



(12) 发明专利

(10) 授权公告号 CN 109844118 B

(45) 授权公告日 2024. 08. 27

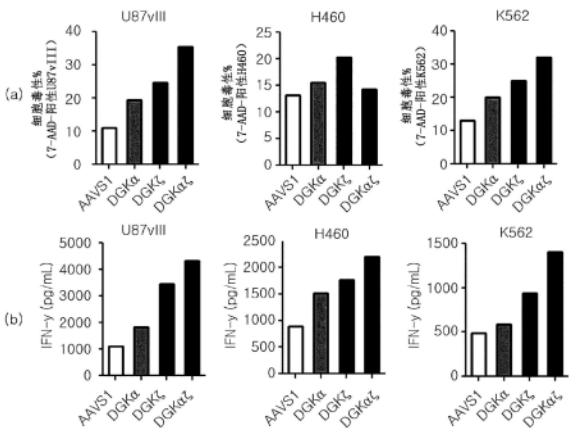
(21) 申请号 201780063250.1	(72) 发明人 金奭中 金润荣 柳浩成 郑仁英 李贞懋
(22) 申请日 2017.08.14	
(65) 同一申请的已公布的文献号 申请公布号 CN 109844118 A	(74) 专利代理机构 北京信慧永光知识产权代理有限公司 11290 专利代理师 洪俊梅 张淑珍
(43) 申请公布日 2019.06.04	
(30) 优先权数据 10-2016-0103308 2016.08.12 KR 62/502,822 2017.05.08 US	(51) Int.Cl. C12N 15/63 (2006.01) C12N 15/85 (2006.01) C12N 9/22 (2006.01) C12N 5/078 (2010.01) A61K 35/17 (2015.01)
(85) PCT国际申请进入国家阶段日 2019.04.12	
(86) PCT国际申请的申请数据 PCT/KR2017/008835 2017.08.14	(56) 对比文件 US 2016120906 A1, 2016.05.05 US 2015224142 A1, 2015.08.13
(87) PCT国际申请的公布数据 W02018/030874 KO 2018.02.15	审查员 冯小叶
(73) 专利权人 株式会社图尔金 地址 韩国首尔	权利要求书4页 说明书162页 序列表45页 附图21页

(54) 发明名称

经操纵的免疫调节因子以及由此改变的免疫力

(57) 摘要

本发明涉及具有改善的免疫效果的经人工操纵的免疫系统。更具体而言,本发明涉及功能被人工改变的免疫系统,所述免疫系统包含经人工操纵的免疫调节因子以及包含该经人工操纵的免疫调节因子的细胞。根据具体实施方式考虑在内的是包含经人工操纵的免疫调节基因(例如PD-1、CTLA-4、A20、DGK α 、DGK ζ 、FAS、EGR2、PPP2R2D、PSGL-1、KDM6A以及TET2)和/或其表达产物的免疫系统。



1. 用于生产经人工操纵的免疫细胞的体外或离体方法,所述经人工操纵的免疫细胞包含选自Dgk α 基因和Dgk ζ 基因的至少一种经人工操纵的免疫调节基因,其中,所述方法包括:

向免疫细胞中引入:

编辑蛋白或编码所述编辑蛋白的核酸,所述编辑蛋白是酿脓链球菌衍生而来的Cas9蛋白,以及

引导RNA或编码所述引导RNA的核酸,

其中,所述引导RNA为单链RNA,并且所述引导RNA以5'至3'方向包含引导结构域、第一互补结构域、接头结构域、第二互补结构域和近端结构域,

其中,所述引导RNA能够与酿脓链球菌衍生而来的Cas9蛋白形成编辑蛋白-引导RNA复合体,

其中,所述引导结构域具有这样的序列,所述序列能够靶向选自Dgk α 基因的外显子区域中的SEQ ID NO:19、SEQ ID NO:20、SEQ ID NO:21和SEQ ID NO:23的靶序列,或者选自Dgk ζ 基因的外显子区域中的SEQ ID NO:109、SEQ ID NO:110、SEQ ID NO:111、SEQ ID NO:112和SEQ ID NO:113的靶序列;以及

在所述免疫细胞的Dgk α 基因和/或Dgk ζ 基因中诱导插入缺失,其中,通过所述编辑蛋白-引导RNA复合体诱导所述插入缺失。

2. 如权利要求1所述的方法,

其中,所述引导结构域具有能够靶向Dgk α 基因的外显子区域中的SEQ ID NO:23的靶序列的序列。

3. 如权利要求1所述的方法,

其中,所述引导结构域具有能够靶向Dgk α 基因的外显子区域中的SEQ ID NO:20的靶序列的序列。

4. 如权利要求1所述的方法,

其中,所述引导结构域具有能够靶向Dgk ζ 基因的外显子区域中的SEQ ID NO:113的靶序列的序列。

5. 如权利要求1所述的方法,

其中,所述引导结构域具有能够靶向Dgk ζ 基因的外显子区域中的SEQ ID NO:111的靶序列的序列。

6. 如权利要求1所述的方法,

其中,所述方法包括向所述免疫细胞中引入第一引导RNA或编码所述第一引导RNA的核酸,以及第二引导RNA或编码所述第二引导RNA的核酸,

其中,所述第一引导RNA为单链RNA,并且所述第一引导RNA包含所述第一引导RNA的引导结构域,

其中,所述第一引导RNA的引导结构域具有能够靶向靶序列的序列,所述靶序列选自Dgk α 基因的外显子区域中的SEQ ID NO:19、SEQ ID NO:20、SEQ ID NO:21和SEQ ID NO:23,并且

其中,所述第二引导RNA为单链RNA,并且所述第二引导RNA包含所述第二引导RNA的引导结构域,

其中,所述第二引导RNA的引导结构域具有能够靶向靶序列的序列,所述靶序列选自

Dgk ζ 基因的外显子区域中的SEQ ID NO:109、SEQ ID NO:110、SEQ ID NO:111、SEQ ID NO:112和SEQ ID NO:113。

7. 如权利要求1所述的方法，

其中，所述方法包括向所述免疫细胞中引入第一引导RNA或编码所述第一引导RNA的核酸，以及第二引导RNA或编码所述第二引导RNA的核酸，

其中，所述第一引导RNA为单链RNA，并且所述第一引导RNA包含所述第一引导RNA的引导结构域，

其中，所述第一引导RNA的引导结构域具有能够靶向靶序列的序列，所述靶序列选自Dgk α 基因的外显子区域中的SEQ ID NO:20和SEQ ID NO:23，并且

其中，所述第二引导RNA为单链RNA，并且所述第二引导RNA包含所述第二引导RNA的引导结构域，

其中，所述第二引导RNA的引导结构域具有能够靶向靶序列的序列，所述靶序列选自Dgk ζ 基因的外显子区域中的SEQ ID NO:111和SEQ ID NO:113。

8. 如权利要求1所述的方法，

其中，所述方法包括向所述免疫细胞中引入第一引导RNA或编码所述第一引导RNA的核酸，以及第二引导RNA或编码所述第二引导RNA的核酸，

其中，所述第一引导RNA为单链RNA，并且所述第一引导RNA包含所述第一引导RNA的引导结构域，

其中，所述第一引导RNA的引导结构域具有能够靶向选自Dgk α 基因的外显子区域中的SEQ ID NO:23的靶序列的序列，并且

其中，所述第二引导RNA为单链RNA，并且所述第二引导RNA包含所述第二引导RNA的引导结构域，

其中，所述第二引导RNA的引导结构域具有能够靶向选自Dgk ζ 基因的外显子区域中的SEQ ID NO:113的靶序列的序列。

9. 如权利要求1所述的方法，

其中，所述引入的步骤通过选自如下方法中的一种或多种方法进行：电穿孔、脂质体、质粒、病毒载体、纳米粒子和蛋白易位结构域融合蛋白法。

10. 如权利要求9所述的方法，

其中，所述病毒载体为选自于如下的组中的至少一种：逆转录病毒、慢病毒、腺病毒、腺相关病毒、痘病毒和单纯疱疹病毒。

11. 如权利要求10所述的方法，

其中，所述病毒载体为痘苗病毒。

12. 如权利要求1所述的方法，

其中，编辑蛋白或编码所述编辑蛋白的核酸以及引导RNA或编码所述引导RNA的核酸通过一个载体引入细胞。

13. 如权利要求1所述的方法，

其中，编辑蛋白或编码所述编辑蛋白的核酸以及引导RNA或编码所述引导RNA的核酸通过两个以上的载体引入细胞。

14. 一种用于生产经人工操纵的免疫细胞的试剂盒，所述经人工操纵的免疫细胞包含

选自Dgk α 基因和Dgk ζ 基因的至少一种经人工操纵的免疫调节基因,

所述试剂盒包含:

(i) 编辑蛋白或编码所述编辑蛋白的核酸,其中,所述编辑蛋白是酿脓链球菌(*Streptococcus pyogenes*)衍生而来的Cas9蛋白,

(ii) 引导RNA或编码所述引导RNA的核酸,

其中,所述引导RNA为单链RNA,并且所述引导RNA以5'至3'方向包含引导结构域、第一互补结构域、接头结构域、第二互补结构域和近端结构域,

其中,所述引导RNA能够与酿脓链球菌衍生而来的Cas9蛋白形成编辑蛋白-引导RNA复合物,

其中,所述引导结构域具有这样的序列,所述序列能够靶向选自Dgk α 基因的外显子区域中的SEQ ID NO:19、SEQ ID NO:20、SEQ ID NO:21和SEQ ID NO:23的靶序列,或者选自Dgk ζ 基因的外显子区域中的SEQ ID NO:109、SEQ ID NO:110、SEQ ID NO:111、SEQ ID NO:112和SEQ ID NO:113的靶序列。

15. 如权利要求14所述的试剂盒,

其中,所述引导结构域具有能够靶向Dgk α 基因的外显子区域中的SEQ ID NO:23的靶序列的序列。

16. 如权利要求14所述的试剂盒,

其中,所述引导结构域具有能够靶向Dgk α 基因的外显子区域中的SEQ ID NO:20的靶序列的序列。

17. 如权利要求14所述的试剂盒,

其中,所述引导结构域具有能够靶向Dgk ζ 基因的外显子区域中的SEQ ID NO:113的靶序列的序列。

18. 如权利要求14所述的试剂盒,

其中,所述引导结构域具有能够靶向Dgk ζ 基因的外显子区域中的SEQ ID NO:111的靶序列的序列。

19. 如权利要求14所述的试剂盒,

其中,所述试剂盒包含:

第一引导RNA或编码所述第一引导RNA的核酸,

其中,所述第一引导RNA为单链RNA,并且所述第一引导RNA包含所述第一引导RNA的引导结构域,

其中,所述第一引导RNA的引导结构域具有能够靶向靶序列的序列,所述靶序列选自Dgk α 基因的外显子区域中的SEQ ID NO:19、SEQ ID NO:20、SEQ ID NO:21和SEQ ID NO:23; 以及

第二引导RNA或编码所述第二引导RNA的核酸,

其中,所述第二引导RNA为单链RNA,并且所述第二引导RNA包含所述第二引导RNA的引导结构域,

其中,所述第二引导RNA的引导结构域具有能够靶向靶序列的序列,所述靶序列选自Dgk ζ 基因的外显子区域中的SEQ ID NO:109、SEQ ID NO:110、SEQ ID NO:111、SEQ ID NO:112和SEQ ID NO:113。

20. 如权利要求14所述的试剂盒，
其中，所述试剂盒包含：
第一引导RNA或编码所述第一引导RNA的核酸，
其中，所述第一引导RNA为单链RNA，并且所述第一引导RNA包含所述第一引导RNA的引导结构域，
其中，所述第一引导RNA的引导结构域具有能够靶向选自Dgk α 基因的外显子区域中的SEQ ID NO:20和SEQ ID NO:23的靶序列的序列；以及
第二引导RNA或编码所述第二引导RNA的核酸，
其中，所述第二引导RNA为单链RNA，并且所述第二引导RNA包含所述第二引导RNA的引导结构域，
其中，所述第二引导RNA的引导结构域具有能够靶向选自Dgk ζ 基因的外显子区域中的SEQ ID NO:111和SEQ ID NO:113的靶序列的序列。
21. 如权利要求14所述的试剂盒，
其中，所述试剂盒包含：
第一引导RNA或编码所述第一引导RNA的核酸，
其中，所述第一引导RNA为单链RNA，并且所述第一引导RNA包含所述第一引导RNA的引导结构域，
其中，所述第一引导RNA的引导结构域具有能够靶向选自Dgk α 基因的外显子区域中的SEQ ID NO:23的靶序列的序列；以及
第二引导RNA或编码所述第二引导RNA的核酸，
其中，所述第二引导RNA为单链RNA，并且所述第二引导RNA包含所述第二引导RNA的引导结构域，
其中，所述第二引导RNA的引导结构域具有能够靶向选自Dgk ζ 基因的外显子区域中的SEQ ID NO:113的靶序列的序列。

经操纵的免疫调节因子以及由此改变的免疫力

技术领域

[0001] 本发明涉及具有改善的免疫效力的经人工操纵的免疫系统。更具体而言,本发明涉及人工修饰免疫系统,所述人工修饰免疫系统包含经人工操纵的免疫调节因子以及包含该经人工操纵的免疫调节因子的免疫细胞。

背景技术

[0002] 细胞治疗剂是使用活细胞诱导再生以修复受损或患病的细胞/组织/实体的药物,其为通过物理、化学或生物操纵(例如对自体细胞、同种异体细胞或异种细胞进行离体培养、增殖、选择等)产生的药物。

[0003] 其中,免疫调节细胞治疗剂是通过使用免疫细胞(例如树突细胞、自然杀伤细胞、T细胞等)调节体内的免疫应答来实现疾病治疗目的的药物。

[0004] 目前,正在开发的免疫调节细胞治疗剂主要靶向癌症治疗作为适应证。与传统用于癌症治疗的手术疗法、抗癌剂和放射疗法不同的是,免疫调节细胞治疗剂具有的治疗机制和效力在于通过将免疫细胞直接给予患者来活化免疫功能,从而获得疗效;免疫调节细胞治疗剂有望在未来的新兴生物学中扮演重要角色。

[0005] 导入细胞的抗原的物理和化学特性根据免疫调节细胞治疗剂的类型而彼此不同。当以病毒载体等形式将外源基因导入免疫细胞时,这些细胞将能够同时具有细胞治疗剂和基因治疗剂的特征。

[0006] 可采用如下方式来实施免疫调节细胞治疗剂的给予:通过利用多种抗体和细胞因子活化多种免疫细胞(例如通过单采(apheresis)从患者分离的外周血单核细胞(PBMC)、T细胞、NK细胞等),随后离体增殖并再次注射至患者中;或者将其中导入有基因(例如T细胞受体(TCR)或嵌合抗原受体(CAR))的免疫细胞再次注射至患者中。

[0007] 过继性免疫疗法涉及离体产生的自体抗原特异性免疫细胞(例如T细胞)的递送,其可能成为治疗多种免疫疾病以及癌症的有希望的策略。

[0008] 最近报道了免疫细胞治疗剂可以以多种方式使用,例如作为自体免疫抑制剂等以及表现出抗癌功能。因此,免疫细胞治疗剂可以通过调控免疫应答而用于多种适应证中。因此,在对用于过继性免疫疗法的经操纵的免疫细胞的治疗效力进行开发和改善方面具有巨大需求。

发明内容

[0009] 技术问题

[0010] 作为示例性实施方式,本发明提供了具有改善的免疫效果的经人工操纵的免疫系统。

[0011] 作为示例性实施方式,本发明提供了经人工操纵的免疫调节因子以及包含该经人工操纵的免疫调节因子的细胞。

[0012] 作为示例性实施方式,本发明提供了用于对免疫细胞的功能进行修饰(例如增强

或抑制)的方法。

[0013] 作为示例性实施方式,本发明提供了对伴有免疫异常的疾病的治疗和/或预防用途,所述用途包含免疫功能经修饰的免疫调节因子和/或免疫细胞作为有效成分。

[0014] 作为示例性实施方式,本发明通过增强免疫细胞的增殖、存活、细胞毒性、浸润和细胞因子释放来提供抗癌功能。

[0015] 作为示例性实施方式,本发明提供了免疫调节基因(例如PD-1、CTLA-4、A20、DGK α 、DGK ζ 、FAS、EGR2、PPP2R2D、PSGL-1、KDM6A、TET2等)和/或所述基因的表达产物。

[0016] 作为示例性实施方式,本发明提供了用于对免疫细胞进行基因组编辑的组合物以及使用所述组合物的方法,所述组合物包含可适用于对免疫调节基因的活性进行调节的引导核酸-编辑蛋白复合体。

[0017] 作为示例性实施方式,本发明提供了能够用于对免疫调节基因(例如PD-1、CTLA-4、A20、DGK α 、DGK ζ 、FAS、EGR2、PPP2R2D、PSGL-1、KDM6A、TET2等)进行操纵的引导核酸-编辑蛋白复合体。

[0018] 技术方案

[0019] 为了解决这些问题,本发明提供了具有改善的免疫效果的经人工操纵的免疫系统。更具体而言,本发明涉及经人工操纵的免疫系统,所述经人工操纵的免疫系统包含经人工操纵的免疫调节因子和包含所述经人工操纵的免疫调节因子的免疫细胞。

[0020] 本发明提供了用于特定目的的经基因操纵或修饰的免疫调节因子。

[0021] 术语“免疫调节因子”是其功能与免疫应答的形成和性能相关的物质,包括所有具有免疫应答调节功能的各种物质,可以为非天然(即经人工操纵的)的物质。例如,免疫调节因子可为在免疫细胞中表达的经基因操纵或修饰的基因或蛋白。

[0022] 免疫调节因子可在使免疫细胞活化或失活方面发挥功能。免疫调节因子可发挥功能来促进免疫应答(例如免疫细胞生长调节因子、免疫细胞死亡调节因子、免疫细胞功能丧失因子或细胞因子分泌元件等)。

[0023] 在本发明的示例性实施方式中,免疫调节因子可为例如经基因操纵或修饰的PD-1基因、CTLA-4基因、TNFAIP3(A20)基因、DGKA基因、DGKZ基因、FAS基因、EGR2基因、PPP2R2D基因、TET2基因、PSGL-1基因或KDM6A基因。

[0024] 在本发明的示例性实施方式中,免疫调节因子可包含两种以上经基因操纵或修饰的基因。例如,可对选自于由如下基因所组成的组中的两种以上基因进行人工操纵或修饰:PD-1基因、CTLA-4基因、TNFAIP3(A20)基因、DGKA基因、DGKZ基因、FAS基因、EGR2基因、PPP2R2D基因、TET2基因、PSGL-1基因或KDM6A基因。

[0025] 作为本发明的优选实例,免疫调节因子可为TNFAIP3(A20)基因、DGKA基因、DGKZ基因、FAS基因、EGR2基因、PSGL-1基因或KDM6A基因。

[0026] 因此,在本发明的示例性实施方式中,提供了一种或多种在核酸序列中进行了修饰的经人工操纵的免疫调节因子,所述免疫调节因子选自于由如下免疫调节因子所组成的组:PD-1基因、CTLA-4基因、TNFAIP3(A20)基因、DGKA基因、DGKZ基因、FAS基因、EGR2基因、PPP2R2D基因、TET2基因、PSGL-1基因以及KDM6A基因。

[0027] 核酸序列中的修饰可不受限制地为通过引导核酸-编辑蛋白复合体进行的人工操纵。

[0028] 术语“引导核酸-编辑蛋白复合体”是指经由引导核酸和编辑蛋白间的相互作用形成的复合体,该核酸-蛋白复合体包含引导核酸和编辑蛋白。

[0029] 引导核酸-编辑蛋白复合体可用于修饰对象物。所述对象物可为靶核酸、基因、染色体或蛋白。

[0030] 例如,所述基因可为通过引导核酸-编辑蛋白复合体进行人工操纵的免疫调节基因。

[0031] 其中,经人工操纵的所述免疫调节基因在组成所述免疫调节基因的核酸序列中的原型间隔区邻近基序(PAM)序列中或在临近所述PAM序列的5'端和/或3'端的连续的1bp-50bp碱基序列区内包含选自于由如下修饰所组成的组中的至少一种修饰:

[0032] 一个或多个核苷酸的删除或插入;

[0033] 利用不同于野生型基因的一个或多个核苷酸进行的置换;以及

[0034] 一个或多个外源核苷酸的插入;

[0035] 在组成所述免疫调节基因的核酸序列中的至少一个核苷酸的化学修饰。

[0036] 核酸修饰可发生于基因的启动子区。

[0037] 核酸修饰可发生于基因的外显子区。在一个示例性实施方式中,50%的修饰可发生于基因编码区的上游部分。

[0038] 核酸修饰可发生于基因的内含子区。

[0039] 核酸修饰可发生于基因的增强子区。

[0040] PAM序列可为例如如下序列中的一种或多种(以5'至3'方向来描述):

[0041] NGG(N为A、T、C或G);

[0042] NNNRYAC(N各自独立地为A、T、C或G;R为A或G;Y为C或T);

[0043] NNAGAAW(N各自独立地为A、T、C或G;W为A或T);

[0044] NNNNGATT(N各自独立地为A、T、C或G);

[0045] NNGRR(T)(N各自独立地为A、T、C或G;R为A或G;Y为C或T);以及

[0046] TTN(N为A、T、C或G)。

[0047] 此外,在另一实施方式中,本发明提供了引导核酸,所述引导核酸能够与选自于由如下基因所组成的组中的至少一种基因的核酸序列中的靶序列SEQ ID NO:1-SEQ ID NO:289分别形成互补结合:PD-1、CTLA-4、A20、DGK α 、DGK ζ 、FAS、EGR2、PPP2R2D、PSGL-1、KDM6A以及TET2。

[0048] 例如,本发明可以提供选自于下文所述组中的一种或多种引导核酸:

[0049] 能够与A20基因核酸序列中的靶序列SEQ ID NO:6和SEQ ID NO:11分别形成互补结合的引导核酸;

[0050] 能够与DGK α 基因核酸序列中的靶序列SEQ ID NO:19、SEQ ID NO:20、SEQ ID NO:21和SEQ ID NO:23分别形成互补结合的引导核酸;

[0051] 能够与EGR2基因核酸序列中的靶序列SEQ ID NO:25形成互补结合的引导核酸;

[0052] 能够与PPP2R2D基因核酸序列中的靶序列SEQ ID NO:64形成互补结合的引导核酸;

[0053] 能够与PD-1基因核酸序列中的靶序列SEQ ID NO:87和SEQ ID NO:89分别形成互补结合的引导核酸;

[0054] 能够与DGK ζ 基因核酸序列中的靶序列SEQ ID NO:109、SEQ ID NO:110、SEQ ID NO:111、SEQ ID NO:112和SEQ ID NO:113分别形成互补结合的引导核酸；

[0055] 能够与TET-2基因核酸序列中的靶序列SEQ ID NO:126、SEQ ID NO:128和SEQ ID NO:129分别形成互补结合的引导核酸；

[0056] 能够与PSGL-1基因核酸序列中的靶序列SEQ ID NO:182形成互补结合的引导核酸；

[0057] 能够与FAS基因核酸序列中的靶序列SEQ ID NO:252、SEQ ID NO:254、SEQ ID NO:257和SEQ ID NO:264分别形成互补结合的引导核酸；以及

[0058] 能够与KDM6A基因核酸序列中的靶序列SEQ ID NO:285形成互补结合的引导核酸。

[0059] 所述引导核酸可不受限制地为18bp-25bp、18bp-24bp、18bp-23bp、19bp-23bp或20bp-23bp的核苷酸。

[0060] 此外,本发明提供了一种经人工操纵的免疫细胞,所述经人工操纵的免疫细胞包含一种或多种经人工操纵的免疫调节基因及其表达产物。

[0061] 所述细胞不受限制地为免疫细胞和干细胞。免疫细胞为参与免疫应答的细胞,包括直接或间接参与免疫应答的所有细胞及它们的分化细胞。

[0062] 干细胞可为胚胎干细胞、成体干细胞、诱导多能干细胞(iPS细胞)或由诱导多能干细胞衍生而来的细胞(例如由受试者产生的经操纵而发生改变的(例如在其中诱导突变)并具有自我复制和分化能力的iPS细胞)。

[0063] 免疫细胞可为CD3阳性细胞。例如,免疫细胞可为T细胞或CAR-T细胞。

[0064] 免疫细胞可为CD56阳性细胞。例如,免疫细胞可为NK细胞,例如NK92、原代NK细胞。

[0065] 在实施方式中,免疫细胞可为CD3和CD56双阳性细胞(CD3/CD56双阳性细胞)。例如,免疫细胞可为自然杀伤T(NKT)细胞或细胞因子诱导的杀伤细胞(CIK)。

[0066] 具体而言,例如,所述细胞为选自于由如下细胞所组成的组中的至少一类:T细胞,例如CD8⁺T细胞(例如CD8⁺初始T细胞、CD8⁺效应T细胞、中央记忆T细胞或效应记忆T细胞)、CD4⁺T细胞、自然杀伤T细胞(NKT细胞)、调节T细胞、干细胞记忆T细胞;淋巴样祖细胞;造血干细胞;自然杀伤细胞(NK细胞);树突细胞;细胞因子诱导的杀伤细胞(CIK);外周血单核细胞(PBMC);单核细胞;巨噬细胞;自然杀伤T(NKT)细胞等。优选地,免疫细胞可为T细胞、CAR-T细胞、NK细胞或NKT细胞。

[0067] 可对免疫细胞进行人工操纵以使免疫调节基因的活性受抑制或失活。

[0068] 免疫细胞可进一步包含嵌合抗原受体(CAR)。

[0069] 作为实例,T细胞可进一步包含嵌合抗原受体(CAR)或经操纵的TCR(T细胞受体)。

[0070] 免疫细胞可进一步包含引导核酸-编辑蛋白复合体或者编码所述引导核酸-编辑蛋白复合体的核酸。

[0071] 编辑蛋白是选自于由如下蛋白所组成的组中的至少一种:酿脓链球菌(*Streptococcus pyogenes*)衍生而来的Cas9蛋白、空肠弯曲杆菌(*Campylobacter jejuni*)衍生而来的Cas9蛋白、嗜热链球菌(*Streptococcus thermophilus*)衍生而来的Cas9蛋白、金黄色葡萄球菌(*Streptococcus aureus*)衍生而来的Cas9蛋白、脑膜炎奈瑟菌(*Neisseria meningitidis*)衍生而来的Cas9蛋白以及Cpf1蛋白。作为实例,编辑蛋白可为酿脓链球菌衍生而来的Cas9蛋白或者空肠弯曲杆菌衍生而来的Cas9蛋白。

[0072] 引导核酸可与选自于由如下基因所组成的组中的一种或多种基因的部分核酸序列形成互补结合:PD-1、CTLA-4、A20、DGK α 、DGK ζ 、FAS、EGR2、PPP2R2D、PSGL-1、KDM6A和TET2。所述结合可产生0-5、0-4、0-3或0-2个错配。作为优选实例,引导核酸可为与表1中的靶序列SEQ ID NO:1-SEQ ID NO:289中的一条或多条形成互补结合的核苷酸。

[0073] 作为示例性实例,引导核酸可为分别与如下靶序列中的一条或多条形成互补结合的核苷酸:SEQ ID NO:6-SEQ ID NO:11(A20);SEQ ID NO:19、SEQ ID NO:20、SEQ ID NO:21和SEQ ID NO:23(DGK α);SEQ ID NO:25(EGR2);SEQ ID NO:64(PPP2R2D);SEQ ID NO:87和SEQ ID NO:89(PD-1);SEQ ID NO:109、SEQ ID NO:110、SEQ ID NO:111、SEQ ID NO:112和SEQ ID NO:113(DGK ζ);SEQ ID NO:126、SEQ ID NO:128和SEQ ID NO:129(TET-2);SEQ ID NO:182(PSGL-1);SEQ ID NO:252、SEQ ID NO:254、SEQ ID NO:257和SEQ ID NO:264(FAS);以及SEQ ID NO:285(KDM6A)。

[0074] 在本发明的示例性实施方式中,免疫细胞包含选自于在核酸序列中进行了修饰的DGK α 基因和DGK ζ 基因中的至少一种经人工操纵的基因。

[0075] 在本发明的另一示例性实施方式中,免疫细胞包含在核酸序列中进行了修饰的经人工操纵的DGK α 基因和DGK ζ 基因。

[0076] 在示例性实施方式中,本发明提供了用于引起期望的免疫应答的组合物。所述组合物可指药物组合物或治疗组合物。

[0077] 在示例性实施方式中,本发明提供了用于基因操纵的组合物,所述组合物包含引导核酸以及编辑蛋白或编码所述编辑蛋白的核酸,所述引导核酸能够与选自于如下基因中的一种或多种基因的核酸序列中的靶序列SEQ ID NO:1-SEQ ID NO:289分别形成互补结合:PD-1、CTLA-4、A20、DGK α 、DGK ζ 、FAS、EGR2、PPP2R2D、PSGL-1、KDM6A以及TET2。

[0078] 相关配置的描述与上文所述相同。

[0079] 在示例性实施方式中,本发明提供了用于对免疫细胞进行人工操纵的方法,所述方法包括将从人体分离的免疫细胞与选自于如下中的至少一种接触:

[0080] (a) 引导核酸,所述引导核酸能够与选自于由如下基因所组成的组中的至少一种基因的核酸序列中的靶序列SEQ ID NO:1-SEQ ID NO:289分别形成互补结合:PD-1、CTLA-4、A20、DGK α 、DGK ζ 、FAS、EGR2、PPP2R2D、PSGL-1、KDM6A以及TET2;以及

[0081] (b) 编辑蛋白,所述编辑蛋白是选自于由如下蛋白所组成的组中的至少一种:酿脓链球菌衍生而来的Cas9蛋白、空肠弯曲杆菌衍生而来的Cas9蛋白、嗜热链球菌衍生而来的Cas9蛋白、金黄色葡萄球菌衍生而来的Cas9蛋白、脑膜炎奈瑟菌衍生而来的Cas9蛋白以及Cpf1蛋白。

[0082] 引导核酸和编辑蛋白可以各自以核酸序列的形式存在于一个或多个载体中,或者可以通过引导核酸和编辑蛋白结合形成复合体的形式存在。

[0083] 接触步骤可以在体内或离体进行。

[0084] 接触步骤可以通过选自于如下方法中的一种或多种方法进行:电穿孔、脂质体、质粒、病毒载体、纳米粒子和蛋白易位结构域融合蛋白法。

[0085] 病毒载体可为选自于如下载体的组中的至少一种:逆转录病毒、慢病毒、腺病毒、腺相关病毒(AAV)、痘苗病毒(vaccinia virus)、痘病毒(poxvirus)和单纯疱疹病毒。

[0086] 在示例性实施方式中,本发明提供了用于提供受试者中的免疫细胞靶位置的序列

信息的方法,所述方法通过对选自于如下基因的组中的至少一种基因进行测序来进行:PD-1、CTLA-4、A20、DGK α 、DGK ζ 、FAS、EGR2、PPP2R2D、PSGL-1、KDM6A以及TET2。

[0087] 此外,本发明提供了用于使用由所述方法提供的信息构建文库的方法。

[0088] 在示例性实施方式中,本发明提供了基因操纵试剂盒,所述试剂盒包含如下成分:

[0089] (a) 引导核酸,所述引导核酸能够与选自于由如下基因所组成的组中的至少一种基因的核酸序列中的靶序列SEQ ID NO:1-SEQ ID NO:289分别形成互补结合:PD-1、CTLA-4、A20、DGK α 、DGK ζ 、FAS、EGR2、PPP2R2D、PSGL-1、KDM6A以及TET2;以及

[0090] (b) 编辑蛋白,所述编辑蛋白包括选自于由如下蛋白所组成的组中的一种或多种蛋白:酿脓链球菌衍生而来的Cas9蛋白、空肠弯曲杆菌衍生而来的Cas9蛋白、嗜热链球菌衍生而来的Cas9蛋白、金黄色葡萄球菌衍生而来的Cas9蛋白、脑膜炎奈瑟菌衍生而来的Cas9蛋白以及Cpf1蛋白。

[0091] 这些试剂盒可用于对期望的基因进行人工操纵。

[0092] 在一个示例性实施方式中,本发明提供了使用免疫治疗方式的疾病的治疗用途的全部方面,其包括向受试者给予经人工操纵的免疫细胞(例如经基因操纵的免疫细胞或干细胞)。其对过继性免疫疗法特别有用。

[0093] 治疗对象可为包括灵长类动物(例如人、猴等)、啮齿类动物(例如小鼠、大鼠等)在内的哺乳动物。

[0094] 有益效果

[0095] 可以通过免疫系统获得有效的免疫细胞疗法产品,通过经人工操纵的免疫调节因子和含有所述经人工操纵的免疫调节因子的细胞对所述免疫系统的功能进行人工操纵。

[0096] 例如,当通过本发明的组合物或方法对免疫调节因子进行人工控制时,可以改善涉及免疫细胞的存活、增殖、持久性、细胞毒性、细胞因子释放和/或浸润等方面的免疫效力。

附图说明

[0097] 图1为显示细胞中CD25的中位荧光强度(median fluorescence intensity,MFI)的图,其中使用针对DGK α 的sgRNA(#11;以DGK α #11表示)将DGK α 基因敲除。

[0098] 图2为显示细胞中CD25的MFI的图,其中使用针对A20的sgRNA(#11;以A20#11表示)将A20基因敲除。

[0099] 图3为显示细胞中CD25的MFI的图,其中使用针对EGR2的sgRNA(#1;以EGR2#1表示)将EGR2基因敲除。

[0100] 图4为显示细胞中CD25的MFI的图,其中使用针对PPP2R2D的sgRNA(#10;以PPP2R2D#10表示)将PPP2R2D基因敲除。

[0101] 图5为显示培养介质中IFN- γ 水平的图,其中分别使用针对DGK α 的sgRNA(#11;以DGK α #11表示)将细胞中的DGK α 基因敲除;使用针对A20的sgRNA(#11;以A20#11表示)将细胞中的A20基因敲除;使用针对EGR2的sgRNA(#1;以EGR2#1表示)将细胞中的EGR2基因敲除(IFN- γ 的单位水平:pg/mL)。

[0102] 图6为显示培养介质中IFN- γ 水平的图,其中分别使用针对DGK α 的sgRNA(#11;以DGK α #11表示)将细胞中的DGK α 基因敲除;使用针对DGK α 的sgRNA(组合使用#8和#11;以DGK

$\alpha\#8+11$ 表示)将细胞中的DGK α 基因敲除;使用针对DGK-zeta的sgRNA(#5;以DGKz#5表示)将细胞中的DGK-zeta基因敲除;使用针对A20的sgRNA(#11;以A20#11表示)将细胞中的A20基因敲除(IFN- γ 的单位水平:pg/mL)。

[0103] 图7为显示培养介质中IL-2水平的图,其中分别使用DGK α #11将细胞中的DGK α 基因敲除;使用DGK α #8+11将细胞中的DGK α 基因敲除;使用DGKz#5将细胞中的DGK-zeta基因敲除;使用A20#11将细胞中的A20基因敲除(IL-2的单位水平:pg/mL)。

[0104] 图8a显示了人原代T细胞中CRISPR/Cas9介导的DGK基因敲除结果,其中(A)确认了人原代T细胞中的基因敲除时间线(通过CD3/CD28珠进行的细胞活化,139 CAR的慢病毒递送以及使用电穿孔d进行的DGK基因敲除)并且(B)借助Mi-seq系统确认了DGK α 和DGK ζ 的插入缺失(indel)效率;图8b显示了表明脱靶分析结果的图。

[0105] 图9a显示了说明通过敲除DGK基因改善CAR-T细胞的增殖以及效应物(effector)的图,其中示出了(A)通过使用流式细胞术测量7-AAD阳性U87vIII细胞来评估139 CAR-T细胞杀伤活性的结果和(B)借助ELISA(IFN- γ 、IL-2试剂盒,Biolegend)的细胞因子分泌能力测定的结果;图9b显示了表明使用流式细胞术来评估139 CAR T细胞增殖能力的结果。

[0106] 图10显示了在敲除DGK后暴露于抗原时,139 CAR表达增强且CD3末端(CD3terminus)信号放大的结果,其中(A)显示了利用CD3/CD28珠刺激的139 CAR-T细胞的磷酸化ERK信号的蛋白质印迹结果;并且(B)显示了使用流式细胞术的139 CAR表达结果(左:依赖于存在抗原暴露的CAR表达;右:暴露于抗原3天后CAR表达的比较)。

[0107] 图11显示了表明DGK敲除不诱导tonic活化和T细胞耗竭的结果的图,其中(A)显示了借助ELISA评估的IFN- γ 的分泌能力;并且(B)显示了CAR阳性T细胞中的耗竭标志物(即,PD-1(左)和TIM-3(右))的分析结果。

[0108] 图12显示了表明DGK敲除的T细胞避免了TGF- β 和PGE2的免疫阻遏效应的图,其中(A)显示了对139 CAR-T细胞和139 DGK $\alpha\zeta$ CAR-T细胞依赖于TGF- β 的存在(10ng/mL)的杀伤活性、IFN- γ 分泌能力和IL-2分泌能力的评估;并且(B)显示了对139 CAR-T细胞和139 DGK $\alpha\zeta$ CAR-T细胞依赖于PGE2的存在(0.5 μ g/mL)的杀伤活性、IFN- γ 分泌能力和IL-2分泌能力的评估。

[0109] 图13显示了说明人NK细胞中CRISPR/Cas9介导的DGK α 敲除的效率及对效应物功能的影响的结果的图,其中(A)和(B)显示了说明利用Mi-seq系统对NK-92细胞和人原代NK细胞中的敲除效率进行分析的图;并且(C)显示了通过测定7-AAD阳性Raji细胞表明NK-92的杀伤活性的图。

[0110] 图14显示了人NKT细胞中CRISPR/Cas9介导的DGK α 和DGK ζ 敲除的效率的结果,其中(A)显示了插入缺失效率的评估结果;(B)显示了细胞生长的评估结果;(C)显示了细胞存活能力的评估结果;并且(D)显示了在蛋白水平对存在表达进行鉴别的蛋白质印迹实验结果。

[0111] 图15显示了说明人NKT细胞中DGK α 和DGK ζ 对效应物功能的影响的图,其中通过ELISA(IFN试剂盒,Biolegend)确认了DGK α 和DGK ζ 分别敲除和同时敲除对(A)杀伤活性和(B)IFN- γ 分泌能力的影响。

[0112] 图16显示了说明为了对在人NKT细胞中敲除DGK α 和DGK ζ 进行功能性评估,在从NKT细胞中敲除PA-1后的(A)插入缺失效率和(B)细胞毒性改善(即杀伤能力改善)的图。

[0113] 图17a-图17c显示了说明在Jurkat细胞中筛选hPSGL-1 sgRNA的分析结果的图,其

中图17a显示了插入缺失效率和敲除后不表达PSGL-1的Jurkat细胞的水平;并且图17b和图17c显示了敲除后在Jurkat细胞表面上表达的PSGL-1的表达水平。

[0114] 图18显示了说明人原代T细胞中hPSGL-1敲除(KO)实验的结果的图,其中(A)显示了插入缺失效率;(B)显示了敲除后不表达PSGL-1的T细胞的水平;并且(C)显示了敲除后在T细胞表面上表达的PSGL-1的表达水平。

具体实施方式

[0115] 除非另有定义,本文使用的全部技术术语和科学术语具有与本发明所属领域的普通技术人员通常理解的含义相同的含义。尽管与本文所述的方法和材料类似或相同的方法和材料可用于本发明的实践或测试中,适合的方法和材料在下文中描述。本文提及的所有出版物、专利申请、专利和其它参考文献都以引用的方式将它们整体并入。另外,材料、方法和实例仅为说明性的,而不旨在进行限制。

[0116] 本发明涉及具有改善的免疫效力的经人工操纵的免疫系统。更具体而言,本发明涉及人工修饰免疫系统,所述人工修饰免疫系统包含经人工操纵的免疫调节因子以及包含该经人工操纵的免疫调节因子的细胞。

[0117] [免疫调节因子]

[0118] 免疫调节因子

[0119] 术语“免疫调节因子”是其功能与免疫应答的形成和性能相关的物质,包括所有能够调节免疫应答的各种物质,可以为非天然(即经人工操纵的)的物质。例如,免疫调节因子可为在免疫细胞中表达的经基因操纵或修饰的基因或蛋白。

[0120] 术语“经人工操纵的”意味着其中施加了人工修饰的状态,而并非其以天然状态出现时的状态。

[0121] 术语“基因操纵”意味着对本发明提及的生物材料或非生物材料实施了人工施加基因修饰的操作的情况,例如,其可为出于特定目的对其中的基因组进行了人工修饰的基因和/或基因产物(例如多肽、蛋白等)。

[0122] 作为优选实例,本发明提供了用于特定目的的经基因操纵或修饰的免疫调节因子。

[0123] 下面列举的元件仅为免疫调节因子的实例,因此并不限制本发明所涵盖的免疫调节因子的类型。下面列出的基因或蛋白可能具有不止一种类型的免疫调节功能,而可具有多种类型的功能。此外,根据需要,可提供两种以上免疫调节因子。

[0124] [免疫细胞活性调节元件]

[0125] 术语“免疫细胞活性调节元件”是发挥对免疫应答的程度或活性进行调节的功能的元件,例如,其可为发挥对免疫应答的程度或活性进行调节的功能的经基因操纵或修饰的基因或蛋白。

[0126] 免疫细胞活性调节元件可执行与免疫细胞的活化或失活相关的功能。

[0127] 免疫细胞活性调节元件可发挥刺激或改善免疫应答的功能。

[0128] 免疫细胞活性调节元件可发挥抑制免疫应答的功能。

[0129] 免疫细胞活性调节元件可与细胞膜的通道蛋白和受体结合,从而执行与调节免疫应答的信号转导相关的功能以及与蛋白合成和分解相关的功能。

[0130] 例如,免疫细胞活性调节元件可为程序性细胞死亡蛋白(PD-1)。

[0131] PD-1基因(也称为PDCD1基因;下文中使用PD-1基因和PDCD1基因来表示相同的基因)是指编码PD-1蛋白(也称为分化簇279(CD279))的基因(全长DNA、cDNA或mRNA)。在实施方式中,PD-1基因可为选自于由如下基因所组成的组中的一种或多种,但不限于此:编码人PD-1的基因(例如NCBI登记号NP_005009.2等),例如以NCBI登记号NM_005018.2、NG_012110.1等表示的PD-1基因。

[0132] 免疫细胞活性调节元件可为细胞毒T淋巴细胞相关蛋白4(CTLA-4)。

[0133] CTLA-4基因是指编码CTLA-4蛋白(也称为分化簇152(CD152))的基因(全长DNA、cDNA或mRNA)。在实施方式中,CTLA-4基因可为选自于由如下基因所组成的组中的一种或多种,但不限于此:编码人CTLA-4的基因(例如NCBI登记号NP_001032720.1、NP_005205.2等),例如以NCBI登记号NM_001037631.2、NM_005214.4、NG_011502.1等表示的CTLA-4基因。

[0134] 免疫细胞活性调节元件可为CBLB。

[0135] 免疫细胞活性调节元件可为PSGL-1。

[0136] 免疫细胞活性调节元件可为ILT2。

[0137] 免疫细胞活性调节元件可为KIR2DL4。

[0138] 免疫细胞活性调节元件可为SHP-1。

[0139] 上述基因可来自于包括灵长类动物(例如人、猴等)、啮齿类动物(例如小鼠、大鼠等)在内的哺乳动物。

[0140] 在一个实施方式中,免疫细胞活性调节元件可发挥刺激免疫应答的功能。

[0141] 免疫细胞活性调节元件可为免疫细胞生长调节元件。

[0142] 术语“免疫细胞生长调节元件”是指通过对免疫细胞中的蛋白合成等进行调节而发挥调节免疫细胞生长的功能的元件,例如,在免疫细胞中表达的基因或蛋白。

[0143] 免疫细胞生长调节元件可在DNA转录、RNA翻译和细胞分化中发挥功能。

[0144] 免疫细胞生长调节元件的实例可为涉及NFAT、I κ B/NF- κ B、AP-1、4E-BP1、eIF4E以及S6的表达通路的基因或蛋白。

[0145] 例如,免疫细胞生长调节元件可为DGK α 。

[0146] DGKA(DGK- α ,DGK α)基因是指编码二酰甘油(diacylglycerol)激酶 α 蛋白(DGKA)的基因(全长DNA、cDNA或mRNA)。在实施方式中,DGKA基因可为选自于由如下基因所组成的组中的一种或多种,但不限于此:编码人DGKA的基因(例如NCBI登记号NP_001336.2、NP_958852.1、NP_958853.1、NP_963848.1等),例如以NCBI登记号NM_001345.4、NM_201444.2、NM_201445.1、NM_201554.1、NC_000012.12等表示的DGKA基因。

[0147] 免疫细胞生长调节元件可为DGK ζ 。

[0148] DGKZ(Dgk-zeta,DGK ζ)基因是指编码二酰甘油激酶 ζ 蛋白(DGKZ)的基因(全长DNA、cDNA或mRNA)。在实施方式中,DGKZ基因可为选自于由如下基因所组成的组中的一种或多种,但不限于此:编码人DGKZ的基因(例如NCBI登记号NP_001099010.1、NP_001186195.1、NP_001186196.1、NP_001186197.1、NP_003637.2、NP_963290.1、NP_963291.2等),例如以NCBI登记号NM_001105540.1、NM_001199266.1、NM_001199267.1、NM_001199268.1、NM_003646.3、NM_201532.2、NM_201533.3、NG_047092.1等表示的DGKZ基因。

[0149] 免疫细胞生长调节元件可为EGR2。

[0150] EGR2基因是指编码早期生长应答蛋白2(EGR2)的基因(全长DNA、cDNA或mRNA)。在实施方式中,EGR2基因可为选自于由如下基因所组成的组中的一种或多种,但不限于此。

[0151] 免疫细胞生长调节元件可为EGR3。

[0152] 免疫细胞生长调节元件可为PPP2R2D。

[0153] 免疫细胞生长调节元件可为A20(TNFAIP3)。

[0154] 免疫细胞生长调节元件可为PSGL-1。

[0155] 上述基因可来自于包括灵长类动物(例如人、猴等)、啮齿类动物(例如小鼠、大鼠等)在内的哺乳动物。

[0156] 免疫细胞活性调节元件可为免疫细胞死亡调节元件。

[0157] 术语“免疫细胞死亡调节元件”是指其功能涉及免疫细胞死亡的元件,其可为在免疫细胞中表达的执行此类功能的基因或蛋白。

[0158] 免疫细胞死亡调节元件可执行与免疫细胞凋亡或坏死相关的功能。

[0159] 在一个实施方式中,免疫细胞死亡调节元件可为胱天蛋白酶级联(caspase cascade)相关蛋白或基因。

[0160] 免疫细胞死亡调节元件可为Fas。下文当提及蛋白或基因时,对本领域普通技术人员而言显而易见的是,可对受体或结合区(所述蛋白或所述基因作用于所述受体或结合区)进行操纵。

[0161] 在另一实施方式中,免疫细胞死亡调节元件可为死亡结构域(death domain)相关蛋白或基因。具体而言,免疫细胞死亡调节元件可为Daxx。

[0162] 免疫细胞死亡调节元件可为Bcl-2家族蛋白。

[0163] 免疫细胞死亡调节元件可为BH3-only家族蛋白。

[0164] 免疫细胞死亡调节元件可为Bim。

[0165] 免疫细胞死亡调节元件可为Bid。

[0166] 免疫细胞死亡调节元件可为BAD。

[0167] 免疫细胞死亡调节元件可为位于免疫细胞外膜的配体或受体。

[0168] 具体而言,免疫细胞死亡调节元件可为PD-1。

[0169] 此外,免疫细胞死亡调节元件可为CTLA-4。

[0170] 免疫细胞活性调节元件可为免疫细胞耗竭调节元件。

[0171] 术语“免疫细胞耗竭调节元件”是执行与免疫细胞功能逐渐丧失相关的功能的元件,其可为在免疫细胞中表达的执行此类功能的基因或蛋白。

[0172] 免疫细胞耗竭调节元件可发挥辅助参与免疫细胞失活的基因的转录或翻译的功能。

[0173] 具体而言,辅助转录的功能可为将相应基因去甲基化的功能。

[0174] 此外,参与免疫细胞失活的基因包括免疫细胞活性调节元件的基因。

[0175] 具体而言,免疫细胞耗竭调节元件可为TET2。

[0176] 其可为编码人TET2的基因(例如NCBI登记号NP_001120680.1、NP_060098.3等),例如以NCBI登记号NM_001127208.2、NM_017628.4、NG_028191.1等表示的TET2基因。

[0177] 免疫细胞耗竭调节元件可发挥参与免疫细胞过度生长的功能。经历过度生长并且未再生的免疫细胞将丧失其功能。

[0178] 具体而言,免疫细胞耗竭调节元件可为Wnt。下文当提及蛋白或基因时,对本领域普通技术人员而言显而易见的是,可对所述蛋白或包含所述蛋白的信号转导通路中的基因以及所述基因作用的受体和结合区进行操纵。

[0179] 此外,免疫细胞耗竭调节元件可为Akt。下文当提及蛋白或基因时,对本领域普通技术人员而言显而易见的是,可对所述蛋白或包含所述蛋白的信号转导通路中的基因以及所述基因作用的受体和结合区进行操纵。

[0180] 免疫细胞活性调节元件可为细胞因子产生调节元件。

[0181] 术语“细胞因子产生调节元件”是参与免疫细胞的细胞因子分泌的基因或蛋白,其可在免疫细胞中表达的此类功能的基因或蛋白。

[0182] 细胞因子是由免疫细胞分泌的蛋白的统称,是在体内发挥重要作用的信号蛋白。细胞因子涉及感染、免疫力、炎症、创伤、溃烂、癌症等。细胞因子可由细胞分泌并随后影响其它细胞,或者影响分泌其自身的细胞。例如,细胞因子可诱导巨噬细胞增殖或促进分泌细胞自身的分化。然而,当细胞因子分泌过量时,可造成诸如攻击正常细胞等问题,因此在免疫应答中细胞因子的适当分泌也是重要的。

[0183] 细胞因子产生调节元件优选可为例如TTNF α 、IFN- γ 、TGF- β 、IL-2、IL-4、IL-10、IL-13、IL-1、IL-6、IL-12、IL-7、IL-15、IL-17和IFN- α 通路中的基因或蛋白。

[0184] 或者,细胞因子可发挥将信号递送至其它免疫细胞以诱导免疫细胞杀死所识别的抗原携带细胞或帮助分化的功能。具体而言,细胞因子产生调节元件优选可为涉及IL-2分泌的基因通路中的基因或蛋白。

[0185] 在实施方式中,免疫调节因子可以指在免疫细胞中表达的一组分子。这些分子可有效地运作以下调/上调或者抑制/促进免疫应答。

[0186] 例如,作为T细胞表达的一组分子,“免疫检查点”可以是但不限于直接抑制免疫细胞的程序性死亡1(PD-1、PDCD1或CD279,登记号NM_005018)、细胞毒T淋巴细胞抗原4(CTLA-4或CD152,GenBank登记号AF414120.1)、LAG3(CD223,登记号NM_002286.5)、Tim3(HAVCR2,GenBank登记号JX049979.1)、BTLA(CD272,登记号NM_181780.3)、BY55(CD160,GenBank登记号:CR541888.1)、TIGIT(IVSTM3,登记号NM_173799)、LAIR1(CD305,GenBank登记号CR542051.1)、SIGLEC10(GeneBank登记号AY358337.1)、2B4(CD244,登记号NM_001166664.1)、PPP2CA、PPP2CB、PTPN6、PTPN22、CD96、CRTAM、SIGLEC7、SIGLEC9、TNFRSF10B、TNFRSF10A、CASP8、CASP10、CASP3、CASP6、CASP7、FADD、FAS、TGFBRII、TGFRBRI、SMAD2、SMAD3、SMAD4、SMAD10、SKI、SKIL、TGIF1、IL10RA、IL10RB、HMOX2、IL6R、IL6ST、EIF2AK4、CSK、PAG1、SIT1、FOXP3、PRDM1、BATF、GUCY1A2、GUCY1A3、GUCY1B2以及GUCY1B3。

[0187] 在本发明的实施方式中,免疫调节因子例如可为经基因操纵或修饰的PD-1基因、CTLA-4基因、TNFAIP3(A20)基因、DGKA基因、DGKZ基因、FAS基因、EGR2基因、PPP2R2D基因、TET2基因、PSGL-1基因以及KDM6A基因。

[0188] 在本发明的实施方式中,免疫调节因子可包含两种以上经基因操纵或修饰的基因。例如,可对选自于由如下基因所组成的组中的两种以上基因进行操纵或修饰:PD-1基因、CTLA-4基因、TNFAIP3(A20)基因、DGKA基因、DGKZ基因、FAS基因、EGR2基因、PPP2R2D基因、TET2基因、PSGL-1基因以及KDM6A基因。

[0189] 在本发明中,这些基因的优选实例可包括经基因操纵或修饰的TNFAIP3、DGKA、

DGKZ、FAS、EGR2、PSGL-1以及KDM6A基因。

[0190] 可以通过在野生型基因的基因组序列的部分或全部区域中引入人工插入、缺失、置换和倒位(inversion)突变来获得基因操纵或修饰。此外,也可以通过对两种以上基因进行融合基因操纵或修饰来获得基因操纵或修饰。

[0191] 例如,可通过此类基因操纵或修饰使得这些基因失活,其结果是阻止这些基因编码的蛋白表达为具有其原始功能的蛋白形式。

[0192] 例如,可通过此类基因操纵或修饰使这些基因进一步活化,从而这些基因编码的蛋白表达为具有比其原始功能更好的功能的蛋白形式。在一个实例中,当由特定基因编码的蛋白的功能为A时,经操纵的基因表达的蛋白的功能可完全不同于A,或者其可具有包含A在内的附加功能(A+B)。

[0193] 例如,基因操纵或修饰可为通过利用功能彼此不同或彼此互补的两种以上基因来表达融合形式的两种以上蛋白。

[0194] 例如,基因操纵或修饰可为通过利用功能彼此不同或彼此互补的两种以上基因来表达各自为独立形式的两种以上蛋白。

[0195] 可从已知数据库(例如美国国立生物技术信息中心(NCBI)的GenBank)获得遗传信息。

[0196] 在实施方式中,可通过如下中的一种或多种诱导进行基因操纵或修饰:

[0197] 删除待修饰的基因(下文中称为“靶基因”)的全部或部分,例如删除靶基因的1bp或更多个核苷酸(例如1-30个核苷酸、1-27个核苷酸、1-25个核苷酸、1-23个核苷酸、1-20个核苷酸、1-15个核苷酸、1-10个核苷酸、1-5个核苷酸、1-3个核苷酸或者1个核苷酸);以及

[0198] 置换靶基因的1bp或更多个核苷酸(例如不同于原始(野生型)核苷酸的1-30个核苷酸、1-27个核苷酸、1-25个核苷酸、1-23个核苷酸、1-20个核苷酸、1-15个核苷酸、1-10个核苷酸、1-5个核苷酸、1-3个核苷酸或者1个核苷酸);以及在靶基因的任意位置插入一个或多个核苷酸(例如1-30个核苷酸、1-27个核苷酸、1-25个核苷酸、1-23个核苷酸、1-20个核苷酸、1-15个核苷酸、1-10个核苷酸、1-5个核苷酸、1-3个核苷酸或者1个核苷酸)。

[0199] 靶基因待被修饰的部分(“靶区域”)可为基因中1bp以上、3bp以上、5bp以上、7bp以上、10bp以上、12bp以上、15bp以上、17bp以上、20bp以上(例如1bp-30bp、3bp-30bp、5bp-30bp、7bp-30bp、10bp-30bp、12bp-30bp、15bp-30bp、17bp-30bp、20bp-30bp、1bp-27bp、3bp-27bp、5bp-27bp、7bp-27bp、10bp-27bp、12bp-27bp、15bp-27bp、17bp-27bp、20bp-27bp、1bp-25bp、3bp-25bp、5bp-25bp、7bp-25bp、10bp-25bp、12bp-25bp、15bp-25bp、17bp-25bp、20bp-25bp、1bp-23bp、3bp-23bp、5bp-23bp、7bp-23bp、10bp-23bp、12bp-23bp、15bp-23bp、17bp-23bp、20bp-23bp、1bp-20bp、3bp-20bp、5bp-20bp、7bp-20bp、10bp-20bp、12bp-20bp、15bp-20bp、17bp-20bp、21bp-25bp、18bp-22bp或者21bp-23bp)的连续核苷酸序列区域。

[0200] [包含免疫调节因子的细胞]

[0201] 本发明的一方面涉及包含经人工操纵的免疫调节因子的细胞。

[0202] 这些细胞为免疫细胞和干细胞,但不限于此。

[0203] 本发明的“免疫细胞”是参与免疫应答的细胞,其包括直接或间接参与免疫应答的全部细胞以及这些细胞的分化前(pre-differentiation)细胞。

[0204] 免疫细胞可具有细胞因子分泌、分化为其它免疫细胞和细胞毒性的功能。免疫细

胞还包括从其自然状态下经历突变的细胞。

[0205] 免疫细胞由骨髓中的造血干细胞分化而来,其主要包括淋巴样祖细胞和髓样祖细胞;还包括如下全部细胞:由淋巴样祖细胞分化而来并负责获得性免疫的T细胞和B细胞;以及由髓样祖细胞分化而来的巨噬细胞、嗜酸性粒细胞、中性粒细胞、嗜碱性粒细胞、巨核细胞、红细胞等。

[0206] 具体而言,所述细胞可为选自于由如下细胞所组成的组中的至少一种:T细胞,例如CD8⁺T细胞(例如CD8⁺初始T细胞、CD8⁺效应T细胞、中央记忆T细胞或效应记忆T细胞)、CD4⁺T细胞、自然杀伤T细胞(NKT细胞)、调节T细胞(Treg)、干细胞记忆T细胞;淋巴样祖细胞;造血干细胞;自然杀伤细胞(NK细胞);树突细胞;细胞因子诱导的杀伤细胞(CIK);外周血单核细胞(PBMC);单核细胞;巨噬细胞;自然杀伤T(NKT)细胞等。当细胞表面的主要组织相容性复合物(MHC)受体与T细胞表面上的TCR相互作用时,巨噬细胞和树突细胞可称为抗原呈递细胞(APC),其为能活化T细胞的特化细胞。或者,可通过引入表达抗原的核酸分子(TCR或其它抗原结合蛋白(例如CAR)识别该抗原)将任何造血干细胞或免疫系统细胞转化为APC。

[0207] 在实施方式中,免疫细胞可为通过对合成与MHC识别和/或免疫功能相关的蛋白(例如免疫检查点蛋白)的基因进行失活或交换而用作免疫疗法的细胞。

[0208] 在实施方式中,免疫细胞可进一步包含编码用于特异性细胞识别的短链和多亚基受体(例如CAR、TCR等)的多核苷酸。

[0209] 在实施方式中,本发明的免疫细胞可为来自健康供体或患者的血(例如外周血)、干细胞(例如胚胎干细胞、诱导多能干细胞等)、脐带血、骨髓等的免疫细胞,或者可为离体操纵的免疫细胞。

[0210] 在实施方式中,免疫细胞可为CD3阳性细胞,例如T细胞或CAR-T细胞。CD3为与TCR和多种蛋白以复合物形式存在于T细胞表面的受体。称为 γ 链、 δ 链、 ϵ 链、 ζ 链和 η 链的五类蛋白构成CD3,这五类蛋白与TCR一起以 $\alpha\beta:\gamma\delta\epsilon\zeta\eta$ 或 $\alpha\beta:\gamma\delta\epsilon\zeta\eta$ 的状态作为TCR/CD3复合物存在。已知其具有在T细胞的抗原识别期间将信号转导至细胞中的功能。

[0211] 在实施方式中,免疫细胞可为CD56阳性细胞,例如NK细胞(例如NK92细胞和原代NK细胞)。

[0212] NK细胞在免疫细胞中数量排名第三,约10%的外周血免疫细胞为NK细胞。NK细胞具有CD56和CD16,在肝或骨髓中成熟。NK细胞攻击病毒感染的细胞或肿瘤细胞。当NK细胞识别出异常细胞时,其在细胞膜上喷洒穿孔素(perforin)使得细胞膜溶解而穿孔,在细胞膜内喷洒颗粒酶(granzyme)使得细胞质分解而造成凋亡,并将水和盐水注入细胞而造成坏死。NK细胞具有杀死多种癌细胞的能力。特别地,众所周知的是,NK细胞是不容易将外源遗传物质引入其中的细胞。

[0213] 在实施方式中,NK细胞可为双阳性细胞,例如自然杀伤T(NKT)细胞或细胞因子诱导的杀伤(CIK)细胞。

[0214] 自然杀伤T(NKT)细胞或细胞因子诱导的杀伤(CIK)细胞为同时表达CD3(即T细胞标志物)和CD56(即自然杀伤细胞(NK细胞)标志物)分子的免疫细胞。由于NKT细胞或CIK细胞衍生自T细胞且都具有NK细胞的特征和功能,其与主要组织相容性复合物(MHC)无关联地杀死肿瘤细胞。具体而言,NKT细胞是表达T细胞受体(TCR)和NK细胞特异性表面标志物NK1.1或NKR-PIA(CD161)的细胞。

[0215] 在一个实例中,NKT细胞可识别由CD1d(结构类似于I类MHC的单型(monomorphic)蛋白)呈递的糖脂。当被配体(例如 α -GalCer)活化时,NKT细胞分泌多种多样的细胞因子(例如IL-4、IL-13、IL-10和IFN- γ)。此外,NKT细胞具有抗肿瘤活性。

[0216] 在另一实例中,CIK细胞为这样的一类免疫细胞:当利用白介素2和CD3抗体处理收集的血液并离体培养2-3周时,CIK细胞增殖,并且CIK细胞为CD3和CD56阳性细胞。CIK细胞产生大量IFN- γ 和TNF- α 。

[0217] 在实施方式中,细胞可为具有自我复制和分化能力的胚胎干细胞、成体干细胞、诱导多能干细胞(iPS)细胞或者从诱导多能干细胞衍生而来的细胞(例如iPS细胞衍生细胞)。

[0218] 作为本发明的优选实施方式,细胞可包含作为免疫调节因子的经操纵或修饰的基因。

[0219] 细胞可包含经操纵或修饰的基因的全部或一部分;或者其表达产物。

[0220] 例如,细胞可为这样的细胞:其中,通过基因操纵或修饰使相应基因失活,从而该基因编码的蛋白不表达为具有其原始功能的蛋白形式。

[0221] 例如,细胞可为这样的细胞:其中,通过此类基因操纵或修饰使相应基因进一步活化,从而该基因编码的蛋白表达为具有比其原始功能更好的功能的蛋白形式。

[0222] 例如,细胞可为这样的细胞:其中,通过此类基因操纵或修饰使该基因编码的蛋白表达为表现出其原始功能和/或附加功能的蛋白形式。

[0223] 例如,细胞可为这样的细胞:其中,通过此类基因操纵或修饰利用功能彼此不同或彼此互补的两种以上基因使得两种以上蛋白表达为修饰形式。

[0224] 例如,细胞可为通过此类基因操纵或修饰而对于三种细胞因子(例如IL-2、TNF- α 和IFN- γ)而言具有高细胞因子产生或分泌能力的免疫细胞。

[0225] 在一个实例中,本发明的细胞可进一步包含如下成分。

[0226] -受体

[0227] 本发明的细胞可包含“免疫受体”。

[0228] 术语“免疫受体”可为存在于经人工操纵或修饰的免疫细胞表面的受体,是指参与免疫应答的物质,例如识别抗原和执行特定功能的功能性实体。

[0229] 受体可处于野生型或经人工操纵的状态。

[0230] 受体可具有抗原亲和性。

[0231] 受体可具有对由MHC结构型蛋白形成的结构以及在结构型蛋白中暴露的抗原的识别能力。

[0232] 受体可产生免疫应答信号。

[0233] 术语“免疫应答信号”是指免疫应答过程中出现的任何信号。

[0234] 免疫应答信号可为与免疫细胞生长和分化相关的信号。

[0235] 免疫应答信号可为与免疫细胞死亡相关的信号。

[0236] 免疫应答信号可为与免疫细胞活性相关的信号。

[0237] 免疫应答信号可为与免疫细胞辅助(aid)相关的信号。

[0238] 免疫应答信号可为对感兴趣的基因的表达进行调节的信号。

[0239] 免疫应答信号可为促进或抑制细胞因子合成的信号。

[0240] 免疫应答信号可为促进或抑制细胞因子分泌的信号。

- [0241] 免疫应答信号可为辅助其它免疫细胞生长或分化的信号。
- [0242] 免疫应答信号可为调控其它免疫细胞活性的信号。
- [0243] 免疫应答信号可为将其它免疫细胞吸引至信号出现的位置的信号。
- [0244] 在实施方式中,受体可为T细胞受体(TCR)。
- [0245] 在实施方式中,细胞可为经修饰的细胞,从而该细胞可包含特定T细胞受体(TCR)基因(例如TRAC或TRBC基因)。在另一实施方式中,TCR可为对肿瘤相关抗原(例如由T细胞识别的黑素瘤抗原1(MART1)、黑素瘤相关抗原3(MAGEA3)、NY-ES01、NYES01、癌胚抗原(CEA)、GP100等)具有结合特异性的受体。
- [0246] 在实施方式中,受体可为Toll样受体(TLR)。
- [0247] 受体可为CD4和CD8,其为参与MHC限制性T细胞活化的共受体。
- [0248] 受体可为CTLA-4(CD152)。
- [0249] 受体可为CD28。
- [0250] 受体可为CD137,4-1BB,其为放大(amplify)T细胞应答的受体。
- [0251] 受体可为CD3 ζ ,其为T细胞抗原受体的信号转导元件。
- [0252] 受体可为嵌合抗原受体(CAR)。
- [0253] 在本发明的实施方式中,受体可为经人工操纵的人工受体。
- [0254] 术语“人工受体”是指人工制备的不是野生型受体的功能性实体,其具有抗原识别和执行特定功能的特定能力。
- [0255] 此类人工受体可以针对特定抗原产生具有改善的或增强的识别的免疫应答信号,因此可有助于改善免疫应答。
- [0256] 作为一个实例,所述人工受体可具有如下成分。
- [0257] (i) 抗原识别部分
- [0258] 人工受体可包含抗原识别部分。
- [0259] 术语“抗原识别部分”是人工受体的一部分,是指识别抗原的区域。
- [0260] 抗原识别部分可为相比野生型受体而言对特定抗原具有改善的识别的抗原识别部分。具体而言,特定抗原可为癌细胞的抗原。此外,特定抗原可为正常体细胞的抗原。
- [0261] 抗原识别部分可具有抗原结合亲和性。
- [0262] 在结合至抗原时,抗原识别部分可产生信号。信号可为电信号。信号可为化学信号。
- [0263] 抗原识别部分可包含信号序列。
- [0264] 信号序列是指在蛋白合成过程中允许将蛋白递送至特定位点的肽序列。
- [0265] 信号序列可位于靠近抗原识别部分的N末端处。具体而言,其与N末端的距离可为约100个氨基酸。信号序列可位于靠近抗原识别部分的C末端处。具体而言,其与C末端的距离可为约100个氨基酸。
- [0266] 抗原识别部分可与第一信号发生部分具有有机功能性关系。
- [0267] 抗原识别部分可与抗体的抗原结合片段(Fab)结构域具有同源性。
- [0268] 抗原识别部分可为单链可变片段(scFv)。
- [0269] 抗原识别部分可通过其自身识别抗原,或者通过形成抗原识别结构来识别抗原。
- [0270] 抗原识别结构可通过建立特定结构来识别抗原,本领域普通技术人员可容易地理

解构成该特定结构的单体单元(monomeric units)和单体单元的结合。此外,抗原识别结构可由一个或两个以上的单体单元组成。

[0271] 抗原识别结构可为其中的单体单元串联连接的结构,或者可为其中的单体单元并联连接的结构。

[0272] 串联连接的结构是指其中的两个以上单体单元在一个方向上连续连接的结构,而并联连接的结构是指其中的两个以上单体单元各自同时例如在不同方向上连接在一个单体单元的末端处的结构。

[0273] 例如,单体单元可为无机物。

[0274] 单体单元可为生化配体。

[0275] 单体单元可与野生型受体的抗原识别部分具有同源性。

[0276] 单体单元可与抗原蛋白具有同源性。

[0277] 单体单元可为免疫球蛋白重链或可与其具有同源性。

[0278] 单体单元可为免疫球蛋白轻链或可与其具有同源性。

[0279] 单体单元可包含信号序列。

[0280] 同时,单体单元可通过化学键连接,或者可通过特定组合部分连接。

[0281] 术语“抗原识别单元组合部分(antigen recognition unit combining part)”是抗原识别单元彼此连接的区域,当存在由两个以上抗原识别单元组成的抗原识别结构时,抗原识别单元组合部分可以以任选组成存在。

[0282] 抗原识别单元组合部分可为肽。具体而言,所述组合部分可具有高丝氨酸和苏氨酸比例。

[0283] 抗原识别单元组合部分可为化学结合。

[0284] 抗原识别单元组合部分可通过具有特定长度来辅助抗原识别单元三维结构的表达。

[0285] 抗原识别单元组合部分可通过在抗原识别单元之间具有特定位置关系来辅助抗原识别结构的功能。

[0286] (ii)受体主体(receptor body)

[0287] 人工受体可包含受体主体。

[0288] 术语“受体主体”是介导抗原识别部分和信号发生部分之间的连接的区域,抗原识别部分和信号发生部分可为物理连接的。

[0289] 受体主体的功能可为对抗原识别部分或信号发生部分中产生的信号进行递送。

[0290] 根据情况,受体主体的结构可同时具有信号发生部分的功能。

[0291] 受体主体的功能可为允许人工受体固定在免疫细胞上。

[0292] 受体主体可包含氨基酸螺旋结构。

[0293] 受体主体的结构可包含与体内存在的正常受体蛋白的一部分具有同源性的部分。同源性可在50%-100%的范围内。

[0294] 受体主体的结构可包含与免疫细胞上的蛋白具有同源性的部分。同源性可在50%-100%的范围内。

[0295] 例如,受体主体可为CD8跨膜结构域。

[0296] 受体主体可为CD28跨膜结构域。具体而言,当第二信号发生部分是CD28时,CD28可

执行第二信号发生部分和受体主体的功能。

[0297] (iii) 信号发生部分

[0298] 人工受体可包含信号发生部分。

[0299] 术语“第一信号发生部分”是人工受体的一部分,是指产生免疫应答信号的部分。

[0300] 术语“第二信号发生部分”是人工受体的一部分,是指通过与第一信号发生部分相互作用产生免疫应答信号或者独立产生免疫应答信号的部分。

[0301] 人工受体可包含第一信号发生部分和/或第二信号发生部分。

[0302] 人工受体可分别包含两个以上的第一信号发生部分和/或第二信号发生部分。

[0303] 第一信号发生部分和/或第二信号发生部分可包含特定序列基序。

[0304] 所述序列基序可与指定簇(cluster of designation, CD) 蛋白的基序具有同源性。

[0305] 具体而言, CD蛋白可为CD3、CD247和CD79。

[0306] 所述序列基序可为氨基酸序列YxxL/I。

[0307] 所述序列基序在第一信号发生部分和/或第二信号发生部分内可为多个。

[0308] 具体而言, 第一序列基序可位于距离第一信号发生部分的起始位置1-200个氨基酸处。第二序列基序可位于距离第二信号发生部分的起始位置1-200个氨基酸处。

[0309] 此外, 各序列基序之间的距离可为1-15个氨基酸。

[0310] 具体而言, 各序列基序之间的优选距离为6-8个氨基酸。

[0311] 例如, 第一信号发生部分和/或第二信号发生部分可为CD3 ζ 。

[0312] 第一信号发生部分和/或第二信号发生部分可为Fc ϵ RI γ 。

[0313] 第一信号发生部分和/或第二信号发生部分可为仅当满足特定条件时才产生免疫应答的信号发生部分。

[0314] 特定条件可为抗原识别部分识别抗原。

[0315] 特定条件可为抗原识别部分与抗原形成结合。

[0316] 特定条件可为当抗原识别部分与抗原形成结合时产生的信号被递送。

[0317] 特定条件可为抗原识别部分识别抗原或者在与抗原结合的情况下抗原识别部分与抗原分离。

[0318] 免疫应答信号可为与免疫细胞的生长和分化相关的信号。

[0319] 免疫应答信号可为与免疫细胞的死亡相关的信号。

[0320] 免疫应答信号可为与免疫细胞的活性相关的信号。

[0321] 免疫应答信号可为与免疫细胞的辅助相关的信号。

[0322] 免疫应答信号可由抗原识别部分产生的信号特异性活化。

[0323] 免疫应答信号可为调节感兴趣的基因的表达的信号。

[0324] 免疫应答信号可为阻遏免疫应答的信号。

[0325] 在实施方式中, 信号发生部分可包含额外信号发生部分。

[0326] 术语“额外信号发生部分”是人工受体的一部分, 是指相对于由第一信号发生部分和/或第二信号发生部分产生的免疫应答信号而言产生额外免疫应答信号的区域。

[0327] 下文中, 将额外信号发生部分按顺序称为第n信号发生部分($n \neq 1$)。

[0328] 除第一信号发生部分外, 人工受体还可包含额外信号发生部分。

- [0329] 人工受体可包含两个以上额外信号发生部分。
- [0330] 额外信号发生部分可为其中可产生4-1BB、CD27、CD28、ICOS和OX40的免疫应答信号或其它信号的结构。
- [0331] 额外信号发生部分产生免疫应答信号的条件及其所产生的免疫应答信号的特征包括与第一信号发生部分和/或第二信号发生部分的免疫应答信号相对应的详细情况。
- [0332] 免疫应答信号可为促进细胞因子合成的信号。免疫应答信号可为促进或抑制细胞因子分泌的信号。具体而言,细胞因子优选可为IL-2、TNF α 或IFN- γ 。
- [0333] 免疫应答信号可为辅助其它免疫细胞生长或分化的信号。
- [0334] 免疫应答信号可为调控其它免疫细胞活性的信号。
- [0335] 免疫应答信号可为将其它免疫细胞吸引至信号出现的位置的信号。
- [0336] 本发明包括人工受体全部可能的结合关系。因此,本发明的人工受体的方面不限于本文所述的这些。
- [0337] 人工受体可由抗原识别部分-受体主体-第一信号发生部分组成。受体主体可为任选包含的。
- [0338] 人工受体可由抗原识别部分-受体主体-第二信号发生部分-第一信号发生部分组成。受体主体可为任选包含的。具体而言,可对第一信号发生部分和第二信号发生部分的位置进行改变。
- [0339] 人工受体可由抗原识别部分-受体主体-第二信号发生部分-第三信号发生部分-第一信号发生部分组成。受体主体可为任选包含的。具体而言,可对第一信号发生部分至第三信号发生部分的位置进行改变。
- [0340] 在人工受体中,信号发生部分的数量不限于1-3个,可包含多于3个。
- [0341] 除上述实施方式外,人工受体还可具有抗原识别部分-信号发生部分-受体主体的结构。当需要产生在具有该人工受体的细胞外发挥作用的免疫应答信号时,该结构可以是有利的。
- [0342] 人工受体可以通过相当于野生型受体的方式发挥功能。
- [0343] 人工受体可通过与特定抗原形成结合而发挥形成特定位置关系的功能。
- [0344] 人工受体可发挥识别抗原并产生免疫应答信号(该免疫应答信号促进针对该特定抗原的免疫应答)的功能。
- [0345] 人工受体可发挥识别体内一般细胞的抗原并在体内抑制针对该细胞的免疫应答的功能。
- [0346] (iv) 信号序列
- [0347] 在实施方式中,人工受体可任选包含信号序列。
- [0348] 当人工受体包含特定蛋白的信号序列时,这可帮助人工受体容易地定位在免疫细胞的膜上。优选地,当人工受体包含跨膜蛋白的信号序列时,这可帮助人工受体穿过免疫细胞膜而定位在免疫细胞的外膜上。
- [0349] 人工受体可包含一个或多个信号序列。
- [0350] 信号序列可包含多个带正电荷的氨基酸。
- [0351] 信号序列可在靠近N末端或C末端的位置处包含带正电荷的氨基酸。
- [0352] 信号序列可为跨膜蛋白的信号序列。

