(19)国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 116529364 A
(43)申请公布日 2023.08.01

(21)申请号 202180073716.2
(22)申请日 2021.08.27
(30)优先权数据
63/072,080 2020.08.28 US
(85)PCT国际申请进入国家阶段日
2023.04.27
(86)PCT国际申请的申请数据
PCT/US2021/048066 2021.08.27
(87)PCT国际申请的公布数据
WO2022/047237 EN 2022.03.03
(71)申请人 阿森纳生物科学公司
地址 美国加利福尼亚州
(72)发明人 A·库珀 M·阮 姚安之
S·桑特罗
(74)专利代理机构 北京市金杜律师事务所 11256
专利代理师 陈文平
(51)Int.Cl.
C12N 9/22(2006.01)

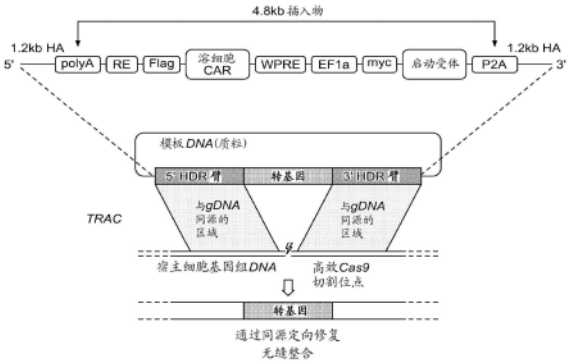
权利要求书7页 说明书51页 附图23页

(54)发明名称

具有启动受体的工程化免疫细胞

(57)摘要

本文提供了使用非病毒编辑方法利用大DNA模板对细胞进行基因编辑的方法。本文还提供了包含至少一个以非病毒方式插入细胞的基因组的靶区中的大DNA模板的细胞。



CN 116529364 A

1.一种原代免疫细胞,所述原代免疫细胞包含至少一个以非病毒方式插入所述细胞的基因组的靶区中的DNA模板,其中所述DNA模板的大小为大于或等于约5千碱基对(kb)。

2.如权利要求1所述的细胞,其中所述原代免疫细胞不包含用于将所述DNA模板引入所述原代免疫细胞中的病毒载体。

3.如权利要求1或2所述的细胞,其中所述DNA模板的大小为大于或等于约5.0kb、5.1kb、5.2kb、5.3kb、5.4kb、5.5kb、5.6kb、5.7kb、5.8kb、5.9kb、6.0kb、6.1kb、6.2kb、6.3kb、6.4kb、6.5kb、6.6kb、6.7kb、6.8kb、6.9kb、7.0kb、7.1kb、7.2kb、7.3kb、7.4kb、7.5kb、7.6kb、7.7kb、7.8kb、7.9kb、8.0kb、8.1kb、8.2kb、8.3kb、8.4kb、8.5kb、8.6kb、8.7kb、8.8kb、8.9kb、9.0kb、9.1kb、9.2kb、9.3kb、9.4kb、9.5kb、9.6kb、9.7kb、9.8kb.9.9kb、10.0kb或介于这些大小之间的任何DNA模板大小。

4.如权利要求1至3中任一项所述的细胞,其中所述DNA模板的大小为约5kb至约10kb、约5kb至约9kb、约5kb至约8kb、约5kb至约7kb、约5kb至约6kb、约kb 6至约10kb、约6kb至约9kb、约6kb至约8kb、约6kb至约7kb、约7kb至约10kb、约7kb至约9kb、约7kb至约8kb、约8kb至约10kb、约8kb至约9kb,或约9kb至约10kb。

5.如权利要求1至4中任一项所述的细胞,其中所述DNA模板是双链DNA模板或单链DNA模板。

6.如权利要求1至5中任一项所述的细胞,其中所述DNA模板是线状DNA模板或环状DNA模板,任选地其中所述环状DNA模板是质粒。

7.如权利要求1至6中任一项所述的细胞,其中所述DNA模板的5'端和3'端包含与所述原代细胞的基因组中的插入位点侧翼的基因组序列同源的核苷酸序列。

8.如权利要求1至7中任一项所述的细胞,其中所述细胞的基因组的所述靶区是T细胞受体α恒定(TRAC)基因座或基因组安全港(GSH)。

9.如权利要求1至8中任一项所述的细胞,其中所述DNA模板包含异源序列。

10.如权利要求1至9中任一项所述的细胞,其中所述DNA模板包含基因。

11.如权利要求1至10中任一项所述的细胞,其中所述DNA模板包含含有转录因子的启动受体。

12.如权利要求1至11中任一项所述的细胞,其中所述DNA模板包含嵌合抗原受体(CAR)。

13.如权利要求1至12中任一项所述的细胞,其中所述DNA模板包含嵌合抗原受体(CAR)和包含转录因子的启动受体。

14.如权利要求1至13中任一项所述的细胞,其中所述DNA模板包含可操作地连接至所述嵌合抗原受体的诱导型启动子。

15.如权利要求1至14中任一项所述的细胞,其中所述DNA模板还包含可操作地连接至所述启动受体的组成型启动子。

16.如权利要求1至15中任一项所述的细胞,其中所述DNA模板还包含可操作地连接至所述嵌合抗原受体的诱导型启动子和可操作地连接至所述启动受体的组成型启动子。

17.如权利要求1至16中任一项所述的细胞,其中所述DNA模板在5'至3'方向上包含:

a.所述诱导型启动子;

b.所述嵌合抗原受体;

c.所述组成型启动子;和

d.所述启动受体。

18.如权利要求1至16中任一项所述的细胞,其中所述DNA模板在5'至3'方向上包含:

a.所述组成型启动子;

b.所述启动受体;

c.所述诱导型启动子;和

d.所述嵌合抗原受体。

19.如权利要求1至24中任一项所述的细胞,其中所述DNA模板还包含自切除2A肽(P2A)。

20.如权利要求1至23中任一项所述的细胞,其中所述P2A核酸处于所述DNA模板的3'端。

21.如权利要求1至20中任一项所述的细胞,其中所述DNA模板还包含土拨鼠肝炎病毒翻译后调控元件(WPRE)。

22.如权利要求21所述的细胞,其中所述WPRE处于编码所述CAR的核酸的3'端和编码所述启动受体的核酸的5'端,或者其中所述WPRE处于编码所述启动受体的核酸的3'端和编码所述CAR的核酸的5'端。

23.如权利要求1至22中任一项所述的细胞,其中所述启动受体在N末端至C末端方向上包含:

a.对抗原具有结合亲和力的细胞外抗原结合结构域;

b.包含一个或多个配体诱导型蛋白水解裂解位点的跨膜结构域;和

c.包含人或人源化转录效应子的细胞内结构域,其中抗原与所述细胞外抗原结合结构域的结合导致在所述配体诱导型蛋白水解裂解位点处的裂解,从而释放所述细胞内结构域。

24.如权利要求23所述的细胞,其中所述启动受体还包含定位于所述跨膜结构域与所述细胞内结构域之间的近膜结构域(JMD)。

25.如权利要求1至24中任一项所述的细胞,其中所述转录因子结合至所述诱导型启动子并诱导所述CAR的表达。

26.如权利要求1至25中任一项所述的细胞,其中所述CAR从N末端至C末端包含:

a.对抗原具有结合亲和力的细胞外抗原结合结构域;

b.跨膜结构域;

c.细胞内共刺激结构域;和

d.细胞内激活结构域。

27.如权利要求1至26中任一项所述的细胞,其中所述启动受体和所述CAR结合不同的抗原。

28.如权利要求1至26中任一项所述的细胞,其中所述启动受体和所述CAR结合相同的抗原。

29.如权利要求1至28中任一项所述的细胞,其中所述免疫细胞是原代人免疫细胞。

30.如权利要求1至29中任一项所述的细胞,其中所述原代免疫细胞是自体免疫细胞。

31.如权利要求1至30中任一项所述的细胞,其中所述原代免疫细胞是自然杀伤(NK)细

胞、T细胞、CD8+T细胞、CD4+T细胞、原代T细胞或T细胞祖细胞。

32.如权利要求1至31中任一项所述的细胞，其中所述原代免疫细胞是原代T细胞。

33.如权利要求1至32中任一项所述的细胞，其中所述原代免疫细胞是原代人T细胞。

34.如权利要求1至33中任一项所述的细胞，其中所述原代免疫细胞是无病毒的。

35.一种细胞群体，所述细胞群体包含多个权利要求1至34中任一项所述的原代免疫细胞。

36.一种原代免疫细胞，所述原代免疫细胞包含至少一个插入所述原代免疫细胞的基因组的靶区中的DNA模板，其中所述DNA模板的大小为大于或等于5千碱基对，并且其中所述原代免疫细胞不包含用于将所述DNA模板引入所述原代免疫细胞中的病毒载体。

37.一种原代免疫细胞，所述原代免疫细胞包含至少一个插入所述原代免疫细胞的基因组的靶区中的DNA模板，所述DNA模板包含嵌合抗原受体(CAR)和包含转录因子的启动受体，其中所述DNA模板的大小为大于或等于5千碱基对，并且其中所述原代免疫细胞不包含用于将所述DNA模板引入所述原代免疫细胞中的病毒载体。

38.一种活的无病毒的原代细胞，所述原代细胞包含核糖核蛋白复合物(RNP)-DNA模板复合物，其中所述RNP包含核酸酶结构域和指导RNA，其中所述DNA模板的大小为大于或等于5千碱基核苷酸的大小，并且其中所述DNA模板的5’端和3’端包含与所述原代细胞的基因组中的插入位点侧翼的基因组序列同源的核苷酸序列。

39.一种活的无病毒的原代细胞，所述原代细胞包含核糖核蛋白复合物(RNP)-DNA模板复合物，其中所述RNP包含核酸酶结构域和指导RNA，其中所述DNA模板的大小为大于或等于5千碱基核苷酸的大小，其中所述DNA模板包含嵌合抗原受体(CAR)和包含转录因子的启动受体，并且其中所述DNA模板的5’端和3’端包含与所述原代细胞的基因组中的插入位点侧翼的基因组序列同源的核苷酸序列。

40.一种治疗受试者的疾病的方法，所述方法包括向所述受试者施用权利要求1至39中任一项所述的原代免疫细胞。

41.如权利要求40所述的方法，其中所述疾病是癌症。

42.一种非病毒载体，所述非病毒载体包含DNA模板，其中所述DNA模板的大小为大于或等于5千碱基核苷酸，并且其中所述DNA模板的5’端和3’端包含与原代细胞的基因组中的插入位点侧翼的基因组序列同源的核苷酸序列。

43.如权利要求42所述的非病毒载体，其中所述DNA模板的大小为大于或等于约5.0kb、5.1kb、5.2kb、5.3kb、5.4kb、5.5kb、5.6kb、5.7kb、5.8kb、5.9kb、6.0kb、6.1kb、6.2kb、6.3kb、6.4kb、6.5kb、6.6kb、6.7kb、6.8kb、6.9kb、7.0kb、7.1kb、7.2kb、7.3kb、7.4kb、7.5kb、7.6kb、7.7kb、7.8kb、7.9kb、8.0kb、8.1kb、8.2kb、8.3kb、8.4kb、8.5kb、8.6kb、8.7kb、8.8kb、8.9kb、9.0kb、9.1kb、9.2kb、9.3kb、9.4kb、9.5kb、9.6kb、9.7kb、9.8kb.9.9kb、10.0kb或介于这些大小之间的任何DNA模板大小。

44.如权利要求42或43所述的非病毒载体，其中所述DNA模板的大小为约5kb至约10kb、约5kb至约9kb、约5kb至约8kb、约5kb至约7kb、约5kb至约6kb、约kb 6至约10kb、约6kb至约9kb、约6kb至约8kb、约6kb至约7kb、约7kb至约10kb、约7kb至约9kb、约7kb至约8kb、约8kb至约10kb、约8kb至约9kb，或约9kb至约10kb。

45.如权利要求42至44中任一项所述的非病毒载体，其中所述细胞的基因组的靶区是T

细胞受体α恒定(TRAC)基因座或基因组安全港(GSH)。

46.如权利要求42至44中任一项所述的非病毒载体,其中所述DNA模板包含异源序列。

47.如权利要求42至44中任一项所述的非病毒载体,其中所述DNA模板包含基因。

48.如权利要求42至47中任一项所述的非病毒载体,其中所述DNA模板包含含有转录因子的启动受体。

49.如权利要求42至48中任一项所述的非病毒载体,其中所述DNA模板包含嵌合抗原受体(CAR)。

50.如权利要求42至49中任一项所述的非病毒载体,其中所述DNA模板包含嵌合抗原受体(CAR)和包含转录因子的启动受体。

51.如权利要求42至50中任一项所述的非病毒载体,其中所述DNA模板包含可操作地连接至所述嵌合抗原受体的诱导型启动子。

52.如权利要求42至51中任一项所述的非病毒载体,其中所述DNA模板还包含可操作地连接至所述启动受体的组成型启动子。

53.如权利要求42至52中任一项所述的非病毒载体,其中所述DNA模板还包含可操作地连接至所述嵌合抗原受体的诱导型启动子和可操作地连接至所述启动受体的组成型启动子。

54.如权利要求42至52中任一项所述的非病毒载体,其中所述DNA模板在5'至3'方向上包含:

a.所述诱导型启动子;

b.所述嵌合抗原受体;

c.所述组成型启动子;和

d.所述启动受体。

55.如权利要求42至52中任一项所述的非病毒载体,其中所述DNA模板在5'至3'方向上包含:

a.所述组成型启动子;

b.所述启动受体;

c.所述诱导型启动子;和

d.所述嵌合抗原受体。

56.如权利要求42至60中任一项所述的非病毒载体,其中所述DNA模板还包含自切除2A肽(P2A)。

57.如权利要求42至56中任一项所述的非病毒载体,其中所述P2A处于所述DNA模板的3'端。

58.如权利要求42至57中任一项所述的非病毒载体,其中所述DNA模板还包含土拨鼠肝炎病毒翻译后调控元件(WPRE)。

59.如权利要求42至53中任一项所述的非病毒载体,其中所述启动受体在N末端至C末端方向上包含:

a.对抗原具有结合亲和力的细胞外抗原结合结构域;

b.包含一个或多个配体诱导型蛋白水解裂解位点的跨膜结构域;和

c.包含人或人源化转录效应子的细胞内结构域,其中抗原与所述细胞外抗原结合结构

域的结合导致在所述配体诱导型蛋白水解裂解位点处的裂解，从而释放所述细胞内结构域。

60.如权利要求59所述的非病毒载体，其中所述启动受体还包含定位于所述跨膜结构域与所述细胞内结构域之间的近膜结构域(JMD)。

61.如权利要求42至58中任一项所述的非病毒载体，其中所述转录因子结合至所述诱导型启动子并诱导所述CAR的表达。

62.如权利要求42至58中任一项所述的非病毒载体，其中所述CAR从N末端至C末端包含：

a.对抗原具有结合亲和力的细胞外抗原结合结构域；

b.跨膜结构域；

c.细胞内共刺激结构域；和

d.细胞内激活结构域。

63.如权利要求42至58中任一项所述的非病毒载体，其中所述启动受体和所述CAR结合不同的抗原。

64.如权利要求42至58中任一项所述的非病毒载体，其中所述启动受体和所述CAR结合相同的抗原。

65.一种编辑原代免疫细胞的方法，所述方法包括：

a.提供核糖核蛋白复合物(RNP)-DNA模板复合物，其中所述RNP包含核酸酶结构域和指导RNA，其中所述DNA模板的大小为大于或等于5千碱基核苷酸的大小，并且其中所述DNA模板的5’端和3’端包含与所述原代免疫细胞的基因组中的插入位点侧翼的基因组序列同源的核苷酸序列；

b.以非病毒方式将所述RNP-DNA模板复合物引入所述原代免疫细胞中，其中所述指导RNA与所述原代免疫细胞的基因组的靶区特异性地杂交，并且其中所述核酸酶结构域使所述靶区裂解以产生在所述原代免疫细胞的基因组中的所述插入位点；以及

c.通过将所述DNA模板插入所述原代免疫细胞的基因组中的所述插入位点中来编辑所述原代免疫细胞。

66.如权利要求65所述的方法，其中以非病毒方式引入包括电穿孔。

67.如权利要求65或66所述的方法，其中所述核酸酶结构域包含CRISPR相关核酸内切酶(Cas)，任选地是Cas9核酸酶。

68.如权利要求65至67中任一项所述的方法，其中所述DNA模板的大小为大于或等于约5.0kb、5.1kb、5.2kb、5.3kb、5.4kb、5.5kb、5.6kb、5.7kb、5.8kb、5.9kb、6.0kb、6.1kb、6.2kb、6.3kb、6.4kb、6.5kb、6.6kb、6.7kb、6.8kb、6.9kb、7.0kb、7.1kb、7.2kb、7.3kb、7.4kb、7.5kb、7.6kb、7.7kb、7.8kb、7.9kb、8.0kb、8.1kb、8.2kb、8.3kb、8.4kb、8.5kb、8.6kb、8.7kb、8.8kb、8.9kb、9.0kb、9.1kb、9.2kb、9.3kb、9.4kb、9.5kb、9.6kb、9.7kb、9.8kb.9.9kb、10.0kb或介于这些大小之间的任何DNA模板大小。

69.如权利要求65至68中任一项所述的方法，其中所述DNA模板的大小为约5kb至约10kb、约5kb至约9kb、约5kb至约8kb、约5kb至约7kb、约5kb至约6kb、约kb 6至约10kb、约6kb至约9kb、约6kb至约8kb、约6kb至约7kb、约7kb至约10kb、约7kb至约9kb、约7kb至约8kb、约8kb至约10kb、约8kb至约9kb，或约9kb至约10kb。

70.如权利要求65至69中任一项所述的方法,其中所述细胞的基因组的所述靶区是T细胞受体α恒定(TRAC)基因座或基因组安全港(GSH)。

71.如权利要求65至70中任一项所述的方法,其中所述DNA模板是双链DNA模板或单链DNA模板。

72.如权利要求65至71中任一项所述的方法,其中所述DNA模板是线状DNA模板或环状DNA模板,任选地其中所述环状DNA模板是质粒。

73.如权利要求65至72中任一项所述的方法,其中所述DNA模板包含异源序列。

74.如权利要求65至73中任一项所述的方法,其中所述DNA模板包含基因。

75.如权利要求65至74中任一项所述的方法,其中所述DNA模板包含含有转录因子的启动受体。

76.如权利要求65至75中任一项所述的方法,其中所述DNA模板包含嵌合抗原受体(CAR)。

77.如权利要求65至76中任一项所述的方法,其中所述DNA模板包含嵌合抗原受体(CAR)和包含转录因子的启动受体。

78.如权利要求65至77中任一项所述的方法,其中所述DNA模板包含可操作地连接至所述嵌合抗原受体的诱导型启动子。

79.如权利要求65至78中任一项所述的方法,其中所述DNA模板还包含可操作地连接至所述启动受体的组成型启动子。

80.如权利要求65至79中任一项所述的方法,其中所述DNA模板还包含可操作地连接至所述嵌合抗原受体的诱导型启动子和可操作地连接至所述启动受体的组成型启动子。

81.如权利要求65至80中任一项所述的方法,其中所述DNA模板在5'至3'方向上包含:

a.所述诱导型启动子;

b.所述嵌合抗原受体;

c.所述组成型启动子;和

d.所述启动受体。

82.如权利要求65至80中任一项所述的方法,其中所述DNA模板在5'至3'方向上包含:

a.所述组成型启动子;

b.所述启动受体;

c.所述诱导型启动子;和

d.所述嵌合抗原受体。

83.如权利要求65至82中任一项所述的方法,其中所述DNA模板还包含自切除2A肽(P2A)。

84.如权利要求83所述的方法,其中所述P2A核酸处于所述DNA模板的3'端。

85.如权利要求65至84中任一项所述的方法,其中所述DNA模板还包含土拨鼠肝炎病毒翻译后调控元件(WPRE)。

86.如权利要求85所述的方法,其中所述WPRE处于编码所述CAR的核酸的3'端和编码所述启动受体的核酸的5'端,或者其中所述WPRE处于编码所述启动受体的核酸的3'端和编码所述CAR的核酸的5'端。

87.如权利要求65至86中任一项所述的方法,其中所述启动受体在N末端至C末端方向

上包含：

a.对抗原具有结合亲和力的细胞外抗原结合结构域；

b.包含一个或多个配体诱导型蛋白水解裂解位点的跨膜结构域;和

c.包含人或人源化转录效应子的细胞内结构域,其中抗原与所述细胞外抗原结合结构域的结合导致在所述配体诱导型蛋白水解裂解位点处的裂解,从而释放所述细胞内结构域。

88.如权利要求87所述的方法,其中所述启动受体还包含定位于所述跨膜结构域与所述细胞内结构域之间的近膜结构域(JMD)。

89.如权利要求65至88中任一项所述的方法,其中所述转录因子结合至所述诱导型启动子并诱导所述CAR的表达。

90.如权利要求65至89中任一项所述的方法,其中所述CAR从N末端至C末端包含：

a.对抗原具有结合亲和力的细胞外抗原结合结构域；

b.跨膜结构域；

c.细胞内共刺激结构域;和

d.细胞内激活结构域。

91.如权利要求65至90中任一项所述的方法,其中所述启动受体和所述CAR结合不同的抗原。

92.如权利要求65至90中任一项所述的方法,其中所述启动受体和所述CAR结合相同的抗原。

93.如权利要求65至92中任一项所述的方法,其中所述免疫细胞是原代人免疫细胞。

94.如权利要求65至93中任一项所述的方法,其中所述原代免疫细胞是自体免疫细胞。

95.如权利要求65至94中任一项所述的方法,其中所述原代免疫细胞是自然杀伤(NK)细胞、T细胞、CD8+T细胞、CD4+T细胞、原代T细胞或T细胞祖细胞。

96.如权利要求65至95中任一项所述的方法,其中所述原代免疫细胞是原代T细胞。

97.如权利要求65至96中任一项所述的方法,其中所述原代免疫细胞是原代人T细胞。

98.如权利要求65至97中任一项所述的方法,其中所述原代免疫细胞是无病毒的。

99.如权利要求65至98中任一项所述的方法,所述方法还包括从患者获得所述免疫细胞以及在体外引入所述质粒。

100.一种编辑原代免疫细胞的方法,所述方法包括：

a.提供核糖核蛋白复合物(RNP)-DNA模板复合物,其中所述RNP包含核酸酶结构域和指导RNA,其中所述DNA模板的大小为大于或等于5千碱基核苷酸的大小,其中所述DNA模板包含嵌合抗原受体(CAR)和包含转录因子的启动受体,并且其中所述DNA模板的5'端和3'端包含与所述原代免疫细胞的基因组中的插入位点侧翼的基因组序列同源的核苷酸序列；

b.以非病毒方式将所述RNP-DNA模板复合物引入所述原代免疫细胞中,其中所述指导RNA与所述原代免疫细胞的基因组的靶区特异性地杂交,并且其中所述核酸酶结构域使所述靶区裂解以产生在所述原代免疫细胞的基因组中的所述插入位点;以及

c.通过将所述DNA模板插入所述原代免疫细胞的基因组中的所述插入位点中来编辑所述原代免疫细胞。

具有启动受体的工程化免疫细胞

[0001] 相关申请的交叉引用

[0002] 本申请要求2020年8月28日提交的美国临时申请第63/072,080号的利益,所述申请据此以引用的方式整体并入本文。

[0003] 序列表

[0004] 本申请含有已通过EFS网提交并且据此以引用的方式整体并入本文的序列表。所述ASCII拷贝于20XX年某月XX日创建,名为XXXXXUS_sequencelisting.txt,并且大小是X,XXX,XXX字节。

背景技术

[0005] 癌症是一种以细胞不受控制的生长为特征的疾病。已经尝试了许多治疗癌症的方法,包括药物和放射疗法。最近的癌症治疗试图利用人体自身的免疫细胞来攻击癌细胞。一种有前景的方法是使用从患者身上提取的T细胞,并通过对T细胞进行基因工程化来产生嵌合抗原受体或CAR,这种受体蛋白赋予T细胞新的靶向特定蛋白质的能力。这些受体是嵌合的,因为它们将抗原结合与T细胞激活功能结合到单个受体中。

[0006] 使用这些CAR-T细胞的免疫疗法是很有前景的,因为经修饰的T细胞有可能识别癌细胞,从而更有效地靶向和摧毁它们。

[0007] 在利用CAR对T细胞进行工程化后,将所得CAR-T细胞引入患者体内以攻击肿瘤细胞。CAR-T细胞可以源自患者自身血液中的T细胞(自体)或来自另一健康供体的T细胞(同种异体)。

[0008] CAR-T细胞一旦注入患者体内,它们就会与细胞上的靶向抗原接触。CAR-T细胞与抗原结合并被激活。抗原结合后,CAR T细胞可以呈指数增殖,启始肿瘤细胞因子的产生以及靶肿瘤细胞的杀伤。然而,基于CAR T细胞的免疫疗法仍然存在一些问题和局限性。一些CAR T细胞可能与表达低水平靶抗原的正常细胞接合,导致脱靶毒性。

[0009] 此外,使用病毒载体很难将大基因构建体插入特定基因靶标中,这是免疫细胞工程化中的一个已知瓶颈。

发明内容

[0010] 在一方面,本文提供了原代免疫细胞,所述原代免疫细胞包含至少一个以非病毒方式插入所述细胞的基因组的靶区中的DNA模板,其中所述DNA模板的大小为大于或等于约5千碱基对(kb)。

[0011] 在一些实施方案中,所述原代免疫细胞不包含用于将所述DNA模板引入所述原代免疫细胞中的病毒载体。

[0012] 在一些实施方案中,所述DNA模板的大小为大于或等于约5.0kb、5.1kb、5.2kb、5.3kb、5.4kb、5.5kb、5.6kb、5.7kb、5.8kb、5.9kb、6.0kb、6.1kb、6.2kb、6.3kb、6.4kb、6.5kb、6.6kb、6.7kb、6.8kb、6.9kb、7.0kb、7.1kb、7.2kb、7.3kb、7.4kb、7.5kb、7.6kb、7.7kb、7.8kb、7.9kb、8.0kb、8.1kb、8.2kb、8.3kb、8.4kb、8.5kb、8.6kb、8.7kb、8.8kb、

8.9kb、9.0kb、9.1kb、9.2kb、9.3kb、9.4kb、9.5kb、9.6kb、9.7kb、9.8kb.9.9kb、10.0kb或介于这些大小之间的任何DNA模板大小。

[0013] 在一些实施方案中，所述DNA模板的大小为约5kb至约10kb、约5kb至约9kb、约5kb至约8kb、约5kb至约7kb、约5kb至约6kb、约kb 6至约10kb、约6kb至约9kb、约6kb至约8kb、约6kb至约7kb、约7kb至约10kb、约7kb至约9kb、约7kb至约8kb、约8kb至约10kb、约8kb至约9kb，或约9kb至约10kb。

[0014] 在一些实施方案中，所述DNA模板是双链DNA模板或单链DNA模板。

[0015] 在一些实施方案中，所述DNA模板是线状DNA模板或环状DNA模板，任选地其中所述环状DNA模板是质粒。

[0016] 在一些实施方案中，所述DNA模板的5’端和3’端包含与所述原代细胞的基因组中的插入位点侧翼的基因组序列同源的核苷酸序列。

[0017] 在一些实施方案中，所述细胞的基因组的所述靶区是T细胞受体α恒定(TRAC)基因座或基因组安全港(GSH)。

[0018] 在一些实施方案中，所述DNA模板包含异源序列。

[0019] 在一些实施方案中，所述DNA模板包含基因。

[0020] 在一些实施方案中，所述DNA模板包含含有转录因子的启动受体。

[0021] 在一些实施方案中，所述DNA模板包含嵌合抗原受体(CAR)。

[0022] 在一些实施方案中，所述DNA模板包含嵌合抗原受体(CAR)和包含转录因子的启动受体。

[0023] 在一些实施方案中，所述DNA模板包含可操作地连接至所述嵌合抗原受体的诱导型启动子。

[0024] 在一些实施方案中，所述DNA模板还包含可操作地连接至所述启动受体的组成型启动子。

[0025] 在一些实施方案中，所述DNA模板还包含可操作地连接至所述嵌合抗原受体的诱导型启动子和可操作地连接至所述启动受体的组成型启动子。

[0026] 在一些实施方案中，所述DNA模板在5’至3’方向上包含：所述诱导型启动子；所述嵌合抗原受体；所述组成型启动子；和所述启动受体。

[0027] 在一些实施方案中，所述DNA模板在5’至3’方向上包含：所述组成型启动子；所述启动受体；所述诱导型启动子；和所述嵌合抗原受体。

[0028] 在一些实施方案中，所述DNA模板还包含自切除2A肽(P2A)。

[0029] 在一些实施方案中，所述P2A核酸处于所述DNA模板的3’端。

[0030] 在一些实施方案中，所述DNA模板还包含土拨鼠肝炎病毒翻译后调控元件(WPRE)。

[0031] 在一些实施方案中，所述WPRE处于编码所述CAR的核酸的3’端和编码所述启动受体的核酸的5’端，或者其中所述WPRE处于编码所述启动受体的核酸的3’端和编码所述CAR的核酸的5’端。

[0032] 在一些实施方案中，所述启动受体在N末端至C末端方向上包含：对抗原具有结合亲和力的细胞外抗原结合结构域；包含一个或多个配体诱导型蛋白水解裂解位点的跨膜结构域；和包含人或人源化转录效应子的细胞内结构域，其中抗原与所述细胞外抗原结合结构域的结合导致在所述配体诱导型蛋白水解裂解位点处的裂解，从而释放所述细胞内结构

域。

[0033] 在一些实施方案中,所述启动受体还包含定位于所述跨膜结构域与所述细胞内结构域之间的近膜结构域(JMD)。

[0034] 在一些实施方案中,所述转录因子结合至所述诱导型启动子并诱导所述CAR的表达。

[0035] 在一些实施方案中,所述CAR从N末端至C末端包含:对抗原具有结合亲和力的细胞外抗原结合结构域;跨膜结构域;细胞内共刺激结构域;和细胞内激活结构域。

[0036] 在一些实施方案中,所述启动受体和所述CAR结合不同的抗原。

[0037] 在一些实施方案中,所述启动受体和所述CAR结合相同的抗原。

[0038] 在一些实施方案中,所述免疫细胞是原代人免疫细胞。

[0039] 在一些实施方案中,所述原代免疫细胞是自体免疫细胞。

[0040] 在一些实施方案中,所述原代免疫细胞是自然杀伤(NK)细胞、T细胞、CD8+T细胞、CD4+T细胞、原代T细胞或T细胞祖细胞。

[0041] 在一些实施方案中,所述原代免疫细胞是原代T细胞。

[0042] 在一些实施方案中,所述原代免疫细胞是原代人T细胞。

[0043] 在一些实施方案中,所述原代免疫细胞是无病毒的。

[0044] 在另一方面,本文提供了包含多个本文公开的原代免疫细胞的细胞群体。

[0045] 在另一方面,本文提供了原代免疫细胞,所述原代免疫细胞包含至少一个插入所述原代免疫细胞的基因组的靶区中的DNA模板,其中所述DNA模板的大小为大于或等于5千碱基对,并且其中所述原代免疫细胞不包含用于将所述DNA模板引入所述原代免疫细胞中的病毒载体。

[0046] 在另一方面,本文提供了原代免疫细胞,所述原代免疫细胞包含至少一个插入所述原代免疫细胞的基因组的靶区中的DNA模板,所述DNA模板包含嵌合抗原受体(CAR)和包含转录因子的启动受体,其中所述DNA模板的大小为大于或等于5千碱基对,并且其中所述原代免疫细胞不包含用于将所述DNA模板引入所述原代免疫细胞中的病毒载体。

[0047] 在另一方面,本文提供了活的无病毒的原代细胞,所述原代细胞包含核糖核蛋白复合物(RNP)-DNA模板复合物,其中所述RNP包含核酸酶结构域和指导RNA,其中所述DNA模板的大小为大于或等于5千碱基核苷酸的大小,并且其中所述DNA模板的5'端和3'端包含与所述原代细胞的基因组中的插入位点侧翼的基因组序列同源的核苷酸序列。

[0048] 在另一方面,本文提供了活的无病毒的原代细胞,所述原代细胞包含核糖核蛋白复合物(RNP)-DNA模板复合物,其中所述RNP包含核酸酶结构域和指导RNA,其中所述DNA模板的大小为大于或等于5千碱基核苷酸的大小,其中所述DNA模板包含嵌合抗原受体(CAR)和包含转录因子的启动受体,并且其中所述DNA模板的5'端和3'端包含与所述原代细胞的基因组中的插入位点侧翼的基因组序列同源的核苷酸序列。

[0049] 在另一方面,本文提供了治疗受试者的疾病的方法,所述方法包括向受试者施用原代免疫细胞。

[0050] 在一些实施方案中,所述疾病是癌症。

[0051] 在另一方面,本文提供了非病毒载体,所述非病毒载体包含DNA模板,其中所述DNA模板的大小为大于或等于5千碱基核苷酸,并且其中所述DNA模板的5'端和3'端包含与原代

细胞的基因组中的插入位点侧翼的基因组序列同源的核苷酸序列。

[0052] 在一些实施方案中,所述DNA模板的大小为大于或等于约5.0kb、5.1kb、5.2kb、5.3kb、5.4kb、5.5kb、5.6kb、5.7kb、5.8kb、5.9kb、6.0kb、6.1kb、6.2kb、6.3kb、6.4kb、6.5kb、6.6kb、6.7kb、6.8kb、6.9kb、7.0kb、7.1kb、7.2kb、7.3kb、7.4kb、7.5kb、7.6kb、7.7kb、7.8kb、7.9kb、8.0kb、8.1kb、8.2kb、8.3kb、8.4kb、8.5kb、8.6kb、8.7kb、8.8kb、8.9kb、9.0kb、9.1kb、9.2kb、9.3kb、9.4kb、9.5kb、9.6kb、9.7kb、9.8kb.9.9kb、10.0kb或介于这些大小之间的任何DNA模板大小。

[0053] 在一些实施方案中,所述DNA模板的大小为约5kb至约10kb、约5kb至约9kb、约5kb至约8kb、约5kb至约7kb、约5kb至约6kb、约kb 6至约10kb、约6kb至约9kb、约6kb至约8kb、约6kb至约7kb、约7kb至约10kb、约7kb至约9kb、约7kb至约8kb、约8kb至约10kb、约8kb至约9kb,或约9kb至约10kb。

[0054] 在一些实施方案中,所述细胞的基因组的所述靶区是T细胞受体α恒定(TRAC)基因座或基因组安全港(GSH)。

[0055] 在一些实施方案中,所述DNA模板包含异源序列。

[0056] 在一些实施方案中,所述DNA模板包含基因。

[0057] 在一些实施方案中,所述DNA模板包含含有转录因子的启动受体。

[0058] 在一些实施方案中,所述DNA模板包含嵌合抗原受体(CAR)。

[0059] 在一些实施方案中,所述DNA模板包含嵌合抗原受体(CAR)和包含转录因子的启动受体。

[0060] 在一些实施方案中,所述DNA模板包含可操作地连接至所述嵌合抗原受体的诱导型启动子。

[0061] 在一些实施方案中,所述DNA模板还包含可操作地连接至所述启动受体的组成型启动子。

[0062] 在一些实施方案中,所述DNA模板还包含可操作地连接至所述嵌合抗原受体的诱导型启动子和可操作地连接至所述启动受体的组成型启动子。

[0063] 在一些实施方案中,所述DNA模板在5’至3’方向上包含:所述诱导型启动子;所述嵌合抗原受体;所述组成型启动子;和所述启动受体。

[0064] 在一些实施方案中,所述DNA模板在5’至3’方向上包含:所述组成型启动子;所述启动受体;所述诱导型启动子;和所述嵌合抗原受体。

[0065] 在一些实施方案中,所述DNA模板还包含自切除2A肽(P2A)。

[0066] 在一些实施方案中,所述P2A处于所述DNA模板的3’端。

[0067] 在一些实施方案中,所述DNA模板还包含土拨鼠肝炎病毒翻译后调控元件(WPRE)。

[0068] 在一些实施方案中,所述启动受体在N末端至C末端方向上包含:对抗原具有结合亲和力的细胞外抗原结合结构域;包含一个或多个配体诱导型蛋白水解裂解位点的跨膜结构域;和包含人或人源化转录效应子的细胞内结构域,其中抗原与所述细胞外抗原结合结构域的结合导致在所述配体诱导型蛋白水解裂解位点处的裂解,从而释放所述细胞内结构域。

