



(12) 发明专利

(10) 授权公告号 CN 107666828 B

(45) 授权公告日 2021. 04. 09

(21) 申请号 201680030625.X

(22) 申请日 2016.04.06

(65) 同一申请的已公布的文献号  
申请公布号 CN 107666828 A

(43) 申请公布日 2018.02.06

(30) 优先权数据  
62/143,687 2015.04.06 US  
62/158,804 2015.05.08 US  
62/186,935 2015.06.30 US

(85) PCT国际申请进入国家阶段日  
2017.11.27

(86) PCT国际申请的申请数据  
PCT/US2016/026260 2016.04.06

(87) PCT国际申请的公布数据  
W02016/164492 EN 2016.10.13

(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  
C·古雷尔 C·基拉特索斯

(74) 专利代理机构 北京市君合律师事务所  
11517

代理人 张怡 孙倩

(51) Int.Cl.  
C07K 16/00 (2006.01)  
C07K 14/705 (2006.01)  
A01K 67/027 (2006.01)

审查员 李杨青

权利要求书9页 说明书78页  
序列表32页 附图34页

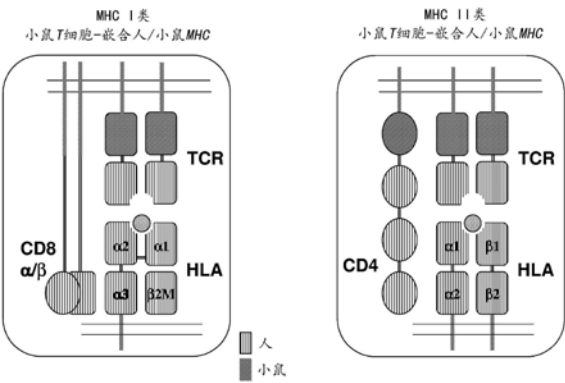
(54) 发明名称

非人动物中的人源化T细胞介导的免疫应答

(57) 摘要

本文公开了非人动物(例如,啮齿动物,例如小鼠或大鼠),所述非人动物经遗传工程改造成表达人源化T细胞辅助受体(例如,人源化CD4和/或CD8(例如,CD8 $\alpha$ 和/或CD8 $\beta$ ))、包含由至少一个人TCR可变区基因区段编码的可变结构域的人或人源化T细胞受体(TCR)、和/或结合所述人源化T细胞辅助受体的人或人源化主要组织相容性复合体(例如,分别为人或人源化MHC II(例如,MHC II $\alpha$ 和/或MHC II $\beta$ 链)和/或MHC I(例如,MHC I $\alpha$ )以及任选的人或人源化 $\beta$ 2微球蛋白)。本发明还提供了表达其的胚胎、组织和细胞。本发明还提供了用于制备经遗传工程改造的动物的方法,所述经遗传工程改造的动物表达至少一种人源化T细胞辅助受体(例如,人源化CD4和/或CD8)、与所述人源化T细胞辅助受体结合的至少一种人源化MHC(例如,分别为人源化MHC II和/或MHC I)和/或所述人源化TCR。本发明还提供了使用所述经遗传工程改造的动物来开发人类治

疗剂的方法,所述经遗传工程改造的动物产生基本上人源化的T细胞免疫应答。



1. 一种经遗传修饰的小鼠基因组, 包含:

(a) 编码嵌合人/小鼠CD4辅助受体的第一核苷酸序列, 其中所述嵌合人/小鼠CD4辅助受体包含人CD4多肽的细胞外部分以及小鼠CD4多肽的至少跨膜结构域和胞质结构域;

(b) 编码嵌合人/小鼠CD8 $\alpha$ 多肽的第二核苷酸序列和编码嵌合人/小鼠CD8 $\beta$ 多肽的第三核苷酸序列,

其中所述嵌合人/小鼠CD8 $\alpha$ 多肽包含人CD8 $\alpha$ 多肽的细胞外部分以及小鼠CD8 $\alpha$ 多肽的至少跨膜结构域和胞质结构域,

其中所述嵌合人/小鼠CD8 $\beta$ 多肽包含人CD8 $\beta$ 多肽的细胞外部分以及小鼠CD8 $\beta$ 多肽的至少跨膜结构域和胞质结构域;

(c) 编码嵌合MHCII $\alpha$ 多肽的第一核酸序列和编码嵌合MHCII $\beta$ 多肽的第二核酸序列,

其中所述嵌合MHC II $\alpha$ 多肽包含人HLA II类 $\alpha$ 多肽的细胞外部分以及小鼠MHCII $\alpha$ 多肽的至少跨膜结构域和胞质结构域,

所述嵌合MHC II $\beta$ 多肽包含人HLA II类 $\beta$ 多肽的细胞外部分以及小鼠MHCII $\beta$ 多肽的至少跨膜结构域和胞质结构域;

(d) 编码嵌合MHC I多肽的第三核酸序列,

其中所述嵌合MHC I多肽包含人HLA I类多肽的细胞外部分以及小鼠MHC I多肽的跨膜结构域和胞质结构域; 以及

(e) 未重排的T细胞受体(TCR) $\alpha$ 可变区序列, 所述未重排的T细胞受体(TCR) $\alpha$ 可变区序列包含至少一个人V $\alpha$ 区段和至少一个人J $\alpha$ 区段, 所述至少一个人V $\alpha$ 区段和所述至少一个人J $\alpha$ 区段能够在T细胞中重排以形成可操作地连接至小鼠TCR $\alpha$ 恒定基因序列的经重排的人V $\alpha$ /J $\alpha$ 序列, 和

未重排的TCR $\beta$ 可变区序列, 所述未重排的TCR $\beta$ 可变区序列包含至少一个人V $\beta$ 区段、至少一个人D $\beta$ 区段和至少一个人J $\beta$ 区段, 所述至少一个人V $\beta$ 区段、所述至少一个人D $\beta$ 区段和所述至少一个人J $\beta$ 区段能够在T细胞中重排以形成可操作地连接至小鼠TCR $\beta$ 恒定基因序列的经重排的人V $\beta$ /D $\beta$ /J $\beta$ 序列,

其中所述可操作地连接至小鼠TCR $\alpha$ 恒定基因序列的经重排的人V $\alpha$ /J $\alpha$ 序列编码人源化TCR $\alpha$ 链, 所述人源化TCR $\alpha$ 链包含可操作地连接至小鼠TCR $\alpha$ 恒定区的人TCR $\alpha$ 可变结构域, 并且

其中所述可操作地连接至小鼠TCR $\beta$ 恒定基因序列的经重排的人V $\beta$ /D $\beta$ /J $\beta$ 序列编码人源化TCR $\beta$ 链, 所述人源化TCR $\beta$ 链包含可操作地连接至小鼠TCR $\beta$ 恒定区的人TCR $\beta$ 可变结构域。

2. 根据权利要求1所述的经遗传修饰的小鼠基因组, 其特征在于, 所述小鼠基因组是种系基因组。

3. 根据权利要求1所述的经遗传修饰的小鼠基因组, 其中

(a) 所述第一核苷酸序列存在于内源CD4 T细胞辅助受体基因座;

(b) 所述第二核苷酸序列存在于内源CD8 $\alpha$ T细胞辅助受体基因座, 并且所述第三核苷酸序列存在于内源CD8 $\beta$ T细胞辅助受体基因座;

(c) 所述第一核酸序列存在于内源小鼠MHCII $\alpha$ 基因座, 并且所述第二核酸序列存在于内源小鼠MHCII $\beta$ 基因座;

(d) 所述第三核酸序列存在于内源小鼠MHC I基因座;和/或

(e) 所述未重排的TCR $\alpha$ 可变区序列存在于内源TCR $\alpha$ 可变基因座,并且所述未重排的TCR $\beta$ 可变区序列存在于内源TCR $\beta$ 可变基因座。

4. 根据权利要求3所述的经遗传修饰的小鼠基因组,其中

(a) 所述第一核苷酸序列存在于内源CD4 T细胞辅助受体基因座并且可操作地连接至内源小鼠CD4辅助受体启动子和调控元件;

(b) 所述第二核苷酸序列存在于内源CD8 $\alpha$ T细胞辅助受体基因座并且可操作地连接至内源小鼠CD8 $\alpha$ 多肽启动子和调控元件,并且所述第三核苷酸序列存在于内源CD8 $\beta$ T细胞辅助受体基因座并且可操作地连接至内源小鼠CD8 $\beta$ 多肽启动子和调控元件;

(c) 所述第一核酸序列存在于内源MHCII $\alpha$ 基因座并且可操作地连接至内源小鼠MHCII $\alpha$ 启动子和调控元件,并且所述第二核酸序列存在于内源MHCII $\beta$ 基因座并且可操作地连接至内源小鼠MHCII $\beta$ 启动子和调控元件;

(d) 所述第三核酸序列存在于内源MHC I基因座并且可操作地连接至内源小鼠MHC I启动子和调控元件;和/或

(e) 所述未重排的TCR $\alpha$ 可变区序列存在于内源TCR $\alpha$ 可变基因座并且可操作地连接至内源TCR $\alpha$ 启动子和调控元件,并且所述未重排的TCR $\beta$ 可变区序列存在于内源TCR $\beta$ 可变基因座并且可操作地连接至内源TCR $\beta$ 启动子和调控元件。

5. 根据权利要求1所述的经遗传修饰的小鼠基因组,其中

(a) 所述嵌合人/小鼠CD4辅助受体包含所述人CD4多肽的D1、D2和D3结构域,所述人CD4多肽的D1、D2和D3结构域可操作地连接至内源小鼠CD4多肽的跨膜结构域和胞质结构域;

(b) 所述嵌合人/小鼠CD8 $\alpha$ 多肽包含所述人CD8 $\alpha$ 多肽的IgV样结构域,所述人CD8 $\alpha$ 多肽的IgV样结构域可操作地连接至内源小鼠CD8 $\alpha$ 多肽的跨膜结构域和胞质结构域,并且所述嵌合CD8 $\beta$ 多肽包含所述人CD8 $\beta$ 多肽的IgV样结构域,所述嵌合CD8 $\beta$ 多肽包含所述人CD8 $\beta$ 多肽的IgV样结构域可操作地连接至内源小鼠CD8 $\beta$ 多肽的跨膜结构域和胞质结构域;

(c) 所述嵌合人/小鼠MHC II $\alpha$ 多肽包含人HLA II类 $\alpha$ 1和 $\alpha$ 2结构域,所述人HLA II类 $\alpha$ 1和 $\alpha$ 2结构域可操作地连接至内源小鼠MHCII $\alpha$ 多肽的小鼠跨膜结构域和胞质结构域,并且所述嵌合MHC II $\beta$ 多肽包含人HLA II类 $\beta$ 1和 $\beta$ 2结构域,所述人HLA II类 $\beta$ 1和 $\beta$ 2结构域可操作地连接至内源小鼠MHCII $\beta$ 多肽的跨膜结构域和胞质结构域;和/或

(d) 所述嵌合人/小鼠MHCI多肽包含人HLAI类 $\alpha$ 1、 $\alpha$ 2和 $\alpha$ 3结构域,所述人HLA I类 $\alpha$ 1、 $\alpha$ 2和 $\alpha$ 3结构域可操作地连接至内源小鼠MHC I多肽的跨膜结构域和胞质结构域。

6. 根据权利要求1所述的经遗传修饰的小鼠基因组,其中所述嵌合MHC II $\alpha$ 多肽的人细胞外部分由选自下组的人白细胞抗原(HLA) II类基因:HLA-DR、HLA-DQ或HLA-DP中任一编码的 $\alpha$ 链基因,其中所述嵌合MHCII $\beta$ 多肽的人细胞外部分由选自下组的人HLA II类基因:HLA-DR、HLA-DQ或人HLA-DP中任一编码的 $\beta$ 链基因,和/或

其中所述嵌合MHC I多肽的细胞外部分由人HLA-A基因、人HLA-B基因或人HLA-C基因编码。

7. 根据权利要求6所述的经遗传修饰的小鼠基因组,其中所述嵌合MHCII $\alpha$ 多肽和所述嵌合MHCII $\beta$ 多肽的人细胞外部分分别由人HLA-DR2基因的 $\alpha$ 链和 $\beta$ 链基因编码,并且

其中所述嵌合MHC I多肽的细胞外部分由人HLA-A基因编码。

8. 根据权利要求7所述的经遗传修饰的小鼠基因组,其中所述嵌合MHC II复合体包含人HLA-DR2蛋白的所述 $\alpha 1$ 、 $\alpha 2$ 、 $\beta 1$ 和 $\beta 2$ 结构域。

9. 根据权利要求7所述的经遗传修饰的小鼠基因组,其中所述嵌合人/小鼠MHCI多肽包含人HLA-A2蛋白的所述 $\alpha 1$ 、 $\alpha 2$ 和 $\alpha 3$ 结构域。

10. 根据权利要求1所述的经遗传修饰的小鼠基因组,其中所述未重排的TCR $\alpha$ 可变区序列包含人V $\alpha$ 基因区段的完整组库和人J $\alpha$ 基因区段的完整组库,和/或所述未重排的TCR $\beta$ 可变区序列包含人V $\beta$ 基因区段的完整组库、人D $\beta$ 基因区段的完整组库以及人J $\beta$ 基因区段的完整组库。

11. 根据权利要求1所述的经遗传修饰的小鼠基因组,其中内源小鼠TCR $\alpha$ 可变基因座缺少全部或基本上全部功能性内源V $\alpha$ 基因区段和/或缺少全部或基本上全部功能性内源J $\alpha$ 基因区段;和/或

其中内源小鼠TCR $\beta$ 可变基因座(a)缺少全部或基本上全部功能性内源V $\beta$ 基因区段,(b)缺少全部或基本上全部功能性内源D $\beta$ 基因区段,(c)缺少全部或基本上全部功能性内源J $\beta$ 基因区段,或(d)(a)、(b)和(c)的任何组合。

12. 根据权利要求1所述的经遗传修饰的小鼠基因组,其中

(a) 所述第一核苷酸序列包含编码人CD4多肽的细胞外部分的序列,所述编码人CD4多肽的细胞外部分的序列(i)替换编码内源小鼠CD4辅助受体多肽的细胞外部分的序列,并且(ii)在内源小鼠CD4辅助受体基因座处可操作地连接至内源小鼠CD4跨膜结构域编码序列和胞质结构域编码序列;

(b) 所述第二核苷酸序列包含编码人CD8 $\alpha$ 多肽的细胞外部分的序列,所述编码人CD8 $\alpha$ 多肽的细胞外部分的序列(i)替换编码内源小鼠CD8 $\alpha$ 多肽的细胞外部分的序列,并且(ii)在内源小鼠CD8 $\alpha$ 多肽基因座处可操作地连接至内源小鼠CD8 $\alpha$ 跨膜结构域编码序列和胞质结构域编码序列,和

所述第三核苷酸序列包含编码人CD8 $\beta$ 多肽的细胞外部分的序列,所述编码人CD8 $\beta$ 多肽的细胞外部分的序列(i)替换编码内源小鼠CD8 $\beta$ 多肽的细胞外部分的序列,并且(ii)在内源CD8 $\beta$ 基因座处可操作地连接至内源小鼠CD8 $\beta$ 跨膜结构域编码序列和胞质结构域编码序列;

(c) 所述第一核酸序列包含编码人HLA II类 $\alpha$ 多肽的细胞外部分的序列,所述编码人HLA II类 $\alpha$ 多肽的细胞外部分的序列(i)替换编码内源小鼠MHCII $\alpha$ 多肽的细胞外部分的序列,并且(ii)在内源小鼠MHCII $\alpha$ 基因座处可操作地连接至内源MHCII $\alpha$ 多肽跨膜结构域编码序列和胞质结构域编码序列,

所述第二核酸序列包含编码人HLA II类 $\beta$ 多肽的细胞外部分的序列,所述编码人HLA II类 $\beta$ 多肽的细胞外部分的序列(i)替换编码内源小鼠MHCII $\beta$ 多肽的细胞外部分的序列,并且(ii)在内源小鼠MHC II $\beta$ 基因座处可操作地连接至内源MHC II $\beta$ 多肽跨膜结构域编码序列和胞质结构域编码序列;

(d) 所述第三核酸序列包含编码人HLA I类多肽的细胞外部分的序列,所述编码人HLAI类多肽的细胞外部分的序列(i)替换编码内源小鼠MHCI多肽的细胞外部分的序列,并且(ii)在内源小鼠MHCI基因座处可操作地连接至内源MHC I多肽跨膜结构域编码序列和胞质结构域编码序列;和/或

(e) 所述未重排的TCR $\alpha$ 可变区序列在内源TCR $\alpha$ 可变区基因座处替换一个或多个内源V $\alpha$

基因区段和/或J $\alpha$ 基因区段,并且所述未重排的TCR $\beta$ 可变区序列在内源TCR $\beta$ 可变区基因座处替换一个或多个内源V $\beta$ 基因区段、D $\beta$ 基因区段和/或J $\beta$ 基因区段

13.根据权利要求1所述的经遗传修饰的小鼠基因组,其中所述小鼠基因组不包含(a)分别位于内源CD4辅助受体基因座和/或CD8辅助受体基因座的编码功能性内源小鼠CD4辅助受体和/或CD8辅助受体的序列,(b)位于内源TCR $\alpha$ 基因座的编码内源TCR $\alpha$ 可变结构域的序列,(c)位于内源TCR $\beta$ 基因座的编码内源TCR $\beta$ 可变结构域的序列,和/或(d)位于其内源MHC基因座的编码在细胞表面上表达的内源MHC I类的胞外结构域和/或在细胞表面上表达的内源MHC II类多肽的胞外结构域的序列。

14.根据权利要求1所述的经遗传修饰的小鼠基因组,还包含编码包含人 $\beta$ 2微球蛋白氨基酸序列的多肽的核苷酸序列。

15.根据权利要求14所述的经遗传修饰的小鼠基因组,其中所述小鼠基因组在内源小鼠 $\beta$ 2微球蛋白基因座处不包括编码功能性内源小鼠 $\beta$ 2微球蛋白多肽的序列。

16.根据权利要求14所述的经遗传修饰的小鼠基因组,其中所述编码包含人 $\beta$ 2微球蛋白氨基酸序列的多肽的核苷酸序列可操作地连接至内源小鼠 $\beta$ 2微球蛋白调控元件。

17.根据权利要求14所述的经遗传修饰的小鼠基因组,其中所述编码包含人 $\beta$ 2微球蛋白氨基酸序列的多肽的核苷酸序列包含人 $\beta$ 2微球蛋白基因的外显子2、外显子3和外显子4中所示的核苷酸序列。

18.根据权利要求17所述的经遗传修饰的小鼠基因组,其中所述编码包含人 $\beta$ 2微球蛋白氨基酸序列的多肽的核苷酸序列还包含小鼠 $\beta$ 2微球蛋白基因的外显子1中所示的核苷酸序列。

19.根据权利要求1所述的经遗传修饰的小鼠基因组,其中所述小鼠基因组包含全部功能性人TCR $\alpha$ 基因区段的至少50%和/或全部功能性人TCR $\beta$ 基因区段的至少50%。

20.经遗传修饰的小鼠基因组,包含:

(a)编码嵌合人/小鼠CD4的第一核苷酸序列,所述嵌合人/小鼠CD4包含人CD4多肽的D1、D2和D3结构域,所述人CD4多肽的D1、D2和D3结构域可操作地连接至小鼠CD4多肽的D4、跨膜结构域和胞质结构域;

(b)编码嵌合人/小鼠CD8 $\alpha$ 多肽的第二核苷酸序列和和编码嵌合人/小鼠CD8 $\beta$ 辅助受体多肽的第三核苷酸序列,它们各自分别包含人CD4辅助受体多肽的细胞外部分、CD8 $\alpha$ 辅助受体多肽的细胞外部分和CD8 $\beta$ 辅助受体多肽的细胞外部分,

其中所述嵌合人/小鼠CD8 $\alpha$ 多肽包含可操作地连接至内源小鼠CD8 $\alpha$ 多肽的跨膜结构域和胞质结构域的人CD8 $\alpha$ 多肽的IgV样结构域,并且所述嵌合人/小鼠CD8 $\beta$ 多肽包含可操作地连接至内源小鼠CD8 $\beta$ 多肽的跨膜结构域和胞质结构域的人CD8 $\beta$ 多肽的IgV样结构域;

(c)编码嵌合人/小鼠MHCII $\alpha$ 多肽的第一核酸序列和编码嵌合人/小鼠MHCII $\beta$ 多肽的第二核酸序列,

其中所述嵌合人/小鼠MHCII $\alpha$ 多肽包含可操作地连接至内源小鼠MHCII $\alpha$ 多肽的小鼠跨膜结构域和胞质结构域的人HLAII类 $\alpha$ 多肽的 $\alpha$ 1和 $\alpha$ 2结构域,并且所述嵌合人/小鼠MHCII $\beta$ 多肽包含可操作地连接至内源小鼠MHCII $\beta$ 多肽的跨膜结构域和胞质结构域的人HLA II类 $\beta$ 多肽的 $\beta$ 1和 $\beta$ 2结构域;

(d)编码嵌合人/小鼠MHC I多肽的第三核酸序列,其中所述嵌合人/小鼠MHC I多肽包

含可操作地连接至内源小鼠MHC I多肽的跨膜结构域和胞质结构域的人HLA I类多肽的 $\alpha 1$ 、 $\alpha 2$ 和 $\alpha 3$ 结构域；

(e) 未重排的T细胞受体(TCR) $\alpha$ 可变区序列,所述未重排的T细胞受体(TCR) $\alpha$ 可变区序列包含至少一个人V $\alpha$ 区段和至少一个人J $\alpha$ 区段,所述至少一个人V $\alpha$ 区段和所述至少一个人J $\alpha$ 区段能够在T细胞中重排以形成可操作地连接至小鼠TCR $\alpha$ 恒定基因序列的经重排的人V $\alpha$ /J $\alpha$ 序列,和

未重排的TCR $\beta$ 可变区序列,所述未重排的TCR $\beta$ 可变区序列包含至少一个人V $\beta$ 区段、至少一个人D $\beta$ 区段和至少一个人J $\beta$ 区段,所述至少一个人V $\beta$ 区段、所述至少一个人D $\beta$ 区段和所述至少一个人J $\beta$ 区段能够在T细胞中重排以形成可操作地连接至小鼠TCR $\beta$ 恒定基因序列的经重排的人V $\beta$ /D $\beta$ /J $\beta$ 序列,

其中所述可操作地连接至小鼠TCR $\alpha$ 恒定基因序列的经重排的人V $\alpha$ /J $\alpha$ 序列编码人源化TCR $\alpha$ 链,所述人源化TCR $\alpha$ 链包含可操作地连接至小鼠TCR $\alpha$ 恒定区的人TCR $\alpha$ 可变结构域,并且

其中所述可操作地连接至小鼠TCR $\beta$ 恒定基因序列的经重排的人V $\beta$ /D $\beta$ /J $\beta$ 序列编码人源化TCR $\beta$ 链,所述人源化TCR $\beta$ 链包含可操作地连接至小鼠TCR $\beta$ 恒定区的人TCR $\beta$ 可变结构域;

(f) 编码人或人源化 $\beta 2$ 微球蛋白多肽的多核苷酸序列,并且所述多核苷酸序列包括核苷酸序列,所述核苷酸序列包含可操作地连接至人 $\beta 2$ 微球蛋白基因的外显子2、外显子3和外显子4中的核苷酸序列的小鼠 $\beta 2$ 微球蛋白基因的外显子1中的核苷酸序列。

21. 根据权利要求20所述的经遗传修饰的小鼠基因组,其中所述嵌合人/小鼠MHC II $\alpha$ 多肽是嵌合HLA-DR/H-2E $\alpha$ 多肽,所述嵌合人/小鼠MHCII $\beta$ 多肽是编码嵌合人/小鼠HLA-DR/H-2E $\beta$ 多肽,并且所述嵌合人/小鼠MHC I多肽编码嵌合人/小鼠HLA-A/H-2K多肽,并且其中所述小鼠表达HLA-A/H-2K和HLA-DR/H-2E蛋白质。

22. 一种制备经遗传修饰的小鼠的方法,包括修饰所述小鼠基因组以包含:

- (a) 编码嵌合CD4辅助受体的第一核苷酸序列;
- (b) 编码嵌合CD8 $\alpha$ 多肽的第二核苷酸序列和编码嵌合CD8 $\beta$ 多肽的第三核苷酸序列,
- (c) 编码嵌合MHCII $\alpha$ 多肽的第一核苷酸序列和编码嵌合MHCII $\beta$ 多肽的第二核苷酸序列,
- (d) 编码嵌合MHC I多肽的第三核苷酸序列,和

(e) 未重排的T细胞受体(TCR) $\alpha$ 可变区序列,所述未重排的T细胞受体(TCR) $\alpha$ 可变区序列包含至少一个人V $\alpha$ 区段和至少一个人J $\alpha$ 区段,所述至少一个人V $\alpha$ 区段和至少一个人J $\alpha$ 区段能够在T细胞中重排以形成可操作地连接至小鼠TCR $\alpha$ 恒定区序列的经重排的人V $\alpha$ /J $\alpha$ 序列,和未重排的TCR $\beta$ 可变区序列,所述未重排的TCR $\beta$ 可变区序列包含至少一个人V $\beta$ 区段、至少一个人D $\beta$ 区段和至少一个人J $\beta$ 区段,所述至少一个人V $\beta$ 区段、所述至少一个人D $\beta$ 区段和所述至少一个人J $\beta$ 区段能够在T细胞中重排以形成可操作地连接至小鼠TCR $\beta$ 恒定区序列的经重排的人V $\beta$ /D $\beta$ /J $\beta$ 序列。

23. 根据权利要求22所述的方法,包括修饰所述小鼠基因组以使其包含编码人或人源化 $\beta 2$ 微球蛋白多肽的核苷酸序列。

24. 根据权利要求22所述的方法,其中修饰小鼠基因组包括在一个或多个小鼠ES细胞中的同源重组,使得以任何顺序在所述小鼠ES细胞的所述基因组中引入所述第一核苷酸序

列、所述第二核苷酸序列和所述第三核苷酸序列、所述未重排的TCR $\alpha$ 可变区序列、所述未重排的TCR $\beta$ 可变区序列、所述第一核苷酸序列、所述第二核苷酸序列和所述第三核苷酸序列。

25. 根据权利要求22所述的方法, 其中

(a) 所述第一核苷酸序列存在于内源CD4 T细胞辅助受体基因座;

(b) 所述第二核苷酸序列存在于内源CD8 $\alpha$ T细胞辅助受体基因座, 并且所述第三核苷酸序列存在于内源CD8 $\beta$ T细胞辅助受体基因座;

(c) 所述第一核苷酸序列存在于内源小鼠MHCII $\alpha$ 基因座, 并且所述第二核苷酸序列存在于内源小鼠MHCII $\beta$ 基因座;

(d) 所述第三核苷酸序列存在于内源小鼠MHC I基因座; 和/或

(e) 所述未重排的TCR $\alpha$ 可变区序列存在于内源TCR $\alpha$ 可变区基因座, 并且所述未重排的TCR $\beta$ 可变区序列存在于内源TCR $\beta$ 可变区基因座。

26. 根据权利要求25所述的方法, 其中

(a) 所述第一核苷酸序列存在于内源CD4 T细胞辅助受体基因座并且可操作地连接至内源小鼠CD4辅助受体启动子和调控元件;

(b) 所述第二核苷酸序列存在于内源CD8 $\alpha$ T细胞辅助受体基因座并且可操作地连接至内源小鼠CD8 $\alpha$ 多肽启动子和调控元件, 并且所述第三核苷酸序列存在于内源CD8 $\beta$ T细胞辅助受体基因座并且可操作地连接至内源小鼠CD8 $\beta$ 多肽启动子和调控元件;

(c) 所述第一核苷酸序列存在于内源小鼠MHCII $\alpha$ 基因座并且可操作地连接至内源小鼠MHCII $\alpha$ 启动子和调控元件, 并且所述第二核苷酸序列存在于内源小鼠MHCII $\beta$ 基因座并且可操作地连接至内源小鼠MHCII $\beta$ 启动子和调控元件;

(d) 所述第三核苷酸序列存在于内源小鼠MHC I基因座并且可操作地连接至内源小鼠MHC I启动子和调控元件; 和/或

(e) 所述未重排的TCR $\alpha$ 可变区序列存在于内源TCR $\alpha$ 可变基因座并且可操作地连接至内源TCR $\alpha$ 启动子和调控元件, 并且所述未重排的TCR $\beta$ 可变区序列存在于内源TCR $\beta$ 可变基因座并且可操作地连接至内源TCR $\beta$ 启动子和调控元件。

27. 根据权利要求22所述的方法, 其中

(a) 所述嵌合CD4辅助受体包含所述人CD4多肽的D1、D2和D3结构域, 所述人CD4多肽的D1、D2和D3结构域可操作地连接至内源小鼠CD4多肽的跨膜结构域和胞质结构域;

(b) 所述嵌合CD8 $\alpha$ 多肽包含所述人CD8 $\alpha$ 多肽的IgV样结构域, 所述人CD8 $\alpha$ 多肽的IgV样结构域可操作地连接至内源小鼠CD8 $\alpha$ 多肽的跨膜结构域和胞质结构域, 并且所述嵌合CD8 $\beta$ 多肽包含所述人CD8 $\beta$ 多肽的IgV样结构域, 所述人CD8 $\beta$ 多肽的IgV样结构域可操作地连接至内源小鼠CD8 $\beta$ 多肽的跨膜结构域和胞质结构域;

(c) 所述嵌合MHCII $\alpha$ 多肽包含人HLAII类 $\alpha$ 1和 $\alpha$ 2结构域, 所述人HLAII类 $\alpha$ 1和 $\alpha$ 2结构域可操作地连接至内源小鼠MHCII $\alpha$ 多肽的跨膜结构域和胞质结构域, 并且所述嵌合MHCII $\beta$ 多肽包含人HLA II类 $\beta$ 1和 $\beta$ 2结构域, 所述人HLA II类 $\beta$ 1和 $\beta$ 2结构域可操作地连接至内源小鼠MHCII $\beta$ 多肽的跨膜结构域和胞质结构域, 和/或

(d) 所述嵌合MHC I多肽包含人HLA I类 $\alpha$ 1、 $\alpha$ 2和 $\alpha$ 3结构域, 所述人HLAI类 $\alpha$ 1、 $\alpha$ 2和 $\alpha$ 3结构域可操作地连接至内源小鼠MHC I多肽的跨膜结构域和胞质结构域。

28. 根据权利要求22所述的方法, 其中

所述嵌合MHCII $\alpha$ 多肽的人细胞外部分由选自下组的人白细胞抗原 (HLA) II类基因编码:HLA-DR、HLA-DQ和HLA-DP中任一的 $\alpha$ 链基因,

其中所述嵌合MHCII $\beta$ 多肽的人细胞外部分由选自下组的人HLA II类基因编码:HLA-DR、HLA-DQ和HLA-DP中任一的 $\beta$ 链基因,和/或

其中所述嵌合MHC I多肽的细胞外部分由人HLA-A基因、人HLA-B基因或人HLA-C基因编码。

29. 根据权利要求28所述的方法,其中所述嵌合MHCII $\alpha$ 多肽和所述嵌合MHC II $\beta$ 多肽的人细胞外部分分别由HLA-DR2基因的 $\alpha$ 链和 $\beta$ 链基因编码,并且

其中所述嵌合MHC I多肽的细胞外部分由人HLA-A基因编码。

30. 根据权利要求29所述的方法,其中所述嵌合MHC II复合体包含人HLA-DR2蛋白的 $\alpha$ 1、 $\alpha$ 2、 $\beta$ 1和 $\beta$ 2结构域。

31. 根据权利要求29所述的方法,其中所述嵌合MHC I多肽包含人HLA-A2蛋白的 $\alpha$ 1、 $\alpha$ 2和 $\alpha$ 3结构域。

32. 根据权利要求22所述的方法,其中所述未重排的TCR $\alpha$ 可变区序列包含人V $\alpha$ 基因区段的完整组库和人J $\alpha$ 基因区段的完整组库,和/或所述未重排的TCR $\beta$ 可变区序列包含人V $\beta$ 基因区段的完整组库、人D $\beta$ 基因区段的完整组库以及人J $\beta$ 基因区段的完整组库。

33. 根据权利要求22所述的方法,其中内源小鼠TCR $\alpha$ 可变基因座缺少全部或基本上全部功能性内源V $\alpha$ 基因区段和/或缺少全部或基本上全部功能性内源J $\alpha$ 基因区段;和/或

其中内源小鼠TCR $\beta$ 可变基因座(a)缺少全部或基本上全部功能性内源V $\beta$ 基因区段,(b)缺少全部或基本上全部功能性内源D $\beta$ 基因区段,(c)缺少全部或基本上全部功能性内源J $\beta$ 基因区段,或(d) (a)、(b)和(c)的任何组合。

34. 根据权利要求22所述的方法,其中

(a) 所述第一核苷酸序列包含编码人CD4多肽的细胞外部分的序列,所述编码人CD4多肽的细胞外部分的序列(i)替换编码内源小鼠CD4辅助受体多肽的细胞外部分的序列,并且(ii)在内源小鼠CD4辅助受体基因座处可操作地连接至内源小鼠CD4跨膜结构域编码序列和胞质结构域编码序列;

(b) 所述第二核苷酸序列包含编码人CD8 $\alpha$ 多肽的细胞外部分的序列,所述编码人CD8 $\alpha$ 多肽的细胞外部分的序列(i)替换编码内源小鼠CD8 $\alpha$ 多肽的细胞外部分的序列,并且(ii)在内源小鼠CD8 $\alpha$ 多肽基因座处可操作地连接至内源小鼠CD8 $\alpha$ 跨膜结构域编码序列和胞质结构域编码序列,和

所述第三核苷酸序列包含编码人CD8 $\beta$ 多肽的细胞外部分的序列,所述编码人CD8 $\beta$ 多肽的细胞外部分的序列(i)替换编码内源小鼠CD8 $\beta$ 多肽的细胞外部分的序列,并且(i)在内源CD8 $\beta$ 基因座处可操作地连接至内源小鼠CD8 $\beta$ 跨膜结构域编码序列和胞质结构域编码序列;

(c) 所述第一核苷酸序列包含编码人HLA II类 $\alpha$ 多肽的细胞外部分的序列,所述编码人HLA II类 $\alpha$ 多肽的细胞外部分的序列(i)替换编码内源小鼠MHCII $\alpha$ 多肽的细胞外部分的序列,并且(ii)在内源小鼠MHCII $\alpha$ 基因座处可操作地连接至内源MHCII $\alpha$ 多肽跨膜结构域编码序列和胞质结构域编码序列,

所述第二核苷酸序列包含编码人HLA II类 $\beta$ 多肽的细胞外部分的序列,所述编码人HLA II类 $\beta$ 多肽的细胞外部分的序列(i)替换编码内源小鼠MHCII $\beta$ 多肽的细胞外部分的序列,并



且(ii)在内源小鼠MHC IIB基因座处可操作地连接至内源MHC IIB多肽跨膜结构域编码序列和胞质结构域编码序列;

(d)所述第三核酸序列包含编码人HLA I类多肽的细胞外部分的序列,所述编码人HLA I类多肽的细胞外部分的序列(i)替换编码内源小鼠MHCI多肽的细胞外部分的序列,并且(ii)在内源小鼠MHCI基因座处可操作地连接至内源MHC I多肽跨膜结构域编码序列和胞质结构域编码序列;和/或

(e)所述未重排的TCR $\alpha$ 可变区序列在内源TCR $\alpha$ 可变区基因座处替换一个或多个内源V $\alpha$ 基因区段和/或J $\alpha$ 基因区段,并且所述未重排的TCR $\beta$ 可变区序列在内源TCR $\beta$ 可变区基因座处替换一个或多个内源V $\beta$ 基因区段、D $\beta$ 基因区段和/或J $\beta$ 基因区段。

35.根据权利要求22所述的方法,其中所述小鼠在细胞表面上(a)分别不从内源CD4辅助受体基因座和/或CD8辅助受体基因座表达功能性内源小鼠CD4辅助受体和/或CD8辅助受体,(b)不从内源TCR $\alpha$ 基因座表达内源TCR $\alpha$ 可变结构域,(c)不从内源TCR $\beta$ 基因座表达内源TCR $\beta$ 可变结构域,和/或(d)不从其内源MHC基因座表达内源MHC I类和/或II类多肽的细胞外结构域。

36.根据权利要求22所述的方法,还包含编码包含人 $\beta$ 2微球蛋白氨基酸序列的多肽的核苷酸序列,其中所述小鼠表达人或人源化 $\beta$ 2微球蛋白多肽。

37.根据权利要求36所述的方法,其中所述小鼠不从内源小鼠 $\beta$ 2微球蛋白基因座表达功能性内源小鼠 $\beta$ 2微球蛋白多肽。

38.根据权利要求36所述的方法,其中所述编码包含人 $\beta$ 2微球蛋白氨基酸序列的多肽的核苷酸序列可操作地连接至内源小鼠 $\beta$ 2微球蛋白调控元件。

39.根据权利要求36所述的方法,其中所述编码包含人 $\beta$ 2微球蛋白氨基酸序列的多肽的核苷酸序列包含人 $\beta$ 2微球蛋白基因的外显子2、外显子3和外显子4中的核苷酸序列。

40.根据权利要求39所述的方法,其中所述编码包含人 $\beta$ 2微球蛋白氨基酸序列的多肽的核酸序列还包含小鼠 $\beta$ 2微球蛋白基因的外显子1中的核苷酸序列。。

41.根据权利要求22所述的方法,还包括从所述一个或多个小鼠ES细胞产生小鼠。

42.一种获得TCR蛋白或编码TCR蛋白的序列的方法,所述TCR蛋白包括抗原特异性的人TCR可变结构域,所述方法包括

从包含根据权利要求1-21中任一项所述的基因组的小鼠中或从根据权利要求22-41中任一项所述的方法制备的小鼠中分离T细胞、结合所述抗原的TCR蛋白、和/或编码结合所述抗原的TCR蛋白的TCR $\alpha$ 可变结构域的核酸序列和编码结合所述抗原的TCR蛋白的TCR $\beta$ 可变结构域的核酸序列。

43.根据权利要求42所述的方法,包括:

从所述小鼠处分离(i)所述编码TCR $\alpha$ 可变结构域的核酸序列和(ii)所述编码TCR $\beta$ 可变结构域的核酸序列,并且

还包括:

向宿主细胞中引入(i)所述编码TCR $\alpha$ 可变结构域的核酸序列,所述TCR $\alpha$ 可变结构域可操作地连接至人TCR $\alpha$ 恒定区,和(ii)所述编码TCR $\beta$ 可变结构域的核酸序列,所述TCR $\beta$ 可变结构域可操作地连接至人TCR $\beta$ 恒定区,以及

在(i)所述编码TCR $\alpha$ 可变结构域的核酸序列和(ii)所述编码TCR $\beta$ 可变结构域的核酸序

列表表达的充分条件下培养宿主细胞,其中所述TCR $\alpha$ 可变结构域可操作地连接至人TCR $\alpha$ 恒定区,所述TCR $\beta$ 可变结构域可操作地连接至人TCR $\beta$ 恒定区,

其中所述编码TCR $\alpha$ 可变结构域的核酸序列和所述编码TCR $\beta$ 可变结构域的核酸序列位于相同或不同的表达载体上。

44. 根据权利要求42所述的方法,其中所述抗原为肿瘤抗原、病毒抗原或细菌抗原。

45. 一种T细胞,所述T细胞根据权利要求42所述的方法分离。

46. 一种杂交瘤,所述杂交瘤由根据权利要求42所述的方法分离的所述T细胞产生。

47. 一种人T细胞受体可变结构域,所述人T细胞受体可变结构域根据权利要求42所述的方法分离。

48. 一种核酸,所述核酸包含根据权利要求42所述的方法分离的编码所述TCR $\alpha$ 可变结构域和/或TCR $\beta$ 可变结构域的核酸序列。

49. 一种宿主细胞,所述宿主细胞包含与根据权利要求48所述的核酸。

50. 一种表达载体,所述表达载体包含根据权利要求48所述的核酸。

51. 根据权利要求50所述的表达载体,其中所述表达载体包含

编码所述人TCR $\alpha$ 可变结构域的核酸序列,所述核酸序列可操作地连接至编码人TCR $\alpha$ 恒定区的序列,和/或

编码所述人TCR $\beta$ 可变结构域的核酸序列,所述核酸序列可操作地连接至编码人TCR $\beta$ 恒定区的序列。

52. 一种组合物,所述组合物包含小鼠的第一细胞和第二细胞,其中所述第一细胞表达

(a) 嵌合人/小鼠T细胞辅助受体,所述嵌合人/小鼠T细胞辅助受体包含人T细胞辅助受体的细胞外部分,所述人T细胞辅助受体的细胞外部分可操作地连接至小鼠T细胞辅助受体的跨膜结构域和胞质结构域,和

(b) 嵌合人/小鼠TCR包括以下一者或两者:(i) 嵌合人/小鼠TCR $\alpha$ 链,所述嵌合人/小鼠TCR $\alpha$ 链包括人TCR $\alpha$ 可变结构域,所述人TCR $\alpha$ 可变结构域可操作地连接至小鼠TCR $\alpha$ 恒定结构域和(ii) 嵌合人/小鼠TCR $\beta$ 链,所述嵌合人/小鼠TCR $\beta$ 链包括人TCR $\beta$ 可变结构域,所述人TCR $\beta$ 可变结构域可操作地连接至小鼠TCR $\beta$ 恒定结构域,并且

其中所述第二细胞表达嵌合人/小鼠MHC多肽,所述嵌合人/小鼠MHC多肽包含人MHC多肽的细胞外部分,所述人MHC多肽的细胞外部分可操作地连接至小鼠MHC多肽的跨膜结构域和胞质结构域,并且

其中所述嵌合人/小鼠MHC多肽与所述嵌合人/小鼠T细胞辅助受体结合。

53. 根据权利要求52所述的组合物,其中所述第一细胞为小鼠T细胞。

54. 根据权利要求52所述的组合物,其中所述第二细胞为小鼠抗原递呈细胞。

## 非人动物中的人源化T细胞介导的免疫应答

[0001] 相关申请的交叉引用

[0002] 本申请要求美国临时申请序列号62/143,687 (2015年4月6日提交)、62/158,804 (2015年5月8日提交)和62/186,935 (2015年6月30日提交)的优先权,每个申请均据此以引用方式并入。

[0003] 序列表

[0004] 通过EFS-Web以电子方式将序列表的正式文本作为ASCII格式的序列表提交,该文件名称为2016-04-06-10145W001-SEQ-LIST\_ST25.txt,创建日期为2016年4月6日,文件大小为56.7千字节,并且该文件与本说明书同时提交。该ASCII格式文档中所含的序列表是本说明书的一部分,并且全文以引用的方式并入本文。

### 技术领域

[0005] 本发明涉及能够产生基本上人(源化)T细胞介导的免疫应答并且表达(i)一种或多种人(源化)T细胞辅助受体(例如,CD4和/或CD8(例如,CD8 $\alpha$ 和/或CD8 $\beta$ 、(ii)与一种或多种人(源化)T细胞辅助受体结合的一种或多种人(源化)主要组织相容性复合体(例如,MHC II(例如,MHC $\alpha$ 和/或MHC II $\beta$ 和/或MHC I(例如,MHC I $\alpha$ 和/或 $\beta$ 2微球蛋白))和/或(iii)人(源化)T细胞受体(TCR)(例如,TCR $\alpha$ 和/或TCR $\beta$ )的非人动物(例如,啮齿动物,例如小鼠或大鼠);从非人动物分离的胚胎、组织、细胞和/或核酸;制备非人动物的方法;以及使用非人动物开发人类治疗剂的方法。

