



## (12)发明专利

(10)授权公告号 CN 105164153 B

(45)授权公告日 2020.04.24

(21)申请号 201480021698.3

(22)申请日 2014.02.20

(65)同一申请的已公布的文献号  
申请公布号 CN 105164153 A

(43)申请公布日 2015.12.16

(30)优先权数据  
61/766,762 2013.02.20 US  
61/890,915 2013.10.15 US

(85)PCT国际申请进入国家阶段日  
2015.10.16

(86)PCT国际申请的申请数据  
PCT/US2014/017395 2014.02.20

(87)PCT国际申请的公布数据  
W02014/130671 EN 2014.08.28

(73)专利权人 瑞泽恩制药公司  
地址 美国纽约州

(72)发明人 L·麦克唐纳 A·J·莫菲 N·图  
V·沃罗宁那 C·古雷尔

(74)专利代理机构 北京商专永信知识产权代理  
事务所(普通合伙) 11400  
代理人 郭玥 葛强

(51)Int.Cl.  
C07K 14/73(2006.01)  
C07K 14/705(2006.01)  
C12N 15/85(2006.01)  
A01K 67/027(2006.01)

(56)对比文件  
US 2005066375 A1,2005.03.24,  
WO 2011004192 A1,2011.01.13,  
Sherman et al..Selecting T Cell  
Receptors with High Affinity for Self-MHC  
by Decreasing the Contribution of CD8.  
《science》.1992,第258卷(第5083期),第815-  
816页.

审查员 段珊

权利要求书3页 说明书29页  
序列表32页 附图7页

### (54)发明名称

人源化的T细胞共受体的小鼠

### (57)摘要

本发明提供了表达嵌合的人/非人T细胞共受体多肽(例如CD4、CD8 $\alpha$ 、CD8 $\beta$ )的基因修饰的非人动物,及包含其的胚胎、细胞和组织。本发明还提供了制备所述基因修饰的动物的构建体和制备其的方法。

1. 一种制备基因修饰的啮齿动物的方法,包括修饰啮齿动物基因组以包含

(a) 在内源CD4基因座并且与内源CD4启动子和调控序列可操作地连接的编码嵌合的人/啮齿动物CD4多肽的核苷酸序列,所述嵌合的人/啮齿动物CD4多肽包含人CD4的至少D1、D2和D3结构域,所述D1、D2和D3结构域可操作地连接啮齿动物CD4的D4、跨膜和胞质结构域,

(b) 在内源MHCII $\alpha$ 基因座并且与内源MHC II $\alpha$ 启动子和调控序列可操作地连接的编码嵌合的人/啮齿动物MHCII $\alpha$ 多肽的核苷酸序列,所述嵌合的人/啮齿动物MHCII $\alpha$ 多肽包含与啮齿动物MHC II $\alpha$ 多肽的跨膜和胞质结构域可操作连接的人MHCII $\alpha$ 多肽的 $\alpha$ 1和 $\alpha$ 2结构域,和

(c) 在内源MHCII $\beta$ 基因座并且与内源MHC II $\beta$ 启动子和调控序列可操作地连接的编码嵌合的人/啮齿动物MHCII $\beta$ 多肽的核苷酸序列,所述嵌合的人/啮齿动物MHCII $\beta$ 多肽包含与啮齿动物MHC II $\beta$ 多肽的跨膜和胞质结构域可操作连接的人MHCII $\beta$ 多肽的 $\beta$ 1和 $\beta$ 2结构域,

其中所述啮齿动物表达所述嵌合的人/啮齿动物CD4多肽和嵌合的人/啮齿动物MHC II复合物,所述嵌合的人/啮齿动物MHC II复合物包含(i)所述嵌合的人/啮齿动物MHCII $\alpha$ 多肽和(ii)所述嵌合的人/啮齿动物MHC II $\beta$ 多肽,并且其中所述嵌合的人/啮齿动物CD4多肽和所述嵌合的人/啮齿动物MHC II复合物相互作用。

2. 根据权利要求1所述的方法,其中在所述啮齿动物的B细胞上不表达所述嵌合的CD4多肽。

3. 根据权利要求1所述的方法,其中在B细胞或CD8单一阳性T细胞上不表达所述嵌合的CD4多肽。

4. 根据权利要求1-3任意一项所述的方法,其中所述啮齿动物是大鼠。

5. 根据权利要求1-3任意一项所述的方法,其中所述啮齿动物是小鼠。

6. 根据权利要求5所述的方法,其中所述啮齿动物基因组是小鼠基因组,并且所述方法包括修饰所述小鼠基因组以包含:

(a) 在内源小鼠CD4基因座并且与内源小鼠CD4启动子和调控序列可操作地连接的编码嵌合的人/小鼠CD4多肽的核苷酸序列,所述嵌合的人/小鼠CD4多肽包含人CD4的D1、D2和D3结构域,所述D1、D2和D3结构域可操作地连接小鼠CD4的D4、跨膜和胞质结构域,

(b) 在内源小鼠MHCII $\alpha$ 基因座并且与内源小鼠MHC II $\alpha$ 启动子和调控序列可操作地连接的编码嵌合的人/小鼠MHCII $\alpha$ 多肽的核苷酸序列,所述嵌合的人/小鼠MHCII $\alpha$ 多肽包含与小鼠MHC II $\alpha$ 多肽的跨膜和胞质结构域可操作连接的人MHCII $\alpha$ 多肽的 $\alpha$ 1和 $\alpha$ 2结构域,

(c) 在内源小鼠MHCII $\beta$ 基因座并且与内源小鼠MHC II $\beta$ 启动子和调控序列可操作地连接的编码嵌合的人/小鼠MHCII $\beta$ 多肽的核苷酸序列,所述嵌合的人/小鼠MHCII $\beta$ 多肽包含与小鼠MHC II $\beta$ 多肽的跨膜和胞质结构域可操作连接的人MHCII $\beta$ 多肽的 $\beta$ 1和 $\beta$ 2结构域,且

其中所述小鼠表达所述嵌合的人/小鼠CD4多肽和嵌合的人/小鼠MHC II复合物,所述嵌合的人/小鼠MHC II复合物包含(i)所述嵌合的人/小鼠MHC II $\alpha$ 多肽和(ii)所述嵌合的人/小鼠MHCII $\beta$ 多肽,并且其中所述嵌合的人/小鼠CD4多肽和所述嵌合的人/小鼠MHC II复合物相互作用。

7. 根据权利要求6所述的方法,其中所述小鼠不表达来自其内源性小鼠CD4基因座的功能性内源性小鼠CD4。

8. 根据权利要求6所述的方法,其中编码所述嵌合的人/小鼠CD4多肽的所述核苷酸序

列包含在所述小鼠的种系中。

9. 根据权利要求6所述的方法,其中修饰所述小鼠基因组包括使用编码所述嵌合的人/小鼠CD4多肽的所述核苷酸序列取代在所述小鼠基因组内源性CD4基因座的编码内源性小鼠CD4多肽的核苷酸序列。

10. 根据权利要求6所述的方法,其中所述嵌合的人/小鼠CD4多肽的人部分如SEQ ID NO:57所示。

11. 根据权利要求10所述的方法,其中所述嵌合的人/小鼠CD4多肽包含如SEQ ID NO:4所示的氨基酸序列。

12. 一种制备基因修饰的啮齿动物的方法,包括修饰啮齿动物基因组以包含  
在于内源CD8 $\alpha$ 基因座并且与内源CD8 $\alpha$ 启动子和调控序列可操作地连接的编码嵌合的人/啮齿动物CD8 $\alpha$ 多肽的核苷酸序列,

在内源CD8 $\beta$ 基因座并且与内源CD8 $\beta$ 启动子和调控序列可操作地连接的编码嵌合的人/啮齿动物CD8 $\beta$ 多肽的核苷酸序列,和

在内源MHC I基因座并且与内源MHC I启动子和调控序列可操作地连接的编码人源化MHC I多肽的核苷酸序列,

其中所述嵌合的人/啮齿动物CD8 $\alpha$ 多肽的人部分如SEQ ID NO:59所示,

其中所述嵌合的人/啮齿动物CD8 $\beta$ 多肽的人部分如SEQ ID NO:58所示,

其中所述人源化MHC I多肽包含与啮齿动物MHC I多肽的跨膜和胞质结构域可操作连接的人MHC I多肽的 $\alpha 1$ 、 $\alpha 2$ 和 $\alpha 3$ 结构域;

其中所述啮齿动物表达所述人源化MHC I多肽和嵌合的人/啮齿动物CD8蛋白,所述嵌合的人/啮齿动物CD8蛋白包含所述嵌合的人/啮齿动物CD8 $\alpha$ 多肽和嵌合的人/啮齿动物CD8 $\beta$ 多肽,且

其中所述嵌合的人/啮齿动物CD8蛋白与所述人源化的MHC I多肽相互作用。

13. 根据权利要求12所述的方法,其中在所述啮齿动物的B细胞上不表达所述嵌合的人/啮齿动物CD8多肽。

14. 根据权利要求12所述的方法,其中在B细胞或CD4单一阳性T细胞上不表达所述嵌合的人/啮齿动物CD8蛋白。

15. 根据权利要求12-14任意一项所述的方法,其中所述啮齿动物是大鼠。

16. 根据权利要求12-14任意一项所述的方法,其中所述啮齿动物是小鼠。

17. 根据权利要求16所述的方法,其中所述啮齿动物基因组是小鼠基因组,并且所述方法包括修饰所述小鼠基因组以包含

在内源小鼠CD8 $\alpha$ 基因座并且与内源小鼠CD8 $\alpha$ 启动子和调控序列可操作地连接的编码嵌合的人/小鼠CD8 $\alpha$ 多肽的第一核苷酸序列,和

在内源小鼠CD8 $\beta$ 基因座并且与内源小鼠CD8 $\beta$ 启动子和调控序列可操作地连接的编码嵌合的人/小鼠CD8 $\beta$ 多肽的第二核苷酸序列,

在内源小鼠MHC I基因座并且与内源小鼠MHC I启动子和调控序列可操作地连接的编码人源化MHC I多肽的核苷酸序列,

其中所述嵌合的人/小鼠CD8 $\alpha$ 多肽包含如SEQ ID NO:59所示的氨基酸序列,其可操作地连接小鼠CD8 $\alpha$ 多肽的至少跨膜和胞质结构域,

其中所述嵌合的人/小鼠CD8 $\beta$ 多肽包含如SEQ ID NO:58所示的氨基酸序列,其可操作地连接小鼠CD8 $\beta$ 多肽的至少跨膜和胞质结构域,

其中所述人源化的MHC I多肽包含与小鼠MHC I多肽的跨膜和胞质结构域可操作连接的人MHC I多肽的 $\alpha 1$ 、 $\alpha 2$ 和 $\alpha 3$ 结构域,且

其中所述小鼠表达嵌合的人/小鼠CD8蛋白,所述嵌合的人/小鼠CD8蛋白包含(i)所述嵌合的人/小鼠CD8 $\alpha$ 多肽和(ii)所述嵌合的人/小鼠CD8 $\beta$ 多肽和所述人源化MHC I多肽,且所述嵌合的人/小鼠CD8蛋白与所述人源化的MHC I多肽相互作用。

18.根据权利要求17所述的方法,其中所述小鼠在其内源性CD8基因座不表达功能性内源性CD8蛋白。

19.根据权利要求17所述的方法,其中所述小鼠还包含人或人源化的 $\beta 2$ 微球蛋白多肽。

20.根据权利要求17所述的方法,其中所述第一和/或第二核苷酸序列包含在所述小鼠的种系中。

21.根据权利要求17所述的方法,其中修饰所述小鼠基因组包括使用编码所述嵌合的人/小鼠CD8 $\alpha$ 多肽的第一核苷酸序列取代在所述小鼠基因组内源性CD8 $\alpha$ 基因座的编码内源性小鼠CD8 $\alpha$ 多肽的核苷酸序列和/或使用编码所述嵌合的人/小鼠CD8 $\beta$ 多肽的第二核苷酸序列取代在所述小鼠基因组内源性CD8 $\beta$ 基因座的编码内源性小鼠CD8 $\beta$ 多肽的核苷酸序列。

22.根据权利要求21所述的方法,其中所述内源性CD8 $\alpha$ 和CD8 $\beta$ 基因座分别编码嵌合的人/小鼠CD8 $\alpha$ 多肽和嵌合的人/小鼠CD8 $\beta$ 多肽,所述嵌合的人/小鼠CD8 $\alpha$ 多肽包含如SEQ ID NO:54所示的氨基酸序列,所述嵌合的人/小鼠CD8 $\beta$ 多肽包含如SEQ ID NO:53所示的氨基酸序列。

## 人源化的T细胞共受体的小鼠

### [0001] 相关申请的交叉引用

[0002] 本申请要求根据35U.S.C.§119(e)的2013年10月15日提交的美国临时专利申请序列号61/890,915和2013年2月20提交的美国临时专利申请序列号61/766,762的优先权,其全部内容在此参考并入。

### [0003] 序列表

[0004] 本申请的说明书中引用了以电子形式于2014年2月20日提交的序列表,其为文件名为“2010794-0441\_ST25”的ascii.txt文件。该.txt文件于2014年2月13日生成,其大小为47kb。

### 发明领域

[0005] 本发明涉及非人动物(例如,啮齿动物,例如小鼠或大鼠),所述非人动物经基因工程改造以表达人源化的T细胞共受体。本发明涉及经基因工程改造以表达人源化的CD4或CD8共受体的非人动物,以及表达其的胚胎、组织和细胞。本发明还涉及经工程改造以共表达人源化的CD4共受体和人源化的主要组织相容性复合物(MHC) II的非人动物。本发明还涉及经工程改造以共表达人源化的CD8共受体和人源化的MHC I的非人动物。本发明还提供了制备经基因工程改造的动物的方法,所述动物表达人源化的T细胞共受体(例如,人源化的CD4或CD8)。本发明还提供了使用表达人源化的T细胞共受体的所述经基因工程改造的动物用于研发人类疗剂的方法。

### [0006] 发明背景

[0007] 在适应性免疫应答中,外来抗原被B淋巴细胞和T淋巴细胞上的受体分子(例如,免疫球蛋白和T细胞受体或TCR)所识别。这些外来抗原被特定蛋白作为肽片段递呈至细胞表面上,通常将所述特定蛋白称为主要组织相容性复合物(MHC)分子。在T细胞介导的应答过程中,由MHC分子递呈的抗原被T细胞受体所识别。然而,有效的免疫应答除了需要MHC-抗原复合物的T细胞受体识别以外,还需要T细胞共受体分子(例如,CD4或CD8)与MHC恒定部分的结合。

[0008] T细胞分为几个种类,包括辅助T细胞和细胞毒性T细胞。辅助T细胞表达共受体CD4并且识别与MHC II分子结合的抗原。CD4+T细胞活化免疫系统中的其他效应细胞,例如活化表达MHC II的B细胞以产生抗体,活化表达MHC II的巨噬细胞以破坏病原体等。CD4和T细胞受体与相同MHC II递呈的外来抗原的结合使得T细胞对该抗原具有更加显著的敏感性。

[0009] 而相反的是,细胞毒性T细胞(CTL)表达共受体CD8并且识别与MHC I分子结合的外来抗原。CTL专门地杀伤携带被其自身膜结合TCR所识别的MHC I结合肽的任意细胞。当细胞展示出来源于非正常存在的(例如,来源于病毒、肿瘤或其他非自身来源)细胞蛋白的肽时,此类肽被CTL所识别,CTL被活化并且杀伤展示出所述肽的细胞。与CD4类似,CD8的结合使得CTL对MHC I递呈的抗原更加敏感。

[0010] 由于耐受机制使得并非所有的抗原均激发T细胞活化。然而,在一些疾病中(例如,癌症、自身免疫性疾病)来源于自身蛋白的肽成为免疫系统细胞成分的靶点,这导致递呈此

类肽的细胞被破坏。其对于识别具有临床意义的抗原(例如,与多种类型的癌症相关的抗原)方面具有显著的推动作用。然而,为了改进对在人类T细胞中激发适宜应答的肽的鉴定和选择,特别是具有临床意义的抗原的肽,仍需要模拟人免疫系统的体内和体外系统。因此,需要能够展示出人免疫系统成分的生物系统(例如,基因修饰的非人动物和细胞)。

#### [0011] 发明概述

[0012] 本申请提供了包含非人细胞的非人动物,所述非人细胞表达在细胞免疫应答中具有功能的人的或人源化的分子。本申请还提供了编码人或人源化的T细胞共受体(例如,CD4和/或CD8)蛋白的人源化的啮齿动物基因座。本申请还提供了表达人或人源化的T细胞共受体(例如,CD4和/或CD8)蛋白的人源化的啮齿动物细胞。本申请提供了体内和体外系统,所述系统包含人源化的啮齿动物细胞,其中所述啮齿动物细胞表达一种或多种人或人源化的免疫系统分子。

[0013] 本申请提供一种基因修饰的非人动物,所述非人动物在其基因组中包含编码人或人源化的T细胞共受体多肽的核苷酸序列。在不同实施方式中,本申请提供了一种基因修饰的非人动物,所述非人动物包含编码嵌合的人/非人T细胞共受体多肽的核苷酸序列。在一个实施方式中,所述核苷酸序列存在于内源性T细胞共受体基因座。在一个实施方式中,所述嵌合的T细胞共受体多肽的人部分包含全部或基本上全部的人T细胞共受体的胞外域,和所述非人动物表达功能性的嵌合的T细胞共受体多肽。在一个实施方式中,所述嵌合的T细胞共受体多肽的非人部分至少包含非人T细胞共受体的跨膜和胞质结构域,和所述非人动物表达功能性的嵌合的T细胞共受体多肽。在本发明的一个方面,所述嵌合的T细胞共受体多肽仅在所述非人动物的T细胞上表达,例如其不在所述非人动物的B细胞上表达。在一个方面,所述动物不表达来自其内源性非人T细胞共受体基因座的功能性非人T细胞共受体。在本发明的一个方面,所述嵌合的T细胞共受体多肽包含在所述非人动物的种系中。在一个方面,所述动物在内源性T细胞共受体基因座包含编码所述嵌合的T细胞共受体多肽的核苷酸序列的一个或两个拷贝;因此,所述动物针对编码嵌合的T细胞共受体多肽的所述核苷酸序列可以是杂合的或纯合的。

[0014] 在一个实施方式中,所述T细胞共受体是CD4。因此,在一个方面,本发明提供了一种基因修饰的非人动物,所述非人动物包含编码嵌合的人/或非人CD4多肽的核苷酸序列。在一个实施方式中,所述核苷酸序列存在于内源性CD4基因座。在一个实施方式中,所述动物是啮齿动物,例如小鼠或大鼠。因此,在一个实施方式中,本发明提供了一种基因修饰的小鼠,所述小鼠在其内源性CD4基因座包含编码嵌合的人/小鼠CD4多肽的核苷酸序列,其中所述嵌合的CD4多肽的人源部分包含全部或基本上全部的人CD4多肽的胞外域,其中所述嵌合的CD4多肽的小鼠部分至少包含小鼠CD4多肽的跨膜和胞质结构域,和其中所述小鼠表达功能性的嵌合的人/小鼠CD4。在一个实施方式中,本申请提供了一种基因修饰的小鼠,所述小鼠在其内源性CD4基因座包含编码嵌合的人/小鼠CD4多肽的核苷酸序列,其中所述嵌合的多肽的人源部分至少包含全部或基本上全部的人CD4多肽的结构域D1-D3,其中所述嵌合的多肽的小鼠部分至少包含小鼠CD4的跨膜和胞质结构域,和其中所述小鼠表达功能性的嵌合的人/小鼠CD4。在一个方面,所述小鼠不表达来自其内源性小鼠CD4基因座的功能性内源性小鼠CD4。在一个实施方式中,编码嵌合的人/小鼠CD4多肽的所述核苷酸序列与内源性小鼠启动子和调控序列可操作地连接。因此,在一个实施方式中,所述小鼠在CD8谱系的B细

胞或T细胞上不表达所述嵌合的CD4蛋白。在一个实施方式中,所述嵌合的CD4蛋白的人源部分包含如SEQ ID NO:57所示的氨基酸序列。在一个实施方式中,所述嵌合的人/小鼠CD4多肽如SEQ ID NO:4所示。

[0015] 在一个方面,包含本申请所述的嵌合的CD4多肽的所述基因修饰的非人动物,例如所述基因修饰的小鼠,还包含人或人源化的MHC II蛋白,其中所述MHC II蛋白包含人MHC II  $\alpha$ 多肽的胞外域和人MHC II  $\beta$ 多肽的胞外域。在一个方面,所述动物包含人源化的MHC II蛋白。在一个实施方式中,所述动物是小鼠并且所述小鼠在内源性MHC II基因座包含(1) 编码嵌合的人/小鼠MHC II  $\alpha$ 多肽的核苷酸序列,其中所述MHC II  $\alpha$ 多肽的人源部分包含人MHC II  $\alpha$ 的胞外域,和内源性小鼠MHC II  $\alpha$ 多肽的跨膜和胞质结构域,以及(2) 编码嵌合的人/小鼠MHC II  $\beta$ 多肽的核苷酸序列,其中所述MHC II  $\beta$ 多肽的人源部分包含人MHC II  $\beta$ 的胞外域,和内源性小鼠MHC II  $\beta$ 多肽的跨膜和胞质结构域。对包含编码嵌合的人/非人(例如人/小鼠)MHC II核苷酸序列的基因修饰的非人动物(例如小鼠)更加详细的描述请参见美国专利申请号13/661,116和13/793,935,其通过引用整体并入本申请。在一个实施方式中,表达人源化的CD4和/或MHC II蛋白的动物通过取代分别位于CD4和/或MHC II基因座的内源性非人(例如小鼠)CD4和/或MHC II基因的部分产生。

[0016] 因此,本申请还提供了一种修饰非人动物(例如啮齿动物,例如小鼠)的CD4基因座以表达嵌合的人/小鼠CD4多肽的方法,所述方法包括使用编码嵌合的人/小鼠CD4多肽的核苷酸序列在内源性CD4基因座取代编码内源性非人(例如小鼠)CD4多肽的核苷酸序列。在一个实施方式中,所述嵌合的人/非人(例如人/啮齿动物,例如人/小鼠)CD4多肽包含至少全部或基本上全部的人CD4多肽的结构域D1-D3和内源性非人(例如啮齿动物,例如小鼠)CD4多肽的至少跨膜和胞质结构域。在一个实施方式中,所表达的嵌合的人/小鼠CD4如SEQ ID NO:4所示。

[0017] 在另一个实施方式中,所述T细胞共受体是CD8。因此,在一个方面,本发明提供了一种基因修饰的非人动物,所述非人动物包含编码嵌合的人/非人CD8多肽的核苷酸序列,例如嵌合的人/非人CD8 $\alpha$ 和/或CD8 $\beta$ 多肽。在一个实施方式中,所述核苷酸序列位于内源性CD8基因座。在一个实施方式中,所述动物是啮齿动物,例如小鼠或大鼠。因此,在一个实施方式中,本申请提供了一种基因修饰的小鼠,所述小鼠在其内源性CD8基因座(例如内源性CD8 $\alpha$ 和/或CD8 $\beta$ 基因座)包含编码嵌合的人/小鼠CD8 $\alpha$ 多肽的第一核苷酸序列和编码嵌合的人/小鼠CD8 $\beta$ 多肽的第二核苷酸序列,其中所述第一核苷酸序列包含编码全部或基本上全部的人CD8 $\alpha$ 多肽的胞外域和小鼠CD8 $\alpha$ 多肽至少跨膜和胞质结构域的序列,和其中所述第二核苷酸序列包含编码全部或基本上全部的人CD8 $\beta$ 多肽的胞外域和小鼠CD8 $\beta$ 多肽至少跨膜和胞质结构域的序列,其中所述小鼠表达功能性的嵌合的人/小鼠CD8蛋白。在一个方面,所述小鼠不表达来自其内源性小鼠CD8基因座的功能性的内源性的CD8多肽。在一个实施方式中,所述第一核苷酸序列与内源性小鼠CD8 $\alpha$ 启动子和调控序列可操作地连接,和所述第二核苷酸序列与内源性小鼠CD8 $\beta$ 启动子和调控序列可操作地连接。因此,在一个实施方式中,所述小鼠在CD4谱系的B细胞或T细胞上不表达嵌合的CD8蛋白。在一个实施方式中,所述嵌合的CD8 $\alpha$ 和/或 $\beta$ 多肽的人源部分包含人CD8 $\alpha$ 和/或 $\beta$ 多肽的免疫球蛋白V-样结构域。在一个实施方式中,所述嵌合的人/小鼠CD8 $\alpha$ 多肽的人源部分包含如SEQ ID NO:59所示的氨基酸序列。在一个实施方式中,所述嵌合的人/小鼠CD8 $\beta$ 多肽的人源部分包含如SEQ ID NO:

58所示的氨基酸序列。在一个实施方式中,所述嵌合的人/小鼠CD8 $\alpha$ 多肽如SEQ ID NO:54所示,和所述嵌合的人/小鼠CD8 $\beta$ 多肽如SEQ ID NO:53所示。

[0018] 在一个方面,包含本申请所述的嵌合的CD8 $\alpha$ 和/或 $\beta$ 多肽的基因修饰的非人动物(例如基因修饰的小鼠)还包含人或人源化的MHC I蛋白,其中所述MHC I蛋白包含人MHC I多肽的胞外域。在一个方面,所述动物包含人源化的MHC I复合物。因此,所述动物可以包含人源化的MHC I蛋白和人或人源化的 $\beta$ 2微球蛋白多肽。在一个实施方式中,所述动物是小鼠和所述小鼠在内源性MHC I基因座包含编码嵌合的人/小鼠MHC I多肽的核苷酸序列,其中所述MHC I多肽的人源部分包含人MHC I多肽的胞外域,和内源性小鼠MHC I多肽的跨膜和胞质结构域;和所述动物在内源性 $\beta$ 2微球蛋白基因座还包含编码人或人源化的 $\beta$ 2微球蛋白的核苷酸序列。对包含编码嵌合的人/非人(例如人/小鼠)MHC I和 $\beta$ 2微球蛋白的核苷酸序列的基因修饰的非人动物(例如小鼠)更加详细的描述请参见美国专利申请号13/661,159和13/793,812,其通过引用整体并入本申请。在一个实施方式中,表达人源化的CD8、MHC I和/或 $\beta$ 2微球蛋白的动物通过取代分别位于CD8、MHC I和/或 $\beta$ 2微球蛋白基因座的内源性非人(例如小鼠)CD8、MHC I和/或 $\beta$ 2微球蛋白基因的部分产生。

[0019] 因此,本申请还提供了一种修饰非人动物(例如啮齿动物,例如小鼠)的CD8基因座以表达嵌合的人/小鼠CD8多肽的方法,所述方法包括使用编码嵌合的人/小鼠CD8多肽的核苷酸序列取代在内源性CD8基因座的编码内源性非人(例如小鼠)CD8多肽的核苷酸序列。在一个方面,所述CD8多肽选自下组:CD8 $\alpha$ 、CD8 $\beta$ 及其组合。在一个实施方式中,所述嵌合的人/非人(例如人/啮齿动物,例如人/小鼠)CD8多肽(CD8 $\alpha$ 和/或CD8 $\beta$ )包含全部或基本上全部的人CD8多肽的胞外域和内源性非人(例如啮齿动物,例如小鼠)CD8多肽的至少跨膜和胞质结构域。