[0069] 在一些实施方案中,所述启动受体还包含定位于所述跨膜结构域与所述细胞内结构域之间的近膜结构域(JMD)。

[0070] 在一些实施方案中,所述转录因子结合至所述诱导型启动子并诱导所述CAR的表达。

[0071] 在一些实施方案中,所述CAR从N末端至C末端包含:对抗原具有结合亲和力的细胞外抗原结合结构域;跨膜结构域;细胞内共刺激结构域;和细胞内激活结构域。

[0072] 在一些实施方案中,所述启动受体和所述CAR结合不同的抗原。

[0073] 在一些实施方案中,所述启动受体和所述CAR结合相同的抗原。

[0074] 在另一方面,本文提供了编辑原代免疫细胞的方法,所述方法包括:提供核糖核蛋白复合物(RNP)-DNA模板复合物,其中所述RNP包含核酸酶结构域和指导RNA,其中所述DNA模板的大小为大于或等于5千碱基核苷酸的大小,并且其中所述DNA模板的5'端和3'端包含与所述原代免疫细胞的基因组中的插入位点侧翼的基因组序列同源的核苷酸序列;以非病毒方式将所述RNP-DNA模板复合物引入所述原代免疫细胞中,其中所述指导RNA与所述原代免疫细胞的基因组的靶区特异性地杂交,并且其中所述核酸酶结构域使所述靶区裂解以产生在所述原代免疫细胞的基因组中的所述插入位点;以及通过将所述DNA模板插入所述原代免疫细胞的基因组中的所述插入位点中来编辑所述原代免疫细胞。

[0075] 在一些实施方案中,以非病毒方式引入包括电穿孔。

[0076] 在一些实施方案中,所述核酸酶结构域包含CRISPR相关核酸内切酶(Cas),任选地是Cas9核酸酶。

[0077] 在一些实施方案中,所述DNA模板的大小为大于或等于约5.0kb、5.1kb、5.2kb、5.3kb、5.4kb、5.5kb、5.6kb、5.7kb、5.8kb、5.9kb、6.0kb、6.1kb、6.2kb、6.3kb、6.4kb、6.5kb、6.6kb、6.7kb、6.8kb、6.9kb、7.0kb、7.1kb、7.2kb、7.3kb、7.4kb、7.5kb、7.6kb、7.7kb、7.8kb、7.9kb、8.0kb、8.1kb、8.2kb、8.3kb、8.4kb、8.5kb、8.6kb、8.7kb、8.8kb、8.9kb、9.0kb、9.1kb、9.2kb、9.3kb、9.4kb、9.5kb、9.6kb、9.7kb、9.8kb.9.9kb、10.0kb或介于这些大小之间的任何DNA模板大小。

[0078] 在一些实施方案中,所述DNA模板的大小为约5kb至约10kb、约5kb至约9kb、约5kb至约8kb、约5kb至约7kb、约5kb至约6kb、约kb 6至约10kb、约6kb至约9kb、约6kb至约8kb、约6kb至约7kb、约7kb至约10kb、约7kb至约9kb、约7kb至约8kb、约8kb至约10kb、约8kb至约9kb,或约9kb至约10kb。

[0079] 在一些实施方案中,所述细胞的基因组的所述靶区是T细胞受体α恒定(TRAC)基因座或基因组安全港(GSH)。

[0080] 在一些实施方案中,所述DNA模板是双链DNA模板或单链DNA模板。

[0081] 在一些实施方案中,所述DNA模板是线状DNA模板或环状DNA模板,任选地其中所述环状DNA模板是质粒。

[0082] 在一些实施方案中,所述DNA模板包含异源序列。

[0083] 在一些实施方案中,所述DNA模板包含基因。

[0084] 在一些实施方案中,所述DNA模板包含含有转录因子的启动受体。

[0085] 在一些实施方案中,所述DNA模板包含嵌合抗原受体(CAR)。

[0086] 在一些实施方案中,所述DNA模板包含嵌合抗原受体(CAR)和包含转录因子的启动受体。

[0087] 在一些实施方案中,所述DNA模板包含可操作地连接至所述嵌合抗原受体的诱导

型启动子。

[0088] 在一些实施方案中,所述DNA模板还包含可操作地连接至所述启动受体的组成型启动子。

[0089] 在一些实施方案中,所述DNA模板还包含可操作地连接至所述嵌合抗原受体的诱导型启动子和可操作地连接至所述启动受体的组成型启动子。

[0090] 在一些实施方案中,所述DNA模板在5'至3'方向上包含:所述诱导型启动子;所述嵌合抗原受体;所述组成型启动子;和所述启动受体。

[0091] 在一些实施方案中,所述DNA模板在5'至3'方向上包含:所述组成型启动子;所述启动受体;所述诱导型启动子;和所述嵌合抗原受体。

[0092] 在一些实施方案中,所述DNA模板还包含自切除2A肽(P2A)。

[0093] 在一些实施方案中,所述P2A核酸处于所述DNA模板的3'端。

[0094] 在一些实施方案中,所述DNA模板还包含土拨鼠肝炎病毒翻译后调控元件(WPRE)。

[0095] 在一些实施方案中,所述WPRE处于编码所述CAR的核酸的3'端和编码所述启动受体的核酸的5'端,或者其中所述WPRE处于编码所述启动受体的核酸的3'端和编码所述CAR的核酸的5'端。

[0096] 在一些实施方案中,所述启动受体在N末端至C末端方向上包含:对抗原具有结合亲和力的细胞外抗原结合结构域;包含一个或多个配体诱导型蛋白水解裂解位点的跨膜结构域;和包含人或人源化转录效应子的细胞内结构域,其中抗原与所述细胞外抗原结合结构域的结合导致在所述配体诱导型蛋白水解裂解位点处的裂解,从而释放所述细胞内结构域。

[0097] 在一些实施方案中,所述启动受体还包含定位于所述跨膜结构域与所述细胞内结构域之间的近膜结构域(JMD)。

[0098] 在一些实施方案中,所述转录因子结合至所述诱导型启动子并诱导所述CAR的表达。

[0099] 在一些实施方案中,所述CAR从N末端至C末端包含:对抗原具有结合亲和力的细胞外抗原结合结构域;跨膜结构域;细胞内共刺激结构域;和细胞内激活结构域。

[0100] 在一些实施方案中,所述启动受体和所述CAR结合不同的抗原。

[0101] 在一些实施方案中,所述启动受体和所述CAR结合相同的抗原。

[0102] 在一些实施方案中,所述免疫细胞是原代人免疫细胞。

[0103] 在一些实施方案中,所述原代免疫细胞是自体免疫细胞。

[0104] 在一些实施方案中,所述原代免疫细胞是自然杀伤(NK)细胞、T细胞、CD8+T细胞、CD4+T细胞、原代T细胞或T细胞祖细胞。

[0105] 在一些实施方案中,所述原代免疫细胞是原代T细胞。

[0106] 在一些实施方案中,所述原代免疫细胞是原代人T细胞。

[0107] 在一些实施方案中,所述原代免疫细胞是无病毒的。

[0108] 在一些实施方案中,所述方法还包括从患者获得所述免疫细胞以及在体外引入所述质粒。

[0109] 在另一方面,本文提供了编辑原代免疫细胞的方法,所述方法包括:提供核糖核蛋白复合物(RNP)-DNA模板复合物,其中所述RNP包含核酸酶结构域和指导RNA,其中所述DNA

模板的大小为大于或等于5千碱基核苷酸的大小，其中所述DNA模板包含嵌合抗原受体(CAR)和包含转录因子的启动受体，并且其中所述DNA模板的5'端和3'端包含与所述原代免疫细胞的基因组中的插入位点侧翼的基因组序列同源的核苷酸序列；以非病毒方式将所述RNP-DNA模板复合物引入所述原代免疫细胞中，其中所述指导RNA与所述原代免疫细胞的基因组的靶区特异性地杂交，并且其中所述核酸酶结构域使所述靶区裂解以产生在所述原代免疫细胞的基因组中的所述插入位点；以及通过将所述DNA模板插入所述原代免疫细胞的基因组中的所述插入位点中来编辑所述原代免疫细胞。

附图说明

[0110] 参考以下描述和附图将更好地理解本公开的这些和其他特征、方面和优点，在附图中：

[0111] 图1示出了本发明的示例性质粒。

[0112] 图2示出了产生工程化免疫细胞的示例性方法。

[0113] 图3示出了实验方案的示意图。

[0114] 图4提供了显示所需受体表达和控制的实验结果。

[0115] 图5示出了实验方案的示意图。

[0116] 图6提供了显示少量不需要的CAR表达的实验结果。

[0117] 图7示出了11号染色体上的GS94整合位点的示意图。

[0118] 图8A包括显示工程化T细胞在与K562-PrimeR细胞共培养48小时后的CAR诱导和primeR表达的图。图8B包括显示工程化T细胞在与K562-primeR/CAR细胞共培养48小时后的细胞毒性和细胞因子分泌水平的图。

[0119] 图9A包括显示用于评价整合位点对细胞毒性的影响的实验概览的示意图。图9B包括显示与K562-primeR/CAR或K562-CAR细胞共培养48小时的工程化T细胞的测量细胞毒性的图。

[0120] 图10A包括显示用于评价整合位点对细胞因子分泌的影响的实验概览的示意图。图10B包括显示与K562-primeR/CAR细胞共培养48小时的工程化T细胞的测量细胞因子分泌的图。

[0121] 图11包括显示为确定整合位点对primeR非依赖性CAR表达的影响而进行的体外实验的示意图。“Flow”是指流式细胞术，“restim”是指对工程化T细胞的重复CD3/CD28刺激。“EP”是指电穿孔。

[0122] 图12A和图12B包括显示当使用所指示整合位点时PrimeR表达随时间的稳定性的图。“Flow”是指流式细胞术，“restim”是指对工程化T细胞的重复CD3/CD28刺激。“EP”是指电穿孔。在图12B中，对PrimeR表达针对使用TRAC整合位点时的表达作归一化。

[0123] 图13A包括显示iGuide-Seq测定技术的示意图。图13B包括显示使用iGuide Seq测定的所指示整合位点的中靶效率的图。图13C包括来自iGuide-Seq分析的示意图，显示在两名供体之间GS94没有可重现的推定脱靶。

[0124] 图14包括显示iGuide-Seq工作流程和数据的示意图。

[0125] 图15包括显示通过iGUIDE-seq和Elevation预测确定的推定脱靶位点的rhAmp-seq分析的图。

[0126] 图16包括显示在GS94、GS102和TRAC敲入primeR/CAR回路的细胞的RNA-seq分析的图。两名供体中在GS94基因座处整合的细胞(y轴)与在TRAC或GS102基因座处整合的细胞(x轴)的基因表达的散点图。浅灰色点对应于ETS1和FLI1。深灰色是使用edgeR发现差异性表达的基因(倍数变化>0,FDR校正的p值<0.01,比较条件下的平均每百万计数至少为2)。

[0127] 图17包括显示在GS94处敲入primeR/CAR回路的细胞中不存在细胞因子非依赖性生长的图。

[0128] 图18示出了包含primeR和CAR的8.3kb转基因回路在K562细胞中的表达。

[0129] 图19A示出了插入GS94安全港基因座的4.6kb盒的图解。图19B示出了插入GS94安全港基因座的8.3kb盒的图解。

具体实施方式

[0130] 本发明提供了新型工程化免疫细胞以及制备和使用此类细胞的方法。工程化免疫细胞包括插入免疫细胞基因组中的转基因,所述转基因编码工程化启动受体和嵌合抗原受体(CAR)的基因。最初,工程化免疫细胞在细胞表面上表达由质粒编码的启动受体,但不表达CAR。启动受体包括可裂解的转录因子。当本发明的免疫细胞遇到靶细胞时,启动受体结合至靶细胞上的同源配体。这导致启动受体释放可裂解的转录因子。转录因子的释放使得工程化免疫细胞在细胞表面上表达由整合的转基因编码的CAR。

[0131] 当在免疫细胞表面上表达时,CAR可用于结合至靶细胞上的同源配体。CAR与其靶配体的结合会触发工程化免疫细胞的细胞毒性应答,从而杀伤靶细胞。

[0132] 本发明的工程化免疫细胞与现有工程化免疫细胞相比具有若干优点。下文描述的实施方案表明,可以使用靶向非病毒递送将合成回路递送至T细胞基因组中,并且可以在基因组整合后保持回路保真度。这增强了安全特性并扩大了细胞的治疗窗口,因为靶向整合到TRAC启动子中降低了脱靶基因插入基因组中的风险,如通过慢病毒基因递送发生的那样。此外,所述回路还提供另外的安全特征。例如,细胞毒性CAR不可用于激活,因为它不表达,直到它遇到具有启动受体同源配体的靶细胞。此外,由于CAR最初未表达,因此细胞不太可能表现出耗竭、分化、强直信号传导和激活诱导的免疫细胞死亡,同时表现出高增殖和持久性潜力。

[0133] 定义

[0134] 除非另外规定,否则权利要求和说明书中所用的术语是如下文所陈述加以定义。

[0135] 如本文所用,术语“基因”是指遗传的基本单位,由沿染色体排列的DNA区段组成,其编码特定蛋白质或蛋白质区段。基因通常包括启动子、5’非翻译区、一个或多个编码序列(外显子)、任选的内含子和3’非翻译区。基因还可包含终止子、增强子和/或沉默子。

[0136] 如本文所用,术语“基因座”是指基因或遗传标志物所在的染色体上特定的、固定的物理位置。

[0137] 术语“安全港基因座”是指基因或遗传元件可以在不破坏相邻基因的表达或调控的情况下被并入的基因座。这些安全港基因座也称为安全港位点(SHS)。如本文所用,安全港基因座是指“整合位点”或“敲入位点”,在该位点上可以插入编码如本文所定义的转基因的序列。在一些实施方案中,插入发生伴随着位于整合位点的序列的置换。在一些实施方案中,插入发生不伴随位于整合位点的序列的置换。设想的整合位点的实例提供于表1中。

[0138] 如本文所用,术语“插入物”是指整合(插入)到靶基因座或安全港位点处的核苷酸序列。插入物可用于指使用例如同源定向修复(HDR)CRISPR/Cas9基因组编辑或本领域普通技术人员已知的将核苷酸序列插入基因组区域中的其他方法被并入靶基因座或安全港位点的基因或遗传元件。

[0139] “CRISPR/Cas”系统是指一类广泛用于防御外来核酸的细菌系统。CRISPR/Cas系统广泛存在于真细菌和古细菌生物体中。CRISPR/Cas系统包括I型、II型和III型亚型。野生型II型CRISPR/Cas系统利用RNA介导的核酸酶Cas9与指导和激活性RNA的复合物来识别和裂解外来核酸。具有指导RNA和激活性RNA两者的活性的指导RNA在本领域中也是已知的。在一些情况下,这种双重活性指导RNA被称为小指导RNA(sgRNA)。

[0140] Cas9同系物存在于多种真细菌中,包括但不限于以下分类群的细菌:放线菌门(Actinobacteria)、产水菌门(Aquificae)、拟杆菌门-绿菌门(Bacteroidetes-Chlorobi)、衣原体-疣微菌门(Chlamydiae-Verrucomicrobia)、绿弯菌门(Chlroflexi)、蓝藻门(Cyanobacteria)、厚壁菌门(Firmicutes)、变形菌门(Proteobacteria)、螺旋体门(Spirochaetes)和热袍菌门(Thermotogae)。示例性的Cas9蛋白是化脓性链球菌(Streptococcus pyogenes)Cas9蛋白。另外的Cas9蛋白及其同系物描述于例如Chylinksi等人,RNA Biol.2013年5月1日;10(5):726-737;Nat.Rev.Microbiol.2011年6月;9(6):467-477;Hou等人,Proc Natl Acad Sci U S A.2013年9月24日;110(39):15644-9;Sampson等人,Nature.2013年5月9日;497(7448):254-7;以及Jinek等人,Science.2012年8月17日;337(6096):816-21。可以优化Cas9核酸酶结构域以提高宿主细胞中的活性或增强稳定性。

[0141] 如本文所用,术语“Cas9”是指RNA介导的核酸酶(例如,细菌或古细菌来源的,或源自它们的)。示例性的RNA介导的核酸酶包括上述Cas9蛋白及其同系物,并且包括但不限于CPF1(参见,例如,Zetsche等人,Cell,第163卷,第3期,p759-771,2015年10月22日)。类似地,如本文所用,术语“Cas9核糖核蛋白”复合物等是指Cas9蛋白与crRNA(例如,指导RNA或小指导RNA)之间的复合物、Cas9蛋白与反式激活性crRNA(tracrRNA)之间的复合物、Cas9蛋白与小指导RNA之间的复合物或它们的组合(例如,含有Cas9蛋白、tracrRNA和crRNA指导RNA的复合物)。

[0142] 如本文所用,术语“T淋巴细胞”和“T细胞”可互换使用,并且是指在胸腺中已完成成熟并识别体内某些外来抗原的细胞。这些术语还指在免疫系统中具有各种作用,包括其他免疫细胞的激活和失活的主要白细胞类型。T细胞可以是任何T细胞,诸如培养的T细胞,例如原代T细胞,或源自培养的T细胞系的T细胞,例如Jurkat、SupT1等,或获自哺乳动物的T细胞。T细胞包括但不限于幼稚T细胞、刺激T细胞、原代T细胞(例如,未培养的)、培养的T细胞、永生化T细胞、辅助性T细胞、细胞毒性T细胞、记忆T细胞、调节性T细胞、自然杀伤T细胞、它们的组合或它们的亚群。T细胞可以是CD3+细胞。T细胞可以是$CD4^+$、$CD8^+$或$CD4^+$和$CD8^+$。T细胞可以是任何类型的T细胞、CD4+/CD8+双阳性T细胞、CD4+辅助性T细胞(例如Th1和Th2细胞)、CD8+T细胞(例如细胞毒性T细胞),外周(包括但不限于)血单核细胞(PBMC)、外周血白细胞(PBL)、肿瘤浸润性淋巴细胞(TIL)、记忆T细胞、幼稚T细胞、调节性T细胞、γδT细胞等。其可以是任何发育阶段的任何T细胞。其他类型的辅助性T细胞包括Th3(Treg)细胞、Th17细胞、Th9细胞或Tfh细胞。其他类型的记忆T细胞包括诸如中央记忆T细胞(Tcm细胞)、效应记

忆T细胞(Tem细胞和TEMRA细胞)的细胞。T细胞还可以指代基因修饰的T细胞,诸如经修饰以表达T细胞受体(TCR)或嵌合抗原受体(CAR)的T细胞。T细胞也可以从干细胞或祖细胞分化而来。

[0143] “CD4+T细胞”是指在其表面上表达CD4并与细胞免疫应答相关的T细胞亚群。CD4+T细胞的特征在于刺激后分泌特性,可包括分泌细胞因子,诸如IFN-γ、TNF-α、IL-2、IL-4和IL-10。“CD4”是一种55kD的糖蛋白,最初被定义为T淋巴细胞上的分化抗原,但也存在于包括单核细胞/巨噬细胞在内的其他细胞上。CD4抗原是免疫球蛋白超家族的成员,并被认为是MHC(主要组织相容性复合物)II类限制性免疫应答中的关联识别元件。在T淋巴细胞上,CD4抗原限定了辅助/诱导子亚群。

[0144] “CD8+T细胞”是指在其表面上表达CD8的T细胞亚群,受MHC I类限制,并作为细胞毒性T细胞发挥作用。“CD8”分子是存在于胸腺细胞以及细胞毒性和抑制性T淋巴细胞上的分化抗原。CD8抗原是免疫球蛋白超家族的成员,并且是主要组织相容性复合物I类限制性相互作用中的关联识别元件。

[0145] 如本文所用,短语“造血干细胞”是指一种可以产生血细胞的干细胞。造血干细胞可以产生骨髓或淋巴谱系或其组合的细胞。造血干细胞主要存在于骨髓中,但是它们可以从外周血或其部分中分离出来。各种细胞表面标志物可用于鉴定、分选或纯化造血干细胞。在一些情况下,造血干细胞被鉴定为$c\text{-}kit^{+}$和lin^{-}。在一些情况下,人造血干细胞被鉴定为$CD34^{+}$、$CD59^{+}$、$Thy1/CD90^{+}$、$CD38^{lo/-}$、$C\text{-}kit/CD117^{+}$、lin^{-}。在一些情况下,人造血干细胞被鉴定为$CD34^{-}$、$CD59^{+}$、$Thy1/CD90^{+}$、$CD38^{lo/-}$、$C\text{-}kit/CD117^{+}$、lin^{-}。在一些情况下,人造血干细胞被鉴定为$CD133^{+}$、$CD59^{+}$、$Thy1/CD90^{+}$、$CD38^{lo/-}$、$C\text{-}kit/CD117^{+}$、lin^{-}。在一些情况下,小鼠造血干细胞被鉴定为$CD34^{lo/-}$、$SCA\text{-}1^{+}$、$Thy1^{+/lo}$、$CD38^{+}$、$C\text{-}kit^{+}$、lin^{-}。在一些情况下,造血干细胞为$CD150^{+}CD48^{-}CD244^{-}$。

[0146] 如本文所用,短语“造血细胞”是指源自造血干细胞的细胞。造血细胞可通过从生物体、系统、器官或组织(例如,血液或其部分)中分离而获得或提供。或者,可以分离造血干细胞并通过分化干细胞来获得或提供造血细胞。造血细胞包括分化成其他细胞类型潜能有限的细胞。此类造血细胞包括但不限于多能祖细胞、谱系限制性祖细胞、普通骨髓祖细胞、粒细胞-巨噬细胞祖细胞或巨核细胞-红系细胞祖细胞。造血细胞包括淋巴和骨髓谱系的细胞,诸如淋巴细胞、红细胞、粒细胞、单核细胞和血小板。

[0147] 如本文所用,短语“免疫细胞”包括产生免疫细胞的所有细胞类型,包括造血细胞、多能干细胞和诱导型多能干细胞(iPSC)。在一些实施方案中,免疫细胞是B细胞、巨噬细胞、自然杀伤(NK)细胞、诱导型多能干细胞(iPSC)、人多能干细胞(HSPC)、T细胞或T细胞祖细胞或树突细胞。在一些实施方案中,细胞是先天免疫细胞。

[0148] 如本文所用,在原代细胞或原代干细胞的上下文中的术语“原代”是指尚未转化或永生化的细胞。此类原代细胞可以经过培养、继代培养或传代有限次数(例如,培养0、1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19或20次)。在一些情况下,原代细胞适应体外培养条件。在一些情况下,原代细胞从生物体、系统、器官或组织中分离出来,任选地经分选和利用,例如不经培养或继代培养而直接使用。在一些情况下,原代细胞经刺激、激活或分化。例如,可以通过接触(例如,培养时存在)CD3、CD28激动剂、IL-2、IFN-γ或它们的组合来激活原代T细胞。

[0149] 如本文所用,术语“离体”一般包括在活组织中或活组织上进行的实验或测量,优选地是在生物体外的人工环境中,优选地是与自然条件的差异极小。

[0150] 如本文所用,术语“构建体”是指分子(包括大分子或多核苷酸)的复合物。

[0151] 如本文所用,术语“整合”是指将构建体的一个或多个核苷酸稳定插入细胞基因组中,即,共价连接至细胞染色体DNA中的核酸序列的过程。其也可以指整合位点处的核苷酸缺失。在插入位点处存在缺失的情况下,“整合”还可包括用一个或多个插入的核苷酸取代缺失的内源序列或核苷酸。

[0152] 如本文所用,术语“外源”是指已引入宿主细胞中且不是该细胞原生的分子或活性。例如,可以通过将编码核酸引入宿主遗传物质中,如通过整合到宿主染色体中,或作为非染色体遗传物质诸如质粒,来引入分子。因此,当与编码核酸的表达结合使用时,该术语是指将编码核酸以可表达的形式引入细胞中。术语“内源”是指在自然、未经编辑的条件下存在于宿主细胞中的分子或活性。类似地,当与编码核酸的表达结合使用时,该术语是指包含在细胞内而不是外源性地引入的编码核酸。

[0153] 术语“异源”指非侧翼序列原生的核酸或多肽序列或结构域,例如,其中未发现异源序列天然偶联至在一端或两端出现的核酸或多肽序列。

[0154] 术语“同源”指侧翼序列原生的核酸或多肽序列或结构域,例如,其中发现同源序列天然偶联至在一端或两端出现的核酸或多肽序列。

[0155] 如本文所用,“多核苷酸供体构建体”是指遗传地插入多核苷酸中并且对该多核苷酸而言是外源的核苷酸序列(例如DNA序列)。多核苷酸供体构建体被转录成RNA并任选地被翻译成多肽。多核苷酸供体构建体可包括原核序列、来自真核mRNA的cDNA、来自真核(例如,哺乳动物)DNA的基因组DNA序列,和合成DNA序列。例如,多核苷酸供体构建体可以是miRNA、shRNA、天然多肽(即,天然存在的多肽)或其片段,或变体多肽(例如,与天然多肽具有小于100%序列同一性的天然多肽)或其片段。

[0156] 如本文所用,术语“互补的”或“互补性”是指核苷酸或核酸之间的特定碱基配对。互补核苷酸一般是A和T(或A和U)以及G和C。本文所述的指导RNA可以包含序列,例如与细胞中的基因组序列完美互补或基本上互补(例如,具有1-4个错配)的DNA靶向序列。

[0157] 如本文所用,术语“转基因”是指已自然转移或通过多种基因工程化技术中的任何一种从一种生物体转移到另一种生物体的多核苷酸。其任选地被翻译成多肽。其任选地被翻译成重组蛋白。“重组蛋白”是一种由基因编码的蛋白质——重组DNA——已在支持基因表达和信使RNA翻译的系统中克隆(参见表达系统)。重组蛋白可以是治疗剂,例如治疗本文公开的疾病或病症的蛋白质。如所用,转基因可以指代编码多肽的多核苷酸。转基因还可以指代非编码序列,诸如但不限于shRNA、miRNA和miR。

[0158] 术语“蛋白质”、“多肽”和“肽”在本文中可互换使用。

[0159] 如本文所用,术语“可操作地连接”或“操作性地连接”是指核酸序列与单个核酸片段结合,使得一种功能受另一种功能影响。例如,如果启动子能够影响编码序列或功能性RNA的表达(即,编码序列或功能性RNA在启动子的转录控制下),则启动子与其可操作地连接。编码序列可以在有义和反义方向上可操作地连接至控制序列。

[0160] 如本文所用,术语“发育细胞状态”是指,例如,细胞所处的非活性、活性表达、分化、衰老等状态。发育细胞状态还可以指处于前体状态的细胞(例如,T细胞前体)。

[0161] 如所用,术语“编码”是指编码目标蛋白质或多肽的核酸序列。核酸序列可以是DNA分子或RNA分子。在优选实施方案中,分子是DNA分子。在其他优选实施方案中,分子是RNA分子。当作为RNA分子存在时,其将包含引导宿主细胞的核糖体开始翻译的序列(例如,起始密码子,ATG)和引导核糖体结束翻译的序列(例如,终止密码子)。在起始密码子与终止密码子之间是开放阅读框(ORF)。此类术语是本领域普通技术人员已知的。

[0162] 术语“插入”是指操纵核苷酸序列以引入非天然序列。这是例如通过使用限制酶和连接酶来完成的,由此可以通过以下方式将通常编码目标基因的目标DNA序列并入另一核酸分子中:用适当的限制酶消化这两个分子以便产生相容的重叠,接着使用连接酶将这些分子接合在一起。本领域的技术人员非常熟悉此类操纵,实例可见于Sambrook等人(Sambrook、Fritsch和Maniatis,“Molecular Cloning:A Laboratory Manual”,2nd ed.,Cold Spring Harbor Laboratory,1989),其据此以引用的方式整体并入本文,包括任何附图、图形和表格。

[0163] 如本文所用,术语“受试者”是指哺乳动物受试者。示例性受试者包括人、猴子、狗、猫、小鼠、大鼠、牛、马、骆驼、山羊、兔子、猪和绵羊。在某些实施方案中,受试者为人。在一些实施方案中,受试者患有可用本文提供的工程化细胞或其群体治疗的疾病或疾患。在一些方面,疾病或疾患为癌症。

[0164] 如本文所用,术语“启动子”是指能够控制编码序列或功能性RNA的表达的核苷酸序列(例如DNA序列)。启动子序列由近端和更远端的上游元件组成,后者元件通常称为增强子。启动子可以完整地源自天然基因,可以由来自天然存在的不同启动子的不同元件组成,并且/或者可包含合成的DNA区段。如本文所设想的,启动子对于目标细胞而言可以是内源的或者对于目标细胞而言是外源的。本领域技术人员理解不同的启动子可以在不同的组织或细胞类型中,或在不同的发育阶段,或响应于不同的环境条件诱导基因表达。如本领域已知的,可以根据启动子的强度和/或启动子具有活性的条件来选择启动子,例如,组成型启动子、强启动子、弱启动子、诱导型/抑制型启动子、组织特异性或发育调控启动子、细胞周期依赖性启动子等。

[0165] 启动子可以是诱导型启动子(例如,热休克启动子、四环素调控启动子、类固醇调控启动子、金属调控启动子、雌激素受体调控启动子等)。启动子可以是组成型启动子(例如,CMV启动子、UBC启动子)。在一些实施方案中,启动子可以是空间限制和/或时间限制的启动子(例如,组织特异性启动子、细胞类型特异性启动子等)。参见例如美国公布20180127786,其公开内容以引用的方式整体并入本文。

[0166] 如本文所设想的,基因编辑可以涉及基因(或核苷酸序列)敲入或敲除。如本文所用,术语“敲入”是指将DNA序列或其片段添加到基因组中。此类待敲入的DNA序列可包括一个或多个完整基因,可包括与基因或前述基因的任何部分或片段相关的调控序列。例如,可将编码重组蛋白的多核苷酸供体构建体插入携带突变基因的细胞的基因组中。在一些实施方案中,敲入策略涉及用提供的序列取代现有序列,例如用野生型拷贝取代突变等位基因。另一方面,术语“敲除”是指基因或基因表达的消除。例如,可以通过缺失或添加导致阅读框破坏的核苷酸序列来敲除基因。再如,可通过用不相关的(例如,非编码的)序列置换基因的一部分来敲除基因。

[0167] 如本文所用,术语“非同源末端接合”或NHEJ是指DNA链的切割或切口末端直接接

合而不需要同源模板核酸的细胞过程。NHEJ可以导致修复位点处一个或多个核苷酸的添加、缺失、取代或它们的组合。

[0168] 如本文所用，术语“同源定向修复”或HDR是指通过从同源模板核酸聚合来修复DNA链的切割或缺口末端的细胞过程。因此，原始序列被模板的序列置换。同源模板核酸可以由基因组中其他位置的同源序列(姐妹染色单体、同源染色体，或相同或不同染色体上的重复区域)提供。或者，可以引入外源模板核酸以获得靶位点处特定HDR诱导的序列的变化。通过这种方式，可以在切割位点引入特定的突变。

[0169] 如本文所用，单链DNA模板或双链DNA模板是指可被细胞用作HDR的模板的DNA寡核苷酸。一般来说，单链DNA模板或双链DNA模板具有至少一个与靶位点同源的区域。在一些情况下，单链DNA模板或双链DNA模板具有两个同源区域，位于包含待插入靶切割位点处的异源序列的区域的两侧。

[0170] 术语“载体”和“质粒”可互换使用，并且如本文所用是指可用于将遗传物质引入细胞中的多核苷酸媒介物。载体可以是线状的或环状的。载体可以整合到宿主细胞的靶基因组中或者在宿主细胞中独立复制。载体可以包含例如复制起点、多克隆位点和/或选择性标志物。表达载体通常包含表达盒。载体和质粒包括但不限于整合载体、原核质粒、真核质粒、植物合成染色体、附加体、粘粒和人工染色体。

[0171] 如本文所用，在引入核酸或包含核酸的复合物例如RNP-DNA模板复合物的上下文中的短语“引入”是指核酸序列或RNP-DNA模板复合物从细胞外到细胞内的易位。在一些情况下，引入是指核酸或复合物从细胞外到细胞核内的易位。设想了这种易位的各种方法，包括但不限于电穿孔、与纳米线或纳米管接触、受体介导的内化、通过细胞穿透肽的易位、脂质体介导的易位等。

[0172] 如本文所用，术语“表达盒”是以重组或合成方式产生的多核苷酸构建体，其包含可操作地连接至所选多核苷酸以促进该所选多核苷酸在宿主细胞中的表达的调控序列。例如，调控序列可以促进所选多核苷酸在宿主细胞中的转录，或所选多核苷酸在宿主细胞中的转录和翻译。表达盒可以例如整合到宿主细胞的基因组中或存在于表达载体中。

[0173] 如本文所用，短语“有需要的受试者”是指表现出和/或被诊断具有如本文所述的疾病或病症的一种或多种症状或体征的受试者。

[0174] “化学治疗剂”是指可用于治疗癌症的化学化合物。化学治疗剂包括用于调节、减少、阻断或抑制可促进癌症生长的激素的影响的“抗激素剂”或“内分泌治疗剂”。

[0175] 术语“组合物”是指含有例如本文所设想的工程化细胞或蛋白质的混合物。在一些实施方案中，组合物可含有另外的组分，诸如佐剂、稳定剂、赋形剂等。术语“组合物”或“药物组合物”是指呈允许包含在其中的活性成分的生物活性对治疗受试者有效的形式的制剂，并且在药物组合物中提供的量中，其不含有对受试者具有不可接受的毒性的其它组分。

[0176] 术语“改善”是指在治疗疾病状态(例如，癌症疾病状态)时的任何治疗方面有益的结果，包括对疾病状态的预防、严重性或进展的减轻、缓解或治愈。

[0177] 术语“原位”是指在活细胞中发生的与活生物体分开生长的过程，例如，在组织培养物中生长。

[0178] 术语“体内”是指在活生物体中发生的过程。

[0179] 如本文所用，术语“哺乳动物”包括人和非人，并且包括但不限于人、非人灵长类动

物、犬科动物、猫科动物、鼠类、牛、马和猪。

[0180] 在两个或更多个核酸或多肽序列的上下文中，术语“同一性”百分比是指当针对最大对应而进行比较和比对时具有指定百分比的为相同的核苷酸或氨基酸残基的两个或更多个序列或子序列，如使用下文描述的序列比较算法之一(例如，BLASTP和BLASTN或其他技术人员可用的算法)或通过视觉检查来测量。根据应用，“同一性”百分比可以存在于所比较的序列的区上，例如，存在于功能性结构域上，或者可替代地存在于待比较的两个序列的全长上。

