聚核苷酸序列的任一条链或两条链。可以使用生物信息学或使用实验方法鉴别PAM。Esvelt等人, 2013, 《自然方法(Nature Methods.)》10(11):1116-1121, 其以全文引用的方式并入本文中。

[0469] 在一个实施例中, Cas核酸酶包含一个或多个异源DNA结合结构域, 例如TALE DNA结合结构域或锌指DNA结合结构域。Cas核酸酶与TALE或锌指DNA结合结构域的融合可以增加DNA裂解效率和特异性。在具体实施例中, Cas核酸酶任选地包含一个或多个连接子和/或其它功能性结构域, 例如末端处理酶的末端处理酶结构域, 所述末端处理酶呈现5-3'核酸外切酶、5-3'碱性核酸外切酶、3-5'核酸外切酶(例如Trex2)、5'翻转核酸内切酶、解螺旋酶或模板独立性DNA聚合酶活性。在具体实施例中, 可以将Cas核酸酶与末端处理酶一起引入T细胞中, 所述末端处理酶呈现5-3'核酸外切酶、5-3'碱性核酸外切酶、3-5'核酸外切酶(例如Trex2)、5'翻转核酸内切酶、解螺旋酶或模板独立性DNA聚合酶活性。Cas核酸酶和3'处理酶可以分开地引入, 例如在不同载体或独立mRNA中, 或一起引入, 例如以融合蛋白质形式, 或在通过病毒自裂解肽或IRES元件隔开的多顺反子构筑体中。

[0470] 在各种实施例中, Cas核酸酶是Cas9或Cpf1。

[0471] 具体实施例中所涵盖的适用于具体实施例中的Cas9多肽的说明性实例可以从细菌物种获得, 所述细菌物种包括(但不限于): 屎肠球菌(*Enterococcus faecium*)、意大利肠球菌(*Enterococcus italicus*)、无害李斯特菌(*Listeria innocua*)、单核细胞增多性李斯特菌(*Listeria monocytogenes*)、西尔李斯特菌(*Listeria seeligeri*)、伊氏李斯特菌(*Listeria ivanovii*)、无乳链球菌(*Streptococcus agalactiae*)、咽峡炎链球菌(*Streptococcus anginosus*)、牛链球菌(*Streptococcus bovis*)、停乳链球菌(*Streptococcus dysgalactiae*)、马链球菌(*Streptococcus equinus*)、解没食子酸链球菌(*Streptococcus gallolyticus*)、猕猴链球菌(*Streptococcus macacae*)、变形链球菌(*Streptococcus mutans*)、假肺炎链球菌(*Streptococcus pseudoporcinus*)、酿脓链球菌

(*Streptococcus pyogenes*)、嗜热链球菌(*Streptococcus thermophilus*)、格氏链球菌(*Streptococcus gordonii*)、婴儿链球菌(*Streptococcus infantarius*)、马氏链球菌(*Streptococcus macedonicus*)、轻型链球菌(*Streptococcus mitis*)、巴氏链球菌(*Streptococcus pasteurianus*)、猪链球菌(*Streptococcus suis*)、前庭链球菌(*Streptococcus vestibularis*)、血链球菌(*Streptococcus sanguinis*)、汗毛链球菌(*Streptococcus downei*)、杆菌奈瑟菌(*Neisseria bacilliformis*)、灰色奈瑟菌(*Neisseria cinerea*)、浅黄奈瑟菌(*Neisseria flavescens*)、丙氨奈瑟菌(*Neisseria lactamica*)、脑膜炎奈瑟菌(*Neisseria meningitidis*)、淡黄色奈瑟菌(*Neisseria subflava*)、短乳杆菌(*Lactobacillus brevis*)、布氏乳杆菌(*Lactobacillus buchneri*)、干酪乳杆菌(*Lactobacillus casei*)、副干酪乳杆菌(*Lactobacillus paracasei*)、发酵乳杆菌(*Lactobacillus fermentum*)、加氏乳杆菌(*Lactobacillus gasseri*)、詹氏乳酸杆菌(*Lactobacillus jensenii*)、约氏乳杆菌(*Lactobacillus johnsonii*)、鼠李糖乳杆菌(*Lactobacillus rhamnosus*)、瘤胃乳杆菌(*Lactobacillus ruminis*)、唾液乳杆菌(*Lactobacillus salivarius*)、旧金山乳杆菌(*Lactobacillus sanfranciscensis*)、拥挤棒杆菌(*Corynebacterium accolens*)、白喉杆菌(*Corynebacterium diphtheriae*)、马氏棒状杆菌(*Corynebacterium matruchotii*)、空肠弯曲杆菌(*Campylobacter jejuni*)、产气荚膜梭菌(*Clostridium perfringens*)、文氏密螺旋体(*Treponema vincentii*)、溃蚀齿密螺旋体(*Treponema phagedenis*)和齿垢密螺旋体(*Treponema denticola*)。

[0472] 具体实施例中所涵盖的适用于具体实施例中的Cpf1多肽的说明性实例可以从细菌物种获得,所述细菌物种尤其包括(但不限于):弗朗西斯氏菌属(*Francisella* spp.)、氨基酸球菌属(*Acidaminococcus* spp.)、普雷沃氏菌属(*Prevotella* spp.)、毛螺菌属(*Lachnospiraceae* spp.)。

[0473] Cas9直系同源物的保守区包括中心HNH核酸内切酶结构域和裂解RuvC/RNase H结构域。Cpf1直系同源物具有RuvC/RNase H结构域,但不具有可辨别的HNH结构域。HNH和RuvC样结构域各自负责裂解双链DNA目标序列的一条链。Cas9核酸酶多肽的HNH结构域裂解与tracrRNA:crRNA或sgRNA互补的DNA链。Cas9核酸酶的RuvC样结构域裂解不与tracrRNA:crRNA或sgRNA互补的DNA链。预测Cpf1充当二聚体,其中Cpf1的每个RuvC样结构域裂解目标位点的互补或非互补链。在具体实施例中,涵盖Cas9核酸酶变异体(例如Cas9切口酶),其包含HNH或RuvC样核酸内切酶结构域中的一个或多个可以降低或消除变异型结构域的核酸酶活性的氨基酸添加、缺失、突变或取代。

[0474] 可以降低或消除结构域中的核酸酶活性的Cas9 HNH突变的说明性实例包括(但不限于):化脓性链球菌(*S. pyogenes*) (D10A);嗜热链球菌(*S. thermophilis*) (D9A);齿垢密螺旋体(*T. denticola*) (D13A);和脑膜炎奈瑟菌(*N. meningitidis*) (D16A)。

[0475] 可以降低或消除结构域中的核酸酶活性的Cas9 RuvC样结构域突变的说明性实例包括(但不限于):化脓性链球菌(D839A、H840A或N863A);嗜热链球菌(D598A、H599A或N622A);齿垢密螺旋体(D878A、H879A或N902A);和脑膜炎奈瑟菌(D587A、H588A或N611A)。

[0476] E. 多肽

[0477] 本文中涵盖各种多肽,包括(但不限于)免疫效能增强子、免疫抑制信号减弱子、被工程改造的抗原受体和被工程改造的核酸酶。除非有相反说明,否则“多肽”、“多肽片段”、

“肽”和“蛋白质”可以互换使用,并且根据常规即,作为氨基酸序列。在一个实施例中,“多肽”包括融合多肽和其它变异体。多肽可以使用多种众所周知的重组型和/或合成技术中的任一种制备。多肽不限于特定长度,例如其可以包含全长蛋白质序列、全长蛋白质的片段或融合蛋白质,并且可以包括多肽的翻译后修饰,例如糖基化、乙酰化、磷酸化等,以及所属领域中已知的天然存在的和非天然存在的其它修饰。

[0478] 具体实施例中所涵盖的多肽的说明性实例包括(但不限于)导向核酸内切酶变异体、megaTAL、球蛋白、抗镰状化球蛋白、BITE、细胞因子、趋化因子、细胞毒素和细胞因子受体、翻转受体、免疫抑制信号减弱子、CAR、DARIC、TCR和 ζ 因子。

[0479] 多肽包括“多肽变异体”。多肽变异体与天然存在的多肽的不同之处可在于一个或多个氨基酸取代、缺失、添加和/或插入。这类变异体可以是天然存在的或可以例如通过修饰上述多肽序列中的一个或多个氨基酸,以合成方式产生。举例来说,在具体实施例中,可能需要通过将一个或多个取代、缺失、添加和/或插入引入多肽中来改良多肽的生物学特性。在具体实施例中,多肽包括与本文中所涵盖的参考序列中的任一个具有至少约65%、70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%氨基酸一致性的多肽,典型地其中变异体保持参考序列的至少一种生物活性。

[0480] 多肽变异体包括生物活性“多肽片段”。生物活性多肽片段的说明性实例包括DNA结合结构域、核酸酶结构域等。如本文中所使用,术语“生物活性片段”或“最小生物活性片段”是指保留至少100%、至少90%、至少80%、至少70%、至少60%、至少50%、至少40%、至少30%、至少20%、至少10%、或至少5%的天然存在的多肽活性的多肽片段。在优选实施例中,所述生物活性是针对目标序列的结合亲和力和/或裂解活性。在某些实施例中,多肽片段可以包含长度是至少5个到约1700个氨基酸的氨基酸链。应了解,在某些实施例中,片段的长度是至少5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、50、55、60、65、70、75、80、85、90、95、100、110、150、200、250、300、350、400、450、500、550、600、650、700、750、800、850、900、950、1000、1100、1200、1300、1400、1500、1600、1700个或更多个氨基酸。在具体实施例中,多肽包含导向核酸内切酶变异体的生物活性片段。在具体实施例中,本文所述的多肽可以包含一个或多个表示为“X”的氨基酸。“X”如果存在于氨基酸SEQ ID NO中,则是指任何氨基酸。一个或多个“X”残基可以存在于本文中所涵盖的具体SEQ ID NO中所述氨基酸序列的N端和C端。如果“X”氨基酸不存在,则SEQ ID NO中所述的其余氨基酸序列可以视为生物活性片段。

[0481] 如上文所述,多肽可以不同方式改变,包括氨基酸取代、缺失、截短和插入。用于这类操控的方法通常是所属领域中已知的。举例来说,参考多肽的氨基酸序列变异体可以通过DNA中的突变来制备。用于突变诱发和核苷酸序列改变的方法是所属领域中众所周知的。参见例如Kunkel (1985,《美国国家科学院院刊》82:488-492);Kunkel等人,(1987,《酶学方法(Methods in Enzymol)》,154:367-382);美国专利案第4,873,192号;Watson, J.D.等人(《基因分子生物学(Molecular Biology of the Gene)》,第四版,Benjamin/Cummings, Menlo Park, Calif., 1987)以及其中引用的参考文献。关于不影响相关蛋白质的生物活性的合适的氨基酸取代的指导可见于Dayhoff等人,(1978)《蛋白序列和结构图集(Atlas of

Protein Sequence and Structure)》(《国家生物医学研究基础(Natl.Biomed.Res.Found.)》,Washington,D.C.)的模型中。

[0482] 在某些实施例中,变异体将含有一个或多个保守性取代。“保守性取代”是氨基酸被另一个具有类似特性的氨基酸取代,使得肽化学方法领域的技术人员将预期多肽的二级结构和亲水性质基本上不变。可以对具体实施例中所涵盖的聚核苷酸和多肽的结构进行修饰,多肽包括具有至少约的多肽且仍获得编码具有所需特征的变异体或衍生物多肽的功能分子。当需要改变多肽的氨基酸序列以产生等效或甚至改良的变异型多肽时,所属领域的技术人员例如可以改变例如根据表1的编码DNA序列的密码子中的一种或多种。

[0483] 表1-氨基酸密码子

[0484]

氨基酸	单字母码	三字码	密码子					
丙氨酸	A	Ala	GCA	GCC	GCG	GCU		
半胱氨酸	C	Cys	UGC	UGU				
天冬氨酸	D	Asp	GAC	GAU				
谷氨酸	E	Glu	GAA	GAG				
苯丙氨酸	F	Phe	UUC	UUU				
甘氨酸	G	Gly	GGA	GGC	GGG	GGU		
组氨酸	H	His	CAC	CAU				
异白氨酸	I	Iso	AUA	AUC	AUU			
赖氨酸	K	Lys	AAA	AAG				
白氨酸	L	Leu	UUA	UUG	CUA	CUC	CUG	CUU
甲硫氨酸	M	Met	AUG					
天冬酰胺	N	Asn	AAC	AAU				
脯氨酸	P	Pro	CCA	CCC	CCG	CCU		
谷氨酰胺	Q	Gln	CAA	CAG				
精氨酸	R	Arg	AGA	AGG	CGA	CGC	CGG	CGU
丝氨酸	S	Ser	AGC	AGU	UCA	UCC	UCG	UCU
苏氨酸	T	Thr	ACA	ACC	ACG	ACU		
缬氨酸	V	Val	GUA	GUC	GUG	GUU		
色氨酸	W	Trp	UGG					
酪氨酸	Y	Tyr	UAC	UAU				

[0485] 可以使用所属领域中众所周知的计算机程序,如DNASTAR、DNA Strider、Geneious、Mac Vector或Vector NTI软件发现有关确定可以被取代、插入或缺失而不消除生物活性的氨基酸残基的指导。优选地,本文中所公开的蛋白质变异体中的氨基酸变化是保守性氨基酸变化,即,带类似电荷或不带电氨基酸的取代。保守性氨基酸变化涉及其侧链相关的氨基酸家族中的一个的取代。天然存在的氨基酸通常被分成四个家族:酸性氨基酸(天冬氨酸、谷氨酸)、碱性氨基酸(赖氨酸、精氨酸、组氨酸)、非极性氨基酸(丙氨酸、缬氨酸、白氨酸、异白氨酸、脯氨酸、苯丙氨酸、甲硫氨酸、色氨酸)和不带电极性氨基酸(甘氨酸、天冬酰胺、谷氨酰胺、半胱氨酸、丝氨酸、苏氨酸、酪氨酸)。苯丙氨酸、色氨酸和酪氨酸有时

共同归类为芳香族氨基酸。在肽或蛋白质中,氨基酸的合适的保守性取代是所属领域的技术人员已知的并且通常在不改变所得分子的生物活性下进行。所属领域的技术人员认识到,一般来说,多肽的非必需区中的单氨基酸取代不显著改变生物活性(参见例如Watson等人《基因分子生物学》,第4版,1987,The Benjamin/Cummings Pub.Co.,第224页)。

[0486] 在需要表达两个或更多个多肽的一个实施例中,编码所述多肽的聚核苷酸序列可以由如本文中其它地方所公开的IRES序列隔开。

[0487] 具体实施例中所涵盖的多肽包括融合多肽。在具体实施例中,提供融合多肽和编码融合多肽的聚核苷酸。融合多肽和融合蛋白质是指具有至少二、三、四、五、六、七、八、九或十个多肽区段的多肽。

[0488] 在另一个实施例中,两个或更多个多肽可以表达为融合蛋白质形式,所述融合蛋白质包含一个或多个如本文中其它地方所公开的自裂解多肽序列。

[0489] 在一个实施例中,本文中所涵盖的融合蛋白质包含一个或多个DNA结合结构域和一个或多个核酸酶,以及一个或多个连接子和/或自裂解多肽。

[0490] 在一个实施例中,本文中所涵盖的融合蛋白质包含核酸酶变异体;连接子或自裂解肽;以及末端处理酶,包括(但不限于)5'-3'核酸外切酶、5'-3'碱性核酸外切酶和3'-5'核酸外切酶(例如Trex2)。

[0491] 融合多肽可以包含一个或多个多肽结构域或区段,包括(但不限于)信号肽、细胞穿透肽结构域(CPP)、DNA结合结构域、核酸酶结构域等、表位标签(例如麦芽糖结合蛋白("MBP")、谷胱甘肽S转移酶(GST)、HIS6、MYC、FLAG、V5、VSV-G以及HA)、多肽连接子以及多肽裂解信号。融合多肽通常是C端连接至N端,不过其还可以是C端连接至C端、N端连接至N端、或N端连接至C端。在具体实施例中,融合蛋白质的多肽可以遵循任何次序。融合多肽或融合蛋白质还可以包括保守修饰变异体、多态变异体、等位基因、突变体、子序列和种间同系物,只要保留融合多肽的所需活性即可。融合多肽可以通过化学合成方法或通过两个部分之间的化学连接制得或通常可以使用其它标准技术制备。构成融合多肽的连接DNA序列可操作地连接到如本文中其它地方所公开的合适的转录或翻译控制元件。

[0492] 融合多肽可以任选地包含连接子,所述连接子可以用于连接一个或多个多肽或多肽内的结构域。肽连接子序列可以用于将任何两个或更多个多肽组件隔开一定距离,该距离足以确保每一个多肽折叠成其适当二级结构和三级结构,从而允许多肽结构域发挥其所需功能。这类肽连接子序列是使用所属领域中的标准技术并入融合多肽中。合适的肽连接子序列可以基于以下因素选择:(1)其呈现柔性延长构形的能力;(2)其无法呈现能与第一多肽和第二多肽上的功能性表位相互作用的二级结构;以及(3)缺乏可能与多肽功能性表位反应的疏水性或带电残基。优选的肽连接子序列含有Gly、Asn和Ser残基。其它接近的中性氨基酸,如Thr和Ala,也可以用于连接子序列中。可以有效地用作连接子的氨基酸序列包括Maratea等人,《基因(Gene)》40:39-46,1985;Murphy等人,《美国国家科学院院刊》83:8258-8262,1986;美国专利案第4,935,233号和美国专利案第4,751,180号中所公开的那些。当具体融合多肽区段含有可以用于隔开功能性结构域并且防止空间干扰的非必需N端氨基酸区域时,不需要连接子序列。优选的连接子典型地是作为重组型融合蛋白质的一部分合成的柔性氨基酸子序列。连接子多肽可以是长度在1个与200个氨基酸之间、长度在1个与100个氨基酸之间或长度在1个与50个氨基酸之间,包括其间的所有整数值。

[0493] 例示性连接子包括(但不限于)以下氨基酸序列:甘氨酸聚合物($(G)_n$);甘氨酸-丝氨酸聚合物($(G_{1-5}S_{1-5})_n$),其中n是至少一、二、三、四或五的整数;甘氨酸-丙氨酸聚合物;丙氨酸-丝氨酸聚合物;GGG (SEQ ID NO:3);DGGGS (SEQ ID NO:4);TGEKP (SEQ ID NO:5) (参见例如Liu等人,《美国国家科学院院刊》5525-5530(1997));GRR (SEQ ID NO:6) (Pomerantz等人,1995,见上文);(GGGS) $_n$,其中n=1、2、3、4或5 (SEQ ID NO:7) (Kim等人,《美国国家科学院院刊》93,1156-1160(1996.));EGKSSGSGSESKVD (SEQ ID NO:8) (Chaudhary等人,1990,《美国国家科学院院刊》87:1066-1070);KESGSVSSEQLAQFRSLD (SEQ ID NO:9) (Bird等人,1988,《科学》242:423-426);GRRGGGS (SEQ ID NO:10);LRQRDGERP (SEQ ID NO:11);LRQKDGGSERP (SEQ ID NO:12);LRQK(DGGGS) $_2$ ERP (SEQ ID NO:13)。或者,柔性连接子可以使用能够模型化DNA-结合位点和肽本身的计算机程序(Desjarlais和Berg,《美国国家科学院院刊》90:2256-2260(1993),《美国国家科学院院刊》91:11099-11103(1994))或通过噬菌体呈现方法合理地设计。

[0494] 融合多肽还可以在本文中所描述的每个多肽结构域之间或在内源性开放阅读框架与由供体修复模板编码的多肽之间包含多肽裂解信号。此外,可将多肽裂解位点置于任何连接肽序列中。例示性多肽裂解信号包括多肽裂解识别位点,如蛋白酶裂解位点、核酸酶裂解位点(例如罕见限制酶识别位点、自裂解核糖核酸酶识别位点)和自裂解病毒寡肽(参见deFelipe和Ryan,2004.《交通(Traffic)》,5(8):616-26)。

[0495] 合适的蛋白酶裂解位点和自裂解肽是所属领域的技术人员已知的(参见例如Ryan等人,1997.《普通病毒学杂志(J. Gener. Virol.)》78,699-722;Scymczak等人(2004)《自然生物技术(Nature Biotech.)》5,589-594)。例示性蛋白酶裂解位点包括(但不限于)马铃薯Y病毒属(potyvirus)NIa蛋白酶(例如烟草蚀刻病毒蛋白酶)、马铃薯Y病毒属HC蛋白酶、马铃薯Y病毒属P1(P35)蛋白酶、byo病毒NIa蛋白酶、byo病毒RNA-2编码的蛋白酶、口疮病毒属(aphthovirus)L蛋白酶、肠病毒(enterovirus)2A蛋白酶、鼻病毒(rhinovirus)2A蛋白酶、细小核糖核酸3C蛋白酶、豇豆花叶病毒组(comovirus)24K蛋白酶、线虫传多面体病毒(nepovirus)24K蛋白酶、RTSV(水稻东格鲁球状病毒(rice tungro spherical virus))3C样蛋白酶、PYVF(欧防风黄点病毒(parsnip yellow fleck virus))3C样蛋白酶、肝素、凝血酶、因子Xa和肠激酶的裂解位点。归因于其高裂解严格度,TEV(烟草蚀刻病毒)蛋白酶裂解位点在一个实施例中是优选的,例如EXXYXQ(G/S)(SEQ ID NO:14),例如ENLYFQG(SEQ ID NO:15)和ENLYFQS(SEQ ID NO:16),其中X表示任何氨基酸(由TEV进行的裂解发生在Q与G或Q与S之间)。

[0496] 在某些实施例中,自裂解多肽位点包含2A或2A样位点、序列或结构域(Donnelly等人,2001.《普通病毒学杂志》82:1027-1041)。在具体实施例中,病毒2A肽是口疮病毒属2A肽、马铃薯Y病毒属2A肽或心病毒属(cardiovirus)2A肽。

[0497] 在一个实施例中,病毒2A肽是选自由以下组成的群组:口蹄疫病毒(foot-and-mouth disease virus;FMDV)(F2A)肽、马鼻炎A病毒(equine rhinitis A virus;ERAV)(E2A)2A肽、明脉扁刺蛾病毒(Thosea asigna virus;TaV)(T2A)肽、猪捷申病毒-1(porcine teschovirus-1;PTV-1)(P2A)肽、泰勒病毒(Theilovirus)2A肽和脑心肌炎病毒2A肽。

[0498] 2A位点的说明性实例提供于表2中。

[0499] 表2:例示性2A位点包括以下序列:

[0500]

SEQ ID NO:17	GSGATNFSLLKQAGDVEENPGP
SEQ ID NO:18	ATNFSLLKQAGDVEENPGP
SEQ ID NO:19	LLKQAGDVEENPGP
SEQ ID NO:20	GSGEGRGSLTTCGDVEENPGP
SEQ ID NO:21	EGRGSLTTCGDVEENPGP
SEQ ID NO:22	LLTCGDVEENPGP
SEQ ID NO:23	GSGQCTNYALLKLAGDVESNPGP
SEQ ID NO:24	QCTNYALLKLAGDVESNPGP
SEQ ID NO:25	LLKLAGDVESNPGP
SEQ ID NO:26	GSGVKQTLNFDLLKLAGDVESNPGP
SEQ ID NO:27	VKQTLNFDLLKLAGDVESNPGP
SEQ ID NO:28	LLKLAGDVESNPGP
SEQ ID NO:29	LLNFDLLKLAGDVESNPGP
SEQ ID NO:30	TLNFDLLKLAGDVESNPGP
SEQ ID NO:31	LLKLAGDVESNPGP
SEQ ID NO:32	NFDLLKLAGDVESNPGP

