



(12)发明专利申请

(10)申请公布号 CN 108026545 A

(43)申请公布日 2018.05.11

(21)申请号 201680044599.6

(22)申请日 2016.06.17

(30)优先权数据

62/181,145 2015.06.17 US

(85)PCT国际申请进入国家阶段日

2018.01.29

(86)PCT国际申请的申请数据

PCT/US2016/038125 2016.06.17

(87)PCT国际申请的公布数据

W02016/205680 EN 2016.12.22

(71)申请人 UAB研究基金会

地址 美国阿拉巴马州

(72)发明人 蒂姆·汤斯 詹家伟 丁磊

(74)专利代理机构 北京英赛嘉华知识产权代理有限公司 11204

代理人 王达佐 洪欣

(51)Int.Cl.

C12N 15/90(2006.01)

C12N 9/22(2006.01)

C12N 15/00(2006.01)

C12N 15/113(2010.01)

A61K 48/00(2006.01)

权利要求书2页 说明书22页

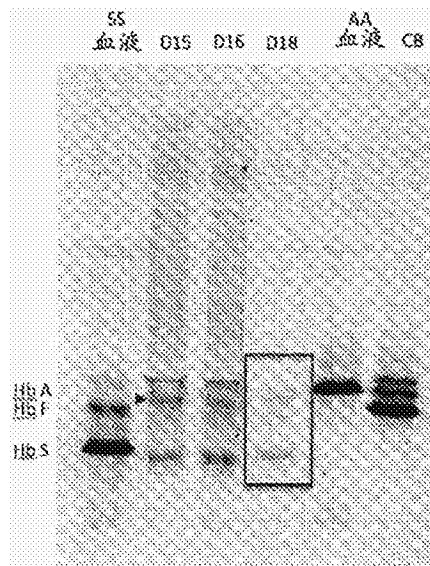
序列表25页 附图10页

(54)发明名称

用于将功能性多肽引入到血细胞谱系细胞中的CRISPR/CAS9复合物

(57)摘要

本文提供CRISPR/Cas9复合物及其使用方法。



1. 一种制备肿瘤特异性T细胞前体细胞的方法,其包括将包含以下项的复合物引入到T细胞前体细胞群体中:

a. 指导序列(gRNA),其包含与所述T细胞前体细胞基因组中的靶DNA杂交的第一核苷酸序列和与定点核酸酶相互作用的第二核苷酸序列;

b. 重组定点核酸酶,其可操作地连接至超荷电蛋白,其中所述定点核酸酶包含与所述gRNA的所述第二核苷酸序列相互作用的RNA结合部分,并且其中所述定点核酸酶特异性结合并切割所述靶DNA以产生双链断裂;以及

c. 供体核酸序列,其包含编码嵌合抗原受体(CAR)的第三核苷酸序列和与所述靶DNA中的所述双链断裂侧翼的基因组序列杂交的第四核苷酸序列,

其中在允许同源性定向修复(HDR)和将所述第三核苷酸序列整合到所述靶DNA中的条件下将所述复合物引入到所述T细胞前体细胞中以形成表达所述CAR的修饰的T细胞前体细胞。

2. 如权利要求1所述的方法,其中所述细胞选自造血干细胞或多能干细胞组成的组。

3. 如权利要求2所述的方法,其中所述多能干细胞是诱导性多能干细胞。

4. 如权利要求1-3所述的方法,其中所述超荷电蛋白可操作地连接至所述核酸酶的氨基末端或羧基末端。

5. 如权利要求1-4中任一项所述的方法,其中可操作地连接至超荷电蛋白的所述重组定点核酸酶还包含可操作地连接至所述定点核酸酶的氨基末端的反式激活转录激活因子(TAT)肽。

6. 如权利要求1-5中任一项所述的方法,其中所述超荷电蛋白的总体正电荷大于其对应的未修饰蛋白。

7. 如权利要求6所述的方法,其中所述总体正电荷为约+5至约+40。

8. 如权利要求1-7中任一项所述的方法,其中所述超荷电蛋白是超荷正电绿色荧光蛋白(GFP)。

9. 如权利要求8所述的方法,其中所述超荷电蛋白是超荷正电的+36GFP。

10. 如权利要求1-9中任一项所述的方法,其中所述核酸酶是Cas9。

11. 如权利要求1-10中任一项所述的方法,其中gRNA与可操作地连接至超荷电蛋白的定点核酸酶与ssODN的摩尔比为约1:1:1至约1.5:1:1。

12. 如权利要求1-11中任一项所述的方法,其中所述T细胞前体细胞群体中同源性定向修复与非同源性末端接合的比率为至少约0.5。

13. 如权利要求1-12中任一项所述的方法,其中所述复合物通过核穿孔引入到所述细胞中。

14. 如权利要求1-13中任一项所述的方法,其中至少5%的所述T细胞前体细胞群体通过HDR进行修饰以形成表达所述CAR的修饰的T细胞前体细胞。

15. 如权利要求1-14中任一项所述的方法,其还包括分离所述修饰的T细胞前体细胞。

16. 如权利要求15所述的方法,其还包括培养所述修饰的T细胞前体细胞。

17. 如权利要求16所述的方法,其中所述修饰的T细胞前体细胞在扩增条件下培养。

18. 如权利要求16或17所述的方法,其还包括在促进所述修饰的T细胞前体细胞分化成表达所述CAR的T细胞的条件下培养所述修饰的T细胞前体细胞。

19. 一种治疗受试者的癌症的方法,其包括将通过如权利要求1-18中任一项所述的方法获得的细胞移植到所述受试者体内。

20. 如权利要求19所述的方法,其中所述移植是自体的。

21. 一种用于制备肿瘤特异性T细胞前体细胞的复合物,其包含:

a. 指导序列(gRNA),其包含与所述T细胞前体细胞基因组中的靶DNA杂交的第一核苷酸序列和与定点核酸酶相互作用的第二核苷酸序列;

b. 重组定点核酸酶,其可操作地连接至超荷电蛋白,其中所述定点核酸酶包含与所述gRNA的所述第二核苷酸序列相互作用的RNA结合部分,并且其中所述定点核酸酶特异性结合并切割所述靶DNA以产生双链断裂;以及

c. 供体核酸序列,其包含编码嵌合抗原受体(CAR)的第三核苷酸序列和与所述靶DNA中的所述双链断裂侧翼的基因组序列杂交的第四核苷酸序列。

22. 如权利要求21所述的复合物,其中所述超荷电蛋白可操作地连接至所述核酸酶的氨基末端或羧基末端。

23. 如权利要求21或22中任一项所述的复合物,其中可操作地连接至超荷电蛋白的所述重组定点核酸酶还包含可操作地连接至所述定点核酸酶的氨基末端的反式激活转录激活因子(TAT)肽。

24. 如权利要求21-23中任一项所述的复合物,其中所述超荷电蛋白的总体正电荷大于其对应的未修饰蛋白。

25. 如权利要求24所述的复合物,其中所述总体正电荷为约+5至约+40。

26. 如权利要求21-25中任一项所述的复合物,其中所述超荷电蛋白是超荷正电绿色荧光蛋白(GFP)。

27. 如权利要求26所述的复合物,其中所述超荷电蛋白是超荷正电的+36GFP。

28. 如权利要求21-27中任一项所述的复合物,其中所述核酸酶是Cas9。

29. 如权利要求21-28中任一项所述的复合物,其中gRNA与可操作地连接至超荷电蛋白的定点核酸酶与ssODN的摩尔比为约1:1:1至约1.5:1:1。

用于将功能性多肽引入到血细胞谱系细胞中的CRISPR/CAS9复合物

[0001] 本申请要求2015年6月17日提交的美国临时申请号62/181,145的权益,所述临时申请特此以引用的方式整体并入。

[0002] 背景

[0003] 成簇规律间隔的短回文重复序列(CRISPR)相关(Cas)系统(CRISPR-Cas9系统)用于在哺乳动物细胞中的期望基因组位点处的基因编辑。在CRISPR-Cas9系统中,Cas9核酸酶通过与指导RNA复合而靶向基因组位点,所述指导RNA与基因组中的靶位点杂交。这导致通过双链或单链DNA修复模板启动基因组DNA的非同源性末端接合(NHEJ)或同源性定向修复(HDR)的双链断裂。然而,通过HDR修复基因组位点效率不高。

[0004] 概述

[0005] 本文提供用于将功能性多肽引入到血谱系细胞中的复合物和方法。本文提供一种制备肿瘤特异性T细胞前体细胞的方法所述方法包括将包含以下项的复合物引入到T细胞前体细胞群体中:(a) 指导序列(gRNA),其包含与T细胞前体细胞基因组中的靶DNA杂交的第一核苷酸序列和与定点核酸酶相互作用的第二核苷酸序列;(b) 重组定点核酸酶,其可操作地连接至超荷电蛋白,其中定点核酸酶包含与gRNA的第二核苷酸序列相互作用的RNA结合部分,并且其中定点核酸酶特异性结合并切割靶DNA以产生双链断裂;以及(c) 供体核酸序列,其包含编码嵌合抗原受体(CAR)的第三核苷酸序列和与靶DNA中的双链断裂侧翼的基因组序列杂交的第四核苷酸序列,其中在允许同源性定向修复(HDR)和将第三核苷酸序列整合到靶DNA中的条件下将复合物引入到T细胞前体细胞中以形成表达CAR的修饰的T细胞前体细胞。所述方法进一步提供修饰的T细胞前体细胞的高细胞存活率。还提供用于制备肿瘤特异性T细胞前体细胞的复合物。

[0006] 附图描述

[0007] 图1A-1C显示,JAK3 C1837T患者iPSC的体外分化重现SCID表型。图1A和1B示出iPSC衍生的T细胞的流式细胞术。将JAK3 WT iPSC(对照)和JAK3缺陷型iPSC(JAK3 C1837T)在OP9基质细胞上分化成CD34⁺细胞,随后在OP9-DL4单层上分化成T细胞。与对照相比,不存在来自JAK3缺陷型iPSC的T细胞分化;未观察到CD3⁺T细胞或CD3⁻CD16⁺CD56⁺NK细胞(图1A),并且未检测到CD4⁺CD8⁺双阳性(DP)、CD4⁺单阳性(SP)或CD8⁺单阳性(SP)T细胞(图1B)。图1C示出在T细胞谱系特化期间调控早期事件的关键基因的转录物的RT-qPCR测定结果。示出关于GAPDH表达的RNA水平。

[0008] 图2A-2C显示,BCL2部分地挽救JAK3缺陷型体外衍生细胞中的T细胞发育缺陷。图2A示出JAK3缺陷型iPSC衍生的T细胞与JAK3 WT对照相比的凋亡。在T细胞诱导第10天(TD10)和第17天(TD17)分析膜联蛋白V阳性细胞。用对照JAK3 WT细胞(对照)进行四个独立实验,并且用JAK3缺陷型细胞(JAK3 C1837T)进行5个独立实验。*P<0.005。图2B示出来自JAK3 WT(对照)和JAK3缺陷型细胞(JAK3 C1837T)的两列(1和2)中抗凋亡BCL2和促凋亡BAX表达的RT-qPCR测定结果。ND,未确定的(由于JAK3qPCR信号不明显)。示出关于GAPDH表达的RNA水平。图2C示出用BCL2-2A-GFP慢病毒转导的JAK3缺陷型iPSC衍生的T细胞的流式细胞

术,以评估对NK (CD16+56+) 和T细胞 (CD3+) 发育以及DP (CD4+CD8+) 至SP (CD4+或CD8+) T细胞成熟的影响。

[0009] 图3A-3D显示,CRISPR/Cas9增强了对患者特异性iPSC中JAK3 C1837T突变的校正。图3A绘示出使用诱导JAK3基因座中双链断裂的CRISPR/Cas9和用于同源性定向修复的模板进行基因组修饰的策略。顶行, JAK3基因的结构。开放框,外显子。星号, C1837T突变。箭头, 指导RNA。图3B顶部示出表明同源重组的PCR分析;指示用于5'和3'分析的引物。(左下方) RT-PCR分析,其表明JAK3 WT (对照)、JAK3缺陷型 (JAK3 C1837T) 和校正型 (JAK3校正型) T细胞中的JAK3 mRNA表达。(右下方) 蛋白质印迹分析,其表明JAK3 WT (对照)、JAK3缺陷型 (JAK3 C1837T) 和校正型 (JAK3校正型) T细胞中的JAK3蛋白质表达。图3C提供对指导RNA的靶向效率的总结。(图3D) 来自亲本JAK3 iPSC (左)、杂合校正型 (中) 和纯合校正型iPSC (右) 的PCR扩增子的Sanger测序。用gRNA2+野生型Cas9校正两个杂合克隆,并用gRNA1+gRNA2+切口酶Cas9 (D10A) 校正纯合克隆。

[0010] 图4A-4C显示, JAK3校正型患者iPSC的体外分化产生具有成熟T细胞的表型特征和功能特征的T细胞。图4A显示JAK3 WT (对照, n=3)、JAK3缺陷型 (JAK3 C1837T, n=5) 和JAK3校正型 (JAK3校正型, n=6) T细胞中的T细胞发育标记物的表达。用指示的抗体对细胞进行染色并在T细胞诱导第14、21、28和35天 (TD 14、21、28和35) 通过流式细胞术进行分析。图4B显示JAK3校正型T细胞的T细胞受体 (TCR) V β 分析。在衍生自校正型SCID患者iPSC的T细胞中表现出高度多样的TCR V β 库。图4C示出表明JAK3校正型T细胞中的T细胞激活的流式细胞术。在分析激活标记物CD25和CD69之前,用抗CD3/28珠将衍生自JAK3 WT (对照) 和JAK3校正型iPSC的T细胞刺激3天。在CD3+群体上对数据进行门控。

[0011] 图5A-5C示出通过与人类骨髓基质细胞 (hMSC) 共培养而由hiPSC对CD34+HSC的体外生成。在分析造血标记物CD34和CD43 (图A) 之前,将人类iPSC在hMSC上培养18天。将CD34+细胞在珠上进行纯化并分化成T细胞 (图B)、红系细胞和骨髓细胞 (图C)。为了生成T细胞,将纯化的CD34+细胞铺在OP9-DL4细胞上3至4周。对于生成骨髓细胞和红系细胞的CFC测定,根据制造商方案将纯化的CD34+细胞铺在MethoCult H4434经典培养基中。这些数据证实,在hMSC上共培养后,hiPSC可高效地分化成多能性HSC。

[0012] 图6A-6C示出通过将hiPSC衍生的CD34+细胞与hMSC-DL4一起培养而进行的体外T细胞生成。为了生成CD7+T祖细胞,将hiPSC衍生的CD34+细胞在hMSC-DL4上共培养3至4周 (图6A)。当将来自图6A的CD7+细胞在磁珠上纯化并在OP9-DL4上共培养时,在10天或更短时间内产生完全成熟的CD4+/CD8+/CD3+/TCR- $\alpha\beta$ +细胞 (图B和C)。这些数据证实,在hMSC-DL4上共培养后,hiPSC可高效地分化成CD7+淋巴样祖细胞。

[0013] 图7示出由hiPSC对 $\gamma\delta$ T细胞的体外生成。将人类iPSC用携带与2A-GFP cDNA片段连接的预重排的人类V $\gamma\delta$ 1 cDNA的慢病毒载体进行转导。与OP9共培养18天后,将hiPSC衍生的CD34+细胞在磁珠上进行纯化。随后将这些细胞铺在OP9-DL4细胞上进行T细胞分化。在第32天采集细胞,并通过FACS分析T细胞表面标记物。GFP+群体表示V δ 1-2A-GFP慢病毒转导的细胞。高百分比的这些GFP阳性细胞表达V δ 1 (66%)。低百分比的GFP阴性细胞表达V δ 1 (1%)。这些结果证实,可由基因修饰的iPSC高效地产生表达重组T细胞受体 (TCR) 的V δ T细胞。表达对肿瘤抗原具有特异性的重组T细胞受体 (TCR) 的V δ T细胞的产生为许多类型的癌症提供强有力的细胞疗法。

[0014] 图8显示,包含指导RNA、修饰的Cas9和单链寡核苷酸供体序列(ssODN)的校正复合物可校正镰状细胞突变。通过核穿孔将复合物引入到镰状iPSC中,并且2天后通过数字PCR(ddPCR)分析基因组DNA并对其进行测序。超过65%的细胞含有至少一种校正基因。如下证实结果。在引入校正复合物两天后,将细胞铺在培养皿中并分离出43个单个iPSC集落。从这些集落中分离基因组DNA,并对 β -珠蛋白基因进行测序。百分之六十五的集落含有至少一种校正的 β -珠蛋白基因(S校正为A)。

[0015] 图9显示,将镰状细胞校正复合物(gRNA修饰的重组Cas9-ssODN)引入到患者原代骨髓CD34+细胞可校正镰状细胞突变。在体外分化十二天后,通过数字PCR(ddPCR)分析DNA并对其进行测序。观察到近似等量的 β A和 β S mRNA。

[0016] 图10是来自图9的校正镰状患者CD34+细胞的体外分化红细胞的等电聚焦(IEF)凝胶,显示HbA(正常血红蛋白)与HbS(具有镰状细胞突变的血红蛋白)比率为约1:3,其足以抑制镰状化并治疗镰状细胞性贫血。

[0017] 详细描述

[0018] 本文提供用于细胞的基因组修饰的CRISPR/Cas9复合物。使用本文提供的复合物的方法导致修饰效率增加、细胞存活率增加和/或细胞中HDR与NHEJ的比率增加。这些复合物和方法可用于治疗目的,例如校正细胞中的突变,其中突变与疾病或病症相关联。

[0019] 本文提供一种用于校正细胞基因组中的突变的复合物,其包含:(a)指导RNA(gRNA),其包含:与细胞基因组中的靶DNA杂交的第一核苷酸序列,其中靶DNA包含突变;以及与定点核酸酶相互作用的第二核苷酸序列;(b)重组定点核酸酶,其可操作地连接至超荷电蛋白,其中定点核酸酶包含与指导RNA的第二核苷酸序列相互作用的RNA结合部分,并且其中定点核酸酶特异性结合并切割靶DNA以产生双链断裂;以及(c)单链供体寡核苷酸(ssODN),其与靶DNA中的双链断裂侧翼的基因组序列杂交并整合到靶DNA中以校正靶DNA中的突变。

[0020] 应当理解,本文所述的包含指导RNA(gRNA)、重组定点核酸酶和供体核苷酸的复合物在自然界中不存在。然而,所述复合物提供有必要具有所需构型和化学计量的元件,以高效和有效地修饰细胞。gRNA分子结合至定点核酸酶并将核酸酶靶向靶DNA内的特定位置。gRNA包含:与细胞基因组中的靶DNA杂交的第一核苷酸序列,其中靶DNA包含突变;以及与定点核酸酶相互作用的第二核苷酸序列。本文所述的复合物可包含一个或两个独立的gRNA。因此,术语指导RNA包括单指导RNA和双指导RNA。可用于校正与镰状细胞性贫血相关联的突变的指导序列的实例在本文中陈述为TAACGGCAGACTTCTCCAC(SEQ ID NO:1)。包含用于Cas9结合的茎环的指导序列的实例在本文中提供为GTAACGGCAGACTTCTCCACGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGTCTTTTTT(SEQ ID NO:2)。应注意,在体外转录期间,通过T7添加了SEQ ID NO:2的5'G。

[0021] 在本文所述的复合物中,重组定点核酸酶可以是RNA指导的定点核酸酶,例如来自任何细菌物种的Cas蛋白或其功能片段。例如,Cas蛋白可以是Cas9蛋白或其功能片段。如本文所用,术语“Cas9”意指存在于编码II型CRISPR/Cas9系统的任何细菌物种中的Cas9蛋白或其片段。参见例如Makarova等人.Nature Reviews, Microbiology, 9:467-477(2011)(包括补充信息),其特此以引用的方式整体并入。例如,Cas9蛋白或其片段可来自酿脓链球菌

(*Streptococcus pyogenes*)。全长Cas9是包含识别结构域和两个核酸酶结构域(分别是HNH和RuvC)的内切核酸酶。在氨基酸序列中,HNH是线性连续的,而RuvC分成三个区域,一个在识别结构域的左侧,并且其他两个识别结构域的右侧、在HNH结构域侧翼。来自酿脓链球菌的Cas9通过与指导RNA相互作用而靶向细胞中的基因组位点,所述指导RNA与紧接于Cas9识别的NGG基序之前的20个核苷酸的DNA序列杂交。这导致由供体核苷酸(例如,ssODN或双链DNA构建体)通过HDR修复的双链断裂,所述供体核苷酸与靶DNA中的双链断裂侧翼的基因组序列杂交并整合到靶DNA中以校正靶DNA中的突变。