- [0353] 信号序列可为位于免疫细胞的外膜上的蛋白的信号序列。
- [0354] 可将信号序列包含于人工受体所具有的结构中,即,抗原识别部分、受体主体、第一信号发生部分以及额外信号发生部分。
- [0355] 具体而言,信号序列可位于靠近各结构的N末端或C末端的位置处。
- [0356] 具体而言,信号序列距N末端或C末端的距离可为约100个氨基酸。
- [0357] 在实施方式中,细胞可为经修饰的细胞,从而该细胞可包含特定T细胞受体 (TCR) 基因。
- [0358] 在另一实施方式中,TCR可为对肿瘤相关抗原 (例如由T细胞识别的黑素瘤抗原1 (MART1)、黑素瘤相关抗原3 (MAGEA3)、NY-ESO1、癌胚抗原 (CEA)、NY-ES-01 (GP100等)、黑素瘤) 具有结合特异性的受体。
- [0359] 在又一实施方式中,细胞可为经修饰的细胞,从而该细胞可包含特定嵌合抗原受体 (CAR)。在实施方式中,CAR可为对肿瘤相关抗原 (例如CD19、CD20、碳酸酐酶 IX (CAIX)、CD171、CEA、ERBB2、GD2、 α 叶酸受体、Lewis Y抗原、前列腺特异性膜抗原 (PSMA) 或肿瘤相关糖蛋白72 (TAG72)) 具有结合特异性的受体。
- [0360] 在又一实施方式中,细胞可为经修饰的细胞,从而该细胞例如可通过TCR或CAR与如下肿瘤抗原中的一种或多种结合。
- [0361] 肿瘤抗原可包括但不限于:AD034、AKT1、BRAP、CAGE、CDX2、CLP、CT-7、CT8/HOM-TES-85、cTAGE-1、EGFR、EGFRvIII、腓骨蛋白-1 (Fibulin-1)、HAGE、HCA587/MAGE-C2、hCAP-G、HCE661、HER2/neu、HLA-Cw、HOM-HD-21/半乳糖凝集素9 (Galectin9)、HOM-MEEL-40/SSX2、HOM-RCC-3.1.3/CAXII、HOXA7、HOXB6、Hu、HUB 1、KM-HN-3、KM-KN-1、KOC1、KOC2、KOC3、KOC3、LAGE-1、MAGE-1、MAGE-4a、MPP1 1、MSLN、NNP-1、NY-BR-1、NY-BR-62、NY-BR-85、NY-CO-37、NY-CO-38、NY-ESO-1、NY-ESO-5、NY-LU-12、NY-REN-10、NY-REN-19/LKB/STK11、NY-REN-21、NY-REN-26/BCR、NY-REN-3/NY-CO-38、NY-REN-33/SNC6、NY-REN-43、NY-REN-65、NY-REN-9、NY-SAR-35、OGFr、PSMA、PSCA、PLU-1、Rab38、RBPJkappa、RHAMM、SCP1、SCP-1、SSX3、SSX4、SSX5、TOP2A、TOP2B以及酪氨酸酶。
- [0362] -抗原结合调节元件
- [0363] 本发明的细胞可进一步包含“抗原结合调节元件”。
- [0364] “抗原结合调节元件”是使得受体和抗原之间能够结合的元件,可为执行此类功能的基因或蛋白。
- [0365] 可使用此类抗原结合调节元件来调节免疫应答。例如,在通过将经过外部操作的细胞添加至活体中来执行处理的情况下,当HVGD (宿主HVGD (移植病;移植物抗宿主病; HostGraft)) (其中针对该经过外部操作的细胞的免疫应答被活化并因此造成疗效消失) 成为问题时,可通过阻遏免疫细胞受体的抗原结合能力来解决该问题。
- [0366] 抗原结合调节元件可为与受体结构相关联的蛋白或基因。
- [0367] 抗原结合调节元件可为与受体结构具有同源性的蛋白或基因。
- [0368] 例如,抗原结合调节元件可为dCK。
- [0369] 抗原结合调节元件可为CD52。
- [0370] 抗原结合调节元件可为B2M。
- [0371] 抗原结合调节元件可为与受体识别的结构相关联的蛋白或基因。

[0372] 例如,抗原结合调节元件可为MHC蛋白。

[0373] 在实施方式中,本发明涉及包含经人工操纵的免疫调节基因或由这些基因表达的蛋白的免疫细胞。

[0374] 在另一实施方式中,本发明涉及包含经人工操纵的免疫调节基因或由这些基因表达的蛋白;以及受体的免疫细胞。

[0375] 在又一实施方式中,本发明涉及包含经人工操纵的免疫调节基因或由这些基因表达的蛋白;受体;以及抗原结合调节元件的免疫细胞。

[0376] 本发明的细胞的代表性实例为免疫细胞。

[0377] 在本发明的一些示例性实施方式中,免疫细胞可为选自于由如下细胞所组成的组中的至少一种细胞:外周血单核细胞(PBMC)、自然杀伤细胞(NK细胞)、单核细胞、T细胞、CAR-T细胞、巨噬细胞、自然杀伤T细胞(NKT细胞)等,优选为T细胞、CAR-T细胞、自然杀伤细胞(NK细胞)或自然杀伤T细胞(NKT细胞)。

[0378] 限制经基因操纵的免疫细胞(例如T细胞、NK细胞和NKT细胞)的效力的因素包括:

[0379] (1) 免疫细胞增殖(例如在过继转移后免疫细胞的繁殖受限);

[0380] (2) 免疫细胞存活(例如肿瘤环境中的因子诱导免疫细胞凋亡);以及

[0381] (3) 免疫细胞功能(例如宿主免疫细胞和癌细胞分泌的抑制因子对细胞毒免疫细胞的功能造成抑制)。

[0382] 出于该目的,通过免疫细胞来调节上述限制因素,所述免疫细胞中表达的一种或多种基因(例如选自于由如下基因所组成的组中的一种或多种基因:PD-1、CTLA-4、TNFAIP3、DGKA、DGKZ、FAS、EGR2、PPP2R2D、PSGL-1、KDM6A和TET2)被失活。

[0383] 在一个实例中,可靶向和操纵免疫细胞中表达的一种或多种基因(例如选自于由如下基因所组成的组中的一种或多种基因:PD-1、CTLA-4、TNFAIP3、DGKA、DGKZ、FAS、EGR2、PPP2R2D、PSGL-1、KDM6A和/或TET2基因),以使其各自独立地被敲除、敲减或敲入,从而影响一个或多个免疫细胞的功能、增殖和存活。

[0384] 在一个实例中,可靶向和操纵免疫细胞中的两种以上基因(选自于由如下基因所组成的组:PD-1、CTLA-4、TNFAIP3、DGKA、DGKZ、FAS、EGR2、PPP2R2D、PSGL-1、KDM6A和/或TET2基因),以使其同时被敲除、敲减或敲入。在实施方式中,同时敲除DGKA和DGKZ。

[0385] 在一个实例中,可通过靶向非编码区或编码区(例如启动子区、增强子、3'UTR和/或多聚腺苷酸信号序列或转录序列(例如内含子或外显子序列)),来靶向和操纵免疫细胞中表达的一种或多种基因(例如选自于由如下基因所组成的组中的一种或多种基因:PD-1、CTLA-4、TNFAIP3、DGKA、DGKZ、FAS、EGR2、PPP2R2D、PSGL-1、KDM6A和/或TET2基因),以使其各自独立地被敲除、敲减或敲入,从而影响一个或多个免疫细胞的功能、增殖和存活。

[0386] 在一个实例中,可通过在序列的一个或多个区域处诱导包括缺失、置换、插入或突变在内的改变,来靶向和操纵免疫细胞中表达的一种或多种基因(例如选自于由如下基因所组成的组中的一种或多种基因:PD-1、CTLA-4、TNFAIP3、DGKA、DGKZ、FAS、EGR2、PPP2R2D、PSGL-1、KDM6A和/或TET2基因),以使其各自独立地被敲除、敲减或敲入,从而影响一个或多个免疫细胞的功能、增殖和存活。

[0387] 具体而言,很明显的是,也可组合和靶向本文未公开的免疫调节基因。

[0388] [免疫系统]

[0389] 此外,本发明的另一方面提供了形成免疫应答机制的免疫系统,所述免疫系统包含经人工操纵的免疫调节因子和/或包含所述经人工操纵的免疫调节因子的细胞。

[0390] 本申请的术语“免疫系统”包括借助经操纵的免疫调节因子的功能改变来影响体内免疫应答(即,参与表现出新免疫效力的机制)的全部现象,并包括与该免疫系统直接或间接相关的全部物质、组合物、方法和用途。例如,免疫系统包含参与先天免疫、适应性免疫、细胞免疫、体液免疫、主动免疫和被动免疫应答的全部基因、免疫细胞和免疫器官/组织。

[0391] 构成此类免疫系统的元件通常统称为“免疫系统因子”。

[0392] 本发明的免疫系统包含经操纵的免疫细胞。

[0393] 经操纵的免疫细胞是指经历了人工操纵而不处于天然状态的免疫细胞。最近积极研究了通过从体内提取免疫细胞以及实施人工操纵来增强免疫力的技术。由于针对某些疾病具有优异的免疫效力,此类经操纵的免疫细胞已被证明是一种新治疗方法。特别地,对经操纵的免疫细胞的研究已与癌症治疗相关联地积极进行。

[0394] 经操纵的免疫细胞可为经功能性操纵的免疫细胞或者增补人工结构的免疫细胞。

[0395] 经功能性操纵的免疫细胞

[0396] 本发明的“经功能性操纵的免疫细胞”是指其中的天然状态下的野生型受体或免疫调节因子被操纵的免疫细胞。

[0397] 下文中,操纵是指全部类型的操纵,包括基因的切割、连接、去除、插入和修饰;或者蛋白的去除、增加或修饰,本领域普通技术人员能够将它们用于操纵蛋白和基因。下文中,免疫细胞不仅包括分化免疫细胞,还包括分化前细胞(例如干细胞)。

[0398] 经功能性操纵的免疫细胞可为其中的野生型受体被操纵的免疫细胞。具体而言,野生型受体可为TCR。

[0399] 经功能性操纵的免疫细胞可为其中的野生型受体缺失或以较低比率存在于表面上的免疫细胞。

[0400] 经功能性操纵的免疫细胞可为其中的野生型受体以较高比率存在于表面上的免疫细胞。

[0401] 经功能性操纵的免疫细胞可为其中的野生型受体对特定抗原具有增强的识别能力的免疫细胞。

[0402] 经功能性操纵的免疫细胞可为其中的免疫调节因子被操纵的免疫细胞。

[0403] 经功能性操纵的免疫细胞可为其中的免疫细胞活性调节元件被操纵的免疫细胞。

[0404] 具体而言,经功能性操纵的免疫细胞可为其中选自于由SHP-1、PD-1、CTLA-4、CBLB、ILT-2、KIR2DL4和PSGL-1所组成的组中的一种或多种失活的免疫细胞。

[0405] 经功能性操纵的免疫细胞可为其中的免疫细胞生长调节元件被操纵的免疫细胞。

[0406] 具体而言,经功能性操纵的免疫细胞可为其中选自于由DGK α 、DGK ζ 、FAS、EGR2、EGR3、PPP2R2D和A20所组成的组中的一种或多种失活的免疫细胞。

[0407] 经功能性操纵的免疫细胞可为其中的免疫细胞死亡调节元件被操纵的免疫细胞。

[0408] 具体而言,经功能性操纵的免疫细胞可为其中选自于由Daxx、Bim、Bid、BAD、PD-1和CTLA-4所组成的组中的一种或多种失活的免疫细胞。

[0409] 此外,经功能性操纵的免疫细胞可为其中插入有诱导自身死亡的元件的免疫细

胞。

[0410] 经功能性操纵的免疫细胞可为其中的免疫细胞耗竭调节元件被操纵的免疫细胞。

[0411] 具体而言,经功能性操纵的免疫细胞可为其中选自于由TET2、Wnt和Akt所组成的组中的一种或多种失活的免疫细胞。

[0412] 经功能性操纵的免疫细胞可为其中的细胞因子产生调节元件被操纵的免疫细胞。

[0413] 经功能性操纵的免疫细胞可为其中的抗原结合调节元件被操纵的免疫细胞。

[0414] 具体而言,经功能性操纵的免疫细胞可为其中选自于由dCK、CD52、B2M和MHC所组成的组中的一种或多种失活的免疫细胞。

[0415] 经功能性操纵的免疫细胞可为其中的不同于上文提到的免疫调节因子的免疫调节因子被操纵的免疫细胞。

[0416] 经功能性操纵的免疫细胞可为其中的一种或多种免疫调节因子被同时操纵的免疫细胞。具体而言,一种或多种类型的免疫调节因子可被操纵。

[0417] 通过对野生型受体和免疫调节因子进行操纵,经功能性操纵的免疫细胞可具有新免疫效力。

[0418] 具体而言,当操纵一种免疫调节因子时,并不一定意味着必须表现出新免疫调节效果。对一种免疫调节因子的操纵可造成多种新免疫效力或抑制多种新免疫效力。

[0419] 新免疫效力可为其中调节了识别特定抗原的能力的免疫效力。

[0420] 新免疫效力可为其中改善了识别特定抗原的能力的免疫效力。

[0421] 具体而言,特定抗原可为疾病抗原,例如癌细胞抗原。

[0422] 新免疫效力可为其中识别特定抗原的能力恶化的免疫效力。

[0423] 新免疫效力可为其中改善了新免疫效力的免疫效力。

[0424] 新免疫效力可为其中调控了免疫细胞生长的免疫效力。具体而言,免疫效力可为其中促进或延迟了生长和分化的免疫效力。

[0425] 新免疫效力可为其中调控了免疫细胞死亡的免疫效力。具体而言,免疫效力可为防止免疫细胞死亡。此外,免疫效力可为在经过合适的时间后引起免疫细胞自杀。

[0426] 新免疫效力可为其中减轻了免疫细胞的功能丧失的免疫效力。

[0427] 新免疫效力可为其中调控了免疫细胞的细胞因子分泌的免疫效力。具体而言,免疫效力可为促进或抑制细胞因子分泌。

[0428] 新免疫效力可为调控免疫细胞中野生型受体的抗原结合能力。具体而言,免疫效力可为改善野生型抗体针对特定抗原的特异性。

[0429] 增补人工结构的免疫细胞

[0430] 术语“增补人工结构的免疫细胞”是指在免疫细胞中增补人工结构。

[0431] 例如,增补人工结构的免疫细胞可为其中增补了人工受体的免疫细胞。

[0432] 所述人工受体可为对特定抗原具有识别能力的人工受体。在一个实例中,增补人工结构的免疫细胞可为CAR-T细胞。

[0433] 此外,可以增补对造成特定疾病的两种以上抗原分别具有识别能力的人工受体。具体而言,各人工受体可以根据情况以时间依赖性方式表达。

[0434] 例如,在将经操纵的免疫细胞用于癌症治疗的情况下,第一人工受体可产生引发第二人工受体基因表达的免疫应答信号,随后第二人工受体可表达。第二人工受体可产生

诱导针对癌症细胞的免疫应答的免疫应答信号。在该情况下,经操纵的免疫细胞攻击癌细胞的能力可得到改善。

[0435] 人工受体可为具有识别经操纵的免疫细胞的能力的人工受体。

[0436] 人工受体可为具有识别体内一般细胞的能力的人工受体。在一个实例中,增补人工结构的免疫细胞可为iCAR-T细胞。

[0437] 人工受体可为具有识别第三物质的能力的人工受体。具体而言,第三物质可具有结合特定疾病的抗原的能力。

[0438] 具体而言,第三物质可能够与人工受体结合并同时与特定疾病的抗原结合。例如,第三物质可具有同时结合人工受体和癌细胞相关抗原的能力。

[0439] 在另一实例中,增补人工结构的免疫细胞可为其中除人工受体外还增补了具有特定功能的人工结构的免疫细胞。

[0440] 在免疫细胞中增补有人工结构(不同于天然状态)的情况下,增补人工结构的免疫细胞可具有新免疫效力。

[0441] 例如,新免疫效力可为其中免疫细胞与特定抗原结合,从而免疫细胞与抗原处于特定位置关系的免疫效力。

[0442] 新免疫效力可为识别和促进针对特定抗原的免疫应答的功能。

[0443] 新免疫效力可为抑制过度免疫应答的功能。

[0444] 新免疫效力可为调节免疫应答的信号转导通路的功能。

[0445] 新免疫效力可为免疫细胞与第三物质形成结合并确认特定疾病的功能。具体而言,第三物质可为特定疾病的生物标志物。

[0446] 上述特定抗原的一个优选实例可为癌细胞抗原。

[0447] 癌细胞抗原可包括但不限于:AD034、AKT1、BRAP、CAGE、CDX2、CLP、CT-7、CT8/HOM-TES-85、cTAGE-1、EGFR、EGFRvIII、腓骨蛋白-1、HAGE、HCA587/MAGE-C2、hCAP-G、HCE661、HER2/neu、HLA-Cw、HOM-HD-21/半乳糖凝集素9、HOM-MEEL-40/SSX2、HOM-RCC-3.1.3/CAXII、HOXA7、HOXB6、Hu、HUB 1、KM-HN-3、KM-KN-1、KOC1、KOC2、KOC3、KOC3、LAGE-1、MAGE-1、MAGE-4a、MPP1 1、MSLN、NNP-1、NY-BR-1、NY-BR-62、NY-BR-85、NY-CO-37、NY-CO-38、NY-ESO-1、NY-ESO-5、NY-LU-12、NY-REN-10、NY-REN-19/LKB/STK1 1、NY-REN-21、NY-REN-26/BCR、NY-REN-3/NY-CO-38、NY-REN-33/SNC6、NY-REN-43、NY-REN-65、NY-REN-9、NY-SAR-35、OGFr、PLU-1、PSMA、PSCA、Rab38、RBPJkappa、RHAMM、SCP1、SCP-1、SSX3、SSX4、SSX5、TOP2A、TOP2B、ROR-1以及酪氨酸酶。

[0448] 经混合型操纵的免疫细胞(hybrid manipulated immune cell)

[0449] 术语“经混合型操纵的免疫细胞”是指其中同时实现免疫调节因子的操纵和人工结构的增补的免疫细胞。

[0450] 在经混合型操纵的免疫细胞中,免疫调节因子的操纵与上文就经功能性操纵的免疫细胞所描述的相同。此外,人工结构的增补与上文就增补人工结构的免疫细胞所描述的相同。

[0451] 当对免疫细胞功能的操纵是基因操纵时,增补人工结构的位置可以与基因上发生功能性操纵的位置相同。

[0452] 经混合型操纵的免疫细胞的新免疫效力可包括经功能性操纵的免疫细胞和增补

人工结构的免疫细胞的新免疫效力,并由于这些细胞之间的相互作用而表现出更为改善的免疫效力。

[0453] 改善的免疫效力可为对特定疾病的特异性和免疫应答得到改善。在优选的实例中,经混合型操纵的免疫细胞可为癌症特异性和免疫应答均得到改善的免疫细胞。

[0454] 本发明的免疫系统包括通过经操纵的免疫调节因子和/或经操纵的免疫细胞实现的期望免疫应答和通过该免疫应答进行疾病治疗的机制。

[0455] 在实施方式中,可将其中使得抑制免疫细胞增殖的基因失活的免疫调节因子和/或经操纵的免疫细胞用于影响免疫细胞的增殖。

[0456] 在实施方式中,可将其中使得介导免疫细胞死亡的基因失活的免疫调节因子和/或经操纵的免疫细胞用于影响免疫细胞的存活。

[0457] 在实施方式中,可将其中使得编码用于阻遏和抑制免疫力的信号转导因子的基因失活的免疫调节因子和/或经操纵的免疫细胞用于影响免疫细胞的功能。

[0458] 本文所述的方法和组合物可单独使用或将它们组合使用,从而影响对作为特定疾病的治疗处理的经基因操纵的免疫细胞的效力具有限制作用的一种或多种因素(例如免疫细胞增殖、免疫细胞存活、免疫细胞功能或它们的任意组合)。

[0459] 同时,术语“免疫调节疗法”是指通过使用经操纵的免疫调节因子和/或经操纵的免疫细胞调节体内免疫应答来治疗疾病。

[0460] 例如,可通过利用免疫细胞(例如树突细胞、自然杀伤细胞、T细胞等)使得体内免疫应答活化或失活来治疗疾病。

[0461] 已将此类免疫调节疗法开发为主要针对癌症疗法作为适应证,由于通过直接将免疫细胞给予患者来活化免疫功能从而引发疗效,免疫调节疗法是不同于用于现有癌症治疗的手术疗法、抗癌剂或放射疗法的治疗机制。

[0462] 在免疫调节疗法的实施方式中,根据所用免疫细胞的特征和制造过程中导入细胞的基因,免疫调节疗法包括树突免疫调节细胞治疗剂、淋巴因子活化杀伤细胞(LAK)、肿瘤浸润性T淋巴细胞(TIL)、T细胞受体修饰的T细胞(TCR-T)、嵌合抗原受体修饰的T细胞(CAR-T)等。

[0463] [基因操纵或修饰]

[0464] 本发明的免疫调节因子、免疫细胞和免疫系统中涉及的物质的操纵或修饰可优选通过经基因操纵来实现。

[0465] 一方面,可通过靶向影响免疫细胞增殖、存活和/或功能的免疫调节基因的非编码区和编码区的全部或部分来提供用于基因操纵的组合物和方法。

[0466] 在实施方式中,为形成期望的免疫系统,所述组合物和方法可对该免疫系统所涉及的一种或多种免疫调节基因进行操纵或修饰。这可通过对构成该基因的核酸进行修饰而实现。作为操纵结果,包括处于敲减、敲除和敲入形式的全部方面。

[0467] 在实施方式中,可靶向启动子区或转录序列(例如内含子序列、外显子序列)。可靶向编码序列(例如编码区和起始编码区)来改变表达或进行敲除。

[0468] 在实施方式中,核酸的改变可为多个核苷酸(例如在1bp-30bp、1bp-27bp、1bp-25bp、1bp-23bp、1bp-20bp、1bp-15bp、1bp-10bp、1bp-5bp、1bp-3bp范围内)或1bp的核苷酸的置换、删除和/或插入。

[0469] 在实施方式中,对于免疫调节基因中一种或多种基因的敲除或一种或多种表达的去除或者对于一个等位基因或两个等位基因中的一个或多个的敲除而言,可靶向所述基因以使得所述免疫调节基因中的一种或多种中包含缺失或突变。

[0470] 在实施方式中,可利用基因敲减来降低不希望的等位基因或转录物的表达。

[0471] 在实施方式中,可通过靶向启动子、增强子、内含子、3'UTR和/或多聚腺苷酸信号的非编码序列来改变影响免疫细胞功能的免疫调节基因。

[0472] 在实施方式中,可通过改变基因的核酸来诱导对免疫调节基因活性的调节(例如活化、失活)。

[0473] 在实施方式中,改变基因的核酸可以通过借助引导核酸-编辑蛋白复合体在靶基因的特定区域中切割单链或双链(即,通过催化核酸链的断裂)来使得靶基因失活。

[0474] 在实施方式中,可借助诸如同源重组、非同源末端接合(NHEJ)等机制来修复核酸链的断裂。

[0475] 在该情况下,当出现NHEJ机制时,在切割位点中诱导DNA序列的改变,并可通过相同的方式使基因失活。借助NHEJ的修复可引起短基因片段的置换、插入或删除,并可用于诱导相应的基因敲除。

[0476] 另一方面,本发明可提供用于上述基因操纵的位置。

[0477] 在实施方式中,当通过NHEJ介导的改变实现所述改变时,用于基因操纵的位置是指基因中导致免疫调节基因产物表达降低或去除的位置。

[0478] 例如,基因中的所述位置可处于起始编码区中。

[0479] 可处于50%上游编码区中。

[0480] 可处于启动子序列中。

[0481] 可处于特定内含子序列中。

[0482] 可处于特定外显子序列中。

[0483] 所述位置可位于基因中影响免疫细胞增殖、存活和/或功能的特定位置。

[0484] 所述位置可位于基因中影响参与免疫应答的蛋白的功能的特定位置。

[0485] 所述位置可位于基因中影响对特定抗原的识别能力的特定位置。

[0486] 所述位置可位于基因中影响对免疫细胞中的细胞因子分泌进行调节的功能的特定位置。

[0487] 所述位置可位于基因中影响对免疫细胞中受体的抗原结合能力进行调节的功能的特定位置。

[0488] 可考虑基因表达的调节过程来实施基因操纵。

[0489] 在实施方式中,可通过选择适于各步骤的操纵方法,在如下步骤中实施基因操纵:转录调节、RNA加工调节、RNA转运调节、RNA降解调节、翻译调节或蛋白修饰调节。

[0490] 在实施方式中,可利用RNA干扰(RNAi)或RNA沉默通过小RNA(sRNA)来阻碍mRNA或降低其稳定性,从而控制遗传信息的表达;并且,在一些情况下,可通过破坏来阻止中间步骤期间蛋白合成信息的递送,从而控制遗传信息的表达。

[0491] 在实施方式中,可使用能够催化DNA或RNA分子水解(切割)、优选能够催化DNA分子中的核酸之间的键水解(切割)的野生型或变体酶。可使用引导核酸-编辑蛋白复合体。

[0492] 例如,可通过使用选自于由如下核酸酶所组成的组中的一种或多种对基因进行操

纵来控制遗传信息的表达:大范围核酸酶、锌指核酸酶、CRISPR/Cas9 (Cas9蛋白)、CRISPR-Cpf1 (Cpf1蛋白) 以及TALE核酸酶。

[0493] 在优选实例中,可使用引导核酸-编辑蛋白复合体(例如CRISPR/Cas系统),借助非同源末端接合(NHEJ)或同源介导的修复(HDR)来介导基因操纵,但不限于此。

[0494] 在实施方式中,影响免疫细胞增殖、存活和/或功能的免疫调节基因的实例可包括PD-1基因、CTLA-4基因、TNFAIP3基因、DGKA基因、DGKZ基因、FAS基因、EGR2基因、PPP2R2D基因、TET2基因、PSGL-1基因或KDM6A基因。

[0495] 下表1中总结了上述基因的靶序列区(即,可发生核酸修饰的位点)(表1中示出的靶序列部分描述为在3'末端含有PAM序列5'-NGG-3')。

[0496] 可同时靶向两种以上类型的靶序列。

[0497] 可同时靶向两种以上类型的基因。

[0498] 可同时靶向同源基因中的两种以上的靶序列或异源基因中的两种以上的靶序列。

[0499] 在示例性实施方式中,可分别靶向DGKa或DGKz。

[0500] 在示例性实施方式中,可同时靶向DGKa和DGKz。

[0501] [表1]靶序列

[0502]

靶基因	DNA 靶序列	序列号
A20	CTTGTGGCGCTGAAAACGAACGG	SEQ ID NO: 1
	ATGCCACTTCTCAGTACATGTGG	SEQ ID NO: 2
	GCCACTTCTCAGTACATGTGGGG	SEQ ID NO: 3
	GCCCCACATGTACTGAGAAGTGG	SEQ ID NO: 4
	TCAGTACATGTGGGGCGTTCAGG	SEQ ID NO: 5
	GGGCGTTCAGGACACAGACTTGG	SEQ ID NO: 6
	CACAGACTTGGTACTGAGGAAGG	SEQ ID NO: 7
	GGCGCTGTTCAGCACGCTCAAGG	SEQ ID NO: 8
	CACGCAACTTTAAATTCCGCTGG	SEQ ID NO: 9
	CGGGGCTTTGCTATGATACTCGG	SEQ ID NO: 10
	GGCTTCCACAGACACACCCATGG	SEQ ID NO: 11
	TGAAGTCCACTTCGGGCCATGGG	SEQ ID NO: 12
靶基因	DNA 靶序列	序列号
DGKα	CTGTACGACACGGACAGAAATGG	SEQ ID NO: 13
	TGTACGACACGGACAGAAATGGG	SEQ ID NO: 14
	CACGGACAGAAATGGGATCCTGG	SEQ ID NO: 15
	GATGCGAGTGGCTGAATACCTGG	SEQ ID NO: 16
	GAGTGGCTGAATACCTGGATTGG	SEQ ID NO: 17
	AGTGGCTGAATACCTGGATTGGG	SEQ ID NO: 18
	ATTGGGATGTGTCTGAGCTGAGG	SEQ ID NO: 19
	ATGAAAGAGATTGACTATGATGG	SEQ ID NO: 20
	CTCTGTCTCTCAAGCTGAGTGGG	SEQ ID NO: 21
	TCTCTCAAGCTGAGTGGGTCCGG	SEQ ID NO: 22
	CTCTCAAGCTGAGTGGGTCCGGG	SEQ ID NO: 23
	CAAGCTGAGTGGGTCCGGGCTGG	SEQ ID NO: 24
靶基因	DNA 靶序列	序列号
	TTGACATGACTGGAGAGAAGAGG	SEQ ID NO: 25
	GACTGGAGAGAAGAGGTCGTTGG	SEQ ID NO: 26
	GAGACGGGAGCAAAGCTGCTGGG	SEQ ID NO: 27

[0503]

EGR2

AGAGACGGGAGCAAAGCTGCTGG	SEQ ID NO: 28
TGGTTTCTAGGTGCAGAGACGGG	SEQ ID NO: 29
TAAGTGAAGGTCTGGTTTCTAGG	SEQ ID NO: 30
TGCCCATGTAAGTGAAGGTCTGG	SEQ ID NO: 31
GAACCTTGCCCATGTAAGTGAAGG	SEQ ID NO: 32
TCCATTGACCCTCAGTACCCTGG	SEQ ID NO: 33
TATGCCTTCTGGGTAGCAGCTGG	SEQ ID NO: 34
TGAGTGCAGGCATCTTGCAAGGG	SEQ ID NO: 35
GAGTGCAGGCATCTTGCAAGGGG	SEQ ID NO: 36
GATGAGGCTGTGGTTGAAGCTGG	SEQ ID NO: 37
CCACTGGCCACAGGACCCCTGGG	SEQ ID NO: 38
GGGACATGGTGCACACACCCAGG	SEQ ID NO: 39
GAGTACAGGTGGTCCAGGTCAGG	SEQ ID NO: 40
GCGGAGAGTACAGGTGGTCCAGG	SEQ ID NO: 41
GCGGTGGCGGAGAGTACAGGTGG	SEQ ID NO: 42
TCTCCTGCACAGCCAGAATAAGG	SEQ ID NO: 43
ACGCAGAAGGGTCCTGGTAGAGG	SEQ ID NO: 44
AGGTGGTGGGTAGGCCAGAGAGG	SEQ ID NO: 45
CCCAAGCCAGCCACGGACCCAGG	SEQ ID NO: 46
ACCTGGGTCCGTGGCTGGCTTGG	SEQ ID NO: 47
AAGAGACCTGGGTCCGTGGCTGG	SEQ ID NO: 48
GGATCATTGGGAAGAGACCTGGG	SEQ ID NO: 49
GGGATCATTGGGAAGAGACCTGG	SEQ ID NO: 50
CAGGATAGTCTGGGATCATTGGG	SEQ ID NO: 51
GGAAAGAATCCAGGATAGTCTGG	SEQ ID NO: 52
CAGTGCCAGAGAGACCTACATGG	SEQ ID NO: 53
CTGTACCATGTAGGTCTCTCTGG	SEQ ID NO: 54
AGAGACCTACATGGTACAGCTGG	SEQ ID NO: 55
CTGGGCCAGCTGTACCATGTAGG	SEQ ID NO: 56
AGGGAAAGGGCTTACGGTCTGGG	SEQ ID NO: 57

[0504]

	CAGGGAAAGGGCTTACGGTCTGG	SEQ ID NO: 58
靶基因	DNA 靶序列	序列号
PPP2R2D	TCTGGAGATCTTCTTGCAACAGG	SEQ ID NO: 59
	CTCCGGTTCATGACTTTGAAAGG	SEQ ID NO: 60
	GTCTTCCATCTTCGTCTTTCAGG	SEQ ID NO: 61
	GAAGACTTCGAGACCCATTTAGG	SEQ ID NO: 62
	TCGAGACCCATTTAGGATCACGG	SEQ ID NO: 63
	GTAGCGCCGTGATCCTAAATGGG	SEQ ID NO: 64
	CGTAGCGCCGTGATCCTAAATGG	SEQ ID NO: 65
	CATTTAGGATCACGGCGCTACGG	SEQ ID NO: 66
	GGTCCCAATATTGAAGCCCATGG	SEQ ID NO: 67
	GATCCATGGGCTTCAATATTGGG	SEQ ID NO: 68
	AGATCCATGGGCTTCAATATTGG	SEQ ID NO: 69
	GCTTCTACCATAAGATCCATGGG	SEQ ID NO: 70
	CGCTTCTACCATAAGATCCATGG	SEQ ID NO: 71
	GCATTTGCAAAAATTCGCCGTGG	SEQ ID NO: 72
	ATGACCTGAGAATTAATTTATGG	SEQ ID NO: 73
	CCATGCACTCCCAGACATCGTGG	SEQ ID NO: 74
	GCACTGGTGCGGGTGGAACCTCGG	SEQ ID NO: 75
	ACACGTTGCACTGGTGCGGGTGG	SEQ ID NO: 76
	CGAACACGTTGCACTGGTGCGGG	SEQ ID NO: 77
	ACGAACACGTTGCACTGGTGCGG	SEQ ID NO: 78
	TGTAGACGAACACGTTGCACTGG	SEQ ID NO: 79
	GCGCATGTCACACAGGCGGATGG	SEQ ID NO: 80
	AGGAGCGCATGTCACACAGGCGG	SEQ ID NO: 81
	CCGAGGAGCGCATGTCACACAGG	SEQ ID NO: 82
	CCTGTGTGACATGCGCTCCTCGG	SEQ ID NO: 83
靶基因	DNA 靶序列	序列号
	CGACTGGCCAGGGCGCCTGTGGG	SEQ ID NO: 84
	ACCGCCCAGACGACTGGCCAGGG	SEQ ID NO: 85

[0505]

PD-1	CACCGCCCAGACGACTGGCCAGG	SEQ ID NO: 86
	GTCTGGGCGGTGCTACAACTGGG	SEQ ID NO: 87
	CTACAACTGGGCTGGCGGCCAGG	SEQ ID NO: 88
	CACCTACCTAAGAACCATCCTGG	SEQ ID NO: 89
	CGGTCACCACGAGCAGGGCTGGG	SEQ ID NO: 90
	GCCCTGCTCGTGGTGACCGAAGG	SEQ ID NO: 91
	CGGAGAGCTTCGTGCTAAACTGG	SEQ ID NO: 92
	CAGCTTGTCCGTCTGGTTGCTGG	SEQ ID NO: 93
	AGGCGGCCAGCTTGTCCGTCTGG	SEQ ID NO: 94
	CCGGGCTGGCTGCGGTCCTCGGG	SEQ ID NO: 95
	CGTTGGGCAGTTGTGTGACACGG	SEQ ID NO: 96
靶基因	DNA 靶序列	序列号
CTLA-4	CATAAAGCCATGGCTTGCCTTGG	SEQ ID NO: 97
	CCTTGGATTTCAGCGGCACAAGG	SEQ ID NO: 98
	CCTTGTGCCGCTGAAATCCAAGG	SEQ ID NO: 99
	CACTCACCTTTGCAGAAGACAGG	SEQ ID NO: 100
	TTCCATGCTAGCAATGCACGTGG	SEQ ID NO: 101
	GGCCACGTGCATTGCTAGCATGG	SEQ ID NO: 102
	GGCCCAGCCTGCTGTGGTACTGG	SEQ ID NO: 103
	AGGTCCGGGTGACAGTGCTTCGG	SEQ ID NO: 104
	CCGGGTGACAGTGCTTCGGCAGG	SEQ ID NO: 105
	CTGTGCGGCAACCTACATGATGG	SEQ ID NO: 106
	CAACTCATTCCCCATCATGTAGG	SEQ ID NO: 107
	CTAGATGATTCCATCTGCACGGG	SEQ ID NO: 108
靶基因	DNA 靶序列	序列号
	GGCTAGGAGTCAGCGACATATGG	SEQ ID NO: 109
	GCTAGGAGTCAGCGACATATGGG	SEQ ID NO: 110
	CTAGGAGTCAGCGACATATGGGG	SEQ ID NO: 111
	GTA CTGTGTAGCCAGGATGCTGG	SEQ ID NO: 112
	ACGAGCACTCACCAGCATCCTGG	SEQ ID NO: 113

[0506]	DGKζ	AGGCTCCAGGAATGTCCGCGAGG	SEQ ID NO: 114
		ACTTACCTCGCGGACATTCCTGG	SEQ ID NO: 115
		CACCCTGGGCACTTACCTCGCGG	SEQ ID NO: 116
		GTGCCGTACAAAGGTTGGCTGGG	SEQ ID NO: 117
		GGTGCCGTACAAAGGTTGGCTGG	SEQ ID NO: 118
		CTCTCCTCAGTACCACAGCAAGG	SEQ ID NO: 119
		CCTGGGGCCTCCGGGCGCGGAGG	SEQ ID NO: 120
		AGTACTCACCTGGGGCCTCCGGG	SEQ ID NO: 121
		AGGGTCTCCAGCGGCCCTCCTGG	SEQ ID NO: 122
		GCAAGTACTTACGCCTCCTTGGG	SEQ ID NO: 123
		TTGCGGTACATCTCCAGCCTGGG	SEQ ID NO: 124
		TTTGCGGTACATCTCCAGCCTGG	SEQ ID NO: 125
	靶基因	DNA 靶序列	序列号
		GCAAAACCTGTCCACTCTTATGG	SEQ ID NO: 126
		TTGGTGCCATAAGAGTGGACAGG	SEQ ID NO: 127
		GGTGCAAGTTTCTTATATGTTGG	SEQ ID NO: 128
		ACCTGATGCATATAATAATCAGG	SEQ ID NO: 129
		ACCTGATTATTATATGCATCAGG	SEQ ID NO: 130
		CAGAGCACCAGAGTGCCGTCTGG	SEQ ID NO: 131
		AGAGCACCAGAGTGCCGTCTGGG	SEQ ID NO: 132
		AGAGTGCCGTCTGGGTCTGAAGG	SEQ ID NO: 133
		AGGAAGGCCGTCCATTCTCAGGG	SEQ ID NO: 134
		GGATAGAACCAACCATGTTGAGG	SEQ ID NO: 135
		TCTGTTGCCCTCAACATGGTTGG	SEQ ID NO: 136
		TTAGTCTGTTGCCCTCAACATGG	SEQ ID NO: 137
		GTCTGGCAAATGGGAGGTGATGG	SEQ ID NO: 138
		CAGAGGTTCTGTCTGGCAAATGG	SEQ ID NO: 139
		TTGTAGCCAGAGGTTCTGTCTGG	SEQ ID NO: 140
		ACTTCTGGATGAGCTCTCTCAGG	SEQ ID NO: 141
		AGAGCTCATCCAGAAGTAAATGG	SEQ ID NO: 142

[0507]

Tet2	TTGGTGTCTCCATTTACTTCTGG	SEQ ID NO: 143
	TTCTGGCTTCCCTTCATACAGGG	SEQ ID NO: 144
	CAGGACTCACACGACTATTCTGG	SEQ ID NO: 145
	CTACTTTCTTGTGTAAAGTCAGG	SEQ ID NO: 146
	GACTTTACACAAGAAAGTAGAGG	SEQ ID NO: 147
	GTCTTTCTCCATTAGCCTTTTGG	SEQ ID NO: 148
	AATGGAGAAAGACGTAACCTTCGG	SEQ ID NO: 149
	ATGGAGAAAGACGTAACCTTCGGG	SEQ ID NO: 150
	TGGAGAAAGACGTAACCTTCGGGG	SEQ ID NO: 151
	TTTGGTTGACTGCTTTCACCTGG	SEQ ID NO: 152
	TCACTCAAATCGGAGACATTTGG	SEQ ID NO: 153
	ATCTGAAGCTCTGGATTTTCAGG	SEQ ID NO: 154
	GCTTCAGATTCTGAATGAGCAGG	SEQ ID NO: 155
	CAGATTCTGAATGAGCAGGAGGG	SEQ ID NO: 156
	AAGGCAGTGCTAATGCCTAATGG	SEQ ID NO: 157
	GCAGAACTGTAGCACCATTAGG	SEQ ID NO: 158
	ACCGCAATGGAAACACAATCTGG	SEQ ID NO: 159
	TGTGGTTTTCTGCACCGCAATGG	SEQ ID NO: 160
	CATAAATGCCATTAACAGTCAGG	SEQ ID NO: 161
	ATTAGTAGCCTGACTGTTAATGG	SEQ ID NO: 162
	CGATGGGTGAGTGATCTCACAGG	SEQ ID NO: 163
	ACTCACCCATCGCATACCTCAGG	SEQ ID NO: 164
	CTCACCCATCGCATACCTCAGGG	SEQ ID NO: 165
靶基因	DNA 靶序列	序列号
	AGCAACAGGAGGAGTTGCAGAGG	SEQ ID NO: 166
	CCAGTAGGATCAGCAACAGGAGG	SEQ ID NO: 167
	CTCCTGTTGCTGATCCTACTGGG	SEQ ID NO: 168
	GGCCCAGTAGGATCAGCAACAGG	SEQ ID NO: 169
	TTGCTGATCCTACTGGGCCCTGG	SEQ ID NO: 170
	TGGCAACAGCTTGCAGCTGTGGG	SEQ ID NO: 171

[0508]

CTTGGGTCCCCTGCTTGCCCGGG	SEQ ID NO: 172
GTCCCCTGCTTGCCCGGGACCGG	SEQ ID NO: 173
CTCCGGTCCCGGGCAAGCAGGGG	SEQ ID NO: 174
TCTCCGGTCCCGGGCAAGCAGGG	SEQ ID NO: 175
GTCTCCGGTCCCGGGCAAGCAGG	SEQ ID NO: 176
GCTTGCCCGGGACCGGAGACAGG	SEQ ID NO: 177
GGTGGCCTGTCTCCGGTCCCGGG	SEQ ID NO: 178
CGGTGGCCTGTCTCCGGTCCCGG	SEQ ID NO: 179
CATATTCGGTGGCCTGTCTCCGG	SEQ ID NO: 180
ATCTAGGTACTCATATTCGGTGG	SEQ ID NO: 181
ATAATCTAGGTACTCATATTCGG	SEQ ID NO: 182
TTATGATTTCCTGCCAGAAACGG	SEQ ID NO: 183
ATTCTGGAGGCTCCGTTTCTGG	SEQ ID NO: 184
ACTGACACCACTCCTCTGACTGG	SEQ ID NO: 185
CTGACACCACTCCTCTGACTGGG	SEQ ID NO: 186
ACCACTCCTCTGACTGGGCCTGG	SEQ ID NO: 187
AACCCCTGAGTCTACCACTGTGG	SEQ ID NO: 188
CTCCACAGTGGTAGACTCAGGGG	SEQ ID NO: 189
GCTCCACAGTGGTAGACTCAGGG	SEQ ID NO: 190
GGCTCCACAGTGGTAGACTCAGG	SEQ ID NO: 191
CCTGCTGCAAGGCGTTCTACTGG	SEQ ID NO: 192
CCAGTAGAACGCCTTGACAGCAGG	SEQ ID NO: 193
CGTTCTACTGGCCTGGATGCAGG	SEQ ID NO: 194
TCTACTGGCCTGGATGCAGGAGG	SEQ ID NO: 195
CCACGGAGCTGGCCAACATGGGG	SEQ ID NO: 196
CGTGGACAGGTTCCCATGTTGG	SEQ ID NO: 197
GTCCACGGATTCAGCAGCTATGG	SEQ ID NO: 198
GACCACTCAACCAGTGCCACGG	SEQ ID NO: 199
GGAGTGGTCTGTGCCTCCGTGGG	SEQ ID NO: 200
GGCACAGACAACCTGACTGACGG	SEQ ID NO: 201

[0509]

PSGL-1

GACAACTCGACTGACGGCCACGG	SEQ ID NO: 202
AACTCGACTGACGGCCACGGAGG	SEQ ID NO: 203
CACAGAACCCAGTGCCACAGAGG	SEQ ID NO: 204
GGTAGTAGGTTCCATGGACAGGG	SEQ ID NO: 205
TGGTAGTAGGTTCCATGGACAGG	SEQ ID NO: 206
TCTTTTGGTAGTAGGTTCCATGG	SEQ ID NO: 207
ATGGAACCTACTACCAAAGAGG	SEQ ID NO: 208
AACAGACCTCTTTTGGTAGTAGG	SEQ ID NO: 209
GGGTATGAACAGACCTCTTTTGG	SEQ ID NO: 210
TGTGTCCTCTGTTACTCACAAGG	SEQ ID NO: 211
GTGTCCTCTGTTACTCACAAGGG	SEQ ID NO: 212
GTAGTTGACGGACAAATTGCTGG	SEQ ID NO: 213
TTTGTCCGTCAACTACCCAGTGG	SEQ ID NO: 214
TTGTCCGTCAACTACCCAGTGGG	SEQ ID NO: 215
TGTCCGTCAACTACCCAGTGGGG	SEQ ID NO: 216
GTCCGTCAACTACCCAGTGGGGG	SEQ ID NO: 217
CTCTGTGAAGCAGTGCCTGCTGG	SEQ ID NO: 218
CCTGCTGGCCATCCTAATCTTGG	SEQ ID NO: 219
CCAAGATTAGGATGGCCAGCAGG	SEQ ID NO: 220
GGCCATCCTAATCTTGGCGCTGG	SEQ ID NO: 221
CACCAGCGCCAAGATTAGGATGG	SEQ ID NO: 222
AGTGCACACGAAGAAGATAGTGG	SEQ ID NO: 223
TATCTTCTTCGTGTGCACTGTGG	SEQ ID NO: 224
CTTCGTGTGCACTGTGGTGCTGG	SEQ ID NO: 225
GGCGGTCCGCCTCTCCCGCAAGG	SEQ ID NO: 226
GCGGTCCGCCTCTCCCGCAAGGG	SEQ ID NO: 227
AATTACGCACGGGGTACATGTGG	SEQ ID NO: 228
TGGGGGAGTAATTACGCACGGGG	SEQ ID NO: 229
GTGGGGGAGTAATTACGCACGGG	SEQ ID NO: 230
GGTGGGGGAGTAATTACGCACGG	SEQ ID NO: 231

[0510]

	TAATTACTCCCCACCGAGATGG	SEQ ID NO: 232
	AGATGCAGACCATCTCGGTGGGG	SEQ ID NO: 233
	GAGATGCAGACCATCTCGGTGGG	SEQ ID NO: 234
	TGAGATGCAGACCATCTCGGTGG	SEQ ID NO: 235
	GGATGAGATGCAGACCATCTCGG	SEQ ID NO: 236
	ATCTCATCCCTGTTGCCTGATGG	SEQ ID NO: 237
	TCATCCCTGTTGCCTGATGGGGG	SEQ ID NO: 238
	CTCACCCCCATCAGGCAACAGGG	SEQ ID NO: 239
	GAGGGCCCCCTCACCCCCATCAGG	SEQ ID NO: 240
	GGGCCCTCTGCCACAGCCAATGG	SEQ ID NO: 241
	CCCTCTGCCACAGCCAATGGGGG	SEQ ID NO: 242
	CCCCCATTGGCTGTGGCAGAGGG	SEQ ID NO: 243
	GCCCCCATTGGCTGTGGCAGAGG	SEQ ID NO: 244
	GGACAGGCCCCCATTGGCTGTGG	SEQ ID NO: 245
	CCGGGCTCTTGGCCTTGGACAGG	SEQ ID NO: 246
	CTGTCCAAGGCCAAGAGCCCGGG	SEQ ID NO: 247
	TGGCGTCAGGCCCGGGCTCTTGG	SEQ ID NO: 248
	CGGGCCTGACGCCAGAGCCCAGG	SEQ ID NO: 249
靶基因	DNA 靶序列	序列号
FAS	CAACAACCATGCTGGGCATCTGG	SEQ ID NO: 250
	GAGGGTCCAGATGCCCAGCATGG	SEQ ID NO: 251
	CATCTGGACCCTCCTACCTCTGG	SEQ ID NO: 252
	AGGGCTCACCAGAGGTAGGAGGG	SEQ ID NO: 253
	GGAGTTGATGTCAGTCACTTGGG	SEQ ID NO: 254
	TGGAGTTGATGTCAGTCACTTGG	SEQ ID NO: 255
	AGTGACTGACATCAACTCCAAGG	SEQ ID NO: 256
	GTGACTGACATCAACTCCAAGGG	SEQ ID NO: 247
	ACTCCAAGGGATTGGAATTGAGG	SEQ ID NO: 258
	CTTCCTCAATTCCAATCCCTTGG	SEQ ID NO: 259
	TACAGTTGAGACTCAGAACTTGG	SEQ ID NO: 260

[0511]		TTGGAAGGCCTGCATCATGATGG	SEQ ID NO: 261
		AGAATTGGCCATCATGATGCAGG	SEQ ID NO: 262
		GACAGGGCTTATGGCAGAATTGG	SEQ ID NO: 263
		TGTAACATACCTGGAGGACAGGG	SEQ ID NO: 264
		GTGTAACATACCTGGAGGACAGG	SEQ ID NO: 265
	靶基因	DNA 靶序列	序列号
	KDM6A	CGTACCTGTGCAACTCCTGTTGG	SEQ ID NO: 266
		GATCTACTGGAATTCCTAATGGG	SEQ ID NO: 267
		GAGTCAGCTGTTGGCCCATTAGG	SEQ ID NO: 268
		CTGCCTACAACTCAGTCTCTGG	SEQ ID NO: 269
		GGGCAGGCAGGACGGACTCCAGG	SEQ ID NO: 270
		GGAGTCCGTCCTGCCTGCCCTGG	SEQ ID NO: 271
		GAGTCCGTCCTGCCTGCCCTGGG	SEQ ID NO: 272
		GAAAAGGGTCCATTGGCCAAAGG	SEQ ID NO: 273
		GCCTGCAGAAAAGGGTCCATTGG	SEQ ID NO: 274
		TTGATGTGCTACAGGGAACATGG	SEQ ID NO: 275
		AGCGTTCTTGATGTGCTACAGGG	SEQ ID NO: 276
		CAGCGTTCTTGATGTGCTACAGG	SEQ ID NO: 277
		CTGTAGCACATCAAGAACGCTGG	SEQ ID NO: 278
		TGTAGCACATCAAGAACGCTGGG	SEQ ID NO: 279
		ATAGGCAATAATCATATAACAGG	SEQ ID NO: 280
		AGTGCGTTTCGCTGCAGGTAAGG	SEQ ID NO: 281
		GAGTGAGTGCGTTTCGCTGCAGG	SEQ ID NO: 282
		GTCAGGTTTGTGCGGTTATGAGG	SEQ ID NO: 283
		CGCTGCTGGTCAGGTTTGTGCGG	SEQ ID NO: 284
		AAACCTGACCAGCAGCGCAGAGG	SEQ ID NO: 285
		CCAGCAGCGCAGAGGAGCCGTGG	SEQ ID NO: 286
		CCACGGCTCCTCTGCGCTGCTGG	SEQ ID NO: 287
		CCAACTATCTAACTCCACTCAGG	SEQ ID NO: 288
		CCTGAGTGGAGTTAGATAGTTGG	SEQ ID NO: 289

[0512] [基因剪刀(经操纵的核酸酶)系统]

[0513] 可使用“引导核酸-编辑蛋白复合体”实现对本发明的免疫调节因子、免疫细胞以及免疫系统中涉及的物质进行的基因操纵或修饰。

[0514] 引导核酸-编辑蛋白复合体

[0515] 术语“引导核酸-编辑蛋白复合体”是指通过引导核酸和编辑蛋白之间的相互作用形成的复合体,该核酸-蛋白复合体包含引导核酸和编辑蛋白。

[0516] 术语“引导核酸”被设置为对引导核酸-蛋白复合体所靶向的蛋白、核酸、基因或染色体进行识别。

[0517] 引导核酸可以以DNA、RNA或者DNA/RNA混合物的形式存在,并可具有5-150个核酸的序列。

[0518] 引导核酸可包含一个或多个结构域。

[0519] 所述结构域可为引导结构域、第一互补结构域、接头结构域、第二互补结构域、近端(proximal)结构域或者尾部(tail)结构域,但不限于此。

[0520] 引导核酸可包含两个以上结构域,所述两个以上结构域可为相同结构域的重复,或可为不同的结构域。

[0521] 引导核酸可为一条连续的核酸序列。

[0522] 例如,所述一条连续的核酸序列可为(N)_m,其中N为A、T、C或G,或为A、U、C或G;m为1-150的整数。

[0523] 引导核酸可为两条以上连续的核酸序列。

[0524] 例如,所述两条以上连续的核酸序列可为(N)_m以及(N)_o,其中N代表A、T、C或G,或代表A、U、C或G;m和o为1-150的整数,并且可彼此相同或彼此不同。

[0525] 术语“编辑蛋白”是指能够与核酸直接结合或无需直接结合而与核酸相互作用的肽、多肽或蛋白。从概念上讲,编辑蛋白也可指“基因剪刀”或RNA引导的核酸内切酶(RGEN)。

[0526] 编辑蛋白可为酶。

[0527] 编辑蛋白可为融合蛋白。

[0528] 此处,“融合蛋白”是指通过将酶与额外的结构域、肽、多肽或蛋白融合而产生的蛋白。

[0529] 术语“酶”是指含有能够切割核酸、基因、染色体或蛋白的结构域的蛋白。

[0530] 所述额外的结构域、肽、多肽或蛋白可以是具有与所述酶相同或不同的功能的功能结构域、肽、多肽或蛋白。

[0531] 融合蛋白可以在酶的氨基末端(N末端)或其附近、羧基末端(C末端)或其附近、酶的中间部分及它们的组合中的一个或多个区域处包含额外的结构域、肽、多肽或蛋白。

[0532] 融合蛋白可以在酶的N末端或其附近、C末端或其附近、酶的中间部分及它们的组合中的一个或多个区域处包含功能结构域、肽、多肽或蛋白。

[0533] 引导核酸-编辑蛋白复合体可用于修饰对象物。

[0534] 所述对象物可为靶核酸、基因、染色体或蛋白。

[0535] 例如,引导核酸-编辑蛋白复合体可造成感兴趣的蛋白表达的最终调控(例如抑制、阻遏、降低、升高或促进)、蛋白的去除或新蛋白的表达。

[0536] 此处,引导核酸-编辑蛋白复合体可在DNA、RNA、基因或染色体水平发挥作用。

[0537] 引导核酸-编辑蛋白复合体可在基因转录和翻译阶段发挥作用。

[0538] 引导核酸-编辑蛋白复合体可在蛋白水平发挥作用。

[0539] 1. 引导核酸

[0540] 引导核酸为能够识别靶核酸、基因、染色体或蛋白并形成引导核酸-蛋白复合体的核酸。

[0541] 此处,引导核酸被设置为对引导核酸-蛋白复合体所靶向的核酸、基因、染色体或蛋白进行识别或靶向。

[0542] 引导核酸可以以DNA、RNA或者DNA/RNA混合物的形式存在,并可具有5-150个核酸的序列。

[0543] 引导核酸可以以线性或环状形式存在。

[0544] 引导核酸可为一条连续的核酸序列。

[0545] 例如,所述一条连续的核酸序列可为 $(N)_m$,其中N为A、T、C或G,或为A、U、C或G;m为1-150的整数。

[0546] 引导核酸可为两条以上连续的核酸序列。

[0547] 例如,所述两条以上连续的核酸序列可为 $(N)_m$ 以及 $(N)_o$,其中N代表A、T、C或G,或代表A、U、C或G;m和o为1-150的整数,并且可彼此相同或彼此不同。

[0548] 引导核酸可包含一个或多个结构域。

[0549] 此处,所述结构域可为引导结构域、第一互补结构域、接头结构域、第二互补结构域、近端结构域或者尾部结构域,但不限于此。

[0550] 引导核酸可包含两个以上结构域,所述两个以上结构域可为相同结构域的重复,或可为不同的结构域。

[0551] 结构域将在下文中描述。

[0552] i) 引导结构域

[0553] 术语“引导结构域”是具有能够与靶基因或核酸上的靶序列形成互补结合的互补引导序列的结构域,功能在于与靶基因或核酸特异性相互作用。

[0554] 引导序列为与靶基因或核酸上的靶序列互补的核酸序列,例如具有至少50%、55%、60%、65%、70%、75%、80%、85%、90%或95%或更高的互补性或完全互补性。

[0555] 引导结构域可为5-50个碱基的序列。

[0556] 在实例中,引导结构域可为5-50个碱基、10-50个碱基、15-50个碱基、20-50个碱基、25-50个碱基、30-50个碱基、35-50个碱基、40-50个碱基或45-50个碱基的序列。

[0557] 在另一实例中,引导结构域可为1-5个碱基、5-10个碱基、10-15个碱基、15-20个碱基、20-25个碱基、25-30个碱基、30-35个碱基、35-40个碱基、40-45个碱基或45-50个碱基的序列。

[0558] 引导结构域可具有引导序列。

[0559] 引导序列可为能够与靶基因或核酸上的靶序列形成互补结合的互补碱基序列。

[0560] 引导序列可为与靶基因或核酸上的靶序列互补的核酸序列,例如具有至少70%、75%、80%、85%、90%或95%或更高的互补性或完全互补性。

[0561] 引导序列可为5-50个碱基的序列。

[0562] 在实例中,引导结构域可为5-50个碱基、10-50个碱基、15-50个碱基、20-50个碱基、25-50个碱基、30-50个碱基、35-50个碱基、40-50个碱基或45-50个碱基的序列。

[0563] 在另一实例中,引导序列可为1-5个碱基、5-10个碱基、10-15个碱基、15-20个碱基、20-25个碱基、25-30个碱基、30-35个碱基、35-40个碱基、40-45个碱基或45-50个碱基的

序列。

[0564] 此外,引导结构域可包含引导序列和额外碱基序列。

[0565] 额外碱基序列可用于提高或降低引导结构域的功能。

[0566] 额外碱基序列可用于提高或降低引导序列的功能。

[0567] 额外碱基序列可为1-35个碱基的序列。

[0568] 在一个实例中,额外碱基序列可为5-35个碱基、10-35个碱基、15-35个碱基、20-35个碱基、25-35个碱基或30-35个碱基的序列。

[0569] 在另一实例中,额外碱基序列可为1-5个碱基、5-10个碱基、10-15个碱基、15-20个碱基、20-25个碱基、25-30个碱基或30-35个碱基的序列。

[0570] 额外碱基序列可位于引导序列的5'端。

[0571] 额外碱基序列可位于引导序列的3'端。

[0572] ii) 第一互补结构域

[0573] 术语“第一互补结构域”是包含与第二互补结构域互补的核酸序列的核酸序列,其具有足够的互补性以与第二互补结构域形成双链。

[0574] 第一互补结构域可为5-35个碱基的序列。

[0575] 在实例中,第一互补结构域可为5-35个碱基、10-35个碱基、15-35个碱基、20-35个碱基、25-35个碱基或30-35个碱基的序列。

[0576] 在另一实例中,第一互补结构域可为1-5个碱基、5-10个碱基、10-15个碱基、15-20个碱基、20-25个碱基、25-30个碱基或30-35个碱基的序列。

[0577] iii) 接头结构域

[0578] 术语“接头结构域”是连接两个以上结构域(两个以上相同或不同的结构域)的核酸序列。接头结构域可借助共价键或非共价键与两个以上结构域连接,或可借助共价键或非共价键连接两个以上结构域。

[0579] 接头结构域可为1-30个碱基的序列。

[0580] 在一个实例中,接头结构域可为1-5个碱基、5-10个碱基、10-15个碱基、15-20个碱基、20-25个碱基或25-30个碱基的序列。

[0581] 在另一实例中,接头结构域可为1-30个碱基、5-30个碱基、10-30个碱基、15-30个碱基、20-30个碱基或25-30个碱基的序列。

[0582] iv) 第二互补结构域

[0583] 术语“第二互补结构域”是包含与第一互补结构域互补的核酸序列的核酸序列,其具有足够的互补性以与第一互补结构域形成双链。

[0584] 第二互补结构域可具有与第一互补结构域互补的碱基序列以及与第一互补结构域没有互补性的碱基序列(例如不与第一互补结构域形成双链的碱基序列),并可具有比第一互补结构域更长的碱基序列。

[0585] 第二互补结构域可具有5-35个碱基的序列。

[0586] 在实例中,第二互补结构域可为1-35个碱基、5-35个碱基、10-35个碱基、15-35个碱基、20-35个碱基、25-35个碱基或30-35个碱基的序列。

[0587] 在另一实例中,第二互补结构域可为1-5个碱基、5-10个碱基、10-15个碱基、15-20个碱基、20-25个碱基、25-30个碱基或30-35个碱基的序列。

[0588] v) 近端结构域

[0589] 术语“近端结构域”是指其位置靠近第二互补结构域的核酸序列。

[0590] 近端结构域中可具有互补碱基序列,可基于互补碱基序列形成双链。

[0591] 近端结构域可为1-20个碱基的序列。

[0592] 在一个实例中,近端结构域可为1-20个碱基、5-20个碱基、10-20个碱基或15-20个碱基的序列。

[0593] 在另一实例中,近端结构域可为1-5个碱基、5-10个碱基、10-15个碱基或15-20个碱基的序列。

[0594] vi) 尾部结构域

[0595] 术语“尾部结构域”为位于引导核酸两个末端中的一个或多个末端处的核酸序列。

[0596] 尾部结构域中可具有互补碱基序列,并可基于互补碱基序列形成双链。

[0597] 尾部结构域可为1-50个碱基的序列。

[0598] 在一个实例中,尾部结构域可为5-50个碱基、10-50个碱基、15-50个碱基、20-50个碱基、25-50个碱基、30-50个碱基、35-50个碱基、40-50个碱基或45-50个碱基的序列。

[0599] 在另一实例中,尾部结构域可为1-5个碱基、5-10个碱基、10-15个碱基、15-20个碱基、20-25个碱基、25-30个碱基、30-35个碱基、35-40个碱基、40-45个碱基或45-50个碱基的序列。

[0600] 同时,所述结构域(即引导结构域、第一互补结构域、接头结构域、第二互补结构域、近端结构域和尾部结构域)中包含的部分或全部核酸序列可任选地或额外地包含化学修饰。

[0601] 化学修饰可为但不限于甲基化、乙酰化、磷酸化、硫代磷酸酯连接、锁核酸(LNA)、2'-O-甲基3'硫代磷酸酯(MS)或2'-O-甲基3'硫代PACE(MSP)。

[0602] 引导核酸包含一个或多个结构域。

[0603] 引导核酸可包含引导结构域。

[0604] 引导核酸可包含第一互补结构域。

[0605] 引导核酸可包含接头结构域。

[0606] 引导核酸可包含第二互补结构域。

[0607] 引导核酸可包含近端结构域。

[0608] 引导核酸可包含尾部结构域。

[0609] 此处,可以存在1、2、3、4、5、6个或更多个结构域。

[0610] 引导核酸可包含1、2、3、4、5、6个或更多个引导结构域。

[0611] 引导核酸可包含1、2、3、4、5、6个或更多个第一互补结构域。

[0612] 引导核酸可包含1、2、3、4、5、6个或更多个接头结构域。

[0613] 引导核酸可包含1、2、3、4、5、6个或更多个第二互补结构域。

[0614] 引导核酸可包含1、2、3、4、5、6个或更多个近端结构域。

[0615] 引导核酸可包含1、2、3、4、5、6个或更多个尾部结构域。

[0616] 此处,在引导核酸中,一种类型的结构域可以是重复的。

[0617] 引导核酸可包含具有或不具有重复的数个结构域。

[0618] 引导核酸可包含相同类型的结构域。此处,相同类型的结构域可具有相同的核酸

序列或不同的核酸序列。

[0619] 引导核酸可包含两种类型的结构域。此处,两种不同类型的结构域可具有不同的核酸序列或相同的核酸序列。

[0620] 引导核酸可包含三种类型的结构域。此处,三种不同类型的结构域可具有不同的核酸序列或相同的核酸序列。

[0621] 引导核酸可包含四种类型的结构域。此处,四种不同类型的结构域可具有不同的核酸序列或相同的核酸序列。

[0622] 引导核酸可包含五种类型的结构域。此处,五种不同类型的结构域可具有不同的核酸序列或相同的核酸序列。

[0623] 引导核酸可包含六种类型的结构域。此处,六种不同类型的结构域可具有不同的核酸序列或相同的核酸序列。

[0624] 例如,引导核酸可由[引导结构域]-[第一互补结构域]-[接头结构域]-[第二互补结构域]-[接头结构域]-[引导结构域]-[第一互补结构域]-[接头结构域]-[第二互补结构域]组成。此处,两个引导结构域可包含针对不同或相同靶标的引导序列;两个第一互补结构域和两个第二互补结构域可具有相同或不同的核酸序列。当引导结构域包含针对不同靶标的引导序列时,引导核酸可与两种不同靶标特异性结合;此处,该特异性结合可以同时进行或顺序进行。此外,接头结构域可被特定的酶切割,在特定的酶的存在下,引导核酸可被分为两个或三个部分。

[0625] 作为本说明书中引导核酸的具体实例,引导核酸如下所述。

[0626] gRNA

[0627] 术语“gRNA”是指能够将gRNA-CRISPR酶复合体(即,CRISPR复合体)导向至靶基因或核酸的核酸。此外,gRNA是可结合至CRISPR酶并将CRISPR酶引导至靶基因或核酸的核酸特异性RNA。

[0628] gRNA可包含多个结构域。基于各结构域,相互作用可出现在三维结构或者gRNA的活性形式的链中或者这些链之间。

[0629] gRNA可指单链gRNA(单个RNA分子)或者双链gRNA(包含多于一个RNA分子,通常为两个独立的RNA分子)。

[0630] 在一个示例性实施方式中,单链gRNA从5'至3'方向可包含引导结构域(即,包含能够与靶基因或核酸形成互补结合的引导序列的结构域)、第一互补结构域、接头结构域、第二互补结构域(该结构域具有与第一互补结构域序列互补的序列,因此与第一互补结构域形成双链核酸)、近端结构域以及任选的尾部结构域。

[0631] 在另一实施方式中,双链gRNA可包含第一链和第二链,所述第一链从5'至3'方向可包含引导结构域(即,包含能够与靶基因或核酸形成互补结合的引导序列的结构域)以及第一互补结构域;所述第二链包含第二互补结构域(该结构域具有与第一互补结构域序列互补的序列,因此与第一互补结构域形成双链核酸)、近端结构域以及任选的尾部结构域。

[0632] 此处,第一链可指crRNA,第二链可指tracrRNA。crRNA可包含引导结构域和第一互补结构域;tracrRNA可包含第二互补结构域、近端结构域和任选的尾部结构域。

[0633] 在又一实施方式中,单链gRNA从3'至5'方向可包含引导结构域(即,包含能够与靶基因或核酸形成互补结合的引导序列的结构域)、第一互补结构域和第二互补结构域(该结

构域具有与第一互补结构域序列互补的序列,因此与第一互补结构域形成双链核酸)。

[0634] 引导结构域

[0635] 引导结构域包含能够与靶基因或核酸上的靶序列形成互补结合的互补引导序列。引导序列可为与靶基因或核酸上的靶序列具有互补性的核酸序列,例如具有至少70%、75%、80%、85%、90%或95%或更高的互补性或完全互补性。认为引导结构域允许gRNA-Cas复合体(即CRISPR复合体)与靶基因或核酸特异性相互作用。

[0636] 引导结构域可为5-50个碱基的序列。

[0637] 作为示例性实施方式,引导结构域可为16、17、18、19、20、21、22、23、24或25个碱基的序列。

[0638] 作为示例性实施方式,引导结构域可包含16、17、18、19、20、21、22、23、24或25个碱基的序列。

[0639] 此处,引导结构域可包含引导序列。

[0640] 引导序列可为能够与靶基因或核酸上的靶序列形成互补结合的互补碱基序列。

[0641] 引导序列可为与靶基因或核酸上的靶序列互补的核酸序列,例如具有至少70%、75%、80%、85%、90%或95%或更高的互补性或完全互补性。

[0642] 引导序列可为5-50个碱基的序列。

[0643] 在一个示例性实施方式中,引导序列可为16、17、18、19、20、21、22、23、24或25个碱基的序列。

[0644] 在一个示例性实施方式中,引导序列可包含16、17、18、19、20、21、22、23、24或25个碱基的序列。

[0645] 此处,引导结构域可包含引导序列和额外碱基序列。

[0646] 额外碱基序列可为1-35个碱基的序列。

[0647] 在一个示例性实施方式中,额外碱基序列可为1、2、3、4、5、6、7、8、9或10个碱基的序列。

[0648] 例如,额外碱基序列可为单碱基序列(鸟嘌呤(G))或者两个碱基的序列(GG)。

[0649] 额外碱基序列可位于引导序列的5'端。

[0650] 额外碱基序列可位于引导序列的3'端。

[0651] 任选地,引导结构域的部分或全部碱基序列可包含化学修饰。化学修饰可为甲基化、乙酰化、磷酸化、硫代磷酸酯连接、锁核酸(LNA)、2'-O-甲基3'硫代磷酸酯(MS)或2'-O-甲基3'硫代PACE(MSP),但本发明不限于此。

[0652] 第一互补结构域

[0653] 第一互补结构域包含与第二互补结构域互补的核酸序列,其具有足够的互补性,从而能够与第二互补结构域形成双链。

[0654] 此处,第一互补结构域可为5-35个碱基的序列。第一互补结构域可包含5-35个碱基的序列。

[0655] 在一个示例性实施方式中,第一互补结构域可为5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24或25个碱基的序列。

[0656] 在另一实施方式中,第一互补结构域可包含5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24或25个碱基的序列。

[0657] 第一互补结构域可与天然第一互补结构域具有同源性,或可由天然第一互补结构域衍生而来。此外,第一互补结构域可取决于天然存在的物种而在第一互补结构域的碱基序列中存在差异、可由天然存在的物种中含有的第一互补结构域衍生而来、或可与天然存在的物种中含有的第一互补结构域具有部分或完全同源性。

[0658] 在一个示例性实施方式中,第一互补结构域可与酿脓链球菌、空肠弯曲杆菌、嗜热链球菌、金黄色葡萄球菌或脑膜炎奈瑟菌的第一互补结构域或由它们衍生而来的第一互补结构域具有部分(即至少50%以上)或完全同源性。

[0659] 例如,当第一互补结构域是酿脓链球菌的第一互补结构域或由其衍生而来的第一互补结构域时,第一互补结构域可为5'-GUUUUAGAGCUA-3'或与5'-GUUUUAGAGCUA-3'具有部分(即至少50%以上)或完全同源性的碱基序列。此处,第一互补结构域可进一步包含 $(X)_n$,使得其为5'-GUUUUAGAGCUA $(X)_n$ -3'。 X 可选自于由碱基A、T、U和G所组成的组; n 可表示碱基数,其为5-15的整数。此处, $(X)_n$ 可为相同碱基的 n 个重复,或者为 n 个碱基A、T、U和G的混合。

[0660] 在另一实施方式中,当第一互补结构域为空肠弯曲杆菌的第一互补结构域或由其衍生而来的第一互补结构域时,第一互补结构域可为5'-GUUUUAGUCCCUUUUAAAUUUCUU-3'或与5'-GUUUUAGUCCCUUUUAAAUUUCUU-3'具有部分(即至少50%以上)或完全同源性的碱基序列。此处,第一互补结构域可进一步包含 $(X)_n$,使得其为5'-GUUUUAGUCCCUUUUAAAUUUCUU $(X)_n$ -3'。 X 可选自于由碱基A、T、U和G所组成的组; n 可表示碱基数,其为5-15的整数。此处, $(X)_n$ 可表示相同碱基的 n 个重复,或者表示 n 个碱基A、T、U和G的混合。

[0661] 在另一实施方式中,第一互补结构域可与如下菌的第一互补结构域或由其衍生而来的第一互补结构域具有部分(即至少50%以上)或完全同源性:俭菌(*Parcubacteria bacterium*) (GWC2011_GWC2_44_17)、毛螺菌(*Lachnospiraceae bacterium*) (MC2017)、*Butyrivibrio proteoclasticus*、*Peregrinibacteria bacterium* (GW2011_GWA_33_10)、氨基酸球菌属(*Acidaminococcus* sp.) (BV3L6)、猕猴卟啉单胞菌(*Porphyromonas macacae*)、毛螺菌 (ND2006)、*Porphyromonas crevioricanis*、解糖脲普雷沃菌(*Prevotella disiens*)、*Moraxella bovoculi* (237)、*Smiihella* sp. (SC_K08D17)、稻田钩端螺旋体(*Leptospira inadai*)、毛螺菌 (MA2020)、新凶手弗朗西斯菌(*Francisella novicida*) (U112)、*Candidatus Methanoplasma termitum*或挑剔真杆菌(*Eubacterium eligens*)。

[0662] 例如,当第一互补结构域是俭菌的第一互补结构域或由其衍生而来的第一互补结构域时,第一互补结构域可为5'-UUUGUAGAU-3'或与5'-UUUGUAGAU-3'具有部分(即至少50%以上)同源性的碱基序列。此处,第一互补结构域可进一步包含 $(X)_n$,使得其为5'- $(X)_n$ UUUGUAGAU-3'。 X 可选自于由碱基A、T、U和G所组成的组; n 可表示碱基数,其为1-5的整数。此处, $(X)_n$ 可表示相同碱基的 n 个重复,或者表示 n 个碱基A、T、U和G的混合。

[0663] 任选地,第一互补结构域的部分或全部碱基序列可具有化学修饰。化学修饰可为甲基化、乙酰化、磷酸化、硫代磷酸酯连接、锁核酸(LNA)、2'-O-甲基3'硫代磷酸酯(MS)或2'-O-甲基3'硫代PACE(MSP),但本发明不限于此。

[0664] 接头结构域

[0665] 接头结构域为连接两个以上结构域(连接两个以上相同或不同的结构域)的核酸序列。借助共价或非共价键,接头结构域可与两个以上结构域连接,或可连接两个以上结构域。

[0666] 接头结构域可为使得第一互补结构域与第二互补结构域连接以产生单链gRNA的核酸序列。

[0667] 接头结构域可通过共价或非共价键与第一互补结构域和第二互补结构域连接。

[0668] 接头结构域可通过共价或非共价键连接第一互补结构域和第二互补结构域。

[0669] 接头结构域可为1-30个碱基的序列。接头结构域可包含1-30个碱基的序列。

[0670] 在示例性实施方式中,接头结构域可为1-5个碱基、5-10个碱基、10-15个碱基、15-20个碱基、20-25个碱基或25-30个碱基的序列。

[0671] 在示例性实施方式中,接头结构域可包含1-5个碱基、5-10个碱基、10-15个碱基、15-20个碱基、20-25个碱基或25-30个碱基的序列。

[0672] 接头结构域适合用于单链gRNA分子中,并可用于借助共价或非共价键与双链gRNA的第一链和第二链连接或者连接第一链和第二链来产生单链gRNA。接头结构域可用于借助共价或非共价键与双链gRNA的crRNA和tracrRNA连接或者连接crRNA和tracrRNA来产生单链gRNA。

[0673] 接头结构域可与天然序列(例如tracrRNA的部分序列)具有同源性,或者可由天然序列衍生而来。

[0674] 任选地,接头结构域的部分或全部碱基序列可具有化学修饰。化学修饰可为甲基化、乙酰化、磷酸化、硫代磷酸酯连接、锁核酸(LNA)、2'-O-甲基3'硫代磷酸酯(MS)或2'-O-甲基3'硫代PACE(MSP),但本发明不限于此。

[0675] 第二互补结构域

[0676] 第二互补结构域包含与第一互补结构域互补的核酸序列,并具有足够的互补性以与第一互补结构域形成双链。第二互补结构域可包含与第一互补结构域互补的碱基序列以及与第一互补结构域没有互补性的碱基序列(例如不与第一互补结构域形成双链的碱基序列),并可具有比第一互补结构域更长的碱基序列。

[0677] 此处,第二互补结构域可为5-35个碱基的序列。第二互补结构域可包含5-35个碱基的序列。

[0678] 在示例性实施方式中,第二互补结构域可为5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24或25个碱基的序列。

[0679] 在示例性实施方式中,第二互补结构域可包含5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24或25个碱基的序列。

[0680] 此外,第二互补结构域可与天然第二互补结构域具有同源性,或可由天然第二互补结构域衍生而来。此外,第二互补结构域可取决于天然存在的物种而在第二互补结构域的碱基序列中存在差异、可由天然存在的物种中含有的第二互补结构域衍生而来、或可与天然存在的物种中含有的第二互补结构域具有部分或完全同源性。

[0681] 在示例性实施方式中,第二互补结构域可与酿脓链球菌、空肠弯曲杆菌、嗜热链球菌、金黄色葡萄球菌或脑膜炎奈瑟菌的第二互补结构域或由它们衍生而来的第二互补结构域具有部分(即至少50%以上)或完全同源性。

[0682] 例如,当第二互补结构域是酿脓链球菌的第二互补结构域或由其衍生而来的第二互补结构域时,第二互补结构域可为5'-UAGCAAGUAAAAU-3'或与5'-UAGCAAGUAAAAU-3'具有部分(即至少50%以上)同源性的碱基序列(下划线标出与第一互补结构域形成双链的碱

基序列)。此处,第二互补结构域可进一步包含 $(X)_n$ 和/或 $(X)_m$,使得其为5'- $(X)_n$ UAGCAAGUAAAAU $(X)_m$ -3'。 X 可选自于由碱基A、T、U和G所组成的组; n 和 m 各自可表示碱基数,其中 n 可为1-15的整数, m 可为1-6的整数。此处, $(X)_n$ 可表示相同碱基的 n 个重复,或者表示 n 个碱基A、T、U和G的混合。此外, $(X)_m$ 可表示相同碱基的 m 个重复,或者表示 m 个碱基A、T、U和G的混合。