[0501]

SEQ ID NO:33	QLLNFDLLKLAGDVESNPGP
SEQ ID NO:34	APVKQTLNFDLLKLAGDVESNPGP
SEQ ID NO:35	VTELLYRMKRAETY CPRLLAIHPTEARHKQKIVAPVKQT
SEQ ID NO:36	LNFDLLKLAGDVESNPGP
SEQ ID NO:37	LLAIHPTEARHKQKIVAPVKQTLNFDLLKLAGDVESNPGP
SEQ ID NO:38	EARHKQKIVAPVKQTLNFDLLKLAGDVESNPGP

[0502] F. 聚核苷酸

[0503] 在具体实施例中,提供DNA供体修复模板聚核苷酸和编码本文中所涵盖的免疫效能增强子、免疫抑制信号减弱子、被工程改造的抗原受体和被工程改造的核酸酶的聚核苷酸。如本文中所使用,术语“聚核苷酸”或“核酸”是指脱氧核糖核酸(DNA)、核糖核酸(RNA)和DNA/RNA杂交体。聚核苷酸可以是单链或双链的并且是重组、合成或分离的。聚核苷酸包括(但不限于):前信使RNA(pre-mRNA)、信使RNA(mRNA)、RNA、短干扰RNA(siRNA)、短发夹RNA(shRNA)、微小RNA(miRNA)、核糖核酸酶、基因组RNA(gRNA)、正链RNA(RNA(+))负链RNA(RNA(-))、tracrRNA、crRNA、单向导RNA(sgRNA)、合成RNA、合成mRNA、基因组DNA(gDNA)、PCR扩增的DNA、互补DNA(cDNA)、合成DNA或重组型DNA。聚核苷酸是指至少5个、至少10个、至少15个、至少20个、至少25个、至少30个、至少40个、至少50个、至少100个、至少200个、至少300个、至少400个、至少500个、至少1000个、至少5000个、至少10000个或至少15000个或更多个核苷酸长度,以及所有中间长度的核苷酸,即核糖核苷酸或脱氧核糖核苷酸或任一类型核苷酸的修饰形式的聚合物形式。应容易理解,在这种情形下,“中间长度”意指所述值之间的任何长度,如6、7、8、9等;101、102、103等;151、152、153等;201、202、203等。在具体实施例中,聚核苷酸或变体与参考序列具有至少或约50%、55%、60%、65%、70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%序列一致性。

[0504] 在具体实施例中,聚核苷酸可以经密码子最佳化。如本文中所使用,术语“密码子最佳化”是指取代编码多肽的聚核苷酸中的密码子以便增加所述多肽的表达、稳定性和/或

活性。影响密码子最佳化的因素包括(但不限于)以下中的一个或多个:(i)两种或更多种生物体或基因之间密码子偏好的变化或以合成方式构筑的偏好表;(ii)生物体、基因或基因集内密码子偏好程度的变化;(iii)包括环境在内的密码子的系统性变化;(iv)根据解码tRNA得到的密码子变化;(v)根据总体或三联体的一个位置中的GC%得到的密码子变化;(vi)与参考序列,例如天然存在的序列的相似程度的变化;(vii)密码子频率截止值的变化;(viii)由DNA序列转录的mRNA的结构特性;(ix)有关作为设计密码子取代集合的基础的DNA序列的功能的先验知识;(x)每个氨基酸的密码子集合的系统性变化;和/或(xi)分离去除的伪翻译起始位点。

[0505] 如本文中所使用,术语“核苷酸”是指与磷酸化糖的N-糖苷键联中的杂环含氮碱基。核苷酸应理解为包括天然碱基,以及多种所属领域中认可的被修饰的碱基。这类碱基通常位于核苷酸糖部分的1'位置处。核苷酸通常包含碱基、糖和磷酸酯基。在核糖核酸(RNA)中,糖是核糖,并且在脱氧核糖核酸(DNA)中,糖是脱氧核糖,即,不含核糖中存在的羟基的糖。例示性天然含氮碱基包括嘌呤,即腺苷(A)和胍(G);以及嘧啶,即胞苷(C)和胸苷(T)(或在RNA情形中是尿嘧啶(U))。脱氧核糖的C-1原子结合于嘧啶的N-1或嘌呤的N-9。核苷酸通常是单磷酸酯、二磷酸酯或三磷酸酯。核苷酸可以是未被修饰的或在糖、磷酸酯和/或碱基部分处被修饰(也可以互换地称为核苷酸类似物、核苷酸衍生物、被修饰的核苷酸、非天然核苷酸和非标准核苷酸;参见例如WO 92/07065和WO 93/15187)。被修饰的核酸碱基的实例概述于Limbach等人(1994,《核酸研究》22,2183-2196)中。

[0506] 核苷酸也可以被视为核苷的磷酸酯,并且酯化是在连接到糖的C-5的羟基上发生。如本文中所使用,术语“核苷”是指与糖的N-糖苷键联中的杂环含氮碱基。核苷在所属领域中公认是包括天然碱基,并且还包含众所周知的被修饰的碱基。这类碱基通常位于核苷糖部分的1'位置处。核苷通常包含碱基和糖基。核苷可以是未被修饰的或在糖和/或碱基部分处被修饰(也可以互换地称为核苷类似物、核苷衍生物、被修饰的核苷、非天然核苷或非标准核苷)。也如上文所提及,被修饰的核酸碱基的实例概述于Limbach等人(1994,《核酸研究》22,2183-2196)中。

[0507] 在各种说明性实施例中,本文中所涵盖的聚核苷酸包括(但不限于)DNA供体修复模板聚核苷酸和编码免疫效能增强子、免疫抑制信号减弱子、被工程改造的抗原受体和被工程改造的核酸酶的聚核苷酸,和包含本文中所涵盖的聚核苷酸的表达载体、病毒载体和转移质粒。

[0508] 如本文中所使用,术语“聚核苷酸变体”和“变体”等是指与参考聚核苷酸序列呈现显著序列一致性的聚核苷酸或在下文中所定义的严格条件下与参考序列杂交的聚核苷酸。这些术语还涵盖由于至少一个核苷酸的添加、缺失、取代或修饰而有别于参考聚核苷酸的聚核苷酸。因此,术语“聚核苷酸变体”和“变体”包括其中一个或多个核苷酸被添加或缺失或修饰,或由不同核苷酸置换的聚核苷酸。在这方面,在所属领域中充分理解,可以对参考聚核苷酸进行包括突变、添加、缺失和取代的某些改变,由此改变的聚核苷酸保留参考聚核苷酸的生物功能或活性。

[0509] 在一个实施例中,聚核苷酸包含在严格条件下与目标核酸序列杂交的核苷酸序列。在“严格条件”下杂交描述使彼此至少60%一致的核苷酸序列保持杂交的杂交方案。通常,选择严格条件比特定序列在既定离子强度和pH值下的热熔点(T_m)低约5°C。 T_m 是50%的

与目标序列互补的探针与目标序列在均衡状态下杂交的温度(在既定离子强度、pH值和核酸浓度下)。因为目标序列在 T_m 下通常过量存在,所以平衡状态下占据50%探针。

[0510] 如本文中所使用,叙述“序列一致性”或例如包含“与……50%一致的序列”是指比较窗口上逐核苷酸计或逐氨基酸计的序列相同程度。因此,“序列一致性百分比”可以通过以下方式计算:在比较窗口内比较两个最佳对准的序列,测定的一致核酸碱基(例如A、T、C、G、I)或一致氨基酸残基(例如Ala、Pro、Ser、Thr、Gly、Val、Leu、Ile、Phe、Tyr、Trp、Lys、Arg、His、Asp、Glu、Asn、Gln、Cys以及Met)在两个序列中出现的位置的数目以得到相配位置的数目,用相配位置的数目除以比较窗口中的位置总数(即,窗口大小)并且将结果乘以100,得到序列一致性百分比。包括与本文中所描述的参考序列中的任一个具有至少约50%、55%、60%、65%、70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列一致性的核苷酸和多肽,典型地其中多肽变异体保持参考多肽的至少一种生物活性。

[0511] 用于描述两个或更多个聚核苷酸或多肽之间的序列关系的术语包括“参考序列”、“比较窗口”、“序列一致性”、“序列一致性百分比”和“大体一致性”。“参考序列”的长度是至少12个,但通常是15到18个,并且通常是至少25个单体单元,包括核苷酸和氨基酸残基。因为两个聚核苷酸可以各自包含(1)两个聚核苷酸之间类似的序列(即完全聚核苷酸序列的仅一部分),和(2)在两个聚核苷酸之间相异的序列,两个(或更多个)聚核苷酸之间的序列比较通常通过在“比较窗口”上比较两个聚核苷酸的序列以鉴别和比较局部区域的序列相似性来进行。“比较窗口”是指至少6个,通常约50到约100个,更通常约100到约150个相邻位置的概念性区段,其中在将两个序列最佳比对后,将序列与具有相同数目的相邻位置的参考序列比较。对于最佳对准的两个序列,比较窗口可以包含与参考序列(不包含添加或缺失的序列)相比,约20%或更低百分比的添加或缺失(即,空位)。用于比对比较窗口的序列最佳比对可以通过计算机化执行算法(美国威斯康星州的威斯康星州遗传学软件包7.0版本,遗传学计算机组,575科学驱动麦迪逊(Wisconsin Genetics Software Package Release 7.0, Genetics Computer Group, 575 Science Drive Madison, WI, USA)中的GAP、BESTFIT、FASTA和TFASTA)或通过多种选择方法中的任一种产生的检查和最佳比对(即在比较窗口上产生最高同源性百分比)来进行。还可以参考例如Altschul等人,1997,《核酸研究》25:3389所公开的BLAST程序家族。序列分析的详细论述可以见于Ausubel等人,《现代分子生物学实验技术(Current Protocols in Molecular Biology)》,John Wiley&Sons Inc., 1994-1998,第15章的19.3单元中。

[0512] 在各种实施例中,聚核苷酸包含编码本文中所涵盖的多肽的mRNA。在某些实施例中,mRNA包含帽、一个或多个核苷酸和聚(A)尾区。

[0513] 如本文所使用,术语“5'帽”或“5'帽结构”或“5'帽部分”是指并入mRNA的5'端的化学修饰。5'帽涉及核输出、mRNA稳定性和翻译。

[0514] 在具体实施例中,本文中所涵盖的mRNA包含5'帽,其包含在末端鸟苷帽残基与mRNA分子的5'端转录的有义核苷酸之间的5'-ppp-5'-三磷酸键联。接着,这个5'-鸟苷酸帽可以被甲基化以产生N7-甲基-鸟苷酸残基。

[0515] 适用于本文中所涵盖的mRNA聚核苷酸的具体实施例中的5'帽的说明性实例包括(但不限于):未甲基化5'帽类似物,例如G(5') ppp(5') G、G(5') ppp(5') C、G(5') ppp(5') A;甲基化5'帽类似物,例如 m^7 G(5') ppp(5') G、 m^7 G(5') ppp(5') C和 m^7 G(5') ppp(5') A;二甲甲基化

5'帽类似物,例如 $m^{2,7}G(5')$ ppp(5')G、 $m^{2,7}G(5')$ ppp(5')C和 $m^{2,7}G(5')$ ppp(5')A;三甲基化5'帽类似物,例如 $m^{2,2,7}G(5')$ ppp(5')G、 $m^{2,2,7}G(5')$ ppp(5')C和 $m^{2,2,7}G(5')$ ppp(5')A;二甲基化对称5'帽类似物,例如 $m^7G(5')$ pppm⁷(5')G、 $m^7G(5')$ pppm⁷(5')C和 $m^7G(5')$ pppm⁷(5')A;和抗逆转5'帽类似物,例如抗逆转帽类似物(ARCA)帽,称为3'-O-Me- $m^7G(5')$ ppp(5')G、2'-O-Me- $m^7G(5')$ ppp(5')G、2'-O-Me- $m^7G(5')$ ppp(5')C、2'-O-Me- $m^7G(5')$ ppp(5')A、 $m^72'd(5')$ ppp(5')G、 $m^72'd(5')$ ppp(5')C、 $m^72'd(5')$ ppp(5')A、3'-O-Me- $m^7G(5')$ ppp(5')C、3'-O-Me- $m^7G(5')$ ppp(5')A、 $m^73'd(5')$ ppp(5')G、 $m^73'd(5')$ ppp(5')C、 $m^73'd(5')$ ppp(5')A和其四磷酸盐衍生物(参见例如Jemielity等人,《RNA》,9:1108-1122(2003))。

[0516] 在具体实施例中,mRNA包含5'帽,其是7-甲基鸟苷酸(“ m^7G ”),通过三磷酸酯桥连接到第一个被转录的核苷酸的5'端,产生 $m^7G(5')$ ppp(5')N,其中N是任何核苷。

[0517] 在一些实施例中,mRNA包含5'帽,其中所述帽是帽0结构(帽0结构不包含连接到碱基1和2的核糖的2'-O-甲基残基)、帽1结构(帽1结构在碱基2处具有2'-O-甲基残基)或帽2结构(帽2结构具有连接到碱基2和3的2'-O-甲基残基)。

[0518] 在一个实施例中,mRNA包含 $m^7G(5')$ ppp(5')G帽。

[0519] 在一个实施例中,mRNA包含ARCA帽。

[0520] 在具体实施例中,本文中所涵盖的mRNA包含一个或多个被修饰的核苷。

[0521] 在一个实施例中,mRNA包含一个或多个选自以下组成的群组的被修饰的核苷:假尿苷、吡啶-4-酮核苷、5-氮杂-尿苷、2-硫-5-氮杂-尿苷、2-硫尿核苷、4-硫-假尿苷、2-硫-假尿苷、5-羟基尿苷、3-甲基尿苷、5-羧基甲基-尿苷、1-羧基甲基-假尿苷、5-丙炔基-尿苷、1-丙炔基-假尿苷、5-牛磺酸甲基尿苷、1-牛磺酸甲基-假尿苷、5-牛磺酸甲基-2-硫-尿苷、1-牛磺酸甲基-4-硫-尿苷、5-甲基-尿苷、1-甲基-假尿苷、4-硫-1-甲基-假尿苷、2-硫-1-甲基-假尿苷、1-甲基-1-脱氮-假尿苷、2-硫-1-甲基-1-脱氮-假尿苷、二氢尿苷、二氢假尿苷、2-硫-二氢尿苷、2-硫-二氢假尿苷、2-甲氧基尿苷、2-甲氧基-4-硫-尿苷、4-甲氧基-假尿苷、4-甲氧基-2-硫-假尿苷、5-氮-胞嘧啶核苷、假异胞苷、3-甲基-胞嘧啶核苷、N4-乙酰基胞嘧啶核苷、5-甲酰基胞嘧啶核苷、N4-甲基胞嘧啶核苷、5-羟基甲基胞苷、1-甲基-假异胞苷、吡咯并胞嘧啶核苷、吡咯并假异胞苷、2-硫-胞嘧啶核苷、2-硫-5-甲基-胞嘧啶核苷、4-硫-假异胞苷、4-硫-1-甲基-假异胞苷、4-硫-1-甲基-1-脱氮-假异胞苷、1-甲基-1-脱氮-假异胞苷、泽布拉恩(zebularine)、5-氮-泽布拉恩、5-甲基-泽布拉恩、5-氮-2-硫-泽布拉恩、2-硫-泽布拉恩、2-甲氧基-胞嘧啶核苷、2-甲氧基-5-甲基-胞嘧啶核苷、4-甲氧基-假异胞苷、4-甲氧基-1-甲基-假异胞苷、2-氨基嘌呤、2,6-二氨基嘌呤、7-脱氮-腺嘌呤、7-脱氮-8-氮杂-腺嘌呤、7-脱氮-2-氨基嘌呤、7-脱氮-8-氮杂-2-氨基嘌呤、7-脱氮-2,6-二氨基嘌呤、7-脱氮-8-氮杂-2,6-二氨基嘌呤、1-甲基腺苷、N6-甲基腺苷、N6-异戊烯基腺苷、N6-(顺-羟基异戊烯基)腺苷、2-甲硫基-N6-(顺-羟基异戊烯基)腺苷、N6-甘氨酸基氨甲酰基腺苷、N6-苏氨酸基氨甲酰基腺苷、2-甲硫基-N6-苏氨酸基氨甲酰基腺苷、N6,N6-二甲基腺苷、7-甲基腺嘌呤、2-甲硫基-腺嘌呤、2-甲氧基-腺嘌呤、肌苷、1-甲基-肌苷、γ-苷、怀丁苷、7-脱氮-鸟苷、7-脱氮-8-氮杂-鸟苷、6-硫-鸟苷、6-硫-7-脱氮-鸟苷、6-硫-7-脱氮-8-氮杂-鸟苷、7-甲基-鸟苷、6-硫-7-甲基-鸟苷、7-甲基肌苷、6-甲氧基-鸟苷、1-甲基鸟苷、N2-甲基鸟苷、N2,N2-二甲基鸟苷、8-氧-鸟苷、7-甲基-8-氧-鸟苷、1-甲基-6-硫-鸟苷、N2-甲基-6-硫-鸟苷和N2,N2-二甲基-6-硫-鸟苷。

[0522] 在一个实施例中,mRNA包含一个或多个选自由以下组成的群组的被修饰的核苷:假尿苷、吡啶-4-酮核苷、5-氮杂-尿苷、2-硫-5-氮杂-尿苷、2-硫尿核苷、4-硫-假尿苷、2-硫-假尿苷、5-羟基尿苷、3-甲基尿苷、5-羧基甲基-尿苷、1-羧基甲基-假尿苷、5-丙炔基至尿苷、1-丙炔基-假尿苷、5-牛磺酸甲基尿苷、1-牛磺酸甲基-假尿苷、5-牛磺酸甲基-2-硫-尿苷、1-牛磺酸甲基-4-硫-尿苷、5-甲基-尿苷、1-甲基-假尿苷、4-硫-1-甲基-假尿苷、2-硫-1-甲基-假尿苷、1-甲基-1-脱氮-假尿苷、2-硫-1-甲基-1-脱氮-假尿苷、二氢尿苷、二氢假尿苷、2-硫-二氢尿苷、2-硫-二氢假尿苷、2-甲氧基尿苷、2-甲氧基-4-硫-尿苷、4-甲氧基-假尿苷和4-甲氧基-2-硫-假尿苷。

[0523] 在一个实施例中,mRNA包含一个或多个选自由以下组成的群组的被修饰的核苷:5-氮杂-胞嘧啶核苷、假异胞苷、3-甲基-胞嘧啶核苷、N4-乙酰基胞嘧啶核苷、5-甲酰基胞嘧啶核苷、N4-甲基胞嘧啶核苷、5-羟基甲基胞苷、1-甲基-假异胞苷、吡咯并胞嘧啶核苷、吡咯并假异胞苷、2-硫-胞嘧啶核苷、2-硫-5-甲基-胞嘧啶核苷、4-硫-假异胞苷、4-硫-1-甲基-假异胞苷、4-硫-1-甲基-1-脱氮-假异胞苷、1-甲基-1-脱氮-假异胞苷、泽布拉恩、5-氮杂-泽布拉恩、5-甲基-泽布拉恩、5-氮杂-2-硫-泽布拉恩、2-硫-泽布拉恩、2-甲氧基-胞嘧啶核苷、2-甲氧基-5-甲基-胞嘧啶核苷、4-甲氧基-假异胞苷和4-甲氧基-1-甲基-假异胞苷。

[0524] 在一个实施例中,mRNA包含一个或多个选自由以下组成的群组的被修饰的核苷:2-氨基嘌呤、2,6-二氨基嘌呤、7-脱氮-腺嘌呤、7-脱氮-8-氮杂-腺嘌呤、7-脱氮-2-氨基嘌呤、7-脱氮-8-氮杂-2-氨基嘌呤、7-脱氮-2,6-二氨基嘌呤、7-脱氮-8-氮杂-2,6-二氨基嘌呤、1-甲基腺苷、N6-甲基腺苷、N6-异戊烯基腺苷、N6-(顺-羟基异戊烯基)腺苷、2-甲硫基-N6-(顺-羟基异戊烯基)腺苷、N6-甘氨酸基氨甲酰基腺苷、N6-苏氨酸基氨甲酰基腺苷、2-甲硫基-N6-苏氨酸基氨甲酰基腺苷、N6,N6-二甲基腺苷、7-甲基腺嘌呤、2-甲硫基-腺嘌呤和2-甲氧基-腺嘌呤。

[0525] 在一个实施例中,mRNA包含一个或多个选自由以下组成的群组的被修饰的核苷:肌苷、1-甲基-肌苷、γ苷、怀丁苷、7-脱氮-鸟苷、7-脱氮-8-氮杂-鸟苷、6-硫-鸟苷、6-硫-7-脱氮-鸟苷、6-硫-7-脱氮-8-氮杂-鸟苷、7-甲基-鸟苷、6-硫-7-甲基-鸟苷、7-甲基肌苷、6-甲氧基-鸟苷、1-甲基鸟苷、N2-甲基鸟苷、N2,N2-二甲基鸟苷、8-氧-鸟苷、7-甲基-8-氧-鸟苷、1-甲基-6-硫-鸟苷、N2-甲基-6-硫-鸟苷和N2,N2-二甲基-6-硫-鸟苷。

[0526] 在一个实施例中,mRNA包含一个或多个假尿苷、一个或多个5-甲基-胞嘧啶和/或一个或多个5-甲基-胞苷。

[0527] 在一个实施例中,mRNA包含一个或多个假尿苷。

[0528] 在一个实施例中,mRNA包含一个或多个5-甲基-胞苷。

[0529] 在一个实施例中,mRNA包含一个或多个5-甲基-胞嘧啶。

[0530] 在具体实施例中,本文中所涵盖的mRNA包含聚(A)尾区以帮助保护mRNA避免核酸外切酶降解,使mRNA稳定和促进翻译。在某些实施例中,mRNA包含3'聚(A)尾区结构。

[0531] 在具体实施例中,聚(A)尾区的长度是至少约10、25、50、75、100、150、200、250、300、350、400、450或至少约500个或更多个腺嘌呤核苷酸或任何中间数目的腺嘌呤核苷酸。在具体实施例中,聚(A)尾区的长度是至少约125、126、127、128、129、130、131、132、133、134、135、136、137、138、139、140、141、142、143、144、145、146、147、148、149、150、151、152、153、154、155、156、157、158、159、160、161、162、163、164、165、166、167、168、169、170、171、

172、173、174、175、176、177、178、179、180、181、182、183、184、185、186、187、188、189、190、191、192、193、194、195、196、197、198、199、200、201、202、202、203、205、206、207、208、209、210、211、212、213、214、215、216、217、218、219、220、221、222、223、224、225、226、227、228、229、230、231、232、233、234、235、236、237、238、239、240、241、242、243、244、245、246、247、248、249、250、251、252、253、254、255、256、257、258、259、260、261、262、263、264、265、266、267、268、269、270、271、272、273、274或275个或更多个腺嘌呤核苷酸。

