



(12)发明专利申请

(10)申请公布号 CN 110938641 A

(43)申请公布日 2020.03.31

(21)申请号 201811113025.7

(22)申请日 2018.09.25

(71)申请人 上海恒润达生生物科技有限公司
地址 201210 上海市浦东新区张江路1238
弄1号楼9楼

(72)发明人 刘雅容

(51)Int.Cl.

C12N 15/62(2006.01)

C12N 15/867(2006.01)

C12N 7/01(2006.01)

C12N 5/10(2006.01)

C07K 19/00(2006.01)

A61K 35/17(2015.01)

A61P 35/00(2006.01)

权利要求书2页 说明书11页
序列表5页 附图1页

(54)发明名称

靶向APRIL的嵌合抗原受体及其用途

(57)摘要

本发明涉及靶向APRIL的嵌合抗原受体及其用途。具体而言,本发明提供一种多核苷酸序列,选自:(1)含有依次连接的抗APRIL单链抗体的编码序列、人CD8 α 铰链区的编码序列、人CD8跨膜区的编码序列、人CD28胞内区的编码序列、人41BB胞内区的编码序列、人CD3 ζ 胞内区的编码序列和任选的EGFR的含胞外结构域III和胞外结构域IV的片段的编码序列的多核苷酸序列;和(2)(1)所述多核苷酸序列的互补序列。本发明还提供相关的融合蛋白、含所述编码序列的载体,以及所述融合蛋白、编码序列、载体的用途。

1. 一种多核苷酸序列,所述多核苷酸序列选自:

(1) 含有依次连接的抗APRIL单链抗体的编码序列、人CD8 α 铰链区的编码序列、人CD8跨膜区的编码序列、人CD28胞内区的编码序列、人41BB胞内区的编码序列、人CD3 ζ 胞内区的编码序列和任选的EGFR的含胞外结构域III和胞外结构域IV的片段的编码序列的多核苷酸序列;和

(2) (1)所述多核苷酸序列的互补序列。

2. 如权利要求1所述的多核苷酸序列,其特征在于,

在所述抗APRIL单链抗体的编码序列前的所述信号肽的编码序列如SEQ ID NO:1第1-63位核苷酸序列所示;和/或

所述抗APRIL单链抗体的编码序列如SEQ ID NO:1第64-501位核苷酸序列所示;和/或

所述人CD8 α 铰链区和CD8跨膜区的编码序列如SEQ ID NO:1第502-708位核苷酸序列所示;和/或

所述人CD28胞内区的编码序列如SEQ ID NO:1第709-831位核苷酸序列所示;和/或

所述人41BB胞内区的编码序列如SEQ ID NO:1第832-975位核苷酸序列所示;和/或

所述人CD3 ζ 胞内区的编码序列如SEQ ID NO:1第976-1308位核苷酸序列所示;和/或

所述EGFR的片段的编码序列如SEQ ID NO:1第1453-2457位核苷酸序列所示。

3. 一种融合蛋白,所述融合蛋白选自:

(1) 含有依次连接的抗APRIL单链抗体、人CD8 α 铰链区、人CD8跨膜区、人CD28胞内区、人41BB胞内区和人CD3 ζ 胞内区的融合蛋白和任选的EGFR的含胞外结构域III和胞外结构域IV的片段的编码序列;和

(2) 在(1)限定的氨基酸序列中经过取代、缺失或添加一个或几个氨基酸且保留活化T细胞活性的由(1)衍生的融合蛋白。

4. 如权利要求3所述的融合蛋白,其特征在于,所述融合蛋白具有以下一个或多个特征:

所述融合蛋白序列在所述抗APRIL单链抗体的编码序列前还含有信号肽的编码序列,优选地,所述信号肽的氨基酸序列如SEQ ID NO:2第1-21位氨基酸所示;和/或

所述抗APRIL单链抗体的氨基酸序列如SEQ ID NO:2第22-167位氨基酸所示;和/或

所述人CD8 α 铰链区和跨膜区的氨基酸序列如SEQ ID NO:2第168-236位氨基酸所示;和/或

所述人CD28胞内区的氨基酸序列如SEQ ID NO:2第237-277位氨基酸所示;和/或

所述人41BB胞内区的氨基酸序列如SEQ ID NO:2第278-325位氨基酸所示;和/或

所述人CD3 ζ 胞内区的氨基酸序列如SEQ ID NO:2第326-436位氨基酸所示;和/或

所述EGFR的片段含有EGFR的胞外结构域III、胞外结构域IV以及跨膜区,或由EGFR的胞外结构域III、胞外结构域IV以及跨膜区组成;优选地,所述片段含有人EGFR的第310-646位氨基酸序列,或由人EGFR的第310-646位氨基酸序列组成;优选地,所述片段的氨基酸序列如SEQ ID NO:2第485-819位氨基酸所示。

5. 一种核酸构建物,所述核酸构建物含有权利要求1-3中任一项所述的多核苷酸序列;

优选地,所述核酸构建物为载体;

更优选地,所述核酸构建物为逆转录病毒载体,含有复制起始位点,3' LTR,5' LTR,以及

权利要求1-2中任一项所述的多核苷酸序列。

6. 一种逆转录病毒,所述逆转录病毒含有权利要求5所述的核酸构建物,优选含有所述载体,更优选含有所述逆转录病毒载体。

7. 一种基因修饰的T细胞或含该基因修饰的T细胞的药物组合物,其特征在于,所述细胞含有权利要求1-2中任一项所述的多核苷酸序列,或含有权利要求5所述的核酸构建物,或感染了权利要求6所述的逆转录病毒,或稳定表达权利要求3-4中任一项所述的融合蛋白和任选的EGFR的胞外结构域III、胞外结构域IV和任选的跨膜区的片段。

8. 权利要求1-2中任一项所述的多核苷酸序列、权利要求3-4中任一项所述的融合蛋白、权利要求5所述的核酸构建物或权利要求6所述的逆转录病毒在制备活化的T细胞中的应用。

9. 权利要求1-2中任一项所述的多核苷酸序列、权利要求3-4中任一项所述的融合蛋白、权利要求5所述的核酸构建物、权利要求6所述的逆转录病毒、或权利要求7所述的基因修饰的T细胞或其药物组合物在制备治疗APRIL介导的疾病的药物中的用途;

优选地,所述APRIL介导的疾病为多发性骨髓瘤。

靶向APRIL的嵌合抗原受体及其用途

技术领域

[0001] 本发明属于嵌合抗原受体领域,具体涉及靶向BCMA的嵌合抗原受体及其用途。

背景技术

[0002] 多发性骨髓瘤是一种恶性浆细胞疾病,表现为骨髓浆细胞恶性克隆性增生,分泌单克隆免疫球蛋白或其片段(M蛋白),导致骨骼、肾脏等相关靶器官或组织损伤,常见临床表现为骨痛、贫血、肾功能不全、感染等[Multiple myeloma.N Engl J Med,2011.364(11):p.1046-60.].目前多发性骨髓瘤为血液系统第二大恶性肿瘤,占血液系统恶性肿瘤的10%,多发病于男性,其发病率随着年龄的增长逐年增高,近几年更是有年轻化的趋势[Siegel,R.,et al.,Cancer statistics,2014.CA Cancer J Clin,2014.64(1):p.9-29.]。