[0683] 在另一实例中,当第二互补结构域是空肠弯曲杆菌的第二互补结构域或由其衍生而来的第二互补结构域时,第二互补结构域可为5'-AAGAAAUUAAAAAGGGACUAAAAU-3'或与5'-AAGAAAUUAAAAAGGGACUAAAAU-3'具有部分(即至少50%以上)同源性的碱基序列(下划线标出与第一互补结构域形成双链的碱基序列)。此处,第二互补结构域可进一步包含 $(X)_n$ 和/或 $(X)_m$,使得其为5'- $(X)_n$ AAGAAAUUAAAAAGGGACUAAAAU $(X)_m$ -3'。 X 可选自于由碱基A、T、U和G所组成的组; n 和 m 各自可表示碱基数,其中 n 可为1-15的整数, m 可为1-6的整数。此处, $(X)_n$ 可表示相同碱基的 n 个重复,或者表示 n 个碱基A、T、U和G的混合。此外, $(X)_m$ 可表示相同碱基的 m 个重复,或者表示 m 个碱基A、T、U和G的混合。

[0684] 在另一实施方式中,第二互补结构域可与如下菌的第二互补结构域或由其衍生而来的第二互补结构域具有部分(即至少50%以上)或完全同源性:俭菌(GWC2011_GWC2_44_17)、毛螺菌(MC2017)、*Butyrivibrio proteoclasticus*、*Peregrinibacteria bacterium* (GW2011_GWA_33_10)、氨基酸球菌属(BV3L6)、猕猴卟啉单胞菌、毛螺菌(ND2006)、*Porphyromonas crevioricanis*、解糖豚普雷沃菌、*Moraxella bovoculi* (237)、*Smiihella* sp. (SC_K08D17)、稻田钩端螺旋体、毛螺菌(MA2020)、新凶手弗朗西斯菌(U112)、*Candidatus Methanoplasma termitum*或挑剔真杆菌。

[0685] 例如,当第二互补结构域是俭菌的第二互补结构域或由其衍生而来的第二互补结构域时,第二互补结构域可为5'-AAAUUCUACU-3'或与5'-AAAUUCUACU-3'具有部分(即至少50%以上)同源性的碱基序列(下划线标出与第一互补结构域形成双链的碱基序列)。此处,第二互补结构域可进一步包含 $(X)_n$ 和/或 $(X)_m$,使得其为5'- $(X)_n$ AAAUUCUACU $(X)_m$ -3'。 X 可选自于由碱基A、T、U和G所组成的组; n 和 m 各自可表示碱基数,其中 n 可为1-10的整数, m 可为1-6的整数。此处, $(X)_n$ 可表示相同碱基的 n 个重复,或者表示 n 个碱基A、T、U和G的混合。此外, $(X)_m$ 可表示相同碱基的 m 个重复,或者表示 m 个碱基A、T、U和G的混合。

[0686] 任选地,第二互补结构域的部分或全部碱基序列可具有化学修饰。化学修饰可为甲基化、乙酰化、磷酸化、硫代磷酸酯连接、锁核酸(LNA)、2'-O-甲基3'硫代磷酸酯(MS)或2'-O-甲基3'硫代PACE(MSP),但本发明不限于此。

[0687] 近端结构域

[0688] 近端结构域是位于靠近第二互补结构域的位置处的1-20个碱基的序列,该结构域位于第二互补结构域的3'端方向。此处,近端结构域可用于在近端结构域内的互补碱基序列之间形成双链。

[0689] 在一个示例性实施方式中,近端结构域可为5、6、7、8、9、10、11、12、13、14或15个碱基的序列。

[0690] 在另一实施方式中,近端结构域可包含5、6、7、8、9、10、11、12、13、14或15个碱基的序列。

[0691] 此外,近端结构域可与天然近端结构域具有同源性,或可由天然近端结构域衍生

而来。此外,近端结构域可取决于天然存在的物种而在碱基序列中存在差异、可由天然存在的物种中含有的近端结构域衍生而来、或可与天然存在的物种中含有的近端结构域具有部分或完全同源性。

[0692] 在示例性实施方式中,近端结构域可与酿脓链球菌、空肠弯曲杆菌、嗜热链球菌、金黄色葡萄球菌或脑膜炎奈瑟菌的近端结构域或由它们衍生而来的近端结构域具有部分(即至少50%以上)或完全同源性。

[0693] 例如,当近端结构域是酿脓链球菌的近端结构域或其衍生而来的近端结构域时,近端结构域可为5'-AAGGCUAGUCCG-3'或与5'-AAGGCUAGUCCG-3'具有部分(即至少50%以上)同源性的碱基序列。此处,近端结构域可进一步包含 $(X)_n$,使其为5'-AAGGCUAGUCCG $(X)_n$ -3'。X可选自于由碱基A、T、U和G所组成的组;n可表示碱基数,其可为1-15的整数。此处, $(X)_n$ 可表示相同碱基的n个重复,或者表示n个碱基A、T、U和G的混合。

[0694] 在又一实施方式中,当近端结构域是空肠弯曲杆菌的近端结构域或其衍生而来的近端结构域时,近端结构域可为5'-AAAGAGUUUGC-3'或与5'-AAAGAGUUUGC-3'具有至少50%或更高同源性的碱基序列。此处,近端结构域可进一步包含 $(X)_n$,使其为5'-AAAGAGUUUGC $(X)_n$ -3'。X可选自于由碱基A、T、U和G所组成的组;n可表示碱基数,其可为1-40的整数。此处, $(X)_n$ 可表示相同碱基的n个重复,或者表示n个碱基A、T、U和G的混合。

[0695] 任选地,近端结构域的部分或全部碱基序列可具有化学修饰。化学修饰可为甲基化、乙酰化、磷酸化、硫代磷酸酯连接、锁核酸(LNA)、2'-O-甲基3'硫代磷酸酯(MS)或2'-O-甲基3'硫代PACE(MSP),但本发明不限于此。

[0696] 尾部结构域

[0697] 尾部结构域为能被选择性添加至单链gRNA或双链gRNA的3'末端的结构域。尾部结构域可为1-50个碱基的序列,或包含1-50个碱基的序列。此处,尾部结构域可用于在尾部结构域内的互补碱基序列之间形成双链。

[0698] 在示例性实施方式中,尾部结构域可为1-5个碱基、5-10个碱基、10-15个碱基、15-20个碱基、20-25个碱基、25-30个碱基、30-35个碱基、35-40个碱基、40-45个碱基或45-50个碱基的序列。

[0699] 在示例性实施方式中,尾部结构域可包含1-5个碱基、5-10个碱基、10-15个碱基、15-20个碱基、20-25个碱基、25-30个碱基、30-35个碱基、35-40个碱基、40-45个碱基或45-50个碱基的序列。

[0700] 此外,尾部结构域可与天然尾部结构域具有同源性,或可由天然尾部结构域衍生而来。此外,尾部结构域可取决于天然存在的物种而在碱基序列中存在差异、可由天然存在的物种中含有的尾部结构域衍生而来、或可与天然存在的物种中含有的尾部结构域具有部分或完全同源性。

[0701] 在一个示例性实施方式中,尾部结构域可与酿脓链球菌、空肠弯曲杆菌、嗜热链球菌、金黄色葡萄球菌或脑膜炎奈瑟菌的尾部结构域或由它们衍生而来的尾部结构域具有部分(即至少50%以上)或完全同源性。

[0702] 例如,当尾部结构域是酿脓链球菌的尾部结构域或其衍生而来的尾部结构域时,尾部结构域可为5'-UUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC-3'或与5'-UUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC-3'具有部分(即至少50%以上)同源性的碱基序列。此处,尾部结

构域可进一步包含 $(X)_n$, 使其为 5'-UUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC $(X)_n$ -3'。X 可选自由碱基 A、T、U 和 G 所组成的组; n 可表示碱基数, 其可为 1-15 的整数。此处, $(X)_n$ 可表示相同碱基的 n 个重复, 或者表示 n 个碱基 (如 A、T、U 和 G) 的混合。

[0703] 在另一实例中, 当尾部结构域是空肠弯曲杆菌的尾部结构域或由其衍生而来的尾部结构域时, 尾部结构域可为 5'-GGGACUCUGCGGGGUUACAAUCCCCUAAAACCGCUUUU-3' 或与 5'-GGGACUCUGCGGGGUUACAAUCCCCUAAAACCGCUUUU-3' 具有部分 (即至少 50% 以上) 同源性的碱基序列。此处, 尾部结构域可进一步包含 $(X)_n$, 使其为 5'-GGGACUCUGCGGGGUUACAAUCCCCUAAAACCGCUUUU $(X)_n$ -3'。X 可选自由碱基 A、T、U 和 G 所组成的组; n 可表示碱基数, 其可为 1-15 的整数。此处, $(X)_n$ 可表示相同碱基的 n 个重复, 或者表示 n 个碱基 A、T、U 和 G 的混合。

[0704] 在另一实施方式中, 尾部结构域可在 3' 端包含参与体外或体内转录方法的 1-10 个碱基的序列。

[0705] 例如, 当将 T7 启动子用于 gRNA 的体外转录时, 尾部结构域可为存在于 DNA 模板 3' 端的任意碱基序列。此外, 当将 U6 启动子用于体内转录时, 尾部结构域可为 UUUUUU; 当将 H1 启动子用于转录时, 尾部结构域可为 UUUU; 并且当使用 pol-III 启动子时, 尾部结构域可包含数个尿嘧啶碱基或可替代的碱基。

[0706] 任选地, 尾部结构域的部分或全部碱基序列可具有化学修饰。化学修饰可为甲基化、乙酰化、磷酸化、硫代磷酸酯连接、锁核酸 (LNA)、2'-O-甲基 3' 硫代磷酸酯 (MS) 或 2'-O-甲基 3' 硫代 PACE (MSP), 但本发明不限于此。

[0707] gRNA 可包含上文所述的多个结构域, 因此可根据 gRNA 中含有的结构域来调整核酸序列的长度; 基于各结构域, 相互作用可出现在三维结构或者 gRNA 的活性形式的链中或者这些链之间。

[0708] gRNA 可指单链 gRNA (单个 RNA 分子) 或者双链 gRNA (包含多于一个 RNA 分子, 通常为两个独立的 RNA 分子)。

[0709] 双链 gRNA

[0710] 双链 gRNA 由第一链和第二链组成。

[0711] 此处, 第一链可由

[0712] 5'-[引导结构域]-[第一互补结构域]-3' 组成; 以及

[0713] 第二链可由

[0714] 5'-[第二互补结构域]-[近端结构域]-3' 或者

[0715] 5'-[第二互补结构域]-[近端结构域]-[尾部结构域]-3' 组成。

[0716] 此处, 第一链可以指 crRNA, 第二链可以指 tracrRNA。

[0717] 第一链

[0718] [引导结构域]

[0719] 在第一链中, 引导结构域包含能够与靶基因或核酸上的靶序列形成互补结合的互补引导序列。引导序列为与靶基因或核酸上的靶序列互补的核酸序列, 例如具有至少 70%、75%、80%、85%、90% 或 95% 或更高的互补性或完全互补性。认为引导结构域允许 gRNA-Cas 复合体 (即 CRISPR 复合体) 与靶基因或核酸特异性相互作用。

[0720] 此处, 引导结构域可为 5-50 个碱基的序列, 或包含 5-50 个碱基的序列。例如, 引导结构域可为 16、17、18、19、20、21、22、23、24 或 25 个碱基的序列, 或包含 16、17、18、19、20、21、

22、23、24或25个碱基的序列。

[0721] 此外,引导结构域可包含引导序列。

[0722] 此处,引导序列可为能够与靶基因或核酸上的靶序列形成互补结合的互补碱基序列,例如具有至少70%、75%、80%、85%、90%或95%或更高的互补性或完全互补性。

[0723] 此处,引导序列可为5-50个碱基的序列,或包含5-50个碱基的序列。例如,引导序列可为16、17、18、19、20、21、22、23、24或25个碱基的序列,或包含16、17、18、19、20、21、22、23、24或25个碱基的序列。

[0724] 任选地,引导结构域可包含引导序列和额外碱基序列。

[0725] 此处,额外碱基序列可为1-35个碱基的序列。例如,额外碱基序列可为1、2、3、4、5、6、7、8、9或10个碱基的序列。

[0726] 在一个示例性实施方式中,额外碱基序列可包含单个碱基(鸟嘌呤(G))或者两个碱基(GG)。

[0727] 此处,额外碱基序列可位于引导结构域的5'端,或位于引导序列的5'端。

[0728] 额外碱基序列可位于引导结构域的3'端,或位于引导序列的3'端。

[0729] [第一互补结构域]

[0730] 第一互补结构域包含与第二链的第二互补结构域互补的核酸序列,该结构域具有足够的互补性以与第二互补结构域形成双链。

[0731] 此处,第一互补结构域可为5-35个碱基的序列,或包含5-35个碱基的序列。例如,第一互补结构域可为5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24或25个碱基的序列,或包含5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24或25个碱基的序列。

[0732] 第一互补结构域可与天然第一互补结构域具有同源性,或可由天然第一互补结构域衍生而来。此外,第一互补结构域可取决于天然存在的物种而在碱基序列中存在差异、可由天然存在的物种中含有的第一互补结构域衍生而来、或可与天然存在的物种中含有的第一互补结构域具有部分或完全同源性。

[0733] 在一个示例性实施方式中,第一互补结构域可与酿脓链球菌、空肠弯曲杆菌、嗜热链球菌、金黄色葡萄球菌或脑膜炎奈瑟菌的第一互补结构域或由它们衍生而来的第一互补结构域具有部分(即至少50%以上)或完全同源性。

[0734] 任选地,第一互补结构域可含有不与第二链的第二互补结构域进行互补结合的额外碱基序列。

[0735] 此处,额外碱基序列可为1-15个碱基的序列。例如,额外碱基序列可为1-5个碱基、5-10个碱基或10-15个碱基的序列。

[0736] 任选地,引导结构域和/或第一互补结构域的部分或全部碱基序列可具有化学修饰。化学修饰可为甲基化、乙酰化、磷酸化、硫代磷酸酯连接、锁核酸(LNA)、2'-O-甲基3'硫代磷酸酯(MS)或2'-O-甲基3'硫代PACE(MSP),但本发明不限于此。

[0737] 因此,第一链可如上所述由5'-[引导结构域]-[第一互补结构域]-3'构成。

[0738] 此外,第一链可任选地包含额外碱基序列。

[0739] 在一个实例中,第一链可为

[0740] 5'-(N_{靶标})-(Q_m)-3';或者

[0741] $5'-(X)_a-(N_{\text{靶标}})-(X)_b-(Q)_m-(X)_c-3'$ 。

[0742] 此处, $N_{\text{靶标}}$ 是能够与靶基因或核酸上的靶序列形成互补结合的碱基序列, 是可根据靶基因或核酸上的靶序列进行改变的碱基序列区域。

[0743] 此处, $(Q)_m$ 是包含第一互补结构域的碱基序列, 其能够与第二链的第二互补结构域形成互补结合。 $(Q)_m$ 可为与天然存在的物种的第一互补结构域具有部分或完全同源性的序列; 根据来源的物种, 可对第一互补结构域的碱基序列进行改变。 Q 可各自独立地选自由 A、U、C 和 G 所组成的组; m 可为碱基数, 其为 5-35 的整数。

[0744] 例如, 当第一互补结构域与酿脓链球菌的第一互补结构域或由酿脓链球菌衍生而来的第一互补结构域具有部分或完全同源性时, $(Q)_m$ 可为 $5'-GUUUUAGAGCUA-3'$ 或与 $5'-GUUUUAGAGCUA-3'$ 具有至少 50% 或更高同源性的碱基序列。

[0745] 在另一实例中, 当第一互补结构域与空肠弯曲杆菌的第一互补结构域或由空肠弯曲杆菌衍生而来的第一互补结构域具有部分或完全同源性时, $(Q)_m$ 可为 $5'-GUUUUAGUCCCUUUUUAAAUUUCUU-3'$ 或与 $5'-GUUUUAGUCCCUUUUUAAAUUUCUU-3'$ 具有至少 50% 或更高同源性的碱基序列。

[0746] 在又一实例中, 当第一互补结构域与嗜热链球菌的第一互补结构域或由嗜热链球菌衍生而来的第一互补结构域具有部分或完全同源性时, $(Q)_m$ 可为 $5'-GUUUUAGAGCUGUGUUGUUUCG-3'$ 或与 $5'-GUUUUAGAGCUGUGUUGUUUCG-3'$ 具有至少 50% 或更高同源性的碱基序列。

[0747] 此外, $(X)_a$ 、 $(X)_b$ 、 $(X)_c$ 各自为任选的额外碱基序列, 其中 X 可各自独立地选自由 A、U、C 和 G 所组成的组; a 、 b 、 c 各自可为碱基数, 其为 0 或 1-20 的整数。

[0748] 第二链

[0749] 第二链可由第二互补结构域和近端结构域组成, 并可任选地包含尾部结构域。

[0750] [第二互补结构域]

[0751] 在第二链中, 第二互补结构域包含与第一链的第一互补结构域互补的核酸序列, 并具有足够的互补性以与第一互补结构域形成双链。第二互补结构域可包含与第一互补结构域互补的碱基序列以及不与第一互补结构域互补的碱基序列 (例如不与第一互补结构域形成双链的碱基序列), 并可具有比第一互补结构域更长的碱基序列。

[0752] 此处, 第二互补结构域可为 5-35 个碱基的序列, 或者包含 5-35 个碱基的序列。例如, 第二互补结构域可为 5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24 或 25 个碱基的序列, 或包含 5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24 或 25 个碱基的序列, 但本发明不限于此。

[0753] 第二互补结构域可与天然第二互补结构域具有同源性, 或可由天然第二互补结构域衍生而来。此外, 第二互补结构域可取决于天然存在的物种而在第二互补结构域的碱基序列中存在差异、可由天然存在的物种中含有的第二互补结构域衍生而来、或可与天然存在的物种中含有的第二互补结构域具有部分或完全同源性。

[0754] 在一个示例性实施方式中, 第二互补结构域可与酿脓链球菌、空肠弯曲杆菌、嗜热链球菌、金黄色葡萄球菌或脑膜炎奈瑟菌的第二互补结构域或由它们衍生而来的第二互补结构域具有部分 (即至少 50% 以上) 或完全同源性。

[0755] 任选地, 第二互补结构域可进一步包含不与第一链的第一互补结构域进行互补结

合的额外碱基序列。

[0756] 此处,额外碱基序列可为1-25个碱基的序列。例如,额外碱基序列可为1-5个碱基、5-10个碱基、10-15个碱基、15-20个碱基或20-25个碱基的序列。

[0757] [近端结构域]

[0758] 在第二链中,近端结构域为1-20个碱基的序列,该结构域位于第二互补结构域的3'端方向。例如,近端结构域可为5、6、7、8、9、10、11、12、13、14或15个碱基的序列,或包含5、6、7、8、9、10、11、12、13、14或15个碱基的序列。

[0759] 此处,近端结构域可在其互补碱基序列之间具有双链结合。

[0760] 此外,近端结构域可与天然近端结构域具有同源性,或可由天然近端结构域衍生而来。此外,近端结构域可取决于天然存在的物种而在碱基序列中存在差异、可由天然存在的物种的近端结构域衍生而来、或可与天然存在的物种的近端结构域具有部分或完全同源性。

[0761] 在一个示例性实施方式中,近端结构域可与酿脓链球菌、空肠弯曲杆菌、嗜热链球菌、金黄色葡萄球菌或脑膜炎奈瑟菌的近端结构域或由它们衍生而来的近端结构域具有部分(即至少50%以上)或完全同源性。

[0762] [尾部结构域]

[0763] 任选地,在第二链中,尾部结构域可为被选择性添加至第二链的3'端的结构域,尾部结构域可为1-50个碱基的序列,或包含1-50个碱基的序列。例如,尾部结构域可为1-5个碱基、5-10个碱基、10-15个碱基、15-20个碱基、20-25个碱基、25-30个碱基、30-35个碱基、35-40个碱基、40-45个碱基或45-50个碱基的序列,或包含1-5个碱基、5-10个碱基、10-15个碱基、15-20个碱基、20-25个碱基、25-30个碱基、30-35个碱基、35-40个碱基、40-45个碱基或45-50个碱基的序列。

[0764] 此处,尾部结构域可在其互补碱基序列之间具有双链结合。

[0765] 此外,尾部结构域可与天然尾部结构域具有同源性,或可由天然尾部结构域衍生而来。此外,尾部结构域可取决于天然存在的物种而在碱基序列中存在差异、可由天然存在的物种中含有的尾部结构域衍生而来、或可与天然存在的物种中含有的尾部结构域具有部分或完全同源性。

[0766] 在一个示例性实施方式中,尾部结构域可与酿脓链球菌、空肠弯曲杆菌、嗜热链球菌、金黄色葡萄球菌或脑膜炎奈瑟菌的尾部结构域或由它们衍生而来的尾部结构域具有部分(即至少50%以上)或完全同源性。

[0767] 在另一实施方式中,尾部结构域可在3'端包含参与体外或体内转录方法的1-10个碱基的序列。

[0768] 例如,当将T7启动子用于gRNA的体外转录时,尾部结构域可为存在于DNA模板3'端的任意碱基序列。此外,当将U6启动子用于体内转录时,尾部结构域可为UUUUUU;当将H1启动子用于转录时,尾部结构域可为UUUU;并且当使用pol-III启动子时,尾部结构域可包含数个尿嘧啶碱基或可替代的碱基。

[0769] 任选地,第二互补结构域、近端结构域和/或尾部结构域的部分或全部碱基序列可具有化学修饰。化学修饰可为甲基化、乙酰化、磷酸化、硫代磷酸酯连接、锁核酸(LNA)、2'-O-甲基3'硫代磷酸酯(MS)或2'-O-甲基3'硫代PACE(MSP),但本发明不限于此。

[0770] 因此,第二链可如上所述由5'-[第二互补结构域]-[近端结构域]-3'或者5'-[第二互补结构域]-[近端结构域]-[尾部结构域]-3'组成。

[0771] 此外,第二链可任选地包含额外碱基序列。

[0772] 在一个示例性实施方式中,第二链可为5'-(Z)_h-(P)_k-3';或者5'-(X)_d-(Z)_h-(X)_e-(P)_k-(X)_f-3'。

[0773] 在另一实施方式中,第二链可为5'-(Z)_h-(P)_k-(F)_i-3';或者5'-(X)_d-(Z)_h-(X)_e-(P)_k-(X)_f-(F)_i-3'。

[0774] 此处,(Z)_h是包含第二互补结构域的碱基序列,其能够与第一链的第一互补结构域形成互补结合。(Z)_h可为与天然存在的物种的第二互补结构域具有部分或完全同源性的序列;根据来源的物种,可对第二互补结构域的碱基序列进行修饰。Z可各自独立地选自于由A、U、C和G所组成的组;h可为碱基数,其可为5-50的整数。

[0775] 例如,当第二互补结构域与酿脓链球菌的第二互补结构域或由酿脓链球菌衍生而来的第二互补结构域具有部分或完全同源性时,(Z)_h可为5'-UAGCAAGUAAAAU-3'或与5'-UAGCAAGUAAAAU-3'具有至少50%或更高同源性的碱基序列。

[0776] 在另一实例中,当第二互补结构域与空肠弯曲杆菌的第二互补结构域或由空肠弯曲杆菌衍生而来的第二互补结构域具有部分或完全同源性时,(Z)_h可为5'-AAGAAUUUAAAAAGGGACUAAAAU-3'或与5'-AAGAAUUUAAAAAGGGACUAAAAU-3'具有至少50%或更高同源性的碱基序列。

[0777] 在又一实例中,当第二互补结构域与嗜热链球菌的第二互补结构域或由嗜热链球菌衍生而来的第二互补结构域具有部分或完全同源性时,(Z)_h可为5'-CGAAACAACACAGCGAGUAAAAU-3'或与5'-CGAAACAACACAGCGAGUAAAAU-3'具有至少50%或更高同源性的碱基序列。

[0778] (P)_k是包含近端结构域的碱基序列,其可与天然存在的物种的近端结构域具有部分或完全同源性;根据来源的物种,可对近端结构域的碱基序列进行修饰。P可各自独立地选自于由A、U、C和G所组成的组;k可为碱基数,其为1-20的整数。

[0779] 例如,当近端结构域与酿脓链球菌的近端结构域或由酿脓链球菌衍生而来的近端结构域具有部分或完全同源性时,(P)_k可为5'-AAGGCUAGUCCG-3'或与5'-AAGGCUAGUCCG-3'具有至少50%或更高同源性的碱基序列。

[0780] 在另一实例中,当近端结构域与空肠弯曲杆菌的近端结构域或由空肠弯曲杆菌衍生而来的近端结构域具有部分或完全同源性时,(P)_k可为5'-AAAGAGUUUGC-3'或与5'-AAAGAGUUUGC-3'具有至少50%或更高同源性的碱基序列。

[0781] 在又一实例中,当近端结构域与嗜热链球菌的近端结构域或由嗜热链球菌衍生而来的近端结构域具有部分或完全同源性时,(P)_k可为5'-AAGGCUUAGUCCG-3'或与5'-AAGGCUUAGUCCG-3'具有至少50%或更高同源性的碱基序列。

[0782] (F)_i可为包含尾部结构域的碱基序列,其可与天然存在的物种的尾部结构域具有部分或完全同源性;根据来源的物种,可对尾部结构域的碱基序列进行修饰。F可各自独立地选自于由A、U、C和G所组成的组;i可为碱基数,其为1-50的整数。

[0783] 例如,当尾部结构域与酿脓链球菌的尾部结构域或由酿脓链球菌衍生而来的尾部结构域具有部分或完全同源性时,(F)_i可为5'-UUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC-3'

或与5'-UUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC-3'具有至少50%或更高同源性的碱基序列。

[0784] 在另一实例中,当尾部结构域与空肠弯曲杆菌的尾部结构域或由空肠弯曲杆菌衍生而来的尾部结构域具有部分或完全同源性时,(F)_i可为5'-GGGACUCUGCGGGGUUACAAUCCC CUA AAAACCGCUUUU-3'或与5'-GGGACUCUGCGGGGUUACAAUCCCCUA AAAACCGCUUUU-3'具有至少50%或更高同源性的碱基序列。

[0785] 在又一实施方式中,当尾部结构域与嗜热链球菌的尾部结构域或由嗜热链球菌衍生而来的尾部结构域具有部分或完全同源性时,(F)_i可为5'-UACUCAACUUGAAAAGGUGGCACCGAUUCGGUGUUUUU-3'或与5'-UACUCAACUUGAAAAGGUGGCACCGAUUCGGUGUUUUU-3'具有至少50%或更高同源性的碱基序列。

[0786] 此外,(F)_i可在3'端包含参与体外或体内转录方法的1-10个碱基的序列。

[0787] 例如,当将T7启动子用于gRNA的体外转录时,尾部结构域可为存在于DNA模板3'端的任意碱基序列。此外,当将U6启动子用于体内转录时,尾部结构域可为UUUUUU;当将H1启动子用于转录时,尾部结构域可为UUUU;并且当使用pol-III启动子时,尾部结构域可包含数个尿嘧啶碱基或可替代的碱基。

[0788] 此外,(X)_d、(X)_e和(X)_f可为任选添加的碱基序列,其中X可各自独立地选自由A、U、C和G所组成的组;d、e、f各自可为碱基数,其为0或1-20的整数。

[0789] 单链gRNA

[0790] 单链gRNA可分为两种类型。

[0791] i) 单链gRNA

[0792] 首先,存在这样的单链gRNA:在该单链gRNA中,借助接头结构域使得双链gRNA的第一链和第二链连接;此处,单链gRNA由5'-[第一链]-[接头结构域]-[第二链]-3'组成。

[0793] 具体而言,单链gRNA可由

[0794] 5'-[引导结构域]-[第一互补结构域]-[接头结构域]-[第二互补结构域]-[近端结构域]-3';或者

[0795] 5'-[引导结构域]-[第一互补结构域]-[接头结构域]-[第二互补结构域]-[近端结构域]-[尾部结构域]-3'组成。

[0796] 除接头结构域外的各结构域与双链gRNA的第一链和第二链中的各结构域的描述相同。

[0797] -接头结构域

[0798] 在单链gRNA中,接头结构域是连接第一链和第二链的结构域,具体而言,是连接第一互补结构域与第二互补结构域以产生单链gRNA的核酸序列。此处,接头结构域可借助共价或非共价键与第一互补结构域和第二互补结构域连接,或可借助共价或非共价键连接第一互补结构域和第二互补结构域。

[0799] 接头结构域可为1-30个碱基的序列,或包含1-30个碱基的序列。例如,接头结构域可为1-5个碱基、5-10个碱基、10-15个碱基、15-20个碱基、20-25个碱基或者25-30个碱基的序列,或包含1-5个碱基、5-10个碱基、10-15个碱基、15-20个碱基、20-25个碱基或者25-30个碱基的序列。

[0800] 接头结构域适合用于单链gRNA分子中,并可用于借助共价或非共价键与双链gRNA

的第一链和第二链连接或者连接第一链和第二链来产生单链gRNA。接头结构域可用于借助共价或非共价键与双链gRNA的crRNA和tracrRNA连接或者连接crRNA和tracrRNA来产生单链gRNA。

[0801] 接头结构域可与天然序列(例如tracrRNA的部分序列)具有同源性,或者可由天然序列衍生而来。

[0802] 任选地,接头结构域的部分或全部碱基序列可具有化学修饰。化学修饰可为甲基化、乙酰化、磷酸化、硫代磷酸酯连接、锁核酸(LNA)、2'-O-甲基3'硫代磷酸酯(MS)或2'-O-甲基3'硫代PACE(MSP),但本发明不限于此。

[0803] 因此,单链gRNA可如上所述由5'-[引导结构域]-[第一互补结构域]-[接头结构域]-[第二互补结构域]-[近端结构域]-3';或者5'-[引导结构域]-[第一互补结构域]-[接头结构域]-[第二互补结构域]-[近端结构域]-[尾部结构域]-3'组成。

[0804] 此外,单链gRNA可任选包含额外碱基序列。

[0805] 在一个示例性实施方式中,单链gRNA可为

[0806] $5'-(N_{\text{靶标}})-(Q)_m-(L)_j-(Z)_h-(P)_k-3'$;或者

[0807] $5'-(N_{\text{靶标}})-(Q)_m-(L)_j-(Z)_h-(P)_k-(F)_i-3'$ 。

[0808] 在另一实施方式中,单链gRNA可为

[0809] $5'-(X)_a-(N_{\text{靶标}})-(X)_b-(Q)_m-(X)_c-(L)_j-(X)_d-(Z)_h-(X)_e-(P)_k-(X)_f-3'$;或者

[0810] $5'-(X)_a-(N_{\text{靶标}})-(X)_b-(Q)_m-(X)_c-(L)_j-(X)_d-(Z)_h-(X)_e-(P)_k-(X)_f-(F)_i-3'$ 。

[0811] 此处, $N_{\text{靶标}}$ 是能够与靶基因或核酸上的靶序列形成互补结合的碱基序列,是可根据靶基因或核酸上的靶序列进行改变的碱基序列区域。

[0812] $(Q)_m$ 含有包含第一互补结构域的碱基序列,其能够与第二互补结构域形成互补结合。 $(Q)_m$ 可为与天然存在的物种的第一互补结构域具有部分或完全同源性的序列;根据来源的物种,可对第一互补结构域的碱基序列进行改变。 Q 可各自独立地选自由A、U、C和G所组成的组; m 可为碱基数,其可为5-35的整数。

[0813] 例如,当第一互补结构域与酿脓链球菌的第一互补结构域或由酿脓链球菌衍生而来的第一互补结构域具有部分或完全同源性时, $(Q)_m$ 可为5'-GUUUUAGAGCUA-3'或与5'-GUUUUAGAGCUA-3'具有至少50%或更高同源性的碱基序列。

[0814] 在另一实例中,当第一互补结构域与空肠弯曲杆菌的第一互补结构域或由空肠弯曲杆菌衍生而来的第一互补结构域具有部分或完全同源性时, $(Q)_m$ 可为5'-GUUUUAGUCCCUUUUAAAUUUCUU-3'或与5'-GUUUUAGUCCCUUUUAAAUUUCUU-3'具有至少50%或更高同源性的碱基序列。

[0815] 在又一实例中,当第一互补结构域与嗜热链球菌的第一互补结构域或由嗜热链球菌衍生而来的第一互补结构域具有部分或完全同源性时, $(Q)_m$ 可为5'-GUUUUAGAGCUGUGUUGUUUCG-3'或与5'-GUUUUAGAGCUGUGUUGUUUCG-3'具有至少50%或更高同源性的碱基序列。

[0816] 此外, $(L)_j$ 是包含接头结构域的碱基序列,它连接第一互补结构域和第二互补结构域,由此产生单链gRNA。此处, L 可各自独立地选自由A、U、C和G所组成的组; j 可为碱基数,其为1-30的整数。

[0817] $(Z)_h$ 是包含第二互补结构域的碱基序列,其能够与第一互补结构域形成互补结

合。(Z)_h可为与天然存在的物种的第二互补结构域具有部分或完全同源性的序列;根据来源的物种,可对第二互补结构域的碱基序列进行改变。Z可各自独立地选自由A、U、C和G所组成的组;h为碱基数,其可为5-50的整数。

[0818] 例如,当第二互补结构域与酿脓链球菌的第二互补结构域或由酿脓链球菌衍生而来的第二互补结构域具有部分或完全同源性时,(Z)_h可为5'-UAGCAAGUAAAAU-3'或与5'-UAGCAAGUAAAAU-3'具有至少50%或更高同源性的碱基序列。

[0819] 在另一实例中,当第二互补结构域与空肠弯曲杆菌的第二互补结构域或由空肠弯曲杆菌衍生而来的第二互补结构域具有部分或完全同源性时,(Z)_h可为5'-AAGAAUUUAAAAAGGGACUAAAAU-3'或与5'-AAGAAUUUAAAAAGGGACUAAAAU-3'具有至少50%或更高同源性的碱基序列。

[0820] 在又一实例中,当第二互补结构域与嗜热链球菌的第二互补结构域或由嗜热链球菌衍生而来的第二互补结构域具有部分或完全同源性时,(Z)_h可为5'-CGAAACAACACAGCGAGUAAAAU-3'或与5'-CGAAACAACACAGCGAGUAAAAU-3'具有至少50%或更高同源性的碱基序列。

[0821] (P)_k是包含近端结构域的碱基序列,其可与天然存在的物种的近端结构域具有部分或完全同源性;根据来源的物种,可对近端结构域的碱基序列进行修饰。P可各自独立地选自由A、U、C和G所组成的组;k可为碱基数,其为1-20的整数。

[0822] 例如,当近端结构域与酿脓链球菌的近端结构域或由酿脓链球菌衍生而来的近端结构域具有部分或完全同源性时,(P)_k可为5'-AAGGCUAGUCCG-3'或与5'-AAGGCUAGUCCG-3'具有至少50%或更高同源性的碱基序列。

[0823] 在另一实例中,当近端结构域与空肠弯曲杆菌的近端结构域或由空肠弯曲杆菌衍生而来的近端结构域具有部分或完全同源性时,(P)_k可为5'-AAAGAGUUUGC-3'或与5'-AAAGAGUUUGC-3'具有至少50%或更高同源性的碱基序列。

[0824] 在又一实例中,当近端结构域与嗜热链球菌的近端结构域或由嗜热链球菌衍生而来的近端结构域具有部分或完全同源性时,(P)_k可为5'-AAGGCUAGUCCG-3'或与5'-AAGGCUAGUCCG-3'具有至少50%或更高同源性的碱基序列。

[0825] (F)_i可以是包含尾部结构域的碱基序列,其可与天然存在的物种的尾部结构域具有部分或完全同源性;根据来源的物种,可对尾部结构域的碱基序列进行修饰。F可各自独立地选自由A、U、C和G所组成的组;i可为碱基数,其为1-50的整数。

[0826] 例如,当尾部结构域与酿脓链球菌的尾部结构域或由酿脓链球菌衍生而来的尾部结构域具有部分或完全同源性时,(F)_i可为5'-UUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC-3'或与5'-UUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC-3'具有至少50%或更高同源性的碱基序列。

[0827] 在另一实例中,当尾部结构域与空肠弯曲杆菌的尾部结构域或由空肠弯曲杆菌衍生而来的尾部结构域具有部分或完全同源性时,(F)_i可为5'-GGGACUCUGCGGGGUUACAAUCCCUAAAAACCGCUUUU-3'或与5'-GGGACUCUGCGGGGUUACAAUCCCUAAAAACCGCUUUU-3'具有至少50%或更高同源性的碱基序列。

[0828] 在又一实例中,当尾部结构域与嗜热链球菌的尾部结构域或由嗜热链球菌衍生而来的尾部结构域具有部分或完全同源性时,(F)_i可为5'-UACUCAACUUGAAAAAGUGGCACCGAUU

CGGUGUUUUU-3'或与5'-UACUCAACUUGAAAAGGUGGCACCGAUUCGGUGUUUUU-3'具有至少50%或更高同源性的碱基序列。

[0829] 此外, (F)_i可在3'端包含参与体外或体内转录方法的1-10个碱基的序列。

[0830] 例如, 当将T7启动子用于gRNA的体外转录时, 尾部结构域可为存在于DNA模板3'端的任意碱基序列。此外, 当将U6启动子用于体内转录时, 尾部结构域可为UUUUUU; 当将H1启动子用于转录时, 尾部结构域可为UUUU; 并且当使用pol-III启动子时, 尾部结构域可包含数个尿嘧啶碱基或可替代的碱基。

[0831] 此外, (X)_a、(X)_b、(X)_c、(X)_d、(X)_e和(X)_f可为任选添加的碱基序列, 其中X可各自独立地选自由A、U、C和G所组成的组; a、b、c、d、e、f各自可为碱基数, 其为0或1-20的整数。

[0832] ii) 单链gRNA

[0833] 其次, 单链gRNA可为由引导结构域、第一互补结构域和第二互补结构域组成的单链gRNA, 此处, gRNA可由

[0834] 5'-[第二互补结构域]-[第一互补结构域]-[引导结构域]-3'; 或者

[0835] 5'-[第二互补结构域]-[接头结构域]-[第一互补结构域]-[引导结构域]-3'组成。

[0836] -引导结构域

[0837] 在单链gRNA中, 引导结构域包含能够与靶基因或核酸上的靶序列形成互补结合的互补引导序列。引导序列可为与靶基因或核酸上的靶序列具有互补性的核酸序列, 例如具有至少70%、75%、80%、85%、90%或95%或更高的互补性或完全互补性。认为引导结构域允许gRNA-Cas复合体(即CRISPR复合体)与靶基因或核酸特异性相互作用。

[0838] 此处, 引导结构域可为5-50个碱基的序列, 或包含5-50个碱基的序列。例如, 引导结构域可为16、17、18、19、20、21、22、23、24或25个碱基的序列, 或包含16、17、18、19、20、21、22、23、24或25个碱基的序列。

[0839] 此外, 引导结构域可包含引导序列。

[0840] 此处, 引导序列可为能够与靶基因或核酸上的靶序列形成互补结合的互补碱基序列, 例如具有至少70%、75%、80%、85%、90%或95%或更高的互补性或完全互补性。

[0841] 此处, 引导序列可为5-50个碱基的序列, 或包含5-50个碱基的序列。例如, 引导序列可为16、17、18、19、20、21、22、23、24或25个碱基的序列, 或包含16、17、18、19、20、21、22、23、24或25个碱基的序列。

[0842] 任选地, 引导结构域可包含引导序列和额外碱基序列。

[0843] 此处, 额外碱基序列可为1-35个碱基的序列。例如, 额外碱基序列可为1、2、3、4、5、6、7、8、9或10个碱基的序列。

[0844] 在一个示例性实施方式中, 额外碱基序列可为单个碱基的序列(鸟嘌呤(G))或者两个碱基的序列(GG)。

[0845] 此处, 额外碱基序列可位于引导结构域的5'端, 或位于引导序列的5'端。

[0846] 额外碱基序列可位于引导结构域的3'端, 或位于引导序列的3'端。

[0847] -第一互补结构域

[0848] 第一互补结构域为包含与第二互补结构域互补的核酸序列的结构域, 该结构域具有足够的互补性以与第二互补结构域形成双链。

[0849] 此处,第一互补结构域可为5-35个碱基的序列,或包含5-35个碱基的序列。例如,第一互补结构域可为5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24或25个碱基的序列,或包含5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24或25个碱基的序列。

[0850] 第一互补结构域可与天然第一互补结构域具有同源性,或可由天然第一互补结构域衍生而来。此外,第一互补结构域可取决于天然存在的物种而在第一互补结构域的碱基序列中存在差异、可由天然存在的物种中含有的第一互补结构域衍生而来、或可与天然存在的物种中含有的第一互补结构域具有部分或完全同源性。

[0851] 在一个示例性实施方式中,第一互补结构域可与如下菌的第一互补结构域或由其衍生而来的第一互补结构域具有部分(即至少50%以上)或完全同源性:俭菌(GWC2011_GWC2_44_17)、毛螺菌(MC2017)、*Butyrivibrio proteoclasticus*、*Peregrinibacterium* (GW2011_GWA_33_10)、氨基酸球菌属(BV3L6)、猕猴卟啉单胞菌、毛螺菌(ND2006)、*Porphyromonas crevioricanis*、解糖脲普雷沃菌、*Moraxella bovoculi* (237)、*Smiihella* sp. (SC_K08D17)、稻田钩端螺旋体、毛螺菌(MA2020)、新凶手弗朗西斯菌(U112)、*Candidatus Methanoplasma termitum*或挑剔真杆菌。

[0852] 任选地,第一互补结构域可包含不与第二互补结构域进行互补结合的额外碱基序列。

[0853] 此处,额外碱基序列可为1-15个碱基的序列。例如,额外碱基序列可为1-5个碱基、5-10个碱基或10-15个碱基的序列。

[0854] -第二互补结构域

[0855] 第二互补结构域包含与第一互补结构域互补的核酸序列,并具有足够的互补性以与第一互补结构域形成双链。第二互补结构域可包含与第一互补结构域互补的碱基序列以及不与第一互补结构域没有互补性的碱基序列(例如不与第一互补结构域形成双链的碱基序列),并可具有比第一互补结构域更长的碱基序列。

[0856] 此处,第二互补结构域可为5-35个碱基的序列,或包含5-35个碱基的序列。例如,第二互补结构域可为5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24或25个碱基的序列,或包含5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24或25个碱基的序列。

[0857] 第二互补结构域可与天然第二互补结构域具有同源性,或可由天然第二互补结构域衍生而来。此外,第二互补结构域可取决于天然存在的物种而在第二互补结构域的碱基序列中存在差异、可由天然存在的物种中含有的第二互补结构域衍生而来、或可与天然存在的物种中含有的第二互补结构域具有部分或完全同源性。

[0858] 在一个示例性实施方式中,第二互补结构域可与如下菌的第二互补结构域或由其衍生而来的第二互补结构域具有部分(即至少50%以上)或完全同源性:俭菌(GWC2011_GWC2_44_17)、毛螺菌(MC2017)、*Butyrivibrio proteoclasticus*、*Peregrinibacterium* (GW2011_GWA_33_10)、氨基酸球菌属(BV3L6)、猕猴卟啉单胞菌、毛螺菌(ND2006)、*Porphyromonas crevioricanis*、解糖脲普雷沃菌、*Moraxella bovoculi* (237)、*Smiihella* sp. (SC_K08D17)、稻田钩端螺旋体、毛螺菌(MA2020)、新凶手弗朗西斯菌(U112)、*Candidatus Methanoplasma termitum*或挑剔真杆菌。

[0859] 任选地,第二互补结构域可包含不与第一互补结构域进行互补结合的额外碱基序列。

[0860] 此处,额外碱基序列可为1-15个碱基的序列。例如,额外碱基序列可为1-5个碱基、5-10个碱基或10-15个碱基的序列。

[0861] -接头结构域

[0862] 任选地,接头结构域是连接第一互补结构域和第二互补结构域以产生单链gRNA的核酸序列。此处,接头结构域可借助共价或非共价键与第一互补结构域和第二互补结构域连接,或可借助共价或非共价键连接第一互补结构域和第二互补结构域。

[0863] 接头结构域可为1-30个碱基的序列,或包含1-30个碱基的序列。例如,接头结构域可为1-5个碱基、5-10个碱基、10-15个碱基、15-20个碱基、20-25个碱基或者25-30个碱基的序列,或包含1-5个碱基、5-10个碱基、10-15个碱基、15-20个碱基、20-25个碱基或者25-30个碱基的序列。

[0864] 任选地,引导结构域、第一互补结构域、第二互补结构域和接头结构域的部分或全部碱基序列可具有化学修饰。化学修饰可为甲基化、乙酰化、磷酸化、硫代磷酸酯连接、锁核酸(LNA)、2'-O-甲基3'硫代磷酸酯(MS)或2'-O-甲基3'硫代PACE(MSP),但本发明不限于此。

[0865] 因此,单链gRNA可如上所述由5'-[第二互补结构域]-[第一互补结构域]-[引导结构域]-3';或者5'-[第二互补结构域]-[接头结构域]-[第一互补结构域]-[引导结构域]-3'组成。

[0866] 此外,单链gRNA可任选地包含额外碱基序列。

[0867] 在一个示例性实施方式中,单链gRNA可为

[0868] $5'-(Z)_h-(Q)_m-(N_{\text{靶标}})-3'$;或者

[0869] $5'-(X)_a-(Z)_h-(X)_b-(Q)_m-(X)_c-(N_{\text{靶标}})-3'$ 。

[0870] 在另一实施方式中,单链gRNA可为

[0871] $5'-(Z)_h-(L)_j-(Q)_m-(N_{\text{靶标}})-3'$;或者

[0872] $5'-(X)_a-(Z)_h-(L)_j-(Q)_m-(X)_c-(N_{\text{靶标}})-3'$ 。

[0873] 此处, $N_{\text{靶标}}$ 是能够与靶基因或核酸上的靶序列形成互补结合的碱基序列,是可根据靶基因或核酸上的靶序列进行改变的碱基序列区域。

[0874] $(Q)_m$ 是包含第一互补结构域的碱基序列,其能够与第二链的第二互补结构域形成互补结合。 $(Q)_m$ 可为与天然存在的物种的第一互补结构域具有部分或完全同源性的序列;根据来源的物种,可对第一互补结构域的碱基序列进行改变。 Q 可各自独立地选自由A、U、C和G所组成的组; m 可为碱基数,其可为5-35的整数。

[0875] 例如,当第一互补结构域与俭菌的第一互补结构域或由其衍生而来的第一互补结构域具有部分或完全同源性时, $(Q)_m$ 可为5'-UUUGUAGAU-3'或与5'-UUUGUAGAU-3'具有至少50%或更高同源性的碱基序列。

[0876] $(Z)_h$ 是包含第二互补结构域的碱基序列,其能够与第一链的第一互补结构域形成互补结合。 $(Z)_h$ 可为与天然存在的物种的第二互补结构域具有部分或完全同源性的序列;根据来源的物种,可对第二互补结构域的碱基序列进行修饰。 Z 可各自独立地选自由A、U、C和G所组成的组; h 可为碱基数,其为5-50的整数。

[0877] 例如,当第二互补结构域与俭菌的第二互补结构域或由俭菌衍生而来的第二互补

结构域具有部分或完全同源性时, $(Z)_h$ 可为5'-AAAUUUCUACU-3'或与5'-AAAUUUCUACU-3'具有至少50%或更高同源性的碱基序列。

[0878] 此外, $(L)_j$ 是包含接头结构域的碱基序列, 它连接第一互补结构域和第二互补结构域。此处, L可各自独立地选自于由A、U、C和G所组成的组; j可为碱基数, 其为1-30的整数。

[0879] 此外, $(X)_a$ 、 $(X)_b$ 和 $(X)_c$ 各自为任选的额外碱基序列, 其中X可各自独立地选自于由A、U、C和G所组成的组; a、b和c可为碱基数, 其为0或1-20的整数。

[0880] 2. 编辑蛋白

[0881] 编辑蛋白是指能够与核酸直接结合或无需直接结合而与核酸相互作用的肽、多肽或蛋白。从概念上讲, 编辑蛋白有时也指“基因剪刀”或RGEN(RNA引导的核酸内切酶)。

[0882] 核酸可为靶核酸、基因或染色体中含有的核酸。

[0883] 核酸可为引导核酸。

[0884] 编辑蛋白可为酶。

[0885] 编辑蛋白可为融合蛋白。

[0886] 此处, 融合蛋白是指通过将酶与额外的结构域、肽、多肽或蛋白融合而产生的蛋白。

[0887] 酶是指含有能够切割核酸、基因、染色体或蛋白的结构域的蛋白。

[0888] 酶可为核酸酶、蛋白酶或限制性酶。

[0889] 所述额外的结构域、肽、多肽或蛋白可以是具有与所述酶相同或不同的功能的功能结构域、肽、多肽或蛋白。

[0890] 融合蛋白可以在酶的N末端或其附近、酶的C末端或其附近、酶的中间部分及它们的组合中的一处或多处包含额外的结构域、肽、多肽或蛋白。

[0891] 融合蛋白可以在酶的N末端或其附近、酶的C末端或其附近、酶的中间部分及它们的组合中的一处或多处包含功能结构域、肽、多肽或蛋白。

[0892] 此处, 功能结构域、肽、多肽或蛋白可为具有甲基化酶活性、去甲基化酶活性、转录激活活性、转录阻遏活性、转录释放因子活性、组蛋白修饰活性、RNA切割活性或核酸结合活性的结构域、肽、多肽或蛋白, 或者为用于纯化和分离蛋白(包括肽)的标签或报告基因, 但本发明不限于此。

[0893] 功能结构域、肽、多肽或蛋白可为脱氨酶。

[0894] 标签包括组氨酸(His)标签、V5标签、FLAG标签、流感血凝素(HA)标签、Myc标签、VSV-G标签和硫氧还蛋白(Trx)标签; 报告基因包括谷胱甘肽硫转移酶(GST)、辣根过氧化物酶(HRP)、氯霉素乙酰转移酶(CAT)、 β -半乳糖苷酶、 β -葡萄糖醛酸酶、萤光素酶、自发荧光蛋白(包括绿色荧光蛋白(GFP)、HcRed、DsRed、青色荧光蛋白(CFP)、黄色荧光蛋白(YFP)和蓝色荧光蛋白(BFP)), 但本发明不限于此。

[0895] 此外, 功能结构域、肽、多肽或蛋白可为核定位序列或信号(NLS)或者核输出序列或信号(nuclear export sequence or signal, NES)。

[0896] NLS可为: 具有氨基酸序列PKKKRKV的SV40病毒大T抗原的NLS; 由核质蛋白衍生而来的NLS(例如具有序列KRPAATKKAGQAKKKK的双分型核质蛋白(nucleoplasmin bipartite) NLS); 具有氨基酸序列PAAKRVKLD或RQRRNELKRSP的c-myc NLS; 具有序列NQSSNFGPMKGGNFGGRSSGPYGGGGQYFAKPRNQGGY的hRNPA1M9NLS; 由输入蛋白 α (importin- α)衍生而来的IBB结构

域序列RMRIZFKNKGKDTAELRRRRVEVSVELRKAKKDEQILKRRNV;肌瘤T蛋白序列VSRKRPRP和PPKKARED;人p53序列POPKKKPL;小鼠c-ab1 IV序列SALIKKKKKMAP;流感病毒NS1序列DRLRR和PKQKKRK;肝炎病毒δ抗原序列RKLKKKIKKL;小鼠Mx1蛋白序列REKKKFLKRR;人多聚(ADP-核糖)聚合酶序列KRKGDEVDGVDEVAKKSKK;或者类固醇激素受体(人)糖皮质激素序列RKCLQAGMNLEARKTKK,但本发明不限于此。

[0897] 编辑蛋白可包括具有完全活性的酶。

[0898] 此处,“具有完全活性的酶”是指具有与野生型酶的功能相同的功能的酶,例如,切割双链DNA的野生型酶具有将双链DNA全部切割的完全酶活性。

[0899] 此外,具有完全活性的酶包括与野生型酶的功能相比具有改善的功能的酶,例如,切割双链DNA的野生型酶的特定修饰或经操纵的形式具有与野生型酶(即切割双链DNA的活性)相比改善的完全酶活性。

[0900] 编辑蛋白可包括具有不完全或部分活性的酶。

[0901] 此处,“具有不完全或部分活性的酶”是指具有野生型酶的部分功能的酶,例如,野生型酶(切割双链DNA)的特定修饰或经操纵的形式具有切割双链一部分(即单链DNA)的不完全或部分的酶活性。

[0902] 编辑蛋白可包括失活的酶。

[0903] 此处,“失活的酶”是指其中野生型酶的功能完全失活的酶。例如,切割双链DNA的野生型酶的特定修饰或经操纵的形式没有活性,从而完全不切割双链DNA。

[0904] 编辑蛋白可为天然的酶或融合蛋白。

[0905] 编辑蛋白可以以部分修饰的天然酶或融合蛋白的形式存在。

[0906] 编辑蛋白可为在天然状态下不存在的人工产生的酶或融合蛋白。

[0907] 编辑蛋白可以以在天然状态下不存在的部分修饰的人工酶或融合蛋白的形式存在。

[0908] 此处,修饰可为对编辑蛋白中含有的氨基酸进行置换、删除、添加或上述修饰的组合。

[0909] 此外,修饰可为对编码编辑蛋白的碱基序列中的部分碱基进行置换、删除、添加或上述修饰的组合。

[0910] 作为本发明编辑蛋白的一个示例性实施方式,将在下文描述CRISPR酶。

[0911] CRISPR酶

[0912] 术语“CRISPR酶”是CRISPR-Cas系统的主要蛋白成分,与gRNA形成复合体,从而产生CRISPR-Cas系统。

[0913] CRISPR酶是具有编码CRISPR酶的序列的核酸或多肽(或蛋白),典型地,II型CRISPR酶或V型CRISPR酶已得到广泛使用。

[0914] II型CRISPR酶为Cas9,Cas9可由多种微生物衍生而来,例如为放线菌(Actinobacteria)Cas9(例如内氏放线菌(*Actinomyces naeslundii*)Cas9);产水菌门(Aquificae)Cas9;拟杆菌(Bacteroidetes)Cas9;衣原体(Chlamydiae)Cas9;绿弯菌门(Chloroflexi)Cas9;蓝细菌(Cyanobacteria)Cas9;迷踪菌门(Elusimicrobia)Cas9;纤维杆菌门(Fibrobacteres)Cas9;厚壁菌门(Firmicutes)Cas9(例如酿脓链球菌Cas9、嗜热链球菌Cas9、无害李斯特菌(*Listeria innocua*)Cas9、无乳链球菌(*Streptococcus*

agalactiae) Cas9、变形链球菌 (*Streptococcus mutans*) Cas9 和屎肠球菌 (*Enterococcus faecium*) Cas9) ; 梭杆菌门 (*Fusobacteria*) Cas9; 变形菌门 (*Proteobacteria*) Cas9 (例如脑膜炎奈瑟菌 Cas9、空肠弯曲杆菌 Cas9) 和螺旋体 (*Spirochaetes*) Cas9 (例如密螺旋体 (*Treponema denticola*) Cas9)。

[0915] 术语“Cas9”是与gRNA结合从而切割或修饰靶基因或核酸上的靶序列或位置的酶, Cas9可由HNH结构域(能够切割与gRNA形成互补结合的核酸链)、RuvC结构域(能够切割与gRNA形成互补结合的核酸链)、REC结构域(识别靶标)以及PI结构域(识别PAM)组成。对于Cas9的具体结构特征,可参考Hiroshi Nishimasu等(2014)Cell 156:935-949。

[0916] 此外,V型CRISPR酶可为Cpf1,Cpf1可由如下微生物衍生而来:链球菌 (*Streptococcus*)、弯曲杆菌 (*Campylobacter*)、Nitratifractor、葡萄球菌 (*Staphylococcus*)、Parvibaculum、罗斯氏菌 (*Roseburia*)、奈瑟菌 (*Neisseria*)、葡萄糖醋杆菌 (*Gluconacetobacter*)、固氮螺菌 (*Azospirillum*)、Sphaerochaeta、乳杆菌 (*Lactobacillus*)、真杆菌 (*Eubacterium*)、棒状杆菌 (*Corynebacter*)、肉食杆菌 (*Carnobacterium*)、红细菌 (*Rhodobacter*)、李斯特菌 (*Listeria*)、Paludibacter、梭菌 (*Clostridium*)、毛螺菌、Clostridiaridium、纤毛菌 (*Leptotrichia*)、弗朗西斯氏菌属 (*Francisella*)、军团杆菌 (*Legionella*)、脂环酸芽孢杆菌 (*Alicyclobacillus*)、Methanomethyophilus、卟啉单胞菌 (*Porphyromonas*)、普雷沃菌 (*Prevotella*)、拟杆菌、创伤球菌 (*Helcococcus*)、钩端螺旋体 (*Letospira*)、脱硫弧菌 (*Desulfovibrio*)、Desulfonatronum、丰佑菌 (*Opitutaceae*)、肿块芽孢杆菌 (*Tuberibacillus*)、芽孢杆菌 (*Bacillus*)、短芽孢杆菌 (*Brevibacillus*)、甲基杆菌 (*Methylobacterium*) 或氨基酸球菌。

[0917] Cpf1可由RuvC结构域(类似于并对应于Cas9的RuvC结构域)、Nuc结构域(而不是Cas9的HNH结构域)、REC结构域和WED结构域(识别靶标)以及PI结构域(识别PAM)组成。对于Cpf1的具体结构特征,可参见Takashi Yamano等(2016)Cell 165:949-962。

[0918] 可从天然存在的微生物中分离或者通过重组或合成方法非天然地产生CRISPR酶(Cas9或Cpf1蛋白)。

[0919] II型CRISPR酶

[0920] 根据对两种以上类型的天然微生物II型CRISPR酶分子的研究(Jinek等,Science, 343(6176):1247997,2014)以及对酿脓链球菌Cas9(SpCas9)与gRNA复合的研究(Nishimasu等,Cell,156:935-949,2014;以及Anders等,Nature,2014,doi:10.1038/nature13579)确定了II型CRISPR酶的晶体结构。

[0921] II型CRISPR酶包含两个叶(lobes),即识别(REC)叶和核酸酶(NUC)叶,各叶包含数个结构域。

[0922] REC叶包含富精氨酸的螺旋桥(BH)结构域、REC1结构域和REC2结构域。

[0923] 此处,BH结构域为长 α 螺旋并且为富精氨酸区域,而REC1和REC2结构域在识别gRNA(例如单链gRNA、双链gRNA或tracrRNA)中形成的双链中起到重要作用。

[0924] NUC叶包含RuvC结构域、HNH结构域和PAM相互作用(PI)结构域。此处,RuvC结构域包括RuvC样结构域,HNH结构域则用于包括HNH样结构域。

[0925] 此处,RuvC结构域与含有II型CRISPR酶的天然存在的微生物家族的成员共享结构相似性,并切割单链(例如靶基因或核酸的非互补链,即不与gRNA形成互补结合的链)。在本

领域中,RuvC结构域有时指RuvCI结构域、RuvCII结构域或RuvCIII结构域,一般称为RuvCI、RuvCII或RuvCIII。例如,在SpCas9的情况下,RuvC结构域由分别位于SpCas9氨基酸序列第1-59位、第718-769位和第909-1098位的三个分离的RuvC结构域(RuvCI、RuvCII和RuvCIII)组装而来。

[0926] HNH结构域与HNH核酸内切酶共享结构相似性,并切割单链(例如靶核酸分子的互补链,即与gRNA形成互补结合的链)。HNH结构域位于RuvCII和RuvCIII基序之间。例如,在SpCas9的情况下,HNH结构域位于SpCas9氨基酸序列的第775-908位。

[0927] PI结构域识别靶基因或核酸中的特定碱基序列(即原型间隔区邻近基序(PAM))或与PAM相互作用。例如,在SpCas9的情况下,PI结构域位于SpCas9氨基酸序列的第1099-1368位。

[0928] 此处,PAM可根据II型CRISPR酶的来源改变。例如,当CRISPR酶是SpCas9时,PAM可为5'-NGG-3';当CRISPR酶是嗜热链球菌Cas9(StCas9)时,PAM可为5'-NNAGAAW-3'(W=A或T);当CRISPR酶是脑膜炎奈瑟菌Cas9(NmCas9)时,PAM可为5'-NNNGATT-3';当CRISPR酶是空肠弯曲杆菌Cas9(CjCas9)时,PAM可为5'-NNNVRYAC-3'(V=G或C或A,R=A或G,Y=C或T),其中,N可为A、T、G或C;或者A、U、G或C。

[0929] V型CRISPR酶

[0930] V型CRISPR酶包含类似的RuvC结构域(对应于II型CRISPR酶的RuvC结构域),并可由Nuc结构域(而不是II型CRISPR酶的HNH结构域)、REC结构域和WED结构域(识别靶标)以及PI结构域(识别PAM)组成。对于V型CRISPR酶的具体结构特征,可参见Takashi Yamano等(2016)Cell 165:949-962。

[0931] V型CRISPR酶可与gRNA相互作用,从而形成gRNA-CRISPR酶复合体,即CRISPR复合体,并且可在gRNA的协作下允许引导序列接近包含PAM序列的靶序列。此处,V型CRISPR酶与靶基因或核酸相互作用的能力依赖于PAM序列。

[0932] PAM序列是存在于靶基因或核酸中的序列,其可被V型CRISPR酶的PI结构域识别。PAM序列可根据V型CRISPR酶的来源改变。即,取决于物种,存在能够被特异性识别的不同PAM序列。

[0933] 在一个实例中,由Cpf1识别的PAM序列可为5'-TTN-3'(N为A、T、C或G)。

[0934] CRISPR酶活性

[0935] CRISPR酶切割靶基因或核酸的双链或单链,并具有造成双链或单链断裂或缺失的核酸酶活性。通常,野生型II型CRISPR酶或V型CRISPR酶切割靶基因或核酸的双链。

[0936] 为了对CRISPR酶的上述核酸酶活性进行操纵或修饰,可对CRISPR酶进行操纵或修饰,例如可以对CRISPR酶进行操纵或修饰来使其改变为具有不完全或部分活性、或者失活的酶。

[0937] 具有不完全或部分活性的酶

[0938] 经修饰而改变其酶活性从而表现出不完全或部分活性的CRISPR酶称为切口酶(nickase)。

[0939] 术语“切口酶”是指经操纵或修饰而仅切割靶基因或核酸双链中的一条链的CRISPR酶,切口酶具有切割单链(例如不与靶基因或核酸的gRNA互补的链或与其互补的链)的核酸酶活性。因此,为了切割双链需要两种切口酶的核酸酶活性。

[0940] 例如,切口酶可具有RuvC结构域的核酸酶活性。即,切口酶可不包含HNH结构域的核酸酶活性,为此可对HNH结构域进行操纵或修饰。

[0941] 在一个实例中,在CRISPR酶是II型CRISPR酶的情况下,当将SpCas9氨基酸序列中的第840位残基由组氨酸突变为丙氨酸时,HNH结构域的核酸酶活性失活而用作切口酶。由于由此产生的切口酶具有RuvC结构域的核酸酶活性,它能够切割靶基因或核酸的非互补链,即不与gRNA形成互补结合的链。

[0942] 在另一示例性实施方式中,当将CjCas9氨基酸序列中的第559位残基由组氨酸突变为丙氨酸时,HNH结构域的核酸酶活性失活而用作切口酶。由于由此产生的切口酶具有RuvC结构域的核酸酶活性,它能够切割靶基因或核酸的非互补链,即不与gRNA形成互补结合的链。

[0943] 例如,切口酶可具有HNH结构域的核酸酶活性。即,切口酶可不包含RuvC结构域的核酸酶活性,为此可对RuvC结构域进行操纵或修饰。

[0944] 在一个实例中,在CRISPR酶是II型CRISPR酶的情况下,在一个示例性实施方式中,当将SpCas9氨基酸序列中的第10位残基由天冬氨酸突变为丙氨酸时,RuvC结构域的核酸酶活性失活而用作切口酶。由于由此产生的切口酶具有HNH结构域的核酸酶活性,因此它能够切割靶基因或核酸的互补链,即与gRNA形成互补结合的链。

[0945] 在另一示例性实施方式中,当将CjCas9氨基酸序列中的第8位残基由天冬氨酸突变为丙氨酸时,RuvC结构域的核酸酶活性失活而用作切口酶。由于由此产生的切口酶具有HNH结构域的核酸酶活性,它能够切割靶基因或核酸的互补链,即与gRNA形成互补结合的链。

[0946] 失活的酶

[0947] 经修饰而使得酶活性完全失活的CRISPR酶称为失活的CRISPR酶。

[0948] 术语“失活的CRISPR酶”是指经修饰而完全不切割靶基因或核酸双链的CRISPR酶,由于野生型CRISPR酶中具有核酸酶活性的结构域中的突变,失活的CRISPR酶不具核酸酶活性。失活的CRISPR酶可为其中RuvC结构域和HNH结构域的核酸酶活性失活的酶。

[0949] 例如,失活的CRISPR酶可为对RuvC结构域和HNH结构域进行操纵或修饰而使得核酸酶活性失活。

[0950] 在一个实例中,在CRISPR酶是II型CRISPR酶的情况下,在一个示例性实施方式中,当分别将SpCas9氨基酸序列中的第10位和第840位残基由天冬氨酸和组氨酸突变为丙氨酸时,RuvC结构域和HNH结构域的核酸酶活性失活,从而可完全不切割靶基因或核酸的双链。

[0951] 在另一示例性实施方式中,当将CjCas9氨基酸序列中的第8位和第559位残基由天冬氨酸和组氨酸突变为丙氨酸时,RuvC结构域和HNH结构域的核酸酶活性失活,从而可完全不切割靶基因或核酸的双链。

[0952] 其它活性

[0953] 除上述核酸酶活性外,CRISPR酶还可具有核酸内切酶活性、核酸外切酶活性或解旋酶活性(即,使得双链核酸的螺旋结构解旋的能力)。

[0954] 此外,可对CRISPR酶进行修饰而使其具有核酸内切酶活性、核酸外切酶活性或解旋酶活性的完全活性、不完全或部分活性。

[0955] CRISPR酶的靶向

[0956] CRISPR酶可与gRNA相互作用,从而形成gRNA-CRISPR酶复合体,即CRISPR复合体,并且可在gRNA的协作下使得引导序列接近包含PAM序列的靶序列。此处,CRISPR酶与靶基因或核酸相互作用的能力依赖于PAM序列。

[0957] PAM序列是存在于靶基因或核酸中的序列,其可被CRISPR酶的PI结构域识别。PAM序列可根据CRISPR酶的来源改变。即,取决于物种,存在能够被特异性识别的多种PAM序列。

[0958] 在一个实例中,在CRISPR酶是II型CRISPR酶的情况下,

[0959] 在SpCas9的情况下,PAM序列可为5'-NGG-3'、5'-NAG-3'和/或5'-NGA-3';

[0960] 在StCas9的情况下,PAM序列可为5'-NGGNG-3'和/或5'-NNAGAAW-3'(W=A或T);

[0961] 在NmCas9的情况下,PAM序列可为5'-NNNGATT-3'和/或5'-NNNGCTT-3';

[0962] 在CjCas9的情况下,PAM序列可为5'-NNVRYAC-3'(V=G、C或A;R=A或G;Y=C或T);

[0963] 在变形链球菌Cas9(SmCas9)的情况下,PAM序列可为5'-NGG-3'和/或5'-NAAR-3'(R=A或G);以及

[0964] 在金黄色葡萄球菌Cas9(SaCas9)的情况下,PAM序列可为5'-NGRR-3'、5'-NGRRT-3'和/或5'-NGRRV-3'(R=A或G;V=G、C或A)。

[0965] 在另一实例中,当CRISPR酶是V型CRISPR酶的情况下,

[0966] 在Cpf1的情况下,PAM序列可为5'-TTN-3'。

[0967] 此处,N可为A、T、G或C;或为A、U、G或C。

[0968] 可利用取决于物种能被特异性识别的PAM序列对能够识别特定PAM序列的CRISPR酶进行操纵或修饰。例如,可利用CjCas9的PI结构域替换SpCas9的PI结构域,使其具有SpCas9的核酸酶活性并识别CjCas9特异性PAM序列,从而产生识别CjCas9特异性PAM序列的SpCas9。可通过对PI结构域进行置换或替换来改变特异性识别的PAM序列。

[0969] CRISPR酶突变体

[0970] 可对CRISPR酶进行修饰来改善或抑制多种特征,例如核酸酶活性、解旋酶活性、与gRNA相互作用的能力、接近靶基因或核酸的能力(例如CRISPR酶的PAM识别能力)。

[0971] 此外,CRISPR酶突变体可为这样的CRISPR酶:其与gRNA相互作用以形成gRNA-CRISPR酶复合体(即CRISPR复合体),并经修饰或操纵以改善靶特异性,使得当接近或定位于靶基因或核酸时,仅切割靶基因或核酸的双链或单链,而不切割与gRNA形成部分互补结合的非靶基因或核酸以及不与gRNA形成互补结合的非靶基因或核酸的双链或单链。

[0972] 此处,将对与gRNA形成部分互补结合的非靶基因或核酸的双链或单链或者不与gRNA形成互补结合的非靶基因或核酸的双链或单链进行切割的效应称为脱靶效应,将与gRNA形成部分互补结合的非靶基因或核酸或者不与gRNA形成互补结合的非靶基因或核酸中的位置或碱基序列称为脱靶靶标。此处,可存在一个或多个脱靶靶标。另一方面,将对靶基因或核酸的双链或单链进行切割的效应称为中靶效应,并将靶基因或核酸中的位置或靶序列称为中靶靶标。

[0973] CRISPR酶突变体可为对天然存在的CRISPR酶的氨基酸中的至少一个进行修饰,与未修饰的CRISPR酶相比,所作修饰例如可改善或抑制各种特征中的一种或多种,例如核酸酶活性、解旋酶活性、与gRNA相互作用的能力、接近靶基因或核酸的能力以及靶特异性。此处,修饰可为氨基酸的置换、删除、添加或它们的混合。

[0974] 在CRISPR酶突变体中，

[0975] 修饰可为对位于存在于天然CRISPR酶中的由带正电的氨基酸组成的区域中的一个或两个以上氨基酸进行的修饰。

[0976] 例如，修饰可为对存在于天然CRISPR酶中的一个或两个以上带正电的氨基酸（例如赖氨酸(K)、精氨酸(R)和组氨酸(H)）进行的修饰。

[0977] 修饰可为对位于存在于天然CRISPR酶中的由不带正电的氨基酸组成的区域中的一个或两个以上氨基酸进行的修饰。

[0978] 例如，修饰可为对存在于天然CRISPR酶中的一个或两个以上不带正电的氨基酸（即天冬氨酸(D)、谷氨酸(E)、丝氨酸(S)、苏氨酸(T)、天冬酰胺(N)、谷氨酰胺(Q)、半胱氨酸(C)、脯氨酸(P)、甘氨酸(G)、丙氨酸(A)、缬氨酸(V)、异亮氨酸(I)、亮氨酸(L)、甲硫氨酸(M)、苯丙氨酸(F)、酪氨酸(Y)和色氨酸(W)）进行的修饰。

[0979] 在另一实例中，修饰可为对存在于天然CRISPR酶中的一个或两个以上非带电氨基酸（即丝氨酸(S)、苏氨酸(T)、天冬酰胺(N)、谷氨酰胺(Q)、半胱氨酸(C)、脯氨酸(P)、甘氨酸(G)、丙氨酸(A)、缬氨酸(V)、异亮氨酸(I)、亮氨酸(L)、甲硫氨酸(M)、苯丙氨酸(F)、酪氨酸(Y)和色氨酸(W)）进行的修饰。

[0980] 此外，修饰可为对存在于天然CRISPR酶中的一个或两个以上带疏水残基的氨基酸进行的修饰。

[0981] 例如，修饰可为对存在于天然CRISPR酶中的甘氨酸(G)、丙氨酸(A)、缬氨酸(V)、异亮氨酸(I)、亮氨酸(L)、甲硫氨酸(M)、苯丙氨酸(F)、酪氨酸(Y)和色氨酸(W)中的一个或两个以上氨基酸进行的修饰。

[0982] 修饰可为对存在于天然CRISPR酶中的一个或两个以上带极性残基的氨基酸进行的修饰。

[0983] 例如，修饰可为对存在于天然CRISPR酶中的丝氨酸(S)、苏氨酸(T)、天冬酰胺(N)、谷氨酰胺(Q)、半胱氨酸(C)、脯氨酸(P)、赖氨酸(K)、精氨酸(R)、组氨酸(H)、天冬氨酸(D)和谷氨酸(E)中的一个或两个以上氨基酸进行的修饰。

[0984] 此外，修饰可为对存在于天然CRISPR酶中的包括赖氨酸(K)、精氨酸(R)和组氨酸(H)在内的一个或两个以上氨基酸进行的修饰。

[0985] 例如，修饰可为对存在于天然CRISPR酶中的包括赖氨酸(K)、精氨酸(R)和组氨酸(H)在内的一个或两个以上氨基酸进行的置换。

[0986] 修饰可为对存在于天然CRISPR酶中的包括天冬氨酸(D)和谷氨酸(E)在内的一个或两个以上氨基酸进行的修饰。

[0987] 例如，修饰可为对存在于天然CRISPR酶中的包括天冬氨酸(D)和谷氨酸(E)在内的一个或两个以上氨基酸进行的置换。

[0988] 修饰可为对存在于天然CRISPR酶中的包括丝氨酸(S)、苏氨酸(T)、天冬酰胺(N)、谷氨酰胺(Q)、半胱氨酸(C)、脯氨酸(P)、甘氨酸(G)、丙氨酸(A)、缬氨酸(V)、异亮氨酸(I)、亮氨酸(L)、甲硫氨酸(M)、苯丙氨酸(F)、酪氨酸(Y)和色氨酸(W)在内的一个或两个以上氨基酸进行的修饰。

[0989] 例如，修饰可为对存在于天然CRISPR酶中的包括丝氨酸(S)、苏氨酸(T)、天冬酰胺(N)、谷氨酰胺(Q)、半胱氨酸(C)、脯氨酸(P)、甘氨酸(G)、丙氨酸(A)、缬氨酸(V)、异亮氨酸

(I)、亮氨酸(L)、甲硫氨酸(M)、苯丙氨酸(F)、酪氨酸(Y)和色氨酸(W)在内的一个或两个以上氨基酸进行的置换。

[0990] 此外,修饰可为对存在于天然CRISPR酶中的一个、两个、三个、四个、五个、六个、七个或更多个氨基酸进行的修饰。

[0991] 此外,在CRISPR酶突变体中,

[0992] 修饰可为对CRISPR酶的RuvC结构域中存在的一个或两个以上氨基酸进行的修饰。此处,RuvC结构域可为RuvCI、RuvCII或RuvCIII结构域。

[0993] 修饰可为对CRISPR酶的HNH结构域中存在的一个或两个以上氨基酸进行的修饰。

[0994] 修饰可为对CRISPR酶的REC结构域中存在的一个或两个以上氨基酸进行的修饰。

[0995] 修饰可为对CRISPR酶的PI结构域中存在的一个或两个以上氨基酸进行的修饰。

[0996] 修饰可为对CRISPR酶的REC、RuvC、HNH和PI结构域中的至少两个或更多个结构域中含有的两个以上氨基酸进行的修饰。

[0997] 在一个实例中,修饰可为对CRISPR酶的REC和RuvC结构域中含有的两个以上氨基酸进行的修饰。

[0998] 在一个示例性实施方式中,在SpCas9突变体中,修饰可为对SpCas9的REC和RuvC结构域中含有的A203、H277、G366、F539、I601、M763、D965和F1038中的至少两个或更多个氨基酸进行的修饰。

[0999] 在另一实例中,修饰可为对CRISPR酶的REC和HNH结构域中含有的两个以上氨基酸进行的修饰。

[1000] 在一个示例性实施方式中,在SpCas9突变体中,修饰可为对SpCas9的REC和HNH结构域中含有的A203、H277、G366、F539、I601和K890中的至少两个或更多个氨基酸进行的修饰。

[1001] 在一个实例中,修饰可为对CRISPR酶的REC和PI结构域中含有的两个以上氨基酸进行的修饰。

[1002] 在一个示例性实施方式中,在SpCas9突变体中,修饰可为对SpCas9的REC和PI结构域中含有的A203、H277、G366、F539、I601、T1102和D1127中的至少两个或更多个氨基酸进行的修饰。

[1003] 在另一实例中,修饰可为对CRISPR酶的REC、RuvC和HNH结构域中含有的三个以上氨基酸进行的修饰。

[1004] 在一个示例性实施方式中,在SpCas9突变体中,修饰可为对SpCas9的REC、RuvC和HNH结构域中含有的A203、H277、G366、F539、I601、M763、K890、D965和F1038中的至少三个或更多个氨基酸进行的修饰。

[1005] 在一个实例中,修饰可为对CRISPR酶的REC、RuvC和PI结构域中含有的三个以上氨基酸进行的修饰。

[1006] 在一个示例性实施方式中,在SpCas9突变体中,修饰可为对SpCas9的REC、RuvC和PI结构域中含有的A203、H277、G366、F539、I601、M763、D965、F1038、T1102和D1127中的至少三个或更多个氨基酸进行的修饰。