[0020] 本申请还提供了来源于本申请所述的非人动物(例如啮齿动物,例如小鼠或大鼠)的细胞,例如T细胞。本申请还提供了来源于本申请所述的非人动物的组织和胚胎。

[0021] 除非另有明示或者从上下文中是显而易见的,本申请所述的任意实施方式和方面均可以用于彼此之间组合。通过对随后详细描述的回顾,其他实施方式对本领域技术人员将是显而易见的。下述的详细描述包括对本发明不同实施方式的示例性表述,其对所主张本发明不具有限制作用。附图构成本说明书的一部分,其与该描述一起,仅用于说明实施方式,并非用于限制本发明。

#### [0022] 附图简述

[0023] 图1是产生人源化CD4基因座策略的示意图(未按照比例)。首先,紧随信号肽之后的小鼠外显子3-6的序列被所述信号肽下游的人外显子3序列所取代(上图),随后通过限制性消化/连接将人外显子4-6插入人外显子3下游。

[0024] 图2显示了使用抗-人CD4和抗-小鼠CD4抗体对来源于WT小鼠或针对人CD4的小鼠杂合子(1766HET)的脾细胞进行的FACS分析(A);和对来源于WT小鼠vs.1766HET小鼠vs.Jurkat人CD4T细胞系的T细胞进行的FACS分析。

[0025] 图3是通过使用人CD8 $\beta$ 外显子2-3取代小鼠CD8 $\beta$ 外显子2-3产生人源化CD8 $\beta$ 基因座(MAID 1737)策略的示意图(未按照比例)。小鼠外显子序列以实心矩形表示,人外显子序列以散列矩形表示。

[0026] 图4是通过使用人外显子2-3取代小鼠外显子1的一部分和外显子2,保留在外显子

1起始的小鼠先导序列产生人源化CD8a基因座(MAID 1738)策略的示意图(未按照比例)。小鼠外显子序列以实心矩形表示,人外显子序列以散列矩形表示。

[0027] 图5是产生包含编码人源化的CD8b和CD8a基因序列的人源化基因座的顺序靶向策略的示意图(未按照比例)。小鼠外显子序列以实心矩形表示,人外显子序列以散列矩形表示。

[0028] 图6是使用小鼠CD8b、小鼠CD8a、人CD8b或人CD8a抗体对来自于WT小鼠或已除去选择性表达盒的针对人CD8b和CD8a的杂合子小鼠(1739Het,1740Het)的脾细胞进行的FACS分析。

[0029] 图7是使用小鼠CD8b、小鼠CD8a、人CD8b、人CD8a或CD4对来自于WT小鼠或1739HET/1740HET小鼠(针对CD8b和CD8a的杂合子小鼠)的胸腺细胞进行的FACS分析。

#### [0030] 发明详述

##### [0031] 定义

[0032] 本发明提供了表达人源化的T细胞共受体多肽的基因修饰的非人动物(例如小鼠、大鼠、家兔等);包含其的胚胎、细胞和组织;制备其的方法;以及使用其的方法。除非另有定义,本申请中使用的所有术语和短语包括所述术语和短语在现有技术中已有的含义,除非存在相反的明确指示或从该术语或短语使用的上下文中是显而易见的。

[0033] 术语“保守的”当用于描述保守的氨基酸取代时,包括氨基酸残基被带有具有相似化学性质(例如电荷或疏水性)的侧链R基团的其他氨基酸残基取代。可以通过修饰核苷酸序列实现保守的氨基酸取代,以引入编码保守的取代的核苷酸改变。在通常情况下,保守的氨基酸取代将基本上不改变目的蛋白的功能性性质,例如CD4或CD8分别与MHC II或MHC I结合的能力,并且增加TCR对MHC提呈的抗原的敏感性。带有具有相似化学性质的侧链的氨基酸基团的示例包括脂肪族侧链如甘氨酸、丙氨酸、缬氨酸、亮氨酸和异亮氨酸;脂肪族-羟基侧链如丝氨酸和苏氨酸;含有酰胺的侧链如天冬酰胺和谷氨酰胺;芳香侧链如苯丙氨酸、酪氨酸和色氨酸;碱性侧链如赖氨酸、精氨酸和组氨酸;酸性侧链如天冬氨酸和谷氨酸;以及含硫侧链如半胱氨酸和甲硫氨酸。保守的氨基酸取代基团包括例如缬氨酸/亮氨酸/异亮氨酸、苯丙氨酸/酪氨酸、赖氨酸/精氨酸、丙氨酸/缬氨酸、谷氨酸/天冬氨酸和天冬酰胺/谷氨酰胺。在一些实施方式中,保守的氨基酸取代可以是使用丙氨酸取代蛋白中的任意天然残基,如用在例如丙氨酸扫描诱变中。在一些实施方式中,Gonnet等((1992) Exhaustive Matching of the Entire Protein Sequence Database, Science 256:1443-45)披露保守的氨基酸取代使得在PAM250log-似然性矩阵中具有正值,其通过引用并入本申请。在一些实施方式中,所述取代是适度保守的取代,其中所述取代在PAM250log-似然性矩阵中具有非负值。

[0034] 因此,本发明还包括一种基因修饰的非人动物,其基因组包含(例如在内源性基因座)编码人源化的T细胞共受体多肽(例如CD4或CD8多肽)的核苷酸序列,其中所述多肽包含本申请所述的氨基酸序列的保守的氨基酸取代。

[0035] 本领域技术人员将理解除了编码本申请所述的人源化的T细胞共受体多肽的核酸残基以外,由于遗传密码的简并性,其他核酸可以编码本发明的多肽。因此,除了在其基因组中包含编码具有保守的氨基酸取代的T细胞共受体多肽(例如CD4或CD8多肽)的核苷酸序列的基因修饰的非人动物以外,本申请还提供了由于遗传密码的简并性使其基因组中包含

不同于本申请所述核苷酸序列的非人动物。

[0036] 术语“同一性”当与序列结合使用时包括根据本领域公知的能够用于测定核苷酸和/或氨基酸序列同一性的多种不同算法确定的同一性。在本申请所述的一些实施方式中,使用开放缺口罚分为10.0,延伸缺口罚分为0.1的ClustalW v.1.83(slow)比对和使用Gonnet相似性矩阵(MacVector™10.0.2,MacVector Inc.,2008)确定同一性。针对序列的同一性比较序列的长度将取决于特定的序列。在不同实施方式中,通过从其N-末端至其C-末端比较成熟蛋白的序列确定其同一性。在不同实施方式中,当将嵌合的人/非人序列与人序列进行比较时,将所述嵌合的人/非人序列的人部分(而不是非人部分)用于进行针对确定人序列与嵌合的人/非人序列的人部分之间的同一性水平目的的比较(例如将嵌合人/小鼠蛋白的人胞外域与人蛋白的人胞外域进行比较)。

[0037] 术语“同源性”或“同源的”涉及序列例如核苷酸或氨基酸序列时指两条序列根据最佳比对和比较例如至少约75%的核苷酸或氨基酸、例如至少约80%的核苷酸或氨基酸、例如至少约90-95%的核苷酸或氨基酸、例如超过97%的核苷酸或氨基酸是一致的。本领域技术人员将理解对于最佳基因靶向所述靶向构建体应含有与内源性DNA序列具有同源性的臂(“即同源臂”);因此,同源重组可以发生在靶向构建体与被靶向的内源性序列之间。

[0038] 术语“可操作地连接”指并列的,其中这样描述的所述组分与允许其以其意向的方式发挥功能相关。正因如此,编码蛋白的核酸序列可以与调控序列(例如,启动子、增强子、沉默子序列等)可操作地连接以保持正确的转录调控。此外,本发明的嵌合或人源化蛋白的不同部分可以可操作地连接以保持合适的折叠、加工、靶向、表达,以及所述蛋白在细胞中的其他功能性性质。除非另有说明,本发明的嵌合或人源化蛋白的不同结构域彼此之间可操作地连接。

[0039] 术语“取代”涉及基因取代时指在内源性基因座放置外源性遗传物质,从而使用直系同源或同源的核酸序列取代全部或部分内源性基因。如下述实施例中所示,编码小鼠CD4或CD8(CD8 $\alpha$ 和/或CD8 $\beta$ )多肽部分的内源性基因座的核酸序列分别被编码人CD4或CD8(CD8 $\alpha$ 和/或CD8 $\beta$ )多肽部分的核苷酸序列所取代。

[0040] 在本申请中使用的“功能性”例如涉及功能性多肽时指与天然蛋白相关的至少一种生物活性保留正常的多肽。例如,在本发明的一些实施方式中,在内源性基因座的取代(例如在内源性非人CD4或CD8基因座的取代)使得在基因座中不能表达功能性的内源性多肽。

[0041] 本申请下文中描述的针对基因修饰的CD4非人动物的若干方面,例如动物类型;动物品系;细胞类型;筛选、检测和其他方法;使用方法等,将适用于基因工程改造的CD8动物。

[0042] 基因修饰的CD4动物

[0043] 在不同实施方式中,本发明通常地提供了基因修饰的非人动物,所述非人动物在其基因组例如在内源性CD4基因座中包含编码人源化的CD4多肽的核苷酸序列;因此所述动物表达人源化的CD4多肽。

[0044] 人CD4基因位于12号染色体,认为其含有10个外显子。CD4基因编码具有氨基末端疏水性信号序列的蛋白,该信号序列由所述基因的外显子2和3所编码。所述蛋白包含4个免疫球蛋白样结构域,通常称为D1-D4结构域。Maddon等,(1987) Structure and expression of the human and mouse T4 genes(人和小鼠T4基因的结构和表达),

Proc.Natl.Acad.Sci.USA 84:9155-59。据信D1结构域由外显子3(信号肽的下游序列)和外显子4编码,而D2、D3和D4分别由彼此独立的外显子——外显子5、6和7编码。Littman(1987) The Structure of the CD4and CD8Genes(CD4和CD8基因的结构),Ann.Rev.Immunol.5:561-84;Hanna等,(1994) Specific Expression of the Human CD4Gene in Mature CD4+CD8-and Immature CD4+CD8+T cells and in Macrophages of Transgenic Mice(人CD4基因在转基因小鼠成熟的CD4+CD8-和未成熟的CD4+CD8+T细胞中和巨噬细胞中的特异性表达),Mol.Cell.Biol.14(2):1084-94;Maddon等,如上所述。在所述蛋白浓度较高的区域,如T细胞与抗原递呈细胞之间相接触的区域,分子趋向于通过反向D4结构域之间的相互作用形成同型二聚体。Zamoyska(1998) CD4and CD8:modulators of T cell receptor recognition of antigen and of immune responses?(CD4和CD8:T细胞受体识别抗原和免疫应答的调节剂?)Curr.Opin.Immunol.10:82-87;Wu等,(1997) Dimeric association and segmental variability in the structure of human CD4(在人CD4结构中的二聚结合和区段可变性),Nature387:527;Moldovan等,(2002) CD4Dimers Constitute the Functional Component Required for T Cell Activation(CD4二聚体构成T细胞活化所需的功能性成分),J.Immunol.169:6261-68。

[0045] CD4的D1结构域类似于免疫球蛋白的可变(V)结构域,其与D2结构域的一部分一起据信与MHC II结合。Huang等,(1997) Analysis of the contact sites on the CD4 Molecule with Class II MHC Molecule(对CD4分子与II类MHC分子上的接触位点的分子),J.Immunol.158:216-25。反过来,MHC II与T细胞共受体CD4在MHC II $\alpha$ 2和 $\beta$ 2结构域之间连接的疏水性裂隙上发生相互作用。Wang和Reinherz(2002) Structural Basis of T Cell Recognition of Peptides Bound to MHC Molecules(肽与MHC分子结合的T细胞识别的结构基础),Molecular Immunology,38:1039-49。

[0046] CD4共受体的D3和D4结构域据信与TCR-CD3复合物相互作用,当这两个结构域被取代时,CD4不能与TCR结合。Vignali等,(1996) The Two Membrane Proximal Domains of CD4Interact with the T Cell Receptor(CD4与T细胞受体相互作用的两个膜近端结构域),J.Exp.Med.183:2097-2107。CD4分子以二聚体的形式存在,在该分子D4结构域中的残基据信负责CD4的二聚化。Moldovan等,(2002) CD4 Dimers Constitute the Functional Components Required for T Cell Activation(CD4二聚体构成T细胞活化所需的功能性成分),J.Immunol.169:6261-68。

[0047] CD4基因的外显子8编码跨膜结构域,而该基因其余的部分编码胞质结构域。CD4的胞质结构域具有多种独特的功能。例如,CD4的胞质结构域募集酪氨酸激酶Lck。Lck是Src家族激酶,其与CD4和CD8胞质结构域结合并且所述共受体和TCR同时与相同的MHC结合导致CD3和TCR复合物 $\xi$ 链酪氨酸磷酸化的增加,其反过来会导致在T细胞活化中具有作用的其他因子的募集。Itano及其同事通过设计和检测包含CD8胞外域和CD4胞质尾的杂交蛋白在转基因小鼠中的表达提出CD4的胞质尾也能够促进CD4+CD8+T细胞分化成CD4+谱系。Itano等,(1996) The Cytoplasmic Domain of CD4 Promotes the Development of CD4 Lineage T Cells(CD4的胞质结构域促进CD4谱系T细胞的发育),J.Exp.Med.183:731-41。所述杂交蛋白的表达导致MHC I特异性CD4谱系T细胞的发育。同上。

[0048] CD4共受体似乎是HIV病毒的主要受体,而CD4+T细胞耗竭是疾病进展的指示剂。

CD4的胞质尾似乎是在HIV诱导的凋亡中将凋亡信号传送至CD4+T细胞所必需的。特别地,CD4与Lck之间的相互作用显示出在这些细胞中加强HIV诱导的凋亡。Corbeil等,(1996) HIV-induced Apoptosis Requires the CD4 Receptor Cytoplasmic Tail and Is Accelerated by Interaction of CD4 with p56lck (HIV诱导的凋亡需要CD4受体胞质尾并且其被CD4与p56lck之间的相互作用所加速),J.Exp.Med.183:39-48。

[0049] T细胞在胸腺中的发育过程为从不成熟的CD4-/CD8- (双重阴性或DN) 胸腺细胞变成CD4+/CD8+ (双重阳性或DP) 胸腺细胞,其最终经历阳性选择成为CD4+或CD8+ (单一阳性或SP) T细胞。通过MHC I限制性TCR接收信号的DP胸腺细胞分化成CD8+T细胞,而通过MHC II限制性TCR接收信号的DP胸腺细胞分化成CD4+T细胞。被DP细胞所接收而导致其分化成CD4+或CD8+T细胞的信号已成为研究非常广泛的主题。已提出了针对CD4/CD8谱系选择的多种模型,Singer等,(2008) Lineage fate and intense debate: myths, models and mechanisms of CD4-versus CD8-lineage choice (谱系命运和激烈的争论: CD4-对比CD8-谱系选择的猜测、模型和机制), Nat.Rev.Immunol.8:788-801. 对其进行了综述。

[0050] 阳性选择导致的特定T细胞共受体的失活是转录调控的产物。对于CD4而言,已发现其增强子位于CD4外显子1上游13kb处,其在CD4+和CD8+T细胞中上调CD4的表达。Killeen等,(1993) Regulated expression of human CD4 rescues helper T cell development in mice lacking expression of endogenous CD4, (人CD4受调控的表达在缺乏内源性CD4表达的小鼠中挽救辅助T细胞的发育) EMBO J.12:1547-53。位于鼠源性CD4基因的第一内含子中的顺式作用转录沉默子的功能为在CD4+T细胞以外的细胞中沉默CD4的表达。Siu等,(1994) A transcriptional silencer control the developmental expression of the CD4 gene (转录沉默子控制CD4基因的发育表达), EMBO J.13:3570-3579。

[0051] 由于在此前研发的表达人CD4的转基因小鼠的若干品系中控制CD4品系选择的重要转录调节子(例如,启动子、增强子、沉默子等) 缺失,使得这些小鼠不能进行正常的T细胞谱系发育,并且产生表达CD4的CD4+T细胞以外的免疫细胞。参见例如, Law等,(1994) Human CD4 Restores Normal T Cell Development and Function in Mice Deficient in CD4 (人CD4使得CD4缺陷小鼠恢复正常的T细胞发育和功能), J.Exp.Med.179:1233-42 (在CD8+T细胞和B细胞中表达CD4); Fugger等,(1994) Expression of HLA-DR4 and human CD4 transgenes in mice determines the variable region  $\beta$ -chain T-cell repertoire and mediates an HLA-D-restricted immune response (HLA-DR4和人CD4转基因在小鼠中的表达决定可变区 $\beta$ -链T-细胞库和介导HLA-D-受限制的免疫应答), Proc.Natl.Acad.Sci.USA,91:6151-55 (CD4在所有CD3+淋巴细胞和B细胞上表达)。因此,在一个实施方式中,开发一种基因修饰的动物使其保留内源性小鼠启动子和其他调控元件以使得所述动物产生能够经历正常的T细胞发育和谱系选择的T细胞可能是有益的。

[0052] 因此,在不同实施方式中,本发明提供了一种基因修饰的非人动物,所述非人动物例如在其内源性T细胞共受体基因座(例如CD4基因座) 包含编码嵌合的人/非人T细胞共受体多肽的核苷酸序列。在一个实施方式中,所述嵌合多肽的人部分包含全部或基本上全部的人T细胞共受体胞外域。在一个实施方式中,所述嵌合多肽的非人部分包含非人T细胞共受体的跨膜和胞质结构域。在一个实施方式中,所述非人动物表达功能性的嵌合的T细胞共受体多肽。因此,在一个方面,本发明提供了一种基因修饰的非人动物,所述非人动物在其

内源性CD4基因座包含编码嵌合的人/非人CD4多肽的核苷酸序列,其中所述嵌合多肽的人部分包含全部或基本上全部的人CD4的胞外域,其中非人部分至少包含非人CD4的跨膜和胞质结构域,和其中所述动物表达功能性的嵌合的CD4多肽。在一个方面,所述非人动物仅表达人源化的CD4多肽,即嵌合的人/非人CD4多肽,并且不表达来自其内源性CD4基因座的功能性的内源性非人CD4蛋白。

[0053] 在一个实施方式中,所述嵌合的人/非人CD4多肽的人部分包含全部或基本上全部的人CD4多肽的胞外域。在另一个实施方式中,所述嵌合的人/非人CD4多肽至少包含全部或基本上全部的所述人CD4多肽的MHC II结合结构域(例如D1和D2结构域的大部分);在一个实施方式中,所述嵌合的人/非人CD4多肽的人部分包含全部或基本上全部的所述人CD4多肽的D1、D2和D3结构域;在又一个实施方式中,所述嵌合的人/非人CD4多肽的人部分包含全部或基本上全部的CD4免疫球蛋白样结构域,例如称为D1、D2、D3和D4的结构域。在又一个实施方式中,所述嵌合的人/非人CD4多肽的人部分在其人部分中包含全部或基本上全部的负责与MHC II和/或T细胞受体的胞外部分相互作用的人CD4序列。在又一个实施方式中,所述嵌合的人/非人CD4多肽的人部分包含全部或基本上全部的负责与MHC II和/或T细胞受体的可变结构域相互作用的人CD4的胞外部分。因此,在一个实施方式中,编码所述嵌合的CD4多肽的人部分的核苷酸序列包含全部或基本上全部的人CD4结构域D1-D2的编码序列(例如,人CD4基因外显子3的一部分和外显子4-5);在另一个实施方式中,其包含全部或基本上全部的人CD4D1-D3的编码序列(例如,人CD4外显子3的一部分和外显子4-6)。因此,在一个实施方式中,编码嵌合的人/非人CD4的核苷酸序列包含编码全部或基本上全部人CD4的D1-D3结构域的核苷酸序列。在另一个实施方式中,编码所述嵌合的CD4多肽的人部分的核苷酸序列包含人CD4基因的D1-D4结构域的编码序列。在另一个实施方式中,所述核苷酸序列可以包含编码小鼠CD4信号肽的核苷酸序列,例如由小鼠基因的外显子2-3的部分编码的区域。在另一个实施方式中,所述核苷酸序列可以包含编码人CD4信号肽的核苷酸序列。在一个实施方式中,所述嵌合的人/非人CD4多肽包含如SEQ ID NO:4所示的氨基酸序列,和所述嵌合多肽的人部分大致跨SEQ ID NO:4的氨基酸27-319(其单独以SEQ ID NO:57表示)。

[0054] 在一个实施方式中,所述非人动物表达嵌合的人/非人CD4多肽序列。在一个实施方式中,所述嵌合的CD4序列的人部分包含一个或多个保守的或非保守的修饰。

[0055] 在一个方面,本申请提供了一种表达人CD4序列的非人动物,其中所述人CD4序列与人CD4序列具有至少约85%、90%、95%、96%、97%、98%或99%的同一性。在一个特定实施方式中,所述人CD4序列与实施例中描述的人CD4序列具有至少约90%、95%、96%、97%、98%或99%的同一性。在一个实施方式中,所述人CD4序列包含一个或多个保守的取代。在一个实施方式中,所述人CD4序列包含一个或多个非保守的取代。

[0056] 在一些实施方式中,一部分例如所述嵌合CD4的人部分可以包含基本上全部的本申请所示的序列(例如,基本上全部的本申请所示的蛋白结构域)。基本上全部序列通常包括据信代表所述蛋白特定部分(例如特定功能性结构域等)的85%、90%、95%、96%、97%、98%或99%的氨基酸。本领域技术人员将理解可以根据所使用的比对和结构域预测方法对功能性结构域的边界进行微小改变。

[0057] 在一个方面,所述嵌合的人/非人CD4多肽的非人部分至少包含所述非人CD4多肽的跨膜和胞质结构域。由于CD4胞质结构域具有重要的功能,因此在基因工程改造的动物中

保留内源性非人(例如小鼠)序列确保保持共受体具有适当的细胞内信号和其他功能。在一个实施方式中,所述非人动物是小鼠,和所述非人CD4多肽是小鼠CD4多肽。尽管在实施例中描述了特定的小鼠CD4序列,但是来源于其的任意适宜序列例如包含保守的/非保守的氨基酸取代的序列均包括在本申请中。在一个实施方式中,所述嵌合的CD4共受体的非人部分包含未被人源化的内源性CD4的任意序列。

[0058] 本申请所述的非人动物在其内源性基因座中可以包含编码嵌合的人/非人CD4多肽的核苷酸序列。在一个方面,其导致内源性CD4基因的一部分被编码人CD4多肽一部分的核苷酸序列取代。在一个实施方式中,此类取代是编码例如全部或基本上全部的非人CD4的胞外域的内源性核苷酸序列,例如编码至少全部或基本上全部的非人CD4的第一免疫球蛋白样结构域(即D1)的序列(例如编码全部或基本上全部的非人CD4的结构域D1-D2的序列,例如编码全部或基本上全部的非人CD4的结构域D1-D3的序列,编码全部或基本上全部的非人CD4的结构域D1-D4的序列)被编码其的人核苷酸序列取代。在一个实施方式中,所述取代产生包含负责与MHC II和/或T细胞受体胞外部分相互作用的人CD4序列的嵌合蛋白。在又一个实施方式中,所述取代产生包含负责与MHC和/或T细胞受体的可变结构域相互作用的人CD4序列的嵌合蛋白。在一个实施方式中,所述取代不包含编码非人CD多肽的至少跨膜和胞质结构域的CD4序列的取代。因此,在一个方面,所述非人动物表达来自内源性非人CD4基因座的嵌合的人/非人CD4多肽。在又一个实施方式中,所述取代产生包含如SEQ ID NO:4所示的多肽序列的蛋白。

[0059] 在一个实施方式中,本申请提供了本申请所述的嵌合的人/非人CD4基因座(例如嵌合的人/啮齿动物CD4基因座,例如嵌合的人/小鼠CD4基因座)的核苷酸序列。在一个方面,由于所述嵌合的人/非人(例如人/啮齿动物,例如人/小鼠)CD4序列被置于内源性非人(例如啮齿动物,例如小鼠)CD4基因座,因此其保留了位于第一个CD4外显子上游的CD4增强子元件。在一个实施方式中,在内源性非人(例如啮齿动物,例如小鼠)CD4基因座的所述取代包含例如编码CD4多肽D1外显子3的一部分以及D1其余部分和D2-D3的外显子4-6的取代;因此,在一个方面,所述嵌合的CD4基因座保留了位于所述非人(例如小鼠)CD4基因内含子1中的顺式作用沉默子。因此,在一个实施方式中,所述嵌合的基因座保留了内源性非人(例如啮齿动物,例如小鼠)CD4启动子和调控元件。在另一个实施方式中,所述嵌合的基因座可以包含人启动子和调控元件的程度为使得适当地实现CD4表达、CD4+T细胞发育、CD4谱系选择和共受体发挥功能。因此,在一些方面,本发明的动物包含不改变正确的谱系选择和T细胞发育的基因修饰。在一个方面,本发明的动物(例如啮齿动物,例如小鼠)在正常表达CD4的细胞以外的免疫细胞上不表达嵌合的CD4多肽。在一个方面,动物在B细胞或CD8+SP T细胞上不表达CD4。在一个实施方式中,所述取代保留了能够适当地在空间和时间上调控CD4表达的元件。

[0060] 本发明基因修饰的非人动物可以选自下组:小鼠、大鼠、家兔、猪、牛(例如奶牛、公牛、水牛)、鹿、绵羊、山羊、鸡、猫、狗、雪貂、灵长类动物(例如狨猴、猕猴)。对于不易于获得适宜的基因修饰的ES细胞的非人动物而言,可以使用其他方法制备包含基因修饰的非人动物。这类方法包括例如修饰非ES细胞基因组(例如成纤维细胞或诱导多能干细胞)以及使用核转移以便将经修饰的基因组转移至适宜细胞例如卵母细胞,并且在适宜条件下在非人动物中使经修饰的细胞(例如经修饰的卵细胞)妊娠以形成胚胎。

[0061] 在一个方面,所述非人动物是哺乳动物。在一个方面,所述非人动物是小型哺乳动物,例如跳鼠科或鼠总科超家族。在一个实施方式中,所述基因修饰的动物是啮齿动物。在一个实施方式中,所述啮齿动物选自小鼠、大鼠和仓鼠。在一个实施方式中,所述啮齿动物选自鼠总科超家族。在一个实施方式中,所述基因修饰的动物来自选自丽仓鼠科(例如小鼠样仓鼠)、仓鼠科(例如仓鼠、新世界大鼠和小鼠、田鼠)、鼠总科(真小鼠和大鼠、沙鼠、刺毛鼠、冠毛大鼠)、马岛鼠科(登山小鼠、岩小鼠、有尾大鼠、马达加斯加大鼠和小鼠)、刺睡鼠科(例如多刺睡鼠)和鼯形鼠科(例如摩尔大鼠、竹大鼠和鼯鼠)家族。在一个特定实施方式中,所述基因修饰的啮齿动物选自真小鼠或大鼠(鼠总科)、沙鼠、刺毛鼠和冠毛大鼠。在一个实施方式中,所述基因修饰的小鼠来自鼠科家族成员。在一个实施方式中,所述动物是啮齿动物。在一个特定实施方式中,所述啮齿动物选自小鼠和大鼠。在一个实施方式中,所述非人动物是小鼠。