[0181] 对于序列比较，通常一个序列充当与测试序列进行比较的参考序列。当使用序列比较算法时，将测试序列和参考序列输入计算机中，必要时指定子序列坐标，并且指定序列算法程序参数。序列比较算法然后基于指定的程序参数计算测试序列相对于参考序列的序列同一性百分比。

[0182] 例如，用于比较的最佳序列比对可以通过Smith&Waterman,Adv.Appl.Math.2:482(1981)的局部同源性算法；通过Needleman&Wunsch,J.Mol.Biol.48:443(1970)的同源性比对算法；通过Pearson&Lipman,Proc.Nat'l.Acad.Sci.USA 85:2444(1988)的相似性搜索方法；通过这些算法的计算机化实施(Wisconsin Genetics Software Package,Genetics Computer Group,575Science Dr.,Madison,Wis.中的GAP、BESTFIT、FASTA和TFASTA)；或通过目视检查(一般参见Ausubel等人，下文)来进行。

[0183] 适用于确定序列同一性和序列相似性百分比的算法的一个实例是Altschul等人，J.Mol.Biol.215:403-410(1990)中描述的BLAST算法。用于进行BLAST分析的软件可通过国家生物技术信息中心(www.ncbi.nlm.nih.gov/)公开获得。

[0184] 术语“足够的量”是指足以产生所需效果的量，例如足以调节细胞内蛋白质聚集的量。

[0185] 术语“治疗有效量”是有效改善疾病症状的量。治疗有效量可以是“防治有效量”，因为防治可以被认为是治疗。

[0186] 如本文所用，术语“有效量”是指化合物(例如，本文所述的组合物、本文所述的细胞)足以实现有益或所需结果的量。有效量可以按一次或多次施用、应用或剂量施用，并且不意图限于特定的制剂或施用途径。如本文所用，术语“治疗”包括导致疾患、疾病、病症等的改善或改善其症状的任何作用，例如减轻、减少、调节、改善或消除。

[0187] 术语“调节(modulate/modulation)”是指降低或抑制，或替代地，激活或增加所叙述变量。

[0188] 术语“增加”和“活化”是指所叙述变量增加10％、20％、30％、40％、50％、60％、70％、75％、80％、85％、90％、95％、100％，增加至2倍、3倍、4倍、5倍、10倍、20倍、50倍、100倍或更大。

[0189] 术语“降低”和“抑制”是指所叙述变量降低10％、20％、30％、40％、50％、60％、70％、75％、80％、85％、90％、95％，降低至1/2、1/3、1/4、1/5、1/10、1/20、1/50、1/100，或降低更多。

[0190] 必须注意的是，除非上下文另外明确规定，否则如说明书和随附权利要求中所用，单数形式“一个/种(a/an)”和“所述(the)”包括复数个指示物。

[0191] 重组核酸和载体

[0192] 在一些实施方案中，本公开设想包含一种或多种转基因的重组核酸插入物。转基因可以编码治疗性蛋白质、抗体、肽、自杀基因、凋亡基因或任何其他目标基因。在一些实施方案中，转基因编码启动受体。在一些实施方案中，转基因编码嵌合抗原受体。在一些实施方案中，插入物包含启动受体转基因和嵌合抗原受体转基因。

[0193] 插入物还可以包含自裂解肽。自裂解肽的实例包括但不限于自裂解病毒2A肽，例如猪捷申病毒-1(P2A)肽、明脉扁刺蛾病毒(T2A)肽、马鼻炎A病毒(E2A)肽或口蹄疫病毒(F2A)肽。自裂解2A肽允许从单个构建体表达多个基因产物。(参见，例如，Chang等人“Cleavage efficient 2A peptides for high level monoclonal antibody expression in CHO cells,”MAbs 7(2):403-412(2015))。

[0194] 插入物还可以包含WPRE元件。WPRE元件总体上描述于Higashimoto,T.等人,Gene Ther 14,1298-1304(2007)；以及Zufferey,R.等人,J Virol.1999年4月;73(4):2886-92.，两者据此以引用的方式并入本文。

[0195] 启动受体

[0196] 在本公开的某些方面，启动受体是基于Notch蛋白的合成回路受体。天然Notch受体与同源配体(诸如来自Delta蛋白家族的配体)的结合会导致膜内蛋白水解，从而使Notch蛋白的细胞内片段裂解。该细胞内片段是转录调节因子，仅在从Notch上裂解下来后才起作用。裂解可通过ADAM金属蛋白酶和γ分泌酶复合物的顺序蛋白水解而发生。该细胞内片段进入细胞核并激活细胞-细胞信号传导基因。与天然Notch蛋白相比，synNotch启动受体用能导致编码CAR的基因在从启动受体释放时激活的片段取代天然Notch细胞内片段。

[0197] Notch受体具有模块化结构域组织。Notch受体的胞外域由一系列负责配体结合的N末端表皮生长因子(EGF)样重复序列组成。在合成的Notch受体或启动受体中，Notch配体结合结构域被结合所选靶配体或抗原的配体结合结构域置换。EGF重复序列之后是三个LIN-12/Notch重复序列(LNR)模块，这些模块是Notch受体所特有的，被广泛报道参与防止受体过早激活。Notch1的异源二聚化(HD)结构域经弗林蛋白酶裂解而分开，使得其N末端部分终止细胞外亚基，而其C末端半部分构成跨膜亚基的起点。在细胞外区域之后，受体具有跨膜区段和细胞内结构域(ICD)，细胞内结构域包括转录调节因子。

[0198] 多种形式的启动受体可用于如本文所述的方法、细胞和核酸。设想用于本文的方法和细胞中的一类启动受体包含异源细胞外配体结合结构域、与包括NRR的Notch受体具有实质序列同一性的连接多肽、TMD和ICD。“Fn Notch”受体包含异源细胞外配体结合结构域、与Robo受体(诸如哺乳动物Robo1、Robo2、Robo3或Robo4)具有实质序列同一性的连接多肽，随后是1、2或3个纤连蛋白重复序列(“Fn”)、TMD和ICD。“Mini Notch”受体包含异源细胞外配体结合结构域、与Notch受体(缺乏NRR)具有实质序列同一性的连接多肽、TMD和ICD。“Minimal Tinker Notch”受体包含异源细胞外配体结合结构域、与Notch受体(例如，合成(GGS)n多肽序列)缺乏实质序列同一性的连接多肽、TMD和ICD。“铰链Notch”受体包含异源细胞外配体结合结构域、包含寡聚化结构域(即，促进与合成受体和/或现有宿主受体的二聚化、三聚化或更高级多聚化的结构域)的铰链序列、TMD和ICD。所有这些受体类别都是合成的、重组的，并且不天然存在。在一些实施方案中，本文公开的非天然存在的受体结合靶细胞表面展示的配体，其触发受体的蛋白水解裂解以及调节细胞中的定制转录程序的转录调节因子的释放。在一些实施方案中，启动受体不包括Notch受体的LIN-12-Notch重复序列

(LNR)和/或异源二聚化结构域(HD)。

[0199] 启动受体细胞外结构域

[0200] 启动受体包含细胞外结构域。在一些实施方案中,细胞外结构域包括受体的配体结合部分。在一些实施方案中,细胞外结构域包括结合至一种或多种靶抗原的抗原结合部分。在一些实施方案中,抗原结合部分包括抗体或其功能性抗原结合片段的一个或多个抗原结合决定簇。在一些实施方案中,抗原结合部分选自抗体、纳米抗体、双抗体、三抗体或微型抗体、F(ab')2片段、Fab片段、单链可变片段(scFv)和单结构域抗体(sdAb)或其功能性片段。在一些实施方案中,抗原结合部分包含scFv。抗原结合部分可以包括天然存在的氨基酸序列或者可经工程化、设计或修饰以提供所需的和/或改进的特性,例如增加的结合亲和力。"选择性地结合"抗原的抗体是以高亲和力结合抗原并且不显著结合其他无关抗原的抗原结合部分。在一些实施方案中,细胞外抗原结合结构域结合至生殖细胞型碱性磷酸酶(ALPG)。WO201061872中描述了另外的抗原。

[0201] 跨膜结构域

[0202] 如上文所述,本公开的嵌合多肽包括含有一个或多个配体诱导型蛋白水解裂解位点的TMD。

[0203] 一般来说,适用于本文公开的嵌合受体的TMD可以是1型跨膜受体的包括至少一个γ分泌酶裂解位点的任何跨膜结构域。γ分泌酶复合物及其底物蛋白(包括淀粉样前体蛋白(APP)和Notch)的结构和功能的详细描述可以例如在Zhang等人最近在Frontiers Cell Neurosci(2014)中发表的一篇综述中找到。来自1型跨膜受体的非限制性适合TMD包括来自CLSTN1、CLSTN2、APLP1、APLP2、LRP8、APP、BTC、TGBR3、SPN、CD44、CSF1R、CXCL16、CX3CL1、DCC、DLL1、DSG2、DAG1、CDH1、EPCAM、EPHA4、EPHB2、EFNB1、EFNB2、ErbB4、GHR、HLA-A和IFNAR2的TMD,其中TMD包括至少一个γ分泌酶裂解位点。适用于本文所述的组合物和方法的另外TMD包括但不限于来自1型跨膜受体IL1R1、IL1R2、IL6R、INSR、ERN1、ERN2、JAG2、KCNE1、KCNE2、KCNE3、KCNE4、KL、CHL1、PTPRF、SCN1B、SCN3B、NPR3、NGFR、PLXDC2、PAM、AGER、ROBO1、SORCS3、SORCS1、SORL1、SDC1、SDC2、SPN、TYR、TYRP1、DCT、YASN、FLT1、CDH5、PKHD1、NECTIN1、PCDHGC3、NRG1、LRP1B、CDH2、NRG2、PTPRK、SCN2B、Nradd和PTPRM的跨膜结构域。在一些实施方案中,本公开的嵌合多肽或Notch受体的TMD是源自钙同线蛋白(calsyntenin)家族成员(诸如alcadeinα和alcadeinγ)的TMD的TMD。在一些实施方案中,本发明的嵌合多肽或Notch受体的TMD是已知用于Notch受体的TMD。在一些实施方案中,本发明的嵌合多肽或Notch受体的TMD是源自不同Notch受体的TMD。例如,在基于人Notch1的Mini Notch中,Notch1 TMD可以用Notch2 TMD、Notch3 TMD、Notch4 TMD或来自非人动物(诸如斑马鱼、黑腹果蝇、非洲爪蟾或原鸡)的Notch TMD取代。

[0204] 在一些实施方案中,启动受体包含Notch裂解位点,诸如S2或S3。适用于本文公开的组合物和方法的另外蛋白水解裂解位点包括但不限于选自以下的MMP的金属蛋白酶裂解位点:胶原酶1、2和3(MMP-1、-8和-13)、明胶酶A和B(MMP-2和-9)、溶基质素1、2和3(MMP-3、-10和-11)、基质溶解素(MMP-7)和膜金属蛋白酶(MT1-MMP和MT2-MMP)。适合蛋白酶裂解位点的另一实例是纤溶酶原激活物裂解位点,例如尿激酶纤溶酶原激活物(uPA)或组织纤溶酶原激活物(tPA)裂解位点。适合蛋白酶裂解位点的另一实例是催乳素裂解位点。uPA和tPA的裂解序列的具体实例包括包含Yal-Gly-Arg的序列。可包含在蛋白水解可裂解接头中的蛋

白酶裂解位点的另一实例是烟草蚀刻病毒(TEV)蛋白酶裂解位点,例如Glu-Asn-Leu-Tyr-Thr-Gln-Ser,其中蛋白酶在谷氨酰胺与丝氨酸之间裂解。可包含在蛋白水解可裂解接头中的蛋白酶裂解位点的另一实例是肠激酶裂解位点,例如Asp-Asp-Asp-Asp-Lys,其中裂解发生在赖氨酸残基之后。可包含在蛋白水解可裂解接头中的蛋白酶裂解位点的另一实例是凝血酶裂解位点,例如Leu-Val-Pro-Arg。包含蛋白酶裂解位点的另外的适合接头包括可被以下蛋白酶裂解的序列:PreScission™蛋白酶(包含人鼻病毒3C蛋白酶和谷胱甘肽-S-转移酶的融合蛋白)、凝血酶、组织蛋白酶B、爱泼斯坦-巴尔病毒蛋白酶、MMP-3(溶基质素)、MMP-7(基质溶解素)、MMP-9;类嗜热菌蛋白酶MMP、基质金属蛋白酶2(MMP-2)、组织蛋白酶L;组织蛋白酶D、基质金属蛋白酶1(MMP-1)、尿激酶型纤溶酶原激活剂、膜1型基质金属蛋白酶(MT-MMP)、溶基质素3(或MMP-11)、嗜热溶素、成纤维细胞胶原酶和溶基质素1、基质金属蛋白酶13(胶原酶3)、组织型纤溶酶原激活剂(tPA)、人前列腺特异性抗原、激肽释放酶(hK3)、中性粒细胞弹性蛋白酶和钙蛋白酶(钙激活中性蛋白酶)。非表达受体的宿主细胞原生的蛋白酶(例如,TEV)可用作进一步的调控机制,其中受体的激活减少,直至蛋白酶被表达或以其他方式提供。此外,蛋白酶可能与肿瘤相关或与疾病相关(表达程度显著高于正常组织中的程度),并充当独立的调控机制。例如,一些基质金属蛋白酶在某些癌症类型中高度表达。

[0205] 在一些实施方案中,TMD内的氨基酸取代包括TMD的“GV”基序内的一个或多个取代。在一些实施方案中,此类取代中的至少一种包括对丙氨酸的取代。另外的序列和取代描述于WO2021061872中,其据此以引用的方式整体并入本文。

[0206] 细胞内结构域

[0207] 在一些实施方案中,启动受体包含一个或多个来自或源自转录调节因子的细胞内结构域。转录调节因子激活或抑制同源启动子的转录。转录激活因子通常在转录启动子附近与其结合并募集RNA聚合酶以直接启始转录。转录抑制因子与转录启动子结合并在空间上阻碍RNA聚合酶的转录起始。其他转录调节因子根据其结合位置和细胞条件充当激活因子或抑制因子。因此,如本文所用,“转录激活结构域”是指与转录控制元件和/或转录调节蛋白(即,转录因子、RNA聚合酶等)相互作用以增加和/或激活一个或多个基因的转录的转录因子的结构域。转录激活结构域的非限制性实例包括:单纯疱疹病毒VP16激活结构域、VP64(其为VP16的四聚体衍生物)、HIV TAT、NFkB p65激活结构域、p53激活结构域1和2、CREB(cAMP应答元件结合蛋白)激活结构域、E2A激活结构域、NFAT(活化T细胞核因子)激活结构域、酵母Gal4、酵母GCN4、酵母HAP1、MLL、RTG3、GLN3、OAF1、PIP2、PDR1、PDR3、PHO4、LEU3糖皮质激素受体转录激活域、B细胞POU同源结构域蛋白Oct2、植物Ap2或本领域技术人员或普通技术人员已知的任何其他结构域。在一些实施方案中,转录调节因子选自Gal4-VP16、Gal4-VP64、tetR-VP64、ZFHD1-YP64、Gal4-KRAB和HAP1-VP16。在一些实施方案中,转录调节因子是Gal4-VP64。转录激活结构域可以包含野生型或天然存在的序列,或者它可以是原始转录激活结构域的修饰、突变或衍生型式,具有所需的增加和/或激活一个或多个基因的转录的能力。在一些实施方案中,转录调节因子还可包括核定位信号。

[0208] DNA结合结构域

[0209] 在本文所述方面的一些实施方案中,合成蛋白质包含一个或多个细胞内“DNA结合结构域”(或“DB结构域”)。此类“DNA结合结构域”是指结合特定DNA序列元件的序列特异性DNA结合结构域。因此,如本文所用,“序列特异性DNA结合结构域”是指具有选择性地结合具

有特定预定序列的DNA的能力的蛋白质结构域部分。序列特异性DNA结合结构域可以包含野生型或天然存在的序列，或者它可以是原始结构域的具有结合所需序列的所需能力的修饰、突变或衍生型式。在一些实施方案中，序列特异性DNA结合结构域经工程化以结合所需序列。可用于本文所述的合成蛋白质中的具有序列特异性DNA结合结构域的蛋白质的非限制性实例包括HNF1a、Gal4、GCN4、反向四环素受体、THY1、SYN1、NSE/RU5′、AGRP、CALB2、CAMK2A、CCK、CHAT、DLX6A、EMX1、锌指蛋白或其结构域、CRISPR/Cas蛋白，诸如Cas9、Cas3、Cas4、Cas5、Cas5e(或CasD)、Cash、Cas6e、Cas6f、Cas7、Cas8a1、Cas8a2、Cas8b、Cas8c、Cas10、Cas10d、CasF、CasG、CasH、Csy1、Csy2、Csy3、Cse1(或CasA)、Cse2(或CasB)、Cse3(或CasE)、Cse4(或CasC)、Csc1、Csc2、Csa5、Csn2、Csm2、Csm3、Csm4、Csm5、Csm6、Cmr1、Cmr3、Cmr4、Cmr5、Cmr6、Csb1、Csb2、Csb3、Csx17、Csx14、Csx10、Csx16、CsaX、Csx3、Csz1、Csx15、Csf1、Csf2、Csf3、Csf4，和Cu196，以及TALES。

[0210] 在使用CRISPR/Cas样蛋白的那些实施方案中，CRISPR/Cas样蛋白可以是野生型CRISPR/Cas蛋白、修饰的CRISPR/Cas蛋白，或野生型或修饰的CRISPR/Cas蛋白的片段。CRISPR/Cas样蛋白可经修饰以增加核酸结合亲和力和/或特异性，改变酶活性，并且/或者改变蛋白质的另一特性。例如，CRISPR/Cas样蛋白的核酸酶(即，DNase、RNase)结构域可以被修饰、缺失或失活。或者，CRISPR/Cas样蛋白可被截短以去除对本文所述系统的功能而言不是必需的结构域。例如，用作DNA结合蛋白或其结构域的CRISPR酶可以相对于相应的野生型酶发生突变，使得突变的CRISPR或其结构域缺乏裂解含有DNA结合结构域靶位点的核酸序列的能力。例如，D10A突变可以与H840A、N854A或N863A突变中的一个或多个组合以产生基本上缺乏所有DNA裂解活性的Cas9酶。

[0211] 近膜结构域

[0212] ECD和TMD，或TMD和ICD，可以用连接多肽诸如近膜结构域彼此连接。“SynNotch”受体包含异源细胞外配体结合结构域、与Notch受体JMD(包括NRR)具有实质序列同一性的连接多肽、TMD和ICD。“Fn Notch”受体包含异源细胞外配体结合结构域、与Robo受体(诸如哺乳动物Robo1、Robo2、Robo3或Robo4)具有实质序列同一性的连接多肽，随后是1、2或3个纤连蛋白重复序列(“Fn”)、TMD和ICD。“Mini Notch”受体包含异源细胞外配体结合结构域、与Notch受体JMD(但缺乏NRR(LIN-12-Notch重复序列(LNR)模块))具有实质序列同一性的连接多肽、TMD和ICD。“Minimal Linker Notch”受体包含异源细胞外配体结合结构域、与Notch受体(例如但不限于，具有合成$(GGS)_n$多肽序列)缺乏实质序列同一性的连接多肽、TMD和ICD。“Hinge Notch”受体包含异源细胞外配体结合结构域、包含寡聚化结构域(即，促进与合成受体和/或现有宿主受体的二聚化、三聚化或更高级多聚化的结构域)的铰链序列、TMD和ICD。

[0213] 在一些实施方案中，启动受体包含介于细胞外结构域与跨膜结构域之间的近膜结构域(JMD)肽。在一些实施方案中，启动受体包含介于跨膜结构域与细胞内结构域之间的近膜结构域(JMD)肽。在一些实施方案中，JMD肽包含LWF基序。LWF基序在受体构建体中的用途描述于美国专利第10,858,443号中，该专利据此以引用的方式整体并入本文。在一些实施方案中，JMD肽与Notch1、Notch2、Notch3和/或Notch4的JMD具有实质序列同一性。在一些实施方案中，JMD肽与Notch1、Notch2、Notch3和/或Notch4 JMD具有实质序列同一性，但不包括Notch受体的LIN-12-Notch重复序列(LNR)和/或异源二聚化结构域(HD)。在一些实施方

案中，JMD肽与Notch1、Notch2、Notch3和/或Notch4 JMD不具有实质序列同一性。在一些实施方案中，JMD肽包括促进受体的二聚体、三聚体或更高级组合的形成的寡聚化结构域。此类JMD肽描述于WO2021061872中，其据此以引用的方式整体并入本文。

[0214] 在Mini Notch受体中，连接多肽源自NRR和HD结构域缺失后的Notch JMD序列。Notch JMD序列可以是来自Notch1、Notch2、Notch3或Notch4的序列，并且可以源自非人同系物，诸如来自果蝇属、原鸡属、鱼丹属等的那些。剩余的Notch序列的4至50个氨基酸残基可用作多肽接头。在一些实施方案中，改变接头多肽序列的长度和氨基酸组成以改变ECD和TMD相对于彼此的方向和/或接近度，从而实现嵌合多肽的所需活性，诸如诱导配体或不存在配体时的信号转导水平。

[0215] 在Minimal Linker Notch受体中，连接多肽与Notch JMD序列(包括来自Notch1、Notch2、Notch3或Notch4的Notch JMD序列，或其非人同系物)不具有实质序列同一性。4至50个氨基酸残基可用作多肽接头。在一些实施方案中，改变接头多肽序列的长度和氨基酸组成以改变ECD和TMD相对于彼此的方向和/或接近度，从而实现本公开的嵌合多肽的所需活性。Minimal Linker序列可以被设计为包括或省略蛋白酶裂解位点，并且可以包括或省略糖基化位点或用于其他类型的翻译后修饰的位点。在一些实施方案中，Minimal Linker接头不包含蛋白酶裂解位点或糖基化位点。

[0216] 在一些实施方案中，启动受体还包含铰链。可用于启动受体中的铰链接头可以包括含有一个或多个通过分子间二硫键键合促进嵌合多肽的寡聚体形成的多肽基序的寡聚化结构域(例如，铰链结构域)。在这些情形下，在本文公开的嵌合受体内，铰链结构域一般包括设置于ECD与TMD之间的柔性多肽连接区。因此，铰链结构域提供了ECD与TMD之间的柔性，并且还提供了供两个或更多个嵌合多肽单体之间的分子间二硫键键合以形成寡聚复合物的位点。在一些实施方案中，铰链结构域包括促进本文公开的嵌合多肽的二聚体形成的基序。在一些实施方案中，铰链结构域包括促进本文公开的嵌合多肽的三聚体形成的基序(例如，源自OX40的铰链结构域)。适用于本公开的组合物和方法的铰链多肽序列可以是天然存在的铰链多肽序列(例如，来自天然存在的免疫球蛋白的那些序列)或者可以经工程化、设计或修饰以提供所需的和/或改进的特性，例如，调节转录。适合铰链多肽序列包括但不限于源自IgA、IgD和IgG亚类的那些，诸如IgG1铰链结构域、IgG2铰链结构域、IgG3铰链结构域和IgG4铰链结构域，或它们的功能性变体。在一些实施方案中，铰链多肽序列含有一个或多个CXXC基序。在一些实施方案中，铰链多肽序列含有一个或多个CPPC基序。

[0217] 铰链多肽序列还可源自CD8α铰链结构域、CD28铰链结构域、CD152铰链结构域、PD-1铰链结构域、CTLA4铰链结构域、OX40铰链结构域，以及它们的功能性变体。在一些实施方案中，铰链结构域包括源自CD8α铰链结构域或其功能性变体的铰链多肽序列。在一些实施方案中，铰链结构域包括源自CD28铰链结构域或其功能性变体的铰链多肽序列。在一些实施方案中，铰链结构域包括源自OX40铰链结构域或其功能性变体的铰链多肽序列。在一些实施方案中，铰链结构域包括源自IgG4铰链结构域或其功能性变体的铰链多肽序列。

[0218] Fn Notch连接多肽源自Robo1 JMD，其含有纤连蛋白重复序列(Fn)结构域，在Fn重复序列与TMD之间有短多肽序列。Fn Notch连接多肽不含Notch负调控区(NRR)或Notch HD结构域。Fn连接多肽可以含有1、2、3、4或5个Fn重复序列。在一些实施方案中，嵌合受体包含具有约1至约5个Fn重复序列、约1至约3个Fn重复序列或约2至约3个Fn重复序列的Fn连接多

肽。Fn重复序列与TMD之间的短多肽序列可以为约2至约30个氨基酸残基。在一些实施方案中，短多肽序列可以介于任何序列的约5与约20个氨基酸之间。在一些实施方案中，短多肽序列可以介于任何序列的约5与约20个天然存在的氨基酸之间。在一些实施方案中，短多肽序列可以介于任何序列的约5与约20个氨基酸之间但不具有超过一个脯氨酸。在一些实施方案中，短多肽序列可以介于约5与约20个氨基酸之间，并且约50％或更多的氨基酸是甘氨酸。在一些实施方案中，短多肽序列可以介于约5与约20个氨基酸之间，其中氨基酸选自甘氨酸、丝氨酸、苏氨酸和丙氨酸。在一些实施方案中，可以改变Fn连接多肽序列的长度和氨基酸组成以改变ECD和TMD相对于彼此的方向和/或接近度，从而实现本公开的嵌合多肽的所需活性。

[0219] 终止转移序列

[0220] 在一些实施方案中，启动受体还包含在跨膜结构域与细胞内结构域之间的终止转移序列(STS)。STS包含带电荷的疏脂序列。不受任何理论的束缚，STS用作膜锚，并且被认为阻止细胞内结构域进入质膜。STS结构域在启动受体中的用途描述于WO2021061872中，其据此以引用的方式整体并入本文。非限制性示例性STS序列包括APLP1、APLP2、APP、TGBR3、CSF1R、CXCL16、CX3CL1、DAG1、DCC、DNER、DSG2、CDH1、GHR、HLA-A、IFNAR2、IGF1R、IL1R1、ERN2、KCNE1、KCNE2、CHL1、LRP1、LRP2、LRP18、PTPRF、SCN1B、SCN3B、NPR3、NGFR、PLXDC2、PAM、AGER、ROBO1、SORCS3、SORCS1、SORL1、SDC1、SDC2、SPN、TYR、TYRP1、DCT、VASN、FLT1、CDH5、PKTFD1、NECTIN1、KL、IL6R、EFNB1、CD44、CLSTN1、LRP8、PCDHGC3、NRG1、LRP1B、JAG2、EFNB2、DLL1、CLSTN2、EPCAM、ErbB4、KCNE3、CDH2、NRG2、PTPRK、BTC、EPHA4、IL1R2、KCNE4、SCN2B、Nradd、PTPRM、Notch1、Notch2、Notch3和Notch4STS序列。在一些实施方案中，STS与跨膜域异源。在一些实施方案中，STS与跨膜域同源。STS序列描述于WO2021061872中，其据此以引用的方式整体并入本文。

[0221] 嵌合抗原受体

[0222] 重组CAR可以是人CAR，其包含完全人序列，例如天然人序列。

[0223] 在一些实施方案中，诸如CAR的重组受体，诸如其抗体部分，还包括间隔区，间隔区可以是或包括免疫球蛋白恒定区或其变体或修饰型式的至少一部分，诸如铰链区(例如，IgG4铰链区)，和/或CH1/CL和/或Fc区。在一些实施方案中，恒定区或部分属于人IgG，诸如IgG4或IgG1。在一些方面，恒定区的部分用作抗原识别组分(例如，scFv)与跨膜结构域之间的间隔区。与不存在间隔区的情况相比，间隔区的长度可增加细胞在抗原结合后的应答性。在一些实例中，间隔区的长度为或约为12个氨基酸，或者长度不超过12个氨基酸。示例性间隔区包括具有至少约10至229个氨基酸、约10至200个氨基酸、约10至175个氨基酸、约10至150个氨基酸、约10至125个氨基酸、约10至100个氨基酸、约10至75个氨基酸、约10至50个氨基酸、约10至40个氨基酸、约10至30个氨基酸、约10至20个氨基酸或约10至15个氨基酸(并且包括任何所列范围的端点之间的任何整数)的间隔区。在一些实施方案中，间隔区具有约12个或更少的氨基酸、约119个或更少的氨基酸，或约229个或更少的氨基酸。示例性间隔区包括单独的IgG4铰链、连接至CH2和CH3结构域的IgG4铰链，或连接至CH3结构域的IgG4铰链。示例性间隔区包括但不限于Hudecek等人(2013)Clin.Cancer Res.,19:3153或国际专利申请公布第WO2014031687号中描述的间隔区。在一些实施方案中，恒定区或部分属于IgD。

[0224] 受体(诸如CAR)的抗原识别结构域可以连接至一种或多种细胞内信号传导组分，诸如在CAR的情况下模拟通过抗原受体复合物(诸如TCR复合物)激活的信号传导组分，和/或经由另一细胞表面受体进行信号传导。因此，在一些实施方案中，细胞外结合组分(例如，配体结合或抗原结合结构域)连接至一个或多个跨膜和细胞内信号传导结构域。在一些实施方案中，跨膜结构域与细胞外结构域融合。在一个实施方案中，使用天然与受体(例如，CAR)中的结构域之一相缔合的跨膜结构域。在一些情况下，通过氨基酸取代选择或修饰跨膜结构域，以避免此类结构域与相同或不同表面膜蛋白的跨膜结构域结合，从而最大限度地减少与受体复合物的其他成员的相互作用。

[0225] 在一些实施方案中，跨膜结构域源自天然来源或合成来源。在来源是天然的情况下，在一些方面结构域源自任何膜结合蛋白或跨膜蛋白。跨膜区包括源自(即，至少包含以下的跨膜区)T细胞受体的α、β或δ链，CD28，CD3ε，CD45，CD4，CD5，CDS，CD9，CD 16，CD22，CD33，CD37，CD64，CD80，CD86，CD 134，CD137和/或CD 154的跨膜区。或者，在一些实施方案中，跨膜结构域是合成的。在一些方面，合成的跨膜结构域主要包含疏水性残基，诸如亮氨酸和缬氨酸。在一些方面，苯丙氨酸、色氨酸和缬氨酸的三联体将存在于合成的跨膜结构域的每一端。在一些实施方案中，连接是通过接头、间隔区和/或跨膜结构域完成的。

[0226] 在细胞内信号传导结构域中，有通过天然抗原受体模拟或近似信号、通过这种受体与共刺激受体的组合模拟或近似信号和/或通过单独的共刺激受体模拟或近似信号的那些。在一些实施方案中，存在短的寡肽或多肽接头，例如长度为2与10个氨基酸之间的接头，诸如含有甘氨酸和丝氨酸的接头，例如甘氨酸-丝氨酸双联体，并且所述接头在受体的跨膜结构域与细胞质信号传导结构域之间形成连接。

[0227] 受体，例如CAR，可以包括至少一种细胞内信号传导组分。在一些实施方案中，受体包括TCR复合物的细胞内组分，诸如介导T细胞激活和细胞毒性的TCR CD3链，例如CD3δ链。因此，在一些方面，细胞外结构域连接至一个或多个细胞信号传导模块。在一些实施方案中，细胞信号传导模块包括CD3跨膜结构域、CD3细胞内信号传导结构域和/或其他CD跨膜结构域。在一些实施方案中，受体，例如CAR，还包括一种或多种另外分子(诸如Fc受体γ、CD8、CD4、CD25或CD16)的一部分。例如，在一些方面，CAR包括在CD3δ或Fc受体γ与CD8、CD4、CD25或CD16之间的嵌合分子。

[0228] 在一些实施方案中，在连结CAR后，受体的细胞质结构域或细胞内信号传导结构域激活免疫细胞(例如，经工程化以表达受体的T细胞)的至少一种正常效应功能或应答。例如，在一些情形下，受体诱导T细胞的功能，诸如细胞溶解活性或T辅助细胞活性，诸如分泌细胞因子或其他因子。在一些实施方案中，例如，如果抗原受体组分或共刺激分子的细胞内信号传导结构域的截短部分转导效应功能信号，则使用其替代完整的免疫刺激链。在一些实施方案中，一个或多个细胞内信号传导结构域包括T细胞受体(TCR)的细胞质序列，并且在一些方面还包括在自然环境中与所述受体协同作用以在抗原受体接合后启始信号转导的共受体的细胞质序列，和/或此类分子的任何衍生物或变体，和/或具有相同功能能力的任何合成序列。

[0229] 在天然TCR的情形下，完全激活一般不仅需要通过TCR进行信号传导，还需要共刺激信号。因此，在一些实施方案中，为了促进完全激活，受体中还包含用于产生次级或共刺激信号的组分。在其他实施方案中，受体不包括用于产生共刺激信号的组分。在一些方面，

另外的受体在同一细胞中表达并提供用于产生次级或共刺激信号的组分。

[0230] 在一些方面,T细胞激活被描述为由两类细胞质信号传导序列介导:通过TCR启始抗原依赖性初级激活的那些(初级细胞质信号传导序列),以及以抗原非依赖性方式起作用以提供次级或共刺激信号的那些(次级细胞质信号传导序列)。在一些方面,受体包括所述信号传导组分中的一者或两者。

[0231] 在一些方面,受体包括调控TCR复合物的初级激活的初级细胞质信号传导序列。以刺激方式起作用的初级细胞质信号传导序列可含有信号传导基序,该基序称为基于免疫受体酪氨酸的激活基序或ITAM。含有初级细胞质信号传导序列的ITAM的实例包括源自TCR或CD3δ、FcRγ、FcRβ、CD3γ、CD3δ、CD3ε、CDS、CD22、CD79a、CD79b和CD66d的那些。在一些实施方案中,CAR中的细胞质信号传导分子含有细胞质信号传导结构域、其部分或源自CD3δ的序列。

[0232] 在一些实施方案中,受体包括共刺激受体(诸如CD28、4-1BB、OX40、DAP10和ICOS)的信号传导结构域和/或跨膜部分。在一些方面,同一受体包括激活性组分和共刺激组分两者。

[0233] 在某些实施方案中,细胞内信号传导结构域包含连接至CD3(例如,CD3-δ)细胞内结构域的CD28跨膜和信号传导结构域。在一些实施方案中,细胞内信号传导结构域包含连接至CD3δ细胞内结构域的嵌合的CD28和CD137(4-1BB,TNFRSF9)共刺激结构域。

[0234] 在一些实施方案中,受体在细胞质部分包含一个或多个,例如两个或更多个共刺激结构域和激活结构域,例如初级激活结构域。示例性受体包括CD3-δ、CD28和4-1BB的细胞内组分。

[0235] 在一些实施方案中,CAR或其他抗原受体还包括标志物,诸如细胞表面标志物,其可用于确认细胞的转导或工程化以表达受体,诸如细胞表面受体的截短型式,诸如截短的EGFR(tEGFR)。在一些方面,标志物包括CD34、神经生长因子受体(NGFR)或表皮生长因子受体(例如,tEGFR)的全部或部分(例如,截短形式)。在一些实施方案中,编码标志物的核酸可操作地连接至编码接头序列(诸如可裂解的接头序列或核糖体跳跃序列例如T2A)的多核苷酸。参见WO2014031687。在一些实施方案中,引入编码由T2A核糖体开关分隔开的CAR和EGFRt的构建体可以表达来自同一构建体的两种蛋白质,使得EGFRt可以用作检测表达这种构建体的细胞的标志物。在一些实施方案中,标志物和任选的接头序列可以是如公布的专利公布第WO2014031687号中所公开的任何标志物和接头序列。例如,标志物可以是截短的EGFR(tEGFR),其任选地连接至接头序列,诸如T2A核糖体跳跃序列。