[1007] 在另一实例中,修饰可为对CRISPR酶的REC、HNH和PI结构域中含有的三个以上氨基酸进行的修饰。

[1008] 在一个示例性实施方式中,在SpCas9突变体中,修饰可为对SpCas9的REC、HNH和PI结构域中含有的A203、H277、G366、F539、I601、K890、T1102和D1127中的至少三个或更多个氨基酸进行的修饰。

[1009] 在一个实例中,修饰可为对CRISPR酶的RuvC、HNH和PI结构域中含有的三个以上氨基酸进行的修饰。

[1010] 在一个示例性实施方式中,在SpCas9突变体中,修饰可为对SpCas9的RuvC、HNH和PI结构域中含有的M763、K890、D965、F1038、T1102和D1127中的至少三个或更多个氨基酸进行的修饰。

[1011] 在另一实例中,修饰可为对CRISPR酶的REC、RuvC、HNH和PI结构域中含有的四个以上氨基酸进行的修饰。

[1012] 在一个示例性实施方式中,在SpCas9突变体中,修饰可为对SpCas9的REC、RuvC、HNH和PI结构域中含有的A203、H277、G366、F539、I601、M763、K890、D965、F1038、T1102和D1127中的至少四个或更多个氨基酸进行的修饰。

[1013] 此外,在CRISPR酶突变体中,

[1014] 修饰可为对参与CRISPR酶的核酸酶活性的一个或两个以上氨基酸进行的修饰。

[1015] 例如,在SpCas9突变体中,修饰可为对由如下氨基酸所组成的组中的一个或两个以上氨基酸进行的修饰:D10、E762、H840、N854、N863以及D986;或者修饰可为对由其它Cas9直系同源物中对应于这些氨基酸的氨基酸所组成的组中的一个或两个以上氨基酸进行的修饰。

[1016] 修饰可为使得CRISPR酶的核酸酶活性部分失活的修饰,从而CRISPR酶突变体可为切口酶。

[1017] 此处,修饰可为使得CRISPR酶的RuvC结构域的核酸酶活性失活的修饰,从而CRISPR酶突变体不能切割靶基因或核酸的非互补链(即不与gRNA形成互补结合的链)。

[1018] 在一个示例性实施方式中,在SpCas9的情况下,当将SpCas9氨基酸序列的第10位残基由天冬氨酸突变为丙氨酸(即D10A突变)时,RuvC结构域的核酸酶活性失活,从而SpCas9可用作切口酶。由此生成的切口酶不能切割靶基因或核酸的非互补链(即不与gRNA形成互补结合的链)。

[1019] 在另一示例性实施方式中,在CjCas9的情况下,当将CjCas9氨基酸序列的第8位残基由天冬氨酸突变为丙氨酸(即D8A突变)时,RuvC结构域的核酸酶活性失活,从而CjCas9可用作切口酶。由此生成的切口酶不能切割靶基因或核酸的非互补链(即不与gRNA形成互补结合的链)。

[1020] 此外,此处,修饰可为使得CRISPR酶的HNH结构域的核酸酶活性失活的修饰,从而CRISPR酶突变体不能切割靶基因或核酸的互补链(即与gRNA形成互补结合的链)。

[1021] 在一个示例性实施方式中,在SpCas9的情况下,当将SpCas9氨基酸序列的第840位残基由组氨酸突变为丙氨酸(即H840A突变)时,HNH结构域的核酸酶活性失活,从而SpCas9可用作切口酶。由此生成的切口酶不能切割靶基因或核酸的互补链(即与gRNA形成互补结合的链)。

[1022] 在另一示例性实施方式中,在CjCas9的情况下,当将CjCas9氨基酸序列的第559位残基由组氨酸突变为丙氨酸(即H559A突变)时,HNH结构域的核酸酶活性失活,从而CjCas9

可用作切口酶。由此生成的切口酶不能切割靶基因或核酸的互补链(即与gRNA形成互补结合的链)。

[1023] 此外,修饰可为使得CRISPR酶的核酸酶活性完全失活的修饰,从而CRISPR酶突变体可为失活的CRISPR酶。

[1024] 此处,修饰可为使得CRISPR酶的RuvC和HNH结构域的核酸酶活性失活的修饰,从而CRISPR酶突变体不能切割靶基因或核酸的双链。

[1025] 在一个示例性实施方式中,在SpCas9的情况下,当将SpCas9氨基酸序列的第10位和840位残基分别由天冬氨酸和组氨酸突变为丙氨酸(即D10A和H840A突变)时,RuvC结构域和HNH结构域的核酸酶活性失活,从而靶基因或核酸的双链可完全不被切割。

[1026] 在另一示例性实施方式中,在CjCas9的情况下,当将CjCas9氨基酸序列的第8位和559位残基分别由天冬氨酸和组氨酸突变为丙氨酸(即D8A和H559A突变)时,RuvC结构域和HNH结构域的核酸酶活性失活,从而靶基因或核酸的双链可完全不被切割。

[1027] 此外,除CRISPR酶的固有特征外,CRISPR酶突变体还可进一步包含任选的功能结构域,从而CRISPR酶突变体还可具有除固有特征外的额外特征。

[1028] 此处,功能结构域可为具有甲基化酶活性、去甲基化酶活性、转录激活活性、转录阻遏活性、转录释放因子活性、组蛋白修饰活性、RNA切割活性或核酸结合活性的结构域,或者为用于纯化和分离蛋白(包括肽)的标签或报告基因,但本发明不限于此。

[1029] 功能结构域、肽、多肽或蛋白可为脱氨酶。

[1030] 例如,不完整或部分的CRISPR酶可额外包含胞苷脱氨酶作为功能结构域。在一个示例性实施方式中,可将胞苷脱氨酶(例如载脂蛋白B编辑复合体1(APOBEC1))添加至SpCas9切口酶,从而生成融合蛋白。由此形成的[SpCas9切口酶]-[APOBEC1]可用于由C到T或U、或者由G到A的编辑或者碱基修复中。

[1031] 标签包括组氨酸(His)标签、V5标签、FLAG标签、流感血凝素(HA)标签、Myc标签、VSV-G标签和硫氧还蛋白(Trx)标签;报告基因包括谷胱甘肽硫转移酶(GST)、辣根过氧化物酶(HRP)、氯霉素乙酰转移酶(CAT)、 β -半乳糖苷酶、 β -葡萄糖醛酸酶、萤光素酶、自发荧光蛋白(包括绿色荧光蛋白(GFP)、HcRed、DsRed、青色荧光蛋白(CFP)、黄色荧光蛋白(YFP)和蓝色荧光蛋白(BFP)),但本发明不限于此。

[1032] 此外,功能结构域可为核定位序列或信号(NLS)或者核输出序列或信号(NES)。

[1033] 在一个实例中,CRISPR酶可包含一个或多个NLS。此处,一个或多个NLS可包含于CRISPR酶的N端或其附近、酶的C端或其附近或者二者的组合。NLS可为由如下NLS衍生而来的NLS序列,但本发明不限于此:具有氨基酸序列PKKKRKV的SV40病毒大T抗原的NLS;来自核质蛋白的NLS(例如具有序列KRPAATKKAGQAKKKK的双分型核质蛋白NLS);具有氨基酸序列PAAKRVKLD或RQRRNELKRSP的c-myc NLS;具有序列NQSSNFGPMKGGNFGGRSSGPYGGGGQYFAKPRNQGGY的hRNPA1M9NLS;来自输入蛋白 α 的IBB结构域的序列RMRIZFKNKGKDTAELRRRRVEVSVELRKAKKDEQILKRRNV;肌瘤T蛋白的序列VSRKRPRP和PPKKARED;人p53的序列POPKKKPL;小鼠c-abl IV的序列SALIKKKKKMAP;流感病毒NS1的序列DRLRR和PKQKKRK;肝炎病毒 δ 抗原的序列RKLKKKIKKL;小鼠Mx1蛋白的序列REKKKFLKRR;人多聚(ADP-核糖)聚合酶的序列KRKGDEVDGVDEVAKKKSKK;或者由类固醇激素受体(人)糖皮质激素的序列衍生而来的NLS序列RKCLQAGMNLEARKTKK。

[1034] 此外,CRISPR酶突变体可包括通过将CRISPR酶分为两个以上部分而制备的拆分型(split-type)CRISPR酶。术语“拆分”是指对蛋白进行功能或结构性划分,或者将蛋白随机划分为两个以上部分。

[1035] 此处,拆分型CRISPR酶可为具有完全活性的酶、具有不完全或部分活性的酶或失活的酶。

[1036] 例如,可在第656位残基(酪氨酸)和第657位残基(苏氨酸)之间将SpCas9分为两部分,从而生成拆分型SpCas9。

[1037] 此外,拆分型CRISPR酶可任选包含用于重构(reconstitution)的额外的结构域、肽、多肽或蛋白。

[1038] 此处,“重构”是指所形成的拆分型CRISPR酶在结构方面与野生型CRISPR酶相同或类似。

[1039] 用于重构的额外的结构域、肽、多肽或蛋白可为FRB和FKBP二聚化结构域;内含肽(intein);ERT和VPR结构域;或者在特定条件下形成异二聚体的结构域。

[1040] 例如,可在第713位残基(丝氨酸)和第714位残基(甘氨酸)之间将SpCas9分为两部分,从而生成拆分型SpCas9。可将FRB结构域连接至两部分中的一个部分,并将FKBP结构域连接至另一部分。在由此产生的拆分型SpCas9中,FRB结构域和FKBP结构域可以在存在雷帕霉素的环境中形成二聚体,从而生成重构的CRISPR酶。

[1041] 本发明所述的CRISPR酶或CRISPR酶突变体可为多肽、蛋白或者具有编码所述多肽、蛋白的序列的核酸,并可针对待导入所述CRISPR酶或CRISPR酶突变体的受试者实施密码子优化。

[1042] 术语“密码子优化”是指对核酸序列的修饰过程,该修饰过程通过在保持天然氨基酸序列的同时将天然序列中的至少一个密码子替换为在宿主细胞中更常或最常使用的密码子来改善在宿主细胞中的表达。多种物种对特定氨基酸的特定密码子具有特定偏好,该密码子偏好(不同生物体间密码子使用的差别)通常与mRNA的翻译效率相关,认为这取决于所翻译的密码子的特征和特定tRNA分子的可获得性。细胞中选择的优势tRNA通常反映了肽合成中最常使用的密码子。因此,可基于密码子优化在给定生物体中通过优化基因表达来对基因进行定制化。

[1043] 3. 靶序列

[1044] 术语“靶序列”是存在于靶基因或核酸中并与引导核酸的引导结构域中含有的引导序列具有互补性的碱基序列。靶序列是可根据靶基因或核酸(即用于基因操纵或修正的对象物)而改变的碱基序列,可根据靶基因或核酸将其设计为多种形式。

[1045] 靶序列可与引导核酸的引导结构域中含有的引导序列形成互补结合,靶序列的长度可与引导序列的长度相同。

[1046] 靶序列可为5-50个碱基的序列。

[1047] 在实施方式中,靶序列可为16、17、18、19、20、21、22、23、24或25个碱基的序列。

[1048] 靶序列可为与引导核酸的引导结构域中含有的引导序列互补的核酸序列,例如具有至少70%、75%、80%、85%、90%或95%或更高的互补性或完全互补性。

[1049] 在一个实例中,靶序列可为1-8个碱基的序列或包含1-8个碱基的序列,其与引导核酸的引导结构域中含有的引导序列不互补。

[1050] 此外,靶序列可为临近能够被编辑蛋白识别的核酸序列的碱基序列。

[1051] 在一个实例中,靶序列可为临近能够被编辑蛋白识别的核酸序列的5'端和/或3'端的连续的5-50个碱基的序列。

[1052] 在一个示例性实施方式中,gRNA-CRISPR酶复合体的靶序列将在下文中描述。

[1053] 当靶基因或核酸被gRNA-CRISPR酶复合体靶向时,

[1054] 靶序列具有对gRNA的引导结构域中含有的引导序列的互补性。靶序列是可根据靶基因或核酸(即用于基因操纵或修正的对象物)而改变的碱基序列,可根据靶基因或核酸将其设计为多种形式。

[1055] 此外,靶序列可为临近能够被CRISPR酶(即Cas9或Cpf1)识别的PAM序列的碱基序列。

[1056] 在一个实例中,靶序列可为临近被CRISPR酶识别的PAM序列的5'端和/或3'端的连续的5-50个碱基的序列。

[1057] 在一个示例性实施方式中,当CRISPR酶是SpCas9时,靶序列可为与5'-NGG-3'、5'-NAG-3'和/或5'-NGA-3'(N=A、T、G或C,或A、U、G或C)序列的5'端和/或3'端临近的连续的16-25个碱基的序列。

[1058] 在另一示例性实施方式中,当CRISPR酶是StCas9时,靶序列可为临近5'-NGGNG-3'和/或5'-NNAGAAW-3'(W=A或T;N=A、T、G或C,或A、U、G或C)序列的5'端和/或3'端的连续的16-25个碱基的序列。

[1059] 在又一示例性实施方式中,当CRISPR酶是NmCas9时,靶序列可为临近5'-NNNGATT-3'和/或5'-NNNGCTT-3'(N=A、T、G或C,或A、U、G或C)序列的5'端和/或3'端的连续的16-25个碱基的序列。

[1060] 在一个示例性实施方式中,当CRISPR酶是CjCas9时,靶序列可为临近5'-NNNRYAC-3'(V=G、C或A;R=A或G;Y=C或T;N=A、T、G或C,或A、U、G或C)序列的5'端和/或3'端的连续的16-25个碱基的序列。

[1061] 在另一示例性实施方式中,当CRISPR酶是SmCas9时,靶序列可为与5'-NGG-3'和/或5'-NAAR-3'(R=A或G;N=A、T、G或C,或A、U、G或C)序列的5'端和/或3'端临近的连续的16-25个碱基的序列。

[1062] 在又一示例性实施方式中,当CRISPR酶是SaCas9时,靶序列可为临近5'-NNGRR-3'、5'-NNGRRT-3'和/或5'-NNGRRV-3'(R=A或G;V=G、C或A;N=A、T、G或C,或A、U、G或C)序列的5'端和/或3'端的连续的16-25个碱基的序列。

[1063] 在一个示例性实施方式中,当CRISPR酶是Cpf1时,靶序列可为临近5'-TTN-3'(N=A、T、G或C,或A、U、G或C)序列的5'端和/或3'端的连续的16-25个碱基的序列。

[1064] 4. 引导核酸-编辑蛋白复合体及其用途

[1065] 引导核酸-编辑蛋白复合体可对靶标进行修饰。

[1066] 靶标可为靶核酸、基因、染色体或蛋白。

[1067] 例如,引导核酸-编辑蛋白复合体可用于对感兴趣的蛋白的表达进行最终调节(例如抑制、阻遏、降低、升高或促进)、去除蛋白、对蛋白活性进行调节(例如抑制、阻遏、降低、升高或促进)或表达新蛋白。

[1068] 此处,引导核酸-编辑蛋白复合体可在DNA、RNA、基因或染色体水平发挥作用。

[1069] 例如,引导核酸-编辑蛋白复合体可通过对靶DNA进行操纵或修饰来对靶DNA编码的蛋白的表达进行调节(例如抑制、阻遏、降低、升高或促进)、去除蛋白、对蛋白活性进行调节(例如抑制、阻遏、降低、升高或促进)或表达经修饰的蛋白。

[1070] 在另一实例中,引导核酸-编辑蛋白复合体可通过对靶RNA进行操纵或修饰来对靶DNA编码的蛋白的表达进行调节(例如抑制、阻遏、降低、升高或促进)、去除蛋白、对蛋白活性进行调节(例如抑制、阻遏、降低、升高或促进)或表达经修饰的蛋白。

[1071] 在一个实例中,引导核酸-编辑蛋白复合体可通过对靶基因进行操纵或修饰来对靶DNA编码的蛋白的表达进行调节(例如抑制、阻遏、降低、升高或促进)、去除蛋白、对蛋白活性进行调节(例如抑制、阻遏、降低、升高或促进)或表达经修饰的蛋白。

[1072] 在另一实例中,引导核酸-编辑蛋白复合体可通过对靶染色体进行操纵或修饰来对靶DNA编码的蛋白的表达进行调节(例如抑制、阻遏、降低、升高或促进)、去除蛋白、对蛋白活性进行调节(例如抑制、阻遏、降低、升高或促进)或表达经修饰的蛋白。

[1073] 引导核酸-编辑蛋白复合体可在基因转录和翻译阶段发挥作用。

[1074] 在一个实例中,引导核酸-编辑蛋白复合体可促进或阻遏靶基因的转录,从而对靶基因编码的蛋白的表达进行调节(例如抑制、阻遏、降低、升高或促进)。

[1075] 在另一实例中,引导核酸-编辑蛋白复合体可促进或阻遏靶基因的翻译,从而对靶基因编码的蛋白的表达进行调节(例如抑制、阻遏、降低、升高或促进)。

[1076] 引导核酸-编辑蛋白复合体可在蛋白水平发挥作用。

[1077] 在一个实例中,引导核酸-编辑蛋白复合体可对靶蛋白进行操纵或修饰,从而去除靶蛋白或对蛋白活性进行调节(例如抑制、阻遏、降低、升高或促进)。

[1078] 作为本发明的引导核酸-编辑蛋白复合体的用途的具体实例,下面对利用gRNA-CRISPR酶复合体对靶DNA、RNA、基因或染色体进行的操纵或修饰进行描述。

[1079] 基因操纵(Gene manipulation)

[1080] 可使用上述gRNA-CRISPR酶复合体(即CRISPR复合体)对靶基因或核酸进行操纵或修正。此处,对靶基因或核酸的操纵或修正包括用于i)对靶基因或核酸进行切割或损伤;以及ii)对损伤的靶基因或核酸进行修复的全部步骤。

[1081] i) 靶基因或核酸的切割或损伤

[1082] i) 靶基因或核酸的切割或损伤可为使用CRISPR复合体对靶基因或核酸进行切割或损伤,具体而言对靶基因或核酸中的靶序列进行切割或损伤。

[1083] 在一个实例中,使用CRISPR复合体对靶基因或核酸进行切割或损伤可为将靶序列的双链完全切割或损伤。

[1084] 在一个示例性实施方式中,当使用野生型SpCas9时,与gRNA形成互补结合的靶序列的双链可被完全切割。

[1085] 在另一示例性实施方式中,当使用SpCas9切口酶(D10A)和SpCas9切口酶(H840A)时,SpCas9切口酶(D10A)可切割与gRNA形成互补结合的靶序列的互补单链,而SpCas9切口酶(H840A)可切割与gRNA形成互补结合的靶序列的非互补单链,可顺序或同时实施切割。

[1086] 在又一示例性实施方式中,当使用SpCas9切口酶(D10A)和SpCas9切口酶(H840A)以及具有不同靶序列的两条gRNA时,SpCas9切口酶(D10A)可切割与第一gRNA形成互补结合的靶序列的互补单链,而SpCas9切口酶(H840A)可切割与第二gRNA形成互补结合的靶序列

的非互补单链,可顺序或同时实施切割。

[1087] 在另一实例中,使用CRISPR复合体对靶基因或核酸进行切割或损伤可为仅对靶序列的单链进行切割或损伤。此处,单链可为与gRNA形成互补结合的靶序列的互补单链或与gRNA形成互补结合的靶序列的非互补单链。

[1088] 在一个示例性实施方式中,当使用SpCas9切口酶(D10A)时,SpCas9切口酶(D10A)可切割与gRNA形成互补结合的靶序列的互补单链,而与gRNA形成互补结合的靶序列的非互补单链可不被切割。

[1089] 在另一示例性实施方式中,当使用SpCas9切口酶(H840A)时,SpCas9切口酶(H840A)可切割与gRNA形成互补结合的靶序列的非互补单链,而与gRNA形成互补结合的靶序列的互补单链可不被切割。

[1090] 在又一实例中,使用CRISPR复合体对靶基因或核酸进行切割或损伤可为部分去除核酸片段。

[1091] 在一个示例性实施方式中,当使用具有不同靶序列的两条gRNA和野生型SpCas9时,可对与第一gRNA形成互补结合的靶序列的双链进行切割,并可对与第二gRNA形成互补结合的靶序列的双链进行切割,从而借助第一gRNA和第二gRNA以及SpCas9去除核酸片段。

[1092] 在另一示例性实施方式中,当使用具有不同靶序列的两条gRNA、野生型SpCas9、SpCas9切口酶(D10A)和SpCas9切口酶(H840A)时,野生型SpCas9可切割与第一gRNA形成互补结合的靶序列的双链,SpCas9切口酶(D10A)可切割与第二gRNA形成互补结合的靶序列的互补单链,而SpCas9切口酶(H840A)可切割非互补单链,从而借助第一gRNA和第二gRNA、野生型SpCas9、SpCas9切口酶(D10A)以及SpCas9切口酶(H840A)去除核酸片段。

[1093] 在又一示例性实施方式中,当使用具有不同靶序列的两条gRNA、SpCas9切口酶(D10A)和SpCas9切口酶(H840A)时,SpCas9切口酶(D10A)可切割与第一gRNA形成互补结合的靶序列的互补单链,SpCas9切口酶(H840A)可切割非互补单链,SpCas9切口酶(D10A)可切割与第二gRNA形成互补结合的靶序列的互补单链,而SpCas9切口酶(H840A)可切割非互补单链,从而借助第一gRNA和第二gRNA、SpCas9切口酶(D10A)以及SpCas9切口酶(H840A)去除核酸片段。

[1094] 在又一示例性实施方式中,当使用具有不同靶序列的三条gRNA、野生型SpCas9、SpCas9切口酶(D10A)和SpCas9切口酶(H840A)时,野生型SpCas9可切割与第一gRNA形成互补结合的靶序列的双链,SpCas9切口酶(D10A)可切割与第二gRNA形成互补结合的靶序列的互补单链,而SpCas9切口酶(H840A)可切割与第三gRNA形成互补结合的靶序列的非互补单链,从而借助第一gRNA、第二gRNA、第三gRNA、野生型SpCas9、SpCas9切口酶(D10A)以及SpCas9切口酶(H840A)去除核酸片段。

[1095] 在又一示例性实施方式中,当使用具有不同靶序列的四条gRNA、SpCas9切口酶(D10A)和SpCas9切口酶(H840A)时,SpCas9切口酶(D10A)可切割与第一gRNA形成互补结合的靶序列的互补单链,SpCas9切口酶(H840A)可切割与第二gRNA形成互补结合的靶序列的非互补单链,SpCas9切口酶(D10A)可切割与第三gRNA形成互补结合的靶序列的互补单链,而SpCas9切口酶(H840A)可切割与第四gRNA形成互补结合的靶序列的非互补单链,从而借助第一gRNA、第二gRNA、第三gRNA、第四gRNA、SpCas9切口酶(D10A)以及SpCas9切口酶(H840A)去除核酸片段。

[1096] ii) 损伤的靶基因或核酸的修复或恢复

[1097] 可借助NHEJ和同源介导的修复 (HDR) 来对CRISPR复合体所切割或损伤的靶基因或核酸进行修复或恢复。

[1098] 非同源末端接合 (NHEJ)

[1099] NHEJ是通过将被切割的双链或单链的两端进行连接对DNA中的双链断裂进行恢复或修复的方法,一般而言,当将通过双链断裂(例如切割)形成的两个相容末端持续彼此接触使得两个末端完全相接,受损的双链得以修复。NHEJ是能够用于整个细胞周期的恢复方法,通常在细胞中无同源基因组作为模板时(例如G1期)发生。

[1100] 在利用NHEJ对损伤的基因或核酸进行修复的过程中,NHEJ修复区中的核酸序列出现一些插入和/或删除(插入缺失,indel),此类插入和/或删除导致读码框位移,产生移码突变的转录组mRNA。其结果是,由于无义介导的衰变(nonsense-mediated decay)或正常蛋白无法合成,造成固有功能丧失。此外,即使读码框保持不变,序列中相当数量的插入或缺失造成的突变也可导致蛋白功能破坏。由于相比蛋白中非重要区域的突变,对重要功能结构域中的突变可能耐受性更低,突变为基因座依赖型的。

[1101] 由于不能预测在天然状态下由NHEJ产生的插入缺失突变,特定的插入缺失序列优选位于指定的受损区中,并可来自于微同源的小区域。常规地,删除的长度范围为1bp-50bp,插入趋向于更短,并通常包含直接包围受损区域的短重复序列。

[1102] 此外,NHEJ是造成突变的过程,当不必须生成特定的最终序列时,可将NHEJ用于对短序列基序进行删除。

[1103] 可通过该NHEJ对CRISPR复合体靶向的基因进行特异性敲除。可使用CRISPR酶(例如Cas9或Cpf1)切割靶基因或核酸的双链或两条单链,并可借助NHEJ使得靶基因或核酸中的受损双链或两条单链具有插入缺失,从而诱导靶基因或核酸的特异性敲除。此处,CRISPR酶所切割的靶基因或核酸的位点可在非编码区或编码区;此外,由NHEJ所恢复的靶基因或核酸的位点可在非编码区或编码区。

[1104] 同源介导的修复 (HDR)

[1105] HDR是无错修正方法,其使用同源序列作为模板对损伤的基因或核酸进行修复或恢复,一般而言,为修复或恢复受损DNA(即,恢复细胞的固有信息),利用未被修饰的互补碱基序列的信息或者姐妹染色单体的信息对受损DNA进行修复。HDR最常见的类型为同源重组(HR)。HDR是通常出现在活跃分裂的细胞的S期或G2/M期的修复或恢复方法。

[1106] 为借助HDR而不使用细胞的姐妹染色单体或互补碱基序列对损伤的DNA进行修复或恢复,可将使用互补碱基序列或同源碱基序列的信息人工合成的DNA模板(即,包含互补碱基序列或同源碱基序列的核酸模板)提供至细胞来修复受损DNA。此处,当进一步将核酸序列或核酸片段添加至核酸模板来修复受损DNA时,可将进一步添加至受损DNA的核酸序列或核酸片段敲入。进一步添加的核酸序列或核酸片段可为对由正常基因或核酸的突变造成的修饰的靶基因或核酸进行修正的核酸序列或核酸片段,或者为待在细胞中表达的基因或核酸,但本发明不限于此。

[1107] 在一个实例中,可使用CRISPR复合体切割靶基因或核酸的双链或单链,可将核酸模板(该核酸模板包含与临近切割位点的碱基序列互补的碱基序列)提供至细胞,并可通过HDR修复或恢复靶基因或核酸中被切割的碱基序列。

[1108] 此处,包含互补碱基序列的核酸模板可具有受损DNA(即,互补碱基序列被切割的双链或单链),并进一步包含待插入至受损DNA的核酸序列或核酸片段。可使用包含互补碱基序列和待插入的核酸序列或核酸片段的核酸模板将额外核酸序列或核酸片段插入至受损DNA(即,靶基因或核酸的切割位点)。此处,待插入的核酸序列或核酸片段以及额外核酸序列或核酸片段可为对由正常基因或核酸的突变造成的修饰的靶基因或核酸进行修正的核酸序列或核酸片段,或者为待在细胞中表达的基因或核酸。互补碱基序列可为与受损DNA具有互补结合的碱基序列(即,靶基因或核酸被切割的双链或单链的左侧或右侧的碱基序列)。或者,互补碱基序列可为与受损DNA具有互补结合的碱基序列(即,靶基因或核酸被切割的双链或单链的5'和3'端)。互补碱基序列可为15-3000个碱基的序列,可根据核酸模板或靶基因的大小对互补碱基序列的长度或大小进行适当设计。此处,作为核酸模板,可使用双链或单链的核酸,或者其可为线性或环状,但本发明不限于此。

[1109] 在另一实例中,可使用CRISPR复合体切割靶基因或核酸的双链或单链,可将核酸模板(该核酸模板包含临近切割位点的碱基序列的同源碱基序列)提供至细胞,并可通过HDR修复或恢复靶基因或核酸中被切割的碱基序列。

[1110] 此处,包含同源碱基序列的核酸模板可为受损DNA(即,被切割的双链或单链同源碱基序列),并进一步包含待插入至受损DNA的核酸序列或核酸片段。可使用包含同源碱基序列和待插入的核酸序列或核酸片段的核酸模板将额外核酸序列或核酸片段插入至受损DNA(即,靶基因或核酸的切割位点)。此处,待插入的核酸序列或核酸片段以及额外核酸序列或核酸片段可为对由正常基因或核酸的突变造成的修饰的靶基因或核酸进行修正的核酸序列或核酸片段,或者为待在细胞中表达的基因或核酸。同源碱基序列可为受损DNA,即,与靶基因或核酸被切割的双链或单链碱基序列左侧和右侧的碱基序列具有同源性的碱基序列。或者,同源碱基序列可为与受损DNA具有同源性的碱基序列,即,与靶基因或核酸被切割的双链或单链的5'和3'端具有同源性的碱基序列。同源碱基序列可为15-3000个碱基的序列,可根据核酸模板或者靶基因或核酸的大小对同源碱基序列的长度或大小进行适当设计。此处,作为核酸模板,可使用双链或单链的核酸,或者其可为线性或环状,但本发明不限于此。

[1111] 存在不同于NHEJ和HDR的对受损DNA进行修复或恢复的方法。

[1112] 单链退火(SSA)

[1113] SSA是对靶核酸中存在的两个重复序列间的双链断裂进行修复的方法,一般使用多于30个碱基的重复序列。可对重复序列进行切割(以产生粘性末端),从而在靶核酸双链的各断裂端产生单链;并且,在切割后利用RPA蛋白对含有重复序列的单链垂悬部分(overhang)进行包覆,来防止重复序列彼此的不适当退火。RAD52结合至垂悬部分上的各重复序列,并排列能够对互补重复序列进行退火的序列。退火后,垂悬部分的单链悬垂(flap)被切割,合成新DNA来填充特定缺口,从而恢复DNA双链。该修复的结果是两个重复间的DNA序列被删除,删除长度可取决于多种因素(包括此处使用的两个重复的位置和切割的路径或进行度)。

[1114] 就对靶核酸序列进行修饰或修正而言,与HDR类似,SSA使用互补序列(即互补重复序列);与HDR不同,SSA不需要核酸模板。

[1115] 单链断裂修复(SSBR)

[1116] SSBR借助与上述修复机制不同的机制对基因组中的单链断裂进行修复。在单链DNA断裂的情况下,PARP1和/或PARP2识别断裂并动员修复机制。PARP1对DNA断裂的结合和活性是暂时的,通过促进受损区域中SSBR蛋白复合体的稳定性来促进SSBR。SSBR复合体中最重要的蛋白是XRCC1,它与促进DNA的3'和5'端加工的蛋白相互作用来稳定DNA。末端加工通常涉及将损伤的3'端修复为羟基化状态和/或将损伤的5'端修复为具有磷酸部分,并在末端加工后发生DNA缺口填充。存在两种DNA缺口填充方法,即短补丁(patch)修复和长补丁修复,短补丁修复涉及单碱基的插入。在DNA缺口填充后,DNA连接酶促进末端连接。

[1117] 错配修复(MMR)

[1118] MMR作用于错配的DNA碱基。MSH2/6或MSH2/3复合体各自具有ATPase活性,并因此在识别错配和引发修复中起到重要作用。MSH2/6主要识别碱基-碱基错配并识别一个或两个碱基的错配,而MSH2/3主要识别更长的错配。

[1119] 碱基切除修复(BER)

[1120] BER是在整个细胞周期中均活跃的修复方法,其用于从基因组中去除较小的非螺旋扭曲碱基损伤区。在损伤的DNA中,通过切割连接碱基与脱氧核糖-磷酸骨架的N-糖苷键去除损伤的碱基,随后切割磷酸二酯键骨架,从而生成单链DNA断裂。去除由此形成的受损单链末端,并利用新的互补碱基填充由于单链去除而造成的缺口,随后利用DNA连接酶将新填充的互补碱基的末端连接至骨架,实现对损伤DNA的修复。

[1121] 核苷酸切除修复(NER)

[1122] NER是对于从DNA中去除较大的螺旋扭曲损伤而言重要的切除机制,当识别到损伤时,去除含有损伤区域的短单链DNA片段,产生22-30个碱基的单链缺口。利用新的互补碱基填充产生的缺口,并利用DNA连接酶将新填充的互补碱基的末端连接至骨架,实现对损伤DNA的修复。

[1123] 基因操纵效果

[1124] 对靶基因或核酸的操纵或修正主要可产生敲除、敲减和敲入的效果。

[1125] 敲除

[1126] 术语“敲除”是指靶基因或核酸的失活,而“靶基因或核酸的失活”是指不发生靶基因或核酸的转录和/或翻译的状态。通过敲除可对造成疾病的基因或具有异常功能的基因的转录和翻译进行抑制,阻止蛋白表达。

[1127] 例如,当使用gRNA-CRISPR酶复合体(即CRISPR复合体)对靶基因或核酸进行编辑或修正时,可使用CRISPR复合体对靶基因或核酸进行切割。可利用CRISPR复合体通过NHEJ对损伤的靶基因或核酸进行修复。由于NHEJ,损伤的靶基因或核酸可具有插入缺失,从而可诱导针对靶基因或核酸的特异性敲除。

[1128] 敲减

[1129] 术语“敲减”是指靶基因或核酸的转录和/或翻译或靶蛋白的表达降低。通过敲减对基因或蛋白的过表达进行调节,可预防发病或可治疗疾病。

[1130] 例如,当利用gRNA-CRISPR失活酶-转录抑制活性结构域复合体(即,包含转录抑制活性结构域的CRISPR失活复合体)对靶基因或核酸进行编辑或修正时,CRISPR失活复合体可特异性地结合至靶基因或核酸,可借助CRISPR失活复合体中包含的转录抑制活性结构域对靶基因或核酸的转录进行抑制,从而诱导敲减(其中,相应基因或核酸的表达被抑制)。

[1131] 敲入

[1132] 术语“敲入”是指将特定核酸或基因插入至靶基因或核酸,具体而言,术语“特定核酸”是指待被插入或者期望被表达的基因或核酸。通过对造成疾病的突变基因进行精确修正或通过插入正常基因来诱导正常基因表达,可将敲入用于疾病的治疗。

[1133] 此外,敲入可需要额外供体(donor)。

[1134] 例如,当借助gRNA-CRISPR酶复合体(即CRISPR复合体)对靶基因或核酸进行编辑或修正时,可利用CRISPR复合体切割靶基因或核酸。可使用CRISPR复合体通过HDR修复损伤的靶基因或核酸。具体而言,可借助供体将特定核酸插入至受损的基因或核酸。

[1135] 术语“供体”是指通过HDR帮助修复损伤的基因或核酸的核酸序列,具体而言,模板可包含特定核酸。

[1136] 供体可为双链核酸或单链核酸。

[1137] 供体可为线性或环状。

[1138] 供体可包含与靶基因或核酸具有同源性的核酸序列。

[1139] 例如,供体可包含与待插入特定核酸的位置(例如损伤核酸的上游和下游)处的核苷酸序列分别具有同源性的核酸序列。具体而言,待插入的特定核酸可位于与损伤核酸的下游核酸序列具有同源性的核酸序列和与损伤核酸的上游核酸序列具有同源性的核酸序列之间。具体而言,具有上述同源性的核酸序列可具有至少50%、55%、60%、65%、70%、75%、80%、85%、90%或95%或更高的同源性或完全同源性。

[1140] 供体可任选包含额外核酸序列。具体而言,额外核酸序列可在增强供体的HDR效率、稳定性或敲入效率方面发挥作用。

[1141] 例如,额外核酸序列可为富含A和T碱基的核酸序列(即富A-T结构域)。或者,额外核酸序列可为支架/基质附着区(S/MAR)。

[1142] 5.其它额外成分

[1143] 任选地,可添加额外成分来增强引导核酸-编辑蛋白复合体的效率或改善对受损基因或核酸的修复效率。

[1144] 任选地,可使用额外成分来改善引导核酸-编辑蛋白复合体的效率。

[1145] 激活子(Activator)

[1146] 额外成分可用作激活子来提高引导核酸-编辑蛋白复合体对靶核酸、基因或染色体的切割效率。

[1147] 术语“激活子”是指作用于使得引导核酸-编辑蛋白复合体与靶核酸、基因或染色体间的结合稳定或者使得引导核酸-编辑蛋白复合体能够更容易地接近靶核酸、基因或染色体的核酸。

[1148] 激活子可为双链核酸或单链核酸。

[1149] 激活子可为线性或环状。

[1150] 激活子可分为使得引导核酸-编辑蛋白复合体与靶核酸、基因或染色体间的结合稳定的“协助部(helper)”以及使得引导核酸-编辑蛋白复合体能够更容易地接近靶核酸、基因或染色体的“护送部(escorter)”。

[1151] 协助部可提高引导核酸-编辑蛋白复合体对靶核酸、基因或染色体的切割效率。

[1152] 例如,协助部包含与靶核酸、基因或染色体具有同源性的核酸序列。因此,当引导

核酸-编辑蛋白复合体与靶核酸、基因或染色体结合时,协助部包含的同源核酸序列可与靶核酸、基因或染色体形成额外的互补结合,来稳定引导核酸-编辑蛋白复合体与靶核酸、基因或染色体间的结合。

[1153] 护送部可提高引导核酸-编辑蛋白复合体对靶核酸、基因或染色体的切割效率。

[1154] 例如,护送部包含与靶核酸、基因或染色体具有同源性的核酸序列。此处,护送部中包含的同源核酸序列可与引导核酸-编辑蛋白复合体的引导核酸形成部分互补结合。因此,与引导核酸-编辑蛋白复合体形成部分互补结合的护送部可与靶核酸、基因或染色体形成部分互补结合,其结果是可允许引导核酸-编辑蛋白复合体精确接近靶核酸、基因或染色体的位置。

[1155] 同源核酸序列可具有至少50%、55%、60%、65%、70%、75%、80%、85%、90%或95%或更高的同源性或完全同源性。

[1156] 此外,额外成分可任选地用于改善受损基因或核酸的修复效率。

[1157] 辅助子(assistor)

[1158] 额外成分可用作辅助子来改善受损基因或核酸的修复效率。

[1159] 术语“辅助子”是指参与受损基因或核酸(例如由引导核酸-编辑蛋白复合体切割的基因或核酸)的修复过程或提高修复效率的核酸。

[1160] 辅助子可为双链核酸或单链核酸。

[1161] 辅助子可以以线性或环状形式存在。

[1162] 根据修复方法,辅助子可分为参与借助NHEJ的修复过程或改善其修复效率的“NHEJ辅助子”以及参与借助HDR的修复过程或改善其修复效率的“HDR辅助子”。

[1163] NHEJ辅助子可参与借助NHEJ对受损基因或核酸进行修复的过程或改善其修复效率。

[1164] 例如,NHEJ辅助子可包含与受损核酸序列的一部分具有同源性的核酸序列。此处,同源核酸序列可包含与受损核酸序列一端(例如3'端)处的核酸序列具有同源性的核酸序列以及与受损核酸序列另一端(例如5'端)处的核酸序列具有同源性的核酸序列。此外,可包含与受损核酸序列上游和下游的碱基序列各自具有同源性的核酸序列。具有此类同源性的核酸序列可帮助受损核酸序列的两部分被置于紧密靠近的位置,从而使得借助NHEJ对受损核酸进行修复的效率增加。

[1165] HDR辅助子可参与借助HDR对受损基因或核酸进行修复的过程或改善其修复效率。

[1166] 例如,HDR辅助子可包含与受损核酸序列的一部分具有同源性的核酸序列。此处,同源核酸序列可包含与受损核酸序列一端(例如3'端)处的核酸序列具有同源性的核酸序列以及与受损核酸序列另一端(例如5'端)处的核酸序列具有同源性的核酸序列。或者,可包含与受损核酸序列上游和下游的碱基序列各自具有同源性的核酸序列。具有此类同源性的核酸序列可作为受损核酸序列的模板,使得借助HDR对受损核酸进行修复的效率增加。

[1167] 在另一实例中,HDR辅助子可包含与受损核酸序列的一部分具有同源性的核酸序列以及特定核酸(例如待插入的核酸或基因)。此处,同源核酸序列可包含与受损核酸序列上游和下游的碱基序列各自具有同源性的核酸序列。特定核酸可位于与受损核酸下游的碱基序列具有同源性的核酸序列和与受损核酸上游的碱基序列具有同源性的核酸序列之间。具有此类同源性的核酸序列和特定核酸可作为供体将特定核酸插入至受损核酸,从而使得

HDR的敲入效率增加。

[1168] 同源核酸序列可具有至少50%、55%、60%、65%、70%、75%、80%、85%、90%或95%或更高的同源性或完全同源性。

[1169] 6.受试者

[1170] 术语“受试者”是指导入引导核酸、编辑蛋白或引导核酸-编辑蛋白复合体的有机体；其中运行(operates)引导核酸、编辑蛋白或引导核酸-编辑蛋白复合体的有机体；或获取自所述有机体的试样或样本。

[1171] 受试者可为包含引导核酸-编辑蛋白复合体的靶核酸、基因、染色体或蛋白的有机体。

[1172] 有机体可为细胞、组织、植物、动物或人。

[1173] 细胞可为原核细胞或真核细胞。

[1174] 真核细胞可为植物细胞、动物细胞或人细胞，但本发明不限于此。

[1175] 组织可为动物体组织或人体组织，例如皮肤、肝、肾、心脏、肺、脑或肌肉组织。

[1176] 受试者可为包含引导核酸-编辑蛋白复合体的靶核酸、基因、染色体或蛋白的试样或样本。

[1177] 试样或样本可获取自包含靶核酸、基因、染色体或蛋白的有机体，并可为唾液、血液、皮肤组织、癌细胞或干细胞。

[1178] 作为本发明受试者的实施方式，下面将描述含有引导核酸-编辑蛋白复合体的靶基因或核酸的受试者。

[1179] 例如，靶基因可为PD-1基因、CTLA-4基因、TNFAIP3基因、DGKA基因、DGKZ基因、Fas基因、EGR2基因、PPP2R2D基因、PSGL-1基因、KDM6A基因和/或TET2基因。

[1180] 在实施方式中，上述各基因的靶序列可为选自于表1中所描述的序列(将PAM序列排除，且将T改为U)中的一种或多种。在引导核酸的设计中将所述靶序列作为基础。

[1181] 也就是说，上述表1中总结了各基因的靶序列区的核苷酸序列以及相应的引导RNA的靶序列区(引导RNA的靶序列区；以及具有能够与靶序列区杂交的核苷酸序列的引导RNA的靶序列区)(表1中示出的靶序列区以在3'端包含PAM序列(5'-NGG-3')的状态描述)。

[1182] 这些靶序列区的特征在于，其是除靶序列外在基因的基因组内不具有任何0bp-2bp的错配区的序列，具有低脱靶效应和高基因修正效率。

[1183] 可同时靶向两种以上类型的靶序列。

[1184] 可同时靶向两种以上类型的基因。

[1185] 可同时靶向同源基因中的两种以上的靶序列或异源基因中的两种以上的靶序列。

[1186] 可靶向基因内的非编码区或编码区(例如启动子区、增强子、3'UTR和/或多聚腺苷酸信号序列或转录序列(例如内含子或外显子序列))。

[1187] 可靶向基因编码区的前50%。

[1188] 在示例性实施方式中，可分别靶向DGKa或DGKz。

[1189] 在示例性实施方式中，可同时靶向DGKa和DGKz。

[1190] 在本发明的实施方式中，为对各基因进行人工操纵，提供了对应于靶序列SEQ ID NO:1-SEQ ID NO:289的引导核酸序列。

[1191] 在本发明的实施方式中，为对各基因进行人工操纵，提供了与对应于靶序列SEQ

ID NO:1-SEQ ID NO:289的引导核酸序列相互作用的编辑蛋白(例如形成复合体的蛋白)。

[1192] 在本发明的实施方式中,提供了各基因的核酸修饰产物(其中对SEQ ID NO:1-SEQ ID NO:289的靶序列区进行了人工操纵)及其表达产物。

[1193] 在本发明的实施方式中,为对各基因进行人工操纵,提供了引导核酸序列和与之相互作用的编辑蛋白的复合体,所述引导核酸序列对应于如下靶序列中的一种或多种靶序列:

[1194] SEQ ID NO:6和SEQ ID NO:11(A20);

[1195] SEQ ID NO:19、SEQ ID NO:20、SEQ ID NO:21和SEQ ID NO:23(DGKa);

[1196] SEQ ID NO:25(EGR2);

[1197] SEQ ID NO:64(PPP2R2D);

[1198] SEQ ID NO:87和SEQ ID NO:89(PD-1);

[1199] SEQ ID NO:109、SEQ ID NO:110、SEQ ID NO:111、SEQ ID NO:112和SEQ ID NO:113(Dgk ζ);

[1200] SEQ ID NO:126、SEQ ID NO:128和SEQ ID NO:129(Tet-2);

[1201] SEQ ID NO:182(PSGL-1);

[1202] SEQ ID NO:252、SEQ ID NO:254、SEQ ID NO:257和SEQ ID NO:264(FAS);以及

[1203] SEQ ID NO:285(KDM6A)。

[1204] 在本发明的实施方式中,提供了各基因的核酸修饰产物及其表达产物,其中,在各基因的核酸修饰产物中对如下靶序列区进行了人工操纵:SEQ ID NO:6和SEQ ID NO:11(A20);SEQ ID NO:19、SEQ ID NO:20、SEQ ID NO:21和SEQ ID NO:23(DGKa);SEQ ID NO:25(EGR2);SEQ ID NO:64(PPP2R2D);SEQ ID NO:87和SEQ ID NO:89(PD-1);SEQ ID NO:109、SEQ ID NO:110、SEQ ID NO:111、SEQ ID NO:112和SEQ ID NO:113(Dgk ζ);SEQ ID NO:126、SEQ ID NO:128和SEQ ID NO:129(Tet-2);SEQ ID NO:182(PSGL-1);SEQ ID NO:252、SEQ ID NO:254、SEQ ID NO:257和SEQ ID NO:264(FAS);以及SEQ ID NO:285(KDM6A)。

[1205] 7.递送

[1206] 可借助多种递送方法以多种形式将引导核酸、编辑蛋白或引导核酸-编辑蛋白复合体递送入或导入受试者。

[1207] 可以以DNA、RNA或混合形式将引导核酸递送入或导入受试者。

[1208] 可以以编码编辑蛋白的DNA、RNA、DNA/RNA混合物、肽、多肽或者蛋白的形式将编辑蛋白递送入或导入受试者。

[1209] 可以以编码各成分(即引导核酸或编辑蛋白)的DNA、RNA或其混合物的形式将引导核酸-编辑蛋白复合体递送入或导入受试者。

[1210] 引导核酸-编辑蛋白复合体可以作为引导核酸(具有DNA、RNA或其混合物的形式)与编辑蛋白(具有肽、多肽或蛋白的形式)的复合体递送入或导入受试者。

[1211] 此外,可借助多种递送方法以多种形式将能够提高或抑制引导核酸-编辑蛋白复合体的效率的额外成分递送入或导入受试者。

[1212] i) 以DNA、RNA或其混合物的形式递送

[1213] 可借助本领域已知的方法将编码引导核酸和/或编辑蛋白的DNA、RNA或其混合物的形式递送入或导入受试者。

[1214] 或者,可借助载体、非载体或其组合将编码引导核酸和/或编辑蛋白的DNA、RNA或其混合物的形式递送入或导入受试者。

[1215] 载体可为病毒载体或非病毒载体(例如质粒)。

[1216] 非载体可为裸DNA、DNA复合体或mRNA。

[1217] 基于载体的导入

[1218] 可借助载体将编码引导核酸和/或编辑蛋白的核酸序列递送入或导入受试者。

[1219] 载体可包含编码引导核酸和/或编辑蛋白的核酸序列。

[1220] 例如,载体可同时包含分别编码引导核酸和编辑蛋白的核酸序列。

[1221] 例如,载体可包含编码引导核酸的核酸序列。

[1222] 作为实例,引导核酸中包含的结构域可全部包含于一个载体中,或可将其分割并随后包含于不同载体中。

[1223] 例如,载体可包含编码编辑蛋白的核酸序列。

[1224] 在一个实例中,在编辑蛋白的情况下,编码编辑蛋白的核酸序列可包含于一个载体中,或可将其分割并随后包含于数个载体中。

[1225] 载体可含有一种或多种调节/控制成分。

[1226] 此处,调节/控制成分可以包括:启动子、增强子、内含子、多聚腺苷酸信号、Kozak共有序列、内部核糖体进入位点(IRES)、剪接受体和/或2A序列。

[1227] 启动子可为由RNA聚合酶II识别的启动子。

[1228] 启动子可为由RNA聚合酶III识别的启动子。

[1229] 启动子可为诱导型启动子。

[1230] 启动子可为受试者特异性启动子。

[1231] 启动子可为病毒或非病毒启动子。

[1232] 就启动子而言,可根据控制区(即,编码引导核酸或编辑蛋白的核酸序列)而使用适当的启动子。

[1233] 例如,可用于引导核酸的启动子可为H1、EF-1a、tRNA或U6启动子。例如,可用于编辑蛋白的启动子可为CMV、EF-1a、EFS、MSCV、PGK或CAG启动子。

[1234] 载体可为病毒载体或重组病毒载体。

[1235] 病毒可为DNA病毒或RNA病毒。

[1236] 此处,DNA病毒可为双链DNA(dsDNA)病毒或单链DNA(ssDNA)病毒。

[1237] 此处,RNA病毒可为单链RNA(ssRNA)病毒。

[1238] 病毒可为逆转录病毒、慢病毒、腺病毒、腺相关病毒(AAV)、痘苗病毒、痘病毒或单纯疱疹病毒,但本发明不限于此。

[1239] 一般说来,病毒可感染宿主(例如细胞),由此将编码病毒遗传信息的核酸导入宿主或将编码遗传信息的核酸插入宿主基因组。可使用具有此类特征的病毒将引导核酸和/或编辑蛋白导入受试者。借助病毒导入的引导核酸和/或编辑蛋白可在受试者(例如细胞)中瞬时表达。或者,借助病毒导入的引导核酸和/或编辑蛋白可在受试者(例如细胞)中长时间持续表达(例如1、2或3周,1、2、3、6或9个月,1或2年,或永久)。

[1240] 根据病毒的类型,病毒的包装能力可在至少2kb至50kb间变化。取决于此类包装能力,可设计包含引导核酸或编辑蛋白的病毒载体或者包含引导核酸和编辑蛋白二者的病毒

载体。或者,可设计包含引导核酸、编辑蛋白和额外成分的病毒载体。

[1241] 在一个实例中,可借助重组慢病毒实施编码引导核酸和/或编辑蛋白的核酸序列的递送或导入。

[1242] 在另一实例中,可借助重组腺病毒实施编码引导核酸和/或编辑蛋白的核酸序列的递送或导入。

[1243] 在又一实施方式中,可借助重组AAV实施编码引导核酸和/或编辑蛋白的核酸序列的递送或导入。

[1244] 在又一实例中,可借助混合病毒(例如本文列出的病毒中的一种以上的混合)实施编码引导核酸和/或编辑蛋白的核酸序列的递送或导入。

[1245] 基于非载体的导入

[1246] 可使用非载体将编码引导核酸和/或编辑蛋白的核酸序列递送入或导入受试者。

[1247] 非载体可包含编码引导核酸和/或编辑蛋白的核酸序列。

[1248] 非载体可为裸DNA、DNA复合体、mRNA或它们的混合物。

[1249] 可借助电穿孔、粒子轰击、声致穿孔(sonoporation)、磁性转染、瞬时细胞压缩或挤压(transient cell compression or squeezing)(例如在文献“Lee等(2012) Nano Lett.,12,6322-6327”中所记载的)、脂质介导的转染、树枝状大分子、纳米粒子、磷酸钙、二氧化硅、硅酸盐(Ormosil)或它们的组合将非载体递送入或导入受试者。

[1250] 作为实例,可通过将细胞与编码引导核酸和/或编辑蛋白的核酸序列在卡盒(cartridge)、腔室(chamber)或比色皿(cuvette)中混合并以预定的持续时间和振幅将电刺激施加至细胞来实施借助电穿孔的递送。

[1251] 在另一实例中,可使用纳米粒子实施非载体递送。纳米粒子可为无机纳米粒子(例如磁性纳米粒子、二氧化硅等)或有机纳米粒子(例如聚乙二醇(PEG)包覆的脂质等)。纳米粒子的外表面可缀合有能够进行附着的带正电的聚合物(例如聚乙烯亚胺、聚赖氨酸、聚丝氨酸等)。

[1252] ii) 以肽、多肽或蛋白的形式递送

[1253] 可借助本领域已知的方法以肽、多肽或蛋白的形式将编辑蛋白递送入或导入受试者。

[1254] 可借助电穿孔、显微注射、瞬时细胞压缩或挤压(例如在文献“Lee等(2012) Nano Lett.,12,6322-6327”中所记载的)、脂质介导的转染、纳米粒子、脂质体、肽介导的递送或它们的组合将肽、多肽或蛋白形式递送入或导入受试者。

[1255] 肽、多肽或蛋白可与编码引导核酸的核酸序列一起递送。

[1256] 在一个实例中,可通过将待导入编辑蛋白的细胞与引导核酸一起(或不与引导核酸一起)在卡盒、腔室或比色皿中混合并以预定的持续时间和振幅将电刺激施加至细胞来实施借助电穿孔的递送。

[1257] iii) 以核酸-蛋白混合物的形式递送

[1258] 可以以引导核酸-编辑蛋白复合体的形式将引导核酸和编辑蛋白递送入或导入受试者。

[1259] 例如,引导核酸可为DNA、RNA或其混合物。编辑蛋白可为肽、多肽或蛋白。

[1260] 在一个实例中,可以以包含RNA型引导核酸和蛋白型编辑蛋白的引导核酸-编辑蛋

白复合体(即核糖核蛋白(RNP))的形式将引导核酸和编辑蛋白递送入或导入受试者。

[1261] 在本发明中,作为用于将引导核酸和/或编辑蛋白递送入受试者的方法的实施方式,下面将对gRNA、CRISPR酶或gRNA-CRISPR酶复合体的递送进行描述。

[1262] 8.转化体(Transformant)

[1263] 术语“转化体”是指其中导入了引导核酸、编辑蛋白或引导核酸-编辑蛋白复合体的有机体;其中表达引导核酸、编辑蛋白或引导核酸-编辑蛋白复合体的有机体;或者获取自所述有机体的试样或样本。

[1264] 转化体可为其中以DNA、RNA或其混合物的形式导入了引导核酸、编辑蛋白或引导核酸-编辑蛋白复合体的有机体。

[1265] 例如,转化体可为其中导入了包含编码引导核酸和/或编辑蛋白的核酸序列的载体的有机体。此处,载体可为非病毒载体、病毒载体或重组病毒载体。

[1266] 在另一实例中,转化体可为其中以非载体形式导入了编码引导核酸和/或编辑蛋白的核酸序列的有机体。此处,非载体可为裸DNA、DNA复合体、mRNA或它们的混合物。

[1267] 转化体可为其中以肽、多肽或蛋白的形式导入了编辑蛋白、引导核酸-编辑蛋白复合体或引导核酸的有机体。

[1268] 转化体可为其中以DNA、RNA、肽、多肽、蛋白或它们的混合物的形式导入了引导核酸、编辑蛋白或引导核酸-编辑蛋白复合体的有机体。

[1269] 例如,转化体可为其中导入了包含RNA型引导核酸和蛋白型编辑蛋白的引导核酸-编辑蛋白复合体的有机体。

[1270] 转化体可为包含引导核酸-编辑蛋白复合体的靶核酸、基因、染色体或蛋白的有机体。

[1271] 有机体可为细胞、组织、植物、动物或人。

[1272] 细胞可为原核细胞或真核细胞。

[1273] 真核细胞可为植物细胞、动物细胞或人细胞,但本发明不限于此。

[1274] 组织可为动物体组织或人体组织,例如皮肤、肝、肾、心脏、肺、脑或肌肉组织。

[1275] 转化体可为其中导入了或表达引导核酸、编辑蛋白或引导核酸-编辑蛋白复合体的有机体,或获取自所述有机体的试样或样本。

[1276] 试样或样本可为唾液、血液、皮肤组织、癌细胞或干细胞。

[1277] 此外,在实施方式中,本发明提供了引导核酸-编辑蛋白复合体,该引导核酸-编辑蛋白复合体用于PD-1、CTLA-4、TNFAIP3、DGKA、DGKZ、Fas、EGR2、PPP2R2D、PSGL-1、KDM6A和/或TET2基因的靶位点中的核酸修饰。

[1278] 具体而言,可提供含有能够与来自基因的靶位点形成互补结合的结构域的gRNA分子(例如分离的或非天然存在的gRNA分子和编码该gRNA分子的DNA)。可对gRNA分子和编码该gRNA分子的DNA的序列进行设计,从而这些序列可与表1的靶位点序列具有互补结合。

[1279] 此外,可构建gRNA分子的靶位点,从而提供第三免疫调节因子,其中第三免疫调节因子与免疫细胞靶位置中的改变(例如双链断裂或单链断裂)相关;或在PD-1、CTLA-4、TNFAIP3、DGKA、DGKZ、Fas、EGR2、PPP2R2D、PSGL-1、KDM6A和/或TET2基因的靶位置中具有特定功能。

[1280] 此外,当利用两个以上gRNA来定位(locate)靶核酸中的两个以上切割事件(例如

双链或单链断裂)时,可利用相同或不同的Cas9蛋白来产生两个以上切割事件。

[1281] gRNA例如

[1282] 可靶向如下基因中的两种以上:PD-1、CTLA-4、TNFAIP3、DGKA、DGKZ、Fas、EGR2、PPP2R2D、PSGL-1、KDM6A和/或TET2基因;

[1283] 可靶向如下各基因内的两个以上位点:PD-1、CTLA-4、TNFAIP3、DGKA、DGKZ、Fas、EGR2、PPP2R2D、PSGL-1、KDM6A和/或TET2基因;

[1284] 可独立诱导如下基因的双链和/或单链的切割:PD-1、CTLA-4、TNFAIP3、DGKA、DGKZ、Fas、EGR2、PPP2R2D、PSGL-1、KDM6A和/或TET2基因;或者

[1285] 可诱导在如下基因的切割位点中插入一个或多个外源核苷酸:PD-1、CTLA-4、TNFAIP3、DGKA、DGKZ、Fas、EGR2、PPP2R2D、PSGL-1、KDM6A和/或TET2基因。

[1286] 此外,在本发明的另一实施方式中,构成引导核酸-编辑蛋白复合体的核酸可包含:

[1287] (a) 编码包含引导结构域的gRNA分子的序列,该引导结构域与本文所述的PD-1、CTLA-4、TNFAIP3、DGKA、DGKZ、Fas、EGR2、PPP2R2D、PSGL-1、KDM6A和/或TET2基因中的靶位点序列互补;以及

[1288] (b) 编码编辑蛋白的序列。

[1289] 具体而言,取决于靶位点,(a)中可存在两种以上;而(b)中可使用同一编辑蛋白或两种以上编辑蛋白。

[1290] 在实施方式中,将核酸构建为靶向与免疫细胞的敲减靶位置足够接近的酶促失活编辑蛋白或其融合蛋白(例如融合有转录阻遏结构域),用于抑制、降低或减少如下基因的表达:PD-1、CTLA-4、TNFAIP3、DGKA、DGKZ、Fas、EGR2、PPP2R2D、PSGL-1、KDM6A和/或TET2基因。

[1291] 此外,在本发明的实施方式中,通过引导核酸-编辑蛋白复合体对免疫细胞表达的基因(例如PD-1、CTLA-4、TNFAIP3、DGKA、DGKZ、Fas、EGR2、PPP2R2D、PSGL-1、KDM6A和/或TET2基因)进行的操纵可由任何机制介导。

[1292] 所述机制的实例包括但不限于非同源末端接合(NHEJ)、微同源介导的末端接合(microhomology-mediated end joining,MMEJ)、同源介导的修复(HDR)、合成依赖性链退火(SDSA)、单链退火或单链渗透。

[1293] 此外,明显的是,可将上文所述的引导核酸-编辑蛋白复合体的结构、功能和用途的全部特征用于对PD-1、CTLA-4、TNFAIP3、DGKA、DGKZ、Fas、EGR2、PPP2R2D、PSGL-1、KDM6A和/或TET2基因进行的操纵。

[1294] 在本发明的实施方式中,使用“引导核酸-编辑蛋白复合体”最终获得的产物免疫系统因子可为例如经操纵的基因;由该经操纵的基因表达的产物;含有它们的细胞、组合物、转化体等。

[1295] 在本发明的实施方式中,免疫系统因子为通过引导核酸-编辑蛋白复合体进行人工操纵的免疫调控基因或其表达的蛋白;以及含有它们的细胞。

[1296] 在本发明的实施方式中,免疫系统因子为通过引导核酸-编辑蛋白复合体进行基因操纵的免疫调节基因或其表达的蛋白;以及含有它们的细胞。

[1297] 在本发明的实施方式中,免疫系统因子为通过引导核酸-编辑蛋白复合体进行基

因操纵的免疫调节基因的核酸序列或氨基酸序列。

[1298] 在本发明的实施方式中,免疫系统因子为通过引导核酸-编辑蛋白复合体进行基因操纵的免疫调节基因;其表达的蛋白;含有经操纵的免疫调节因子和/或蛋白的细胞;或者含有经操纵的免疫调节因子、蛋白和/或细胞的组合物。

[1299] 在本发明的实施方式中,免疫系统因子为通过导入如下中的一种或多种而形成的转化体:通过引导核酸-编辑蛋白复合体进行基因操纵的免疫调节基因;其表达的蛋白;含有经操纵的免疫调节因子和/或蛋白的细胞;或者含有经操纵的免疫调节因子、蛋白和/或细胞的组合物。

[1300] 利用引导核酸-编辑蛋白复合体最终获得的产物免疫因子可独立地包含两种以上各因子,且每种因子可进一步包含两种以上。

[1301] 例如,

[1302] 免疫因子可以提供为同时包含经人工操纵的免疫调节基因、其表达的蛋白和含有它们的细胞中的两种以上的形式;

[1303] 可同时提供经人工操纵的一种或两种以上免疫调节基因;

[1304] 可同时提供经人工操纵的一种或两种以上免疫调节蛋白;

[1305] 可同时提供经人工操纵的一种或两种以上免疫细胞;以及

[1306] 可同时提供经人工操纵的免疫因子中的两种以上的组合。

[1307] 使用“引导核酸-编辑蛋白复合体”获得的产物免疫系统因子的优选实例可具有如下组成。

[1308] 在实施方式中,当免疫调节因子是基因时,通过引导核酸-编辑蛋白复合体进行人工操纵的免疫调节基因的组成可以

[1309] 在构成免疫调节基因的核酸序列中的原型间隔区邻近基序(PAM)序列中或在临近其5'端和/或3'端的连续的1bp-50bp、1bp-40bp、1bp-30bp并优选3bp-25bp的核苷酸序列区中,

[1310] 包含如下中的一种或多种核酸修饰:

[1311] 一个或多个核苷酸的删除或插入;

[1312] 利用不同于野生型基因的一个或多个核苷酸进行的置换;以及

[1313] 一个或多个外源核苷酸的插入。

[1314] 此外,通过引导核酸-编辑蛋白复合体进行人工操纵的免疫调节基因的组成可包括在构成免疫调节基因的核酸序列中的一个或多个核苷酸的化学修饰。

[1315] 具体而言,术语“外源核苷酸”是包括全部产生自外部(例如由异源有机体衍生而来的核苷酸或人工合成的核苷酸),而并非免疫调节基因所具有的那些核苷酸的概念。外源核苷酸不仅包括小尺寸寡核苷酸(50bp或更短),还包括用于表达具有特定功能的蛋白的大尺寸核苷酸(例如几百bp、几千bp或几万bp)。此类“外源核苷酸”也称为供体。

[1316] 化学修饰包括甲基化、乙酰化、磷酸化、泛素化、ADP-核糖基化、十四烷基化和糖基化,例如,核苷酸的部分官能团被氢原子、氟原子、-O-烷基、-O-酰基和氨基中的任一者取代,但不限于此。另外,为了增加核酸分子的转移能力,核苷酸的部分官能团可以被-Br、-Cl、-R、-R'OR、-SH、-SR、-N₃和-CN(R=烷基、芳基、亚烷基)中的任一者取代。另外,至少一个核苷酸的磷酸骨架被烷基膦酸酯(alkylphosphonate)形式、氨基磷酸酯

(phosphoroamidate)形式和硼代磷酸盐(boranophosphate)形式中的任一者取代。另外,化学修饰的特征可在于核酸分子中包含的至少一个核苷酸被锁核酸(LNA)、解锁核酸(UNA)、吗啉基和肽核酸(PNA)中的任一者取代,并且化学修饰的特征可在于核酸分子与选自于由脂质、细胞穿透肽和细胞靶向配体所组成的组中的一种或多种结合。

[1317] 为形成期望的免疫系统,可利用引导核酸-编辑蛋白复合体对构成人工免疫调节基因的核酸进行修饰。

[1318] 包含免疫调节基因的核酸修饰的位点(其能够形成期望的免疫系统)称为靶序列或靶位点。

[1319] “靶序列”可为引导核酸-编辑蛋白复合体的靶标,靶序列可包括编辑蛋白所识别的原型间隔区邻近基序(PAM)序列,但不限于此。靶序列可为实施者提供用于设计引导核酸的重要标准。

[1320] 此类核酸修饰包括核酸的“切割”。

[1321] 靶位点处的“切割”是指多聚核苷酸共价骨架的断裂。切割可包括但不限于磷酸二酯键的酶促水解或化学水解,并可通过多种其它方法进行。单链切割和双链切割都是可能的,双链切割可以作为两条不同的单链切割的结果而发生。双链切割可产生平末端或交错(staggered)末端。

[1322] 当使用非活性编辑蛋白时,可将具有特定功能的因子诱导定位至靠近靶位点或免疫调节基因的任何部分,不进行切割过程。取决于该特定功能,免疫调节基因的核酸序列中可包含一个或多个核苷酸的化学修饰。

[1323] 在实施方式中,可经由引导核酸-编辑蛋白复合体进行的核酸切割通过靶活性和非靶活性产生多个插入和删除(插入缺失,indel)。

[1324] 术语“插入缺失”统称此类突变:其中,在DNA的核苷酸序列中插入或删除了一些核苷酸。

[1325] 如上所述,当引导核酸-编辑蛋白复合体切割免疫调节基因的核酸(DNA、RNA)时,插入缺失可为在通过同源重组或非同源末端接合(NHEJ)机制进行修复的过程中引入靶序列的插入缺失。

[1326] 本发明的经人工操纵的免疫调节基因是指原始基因的核酸序列通过核酸的切割和插入缺失、供体的插入等而进行了修饰的基因;经人工操纵的免疫调节基因有助于建立期望的免疫系统(例如表现出促进或阻遏或补充特定免疫功能的效果)。

[1327] 例如,可借助经人工操纵的免疫调节基因促进特定蛋白的表达和活性。

[1328] 可借助经人工操纵的免疫调节基因使特定蛋白失活。

[1329] 在一个实例中,可对基因组中下调免疫应答的免疫调节基因(例如PD-1、CTLA-4、TNFAIP3、DGKA(Dgk α)、DGKZ(Dgk ζ)、Fas、EGR2、PPP2R2D、PSGL-1和/或TET2基因)的特定靶位点进行切割来敲减或敲除这些基因。

[1330] 在另一实例中,为改变转录,例如阻断、降低或减少PD-1、CTLA-4、TNFAIP3、DGKA(Dgk α)、DGKZ(Dgk ζ)、Fas、EGR2、PPP2R2D、PSGL-1和/或TET2基因的转录,可通过靶向编辑蛋白来介导靶向敲减,该编辑蛋白与转录阻遏结构域或染色质修饰蛋白融合并且是酶促失活的。

[1331] 可通过经人工操纵的免疫调节基因来调节免疫细胞的活性。可对免疫细胞中的增

殖、存活、细胞毒性、浸润以及细胞因子释放等进行调节。

[1332] 可通过经人工操纵的免疫调节基因来获得治疗效果(例如免疫功能、抗肿瘤功能、抗炎功能等)。

[1333] 取决于引导核酸-编辑蛋白复合体的构成特征,可改变免疫调节基因的靶位点所具有的主要PAM序列。

[1334] 下文中,本发明将就编辑蛋白和免疫调节基因的代表性实例进行描述,然而这些实施方式仅用于具体说明目的,本发明不限于这些实施方式。

[1335] 例如,当编辑蛋白是由酿脓链球菌衍生而来的Cas9蛋白时,PAM序列可为5'-NGG-3'(N为A、T、G或C);待切割的核苷酸序列区(靶位点)可为靶基因内临近5'-NGG-3'序列的5'端和/或3'端的连续的1bp-25bp(例如17bp-23bp或21bp-23bp)的核苷酸序列区。

[1336] 可提供因免疫调节基因中的如下修饰而产生的经人工操纵的免疫调节基因(例如经人工操纵的PD-1基因、CTLA-4基因、TNFAIP3基因、DGKA基因、DGKZ基因、Fas基因、EGR2基因、PPP2R2D基因、PSGL-1基因、KDM6A基因以及TET2基因):

[1337] a) 删除临近'NGG'(N为A、T、C或G)序列的5'端和/或3'端的连续的1bp-25bp(例如17bp-23bp)的核苷酸序列区中的一个或多个核苷酸;

[1338] b) 利用不同于野生型基因的一个或多个核苷酸置换临近'NGG'序列的5'端和/或3'端的连续的1bp-25bp(例如17bp-23bp)的核苷酸序列区中的一个或多个核苷酸;

[1339] c) 将一个或多个核苷酸插入至临近'NGG'序列的5'端和/或3'端的连续的1bp-25bp(例如17bp-23bp)的核苷酸序列区;或者

[1340] d) 选自a)-c)中的两者以上的组合。

[1341] 例如,当编辑蛋白是由空肠弯曲杆菌衍生而来的Cas9蛋白时,PAM序列可为5'-NNNNRYAC-3'(N各自独立地为A、T、G或C;R为A或G;Y为C或T);待切割的核苷酸序列区(靶位点)可为靶基因内临近5'-NNNNRYAC-3'序列的5'端和/或3'端的连续的1bp-25bp(例如17bp-23bp或21bp-23bp)的核苷酸序列区。

[1342] 可提供因免疫调节基因中的如下修饰而产生的经人工操纵的免疫调节基因(例如经人工操纵的PD-1基因、CTLA-4基因、TNFAIP3基因、DGKA基因、DGKZ基因、Fas基因、EGR2基因、PPP2R2D基因、PSGL-1基因、KDM6A基因以及TET2基因):

[1343] a') 删除临近'NNNNRYAC'(N各自独立地为A、T、G或C;R为A或G;Y为C或T)序列的5'端和/或3'端的连续的1bp-25bp(例如17bp-23bp)的核苷酸序列区中的一个或多个核苷酸;

[1344] b') 利用不同于野生型基因的一个或多个核苷酸置换临近'NNNNRYAC'序列的5'端和/或3'端的连续的1bp-25bp(例如17bp-23bp)的核苷酸序列区中的一个或多个核苷酸;

[1345] c') 将一个或多个核苷酸插入至临近'NNNNRYAC'序列的5'端和/或3'端的连续的1bp-25bp(例如17bp-23bp)的核苷酸序列区;或者

[1346] d') 选自a')-c')中的两者以上的组合。

[1347] 例如,当编辑蛋白是由嗜热链球菌衍生而来的Cas9蛋白时,PAM序列可为5'-NNAGAAW-3'(N各自独立地为A、T、G或C;W为A或T);待切割的核苷酸序列区(靶位点)可为靶基因内临近5'-NNAGAAW-3'序列的5'端和/或3'端的连续的1bp-25bp(例如17bp-23bp或21bp-23bp)的核苷酸序列区。

[1348] 可提供因免疫调节基因中的如下修饰而产生的经人工操纵的免疫调节基因(例如

经人工操纵的PD-1基因、CTLA-4基因、TNFAIP3基因、DGKA基因、DGKZ基因、Fas基因、EGR2基因、PPP2R2D基因、PSGL-1基因、KDM6A基因以及TET2基因)：

[1349] a”)删除临近’NNAGAAW’(N各自独立地为A、T、G或C；W为A或T)序列的5’端和/或3’端的连续的1bp-25bp(例如17bp-23bp)的核苷酸序列区中的一个或多个核苷酸；

[1350] b”)利用不同于野生型基因的一个或多个核苷酸置换临近’NNAGAAW’序列的5’端和/或3’端的连续的1bp-25bp(例如17bp-23bp)的核苷酸序列区中的一个或多个核苷酸；

[1351] c”)将一个或多个核苷酸插入至临近’NNAGAAW’序列的5’端和/或3’端的连续的1bp-25bp(例如17bp-23bp)的核苷酸序列区；或者

[1352] d”)选自a”)-c”)中的两者以上的组合。

[1353] 例如，当编辑蛋白是由脑膜炎奈瑟菌衍生而来的Cas9蛋白时，PAM序列可为5’-NNNNGATT-3’(N各自独立地为A、T、G或C)；待切割的核苷酸序列区(靶位点)可为靶基因内临近5’-NNNNGATT-3’序列的5’端和/或3’端的连续的1bp-25bp(例如17bp-23bp或21bp-23bp)的核苷酸序列区。

[1354] 可提供因免疫调节基因中的如下修饰而产生的经人工操纵的免疫调节基因(例如经人工操纵的PD-1基因、CTLA-4基因、TNFAIP3基因、DGKA基因、DGKZ基因、Fas基因、EGR2基因、PPP2R2D基因、PSGL-1基因、KDM6A基因以及TET2基因)：

[1355] a”)删除临近’NNNNGATT’(N各自独立地为A、T、G或C)序列的5’端和/或3’端的连续的1bp-25bp(例如17bp-23bp)的核苷酸序列区中的一个或多个核苷酸；

[1356] b”)利用不同于野生型基因的一个或多个核苷酸置换临近’NNNNGATT’序列的5’端和/或3’端的连续的1bp-25bp(例如17bp-23bp)的核苷酸序列区中的一个或多个核苷酸；

[1357] c”)将一个或多个核苷酸插入至临近’NNNNGATT’序列的5’端和/或3’端的连续的1bp-25bp(例如17bp-23bp)的核苷酸序列区；或者

[1358] d”)选自a”)-c”)中的两者以上的组合。

[1359] 例如，当编辑蛋白是由金黄色葡萄球菌衍生而来的Cas9蛋白时，PAM序列可为5’-NNGRR(T)-3’(N各自独立地为A、T、G或C；R为A或G；(T)为可任选包含的任何序列)；待切割的核苷酸序列区(靶位点)可为靶基因内临近5’-NNGRR(T)-3’序列的5’端和/或3’端的连续的1bp-25bp(例如17bp-23bp或21bp-23bp)的核苷酸序列区。

[1360] 可提供因免疫调节基因中的如下修饰而产生的经人工操纵的免疫调节基因(例如经人工操纵的PD-1基因、CTLA-4基因、TNFAIP3基因、DGKA基因、DGKZ基因、Fas基因、EGR2基因、PPP2R2D基因、PSGL-1基因、KDM6A基因以及TET2基因)：

[1361] a”)删除临近5’-NNGRR(T)-3’(N各自独立地为A、T、G或C；R为A或G；(T)为可任选包含的任何序列)序列的5’端和/或3’端的连续的1bp-25bp(例如17bp-23bp)的核苷酸序列区中的一个或多个核苷酸；

[1362] b”)利用不同于野生型基因的一个或多个核苷酸置换临近5’-NNGRR(T)-3’序列的5’端和/或3’端的连续的1bp-25bp(例如17bp-23bp)的核苷酸序列区中的一个或多个核苷酸；

[1363] c”)将一个或多个核苷酸插入至临近5’-NNGRR(T)-3’序列的5’端和/或3’端的连续的1bp-25bp(例如17bp-23bp)的核苷酸序列区；或者

[1364] d”)选自a”)-c”)中的两者以上的组合。

[1365] 例如,当利用Cpf1蛋白作为编辑蛋白时,PAM序列可为5'-TTN-3'(N为A、T、G或C);待切割的核苷酸序列区(靶位点)可为靶基因内临近5'-TTN-3'序列的5'端和/或3'端的连续的10bp-30bp(例如15bp-26bp、17bp-30bp或17bp-26bp)的核苷酸序列区。

[1366] Cpf1蛋白可为由如下微生物衍生而来的Cpf1蛋白:例如俭菌(GWC2011_GWC2_44_17)、毛螺菌(MC2017)、*Butyrivibrio proteoclasticus*、*Peregrinibacteria bacterium*(GW2011_GWA_33_10)、氨基酸球菌属(BV3L6)、猕猴卟啉单胞菌、毛螺菌(ND2006)、*Porphyromonas crevioricanis*、解糖脲普雷沃菌、*Moraxella bovoculi*(237)、*Smiihella* sp.(SC_K08D17)、稻田钩端螺旋体、毛螺菌(MA2020)、新凶手弗朗西斯菌(U112)、*Candidatus Methanoplasma termitum*或挑剔真杆菌等;例如为由如下微生物衍生而来的Cpf1蛋白:俭菌(GWC2011_GWC2_44_17)、*Peregrinibacteria bacterium*(GW2011_GWA_33_10)、氨基酸球菌属(BV3L6)、猕猴卟啉单胞菌、毛螺菌(ND2006)、*Porphyromonas crevioricanis*、解糖脲普雷沃菌、*Moraxella bovoculi*(237)、稻田钩端螺旋体、毛螺菌(MA2020)、新凶手弗朗西斯菌(U112)、*Candidatus Methanoplasma termitum*或挑剔真杆菌,但所述微生物不限于此。