### 背景技术

[0006] 在适应性免疫应答中,外来抗原被B淋巴细胞(例如免疫球蛋白)和T淋巴细胞(例如,T细胞受体,也称为TCR)上的受体分子识别。这些外来抗原通过专门的蛋白质(通常称为主要组织相容性复合体(MHC)分子)以肽片段的形式递呈于细胞表面上,并且在人类中具体地称为人白细胞抗原(HLA)。在T细胞介导的应答期间,由MHC分子递呈的抗原被T细胞受体识别。然而,对于有效的免疫应答,不仅仅需要MHC-抗原复合体的T细胞受体识别。还需要T细胞辅助受体分子(例如CD4或CD8)与MHC的不变部分的结合。

[0007] T细胞有若干种类,包括辅助T细胞和细胞毒性T细胞。辅助T细胞表达辅助受体CD4并识别与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细胞而产生抗体,活化表达MHC II的巨噬细胞而破坏病原体等等。CD4和T细胞受体与相同MHC II递呈的外来抗原的结合使T细胞对该抗原明显更敏感。

[0008] 相比之下,细胞毒性T细胞(CTL)表达辅助受体CD8并识别与MHC I分子结合的外来抗原。CTL专门杀死任何携带由其自身的膜结合TCR识别的MHC I结合肽的细胞。当细胞展示来源于通常不存在的细胞蛋白(例如病毒、肿瘤或其他非自身起源)的肽时,这些肽被CTL识别,CTL被活化并杀死展示该肽的细胞。与CD4类似,CD8的参与使CTL对MHC I递呈的抗原更敏感。

[0009] 由于存在耐受机制,并非所有抗原都会引起T细胞活化。然而,在某些疾病(例如,

癌症、自身免疫疾病)中,来源于自身蛋白质的肽成为免疫系统的细胞组分的靶标,这导致递呈此类肽的细胞受到破坏。在识别临床上重要的抗原(例如与各种类型的癌症相关的抗原)和/或结合临床上重要的抗原的TCR序列方面已经取得了显著进步。然而,为了改善将在人T细胞中引发适当的应答的临床上重要的肽和/或能够结合临床上重要的抗原的TCR的鉴定和选择(例如,对于癌症的过继性免疫疗法、防治自身免疫的T细胞疫苗接种等等),仍然需要模拟人免疫系统诸方面的体内和体外系统。因此,需要能够展示人免疫系统的组分,特别是T细胞免疫应答的组分的生物系统(例如,经遗传修饰的非人动物和细胞)。

## 发明内容

[0010] 如本文所公开,包含基本上人源化T细胞免疫系统的经遗传修饰的非人动物的胸腺具有与对照动物相似的胸腺细胞和CD3+T细胞绝对数量。另外,这些细胞显示出与对照动物相当的向单一阳性T细胞的发育,并且能够产生针对抗原(例如病毒抗原)的强烈人细胞应答。非人动物的人细胞应答通常包括表达人或人源化T细胞受体(TCR)可变结构域的活化的非人T细胞,这些可变结构域识别由人白细胞抗原(HLA)细胞外结构域形成的肽结合槽中递呈的抗原,该肽结合槽可以在非人抗原递呈细胞的表面上表达。在一些实施方案中,基本上人源化的T细胞免疫系统包含:

[0011] (A) 表达以下物质的非人T细胞:

[0012] (i) 包含人T细胞辅助受体的细胞外部分的一部分或全部的T细胞辅助受体多肽,例如包含一个或多个人T细胞辅助受体细胞外结构域的T细胞辅助受体多肽,使得T细胞辅助受体多肽能够与以下结构域进行结合和/或与以下结构域结合:

[0013] (a) 人或人源化HLA分子的一个或多个细胞外结构域(例如,作为T细胞辅助受体多肽的结合位点的第一人HLA细胞外结构域和/或形成肽结合槽(例如与第三人HLA细胞外结构域)的第二人HLA细胞外结构域),

[0014] (b) 人或人源化TCR可变结构域(例如,分别由至少一个人TCR $\alpha$ 和/或TCR $\beta$ 可变区基因区段编码的人或人源化TCR $\alpha$ 可变结构域和/或人或人源化TCR $\beta$ 可变结构域)的细胞外结构域,和/或

[0015] (c) 人TCR恒定结构域的细胞外结构域,和

[0016] (ii) 包含至少人TCR可变结构域的T细胞受体(TCR);和任选地

[0017] (B) 在人HLA背景下递呈抗原的非人抗原递呈细胞,例如在其细胞表面上表达至少一个MHC分子的非人抗原递呈细胞,所述MHC分子包含由两个人HLA细胞外结构域形成的肽结合槽,并且能够进行活化和/或活化非人T细胞。

[0018] 在一个方面,非人T细胞和非人抗原递呈细胞存在于相同的非人动物中或分离自相同的非人动物。

[0019] 因此,本文提供了经遗传工程改造而表达以下物质的非人动物(例如,啮齿动物,例如小鼠或大鼠):

[0020] (A) 人或人源化T细胞辅助受体(例如,人或人源化CD4和/或人或人源化CD8(例如,人或人源化CD8 $\alpha$ 和/或人或人源化CD8 $\beta$ )),

[0021] (B) 与人或人源化T细胞辅助受体相连的人或人源化主要组织相容性复合体(例如,结合人或人源化CD4的人或人源化MHC II(例如,人或人源化MHC II $\alpha$ 和/或人或人源化

MHC II $\beta$ ) 和/或结合人或人源化CD8的人或人源化MHC I (例如,人或人源化MHC I $\alpha$ 和任选的人或人源化 $\beta$ 2微球蛋白),和/或

[0022] (C) 人或人源化T细胞受体 (TCR);

[0023] 以及表达这些物质的胚胎、组织和细胞,及编码这些物质的核酸。还提供了制备和使用所公开的非人动物的方法。

[0024] 在一个方面,提供了一种经遗传修饰的非人动物,其包含:

[0025] (A) 人源化CD4辅助受体和/或包含人源化CD8 $\alpha$ 多肽和人源化CD8 $\beta$ 多肽的人源化CD8辅助受体 (例如,该非人动物例如在其种系基因组中包含编码嵌合人/非人CD4多肽的第一核苷酸序列,和/或编码嵌合人/非人CD8 $\alpha$ 多肽的第二核苷酸序列以及编码嵌合人/非人CD8 $\beta$ 多肽的第三核苷酸序列),

[0026] 其中每种人源化T细胞辅助受体多肽包含非人T细胞辅助受体的至少跨膜结构域和胞质结构域,例如,其中所述人源化CD4辅助受体包含非人CD4辅助受体的至少跨膜结构域和胞质结构域,和/或人源化CD8辅助受体包含非人CD8 $\alpha$ 和非人CD8 $\beta$ 多肽的至少跨膜结构域和胞质结构域,

[0027] 其中每种嵌合T细胞辅助受体多肽包含人T细胞辅助受体的细胞外部分的一部分或全部,例如人T细胞辅助受体的一个或多个细胞外结构域,例如与HLA分子相连的人T细胞辅助受体的至少细胞外结构域,例如其中人源化CD4辅助受体包含负责与MHC II、T细胞受体可变结构域、T细胞受体恒定结构域或它们的组合相互作用的人CD4的细胞外部分 (或其部分,例如一个或多个细胞外结构域),和/或例如其中人源化CD8辅助受体包含负责与MHC I、T细胞受体可变结构域、T细胞受体恒定结构域或它们的组合相互作用的人CD8 $\alpha$ 和人CD8 $\beta$ 的细胞外部分 (或其部分,例如细胞外结构域);

[0028] (B) 人 (源化) TCR (例如,非人动物例如在其种系基因组中包含可操作地连接至非人TCR $\alpha$ 恒定基因序列的未重排T细胞受体 (TCR)  $\alpha$ 可变基因座和/或可操作地连接至非人TCR $\beta$ 恒定基因序列的未重排TCR $\beta$ 可变基因座,所述未重排T细胞受体 (TCR)  $\alpha$ 可变基因座包含至少一个人V $\alpha$ 区段和至少一个人J $\alpha$ 区段,所述未重排TCR $\beta$ 可变基因座包含至少一个人V $\beta$ 区段、至少一个人D $\beta$ 区段和至少一个人J $\beta$ 区段);以及任选地,

[0029] (C) 与人源化CD4辅助受体结合的人 (源化) MHC II复合体和/或与人源化CD8辅助受体结合的人 (源化) MHC I复合体 (例如,非人动物例如在其种系基因组中包含编码嵌合人/非人MHC II $\alpha$ 多肽的第一核酸序列和编码嵌合人/非人MHC II $\beta$ 多肽的第二核酸序列,和/或编码嵌合人/非人MHC I多肽的第三核酸序列),

[0030] 其中每种嵌合MHC多肽包含人MHC多肽 (例如,HLA多肽) 的至少细胞外部分 (或其一部分),其单独存在 (例如,MHC I) 或与另一嵌合MHC多肽 (例如,MHC II $\alpha$ 和MHC II $\beta$ 复合时) 分别能够与人 (源化) CD8辅助受体或人 (源化) CD4辅助受体结合并且在HLA背景下递呈肽,例如,其中人源化MHC II复合体包含 (i) 嵌合人/非人MHC II $\alpha$ 多肽,其包含人HLA II类 $\alpha$ 多肽的 $\alpha$ 1和 $\alpha$ 2结构域以及非人HLA II类 $\alpha$ 多肽的跨膜结构域和胞质结构域,和 (ii) 嵌合人/非人MHC II $\beta$ 多肽包含人HLA II类 $\beta$ 多肽的 $\beta$ 1和 $\beta$ 2结构域、非人HLA II类 $\beta$ 多肽的跨膜结构域和胞质结构域,和/或其中人源化MHC I复合体包含人MHC I多肽的 $\alpha$ 1、 $\alpha$ 2和 $\alpha$ 3结构域,以及任选的人 (源化)  $\beta$ 2微球蛋白。

[0031] 在一些实施方案中,非人动物包含:

[0032] (A) 人源化CD8辅助受体和包含人源化CD8 $\alpha$ 多肽和人源化CD8 $\beta$ 多肽的人源化CD4辅助受体(例如,非人动物例如在其种系基因组中包含编码嵌合人/非人CD4多肽的第一核苷酸序列、编码嵌合人/非人CD8 $\alpha$ 多肽的第二核苷酸序列以及编码嵌合人/非人CD8 $\beta$ 多肽的第三核苷酸序列),

[0033] 其中每种人源化T细胞辅助受体多肽包含非人T细胞辅助受体的至少跨膜结构域和胞质结构域,例如,其中所述人源化CD4辅助受体包含非人CD4辅助受体的至少跨膜结构域和胞质结构域,并且人源化CD8辅助受体包含非人CD8 $\alpha$ 和非人CD8 $\beta$ 多肽的至少跨膜结构域和胞质结构域,

[0034] 其中每种嵌合T细胞辅助受体多肽包含人T细胞辅助受体的细胞外部分的一部分或全部,例如人T细胞辅助受体的一个或多个细胞外结构域,例如与HLA分子结合的人T细胞辅助受体的至少细胞外结构域,例如其中人源化CD4辅助受体包含负责与MHC II、T细胞受体可变结构域、T细胞受体恒定结构域或它们的组合相互作用的人CD4的细胞外部分(或其部分,例如一个或多个细胞外结构域),和/或例如其中人源化CD8辅助受体包含负责与MHC I、T细胞受体可变结构域、T细胞受体恒定结构域或它们的组合相互作用的人CD8 $\alpha$ 和人CD8 $\beta$ 的细胞外部分(或其部分,例如细胞外结构域);

[0035] (B) 人源化TCR(例如,非人动物例如在其种系基因组中包含具有至少一个人V $\alpha$ 区段和至少一个人J $\alpha$ 区段并可操作地连接至非人TCR $\alpha$ 恒定基因序列的未重排T细胞受体(TCR) $\alpha$ 可变基因座,和/或具有至少一个人V $\beta$ 区段、至少一个人D $\beta$ 区段和至少一个人J $\beta$ 区段并可操作地连接至非人TCR $\beta$ 恒定基因序列的未重排TCR $\beta$ 可变基因座);以及

[0036] (C) 与人源化CD4辅助受体结合的人源化MHC II复合体和与人源化CD8辅助受体结合的人源化MHC I复合体(例如,非人动物例如在其种系基因组中包含编码嵌合人/非人MHC II $\alpha$ 多肽的第一核苷酸序列、编码嵌合人/非人MHC II $\beta$ 多肽的第二核苷酸序列以及编码嵌合人/非人MHC I多肽的第三核苷酸序列),

[0037] 其中每种嵌合MHC多肽包含人MHC多肽(例如,HLA多肽)的至少细胞外部分(或其一部分),其单独存在(例如,MHC I)或与另一嵌合MHC多肽(例如,MHC II $\alpha$ 和MHC II $\beta$ 复合时)分别能够与人源化CD8辅助受体或人源化CD4辅助受体结合并且在HLA背景下递呈肽,例如,其中人源化MHC II复合体包含(i)嵌合人/非人MHC II $\alpha$ 多肽,其包含人HLA II类 $\alpha$ 多肽的 $\alpha$ 1和 $\alpha$ 2结构域以及非人HLA II类 $\alpha$ 多肽的跨膜结构域和胞质结构域,和(ii)嵌合人/非人MHC II $\beta$ 多肽包含人HLA II类 $\beta$ 多肽的 $\beta$ 1和 $\beta$ 2结构域、非人HLA II类 $\beta$ 多肽的跨膜结构域和胞质结构域,以及(iii)人源化MHC I复合体包含人MHC I多肽的 $\alpha$ 1、 $\alpha$ 2和 $\alpha$ 3结构域,以及任选的人(源化) $\beta$ 2微球蛋白(例如,非人动物还包含编码多肽的 $\beta$ 2微球蛋白基因座,所述多肽包含人 $\beta$ 2微球蛋白氨基酸序列或其一部分)。

[0038] 在一些实施方案中,编码嵌合T细胞CD4辅助受体多肽的第一核苷酸序列存在于内源CD4T细胞辅助受体基因座,和/或编码嵌合T细胞CD8 $\alpha$ 辅助受体多肽的第二核苷酸序列存在于内源CD8 $\alpha$ T细胞辅助受体基因座,以及编码嵌合T细胞CD8 $\beta$ 辅助受体多肽的第三核苷酸序列存在于内源CD8 $\beta$ T细胞辅助受体基因座。另外的实施方案包括由图5A所示的基因编码的嵌合人/非人CD4多肽(例如,其中所得嵌合人/非人CD4T细胞辅助受体多肽的人部分包含至少人Ig1、人Ig2和人Ig3结构域,也分别称为D1、D2和D3结构域)和/或由图5B所示的基因编码的嵌合CD8辅助受体(例如,其中嵌合CD8辅助受体的人部分包含人CD8多肽的全部或基

本上全部细胞外部分(例如,CD8 $\alpha$ 和/或CD8 $\beta$ ),包括人免疫球蛋白V(IgV)样 $\alpha$ 和 $\beta$ 结构域。在一些实施方案中,嵌合CD4T细胞辅助受体多肽的人部分包含人CD4多肽的一个或多个细胞外结构域(例如,D1、D2、D3、D4或它们的任何组合)并且嵌合CD4T细胞辅助受体多肽的非人部分包含非人CD4T细胞辅助受体的跨膜结构域和胞质结构域,嵌合CD8 $\alpha$ 多肽的人部分包含人CD8 $\alpha$ 多肽的细胞外结构域(例如,IgV样结构域)并且嵌合CD8 $\alpha$ 多肽的非人部分包含非人CD8 $\alpha$ 多肽的跨膜结构域和胞质结构域,和/或CD8 $\beta$ 多肽的人部分包含人CD8 $\beta$ 多肽的细胞外结构域(例如,IgV样结构域)并且嵌合CD8 $\beta$ T细胞辅助受体多肽的非人部分包含非人CD8 $\beta$ 多肽的跨膜结构域和胞质结构域。

[0039] 在一些实施方案中,编码人(源化)MHC II $\alpha$ 的第一核酸序列存在于内源非人MHC II $\alpha$ 基因座,并且编码人(源化)MHC II $\beta$ 的第二核酸序列存在于内源非人MHC II $\beta$ 基因座,和/或编码人(源化)MHC I的第三核酸序列存在于内源非人MHC I基因座。在一个方面,人(源化)MHC II $\alpha$ 多肽包含人MHC II $\alpha$ 多肽(例如,HLA II类 $\alpha$ 多肽)的细胞外部分(或其一部分),人(源化)MHC II $\beta$ 多肽包含人MHC II $\beta$ 多肽(例如,HLA I类 $\beta$ 多肽)的细胞外部分(或其一部分),和/或人(源化)MHC I多肽包含人MHC I多肽(例如,HLA I类多肽)的细胞外部分(或其一部分)。在一些实施方案中,人源化MHC II $\alpha$ 多肽包含人MHC II $\alpha$ 1和 $\alpha$ 2结构域,人源化MHC II $\beta$ 多肽包含人MHC II $\beta$ 1和 $\beta$ 2结构域,和/或人源化MHC I多肽包含人MHC I $\alpha$ 1、 $\alpha$ 2和 $\alpha$ 3结构域。在一些实施方案中,编码嵌合人/非人MHC II $\alpha$ 多肽的第一核酸序列在内源非人MHC II $\alpha$ 启动子和调控元件的调控下表达,编码嵌合人/非人MHC II $\beta$ 多肽的第二核酸序列在内源非人MHC II $\beta$ 启动子和调控元件的调控下表达,和/或编码嵌合人/非人MHC I多肽的第三核酸序列在内源非人MHC I启动子和调控元件的调控下表达。在另外的实施方案中,嵌合人/非人MHC II $\alpha$ 多肽的非人部分包含内源非人MHC II $\alpha$ 多肽的跨膜结构域和胞质结构域,嵌合人/非人MHC II $\beta$ 多肽的非人部分包含内源非人MHC II $\beta$ 多肽的跨膜结构域和胞质结构域,和/或嵌合人/非人MHC I多肽的非人部分包含内源非人MHC I多肽的跨膜结构域和胞质结构域。实施方案包括非人动物,其中嵌合人/非人MHC II复合体的蛋白质的人部分来源于选自HLA-DR、HLA-DQ和HLA-DP的对应人HLA II类蛋白质,和/或其中第三嵌合人/非人MHC I多肽的人部分来源于人HLA-A、人HLA-B或人HLA-C。作为非限制性示例,在一些实施方案中,嵌合MHC II $\alpha$ 多肽包含HLA-DR $\alpha$ 蛋白质、HLA-DQ $\alpha$ 蛋白质或HLA-DP $\alpha$ 蛋白质的细胞外部分或其一部分,嵌合MHC II $\beta$ 多肽包含HLA-DR $\beta$ 蛋白质、HLA-DQ $\beta$ 蛋白质或HLA-DP $\beta$ 蛋白质的细胞外部分或其一部分,和/或嵌合MHC I多肽包含人HLA-A蛋白质、人HLA-B蛋白质或人HLA-C蛋白质的细胞外部分或其一部分。还提供了非人动物,其中嵌合人/非人MHC II蛋白质的人部分来源于对应人HLA-DR蛋白质,例如,人/非人MHC II $\alpha$ 多肽的人部分包含HLA-DR2的 $\alpha$ 链的 $\alpha$ 1和 $\alpha$ 2结构域,并且人/非人MHC II $\beta$ 多肽的人部分包含HLA-DR2的 $\beta$ 链的 $\beta$ 1和 $\beta$ 2结构域,和/或其中MHC I多肽的人部分来源于人HLA-A多肽,例如,人/非人MHC I多肽的人部分包含人HLA-A2多肽的 $\alpha$ 1、 $\alpha$ 2和 $\alpha$ 3结构域,例如人HLA-A2.1多肽的 $\alpha$ 1、 $\alpha$ 2和 $\alpha$ 3结构域。还提供了非人动物,其中MHC II复合体的非人部分来源于鼠H-2E编码序列和/或其中MHC I多肽的非人部分来源于鼠H-2K编码序列。例如,嵌合MHC II $\alpha$ 多肽包含鼠H-2E $\alpha$ 多肽的跨膜结构域和胞质结构域,嵌合MHC II $\beta$ 多肽包含鼠H-2E $\beta$ 多肽的跨膜结构域和胞质结构域,并且嵌合MHC I多肽包含鼠H-2K多肽的跨膜结构域和胞质结构域。

[0040] 在一些实施方案中,未重排TCR $\alpha$ 可变基因座存在于内源TCR $\alpha$ 可变基因座,并且未

重排TCR $\beta$ 可变基因座存在于内源TCR $\beta$ 可变基因座。在一些方面,未重排TCR $\alpha$ 可变基因座包含人未重排V $\alpha$ 基因区段的完整组库和人未重排J $\alpha$ 基因区段的完整组库,和/或未重排TCR $\beta$ 可变基因座包含人未重排V $\beta$ 基因区段的完整组库、人未重排D $\beta$ 基因区段的完整组库和人未重排J $\beta$ 基因区段的完整组库。在一些实施方案中,人未重排V $\alpha$ 和J $\alpha$ 基因区段重排形成重排的人V $\alpha$ /J $\alpha$ 序列,和/或人未重排V $\beta$ 、D $\beta$ 和J $\beta$ 基因区段重排形成重排的人V $\beta$ /D $\beta$ /J $\beta$ 序列。在一些实施方案中,本文所公开的非人动物在T细胞的表面上表达包含人TCR $\alpha$ 可变区和/或人TCR $\beta$ 可变区的T细胞受体。在一些实施方案中,内源非人V $\alpha$ 和J $\alpha$ 区段不能重排形成重排的V $\alpha$ /J $\alpha$ 序列,和/或内源非人V $\beta$ 、D $\beta$ 和J $\beta$ 区段不能重排形成重排的V $\beta$ /D $\beta$ /J $\beta$ 序列,例如,动物可能缺少功能性内源非人TCR $\alpha$ 可变基因座,和/或动物可能缺少功能性内源非人TCR $\beta$ 可变基因座,例如动物包含(a)全部或基本上全部功能性内源V $\alpha$ 基因区段的缺失,(b)全部或基本上全部功能性内源J $\alpha$ 基因区段的缺失,(c)全部或基本上全部功能性内源V $\beta$ 基因区段的缺失,(d)全部或基本上全部功能性内源D $\beta$ 基因区段的缺失,(e)全部或基本上全部功能性内源J $\beta$ 基因区段的缺失,和/或(f)它们的组合。在一些实施方案中,内源非人TCR $\alpha$ 可变基因座缺少全部或基本上全部功能性内源V $\alpha$ 基因区段和/或缺少全部或基本上全部功能内源J $\alpha$ 基因区段;和/或内源非人TCR $\beta$ 可变基因座(a)缺少全部或基本上全部功能性内源V $\beta$ 基因区段,(b)缺少全部或基本上全部功能性内源D $\beta$ 基因区段,(c)缺少全部或基本上全部功能性内源J $\beta$ 基因区段,或(d)(a)、(b)和(c)的任何组合。

[0041] 在一些实施方案中,分别编码嵌合T细胞CD4、CD8 $\alpha$ 和/或CD8 $\beta$ 辅助受体多肽的第一、第二和/或第三核苷酸序列存在于内源T细胞辅助受体基因座,例如分别存在于内源CD4、CD8 $\alpha$ 和/或CD8 $\beta$ 辅助受体基因座;未重排TCR $\alpha$ 可变基因座存在于内源TCR $\alpha$ 可变基因座;未重排TCR $\beta$ 可变基因座存在于内源TCR $\beta$ 可变基因座;和/或分别编码嵌合MHC II $\alpha$ 、MHC II $\beta$ 和/或MHC I多肽的第一、第二和/或第三核苷酸序列存在于内源MHC基因座;例如分别存在于MHC II $\alpha$ 、MHC II $\beta$ 和/或MHC I基因座。在一些实施方案中,编码嵌合T细胞辅助受体的核苷酸序列、未重排TCR $\alpha$ 可变基因座、未重排TCR $\beta$ 可变基因座和/或编码嵌合MHC分子的核苷酸序列可以可操作地连接至非人启动子和调控序列。例如,第一核苷酸序列可在内源非人CD4启动子和调控元件的调控下表达,第二核苷酸序列可在内源非人CD8 $\alpha$ 启动子和调控元件的调控下表达,和/或第三核苷酸序列可在内源非人CD8 $\beta$ 启动子和调控元件的调控下表达;未重排TCR $\alpha$ 可变基因座可在内源TCR $\alpha$ (可变)调控元件和启动子元件的调控下表达,并且未重排TCR $\beta$ 可变基因座可在内源TCR $\beta$ (可变)调控元件和启动子元件的调控下表达;第一核苷酸序列可在内源非人MHC II $\alpha$ 启动子和调控元件的调控下表达,第二核苷酸序列可在内源非人MHC II $\beta$ 启动子和调控元件的调控下表达,并且第三核苷酸序列可在内源非人MHC I启动子和调控元件的调控下表达。

[0042] 在一些实施方案中,编码人CD4多肽的细胞外部分(或其部分,例如D1、D2、D3和/或D4)的核苷酸序列替换编码内源非人(小鼠)CD4辅助受体多肽的细胞外部分(或其部分,例如D1、D2、D3和/或D4)的序列,并且可在内源非人(小鼠)CD4辅助受体基因座处可操作地连接至内源非人(小鼠)CD4跨膜结构域和胞质结构域编码序列;编码人CD8 $\alpha$ 多肽的细胞外部分的全部或一部分的核苷酸序列替换编码内源非人(小鼠)T细胞CD8 $\alpha$ 多肽的细胞外部分的全部或一部分的序列,并且可在内源非人(小鼠)CD8 $\alpha$ 基因座处可操作地连接至内源非人(小鼠)CD8 $\alpha$ 跨膜结构域和胞质结构域编码序列;编码人CD8 $\beta$ 多肽的细胞外结构域的全部或



一部分的核苷酸序列替换编码内源非人(小鼠)T细胞CD8 $\beta$ 多肽的细胞外结构域的全部或一部分的序列,并且可在内源CD8 $\beta$ 基因座处可操作地连接至内源非人CD8 $\beta$ 跨膜结构域和胞质结构域编码序列;未重排TCR $\alpha$ 可变基因座在内源非人(小鼠)TCR $\alpha$ 可变基因座处替换一个或多个内源V $\alpha$ 和/或J $\alpha$ 基因区段;未重排TCR $\beta$ 可变基因座在内源非人(小鼠)TCR $\beta$ 可变基因座处替换一个或多个内源V $\beta$ 、D $\beta$ 和/或J $\beta$ 基因区段;编码人MHC II $\alpha$ 多肽的细胞外部分(或其部分,例如 $\alpha$ 1和 $\alpha$ 2结构域)的核苷酸序列替换编码内源非人(小鼠)MHC II $\alpha$ 多肽的细胞外部分(或其部分,例如 $\alpha$ 1和 $\alpha$ 2结构域)的序列,并且可在内源非人(小鼠)MHC II $\alpha$ 基因座处可操作地连接至内源非人(小鼠)MHC II $\alpha$ 跨膜结构域和胞质结构域编码序列;编码人MHC II $\beta$ 多肽的细胞外部分(或其部分,例如 $\beta$ 1和 $\beta$ 2结构域)的核苷酸序列替换编码内源非人(小鼠)MHC II $\beta$ 多肽的细胞外部分(或其部分,例如 $\beta$ 1和 $\beta$ 2结构域)的序列,并且可在内源非人(小鼠)MHC II $\beta$ 基因座处可操作地连接至内源非人(小鼠)MHC II $\beta$ 跨膜结构域和胞质结构域编码序列;和/或编码人MHC I多肽的细胞外部分(或其部分,例如 $\alpha$ 1、 $\alpha$ 2和/或 $\alpha$ 3结构域)的核苷酸序列替换编码内源非人(小鼠)MHC I多肽的细胞外部分(或其部分,例如 $\alpha$ 1、 $\alpha$ 2和/或 $\alpha$ 3结构域)的序列,并且可在内源非人(小鼠)MHC I基因座处可操作地连接至内源非人(小鼠)MHC I跨膜结构域和胞质结构域编码序列。

[0043] 在一些实施方案中,本文所公开的经遗传修饰的非人动物不从其内源基因座表达功能性内源非人T细胞CD4辅助受体,不从其内源CD8基因座表达功能性内源非人T细胞CD8辅助受体,不从内源TCR $\alpha$ 可变基因座表达功能性TCR $\alpha$ 可变结构域,不从内源TCR $\beta$ 可变基因座表达功能性TCR $\beta$ 可变结构域,不从内源MHC II基因座表达内源MHC II复合体的细胞外结构域(例如,细胞表面上)和/或不从内源MHC I基因座表达内源MHC I多肽的细胞外结构域(例如,细胞表面上)。

[0044] 本文所公开的任何非人动物还可以包含编码包含人或人源化 $\beta$ 2微球蛋白氨基酸序列的多肽的 $\beta$ 2微球蛋白基因座,其中非人动物表达人或人源化 $\beta$ 2微球蛋白多肽。在一些实施方案中,非人动物不从内源非人 $\beta$ 2微球蛋白基因座表达功能性内源非人动物 $\beta$ 2微球蛋白多肽。在一些实施方案中, $\beta$ 2微球蛋白基因座可操作地连接至内源非人 $\beta$ 2微球蛋白调控元件。在一个实施方案中, $\beta$ 2微球蛋白基因座包含人 $\beta$ 2微球蛋白基因的外显子2、外显子3和外显子4(例如外显子2至外显子4)中所示的核苷酸序列,并且任选地, $\beta$ 2微球蛋白基因座还包含非人(例如啮齿动物) $\beta$ 2微球蛋白基因的外显子1中所示的核苷酸序列。

[0045] 本文所提供的非人动物可以是啮齿动物,例如小鼠或大鼠。

[0046] 本文还提供了小鼠,所述小鼠表达嵌合人/鼠T细胞CD4、CD8 $\alpha$ 和CD8 $\beta$ 辅助受体多肽,它们各自分别包含鼠CD4、CD8 $\alpha$ 和CD8 $\beta$ 跨膜结构域和胞质结构域;T细胞表面上的T细胞受体,其包含人TCR $\alpha$ 可变区和人TCR $\beta$ 可变区;嵌合人/鼠MHC II $\alpha$ 、MHC II $\beta$ 和MHC I多肽,它们各自分别包含人MHC II $\alpha$ 的细胞外结构域(例如,人HLA II类 $\alpha$ 1和 $\alpha$ 2结构域、MHC II $\beta$ 的细胞外结构域人HLA II类 $\beta$ 1和 $\beta$ 2结构域)和MHC I多肽的细胞外结构域(例如,人HLA I类 $\alpha$ 1、 $\alpha$ 2和 $\alpha$ 结构域);以及任选的人或人源化 $\beta$ 2微球蛋白多肽。在一个实施方案中,本文提供了非人动物,例如小鼠,其中第一核苷酸序列编码嵌合人/鼠HLA-DR/H-2E多肽的 $\alpha$ 链,第二核苷酸序列编码嵌合HLA-DR/H-2E多肽的 $\beta$ 链,并且第三核苷酸序列编码嵌合人/鼠HLA-A/H-2K多肽,并且其中小鼠表达HLA-A/H-2K和HLA-DR/H-2E蛋白质。

[0047] 本文还提供了包含基本上人源化的T细胞免疫系统的非人动物,例如其中基本上

人源化的T细胞免疫系统产生对抗抗原的基本上人源化的T细胞免疫应答。在一些实施方案中,基本上人源化的T细胞免疫应答包括表达人T细胞受体 (TCR) 可变结构域的活化的T细胞,这些可变结构域识别在人白细胞抗原 (HLA) 细胞外结构域的背景下递呈的抗原和/或在HLA细胞外结构域的背景下递呈抗原的抗原递呈细胞。在一些实施方案中,基本上人源化的T细胞免疫系统包含:(a) 非人T细胞,其表达包含与人HLA分子结合的人T细胞辅助受体结构域的T细胞辅助受体多肽和/或包含由至少一个人TCR可变区基因区段编码的TCR可变结构域的T细胞受体 (TCR);和(b) 非人抗原递呈细胞,其在人HLA的背景下递呈抗原并活化非人T细胞。

[0048] 还提供了制备和使用本文所公开的非人动物的方法。通常,本文所公开的制备经遗传修饰的非人动物的方法包括(a) 将编码嵌合人/非人T细胞辅助受体多肽(例如,嵌合CD4多肽)的第一核苷酸序列、和/或编码第二嵌合人/非人T细胞辅助受体多肽(例如,嵌合CD8 $\alpha$ 多肽)的第二核苷酸序列和编码第三嵌合人/非人T细胞辅助受体多肽(例如,CD8 $\beta$ 多肽)的第三核苷酸序列引入非人动物的基因组中,其中每种嵌合T细胞辅助受体多肽的非人部分均包含非人T细胞辅助受体的至少跨膜结构域和胞质结构域,并且其中每种嵌合多肽的人部分均包含人T细胞辅助受体的细胞外部分(或其一部分,例如一个或多个结构域);(b) 将包含至少一个人V $\alpha$ 区段和至少一个人J $\alpha$ 区段并可操作地连接至非人TCR $\alpha$ 恒定基因序列的未重排T细胞受体 (TCR)  $\alpha$ 可变基因座和/或包含至少一个人V $\beta$ 区段、至少一个人D $\beta$ 区段和至少一个人J $\beta$ 区段并可操作地连接至非人TCR $\beta$ 恒定基因序列的未重排TCR $\beta$ 可变基因座插入非人动物的基因组中;以及任选地(c) 将编码第一嵌合人/非人MHC多肽(例如,嵌合MHC II $\alpha$ 多肽)的第一核苷酸序列、编码第二嵌合人/非人MHC多肽(例如,嵌合MHC II $\beta$ 多肽)的第二核苷酸序列和/或编码第三嵌合人/非人MHC多肽(例如,嵌合MHC I多肽)的第三核苷酸序列放置于所述基因组中,和/或(d) 将编码人或人源化 $\beta$ 2微球蛋白多肽的 $\beta$ 2微球蛋白基因座添加到非人动物的基因组中。在一些实施方案中,第一核苷酸序列编码可操作地连接至非人CD4辅助受体的至少跨膜结构域和胞质结构域的人CD4的细胞外部分或其一部分,第二核苷酸序列编码人CD8 $\alpha$ 的细胞外部分或其一部分以及非人CD8 $\alpha$ 的至少跨膜结构域和胞质结构域,第三核苷酸序列编码人CD8 $\beta$ 的细胞外部分或其一部分以及非人CD8 $\beta$ 的至少跨膜结构域和胞质结构域,第一核苷酸序列编码人HLA II类 $\alpha$ 多肽的细胞外部分(或其一部分)以及非人MHC II $\alpha$ 多肽的至少跨膜结构域和胞质结构域,第二核苷酸序列编码人HLA II类 $\beta$ 多肽的细胞外部分(或其一部分)以及非人MHC II $\beta$ 多肽的至少跨膜结构域和胞质结构域,第三核苷酸序列编码人HLA I类多肽的细胞外部分(或其一部分)以及非人MHC I多肽的跨膜结构域和胞质结构域,并且 $\beta$ 2微球蛋白基因座包含人 $\beta$ 2微球蛋白基因的外显子2至4中所示的核苷酸序列,例如人 $\beta$ 2微球蛋白基因的外显子2、3和4中所示的核苷酸序列。

[0049] 制备非人动物的方法包括实施方案,其中(a) 将编码嵌合T细胞辅助受体多肽的第一、第二和/或第三核苷酸序列引入非人动物的基因组中包括在内源CD4基因座处用编码嵌合人/非人CD4多肽的核苷酸序列替换编码内源非人CD4多肽的核苷酸序列,和/或在内源CD8 $\alpha$ 基因座处用编码嵌合人/非人CD8 $\alpha$ 多肽的核苷酸序列替换编码内源非人CD8 $\alpha$ 多肽的核苷酸序列,以及在内源CD8 $\beta$ 基因座处用编码嵌合人/非人CD8 $\beta$ 多肽的核苷酸序列替换编码内源非人CD8 $\beta$ 多肽的核苷酸序列;(b) 将未重排TCR $\alpha$ 基因座和/或未重排TCR $\beta$ 基因座插入动物的基因组中包括用包含至少一个人V $\alpha$ 区段和至少一个人J $\alpha$ 区段的未重排人源化TCR $\alpha$ 可

变基因座替换内源非人TCR $\alpha$ 可变基因座以产生人源化TCR $\alpha$ 可变基因座,其中人源化TCR $\alpha$ 可变基因座可操作地连接至内源非人TCR $\alpha$ 恒定区,和/或用包含至少一个人VB区段、至少一个人DB区段和至少一个人JB区段的未重排人源化TCR $\beta$ 可变基因座替换内源非人TCR $\beta$ 可变基因座以产生人源化TCR $\beta$ 可变基因座,其中人源化TCR $\beta$ 可变基因座可操作地连接至内源非人TCR $\beta$ 恒定区;(c)将编码嵌合MHC多肽的第一、第二和/或第三核酸序列放置于非人动物的基因组中包括在内源非人MHC II基因座处用编码嵌合人/非人MHC II复合体的核苷酸序列替换编码非人MHC II复合体的核苷酸序列,以及在内源非人MHC I基因座处用编码嵌合人/非人MHC I多肽的核苷酸序列替换编码非人MHC I多肽的核苷酸序列,和/或(d)将编码人或人源化 $\beta$ 2微球蛋白多肽的 $\beta$ 2微球蛋白基因座添加到非人动物的基因组中包括在内源非人 $\beta$ 2微球蛋白基因座处用编码人或人源化 $\beta$ 2微球蛋白多肽的核苷酸序列替换编码非人 $\beta$ 2微球蛋白多肽的核苷酸序列。

[0050] 在一些实施方案中,(a)将第一、第二和/或第三核苷酸序列分别引入非人动物的基因组中包括(i)在内源CD4基因座处用与编码内源非人CD4跨膜结构域和胞质结构域的序列可操作地连接的编码人CD4多肽的细胞外部分(或其一部分)的核苷酸序列替换编码内源非人CD4多肽的细胞外部分(或其一部分)的核苷酸序列,(ii)在内源CD8 $\alpha$ 基因座处用与编码内源非人CD8 $\alpha$ 跨膜结构域和胞质结构域的序列可操作地连接的编码人CD8 $\alpha$ 多肽的细胞外部分(或其一部分)的核苷酸序列替换编码内源非人CD8 $\alpha$ 多肽的细胞外部分(或其一部分)的核苷酸序列,和/或(iii)在内源CD8 $\beta$ 基因座处用与编码内源非人CD8 $\beta$ 跨膜结构域和胞质结构域的序列可操作地连接的编码人CD8 $\beta$ 多肽的细胞外部分(或其一部分)的核苷酸序列替换编码内源非人CD8 $\beta$ 多肽的细胞外部分(或其一部分)的核苷酸序列;(b)将未重排TCR $\alpha$ 基因座和/或未重排TCR $\beta$ 基因座分别插入动物的基因组中包括(i)用包含至少一个人V $\alpha$ 区段和至少一个人J $\alpha$ 区段的未重排人源化TCR $\alpha$ 可变基因座替换内源非人TCR $\alpha$ 可变基因座以产生人源化TCR $\alpha$ 可变基因座,其中人源化TCR $\alpha$ 可变基因座可操作地连接至内源非人TCR $\alpha$ 恒定区,和/或(ii)用包含至少一个人VB区段、至少一个人DB区段和至少一个人JB区段的未重排人源化TCR $\beta$ 可变基因座替换内源非人TCR $\beta$ 可变基因座以产生人源化TCR $\beta$ 可变基因座,其中人源化TCR $\beta$ 可变基因座可操作地连接至内源非人TCR $\beta$ 恒定区;(c)将第一、第二和/或第三核酸序列分别放置于非人动物的基因组中包括(i)在内源非人MHC II $\alpha$ 基因座处用与编码内源非人MHC II $\alpha$ 跨膜结构域和胞质结构域的序列可操作地连接的编码人HLA II类 $\alpha$ 多肽的细胞外部分(或其一部分)的核苷酸序列替换编码非人MHC II $\alpha$ 多肽的细胞外部分(或其一部分)的核苷酸序列,(ii)在内源非人MHC II $\beta$ 基因座处用与编码内源非人MHC II $\beta$ 跨膜结构域和胞质结构域的序列可操作地连接的编码人HLA II类 $\beta$ 多肽的细胞外部分(或其一部分)的核苷酸序列替换编码非人MHC II $\beta$ 多肽的细胞外部分(或其一部分)的核苷酸序列,和/或(iii)在内源非人MHC I基因座处用与编码内源非人MHC I跨膜结构域和胞质结构域的序列可操作地连接的编码人HLA I类多肽的细胞外部分(或其一部分)的核苷酸序列替换编码非人MHC I多肽的细胞外部分(或其一部分)的核苷酸序列;和/或在内源 $\beta$ 2微球蛋白基因座处用包含人 $\beta$ 2微球蛋白基因的外显子2、3和4的核苷酸序列替换外显子2-外显子4中所示的核苷酸序列。

[0051] 在一个实施方案中,引入步骤包括在第一非人动物中在内源CD4基因座处用编码嵌合人/非人CD4多肽的核苷酸序列替换编码内源非人CD4多肽的核苷酸序列,在第二非人

动物中在内源CD8 $\alpha$ 基因座处用编码嵌合人/非人CD8 $\alpha$ 多肽的核苷酸序列替换编码内源非人CD8 $\alpha$ 多肽的核苷酸序列,并且在内源CD8 $\beta$ 基因座处用编码嵌合人/非人CD8 $\beta$ 多肽的核苷酸序列替换编码内源非人CD8 $\beta$ 多肽的核苷酸序列。在一些实施方案中,引入步骤包括在第一非人动物中在内源CD4基因座处用与编码内源非人CD4跨膜结构域和胞质结构域的序列可操作地连接的编码人CD4多肽的细胞外部分(或其一部分)的核苷酸序列替换编码内源非人CD4多肽的细胞外部分(或其一部分)的核苷酸序列,在第二非人动物中在内源CD8 $\alpha$ 基因座处用与编码内源非人CD8 $\alpha$ 跨膜结构域和胞质结构域的序列可操作地连接的编码人CD8 $\alpha$ 多肽的细胞外部分(或其一部分)的核苷酸序列替换编码内源非人CD8 $\alpha$ 多肽的细胞外部分(或其一部分)的核苷酸序列,以及在内源CD8 $\beta$ 基因座处用与编码内源非人CD8 $\beta$ 跨膜结构域和胞质结构域的序列可操作地连接的编码人CD8 $\beta$ 多肽的细胞外部分(或其一部分)的核苷酸序列替换编码内源非人CD8 $\beta$ 多肽的细胞外部分(或其一部分)的核苷酸序列。在一些实施方案中,这些替换步骤同时或以任何顺序执行。

[0052] 在一些实施方案中,插入步骤包括在第三非人动物中用包含至少一个人V $\alpha$ 区段和至少一个人J $\alpha$ 区段的未重排人源化TCR $\alpha$ 可变基因座替换内源非人TCR $\alpha$ 可变基因座以产生人源化TCR $\alpha$ 可变基因座,其中人源化TCR $\alpha$ 可变基因座可操作地连接至内源非人TCR $\alpha$ 恒定区;在第四非人动物中用包含至少一个人V $\beta$ 区段、至少一个人D $\beta$ 区段和至少一个人J $\beta$ 区段的未重排人源化TCR $\beta$ 可变基因座替换内源非人TCR $\beta$ 可变基因座以产生人源化TCR $\beta$ 可变基因座,其中人源化TCR $\beta$ 可变基因座可操作地连接至内源非人TCR $\beta$ 恒定区。在一些实施方案中,这些替换步骤同时或以任何顺序执行。