[0236] 在一些实施方案中,标志物是分子,例如细胞表面蛋白,非天然存在于T细胞上或非天然存在于T细胞表面上,或其部分。

[0237] 在一些实施方案中,分子是非自身分子,例如非自身蛋白质,即不被细胞将过继转移到其中的宿主的免疫系统识别为“自身”的分子。

[0238] 在一些实施方案中,标志物不提供治疗功能和/或不产生别的作用,除了用作基因工程化的标志物,例如,用于选择成功工程化的细胞。在其他实施方案中,标志物可以是治疗性分子或以其他方式发挥某些所需作用的分子,诸如在体内遇到的细胞的配体,诸如增强和/或阻抑细胞在过继转移和遇到配体后的应答的共刺激或免疫检查点分子。

[0239] CAR可包含一个或修饰的合成氨基酸来代替一个或多个天然存在的氨基酸。示例

性的修饰的氨基酸包括但不限于氨基环己烷羧酸、正亮氨酸、α-氨基正癸酸、高丝氨酸、S-乙酰氨基甲基半胱氨酸、反式-3-羟脯氨酸和反式-4-羟脯氨酸、4-氨基苯丙氨酸、4-硝基苯丙氨酸、4-氯苯丙氨酸、4-羧基苯丙氨酸、(3-苯基丝氨酸(3-羟基苯丙氨酸、苯基甘氨酸、α-萘基丙氨酸、环己基丙氨酸、环己基甘氨酸、二氢吲哚-2-羧酸、1,2,3,4-四氢异喹啉-3-羧酸、氨基丙二酸、氨基丙二酸单酰胺、N'-苄基-N'-甲基-赖氨酸、N',N'-二苄基-赖氨酸、6-羟赖氨酸、鸟氨酸、α-氨基环戊烷羧酸、α-氨基环己烷羧酸、α-氨基环庚烷羧酸、α-(2-氨基-2-降冰片烷)-羧酸、α,γ-二氨基丁酸、α,γ-二氨基丙酸、高苯丙氨酸和α-叔丁基甘氨酸。

[0240] 在一些情况下,CAR被称为第一代、第二代和/或第三代CAR。在一些方面,第一代CAR是在抗原结合后仅提供CD3链诱导的信号的CAR;在一些方面,第二代CAR是提供这种信号和共刺激信号的CAR,诸如包括来自共刺激受体(诸如CD28或CD137)的细胞内信号传导结构域的CAR;在一些方面,第三代CAR在某些方面是包括不同共刺激受体的多个共刺激结构域的CAR。

[0241] 在一些实施方案中,嵌合抗原受体包括含有本文所述的抗体或片段的细胞外部分。在一些方面,嵌合抗原受体包括含有细胞内信号传导结构域和本文所述的抗体或片段的细胞外部分。在一些实施方案中,抗体或片段包括scFv或单结构域VH抗体并且细胞内结构域含有ITAM。在一些方面,细胞内信号传导结构域包括CD3-δ(CD3)链的δ链的信号传导结构域。在一些实施方案中,嵌合抗原受体包括连接细胞外结构域和细胞内信号传导结构域的跨膜结构域。

[0242] 在一些方面,跨膜结构域含有CD28的跨膜部分。细胞外结构域和跨膜可以直接或间接连接。在一些实施方案中,细胞外结构域和跨膜由间隔区连接,诸如本文所述的任何间隔区。在一些实施方案中,嵌合抗原受体如在跨膜结构域与细胞内信号传导结构域之间含有T细胞共刺激分子的细胞内结构域。在一些方面,T细胞共刺激分子是CD28或41BB。

[0243] 在一些实施方案中,CAR含有抗体,例如抗体片段;为或含有CD28或其功能性变体的跨膜部分的跨膜结构域;以及含有CD28或其功能性变体的信号传导部分和CD3δ或其功能性变体的信号传导部分的细胞内信号传导结构域。在一些实施方案中,CAR含有抗体,例如抗体片段;为或含有CD28或其功能性变体的跨膜部分的跨膜结构域;以及含有4-1BB或其功能性变体的信号传导部分和CD3δ或其功能性变体的信号传导部分的细胞内信号传导结构域。在一些此类实施方案中,受体还包括含有Ig分子(诸如人Ig分子)的部分(诸如Ig铰链,例如IgG4铰链)的间隔区,诸如仅铰链间隔区。

[0244] 在一些实施方案中,受体(例如,CAR)的跨膜结构域是人CD28或其变体的跨膜结构域,例如人CD28(登录号:P10747.1)的27个氨基酸的跨膜结构域。

[0245] 在一些实施方案中,嵌合抗原受体含有T细胞共刺激分子的细胞内结构域。在一些方面,T细胞共刺激分子是CD28或4-1BB。

[0246] 在一些实施方案中,细胞内信号传导结构域包含人CD28或其功能性变体或部分的细胞内共刺激信号传导结构域,诸如其41个氨基酸的结构域和/或在天然CD28蛋白的位置186-187具有LL至GG取代的这种结构域。在一些实施方案中,细胞内结构域包含41BB或其功能性变体或部分的细胞内共刺激信号传导结构域,诸如人4-1BB(登录号Q07011.1)或其功能性变体或部分的42个氨基酸的细胞质结构域。

[0247] 在一些实施方案中,细胞内信号传导结构域包含人CD3δ刺激性信号传导结构域或

其功能性变体，诸如人CD3δ(登录号:P20963.2)的同种型3的112AA细胞质结构域或如美国专利第7,446,190号或美国专利第8,911,993号中描述的CD3δ信号传导结构域。

[0248] 在一些方面，间隔区仅含有IgG的铰链区，诸如仅含有IgG4或IgG1的铰链。在其他实施方案中，间隔区是连接至CH2和/或CH3结构域的Ig铰链，例如，和IgG4铰链。在一些实施方案中，间隔区是连接至CH2和CH3结构域的Ig铰链，例如，IgG4铰链。在一些实施方案中，间隔区是仅连接至CH3结构域的Ig铰链，例如，IgG4铰链。在一些实施方案中，间隔区是或包含富含甘氨酸-丝氨酸的序列或其他柔性接头，诸如已知的柔性接头。

[0249] 例如，在一些实施方案中，CAR包括抗体或其片段，包括本文所述的单链抗体(sdAb，例如仅含有VH区)和scFv；间隔区，诸如含Ig铰链的任何间隔区；CD28跨膜结构域；CD28细胞内信号传导结构域；和CD3δ信号传导结构域。在一些实施方案中，CAR包括抗体或片段，包括本文所述的sdAb和scFv；间隔区，诸如含Ig铰链的任何间隔区；CD28跨膜结构域；CD28细胞内信号传导结构域；和CD3δ信号传导结构域。

[0250] 嵌合抗原受体(CAR)T细胞是经过基因工程化以产生用于免疫疗法的人工T细胞受体的T细胞。嵌合抗原受体是经过工程化以赋予T细胞靶向特定蛋白质的能力的受体蛋白。通过并入例如CAR对淋巴细胞(例如T细胞)进行基因修饰，并将工程化细胞施用至受试者是“过继细胞疗法”的一个实例。如本文所用，术语“过继细胞疗法”是指用于输注自体或同种异体淋巴细胞(称为T细胞或B细胞)的基于细胞的免疫疗法。在这种CAR治疗方法中，在输注前对细胞进行离体扩增和培养，并进行基因修饰。

[0251] CAR的表达允许工程化T细胞靶向并结合特定蛋白质，例如肿瘤抗原。在CAR疗法中，T细胞是从受试者采集的，这些细胞可以是来自受试者自己的血液或来自将不会接受CAR疗法的供体的自体T细胞。分离后，利用CAR对T细胞进行基因修饰，离体扩增，并通过例如输注施用至受试者(即患者)。

[0252] 可使用例如位点特异性技术将CAR引入T细胞。通过转基因(例如CAR)的位点特异性整合，转基因可以靶向安全性港基因座或TRAC。用于整合到安全港基因座中的位点特异性技术的实例包括但不限于使用核酸酶的同源依赖性工程化和使用Cas9的同源非依赖性靶向插入。

[0253] 工程化的CAR T细胞可应用于免疫肿瘤学。例如，可以选择CAR来靶向特定的肿瘤抗原。可以使用CAR T细胞有效靶向的癌症的实例是血癌。在一些实施方案中，CAR T细胞疗法可用于治疗实体瘤。

[0254] 重组细胞

[0255] 本文还提供了一种重组原代免疫细胞，所述重组原代免疫细胞包含至少一个以病毒方式插入细胞的基因组的靶区中的DNA模板，其中DNA模板的大小为大于或等于约4.5或5千碱基对(kb)。

[0256] 如本公开所述的在靶基因座或安全港位点处包含插入物的细胞可称为工程化细胞。在一些实施方案中，免疫细胞是可以产生多能免疫细胞的任何细胞。在一些实施方案中，免疫细胞可以是诱导型多能干细胞(iPSC)或人多能干细胞(HSPC)。在一些实施方案中，免疫细胞包括原代造血细胞或原代造血干细胞。在一些实施方案中，工程化细胞是干细胞、人细胞、原代细胞、造血细胞、适应性免疫细胞、先天免疫细胞、自然杀伤(NK)细胞、T细胞、CD8+细胞、CD4+细胞或T细胞祖细胞。在一些实施方案中，免疫细胞是T细胞。在一些实施方

案中，T细胞是调节性T细胞、效应T细胞或幼稚T细胞。在一些实施方案中，T细胞是$CD8^+$T细胞。在一些实施方案中，T细胞是$CD4^+$T细胞。在一些实施方案中，T细胞是$CD4^+CD8^+$T细胞。

[0257] 在一些实施方案中，工程化细胞是干细胞、人细胞、原代细胞、造血细胞、适应性免疫细胞、先天免疫细胞、T细胞或T细胞祖细胞。本公开中设想的免疫细胞的非限制性实例包括T细胞、B细胞、自然杀伤(NK)细胞、NKT/iNKT细胞、巨噬细胞、骨髓细胞和树突细胞。本公开中设想的干细胞的非限制性实例包括多能干细胞(PSC)、胚胎干细胞(ESC)、诱导型多能干细胞(iPSC)、通过核移植获得的胚胎来源的胚胎干细胞(ntES;核移植ES)、雄性生殖系干细胞(GS细胞)、胚胎生殖细胞(EG细胞)、造血干细胞/祖细胞干细胞(HSPC)、体干细胞(成体干细胞)、成血管细胞、神经干细胞、间充质干细胞和其他细胞(包括骨细胞、软骨细胞、肌细胞、心肌细胞、神经元、肌腱细胞、脂肪细胞、胰腺细胞、肝细胞、肾细胞、滤泡细胞等)的干细胞。在一些实施方案中，工程化细胞是T细胞、NK细胞、iPSC和HSPC。在一些实施方案中，本公开中使用的工程化细胞是体外生长的人细胞系(例如，故意永生化的细胞系、癌细胞系等)。

[0258] 本文还提供了包含多个原代免疫细胞的细胞群体。在一些实施方案中，细胞的至少20％、30％、40％、50％、60％、70％、80％、90％、95％、99％或更多的基因组包含异源DNA模板的靶向插入，其中DNA模板的大小至少约为5kb。

[0259] 编辑细胞的方法

[0260] 如本文所用，术语“基因编辑”或“基因组编辑”是指使用人工操纵的核酸酶或“分子剪刀”从基因组中插入、置换或移除DNA的一种基因操纵。它是阐明序列特异性基因或蛋白质的功能和作用或改变细胞行为(例如用于治疗目的)的有用工具。

[0261] 目前可用的基因组编辑工具包括锌指核酸酶(ZFN)和转录激活因子样效应核酸酶(TALEN)，以在安全港基因座(例如腺相关病毒整合位点1(AAVS1)安全港基因座)处插入基因。利用phiC31整合酶和Bxb1整合酶的DICE(双整合酶盒交换)系统是一种用于靶标整合的工具。此外，成簇规律间隔的短回文重复序列/Cas9(CRISPR/Cas9)技术可用于靶向基因插入。

[0262] 位点特异性基因编辑方法可以包括同源依赖性机制或同源非依赖性机制。

[0263] 在本文所述的方法中设想本领域已知的用于靶向插入基因序列的所有方法以在基因靶标或安全港基因座处插入构建体。

[0264] 本文提供了在不存在病毒载体的情况下将长度大于约5千碱基的核苷酸序列插入细胞基因组中的方法。在一些实施方案中，在不存在病毒载体的情况下，可以将长度大于约5千碱基的核苷酸序列插入原代免疫细胞的基因组中。

[0265] 将大核酸(例如，大小为大于5千碱基的核酸)整合到细胞中可能受到低整合效率、脱靶效应和/或细胞活力损失的限制。本文描述了用于实现将核苷酸序列(例如，大小为大于约5千碱基的核苷酸序列)整合到细胞的基因组中的方法和组合物。在一些方法中，整合效率增加，脱靶效应减少并且/或者细胞活力损失减少。

[0266] 图1示出了编码启动受体和CAR的示例性质粒，以及将其并入免疫细胞的基因组的示意图。利用核酸酶，诸如CRISPR相关系统(Cas)将质粒引入免疫细胞中。核酸酶可以按核糖核蛋白形式与靶向免疫细胞基因组上的特定位点的指导RNA(gRNA)一起引入。核酸酶在该特定位点处切割基因组DNA。该特定位点可以是编码内源免疫细胞受体的基因组的一部分。因此，在该位点切割基因组将导致免疫细胞不再表达内源免疫细胞受体。

[0267] 如图1所示，质粒可包括与免疫细胞基因组上的特定位点处的序列互补的5’和3’同源定向修复臂。互补序列位于被核酸酶切割的位点的两侧，这使得质粒可以并入免疫细胞基因组上的特定插入位点处。一旦并入质粒，细胞将表达启动受体。然而，正如所解释的，转基因盒的设计确保以病毒方式递送的回路受体在启动受体与其同源配体结合并释放可裂解的转录因子之前不表达CAR。

[0268] 图2示出了产生本公开的工程化免疫细胞的示例性方法的示意图。最初，激活T细胞。T细胞可获自患者。因此，本公开提供了从患者收获免疫细胞诸如T细胞的方法。接着，将编码CAR和启动受体的质粒引入T细胞中。有利地，可以使用电穿孔引入本公开的质粒。当通过电穿孔引入质粒时，也可以引入核酸酶。通过使用电穿孔，本公开的方法避免了使用病毒载体来引入转基因，这是免疫细胞工程化中的一个已知瓶颈。接着将T细胞扩增并共同培养，以产生足够数量的工程化免疫细胞，用作治疗性治疗。

[0269] 用于编辑细胞的基因组的方法可以包括：a)提供Cas9核糖核蛋白复合物(RNP)-DNA模板复合物，所述RNP-DNA模板复合物包含：(i)RNP，其中所述RNP包含Cas9核酸酶结构域和指导RNA，其中所述指导RNA与细胞的基因组的靶区特异性地杂交，并且其中所述Cas9核酸酶结构域使靶区裂解以在细胞的基因组中产生插入位点；和(ii)双链或单链DNA模板，其中所述DNA模板的大小为大于约200个核苷酸，其中所述DNA模板的5’端和3’端包含与插入位点侧翼的基因组序列同源的核苷酸序列，并且其中所述复合物中RNP与DNA模板的摩尔比为约3:1至约100:1；以及b)将所述RNP-DNA模板复合物引入细胞中。

[0270] 在一些实施方案中，本文所述的方法提供至少约20％、25％、30％、35％、40％、45％、50％、55％、60％、65％、70％、75％、80％、85％、90％、95％、97.5％、99％、99.5％、99％或更高的RNP-DNA模板复合物递送效率。在一些情况下，效率是相对于将RNP-DNA模板引入细胞中后存活的细胞来确定的。在一些情况下，效率是相对于将RNP-DNA模板引入细胞中的细胞(活细胞或非活细胞)的总数确定的。

[0271] 作为另一实例，可以通过定量细胞群体中经基因组编辑的细胞的数量来确定递送效率(与引入步骤后获得的总细胞数或总活细胞数相比较)。可以利用各种定量基因组编辑的方法。这些方法包括但不限于使用错配特异性核酸酶，诸如T7核酸内切酶I；对一个或多个靶基因座的测序(例如，通过对克隆的靶基因座扩增片段的桑格测序)；以及高通量深度测序。

[0272] 在一些实施方案中，与将裸DNA引入细胞中或使用病毒载体将DNA引入细胞中后的细胞活力损失相比，细胞活力的损失减少。减少可以是减少至少10％、20％、30％、40％、50％、60％、70％、80％、90％、100％或这些百分比之间的任何百分比。在一些实施方案中，与将裸DNA引入细胞中或使用病毒载体将DNA引入细胞中后的脱靶整合相比，脱靶整合效应减少。减少可以是减少至少10％、20％、30％、40％、50％、60％、70％、80％、90％、100％或这些百分比之间的任何百分比。

[0273] 在一些情况下，本文所述的方法提供了已引入RNP-DNA模板的细胞的高细胞活力。在一些情况下，已引入RNP-DNA模板的细胞的活力为至少约20％、25％、30％、35％、40％、45％、50％、55％、60％、65％、70％、75％、80％、85％、90％、95％、97.5％、99％、99.5％、99％或更高。在一些情况下，已引入RNP-DNA模板的细胞的活力为约20％至约99％、约30％至约90％、约35％至约85％或90％或更高、约40％至约85％或90％或更高、约50％至约85％

或90％或更高、约50％至约85％或90％或更高、约60％至约85％或90％或更高，或约70％至约85％或90％或更高。

[0274] 在本文提供的方法中，RNP与DNA模板的摩尔比可为约3:1至约100:1。例如，摩尔比可为约5:1至10:1、约5:1至约15:1、5:1至约20:1、5:1至约25:1；约8:1至约12:1、约8:1至约15:1、约8:1至约20:1，或约8:1至约25:1。

[0275] 在一些实施方案中，DNA模板的浓度为约2.5pM至约25pM。例如，DNA模板的浓度可为约2.5、3、3.5、4、4.5、5、5.5、6、6.5、7、7.5、8、8.5、9、9.5、10、10.5、11、11.5、12、12.5、13、13.5、14、14.5、15、15.5、16、16.5、17、17.5、18、18.5、19、19.5、20、20.5、21、21.5、22、22.5、23、23.5、24、24.5、25pM或这些浓度之间的任何浓度。

[0276] 在一些实施方案中，DNA模板的大小或长度为大于约4.5kb、5.0kb、5.1kb、5.2kb、5.3kb、5.4kb、5.5kb、5.6kb、5.7kb、5.8kb、5.9kb、6.0kb、6.1kb、6.2kb、6.3kb、6.4kb、6.5kb、6.6kb、6.7kb、6.8kb、6.9kb、7.0kb、7.1kb、7.2kb、7.3kb、7.4kb、7.5kb、7.6kb、7.7kb、7.8kb、7.9kb、8.0kb、8.1kb、8.2kb、8.3kb、8.4kb、8.5kb、8.6kb、8.7kb、8.8kb、8.9kb、9.0kb、9.1kb、9.2kb、9.3kb、9.4kb、9.5kb、9.6kb、9.7kb、9.8kb、9.9kb或10kb或介于这些大小之间的任何DNA模板大小。例如，DNA模板的大小可为约4.5kb至约10kb、约5kb至约10kb、约5kb至约9kb、约5kb至约8kb、约5kb至约7kb、约5kb至约6kb、约kb 6至约10kb、约6kb至约9kb、约6kb至约8kb、约6kb至约7kb、约7kb至约10kb、约7kb至约9kb、约7kb至约8kb、约8kb至约10kb、约8kb至约9kb，或约9kb至约10kb。

[0277] 在一些实施方案中，DNA模板的量为约1μg至约10μg。例如，DNA模板的量可为约1μg至约2μg、约1μg至约3μg、约1μg至约4μg、约1μg至约5μg、约1μg至约6μg、约1μg至约7μg、约1μg至约8μg、约1μg至约9μg、约1μg至约10μg。在一些实施方案中，DNA模板的量为约2μg至约3μg、约2μg至约4μg、约2μg至约5μg、约2μg至约6μg、约2μg至约7μg、约2μg至约8μg、约2μg至约9μg，或2μg至约10μg。在一些实施方案中，DNA模板的量为约3μg至约4μg、约3μg至约5μg、约3μg至约6μg、约3μg至约7μg、约3μg至约8μg、约3μg至约9μg，或约3μg至约10μg。在一些实施方案中，DNA模板的量为约4μg至约5μg、约4μg至约6μg、约4μg至约7μg、约4μg至约8μg、约4μg至约9μg，或约4μg至约10μg。在一些实施方案中，DNA模板的量为约5μg至约6μg、约5μg至约7μg、约5μg至约8μg、约5μg至约9μg，或约5μg至约10μg。在一些实施方案中，DNA模板的量为约6μg至约7μg、约6μg至约8μg、约6μg至约9μg，或约6μg至约10μg。在一些实施方案中，DNA模板的量为约7μg至约8μg、约7μg至约9μg，或约7μg至约10μg。在一些实施方案中，DNA模板的量为约8μg至约9μg，或约8μg至约10μg。在一些实施方案中，DNA模板的量为约9μg至约10μg。

[0278] 在一些情况下，DNA模板的大小和数量足够大，足以作为裸DNA致死。在一些实施方案中，DNA模板编码异源蛋白质或其片段。在一些实施方案中，DNA模板编码至少一个基因。在一些实施方案中，DNA模板编码至少两个基因。在一些实施方案中，DNA模板编码一个、两个、三个、四个、五个、六个、七个、八个、九个、十个或更多个基因。

[0279] 在一些实施方案中，DNA模板包括调控序列，例如启动子序列和/或增强子序列，以在插入细胞基因组中后调控异源蛋白质或其片段的表达。

[0280] 在一些情况下，DNA模板是线状DNA模板。在一些情况下，DNA模板是单链DNA模板。在一些情况下，单链DNA模板是纯单链DNA模板。如本文所用，“纯单链DNA”是指实质上缺乏另一条或相反DNA链的单链DNA。“实质上缺乏”是指纯单链DNA中的一条链比另一条DNA链缺

乏至少100倍。

[0281] 在一些情况下，RNP-DNA模板复合物是通过将RNP与DNA模板在约20℃至约25℃的温度下孵育少于约一分钟至约三十分钟而形成的。例如，可以将RNP与DNA模板在约20℃、21℃、22℃、23℃、24℃或25℃的温度下孵育约5秒、10秒、15秒、20秒、25秒、30秒、35秒、40秒、45秒、50秒、55秒、1分钟、2分钟、3分钟、4分钟、5分钟、6分钟、7分钟、8分钟、9分钟、10分钟、11分钟、12分钟、13分钟、14分钟、15分钟、16分钟、17分钟、18分钟、19分钟、20分钟、21分钟、22分钟、23分钟、24分钟、25分钟、26分钟、27分钟、28分钟、29分钟或30分钟或这些时间之间的任何时间量。在另一实例中，可以将RNP与DNA模板在约20℃至约25℃的温度下孵育少于约一分钟至约一分钟、少于约一分钟至约5分钟、少于约1分钟至约10分钟、约5分钟至10分钟、约5分钟至15分钟、约10至约15分钟、约10分钟至约20分钟，或约10分钟至约30分钟。在一些实施方案中，在将RNP-DNA模板复合物引入细胞中之前将RNP-DNA模板复合物和细胞混合。

[0282] 在一些实施方案中，引入RNP-DNA模板复合物包括电穿孔。用于电穿孔细胞以引入RNP-DNA模板复合物的方法、组合物和装置可包括本文实施例中描述的那些。用于电穿孔细胞以引入RNP-DNA模板复合物的另外的或替代的方法、组合物和装置可包括WO/2006/001614或Kim,J.A.等人Biosens.Bioelectron.23,1353-1360(2008)中描述的那些。用于电穿孔细胞以引入RNP-DNA模板复合物的另外的或替代的方法、组合物和装置可包括美国专利申请公布第2006/0094095号;第2005/0064596号;或第2006/0087522号中描述的那些。用于电穿孔细胞以引入RNP-DNA模板复合物的另外的或替代的方法、组合物和装置可包括Li,L.H.等人Cancer Res.Treat.1,341-350(2002);美国专利第6,773,669号;第7,186,559号;第7,771,984号;第7,991,559号;第6485961号;第7029916号;以及美国专利申请公布第2014/0017213号;和第2012/0088842号中描述的那些，所有这些参考文献据此以引用的方式并入本文。用于电穿孔细胞以引入RNP-DNA模板复合物的另外的或替代的方法、组合物和装置可包括Geng,T.等人J.Control Release 144,91-100(2010);以及Wang,J.等人Lab.Chip 10,2057-2061(2010)中描述的那些，所有这些参考文献据此以引用的方式并入本文。

[0283] 在一些实施方案中，Cas9蛋白可以呈活性核酸内切酶形式，使得当作为与指导RNA的复合物的一部分或作为与DNA模板的复合物的一部分与靶核酸结合时，将双链断裂引入靶核酸中。双链断裂可以通过NHEJ修复以引入随机突变，或通过HDR修复以引入特定突变。在本文所述的方法中可以利用多种Cas9核酸酶。例如，可以利用Cas9核酸酶，该酶需要紧邻指导RNA所靶向的区域的3'处的NGG原型间隔区相邻基序(PAM)。此类Cas9核酸酶可以靶向含有NGG序列的基因组的任何区域。作为另一实例，具有正交PAM基序要求的Cas9蛋白可用于靶向没有相邻NGG PAM序列的序列。具有正交PAM序列特异性的示例性Cas9蛋白包括但不限于CFP1;Nature Methods 10,1116-1121(2013)中描述的那些;以及Zetsche等人,Cell,第163卷,第3期,p759-771,2015年10月22日中描述的那些，两者均据此以引用的方式并入本文。

[0284] 在一些情况下，Cas9蛋白是切口酶，使得当作为与指导RNA的复合物的一部分与靶核酸结合时，将单链断裂或切口引入靶核酸中。一对Cas9切口酶(每个切口酶都与结构不同的指导RNA结合)可以靶向靶基因组区域的两个近端位点，从而将一对近端单链断裂引入靶

基因组区域中。切口酶对可以提供增强的特异性,因为脱靶效应可能导致单个切口,这些切口一般通过碱基切除修复机制修复而没有损伤。示例性的Cas9切口酶包括具有D10A或H840A突变的Cas9核酸酶。

[0285] 在一些实施方案中,RNP包含Cas9核酸酶。在一些实施方案中,RNP包含Cas9切口酶。在一些实施方案中,RNP-DNA模板复合物包含至少两个结构不同的RNP复合物。在一些实施方案中,所述至少两个结构不同的RNP复合物含有结构不同的Cas9核酸酶结构域。在一些实施方案中,所述至少两个结构不同的RNP复合物含有结构不同的指导RNA。在一些实施方案中,其中所述至少两个结构不同的RNP复合物含有结构不同的指导RNA,每个结构不同的RNP复合物包含Cas9切口酶,并且结构不同的指导RNA与靶区的相反链杂交。

[0286] 在一些情况下,将包含结构不同的核糖核蛋白复合物的多个RNP-DNA模板引入细胞中。例如,Cas9蛋白可以与多个(例如,2个、3个、4个、5个或更多,例如,2-10个、5-100个、20-100个)结构不同的指导RNA复合,以在多个结构不同的靶基因组区域处靶向插入DNA模板。

[0287] 在本文提供的方法和组合物中,细胞包括但不限于真核细胞、原核细胞、动物细胞、植物细胞、真菌细胞等。任选地,细胞是哺乳动物细胞,例如人细胞。细胞可以在体外,离体或在体内。细胞还可以是原代细胞、生殖细胞、干细胞或前体细胞。前体细胞可以是例如多能干细胞或造血干细胞。在一些实施方案中,细胞是原代造血细胞或原代造血干细胞。在一些实施方案中,原代造血细胞是免疫细胞。在一些实施方案中,免疫细胞是T细胞。在一些实施方案中,T细胞是调节性T细胞、效应T细胞或幼稚T细胞。在一些实施方案中,T细胞是$CD4^+$T细胞。在一些实施方案中,T细胞是$CD8^+$T细胞。在一些实施方案中,T细胞是$CD4^+CD8^+$T细胞。在一些实施方案中,T细胞是$CD4^-CD8^-$T细胞。还提供了通过本文所述的任何方法修饰的任何细胞的群体。在一些实施方案中,所述方法还包括扩大修饰的细胞的群体。

[0288] 在一些情况下,将细胞从受试者中取出,使用本文所述的任何方法进行修饰并施用至患者。在其他情况下,本文所述的任何构建体递送至患者体内。参见,例如,美国专利第9737604号和Zhang等人"Lipid nanoparticle-mediated efficient delivery of CRISPR/Cas9 for tumor therapy,"NPG Asia Materials第9期,第e441页(2017),两者均据此以引用的方式并入本文。

[0289] 在一些实施方案中,将RNP-DNA模板复合物引入约$1x\ 10^5$至约$2x\ 10^6$个细胞中。例如,可以将RNP-DNA模板复合物引入约$1x\ 10^5$至约$5x\ 10^5$个细胞、约$1x\ 10^5$至约$1x\ 10^6$个细胞、$1x\ 10^5$至约$1.5x\ 10^6$个细胞、$1x\ 10^5$至约$2x\ 10^6$个细胞、约$1x\ 10^6$至约$1.5x\ 10^6$个细胞或约$1x\ 10^6$至约$2x\ 10^6$个细胞中。

[0290] 在一些情况下,本文所述的方法和组合物可用于产生、修饰、使用或控制重组T细胞,诸如嵌合抗原受体T细胞(CAR T细胞)。此类CAR T细胞可用于治疗或预防受试者的癌症、传染病或自身免疫性疾病。例如,在一些实施方案中,将一个或多个基因产物插入或敲入T细胞以表达异源蛋白(例如,嵌合抗原受体(CAR)或启动受体)。

[0291] 插入位点

[0292] 用于编辑T细胞基因组的方法,具体地包括编辑人T细胞基因组的方法,包括将核酸序列或构建体插入人T细胞中的TCR-α亚基(TRAC)基因的外显子1的靶区中。在一些实施方案中,靶区位于TRAC基因的恒定结构域的外显子1中。在其他实施方案中,靶区位于外显

子1、外显子2或外显子3中，在编码TCR-α跨膜结构域的序列开始之前。

[0293] 用于编辑T细胞基因组的方法还包括编辑人T细胞基因组的方法，包括将核酸序列或构建体插入人T细胞中的TCR-β亚基(TRBC)基因的外显子1的靶区中。在一些实施方案中，靶区位于TRBC1或TRBC2基因的外显子1中。

[0294] 用于编辑T细胞基因组的方法，具体地包括编辑人T细胞基因组的方法，包括将核酸序列或构建体插入基因组安全港(GSH)的靶区中。

[0295] 基因编辑疗法包括，例如，载体整合和位点特异性整合。位点特异性整合是病毒载体随机整合的一种有前景的替代方法，因为它降低了插入诱变或插入瘤形成的风险(Kolb等人Trends Biotechnol.2005 23:399-406;Porteus等人Nat Biotechnol.2005 23:967-973;Paques等人Curr Gen Ther.2007 7:49-66)。然而，位点特异性整合仍然面临着诸如低敲入效率、插入瘤形成风险、相邻基因或转基因的不稳定和/或异常表达、低可及性(例如相邻基因的20kB以内)等挑战。这些挑战可以部分地通过确定和使用安全港基因座或安全港位点(SHS)来解决，这些位点是可以在不破坏相邻基因的表达或调控的情况下整合基因或遗传元件的位点。

[0296] 最广泛使用的推定人安全港位点是染色体19q上的AAVS1位点，该位点最初被确定为用于复发性腺相关病毒插入的位点。已在同源性的基础上鉴定出其他潜在的SHS，这些位点首先是在其他物种(例如，允许的鼠Rosa26基因座的人同系物)或在越来越多的在某些情况下显得非必需的人基因中鉴定出来的。这种类型的一种推定的SHS是CCR5趋化因子受体基因，该基因当被破坏时赋予对人免疫缺陷病毒感染的抵抗力。在病毒整合位点作图或基因陷阱分析的基础上，已在人和其他细胞类型中鉴定出其他潜在的基因组SHS，就像最初的小鼠Rosa26基因座一样。三个顶级SHS、AAVS1、CCR5和Rosa26与许多蛋白质编码基因和调控元件非常接近。(参见Sadelain,M.等人(2012).Safe harbours for the integration of new DNA in the human genome.Nature reviews Cancer,12(1),51-58，其相关公开内容以引用的方式整体并入本文)。

[0297] 人19号染色体上的AAVS1(也称为PPP1R12C基因座)是一个已知的用于承载具有预期功能的转基因(例如DNA转基因)的SHS。它位于位置19q13.42。它具有开放的染色质结构并且具有转录能力。AAVS1的规范SHS基因座是chr19:55,625,241-55,629,351。参见Pellenz等人“New Human Chromosomal Sites with″Safe Harbor″Potential for Targeted Transgene Insertion.”Human gene therapy vol.30,7(2019):814-828，其相关公开内容以引用的方式并入本文。下文提供了示例性AAVS1靶gRNA和靶序列：

[0298] ●AAVS1-gRNA序列:ggggccactagggacaggatGTTTTAGAGCTAGAAA TAGCAAGTTAAAAT AAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCA CCGAGTCGGTGCTTTTTTT

[0299] ●AAVS1靶序列:ggggccactagggacaggat

[0300] CCR5位于3号染色体的位置3p21.31，编码HIV-1的主要共受体。CCR5基因中该位点的破坏有利于HIV/AIDS疗法，并促进了靶向其第三外显子的锌指核酸酶的开发。CCR5的规范SHS基因座是chr3:46,414,443-46,414,942。参见Pellenz等人“New Human Chromosomal Sites with″Safe Harbor″Potential for Targeted Transgene Insertion.”Human gene therapy vol.30,7(2019):814-828，其相关公开内容以引用的方式并入本文。

[0301] 小鼠Rosa26基因座对于基因修饰特别有用，因为它可以被高效靶向并在大多数测

试细胞类型中表达。Irion等人2007("Identification and targeting of the ROSA26 locus in human embryonic stem cells."Nature biotechnology 25.12(2007):1477-1482,其相关公开内容以引用的方式并入本文)鉴定了3号染色体(位置3p25.3)中的人同系物,人ROSA26。人Rosa26(hRosa26)的规范SHS基因座是chr3:9,415,082 -9,414,043。参见Pellenz等人"New Human Chromosomal Sites with"Safe Harbor"Potential for Targeted Transgene Insertion."Human gene therapy vol.30,7(2019):814-828,其相关公开内容以引用的方式并入本文。

[0302] 安全港位点的另外实例提供于Pellenz等人"New Human Chromosomal Sites with"Safe Harbor"Potential for Targeted Transgene Insertion."Human gene therapy vol.30,7(2019):814-828,其相关公开内容以引用的方式并入本文。另外的整合位点的实例提供于表1中。

[0303] 在一些实施方案中,安全港位点允许转基因高表达(足以实现转基因功能性或对目标疾病的治疗)和转基因在数天、数周或数月内的稳定表达。在一些实施方案中,安全港基因座处基因的敲除赋予细胞功能益处,或者安全港基因座处的基因在细胞内没有已知功能。在一些实施方案中,安全港基因座使得:在有或没有CD3/CD28刺激的情况下转基因在体外稳定表达,如通过iGuide-Seq或CRISPR-Seq所检测脱靶裂解可忽略不计,如通过iGuide-Seq或CRISPR-Seq所检测相对于其他基因座脱靶裂解减少,转基因非依赖性细胞毒性可忽略不计,转基因非依赖性细胞因子表达可忽略不计,转基因非依赖性嵌合抗原受体表达可忽略不计,附近基因的失调控或沉默可忽略不计,以及定位于癌相关基因之外。