[1367] 可提供因免疫调节基因中的如下修饰而产生的经人工操纵的免疫调节基因(例如经人工操纵的PD-1基因、CTLA-4基因、TNFAIP3基因、DGKA基因、DGKZ基因、Fas基因、EGR2基因、PPP2R2D基因、PSGL-1基因、KDM6A基因以及TET2基因):

[1368] a'''')删除临近5'-TTN-3'(N为A、T、G或C)序列的5'端和/或3'端的连续的10bp-30bp(例如15bp-26bp)的核苷酸序列区中的一个或多个核苷酸;

[1369] b'''')利用不同于野生型基因的一个或多个核苷酸置换临近5'-TTN-3'序列的5'端和/或3'端的连续的10bp-30bp(例如15bp-26bp)的核苷酸序列区中的一个或多个核苷酸;

[1370] c'''')将一个或多个核苷酸插入至临近5'-TTN-3'序列的5'端和/或3'端的连续的10bp-30bp(例如15bp-26bp)的核苷酸序列区;或者

[1371] d'''')选自a'''')-c'''')中的两者以上的组合。

[1372] 在另一实施方式中,当免疫调节因子为蛋白时,

[1373] 经人工操纵的蛋白可包括参与由引导核酸-编辑蛋白复合体的直接/间接作用形成的新免疫应答或改变的免疫应答的全部蛋白。

[1374] 例如,免疫调节因子可为但不限于由通过引导核酸-编辑蛋白复合体进行人工操纵的免疫调节基因表达的蛋白,或者因所述蛋白的活性影响而导致表达升高或降低的其它蛋白。

[1375] 经人工操纵的免疫调节蛋白可具有与经人工操纵的免疫调节基因相对应的氨基酸组成和活性。

[1376] 在实施方式中,可提供(i)表达特征被改变的经人工操纵的蛋白。

[1377] 例如,可在免疫调节基因的核酸序列中位于原型间隔区邻近基序(PAM)序列中或临近PAM序列的5'端和/或3'端的连续的1bp-50bp、1bp-40bp、1bp-30bp并优选3bp-25bp的核苷酸序列区中包含具有如下特征中的一种的蛋白修饰:

[1378] 由于一个或多个核苷酸删除或插入而造成表达量降低或升高;

[1379] 由于被不同于野生型基因的一个或多个核苷酸置换而造成表达量降低或升高;

[1380] 由于一个或多个外源核苷酸插入而造成表达水平降低或升高,或者融合蛋白表达或特定蛋白独立表达;以及

[1381] 受上述蛋白的表达特征影响的第三蛋白的表达水平降低或升高。

[1382] 可提供(ii)结构特征被改变的经人工操纵的蛋白。

[1383] 例如,可在免疫调节基因的核酸序列中位于原型间隔区邻近基序(PAM)序列中或临近PAM序列的5'端和/或3'端的连续的1bp-50bp、1bp-40bp、1bp-30bp并优选3bp-25bp的核苷酸序列区中包含具有如下特征中的一种的蛋白修饰:

[1384] 由于一个或多个核苷酸删除或插入而造成密码子改变、氨基酸改变和三维结构改变;

[1385] 由于被不同于野生型基因的一个或多个核苷酸置换而造成密码子改变、氨基酸改变和随后的三维结构改变;

[1386] 由于一个或多个外源核苷酸插入而造成密码子改变、氨基酸改变和三维结构改变,或者与特定蛋白的融合结构或特定蛋白分离的独立结构;以及

[1387] 受上述结构特征改变的蛋白影响的第三蛋白的密码子改变、氨基酸改变和三维结构改变。

[1388] 可提供(iii)免疫功能特征被改变的经人工操纵的蛋白。

[1389] 例如,可在免疫调节基因的核酸序列中位于原型间隔区邻近基序(PAM)序列中或临近PAM序列的5'端和/或3'端的连续的1bp-50bp、1bp-40bp、1bp-30bp并优选3bp-25bp的核苷酸序列区中包含具有如下特征中的一种的蛋白修饰:

[1390] 由于一个或多个核苷酸删除或插入而造成的蛋白修饰使得特定免疫功能活化或失活或者引入新免疫功能;

[1391] 由于被不同于野生型基因的一个或多个核苷酸置换而造成的蛋白修饰使得特定免疫功能活化或失活或者引入新免疫功能;

[1392] 由于一个或多个外源核苷酸插入而造成的蛋白修饰使得特定免疫功能活化或失活或者引入新免疫功能(具体而言,可通过特定蛋白的融合表达或独立表达而在现有免疫功能中引入第三功能);以及

[1393] 受上述免疫功能特征改变的蛋白影响的第三蛋白的功能改变。

[1394] 此外,可包含通过在构成免疫调节基因的核酸序列中的一个或多个核苷酸的化学修饰进行人工操纵的蛋白。

[1395] 例如,可通过甲基化、乙酰化、磷酸化、泛素化、ADP-核糖基化、十四烷基化和糖基化改变蛋白的表达特征、结构特征和免疫功能特征中的一种或多种特征。

[1396] 例如,可通过核苷酸的化学修饰使第三蛋白与基因的核酸序列结合,从而赋予第三结构和功能。

[1397] 在另一实施方式中,提供了经人工操纵的细胞(其为利用“引导核酸-编辑蛋白复合体”获得的作为产物的免疫系统因子)。

[1398] 经人工操纵的细胞可为包含如下中的一种或多种的细胞:

[1399] 通过引导核酸-编辑蛋白复合体进行人工操纵的免疫调节基因;以及

[1400] 参与由引导核酸-编辑蛋白复合体的直接/间接作用形成的新免疫应答或改变的免疫应答的蛋白。在实施方式中,细胞可为免疫细胞或干细胞。

[1401] 这些细胞具有由上述经人工操纵的免疫调节基因和/或蛋白表现出的免疫功能以及参与与之对应的细胞内机制的后续功能。

[1402] 在又一实施方式中,提供了诱导期望的免疫应答的组合物(其为利用“引导核酸-编辑蛋白复合体”获得的作为产物的免疫系统因子)。该组合物可称为药物组合物或治疗组合物。

[1403] 诱导期望的免疫应答的组合物可包含如下中的一种或多种作为活性成分:

[1404] 通过引导核酸-编辑蛋白复合体进行人工操纵的免疫调节基因;

[1405] 参与由引导核酸-编辑蛋白复合体的直接/间接作用形成的新免疫应答或改变的免疫应答的蛋白;以及

[1406] 包含所述免疫调节基因和/或蛋白的细胞。

[1407] 这些组合物具有由上述经人工操纵的免疫调节基因、蛋白和/或细胞表现出的免疫功能以及参与与之对应的机体中多种机制的后续功能。

[1408] 可将组合物(例如细胞治疗剂)用于免疫相关疾病(例如癌症)的预防和/或治疗。

[1409] [制备方法]

[1410] 作为本发明的一个实施方式,提供了经人工操纵的免疫调节因子以及用于制备包含该经人工操纵的免疫调节因子的免疫细胞的方法。

[1411] 关于经人工操纵的免疫调节因子的描述可参照上文。下文中将以前操纵的免疫细胞的代表性实施方式作为重点对上述方法进行描述。

[1412] -细胞培养

[1413] 为产生经操纵的免疫细胞,首先从健康供体收集细胞,并对其进行培养。例如,使用已知方法从供体收集免疫细胞(例如T细胞、NK细胞、NKT细胞等)并在合适的细胞培养基中进行培养。

[1414] 如之后所描述的,对由所培养的免疫细胞表达的一些免疫调节因子进行选择 and 人工操纵。例如,对PD-1、CTLA-4、TNFAIP3、DGKA (Dgk α)、DGKZ (Dgk ζ)、Fas、EGR2、PPP2R2D、PSGL-1和/或TET2基因进行基因操纵。可参考上文所述的关于基因操纵的具体描述。

[1415] 或者,可对免疫细胞进行转染并随后进行培养,来生成经操纵的免疫细胞。

[1416] -产生经功能性操纵的免疫细胞的方法

[1417] 可通过插入或去除作为免疫调节因子的蛋白来产生经功能性操纵的免疫细胞。

[1418] 可通过对作为免疫调节因子的基因进行修饰来产生经功能性操纵的免疫细胞。

[1419] 可通过野生型受体或免疫调节基因的敲减(KD)或敲除(KO)来产生经功能性操纵的免疫细胞。敲减或敲除是指通过靶基因切割、DNA转录抑制剂和RNA翻译抑制剂(例如互补microRNA等)等来阻遏基因表达。

[1420] 敲减或敲除可通过microRNA实现。

[1421] 敲减或敲除可优选借助本发明的引导核酸-编辑蛋白复合体实现。

[1422] 敲减或敲除可借助基因剪刀通过NHEJ实现。

[1423] 敲减或敲除可借助基因剪刀和核苷酸模板通过HR实现。

[1424] 在一个实例中,敲减或敲除可通过切割PD-1、CTLA-4、TNFAIP3、DGKA (Dgk α)、DGKZ (Dgk ζ)、Fas、EGR2、PPP2R2D、PSGL-1和/或TET2基因的特定靶位点来实现。

[1425] 经功能性操纵的免疫细胞可包含靶区域的修饰,例如:

- [1426] 基因编码区内或紧邻基因编码区的一个或多个核苷酸的插入或删除(例如NHEJ介导的插入或删除);
- [1427] 包含基因的至少一部分的基因组序列的删除(例如NHEJ介导的删除);以及
- [1428] 通过靶向基因的非编码区(例如启动子区)由酶促失活编辑蛋白介导的基因的敲减或敲除的修饰。
- [1429] 此外,可通过转染野生型受体或免疫调节基因来产生经功能性操纵的免疫细胞。
- [1430] 转染方法包括插入含有靶基因的游离体(episome)或融合至基因组中。
- [1431] 转染可通过插入游离体实现。游离体载体是指在真核生物的细胞核中作为外源基因发挥作用的载体,其不与基因组融合。具体而言,游离体可为质粒。
- [1432] 转染可借助HR利用引导核酸-编辑蛋白复合体和核苷酸模板实现。
- [1433] 此外,可通过在转染不同的野生型受体或免疫调节基因的同时敲除野生型受体或免疫调节基因来产生经功能性操纵的免疫细胞。转染方法包括插入含有靶基因的游离体或融合至基因组中。
- [1434] 具体而言,待转染的基因可融合至待敲除的基因的位置处。
- [1435] 转染可通过插入游离体实现。
- [1436] 转染可借助HR利用引导核酸-编辑蛋白复合体和核苷酸模板实现。
- [1437] -产生增补人工结构的免疫细胞的方法
- [1438] 可通过直接将人工结构以蛋白的形式增补至免疫细胞来产生增补人工结构的免疫细胞。
- [1439] 可通过转染编码人工结构的基因来产生增补人工结构的免疫细胞。
- [1440] 转染方法包括插入含有靶基因的游离体或融合至基因组中。
- [1441] 转染可通过插入游离体实现。
- [1442] 转染可借助HR利用引导核酸-编辑蛋白复合体和核苷酸模板实现。
- [1443] 在实施方式中,提供了使得免疫细胞中一个或多个免疫调节基因失活的方法,该方法包括将引导核酸和编辑蛋白导入免疫细胞(转染)。
- [1444] 在实施方式中,提供了制备转染的免疫细胞的方法,该方法包括将引导核酸和编辑蛋白导入免疫细胞(转染)。
- [1445] -产生经混合型操纵的免疫细胞的方法
- [1446] 可通过在产生经功能性操纵的免疫细胞的方法和产生增补人工结构的免疫细胞的方法中描述的对蛋白或基因进行操纵的方法来产生经混合型操纵的免疫细胞。
- [1447] 产生经混合型操纵的免疫细胞的方法包括敲除野生型受体或免疫调节因子,或者实施转染。该步骤可根据在产生经功能性操纵的免疫细胞的方法中描述的方法来实现。
- [1448] 产生经混合型操纵的免疫细胞的方法包括转染人工结构。该步骤可根据在产生增补人工结构的免疫细胞的方法中描述的方法来实现。
- [1449] 产生经混合型操纵的免疫细胞的方法的优选方面是在实施人工结构转染的同时敲除免疫细胞的野生型受体。
- [1450] 在一个实例中,所述方法为在实施人工结构转染的同时敲除免疫细胞中的PD-1和CTLA-4。
- [1451] 在另一实例中,产生经混合型操纵的免疫细胞的方法为在实施人工结构转染的同

时敲除TNFAIP3 (A20)、DGK- α 、DGK- ζ 、Fas、EGR2、PPP2R2D、PSGL-1、KDM6A和/或TET2。

[1452] 具体而言,待转染的基因可融合至与待敲除的基因相同的位置处。

[1453] 可使用已知方法(例如通常利用重组载体的方法)产生上述经操纵的免疫细胞。

[1454] -用于免疫细胞的重组表达载体

[1455] 术语“表达靶序列”是指用于修饰靶细胞的蛋白或基因的手段,或者编码新表达的基因的核苷酸序列。在本发明的实施方式中,表达靶序列可包含编码引导核酸和编辑蛋白的序列以及用于表达引导核酸和编辑蛋白的额外序列。

[1456] 术语“重组载体”是指其功能在于将表达靶序列运送至靶细胞的运送体,包括例如质粒、游离体载体、病毒载体等。

[1457] 术语“重组表达载体”为重组载体的实施方式,是指人工构建的载体,该载体表现出在靶细胞中表达与重组载体相连的表达靶序列的功能。

[1458] 用于免疫细胞的重组表达载体是一种重组表达载体,是用于对免疫细胞的基因或蛋白进行修饰从而将免疫细胞作为经操纵的免疫细胞进行表达的手段,或者是用于编码新表达的基因的重组表达载体。

[1459] 用于免疫细胞的重组表达载体包括如上所述的用于表达引导核酸-编辑蛋白复合体的重组表达载体。

[1460] 在实施方式中,可通过仅转染一种类型的用于免疫细胞的重组表达载体获得经操纵的免疫细胞。

[1461] 可通过转染两种以上类型的用于免疫细胞的重组表达载体获得经操纵的免疫细胞。

[1462] 可根据最终表达的核苷酸序列的大小将重组表达载体分为适当数目的重组表达载体来设计用于免疫细胞的重组表达载体。

[1463] (经功能性操纵的重组表达载体)

[1464] 在实施方式中,提供了用于制备经功能性操纵的免疫细胞的重组表达载体。

[1465] 在实施方式中,经功能性操纵的重组表达载体包含用于敲除野生型受体或免疫调节因子基因的重组核苷酸序列。

[1466] 用于敲除基因的重组表达载体包括如上所述的用于表达引导核酸-编辑蛋白复合体的重组表达载体。具体而言,gRNA的靶序列可与野生型受体的核苷酸序列或免疫调节因子的核苷酸序列具有互补性。此外,如有需要,重组表达载体可包含待插入至通过引导核酸-编辑蛋白复合体进行切割的位置处的核苷酸模板。

[1467] 在实施方式中,经功能性操纵的重组表达载体包含用于转染野生型受体或免疫调节因子基因的重组核苷酸序列。

[1468] 具体而言,经功能性操纵的重组表达载体可为游离体载体。游离体载体可包含用于基因表达的启动子。

[1469] 在实施方式中,经功能性操纵的重组表达载体可为具有与活体基因组融合的功能的载体。具体而言,经功能性操纵的重组表达载体可为病毒载体。具体而言,优选的病毒载体可为腺相关病毒载体。

[1470] 在实施方式中,经功能性操纵的重组表达载体可包含与插入靶位点具有同源性的核苷酸序列。该核苷酸序列可为待在HR过程中插入的核苷酸模板。该核苷酸模板可与待通

过引导核酸-编辑蛋白复合体进行切割的区域的序列具有同源性。

[1471] 在实施方式中,经功能性操纵的重组表达载体可以在相同的载体中或者在不同的载体中独立地包含用于表达如上所述的引导核酸-编辑蛋白复合体的序列。

[1472] 另一方面,经功能性操纵的重组表达载体包含用于敲除野生型受体或免疫调节因子基因的重组核苷酸序列,或者用于转染不同的野生型受体或免疫调节因子基因的重组核苷酸序列。

[1473] 用于敲除基因的重组核苷酸序列包含如上所述的用于表达引导核酸-编辑蛋白复合体的重组表达载体的核苷酸序列。具体而言,gRNA的靶序列可与免疫调节因子的核苷酸序列具有互补性。

[1474] 用于转染的重组表达载体可为游离体载体。具体而言,游离体载体可包含用于基因表达的启动子。

[1475] 用于转染的重组表达载体可具有与活体基因组融合的功能。

[1476] 用于转染的重组表达载体可为病毒载体。具体而言,优选的病毒载体可为腺相关病毒载体。

[1477] 用于转染的重组表达载体可包含与插入靶位点具有同源性的核苷酸序列。该核苷酸序列可为待在HR过程中插入的核苷酸模板。该核苷酸模板可与待通过引导核酸-编辑蛋白复合体进行切割的区域的序列具有同源性。

[1478] 此外,经功能性操纵的重组表达载体可包括如上所述的用于表达引导核酸-编辑蛋白复合体的重组表达载体。

[1479] (增补人工结构的重组表达载体)

[1480] 在实施方式中,提供了用于产生增补人工结构的免疫细胞的重组表达载体。

[1481] 增补人工结构的重组表达载体包含用于转染免疫调节因子基因的重组核苷酸序列。

[1482] 在一个实例中,增补人工结构的重组表达载体可为游离体载体。游离体载体是指在真核生物的细胞核中作为外源基因发挥作用的载体,其不融合至基因组。具体而言,游离体载体可包含用于基因表达的启动子。

[1483] 在另一实例中,增补人工结构的重组表达载体可具有融合至活体基因组的功能。

[1484] 具体而言,增补人工结构的重组表达载体可为病毒载体。具体而言,优选的病毒载体可为腺相关病毒载体。

[1485] 此外,增补人工结构的重组表达载体可包含与插入位点具有同源性的核苷酸序列。该核苷酸序列可为待在HR过程中插入的核苷酸模板。该核苷酸模板可与待通过引导核酸-编辑蛋白复合体进行切割的序列具有同源性。

[1486] 此外,增补人工结构的重组表达载体可包括如上所述的用于表达引导核酸-编辑蛋白复合体的重组表达载体。

[1487] (经混合型操纵的重组表达载体)

[1488] 经混合型操纵的重组表达载体可包含用于敲除野生型受体或免疫调节因子基因以及转染具有人工结构的不同基因的重组核苷酸序列。

[1489] 用于敲除基因的重组核苷酸序列可包含如上所述的用于表达引导核酸-编辑蛋白复合体的重组表达载体的核苷酸序列。具体而言,gRNA的靶序列可与免疫调节因子的核苷

酸序列具有互补性。

[1490] 在一个实例中,用于转染的重组表达载体可为游离体载体。具体而言,游离体载体可包含用于基因表达的启动子。

[1491] 在另一实例中,用于转染的重组表达载体可具有融合至活体基因组的功能。

[1492] 具体而言,用于转染的重组表达载体可为病毒载体。具体而言,优选的病毒载体可为腺相关病毒载体。

[1493] 此外,用于转染的重组表达载体可包含与插入靶位点具有同源性的核苷酸序列。该核苷酸序列可为待在HR过程中插入的核苷酸模板。该核苷酸模板可与待通过引导核酸-编辑蛋白复合体进行切割的序列具有同源性。

[1494] 此外,经功能性操纵的重组表达载体可包括如上所述的用于表达引导核酸-编辑蛋白复合体的重组表达载体。

[1495] 同时,在本发明的具体示例性实施方式中,提供了用于制备免疫细胞的方法,所述免疫细胞包含通过引导核酸-编辑蛋白复合体进行人工操纵的免疫调节因子。

[1496] 在实施方式中,所述方法可为对所述细胞中靶核酸的序列进行改变来制备经操纵的免疫细胞的方法,所述方法包括使细胞与(a)靶向PD-1、CTLA-4、TNFAIP3、DGKA (Dgk α)、DGKZ (Dgk ζ)、Fas、EGR2、PPP2R2D、PSGL-1和/或TET2基因的一种或多种引导核酸(例如gRNA)以及(b)编辑蛋白(例如Gas9蛋白)接触。

[1497] 接触方法可为借助传统方法将引导核酸和编辑蛋白直接导入免疫细胞。

[1498] 接触方法可为将编码引导核酸和编辑蛋白的各个DNA分子以包含于一个载体或不同载体中的状态下导入免疫细胞。

[1499] 接触方法可使用载体实现。载体可为病毒载体。病毒载体可为例如逆转录病毒、腺相关病毒载体。

[1500] 在该方法中,可采用本领域已知的多种方法(例如电穿孔、脂质体、病毒载体、纳米粒子以及蛋白易位结构域(PTD)融合蛋白方法等),用于运送至免疫细胞中。

[1501] 所述方法可进一步包括:将靶向不同基因的gRNA导入细胞,或将编码此类gRNA的核酸导入细胞。

[1502] 所述方法可在体内或在体外实施,例如离体实施。

[1503] 例如,所述接触可在体外实施,并可在接触后使经接触的细胞返回至受试者体内。

[1504] 所述方法可在体内使用有机体或免疫细胞,例如从人体分离的免疫细胞或人工产生的免疫细胞。在一个实例中,可包括使来自患有癌症的受试者的细胞进行接触。

[1505] 用于上述方法的免疫细胞可为由包括灵长类动物(例如人、猴等)和啮齿类动物(例如小鼠、大鼠等)在内的哺乳动物衍生而来的免疫细胞。例如,免疫细胞可为NKT细胞、NK细胞、T细胞等。具体而言,免疫细胞可为增补了免疫受体的经操纵的免疫细胞(例如增补了嵌合抗体受体(CAR)或经操纵的T细胞受体(TCR))。可对免疫细胞进行操纵,使得在如下基因中的一种或多种基因中导入免疫细胞的靶位置突变之前、之后或同时表达免疫受体(例如TCR或CAR):PD-1、CTLA-4、TNFAIP3、DGKA (Dgk α)、DGKZ (Dgk ζ)、Fas、EGR2、PPP2R2D、PSGL-1和/或TET2基因。

[1506] 可在适合免疫细胞的培养基中实施所述方法,所述培养基可含有血清(例如胎牛血清或人血清)、白细胞介素-2(IL-2)、胰岛素、IFN- γ 、IL-4、IL-7、GM-CSF、IL-10、IL-15、

TGF- β 以及TNF- α ;或者合适的培养基可含有对于增殖和存活而言必要的因子,包括本领域技术人员已知的其它细胞生长添加剂(例如最小必需培养基、RPMI 1640培养基或X-vivo-10、X-vivo-15、X-vivo-20(Lonza)),但所述培养基不限于此。

[1507] [用途]

[1508] 在实施方式中,本发明涉及使用免疫疗法手段用于疾病治疗的用途,所述免疫疗法手段包括向受试者给予经人工操纵的细胞(例如经基因操纵的免疫细胞或干细胞)。

[1509] 待治疗的受试者可为包括灵长类动物(例如人、猴等)和啮齿类动物(例如小鼠、大鼠等)在内的哺乳动物。

[1510] 药物组合物

[1511] 本发明的一个实施方式为利用免疫应答用于治疗疾病的组合物,例如为含有经人工操纵的免疫调节基因或包含该免疫调节基因的免疫细胞的组合物。组合物可称为治疗组合物、药物组合物或细胞治疗剂。

[1512] 在实施方式中,组合物可含有免疫细胞。

[1513] 在实施方式中,组合物可含有用于免疫调节的经人工操纵的基因和/或其表达的蛋白。

[1514] 免疫细胞可为已分化的免疫细胞。

[1515] 免疫细胞可从骨髓或脐带血提取。

[1516] 免疫细胞可为干细胞。具体而言,干细胞可为造血干细胞。

[1517] 组合物可含有经操纵的免疫细胞。

[1518] 组合物可含有经功能性操纵的免疫细胞。

[1519] 组合物可含有增补人工结构的免疫细胞。

[1520] 在另一实施方式中,组合物可进一步含有额外因子。

[1521] 组合物可含有抗原结合剂。

[1522] 组合物可含有细胞因子。

[1523] 组合物可含有细胞因子的促分泌素(secretagogue)或抑制剂。

[1524] 组合物可含有适用于将经操纵的免疫细胞递送至体内的运载体。

[1525] 组合物中含有的免疫细胞对患者而言可为同种异体的。

[1526] 治疗方法

[1527] 本发明的另一实施方式为在患者中对疾病进行治疗的方法,所述方法包括将所述组合物(其中,所述组合物的产生和组合物的有效量已在上文中进行描述)给予有需要的患者。

[1528] 在实施方式中,所述方法可为使用过继性免疫疗法的方法。

[1529] -待治疗的疾病

[1530] 过继性免疫疗法可用于治疗任何特定疾病。

[1531] 所述任何特定疾病可为免疫疾病。具体而言,免疫疾病可为其中免疫能力(immune competence)恶化的疾病。

[1532] 免疫疾病可为自身免疫疾病。

[1533] 例如,自身免疫疾病可包括移植物抗宿主病(GVHD)、系统性红斑狼疮、乳糜泻、I型糖尿病、graves病、炎症性肠病、牛皮癣、类风湿性关节炎、多发性硬化症等。

[1534] 免疫疾病可为增生性疾病。

[1535] 例如,免疫疾病可为恶性血液病(hematologic malignancy)或实体癌(solid cancer)。代表性恶性血液病包括急性淋巴性白血病(ALL)、急性髓性白血病(AML)、慢性髓性白血病(CML)、慢性嗜酸性粒细胞白血病(CEL)、骨髓增生异常综合征(MDS)、非霍奇金淋巴瘤(NHL)和多发性骨髓瘤(MM)。实体瘤的实例包括胆道癌、膀胱癌、骨和软组织癌、脑肿瘤、乳腺癌、子宫颈癌、结肠癌、结肠腺癌、结肠直肠癌、硬纤维瘤、胚胎癌、子宫内膜癌、食道癌、胃癌、胃腺癌、多形性成胶质细胞瘤、妇科肿瘤、头颈部鳞状细胞癌、肝癌、肺癌、恶性黑色素瘤、骨肉瘤、卵巢癌、胰腺癌、胰腺管腺癌、原发性星形细胞肿瘤、原发性甲状腺癌、前列腺癌、肾癌、肾细胞癌、横纹肌肉瘤、皮肤癌、软组织肉瘤、睾丸生殖细胞瘤、尿路上皮细胞癌、子宫肉瘤、子宫癌等。

[1536] 待治疗的主题疾病可为多种癌症(包括实体恶性肿瘤和恶性血液病)。

[1537] 例如,可治疗的癌症类型包括:乳腺癌、前列腺癌、胰腺癌、结肠和直肠腺癌;所有形式的支气管肺癌(包括鳞状细胞癌、腺癌、小细胞肺癌和非小细胞肺癌);骨髓瘤;黑色素瘤;肝细胞瘤;神经母细胞瘤;乳头状瘤;胺前体摄取和脱羧细胞瘤(apudoma);迷芽瘤(choristoma);鳃裂囊肿;恶性类癌综合征(malignant carcinoid syndrome);类癌心脏病;以及癌(carcinoma)(例如Walker、基底细胞癌、基底鳞状癌、Brown-Pierce、导管癌、Ehrlich肿瘤、Krebs-2癌、Merkel细胞癌、粘液性(mucinous)癌、非小细胞肺癌、燕麦细胞癌、乳头状癌、硬性癌(scirrhous)、细支气管癌、支气管原性(bronchogenic)癌、鳞状细胞癌和移行细胞癌)。

[1538] 例如,可治疗的其它类型的癌症包括:组织细胞增生(histiocytocytic)失调;白血病;恶性组织细胞增生(histiocytosis);霍奇金病;非霍奇金淋巴瘤;浆细胞瘤、网状内皮瘤;黑色素瘤;肾细胞癌;软骨母细胞瘤;软骨瘤;软骨肉瘤;纤维瘤;纤维肉瘤;巨细胞瘤;组织细胞瘤、脂肪瘤、脂肪肉瘤;间皮瘤;粘液瘤;粘液肉瘤;骨瘤;骨肉瘤;脊索瘤、颅咽管瘤;无性细胞瘤;错构瘤(hamartoma);间质瘤;中肾瘤;肌肉瘤;adamantio;牙骨质瘤;牙瘤;畸胎瘤;胸腺瘤;和滋养细胞瘤。

[1539] 进而,认为也可治疗如下类型的癌症:腺瘤;胆管瘤;胆脂瘤;圆柱瘤;囊腺癌;囊腺瘤;颗粒细胞瘤;两性母细胞瘤;肝细胞瘤;汗腺瘤;胰岛细胞瘤;Leydig细胞瘤;乳头状瘤;Sertoli细胞瘤;卵泡膜细胞(theca cell)瘤;子宫平滑肌瘤;子宫肉瘤;成肌细胞瘤;肌瘤;肌肉瘤;横纹肌瘤;横纹肌肉瘤;室管膜瘤;神经节瘤;神经胶质瘤;成神经管细胞瘤;脑膜瘤;神经鞘瘤;神经母细胞瘤;神经上皮瘤;神经纤维瘤;神经瘤;副神经节瘤;非嗜铬性副神经节瘤;和多形性成胶质细胞瘤。

[1540] 可被治疗的癌症类型还包括:血管角化瘤;血管淋巴样增生伴嗜酸粒细胞增多;血管硬化;血管瘤病;血管球瘤;血管内皮瘤;血管瘤;血管外皮细胞瘤;血管肉瘤;淋巴管瘤;淋巴管肌瘤;淋巴管肉瘤;松果体瘤;癌肉瘤;软骨肉瘤;叶状囊肉瘤;纤维肉瘤;血管肉瘤;平滑肌肉瘤;白血病性肉瘤;脂肪肉瘤;淋巴管肉瘤;肌肉瘤;粘液肉瘤;卵巢癌;横纹肌肉瘤;肉瘤;赘生物;神经纤维瘤病;和宫颈非典型增生。

[1541] 此外,任何特定疾病可为其病原体已知但治疗未知的难治性疾病。

[1542] 难治性疾病可为病毒感染疾病。

[1543] 难治性疾病可为由朊病毒病原体引起的疾病。

[1544] 任何特定疾病可为细菌疾病。

[1545] 任何特定疾病可为炎性疾病。

[1546] 任何特定疾病可为衰老相关疾病。

[1547] -增强免疫力的治疗

[1548] 对于免疫力显著降低的患者而言,即便轻微感染也可导致致命后果。免疫力降低由免疫细胞功能下降、免疫细胞产量减少等造成。作为用于增强免疫力来治疗免疫功能恶化的方法,其可为活化正常免疫细胞产生的永久性治疗方法,也可对免疫细胞进行瞬时注入的暂时治疗方法。

[1549] 增强免疫力的治疗可意在将治疗组合物注入患者体内来永久增强免疫力。

[1550] 增强免疫力的治疗可为将治疗组合物注入患者的特定身体部分的方法。具体而言,特定的身体部分可为具有供应免疫细胞源的组织的部分。

[1551] 增强免疫力的治疗可为在患者体内制造新免疫细胞源。具体而言,在一个实例中,治疗组合物可包含干细胞。具体而言,干细胞可为造血干细胞。

[1552] 增强免疫力的治疗可意在将治疗组合物注入患者体内来暂时增强免疫力。

[1553] 增强免疫力的治疗可为将治疗组合物注入患者体内。

[1554] 具体而言,优选的治疗组合物可含有分化的免疫细胞。

[1555] 用于增强免疫力的治疗的治疗组合物可含有特定数量的免疫细胞。

[1556] 可根据免疫力的恶化程度来改变所述特定数量。

[1557] 可根据身体体积来改变所述特定数量。

[1558] 可根据患者释放的细胞因子量来调整所述特定数量。

[1559] -难治性疾病的治疗

[1560] 免疫细胞操纵技术可提供针对例如HIV、朊病毒以及癌症等的病原体(其完全治疗尚且未知)的疾病治疗方法。虽然这些疾病的病原体已知,但在很多情况下由于存在抗体难以形成、疾病迅速发展且患者免疫系统失活以及病原体在体内具有潜伏期等问题,难以对这些疾病进行治疗。经操纵的免疫细胞可为解决这些问题的强有力手段。

[1561] 可通过将治疗组合物注入体内来实施难治性疾病的治疗。具体而言,优选的治疗组合物可含有经操纵的免疫细胞。此外,可将治疗组合物注入特定的身体部分。

[1562] 经操纵的免疫细胞可为对目标疾病的病原体具有改善的识别能力的免疫细胞。

[1563] 经操纵的免疫细胞可为免疫应答强度或活性增强的免疫细胞。

[1564] -基因修正治疗

[1565] 除借助外源提取的免疫细胞的治疗方法外,还存在通过操纵活体基因直接影响免疫细胞表达的治疗方法。可通过将用于基因操纵的基因修正组合物直接注入体内来实施该治疗方法。

[1566] 基因修正组合物可含有引导核酸-编辑蛋白复合体。

[1567] 可将基因修正组合物注入特定的身体部分。

[1568] 特定的身体部分可为免疫细胞源,例如骨髓。

[1569] 本发明的一个实施方式涉及通过将有效量的含有如上所述的经人工操纵的免疫系统的成分的组合物给予受试者来治疗免疫相关疾病的方法。

[1570] 在任何实施方式中,所述治疗方法提供以离体重组方式(例如借助病毒载体)操纵

或修饰的细胞群的用途。在进一步的实施方式中,经修饰的细胞群为同源、同种异体或自体细胞。在上述任何实施方式中,如上所述,可进一步将经操纵或修饰的细胞群与药学上可接受的运载体、稀释剂或赋形剂进行配制。

[1571] 待给药的受试者可为包括灵长类动物(例如人、猴等)和啮齿类动物(例如小鼠、大鼠等)在内的哺乳动物。

[1572] 给药是指将目标递送至受试者,而不考虑递送的途径或模式。给药可连续或间歇地进行,并可通过胃肠外方式进行。

[1573] 在某些实施方式中,与辅助治疗剂共同给药可包括以任何顺序和任何剂量方案(dosage regimen)同时和/或顺序递送多种试剂(例如,一种或多种细胞因子与抗原特异性重组宿主T细胞和抗原表达细胞一起;免疫阻遏疗法,例如钙调磷酸酶抑制剂、皮质类固醇、微管抑制剂、低剂量霉酚酸前药,或它们的任何组合)。

[1574] 在某些实施方式中,给药可重复多次,并进行数周、数月或长达两年的时间。

[1575] 如药物领域技术人员所确定的,可将组合物以适于待治疗或预防的疾病或病症的方式进行给药。根据例如患者的健康状况、患者的大小(即体重、质量、身体面积)、患者疾病的类型和严重程度、活性成分的特定形式和给药方法等因素确定组合物给药的合适剂量、合适持续时间和频率。

[1576] 例如,可以通过任何方便的方式(例如注射、输液、植入、移植等)实施组合物给药。给药途径可选自于皮下、皮内、瘤内、结内(intranodal)、髓内、肌内、静脉内、淋巴管内和腹膜内给药等。

[1577] 组合物的单次剂量(用于实现期望效果的药物有效量)可选自于约 10^4 - 10^9 个细胞/kg待给药的受试者的体重的范围内的全部整数数值(例如约 10^5 - 10^6 个细胞/kg体重,但所述剂量不限于此;可考虑待给药的受试者的年龄、健康状况和体重;同时进行治疗的种类,(如果有的话)治疗频率和期望效果的性质,来适当规定组合物的单次剂量。

[1578] 当经人工操纵的免疫调节因子由本申请文件中的方法或组合物调控时,涉及免疫细胞的存活、增殖、持久性、细胞毒性、细胞因子释放和/或浸润等方面的免疫效力可得到改善。

[1579] 实施例

[1580] 在下文中,将参考实施例更详细地描述本发明。

[1581] 本领域技术人员应理解的是,这些实施例仅用于更详细地描述本发明,而本发明的范围不受这些实施例的限制。

[1582] 实施例1:细胞的制备(活化和培养)和转染

[1583] 在补充有10% (v/v) 胎牛血清(GeneA11)的RPMI 1640培养基中培养Jurkat细胞(ATCC TIB-152;人T细胞永生化细胞系)。在37°C和5%CO₂的条件下在培养箱中对细胞进行孵育。

[1584] 在补充有10% (v/v) 胎牛血清(GeneA11)和/或IL-2 (50U/mL)、IL-7 (5ng/mL)以及IL-15 (5ng/mL) (PEPROTECH)的X-VIVO 15培养基(Lonza)中对人初始(naive)T细胞(STEMCELL Technology)进行培养。为活化细胞,将培养基中的细胞浓度分别保持为 1×10^6 个细胞/mL。

[1585] 以3:1的比例(珠:细胞;珠和细胞的数目)添加CD2/CD3/CD28珠(抗CD2/3/

CD28Dynabeads;Miltenyi Biotec),并在37℃和5%CO₂的条件下在培养箱中对细胞进行孵育。在进行细胞活化72小时后,用磁铁去除CD2/CD3/CD28珠,并在无珠条件下进一步对细胞培养12-24小时。

[1586] 为了寻找能够高效敲除特定基因的gRNA,如下文实施例2和实施例3所述,通过电转将1μg体外转录的sgRNA和4μg Cas9蛋白(Toolgen,韩国)导入1×10⁶个Jurkat细胞(体外)。使用Neon转染系统(ThermoFisher Scientific,Grand Island,NY)的10μL枪头在如下条件下导入基因:

[1587] Jurkat(缓冲液R):1,400V,20ms,2个脉冲。

[1588] 类似地,通过电转将1μg gRNA和4μg Cas9蛋白(Toolgen,韩国)导入1×10⁶个人原代细胞来敲除T细胞中的特定基因。该研究中使用的gRNA为体外转录并经AP(碱性磷酸酶)处理的sgRNA;或者化学合成的crRNA和tracrRNA复合体(Integrated DNA Technologies)。对于电转,使用Neon转染系统(ThermoFisher Scientific,Grand Island,NY)的10μL枪头在如下条件下导入基因:

[1589] 人原代T细胞(缓冲液T):1,550V,10ms,3个脉冲。

[1590] 将细胞涂板于500μL无抗生素培养基上,并在37℃和5%CO₂下在培养箱中进行培养。

[1591] 实施例2:sgRNA的设计和合成

[1592] 2.1 sgRNA的设计

[1593] 使用CRISPR RGEN工具(Institute for Basic Science,韩国)选择如下基因的CRISPR/Cas9靶区域并通过脱靶测试进行评估:人PD-1基因(PDCD1;NCBI登记号NM_005018.2)、CTLA-4基因(NCBI登记号NM_001037631.2)、A20基因(TNFAIP3;NCBI登记号NM_001270507.1)、DGKα基因(NCBI登记号NM_001345.4)、DGKζ基因(NCBI登记号NM_001105540.1)、EGR2基因(NCBI登记号NM_000399.4)、PPP2R2D基因(NCBI登记号NM_001291310.1)、PSGL-1基因(NCBI登记号NM_003006.4)、TET2基因(NCBI登记号NM_017628.4)、FAS基因(NCBI登记号XM_006717819.3、XM_011539764.2、NM_152871.3或NM_152872.3)以及KDM6A基因(NCBI登记号NM_001291415.1、NM_001291416.1、NM_001291418.1、NM_001291417.1、NM_001291421.1或NM_021140.3)。对于CRISPR/Cas9靶区域,在人基因组(GRCh38/hg38)内选择除中靶序列区之外的不具有0bp、1bp或2bp错配位点的DNA序列作为sgRNA的靶区域。

[1594] 2.2 sgRNA的合成

[1595] 通过对两条互补寡核苷酸进行退火和延伸来对sgRNA合成模板进行PCR扩增。

[1596] 此时使用的靶区域序列、用于对其进行扩增的引物序列和由此获得的sgRNA所靶向的DNA靶序列(包含PAM)在下表2中描述。

[1597] 使用T7 RNA聚合酶(New England Biolabs)对模板DNA(除去靶序列3'端的'NGG')实施体外转录,根据制造商的说明合成RNA,随后使用Turbo DNAase(Ambion)去除模板DNA。利用Expin Combo试剂盒(GeneA11)和异丙醇沉淀纯化转录的RNA。

[1598] 在使用T细胞的实验中,为将sgRNA的免疫原性和降解最小化,使用碱性磷酸酶(New England Biolabs)从利用上述方法合成的sgRNA去除5'端磷酸残基,随后利用Expin Combo试剂盒(GeneA11)和异丙醇沉淀再次纯化RNA。此外,在一些T细胞实验中使用了化学

合成的sgRNA(Trilink)以及化学合成的dgRNA(Integrated DNA Technologies)。

[1599] 在某些实施例中使用的化学合成的sgRNA为具有2'-OMe和硫代磷酸酯修饰的sgRNA。

[1600] 例如,该实施例中使用的DGK α sgRNA#11具有5'-2'OMe(**C(ps)U(ps)**

C(ps)) UCA AGC UGA GUG GGU CCG UUU UAG AGC UAG AAA
UAG CAA GUU AAA AUA AGG CUA GUC CGU UAU CAA CUU
GAA AAA GUG GCA CCG AGU CGG UGC 2'OMe(U(ps)U(ps)U(ps)
U-3' (2'OMe=2'-甲基RNA;ps=硫代磷酸酯)的结构。

[1601] 在另一实例中,该实施例中使用的A20 sgRNA#1为**GCUUGUGG**

CGCUGAAAACGAAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAU
AAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGU
GCUUUUUUUU (粗体部分为与靶序列区杂交的序列;用于其它靶基因或其它靶

序列的sgRNA具有该粗体序列作为靶序列(只是T变为U);或者,可对上述序列进行修饰,其中,向序列3'端的三个核苷酸和5'端的三个核苷酸进行2'-OMe修饰和硫代磷酸酯骨架导入。

[1602] [表2]

[1603]

基因	#	DNA 靶序列	正向引物序列	反向引物	序列号
A20	1	CTTGTGGCGCTGAAAAC GAACGG	GAAATTAATACGACTCA CTATAGCTTGTGGCGCT GAAAACGAAGTTTTAGA GCTAGAAATAGC	AAAAAAGCAC CGACTCGGTGC CACTTTTCAAG TTGATAACGGA CTAGCCTATTT TAACTTGCTATT TCTAGCTCTAAA AC	SEQ ID NO 1
	2	ATGCCACTTCTCAGTAC ATGTGG	GAAATTAATACGACTCA CTATAGATGCCACTTCTC AGTACATGGTTTTAGAG CTAGAAATAGC		SEQ ID NO 2
	3	GCCACTTCTCAGTACAT GTGGGG	GAAATTAATACGACTCA CTATAGGCCACTTCTCA GTACATGTGGTTTTAGA GCTAGAAATAGC		SEQ ID NO 3
	4	GCCCCACATGTACTGAG AAGTGG	GAAATTAATACGACTCA CTATAGGCCCCACATGT ACTGAGAAGGTTTTAGA GCTAGAAATAGC		SEQ ID NO 4
	5	TCAGTACATGTGGGGCG TTCAGG	GAAATTAATACGACTCA CTATAGTCAGTACATGT GGGGCGTTCGTTTTAGA GCTAGAAATAGC		SEQ ID NO 5
	6	GGGCGTTCAGGACACA GACTTGG	GAAATTAATACGACTCA CTATAGGGGCGTTCAGG ACACAGACTGTTTTAGA GCTAGAAATAGC		SEQ ID NO 6
	7	CACAGACTTGGTACTGA GGAAGG	GAAATTAATACGACTCA CTATAGCACAGACTTGG TACTGAGGAGTTTTAGA GCTAGAAATAGC		SEQ ID NO 7
	8	GGCGCTGTTCAAGCACGC TCAAGG	GAAATTAATACGACTCA CTATAGGGCGCTGTTCA GCACGCTCAGTTTTAGA GCTAGAAATAGC		SEQ ID NO 8

[1604]

	9	CACGCAACTTTAAATTC CGCTGG	GAAATTAATACGACTCA CTATAGCACGCAACTTT AAATTCGCGTTTTAGA GCTAGAAATAGC		SEQ ID NO 9
	10	CGGGGCTTTGCTATGAT ACTCGG	GAAATTAATACGACTCA CTATAGCGGGGCTTTGC TATGATACTGTTTTAGAG CTAGAAATAGC		SEQ ID NO 10
	11	GGCTTCCACAGACACA CCCATGG	GAAATTAATACGACTCA CTATAGGGCTTCCACAG ACACACCCAGTTTTAGA GCTAGAAATAGC		SEQ ID NO 11
	12	TGAAGTCCACTTCGGGC CATGGG	GAAATTAATACGACTCA CTATAGTGAAGTCCACT TCGGGCCATGTTTTAGA GCTAGAAATAGC		SEQ ID NO 12
基因	#	DNA 靶序列	正向引物序列	反向引物	序列号
	1	CTGTACGACACGGACA GAAATGG	GAAATTAATACGACTCA CTATAGCTGTACGACAC GGACAGAAAGTTTTAGA GCTAGAAATAGC		SEQ ID NO 13
	2	TGTACGACACGGACAG AAATGGG	GAAATTAATACGACTCA CTATAGTGTACGACACG GACAGAAATGTTTTAGA GCTAGAAATAGC		SEQ ID NO 14
	3	CACGGACAGAAATGGG ATCCTGG	GAAATTAATACGACTCA CTATAGCACGGACAGAA ATGGGATCCGTTTTAGA GCTAGAAATAGC		SEQ ID NO 15
	4	GATGCGAGTGGCTGAAT ACCTGG	GAAATTAATACGACTCA CTATAGGATGCGAGTGG CTGAATACCGTTTTAGA GCTAGAAATAGC		SEQ ID NO 16

[1605]

DGK α	5	GAGTGGCTGAATACCTG GATTGG	GAAATTAATACGACTCA CTATAGGAGTGGCTGAA TACCTGGATGTTTTAGA GCTAGAAATAGC	AAAAAAGCAC CGACTCGGTGC CACTTTTCAAG TTGATAACGGA CTAGCCTTATT TAACTTGCTATT TCTAGCTCTAAA AC	SEQ ID NO 17
	6	AGTGGCTGAATACCTGG ATTGGG	GAAATTAATACGACTCA CTATAGAGTGGCTGAAT ACCTGGATTGTTTTAGA GCTAGAAATAGC		SEQ ID NO 18
	7	ATTGGGATGTGTCTGA GCTGAGG	GAAATTAATACGACTCA CTATAGATTGGGATGTG TCTGAGCTGGTTTTAGA GCTAGAAATAGC		SEQ ID NO 19
	8	ATGAAAGAGATTGACT ATGATGG	GAAATTAATACGACTCA CTATAGATGAAAGAGAT TGACTATGAGTTTTAGA GCTAGAAATAGC		SEQ ID NO 20
	9	CTCTGTCTCTCAAGCTG AGTGGG	GAAATTAATACGACTCA CTATAGCTCTGTCTCTCA AGCTGAGTGTTTTAGAG CTAGAAATAGC		SEQ ID NO 21
	10	TCTCTCAAGCTGAGTGG GTCCGG	GAAATTAATACGACTCA CTATAGTCTCTCAAGCT GAGTGGGTCGTTTTAGA GCTAGAAATAGC		SEQ ID NO 22
	11	CTCTCAAGCTGAGTGG GTCCGGG	GAAATTAATACGACTCA CTATAGCTCTCAAGCTG AGTGGGTCCGTTTTAGA GCTAGAAATAGC		SEQ ID NO 23
	12	CAAGCTGAGTGGGTCC GGGCTGG	GAAATTAATACGACTCA CTATAGCAAGCTGAGTG GGTCCGGGCGTTTTAGA GCTAGAAATAGC		SEQ ID NO 24
基因	#	DNA 靶序列	正向引物序列	反向引物	序列号

[1606]

1	TTGACATGACTGGAGA GAAGAGG	GAAATTAATACGACTCA CTATAGTTGACATGACT GGAGAGAAGGTTT TAGA GCTAGAAATAGC	SEQ ID NO 25
2	GACTGGAGAGAAGAGG TCGTTGG	GAAATTAATACGACTCA CTATAGGACTGGAGAGA AGAGGTCGTGTTT TAGA GCTAGAAATAGC	SEQ ID NO 26
3	GAGACGGGAGCAAAGC TGCTGGG	GAAATTAATACGACTCA CTATAGGAGACGGGAG CAAAGCTGCTGTTT TAG AGCTAGAAATAGC	SEQ ID NO 27
4	AGAGACGGGAGCAAAG CTGCTGG	GAAATTAATACGACTCA CTATAGAGAGACGGGA GCAAAGCTGCGTTT TAG AGCTAGAAATAGC	SEQ ID NO 28
5	TGGTTTCTAGGTGCAGA GACGGG	GAAATTAATACGACTCA CTATAGTGGTTTCTAGGT GCAGAGACGTTT TAGAG CTAGAAATAGC	SEQ ID NO 29
6	TAAGTGAAGGTCTGGTT TCTAGG	GAAATTAATACGACTCA CTATAGTAAGTGAAGGT CTGGTTTCTGTTT TAGAG CTAGAAATAGC	SEQ ID NO 30
7	TGCCCATGTAAGTGAAG GTCTGG	GAAATTAATACGACTCA CTATAGTGCCCATGTAA GTGAAGGTCGTTT TAGA GCTAGAAATAGC	SEQ ID NO 31
8	GAACTTGCCCATGTAAG TGAAGG	GAAATTAATACGACTCA CTATAGGAAGTTGCCCA TGTAAGTGAGTTT TAGA GCTAGAAATAGC	SEQ ID NO 32

[1607]

9	TCCATTGACCCTCAGTACCCTGG	GAAATTAATACGACTCACTATAGTCCATTGACCCTCAGTACCCGTTTTAGAGCTAGAAATAGC	SEQ ID NO 33
10	TATGCCTTCTGGGTAGCAGCTGG	GAAATTAATACGACTCACTATAGTATGCCTTCTGGGTAGCAGCGTTTTAGAGCTAGAAATAGC	SEQ ID NO 34
11	TGAGTGCAGGCATCTTGCAAGGG	GAAATTAATACGACTCACTATAGTGAGTGCAGGCATCTTGCAAGTTTTAGAGCTAGAAATAGC	SEQ ID NO 35
12	GAGTGCAGGCATCTTGAAGGGG	GAAATTAATACGACTCACTATAGGAGTGCAGGCATCTTGCAAGTTTTAGAGCTAGAAATAGC	SEQ ID NO 36
13	GATGAGGCTGTGGTTGAGCTGG	GAAATTAATACGACTCACTATAGGATGAGGCTGTGGTTGAAGCGTTTTAGAGCTAGAAATAGC	SEQ ID NO 37
14	CCACTGGCCACAGGACCCTGGG	GAAATTAATACGACTCACTATAGCCACTGGCCACAGGACCCCTGTTTTAGAGCTAGAAATAGC	SEQ ID NO 38
15	GGGACATGGTGCACACACCCAGG	GAAATTAATACGACTCACTATAGGGGACATGGTGACACACCCGTTTTAGAGCTAGAAATAGC	SEQ ID NO 39
16	GAGTACAGGTGGTCCAGGTCAGG	GAAATTAATACGACTCACTATAGGAGTACAGGTGTCCAGGTCGTTTTAGAGCTAGAAATAGC	SEQ ID NO 40

[1608]

EGR2	17	GCGGAGAGTACAGGTG GTCCAGG	GAAATTAATACGACTCA CTATAGGCGGAGAGTAC AGGTGGTCCGTTT TAGA GCTAGAAATAGC	AAAAAAGCAC CGACTCGGTGC CACTTTTCAAG TTGATAACGGA CTAGCCTTATTT TAACTTGCTATT TCTAGCTCTAAA AC	SEQ ID NO 41
	18	GCGGTGGCGGAGAGTA CAGGTGG	GAAATTAATACGACTCA CTATAGGCGGTGGCGG AGAGTACAGGGTTT TAG AGCTAGAAATAGC		SEQ ID NO 42
	19	TCTCCTGCACAGCCAGA ATAAGG	GAAATTAATACGACTCA CTATAGTCTCCTGCACA GCCAGAATAGTTT TAGA GCTAGAAATAGC		SEQ ID NO 43
	20	ACGCAGAAGGGTCCTG GTAGAGG	GAAATTAATACGACTCA CTATAGACGCAGAAGG GTCCTGGTAGGTTT TAG AGCTAGAAATAGC		SEQ ID NO 44
	21	AGGTGGTGGGTAGGCC AGAGAGG	GAAATTAATACGACTCA CTATAGAGGTGGTGGGT AGGCCAGAGGTTT TAGA GCTAGAAATAGC		SEQ ID NO 45
	22	CCCAAGCCAGCCACGG ACCCAGG	GAAATTAATACGACTCA CTATAGCCCAAGCCAGC CACGGACCCGTTT TAGA GCTAGAAATAGC		SEQ ID NO 46
	23	ACCTGGGTCCGTGGCTG GCTTGG	GAAATTAATACGACTCA CTATAGACCTGGGTCCG TGGCTGGCTGTTT TAGA GCTAGAAATAGC		SEQ ID NO 47
	24	AAGAGACCTGGGTCCG TGGCTGG	GAAATTAATACGACTCA CTATAGAAGAGACCTGG GTCCGTGGCGTTT TAGA GCTAGAAATAGC		SEQ ID NO 48

[1609]

25	GGATCATTGGGAAGAG ACCTGGG	GAAATTAATACGACTCA CTATAGGGATCATTGGG AAGAGACCTGTTTTAGA GCTAGAAATAGC	SEQ ID NO 49
26	GGGATCATTGGGAAGA GACCTGG	GAAATTAATACGACTCA CTATAGGGGATCATTGG GAAGAGACCGTTTTAGA GCTAGAAATAGC	SEQ ID NO 50
27	CAGGATAGTCTGGGATC ATTGGG	GAAATTAATACGACTCA CTATAGCAGGATAGTCT GGGATCATTGTTTTAGA GCTAGAAATAGC	SEQ ID NO 51
28	GGAAAGAATCCAGGAT AGTCTGG	GAAATTAATACGACTCA CTATAGGGGAAAGAATCC AGGATAGTCGTTTTAGA GCTAGAAATAGC	SEQ ID NO 52
29	CAGTGCCAGAGAGACC TACATGG	GAAATTAATACGACTCA CTATAGCAGTGCCAGAG AGACCTACAGTTTTAGA GCTAGAAATAGC	SEQ ID NO 53
30	CTGTACCATGTAGGTCT CTCTGG	GAAATTAATACGACTCA CTATAGCTGTACCATGT AGGTCTCTCGTTTTAGA GCTAGAAATAGC	SEQ ID NO 54
31	AGAGACCTACATGGTAC AGCTGG	GAAATTAATACGACTCA CTATAGAGAGACCTACA TGGTACAGCGTTTTAGA GCTAGAAATAGC	SEQ ID NO 55
32	CTGGGCCAGCTGTACCA TGTAGG	GAAATTAATACGACTCA CTATAGCTGGGCCAGCT GTACCATGTGTTTTAGA GCTAGAAATAGC	SEQ ID NO 56

[1610]

	33	AGGGAAAGGGCTTACG GTCTGGG	GAAATTAATACGACTCA CTATAGAGGGAAAGGG CTTACGGTCTGTTTAGA GCTAGAAATAGC		SEQ ID NO 57
	34	CAGGGAAAGGGCTTAC GGTCTGG	GAAATTAATACGACTCA CTATAGCAGGGAAAGG GCTTACGGTCGTTTATAG AGCTAGAAATAGC		SEQ ID NO 58
基因	#	DNA 靶序列	正向引物序列	反向引物	序列号
	5	TCTGGAGATCTTCTTGC AACAGG	GAAATTAATACGACTCA CTATAGTCTGGAGATCT TCTTGCAACGTTTATAGA GCTAGAAATAGC		SEQ ID NO 59
	6	CTCCGGTTCATGACTTT GAAAGG	GAAATTAATACGACTCA CTATAGCTCCGGTTCAT GACTTTGAAGTTTATAGA GCTAGAAATAGC		SEQ ID NO 60
	7	GTCTTCCATCTTCGTCTT TCAGG	GAAATTAATACGACTCA CTATAGGTCTTCCATCTT CGTCTTTCGTTTATAGAGC TAGAAATAGC		SEQ ID NO 61
	8	GAAGACTTCGAGACCC ATTTAGG	GAAATTAATACGACTCA CTATAGGAAGACTTCGA GACCCATTTGTTTATAGA GCTAGAAATAGC		SEQ ID NO 62
	9	TCGAGACCCATTTAGGA TCACGG	GAAATTAATACGACTCA CTATAGTCGAGACCCAT TTAGGATCAGTTTATAGA GCTAGAAATAGC		SEQ ID NO 63
	10	GTAGCGCCGTGATCCT AAATGGG	GAAATTAATACGACTCA CTATAGGTAGCGCCGTG ATCCTAAATGTTTATAGA GCTAGAAATAGC		SEQ ID NO 64

[1611]

PPP2R2D	11	CGTAGCGCCGTGATCCT AAATGG	GAAATTAATACGACTCA CTATAGCGTAGCGCCGT GATCCTAAAGTTTTAGA GCTAGAAATAGC	SEQ ID NO 65
	12	CATTTAGGATCACGGCG CTACGG	GAAATTAATACGACTCA CTATAGCATTAGGATC ACGGCGCTAGTTTTAGA GCTAGAAATAGC	SEQ ID NO 66
	13	GGTCCCAATATTGAAGC CCATGG	GAAATTAATACGACTCA CTATAGGGTCCCAATAT TGAAGCCCAGTTTTAGA GCTAGAAATAGC	SEQ ID NO 67
	14	GATCCATGGGCTTCAAT ATTGGG	GAAATTAATACGACTCA CTATAGGATCCATGGGC TTCAATATTGTTTTAGAG CTAGAAATAGC	SEQ ID NO 68
	15	AGATCCATGGGCTTCAA TATTGG	GAAATTAATACGACTCA CTATAGAGATCCATGGG CTTCAATATGTTTTAGAG CTAGAAATAGC	SEQ ID NO 69
	16	GCTTCTACCATAAGATC CATGGG	GAAATTAATACGACTCA CTATAGGCTTCTACCAT AAGATCCATGTTTTAGA GCTAGAAATAGC	SEQ ID NO 70
	17	CGCTTCTACCATAAGAT CCATGG	GAAATTAATACGACTCA CTATAGCGCTTCTACCAT AAGATCCAGTTTTAGAG CTAGAAATAGC	SEQ ID NO 71
	18	GCATTTGCAAAAATTCG CCGTGG	GAAATTAATACGACTCA CTATAGGCATTTGCAAA AATTCGCCGTTTTAGA GCTAGAAATAGC	SEQ ID NO 72

[1612]

19	ATGACCTGAGAATTAAT TTATGG	GAAATTAATACGACTCA CTATAGATGACCTGAGA ATTAATTTAGTTTTAGAG CTAGAAATAGC	SEQ ID NO 73
20	CCATGCACTCCCAGACA TCGTGG	GAAATTAATACGACTCA CTATAGCCATGCACTCC CAGACATCGGTTTTAGA GCTAGAAATAGC	SEQ ID NO 74
21	GCACTGGTGCGGGTGG AACTCGG	GAAATTAATACGACTCA CTATAGGCACTGGTGCG GGTGGAAGTGTTTTAGA GCTAGAAATAGC	SEQ ID NO 75
22	ACACGTTGCACTGGTGC GGGTGG	GAAATTAATACGACTCA CTATAGACACGTTGCAC TGGTGCGGGTTTTAGA GCTAGAAATAGC	SEQ ID NO 76
23	CGAACACGTTGCACTGG TGCGGG	GAAATTAATACGACTCA CTATAGCGAACACGTTG CACTGGTGCGTTTTAGA GCTAGAAATAGC	SEQ ID NO 77
24	ACGAACACGTTGCACTG GTGCGG	GAAATTAATACGACTCA CTATAGACGAACACGTT GCACTGGTGTTTTAGA GCTAGAAATAGC	SEQ ID NO 78
25	TGTAGACGAACACGTTG CACTGG	GAAATTAATACGACTCA CTATAGTGTAGACGAAC ACGTTGCACGTTTTAGA GCTAGAAATAGC	SEQ ID NO 79
26	GCGCATGTCACACAGG CGGATGG	GAAATTAATACGACTCA CTATAGGCGCATGTCAC ACAGGCGGAGTTTTAGA GCTAGAAATAGC	SEQ ID NO 80

[1613]

	27	AGGAGCGCATGTCACA CAGGCGG	GAAATTAATACGACTCA CTATAGAGGAGCGCATG TCACACAGGGTTTTAGA GCTAGAAATAGC		SEQ ID NO 81
	28	CCGAGGAGCGCATGTC ACACAGG	GAAATTAATACGACTCA CTATAGCCGAGGAGCG CATGTCACACGTTTTAG AGCTAGAAATAGC		SEQ ID NO 82
	29	CCTGTGTGACATGCGCT CCTCGG	GAAATTAATACGACTCA CTATAGCCTGTGTGACA TGCGCTCCTGTTTTAGA GCTAGAAATAGC		SEQ ID NO 83
基因	#	DNA 靶序列	正向引物序列	反向引物	序列号
	1	CGACTGGCCAGGGCGC CTGTGGG	GAAATTAATACGACTCA CTATAGCGACTGGCCAG GGCGCCTGTGTTTTAGA GCTAGAAATAGC		SEQ ID NO 84
	2	ACCGCCCAGACGACTG GCCAGGG	GAAATTAATACGACTCA CTATAGACCGCCCAGAC GACTGGCCAGTTTTAGA GCTAGAAATAGC		SEQ ID NO 85
	3	CACCGCCCAGACGACT GGCCAGG	GAAATTAATACGACTCA CTATAGCACCGCCCAGA CGACTGGCCGTTTTAGA GCTAGAAATAGC		SEQ ID NO 86
	4	GTCTGGGCGGTGCTAC AACTGGG	GAAATTAATACGACTCA CTATAGGTCTGGGCGGT GCTACAACGTGTTTTAGA GCTAGAAATAGC		SEQ ID NO 87
	5	CTACAAC TGGGCTGGCG GCCAGG	GAAATTAATACGACTCA CTATAGCTACAAC TGGG CTGGCGGCCGTTTTAGA GCTAGAAATAGC		SEQ ID NO 88

[1614]

PD-1	6	CACCTACCTAAGAACC ATCCTGG	GAAATTAATACGACTCA CTATAGCACCTACCTAA GAACCATCCGTTTTAGA GCTAGAAATAGC	AAAAAAGCAC CGACTCGGTGC CACTTTTCAAG TTGATAACGGA CTAGCCTTATT TAACTTGCTATT TCTAGCTCTAAA AC	SEQ ID NO 89
	7	CGGTCACCACGAGCAG GGCTGGG	GAAATTAATACGACTCA CTATAGCGGTCACCACG AGCAGGGCTGTTTTAGA GCTAGAAATAGC		SEQ ID NO 90
	8	GCCCTGCTCGTGGTGAC CGAAGG	GAAATTAATACGACTCA CTATAGGCCCTGCTCGT GGTGACCGAGTTTTAGA GCTAGAAATAGC		SEQ ID NO 91
	9	CGGAGAGCTTCGTGCTA AACTGG	GAAATTAATACGACTCA CTATAGCGGAGAGCTTC GTGCTAAACGTTTTAGA GCTAGAAATAGC		SEQ ID NO 92
	10	CAGCTTGTCCTGCTGGT TGCTGG	GAAATTAATACGACTCA CTATAGCAGCTTGCCG TCTGGTTGCGTTTTAGA GCTAGAAATAGC		SEQ ID NO 93
	11	AGGCGGCCAGCTTGTC GTCTGG	GAAATTAATACGACTCA CTATAGAGGCGGCCAG CTTGTCCTCGTTTTAGA GCTAGAAATAGC		SEQ ID NO 94
	12	CCGGGCTGGCTGCGGT CCTCGGG	GAAATTAATACGACTCA CTATAGCCGGGCTGGCT GCGGTCCTCGTTTTAGA GCTAGAAATAGC		SEQ ID NO 95
	13	CGTTGGGCAGTTGTGTG ACACGG	GAAATTAATACGACTCA CTATAGCGTTGGGCAGT TGTGTGACAGTTTTAGA GCTAGAAATAGC		SEQ ID NO 96
基因	#	DNA 靶序列	正向引物序列	反向引物	序列号

[1615]

CTLA-4	1	CATAAGCCATGGCTG CCTTGG	GAAATTAATACGACTCA CTATAGCATAAAGCCAT GGCTTGCCTGTTTATAG GCTAGAAATAGC	AAAAAAGCAC CGACTCGGTGC CACTTTTCAAG TTGATAACGGA CTAGCCTATTT TAACTTGCTATT TCTAGCTCTAAA AC	SEQ ID NO 97
	2	CCTTGGATTTCAGCGG ACAAGG	GAAATTAATACGACTCA CTATAGCCTTGGATTTC GCGGCACAGTTTATAG CTAGAAATAGC		SEQ ID NO 98
	3	CCTTGTGCCGCTGAAAT CCAAGG	GAAATTAATACGACTCA CTATAGCCTTGTGCCGC TGAAATCCAGTTTATAG GCTAGAAATAGC		SEQ ID NO 99
	4	CACTCACCTTTCAGAA GACAGG	GAAATTAATACGACTCA CTATAGCACTCACCTTTC CAGAAGACGTTTATAG CTAGAAATAGC		SEQ ID NO 100
	5	TTCCATGCTAGCAATGC ACGTGG	GAAATTAATACGACTCA CTATAGTTCCATGCTAG CAATGCACGGTTTATAG GCTAGAAATAGC		SEQ ID NO 101
	6	GGCCACGTGCATTGCTA GCATGG	GAAATTAATACGACTCA CTATAGGGCCACGTGCA TTGCTAGCAGTTTATAG GCTAGAAATAGC		SEQ ID NO 102
	7	GGCCAGCCTGCTGTGG TACTGG	GAAATTAATACGACTCA CTATAGGGCCAGCCTG CTGTGGTACGTTTATAG GCTAGAAATAGC		SEQ ID NO 103
	8	AGGTCCGGGTGACAGT GCTTCGG	GAAATTAATACGACTCA CTATAGAGGTCCGGGTG ACAGTGCTTGTATAG GCTAGAAATAGC		SEQ ID NO 104

[1616]

	9	CCGGGTGACAGTGCTTC GGCAGG	GAAATTAATACGACTCA CTATAGCCGGGTGACAG TGCTTCGGCGTTTTAGA GCTAGAAATAGC		SEQ ID NO 105
	10	CTGTGCGGCAACCTACA TGATGG	GAAATTAATACGACTCA CTATAGCTGTGCGGCAA CCTACATGAGTTTTAGA GCTAGAAATAGC		SEQ ID NO 106
	11	CAACTATTCCCATC ATGTAGG	GAAATTAATACGACTCA CTATAGCAACTCATTCC CCATCATGTGTTTTAGA GCTAGAAATAGC		SEQ ID NO 107
	12	CTAGATGATTCCATCTG CACGGG	GAAATTAATACGACTCA CTATAGCTAGATGATTC CATCTGCACGTTTTAGA GCTAGAAATAGC		SEQ ID NO 108
基因	#	DNA 靶序列	正向引物序列	反向引物	序列号
	1	GGCTAGGAGTCAGCGA CATATGG	GAAATTAATACGACTCA CTATAGGGCTAGGAGTC AGCGACATAGTTTTAGA GCTAGAAATAGC		SEQ ID NO 109
	2	GCTAGGAGTCAGCGAC ATATGGG	GAAATTAATACGACTCA CTATAGGCTAGGAGTCA GCGACATATGTTTTAGA GCTAGAAATAGC		SEQ ID NO 110
	3	CTAGGAGTCAGCGACA TATGGGG	GAAATTAATACGACTCA CTATAGCTAGGAGTCAG CGACATATGGTTTTAGA GCTAGAAATAGC		SEQ ID NO 111
	4	GTA CTGTGTAGCCAGG ATGCTGG	GAAATTAATACGACTCA CTATAGGTACTGTGTAG CCAGGATGCGTTTTAGA GCTAGAAATAGC		SEQ ID NO 112

[1617]

D G K Z	5	ACGAGCACTCACCAGC ATCCTGG	GAAATTAATACGACTCA CTATAGACGAGCACTCA CCAGCATCCGTTTTAGA GCTAGAAATAGC	AAAAAAGCAC CGACTCGGTGC CACTTTTCAAG TTGATAACGGA CTAGCCTATTT TAACTTGCTATT TCTAGCTCTAAA AC	SEQ ID NO 113
	6	AGGCTCCAGGAATGTCC GCGAGG	GAAATTAATACGACTCA CTATAGAGGCTCCAGGA ATGTCCGCGGTTTTAGA GCTAGAAATAGC		SEQ ID NO 114
	7	ACTTACCTCGCGGACAT TCCTGG	GAAATTAATACGACTCA CTATAGACTTACCTCGC GGACATTCCGTTTTAGA GCTAGAAATAGC		SEQ ID NO 115
	8	CACCCTGGGCACTTACC TCGCGG	GAAATTAATACGACTCA CTATAGCACCCTGGGCA CTTACCTCGGTTTTAGAG CTAGAAATAGC		SEQ ID NO 116
	9	GTGCCGTACAAAGGTTG GCTGGG	GAAATTAATACGACTCA CTATAGGTGCCGTACAA AGGTTGGCTGTTTTAGA GCTAGAAATAGC		SEQ ID NO 117
	10	GGTGCCGTACAAAGGTT GGCTGG	GAAATTAATACGACTCA CTATAGGGTGCCGTACA AAGGTTGGCGTTTTAGA GCTAGAAATAGC		SEQ ID NO 118
	11	CTCTCCTCAGTACCACA GCAAGG	GAAATTAATACGACTCA CTATAGCTCTCCTCAGTA CCACAGCAGTTTTAGAG CTAGAAATAGC		SEQ ID NO 119
	12	CCTGGGGCCTCCGGGC GCGGAGG	GAAATTAATACGACTCA CTATAGCCTGGGGCCTC CGGGCGCGGTTTTAGA GCTAGAAATAGC		SEQ ID NO 120

[1618]

	13	AGTACTCACCTGGGGCC TCCGGG	GAAATTAATACGACTCA CTATAGAGTACTCACCT GGGGCCTCCGTTT TAGA GCTAGAAATAGC		SEQ ID NO 121
	14	AGGGTCTCCAGCGGCC CTCCTGG	GAAATTAATACGACTCA CTATAGAGGGTCTCCAG CGGCCCTCCGTTT TAGA GCTAGAAATAGC		SEQ ID NO 122
	15	GCAAGTACTTACGCCTC CTTGGG	GAAATTAATACGACTCA CTATAGGCAAGTACTTA CGCCTCCTTGTTT TAGAG CTAGAAATAGC		SEQ ID NO 123
	16	TTGCGGTACATCTCCAG CCTGGG	GAAATTAATACGACTCA CTATAGTTGCGGTACAT CTCCAGCCTGTTT TAGA GCTAGAAATAGC		SEQ ID NO 124
	17	TTTGCGGTACATCTCCA GCCTGG	GAAATTAATACGACTCA CTATAGTTTGCGGTACA TCTCCAGCCGTTT TAGA GCTAGAAATAGC		SEQ ID NO 125
基因	#	DNA 靶序列	正向引物序列	反向引物	序列号
	1	GCAAAACCTGTCCACT CTTATGG	GAAATTAATACGACTCA CTATAGGCAAAACCTGT CCACTCTTAGTTT TAGAG CTAGAAATAGC		SEQ ID NO 126
	2	TTGGTGCCATAAGAGTG GACAGG	GAAATTAATACGACTCA CTATAGTTGGTGCCATA AGAGTGGACGTTT TAGA GCTAGAAATAGC		SEQ ID NO 127
	3	GGTGCAAGTTTCTTAT ATGTTGG	GAAATTAATACGACTCA CTATAGGGTGCAAGTTT CTTATATGTGTTT TAGAG CTAGAAATAGC		SEQ ID NO 128

[1619]

4	ACCTGATGCATATAAT AATCAGG	GAAATTAATACGACTCA CTATAGACCTGATGCAT ATAATAATCGTTTTAGA GCTAGAAATAGC	SEQ ID NO 129
5	ACCTGATTATTATATGC ATCAGG	GAAATTAATACGACTCA CTATAGACCTGATTATTA TATGCATCGTTTTAGAG CTAGAAATAGC	SEQ ID NO 130
6	CAGAGCACCAGAGTGC CGTCTGG	GAAATTAATACGACTCA CTATAGCAGAGCACCAG AGTGCCGTCGTTTTAGA GCTAGAAATAGC	SEQ ID NO 131
7	AGAGCACCAGAGTGCC GTCTGGG	GAAATTAATACGACTCA CTATAGAGAGCACCAGA GTGCCGTCTGTTTTAGA GCTAGAAATAGC	SEQ ID NO 132
8	AGAGTGCCGTCTGGGTC TGAAGG	GAAATTAATACGACTCA CTATAGAGAGTGCCGTC TGGGTCTGAGTTTTAGA GCTAGAAATAGC	SEQ ID NO 133
9	AGGAAGGCCGTCATTCT CAGGG	GAAATTAATACGACTCA CTATAGAGGAAGGCCGT CCATTCTCAGTTTTAGAG CTAGAAATAGC	SEQ ID NO 134
10	GGATAGAACCAACCAT GTTGAGG	GAAATTAATACGACTCA CTATAGGGATAGAACCA ACCATGTTGGTTTTAGA GCTAGAAATAGC	SEQ ID NO 135
11	TCTGTTGCCCTCAACAT GGTTGG	GAAATTAATACGACTCA CTATAGTCTGTTGCCCTC AACATGGTGTGTTTTAGAG CTAGAAATAGC	SEQ ID NO 136

[1620]

12	TTAGTCTGTTGCCCTCAACATGG	GAAATTAATACGACTCACTATAGTTAGTCTGTTGCCCTCAACAGTTTATAGAGCTAGAAATAGC	SEQ ID NO 137
13	GTCTGGCAAATGGGAGGTGATGG	GAAATTAATACGACTCACTATAGGTCTGGCAAATGGGAGGTGAGTTTATAGAGCTAGAAATAGC	SEQ ID NO 138
14	CAGAGGTTCTGTCTGGCAAATGG	GAAATTAATACGACTCACTATAGCAGAGGTTCTGTCTGGCAAAGTTTATAGAGCTAGAAATAGC	SEQ ID NO 139
15	TTGTAGCCAGAGGTTCTGTCTGG	GAAATTAATACGACTCACTATAGTTGTAGCCAGAGGTTCTGTCTGTTTATAGAGCTAGAAATAGC	SEQ ID NO 140
16	ACTTCTGGATGAGCTCTCTCAGG	GAAATTAATACGACTCACTATAGACTTCTGGATGAGCTCTCTGTTTATAGAGCTAGAAATAGC	SEQ ID NO 141
17	AGAGCTCATCCAGAAGTAAATGG	GAAATTAATACGACTCACTATAGAGAGCTCATCCAGAAGTAAAGTTTATAGAGCTAGAAATAGC	SEQ ID NO 142
18	TTGGTGTCTCCATTTACTTCTGG	GAAATTAATACGACTCACTATAGTTGGTGTCTCCATTTACTTCGTTTATAGAGCTAGAAATAGC	SEQ ID NO 143
19	TTCTGGCTTCCCTTCATACAGGG	GAAATTAATACGACTCACTATAGTTCTGGCTTCCCTTCATACAGTTTATAGAGCTAGAAATAGC	SEQ ID NO 144

[1621]

Tet2	20	CAGGACTCACACGACTA TTCTGG	GAAATTAATACGACTCA CTATAGCAGGACTCACA CGACTATTCGTTTTAGA GCTAGAAATAGC	AAAAAAGCAC CGACTCGGTGC CACTTTTCAAG TTGATAACGGA CTAGCCTTATTT TAACTTGCTATT TCTAGCTCTAAA AC	SEQ ID NO 145
	21	CTACTTTCTTGTGTAAAG TCAGG	GAAATTAATACGACTCA CTATAGCTACTTTCTTGT GTAAAGTCGTTTTAGAG CTAGAAATAGC		SEQ ID NO 146
	22	GACTTTACACAAGAAAG TAGAGG	GAAATTAATACGACTCA CTATAGGACTTTACACA AGAAAGTAGGTTTTAGA GCTAGAAATAGC		SEQ ID NO 147
	23	GTCTTTCTCCATTAGCCT TTTGG	GAAATTAATACGACTCA CTATAGGTCTTCTCCAT TAGCCTTTGTTTTAGAGC TAGAAATAGC		SEQ ID NO 148
	24	AATGGAGAAAGACGTA ACTTCGG	GAAATTAATACGACTCA CTATAGAATGGAGAAAG ACGTAACCTGTTTTAGA GCTAGAAATAGC		SEQ ID NO 149
	25	ATGGAGAAAGACGTAA CTTCGGG	GAAATTAATACGACTCA CTATAGATGGAGAAAGA CGTAACCTCGTTTTAGA GCTAGAAATAGC		SEQ ID NO 150
	26	TGGAGAAAGACGTAA CTTCGGG	GAAATTAATACGACTCA CTATAGTGGAGAAAGAC GTAACCTCGTTTTAGA GCTAGAAATAGC		SEQ ID NO 151
	27	TTTGGTTGACTGCTTTCA CCTGG	GAAATTAATACGACTCA CTATAGTTTGGTTGACTG CTTTCACCGTTTTAGAGC TAGAAATAGC		SEQ ID NO 152