[0022] 在本文提供的复合物中,gRNA与可操作地连接至超荷电蛋白的定点核酸酶与ssODN的摩尔比可以为约1:1:0.2至约1.5:1:2.0。例如,gRNA与可操作地连接至超荷电蛋白的定点核酸酶与ssODN的摩尔比可以为约1:1:1、1.1:1:1、1.1:1:1.15、1.1:1:1.25、1.1:1:1.30、1:1:1.35、1.1:1:1.40、1.1:1:1.50、1.2:1:1、1.3:1:1、1.4:1:1、1.5:1:1、1.5:1:1.15、1.5:1:1.25、1.5:1:1.35、1.5:1:1.40、1.5:1:1.45、1.5:1:1.50、1.5:1:1.55、1.5:1:1.60、1.5:1:1.65、1.5:1:1.70、1.5:1:1.75、1.5:1:1.80、1.5:1:1.85、1.5:1:1.90、1.5:1:1.95、1.5:1:2.0或这些比率之间的任何比率。具有这些摩尔比的复合物可用于本文所述的任一种方法中。用于在将复合物引入到细胞或细胞群体之前制备复合物的方法在实施例中进行陈述。

[0023] 如本文所用,超荷电蛋白可以是总体正电荷大于其对应未修饰蛋白的超荷正电蛋白。例如,超荷正电蛋白可以是总体正电荷为约+5至约+40的超荷正电绿色荧光蛋白(GFP)。例如,总体正电荷可以为约+5、+6、+7、+8、+9、+10、+11、+12、+13、+14、+15、+16、+17、+18、+19、+20、+21、+22、+23、+24、+25、+26、+27、+28、+29、+30、+31、+32、+33、+34、+35、+36、+37、+38、+39或+40。

[0024] 超荷电蛋白可以可操作地连接至核酸酶的氨基末端或羧基末端。还预期超荷电蛋白可与核酸酶缔合,而不一定共价连接至核酸酶。超荷电蛋白的实例是超荷正电的GFP,例如+36GFP。+36GFP可以可操作地连接至Cas9的氨基或羧基末端或其功能片段。参见例如McNaughton等人,“Mammalian cell penetration, siRNA transfection, and DNA transfection by supercharged proteins,”PNAS 106 (15):6111-6116。包含可操作地连接至Cas9羧基末端的+36GFP的多肽的实例在本文中提供为SEQ ID NO:3。

[0025] 核酸酶还可以可操作地连接至超荷电蛋白,并且一个或多个荷正电肽(例如,一个或多个反式激活转录激活因子(TAT)肽)可连接至核酸酶的氨基末端或羧基末端。例如但不限于,超荷正电蛋白可以可操作地连接至核酸酶的羧基末端,并且一种或多种TAT肽(例如,1X TAT、2X TAT、3X TAT、4X TAT等)可以可操作地连接至核酸酶的氨基末端。包含可操作地连接至核酸酶的氨基末端的TAT肽和可操作地连接至核酸酶的羧基末端的超荷正电GFP的多肽的实例在本文中提供为SEQ ID NO:4。还提供与SEQ ID NO:3或SEQ ID NO:4至少约75%相同的多肽序列。例如,还提供至少约75%、80%、85%、90%、95%、99%或其间任何百分比的多肽序列。

[0026] 核酸酶还可以可操作地连接至超荷电蛋白,并且一个或多个荷负电肽(例如,长度为约10个至约25个氨基酸的荷负电肽,例如SEQ ID NO:50)可以可操作地连接至定点核酸酶的羧基末端。例如但不限于,超荷正电蛋白可以可操作地连接至核酸酶的羧基末端,并且荷负电肽可以可操作地连接至超荷正电蛋白的羧基末端。

[0027] 如全文所用,重组是两个多核苷酸之间交换遗传信息的过程。同源性定向修复

(HDR)是指例如在细胞中双链断裂的修复期间发生的DNA修复。这个过程需要核苷酸序列同源性,并使用供体分子(例如,单链或双链核苷酸序列)作为用于修复靶基因组序列(即具有双链断裂的基因组序列)的模板,并使得遗传信息从供体转移至靶基因组序列。同源性定向修复可导致对靶基因组序列的序列的修饰。例如,HDR可导致靶基因组序列中的插入、缺失或突变。可将供体多核苷酸的部分或全部序列掺入到靶DNA中。还预期将供体多核苷酸、供体多核苷酸的一部分、供体多核苷酸的拷贝或供体多核苷酸拷贝的一部分整合到靶DNA中。

[0028] 如全文所用,非同源性末端接合(NHEJ)意指在不需要同源模板(与需要同源序列来指导修复的同源性定向修复相比)的情况下通过将断裂末端直接彼此连接来修复DNA中的双链断裂。

[0029] 本文提供的复合物和方法可用于通过HDR校正靶DNA中的任何突变。例如但不限于,所述复合物可用于在基因组中的特定位置处用正确的核苷酸序列替换不正确的核苷酸序列(例如,以恢复由于功能丢失突变(即SNP)而受损的靶多核苷酸序列的功能)。这些突变可与自身免疫性疾病、遗传性疾病、血液病症、T细胞病症、单基因病症、癌症、神经退行性疾病、心血管疾病或感染性疾病相关联。例如但不限于,本文提供的复合物和方法可用于校正与以下项相关联的突变:镰状细胞疾病(即,血红蛋白基因中的突变,例如导致谷氨酸至缬氨酸替代的 β -珠蛋白基因密码子6处的GAG至GTG突变)、重症联合免疫缺陷(SCID)(例如,JAK3中的突变)、 β 地中海贫血或威斯科特-奥尔德里奇综合征。

[0030] 单一突变或多个突变的校正可用一种或多种复合物进行。本文提供的复合物和方法还可用于将序列插入到基因组中的特定位置中以校正缺失,而不是进行校正或替代。本文提供的复合物和方法还可用于将编码功能性多肽的核苷酸序列插入到细胞基因组中的特定位置中,以便在细胞中表达功能性多肽。功能性多肽可以是细胞内源性多肽(即,通常由细胞表达)或外源性多肽(即,通常不由细胞表达)。例如,可将嵌合抗原受体(CAR)序列插入到T细胞前体的基因组中,以便生成用于治疗癌症的癌症特异性T细胞。在另一个实例中,本文提供的复合物和方法可用于抑制细胞基因组中特定位置处的基因的活性。例如,本文提供的复合物和方法可用于将序列插入到CXCR4或CCR5受体中以治疗或预防HIV感染。

[0031] 与常规的CRISPR/Cas系统相比,本文提供的复合物可以惊人的高效率修饰或改变靶DNA。细胞群体中的改变效率可以为至少约5%、10%、15%、20%、25%、30%、35%、40%、45%、50%、55%、60%、65%、70%、75%或80%或更高或这些百分比之间的任何百分比。改变的效率也可大于或等于约80%。因此,本文还提供细胞群体,其中至少5%、10%、15%、20%、25%、30%、35%、40%、45%、50%、55%、60%、65%、70%、75%或80%或更高或其间的任何百分比发生改变。例如,已经校正了与镰状细胞疾病或另一种病症相关联的突变。如果将包含与镰状细胞疾病相关联的突变的细胞群体与本文所述的CRISPR/Cas复合物接触,并且在约5%的细胞中校正突变,则修饰或改变的效率为约5%。任选地,在移植到患有镰状细胞疾病的受试者体内后,其中在约30%的细胞(包括例如27%、28%和29%)中校正了与镰状细胞疾病相关联的突变的细胞群体足以治疗镰状细胞疾病。任选地,在群体中约40%、50%、60%、70%、80%、90%或更高或其间任何百分比的细胞中校正了与镰状细胞疾病相关联的突变。

[0032] 除了以高效率改变靶DNA之外,本文提供的复合物还可增加与复合物接触的细胞群体中HDR与NHEJ的比率。HDR/NHEJ比率可以为约10至约0.5。例如,HDR/NHEJ比率可以为约

10、9、8、7、6、5、4、3、2、1、0.5或更小或这些比率之间的任何比率。除了校正的高效率和HDR与NHEJ的高比率之外,校正型细胞的细胞存活率可以为至少约50%、60%、70%、80%、90%或更高以及其间的任何百分比。

[0033] 可使用本文所述的复合物修饰或衍生任何细胞。将复合物引入到细胞中可以是细胞周期依赖性或细胞周期独立性的。使细胞同步化以增加特定时期(例如,S期)中的细胞比例的方法是本领域已知的。参见例如Takahashi等人。“Efficient introduction of a gene into hematopoietic cells in S-phase by electroporation,”*Exp.Hematol.*19(5):343-346(1991)。根据待修饰的细胞类型,本领域技术人员可容易地确定细胞周期同步化是否必要。

[0034] 细胞可以是真核细胞,例如哺乳动物细胞。细胞也可以是原核细胞或植物细胞。细胞可以是人类细胞。细胞可以是生殖细胞、体细胞、干细胞、前体细胞或祖细胞。前体细胞可以是例如多能干细胞或多能性干细胞,如造血干细胞。如全文所用,多能细胞包括诱导性多能干细胞。制备诱导性多能干细胞的方法是本领域已知的并在实施例中进行描述。细胞也可以是任选地衍生自诱导性多能干细胞的CD34+细胞。CD34+细胞可选自由原代CD34+造血祖细胞、CD34+外周血细胞、CD34+脐血细胞和CD34+骨髓细胞组成的组。细胞也可以是原代细胞,例如原代CD34+造血祖细胞。细胞是非癌细胞的细胞、非肿瘤细胞的细胞或非转化细胞的细胞。可在校正之前或之后针对不期望的遗传特征的迹象对细胞进行筛查。还提供一种包含本文所述的任何复合物的细胞。细胞可以是体外、离体或体内的。

[0035] 还提供一种位点特异性修饰细胞群体中的靶DNA的方法,其包括将本文所述的任何复合物引入到细胞中,其中在允许同源性定向修复(HDR)和将供体核苷酸(例如,ssODN或双链核苷酸序列)整合到靶DNA中的条件下将复合物引入到细胞中。可通过核穿孔将复合物引入到细胞中。用于核穿孔的方法是本领域已知的。参见例如Maasho等人“Efficient gene transfer into the human natural killer cell line,NKL,using the amaxa nucleofection system,”*Journal of Immunological Methods* 284(1-2):133-140(2004)和Aluigi等人“Nucleofection is an efficient non-viral transduction technique for human bone marrow derived mesenchymal stem cells,”*Stem Cells* 24(2):454-461(2006)),所述两篇文献均以引用的方式整体并入本文。

[0036] 在本文提供的一些方法中,供体核苷酸(例如,ssODN或双链核苷酸序列)整合到靶DNA中并校正靶DNA中的突变。在本文提供的方法中,细胞群体中HDR与NHEJ的比率相对于其他CRISPR-Cas9递送方法增加。HDR/NHEJ比率可以为约10至约0.5。例如,HDR/NHEJ比率可以为约10、9、8、7、6、5、4、3、2、1、0.5或更小或这些比率之间的任何比率。在本文提供的方法中,通过HDR改变的效率可以为至少约5%、10%、15%、20%、25%、30%、35%、40%、45%、50%、55%、60%、65%、70%、75%、80%或更大或这些百分比之间的任何百分比。通过HDR改变的效率也可大于或等于约80%。例如,如果将包含与镰状细胞性贫血相关联的突变的细胞群体与本文所述的CRISPR/Cas复合物接触,并且在约5%的细胞中校正突变,则通过HDR改变的效率为约5%。细胞群体可从患有病症的受试者获得,使得至少约5%、10%、15%、20%、25%、30%、35%、40%、45%、50%、55%、60%、65%、70%、75%或80%或更大或这些百分比之间的任何百分比的细胞经受HDR,以校正与病症相关联的突变。在一些情况下,来自受试者的大于80%的细胞将经受HDR,以校正与病症相关联的突变。在本文所述的

方法中,在引入复合物之后,约50%与99%之间的细胞存活。例如,在引入复合物之后,大于约50%、60%、70%、80%、90%、95%、99%或这些百分比之间的任何百分比的校正型细胞存活。

[0037] 还提供一种治疗受试者的与编码血红蛋白的基因组序列中的突变相关联的疾病的方法,其包括:(a)将包含以下项的复合物引入到从受试者获得的细胞群体中:(1)指导RNA(gRNA),其包含:与细胞基因组中的靶DNA杂交的第一核苷酸序列,其中靶DNA是包含突变的血红蛋白基因;以及与定点核酸酶相互作用的第二核苷酸序列;(2)重组定点核酸酶,其可操作地连接至超荷电蛋白,其中定点核酸酶包含与指导RNA的第二核苷酸序列相互作用的RNA结合部分,并且其中定点核酸酶特异性结合并切割靶DNA以产生双链断裂;以及(3)单链供体寡核苷酸(ssODN),其与靶DNA中的双链断裂侧翼的基因组序列杂交并整合到靶DNA中以校正血红蛋白基因中的突变;以及(b)将校正型细胞移植到受试者体内。

[0038] 在用于治疗受试者的与编码血红蛋白的基因组序列中的突变相关联的疾病(例如,镰状细胞性贫血)的方法中,患有镰状细胞性贫血的受试者可任选地是输血依赖性受试者或患有至少一种无症状梗塞的受试者。受试者的年龄也可小于约十二个月、十一个月、十个月、九个月、八个月、七个月、六个月、五个月、四个月、三个月、两个月或一个月。因为对婴儿例行筛查镰状细胞疾病,所以可在疾病症状出现之前对婴儿进行治疗。本文提供的方法还可包括诊断患有病症(例如,镰状细胞疾病)的受试者。

[0039] 如上所述,可从患有疾病的受试者或从相关供体获得细胞。例如,可从受试者获得或采集骨髓细胞。骨髓采集包括用置入骨头的软中心即髓的针收集干细胞。骨髓可例如从受试者的髌骨或胸骨采集。可从受试者获得约500ml至约1升的骨髓。

[0040] 在本文提供的任何方法中,细胞可以是真核细胞,例如人类细胞。细胞可以是生殖细胞、干细胞、前体细胞。前体细胞可以是例如多能干细胞或造血干细胞。如全文所用,多能细胞包括诱导性多能干细胞。制备诱导性多能干细胞的方法是本领域已知的并在实施例中进行描述。细胞也可以是CD34+细胞。CD34+细胞可选自自由原代CD34+造血祖细胞、CD34+外周血细胞、CD34+脐血细胞和CD34+骨髓细胞组成的组。细胞也可以是原代细胞,例如原代CD34+造血祖细胞。细胞是非癌细胞的细胞、非肿瘤细胞的细胞或非转化细胞的细胞。细胞可以是体外或离体的。细胞也可处于药学上可接受的组合物中。

[0041] 本文提供的方法还可包括培养用HDR校正的细胞。例如,可在扩增条件下或在促进校正型细胞分化成T细胞的条件下培养细胞。例如但不限于,使用本文提供的方法,在通过HDR在诱导性多能干细胞中校正突变后,可将校正型细胞与人类骨髓基质细胞共培养以生成CD34+细胞。然后可在使CD34+细胞分化成T细胞的条件下培养CD34+细胞。

[0042] 本文提供的方法还可包括在移植之前针对适当校正、其他突变或NEJ筛查校正型细胞。任选地,可筛查细胞以检测具有一种或多种校正的细胞。

[0043] 在本文提供的方法中,可在修饰之后,例如在通过HDR校正突变之后,将细胞移植到受试者体内。可在分化或未分化的情况下将细胞移植到受试者体内。例如,修饰的造血干细胞(HSC)可在骨髓移植中施用,其中允许HSC在受试者体内分化和成熟可替代地,可在移植之前将修饰的细胞分化成期望的细胞群体。

[0044] 如本文所用,将细胞移植、引入或施用给受试者是指将细胞置入受试者体内。例如,本文所述的包含根据本文所述方法校正或修饰的靶DNA序列的细胞可通过适当的途径

移植到受试者体内,所述途径使得移植细胞至少部分地定位在期望位点。细胞可直接植入到期望位点,或者可通过使得递送至受试者体内的期望位置的任何适当途径施用,在所述期望位点植入细胞的至少一部分保持活力。例如,可通过静脉内输注全身施用细胞。施用给受试者之后细胞的存活期可短达数小时(例如,二十四小时至数天)至长达若干年。

[0045] 对于离体方法,细胞可以是自体细胞,即取自需要修饰一种或多种细胞中的靶DNA的受试者的一种或多种细胞(即,供体和受体是相同个体)。如本文所述,修饰可以是例如校正突变、插入抑制蛋白质活性的序列或插入增加蛋白质表达的序列,例如插入编码可用于靶向癌细胞的嵌合抗原受体的序列。自体细胞可用于避免可导致细胞排斥的免疫反应。换句话说,当使用自体细胞时,供体和受体是相同受试者。可替代地,细胞可以是异源的,例如取自供体,优选为相关供体。第二受试者可以是相同或不同的物种。通常,当细胞来自供体时,它们将来自与受体充分免疫相容的供体,以减少移植排斥的可能性和/或减少对免疫抑制疗法的需要。细胞也可获自异种来源,即已经基因工程化以与受体或受体物种充分免疫相容的非人类哺乳动物。治疗本文描述的病症的任何方法还可包括向受试者施用一种或多种免疫抑制剂。

[0046] 在包括移植的方法中,在移植本文所述的任何细胞之前,受试者任选地经受骨髓性疗法。骨髓性疗法可包括施用一个或多个剂量的化学疗法、放射疗法或两者,其导致健康骨髓细胞严重或完全耗尽。在另一个实例中,受试者可经受亚骨髓性疗法,其包括施用耗尽健康骨髓细胞的一部分的一个或多个剂量的化学疗法、放射疗法或两者。也可将细胞移植到已经受非消融性化学疗法的受试者体内。例如,可将细胞移植到已经用白消安、氟达拉滨和/或苏消安治疗过的受试者体内。

[0047] 在包括移植的方法中,向受试者施用有效剂量或有效量的校正型细胞。术语有效量和有效剂量可互换使用。术语有效量定义为产生期望生理反应所需的任何量。在一些方法中,可施用约 1×10^6 至约 7×10^6 个校正型细胞/kg,但是这个量可根据相关联的病症而变化。校正型细胞的百分比用于施用细胞的有效量和时间表可凭经验确定,并且进行此类确定在本领域技术范围内。施用的剂量范围是大得足以产生期望作用(例如,治疗疾病,例如镰状细胞性贫血)的那些剂量范围。剂量不应太大,以致于导致大量的不良副作用,诸如不希望的交叉反应、过敏反应等。通常,剂量将随着年龄、病状、性别、疾病类型、疾病或病症的程度、施用途径或者方案中是否包括其他药物而变化,并且可由本领域技术人员确定。在任何禁忌症的情况下,可由个体医师调整剂量。剂量可变化,并且可将药剂以每天一次或多次剂量施用进行施用,根据需要持续一天或多天。

[0048] 如全文所用,受试者可以是脊椎动物,更具体地是哺乳动物(例如,人、马、猫、狗、牛、猪、绵羊、山羊、小鼠、兔、大鼠和豚鼠)。所述术语不指示特定年龄或性别。因此,意图涵盖成年和新生的受试者,无论是男性还是女性。如本文所用,患者或受试者可互换使用,并且可以指患有病症或处于患上病症的风险中的受试者。术语患者或受试者包括人类和兽类受试者。

[0049] 如本文所用,术语治疗(treatment/treat/treating)是指通过引发受试者体内的免疫反应来减少病症的一种或多种作用或病症(例如,镰状细胞疾病)的一种或多种症状的方法。因此,在所公开的方法中,治疗可以指镰状细胞疾病或其他病症的严重程度降低10%、20%、30%、40%、50%、60%、70%、80%、90%或100%。例如,如果与对照相比,受试

者感染的一种或多种症状减少10%，则认为用于治疗镰状细胞疾病的方法是一种疗法。因此，与固有或对照水平相比，降低可以是10%、20%、30%、40%、50%、60%、70%、80%、90%、100%或10%与100%之间的任何百分比的降低。应当理解，治疗不一定是指治愈或完全消除病症或病症的症状。

[0050] 还提供一种校正与T细胞病症相关联的突变的方法，其包括将包含以下项的复合物引入到从患有T细胞病症的受试者获得的细胞群体中：(a) 指导RNA (gRNA)，其包含：与细胞基因组中的靶DNA杂交的第一核苷酸序列，其中靶DNA包含与T细胞病症相关联的突变；以及与定点核酸酶相互作用的第二核苷酸序列；(b) 重组定点核酸酶，其可操作地连接至超荷电蛋白，其中定点核酸酶包含与gRNA的第二核苷酸序列相互作用的RNA结合部分，并且其中定点核酸酶特异性结合并切割包含与T细胞病症相关联的突变的靶DNA以在靶DNA中产生双链断裂；以及(c) 单链供体寡核苷酸 (ssODN)，其包含与靶DNA中的双链断裂侧翼的基因组序列杂交并整合到靶DNA中以校正与T细胞病症相关联的突变的第三核苷酸序列，其中在允许同源定向修复 (HDR) 的条件下将复合物引入到细胞中以校正与T细胞病症相关联的突变。

[0051] 在本文提供的方法中，包含与T细胞病症相关联的突变的靶DNA可以是编码与T淋巴细胞发育相关联的蛋白质的靶DNA。例如，靶DNA可编码JAK3。此类校正型细胞可用于例如治疗SCID。