[0532] 在具体实施例中,聚(A)尾区的长度是约10到约500个腺嘌呤核苷酸、约50到约500个腺嘌呤核苷酸、约100到约500个腺嘌呤核苷酸、约150到约500个腺嘌呤核苷酸、约200到约500个腺嘌呤核苷酸、约250到约500个腺嘌呤核苷酸、约300到约500个腺嘌呤核苷酸、约50到约450个腺嘌呤核苷酸、约50到约400个腺嘌呤核苷酸、约50到约350个腺嘌呤核苷酸、约100到约500个腺嘌呤核苷酸、约100到约450个腺嘌呤核苷酸、约100到约400腺嘌呤核苷酸、约100到约350个腺嘌呤核苷酸、约100到约300个腺嘌呤核苷酸、约150到约500个腺嘌呤核苷酸、约150到约450个腺嘌呤核苷酸、约150到约400个腺嘌呤核苷酸、约150到约350个腺嘌呤核苷酸、约150到约300个腺嘌呤核苷酸、约150到约250个腺嘌呤核苷酸、约150到约200个腺嘌呤核苷酸、约200到约500个腺嘌呤核苷酸、约200到约450个腺嘌呤核苷酸、约200到约400个腺嘌呤核苷酸、约200到约350个腺嘌呤核苷酸、约200到约300个腺嘌呤核苷酸、约250到约500个腺嘌呤核苷酸、约250到约450个腺嘌呤核苷酸、约250到约400个腺嘌呤核苷酸、约250到约350个腺嘌呤核苷酸或约250到约300个腺嘌呤核苷酸或任何中间范围的腺嘌呤核苷酸。

[0533] 描述聚核苷酸的定向的术语包括:5' (通常是聚核苷酸的具有游离磷酸酯基的末端) 和3' (通常是聚核苷酸的具有游离羟基(OH)的末端)。聚核苷酸序列可以按5'到3'定向或3'到5'定向标注。对于DNA和mRNA,5'到3'链称为“有义”链、“正”链或“编码”链,因为其序列与前信使(前mRNA)的序列一致[不过在RNA中是尿嘧啶(U),而在DNA中是胸腺嘧啶(T)]。对于DNA和mRNA,互补3'到5'链(其是由RNA聚合酶转录的链)称为“模板”、“反义”、“负”或“非编码”链。如本文中所使用,术语“逆定向”是指以3'到5'定向书写的5'到3'序列或以5'到3'定向书写的3'到5'序列。

[0534] 术语“互补”和“互补性”是指通过碱基配对规则相关联的聚核苷酸(即,核苷酸序列)。举例来说,DNA序列5' A G T C A T G 3'的互补链是3' T C A G T A C 5'。后一序列通常写成反向互补序列,5'端在左侧并且3'端在右侧,5' C A T G A C T 3'。与其反向互补序列相等的序列被称为回文序列。互补可以“部分”,其中仅一些核酸的碱基根据碱基配对规则匹配。或者,核酸之间可以存在“完全的”或“全部的”互补性。

[0535] 如本文中所使用的术语“核酸盒”或“表达盒”是指可以表达RNA并且随后表达多肽的载体内的基因序列。在一个实施例中,核酸盒含有相关基因,例如相关聚核苷酸。在另一个实施例中,核酸盒含有一个或多个表达控制序列,例如启动子、增强子、聚(A)序列以及相关基因,例如相关聚核苷酸。载体可以包含1、2、3、4、5、6、7、8、9或10个或更多个核酸盒。核酸盒位置上并且依序地定向在载体内,使得盒中的核酸可以转录成RNA,并且当需要时翻译成蛋白质或多肽,经历转化的细胞的活性所需要的适当翻译后修饰,并且通过靶向适当胞内区室或分泌到胞外区室中而易位到适当区室以获得生物活性。优选地,盒使其3'和5'端适合于简便插入到载体中,例如其在每个端具有限制性核酸内切酶位点。在优选实施例中,

核酸盒含有用于治疗、预防或改善遗传病症的治疗基因序列。盒可以按单一单元的形式移出并且插入到质粒或病毒载体中。

[0536] 聚核苷酸包括相关聚核苷酸。如本文中所使用,术语“相关聚核苷酸”是指如本文所预期,编码多肽或融合多肽的聚核苷酸或用作模板以转录抑制性聚核苷酸的聚核苷酸。

[0537] 此外,所属领域的一般技术人员应了解,如本文所预期,由于遗传密码的简并性,存在许多可以编码多肽或其变异体片段的核苷酸序列。这些聚核苷酸中的一些与任何天然基因的核苷酸序列具有最小同源性。尽管如此,在具体实施例中专门涵盖归因于密码子使用差异而变化的聚核苷酸,例如针对人类和/或灵长类动物密码子选择而最佳化的聚核苷酸。在一个实施例中,提供包含具体等位基因序列的聚核苷酸。等位基因是因核苷酸的一个或多个突变,如缺失、添加和/或取代而改变的内源性聚核苷酸序列。

[0538] 在某一实施例中,相关聚核苷酸包含供体修复模板。

[0539] 在某一实施例中,相关聚核苷酸包含抑制性聚核苷酸,包括(但不限于) siRNA、miRNA、shRNA、核糖核酸酶或另一抑制性RNA。

[0540] 在一个实施例中,包含抑制性RNA的供体修复模板包含一个或多个调节序列,例如强组成型pol III,例如人类或小鼠U6 snRNA启动子、人类和小鼠H1 RNA启动子或人类tRNA-val启动子,或强组成型pol II启动子,如本文中其它地方所描述。

[0541] 具体实施例中涵盖的聚核苷酸,不管编码序列本身的长度如何,可以与其它DNA序列组合,如本文中其它地方所公开或如所属领域中已知,所述序列是如启动子和/或增强子、非翻译区(UTR)、Kozak序列、聚腺苷酸化信号、其它限制酶位点、多克隆位点、内部核糖体进入位点(IRES)、重组酶识别位点(例如LoxP、FRT和Att位点)、终止密码子、转录终止信号、转录后反应元件以及编码自裂解多肽的聚核苷酸、表位标签,使得其总长度可能显著变化。因此,具体实施例中预期可以使用几乎任何长度的聚核苷酸片段,并且总长度优选受预定重组型DNA方案中的制备简易性和用途限制。

[0542] 聚核苷酸可以使用所属领域中已知并且可用的多种公认技术中的任一种制备、操作、表达和/或递送。为了表达所需多肽,编码多肽的核苷酸序列可以插入适当载体中。也可以通过将编码所需多肽的mRNA递送到细胞中来表达所述多肽。

[0543] 载体的说明性实例包括(但不限于)质粒、自主复制序列和转座因子,例如Sleeping Beauty、PiggyBac。

[0544] 载体的其它说明性实例包括(但不限于)质粒;噬菌粒;粘粒;人工染色体,如酵母人工染色体(YAC)、细菌人工染色体(BAC)或P1衍生的人工染色体(PAC);噬菌体,如 λ 噬菌体或M13噬菌体;和动物病毒。

[0545] 适用作载体的病毒的说明性实例包括(但不限于)逆转录病毒(包括慢病毒)、腺病毒、腺相关病毒、疱疹病毒(例如单纯疱疹病毒)、痘病毒、杆状病毒、乳头瘤病毒以及乳多空病毒(例如SV40)。

[0546] 表达载体的说明性实例包括(但不限于)用于在哺乳动物细胞中表达的pCIneo载体(Promega);用于慢病毒介导的基因转移和在哺乳动物细胞中表达的pLenti4/V5-DESTTM、pLenti6/V5-DESTTM和pLenti6.2/V5-GW/lacZ(Invitrogen)。在具体实施例中,本文中所公开的多肽的编码序列可以接合到用于表达哺乳动物细胞中的多肽的这类表达载体中。

[0547] 在具体实施例中,载体是游离型载体或维持在染色体外的载体。如本文中所使用,

术语“游离型”是指载体能够复制而不整合到宿主的染色体DNA中并且不由分裂宿主细胞逐渐丧失,还意指所述载体在染色体外或游离地复制。

[0548] 表达载体中存在的“表达控制序列”、“控制元件”或“调节序列”是载体的非翻译区域,即复制起点、选择盒、启动子、增强子、翻译起始信号(Shine Dalgarno序列或Kozak序列)内含子、转录后调节元件、聚腺苷酸化序列、5'和3'非翻译区,所述非翻译区与宿主细胞蛋白质相互作用以进行转录和翻译。这类元件可以在其强度和特异性方面变化。取决于所利用的载体系统和宿主,可以使用任何数目的合适的转录与翻译元件,包括遍在启动子和诱导型启动子。

[0549] 在具体实施例中,聚核苷酸包含载体,包括(但不限于)表达载体和病毒载体。载体可以包含一个或多个外源、内源或异源控制序列,如启动子和/或增强子。“内源控制序列”是天然地与基因组中的既定基因连接的控制序列。“外源控制序列”是借助于遗传操作(即,分子生物技术)与基因并接定位使得所述基因的转录通过所连接的增强子/启动子引导的控制序列。“异源控制序列”是来自与所遗传操作的细胞不同的物种的外源序列。“合成”控制序列可以包含具有一个或多个内源和/或外源序列和/或在体外或通过计算机模拟确定的序列的元件,其为具体疗法提供最佳的启动子和/或增强子活性。

[0550] 如本文中所使用,术语“启动子”是指RNA聚合酶所结合的聚核苷酸(DNA或RNA)的识别位点。RNA聚合酶起始并且转录可操作地连接到启动子的聚核苷酸。在具体实施例中,在哺乳动物细胞中可操作的启动子包含位于转录起始位点上游约25到30个碱基处的富AT区,和/或见于转录起始上游70到80个碱基处的另一序列,N可以是任何核苷酸的CNCAAT区。

[0551] 术语“增强子”是指含有能够提供增强的转录的序列并且在一些情况下可以独立于其相对于另一个控制序列的定向起作用的一段DNA。增强子可以与启动子和/或其它增强子元件协作地或叠加地起作用。术语“启动子/增强子”是指含有能够提供启动子和增强子功能的序列的一段DNA。

[0552] 术语“可操作地连接”是指所描述的组件处于允许其以其预期方式起作用的关系的并接。在一个实施例中,所述术语是指核酸表达控制序列(如启动子和/或增强子)与第二聚核苷酸序列(例如相关聚核苷酸)之间的功能性连接,其中表达控制序列引导对应于第二序列的核酸的转录。

[0553] 如本文中所使用,术语“组成型表达控制序列”是指不断地或连续地引起可操作地连接的序列转录的启动子、增强子或启动子/增强子。组成型表达控制序列可以是允许在多种细胞和组织类型中表达的“遍在”启动子、增强子或启动子/增强子,或允许分别在多种限定细胞和组织类型中表达的“细胞特异性”、“细胞类型特异性”、“细胞谱系特异性”或“组织特异性”启动子、增强子或启动子/增强子。

[0554] 适用于具体实施例中的说明性遍在表达控制序列包括(但不限于)巨细胞病毒(CMV)立即早期启动子、病毒猿猴病毒40(SV40)(例如早期或晚期)、莫洛尼鼠白血病毒(Moloney murine leukemia virus;MoMLV)LTR启动子、劳氏肉瘤病毒(Rous sarcoma virus;RSV)LTR、单纯疱疹病毒(HSV)(胸苷激酶)启动子、来自牛痘病毒的H5、P7.5和P11启动子、短延伸因子1- α (EF1a-短)启动子、长延伸因子1- α (EF1a-长)启动子、早期生长反应1(EGR1)、铁蛋白H(FerH)、铁蛋白L(FerL)、甘油醛3-磷酸脱氢酶(GAPDH)、真核翻译起始因子4A1(EIF4A1)、热休克70kDa蛋白质5(HSPA5)、热休克蛋白90kDa β 成员1(HSP90B1)、热休克蛋

白70kDa (HSP70)、 β -驱动蛋白(β -KIN)、人类ROSA 26基因座(Irions等人,《自然生物技术(Nature Biotechnology)》25,1477-1482(2007))、泛素C启动子(UBC)、磷酸甘油酸激酶-1(PGK)启动子、巨细胞病毒增强子/鸡 β -肌动蛋白(CAG)启动子、 β -肌动蛋白启动子和骨髓增生肉瘤病毒增强子、负调控区缺失、dl587rev引物结合位点取代(MND)的启动子(Challita等人,《病毒学杂志(J Virol.)》69(2):748-55(1995))。

[0555] 在具体实施例中,可能需要使用细胞、细胞类型、细胞谱系或组织特异性表达控制序列来实现所需聚核苷酸序列的细胞类型特异性、谱系特异性或组织特异性表达(例如,仅在细胞类型、细胞谱系或组织的子集中或在特定发展阶段期间表达编码多肽的具体核酸)。

[0556] 如本文中所使用,“条件表达”可以指任何类型的条件表达,包括(但不限于)诱导型表达;阻抑型表达;在具有具体生理、生物或疾病状态的细胞或组织中的表达;等。这一定义不打算排除细胞类型或组织特异性表达。某些实施例提供相关聚核苷酸的条件表达,例如通过使细胞、组织、生物体等经历可以引起聚核苷酸表达或引起由相关聚核苷酸编码的聚核苷酸的表达增加或减少的处理或条件来控制表达。

[0557] 诱导型启动子/系统的说明性实例包括(但不限于)类固醇诱导型启动子,如用于编码糖皮质激素或雌激素受体的基因的启动子(可以通过用相应激素处理来诱导);金属硫蛋白启动子(可以通过用各种重金属处理来诱导);MX-1启动子(可以通过干扰素来诱导);“GeneSwitch”米非司酮可调节系统(Sirin等人,2003,《基因》,323:67);cumate诱导型基因开关(WO 2002/088346);四环素依赖性调节系统;等。

[0558] 条件表达还可以通过使用位点特异性DNA重组酶来实现。根据某些实施例,聚核苷酸包含至少一个(典型地两个)用于由位点特异性重组酶介导的重组的位点。如本文中所使用,术语“重组酶”或“位点特异性重组酶”包括参与涉及一个或多个(例如二、三、四、五、六、七、八、九、十个或更多个)重组位点的重组反应的切除型或整合型蛋白质、酶、辅因子或相关蛋白质,其可以是野生型蛋白质(参见Landy,《生物技术新见(Current Opinion in Biotechnology)》3:699-707(1993))或其突变体、衍生物(例如含有重组蛋白质序列或其片段的融合蛋白质)、片段和变异体。适用于具体实施例中的重组酶的说明性实例包括(但不限于):Cre、Int、IHF、Xis、Flp、Fis、Hin、Gin、 Φ C31、Cin、Tn3解离酶、TndX、XerC、XerD、TnpX、Hjc、Gin、SpCCE1以及ParA。

[0559] 聚核苷酸可以包含用于多种位点特异性重组酶中的任一种的一个或多个重组位点。应理解,除了载体(例如逆转录病毒载体或慢病毒载体)的整合所需的任何位点以外,还需要用于位点特异性重组酶的目标位点。如本文中所使用,术语“重组序列”、“重组位点”或“位点特异性重组位点”是指重组酶所识别并且结合的具体核酸序列。

[0560] 举例来说,Cre重组酶的一个重组位点是loxP,其是34个碱基对的序列,包含侧接8个碱基对的核心序列的两个13个碱基对的反向重复序列(充当重组酶结合位点)(参看Sauer B.,《生物技术新见》5:521-527(1994)中的图1)。其它例示性loxP位点包括(但不限于):lox511(Hoess等人,1996;Bethke和Sauer,1997)、lox5171(Lee和Saito,1998)、lox2272(Lee和Saito,1998)、m2(Langer等人,2002)、lox71(Albert等人,1995)以及lox66(Albert等人,1995)。

[0561] 适用于FLP重组酶的识别位点包括(但不限于):FRT(McLeod等人,1996)、F₁、F₂、F₃(Schlake和Bode,1994)、F₄、F₅(Schlake和Bode,1994)、FRT(LE)(Senecoff等人,1988)、FRT

(RE) (Senecoff等人,1988)。

[0562] 识别序列的其它实例是attB、attP、attL和attR序列,其由重组酶 λ 整合酶,例如 ϕ -c31识别。 ϕ C31 SSR介导仅异型位点attB(长度是34bp)与attP(长度是39bp)之间的重组(Groth等人,2000)。分别针对细菌和噬菌体基因组上的噬菌体整合酶的连接位点而命名的attB和attP都含有可能与 ϕ C31同源二聚体结合的不完美反向重复序列(Groth等人,2000)。产物位点attL和attR对进一步 ϕ C31介导的重组实际上是惰性的(Belteki等人,2003),使得反应不可逆。对于催化插入,已经发现,与attP位点插入基因组attB位点中相比,携有attB的DNA更容易插入基因组attP位点中(Thyagarajan等人,2001;Belteki等人,2003)。因此,典型策略通过同源重组将携有attP的“对接位点”安置到既定基因座中,所述基因座接着与携有attB的进入序列搭配用于插入。

[0563] 在一个实施例中,本文中所涵盖的聚核苷酸包含侧接一对重组酶识别位点的供体修复模板聚核苷酸。在具体实施例中,修复模板聚核苷酸侧接LoxP位点、FRT位点或att位点。

[0564] 在具体实施例中,本文中所涵盖的聚核苷酸包括编码一个或多个多肽的一个或多个相关聚核苷酸。在具体实施例中,为了实现多个多肽中的每一个的有效翻译,聚核苷酸序列可以由一个或多个编码自裂解多肽的IRES序列或聚核苷酸序列隔开。

[0565] 如本文中所使用,“内部核糖体进入位点”或“IRES”是指促进直接内部核糖体进入顺反子(蛋白质编码区)的起始密码子(如ATG),由此引起基因的与帽无关的翻译的元件。参见例如Jackson等人,1990.《生化科学趋势(Trends Biochem Sci)》15(12):477-83,和Jackson和Kaminski.1995.《RNA》1(10):985-1000。所属领域的技术人员通常使用的IRES的实例包括美国专利案第6,692,736号中所描述的那些。所属领域中已知的“IRES”的其它实例包括(但不限于)获自小核糖核酸病毒的IRES(Jackson等人,1990)和获自病毒或细胞mRNA源,例如免疫球蛋白重链结合蛋白质(BiP)、血管内皮生长因子(VEGF)的IRES(Huez等人1998.《分子与细胞生物学(Mol.Cell.Biol.)》18(11):6178-6190)、成纤维细胞生长因子2(FGF-2)和胰岛素样生长因子(IGFII)、翻译起始因子eIF4G和酵母菌转录因子TFIID和HAP4、可以购自Novagen的脑心肌炎病毒(EMCV)(Duke等人,1992.《病毒学杂志》66(3):1602-9)和VEGF IRES(Huez等人,1998.《分子细胞生物学(Mol Cell Biol)》18(11):6178-90)。IRES在小核糖核酸病毒科(Picornaviridae)、二顺反子病毒科(Dicistroviridae)和黄病毒科(Flaviviridae)物种的病毒基因组中以及在HCV、弗云德鼠白血病病毒(Friend murine leukemia virus;FrMLV)和莫洛尼鼠白血病病毒(MoMLV)中也已有报导。

[0566] 在一个实施例中,用于本文中所涵盖的聚核苷酸中的IRES是EMCV IRES。

[0567] 在具体实施例中,聚核苷酸包含具有共同Kozak序列且编码所需多肽的聚核苷酸。如本文中所使用,术语“Kozak序列”是指极大地促进mRNA与核糖体的小型子单元的初始结合并且增加翻译的短核苷酸序列。共同Kozak序列是(GCC)RCCATGG(SEQ ID NO:39),其中R是嘌呤(A或G)(Kozak,1986.《细胞》.44(2):283-92,和Kozak,1987.《核酸研究》15(20):8125-48)。

[0568] 引导异源核酸转录物的有效终止和聚腺苷酸化的元件增加异源基因表达。转录终止信号通常见于聚腺苷酸化信号的下游。在具体实施例中,载体包含编码待表达的多肽的聚核苷酸的聚腺苷酸化序列3'。如本文中所使用,术语“聚腺苷酸位点(polyA site)”、“聚

腺苷酸序列 (polyA sequence)”、“聚(A)位点 (poly(A) site)”或“聚(A)序列 (poly(A) sequence)”表示引导由RNA聚合酶II引起的初生RNA转录物的终止和聚腺苷酸化的DNA序列。聚腺苷酸化序列可以通过添加聚(A)聚(A)到编码序列的3'端来提升mRNA稳定性,并且因此促进翻译效率提高。重组型转录物的有效聚腺苷酸化是合乎需要的,因为不含聚(A)尾区的转录物不稳定并且快速降解。可以用于载体中的聚(A)信号的说明性实例包括理想的聚(A)序列(例如AATAAA、ATTAAA、AGTAAA)、牛生长激素聚(A)序列(BGHpA)、兔 β -球蛋白聚(A)序列(r β gpA)或所属领域中已知的另一合适的异源或内源聚(A)序列。

[0569] 在一些实施例中,聚核苷酸或具有聚核苷酸的细胞利用自杀基因(包括诱导型自杀基因)来降低直接毒性和/或不受控增殖的风险。在具体实施例中,自杀基因对含有所述聚核苷酸或细胞的宿主不具有免疫原性。可以使用的自杀基因的某一实例是卡斯蛋白酶-9或卡斯蛋白酶-8或胞嘧啶脱氨酶。卡斯蛋白酶-9可以使用特定化学二聚诱导剂(CID)活化。

[0570] 在某些实施例中,聚核苷酸包含使本文中所涵盖的被基因修饰的细胞易于在体内发生阴性选择的基因区段。“阴性选择”是指输注的细胞可以由于个体体内条件的变化而消除。阴性可选表型可以由赋予对所给与的试剂(例如化合物)的敏感性的基因的插入产生。阴性选择基因是所属领域中已知的,并且包括(但不限于):赋予更昔洛韦(ganciclovir)敏感性的I型单纯疱疹病毒胸苷激酶(HSV-I TK)基因;细胞次黄嘌呤磷酸核糖转移酶(HPRT)基因、细胞腺嘌呤磷酸核糖转移酶(APRT)基因以及细菌胞嘧啶脱氨酶。

[0571] 在一些实施例中,被基因修饰的细胞包含聚核苷酸,其进一步包含能够在体外选择具有阴性可选表型的细胞的阳性标记物。阳性可选标记物可以是在引入宿主细胞中后表达显性表型的基因,所述显性表型允许对带有所述基因的细胞进行阳性选择。这种类型的基因是所属领域中已知的并且包括(但不限于):赋予对潮霉素B(hygromycin B)的抗性的潮霉素-B磷酸转移酶基因(hph);编码对抗生素G418的抗性的来自Tn5的氨基糖苷磷酸转移酶基因(neo或aph);二氢叶酸还原酶(DHFR)基因;腺苷脱氨酶基因(ADA);以及多药耐药性(MDR)基因。

[0572] 在一个实施例中,阳性可选标记物和阴性可选元件被连接起来,使得阴性可选元件的丧失必然还伴随阳性可选标记物的丧失。在具体实施例中,阳性和阴性可选标记物融合,使得其中之一的丧失强制性地引起另一个的丧失。以表达产物的形式产生赋予上文所述的所需阳性和阴性选择特征的多肽的融合聚核苷酸的实例是潮霉素磷酸转移酶胸苷激酶融合基因(HyTK)。这一基因的表达产生赋予潮霉素B抗性用于体外阳性选择和更昔洛韦敏感性用于体内阴性选择的多肽。还参见S.D.Lupton的PCT US91/08442和PCT/US94/05601的公开案,其描述由融合显性阳性可选标记物与阴性可选标记物而衍生的双功能可选融合基因的用途。