[0053] 在一些实施方案中,放置步骤包括不按特定顺序,在第五非人动物中在内源非人MHC II基因座处用编码嵌合人/非人MHC II复合体的一个或多个核苷酸序列替换编码非人MHC II复合体的一个或多个核苷酸序列;以及在第五非人动物中在内源非人MHC I基因座处用编码嵌合人/非人MHC I多肽的核苷酸序列替换编码非人MHC I多肽的核苷酸序列。在一些实施方案中,放置步骤包括在第五非人动物中在内源非人MHC II $\alpha$ 基因座处用与编码内源非人MHC II $\alpha$ 跨膜结构域和胞质结构域的序列可操作地连接的编码人MHC II $\alpha$ 多肽的细胞外部分(或其一部分)的核苷酸序列替换编码非人MHC II $\alpha$ 多肽的细胞外部分(或其一部分)的核苷酸序列,以及在内源非人MHC II $\beta$ 基因座处用与编码内源非人MHC II $\beta$ 跨膜结构域和胞质结构域的序列可操作地连接的编码人MHC II $\beta$ 多肽的细胞外部分(或其一部分)的核苷酸序列替换编码非人MHC II $\beta$ 多肽的细胞外部分(或其一部分)的核苷酸序列,以及在第五非人动物中在内源非人MHC I基因座处用与编码内源非人MHC I跨膜结构域和胞质结构域的序列可操作地连接的编码人MHC I多肽的细胞外部分(或其一部分)的核苷酸序列替换编码非人MHC I多肽的细胞外部分(或其一部分)的核苷酸序列。在一些实施方案中,这些替换步骤同时或以任何顺序执行。

[0054] 在一些实施方案中,添加步骤包括在第六非人动物中在内源非人 $\beta$ 2微球蛋白基因座处用编码人或人源化 $\beta$ 2微球蛋白多肽的核苷酸序列替换编码非人 $\beta$ 2微球蛋白多肽的核苷酸序列。在一些实施方案中,人或人源化 $\beta$ 2微球蛋白多肽由人 $\beta$ 2微球蛋白基因的外显子2、外显子3和外显子4中所示的核苷酸序列编码。

[0055] 本文所公开的方法包括实施方案,其中通过使包含如本文所述的遗传修饰中的一者或多者的非人动物与包含其余遗传修饰的相同物种的另一(或更多)非人动物交配来引

入编码嵌合T细胞辅助受体多肽的第一、第二和/或第三核苷酸序列；插入TCR $\alpha$ 基因座和/或未重排TCR $\beta$ 基因座；放置编码嵌合MHC多肽的第一、第二和/或第三核苷酸序列；和/或添加 $\beta$ 2微球蛋白基因座。非限制性实施方案包括以任何顺序繁育如上所述的第一、第二、第三、第四、第五和第六非人动物。

[0056] 本文所公开的方法可包括非人胚胎干 (ES) 细胞中的同源重组。本文所公开的方法可用于产生本文所公开的小鼠。表达嵌合人/非人CD4、CD8 $\alpha$ 和/或CD8 $\beta$ T细胞辅助受体多肽、人(源化)TCR $\alpha/\beta$ 蛋白质以及嵌合MHC II复合体和MHC I (含人或人源化 $\beta$ 2微球蛋白)的非人动物可通过以下方式产生：(a) 首先通过分别在单独ES细胞中同源重组来引入每个单独的人(源化)基因，并从这样的ES细胞产生每个单独的非人动物，随后以任何顺序繁育每个产生的非人动物，(b) 通过在单个ES细胞中连续同源重组来引入所有人(源化)基因，然后从这样的ES细胞产生非人动物，或(c) 在ES细胞中的一些基因座处连续同源重组和繁育的组合。如本文所公开的动物还可视情况通过初始繁育的后代与其他动物交配来产生。繁育和/或同源重组可以以任何优选的顺序完成。

[0057] 还提供了从非人动物分离抗原特异性的人TCR可变结构域的方法，所述方法包括从本文提供的非人动物分离或根据本文所公开的方法制备与抗原结合的T细胞或TCR蛋白质。在一些实施方案中，所述方法还可包括鉴定编码与抗原结合的TCR $\alpha$ 和/或TCR $\beta$ 可变结构域的第一和/或第二核苷酸，和/或在载体表达的充分条件下培养包含一种或多种载体的细胞，其中所述载体包含分别与第一和/或第二核苷酸相同或基本上相同的第三和/或第四核苷酸，并且其中第三和/或第四核苷酸分别与例如人TCR恒定区基因(例如，TCR $\alpha$ 恒定区基因和/或TCR $\beta$ 恒定区基因)进行框内克隆。还提供了包含本文所公开的遗传修饰(其可包括重排人TCR $\alpha$ 和/或TCR $\beta$ 可变区基因)的组织和细胞，以及编码由分离自本文所述经修饰的非人动物的此类组织或细胞表达的此类人TCR可变结构域的核酸。还包括(1) 重组核酸，例如表达载体，其包含分别与适当人TCR恒定区基因例如TCR $\alpha$ 恒定区基因或TCR $\beta$ 恒定区基因框内克隆的、编码本文所公开的人TCR可变结构域的核苷酸序列例如人重排TCR $\alpha$ 或人重排TCR $\beta$ 可变区基因，(2) 包含此类核酸(例如，表达载体)的宿主细胞以及(3) 由宿主细胞表达的TCR。在一些实施方案中，本文所提供的重组核酸包含人重排TCR $\delta$ 可变区基因或TCR $\gamma$ 可变区基因，其例如来源于本文所公开的经遗传修饰的非人动物或从其分离的组织，分别与人TCR $\delta$ 恒定区基因或TCR $\gamma$ 恒定区基因框内克隆。

[0058] 还提供了在非人动物中产生人源化T细胞应答的方法，所述方法通常包括用抗原例如人抗原如人肿瘤抗原、人细菌病原体、人病毒病原体等对非人动物、经遗传修饰的或具有本文所述基本上人源化的T细胞免疫系统的非人动物进行免疫。在一些实施方案中，经免疫的非人动物表达全部功能性人TCRV $\alpha$ 基因区段的至少50%和/或全部功能性人TCRV $\beta$ 基因区段的至少50%，和/或包含全部或基本上全部功能性人TCRV $\alpha$ 基因区段和/或全部或基本上全部功能性人TCRV $\beta$ 基因区段。

[0059] 还提供了分离抗原特异性的人TCR的体外方法，所述方法通常包括检测非人动物的第一细胞在(a) 与非人动物的第二细胞接触及(b) 与抗原一起温育之后的活化；其中第一细胞表达嵌合人/非人T细胞辅助受体以及以下一者或两者：(i) 嵌合人/非人TCR $\alpha$ 链和(ii) 嵌合人/非人TCR $\beta$ 链，并且其中第二细胞表达嵌合人/非人MHC多肽。所述方法还可包括从第一细胞分离TCR或编码其的核酸。

[0060] 在本文所公开的体外方法中,抗原可以是肿瘤抗原、病毒抗原、自身抗原或细菌抗原。在一些实施方案中,非人动物是啮齿动物,例如大鼠或小鼠。本文还提供了分离自经遗传修饰的或具有本文所述基本上人源化的T细胞免疫系统的非人动物的组织、T细胞、TCR(例如,可溶性TCR)或编码全部或部分TCR的核酸、来源于这种T细胞的杂交瘤或四源杂交瘤。

[0061] 还提供了组合物,例如其包含非人动物的第一和第二细胞;其中第一细胞表达嵌合人/非人T细胞辅助受体,以及任选地以下一者或两者:(i)嵌合人/非人TCR $\alpha$ 链和(ii)嵌合人/非人TCR $\beta$ 链,并且其中第二细胞表达与嵌合人/非人T细胞辅助受体结合的嵌合人/非人MHC多肽。在一些实施方案中,第一细胞是非人T细胞。在其他实施方案中,第二细胞是非人抗原递呈细胞。

## 附图说明

[0062] 图1是包含人源化TCR $\alpha$ 和 $\beta$ 蛋白质、与人源化 $\beta$ 2微球蛋白复合的人源化MHC I类和人源化CD8异源二聚体的人源化T细胞受体复合体(左图);以及包含人源化TCR $\alpha$ 和 $\beta$ 蛋白质、人源化MHC II类异源二聚体和人源化CD4的T细胞受体复合体(右图)的示意图(未按比例)。由人源化MHC递呈的抗原被描绘为圆圈。小鼠区域被描绘为填充形状,而人区域被描绘为条纹形状。

[0063] 图2A-C提供了示例性嵌合MHC I和MHC II基因座的示意图(未按比例),例如嵌合HLA-A2/H-2K基因座(图2A)、嵌合HLA-DR2/H-2E基因座(图2B)和人源化 $\beta$ 2M基因座(图2C)。除非另外指明,否则人序列被描绘为空心形状,而小鼠序列被描绘为填充形状。条纹形状表示来源于与内源基因座不同的小鼠品系的H-2E的外显子1(参见实施例1.3和图3B)。用相应标记的箭头描绘Floxed新霉素磷酸转移酶盒。

[0064] 图3A-C描绘了用于产生包含人源化MHC I和MHC II基因的人源化MHC基因座的策略。在图3A中所描绘的特定实施方案中,所产生的小鼠的MHC基因座包含嵌合HLA-A2/H-2K和HLA-DR2/H-2E序列(H2-K<sup>+/1666</sup>MHC-II<sup>+/6112</sup>)并缺少H2-D序列(H2-D<sup>+/缺失</sup>)和H-2A序列(基因工程方案也会导致H-2A缺失,参见实施例1.2)。箭头右侧描绘了在人源化的每个阶段引入到ES细胞中的大靶向载体(LTVEC)或Cre重组酶构建体。MAID或4位数字表示修改后的等位基因ID号。图3B是示例性HLA-DR2/H-2E大靶向载体的示意图(未按比例)。除非另外指明,否则人序列被描绘为空心形状,而小鼠序列被描绘为填充形状。条纹形状表示来源于与内源基因座不同的小鼠品系的H-2E的外显子1(参见实施例1.3)。floxed潮霉素盒被描绘为相应标记的箭头。图3C是嵌合人/小鼠MHC基因座的示例性基因型的示意图(未按比例)(\*\*表示不存在于所有小鼠品系中例如不存在于C57BL/6或129小鼠品系中的H-2L基因),其中内源小鼠H-2K和H-2E基因座分别被嵌合人/小鼠HLA-A2/H-2K和HLA-DR2/H-2E基因座(条纹形状)替换,H-2A和H-2D基因座缺失(用虚线勾画出的空心形状),并且其余基因座是内源小鼠基因(用实线勾画出的实心形状)。

[0065] 图4A(未按比例)描绘了用于小鼠TCR $\alpha$ 基因座人源化的渐进式策略,其中TCR $\alpha$ 可变区基因区段依次在所缺失的小鼠基因座(MAID1540)的初始人源化的上游添加。小鼠序列由填充形状指示;人序列由空心形状指示。MAID是指修改后的等位基因ID号。TRAV=TCR V $\alpha$ 区段,TRAJ=TCR J $\alpha$ 区段(hTRAJ=人TRAJ),TRAC=TCR C $\alpha$ 结构域,TCRD=TCR $\delta$ 。图4B(未按比

例) 描绘了用于小鼠TCRB基因座人源化的渐进式策略, 其中TCRB可变区基因区段依次添加到缺失的小鼠TCRB可变基因座。小鼠序列由填充形状指示; 人序列由空心形状指示。MAID是指修改后的等位基因ID号。TRBV或TCRBV=TCRBV区段。

[0066] 图5A描绘了嵌合CD4基因座的示意图(未按比例)。人编码外显子由条纹形状呈现, 小鼠编码外显子由填充形状呈现, 并且非编码外显子由空心形状呈现。指示了免疫球蛋白样结构域(Ig)、跨膜(TM)、胞质(CYT)和信号肽(信号)编码外显子以及3'非翻译区(UTR)。用相应标记的箭头描绘floxed(loxP)新霉素磷酸转移酶(Pgk-neo)盒。图5B描绘了嵌合CD8a和CD8b基因座的示意图(未按比例)。人编码外显子由条纹形状呈现, 小鼠编码外显子由填充形状呈现, 并且非编码外显子由空心形状呈现。指示了免疫球蛋白样结构域(IgV)、跨膜(TM)、胞质(CYT)和信号肽(信号)编码外显子以及3'非翻译区(UTR)。用相应标记的箭头描绘Floxed(loxP)潮霉素(Hyg)和新霉素磷酸转移酶(Pgk-neo)盒。

[0067] 图6A-C是从对照小鼠或包含人源化MHC I、MHC II $\alpha$ 和 $\beta$ 、TCR $\alpha$ 和 $\beta$ 、CD4、CD8 $\alpha$ 和 $\beta$ 以及 $\beta$ 2M(TM I/II B C4/8)基因座的小鼠分离、对单一细胞设门并且用(图6A)抗小鼠CD19和抗小鼠CD3抗体、(图6B)抗小鼠CD19和抗小鼠F4/80抗体、或(图6C)抗小鼠CD8 $\alpha$ 和抗小鼠CD4抗体(左图)或抗人CD8 $\alpha$ 和抗人CD4抗体(右图)染色的胸腺细胞的FACS等高线图。

[0068] 图7A-G是从对照小鼠或包含人源化MHC I、MHC II $\alpha$ 和 $\beta$ 、TCR $\alpha$ 和 $\beta$ 、CD4、CD8 $\alpha$ 和 $\beta$ 以及 $\beta$ 2M(TM I/II B C4/8)基因座的小鼠分离、对CD19<sup>+</sup>细胞、F4/80<sup>+</sup>细胞或CD3<sup>+</sup>细胞设门并且用(图7A、图7B)抗人 $\beta$ 2M或抗小鼠H-2D抗体; (图7C、图7D)抗HLA-A2或抗HLA-DR抗体; (图7E、图7F)抗H-2D和抗I<sup>A</sup>E抗体; 或(图7G)抗小鼠CD4和抗人CD4抗体(顶部)、抗小鼠CD8 $\alpha$ 和抗人CD8 $\alpha$ 抗体(中部)以及抗小鼠CD8 $\beta$ 和抗人CD8 $\beta$ 抗体(底部)染色的胸腺细胞的FACS等高线图。

[0069] 图8提供了从对照小鼠或包含人源化MHC I、MHC II $\alpha$ 和 $\beta$ 、TCR $\alpha$ 和 $\beta$ 、CD4、CD8 $\alpha$ 和 $\beta$ 以及 $\beta$ 2M(TM I/II B C4/8)的小鼠分离、对CD3<sup>+</sup>CD4<sup>+</sup>细胞设门并且用抗小鼠FoxP3和抗小鼠CD25抗体染色的胸腺细胞的FACS等高线图。

[0070] 图9A-E是从对照小鼠或包含人源化MHC I、MHC II $\alpha$ 和 $\beta$ 、TCR $\alpha$ 和 $\beta$ 、CD4、CD8 $\alpha$ 和 $\beta$ 以及 $\beta$ 2M(TM I/II B C4/8)基因座的小鼠分离、对单一细胞、CD3<sup>+</sup>细胞、CD4<sup>+</sup>T细胞或CD8<sup>+</sup>T细胞设门并且用(图9A)抗小鼠CD19和抗小鼠CD3、(图9B)抗小鼠CD19和抗小鼠F4/80抗体、(图9C)或抗小鼠CD4和抗小鼠CD8 $\alpha$ 抗体(左)或抗人CD4和抗人CD8 $\alpha$ 抗体(右)或(图9D、图9E)抗小鼠CD44和抗小鼠CD62L抗体染色的脾细胞的FACS等高线图。

[0071] 图10A-G是从对照小鼠或包含人源化MHC I、MHC II $\alpha$ 和 $\beta$ 、TCR $\alpha$ 和 $\beta$ 、CD4、CD8 $\alpha$ 和 $\beta$ 以及 $\beta$ 2M(TM I/II B C4/8)基因座的小鼠分离、对CD19<sup>+</sup>细胞、F4/80<sup>+</sup>细胞或CD3<sup>+</sup>细胞设门并且用(图10A、图10B)抗人 $\beta$ 2M或抗小鼠H-2D抗体、(图10C、图10D)抗HLA-A2或抗HLA-DR抗体、(图10E、图10F)抗H-2D和抗I<sup>A</sup>E抗体、或(图10G)抗小鼠CD4和抗人CD4抗体(顶部)、抗小鼠CD8 $\alpha$ 和抗人CD8 $\alpha$ 抗体(中部)以及抗小鼠CD8 $\beta$ 和抗人CD8 $\beta$ 抗体(底部)染色的脾细胞的FACS等高线图。

[0072] 图11提供了从对照小鼠或包含人源化MHC I、MHC II $\alpha$ 和 $\beta$ 、TCR $\alpha$ 和 $\beta$ 、CD4、CD8 $\alpha$ 和 $\beta$ 以及 $\beta$ 2M(TM I/II B C4/8)的小鼠分离、对CD3<sup>+</sup>CD4<sup>+</sup>细胞设门并且用抗小鼠FoxP3和抗小鼠CD25抗体染色的脾细胞的FACS等高线图。

[0073] 图12提供了在从对照小鼠或包含人源化MHC I、MHC II $\alpha$ 和 $\beta$ 、TCR $\alpha$ 和 $\beta$ 、CD4、CD8 $\alpha$ 和



$\beta$ 以及 $\beta 2M$ (TM I/II B C4/8)基因座的小鼠分离并且在不存在肽(仅200k细胞;x轴)或存在10 $\mu$ g/ml或1 $\mu$ g/ml MAG-E-A3肽(x轴)的情况下温育之后,在酶联免疫吸附斑点测定法中得出的产生IFN- $\gamma$ 的脾细胞数(每孔斑点数(平均值+SD);y轴)。

[0074] 图13A描绘了在对照或包含人源化MHC I、MHC II $\alpha$ 和 $\beta$ 、TCR $\alpha$ 和 $\beta$ 、CD4、CD8 $\alpha$ 和 $\beta$ 以及 $\beta 2M$ (TM I/II B C4/8)基因座的小鼠中急性Armstrong株病毒感染的进展;在该图的顶部描绘了实验的时间线,并且在底图中描绘了这两种小鼠品系感染后不同天数的病毒滴度的测量值。图13B描绘了在对照或包含人源化MHC I、MHC II $\alpha$ 和 $\beta$ 、TCR $\alpha$ 和 $\beta$ 、CD4、CD8 $\alpha$ 和 $\beta$ 以及 $\beta 2M$ (TM I/II B C4/8)基因座的小鼠中慢性克隆13株病毒感染的进展;在该图的顶部描绘了实验的时间线,并且在底图中描绘了这两种小鼠品系感染后第21天的病毒滴度的测量值。来自未感染或慢性感染的TM I/II B C4/8或对照B6小鼠的T细胞用抗PD1、抗Lag3和抗Tim3抗体染色(图13C;x轴);该图提供了染色阳性细胞的定量(阳性细胞%;y轴)。

[0075] 图14描绘了在事先急性Armstrong株感染之后对照或TM I/II B C4/8小鼠中慢性克隆13株病毒感染的进展;在该图的顶部描绘了实验的时间线,并且在底图中描绘了感染后第31天的病毒滴度的测量值。模拟感染的小鼠作为附加对照包括在实验中。

[0076] 图15A-B描绘了CD8<sup>+</sup>细胞数(y轴;IFN- $\gamma$ 阳性细胞),这些细胞对HLA-A2限制性(GPC10-18;N69-77;Z49-58)、H2D<sup>b</sup>限制性(GP33-41)的LCMV肽、卵清蛋白或单独温育作出应答而产生IFN- $\gamma$ ,并且从对照动物(图15A)或包含人源化MHC I、MHC II $\alpha$ 和 $\beta$ 、TCR $\alpha$ 和 $\beta$ 、CD4、CD8 $\alpha$ 和 $\beta$ 以及 $\beta 2M$ (TM I/II B C4/8)基因座的小鼠(图15B)分离,每只动物接受模拟感染(模拟;每组n=1)或急性Armstrong株感染(Arm;每组n=3)。在包含人源化MHC I、MHC II $\alpha$ 和 $\beta$ 、TCR $\alpha$ 和 $\beta$ 、CD4、CD8 $\alpha$ 和 $\beta$ 以及 $\beta 2M$ (TM I/II B C4/8)基因座的小鼠或对照B6动物中,在感染的时间进程(感染后天数;x轴)期间用所指示的肽(OVA、GP33、NP69、GPC10、GPC447或Z49)刺激之后IFN- $\gamma$ +CD8<sup>+</sup>淋巴细胞的%(y轴)分别示于图15C和图15D中。

## 具体实施方式

[0077] 本文公开了经遗传工程改造成表达人源化T细胞辅助受体(例如,人源化CD4和/或CD8(例如,CD8 $\alpha$ 和/或CD8 $\beta$ ))、结合人源化T细胞辅助受体的人或人源化主要组织相容性复合体(MHC)(例如,人或人源化MHC II(例如,MHC II $\alpha$ 和/或MHC II $\beta$ 链)和/或MHC I(例如,MHC I $\alpha$ )以及任选的人或人源化 $\beta 2$ 微球蛋白)和/或人或人源化T细胞受体(TCR)的非人动物(例如,啮齿动物,例如小鼠或大鼠),以及表达这些物质的胚胎、组织和细胞。本文所公开的非人动物的免疫系统的细胞免疫的发展与对照动物相当,例如,胸腺和脾包含相似的胸腺细胞和CD3<sup>+</sup>细胞绝对数量。这与经修饰以包含人TCR( $\alpha$ 和 $\beta$ )和嵌合人/小鼠MHC I分子两者的其他非人动物大相径庭,参见例如Li(2010)Nature Medicine 16:1029-1035(Li,2010年,《自然医学》,第16卷,第1029-1035页)和补充材料。不仅与野生型对照动物相比,而且与仅用人TCR修饰的动物和仅用嵌合人/小鼠MHC I分子修饰的动物相比,这样的动物都显示出T细胞群的减少,同上。因此,本文提供了非人动物,所述非人动物经工程改造成共表达人源化CD4辅助受体和人源化MHC II,和/或人源化CD8辅助受体和人源化MHC I,以及任选的人源化TCR。还提供了用于制备经遗传工程改造的动物的方法,所述经遗传工程改造的动物表达至少一种人源化T细胞辅助受体(例如,人源化CD4和/或CD8)、与人源化T细胞辅助受体结合的至少一种人源化MHC(例如,分别与人源化CD4和/或CD8结合的人源化MHC II和/或



MHC I) 和/或人源化TCR。还提供了使用经遗传工程改造的动物开发人类治疗剂的方法,所述经遗传工程改造的动物产生基本上人源化的T细胞免疫应答。

**[0078] 基本人源化的T细胞免疫应答**

**[0079]** 本文公开了经遗传修饰以产生基本上人源化的T细胞免疫应答的非人动物。本文所公开的小鼠表达至少一种人或人源化T细胞辅助受体,能够与所述至少一种人或人源化T细胞辅助受体结合的至少一种人或人源化主要组织相容性复合体(MHC),和/或优选能够识别在与人或人源化T细胞辅助受体结合的人或人源化MHC的背景下递呈的抗原的人或人源化T细胞受体(TCR),并且向表达人或人源化TCR的非人细胞例如非人T细胞提供活化信号。人或人源化T细胞辅助受体、人或人源化TCR和/或人或人源化MHC可以由非人动物的基因组编码。在优选的实施方案中,当用抗原免疫时,非人动物将抗原的HLA限制性表位递呈到来源于人TCR基因区段的TCR,例如人TCR $\alpha$ V区段、人TCR $\alpha$ J区段、人TCR $\beta$ V区段、人TCR $\beta$ D区段和/或人TCR $\beta$ J区段。

**[0080]** 因此,本发明涵盖经遗传修饰的非人动物,其基因组包含(例如,在内源基因座处)编码人源化T细胞辅助受体多肽(例如,CD4或CD8多肽)的核苷酸序列,其中嵌合T细胞辅助受体多肽包含本文所述的氨基酸序列的保守氨基酸取代,和/或编码与人源化T细胞辅助受体多肽结合的人源化MHC多肽的核酸序列,其中人源化MHC多肽包含本文所述的氨基酸序列的保守氨基酸取代。

**[0081]** 保守氨基酸取代包括用具有的侧链R基团的化学性质(例如电荷或疏水性)相似的另一个氨基酸残基取代某个氨基酸残基。保守氨基酸取代可通过修饰核苷酸序列以引入将编码保守取代的核苷酸变化来实现。通常,保守氨基酸取代基本上不会改变蛋白质的目标功能性质,例如,CD4或CD8分别与MHC II或MHC I结合(例如结合)的能力,并且可例如提高TCR对MHC递呈的抗原的敏感性。具有化学性质相似的侧链的氨基酸组的示例包括脂肪族侧链,诸如甘氨酸、丙氨酸、缬氨酸、亮氨酸和异亮氨酸;脂族羟基侧链,诸如丝氨酸和苏氨酸;含酰胺侧链,诸如天冬酰胺和谷氨酰胺;芳香族侧链,诸如苯丙氨酸、酪氨酸和色氨酸;碱性侧链,诸如赖氨酸、精氨酸和组氨酸;酸性侧链,诸如天冬氨酸和谷氨酸;以及含硫侧链,诸如半胱氨酸和甲硫氨酸。保守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。在一些实施方案中,保守氨基酸取代可以用丙氨酸取代蛋白质中的任何天然残基,例如在丙氨酸扫描诱变中所用。在一些实施方案中,所进行的保守取代在Gonnet等人((1992) Exhaustive Matching of the Entire Protein Sequence Database, Science 256:1443-45 (1992年,《整个蛋白质序列数据库的穷尽匹配》))公开的PAM250对数似然矩阵中具有正值,该文献据此以引用方式并入。在一些实施方案中,所述取代是适度保守的取代,其中所述取代在PAM250对数似然矩阵中具有非负值。

**[0082]** 本领域技术人员应当理解,除了本文所述编码人源化T细胞辅助受体多肽、人源化MHC多肽和/或TCR可变区的核酸残基之外,由于遗传密码的简并性,其他核酸也可以编码本发明的多肽。因此,除了在其基因组中包含编码人源化T细胞辅助受体多肽(例如,CD4或CD8多肽)的核苷酸序列、包含人未重排基因区段的未重排T细胞受体可变基因座(例如,TCR $\alpha$ 和/或TCR $\beta$ )、和/或编码能够与具有保守氨基酸取代的人源化T细胞辅助受体多肽结合的人源化MHC多肽的核酸序列的经遗传修饰的非人动物之外,还提供了这样的非人动物,其基因

组包含编码人源化T细胞辅助受体多肽(例如,CD4或CD8多肽)的核苷酸序列、包含人未重排基因区段的未重排T细胞受体可变基因座(例如,TCR $\alpha$ 和/或TCR $\beta$ )、和/或由于遗传密码的简并性而与本文所述有所不同的编码能够与人源化T细胞辅助受体多肽结合的人源化MHC多肽的核酸序列。

[0083] 序列的同一性可由本领域已知的可用于测量核苷酸和/或氨基酸序列同一性的多种不同算法确定。在本文所述的一些实施方案中,同一性使用ClustalW v.1.83(慢)比对法进行确定,该比对法采用10.0的开放空位罚分、0.1的扩展空位罚分,并使用Gonnet相似矩阵(MacVector公司的MacVector™10.0.2,2008年(MacVector™10.0.2,MacVector Inc., 2008))。针对序列同一性进行比较的序列长度取决于具体序列。在各种实施方案中,同一性是通过将成熟蛋白从N端到C端的序列进行比较来确定的。在各种实施方案中,当将嵌合人/非人序列与人序列进行比较时,使用嵌合人/非人序列的人部分(而不是非人部分)进行比较以确定人序列和嵌合人/非人序列的人部分之间的同一性水平(例如,将嵌合人/小鼠蛋白质的人胞外域与人蛋白质的人胞外域进行比较)。

[0084] 指涉序列例如核苷酸或氨基酸序列的术语“同源性”或“同源”是指在最佳比对和比较时例如至少约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%核苷酸或氨基酸、例如大于97%核苷酸或氨基酸相同的两个序列。本领域技术人员应当理解,为了实现最佳基因靶向,靶向构建体应包含与内源DNA序列同源的臂(即“同源臂”);因此,靶向构建体和所靶向内源序列之间可能发生同源重组。

[0085] 术语“可操作地连接”是指并列,其中如此描述的组分处于允许它们以其预期方式起作用的关系。因此,编码蛋白的核酸序列可以可操作地连接至调控序列(例如,启动子、增强子、沉默子序列等)以保持适当的转录调控。此外,本发明的嵌合或人源化蛋白质的各个部分可以可操作地连接以保持该蛋白在细胞中正确的折叠、加工、靶向、表达和其他功能性质。除非另有说明,否则本发明的嵌合或人源化蛋白质的各个结构域彼此可操作地连接。

[0086] 指涉及基因替换的术语“替换”是指将外源遗传物质放置于内源遗传基因座,从而用直系同源或同源核酸序列替换全部或一部分内源基因。如以下实施例中所证实,在一个实施方案中,编码部分小鼠CD4或CD8(CD8 $\alpha$ 和/或CD8 $\beta$ )多肽的内源基因座的核酸序列分别被替换为编码部分人CD4或CD8(CD8 $\alpha$ 和/或CD8 $\beta$ )多肽的核苷酸序列。

[0087] 如本文所用,例如指涉功能性多肽的“功能性”是指保留正常情况下与天然蛋白质相关的至少一种生物学活性的多肽。例如,在本发明的一些实施方案中,内源基因座处的替换(例如,内源非人CD4或CD8基因座处的替换)导致基因座不能表达功能性内源多肽。

#### [0088] 人源化T细胞辅助受体

[0089] 本文公开了表达至少一种人或人源化T细胞辅助受体例如CD4、CD8 $\alpha$ 和/或CD8 $\beta$ 的非人动物。因此,本文所公开的非人动物包含第一、第二和/或第三核苷酸序列中的至少一者,每个核苷酸序列编码选自人或人源化CD4多肽、人或人源化CD8 $\alpha$ 多肽和人或人源化CD8 $\beta$ 多肽的不同人或嵌合人/非人T细胞辅助受体多肽。本文中使用第一、第二、第三名称不应被解释为将本文所公开的非人动物限制为需要所有三个核苷酸序列或以任何顺序存在任何辅助受体核苷酸序列。因此,本文所公开的非人动物可包含编码人或人源化CD4和/或人或人源化CD8(例如,人或人源化CD8 $\alpha$ 和/或CD8 $\beta$ )多肽的一个或多个核酸序列。

[0090] 在一个实施方案中,本文所公开的非人动物包含编码人或人源化CD4多肽的第一

核苷酸序列。在另一个实施方案中,本文所公开的非人动物包含编码人或人源化CD8 $\alpha$ 多肽的第一核苷酸序列和编码人或人源化CD8 $\beta$ 多肽的第二核苷酸序列。在另一个实施方案中,本文所公开的非人动物包含编码人或人源化CD8 $\alpha$ 和CD8 $\beta$ 多肽的第一和第二核苷酸序列,并且还包含编码人或人源化CD4多肽的第三核苷酸序列。

#### [0091] 人或人源化CD4

[0092] 在各种实施方案中,本发明通常提供在其基因组中例如在内源CD4基因座处包含编码人或人源化CD4多肽的核苷酸序列的经遗传修饰的非人动物;因此,所述动物表达人或人源化CD4多肽。

[0093] 人CD4基因定位于12号染色体,且被认为含有10个外显子。CD4基因编码具有氨基末端疏水信号序列的蛋白质,该氨基末端疏水信号序列由该基因的外显子2和3编码。该蛋白质包含四个细胞外免疫球蛋白样结构域Ig1-Ig4,通常也分别称为D1-D4结构域。Maddon et al. (1987) Structure and expression of the human and mouse T4 genes, Proc. Natl. Acad. Sci. USA 84:9155-59 (Maddon等人,1987年,人和小鼠T4基因的结构和表达,《美国国家科学院院刊》,第84卷,第9155-9159页)。D1结构域被认为是由外显子3(信号肽下游的序列)和外显子4编码,而D2、D3和D4分别由外显子5、6和7各单独外显子编码(参见图5A:D1、D2、D3和D4结构域分别由指定为Ig1、Ig2、Ig3和Ig4的序列编码)。Littman (1987) The Structure of the CD4 and CD8 Genes, Ann. Rev. Immunol. 5:561-84 (Littman,1987年,CD4和CD8基因的结构,《免疫学年评》,第5卷,第561-584页); Hanna et al. (1994) Specific Expression of the Human CD4 Gene in Mature CD4+CD8- and Immature CD4+CD8+ T cells and in Macrophages of Transgenic Mice, Mol. Cell. Biol. 14 (2):1084-94 (Hanna等人,1994年,人CD4基因在转基因小鼠的成熟CD4+CD8-和未成熟CD4+CD8+T细胞中及在巨噬细胞中的特异性表达,《分子与细胞生物学》,第14卷,第2期,第1084-1094页); Maddon等人,出处同上。在高蛋白质浓度的区域,诸如T细胞和抗原递呈细胞之间的接触区域处,分子趋向于通过相对D4结构域之间的相互作用而发生同型二聚。Zamoyska (1998) CD4 and CD8: modulators of T cell receptor recognition of antigen and of immune responses? Curr. Opin. Immunol. 10:82-87 (Zamoyska,1998年,CD4和CD8:T细胞受体识别抗原和免疫应答的调节剂?,《免疫学新见》,第10卷,第82-87页); Wu et al. (1997) Dimeric association and segmental variability in the structure of human CD4, Nature 387:527 (Wu等人,1997年,人CD4的二聚缔合和结构的节段性变化,《自然》,第387卷,第527页); Moldovan et al. (2002) CD4 Dimers Constitute the Functional Component Required for T Cell Activation, J. Immunol. 169:6261-68 (Moldovan等人,2002年,CD4二聚体构成T细胞活化所需的功能组分,《免疫学杂志》,第169卷,第6261-6268页)。

[0094] CD4的D1结构域类似于免疫球蛋白可变(V)结构域,并且被认为与D2结构域的一部分一起例如在MHC II辅助受体结合位点处结合MHC II(与MHC II结合)。Huang et al. (1997) Analysis of the contact sites on the CD4 molecule with Class II MHC molecule, J. Immunol. 158:216-25 (Huang等人,1997年,CD4分子上与II类MHC分子接触的位点的分析,《免疫学杂志》,第158卷,第216-225页)。继而,MHC II在MHC II $\alpha$ 2和 $\beta$ 2结构域之间的连接处的疏水缝隙处与T细胞辅助受体CD4相互作用。Wang and Reinherz (2002) Structural Basis of T Cell Recognition of Peptides Bound to MHC Molecules,

Molecular Immunology, 38:1039-49 (Wang和Reinherz, 2002年, T细胞识别与MHC分子结合的肽的结构基础,《分子免疫学》,第38卷,第1039-1049页)。

[0095] CD4辅助受体的结构域D3和D4被认为与TCR-CD3复合体相互作用,因为这两个结构域的取代消除了CD4与TCR结合的能力。Vignali et al. (1996) The Two Membrane Proximal Domains of CD4 Interact with the T Cell Receptor, J. Exp. Med. 183:2097-2107 (Vignali等人, 1996年, CD4的两个近膜结构域与T细胞受体相互作用,《实验医学杂志》,第183卷,第2097-2107页)。CD4分子作为二聚体存在,并且该分子的D4结构域中的残基被认为是CD4二聚化的原因。Moldovan et al. (2002) CD4 Dimers Constitute the Functional Components Required for T Cell Activation, J. Immunol. 169:6261-68 (Moldovan等人, 2002年, CD4二聚体构成T细胞活化所需的功能组分,《免疫学杂志》,第169卷,第6261-6268页)。

[0096] CD4基因的外显子8编码跨膜结构域,而其余基因编码胞质结构域。CD4胞质结构域具有许多不同的功能。例如,CD4的胞质结构域募集酪氨酸激酶Lck。Lck是与CD4和CD8胞质结构域结合的Src家族激酶,并且辅助受体和TCR同时结合于相同的MHC导致TCR复合体的CD3和 $\zeta$ 链的酪氨酸磷酸化升高,这继而引起在T细胞活化中起作用的其他因子的募集。Itano及其同事通过设计包含CD8细胞外结构域和CD4胞质尾的杂合蛋白质并测试该杂合蛋白质在转基因小鼠中的表达,提出了CD4的胞质尾也能促进CD4+CD8+T细胞分化为CD4+谱系。Itano et al. (1996) The Cytoplasmic Domain of CD4 Promotes the Development of CD4 Lineage T Cells, J. Exp. Med. 183:731-41 (Itano等人, 1996年, CD4的胞质结构域促进CD4谱系T细胞的发育,《实验医学杂志》,第183卷,第731-741页)。杂合蛋白质的表达引起MHC I特异性CD4谱系T细胞的发育。同上。

[0097] CD4辅助受体似乎是HIV病毒的主要受体,其中CD4+T细胞清除是疾病进展的指标。CD4的胞质尾似乎对在HIV诱导的细胞凋亡中将细胞凋亡信号递送到CD4+T细胞至关重要。具体地讲,CD4和Lck的相互作用显示出增强这些细胞中HIV诱导的细胞凋亡。Corbeil et al. (1996) HIV-induced Apoptosis Requires the CD4 Receptor Cytoplasmic Tail and Is Accelerated by Interaction of CD4 with p56lck, J. Exp. Med. 183:39-48 (Corbeil等人, 1996年, HIV诱导的细胞凋亡需要CD4受体胞质尾并通过CD4与p56lck的相互作用加速,《实验医学杂志》,第183卷,第39-48页)。

[0098] T细胞在胸腺中发育,从未成熟CD4-/CD8- (双阴性或DN) 胸腺细胞发展为CD4+/CD8+ (双阳性或DP) 胸腺细胞,最后经过阳性选择而成为CD4+或CD8+ (单阳性或SP) T细胞。通过MHC I限制性TCR接收信号的DP胸腺细胞分化为CD8+T细胞,而通过MHC II限制性TCR接收信号的DP胸腺细胞分化为CD4+T细胞。由DP细胞接收的使其分化为CD4+或CD8+T细胞的信号已成为许多研究的课题。已经提出了各种用于CD4/CD8谱系选择的模型,并在以下文献中综述:Singer et al. (2008) Lineage fate and intense debate: myths, models and mechanisms of CD4-versus CD8-lineage choice, Nat. Rev. Immunol. 8:788-801 (Singer等人, 2008年, 谱系命运和激烈的争论: CD4与CD8谱系选择的流言、模型和机制,《自然综述: 免疫学》,第8卷,第788-801页)。

[0099] 特异性T细胞辅助受体因阳性选择所致的失活是转录调节的产物。对于CD4, 已证实位于CD4的外显子1上游13kb的增强子上调CD4+和CD8+T细胞中的CD4表达。Killeen et

al. (1993) Regulated expression of human CD4 rescues helper T cell development in mice lacking expression of endogenous CD4, *EMBO J.* 12:1547-53 (Killeen等人, 1993年, 人CD4的调节表达拯救缺少内源CD4表达的小鼠中的辅助T细胞发育, 《欧洲分子生物学学会杂志》, 第12卷, 第1547-1553页)。位于鼠CD4基因的第一内含子内的顺式作用转录沉默子起到使除CD4+T细胞之外的细胞中的CD4表达沉默的作用。Siu et al. (1994) A transcriptional silencer control the developmental expression of the CD4 gene, *EMBO J.* 13:3570-3579 (Siu等人, 1994年, 转录沉默子控制CD4基因的发育表达, 《欧洲分子生物学学会杂志》, 第13卷, 第3570-3579页)。

[0100] 由于先前开发的表达人CD4的转基因小鼠的数种品系中缺少控制CD4谱系选择的重要转录调节因子(例如, 启动子、增强子、沉默子等), 这些小鼠不能重新恢复正常T细胞谱系发育, 并产生除表达CD4的CD4+T细胞之外的免疫细胞。参见例如Law et al. (1994) Human CD4 Restores Normal T Cell Development and Function in Mice Deficient in CD4, *J. Exp. Med.* 179:1233-42 (Law等人, 1994年, 人CD4恢复CD4缺陷小鼠中的正常T细胞发育和功能, 《实验医学杂志》, 第179卷, 第1233-1242页) (CD4在CD8+T细胞和B细胞中表达); Fugger et al. (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, *Proc. Natl. Acad. Sci. USA*, 91:6151-55 (Fugger等人, 1994年, HLA-DR4和人CD4转基因在小鼠中的表达决定了可变区 $\beta$ 链T细胞组库并介导HLA-D限制性免疫应答, 《美国国家科学院院刊》, 第91卷, 第6151-6155页) (CD4在所有CD3+胸腺细胞和B细胞上表达)。因此, 在一个实施方案中, 可能有益的是开发保留内源小鼠启动子和其他调控元件的经遗传修饰的动物, 以便使该动物产生能够经历T细胞发育和谱系选择的T细胞。

[0101] 因此, 在各种实施方案中, 本发明提供了一种经遗传修饰的非人动物, 所述非人动物例如在其内源T细胞辅助受体基因座(例如, CD4基因座)处包含编码嵌合人/非人T细胞辅助受体多肽的核苷酸序列。在一个实施方案中, 嵌合多肽的人部分包含人T细胞辅助受体的全部或基本上全部细胞外部分(或其一部分, 例如一个或多个细胞外结构域, 例如至少两个连续细胞外结构域)。在一个实施方案中, 嵌合多肽的非人部分包含非人T细胞辅助受体的跨膜结构域和胞质结构域。在一个实施方案中, 非人动物表达功能性嵌合T细胞辅助受体多肽。因此, 在一个方面, 本发明提供了一种经遗传修饰的非人动物, 所述非人动物在其内源CD4基因座处包含编码嵌合人/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人CD4的全部或基本上全部细胞外部分, 其中非人部分包含非人CD4的至少跨膜结构域和胞质结构域, 并且其中所述动物表达功能性嵌合CD4多肽。在一个方面, 非人动物仅表达人源化CD4多肽, 即嵌合人/非人CD4多肽, 并且不从其内源CD4基因座表达功能性内源非人CD4蛋白质。

[0102] 在一个实施方案中, 嵌合人/非人CD4多肽的人部分包含人CD4多肽的全部或基本上全部细胞外部分。在另一个实施方案中, 嵌合人/非人CD4多肽的人部分包含人CD4多肽的至少全部或基本上全部MHC II结合结构域(例如人D1和D2结构域的大部分); 在一个实施方案中, 嵌合人/非人CD4多肽的人部分包含人CD4多肽的全部或基本上全部D1、D2和D3结构域; 在又一个实施方案中, 嵌合人/非人CD4多肽的人部分包含CD4的全部或基本上全部免疫

球蛋白样结构域,例如称为D1、D2、D3和D4的结构域。在又一个实施方案中,嵌合人/非人CD4多肽的人部分在其人部分中包含负责与MHC II和/或T细胞受体的细胞外部分相互作用的全部或基本上全部人CD4序列。在又一个实施方案中,嵌合人/非人CD4多肽的人部分包含负责与MHC II和/或T细胞受体的可变结构域相互作用的人CD4的全部或基本上全部细胞外部分。因此,在一个实施方案中,编码嵌合CD4多肽的人部分的核苷酸序列包含人CD4的结构域D1-D2的全部或基本上全部编码序列(例如,人CD4基因的外显子3和外显子4-5的一部分);在另一个实施方案中,其包含人CD4的D1-D3的全部或基本上全部编码序列(例如,人CD4的外显子3和外显子4-6的一部分)。因此,在一个实施方案中,编码嵌合人/非人CD4的核苷酸序列包含编码人CD4的全部或基本上全部D1-D3结构域的核苷酸序列。在另一个实施方案中,编码嵌合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。在另一个实施方案中,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。在另一个实施方案中,核苷酸序列可以包含编码人CD4信号肽的核苷酸序列。在一个实施方案中,嵌合人/非人CD4多肽包含SEQ ID NO:78所示的氨基酸序列,并且嵌合多肽的人部分跨越SEQ ID NO:78的大约第27-319位氨基酸(在SEQ ID NO:79中单独示出)。