[0062] 在一个特定实施方式中,所述非人动物是啮齿动物,其为选自C57BL/A、C57BL/An、C57BL/GrFa、C57BL/KaLwN、C57BL/6、C57BL/6J、C57BL/6ByJ、C57BL/6NJ、C57BL/10、C57BL/10ScSn、C57BL/10Cr和C57BL/01a的C57BL品系的小鼠。在另一个实施方式中,所述小鼠是选自由以下品系构成的组的129品系:129P1、129P2、129P3、129X1、129S1(例如129S1/SV、129S1/SvIm)、129S2、129S4、129S5、129S9/SvEvH、129S6(129/SvEvTac)、129S7、129S8、129T1、129T2(参见例如Festing等(1999) Revised nomenclature for strain 129mice(对129小鼠品系修订的命名法), Mammalian Genome 10:836, 亦参见Auerbach等(2000) Establishment and Chimera Analysis of 129/SvEv-and C57BL/6-Derived Mouse Embryonic Stem Cell Lines(129/SvEv-和C57BL/6-来源的小鼠胚胎干细胞系的建立和嵌合体分析))。在一个特定实施方式中,所述基因修饰的小鼠是前述129品系和前述C57BL/6品系的混合。在另一个特定实施方式中,所述小鼠是前述129品系的混合或前述BL/6品系的混合。在一个特定实施方式中,所述混合的129品系是129S6(129/SvEvTac)品系。在另一个实施方式中,所述小鼠是BALB品系,例如BALB/c品系。在又另一个实施方式中,所述小鼠是BALB品系和另一个前述品系的混合。

[0063] 在一个实施方式中,所述非人动物是大鼠。在一个实施方式中,所述大鼠选自Wistar大鼠、LEA品系、Sprague Dawley品系、Fischer品系、F344、F6和黑刺鼠。在一个实施方式中,所述大鼠品系是两种或多种选自下组的品系的混合:Wistar、LEA、Sprague Dawley、Fischer、F344、F6和黑刺鼠。

[0064] 因此,在一个实施方式中,本发明提供了一种基因修饰的小鼠,所述小鼠在其基因组例如在其内源性CD4基因座中包含编码嵌合的人/小鼠CD4多肽的核苷酸序列,其中所述嵌合多肽的小鼠部分包含小鼠CD4多肽的至少跨膜和胞质结构域,和其中所述小鼠表达嵌合的人/小鼠CD4。在一个实施方式中,所述嵌合多肽的人部分至少包含全部或基本上全部的人CD4多肽的胞外域。在一个实施方式中,所述嵌合多肽的人部分至少包含全部或基本上全部的人CD4蛋白的D1结构域。在一个实施方式中,所述嵌合多肽的人部分至少包含全部或基本上全部的人CD4蛋白的D1-D2结构域,例如至少全部或基本上全部的人CD4蛋白的D1-D3结构域,例如全部或基本上全部的人CD4蛋白的D1-D4结构域。因此,在一个实施方式中,所述小鼠在其内源性CD4基因座包含至少含有全部或基本上全部的人CD4基因的外显子4、5和6的核苷酸序列,例如编码人CD4D1结构域的一部分的人CD4基因的外显子3和人CD4基因的

外显子4-6的序列。在一个实施方式中,所述小鼠在内源性CD4基因座包含嵌合的人/小鼠CD4,其包含负责与MHC II和/或T细胞受体的胞外部分相互作用的人CD4序列。在另一个实施方式中,所述小鼠在内源性CD4基因座包含嵌合的人/小鼠CD4,其包含负责与MHC II和/或T细胞受体的可变结构域相互作用的人CD4序列。在一个实施方式中,所述核苷酸序列包含编码小鼠CD4信号肽的序列。在一个实施方式中,所述小鼠包含编码小鼠CD4胞外域的核苷酸序列被编码人CD4胞外域的核苷酸序列的取代。在另一个实施方式中,所述小鼠包含编码至少全部或基本上全部的小鼠CD4D1结构域,例如编码至少全部或基本上全部的小鼠CD4 D1-D2结构域的核苷酸序列,例如编码至少全部或基本上全部的小鼠CD4 D1-D3结构域的核苷酸序列,被编码其的人核苷酸序列的取代。在一个实施方式中,所述小鼠不表达来自其内源性小鼠CD4基因座的功能性内源性小鼠CD4。在一个实施方式中,本申请所述的小鼠在小鼠种系中包含嵌合的人/小鼠CD4核苷酸序列。在一个实施方式中,所述小鼠保留了未被人源化的任意内源性序列,例如在所述实施方式中其中所述小鼠包含编码全部或基本上全部D1-D3结构域的核苷酸序列的取代,所述小鼠保留了编码小鼠CD4 D4结构域的内源性核苷酸序列和编码小鼠CD4跨膜和胞质结构域的核苷酸序列。

[0065] 在一个方面,表达嵌合的人/小鼠CD4蛋白的小鼠保留了小鼠CD4启动子和调控序列,例如在小鼠中编码与内源性小鼠CD4启动子和调控序列可操作地连接的嵌合的人/小鼠CD4的核苷酸序列。在一个方面,在本发明的基因工程改造的动物中保留的这些小鼠调控序列包括调控所述嵌合蛋白在T细胞发育的适当阶段表达的序列。因此,在一个方面,所述小鼠在CD8谱系的B细胞或T细胞上不表达嵌合的CD4。在一个方面,所述小鼠还在任意细胞类型上均不表达嵌合的CD4,例如任意的免疫细胞类型,其正常情况下不表达内源性CD4。

[0066] 在不同实施方式中,如本申请所述的表达来自嵌合的CD4基因座的功能性嵌合的CD4蛋白的非人动物(例如啮齿动物,例如小鼠或大鼠)在细胞表面上展示了嵌合蛋白,例如T细胞表面。在一个实施方式中,所述非人动物在细胞表面上表达的嵌合的CD4蛋白的细胞分布与在人中观察到的一致。在一个方面,本发明的CD4蛋白能够与在另一种细胞的表面上表达的MHC II蛋白相互作用,例如抗原提呈细胞(APC)。

[0067] 在一个实施方式中,本发明的非人动物(例如啮齿动物,例如小鼠)还包含编码人或人源化MHC II蛋白的核苷酸序列,这样在所述动物的T细胞表面上表达的所述嵌合的CD4蛋白能够与在另一种细胞(例如抗原递呈细胞)表面上表达的人或人源化的MHC II相互作用。在一个实施方式中,所述MHC II蛋白包含人MHC II  $\alpha$ 多肽的胞外域和人MHC II  $\beta$ 多肽的胞外域。在2012年10月26日提交的美国专利申请号13/661,116和2013年3月11日提交的美国专利申请号13/793,935中描述了示例性的表达人或人源化的MHC II多肽的基因修饰的动物,其通过引用整体并入本申请。因此,在一个实施方式中,包含本申请所述的嵌合的CD4蛋白的动物还可以包含人源化的MHC II蛋白,其中所述人源化的MHC II蛋白包含:(1)包含人MHC II  $\alpha$ 的胞外域以及内源性(例如小鼠)MHC II的跨膜和胞质结构域的人源化的MHC II  $\alpha$ 多肽,其中所述人MHC II  $\alpha$ 胞外域包含人MHC II  $\alpha$ 的 $\alpha 1$ 和 $\alpha 2$ 结构域和(2)包含人MHC II  $\beta$ 的胞外域以及内源性(例如小鼠)MHC II的跨膜和胞质结构域的人源化的MHC II  $\beta$ 多肽,其中所述人MHC II  $\beta$ 胞外域包含人MHC II  $\beta$ 的 $\beta 1$ 和 $\beta 2$ 结构域。在一个方面,人源化的MHC II  $\alpha$ 和 $\beta$ 多肽均分别由位于内源性MHC II  $\alpha$ 和 $\beta$ 基因座的核苷酸序列所编码;在一个方面,所述动物不表达功能性的内源性MHC II  $\alpha$ 和 $\beta$ 多肽。因此,所述动物表达的MHC II可

以是嵌合的人/非人,例如人/啮齿动物(例如人/小鼠)MHC II蛋白。所述嵌合的MHC II蛋白的人部分可以来源于选自下组的人II类HLA蛋白:HLA-DR、HLA-DQ、HLA-DP,例如HLA-DR4、HLA-DR2、HLA-DQ2.5、HLA-DQ8,或者在人类群体中的任意其他的II类HLA分子。在所述实施方式中,其中所述动物是小鼠,所述嵌合的MHC II多肽的非人(即小鼠)部分可以来源于选自H-2E和H-2A的小鼠MHC II蛋白。在一个方面,包含嵌合的人/非人CD4和美国专利申请号13/661,116和13/793,935中所述的人源化的MHC II的所述非人动物可以通过将包含如本申请所述的嵌合CD4基因座的动物与包含人源化MHC II基因座的动物繁殖产生。所述动物还可以通过将编码嵌合的CD4(例如在内源性CD4基因座取代)的核苷酸序列引入包含人源化的MHC II基因座的动物的ES细胞中;或者将编码人源化的MHC II的核苷酸序列引入包含嵌合CD4基因座的动物的ES细胞中产生。

[0068] 在一个实施方式中,包含嵌合的人/非-人CD4和人或人源化的MHC II的所述基因修饰的非人动物(例如小鼠)可以包含一个或两个拷贝的编码这些蛋白的基因;因此,针对编码嵌合的CD4和MHC II(即,MHC II  $\alpha$ 和/或MHC II  $\beta$ )的基因可以分别是杂合子或纯合子。

[0069] 除了基因工程改造的非人动物以外,本申请还提供了非人胚胎(例如啮齿动物,例如小鼠或大鼠胚胎),其中所述胚胎包含来源于如本申请所述的非人动物(例如,啮齿动物,例如小鼠或大鼠)的供体ES细胞。在一个方面,所述胚胎包含含有所述嵌合的CD4基因的ES供体细胞和宿主胚胎细胞。

[0070] 本申请提供了一种组织,其中所述组织来源于如本申请所述的非人动物(例如啮齿动物,例如小鼠或大鼠),并且表达所述嵌合的CD4蛋白。

[0071] 此外,本申请提供了来源于如本申请所述的非人动物的非人细胞。在一个实施方式中,所述细胞是ES细胞。在一个实施方式中,所述细胞是T细胞,例如CD4<sup>+</sup>T细胞。在一个实施方式中,所述细胞是辅助T细胞(T<sub>H</sub>细胞)。在一个实施方式中,所述T<sub>H</sub>细胞是效应器T<sub>H</sub>细胞,例如T<sub>H</sub>1细胞或T<sub>H</sub>2细胞。

[0072] 本申请提供了一种非人细胞,所述非人细胞包含如本申请所述的非人动物的染色体或其片段。在一个实施方式中,所述非人动物细胞包含如本申请所述的非人动物的细胞核。在一个实施方式中,所述非人细胞包含核转移得到的染色体或其片段。

[0073] 在一个方面,本申请提供了一种非人的诱导多能细胞,所述干细胞包含编码如本申请所述的嵌合的CD4多肽的基因。在一个实施方式中,所述诱导多能细胞来源于如本申请所述的非人动物。

[0074] 在一个方面,本申请提供了一种杂交瘤或四价体瘤,其来源于如本申请所述的非人动物细胞。在一个实施方式中,所述非人动物是小鼠或大鼠。

[0075] 在一个方面,本申请提供了一种体外制备物,所述体外制备物包含在其表面携带嵌合CD4蛋白的T细胞和与所述嵌合CD4结合的另一种细胞。在一个实施方式中,所述另一种细胞是表达MHC II多肽(例如嵌合的人/非人MHC II蛋白)的细胞(例如APC)。在一个实施方式中,在第一种细胞表面上的嵌合的CD4与在第二种细胞表面上的嵌合的MHC II相互作用。在一个实施方式中,所述嵌合的CD4蛋白保留了与内源性胞质分子的相互作用,例如内源性胞质信号分子(例如内源性Lck等)。

[0076] 本申请还提供了一种用于制备本申请所述的基因工程改造的非人动物(例如基因

工程改造的啮齿动物,例如小鼠或大鼠)的方法。在一个实施方式中,用于制备基因工程改造的非人动物的方法使得所述动物在内源性CD4基因座包含编码嵌合的人/非人CD4多肽的核苷酸序列。在一些实施方式中,如在实施例描述的,所述方法利用使用 **VELOCIGENE®** 技术制备的靶向构建体,将所述构建体引入ES细胞,并使用 **VELOCIMOUSE®** 技术将经靶向的ES细胞克隆引入小鼠的胚胎中。

[0077] 在一个实施方式中,本发明包括修饰非人动物的CD4基因座以表达本申请所述的嵌合的人/非人CD4多肽的方法。在一个实施方式中,本发明提供了一种修饰小鼠的CD4基因座以表达嵌合的人/小鼠CD4多肽的方法,所述方法包括使用编码嵌合的人/小鼠CD4多肽的核苷酸序列在小鼠的内源性CD4基因座取代编码内源性小鼠CD4多肽的核苷酸序列。在所述方法的一个方面,所述嵌合的人/小鼠CD4多肽包含全部或基本上全部的人CD4多肽的胞外域和内源性小鼠CD4多肽的至少跨膜和胞质结构域。在所述方法的另一个方面,所述嵌合的人/小鼠CD4多肽包含全部或基本上全部的人CD4多肽的D1-D2结构域。在又一个实施方式中,所述嵌合的人/小鼠CD4多肽包含全部或基本上全部的人CD4多肽的D1-D3结构域。在又一个实施方式中,所述嵌合的人/小鼠CD4多肽包含全部或基本上全部的负责与MHC II和/或T细胞受体的胞外域相互作用的人CD4的氨基酸序列。在又一个实施方式中,所述嵌合的人/小鼠CD4多肽包含全部或基本上全部的负责与MHC II和/或T细胞受体的可变结构域相互作用的人CD4的氨基酸序列。

[0078] 因此,本申请还提供了用于产生本申请所述的基因修饰动物的核苷酸构建体。在一个方面,所述核苷酸序列包含5' 和3' 同源臂、包含人CD4基因序列的DNA片段(例如人CD4胞外域基因序列,例如全部或基本上全部的人CD4的结构域D1-D2的基因序列,例如全部或基本上全部的人CD4的结构域D1-D3和/或D2-D3的基因序列,例如全部或基本上全部的人CD4的结构域D1-D4的基因序列)和翼侧为重组位点的选择性表达盒。在一个实施方式中,人CD4基因序列是包含人CD4的内含子和外显子的基因组序列。在一个实施方式中,同源臂与非人(例如小鼠)CD4基因组序列是同源的。本发明的示例性构建体如图1中的下图所示。

[0079] 选择性表达盒是插入靶向构建体以便于选择具有已整合了目标构建体的细胞(例如ES细胞)的核苷酸序列。许多适宜的选择性表达盒是本领域公知的。通常,在存在特定抗生素(例如Neo、Hyg、Pur、CM、SPEC等)的情况下能够对选择性表达盒进行阳性选择。此外,选择性表达盒的翼侧可以是重组位点,其能够使所述选择性表达盒在经过重组酶处理后缺失。通常使用的重组位点是loxP和Frt,其分别被Cre和Flp酶所识别,但是,其他的也是本领域公知的。选择性表达盒可以位于构建体编码区以外的任意位置。在一个实施方式中,所述选择性表达盒位于人CD4序列的外显子3和外显子4之间。

[0080] 在完成基因靶向后,对ES细胞或基因修饰的非人动物进行筛选以确证目标外源性核苷酸序列的成功掺入或外源性多肽的表达。多种技术是本领域技术人员公知的,并且包括(但不限于)Southern印迹、长PCR、定量PCR(例如使用**TAQMAN®**的实时PCR)、荧光原位杂交、Northern印迹、流式细胞术、Western分析、免疫细胞化学、免疫组织化学等。在一个例子中,可以通过使用Valenzuela等,(2003) High-throughput engineering of the mouse genome coupled with high-resolution expression analysis(与高分辨率表达分析偶联的小鼠基因组的高通量工程改造),Nature Biotech.21(6):652-659描述的改良的等位

基因测定对小鼠等位基因的丢失和/或人等位基因的获得情况进行筛选对携带目标基因修饰的非人动物(例如小鼠)进行鉴定。鉴定基因修饰动物中的特定核苷酸或氨基酸序列的其他测定是本领域技术人员公知的。

[0081] 在一个方面,本申请提供了一种制备嵌合的人/或非人CD4分子的方法,所述方法包括由如本申请所述的核苷酸构建体在单一细胞中表达嵌合的CD4蛋白。在一个实施方式中,所述核苷酸构建体是病毒载体;在一个特定的实施方式中,所述病毒载体是慢病毒载体。在一个实施方式中,所述细胞选自CHO、COS、293、HeLa、以及表达病毒核酸序列的视网膜细胞(例如PERC.6<sup>TM</sup>细胞)。

[0082] 在一个方面,本申请提供了表达嵌合的CD4蛋白的细胞。在一个实施方式中,所述细胞包含含有如本申请所述的嵌合的CD4序列的表达载体。在一个实施方式中,所述细胞选自CHO、COS、293、HeLa、以及表达病毒核酸序列的视网膜细胞(例如PERC.6<sup>TM</sup>细胞)。

[0083] 本申请还提供了由本申请所述的非人动物制备的嵌合的CD4分子,其中,在一个实施方式中,所述嵌合的CD4分子包含全部或基本上全部的人CD4蛋白的胞外域的氨基酸序列,以及非人CD4蛋白(例如小鼠CD4蛋白)的至少跨膜和胞质结构域的氨基酸序列。在另一个实施方式中,本申请提供了由本申请所述的非人动物制备的嵌合的CD4分子,其中所述嵌合的CD4分子包含至少全部或基本上全部的人CD4的D1结构域的氨基酸序列,例如至少全部或基本上全部的人CD4的D1-D2结构域,例如至少全部或基本上全部的人CD4的D1-D3结构域,例如负责与MHC II和/或TCR的胞外域结合的人CD4的氨基酸序列,例如负责与MHC II和/或TCR的可变结构域结合的人CD4的氨基酸序列;和其中所述蛋白其余部分(例如跨膜结构域、胞质结构域和未被人源化的胞外域的任意部分)来源于内源性的非人蛋白序列。

[0084] 本申请上文描述的涉及表达嵌合的CD4蛋白的动物,以及包含其的细胞和组织的不同实施方式可以应用于描述表达本申请下文所述的嵌合的人/非人CD8的非人动物,以及表达其他重要嵌合的人/非人T细胞共受体的动物的实施方式。

[0085] 基因修饰的CD8动物

[0086] 在不同实施方式中,本发明通常提供了基因修饰的非人动物,所述非人动物在其基因组(例如内源性CD8基因座)中包含编码人源化的CD8多肽的核苷酸序列;因此,所述动物表达人源化的CD8多肽。在不同实施方式中,本发明提供了非人动物,所述非人动物在其基因组(例如内源性CD8基因座)中包含编码人源化的CD8 $\alpha$ 多肽的核苷酸序列和/或编码人源化的CD8 $\beta$ 多肽的核苷酸序列。因此,本发明的基因修饰的非人动物表达人源化的CD8 $\alpha$ 和/或人源化的CD8 $\beta$ 多肽。

[0087] 人CD8蛋白通常以两种多肽CD8 $\alpha$ 和CD8 $\beta$ 的异源二聚体的形式在细胞表面表达,但是也检测到了二硫键连接的同源二聚体和同源多聚体(例如在NK细胞和肠 $\gamma\delta$ T细胞中,其表达CD8 $\alpha\alpha$ )。编码人CD8 $\alpha$ 和CD8 $\beta$ 的基因位于2号染色体上彼此接近的位置。Nakayama等,(1992)Recent Duplication of the Two Human CD8 $\beta$ -chain genes(两个人CD8 $\beta$ -链基因的最新复制),J.Immunol.148:1919-27。CD8 $\alpha$ 蛋白含有先导肽、免疫球蛋白V-样区、铰链区、跨膜结构域和胞质尾。Norment等,(1989)Alternatively Spliced mRNA Encodes a Secreted Form of Human CD8 $\alpha$ (替代性剪接mRNA编码分泌形式的人CD8 $\alpha$ ) Characterization of the Human CD8 $\alpha$  gene(人CD8 $\alpha$ 基因的特征),J.Immunol.142:3312-19。CD8 $\alpha$ 基因外显子/内含子的示意性描述参见图4和5。

[0088] 人CD8 $\beta$ 基因位于2号染色体上CD8 $\alpha$ 基因的上游。已报道了通过CD8 $\beta$ 的替代性剪接产生的多种亚型,其中一种亚型预计缺乏跨膜结构域并且产生分泌蛋白。Norment等,(1988) A second subunit of CD8 is expressed in human T cells (在人T细胞中表达CD8的另一种亚基), EMBO J. 7:3433-39。CD8 $\beta$ 基因外显子/内含子的示意性描述参见图3和5。

[0089] 膜结合的CD8 $\beta$ 蛋白含有N-末端信号序列,随后是免疫球蛋白V-样结构域、短的胞外铰链区、跨膜结构域和胞质尾。参见,Littman(1987) The structure of the CD4 and CD8 genes (CD4和CD8基因的结构), Ann Rev. Immunol. 5:561-84。铰链区是广泛的糖基化位点,认为该区域维持了其构象并且保护所述蛋白免于被蛋白酶裂解。Leahy (1995) A structural view of CD4 and CD8 (CD4和CD8的结构综述), FASEB J. 9:17-25。

[0090] CD8通常在细胞毒性T细胞上表达,并且与MHC I分子相互作用。所述相互作用由CD8与MHC I的 $\alpha 3$ 结构域的结合所介导。尽管I类MHC与CD8的结合比TCR与I类MHC的结合弱约100倍,但是CD8结合增强了TCR结合的亲和性。Wooldridge等,(2010) MHC Class I Molecules with Superenhanced CD8 Binding Properties Bypass the Requirement for Cognate TCR Recognition and Nonspecifically Activate CTLs (具有超强提升的CD8结合性质的I类MHC分子绕开对同源性TCR识别的需求并且非特异性活化CTL), J. Immunol. 184:3357-3366。

[0091] CD8与I类MHC分子的结合具有物种特异性;CD8、Lyt-2的小鼠同源物显示出在 $\alpha 3$ 结构域与H-2D<sup>d</sup>分子的结合,但是其不与HLA-A分子结合。Connolly等,(1988) The Lyt-2 Molecule Recognizes Residues in the Class I  $\alpha 3$  Domain in Allogeneic Cytotoxic T Cell Responses (在同种异体细胞毒性T细胞应答中Lyt-2分子识别I类 $\alpha 3$ 结构域中的残基), J. Exp. Med. 168:325-341。据推测差异性结合是由于CD8上的CDR-样决定簇(CDR1-和CDR2-样)在人和小鼠之间不是保守的导致的。Sanders等,(1991) Mutations in CD8 that Affect Interactions with HLA Class I and Monoclonal Anti-CD8 Antibodies (在CD8中影响I类HLA与单克隆抗-CD8抗体相互作用的突变), J. Exp. Med. 174:371-379; Vitiello等,(1991) Analysis of the HLA-restricted Influenza-specific Cytotoxic T Lymphocyte Response in Transgenic Mice Carrying a Chimeric Human-Mouse Class I Major Histocompatibility Complex (在携带嵌合的人-小鼠I类主要组织相容性复合物的转基因小鼠中对HLA-限制性流感特异性细胞毒性T淋巴细胞应答的分析), J. Exp. Med. 173:1007-1015; 和Gao等,(1997) Crystal structure of the complex between human CD8 $\alpha\alpha$  and HLA-A2 (人CD8 $\alpha\alpha$ 和HLA-A2之间复合物的晶体结构), Nature 387:630-634。已报道CD8在 $\alpha 3$ 结构域的保守区(在位置223-229)与HLA-A2结合。在HLA-A中的单取代(V245A)降低了CD8与HLA-A的结合,同时伴随了大大减少的T细胞介导的裂解。Salter等,(1989), Polymorphism in the  $\alpha 3$  domain of HLA-A molecules affects binding to CD8 (HLA-A分子在 $\alpha 3$ 结构域中的多态性影响其与CD8的结合), Nature 338:345-348。在通常情况下,HLA-A分子在 $\alpha 3$ 结构域中的多态性还影响其与CD8的结合。Id。在小鼠中,在H-2D<sup>d</sup>的残基227中的氨基酸取代影响小鼠Lyt-2与H-2D<sup>d</sup>的结合,并且转染突变的H-2D<sup>d</sup>的细胞不会被CD8+T细胞所裂解。Potter等,(1989) Substitution at residue 227 of H-2 class I molecules abrogates recognition by CD8-dependent, but not CD8-

independent, cytotoxic T lymphocytes (在I类H-2分子残基227的取代消除CD8依赖性识别,但不影响非CD8依赖的细胞毒性T淋巴细胞), Nature 337:73-75。因此,人或人源化CD8的表达可能对于研究针对人或人源化的MHC I递呈抗原的T细胞应答是有益的。

[0092] 与CD4类似,CD8的胞质结构域与酪氨酸激酶Lck相互作用,进而导致T细胞活化。尽管Lck似乎与CD8 $\alpha$ 的胞质结构域相互作用,但是这种相互作用被存在CD8 $\beta$ 的胞质结构域所调控,因为CD8 $\beta$ 胞质结构域的突变或缺失会导致CD8 $\alpha$ 相关的Lck活性的降低。Irie等,(1998) The cytoplasmic domain of CD8 $\beta$  Regulates Lck Kinase Activation and CD8 T cell Development (CD8 $\beta$ 的胞质结构域调控Lck激酶活化和CD8T细胞发育), J. Immunol. 161:183-91。Lck活性的降低与T细胞发育受损相关, Id。

[0093] CD8在适宜细胞(例如细胞毒性T细胞)上的表达被位于整个CD8基因座的多种增强子元件紧密地调控。例如,在CD8基因座已鉴定得到了至少4个DNase I-高敏感性区域,这些区域通常与调控结合相关。Hosert等,(1997) A CD8 genomic fragment that directs subset-specific expression of CD8 in transgenic mice (在转基因小鼠中涉及CD8的子集特异性表达的CD8基因组片段), J. Immunol. 158:4270-81。由于在CD8基因座的这些DNase I-高敏感性区域的发现,已鉴定得到的至少5个增强子元件,其分布在整个CD8基因座,其在T细胞的不同谱系中调控CD8 $\alpha$ 和/或 $\beta$ 的表达,包括DP、CD8 SP T细胞或表达 $\gamma\delta$  TCR的细胞。参见例如Kioussis等,(2002) Chromatin and CD4, CD8 $\alpha$ , and CD8 $\beta$  gene expression during thymic differentiation (在胸腺分化过程中的染色质以及CD4、CD8 $\alpha$ 和CD8 $\beta$ 的基因表达), Nature Rev. 2:909-919和网络勘误; Ellmeier等,(1998) Multiple Development Stage-Specific Enhancers Regulate CD8 Expression in Developing Thymocytes and in Thymus-Independent T cells (在发育中的胸腺细胞和不依赖性胸腺的T细胞中的多个发育阶段特异性的增强子调控CD8的表达), Immunity 9:485-96。

[0094] 因此,与在人或人源化的CD4基因修饰的动物中保留内源性CD4启动子和调控元件的益处相似,在一些实施方式中,研发保留控制人或人源化CD8的表达的内源性小鼠启动子和调控元件的基因修饰的非人动物可能是有益的。其在形成如本申请所述的包含使用编码人或人源化的CD8 $\alpha$ 和/或 $\beta$ 蛋白的那些取代编码CD8 $\alpha$ 和/或 $\beta$ 蛋白的内源性非人序列的基因修饰的动物中可能是特别有益的。