[0573] 优选的阳性可选标记物是来源于选自由以下组成的群组的基因:hph、neo和gpt,并且优选的阴性可选标记物是来源于选自由以下组成的群组的基因:胞嘧啶脱氨酶、HSV-I TK、VZV TK、HPRT、APRT和gpt。在具体实施例中涵盖的例示性双功能可选融合基因包括(但不限于)其中阳性可选标记物来源于hph或neo,且阴性可选标记物来源于胞嘧啶脱氨酶或TK基因或可选标记物的基因。

[0574] 在具体实施例中,可以通过非病毒和病毒方法将聚核苷酸引入造血细胞(例如T细胞)。在具体实施例中,可以通过相同方法或通过不同方法,和/或通过相同载体或通过不同

载体递送一个或多个聚核苷酸。

[0575] 术语“载体”在本文中用于指能够转移或输送另一个核酸分子的核酸分子。被转移的核酸通常连接到载体核酸分子，例如插入载体核酸分子中。载体可以包括在细胞中直接自主复制的序列，或可以包括足以允许整合到宿主细胞DNA中的序列。在具体实施例中，使用非病毒载体将本文中所涵盖的一个或多个聚核苷酸递送到T细胞中。

[0576] 非病毒载体的说明性实例包括(但不限于)质粒(例如DNA质粒或RNA质粒)、转座子、粘粒和细菌人工染色体。

[0577] 在具体实施例中涵盖的聚核苷酸的非病毒递送的说明性方法包括(但不限于):电致孔、声致穿孔、脂质体转染、微注射、基因枪法(biolistics)、病毒粒子、脂质体、免疫脂质体、纳米粒子、聚阳离子或脂质:核酸结合物、裸DNA、人工病毒粒子、DEAE-聚葡萄糖介导的转移、基因枪(gene gun)和热休克。

[0578] 具体实施例中所涵盖的适用于具体实施例中的聚核苷酸递送系统的说明性实例包括(但不限于)由Amaxa Biosystems、Maxcyte, Inc.、BTX Molecular Delivery Systems和Copernicus Therapeutics Inc.提供的那些。脂质体转染试剂是市售的(例如Transfectam™和Lipofectin™)。适用于聚核苷酸的有效受体识别脂质体转染的阳离子性和中性脂质已描述于文献中。参见例如Liu等人(2003)《基因疗法》10:180-187;和Balazs等人(2011)《药物递送杂志(Journal of Drug Delivery.)》2011:1-12。抗体靶向、细菌衍生、基于非生命纳米单元的递送也涵盖于具体实施例中。

[0579] 包含具体实施例中所涵盖的聚核苷酸的病毒载体可以通过向各个患者给药来体内递送,通常通过全身给药(例如静脉内、腹膜内、肌肉内、皮下或颅内输注)或局部施用,如下文所描述。或者,可以将载体递送到离体细胞,如从各个患者取出的细胞(例如动员后外周血、淋巴细胞、骨髓穿刺液、组织活检样品等)或全适供体造血干细胞,随后将所述细胞再植入患者体内。

[0580] 在一个实施例中,向生物体直接给予编码DNA供体修复模板的病毒载体以用于体内细胞转导。或者,可以给予裸DNA。给药是通过通常用于引入分子以与血液或组织细胞最终接触的途径中的任一种进行,包括(但不限于)注射、输注、表面施用和电致孔。给予这类核酸的合适的方法是所属领域的技术人员可用且众所周知的,并且尽管可以使用超过一种途径给予具体组合物,但具体途径通常可以提供比另一途径更迅速且更有效的反应。

[0581] 适用于本文中所涵盖的具体实施例中的病毒载体系统的说明性实例包括(但不限于)腺相关病毒(AAV)、逆转录病毒、单纯疱疹病毒、腺病毒和牛痘病毒载体。

[0582] 在各种实施例中,通过用包含聚核苷酸的重组型腺相关病毒(rAAV)转导细胞来将编码DNA供体修复模板的聚核苷酸引入造血细胞,例如T细胞。

[0583] AAV是小型(约26nm)复制缺陷型,主要是游离型、无包膜病毒。AAV可以感染分裂细胞和非分裂细胞并且可以将其基因组并入宿主细胞的基因组中。重组型AAV(rAAV)通常至少由转基因和其调节序列,以及5'和3' AAV反向末端重复(ITR)构成。ITR序列的长度是约145bp。在具体实施例中,rAAV包含ITR和从AAV1、AAV2、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAV9或AAV10分离的衣壳序列。

[0584] 在一些实施例中,嵌合rAAV使用从一种AAV血清型分离ITR序列和从不同AAV血清型分离的衣壳序列。举例来说,具有来源于AAV2的ITR序列和来源于AAV6的衣壳序列的rAAV

称为AAV2/AAV6。在具体实施例中，rAAV载体可以包含来自AAV2的ITR，以及来自AAV1、AAV2、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAV9或AAV10中的任一个的衣壳蛋白。在优选实施例中，rAAV包含来源于AAV2的ITR序列和来源于AAV6的衣壳序列。在优选实施例中，rAAV包含来源于AAV2的ITR序列和来源于AAV2的衣壳序列。

[0585] 在一些实施例中，可以将工程改造和选择方法应用于AAV衣壳以使其更可能转导相关细胞。

[0586] rAAV载体的构造、制造和其纯化已经公开于例如美国专利案第9,169,494号；第9,169,492号；第9,012,224号；第8,889,641号；第8,809,058号；以及第8,784,799号中，其各自以全文引用的方式并入本文中。

[0587] 在各种实施例中，通过电致孔将编码多个核酸酶变异体的聚核苷酸引入造血细胞，并且通过用包含供体修复模板的逆转录病毒（例如慢病毒）转导细胞来将本文中所涵盖的供体修复模板引入造血细胞。

[0588] 如本文中所使用，术语“逆转录病毒”是指将其基因组RNA逆转录成线性双链DNA拷贝并且随后将其基因组DNA共价整合到宿主基因组中的RNA病毒。适用于具体实施例中的说明性逆转录病毒包括（但不限于）：莫洛尼鼠白血病毒（M-MuLV）、莫洛尼鼠肉瘤病毒（MoMSV）、哈维鼠（Harvey murine）肉瘤病毒（HaMuSV）、鼠乳腺肿瘤病毒（MuMTV）、长臂猿白血病毒（GaLV）、猫白血病毒（FLV）、泡沫病毒、弗兰德鼠（Friend murine）白血病毒、鼠干细胞病毒（MSCV）和劳斯肉瘤病毒（RSV）和慢病毒。

[0589] 如本文中所使用，术语“慢病毒”是指复杂逆转录病毒组（或属）。说明性慢病毒包括（但不限于）：HIV（人类免疫缺陷病毒；包括1型HIV和2型HIV）；维斯纳-梅迪病毒（visna-maedi virus；VMV）病毒；山羊关节炎-脑炎病毒（CAEV）；马传染性贫血病毒（EIAV）；猫免疫缺陷病毒（FIV）；牛免疫缺陷病毒（BIV）；和猿猴免疫缺陷病毒（SIV）。在一个实施例中，基于HIV的载体主链（即，HIV顺式作用序列元件）是优选的。

[0590] 在各种实施例中，本文中所涵盖的慢病毒载体包含一种或多种LTR，和以下辅助元件中的一种或多种或全部：cPPT/FLAP、Psi（ Ψ ）包装信号、导出元件、聚（A）序列，且可以任选地包含如本文中其它地方所论述的WPRE或HPRE、绝缘子元件、可选标记物和细胞自杀基因。

[0591] 在具体实施例中，本文中所涵盖的慢病毒载体可以是整合或非整合或整合缺陷型慢病毒。如本文中所使用，术语“整合缺陷型慢病毒”或“IDLV”是指具有不能将病毒基因组整合到宿主细胞的基因组中的整合酶的慢病毒。整合不胜任型病毒载体已描述于专利申请案W0 2006/010834中，其以全文引用的方式并入本文中。

[0592] 适合于降低整合酶活性的HIV-1pol基因中的说明性突变包括（但不限于）：H12N、H12C、H16C、H16V、S81 R、D41A、K42A、H51A、Q53C、D55V、D64E、D64V、E69A、K71A、E85A、E87A、D116N、D116I、D116A、N120G、N120I、N120E、E152G、E152A、D35E、K156E、K156A、E157A、K159E、K159A、K160A、R166A、D167A、E170A、H171A、K173A、K186Q、K186T、K188T、E198A、R199c、R199T、R199A、D202A、K211A、Q214L、Q216L、Q221 L、W235F、W235E、K236S、K236A、K246A、G247W、D253A、R262A、R263A和K264H。

[0593] 在一个实施例中，HIV-1整合酶缺陷型pol基因包含D64V、D116I、D116A、E152G或E152A突变；D64V、D116I和E152G突变；或D64V、D116A和E152A突变。

[0594] 在一个实施例中,HIV-1整合酶缺陷型pol基因包含D64V突变。

[0595] 术语“长末端重复序列(LTR)”是指位于逆转录病毒DNA的末端处的碱基对结构域, 其在天然序列情形下是直接重复序列并且含有U3、R和U5区。

[0596] 如本文中所使用,术语“FLAP元件”或“cPPT/FLAP”是指满足以下条件的核酸:其序列包括逆转录病毒(例如HIV-1或HIV-2)的中心多嘌呤段和中心终止序列(cPPT和CTS)。合适的FLAP元件描述于美国专利案第6,682,907号和Zennou等人,2000,《细胞》,101:173中。在另一个实施例中,慢病毒载体含有在cPPT和/或CTS元件中具有一个或多个突变的FLAP元件。在另一个实施例中,慢病毒载体包含cPPT或CTS元件。在另一个实施例中,慢病毒载体不包含cPPT或CTS元件。

[0597] 如本文中所使用,术语“包装信号”或“包装序列”是指位于逆转录病毒基因组内的psi[Ψ]序列,其是病毒RNA插入病毒衣壳或粒子中所需的,参见例如Clever等人,1995.《病毒学杂志》,第69卷,第4期;第2101-2109页。

[0598] 术语“导出元件”是指调节RNA转录物从细胞的细胞核输送到细胞质的顺式作用转录后调节元件。RNA导出元件的实例包括(但不限于)人类免疫缺陷病毒(HIV) rev反应元件(RRE)(参见例如Cullen等人,1991.《病毒学杂志》65:1053;和Cullen等人,1991.《细胞》58:423)和乙型肝炎病毒转录后调节元件(HPRE)。

[0599] 在具体实施例中,病毒载体中异源序列的表达通过将转录后调节元件、有效聚腺苷酸化位点和任选地转录终止信号并入载体中而增加。多种转录后调节元件可以增加异源核酸在蛋白质处的表达,例如土拨鼠肝炎病毒转录后调节元件(WPRE;Zufferey等人,1999,《病毒学杂志》,73:2886);存在于乙型肝炎病毒中的转录后调节元件(HPRE)(Huang等人,《分子与细胞生物学》,5:3864);等(Liu等人,1995,《基因与发育(Genes Dev.)》,9:1766)。

[0600] 慢病毒载体优选含有因修饰LTR而引起的若干安全性增强。“自我失活”(SIN)载体是指复制缺陷型载体,例如其中称为U3区的右(3')LTR增强子-启动子区已被修饰(例如通过缺失或取代)以防止病毒转录超过第一轮病毒复制。其它安全性增强是通过用异源启动子置换5'LTR的U3区以在病毒粒子产生期间驱动病毒基因组的转录而提供。可以使用的异源启动子的实例包括例如病毒猿猴病毒40(SV40)(例如早期或晚期)、巨细胞病毒(CMV)(例如立即早期)、莫洛尼鼠白血病毒(MoMLV)、劳氏肉瘤病毒(RSV)以及单纯疱疹病毒(HSV)(胸苷激酶)启动子。

[0601] 如本文中所使用,术语“假型”或“假型化”是指病毒包膜蛋白质已经被另一个具有优选特征的病毒的病毒包膜蛋白质取代的病毒。举例来说,HIV可以由水泡性口炎病毒G蛋白质(VSV-G)包膜蛋白质假型化,所述包膜蛋白质允许HIV感染更广泛范围的细胞,因为HIV包膜蛋白质(由env基因编码)通常使病毒靶向CD4⁺呈现细胞。

[0602] 在某些实施例中,根据已知方法制备慢病毒载体。参见例如Kutner等人,《BMC生物技术(BMC Biotechnol.)》2009;9:10.doi:10.1186/1472-6750-9-10;Kutner等人.《自然实验手册》2009;4(4):495-505.doi:10.1038/nprot.2009.22。

[0603] 根据本文中所涵盖的某些特定实施例,大部分或所有病毒载体主链序列都来源于慢病毒,例如HIV-1。然而,应理解,可以使用多种不同来源的逆转录病毒和/或慢病毒序列,或可以在不削弱转移载体执行本文中所描述的功能的能力的情况下对某些慢病毒序列提供组合并且众多的取代和变化。此外,多种慢病毒载体是所属领域中已知的,参见Naldini

等人, (1996a, 1996b和1998); Zufferey等人, (1997); Dull等人, 1998; 美国专利案第6, 013, 516号和第5, 994, 136号, 许多所述专利申请案适用于产生本文中所涵盖的病毒载体或转移质粒。

[0604] 在各种实施例中, 通过电致孔将编码多个核酸酶变异体的聚核苷酸引入造血细胞, 并且通过用包含供体修复模板的腺病毒转导细胞来将本文中所涵盖的供体修复模板引入造血细胞。

[0605] 基于腺病毒的载体能够在许多细胞类型中实现极高转导效率且不需要细胞分裂。利用这类载体, 已获得高效价和高水平表达。这种载体可以在相对简单的系统中大量制备。大部分腺病毒载体被工程改造使得转基因置换Ad E1a、E1b和/或E3基因; 接着, 复制缺陷型载体在供应反式缺失基因功能的人类293细胞中繁殖。Ad载体可以在体内转导多种类型的组织, 包括非分裂、分化细胞, 如在肝、肾和肌肉中发现的那些。常规的Ad载体具有较大运载能力。

[0606] 当前复制缺乏型腺病毒载体的产生和传送可以利用称为293的独特辅助细胞系, 这种细胞系利用Ad5 DNA片段转型人胚肾细胞且组成性表达E1蛋白质 (Graham等人, 1977)。由于E3区在腺病毒基因组中是可有可无的 (Jones和Shenk, 1978), 所以当前的腺病毒载体借助于293细胞在E1、D3或两个区域中运载外来DNA (Graham和Prevec, 1991)。腺病毒载体已用于真核基因表达 (Levrero等人, 1991; Gomez-Foix等人, 1992) 和疫苗研究 (Grunhaus和Horwitz, 1992; Graham和Prevec, 1992) 中。向不同组织施用重组型腺病毒的研究包括气管灌输 (Rosenfeld等人, 1991; Rosenfeld等人, 1992)、肌肉注射 (Ragot等人, 1993)、周边静脉注射 (Herz和Gerard, 1993) 和立体定向接种到大脑 (Le Gal La Salle等人, 1993)。在临床试验中使用Ad载体的实例涉及使用肌肉内注射进行抗肿瘤免疫接种的聚核苷酸疗法 (Sterman等人, 《人类基因治疗 (Hum. Gene Ther.)》7:1083-9 (1998))。

[0607] 在各种实施例中, 通过电致孔将编码多个核酸酶变异体的聚核苷酸引入造血细胞, 并且通过用包含供体修复模板的单纯疱疹病毒 (例如HSV-1、HSV-2) 转导细胞来将本文中所涵盖的供体修复模板引入造血细胞。

[0608] 成熟HSV病毒粒子由包膜二十面体衣壳和病毒基因组组成, 所述病毒基因组由152kb线性双链DNA分子组成。在一个实施例中, 基于HSV的病毒载体缺乏一种或多种必需或非必需HSV基因。在一个实施例中, 基于HSV的病毒载体是复制缺陷型。大部分复制缺乏型HSV载体含有缺失以移除一种或多种中早期、早期或晚期HSV基因从而防止复制。举例来说, HSV载体可以缺乏选自由以下组成的群组的中早期基因: ICP4、ICP22、ICP27、ICP47和其组合。HSV载体的优势是其进入可以引起长期DNA表达的的潜伏期的能力以及其可以容纳多达25kb的外源DNA插入物的大型病毒DNA基因组。基于HSV的载体描述于例如美国专利案第5, 837, 532号、第5, 846, 782号和第5, 804, 413号, 以及国际专利申请案W0 91/02788、W0 96/04394、W0 98/15637和W0 99/06583中, 其各自以全文引用的方式并入本文中。

[0609] G. 基因组编辑的细胞

[0610] 通过具体实施例中所涵盖的方法制造的被基因组编辑的细胞提供改良的细胞疗法组合物。

[0611] 在各种实施例中, 被基因组编辑的细胞包含造血干细胞或祖细胞, 例如CD34⁺细胞。在具体实施例中, 被基因组编辑的造血干细胞或祖细胞在以下中包含多个编辑物: 一

个、两个、三个或更多个有助于抑制 γ -球蛋白基因表达和HbF的基因,包括(但不限于)BCL11A基因座、KLF1基因座、SOX6基因座、GATA1基因座和LSD1基因座; β -球蛋白基因座的地中海型贫血等位基因;或 β -球蛋白基因座的镰状化等位基因。

[0612] 在具体实施例中,编辑一个或多个造血干细胞或祖细胞中的多个目标位点的方法包含将多个被工程改造的核酸酶引入细胞群体;用包含供体修复模板的载体转导细胞群体,所述供体修复模板包含对应于基因组中的多个目标位点的多对同源臂,其中每对同源臂侧接一个或多个转基因;其中被工程改造的核酸酶的表达引起基因组中的目标位点处的双链断裂,并且供体修复模板的同源臂使得能够通过双链断裂(DSB)位点处的同源定向修复(HDR)将一个或多个转基因并入目标位点中。

[0613] 在具体实施例中,编辑一个或多个造血干细胞或祖细胞中的多个目标位点的方法包含将多个被工程改造的核酸酶引入细胞群体;用包含供体修复模板的载体转导细胞群体,所述供体修复模板包含对应于基因组中的多个目标位点的多对同源臂,其中每对同源臂侧接不同转基因;其中被工程改造的核酸酶的表达引起基因组中的目标位点处的双链断裂,并且供体修复模板的同源臂使得能够通过双链断裂(DSB)位点处的同源定向修复(HDR)将一个或多个转基因并入目标位点中。

[0614] 在具体实施例中,编辑一个或多个造血干细胞或祖细胞中的多个目标位点的方法包含将多个被工程改造的核酸酶引入细胞群体;用包含供体修复模板的载体转导细胞群体,所述供体修复模板包含对应于基因组中的多个目标位点的多对同源臂,其中每对同源臂侧接一个或多个转基因,其中至少两对同源臂侧接相同转基因;其中被工程改造的核酸酶的表达引起基因组中的目标位点处的双链断裂,并且供体修复模板的同源臂使得能够通过双链断裂(DSB)位点处的同源定向修复(HDR)将一个或多个转基因并入目标位点中。

[0615] 在各种实施例中,被基因组编辑的细胞包含免疫效应细胞,例如T细胞,其包含多个基因组编辑物。不希望受任何具体理论约束,相信由本文中所涵盖的方法制备的多重基因组编辑的免疫效应细胞具有优良特性,包括增加的改良的体内安全性、功效和持久性。

[0616] 在具体实施例中,被基因组编辑的细胞在以下中包含多个编辑物:编码TCR信号传导组件(例如TCR α 基因的恒定区)的基因;免疫系统检查点基因,包括(但不限于):PD-1、LAG-3、TIM-3、CTLA-4、BTLA、TIGIT、VISTA和KIR基因;或免疫抑制性信号传导组件,包括(但不限于):IL-10R α 、TGF β R1、TGF β R2、AHR、SGK1、TSC2、VHL、A2AR和CBLB。

[0617] 在具体实施例中,被基因组编辑的细胞包含插入至少两个选自由以下组成的群组的基因组目标位点中的被工程改造的抗原受体:TCR α 基因的恒定区、PD-1、LAG-3、TIM-3、CTLA-4、BTLA、TIGIT、VISTA、KIR基因、IL-10R α 、TGF β R1、TGF β R2、AHR、SGK1、A2AR、TSC2、VHL和CBLB。

[0618] 在具体实施例中,被基因组编辑的细胞包含插入TCR α 基因的恒定区和选自由以下组成的群组的一个或多个基因中的被工程改造的抗原受体:PD-1、LAG-3、TIM-3、CTLA-4、BTLA、TIGIT、VISTA、KIR基因、IL-10R α 、TGF β R1、TGF β R2、AHR、SGK1、A2AR、TSC2、VHL和CBLB。

[0619] 在具体实施例中,编辑一个或多个T细胞中的多个目标位点的方法包含活化T细胞群体和刺激T细胞群体以使其增殖;将多个被工程改造的核酸酶引入T细胞群体;用包含供体修复模板的载体转导T细胞群体,所述供体修复模板包含对应于基因组中的多个目标位点的多对同源臂,其中每对同源臂侧接一个或多个转基因;其中被工程改造的核酸酶的表

达引起基因组中的目标位点处的双链断裂,并且供体修复模板的同源臂使得能够通过双链断裂(DSB)位点处的同源定向修复(HDR)将一个或多个转基因并入目标位点中。

[0620] 在具体实施例中,编辑一个或多个T细胞中的多个目标位点的方法包含活化T细胞群体和刺激T细胞群体以使其增殖;将多个被工程改造的核酸酶引入T细胞群体;用包含供体修复模板的载体转导T细胞群体,所述供体修复模板包含对应于基因组中的多个目标位点的多对同源臂,其中每对同源臂侧接不同转基因;其中被工程改造的核酸酶的表达引起基因组中的目标位点处的双链断裂,并且供体修复模板的同源臂使得能够通过双链断裂(DSB)位点处的同源定向修复(HDR)将一个或多个转基因并入目标位点中。

[0621] 在具体实施例中,编辑一个或多个T细胞中的多个目标位点的方法包含活化T细胞群体和刺激T细胞群体以使其增殖;将多个被工程改造的核酸酶引入T细胞群体;用包含供体修复模板的载体转导T细胞群体,所述供体修复模板包含对应于基因组中的多个目标位点的多对同源臂,其中每对同源臂侧接一个或多个转基因,其中至少两对同源臂侧接相同转基因;其中被工程改造的核酸酶的表达引起基因组中的目标位点处的双链断裂,并且供体修复模板的同源臂使得能够通过双链断裂(DSB)位点处的同源定向修复(HDR)将一个或多个转基因并入目标位点中。

[0622] H. 组合物和配制物

[0623] 具体实施例中所涵盖的组合物可以包含本文所涵盖的一种或多种多肽、聚核苷酸、包含其的载体以及基因组编辑组合物和被基因组编辑的细胞组合物。具体实施例中所涵盖的基因组编辑组合物和方法适用于使用单一DNA供体修复模板编辑细胞或细胞群体中的多个目标位点。在优选实施例中,使用被工程改造的核酸酶和DNA供体修复模板在造血细胞(例如造血干细胞或祖细胞)、CD34⁺细胞、T细胞或免疫效应细胞的基因组中引入多个编辑物。