[1622]

28	TCAC TCAAATCGGAGAC ATTGG	GAAATTAATACGACTCA CTATAGTCACTCAAATC GGAGACATTGTTTTAGA GCTAGAAATAGC	SEQ ID NO 153
29	ATCTGAAGCTCTGGATT TTCAGG	GAAATTAATACGACTCA CTATAGATCTGAAGCTC TGGATTTTCGTTTTAGAG CTAGAAATAGC	SEQ ID NO 154
30	GCTTCAGATTCTGAATG AGCAGG	GAAATTAATACGACTCA CTATAGGCTTCAGATTCT GAATGAGCGTTTTAGAG CTAGAAATAGC	SEQ ID NO 155
31	CAGATTCTGAATGAGCA GGAGGG	GAAATTAATACGACTCA CTATAGCAGATTCTGAA TGAGCAGGAGTTTTAGA GCTAGAAATAGC	SEQ ID NO 156
32	AAGGCAGTGCTAATGCC TAATGG	GAAATTAATACGACTCA CTATAGAAGGCAGTGCT AATGCCTAAGTTTTAGA GCTAGAAATAGC	SEQ ID NO 157
33	GCAGAAACTGTAGCAC CATTAGG	GAAATTAATACGACTCA CTATAGGCAGAACTGT AGCACCATTGTTTTAGA GCTAGAAATAGC	SEQ ID NO 158
34	ACCGCAATGGAAACAC AATCTGG	GAAATTAATACGACTCA CTATAGACCGCAATGGA AACACAATCGTTTTAGA GCTAGAAATAGC	SEQ ID NO 159
35	TGTGGTTTTCTGCACCG CAATGG	GAAATTAATACGACTCA CTATAGTGTGGTTTTCTG CACCGCAAGTTTTAGAG CTAGAAATAGC	SEQ ID NO 160

[1623]

	36	CATAAATGCCATTAACA GTCAGG	GAAATTAATACGACTCA CTATAGCATAAATGCCA TTAACAGTCGTTTTAGA GCTAGAAATAGC		SEQ ID NO 161
	37	ATTAGTAGCCTGACTGT TAATGG	GAAATTAATACGACTCA CTATAGATTAGTAGCCT GACTGTAAAGTTTTAGA GCTAGAAATAGC		SEQ ID NO 162
	38	CGATGGGTGAGTGATCT CACAGG	GAAATTAATACGACTCA CTATAGCGATGGGTGAG TGATCTCACGTTTTAGA GCTAGAAATAGC		SEQ ID NO 163
	39	ACTCACCCATCGCATAC CTCAGG	GAAATTAATACGACTCA CTATAGACTCACCCATC GCATACCTCGTTTTAGA GCTAGAAATAGC		SEQ ID NO 164
	40	CTCACCCATCGCATACC TCAGGG	GAAATTAATACGACTCA CTATAGCTCACCCATCG CATACCTCAGTTTTAGA GCTAGAAATAGC		SEQ ID NO 165
基因	#	DNA 靶序列	正向引物序列	反向引物	序列号
	1	AGCAACAGGAGGAGTT GCAGAGG	GAAATTAATACGACTCA CTATAGAGCAACAGGA GGAGTTGCAGGTTTTAG AGCTAGAAATAGC		SEQ ID NO 166
	2	CCAGTAGGATCAGCAA CAGGAGG	GAAATTAATACGACTCA CTATAGCCAGTAGGATC AGCAACAGGGTTTTAGA GCTAGAAATAGC		SEQ ID NO 167
	3	CTCCTGTTGCTGATCCTA CTGGG	GAAATTAATACGACTCA CTATAGCTCCTGTTGCTG ATCCTACTGTTTTAGAGC TAGAAATAGC		SEQ ID NO 168

[1624]

4	GGCCCAGTAGGATCAG CAACAGG	GAAATTAATACGACTCA CTATAGGGCCCAGTAGG ATCAGCAACGTTTTAGA GCTAGAAATAGC	SEQ ID NO 169
5	TTGCTGATCCTACTGGG CCCTGG	GAAATTAATACGACTCA CTATAGTTGCTGATCCTA CTGGGCCCCGTTTTAGAG CTAGAAATAGC	SEQ ID NO 170
6	TGGCAACAGCTTGCAGC TGTGGG	GAAATTAATACGACTCA CTATAGTGGCAACAGCT TGCAGCTGTGTTTTAGA GCTAGAAATAGC	SEQ ID NO 171
7	CTTGGGTCCCCTGCTTG CCCGGG	GAAATTAATACGACTCA CTATAGCTTGGGTCCCC TGCTTGCCCCGTTTTAGA GCTAGAAATAGC	SEQ ID NO 172
8	GTCCCCTGCTTGCCCGG GACCGG	GAAATTAATACGACTCA CTATAGGTCCCCTGCTT GCCCCGGGACGTTTTAGA GCTAGAAATAGC	SEQ ID NO 173
9	CTCCGGTCCCGGGCAA GCAGGGG	GAAATTAATACGACTCA CTATAGCTCCGGTCCCG GGCAAGCAGGTTTTAGA GCTAGAAATAGC	SEQ ID NO 174
10	TCTCCGGTCCCGGGCAA GCAGGG	GAAATTAATACGACTCA CTATAGTCTCCGGTCCC GGGCAAGCAGTTTTAGA GCTAGAAATAGC	SEQ ID NO 175
11	GTCTCCGGTCCCGGGCA AGCAGG	GAAATTAATACGACTCA CTATAGGTCTCCGGTCC CGGGCAAGCGTTTTAGA GCTAGAAATAGC	SEQ ID NO 176

[1625]

12	GCTTGCCCGGGACCGG AGACAGG	GAAATTAATACGACTCA CTATAGGCTTGCCCGGG ACCGGAGACGTTTTAGA GCTAGAAATAGC	SEQ ID NO 177
13	GGTGGCCTGTCTCCGGT CCCGGG	GAAATTAATACGACTCA CTATAGGGTGGCCTGTC TCCGGTCCCGTTTTAGA GCTAGAAATAGC	SEQ ID NO 178
14	CGGTGGCCTGTCTCCGG TCCCGG	GAAATTAATACGACTCA CTATAGCGGTGGCCTGT CTCCGGTCCCGTTTTAGA GCTAGAAATAGC	SEQ ID NO 179
15	CATATTCGGTGGCCTGT CTCCGG	GAAATTAATACGACTCA CTATAGCATATTCGGTG GCCTGTCTCGTTTTAGA GCTAGAAATAGC	SEQ ID NO 180
16	ATCTAGGTAATCATATTC GGTGG	GAAATTAATACGACTCA CTATAGATCTAGGTAAT CATATTCGGGTTTTAGA GCTAGAAATAGC	SEQ ID NO 181
17	ATAATCTAGGTAATCA TATTCGG	GAAATTAATACGACTCA CTATAGATAATCTAGGT ACTCATATTGTTTTAGAG CTAGAAATAGC	SEQ ID NO 182
18	TTATGATTCCTGCCAG AAACGG	GAAATTAATACGACTCA CTATAGTTATGATTCCT GCCAGAAAGTTTTAGAG CTAGAAATAGC	SEQ ID NO 183
19	ATTTCTGGAGGCTCCGT TTCTGG	GAAATTAATACGACTCA CTATAGATTTCTGGAGG CTCCGTTTCGTTTTAGAG CTAGAAATAGC	SEQ ID NO 184

[1626]

20	ACTGACACCACTCCTCT GACTGG	GAAATTAATACGACTCA CTATAGACTGACACCAC TCCTCTGACGTTTTAGAG CTAGAAATAGC	SEQ ID NO 185
21	CTGACACCACTCCTCTG ACTGGG	GAAATTAATACGACTCA CTATAGCTGACACCACT CCTCTGACTGTTTTAGAG CTAGAAATAGC	SEQ ID NO 186
22	ACCACTCCTCTGACTGG GCCTGG	GAAATTAATACGACTCA CTATAGACCACTCCTCT GACTGGGCCGTTTTAGA GCTAGAAATAGC	SEQ ID NO 187
23	AACCCCTGAGTCTACCA CTGTGG	GAAATTAATACGACTCA CTATAGAACCCCTGAGT CTACCACTGGTTTTAGA GCTAGAAATAGC	SEQ ID NO 188
24	CTCCACAGTGGTAGACT CAGGGG	GAAATTAATACGACTCA CTATAGCTCCACAGTGG TAGACTCAGTTTTAGA GCTAGAAATAGC	SEQ ID NO 189
25	GCTCCACAGTGGTAGAC TCAGGG	GAAATTAATACGACTCA CTATAGGCTCCACAGTG GTAGACTCAGTTTTAGA GCTAGAAATAGC	SEQ ID NO 190
26	GGCTCCACAGTGGTAG ACTCAGG	GAAATTAATACGACTCA CTATAGGCTCCACAGT GGTAGACTCGTTTTAGA GCTAGAAATAGC	SEQ ID NO 191
27	CCTGCTGCAAGGCGTTC TACTGG	GAAATTAATACGACTCA CTATAGCCTGCTGCAAG GCGTTCACGTTTTAGA GCTAGAAATAGC	SEQ ID NO 192

[1627]

28	CCAGTAGAACGCCTTGC AGCAGG	GAAATTAATACGACTCA CTATAGCCAGTAGAACG CCTTGACGCGTTTTAGA GCTAGAAATAGC	SEQ ID NO 193
29	CGTTCTACTGGCCTGGA TGCAGG	GAAATTAATACGACTCA CTATAGCGTTCTACTGG CCTGGATGCGTTTTAGA GCTAGAAATAGC	SEQ ID NO 194
30	TCTACTGGCCTGGATGC AGGAGG	GAAATTAATACGACTCA CTATAGTCTACTGGCCT GGATGCAGGGTTTTAGA GCTAGAAATAGC	SEQ ID NO 195
31	CCACGGAGCTGGCCAA CATGGGG	GAAATTAATACGACTCA CTATAGCCACGGAGCTG GCCAACATGGTTTTAGA GCTAGAAATAGC	SEQ ID NO 196
32	CGTGGACAGGTTCCCCA TGTTGG	GAAATTAATACGACTCA CTATAGCGTGGACAGGT TCCCCATGTGTTTTAGAG CTAGAAATAGC	SEQ ID NO 197
33	GTCCACGGATTCAGCAG CTATGG	GAAATTAATACGACTCA CTATAGGTCCACGGATT CAGCAGCTAGTTTTAGA GCTAGAAATAGC	SEQ ID NO 198
34	GACCACTCAACCAGTGC CCACGG	GAAATTAATACGACTCA CTATAGGACCACTCAAC CAGTGCCAGTTTTAGA GCTAGAAATAGC	SEQ ID NO 199
35	GGAGTGGTCTGTGCCTC CGTGGG	GAAATTAATACGACTCA CTATAGGGAGTGGTCTG TGCCTCCGTGTTTTAGA GCTAGAAATAGC	SEQ ID NO 200

[1628]

PSGL-1	36	GGCACAGACAACTCGA CTGACGG	GAAATTAATACGACTCA CTATAGGGCACAGACAA CTCGACTGAGTTTTAGA GCTAGAAATAGC	AAAAAAGCAC CGACTCGGTGC CACTTTTCAAG TTGATAACGGA CTAGCCTATTT TAACTTGCTATT TCTAGCTCTAAA AC	SEQ ID NO 201
	37	GACAACTCGACTGACG GCCACGG	GAAATTAATACGACTCA CTATAGGACAACCTCGAC TGACGGCCAGTTTTAGA GCTAGAAATAGC		SEQ ID NO 202
	38	AACTCGACTGACGGCCA CGGAGG	GAAATTAATACGACTCA CTATAGAACTCGACTGA CGGCCACGGGTTTTAGA GCTAGAAATAGC		SEQ ID NO 203
	39	CACAGAACCCAGTGCC ACAGAGG	GAAATTAATACGACTCA CTATAGCACAGAACCCA GTGCCACAGTTTTAGA GCTAGAAATAGC		SEQ ID NO 204
	40	GGTAGTAGGTTCCATGG ACAGGG	GAAATTAATACGACTCA CTATAGGGTAGTAGGTT CCATGGACAGTTTTAGA GCTAGAAATAGC		SEQ ID NO 205
	41	TGGTAGTAGGTTCCATG GACAGG	GAAATTAATACGACTCA CTATAGTGGTAGTAGGT TCCATGGACGTTTTAGA GCTAGAAATAGC		SEQ ID NO 206
	42	TCTTTTGGTAGTAGGTT CATGG	GAAATTAATACGACTCA CTATAGTCTTTTGGTAGT AGGTTCCAGTTTTAGAG CTAGAAATAGC		SEQ ID NO 207
	43	ATGGAACCTACTACCAA AAGAGG	GAAATTAATACGACTCA CTATAGATGGAACCTAC TACCAAAAGGTTTTAGA GCTAGAAATAGC		SEQ ID NO 208

[1629]

44	AACAGACCTCTTTTGGT AGTAGG	GAAATTAATACGACTCA CTATAGAACAGACCTCT TTTGGTAGTGTTTAGAG CTAGAAATAGC	SEQ ID NO 209
45	GGGTATGAACAGACCTC TTTTGG	GAAATTAATACGACTCA CTATAGGGGTATGAACA GACCTCTTTGTTTAGAG CTAGAAATAGC	SEQ ID NO 210
46	TGTGTCCTCTGTTACTCA CAAGG	GAAATTAATACGACTCA CTATAGTGTGTCCTCTGT TACTCACAGTTTAGAG CTAGAAATAGC	SEQ ID NO 211
47	GTGTCCTCTGTTACTCAC AAGGG	GAAATTAATACGACTCA CTATAGGTGTCCTCTGTT ACTCACAAGTTTAGAG CTAGAAATAGC	SEQ ID NO 212
48	GTAGTTGACGGACAAAT TGCTGG	GAAATTAATACGACTCA CTATAGGTAGTTGACGG ACAAATTGCGTTTAGA GCTAGAAATAGC	SEQ ID NO 213
49	TTTGTCGGTCAACTACC CAGTGG	GAAATTAATACGACTCA CTATAGTTTGTCGGTCAA CTACCCAGGTTTAGAG CTAGAAATAGC	SEQ ID NO 214
50	TTGTCCGTCAACTACCC AGTGGG	GAAATTAATACGACTCA CTATAGTTGTCCGTCAA CTACCCAGTGTTTAGA GCTAGAAATAGC	SEQ ID NO 215
51	TGTCCGTCAACTACCCA GTGGGG	GAAATTAATACGACTCA CTATAGTGTCCGTCAAC TACCCAGTGTTTAGA GCTAGAAATAGC	SEQ ID NO 216

[1630]

52	GTCCGTCAACTACCCAG TGGGGG	GAAATTAATACGACTCA CTATAGGTCCGTCAACT ACCCAGTGGGTTTTAGA GCTAGAAATAGC	SEQ ID NO 217
53	CTCTGTGAAGCAGTGCC TGCTGG	GAAATTAATACGACTCA CTATAGCTCTGTGAAGC AGTGCCTGCGTTTTAGA GCTAGAAATAGC	SEQ ID NO 218
54	CCTGCTGGCCATCCTAA TCTTGG	GAAATTAATACGACTCA CTATAGCCTGCTGGCCA TCCTAATCTGTTTTAGAG CTAGAAATAGC	SEQ ID NO 219
55	CCAAGATTAGGATGGCC AGCAGG	GAAATTAATACGACTCA CTATAGCCAAGATTAGG ATGGCCAGCGTTTTAGA GCTAGAAATAGC	SEQ ID NO 220
56	GGCCATCCTAATCTTGG CGCTGG	GAAATTAATACGACTCA CTATAGGGCCATCCTAA TCTTGGCGCGTTTTAGA GCTAGAAATAGC	SEQ ID NO 221
57	CACCAGCGCCAAGATTA GGATGG	GAAATTAATACGACTCA CTATAGCACCAGCGCCA AGATTAGGAGTTTTAGA GCTAGAAATAGC	SEQ ID NO 222
58	AGTGACACGAAGAAG ATAGTGG	GAAATTAATACGACTCA CTATAGAGTGACACGA AGAAGATAGGTTTTAGA GCTAGAAATAGC	SEQ ID NO 223
59	TATCTTCTTCGTGTGCAC TGTGG	GAAATTAATACGACTCA CTATAGTATCTTCTTCGT GTGCACTGGTTTTAGAG CTAGAAATAGC	SEQ ID NO 224

[1631]

60	CTTCGTGTGCACTGTGG TGCTGG	GAAATTAATACGACTCA CTATAGCTTCGTGTGCA CTGTGGTGCGTTTTAGA GCTAGAAATAGC	SEQ ID NO 225
61	GGCGGTCCGCCTCTCCC GCAAGG	GAAATTAATACGACTCA CTATAGGGCGGTCCGCC TCTCCCGCAGTTTTAGA GCTAGAAATAGC	SEQ ID NO 226
62	GCGGTCCGCCTCTCCCG CAAGGG	GAAATTAATACGACTCA CTATAGGCGGTCCGCCT CTCCCGCAAGTTTTAGA GCTAGAAATAGC	SEQ ID NO 227
63	AATTACGCACGGGGTAC ATGTGG	GAAATTAATACGACTCA CTATAGAATTACGCACG GGGTACATGGTTTTAGA GCTAGAAATAGC	SEQ ID NO 228
64	TGGGGGAGTAATTACGC ACGGGG	GAAATTAATACGACTCA CTATAGTGGGGGAGTAA TTACGCACGGTTTTAGA GCTAGAAATAGC	SEQ ID NO 229
65	GTGGGGGAGTAATTAC GCACGGG	GAAATTAATACGACTCA CTATAGGTGGGGGAGTA ATTACGCACGTTTTAGA GCTAGAAATAGC	SEQ ID NO 230
66	GGTGGGGGAGTAATTA CGCACGG	GAAATTAATACGACTCA CTATAGGGTGGGGGAG TAATTACGCAGTTTTAG AGCTAGAAATAGC	SEQ ID NO 231
67	TAATTACTCCCCCACC GAGATGG	GAAATTAATACGACTCA CTATAGTAATTACTCCCC CACCAGAGGTTTTAGAG CTAGAAATAGC	SEQ ID NO 232

[1632]

68	AGATGCAGACCATCTCG GTGGGG	GAAATTAATACGACTCA CTATAGAGATGCAGACC ATCTCGGTGGTTTTAGA GCTAGAAATAGC	SEQ ID NO 233
69	GAGATGCAGACCATCTC GGTGGG	GAAATTAATACGACTCA CTATAGGAGATGCAGAC CATCTCGGTGTTTTAGA GCTAGAAATAGC	SEQ ID NO 234
70	TGAGATGCAGACCATCT CGGTGG	GAAATTAATACGACTCA CTATAGTGAGATGCAGA CCATCTCGGGTTTTAGA GCTAGAAATAGC	SEQ ID NO 235
71	GGATGAGATGCAGACC ATCTCGG	GAAATTAATACGACTCA CTATAGGGATGAGATGC AGACCATCTGTTTTAGA GCTAGAAATAGC	SEQ ID NO 236
72	ATCTCATCCCTGTTGCCT GATGG	GAAATTAATACGACTCA CTATAGATCTCATCCCTG TTGCCTGAGTTTTAGAG CTAGAAATAGC	SEQ ID NO 237
73	TCATCCCTGTTGCCTGAT GGGGG	GAAATTAATACGACTCA CTATAGTCATCCCTGTTG CCTGATGGGTTTTAGAG CTAGAAATAGC	SEQ ID NO 238
74	CTCACCCCATCAGGCA ACAGGG	GAAATTAATACGACTCA CTATAGCTCACCCCAT CAGGCAACAGTTTTAGA GCTAGAAATAGC	SEQ ID NO 239
75	GAGGGCCCCTCACCCC CATCAGG	GAAATTAATACGACTCA CTATAGGAGGGCCCCTC ACCCCATCGTTTTAGA GCTAGAAATAGC	SEQ ID NO 240

[1633]

76	GGGCCCTCTGCCACAGC CAATGG	GAAATTAATACGACTCA CTATAGGGGCCCTCTGC CACAGCCAAGTTTTAGA GCTAGAAATAGC	SEQ ID NO 241
77	CCCTCTGCCACAGCCAA TGGGGG	GAAATTAATACGACTCA CTATAGCCCTCTGCCAC AGCCAATGGGTTTTAGA GCTAGAAATAGC	SEQ ID NO 242
78	CCCCCATTGGCTGTGGC AGAGGG	GAAATTAATACGACTCA CTATAGCCCCCATTGGC TGTGGCAGAGTTTTAGA GCTAGAAATAGC	SEQ ID NO 243
79	GCCCCCATTGGCTGTGG CAGAGG	GAAATTAATACGACTCA CTATAGGCCCCCATTTGG CTGTGGCAGGTTTTAGA GCTAGAAATAGC	SEQ ID NO 244
80	GGACAGGCCCCCATTG GCTGTGG	GAAATTAATACGACTCA CTATAGGGACAGGCCCC CATTGGCTGGTTTTAGA GCTAGAAATAGC	SEQ ID NO 245
81	CCGGGCTCTTGGCCTTG GACAGG	GAAATTAATACGACTCA CTATAGCCGGGCTCTTG GCCTTGGACGTTTTAGA GCTAGAAATAGC	SEQ ID NO 246
82	CTGTCCAAGGCCAAGA GCCCCGG	GAAATTAATACGACTCA CTATAGCTGTCCAAGGC CAAGAGCCCGTTTTAGA GCTAGAAATAGC	SEQ ID NO 247
83	TGGCGTCAGGCCCGGG CTCTTGG	GAAATTAATACGACTCA CTATAGTGGCGTCAGGC CCGGGCTCTGTTTTAGA GCTAGAAATAGC	SEQ ID NO 248

[1634]

	84	CGGGCCTGACGCCAGA GCCCAGG	GAAATTAATACGACTCA CTATAGCGGGCCTGACG CCAGAGCCCCGTTT TAGA GCTAGAAATAGC		SEQ ID NO 249
基因	#	DNA 靶序列	正向引物序列	反向引物	序列号
	1	CAACAACCATGCTGGGC ATCTGG	GAAATTAATACGACTCA CTATAGCAACAACCATG CTGGGCATCGTTT TAGA GCTAGAAATAGC		SEQ ID NO 250
	2	GAGGGTCCAGATGCCC AGCATGG	GAAATTAATACGACTCA CTATAGGAGGGTCCAGA TGCCCAGCAGTTT TAGA GCTAGAAATAGC		SEQ ID NO 251
	3	CATCTGGACCCTCCTA CCTCTGG	GAAATTAATACGACTCA CTATAGCATCTGGACCC TCCTACCTCGTTT TAGAG CTAGAAATAGC		SEQ ID NO 252
	4	AGGGCTCACCAGAGGT AGGAGGG	GAAATTAATACGACTCA CTATAGAGGGCTCACCA GAGGTAGGAGTTT TAGA GCTAGAAATAGC		SEQ ID NO 253
	5	GGAGTTGATGTCAGTC ACTTGGG	GAAATTAATACGACTCA CTATAGGGAGTTGATGT CAGTCACTTGTTT TAGA GCTAGAAATAGC		SEQ ID NO 254
	6	TGGAGTTGATGTCAGTC ACTTGG	GAAATTAATACGACTCA CTATAGTGGAGTTGATG TCAGTCACTGTTT TAGA GCTAGAAATAGC		SEQ ID NO 255
	7	AGTGA CTGACATCAACT CCAAGG	GAAATTAATACGACTCA CTATAGAGTGACTGACA TCAACTCCAGTTT TAGA GCTAGAAATAGC		SEQ ID NO 256

[1635]

FAS	8	GTGACTGACATCAACT CCAAGGG	GAAATTAATACGACTCA CTATAGGTGACTGACAT CAACTCCAAGTTTTAGA GCTAGAAATAGC	AAAAAAGCAC CGACTCGGTGC CACTTTTCAAG TTGATAACGGA CTAGCCTTATTT TAACTTGCTATT TCTAGCTCTAAA AC	SEQ ID NO 257
	9	ACTCCAAGGGATTGGAA TTGAGG	GAAATTAATACGACTCA CTATAGACTCCAAGGGA TTGGAATTGGTTTTAGA GCTAGAAATAGC		SEQ ID NO 258
	10	CTTCCTCAATTCCAATCC CTTGG	GAAATTAATACGACTCA CTATAGCTTCCTCAATTC CAATCCCTGTTTTAGAG CTAGAAATAGC		SEQ ID NO 259
	11	TACAGTTGAGACTCAGA ACTTGG	GAAATTAATACGACTCA CTATAGTACAGTTGAGA CTCAGAACTGTTTTAGA GCTAGAAATAGC		SEQ ID NO 260
	12	TTGGAAGGCCTGCATCA TGATGG	GAAATTAATACGACTCA CTATAGTTGGAAGGCCT GCATCATGAGTTTTAGA GCTAGAAATAGC		SEQ ID NO 261
	13	AGAATTGGCCATCATGA TGCAGG	GAAATTAATACGACTCA CTATAGAGAATTGGCCA TCATGATGCGTTTTAGA GCTAGAAATAGC		SEQ ID NO 262
	14	GACAGGGCTTATGGCA GAATTGG	GAAATTAATACGACTCA CTATAGGACAGGGCTTA TGGCAGAATGTTTTAGA GCTAGAAATAGC		SEQ ID NO 263
	15	TGTAACATACCTGGAG GACAGGG	GAAATTAATACGACTCA CTATAGTGTAACATACC TGGAGGACAGTTTTAGA GCTAGAAATAGC		SEQ ID NO 264

[1636]

	16	GTGTAACATACCTGGAG GACAGG	GAAATTAATACGACTCA CTATAGGTGTAACATAC CTGGAGGACGTTTTAGA GCTAGAAATAGC		SEQ ID NO 265
基因	#	DNA 靶序列	正向引物序列	反向引物	序列号
	1	CGTACCTGTGCAACTCC TGTTGG	GAAATTAATACGACTCA CTATAGCGTACCTGTGC AACTCCTGTGTTTTAGA GCTAGAAATAGC		SEQ ID NO 266
	2	GATCTACTGGAATTCCT AATGGG	GAAATTAATACGACTCA CTATAGGATCTACTGGA ATTCCTAATGTTTTAGAG CTAGAAATAGC		SEQ ID NO 267
	3	GAGTCAGCTGTTGGCCC ATTAGG	GAAATTAATACGACTCA CTATAGGAGTCAGCTGT TGGCCATTGTTTTAGA GCTAGAAATAGC		SEQ ID NO 268
	4	CTGCCTACAACTCAGT CTCTGG	GAAATTAATACGACTCA CTATAGCTGCCTACAAA CTCAGTCTCGTTTTAGAG CTAGAAATAGC		SEQ ID NO 269
	5	GGGCAGGCAGGACGGA CTCCAGG	GAAATTAATACGACTCA CTATAGGGGCAGGCAG GACGGACTCCGTTTTAG AGCTAGAAATAGC		SEQ ID NO 270
	6	GGAGTCCGTCCTGCCTG CCCTGG	GAAATTAATACGACTCA CTATAGGGAGTCCGTCC TGCCTGCCCGTTTTAGA GCTAGAAATAGC		SEQ ID NO 271
	7	GAGTCCGTCCTGCCTGC CCTGGG	GAAATTAATACGACTCA CTATAGGAGTCCGTCC GCCTGCCCTGTTTTAGA GCTAGAAATAGC		SEQ ID NO 272

[1637]

KDM6A	8	GAAAAGGGTCCATTGG CCAAAGG	GAAATTAATACGACTCA CTATAGGAAAAGGGTCC ATTGGCCAAGTTTTAGA GCTAGAAATAGC	AAAAAAGCAC CGACTCGGTGC CACTTTTCAAG TTGATAACGGA CTAGCCTTATTT TAACTTGCTATT TCTAGCTCTAAA AC	SEQ ID NO 273
	9	GCCTGCAGAAAAGGGT CCATTGG	GAAATTAATACGACTCA CTATAGGCTGCAGAAA AGGGTCCATGTTTTAGA GCTAGAAATAGC		SEQ ID NO 274
	10	TTGATGTGCTACAGGGA ACATGG	GAAATTAATACGACTCA CTATAGTTGATGTGCTA CAGGGAACAGTTTTAGA GCTAGAAATAGC		SEQ ID NO 275
	11	AGCGTTCCTTGATGTGCT ACAGGG	GAAATTAATACGACTCA CTATAGAGCGTTCCTTGA TGTGCTACAGTTTTAGA GCTAGAAATAGC		SEQ ID NO 276
	12	CAGCGTTCCTTGATGTGC TACAGG	GAAATTAATACGACTCA CTATAGCAGCGTTCCTTG ATGTGCTACGTTTTAGA GCTAGAAATAGC		SEQ ID NO 277
	13	CTGTAGCACATCAAGAA CGCTGG	GAAATTAATACGACTCA CTATAGCTGTAGCACAT CAAGAACGCGTTTTAGA GCTAGAAATAGC		SEQ ID NO 278
	14	TGTAGCACATCAAGAAC GCTGGG	GAAATTAATACGACTCA CTATAGTGTAGCACATC AAGAACGCTGTTTTAGA GCTAGAAATAGC		SEQ ID NO 279
	15	ATAGGCAATAATCATAT AACAGG	GAAATTAATACGACTCA CTATAGATAGGCAATAA TCATATAACGTTTTAGA GCTAGAAATAGC		SEQ ID NO 280

[1638]

16	AGTGCGTTTCGCTGCAG GTAAGG	GAAATTAATACGACTCA CTATAGAGTGCGTTTCG CTGCAGGTAGTTTTAGA GCTAGAAATAGC	SEQ ID NO 281
17	GAGTGAGTGCGTTTCGC TGCAGG	GAAATTAATACGACTCA CTATAGGAGTGAGTGCG TTTCGCTGCGTTTTAGAG CTAGAAATAGC	SEQ ID NO 282
18	GTCAGGTTTGTGCGGTT ATGAGG	GAAATTAATACGACTCA CTATAGGTCAGGTTTGT GCGGTTATGGTTTTAGA GCTAGAAATAGC	SEQ ID NO 283
19	CGCTGCTGGTCAGGTTT GTGCGG	GAAATTAATACGACTCA CTATAGCGCTGCTGGTC AGGTTTGTGGTTTTAGA GCTAGAAATAGC	SEQ ID NO 284
20	AAACCTGACCAGCAGC GCAGAGG	GAAATTAATACGACTCA CTATAGAAACCTGACCA GCAGCGCAGGTTTTAGA GCTAGAAATAGC	SEQ ID NO 285
21	CCAGCAGCGCAGAGGA GCCGTGG	GAAATTAATACGACTCA CTATAGCCAGCAGCGCA GAGGAGCCGGTTTTAGA GCTAGAAATAGC	SEQ ID NO 286
22	CCACGGCTCCTCTGCGC TGCTGG	GAAATTAATACGACTCA CTATAGCCACGGCTCCT CTGCGCTGCGTTTTAGA GCTAGAAATAGC	SEQ ID NO 287
23	CCAACTATCTAACTCCA CTCAGG	GAAATTAATACGACTCA CTATAGCCAACTATCTA ACTCCACTCGTTTTAGA GCTAGAAATAGC	SEQ ID NO 288
24	CCTGAGTGGAGTTAGAT AGTTGG	GAAATTAATACGACTCA CTATAGCCTGAGTGGAG TTAGATAGTGTTTAGA GCTAGAAATAGC	SEQ ID NO 289

[1639] 2.3深度测序

[1640] 使用Hipi Plus DNA聚合酶(Elpis-bio)将中靶(on-target)和脱靶(off-target)位点PCR扩增为200-300bp大小。利用Mi-seq仪器(Illumina)对通过上述方法获得的PCR产物进行测序,并用CRISPR RGEN工具(www.rgenome.net)的Cas Analyzer实施分析。认为距

CRISPR/Cas9切割位点5bp的插入/删除为由RGEN诱导的突变。

[1641] 如表4和表6所示,作为深度测序的结果,证实了当递送CRISPR-Cas9时在多种免疫细胞中高效产生插入缺失突变。

[1642] 实施例3:sgRNA的制备

[1643] 3.1在Jurkat细胞中筛选sgRNA

[1644] 在Jurkat细胞中测试由实施例2中所述的方法获得的靶向A20、DGK α 、EGR2、PPP2R2D、PD-1、CTLA-4、DGK ζ 、PSGL-1、KDM6A、FAS和TET2的外显子的sgRNA的活性。

[1645] 通过对用实施例1的方法转染Cas9的Jurkat细胞和未转染的Jurkat细胞中插入缺失的比例进行比较,对实施例2中获得的各sgRNA进行测试。表3示出了在人基因组和CRISPR/Cas9靶序列中具有相似靶序列的错配位点的数目,表4示出了各sgRNA的插入缺失比例。在靶向各基因的gRNA中,将具有优良活性的DNA靶区域以粗体标示。

[1646] [表3]

[1647]

基因	#	DNA 靶序列	错配		
			0 bp	1 bp	2 bp
A20	1	CTTGTGGCGCTGAAAACGAACGG	1	0	0
	2	ATGCCACTTCTCAGTACATGTGG	1	0	0
	3	GCCACTTCTCAGTACATGTGGGG	1	0	0
	4	GCCCCACATGTACTGAGAAGTGG	1	0	0
	5	TCAGTACATGTGGGGCGTTCAGG	1	0	0
	6	GGGCGTTCAGGACACAGACTTGG	1	0	0
	7	CACAGACTTGGTACTGAGGAAGG	1	0	0
	8	GGCGCTGTTCAAGCACGCTCAAGG	1	0	0
	9	CACGCAACTTTAAATTCCGCTGG	1	0	0
	10	CGGGGCTTTGCTATGATACTCGG	1	0	0
	11	GGCTTCCACAGACACACCCATGG	1	0	0
	12	TGAAGTCCACTTCGGGCCATGGG	1	0	0
基因	#	DNA 靶序列	错配		
			0 bp	1 bp	2 bp
DGK α	1	CTGTACGACACGGACAGAAATGG	1	0	0
	2	TGTACGACACGGACAGAAATGGG	1	0	0
	3	CACGGACAGAAATGGGATCCTGG	1	0	0
	4	GATGCGAGTGGCTGAATACCTGG	1	0	0
	5	GAGTGGCTGAATACCTGGATTGG	1	0	0
	6	AGTGGCTGAATACCTGGATTGGG	1	0	0
	7	ATTGGGATGTGTCTGAGCTGAGG	1	0	0
	8	ATGAAAGAGATTGACTATGATGG	1	0	0
	9	CTCTGTCTCTCAAGCTGAGTGGG	1	0	0
	10	TCTCTCAAGCTGAGTGGGTCCGG	1	0	0
	11	CTCTCAAGCTGAGTGGGTCCGGG	1	0	0
	12	CAAGCTGAGTGGGTCCGGGCTGG	1	0	0
基因	#	DNA 靶序列	错配		
			0 bp	1 bp	2 bp
	1	TTGACATGACTGGAGAGAAGAGG	1	0	0

[1648]

EGR2

2	GACTGGAGAGAAGAGGTCGTTGG	1	0	0
3	GAGACGGGAGCAAAGCTGCTGGG	1	0	0
4	AGAGACGGGAGCAAAGCTGCTGG	1	0	0
5	TGGTTTCTAGGTGCAGAGACGGG	1	0	0
6	TAAGTGAAGGTCTGGTTTCTAGG	1	0	0
7	TGCCCATGTAAGTGAAGGTCTGG	1	0	0
8	GAAGTTGCCCATGTAAGTGAAGG	1	0	0
9	TCCATTGACCCTCAGTACCCTGG	1	0	0
10	TATGCCTTCTGGGTAGCAGCTGG	1	0	0
11	TGAGTGCAGGCATCTTGCAAGGG	1	0	0
12	GAGTGCAGGCATCTTGCAAGGGG	1	0	0
13	GATGAGGCTGTGGTTGAAGCTGG	1	0	0
14	CCACTGGCCACAGGACCCCTGGG	1	0	0
15	GGGACATGGTGCACACACCCAGG	1	0	0
16	GAGTACAGGTGGTCCAGGTCAGG	1	0	0
17	GCGGAGAGTACAGGTGGTCCAGG	1	0	0
18	GCGGTGGCGGAGAGTACAGGTGG	1	0	0
19	TCTCCTGCACAGCCAGAATAAGG	1	0	0
20	ACGCAGAAGGGTCCTGGTAGAGG	1	0	0
21	AGGTGGTGGGTAGGCCAGAGAGG	1	0	0
22	CCCAAGCCAGCCACGGACCCAGG	1	0	0
23	ACCTGGGTCCGTGGCTGGCTTGG	1	0	0
24	AAGAGACCTGGGTCCGTGGCTGG	1	0	0
25	GGATCATTGGGAAGAGACCTGGG	1	0	0
26	GGGATCATTGGGAAGAGACCTGG	1	0	0
27	CAGGATAGTCTGGGATCATTGGG	1	0	0
28	GGAAAGAATCCAGGATAGTCTGG	1	0	0
29	CAGTGCCAGAGAGACCTACATGG	1	0	0
30	CTGTACCATGTAGGTCTCTCTGG	1	0	0
31	AGAGACCTACATGGTACAGCTGG	1	0	0
32	CTGGGCCAGCTGTACCATGTAGG	1	0	0

[1649]

	33	AGGGAAAGGGCTTACGGTCTGGG	1	0	0
	34	CAGGGAAAGGGCTTACGGTCTGG	1	0	0
基因	#	DNA 靶序列	错配		
			0 bp	1 bp	2 bp
PPP2R2D	5	TCTGGAGATCTTCTTGCAACAGG	1	0	0
	6	CTCCGGTTCATGACTTTGAAAGG	1	0	0
	7	GTCTCCATCTTCGTCTTTCAGG	1	0	0
	8	GAAGACTTCGAGACCCATTTAGG	1	0	0
	9	TCGAGACCCATTTAGGATCACGG	1	0	0
	10	GTAGCGCCGTGATCCTAAATGGG	1	0	0
	11	CGTAGCGCCGTGATCCTAAATGG	1	0	0
	12	CATTTAGGATCACGGCGCTACGG	1	0	0
	13	GGTCCCAATATTGAAGCCCATGG	1	0	0
	14	GATCCATGGGCTTCAATATTGGG	1	0	0
	15	AGATCCATGGGCTTCAATATTGG	1	0	0
	16	GCTTCTACCATAAGATCCATGGG	1	0	0
	17	CGCTTCTACCATAAGATCCATGG	1	0	0
	18	GCATTTGCAAAAATTCGCCGTGG	1	0	0
	19	ATGACCTGAGAATTAATTTATGG	1	0	0
	20	CCATGCACTCCCAGACATCGTGG	1	0	0
	21	GCACTGGTGCGGGTGGAACCTCGG	1	0	0
	22	ACACGTTGCACTGGTGCGGGTGG	1	0	0
	23	CGAACACGTTGCACTGGTGCGGG	1	0	0
	24	ACGAACACGTTGCACTGGTGCGG	1	0	0
	25	TGTAGACGAACACGTTGCACTGG	1	0	0
	26	GCGCATGTGACACAGGCGGATGG	1	0	0
	27	AGGAGCGCATGTGACACAGGCGG	1	0	0
	28	CCGAGGAGCGCATGTGACACAGG	1	0	0
	29	CCTGTGTGACATGCGCTCCTCGG	1	0	0
基因	#	DNA 靶序列	错配		
			0 bp	1 bp	2 bp

[1650]

PD-1	1	CGACTGGCCAGGGCGCCTGTGGG	1	0	0
	2	ACCGCCCAGACGACTGGCCAGGG	1	0	0
	3	CACCGCCCAGACGACTGGCCAGG	1	0	0
	4	GTCTGGGCGGTGCTACAACCTGGG	1	0	0
	5	CTACAACCTGGGCTGGCGGCCAGG	1	0	0
	6	CACCTACCTAAGAACCATCCTGG	1	0	0
	7	CGGTCACCACGAGCAGGGCTGGG	1	0	0
	8	GCCCTGCTCGTGGTGACCGAAGG	1	0	0
	9	CGGAGAGCTTCGTGCTAAACTGG	1	0	0
	10	CAGCTTGTCCTGTCTGGTTGCTGG	1	0	0
	11	AGGCGGCCAGCTTGTCCTGTCTGG	1	0	0
	12	CCGGGCTGGCTGCGGTCTCTCGGG	1	0	0
	13	CGTTGGGCAGTTGTGTGACACGG	1	0	0
基因	#	DNA 靶序列	错配		
			0 bp	1 bp	2 bp
CTLA-4	1	CATAAAGCCATGGCTTGCCTTGG	1	0	0
	2	CCTTGGATTTAGCGGCACAAGG	1	0	0
	3	CCTTGTGCCGCTGAAATCCAAGG	1	0	0
	4	CACTCACCTTTGCAGAAGACAGG	1	0	0
	5	TTCCATGCTAGCAATGCACGTGG	1	0	0
	6	GGCCACGTGCATTGCTAGCATGG	1	0	0
	7	GGCCCAGCCTGCTGTGGTACTGG	1	0	0
	8	AGGTCCGGGTGACAGTGCTTCGG	1	0	0
	9	CCGGGTGACAGTGCTTCGGCAGG	1	0	0
	10	CTGTGCGGCAACCTACATGATGG	1	0	0
	11	CAACTCATTCCCCATCATGTAGG	1	0	0
	12	CTAGATGATTCCATCTGCACGGG	1	0	0
基因	#	DNA 靶序列	错配		
			0 bp	1 bp	2 bp
	1	GGCTAGGAGTCAGCGACATATGG	1	0	0
	2	GCTAGGAGTCAGCGACATATGGG	1	0	0
	3	CTAGGAGTCAGCGACATATGGGG	1	0	0

[1651]

DGKζ	4	GTACTGTGTAGCCAGGATGCTGG	1	0	0
	5	ACGAGCACTCACCAGCATCCTGG	1	0	0
	6	AGGCTCCAGGAATGTCCGCGAGG	1	0	0
	7	ACTTACCTCGCGGACATTCCTGG	1	0	0
	8	CACCCTGGGCACTTACCTCGCGG	1	0	0
	9	GTGCCGTACAAAGGTTGGCTGGG	1	0	0
	10	GGTGCCGTACAAAGGTTGGCTGG	1	0	0
	11	CTCTCCTCAGTACCACAGCAAGG	1	0	0
	12	CCTGGGGCCTCCGGGCGCGGAGG	1	0	0
	13	AGTACTCACCTGGGGCCTCCGGG	1	0	0
	14	AGGGTCTCCAGCGGCCCTCCTGG	1	0	0
	15	GCAAGTACTTACGCCTCCTTGGG	1	0	0
	16	TTGCGGTACATCTCCAGCCTGGG	1	0	0
	17	TTTGCGGTACATCTCCAGCCTGG	1	0	0
基因	#	DNA 靶序列	错配		
			0 bp	1 bp	2 bp
	1	GCAAAACCTGTCCACTCTTATGG	1	0	0
	2	TTGGTGCCATAAGAGTGGACAGG	1	0	0
	3	GGTGCAAGTTTCTTATATGTTGG	1	0	0
	4	ACCTGATGCATATAATAATCAGG	1	0	0
	5	ACCTGATTATTATATGCATCAGG	1	0	0
	6	CAGAGCACCAGAGTGCCGTCTGG	1	0	0
	7	AGAGCACCAGAGTGCCGTCTGGG	1	0	0
	8	AGAGTGCCGTCTGGGTCTGAAGG	1	0	0
	9	AGGAAGGCCGTCCATTCTCAGGG	1	0	0
	10	GGATAGAACCAACCATGTTGAGG	1	0	0
	11	TCTGTTGCCCTCAACATGGTTGG	1	0	0
	12	TTAGTCTGTTGCCCTCAACATGG	1	0	0
	13	GTCTGGCAAATGGGAGGTGATGG	1	0	0
	14	CAGAGGTTCTGTCTGGCAAATGG	1	0	0
	15	TTGTAGCCAGAGGTTCTGTCTGG	1	0	0

[1652]

Tet2	16	ACTTCTGGATGAGCTCTCTCAGG	1	0	0
	17	AGAGCTCATCCAGAAGTAAATGG	1	0	0
	18	TTGGTGTCTCCATTTACTTCTGG	1	0	0
	19	TTCTGGCTTCCCTTCATACAGGG	1	0	0
	20	CAGGACTCACACGACTATTCTGG	1	0	0
	21	CTACTTTCTTGTGTAAAGTCAGG	1	0	0
	22	GACTTTACACAAGAAAGTAGAGG	1	0	0
	23	GTCTTTCTCCATTAGCCTTTTGG	1	0	0
	24	AATGGAGAAAGACGTAACCTCGG	1	0	0
	25	ATGGAGAAAGACGTAACCTCGGG	1	0	0
	26	TGGAGAAAGACGTAACCTCGGGG	1	0	0
	27	TTTGGTTGACTGCTTTCACCTGG	1	0	0
	28	TCACTCAAATCGGAGACATTTGG	1	0	0
	29	ATCTGAAGCTCTGGATTTTCAGG	1	0	0
	30	GCTTCAGATTCTGAATGAGCAGG	1	0	0
	31	CAGATTCTGAATGAGCAGGAGGG	1	0	0
	32	AAGGCAGTGCTAATGCCTAATGG	1	0	0
	33	GCAGAACTGTAGCACCATTAGG	1	0	0
	34	ACCGCAATGGAAACACAATCTGG	1	0	0
	35	TGTGGTTTTCTGCACCGCAATGG	1	0	0
	36	CATAAATGCCATTAACAGTCAGG	1	0	0
	37	ATTAGTAGCCTGACTGTTAATGG	1	0	0
	38	CGATGGGTGAGTGATCTCACAGG	1	0	0
	39	ACTCACCCATCGCATACCTCAGG	1	0	0
	40	CTCACCCATCGCATACCTCAGGG	1	0	0
基因	#	DNA 靶序列	错配		
			0 bp	1 bp	2 bp
	1	AGCAACAGGAGGAGTTGCAGAGG	1	0	0
	2	CCAGTAGGATCAGCAACAGGAGG	1	0	0
	3	CTCCTGTTGCTGATCCTACTGGG	1	0	0
	4	GGCCCAGTAGGATCAGCAACAGG	1	0	0

[1653]

5	TTGCTGATCCTACTGGGCCCTGG	1	0	0
6	TGGCAACAGCTTGCAGCTGTGGG	1	0	0
7	CTTGGGTCCCCTGCTTGCCCGGG	1	0	0
8	GTCCCCTGCTTGCCCGGGACCGG	1	0	0
9	CTCCGGTCCCGGGCAAGCAGGGG	1	0	0
10	TCTCCGGTCCCGGGCAAGCAGGG	1	0	0
11	GTCTCCGGTCCCGGGCAAGCAGG	1	0	0
12	GCTTGCCCGGGACCGGAGACAGG	1	0	0
13	GGTGGCCTGTCTCCGGTCCCGGG	1	0	0
14	CGGTGGCCTGTCTCCGGTCCCGG	1	0	0
15	CATATTCGGTGGCCTGTCTCCGG	1	0	0
16	ATCTAGGTACTCATATTCGGTGG	1	0	0
17	ATAATCTAGGTACTCATATTCGG	1	0	0
18	TTATGATTTCCTGCCAGAAACGG	1	0	0
19	ATTCTGGAGGCTCCGTTTCTGG	1	0	0
20	ACTGACACCACTCCTCTGACTGG	1	0	0
21	CTGACACCACTCCTCTGACTGGG	1	0	0
22	ACCACTCCTCTGACTGGGCCTGG	1	0	0
23	AACCCCTGAGTCTACCACTGTGG	1	0	0
24	CTCCACAGTGGTAGACTCAGGGG	1	0	0
25	GCTCCACAGTGGTAGACTCAGGG	1	0	0
26	GGCTCCACAGTGGTAGACTCAGG	1	0	0
27	CCTGCTGCAAGGCGTTCTACTGG	1	0	0
28	CCAGTAGAACGCCTTGCAGCAGG	1	0	0
29	CGTTCTACTGGCCTGGATGCAGG	1	0	0
30	TCTACTGGCCTGGATGCAGGAGG	1	0	0
31	CCACGGAGCTGGCCAACATGGGG	1	0	0
32	CGTGGACAGGTTCCCCATGTTGG	1	0	0
33	GTCCACGGATTCAGCAGCTATGG	1	0	0
34	GACCACTCAACCAGTGCCACCGG	1	0	0
35	GGAGTGGTCTGTGCCTCCGTGGG	1	0	0

[1654]

PSGL-1

36	GGCACAGACAACTCGACTGACGG	1	0	0
37	GACAACTCGACTGACGGCCACGG	1	0	0
38	AACTCGACTGACGGCCACGGAGG	1	0	0
39	CACAGAACCCAGTGCCACAGAGG	1	0	0
40	GGTAGTAGGTTCCATGGACAGGG	1	0	0
41	TGGTAGTAGGTTCCATGGACAGG	1	0	0
42	TCTTTTGGTAGTAGGTTCCATGG	1	0	0
43	ATGGAACCTACTACCAAAAGAGG	1	0	0
44	AACAGACCTCTTTTGGTAGTAGG	1	0	0
45	GGGTATGAACAGACCTCTTTTGG	1	0	0
46	TGTGTCCTCTGTTACTCACAAGG	1	0	0
47	GTGTCCTCTGTTACTCACAAGGG	1	0	0
48	GTAGTTGACGGACAAATTGCTGG	1	0	0
49	TTGTCCGTCAACTACCCAGTGG	1	0	0
50	TTGTCCGTCAACTACCCAGTGGG	1	0	0
51	TGTCCGTCAACTACCCAGTGGGG	1	0	0
52	GTCCGTCAACTACCCAGTGGGGG	1	0	0
53	CTCTGTGAAGCAGTGCCTGCTGG	1	0	0
54	CCTGCTGGCCATCCTAATCTTGG	1	0	0
55	CCAAGATTAGGATGGCCAGCAGG	1	0	0
56	GGCCATCCTAATCTTGCGCTGG	1	0	0
57	CACCAGCGCCAAGATTAGGATGG	1	0	0
58	AGTGCACACGAAGAAGATAGTGG	1	0	0
59	TATCTTCTTCGTGTGCACTGTGG	1	0	0
60	CTTCGTGTGCACTGTGGTGCTGG	1	0	0
61	GGCGGTCCGCCTCTCCCGCAAGG	1	0	0
62	GCGGTCCGCCTCTCCCGCAAGGG	1	0	0
63	AATTACGCACGGGGTACATGTGG	1	0	0
64	TGGGGGAGTAATTACGCACGGGG	1	0	0
65	GTGGGGGAGTAATTACGCACGGG	1	0	0
66	GGTGGGGGAGTAATTACGCACGG	1	0	0

[1655]

	67	TAATTACTCCCCACCGAGATGG	1	0	0
	68	AGATGCAGACCATCTCGGTGGGG	1	0	0
	69	GAGATGCAGACCATCTCGGTGGG	1	0	0
	70	TGAGATGCAGACCATCTCGGTGG	1	0	0
	71	GGATGAGATGCAGACCATCTCGG	1	0	0
	72	ATCTCATCCCTGTTGCCTGATGG	1	0	0
	73	TCATCCCTGTTGCCTGATGGGGG	1	0	0
	74	CTCACCCCCATCAGGCAACAGGG	1	0	0
	75	GAGGGCCCCTACCCCCATCAGG	1	0	0
	76	GGGCCCTCTGCCACAGCCAATGG	1	0	0
	77	CCCTCTGCCACAGCCAATGGGGG	1	0	0
	78	CCCCATTGGCTGTGGCAGAGGG	1	0	0
	79	GCCCCATTGGCTGTGGCAGAGG	1	0	0
	80	GGACAGGCCCCATTGGCTGTGG	1	0	0
	81	CCGGGCTCTTGGCCTTGGACAGG	1	0	0
	82	CTGTCCAAGGCCAAGAGCCCGGG	1	0	0
	83	TGGCGTCAGGCCCGGGCTCTTGG	1	0	0
	84	CGGGCCTGACGCCAGAGCCCAGG	1	0	0
基因	#	DNA 靶序列	错配		
			0 bp	1 bp	2 bp
FAS	1	CAACAACCATGCTGGGCATCTGG	1	0	0
	2	GAGGGTCCAGATGCCCAGCATGG	1	0	0
	3	CATCTGGACCCTCCTACCTCTGG	1	0	0
	4	AGGGCTCACCAGAGGTAGGAGGG	1	0	0
	5	GGAGTTGATGTCAGTCACTTGGG	1	0	0
	6	TGGAGTTGATGTCAGTCACTTGG	1	0	0
	7	AGTGACTGACATCAACTCCAAGG	1	0	0
	8	GTGACTGACATCAACTCCAAGGG	1	0	0
	9	ACTCCAAGGGATTGGAATTGAGG	1	0	0
	10	CTTCCTCAATTCCAATCCCTTGG	1	0	0
	11	TACAGTTGAGACTCAGAACTTGG	1	0	0

[1656]

	12	TTGGAAGGCCTGCATCATGATGG	1	0	0
	13	AGAATTGGCCATCATGATGCAGG	1	0	0
	14	GACAGGGCTTATGGCAGAATTGG	1	0	0
	15	TGTAACATACCTGGAGGACAGGG	1	0	0
	16	GTGTAACATACCTGGAGGACAGG	1	0	0
基因	#	DNA 靶序列	错配		
			0 bp	1 bp	2 bp
KDM6A	1	CGTACCTGTGCAACTCCTGTTGG	1	0	0
	2	GATCTACTGGAATTCCTAATGGG	1	0	0
	3	GAGTCAGCTGTTGGCCATTAGG	1	0	0
	4	CTGCCTACAACTCAGTCTCTGG	1	0	0
	5	GGGCAGGCAGGACGGACTCCAGG	1	0	0
	6	GGAGTCCGTCCTGCCTGCCCTGG	1	0	0
	7	GAGTCCGTCCTGCCTGCCCTGGG	1	0	0
	8	GAAAAGGGTCCATTGGCCAAAGG	1	0	0
	9	GCCTGCAGAAAAGGGTCCATTGG	1	0	0
	10	TTGATGTGCTACAGGGAACATGG	1	0	0
	11	AGCGTTCTTGATGTGCTACAGGG	1	0	0
	12	CAGCGTTCTTGATGTGCTACAGG	1	0	0
	13	CTGTAGCACATCAAGAACGCTGG	1	0	0
	14	TGTAGCACATCAAGAACGCTGGG	1	0	0
	15	ATAGGCAATAATCATATAACAGG	1	0	0
	16	AGTGC GTTTCGCTGCAGGTAAGG	1	0	0
	17	GAGTGAGTGCGTTTCGCTGCAGG	1	0	0
	18	GTCAGGTTTGTGCGGTTATGAGG	1	0	0
	19	CGCTGCTGGTCAGGTTTGTGCGG	1	0	0
	20	AAACCTGACCAGCAGCGCAGAGG	1	0	0
	21	CCAGCAGCGCAGAGGAGCCGTGG	1	0	0
	22	CCACGGCTCCTCTGCGCTGCTGG	1	0	0
	23	CCA ACTATCTAACTCCACTCAGG	1	0	0
	24	CCTGAGTGGAGTTAGATAGTTGG	1	0	0

[1657] [表4] 各sgRNA在Jurkat细胞上对靶序列的活性

[1658]

基因	#	Cas9/sgRNA 未转染				Cas9/sgRNA 转染			
		总读段	插入	删除	插入缺失比例 (%)	总读段	插入	删除	插入缺失比例 (%)
A20	1	58003	46	55	0.20%	63455	17711	9469	42.80%
	2	40652	0	18	0.00%	46245	12025	6331	39.70%
	3	40652	0	18	0.00%	41702	301	92	0.90%
	4	40652	0	18	0.00%	4	2	2	0.00%
	5	40652	0	18	0.00%	52838	36339	4989	78.20%
	6	40652	0	18	0.00%	10641	5864	3460	87.60%
	7	40652	0	18	0.00%	40168	10298	4194	36.10%
	8	40652	0	18	0.00%	43044	9494	13398	53.20%
	9	40652	0	18	0.00%	46853	6629	2620	19.70%
	10	40652	0	18	0.00%	44573	17644	5168	51.20%
	11	63969	37	103	0.20%	61003	26844	22740	81.30%
	12	63969	37	103	0.20%	63321	949	1464	3.80%
基因	#	Cas9/sgRNA 未转染				Cas9/sgRNA 转染			
		总读段	插入	删除	插入缺失比例 (%)	总读段	插入	删除	插入缺失比例 (%)
DGK α	1	61246	0	4	0.00%	70438	4171	793	7.00%
	2	61246	0	4	0.00%	55262	7413	662	14.60%
	3	61246	0	4	0.00%	62354	19424	1546	33.60%
	4	59349	0	44	0.10%	58402	20072	5137	43.20%
	5	59349	0	44	0.10%	60718	14921	2484	28.70%
	6	59349	0	44	0.10%	67024	18760	2365	31.50%
	7	49807	0	0	0.00%	49459	26142	2877	58.70%
	8	49807	0	0	0.00%	65141	29740	3324	50.80%
	9	49807	0	0	0.00%	50760	30324	3742	67.10%
	10	49807	0	0	0.00%	61315	8953	4772	22.40%
	11	49807	0	0	0.00%	78876	61415	8416	88.50%
	12	49807	0	0	0.00%	64641	12255	1780	21.70%
		Cas9/sgRNA 未转染				Cas9/sgRNA 转染			
		总读段	插入	删除	插入缺失比例 (%)	总读段	插入	删除	插入缺失比例 (%)

[1659]

基因	#	总读段	插入	删除	插入缺失比例 (%)	总读段	插入	删除	插入缺失比例 (%)
EGR2	1	37189	0	0	0.00%	53321	11060	4974	30.10%
	2	37189	0	0	0.00%	48475	6809	1965	18.10%
	3	37189	0	0	0.00%	43800	8688	7796	37.60%
	4	37189	0	0	0.00%	43670	2921	569	8.00%
	5	37189	0	0	0.00%	34730	3002	497	10.10%
	6	37189	0	0	0.00%	46018	10502	1408	25.90%
	7	37189	0	0	0.00%	48537	5271	2475	16.00%
	8	37189	0	0	0.00%	36551	6457	686	19.50%
	9	37189	0	0	0.00%	37903	6210	1671	20.80%
	10	37189	0	0	0.00%	44855	9524	2320	26.40%
	11	37189	0	0	0.00%	39615	9368	2622	30.30%
	12	37189	0	0	0.00%	43995	2542	563	7.10%
	13					46228	289	62	0.76%
	14					50220	1323	821	4.27%
	15					33478	5638	1156	20.29%
	16					20489	1731	483	10.81%
	17					26353	3835	495	16.43%
	18					23901	1456	896	9.84%
	19					24352	3956	1672	23.11%
	20					11	0	0	0.00%
	21					34764	1522	359	5.41%
	22					31546	91	0	0.29%
	23					42734	10	0	0.02%
	24					32492	59	0	0.18%
	25					32243	1917	304	6.89%
	26					39333	868	328	3.04%
	27					36373	806	556	3.74%
	28					45819	2	26	0.06%
	29					53425	1159	584	3.26%
	30					36877	169	47	0.59%
	31					36317	0	76	0.21%

[1660]

	32					37941	829	122	2.51%
	33					47730	167	2	0.35%
	34					38753	347	62	1.06%
基因	#	Cas9/sgRNA 未转染				Cas9/sgRNA 转染			
		总读段	插入	删除	插入缺失比例 (%)	总读段	插入	删除	插入缺失比例 (%)
PPP2R2 D	5	38644	0	31	0.10%	48997	2891	240	6.40%
	6	50653	2	19	0.00%	48327	7669	1403	18.80%
	7	36764	0	0	0.00%	54465	670	70	1.40%
	8	36764	0	0	0.00%	45004	11382	1569	28.80%
	9	36764	0	0	0.00%	54094	17825	3635	39.70%
	10	36764	0	0	0.00%	47800	19253	3432	47.50%
	11	36764	0	0	0.00%	50362	966	129	2.20%
	12	36764	0	0	0.00%	42667	12810	2318	35.50%
	13					67258	1380	1050	3.61%
	14					69925	13321	3599	24.20%
	15					104102	21836	3254	24.10%
	16					77282	19219	7372	34.41%
	17					66732	3687	2227	8.86%
	18					96593	9524	1111	11.01%
	19					63082	11415	4155	24.68%
	20					57937	4360	676	8.69%
	21					67752	20314	4900	37.22%
	22					72814	2244	1198	4.73%
	23					79305	14047	1175	19.19%
	24					73629	2914	571	4.73%
	25					85222	5472	1905	8.66%
	26					73094	1937	288	3.04%
	27					94017	9895	6171	17.09%
	28					93118	8847	2464	12.15%
	29					77821	5007	1962	8.96%
		Cas9/sgRNA 未转染				Cas9/sgRNA 转染			

[1661]

基因	#	总读段	插入	删除	插入缺失比例 (%)	总读段	插入	删除	插入缺失比例 (%)
PD-1	1	68258	581	105	1.00%	77910	29123	7725	47.30%
	2	68258	581	105	1.00%	77866	1270	3816	6.50%
	3	68258	581	105	1.00%	66362	912	94	1.50%
	4	68258	581	105	1.00%	55936	41594	10324	92.80%
	5	68258	581	105	1.00%	65077	2554	192	4.20%
	6	68258	581	105	1.00%	71898	50678	10542	85.10%
	7	68258	581	105	1.00%	83902	17154	3246	24.30%
	8	68258	581	105	1.00%	79724	28304	7542	45.00%
	9	68258	581	105	1.00%	65936	10471	649	16.90%
	10	68258	581	105	1.00%	66937	0	29	0.00%
	11	68258	581	105	1.00%	77994	1135	754	2.40%
	12	68258	581	105	1.00%	67631	0	8	0.00%
	13	68258	581	105	1.00%	67161	30099	8037	56.80%
基因	#	Cas9/sgRNA 未转染				Cas9/sgRNA 转染			
		总读段	插入	删除	插入缺失比例 (%)	总读段	插入	删除	插入缺失比例 (%)
CTLA-4	7	68230	0	0	0	51173	3216	714	7.70%
	10	53694	3	18	0	40995	11760	1803	33.10%
	11	53694	3	18	0	55767	33107	3935	66.40%
	12	53333	0	0	0	54992	19469	8396	50.70%
基因	#	Cas9/sgRNA 未转染				Cas9/sgRNA 转染			
		总读段	插入	删除	插入缺失比例 (%)	总读段	插入	删除	插入缺失比例 (%)
	1	26039	3	2	0.00%	25450	10061	2453	49.20%
	2	26039	3	2	0.00%	24907	17380	2591	80.20%
	3	26039	3	2	0.00%	21950	14819	3291	82.50%
	4	26039	3	2	0.00%	20959	17708	1027	89.40%
	5	26039	3	2	0.00%	29570	26290	2120	96.10%
	6	37268	0	0	0.00%	32463	3663	1878	17.10%

[1662]

DGKζ	7	37268	0	0	0.00%	34154	6884	1706	25.20%
	8	37268	0	0	0.00%	32920	13190	4952	55.10%
	9	22544	7	12	0.10%	40374	5391	1209	16.30%
	10	22544	7	12	0.10%	28637	879	702	5.50%
	11	21780	0	0	0.00%	27636	9279	1859	40.30%
	12	21780	0	0	0.00%	20548	9474	2164	56.60%
	13	21780	0	0	0.00%	19161	9909	3016	67.50%
	14	53786	0	6	0.00%	36736	13	45	0.20%
	15	24528	0	10	0.00%	24319	12791	1446	58.50%
	16	24528	0	10	0.00%	20768	1520	140	8.00%
	17	24528	0	10	0.00%	26158	301	56	1.40%
基因	#	Cas9/sgRNA 未转染				Cas9/sgRNA 转染			
		总读段	插入	删除	插入缺失比例 (%)	总读段	插入	删除	插入缺失比例 (%)
	1	42428	375	573	2.23%	48887	35150	5438	83.02%
	2	42428	375	573	2.23%	44082	852	1852	6.13%
	3	42428	375	573	2.23%	49662	24418	7469	64.21%
	4	42428	375	573	2.23%	39571	20708	6428	68.58%
	5	42428	375	573	2.23%	52562	11325	2524	26.35%
	6	38575	7	14	0.10%	38990	3873	6433	26.43%
	7	38575	7	14	0.10%	36884	8795	1143	26.94%
	8	38575	7	14	0.10%	34674	5096	1843	20.01%
	9	38575	7	14	0.10%	38693	16101	4895	54.26%
	10					17614	4770	780	31.51%
	11					19411	1855	1416	16.85%
	12					14049	6887	1565	60.16%
	13					16272	2960	2087	31.02%
	14					18553	110	79	1.02%
	15					18062	1434	591	11.21%
	16					12053	2969	2423	44.74%
	17					14802	738	444	7.99%
	18					16943	395	154	3.24%
	19					18051	2953	1070	22.29%

[1663]

Tet2	20					14729	3041	474	23.86%
	21					18590	1074	320	7.50%
	22					19329	3304	1481	24.76%
	23					17420	36	19	0.32%
	24					20994	5582	1354	33.04%
	25					16860	2573	370	17.46%
	26					15137	1509	998	16.56%
	27					16035	635	185	5.11%
	28					14636	2734	1750	30.64%
	29					18893	133	45	0.94%
	30					15959	0	0	0.00%
	31					22627	216	126	1.51%
	32					15361	368	361	4.75%
	33					14501	1358	1939	22.74%
	34					3225	171	21	5.95%
	35					20968	725	209	4.45%
	36					15689	147	155	1.92%
	37					17405	239	18	1.48%
	38					20122	166	134	1.49%
	39					12585	370	106	3.78%
	40					15027	344	378	4.80%
基因	#	Cas9/sgRNA 未转染				Cas9/sgRNA 转染			
		总读段	插入	删除	插入缺失比例 (%)	总读段	插入	删除	插入缺失比例 (%)
PSGL-1	5	29368	0	9	0.03%	36584	8978	2453	31.25%
	6	29368	0	9	0.03%	35183	6859	639	21.31%
	7	33707	125	13	0.41%	24237	14697	2248	69.91%
	9	33707	125	13	0.41%	23911	9948	2001	49.97%
	10	33707	125	13	0.41%	30152	804	207	3.35%
	11	33707	125	13	0.41%	28425	95	6	0.36%
	12	33707	125	13	0.41%	25153	8931	1355	40.89%
	15	33707	125	13	0.41%	24798	2996	414	13.75%
	16	33707	125	13	0.41%	23116	8737	1192	42.95%

[1664]

		17	33707	125	13	0.41%	19094	10638	2066	66.53%
		27	29168	0	3	0.41%	29561	9316	1202	35.58%
		29	29168	0	3	0.01%	36720	5836	396	16.97%
		30	29168	0	3	0.01%	41685	3815	976	11.49%
基因	#	Cas9/sgRNA 未转染				Cas9/sgRNA 转染				
		总读段	插入	删除	插入缺失比例 (%)	总读段	插入	删除	插入缺失比例 (%)	
FAS	1					33594	14802	6170	62.43%	
	2					24634	7187	2668	40.01%	
	3					32994	21062	10555	95.83%	
	4					30374	1328	529	6.11%	
	5					40549	33991	4118	93.98%	
	6					51209	7460	1737	17.96%	
	7					24583	8997	9498	75.23%	
	8					28815	20681	6053	92.78%	
	9					29188	17689	4990	77.70%	
	10					25433	10120	9482	77.07%	
	11					29184	15700	7500	79.50%	
	12					25410	18254	1737	78.67%	
	13					28564	18560	1575	70.49%	
	14					2482	1241	325	63.09%	
	15					29819	14067	10479	82.32%	
	16					31325	8422	3600	38.38%	
基因	#	Cas9/sgRNA 未转染				Cas9/sgRNA 转染				
		总读段	插入	删除	插入缺失比例 (%)	总读段	插入	删除	插入缺失比例 (%)	
	1					33935	4337	1753	17.95%	
	2					42016	10713	3625	34.13%	
	3					56988	1195	951	3.77%	
	4					25006	3298	1295	18.37%	
	5					38511	43	16	0.15%	
	6					20361	598	340	4.61%	

[1665]	KDM6A	7				32084	2785	1161	12.30%
		8				31373	1616	523	6.82%
		9				5215	199	228	8.19%
		10				32955	4524	1097	17.06%
		11				38820	5726	1940	19.75%
		12				24536	72	12	0.34%
		13				42251	2640	475	7.37%
		14				44333	2018	628	5.97%
		15				33618	722	290	3.01%
		16				36221	466	250	1.98%
		17				40214	1357	261	4.02%
		18				31381	1958	714	8.51%
		19				40205	345	151	1.23%
		20				32494	9665	1761	35.16%
		21				37911	1286	381	4.40%
		22				30751	677	103	2.54%
		23				38635	8932	2445	29.45%
		24				44475	1263	978	5.04%

[1666] 3.2在人原代T细胞中筛选sgRNA

[1667] 基于上述实施例3.1中获得的Jurkat细胞中sgRNA活性的结果,选择在Jurkat细胞中具有较高活性的sgRNA(参见表3和表4中粗体示出的内容)在人原代T细胞中进行测试。

[1668] 将单链或双链的gRNA以及Cas9转化入人原代T细胞。表5中示出了测试的CRISPR/Cas9靶序列,表6中总结了各sgRNA的插入缺失比例。

[1669] [表5]人原代T细胞中的靶序列和错配

[1670]

基因	#	DNA 靶序列	序列号	错配		
				0 bp	1 bp	2 bp
A20	6	GGGCGTTCAGGACA CAGACTTGG	SEQ ID NO: 6	1	0	0
	11	GGCTTCCACAGACA CACCCATGG	SEQ ID NO: 11	1	0	0
DGKα	7	ATTGGGATGTGTCT GAGCTGAGG	SEQ ID NO: 19	1	0	0
	8	ATGAAAGAGATTGA CTATGATGG	SEQ ID NO: 20	1	0	0
	9	CTCTGTCTCTCAAGC TGAGTGGG	SEQ ID NO: 21	1	0	0
	11	CTCTCAAGCTGAGT GGGTCCGGG	SEQ ID NO: 23	1	0	0
	8+11	ATGAAAGAGATTGA CTATGATGG + CTCTCAAGCTGAGT GGGTCCGGG	SEQ ID NO: 20 + SEQ ID NO: 23	1	0	0
	9+11	CTCTGTCTCTCAAGC TGAGTGGG + CTCTCAAGCTGAGT GGGTCCGGG	SEQ ID NO: 21 + SEQ ID NO: 23	1	0	0
EGR2	1	TTGACATGACTGGA GAGAAGAGG	SEQ ID NO: 25	1	0	0
PPP2R2D	10	GTAGCGCCGTGATC CTAAATGGG	SEQ ID NO: 64	1	0	0
PD-1	4	GTCTGGGCGGTGCT ACAACTGGG	SEQ ID NO: 87	1	0	0
	6	CACCTACCTAAGAA CCATCCTGG	SEQ ID NO: 89	1	0	0
	1	GGCTAGGAGTCAGC GACATATGG	SEQ ID NO: 109	1	0	0
	2	GCTAGGAGTCAGCG ACATATGGG	SEQ ID NO: 110	1	0	0

[1671]	DGKζ	3	CTAGGAGTCAGCGA CATATGGGG	SEQ ID NO: 111	1	0	0
		4	GTACTGTGTAGCCA GGATGCTGG	SEQ ID NO: 112	1	0	0
		5	ACGAGCACTCACCA GCATCCTGG	SEQ ID NO: 113	1	0	0

[1672] [表6] 各gRNA在人原代T免疫细胞上对靶序列的活性

基因	#	Cas9/sgRNA 未转染				Cas9/sgRNA 转染			
		总读段	插入	删除	插入缺失比例 (%)	总读段	插入	删除	插入缺失比例 (%)
A20	6	32158	0	26	0.10%	31976	190	3865	12.68%
	11	32158	0	26	0.10%	30008	354	3324	12.26%
DGKα	7	35903	15	7	0.10%	29446	332	4465	16.29%
	8	35903	15	7	0.10%	40656	395	13739	34.76%
	9	35903	15	7	0.10%	48602	353	3263	7.44%
	11	35903	15	7	0.10%	43261	1222	17621	43.56%
	8+11	35903	15	7	0.10%	42504	184	21684	51.45%
	9+11	35903	15	7	0.10%	42025	41	5546	13.29%
EGR2	1	55074	26	67	0.20%	42275	986	5176	14.58%
PPP2R2D	10	35903	15	7	0.10%	46205	1505	5532	15.23%
PD-1	4	31063	0	13	0.00%	62882	8104	23113	49.64%
	6	31063	0	13	0.00%	93252	2431	8707	11.94%
DGKζ	1	20278	0	11	0.10%	56415	1384	3898	9.36%
	2	20278	0	11	0.10%	49114	2390	4923	14.89%
	3	20278	0	11	0.10%	65225	6738	3929	16.35%
	4	20278	0	11	0.10%	36502	1303	3477	13.10%
	5	20278	0	11	0.10%	28580	2945	10392	46.67%

[1674] 类似地,基于上述实施例3.1中获得的Jurkat细胞中sgRNA活性的结果,选择在Jurkat细胞中具有较高活性的PSGL-1#17 sgRNA,测试其在人原代T细胞中的活性。

[1675] 此外,通过电穿孔(Neon,Thermo Scientific)用4μg SpCas9蛋白以及1μg体外转录并经AP处理的sgRNA转染活化的人原代T细胞。5天后从各T细胞中分离提取gDNA,并通过靶向深度测序分析插入缺失效率(图18A)。此外,借助流式细胞术(Attune Flow cytometry,Thermo Scientific)分析T细胞表面上的PSGL-1表达,从而确认PSGL-1敲除(图18B和图18C)。

[1676] 图17a-图17c示出了在Jurkat细胞中用于hPSGL-1 sgRNA筛选的分析结果。这些图

是示出了敲除后不表达PSGL-1的Jurkat细胞(hPSGL-1阴性细胞)的水平和插入缺失效率(图17a)以及敲除后Jurkat细胞表面上PSGL-1的表达水平(图17b和图17c)的图。

[1677] 图18示出了在人原代T细胞中hPSGL-1敲除(KO)实验的结果,其中,(A)示出了插入缺失效率;(B)示出了敲除后不表达PSGL-1的T细胞的水平;(C)示出了敲除后T细胞表面上PSGL-1的表达水平。其结果是,证实了通过递送Cas9蛋白和gRNA复合体有效敲除了PSGL-1,因此表面蛋白PSGL-1不能通过流式细胞术观察到。

[1678] 实施例4:Jurkat细胞的活化和细胞因子分泌的促进

[1679] 在其中导入有Cas9蛋白和sgRNA的Jurkat细胞中,与导入的sgRNA的靶区域对应的基因组DNA序列被切割,并且被切割的DNA序列附近的区域通过经由NHEJ进行的删除、插入和/或置换而突变,从而引起基因(被切割的DNA序列位于该基因上)的敲除。

[1680] 对于如实施例1所述利用电穿孔用Cas9蛋白和sgRNA转染的Jurkat细胞,在电穿孔后将所述细胞培养7天并使用CD3dynabeads(Miltenyi Biotec)或CD3/28 dynabeads(Miltenyi Biotec)活化所述细胞。

[1681] 24小时后分别利用流式细胞术和ELISA对CD25(其为IL-2受体)的表达和IFN- γ 的释放水平进行分析。

[1682] 首先利用流式细胞术对CD25(IL-2受体)的表达水平进行测量。在实施导入后对用Cas9蛋白和sgRNA转染的Jurkat细胞培养7天,并以3:1的比例(珠:细胞;数目)使用CD3或CD3/28 dynabeads(Miltenyi Biotec)再次刺激,随后测量CD25的表达。

[1683] 在细胞活化后1天实施表型分析。用补充有1% (v/v) 胎牛血清(FBS)的PBS(磷酸盐缓冲盐水)对经珠-再次刺激(活化)的细胞进行洗涤,并用PE缀合的抗CD25抗体(BD Bioscience)在4℃染色30min。

[1684] 对获得的细胞进行洗涤并在PBS中重悬,随后在BD ACCURI C6(BD Biosciences)上实施流式细胞术,并以中位荧光强度(median fluorescence intensity,MFI)对CD25表达的水平进行测量。

[1685] 为进行比较,以相同的方式针对未导入Cas9蛋白和sgRNA的野生型细胞和未利用CD3或CD3/28 dynabeads处理的细胞实施流式细胞术。

[1686] 获得的CD25表达水平(CD25 MFI)示于图1-图4中。

[1687] 分别地,图1示出了细胞中CD25的MFI,所述细胞中使用针对DGK α 的sgRNA(#11;以DGK α #11表示)将DGK α 基因敲除;

[1688] 图2示出了细胞中CD25的MFI,所述细胞中使用针对A20的sgRNA(#11;以A20#11表示)将A20基因敲除;

[1689] 图3示出了细胞中CD25的MFI,所述细胞中使用针对EGR2的sgRNA(#1;以EGR2#1表示)将EGR2基因敲除;