[0003] 增殖诱导配体(APRIL),是BCMA 2个已经确定的配体中的1个,与B细胞激活因子(BAFF)相比其与浆细胞(PC)上的受体的亲和力更强特异性更高,但不存在于其他B和T谱系细胞中。APRIL除了与BCMA结合外,还与跨膜激活剂和钙调节剂受体(TACI)结合,TACI在恶性PC肿瘤患者上过表达。APRIL主要由骨髓细胞分泌,其具有维持多发性骨髓瘤浸润的异常骨髓细胞生长的作用。在离体培养实验中,破骨细胞(OC)刺激多发性骨髓瘤生长同时也刺激了疾病进展期间的骨病变,比CD14+未刺激的单核细胞分泌更多的APRIL。与BAFF相比,APRIL除了受体介导的效应,还可以促进恶性浆细胞的存活通过与硫酸乙酰肝素蛋白多糖CD138相互作用,那么CD138具有调节生长因子信号传导,细胞骨架组成和细胞粘附和迁移等作用。APRIL可以拯救白细胞介素6(IL-6)依赖的多发性骨髓瘤细胞系在培养过程中剥夺了IL-6免受凋亡。总的来说,这些结果表明存在一种潜在的治疗策略,即用APRIL单克隆抗体来防止骨髓瘤细胞中的BCMA活化。

[0004] 靶向BCMA的作用可以被补充通过阻断其与配体APRIL和BAFF的结合,由于MM患者的亚细胞的也表达TACI,其能够与APRIL和BAFF相结合。但是,TACI表达模式并不受BCMA的限制其在正常T和B细胞中也存在。中和BAFF单克隆抗体(LY2127355)并联合硼替佐米处理的MM患者,检测实验具有良好的耐受性和安全概况。Atacicept,是一种融合蛋白,它结合并中和BAFF和APRIL,旨在通过TACI阻断B细胞的活化。它能够抑制患者MM细胞增殖和诱导细胞凋亡。第一个治疗性的APRIL中和单克隆抗体已经用于在MM的临床前研究。具体机制是阻断APRIL诱导的信号级联通过BCMA,在骨髓环境进一步抑制MM细胞的生长和存活。抗APRIL单克隆抗体可以与抗-BCMA抗体药物结合或与裸的单克隆抗体结合以进一步增强其中任一细胞的毒性,并克服骨髓微环境中的耐药性。

[0005] 嵌合抗原受体(Chimeric Antigen Receptor-T cell,CAR-T)T细胞是指经基因修饰后,能以MHC非限制性方式识别特定目的抗原,并且持续活化扩增的T细胞。2012年国际细胞治疗协会年会中指出生物免疫细胞治疗已经成为手术、放疗、化疗外的第四种治疗肿瘤的手段,并将成为未来肿瘤治疗必选手段。CAR-T细胞回输治疗是当前肿瘤治疗中最明确有效的免疫治疗形式。大量研究表明,CAR-T细胞可以有效的识别肿瘤抗原,引起特异性的抗

肿瘤免疫应答,显著改善患者的生存状况。

[0006] 嵌合抗原受体(CAR)是CAR-T的核心部件,赋予T细胞HLA非依赖的方式识别肿瘤抗原的能力,这使得经过CAR改造的T细胞相较于天然T细胞表面受体TCR能够识别更广泛的目标。CAR的基础设计中包括一个肿瘤相关抗原(tumor-associated antigen,TAA)结合区(通常来源于单克隆抗体抗原结合区域的scFV段),一个胞外铰链区,一个跨膜区和一个胞内信号区。目标抗原的选择对于CAR的特异性、有效性以及基因改造T细胞自身的安全性来讲都是关键的决定因素。

[0007] 随着嵌合抗原受体T细胞(Chimeric Antigen Receptor-T cell,CAR-T)技术的发展,目前CAR-T主要可划分为四代。

[0008] 第一代CAR-T细胞由胞外结合区-单链抗体(single chain fragment variable, scFv)、跨膜区(transmembrane region, TM)和胞内信号区-免疫受体酪氨酸活化基序(immunoreceptor tyrosine based activation motif, ITAM)组成,其中嵌合抗原受体各部分按如下形式连接:scFv-TM-CD3 ζ 。第一代CAR虽然能够看到一些特异性的细胞毒性,但2006年对其进行临床试验总结的时候却发现疗效差强人意。究其原因是因为第一代CAR-T细胞在病人体内很快就会耗竭,其持久性很差,以至于CAR-T细胞还没有来得及接触到大量的肿瘤细胞时就已经凋亡了该种CAR-T细胞可以激发抗肿瘤的细胞毒性效应,但是细胞因子分泌比较少,但其在体内的存活期较短不能激发持久的抗肿瘤效应(Chimeric NKGD2-modified T cells inhibit systemic T-cell lymphoma growth in a manner involving multiple cytokines and cytotoxic pathways, Cancer Res 2007, 67 (22): 11029-11036)。

[0009] 第二代CAR-T细胞优化CAR设计中T细胞活化信号区仍然是研究的热点。T细胞的完全活化有赖于双信号和细胞因子的作用。其中第一信号为特异性信号,由TCR识别抗原递呈细胞表面的抗原肽-MHC复合物所启动;第二信号为协同刺激信号。早在1998年就出现了第二代CAR(J Immunol. 1998; 161 (6): 2791-7)。第2代CAR在胞内信号肽区添加了一个协同刺激分子,即把协同刺激信号组装到CAR里面,能够更好的为CAR-T细胞提供活化信号,这样CAR识别肿瘤细胞后能够同时活化协同刺激分子和胞内信号,实现双重活化,能明显提高T细胞增殖分泌能力和抗肿瘤效应。第一个被详细研究的T细胞共刺激信号受体是CD28,它能够与靶细胞表面的B7家族成员结合。CD28的共刺激能够促进T细胞的增殖,IL-2的合成和表达以及增强T细胞抵抗凋亡的能力。随后又出现了CD134(OX40)和41BB(4-1BB)等共刺激分子,以提高T细胞的细胞毒性、增殖活性,维持T细胞应答,延长T细胞存活时间等。这样的第二代CAR在随后的临床试验中产生了意想不到的效果,从2010年起基于第二代CAR的临床报道屡次引发震动,特别是对于复发性、难治性的ALL病人,其完全缓解率高达90%以上。

[0010] 第三代CAR信号肽区整合2个以上的协同刺激分子,可使T细胞持续活化增殖,细胞因子持续分泌,T细胞杀伤肿瘤细胞的能力更加显著,即新一代的CAR可获得更强的抗肿瘤应答(Mol Ther., 2005, 12 (5): 933-941)。最典型的就是U Pen Carl June在CD28刺激因子的作用下又加了一个41BB的刺激因子。

[0011] 第四代的CAR-T细胞则加入了细胞因子或共刺激配体,例如四代CAR可以产生IL-12,其能够调节免疫微环境-增加T细胞的激活,同时激活固有免疫细胞使其发挥作用来清除靶抗原阴性的癌细胞,从而达到双向调节的作用(TRUCKs: the fourth generation of

CARs, *Expert Opin Biol Ther.*, 2015; 15 (8) : 1145-54)。

[0012] CAR-T细胞的一大优点是它们是活性药物,一旦输入,生理机制会调控T细胞的平衡、记忆形成和抗原驱动的扩增。然而,这种治疗尚未完善,T细胞会脱靶而攻击其他的组织,或扩增量过高,超出治疗所需。鉴于CAR-T细胞已被纳入标准治疗范围,设计病人或药物可控的启动或关闭机制来调控CAR-T细胞的存在是非常有用的。由于技术原因,关闭机制更易应用于T细胞。作为其中之一,iCas9系统正在临床研究当中。细胞在表达iCas9时,使用小分子化合物可诱导iCas9前体分子形成二聚体,激活凋亡途径,从而实现清除细胞的目的。在移植物抗宿主病中,小分子AP1903已被用于诱导iCas9二聚体和清除T细胞,表明了这种方法的可行性(Making Better Chimeric Antigen Receptors for Adoptive T-cell Therapy, *Clin Cancer Res*; 22 (8), 2016年4月15日)。

[0013] 另外,还可利用临床上已经使用的清除性抗体,使CAR-T细胞同时表达这些抗体针对的蛋白,如tEGFR,在治疗相关的毒性反应产生或是治疗已经完成后,通过给予抗体药物清除相应的CAR-T细胞(Rational development and characterization of humanized anti-EGFR variant III chimeric antigen receptor T cells for glioblastoma, *Sci Transl Med* 2015; 7: 275ra22)。