[0304] 如所用,"附近基因"可以指代距离安全港基因座(整合位点)约100kB、约125kB、约150kB、约175kB、约200kB、约225kB、约250kB、约275kB、约300kB、约325kB、约350kB、约375kB、约400kB、约425kB、约450kB、约475kB、约500kB、约525kB、约550kB内的基因。

[0305] 在一些实施方案中,本公开设想包含一个或多个转基因的插入物。转基因可以编码治疗性蛋白质、抗体、肽、自杀基因、凋亡基因或任何其他目标基因。转基因整合可以导致例如增强的治疗特性。如本文所用,这些增强的治疗特性是指与相同正常细胞类型的典型免疫细胞相比,细胞的增强的治疗特性。例如,与典型的、未修饰的和/或天然存在的NK细胞相比,具有"增强的治疗特性"的NK细胞具有增强的、改善的和/或增加的治疗结果。免疫细胞的治疗特性可以包括但不限于细胞移植、运输、归巢、活力、自我更新、持久性、免疫应答控制和调控、存活和细胞毒性。免疫细胞的治疗特性还表现为:抗原靶向受体表达;HLA呈递或缺乏;对瘤内微环境的耐受性;旁观者免疫细胞的诱导和免疫调节;减量下靶标特异性提高;对治疗诸如化学疗法的抗性。

[0306] 如本文所用,术语"插入物大小"是指整合(插入)到靶基因座或安全港位点处的核苷酸序列的长度。在一些实施方案中,插入物大小包括至少约4.5千碱基对(kb)至约10千碱基对(kb)。在一些实施方案中,插入物大小包括约5000个核苷酸或更多碱基对。在一些实施方案中,插入物大小包括多达4.5、4.8、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20kbp(千碱基对)或介于之间的大小。在一些实施方案中,插入物大小为大于4.5、4.8、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20kbp或介于之间的大小。在一些实施方案中,插入物大小在4.5-15kbp的范围内或者是该范围内的任何数字。在一些实施方案中,插入物大小在4.8-8.3kbp的范围内或者是该范围内的任何数字。在一些实施方案中,插入物大小在

5-8.3kbp的范围内或者是该范围内的任何数字。在一些实施方案中，插入物大小在5-15kbp的范围内或者是该范围内的任何数字。在一些实施方案中，插入物大小在4.5-20kbp的范围内或者是该范围内的任何数字。在一些实施方案中，插入物大小为5-10kbp。在一些实施方案中，插入物大小为4.5-10、5-10、6-10、7-10、8-10、9-10kbp。在一些实施方案中，插入物大小为4.5-11、6-11、7-11、8-11、9-11或10-11kbp。在一些实施方案中，插入物大小为4.5-12、6-12、7-12、8-12、9-12、10-12或11-12kbp。在一些实施方案中，插入物大小为4.5-13、6-13、7-13、8-13、9-13、10-13、11-13或12-13kbp。在一些实施方案中，插入物大小为4.5-14、6-14、7-14、8-14、9-14、10-14、11-14、12-14或13-14kbp。在一些实施方案中，插入物大小为4.5-15、6-15、7-15、8-15、9-15、10-15、11-15、12-15、13-15或14-15kbp。在一些实施方案中，插入物大小为4.5-16、6-16、7-16、8-16、9-16、10-16、11-16、12-16、13-16、14-16或15-16kbp。在一些实施方案中，插入物大小为4.5-17、6-17、7-17、8-17、9-17、10-17、11-17、12-17、13-17或14-17、15-17或16-17kbp。在一些实施方案中，插入物大小为4.5-18、6-18、7-18、8-18、9-18、10-18、11-18、12-18、13-18、14-18、15-18、16-18或17-18kbp。在一些实施方案中，插入物大小为4.5-19、6-19、7-19、8-19、9-19、10-19、11-19、12-19、13-19、14-19、15-19、16-19、17-19或18-19kbp。在一些实施方案中，插入物大小为4.5-20、6-20、7-20、8-20、9-20、10-20、11-20、12-20、13-20、14-20、15-20、16-20、17-20、18-20或19-20kbp。

[0307] 本公开的插入物是指插入到靶基因座或安全港位点处的核酸分子或多核苷酸。在一些实施方案中，核苷酸序列是DNA分子例如基因组DNA，或者包含脱氧核糖核苷酸。在一些实施方案中，插入物包含较小的DNA片段，诸如质体DNA、线粒体DNA，或以质粒、福斯质粒、粘粒、细菌人工染色体(BAC)、酵母人工染色体(YAC)和/或DNA的任何其他亚基因组区段形式分离的DNA。在一些实施方案中，插入物是RNA分子或者包含核糖核苷酸。插入物中的核苷酸被认为是天然存在的核苷酸、非天然存在的核苷酸和修饰的核苷酸。如本领域技术人员将容易理解的，核苷酸可进行化学或生物化学修饰，或者可含有非天然或衍生化的核苷酸碱基。此类修饰包括例如标记、甲基化、用类似物取代一个或多个天然存在的核苷酸、核苷酸间修饰。多核苷酸可以是任何拓扑构象，包括单链构象、双链构象、部分双链构象、三链构象、发夹式构象、环状象和本领域所设想的其他三维构象。

[0308] 插入物可以具有编码和/或非编码区。插入物可以包含非编码序列(例如，控制元件，例如启动子序列)。在一些实施方案中，插入物编码转录因子。在一些实施方案中，插入物编码抗原结合受体，诸如单一受体、T细胞受体(TCR)、启动受体、CAR、mAb等。在一些实施方案中，插入物是RNAi分子，包括但不限于miRNA、siRNA、shRNA等。在一些实施方案中，插入物是人序列。在一些实施方案中，插入物是模拟物。在一些实施方案中，插入物是多基因/多模块治疗性盒。多基因/多模块治疗性盒是指具有一种或超过一种受体(例如，合成受体)、其他外源蛋白编码序列、非编码RNA、转录调控元件和/或隔离子序列(insulator sequence)等的插入物或盒。

[0309] 在一些实施方案中，将核酸序列通过非病毒递送插入T细胞的基因组中。在非病毒递送方法中，核酸可以是裸DNA，或在非病毒质粒或载体中。如本文所述或本领域普通技术人员已知的，非病毒递送技术可以是位点特异性整合技术。用于整合到安全港基因座中的位点特异性技术的实例包括但不限于使用核酸酶的同源依赖性工程化和使用Cas9或其他CRISPR核酸内切酶的同源非依赖性靶向插入。

[0310] 在一些实施方案中,通过向工程化细胞中引入以下各项将插入物整合到安全港位点处:(a)使安全港位点中的靶区裂解而产生插入位点的靶向核酸酶;和(b)核酸序列(插入物),其中插入物通过例如HDR并入插入位点。可用于本公开的方法中的非病毒递送技术的实例提供于美国申请第16/568,116号和第16/622,843号中,其相关公开内容以引用的方式整体并入本文。

[0311] CRISPR-Cas编辑

[0312] 基因编辑的一个有效实例是CRISPR-Cas方法(例如CRISPR-Cas9)。该方法结合使用指导多核苷酸(例如指导核糖核酸或gRNA)和cas核酸内切酶(例如Cas9核酸内切酶)。

[0313] 如本文所用,称为"Cas核酸内切酶"或具有"Cas核酸内切酶活性"的多肽是指由Cas基因编码的CRISPR相关(Cas)多肽,其中Cas多肽是当可操作地连接至一个或多个指导多核苷酸时可以被裂解的靶DNA序列(参见,例如,美国专利第8,697,359号)。该定义还包括保留指导多核苷酸依赖性核酸内切酶活性的Cas核酸内切酶的变体。在本文详述的供体DNA插入方法中使用的Cas核酸内切酶是将双链断裂引入DNA中的靶位点处(例如,靶基因座内或安全港位点处)的核酸内切酶。

[0314] 如本文所用,术语"指导多核苷酸"涉及能够与Cas核酸内切酶复合并允许Cas核酸内切酶识别并裂解DNA靶位点的多核苷酸序列。指导多核苷酸可以是单分子或双分子。指导多核苷酸序列可以是RNA序列、DNA序列或它们的组合(RNA-DNA组合序列)。仅包含核糖核酸的指导多核苷酸也称为"指导RNA"。在一些实施方案中,使用指导RNA(gRNA)与cas核酸内切酶(例如Cas9核酸内切酶)的组合将多核苷酸供体构建体插入安全港基因座处。

[0315] 指导多核苷酸包括与靶DNA中的核苷酸序列互补的第一核苷酸序列结构域(也称为可变靶向结构域或VT结构域),以及与Cas核酸内切酶多肽相互作用的第二核苷酸。它可以是包含序列结构域(称为Cas核酸内切酶识别结构域或CER结构域)的双分子(也称为双链指导多核苷酸)。这种双分子指导多核苷酸的CER结构域包含两个沿着互补区杂交的独立分子。两个独立分子可以是RNA序列、DNA序列和/或RNA-DNA组合序列。

[0316] 使用CRISPR-Cas方法的基因组编辑依赖于由RNA指导的Cas核酸内切酶(例如Cas9核酸内切酶)诱导的位点特异性DNA双链断裂(DSB)的修复。对这些DSB的同源定向修复(HDR)使得能够通过引入限定的基因组变化(包括碱基取代、序列插入和缺失)实现基因组的精确编辑。常规的基于HDR的CRISPR/Cas9基因组编辑涉及用Cas9、gRNA和含有与目标基因组位点匹配的同源臂的供体DNA转染细胞。

[0317] HITI(同源非依赖性靶向插入)使用基于非同源末端接合(NHEJ)的同源非依赖性策略,并且该方法比HDR更有效。指导RNA(gRNA)靶向插入位点。对于HITI,供体质粒缺乏同源臂,并且DSB修复不会通过HDR途径发生。供体多核苷酸构建体可以经工程化以包括位于待插入基因或序列侧翼的Cas9裂解位点。这导致Cas9在供体质粒和基因组靶序列处裂解。靶标和供体都有平端,线性化的供体DNA质粒被NHEJ途径所用,导致整合到基因组DSB位点中。(参见,例如,Suzuki,K.等人(2016).In vivo genome editing via CRISPR/Cas9 mediated homology-independent targeted integration.Nature,540(7631),144-149,其相关公开内容整体并入本文)。

[0318] 使用CRISPR-Cas方法进行基因编辑的方法是本领域普通技术人员已知的。(参见,例如,美国申请第US16/312,676号、第US15/303,722号和第US15/628,533号,其公开内容以

引用的方式整体并入本文)。此外,核酸内切酶用于将转基因插入安全港基因座中的用途描述于例如美国申请第13/036,343号中,其公开内容以引用的方式整体并入本文。

[0319] 编码核酸内切酶的指导RNA和/或mRNA(或DNA)可以化学连接至一个或多个增强寡核苷酸的活性、细胞分布或细胞摄取的部分或缀合物。此类部分的非限制性实例包括脂质部分(诸如胆固醇部分)、胆酸、硫醚、巯基胆固醇、脂族链(例如,十二烷二醇或十一烷基残基)、磷脂(例如二-十六烷基-外消旋-甘油或三乙基铵1,2-二-O-十六烷基-外消旋-甘油-3-H-膦酸酯)、聚胺或聚乙二醇链、金刚烷乙酸、棕榈基部分和十八胺或己基氨基-羰基-t-羟胆固醇部分。参见例如美国专利公布第20180127786号,其公开内容以引用的方式整体并入本文。

[0320] 治疗应用

[0321] 对于治疗应用,将工程化细胞、其群体或其组合物以有效量施用至受试者,一般是哺乳动物,一般是人。

[0322] 可通过输注(例如,在一段时间内连续输注)或本领域普通技术人员已知的其他施用方式将工程化细胞施用至受试者。

[0323] 本文提供的工程化细胞不仅可用于基因疗法,还可用于非药物用途,诸如例如动物模型的产生以及表达目标蛋白质的重组细胞系的产生。

[0324] 本公开的工程化细胞可以是已通过在本文所述的安全港基因座处整合转基因而经修饰的任何细胞,一般是哺乳动物细胞,一般是人细胞。在一些实施方案中,工程化细胞是免疫细胞。在一些实施方案中,工程化细胞是淋巴细胞。在一些实施方案中,工程化细胞是T细胞或T细胞前体。

[0325] 本公开的工程化细胞、组合物和方法可用于治疗应用,诸如CAR T细胞疗法和TCR T细胞疗法。在一些实施方案中,在安全港基因座内插入编码转基因的序列相对于没有插入时的情况保持TCR表达,并在保持TCR功能的同时实现转基因表达。

[0326] 在一些实施方案中,本公开提供了通过向需要治疗的受试者施用包含本文所述的任何工程化细胞的组合物来治疗受试者的方法。如所使用,术语“治疗(treat/treatment)”等一般指获得所需药理和/或生理作用。该作用就完全或部分预防疾病而言是预防性的并且/或者在部分或完全治愈疾病和/或由疾病引起的副作用而言是治疗性的。如本文所用,术语“治疗”涵盖对受试者(例如,哺乳动物,例如人)的疾病的任何治疗。治疗还可指代将本文提供的工程化细胞施用至对疾病易感但尚未被诊断为罹患该疾病的受试者,包括预防疾病的发生;抑制疾病进展;或减轻疾病(即,导致疾病消退)。此外,治疗可以稳定或减少受试者(例如,患者)的不良临床症状。本文提供的细胞、其群体或其组合物可在疾病或损伤发生之前、期间或之后施用。

[0327] 在某些实施方案中,受试者患有可以通过细胞疗法治疗和/或改善的疾病、疾患和/或损伤。在一些实施方案中,需要细胞疗法的受试者是具有损伤、疾病或疾患的受试者,从而引起细胞疗法(例如,将细胞物质施用至受试者的疗法)。然而,设想有可能治疗、改善与损伤、疾病或疾患相关的至少一种症状和/或降低其严重性。在某些实施方案中,需要细胞疗法的受试者包括但不限于骨髓移植或干细胞移植候选者、已接受化学疗法或放射疗法的受试者、过度增殖性疾病或癌症(例如,造血系统)、患有或有风险发展过度增殖性疾病或癌症的受试者、患有或有风险发展肿瘤(例如,实体瘤)、病毒感染或病毒的受试者。还意图

涵盖罹患或有风险罹患与感染相关的疾病的受试者。

[0328] 在一些实施方案中，本公开提供本公开的组合物连同使用说明。使用说明可以作为包装插页存在于药盒中，存在于药盒容器或其部件的标签中，或者可以呈数字形式(例如，在CD-ROM上，通过互联网上的链接)。药盒可以包括一种或多种基因组靶向核酸、编码基因组靶向核酸的多核苷酸、定点多肽和/或编码定点多肽的多核苷酸。还设想了药盒内的另外组分，例如缓冲液(诸如复原缓冲液、稳定缓冲液、稀释缓冲液)和/或一种或多种对照载体。

[0329] 药物组合物

[0330] 本文提供的工程化重组细胞可以作为药物组合物的一部分来施用。除重组细胞中的一者或多者之外，这些组合物还可包含药学上可接受的赋形剂、载体、缓冲剂、稳定剂或为本领域技术人员所熟知的其他物质。所述物质应是无毒的，并且不应干扰活性成分的功效。载体或其他物质的确切性质可取决于施用途径，例如口服、静脉内、经皮肤或皮下、经鼻、肌肉内、腹膜内途径。药物组合物可包含一种或多种药物赋形剂。可使用任何适合药物赋形剂，并且本领域普通技术人员能够选择适合药物赋形剂。因此，以下提供的药物赋形剂旨在说明而非限制。另外的药物赋形剂包括例如描述于以下中的那些:Handbook of Pharmaceutical Excipients,Rowe等人(编)第6版(2009)，其以引用的方式整体并入。

[0331] 本文设想了施用另外治疗剂的各种模式。在一些实施方案中，通过任何适合施用模式施用另外的治疗剂。一般来说，施用模式包括但不限于玻璃体内、视网膜下、脉络膜上、动脉内、真皮内、肌肉内、腹膜内、静脉内、经鼻、肠胃外、外用、经肺和皮下途径。

[0332] 用于口服施用的药物组合物可呈片剂、胶囊、粉剂或液体形式。片剂可包括固体载体诸如明胶或佐剂。液体药物组合物通常包括液体载体诸如水、石油、动物油或植物油、矿物油或合成油。可包括生理盐水溶液、右旋糖或其他糖溶液或二醇诸如乙二醇、丙二醇或聚乙二醇。

[0333] 对于静脉内、经皮肤或皮下注射，或在病痛部位处的注射，活性成分将呈胃肠外可接受的水溶液形式，所述水溶液是无热原的，并且具有适合pH、等张性和稳定性。本领域相关技术人员完全能够使用例如等张媒介物诸如氯化钠注射液，林格氏注射液(Ringer's Injection)、乳酸林格氏注射液来制备适合溶液。防腐剂、稳定剂、缓冲剂、抗氧化剂和/或其他添加剂可根据需要加以包括。

[0334] 无论对个体给与的是多肽、细胞或核酸还是根据本公开的其他药学上有用的化合物，施用都优选地以“治疗有效量”或“防治有效量”(视情况而定，尽管防治可被视为治疗)进行，这足以显示对所述个体的益处。实际施用量以及施用的速率和时程将取决于所治疗的蛋白质聚集疾病的性质和严重性。对治疗的指定例如关于剂量的决定等在普通从业者和其他医师的职责范围内，并且通常考虑待治疗的疾患、单个患者的状况、递送部位、施用方法以及为从业者所知的其他因素。以上提及的技术和方案的实例可见于Remington's Pharmaceutical Sciences,第16版,Osol,A.(编)，1980中。

[0335] 视待治疗的疾患而定，组合物可单独施用，或与其他治疗进行组合，同时或依序加以施用。

[0336] 实施例

[0337] 以下是用于进行本公开的具体实施方案的实施例。实施例仅出于说明性目的而提

供,并且不意图以任何方式限制本公开的范围。已努力确保关于所用数值(例如,数量、温度等)的准确性,但当然应允许一些实验误差和偏差。

[0338] 除非另外指示,否则本公开的实施将采用属于本领域的技能的常规蛋白质化学、生物化学、重组DNA技术和药理学方法。此类技术充分说明于文献中。参见,例如,T.E.Creighton,Proteins:Structures and Molecular Properties(W.H.Freeman and Company,1993);A.L.Lehninger,Biochemistry(Worth Publishers,Inc.,本期新增);Sambrook等人,Molecular Cloning:A Laboratory Manual(第2版,1989);Methods In Enzymology(S.Colowick和N.Kaplan编,Academic Press,Inc.);Remington's Pharmaceutical Sciences,第18版(Easton,Pennsylvania:Mack Publishing Company,1990);Carey和Sundberg Advanced Organic Chemistry第3版(Plenum Press)第A和B卷(1992)。

[0339] 实施例1:利用大DNA插入物对免疫细胞的非病毒编辑.

[0340] 由于i)表达盒的大小限制,ii)与病毒载体生物学冲突所导致的转基因布局限制,以及iii)转基因在随机整合位点处的不良且不可预测的性能,因此基因回路或系统的病毒递送很困难。以下实施例演示了将现有病毒载体无法实现的大盒递送至限定插入位点,在那里它们具有可预测的高性能。

[0341] 构建体产生

[0342] 为了产生用于敲入的质粒构建体,从Twist、IDT和GENEWIZ订购合成的DNA,并通过Gibson Assembly和Golden Gate Assembly进行组装。质粒含有与基因组中CRISPR靶位点侧翼的序列同源的同源臂,长度为1.2kb。

[0343] T细胞工程化

[0344] 使用Lymphoprep(STEMCELL Technologies)和EasySep人T细胞分离试剂盒(STEMCELL Technologies)从获自正常供体Leukopaks(STEMCELL Technologies)获得的外周血单核细胞(PBMC)富集T细胞。随后将T细胞在补充有3%人AB血清(Gemini Bio)和12.5ng/ml人IL-7和IL-15(Miltenyi优质级)的TexMACS培养基(Miltenyi 130-197-196)中用CD3/CD28 Dynabead以1:1珠细胞比(ThermoFisher,40203D)激活,在37℃、5% CO2下培养48小时,然后电穿孔。

[0345] 通过将120μM靶向DNA序列的sgRNA(Synthego)GTCAGGGTTCTGGATATCTG(TRAC,SEQ ID NO:79)、62.5μM sNLS-SpCas9-sNLS(Aldevron)和P3缓冲液(Lonza)以5:1:3:6的体积比组合来制备CRISPR RNP,并将其在室温下孵育15分钟。将通过剂量滴定实验确定的优化量的质粒DNA(范围为0.5-3微克)与3.5μl RNP混合。对T细胞进行计数、去珠、以90X G离心10分钟,并以10^6个细胞/14.5μl添加有补充剂的P3(Lonza)重悬。将14.5μl T细胞悬浮液添加到DNA/RNP混合物中,转移到Lonza 16孔比色皿条中,并在代码为EH-115的Lonza 4D Nucleofector System中进行脉冲处理。使细胞在室温下静置15分钟,接着转移至96孔板(Sarstedt)中的补充有12.5ng/ml人IL-7和IL-15(Miltenyi优质级)的TexMACS培养基中。

[0346] 通过用抗Myc抗体(Cell Signaling Technology,克隆9B11)和抗Flag抗体(RnD systems,克隆1042E)染色检测转基因表达,并在Attune NxT流式细胞仪上进行分析。所使用的其他抗体是活/死Fixable Near-IR(Thermo Fisher)、TCRα/β抗体(BioLegend,克隆IP26)、CD4抗体(BioLegend,克隆RPA-T4)、CD8抗体(BioLegend,克隆SK1)。

[0347] 启动受体诱导

[0348] 为了评估转基因(即合成回路)的功能活性,将经编辑的T细胞与表达启动抗原的靶细胞系以1:1的E:T比共培养,并孵育24小时。收获T细胞并用抗Myc和抗Flag抗体染色,以分别评估启动受体和CAR表达。

[0349] 图3提供了用于展示使用如图1所示的定制转基因盒和非病毒基因编辑技术引入T细胞的TRAC基因座的合成回路的保真度的实验方案的概览。T细胞包括一个与CAR共表达的FLAG标签和一个与启动受体(primeR)共表达的myc标签。制备了表达TRAC基因座中的转基因插入物的工程化免疫细胞,并最初表达了启动受体。接着将工程化免疫细胞与表达启动抗原的K562靶细胞以1:1的E:T比孵育24小时。这些靶细胞表达启动受体的同源配体。收获T细胞并用抗Myc和抗Flag抗体染色,以分别评估启动受体和CAR表达。预期仅在存在启动抗原的情况下实现对经编辑的T细胞的CAR诱导。因此,在与靶细胞充分孵育后,工程化免疫细胞表达CAR。

[0350] 图4提供了实验结果。在Attune NxT流式细胞仪上读出分析结果。如图所示,当T细胞在不存在启动受体配体的情况下单独孵育或与靶细胞K562一起孵育时,CAR表达非常少,但启动受体表达水平很高。这指示启动受体如预期得到特异性表达和激活。这还指示,在不存在启动受体配体的情况下,很少有细胞表达CAR。因此,回路受体无泄漏。当T细胞与具有启动受体同源蛋白($K562^{ALPG}$)的靶细胞一起孵育时,许多T细胞表现出CAR的表达。这指示启动受体的激活诱导CAR的表达。由于γ分泌酶的裂解,启动受体的表达也相应减少,γ分泌酶也释放转录因子。因此,CAR诱导仅在存在启动抗原的情况下观察到,并且与启动受体表达的缺失相关。获得的结果表明,在使用CRISPR技术递送转基因后,回路的保真度得以维持。

[0351] 图5提供了用于展示使用如图1所示的定制转基因盒和非病毒基因编辑技术引入T细胞的合成回路的保真度的实验方案的概览。T细胞包括一个与CAR共表达的FLAG标签和一个与启动受体(primeR)共表达的myc标签。制备了工程化免疫细胞,并最初表达了启动受体。接着将工程化免疫细胞与表达启动抗原的靶细胞以1:1的E:T比孵育24小时。这些靶细胞表达启动受体的同源配体。收获T细胞并用抗Myc和抗Flag抗体染色,以分别评估启动受体和CAR表达。预期仅在存在启动抗原的情况下实现对经编辑的T细胞的CAR诱导。因此,在与靶细胞充分孵育后,工程化免疫细胞表达CAR。

[0352] 图6示出了来自两名供体的进一步定量数据,这些数据表明在不存在启动受体激活的情况下,非常小比例的T细胞表达CAR。使用CRISPR技术将合成回路递送至源自2名健康供体的T细胞。将经编辑的T细胞与表达同源引物的K562细胞系(1:1的E:T比)、阴性对照K562亲本细胞系或作为阳性对照的Myc珠共培养,并孵育24小时。收获T细胞并用抗Myc和抗Flag抗体染色,以分别评估启动受体和CAR表达。在Attune NxT流式细胞仪上读出分析结果。超过30%的启动受体阳性T细胞在存在靶同源抗原或myc珠对照的情况下表达CAR。在基础水平或与K562亲本细胞系一起培养时观察到极低CAR表达。综上所述,结果表明将合成回路敲入TRAC基因座不会损害其功能。这还指示编码CAR的DNA盒无泄漏。

[0353] 实施例2:评价非病毒插入后的大回路表达和功能

[0354] GS94是位于11号染色体远端q臂上的候选整合位点。它在ETS1和FLI1的启动子的180-350kb范围内(图7),但它被认为对于整合载体基因疗法是低风险的。评价了GS94基因

的回路表达和功能潜力。

[0355] 构建体产生

[0356] 为了产生用于敲入的质粒构建体,从Twist、IDT和GENEWIZ订购合成的DNA,并通过Gibson Assembly和Golden Gate Assembly进行组装。质粒含有与基因组中CRISPR靶位点侧翼的序列同源的同源臂,长度为1.2kb或450bp。回路盒长度为4558bp。盒的图解示于图19A中。

[0357] T细胞在GS79(TRAC)整合位点、GS94整合位点和GS102整合位点与PrimeR进行回路盒整合。用于插入的sgRNA序列示于表1中。将细胞与K562细胞共培养48小时,接着将PrimeR诱导的CAR MFI与PrimeR MFI进行比较。

[0358] 将CD3-CD28 Dynabead激活的T细胞用靶向GS94的sgRNA/Cas9 RNP以及引导HDR介导的整合进入所指示位点的具有同源臂的HDRT进行电穿孔。在电穿孔后第7天,将T细胞与表达primeR靶标的K562细胞(K562单表达细胞)共培养。接着在共培养开始后48小时用抗FLAG抗体对细胞进行染色,并在Attune NxT流式细胞仪上进行分析。结果显示,在与K562细胞共培养48小时后,GS94产生了优异的CAR诱导与高prime R表达。参见图8A。GS94导致启动抗原依赖性CAR表达,比其他几个候选整合位点以及TRAC整合位点的表达高大约两倍。此外,平均而言,启动受体表面表达水平不低于使用TRAC整合位点时表达水平的50%。

[0359] 细胞毒性和细胞因子分泌

[0360] 为了评价候选整合位点对细胞毒性和细胞因子分泌的影响,将在GS79(TRAC)整合位点和GS94整合位点与PrimeR进行回路盒整合的T细胞与表达primeR靶标的K562细胞(K562单表达细胞)共培养48小时。以1:1的效应物:靶细胞比(1:1E:T)处理细胞。简而言之,在电穿孔后第7天,将如上文所述产生的T细胞与双重表达primeR靶标和CAR靶标的K562细胞(K562双表达细胞)共培养。共培养开始后48小时,收集上清液并通过Luminex分析细胞因子水平。细胞因子量度是IL-2、INFg和TNF。通过测量48小时后剩余靶细胞的荧光素酶活性来分析细胞毒性。图8B中的每个数据点代表两个重复,线条代表重复的细胞毒性范围。如图8B所示,GS94整合位点在与K562双表达细胞共培养48小时后导致优异的细胞毒能力和细胞因子分泌。

[0361] 引物非依赖性细胞毒性

[0362] 为了比较候选整合位点对细胞毒性与引物非依赖性细胞毒性的影响,在电穿孔后第7天,将在GS79(TRAC)整合位点和GS94整合位点与PrimeR进行回路盒整合的T细胞与K562双表达细胞(“K562 primeR/CAR)或仅表达CAR靶标的K562细胞(K562 CAR)共培养48小时。检测0.3、1.0和3.0的E:T细胞比。参见图9A和图9B。在共培养开始后48小时,通过测量剩余靶细胞的荧光素酶活性来分析细胞毒性。如图9B所示,GS94整合位点产生与TRAC整合位点等同的细胞毒性潜力,并且不存在引物非依赖性细胞毒性。

[0363] 引物非依赖性细胞因子分泌

[0364] 为了比较候选整合位点对细胞毒性与引物非依赖性细胞毒性的影响,将上文所述的在GS79(TRAC)整合位点和GS94整合位点与PrimeR进行回路盒整合的T细胞与K562双表达细胞共培养。将仅具有靶细胞的组(E:T=0;仅靶标)与E:T细胞比为1的组进行比较。在与K562双表达细胞共培养48小时后,收集上清液并通过Luminex分析细胞因子水平以测量IL-2、INFg和TNF细胞因子的分泌。参见图10A和图10B。如图10B所示,GS94整合位点导致与TRAC

整合位点等同的细胞因子分泌,并且不存在引物非依赖性的IL-2、INFg或TNF分泌。

[0365] 引物非依赖性CAR表达

[0366] 为了评价候选整合位点对引物非依赖性CAR表达的影响,将在GS79(TRAC)整合位点、GS94整合位点和GS102整合位点与PrimeR进行回路盒整合的T细胞体外培养32天。在实验的第5、12、19和28天,对细胞进行重复CD3/CD28刺激处理。在第16天,使用流式细胞术测定评价细胞的CAR表达。如图11所示,通过TCR激活T细胞没有导致候选整合位点处回路盒整合的primeR非依赖性CAR表达。

[0367] 将如上文所述产生的T细胞在96孔板中培养,每2天更换一次T细胞生长培养基。在第5、12、19和28天,用1:1CD3/CD28 Dynabead刺激T细胞。通过myc表位标签染色分析细胞的PrimeR表达,并在所指示时间点通过FLAG表位标签染色分析CAR表达。在Attune NxT流式细胞仪上进行流式分析。

[0368] 实施例3:评价数周内来自大插入物的启动受体表达的稳定性

[0369] 为了评价候选整合位点对PrimeR稳定(持续)表达的影响,将在图12A所指示的整合位点与PrimeR进行回路盒整合的T细胞体外培养32天。简单来说,将如上文所述产生的T细胞在96孔板中培养,每2天更换一次T细胞生长培养基。在第5、12、19和28天,用1:1CD3/CD28 Dynabead重复刺激刺激T细胞。在第16天和第32天使用Attune NxT流式细胞仪进行流式细胞术测定。通过myc表位标签染色分析细胞的PrimeR表达。如图12A和图12B所示,GS94整合位点导致至少4周时间内稳定的PrimeR表达。

[0370] 实施例4:评价使用非病毒插入的中靶编辑效率

[0371] 为了评价候选敲入位点的中靶编辑效率,使用iGUIDE-Seq测定。用于进行iGUIDE-Seq测定的方法示于图13A并提供于Nobles等人,Genome Biology(2019)中,其据此以引用的方式整体并入本文。如图13B所示,GS94整合位点在评价的候选整合位点中具有最高的中靶编辑效率。如图13C所示,GS94没有导致如在两名供体中观察到的推定的脱靶编辑。

[0372] 实施例5:评价大插入物的各种GSH非病毒敲入

[0373] 方法

[0374] Elevation预测:使用Elevation-search(Listgarten等人2018.Prediction of off-target activities for the end-to-end design of CRISPR guide RNAs.Nat Biomed Engr 2,37-48中描述的算法;获自https://github.com/Microsoft/Elevation的软件)对来自(gs94)的潜在脱靶位点进行计算预测。使用rhAmp-seq对通过Elevation-search确定的所有位点进行分析。

[0375] rhAmpSeq:49个通过iGUIDE或Elevation预测算法确定的GS94候选脱靶位点和GS94靶位点通过rhAmpSeq(Integrated DNA Technologies,Inc.)来表征。这种有针对性的扩增能够对同时发生在多个位点的编辑进行基于NGS的定量。用GenFind V3 DNA纯化系统(Beckman Coulter)分离至少2名供体的T细胞的基因组DNA,这些供体分别用以下7种指导物中的每一种单独处理过:GS84、GS94、GS95、GS96、GS102、GS108和GS138(表1)。使用两个单独的rhAmpSeq扩增池来覆盖50个基因座,按照Integrated DNA Technologies的建议对每个样本进行该程序。在MiniSeq上用Mid Output Kit(300个周期)(Illumina)对rhAmpSeq文库进行测序。使用CRISPResso2算法(https://github.com/pinellolab/CRISPResso2)用于确定每个扩增基因座的插入和缺失百分比。使用卡方检验仅在GS94位点观察到统计显著性

(FDR调整后的p值≤0.001)。

[0376] RNA-seq:为了评价GS94整合在转录水平上诱导的变化,在GS79(TRAC)、GS94和GS102整合位点处整合了primeR/CAR回路。整合后第6天,使用BD FACSAria基于转基因表达对1e6个经编辑的细胞进行分类。使用RNeasy试剂盒(Qiagen)从分选的T细胞中分离出RNA。使用TruSeq RNA文库制备试剂盒v2(Illumina)将纯化的RNA转化为NGS文库。在NovaSeq6000或NextSeq 550仪器(Illumina)上对文库进行测序。使用STAR 2.7.3a比对器(Dobin A.等人Bioinformatics.2013.29:15-21)将RNA-seq数据与参考人GRCh38转录组进行比对,并获得基因水平的读计数。使用edgeR(Robinson MD等人Bioinformatics.2010.26:139-140)计算差异表达,结合两名供体的数据。与在任何其他两个基因座处整合的细胞相比,GS94位点300Kb内仅有的基因ETS1和FLI1于在GS94整合位点处整合的细胞中无差异性表达。在FDR调整后的p值截断值为0.01时,差异性表达的基因的数量最少(全基因组<100个基因)。

[0377] 细胞因子非依赖性生长测定:为了评价具有GS94基因座KI的原代T细胞的安全性,进行了细胞因子非依赖性生长测定以评价致癌转化的可能性。简单来说,将在GS94基因座处经历primeR/CAR回路盒整合的原代人T细胞解冻并恢复过夜。接着将$1x10^6$个细胞接种到24孔GRex板的一个孔中,在含或不含细胞因子的培养基中培养5天。在第0、3和5天记录细胞数量和活力。作为阳性对照,在不含细胞因子的培养基中平行培养$1x10^6$个Jurkat细胞。如图17所示,GS94 KI T细胞在有细胞因子的情况下培养时保持良好的活力和总细胞计数,而当在没有细胞因子的情况下培养时,GS94 KI T细胞的活力在5天内急剧下降,并且在第5天已无活细胞。在整个测定过程中,阳性对照Jurkat细胞在没有细胞因子的情况下保持良好的活力和扩增。综上所述,该数据表明GS94编辑的原代人T细胞的生长、存活和扩增仍然依赖于外源细胞因子,因此不存在细胞转化的问题。

[0378] 结果

[0379] 通过iGUIDE-seq评价靶向候选基因座(包括GS94)的CRISPR试剂(例如,与sgRNA复合的SpCas9)的特异性(图14)。靶向GS94的CRISPR RNP在评价的所有候选者中显示出最高百分比的iGUIDE-seq寡核苷酸盒捕获事件,并且iGUIDE-seq论文中的对照sgRNA序列显示出与原始出版物中报告的相似的特异性,这表明测定结果符合预期。