[0095] 在不同实施方式中,本发明提供了一种基因修饰的非人动物,所述非人动物在其基因组(例如内源性CD8基因座)中包含至少一个编码嵌合的人/非人CD8多肽(例如CD8 $\alpha$ 和/或 $\beta$ 多肽)的核苷酸序列,其中所述多肽的人源部分包含全部或基本上全部的人CD8(例如CD8 $\alpha$ 和/或 $\beta$ )的胞外域,其中非人部分包含非人CD8(例如CD8 $\alpha$ 和/或 $\beta$ )的至少跨膜和胞质结构域,和其中所述动物表达嵌合的CD8多肽(例如CD8 $\alpha$ 和/或 $\beta$ 多肽)。因此,在一个实施方式中,本发明提供了一种基因修饰的非人动物,所述非人动物在其内源性非人CD8基因座包含编码嵌合的人/非人CD8 $\alpha$ 多肽的第一核苷酸序列和编码嵌合的人/非人CD8 $\beta$ 多肽的第二核苷酸序列,其中所述第一核苷酸序列包含编码全部或基本上全部的人CD8 $\alpha$ 多肽胞外域以及非人CD8 $\alpha$ 多肽的至少跨膜和胞质结构域的序列,和其中所述第二核苷酸序列包含全部或基本上全部的人CD8 $\beta$ 多肽胞外域以及非人CD8 $\beta$ 多肽的至少跨膜和胞质结构域的序列,其中所述动物表达功能性的嵌合人/非人CD8蛋白。在一个方面,所述非人动物仅表达人源化的CD8多肽(例如嵌合的人/非人CD8 $\alpha$ 和/或 $\beta$ 多肽),并且不表达来自内源性CD8基因座的相应的功

能性非人CD8多肽。

[0096] 在一个实施方式中,所述嵌合的人/非人CD8 $\alpha$ 多肽在其人部分包含全部或基本上全部的人CD8 $\alpha$ 多肽的胞外域。在一个实施方式中,所述嵌合的CD8 $\alpha$ 多肽的人部分至少包含人CD8 $\alpha$ 多肽的MHC I结合结构域。在一个实施方式中,所述嵌合的CD8 $\alpha$ 多肽的人部分包含至少全部或基本上全部的人CD8 $\alpha$ 免疫球蛋白V-样结构域的序列。在一个实施方式中,编码所述嵌合的CD8 $\alpha$ 多肽的人部分的核苷酸序列至少包含编码人CD8 $\alpha$ 多肽的胞外域的外显子。在一个实施方式中,所述核苷酸序列至少包含编码Ig V-样结构域的外显子。在一个实施方式中,人CD8 $\alpha$ 多肽的胞外域是包含所述多肽非跨膜或胞质结构域的结构域的区域。在一个实施方式中,编码所述嵌合的人/非人CD8 $\alpha$ 多肽的核苷酸序列包含编码非人(例如啮齿动物,例如小鼠)CD8 $\alpha$ 信号肽的序列。或者,所述核苷酸序列可以包含编码人CD8 $\alpha$ 信号肽的序列。在一个实施方式中,所述嵌合的人/非人CD8 $\alpha$ 多肽包含如SEQ ID NO:54所示的氨基酸序列,和所述嵌合多肽的人部分如SEQ ID NO:54第28-179位的氨基酸所示(单独用SEQ ID NO:59表示)。

[0097] 类似地,在一个实施方式中,所述嵌合的人/非人CD8 $\beta$ 多肽在其人部分中包含全部或基本上全部的人CD8 $\beta$ 多肽的胞外域。在一个实施方式中,所述嵌合的CD8 $\beta$ 多肽的人部分包含全部或基本上全部的CD8 $\beta$ 免疫球蛋白V-样结构域的序列。在一个实施方式中,编码所述嵌合的CD8 $\beta$ 多肽的人部分的核苷酸序列至少包含编码人CD8 $\beta$ 多肽的胞外域的外显子。在一个实施方式中,所述嵌合的人/非人CD8 $\beta$ 多肽的人部分至少包含编码CD8 $\beta$  IgG V-样结构域的外显子。在一个实施方式中,所述编码嵌合的人/非人CD8 $\beta$ 多肽的核苷酸序列包含编码非人(例如啮齿动物,例如小鼠)CD8 $\beta$ 信号肽的序列。或者,所述核苷酸序列可以包含编码人CD8 $\beta$ 信号肽的序列。在一个实施方式中,所述嵌合的人/非人CD8 $\beta$ 多肽包含如SEQ ID NO:53所示的氨基酸序列,和所述嵌合多肽的人部分如SEQ ID NO:53第15-165位的氨基酸所示(单独用SEQ ID NO:58表示)。

[0098] 在一个实施方式中,所述非人动物表达嵌合的人/非人CD8 $\alpha$ 和/或CD8 $\beta$ 多肽。在一些实施方式中,所述嵌合的人/非人CD8 $\alpha$ 和/或 $\beta$ 多肽的人部分包含一个或多个保守的或非保守的修饰。

[0099] 在一个方面,本申请提供了一种表达人CD8 $\alpha$ 和/或 $\beta$ 多肽序列的非人动物,其中所述人CD8 $\alpha$ 和/或 $\beta$ 多肽的序列分别与人CD8 $\alpha$ 和/或 $\beta$ 多肽的序列具有至少约85%、90%、95%、96%、97%、98%或99%的一致性。在一个特定实施方式中,所述人CD8 $\alpha$ 和/或 $\beta$ 多肽的序列分别与实施例中描述的人CD8 $\alpha$ 和/或 $\beta$ 多肽的序列具有至少约90%、95%、96%、97%、98%或99%的同一性。在一个实施方式中,所述人CD8 $\alpha$ 和/或 $\beta$ 多肽的序列包含一个或多个保守的取代。在一个实施方式中,所述人CD8 $\alpha$ 和/或 $\beta$ 多肽的序列包含一个或多个非保守的取代。

[0100] 在一些实施方式中,一部分例如所述嵌合CD8的人部分可以包含基本上全部的本申请所示的序列(例如,基本上全部的本申请所示的蛋白结构域)。基本上全部序列通常包括据信代表所述蛋白特定部分(例如特定功能性结构域等)的85%、90%、95%、96%、97%、98%或99%的氨基酸。本领域技术人员将理解可以根据所使用的比对和结构域预测方法的不同,功能性结构域的边界可能发生变化。

[0101] 在一个方面,所述嵌合的人/非人CD8 $\alpha$ 和/或 $\beta$ 多肽的非人部分至少分别包含所述非人CD8 $\alpha$ 和/或 $\beta$ 多肽的跨膜和/或胞质结构域。由于CD8胞质结构域具有重要的功能,因此

在基因工程改造的动物中保留内源性非人(例如小鼠)序列确保保持共受体具有正确的细胞内信号和其他功能。在一个实施方式中,所述非人动物是小鼠,和所述非人CD8 $\alpha$ 和/或 $\beta$ 多肽分别是小鼠CD8 $\alpha$ 和/或 $\beta$ 多肽。尽管在实施例中描述了特定的小鼠CD8 $\alpha$ 和/或 $\beta$ 序列,但是来源于其的任意适宜序列例如包含保守的/非保守的氨基酸取代的序列均包括在本申请中。在一个实施方式中,所述非人动物(例如啮齿动物,例如小鼠)保留了未被人源化的任意内源性序列。

[0102] 本申请所述的非人动物可以在其内源性基因座包含编码嵌合的人/非人CD8 $\alpha$ 和/或 $\beta$ 多肽的核苷酸序列。在一个方面,其导致内源性CD8 $\alpha$ 基因的一部分被编码人CD8 $\alpha$ 多肽的一部分的核苷酸序列所取代,和/或内源性CD8 $\beta$ 基因的一部分被编码人CD8 $\beta$ 多肽的一部分的核苷酸序列所取代。在一个实施方式中,此类取代是编码全部或基本上全部的非人CD8 $\alpha$ 和/或 $\beta$ 的胞外域的内源性核苷酸序列被编码其的人核苷酸序列的取代。在一个实施方式中,此类取代是编码全部或基本上全部的非人CD8 $\alpha$ 和/或 $\beta$ 的免疫球蛋白V-样结构域的序列被编码其的人核苷酸序列的取代。在一个实施方式中,所述取代不包含对编码非人CD8 $\alpha$ 和/或 $\beta$ 多肽跨膜和胞质结构域的CD8 $\alpha$ 和/或 $\beta$ 序列的取代。因此,所述非人动物表达来自内源性非人CD8基因座的嵌合的人/非人CD8 $\alpha$ 和/或 $\beta$ 多肽。在又一个实施方式中,所述取代产生包含分别如SEQ ID NO:54和/或53所示的多肽序列的CD8 $\alpha$ 和/或 $\beta$ 蛋白。

[0103] 在一个实施方式中,本申请提供了所述嵌合的人/非人CD8基因座(例如嵌合的啮齿动物CD8基因座,例如嵌合的小鼠CD8基因座)的核苷酸序列。在一个方面,由于所述嵌合的人/非人(例如人/啮齿动物,例如人/小鼠)CD8 $\alpha$ 和/或 $\beta$ 序列分别位于内源性非人(例如啮齿动物,例如小鼠)的CD8 $\alpha$ 和/或 $\beta$ 基因座,因而其保留了内源性CD8 $\alpha$ 和/或 $\beta$ 的启动子和调控元件。在另一个实施方式中,所述嵌合的基因座可以包含人CD8 $\alpha$ 和/或 $\beta$ 启动子和调控元件的程度为使得恰当地实现CD8 $\alpha$ 和/或 $\beta$ 表达(恰当的空间和时间上的蛋白表达)、CD8 $^+$ T细胞发育、CD8谱系选择和共受体功能。因此,在一个方面,本发明的动物包含不改变恰当的谱系选择和T细胞发育的基因修饰。在一个方面,本发明的动物(例如啮齿动物,例如小鼠)在正常表达CD8的细胞以外的免疫细胞上不表达嵌合的CD8蛋白,例如动物在B细胞或CD4 $^+$ SP T细胞上不表达CD8。在一个实施方式中,所述取代保留了能够恰当地在空间和时间上调控CD8 $\alpha$ 和/或 $\beta$ 表达的元素。

[0104] 包含本申请所述的人或人源化的CD8多肽的基因修饰的非人动物可以选自在上文中与人源化的CD4动物的相关章节描述的任意动物。在一个实施方式中,所述非人动物可以是啮齿动物,例如大鼠或小鼠。

[0105] 因此,在一个实施方式中,本发明提供了一种基因修饰的小鼠,所述小鼠在其基因组例如内源性CD8基因座中包含编码嵌合的人/小鼠CD8 $\alpha$ 多肽的第一核苷酸序列和编码嵌合的人/小鼠CD8 $\beta$ 多肽的第二核苷酸序列。在一个实施方式中,所述第一核苷酸序列包含编码全部或基本上全部的人CD8 $\alpha$ 多肽的胞外域和小鼠CD8 $\alpha$ 多肽的至少跨膜和胞质结构域的序列,和所述第二核苷酸序列包含编码全部或基本上全部的人CD8 $\beta$ 多肽的胞外域和小鼠CD8 $\beta$ 多肽的至少跨膜和胞质结构域的序列,和其中所述小鼠表达功能性的嵌合的人/小鼠CD8蛋白。在一个实施方式中,所述第一核苷酸序列包含编码人CD8 $\alpha$ 多肽的至少免疫球蛋白V-样结构域的序列和小鼠CD8 $\alpha$ 多肽的其余序列,和所述第二核苷酸序列包含编码人CD8 $\beta$ 多肽的至少免疫球蛋白V-样结构域的序列和小鼠CD8 $\beta$ 多肽的其余序列。在一个实施方式中,

所述第一核苷酸序列至少包含人CD8 $\alpha$ 多肽的MHC I结合结构域。在一个实施方式中,所述第一和第二核苷酸序列至少包含分别编码人CD8 $\alpha$ 多肽和/或CD8 $\beta$ 多肽胞外域的外显子的序列。在一个实施方式中,所述人CD8 $\alpha$ 多肽和/或CD8 $\beta$ 多肽的胞外域是包含非跨膜或胞质结构域的人CD8 $\alpha$ 多肽和/或CD8 $\beta$ 多肽结构域的区域。在一个实施方式中,所述人CD8 $\alpha$ 多肽的结构域如图4和5中的示意图。在一个实施方式中,所述人CD8 $\beta$ 多肽的结构域如图3和5中示意图。在一个实施方式中,编码所述嵌合的人/小鼠CD8 $\alpha$ 多肽和CD8 $\beta$ 多肽的核苷酸序列包含分别编码小鼠CD8 $\alpha$ 和/或CD8 $\beta$ 信号肽的序列。或者,所述核苷酸序列可以包含编码人CD8 $\alpha$ 和/或 $\beta$ 信号序列的序列。在一个实施方式中,所述小鼠包含编码全部或基本上全部的小鼠CD8 $\alpha$ 和/或CD8 $\beta$ 胞外域的核苷酸序列分别被编码全部或基本上全部的人CD8 $\alpha$ 和/或CD8 $\beta$ 胞外域的核苷酸序列的取代。

[0106] 在一个实施方式中,所述小鼠不表达来自其内源性CD8基因座的功能性内源性小鼠CD8 $\alpha$ 和/或CD8 $\beta$ 多肽。在一个实施方式中,本申请所述的小鼠在其种系中包含嵌合的人/小鼠CD8序列。

[0107] 在一个方面,表达嵌合的人/小鼠CD8 $\alpha$ 和/或CD8 $\beta$ 多肽的小鼠保留了小鼠CD8 $\alpha$ 和/或CD8 $\beta$ 的启动子和调控序列,例如在所述小鼠中编码嵌合的人/小鼠CD8的核苷酸序列与内源性的小鼠CD8启动子和调控序列可操作地连接。在一个方面,在所述小鼠中保留的这些调控序列包括在T细胞发育的合适阶段调控CD8蛋白表达的序列。在一个方面,所述基因修饰的小鼠在CD4谱系的B细胞或T细胞,或正常不表达内源性CD8的任意细胞(例如免疫细胞)上不表达嵌合的CD8。

[0108] 在不同实施方式中,表达来自如本申请所述的嵌合的CD8基因座的功能性嵌合CD8蛋白(例如CD8 $\alpha\beta$ 或CD8 $\alpha\alpha$ )的非人动物(例如啮齿动物,例如小鼠或大鼠)在细胞表面上显示出嵌合的蛋白。在一个实施方式中,所述非人动物表达的嵌合的CD8蛋白在细胞表面上的细胞分布与在人中观察到的相同。在一个方面,本发明的CD8蛋白能够与另一种细胞表面上表达的MHC I蛋白相互作用。

[0109] 在一个实施方式中,本发明的非人动物(例如啮齿动物,例如小鼠)还包含编码人或人源化的MHC I蛋白的核苷酸序列,以使得在所述动物的T细胞表面上表达的嵌合的CD8蛋白能够与另一种细胞(例如抗原递呈细胞)表面上表达的人或人源化的MHC I相互作用。在一个实施方式中,所述MHC I蛋白包含人MHC I多肽的胞外域。在一个实施方式中,所述动物还包含人或人源化的 $\beta$ -2微球蛋白多肽。在2012年10月26日提交的美国专利申请号13/661,159和2013年3月11日提交的美国专利申请号13/793,812中描述了表达人或人源化的MHC I多肽和/或 $\beta$ -2微球蛋白多肽的示例性基因修饰的动物,其全部内容通过引入并入本申请。因此,在一个实施方式中,包含本申请所述的嵌合的CD8蛋白的动物还可以包含人源化的MHC I复合物,其中所述人源化的MHC I复合物包含:(1)人源化的MHC I多肽,例如其中所述人源化的MHC I多肽包含人MHC I胞外域以及内源性(例如小鼠)MHC I的跨膜和胞质结构域,例如其中所述人源化的MHC I包含人MHC I多肽的 $\alpha$ 1、 $\alpha$ 2和 $\alpha$ 3结构域,和(2)人或人源化的 $\beta$ 2微球蛋白多肽(例如所述动物在其基因组中包含如人 $\beta$ 2微球蛋白的外显子2、3和4所示的核苷酸序列)。在一个方面,人源化的MHC I以及人或人源化的 $\beta$ 2微球蛋白多肽分别由位于内源性MHC I和 $\beta$ 2微球蛋白基因座的核苷酸序列编码;在一个方面,所述动物不表达功能性的内源性的MHC I和 $\beta$ 2微球蛋白多肽。因此,所述动物表达的MHC I可以是嵌合的人/非

人,例如人/啮齿动物(例如人/小鼠)MHC I多肽。所述嵌合的MHC I多肽的人部分可以来源于选自下组的人I类HLA蛋白:HLA-A、HLA-B和HLA-C,例如HLA-A2、HLA-B27、HLA-B7、HLA-Cw6,或者在人类群体中存在的任意其他I类HLA。在所述实施方式中,其中所述动物是小鼠,所述嵌合的MHC I多肽的非人(即小鼠)部分可以来源于选自H-2D、H-2K和H-2L的小鼠MHC I蛋白。在一个方面,包含本申请所述的嵌合的人/非人CD8和美国专利申请号13/661,159和13/793,812中所述的人源化的MHC I和/或 $\beta$ 2微球蛋白的所述非人动物可以通过将包含如本申请所述的嵌合CD8基因座(例如嵌合的CD8 $\alpha$ 和/或 $\beta$ )的动物与包含人源化MHC I和/或 $\beta$ -2微球蛋白基因座的动物繁殖产生。所述动物还可以通过将编码嵌合的CD8(例如嵌合的CD8 $\alpha$ 和/或 $\beta$ ) (例如用于在内源性CD8基因座(例如嵌合的CD8 $\alpha$ 和/或 $\beta$ 基因座)进行取代)的核苷酸序列引入包含人源化的MHC I和/或 $\beta$ -2微球蛋白基因座的动物的ES细胞中产生;或者将编码人源化的MHC I和/或 $\beta$ -2微球蛋白的核苷酸序列引入包含嵌合CD8基因座(例如嵌合的CD8 $\alpha$ 和/或 $\beta$ 基因座)的动物的ES细胞中产生。

[0110] 除了基因工程改造的非人动物以外,本申请还提供了非人胚胎(例如内齿动物,例如小鼠或大鼠胚胎),其中所述胚胎包含来源于如本申请所述的非人动物(例如,啮齿动物,例如小鼠或大鼠)的供体ES细胞。在一个方面,所述胚胎包含含有所述嵌合的CD8基因的ES供体细胞和宿主胚胎细胞。

[0111] 本申请还提供了一种组织,其中所述组织来源于如本申请所述的非人动物(例如啮齿动物,例如小鼠或大鼠),并且表达所述嵌合的CD8蛋白。

[0112] 此外,本申请提供了分离自如本申请所述的非人动物的非人细胞。在一个实施方式中,所述细胞是ES细胞。在一个实施方式中,所述细胞是T细胞,例如CD8+T细胞。在一个实施方式中,所述细胞是细胞毒性T细胞。

[0113] 本申请还提供了一种非人细胞,所述非人细胞包含如本申请所述的非人动物的染色体或其片段。在一个实施方式中,所述非人细胞包含如本申请所述的非人动物的核。在一个实施方式中,所述非人细胞包含核转移得到的染色体或其片段。

[0114] 在一个方面,本申请提供了一种非人的诱导的多能细胞,所述多能细胞包含编码如本申请所述的嵌合的CD8多肽(例如CD8 $\alpha$ 和/或CD8 $\beta$ 多肽)的基因。在一个实施方式中,所述诱导多能细胞来源于如本申请所述的非人动物。

[0115] 在一个方面,本申请提供了一种杂交瘤或四价体瘤,其来源于如本申请所述的非人动物细胞。在一个实施方式中,所述非人动物是小鼠或大鼠。

[0116] 在一个方面,本申请提供了一种体外制备物,所述体外制备物包含在其表面携带嵌合CD8蛋白的T细胞和与所述嵌合CD8结合的另一种细胞。在一个实施方式中,所述另一种细胞是表达MHC I多肽的细胞,例如嵌合的人/非人MHC I蛋白。在一个实施方式中,在第一种细胞表面上的嵌合的CD8与在第二种细胞表面上的嵌合的MHC I相互作用。在一个实施方式中,所述嵌合的CD8蛋白保留了与内源性胞质分子的相互作用,例如内源性胞质信号分子(例如内源性Lck等)。

[0117] 本申请还提供了一种用于制备本申请所述的基因工程改造的非人动物的方法。所述方法使得所述动物在内源性CD8基因座包含编码嵌合的人/非人CD8 $\alpha$ 和/或CD8 $\beta$ 多肽的核苷酸序列。如实施例中所描述的,所述方法利用使用VELOCIGENE®技术制备靶向构建体,将所述构建体引入ES细胞,并使用VELOCIMOUSE®技术将经靶向的ES细胞克隆引入小鼠

的胚胎中。

[0118] 在一个实施方式中,本发明提供了一种修饰非人动物的CD8基因座以表达本申请所述的嵌合的人/非人CD8多肽的方法。在一个方面,提供了一种修饰小鼠的CD8基因座以表达嵌合的人/小鼠CD8多肽的方法,所述方法包括使用编码嵌合的人/小鼠CD8多肽的核苷酸序列在内源性CD8基因座取代编码内源性小鼠CD8多肽的小鼠核苷酸序列。所述CD8多肽可以选自CD8 $\alpha$ 、CD8 $\beta$ 及其组合。在一个方面,所述嵌合的多肽包含全部或基本上全部的人CD8多肽的胞外域和内源性小鼠CD8多肽的至少跨膜和胞质结构域。

[0119] 因此,本申请还提供了用于产生本申请所述的基因修饰的动物的核苷酸构建体。在一个方面,所述核苷酸构建体的序列包含5' 和3' 同源臂、包含人CD8 $\alpha$ 或CD8 $\beta$ 序列的DNA片段和翼侧为重组位点的选择性表达盒。在一些实施方式中,所述人序列包含人CD8 $\alpha$ 或CD8 $\beta$ 的内含子和外显子,例如分别编码人CD8 $\alpha$ 或CD8 $\beta$ 胞外域的外显子。在一个实施方式中,同源臂与非人CD8 $\alpha$ 或CD8 $\beta$ 序列是同源的。CD8 $\alpha$ 或CD8 $\beta$ 的示例性构建体分别如图4和3所示。

[0120] 由于编码CD8 $\alpha$ 或CD8 $\beta$ 的基因具有接近的染色体位置,因此对这两个基因顺序靶向提高了成功的人源化的机会。在一个实施方式中,所述靶向策略包含将本申请所述的嵌合的CD8 $\beta$ 构建体引入ES细胞,由被靶向的ES细胞产生小鼠,从所述小鼠获得基因修饰的ES细胞,以及将本申请所述的嵌合的CD8 $\alpha$ 构建体引入所述基因修饰的ES细胞。在另一个实施方式中,所述靶向策略包含将本申请所述的嵌合的CD8 $\beta$ 构建体引入ES细胞,选择已掺入所述嵌合的CD8 $\beta$ 构建体的所述细胞,将本申请所述的嵌合的CD8 $\alpha$ 构建体引入已掺入和携带嵌合的CD8 $\beta$ 构建体的ES细胞,以及选择已掺入嵌合的CD8 $\beta$ 和CD8 $\alpha$ 的细胞。在本实施方式的一个方面,利用不同的选择标记物进行选择步骤。在其他实施方式中,可以首先完成CD8 $\alpha$ 的人源化。在基因靶向完成后,可以通过多种方法(例如上文所述的针对CD4人源化的方法,例如Valenzuela等描述的改良的等位基因测定,如上所述)对基因修饰的非人动物的ES细胞进行筛选以确证成功掺入了目标外源性核苷酸序列或表达外源性多肽。

[0121] 在一个方面,本申请提供了一种用于制备嵌合的人/非人CD8分子(例如CD8 $\alpha$ 和/或CD8 $\beta$ )的方法,所述方法包括在单一细胞中表达来自如上文所述的核苷酸构建体的嵌合的CD8多肽。在一个实施方式中,所述核苷酸构建体是病毒载体;在一个特定的实施方式中,所述病毒载体是慢病毒载体。在一个实施方式中,所述细胞选自CHO、COS、293、HeLa和表达病毒核酸序列的视网膜细胞(例如PERC.6<sup>TM</sup>细胞)。

[0122] 在一个方面,本申请提供了一种表达嵌合的CD8蛋白的细胞。在一个实施方式中,所述细胞包含含有如本申请所述的嵌合的CD8序列的表达载体。在一个实施方式中,所述细胞选自CHO、COS、293、HeLa和表达病毒核酸序列的视网膜细胞(例如PERC.6<sup>TM</sup>细胞)。

[0123] 本申请还提供了一种由如本申请所述的非人动物制备的嵌合的CD8分子,其中所述嵌合的CD8分子包含全部或基本上全部的来自人CD8蛋白(例如CD8 $\alpha$ 和/或CD8 $\beta$ )的胞外域,和来自非人CD8蛋白(例如小鼠CD8蛋白)的至少跨膜和胞质结构域。示例性的嵌合的CD8 $\alpha$ 多肽如SEQ ID NO:54所示,和示例性的嵌合的CD8 $\beta$ 蛋白如SEQ ID NO:53所示。

[0124] 基因修饰的CD4和CD8动物的用途

[0125] 包含人源化的CD4和MHC II或者人源化的CD8和MHC I的基因修饰的非人动物,例如啮齿动物,例如小鼠或大鼠,以与人类似的方式将肽递呈至T细胞(分别为CD4<sup>+</sup>或CD8<sup>+</sup>T细胞),这是因为所述复合物基本上全部的的成分是人或人源化的。本发明的基因修饰的非人

动物能够用于在人源化的动物中研究人免疫系统的功能；用于鉴定激发免疫应答的抗原和抗原表位（例如T细胞表位，例如独特的人癌症表位），例如用于疫苗研发；用于鉴定人病原体或癌症抗原的高亲和性T细胞（即在具有高亲和力的人MHC I复合物的背景下与抗原结合的T细胞），例如用于适应性T细胞疗法；用于评估疫苗候选物和其他疫苗策略；用于研究人的自身免疫；用于研究人的感染性疾病；以及用于设计基于人MHC和CD4/CD8表达的更好的治疗策略的其他方面。

[0126] 因此，在不同实施方式中，其中，本发明的基因修饰的动物在用于评估抗原在人体内激发免疫应答的能力，以及产生多样性的抗原和鉴定可以用于人用疫苗研发的特定抗原方面是有益的。

[0127] 在一个方面，本申请提供了一种用于确定是否肽将在人体内激发细胞免疫应答的方法，所述方法包括将本申请所述的基因修饰的非人动物暴露于所述肽，使所述非人动物产生免疫应答，以及在与由如本申请所述的嵌合的人/非人MHC I或II分子提呈的所述肽的序列结合的所述非人动物的细胞（例如分别为CD8+或CD4+T细胞，包括人CD8或CD4）中进行检测。在一个实施方式中，暴露后的所述非人动物包含与所述肽结合的I类MHC限制性CD8+细胞毒性T淋巴细胞（CTL）。在另一个实施方式中，暴露后的所述非人动物包含与所述肽结合的MHC II限制性CD4+T细胞。

[0128] 在一个方面，本申请提供了一种用于鉴定人T细胞表位的方法，所述方法包括将如本申请所述的非人动物暴露于包含假定的T细胞表位的抗原，使所述非人动物产生免疫应答，从所述非人动物中分离与所述表位结合的I类MHC或II类MHC限制性T细胞，以及鉴定被所述T细胞结合的表位。