发明内容

[0014] 本发明第一方面提供一种多核苷酸序列,所述多核苷酸序列选自:

[0015] (1) 含有依次连接的抗APRIL单链抗体的编码序列、人CD8 α 铰链区的编码序列、人CD8跨膜区的编码序列、人CD28胞内区的编码序列、人41BB胞内区的编码序列、人CD3 ζ 胞内区的编码序列和任选的EGFR的胞外结构域III和胞外结构域IV的片段的编码序列的多核苷酸序列;和

[0016] (2) (1)所述多核苷酸序列的互补序列。

[0017] 在一个或多个实施方案中,所述多核苷酸序列在所述抗APRIL单链抗体的编码序列前还含有信号肽的编码序列。在一个或多个实施方案中,所述信号肽的氨基酸序列如SEQ ID NO:2第1-21位氨基酸所示。在一个或多个实施方案中,所述抗APRIL单链抗体的轻链可变区的氨基酸序列如SEQ ID NO:2第22-167位氨基酸所示。在一个或多个实施方案中,所述人CD8 α 铰链区和CD8跨膜区的氨基酸序列如SEQ ID NO:2第168-236位氨基酸所示。在一个或多个实施方案中,所述人CD28胞内区的氨基酸序列如SEQ ID NO:2第237-277位氨基酸所示。在一个或多个实施方案中,所述人41BB胞内区的氨基酸序列如SEQ ID NO:2第278-325位氨基酸所示。在一个或多个实施方案中,所述人CD3 ζ 胞内区的氨基酸序列如SEQ ID NO:2第326-436位氨基酸所示。在一个或多个实施方案中,所述EGFR的片段含有EGFR的胞外结构域III、胞外结构域IV以及跨膜区,或由EGFR的胞外结构域III、胞外结构域IV以及跨膜区组成。在一个或多个实施方案中,所述EGFR的片段含有人EGFR的第310-646位氨基酸序列,或由人EGFR的第310-646位氨基酸序列组成,所述片段的氨基酸序列如SEQ ID NO:2第485-819位氨基酸所示。

[0018] 在一个或多个实施方案中,在所述抗APRIL单链抗体的编码序列前的所述信号肽的编码序列如SEQ ID NO:1第1-63位核苷酸序列所示。在一个或多个实施方案中,所述抗APRIL单链抗体的编码序列如SEQ ID NO:1第64-501位核苷酸序列所示。在一个或多个实施

方案中,所述人CD8 α 铰链区和CD8跨膜区的编码序列如SEQ ID NO:1第502-708位核苷酸序列所示。在一个或多个实施方案中,所述人CD28胞内区的编码序列如SEQ ID NO:1第709-831位核苷酸序列所示。在一个或多个实施方案中,所述人41BB胞内区的编码序列如SEQ ID NO:1第832-975位核苷酸序列所示。在一个或多个实施方案中,所述人CD3 ζ 胞内区的编码序列如SEQ ID NO:1第976-1308位核苷酸序列所示。在一个或多个实施方案中,所述EGFR的片段的编码序列如SEQ ID NO:1第1453-2457位核苷酸序列所示。

[0019] 本发明第二方面提供一种融合蛋白,所述融合蛋白选自:

[0020] (1) 含有依次连接的抗APRIL单链抗体、人CD8 α 铰链区、人CD8跨膜区、人CD28胞内区、人41BB胞内区和人CD3 ζ 胞内区的融合蛋白和任选的EGFR的含胞外结构域III和胞外结构域IV的片段的编码序列;和

[0021] (2) 在(1)限定的氨基酸序列中经过取代、缺失或添加一个或几个氨基酸且保留活化T细胞活性的由(1)衍生的融合蛋白。

[0022] 在一个或多个实施方案中,所述融合蛋白在所述抗APRIL单链抗体的N端还含有信号肽。在一个或多个实施方案中,所述信号肽的氨基酸序列如SEQ ID NO:2第1-21位氨基酸所示。在一个或多个实施方案中,所述抗APRIL单链抗体的氨基酸序列如SEQ ID NO:2第22-167位氨基酸所示。在一个或多个实施方案中,所述人CD8 α 铰链区和CD8跨膜区的氨基酸序列如SEQ ID NO:2第168-236位氨基酸所示。在一个或多个实施方案中,所述人CD28胞内区的氨基酸序列如SEQ ID NO:2第237-277位氨基酸所示。在一个或多个实施方案中,所述人41BB胞内区的氨基酸序列如SEQ ID NO:2第278-325位氨基酸所示。在一个或多个实施方案中,所述人CD3 ζ 胞内区的氨基酸序列如SEQ ID NO:2第326-436位氨基酸所示。在一个或多个实施方案中,所述EGFR的片段含有EGFR的胞外结构域III、胞外结构域IV以及跨膜区,或由EGFR的胞外结构域III、胞外结构域IV以及跨膜区组成。在一个或多个实施方案中,所述EGFR片段含有人EGFR的第310-646位氨基酸序列,或由人EGFR的第310-646位氨基酸序列组成。在一个或多个实施方案中,所述EGFR片段的氨基酸序列如SEQ ID NO:2第485-819位氨基酸所示。

[0023] 本发明第三方面提供一种核酸构建物,所述核酸构建物含有本文所述的多核苷酸序列。

[0024] 在一个或多个实施方案中,所述核酸构建物为载体。在一个或多个实施方案中,所述核酸构建物为逆转录病毒载体,含有复制起始位点,3' LTR,5' LTR,本文所述的多核苷酸序列,以及任选的可选择的标记。

[0025] 本发明第四方面提供一种逆转录病毒,所述逆转录病毒含有本文所述的核酸构建物,优选含有所述载体,更优选含有所述逆转录病毒载体。

[0026] 本发明第五方面提供一种基因修饰的T细胞,所述细胞含有本文所述的多核苷酸序列,或含有本文所述的核酸构建物,或感染了本文所述的逆转录病毒,或稳定表达本文所述的融合蛋白和任选的EGFR的含胞外结构域III、胞外结构域IV和任选的跨膜区的片段。

[0027] 本发明第六方面提供一种含本文所述的基因修饰的T细胞的药物组合物。

[0028] 本发明第七方面提供本文所述的多核苷酸序列、融合蛋白、核酸构建物或逆转录病毒在制备活化的T细胞中的应用。

[0029] 本发明第八方面提供本文所述的多核苷酸序列、融合蛋白、核酸构建物、逆转录病

毒、或基因修饰的T细胞或其药物组合物在制备治疗APRIL介导的疾病的药物中的用途。

[0030] 在一个或多个实施方案中,所述APRIL介导的疾病为多发性骨髓瘤。

附图说明

[0031] 图1为APRIL-CAR逆转录病毒表达载体 (APRIL-28BBz-tEGFR) 示意图。

[0032] 图2为流式细胞仪显示逆转录病毒感染T细胞72小时的APRIL-28BBz-tEGFR CAR+表达效率。

具体实施方式

[0033] 本发明提供一种靶向APRIL的嵌合抗原受体 (CAR)。该CAR含有依次连接的抗APRIL单链抗体、人CD8 α 铰链区、人CD8跨膜区、人CD28胞内区、人41BB胞内区、人CD3 ζ 胞内区和任选的EGFR的胞外结构域III和胞外结构域IV的片段。

[0034] 任选地,所述轻链可变区和重链可变区可通过接头序列连接在一起。在某些实施方案中,所述抗APRIL单链抗体的氨基酸序列如SEQ ID NO:2的第22-167位氨基酸残基所示。

[0035] 适用于本发明的人CD8 α 铰链区和CD8跨膜区的氨基酸序列可如SEQ ID NO:2第168-236位氨基酸所示。

[0036] 适用于本发明的CD28可以是本领域已知的各种用于CAR的CD28。作为示范性例子,本发明使用SEQ ID NO:2第237-277位氨基酸序列所示的CD28。