[0052] 除了校正细胞基因组中的突变之外，本文提供的复合物和方法还可用于在细胞基因组的特定位点处插入功能性多肽，使得细胞表达多肽。可在细胞中或细胞表面上表达多肽。

[0053] 还提供一种制备肿瘤特异性T细胞前体细胞的方法，其包括将包含以下项的复合物引入到T细胞前体细胞群体中：(a) 指导序列 (gRNA)，其包含与T细胞前体细胞基因组中的靶DNA杂交的第一核苷酸序列和与定点核酸酶相互作用的第二核苷酸序列；(b) 重组定点核酸酶，其可操作地连接至超荷电蛋白，其中定点核酸酶包含与gRNA的第二核苷酸序列相互作用的RNA结合部分，并且其中定点核酸酶特异性结合并切割靶DNA以产生双链断裂；以及(c) 供体核苷酸序列，其包含编码嵌合抗原受体 (CAR) 的第三核苷酸序列和与靶DNA中的双链断裂侧翼的基因组序列杂交的第四核苷酸序列，其中在允许同源定向修复 (HDR) 和将第三核苷酸序列整合到靶DNA中的条件下将复合物引入到T细胞前体细胞中以形成表达CAR的修饰的T细胞前体细胞。

[0054] T细胞前体细胞可从患有癌症的受试者获得。如上所述，HDR/NHEJ比率可以为约10至约0.5。例如，HDR/NHEJ比率可以为约10、9、8、7、6、5、4、3、2、1、0.5或这些比率之间的任何比率。在本文提供的方法中，通过HDR改变的效率可以为至少约5%、10%、15%、20%、25%、30%、35%、40%、45%、50%、55%、60%、65%、70%、75%、80%或这些百分比之间的任何百分比。通过HDR改变的效率也可大于或等于约80%。例如，当使用本文所述的方法时，如果将编码功能性多肽的核苷酸序列（例如，编码CAR的核苷酸序列）插入约5%的细胞中时，则通过HDR改变的效率为约5%。细胞群体可从患有癌症的受试者获得，使得至少约5%、10%、15%、20%、25%、30%、35%、40%、45%、50%、55%、60%、65%、70%、75%或80%或这些百分比之间的任何百分比的细胞经受HDR，以插入编码嵌合抗原受体 (CAR) 的核苷酸序列并形成表达CAR的细胞。在一些情况下，来自受试者的大于80%的细胞将经受HDR，以校正与病症相关联的突变。

[0055] 表达CAR的修饰的T细胞前体细胞可移植到患有癌症的受试者体内。如本文所用，癌症是特征在于异常细胞的快速且不受控制的生长的疾病。癌细胞可局部扩散，或通过血流和淋巴系统扩散到身体的其他部位。癌症的实例包括但不限于乳腺癌、前列腺癌、卵巢癌、宫颈癌、皮肤癌、胰腺癌、结肠直肠癌、肾癌、肝癌、脑癌、淋巴瘤、白血病、肺癌等。当抗原结合结构域结合至其对应的抗原时，表达CAR的修饰的T细胞前体细胞表现出抗肿瘤免疫性。

[0056] 所公开的材料、组合物以及组分可用于所公开的方法和组合物，可与所公开的方法和组合物联用，可用来制备所公开的组合物，或者是所公开的方法和组合物的产品。本文公开了这些材料和其他材料，应当理解，如果公开了这些材料的组合、子集、相互作用、群组等，而没有明确地公开对各种个体和集体组合以及这些化合物的排列的具体提及，则每种情况均得到本文的明确涵盖和描述。例如，如果公开并讨论了方法，并且讨论可对多种分子（包括所述方法）进行的多重修改，则除非明确相反地指示，否则明确涵盖所述方法的每一种组合和排列以及可能的修改。同样地，还具体地涵盖并公开这些修改的任何子集或组合。这个概念适用于本公开的所有方面，包括但不限于使用所公开的组合物方法中的步骤。因此，如果有多种另外的步骤可进行，那么应当理解，这些另外的步骤中的每一个均可与所公开方法的任何具体方法步骤或方法步骤的组合一起进行，并且每种这样的组合或组合的子集均得到明确涵盖并应视为得到公开。

[0057] 本文所引用的出版物以及引用这些出版物所涉及到的材料特此以引用的方式明确地整体并入。

实施例

[0058] 实施例1

[0059] 通过CRISPR/Cas9增强基因置换校正SCID

[0060] Janus家族激酶JAK3基因的突变导致重症联合免疫缺陷 (SCID)。人类JAK3缺陷的特征在于缺乏循环的T细胞和自然杀伤 (NK) 细胞、具有正常数量的功能低下的B细胞 (T-B+ NK-)。如本文所示，使用SCID患者特异性诱导性多能干细胞 (iPSC) 和T细胞体外分化系统，证实JAK3缺陷型细胞的早期T细胞发育的完整阻断。通过CRISPR/Cas9增强基因置换校正新型JAK3突变恢复了正常的T细胞发育，包括产生具有广泛的T细胞受体 (TCR) 库的成熟T细胞群体。校正型细胞的全基因组测序证实没有CRISPR/Cas9脱靶修饰。因此，本文提供一种用于研究人类淋巴细胞生成的新方法和一种用于具有免疫缺陷的人类中的基因置换疗法的方法。

[0061] 异基因造血干细胞 (HSC) 移植是目前建立的唯一的SCID疗法；然而，免疫恢复延迟和移植物抗宿主疾病风险存在相当大的危险。对于X连锁的SCID，已经成功地证明基于逆转录病毒的基因疗法进行的治疗。然而，在逆转录病毒基因疗法的情况下已经观察到插入诱变的严重不良作用。自身灭活的慢病毒载体已被有效地用于最近的临床试验，但需要长期随访以彻底解决安全问题。

[0062] 本文提供一种替代治疗策略，其中衍生得到患者特异性诱导性多能干细胞 (iPSC)，并且通过使用CRISPR-Cas9复合物进行基因置换来校正致病性突变。这些校正的iPSC可任选地分化成用于移植到患者体内以治疗疾病的造血祖细胞 (Hanna等人，

“Treatment of sickle cell anemia mouse model with iPS cells generated from autologous skin,” *Science* 318:1920-1923 (2007))。如本文所示, JAK3缺陷型人类T细胞的分化在早期发育阶段被阻断。还证实通过CRISPR/Cas9增强基因置换校正人类JAK3突变恢复了早期T细胞祖细胞的分化潜能。这些校正的祖细胞能够产生表达广泛的T细胞抗原受体 (TCR) 库的NK细胞和成熟T细胞群体。这些研究建立用于确定人类SCID患者的免疫缺陷机制以及测试病症的药理学和基因疗法的强有力系统。

[0063] 患者信息

[0064] 根据赫尔辛基宣言, 将男性患者招入机构审查委员会批准的研究中。家族史为免疫缺陷阴性。对于年龄的前8个月, 他体重增加缓慢, 患有腹泻和需要频繁住院的复发性毛细支气管炎。他在8个月大时患有严重的呼吸窘迫和鹅口疮而需住院。用支气管肺泡灌洗液进行的支气管镜检证实细菌生物体 (假单胞菌、流感嗜血杆菌、肺炎链球菌) 和病毒生物体 (呼吸道合胞病毒)。免疫学评估证实严重的低丙球蛋白血症, 其中IgE < 3, IgA < 4, IgG = 29, IgM = 26。外周血的免疫表型分析证实完全缺乏CD3+T细胞和NK细胞, 但是存在B细胞 (绝对B细胞计数 = 875)。有丝分裂原研究证实对伴刀豆球蛋白A、美洲商陆有丝分裂原 (pokeweed mitogen) 和植物凝集素A完全缺乏反应。通过基因测试证实了SCID的诊断, 在JAK3基因外显子14中存在纯合C > T核苷酸替代, 这导致在氨基酸位置613处的精氨酸密码子 (CGA) 被终止密码子 (TGA) 置换。这是将这种JAK3变体 (rs149316157) 与SCID临床案例联系起来的第一份报告。患者经受了降低强度预调理、匹配的非血缘骨髓移植, 并且在停止疗法两年目前很健康, 免疫完全重建。

[0065] 人类iPSC重编程和表征

[0066] 对于iPSC诱导, 将 5×10^4 个原代角质细胞接种到6孔板的一个孔中。在第二天, 用1mL病毒上清液和1mL含有最终浓度为4 μ g/mL的聚凝胺的人类角质细胞培养基转导角质细胞。将角质细胞以800x g旋转感染45分钟 (第1天)。第二天再次重复转导程序。在第3天, 将细胞更换至新鲜的人类角质细胞培养基并再培养两天。在第5天, 将角质细胞用胰蛋白酶消化并转移到预先接种有丝裂霉素C处理的鼠胚胎成纤维细胞 (MEF) 的10cm皿, 并在人类角质细胞培养基中培养。在第7天, 将细胞更换至人类ES培养基并在相同皿中连续培养3-4周。每天更换ES培养基。2-3周后可见潜在的iPSC集落。将这些集落单个挑取出并在MEF上扩增进行分析。为了去除整合的慢病毒和多顺反子序列, 用表达Cre的腺病毒 (rAd-Cre-IE) 感染iPSC。将单个集落挑取出并使用引物gctaattcactcccaagaagacaag (SEQ ID NO:5) 和cttcagcaagccgagtctctg (SEQ ID NO:6) 通过PCR验证Cre介导的floxed序列去除。

[0067] 通过OP9共培养生成CD34+细胞和T细胞

[0068] 程序先前有所描述 (Chang等人, “Broad T-cell receptor repertoire in T-lymphocytes derived from human induced pluripotent stem cells,” *PloS one* 9, e97335 (2014))。在伴有以下修改的情况下使用这种方法。如Ohnuki等人 (Ohnuki M, “Generation and characterization of human induced pluripotent stem cells. *Curr Protoc Stem Cell Biol* 第4章: 第4A单元2 (2009)) 所述用CTK溶液处理6孔板的一个孔中的hiPSC培养物, 以制备小细胞团块。然后将细胞团块转移至10cm板, 所述板预先接种在含有10%FBS、1X青霉素/链霉素和100 μ M单硫代甘油的基于 α -MEM的培养基中的2天龄OP9细胞。每隔一天更换培养基, 并且将细胞培养18天而无分裂。共培养18天后, 通过用解离溶液

(0.15% 胶原酶 IV 和 0.015% 透明质酸酶的 α -MEM 培养基溶液) 处理约 30 分钟、之后再用 0.25% 胰蛋白酶处理 30 分钟来采集细胞。然后在抗 CD34+ 磁珠 (MicroBead 试剂盒; Miltenyi Biotec, Bergisch Gladbach, Germany) 上纯化 CD34+ 细胞。对于 T 细胞分化, 将这些 CD34+ 细胞铺到 OP9-DL4 细胞上, 并与含有 20% FBS、5ng/mL LhFlt3-L、5ng/mL hIL-7 和 10ng/mL hSCF 的 α -MEM 培养基一起培养。每隔一天更换培养基, 并且每 4 天将细胞转移至新的 OP9-DL4 板。

[0069] T 细胞刺激

[0070] 如先前所述 (Chang 等人, 2014), 在通过流式细胞术分析之前, 根据制造商方案, 通过与 CD3/28 珠 (Invitrogen, Carlsbad, CA) 一起孵育 3 天来刺激体外衍生自 hiPSC 的 T 细胞。

[0071] 流式细胞术

[0072] 采集细胞并将其在用 LSRFortessa 细胞分析仪 (BD Bioscience, San Jose, CA) 分析之前洗涤。对于细胞表面染色, 使用碘化丙啶 (PI, Sigma-Aldrich, St. Louis, MO) 来排除死细胞。对于凋亡测定, 首先用细胞表面抗体将采集的细胞染色 30min。用 1X PBS 洗涤一次后, 将细胞重悬于 100 μ L 含有膜联蛋白 V-647 (Invitrogen, Carlsbad, CA) 和 PI 的膜联蛋白结合缓冲液 (Invitrogen, Carlsbad, CA) 中, 并孵育 15min, 然后添加 400 μ L 含 PI 的膜联蛋白结合缓冲液。除非另外指示, 否则从 BD Biosciences 获得抗体: CD3 (Percp-Cy5-5, 克隆 UCHT1)、CD4 (PE-Cy7, 克隆 SK3)、CD7 (APC, BV510, 克隆 M-T701)、CD8 (APC-Cy7, 克隆 SK1)、CD16 (PE, 克隆 B73.1)、CD25 (FITC, 克隆 2A3)、CD34 (PE-Cy7, 克隆 WM59)、CD43 (PE, 克隆 1G10)、CD56-PE (克隆 MY31)、CD69 (FITC, 克隆 L78)、NKG2D-PE (克隆 1D11)、TCR- $\alpha\beta$ (FITC, PE, 克隆 T10B9.1A-31)、TCR-V δ 1-FITC (Fisher Scientific, Pittsburgh, PA, 克隆 TS8.2)、TCR-V δ 2-PE (克隆 B6)、TCRV γ 9-FITC (克隆 B3)、TNF- α -PE-Cy7 (克隆 MAB11)、Beta Mark TCR 试剂盒 (Beckman Coulter, Atlanta, GA)。

[0073] 载体构建

[0074] 先前描述了多顺反子 OSKM 载体 (Chang 等人, “Polycistronic lentiviral vector for “hit and run” reprogramming of adult skin fibroblasts to induced pluripotent stem cells,” Stem cells 27:1042-1049 (2009))。通过将 PCR 扩增的人类 DL4 cDNA (Open Biosystems, LaFayette, CO)、IRES 片段 (Open Biosystems) 和 mCherry cDNA 克隆到含有 EF1 α 启动子的慢病毒载体 (pDL171) 中来构建 Lenti-hDL4-mCherry 质粒。使用 PrimeStar 聚合酶 (Takara, Mountain View) 进行 PCR 反应。

[0075] 为了构建 CRISPR 质粒, 设计 gRNA 寡核苷酸并遵循 Zhang 实验室方案 (Addgene, Cambridge, MA) 将其引入到 pX330 和 pX335 质粒中。为了构建 JAK3 修复质粒, 使用 JAK3 引物组 (5' 臂: gtcgacgtcgacgctcagtgaaagctgaagtattccttctgcttcacagggcgaccactac (SEQ ID NO: 7) 和 attttaaattcctcccctgaacccttaccaaactcctatgcatactacag (SEQ ID NO: 8); 3' 臂: ttaattaattaattagcatttttaggttcaggttgtagaactagaagagaacaagtca (SEQ ID NO: 9) 和 gtatacgtatacgcatacctggagaggggacaaggctcttgagatgcagggt (SEQ ID NO: 10)) 对野生型人类基因组 DNA 进行 PCR 扩增。用酶 (5' 臂: SalI 和 SmaI; 3' 臂: PacI 和 BstZ17I) 消化后, 将 PCR 产物克隆到含有 LoxP-PGK-Neo-LoxP 片段的质粒中。本研究中使用的所有寡核苷酸均由 Integrated DNA Technologies (IDT, Coralville, IA) 合成。为了构建 BCL2 慢病毒质粒, 使用引物组 (正向: agccaccttaattaagccaccatggcgcacgctgggagaacggggtacgata (SEQ ID NO: 11) 和反向: taacagagagaagttcgtggctccgatcccttgtggcccagatagcaccagggtgat (SEQ ID

NO:12)) 来扩增人类BCL2 cDNA (Open Biosystems) 片段。通过PCR使产物通过2A序列与GFP连接,并将其克隆到pDL171载体中。gRNA-F1 caccGTG AGA TAC AGA TAC AGA CA (SEQ ID NO:13) gRNA-R1 aaacTGT CTG TAT CTG TAT CTC AC (SEQ ID NO:14) gRNA-F2 caccgAAT GAT TTG CCT GGA ATG CC (SEQ ID NO:14) gRNA-R2 aaacGGC ATT CCA GGC AAA TCA Ttc (SEQ ID NO:15) gRNA-F3 caccgCAG CCT AGG CAA AGG CCTGC (SEQ ID NO:16) gRNA-R3 aaacGCA GGC CTT TGC CTAGGC TGc (SEQ ID NO:17) gRNA-F4 caccgTGC CAA CAG AAC TGC CTG AT (SEQ ID NO:18) gRNA-R4 aaacATC AGG CAG TTC TGT TGG Cac (SEQ ID NO:19) gRNA-F5 caccGAC CAG GGT GCA AGT GTG GA (SEQ ID NO:20) gRNA-R5 aaacTCC ACA CTT GCA CCC TGG TC (SEQ ID NO:21) gRNA-F6 caccGCT CCT CAG CCT GGC ATT CA (SEQ ID NO:22) gRNA-R6 aaacTGA ATG CCA GGC TGA GGA GC (SEQ ID NO:23)

[0076] 细胞培养

[0077] 将IPSC在ES细胞培养基中、在衍生自E14.5CF-1胚胎的丝裂霉素C处理的MEF上培养,所述ES细胞培养基由DMEM F-12组成、补充有1X非必需氨基酸、1X青霉素-链霉素、1X L-谷氨酰胺(全部来自Mediatech, Corning, NY)、20% KnockOut血清替代品 (Invitrogen)、2-β ME (Sigma) 和5-10ng/mL bFGF (Invitrogen)。在DermaLife K培养基完整试剂盒 (LifeLine Cell Technology, Frederick, MD) 中培养人类原代角质细胞。OP9细胞购自ATCC,并在含20%FBS和青霉素-链霉素的α-MEM培养基中生长。通过用含有hDL4和mCherry的慢病毒转导OP9细胞来建立OP9-DL4细胞。

[0078] 病毒制备

[0079] 为了制备慢病毒,通过Fugene 6 (Roche, Nutley, NJ)或Promega, Madison, WI)将10μg慢病毒载体、2.5μg包膜质粒 (pMDG) 和7.5μg包装质粒 (pCMBVdR8.9.1) 共转染到5x10⁶个293T细胞中。转染2天后收集含有病毒的上清液,并使其通过0.45μm过滤器。

[0080] 基因靶向

[0081] 用0.25%胰蛋白酶将IPSC处理5分钟以生成单细胞悬液。用1X PBS洗涤两次后,将100至200万个细胞与5μg JAK3修复质粒和5μg pX330-JAK3或pX335-JAK3质粒混合用于核转染(人类干细胞核转染试剂盒,程序A-023, Lonza, Alpharetta, GA) 并铺到MEF上。二至四天后,将含有30μg/mL G418的hES培养基添加至板以选择耐药集落。3至4周后挑取集落并扩增用于基因组DNA提取。对于PCR基因分型,使用5'引物组 (tgctaaagcgcagctccagact (SEQ ID NO:24) 和gtcttcatctcagggctcggct (SEQ ID NO:25)) 和3'引物组 (cctctctgtgcattatggcag (SEQ ID NO:26) 和gccttctatgccttcttg (SEQ ID NO:27))。为了去除Neo选择标记物,用表达Cre的腺病毒 (rAd-Cre-IE) 感染hiPSC。

[0082] RT-PCR

[0083] 用Trizol试剂 (Invitrogen, Carlsbad, CA) 从体外衍生的细胞中分离总RNA。根据制造商说明书使用Superscript第一链合成系统 (Invitrogen), 用0.5至2μg的总RNA合成cDNA。根据制造商说明书,将SYBR Green PCR预混液 (Life Technologies, Carlsbad, CA) 用于qPCR。用于qPCR的引物组是:GAPDH (F:actcctccacctttgacgct (SEQ ID NO:28), R: tcccctcttcaagggctctacatg (SEQ ID NO:29)); PU.1 (F:gtgcaaaatggaagggtttc (SEQ ID NO:30), R:ggagctccgtgaagttgttc (SEQ ID NO:31)); GATA3 (F:tgtttcctttcactggccaca (SEQ ID NO:32), R:aacggcaactggtgaacggta (SEQ ID NO:33)); BCL11B (F:

ggcgatgccagaatagatgccg (SEQ ID NO:34), R:ccaggccacttggctcctctatctccaga (SEQ ID NO:35); RAG1 (F:ccttactgttgagactgcaatattcc (SEQ ID NO:36), R:ctgaagtcccagtatatacttcacac (SEQ ID NO:37)); RAG2 (F:cccagaagcagtaataatcatcgag (SEQ ID NO:38), R:atgtgggatgtagtagatcttgc (SEQ ID NO:39)); pTa (F:gggtcttacctcagcagttac (SEQ ID NO:40), R:cctcacacagtgtgacgcag (SEQ ID NO:41)); BCL2 (F:gactgagtacctgaaccggc (SEQ ID NO:42), R:gggccaaactgagcagagtc (SEQ ID NO:43)); BAX (F:aagaccagggtggttgggac (SEQ ID NO:44), R:gtaagaaaaatgccacgtc (SEQ ID NO:45)); 以及 JAK3 (F:agtcagacgtctggagcttc (SEQ ID NO:46), R:gtgagcagtgaagcatgagtc (SEQ ID NO:47))。所有值都相对于GAPDH表达进行归一化。