[0624] 在具体实施例中,本文中所涵盖的组合物包含DNA供体修复模板。

[0625] 在具体实施例中,本文中所涵盖的组合物包含编码DNA供体修复模板的载体。

[0626] 在各种实施例中,本文中所涵盖的组合物包含被工程改造的核酸酶和/或DNA供体修复模板。核酸酶变异体可以呈通过本文中所公开的聚核苷酸递送方法(例如电致孔、脂质纳米粒子、转导等)引入细胞中的mRNA形式。在一个实施例中,通过本文中所公开的聚核苷酸递送方法将组合物引入细胞中,所述组合物包含编码导向核酸内切酶变异体或megaTAL的mRNA和DNA供体修复模板。可以使用在供体修复模板存在下基因编辑酶的表达,通过HDR在基因组编辑的细胞或基因组编辑的细胞群体中的多个目标位点处引入基因组编辑物。在具体实施例中,细胞群体包含被基因修饰的免疫效应细胞。

[0627] 组合物包括(但不限于)医药组合物。“药物组合物”是指配制于药学上可接受的或生理学上可接受的溶液中以用于单独或与一种或多种其它治疗模式组合给予细胞或动物的组合物。还应理解,必要时,组合物还可以与其它试剂(例如细胞因子、生长因子、激素、小分子、化学治疗剂、前药、药物、抗体或其它各种医药学活性剂)组合给予。对于所述组合物中还可以包括的其它组分几乎无限制,只要这些其它试剂不会不利地影响所述组合物即可。

[0628] 短语“药学上可接受的”在本文中用于指在合理医学判断范围内,适用于与人类和动物的组织接触而无过量毒性、刺激、过敏反应或其它问题或并发症,与合理效益/风险比

相称的化合物、物质、组合物和/或剂型。

[0629] 术语“医药学上可接受的载剂”是指与治疗性细胞一起给予的稀释剂、佐剂、赋形剂或媒剂。药物载剂的说明性实例可以是无菌液体，如细胞培养基、水和油，包括石油、动物、植物或合成来源的油，如花生油、大豆油、矿物油、芝麻油等。生理食盐水溶液和右旋糖水溶液和甘油溶液也可以用作液体载剂，特别是用于可注射溶液。在具体实施例中，合适的药物赋形剂包括淀粉、葡萄糖、乳糖、蔗糖、明胶、麦芽、稻谷、面粉、白垩、硅胶、硬脂酸钠、单硬脂酸甘油酯、滑石、氯化钠、无水脱脂奶粉、甘油、丙烯、乙二醇、水、乙醇等。除非任何常规介质或试剂与活性成分不相容，否则涵盖将其用于治疗组合中。还可以将补充性活性成分并入组合中。

[0630] 在一个实施例中，包含医药学上可接受的载剂的组合物适于给予个体。在具体实施例中，包含载剂的组合物适合非经肠给药，例如血管内（静脉内或动脉内）、腹膜内或肌肉内给药。在具体实施例中，包含医药学上可接受的载剂的组合物适于室内、脊柱内或鞘内给药。医药学上可接受的载剂包括无菌水溶液、细胞培养基或分散液。这类介质和药剂用于医药学活性物质的用途是所属领域中众所周知的。除非任何常规介质或试剂与被转导的细胞不相容，否则涵盖将其用于药物组合中。

[0631] 在具体实施例中，本文中所涵盖的组合物包含被基因修饰的T细胞和医药学上可接受的载剂。包含本文中所涵盖的基于细胞的组合物的组合物可以通过肠内或非经肠给药方法单独给予或与其它适合的化合物组合给予以实现所需治疗目标。

[0632] 医药学上可接受的载剂必须具有足够高的纯度和足够低的毒性，以使其适合对所治疗的人类个体给药。其还应该保持或提高组合物的稳定性。医药学上可接受的载剂可以是液体或固体，并且考虑到计划的给药方式，当与组合物的其它组分组合时，将医药学上可接受的载剂选择成能提供所需松密度(bulk)、稠度等。举例来说，医药学上可接受的载剂可以是(但不限于)粘合剂(例如预胶凝化玉米淀粉、聚乙烯吡咯烷酮或羟丙基甲基纤维素等)、填充剂(例如乳糖和其它糖、微晶纤维素、果胶、明胶、硫酸钙、乙基纤维素、聚丙烯酸酯、磷酸氢钙等)、润滑剂(例如硬脂酸镁、滑石、二氧化硅、胶态二氧化硅、硬脂酸、金属硬脂酸盐、氢化植物油、玉米淀粉、聚乙二醇、苯甲酸钠、乙酸钠等)、崩解剂(例如淀粉、羟基乙酸淀粉钠等)或润湿剂(例如月桂基硫酸钠等)。其它适用于本文中所涵盖的组合物医药学上可接受的载剂包括(但不限于)水、盐溶液、醇、聚乙二醇、明胶、直链淀粉、硬脂酸镁、滑石、硅酸、粘性石蜡、羟甲基纤维素、聚乙烯基吡咯烷酮等。

[0633] 这类载剂溶液还可以含有缓冲剂、稀释剂和其它合适的添加剂。如本文中所使用，术语“缓冲剂”是指其化学组成能中和酸或碱但不引起pH值显著改变的溶液或液体。本文中所涵盖的缓冲剂的实例包括(但不限于)杜尔贝科氏磷酸盐缓冲生理食盐水(Dulbecco's phosphate buffered saline)(PBS)、林格氏溶液(Ringer's solution)、含5%右旋糖水(D5W)、正常/生理食盐水(0.9%NaCl)。

[0634] 医药学上可接受的载剂可以按足以保持组合物的pH值是约7的量存在。或者，组合物的pH值在约6.8到约7.4范围内，例如6.8、6.9、7.0、7.1、7.2、7.3和7.4。在另一实施例中，组合物的pH值是约7.4。

[0635] 本文中所涵盖的组合物可以包含无毒的医药学上可接受的介质。组合物可以是悬浮液。如本文中所使用，术语“悬浮液”是指非贴壁条件，其中细胞不附着于固体载体。举例

来说,可以搅拌或搅动保持悬浮状态的细胞并且使其不粘附到载体,如培养皿。

[0636] 在具体实施例中,本文中所涵盖的组合物被配制成悬浮液形式,其中被基因组编辑的细胞分散于静脉内(IV)袋等中的可接受的液体介质或溶液,例如生理食盐水或无血清培养基内。可接受的稀释剂包括(但不限于)水、PlasmaLyte、林格氏溶液、等张氯化钠(生理食盐水)溶液、无血清细胞培养基以及适用于低温储存的培养基,例如Cryostor®培养基。

[0637] 在某些实施例中,医药学上可接受的载剂基本上不含人类或动物来源的天然蛋白质,并且适于储存包含被基因组编辑的细胞群体的组合物。治疗组合物旨在给予人类患者,因此基本上不含细胞培养组分,如牛血清白蛋白、马血清和胎牛血清。

[0638] 在一些实施例中,医药学上可接受的细胞培养基中配制组合物。这类组合物适用于对人类个体给药。在具体实施例中,医药学上可接受的细胞培养基是无血清培养基。

[0639] 与含有血清的培养基相比,无血清培养基具有若干优点,包括组成简化且更明确、污染程度降低、消除潜在感染物来源并且成本更低。在各种实施例中,无血清培养基是无动物成分的,并且可以任选地是无蛋白质的。任选地,培养基可以含有生物药学上可接受的重组型蛋白质。“无动物成分”培养基是指其中组分来源于非动物来源的培养基。重组型蛋白质置换无动物成分培养基中的原生动物蛋白质且养分是从合成、植物或微生物来源获得。相比之下,“无蛋白质”培养基被定义为基本上不含蛋白质。

[0640] 具体组合物中所用的无血清培养基的说明性实例包括(但不限于)QBSF-60(Quality Biological, Inc.)、StemPro-34(Life Technologies)和X-VIVO 10。

[0641] 在优选实施例中,包含被基因组编辑的细胞的组合物是在PlasmaLyte中配制。

[0642] 在各种实施例中,包含被基因组编辑的细胞的组合物是在低温保存培养基中配制。例如,可以使用含有低温保存剂的低温保存培养基在解冻后维持高细胞活力结果。具体组合物中所用的低温保存培养基的说明性实例包括(但不限于)CryoStor CS10、CryoStor CS5和CryoStor CS2。

[0643] 在一个实施例中,组合物是在包含50:50PlasmaLyte A:CryoStor CS10的溶液中配制。

[0644] 在具体实施例中,组合物基本上不含支原体、内毒素和微生物污染物。相对于内毒素而言,“基本上不含”意指每剂细胞的内毒素低于FDA对生物制剂所允许的内毒素,即5EU/千克体重/天的总内毒素,对于平均70kg的人来说,即为350EU/总细胞剂量。在具体实施例中,包含用本文中所涵盖的逆转录病毒载体转导的造血干细胞或祖细胞的组合物含有约0.5EU/mL到约5.0EU/mL、或约0.5EU/mL、1.0EU/mL、1.5EU/mL、2.0EU/mL、2.5EU/mL、3.0EU/mL、3.5EU/mL、4.0EU/mL、4.5EU/mL或5.0EU/mL。

[0645] 在某些实施例中,所涵盖的适用于递送聚核苷酸的组合物和配制物包括(但不限于)编码一种或多种重编程的核酸酶以及任选地末端酶类酶的一个或多个mRNA。

[0646] 用于离体递送的例示性配制物还可以包括使用所属领域中已知的各种转染剂,如磷酸钙、电致孔、热休克以及各种脂质体配制物(即,脂质介导的转染)。脂质体是包裹一部分水性流体的脂质双层。DNA自发地与阳离子脂质体的外表面缔合(借助于其电荷),并且这些脂质体将与细胞膜相互作用。

[0647] 在具体实施例中,药学上可接受的载剂溶液的配制物,以及在多种治疗方案中使用本文中所描述的具体组合物的适合剂量和治疗方案的研究是所属领域的技术人员众所

周知的,包括例如肠内和非经肠,例如血管内、静脉内、动脉内、骨内、室内、脑内、颅内、脊柱内、鞘内和髓内给药和配制物。所属领域技术人员应理解,本文中所涵盖的具体实施例可以包含其它配制物,如医药学领域中众所周知的那些,并且描述于例如《雷明顿:制药科学和实践(Remington:The Science and Practice of Pharmacy)》,第I卷和第II卷,第22版,Loyd V.Allen Jr.编,Philadelphia,PA:Pharmaceutical Press;2012中,其以全文引用的方式并入本文中。

[0648] I. 被基因组编辑的细胞疗法

[0649] 由本文中所涵盖的组合物和方法制造的被基因组编辑的细胞提供用于预防、治疗或改善血红蛋白病、癌症、GVHD、感染性疾病、自身免疫疾病、炎症性疾病或免疫缺陷的至少一种症状的药物产品。如本文中所使用,术语“药物产品”是指使用本文中所涵盖的组合物和方法产生的被基因修饰的细胞。

[0650] 在具体实施例中,药物产品包含被基因编辑的造血干细胞或祖细胞或CD34⁺细胞。在具体实施例中,给予个体有效量的包含基因组中的多个编辑物的被基因组编辑的造血干细胞或祖细胞或CD34⁺细胞,以预防、治疗或改善血红蛋白病(包括(但不限于) β -地中海贫血和镰状细胞病)的至少一种症状。

[0651] 在具体实施例中,药物产品包含被基因编辑的免疫效应细胞或T细胞。此外,具体实施例中所涵盖的被基因组编辑的T细胞提供更安全且更有效的过继性细胞疗法,因为这些细胞对T细胞耗竭具有抗性并且在肿瘤微环境显示增加的耐久性和持久性,使得可以持续进行治疗。

[0652] 在具体实施例中,将有效量的包含基因组中的多个编辑物的被基因组编辑的免疫效应细胞或T细胞给予个体以预防、治疗或改善癌症、GVHD、感染性疾病、自身免疫疾病、炎症性疾病或免疫缺陷的至少一种症状。

[0653] 在具体实施例中,被基因组编辑的细胞包含插入基因中的多个目标位点中的一个或多个免疫效能增强子、免疫抑制信号减弱子或被工程改造的抗原受体,其增加对T细胞耗竭的抗性、提高体内T细胞耐久性和/或提高体内T细胞持久性。

[0654] 在具体实施例中,被基因组编辑的细胞包含插入选自以下组成的群组的多个目标位点中的CAR:TCR α 恒定区、免疫系统检查点基因和/或编码免疫抑制性信号传导组件的基因。

[0655] 在具体实施例中,被基因组编辑的细胞包含插入TCR α 恒定区中的CAR和一个或多个插入一个或多个免疫系统检查点基因和/或编码免疫抑制性信号传导组件的基因中的转基因,其编码免疫效能增强子、免疫抑制信号减弱子或另一被工程改造的抗原受体。

[0656] 在具体实施例中,被基因组编辑的细胞包含插入TCR α 恒定区、一个或多个免疫系统检查点基因和/或编码免疫抑制性信号传导组件的基因中的CAR或翻转受体。

[0657] 在具体实施例中,被基因组编辑的细胞包含一个或多个插入TCR α 恒定区、一个或多个免疫系统检查点基因和/或编码免疫抑制性信号传导组件的基因中的翻转受体。

[0658] 在具体实施例中,本文中所涵盖的被基因组编辑的细胞用于治疗实体肿瘤或癌症。

[0659] 在具体实施例中,本文中所涵盖的被基因组编辑的细胞是用于治疗实体肿瘤或癌症,包括(但不限于):肾上腺癌、肾上腺皮质癌、肛门癌、阑尾癌、星形细胞瘤、非典型畸胎

样/横纹肌样肿瘤、基底细胞癌、胆管癌、膀胱癌、骨癌、脑癌/CNS癌、乳癌、支气管肿瘤、心脏肿瘤、子宫颈癌、胆管癌、软骨肉瘤、脊索瘤、结肠癌、结肠直肠癌、颅咽管瘤、乳腺管原位癌(DCIS)、子宫内膜癌、室管膜瘤、食道癌、鼻腔神经胶质瘤、尤文氏肉瘤(Ewing's sarcoma)、颅外生殖细胞肿瘤、性腺外生殖细胞肿瘤、眼癌、输卵管癌、纤维组织肉瘤、纤维肉瘤、胆囊癌、胃癌、胃肠道类癌、胃肠道间质瘤(GIST)、生殖细胞肿瘤、神经胶质瘤、神经胶母细胞瘤、头颈癌、成血管细胞瘤、肝细胞癌、下咽癌、眼内黑素瘤、卡波西肉瘤(kaposi sarcoma)、肾癌、喉癌、平滑肌肉瘤、唇癌、脂肪肉瘤、肝癌、肺癌、非小细胞肺癌、肺类癌肿瘤、恶性间皮瘤、髓质癌、神经管胚细胞瘤、脑膜瘤、黑素瘤、梅克尔细胞癌(Merkel cell carcinoma)、中线消化道癌、嘴部癌症、粘液肉瘤、骨髓发育不良综合症、骨髓增生赘瘤、鼻腔和副鼻窦癌、鼻咽癌、神经母细胞瘤、少突神经胶质瘤、口腔癌(oral cancer)、口腔癌症(oral cavity cancer)、口咽癌、骨肉瘤、卵巢癌、胰腺癌、胰腺胰岛细胞瘤、乳头状癌、副神经节瘤、副甲状腺癌、阴茎癌、咽癌、嗜铬细胞瘤、松果体瘤、垂体肿瘤、胸膜肺母细胞瘤、原发性腹膜癌、前列腺癌、直肠癌、视网膜母细胞瘤、肾细胞癌、肾盂和输尿管癌、横纹肌肉瘤、唾液腺癌、皮脂腺癌、皮肤癌、软组织肉瘤、鳞状细胞癌、小细胞肺癌、小肠癌、胃癌、汗腺癌、滑膜瘤、睾丸癌、喉癌、胸腺癌、甲状腺癌、尿道癌、子宫癌、子宫肉瘤、阴道癌、血管癌、外阴癌和威尔姆斯氏肿瘤。

[0660] 在具体实施例中,本文中所涵盖的被基因组编辑的细胞是用于治疗实体肿瘤或癌症,包括(但不限于)肝癌、胰腺癌、肺癌、乳癌、膀胱癌、脑癌、骨癌、甲状腺癌、肾癌或皮肤癌。

[0661] 在具体实施例中,本文中所涵盖的被基因组编辑的细胞是用于治疗各种癌症,包括(但不限于)胰腺癌、膀胱癌和肺癌。

[0662] 在具体实施例中,本文中所涵盖的被基因组编辑的细胞是用于治疗液体癌症或血液癌。

[0663] 在具体实施例中,本文中所涵盖的被基因组编辑的细胞是用于治疗B细胞恶性病,包括(但不限于):白血病、淋巴瘤和多发性骨髓瘤。

[0664] 在具体实施例中,本文中所涵盖的被基因组编辑的细胞是用于治疗液体癌症,包括(但不限于)白血病、淋巴瘤和多发性骨髓瘤:急性淋巴细胞性白血病(ALL)、急性骨髓性白血病(AML)、骨髓母细胞白血病、早幼粒细胞性白血病、骨髓单核细胞性白血病、单核细胞性白血病、红白血病、毛细胞白血病(HCL)、慢性淋巴细胞性白血病(CLL)和慢性骨髓白血病(CML)、慢性骨髓单核细胞性白血病(CMML)和真性红细胞增多症、霍奇金氏淋巴瘤(Hodgkin lymphoma)、结节性淋巴细胞为主型霍奇金氏淋巴瘤、伯基特氏淋巴瘤(Burkitt lymphoma)、小淋巴细胞性淋巴瘤(SLL)、弥漫性大型B细胞淋巴瘤、滤泡性淋巴瘤、成免疫细胞性大细胞淋巴瘤、前体B成淋巴细胞性淋巴瘤、套细胞淋巴瘤、边缘区淋巴瘤、蕈样真菌病、多形性大细胞淋巴瘤、塞扎莱氏综合症(Sézary syndrome)、前体T成淋巴细胞性淋巴瘤、多发性骨髓瘤、明显多发性骨髓瘤、郁积性多发性骨髓瘤、浆细胞白血病、非分泌性骨髓瘤、IgD骨髓瘤、骨硬化性骨髓瘤、孤立性骨浆细胞瘤以及髓外浆细胞瘤。

[0665] 用于本文中所涵盖的基因组编辑方法中的优选细胞包括自体/自体性(“自身”)细胞,优选造血细胞。

[0666] 在具体实施例中,提供一种方法,其包含向有需要的患者单独或与一种或多种治

疗剂组合给予治疗有效量的本文中所涵盖的被基因组编辑的细胞或包含这些细胞的组合物。在某些实施例中,细胞是用于治疗具有产生血红蛋白病、癌症、GVHD、感染性疾病、自身免疫疾病、发炎性疾病或免疫缺陷的风险的患者。因此,具体实施例包含治疗或预防或改善血红蛋白病、癌症、感染性疾病、自身免疫疾病、发炎性疾病或免疫缺陷的至少一种症状,其包含向有需要的个体给予治疗有效量的本文中所涵盖的被基因组编辑的细胞。

[0667] 在一个实施例中,治疗有需要的个体中的血红蛋白病、癌症、GVHD、感染性疾病、自身免疫疾病、发炎性疾病或免疫缺陷的方法包含给予有效量,例如治疗有效量的包含本文中所涵盖的被基因组编辑的细胞的组合物。给药量和频率由如患者的病状以及患者疾病的类型和严重程度等因素决定,但适当剂量可以通过临床试验确定。

[0668] 在一个说明性实施例中,提供给个体的被基因组编辑的细胞的有效量是至少 2×10^6 个细胞/千克、至少 3×10^6 个细胞/千克、至少 4×10^6 个细胞/千克、至少 5×10^6 个细胞/千克、至少 6×10^6 个细胞/千克、至少 7×10^6 个细胞/千克、至少 8×10^6 个细胞/千克、至少 9×10^6 个细胞/千克或至少 10×10^6 个细胞/千克,或每千克更多个细胞,包括所有中间细胞剂量。

[0669] 在另一个说明性实施例中,提供给个体的被基因组编辑的细胞的有效量是约 2×10^6 个细胞/千克、约 3×10^6 个细胞/千克、约 4×10^6 个细胞/千克、约 5×10^6 个细胞/千克、约 6×10^6 个细胞/千克、约 7×10^6 个细胞/千克、约 8×10^6 个细胞/千克、约 9×10^6 个细胞/千克或约 10×10^6 个细胞/千克,或每千克更多个细胞,包括所有中间细胞剂量。

[0670] 在另一个说明性实施例中,提供给个体的被基因组编辑的细胞的有效量是约 2×10^6 个细胞/千克到约 10×10^6 个细胞/千克、约 3×10^6 个细胞/千克到约 10×10^6 个细胞/千克、约 4×10^6 个细胞/千克到约 10×10^6 个细胞/千克、约 5×10^6 个细胞/千克到约 10×10^6 个细胞/千克、 2×10^6 个细胞/千克到约 6×10^6 个细胞/千克、 2×10^6 个细胞/千克到约 7×10^6 个细胞/千克、 2×10^6 个细胞/千克到约 8×10^6 个细胞/千克、 3×10^6 个细胞/千克到约 6×10^6 个细胞/千克、 3×10^6 个细胞/千克到约 7×10^6 个细胞/千克、 3×10^6 个细胞/千克到约 8×10^6 个细胞/千克、 4×10^6 个细胞/千克到约 6×10^6 个细胞/千克、 4×10^6 个细胞/千克到约 7×10^6 个细胞/千克、 4×10^6 个细胞/千克到约 8×10^6 个细胞/千克、 5×10^6 个细胞/千克到约 6×10^6 个细胞/千克、 5×10^6 个细胞/千克到约 7×10^6 个细胞/千克、 5×10^6 个细胞/千克到约 8×10^6 个细胞/千克或 6×10^6 个细胞/千克到约 8×10^6 个细胞/千克,包括所有中间细胞剂量。

[0671] 所属领域的一般技术人员应认识到,可能需要多次给予具体实施例中所涵盖的组合物来实现所需疗法。举例来说,组合物可以跨越1周、2周、3周、1个月、2个月、3个月、4个月、5个月、6个月、1年、2年、5年、10年或更久的时间间隔给予1、2、3、4、5、6、7、8、9或10次或更多次。

[0672] 在某些实施例中,可能需要向个体给予被活化的T细胞且随后相继地再抽取血液(或进行析离术),活化其中的T细胞,并且将这些被活化和扩增的T细胞再输注给患者。这一过程可以每隔数周进行多次。在某些实施例中,可以活化10cc到400cc的抽血中的T细胞。在某些实施例中,活化20cc、30cc、40cc、50cc、60cc、70cc、80cc、90cc、100cc、150cc、200cc、250cc、300cc、350cc或400cc或更多的抽血中的T细胞。不受理论约束,使用这种多次抽血/多次再输注方案可以用于选出某些T细胞群体。