[0037] 适用于本发明的41BB可以是本领域已知的各种用于CAR的41BB。作为示范性例子,本发明使用SEQ ID NO:2第278-325位氨基酸序列所示的41BB。

[0038] 适用于本发明的人CD3 ζ 胞内区可以是本领域常规用于CAR的各种人CD3 ζ 胞内区。在某些实施方案中,所述人CD3 ζ 胞内区的氨基酸序列如SEQ ID NO:2第326-436位氨基酸所示。

[0039] 形成本发明的融合蛋白的上述各部分,如抗APRIL单链抗体的可变区、人CD8 α 铰链区、人CD8跨膜区、CD28、41BB和人CD3 ζ 胞内区等,相互之间可直接连接,或者可通过接头序列连接。接头序列可以是本领域周知的适用于抗体的接头序列,例如含G和S的接头序列。通常,接头含有一个或多个前后重复的基序。例如,该基序可以是GGGS、GGGGS、SSSSG、GSGSA和GGSGG。优选地,该基序在接头序列中是相邻的,在重复之间没有插入氨基酸残基。接头序列可以包含1、2、3、4或5个重复基序组成。接头的长度可以是3~25个氨基酸残基,例如3~15、5~15、10~20个氨基酸残基。在某些实施方案中,接头序列是多甘氨酸接头序列。接头序列中甘氨酸的数量无特别限制,通常为2~20个,例如2~15、2~10、2~8个。除甘氨酸和丝氨酸来,接头中还可含有其它已知的氨基酸残基,例如丙氨酸(A)、亮氨酸(L)、苏氨酸(T)、谷氨酸(E)、苯丙氨酸(F)、精氨酸(R)、谷氨酰胺(Q)等。

[0040] 在某些实施方案中,本发明CAR的氨基酸序列如SEQ ID NO:2第22-436位氨基酸所示或如SEQ ID NO:2第1-436位氨基酸所示。在某些实施方案中,本发明的CAR的氨基酸系列中还包含如下文所述的EGFR的胞外结构域III和胞外结构域IV的片段,其信号肽以及接头序列。

[0041] 应理解,在基因克隆操作中,常常需要设计合适的酶切位点,这势必在所表达的氨

氨基酸序列末端引入了一个或多个不相干的残基,而这并不影响目的序列的活性。为了构建融合蛋白、促进重组蛋白的表达、获得自动分泌到宿主细胞外的重组蛋白、或利于重组蛋白的纯化,常常需要将一些氨基酸添加至重组蛋白的N-末端、C-末端或该蛋白内的其它合适区域内,例如,包括但不限于,适合的接头肽、信号肽、前导肽、末端延伸等。因此,本发明的融合蛋白(即所述CAR)的氨基端或羧基端还可含有一个或多个多肽片段,作为蛋白标签。任何合适的标签都可以用于本文。例如,所述的标签可以是FLAG,HA,HA1,c-Myc,Poly-His,Poly-Arg,Strep-TagII,AU1,EE,T7,4A6, ϵ ,B,gE以及Ty1。这些标签可用于对蛋白进行纯化。

[0042] 本发明也包括SEQ ID NO:2第22-436位氨基酸序列所示的CAR、SEQ ID NO:2第1-436位氨基酸序列所示的CAR或SEQ ID NO:2所示的CAR的突变体。这些突变体包括:与该CAR具有至少80%,优选至少85%,优选至少90%,优选至少95%,优选至少97%的序列相同性并保留该CAR的生物学活性(如活化T细胞)的氨基酸序列。可采用例如NCBI的BLASTp计算两条比对的序列之间的序列相同性。

[0043] 突变体还包括:在SEQ ID NO:2第22-436位所示的氨基酸序列、SEQ ID NO:2第1-436位所示的氨基酸序列或SEQ ID NO:2所示的氨基酸序列中具有一个或数个突变(插入、缺失或取代)、同时仍保留该CAR的生物学活性的氨基酸序列。所述数个突变通常指1-10个以内,例如1-8个、1-5个或1-3个。取代优选是保守性取代。例如,在本领域中,用性能相近或相似的氨基酸进行保守性取代时,通常不会改变蛋白质或多肽的功能。“性能相近或相似的氨基酸”包括例如,具有相似侧链的氨基酸残基的家族,这些家族包括具有碱性侧链的氨基酸(例如赖氨酸、精氨酸、组氨酸)、具有酸性侧链的氨基酸(例如天冬氨酸、谷氨酸)、具有不带电荷的极性侧链的氨基酸(例如甘氨酸、天冬酰胺、谷氨酰胺、丝氨酸、苏氨酸、酪氨酸、半胱氨酸)、具有非极性侧链的氨基酸(例如丙氨酸、缬氨酸、亮氨酸、异亮氨酸、脯氨酸、苯丙氨酸、甲硫氨酸、色氨酸)、具有 β -分支侧链的氨基酸(例如苏氨酸、缬氨酸、异亮氨酸)和具有芳香侧链的氨基酸(例如酪氨酸、苯丙氨酸、色氨酸、组氨酸)。因此,在本发明多肽中用来自同一侧链类的另一氨基酸残基替换一个或几个位点,将不会在实质上影响其活性。

[0044] 本发明包括编码本发明融合蛋白的多核苷酸序列。本发明的多核苷酸序列可以是DNA形式或RNA形式。DNA形式包括cDNA、基因组DNA或人工合成的DNA。DNA可以是单链的或是双链的。DNA可以是编码链或非编码链。本发明也包括编码融合蛋白的多核苷酸序列的简并变异体,即编码相同的氨基酸序列但核苷酸序列有所不同的核苷酸序列。

[0045] 本文所述的多核苷酸序列通常可以用PCR扩增法获得。具体而言,可根据本文所公开的核苷酸序列,尤其是开放阅读框序列来设计引物,并用市售的cDNA库或按本领域技术人员已知的常规方法所制备的cDNA库作为模板,扩增而得有关序列。当序列较长时,常常需要进行两次或多次PCR扩增,然后再将各次扩增出的片段按正确次序拼接在一起。例如,在某些实施方案中,编码本文所述融合蛋白的多核苷酸序列如SEQ ID NO:1第64-1308位核苷酸所示,或如SEQ ID NO:1第1-1308位核苷酸所示。

[0046] 在某些实施方案中,本发明的多核苷酸序列还包含编码EGFR的片段的核苷酸序列。

[0047] 适用于本发明的EGFR可以是本领域周知的EGFR,例如来自人的EGFR。EGFR含有N末端胞外结构域I和II,胞外结构域III,胞外结构域IV,跨膜区,近膜区结构域以及酪氨酸激

酶结构域。本发明优选使用截短的EGFR (“tEGFR”, 即本文所述的EGFR的片段), 尤其是不包括其胞内区域(近膜区结构域及酪氨酸激酶结构域)的截短的EGFR。在某些实施方案中, 还可以进一步将不包括胞内区域的EGFR进一步截短成不包括胞外结构域I和II。因此, 在某些实施方案中, 本发明使用的tEGFR含有EGFR的胞外结构域III、胞外结构域IV以及跨膜区, 或由EGFR的胞外结构域III、胞外结构域IV以及跨膜区组成。在某些实施方案中, 所述tEGFR含有人EGFR的第310-646位氨基酸序列, 或由人EGFR的第310-646位氨基酸序列组成, 其中第310-480位氨基酸序列为人EGFR的胞外结构域III, 第481-620为人EGFR的胞外结构域IV, 第621-646位氨基酸为人EGFR的跨膜区。在某些实施例中, 所述tEGFR的氨基酸序列的胞外结构域III和IV如SEQ ID NO:2第485-819位氨基酸所示的氨基酸序列。

[0048] 本发明也涉及核酸构建物, 该核酸构建物含有本文所述的多核苷酸序列, 以及与这些序列操作性连接的一个或多个调控序列。本发明所述的多核苷酸序列可以多种方式被操作以保证所述融合蛋白(CAR和/或tEGFR)的表达。在将核酸构建物插入载体之前可根据表达载体的不同或要求而对核酸构建物进行操作。利用重组DNA方法来改变多核苷酸序列的技术是本领域已知的。