[0084] 全基因组测序和分析

[0085] 使用Covaris S2聚焦超声发生器剪切来自iPSC的DNA:使微型管中的130 μ L样品在频率扫描模式下经受两个40秒周期,其中占空比为10%、强度为4和每突发200个循环。使用Agilent 2100生物分析仪进行的DNA芯片(DNA 1000试剂盒;Agilent Technologies,Santa Clara,CA)分析表明平均片段大小为400bp。使用用于Illumina的NEBNext Ultra DNA文库制备试剂盒(NEB#E7370)进行文库制备,并且使用KAPA Illumina文库定量试剂盒(KK4835;KAPA Biosystems,Wilmington,MA)和Applied Biosystems ViiA 7实时PCR系统(Life Technologies)通过qPCR确定最终文库浓度。使用Illumina TruSeq PECluster Kit v3-cBot-HS(PE-401-3001)和Illumina cBot在流动池上产生测序簇。使用Illumina TruSeq SBS Kit v3-HS-200个循环(FC-401-3001)和Illumina HiSeq 2500升级版进行WGS,以生成用于生物信息学分析的2x100单索引配对末端读取。使用EMBOSS fuzznuc软件(v6.6.0.0)(Rice等人,“EMBOSS:the European Molecular Biology Open Software Suite,”Trends in Genetics:TIG 16:276-277(2000)),通过将CRISPR/Cas9指导序列与hg19参考基因组比对并允许最多三个错配来鉴定可能的脱靶位点;对于第一指导序列(GTGAGATACAGATACAGACA)(SEQ ID NO:48)预测了1193个位点,并且对于第二指导序列(AATGATTTGCCTGGAATGCC)(SEQ ID NO:49)预测了257个位点。使用BWA(v0.7.5a)mem算法将每个样品的来自WGS的所有读取都映射到hg19参考基因组(Li和Durbin,“Fast and accurate long-read alignment with Burrows-Wheeler transform,”Bioinformatics 26:589-595(2010)),并使用Picard-tools(v1.100)(<http://picard.sourceforge.net>)去除重复读取。使用GATK(v2.7-2)进行局部重新比对和碱基质量重校准(McKenna等人,“The Genome Analysis Toolkit:a MapReduce framework for analyzing next-generation DNA sequencing data,”Genome research 20:1297-1303(2010))。使用GATK HaplotypeCaller来调用SNV和插入缺失。另外,如GATK最佳实践所述分别重校准SNV和插入缺失,并应用质量过滤器。将为所有四种iPSC样品共有的来自参考基因组的变体从CRISPR/Cas9脱靶分析中排除。使用Bedtools(v2.17.0)(Quinlan和Hall,“BEDTools:a flexible suite of utilities for comparing genomic features,”Bioinformatics 26:841-842(2010))筛查未排除的变体以确定它们是否属于可能的脱靶位点。分析显示这些变体都不存在于脱靶位点,并表示这些突变是随机积累的。然后使用ANNOVAR软件包对具有低等位基因频率(<1%,dbSNP 138)的所有功能变体(排除的和未排除的)进行注释,并筛查与HGMD和ClinVar(v20140902)中的疾病的已知相关性;另外,还使用GWAS目录和COSMIC(v70)对

CADD评分高 (CADD ≥ 20) 的所有命中筛查与复杂疾病的相关性。在所查询的数据库中未鉴定出验证的疾病相关变体。特别感兴趣的是,也未验证出JAK3 C1837T (p.R613X) 突变与疾病相关联,但是预测SNP (rs149316157) 明显是有害的,其中GERP评分为3.85,并且CADD评分 (CADD phred样评分) 为38。因此,首次将JAK3 C1837T变体与SCID的临床案例相关联。

[0086] 登录代码

[0087] 可在NCBI SRA数据库中访问WGS数据,登录号为SRP056149。

[0088] JAK3缺陷型人类T细胞表达低水平的BCL2并在早期发育阶段死亡

[0089] 由对于JAK3基因的外显子14中C>T核苷酸替代纯合的SCID患者的皮肤角质细胞 (Chang等人,2009) 生成iPSC。这种突变使CGA密码子 (613处的精氨酸) 被TGA终止密码子 (p.R613X) 置换。如上所述,4个月大的患者存在T-B+NK-临床表型。为了确定这种SCID表型是否可体外重现,试图使用两步OP9和OP9-DL4系统 (Chang等人,2014) 将患者特异性iPSC分化成T淋巴细胞。JAK3缺陷型iPSC的生长速率与衍生自健康供体的对照iPSC相当,并且这些iPSC在OP9基质细胞单层上高效地分化成CD34+造血祖细胞 (HP)。然而,当将JAK3缺陷型iPSC衍生的CD34+HP铺在OP9-DL4基质单层上时,与对照相比缺乏T-细胞分化 (图1)。未观察到CD3+T细胞或CD3-CD16+CD56+NK细胞 (图1A),并且未检测到CD4+CD8+双阳性 (DP)、CD4+单阳性 (SP) 或CD8+单阳性 (SP) T细胞 (图1B)。Jak3敲除 (KO) 小鼠具有小胸腺,这是由于在生产性TCR重排之前CD4-CD8-双阴性 (DN) 阶段的胸腺细胞分化阻断所致。为了进一步理解由人类JAK3突变导致的发育缺陷,测定了JAK3缺陷型细胞与正常JAK3WT对照相比的T谱系定型和成熟。将IPSC衍生的CD34+细胞铺在OP9-DL4单层上,并且采集细胞并在T细胞诱导第14、21、28和35天 (TD) 分析淋巴细胞标记物 (图4A)。在正常对照中,在TD14时,由 $1-2 \times 10^6$ 个CD34+细胞生成 1.2×10^7 个CD7+细胞 (淋巴门中计数的细胞的84%)。T细胞成熟后,依次检测到T细胞标记物CD4、CD8、CD3和TCR $\alpha\beta$ 。在TD35时,超过50%的群体是CD8 SP细胞。在JAK3缺陷型细胞中,在TD14时,由 $1-2 \times 10^6$ 个CD34+细胞仅生成 4.5×10^4 个CD7+细胞 (淋巴门中计数的细胞的38.9%)。延长培养期间CD7+细胞数量减少,并且T细胞标记物CD3、CD4、CD8和TCR $\alpha\beta$ 未明显表达。在通过早期T细胞祖细胞 (ETP) 转变期间,通过精确激活和抑制特定转录因子来引导CD4-CD8- (DN) 至CD4+CD8+ (DP) 阶段。在对照细胞中,PU.1的沉默以及GATA3和BCL11B的诱导 (图1C) 表明这些细胞继续发生T谱系定型 (DN2至DN3),然后是TCR重排。相比之下,在JAK3缺陷型细胞中,PU.1积累,并且GATA3和BCL11B水平降低 (图1C)。这些数据表明人类JAK3缺陷型细胞在DN2阶段之前或之时停滞,这与T细胞在Jak3 KO小鼠中死亡的阶段相似。有趣的是,人类JAK3缺陷型细胞可表达足够的RAG1、RAG2和PTCRA (图1C) 以进行TCR重排,但细胞存活时间不足以进入这个重要的发育阶段。JAK3缺陷型细胞的淋巴细胞发育中的这些严重缺陷可通过在淋巴祖细胞存活和分化中起重要作用的IL-7信号传导的缺乏来解释。IL-7/JAK3信号传导通过调控凋亡调节因子的BCL2家族来保持胸腺细胞稳态。来自Jak3 KO小鼠的胸腺细胞和外周T细胞部分通过选择性升高促凋亡因子BAX并通过降低抗凋亡因子BCL2的表达而具有高凋亡指数。类似地,在这些研究中,体外衍生的人类JAK3缺陷型细胞的凋亡与对照相比在TD10 (9%相比于2.2%) 和TD17 (7%相比于1.9%) 时增加 (图2A)。与这种表型一致,与对照相比,JAK3缺陷型细胞中的BAX水平增加并且BCL2水平降低 (图2B)。Bc12的强制表达挽救了 γc 缺陷型小鼠中的T细胞 (而非B或NK细胞) 发育 (Kondo等人,Immunity7:155-162 (1997))。移植有表达Bc12的Jak3 KO骨髓细胞的Jak3 KO小鼠还提高外

周T细胞数量(Wen等人, *Molecular and cellular biology* 21:678-689(2001))。为了确定BCL2的过表达是否能挽救人类JAK3缺陷型细胞的T细胞发育缺陷,用含有由EF1a启动子驱动的BCL2-2A-GFP多顺反子的慢病毒转导体外衍生的JAK3缺陷型CD34+细胞。转导后,将CD34+细胞铺在OP9-DL4单层上,并在TD 28时测定NK和T细胞标记物。在GFP-细胞(JAK3-; BCL2低)或GFP+细胞(JAK3-; BCL2+)中未发现CD3-CD16+CD56+NK细胞(图2C)。这些发现表明BCL2在JAK3缺陷型细胞的DN阶段释放阻断。有趣的是,在DP阶段第二次发育停滞明显;在GFP+细胞中未观察到CD8+CD4+DP阳性细胞的进一步分化(图2C)。总之,上述研究证实人类SCID表型可用患者衍生的iPSC体外重现。JAK3缺陷导致DN胸腺细胞的增殖性缺陷。BCL2的强制表达提高DN细胞的存活率,所述DN细胞进一步分化成DP胸腺细胞。然而,DP胸腺细胞不能成熟为SP T细胞,并且这种缺陷可能是由于缺乏IL7/JAK3信号传导所致。

[0090] 通过CRISPR/Cas9增强基因置换校正SCID hiPSC中的JAK3缺陷

[0091] 为了确定在JAK3缺陷型SCID患者细胞中正常T细胞发育是否可恢复,在iPSC中通过CRISPR/Cas9增强基因置换来校正JAK3突变。外显子14的上游和下游的内含子内的6个指导RNA被设计来靶向靠近C1837T突变的wtCas9或nCas9,并且使用校正模板进行基因置换(图3A)。用表达D10A Cas9切口酶和配对指导RNA的两种质粒或表达野生型Cas9和单一指导RNA的单一质粒对IPSC进行核转染。核转染后,细胞在含有G418的培养基中生长2周。挑取单个集落、进行扩增,并通过PCR进行基因分型(图3B顶部)。CRISPR/Cas9介导的JAK3基因校正的效率在图3C中显示。通过Sanger测序进一步验证来自WT Cas9+gRNA#1的三个克隆、来自WT Cas9+gRNA#2的3个克隆以及来自Cas9切口酶+配对的gRNAs#1和#2的6个克隆。在12个测序的克隆中,鉴定了2个纯合校正克隆(来自Cas9切口酶+配对的gRNA#1和#2的1个克隆,以及来自WT Cas9+gRNA#1的1个克隆)和10个杂合校正克隆(图3D)。通过RT-PCR(JAK3 mRNA)(图3B;左下图)和蛋白质印迹(JAK3蛋白)(图3B;右下方)证实JAK3基因表达恢复。

[0092] CRISPR/Cas9定向JAK3校正的特异性

[0093] 脱靶的CRISPR/Cas9定向基因组修饰的可能性引起在人类中使用这种方法进行治疗的担忧。在癌细胞系中,已经描述了通过Cas9-gRNA进行的相对高水平的脱靶诱变。为了确定人类SCID iPSC中CRISPR/Cas9定向JAK3校正的特异性,在基因置换之前和之后进行全基因组测序。对两个杂合校正克隆和一个纯合校正克隆的基因组进行测序。用gRNA#2+野生型Cas9校正两个杂合克隆,并用gRNA#1+gRNA#2+切口酶Cas9(D10A)校正纯合克隆。将20个碱基的CRISPR指导序列映射到人类参考基因组(允许最多至3个错配),以便鉴定可能的脱靶位点。然后在CRISPR/Cas9定向基因置换后的iPSC样品中分析这些位点的变化。对一个纯合校正型iPSC系和两个杂合校正型iPSC系的WGS分析证实没有突变(SNV或插入缺失)被引入到预测的脱靶位点中,从而表明CRISPR/Cas9定向基因置换的强特异性。

[0094] CRISPR/Cas9定向JAK3校正后的T细胞发育恢复

[0095] 为了确定JAK3基因校正后T细胞发育是否恢复,测定了T细胞谱系定型和成熟。T细胞分化依序通过体内观察到的以下中间物:CD34+CD7+T/NK定型阶段;CD7+CD4+CD8-未成熟,SP阶段;CD4+CD8+DP阶段;以及最后的CD3+CD8+TCR $\alpha\beta$ 成熟阶段。成熟T细胞是多克隆的,其响应于有丝分裂原增殖并分泌细胞因子。因此,将JAK3校正型hiPSC在OP9单层上分化成造血祖细胞,并且在抗CD34磁珠上正向选择CD34+细胞。将这些细胞铺在OP9-DL4单层上,并且在TD14、21、28和35时分析非贴壁细胞的淋巴细胞标记物(图4)。与对照细胞类似,在TD14

时,将 $1-2 \times 10^6$ 个CD34+JAK3校正型细胞分化成 4.7×10^6 个CD7+细胞(淋巴门中计数的细胞的91%)。进一步分化至TD21、TD28和TD35后,大量观察到T细胞成熟标记物CD3、CD4、CD8和TCR $\alpha\beta$ (图4A)。为了确定是否在JAK3校正型T细胞中重建TCR重排,通过流式细胞术进行TCR V β 分型,并总结在图4B中。JAK3校正型T细胞表达我们测试的所有V β 片段(25中的19个);因此,广泛的TCR库得到恢复。最后,通过测量抗CD3/CD28刺激后的细胞表面激活标记物来检查JAK3校正型T细胞中TCR信号传导通路(T细胞功能的替代物)的完整性。在刺激后第3天,JAK3校正型T细胞中CD3+CD25+CD69+T细胞的百分比从0.68%增加至59.7%,这与对照细胞中观察到的增加类似(0.01%至37.6%)(图4C)。上述这些数据 and 结果证实,在体外iPSC模型系统中通过CRISPR/Cas9增强基因置换校正JAK3 C1837T(p.R613X)突变恢复了正常的T细胞发育,从而具有产生具有广泛TCR库的功能性成熟T细胞群体的能力。

[0096] 在人类中,虽然SCID患者外周血中淋巴细胞的表型已经得到充分描述,但关于淋巴定型和胸腺细胞发育的关键步骤的研究很难进行。从未经治疗的SCID患者获取骨髓和胸腺细胞样品很具有挑战性,因为这些病状非常罕见,并且婴儿通常存在危及生命的感染,需要紧急HSC移植才能存活。本文描述的用于研究人类SCID的策略绕过这些限制;可在体外由患者特异性iPSC产生大量造血祖细胞,并且可精确确定造成免疫缺陷的机制。本文证实人类JAK3缺陷型SCID中的T细胞发育在CD4-CD8-(DN2)阶段之前或之时被完全阻断。有趣的是,BCL2的强制表达提高DN细胞的存活率,所述DN细胞进一步分化成DP胸腺细胞。然而,DP胸腺细胞不能成熟为SP T细胞,并且这种缺陷可能是由于缺乏IL7/JAK3信号传导所致。还证实通过CRISPR/Cas9增强基因置换校正人类JAK3突变恢复了早期T细胞祖细胞的分化潜能。校正型祖细胞能够产生表达广泛的TCR库的NK细胞和成熟T细胞群体。对一个纯合校正iPSC系和两个杂合校正iPSC系的全基因组测序分析证实没有突变(SNV或插入缺失)被引入到预测的脱靶位点中,从而表明CRISPR/Cas9定向基因置换的强特异性。

[0097] 在本文所述的方法中,可通过与人类骨髓基质干细胞(hMSC)共培养,由hiPSC生成CD34+HSC(参见图5)。可将通过这种方法由基因校正/修饰后的患者特异性iPSC产生的HSC移植回患者体内,以治疗诸如镰状细胞疾病(SCD)、SCID或癌症的疾病。在本文所述的方法中,可通过将hiPSC衍生的CD34+细胞与hMSC-DL4共培养来培养hiPSC衍生的CD34+细胞而生成T细胞(参见图6)。可将通过这种方法由校正/修饰后的患者特异性iPSC产生的HSC移植回患者体内,以治疗疾病。T细胞可包括 $\gamma\delta$ T细胞。如图7所示,可由基因修饰的iPSC高效地产生表达重组T细胞受体(TCR)的 $\gamma\delta$ T细胞。产生表达对肿瘤抗原具有特异性的TCR的 $\gamma\delta$ T细胞为癌症提供细胞疗法。

[0098] 实施例2

[0099] 通过CRISPR/Cas9增强基因置换校正与镰状细胞性贫血相关联的突变

[0100] 载体构建

[0101] 将人类密码子优化的具有N末端和C末端核定位序列(nls-Cas9-nls)的酿脓链球菌Cas9从px330载体(Addgene ID:42230)PCR克隆到N末端具有His6-SUMO标签的修饰的pET-28b(EMD Biosciences)载体中。将含有短接头肽,然后是带有净电荷+36和23个氨基酸的流感病毒血凝素HA-2变体肽INF7INF7(GLFEAIEGFIENGWEGMIDGWYG)(SEQ ID NO:50)的超电荷GFP的基因嵌盒针对大肠杆菌进行密码子优化并且进行合成(IDT DNA)和克隆以与nls-Cas9-nls的C末端融合。还对HIV-TAT肽(YGRKKRRQRRRPPQ)(SEQ ID NO:51)编码序列

进行合成 (IDT DNA) 和克隆以与nls-Cas9-nls的N末端融合。

[0102] 蛋白质过表达和纯化

[0103] 在LB培养基中,将pET-SUMO-scCas9质粒转化到大肠杆菌菌株Rosetta™2 (DE3) 细胞 (EMD Millipore, Billerica, MA) 中。将细胞在37℃下生长,直至600nm处的光密度达到0.6。通过添加0.5mM异丙基-1-硫代-β-D-吡喃半乳糖苷 (IPTG) 并在18℃下在摇动器中培养过夜来实现对蛋白质过表达的诱导。将采集的细胞重悬于Ni-结合缓冲液 (20mM Tris-HCl pH 8.0, 1.5M NaCl, 25mM咪唑和0.2mM TCEP) 中并通过Emulsiflex C3高压均化器 (Avestin) 进行裂解。将最终浓度为0.4%的聚乙烯亚胺 (PEI) 添加到澄清裂解液中以沉淀核酸。然后将离心后上清液中的蛋白质通过硫酸铵沉淀以去除PEI, 并且重新溶解在Ni-结合缓冲液中。首先通过HisTrap镍亲和柱 (GE Healthcare) 纯化蛋白质, 之后在4℃下用SUMO蛋白酶Ulp1过夜消化。然后通过第二HisTrap柱去除切割的His-SUMO标签。将含有scCas9蛋白的通过流稀释以达到0.5M的最终NaCl浓度, 并在HiTrap肝素柱 (GE Healthcare) 上通过用含有20mM Tris-HCl pH 8.0、2.0M NaCl和0.2mM TCEP的缓冲液进行梯度洗脱来纯化。所洗脱的scCas9蛋白通过凝胶过滤缓冲液 (20mM Tris-HCl pH 8.0, 0.5M NaCl和0.2mM TCEP) 中的尺寸排阻柱Superdex 200 16/600 (GE Healthcare) 进一步纯化, 通过穿过0.22μm过滤器进行灭菌, 并通过具有100kDa截留值的Amicon离心单元 (EMD Millipore) 进行浓缩。浓缩的蛋白质通过UV分光光度计定量并在液氮中快速冷冻。

[0104] 指导RNA制备

[0105] 用添加T7启动子和polyA序列的引物组通过PCR生成用于sgRNA转录的模板DNA。根据制造商手册, 使用T7Ribomax Express系统 (Promega, Madison, WI) 通过T7RNA聚合酶体外转录sgRNA。通过苯酚:氯仿提取、乙醇沉淀、之后用MEGAclean™转录清洁试剂盒 (Ambion, Austin, TX) 进行柱纯化来纯化转录的RNA。纯化的gRNA通过UV分光光度计定量并储存在-80℃冷冻机中。

[0106] 单链DNA供体

[0107] 如下所示, 通过IDT DNA合成单链DNA (ssODN) 供体。

用于 HBB 镰状校正的单链供体 DNA	
[0108] HBB-T2-ss ODN	ATCCACGTTACCTTGCCCCACAGGGCAGTAACGGCAGACCTTCTCCtCAGGAGTCAGGTGCACCATGGTGTCTGTTTGA GGTTGCTAGTGA (SEQ ID NO: 52)
HBB-T2-ss ODN- 摆动 (wobble)	CTTCATCCACGTTACCTTGCCCCACAGGGCAGTAACGGCAGAtTTtTCCtCAGGAGTCAGGTGCACCATGGTGTCTGTTT GAGGTTGCTAGTGA (SEQ ID NO: 53)

[0109] 细胞培养

[0110] 人类镰状患者iPSC衍生自皮肤成纤维细胞, 并在含青霉素/链霉素的mTeSR™1培养基 (Stem Cell Technologies, Vancouver, CA) 中保持在Matrigel (BD) 上。