[0129] 在一个方面，本申请提供了一种用于鉴定在人体内产生T细胞应答的抗原的方法，所述方法包括将假定抗原暴露于如本申请所述的小鼠，使所述小鼠产生免疫应答，以及鉴定被I类或II类HLA限制性分子结合的抗原情况。

[0130] 在一个方面，本申请提供了一种用于确定是否假定的抗原含有当暴露于人免疫系统后将产生I类或II类HLA限制性免疫应答的表位的方法，所述方法包括将如本申请所述的小鼠暴露于所述假定的抗原并且测定在小鼠中的抗原特异性I类HLA或II类HLA限制性免疫应答。

[0131] 此外，本申请所述的基因工程改造的非人动物可以用于鉴定识别目标抗原（例如肿瘤或其他疾病抗原）的T细胞受体（例如高亲和力的T细胞受体）。所述方法可以包括：将本申请所述的非人动物暴露于抗原，使所述非人动物针对所述抗原产生免疫应答，从所述非人动物中分离包含与由人或人源化的MHC I或MHC II递呈的抗原结合的T细胞受体的T细胞，以及确定所述T细胞受体的序列。

[0132] 在一个实施方式中，本申请提供了一种通过假定的人用治疗确定T细胞活化的方法，所述方法包括将如本申请所述的基因修饰的动物暴露于假定的人用治疗（或者例如将此类动物的表达人或人源化MHC II或MHC I的细胞暴露于所述假定治疗的肽序列），将所述基因修饰的动物的显示出人或人源化MHC/肽复合物的细胞暴露于包含能够与所述基因修饰的动物的细胞结合的嵌合的人/非人（例如人/小鼠）CD4或CD8的T细胞，以及测定由所述基因修饰的动物的展示出肽的细胞诱导的T细胞活化。

[0133] 除了鉴定来自人病原体或肿瘤的抗原和抗原表位的能力以外，本发明的基因修饰

的动物能够用于鉴定与人自身免疫性疾病相关的自身抗原,例如I型糖尿病、多发性硬化等。而且,本发明的基因修饰的动物能够用于研究人自身免疫性疾病的多个方面,并且可以用作自身免疫性疾病的模型。

[0134] 本申请所述的基因修饰动物(即包含人或人源化的T细胞共受体(例如嵌合的人/非人CD4或CD8)的动物)的其他用途任选地还包括人或人源化的MHC II或I蛋白,其从本公开中将是显而易见的。

## 实施例

[0135] 将通过下述非限制性实施例对本发明进行进一步解释。列出这些实施例用于帮助理解本发明,其不以任何方式构成对其保护范围的限制。所述实施例不包括对本领域普通技术人员熟知的常规方法(分子克隆技术等)的详细描述。除非另有明示,以重量份表示重量,分子量是平均分子量,以摄氏度表示温度和压力为处于或接近大气压。

[0136] 实施例1:基因修饰的CD4小鼠的构建和表征

[0137] 实施例1.1:工程化制备嵌合的CD4基因座

[0138] 通过使用 VELOCIGENE®技术(参见例如美国专利号6,586,251和Valenzuela等,(2003)High-throughput engineering of the mouse genome coupled with high-resolution expression analysis(与高分辨率表达分析偶联的小鼠基因组的高通量工程改造),Nature Biotech.21(6):652-659)在一步中从人和小鼠的细菌人工染色体(BAC)DNA构建独特的靶向载体将小鼠CD4基因座人源化。为产生靶向载体,使用细菌人工染色体(BAC)DNA进行一系列细菌同源重组(BHR)以及其他的基因工程步骤。

[0139] 简言之,四个DNA片段:(1)含有小鼠信号肽的片段(由小鼠CD4基因的外显子2和3编码),(2)含有小鼠信号肽下游的人外显子3的片段,(3)含有翼侧为Asc I和PI-SceI位点的SPEC抗性表达盒的片段,和(4)含有160bp的小鼠CD4内含子6的片段(在外显子6和7之间的内含子),其从小鼠CD4外显子6下游约200bp开始通过注入连接(Clontech)进行连接。所得到的DNA片段从5'至3'含有:小鼠外显子2、小鼠内含子2、含有信号肽的小鼠外显子3的部分、人信号肽下游的人外显子3、人内含子3的一部分、SPEC表达盒、小鼠内含子6的一部分。该DNA片段用在BHR中对小鼠BAC克隆BMQ391F08进行修饰,以缺失编码小鼠CD4Ig样结构域1-3的小鼠序列并引入人CD4的外显子3。使用SPEC表达盒取代BAC的CM表达盒得到第一BAC载体(图1,上图)。

[0140] 利用BHR对人BAC RP11-101F21进行修饰以便在人外显子3下游60bp处引入AscI-LoxP-PGK-neo-loxP表达盒,和在外显子6下游约100bp处引入PI-SceI限制性位点和SPEC表达盒,得到第二BAC载体(图1,中图)。随后在经过AscI和PI-SceI限制性内切酶消化后将所述第一和第二BAC载体进行BAC与BAC连接以产生CD4靶向载体(图1,下图)。小鼠-人和人-小鼠序列的上游和下游连接处分别列于下表1和如序列表所示。跨人内含子3-lox-neo表达盒连接处的序列(所述表达盒的5'端)如SEQ ID NO:55所示,和跨lox-neo表达盒-人内含子3连接处的序列(所述表达盒的3'端)如SEQ ID NO:56所示;这两条序列也列于表1中。图1中显示的包括pgk-neo表达盒的完整的人源化CD4片段的核酸序列如SEQ ID NO:3所示。所述Pgk-neo表达盒跨SEQ ID NO:3的残基307-2176,两个lox位点位于残基267-300和2182-2215,而人序列跨残基1-234和2222-18263。完整的人源化CD4蛋白的氨基酸序列如SEQ ID

NO:4所示,其人序列跨氨基酸27-319(如SEQ ID NO:57所示)。

[0141] 表1:嵌合的CD4靶向载体的连接处序列

连接	序列	SEQ ID NO
5'小鼠/ 人连接处	AGGGGAAACCCGCAAAGGATGGGACATAGGGAGACAGCTG TTAACATCTGAAACATGACCTTCTTTTCTGTGCAGCACAACT CCTAGCTGTCACTCAAGGG(AAGAAAGTGGTGCTGGGCAAA AAAGGGGATACAGTGGAAGTACCTGTACAGCTTCCCAGAA GAAGAGCATACAATTCCACTGGAAAACTCCAACCAGAT)	1
3'人/ 小鼠连接处	(CTGGTCACCTGGATGAAGTGAGGGAGGGCCCTCTGGGTTTG GGGCTGGTTTTGAAGTGAGACATCCATGAGCCAGCCTGGGG CTGGCTTCACTGAAGATC) <b>ATCTATGTCGGGTGCGGAGAA</b> <b>AGAGGTAATGAAATGGCACATGCTATGTACAAACTCTATT</b> GCTGAGCAGCACCCAGTCTGAGCTGGCTCTGAATTGAGGG TGAAATTCACACATTCTCCCCAACATCTATAATCTGG	2
人/5' lox 位点	(TATGGAGTGAAAGCCTTTGGTGTCTGAGATCTGGTCTTAGT TAAACTCTGGGATC) <i>GGCGCGCCGAATTCCTGCAGCCCGGGCT</i> <i>CGAGATAACTTCGTATAATGTATGCTATACGAAGTTATATGCATC</i> <i>CGGGTAGGGGAGGCGCTTTTCCC</i>	55
3' lox 位点/人	<i>AGTATTGTTTTGCCAAGTTCTAATTCATCAGACCTCGACCTGCA</i> <i>GCCCTAGATAACTTCGTATAATGTATGCTATACGAAGTTATCCTA</i> <i>GG(CCAGAGGGCTTGGGTTGACAGAACTCAGTGGCATTCTT</i> <i>ATCCAGAGTTTCTCTACACC)</i>	56

[0143] 人序列在括号中和含有限制性内切酶位点 (PI-Sce I) 的序列以粗体表示。选择性表达盒序列以斜体表示。

[0144] 使用NotI将人CD4靶向载体线性化并且经电穿孔引入F1H4小鼠ES细胞。利用使用改良的等位基因测定 (Valenzuela等) 进行基因分型鉴定携带人源化CD4基因座的经靶向的ES细胞,所述测定检测到存在新霉素表达盒和人CD4基因,以及一个拷贝的小鼠CD4基因。在该测定中使用的引物和探针如表2所示并且列于序列表中。

[0145] 表2:用于对嵌合的CD4进行基因分型的引物和探针

检测区域	测定	5'引物	3'引物	探针
小鼠 CD4 内含子 3 (1765m2)	LOA	GAAGTGGGTGTG CCATTCAGA (SEQ ID NO:5)	AAAGCTCAGAAG CAGACAGAGTCA (SEQ ID NO:6)	TTCCAAAAGCCTA CAGCAGGCCAG (SEQ ID NO:7)
小鼠 CD4 内含子 5 (1765m4)	LOA	TCATCTCCCCTTC CTGAACCT (SEQ ID NO:8)	CCCAGCCACAAG AAGAAGAAA (SEQ ID NO:9)	CTTCCCCGCATCC ATTTTCTGTTC (SEQ ID NO:10)
在外显子 3 和 4 之间 的人 CD4 内含子 (在 pgk-neo 下游) (1765h1)	GOA	GGTCTCGAACTCA TGAGCTCAA (SEQ ID NO:11)	GGCATAGTGACA CACACCTGTAATT (SEQ ID NO:12)	TGATCCACTCACCT TGGCCTCTCAGAG (SEQ ID NO:13)
在外显子 5 和 6 之间 的人 CD4 内含子 (1765h2)	GOA	GTCAGGGAGCTT ACTTCTTTGTTG (SEQ ID NO:14)	TGTTAGTGTCCCT GAGTAAGTGGAT T (SEQ ID NO:15)	CTCAGCTCCACAC CCCTACCAAGTTG G (SEQ ID NO:16)

[0148] 通过电穿孔将表达Cre重组酶的质粒转移至含有人源化CD4基因座的ES细胞中除去翼侧为loxP的新霉素抗性表达盒。

[0149] 利用基因分型鉴定携带人源化的CD4基因座且无抗性标记物的被靶向的ES细胞,

其检测不到新霉素表达盒的存在,存在一个拷贝的人CD4基因和一个拷贝的小鼠CD4基因。

[0150] 使用上文所述的被靶向的ES细胞作为供体ES细胞并通过 **VELOCIMOUSE®**方法(参见例如美国专利号7,294,754和Poueymirou等,(2007) F0 generation mice that are essentially fully derived from the donor gene-targeted ES cells allowing immediate phenotypic analyses(能够直接进行表型分析的基本上完全来源于供体基因靶向ES细胞的F0代小鼠)Nature Biotech.25 (1):91-99)将其引入8-细胞阶段的小鼠胚胎。通过使用检测独特的人CD4基因序列存在情况的改良的等位基因测定(Valenzuela等,如上所述)进行基因分析鉴定得到 **VELOCIMICE®**(完全来源于供体ES细胞的F0小鼠)独立地携带嵌合的CD4基因。

[0151] 实施例1.2:嵌合的CD4在基因工程小鼠中的表达情况

[0152] 使用胶原酶D(Roche Bioscience)灌流来自WT或杂合的人源化CD4小鼠(“1766HET”)的脾脏并使用ACK裂解缓冲液裂解红细胞。分别使用抗-人CD4或抗-小鼠CD4抗体通过FACS分析人CD4或小鼠CD4在细胞表面的表达情况。如图2A和2B所示,在来自针对本申请所述的人源化CD4的小鼠杂合子的T细胞表面表达人CD4。

[0153] 实施例2:基因修饰的CD8小鼠的构建和表征

[0154] CD8蛋白以二硫键连接的两个亚基的同源二聚体(例如CD8 $\alpha$ 的同源二聚体)或同源多聚体,或者两个蛋白CD8 $\alpha$ (CD8a)和CD8 $\beta$ (CD8b)的异源二聚体出现。CD8 $\alpha$ 和CD8 $\beta$ 基因共同定位于基因组中,例如在小鼠的6号染色体上,其彼此之间距离约37kb。这种紧密连锁使其非常难以通过繁殖产生在CD8 $\alpha$ 和CD8 $\beta$ 基因座均包含人源化的基因修饰的小鼠。因此,进行通过先引入一个基因例如CD8 $\beta$ ,再引入第二个基因例如CD8 $\alpha$ 的顺序靶向。

[0155] 实施例2.1:工程化制备嵌合的CD8 $\beta$ 基因座

[0156] 通过使用**VELOCIGENE®**技术(参见例如美国专利号6,586,251和Valenzuela等,(2003)High-throughput engineering of the mouse genome coupled with high-resolution expression analysis(与高分辨率表达分析偶联的小鼠基因组的高通量工程改造),Nature Biotech.21 (6):652-659)在一步中从小鼠的细菌人工染色体(BAC)DNA构建独特的靶向载体将小鼠CD8 $\beta$ 基因座人源化。利用BHR对来自BAC RP23-431M6的DNA进行修饰以产生较大的靶向载体(LTVEC)MAID 1737,以含有编码CD8胞外域的小鼠外显子2-3(从在内含子1中的5'连接处至在内含子3中的3'连接处)被同源的人序列的取代(图3)。将loxP-Ub-Hyg表达盒插入内含子3中的3'连接处。在所得到的载体不同连接处的核苷酸序列列于表3并且如序列表所示。完整的人源化CD8 $\beta$ 蛋白的氨基酸序列如SEQ ID NO:53所示;其具有跨距氨基酸15-165的人序列(如SEQ ID NO:58所示)。

[0157] 表3:嵌合的CD8 $\beta$ 靶向载体的连接处序列

连接处	序列	SEQ ID NO
在内含子 1 中的小鼠/人	TGTTTGCCTGTGACATGAACCTATTGTGACACAAACCACTGTGCTAGGGGGGATCCACTAGTAACGGCCGCCAGTGTGCTGGAATTCGCCC(TCGCAAGGGCCAGGCATATAAGTACACAATAAAACAAATGGCAGCTCTCTCC)	17
在内含子 3 中的人/lox 位点的 5'端	(CCCCCTCCTTCCCTCCCCAGGCACCTTTCCAAGTGTCAACTCTAGAGCCTAT)CGCGGCCGCACCGGTATAACTTCGTATAATGTATGCTATACGAAGTTAT	18
在内含子 3 中的 lox 位点的 3'端/小鼠	ATAACTTCGTATAATGTATGCTATACGAAGTTATGTCGACGTAGCCTATTTCTCTAGATCCAAAATGATGACAACAAAAGGTACCTTGTG	19

[0159] 人序列在括号中, lox位点以斜体表示以及限制性内切酶位点、多重克隆位点和来源于载体的序列以粗体表示。

[0160] 将靶向载体经电穿孔引入F1H4小鼠的ES细胞(图5, 左图)。通过使用改良的等位基因测定 (Valenzuela等) 进行基因分型鉴定携带人源化CD8b基因座的被靶向的ES细胞, 其检测到人CD8b基因的存在。在该测定中使用的引物和探针如表4所示并且列于序列表中。

[0161] 表4: 用于对嵌合的CD8 $\beta$ 进行基因分型的引物和探针

检测区域	测定	5'引物	3'引物	探针
小鼠外显子 2	LOA	GCAGCTCTGCCCTCATTCAG (SEQ ID NO:20)	CATCTTTGCCGTA TGGTTGGT (SEQ ID NO:21)	CCCCCTCGTCCCTG CTGGTTCA (SEQ ID NO:22)
小鼠外显子 3	LOA	CAAGAAGACTACCCTGAAGATGAAGA (SEQ ID NO:23)	TGTGAGTGCAACA ATGGAAACT (SEQ ID NO:24)	CGTCCCCCACCCA GAGACCCA (SEQ ID NO:25)
人外显子 2	GOA	GGCACCGAGCAGTGACAGT (SEQ ID NO:26)	TTCACCGTGGATA GTCCCTTTT (SEQ ID NO:27)	AGTTCCTGGCCCTC TGGGATTCCG (SEQ ID NO:28)
人外显子 3	GOA	TTGCTTTCTTTCTGTAGTTGATTTC (SEQ ID NO:29)	CCGGCACACTCTCTTCTTGAG (SEQ ID NO:30)	TCCACCACTGCCAGCCCA (SEQ ID NO:31)

检测区域	测定	5'引物	3'引物	探针
Hyg	GOA	TGCGGCCGATCTTAGC C (SEQ ID NO:32)	TTGACCGATTCTT TGCGG (SEQ ID NO:33)	ACGAGCGGGTTTCG GCCCATTC (SEQ ID NO:34)

[0164] LOA=缺失等位基因; GOA=获得等位基因。

[0165] 使用上文所述的被靶向的ES细胞作为供体ES细胞并通过 VELOCIMOUSE®方法 (参见例如美国专利号7,294,754和Poueymirou等, (2007) F0 generation mice that are essentially fully derived from the donor gene-targeted ES cells allowing immediate phenotypic analyses (能够直接进行表型分析的基本上完全来源于供体基因靶向ES细胞的F0代小鼠) Nature Biotech.25 (1):91-99) 将其引入8-细胞阶段的小鼠胚胎。通过使用检测独特的人CD8b基因序列存在情况的改良的等位基因测定 (Valenzuela等, 如上所述) 进行基因分析鉴定得到 VELOCIMICE® (完全来源于供体ES细胞的F0小鼠) 独立地携带嵌合的CD8b基因。

[0166] 可以采用本领域技术人员公知的方法除去选择性表达盒。例如, 可以通过使用表达Cre的构建体转染携带嵌合的人/小鼠CD8b基因座的ES细胞以除去通过插入含有人CD8b

基因序列的靶向构建体引入的两端侧翼为LoxP的潮霉素表达盒。可以任选地通过与表达Cre重组酶的小鼠繁殖除去潮霉素表达盒。任选地,将所述潮霉素表达盒保留在所述小鼠中。在一个实施方式中,将所述表达盒缺失以产生MAID 1739。

[0167] 实施例2.2:工程化制备嵌合的CD8 $\alpha$ 基因座

[0168] 通过使用VELOCIGENE®技术(参见例如美国专利号6,586,251和Valenzuela等,如上文所示)在一步中从小鼠的细菌人工染色体(BAC) DNA构建独特的靶向载体将小鼠CD8 $\alpha$ 基因座人源化。利用BHR对来自BAC RP23-431M6的DNA进行修饰以产生较大的靶向载体(LTVEC) MAID 1738,以含有编码CD8 $\alpha$ 胞外域的小鼠外显子1-2(从小鼠外显子1中的Ala密码子27的5'连接处至在小鼠内含子2中的3'连接处)被同源的人序列的取代(从在人外显子2中的5'连接处至在内含子3中的3'连接处(图4))。其保留了在外显子1起始处的小鼠先导序列。在人/小鼠序列的3'连接处插入lox2372-Ub-Neo表达盒。在所得到的载体不同连接处的核苷酸序列列于表5并且如序列表所示。完整的人源化CD8 $\alpha$ 多肽的氨基酸序列如SEQ ID NO:54所示,其具有人序列跨氨基酸28-179(如SEQ ID NO:59所示)。

[0169] 表5:嵌合的CD8 $\alpha$ 靶向载体的连接处序列

[0170]	连接处	序列	SEQ ID NO
	在外显子 1 (小鼠) 和外显子 2 (人) 的小鼠/人	TGAACCTGCTGCTGCTGGGTGAGTCGATTATCCTGGGGAGTGAGAAAGCT(AGGCCGAGCCAGTTCCGGGTGTCGCGCTGGATCGGACCTGGAACCTGGG)	35
[0171]	人/lox 2372 位点的 5'端	(ATGCCAGGGACAGCCCTGATACTGTAGGTAGAGTCAAGGGCTGTCCAAGT) <b>ACCGGTATAACTTCGTATAAGGTATCTATACGAAGTTAT</b>	36
	lox 2372 位点的 3'端/小鼠	<b>ATAACTTCGTATAAGGTATCCTATACGAAGTTATCTCGACC</b> TGATCTTGGAGGGAGACCTGGACCGGGAGACGTGCTGGGGCAGGGTT	37

[0172] 人序列在括号中,lox位点以斜体表示以及限制性内切酶位点、多重克隆位点和来源于载体的序列以粗体表示。

[0173] 将上文所述的人源化的CD8 $\alpha$ 靶向载体经电穿孔引入含有人源化的CD8 $\beta$ 基因座的小鼠ES细胞以形成包含人源化的CD8 $\beta$ 和CD8 $\alpha$ 基因座的经修饰的ES细胞(图5)。通过使用改良的等位基因测定(Valenzuela等)进行基因分型鉴定携带人源化CD8 $\alpha$ 基因座的被靶向的ES细胞,其检测到人CD8 $\alpha$ 基因的存在。在该测定中使用的引物和探针如表6所示并且列于序列表中。

[0174] 表6:用于对嵌合的CD8 $\alpha$ 进行基因分型的引物和探针

[0175]

检测区域	测定	5'引物	3'引物	探针
小鼠外显子 1	LOA	GATGCTCTTGGCTCT TCCAGAA (SEQ ID NO:38)	ATGAAGCCATATAG ACAACGAAGGT (SEQ ID NO:39)	CCAGCTCCAAACTCC CCCAGCC (SEQ ID NO:40)
小鼠外显子 2	LOA	TCAGCCCCAGAGAC CAGAAG (SEQ ID NO:41)	TCAATCGCTTGAGA GCACCTAA (SEQ ID NO:42)	TTGTCGGCCCCGTGG CTCA (SEQ ID NO:43)
人外显子 2	GOA	GCGGTTCTCGGGCA AGA (SEQ ID NO:44)	TCAGGGCCGAGCAG AAATAG (SEQ ID NO:45)	ACACCTTCGTCCTCA CCCTGAGCGA (SEQ ID NO:46)
人内含子 2	GOA	GGTTCACCTCAACCT GTTTTCC (SEQ ID NO:47)	CGCTTCCAGGTGCG CTAA (SEQ ID NO:48)	ACCTGGGCCCTGCTT TCAAGCC (SEQ ID NO:49)
Neo	GOA	GGTGGAGAGGCTAT TCGGC (SEQ ID NO:50)	GAACACGGCGGCAT CAG (SEQ ID NO:51)	TGGGCACAACAGAC AATCGGCTG (SEQ ID NO:52)

[0176] LOA=缺失等位基因;GOA=获得等位基因。

[0177] 使用上文所述的被靶向的ES细胞作为供体ES细胞并通过 **VELOCIMOUSE®** 方法 (参见例如美国专利号7,294,754和Poueymirou等,如上文所述) 将其引入8-细胞阶段的小鼠胚胎。通过使用检测独特的人CD8b和CD8a基因序列存在情况的改良的等位基因测定 (Valenzuela等,如上所述) 进行基因分析鉴定得到 **VELOCIMICE®** (完全来源于所述供体ES细胞的F0小鼠) 携带嵌合的CD8b基因和嵌合的CD8a基因。

[0178] 或者,将本申请所述的人源化的CD8a靶向载体经电穿孔引入不含有源化的CD8b基因座的小鼠ES细胞中以产生仅包含人源化的CD8a基因座的小鼠。

[0179] 可以采用本领域技术人员公知的方法除去CD8a基因座中的选择性表达盒,例如上文的实施例2.1中所述的。在一个实施方式中,缺失选择性表达盒以产生MAID 1740小鼠。

[0180] 实施例2.3:嵌合的CD8在基因工程化的小鼠中的表达

[0181] 针对人CD8的表达情况对针对人源化的CD8a (MAID 1740) 和CD8b (MAID 1739) 基因的小鼠杂合子进行分析。

[0182] 在来源于杂合子脾细胞的CD3+CD4-T细胞表面上能够清楚地检测到人CD8a和CD8b的表达,但是在野生型动物中不能检测到 (图6)。

[0183] 在从杂合子获得的胸腺细胞表面上也能够检测到人CD8a和CD8b的表达,但是在野生型动物中不能检测到 (图7)。

[0184] 等同物

[0185] 本领域技术人员将认识到,或者能够使用不超过常规的实验确定本申请所述的发明的特定实施方式的多种等同物。这类等同物旨在包含在下述权利要求中。

[0186] 本申请通篇引用的非专利文件、专利申请和专利的全部内容通过引用整体并入本申请。

## 序列表

	<110> 瑞恩泽制药公司	
	<120> 人源化的T细胞共受体小鼠	
	<130> 2010794-0441	
	<150> 61/766, 762	
	<151> 2013-02-20	
	<150> 61/890, 915	
	<151> 2013-10-15	
	<160> 59	
	<170> PatentIn 3.5版	
	<210> 1	
	<211> 202	
	<212> DNA	
	<213> 人工的	
	<220>	
	<223> 合成的	
	<400> 1	
[0001]	aggggaaacc cgcaaaggat gggacatagg gagacagctg ttaacatctg aaacatgacc	60
	ttcttttctg tgcagcaciaa ctctagctg tcaactcaagg gaagaaagtg gtgctgggca	120
	aaaaagggga tacagtggaa ctgacctgta cagcttccca gaagaagagc atacaattcc	180
	actggaaaaa ctccaaccag at	202
	<210> 2	
	<211> 240	
	<212> DNA	
	<213> 人工的	
	<220>	
	<223> 合成的	
	<400> 2	
	ctgggtcacct ggatgaagtg agggagggcc ctctgggttt ggggctggtt ttgaactgag	60
	acatccatga gccagcctgg ggctggcttc actgaagatc atctatgtcg ggtgcggaga	120
	aagaggtaat gaaatggcac atgctatgta caaactctat tgctgagcag caccagtc	180
	tgagctggct ctgaattgag ggtgaaattc acacattctc cccaacatc tataatctgg	240
	<210> 3	
	<211> 18263	

	<212> DNA	
	<213> 人工的	
	<220>	
	<223> 合成的	
	<220>	
	<221> misc 特征	
	<222> (2957)..(2957)	
	<223> n是a, c, g, or t	
	<220>	
	<221> misc 特征	
	<222> (3193)..(3193)	
	<223> n是a, c, g, or t	
	<400> 3	
	aagaaagtgg tgctgggcaa aaaaggggat acagtggaac tgacctgtac agcttcccag	60
	aagaagagca tacaattcca ctggaaaaac tccaaccaga taaagattct gggaaatcag	120
	ggctccttct taactaaagg tagggttgcc tggctcccca tccaggaggagg aaaacacact	180
	atggagtgaag agccttttgt gtctgagatc tggctcttagt taaactctgg gatcggcgcg	240
	ccgaattcct gcagcccggg ctcgagataa cttcgtataa tgtatgctat acgaagttat	300
[0002]	atgcatccgg gtaggggagg cgcttttccc aaggcagtct ggagcatgcg ctttagcagc	360
	cccgttgggc acttggcgct acacaagtgg cctctggcct cgcacacatt ccacatccac	420
	cggtaggcgc caaccggctc cgcttcttgg tggcccttc gcgccacctt ctactcctcc	480
	cctagtcagg aagtcccccc ccgccccgca gctcgcgtcg tgcaggacgt gacaaatgga	540
	agtagcacgt ctactagtc tcgtgcagat ggacagcacc gctgagcaat ggaagcgggt	600
	aggcctttgg ggcagcggcc aatagcagct ttgctccttc gctttctggg ctgagaggt	660
	gggaaggggt ggggtccggg gcgggctcag gggcgggctc aggggcgggg cgggcgcccg	720
	aaggtcctcc ggaggcccgg cattctgcac gcttcaaaag cgcacgtctg ccgcgtgtt	780
	ctcctcttcc tcatctccgg gcttttcgac ctgcagccaa ttgttgacaa ttaatcatcg	840
	gcatagtata tcggcatagt ataatacgac aaggtgagga actaaaccat gggatcggcc	900
	attgaacaag atggattgca cgcaggttct ccggccgctt gggtggagag gctattcggc	960
	tatgactggg cacaacagac aatcggtgc tctgatgccg ccgtgttccg gctgtcagcg	1020
	caggggcgcc cggttctttt tgtcaagacc gacctgtccg gtgccctgaa tgaactgcag	1080
	gacgaggcag cgcggctatc gtggctggcc acgacgggcg ttccttgcgc agctgtgctc	1140