[0673] 具体实施例中所涵盖的组合物的给药可以按任何适宜方式进行,包括通过气溶胶

吸入、注射、摄入、输注、植入或移植。在优选实施例中，组合物是非经肠给予。如本文中所使用，短语“非经肠给药”和“非经肠给予”是指除经肠和局部给药以外的给药模式，通常通过注射并且包括（但不限于）血管内、静脉内、肌肉内、动脉内、鞘内、囊内、眶内、瘤内、心内、皮内、腹膜内、经气管、皮下、表皮下、关节内、囊下、蛛网膜下、脊椎内和胸骨内注射和输注。在一个实施例中，本文中所涵盖的组合物通过直接注射到肿瘤、淋巴结或感染位点中来给予个体。

[0674] 在一个实施例中，治疗诊断患有血红蛋白病的个体的方法包含从个体移出外周血液细胞或骨髓细胞、编辑所述细胞的基因组并且产生被基因组编辑的造血干细胞或祖细胞群体以及向同一名个体给予被基因组编辑的细胞群体。在优选实施例中，免疫效应细胞包含CD34⁺细胞。

[0675] 在一个实施例中，治疗诊断患有癌症的个体的方法包含从个体移出免疫效应细胞、编辑所述免疫效应细胞的基因组并且产生被基因组编辑过的免疫效应细胞群体，以及向同一名个体给予被基因组编辑的免疫效应细胞群体。在优选实施例中，免疫效应细胞包含T细胞。

[0676] 用于给予具体实施例中涵盖的细胞组合物的方法包括可以有效再引入被离体基因组编辑的细胞，或再引入被基因组编辑的细胞的祖细胞，使其在引入个体后分化成成熟细胞的任何方法。一种方法包含离体基因组编辑外周血液细胞或骨髓细胞并且将被转导的细胞递送回个体体内。

[0677] 本说明书中所引用的所有公开案、专利申请案和授权专利案都以引用的方式并入本文中，其引用的程度如同每个个别公开案、专利申请案或授权专利案经特定并且独立地指示以引用的方式并入一般。

[0678] 尽管已经出于清楚理解的目的，通过说明和实例相当详细地描述了前述实施例，但根据本文所涵盖的教导内容，所属领域的普通技术人员将易于了解，可以在不脱离所附权利要求书的精神或范围的情况下对其作出某些改变和修改。提供以下实例仅作为说明并且不具有限制性。所属领域的技术人员将容易地识别出多种可被改变或修改以产生基本类似结果的非关键参数。

[0679] 实例

[0680] 实例1

[0681] 使用单一AAV HDR DNA供体修复模板将转基因靶向整合到两个不同基因座中

[0682] 设计和构筑腺相关病毒(AAV)质粒，其包含启动子、编码绿色荧光蛋白质的转基因和聚腺苷酸化信号(SEQ ID NO:1)。用XmaI消化确认AAV ITR元件的完整性。转基因盒由与人类程序性死亡受体1(PD-1)同源的区域和T细胞受体α(TCRα)基因的恒定区侧接。PD-1同源臂含有侧接PD-1外显子1中的megaTAL裂解位点的300bp区域。TCRα同源臂含有侧接TCRα基因的恒定区中的megaTAL裂解位点的300bp区域。串联安置同源臂，使得TCRα同源臂紧邻转基因盒的5'和3'安置，而PD-1同源臂紧邻TCRα同源臂的5'和3'安置(图3A)。例示性转基因表达盒含有骨髓增生肉瘤病毒增强子、负调控区缺失、d1587rev引物结合位点取代(MND)的启动子，所述启动子可操作地连接到编码荧光多肽(例如蓝色荧光蛋白质(BFP)、红色荧光蛋白质(RFP)、青色荧光蛋白质(CFP)、绿色荧光蛋白质(GFP)等)的聚核苷酸。表达盒还含有SV40晚期聚腺苷酸化信号。

[0683] 通过用一个或多个提供所需复制、衣壳和腺病毒辅助元件的质粒短暂共转染 HEK293T 细胞来制备重组型 AAV (rAAV)。使用超离心法,以基于碘克沙醇 (iodixanol) 的梯度从被共转染的 HEK 293T 细胞培养物纯化 rAAV。

[0684] 在用 CD3 和 CD28 活化并且在补充有 IL-2 的完全培养基中培养的原代人类 T 细胞中评估 MegaTAL 诱导的同源重组。在 3 天之后,洗涤 T 细胞并且用体外被转录的 mRNA 进行电致孔,所述 mRNA 编码靶向 PD-1 基因的外显子 1 的 megaTAL、靶向 TCR α 基因的恒定区的 megaTAL 或同时编码这两种 megaTAL。接着,用被纯化的重组型 AAV 靶向载体转导 T 细胞,所述载体包含编码 MND-GFP 转基因盒的 DNA 供体修复模板。对照物包括用单独的 rAAV 靶向载体处理的 T 细胞。在多个时间点使用流式细胞术以测量表达荧光蛋白质的 T 细胞的出现率以及区分未被整合的 rAAV 靶向载体的荧光蛋白质的瞬时表达。通过针对 CD3 表达进行染色来检测 MegaTAL 介导的 TCR α 破坏,而通过测序和在多克隆 T 细胞活化后 PD-1 表达的丧失来检测 PD-1 基因的破坏。

[0685] 在 40-70% 的用 megaTAL 和 rAAV 靶向载体处理的 T 细胞中观察到长期转基因表达 (图 3B)。对照样品产生不同程度的短暂荧光蛋白质表达和极少量 (<1%) 长期荧光蛋白质表达,与基因组中的最小随机整合一致。MegaTAL 介导的 TCR α 基因破坏在 70-90% 之间的范围内 (图 4),而 megaTAL 介导的 PD-1 基因座破坏在 60% 到 80% 范围内 (图 5)。每个基因座处的裂解是独立的,并且组合靶向 PD-1 和 TCR α 的 megaTAL 不会显著影响使用任一种单独的 megaTAL 时 PD-1 或 TCR α 目标位点处的裂解率。用相同的 rAAV 同源模板,在 PD-1 或 TCR α 基因座处观察到类似的同源重组率。组合用 PD-1 和 TCR α megaTAL 与 rAAV 模板进行的处理可以产生 60-70% 稳定 GFP 表达,和增加的 GFP 表达可变性。增加的 GFP 表达可变性指示两个不同遗传位置中的异源程度更高的转基因整合模式。结果在针对从若干独立供体分离的 T 细胞进行的实验中得到证实。

[0686] 实例 2

[0687] 使用单一 AAV HDR DNA 供体修复模板将 BCMA 特异性嵌合抗原受体靶向整合到两个不同基因座中

[0688] 设计、构筑和检验腺相关病毒 (AAV) 质粒,其含有启动子、编码嵌合抗原受体 (CAR) 的转基因和聚腺苷酸化信号 (SEQ ID NO. 2)。CAR 表达盒含有 MND 启动子,其可操作地连接到包含以下的 CAR: CD8 α 衍生的信号肽、靶向 B 细胞成熟抗原 (BCMA) 的单链可变片段 (scFv)、CD8 α 衍生的铰链区和跨膜结构域、细胞内 4-1BB 共刺激域和 CD3 ζ 信号传导结构域。串联安置同源臂,引起 TCR α 同源臂紧邻转基因盒的 5' 和 3' 安置,而 PD-1 同源臂紧邻 TCR α 同源臂的 5' 和 3' 安置 (图 6A)。

[0689] 用 CD3 和 CD28 活化原代人类 T 细胞,如实例 1 中所描述。使用被活化的原代人类 T 细胞评估使用编码抗 BCMA CAR 的 AAV 供体修复模板的 PD-1 和 TCR α 基因的 MegaTAL 诱导的同源定向修复,所述被活化的原代人类 T 细胞用体外被转录的 mRNA 电致孔,所述 mRNA 编码靶向 PD-1 基因的外显子 1 的 megaTAL、靶向 TCR α 基因的恒定区的 megaTAL 或同时靶向这两种 megaTAL。被电致孔的 T 细胞用 AAV 靶向载体转导并且在 IL2 存在下,在 30°C 下培养过夜且接着转移到 37°C 下,所述 AAV 靶向载体包含编码抗 BCMA CAR 的 DNA 供体修复模板。在电致孔之后 7 天进行 CAR 染色 (总共培养 10 天)。对照物包括含有单独的 megaTAL 或 AAV 处理物的 T 细胞。通过用 PE 结合的 BCMA-Fc 染色,通过流式细胞术分析抗 BCMA CAR 表达。

[0690] 用靶向megaTAL mRNA的PD-1和编码抗BCMA CAR的rAAV DNA供体修复模板处理的T细胞在20-30%的全部细胞中显示CAR表达,而用靶向megaTAL的TCR α 和编码抗BCMA CAR的rAAV DNA供体修复模板处理的细胞在30-50%的全部细胞中显示CAR表达。用PD-1和TCR α megaTAL进行的组合处理在40-55%的细胞中引起抗BCMA CAR表达(图6B)。在用PD-1和靶向megaTAL的TCR α 处理的细胞中观察到CAR表达的更多变化,表明两个不同遗传位置中异源程度更高的转基因整合模式。在未被处理、用AAV处理和用megaTAL/rAAV抗BCMA CAR处理的T细胞之间观察到类似的T细胞扩增率和类似的T细胞表型。

[0691] 通常,在以下权利要求书中,所使用的术语不应理解为将权利要求书限制于本说明书和权利要求书中所公开的具体实施例,但应理解为包括所有可能的实施例以及这份权利要求书所有权获得的等效物的全部范围。因此,权利要求书不受本发明限制。

序列表

<110> 蓝鸟生物公司 (bluebird bio, Inc.)

李百胜 (Lee, Baekseung)

亚历山大·阿斯特拉罕 (Astrakhan, Alexander)

<120> 供体修复模板多重基因组编辑

<130> BLBD-084/01W0 315698-2640

<150> US 62/459,203

<151> 2017-02-15

<160> 39

<170> PatentIn version 3.5

<210> 1

<211> 7317

<212> DNA

<213> 人工序列 (Artificial Sequence)

<220>

<223> 合成序列pBW1411

<400> 1

```

cagctgcgcg ctcgctcgt cactgaggcc gcccgggcaa agcccgggcg tcgggcgacc 60
tttggtcgcc cggcctcagt gagcgagcga gcgcgcagag agggagtggc caactccatc 120
actaggggtt ccttgtagtt aatgattaac ccgcatgct acttatctac gtaaggctgt 180
tgcaggcatc acacggtgga aagatctgga actgtggcca tgggtgtgagg ccatccacaa 240
ggtggaagct ttgaggggga gccgattagc catggacagt tgtcattcag tagggtcacc 300
tgtgccccag cgaaggggga tgggccggga aggcagaggc caggcacctg cccccagcag 360
gggcagaggc tgtgggcagc cgggaggctc ccagaggctc cgacagaatg ggagtggggt 420
tgagcccacc cctcactgca gcccaggaac ctgagcccag agggggccac ccaccttccc 480
caggcagggga ggcccggccc ccaggagat gggggggatg ggggaggaga agggcctgcc 540
cccacccggc agcctcagga ggggcagctc gggcgggata tggaaagagg ccacagcagt 600
gagcagagac acagaggagg aaggggccct gagctgggga gacccccacg gggtagggcg 660
tgggggccac gggcccacct cctccccatc tctctgtct cctgtctct gtctctctct 720
ccctccccca cctctcccc agtcttacc cctcttacc cctctcccc cagcactgcc 780
tctgtcactc tcgcccacgt ggatgtggag gaagaggggg cgggagcaag gggcgggcac 840
cctcccttca acctgacctg ggacagttc ctttcgctc acctccgct gagcagtgga 900
gaaggcggca ctctggtggg gctgtctcaa atcatggcct cttggccaag attgatagct 960
tgtgcctgtc cctgagtccc agtccatcac gagcagctgg tttctaagat gctatttccc 1020
gtataaagca tgagaccgtg acttgccagc cccacagagc cccgcccttg tccatcactg 1080
gcactctggac tccagcctgg gttggggcaa agagggaaat gagatcatgt cctaaccctg 1140
atcctcttgt cccacagata tccagaacct tgacctgcc gtgtaccagc tgagagactc 1200
taaatccagt gacaagtctg tctgcctatc gcgtgaacag agaaacagga gaatatgggc 1260

```

caaacaggat atctgtggta agcagttcct gccccggctc agggccaaga acagttggaa 1320
cagcagaata tgggccaac aggatatctg tggtaaagcag ttctgcccc ggctcagggc 1380
caagaacaga tgggtccccag atgcgggtccc gccctcagca gtttctagag aaccatcaga 1440
tgtttccagg gtgccccaaag gacctgaaat gaccctgtgc cttatttgaa ctaaccaatc 1500
agttcgcctt tcgcttctgt tcgcgcgctt ctgctcccc agctctatat aagcagagct 1560
cgtttagtga accgtcagat cgcctggaga cgccatccac gctgttttga cttccataga 1620
aggatctcga ggccaccatg gtgagcaagg gcgaggagct gttcaccggg gtgggtgccc 1680
tcctggtcga gctggacggc gacgtaaagc gccacaagtt cagcgtgtcc ggcgaggggc 1740
agggcgatgc cacctacggc aagctgacc tgaagttcat ctgcaccacc ggcaagctgc 1800
ccgtgccctg gcccaccctc gtgaccacc tgacctagg cgtgcagtgc ttcagccgct 1860
accccgacca catgaagcag cacgacttct tcaagtcgc catgcccga ggctacgtcc 1920
aggagcgcac catcttcttc aaggacgac gcaactaca gaccgcgcc gaggtgaagt 1980
tcgagggcga caccctgggt aaccgcatcg agctgaagg catcgactt aaggaggacg 2040
gcaacatcct ggggcacaag ctggagtaca actacaacag ccacaacgtc tatatcatgg 2100
ccgacaagca gaagaacggc atcaaggtga acttcaagat ccgccacaac atcgaggacg 2160
gcagcgtgca gctcggcag cactaccagc agaacacccc catcggcgac ggccccgtgc 2220
tgctgcccga caaccactac ctgagcacc agtccgcct gagcaaagac cccaacgaga 2280
agcgcgatca catggtcctg ctggagtctg tgaccgccg cgggatcact ctcgcatgg 2340
acgagctgta caagtaagc gccgcgctt atttgtgaa tttgtgatgc tattgcttta 2400
tttgtaacca ttataagctg caataaaca gttaacaaca acaattgcat tcattttatg 2460
tttcaggttc agggggagat gtgggaggt ttttaaagc tctttgattc tcaaaaaat 2520
gtgtcacaaa gtaaggattc tgatgtgtat atcacagaca aaactgtgct agacatgagg 2580
tctatggact tcaagagcaa cagtgtgtg gcctggagca acaaatctga ctttgcatgt 2640
gcaaacgcct tcaacaacag cattattcca gaagacacct tcttcccag cccaggtaag 2700
ggcagctttg gtgccttcgc aggctgtttc cttgcttcag gaatggccag gttctgccc 2760
gagctctggc caatgatgtc taaaactcct ctgattgggt gtaccggtt tggcggtgc 2820
tacaactggg ctggcggcca ggatggttct taggtagggt gggtcggcgg tcagggttc 2880
cagagccagg ggtctggagg gacctccac cctcagtccc tggcaggctg ggggtgctg 2940
aggcgggcct ggccctggca gccagggggt cccggagcga ggggtctgga gggaccttc 3000
actctcagtc cctggcaggt cggggggtgc tgtggcaggc ccagccttg ccccagctc 3060
tgccccttac cctgagctgt gtggctttgg gcagctcga ctctgggtt cctctctggg 3120
ccccaaactc tcccctggcc caagtcctt cttgtctct gggcaggcag gacctctgtc 3180
ccctctcagc cggctccttg ggctgcgtgt ttctgtagaa tgacgggtca ggctggccag 3240
aaccceaaac cttggcctg gggagtctgc gtggcggtc tgccctgccc aggcatectt 3300
ggctcctact cgagtttcc taaggatggg atgagccca tgtgggacta acctggctt 3360
tacgacgtca aagtttagat gagctgggtg ttttttctc attatatcca aagtgtacct 3420
gttcgagtga ggacagttct tctgtctcca ggatccctc tgggtgggga ttgtgcccgc 3480
ctgggtctct gccagattc cagggtctc cccgagcct gttcagacca tccgtggggg 3540
aggccttggc ctcaactcta cgtagataag tagcatggcg ggttaatcat taactacaag 3600

gaaccctag tgatggagtt ggccactccc tctctgcegc ctcgctcgct cactgaggcc 3660
 gggcgaccaa aggtcgcccg acgcccgggc tttgcccggg cggcctcagt gagcgagcga 3720
 gcgcgccagc tggcgtaata gcgaagaggc ccgcaccgat cgcccttccc aacagttgcg 3780
 cagcctgaat ggcgaaatggc gattccgttg caatggctgg cggtaaatatt gttctggata 3840
 ttaccagcaa ggccgatagt ttgagttctt ctactcaggc aagtgatgtt attactaatc 3900
 aaagaagtat tgcgacaacg gttaatttgc gtgatggaca gactctttta ctcggtggcc 3960
 tcaactgatta taaaaaact tctcaggatt ctggcgtacc gttcctgtct aaaatccctt 4020
 taatcggcct cctgtttagc tcccgtctg attctaacga ggaaagcacg ttatacgtgc 4080
 tcgtcaaagc aaccatagta cgcgcctgt agcggcgcac taagcgcggc ggggtgtggg 4140
 gttacgcgca gcgtgaccgc tacacttgc agcgcctag cgcccgtcc tttcgctttc 4200
 ttccccttct tctcgcacc gttcgcggc tttcccgc taaactctaa tcgggggctc 4260
 cctttagggt tccgatttag tgetttacgg cacctcgacc ccaaaaaact tgattagggt 4320
 gatggttcac gtagtgggccc atcgccctga tagacggttt ttcgcccctt gacgttggag 4380
 tccacgttct ttaatagtgg actcttgttc caaactggaa caaactcaa ccctatctcg 4440
 gtctattctt ttgatttata agggatttt ccgatttcgg cctattgggt aaaaaatgag 4500
 ctgatttaac aaaaatttaa cgcgaaattt aacaaaatat taacgtttac aatttaaata 4560
 tttgcttata caatcttctt gtttttgggg cttttctgat tatcaaccgg ggtacatatg 4620
 attgacatgc tagttttacg attaccgttc atcgattctc ttgtttgctc cagactctca 4680
 ggcaatgacc tgatagcctt tgtagagacc tctcaaaaat agctaccctc tccggcatga 4740
 atttatcagc tagaacggtt gaatatcata ttgatggtga tttgactgtc tccggccttt 4800
 ctcaccggtt tgaatcttta cctacacatt actcaggcat tgcatttaa atatatgagg 4860
 gttctaaaaa tttttatcct tgcgttgaat taaaggcttc tcccgcaaaa gtattacagg 4920
 gtcataatgt ttttggtaca accgatttag ctttatgctc tgaggcttta ttgcttaatt 4980
 ttgctaattc tttgccttgc ctgtatgatt tattggatgt tggaatcgcc tgatgcggta 5040
 ttttctcctt acgcatctgt gcggtatttc acaccgcata tgggtgactc tcagtacaat 5100
 ctgctctgat gccgcatagt taagccagcc ccgacaccg ccaacaccg ctgacgcgcc 5160
 ctgacgggct tgtctgctcc cggcatccgc ttacagacaa gctgtgaccg tctccgggag 5220
 ctgcatgtgt cagaggtttt cacctcacc accgaaacgc gcgagacgaa agggcctcgt 5280
 gatacgccta tttttatagg ttaatgtcat gataataat gtttcttaga cgtcagggtg 5340
 cacttttcgg ggaaatgtgc gcggaacccc tatttgttta tttttctaaa tacattcaaa 5400
 tatgtatccg ctcatgagac aataaccctg ataatgctt caataatatt gaaaaaggaa 5460
 gagtatgagt attcaacatt tccgtgtcgc cttatctcc ttttttgcgg cattttgcct 5520
 tcctgttttt gctcaccag aaacgctggt gaaagtaaaa gatgctgaag atcagttggg 5580
 tgcacgagtg ggttacatcg aactggatct caacagcggc aagatcctg agagttttcg 5640
 ccccgaaaga cgttttccaa tgatgagcac ttttaaagtt ctgctatgtg gcgcggtatt 5700
 atcccgtatt gagcgcgggc aagagcaact cggtcgccgc atacactatt ctcagaatga 5760
 cttggttgag tactcaccag tcacagaaaa gcactctac gatggcatga cagtaagaga 5820
 attatgcagt gctgccataa ccatgagtga taactctgc gccacttac ttctgacaac 5880
 gatcggagga ccgaaggagc taaccgcttt tttgcacaac atgggggatc atgtaactcg 5940

ccttgatcgt tgggaaccgg agctgaatga agccatacca aacgacgagc gtgacaccac 6000
 gatgcctgta gcaatggcaa caacgttgcg caaactatta actggcgaac tacttactct 6060
 agcttcccgg caacaattaa tagactggat ggaggcggat aaagttgcag gaccacttct 6120
 gcgctcggcc cttccggctg gctggtttat tgctgataaa tctggagccg gtgagcgtgg 6180
 gtctcgcggt atcattgcag cactggggcc agatggtaag ccctcccgta tcgtagttat 6240
 ctacacgacg gggagtcagg caactatgga tgaacgaaat agacagatcg ctgagatagg 6300
 tgccctactg attaagcatt ggtaactgtc agaccaagtt tactcatata tactttagat 6360
 tgatttaaaa cttcattttt aatttaaaag gatctaggtg aagatccttt ttgataatct 6420
 catgacaaaa atcccttaac gtgagttttc gttccactga gcgtcagacc ccgtagaaaa 6480
 gatcaaagga tcttcttgag atcctttttt tctgcgcgta atctgctgct tgcaaacaaa 6540
 aaaaccaccg ctaccagcgg tggtttgttt gccggatcaa gagctaccaa ctctttttcc 6600
 gaaggtaact ggcttcagca gagcgcagat accaaatact gtccttctag tgtagccgta 6660
 gttagccac cacttcaaga actctgtagc accgcctaca tacctcgtc tgctaatect 6720
 gttaccagtg gctgctgcca gtggcgataa gtcgtgtctt accgggttg actcaagacg 6780
 atagttaccg gataaggcgc agcggtcggg ctgaacgggg ggttcgtgca cacagcccag 6840
 cttggagcga acgacctaca ccgaactgag atacctacag cgtgagctat gagaaagcgc 6900
 cacgcttccc gaagggagaa aggcggacag gtatccgta agcggcaggg tcggaacagg 6960
 agagcgcacg agggagcttc cagggggaaa cgcttggtat ctttatagtc ctgtcgggtt 7020
 tcgccacctc tgacttgagc gtcgattttt gtgatgctcg tcaggggggc ggagcctatg 7080
 gaaaaacgcc agcaacgcgg cttttttacg gttcctggcc ttttgctggc cttttgctca 7140
 catgttcttt cctgcgttat cccctgattc tgtggataac cgtattaccg cttttgagtg 7200
 agctgatacc gctcggcga gccgaacgac cgagcgcagc gagtcagtga gcgaggaagc 7260
 ggaagagcgc ccaatacgca aaccgcctct ccccgcgcgt tggccgattc attaatg 7317

<210> 2

<211> 7413

<212> DNA

<213> 人工序列(Artificial Sequence)