[0111] scCas9-sgRNA-ssODN复合物制备和核转染

[0112] 将1/10体积的10x PBS添加到sgRNA中以达到1x最终浓度。使sgRNA在PCR热循环仪上退火, 温度从95℃缓慢降低至4℃。退火后, 将scCas9蛋白以1:1.5的蛋白质与RNA摩尔比添加到sgRNA中, 并通过轻拍管快速混合, 直至所有瞬时沉淀消失。将混合物在黑暗中在室温下孵育10分钟。然后, 将1摩尔比量的ssODN添加到混合物中并在黑暗中再孵育10分钟以

形成scCas9-sgRNA-ssODN复合物。

[0113] 核转染前一天,通过Accutase (Stem Cell Technologies) 分离细胞,并以 1×10^6 个细胞/孔将细胞接种在具有 $10 \mu\text{M}$ Rock抑制剂(Y-27632) (EMD Millipore) 的6孔板上。对于每项实验,将 5×10^5 个hsIPSC作为单细胞重悬于 $100 \mu\text{l}$ 补充的人类干细胞核转染溶液1 (Lonza) 中,然后将scCas9-sgRNA-ssODN复合物与细胞溶液混合。使用Nucleofector II装置 (Lonza, Basel, Switzerland), 用程序A-023对细胞进行核转染。核转染两天后,通过ddPCR分析HBB基因组校正的效率。

[0114] 通过ddPCR检测镰状校正

[0115] 遵循制造商手册,通过prepGEM组织DNA提取试剂 (ZyGEM, Hamilton, NZ) 裂解用scCas9-sgRNA-ssODN复合物核转染的细胞,并用水进行1:3稀释。在 $22 \mu\text{l}$ ddPCR反应中,将 $11 \mu\text{l}$ 2x ddPCR混合物 (Bio-rad) 与下文所述的 $5 \mu\text{M}$ 等位基因特异性FAM或VIC Taqman探针各 $1 \mu\text{l}$ 、 $100 \mu\text{M}$ 正向引物和反向引物各 $0.2 \mu\text{l}$ 以及 $8.6 \mu\text{l}$ 稀释的基因组DNA混合。根据制造商手册,通过QX200液滴生成器 (Bio-rad, Hercules, CA) 生成液滴。然后将反应混合物转移到96孔PCR板中,并在标准热循环仪 (Bio-rad) 上进行PCR。PCR程序为:步骤1: 95°C 10min; 步骤2: 95°C 30s; 步骤3: 55°C 1min; 将步骤2-3重复39次; 步骤4: 98°C 10min; 步骤5: 8°C 保持。PCR完成后,通过QX200液滴读取器 (Bio-rad) 分析板。

T7-sgRNA 转录模板引物	
[0116] T7-T2-F	TAATACGACTCACTATAGGGTAACGGCAGACTTCTCCAC (SEQ ID NO: 54)
T7-polyA-R	AAAAGCACCGACTCGGTGCC (SEQ ID NO: 55)

[0117] TaqMan探针:

[0118] HBB-wb-FAM-TM	FAM-TCCTGaGGAaAAaT-MGB (SEQ ID NO: 56)
[0119] HBB-wt-FAM-TM	FAM-TGACTCCTGAGGAGAA-MGB (SEQ ID NO: 57)
HBB-sk-VIC-TM	VIC-ACTCCTGTGGAGAAG-MGB (SEQ ID NO: 58)

[0120] ddPCR引物:

[0121]

R196	HBB-TaQM-f2	CAGAGCCATCTATTGCTTACATTTG (SEQ ID NO: 59)
R197	HBB-TaQM-r1	GGCTCACCACCAACTTCAT (SEQ ID NO: 60)

[0122] 如上所述,可将包含指导RNA (gRNA)、修饰的重组Cas9蛋白 (mrCas9) 和单链寡脱氧核糖核苷酸 (ssODN) 的复合物引入到人类干细胞或其衍生物中,以校正导致疾病的单碱基突变。表1和图8示出来自将镰状细胞校正复合物 (gRNA-mrCas9-ssODN) 引入到衍生自镰状细胞患者的皮肤细胞的诱导性多能干细胞 (iPSC) 中的结果。如实施例1所述那样衍生IPSC。通过核穿孔将校正复合物引入到镰状iPSC中,并且2天后使用上述引物通过数字PCR分析基因组DNA并对其进行测序。超过65%的细胞含有至少一种校正基因。一种校正基因足以治愈疾病。如下证实结果。在引入校正复合物两天后,将细胞铺在培养皿中并分离出43个单个iPSC集落。从这些集落中分离基因组DNA,并对 β -珠蛋白基因进行测序。百分之六十五的集落含有至少一种校正的 β -珠蛋白基因 (S校正为A)。

[0123] 表1

		gRNA-mrCas9-ssODN	
	合并的 ddPCR 结果(2 天)	68.6%	
	2 周后挑取的总集落	48	
	混合的集落	5	
	总单一集落	43	
	A/A	14	32.6%
	A/S	4	9.3%
[0124]	S/S	3	7.0%
	A/插入缺失	10	23.3%
	S/插入缺失	6	14.0%
	插入缺失/插入缺失	6	14.0%
	至少 1 个等位基因校正的克隆	28	65.1%
	具有插入缺失的克隆	22	51.2%
	具有基因组修饰的克隆	40	93.0%
	等位基因总数	86	
	总“A”等位基因(校正的)	42	48.8%
	总“S”等位基因(未校正的)	16	18.6%
	总“插入缺失”等位基因	28	32.6%
[0125]	A:(A+S)		
	*相当于 ddPCR 结果	42/58 = 72.4%	
	HR:NHEJ (A:插入缺失)比率	1.50	

[0126] 用患者原代骨髓CD34+细胞进行类似的研究。方案如下。按照IRB批准的方案从镰状患者获得骨髓。在Miltenyi抗CD34+珠(Miltenyi, Bergisch Gladbach, Germany)上纯化CD34+细胞。用如上所述制备的复合物对细胞进行核穿孔。在核穿孔后,在两周后挑取出铺在methycult和BFU-E、CFU-E和CFU-GEMM集落中的细胞,并分析校正的等位基因。表2和图9示出来自将镰状细胞校正复合物(gRNA-mrCas9-ssODN)引入到患者原代骨髓CD34+细胞中的结果。在体外分化十二天后,通过数字PCR(ddPCR)分析DNA并对其进行测序。观察到近似等量的BA和BS mRNA(参见图9)。紧接核穿孔之后,将一些细胞在红系分化培养基中培养多至十八天,并分析去核红细胞的HbA。来自校正镰状患者CD34+细胞的体外分化红细胞的等电聚焦(IEF)凝胶显示HbA(正常血红蛋白)与HbS(具有镰状细胞突变的血红蛋白)比率为约1:3,其足以抑制镰状化并治疗疾病(参见图10)。

[0127] 表2

核转染复合物 核转染程序		Cas9wt-36GFP-T2-ssODN	
		P4 DN-100	
[0128]	在 D10 和 D15 时挑取的 BFU-E/CFU-E/GEMM 集落	21/23/7	
	总集落*	51	
	A/A	2	4%
	A/S	4	8%
	S/S	19	37%
	A/插入缺失	5	10%
	S/插入缺失	15	29%
	插入缺失/插入缺失	6	12%
	至少 1 个等位基因校正的克隆	11	22%
	具有插入缺失的克隆	24	47%
	具有基因组修饰的克隆	29	57%
	等位基因总数	102	
	[0129]	总“A”等位基因(校正的)	13
总“S”等位基因(未校正的)		57	56%
总“插入缺失”等位基因		32	31%
A:(A+S) *相当于 ddPCR 结果		13/70=18.6%	
HR:NHEJ (A;插入缺失)比率		0.41	

[0130] 序列

[0131] SEQ ID NO:1

[0132] TAACGGCAGACTTCTCCAC

[0133] SEQ ID NO:2

[0134] GTAACGGCAGACTTCTCCACGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAA
CTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTTT

[0135] SEQ ID NO:3

[0136] Cas9超荷电GFP构建体

[0137]

mdykdhdgdykdhdidykdddmapkkkrkvgihgypaadkkysigldigtntsvgwavitdeykvpskkfklvlgnt
drhsikknligallfdsetaetrkrtrrrytrrknricylqeifsnemakvddsfhrleesflveedkkher
hpifgnivdevayhekyptiyhrkkldvstdkadrliylalahmikfrghfliedlnpdnsdvdklfiqlvqty
nqlfeenpinasgvdakailsarlsksrrlenliaqlpgeknglfgnlialslgltpnfksnfdlaedaklqlskd
tyddldnllaqidqyadflaaknlsdaildsilrvnteitkaplsasmikrydehhqdltlkalvrqqlpek
ykeiffdqskngyagyidggasqeefykfikpilekmdgteellvklredllrkqrtfdngsiphqihlgel hail
rrqedfypflkdnrekiekiltfripyyvgplargnsrfawmtrksetitpwnfeevdkgasafsiermtnfdk
nlpnekvlpkhslllyeyftvyneltkvkyvtegmrkpaflsgeqkkaivdlfktrkvtvkqlkedyfkkiecfds
veisgvedrfnaslgyhdllkiikdkdfl dneenediledivltltl fedremieerlktyahlfddkvmkqlkrr
rytgwgrlsrklingirdkqsgktildflksdganrnfmlihddsltfkediqaqvsgqgds lhehianlagsp
aikkgilqtvkvvdvlkvmgrhkeniviemarenqttqkgqksrermkrieegikelgsqilkehpventqlqn

eklylyylqngrdmyvdqeldinrlsdydvdhivpqsflkddsidnkvltrsdknrgksdnpseevvkkmknywrq
llnaklitqrkfdnlткаergglseldkagfikrqlvetrqi tkhvaqildsrmtkydendklirevkvitlkskl
vsdfrkdfqfykvreinnyhhahdaylnavvgtalikkyplesefvgydykvydvrmiaakseiqgatakayffy
snimffkteitlangeirkrplietngetgeivwdkgrdfatvrkvlsmqpvnivkktevqtggfskesilpkrns
dkliarkkdwdpkkyggfdsptvaysvlvakvekgkssklksvkellgitimerssfeknpidfle

[0138]

akgykevkkdliiklpkyslfelengrkrmlasagelqkgnelalpskyvnflylashyeklkgspedneqkqlfve
qhkhyldiieqisefskrviladanldkvl saynkhrdkpireqaenihlftltnlgapaafkyfdttidrkryt
stkevldatlihqsitglyetridlsqggdkrpaatkkagqakkkkgsgsngssgsaskgerlfrgkvpilvelkg
dvnghkfsvrgkgkdatrgkltlkficttgklpvpwptlvttltygvqcf srypkhmkrhdfkssampkgyvqert
isfkkgdkyktraevkfegrtlvnriklgdrdfkekgnilghklryfnshkvyitadkrkngikakfkirhvnkdg
svqladhyqqntpigrgpvllprnhylstrsklskdpkekrdhmvllfvtaagikhgrderyk

[0139] SEQ ID NO:4

[0140] TAT-Cas9超荷电GFP构建体

[0141]

ygrkrrrqrppqaggsmdykdhdgdykdhdidykdddmapkkkrkvghgvpaadkkysigldigtntsvgwav
itdcykvpskkfvlgntrhsikknligallfdsgetaeatrkrtrrrytrrknricylqeifsnemakvdds f
fhrleesflveedkkherhphifgnivdevayhekyptiyhlrkkldvstdkadlrliylalahmikfrghfliedl
npdndsvdklfiqlvqtnqlfeenpinasgvdakailsarlsksrrenliaqlpgekknglfgnlialslgltpn
fksnfdlaedaklqlskdtydddlnllaqidqyadflaaknlsdailldilrvnteitkaplsasmikrydeh
hqdltlkalkvrqqpkykeiffdqskngyagyidggasqeefykfikpilekmdgteallvklredllrkqrtf
dngsiphqihlgel hailrrqedfypflkdnrekiekiltfripyyvgplargnsrfawmtrkseetitpwnfeev
dkgasaqsfiermtnfdknlpnekvlpkshlyeyftvyneltkvkvtegmrkpaf lsgeqkkaivdllfktrkv
tvkqlkedyfkkiecfdsveisgvedrfnaslgtyhdllkiikdkdfl dneenediledivltl tlfedremieerl
ktyahlfddkvmkqlkrrrytgwgrlsrklingirdkqsgktildflksdganrnmqlihdds ltfkediqaqv
sgqgds l hehianlagspaikkgilqtvkvvdelvkmgrhkp eniviemarenqttqkgqknsrermkrieegike
lgsqilkehpventqlqneklylyylqngrdmyvdqeldinrlsdydvdhivpqsflkddsidnkvltrsdknrgks
dnpseevvkkmknywrqllnaklitqrkfdnlткаergglseldkagfikrqlvetrqi tkhvaqildsrmtkyd
endklirevkvitlksklvsdfrkdfqfykvreinnyhhahdaylnavvgtalikkyplesefvgydykvydvrm
iakseiqgatakayffysnimffkteitlangeirkrplietngetgeivwdkgrdfatvrkvlsmqpvnivkkt
evqtggfskesilpkrnsdkliarkkdwdpkkyggfdsptvaysvlvakvekgkssklksvkellgitimerssfe
knpldfleakgykevkkdliiklpkyslfelengrkrmlasagelqkgnelalpskyvnflylashyeklkgspedn
eqkqlfveqhkhyldiieqisefskrviladanldkvl saynkhrdkpireqaenihlftltnlgapaafkyfdt
tidrkrytstkevldatlihqsilglyetridlsqggdkrpaatkkagqakkkkgsgsngssgsaskgerlfrgkv
p ilvelkgdvnghkfsvrgkgkdatrgkltlkficttgklpvpwptlvttltygvqcf srypkhmkrhdfkssamp
kgyvqertisfkkgdkyktraevkfegrtlvnriklgdrdfkekgnilghklryfnshkvyitadkrkngikakfk
irhvnkdgs vqladhyqqntpigrgpvllprnhylstrsklskdpkekrdhmvllfvtaagikhgrderyk gsgg
svdglfeaiegfiengwegmidgwyg。

序列表

- <110> UAB研究基金会 (UAB Research Foundation)
Townes, Tim
Ding, Lei
Chang, Chia-Wei
- <120> 用于将功能性多肽引入到血细胞谱系细胞中的CRISPR/CAS9复合物
- <130> 035979-1014657 (178W01)
- <150> US 62/181,145
<151> 2015-06-17
- <160> 60
- <170> PatentIn 3.5版
- <210> 1
<211> 19
<212> DNA
<213> 人工序列
- <220>
<223> 合成构建体
- <400> 1
taacggcaga cttctccac 19
- [0001] <210> 2
<211> 103
<212> DNA
<213> 人工序列
- <220>
<223> 合成构建体
- <400> 2
gtaacggcag acttctccac gttttagagc tagaaatagc aagttaaaat aaggctagtc 60
cgttatcaac ttgaaaaagt ggcaccgagt cgggtctttt ttt 103
- <210> 3
<211> 1671
<212> PRT
<213> 人工序列
- <220>
<223> 合成构建体
- <400> 3
- Met Asp Tyr Lys Asp His Asp Gly Asp Tyr Lys Asp His Asp Ile Asp
1 5 10 15
- Tyr Lys Asp Asp Asp Asp Lys Met Ala Pro Lys Lys Lys Arg Lys Val
20 25 30
- Gly Ile His Gly Val Pro Ala Ala Asp Lys Lys Tyr Ser Ile Gly Leu
35 40 45

Asp Ile Gly Thr Asn Ser Val Gly Trp Ala Val Ile Thr Asp Glu Tyr
 50 55 60

Lys Val Pro Ser Lys Lys Phe Lys Val Leu Gly Asn Thr Asp Arg His
 65 70 75 80

Ser Ile Lys Lys Asn Leu Ile Gly Ala Leu Leu Phe Asp Ser Gly Glu
 85 90 95

Thr Ala Glu Ala Thr Arg Leu Lys Arg Thr Ala Arg Arg Arg Tyr Thr
 100 105 110

Arg Arg Lys Asn Arg Ile Cys Tyr Leu Gln Glu Ile Phe Ser Asn Glu
 115 120 125

Met Ala Lys Val Asp Asp Ser Phe Phe His Arg Leu Glu Glu Ser Phe
 130 135 140

Leu Val Glu Glu Asp Lys Lys His Glu Arg His Pro Ile Phe Gly Asn
 145 150 155 160

Ile Val Asp Glu Val Ala Tyr His Glu Lys Tyr Pro Thr Ile Tyr His
 165 170 175

[0002]

Leu Arg Lys Lys Leu Val Asp Ser Thr Asp Lys Ala Asp Leu Arg Leu
 180 185 190

Ile Tyr Leu Ala Leu Ala His Met Ile Lys Phe Arg Gly His Phe Leu
 195 200 205

Ile Glu Gly Asp Leu Asn Pro Asp Asn Ser Asp Val Asp Lys Leu Phe
 210 215 220

Ile Gln Leu Val Gln Thr Tyr Asn Gln Leu Phe Glu Glu Asn Pro Ile
 225 230 235 240

Asn Ala Ser Gly Val Asp Ala Lys Ala Ile Leu Ser Ala Arg Leu Ser
 245 250 255

Lys Ser Arg Arg Leu Glu Asn Leu Ile Ala Gln Leu Pro Gly Glu Lys
 260 265 270

Lys Asn Gly Leu Phe Gly Asn Leu Ile Ala Leu Ser Leu Gly Leu Thr
 275 280 285

Pro Asn Phe Lys Ser Asn Phe Asp Leu Ala Glu Asp Ala Lys Leu Gln
 290 295 300

Leu Ser Lys Asp Thr Tyr Asp Asp Asp Leu Asp Asn Leu Leu Ala Gln
 305 310 315 320

Ile Gly Asp Gln Tyr Ala Asp Leu Phe Leu Ala Ala Lys Asn Leu Ser
 325 330 335

Asp Ala Ile Leu Leu Ser Asp Ile Leu Arg Val Asn Thr Glu Ile Thr
 340 345 350

Lys Ala Pro Leu Ser Ala Ser Met Ile Lys Arg Tyr Asp Glu His His
 355 360 365

Gln Asp Leu Thr Leu Leu Lys Ala Leu Val Arg Gln Gln Leu Pro Glu
 370 375 380

Lys Tyr Lys Glu Ile Phe Phe Asp Gln Ser Lys Asn Gly Tyr Ala Gly
 385 390 395 400

Tyr Ile Asp Gly Gly Ala Ser Gln Glu Glu Phe Tyr Lys Phe Ile Lys
 405 410 415

Pro Ile Leu Glu Lys Met Asp Gly Thr Glu Glu Leu Leu Val Lys Leu
 420 425 430

[0003]

Asn Arg Glu Asp Leu Leu Arg Lys Gln Arg Thr Phe Asp Asn Gly Ser
 435 440 445

Ile Pro His Gln Ile His Leu Gly Glu Leu His Ala Ile Leu Arg Arg
 450 455 460

Gln Glu Asp Phe Tyr Pro Phe Leu Lys Asp Asn Arg Glu Lys Ile Glu
 465 470 475 480

Lys Ile Leu Thr Phe Arg Ile Pro Tyr Tyr Val Gly Pro Leu Ala Arg
 485 490 495

Gly Asn Ser Arg Phe Ala Trp Met Thr Arg Lys Ser Glu Glu Thr Ile
 500 505 510

Thr Pro Trp Asn Phe Glu Glu Val Val Asp Lys Gly Ala Ser Ala Gln
 515 520 525

Ser Phe Ile Glu Arg Met Thr Asn Phe Asp Lys Asn Leu Pro Asn Glu
 530 535 540

Lys Val Leu Pro Lys His Ser Leu Leu Tyr Glu Tyr Phe Thr Val Tyr
 545 550 555 560

Asn Glu Leu Thr Lys Val Lys Tyr Val Thr Glu Gly Met Arg Lys Pro
 565 570 575

Ala Phe Leu Ser Gly Glu Gln Lys Lys Ala Ile Val Asp Leu Phe
 580 585 590

Lys Thr Asn Arg Lys Val Thr Val Lys Gln Leu Lys Glu Asp Tyr Phe
 595 600 605

Lys Lys Ile Glu Cys Phe Asp Ser Val Glu Ile Ser Gly Val Glu Asp
 610 615 620

Arg Phe Asn Ala Ser Leu Gly Thr Tyr His Asp Leu Leu Lys Ile Ile
 625 630 635 640

Lys Asp Lys Asp Phe Leu Asp Asn Glu Glu Asn Glu Asp Ile Leu Glu
 645 650 655

Asp Ile Val Leu Thr Leu Thr Leu Phe Glu Asp Arg Glu Met Ile Glu
 660 665 670

Glu Arg Leu Lys Thr Tyr Ala His Leu Phe Asp Asp Lys Val Met Lys
 675 680 685

[0004] Gln Leu Lys Arg Arg Arg Tyr Thr Gly Trp Gly Arg Leu Ser Arg Lys
 690 695 700