[0049] 调控序列可以是合适的启动子序列。启动子序列通常与待表达蛋白的编码序列操作性连接。启动子可以是在所选择的宿主细胞中显示转录活性的任何核苷酸序列, 包括突变的、截短的和杂合启动子, 并且可以从编码与该宿主细胞同源或异源的胞外或胞内多肽的基因获得。调控序列也可以是合适的转录终止子序列, 由宿主细胞识别以终止转录的序列。终止子序列与编码该多肽的核苷酸序列的3'末端操作性连接。在选择的宿主细胞中有功能的任何终止子都可用于本发明。调控序列也可以是合适的前导序列, 对宿主细胞翻译重要的mRNA的非翻译区。前导序列与编码该多肽的核苷酸序列的5'末端可操作连接。在选择的宿主细胞中有功能的任何终止子都可用于本发明。

[0050] 在某些实施方案中, 所述核酸构建物是载体。通常通过可操作地连接本发明的多核苷酸序列至启动子, 并将构建体并入表达载体, 实现本发明多核苷酸序列的表达。该载体对于复制和整合真核细胞可为合适的。典型的克隆载体包含可用于调节期望核酸序列表达的转录和翻译终止子、起始序列和启动子。

[0051] 本发明的多核苷酸序列可被克隆入许多类型的载体。例如, 可被克隆入质粒、噬菌粒、噬菌体衍生物、动物病毒和粘粒。进一步地, 载体是表达载体。表达载体可以以病毒载体形式提供给细胞。病毒载体技术在本领域中是公知的并在例如Sambrook等(2001, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, New York)和其他病毒学和分子生物学手册中进行了描述。可用作载体的病毒包括但不限于逆转录病毒、腺病毒、腺伴随病毒、疱疹病毒和慢病毒。

[0052] 通常, 合适的载体包含在至少一种有机体中起作用的复制起点、启动子序列、方便的限制酶位点和一个或多个可选择的标记(例如, WO 01/96584; WO01/29058; 和美国专利号6,326,193)。

[0053] 例如, 在某些实施方案中, 本发明使用逆转录病毒载体, 该逆转录病毒载体含有复制起始位点, 3' LTR, 5' LTR, 本文所述的多核苷酸序列, 以及任意的可选择的标记。

[0054] 合适的启动子的一个例子为即时早期巨细胞病毒(CMV)启动子序列。该启动子序列是能够驱动可操作地连接至其上的任何多核苷酸序列高水平表达的强组成型启动子序

列。合适的启动子的另一个例子为延伸生长因子-1 α (EF-1 α)。然而,也可使用其他组成型启动子序列,包括但不限于类人猿病毒40 (SV40) 早期启动子、小鼠乳癌病毒 (MMTV)、人免疫缺陷病毒 (HIV) 长末端重复 (LTR) 启动子、MoMuLV启动子、鸟类白血病毒启动子、EB病毒即时早期启动子、鲁斯氏肉瘤病毒启动子、以及人基因启动子,诸如但不限于肌动蛋白启动子、肌球蛋白启动子、血红素启动子和肌酸激酶启动子。进一步地,也可考虑使用诱导型启动子。诱导型启动子的使用提供了分子开关,其能够在期限表达时打开可操作地连接诱导型启动子的多核苷酸序列的表达,而在当表达是不期望的时关闭表达。诱导型启动子的例子包括但不限于金属硫蛋白启动子、糖皮质激素启动子、孕酮启动子和四环素启动子。

[0055] 为了评估CAR多肽或其部分的表达,被引入细胞的表达载体也可包含可选择的标记基因或报道基因中的任一个或两者,以便于从通过病毒载体寻求被转染或感染的细胞群中鉴定和选择表达细胞。在其他方面,可选择的标记可被携带在单独一段DNA上并用于共转染程序。可选择的标记和报道基因两者的侧翼都可具有适当的调节序列,以便能够在宿主细胞中表达。有用的可选择标记包括例如抗生素抗性基因,诸如neo等等。

[0056] 报道基因用于鉴定潜在转染的细胞并用于评价调节序列的功能性。在DNA已经被引入受体细胞后,报道基因的表达在合适的时间下进行测定。合适的报道基因可包括编码荧光素酶、 β -半乳糖苷酶、氯霉素乙酰转移酶、分泌型碱性磷酸酶或绿色荧光蛋白基因的基因。合适的表达系统是公知的并可利用已知技术制备或从商业上获得。

[0057] 将基因引入细胞和将基因表达入细胞的方法在本领域中是已知的。载体可通过在本领域中的任何方法容易地引入宿主细胞,例如,哺乳动物、细菌、酵母或昆虫细胞。例如,表达载体可通过物理、化学或生物学手段转移入宿主细胞。

[0058] 将多核苷酸引入宿主细胞的物理方法包括磷酸钙沉淀、脂质转染法、粒子轰击、微注射、电穿孔等等。将感兴趣的多核苷酸引入宿主细胞的生物学方法包括使用DNA和RNA载体。将多核苷酸引入宿主细胞的化学手段包括胶体分散系统,诸如大分子复合物、纳米胶囊、微球、珠;和基于脂质的系统,包括水包油乳剂、胶束、混合胶束和脂质体。

[0059] 将多核苷酸引入宿主细胞的生物学方法包括使用病毒载体,特别是逆转录病毒载体,这已经成为最广泛使用的将基因插入哺乳动物例如人细胞的方法。其他病毒载体可源自慢病毒、痘病毒、单纯疱疹病毒I、腺病毒和腺伴随病毒等等。已经开发了许多基于病毒的系统,用于将基因转移入哺乳动物细胞。例如,逆转录病毒提供了用于基因传递系统的方便的平台。可利用本领域中已知的技术将选择的基因插入载体并包装入逆转录病毒颗粒。该重组病毒可随后被分离和传递至体内或离体的对象细胞。许多反转录病毒系统在本领域中是已知的。在一些实施方案中,使用腺病毒载体。许多腺病毒载体在本领域中是已知的。在一个实施方案中,使用慢病毒载体。

[0060] 因此,在某些实施方案中,本发明还提供用于活化T细胞的逆转录病毒,该病毒含有本文所述的逆转录病毒载体以及相应的包装基因,如gag、pol和vsvg。

[0061] 适用于本发明的T细胞可以是各种来源的各种类型的T细胞。例如,T细胞可来源于B细胞恶性肿瘤患者的PBMC。

[0062] 在某些实施方案中,获得T细胞后,可先用适量的(例如30~80ng/ml,如50ng/ml)的CD3抗体刺激活化,然后在含有适量的(例如30~80IU/ml,如50IU/ml)的IL2培养基进行培养备用。

[0063] 因此,在某些实施方案中,本发明提供一种基因修饰的T细胞,该基因修饰的T细胞含有本文所述的多核苷酸序列,或含有本文所述的逆转录病毒载体,或感染了本文所述的逆转录病毒,或采用本文所述的方法制备得到,或稳定表达本文所述的融合蛋白和任选的tEGFR。

[0064] 本发明的CAR-T细胞可经历稳固的体内T细胞扩展并在血液和骨髓中以高水平持续延长的时间量,并形成特异性记忆T细胞。不希望被任何具体的理论所束缚,在遇到并随后消除表达替代抗原的靶细胞后,本发明的CAR-T细胞可体内分化成中心记忆样状态。

[0065] 本发明还包括一类细胞疗法,其中T细胞被基因修饰以表达本文所述的CAR和任选的tEGFR,和CAR-T细胞被注入需要其的接受者中。注入的细胞能够杀死接受者的肿瘤细胞。不像抗体疗法,CAR-T细胞能够体内复制,产生可导致持续肿瘤控制的长期持久性。