[1690] 图4示出了细胞中CD25的MFI,所述细胞中使用针对PPP2R2D的sgRNA(#10;以PPP2R2D#10表示)将PPP2R2D基因敲除。

[1691] 如图1-图4所示,在未用CD3或CD3/28 dynabeads处理细胞的情况下,基因敲除存在与否不影响CD25的表达水平;而在利用CD3或CD3/28 dynabeads处理细胞的情况下,确认了与野生型相比而言,当基因被敲除时CD25的表达水平显著上升。

[1692] 此外,利用ELISA测试了IFN- γ (一种细胞因子)的分泌水平。

[1693] 如前所述,在将通过CD3或CD3/28 dynabeads进行再次刺激的Jurkat细胞活化36小时后,收集培养基并使用稀释缓冲液(由ELISA试剂盒提供,Biolegend)将其稀释至1/100或1/200的比例(w/v),随后使用ELISA试剂盒(Biolegend)进行显色,并使用分光光度计(MULTISCAN GO, Thermo Scientific)进行定量。

[1694] 为进行比较,以相同的方式针对其中未导入Cas9蛋白和sgRNA的野生型细胞实施ELISA。

[1695] 获得的结果示于图5中。

[1696] 图5显示了:其中使用针对DGK α 的sgRNA(#11;以DGK α #11表示)敲除DGK α 基因的细胞培养基中的IFN- γ 水平;其中使用针对A20的sgRNA(#11;以A20#11表示)敲除A20基因的细胞培养基中的IFN- γ 水平;其中使用针对EGR2的sgRNA(#1;以EGR2#1表示)敲除EGR2基因的细胞培养基中的IFN- γ 水平(IFN- γ 的水平单位:pg/mL)。

[1697] 如图5所示,确认了与野生型相比而言,当基因被敲除时IFN- γ 的分泌量显著升高。

[1698] 实施例5:人原代T细胞的活化和细胞因子分泌的促进

[1699] 参照上文实施例4中所述的方法,利用CD3珠活化用Cas9蛋白和sgRNA转染的人原代T细胞(珠:细胞比例分别为1:1、2:1和3:1)。2天后,通过ELISA(IFN- γ 或IL-2ELISA试剂盒,Biolegend)测量IFN- γ 和IL-2的分泌水平。

[1700] 获得的结果示于图6和图7中。

[1701] 图6显示了:其中使用针对DGK α 的sgRNA(#11;以DGK α #11表示)敲除DGK α 基因的细胞培养基中的IFN- γ 水平;其中使用针对DGK α 的sgRNA(组合使用#8和#11;以DGK α #8+11表示)敲除DGK α 基因的细胞培养基中的IFN- γ 水平;其中使用针对DGK ζ 的sgRNA(#5;以DGK ζ #5表示)敲除DGK ζ 基因的细胞培养基中的IFN- γ 水平;其中使用针对A20的sgRNA(#11;以A20#11表示)敲除A20基因的细胞培养基中的IFN- γ 水平(IFN- γ 的水平单位:pg/mL)。

[1702] 图7显示了:其中使用DGK α #11敲除DGK α 基因的细胞培养基中的IL-2水平;其中使用DGK α #8+11敲除DGK α 基因的细胞培养基中的IL-2水平;其中使用DGK ζ #5敲除DGK ζ 基因的细胞培养基中的IL-2水平;其中使用A20#11敲除A20基因的细胞培养基中的IL-2水平(IL-2的水平单位:pg/mL)。

[1703] 在图6和图7中,使用“AAVS1”作为细胞的阴性对照,其中的AAVS1位点被CRISPR系统切割。

[1704] 如图6和图7所示,确认了与野生型相比而言,当基因被敲除时细胞因子(例如IFN- γ 和IL-2)的分泌量显著升高。

[1705] 显示了Jurkat细胞和人原代T细胞中CD25表达和细胞因子分泌增加的这些结果表明当基因被敲除时TCR介导的活化信号增强,并且T细胞的免疫功能能够由于该活性升高而增强。

[1706] 实施例6:CAR-T细胞的活化和细胞因子分泌的促进

[1707] 人外周血T细胞(pan-T细胞)购自STEMCELL TECHNOLOGIES。使用补充有50U/mL hIL-2和5ng/mL hIL-7的X-VIVO 15培养基进行细胞培养。使用抗CD3/28 Dynabeads(ThermoFisher Scientific)对细胞进行活化(珠:细胞比为3:1)。

[1708] 活化24小时后,在包覆有纤维连接蛋白的板上将T细胞与139-CAR慢病毒混合48小

时。139-CAR是能够特异性识别EGFRvIII并诱导免疫应答的CAR。随后利用4D-Nucleofector (Lonza) 通过电穿孔将40 μ g重组酿脓链球菌Cas9蛋白 (Toolgen, Korea) 和10 μ g化学合成的tracr/crRNA(Integrated DNA Technologies) 导入细胞。

[1709] 对于体外实验,利用Cell Trace(ThermoFisher Scientific)对U87vIII癌细胞进行预染色,将其与139 CAR-T细胞以合适的比例共培养。此时添加或不添加10ng/mL TGF- β 1或0.5 μ g/mL PGE2实施培养。与癌细胞系共培养后,利用7-氨基放线菌素D(7-AAD)对细胞进行染色,来进行细胞毒性测试实验。在Attune NxT Acoustic Focusing Cytometer上收集经染色的样品,并利用FlowJo进行分析。

[1710] 利用公式[(裂解样品值% - 裂解最小值%) / (裂解最大值% [100%] - 裂解最小值%)] \times 100%计算细胞毒性。此外,还利用ELISA试剂盒(Biolegend)进行共培养上清液分析,来确定IL-2和IFN- γ 的含量。对于139 CAR-T细胞的细胞增殖实验,利用Cell Trace对139 CAR-T细胞进行染色,并将其与靶癌细胞系U87vIII共培养,随后使用流式细胞术对139 CAR-T细胞中Cell Trace的稀释程度进行测量。

[1711] 根据实验设计(图8a,A),利用靶向DGK α 或DGK ζ 的单一Cas9/gRNA核糖核蛋白(RNP)复合体进行递送的139 CAR-T细胞上的DGK α 和DGK ζ 的插入缺失效率分别为75.9%和93.5%(图8a,B)。

[1712] 通过电转将分别靶向DGK α 和DGK ζ 的两种gRNA导入细胞来生成对于DGK α 和DGK ζ 而言双阴性的139 CAR-T细胞。其结果是DGK α 和DGK ζ 的插入缺失效率分别为49.2%和92.4%(图8a,B)。

[1713] 使用靶向深度测序确认了各DGK α 和DGK ζ 的gRNA无显著脱靶效应(图8b)。

[1714] 此外,观察到与野生型139 CAR-T细胞相比而言,DGK α 、DGK ζ 和DGK $\alpha\zeta$ KO(敲除)的139 CAR-T细胞具有显著升高的细胞毒性、细胞因子产生能力和增殖能力(图9a A、B和图9b)。

[1715] 有意思的是,与DGK α 或DGK ζ 单敲除的139 CAR-T细胞相比而言,DGK $\alpha\zeta$ KO 139 CAR-T细胞显示出更显著的细胞因子释放升高,认为这是DGK α 和DGK ζ 的协同效应。认为这些敲除DGK的139 CAR-T细胞中效应物功能的升高归因于CD3末端信号的升高,即暴露于抗原后CAR的高表达和ERK1/2的升高(图10A、B)。

[1716] 此外,虽然在敲除DGK的139 CAR-T细胞中信号被显著活化,在不存在靶癌细胞时没有观察到基础细胞因子升高,表明DGK敲除的高安全性(图11A)。进而,与139 CAR-T细胞相比而言,在敲除DGK的139CAR-T细胞中PD-11和TIM-33(它们为耗竭标志物)的表达并未升高(图11B)。这些结果表明,即便在长时间暴露于抗原后DGK的敲除也不会促进T细胞耗竭(图11B)。

[1717] 通过用信号1免疫阻遏性抑制剂(例如TGF- β 1和PGE2)进行处理,139 CAR-T细胞的抗癌效果显著受损;而在DGK $\alpha\zeta$ KO 139 CAR-T细胞的情况下,确认了即便在抑制性细胞溶解因子的存在下,细胞毒性和细胞因子释放也得到保持(图12A、图12B)。

[1718] 这些结果表明可通过使用CRISPR/Cas9使DGK基因失活来活化T细胞功能。

[1719] 换言之,确认了DGK基因的失活能增强CD3末端信号,从而增强CAR-T细胞的增殖和抗癌功能。

[1720] 此外,CAR-T细胞中DGK $\alpha\zeta$ (两种亚型)的敲除(KO)没有表现出耗竭标志物的显著降

低,并且对于免疫阻遏性细胞溶解因子(例如TGF- β 和前列腺素E2(PGE2))的应答性较低。

[1721] 因此,证实了通过CRISPR/Cas9进行的DGK KO能够增强T细胞效应物功能的升高。

[1722] 实施例7:NK(自然杀伤)细胞的活化和细胞因子分泌的促进

[1723] 7.1 NK 92细胞系和人原代NK细胞的培养

[1724] NK92细胞系购自ATCC(CRL-2407),原代NK细胞购自STEMCELL TECHNOLOGY,并根据提供的方案进行培养。

[1725] 在含有10%FCS(胎牛血清)且补充有100 μ g/ml链霉素、100U/ml青霉素、2mM UltraGlutamine I、200-300U/ml IL-2和10U/ml IL-15的RPMI 1640培养基(WellGene)中培养NK92细胞。

[1726] 7.2通过电穿孔进行导入

[1727] 为敲除NK92细胞系中的DGK α 和DGK ζ ,利用Neon电转仪(Thermo Fisher Scientific)在1200V、10ms和3个脉冲下实施电转。对于原代NK细胞,使用1200V、20ms和3个脉冲。

[1728] 将4 μ g重组酿脓链球菌Cas9蛋白(Toolgen,Korea)和1 μ g化学合成的tracr/crRNA(Integrated DNA Technologies)孵育20分钟,从而获得Cas9 RNP复合体。

[1729] 将在R缓冲液中重悬的 2×10^5 个NK92细胞加入至(接触)预孵育的Cas9 RNP复合体以实施电穿孔。在此之后,以 4×10^5 个细胞/mL的浓度将细胞铺板于培养基中。

[1730] 实验中使用的crRNA靶向序列如下:

[1731] DGK α :CTCTCAAGCTGAGTGGGTCC

[1732] DGK ζ :ACGAGCACTCACCAGCATCC。

[1733] 7.3体外杀伤测定

[1734] 为分析NK92细胞和原代NK细胞的细胞毒性,在U形底96孔板中将该细胞与经CellTrace远红外(CellTrace Far Red,Invitrogen)染色的Raji细胞或 1×10^5 个K562细胞共培养。共培养18小时后收集细胞并用7-AAD染色,随后用流式细胞术进行分析。全部细胞毒性实验实施3次。

[1735] 结果示于图13中。确认了在NK92细胞和原代NK细胞中DGK α 的敲除效率(KO效率)优异(图13A和图13B)。此外,通过测量7-AAD阳性Raji细胞证实了NK-92的杀伤活性,表明DGK α 敲除使得细胞毒性增加。

[1736] 具体而言,这些结果证实了即便是对于已知难以进行基因操纵的NK细胞而言,也可以有效地操纵其免疫功能。

[1737] 实施例8:NKT(自然杀伤T)细胞的活化和细胞因子分泌的促进

[1738] 8.1 NKT细胞培养

[1739] 人PBMC购自STEMCELL TECHNOLOGY(Canada)。以 1×10^6 个细胞/mL的浓度将这些细胞铺板于补充有10%FBS且添加有1000U/ml干扰素- γ (Pepro Tech)的RPMI培养基。在培养基中添加5天50ng/ml的抗人OKT-3(Biolegend),添加20天400U/ml的IL-2(Pepro Tech)。

[1740] 8.2通过电穿孔进行导入

[1741] 为敲除NKT细胞系中的DGK α 、DGK ζ 和PD1,利用Neon电转仪(Thermo Fisher Scientific)在1550V、10ms和3个脉冲下实施电转。

[1742] 将4 μ g重组酿脓链球菌Cas9蛋白(Toolgen,Korea)和1 μ g化学合成的tracr/crRNA

(Integrated DNA Technologies) 孵育20分钟,从而获得Cas9 RNP复合体。

[1743] 将在R缓冲液中重悬的 2×10^5 个NKT细胞加入至(接触)预孵育的Cas9 RNP复合体以实施电穿孔。在此之后,以 4×10^5 个细胞/mL的浓度将细胞接种于培养基中。

[1744] 实验中使用的crRNA靶向序列如下:

[1745] DGK α :CTCTCAAGCTGAGTGGGTCC

[1746] DGK ζ :ACGAGCACTCACCAGCATCC

[1747] PD-11:GTCTGGGCGGTGCTACAACCTGGG。

[1748] 8.3体外杀伤测定

[1749] 为分析NKT细胞的细胞毒性,在U形底96孔板中将NKT细胞与经CellTrace远红外(Invitrogen)染色的 2×10^4 个U87vIII细胞共培养。共培养18小时后收集细胞并用7-AAD染色,随后用流式细胞术进行分析。全部细胞毒性实验实施3次。

[1750] 其结果是,如图14所示,确认了CRISPR/Cas9系统有效实施了人NKT细胞中DGK α 和DGK ζ 的敲除。

[1751] 借助深度测序确认了插入缺失效率(图14A);使用台盼兰染色对经CRISPR/Cas9处理的NKT细胞进行分析,确认了细胞生长(图14B)和细胞存活力保持良好(存活力=存活细胞数/总细胞数)。此外,通过Western印迹确认了DGK α 和DGK ζ 敲除也在蛋白水平发生(图14D)。

[1752] 进而,如图15所示,确认了DGK α 和DGK ζ 敲除改善了NKT细胞的效应物功能。

[1753] 利用CellTrace(Thermo fisher)处理U87vIII、H460和K562细胞并在96孔板中以E:T(效应物细胞:靶细胞比例)为20:1的比例对细胞培养18小时。通过流式细胞术分析7-AAD阳性癌细胞凋亡水平的结果表明DGK α 和DGK ζ 敲除提高了相应NKT细胞的NKT杀伤活性。DGK α 和DGK ζ 分别敲除也具有提高杀伤活性的效果,然而进一步证明了同时敲除两个基因更显著地增强了杀伤活性(图15A)。

[1754] 另一方面,借助ELISA(IFN试剂盒,Biolegend)测定IFN分泌的结果显示DGK α 和DGK ζ 敲除增强相应细胞的IFN释放能力。虽然DGK α 和DGK ζ 分别敲除也具有增强IFN分泌的优异效果,确认了同时敲除两个基因进一步增强IFN分泌(图15B)。

[1755] 此外,如图16所示,确认了人NKT细胞中CRISPR/Cas9介导的PD-1敲除增强了NKT细胞的效应物功能。在NKT细胞中使用CRISPR-Cas9诱导PD-1敲除,借助靶向深度测序分析PD-1的敲除效率。此外,U87vIII细胞和NKT细胞共培养证实了通过PD-1敲除使得NKT细胞作为抗癌效应物的功能。在96孔板中以E:T(效应物细胞:靶细胞)=50:1的比例利用Cell Trace(Thermo fisher)对U87vIII细胞处理18小时,并利用流式细胞术分析7-AAD阳性癌细胞来分析杀伤活性。

[1756] 其结果是,证实了归因于CRISPR/Cas9的PD-1基因中的高插入缺失效率(图16A),由此改善了其细胞毒性(图16B)。

[1757] 总之,上述结果表明CRISPR/Cas9介导的免疫调节基因(例如DGK)敲除可在多种类型的免疫细胞中具有显著的免疫增强效果。

[1758] DGK基因敲除的这些生物效果显示,可以通过改善免疫功能来将免疫细胞(包括T细胞、NK细胞和NKT细胞等)开发为临床可用细胞形式的免疫治疗剂。

[1759] 工业实用性

[1760] 通过经修饰的免疫系统(其中,基于经人工操纵的免疫调节因子和含有该经人工操纵的免疫调节因子的细胞,对该经修饰的免疫系统的功能进行了人工操纵)可获得有效的免疫细胞治疗剂。

[1761] 例如,当通过本发明的组合物或方法对免疫调节因子进行人工控制时,可以改善涉及免疫细胞的存活、增殖、持久性、细胞毒性、细胞因子释放和/或浸润等方面的免疫效力。

[0001]	<110>	株式会社图尔金	
[0002]	<120>	经操纵的免疫调节因子以及由此改变的免疫力	
[0003]	<130>	OPP17-041	
[0004]	<150>	KR 10-2016-0103308	
[0005]	<151>	2016-08-12	
[0006]	<150>	US 62/502,822	
[0007]	<151>	2017-05-08	
[0008]	<160>	289	
[0009]	<170>	KoPatentIn 3.0	
[0010]	<210>	1	
[0011]	<211>	23	
[0012]	<212>	DNA	
[0013]	<213>	Homo sapiens(智人)	
[0014]	<400>	1	
[0015]		cttgtggcgc tgaaaacgaa cgg	23
[0016]	<210>	2	
[0017]	<211>	23	
[0018]	<212>	DNA	
[0019]	<213>	Homo sapiens(智人)	
[0020]	<400>	2	
[0021]		atgccacttc tcagtacatg tgg	23
[0022]	<210>	3	
[0023]	<211>	23	
[0024]	<212>	DNA	
[0025]	<213>	Homo sapiens(智人)	
[0026]	<400>	3	
[0027]		gccacttctc agtacatgtg ggg	23
[0028]	<210>	4	
[0029]	<211>	23	
[0030]	<212>	DNA	
[0031]	<213>	Homo sapiens(智人)	
[0032]	<400>	4	
[0033]		gccccacatg tactgagaag tgg	23
[0034]	<210>	5	
[0035]	<211>	23	
[0036]	<212>	DNA	
[0037]	<213>	Homo sapiens(智人)	
[0038]	<400>	5	

[0039]	tcagtacatg tggggcgttc agg	23
[0040]	<210> 6	
[0041]	<211> 23	
[0042]	<212> DNA	
[0043]	<213> Homo sapiens(智人)	
[0044]	<400> 6	
[0045]	gggcgttcag gacacagact tgg	23
[0046]	<210> 7	
[0047]	<211> 23	
[0048]	<212> DNA	
[0049]	<213> Homo sapiens(智人)	
[0050]	<400> 7	
[0051]	cacagacttg gtactgagga agg	23
[0052]	<210> 8	
[0053]	<211> 23	
[0054]	<212> DNA	
[0055]	<213> Homo sapiens(智人)	
[0056]	<400> 8	
[0057]	ggcgctgttc agcacgctca agg	23
[0058]	<210> 9	
[0059]	<211> 23	
[0060]	<212> DNA	
[0061]	<213> Homo sapiens(智人)	
[0062]	<400> 9	
[0063]	cacgcaactt taaattccgc tgg	23
[0064]	<210> 10	
[0065]	<211> 23	
[0066]	<212> DNA	
[0067]	<213> Homo sapiens(智人)	
[0068]	<400> 10	
[0069]	cggggccttg ctatgatact cgg	23
[0070]	<210> 11	
[0071]	<211> 23	
[0072]	<212> DNA	
[0073]	<213> Homo sapiens(智人)	
[0074]	<400> 11	
[0075]	ggcttcaca gacacacca tgg	23
[0076]	<210> 12	
[0077]	<211> 23	

[0078]	<212>	DNA	
[0079]	<213>	Homo sapiens(智人)	
[0080]	<400>	12	
[0081]	tgaagtccac	ttcgggccat ggg	23
[0082]	<210>	13	
[0083]	<211>	23	
[0084]	<212>	DNA	
[0085]	<213>	Homo sapiens(智人)	
[0086]	<400>	13	
[0087]	ctgtacgaca	cggacagaaa tgg	23
[0088]	<210>	14	
[0089]	<211>	23	
[0090]	<212>	DNA	
[0091]	<213>	Homo sapiens(智人)	
[0092]	<400>	14	
[0093]	tgtacgacac	ggacagaaat ggg	23
[0094]	<210>	15	
[0095]	<211>	23	
[0096]	<212>	DNA	
[0097]	<213>	Homo sapiens(智人)	
[0098]	<400>	15	
[0099]	cacggacaga	aatgggatcc tgg	23
[0100]	<210>	16	
[0101]	<211>	23	
[0102]	<212>	DNA	
[0103]	<213>	Homo sapiens(智人)	
[0104]	<400>	16	
[0105]	gatgcgagtg	gctgaatacc tgg	23
[0106]	<210>	17	
[0107]	<211>	23	
[0108]	<212>	DNA	
[0109]	<213>	Homo sapiens(智人)	
[0110]	<400>	17	
[0111]	gagtggctga	atacctggat tgg	23
[0112]	<210>	18	
[0113]	<211>	23	
[0114]	<212>	DNA	
[0115]	<213>	Homo sapiens(智人)	
[0116]	<400>	18	

[0117]	agtggctgaa tacctggatt ggg	23
[0118]	<210> 19	
[0119]	<211> 23	
[0120]	<212> DNA	
[0121]	<213> Homo sapiens(智人)	
[0122]	<400> 19	
[0123]	attgggatgt gtctgagctg agg	23
[0124]	<210> 20	
[0125]	<211> 23	
[0126]	<212> DNA	
[0127]	<213> Homo sapiens(智人)	
[0128]	<400> 20	
[0129]	atgaaagaga ttgactatga tgg	23
[0130]	<210> 21	
[0131]	<211> 23	
[0132]	<212> DNA	
[0133]	<213> Homo sapiens(智人)	
[0134]	<400> 21	
[0135]	ctctgtctct caagctgagt ggg	23
[0136]	<210> 22	
[0137]	<211> 23	
[0138]	<212> DNA	
[0139]	<213> Homo sapiens(智人)	
[0140]	<400> 22	
[0141]	tctctcaagc tgagtgggtc cgg	23
[0142]	<210> 23	
[0143]	<211> 23	
[0144]	<212> DNA	
[0145]	<213> Homo sapiens(智人)	
[0146]	<400> 23	
[0147]	ctctcaagct gagtgggtcc ggg	23
[0148]	<210> 24	
[0149]	<211> 23	
[0150]	<212> DNA	
[0151]	<213> Homo sapiens(智人)	
[0152]	<400> 24	
[0153]	caagctgagt ggggtccgggc tgg	23
[0154]	<210> 25	
[0155]	<211> 23	

[0156]	<212>	DNA	
[0157]	<213>	Homo sapiens(智人)	
[0158]	<400>	25	
[0159]	ttgacatgac tggagagaag agg		23
[0160]	<210>	26	
[0161]	<211>	23	
[0162]	<212>	DNA	
[0163]	<213>	Homo sapiens(智人)	
[0164]	<400>	26	
[0165]	gactggagag aagaggtcgt tgg		23
[0166]	<210>	27	
[0167]	<211>	23	
[0168]	<212>	DNA	
[0169]	<213>	Homo sapiens(智人)	
[0170]	<400>	27	
[0171]	gagacgggag caaagctgct ggg		23
[0172]	<210>	28	
[0173]	<211>	23	
[0174]	<212>	DNA	
[0175]	<213>	Homo sapiens(智人)	
[0176]	<400>	28	
[0177]	agagacggga gcaaagctgc tgg		23
[0178]	<210>	29	
[0179]	<211>	23	
[0180]	<212>	DNA	
[0181]	<213>	Homo sapiens(智人)	
[0182]	<400>	29	
[0183]	tggtttctag gtgcagagac ggg		23
[0184]	<210>	30	
[0185]	<211>	23	
[0186]	<212>	DNA	
[0187]	<213>	Homo sapiens(智人)	
[0188]	<400>	30	
[0189]	taagtgaagg tctggtttct agg		23
[0190]	<210>	31	
[0191]	<211>	23	
[0192]	<212>	DNA	
[0193]	<213>	Homo sapiens(智人)	
[0194]	<400>	31	

[0195]	tgcccatgta agtgaaggtc tgg	23
[0196]	<210> 32	
[0197]	<211> 23	
[0198]	<212> DNA	
[0199]	<213> Homo sapiens(智人)	
[0200]	<400> 32	
[0201]	gaacttgccc atgtaagtga agg	23
[0202]	<210> 33	
[0203]	<211> 23	
[0204]	<212> DNA	
[0205]	<213> Homo sapiens(智人)	
[0206]	<400> 33	
[0207]	tccattgacc ctcagtaccc tgg	23
[0208]	<210> 34	
[0209]	<211> 23	
[0210]	<212> DNA	
[0211]	<213> Homo sapiens(智人)	
[0212]	<400> 34	
[0213]	tatgccttct gggtagcagc tgg	23
[0214]	<210> 35	
[0215]	<211> 23	
[0216]	<212> DNA	
[0217]	<213> Homo sapiens(智人)	
[0218]	<400> 35	
[0219]	tgagtgcagg catcttgcaa ggg	23
[0220]	<210> 36	
[0221]	<211> 23	
[0222]	<212> DNA	
[0223]	<213> Homo sapiens(智人)	
[0224]	<400> 36	
[0225]	gagtgcaggc atcttgcaag ggg	23
[0226]	<210> 37	
[0227]	<211> 23	
[0228]	<212> DNA	
[0229]	<213> Homo sapiens(智人)	
[0230]	<400> 37	
[0231]	gatgaggctg tggttgaagc tgg	23
[0232]	<210> 38	
[0233]	<211> 23	

[0234]	<212>	DNA	
[0235]	<213>	Homo sapiens(智人)	
[0236]	<400>	38	
[0237]	ccactggcca caggaccct ggg		23
[0238]	<210>	39	
[0239]	<211>	23	
[0240]	<212>	DNA	
[0241]	<213>	Homo sapiens(智人)	
[0242]	<400>	39	
[0243]	gggacatggt gcacacacc agg		23
[0244]	<210>	40	
[0245]	<211>	23	
[0246]	<212>	DNA	
[0247]	<213>	Homo sapiens(智人)	
[0248]	<400>	40	
[0249]	gagtacaggt ggtccaggtc agg		23
[0250]	<210>	41	
[0251]	<211>	23	
[0252]	<212>	DNA	
[0253]	<213>	Homo sapiens(智人)	
[0254]	<400>	41	
[0255]	gcggagagta caggtggtcc agg		23
[0256]	<210>	42	
[0257]	<211>	23	
[0258]	<212>	DNA	
[0259]	<213>	Homo sapiens(智人)	
[0260]	<400>	42	
[0261]	gcggtggcgg agagtacagg tgg		23
[0262]	<210>	43	
[0263]	<211>	23	
[0264]	<212>	DNA	
[0265]	<213>	Homo sapiens(智人)	
[0266]	<400>	43	
[0267]	tctcctgcac agccagaata agg		23
[0268]	<210>	44	
[0269]	<211>	23	
[0270]	<212>	DNA	
[0271]	<213>	Homo sapiens(智人)	
[0272]	<400>	44	

[0273]	acgcagaagg gtcctggtag agg	23
[0274]	<210> 45	
[0275]	<211> 23	
[0276]	<212> DNA	
[0277]	<213> Homo sapiens(智人)	
[0278]	<400> 45	
[0279]	aggtggtggg taggccagag agg	23
[0280]	<210> 46	
[0281]	<211> 23	
[0282]	<212> DNA	
[0283]	<213> Homo sapiens(智人)	
[0284]	<400> 46	
[0285]	cccaagccag ccacggaccc agg	23
[0286]	<210> 47	
[0287]	<211> 23	
[0288]	<212> DNA	
[0289]	<213> Homo sapiens(智人)	
[0290]	<400> 47	
[0291]	acctgggtcc gtggctggct tgg	23
[0292]	<210> 48	
[0293]	<211> 23	
[0294]	<212> DNA	
[0295]	<213> Homo sapiens(智人)	
[0296]	<400> 48	
[0297]	aagagacctg ggtccgtggc tgg	23
[0298]	<210> 49	
[0299]	<211> 23	
[0300]	<212> DNA	
[0301]	<213> Homo sapiens(智人)	
[0302]	<400> 49	
[0303]	ggatcattgg gaagagacct ggg	23
[0304]	<210> 50	
[0305]	<211> 23	
[0306]	<212> DNA	
[0307]	<213> Homo sapiens(智人)	
[0308]	<400> 50	
[0309]	gggatcattg ggaagagacc tgg	23
[0310]	<210> 51	
[0311]	<211> 23	

[0312]	<212>	DNA	
[0313]	<213>	Homo sapiens(智人)	
[0314]	<400>	51	
[0315]	caggatagtc tgggatcatt ggg		23
[0316]	<210>	52	
[0317]	<211>	23	
[0318]	<212>	DNA	
[0319]	<213>	Homo sapiens(智人)	
[0320]	<400>	52	
[0321]	ggaaagaatc caggatagtc tgg		23
[0322]	<210>	53	
[0323]	<211>	23	
[0324]	<212>	DNA	
[0325]	<213>	Homo sapiens(智人)	
[0326]	<400>	53	
[0327]	cagtgccaga gagacctaca tgg		23
[0328]	<210>	54	
[0329]	<211>	23	
[0330]	<212>	DNA	
[0331]	<213>	Homo sapiens(智人)	
[0332]	<400>	54	
[0333]	ctgtaccatg taggtctctc tgg		23
[0334]	<210>	55	
[0335]	<211>	23	
[0336]	<212>	DNA	
[0337]	<213>	Homo sapiens(智人)	
[0338]	<400>	55	
[0339]	agagacctac atggtacagc tgg		23
[0340]	<210>	56	
[0341]	<211>	23	
[0342]	<212>	DNA	
[0343]	<213>	Homo sapiens(智人)	
[0344]	<400>	56	
[0345]	ctgggccagc tgtaccatgt agg		23
[0346]	<210>	57	
[0347]	<211>	23	
[0348]	<212>	DNA	
[0349]	<213>	Homo sapiens(智人)	
[0350]	<400>	57	

[0351]	agggaaaggg cttacggtct ggg	23
[0352]	<210> 58	
[0353]	<211> 23	
[0354]	<212> DNA	
[0355]	<213> Homo sapiens(智人)	
[0356]	<400> 58	
[0357]	cagggaaagg gcttacggtc tgg	23
[0358]	<210> 59	
[0359]	<211> 23	
[0360]	<212> DNA	
[0361]	<213> Homo sapiens(智人)	
[0362]	<400> 59	
[0363]	tctggagatc ttcttgcaac agg	23
[0364]	<210> 60	
[0365]	<211> 23	
[0366]	<212> DNA	
[0367]	<213> Homo sapiens(智人)	
[0368]	<400> 60	
[0369]	ctccggttca tgactttgaa agg	23
[0370]	<210> 61	
[0371]	<211> 23	
[0372]	<212> DNA	
[0373]	<213> Homo sapiens(智人)	
[0374]	<400> 61	
[0375]	gtcttccatc ttcgtctttc agg	23
[0376]	<210> 62	
[0377]	<211> 23	
[0378]	<212> DNA	
[0379]	<213> Homo sapiens(智人)	
[0380]	<400> 62	
[0381]	gaagacttcg agaccattt agg	23
[0382]	<210> 63	
[0383]	<211> 23	
[0384]	<212> DNA	
[0385]	<213> Homo sapiens(智人)	
[0386]	<400> 63	
[0387]	tcgagacca tttaggatca cgg	23
[0388]	<210> 64	
[0389]	<211> 23	

[0390]	<212>	DNA	
[0391]	<213>	Homo sapiens(智人)	
[0392]	<400>	64	
[0393]	gtagcgccgt gtcctaaat ggg		23
[0394]	<210>	65	
[0395]	<211>	23	
[0396]	<212>	DNA	
[0397]	<213>	Homo sapiens(智人)	
[0398]	<400>	65	
[0399]	cgtagcgccg tgatcctaaa tgg		23
[0400]	<210>	66	
[0401]	<211>	23	
[0402]	<212>	DNA	
[0403]	<213>	Homo sapiens(智人)	
[0404]	<400>	66	
[0405]	catttaggat cacggcgcta cgg		23
[0406]	<210>	67	
[0407]	<211>	23	
[0408]	<212>	DNA	
[0409]	<213>	Homo sapiens(智人)	
[0410]	<400>	67	
[0411]	ggtcccaata ttgaagccca tgg		23
[0412]	<210>	68	
[0413]	<211>	23	
[0414]	<212>	DNA	
[0415]	<213>	Homo sapiens(智人)	
[0416]	<400>	68	
[0417]	gatccatggg cttcaatatt ggg		23
[0418]	<210>	69	
[0419]	<211>	23	
[0420]	<212>	DNA	
[0421]	<213>	Homo sapiens(智人)	
[0422]	<400>	69	
[0423]	agatccatgg gcttcaatat tgg		23
[0424]	<210>	70	
[0425]	<211>	23	
[0426]	<212>	DNA	
[0427]	<213>	Homo sapiens(智人)	
[0428]	<400>	70	

[0429]	gcttctacca taagatccat ggg	23
[0430]	<210> 71	
[0431]	<211> 23	
[0432]	<212> DNA	
[0433]	<213> Homo sapiens(智人)	
[0434]	<400> 71	
[0435]	cgcttctacc ataagatcca tgg	23
[0436]	<210> 72	
[0437]	<211> 23	
[0438]	<212> DNA	
[0439]	<213> Homo sapiens(智人)	
[0440]	<400> 72	
[0441]	gcatttgcaa aaattcgccg tgg	23
[0442]	<210> 73	
[0443]	<211> 23	
[0444]	<212> DNA	
[0445]	<213> Homo sapiens(智人)	
[0446]	<400> 73	
[0447]	atgacctgag aattaattta tgg	23
[0448]	<210> 74	
[0449]	<211> 23	
[0450]	<212> DNA	
[0451]	<213> Homo sapiens(智人)	
[0452]	<400> 74	
[0453]	ccatgcactc ccagacatcg tgg	23
[0454]	<210> 75	
[0455]	<211> 23	
[0456]	<212> DNA	
[0457]	<213> Homo sapiens(智人)	
[0458]	<400> 75	
[0459]	gcactggtgc gggtggaact cgg	23
[0460]	<210> 76	
[0461]	<211> 23	
[0462]	<212> DNA	
[0463]	<213> Homo sapiens(智人)	
[0464]	<400> 76	
[0465]	acacgttgca ctggtgcggg tgg	23
[0466]	<210> 77	
[0467]	<211> 23	

[0468]	<212>	DNA	
[0469]	<213>	Homo sapiens(智人)	
[0470]	<400>	77	
[0471]	cgaacacggtt gcactggtgc ggg		23
[0472]	<210>	78	
[0473]	<211>	23	
[0474]	<212>	DNA	
[0475]	<213>	Homo sapiens(智人)	
[0476]	<400>	78	
[0477]	acgaacacgt tgcactggtg cgg		23
[0478]	<210>	79	
[0479]	<211>	23	
[0480]	<212>	DNA	
[0481]	<213>	Homo sapiens(智人)	
[0482]	<400>	79	
[0483]	tgtagacgaa cacgttgac tgg		23
[0484]	<210>	80	
[0485]	<211>	23	
[0486]	<212>	DNA	
[0487]	<213>	Homo sapiens(智人)	
[0488]	<400>	80	
[0489]	gcgcatgtca cacaggcgga tgg		23
[0490]	<210>	81	
[0491]	<211>	23	
[0492]	<212>	DNA	
[0493]	<213>	Homo sapiens(智人)	
[0494]	<400>	81	
[0495]	aggagcgcat gtcacacagg cgg		23
[0496]	<210>	82	
[0497]	<211>	23	
[0498]	<212>	DNA	
[0499]	<213>	Homo sapiens(智人)	
[0500]	<400>	82	
[0501]	ccgaggagcg catgtcacac agg		23
[0502]	<210>	83	
[0503]	<211>	23	
[0504]	<212>	DNA	
[0505]	<213>	Homo sapiens(智人)	
[0506]	<400>	83	

[0507]	cctgtgtgac atgcgctcct cgg	23
[0508]	<210> 84	
[0509]	<211> 23	
[0510]	<212> DNA	
[0511]	<213> Homo sapiens(智人)	
[0512]	<400> 84	
[0513]	cgactggcca gggcgcctgt ggg	23
[0514]	<210> 85	
[0515]	<211> 23	
[0516]	<212> DNA	
[0517]	<213> Homo sapiens(智人)	
[0518]	<400> 85	
[0519]	accgcccaga cgactggcca ggg	23
[0520]	<210> 86	
[0521]	<211> 23	
[0522]	<212> DNA	
[0523]	<213> Homo sapiens(智人)	
[0524]	<400> 86	
[0525]	caccgcccag acgactggcc agg	23
[0526]	<210> 87	
[0527]	<211> 23	
[0528]	<212> DNA	
[0529]	<213> Homo sapiens(智人)	
[0530]	<400> 87	
[0531]	gtctgggcgg tgctacaact ggg	23
[0532]	<210> 88	
[0533]	<211> 23	
[0534]	<212> DNA	
[0535]	<213> Homo sapiens(智人)	
[0536]	<400> 88	
[0537]	ctacaactgg gctggcggcc agg	23
[0538]	<210> 89	
[0539]	<211> 23	
[0540]	<212> DNA	
[0541]	<213> Homo sapiens(智人)	
[0542]	<400> 89	
[0543]	cacctaccta agaaccatcc tgg	23
[0544]	<210> 90	
[0545]	<211> 23	

[0546]	<212>	DNA	
[0547]	<213>	Homo sapiens(智人)	
[0548]	<400>	90	
[0549]	cggtcaccac gagcagggct ggg		23
[0550]	<210>	91	
[0551]	<211>	23	
[0552]	<212>	DNA	
[0553]	<213>	Homo sapiens(智人)	
[0554]	<400>	91	
[0555]	gccctgctcg tggtagaccga agg		23
[0556]	<210>	92	
[0557]	<211>	23	
[0558]	<212>	DNA	
[0559]	<213>	Homo sapiens(智人)	
[0560]	<400>	92	
[0561]	cggagagctt cgtgctaaac tgg		23
[0562]	<210>	93	
[0563]	<211>	23	
[0564]	<212>	DNA	
[0565]	<213>	Homo sapiens(智人)	
[0566]	<400>	93	
[0567]	cagcttggtcc gtctggttgc tgg		23
[0568]	<210>	94	
[0569]	<211>	23	
[0570]	<212>	DNA	
[0571]	<213>	Homo sapiens(智人)	
[0572]	<400>	94	
[0573]	aggcggccag cttgtccgtc tgg		23
[0574]	<210>	95	
[0575]	<211>	23	
[0576]	<212>	DNA	
[0577]	<213>	Homo sapiens(智人)	
[0578]	<400>	95	
[0579]	ccgggctggc tgcggtcctc ggg		23
[0580]	<210>	96	
[0581]	<211>	23	
[0582]	<212>	DNA	
[0583]	<213>	Homo sapiens(智人)	
[0584]	<400>	96	

[0585]	cgttgggcag ttgtgtgaca cgg	23
[0586]	<210> 97	
[0587]	<211> 23	
[0588]	<212> DNA	
[0589]	<213> Homo sapiens(智人)	
[0590]	<400> 97	
[0591]	cataaagcca tggcttgcct tgg	23
[0592]	<210> 98	
[0593]	<211> 23	
[0594]	<212> DNA	
[0595]	<213> Homo sapiens(智人)	
[0596]	<400> 98	
[0597]	ccttgattt cageggcaca agg	23
[0598]	<210> 99	
[0599]	<211> 23	
[0600]	<212> DNA	
[0601]	<213> Homo sapiens(智人)	
[0602]	<400> 99	
[0603]	ccttgtgccg ctgaaatcca agg	23
[0604]	<210> 100	
[0605]	<211> 23	
[0606]	<212> DNA	
[0607]	<213> Homo sapiens(智人)	
[0608]	<400> 100	
[0609]	cactcacctt tgcagaagac agg	23
[0610]	<210> 101	
[0611]	<211> 23	
[0612]	<212> DNA	
[0613]	<213> Homo sapiens(智人)	
[0614]	<400> 101	
[0615]	ttccatgcta gcaatgcacg tgg	23
[0616]	<210> 102	
[0617]	<211> 23	
[0618]	<212> DNA	
[0619]	<213> Homo sapiens(智人)	
[0620]	<400> 102	
[0621]	ggccacgtgc attgctagca tgg	23
[0622]	<210> 103	
[0623]	<211> 23	

[0624]	<212>	DNA	
[0625]	<213>	Homo sapiens(智人)	
[0626]	<400>	103	
[0627]		ggcccagcct gctgtggtac tgg	23
[0628]	<210>	104	
[0629]	<211>	23	
[0630]	<212>	DNA	
[0631]	<213>	Homo sapiens(智人)	
[0632]	<400>	104	
[0633]		aggtccgggt gacagtgctt cgg	23
[0634]	<210>	105	
[0635]	<211>	23	
[0636]	<212>	DNA	
[0637]	<213>	Homo sapiens(智人)	
[0638]	<400>	105	
[0639]		ccgggtgaca gtgcttcggc agg	23
[0640]	<210>	106	
[0641]	<211>	23	
[0642]	<212>	DNA	
[0643]	<213>	Homo sapiens(智人)	
[0644]	<400>	106	
[0645]		ctgtgcggca acctacatga tgg	23
[0646]	<210>	107	
[0647]	<211>	23	
[0648]	<212>	DNA	
[0649]	<213>	Homo sapiens(智人)	
[0650]	<400>	107	
[0651]		caactcattc cccatcatgt agg	23
[0652]	<210>	108	
[0653]	<211>	23	
[0654]	<212>	DNA	
[0655]	<213>	Homo sapiens(智人)	
[0656]	<400>	108	
[0657]		ctagatgatt ccatctgcac ggg	23
[0658]	<210>	109	
[0659]	<211>	23	
[0660]	<212>	DNA	
[0661]	<213>	Homo sapiens(智人)	
[0662]	<400>	109	

[0663]	ggctaggagt cagcgacata tgg	23
[0664]	<210> 110	
[0665]	<211> 23	
[0666]	<212> DNA	
[0667]	<213> Homo sapiens(智人)	
[0668]	<400> 110	
[0669]	gctaggagtc agcgacatat ggg	23
[0670]	<210> 111	
[0671]	<211> 23	
[0672]	<212> DNA	
[0673]	<213> Homo sapiens(智人)	
[0674]	<400> 111	
[0675]	ctaggagtca gcgacatatg ggg	23
[0676]	<210> 112	
[0677]	<211> 23	
[0678]	<212> DNA	
[0679]	<213> Homo sapiens(智人)	
[0680]	<400> 112	
[0681]	gtactgtgta gccaggatgc tgg	23
[0682]	<210> 113	
[0683]	<211> 23	
[0684]	<212> DNA	
[0685]	<213> Homo sapiens(智人)	
[0686]	<400> 113	
[0687]	acgagcactc accagcatcc tgg	23
[0688]	<210> 114	
[0689]	<211> 23	
[0690]	<212> DNA	
[0691]	<213> Homo sapiens(智人)	
[0692]	<400> 114	
[0693]	aggctccagg aatgtccgcg agg	23
[0694]	<210> 115	
[0695]	<211> 23	
[0696]	<212> DNA	
[0697]	<213> Homo sapiens(智人)	
[0698]	<400> 115	
[0699]	acttacctcg cggacattcc tgg	23
[0700]	<210> 116	
[0701]	<211> 23	

[0702]	<212>	DNA	
[0703]	<213>	Homo sapiens(智人)	
[0704]	<400>	116	
[0705]	caccctgggc acttacctcg cgg		23
[0706]	<210>	117	
[0707]	<211>	23	
[0708]	<212>	DNA	
[0709]	<213>	Homo sapiens(智人)	
[0710]	<400>	117	
[0711]	gtgccgtaca aaggttggt ggg		23
[0712]	<210>	118	
[0713]	<211>	23	
[0714]	<212>	DNA	
[0715]	<213>	Homo sapiens(智人)	
[0716]	<400>	118	
[0717]	ggtgccgtac aaaggttggt tgg		23
[0718]	<210>	119	
[0719]	<211>	23	
[0720]	<212>	DNA	
[0721]	<213>	Homo sapiens(智人)	
[0722]	<400>	119	
[0723]	ctctcctcag taccacagca agg		23
[0724]	<210>	120	
[0725]	<211>	23	
[0726]	<212>	DNA	
[0727]	<213>	Homo sapiens(智人)	
[0728]	<400>	120	
[0729]	cctggggcct ccgggcgcgg agg		23
[0730]	<210>	121	
[0731]	<211>	23	
[0732]	<212>	DNA	
[0733]	<213>	Homo sapiens(智人)	
[0734]	<400>	121	
[0735]	agtactcacc tggggcctcc ggg		23
[0736]	<210>	122	
[0737]	<211>	23	
[0738]	<212>	DNA	
[0739]	<213>	Homo sapiens(智人)	
[0740]	<400>	122	

[0741]	agggtctcca gcggccctcc tgg	23
[0742]	<210> 123	
[0743]	<211> 23	
[0744]	<212> DNA	
[0745]	<213> Homo sapiens(智人)	
[0746]	<400> 123	
[0747]	gcaagtactt acgcctcctt ggg	23
[0748]	<210> 124	
[0749]	<211> 23	
[0750]	<212> DNA	
[0751]	<213> Homo sapiens(智人)	
[0752]	<400> 124	
[0753]	ttgcggtaca tctccagcct ggg	23
[0754]	<210> 125	
[0755]	<211> 23	
[0756]	<212> DNA	
[0757]	<213> Homo sapiens(智人)	
[0758]	<400> 125	
[0759]	tttgcggtac atctccagcc tgg	23
[0760]	<210> 126	
[0761]	<211> 23	
[0762]	<212> DNA	
[0763]	<213> Homo sapiens(智人)	
[0764]	<400> 126	
[0765]	gcaaaacctg tccactctta tgg	23
[0766]	<210> 127	
[0767]	<211> 23	
[0768]	<212> DNA	
[0769]	<213> Homo sapiens(智人)	
[0770]	<400> 127	
[0771]	ttggtgcat aagagtggac agg	23
[0772]	<210> 128	
[0773]	<211> 23	
[0774]	<212> DNA	
[0775]	<213> Homo sapiens(智人)	
[0776]	<400> 128	
[0777]	ggtgcaagtt tcttatatgt tgg	23
[0778]	<210> 129	
[0779]	<211> 23	

[0780]	<212>	DNA	
[0781]	<213>	Homo sapiens(智人)	
[0782]	<400>	129	
[0783]	acctgatgca tataataatc agg		23
[0784]	<210>	130	
[0785]	<211>	23	
[0786]	<212>	DNA	
[0787]	<213>	Homo sapiens(智人)	
[0788]	<400>	130	
[0789]	acctgattat tatatgcatc agg		23
[0790]	<210>	131	
[0791]	<211>	23	
[0792]	<212>	DNA	
[0793]	<213>	Homo sapiens(智人)	
[0794]	<400>	131	
[0795]	cagagcacca gagtgccgtc tgg		23
[0796]	<210>	132	
[0797]	<211>	23	
[0798]	<212>	DNA	
[0799]	<213>	Homo sapiens(智人)	
[0800]	<400>	132	
[0801]	agagcaccag agtgccgtct ggg		23
[0802]	<210>	133	
[0803]	<211>	23	
[0804]	<212>	DNA	
[0805]	<213>	Homo sapiens(智人)	
[0806]	<400>	133	
[0807]	agagtgccgt ctgggtctga agg		23
[0808]	<210>	134	
[0809]	<211>	23	
[0810]	<212>	DNA	
[0811]	<213>	Homo sapiens(智人)	
[0812]	<400>	134	
[0813]	aggaaggccg tccattctca ggg		23
[0814]	<210>	135	
[0815]	<211>	23	
[0816]	<212>	DNA	
[0817]	<213>	Homo sapiens(智人)	
[0818]	<400>	135	

[0819]	ggatagaacc aaccatgttg agg	23
[0820]	<210> 136	
[0821]	<211> 23	
[0822]	<212> DNA	
[0823]	<213> Homo sapiens(智人)	
[0824]	<400> 136	
[0825]	tctgttgccc tcaacatggg tgg	23
[0826]	<210> 137	
[0827]	<211> 23	
[0828]	<212> DNA	
[0829]	<213> Homo sapiens(智人)	
[0830]	<400> 137	
[0831]	ttagtctgtt gccctcaaca tgg	23
[0832]	<210> 138	
[0833]	<211> 23	
[0834]	<212> DNA	
[0835]	<213> Homo sapiens(智人)	
[0836]	<400> 138	
[0837]	gtctggcaaa tgggaggtga tgg	23
[0838]	<210> 139	
[0839]	<211> 23	
[0840]	<212> DNA	
[0841]	<213> Homo sapiens(智人)	
[0842]	<400> 139	
[0843]	cagaggttct gtctggcaaa tgg	23
[0844]	<210> 140	
[0845]	<211> 23	
[0846]	<212> DNA	
[0847]	<213> Homo sapiens(智人)	
[0848]	<400> 140	
[0849]	ttgtagccag aggttctgtc tgg	23
[0850]	<210> 141	
[0851]	<211> 23	
[0852]	<212> DNA	
[0853]	<213> Homo sapiens(智人)	
[0854]	<400> 141	
[0855]	acttctggat gagctctctc agg	23
[0856]	<210> 142	
[0857]	<211> 23	

[0858]	<212>	DNA	
[0859]	<213>	Homo sapiens(智人)	
[0860]	<400>	142	
[0861]	agagctcatc cagaagtaaa tgg		23
[0862]	<210>	143	
[0863]	<211>	23	
[0864]	<212>	DNA	
[0865]	<213>	Homo sapiens(智人)	
[0866]	<400>	143	
[0867]	ttggtgtctc catttacttc tgg		23
[0868]	<210>	144	
[0869]	<211>	23	
[0870]	<212>	DNA	
[0871]	<213>	Homo sapiens(智人)	
[0872]	<400>	144	
[0873]	ttctggcttc ccttcataca ggg		23
[0874]	<210>	145	
[0875]	<211>	23	
[0876]	<212>	DNA	
[0877]	<213>	Homo sapiens(智人)	
[0878]	<400>	145	
[0879]	caggactcac acgactattc tgg		23
[0880]	<210>	146	
[0881]	<211>	23	
[0882]	<212>	DNA	
[0883]	<213>	Homo sapiens(智人)	
[0884]	<400>	146	
[0885]	ctactttctt gtgtaaagtc agg		23
[0886]	<210>	147	
[0887]	<211>	23	
[0888]	<212>	DNA	
[0889]	<213>	Homo sapiens(智人)	
[0890]	<400>	147	
[0891]	gactttacac aagaaagtag agg		23
[0892]	<210>	148	
[0893]	<211>	23	
[0894]	<212>	DNA	
[0895]	<213>	Homo sapiens(智人)	
[0896]	<400>	148	

[0897]	gtctttctcc attagccttt tgg	23
[0898]	<210> 149	
[0899]	<211> 23	
[0900]	<212> DNA	
[0901]	<213> Homo sapiens(智人)	
[0902]	<400> 149	
[0903]	aatggagaaa gacgtaactt cgg	23
[0904]	<210> 150	
[0905]	<211> 23	
[0906]	<212> DNA	
[0907]	<213> Homo sapiens(智人)	
[0908]	<400> 150	
[0909]	atggagaaag acgtaacttc ggg	23
[0910]	<210> 151	
[0911]	<211> 23	
[0912]	<212> DNA	
[0913]	<213> Homo sapiens(智人)	
[0914]	<400> 151	
[0915]	tggagaaaga cgtaacttcg ggg	23
[0916]	<210> 152	
[0917]	<211> 23	
[0918]	<212> DNA	
[0919]	<213> Homo sapiens(智人)	
[0920]	<400> 152	
[0921]	tttggttgac tgctttcacc tgg	23
[0922]	<210> 153	
[0923]	<211> 23	
[0924]	<212> DNA	
[0925]	<213> Homo sapiens(智人)	
[0926]	<400> 153	
[0927]	tcactcaaatt cggagacatt tgg	23
[0928]	<210> 154	
[0929]	<211> 23	
[0930]	<212> DNA	
[0931]	<213> Homo sapiens(智人)	
[0932]	<400> 154	
[0933]	atctgaagct ctggattttc agg	23
[0934]	<210> 155	
[0935]	<211> 23	

[0936]	<212>	DNA	
[0937]	<213>	Homo sapiens(智人)	
[0938]	<400>	155	
[0939]		gcttcagatt ctgaatgagc agg	23
[0940]	<210>	156	
[0941]	<211>	23	
[0942]	<212>	DNA	
[0943]	<213>	Homo sapiens(智人)	
[0944]	<400>	156	
[0945]		cagattctga atgagcagga ggg	23
[0946]	<210>	157	
[0947]	<211>	23	
[0948]	<212>	DNA	
[0949]	<213>	Homo sapiens(智人)	
[0950]	<400>	157	
[0951]		aaggcagtgc taatgcctaa tgg	23
[0952]	<210>	158	
[0953]	<211>	23	
[0954]	<212>	DNA	
[0955]	<213>	Homo sapiens(智人)	
[0956]	<400>	158	
[0957]		gcagaaactg tagcaccatt agg	23
[0958]	<210>	159	
[0959]	<211>	23	
[0960]	<212>	DNA	
[0961]	<213>	Homo sapiens(智人)	
[0962]	<400>	159	
[0963]		accgcaatgg aaacacaatc tgg	23
[0964]	<210>	160	
[0965]	<211>	23	
[0966]	<212>	DNA	
[0967]	<213>	Homo sapiens(智人)	
[0968]	<400>	160	
[0969]		tgtgggttttc tgcaccgcaa tgg	23
[0970]	<210>	161	
[0971]	<211>	23	
[0972]	<212>	DNA	
[0973]	<213>	Homo sapiens(智人)	
[0974]	<400>	161	

[0975]	cataaatgcc attaacagtc agg	23
[0976]	<210> 162	
[0977]	<211> 23	
[0978]	<212> DNA	
[0979]	<213> Homo sapiens(智人)	
[0980]	<400> 162	
[0981]	attagtagcc tgactgttaa tgg	23
[0982]	<210> 163	
[0983]	<211> 23	
[0984]	<212> DNA	
[0985]	<213> Homo sapiens(智人)	
[0986]	<400> 163	
[0987]	cgatgggtga gtgatctcac agg	23
[0988]	<210> 164	
[0989]	<211> 23	
[0990]	<212> DNA	
[0991]	<213> Homo sapiens(智人)	
[0992]	<400> 164	
[0993]	actcacccat cgcatacctc agg	23
[0994]	<210> 165	
[0995]	<211> 23	
[0996]	<212> DNA	
[0997]	<213> Homo sapiens(智人)	
[0998]	<400> 165	
[0999]	ctcacccatc gcatacctca ggg	23
[1000]	<210> 166	
[1001]	<211> 23	
[1002]	<212> DNA	
[1003]	<213> Homo sapiens(智人)	
[1004]	<400> 166	
[1005]	agcaacagga ggagttgcag agg	23
[1006]	<210> 167	
[1007]	<211> 23	
[1008]	<212> DNA	
[1009]	<213> Homo sapiens(智人)	
[1010]	<400> 167	
[1011]	ccagtaggat cagcaacagg agg	23
[1012]	<210> 168	
[1013]	<211> 23	

[1014]	<212>	DNA	
[1015]	<213>	Homo sapiens(智人)	
[1016]	<400>	168	
[1017]	ctcctgttgc tgatcctact ggg		23
[1018]	<210>	169	
[1019]	<211>	23	
[1020]	<212>	DNA	
[1021]	<213>	Homo sapiens(智人)	
[1022]	<400>	169	
[1023]	ggcccagtag gatcagcaac agg		23
[1024]	<210>	170	
[1025]	<211>	23	
[1026]	<212>	DNA	
[1027]	<213>	Homo sapiens(智人)	
[1028]	<400>	170	
[1029]	ttgctgatcc tactgggccc tgg		23
[1030]	<210>	171	
[1031]	<211>	23	
[1032]	<212>	DNA	
[1033]	<213>	Homo sapiens(智人)	
[1034]	<400>	171	
[1035]	tggcaacagc ttgcagctgt ggg		23
[1036]	<210>	172	
[1037]	<211>	23	
[1038]	<212>	DNA	
[1039]	<213>	Homo sapiens(智人)	
[1040]	<400>	172	
[1041]	cttgggtccc ctgcttgccc ggg		23
[1042]	<210>	173	
[1043]	<211>	23	
[1044]	<212>	DNA	
[1045]	<213>	Homo sapiens(智人)	
[1046]	<400>	173	
[1047]	gtccccctgct tgcccgggac cgg		23
[1048]	<210>	174	
[1049]	<211>	23	
[1050]	<212>	DNA	
[1051]	<213>	Homo sapiens(智人)	
[1052]	<400>	174	

[1053]	ctccggtccc gggcaagcag ggg	23
[1054]	<210> 175	
[1055]	<211> 23	
[1056]	<212> DNA	
[1057]	<213> Homo sapiens(智人)	
[1058]	<400> 175	
[1059]	tctccggtcc cgggcaagca ggg	23
[1060]	<210> 176	
[1061]	<211> 23	
[1062]	<212> DNA	
[1063]	<213> Homo sapiens(智人)	
[1064]	<400> 176	
[1065]	gtctccggtc cgggcaagc agg	23
[1066]	<210> 177	
[1067]	<211> 23	
[1068]	<212> DNA	
[1069]	<213> Homo sapiens(智人)	
[1070]	<400> 177	
[1071]	gcttgcccgg gaccggagac agg	23
[1072]	<210> 178	
[1073]	<211> 23	
[1074]	<212> DNA	
[1075]	<213> Homo sapiens(智人)	
[1076]	<400> 178	
[1077]	ggtggcctgt ctccggtccc ggg	23
[1078]	<210> 179	
[1079]	<211> 23	
[1080]	<212> DNA	
[1081]	<213> Homo sapiens(智人)	
[1082]	<400> 179	
[1083]	cgggtggcctg tctccggtcc cgg	23
[1084]	<210> 180	
[1085]	<211> 23	
[1086]	<212> DNA	
[1087]	<213> Homo sapiens(智人)	
[1088]	<400> 180	
[1089]	catattcggt ggcctgtctc cgg	23
[1090]	<210> 181	
[1091]	<211> 23	

[1092]	<212>	DNA	
[1093]	<213>	Homo sapiens(智人)	
[1094]	<400>	181	
[1095]		atctaggtac tcatattcgg tgg	23
[1096]	<210>	182	
[1097]	<211>	23	
[1098]	<212>	DNA	
[1099]	<213>	Homo sapiens(智人)	
[1100]	<400>	182	
[1101]		ataatctagg tactcatatt cgg	23
[1102]	<210>	183	
[1103]	<211>	23	
[1104]	<212>	DNA	
[1105]	<213>	Homo sapiens(智人)	
[1106]	<400>	183	
[1107]		ttatgatttc ctgccagaaa cgg	23
[1108]	<210>	184	
[1109]	<211>	23	
[1110]	<212>	DNA	
[1111]	<213>	Homo sapiens(智人)	
[1112]	<400>	184	
[1113]		atttctggag gctccgtttc tgg	23
[1114]	<210>	185	
[1115]	<211>	23	
[1116]	<212>	DNA	
[1117]	<213>	Homo sapiens(智人)	
[1118]	<400>	185	
[1119]		actgacacca ctctctgac tgg	23
[1120]	<210>	186	
[1121]	<211>	23	
[1122]	<212>	DNA	
[1123]	<213>	Homo sapiens(智人)	
[1124]	<400>	186	
[1125]		ctgacaccac tcctctgact ggg	23
[1126]	<210>	187	
[1127]	<211>	23	
[1128]	<212>	DNA	
[1129]	<213>	Homo sapiens(智人)	
[1130]	<400>	187	

[1131]	accactcctc tgactgggcc tgg	23
[1132]	<210> 188	
[1133]	<211> 23	
[1134]	<212> DNA	
[1135]	<213> Homo sapiens(智人)	
[1136]	<400> 188	
[1137]	aacccctgag tctaccactg tgg	23
[1138]	<210> 189	
[1139]	<211> 23	
[1140]	<212> DNA	
[1141]	<213> Homo sapiens(智人)	
[1142]	<400> 189	
[1143]	ctccacagt gtagactcag ggg	23
[1144]	<210> 190	
[1145]	<211> 23	
[1146]	<212> DNA	
[1147]	<213> Homo sapiens(智人)	
[1148]	<400> 190	
[1149]	gctccacagt gtagactca ggg	23
[1150]	<210> 191	
[1151]	<211> 23	
[1152]	<212> DNA	
[1153]	<213> Homo sapiens(智人)	
[1154]	<400> 191	
[1155]	ggctccacag tggtagactc agg	23
[1156]	<210> 192	
[1157]	<211> 23	
[1158]	<212> DNA	
[1159]	<213> Homo sapiens(智人)	
[1160]	<400> 192	
[1161]	cctgctgcaa ggcgttctac tgg	23
[1162]	<210> 193	
[1163]	<211> 23	
[1164]	<212> DNA	
[1165]	<213> Homo sapiens(智人)	
[1166]	<400> 193	
[1167]	ccagtagaac gccttgccagg agg	23
[1168]	<210> 194	
[1169]	<211> 23	

[1170]	<212>	DNA	
[1171]	<213>	Homo sapiens(智人)	
[1172]	<400>	194	
[1173]	cgttctactg gcctggatgc agg		23
[1174]	<210>	195	
[1175]	<211>	23	
[1176]	<212>	DNA	
[1177]	<213>	Homo sapiens(智人)	
[1178]	<400>	195	
[1179]	tctactggcc tggatgcagg agg		23
[1180]	<210>	196	
[1181]	<211>	23	
[1182]	<212>	DNA	
[1183]	<213>	Homo sapiens(智人)	
[1184]	<400>	196	
[1185]	ccacggagct ggccaacatg ggg		23
[1186]	<210>	197	
[1187]	<211>	23	
[1188]	<212>	DNA	
[1189]	<213>	Homo sapiens(智人)	
[1190]	<400>	197	
[1191]	cgtggacagg ttcccatgt tgg		23
[1192]	<210>	198	
[1193]	<211>	23	
[1194]	<212>	DNA	
[1195]	<213>	Homo sapiens(智人)	
[1196]	<400>	198	
[1197]	gtccacggat tcagcagcta tgg		23
[1198]	<210>	199	
[1199]	<211>	23	
[1200]	<212>	DNA	
[1201]	<213>	Homo sapiens(智人)	
[1202]	<400>	199	
[1203]	gaccactcaa ccagtgccca cgg		23
[1204]	<210>	200	
[1205]	<211>	23	
[1206]	<212>	DNA	
[1207]	<213>	Homo sapiens(智人)	
[1208]	<400>	200	

[1209]	ggagtgggtct gtgcctccgt ggg	23
[1210]	<210> 201	
[1211]	<211> 23	
[1212]	<212> DNA	
[1213]	<213> Homo sapiens(智人)	
[1214]	<400> 201	
[1215]	ggcacagaca actcgactga cgg	23
[1216]	<210> 202	
[1217]	<211> 23	
[1218]	<212> DNA	
[1219]	<213> Homo sapiens(智人)	
[1220]	<400> 202	
[1221]	gacaactcga ctgacggcca cgg	23
[1222]	<210> 203	
[1223]	<211> 23	
[1224]	<212> DNA	
[1225]	<213> Homo sapiens(智人)	
[1226]	<400> 203	
[1227]	aactcgactg acggccacgg agg	23
[1228]	<210> 204	
[1229]	<211> 23	
[1230]	<212> DNA	
[1231]	<213> Homo sapiens(智人)	
[1232]	<400> 204	
[1233]	cacagaaccc agtgccacag agg	23
[1234]	<210> 205	
[1235]	<211> 23	
[1236]	<212> DNA	
[1237]	<213> Homo sapiens(智人)	
[1238]	<400> 205	
[1239]	ggtagtaggt tccatggaca ggg	23
[1240]	<210> 206	
[1241]	<211> 23	
[1242]	<212> DNA	
[1243]	<213> Homo sapiens(智人)	
[1244]	<400> 206	
[1245]	tggtagtagg ttccatggac agg	23
[1246]	<210> 207	
[1247]	<211> 23	

[1248]	<212>	DNA	
[1249]	<213>	Homo sapiens(智人)	
[1250]	<400>	207	
[1251]	tcttttggta gtaggttcca tgg		23
[1252]	<210>	208	
[1253]	<211>	23	
[1254]	<212>	DNA	
[1255]	<213>	Homo sapiens(智人)	
[1256]	<400>	208	
[1257]	atggaaccta ctacaaaag agg		23
[1258]	<210>	209	
[1259]	<211>	23	
[1260]	<212>	DNA	
[1261]	<213>	Homo sapiens(智人)	
[1262]	<400>	209	
[1263]	aacagacctc ttttggtagt agg		23
[1264]	<210>	210	
[1265]	<211>	23	
[1266]	<212>	DNA	
[1267]	<213>	Homo sapiens(智人)	
[1268]	<400>	210	
[1269]	gggtatgaac agacctcttt tgg		23
[1270]	<210>	211	
[1271]	<211>	23	
[1272]	<212>	DNA	
[1273]	<213>	Homo sapiens(智人)	
[1274]	<400>	211	
[1275]	tgtgtcctct gttactcaca agg		23
[1276]	<210>	212	
[1277]	<211>	23	
[1278]	<212>	DNA	
[1279]	<213>	Homo sapiens(智人)	
[1280]	<400>	212	
[1281]	gtgtcctctg ttactcaca ggg		23
[1282]	<210>	213	
[1283]	<211>	23	
[1284]	<212>	DNA	
[1285]	<213>	Homo sapiens(智人)	
[1286]	<400>	213	

[1287]	gtagttgacg gacaaattgc tgg	23
[1288]	<210> 214	
[1289]	<211> 23	
[1290]	<212> DNA	
[1291]	<213> Homo sapiens(智人)	
[1292]	<400> 214	
[1293]	tttgtccgtc aactacccag tgg	23
[1294]	<210> 215	
[1295]	<211> 23	
[1296]	<212> DNA	
[1297]	<213> Homo sapiens(智人)	
[1298]	<400> 215	
[1299]	ttgtccgtca actacccagt ggg	23
[1300]	<210> 216	
[1301]	<211> 23	
[1302]	<212> DNA	
[1303]	<213> Homo sapiens(智人)	
[1304]	<400> 216	
[1305]	tgtccgtcaa ctacccagt ggg	23
[1306]	<210> 217	
[1307]	<211> 23	
[1308]	<212> DNA	
[1309]	<213> Homo sapiens(智人)	
[1310]	<400> 217	
[1311]	gtccgtcaac taccagtg ggg	23
[1312]	<210> 218	
[1313]	<211> 23	
[1314]	<212> DNA	
[1315]	<213> Homo sapiens(智人)	
[1316]	<400> 218	
[1317]	ctctgtgaag cagtcctgc tgg	23
[1318]	<210> 219	
[1319]	<211> 23	
[1320]	<212> DNA	
[1321]	<213> Homo sapiens(智人)	
[1322]	<400> 219	
[1323]	cctgctggcc atcctaattc tgg	23
[1324]	<210> 220	
[1325]	<211> 23	

[1326]	<212>	DNA	
[1327]	<213>	Homo sapiens(智人)	
[1328]	<400>	220	
[1329]		ccaagattag gatggccagc agg	23
[1330]	<210>	221	
[1331]	<211>	23	
[1332]	<212>	DNA	
[1333]	<213>	Homo sapiens(智人)	
[1334]	<400>	221	
[1335]		ggccatccta atcttggcgc tgg	23
[1336]	<210>	222	
[1337]	<211>	23	
[1338]	<212>	DNA	
[1339]	<213>	Homo sapiens(智人)	
[1340]	<400>	222	
[1341]		caccagcgcc aagattagga tgg	23
[1342]	<210>	223	
[1343]	<211>	23	
[1344]	<212>	DNA	
[1345]	<213>	Homo sapiens(智人)	
[1346]	<400>	223	
[1347]		agtgcacacg aagaagatag tgg	23
[1348]	<210>	224	
[1349]	<211>	23	
[1350]	<212>	DNA	
[1351]	<213>	Homo sapiens(智人)	
[1352]	<400>	224	
[1353]		tatcttcttc gtgtgcactg tgg	23
[1354]	<210>	225	
[1355]	<211>	23	
[1356]	<212>	DNA	
[1357]	<213>	Homo sapiens(智人)	
[1358]	<400>	225	
[1359]		cttcgtgtgc actgtggtgc tgg	23
[1360]	<210>	226	
[1361]	<211>	23	
[1362]	<212>	DNA	
[1363]	<213>	Homo sapiens(智人)	
[1364]	<400>	226	

[1365]	ggcgggtccgc ctctcccgca agg	23
[1366]	<210> 227	
[1367]	<211> 23	
[1368]	<212> DNA	
[1369]	<213> Homo sapiens(智人)	
[1370]	<400> 227	
[1371]	gcgggtccgcc tctcccgcaa ggg	23
[1372]	<210> 228	
[1373]	<211> 23	
[1374]	<212> DNA	
[1375]	<213> Homo sapiens(智人)	
[1376]	<400> 228	
[1377]	aattacgcac ggggtacatg tgg	23
[1378]	<210> 229	
[1379]	<211> 23	
[1380]	<212> DNA	
[1381]	<213> Homo sapiens(智人)	
[1382]	<400> 229	
[1383]	tgggggagta attacgcacg ggg	23
[1384]	<210> 230	
[1385]	<211> 23	
[1386]	<212> DNA	
[1387]	<213> Homo sapiens(智人)	
[1388]	<400> 230	
[1389]	gtgggggagat aattacgcac ggg	23
[1390]	<210> 231	
[1391]	<211> 23	
[1392]	<212> DNA	
[1393]	<213> Homo sapiens(智人)	
[1394]	<400> 231	
[1395]	ggtgggggag taattacgca cgg	23
[1396]	<210> 232	
[1397]	<211> 23	
[1398]	<212> DNA	
[1399]	<213> Homo sapiens(智人)	
[1400]	<400> 232	
[1401]	taattactcc cccaccgaga tgg	23
[1402]	<210> 233	
[1403]	<211> 23	

[1404]	<212>	DNA	
[1405]	<213>	Homo sapiens(智人)	
[1406]	<400>	233	
[1407]	agatgcagac catctcgggtg ggg		23
[1408]	<210>	234	
[1409]	<211>	23	
[1410]	<212>	DNA	
[1411]	<213>	Homo sapiens(智人)	
[1412]	<400>	234	
[1413]	gagatgcaga ccatctcgggt ggg		23
[1414]	<210>	235	
[1415]	<211>	23	
[1416]	<212>	DNA	
[1417]	<213>	Homo sapiens(智人)	
[1418]	<400>	235	
[1419]	tgagatgcag accatctcgg tgg		23
[1420]	<210>	236	
[1421]	<211>	23	
[1422]	<212>	DNA	
[1423]	<213>	Homo sapiens(智人)	
[1424]	<400>	236	
[1425]	ggatgagatg cagaccatct cgg		23
[1426]	<210>	237	
[1427]	<211>	23	
[1428]	<212>	DNA	
[1429]	<213>	Homo sapiens(智人)	
[1430]	<400>	237	
[1431]	atctcatccc tgttcctga tgg		23
[1432]	<210>	238	
[1433]	<211>	23	
[1434]	<212>	DNA	
[1435]	<213>	Homo sapiens(智人)	
[1436]	<400>	238	
[1437]	tcattccctgt tgccatgatgg ggg		23
[1438]	<210>	239	
[1439]	<211>	23	
[1440]	<212>	DNA	
[1441]	<213>	Homo sapiens(智人)	
[1442]	<400>	239	

[1443]	ctcaccccca tcaggcaaca ggg	23
[1444]	<210> 240	
[1445]	<211> 23	
[1446]	<212> DNA	
[1447]	<213> Homo sapiens(智人)	
[1448]	<400> 240	
[1449]	gagggcccct caccgccatc agg	23
[1450]	<210> 241	
[1451]	<211> 23	
[1452]	<212> DNA	
[1453]	<213> Homo sapiens(智人)	
[1454]	<400> 241	
[1455]	ggggccctctg ccacagccaa tgg	23
[1456]	<210> 242	
[1457]	<211> 23	
[1458]	<212> DNA	
[1459]	<213> Homo sapiens(智人)	
[1460]	<400> 242	
[1461]	ccctctgccca cagccaatgg ggg	23
[1462]	<210> 243	
[1463]	<211> 23	
[1464]	<212> DNA	
[1465]	<213> Homo sapiens(智人)	
[1466]	<400> 243	
[1467]	ccccattgg ctgtggcaga ggg	23
[1468]	<210> 244	
[1469]	<211> 23	
[1470]	<212> DNA	
[1471]	<213> Homo sapiens(智人)	
[1472]	<400> 244	
[1473]	gccccattg gctgtggcag agg	23
[1474]	<210> 245	
[1475]	<211> 23	
[1476]	<212> DNA	
[1477]	<213> Homo sapiens(智人)	
[1478]	<400> 245	
[1479]	ggacaggccc ccattggctg tgg	23
[1480]	<210> 246	
[1481]	<211> 23	

[1482]	<212>	DNA	
[1483]	<213>	Homo sapiens(智人)	
[1484]	<400>	246	
[1485]	ccgggctctt ggccttggac agg		23
[1486]	<210>	247	
[1487]	<211>	23	
[1488]	<212>	DNA	
[1489]	<213>	Homo sapiens(智人)	
[1490]	<400>	247	
[1491]	ctgtccaagg ccaagagccc ggg		23
[1492]	<210>	248	
[1493]	<211>	23	
[1494]	<212>	DNA	
[1495]	<213>	Homo sapiens(智人)	
[1496]	<400>	248	
[1497]	tggcgtcagg cccgggctct tgg		23
[1498]	<210>	249	
[1499]	<211>	23	
[1500]	<212>	DNA	
[1501]	<213>	Homo sapiens(智人)	
[1502]	<400>	249	
[1503]	cgggcctgac gccagagccc agg		23
[1504]	<210>	250	
[1505]	<211>	23	
[1506]	<212>	DNA	
[1507]	<213>	Homo sapiens(智人)	
[1508]	<400>	250	
[1509]	caacaacat gctgggcatc tgg		23
[1510]	<210>	251	
[1511]	<211>	23	
[1512]	<212>	DNA	
[1513]	<213>	Homo sapiens(智人)	
[1514]	<400>	251	
[1515]	gagggtccag atgccagca tgg		23
[1516]	<210>	252	
[1517]	<211>	23	
[1518]	<212>	DNA	
[1519]	<213>	Homo sapiens(智人)	
[1520]	<400>	252	

[1521]	catctggacc ctcctacctc tgg	23
[1522]	<210> 253	
[1523]	<211> 23	
[1524]	<212> DNA	
[1525]	<213> Homo sapiens(智人)	
[1526]	<400> 253	
[1527]	agggctcacc agaggtagga ggg	23
[1528]	<210> 254	
[1529]	<211> 23	
[1530]	<212> DNA	
[1531]	<213> Homo sapiens(智人)	
[1532]	<400> 254	
[1533]	ggagttgatg tcagtcactt ggg	23
[1534]	<210> 255	
[1535]	<211> 23	
[1536]	<212> DNA	
[1537]	<213> Homo sapiens(智人)	
[1538]	<400> 255	
[1539]	tggagttgat gtcagtcact tgg	23
[1540]	<210> 256	
[1541]	<211> 23	
[1542]	<212> DNA	
[1543]	<213> Homo sapiens(智人)	
[1544]	<400> 256	
[1545]	agtgactgac atcaactcca agg	23
[1546]	<210> 257	
[1547]	<211> 23	
[1548]	<212> DNA	
[1549]	<213> Homo sapiens(智人)	
[1550]	<400> 257	
[1551]	gtgactgaca tcaactccaa ggg	23
[1552]	<210> 258	
[1553]	<211> 23	
[1554]	<212> DNA	
[1555]	<213> Homo sapiens(智人)	
[1556]	<400> 258	
[1557]	actccaaggg attggaattg agg	23
[1558]	<210> 259	
[1559]	<211> 23	

[1560]	<212>	DNA	
[1561]	<213>	Homo sapiens(智人)	
[1562]	<400>	259	
[1563]	cttcctcaat tccaatccct tgg		23
[1564]	<210>	260	
[1565]	<211>	23	
[1566]	<212>	DNA	
[1567]	<213>	Homo sapiens(智人)	
[1568]	<400>	260	
[1569]	tacagttgag actcagaact tgg		23
[1570]	<210>	261	
[1571]	<211>	23	
[1572]	<212>	DNA	
[1573]	<213>	Homo sapiens(智人)	
[1574]	<400>	261	
[1575]	ttggaaggcc tgcacatga tgg		23
[1576]	<210>	262	
[1577]	<211>	23	
[1578]	<212>	DNA	
[1579]	<213>	Homo sapiens(智人)	
[1580]	<400>	262	
[1581]	agaattggcc atcatgatgc agg		23
[1582]	<210>	263	
[1583]	<211>	23	
[1584]	<212>	DNA	
[1585]	<213>	Homo sapiens(智人)	
[1586]	<400>	263	
[1587]	gacagggctt atggcagaat tgg		23
[1588]	<210>	264	
[1589]	<211>	23	
[1590]	<212>	DNA	
[1591]	<213>	Homo sapiens(智人)	
[1592]	<400>	264	
[1593]	tgtaacatac ctggaggaca ggg		23
[1594]	<210>	265	
[1595]	<211>	23	
[1596]	<212>	DNA	
[1597]	<213>	Homo sapiens(智人)	
[1598]	<400>	265	