[0103] 在一个实施方案中,非人动物表达嵌合人/非人CD4多肽序列。在一个实施方案中,嵌合CD4序列的人部分包含一个或多个保守或非保守修饰。

[0104] 在一个方面,提供了表达人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。在一个具体实施方案中,所述人CD4序列与实施例所述的人CD4序列具有至少约90%、95%、96%、97%、98%或99%的同一性。在一个实施方案中,人CD4序列包含一个或多个保守取代。在一个实施方案中,人CD4序列包含一个或多个非保守取代。

[0105] 在一些实施方案中,嵌合CD4的一部分例如人部分可以包含基本上全部本文所指示的序列(例如,基本上全部本文所指示的蛋白质结构域)。基本上全部序列通常包含据信代表蛋白质的特定部分(例如,特定功能性结构域等)的氨基酸的85%、90%、95%、96%、97%、98%或99%。本领域技术人员应当理解,功能性结构域的边界可能略有变化,具体取决于所使用的比对和结构域预测方法。

[0106] 在一个方面,嵌合人/非人CD4多肽的非人部分包含非人CD4多肽的至少跨膜结构域和胞质结构域。由于CD4胞质结构域所提供的重要功能,内源非人(例如,小鼠)序列在经遗传工程改造的动物中的保留确保了辅助受体的正确细胞内信号传导和其他功能的保存。在一个实施方案中,非人动物是小鼠,并且非人CD4多肽是小鼠CD4多肽。虽然实施例中描述了特定小鼠CD4序列,但是本文涵盖来源于其的任何合适的序列,例如包含保守/非保守氨基酸取代的序列。在一个实施方案中,嵌合CD4辅助受体的非人部分包含未被人源化的内源CD4的任何序列。

[0107] 本文所述的非人动物可在其内源基因座处包含编码嵌合人/非人CD4多肽的核苷酸序列。在一个方面,这使得内源CD4基因的一部分被替换为编码人CD4多肽的一部分的核苷酸序列。在一个实施方案中,这种替换是用编码其的人核苷酸序列替换编码例如非人CD4的全部或基本上全部细胞外结构域的内源核苷酸序列,例如编码非人CD4的至少全部或基本上全部第一免疫球蛋白样结构域(即,D1)的序列(例如,编码非人CD4的全部或基本上全

部结构域D1-D2的序列,例如编码非人CD4的全部或基本上全部结构域D1-D3的序列,例如编码非人CD4的全部或基本上全部结构域D1-D4的序列)。在一个实施方案中,该替换得到包含负责与MHC II和/或T细胞受体的细胞外部分相互作用的人CD4序列的嵌合蛋白质。在又一个实施方案中,该替换得到包含负责与MHC II和/或T细胞受体的可变结构域相互作用的人CD4序列的嵌合蛋白质。在一个实施方案中,该替换不包括编码非人CD4多肽的至少跨膜结构域和胞质结构域的CD4序列的替换。因此,在一个方面,非人动物从内源非人CD4基因座表达嵌合人/非人CD4多肽。在又一个实施方案中,该替换得到包含SEQ ID NO:78所示的多肽序列的蛋白质。

[0108] 在一个实施方案中,提供了本文所述的嵌合人/非人CD4基因座(例如,嵌合人/啮齿动物CD4基因座,例如嵌合人/小鼠CD4基因座)的核苷酸序列。在一个方面,因为嵌合人/非人(例如,人/啮齿动物,例如人/小鼠)CD4序列放置于内源非人(例如,啮齿动物,例如小鼠)CD4基因座处,因此其保留位于第一CD4外显子上游的CD4增强子元件。在一个实施方案中,在内源非人(例如,啮齿动物,例如小鼠)CD4基因座处的替换包括例如编码CD4多肽的D1一部分的外显子3及编码D1其余部分和D2-D3的外显子4-6的替换;因此,在一个方面,嵌合CD4基因座保留位于非人(例如小鼠)CD4基因的内含子1中的顺式作用沉默子。因此,在一个实施方案中,嵌合基因座保留内源非人(例如,啮齿动物,例如小鼠)CD4启动子和调控元件。在另一个实施方案中,嵌合基因座可含有人启动子和调控元件,只要这些元件允许正确的CD4表达、CD4+T细胞发育、CD4谱系选择和辅助受体功能即可。因此,在一些方面,本发明的动物包含不改变T细胞的正确谱系选择和发育的遗传修饰。在一个方面,本发明的动物(例如,啮齿动物,例如小鼠)不在除正常表达CD4的细胞之外的免疫细胞上表达嵌合CD4多肽。在一个方面,该动物不在B细胞或成熟CD8+T细胞上表达CD4。在一个实施方案中,该替换得以保留允许对CD4表达的正确时空调节的元件。

[0109] 在各种实施方案中,从本文所述嵌合CD4基因座表达功能性嵌合CD4蛋白质的非人动物(例如,啮齿动物,例如小鼠或大鼠)在细胞表面例如T细胞表面上展示嵌合蛋白质。在一个实施方案中,非人动物按照与人中观察到的相同的细胞分布在细胞表面上表达嵌合CD4蛋白质。在一个方面,本发明的CD4蛋白质能够与在第二细胞例如抗原递呈细胞(APC)的表面上表达的MHC II蛋白质相互作用。

#### [0110] 人或人源化CD8

[0111] 在各种实施方案中,本发明通常提供在其基因组中例如在内源CD8基因座处包含编码人或人源化CD8多肽的核苷酸序列的经遗传修饰的非人动物;因此,所述动物表达人或人源化CD8多肽。在各种实施方案中,本发明提供在其基因组中例如在内源CD8基因座处包含编码人或人源化CD8 $\alpha$ 多肽的核苷酸序列和/或编码人或人源化CD8 $\beta$ 多肽的核苷酸序列的非人动物。因此,本发明的经遗传修饰的非人动物表达人或人源化CD8 $\alpha$ 和/或人或人源化CD8 $\beta$ 多肽。

[0112] 人CD8蛋白质通常在细胞表面上表达为两种多肽CD8 $\alpha$ 和CD8 $\beta$ 的异源二聚体,但也已检测到二硫键连接的同源二聚体和同源多聚体(例如,在表达CD8 $\alpha\alpha$ 的NK细胞和肠 $\gamma\delta$ T细胞中)。编码人CD8 $\alpha$ 和CD8 $\beta$ 的基因彼此紧邻地位于2号染色体上。Nakayama et al. (1992) Recent Duplication of the Two Human CD8 $\beta$ -chain genes, J. Immunol. 148:1919-27 (Nakayama等人,1992年,两种人CD8 $\beta$ 链基因的最近复制,《免疫学杂志》,第148卷,第1919-

1927页)。CD8 $\alpha$ 蛋白质含有前导肽、免疫球蛋白V样区、铰链区、跨膜结构域和胞质尾。Norment et al. (1989) Alternatively Spliced mRNA Encodes a Secreted Form of Human CD8 $\alpha$ . Characterization of the Human CD8 $\alpha$  gene, J. Immunol. 142:3312-19 (Norment等人, 1989年, 选择性剪接的mRNA编码人CD8 $\alpha$ 的分泌形式。人CD8 $\alpha$ 基因的特征,《免疫学杂志》,第142卷,第3312-3319页)。CD8 $\alpha$ 基因的外显子/内含子示意性地描绘于图5B中。

[0113] 人CD8 $\beta$ 基因位于2号染色体上的CD8 $\alpha$ 基因上游。已经报道了通过CD8 $\beta$ 基因的选择性剪接而产生的多种同种型,其中一种同种型经预测得知缺少跨膜结构域并产生分泌蛋白质。Norment et al. (1988) A second subunit of CD8 is expressed in human T cells, EMBO J. 7:3433-39 (Norment等人, 1988年, CD8的第二亚基在人T细胞中表达,《欧洲分子生物学学会杂志》,第7卷,第3433-3439页)。CD8 $\beta$ 基因的外显子/内含子也示意性地描绘于图5B中。

[0114] 膜结合的CD8 $\beta$ 蛋白质含有N端信号序列,后接免疫球蛋白V样结构域、短细胞外铰链区、跨膜结构域和胞质尾。参见, Littman (1987) The structure of the CD4 and CD8 genes, Ann Rev. Immunol. 5:561-84 (Littman, 1987年, CD4和CD8基因的结构,《免疫学年评》,第5卷,第561-584页)。铰链区是广泛糖基化的位点,其被认为保持其构象并保护蛋白质不被蛋白酶裂解。Leahy (1995) A structural view of CD4 and CD8, FASEB J. 9:17-25 (Leahy, 1995年, CD4和CD8的结构视图,《美国实验生物学学会联合会杂志》,第9卷,第17-25页)。

[0115] CD8蛋白质通常在细胞毒性T细胞上表达,并与MHC I分子相互作用。该相互作用是通过CD8结合于MHC I的 $\alpha_3$ 结构域来介导的。虽然MHC I类与CD8的结合比TCR与MHC I类的结合弱约100倍,但CD8结合增强了TCR结合的亲和力。Wooldridge et al. (2010) MHC Class I Molecules with Superenhanced CD8 Binding Properties Bypass the Requirement for Cognate TCR Recognition and Nonspecifically Activate CTLs, J. Immunol. 184:3357-3366 (Wooldridge等人, 2010年, 具有超增强CD8结合特性的MHC I类分子绕过同源TCR识别的要求并且非特异性活化CTL,《免疫学杂志》,第184卷,第3357-3366页)。

[0116] CD8与MHC I类分子的结合是物种特异性的;CD8的小鼠同源物Lyt-2显示出在 $\alpha_3$ 结构域处结合H-2D<sup>d</sup>分子,但未结合HLA-A分子。Connolly et al. (1988) The Lyt-2 Molecule Recognizes Residues in the Class I  $\alpha_3$  Domain in Allogeneic Cytotoxic T Cell Responses, J. Exp. Med. 168:325-341 (Connolly等人, 1988年, Lyt-2分子在同种异体细胞毒性T细胞应答中识别I类 $\alpha_3$ 结构域中的残基,《实验医学杂志》,第168卷,第325-341页)。差异性结合大概是由于CD8上的CDR样决定簇(CDR1和CDR2样)在人和小鼠之间不保守。Sanders et al. (1991) Mutations in CD8 that Affect Interactions with HLA Class I and Monoclonal Anti-CD8 Antibodies, J. Exp. Med. 174:371-379 (Sanders等人, 1991年, CD8中影响与HLA I类和单克隆抗CD8抗体的相互作用的突变,《实验医学杂志》,第174卷,第371-379页); Vitiello et al. (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, J. Exp. Med. 173:1007-1015 (Vitiello等人, 1991年, 携带嵌合人-小鼠I类主要组织相容性复合体的转基因小鼠中HLA限制性流感特异性细胞毒性T淋巴细胞应答的分析,《实验医学杂志》,第173卷,第1007-



1015页);以及Gao et al.(1997)Crystal structure of the complex between human CD8 $\alpha$  and HLA-A2, *Nature* 387:630-634 (Gao等人,1997年,人CD8 $\alpha$ 与HLA-A2之间的复合体的晶体结构,《自然》,第387卷,第630-634页)。已经报道CD8在 $\alpha$ 3结构域的保守区中(在第223-229位)结合HLA-A2。HLA-A中的单取代(V245A)降低了CD8与HLA-A的结合,伴随着T细胞介导的裂解的大幅减少。Salter et al.(1989),Polymorphism in the  $\alpha$ 3 domain of HLA-A molecules affects binding to CD8, *Nature* 338:345-348 (Salter等人,1989年,HLA-A分子的 $\alpha$ 3结构域中的多态性影响与CD8的结合,《自然》,第338卷,第345-348页)。一般来讲,HLA-A分子的 $\alpha$ 3结构域中的多态性也影响与CD8的结合。同上。在鼠中,H-2D<sup>d</sup>中的残基227处的氨基酸取代影响小鼠Lyt-2与H-2D<sup>d</sup>的结合,并且用突变体H-2D<sup>d</sup>转染的细胞未被CD8+T细胞裂解。Potter et al.(1989)Substitution at residue 227 of H-2 class I molecules abrogates recognition by CD8-dependent, but not CD8-independent, cytotoxic T lymphocytes, *Nature* 337:73-75 (Potter等人,1989年,H-2I类分子的残基227处的取代消除CD8依赖性而不是CD8非依赖性细胞毒性T淋巴细胞的识别,《自然》,第337卷,第73-75页)。因此,人或人源化CD8的表达可能有利于研究对由人或人源化MHC I递呈的抗原的T细胞应答。

[0117] 与CD4类似,CD8的胞质结构域与酪氨酸激酶Lck相互作用,继而引起T细胞活化。虽然Lck似乎与CD8 $\alpha$ 的胞质结构域相互作用,但似乎这种相互作用受到CD8 $\beta$ 胞质结构域的存在调节,因为CD8 $\beta$ 胞质结构域的突变或缺失导致CD8 $\alpha$ 结合的Lck活性降低。Irie et al.(1998)The cytoplasmic domain of CD8 $\beta$  Regulates Lck Kinase Activation and CD8 T cell Development, *J. Immunol.* 161:183-91 (Irie等人,1998年,CD8 $\beta$ 的胞质结构域调节Lck激酶活化和CD8 T细胞发育,《免疫学杂志》,第161卷,第183-191页)。Lck活性的降低与T细胞发育的损害有关。同上。

[0118] CD8在合适的细胞(例如细胞毒性T细胞)上的表达受到位于整个CD8基因座中的多种增强子元件的严格调节。例如,已在CD8基因座处鉴定了DNA酶I超敏反应的至少4个区域,即通常与调节因子结合相关的区域。Hosert et al.(1997)A CD8 genomic fragment that directs subset-specific expression of CD8 in transgenic mice, *J. Immunol.* 158:4270-81 (Hosert等人,1997年,指导转基因小鼠中CD8的亚群特异性表达的CD8基因组片段,《免疫学杂志》,第158卷,第4270-4281页)。自在CD8基因座处发现这些DNA酶I超敏反应区以来,已经鉴定出至少5个增强子元件,这些增强子元件遍及CD8基因座,调节CD8 $\alpha$ 和/或 $\beta$ 在各种谱系的T细胞(包括DP、CD8SP T细胞或表达 $\gamma$   $\delta$ TCR的细胞)中的表达。参见例如Kioussis et al.(2002)Chromatin and CD4, CD8A, and CD8B gene expression during thymic differentiation, *Nature Rev.* 2:909-919 (Kioussis等人,2002年,胸腺分化期间的染色质及CD4、CD8A和CD8B基因表达,《自然评论》,第2卷,第909-919页)和在线勘误;Ellmeier et al.(1998)Multiple Development Stage-Specific Enhancers Regulate CD8 Expression in Developing Thymocytes and in Thymus-Independent T cells, *Immunity* 9:485-96。(Ellmeier等人,1998年,多个发育阶段特异性增强子调节发育胸腺细胞中和胸腺非依赖性T细胞中的CD8表达,《免疫力》,第9卷,第485-496页)。

[0119] 因此,与从人或人源化CD4遗传修饰动物保留内源CD4启动子和调控元件获得的益处相似,在一些实施方案中,可能有利的是开发保留将控制人或人源化CD8的表达的内源小

鼠启动子和调控元件的经遗传修饰的非人动物。可能特别有利的是形成这样经遗传修饰的动物,其包含用编码人或人源化CD8 $\alpha$ 和/或 $\beta$ 蛋白质的那些序列替换编码CD8 $\alpha$ 和/或 $\beta$ 蛋白质的内源非人序列,如本文所述。

[0120] 在各种实施方案中,本发明提供了一种经遗传修饰的非人动物,该非人动物在其基因组中例如在其内源CD8基因座处包含编码嵌合人/非人CD8多肽(例如CD8 $\alpha$ 和/或 $\beta$ 多肽)的至少一个核苷酸序列,其中所述多肽的人部分包含人CD8多肽(例如,CD8 $\alpha$ 和/或 $\beta$ )的全部或基本上全部细胞外部分(例如,或其一部分,例如细胞外结构域),其中非人部分包含非人CD8(例如CD8 $\alpha$ 和/或 $\beta$ )的至少跨膜结构域和胞质结构域,并且其中该动物表达嵌合CD8多肽(例如,CD8 $\alpha$ 和/或 $\beta$ 多肽)。因此,在一个实施方案中,本发明提供了一种经遗传修饰的非人动物,该非人动物在其内源非人CD8基因座处包含编码嵌合人/非人CD8 $\alpha$ 多肽的第一核苷酸序列和编码嵌合人/非人CD8 $\beta$ 多肽的第二核苷酸序列,其中第一核苷酸序列包含编码人CD8 $\alpha$ 多肽的全部或基本上全部细胞外部分以及非人CD8 $\alpha$ 多肽的至少跨膜结构域和胞质结构域的序列,并且其中第二核苷酸序列包含编码人CD8 $\beta$ 多肽的全部或基本上全部细胞外部分以及非人CD8 $\beta$ 多肽的至少跨膜结构域和胞质结构域的序列,其中该动物表达功能性嵌合人/非人CD8蛋白质。在一个方面,非人动物仅表达人源化CD8多肽(例如,嵌合人/非人CD8 $\alpha$ 和/或 $\beta$ 多肽),并且不从内源CD8基因座表达对应的功能性非人CD8多肽。

[0121] 在一个实施方案中,嵌合人/非人CD8 $\alpha$ 多肽在其人部分中包含人CD8 $\alpha$ 多肽的全部或基本上全部细胞外部分。在一个实施方案中,嵌合CD8 $\alpha$ 多肽的人部分包含人CD8 $\alpha$ 多肽的至少MHC I结合结构域。在一个实施方案中,嵌合CD8 $\alpha$ 多肽的人部分包含人CD8 $\alpha$ 的至少全部或基本上全部免疫球蛋白V样结构域的序列。在一个实施方案中,编码嵌合CD8 $\alpha$ 多肽的人部分的核苷酸序列至少包含编码人CD8 $\alpha$ 多肽的细胞外部分的外显子。在一个实施方案中,核苷酸序列至少包含编码Ig V样结构域的外显子。在一个实施方案中,人CD8 $\alpha$ 多肽的细胞外部分是包含并非跨膜或胞质结构域的多肽部分的区域。在一个实施方案中,编码嵌合人/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,该核苷酸序列可以包含编码人CD8 $\alpha$ 信号肽的序列。在一个实施方案中,嵌合人/非人CD8 $\alpha$ 多肽包含SEQ ID NO:88所示的氨基酸序列,并且嵌合多肽的人部分在SEQ ID NO:88的第28-179位氨基酸处示出(在SEQ ID NO:89中单独表示)。

[0122] 相似地,在一个实施方案中,嵌合人/非人CD8 $\beta$ 多肽在其人部分中包含人CD8 $\beta$ 多肽的全部或基本上全部细胞外部分。在一个实施方案中,嵌合CD8 $\beta$ 多肽的人部分包含人CD8 $\beta$ 的全部或基本上全部免疫球蛋白V样结构域的序列。在一个实施方案中,编码嵌合CD8 $\beta$ 多肽的人部分的核苷酸序列至少包含编码人CD8 $\beta$ 多肽的细胞外部分的外显子。在一个实施方案中,编码嵌合人/非人CD8 $\beta$ 多肽的人部分的核苷酸序列至少包含编码人CD8 $\beta$ 的IgG V样结构域的外显子。在一个实施方案中,编码嵌合人/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,该核苷酸序列可以包含编码人CD8 $\beta$ 信号肽的序列。在一个实施方案中,嵌合人/非人CD8 $\beta$ 多肽包含SEQ ID NO:83所示的氨基酸序列,并且嵌合多肽的人部分在SEQ ID NO:83的第15-165位氨基酸处示出(在SEQ ID NO:84中单独表示)。

[0123] 在一个实施方案中,非人动物表达嵌合人/非人CD8 $\alpha$ 和/或CD8 $\beta$ 多肽。在一些实施方案中,嵌合人/非人CD8 $\alpha$ 和/或 $\beta$ 多肽的人部分包含一个或多个保守或非保守修饰。

[0124] 在一个方面,提供了表达人CD8 $\alpha$ 和/或 $\beta$ 多肽序列的非人动物,其中人CD8 $\alpha$ 和/或 $\beta$ 多肽序列分别与人CD8 $\alpha$ 和/或 $\beta$ 多肽序列具有至少约85%、90%、95%、96%、97%、98%或99%的同一性。在一个具体实施方案中,人CD8 $\alpha$ 和/或 $\beta$ 多肽序列与实施例中所描述的相应人CD8 $\alpha$ 和/或 $\beta$ 多肽序列具有至少约90%、95%、96%、97%、98%或99%的同一性。在一个实施方案中,人CD8 $\alpha$ 和/或 $\beta$ 多肽序列包含一个或多个保守取代。在一个实施方案中,人CD8 $\alpha$ 和/或 $\beta$ 多肽序列包含一个或多个非保守取代。

[0125] 在一些实施方案中,嵌合CD8的一部分例如人部分可以包含基本上全部本文所指示的序列(例如,基本上全部本文所指示的蛋白质结构域)。基本上全部序列通常包含据信代表蛋白质的特定部分(例如,特定功能性结构域等)的氨基酸的85%、90%、95%、96%、97%、98%或99%。本领域技术人员应当理解,功能性结构域的边界可能略有变化,具体取决于所使用的比对和结构域预测方法。

[0126] 在一个方面,嵌合人/非人CD8 $\alpha$ 和/或 $\beta$ 多肽的非人部分分别包含非人CD8 $\alpha$ 和/或 $\beta$ 多肽的至少跨膜和/或胞质结构域。由于CD8胞质结构域所提供的重要功能,内源非人(例如,小鼠)序列在经遗传工程改造的动物中的保留确保了辅助受体的正确细胞内信号传导和其他功能的保存。在一个实施方案中,非人动物是小鼠,并且非人CD8 $\alpha$ 和/或 $\beta$ 多肽分别是小鼠CD8 $\alpha$ 和/或 $\beta$ 多肽。虽然实施例中描述了特定小鼠CD8 $\alpha$ 和 $\beta$ 序列,但是本文涵盖来源于其的任何合适的序列,例如包含保守/非保守氨基酸取代的序列。在一个实施方案中,非人动物(例如,啮齿动物,例如小鼠)保留未被人源化的任何内源序列。

[0127] 本文所述的非人动物可在其内源基因座处包含编码嵌合人/非人CD8 $\alpha$ 和/或 $\beta$ 多肽的核苷酸序列。在一个方面,这使得内源CD8 $\alpha$ 基因的一部分被替换为编码人CD8 $\alpha$ 多肽的一部分的核苷酸序列,和/或内源CD8 $\beta$ 基因的一部分被替换为编码人CD8 $\beta$ 多肽的一部分的核苷酸序列。在一个实施方案中,这种替换是用具有编码其的人核苷酸序列的人核苷酸替换编码非人CD8 $\alpha$ 和/或 $\beta$ 的全部或基本上全部细胞外部分的内源核苷酸序列。在一个实施方案中,这种替换是用编码其的人核苷酸序列替换编码非人CD8 $\alpha$ 和/或 $\beta$ 的至少全部或基本上全部免疫球蛋白V样结构域的序列。在一个实施方案中,该替换不包括编码非人CD8 $\alpha$ 和/或 $\beta$ 多肽的跨膜结构域和胞质结构域的CD8 $\alpha$ 和/或 $\beta$ 序列的替换。因此,非人动物从内源非人CD8基因座表达嵌合人/非人CD8 $\alpha$ 和/或 $\beta$ 多肽。在又一个实施方案中,该替换得到分别包含SEQ ID NO:88和/或84所示的多肽序列的CD8 $\alpha$ 和/或 $\beta$ 蛋白质。

[0128] 在一个实施方案中,提供了嵌合人/非人CD8基因座(例如,嵌合啮齿动物CD8基因座,例如嵌合小鼠CD8基因座)的核苷酸序列。在一个方面,由于嵌合人/非人(例如,人/啮齿动物,例如人/小鼠)CD8 $\alpha$ 和/或 $\beta$ 序列放置在相应内源非人(例如,啮齿动物,例如小鼠)CD8 $\alpha$ 和/或 $\beta$ 基因座,其保留内源CD8 $\alpha$ 和/或 $\beta$ 启动子和调控元件。在另一个实施方案中,嵌合基因座可含有人CD8 $\alpha$ 和/或 $\beta$ 启动子和调控元件,只要这些元件允许正确的CD8 $\alpha$ 和/或 $\beta$ 表达(正确的时空蛋白质表达)、CD8+T细胞发育、CD8谱系选择和辅助受体功能即可。因此,在一个方面,本发明的动物包含不改变T细胞的正确谱系选择和发育的遗传修饰。在一个方面,本发明的动物(例如,啮齿动物,例如小鼠)不在除正常表达CD8的细胞之外的免疫细胞上表达嵌合CD8蛋白质,例如,动物不在B细胞或成熟CD4+T细胞上表达CD8。在一个实施方案中,该替换得以保留允许对CD8 $\alpha$ 和/或 $\beta$ 表达的正确时空调节的元件。

[0129] 在各种实施方案中,从本文所述嵌合CD8基因座表达功能性嵌合CD8蛋白质(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)在细胞表面上展示嵌合蛋白质。在一个实施方案中,非人动物按照与人中观察到的相同的细胞分布在细胞表面上表达嵌合CD8蛋白质。在一个方面,本发明的CD8蛋白质能够与在第二细胞的表面上表达的MHC I蛋白质相互作用。

#### [0130] 人或人源化T细胞受体

[0131] 本文公开了包含基本上人源化的T细胞免疫系统的经遗传修饰的非人动物。在某个实施方案中,本文所公开的非人动物例如在其基因组中包含(a)编码嵌合人/非人T细胞辅助受体的核苷酸序列,其中嵌合T细胞辅助受体多肽的人部分由编码人T细胞辅助受体的细胞外结构域的序列编码,并且其中编码人T细胞辅助受体的细胞外结构域的序列可操作地连接至包含编码非人T细胞辅助受体跨膜和/或胞质结构域的序列的核苷酸;(b)包含至少一个人V区段、任选的至少一个人D区段和至少一个人J区段的未重排T细胞受体(TCR)可变基因区,其中TCR可变区基因的未重排V、任选的D和J区段可重组形成可操作地连接至非人TCR恒定基因序列的重排基因;以及(c)编码嵌合人/非人MHC多肽的核苷酸序列,其中嵌合MHC多肽的人部分包含与嵌合T细胞辅助受体多肽的人部分结合的人MHC多肽的细胞外结构域。任选地,非人动物还包含人或人源化 $\beta$ 2微球蛋白多肽。

[0132] 因此,在各种实施方案中,本发明通常提供经遗传修饰的非人动物,其中非人动物在基因组中包含未重排人源化TCR可变基因座,例如,包含能够重组形成重排TCR可变基因序列的人TCR可变区段的未重排人TCR可变基因区。如本文所用,TCR基因座或TCR基因座位(例如,TCR $\alpha$ 基因座或TCR $\beta$ 基因座)是指包含TCR编码区的基因组DNA,所述TCR编码区包括整个TCR编码区,包括未重排V(D)J序列、增强子序列、恒定序列和任何上游或下游(UTR、调控区等)或间插DNA序列(内含子等)。TCR可变基因座、TCR可变区或TCR可变基因座位(例如,TCR $\alpha$ 可变基因座或TCR $\beta$ 可变基因座)是指包含TCR可变区区段(V(D)J区)但不包含TCR恒定序列及在各种实施方案中增强子序列的基因组DNA。其他序列可以包含在TCR可变基因座中以便进行遗传操纵(例如,选择盒、限制性位点等),并且这些序列涵盖于本文中。

[0133] T细胞结合在抗原递呈细胞表面上与主要组织相容性复合体(MHC;在鼠中)或人白细胞抗原(HLA;在人中)复合体结合的小抗原决定簇上的表位。T细胞通过T细胞表面上的T细胞受体(TCR)复合体结合这些表位。T细胞受体是由以下两种类型的链构成的异源二聚体结构: $\alpha$ (alpha)和 $\beta$ (beta)链,或 $\gamma$ (gamma)和 $\delta$ (delta)链。 $\alpha$ 链由位于 $\alpha$ 基因座(在人或小鼠14号染色体上,也涵盖整个 $\delta$ 基因座)内的核苷酸序列编码,并且 $\beta$ 链由位于 $\beta$ 基因座(小鼠6号染色体或人7号染色体上)内的核苷酸序列编码。大多数T细胞具有 $\alpha\beta$ TCR;而少数T细胞携带 $\gamma\delta$ TCR。TCR与MHC I类(递呈到CD8 $^{+}$ T细胞)和MHC II类(递呈到CD4 $^{+}$ T细胞)分子的相互作用示于图1中(闭合符号表示非人序列;条纹符号表示人序列,示出了本发明的TCR蛋白质的一个特定实施方案)。

[0134] T细胞受体 $\alpha$ 和 $\beta$ 多肽(及类似的 $\gamma$ 和 $\delta$ 多肽)彼此通过二硫键连接。构成TCR的这两种多肽中的每一种含有包含恒定和可变区的细胞外结构域、跨膜结构域和胞质尾(跨膜结构域和胞质尾也是恒定区的一部分)。TCR的可变区决定其抗原特异性,并且与免疫球蛋白类似,也包含三个互补决定区(CDR)。也与免疫球蛋白基因类似,T细胞受体可变基因座(例如,TCR $\alpha$ 和TCR $\beta$ 基因座)含有许多未重排V(D)J区段(可变(V)、连接(J)及TCR $\beta$ 和 $\delta$ 中的多样性(D)区段)。在胸腺中的T细胞发育期间,TCR $\alpha$ 可变基因座经历重排,使得所得的TCR $\alpha$ 链由

VJ区段的特定组合 (Va/J $\alpha$ 序列) 编码;并且TCR $\beta$ 可变基因座经历重排,使得所得的TCR $\beta$ 链由VDJ区段的特定组合 (V $\beta$ /D $\beta$ /J $\beta$ 序列) 编码。

[0135] 与胸腺基质的相互作用触发胸腺细胞经历数个发育阶段,其特征在于各种细胞表面标记的表达。胸腺中各个发育阶段的特征性细胞表面标记的总结在表1中给出。TCR $\beta$ 可变基因座处的重排在DN2阶段开始并在DN4阶段结束,而TCR $\alpha$ 可变基因座的重排在DP阶段发生。在TCR $\beta$ 基因座重排完成后,所述细胞在细胞表面同时表达TCR $\beta$ 链和替代 $\alpha$ 链pT $\alpha$ 。参见, Janeway's Immunobiology, Chapter 7, 7<sup>th</sup> Ed., Murphy et al. eds., Garland Science, 2008 (《詹韦氏免疫生物学》,第7章,第7版, Murphy等人编辑,加兰科学出版社,2008年)。

[0136] 表1:胸腺中的T细胞的发育阶段

[0137]	发育阶段	DN1	DN2	DN3	DN4	DP	SP
	标记	CD44+/CD25-	CD44+/CD25+	CD44 <sup>低</sup> /CD25+	CD44- / CD25-	CD4+/CD8+	CD4+ 或 CD8+

[0138] 初始CD4<sup>+</sup>和CD8<sup>+</sup>T细胞离开胸腺并进入外周淋巴器官(例如,脾),它们在此暴露于抗原并被活化,从而以无性繁殖方式扩增并分化为多种效应T细胞(Teff),例如细胞毒性T细胞、T<sub>REG</sub>细胞、T<sub>H</sub>17细胞、T<sub>H</sub>1细胞、T<sub>H</sub>2细胞等。感染之后,多种T细胞作为记忆T细胞存留,并且被归类为中枢记忆T细胞(Tcm)或效应记忆T细胞(Tem)。Sallusto et al. (1999) Two subsets of memory T lymphocytes with distinct homing potentials and effector functions, Nature 401:708-12 (Sallusto等人,1999年,具有不同归巢潜能和效应功能的两种记忆T淋巴细胞亚群,《自然》,第401卷,第708-712页)以及Mackay (1999) Dual personality of memory T cells, Nature 401:659-60 (Mackay,1999年,记忆T细胞的两面性,《自然》,第401卷,第659-660页)评论性文章。Sallusto及其同事提出,在初始感染后, Tem细胞代表具有效应功能的外周组织中易于获得的抗原致敏记忆T细胞库,而Tcm细胞代表外周淋巴器官中的抗原致敏记忆T细胞,它们在二次激发后可以成为新的效应T细胞。虽然所有记忆T细胞都表达CD45的CD45R0同种型(初始T细胞表达CD45RA同种型),但Tcm的特征在于表达L-选择素(也称为CD62L)和CCR7<sup>+</sup>,它们对结合于外周淋巴器官和淋巴结及其中的信号传导很重要。同上。因此,存在于外周淋巴器官中的所有T细胞(例如,初始T细胞、Tcm细胞等)都表达CD62L。除CD45R0之外,已知所有记忆T细胞还表达许多不同的细胞表面标记,例如CD44。关于T细胞上各种细胞表面标记的总结,参见Janeway's Immunobiology (《詹韦氏免疫生物学》),第10章,出处同上。

[0139] 虽然TCR可变区主要在抗原识别中起作用,但TCR的恒定结构域的细胞外部分以及跨膜结构域和胞质结构域也发挥着重要功能。完整的TCR受体复合体不止需要 $\alpha$ 和 $\beta$ 或 $\gamma$ 和 $\delta$ 多肽;另外所需的分子包括CD3 $\gamma$ 、CD3 $\delta$ 和CD3 $\epsilon$ 以及 $\zeta$ 链同源二聚体( $\zeta\zeta$ )。在完成TCR $\beta$ 重排时,若细胞表达TCR $\beta$ /pT $\alpha$ ,则这种前TCR复合体与CD3一起存在于细胞表面上。细胞表面上的TCR $\alpha$ (或pT $\alpha$ )在其跨膜结构域中具有两个碱性残基,其中一个募集CD3 $\gamma\epsilon$ 异源二聚体,并且另一个通过其各自的酸性残基募集 $\zeta\zeta$ 。TCR $\beta$ 在其跨膜结构域中还有一个额外的碱性残基,该碱性残基被认为募集CD3 $\delta\epsilon$ 异源二聚体。参见,例如, Kuhns et al. (2006) Deconstructing the Form and Function of the TCR/CD3 Complex, Immunity 24:133-39 (Kuhns等人,2006年,解构TCR/CD3复合体的形态和功能,《免疫力》,第24卷,第133-139页);

Wucherpennig et al. (2009) Structural Biology of the T-cell Receptor: Insights into Receptor Assembly, Ligand Recognition, and Initiation of Signaling, Cold Spring Harb. Perspect. Biol. 2: a005140 (Wucherpennig等人, 2009年, T细胞受体的结构生物学: 洞察受体组装、配体识别和信号引发, 《冷泉港实验室生物学展望》, 第2卷, 第a005140页)。包含TCR $\alpha\beta$ 异源二聚体、CD3 $\gamma\epsilon$ 、CD3 $\delta\epsilon$ 和 $\zeta\zeta$ 的组装复合体在T细胞表面上表达。已提议跨膜结构域中的极性残基用作离开内质网的质量控制; 已证实, 在不存在CD3亚基的情况下, TCR链保留在ER中并被靶向降解。参见, 例如, Call and Wucherpennig (2005) The T Cell Receptor: Critical Role of the Membrane Environment in Receptor Assembly and Function, Annu. Rev. Immunol. 23: 101-25 (Call和Wucherpennig, 2005年, T细胞受体: 膜环境在受体组装和功能中的关键作用, 《免疫学年评》, 第23卷, 第101-125页)。

[0140] 组装复合体的CD3和 $\zeta$ 链为TCR信号传导提供组分, 因为TCR $\alpha\beta$ 异源二聚体(或TCR $\gamma\delta$ 异源二聚体)本身缺少信号转导活性。CD3链各有一个免疫受体酪氨酸活化基序(ITAM), 而 $\zeta$ 链含有三个串联ITAM。ITAM含有能够被相关激酶磷酸化的酪氨酸残基。因此, 组装的TCR-CD3复合体含有10个ITAM基序。参见, 例如, Love and Hayes (2010) ITAM-Mediated Signaling by the T-Cell Antigen Receptor, Cold Spring Harb. Perspect. Biol. 2: e002485 (Love和Hayes, 2010年, 由T细胞抗原受体引起的ITAM介导的信号传导, 《冷泉港实验室生物学展望》, 第2卷, 第e002485页)。在TCR结合后, ITAM基序被Src家族酪氨酸激酶Lck和Fyn磷酸化, 这会引发信号级联, 导致Ras活化、钙动员、肌动蛋白细胞骨架重排和转录因子的活化, 所有这些最终引起T细胞分化、增殖和效应作用。同上, 还可参见(《詹韦氏免疫生物学》), 出处同上; 这两篇文献以引用方式并入本文。

[0141] 另外, TCR $\beta$ 跨膜结构域和胞质结构域被认为在线粒体靶向和细胞凋亡诱导中起作用; 事实上, 天然存在的N端截短TCR $\beta$ 分子存在于胸腺细胞中。Shani et al. (2009) Incomplete T-cell receptor-- $\beta$ peptides target the mitochondrion and induce apoptosis, Blood 113: 3530-41 (Shani等人, 2009年, 不完全T细胞受体 $\beta$ 肽靶向线粒体并诱导细胞凋亡, 《血液》, 第113卷, 第3530-3541页)。因此, TCR恒定区(其在各种实施方案中包含细胞外以及跨膜结构域和胞质结构域的一部分)提供了若干重要功能; 并且在各种实施方案中, 当设计人源化TCR或表达其的经遗传修饰的非人动物时, 应考虑该区域的结构。

[0142] 重排T细胞受体序列转基因小鼠是本领域已知的。本发明涉及包含能够重排形成编码人T细胞受体可变结构域的核酸序列的未重排人或人源化T细胞可变基因座的、经遗传修饰的非人动物(例如, 啮齿动物, 例如大鼠、小鼠), 包括包含具有重排人可变结构域和非人(例如, 小鼠或大鼠)恒定区的T细胞的动物。本发明还提供了能够产生人T细胞受体可变区序列的多样性组库的非人动物(例如, 啮齿动物, 例如大鼠、小鼠); 因此, 本发明提供了非人动物, 其对目标抗原作出应答而表达具有全人可变结构域的TCR并且结合目标抗原的表位。在一些实施方案中, 提供了产生能够与各种抗原(包括但不限于APC所递呈的抗原)反应的多样性T细胞受体组库的非人动物。

[0143] 在一个实施方案中, 本发明提供了在其基因组中包含未重排人TCR可变区区段(V(D)J区段)的经遗传修饰的非人动物(例如, 啮齿动物, 例如大鼠、小鼠), 其中未重排人TCR可变区区段在内源非人(例如, 啮齿动物)TCR可变基因座(例如TCR $\alpha$ 、 $\beta$ 、 $\delta$ 和/或 $\gamma$ 可变基因座)处替换内源非人TCR可变区区段。在一个实施方案中, 未重排人TCR可变基因座替换内源

非人TCR可变基因座。

[0144] 在另一个实施方案中,本发明提供了在其基因组中包含未重排人TCR可变区区段(V(D)J区段)的经遗传修饰的非人动物(例如,啮齿动物,例如大鼠、小鼠),其中未重排人TCR可变区区段可操作地连接至非人TCR恒定区基因序列,得到人源化TCR基因座,其中人源化TCR基因座位于基因组中除内源非人TCR基因座之外的位点。因此,在一个实施方案中,还提供了包含转基因的非人动物(例如,啮齿动物,例如小鼠、大鼠),该转基因包含可操作地连接至非人TCR恒定区基因序列的未重排人TCR可变区区段。

[0145] 在一个方面,本发明的经遗传修饰的非人动物在其基因组中包含人TCR可变区区段,同时保留编码TCR恒定结构域的非人(例如,啮齿动物,例如小鼠、大鼠)TCR恒定基因序列。在各种实施方案中,TCR恒定结构域包括TCR的跨膜结构域和胞质尾。因此,在本发明的各种实施方案中,经遗传修饰的非人动物保留内源非人TCR跨膜结构域和胞质尾。在其他实施方案中,非人动物包含非人非内源TCR恒定基因序列,例如其编码非人非内源TCR跨膜结构域和胞质尾。如上所指出,TCR的恒定结构域参与在抗原致敏T细胞活化过程中引发的信号级联;因此,内源TCR恒定结构域与T细胞中的多种非人锚定蛋白质和信号传导蛋白质相互作用。因此,在一个方面,本发明的经遗传修饰的非人动物表达人源化T细胞受体,其保留募集多种内源非人锚定分子或信号传导分子的能力,所述分子例如为CD3分子(例如,CD3 $\gamma$ 、CD3 $\delta$ 、CD3 $\epsilon$ )、 $\zeta$ 链、Lck、Fyn、ZAP-70等。被募集到TCR复合体的分子的非限制性列表描述于Janeway's Immunobiology(《詹韦氏免疫生物学》),出处同上。据信,非人动物中的T细胞发育和T细胞分化过程能够推进并允许强效的免疫应答,这可能至少部分地归因于可变区在内源小鼠基因座处的放置和小鼠恒定结构域的保留。

[0146] 在一些实施方案中,提供了非人动物,该非人动物在其基因组中包含未重排人TCR $\alpha$ 可变区区段,其中未重排人TCR $\alpha$ 可变区区段可操作地连接至非人TCR $\alpha$ 恒定区基因序列,得到人源化TCR $\alpha$ 基因座。在一个实施方案中,人源化TCR $\alpha$ 基因座位于基因组中除内源非人TCR $\alpha$ 基因座之外的位点。在另一个实施方案中,未重排人TCR $\alpha$ 可变区区段替换内源非人TCR $\alpha$ 可变区区段,同时保留内源非人TCR $\alpha$ 恒定区基因序列。在一个实施方案中,未重排人TCR $\alpha$ 可变基因座替换内源非人TCR $\alpha$ 可变基因座。在一些实施方案中,用未重排人TCR $\alpha$ 可变基因座替换内源非人TCR $\alpha$ 可变区基因座包括TCR $\delta$ 可变基因座的缺失或失活。在其他实施方案中,用未重排人TCR $\alpha$ 基因座替换内源非人TCR $\alpha$ 可变区基因座包括用未重排人TCR $\delta$ 可变区区段替换内源TCR $\delta$ 可变基因座。在一些实施方案中,该动物保留内源非人TCR $\beta$ 可变区和恒定区基因序列。因此,该动物表达包含嵌合人/非人(即人源化)TCR $\alpha$ 链和非人TCR $\beta$ 链的TCR。

[0147] 在一些实施方案中,提供了非人动物,该非人动物在其基因组中包含未重排人TCR $\delta$ 可变区区段,其中未重排人TCR $\delta$ 可变区区段可操作地连接至非人TCR $\delta$ 恒定区基因序列,得到人源化TCR $\delta$ 基因座。在一个实施方案中,人源化TCR $\delta$ 基因座位于基因组中除内源非人TCR $\delta$ 基因座之外的位点。在另一个实施方案中,未重排人TCR $\delta$ 可变区区段替换内源非人TCR $\delta$ 可变区区段,同时保留内源非人TCR $\delta$ 恒定区基因序列。在一个实施方案中,未重排人TCR $\delta$ 可变基因座替换内源非人TCR $\delta$ 可变基因座。

[0148] 在其他实施方案中,提供了非人动物,该非人动物在其基因组中包含未重排人TCR $\beta$ 可变区区段,其中未重排人TCR $\beta$ 可变区区段可操作地连接至非人TCR $\beta$ 恒定区基因序列,得到人源化TCR $\beta$ 基因座。在一个实施方案中,人源化TCR $\beta$ 基因座位于基因组中除内源非人TCR



β基因座之外的位点。在另一个实施方案中,未重排人TCRβ可变区区段替换内源非人TCRβ可变区区段,同时保留内源非人TCRβ恒定区基因序列。在一个实施方案中,未重排人TCRβ可变基因座替换内源非人TCRβ可变基因座。在一些实施方案中,该动物保留内源非人TCRα可变区和恒定区基因序列。因此,该动物表达包含嵌合人/非人(即人源化)TCRβ链和非人TCRα链的TCR。