[0066] 由CAR-T细胞引起的抗肿瘤免疫应答可为主动或被动免疫应答。另外,CAR介导的免疫应答可为过继免疫疗法步骤的一部分,其中CAR-T细胞诱导对CAR中的抗原结合部分特异性的免疫应答。

[0067] 因此,可采用本发明的CAR、其编码序列、核酸构建物、表达载体、病毒以及CAR-T细胞治疗的疾病优选为BCMA介导的疾病。

[0068] 本发明的CAR-修饰的T细胞可被单独施用或作为药物组合物与稀释剂和/或其他组分诸如相关的细胞因子或细胞群结合施用。简单地说,本发明的药物组合物可包括如本文所述的CAR-T细胞,与一种或多种药学或生理学上可接受载体、稀释剂或赋形剂结合。这样的组合物可包括缓冲液诸如中性缓冲盐水、硫酸盐缓冲盐水等等;碳水化合物诸如葡萄糖、甘露糖、蔗糖或葡聚糖、甘露醇;蛋白质;多肽或氨基酸诸如甘氨酸;抗氧化剂;螯合剂诸如EDTA或谷胱甘肽;佐剂(例如,氢氧化铝);和防腐剂。

[0069] 本发明的药物组合物可以以适于待治疗(或预防)的疾病的方式施用。施用的数量和频率将由这样的因素确定,如患者的病症、和患者疾病的类型和严重程度。

[0070] 当指出“免疫学上有效量”、“抗肿瘤有效量”、“肿瘤-抑制有效量”或“治疗量”时,待施用的本发明组合物的精确量可由医师确定,其考虑患者(对象)的年龄、重量、肿瘤大小、感染或转移程度和病症的个体差异。可通常指出:包括本文描述的T细胞的药物组合物可以以 10^4 至 10^9 个细胞/kg体重的剂量,优选 10^5 至 10^6 个细胞/kg体重的剂量。T细胞组合物也可以以这些剂量多次施用。细胞可通过使用免疫疗法中公知的注入技术(见例如Rosenberg等,New Eng.J.of Med.319:1676,1988)施用。对于具体患者的最佳剂量和治疗方案可通过监测患者的疾病迹象并因此调节治疗由医学领域技术人员容易地确定。

[0071] 对象组合物的施用可以以任何方便的方式进行,包括通过喷雾法、注射、吞咽、输液、植入或移植。本文描述的组合物可被皮下、皮内、瘤内、结内、脊髓内、肌肉内、通过静脉内注射或腹膜内施用给患者。在一个实施方案中,本发明的T细胞组合物通过皮内或皮下注射被施用给患者。在另一个实施方案中,本发明的T细胞组合物优选通过静脉注射施用。T细胞的组合物可被直接注入肿瘤、淋巴结或感染位置。

[0072] 在本发明的一些实施方案中,本发明的CAR-T细胞或其组合物可与本领域已知的其它疗法结合。所述疗法包括但不限于化疗、放疗和免疫抑制剂。例如,可结合本领域周知的治疗BCMA介导的疾病的放疗或化疗制剂进行治疗。

[0073] 本文中,“抗肿瘤效应”指一种生物学效应,其可由肿瘤体积的减少、肿瘤细胞数的

减少、转移数的减少、预期寿命的增加或与癌相关的各种生理症状的改善表示。

[0074] “患者”、“对象”、“个体”等等在本文中可交换使用,指可引起免疫应答的活有机体,如哺乳动物。例子包括但不限于人、狗、猫、小鼠、大鼠和其转基因物种。

[0075] 本发明采用抗APRIL抗体的基因序列,并从NCBI GenBank数据库中搜索到人的CD8 α 较链区、人的CD8跨膜区、人的CD28胞内区、人的41BB胞内区和人的CD3 ζ 胞内区基因序列信息,全基因合成嵌合抗原受体抗APRIL scFv-CD8较链区-CD8TM-CD28-41BB-CD3 ζ -tEGFR的基因片段,插入到逆转录病毒载体中。重组质粒在293T细胞中包装病毒,感染T细胞,使T细胞表达该嵌合抗原受体。本发明实现嵌合抗原受体基因修饰的T淋巴细胞的转化方法是基于逆转录病毒转化方法。该方法具有转化效率高,外源基因能够稳定表达,且可以缩短体外培养T淋巴细胞到达临床级数量的时间等优点。在该转基因T淋巴细胞表面,转化的核酸通过转录、翻译表达在其上。更进一步地,本发明的CAR还携带tEGFR组件,该组件的空间构象可与药物级别的抗EGFR单克隆抗体西妥昔单抗紧密结合,可以作为细胞表面的标记,同时也适合T细胞的体内追踪(可通过流式和免疫组化检测);也可以在体内被妥西单抗(cetuximab)清除,即不希望本发明的CAR发挥作用时可加入妥西单抗,安全有效的控制该CAR-T细胞在体内发挥作用。因此,本发明的CAR还具有体内示踪和安全开关的作用。

[0076] 本发明通过参考以下实验实施例进一步详细地进行描述。这些实施例仅出于说明性的目的提供,并不意欲为限制性的,除非另有规定。因此,本发明决不应被解释为限于以下实施例,而是应被解释为包括由于本文提供的教导变得显而易见的任何和全部的变化。实施例中所用的方法和试剂,除非另有说明,否则为本领域常规的方法和试剂。

[0077] 实施例1:APRILscFv-CD8 α -CD28-41BB-CD3 ζ -tEGFR基因序列的确定

[0078] 从NCBI网站数据库搜索到人CD8 α 较链区、人CD8 α 跨膜区、CD28、41BB胞内区和人CD3 ζ 胞内区、和截短的EGFR基因序列信息这些序列在网站<http://sg.idtdna.com/site>上进行密码子优化,保证在编码氨基酸序列不变的情况下更适合人类细胞表达。

[0079] 采用重叠PCR将上述序列依次按抗APRIL scFv、人CD8 α 较链区基因、人CD8 α 跨膜区基因、CD28胞内区基因、41BB胞内区基因、人CD3 ζ 胞内区、和截短的EGFR基因序列进行连接,在各序列连接处引入不同酶切位点,形成完整的APRIL-28BBz-tEGFR基因序列信息。

[0080] 用NotI (NEB) 和EcoRI (NEB) 双酶切该CAR分子的核苷酸序列,经T4连接酶(NEB) 连接插入逆转录病毒MSCV (Addgene) 的NotI-EcoRI位点,转化到感受态大肠杆菌(DH5 α)。

[0081] 将重组质粒送上海生工生物技术有限公司进行测序,将测序结果与拟合成的APRIL-28BBz-tEGFRCAR序列比对来验证序列是否正确。测序引物为:

[0082] 正义:AGCATCGTTCTGTGTTGTCTC (SEQ ID NO:3);

[0083] 反义:TGTTTGTCTTGTGGCAATACAC (SEQ ID NO:4)。

[0084] 经测序正确后,使用Qiagen公司的质粒纯化试剂盒提取并纯化质粒,纯化质粒的质粒磷酸钙法转染293T细胞进行逆转录病毒包装实验。

[0085] 本实施例所构建得到的质粒图谱如图1所示。

[0086] 实施例2:逆转录病毒包装

[0087] 1、第1天:293T细胞应是小于20代,不过分长满的。以 0.6×10^6 细胞/ml铺板,10cm皿添加10ml的DMEM培养基,充分混匀细胞,37 $^{\circ}$ C培养过夜;

[0088] 2、第2天:293T细胞融合度达到90%左右进行转染(通常是铺板14-18h左右);准备

质粒复合物,各种质粒的量为MSCV骨架12.5ug,Gag-pol 10ug,VSVg 6.25ug,CaCl₂ 250ul,H₂O 1ml,总体积为1.25ml;在另一个管里添加跟质粒复合物等体积的HBS,边加质粒复合物边涡旋震荡20s。温柔地将混合物沿着边加入到293T皿中,37℃培养4h,去除培养基,PBS洗一遍,重新加入预热的新鲜培养基。