[1599]	gtgtaacata cctggaggac agg	23
[1600]	<210> 266	
[1601]	<211> 23	
[1602]	<212> DNA	
[1603]	<213> Homo sapiens(智人)	
[1604]	<400> 266	
[1605]	cgtacctgtg caactcctgt tgg	23
[1606]	<210> 267	
[1607]	<211> 23	
[1608]	<212> DNA	
[1609]	<213> Homo sapiens(智人)	
[1610]	<400> 267	
[1611]	gatctactgg aattcctaata ggg	23
[1612]	<210> 268	
[1613]	<211> 23	
[1614]	<212> DNA	
[1615]	<213> Homo sapiens(智人)	
[1616]	<400> 268	
[1617]	gagtcagctg ttggcccatt agg	23
[1618]	<210> 269	
[1619]	<211> 23	
[1620]	<212> DNA	
[1621]	<213> Homo sapiens(智人)	
[1622]	<400> 269	
[1623]	ctgcctacaa actcagtctc tgg	23
[1624]	<210> 270	
[1625]	<211> 23	
[1626]	<212> DNA	
[1627]	<213> Homo sapiens(智人)	
[1628]	<400> 270	
[1629]	gggcaggcag gacggactcc agg	23
[1630]	<210> 271	
[1631]	<211> 23	
[1632]	<212> DNA	
[1633]	<213> Homo sapiens(智人)	
[1634]	<400> 271	
[1635]	ggagtcgctc ctgcctgccc tgg	23
[1636]	<210> 272	
[1637]	<211> 23	

[1638]	<212>	DNA	
[1639]	<213>	Homo sapiens(智人)	
[1640]	<400>	272	
[1641]		gagtccgtcc tgcctgccct ggg	23
[1642]	<210>	273	
[1643]	<211>	23	
[1644]	<212>	DNA	
[1645]	<213>	Homo sapiens(智人)	
[1646]	<400>	273	
[1647]		gaaaagggtc cattggccaa agg	23
[1648]	<210>	274	
[1649]	<211>	23	
[1650]	<212>	DNA	
[1651]	<213>	Homo sapiens(智人)	
[1652]	<400>	274	
[1653]		gcctgcagaa aagggtccat tgg	23
[1654]	<210>	275	
[1655]	<211>	23	
[1656]	<212>	DNA	
[1657]	<213>	Homo sapiens(智人)	
[1658]	<400>	275	
[1659]		ttgatgtgct acaggaaca tgg	23
[1660]	<210>	276	
[1661]	<211>	23	
[1662]	<212>	DNA	
[1663]	<213>	Homo sapiens(智人)	
[1664]	<400>	276	
[1665]		agcgttcttg atgtgctaca ggg	23
[1666]	<210>	277	
[1667]	<211>	23	
[1668]	<212>	DNA	
[1669]	<213>	Homo sapiens(智人)	
[1670]	<400>	277	
[1671]		cagcgttctt gatgtgctac agg	23
[1672]	<210>	278	
[1673]	<211>	23	
[1674]	<212>	DNA	
[1675]	<213>	Homo sapiens(智人)	
[1676]	<400>	278	

[1677]	ctgtagcaca tcaagaacgc tgg	23
[1678]	<210> 279	
[1679]	<211> 23	
[1680]	<212> DNA	
[1681]	<213> Homo sapiens(智人)	
[1682]	<400> 279	
[1683]	tgtagcacat caagaacgct ggg	23
[1684]	<210> 280	
[1685]	<211> 23	
[1686]	<212> DNA	
[1687]	<213> Homo sapiens(智人)	
[1688]	<400> 280	
[1689]	ataggcaata atcatataac agg	23
[1690]	<210> 281	
[1691]	<211> 23	
[1692]	<212> DNA	
[1693]	<213> Homo sapiens(智人)	
[1694]	<400> 281	
[1695]	agtgcgtttc gctgcaggta agg	23
[1696]	<210> 282	
[1697]	<211> 23	
[1698]	<212> DNA	
[1699]	<213> Homo sapiens(智人)	
[1700]	<400> 282	
[1701]	gagtgagtgc gtttcgctgc agg	23
[1702]	<210> 283	
[1703]	<211> 23	
[1704]	<212> DNA	
[1705]	<213> Homo sapiens(智人)	
[1706]	<400> 283	
[1707]	gtcaggtttg tgcggttatg agg	23
[1708]	<210> 284	
[1709]	<211> 23	
[1710]	<212> DNA	
[1711]	<213> Homo sapiens(智人)	
[1712]	<400> 284	
[1713]	cgctgctggt caggtttgtg cgg	23
[1714]	<210> 285	
[1715]	<211> 23	

[1716]	<212>	DNA	
[1717]	<213>	Homo sapiens(智人)	
[1718]	<400>	285	
[1719]	aaacctgacc agcagcgagcagg		23
[1720]	<210>	286	
[1721]	<211>	23	
[1722]	<212>	DNA	
[1723]	<213>	Homo sapiens(智人)	
[1724]	<400>	286	
[1725]	ccagcagcgc agaggagccg tgg		23
[1726]	<210>	287	
[1727]	<211>	23	
[1728]	<212>	DNA	
[1729]	<213>	Homo sapiens(智人)	
[1730]	<400>	287	
[1731]	ccacggctcc tctgcgctgc tgg		23
[1732]	<210>	288	
[1733]	<211>	23	
[1734]	<212>	DNA	
[1735]	<213>	Homo sapiens(智人)	
[1736]	<400>	288	
[1737]	ccaactatct aactccactc agg		23
[1738]	<210>	289	
[1739]	<211>	23	
[1740]	<212>	DNA	
[1741]	<213>	Homo sapiens(智人)	
[1742]	<400>	289	
[1743]	cctgagtgga gttagatagt tgg		23

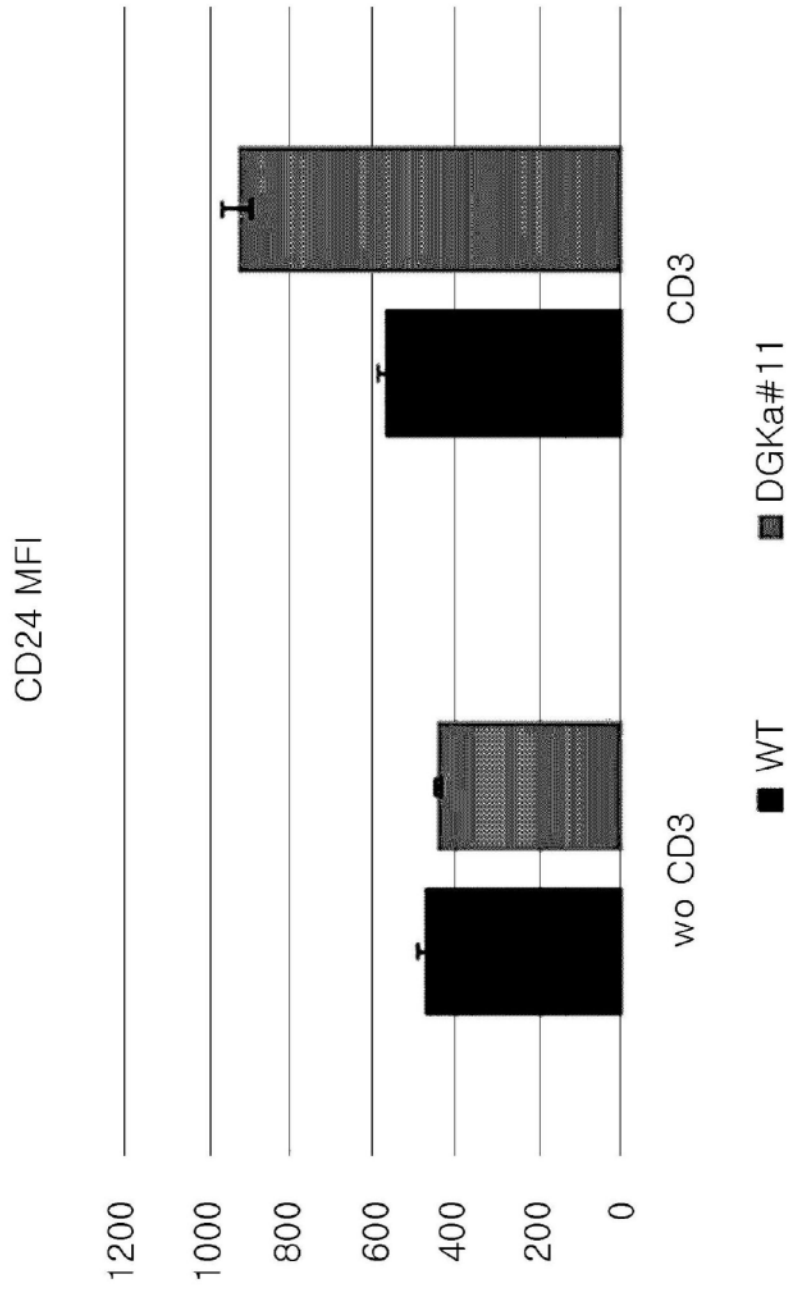


图1

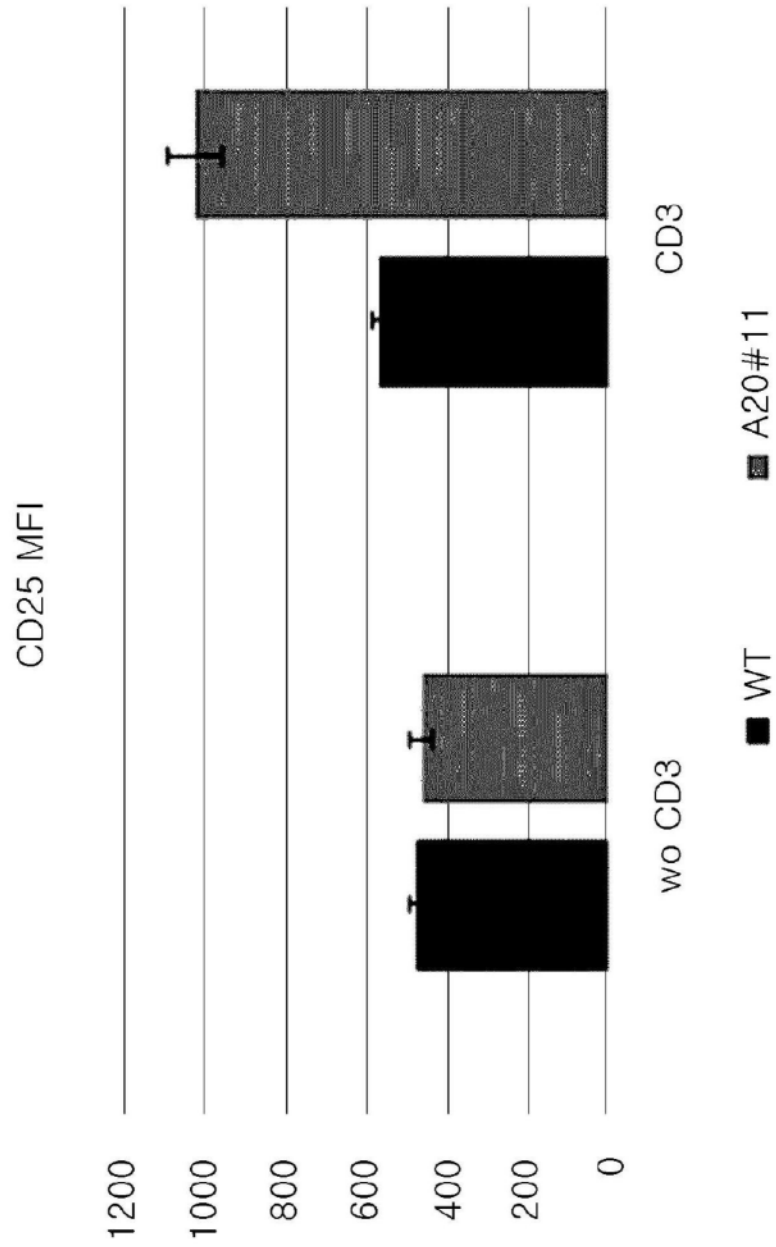


图2

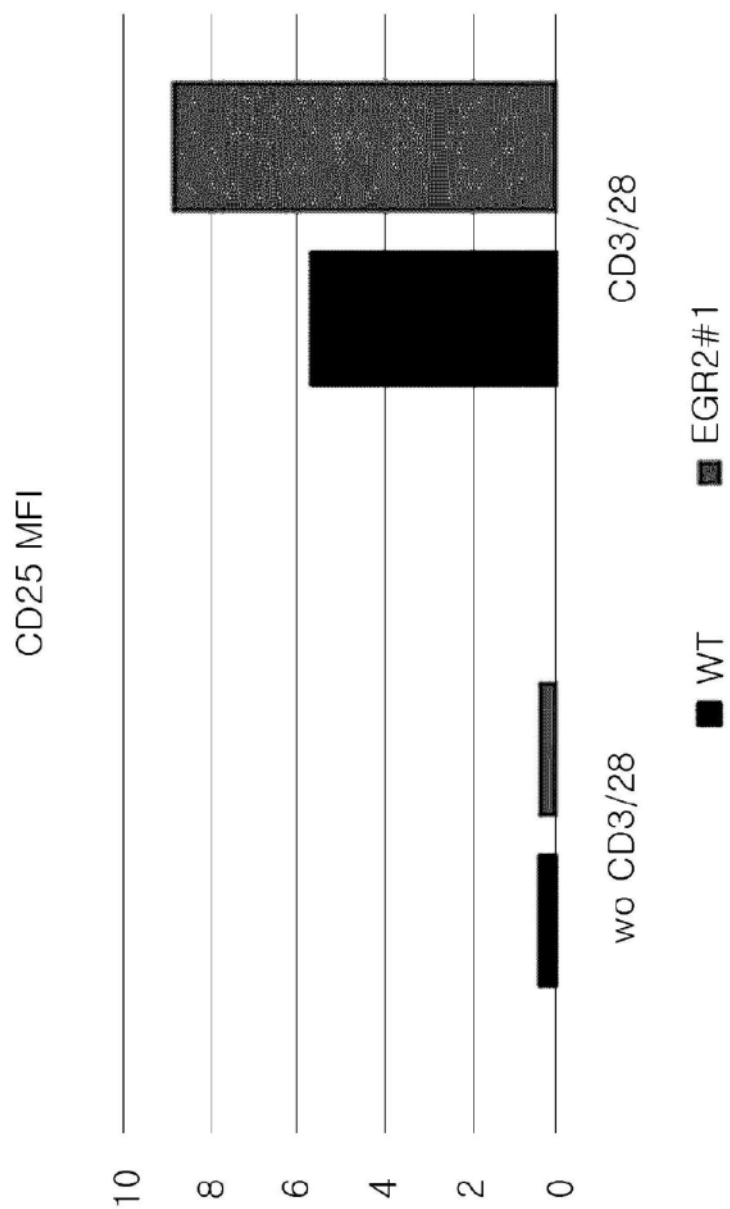


图3

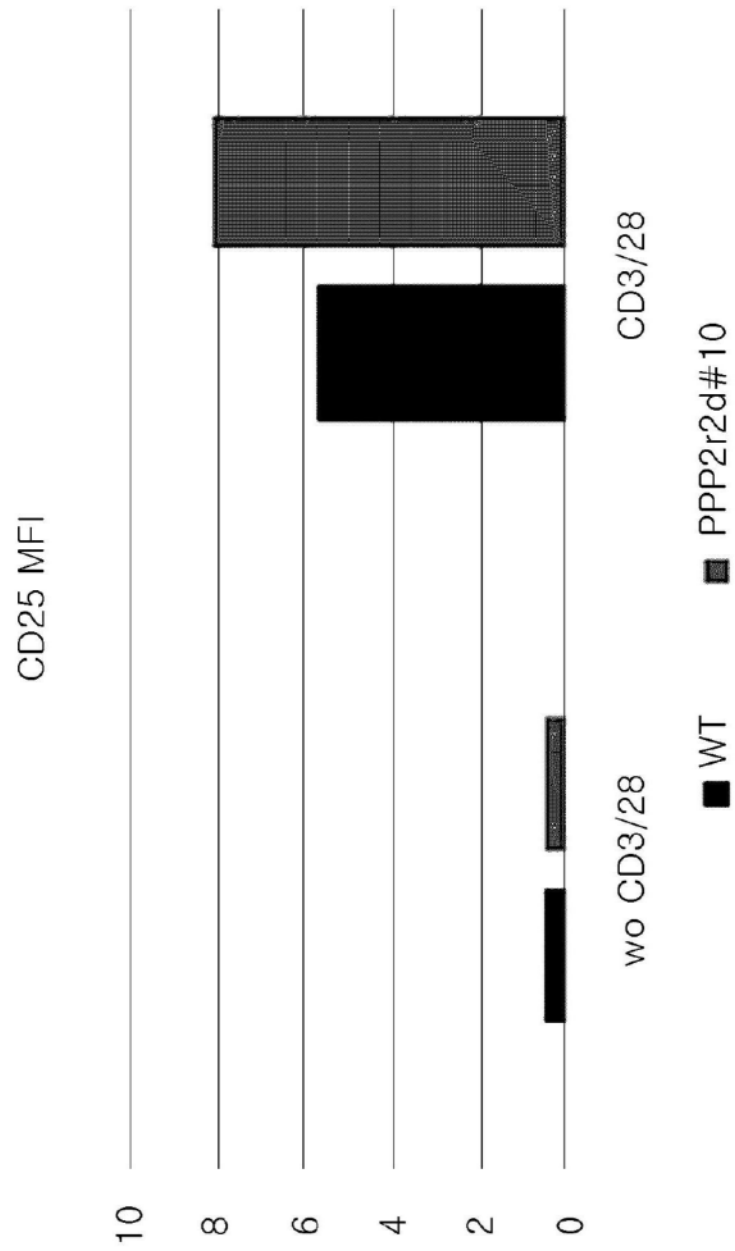


图4

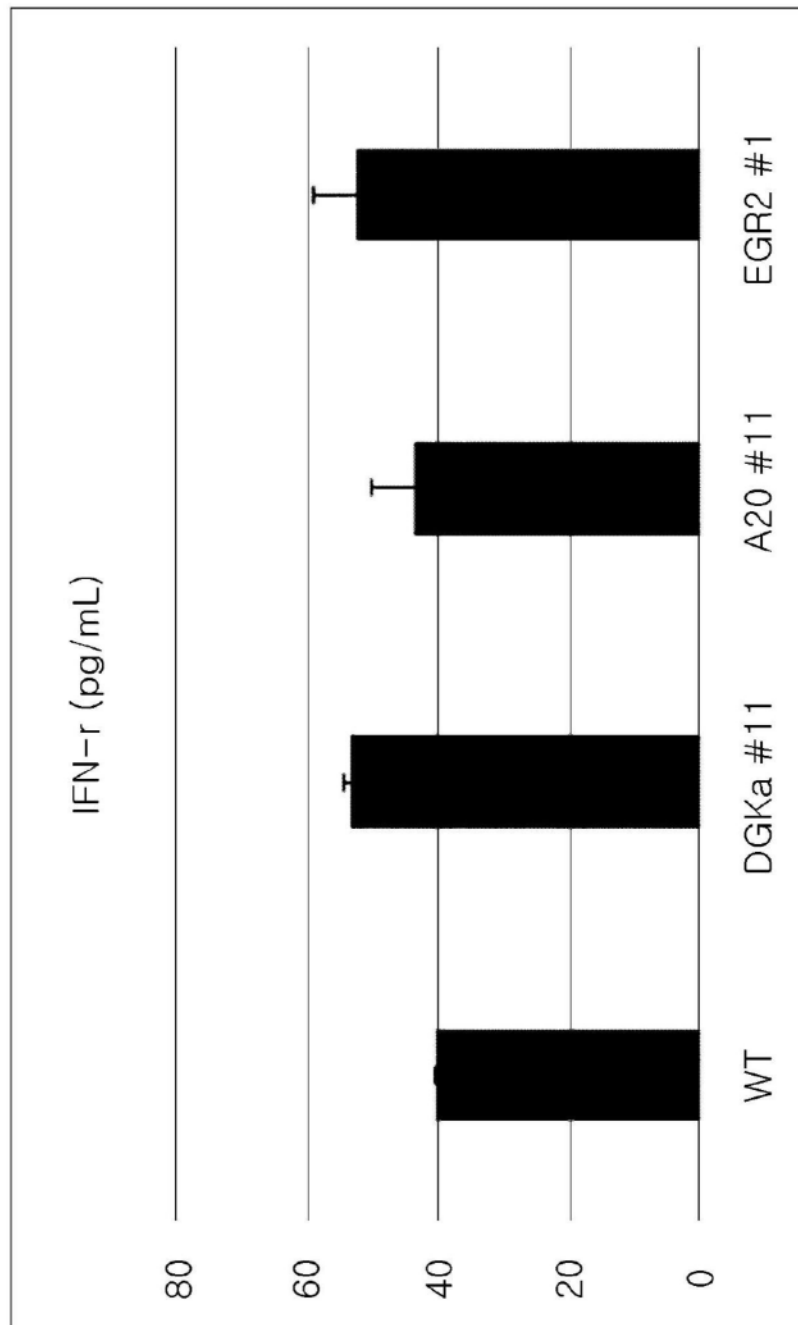


图5

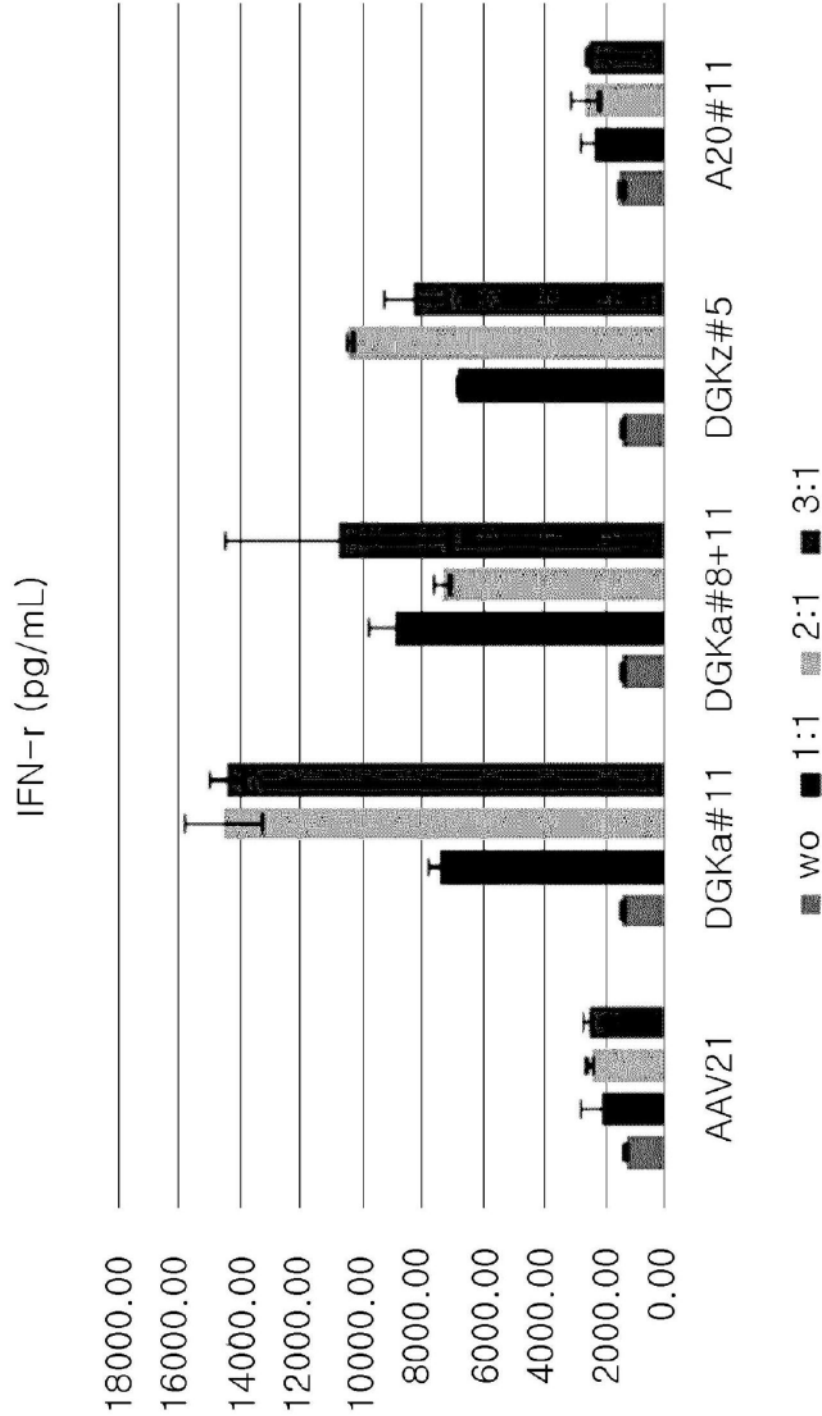


图6

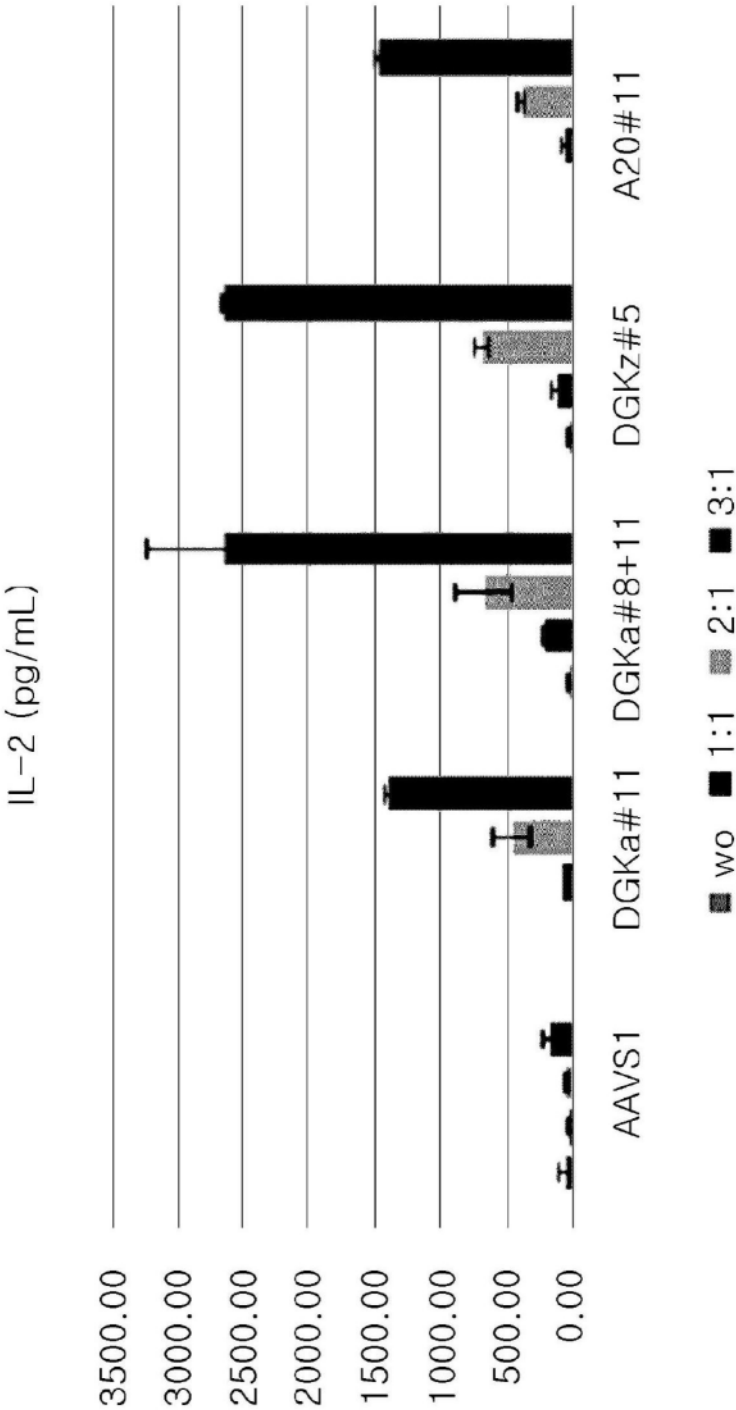
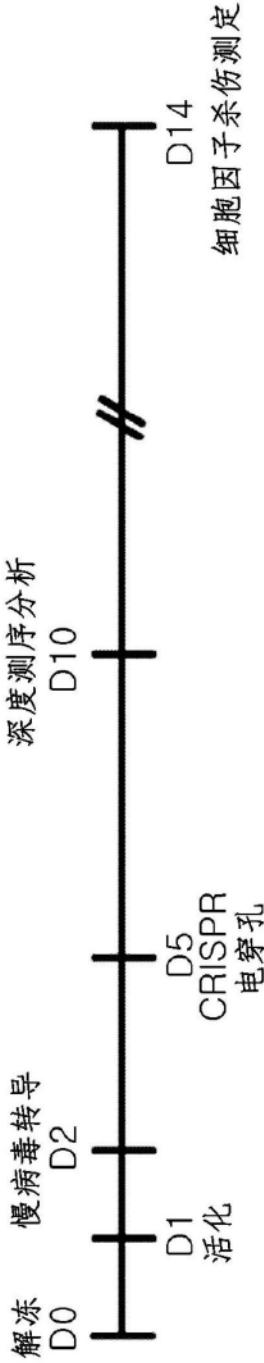


图7

(a)



(b)

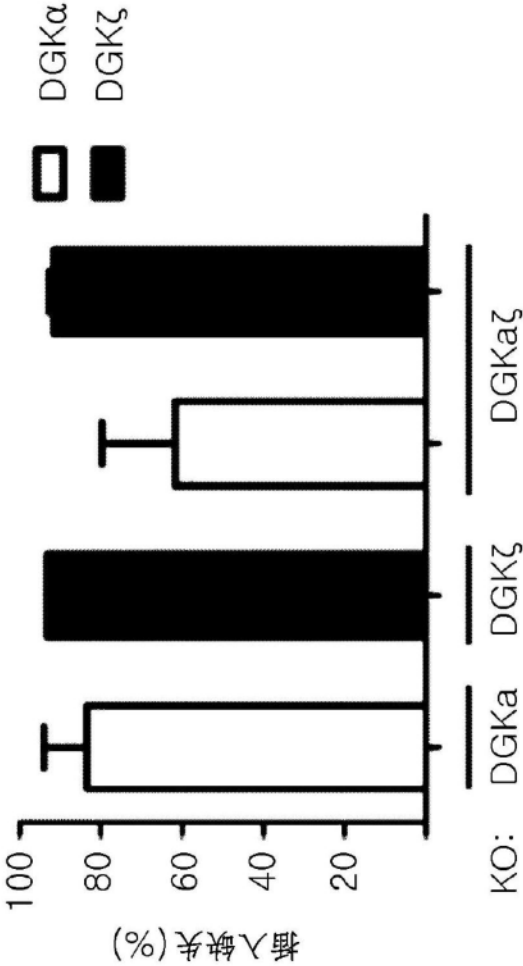


图8a

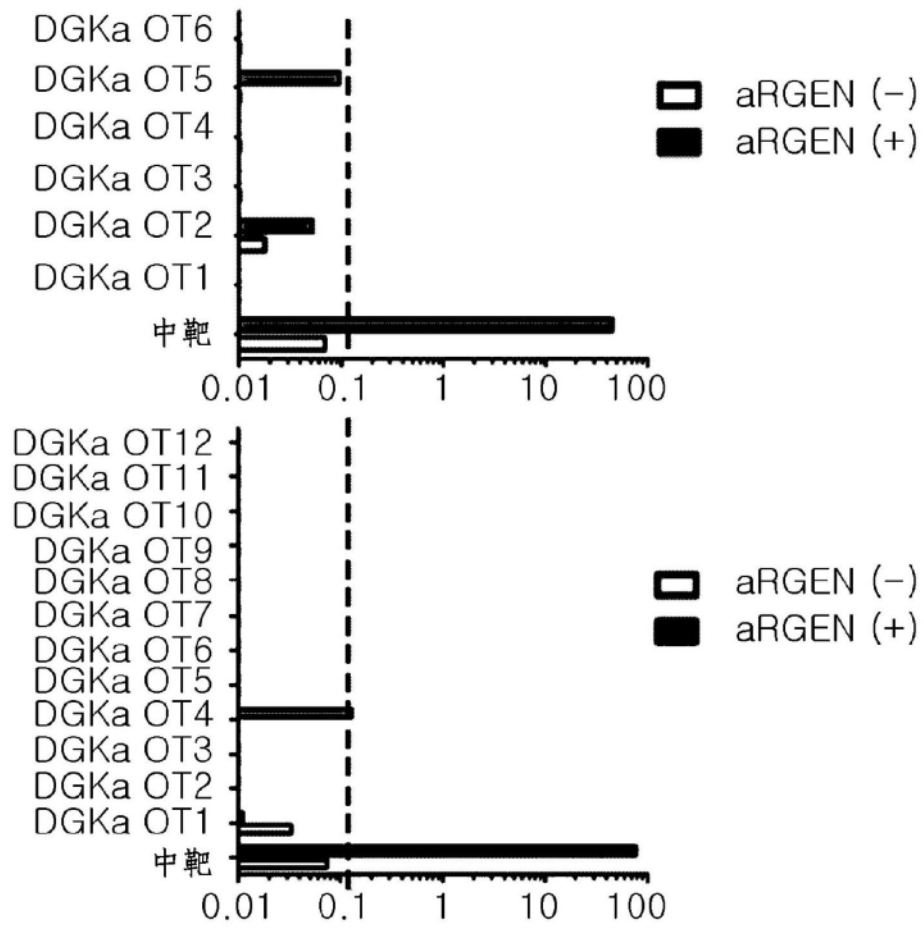


图8b

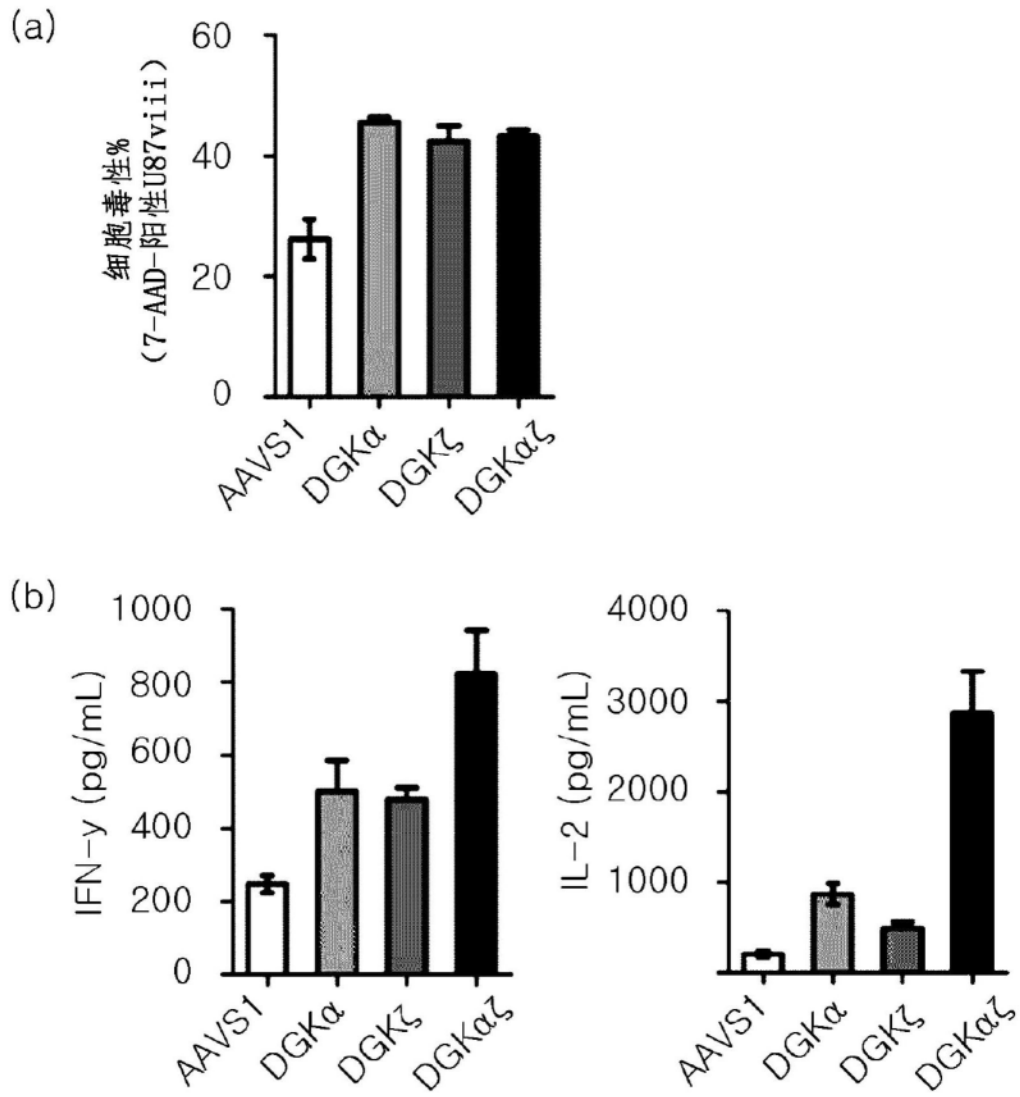


图9a

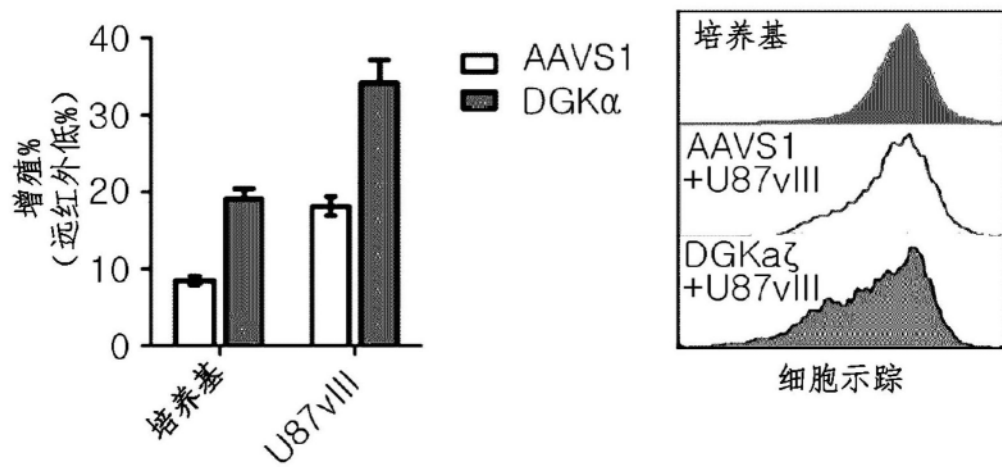


图9b

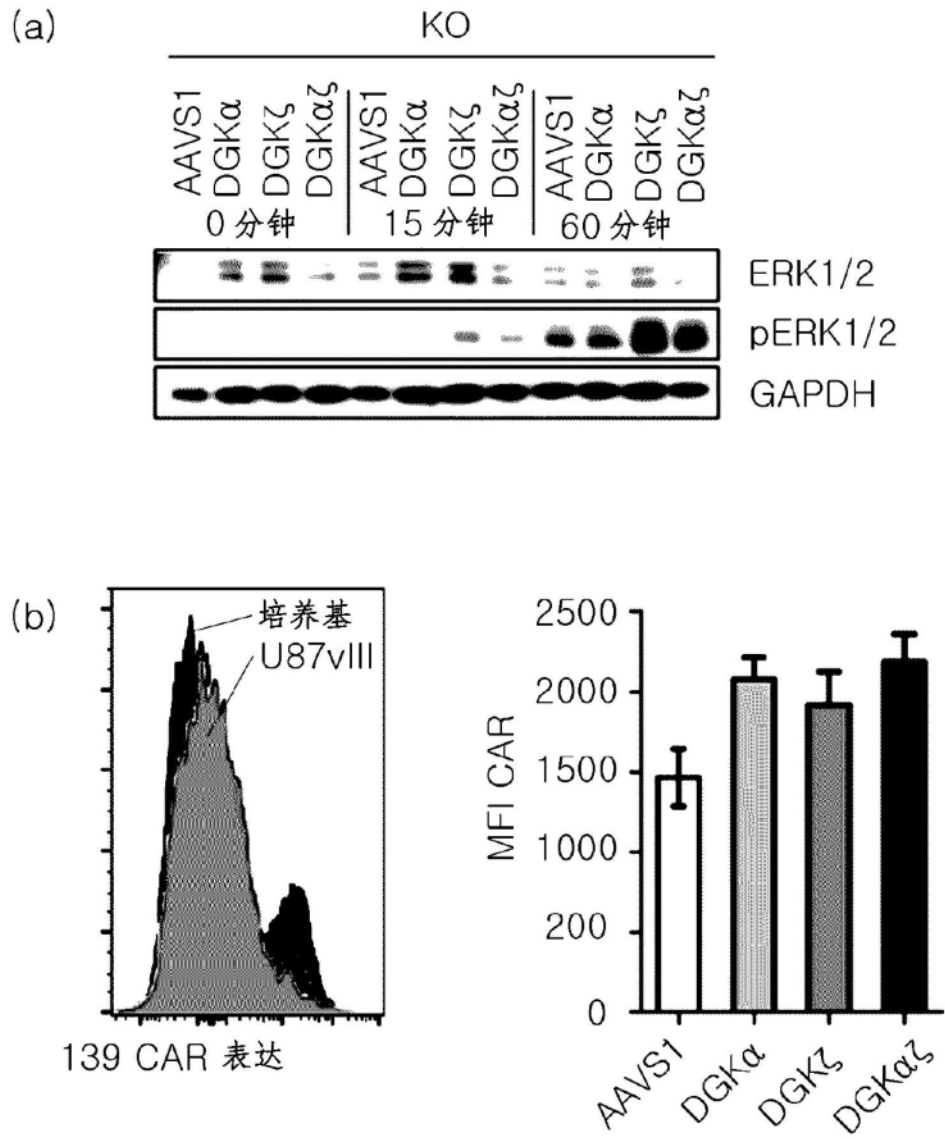


图10

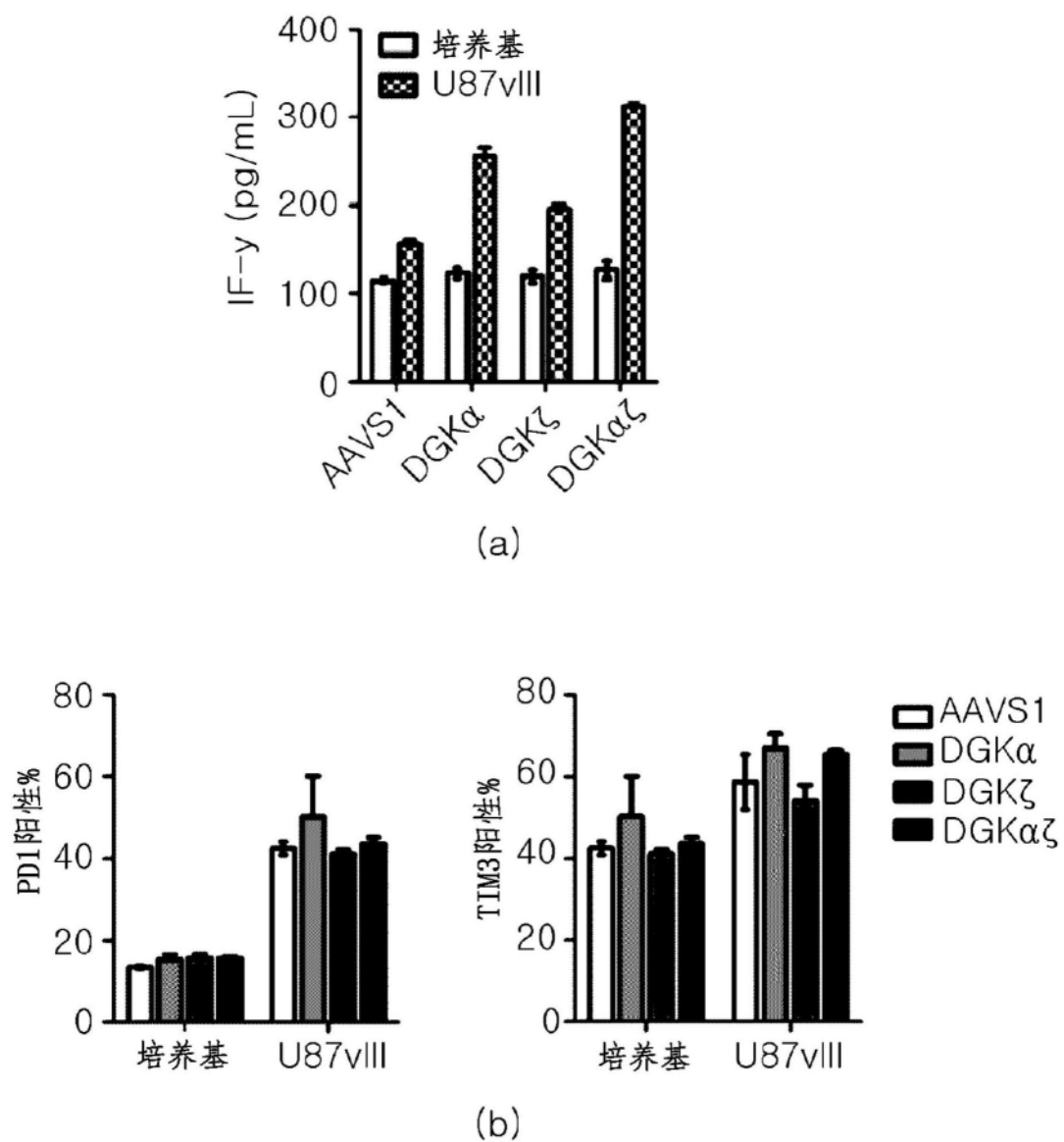


图11

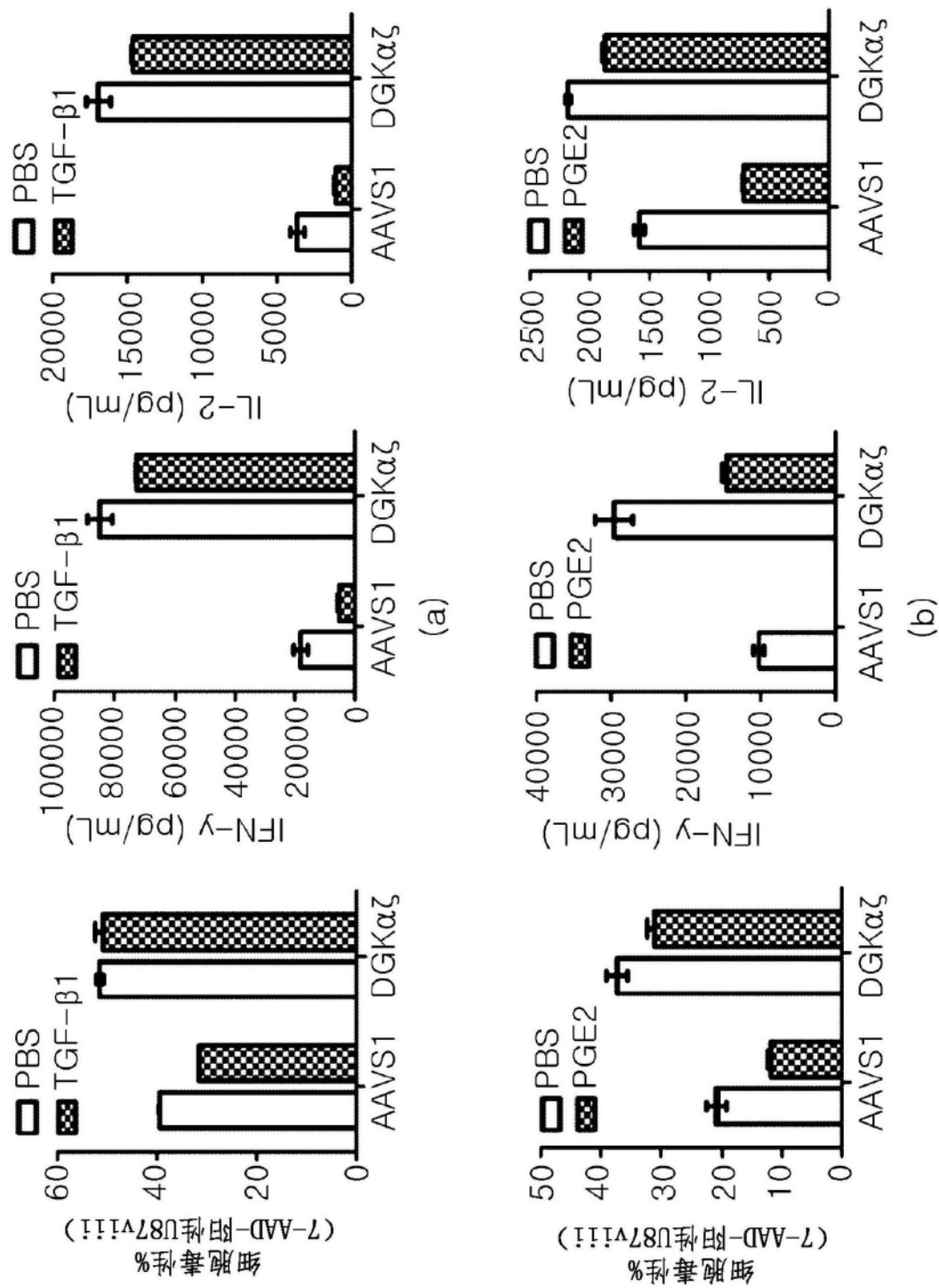


图12

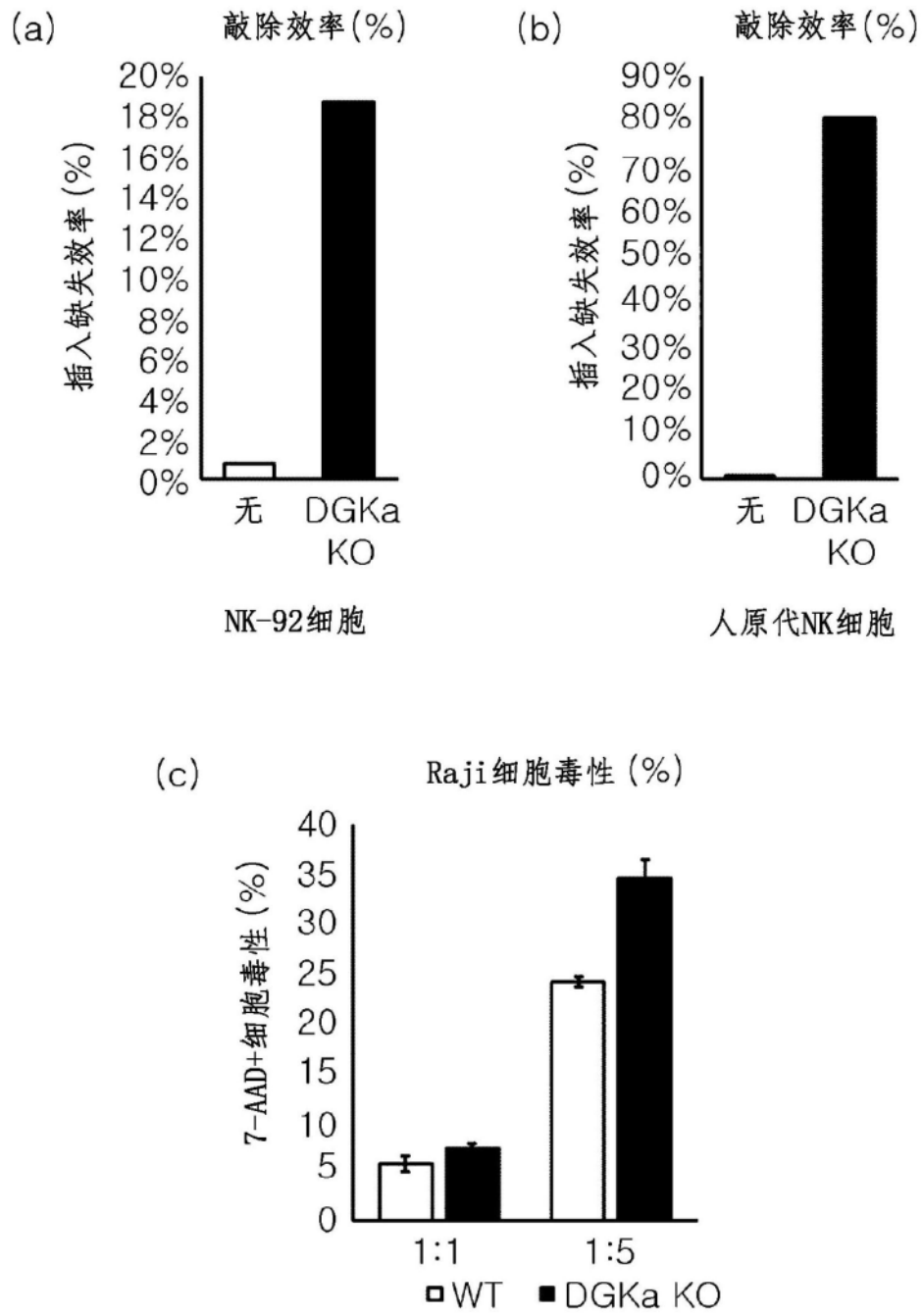


图13

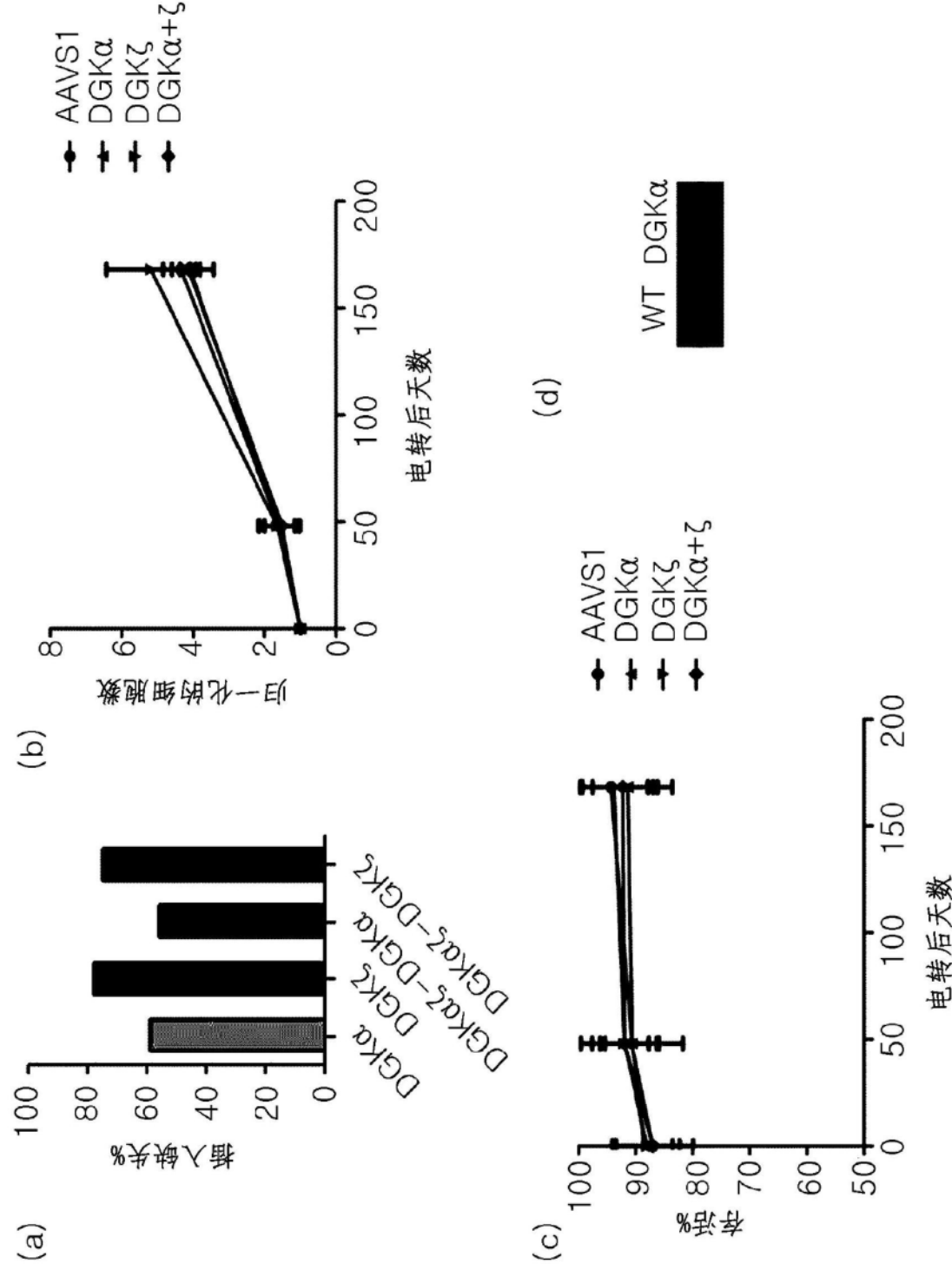


图14

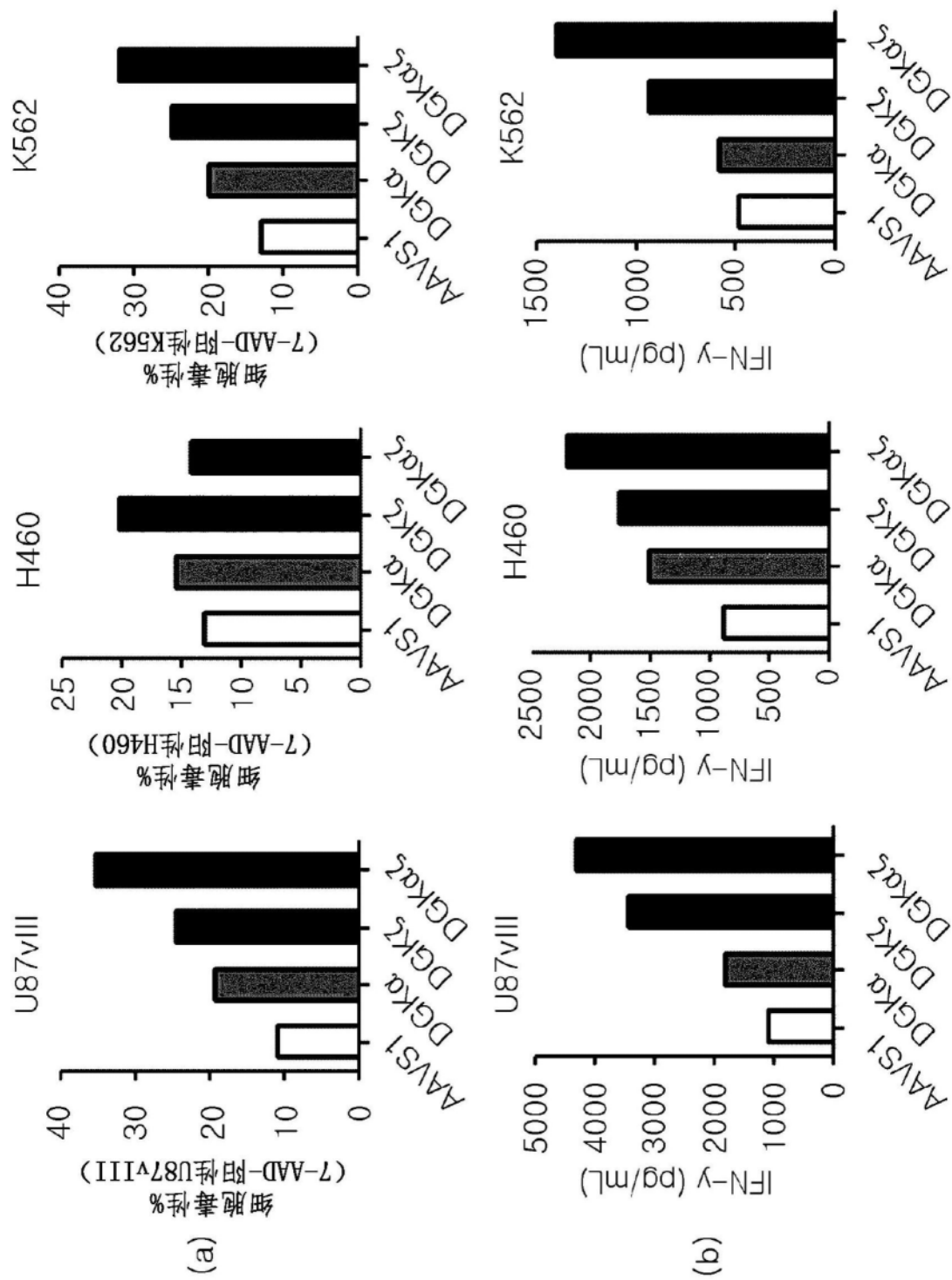


图15

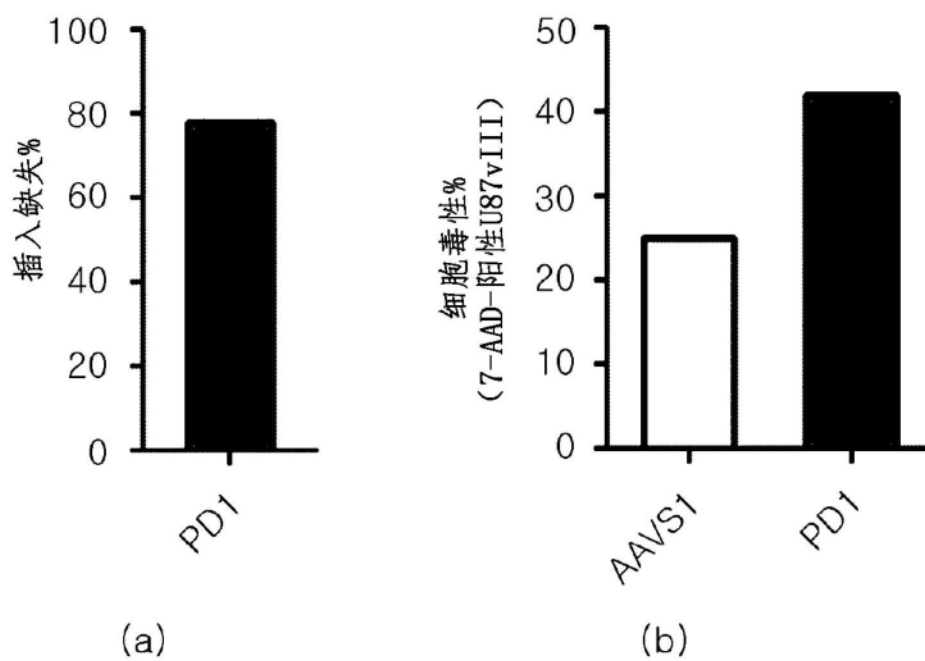


图16

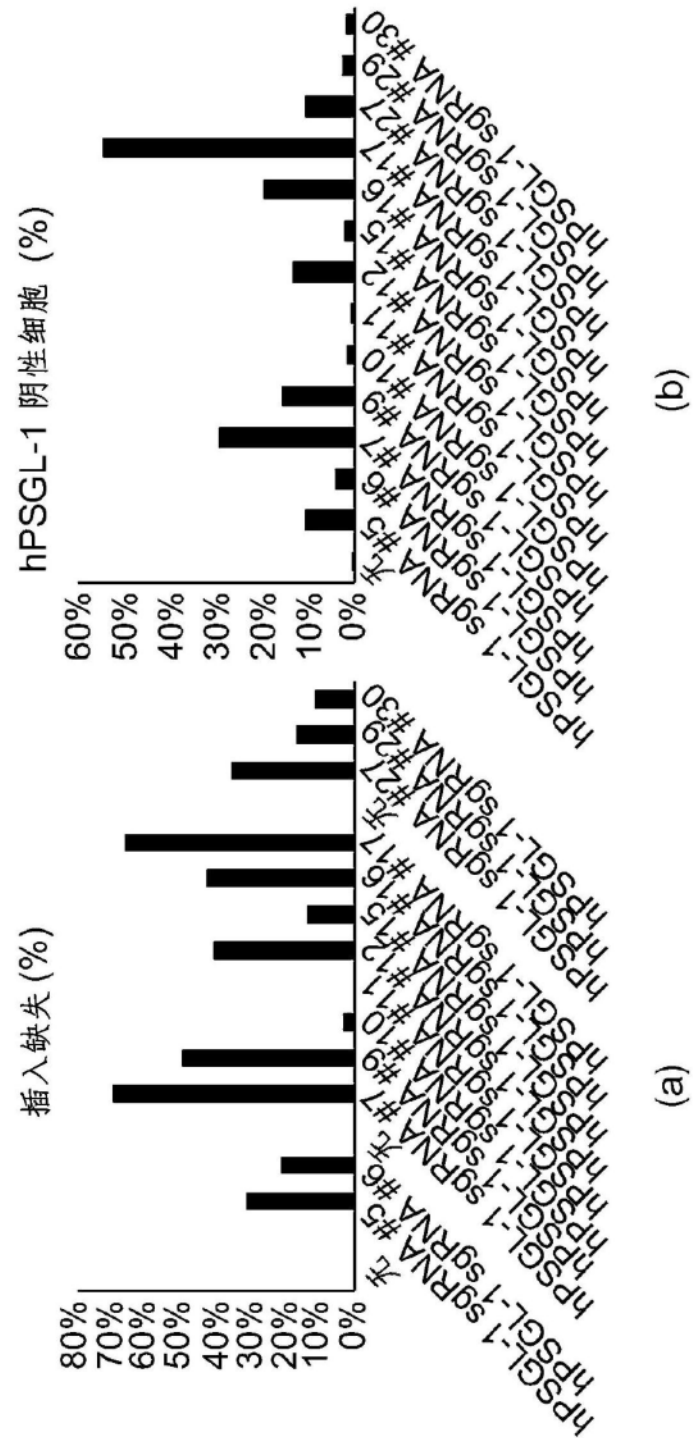


图17a

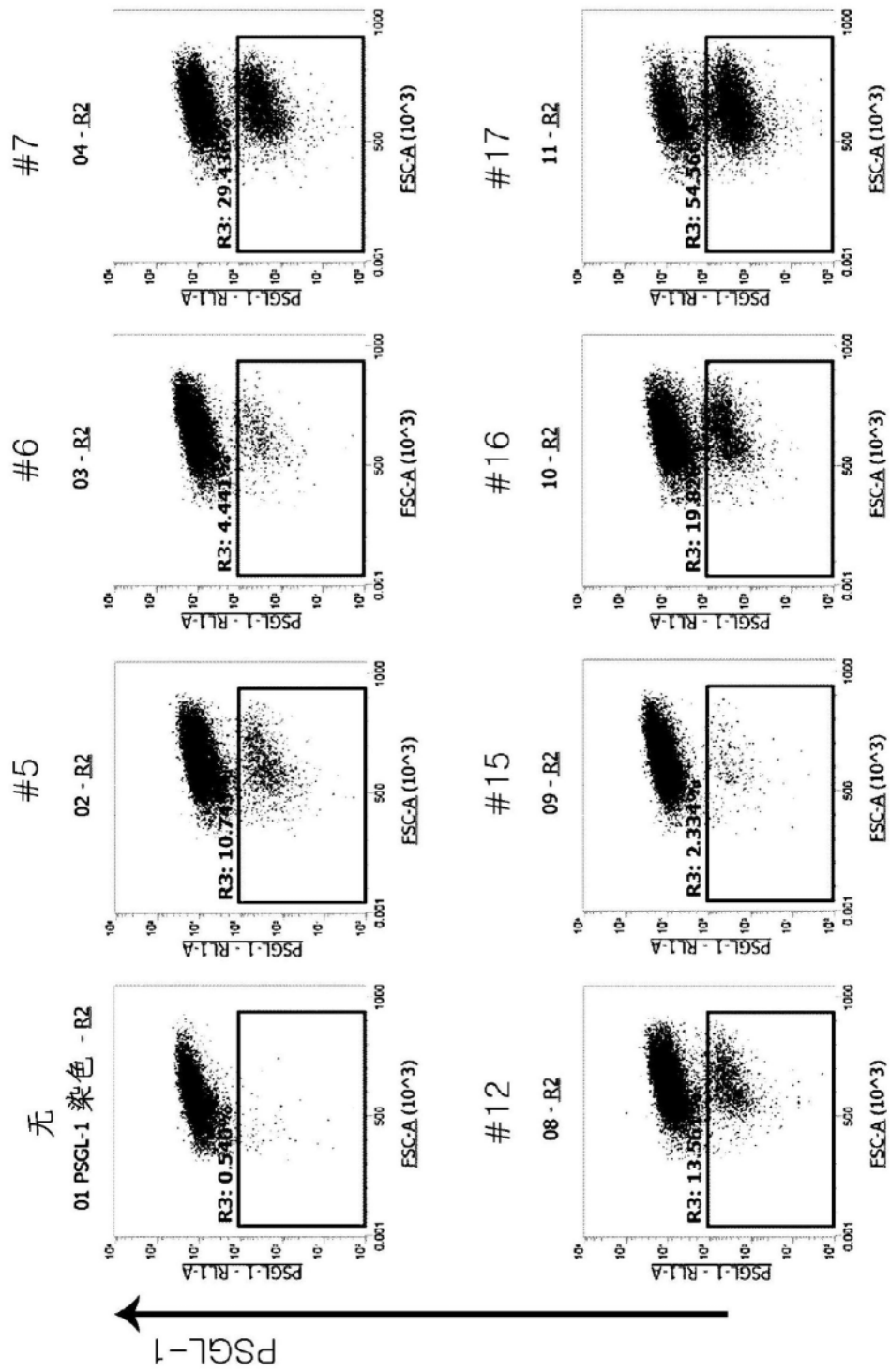


图17b

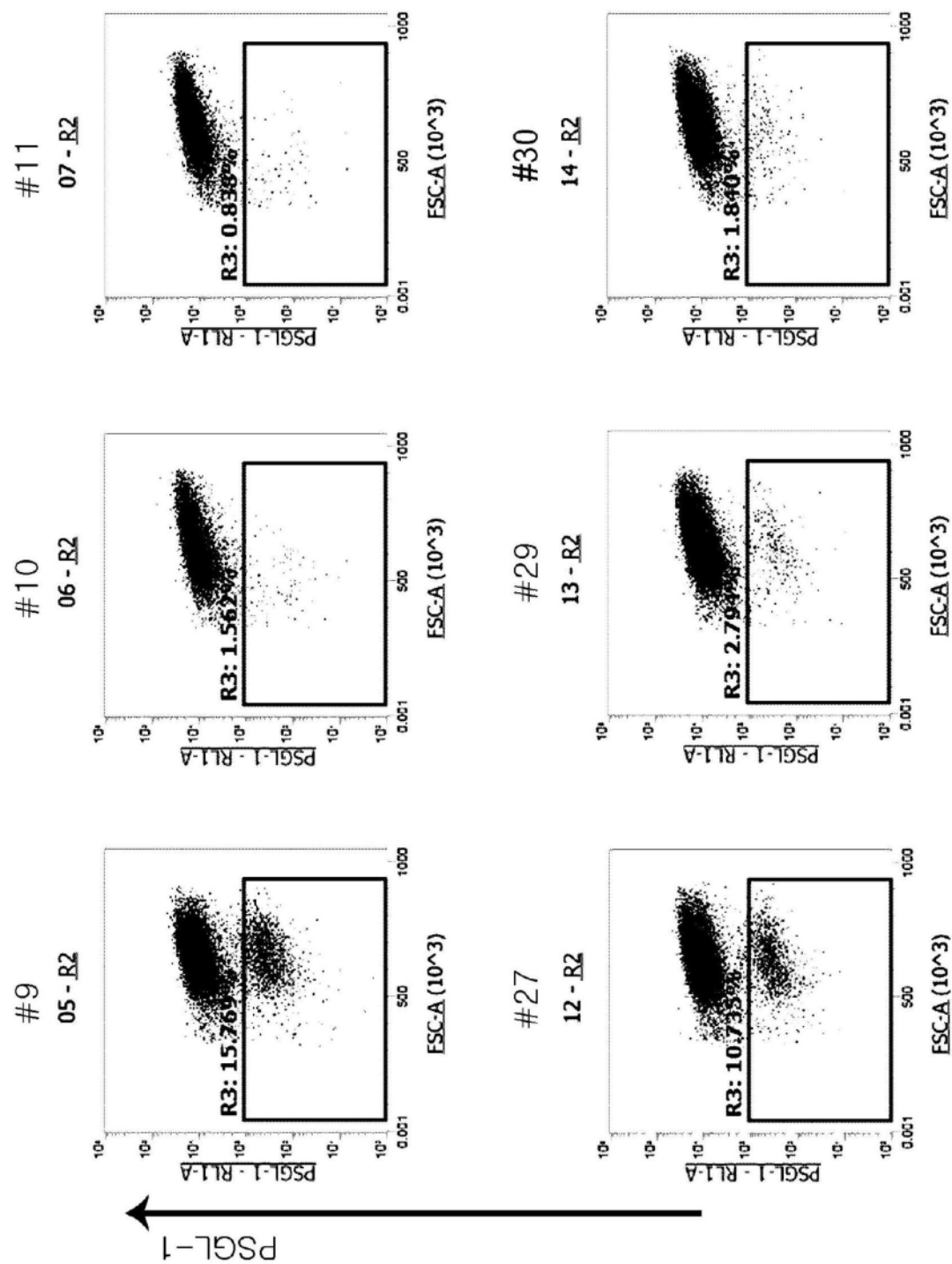


图17c

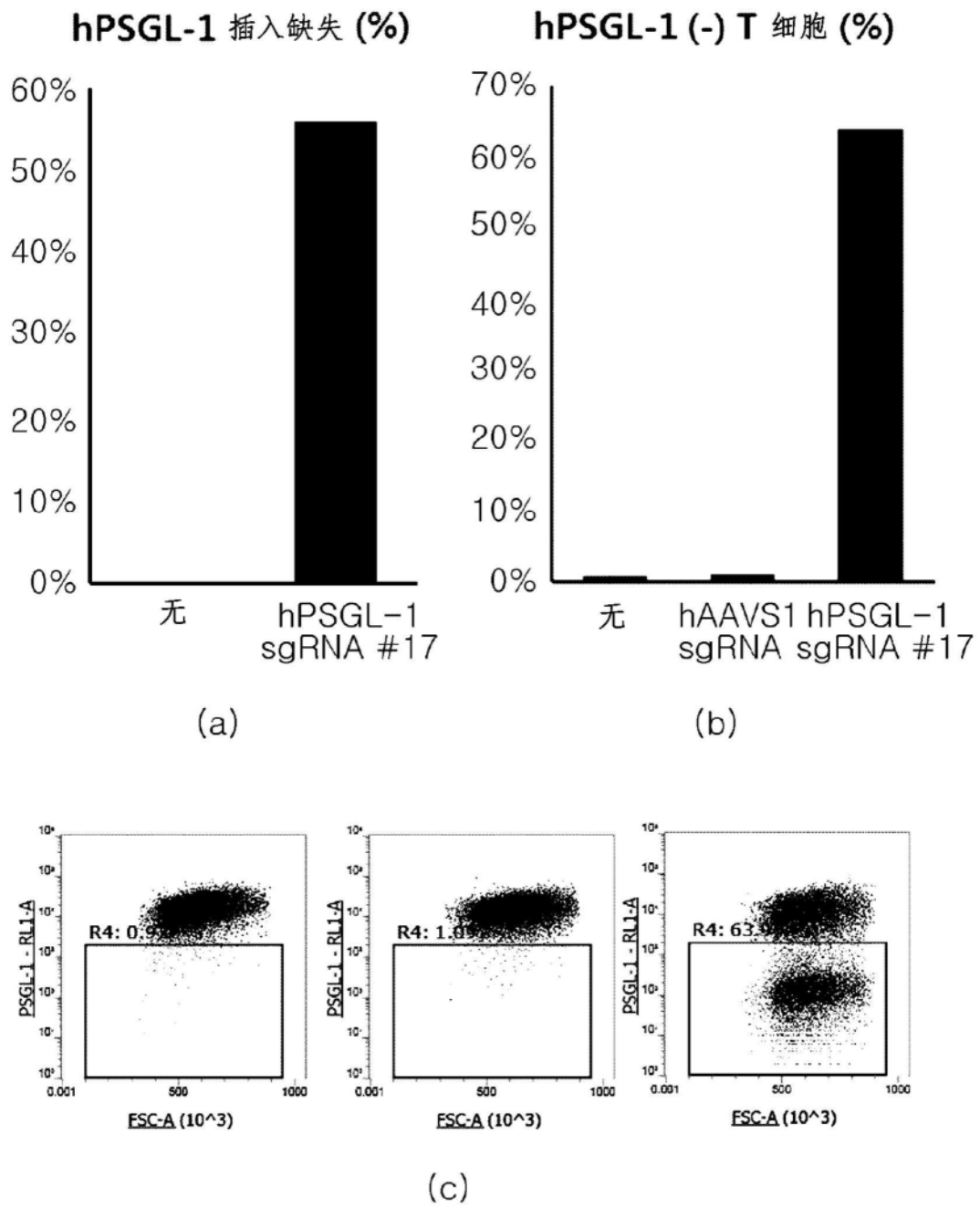


图18