Leu Ile Asn Gly Ile Arg Asp Lys Gln Ser Gly Lys Thr Ile Leu Asp
 705 710 715

Phe Leu Lys Ser Asp Gly Phe Ala Asn Arg Asn Phe Met Gln Leu Ile
 725 730 735

His Asp Asp Ser Leu Thr Phe Lys Glu Asp Ile Gln Lys Ala Gln Val
 740 745 750

Ser Gly Gln Gly Asp Ser Leu His Glu His Ile Ala Asn Leu Ala Gly
 755 760 765

Ser Pro Ala Ile Lys Lys Gly Ile Leu Gln Thr Val Lys Val Val Asp
 770 775 780

Glu Leu Val Lys Val Met Gly Arg His Lys Pro Glu Asn Ile Val Ile
 785 790 795 800

Glu Met Ala Arg Glu Asn Gln Thr Thr Gln Lys Gly Gln Lys Asn Ser
 805 810 815

Arg Glu Arg Met Lys Arg Ile Glu Glu Gly Ile Lys Glu Leu Gly Ser

	820	825	830
	Gln Ile Leu Lys Glu His Pro Val Glu Asn Thr Gln Leu Gln Asn Glu 835 840 845		
	Lys Leu Tyr Leu Tyr Tyr Leu Gln Asn Gly Arg Asp Met Tyr Val Asp 850 855 860		
	Gln Glu Leu Asp Ile Asn Arg Leu Ser Asp Tyr Asp Val Asp His Ile 865 870 875 880		
	Val Pro Gln Ser Phe Leu Lys Asp Asp Ser Ile Asp Asn Lys Val Leu 885 890 895		
	Thr Arg Ser Asp Lys Asn Arg Gly Lys Ser Asp Asn Val Pro Ser Glu 900 905 910		
	Glu Val Val Lys Lys Met Lys Asn Tyr Trp Arg Gln Leu Leu Asn Ala 915 920 925		
	Lys Leu Ile Thr Gln Arg Lys Phe Asp Asn Leu Thr Lys Ala Glu Arg 930 935 940		
[0005]	Gly Gly Leu Ser Glu Leu Asp Lys Ala Gly Phe Ile Lys Arg Gln Leu 945 950 955 960		
	Val Glu Thr Arg Gln Ile Thr Lys His Val Ala Gln Ile Leu Asp Ser 965 970 975		
	Arg Met Asn Thr Lys Tyr Asp Glu Asn Asp Lys Leu Ile Arg Glu Val 980 985 990		
	Lys Val Ile Thr Leu Lys Ser Lys Leu Val Ser Asp Phe Arg Lys Asp 995 1000 1005		
	Phe Gln Phe Tyr Lys Val Arg Glu Ile Asn Asn Tyr His His Ala 1010 1015 1020		
	His Asp Ala Tyr Leu Asn Ala Val Val Gly Thr Ala Leu Ile Lys 1025 1030 1035		
	Lys Tyr Pro Lys Leu Glu Ser Glu Phe Val Tyr Gly Asp Tyr Lys 1040 1045 1050		
	Val Tyr Asp Val Arg Lys Met Ile Ala Lys Ser Glu Gln Glu Ile 1055 1060 1065		
	Gly Lys Ala Thr Ala Lys Tyr Phe Phe Tyr Ser Asn Ile Met Asn 1070 1075 1080		

Phe Phe Lys Thr Glu Ile Thr Leu Ala Asn Gly Glu Ile Arg Lys
 1085 1090 1095

Arg Pro Leu Ile Glu Thr Asn Gly Glu Thr Gly Glu Ile Val Trp
 1100 1105 1110

Asp Lys Gly Arg Asp Phe Ala Thr Val Arg Lys Val Leu Ser Met
 1115 1120 1125

Pro Gln Val Asn Ile Val Lys Lys Thr Glu Val Gln Thr Gly Gly
 1130 1135 1140

Phe Ser Lys Glu Ser Ile Leu Pro Lys Arg Asn Ser Asp Lys Leu
 1145 1150 1155

Ile Ala Arg Lys Lys Asp Trp Asp Pro Lys Lys Tyr Gly Gly Phe
 1160 1165 1170

Asp Ser Pro Thr Val Ala Tyr Ser Val Leu Val Val Ala Lys Val
 1175 1180 1185

Glu Lys Gly Lys Ser Lys Lys Leu Lys Ser Val Lys Glu Leu Leu
 1190 1195 1200

[0006]

Gly Ile Thr Ile Met Glu Arg Ser Ser Phe Glu Lys Asn Pro Ile
 1205 1210 1215

Asp Phe Leu Glu Ala Lys Gly Tyr Lys Glu Val Lys Lys Asp Leu
 1220 1225 1230

Ile Ile Lys Leu Pro Lys Tyr Ser Leu Phe Glu Leu Glu Asn Gly
 1235 1240 1245

Arg Lys Arg Met Leu Ala Ser Ala Gly Glu Leu Gln Lys Gly Asn
 1250 1255 1260

Glu Leu Ala Leu Pro Ser Lys Tyr Val Asn Phe Leu Tyr Leu Ala
 1265 1270 1275

Ser His Tyr Glu Lys Leu Lys Gly Ser Pro Glu Asp Asn Glu Gln
 1280 1285 1290

Lys Gln Leu Phe Val Glu Gln His Lys His Tyr Leu Asp Glu Ile
 1295 1300 1305

Ile Glu Gln Ile Ser Glu Phe Ser Lys Arg Val Ile Leu Ala Asp
 1310 1315 1320

Ala Asn Leu Asp Lys Val Leu Ser Ala Tyr Asn Lys His Arg Asp
1325 1330 1335

Lys Pro Ile Arg Glu Gln Ala Glu Asn Ile Ile His Leu Phe Thr
1340 1345 1350

Leu Thr Asn Leu Gly Ala Pro Ala Ala Phe Lys Tyr Phe Asp Thr
1355 1360 1365

Thr Ile Asp Arg Lys Arg Tyr Thr Ser Thr Lys Glu Val Leu Asp
1370 1375 1380

Ala Thr Leu Ile His Gln Ser Ile Thr Gly Leu Tyr Glu Thr Arg
1385 1390 1395

Ile Asp Leu Ser Gln Leu Gly Gly Asp Lys Arg Pro Ala Ala Thr
1400 1405 1410

Lys Lys Ala Gly Gln Ala Lys Lys Lys Lys Gly Ser Gly Ser Asn
1415 1420 1425

Gly Ser Ser Gly Ser Ala Ser Lys Gly Glu Arg Leu Phe Arg Gly
1430 1435 1440

[0007]

Lys Val Pro Ile Leu Val Glu Leu Lys Gly Asp Val Asn Gly His
1445 1450 1455

Lys Phe Ser Val Arg Gly Lys Gly Lys Gly Asp Ala Thr Arg Gly
1460 1465 1470

Lys Leu Thr Leu Lys Phe Ile Cys Thr Thr Gly Lys Leu Pro Val
1475 1480 1485

Pro Trp Pro Thr Leu Val Thr Thr Leu Thr Tyr Gly Val Gln Cys
1490 1495 1500

Phe Ser Arg Tyr Pro Lys His Met Lys Arg His Asp Phe Phe Lys
1505 1510 1515

Ser Ala Met Pro Lys Gly Tyr Val Gln Glu Arg Thr Ile Ser Phe
1520 1525 1530

Lys Lys Asp Gly Lys Tyr Lys Thr Arg Ala Glu Val Lys Phe Glu
1535 1540 1545

Gly Arg Thr Leu Val Asn Arg Ile Lys Leu Lys Gly Arg Asp Phe
1550 1555 1560

Lys Glu Lys Gly Asn Ile Leu Gly His Lys Leu Arg Tyr Asn Phe
1565 1570 1575

Asn Ser His Lys Val Tyr Ile Thr Ala Asp Lys Arg Lys Asn Gly
1580 1585 1590

Ile Lys Ala Lys Phe Lys Ile Arg His Asn Val Lys Asp Gly Ser
1595 1600 1605

Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Arg
1610 1615 1620

Gly Pro Val Leu Leu Pro Arg Asn His Tyr Leu Ser Thr Arg Ser
1625 1630 1635

Lys Leu Ser Lys Asp Pro Lys Glu Lys Arg Asp His Met Val Leu
1640 1645 1650

Leu Glu Phe Val Thr Ala Ala Gly Ile Lys His Gly Arg Asp Glu
1655 1660 1665

Arg Tyr Lys
1670

[0008]

<210> 4
<211> 1720
<212> PRT
<213> 人工序列

<220>
<223> 合成构建体

<400> 4

Tyr Gly Arg Lys Lys Arg Arg Gln Arg Arg Arg Pro Pro Gln Ala Gly
1 5 10 15

Gly Ser Met Asp Tyr Lys Asp His Asp Gly Asp Tyr Lys Asp His Asp
20 25 30

Ile Asp Tyr Lys Asp Asp Asp Asp Lys Met Ala Pro Lys Lys Lys Arg
35 40 45

Lys Val Gly Ile His Gly Val Pro Ala Ala Asp Lys Lys Tyr Ser Ile
50 55 60

Gly Leu Asp Ile Gly Thr Asn Ser Val Gly Trp Ala Val Ile Thr Asp
65 70 75 80

Glu Tyr Lys Val Pro Ser Lys Lys Phe Lys Val Leu Gly Asn Thr Asp
85 90 95

Arg His Ser Ile Lys Lys Asn Leu Ile Gly Ala Leu Leu Phe Asp Ser
100 105 110

Gly Glu Thr Ala Glu Ala Thr Arg Leu Lys Arg Thr Ala Arg Arg Arg
115 120 125

Tyr Thr Arg Arg Lys Asn Arg Ile Cys Tyr Leu Gln Glu Ile Phe Ser
130 135 140

Asn Glu Met Ala Lys Val Asp Asp Ser Phe Phe His Arg Leu Glu Glu
145 150 155 160

Ser Phe Leu Val Glu Glu Asp Lys Lys His Glu Arg His Pro Ile Phe
165 170 175

Gly Asn Ile Val Asp Glu Val Ala Tyr His Glu Lys Tyr Pro Thr Ile
180 185 190

Tyr His Leu Arg Lys Lys Leu Val Asp Ser Thr Asp Lys Ala Asp Leu
195 200 205

Arg Leu Ile Tyr Leu Ala Leu Ala His Met Ile Lys Phe Arg Gly His
210 215 220

[0009]

Phe Leu Ile Glu Gly Asp Leu Asn Pro Asp Asn Ser Asp Val Asp Lys
225 230 235 240

Leu Phe Ile Gln Leu Val Gln Thr Tyr Asn Gln Leu Phe Glu Glu Asn
245 250 255

Pro Ile Asn Ala Ser Gly Val Asp Ala Lys Ala Ile Leu Ser Ala Arg
260 265 270

Leu Ser Lys Ser Arg Arg Leu Glu Asn Leu Ile Ala Gln Leu Pro Gly
275 280 285

Glu Lys Lys Asn Gly Leu Phe Gly Asn Leu Ile Ala Leu Ser Leu Gly
290 295 300

Leu Thr Pro Asn Phe Lys Ser Asn Phe Asp Leu Ala Glu Asp Ala Lys
305 310 315 320

Leu Gln Leu Ser Lys Asp Thr Tyr Asp Asp Asp Leu Asp Asn Leu Leu
325 330 335

Ala Gln Ile Gly Asp Gln Tyr Ala Asp Leu Phe Leu Ala Ala Lys Asn
340 345 350

Leu Ser Asp Ala Ile Leu Leu Ser Asp Ile Leu Arg Val Asn Thr Glu
 355 360 365

Ile Thr Lys Ala Pro Leu Ser Ala Ser Met Ile Lys Arg Tyr Asp Glu
 370 375 380

His His Gln Asp Leu Thr Leu Leu Lys Ala Leu Val Arg Gln Gln Leu
 385 390 395 400

Pro Glu Lys Tyr Lys Glu Ile Phe Phe Asp Gln Ser Lys Asn Gly Tyr
 405 410 415

Ala Gly Tyr Ile Asp Gly Gly Ala Ser Gln Glu Glu Phe Tyr Lys Phe
 420 425 430

Ile Lys Pro Ile Leu Glu Lys Met Asp Gly Thr Glu Glu Leu Leu Val
 435 440 445

Lys Leu Asn Arg Glu Asp Leu Leu Arg Lys Gln Arg Thr Phe Asp Asn
 450 455 460

Gly Ser Ile Pro His Gln Ile His Leu Gly Glu Leu His Ala Ile Leu
 465 470 475 480

[0010] Arg Arg Gln Glu Asp Phe Tyr Pro Phe Leu Lys Asp Asn Arg Glu Lys
 485 490 495

Ile Glu Lys Ile Leu Thr Phe Arg Ile Pro Tyr Tyr Val Gly Pro Leu
 500 505 510

Ala Arg Gly Asn Ser Arg Phe Ala Trp Met Thr Arg Lys Ser Glu Glu
 515 520 525

Thr Ile Thr Pro Trp Asn Phe Glu Glu Val Val Asp Lys Gly Ala Ser
 530 535 540

Ala Gln Ser Phe Ile Glu Arg Met Thr Asn Phe Asp Lys Asn Leu Pro
 545 550 555 560

Asn Glu Lys Val Leu Pro Lys His Ser Leu Leu Tyr Glu Tyr Phe Thr
 565 570 575

Val Tyr Asn Glu Leu Thr Lys Val Lys Tyr Val Thr Glu Gly Met Arg
 580 585 590

Lys Pro Ala Phe Leu Ser Gly Glu Gln Lys Lys Ala Ile Val Asp Leu
 595 600 605

Leu Phe Lys Thr Asn Arg Lys Val Thr Val Lys Gln Leu Lys Glu Asp

	610	615	620																
	Tyr 625	Phe	Lys	Lys	Ile	Glu 630	Cys	Phe	Asp	Ser	Val 635	Glu	Ile	Ser	Gly	Val 640			
	Glu	Asp	Arg	Phe	Asn 645	Ala	Ser	Leu	Gly	Thr 650	Tyr	His	Asp	Leu	Leu	Lys 655			
	Ile	Ile	Lys	Asp 660	Lys	Asp	Phe	Leu	Asp 665	Asn	Glu	Glu	Asn	Glu	Asp	Ile 670			
	Leu	Glu	Asp 675	Ile	Val	Leu	Thr	Leu	Thr 680	Leu	Phe	Glu	Asp 685	Arg	Glu	Met			
	Ile	Glu	Glu	Arg	Leu	Lys	Thr 695	Tyr	Ala	His	Leu	Phe	Asp 700	Asp	Lys	Val			
	Met 705	Lys	Gln	Leu	Lys	Arg 710	Arg	Arg	Tyr	Thr	Gly 715	Trp	Gly	Arg	Leu	Ser 720			
	Arg	Lys	Leu	Ile	Asn 725	Gly	Ile	Arg	Asp	Lys 730	Gln	Ser	Gly	Lys	Thr 735	Ile			
[0011]	Leu	Asp	Phe	Leu 740	Lys	Ser	Asp	Gly	Phe 745	Ala	Asn	Arg	Asn	Phe	Met	Gln 750			
	Leu	Ile	His 755	Asp	Asp	Ser	Leu	Thr 760	Phe	Lys	Glu	Asp	Ile 765	Gln	Lys	Ala			
	Gln 770	Val	Ser	Gly	Gln	Gly	Asp 775	Ser	Leu	His	Glu	His 780	Ile	Ala	Asn	Leu			
	Ala 785	Gly	Ser	Pro	Ala	Ile 790	Lys	Lys	Gly	Ile	Leu 795	Gln	Thr	Val	Lys	Val 800			
	Val	Asp	Glu	Leu 805	Val	Lys	Val	Met	Gly	Arg 810	His	Lys	Pro	Glu	Asn 815	Ile			
	Val	Ile	Glu	Met 820	Ala	Arg	Glu	Asn	Gln 825	Thr	Thr	Gln	Lys	Gly 830	Gln	Lys			
	Asn	Ser	Arg 835	Glu	Arg	Met	Lys	Arg 840	Ile	Glu	Glu	Gly	Ile 845	Lys	Glu	Leu			
	Gly 850	Ser	Gln	Ile	Leu	Lys	Glu 855	His	Pro	Val	Glu	Asn 860	Thr	Gln	Leu	Gln			
	Asn 865	Glu	Lys	Leu	Tyr	Leu	Tyr 870	Tyr	Leu	Gln	Asn 875	Gly	Arg	Asp	Met	Tyr 880			

Val Asp Gln Glu Leu Asp Ile Asn Arg Leu Ser Asp Tyr Asp Val Asp
 885 890 895

His Ile Val Pro Gln Ser Phe Leu Lys Asp Asp Ser Ile Asp Asn Lys
 900 905 910

Val Leu Thr Arg Ser Asp Lys Asn Arg Gly Lys Ser Asp Asn Val Pro
 915 920 925

Ser Glu Glu Val Val Lys Lys Met Lys Asn Tyr Trp Arg Gln Leu Leu
 930 935 940

Asn Ala Lys Leu Ile Thr Gln Arg Lys Phe Asp Asn Leu Thr Lys Ala
 945 950 955 960

Glu Arg Gly Gly Leu Ser Glu Leu Asp Lys Ala Gly Phe Ile Lys Arg
 965 970 975

Gln Leu Val Glu Thr Arg Gln Ile Thr Lys His Val Ala Gln Ile Leu
 980 985 990

Asp Ser Arg Met Asn Thr Lys Tyr Asp Glu Asn Asp Lys Leu Ile Arg
 995 1000 1005

[0012]

Glu Val Lys Val Ile Thr Leu Lys Ser Lys Leu Val Ser Asp Phe
 1010 1015 1020

Arg Lys Asp Phe Gln Phe Tyr Lys Val Arg Glu Ile Asn Asn Tyr
 1025 1030 1035

His His Ala His Asp Ala Tyr Leu Asn Ala Val Val Gly Thr Ala
 1040 1045 1050

Leu Ile Lys Lys Tyr Pro Lys Leu Glu Ser Glu Phe Val Tyr Gly
 1055 1060 1065

Asp Tyr Lys Val Tyr Asp Val Arg Lys Met Ile Ala Lys Ser Glu
 1070 1075 1080

Gln Glu Ile Gly Lys Ala Thr Ala Lys Tyr Phe Phe Tyr Ser Asn
 1085 1090 1095

Ile Met Asn Phe Phe Lys Thr Glu Ile Thr Leu Ala Asn Gly Glu
 1100 1105 1110

Ile Arg Lys Arg Pro Leu Ile Glu Thr Asn Gly Glu Thr Gly Glu
 1115 1120 1125

Ile Val Trp Asp Lys Gly Arg Asp Phe Ala Thr Val Arg Lys Val
1130 1135 1140

Leu Ser Met Pro Gln Val Asn Ile Val Lys Lys Thr Glu Val Gln
1145 1150 1155

Thr Gly Gly Phe Ser Lys Glu Ser Ile Leu Pro Lys Arg Asn Ser
1160 1165 1170

Asp Lys Leu Ile Ala Arg Lys Lys Asp Trp Asp Pro Lys Lys Tyr
1175 1180 1185

Gly Gly Phe Asp Ser Pro Thr Val Ala Tyr Ser Val Leu Val Val
1190 1195 1200

Ala Lys Val Glu Lys Gly Lys Ser Lys Lys Leu Lys Ser Val Lys
1205 1210 1215

Glu Leu Leu Gly Ile Thr Ile Met Glu Arg Ser Ser Phe Glu Lys
1220 1225 1230

Asn Pro Ile Asp Phe Leu Glu Ala Lys Gly Tyr Lys Glu Val Lys
1235 1240 1245

[0013]