[0089] 3、第4天:转染48h后收集上清并用0.45um滤器过滤后分装保存于-80℃,继续添加预热的新鲜DMEM培养基。

[0090] 实施例3:逆转录病毒感染人的T细胞

[0091] 1、用Ficcol分离液(天津灏洋)分离获得较纯的CD3+T细胞,用含5%AB血清X-VIVO(LONZA)培养基调整细胞密度为 1×10^6 /mL。将细胞以1ml/孔接种到预先用抗人50ng/ml CD3抗体(北京同立海元)和50ng/ml CD28抗体(北京同立海元),再加入100IU/ml的白细胞介素2(北京双鹭),刺激培养48小时后用实施例3制备得到的病毒感染;

[0092] 2、T细胞活化培养后隔天,PBS稀释至终浓度为15μg/ml的Retronectin(Takara)包被非组织处理培养板,24孔板每孔250μl。避光,4℃过夜备用。

[0093] 3、T细胞活化培养两天后,取出2块包被好的24孔板,吸弃包被液,加入含2%BSA的HBSS室温封闭30min。封闭液体积为每孔500μl,吸弃封闭液,用含2.5%HEPES的HBSS洗板两次。

[0094] 4、将用实施例3制备得到的病毒液加入孔内,每孔加2ml病毒液,32℃,2000g,离心2h。

[0095] 5、弃去上清液,24孔板每孔加入活化后的T细胞 1×10^6 个,体积1ml,培养基为T细胞培养基中添加IL-2 200IU/ml。30℃,1000g,离心10min。

[0096] 6、离心完毕后,将培养板置于37℃,5%CO₂培养箱中培养。

[0097] 7、感染后24h,将细胞悬液吸出,1200rpm,4℃,离心7min。

[0098] 8、细胞感染后,每天观察细胞的密度,适时补加含IL-2 100IU/ml的T细胞培养液,使T细胞的密度维持在 5×10^5 /ml左右,使细胞扩增。

[0099] 由此获得分别感染了实施例2所示逆转录病毒的CART细胞,命名为APRIL-28BBz-tEGFR CART细胞(表达实施例1的APRIL-28BBz-tEGFR CAR)。

[0100] 实施例4:流式细胞仪检测感染后T淋巴细胞的比例及表面CAR蛋白的表达

[0101] 分别离心收集感染后72小时的实施例4制备得到的CAR-T细胞和NT细胞(对照组),PBS洗涤1次后弃上清,加入相应的抗体避光30min后PBS洗涤,重悬,最后流式细胞仪检测。CAR+由抗小鼠IgG F(ab')₂抗体(Jackson Immunoresearch)检测。

[0102] 图2显示,使用实施例3制备得到的逆转录病毒感染T细胞72小时后,APRIL-28BBz-tEGFR CAR+的表达效率达37.2%。

序列表

- <110> 上海恒润达生生物科技有限公司
 <120> 靶向APRIL的嵌合抗原受体及其用途
 <160> 4
 <170> SIPOSequenceListing 1.0
 <210> 1
 <211> 2628
 <212> DNA
 <213> 人工序列 (Homo sapiens)
 <400> 1

```

atggctctgc ctgtgaccgc cctgctgctg cctctggctc tgctgctgca cgccgctcgg 60
cctgacatcg ttttgacaca atctctcgcg tcattggcca tgagtctcgg gaagcgcgca 120
acaatatcct gtcgcgccag tgaatctgtg tctgtgatag gagcgcactt gatccattgg 180
tatcagcaga aacctggaca acctcccaag ctgctcatct acctgccag taaccttgaa 240
acaggagtac ctgctcgggtt ttcaggttcc gggtcaggga cggatttcac tttgactatc 300
gaccagttg aggaagacga cgtagccata tatagctgcc tgcagtctcg gatcttcccg 360
cgcacgttcg ggggaggaac taagctggag attaagggcg gcgggggttc tgggtggcggc 420
ggcagcggcg gtggaggatc acaaatccaa ctggttcagt ccggtccaga actgaaaaag 480
ccgggggaga cggtgaaaat ctctgtaag gcctcaggtt ataccttcac cgattacagc 540
atcaattggg taaagcgggc tccaggaaa ggtctgaaat ggatgggttg gatcaacaca 600
gaaacccgag aaccagccta tgcttacgac tttcgagtc gattcgcttt ttccttgaa 660
acttccgcaa gcacagccta tctgcaaatc aacaatctca agtacgaaga tacggccacg 720
tatttttgtg ccctggatta cagctatgca atggattact ggggtcaggg gacgtctgtt 780
acagtttcta gtactacaac tccagcacc agaccccta cacctgctcc aactatcgca 840
agtcagcccc tgtcactgcg ccctgaagcc tgtcgccctg ctgccggggg agctgtgcat 900
actcggggac tggactttgc ctgtgatata tacatctggg cgcccttggc cgggacttgt 960
ggggctcctt tcctgtcact ggttatcacc ctttactgca ggttcagtgt cgtgaagaga 1020
ggccggaaga agctgctgta catcttcaag cagccttca tgaggcccgt gcagactacc 1080
caggaggaag atggatgcag ctgtagattc cctgaagagg aggaaggagg ctgtgagctg 1140
agagtgaagt tctcccgaag cgcagatgcc ccagcctatc agcagggaca gaatcagctg 1200
tacaacgagc tgaacctggg aagacgggag gaatacgatg tgctggacaa aaggcggggc 1260
agagatcctg agatgggcgg caaaccaaga cggaagaacc cccaggaagg tctgtataat 1320
gagctgcaga aagacaagat ggctgaggcc tactcagaaa tcgggatgaa gggcgaaaga 1380
aggagaggaa aaggccacga cggactgtac caggggctga gtacagcaac aaaagacacc 1440
tatgacgctc tgcacatgca ggetctgcca ccaagacgag ctaaacgagg ctcaggcgcg 1500
acgaacttta gtttgctgaa gcaagctggg gatgtagagg aaaatccggg tcccatgttg 1560
ctccttgtga cgagcctcct getctgcgag ctgccccatc cagccttct cctcatcccg 1620
cggaagggtg gcaatggcat aggcattggc gagtttaaag attctctgag cataaatgct 1680

```

acgaatatta agcatttcaa gaattgtact tctattagtg gcgacctcca tattcttccg 1740
 gttgccttca ggggtgactc tttcaccac acacctccat tggatccaca agaacttgac 1800
 atcctgaaga cggttaaaga gattacaggc ttctcctta tccaagcgtg gcccgagaac 1860
 agaacggact tgcacgcctt tgagaacctc gaaataatac ggggtcggac gaagcaacac 1920
 ggccaattta gccttgcggt tgtagtctg aacattactt ctctcggcct tcgctctttg 1980
 aaagaaatca gcgacggaga tgtcatcatt agtggaaaca agaacctgtg ctacgcgaac 2040
 acaatcaact ggaagaagct cttcggctact tcaggccaaa agacaaagat tattagtaac 2100
 agaggagaga atagctgtaa ggctaccgga caagtttgtc acgccttgtg tagtccagag 2160
 ggttgctggg gaccggaacc aagggttgc gtcagttgcc ggaacgtgag tcgcgagcgc 2220
 gagtgtgtgg ataagtcaa tcttctggaa ggggaaccgc gagagtttgt agaaaattcc 2280
 gaatgtatac agtgtcatcc cgagtgtctt ccacaagca tgaatatcac atgtacaggg 2340
 aggggtcctg ataactgtat ccaatgtgca cactacatag atggctctca ctgtgtaaag 2400
 acgtgccccg ccggagtaat gggtgaaaac aacacctcg tgtggaagta cgccgatgcc 2460
 gggcatgtct gtcatttgtg tcatcccaac tgcacatatg gctgtaccgg tcttgattg 2520
 gagggctgtc caacaaacgg gccgaaaata ccgagtatcg caacaggcat ggtgggagca 2580
 cttttgcttc tctcgttgt cgcctgggc atcggcttgt tcattgtga 2628