[0380] 推定的脱靶位点取自iGUIDE-seq输出,其已表明推定的位点是虚假的。通过计算方法(Elevation软件包)预测另外的靶站点。使用rhAmp-seq为每个推定的脱靶位点制备高通量测序文库,并将该方法应用于使用靶向候选靶位点的CRISPR RNP进行电穿孔的T细胞的DNA样本。将所得NGS数据用CRISPResso2软件进行处理,插入和缺失(插入缺失)的频率视为CRISPR裂解活性的指标,这在该领域很常见。与靶向其他位点的CRISPR RNP处理的T细胞相比,用靶向GS94的CRISPR RNP电穿孔的T细胞在一组推定脱靶位点上显示的插入缺失频率并没有更高,这与靶向GS94的CRISPR RNP没有相应的或可检测到的脱靶活性一致,因此是评价的一组中最具特异性的(图15)。

[0381] GS94位点处的转基因整合对T细胞转录组调控的潜在影响是通过将大盒敲入该位点,使T细胞生长数天,对表达该盒内转基因的细胞进行分类,接着从细胞中收集RNA来评价的。制备RNA-seq文库并对其进行测序,对所得Illumina测序数据的分析显示,与TRAC或GS102位点整合的细胞相比,GS94位点整合的细胞中GS94位点300kb内的任何基因的表达在

生物学上或统计学上无显著差异(图16)。此外,达到统计显著性的其他基因表达差异在数量和效应量上都是最小的,这与它们在比较中是噪声一致。

[0382] 为了评估GS94处的转基因整合是否赋予转化的表型,将在GS94位点处整合的细胞在有和没有细胞因子的情况下进行体外培养。细胞在添加细胞因子的情况下保持存活和活力,但在没有细胞因子补充的情况下死亡并失去其活力(图17)。阳性对照Jurkat细胞保持活力并增殖。总的来说,这指示在GS94处的转基因整合不赋予细胞因子非依赖性生长的能力,这是T细胞转化的一个标志。

[0383] 实施例6:评价GS94中8.3kb大表达盒的非病毒插入

[0384] 接下来,使用如先前在实施例1中描述的材料/方法将8.3kb插入物插入T细胞中的GS94安全港基因座中。插入物中增加的长度是盒中另外的蛋白质编码序列所致。盒的图解提供于图19B中。

[0385] 构建体产生

[0386] 为了产生用于敲入的质粒构建体,从Twist、IDT和GENEWIZ订购合成的DNA,并通过Gibson Assembly和Golden Gate Assembly进行组装。质粒含有与基因组中CRISPR靶位点侧翼的序列同源的同源臂,长度为1.2kb或450bp。

[0387] T细胞工程化

[0388] 使用Lymphoprep(STEMCELL Technologies)和EasySep人T细胞分离试剂盒(STEMCELL Technologies)从获自正常供体Leukopaks(STEMCELL Technologies)获得的外周血单核细胞(PBMC)富集T细胞。随后将T细胞在补充有3%人AB血清(Gemini Bio)和12.5ng/ml人IL-7和IL-15(Miltenyi优质级)的TexMACS培养基(Miltenyi 130-197-196)中用CD3/CD28 Dynabead以1:1珠细胞比(ThermoFisher,40203D)激活,在37℃、5% CO2下培养48小时,然后电穿孔。

[0389] 通过将120μM靶向DNA序列的sgRNA(Synthego)GAGCCATGCTTGGCTTACGA(GS94,SEQ ID NO:94)、62.5μM sNLS-SpCas9-sNLS(Aldevron)和P3缓冲液(Lonza)以5:1:3:6的体积比组合来制备CRISPR RNP,并将其在室温下孵育15分钟。将通过剂量滴定实验确定的优化量的质粒DNA(范围为0.5-3微克)与3.5μl RNP混合。对T细胞进行计数、去珠、以90X G离心10分钟,并以10^6个细胞/14.5μl添加有补充剂的P3(Lonza)重悬。将14.5μl T细胞悬浮液添加到DNA/RNP混合物中,转移到Lonza 384孔核比色皿条中,并在代码为EH-115的Lonza HT Nucleofector System中进行脉冲处理。使细胞在室温下静置15分钟,接着转移至96孔板(Sarstedt)中的补充有12.5ng/ml人IL-7和IL-15(Miltenyi优质级)的TexMACS培养基中。

[0390] 通过用抗Myc抗体(Cell Signaling Technology,克隆9B11)和抗Flag抗体(RnD systems,克隆1042E)染色检测转基因表达,并在Attune NxT流式细胞仪上进行分析。所使用的其他抗体是活/死Fixable Near-IR(Thermo Fisher)、TCRα/β抗体(BioLegend,克隆IP26)、CD4抗体(BioLegend,克隆RPA-T4)、CD8抗体(BioLegend,克隆SK1)。

[0391] 启动受体诱导

[0392] 为了评估转基因(即合成回路)的功能活性,将经编辑的T细胞与表达启动抗原的靶细胞系以1:1的E:T比共培养,并孵育24小时。收获T细胞并用抗Myc和抗Flag抗体染色,以分别评估启动受体和CAR表达。

[0393] 为了评估GS94处的8.3kb转基因整合是否导致功能性敲入,将细胞与亲本K562细

胞和表达同源启动抗原的K562细胞以1:1的E:T细胞比培养。如前所述,48小时后通过流式细胞术测定细胞。带有启动抗原的K562细胞诱导CAR表达,而对照亲本K562细胞则没有诱导CAR表达(图18)。总的来说,这指示在插入8.3kb转基因回路后,PrimeR诱导CAR表达。

[0394] 实施例7:体内使用包含CAR表达盒的T细胞

[0395] 针对经工程化表达CAR抗原或表达启动受体和CAR两者的抗原的人肿瘤细胞(诸如K562),对具有表达识别肿瘤抗原的CAR、或在识别肿瘤解剖结构附近的抗原的启动受体控制下表达识别肿瘤抗原的CAR的转基因盒的T细胞的体内功效进行评估。将肿瘤细胞(例如1e6个)皮下注射到NSG小鼠(Jackson Laboratories)的侧腹。每2-4天用卡尺通过尺寸测量评估肿瘤生长。当肿瘤体积达到约100立方毫米时,向小鼠静脉内注射:5e6个通过CRISPR介导的插入在特定位点整合CAR或primeR-CAR回路盒的T细胞,或经单独的CRISPR RNP工程化的T细胞,或作为假注射的单独的PBS。监测肿瘤生长,当肿瘤体积达到2000立方毫米时对小鼠实施安乐死。通过眶后手术从小鼠抽取外周血,并使用流式细胞术和/或ddPCR观察工程化T细胞随时间的扩增。在安乐死时,通过流式细胞术、ddPCR和/或免疫组织化学分析脾脏、血液、肿瘤和/或其他组织中工程化T细胞的存在。结果表明,与未进行盒整合的T细胞相比,在限定的基因组位点之一进行盒整合的工程化T细胞导致注射小鼠的肿瘤消退和清除,并且在注射小鼠的外周血和组织中可检测到工程化T细胞。

[0396] 表1:sgRNA序列

[0397]

sgRNA ID	sgRNA 序列	sgRNA 起始坐标 GRCH38	sgRNA 靶基因座	整合位点	中位数(修饰%)，总结自 2 名供体，2 个引物组
sgRNA_1	GCACCTGAATACCACGCCTG (SEQ ID NO:1)	chr16:88811818	APRT	APRT	79.28
sgRNA_2	CGCCTGCGATGTAGTCGATG (SEQ ID NO:2)	chr16:88811551	APRT	APRT	78.60
sgRNA_3	CAGGACGGGCGAGATGTCCC (SEQ ID NO:3)	chr16:88811640	APRT	APRT	85.25
sgRNA_4	CTGAATCTTTGGAGTACCTG (SEQ ID NO:4)	chr15:44715425	B2M	B2M	78.51
sgRNA_5	GGCCACGGAGCGAGACATCT (SEQ ID NO:5)	chr15:44711550	B2M	B2M	94.75
sgRNA_6	AAGTCAACTTCAATGTCGGA (SEQ ID NO:6)	chr15:44715515	B2M	B2M	70.97
sgRNA_7	GCTTGGAGGCCTGATCAGCG (SEQ ID NO:7)	chr19:36141111	CAPNS1	CAPNS1	89.34
sgRNA_8	CTTATCTCTTCGCAGCGAGG (SEQ ID NO:8)	chr19:36142301	CAPNS1	CAPNS1	91.09
sgRNA_9	CACACATTACTCCAACATTG (SEQ ID NO:9)	chr19:36142676	CAPNS1	CAPNS1	71.98
sgRNA_10	TTCCGCAAAATAGAGCCCCA (SEQ ID NO:10)	chr3:105746019	CBLB	CBLB	91.55
sgRNA_11	TGCACAGAACTATCGTACCA (SEQ ID NO:11)	chr3:105751622	CBLB	CBLB	91.43
sgRNA_12	GCAATAAGACTCTTTAAAGA (SEQ ID NO:12)	chr3:105853470	CBLB	CBLB	76.18

[0398]

sgRNA_13	CAAAGAGATTACGAATGCCT (SEQ ID NO:13)	chr1:116754658	CD2	CD2	89.80
sgRNA_14	CAAGGCACCCCAGGTTTCCA (SEQ ID NO:14)	chr1:116754663	CD2	CD2	92.70
sgRNA_15	TTACGAATGCCTTGGAAACC (SEQ ID NO:15)	chr1:116754666	CD2	CD2	92.82
sgRNA_16	CAGAGACGCATCTGACCCTC (SEQ ID NO:16)	chr11:118315540	CD3E	CD3E	90.96
sgRNA_17	CATGCAGTTCTCACACACTG (SEQ ID NO:17)	chr11:118313715	CD3E	CD3E	87.47
sgRNA_18	GTGTGAGAACTGCATGGAGA (SEQ ID NO:18)	chr11:118313715	CD3E	CD3E	86.65
sgRNA_19	TCTCATTTCAGGAAACCACT (SEQ ID NO:19)	chr11:118349748	CD3G	CD3G	87.24
sgRNA_20	AGTCATACACCTTAACCAAG (SEQ ID NO:20)	chr11:118349754	CD3G	CD3G	87.99
sgRNA_21	TTCAAGGAAACCAGTTGAGG (SEQ ID NO:21)	chr11:118352458	CD3G	CD3G	86.55
sgRNA_22	GAGCCTTGCCTGGAAATCTG (SEQ ID NO:22)	chr11:61118177	CD5	CD5	84.03
sgRNA_23	AAGCGTCAAAAGTCTGCCAG (SEQ ID NO:23)	chr11:61118324	CD5	CD5	89.19
sgRNA_24	CGTTCCAACTCGAAGTGCCA (SEQ ID NO: 24)	chr11:61118121	CD5	CD5	83.11
sgRNA_25	GAGCGACTGGGACACGGTGA (SEQ ID NO:25)	chr9:136866246	EDF1	EDF1	88.84
sgRNA_26	GCTGCGCAAGAAGGGCCCTA (SEQ ID NO:26)	chr9:136866211	EDF1	EDF1	91.04
sgRNA_27	TTGTTCTGGCCAGCAGCCCC (SEQ ID NO: 27)	chr9:136863433	EDF1	EDF1	85.98
sgRNA_28	CTTCCAGAGCCACATCATCG (SEQ ID NO:28)	chr19:48965791	FTL	FTL	93.10
sgRNA_29	GGGACTCACCAGAGAGAGGT (SEQ ID	chr19:48965601	FTL	FTL	88.86

[0399]

	NO:29)				
sgRNA_30	CGGTCGAAATAGAAGCCCTA (SEQ ID NO:30)	chr19:48965770	FTL	FTL	93.14
sgRNA_31	AAAAGGATATTGTGCAACTG (SEQ ID NO:31)	chr10:87933015	PTEN	PTEN	92.37
sgRNA_32	TGTGCATATTTATTACATCG (SEQ ID NO:32)	chr10:87933183	PTEN	PTEN	90.64
sgRNA_33	TTTGTGAAGATCTTGACCAA (SEQ ID NO:33)	chr10:87933087	PTEN	PTEN	85.36
sgRNA_34	TGTCATGCTGAACCGCATTG (SEQ ID NO:34)	chr18:12830972	PTPN2	PTPN2	87.94
sgRNA_35	CCACTCTATGAGGATAGTCA (SEQ ID NO:35)	chr18:12859219	PTPN2	PTPN2	92.45
sgRNA_36	TTGACATAGAAGAGGCACAA (SEQ ID NO:36)	chr18:12836828	PTPN2	PTPN2	93.96
sgRNA_37	GAGTACTACACTCAGCAGCA (SEQ ID NO:37)	chr12:6952098	PTPN6	PTPN6	89.61
sgRNA_38	TCACGCACAAGAAACGTCCA (SEQ ID NO:38)	chr12:6954872	PTPN6	PTPN6	82.74
sgRNA_39	AGGTCTCGGTGAAACCACCT (SEQ ID NO:39)	chr12:6951610	PTPN6	PTPN6	91.27
sgRNA_40	AGCATTATCCAAAGAGTCCG (SEQ ID NO:40)	chr1:198696873	PTPRC	PTPRC	88.88
sgRNA_41	ATATTAATTCTTACCAGTGG (SEQ ID NO:41)	chr1:198692370	PTPRC	PTPRC	88.95
sgRNA_42	AGCTTTAAATCAAGGTTCAT (SEQ ID NO:42)	chr1:198756176	PTPRC	PTPRC	96.89
sgRNA_43	ATCCCGAGCCCTAAGGTGCA (SEQ ID NO:43)	chr11:67436325	PTPRCAP	PTPRCAP	84.08
sgRNA_44	GGCAGCGCGGAGGACAGCGT (SEQ ID NO:44)	chr11:67436285	PTPRCAP	PTPRCAP	97.74
sgRNA_45	CTCAGGGGGCTACTACCACC (SEQ ID NO:45)	chr11:67436170	PTPRCAP	PTPRCAP	91.50
sgRN	GTCACCGACGAGAC	chr5:8227	RPS23	RPS2	79.40

[0400]

A_46	CAGAAG (SEQ ID NO:46)	7810		3	
sgRNA_47	GTCGTGGACTTCGTACTGCT (SEQ ID NO:47)	chr5:82277843	RPS23	RPS23	83.07
sgRNA_48	TAATTTTTAGGCAAGTGTCG (SEQ ID NO:48)	chr5:82277860	RPS23	RPS23	61.94
sgRNA_49	TTAGCTGTTAGACTTGAATA (SEQ ID NO:49)	chr14:51993810	RTRAF	RTRAF	85.50
sgRNA_50	CGAGAGCCGTCAACTTGCGT (SEQ ID NO:50)	chr14:51989652	RTRAF	RTRAF	85.64
sgRNA_51	CGGCTTCAACTGCAAAGGTG (SEQ ID NO:51)	chr14:51989700	RTRAF	RTRAF	88.77
sgRNA_52	TATGAAAAAGCAGAGCGACT (SEQ ID NO:52)	chr15:43793025	SERF2	SERF2	89.61
sgRNA_53	TCTGGCGGGCGAGCTCACGC (SEQ ID NO:53)	chr15:43792989	SERF2	SERF2	86.73
sgRNA_54	CTCACGCTGGTTACCGCCTA (SEQ ID NO:54)	chr15:43792977	SERF2	SERF2	80.57
sgRNA_55	AAAGATTACGAACTTCCCTG (SEQ ID NO:55)	chr12:46207559	SLC38A1	SLC38A1	92.24
sgRNA_56	GTTAAAAACAGACATGCCTA (SEQ ID NO:56)	chr12:46229232	SLC38A1	SLC38A1	91.51
sgRNA_57	ATGCCTAAGGAGGTTGTACC (SEQ ID NO:57)	chr12:46229246	SLC38A1	SLC38A1	79.48
sgRNA_58	CTCCAGGTATCCCATCGAAA (SEQ ID NO:58)	chr18:47869418	SMAD2	SMAD2	79.53
sgRNA_59	CACCAAATACGATAGATCAG (SEQ ID NO:59)	chr18:47870532	SMAD2	SMAD2	86.61
sgRNA_60	TGGCGGCGTGAATGGCAAGA (SEQ ID NO:60)	chr18:47896729	SMAD2	SMAD2	82.91
sgRNA_61	TAGGATGGTAGCACACAACC (SEQ ID NO:61)	chr16:11255478	SOCS1	SOCS1	92.25
sgRNA_62	CAGCAGCAGAGCCCCGACGG (SEQ ID NO:62)	chr16:11255432	SOCS1	SOCS1	83.79

[0401]

sgRNA_63	CGGCGTGCGAACGGAATGTG (SEQ ID NO:63)	chr16:11255296	SOCS1	SOCS1	84.24
sgRNA_64	TATAGACGCTGCCCGACGTC (SEQ ID NO:64)	chr15:40038895	SRP14	SRP14	95.12
sgRNA_65	TCCAAAGAAGGGTACTGTGG (SEQ ID NO:65)	chr15:40038368	SRP14	SRP14	92.14
sgRNA_66	ACAGTACCCTTCTTTGGAAT (SEQ ID NO:66)	chr15:40038358	SRP14	SRP14	65.82
sgRNA_67	GCGACGGGCGCATCTACGTG (SEQ ID NO:67)	chr12:120469572	SRSF9	SRSF9	83.68
sgRNA_68	CCCGACCTCCATAAGTCCTG (SEQ ID NO:68)	chr12:120465700	SRSF9	SRSF9	92.56
sgRNA_69	GGGGTCCTCGAAGCGCACGA (SEQ ID NO:69)	chr12:120469426	SRSF9	SRSF9	89.94
sgRNA_70	TGCTCTGTTTAGAAGATGAC (SEQ ID NO:70)	chr5:32591641	SUB1	SUB1	79.36
sgRNA_71	ATATTCTTTTCTAGTTAAAG (SEQ ID NO:71)	chr5:32591566	SUB1	SUB1	70.93
sgRNA_72	CCTGTAAAGAAACAAAAGAC (SEQ ID NO:72)	chr5:32591614	SUB1	SUB1	93.66
sgRNA_73	TGGAGAAAGACGTAACTTCG (SEQ ID NO:73)	chr4:105234315	TET2	TET2	83.53
sgRNA_74	TCTGCCCTGAGGTATGCGAT (SEQ ID NO:74)	chr4:105234747	TET2	TET2	90.97
sgRNA_75	ATTCCGCTTGGTGAAAACGA (SEQ ID NO:75)	chr4:105235656	TET2	TET2	89.62
sgRNA_76	CAGGCACAATAGAAACAACG (SEQ ID NO:76)	chr3:114295571	TIGIT	TIGIT	92.65
sgRNA_77	CCATTTGTAATGCTGACTTG (SEQ ID NO:77)	chr3:114295700	TIGIT	TIGIT	60.75
sgRNA_78	CTGGGTCACTTGTGCCGTGG (SEQ ID NO:78)	chr3:114295634	TIGIT	TIGIT	87.99
sgRNA_79	GTCAGGGTTCTGGATATCTG (SEQ ID	chr14:22547508	TRAC	TRAC	98.20

[0402]

	NO:79)				
sgRNA_80	TGGATTTAGAGTCTCTCAGC (SEQ ID NO:80)	chr14:22547541	TRAC	TRAC	88.15
sgRNA_81	CTGCGGCTGTGGTCCAGCTG (SEQ ID NO:81)	chr14:22550661	TRAC	TRAC	94.77
sgRNA_82	ACAAAACTGTGCTAGACATG (SEQ ID NO:82)	chr14:22547658	TRAC	TRAC	87.86
sgRNA_83	TTCTTCCCCAGCCCAGGTAA (SEQ ID NO:83)	chr14:22547778	TRAC	TRAC	89.85
sgRNA_84	CGTCATGAGCAGATTAAACC (SEQ ID NO:84)	chr14:22550625	TRAC	TRAC	95.81
sgRNA_85	GAGAGCGCCTGCGACCCGAG (SEQ ID NO:85)	chr19:58544980	TRIM28	TRIM28	89.44
sgRNA_86	CCAGCGGGTGAAGTACACCA (SEQ ID NO:86)	chr19:58544869	TRIM28	TRIM28	94.79
sgRNA_87	GGAGCGCTTTTCGCCGCCAG (SEQ ID NO:87)	chr19:58544839	TRIM28	TRIM28	91.81
sgRNA_88	TGAGGCCTGGACCTTATGCA (SEQ ID NO:88)	chr10:33134193	chr10:33130000-33140000	desert_1 (GS88)	69.44
sgRNA_89	CCTGGTGGAGTGAACCATGA (SEQ ID NO:89)	chr10:33132917	chr10:33130000-33140000	desert_1 (GS89)	95.25
sgRNA_90	CAAGCACTTAGGTTCCCCTG (SEQ ID NO:90)	chr10:33134633	chr10:33130000-33140000	desert_1 (GS90)	91.13
sgRNA_91	GGTCTCCCTACAATTCAGCG (SEQ ID NO:91)	chr10:72294568	chr10:72290000-72300000	desert_2 (GS91)	92.02
sgRNA_92	CACAGCGCGTGACTGCAATG (SEQ ID NO:92)	chr10:72298268	chr10:72290000-72300000	desert_2 (GS92)	90.22
sgRNA_93	TCTGGGGCACCAATTCTAGG (SEQ ID NO:93)	chr10:72292786	chr10:72290000-72300000	desert_2 (GS93)	86.35
sgRN	GAGCCATGCTTGGC	chr11:128	chr11:128	desert	91.24

[0403]

A_94	TTACGA (SEQ ID NO:94)	342576	340000-128350000	_3 (GS94)	
sgRNA_95	GTACAAGTACTTATCTCATG (SEQ ID NO:95)	chr11:128343592	chr11:128340000-128350000	desert_3 (GS95)	89.02
sgRNA_96	GAGATAACAACATAACAACA (SEQ ID NO:96)	chr11:128347170	chr11:128340000-128350000	desert_3 (GS96)	96.47
sgRNA_97	CATATTCCATAGTCTTTGGG (SEQ ID NO:97)	chr11:65425000	chr11:65425000-65427000 (NEAT1)	desert_4 (GS97)	88.54
sgRNA_98	CTGCCCCTTAGCAACTTAGG (SEQ ID NO:98)	chr11:65425507	chr11:65425000-65427000 (NEAT1)	desert_4 (GS98)	92.76
sgRNA_99	TGTTTAAAAATATGTTGACA (SEQ ID NO:99)	chr11:65426264	chr11:65425000-65427000 (NEAT1)	desert_4 (GS99)	90.76
sgRNA_100	CCAGGAATGGAAACTCACGC (SEQ ID NO:100)	chr15:92830315	chr15:92830000-92840000	desert_5 (GS100)	87.84
sgRNA_101	GAGGCCGCTGAATTAACCCG (SEQ ID NO:101)	chr15:92831850	chr15:92830000-92840000	desert_5 (GS101)	85.32
sgRNA_102	ATACACGCACACTTGCAGAA (SEQ ID NO:102)	chr15:92831131	chr15:92830000-92840000	desert_5 (GS102)	99.92
sgRNA_103	GAGCAGACAGAAACCCAGGG (SEQ ID NO:103)	chr16:11225670	chr16:11220000-11230000	desert_6 (GS103)	87.92
sgRNA_104	TGAGTCTCCAAACAGAACAG (SEQ ID NO:104)	chr16:11226284	chr16:11220000-11230000	desert_6 (GS104)	88.53
sgRNA_105	TAATATCACTGACTTCACGG (SEQ ID NO:105)	chr16:11225029	chr16:11220000-11230000	desert_6 (GS105)	87.65
sgRNA_106	TACACACAATGTAAGCAGCA (SEQ ID NO:106)	chr2:87467461	chr2:87460000-87470000	desert_7 (GS1	71.79

[0404]

				06)	
sgRNA_107	GGGAGCTCAATTCGAAACCA (SEQ ID NO:107)	chr2:87468809	chr2:87460000-87470000	desert_7 (GS107)	65.89
sgRNA_108	TTGGACAGGTGAGACAGTCG (SEQ ID NO:108)	chr2:87467001	chr2:87460000-87470000	desert_7 (GS108)	72.64
sgRNA_109	AAGCTCACTCAGATAGTGTG (SEQ ID NO:109)	chr3:186511316	chr3:186510000-186520000	desert_8 (GS109)	76.89
sgRNA_110	CAGGAGAACCACCTTACACG (SEQ ID NO:110)	chr3:186515260	chr3:186510000-186520000	desert_8 (GS110)	86.31
sgRNA_111	GGACAGACCCTGATTCACAA (SEQ ID NO:111)	chr3:186519655	chr3:186510000-186520000	desert_8 (GS111)	85.47
sgRNA_112	ACATGGCAGTCTATGAACAG (SEQ ID NO:112)	chr3:594511154	chr3:594500000-59460000	desert_9 (GS112)	87.77
sgRNA_113	CCTATAGAGAGTACTACTTG (SEQ ID NO:113)	chr3:594516416	chr3:594500000-59460000	desert_9 (GS113)	79.33
sgRNA_114	CCAACCGGGTCTTCATTACG (SEQ ID NO:114)	chr3:594517029	chr3:594500000-59460000	desert_9 (GS114)	92.21
sgRNA_115	TCAAGCGTAGAGTTCCGAGT (SEQ ID NO:115)	chr8:127993006	chr8:127980000-128000000	desert_10 (GS115)	93.07
sgRNA_116	TCATGCAATTATGGACCCAG (SEQ ID NO:116)	chr8:127994663	chr8:127980000-128000000	desert_10 (GS116)	89.40
sgRNA_117	CGGGAAAGTGACTGGCCATG (SEQ ID NO:117)	chr8:127996766	chr8:127980000-128000000	desert_10 (GS117)	87.45
sgRNA_118	TGAGATTGAAATCAAATCGG (SEQ ID NO:118)	chr9:7974159	chr9:7970000-7980000	desert_11 (GS118)	84.84
sgRN	TATGCAATATTCAT	chr9:7977	chr9:7970	desert	85.44

[0405]

A_119	CACGCG (SEQ ID NO:119)	914	000-7980 000	_11 (GS1 19)	
sgRN A_120	AATGTGTTAAATCA AATGCA (SEQ ID NO:120)	chr9:7976 895	chr9:7970 000-7980 000	desert _11 (GS1 20)	83.48

[0406] 尽管本发明已参考优选实施方案和各种替代性实施方案进行特定显示和描述，但相关领域技术人员将了解可在不脱离本发明的精神和范围下在其中进行各种形式和细节变化。

[0407] 在本说明书的主体内引用的所有参考文献、授权专利和专利申请都据此出于所有目的以引用的方式整体并入本文。

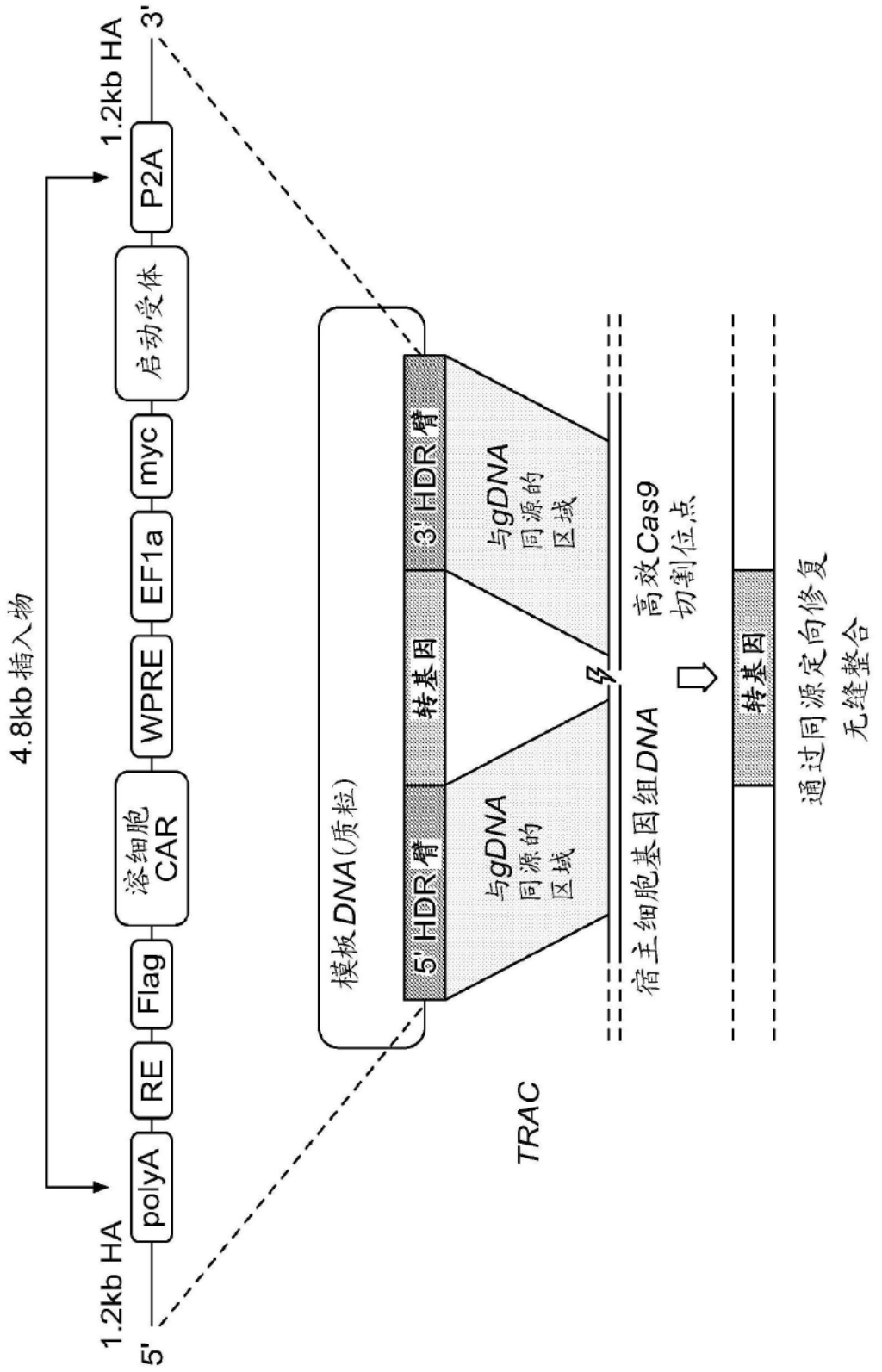

1.2kb HA
5'
polyA
RE
Flag
溶细胞 CAR
WPRE
EF1a
myc
启动受体
P2A
1.2kb HA
3'
4.8kb 插入物
模板 DNA(质粒)
5' HDR 臂
转基因
3' HDR 臂
与gDNA 同源的 区域
与gDNA 同源的 区域
TRAC
宿主细胞基因组DNA
高效Cas9 切割位点
转基因
通过同源定向修复 无缝整合

图1

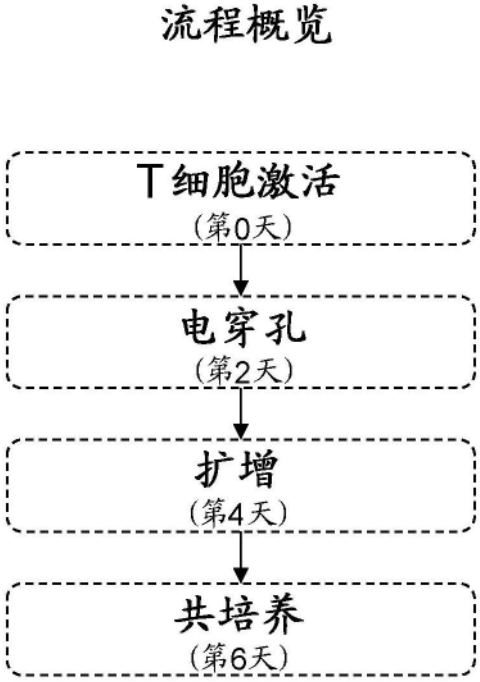

流程概览
T细胞激活
(第0天)
电穿孔
(第2天)
扩增
(第4天)
共培养
(第6天)

图2

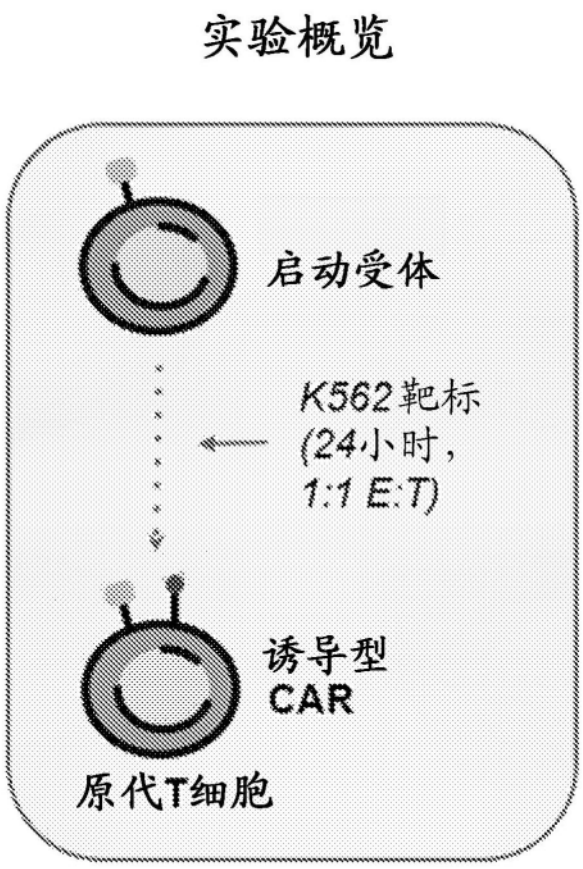

实验概览
启动受体
K562靶标
(24小时,
1:1 E:T)
诱导型
CAR
原代T细胞

图3

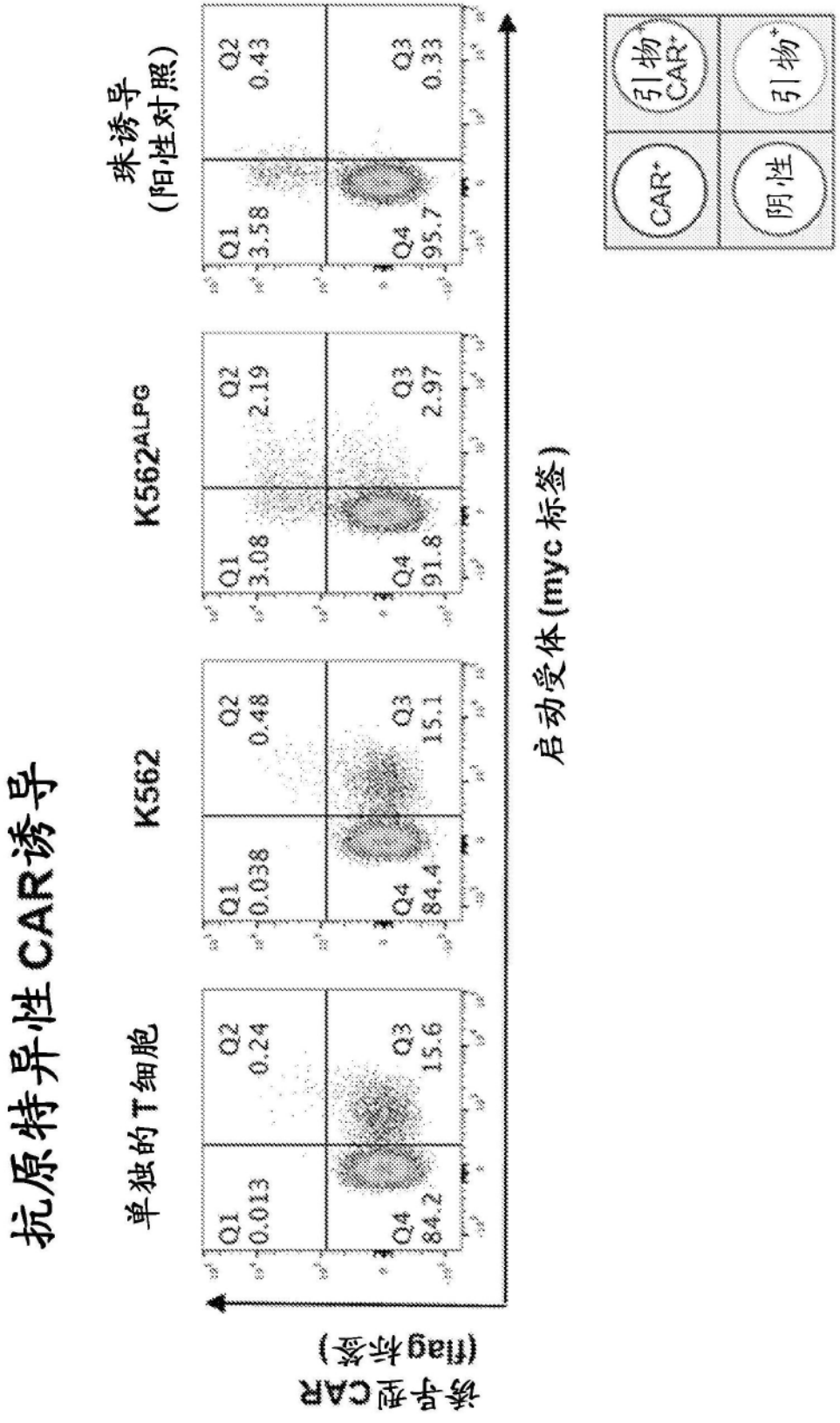

抗原特异性CAR诱导
单独的T细胞
K562
K562ALPG
珠诱导
(阳性对照)
Q1 0.013
Q2 0.24
Q4 84.2
Q3 15.6
Q1 0.038
Q2 0.48
Q4 84.4
Q3 15.1
Q1 3.08
Q2 2.19
Q4 91.8
Q3 2.97
Q1 3.58
Q2 0.43
Q4 95.7
Q3 0.33
诱导型CAR (flag标签)
启动受体(myc标签)
CAR+
引物+ CAR+
阴性
引物+