[0149] 在一些具体实施方案中,本发明提供了一种经遗传修饰的非人动物(例如,啮齿动物,例如小鼠或大鼠),该非人动物在其基因组中包含(a)包含至少一个人Vα区段和至少一个人Jα区段并可操作地连接至内源非人(例如,啮齿动物,例如小鼠或大鼠)TCRα恒定基因序列的未重排T细胞受体(TCR)α可变基因座,(b)包含至少一个人Vβ区段、至少一个人Dβ区段和至少一个人Jβ区段并可操作地连接至内源非人(例如,啮齿动物,例如小鼠或大鼠)TCRβ恒定区基因序列的未重排TCRβ可变基因座,和/或(c)包含至少一个人Vδ区段、至少一个人Dδ区段和至少一个人Jδ区段并可操作地连接至内源非人(例如,啮齿动物,例如小鼠或大鼠)TCRδ恒定区基因序列的未重排TCRδ可变基因座。本文所提供的另一非人动物在其基因组中包含(a)包含至少一个人Vα区段和至少一个人Jα区段并可操作地连接至内源非人(例如,啮齿动物,例如小鼠或大鼠)TCRα恒定基因序列的未重排T细胞受体(TCR)α可变基因座,(b)包含至少一个人Vβ区段、至少一个人Dβ区段和至少一个人Jβ区段并可操作地连接至内源非人(例如,啮齿动物,例如小鼠或大鼠)TCRβ恒定区基因序列的未重排TCRβ可变基因座,(c)包含至少一个人Vδ区段、至少一个人Dδ区段和至少一个人Jδ区段并可操作地连接至内源非人(例如,啮齿动物,例如小鼠或大鼠)TCRδ恒定区基因序列的未重排TCRδ可变基因座,和/或(d)包含至少一个人Vγ区段和至少一个人Jγ区段并可操作地连接至内源非人(例如,啮齿动物,例如小鼠或大鼠)TCRγ恒定区基因序列的未重排TCRγ可变基因座。

[0150] 在本发明的各种实施方案中,未重排人或人源化TCR可变基因座(例如TCRα、TCRβ和/或TCRδ可变基因座)包含在非人动物(例如,啮齿动物,例如小鼠或大鼠)的种系中。在各种实施方案中,TCR V(D)J区段替换为未重排人TCR V(D)J区段(例如,Vα和Jα;Vβ和Dβ和Jβ;Vδ和Dδ和Jδ;Vγ和Jγ区段)在一个(或多个)内源非人TCR可变基因座处进行,其中未重排人V和J和/或V和D和J区段可操作地连接至非人TCR恒定区基因序列。

[0151] 在本发明的一些实施方案中,非人动物包含未重排人或人源化TCRα可变基因座的两个拷贝、未重排人或人源化TCRβ可变基因座的两个拷贝和/或未重排人或人源化TCRδ可变基因座的两个拷贝。因此,非人动物对于一个或多个未重排人或人源化TCRα、TCRβ和/或TCRδ可变基因座是纯合的。在本发明的一些实施方案中,非人动物包含未重排人或人源化TCRα可变基因座的一个拷贝未重排人或人源化TCRβ可变基因座的一个拷贝和/或未重排人或人源化TCRδ可变基因座的一个拷贝。因此,非人动物对于未重排人或人源化TCRα、TCRβ和/或TCRδ可变基因座是杂合的。在其他实施方案中,非人动物对于未重排人或人源化TCRγ可变基因座是杂合的或纯合的。

[0152] 在一个实施方案中,包含人可变区区段(例如,人Vα和Jα区段)的未重排TCRα可变基因座定位在非人基因组中,使得人可变区区段替换对应的非人可变区区段。在一个实施方案中,包含人可变区区段的未重排TCRα可变基因座替换内源TCRα可变基因座。在一个方面,内源非人Vα和Jα区段不能重排形成重排的Vα/Jα序列。因此,在一个方面,未重排TCRα可变基因座中的人Vα和Jα区段能够重排形成重排的人Vα/Jα序列。



[0153] 相似地,在一个实施方案中,包含人可变区区段(例如,人V $\beta$ 、D $\beta$ 和J $\beta$ 区段)的未重排TCR $\beta$ 可变基因座定位在非人基因组中,使得人可变区区段替换对应的非人可变区区段。在一个实施方案中,包含人可变区区段的未重排TCR $\beta$ 可变基因座替换内源TCR $\beta$ 可变基因座。在一个方面,内源非人V $\beta$ 、D $\beta$ 和J $\beta$ 区段不能重排形成重排的V $\beta$ /D $\beta$ /J $\beta$ 序列。因此,在一个方面,未重排TCR $\beta$ 可变基因座中的人V $\beta$ 、D $\beta$ 和J $\beta$ 区段能够重排形成重排的人V $\beta$ /D $\beta$ /J $\beta$ 序列。

[0154] 在一个实施方案中,包含人可变区区段(例如,人V $\delta$ 、D $\delta$ 和J $\delta$ 区段)的未重排TCR $\delta$ 可变基因座定位在非人基因组中,使得人可变区区段替换对应的非人可变区区段。在一个实施方案中,包含人可变区区段的未重排TCR $\delta$ 可变基因座替换内源TCR $\delta$ 可变基因座。在一个方面,内源非人V $\delta$ 、D $\delta$ 和J $\delta$ 区段不能重排形成重排的V $\delta$ /D $\delta$ /J $\delta$ 序列。因此,在一个方面,未重排TCR $\delta$ 可变基因座中的人V $\delta$ 、D $\delta$ 和J $\delta$ 区段能够重排形成重排的人V $\delta$ /D $\delta$ /J $\delta$ 序列。

[0155] 在一个实施方案中,包含人可变区区段(例如,人V $\gamma$ 和J $\gamma$ 区段)的未重排TCR $\gamma$ 可变基因座定位在非人基因组中,使得人可变区区段替换对应的非人可变区区段。在一个实施方案中,包含人可变区区段的未重排TCR $\gamma$ 可变基因座替换内源TCR $\gamma$ 可变基因座。在一个方面,内源非人V $\alpha$ 和J $\alpha$ 区段不能重排形成重排的V $\gamma$ /J $\gamma$ 序列。因此,在一个方面,未重排TCR $\gamma$ 可变基因座中的人V $\gamma$ 和J $\gamma$ 区段能够重排形成重排的人V $\gamma$ /J $\gamma$ 序列。

[0156] 在又一个实施方案中,包含人可变区区段的未重排TCR $\alpha$ 、 $\beta$ 、 $\delta$ 和/或 $\gamma$ 可变基因座替换相应的内源TCR $\alpha$ 、 $\beta$ 、 $\delta$ 和 $\gamma$ 可变基因座。在一个方面,内源非人V $\alpha$ 和J $\alpha$ 区段不能重排形成重排的V $\alpha$ /J $\alpha$ 序列,内源非人V $\beta$ 、D $\beta$ 和J $\beta$ 区段不能重排形成重排的V $\beta$ /D $\beta$ /J $\beta$ 序列,内源V $\delta$ 、D $\delta$ 和J $\delta$ 区段不能重排形成重排的V $\delta$ /D $\delta$ /J $\delta$ 序列,和/或内源非人V $\gamma$ 和J $\gamma$ 区段不能重排形成重排的V $\gamma$ /J $\gamma$ 序列。因此,在一个方面,未重排TCR $\alpha$ 可变基因座中的人V $\alpha$ 和J $\alpha$ 区段能够重排形成重排的人V $\alpha$ /J $\alpha$ 序列,未重排TCR $\beta$ 可变基因座中的人V $\beta$ 、D $\beta$ 和J $\beta$ 区段能够重排形成重排的人V $\beta$ /D $\beta$ /J $\beta$ 序列,未重排TCR $\delta$ 可变基因座中的人V $\delta$ 、D $\delta$ 和J $\delta$ 区段能够重排形成重排的人V $\delta$ /D $\delta$ /J $\delta$ 序列,和/或未重排TCR $\alpha$ 可变基因座中的人V $\gamma$ 和J $\gamma$ 区段能够重排形成重排的人V $\gamma$ /J $\gamma$ 序列。

[0157] 在本发明的一些方面,包含人源化TCR $\alpha$ 、TCR $\beta$ 和/或TCR $\delta$ 基因座(包含未重排人TCR $\alpha$ 、TCR $\beta$ 和/或TCR $\delta$ 可变基因座)的非人动物保留内源非人TCR $\alpha$ 、TCR $\beta$ 和/或TCR $\delta$ 可变基因座。在一个实施方案中,内源非人TCR $\alpha$ 、TCR $\beta$ 和/或TCR $\delta$ 可变基因座是非功能性基因座。在一个实施方案中,非功能性基因座是失活的基因座,例如倒位的基因座(例如,可变基因座的编码核酸序列相对于恒定区序列处于倒位的取向,因此无法利用来自倒位的基因座的可变区区段进行成功的重排)。在一个实施方案中,人源化TCR $\alpha$ 、TCR $\beta$ 和/或TCR $\delta$ 可变基因座分别定位在内源非人TCR $\alpha$ 、TCR $\beta$ 和/或TCR $\delta$ 可变基因座与内源非人TCR $\alpha$ 、TCR $\beta$ 和/或TCR $\delta$ 恒定基因座之间。可以进行类似的染色体排列以便将人或人源化TCR $\gamma$ 放置在非人动物的基因组中,例如TCR $\gamma$ 基因座处。

[0158] 可以使用在国际免疫遗传学信息系统(IMG)的网站上公开的IMGT数据库确定人和小鼠TCR基因座的V和J和/或V、D和J区段的数量、命名、位置以及其他方面。小鼠TCR $\alpha$ 可变基因座为大约1.5兆碱基,并且总共包含110个V $\alpha$ 和60个J $\alpha$ 区段。人TCR $\alpha$ 可变基因座为大约1兆碱基,并且总共包含54个V $\alpha$ 和61个J $\alpha$ 区段,其中45个V $\alpha$ 和50个J $\alpha$ 被认为有功能。除非另有说明,否则整个说明书中提及的人V(D)J区段的数量是指V(D)J区段的总数。在本发明的一个实施方案中,经遗传修饰的非人动物(例如,啮齿动物,例如小鼠或大鼠)包含至少一个人

V $\alpha$ 和至少一个人J $\alpha$ 区段。在一个实施方案中,非人动物包含具有1、2、3、4、5、6、7、8、9、10、15、20、23、25、30、35、40、45、48、50个或至多54个人V $\alpha$ 区段的人源化TCR $\alpha$ 基因座。在一些实施方案中,人源化TCR $\alpha$ 基因座包含2、8、23、35、48或54个人V $\alpha$ 区段。因此,在一些实施方案中,非人动物中的人源化TCR $\alpha$ 基因座可以包含5%、10%、15%、20%、25%、30%、35%、40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%的人V $\alpha$ ;在一些实施方案中,其可以包含约2%、约3%、约15%、约65%、约90%或100%的人V $\alpha$ 。

[0159] 在一个实施方案中,非人动物包含人源化TCR $\alpha$ 基因座,该基因座包含具有人V $\alpha$ 40至V $\alpha$ 41(V $\alpha$ 区段也称为“TRAV”或“TCRAV”)的连续人序列的DNA片段,以及具有61个人J $\alpha$ 区段(J $\alpha$ 区段也称为“TRAJ”或“TCRAJ”)的连续人序列的DNA片段。在一个实施方案中,非人动物包含人源化TCR $\alpha$ 基因座,该基因座包含具有人TRAV35至TRAV41的连续人序列的DNA片段,以及具有61个人TRAJ的连续人序列的DNA片段。在一个实施方案中,非人动物包含人源化TCR $\alpha$ 基因座,该基因座包含具有人TRAV22至TRAV41的连续人序列的DNA片段,以及具有61个人TRAJ的连续人序列的DNA片段。在一个实施方案中,非人动物包含人源化TCR $\alpha$ 基因座,该基因座包含具有人TRAV13-2至TRAV41的连续人序列的DNA片段,以及具有61个人TRAJ的连续人序列的DNA片段。在一个实施方案中,非人动物包含人源化TCR $\alpha$ 基因座,该基因座包含具有人TRAV6至TRAV41和61个人TRAJ的连续人序列的DNA片段。在一个实施方案中,非人动物包含人源化TCR $\alpha$ 基因座,该基因座包含具有人TRAV1-1至TRAV 41和61个人TRAJ的连续人序列的DNA片段。在各种实施方案中,具有人TCR $\alpha$ 可变区区段的连续人序列的DNA片段还包含插入以有利于基因座人源化过程中的克隆和选择的限制性内切酶位点、选择盒、内切核酸酶位点或其他位点。在各种实施方案中,这些附加位点不会干扰TCR $\alpha$ 基因座处的各种基因的正常功能(例如,重排、剪接等)。

[0160] 在一个实施方案中,人源化TCR $\alpha$ 基因座包含61个人J $\alpha$ 区段或100%人J $\alpha$ 区段。在一个特定实施方案中,人源化TCR $\alpha$ 基因座包含8个人V $\alpha$ 区段和61个人J $\alpha$ 区段;在另一个特定实施方案中,人源化TCR $\alpha$ 基因座包含23个人V $\alpha$ 区段和61个人J $\alpha$ 区段。在另一个特定实施方案中,人源化TCR $\alpha$ 基因座包含人V $\alpha$ 和J $\alpha$ 区段的完整组库,即所有由 $\alpha$ 基因座编码的人可变 $\alpha$ 区基因区段,或54个人V $\alpha$ 和61个人J $\alpha$ 区段。在各种实施方案中,非人动物在TCR $\alpha$ 基因座处不包含任何内源非人V $\alpha$ 或J $\alpha$ 区段。

[0161] 小鼠TCR $\beta$ 可变基因座为大约0.6兆碱基,并且总共包含33个V $\beta$ 、2个D $\beta$ 和14个J $\beta$ 区段。人TCR $\beta$ 可变基因座为大约0.6兆碱基,并且总共包含67个V $\beta$ 、2个D $\beta$ 和14个J $\beta$ 区段。在本发明的一个实施方案中,经遗传修饰的非人动物(例如,啮齿动物,例如小鼠或大鼠)包含至少一个人V $\beta$ 、至少一个人D $\beta$ 和至少一个人J $\alpha$ 区段。在一个实施方案中,非人动物包含具有1、2、3、4、5、6、7、8、9、10、15、20、23、25、30、35、40、45、48、50、55、60个或至多67个人V $\beta$ 区段的人源化TCR $\beta$ 基因座。在一些实施方案中,人源化TCR $\beta$ 基因座包含8、14、40、66或67个人V $\beta$ 区段。因此,在一些实施方案中,非人动物中的人源化TCR $\beta$ 基因座可以包含5%、10%、15%、20%、25%、30%、35%、40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%的人V $\beta$ ;在一些实施方案中,其可以包含约20%、约60%、约15%、约98%或100%的人V $\beta$ 。

[0162] 在一个实施方案中,非人动物包含人源化TCR $\beta$ 基因座,该基因座包含具有人V $\beta$ 18

至VB29-1 (VB区段也称为“TRBV”或“TCRBV”)的连续人序列的DNA片段。在一个实施方案中,非人动物包含人源化TCRB基因座,该基因座包含具有人TRBV18至TRBV29-1的连续人序列的DNA片段,具有人DB1-JB1 (即,人DB1-JB1-1-JB1-6区段)的连续人序列的单独DNA片段,以及具有人DB2-JB2 (即,人DB2-JB2-1-JB2-7区段)的连续人序列的单独DNA片段。在一个实施方案中,非人动物包含人源化TCRB基因座,该基因座包含具有人TRBV6-5至TRBV29-1的连续人序列的DNA片段,具有人DB1-JB1 (即,人DB1-JB1-1-JB1-6区段)的连续人序列的单独DNA片段,以及具有人DB2-JB2 (即,人DB2-JB2-1-JB2-7区段)的连续人序列的单独DNA片段。在一个实施方案中,非人动物包含人源化TCRB基因座,该基因座包含具有人TRBV1至TRBV29-1的连续人序列的DNA片段,具有人DB1-JB1的连续人序列的单独DNA片段,以及具有人DB2-JB2的连续人序列的单独DNA片段。在一个实施方案中,非人动物包含人源化TCRB基因座,该基因座包含具有人TRBV1至TRBV29-1的连续人序列的DNA片段,具有人DB1-JB1的连续人序列的单独DNA片段,具有人DB2-JB2的连续人序列的单独DNA片段,以及具有人TRBV30的序列的单独DNA片段。在各种实施方案中,具有人TCRB可变区区段的连续人序列的DNA片段还包含插入以有利于基因座人源化过程中的克隆和选择的限制性内切酶位点、选择盒、内切核酸酶位点或其他位点。在各种实施方案中,这些附加位点不会干扰TCRB基因座处的各种基因的正常功能(例如,重排、剪接等)。

[0163] 在一个实施方案中,人源化TCRB基因座包含14个人JB区段或100%人JB区段,以及2个人DB区段或100%人JB区段。在另一个实施方案中,人源化TCRB基因座包含至少一个人VB区段,例如14个人VB区段,以及所有小鼠DB和JB区段。在一个特定实施方案中,人源化TCRB基因座包含14个人VB区段、2个人DB区段和14个人JB区段。在另一个特定实施方案中,人源化TCRB基因座包含人VB、DB和JB区段的完整组库,即所有由β基因座编码的人可变β区基因区段,或67个人VB、2个人DB和14个人JB区段。在一个实施方案中,非人动物在人源化TCRB基因座处包含一个(例如5')非人VB区段。在各种实施方案中,非人动物在TCRB基因座处不包含任何内源非人VB、DB或JB区段。

[0164] 在其中非人动物(例如,啮齿动物)包含人TCRα和TCRB (以及任选的人TCRδ和TCRγ)可变区区段的组库(例如,可变区区段的完整组库)的各种实施方案中,各个区段的组库(例如,各个区段的完整组库)被该动物用于产生针对各种抗原的TCR分子多样性组库。

[0165] 在各个方面,非人动物包含人基因组TCR可变基因座的连续部分,该基因座包含如未重排人基因组可变基因座中那样排列的V、D和J,或D和J,或V和J,或V区段,例如包含如人基因组TCR可变基因座中那样排列的启动子序列、前导序列、基因间序列、调控序列等。在其他方面,如未重排非人基因组TCR可变基因座中那样排列各个片段。在人源化TCRα、β、δ和/或γ基因座的各种实施方案中,人源化基因座可以包含两个或更多个人基因组区段,它们在人基因组中不出现并列,例如,在人基因组中位于恒定区附近的人可变基因座的V区段中的一个片段,与在人基因组中位于人可变基因座上游端的人可变基因座的V区段中的一个片段并列。

[0166] 在小鼠和人中,TCRδ基因区段与TCRα基因座位于一一起(参见图4A,顶部,加框的TCRD区)。TCRδJ和D区段位于Vα和Jα区段之间,而TCRδV区段散布在整个TCRα基因座中,大多数位于各个Vα区段之中。各个TCRδ区段的数量和位置可以由IMGT数据库确定。由于TCRα基因座内的TCRδ基因区段的基因组排列,TCRα基因座处的成功重排可使TCRδ基因区段缺失或

失活。

[0167] 在本发明的一些实施方案中,包含未重排人TCR $\alpha$ 可变基因座的非人动物还包含至少一个人V $\delta$ 区段,例如至多人V $\delta$ 区段的完整组库。因此,在一些实施方案中,内源TCR $\alpha$ 可变基因座的替换导致至少一个非人V $\delta$ 区段被替换为人V $\delta$ 区段。在其他实施方案中,本发明的非人动物在未重排人源化TCR $\alpha$ 基因座处包含人V $\delta$ 、D $\delta$ 和J $\delta$ 区段的完整组库;在其他实施方案中,非人动物在未重排人源化TCR $\alpha$ 基因座(即,包含人可变区区段以及人增强子和恒定区的TCR $\delta$ 基因座)处包含完整的未重排人TCR $\delta$ 基因座。美国专利No.9,113,616中描述了用于构建包含完整的未重排TCR $\delta$ 基因座的未重排人源化TCR $\alpha$ 基因座的示例性实施方案,该专利以引用方式并入本文。

[0168] 在又一个实施方案中,本发明的非人动物还包含未重排人源化TCR $\gamma$ 基因座,例如包含至少一个人V $\gamma$ 和至少一个人J $\gamma$ 区段(例如,人V $\gamma$ 和人J $\gamma$ 可变区区段的完整组库)的TCR $\gamma$ 基因座。人TCR $\gamma$ 基因座位于人7号染色体上,而小鼠TCR $\gamma$ 基因座位于小鼠13号染色体上。有关TCR $\gamma$ 基因座的更多详细信息,参见IMGT数据库。

[0169] 在一个方面,包含本文所述人源化TCR $\alpha$ 和 $\beta$ 可变基因座(以及任选的人源化TCR $\delta$ / $\gamma$ 可变基因座)的非人动物(例如,啮齿动物,例如小鼠或大鼠)在T细胞的表面上表达包含人可变区和非人(例如,啮齿动物,例如小鼠或大鼠)恒定区的人源化T细胞受体。在一些方面,非人动物能够表达识别多种递呈的抗原的人源化T细胞受体多样性组库。

[0170] 在本发明的各种实施方案中,本文所述的人源化T细胞受体多肽包含人前导序列。在替代实施方案中,对人源化TCR受体核酸序列进行工程改造,使得人源化TCR多肽包含非人前导序列。

[0171] 本文所述的人源化TCR多肽可以在内源非人调控元件(例如,啮齿动物调控元件)(例如,启动子、沉默子、增强子等)的控制下表达。本文所述的人源化TCR多肽也可以在人调控元件的控制下表达。在各种实施方案中,本文所述的非人动物还包含通常原位存在于人基因组中的所有调控序列和其他序列。

[0172] 在各种实施方案中,人源化TCR蛋白质的人可变区能够与相同细胞或另一细胞表面上的各种蛋白质相互作用。在一个实施方案中,人源化TCR的人可变区与在第二细胞(例如抗原递呈细胞(APC))的表面上递呈抗原的MHC蛋白质(例如,MHC I类或II类蛋白质)相互作用。在一些实施方案中,MHC I或II蛋白质是非人(例如,啮齿动物,例如小鼠或大鼠)蛋白质。在其他实施方案中,MHC I或II蛋白质是人(源化)蛋白质。在一个方面,第二细胞(例如APC)是表达人或人源化MHC分子的内源非人细胞。在不同的实施方案中,第二细胞是表达人MHC分子的人细胞。

[0173] 在一个方面,非人动物在T细胞表面上表达具有非人恒定区的人源化T细胞受体,其中所述受体能够与非人分子相互作用,例如在T细胞中表达的锚定分子或信号传导分子(例如,CD3分子、 $\zeta$ 链,或通过CD3分子或 $\zeta$ 链锚定于TCR的其他蛋白质)。因此,在一个方面,提供了一种细胞复合体,该细胞复合体包含(a)非人T细胞,其表达(i)包含如本文所述的人源化TCR $\alpha$ 链和如本文所述的人源化TCR $\beta$ 链的TCR以及(ii)如本文所述的嵌合辅助受体,和(b)非人抗原递呈细胞,其包含与如本文所述的嵌合MHCI和/或嵌合MHC II结合的抗原。在一个实施方案中,非人恒定TCR $\alpha$ 和TCR $\beta$ 链与非人zeta( $\zeta$ )链同源二聚体和CD3异源二聚体复合。在一个实施方案中,细胞复合体是体内细胞复合体。在一个实施方案中,细胞复合体是体外

细胞复合体。

[0174] 在各种实施方案中,本文所述的非人动物(例如,啮齿动物,例如小鼠或大鼠)产生能够经历胸腺发育、从DN1进展至DN2至DN3至DN4至DP以及至CD4或CD8SP T细胞的T细胞。本发明的非人动物的此类T细胞表达通常在胸腺发育的特定阶段期间由T细胞产生的细胞表面分子(例如,CD25、CD44、Kit、CD3、pTα等)。因此,在一个实施方案中,本文所述的非人动物可以在胸腺发育的DN3阶段表达与TCRβ复合的pTα。本文所述的非人动物表达能够经历胸腺发育而产生CD4+和CD8+T细胞的T细胞。

[0175] 在各种实施方案中,本文所述的非人动物产生能够在外周经历T细胞分化的T细胞。在一些实施方案中,本文所述的非人动物能够产生效应T细胞的组库,例如CTL(细胞毒性T淋巴细胞)、T<sub>H</sub>1、T<sub>H</sub>2、T<sub>REG</sub>、T<sub>H</sub>17等。因此,在这些实施方案中,本文所述的非人动物产生效应T细胞,其执行特定T细胞类型所特有的不同功能,例如识别、结合外来抗原并作出应答。在各种实施方案中,本文所述的非人动物产生效应T细胞,其杀死展示在MHC I分子背景下表达的细胞溶质病原体的肽片段的细胞;识别来源于细胞内囊泡中降解的抗原且由MHC II分子递呈于巨噬细胞表面上的肽,并诱导巨噬细胞杀死微生物;产生促进B细胞分化的细胞因子;活化B细胞以产生调理素化抗体;诱导上皮细胞产生将中性粒细胞募集到感染部位的趋化因子;等等。

[0176] 在另外的实施方案中,本文所述的非人动物包含外周中(例如脾中)的CD3+T细胞。在其他方面,本文所述的非人动物能够对目标抗原作出应答而产生记忆T细胞群。例如,非人动物产生针对抗原例如目标抗原(例如,被测试用于疫苗开发的抗原等)的中枢记忆T细胞(T<sub>cm</sub>)和效应记忆T细胞(T<sub>em</sub>)两者。

[0177] 不能接收足够信号(例如,Notch信号)的DN1和DN2细胞可发育成B细胞、骨髓细胞(例如,树突细胞)、肥大细胞和NK细胞。参见,例如,Yashiro-Ohtani et al. (2010) Notch regulation of early thymocyte development, *Seminars in Immunology* 22:261-69 (Yashiro-Ohtani等人,2010年,早期胸腺细胞发育的Notch调节,《免疫学研究文辑》,第22卷,第261-269页)。在一些实施方案中,本文所述的非人动物形成B细胞、骨髓细胞(例如树突细胞)、肥大细胞和NK细胞。在一些实施方案中,本文所述的非人动物在胸腺中形成树突细胞群。

[0178] 在T细胞表面上表达的T细胞受体的主要类型是TCRα/β,少数细胞表达TCRδ/γ。在本发明的一些实施方案中,包含人源化TCRα和/或β基因座的非人动物的T细胞表现出TCRα/β和TCRδ/γ基因座的利用,例如与野生型动物类似的TCRα/β和TCRδ/γ基因座的利用(例如,本文所述非人动物的T细胞以与野生型动物表达的蛋白质相当的比例表达TCRα/β和TCRδ/γ蛋白质)。因此,在一些实施方案中,包含人源化TCRα/β和内源非人TCRδ/γ基因座的非人动物表现出所有基因座的利用。

[0179] 人或人源化MHC分子

[0180] 在各种实施方案中,本文提供了经遗传修饰的非人动物,所述非人动物共表达至少一种人源化T细胞辅助受体、与人源化T细胞辅助受体结合的至少一种人源化MHC,以及任选的人源化TCR,所述人源化TCR在识别和结合由人源化MHC递呈的肽后,与人源化辅助受体联合向表达人源化TCR和嵌合T细胞辅助受体多肽的细胞提供活化信号。因此,如本文所公开的非人动物包含第一、第二和/或第三核酸序列中的至少一者,每个核酸序列编码不同的

人或人源化MHC多肽,所述人或人源化MHC多肽选自人或人源化MHC II $\alpha$ 多肽、人或人源化MHC II $\beta$ 多肽以及人或人源化MHC I $\alpha$ 多肽;该非人动物还任选地包含人或人源化 $\beta$ 2微球蛋白。本文中使用第一、第二和第三名称不应被解释为将本文所公开的非人动物限制为需要所有三个核酸序列或以任何特定顺序存在任何人或人源化MHC多肽。

[0181] 因此,在一些实施方案中,如本文所公开的非人动物可包含例如编码例如人或嵌合CD8 $\alpha$ 多肽和人或嵌合CD8 $\beta$ 多肽的第一和第二核苷酸序列,包含至少一个人V $\alpha$ 区段和至少一个人J $\alpha$ 区段并可操作地连接至非人TCR $\alpha$ 恒定基因序列的未重排T细胞受体(TCR) $\alpha$ 可变基因座,和/或包含至少一个人V $\beta$ 区段、至少一个人D $\beta$ 区段和至少一个人J $\beta$ 区段并可操作地连接至非人TCR $\beta$ 恒定基因序列的未重排TCR $\beta$ 可变基因座,以及任选地编码例如人或人源化MHC I $\alpha$ 多肽和人或人源化 $\beta$ 2-微球蛋白多肽的第一和第二核酸序列。在其他实施方案中,如本文所公开的非人动物可包含例如编码例如嵌合CD4多肽的第一核苷酸序列,包含至少一个人V $\alpha$ 区段和至少一个人J $\alpha$ 区段并可操作地连接至非人TCR $\alpha$ 恒定基因序列的未重排T细胞受体(TCR) $\alpha$ 可变基因座,和/或包含至少一个人V $\beta$ 区段、至少一个人D $\beta$ 区段和至少一个人J $\beta$ 区段并可操作地连接至非人TCR $\beta$ 恒定基因序列的未重排TCR $\beta$ 可变基因座;以及任选地编码例如人或人源化MHC II $\alpha$ 多肽和人或人源化MHC II $\beta$ 多肽的第一和第二核酸序列。在一些实施方案中,如本文所公开的非人动物可以包含例如编码例如嵌合CD4多肽、嵌合CD8 $\alpha$ 多肽和嵌合CD8 $\beta$ 多肽的第一、第二和第三核苷酸序列;包含至少一个人V $\alpha$ 区段和至少一个人J $\alpha$ 区段并可操作地连接至非人TCR $\alpha$ 恒定基因序列的未重排T细胞受体(TCR) $\alpha$ 可变基因座,和/或包含至少一个人V $\beta$ 区段、至少一个人D $\beta$ 区段和至少一个人J $\beta$ 区段并可操作地连接至非人TCR $\beta$ 恒定基因序列的未重排TCR $\beta$ 可变基因座;以及任选地编码例如人或人源化MHC II $\alpha$ 多肽、人源或人源化MHC II $\beta$ 多肽、人或人源化MHC I $\alpha$ 多肽和人或人源化 $\beta$ 2-微球蛋白多肽的第一、第二、第三和第四核酸序列。

[0182] 在各种实施方案中,本文提供了经遗传修饰的非人动物,例如在其基因组中包含编码人或人源化MHC I多肽的核酸序列和/或编码人或人源化MHC II蛋白质的核酸序列的啮齿动物(例如,小鼠或大鼠)。MHC I核酸序列可以编码部分人和部分非人的MHC I多肽,例如嵌合人/非人MHC I多肽,并且MHC II核酸序列可以编码部分人和部分非人的MHC II蛋白质,例如嵌合人/非人MHC II蛋白质(例如,包含嵌合人/非人MHC II $\alpha$ 和 $\beta$ 多肽)。在一些方面,该动物不在细胞表面上表达内源MHC I和/或内源MHC II多肽,例如功能性内源MHC I和/或MHC II多肽。在一些实施方案中,在该动物的细胞表面上表达的唯一MHC I和/或MHC II分子是嵌合MHC I和/或MHC II分子。

[0183] 美国专利公布No.20130111617和20130185819中公开了一种经遗传修饰的非人动物,该非人动物在其基因组中例如在内源基因座处包含编码嵌合人/非人MHC I多肽的核酸序列,所述专利公布全文以引用方式并入本文。美国专利No.8,847,005和美国专利公布No.20130185820中公开了一种经遗传修饰的非人动物,该非人动物在其基因组中例如在内源基因座处包含编码人源化例如嵌合人/非人MHC II多肽的核酸序列,所述专利和专利公布全文均以引用方式并入本文。美国专利公布No.20140245467中公开了一种经遗传修饰的非人动物,该非人动物在其基因组中例如在内源基因座处包含编码嵌合人/非人MHC I多肽的核酸序列,并且在其基因组中例如在内源基因座处包含编码人源化例如嵌合人/非人MHC II多肽的核酸序列,所述专利公布全文以引用方式并入本文。

[0184] 在各种实施方案中,本文提供了经遗传修饰的非人动物,该非人动物在其基因组中例如在一个或多个内源MHC基因座处包含编码嵌合人/非人MHC I多肽的第一核酸序列,其中嵌合MHC I多肽的人部分包含人MHC I多肽的细胞外部分(或其一部分,例如一个或多个细胞外结构域);编码嵌合人/非人MHC II $\alpha$ 多肽的第二核酸序列,其中嵌合MHC II $\alpha$ 多肽的人部分包含人MHC II $\alpha$ 多肽的细胞外部分(或其一部分,例如一个或多个细胞外结构域);和/或编码嵌合人/非人MHC II $\beta$ 多肽的第三核酸序列,其中嵌合MHC II $\beta$ 多肽的人部分包含人MHC II $\beta$ 多肽的细胞外部分(或其一部分,例如一个或多个细胞外结构域);其中该非人动物从其内源非人MHC基因座表达功能性嵌合人/非人MHC I和MHC II蛋白质。在一个实施方案中,第一、第二和/或第三核酸序列分别位于内源非人MHC I、MHC II $\alpha$ 和MHC II $\beta$ 基因座。在其中非人动物是小鼠的一个实施方案中,第一、第二和/或第三核酸序列位于小鼠17号染色体上的内源小鼠MHC基因座。在一个实施方案中,第一核酸序列位于内源非人MHC I基因座。在一个实施方案中,第二核酸序列位于内源非人MHC II $\alpha$ 基因座。在一个实施方案中,第三核酸序列位于内源非人MHC II $\beta$ 基因座。

[0185] 在一个实施方案中,非人动物仅从内源非人MHC基因座表达嵌合人/非人MHC I、MHC II $\alpha$ 和/或MHC II $\beta$ 多肽,并且不表达内源非人MHC多肽(例如,功能性内源MHC I、II $\alpha$ 和/或II $\beta$ 多肽)。在一个实施方案中,本文所述的动物在其细胞例如抗原递呈细胞等的表面上表达功能性嵌合MHC I和功能性嵌合MHC II。在一个实施方案中,该动物在细胞表面上表达的唯一MHC I和MHC II是嵌合MHC I和嵌合MHC II,并且该动物在细胞表面上不表达任何内源MHC I和MHC II。

[0186] 在一个实施方案中,嵌合人/非人MHC I多肽在其人部分中包含例如人MHC I多肽的肽结合槽。在一个方面,嵌合多肽的人部分包含人MHC I的细胞外部分。在该实施方案中,嵌合多肽的人部分包含人MHC I的 $\alpha$ 链的细胞外结构域。在一个实施方案中,嵌合多肽的人部分包含人MHC I的 $\alpha 1$ 和 $\alpha 2$ 结构域。在另一个实施方案中,嵌合多肽的人部分包含人MHC I的 $\alpha 1$ 、 $\alpha 2$ 和 $\alpha 3$ 结构域。

[0187] 在一个方面,嵌合MHC II $\alpha$ 多肽的人部分和/或嵌合MHC II $\beta$ 多肽的人部分分别包含人MHC II $\alpha$ 多肽和/或人MHC II $\beta$ 多肽的肽结合结构域。在一个方面,嵌合MHC II $\alpha$ 和/或 $\beta$ 多肽的人部分分别包含人MHC II $\alpha$ 和/或 $\beta$ 多肽的细胞外部分。在一个实施方案中,嵌合MHC II $\alpha$ 多肽的人部分包含人MHC II $\alpha$ 多肽的 $\alpha 1$ 结构域;在另一个实施方案中,嵌合MHC II $\alpha$ 多肽的人部分包含人MHC II $\alpha$ 多肽的 $\alpha 1$ 和 $\alpha 2$ 结构域。在另外的实施方案中,嵌合MHC II $\beta$ 多肽的人部分包含人MHC II $\beta$ 多肽的 $\beta 1$ 结构域;在另一个实施方案中,嵌合MHC II $\beta$ 多肽的人部分包含人MHC II $\beta$ 多肽的 $\beta 1$ 和 $\beta 2$ 结构域。

[0188] 在一些实施方案中,人或人源化MHC I多肽可以来源于由HLA-A、HLA-B、HLA-C、HLA-E、HLA-F或HLA-G基因座中的任何一者编码的功能性人HLA分子。人或人源化MHC II多肽可以来源于由HLA-DP、-DQ和-DR基因座编码的功能性人HLA分子。常用的HLA抗原和等位基因的列表在Shankarkumar等人((2004) The Human Leukocyte Antigen (HLA) System, Int. J. Hum. Genet. 4 (2): 91-103 (2004年,人白细胞抗原(HLA)系统,《国际人类遗传学杂志》,第4卷,第2期,第91-103页))中有所描述,该文献以引用方式并入本文Shankarkumar等人还给出了本领域中使用的HLA命名的简要说明。关于HLA命名和各种HLA等位基因的另外信息可见于Holdsworth et al. (2009) The HLA dictionary 2008: a summary of HLA-A,-

B,-C,-DRB1/3/4/5,and DQB1alleles and their association with serologically defined HLA-A,-B,-C,-DR,and-DQ antigens,Tissue Antigens 73:95-170 (Holdsworth 等人,2009年,HLA字典2008:HLA-A,-B,-C,-DRB1/3/4/5和DQB1等位基因及其与血清学鉴定 HLA-A,-B,-C,DR和-DQ抗原的关联的总结,《组织抗原》,第73卷,第95-170页),以及Marsh et al. (2010) Nomenclature for factors of the HLA system,2010,Tissue Antigens 75:291-455 (Marsh等人,2010年,HLA系统各因子的命名,2010,《组织抗原》,第75卷,第291-455页)最近的更新,这两篇文献均以引用的方式并入本文。在一些实施方案中,MHC I或MHC II多肽可以来源于任何功能性人HLA-A、B、C、DR或DQ分子。因此,人或人源化MHC I和/或II多肽可以来源于其中所述的任何功能性人HLA分子。在一些实施方案中,在细胞表面上表达的所有MHC I和MHC II多肽包含来源于人HLA分子的部分。

[0189] 特别要关注的是人HLA分子,已知特异性多态性HLA等位基因与许多人疾病例如人自身免疫疾病相关。事实上,已经鉴定出HLA基因座中的特异性多态性与类风湿性关节炎、I型糖尿病、桥本氏甲状腺炎、多发性硬化症、重症肌无力、格雷夫斯病、系统性红斑狼疮、乳糜泻、克罗恩病、溃疡性结肠炎和其他自身免疫疾病相关。参见例如Wong and Wen (2004) What can the HLA transgenic mouse tell us about autoimmune diabetes?, Diabetologia 47:1476-87 (Wong和Wen,2004年,HLA转基因小鼠能够告诉我们关于自身免疫糖尿病的什么信息?《糖尿病学》,第47卷,第1476-1487页);Taneja and David (1998) HLA Transgenic Mice as Humanized Mouse Models of Disease and Immunity, J.Clin.Invest.101:921-26 (Taneja和David,1998年,作为疾病和免疫力的人源化小鼠模型的HLA转基因小鼠,《临床研究杂志》,第101卷,第921-926页);Bakker et al. (2006) ,A high-resolution HLA and SNP haplotype map for disease association studies in the extended human MHC,Nature Genetics 38:1166-72 (Bakker等人,2006年,用于在延伸的人MHC中进行疾病相关性研究的高分辨率HLA和SNP单体型图谱,第38卷,第1166-1172页)和补充信息;以及International MHC and Autoimmunity Genetics Network (2009) Mapping of multiple susceptibility variants within the MHC region for 7immune-mediated diseases,Proc.Natl.Acad.Sci.USA 106:18680-85 (国际MHC和自身免疫性遗传学网络,2009年,MHC区内定位7种免疫介导性疾病的多种易感变体,《美国国家科学院院刊》,第106卷,第18680-18685页)。因此,人或人源化MHC I和/或II多肽可以来源于已知与特定疾病例如自身免疫疾病相关的人HLA分子。

[0190] 在一个具体方面,人或人源化MHC I多肽来源于人HLA-A。在一个具体实施方案中,HLA-A多肽是HLA-A2多肽(例如和HLA-A2.1多肽)。在一个实施方案中,HLA-A多肽是由HLA-A\*0201等位基因例如HLA-A\*02:01:01:01等位基因编码的多肽。HLA-A\*0201等位基因常在北美人群中使用。虽然本发明实施例描述了该特定HLA序列,但是本文涵盖任何合适的HLA-A序列,例如在人群中显示的HLA-A2的多态性变体、具有一个或多个保守或非保守氨基酸修饰的序列、由于遗传密码的简并性而与本文所述序列有所不同的核酸序列等。

[0191] 在另一具体方面,嵌合MHC I多肽的人部分来源于选自HLA-B和HLA-C的人MHC I。在一个方面,其来源于HLA-B,例如HLA-B27。在另一方面,其来源于HLA-A3、-B7、-Cw6等。

[0192] 在一个具体方面,本文所述人源化MHC II $\alpha$ 和 $\beta$ 多肽的人部分来源于人HLA-DR,例如HLA-DR2。通常,HLA-DR $\alpha$ 链是单态的,例如,HLA-DR复合体的 $\alpha$ 链由HLA-DRA基因(例如,



HLA-DR $\alpha$ \*01基因)编码。另一方面,HLA-DR $\beta$ 链是多态的。因此,HLA-DR2包含由HLA-DRA基因编码的 $\alpha$ 链以及由HLA-DR1 $\beta$ \*1501基因编码的 $\beta$ 链。虽然本发明实施例描述了这些特定HLA序列,但是本文涵盖任何合适的HLA-DR序列,例如在人群中显示的多态性变体、具有一个或多个保守或非保守氨基酸修饰的序列、由于遗传密码的简并性而与本文所述序列有所不同的核酸序列等。

[0193] 嵌合MHC II $\alpha$ 和/或 $\beta$ 多肽的人部分可由已知与常见人疾病相关的HLA等位基因的核酸序列编码。这样的HLA等位基因包括但不限于HLA-DRB1\*0401、-DRB1\*0301、-DQA1\*0501、-DQB1\*0201、DRB1\*1501、-DRB1\*1502、-DQB1\*0602、-DQA1\*0102、-DQA1\*0201、-DQB1\*0202、-DQA1\*0501以及它们的组合。关于HLA等位基因/疾病相关性的总结,参见Bakker等人(2006),出处同上,该文献以引用方式并入本文。

[0194] 在一个方面,嵌合人/非人MHC I、MHC II $\alpha$ 和/或MHC II $\beta$ 多肽的非人部分分别包含内源非人(例如啮齿动物,例如小鼠、大鼠等)MHC I、MHC II $\alpha$ 和/或MHC II $\beta$ 多肽的跨膜和/或胞质结构域。因此,嵌合人/非人MHC I多肽的非人部分可以包含内源非人MHC I多肽的跨膜和/或胞质结构域。嵌合MHC II $\alpha$ 多肽的非人部分可以包含内源非人MHC II $\alpha$ 多肽的跨膜和/或胞质结构域。嵌合人/非人MHC II $\beta$ 多肽的非人部分可以包含内源非人MHC II $\beta$ 多肽的跨膜和/或胞质结构域。在一个方面,该非人动物是小鼠,并且嵌合MHC I多肽的非人部分来源于小鼠H-2K蛋白质。在一个方面,该动物是小鼠,并且嵌合MHC II $\alpha$ 和 $\beta$ 多肽的非人部分来源于小鼠H-2E蛋白质。因此,嵌合MHC I多肽的非人部分可以包含来源于小鼠H-2K的跨膜结构域和胞质结构域,并且嵌合MHC II $\alpha$ 和 $\beta$ 多肽的非人部分可以包含来源于小鼠H-2E蛋白质的跨膜结构域和胞质结构域。虽然在实施例中考了具体的H-2K和H-2E序列,但是本文涵盖任何合适的序列,例如多态变体、保守/非保守氨基酸取代等。在一个方面,非人动物是小鼠,并且该小鼠不从其H-2D基因座表达功能性内源MHC多肽。在一些实施方案中,该小鼠经工程改造成缺少内源H-2D基因座的全部或一部分。在其他方面,该小鼠在细胞表面上不表达任何功能性内源小鼠MHC I和MHC II。