<210> 2

<211> 819

<212> PRT

<213> 人工序列 (Homo sapiens)

<400> 2

Met Ala Leu Pro Val Thr Ala Leu Leu Leu Pro Leu Ala Leu Leu Leu
 1 5 10 15
 His Ala Ala Ala Pro Ala Val Leu Thr Gly Leu Gly Leu Leu Gly His
 20 25 30
 Ser Val Leu His Leu Val Pro Ile Ala Ala Thr Ser Leu Ala Ala Ser
 35 40 45
 Ala Val Thr Gly Val Met Thr Gly Pro Ala Leu Ala Ala Gly Ala Gly
 50 55 60
 Leu Gly Ala Gly Gly Thr Gly Val Ala Ile Gly Ala Ala Gly Val Thr
 65 70 75 80
 Leu Leu Thr Ser Gly Val Leu Pro Gly Ala Val Thr Pro Thr Met Gly
 85 90 95
 Gly Val Val Ser Ala Gly Gly Gly Gly Ala Gly Gly Thr Leu Pro Ala
 100 105 110
 Cys Ile Ala Ser Met Pro Ser His Pro Ala Ala Ala Thr Ala Ser Cys
 115 120 125
 Thr Ser Ala Gly Val Pro His Leu His Gly Gly Ala Ile Leu Ser Val
 130 135 140

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

450	455	460
Leu Leu Val Thr Ser	Leu Leu Leu Cys Gly	Leu Pro His Pro Ala Pro
465	470	475
Leu Leu Ile Pro Ala	Leu Val Cys Ala Gly	Ile Gly Ile Gly Gly Pro
485	490	495
Leu Ala Ser Leu Ser Ile	Ala Ala Thr Ala	Ile Leu His Pro Leu Ala
500	505	510
Cys Thr Ser Ile Ser Gly	Ala Leu His Ile Leu	Pro Val Ala Pro Ala
515	520	525
Gly Ala Ser Pro Thr His	Thr Pro Pro Leu Ala	Pro Gly Gly Leu Ala
530	535	540
Ile Leu Leu Thr Val Leu	Gly Ile Thr Gly Pro	Leu Leu Ile Gly Ala
545	550	555
Thr Pro Gly Ala Ala Thr	Ala Leu His Ala Pro	Gly Ala Leu Gly Ile
565	570	575
Ile Ala Gly Ala Thr Leu	Gly His Gly Gly Pro	Ser Leu Ala Val Val
580	585	590
Ser Leu Ala Ile Thr Ser	Leu Gly Leu Ala Ser	Leu Leu Gly Ile Ser
595	600	605
Ala Gly Ala Val Ile Ile	Ser Gly Ala Leu Ala	Leu Cys Thr Ala Ala
610	615	620
Thr Ile Ala Thr Leu Leu	Leu Pro Gly Thr Ser	Gly Gly Leu Thr Leu
625	630	635
Ile Ile Ser Ala Ala Gly	Gly Ala Ser Cys Leu	Ala Thr Gly Gly Val
645	650	655
Cys His Ala Leu Cys Ser	Pro Gly Gly Cys Thr	Gly Pro Gly Pro Ala
660	665	670
Ala Cys Val Ser Cys Ala	Ala Val Ser Ala Gly	Ala Gly Cys Val Ala
675	680	685
Leu Cys Ala Leu Leu Gly	Gly Gly Pro Ala Gly	Pro Val Gly Ala Ser
690	695	700
Gly Cys Ile Gly Cys His	Pro Gly Cys Leu Pro	Gly Ala Met Ala Ile
705	710	715
Thr Cys Thr Gly Ala Gly	Pro Ala Ala Cys Ile	Gly Cys Ala His Thr
725	730	735
Ile Ala Gly Pro His Cys	Val Leu Thr Cys Pro	Ala Gly Val Met Gly
740	745	750
Gly Ala Ala Thr Leu Val	Thr Leu Thr Ala Ala	Ala Gly His Val Cys
755	760	765

His Leu Cys His Pro Ala Cys Thr Thr Gly Cys Thr Gly Pro Gly Leu
 770 775 780

Gly Gly Cys Pro Thr Ala Gly Pro Leu Ile Pro Ser Ile Ala Thr Gly
 785 790 795 800

Met Val Gly Ala Leu Leu Leu Leu Val Val Ala Leu Gly Ile Gly
 805 810 815

Leu Pro Met

<210> 3

<211> 21

<212> DNA

<213> 人工序列 (Homo sapiens)

<400> 3

agcatcgttc tgtgtgtct c 21

<210> 4

<211> 22

<212> DNA

<213> 人工序列 (Homo sapiens)

<400> 4

tgtttgcctt gtggcaatac ac 22

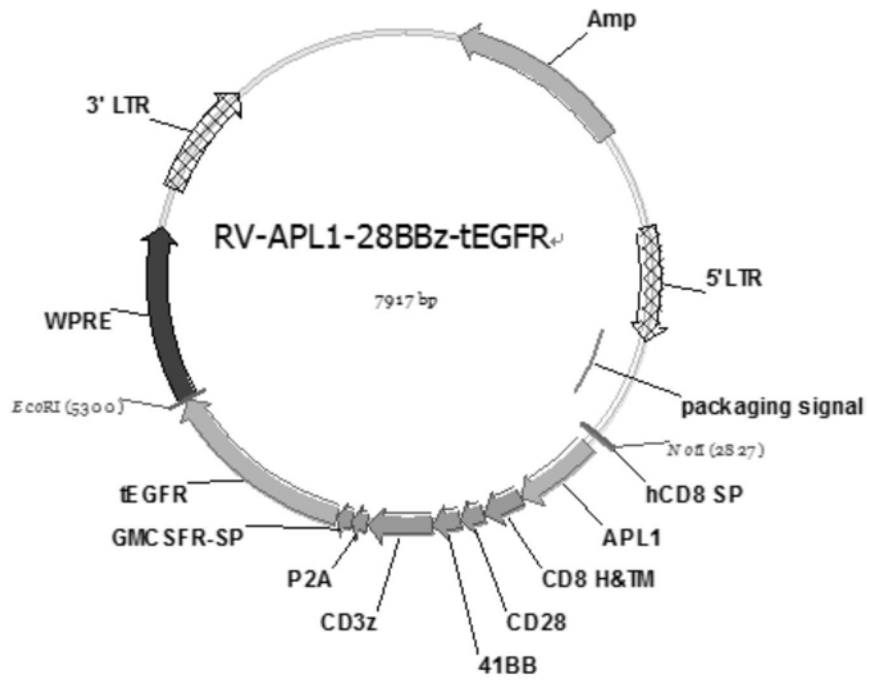


图1

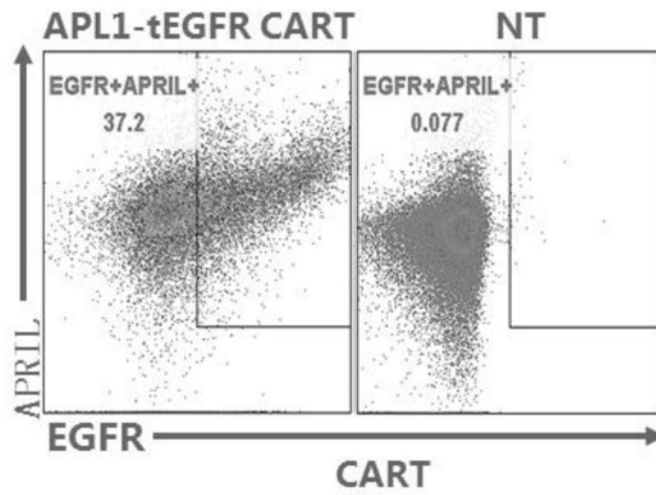


图2