图4

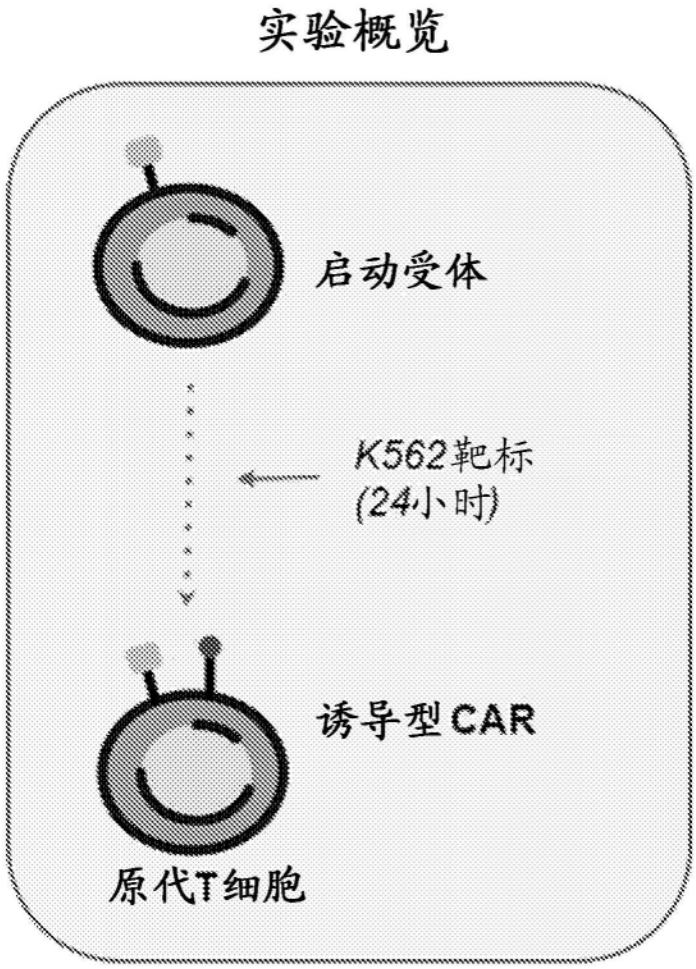


图5

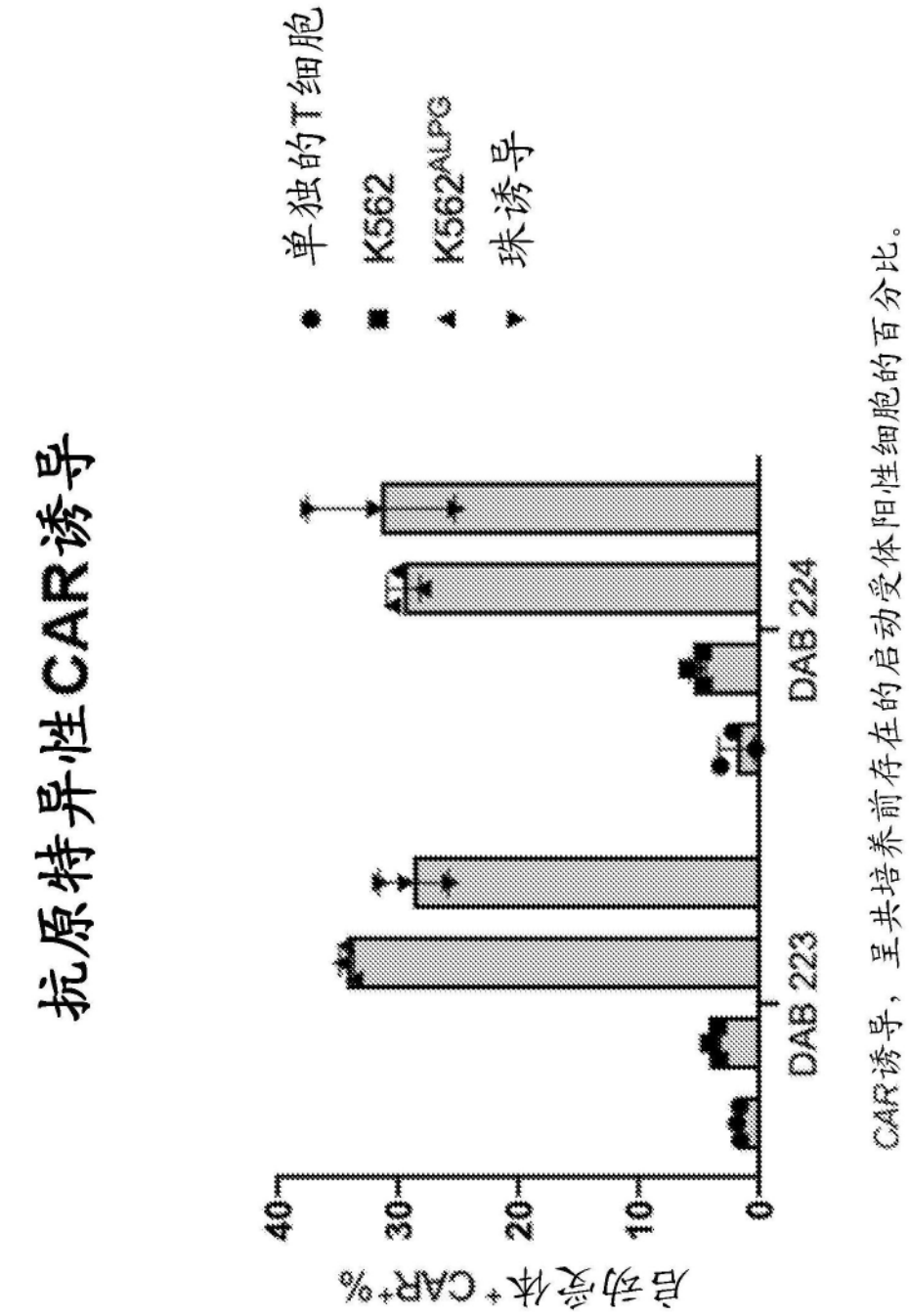

抗原特异性CAR诱导
单独的T细胞
K562
K562ALPG
珠诱导
启动受体+ CAR+%
40
30
20
10
0
DAB 223
DAB 224
CAR诱导，呈共培养前存在的启动受体阳性细胞的百分比。

图6

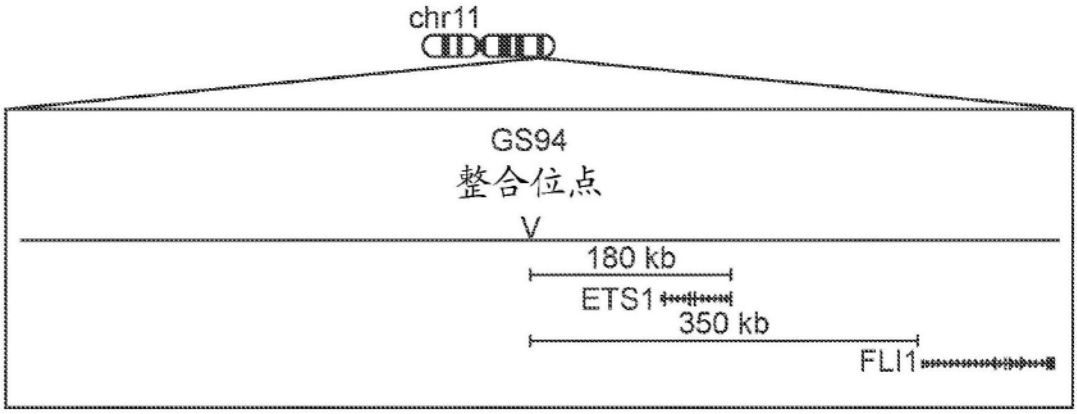


图7

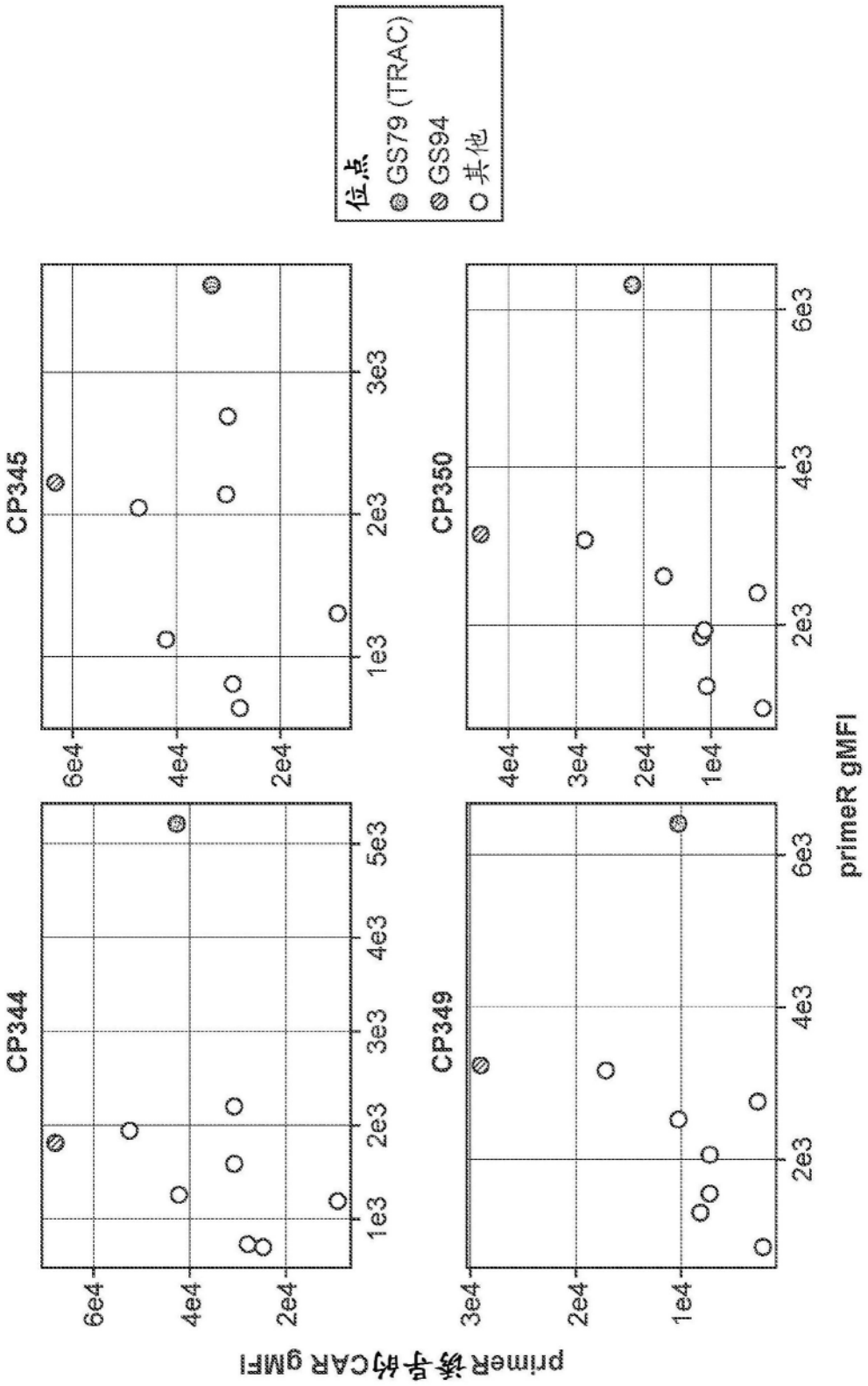

位点
GS79 (TRAC)
GS94
其他
CP345
CP350
CP344
CP349
primeR诱导的CAR gMFI
primeR gMFI
6e4
4e4
2e4
3e4
1e4
1e3
2e3
3e3
4e3
5e3
6e3

图8A

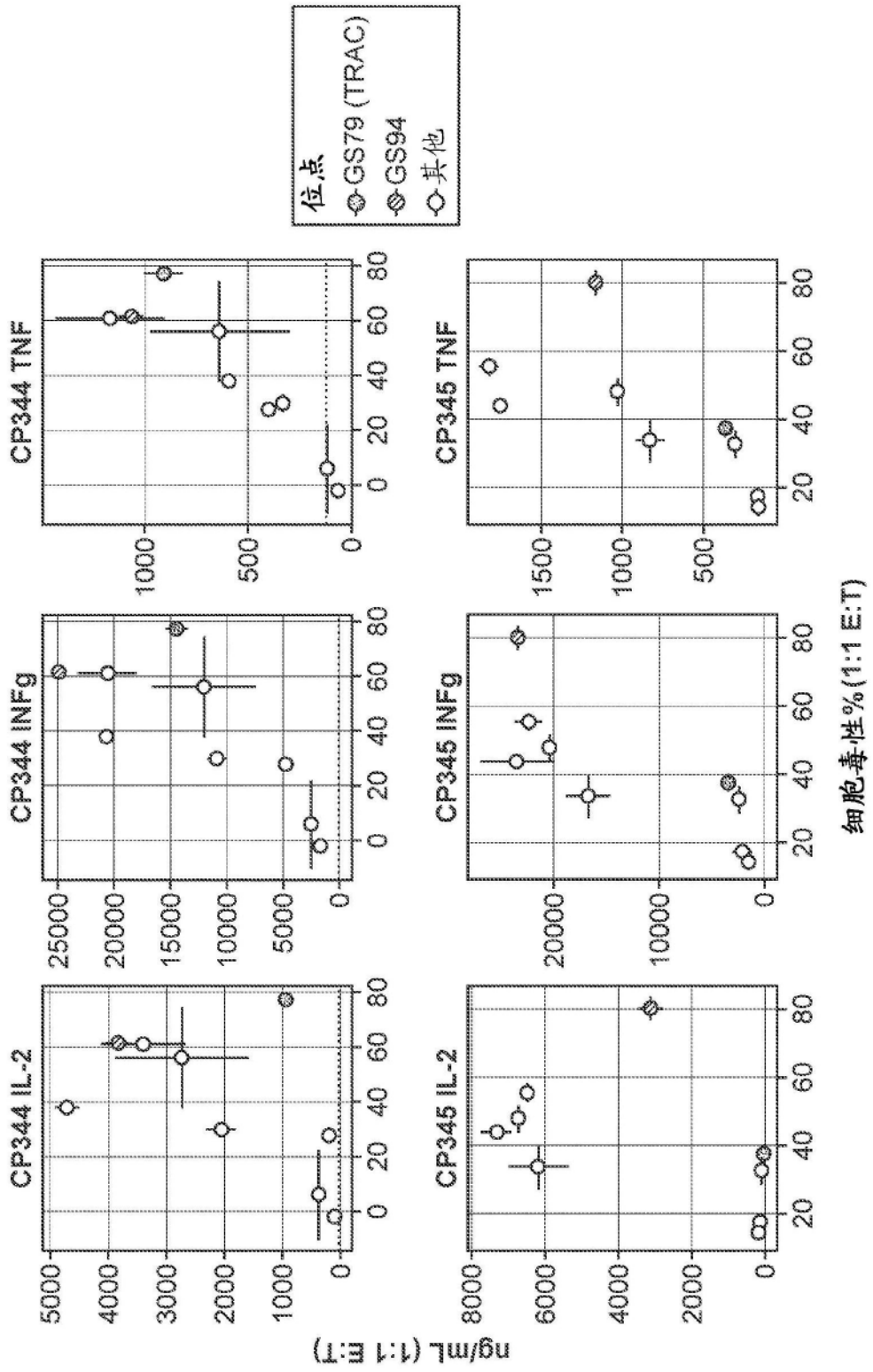

位点
GS79 (TRAC)
GS94
其他
CP344 TNF
CP345 TNF
CP344 INFg
CP345 INFg
CP344 IL-2
CP345 IL-2
ng/mL (1:1 E:T)
细胞毒性% (1:1 E:T)

图8B

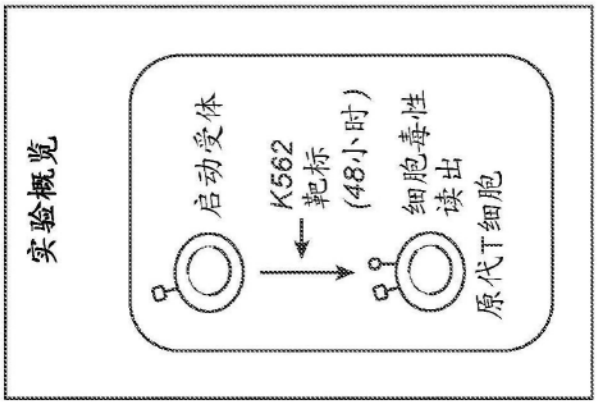

实验概览
启动受体
K562
靶标
(48小时)
细胞毒性
读出
原代T细胞

图9A

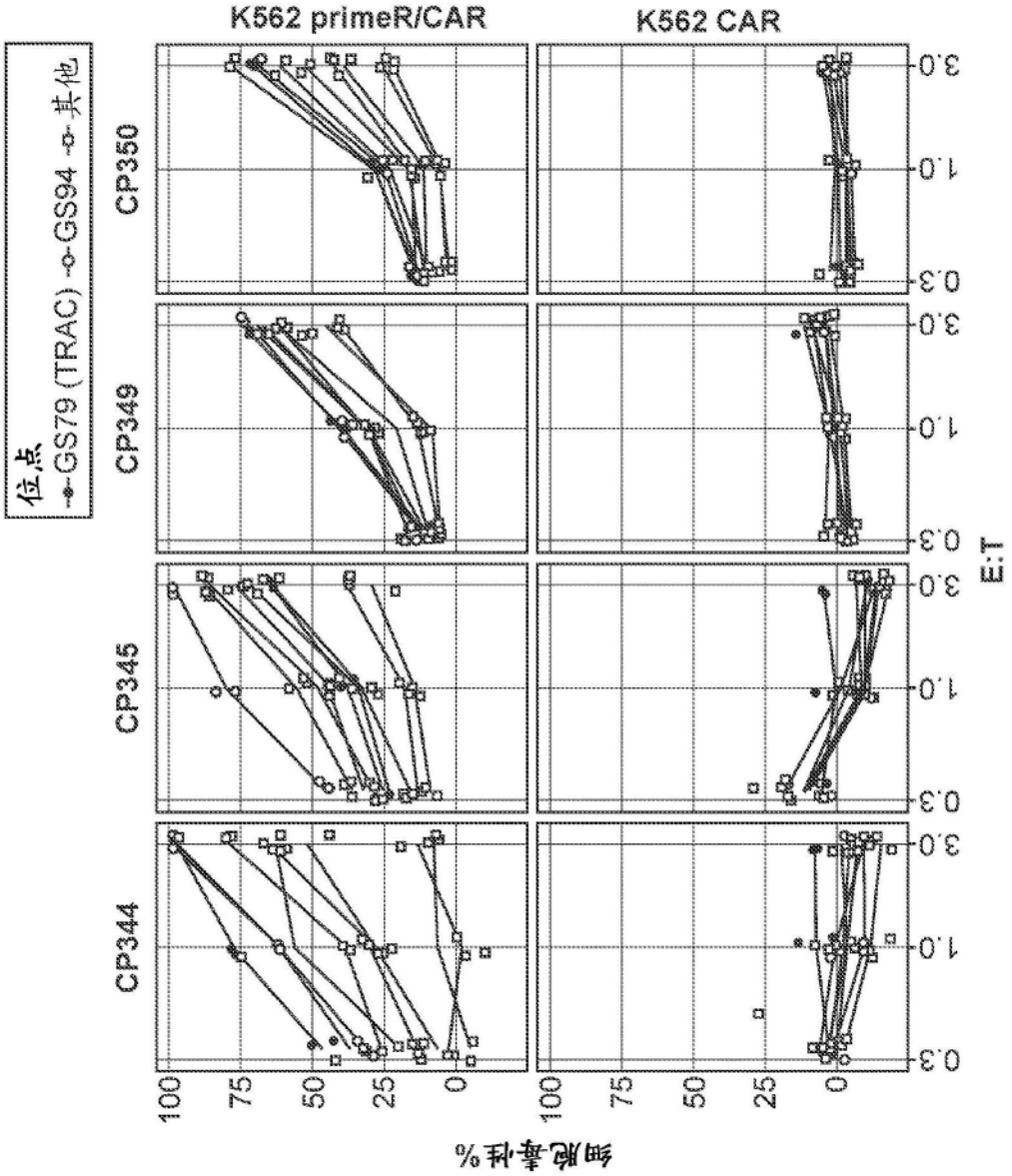

位点
GS79 (TRAC)
GS94
其他
K562 primeR/CAR
K562 CAR
CP350
CP349
CP345
CP344
E:T
细胞毒性%

图9B

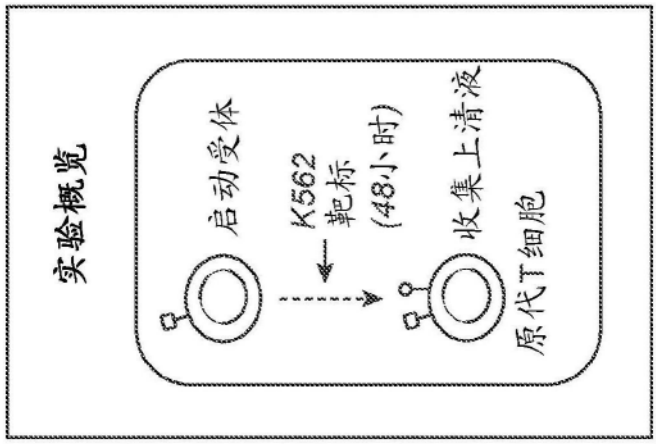


图10A

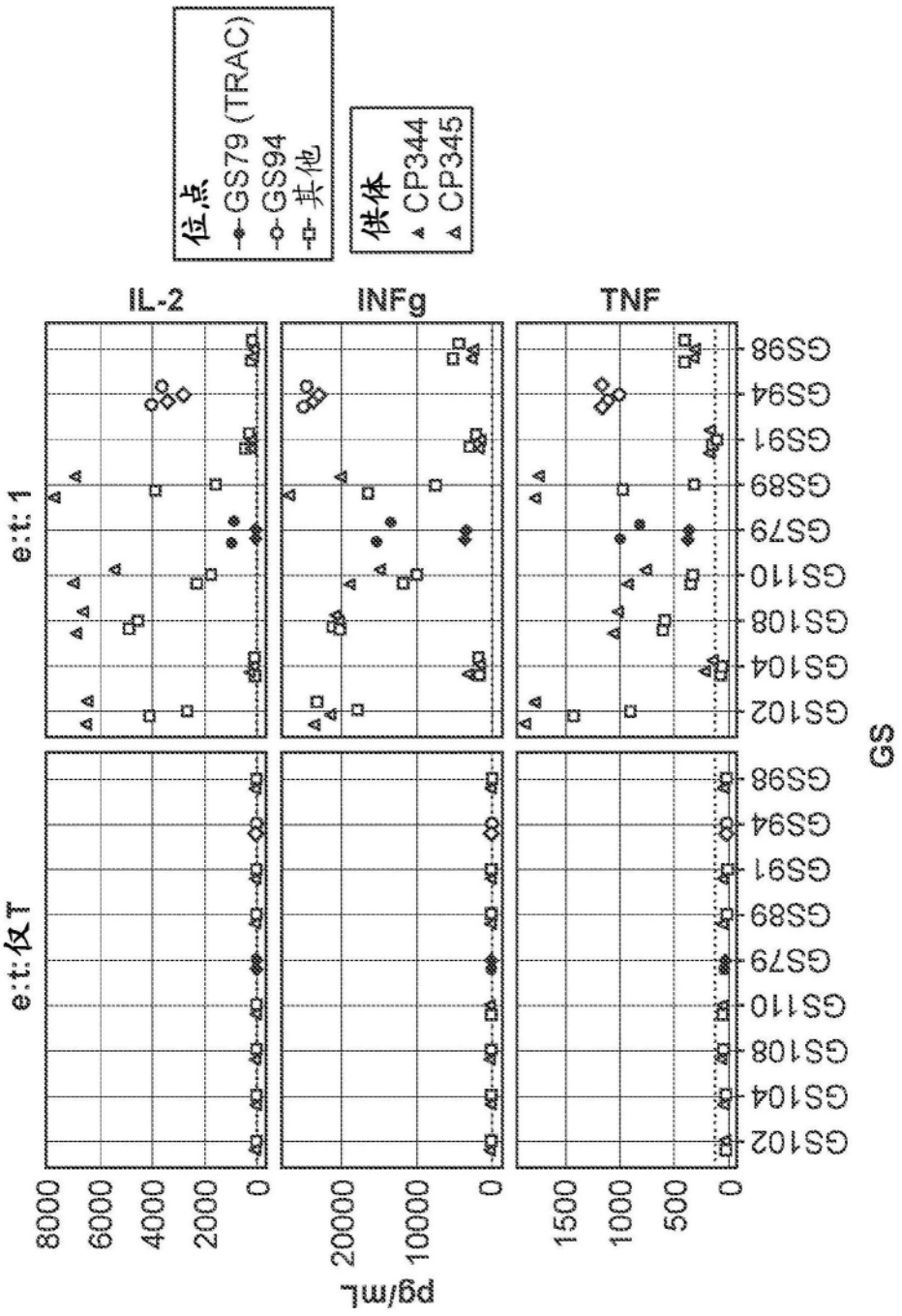

位点
GS79 (TRAC)
GS94
其他
供体
CP344
CP345
IL-2
INFg
TNF
e:t:1
e:t:仅T
pg/mL
GS
GS102
GS104
GS108
GS110
GS79
GS89
GS91
GS94
GS98
8000
6000
4000
2000
0
20000
10000
1500
1000
500

图10B

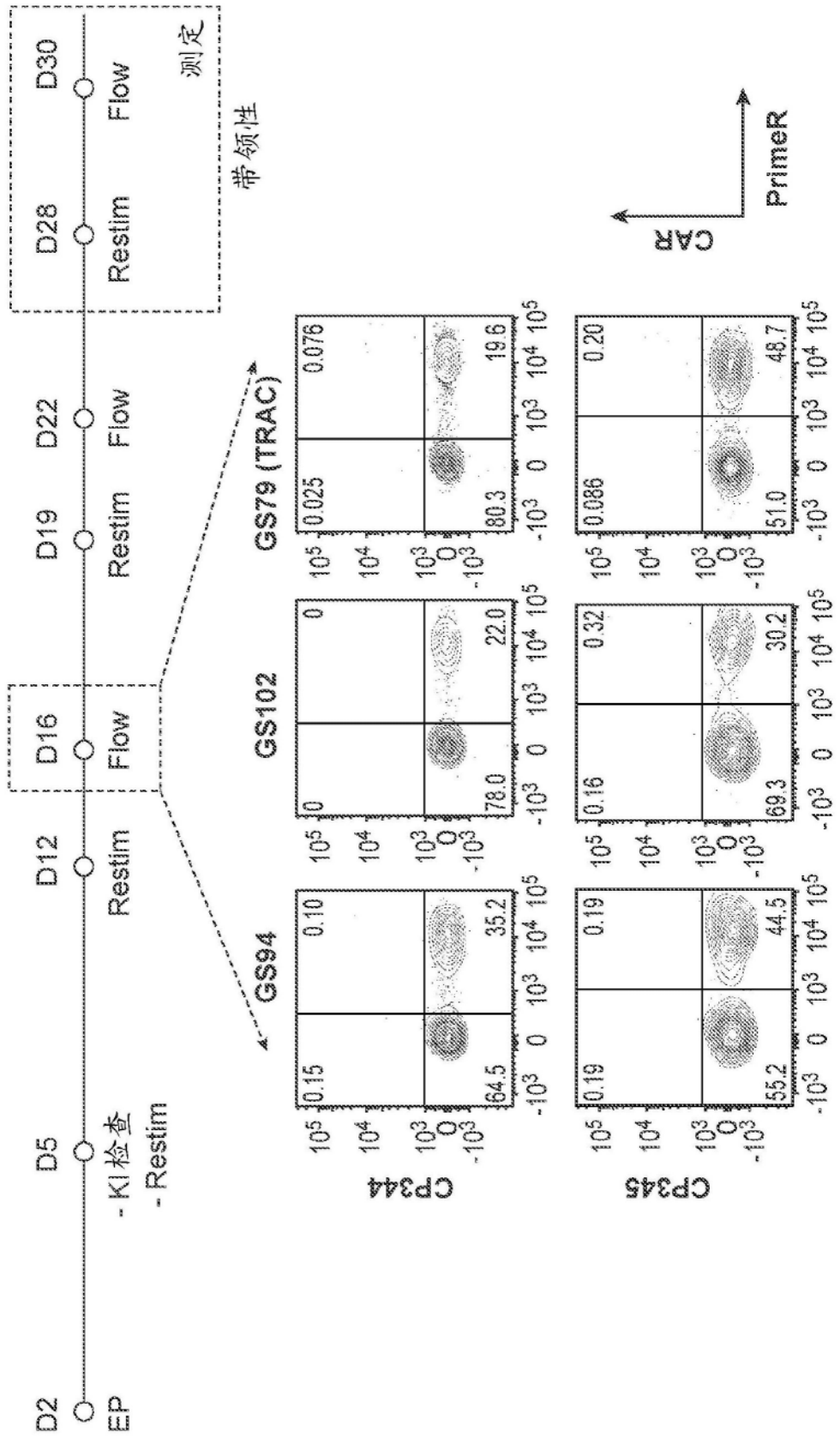

D2
EP
D5
- KI检查
- Restim
D12
Restim
D16
Flow
D19
Restim
D22
Flow
D28
Restim
D30
Flow
测定
带领性
GS94
GS102
GS79 (TRAC)
CP344
CP345
CAR
PrimeR

图11

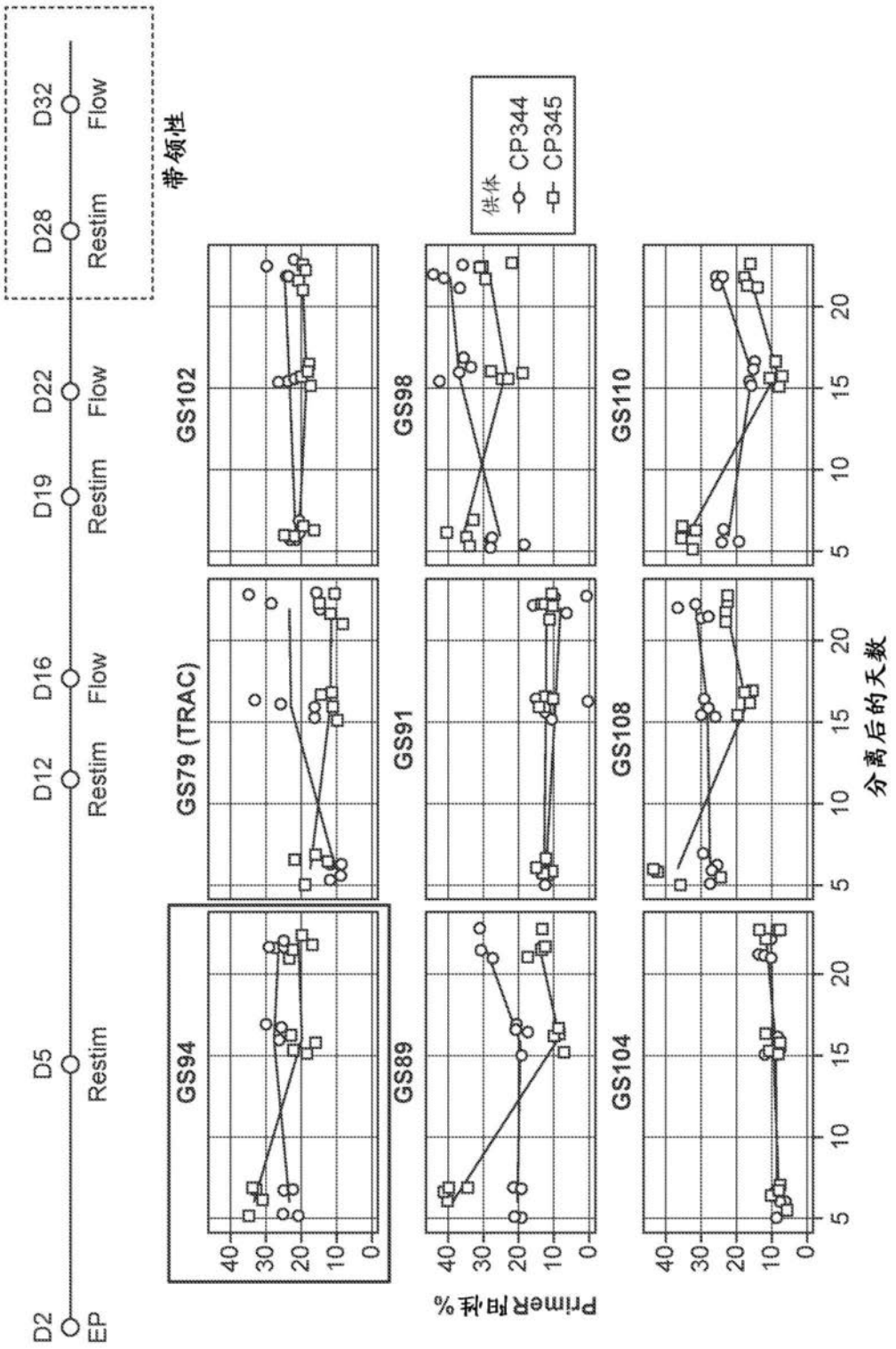

D2
EP
D5
Restim
D12
Restim
D16
Flow
D19
Restim
D22
Flow
D28
Restim
D32
Flow
带领性
GS94
GS79 (TRAC)
GS102
GS89
GS91
GS98
GS104
GS108
GS110
供体
CP344
CP345
Primer阳性%
分离后的天数