[0003]	gacgttgtca ctgaagcggg aagggactgg ctgctattgg gcgaagtgcc ggggcaggat	1200
	ctcctgtcat ctcaccttgc tcctgccgag aaagtatcca tcatggctga tgcaatgcgg	1260
	eggctgcata cgcttgatcc ggctacctgc ccattcgacc accaagcgaa acatcgcatc	1320
	gagcgagcac gtactcggat ggaagccggt cttgtcgatc aggatgatct ggacgaagag	1380
	catcaggggc tcgcgccagc cgaactgttc gccaggctca aggcgcgcat gcccacggc	1440
	gatgatctcg tcgtgaccca tggcgatgcc tgcttgccga atatcatggt ggaaaatggc	1500
	cgcttttctg gattcatcga ctgtggccgg ctgggtgtgg cggaccgcta tcaggacata	1560
	gcgttggcta cccgtgatat tgctgaagag cttggcggcg aatgggctga ccgcttcctc	1620
	gtgctttacg gtatcgccgc tcccattcg cagcgcacgc cttctatcg cttcttgac	1680
	gagttcttct gaggggatcc gctgtaagtc tgcagaaatt gatgatctat taaacaataa	1740
	agatgtccac taaaatggaa gtttttctg tcatactttg ttaagaaggg tgagaacaga	1800
	gtacctacat tttgaatgga aggattggag ctacgggggt gggggtgggg tgggattaga	1860
	taaatgcctg ctctttactg aaggctcttt actattgctt tatgataatg tttcatagtt	1920
	ggatatcata atttaaaca gcaaaaccaa attaagggcc agctcattcc tcccactcat	1980
	gatctataga tctatagatc tctcgtggga tcattgtttt tctcttgatt cccactttgt	2040
	ggttctaagt actgtggttt ccaaattgtg cagtttcata gcctgaagaa cgagatcagc	2100
	agcctctgtt ccacatacac ttcatctca gtattgtttt gccaagttct aattccatca	2160
	gacctcgacc tgcagcccta gataacttcg tataatgtat gctatacgaa gttatcctag	2220
	gccagagggc ttgggttgac agaaactcag tggcattctt atccagagtt tctctacacc	2280
	aactgctggt ggcccaggga aaggtggtat gtgaatttca atattttaat atttaatat	2340
	catgaactta ttttagtgag ttttagaaca atcactatca cttaaaaccc gtgatttctt	2400
	gagtattgtt gctacagacc tatgtagata atactttgca cagtgactca tatgtataat	2460
	cctagcactg tgggaggctg aggccggagg attgcttgag tccaggagtt caagaccagc	2520
	ctgaacaaca tagtgagact ctgtctctat gaaaaaaaat atatatatat ttttttgga	2580
	gacaaggtct agttctatca cccaggctcc agtgcagtgg tgtgatctcg gctcactgca	2640
	atctccacct cccaggetca agtcatcatc ccacctcagc ctcccaagta gctgggacta	2700
	caggcatgca ccacatgcc aggctaattt ttgtattttt tatagagaca gggtttcacc	2760
	atgttggcca ggctggtctc gaactcatga gctcaagtga tccactcacc ttggcctctc	2820

[0004]	agagtgcctgg aattacaggt gtgtgtcact atgcctagcc aaaaaaaatt tttttaatta	2880
	aaaaaaaaa ggccggctgt agtggctcac acctgtaatc cagaactttg ggagtttgag	2940
	gtgggcagat caccgnggt caggagtca agaccagtct ggccaacatg gtgaaacccg	3000
	gtctctacta aaaatacaaa aattagccag gtgtgggggt gcagtcctgt acttcagct	3060
	actcaggagg ctgaggcagg agactcgctt gaacctggga ggcaaaggct gcagtgaagt	3120
	gagattgcac cactgcactc cagcctgggt gacagagcaa gacttcatct caaaaaaaaa	3180
	aaaaaagctg canatttatt attattatta ttagtttatt tattttattt ttgagacag	3240
	agtctcgttc tgtcgcccag gctggagtgc ggtggcgtga tcttggtca ttgcaacctc	3300
	cacctcccg gttcaagtga ttctcctgcc tcagcctccc gagtagctgg gactacaggc	3360
	gtatgccacc atgcctggct aattttttgt acttttagta gagacagagt ttcacggtgt	3420
	tagccaggt ggtcttgatc tcctgacctc gtgatttacc ctcttggtcc tcccaaagt	3480
	ctgggattac aggctgagt cactgtgcc ggccagaat catTTTTTt tactTTTTt	3540
	TTTTtgaggc aaactctga tctgttccc aggctggagt gcagtgggca tgatcttggc	3600
	tactgcaag ctctgcctcc caggttcaag caattctct gcctcagct cctgagtagc	3660
	tgggactaca ggcgtgtgcc accatgccc gctaatttgc gtatttttag tagagaccgg	3720
	ttttcatcat attggccagg ctggtcttga actcctgacc tcaagtgatt ctcccacctt	3780
	agcctccaa agtgcctggga ttacaggcat gagctactgc acttggcctt ttctcctggt	3840
	tttaaaacta ttatatgctc attacaaaat atttggtaaa tgaagaaaag aatatggaag	3900
	aaaatcaaat gcatgcatac ttctatcact cagagataac ctctgctaac attttgattg	3960
	attttcttcc aatctTTTTt tttttttt tttttgagac aggtctcac tctgctgcc	4020
	aggctggagt acagtggcat gaccacaaca catcacagcc tcaagtgatc ttcccacttc	4080
	agccttccca gtagctggga ctacaggtgc acgccacat gttcacctaa ttttttactt	4140
	ttttagaga tgagacttca ccatgttget caggctggtc ttgaattcct aggtcaagt	4200
	gatcttccc ctttggcctc ccaaagtgt gggattatag gtatgagcca ctgcatgtgg	4260
	cctattttct tccactgttg ttcggcgtgg agaataatt atacataatt acgtaaata	4320
	tatcatactg tatatacctt ttttctact ctttcttaa gttatatcat aatgagacta	4380
	ccaattatta gactTTTTt cttttttt agacggagtc tcggtctgtc acctaggctg	4440
	gagtgaatg gcgcgatctc agctcgtgc aacctctgcc tcccaggttc aagcaattct	4500

[0005]	gcctcagcct cccgagtagc tgggactaca gacacgtgcc accatgccc gctaactttt	4560
	ttattttttt attagagaca gggttccacc atgctagcag gatgggtctca atctctcgac	4620
	ttcgtgatca gcccggttg gcctcccaaa gtgctgggag tacagggtgtg agccaccgca	4680
	ctcggcctag actaactatt taaagtaatc tggcaatgtt taacgaatac aaaactctaa	4740
	aacccttgga cctaataata gctatttttg aaagtctact tgacagaaat aaaattgtga	4800
	atattctttt ttgttgtttt tttagacag agtctcattt ggacgcctag gctggagtgc	4860
	agtggcatga tctcggttaa ctgcaacctc cacctcctgg gttcaagtga ttctcctgcc	4920
	tcagcctcct gagcagctgg gattacaggt gtgcaccacc atgtctggct aatttttgca	4980
	tttttagtag atggggtttc accatgttga ccagggtgggt ctggaacttc taccctcaag	5040
	tgatctaccc accttggcct cccaaagtgc tgggattaca ggtgtgagcc accacgcctg	5100
	accagtgaac acttaataat atctatggaa aggtgttatt ataagaattg cttgtggggc	5160
	cgggcgtgggt ggctcacgcc tgtaatccca gcactttggg aggctgtggc aggcggatca	5220
	cgaggtcagg agatcaagat catcctggct aacacggtga aaccccgctct ctactaaaaa	5280
	taccaaaaaa ttagccaggc gtggtggcgg gcacttgtaa tcccagctat ccaggaggct	5340
	gaggcaggag aattgcgtga acccaggagg cggaggtcgc agtgagctga gaccgtgcca	5400
	ttgcactcca gcctgagtga cagagtgaga ctccatcaca aaaaataaat aaataaataa	5460
	ataaaatata aataagtaaa taaaggtcag gagtgtgtgg tcacgcctgt aatcccagca	5520
	ctttgggagg ccgaggtgga cagatcatga ggtcatgaga tcaagaccat cctggctaac	5580
	acagtgaac cctgcctcta ctaaaaatac aaaaagtcac ccaggtgtgg tggcacacac	5640
	ctatagtccc agctacttgg gaggtgagg caggagaatc acttgaaccc aggaggcaga	5700
	ggttgcagtg agctgagatc gcgccactac actccagcct aggcgacaga gcaagactct	5760
	gtctcaaaat aaataaataa ataaataaat aaataaataa ataaataaaa taaaaagcac	5820
	acacacacac acacacacac acacacaatg caaaagaccc accctactac aactaacatt	5880
	atatttaatg gtgaaaaact gaattctttc tccctaagt caggaataag acaaagatgt	5940
	ctgctcttac tactcttatt caacataata ctgcaatccc ttgccagtgc aataaggcaa	6000
	gaaaaatgaa ataaaaggaa aactgatcag aaagaaagaa ataaaactgt tcctatttgt	6060
	ggatgacatg attacataga aaatctcaaa gaatctgtaa gaaacttctt agaattaata	6120
	aatgaattca tcaaggttgc agaataaag ataaacataa aaaatctatt gtatttctat	6180

[0006]	atattagcaa ggaacatgtg tacacagaaa ttaaaactac aataccattt ataattgctc	6240
	aaaaaggcca ggcattggtg ctcacacctg taattcctgc actttgggag gccaaaggtg	6300
	gaagattgct taagcccagg agttcaagac cagcccgggc aacatagtga gacctgtct	6360
	ctacaaaaag taaaaaatta gctgagcatg gccgggtgca gtggctcact cctgtaaccc	6420
	caacactttg ggaggctgag gcgggcggat catgaggtca ggagatcgag accatcctgg	6480
	ctaacacggt gaaaccctgt ctctactaaa aacacaaaaa attagctgga tgtggtggca	6540
	ggcgctgtga gaccagcta ctcgggaagc tgaggcagga gaatggcgtg aacctgggag	6600
	gcggagcttg cagtgagctg agattgtgcc actgcactcc agcctgggtg acacagtgag	6660
	actacgtctc aaaaaaaaaa aaaaaaaaaat tagctgagca ttatggtgta tgcctgtagt	6720
	cccagctact ggggaggctg aggtgggagg attgcttgag ccctaggagg gcaaggctgc	6780
	agtgagccat gatcacacca ctgctttcca gcctcggtag gagagcaaga ccctatctca	6840
	aaaaaaaaaa aaaaaaaaaa agaaaagaaa agaaaagaaa agaaaaagaa agagagaaag	6900
	aaatacttag gtgtaaatct aaaaaacatg cgtagggcca ggtgcagtgg ctcatgcctg	6960
	taatcccagc actttgggaa gttgaggctg gcggatcact tgaagtcggg agtttgagac	7020
	cagcctggcc aacatggtga aaccccgtct ctactaaaaa tgcaaaaatt aggaggtgt	7080
	tgtggcgcat gcctgatccc agctactttg gaggtgagg caggagaatt gcttcaaccc	7140
	gggaggcaga ggttcagtg agccaagact gttccactgc actccagcct gggcaacaga	7200
	gtaagagtct gtctcccgaa aaaaaaaaaa agaaaaaaga aagcattgaa ttgtatgcta	7260
	aaaactacac gatgctgatt aaagaagtca aagaagatct aaatatatgg agagacatgc	7320
	tgtactcatg gattgatgga ttggaagact caacataaga cagatatcaa ttttcccaa	7380
	attaatatac aagtttaatc caattcctat aaaaatacca gcaagatttt ttgtagatat	7440
	aaacaagttg gccagggtga gtggcttaca cctgtaatcc tagcactttg ggaggctgag	7500
	gtgggaagat cgcttgagcc caggtgttca cgactgcagt gagctatgat tgtgtcactg	7560
	cattccagct ggcactccag cctaagtgaac aaaggagagac cctgtctcaa aaacaaaaac	7620
	aaaacaaaaa taattttgct ctgcaaaatc cctattaaga agaagaaaag aggctgggca	7680
	cagtggctca ccgctgtaat ccagcacgt tgggaggctg aggcaggctg atcacttcag	7740
	cccagaagtt tgagatcagc ctgggcaaca tgaggaaacc ccgtctctac caaaaaaaaa	7800
	aaaaggtaca tacacacaca cacacacaca cacacacata cacaagtata tacacatata	7860

	tatacacata caggtgaata gatgtatata catctattta ttgtgaatat acatctatac	7920
	acacacgtgt gtgtacacat atatttataaa tttattttta tttatttatt tatttttgag	7980
	acagagtctt gctctgtcac ccaggctggg tgcacctgta ttcccaacga cacaggaggc	8040
	tgaggtggga gaatcactga gccaggaggag cagaggttgc agtgagccaa gatgttgcct	8100
	ggttgcctgg gcaacagagc gagaccctat atcaaaaaag aagaataata agaaaagaca	8160
	gtttacagaa tataagaaaa tatattcaca atccacatac ttagcaaagg actggtatct	8220
	agaatatgat aaacaactct caaaactcaa aaccaaaaaa atgaacaatt caattagaaa	8280
	acaggccgaa aaggacatac agttggcaaa taagcacatg aaaagttgtt caacatcatt	8340
	aatcattagg gatattgtaca ttaaaaccac aataggctat cactaaacct atcagaatgg	8400
	ctaaatacaa aattggaaca ccaccaaag ctgatgagga tgtggagaaa ctgggtcatt	8460
	cttccaatat tgggtgggagg ctaaaatggc aaagccactc tggaaaacag tttgatagtt	8520
	tcttataaaa caaaacatgc ggccgggcgc ggtagctcac gcctgtaate ccagcacttt	8580
	gggaggccga ggccgggtgga tcacagggtc aggagatcga gaccatcctg gctaacacgg	8640
	tgaaaccctg tctctactaa aaatacaaaa aattagccgg gcgtggtggc gggcgcctgt	8700
[0007]	agtccagct actcgggagg ctgaggcagg agaattggtgt gaaccggga ggccggagctt	8760
	gcagttagcc gagatcgcc cattgcactc caacctggga gacggaggga gactccgtct	8820
	caaaaaaaca aaaacaaca acaaaaaaac atgcaacaat ccagcaatat tgcacccta	8880
	ggcatttatc ctagagcaat gaagacttat gccacacaa aaagctgcac acaaatgttc	8940
	atagcagctt tattcatggt agccaacaat tagaaacaat ctagatgtcc ttcaactggt	9000
	gaatgattac atccatacca cgaaatactt ttcagcaata aaaaggatga atcatagtac	9060
	acaccacaac ctggatgaat ctccaggga ttatgtgag tgaaaaaag ccaatctcaa	9120
	aaghtaatat actgtattaa tccatttata taacattctt aaaataacta attatagaaa	9180
	tggagaacag atgagtgatt gccagggtt aaggggtca gggatggga ggggaagggg	9240
	tatggctaca aaaagcaaca accttatggc gccggaaatg ttctgtattc tgattgtgtc	9300
	aatgtgagca tactgggtga gatatagtgc tacagttttg caagttatta ccatcagagt	9360
	aaactggata gagggcacat aggatttctc tgtattactt cttacaactg caagtgaatc	9420
	tacaattatc tcaaaataat aagtttagtt taatgctagg cgtggtggct cacatctgta	9480
	atctcagctc tttgggaggc tgagacgggt ggatggcttg agtccaggag ttcgagacca	9540

[0008]	gcctggccaa catggcaaaa ccggtctcta ctaaaaatac aaaaattagc tgggcgtggt	9600
	ggcaagtgcc tgtagtccca gctactcggg aggctgaggc aggagaattg cttgaacccg	9660
	ggaggtggag gttgcagtga gccgagatca cgccactaca ctgtagcttg ggcgacagag	9720
	tgaggctctt tctcaaaaaa aaaaaaaaaa aaaaaaaagc aggcaggcag ggccaggaaa	9780
	gcgtataatt tttgtagttc aaatgactaa cctaaaaagt gaagattggc caggcgcagt	9840
	ggctcacgcc tgtaatccca gcactttggg aggccaaggc ggggtggatca cgaggtcagg	9900
	agattgagcc actctggcta acacagtga accccgtctc tactaaaata caaaaaatta	9960
	gctgggcgtg gtggcacccg cctgtagttg cagctacttg ggaggctgag gcaggagaat	10020
	cacttgaacc caggaggcga agttgcagcg agccgagatc acactactgc actccagcct	10080
	gggtgacaaa gtgagattct gtctcaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaagt	10140
	aagtttacct ttttttttaa atttttcttc ttttccttcc ctactttgtg agataatttt	10200
	cttcttttta aaaagccaag agcttacttc tgtaagtaaa gattatctta agacaactta	10260
	gaaatgtata ttattagtat tttctatttc attgtaagtt atttgtaaatt attggttttg	10320
	gtgctaacct agaattccat caaattaatt gtcccctaata atatggccat tatcattttg	10380
	tctaacattg tatcetatta acaatgctgt aagtattatt tttgtagcta aattatgggt	10440
	tgcattttta aattattgtt ttaaggataa agttccagaa atgaaattaa ggatatgaac	10500
	tttttgagca catcttgtca gcactgagta gtattattta aaacttttgg ggggggcaat	10560
	tttataattg aaaaatatat cattgtttta atttgcatth ctttactgc ctatgagatt	10620
	aaaacaatgc actactttcc aaaaattctt aagtcttttg tgttgatgct tttgttctgt	10680
	ttctatggat ctcatcttcc ttcagaacag ctccccttcc caacttctg atttctaaca	10740
	ataacagtat caccctcctt gttctcccaa tttctgaaac acagagtcatt gttttttct	10800
	ctgcttcaat ccctggtttc ctatcgatc caattatgac ctttcttgc tttgaaagt	10860
	ttttgggccc ggcatgatgt ctgccacctt ttgtaatect agcacttttg gaggtgagg	10920
	cggctggatg acttgacctg aggatcttga gaccagcctg ggcaacaggg cgaaacctcg	10980
	tctctacaaa aaatacaaaa gttagtcggg agtggtggca catgcttgta gtcccagtta	11040
	cttggggggc tgaggtggca ggatctcttg agcccacgag gtagatgttg cagtgagccg	11100
	tgattgcgcc actgcacccc agcctaggtg acagagttag accctgtctc aaaaaaaaaa	11160
	aaatgttcta gtttcttctt cttctttgtt cccatgggaa tgccaccatc accagccaag	11220

[0009]	gctcacatac ctcccacctg gattacagtg agcttccagg taatttggtc tgctactagt	11280
	ctcgectact tggatttccc ttccccctgc tgcagcattg ccttccaaag ccatgctttg	11340
	cacatgccac atcctagccc attagactaa gcctagaagc ctctgcagga cgttcacctt	11400
	ctcagcgcca ctgctcagtt tcccagtga aacctctgca cccaggaggt ttccccacag	11460
	cttgccctgtg ctgcctctct ggagcttttc tcccttctctg taatgtcctt gctgctcccc	11520
	gtctctagtc cattgcctat acctcttttt tttttttttt ttgagatgga gtctctctct	11580
	ctcatccagg ctggagtga gtggcgcat ctcggtcac tgcaaccttt gtctcctggg	11640
	ttcaagggat tctcctgcct cagcctcccg agtaactggg attacaggcg tgcaccacca	11700
	ttcctggcta atttttgtat ttttagtaaa gactgggttt caccatgttg gccaggctgg	11760
	tcttgaactc ctgccctcag gtgatccacc tgcctcggtc tcccagagtg ctgggattac	11820
	aggcgtgagc caccgcacct gccacaggcc catacctctt ttaagtcttc attcaatacc	11880
	agttgtccca tgaatttgct ccagactcac tcatatgctt agacctttca tattatcttg	11940
	ccatagcttt ttcaaagtat gggacagcat ggacaagcag gccatggttt tcttttgaag	12000
	agaagcaagg aggcagagtt attttaggag gagggttata catttcattt tgaaccaatt	12060
	gcgtttgggg tgatggcagg atattaacat aaacttattt cttggacctt tggaaatgtg	12120
	tgcttagaac tgaggagaga ggtcagggtt ggcagtaaca acttggccac aatctgcaga	12180
	gctgactggg gatgaggtgg aatttagaat gtctgtagaa acggggaaga gaaccaaaga	12240
	cagagtctgg gacaacacct aaatgtagat gtcagagcaa gagttcaaga cgaagaaaaa	12300
	cgaatcatac ttagaaatgg aggggaggaa caaaagaggc ggagcaaagt ggggcagaac	12360
	cagagtaggc cacgctttta agaagtttgg taaaggaact gtgaaaggaa tgtagttgaa	12420
	tttcagggta agctggggaa ttaaagcagt gtgtagatcc agggcaaaca gcaagtaggg	12480
	caggaaccac tgaaggaaaca aataaagggg gaggttgggt ccaggttgctc ttgagtaggg	12540
	aagttttttt aaaaagtgtg aaactgaagg tgtggggtgg attgggtgcc tgccgtgctc	12600
	tgaggaagct tggggcaact gtgtgctgag gctgtgaggt tgtctggaag gggctcctgg	12660
	acagtaagag ctgagcagtg gggaagagga ctgtgtggtc tggaagagga gagaaaggag	12720
	agtgagtgac tgaactggta tccaggtcc cacaccaagg cagaaagagg gagaggacct	12780
	gggcatctca gggaggcaga ggcagtacca agcagggtga gaggcttttag tcttagccac	12840
	ctttgcccc a ttccetcaaa tatacattct aagtaaaaac aaaacaaaac agaactgttt	12900

[0010]	gctatgtaaa tttagcttct aaagccctgt tctacagaga ttttgagct tccactgcac	12960
	ccagaaaatg cacagctaaa gagaaaactt cccttggtga tggttattag attttacaag	13020
	aagaggccaa aggagacaca tacttatgcc agaagaactt tccagagata gcattgcata	13080
	gcgaaatagc ctgaattatt tttatttttt aaaacatttt ttcttttctt ttttcttttc	13140
	tttttctttt tttttttttt tttttgagac agagtctcac tctgtcaccc aggctggagt	13200
	gcagtggcgt gatcttggct cactgcaatc tccacctccc gggttcaagc cattctcctc	13260
	cctcagcctc ccaagtagct gggattacag gcatgcgtca ctatgctctg gctaattttt	13320
	tttttctttt ttttttttgt atttttagta gagatgggggt ttcacatgt tggccaggct	13380
	ggtcttgaac tcctgacctc aagtgatcca cgccttggc ctcccaaagt gctgggattt	13440
	caggcgtgag ccaccgcacc cggccaaaaa tttcttttct ttaagatgag gcctcactct	13500
	gttgcccagg ctggagtga gtgttacaat catagctcac tgtaactttg aactcctggg	13560
	ctcaagtgat cctcctgctt cagcctctca agtagctggg attacaggca tgtgccacca	13620
	caccagcta atttttttta aaataatttt ttttagagac gaggtctctg attggtgcc	13680
	taggttggtc ccagactcct gacgggctgc attttaatcc tagctccacc acttacggga	13740
	gtcaaaattc aaaagataga aaagggcata taggctgggt gcagtggctc acacctgcaa	13800
	tcccagcaat ttgggaggct gaggtgggcg gggtgcttga ggtcaggagt tcgagatcag	13860
	cctgggcaac atggcaaaac ttgtatctac taaaaataca aaaattagcc agatgtggtg	13920
	gtgtacacct gtaatcccag ctactccgaa ggctgaggca agagaatccc ttgaactcag	13980
	gaggcagagg ttacaatgag cagagatcga aactcagact ccataaaaaac aaacaaacaa	14040
	aaaaagaaag caggctgggt gtggtggctc acgcctgtaa cccagcact tcgggaggcc	14100
	aaggcgagcg gatcacctga gggtgggcat tcgagaccag cctgaccaac aaggagaaac	14160
	cctgtctcta ctgaaaatac aaaattagcc gggcttggtt gcgcatgcct gtaatctcag	14220
	ctactcgga ggcagaggca agataattgc ttgaaccgg gaggcggagg ttgcggtgag	14280
	ccaagatcat gccattgcac tccaacctgg gcaacaatag cgaaactcca tctcaaaaaa	14340
	aaaaaagcaa agggcatata gtgaaaagct ttcttcttac acatgagtat tcacttctc	14400
	ttcctagagg caaccaaggt tatttttggt tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg	14460
	tgttttggga cagtctcact ctctcaccaa ggctggaatg cagtgggtgc atctcactgc	14520
	aaactctgcc tcccagtctc aagcgatctt gtgcctcagc ctcccagttt tttttcttt	14580

[0011]	taaatggggt ctcattctgt cgcccagggt ggagtgcagt ggcatgatca tagctcactg	14640
	cagcctcgac ctcttggtc aggttatcct cccacctcag cctccggcat agctggggct	14700
	actggcatgc accaccacac tcagttaatt ttttttcttt tttagacag agtctcactc	14760
	tgtcacctag actggagtgc agtggtgcca tctcatttgt ttcactgcaa ctttgactt	14820
	ctgggctcaa gtgattctcc cacctcagcc tcccaaggcg gctaattaaa aaaaattttt	14880
	tttttttttt ttttagagat ggggtttcgc catgttgccc aggetgatct cgaactcctg	14940
	ggcacaaca atctttccac ctcgatcttt caaagagctg ggatgagaga tttccaccat	15000
	gcctggcctc attttctttt ttaatttttt ttttagacatt atagctcttt ttaatggcct	15060
	cattttctta tgtttaattc gagaattatt cttttcatat acaaagaata tattttctcc	15120
	acctttaaaa acaaatagta gactgttta catctcgctt tattcagtta gtgatgttcc	15180
	ttagatacgg gtccaaatta gtacacaaag cacttcctca ttctctctt acggctgcat	15240
	agcagtcac tgaatgggtg agctatgac tatttaacct attctttatt gatggacatt	15300
	tggttttgta tatacatttg taattctgta tagattacaa atcaccatcc aaagaaattg	15360
	tactggttta ttctctaca atgtgtgaga gttgggtaat tacttaatct caatatgtga	15420
	gagtttaggc agttacctaa tctctctgag tctcagttc tctatctgca aaataaaca	15480
	aacagtgttg acagtatcta tttctcgga ttattgtgga gattactgag atgatgcctg	15540
	taaagtatth ggcatgtagg agttggtgct ctccaaataa ggatatgatt ttatttgtat	15600
	ttgtgagcta ctgtcccagc caggtaaag gatatgatga gacctcctg ccagaccggg	15660
	tttctctgat tagaacgagg agcagatgtt gcaggaaatt agcaactgat atcagaagag	15720
	ccgtgggcat tctcttgcca gaggtgccct gtctccaggg cgctcagtc ccccccata	15780
	tgtcttctgc tcccaggtec atccaagctg aatgatecgc ctgactcaag aagaagcctt	15840
	tgggaccaag gaaactttcc cctgatcatc aagaatctta agatagaaga ctcagatact	15900
	tacatctgtg aagtggagga ccagaaggag gaggtgcaat tgctagtgtt cggatgtgag	15960
	tggggcaggt ggggatgagg atacctcctg cctggttccc ttccccacta ctcccacccc	16020
	tgcaccaa at ccagcctgag ctggtgatac cgcagcagcc ccaagaggac caggctgtca	16080
	aactggcctc caaatgtctt aaaacccttc ttgatcaggt gagggatgct ggtgggcgga	16140
	ggagggaaga ggccttggga aaaggaaaga aaagggaagg aggcaaggga aggagggaga	16200
	gagactgggg aagagaggat gaggggagag gaggaaagaa gagagagagg aggggagagg	16260