[0195] 嵌合人/非人多肽可以使得其包含人或非人前导(信号)序列。在一个实施方案中,嵌合MHC I多肽包含内源MHC I多肽的非人前导序列。在一个实施方案中,嵌合MHC II $\alpha$ 多肽包含内源MHC II $\alpha$ 多肽的非人前导序列。在一个实施方案中,嵌合MHC II $\beta$ 多肽包含内源MHC II $\beta$ 多肽的非人前导序列。在替代实施方案中,嵌合MHC I、MHC II $\alpha$ 和/或MHC II $\beta$ 多肽分别包含来自另一非人动物例如另一啮齿动物或另一小鼠品系的MHC I、MHC II $\alpha$ 和/或MHC II $\beta$ 多肽的非人前导序列。因此,编码嵌合MHC I、MHC II $\alpha$ 和/或MHC II $\beta$ 多肽的核酸序列可以分别可操作地连接至编码非人MHC I、MHC II $\alpha$ 和/或MHC II $\beta$ 前导序列的核酸序列。在又一个实施方案中,嵌合MHC I、MHC II $\alpha$ 和/或MHC II $\beta$ 多肽分别包含人MHC I、人MHC II $\alpha$ 和/或人MHC II $\beta$ 多肽的人前导序列(例如,分别包含人HLA-A2、人HLA-DR $\alpha$ 和/或人HLA-DR $\beta$ 1\*1501的前导序列)。

[0196] 嵌合人/非人MHC I、MHC II $\alpha$ 和/或MHC II $\beta$ 多肽可分别在其人部分中包含人MHC I、人MHC II $\alpha$ 和/或人MHC II $\beta$ 多肽的完整或基本上完整的细胞外结构域。因此,人部分可以包含编码人MHC I、人MHC II $\alpha$ 和/或人MHC II $\beta$ 多肽(例如,人HLA-A2、人HLA-DR $\alpha$ 和/或人HLA-DR $\beta$ 1\*1501)的细胞外结构域的氨基酸的至少80%、优选至少85%、更优选至少90%、例如95%或更多。在一个示例中,人MHC I、人MHC II $\alpha$ 和/或人MHC II $\beta$ 多肽的基本上完整的细

胞外结构域缺少人前导序列。在另一个示例中,嵌合人/非人MHC I、嵌合人/非人MHC II $\alpha$ 和/或嵌合人/非人MHC II $\beta$ 多肽包含人前导序列。

[0197] 此外,嵌合MHC I、MHC II $\alpha$ 和/或MHC II $\beta$ 多肽可以分别可操作地连接至内源非人启动子和调控元件例如小鼠MHC I、MHC II $\alpha$ 和/或MHC II $\beta$ 调控元件(例如,可在这些元件的调控下表达)。这种排列将促进嵌合MHC I和/或MHC II多肽在非人动物中的适当表达,例如在非人动物中的免疫应答期间。

[0198] 在另一实施方案中,本发明的非人动物例如啮齿动物如小鼠包含(例如,在内源 $\beta$ 2微球蛋白基因座处)编码人或人源化 $\beta$ 2微球蛋白的核酸序列。 $\beta$ 2微球蛋白或MHC I类复合体的轻链(也缩写为“ $\beta$ 2M”)是小分子(12kDa)非糖基化蛋白质,其主要起到稳定MHC I $\alpha$ 链的作用。人或人源化 $\beta$ 2微球蛋白动物的产生在美国专利公布No.20130111617中详细描述,并且以引用的方式并入本文。

[0199] 编码人或人源化 $\beta$ 2微球蛋白多肽的核苷酸序列可以包含对应于整个人 $\beta$ 2微球蛋白基因的核酸残基。或者,该核苷酸序列可以包含编码人 $\beta$ 2微球蛋白蛋白质的第21-119位氨基酸(即,对应于成熟人 $\beta$ 2微球蛋白的氨基酸残基)所示的氨基酸序列的核酸残基。在替代实施方案中,该核苷酸序列可以包含编码人 $\beta$ 2微球蛋白蛋白质的第23-115位氨基酸所示的氨基酸序列(例如人 $\beta$ 2微球蛋白蛋白质的第23-119位氨基酸所示的氨基酸序列)的核酸残基。人 $\beta$ 2微球蛋白的核酸和氨基酸序列描述于Gussow等人,出处同上,该文献以引用方式并入本文。

[0200] 因此,人或人源化 $\beta$ 2微球蛋白多肽可以包含人 $\beta$ 2微球蛋白多肽的第23-115位氨基酸所示的氨基酸序列,例如人 $\beta$ 2微球蛋白多肽的第23-119位氨基酸所示的氨基酸序列,例如人 $\beta$ 2微球蛋白多肽的第21-119位氨基酸所示的氨基酸序列。或者,人 $\beta$ 2微球蛋白可以包含人 $\beta$ 2微球蛋白多肽的第1-119位氨基酸。

[0201] 在一些实施方案中,编码人或人源化 $\beta$ 2微球蛋白的核苷酸序列包含人 $\beta$ 2微球蛋白基因的外显子2至外显子4中所示的核苷酸序列。或者,该核苷酸序列包含人 $\beta$ 2微球蛋白基因的外显子2、3和4中所示的核苷酸序列。在该实施方案中,外显子2、3和4中所示的核苷酸序列可操作地连接以允许基因的正常转录和翻译。因此,在一个实施方案中,该人序列包含对应于人 $\beta$ 2微球蛋白基因的外显子2至外显子4的核苷酸序列。在一个具体实施方案中,该人序列包含对应于人 $\beta$ 2微球蛋白基因的外显子2至在外显子4之后的约267bp的核苷酸序列。在一个具体实施方案中,该人序列包含约2.8kb的人 $\beta$ 2微球蛋白基因。

[0202] 因此,人或人源化 $\beta$ 2微球蛋白多肽可以由核苷酸序列编码,所述核苷酸序列包含人 $\beta$ 2微球蛋白的外显子2至外显子4中所示的核苷酸序列,例如对应于人 $\beta$ 2微球蛋白基因的外显子2至外显子4的核苷酸序列。或者,所述多肽可以由包含人 $\beta$ 2微球蛋白基因的外显子2、3和4中所示的核苷酸序列的核苷酸序列编码。在一个具体实施方案中,人或人源化 $\beta$ 2微球蛋白多肽由对应于人 $\beta$ 2微球蛋白基因的外显子2至在外显子4之后的约267bp的核苷酸序列编码。在另一个具体实施方案中,人或人源化多肽由包含约2.8kb的人 $\beta$ 2微球蛋白基因的核苷酸序列编码。由于 $\beta$ 2微球蛋白基因的外显子4含有5'非翻译区,人或人源化多肽可以由包含 $\beta$ 2微球蛋白基因的外显子2和3的核苷酸序列编码。

[0203] 本领域普通技术人员应当理解,虽然本文描述了用于产生经遗传工程改造的动物的特定核酸和氨基酸序列,但是还提供了一个或多个保守或非保守氨基酸取代的序列或由

于遗传密码的简并性而与本文所述序列不同的序列。

[0204] 因此,提供了表达人 $\beta 2$ 微球蛋白序列的非人动物,其中 $\beta 2$ 微球蛋白序列与人 $\beta 2$ 微球蛋白序列具有至少约85%、90%、95%、96%、97%、98%或99%同一性。在一个具体实施方案中, $\beta 2$ 微球蛋白序列与本文所述的人 $\beta 2$ 微球蛋白序列具有至少约90%、95%、96%、97%、98%或99%同一性。在一个实施方案中,人 $\beta 2$ 微球蛋白序列包含一个或多个保守取代。在一个实施方案中,人 $\beta 2$ 微球蛋白序列包含一个或多个非保守取代。

[0205] 此外,还提供了非人动物,其中编码人或人源化 $\beta 2$ 微球蛋白蛋白质的核苷酸序列还包含非人 $\beta 2$ 微球蛋白基因的外显子1中所示的核苷酸序列。因此,在一个具体实施方案中,非人动物在其基因组中包含编码人或人源化 $\beta 2$ 微球蛋白的核苷酸序列,其中所述核苷酸序列包含非人 $\beta 2$ 微球蛋白的外显子1以及人 $\beta 2$ 微球蛋白基因的外显子2、3和4。因此,人或人源化 $\beta 2$ 微球蛋白多肽由非人 $\beta 2$ 微球蛋白基因的外显子1以及人 $\beta 2$ 微球蛋白基因的外显子2、3和4(例如,人 $\beta 2$ 微球蛋白基因的外显子2和3)编码。

[0206] 在一个实施方案中,除了编码嵌合CD8蛋白质的核苷酸序列之外,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 I蛋白质的核酸序列,使得在该动物的T细胞表面上表达的嵌合CD8蛋白质能够与在第二细胞例如抗原递呈细胞的表面上表达的人或人源化MHC I结合、结合和/或相互作用。在一个实施方案中,MHC I蛋白质包含人MHC I多肽的细胞外结构域。在一个实施方案中,该动物还包含人或人源化 $\beta 2$ 微球蛋白多肽。美国专利公布No.20130111617和20130185819中描述了表达人或人源化MHC I多肽和/或 $\beta 2$ 微球蛋白多肽的示例性经遗传修饰的动物,这两份专利公布全文均以引用方式并入本文。因此,在一个实施方案中,包含本文所述嵌合CD8蛋白质的动物还可包含人源化MHC I复合体,其中所述人源化MHC I复合体包含:(1)人源化MHC I多肽,例如,其中人源化MHC I多肽包含人MHC I细胞外结构域以及内源(例如,小鼠)MHC I的跨膜结构域和胞质结构域,例如,其中人源化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 $\beta 2$ 微球蛋白多肽(例如,该动物在其基因组中包含人 $\beta 2$ 微球蛋白的外显子2、3和4中所示的核苷酸序列。在一个方面,人源化MHC I和人或人源化 $\beta 2$ 微球蛋白多肽分别由位于内源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,该动物不表达功能性内源MHC I和 $\beta 2$ 微球蛋白多肽。因此,由这些动物表达的MHC I可以是嵌合人/非人例如人/啮齿动物(例如,人/小鼠)MHC I多肽。嵌合MHC I多肽的人部分可以来源于人HLA I类蛋白质,该人HLA I类蛋白质选自HLA-A、HLA-B和HLA-C,例如HLA-A2、HLA-B27、HLA-B7、HLA-Cw6或存在于人群中的任何其他HLA I类分子。在其中动物是小鼠的实施方案中,嵌合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蛋白质。

[0207] 在一个实施方案中,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蛋白质的核苷酸序列,使得在该动物的T细胞表面上表达的嵌合CD4蛋白质能够与在第二细胞例如抗原递呈细胞的表面上表达的人或人源化MHC II相互作用。在一个实施方案中,MHC II蛋白质包含人MHC II $\alpha$ 多肽的细胞外结构域和人MHC II $\beta$ 多肽的细胞外结构域。表达人或人源化MHC II多肽的示例性经遗传修饰的动物描述于2014年9月30日发布的美国专利No.8,847,005和美国专利公布No.20130185820中,这两份专利全文以引用方式并入本文。因此,在一个实施方案中,包含本文所述嵌合CD4蛋白质的动物还可包含人源化MHC II蛋白质,其中所述人源化MHC II蛋白质包含:(1)包含人MHC II $\alpha$ 细胞外结构域以

及内源例如小鼠MHC II的跨膜结构域和胞质结构域的人源化MHC II $\alpha$ 多肽,其中人MHC II $\alpha$ 细胞外结构域包含人MHC II $\alpha$ 的 $\alpha$ 1和 $\alpha$ 2结构域,和(2)包含人MHC IIB细胞外结构域以及内源例如小鼠MHC II的跨膜结构域和胞质结构域的人源化MHC IIB多肽,其中人MHC IIB细胞外结构域包含人MHC IIB的 $\beta$ 1和 $\beta$ 2结构域。在一个方面,人源化MHC II $\alpha$ 和 $\beta$ 多肽分别由位于内源MHC II $\alpha$ 和 $\beta$ 基因座的核酸序列编码;在一个方面,该动物不表达功能性内源MHC II $\alpha$ 和 $\beta$ 多肽。因此,由这些动物表达的MHC II可以是嵌合人/非人例如人/啮齿动物(例如,人/小鼠)MHC II蛋白质。嵌合MHC II蛋白质的人部分可以来源于人HLA II类蛋白质,该人HLA II类蛋白质选自HLA-DR、HLA-DQ和HLA-DP,例如HLA-DR4、HLA-DR2、HLA-DQ2.5、HLA-DQ8或存在于人群中的任何其他HLA II类分子。在其中动物是小鼠的实施方案中,嵌合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蛋白质。

[0208] 根据本公开内容以及美国专利公布No.20130111617、20130185819和20130185820及美国专利No.8,847,005的公开内容,经遗传修饰的非人动物(例如,啮齿动物,例如大鼠或小鼠)的各种其他实施方案对于本领域技术人员将是显而易见的,这些专利以引用方式并入本文。

[0209] 在各种实施方案中,本文所述经遗传修饰的非人动物产生在细胞表面上具有人或人源化MHC I和II的细胞例如APC,从而以类似人类的方式将肽递呈为T细胞的表位,这是因为复合体的基本上所有组分都是人或人源化的。本发明的经遗传修饰的非人动物可用于研究人源化动物中的人免疫系统的功能;用于鉴定引起免疫应答的抗原和抗原表位(例如,T细胞表位,例如独特人癌表位),例如用于开发疫苗;用于评估候选疫苗和其他疫苗策略;用于研究人自身免疫;用于研究人类传染病;也用于基于人MHC表达来制定更好的治疗策略。

#### [0210] 非人动物、组织和细胞

[0211] 本发明的经遗传修饰的非人动物可以选自小鼠、大鼠、兔、猪、牛(例如母牛、公牛、水牛)、鹿、绵羊、山羊、鸡、猫、狗、白鼬、灵长类动物(例如,狨猴、恒河猴)。对于不易获得合适的可遗传修饰的ES细胞的非人动物而言,采用其他方法制备包含所述遗传修饰的非人动物。此类方法包括例如对非ES细胞基因组(例如,成纤维细胞或诱导多能细胞)进行修饰并采用核移植将经修饰的基因组转移到合适的细胞,例如卵母细胞,并且在合适的条件下在非人动物中孕育经修饰的细胞(例如,经修饰的卵母细胞)以形成胚胎。

[0212] 在一个方面,非人动物是哺乳动物。在一个方面,非人动物是例如跳鼠总科(Dipodoidea)或鼠总科(Muroidea)的小型哺乳动物。在一个实施方案中,经遗传修饰的动物是啮齿动物。在一个实施方案中,啮齿动物选自小鼠、大鼠和仓鼠。在一个实施方案中,啮齿动物选自鼠总科。在一个实施方案中,经遗传修饰的动物来自选自以下项的科:丽仓鼠科(Calomyscidae)(例如,丽仓鼠)、仓鼠科(Cricetidae)(例如,仓鼠、新世界鼠(New World rats and mice)、田鼠)、鼠科(Muridae)(例如,真小鼠和大鼠(true mice and rats)、沙鼠、非洲刺毛鼠、冠鼠)、马岛鼠科(Nesomyidae)(例如,攀鼠、岩鼠、白尾鼠、马岛鼠(Malagasy rats and mice))、刺睡鼠科(Platacanthomyidae)(例如,刺睡鼠)以及鼯形鼠科(Spalacidae)(例如,鼯形鼠、竹鼠和鼯鼠)。在一个具体实施方案中,经遗传修饰的啮齿动物选自真小鼠或大鼠(鼠科)、沙鼠、非洲刺毛鼠和冠鼠。在一个实施方案中,经遗传修饰的小鼠来自鼠科的成员。在一个实施方案中,动物是啮齿动物。在一个具体实施方案中,啮齿动物选自小鼠和大鼠。在一个实施方案中,非人动物为小鼠。

[0213] 在一个具体实施方案中,非人动物是啮齿动物,该啮齿动物为C57BL品系小鼠,其选自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。在另一个实施方案中,小鼠是129品系,其选自为129P1、129P2、129P3、129X1、129S1 (例如,129S1/SV、129S1/SvIm)、129S2、129S4、129S5、129S9/SvEvH、129S6 (129/SvEvTac)、129S7、129S8、129T1、129T2的品系 (参见例如,Festing et al. (1999) Revised nomenclature for strain 129mice, Mammalian Genome 10:836 (Festing等人,1999年,“品系129小鼠的修订命名法”,《哺乳动物基因组》,第10卷,第836页),还可参见,Auerbach et al (2000) Establishment and Chimera Analysis of 129/SvEv-and C57BL/6-Derived Mouse Embryonic Stem Cell Lines (Auerbach等人,2000年,129/SvEv和C57BL/6来源小鼠胚胎干细胞系的建立和嵌合体分析))。在一个实施方案中,经遗传修饰的小鼠为上述129品系与上述C57BL/6品系的混合型。在另一个具体实施方案中,小鼠为上述129品系的混合型,或上述BL/6品系的混合型。在一个具体实施方案中,混合型的129品系为129S6 (129/SvEvTac) 品系。在另一个实施方案中,小鼠为BALB品系,例如BALB/c品系。在又一个实施方案中,小鼠为BALB品系和另一种上述品系的混合型。本文所提供的非人动物可以是来源于上述品系的任何组合的小鼠。

[0214] 在一个实施方案中,非人动物为大鼠。在一个实施方案中,大鼠选自Wistar大鼠、LEA品系、Sprague Dawley品系、Fischer品系、F344、F6和Dark Agouti。在一个实施方案中,大鼠品系为选自Wistar、LEA、Sprague Dawley、Fischer、F344、F6和Dark Agouti中的两种或更多种品系的混合型。

[0215] 因此,在本发明的一个实施方案中,提供了经遗传修饰的小鼠,其中该小鼠例如在其基因组中例如在其种系基因组中包含 (a) 编码第一嵌合人/鼠T细胞辅助受体多肽 (例如,CD4) 的第一核苷酸序列、编码第二嵌合人/鼠T细胞辅助受体多肽 (例如,CD8 $\alpha$ ) 的第二核苷酸序列、和/或编码第三嵌合人/鼠T细胞辅助受体多肽 (例如,CD8 $\beta$ ) 的第三核苷酸序列,其中每种嵌合T细胞辅助受体多肽的鼠部分包含鼠T细胞辅助受体的至少跨膜结构域和胞质结构域,其中每种嵌合多肽的人部分包含人T细胞辅助受体的细胞外部分 (或其一部分,例如一个或多个细胞外结构域),并且其中该小鼠表达第一、第二和/或第三嵌合T细胞辅助受体多肽; (b) 包含至少一个人V $\alpha$ 区段和至少一个人J $\alpha$ 区段并可操作地连接至鼠TCR $\alpha$ 恒定基因序列的未重排T细胞受体 (TCR)  $\alpha$ 可变基因座,和/或包含至少一个人V $\beta$ 区段、至少一个人D $\beta$ 区段和至少一个人J $\beta$ 区段并可操作地连接至鼠TCR $\beta$ 恒定基因序列的未重排TCR $\beta$ 可变基因座;以及任选地 (c) 编码第一嵌合人/鼠MHC多肽 (例如,MHC II $\alpha$ 的第一核苷酸序列、编码第二嵌合人/鼠MHC多肽 (例如,MHC II $\beta$ 的第二核苷酸序列和/或编码第三嵌合人/鼠MHC多肽 (例如,MHC I) 的第三核苷酸序列以及编码人或人源化 $\beta$ 2微球蛋白的 $\beta$ 2微球蛋白基因座,其中每种嵌合MHC多肽的人部分包含与第一、第二和/或第三嵌合T细胞辅助受体多肽结合的人MHC多肽的细胞外结构域 (例如,其中嵌合MHC II复合体的人部分 (例如,人源化MHC II $\alpha$ 和 $\beta$ 多肽) 与嵌合CD4多肽结合,和/或嵌合MHC I多肽的人部分 (或MHC I复合体,例如人源化MHC I $\alpha$ 和人 (源化)  $\beta$ 2微球蛋白) 与嵌合CD8辅助受体 (例如,人源化CD8 $\alpha$ 和 $\beta$ 多肽) 结合。

[0216] 本文提供了经遗传修饰的小鼠,该小鼠在其基因组中例如在其内源CD4基因座处包含编码嵌合人/小鼠CD4多肽的核苷酸序列,其中该嵌合多肽的小鼠部分包含小鼠CD4多肽的至少跨膜结构域和胞质结构域,并且其中该小鼠表达嵌合人/小鼠CD4。在一个实施方

案中,嵌合多肽的人部分包含人CD4多肽的至少全部或基本上全部细胞外结构域。在一个实施方案中,嵌合多肽的人部分包含人CD4蛋白质的至少全部或基本上全部D1结构域。在一个实施方案中,嵌合多肽的人部分包含人CD4蛋白质的至少全部或基本上全部D1-D2结构域,例如人CD4蛋白质的至少全部或基本上全部D1-D3结构域,例如人CD4蛋白质的全部或基本上全部D1-D4结构域。因此,在一个实施方案中,该小鼠在内源CD4基因座处包含具有人CD4基因的至少全部或基本上全部外显子4、5和6的核苷酸序列,例如编码人CD4的D1结构域一部分的人CD4基因的外显子3和人CD4基因的外显子4-6的序列。在一个实施方案中,该小鼠在内源CD4基因座处包含嵌合人/小鼠CD4,其包含负责与MHC II和/或T细胞受体的细胞外部分相互作用的人CD4序列。在另一个实施方案中,该小鼠在内源CD4基因座处包含嵌合人/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。在一个实施方案中,该核苷酸序列包含编码小鼠CD4信号肽的序列。在一个实施方案中,该小鼠包含用编码人CD4细胞外结构域的核苷酸序列替换编码小鼠CD4细胞外结构域的核苷酸序列。在另一个实施方案中,该小鼠包含用编码其的人核苷酸序列替换编码至少全部或基本上全部小鼠CD4D1结构域的核苷酸序列,例如编码至少全部或基本上全部小鼠CD4D1-D2结构域的核苷酸序列,例如编码至少全部或基本上全部小鼠CD4D1-D3结构域的核苷酸序列。在一个实施方案中,嵌合CD4多肽的结构域由在图5A中示意性表示的核苷酸序列编码。

[0217] 在一个实施方案中,该小鼠不从其内源小鼠CD4基因座表达功能性内源小鼠CD4。在一个实施方案中,本文所述的小鼠在小鼠种系中包含嵌合人/小鼠CD4核苷酸序列。

[0218] 在一个实施方案中,该小鼠保留未被人源化的任何内源序列,例如在其中该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替换的实施方案中,该小鼠保留编码小鼠CD4D4结构域的内源核苷酸序列以及编码小鼠CD4的跨膜结构域和胞质结构域的核苷酸序列。

[0219] 在一个方面,表达嵌合人/小鼠CD4蛋白质的小鼠保留小鼠CD4启动子和调控序列,例如编码嵌合人/小鼠CD4的小鼠中的核苷酸序列可操作地连接至内源小鼠CD4启动子和调控序列。在一个方面,本发明的经遗传工程改造的动物中保留的这些小鼠调控序列包括在T细胞发育过程的适当阶段调节嵌合蛋白质表达的序列。因此,在一个方面,该小鼠不在B细胞或成熟CD8<sup>+</sup>T细胞上表达嵌合CD4。在一个方面,该小鼠也不在通常不表达内源CD4的任何细胞类型例如任何免疫细胞类型上表达嵌合CD4。

[0220] 本文所公开的经遗传修饰的小鼠可在其基因组中例如在其内源CD8基因座处包含编码嵌合人/小鼠CD8 $\alpha$ 多肽的第一核苷酸序列和编码嵌合人/小鼠CD8 $\beta$ 多肽的第二核苷酸序列。在一个实施方案中,第一核苷酸序列包含编码人CD8 $\alpha$ 多肽的全部或基本上全部细胞外部分以及小鼠CD8 $\alpha$ 多肽的至少跨膜结构域和胞质结构域的序列,并且第二核苷酸序列包含编码人CD8 $\beta$ 多肽的全部或基本上全部细胞外部分以及小鼠CD8 $\beta$ 多肽的至少跨膜结构域和胞质结构域的序列,并且其中该小鼠表达功能性嵌合人/小鼠CD8蛋白质。在一个实施方案中,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样结构域和小鼠CD8 $\alpha$ 多肽的其余序列的序列,并且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样结构域和小鼠CD8 $\beta$ 多肽的其余序列的序列。在一个实施方案中,第一核苷酸序列包含人CD8 $\alpha$ 多肽的至少MHC I结合结构域。在一个实施方案中,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 $\beta$ 多肽的细胞外部分的外显子。在一个实施方案中,人CD8 $\alpha$ 多肽和/或

CD8 $\beta$ 多肽的细胞外部分是包含并非跨膜或胞质结构域的人CD8 $\alpha$ 多肽和/或CD8 $\beta$ 多肽的部分的区域。在一个实施方案中,嵌合CD8 $\alpha$ 多肽的结构域由在图5B中示意性表示的核苷酸序列编码。在一个实施方案中,嵌合CD8 $\beta$ 多肽的结构域由在图5B中示意性表示的核苷酸序列编码。在一个实施方案中,编码嵌合人/小鼠CD8 $\alpha$ 多肽和/或CD8 $\beta$ 多肽的核苷酸序列分别包含编码小鼠CD8 $\alpha$ 和/或CD8 $\beta$ 信号肽的序列。或者,该核苷酸序列可以包含编码人CD8 $\alpha$ 和/或CD8 $\beta$ 信号序列的序列。在一个实施方案中,该小鼠包含分别用编码全部或基本上全部人CD8 $\alpha$ 和/或CD8 $\beta$ 细胞外结构域的核苷酸序列替换编码全部或基本上全部小鼠CD8 $\alpha$ 和/或CD8 $\beta$ 细胞外结构域的核苷酸序列。

[0221] 在一个实施方案中,该小鼠不从其内源CD8基因座表达功能性内源小鼠CD8 $\alpha$ 和/或CD8 $\beta$ 多肽。在一个实施方案中,如本文所述的小鼠在其种系中包含嵌合人/小鼠CD8序列。

[0222] 在一个方面,表达嵌合人/小鼠CD8 $\alpha$ 和/或CD8 $\beta$ 多肽的小鼠保留小鼠CD8 $\alpha$ 和/或CD8 $\beta$ 启动子和调控序列,例如编码嵌合人/小鼠CD8的小鼠中的核苷酸序列可操作地连接至内源小鼠CD8启动子和调控序列。在一个方面,该小鼠中保留的这些调控序列包括在T细胞发育的适当阶段调节CD8蛋白质表达的序列。在一个方面,经遗传修饰的小鼠不在B细胞或成熟CD4<sup>+</sup>T细胞或通常不表达内源CD8的任何细胞例如免疫细胞上表达嵌合CD8。

[0223] 本发明还提供了一种经遗传修饰的小鼠,该小鼠在其基因组中包含未重排人或人源化TCR可变基因座,例如TCR $\alpha$ 、TCR $\beta$ 、TCR $\delta$ 和/或TCR $\gamma$ 可变基因座。在一些实施方案中,未重排人或人源化TCR可变基因座替换内源小鼠TCR可变基因座。在其他实施方案中,未重排人或人源化TCR可变基因座位于基因组中除对应内源小鼠TCR基因座之外的位点。在一些实施方案中,人或人源化未重排TCR可变基因座可操作地连接至小鼠TCR恒定区。

[0224] 在一个实施方案中,提供了经遗传修饰的小鼠,其中该小鼠在其基因组中包含具有至少一个人V $\alpha$ 区段和至少一个人J $\alpha$ 区段并可操作地连接至小鼠TCR $\alpha$ 恒定基因序列的未重排T细胞受体(TCR) $\alpha$ 可变基因座,以及具有至少一个人V $\beta$ 区段、至少一个人D $\beta$ 区段和至少一个人J $\beta$ 区段并可操作地连接至小鼠TCR $\beta$ 恒定基因序列的未重排TCR $\beta$ 可变基因座。在一个具体实施方案中,该小鼠在其基因组中包含具有人V $\alpha$ 区段的完整组库和人J $\alpha$ 区段的完整组库并可操作地连接至小鼠TCR $\alpha$ 恒定基因序列的未重排TCR $\alpha$ 可变基因座,以及具有人V $\beta$ 区段的完整组库、人D $\beta$ 区段的完整组库和人J $\beta$ 区段的完整组库并可操作地连接至小鼠TCR $\beta$ 恒定基因序列的未重排TCR $\beta$ 可变基因座。

[0225] 在一些实施方案中,包含人TCR $\alpha$ 可变区段或未重排TCR $\alpha$ 可变基因座替换内源小鼠TCR $\alpha$ 可变基因座,并且包含人TCR $\beta$ 可变区段或未重排TCR $\beta$ 可变基因座替换内源小鼠TCR $\beta$ 可变基因座。在一些实施方案中,内源小鼠V $\alpha$ 和/或J $\alpha$ 区段不能重排形成重排的V $\alpha$ /J $\alpha$ 序列,并且内源小鼠V $\beta$ 、D $\beta$ 和J $\beta$ 区段不能重排形成重排的V $\beta$ /D $\beta$ /J $\beta$ 序列。在一些实施方案中,人V $\alpha$ 和J $\alpha$ 区段重排形成重排的人V $\alpha$ /J $\alpha$ 序列,并且人V $\beta$ 、D $\beta$ 和J $\beta$ 区段重排形成重排的人V $\beta$ /D $\beta$ /J $\beta$ 序列。

[0226] 本发明还涉及一种经遗传修饰的小鼠,该小鼠在其基因组中包含编码嵌合MHC多肽的核酸序列,其中嵌合MHC多肽的人部分与如本文所公开的嵌合T细胞辅助受体的人细胞外结构域结合。如本文所公开的经遗传修饰的小鼠可以包含编码嵌合人/小鼠MHC I的第一核酸序列、编码嵌合人/小鼠MHC II $\alpha$ 的第二核酸序列和/或编码嵌合人/小鼠MHC II $\beta$ 多肽的第三核酸序列。嵌合MHC I、MHC II $\alpha$ 和/或MHC II $\beta$ 的人部分可分别包含人MHC I、MHC II $\alpha$



和MHC II $\beta$ 的细胞外结构域。在一个实施方案中,该小鼠从其内源小鼠MHC基因座表达功能性嵌合人/小鼠MHC I、MHC II $\alpha$ 和MHC II $\beta$ 多肽。在一个实施方案中,该小鼠不从其内源小鼠MHC基因座表达功能性小鼠MHC多肽,例如功能性小鼠MHC I、MHC II $\alpha$ 和MHC II $\beta$ 多肽。在其他实施方案中,该小鼠在细胞表面上表达的唯一MHC I和MHC II是嵌合MHC I和II。

[0227] 在一个实施方案中,嵌合人/小鼠MHC I多肽的人部分包含人MHC I的肽结合结构域或细胞外结构域(例如人HLA-A,如人HLA-A2,如人HLA-A2.1)。在一些实施方案中,该小鼠不从其内源小鼠MHC I基因座表达内源小鼠MHC I多肽的肽结合结构域或细胞外结构域。人MHC I的肽结合结构域可包含 $\alpha$ 1和 $\alpha$ 2结构域。或者,人MHC I的肽结合结构域可包含 $\alpha$ 1、 $\alpha$ 2和 $\alpha$ 3结构域。在一个方面,人MHC I的细胞外结构域包含人MHC I $\alpha$ 链的细胞外结构域。在一个实施方案中,内源小鼠MHC I基因座是H-2K(例如,H-2Kb)基因座,并且嵌合MHC I多肽的小鼠部分包含小鼠H-2K(例如,H-2Kb)多肽的跨膜结构域和胞质结构域。因此,在一个实施方案中,本发明的小鼠在其内源小鼠MHC I基因座处包含编码嵌合人/小鼠MHC I的核酸序列,其中所述嵌合多肽的人部分包含人HLA-A2(例如,HLA-A2.1)多肽的细胞外结构域,而小鼠部分包含小鼠H-2K(例如H-2Kb)多肽的跨膜结构域和胞质结构域,并且小鼠表达嵌合人/小鼠HLA-A2/H-2K蛋白质。在其他实施方案中,嵌合MHC I多肽的小鼠部分可以来源于其他小鼠MHC I,例如H-2D、H-2L等;并且嵌合MHC I多肽的人部分可以来源于其他人MHC I,例如HLA-B、HLA-C等。在一个方面,该小鼠不从其内源小鼠H-2K基因座表达功能性内源H-2K多肽。在一个实施方案中,该小鼠不从其H-2D基因座表达功能性内源MHC多肽。在一些实施方案中,该小鼠经工程改造成缺少内源H-2D基因座的全部或一部分。在其他实施方案中,该小鼠在细胞表面上表达的唯一MHC I多肽是嵌合人/小鼠MHC I多肽。

[0228] 在一个实施方案中,嵌合人/小鼠MHC II $\alpha$ 多肽的人部分包含人MHC II $\alpha$ 肽结合结构域或细胞外结构域,并且嵌合人/小鼠MHC II $\beta$ 多肽的人部分包含人MHC II $\beta$ 肽结合结构域或细胞外结构域。在一些实施方案中,该小鼠不从不内源小鼠基因座(例如,H-2A和/或H-2E基因座)表达内源小鼠 $\alpha$ 和/或 $\beta$ 多肽的肽结合结构域或细胞外结构域。在一些实施方案中,该小鼠包含缺少编码功能性MHC II类分子的基因的基因组,该功能性MHC II类分子包含H-2Ab1、H-2Aa、H-2Eb1、H-2Eb2、H-2Ea以及它们的组合。在一些实施方案中,该小鼠在细胞表面上表达的唯一MHC II多肽是嵌合人/小鼠MHC II多肽。人MHC II $\alpha$ 多肽的肽结合结构域可以包含 $\alpha$ 1结构域,并且人MHC II $\beta$ 多肽的肽结合结构域可以包含 $\beta$ 1结构域;因此,嵌合MHC II复合体的肽结合结构域可以包含人 $\alpha$ 1和 $\beta$ 1结构域。人MHC II $\alpha$ 多肽的细胞外结构域可以包含 $\alpha$ 1和 $\alpha$ 2结构域,并且人MHC II $\beta$ 多肽的细胞外结构域可以包含 $\beta$ 1和 $\beta$ 2结构域;因此,嵌合MHC II复合体的细胞外结构域可以包含人 $\alpha$ 1、 $\alpha$ 2、 $\beta$ 1和 $\beta$ 2结构域。在一个实施方案中,嵌合MHC II复合体的小鼠部分包含小鼠MHC II的跨膜结构域和胞质结构域,例如小鼠H-2E(例如,小鼠H-2E $\alpha$ 和 $\beta$ 链的跨膜结构域和胞质结构域)。因此,在一个实施方案中,本发明的小鼠在其内源小鼠MHC II基因座处包含编码嵌合人/小鼠MHC II $\alpha$ 的核酸序列,其中嵌合MHC II $\alpha$ 多肽的人部分包含来源于人MHC II的 $\alpha$ 链(例如,HLA-DR2的 $\alpha$ 链)的细胞外结构域,且小鼠部分包含来源于小鼠MHC II的 $\alpha$ 链(例如,H-2E)的跨膜结构域和胞质结构域;并且小鼠在其内源小鼠MHC II基因座处包含编码嵌合人/小鼠MHC II $\beta$ 的核酸序列,其中嵌合MHC II $\beta$ 多肽的人部分包含来源于人MHC II的 $\beta$ 链(例如,HLA-DR2的 $\beta$ 链)的细胞外结构域,且小鼠部分包含来源于小鼠MHC II的 $\beta$ 链(例如,H-2E)的跨膜结构域和胞质结构域;例如,其中



该小鼠表达嵌合人/小鼠HLA-DR2/H-2E蛋白质。在其他实施方案中,嵌合MHC II蛋白质的小鼠部分可以来源于其他小鼠MHC II,例如H-2A等;并且嵌合MHC II蛋白质的人部分可以来源于其他人MHC II,例如HLA-DQ等。在一个方面,该小鼠不从其内源小鼠基因座表达功能性内源H-2A和H-2E多肽(例如,该小鼠不表达H-2Ab1、H-2Aa、H-2Eb1、H-2Eb2和H-2Ea多肽)。在一些实施方案中,该小鼠缺少细胞表面上任何内源MHC I或MHC II分子的表达。

[0229] 在各个方面,由经遗传修饰的非人动物或来源于非人动物的细胞、胚胎或组织表达的人或人源化 $\beta$ 2微球蛋白保留内源和/或人 $\beta$ 2微球蛋白的所有功能方面。例如,优选的是人或人源化 $\beta$ 2微球蛋白结合MHC I多肽(例如,内源非人或人MHC I多肽)的 $\alpha$ 链。人或人源化 $\beta$ 2微球蛋白多肽可结合、募集或以其他方式连接任何其他分子,例如与内源非人和/或人 $\beta$ 2微球蛋白结合的受体、锚定分子或信号传导分子(例如,HFE等)。

[0230] 除了经遗传修饰的动物(例如,啮齿动物,例如小鼠或大鼠)之外,还提供了组织或细胞,其中该组织或细胞来源于如本文所述的非人动物,并且包含异源 $\beta$ 2微球蛋白基因或 $\beta$ 2微球蛋白序列,即核苷酸和/或氨基酸序列。在一个实施方案中,异源 $\beta$ 2微球蛋白基因或 $\beta$ 2微球蛋白序列是人或人源化 $\beta$ 2微球蛋白基因或者人或人源化 $\beta$ 2微球蛋白序列。优选地,该细胞是有核细胞。该细胞可以是已知表达MHC I复合体的任何细胞,例如抗原递呈细胞。由所述细胞表达的人或人源化 $\beta$ 2微球蛋白多肽可以与内源非人MHC I(例如,啮齿动物MHC I)相互作用以形成功能性MHC I复合体。所得MHC I复合体可能能够与T细胞例如细胞毒性T细胞相互作用。因此,还提供了来自本文所述非人动物的细胞与T细胞的体外复合体。

[0231] 还提供非人细胞,其包含人或人源化 $\beta$ 2微球蛋白基因或序列,以及另外的人或人源化序列,例如本发明公开的嵌合MHC I多肽。在这种情况下,人或人源化 $\beta$ 2微球蛋白多肽可与例如嵌合人/非人MHC I多肽相互作用,并且可形成功能性MHC I复合体。在一些方面,这样的复合体能够与T细胞例如人或非人T细胞上的TCR相互作用。因此,还提供了来自本文所述非人动物的细胞与人或非人T细胞的体外复合体。

[0232] 本公开的另一方面是包含本文所述异源 $\beta$ 2微球蛋白基因或 $\beta$ 2微球蛋白序列的啮齿动物胚胎(例如,小鼠或大鼠胚胎)。在一个实施方案中,该胚胎包含ES供体细胞和宿主胚胎细胞,该ES供体细胞包含异源 $\beta$ 2微球蛋白基因或 $\beta$ 2微球蛋白序列。异源 $\beta$ 2微球蛋白基因或 $\beta$ 2微球蛋白序列是人或人源化 $\beta$ 2微球蛋白基因或 $\beta$ 2微球蛋白序列。

[0233] 本发明还涵盖包含本文所述非人动物的染色体或其片段的非人细胞(例如,其中该染色体或其片段包含编码人或人源化 $\beta$ 2微球蛋白多肽的核苷酸序列)。该非人细胞可以包含本文所述非人动物的细胞核。在一个实施方案中,作为核移植的结果,该非人细胞包含染色体或其片段。

[0234] 在一个方面,提供了包含异源 $\beta$ 2微球蛋白基因或 $\beta$ 2微球蛋白序列的非人诱导多能细胞。在一个实施方案中,该诱导多能细胞来源于本文所述的非人动物。在一个实施方案中,该异源 $\beta$ 2微球蛋白基因或 $\beta$ 2微球蛋白序列是人或人源化基因或序列。

[0235] 在本发明的一些实施方案中,本文所述的小鼠仅在小鼠的专职抗原递呈细胞例如B细胞、单核细胞/巨噬细胞和/或树突细胞上表达嵌合人/小鼠MHC II。在一些实施方案中,本文所述的小鼠对一种或多种人抗原引发免疫应答,例如细胞免疫应答。在一些实施方案中,本文所述的小鼠对一种或多种人抗原引发人源化T细胞应答。

[0236] 除了经遗传工程改造的非人动物,还提供了非人胚胎(例如啮齿动物,例如小鼠或

大鼠胚胎),其中该胚胎包含如本文所述来源于非人动物(例如啮齿动物,例如小鼠或大鼠)的供体ES细胞。在一个方面,该胚胎包含ES供体细胞和宿主胚胎细胞,该ES供体细胞包含嵌合CD4基因、嵌合CD8(例如,CD8 $\alpha$ 和/或CD8 $\beta$ )基因、人源化MHC I(例如,MHC I $\alpha$ )核酸序列、人源化MHC II(例如,MHC II $\alpha$ 和/或MHC II $\beta$ )核酸序列、未重排人源化TCR(例如,TCR $\alpha$ 和/或TCR $\beta$ 、或TCR $\delta$ 、和/或TCR $\gamma$ )基因座和/或人或人源化 $\beta$ 2微球蛋白基因序列。

[0237] 还提供了组织,其中该组织来源于如本文所述的非人动物(例如,啮齿动物,例如小鼠或大鼠),并且表达嵌合CD4蛋白质、嵌合CD8蛋白质(例如,嵌合CD8 $\alpha$ 和/或CD8 $\beta$ 蛋白质)、人源化TCR多肽(例如,TCR $\alpha$ 和/或TCR $\beta$ 、或TCR $\delta$ 、和/或TCR $\gamma$ 多肽)、人源化MHC I多肽(例如,MHC I $\alpha$ )、人源化MHC II多肽(例如,MHC II $\alpha$ 和/或MHC II $\beta$ 多肽)和/或人或人源化 $\beta$ 2微球蛋白。

[0238] 在一个方面,提供了用于制备嵌合人/非人CD4分子的方法,该方法包括在单细胞中从如本文所述的核苷酸构建体表达嵌合CD4蛋白质。在一个实施方案中,核苷酸构建体是病毒载体;在一个具体实施方案中,病毒载体是慢病毒载体。在一个实施方案中,细胞选自表达病毒核酸序列的CHO、COS、293、HeLa和视网膜细胞(例如PERC.6<sup>TM</sup>细胞)。

[0239] 在一个方面,提供了表达嵌合CD4蛋白质的细胞。在一个实施方案中,该细胞包含表达载体,该表达载体包含如本文所述的嵌合CD4序列。在一个实施方案中,细胞选自表达病毒核酸序列的CHO、COS、293、HeLa和视网膜细胞(例如PERC.6<sup>TM</sup>细胞)。

[0240] 还提供了由本文所述非人动物制备的嵌合CD4分子,其中在一个实施方案中,嵌合CD4分子包含人CD4蛋白质的全部或基本上全部细胞外结构域以及来自非人CD4蛋白质例如小鼠CD4蛋白质的至少跨膜结构域和胞质结构域的氨基酸序列。在另一个实施方案中,提供了由本文所述非人动物制备的嵌合CD4分子,其中该嵌合CD4分子包含人CD4的至少全部或基本上全部D1结构域,例如人CD4的至少全部或基本上全部D1-D2结构域,例如人CD4的至少全部或基本上全部D1-D3结构域的氨基酸序列,例如负责结合MHC II和/或TCR的细胞外结构域的人CD4的氨基酸序列,例如负责结合MHC II和/或TCR的可变结构域的人CD4的氨基酸序列;并且其中该蛋白质的其余部分(例如,跨膜结构域、胞质结构域、未被人源化的细胞外结构域的任何部分)来源于内源非人蛋白质序列。示例性嵌合人/非人CD4多肽包含SEQ ID NO:78所示的氨基酸序列,并且嵌合多肽的人部分跨越SEQ ID NO:78的大约第27-319位氨基酸(在SEQ ID NO:79中单独示出)。