图12A

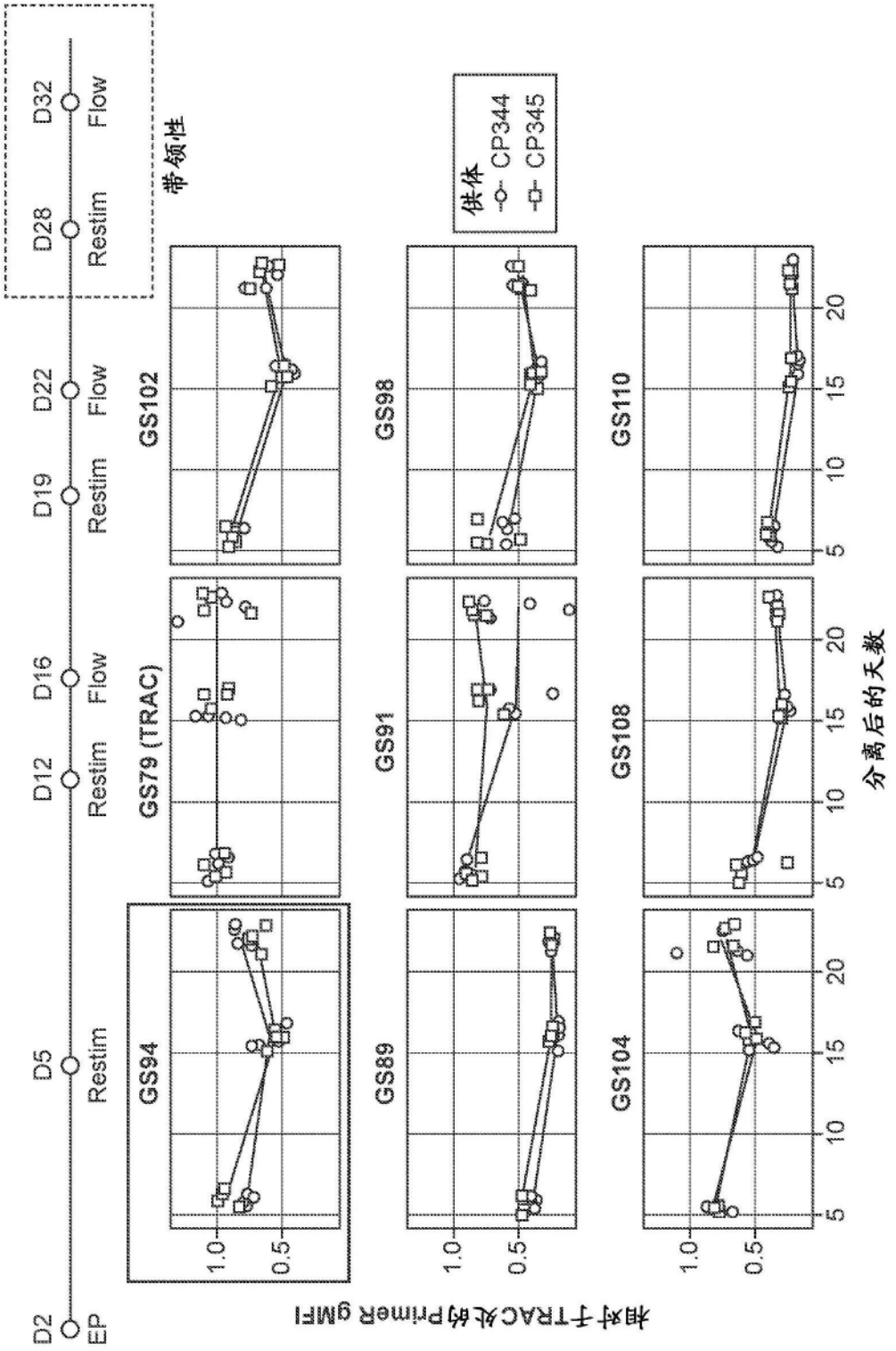

D2
EP
D5
Restim
D12
Restim
D16
Flow
D19
Restim
D22
Flow
D28
Restim
D32
Flow
带领性
供体
CP344
CP345
GS94
GS79 (TRAC)
GS102
GS89
GS91
GS98
GS104
GS108
GS110
1.0
0.5
5
10
15
20
相对于TRAC处的Primer gMFI
分离后的天数

图12B

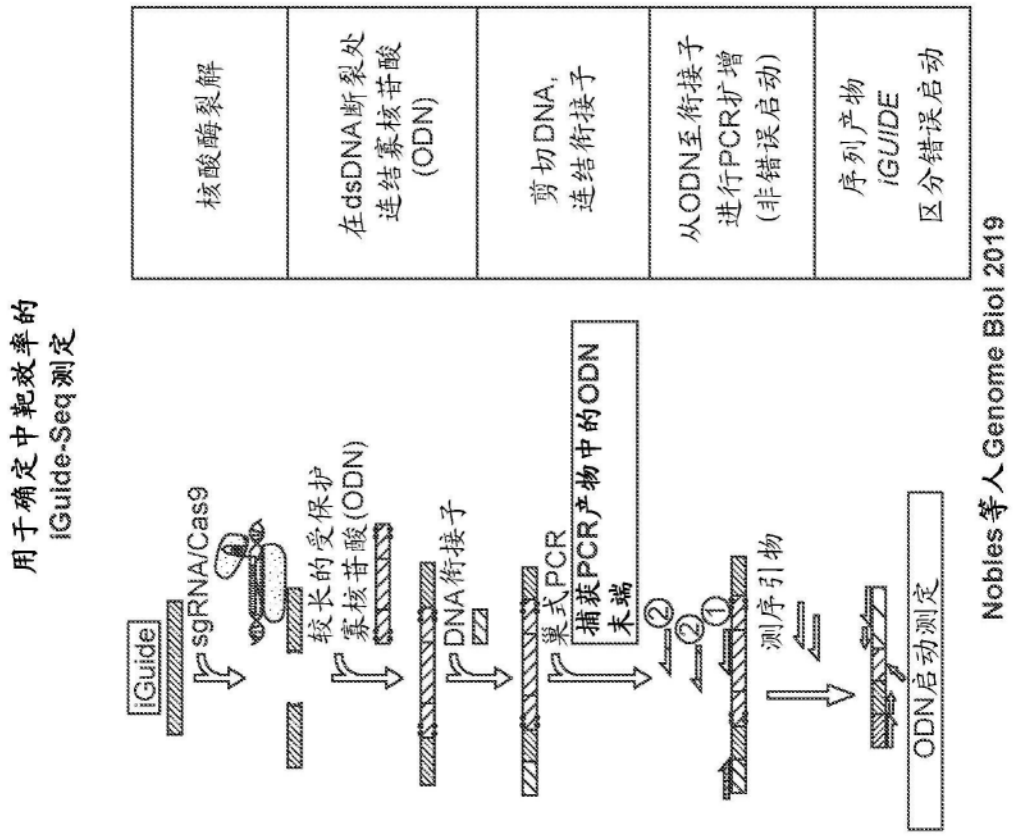

用于确定中靶效率的
iGuide-Seq测定
核酸酶裂解
在dsDNA断裂处
连结寡核苷酸
(ODN)
剪切DNA,
连结衔接子
从ODN至衔接子
进行PCR扩增
(非错误启动)
序列产物
iGUIDE
区分错误启动
iGuide
sgRNA/Cas9
较长的受保护
寡核苷酸(ODN)
DNA衔接子
巢式PCR
捕获PCR产物中的ODN
末端
测序引物
ODN启动测定
Nobles等人 Genome Biol 2019

图13A

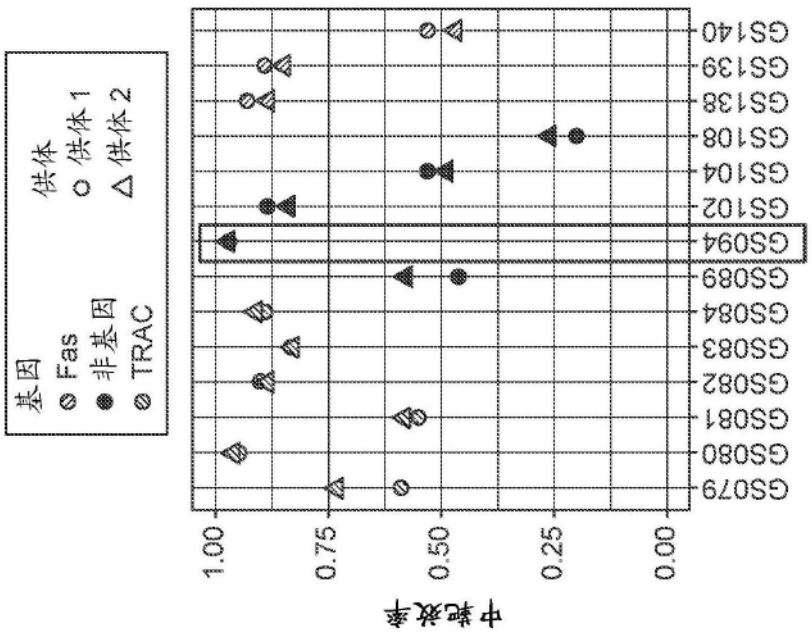
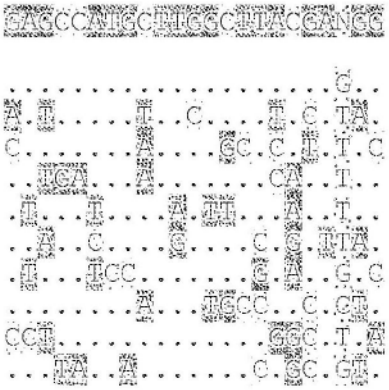

基因
Fas
非基因
TRAC
供体
供体1
供体2
中靶效率
1.00
0.75
0.50
0.25
0.00
GS079
GS080
GS081
GS082
GS083
GS084
GS089
GS094
GS102
GS104
GS108
GS138
GS139
GS140

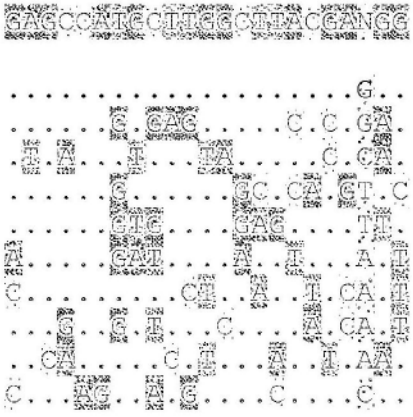
图13B

iGUIDE-seq 20pm 96W sgRNA_94 (G)
靶标: sgRNA_94



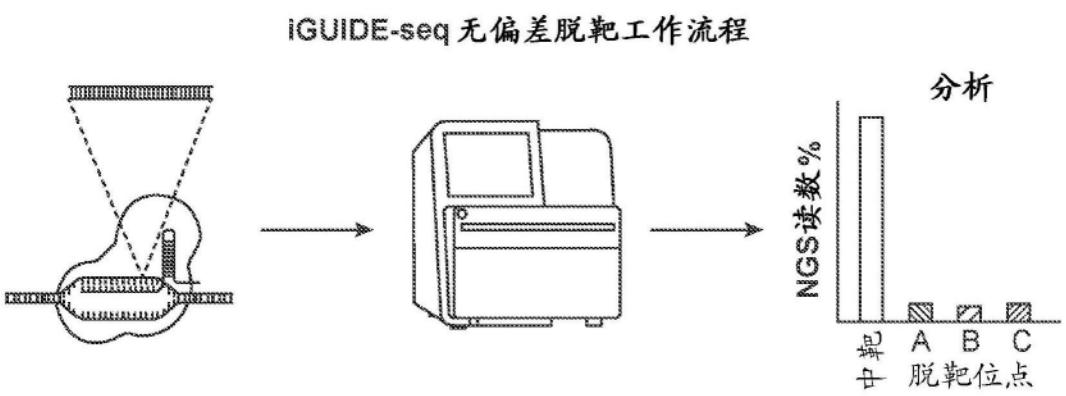
错配	靶标	丰度	MESL	基因_ID
0	中靶	2,336	95.1	ETS1~
6	脱靶	7	23.9	RNU6-259P
6	脱靶	5	32.3	RN7SKP254
6	脱靶	5	23.7	CYCSP17
6	脱靶	4	95.1	RNU4-8P
6	脱靶	4	36.6	TBC1D14*
6	脱靶	4	35.4	MDFI*
6	脱靶	3	32.3	L0C101928823*
6	脱靶	3	30.0	MYBBP1A*~
6	脱靶	3	29.4	HSPD1P15

iGUIDE-seq 20pm 96W sgRNA_94 (供体2) (G)
靶标: sgRNA_94



错配	靶标	丰度	MESL	基因_ID
0	中靶	2,164	95.5	ETS1~
6	脱靶	4	37.3	ZFPM1*
6	脱靶	2	27.3	0R6D1P
6	脱靶	2	25.2	NPHP4*
6	脱靶	1	80.5	ELN~
6	脱靶	1	56.9	RAPGEF4*
6	脱靶	1	43.2	TMEM181*
6	脱靶	1	37.3	LDHD
6	脱靶	1	37.3	LINC01392
6	脱靶	1	32.6	C14orf180

图13C


iGUIDE-seq 无偏差脱靶工作流程
分析
NGS读数 %
A
B
C
脱靶位点

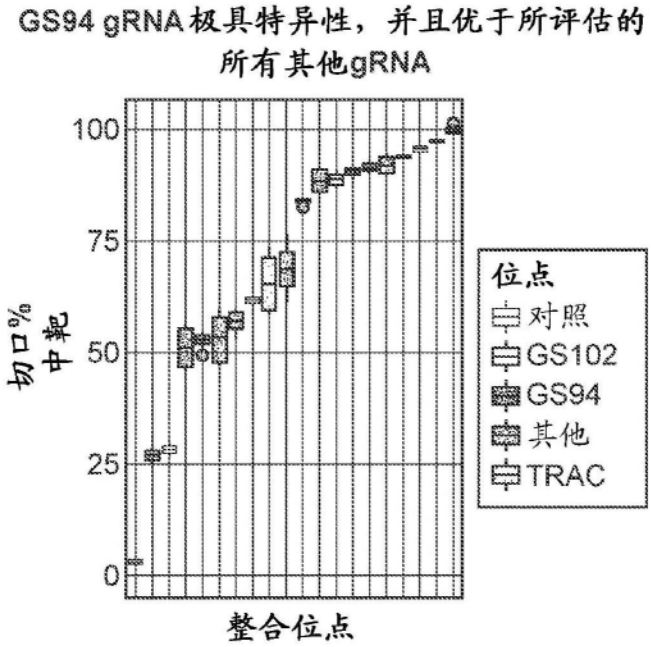

GS94 gRNA 极具特异性，并且优于所评估的所有其他gRNA
切口 % 中靶
100
75
50
25
0
整合位点
位点
对照
GS102
GS94
其他
TRAC

图14

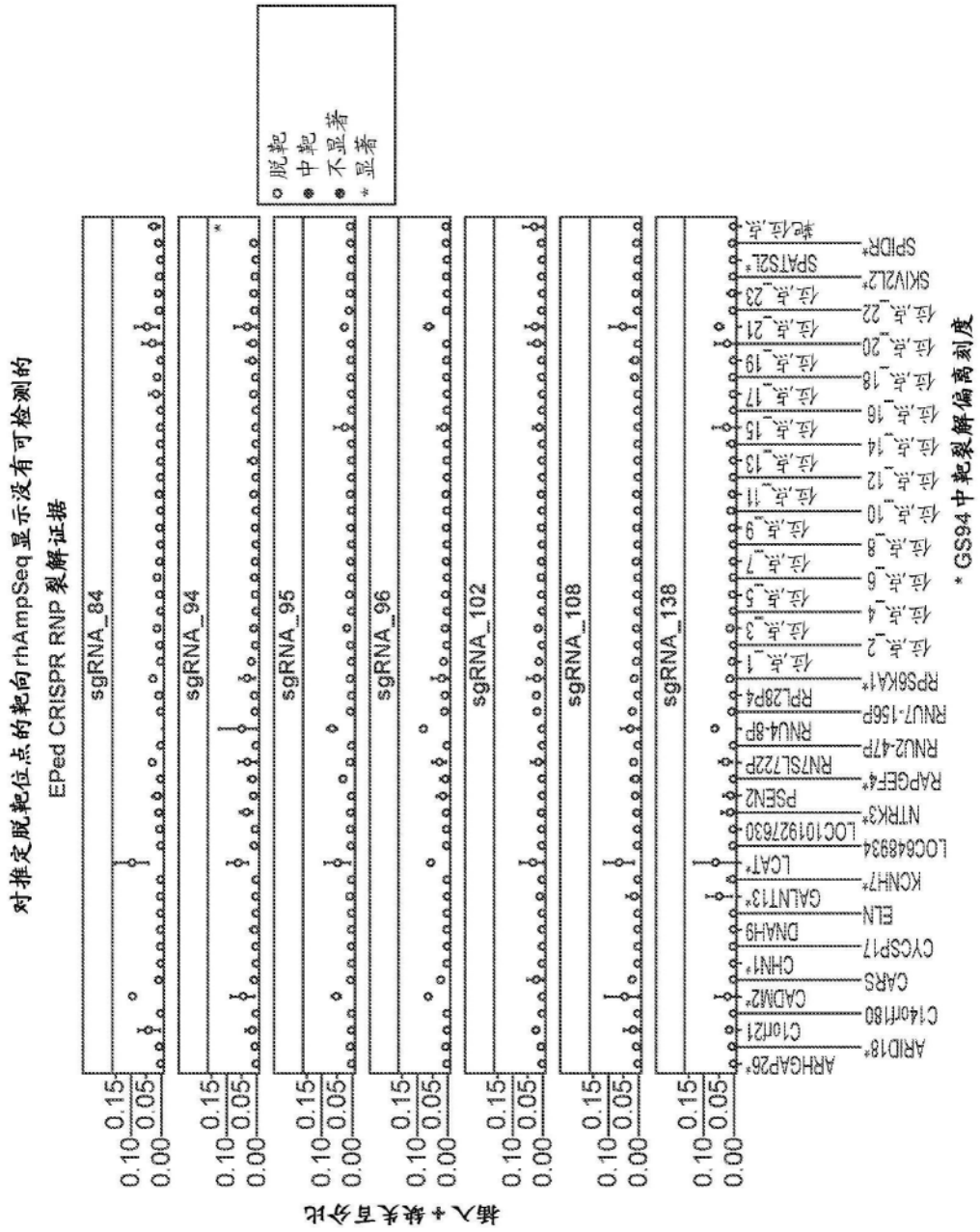

对推定脱靶位点的靶向rhAmpSeq显示没有可检测的
EPed CRISPR RNP 裂解证据
脱靶
中靶
不显著
显著
sgRNA_84
sgRNA_94
sgRNA_95
sgRNA_96
sgRNA_102
sgRNA_108
sgRNA_138
插入+缺失百分比
0.15
0.10
0.05
0.00
ARHGAP26*
ARID1B*
C10rf21
C14orf180
CADM2*
CARS
CHN1*
CYCSP17
DNAH9
ELN
GALNT13*
KCNH7*
LCAT*
LOC64893
LOC101927630
NTRK3*
PSEN2
RAPGEF4*
RN7SL722P
RNU2-47P
RNU4-8P
RNU7-156P
RPL28P4
RPS6KA1*
在点_1
在点_2
在点_3
在点_4
在点_5
在点_6
在点_7
在点_8
在点_9
在点_10
在点_11
在点_12
在点_13
在点_14
在点_15
在点_16
在点_17
在点_18
在点_19
在点_20
在点_21
在点_22
在点_23
SKIV2L2*
SPATS2L*
SPIDR*
靶点
* GS94中靶裂解偏离刻度

图15

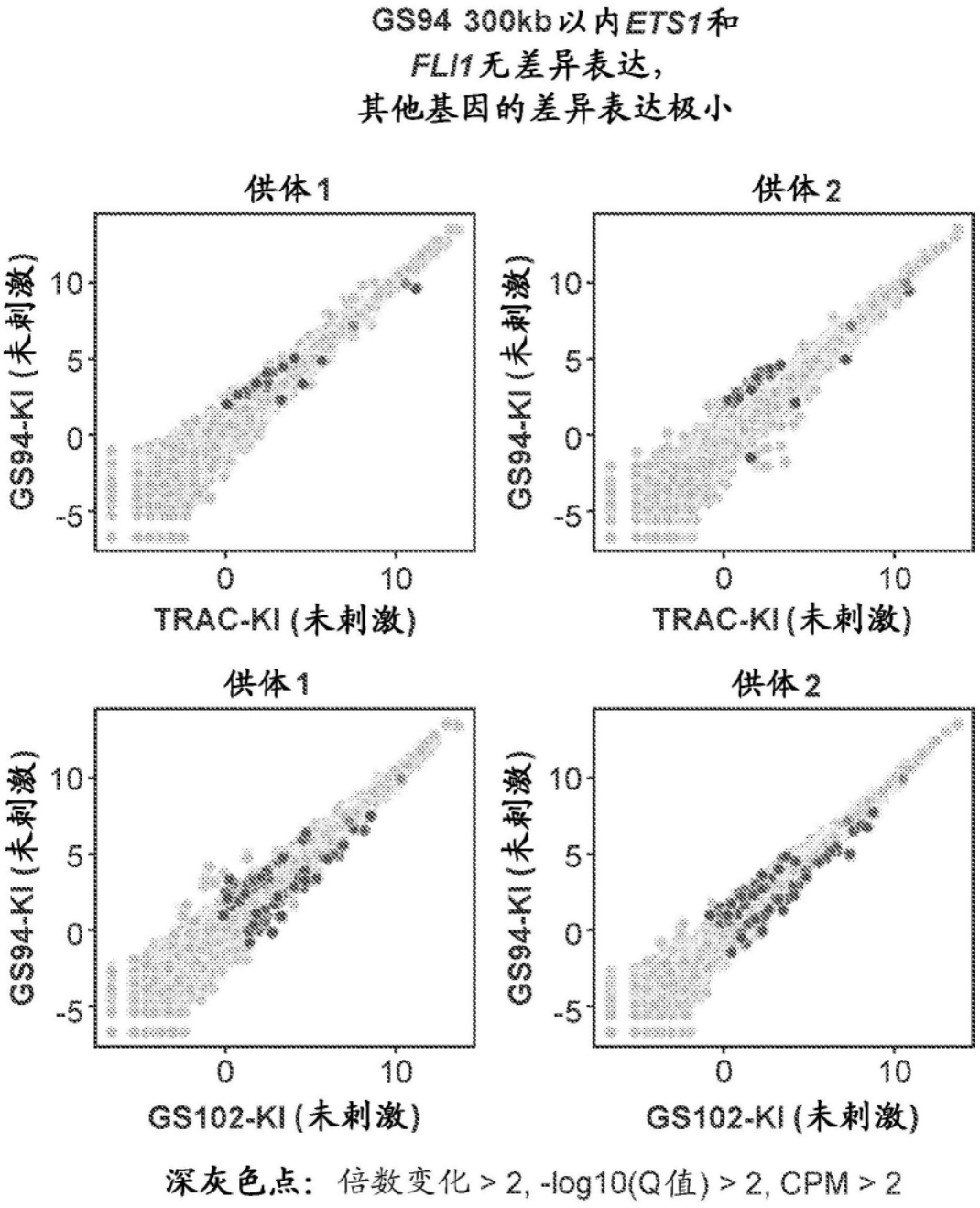

GS94 300kb以内ETS1和
FLI1无差异表达，
其他基因的差异表达极小
供体1
供体2
GS94-KI (未刺激)
TRAC-KI (未刺激)
GS102-KI (未刺激)
10
5
0
-5
0
10
深灰色点：倍数变化 > 2, -log10(Q值) > 2, CPM > 2

图16

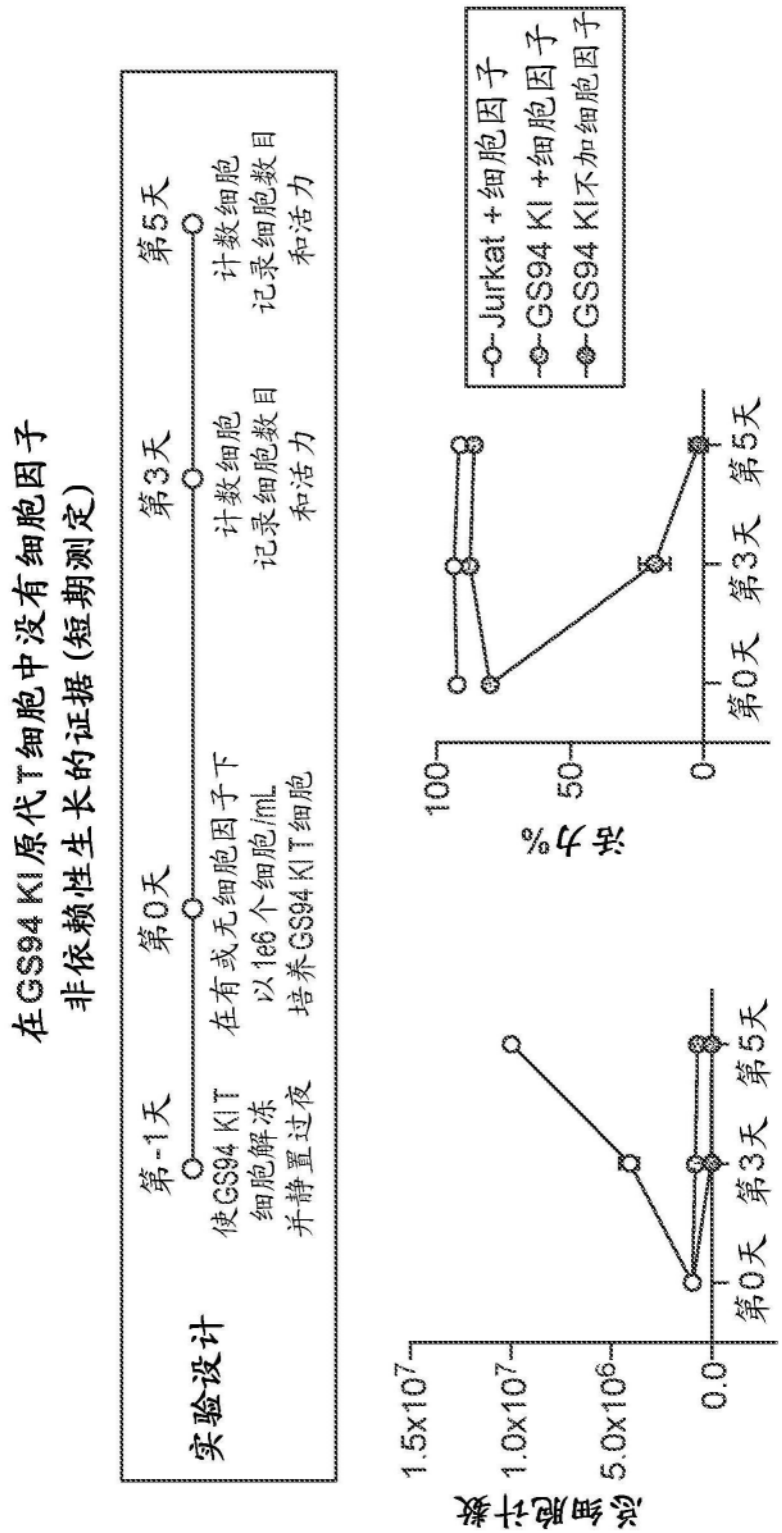

在GS94 KI原代T细胞中没有细胞因子
非依赖性生长的证据(短期测定)
实验设计
第-1天
第0天
第3天
第5天
使GS94 KIT细胞解冻并静置过夜
在有或无细胞因子下以1e6个细胞/mL培养GS94 KIT细胞
计数细胞记录细胞数目和活力
计数细胞记录细胞数目和活力
Jurkat +细胞因子
GS94 KI +细胞因子
GS94 KI不加细胞因子
总细胞计数
1.5x10^7
1.0x10^7
5.0x10^6
0.0
第0天 第3天 第5天
活力%
100
50
0
第0天 第3天 第5天

图17

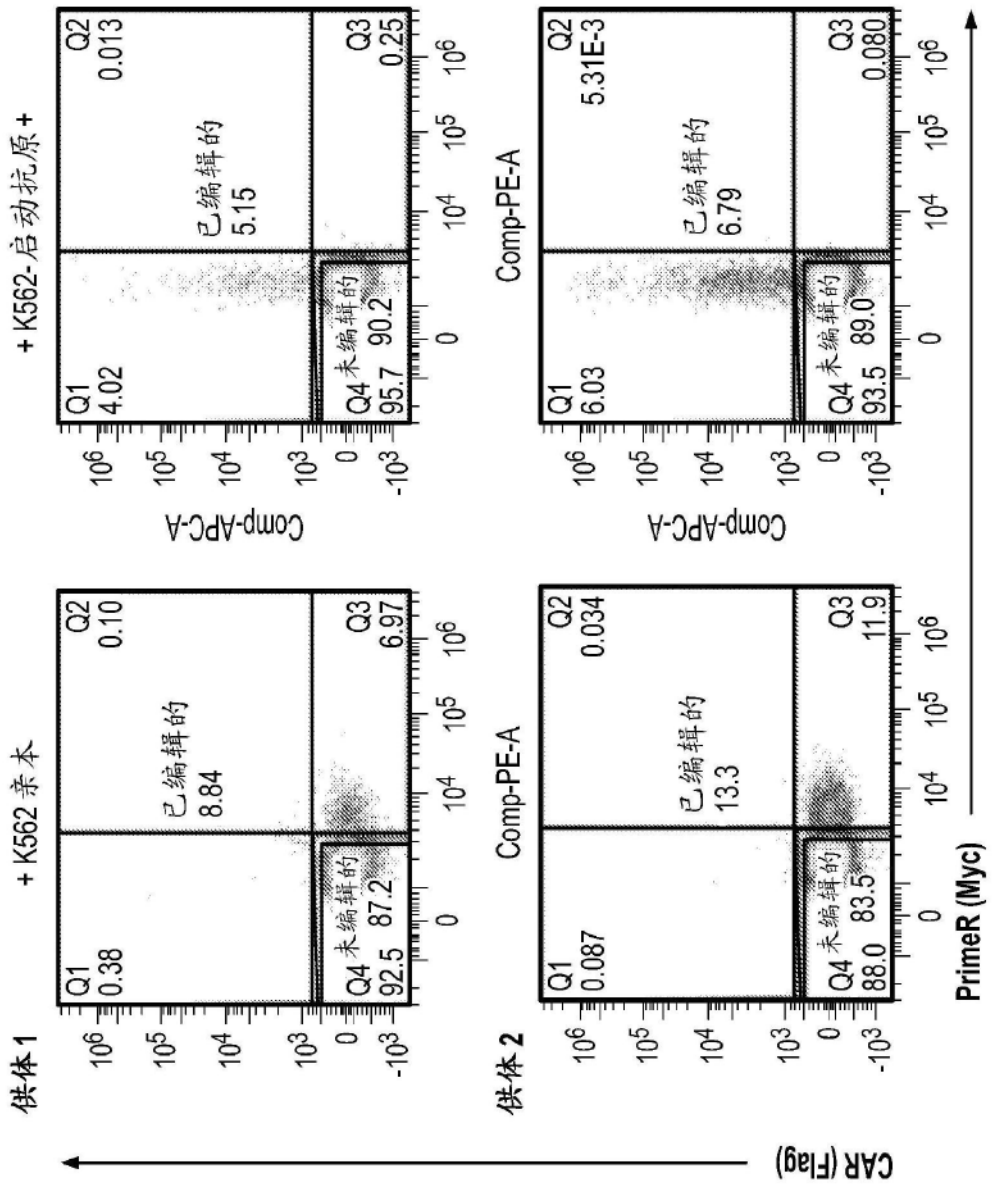

+ K562-启动抗原 +
+ K562 亲本
供体 1
供体 2
Comp-APC-A
Comp-PE-A
CAR (Flag)
PrimeR (Myc)
Q1 4.02
Q2 0.013
已编辑的 5.15
Q4 未编辑的 90.2 95.7
Q3 0.25
Q1 6.03
Q2 5.31E-3
已编辑的 6.79
Q4 未编辑的 89.0 93.5
Q3 0.080
Q1 0.38
Q2 0.10
已编辑的 8.84
Q4 未编辑的 87.2 92.5
Q3 6.97
Q1 0.087
Q2 0.034
已编辑的 13.3
Q4 未编辑的 83.5 88.0
Q3 11.9

图18

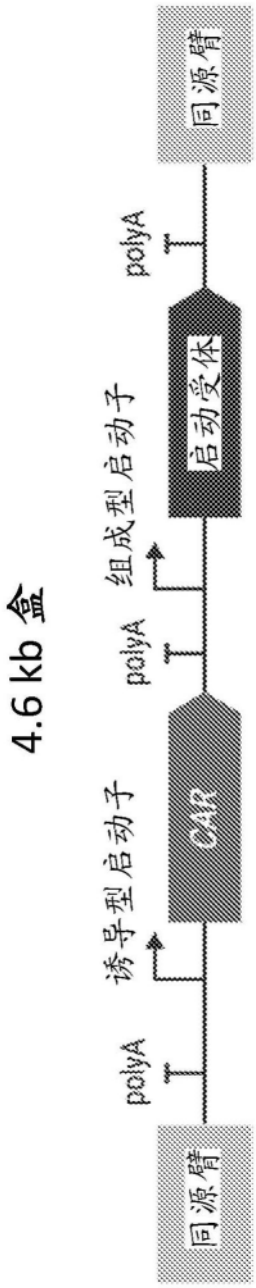


图19A

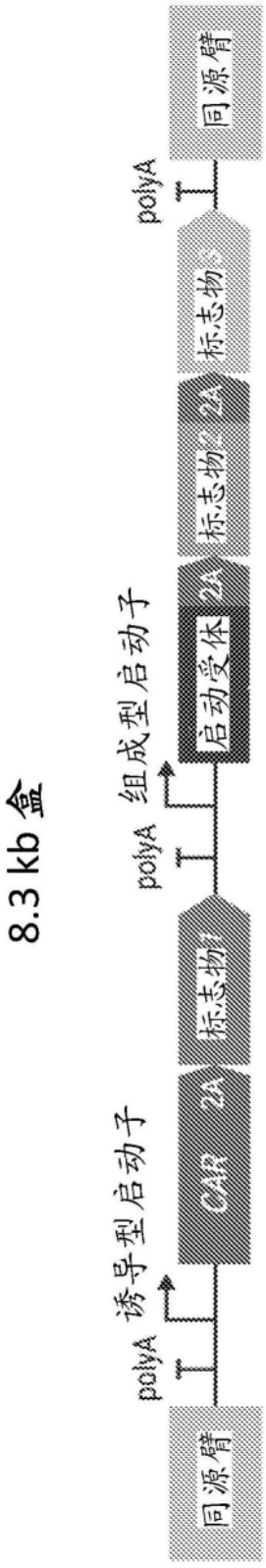

8.3 kb 盒
同源臂
polyA
诱导型启动子
CAR
2A
标志物1
polyA
组成型启动子
启动受体
2A
标志物2
2A
标志物3
polyA
同源臂

图19B