[0012]	gaaaccctat cttggctggg ggtgcgcagc tgggtgctgg gaggaaggag atgttgggac	16320
	ggcgataatg gagagatgtt gttggtttcc tgttgtctgc ctttctcctt ggggatggta	16380
	tgtgtgtgac acagctggcc tttccctcca cagtgaactgc caactctgac acccacctgc	16440
	ttcaggggca gaggctgacc ctgaccttgg agagcccccc tggtagtagc ccctcagtgc	16500
	aatgtaggag tccaaggggt aaaaacatac agggggggaa gaccctctcc gtgtctcagc	16560
	tggagctcca ggatagtggc acctggacat gcactgtctt gcagaaccag aagaagggtg	16620
	agttcaaaat agacatcgtg gtgctaggta agggaagccc ctcttcgcgc agtctcctcc	16680
	ctgccccagg ggctgacagc ccctccctct gctctgactg ccctgtttct ggttctgggtg	16740
	ctgggaggtc aggagtggag aagactaggt cccctagagc tgaggcctgt cttgaaggac	16800
	tcactggggc cctcatcctc agggggctga ttggcagcca cccctcagtg tggtagcat	16860
	ggagaaagga aaggctgggg aaggtaaagga tgctagaggc ccgagtctcc tttggaggcc	16920
	ccaaaggagg aatgtcaggg agcttacttt ctttgttgcc tcagctccac acccctacca	16980
	agttggcaaa tccacttact cagggacact aacaccagta agccaaccct gatgatgttc	17040
	tatgttgtac ctctggacct ctaagccagg cactgtggg gagaccaagg tcctaccca	17100
	gacctgtcc cctgggtgct tatgtgactt aaggtagaca taaggtagtg tgccagttta	17160
	gtgcatgtac gctgattgaa atcctggttc tgccacaacc atgtgacctt gggtagtga	17220
	ctaaacctct ctgcaccttg gtttcagcct ctgtgaaatg gggatgatgt taactgccat	17280
	agtgactacc tcgtattaag ttgaggactg atatacgtaa ggcactgaaa atggtgcctg	17340
	gcacagagta agccctagtt aagtgttcgc tgttattttg tgaagggtga tgaatacgcc	17400
	tctaaggagt ggaggccaaa tggtctctgt ggtccaggaa tcctaaggac agcaaggatc	17460
	ccctgtggct gggctgctct gtgatggctt ccgggaggag ggaggtggcc tgctgtagga	17520
	aatgctggg tggaagaagg gagagaaggc tggagaggta ggaaggaact gaagtatctg	17580
	aagtgacaag gtgggtgtct ggactcgtcg ggtcccttc catctccctg ctgcctccac	17640
	atgccaacct cactcgtgca cctcatctt cctatctcct caccagggt ctctcccttc	17700
	ccacctccag ctttccagaa ggcctccagc atagtctata agaaaggagg ggaacagggtg	17760
	gagttctcct tcccactcgc ctttacagtt gaaaagctga cgggcagtgg cgagctgtgg	17820
	tggcaggcgg agagggttc ctctccaag tcttgatca cctttgacct gaagaacaag	17880
	gaagtgtctg taaaacgggt taccaggac cctaagctcc agatgggcaa gaagctcccg	17940

```

ctccacctca ccctgccccca ggccttgcct cagtatgctg gctctggaaa cctcaccctg 18000
gcccttgaag cgaaaacagg aaagttgcat caggaagtga acctgggtgt gatgagaggt 18060
gagggggccag gccagggagg ggtggggcagg ggaaggagtt ggaggggcct ggcccagggc 18120
tccctctgag gcaagccagg cccaagagg ggatgcctag gccctggtca cctggatgaa 18180
gtgagggagg gccctctggg tttggggctg gttttgaact gagacatcca tgagccagcc 18240
tggggctggc ttactgaag atc 18263

```

<210> 4  
 <211> 458  
 <212> PRT  
 <213> 人工的

<220>  
 <223> 合成的

<400> 4

```

Met Cys Arg Ala Ile Ser Leu Arg Arg Leu Leu Leu Leu Leu Gln
1           5           10           15

```

[0013] Leu Ser Gln Leu Leu Ala Val Thr Gln Gly Lys Lys Val Val Leu Gly  
                   20                  25                  30

```

Lys Lys Gly Asp Thr Val Glu Leu Thr Cys Thr Ala Ser Gln Lys Lys
35           40           45

```

```

Ser Ile Gln Phe His Trp Lys Asn Ser Asn Gln Ile Lys Ile Leu Gly
50           55           60

```

```

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

```

```

Ala Asp Ser Arg Arg Ser Leu Trp Asp Gln Gly Asn Phe Pro Leu Ile
85           90           95

```

```

Ile Lys Asn Leu Lys Ile Glu Asp Ser Asp Thr Tyr Ile Cys Glu Val
100          105          110

```

```

Glu Asp Gln Lys Glu Glu Val Gln Leu Leu Val Phe Gly Leu Thr Ala
115          120          125

```

```

Asn Ser Asp Thr His Leu Leu Gln Gly Gln Ser Leu Thr Leu Thr Leu

```

	130		135				140										
	Glu	Ser	Pro	Pro	Gly	Ser	Ser	Pro	Ser	Val	Gln	Cys	Arg	Ser	Pro	Arg	
	145					150					155					160	
	Gly	Lys	Asn	Ile	Gln	Gly	Gly	Lys	Thr	Leu	Ser	Val	Ser	Gln	Leu	Glu	
					165					170					175		
	Leu	Gln	Asp	Ser	Gly	Thr	Trp	Thr	Cys	Thr	Val	Leu	Gln	Asn	Gln	Lys	
				180					185					190			
	Lys	Val	Glu	Phe	Lys	Ile	Asp	Ile	Val	Val	Leu	Ala	Phe	Gln	Lys	Ala	
			195					200					205				
	Ser	Ser	Ile	Val	Tyr	Lys	Lys	Glu	Gly	Glu	Gln	Val	Glu	Phe	Ser	Phe	
		210					215					220					
	Pro	Leu	Ala	Phe	Thr	Val	Glu	Lys	Leu	Thr	Gly	Ser	Gly	Glu	Leu	Trp	
	225					230					235					240	
[0014]	Trp	Gln	Ala	Glu	Arg	Ala	Ser	Ser	Ser	Lys	Ser	Trp	Ile	Thr	Phe	Asp	
					245					250					255		
	Leu	Lys	Asn	Lys	Glu	Val	Ser	Val	Lys	Arg	Val	Thr	Gln	Asp	Pro	Lys	
				260					265					270			
	Leu	Gln	Met	Gly	Lys	Lys	Leu	Pro	Leu	His	Leu	Thr	Leu	Pro	Gln	Ala	
			275					280					285				
	Leu	Pro	Gln	Tyr	Ala	Gly	Ser	Gly	Asn	Leu	Thr	Leu	Ala	Leu	Glu	Ala	
		290					295					300					
	Lys	Thr	Gly	Lys	Leu	His	Gln	Glu	Val	Asn	Leu	Val	Val	Met	Arg	Val	
	305					310					315					320	
	Ala	Gln	Leu	Asn	Asn	Thr	Leu	Thr	Cys	Glu	Val	Met	Gly	Pro	Thr	Ser	
					325					330					335		
	Pro	Lys	Met	Arg	Leu	Thr	Leu	Lys	Gln	Glu	Asn	Gln	Glu	Ala	Arg	Val	
				340					345					350			
	Ser	Glu	Glu	Gln	Lys	Val	Val	Gln	Val	Val	Ala	Pro	Glu	Thr	Gly	Leu	

355	360	365
Trp Gln Cys Leu Leu Ser	Glu Gly Asp Lys Val	Lys Met Asp Ser Arg
370	375	380
Ile Gln Val Leu Ser Arg	Gly Val Asn Gln Thr	Val Phe Leu Ala Cys
385	390	395
Val Leu Gly Gly Ser Phe	Gly Phe Leu Gly Phe	Leu Gly Leu Cys Ile
405	410	415
Leu Cys Cys Val Arg Cys	Arg His Gln Gln Arg	Gln Ala Ala Arg Met
420	425	430
Ser Gln Ile Lys Arg Leu	Leu Ser Glu Lys Lys Thr	Cys Gln Cys Pro
435	440	445
His Arg Met Gln Lys Ser	His Asn Leu Ile	
450	455	

[0015]

<210> 5  
 <211> 21  
 <212> DNA  
 <213> 人工的

<220>  
 <223> 合成的

<400> 5  
 gaagtgggtg tgccattcag a

21

<210> 6  
 <211> 24  
 <212> DNA  
 <213> 人工的

<220>  
 <223> 合成的

<400> 6  
 aaagctcaga agcagacaga gtca

24

<210> 7  
 <211> 25  
 <212> DNA  
 <213> 人工的

	<220>		
	<223>	合成的	
	<400>	7	
		ttccaaaagc ctacagcagg cccag	25
	<210>	8	
	<211>	21	
	<212>	DNA	
	<213>	人工的	
	<220>		
	<223>	合成的	
	<400>	8	
		tcattctcccc ttccctgaacc t	21
	<210>	9	
	<211>	21	
	<212>	DNA	
	<213>	人工的	
	<220>		
	<223>	合成的	
[0016]	<400>	9	
		cccagccaca agaagaagaa a	21
	<210>	10	
	<211>	26	
	<212>	DNA	
	<213>	人工的	
	<220>		
	<223>	合成的	
	<400>	10	
		cttcccccgcc atccattttt ctgttc	26
	<210>	11	
	<211>	22	
	<212>	DNA	
	<213>	人工的	
	<220>		
	<223>	合成的	
	<400>	11	
		ggtctcgaac tcatgagctc aa	22
	<210>	12	

	<p>&lt;211&gt; 25 &lt;212&gt; DNA &lt;213&gt; 人工的</p> <p>&lt;220&gt; &lt;223&gt; 合成的</p> <p>&lt;400&gt; 12 ggcatagtga cacacacctg taatt</p>	25
	<p>&lt;210&gt; 13 &lt;211&gt; 27 &lt;212&gt; DNA &lt;213&gt; 人工的</p> <p>&lt;220&gt; &lt;223&gt; 合成的</p> <p>&lt;400&gt; 13 tgatccactc accttggcct ctcagag</p>	27
[0017]	<p>&lt;210&gt; 14 &lt;211&gt; 25 &lt;212&gt; DNA &lt;213&gt; 人工的</p> <p>&lt;220&gt; &lt;223&gt; 合成的</p> <p>&lt;400&gt; 14 gtcagggagc ttacttttctt tgttg</p>	25
	<p>&lt;210&gt; 15 &lt;211&gt; 26 &lt;212&gt; DNA &lt;213&gt; 人工的</p> <p>&lt;220&gt; &lt;223&gt; 合成的</p> <p>&lt;400&gt; 15 tgttagtgtc cctgagtaag tggatt</p>	26
	<p>&lt;210&gt; 16 &lt;211&gt; 27 &lt;212&gt; DNA &lt;213&gt; 人工的</p> <p>&lt;220&gt; &lt;223&gt; 合成的</p> <p>&lt;400&gt; 16</p>	

	ctcagctcca caccctacc aagttgg	27
	<210> 17	
	<211> 142	
	<212> DNA	
	<213> 人工的	
	<220>	
	<223> 合成的	
	<400> 17	
	tgtttgctg tgacatgaac tcattgtgac acaaaccact gtgctagggg ggatccacta	60
	gtaacggccg ccagtgtgct ggaattcgcc ctcgcaaggg ccaggcatat aagtacacaa	120
	taaacaaatg gcagctctct cc	142
	<210> 18	
	<211> 99	
	<212> DNA	
	<213> 人工的	
	<220>	
	<223> 合成的	
[0018]	<400> 18	
	cccctccttc cttccccagg cactttccaa gtgtcaactc tagagcctat cgcggccgca	60
	ccggtataac ttcgtataat gtatgctata cgaagttat	99
	<210> 19	
	<211> 90	
	<212> DNA	
	<213> 人工的	
	<220>	
	<223> 合成的	
	<400> 19	
	ataacttcgt ataatgtatg ctatacgaag ttatgtcgac gtagcctatt tctctagatc	60
	caaaatgatg acaacaaaag gtaccttggtg	90
	<210> 20	
	<211> 20	
	<212> DNA	
	<213> 人工的	
	<220>	
	<223> 合成的	
	<400> 20	

	gcagctctgc cctcattcag	20
	<210> 21	
	<211> 21	
	<212> DNA	
	<213> 人工的	
	<220>	
	<223> 合成的	
	<400> 21	
	catctttgcc gtatggttgg t	21
	<210> 22	
	<211> 22	
	<212> DNA	
	<213> 人工的	
	<220>	
	<223> 合成的	
	<400> 22	
	ccccctcgtc cctgctggtt ca	22
[0019]	<210> 23	
	<211> 26	
	<212> DNA	
	<213> 人工的	
	<220>	
	<223> 合成的	
	<400> 23	
	caagaagact accctgaaga tgaaga	26
	<210> 24	
	<211> 23	
	<212> DNA	
	<213> 人工的	
	<220>	
	<223> 合成的	
	<400> 24	
	tgtgagtgc acaatggaaa act	23
	<210> 25	
	<211> 22	
	<212> DNA	
	<213> 人工的	

	<220>		
	<223>	合成的	
	<400>	25	
		cgttccccca cccagagacc ca	22
	<210>	26	
	<211>	19	
	<212>	DNA	
	<213>	人工的	
	<220>		
	<223>	合成的	
	<400>	26	
		ggcaccgagc agtgacagt	19
	<210>	27	
	<211>	22	
	<212>	DNA	
	<213>	人工的	
	<220>		
	<223>	合成的	
[0020]	<400>	27	
		ttcaccgtgg atagtcctt tt	22
	<210>	28	
	<211>	24	
	<212>	DNA	
	<213>	人工的	
	<220>		
	<223>	合成的	
	<400>	28	
		agttcctggc cctctgggat tccg	24
	<210>	29	
	<211>	26	
	<212>	DNA	
	<213>	人工的	
	<220>		
	<223>	合成的	
	<400>	29	
		ttgctttctt tctgtagttg atttcc	26
	<210>	30	

	<p>&lt;211&gt; 21 &lt;212&gt; DNA &lt;213&gt; 人工的</p> <p>&lt;220&gt; &lt;223&gt; 合成的</p> <p>&lt;400&gt; 30 ccggcacact ctcttcttga g</p>	21
	<p>&lt;210&gt; 31 &lt;211&gt; 20 &lt;212&gt; DNA &lt;213&gt; 人工的</p> <p>&lt;220&gt; &lt;223&gt; 合成的</p> <p>&lt;400&gt; 31 tcccaccact gcccagccca</p>	20
[0021]	<p>&lt;210&gt; 32 &lt;211&gt; 17 &lt;212&gt; DNA &lt;213&gt; 人工的</p> <p>&lt;220&gt; &lt;223&gt; 合成的</p> <p>&lt;400&gt; 32 tgcggccgat cttagcc</p>	17
	<p>&lt;210&gt; 33 &lt;211&gt; 18 &lt;212&gt; DNA &lt;213&gt; 人工的</p> <p>&lt;220&gt; &lt;223&gt; 合成的</p> <p>&lt;400&gt; 33 ttgaccgatt ccttgcg</p>	18
	<p>&lt;210&gt; 34 &lt;211&gt; 21 &lt;212&gt; DNA &lt;213&gt; 人工的</p> <p>&lt;220&gt; &lt;223&gt; 合成的</p> <p>&lt;400&gt; 34</p>	

	acgagcgggt tcggcccatt c	21
	<210> 35	
	<211> 100	
	<212> DNA	
	<213> 人工的	
	<220>	
	<223> 合成的	
	<400> 35	
	tgaacctgct gctgctgggt gagtcgatta tcctggggag tggagaagct aggccgagcc	60
	agttccgggt gtcgccgctg gatcggacct ggaacctggg	100
	<210> 36	
	<211> 90	
	<212> DNA	
	<213> 人工的	
	<220>	
	<223> 合成的	
	<400> 36	
[0022]	atgccaggga cagccctgat actgtaggta gagtcaaggg ctgtccaagt accggtataa	60
	cttcgtataa ggtatcctat acgaagttat	90
	<210> 37	
	<211> 89	
	<212> DNA	
	<213> 人工的	
	<220>	
	<223> 合成的	
	<400> 37	
	ataacttcgt ataaggtatc ctatacgaag ttatctcgac ctgatcttgg aggagacct	60
	ggaccgggag acgtgctggg ggcagggtt	89
	<210> 38	
	<211> 22	
	<212> DNA	
	<213> 人工的	
	<220>	
	<223> 合成的	
	<400> 38	
	gatgctcttg gctcttccag aa	22

	<210> 39	
	<211> 25	
	<212> DNA	
	<213> 人工的	
	<220>	
	<223> 合成的	
	<400> 39	
	atgaagccat atagacaacg aaggt	25
	<210> 40	
	<211> 22	
	<212> DNA	
	<213> 人工的	
	<220>	
	<223> 合成的	
	<400> 40	
	ccagctcaa actccccag cc	22
[0023]	<210> 41	
	<211> 20	
	<212> DNA	
	<213> 人工的	
	<220>	
	<223> 合成的	
	<400> 41	
	tcagccccag agaccagaag	20
	<210> 42	
	<211> 22	
	<212> DNA	
	<213> 人工的	
	<220>	
	<223> 合成的	
	<400> 42	
	tcaatcgctt gagagcacct aa	22
	<210> 43	
	<211> 19	
	<212> DNA	
	<213> 人工的	
	<220>	
	<223> 合成的	

	<400> 43 ttgtcggccc cgtggctca	19
	<210> 44 <211> 17 <212> DNA <213> 人工的	
	<220> <223> 合成的	
	<400> 44 gcggttctcg ggcaaga	17
	<210> 45 <211> 20 <212> DNA <213> 人工的	
	<220> <223> 合成的	
	<400> 45 tcagggccga gcagaaatag	20
[0024]	<210> 46 <211> 25 <212> DNA <213> 人工的	
	<220> <223> 合成的	
	<400> 46 acaccttcgt cctcaccctg agcga	25
	<210> 47 <211> 22 <212> DNA <213> 人工的	
	<220> <223> 合成的	
	<400> 47 ggttcacctc aacctgtttt cc	22
	<210> 48 <211> 18 <212> DNA	

	<213> 人工的	
	<220>	
	<223> 合成的	
	<400> 48	
	cgcttccagg tgcgctaa	18
	<210> 49	
	<211> 22	
	<212> DNA	
	<213> 人工的	
	<220>	
	<223> 合成的	
	<400> 49	
	acctggggccc tgctttcaag cc	22
	<210> 50	
	<211> 19	
	<212> DNA	
	<213> 人工的	
	<220>	
[0025]	<223> 合成的	
	<400> 50	
	ggtggagagg ctattcggc	19
	<210> 51	
	<211> 17	
	<212> DNA	
	<213> 人工的	
	<220>	
	<223> 合成的	
	<400> 51	
	gaacacggcg gcatcag	17
	<210> 52	
	<211> 23	
	<212> DNA	
	<213> 人工的	
	<220>	
	<223> 合成的	
	<400> 52	
	tgggcacaac agacaatcgg ctg	23

<210> 53  
 <211> 210  
 <212> PRT  
 <213> 人工的

<220>  
 <223> 合成的

<400> 53

Met Gln Pro Trp Leu Trp Leu Val Phe Ser Met Lys Leu Ala Val Leu  
 1 5 10 15

His Gly Asn Ser Val Leu Gln Gln Thr Pro Ala Tyr Ile Lys Val Gln  
 20 25 30

Thr Asn Lys Met Val Met Leu Ser Cys Glu Ala Lys Ile Ser Leu Ser  
 35 40 45

Asn Met Arg Ile Tyr Trp Leu Arg Gln Arg Gln Ala Pro Ser Ser Asp  
 50 55 60

[0026] Ser His His Glu Phe Leu Ala Leu Trp Asp Ser Ala Lys Gly Thr Ile  
 65 70 75 80

His Gly Glu Glu Val Glu Gln Glu Lys Ile Ala Val Phe Arg Asp Ala  
 85 90 95

Ser Arg Phe Ile Leu Asn Leu Thr Ser Val Lys Pro Glu Asp Ser Gly  
 100 105 110

Ile Tyr Phe Cys Met Ile Val Gly Ser Pro Glu Leu Thr Phe Gly Lys  
 115 120 125

Gly Thr Gln Leu Ser Val Val Asp Phe Leu Pro Thr Thr Ala Gln Pro  
 130 135 140

Thr Lys Lys Ser Thr Leu Lys Lys Arg Val Cys Arg Leu Pro Arg Pro  
 145 150 155 160

Glu Thr Gln Lys Gly Leu Thr Cys Ser Leu Thr Thr Leu Ser Leu Leu  
 165 170 175

Val Val Cys Ile Leu Leu Leu Leu Ala Phe Leu Gly Val Ala Val Tyr

	180	185	190
	Phe Tyr Cys Val Arg Arg Arg Ala Arg Ile His Phe Met Lys Gln Phe 195 200 205		
	His Lys 210		
	<210> 54 <211> 243 <212> PRT <213> 人工的		
	<220> <223> 合成的		
	<400> 54		
	Met Ala Ser Pro Leu Thr Arg Phe Leu Ser Leu Asn Leu Leu Leu Leu 1 5 10 15		
	Gly Glu Ser Ile Ile Leu Gly Ser Gly Glu Ala Arg Pro Ser Gln Phe 20 25 30		
[0027]	Arg Val Ser Pro Leu Asp Arg Thr Trp Asn Leu Gly Glu Thr Val Glu 35 40 45		
	Leu Lys Cys Gln Val Leu Leu Ser Asn Pro Thr Ser Gly Cys Ser Trp 50 55 60		
	Leu Phe Gln Pro Arg Gly Ala Ala Ala Ser Pro Thr Phe Leu Leu Tyr 65 70 75 80		
	Leu Ser Gln Asn Lys Pro Lys Ala Ala Glu Gly Leu Asp Thr Gln Arg 85 90 95		
	Phe Ser Gly Lys Arg Leu Gly Asp Thr Phe Val Leu Thr Leu Ser Asp 100 105 110		
	Phe Arg Arg Glu Asn Glu Gly Tyr Tyr Phe Cys Ser Ala Leu Ser Asn 115 120 125		
	Ser Ile Met Tyr Phe Ser His Phe Val Pro Val Phe Leu Pro Ala Lys 130 135 140		

Pro Thr Thr Thr Pro Ala Pro Arg Pro Pro Thr Pro Ala Pro Thr Ile  
145 150 155 160

Ala Ser Gln Pro Leu Ser Leu Arg Pro Glu Ala Cys Arg Pro Ala Ala  
165 170 175

Gly Gly Ala Val Lys Gly Thr Gly Leu Asp Phe Ala Cys Asp Ile Tyr  
180 185 190

Ile Trp Ala Pro Leu Ala Gly Ile Cys Val Ala Leu Leu Leu Ser Leu  
195 200 205

Ile Ile Thr Leu Ile Cys Tyr His Arg Ser Arg Lys Arg Val Cys Lys  
210 215 220

Cys Pro Arg Pro Leu Val Arg Gln Glu Gly Lys Pro Arg Pro Ser Glu  
225 230 235 240

Lys Ile Val

[0028]

<210> 55  
<211> 151  
<212> DNA  
<213> 人工的  
  
<220>  
<223> 合成的  
  
<400> 55  
tatggagtga aagccttttg tgtctgagat ctggtcttag ttaaactctg ggatcggcgc 60  
gccgaattcc tgcagcccgg gctcgagata acttcgtata atgtatgcta tacgaagtta 120  
tatgcatccg ggtaggggag gcgcttttcc c 151

<210> 56  
<211> 151  
<212> DNA  
<213> 人工的  
  
<220>  
<223> 合成的  
  
<400> 56  
agtattgttt tgccaagttc taattccatc agacctcgac ctgcagccct agataacttc 60

gtataatgta tgctatacga agttatccta ggccagaggg cttgggttga cagaaactca 120  
gtggcattct tatccagagt ttctctacac c 151

<210> 57  
<211> 293  
<212> PRT  
<213> 人

<400> 57

Lys Lys Val Val Leu Gly Lys Lys Gly Asp Thr Val Glu Leu Thr Cys  
1 5 10 15

Thr Ala Ser Gln Lys Lys Ser Ile Gln Phe His Trp Lys Asn Ser Asn  
20 25 30

Gln Ile Lys Ile Leu Gly Asn Gln Gly Ser Phe Leu Thr Lys Gly Pro  
35 40 45

Ser Lys Leu Asn Asp Arg Ala Asp Ser Arg Arg Ser Leu Trp Asp Gln  
50 55 60

[0029]

Gly Asn Phe Pro Leu Ile Ile Lys Asn Leu Lys Ile Glu Asp Ser Asp  
65 70 75 80

Thr Tyr Ile Cys Glu Val Glu Asp Gln Lys Glu Glu Val Gln Leu Leu  
85 90 95

Val Phe Gly Leu Thr Ala Asn Ser Asp Thr His Leu Leu Gln Gly Gln  
100 105 110

Ser Leu Thr Leu Thr Leu Glu Ser Pro Pro Gly Ser Ser Pro Ser Val  
115 120 125

Gln Cys Arg Ser Pro Arg Gly Lys Asn Ile Gln Gly Gly Lys Thr Leu  
130 135 140

Ser Val Ser Gln Leu Glu Leu Gln Asp Ser Gly Thr Trp Thr Cys Thr  
145 150 155 160

Val Leu Gln Asn Gln Lys Lys Val Glu Phe Lys Ile Asp Ile Val Val  
165 170 175

Leu Ala Phe Gln Lys Ala Ser Ser Ile Val Tyr Lys Lys Glu Gly Glu  
180 185 190

Gln Val Glu Phe Ser Phe Pro Leu Ala Phe Thr Val Glu Lys Leu Thr  
195 200 205

Gly Ser Gly Glu Leu Trp Trp Gln Ala Glu Arg Ala Ser Ser Ser Lys  
210 215 220

Ser Trp Ile Thr Phe Asp Leu Lys Asn Lys Glu Val Ser Val Lys Arg  
225 230 235 240

Val Thr Gln Asp Pro Lys Leu Gln Met Gly Lys Lys Leu Pro Leu His  
245 250 255

Leu Thr Leu Pro Gln Ala Leu Pro Gln Tyr Ala Gly Ser Gly Asn Leu  
260 265 270

Thr Leu Ala Leu Glu Ala Lys Thr Gly Lys Leu His Gln Glu Val Asn  
275 280 285

[0030]