[0241] 在一个方面,提供了用于制备嵌合人/非人CD8分子(例如,CD8 $\alpha$ 和/或CD8 $\beta$ )的方法,该方法包括在单细胞中从本文所述的核苷酸构建体表达嵌合CD8多肽。在一个实施方案中,核苷酸构建体是病毒载体;在一个具体实施方案中,病毒载体是慢病毒载体。在一个实施方案中,细胞选自表达病毒核酸序列的CHO、COS、293、HeLa和视网膜细胞(例如PERC.6<sup>TM</sup>细胞)。

[0242] 在一个方面,提供了表达嵌合CD8蛋白质的细胞。在一个实施方案中,该细胞包含表达载体,该表达载体包含如本文所述的嵌合CD8序列。在一个实施方案中,该细胞选自表达病毒核酸序列的CHO、COS、293、HeLa和视网膜细胞(例如PERC.6<sup>TM</sup>细胞)。

[0243] 还提供了由本文所述非人动物制备的嵌合CD8分子,其中该嵌合CD8分子包含来自人CD8蛋白质(例如,CD8 $\alpha$ 和/或CD8 $\beta$ )的全部或基本上全部细胞外结构域,以及来自非人CD8蛋白质例如小鼠CD8蛋白质的至少跨膜结构域和胞质结构域。示例性嵌合CD8 $\alpha$ 多肽在SEQ

ID NO:88中示出,并且示例性嵌合CD8 $\beta$ 蛋白质在SEQ ID NO:83中示出。

[0244] 还提供了由本文所述非人动物(例如,啮齿动物,例如小鼠或大鼠)制备的人源化TCR蛋白质,其中该人源化TCR蛋白质包含人可变区和非人恒定区。因此,人源化TCR蛋白质在其可变结构域和非人恒定区中包含人互补决定区(即人CDR1、2和3)。还提供了编码由本文所述非人动物产生的人TCR可变结构域的核酸。

[0245] 此外,提供了分离自如本文所述的非人动物的非人细胞。在一个实施方案中,该细胞为ES细胞。在一个实施方案中,该细胞为T细胞,例如CD4<sup>+</sup>T细胞。在一个实施方案中,该细胞为辅助T细胞(T<sub>H</sub>细胞)。在一个实施方案中,该T<sub>H</sub>细胞为效应T<sub>H</sub>细胞,例如T<sub>H</sub>1细胞或T<sub>H</sub>2细胞。在一个实施方案中,该细胞为CD8<sup>+</sup>T细胞。在一个实施方案中,该细胞为细胞毒性T细胞。还提供了表达包含人可变区和非人恒定区的TCR蛋白质的非人细胞。该TCR蛋白质可以包含TCR $\alpha$ 、TCR $\beta$ 或它们的组合。在一个实施方案中,该细胞为T细胞,例如CD4<sup>+</sup>或CD8<sup>+</sup>T细胞。另外,本文所提供的非人T细胞可以在其细胞表面上表达(a)包含可操作地连接至非人T细胞辅助受体跨膜和/或细胞内结构域的人T细胞辅助受体细胞外结构域的嵌合人/非人T细胞辅助受体,例如嵌合CD4多肽或嵌合CD8多肽;和(b)包含人可变区和非人恒定区的TCR蛋白质。

[0246] 在另一个实施方案中,该细胞为抗原递呈细胞。在一个实施方案中,该抗原递呈细胞在人源化MHC I分子上递呈抗原。在另一个实施方案中,该抗原递呈细胞为专职抗原递呈细胞,例如B细胞、树突细胞和巨噬细胞。在另一个实施方案中,抗原递呈细胞在人源化MHC I和/或人源化MHC II分子上递呈抗原。

[0247] 在一个方面,提供了表达嵌合人/非人MHC I和MHC II蛋白质(例如HLA-A2/H-2K和HLA-DR2/H-2E蛋白质)的细胞。在一个方面,该细胞是不从其H-2D基因座表达功能性内源MHC多肽的小鼠细胞。在一些实施方案中,该细胞是经工程改造成缺少内源H-2D基因座的全部或一部分的小鼠细胞。在一些实施方案中,该细胞是在其表面上不表达任何功能性内源MHC I和MHC II多肽的小鼠细胞。在一个实施方案中,该细胞包含表达载体,该表达载体包含本文所述的嵌合MHC I类序列和嵌合MHC II类序列。在一个实施方案中,该细胞选自表达病毒核酸序列的CHO、COS、293、HeLa和视网膜细胞(例如PERC.6<sup>TM</sup>细胞)。

[0248] 可以通过抗HLA-DR抗体检测包含本文所述HLA-DR2的细胞外结构域的嵌合MHC II复合体。因此,可以使用抗HLA-DR抗体检测和/或选择展示嵌合人/非人MHC II多肽的细胞。可以使用抗HLA-A例如抗HLA-A2抗体检测包含本文所述HLA-A2的细胞外结构域的嵌合MHC I复合体。因此,可以使用抗HLA-A抗体检测和/或选择展示嵌合人/非人MHC I多肽的细胞。识别其他HLA等位基因的抗体可商购获得或可以产生,并且可用于检测/选择。

[0249] 虽然下面实施例描述的经遗传工程改造的动物的基因组包含分别用编码嵌合人/小鼠HLA-A2/H-2K和HLA-DR2/H-2E蛋白质的核酸序列替换编码小鼠H-2K、H-2A和H-2E蛋白质的核酸序列,但本领域技术人员应当理解,可以使用类似的策略引入包含其他人MHC I和II基因(其他HLA-A、HLA-B和HLA-C;以及其他HLA-DR、HLA-DP和HLA-DQ基因)的嵌合体。还提供了在内源MHC基因座处包含多个嵌合人/非人(例如,人/啮齿动物,例如人/小鼠)MHC I和MHC II基因的此类动物。此类嵌合MHC I和MHC II蛋白质的示例描述于美国专利公布No.20130111617、20130185819、20130185820和20140245467以及美国专利No.8,847,005中,这些专利均以引用方式并入本文。

[0250] 还提供了包含本文所述非人动物的染色体或其片段的非人细胞。在一个实施方案中,该非人细胞包含本文所述非人动物的细胞核。在一个实施方案中,作为核移植的结果,该非人细胞包含染色体或其片段。

[0251] 在一个方面,还提供了非人诱导多能细胞,该非人诱导多能细胞包含编码嵌合CD4多肽的基因、编码嵌合CD8多肽(例如,CD8 $\alpha$ 和/或CD8 $\beta$ 多肽)的基因、编码人源化MHC I多肽(例如,MHC I $\alpha$ 和/或 $\beta$ 2微球蛋白)的基因、编码人源化MHC II多肽(例如,MHC II $\alpha$ 和/或MHC II $\beta$ )的基因和/或编码本文所述人源化TCR $\alpha$ 和/或TCR $\beta$ 多肽的未重排人源化TCR基因座。在一个实施方案中,该诱导多能细胞来源于如本文所述的非人动物。

[0252] 在一个方面,提供了来源于如本文所述非人动物的细胞的杂交瘤或四源杂交瘤。在一个实施方案中,该非人动物为小鼠或大鼠。

[0253] 制备产生基本人源化T细胞免疫应答的经遗传修饰非人动物

[0254] 还提供了一种用于制备本文所述经遗传工程改造的非人动物(例如,经遗传工程改造的啮齿动物,例如小鼠或大鼠)的方法。通常,所述方法包括(a)将编码嵌合人/非人T细胞辅助受体多肽的第一核苷酸序列、编码第二嵌合人/非人T细胞辅助受体多肽的第二核苷酸序列和/或编码第三嵌合人/非人T细胞辅助受体多肽的第三核苷酸序列引入非人动物的基因组,其中每种嵌合T细胞辅助受体多肽的非人部分均包含非人T细胞辅助受体的至少跨膜结构域和胞质结构域,并且其中每种嵌合多肽的人部分均包含人T细胞辅助受体的细胞外部分(或其一部分);(b)将包含至少一个人V $\alpha$ 区段和至少一个人J $\alpha$ 区段并可操作地连接至非人TCR $\alpha$ 恒定基因序列的未重排T细胞受体(TCR) $\alpha$ 可变基因座以及/或者包含至少一个人V $\beta$ 区段、至少一个人D $\beta$ 区段和至少一个人J $\beta$ 区段并可操作地连接至非人TCR $\beta$ 恒定基因序列的未重排TCR $\beta$ 可变基因座插入非人动物的基因组中 $\beta$ ;以及任选地(c)将编码第一嵌合人/非人MHC多肽的第一核苷酸序列、编码第二嵌合人/非人MHC多肽的第二核苷酸序列和/或编码第三嵌合人/非人MHC多肽的第三核苷酸序列放置到所述基因组中和/或(d)将编码人或人源化 $\beta$ 2微球蛋白多肽的 $\beta$ 2微球蛋白基因座添加到非人动物的基因组中。在一些实施方案中,引入、插入和/或放置步骤包括靶向编码T细胞辅助受体的细胞外结构域、TCR的可变结构域、MHC多肽的细胞外结构域或 $\beta$ 2微球蛋白一部分的序列,并分别用编码人T细胞辅助受体细胞外结构域、人TCR可变结构域、人MHC细胞外结构域和/或 $\beta$ 2微球蛋白的人部分的序列替换上述序列。

[0255] 在其他实施方案中,引入、插入、放置和/或添加可包括使相同物种的动物交配(例如性交)。在其他实施方案中,引入、插入、放置和/或添加包括在ES细胞中连续同源重组。在一些实施方案中,ES细胞来源于经遗传修饰以包含一种或多种但非全部期望的基因修饰的非人动物,并且在这类ES细胞中的同源重组完成了所述基因修饰。在其他实施方案中,引入、插入、放置和/或添加可包括交配和在ES细胞中同源重组的组合,例如使动物交配产生相同物种的另一种(或多种)动物,其中这些动物中的一些或全部可由通过单个同源重组或连续同源重组事件被遗传修饰的ES细胞产生,并且其中某些ES细胞可分离自包含本文所公开的一种或多种基因修饰的非人动物。

[0256] 如实施例所述,在一些实施方案中,所述方法利用通过VELOCIGENE<sup>®</sup>技术制备的靶向构建体,将该构建体引入ES细胞中,并采用VELOCIMOUSE<sup>®</sup>技术将靶向的ES细胞克隆引入小鼠胚胎中。靶向构建体可包含靶向待替换的内源性序列的5'同源臂和/

或3'同源臂、插入序列(其替换内源性序列)以及一个或多个选择盒。选择盒是插入靶向构建体中以便于选择整合了目标构建体的细胞(例如ES细胞)的核苷酸序列。许多合适的选择盒在本领域中是已知的。通常,选择盒可以在特定抗生素(例如Neo,Hyg,Pur,CM,SPEC等)的存在下进行正选择。此外,选择盒两侧可侧接重组位点,其允许在用重组酶处理后缺失选择盒。常用的重组位点是loxP和Frt,分别由Cre和Flp酶识别,但本领域已知其他重组位点。选择表达盒可位于构建体中编码区外的任何位置。在一个实施方案中,选择盒位于人DNA片段的5'端。在另一个实施方案中,选择盒位于人DNA片段的3'端。在另一个实施方案中,选择盒位于人DNA片段内。在另一个实施方案中,选择盒位于人DNA片段的内含子内。在另一个实施方案中,选择盒位于人DNA片段和小鼠DNA片的连接处。

[0257] 在一个实施方案中,用于制备经遗传工程改造的非人动物的方法得到在内源性CD4基因座处包含编码嵌合人/非人CD4多肽的核苷酸序列的动物。在一个实施方案中,本发明包括修饰非人动物的CD4基因座以表达本文所述的嵌合人/非人CD4多肽的方法。在一个实施方案中,本发明提供了修饰小鼠的CD4基因座以表达嵌合人/小鼠CD4多肽的方法,所述方法包括在非人动物(如小鼠)的内源性CD4基因座处引入编码嵌合人/小鼠CD4多肽的核苷酸序列,例如用编码嵌合人/小鼠CD4多肽的核苷酸序列替换编码内源非人CD4多肽的核苷酸序列。在所述方法的一个方面,嵌合人/小鼠CD4多肽包含人CD4多肽的全部或基本上全部的细胞外结构域,以及内源性小鼠CD4多肽的至少跨膜结构域和胞质结构域。在所述方法的另一方面,嵌合人/小鼠CD4多肽包含人CD4多肽的全部或基本上全部的D1-D2结构域。在又一个实施方案中,嵌合人/小鼠CD4多肽包含人CD4多肽的全部或基本上全部的D1-D3结构域。在又一个实施方案中,嵌合人/小鼠CD4多肽包含负责与MHC II相互作用的人CD4的全部或基本上全部的氨基酸序列和/或T细胞受体的细胞外结构域。在又一个实施方案中,嵌合人/小鼠CD4多肽包含负责与MHC II相互作用的人CD4的全部或基本上全部的氨基酸序列和/或T细胞受体的可变结构域。

[0258] 因此,提供了一种用于产生包含嵌合人/非人CD4的经遗传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'同源臂和3'同源臂、包含人CD4基因序列(例如,人CD4细胞外结构域基因序列,例如人CD4的全部或基本上全部的结构域D1-D2的基因序列,例如人CD4的全部或基本上全部的结构域D1-D3和/或D2-D3的基因序列,例如人CD4的全部或基本上全部的结构域D1-D4的基因序列)的DNA片段以及侧接重组位点的选择盒。在一个实施方案中,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。在一个实施方案中,同源臂与非人(例如小鼠)CD4基因组序列同源。图5A中示出了本发明的示例性构建体。

[0259] 在一些实施方案中,所述方法得到在内源性CD8基因座处包含编码嵌合人/非人CD8 $\alpha$ 和/或CD8 $\beta$ 多肽的核苷酸序列的动物。在一个实施方案中,本发明提供了修饰非人动物的CD8基因座以表达本文所述的嵌合人/非人CD8多肽的方法。在一个方面,提供了修饰小鼠的CD8基因座以表达嵌合人/小鼠CD8多肽的方法,所述方法包括在非人动物(如小鼠)的内源性CD8基因座处引入编码嵌合人/小鼠CD8多肽的核苷酸序列,例如用编码嵌合人/小鼠CD8多肽的核苷酸序列替换编码内源非人CD8多肽的核苷酸序列。CD8多肽可选自CD8 $\alpha$ 、CD8 $\beta$ 以及它们的组合。在一个方面,嵌合多肽包含人CD8多肽的全部或基本上全部的细胞外结构域,以及内源性小鼠CD8多肽的至少跨膜结构域和胞质结构域。

[0260] 因此,也提供了一种用于产生包含嵌合人/非人CD8的经遗传修饰动物的核苷酸构

建体。在一个方面,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片段以及侧接重组位点的选择盒。在一些实施方案中,这种人序列包含人CD8 $\alpha$ 或CD8 $\beta$ 的内含子和外显子,例如,分别编码人CD8 $\alpha$ 或CD8 $\beta$ 的细胞外结构域的外显子。在一个实施方案中,同源臂与非人CD8 $\alpha$ 或CD8 $\beta$ 序列同源。图5B中示出了用于CD8 $\alpha$ 和CD8 $\beta$ 的示例性构建体。

[0261] 由于编码CD8 $\alpha$ 和CD8 $\beta$ 的基因的染色体定位靠近,因此两个基因的顺序靶向提高了成功人源化的可能性。在一个实施方案中,靶向策略包括将本文所述的嵌合CD8 $\beta$ 构建体引入ES细胞中,由靶向的ES细胞产生小鼠,从所述小鼠中得到经遗传修饰的ES细胞,并将本文所述的嵌合CD8 $\alpha$ 构建体引入所述经遗传修饰的ES细胞中。在另一个实施方案中,靶向策略包括将本文所述的嵌合CD8 $\beta$ 构建体引入ES细胞中,选择已结合嵌合CD8 $\beta$ 构建体的细胞,将本文所述的嵌合CD8 $\alpha$ 构建体引入已结合并携带嵌合CD8 $\beta$ 构建体的ES细胞中,以及选择已结合嵌合CD8 $\beta$ 和CD8 $\alpha$ 两者的细胞。在该实施方案的一个方面,利用不同的选择标记来实施选择步骤。在另选的实施方案中,可首先完成CD8 $\alpha$ 人源化。在完成基因靶向之后,可以筛选经遗传修饰的非人动物的ES细胞,以通过本领域已知的多种方法(例如,在Valenzuela et al. (2003) High-throughput engineering of the mouse genome coupled with high-resolution expression analysis, Nature Biotech. 21 (6) :652-659 (Valenzuela等人, 2003年,“小鼠基因组的高通量工程与高分辨率表达分析相结合”,《自然生物技术》,第21卷第6期,第652-659页)中描述的等位基因修饰测定法) 确认成功结合目标外源性核苷酸序列或成功表达外源性多肽。

[0262] 在一些实施方案中,用于制备经遗传修饰的非人动物的方法得到其基因组包含人源化未重排TCR基因座(例如,人源化未重排TCR $\alpha$ 、TCR $\beta$ 、TCR $\delta$ 和/或TCR $\gamma$ 基因座)的动物。在一个实施方案中,提供了用于制备经遗传修饰的非人动物(例如,啮齿动物如小鼠或大鼠)的方法,这种动物在T细胞表面上表达包含人可变区和非人(例如,啮齿动物如小鼠或大鼠)恒定区的T细胞受体,其中所述方法包括在第一非人动物中插入包含至少一个人V $\alpha$ 区段和至少一个人J $\alpha$ 区段的未重排人源化TCR $\alpha$ 可变基因座,例如,用包含至少一个人V $\alpha$ 区段和至少一个人J $\alpha$ 区段的未重排人源化TCR $\alpha$ 可变基因座替换内源性非人TCR $\alpha$ 可变基因座,其中人源化TCR $\alpha$ 可变基因座可操作地连接至内源性TCR $\alpha$ 恒定区;在第二非人动物中插入包含至少一个人V $\beta$ 区段、一个人D $\beta$ 区段和一个人J $\beta$ 区段的未重排人源化TCR $\beta$ 可变基因座,例如,用包含至少一个人V $\beta$ 区段、一个人D $\beta$ 区段和一个人J $\beta$ 区段的未重排人源化TCR $\beta$ 可变基因座替换内源性非人TCR $\beta$ 可变基因座,其中人源化TCR $\beta$ 可变基因座可操作地连接至内源性TCR $\beta$ 恒定区;以及使第一非人动物和第二非人动物交配以获得表达包含人可变区和非人恒定区的T细胞受体的非人动物。在其他实施方案中,本发明提供了制备其基因组包含人源化未重排TCR $\alpha$ 基因座的经遗传修饰非人动物或其基因组包含人源化未重排TCR $\beta$ 基因座的非人动物的方法。在各种实施方案中,替换是在内源性基因座处进行的。在各种实施方案中,所述方法包括渐进式人源化策略,其中将包含另外的可变区区段的构建体在人源化的每个后续步骤引入ES细胞中,最终得到包含人可变区区段的完整组库的小鼠(参见例如,图4A和4B)。

[0263] 本公开还提供了一种修饰非人动物的TCR可变基因座(例如,TCR $\alpha$ 、TCR $\beta$ 、TCR $\delta$ 和/或TCR $\gamma$ 基因座)以表达本文所述的人源化TCR蛋白的方法。在一个实施方案中,本发明提供了修饰TCR可变基因座以在T细胞表面上表达人源化TCR蛋白的方法,其中所述方法包括在非人动物中插入未重排人源化TCR可变基因座,例如用未重排人源化TCR可变基因座替换内

源性非人TCR可变基因座。在其中TCR可变基因座是TCR $\alpha$ 可变基因座的一个实施方案中,未重排人源化TCR可变基因座包含至少一个人Va区段和至少一个人Ja区段。在其中TCR可变基因座是TCR $\beta$ 可变基因座的一个实施方案中,未重排人源化TCR可变基因座包含至少一个人V $\beta$ 区段、至少一个人D $\beta$ 区段和至少一个人J $\beta$ 区段。在各个方面,未重排人源化TCR可变基因座可操作地连接至对应内源性非人TCR恒定区。

[0264] 因此,还提供了用于产生包含人源化TCR可变区基因的经遗传修饰动物的核苷酸构建体。在一个方面,核苷酸构建体包含:5'同源臂和3'同源臂、包含人TCR可变区基因区段的人DNA片段以及侧接重组位点的选择盒。在一个实施方案中,人DNA片段是TCR $\alpha$ 基因片段,并且其包含至少一个人TCR $\alpha$ 可变区区段。在另一个实施方案中,人DNA片段是TCR $\beta$ 片段,并且其包含至少一个人TCR $\beta$ 可变区基因区段。在一个方面,至少一个同源臂是非人同源臂,并且它与非人TCR基因座(例如,非人TCR $\alpha$ 或TCR $\beta$ 基因座)同源。

[0265] 在本发明的各个方面,编码嵌合人/非人MHC I和MHC II多肽的序列位于内源性非人MHC基因座(例如,小鼠H-2K和/或H-2E基因座)处。在一个实施方案中,这导致放置编码人或人源化MHC I多肽的核酸序列,例如用编码人或人源化MHC I多肽的核酸序列替换内源性MHC基因或其一部分。由于编码MHC I、MHC II $\alpha$ 和MHC II $\beta$ 多肽的核酸序列在染色体上彼此邻近,为了在一只动物中的MHC I和MHC II两者的人源化中取得最大成功,应按顺序靶向MHC I和MHC II基因座。因此,本文还提供了产生包含编码如本文所述的嵌合人/非人MHC I、MHC II $\alpha$ 和MHC II $\beta$ 多肽的核酸序列的经遗传修饰非人动物的方法。

[0266] 因此,提供了一种用于产生包含嵌合人/非人MHC的经遗传修饰动物的核苷酸构建体。在一个方面,核酸构建体包含:5'非人同源臂和3'非人同源臂、包含人MHC基因序列(例如人HLA-A2或人HLA-DRs基因序列)的人DNA片段以及侧接重组位点的选择盒。在一个实施方案中,人DNA片段是包含人MHC基因的内含子和外显子两者的基因组片段(例如人HLA-A2或HLA-DR2基因)。在一个实施方案中,非人同源臂与非人MHC基因座(例如,MHC I或MHC II基因座)同源。

[0267] 在一个实施方案中,5'非人同源臂和3'非人同源臂分别在非人(例如鼠科动物)MHC I类或II类基因座的5'和3'位置(例如,第一前导序列的5'和小鼠MHC I基因的外显子 $\alpha 3$ 的3',或者小鼠H-2Ab1基因的上游和小鼠H-2Ea基因的下游)处包含基因组序列。在一个实施方案中,内源性MHC I类基因座选自小鼠H-2K、H-2D和H-2L。在具体的实施方案中,内源性MHC I类基因座是小鼠H-2K。在一个实施方案中,内源性MHC II基因座选自小鼠H-2E和H-2A。在一个实施方案中,经工程改造的MHC II构建体允许替换小鼠H-2E和H-2A两种基因。在一个实施方案中,小鼠不从其H-2D基因座表达功能性内源性MHC多肽。在一些实施方案中,小鼠经工程改造而缺少内源性H-2D基因座的全部或一部分。在另一个实施方案中,小鼠在细胞表面上不表达任何功能性内源性MHC I和MHC II多肽。在一个实施方案中,小鼠在细胞表面上表达的唯一MHC I和MHC II是嵌合人/小鼠MHC I和MHC II。

[0268] 本公开还提供了用于制备经遗传工程改造的非人动物(例如,经遗传工程改造的啮齿动物如小鼠或大鼠)的方法,这种动物的基因组包含编码人或人源化 $\beta 2$ 微球蛋白多肽的 $\beta 2$ 微球蛋白基因座。在一个方面,所述方法得到经遗传工程改造的啮齿动物(例如小鼠),其基因组在内源性 $\beta 2$ 微球蛋白基因座处包含编码人或人源化 $\beta 2$ 微球蛋白多肽的核苷酸序列。在某些情况下,小鼠不从内源性小鼠 $\beta 2$ 微球蛋白基因座表达功能性小鼠 $\beta 2$ 微球蛋白。如



本文所述,在一些方面,所述方法利用例如通过VELOCIGENE<sup>®</sup>技术制备的靶向构建体,将该构建体引入ES细胞中,并例如采用VELOCIMOUSE<sup>®</sup>技术将靶向的ES细胞克隆引入小鼠胚胎中。

[0269] 还提供了用于产生经遗传工程改造的非人动物的核苷酸构建体。这种核苷酸构建体可包含:5'非人同源臂和3'非人同源臂、包含人 $\beta$ 2微球蛋白序列的人DNA片段以及侧接重组位点的选择盒。在一个实施方案中,人DNA片段是包含人 $\beta$ 2微球蛋白基因的内含子和外显子两者的基因组片段。在一个实施方案中,非人同源臂与非人 $\beta$ 2微球蛋白基因座同源。基因组片段可包含人 $\beta$ 2微球蛋白基因的外显子2、3和4。在一种情况下,基因组片段包含从5'至3':外显子2、内含子、外显子3、内含子和外显子4,全部人 $\beta$ 2微球蛋白序列。选择盒可位于构建体中 $\beta$ 2微球蛋白编码区外的任何位置,例如它可位于人 $\beta$ 2微球蛋白的外显子4的3'处。5'非人同源臂和3'非人同源臂可分别包含内源性非人 $\beta$ 2微球蛋白基因的基因组序列5'和3'。在另一个实施方案中,5'非人同源臂和3'非人同源臂分别包含内源性非人基因的外显子2的基因组序列5'和外显子4的基因组序列3'。

[0270] 本发明的另一方面涉及一种修饰非人动物(例如,啮齿动物如小鼠或大鼠)的 $\beta$ 2微球蛋白基因座以表达本文所述的人或人源化 $\beta$ 2微球蛋白多肽的方法。一种修饰非人动物(例如小鼠)的 $\beta$ 2微球蛋白基因座以表达人或人源化 $\beta$ 2微球蛋白多肽的方法包括在内源性 $\beta$ 2微球蛋白基因座处用编码人或人源化 $\beta$ 2微球蛋白多肽的核苷酸序列替换编码小鼠 $\beta$ 2微球蛋白的核苷酸序列。在这类方法的一个实施方案中,非人动物(例如小鼠)不从内源性非人(例如小鼠) $\beta$ 2微球蛋白基因座表达功能性 $\beta$ 2微球蛋白多肽。在一些具体实施方案中,编码人或人源化 $\beta$ 2微球蛋白多肽的核苷酸序列包含在人 $\beta$ 2微球蛋白基因的外显子2至4中示出的核苷酸序列。在其他实施方案中,编码人或人源化 $\beta$ 2微球蛋白多肽的核苷酸序列包含在人 $\beta$ 2微球蛋白基因的外显子2、3和4中示出的核苷酸序列。

[0271] 图2-5中示出了本文所述的人源化基因座的各种示例性实施方案。

[0272] 在完成基因靶向后,对ES细胞或经遗传修饰的非人动物进行筛选以确认目标外源核苷酸序列的成功整合或者外源多肽的成功表达。许多技术是本领域技术人员已知的,包括(但不限于)Southern印迹、长链PCR、定量PCR(例如,利用TAQMAN<sup>®</sup>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。在一个示例中,携带目标基因修饰的非人动物(例如,小鼠)可通过使用在Valenzuela et al. (2003) High-throughput engineering of the mouse genome coupled with high-resolution expression analysis, Nature Biotech. 21 (6) :652-659 (Valenzuela等人, 2003年,“小鼠基因组的高通量工程与高分辨率表达分析相结合”,《自然生物技术》,第21卷第6期,第652-659页)中描述的等位基因修饰测定法筛选小鼠等位基因的丢失和/或人等位基因的获得来鉴定。鉴定经遗传修饰动物中特定核苷酸或氨基酸序列的其他测定法是本领域技术人员已知的。

[0273] 在一些实施方案中,通过交配在本文中产生动物。

[0274] 在一个非限制性方面,例如,包含本文所述嵌合人/非人CD8和人源化MHC I和/或 $\beta$ 2微球蛋白的非人动物可通过使包含如本文所述的嵌合CD8基因座(例如,嵌合CD8 $\alpha$ 和/或 $\beta$ 基因座)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交配而产生。这种动物



也可通过将例如用于在内源性CD8基因座(例如,嵌合CD8 $\alpha$ 和/或 $\beta$ 基因座)处替换的编码嵌合CD8(例如,嵌合CD8 $\alpha$ 和/或 $\beta$ )的核苷酸序列引入包含人源化MHC I和/或 $\beta$ 2微球蛋白基因座的动物的ES细胞中;或者将编码人源化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CD8 $\alpha$ 和/或 $\beta$ 基因座)的动物的ES细胞中而产生。

[0275] 在一些实施方案中,包含嵌合CD8基因座的动物可首先与包含人源化TCR可变基因座的动物交配以产生包含人源化CD8和TCR可变区基因座的动物,然后可使其与包含人源化MHC I和/或 $\beta$ 2微球蛋白基因座的动物交配,以产生包含人源化MHC I、TCR可变基因和/或 $\beta$ 2微球蛋白基因座的动物。或者,包含人源化MHC I和/或 $\beta$ 2微球蛋白基因座的动物可分别首先与包含人源化TCR可变基因座的动物交配以产生包含人源化MHC I和TCR可变区基因座的动物,然后可使其与包含嵌合CD8基因座的动物交配,以产生包含人源化MHC I、TCR可变基因和/或 $\beta$ 2微球蛋白基因座的动物。

[0276] 在一个方面,包含嵌合人/非人CD4和人源化MHC II的非人动物可通过使包含如本文所述嵌合CD4基因座的动物与包含人源化MHC II基因座的动物交配而产生。这种动物也可通过将例如用于在内源性CD4基因座处替换的编码嵌合CD4的核苷酸序列引入包含人源化MHC II基因座的动物的ES细胞中;或者将编码人源化MHC II的核苷酸序列引入包含嵌合CD4基因座的动物的ES细胞中而产生。

[0277] 在一些实施方案中,包含嵌合CD4基因座的动物可首先与包含人源化TCR可变基因座的动物交配以产生包含人源化CD4和TCR可变区基因座的动物,然后可使其与包含人源化MHC II基因座的动物交配以产生包含人源化CD4、MHC II和TCR可变基因座的动物。或者,包含人源化MHC II基因座的动物可分别首先与包含人源化TCR可变基因座的动物交配以产生包含人源化MHC II和TCR可变区基因座的动物,然后可使其与包含嵌合CD4基因座的动物交配以产生包含人源化MHC II、TCR可变基因和/或 $\beta$ 2微球蛋白基因座的动物。

[0278] 在一些实施方案中,使包含本文所述嵌合人/非人CD8和人源化MHC I和/或 $\beta$ 2微球蛋白的非人动物与包含如本文所述的嵌合CD4基因座的动物以及包含人源化MHC II基因座的动物交配以产生包含嵌合CD4和CD8多肽及人源化MHC I(和/或 $\beta$ 2微球蛋白)和MHC II分子的非人动物。在一些实施方案中,使包含嵌合人/非人CD4和CD8多肽及人源化MHC I和MHC II分子的动物与包含人源化TCR可变结构域的动物交配以产生包含基本上人源化T细胞免疫系统(例如,嵌合人/非人CD4和CD8多肽、人源化MHC I(和/或 $\beta$ 2微球蛋白)和MHC II分子以及人源化TCR可变结构域)的动物。

[0279] 本文所述的任何经遗传修饰的非人动物(例如,小鼠)均可包含编码嵌合人/非人CD8(例如,CD8 $\alpha$ 和/或CD8 $\beta$ );嵌合人/非人CD4;人或人源化MHC I;人或人源化 $\beta$ 2微球蛋白;人或人源化MHC II(例如,MHC II $\alpha$ 和/或MHC II $\beta$ );以及人或人源化TCR(例如,TCR $\alpha$ 和/或TCR $\beta$ )的基因的一个或两个拷贝。因此,这种动物对任何或所有这些基因可以是杂合的或纯合的。

[0280] 使用产生基本上人源化T细胞免疫应答的经遗传修饰的非人动物

[0281] 包含人源化CD4和MHC II或人源化CD8和MHC I(以及 $\beta$ 2微球蛋白)或两者的经遗传修饰非人动物(例如,啮齿动物如小鼠或大鼠)以类似于人的方式向T细胞(分别向CD4+或CD8+T细胞)递呈肽,这是因为该复合体的基本上全部组分是人或人源化的。本发明的经遗传修饰的非人动物可用于研究人源化动物中人免疫系统的功能;用于鉴定引起免疫应答的

抗原和抗原表位(例如,T细胞表位,例如独特人癌症表位),例如用于开发疫苗;用于鉴定对人病原体或癌症抗原具有高亲和力的T细胞(也就是说,在人MHC I复合体的背景下以高亲和力与抗原结合的T细胞),例如用于适应性T细胞疗法;用于评估候选疫苗和其他疫苗策略;用于研究人自体免疫;用于研究人类传染病;也用于制定基于人MHC和CD4/CD8表达的更好治疗策略。

[0282] 因此,在各种实施方案中,本发明的经遗传工程改造的动物尤其可用于评估抗原在人体中引发免疫应答的能力,以及用于产生多种抗原并鉴定可用于开发人类疫苗的特定抗原。

[0283] 在一个方面,提供了用于确定肽是否在人体中引起细胞免疫应答的方法,所述方法包括将如本文所述经遗传修饰的非人动物暴露于肽中,允许非人动物产生免疫应答,以及在非人动物中检测如本文所述结合由嵌合人/非人MHC I或II分子递呈的肽序列的细胞(例如分别包含人CD8或CD4的CD8+或CD4+T细胞)。在一个实施方案中,非人动物在暴露之后包含结合肽的MHC I类限制性CD8+细胞毒性T淋巴细胞(CTL)。在另一个实施方案中,非人动物在暴露之后包含结合肽的MHC II限制性CD4+T细胞。

[0284] 在一个方面,提供了用于鉴定人T细胞表位的方法,所述方法包括将如本文所述的非人动物暴露于包含推定T细胞表位的抗原,允许非人动物产生免疫应答,从非人动物分离出结合所述表位的MHC I类或MHC II类限制性T细胞,以及鉴定由所述T细胞结合的表位。

[0285] 在一个方面,提供了用于鉴定在人体中产生T细胞应答的抗原的方法,所述方法包括将如本文所述的推定抗原暴露于小鼠中,允许小鼠产生免疫应答,以及鉴定由HLA I类或II类限制性分子结合的抗原。

[0286] 在一个方面,提供了用于确定推定抗原是否包含在暴露于人免疫系统时将产生HLA I类或II类限制性免疫应答的表位的方法,所述方法包括将如本文所述的小鼠暴露于推定抗原,以及测量小鼠中抗原特异性HLA I类或HLA II类限制性免疫应答。

[0287] 此外,本文所述经遗传工程改造的非人动物可用于鉴定识别目标抗原(例如肿瘤或另一种疾病抗原)的T细胞受体,例如高亲和力T细胞受体。所述方法可包括:将本文所述的非人动物暴露于抗原,允许非人动物对抗原产生免疫应答,从非人动物中分离出包含T细胞受体的T细胞,该T细胞受体结合由人或人源化MHC I或MHC II递呈的抗原,以及确定所述T细胞受体的序列。

[0288] 表达功能性人TCR V(D)J基因区段的多样性组库的非人动物可用于研究人类疾病。因此,在一个实施方案中,本文所述的经遗传工程改造的非人动物可表达与在人体中表达的TCR组库基本相似的TCR组库,例如本文公开的非人动物的TCR组库可来源于至少约50%、至少约60%、至少约70%、至少约75%、至少约80%、至少约85%、至少约90%、至少约95%、至少约97%或至少约99%的所有功能性人TCR $\alpha$ 、TCR $\beta$ 、TCR $\gamma$ 和/或TCR $\delta$ 基因区段。在一些实施方案中,如所公开的非人动物表达来源于

[0289] (I) 至少约50%、至少约60%、至少约70%、至少约75%、至少约80%、至少约85%、至少约90%、至少约95%、至少约97%或至少约99%的所有功能性人TCR V $\alpha$ 基因区段;

[0290] (ii) 至少约50%、至少约60%、至少约70%、至少约75%、至少约80%、至少约85%、至少约90%、至少约95%、至少约97%或至少约99%的所有功能性人TCR J $\alpha$ 基因区段;

[0291] (iii) 至少约50%、至少约60%、至少约70%、至少约75%、至少约80%、至少约85%、至少约90%、至少约95%、至少约97%或至少约99%的所有功能性人TCR V $\beta$ 基因区段;

[0292] (iv) 至少约50%、至少约60%、至少约70%、至少约75%、至少约80%、至少约85%、至少约90%、至少约95%、至少约97%或至少约99%的所有功能性人TCR D $\beta$ 基因区段;和/或

[0293] (v) 至少约50%、至少约60%、至少约70%、至少约75%、至少约80%、至少约85%、至少约90%、至少约95%、至少约97%或至少约99%的所有功能性人TCR J $\beta$ 基因区段的TCR组库。

[0294] 在一个实施方案中,小鼠产生包含所有或基本上所有功能性人TCRV $\alpha$ 基因区段,并包含所有或基本上所有功能性人TCR V $\beta$ 基因区段的T细胞组库。在一个实施方案中,本文提供的小鼠利用人TCR V $\alpha$ 和/或V $\beta$ 基因,这些基因的频率分别与由人体中人T细胞利用的人TCR V $\alpha$ 和/或V $\beta$ 基因的频率近似。检测在非人动物的TCR组库中表达的基因区段的方法包括例如流式细胞术和/或测序方法(例如,实时PCR、下一代测序等)。

[0295] 在一个实施方案中,提供了用于通过推定人治疗剂确定T细胞活化的方法,所述方法包括将如本文所述经遗传修饰的动物暴露于推定人类治疗剂(或例如,将这种动物的人或人源化MHC II-或MHC I-表达细胞暴露于推定治疗剂的肽序列中),将展示人或人源化MHC/肽复合体的经遗传修饰动物的细胞暴露于能够结合经遗传修饰动物细胞的包含嵌合人/非人(例如人/小鼠)CD4或CD8的T细胞,以及测量由经遗传修饰动物的展示肽细胞诱导的T细胞的活化。

[0296] 除了能够鉴定人病原体或肿瘤的抗原和抗原表位之外,本发明的经遗传修饰动物还可用于鉴定与人自身免疫疾病(例如I型糖尿病、多发性硬化症等)相关的自身抗原。另外,本发明的经遗传修饰动物可用于研究人自身免疫疾病的各个方面,并可用作自身免疫疾病模型。

[0297] 在各种实施方案中,本发明的经遗传修饰非人动物产生在其表面上具有人源化TCR分子的T细胞,因此,将以类似于人的方式识别由MHC复合体递呈给它们的肽。本文所述的经遗传修饰非人动物可用于研究人T细胞的发育和功能以及免疫耐受过程;用于测试人候选疫苗;用于产生对TCR基因疗法具有某些特异性的TCR;用于产生疾病相关抗原(例如肿瘤相关抗原(TAA)等)的TCR文库。

[0298] 本领域中的T细胞疗法越来越受到关注,因为T细胞(例如,细胞毒性T细胞)可涉及攻击目标抗原(例如病毒抗原、细菌抗原、肿瘤抗原等)或递呈它的细胞并导致目标抗原或递呈它的细胞受破坏。对癌症T细胞疗法的初步研究旨在从肿瘤细胞团分离出肿瘤浸润淋巴细胞(TIL;肿瘤块中大概包含T细胞反应性抗肿瘤抗原的淋巴细胞群),使用T细胞生长因子在体外扩增它们,并在称为过继性T细胞转移的过程中将它们转移回到患者身上。参见例如,Restifo et al. (2012) Adoptive immunotherapy for cancer: harnessing the T cell response, Nature Reviews 12:269-81 (Restifo等人,2012年,“用于癌症的过继性免疫疗法:利用T细胞应答”,《自然评论》,第12卷,第269-281页);Linnermann et al. (2011) T-Cell Receptor Gene Therapy: Critical Parameters for Clinical Success, J. Invest. Dermatol. 131:1806-16 (Linnermann等人,2011年,T细胞受体基因疗法:临床成

功的关键参数,《研究性皮肤病学杂志》,第131卷,第1806-1816页)。然而,这些疗法的成功迄今仅限于黑素瘤和肾细胞癌;并且TIL过继性转移并不特别针对限定的肿瘤相关抗原(TAA)。Linnermann等人,出处同上。

[0299] 已经尝试启用TCR基因疗法,在这种疗法中选择或编程T细胞以靶向目标抗原(例如TAA)。目前的TCR基因疗法依赖于对针对特定抗原(例如肿瘤相关抗原)的TCR序列的鉴定。例如,Rosenberg及其同事已经发表了几项研究,在这些研究中他们用黑素瘤相关抗原MART-1表位特异性的编码TCR $\alpha$ 和 $\beta$ 链的基因转导来源于黑素瘤患者的外周血淋巴细胞,并使用所得的扩增淋巴细胞进行过继性T细胞治疗。Johnson et al. (2009) Gene therapy with human and mouse T-cell receptors mediates cancer regression and targets normal tissues expressing cognate antigen, Blood 114:535-46 (Johnson等人,2009年,“采用人和小鼠T细胞受体的基因疗法介导癌症消退并靶向表达同源抗原的正常组织”,《血液》,第114卷:第535-546页);Morgan et al. (2006) Cancer Regression in Patients After Transfer of Genetically Engineered Lymphocytes, Science 314:126-29 (Morgan等人,2006年,“经遗传工程改造的淋巴细胞转移之后患者身上的癌症消退”,《科学》,第314卷:第126-129页)。从TIL治疗后经历肿瘤消退的患者中分离出MART-1特异性TCR。然而,这类TCR,尤其高亲合力TCR(其最有可能是治疗上有用的)的鉴定由于多数肿瘤抗原是自身抗原这一事实而变得复杂,并且靶向这些抗原的TCR通常是缺失的或主要由于免疫耐受性而具有次优亲和力。

[0300] 在各种实施方案中,本发明通过提供在其基因组中包含未重排人TCR可变基因座的经遗传工程改造的非人动物来解决该问题。本文所述的非人类动物能够产生具有人源化T细胞受体的多样性组库的T细胞。因此,本文所述的非人动物可以是人源化T细胞受体(例如,用于过继性T细胞转移中的高亲合力人源化T细胞受体)的多样性组库的来源。

[0301] 因此,在一个实施方案中,本发明提供了产生人抗原的T细胞受体的方法,所述方法包括用目标抗原对本文所述的非人动物(例如,啮齿动物如小鼠或大鼠)进行免疫,允许所述动物产生免疫应答,从所述动物中分离出目标抗原特异性的活化T细胞,以及确定由抗原特异性T细胞表达的T细胞受体的核酸序列。

[0302] 在一个实施方案中,本发明提供了产生目标抗原(例如疾病相关抗原)特异性人T细胞受体的方法,所述方法包括用目标抗原对本文所述的非人动物进行免疫;允许所述动物产生免疫应答;从所述动物中分离出对目标抗原具有反应性的T细胞;确定由T细胞表达的人TCR可变区的核酸序列;将人TCR可变区克隆到包含人TCR恒定区的核酸序列的核苷酸构建体中,以使得人TCR可变区可操作地连接至人TCR恒定区;以及从该构建体中表达目标抗原特异性人T细胞受体。在一个实施方案中,分离T细胞,确定由T细胞表达的人TCR可变区的核酸序列,将人TCR可变区克隆到包含人TCR恒定区的核酸序列的核苷酸构建体中,以及表达人T细胞受体的步骤是利用本领域技术人员已知的标准技术执行的。

[0303] 在一个实施方案中,编码目标抗原特异性T细胞受体的核苷酸序列在细胞中表达。在一个实施方案中,表达TCR的细胞选自CHO、COS、293、HeLa、PERC.6<sup>TM</sup>细胞等。