Lys Asp Leu Ile Ile Lys Leu Pro Lys Tyr Ser Leu Phe Glu Leu
1250 1255 1260

Glu Asn Gly Arg Lys Arg Met Leu Ala Ser Ala Gly Glu Leu Gln
1265 1270 1275

Lys Gly Asn Glu Leu Ala Leu Pro Ser Lys Tyr Val Asn Phe Leu
1280 1285 1290

Tyr Leu Ala Ser His Tyr Glu Lys Leu Lys Gly Ser Pro Glu Asp
1295 1300 1305

Asn Glu Gln Lys Gln Leu Phe Val Glu Gln His Lys His Tyr Leu
1310 1315 1320

Asp Glu Ile Ile Glu Gln Ile Ser Glu Phe Ser Lys Arg Val Ile
1325 1330 1335

Leu Ala Asp Ala Asn Leu Asp Lys Val Leu Ser Ala Tyr Asn Lys
1340 1345 1350

His Arg Asp Lys Pro Ile Arg Glu Gln Ala Glu Asn Ile Ile His
1355 1360 1365

	Leu Phe Thr Leu Thr Asn Leu Gly Ala Pro Ala Ala Phe Lys Tyr 1370 1375 1380
	Phe Asp Thr Thr Ile Asp Arg Lys Arg Tyr Thr Ser Thr Lys Glu 1385 1390 1395
	Val Leu Asp Ala Thr Leu Ile His Gln Ser Ile Thr Gly Leu Tyr 1400 1405 1410
	Glu Thr Arg Ile Asp Leu Ser Gln Leu Gly Gly Asp Lys Arg Pro 1415 1420 1425
	Ala Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys Lys Lys Gly Ser 1430 1435 1440
	Gly Ser Asn Gly Ser Ser Gly Ser Ala Ser Lys Gly Glu Arg Leu 1445 1450 1455
	Phe Arg Gly Lys Val Pro Ile Leu Val Glu Leu Lys Gly Asp Val 1460 1465 1470
	Asn Gly His Lys Phe Ser Val Arg Gly Lys Gly Lys Gly Asp Ala 1475 1480 1485
[0014]	Thr Arg Gly Lys Leu Thr Leu Lys Phe Ile Cys Thr Thr Gly Lys 1490 1495 1500
	Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr Leu Thr Tyr Gly 1505 1510 1515
	Val Gln Cys Phe Ser Arg Tyr Pro Lys His Met Lys Arg His Asp 1520 1525 1530
	Phe Phe Lys Ser Ala Met Pro Lys Gly Tyr Val Gln Glu Arg Thr 1535 1540 1545
	Ile Ser Phe Lys Lys Asp Gly Lys Tyr Lys Thr Arg Ala Glu Val 1550 1555 1560
	Lys Phe Glu Gly Arg Thr Leu Val Asn Arg Ile Lys Leu Lys Gly 1565 1570 1575
	Arg Asp Phe Lys Glu Lys Gly Asn Ile Leu Gly His Lys Leu Arg 1580 1585 1590
	Tyr Asn Phe Asn Ser His Lys Val Tyr Ile Thr Ala Asp Lys Arg 1595 1600 1605
	Lys Asn Gly Ile Lys Ala Lys Phe Lys Ile Arg His Asn Val Lys

1610 1615 1620
 Asp Gly Ser Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro
 1625 1630 1635
 Ile Gly Arg Gly Pro Val Leu Leu Pro Arg Asn His Tyr Leu Ser
 1640 1645 1650
 Thr Arg Ser Lys Leu Ser Lys Asp Pro Lys Glu Lys Arg Asp His
 1655 1660 1665
 Met Val Leu Leu Glu Phe Val Thr Ala Ala Gly Ile Lys His Gly
 1670 1675 1680
 Arg Asp Glu Arg Tyr Lys Gly Gly Ser Gly Gly Ser Val Asp Gly
 1685 1690 1695
 Leu Phe Glu Ala Ile Glu Gly Phe Ile Glu Asn Gly Trp Glu Gly
 1700 1705 1710
 Met Ile Asp Gly Trp Tyr Gly
 1715 1720

[0015]

<210> 5
 <211> 26
 <212> DNA
 <213> 人工序列
 <220>
 <223> 合成构建体
 <400> 5
 gctaattcac tcccaaagaa gacaag 26
 <210> 6
 <211> 20
 <212> DNA
 <213> 合成构建体
 <400> 6
 cttcagcaag ccgagtcctg 20
 <210> 7
 <211> 60
 <212> DNA
 <213> 人工序列
 <220>
 <223> 合成构建体
 <400> 7
 gtcgacgtcg acgctcagtg aagctgaagt attccttctg cttcacaggg cgaccactac 60
 <210> 8

	<211> 49	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成构建体	
	<400> 8	
	atttaaattcc tcccctcgaa cccttaccaa actcctatgc atactacag	49
	<210> 9	
	<211> 60	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成构建体	
	<400> 9	
	ttaattaatt aattagcatt ttaggttcag gttgtgagaa cactagaaga gaacaagtca	60
	<210> 10	
	<211> 52	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成构建体	
[0016]	<400> 10	
	gtatacgtat acgcatacct ggagagggga caaggtcttg agatgcgagg gt	52
	<210> 11	
	<211> 52	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成构建体	
	<400> 11	
	agccacctta attaagccac catggcgcac gctgggagaa cggggtacga ta	52
	<210> 12	
	<211> 60	
	<212> DNA	
	<213> 合成构建体	
	<400> 12	
	taacagagag aagttcgtgg ctccgatcc cttgtggccc agataggcac ccagggtgat	60
	<210> 13	
	<211> 24	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成构建体	
	<400> 13	

	caccgtgaga tacagataca gaca	24
	<210> 14	
	<211> 24	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成构建体	
	<400> 14	
	aaactgtctg tatctgtatc tcac	24
	<210> 15	
	<211> 25	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成构建体	
	<400> 15	
	aaacggcatt ccaggcaaat cattc	25
	<210> 16	
	<211> 25	
	<212> DNA	
	<213> 人工序列	
[0017]	<220>	
	<223> 合成构建体	
	<400> 16	
	caccgcagcc taggcaaagg cctgc	25
	<210> 17	
	<211> 25	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成构建体	
	<400> 17	
	aaacgcaggc cttgcctag gctgc	25
	<210> 18	
	<211> 25	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成构建体	
	<400> 18	
	caccgtgcca acagaactgc ctgat	25
	<210> 19	
	<211> 25	

	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成构建体	
	<400> 19	
	aaacatcagg cagttctggt ggcac	25
	<210> 20	
	<211> 24	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成构建体	
	<400> 20	
	caccgaccag ggtgcaagtg tgga	24
	<210> 21	
	<211> 24	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成构建体	
	<400> 21	
[0018]	aaactccaca ctgaccct ggtc	24
	<210> 22	
	<211> 24	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成构建体	
	<400> 22	
	caccgctcct cagcctggca ttca	24
	<210> 23	
	<211> 24	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成构建体	
	<400> 23	
	aaactgaatg ccaggctgag gage	24
	<210> 24	
	<211> 23	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成构建体	

	<400> 24 tgctaaagcg catgctccag act	23
	<210> 25 <211> 21 <212> DNA <213> 人工序列	
	<220> <223> 合成构建体	
	<400> 25 gtcttcatct cagggtcggc t	21
	<210> 26 <211> 21 <212> DNA <213> 人工序列	
	<220> <223> 合成构建体	
	<400> 26 cctctctgtg cattatggca g	21
[0019]	<210> 27 <211> 19 <212> DNA <213> 人工序列	
	<220> <223> 合成构建体	
	<400> 27 gccttctatc gccttcttg	19
	<210> 28 <211> 20 <212> DNA <213> 人工序列	
	<220> <223> 合成构建体	
	<400> 28 actcctccac ctttgacgct	20
	<210> 29 <211> 23 <212> DNA <213> 人工序列	
	<220> <223> 合成构建体	
	<400> 29 tcccctcttc aagggtctac atg	23

	<210> 30	
	<211> 20	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成构建体	
	<400> 30	
	gtgcaaaatg gaagggttc	20
	<210> 31	
	<211> 20	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成构建体	
	<400> 31	
	ggagctccgt gaagttgttc	20
	<210> 32	
	<211> 21	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成构建体	
[0020]	<400> 32	
	tgtttccttt cactggccac a	21
	<210> 33	
	<211> 21	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成构建体	
	<400> 33	
	aacggcaact ggtgaacggt a	21
	<210> 34	
	<211> 22	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成构建体	
	<400> 34	
	ggcgatgcca gaatagatgc cg	22
	<210> 35	
	<211> 29	
	<212> DNA	
	<213> 人工序列	

	<220>		
	<223>	合成构建体	
	<400>	35	
		ccaggccact tggctcctct atctccaga	29
	<210>	36	
	<211>	25	
	<212>	DNA	
	<213>	人工序列	
	<220>		
	<223>	合成构建体	
	<400>	36	
		ccttactggt gagactgcaa tatcc	25
	<210>	37	
	<211>	26	
	<212>	DNA	
	<213>	人工序列	
	<220>		
	<223>	合成构建体	
	<400>	37	
		ctgaagtccc agtatatact tcacac	26
[0021]	<210>	38	
	<211>	25	
	<212>	DNA	
	<213>	人工序列	
	<220>		
	<223>	合成构建体	
	<400>	38	
		cccagaagca gtaataatca tcgag	25
	<210>	39	
	<211>	23	
	<212>	DNA	
	<213>	人工序列	
	<220>		
	<223>	合成构建体	
	<400>	39	
		atgtgggatg tagtagatct tgc	23
	<210>	40	
	<211>	21	
	<212>	DNA	
	<213>	人工序列	
	<220>		
	<223>	合成构建体	
	<400>	40	
		gggtcttacc tcagcagtta c	21

	<210> 41	
	<211> 20	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成构建体	
	<400> 41	
	cctcacacag tgtgacgcag	20
	<210> 42	
	<211> 20	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成构建体	
	<400> 42	
	gactgagtac ctgaaccggc	20
	<210> 43	
	<211> 20	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成构建体	
[0022]	<400> 43	
	gggccaaact gagcagagtc	20
	<210> 44	
	<211> 20	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成构建体	
	<400> 44	
	aagaccaggg tggttgggac	20
	<210> 45	
	<211> 20	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成构建体	
	<400> 45	
	gtaagaaaaa tgcccacgtc	20
	<210> 46	
	<211> 20	
	<212> DNA	

	<213> Artificial sequence	
	<400> 46	
	agtcagacgt ctggagcttc	20
	<210> 47	
	<211> 22	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成构建体	
	<400> 47	
	gtgagcagtg aaggcatgag tc	22
	<210> 48	
	<211> 20	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成构建体	
	<400> 48	
	gtgagataca gatacagaca	20
[0023]	<210> 49	
	<211> 20	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成构建体	
	<400> 49	
	aatgatttgc ctggaatgcc	20
	<210> 50	
	<211> 23	
	<212> PRT	
	<213> 人工序列	
	<220>	
	<223> 合成构建体	
	<400> 50	
	Gly Leu Phe Glu Ala Ile Glu Gly Phe Ile Glu Asn Gly Trp Glu Gly	
	1 5 10 15	
	Met Ile Asp Gly Trp Tyr Gly	
	20	
	<210> 51	
	<211> 14	
	<212> PRT	
	<213> 人工序列	

<220>	
<223> 合成构建体	
<400> 51	
Tyr Gly Arg Lys Lys Arg Arg Gln Arg Arg Arg Pro Pro Gln	
1 5 10	
<210> 52	
<211> 91	
<212> DNA	
<213> 人工序列	
<220>	
<223> 合成构建体	
<400> 52	
atccacgttc accttgcccc acagggcagt aacggcagac ttctcctcag gagtcaggtg	60
caccatggtg tctgtttgag gttgctagtg a	91
<210> 53	
<211> 95	
<212> DNA	
<213> 人工序列	
<220>	
<223> 合成构建体	
[0024] <400> 53	
cttcatccac gttcaccttg ccccacaggg cagtaacggc agatttttcc tcaggagtca	60
ggtgcaccat ggtgtctggt tgaggttgct agtga	95
<210> 54	
<211> 39	
<212> DNA	
<213> 人工序列	
<220>	
<223> 合成构建体	
<400> 54	
taatacgact cactataggg taacggcaga cttctccac	39
<210> 55	
<211> 21	
<212> DNA	
<213> 人工序列	
<220>	
<223> 合成构建体	
<400> 55	
aaaaagcacc gactcgggtg c	21
<210> 56	
<211> 14	
<212> DNA	
<213> 人工序列	

	<220>		
	<223> 合成构建体		
	<400> 56		
	tcctgaggaa aaat		14
	<210> 57		
	<211> 16		
	<212> DNA		
	<213> 人工序列		
	<220>		
	<223> 合成构建体		
	<400> 57		
	tgactcctga ggagaa		16
	<210> 58		
	<211> 15		
	<212> DNA		
	<213> 人工序列		
[0025]	<220>		
	<223> 合成构建体		
	<400> 58		
	actcctgtgg agaag		15
	<210> 59		
	<211> 25		
	<212> DNA		
	<213> 人工序列		
	<220>		
	<223> 合成构建体		
	<400> 59		
	cagagccatc tattgcttac atttg		25
	<210> 60		
	<211> 20		
	<212> DNA		
	<213> 人工序列		
	<220>		
	<223> 合成构建体		
	<400> 60		
	ggcctcacca ccaacttcat		20