<220>

<223> 合成序列pBW1490

<400> 2

cagctgcgcg ctcgctcgt cactgaggcc gcccgggcaa agcccgggcg tcgggcgacc 60
 tttggtcgcc cggcctcagt gagcgcgca gcgcgcagag agggagtggc caactccatc 120
 actaggggtt cctttagatt aatgattaac ccgcatgct acttatctac gtaatggggg 180
 aggagaagg cctgccccca cccggcagcc tcaggagggg cagctcgggc gggatatgga 240
 aagaggccac agcagtgagc agagacacag aggaggaagg ggccctgagc tggggagacc 300
 cccacggggg agggcgtggg ggccacgggc ccacctctc cccatctct ctgtctcct 360
 gtctctgtct ctctctcct cccccacct ctccccagtc ctaccctc ctcaccctc 420
 ctccccagc actgcctctg tctctctgc ccacgtggat gtggaggaag agggggcggg 480
 agcaaggggc gggcaccctc cttcaacct gacctgggac agtttccctt ccgctcacct 540

ccgcctgagc agtggagaag gcggcactct ggtggggctg ctccaaatca tggcctcttg 600
gccaaagattg atagcttgtg cctgtccctg agtcccagtc catcacgagc agctggtttc 660
taagatgcta tttcccgtat aaagcatgag accgtgactt gccagcccca cagagccccg 720
cccttgtcca tcaactggcat ctggactcca gcctggggtg gggcaaagag ggaaatgaga 780
tcatgtccta accctgatcc tcttgtccca cagatatcca gaaccctgac cctgccgtgt 840
accagctgag agactctaaa tccagtgaca agtctgtctg cctatacgcg taatgaaaga 900
ccccacctgt aggtttggca agctaggatc aaggttagga acagagagac agcagaatat 960
gggccaaca gataatctgt ggtaagcagt tctgccccg gctcagggcc aagaacagtt 1020
ggaacagcag aatatgggcc aaacaggata tctgtggtaa gcagttcctg ccccggtca 1080
gggccaagaa cagatggtec ccagatgcgg tcccgcctc agcagtttct agagaacct 1140
cagatgttcc cagggtgccc caaggacctg aatgacct gtgccttatt tgaactaacc 1200
aatcagttcg cttctcgett ctgttcgcgc gcttctgctc cccgagctca ataaaagagc 1260
ccacaacccc tcaactcgcg cgattcacct gcgctctac gccaccatgg cactccccgt 1320
caccgccctt ctcttgcctc tcgcctgct gctgcatgct gccaggcccc acattgtgct 1380
cactcagtca cctcccagcc tggccatgag cctgggaaaa agggccacca tctcctgtag 1440
agccagtgag tccgtcacia tcttggggag ccatcttatt cactggtatc agcagaagcc 1500
cgggcagcct ccaaccctc ttattcagct cgcgtcaaac gtccagacgg gtgtacctgc 1560
cagattttct ggtagcgggt cccgcaactg ttttactg accatagatc cagtggaaga 1620
agacgatgtg gccgtgtatt attgtctgca gagcagaacg attcctcgca catttggtgg 1680
gggtactaag ctggagatta agggaagcac gtccggctca ggaagccgg gctccggcga 1740
gggaagcacg aaggggcaaa ttcagctggt ccagagcgga cctgagctga aaaaaccgg 1800
cgagactggt aagatcagtt gtaaagcatc tggtataacc ttcaccgact acagcataaa 1860
ttgggtgaaa cgggcccctg gaaaggcct caaatggatg ggttgatca ataccaaac 1920
tagggagcct gcttatgcat atgacttccg cgggagatc gccttttcac tcgagacatc 1980
tgcccttact gcttacctc aaataaacia cctcaagat gaagatacag ccacttactt 2040
ttgcgccctc gactatagtt acgcatgga ctactgggga cagggaacct ccgttacctg 2100
cagttccgcg gccgcaacca caacacctgc tccaaggccc cccacaccgg ctccaactat 2160
agccagccaa ccattgagcc tcagacctga agcttgacag cccgcagcag gaggcgcctg 2220
ccatacgca ggccctggact tcgctgtgta tatttatatt tgggcccctt tggccggaac 2280
atgtgggggtg ttgcttctct ccttgtgat cactctgtat tgtaagcgcg ggagaaagaa 2340
gctcctgtac atcttcaage agccttttat gcgacctgtg caaacctc aggaagaaga 2400
tgggtgttca tgccgcttc ccgaggagga agaaggaggg tgtgaactga gggtgaaatt 2460
ttctagaagc gccgatgctc ccgcatatca gcagggtcag aatcagctct acaatgaatt 2520
gaatctcggc aggcgagaag agtacgatgt tctggacaag agacgggca gggatccga 2580
gatgggggga aagccccgga gaaaaaatcc tcaggagggg ttgtacaatg agctgcagaa 2640
ggacaagatg gctgaagcct atagcgagat cggaatgaaa ggcgaaagac gcagaggcaa 2700
ggggcatgac ggtctgtacc aggtctctc tacagccacc aaggacctt atgatgcgtt 2760
gcatatgcaa gccttgccac cccgctaagc ggccgcgctt ttttgtgaa atttgtgatg 2820
ctattgcttt atttgaacc attataagct gcaataaaca agttaacaac acaattgca 2880

ttcatTTTTat gtttcaggTt cagggggaga tgtgggaggt tttttaagc tcaccggttt 2940
 tgattctcaa acaaagtgt cacaaagtaa ggattctgat gtgtatatca cagacaaaac 3000
 tgtgctagac atgaggTcta tggactTcaa gagcaacagt gctgtggcct ggagcaacaa 3060
 atctgacttt gcatgtgcaa acgcctTcaa caacagcatt attccagaag acaccttctt 3120
 ccccagccca ggtaagggca gctttggTgc cttcgaggc tgtttccttg cttcaggaat 3180
 ggccaggTtc tgcccagagc tctggTcaat gatgtctaaa actcctctga ttggTggTac 3240
 cggttctggg cggTgctaca actgggctgg cggccaggat ggttcttagg taggtggggT 3300
 cggcggtcag gtgtcccaga gccagggTc tggagggacc ttccacctc agtccctggc 3360
 aggtcggggg gtgctgaggc gggcctggc ctggcagccc aggggtcccg gagcgagggg 3420
 tctggagggga cctttcactc tcagtccctg gcaggTcggg gggTgctgtg gcaggccag 3480
 ccttgccccc cagctctgcc ccttacctg agctgtgtgg ctttgggcag ctgaactcc 3540
 tgggttctc tctgggcccc aactcctccc ctggcccaag tcccctctt gctcctgggc 3600
 aggcaggacc tctgtcccct ctccagccggT ccttggggct gcgtgtttct gtagtacgta 3660
 gataagtagc atggcgggtt aatcattaac tacaaggaac ccctagtgat ggagttggcc 3720
 actccctctc tgcgcgctcg ctccctcact gaggccgggc gaccaaaggt cgcgccagc 3780
 ccgggctttg cccgggcggc ctccagtgagc gagcgagcgc gccagctggc gtaatagcga 3840
 agaggcccg accgatcgcc cttcccacaa gttgcgcagc ctgaatggcg aatggcgatt 3900
 ccgttgcaat ggctggcggt aatattgtt tggatattac cagcaaggcc gatagtttga 3960
 gttcttctac tcaggcaagt gatgttatta ctaatcaaag aagtattgcg acaacggtta 4020
 atttgcgtga tggacagact cttttactcg gtggcctcac tgattataaa aacacttctc 4080
 aggattctgg cgtaccgtt ctgtctaaaa tccctttaat cggcctcctg ttagctccc 4140
 gctctgattc taacgaggaa agcacgttat acgtgctcgt caaagcaacc atagtacgcg 4200
 ccctgtagcg gcgcattaag cgcggcgggT gtggTggTta cgcgcagcgt gaccgctaca 4260
 cttgccagcg ccctagcgcc cgtcctttc gttttcttc cttcctttct cgcacgctc 4320
 gccggctttc cccgtcaagc tctaaatcgg gggctccct tagggttccg atttagtgt 4380
 ttacggcacc tcgaccccaa aaaacttgat taggTgatg gttcacgtag tggccatcg 4440
 ccctgataga cggTTTTcgc cctttgacg ttggagtcca cgttctttaa tagtggactc 4500
 ttgttccaaa ctggaacaac actcaacct atctcggTct attcttttga tttataagg 4560
 attttgccga tttcggccta ttggttaaaa aatgagctga tttacaaaa atttaacgcg 4620
 aatTTtaaca aaatattaac gtttacaatt taatatttg cttatacaat cttcctgttt 4680
 ttggggcttt tctgattatc aaccggggta catatgattg acatgctagt tttacgatta 4740
 ccgttcatcg attctcttgt ttgtccaga ctctcaggca atgacctgat agcctttgta 4800
 gagacctctc aaaaatagct accctctccg gcattgaatt atcagctaga acggttgaat 4860
 atcatattga tggTgattg actgtctccg gcctttctca cccgtttgaa tctttaccta 4920
 cacattactc aggcatTgca tttaaaatat atgagggtt taaaaatTT tctccttgcg 4980
 ttgaaataaa ggcttctccc gcaaaagtat tacaggTca taatgtttt ggtacaaccg 5040
 atttagcttt atgctctgag gctttattgc ttaattttgc taattctttg cttgctgt 5100
 atgatttatt ggatgttga atcgctgat gcggtatTT ctcttacgc atctgtcgg 5160
 tatttcacac cgcatatggt gcactctcag tacaatctgc tctgatgccg catagttaag 5220

ccagccccga caccgcca caccgctga cgcgcctga cgggcttgtc tgctcccggc 5280
 atccgcttac agacaagctg tgaccgtctc cgggagctgc atgtgtcaga ggttttcacc 5340
 gtcatcaccg aaacgcgcga gacgaaagg cctcgtgata cgcctatfff tataggttaa 5400
 tgtcatgata ataatggttt cttagacgtc aggtggcact tttcggggaa atgtgcgcgg 5460
 aaccctatt tgtttatfff tctaaataca ttcaaatac tatccgctca tgagacaata 5520
 accctgataa atgcttcaat aatattgaaa aaggaagagt atgagtattc aacatttccg 5580
 tgtcgccctt attccctfff ttgcggcatt ttgccttctt gtttttgctc acccagaaac 5640
 gctggtgaaa gtaaaagatg ctgaagatca gttgggtgca cgagtgggtt acatcgaact 5700
 ggatctcaac agcggtaaga tccttgagag ttttcgccc gaagaacgtt ttccaatgat 5760
 gagcactfff aaagttctgc tatgtggcgc ggtattatcc cgtattgacg ccgggcaaga 5820
 gcaactcggc cgccgcatac actattctca gaatgacttg gttgagtact caccagtcaac 5880
 agaaaagcat cttacggatg gcatgacagt aagagaatta tgcagtgtcg ccataacat 5940
 gagtgataac actgcggcca acttacttct gacaacgatc ggaggaccga aggagctaac 6000
 cgcttttttg cacaacatgg gggatcatgt aactcgcctt gatcgttggg aaccggagct 6060
 gaatgaagcc ataccaaaac acgagcgtga caccacgatg cctgtagcaa tggcaacaac 6120
 gttgcgcaaa ctattaactg gcgaactact tactctagct tcccggcaac aattaataga 6180
 ctggatggag gcggataaag ttgcaggacc acttctgcgc tcggcccttc cggctggctg 6240
 gtttattgct gataaatctg gagccggtga gcgtgggtct cgcggtatca ttgcagcact 6300
 ggggccagat ggtaagccct cccgtatcgt agttatctac acgacgggga gtcaggcaac 6360
 tatggatgaa cgaaatagac agatcgctga gataggtgcc tcaactgatta agcattggta 6420
 actgtcagac caagtttact catatatac ttagattgat ttaaacttc atttttaatt 6480
 taaaaggatc taggtgaaga tcctttttga taatctcatg accaaaatcc cttaacgtga 6540
 gttttcgttc cactgagcgt cagacccctg agaaaagatc aaaggatctt cttgagatcc 6600
 tttttttctg cgcgtaatct gctgcttgca acaaaaaaa ccaccgctac cagcgggtgt 6660
 ttgtttgccg gatcaagagc taccaactct tttccgaag gtaactggct tcagcagagc 6720
 gcagatacca aatactgtcc ttctagtgtg gccgtagtta ggccaccact tcaagaactc 6780
 tgtagcaccg cctacatacc tcgctctgct aatcctgtta ccagtggctg ctgccagtgg 6840
 cgataagtcg tgccttaccg ggttgactc aagacgatag ttaccggata aggcgagcg 6900
 gtcgggctga acggggggtt cgtgcacaca gccagcttg gagcgaacga cctacaccga 6960
 actgagatac ctacagcgtg agctatgaga aagcggcagc cttcccgaag ggagaaaggc 7020
 ggacaggatc ccgtaagcg gcagggtcgg aacaggagag cgcacgaggg agcttccagg 7080
 gggaaacgcc tggatcttt atagtctgt cgggtttcgc cacctctgac ttgagcgtcg 7140
 atttttgtga tgctcgtcag gggggcggag cctatggaaa aacgccagca acgcggcctt 7200
 tttacggttc ctggcctfff gctggccttt tgctcacatg ttctttctg cgttatcccc 7260
 tgattctgtg gataaccgta ttaccgctt tgagtgagct gataccgctc gccgcagccg 7320
 aacgaccgag cgcagcagat cagtgagcga ggaagcggaa gagcgcceaa tacgcaaacc 7380
 gcctctcccc gcgcgttggc cgattcatta atg 7413

<210> 3

<211> 3

<212> PRT
<213> 人工序列 (Artificial Sequence)
<220>
<223> 例示性连接子序列
<400> 3
Gly Gly Gly
1
<210> 4
<211> 5
<212> PRT
<213> 人工序列 (Artificial Sequence)
<220>
<223> 例示性连接子序列
<400> 4
Asp Gly Gly Gly Ser
1 5
<210> 5
<211> 5
<212> PRT
<213> 人工序列 (Artificial Sequence)
<220>
<223> 例示性连接子序列
<400> 5
Thr Gly Glu Lys Pro
1 5
<210> 6
<211> 4
<212> PRT
<213> 人工序列 (Artificial Sequence)
<220>
<223> 例示性连接子序列
<400> 6
Gly Gly Arg Arg
1
<210> 7
<211> 5
<212> PRT
<213> 人工序列 (Artificial Sequence)
<220>

<223> 例示性连接子序列

<400> 7

Gly Gly Gly Gly Ser

1 5

<210> 8

<211> 14

<212> PRT

<213> 人工序列 (Artificial Sequence)

<220>

<223> 例示性连接子序列

<400> 8

Glu Gly Lys Ser Ser Gly Ser Gly Ser Glu Ser Lys Val Asp

1 5 10

<210> 9

<211> 18

<212> PRT

<213> 人工序列 (Artificial Sequence)

<220>

<223> 例示性连接子序列

<400> 9

Lys Glu Ser Gly Ser Val Ser Ser Glu Gln Leu Ala Gln Phe Arg Ser

1 5 10 15

Leu Asp

<210> 10

<211> 8

<212> PRT

<213> 人工序列 (Artificial Sequence)

<220>

<223> 例示性连接子序列

<400> 10

Gly Gly Arg Arg Gly Gly Gly Ser

1 5

<210> 11

<211> 9

<212> PRT

<213> 人工序列 (Artificial Sequence)

<220>

<223> 例示性连接子序列

<400> 11

Glu Xaa Xaa Tyr Xaa Gln Xaa

1 5

<210> 15

<211> 7

<212> PRT

<213> 人工序列 (Artificial Sequence)

<220>

<223> 由TEV蛋白酶获得的裂解序列

<400> 15

Glu Asn Leu Tyr Phe Gln Gly

1 5

<210> 16

<211> 7

<212> PRT

<213> 人工序列 (Artificial Sequence)

<220>

<223> 由TEV蛋白酶获得的裂解序列

<400> 16

Glu Asn Leu Tyr Phe Gln Ser

1 5

<210> 17

<211> 22

<212> PRT

<213> 人工序列 (Artificial Sequence)

<220>

<223> 包含2A位点的自裂解多肽

<400> 17

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> 18

<211> 19

<212> PRT

<213> 人工序列 (Artificial Sequence)

<220>

<223> 包含2A位点的自裂解多肽

<400> 18

Ala Thr Asn Phe Ser Leu Leu Lys Gln Ala Gly Asp Val Glu Glu Asn

1	5	10	15
Pro Gly Pro			
<210> 19			
<211> 14			
<212> PRT			
<213> 人工序列 (Artificial Sequence)			
<220>			
<223> 包含2A位点的自裂解多肽			
<400> 19			
Leu Leu Lys Gln Ala Gly Asp Val Glu Glu Asn Pro Gly Pro			
1	5	10	
<210> 20			
<211> 21			
<212> PRT			
<213> 人工序列 (Artificial Sequence)			
<220>			
<223> 包含2A位点的自裂解多肽			
<400> 20			
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> 21			
<211> 18			
<212> PRT			
<213> 人工序列 (Artificial Sequence)			
<220>			
<223> 包含2A位点的自裂解多肽			
<400> 21			
Glu Gly Arg Gly Ser Leu Leu Thr Cys Gly Asp Val Glu Glu Asn Pro			
1	5	10	15
Gly Pro			
<210> 22			
<211> 13			
<212> PRT			
<213> 人工序列 (Artificial Sequence)			
<220>			
<223> 包含2A位点的自裂解多肽			
<400> 22			

Leu Leu Thr Cys Gly Asp Val Glu Glu Asn Pro Gly Pro

1 5 10

<210> 23

<211> 23

<212> PRT

<213> 人工序列(Artificial Sequence)

<220>

<223> 包含2A位点的自裂解多肽

<400> 23

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> 24

<211> 20

<212> PRT

<213> 人工序列(Artificial Sequence)

<220>

<223> 包含2A位点的自裂解多肽

<400> 24

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> 25

<211> 14

<212> PRT

<213> 人工序列(Artificial Sequence)

<220>

<223> 包含2A位点的自裂解多肽

<400> 25

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

1 5 10

<210> 26

<211> 25

<212> PRT

<213> 人工序列(Artificial Sequence)

<220>

<223> 包含2A位点的自裂解多肽

<220>

<223> 包含2A位点的自裂解多肽

<400> 30

Thr Leu Asn Phe Asp Leu Leu Lys Leu Ala Gly Asp Val Glu Ser Asn
1 5 10 15

Pro Gly Pro

<210> 31

<211> 14

<212> PRT

<213> 人工序列 (Artificial Sequence)

<220>

<223> 包含2A位点的自裂解多肽

<400> 31

Leu Leu Lys Leu Ala Gly Asp Val Glu Ser Asn Pro Gly Pro
1 5 10

<210> 32

<211> 17

<212> PRT

<213> 人工序列 (Artificial Sequence)

<220>

<223> 包含2A位点的自裂解多肽

<400> 32

Asn Phe Asp Leu Leu Lys Leu Ala Gly Asp Val Glu Ser Asn Pro Gly
1 5 10 15

Pro

<210> 33

<211> 20

<212> PRT

<213> 人工序列 (Artificial Sequence)

<220>

<223> 包含2A位点的自裂解多肽

<400> 33

Gln Leu Leu Asn Phe Asp Leu Leu Lys Leu Ala Gly Asp Val Glu Ser
1 5 10 15

Asn Pro Gly Pro

20

<210> 34

<211> 24

<212> PRT

<213> 人工序列 (Artificial Sequence)

<220>

<223> 包含2A位点的自裂解多肽

<400> 34

Ala Pro Val Lys Gln Thr Leu Asn Phe Asp Leu Leu Lys Leu Ala Gly

1 5 10 15

Asp Val Glu Ser Asn Pro Gly Pro

20

<210> 35

<211> 40

<212> PRT

<213> 人工序列 (Artificial Sequence)

<220>

<223> 包含2A位点的自裂解多肽

<400> 35

Val Thr Glu Leu Leu Tyr Arg Met Lys Arg Ala Glu Thr Tyr Cys Pro

1 5 10 15

Arg Pro Leu Leu Ala Ile His Pro Thr Glu Ala Arg His Lys Gln Lys

20 25 30

Ile Val Ala Pro Val Lys Gln Thr

35 40

<210> 36

<211> 18

<212> PRT

<213> 人工序列 (Artificial Sequence)

<220>

<223> 包含2A位点的自裂解多肽

<400> 36

Leu Asn Phe Asp Leu Leu Lys Leu Ala Gly Asp Val Glu Ser Asn Pro

1 5 10 15

Gly Pro

<210> 37

<211> 40

<212> PRT

<213> 人工序列 (Artificial Sequence)

<220>

<223> 包含2A位点的自裂解多肽

<400> 37

Leu Leu Ala Ile His Pro Thr Glu Ala Arg His Lys Gln Lys Ile Val

1 5 10 15
 Ala Pro Val Lys Gln Thr Leu Asn Phe Asp Leu Leu Lys Leu Ala Gly
 20 25 30

Asp Val Glu Ser Asn Pro Gly Pro
 35 40

<210> 38

<211> 33

<212> PRT

<213> 人工序列 (Artificial Sequence)

<220>

<223> 包含2A位点的自裂解多肽

<400> 38

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

Pro

<210> 39

<211> 10

<212> DNA

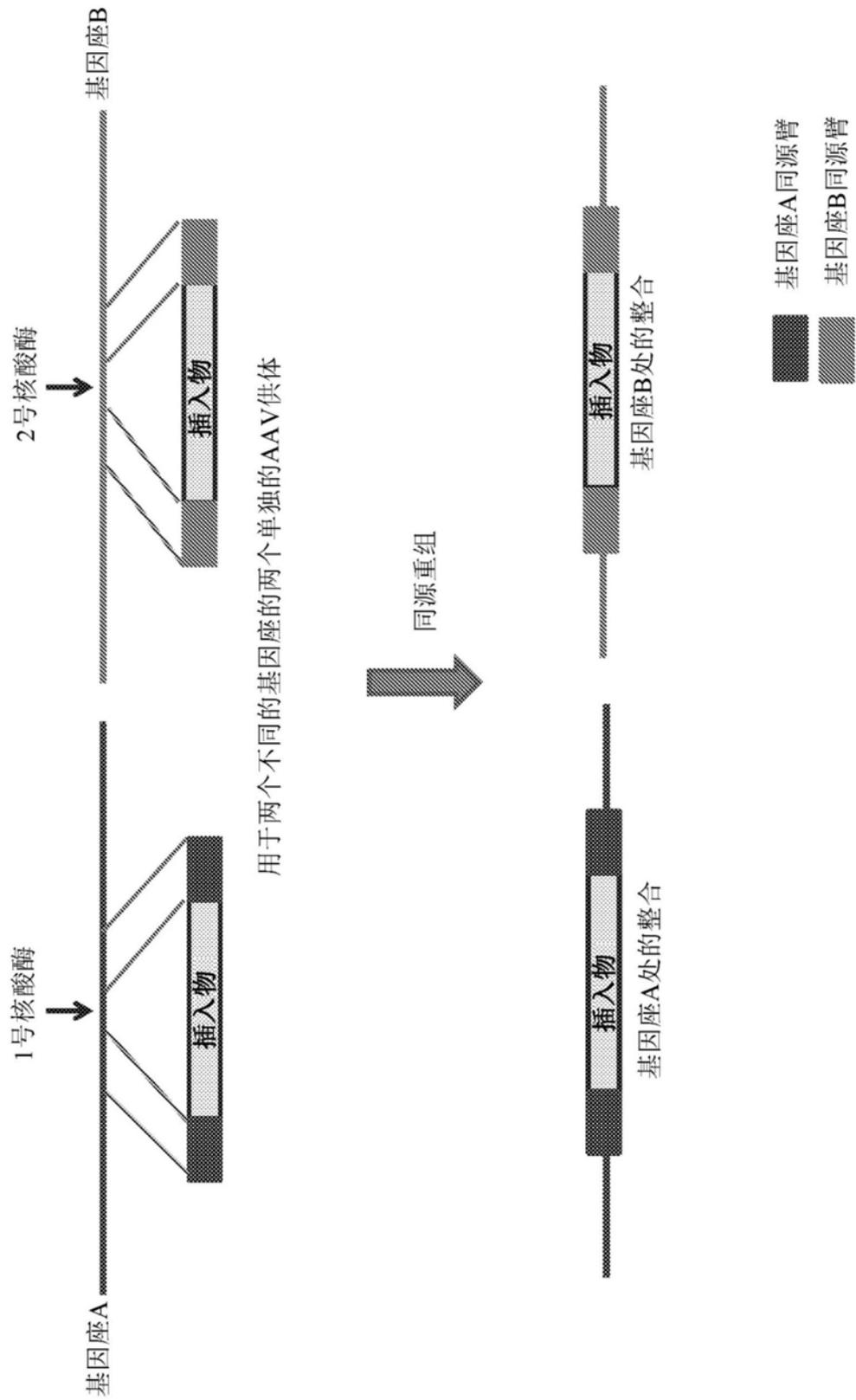
<213> 人工序列 (Artificial Sequence)