Leu Val Val Met Arg  
290

<210> 58

<211> 151

<212> PRT

<213> 人

<400> 58

Val Leu His Gly Asn Ser Val Leu Gln Gln Thr Pro Ala Tyr Ile Lys  
1 5 10 15

Val Gln Thr Asn Lys Met Val Met Leu Ser Cys Glu Ala Lys Ile Ser  
20 25 30

Leu Ser Asn Met Arg Ile Tyr Trp Leu Arg Gln Arg Gln Ala Pro Ser  
35 40 45

Ser Asp Ser His His Glu Phe Leu Ala Leu Trp Asp Ser Ala Lys Gly  
50 55 60

Thr Ile His Gly Glu Glu Val Glu Gln Glu Lys Ile Ala Val Phe Arg

65	70	75	80
Asp Ala Ser Arg Phe Ile Leu Asn Leu Thr Ser Val Lys Pro Glu Asp	85	90	95
Ser Gly Ile Tyr Phe Cys Met Ile Val Gly Ser Pro Glu Leu Thr Phe	100	105	110
Gly Lys Gly Thr Gln Leu Ser Val Val Asp Phe Leu Pro Thr Thr Ala	115	120	125
Gln Pro Thr Lys Lys Ser Thr Leu Lys Lys Arg Val Cys Arg Leu Pro	130	135	140
Arg Pro Glu Thr Gln Lys Gly	145	150	
<210> 59			
<211> 152			
<212> PRT			
<213> 人			
[0031] <400> 59			
Arg Pro Ser Gln Phe Arg Val Ser Pro Leu Asp Arg Thr Trp Asn Leu	1	5	10
Gly Glu Thr Val Glu Leu Lys Cys Gln Val Leu Leu Ser Asn Pro Thr	20	25	30
Ser Gly Cys Ser Trp Leu Phe Gln Pro Arg Gly Ala Ala Ala Ser Pro	35	40	45
Thr Phe Leu Leu Tyr Leu Ser Gln Asn Lys Pro Lys Ala Ala Glu Gly	50	55	60
Leu Asp Thr Gln Arg Phe Ser Gly Lys Arg Leu Gly Asp Thr Phe Val	65	70	75
Leu Thr Leu Ser Asp Phe Arg Arg Glu Asn Glu Gly Tyr Tyr Phe Cys	85	90	95
Ser Ala Leu Ser Asn Ser Ile Met Tyr Phe Ser His Phe Val Pro Val	100	105	110

---

	Phe	Leu	Pro	Ala	Lys	Pro	Thr	Thr	Thr	Pro	Ala	Pro	Arg	Pro	Pro	Thr
			115					120					125			
[0032]	Pro	Ala	Pro	Thr	Ile	Ala	Ser	Gln	Pro	Leu	Ser	Leu	Arg	Pro	Glu	Ala
		130					135					140				
	Cys	Arg	Pro	Ala	Ala	Gly	Gly	Ala								
	145					150										

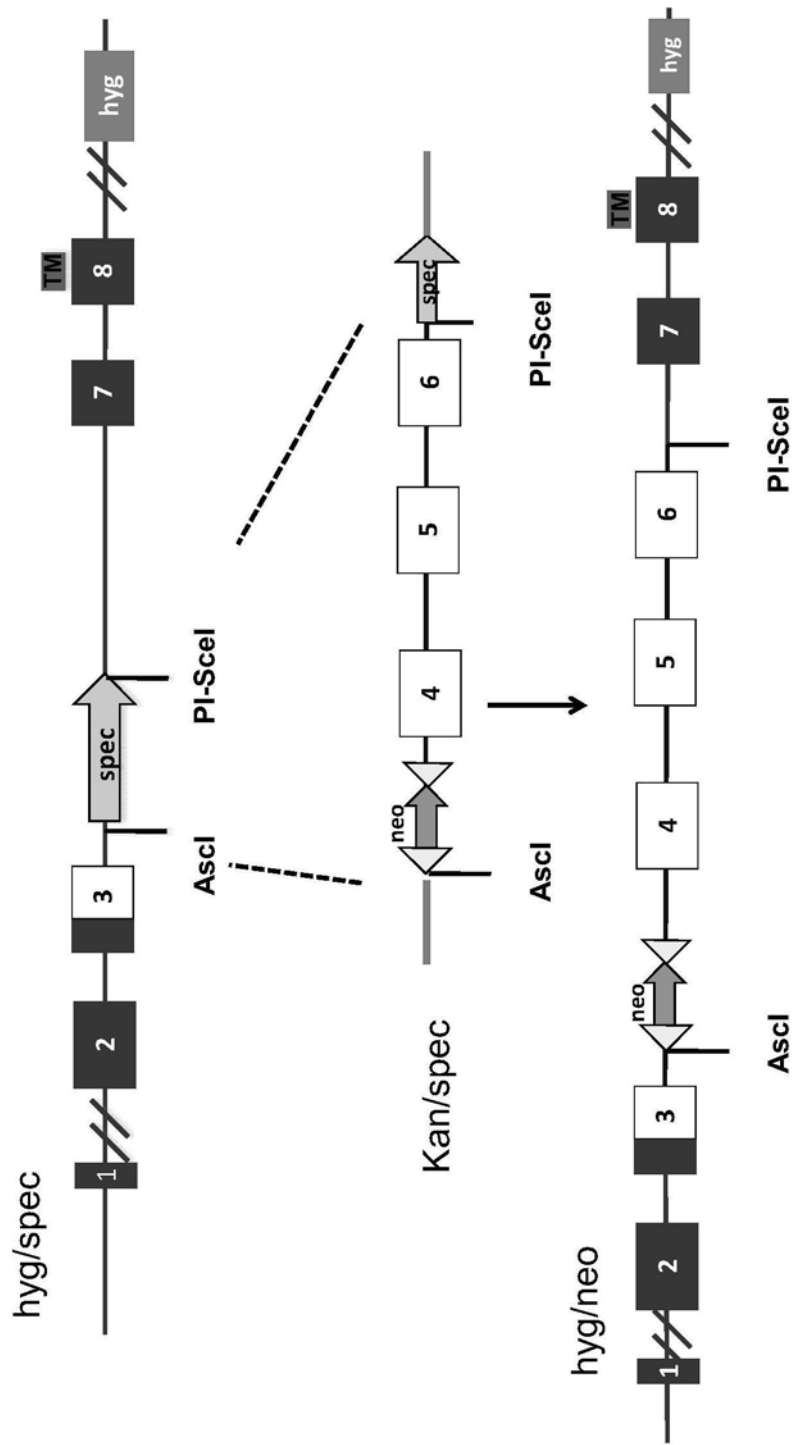


图1

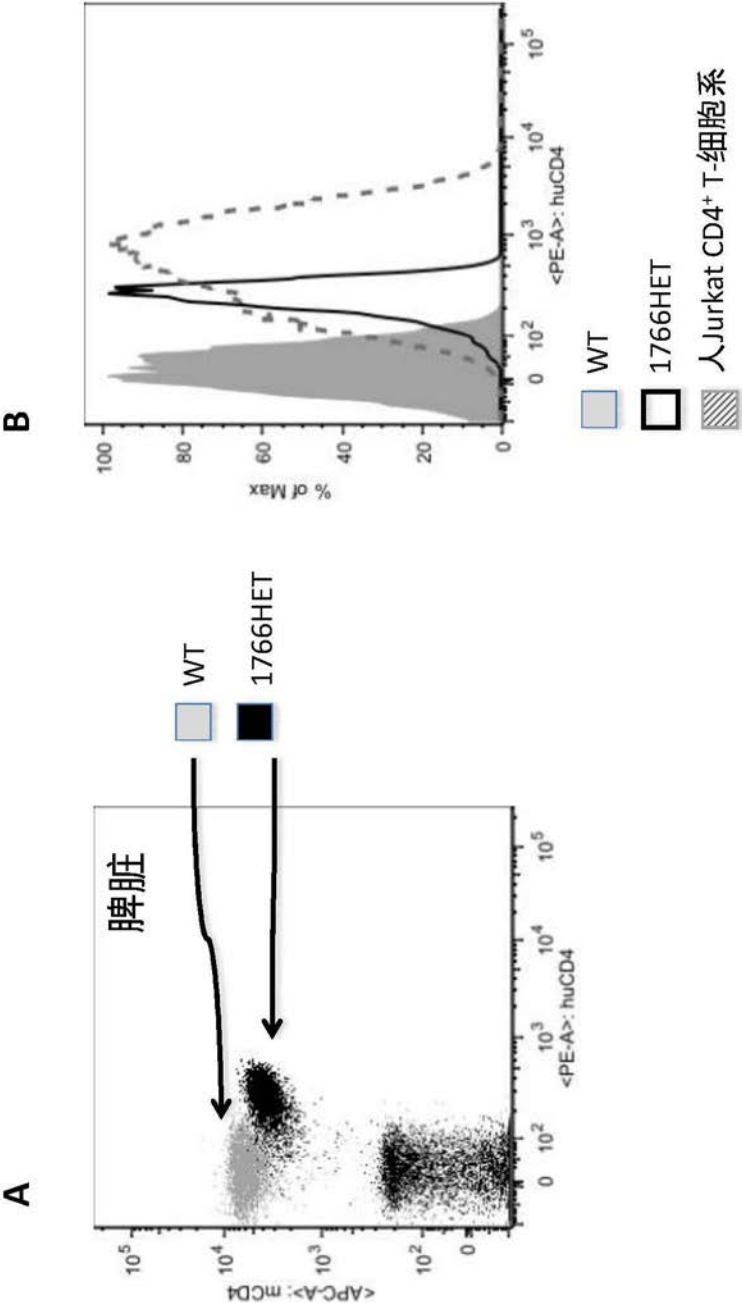


图2

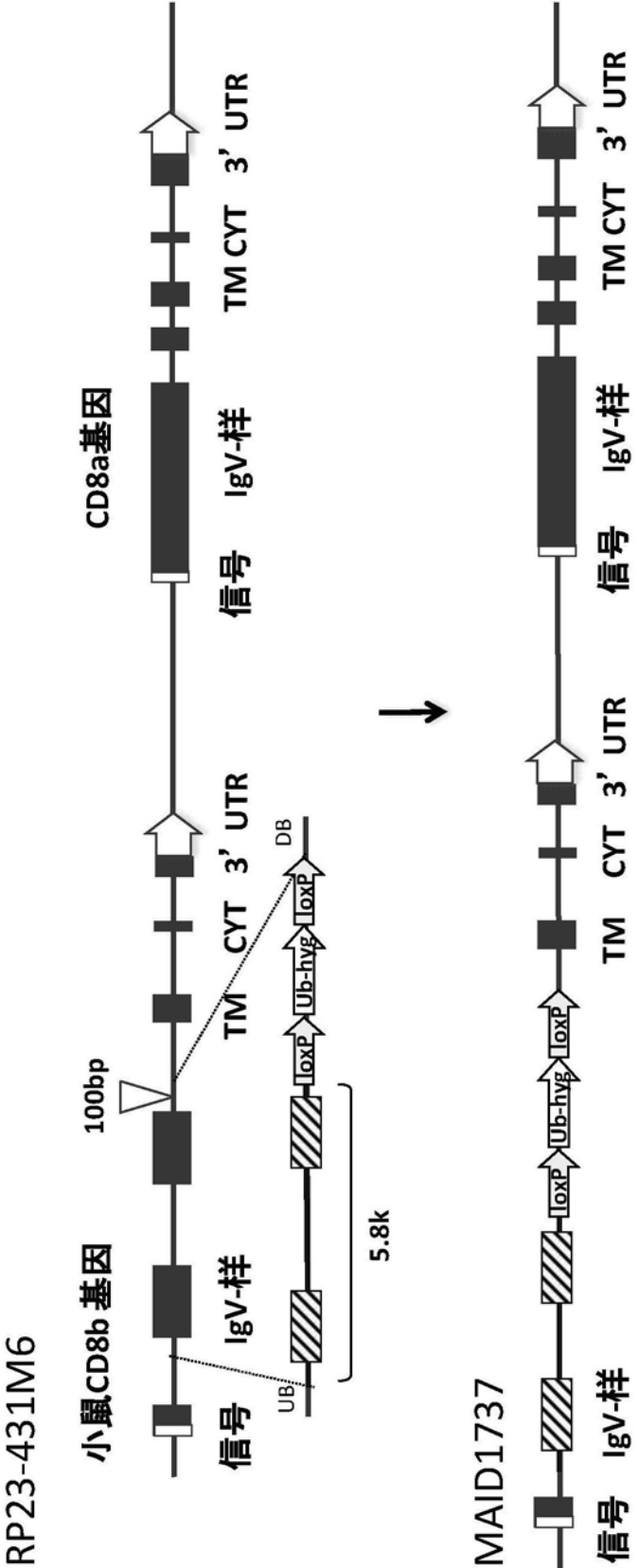


图3

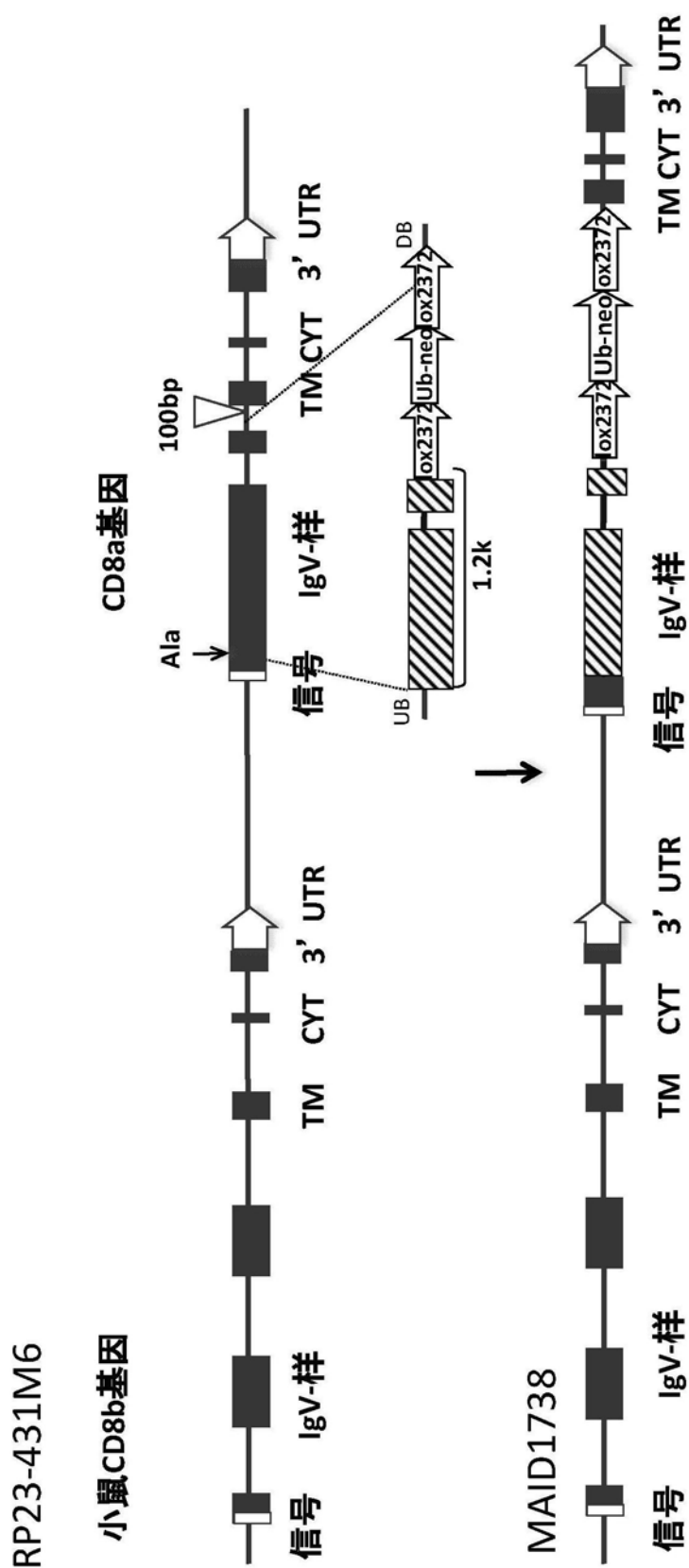


图4

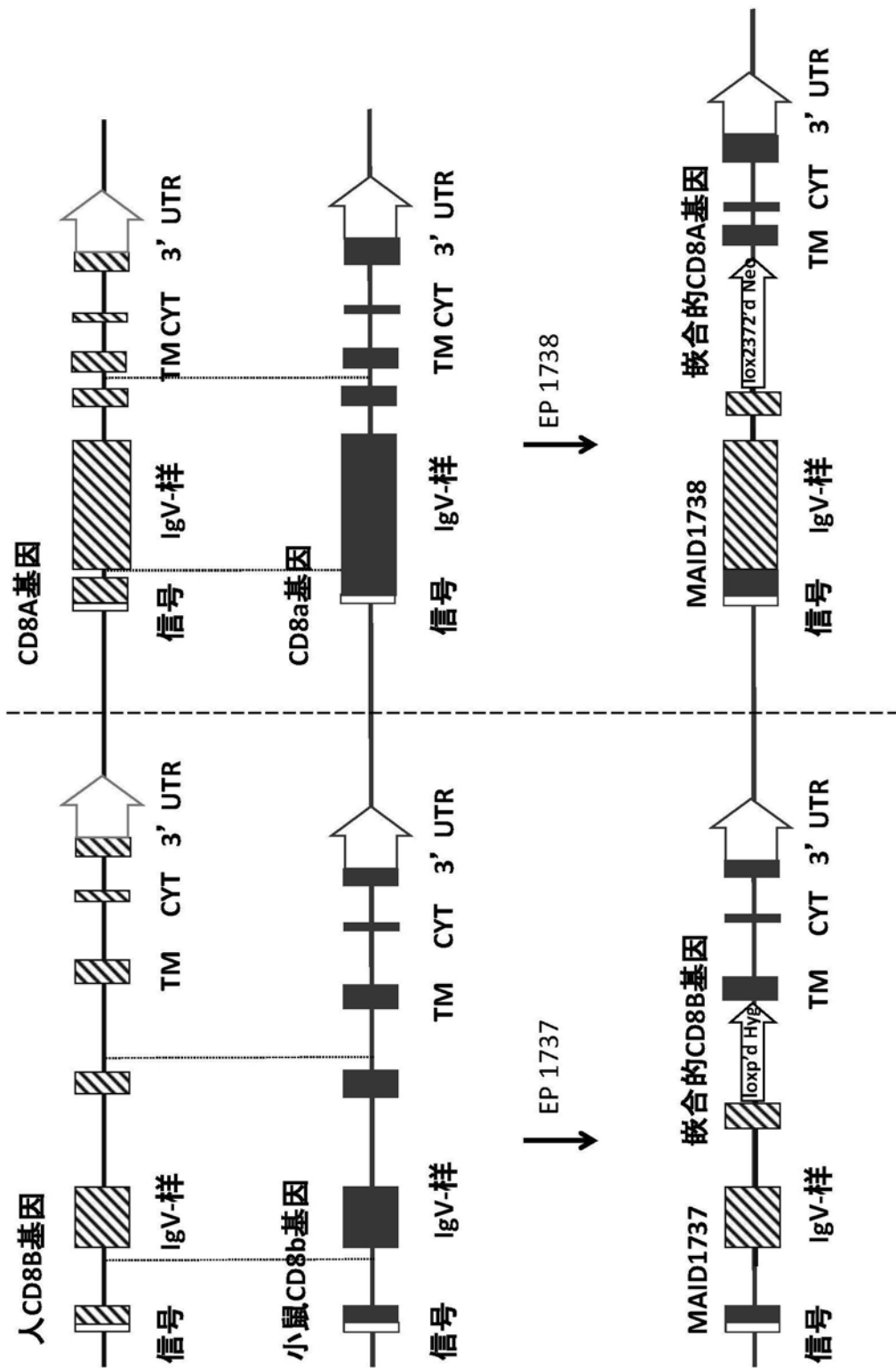


图5

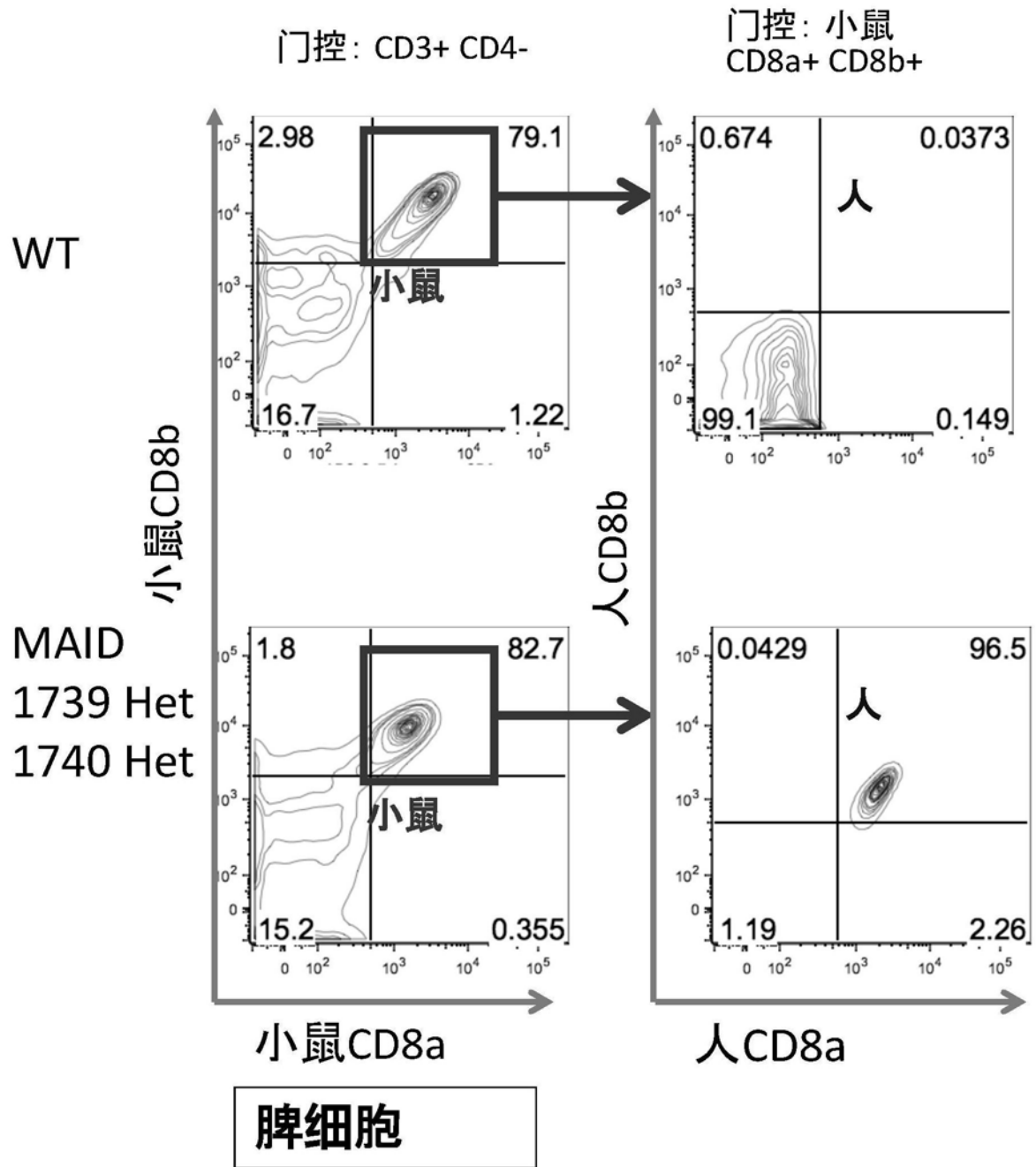


图6

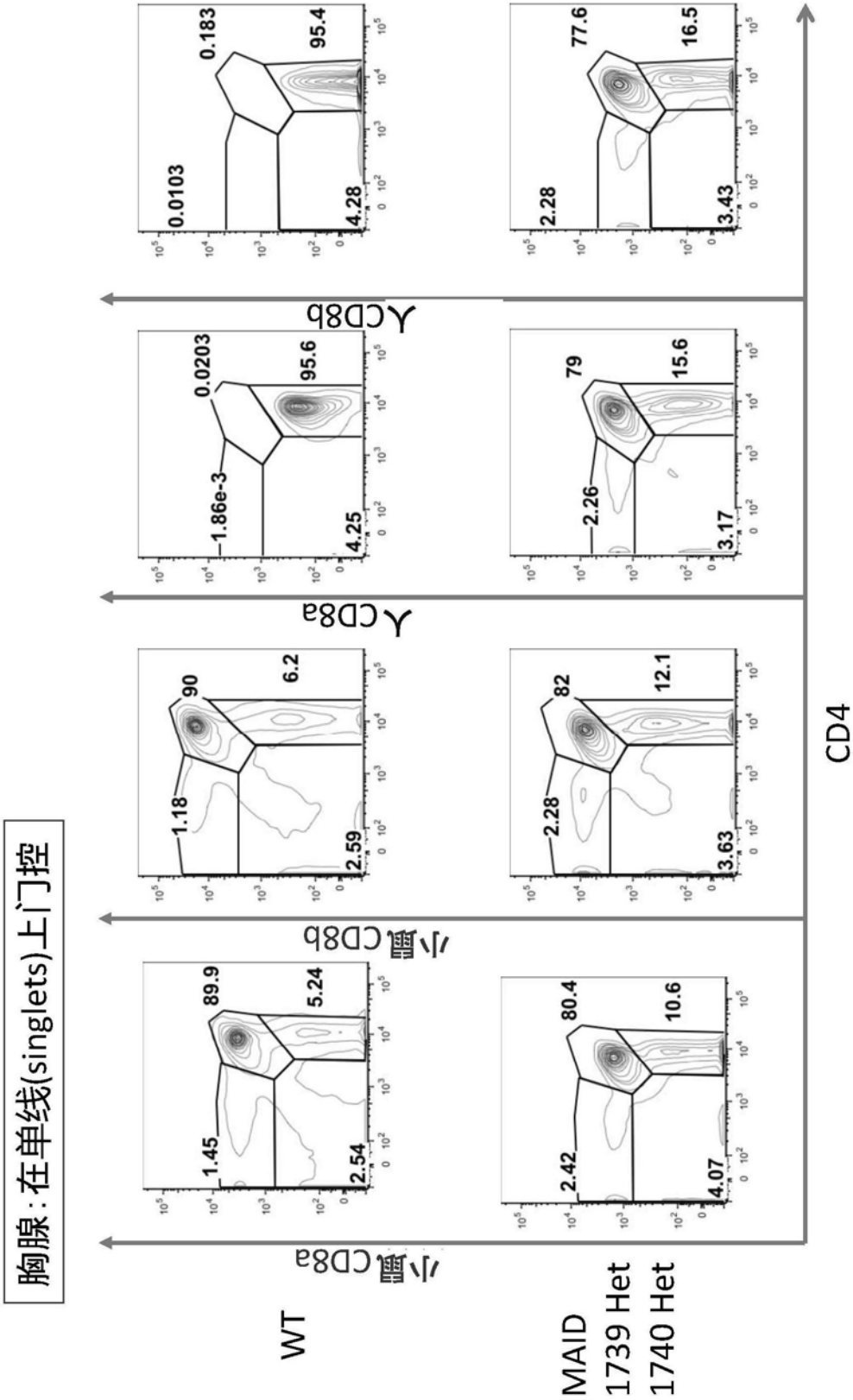


图7