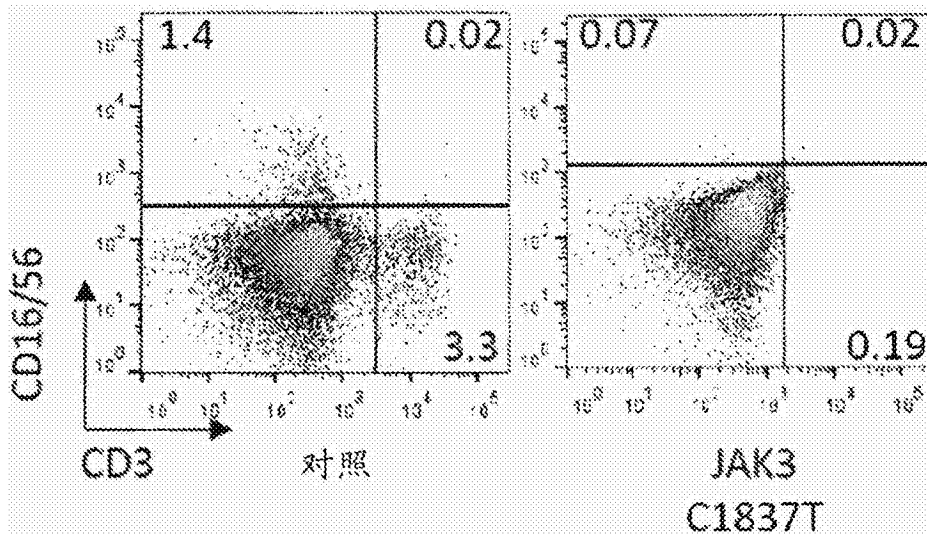


图1A

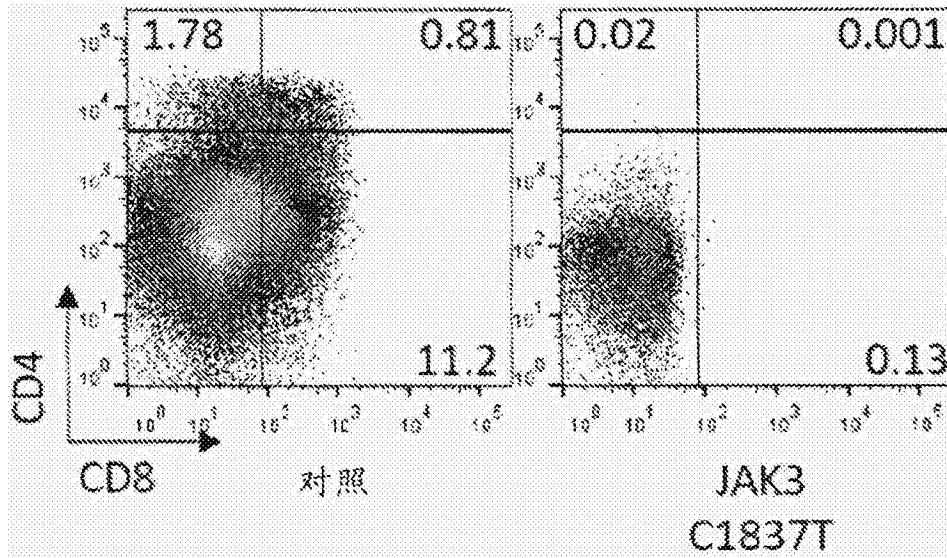


图1B

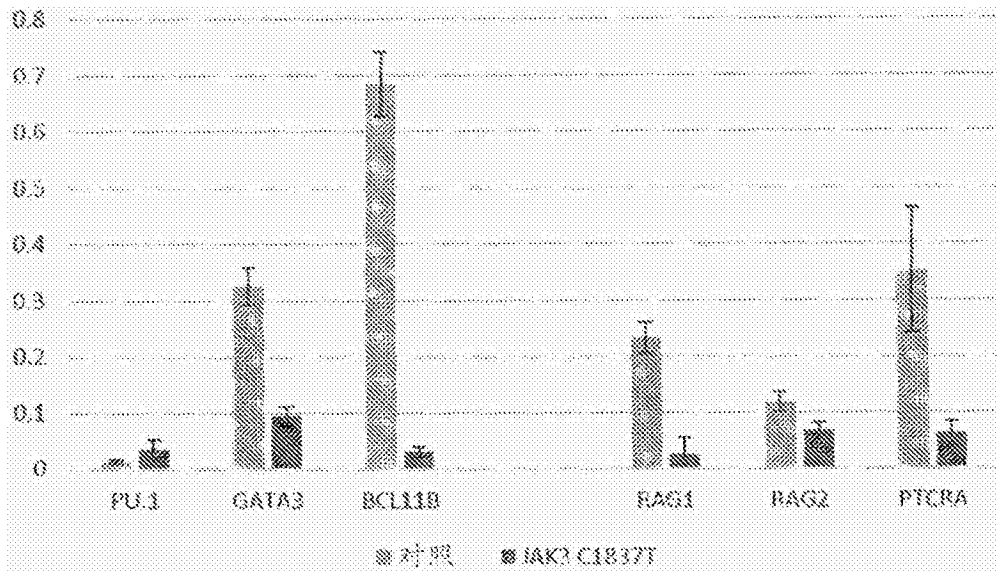


图1C

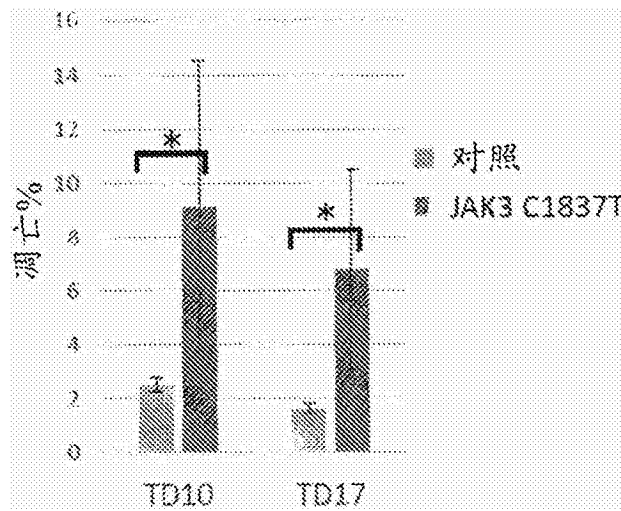


图2A

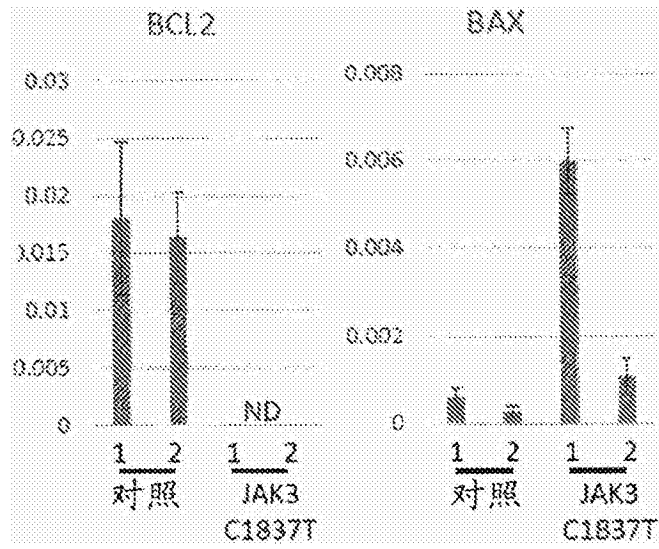


图2B

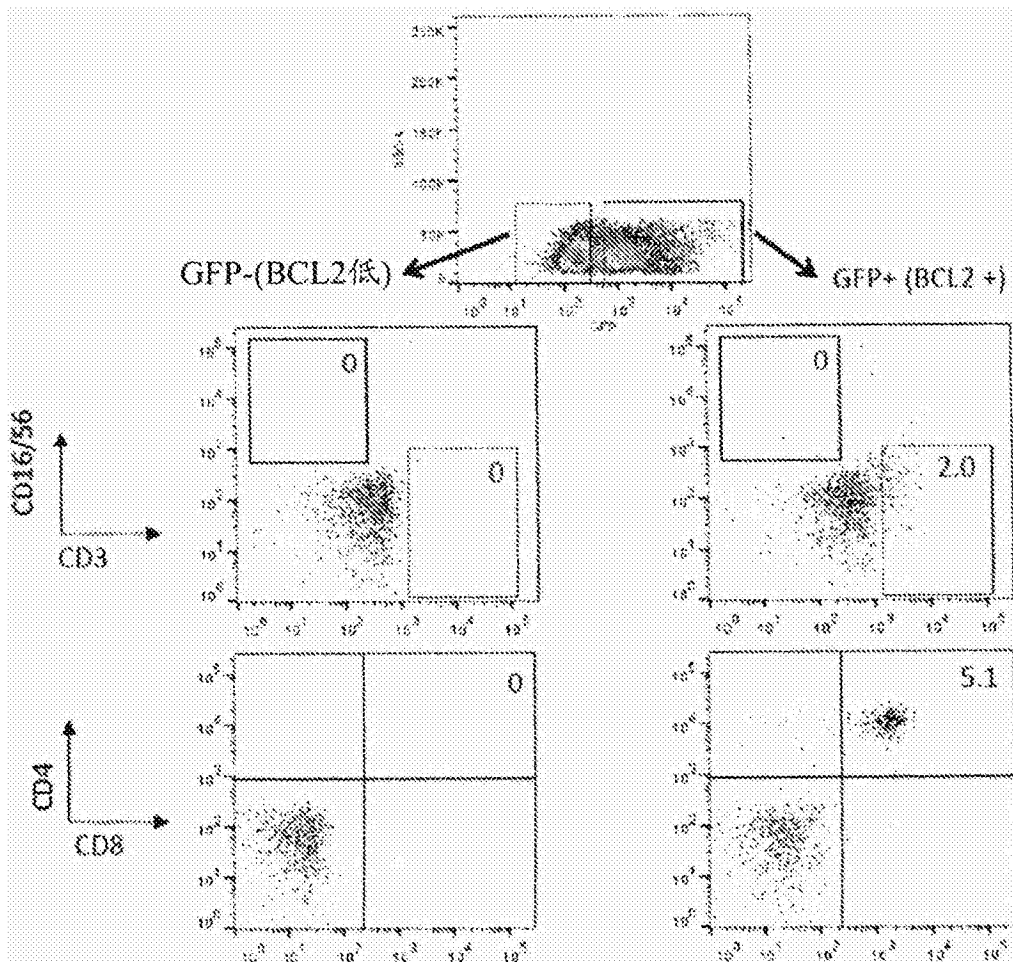


图2C

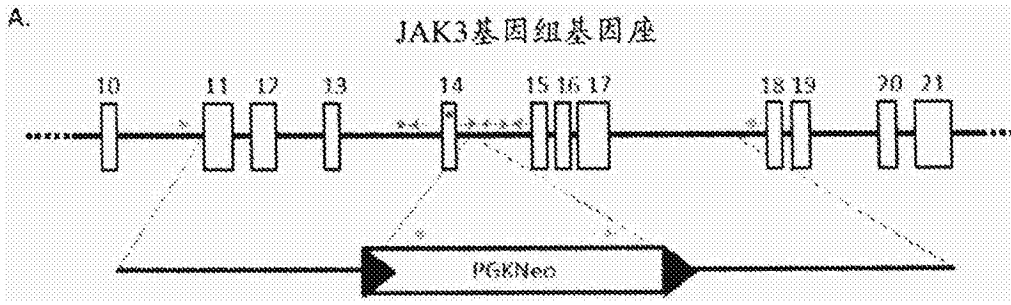


图3A

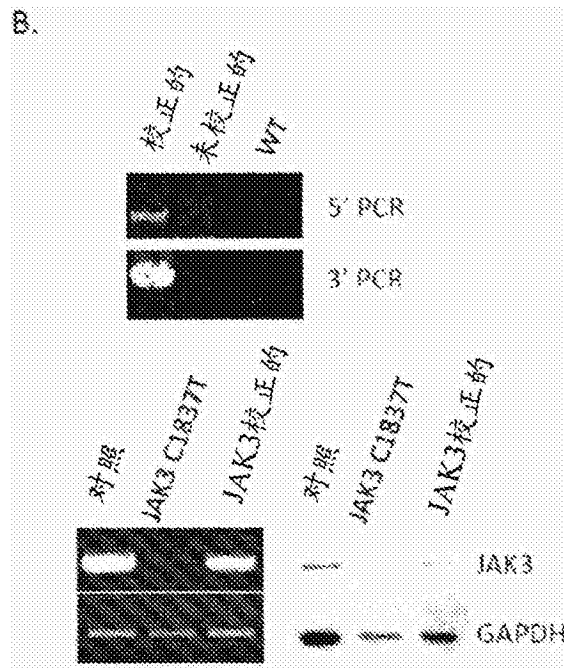


图3B

	检查的集落	PCR 阳性集落	%
gRNA #1	39	9	23
gRNA #2	45	33	73.3
gRNA #3	16	1	6.25
gRNA #4	9	3	33.3
gRNA #5	3	0	0
gRNA #6	7	0	0
gRNA #1 + #2	14	14	100
gRNA #3 + #4	3	0	0
gRNA #5 + #6	4	0	0

图3C

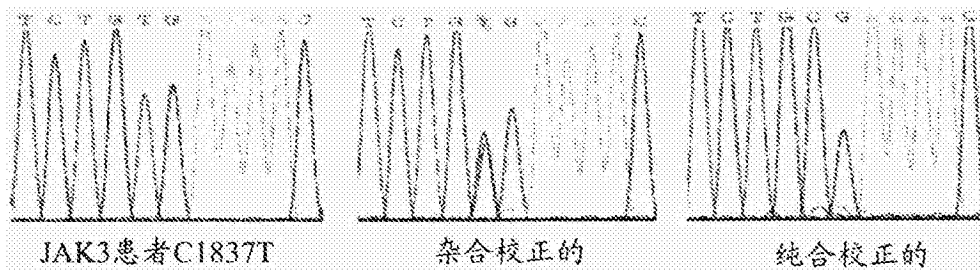


图3D

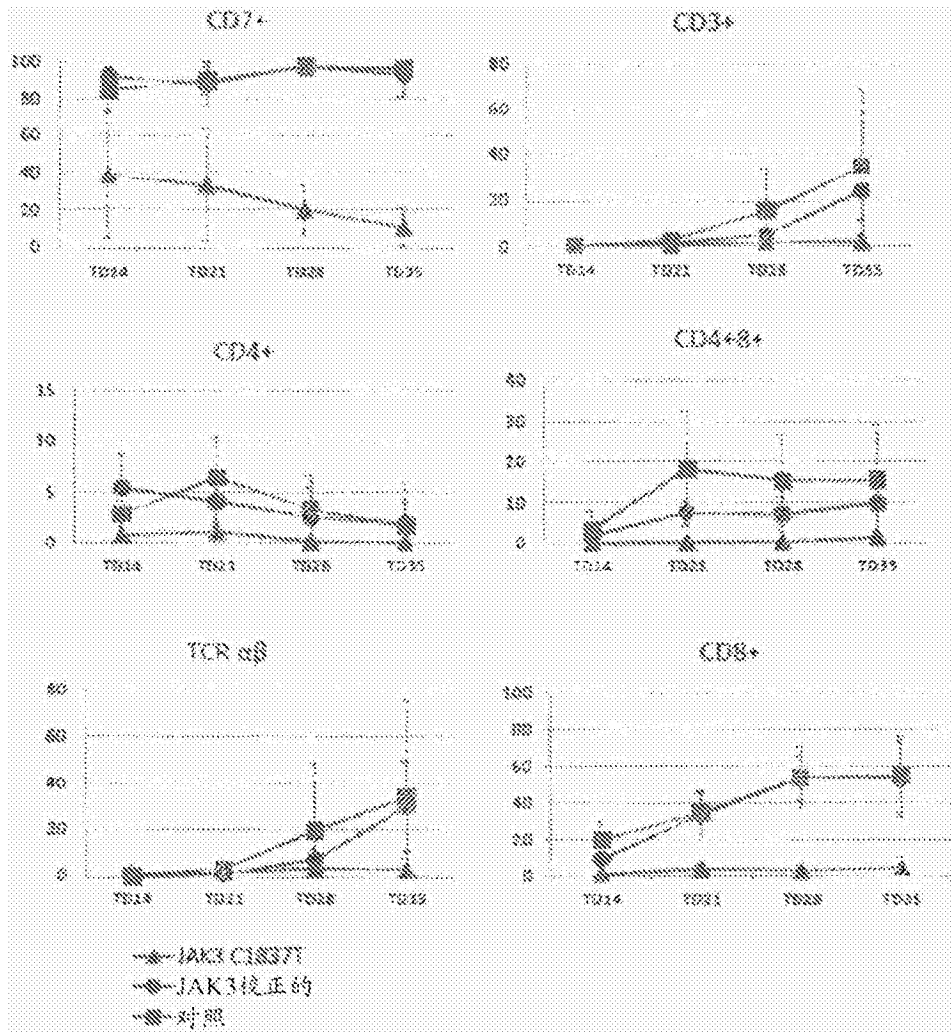


图4A

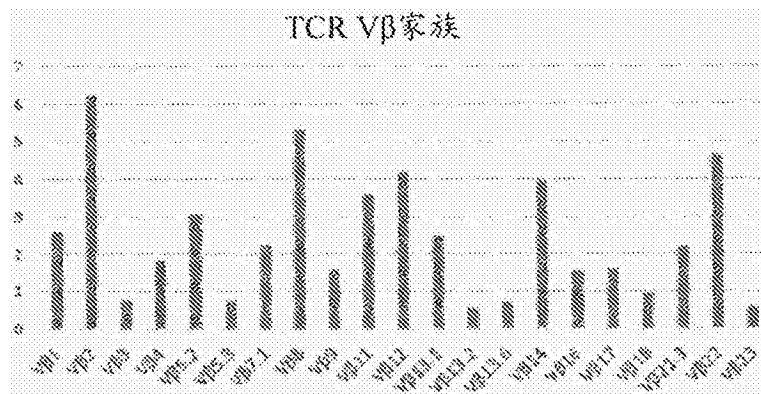


图4B

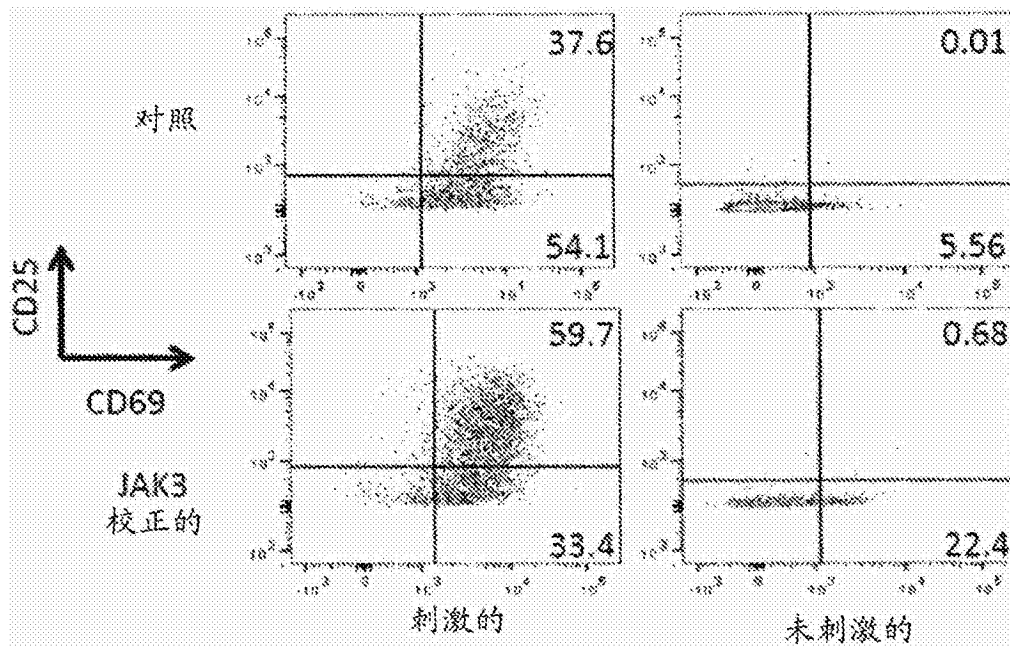


图4C

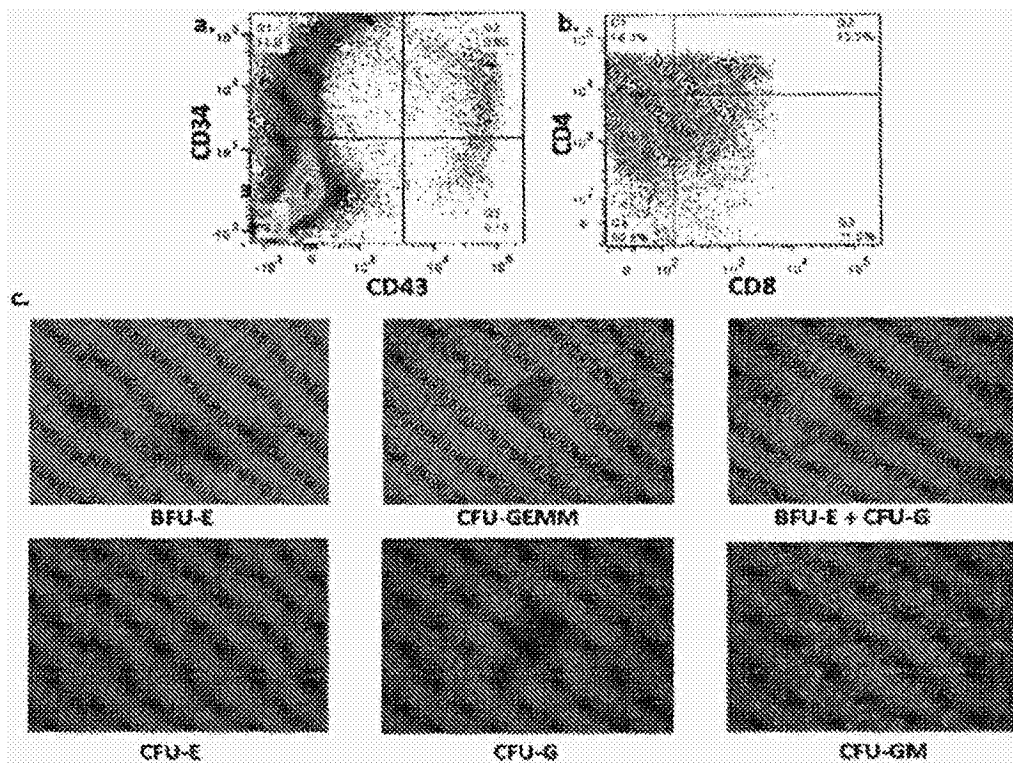


图5

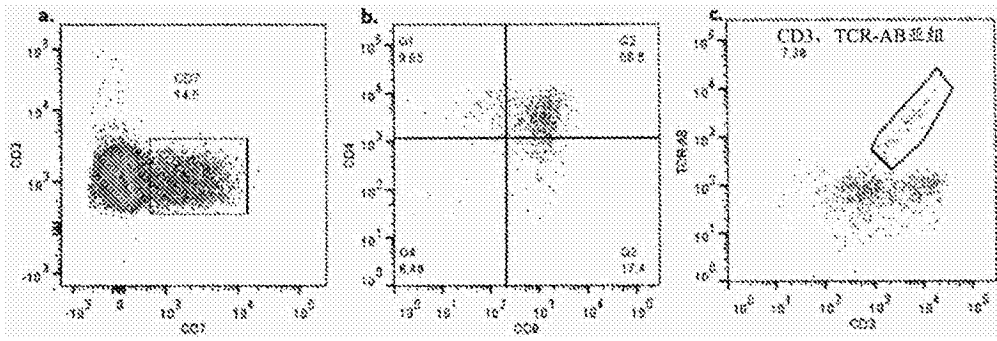


图6

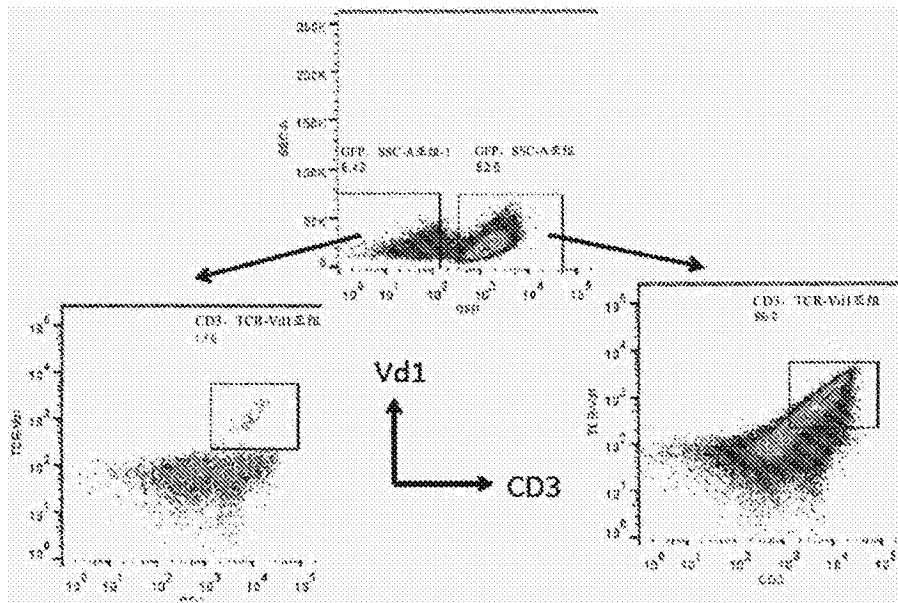


图7

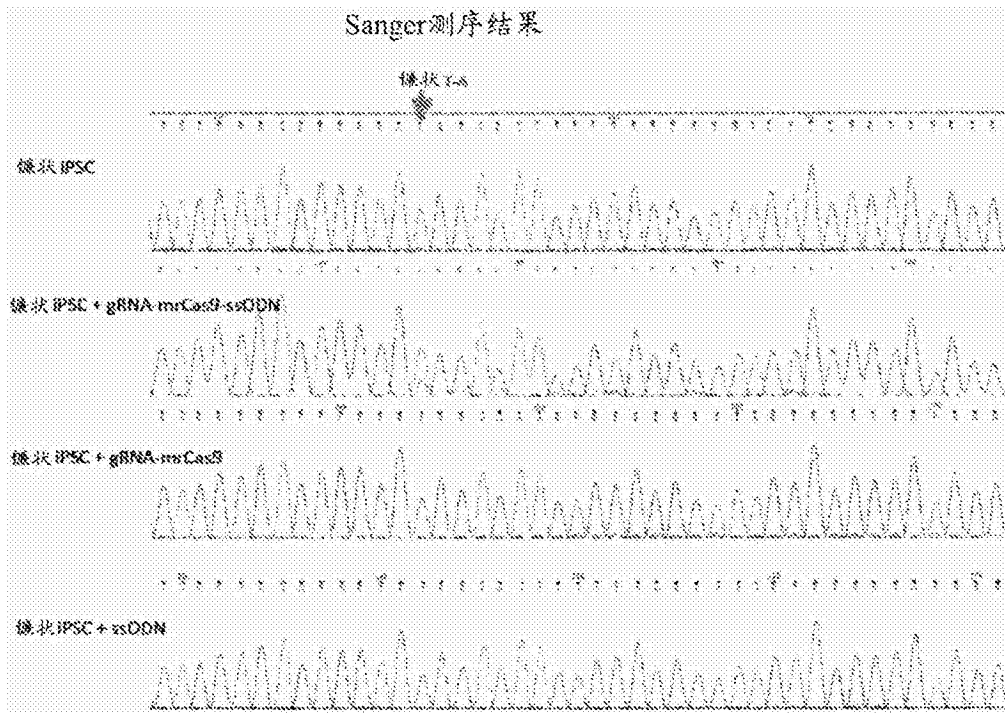


图8

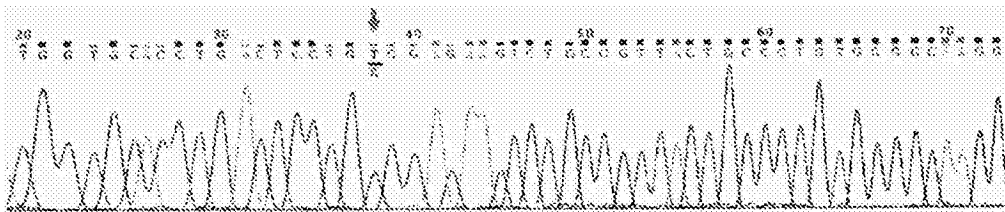


图9

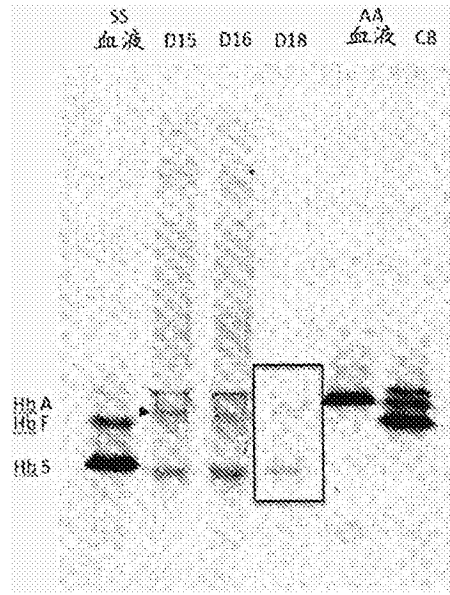


图10