<220>

<223> 共同Kozak序列

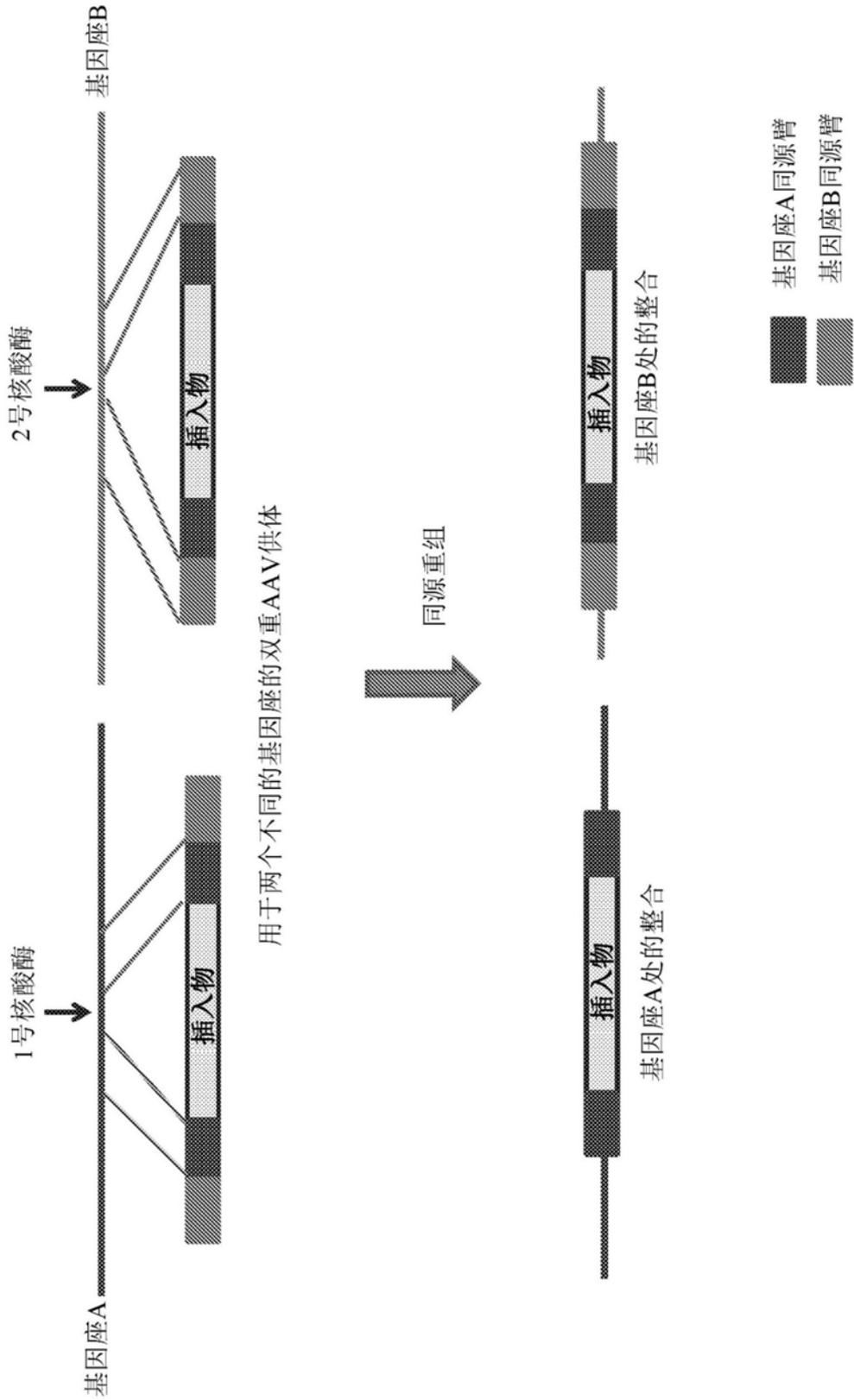
<400> 39

gccrccatgg 10



用于两个不同的基因座的两个单独的AAV供体

图1



用于两个不同的基因座的双重AAV供体

图2

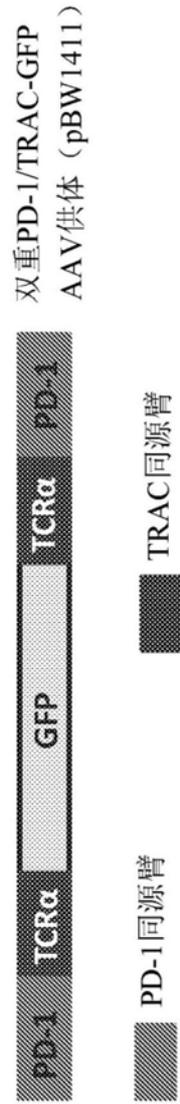


图3A

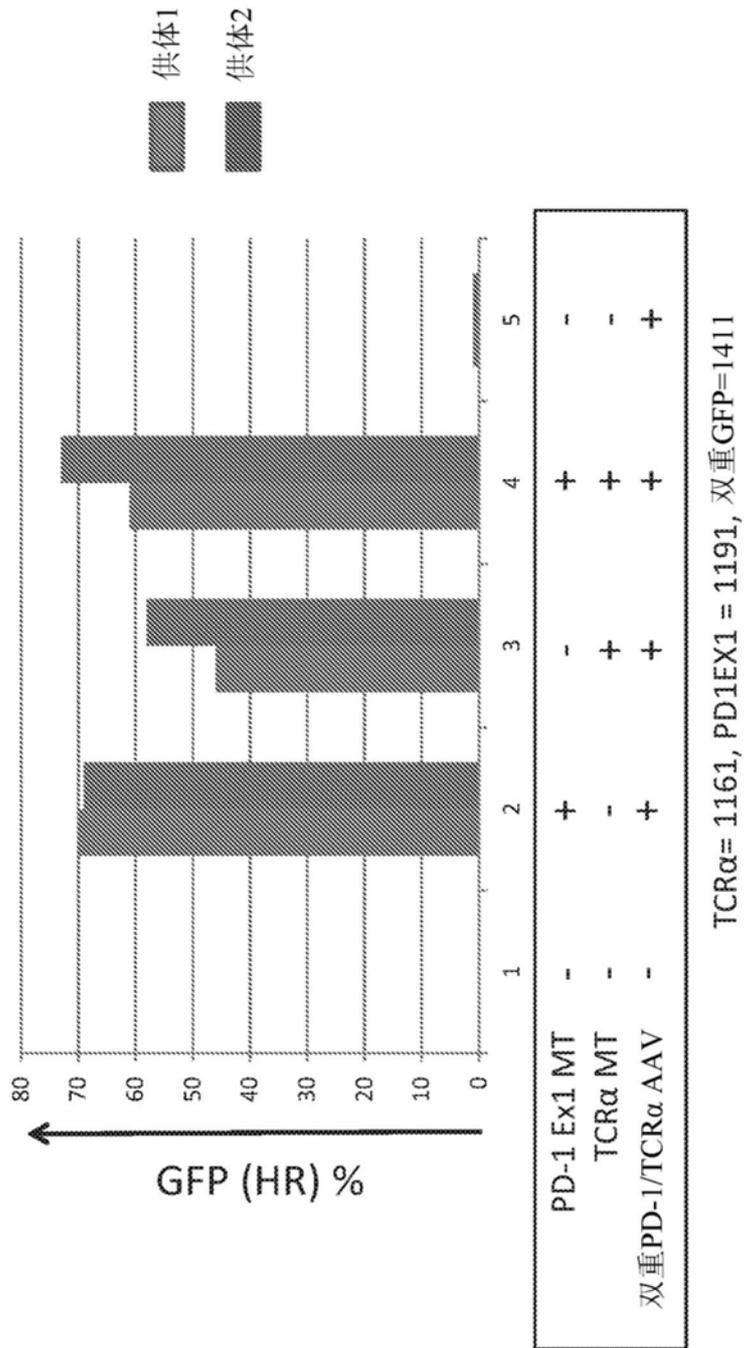


图3B

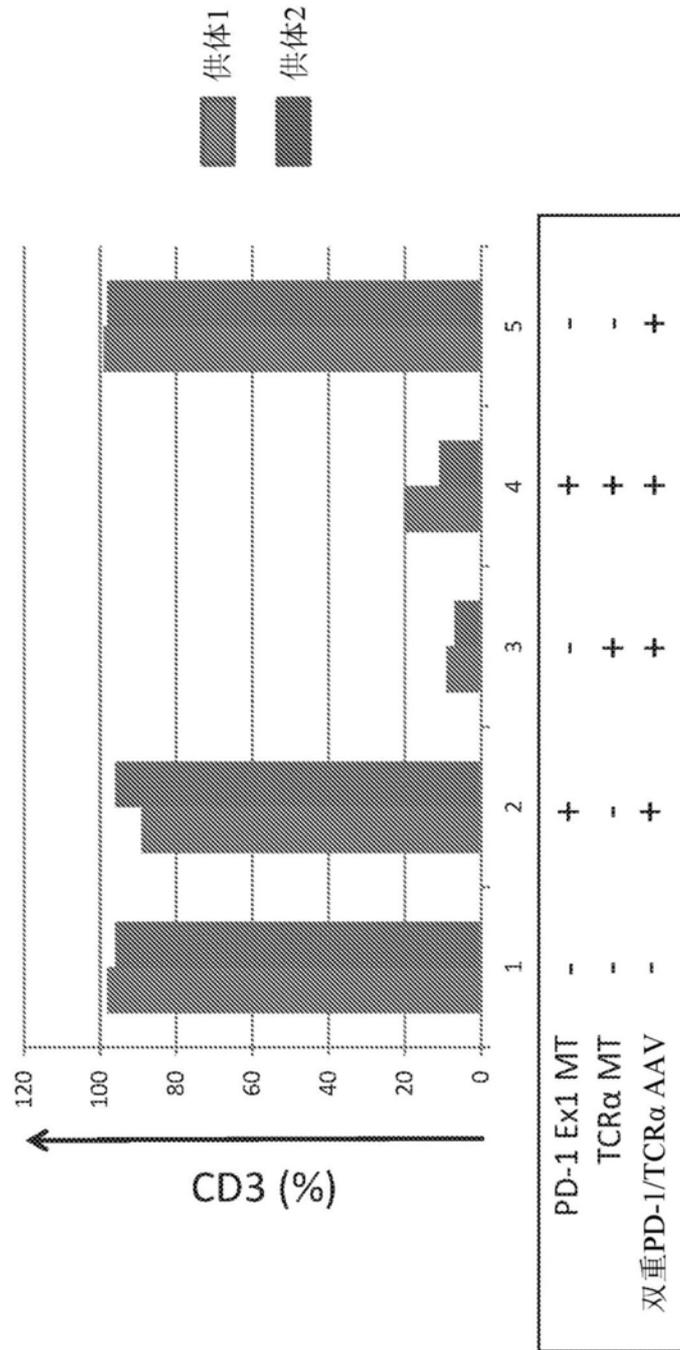


图4

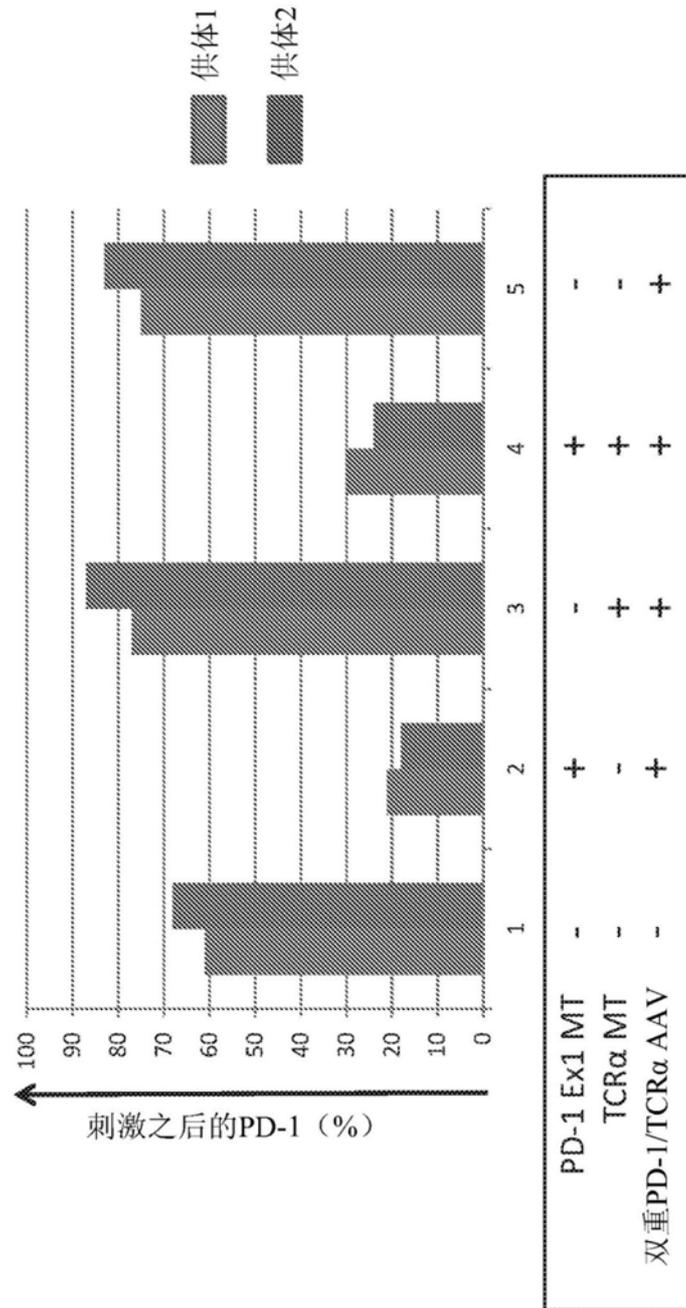


图5

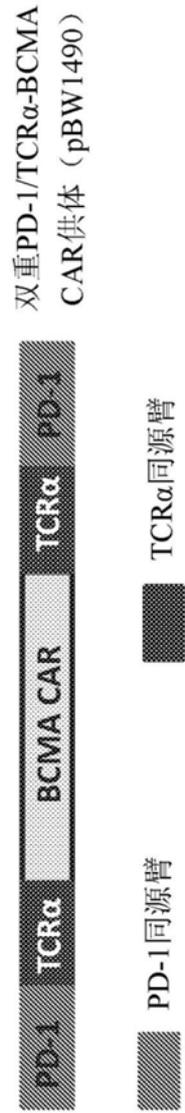


图6A

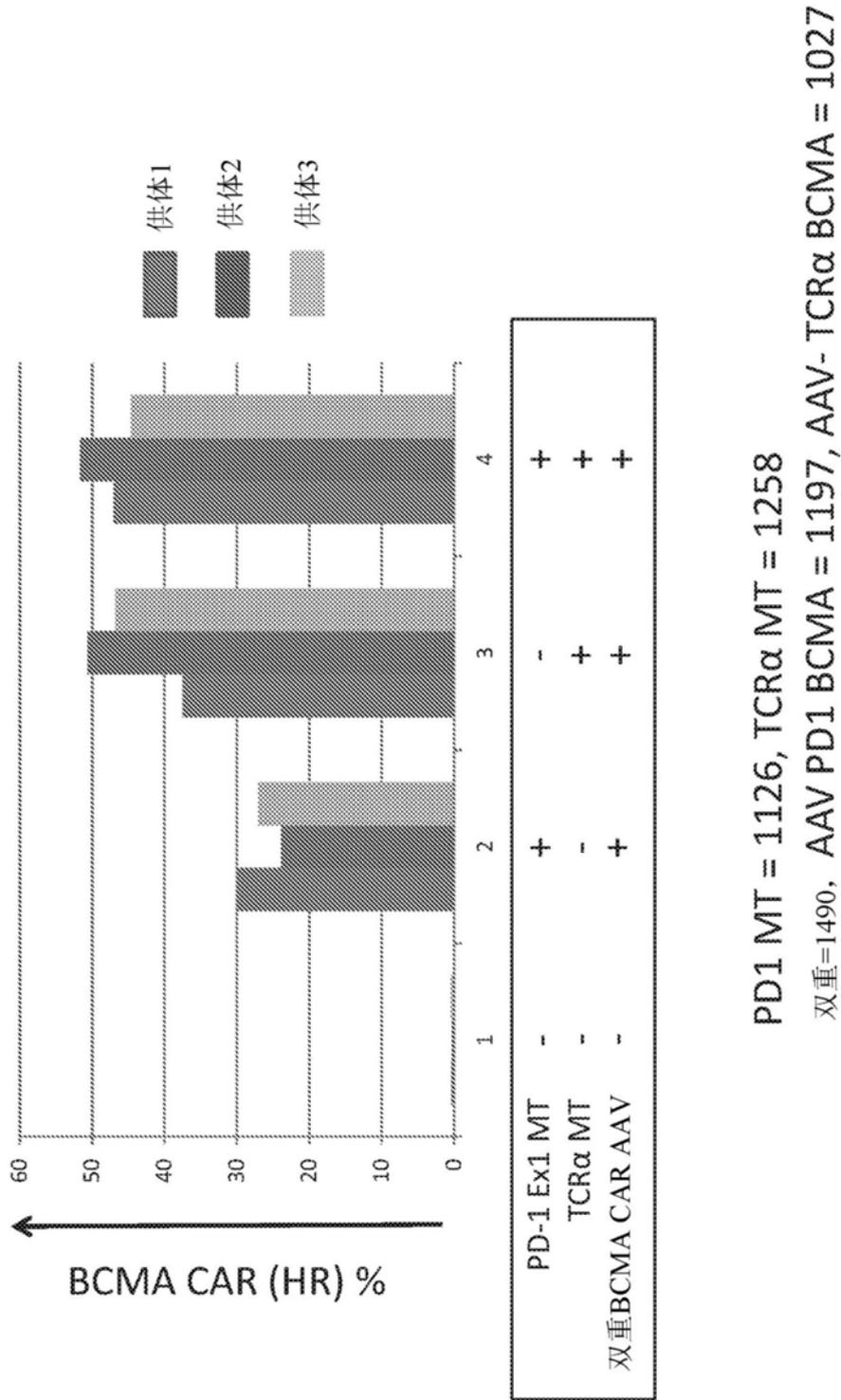


图6B

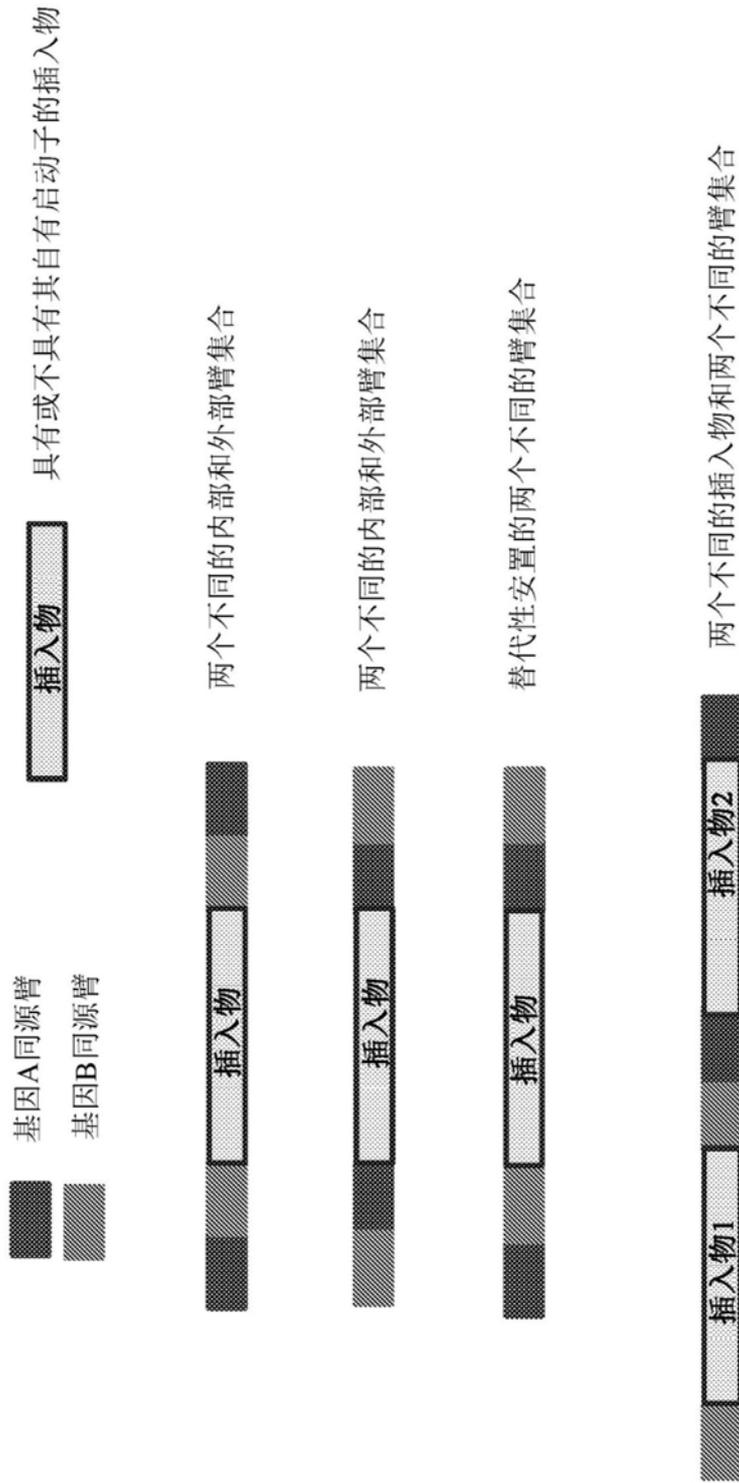


图7