[0304] 目标抗原可以是已知引起疾病或病症或与疾病或病症相关的任何抗原(例如肿瘤相关抗原;病毒、细菌或其他致病源的抗原等)。许多肿瘤相关抗原是本领域已知的。一组肿瘤相关抗原示于《癌症免疫力》(癌症研究所杂志)肽数据库

([archive.cancerimmunity.org/peptidedatabase/Tcellepitopes.htm](http://archive.cancerimmunity.org/peptidedatabase/Tcellepitopes.htm))中。在本发明的一些实施方案中,目标抗原是人抗原,例如人肿瘤相关抗原。在一些实施方案中,抗原是细胞型特异性细胞内抗原,并且T细胞受体用于杀死表达该抗原的细胞。

[0305] 在一个实施方案中,本文提供了鉴定具有抗目标抗原(例如肿瘤相关抗原)特异性的T细胞的方法,所述方法包括用目标抗原对本文所述的非人动物进行免疫,允许这种动物产生免疫应答,并从非人动物中分离出抗原特异性T细胞。

[0306] 本发明提供了用于过继性T细胞治疗的新方法。因此,本文提供了一种治疗或减轻受试者(例如,哺乳动物受试者如人受试者)中疾病或病症(例如,癌症)的方法,所述方法包括用与疾病或病症相关的抗原对本文所述的非人动物进行免疫,允许这种动物产生免疫应答,从这种动物中分离出一群抗原特异性T细胞,以及将分离的抗原特异性T细胞输注给受试者。在一个实施方案中,本发明提供了一种治疗或减轻人受试者中疾病或病症的方法,所述方法包括用目标抗原(例如,疾病或病症相关抗原如肿瘤相关抗原)对本文所述的非人动物进行免疫,允许这种动物产生免疫应答,从这种动物中分离出一群抗原特异性T细胞,确定T细胞受体的核酸序列(例如,编码人重排TCR $\alpha$ 和/或人重排TCR $\beta$ 可变区基因的第一和/或第二核酸序列;由抗原特异性T细胞表达的编码人重排TCR $\delta$ 可变区基因或TCR $\gamma$ 可变区基因的第三和/或第四核酸序列),将T细胞受体的核酸序列(例如,分别编码人重排TCR $\alpha$ 可变区基因、人重排TCR $\beta$ 可变区基因、TCR $\delta$ 可变区基因或TCR $\gamma$ 可变区基因的第一、第二、第三和/或第四核酸序列)克隆到表达载体(例如逆转录病毒载体)中,例如,任选地,其中分别编码人重排TCR $\alpha$ 可变区基因、人重排TCR $\beta$ 可变区基因、TCR $\delta$ 可变区基因或TCR $\gamma$ 可变区基因的第一、第二、第三和/或第四核酸序列分别与人TCR $\alpha$ 恒定基因、人TCR $\beta$ 恒定基因、TCR $\delta$ 恒定基因或TCR $\gamma$ 恒定基因进行框内克隆,将载体引入来源于受试者的T细胞以使得T细胞表达抗原特异性T细胞受体,以及将T细胞输注给受试者。在一个实施方案中,T细胞受体核酸序列在引入来源于受试者的T细胞中之前被进一步人源化,例如,将编码非人恒定区的序列修饰成更加类似于人TCR恒定区(例如,用人恒定区替换非人恒定区)。在一些实施方案中,疾病或病症是癌症。在一些实施方案中,抗原特异性T细胞群在输注给受试者之前进行扩增。在一些实施方案中,在输注抗原特异性T细胞之前将受试者的免疫细胞群免疫耗竭。在一些实施方案中,抗原特异性TCR是高亲合力TCR,例如肿瘤相关抗原的高亲合力TCR。在一些实施方案中,T细胞是细胞毒性T细胞。在其他实施方案中,疾病或病症是由病毒或细菌引起的。

[0307] 在另一个实施方案中,疾病或病症是自身免疫疾病。TREG细胞是保持对自身抗原的耐受性并抑制病理自身反应性的T细胞亚群。因此,本文还提供了治疗自身免疫疾病的方法,这些方法依赖于在本文所述的本发明非人动物中产生抗原特异性TREG细胞。

[0308] 本文还提供了一种治疗或改善受试者中疾病或病症(例如,癌症)的方法,所述方法包括将受疾病或病症影响的细胞(例如癌细胞)从受试者引入非人动物中,允许动物对这些细胞产生免疫应答,从动物中分离出对这些细胞具有反应性的一群T细胞,确定由T细胞表达的T细胞受体可变结构域的核酸序列,将T细胞受体可变结构域编码序列克隆到载体中(例如,框内克隆并可操作地连接至人TCR恒定基因),将载体引入来源于受试者的T细胞中,以及将携带T细胞受体的受试者T细胞输注给受试者。

[0309] 本文还提供了使用如本文所述的非人动物制备编码人TCR可变结构域(例如,TCR $\alpha$ 和/或 $\beta$ 可变结构域)的核酸序列。在一个实施方案中,提供了一种用于制备编码人TCR可变

结构域的核酸序列的方法,所述方法包括用目标抗原对如本文所述的非人动物进行免疫,允许非人动物对目标抗原产生免疫应答,以及从中获得编码与目标抗原结合的人TCR可变结构域的核酸序列。在一个实施方案中,所述方法还包括制备编码可操作地连接至非人TCR恒定区的人TCR可变结构域的核酸序列,其包括从本文所述的非人动物中分离出T细胞并从中获得编码连接至非人恒定区TCR恒定区的TCR可变结构域的核酸序列,以及将编码TCR可变结构域的核酸序列(例如,分别编码人重排TCR $\alpha$ 可变区基因、人重排TCR $\beta$ 可变区基因、TCR $\delta$ 可变区基因或TCR $\gamma$ 可变区基因的第一、第二、第三或第四核酸序列)与合适的人恒定区(例如,分别与人TCR $\alpha$ 恒定区基因、人TCR $\beta$ 恒定区基因、TCR $\delta$ 恒定区基因或TCR $\gamma$ 可变区基因)进行框内克隆。

[0310] 因此,本文提供了TCR可变区核酸序列,诸如重排TCR可变核酸序列,例如在本文所述的非人类动物中产生并分别由例如人重排V $\alpha$ /J $\alpha$ 基因序列和重排人V $\beta$ D $\beta$ J $\beta$ 基因序列编码的重排TCR $\alpha$ 和/或TCR $\beta$ 可变区核酸序列。此外,提供了由这类重排TCR可变区核酸序列编码的TCR可变区氨基酸序列。在本文所述的非人动物中获得的这类重排TCR可变区核酸序列(TCR $\alpha$ 和/或TCR $\beta$ 可变区核酸序列)可以与人TCR恒定区(TCR $\alpha$ 和/或TCR $\beta$ 恒定区)可操作地连接的方式进行克隆,并用于本文所述的各种用途,例如在人体中作为人治疗剂。

[0311] 本文还提供了使用如本文所述的非人动物制备人治疗剂的用途,包括用目标抗原(例如,肿瘤相关抗原)对所述非人动物进行免疫,允许所述非人动物产生免疫应答,从所述动物中获得对目标抗原具有反应性的T细胞,从这些T细胞中获得编码与目标抗原结合的人源化TCR蛋白或人TCR可变结构域的核酸序列,以及在人治疗剂中利用编码人源化TCR蛋白或人TCR可变结构域的核酸序列。

[0312] 因此,还提供了一种用于制备人治疗剂的方法,所述方法包括用目标抗原对如本文所述的非人动物进行免疫,允许非人类动物产生免疫应答,从所述动物中获得对目标抗原具有反应性的T细胞,从这些T细胞中获得编码与目标抗原结合的人源化T细胞受体的核酸序列,以及在人治疗剂中采用人源化(或全人)T细胞受体。

[0313] 在一个实施方案中,人治疗剂是携带目标核酸序列的T细胞(例如,人T细胞如来源于人受试者的T细胞),(例如,转染或转导或以其他方式与目标核酸一起引入)以使得这些T细胞表达对目标抗原具有亲和力的人源化TCR蛋白。在一个方面,采用治疗剂的受试者需要针对特定疾病或病症进行治疗,并且抗原与这种疾病或病症相关。在一个方面,T细胞是细胞毒性T细胞,抗原是肿瘤相关抗原,并且该疾病或病症是癌症。在一个方面,T细胞来源于受试者。

[0314] 在另一个实施方案中,人治疗剂是T细胞受体。在一个实施方案中,治疗剂受体是可溶性T细胞受体。已经作出大量努力以产生用于治疗剂的可溶性T细胞受体或TCR可变区。可溶性T细胞受体的产生取决于获得重排TCR可变区。一种方法是设计出包含TCR $\alpha$ 和TCR $\beta$ 的单链TCR,并以与scFv免疫球蛋白类似的形式,经由连接子将它们融合在一起(参见例如,国际申请No.WO 2011/044186)。如果与scFv相似,所得scTv将提供热稳定且可溶形式的TCR $\alpha$ / $\beta$ 结合蛋白。替代方法包括:设计出具有TCR $\beta$ 恒定结构域的可溶性TCR(参见例如,Chung et al., (1994) Functional three-domain single-chain T-cell receptors, Proc.Natl.Acad.Sci.USA.91:12654-58 (Chung等人,1994年,“功能性三结构域单链T细胞受体”,《美国国家科学院院刊》,第91卷,第12654-12658页));以及将非天然二硫键改造成

TCR恒定结构域之间的界面(在Boulter and Jakobsen (2005) Stable, soluble, high-affinity, engineered T cell receptors: novel antibody-like proteins for specific targeting of peptide antigens, *Clinical and Experimental Immunology* 142:454-60 (Boulter和Jakobsen, 2005年, “稳定、可溶、高亲和力的经工程改造T细胞受体: 用于特异性靶向肽抗原的新型抗体样蛋白”, 《临床与实验免疫学》, 第142卷: 第454-460页) 中综述; 另见美国专利No. 7,569,664)。已经描述了其他形式的可溶性T细胞受体。本文所述的非人动物可用于确定以高亲和力结合至目标抗原的T细胞受体的序列, 并随后基于该序列设计出可溶性T细胞受体。

[0315] 来源于由非人动物表达的TCR受体序列的可溶性T细胞受体可用于阻断目标蛋白(例如病毒、细菌或肿瘤相关蛋白)的功能。或者, 可溶性T细胞受体可融合到可以杀死感染细胞或癌细胞的部分, 例如细胞毒性分子(例如化学治疗剂)、毒素、放射性核素、前体药、抗体等。可溶性T细胞受体也可融合到免疫调节分子, 例如细胞因子、趋化因子等。可溶性T细胞受体也可融合到免疫抑制分子, 例如抑制T细胞杀死其他携带由T细胞识别的抗原的细胞的分子。这类融合到免疫抑制分子的可溶性T细胞受体可用于例如阻断自身免疫。可以融合到可溶性T细胞受体的各种示例性免疫抑制分子在Ravetch and Lanier (2000) *Immune Inhibitory Receptors*, *Science* 290:84-89 (Ravetch和Lanier, 2000年, “免疫抑制受体”, 《科学》, 第290卷: 第84-89页) 中综述, 该文献以引用方式并入本文。

[0316] 本发明还提供了用于研究在人TCR背景下的免疫学应答(包括人TCR重排、T细胞发育、T细胞活化、免疫学耐受等)的方法。

[0317] 还提供了测试候选疫苗的方法。在一个实施方案中, 本文提供了确定疫苗是否会活化免疫学应答(例如, T细胞增殖、细胞因子释放等)并导致效应子以及记忆T细胞(例如中心记忆T细胞和效应记忆T细胞)产生的方法。

[0318] 在一个方面, 提供了一种体外制剂, 其包含在其表面上具有嵌合CD8蛋白的T细胞和结合嵌合CD8的第二细胞。在一个实施方案中, 第二细胞是表达MHC I多肽(例如嵌合人/非人MHC I蛋白)的细胞。在一个实施方案中, 第一细胞表面上的嵌合CD8与第二细胞表面上的嵌合MHC I相互作用。在一个实施方案中, 嵌合CD8蛋白保留与内源性胞质分子例如内源性胞质信号传导分子(例如内源性Lck等)的相互作用。

[0319] 在一个方面, 提供了一种体外制剂, 其包含在其表面上具有嵌合CD4蛋白的T细胞和结合嵌合CD4的第二细胞。在一个实施方案中, 第二细胞是表达MHC II多肽(例如嵌合人/非人MHC II蛋白)的细胞, 例如APC。在一个实施方案中, 第一细胞表面上的嵌合CD4与第二细胞表面上的嵌合MHC II相互作用。在一个实施方案中, 嵌合CD4蛋白保留与内源性胞质分子例如内源性胞质信号传导分子(例如内源性Lck等)的相互作用。

[0320] 本文所述的经遗传修饰动物(即包含人或人源化T细胞辅助受体(例如, 嵌合人/非人CD4或CD8), 任选地还包含人或人源化MHC II或I蛋白的动物)的其他用途从本公开中将是显而易见的。

#### [0321] 实施例

[0322] 提供了下面的实施例以便向本领域的技术人员描述如何制备本发明的组合物和如何使用本发明的方法, 并且这些实施例并非旨在限制本发明人视作其发明的范围。已尽量确保所使用的数字(例如量、温度等)的准确性, 但应考虑到一些实验误差和偏差。实施例

不包括本领域普通技术人员所熟知的常规方法的详细描述(分子克隆技术等)。除非另有说明,份数为重量份,分子量为平均分子量,温度以摄氏度表示,压力为或接近大气压。

#### [0323] 实施例1:人源化MHC小鼠的产生

[0324] 涉及对包含人源化MHC I和MHC II基因座并具有对应和另外的内源性MHC I和MHC II基因座缺失(HLA-A2/H-2K、HLA-DR2/H-2E、H-2A-del、H-2D-del)的小鼠进行工程改造的各个步骤在图3A中示出。对这些步骤的详细描述如下。

#### [0325] 实施例1.1:人源化MHC I小鼠的产生和表征

[0326] 此前已在美国专利公布No. 20130111617中描述了人源化MHC I小鼠的产生,该专利以引用方式并入本文。简言之,通过采用VELOCIGENE<sup>®</sup>技术由人和小鼠细菌人工染色体(BAC) DNA构建独特的靶向载体,在单一步骤中将小鼠H-2K基因人源化(参见例如美国专利No. 6,586,251和Valenzuela et al. (2003) High-throughput engineering of the mouse genome coupled with high-resolution expression analysis. Nat. Biotech. 21 (6): 652-659 (Valenzuela等人, 2003年, “小鼠基因组的高通量工程与高分辨率表达分析相结合”, 《自然生物技术》, 第21卷第6期: 第652-659页))。来自小鼠BAC克隆RP23-173k21 (英杰生命技术有限公司 (Invitrogen)) 的DNA通过同源重组进行修饰, 以用编码人HLA-A基因的 $\alpha 1$ 、 $\alpha 2$ 和 $\alpha 3$ 亚单元的人基因组DNA替换编码小鼠H-2K基因的 $\alpha 1$ 、 $\alpha 2$ 和 $\alpha 3$ 结构域的基因组DNA (图2A)。

[0327] 具体地讲, 使用包含潮霉素盒的靶向载体, 在单个靶向事件中用编码人HLA-A\*0201基因的 $\alpha 1$ 、 $\alpha 2$ 和 $\alpha 3$ 结构域的人基因组DNA替换编码小鼠H-2K基因的 $\alpha 1$ 、 $\alpha 2$ 和 $\alpha 3$ 亚单元的基因组序列, 该潮霉素盒侧接loxP位点, 其中5'小鼠同源臂包含具有5'非翻译区(UTR)的小鼠H-2K基因座的序列5', 并且3'小鼠同源臂包含小鼠H-2K $\alpha 3$ 编码序列的基因组序列3'。

[0328] 用于靶向从5'到3'的内源性H-2K基因座的最终构建体包括(1) 5'同源臂, 其含有包括5'UTR的内源性H-2K基因的约200bp小鼠基因组序列5', (2) 约1339bp的人基因组序列, 其包括HLA-A\*0201前导序列、HLA-A\*0201前导/ $\alpha 1$ 内含子、HLA-A\*0201 $\alpha 1$ 外显子、HLA-A\*0201 $\alpha 1$ - $\alpha 2$ 内含子、HLA-A\*0201 $\alpha 2$ 外显子、 $\alpha 2$ - $\alpha 3$ 内含子约316bp的5'端, (3) 5' loxP位点, (4) 潮霉素盒, (5) 3' loxP位点, (6) 约580bp的人基因组序列, 其包括 $\alpha 2$ - $\alpha 3$ 内含子约304bp的3'端、HLA-A\*0201 $\alpha 3$ 外显子, 以及(7) 3'同源臂, 其含有包括小鼠H-2K $\alpha 3$ 和跨膜编码序列之间的内含子的约200bp小鼠基因组序列。在靶向载体的5'处的小鼠/人序列连接处的149个核苷酸的序列在SEQ ID NO: 90中示出, 并且在靶向载体的3'处的人/小鼠序列连接处的159个核苷酸的序列在SEQ ID No: 91中示出。与该靶向载体的同源重组生成经修饰小鼠H-2K基因座, 其含有编码可操作地连接至内源性小鼠H-2K跨膜结构域和胞质结构域编码序列的HLA-A\*0201基因的 $\alpha 1$ 、 $\alpha 2$ 和 $\alpha 3$ 结构域的人基因组DNA, 其翻译后导致形成嵌合的人/小鼠MHC I类蛋白。可采用本领域已知的各种方法随后除去存在于靶向构建体中的选择盒。

[0329] 所靶向BAC DNA用于电穿孔小鼠F1H4Es细胞以生成经修饰的ES细胞, 用于产生在有核细胞的表面上表达嵌合MHC I类蛋白的小鼠(例如, T和B淋巴细胞、巨噬细胞、中性粒细胞)(参见例如, 图3A所示方案中的步骤1)。通过定量TAQMAN<sup>™</sup>测定法鉴定包含人HLA序列插入的ES细胞(Valenzuela等人, 2003年, 出处同上)。

[0330] 为了产生表达嵌合MHC I的小鼠, 使用本文所述的所靶向ES细胞作为供体ES细胞, 并通过VELOCIMOUSE<sup>®</sup>方法将其引入8细胞期小鼠胚胎中(参见例如, 美国专利No. 7,



294,754和Poueymirou et al. (2007) F0 generation mice that are essentially fully derived from the donor gene-targeted ES cells allowing immediate phenotypic analyses Nature Biotech. 25 (1) :91-99 (Poueymirou等人, 2007年, “基本上完全来源于供体基因靶向的ES细胞的F0代小鼠实现了即时表型分析”, 《自然生物技术》, 第25卷第1期, 第91-99页)。采用等位基因修饰测定法 (Valenzuela等人, 出处同上) (检测是否存在独特的人HLA-A\*0201基因序列), 通过基因分型鉴定独立携带嵌合MHC I类基因的VELOCIMICE<sup>®</sup> (完全来源于供体ES细胞的F0小鼠)。将通过该方法产生的杂合小鼠繁育为纯合小鼠。使用HLA-A和H-2K特异性抗体, 通过流式细胞术确认嵌合HLA-A2/H-2K的表达。

[0331] 包含嵌合HLA-A2/H-2K的上述所靶向ES细胞用于实施例1.2-1.3中所述的另外的遗传工程步骤, 以产生包含人源化MHC I和MHC II基因座两者但缺少内源性MHC I和MHC II基因座的小鼠 (见图3A)。

[0332] 实施例1.2: 包含MHC I和MHC II基因座缺失的小鼠ES细胞的产生

[0333] 在美国专利申请号20130111616中描述了内源性MHC II基因座的缺失, 该专利申请通过引用方式并入本文。简言之, 采用VELOCIGENE<sup>®</sup>遗传工程技术制备用于引入内源性MHC II类H-2Ab1、H-2Aa、H-2Eb1、H-2Eb2和H-2Ea基因的缺失的靶向载体 (参见例如, 美国专利No. 6,586,251和Valenzuela等人, 出处同上)。细菌人工染色体 (BAC) RP23-458i22 (英杰生命技术有限公司 (Invitrogen)) DNA经修饰以使内源性MHC II类基因H-2Ab1、H-2Aa、H-2Eb1、H-2Eb2和H-2Ea缺失。

[0334] 具体地讲, 通过分别来自H-2Ab1基因5'位置和H-2Ea基因3'位置的小鼠BAC DNA的PCR得到上游和下游同源臂。这些同源臂用于通过细菌同源重组 (BHR) 制备缺失了包含MHC II类基因座的基因H-2Ab1、H-2Aa、H-2Eb1、H-2Eb2和H-2Ea约79kb的RP23-458i22的盒。该区域用侧接1ox2372位点的新霉素盒替换。从5'到3'的最终靶向载体包括: 包含内源性MHC II类基因座的H-2Ab1基因的小鼠基因组序列5'的26kb同源臂、5' 1ox2372位点、新霉素盒、3' 1ox2372位点以及包含内源性MHC II类基因座的H-2Ea基因的小鼠基因组序列3'的63kb同源臂。

[0335] (上述) BAC DNA靶向载体用于电穿孔包含人源化MHC I基因座的小鼠ES细胞 (来自上述实施例1.1; 参见例如, 图3A中的步骤2) 以生成包含内源性MHC II类基因座缺失 (H-2A和H-2E均已缺失) 的经修饰ES细胞。使用TAQMAN<sup>™</sup>探针, 通过定量PCR测定法鉴定含有所缺失内源性MHC II类基因座的阳性ES细胞 (Lie and Petropoulos (1998) Curr. Opin. Biotechnology 9:43-48 (Lie和Petropoulos, 1998年, 《生物技术当前述评》, 第9卷: 第43-48页))。使用引物5111U F (CAGAACGCCAGGCTGTAAC; SEQ ID NO:1) 和5111U R (GGAGAGCAGGGTCAGTCAAC; SEQ ID NO:2) 以及探针5111U P (CACCGCCACTCACAGCTCCTTACA; SEQ ID NO:3), 通过PCR确认所缺失基因座的上游区域, 而使用引物5111D F (GTGGGCACCATCTTCATCATTC; SEQ ID NO:4) 和5111D R (CTTCCTTTCCAGGGTGTGACTC; SEQ ID NO:5) 以及探针5111D P (AGGCCTGCGATCAGGTGGCACCT; SEQ ID NO:6) 确认所缺失基因座的下游区域。使用引物NEOF (GGTGGAGAGGCTATTCGGC; SEQ ID NO:7) 和NEOR (GAACACGGCGGCATCAG; SEQ ID NO:8) 以及探针NEOP (TGGGCACAACAGACAATCGGCTG; SEQ ID NO:9) 确认来自靶向载体的新霉素盒是否存在。在上游缺失点上的核苷酸序列 (SEQ ID No:10) (其表示与存在于缺失点处的盒序列连续连接的缺失点的内源性小鼠序列上游, 包括在下面的括号中) 包括下

列：(TTTGTAACA AAGTCTACCC AGAGACAGAT GACAGACTTC AGCTCCAATG CTGATTGGTT CCTCACTTGG GACCAACCCT)ACCGGTATAA CTTCGTATAA GGTATCCTAT ACGAAGTTAT ATGCATGGCC TCCGCGCCGG。在下游缺失点(SEQ ID NO:11)上的核苷酸序列(其表示与缺失点的内源性小鼠序列下游邻接的盒序列,包括在下面的括号中)包括下列:CGACCTGCAG CCGGCGCGCC ATAACCTCGT ATAAGGTATC CTATACGAAG TTATCTCGAG (CACAGGCATT TGGGTGGGCA GGGATGGACG GTGACTGGGA CAATCGGGAT GGAAGAGCAT AGAATGGGAG TTAGGGAAGA)。

[0336] 生成上述包含MHC I人源化和内源性MHC II缺失两者的ES细胞之后,使用CRE除去loxed新霉素盒(参见例如,图3A中的步骤3)。具体地讲,将编码Cre重组酶的质粒电穿孔到ES细胞中以除去新霉素盒。也可使用本领域已知的其他方法除去新霉素盒。

[0337] 为了使小鼠H-2D基因座缺失,使用BHR对小鼠BAC克隆bMQ-218H21(桑格研究所(Sanger Institute))进行修饰,用从5'至3'含有:与5' loxp位点同框的LacZ基因、UbC启动子、新霉素基因和3' loxp位点的6,085bp盒替换3756bp的H2-D基因(从ATG起始密码子到下游3bp的TGA终止密码子,小鼠H-2D的外显子1-8)。

[0338] (上述)BAC DNA靶向载体用于电穿孔上述包含人源化MHC I基因座和小鼠MHC II缺失的小鼠ES细胞(参见例如,图3A中的步骤4)。如上所述,通过定量PCR测定法鉴定含有所缺失内源性H-2D基因座的阳性ES细胞。表2示出用于定量PCR测定法的引物和探针。

[0339] 表2:用于检测所缺失H-2D基因座的TAQMAN™等位基因丢失测定法引物和探针

名称(位置)	正向引物	反向引物	探针
[0340] 5152 mTU (上游)	CGAGGAGCCCCGGT ACA (SEQ ID NO:12)	AAGCGCACGAAGTC CTTGTT (SEQ ID NO:13)	CTCTGTCGGCTATGT GG (SEQ ID NO:14)
5152 mTD (下游)	GGACTCCCAGAATCT CCTGAGA (SEQ ID NO:15)	GAGTCATGAACCAT CACTGTGAAGA (SEQ ID NO:16)	TGGTGGGTTGCTGGA A (SEQ ID NO:17)

[0341] 实施例1.3:嵌合人/小鼠MHC II基因座的引入

[0342] 为了生成包含人源化HLA-DR2/H-2E的载体,首先,根据2014年9月30日公布的美国专利No.8,847,005(该专利以引用方式并入本文)中的描述对小鼠H-2Ea基因进行修饰,以生成包含编码嵌合H-2Ea/HLA-DRA1\*01蛋白的序列的载体。

[0343] 对于小鼠H-2Eb基因,合成的人HLA-DR2β链(DRB1\*1501)用于生成包含DRB1\*02(1501)外显子和内含子的载体,并使用细菌同源重组交换到包含嵌合H-2Ea/HLA-DRA1\*01蛋白的载体中。基本上按照美国专利公布No.20130185820和美国专利No.8,847,005中所述那样对H-2Eb1基因进行修饰,这些专利中每一篇均以引用方式并入本文。使用潮霉素选择盒。

[0344] 所得HLA-DR2/H-2E大靶向载体(LTVEC)示于图2B和3B中。所得LTVEC的各种核苷酸序列连接(例如小鼠/人序列连接、人/小鼠序列连接,或具有选择盒的小鼠或人序列的连接)汇总在下表3中并在序列列表中列出;它们的位置示于图3B的示意图中。在下表3中,除了标有星号(\*,参见表格说明)的序列之外,小鼠序列为常规字体;人序列在括号中;Lox序列为斜体;并且在克隆步骤中引入的限制性位点和其他基于载体的序列(例如,多克隆位点等)以粗体表示。

[0345] 表3:嵌合HLA-DR2/H-2E基因座的核苷酸序列连接

[0346]

SEQ ID NO:	核苷酸序列
18	CTGTTTCTTC CCTAACTCCC ATTCTATGCT CTTCCATCCC GA <b>CCGCGG</b> (CCCA ATCTCTCTCC ACTACTTCCT GCCTACATGT ATGTAGGT)
19	(CAAGGTTTCC TCCTATGATG CTTGTGTGAA ACTCGG) <b>GGCC GGCC</b> AGCATTTAAC AGTACAGGGA TGGGAGCACA GCTCAC
20*	(GAAAGCAGTC TTCCCAGCCT TCACACTCAG AGGTACAAAT) CCCCATTTTC ATATTAGCGA TTTTAATTTA TTCTAGCCTC
21*	TCTTCCCTAA CTCCCATTCT ATGCTCTTCC ATCCCGA <b>CCG CGG</b> (CCCAATC TCTCTCCACT ACTTCCTGCC TACATGTATG)
22	GAGTTCCTCCATCACTTCACTGGGTAGCACAGCTGTAAGTGTCCAGCCTG (TCCTGGGCTGCAGGTGGTGGGCGTTGCGGGTGGGGCCGGTTAAGGTTCCA)
23	(TCCCACATCCTATTTTAATTTGCTCCATGTTCTCATCTCCATCAGCACAG) <b>CTCGAG ATAACCTTCGTATAATGTATGCTATACGAAGTTAT ATGCATGGCC</b>
24	<i>ATACGAAGTTAT</i> <b>GCTAGTAACTATAACGGTCCTAAGGTAGCGAGTGGCTT</b> ACAGGTAGGTGCGTGAAGCTTCTACAAGCACAGTTGCCCCCTGGGAAGCA

[0347] 标有星号的序列是C57BL/6-BALB/c连接序列,其中C57BL/6序列在括号中。在嵌合H-2Ea基因的克隆过程中,H-2Ea的C57BL/6等位基因的外显子1和内含子1的其余部分来自H-2Ea的BALB/c等位基因的等同2616bp区域替换。这是因为H-2Ea的C57BL/6等位基因的外显子1含有使该基因无功能的缺失,而H-2Ea的BALB/c等位基因的外显子1有功能。关于更详细的描述,参见美国专利No.8,847,005,该专利以引用方式并入本文。

[0348] 上述所靶向BAC DNA用于电穿孔包含人源化MHC I (HLA-A2) 以及MHC II和H-2D缺失的小鼠ES细胞,以生成经修饰的ES细胞,用于产生表达嵌合MHC I和MHC II基因并缺少功能性内源小鼠H-2E、H-2A、H-2K和H-2D基因座的小鼠(参见例如,图3A中的步骤5)。使用表4中的引物和探针,通过定量PCR (TAQMAN™) 测定法鉴定含有人HLA序列插入的ES细胞。

[0349] 表4:用于检测MHC I和MHC II基因座人源化的TAQMAN™引物和探针序列

[0350]

名称 (位置)	正向引物	反向引物	探针
Hyg 盒	TGCGGCCGATCT TAGCC (SEQ ID NO:25)	TTGACCGATTCC TTGCGG (SEQ ID NO:26)	ACGAGCGGGTTC GGCCCATTC (SEQ ID NO:27)
7092 hTUP1 (DRB1*1501 的外显子 2)	CCCCACAGCACG TTTCCT (SEQ ID NO:28)	CGTCCCATTGAA GAAATGACACT (SEQ ID NO:29)	TGGCAGCCTAAG AGG (SEQ ID NO:30)
7092 hTUP2 (DRB1*1501 的外显子 2)	CCCCACAGCACG TTTCCT (SEQ ID NO:31)	ACCCGCTCCGTC CCATT (SEQ ID NO:32)	AGCCTAAGAGG GAGTGTC (SEQ ID NO:33)
7092 hTDP1 (DRB1*1501 的外显子 3)	AGACCCTGGTGA TGCTGGAA (SEQ ID NO:34)	CGCTTGGGTGCT CCACTT (SEQ ID NO:35)	TCGAAGTGGAGA GGTTTA (SEQ ID NO:36)
7092 hTDP2 (DRB1*1501 的外显子 3)	TGGAATGGAGTG AGCAGCTTT (SEQ ID NO:37)	GCACGGTCCCCT TCTTAGTG (SEQ ID NO:38)	TGACTTCCTAAA TTTCTC (SEQ ID NO:39)
hDRAIU (DRA 的外显子 2)	CTGGCGGCTTGA AGAATTTGG (SEQ ID NO:40)	CATGATTTCCAG GTTGGCTTTGTC (SEQ ID NO:41)	CGATTTGCCAGC TTTGAGGCTCAA GG (SEQ ID NO:42)
1751jxn2 <sup>1</sup> (等位基因丢失测定 法, 仅存在于 H-2A 和 H-2E 缺失中的序列)	CCTCACTTGGGA CCAACCCTA (SEQ ID NO:43)	TTGTCCCAGTCA CCGTCCAT (SEQ ID NO:44)	TGCATCTCGAGC ACAGGCATTTGG (SEQ ID NO:45)

[0351] <sup>1</sup>除此之外的所有序列均用于等位基因获得测定法。

[0352] 可通过本领域技术人员已知的方法除去选择盒。例如,携带嵌合人/小鼠MHC I类基因座的ES细胞可用表达Cre的构建体转染,以便除去通过插入靶向构建体引入的“loxed”选择盒(参见例如,图3A中的步骤6)。选择盒可任选地通过繁育到表达Cre重组酶的小鼠中除去。任选地,选择盒保留在小鼠体内。

[0353] 使用本文所述用于各个修饰的引物/探针组,采用上述定量TAQMAN<sup>®</sup>测定法验证含有本文所述所有修饰(图3A的HLA-A2/H-2K、HLA-DR2/H-2E、H-2A-del、H-2D-del)的所靶向ES细胞。使用另一引物/探针组来确定,在盒缺失步骤中,是否由于lox位点以相反取向存在而未生成反向克隆。

[0354] 上述所靶向的ES细胞用作供体ES细胞,并通过VELOCIMOUSE<sup>®</sup>方法被引入8细胞期小鼠胚胎中(参见例如,美国专利No.7,294,754和Poueymirou等人,2007年,出处同上)。采用等位基因修饰测定法(Valenzuela等人,出处同上)(检测是否存在独特的人基因序列),通过基因分型鉴定独立携带嵌合MHC I类和MHC II基因的VELOCIMICE<sup>®</sup>(完全来源于供体ES细胞的F0小鼠)。所得小鼠中MHC基因座的基因型的示意图示于图3C中(\*\*代表所有小鼠品系中均不存在的H-2L基因)。使用HLA-DR2和HLA-A2特异性抗体来确认嵌合人/小鼠MHC I和MHCII两种蛋白的表达。将杂合小鼠繁育为纯合小鼠。

[0355] 实施例1.4:人源化β2微球蛋白小鼠的生成

[0356] 美国专利申请公布No.20130111617中描述了β2微球蛋白小鼠的生成,该专利申请以引用方式并入本文。简言之,通过采用VELOCIGENE<sup>®</sup>技术,由人和小鼠细菌人工染色

体 (BAC) DNA 构建独特的靶向载体, 在单一步骤中将小鼠  $\beta 2$  微球蛋白 ( $\beta 2m$ ) 基因人源化 (参见例如, 美国专利 No. 6, 586, 251 和 Valenzuela 等人, 出处同上)。

[0357] 具体地讲, 通过细菌同源重组生成包含来自 RPCI-23 文库的 BAC 克隆 89C24 (英杰生命技术有限公司 (Invitrogen)) 的小鼠  $\beta 2m$  上游和下游同源臂的靶向载体。小鼠同源臂经工程改造而侧接从外显子 2 延伸至非编码外显子 4 下游约 267 个核苷酸的 2.8kb 人  $\beta 2m$  DNA 区段 (图 2C)。侧接重组酶识别位点 (例如, loxP 位点) 的药物选择盒 (新霉素) 经工程改造到靶向载体中以允许后续选择。将最终的靶向载体线性化并电穿孔到 F1H4 小鼠 ES 细胞系中 (Valenzuela 等人, 出处同上)。

[0358] 通过 VELOCIMOUSE<sup>®</sup> 方法, 将去除药物盒的所靶向 ES 细胞克隆 (通过引入 Cre 重组酶) 引入 8 细胞期小鼠胚胎 (参见例如, 美国专利 No. 7, 294, 754 和 Poueymirou 等人, 出处同上)。采用等位基因修饰测定法, 通过筛选小鼠等位基因丢失和人等位基因获得来鉴定携带人源化  $\beta 2m$  基因的 VELOCIMICE<sup>®</sup> (完全来自供体 ES 细胞的 F0 小鼠) (Valenzuela 等人, 出处同上)。将杂合小鼠繁育为纯合小鼠。使用人  $\beta 2$  微球蛋白特异性抗体, 通过流式细胞术确定人  $\beta 2$  微球蛋白的表达。

[0359] 实施例 2: 人源化 T 细胞受体小鼠的产生

[0360] 采用 VELOCIGENE<sup>®</sup> 遗传工程技术制备包含内源性 TCR ( $\alpha$  或  $\beta$ ) 可变基因座缺失和内源性 V 和 J 或 V、D 和 J 区段替换的小鼠 (参见例如, 美国专利 No. 6, 586, 251 和 Valenzuela, D.M. 等人, 2003 年, 出处同上), 其中使用细菌同源重组来源于 BAC 文库的人序列用于制备大靶向载体 (LTVEC), 该大靶向载体包含侧接靶向臂的人 TCR 可变基因座的基因组片段以使 LTVEC 靶向小鼠 ES 细胞中的内源性小鼠 TCR 可变基因座。对 TCR $\alpha$  和  $\beta$  基因座人源化的详细描述在美国专利 9, 113, 616 中有所描述, 该专利以引用方式并入本文。根据 Valenzuela 等人, 将 LTVEC 线性化并电穿孔到小鼠 ES 细胞系中。选择具有潮霉素或新霉素抗性的 ES 细胞, 并筛选小鼠等位基因丢失或人等位基因获得。

[0361] 通过 VELOCIMOUSE<sup>®</sup> 方法将所靶向 ES 细胞克隆引入 8 细胞期 (或更早) 小鼠胚胎中 (Poueymirou, W.T. 等人, 2007 年, 出处同上)。采用等位基因修饰测定法, 通过筛选内源性 TCR 可变等位基因丢失和人等位基因获得来鉴定携带人源化 TCR 基因座的 VELOCIMICE<sup>®</sup> (完全来源于供体 ES 细胞的 F0 小鼠) (Valenzuela 等人, 出处同上)。将 F0 幼崽进行基因分型并繁育为纯合幼崽。如本文所述制备对于人源化 TCR $\alpha$  和/或 TCR $\beta$  可变基因座是纯合的小鼠。

[0362] 实施例 2.1: TCR $\alpha$  基因座的人源化

[0363] 采用在图 4A 中汇总并在美国专利 No. 9, 113, 616 中描述的渐进式人源化策略, 用对应于 54V 和 61J 人 TCR $\alpha$  区段的 1 兆碱基 DNA 替换在小鼠 TCR $\alpha$  基因座处对应于 110V 和 60J 小鼠区段的 1.5 兆碱基 DNA。用于 TCR $\alpha$  基因座的渐进式人源化策略的各种靶向载体的接合核酸序列汇总在表 5 中, 并包括在序列表中。

[0364] 表 5: 用于各种 TCR $\alpha$  基因座靶向载体的接合核酸序列

[0365]	MAID 号	SEQ ID NO	说明
	1626	46	TCR $\alpha$ 可变基因座上游小鼠序列的 3' 端和 loxP-Ub-Hyg-loxP 盒的

[0366]

		5'端之间的接合核酸序列。
	47	<i>loxP</i> -Ub-Hyg- <i>loxP</i> 盒的 3'端和人 TCRV $\alpha$ 40-TCRV $\alpha$ 41-TCRJ $\alpha$ 1 插入的 5'端之间的接合核酸序列, 包括 AsiSI 位点。
	48	人 TCRV $\alpha$ 40-TCRV $\alpha$ 41-TCRJ $\alpha$ 1 插入的 3'端和人 TCR $\alpha$ 可变基因座下游小鼠序列的 5'端之间的接合核酸序列, 包括 NotI 位点。
1767	49	TCR $\alpha$ 可变基因座上游小鼠序列的 3'端和 <i>loxP</i> -Ub-Neo- <i>loxP</i> 盒的 5'端之间的接合核酸序列。
	50	<i>loxP</i> -Ub-Neo- <i>loxP</i> 盒的 3'端和人 TCRV $\alpha$ 35-TCRV $\alpha$ 39 插入的 5'端之间的接合核酸序列, 包括 AsiSI 位点。
1979	51	TCR $\alpha$ 可变基因座上游小鼠序列的 3'端和 <i>frt</i> -Pgk-Hyg- <i>frt</i> 盒的 5'端之间的接合核酸序列。
	52	<i>frt</i> -Pgk-Hyg- <i>frt</i> 盒的 3'端和人 TCRV $\alpha$ 22-TCRV $\alpha$ 34 插入的 5'端之间的接合核酸序列, 包括 AsiSI 位点。
1769	53	TCR $\alpha$ 可变基因座上游小鼠序列的 3'端和 <i>loxP</i> -Ub-Neo- <i>loxP</i> 盒的 5'端之间的接合核酸序列。
	54	<i>loxP</i> -Ub-Neo- <i>loxP</i> 盒的 3'端和人 TCRV $\alpha$ 13-2-TCRV $\alpha$ 21 插入的 5'端之间的接合核酸序列, 包括 AsiSI 位点。
1770	55	TCR $\alpha$ 可变基因座上游小鼠序列的 3'端和 <i>loxP</i> -Ub-Hyg- <i>loxP</i> 盒的 5'端之间的接合核酸序列。
	56	<i>loxP</i> -Ub-Hyg- <i>loxP</i> 盒的 3'端和人 TCRV $\alpha$ 6-TCRV $\alpha$ 8-5 插入的 5'端之间的接合核酸序列, 包括 AsiSI 位点。
1771	57	TCR $\alpha$ 可变基因座上游小鼠序列的 3'端至 <i>loxP</i> -Ub-Neo- <i>loxP</i> 盒的 5'端之间的接合核酸序列。
	58	<i>loxP</i> -Ub-Neo- <i>loxP</i> 盒的 3'端和人 TCRV $\alpha$ 1-1-TCRV $\alpha$ 5 插入的 5'端之间的接合核酸序列, 包括 AsiSI 位点。

[0367] 人TCR $\alpha$ 可变区区段按IMGT数据库中的方式编号。每一接合区处的至少100bp (距每端约50bp) 包括在序列表中。

[0368] 首先, 通过同源重组对来自小鼠BAC克隆RP23-6A14 (英杰生命技术有限公司 (Invitrogen)) 的DNA进行修饰, 然后将其用作靶向载体以用Ub-潮霉素盒, 然后用*loxP*位点替换内源性小鼠TCR $\alpha$ 基因座的TCRAJ1-TCRAJ28区域。通过同源重组对来自小鼠BAC克隆RP23-117i19 (英杰生命技术有限公司 (Invitrogen)) 的DNA进行修饰, 然后将其用作靶向载体以用PGK-新霉素盒, 然后用*loxP*位点替换内源性小鼠TCR $\alpha$ 和 $\delta$ 基因座的TCRAV1周围 (并包括) 约15kb区域。通过本领域已知的核型分析和筛选方法 (例如, TAQMAN™) 确认携带双靶向染色体的ES细胞 (即, 用这两种靶向载体靶向的单个内源性小鼠TCR $\alpha$ 基因座)。用CRE重组酶处理经修饰的ES细胞, 从而介导两个*loxP*位点之间区域 (即, 由从TCRAV1到TCRAJ1内源性小鼠TCR $\alpha$ 基因座组成的区域) 的缺失, 并且只留下单个*loxP*位点、新霉素盒以及小鼠恒定区和增强子区域。这一策略生成了缺失的小鼠TCR $\alpha$ / $\delta$ 基因座 (MAID 1540, 图4A, 第二个图)。

[0369] 用于TCR $\alpha$ 的第一人靶向载体具有来自CTD2216p1和CTD2285m07BAC克隆 (英杰生命技术有限公司 (Invitrogen)) 的191,660bp人DNA, 这些克隆包含前两个连续人TCR $\alpha$ V基因区段 (TRAV40和41) 和61TCR $\alpha$ J (50个功能性) 基因区段。通过同源重组对该BAC进行修饰以包含用于连接3'小鼠同源臂的TCRAJ1基因区段下游 (3') 403bp的Not1位点以及用于连接5'小鼠同源臂的5' AsiSI位点。两个不同的同源臂用于连接至该人片段: 3' 同源臂包含来自RP23-6A14BAC克隆的内源性小鼠TCR $\alpha$ 序列, 并且5' 同源臂包含来自小鼠BAC克隆RP23-117i19的

小鼠TCR $\alpha$ V的内源性TCR $\alpha$ 序列5'。使用该小鼠-人嵌合BAC作为靶向载体(MAID 1626),用于在小鼠TCR $\alpha$ 基因座处初步插入人TCR $\alpha$ 基因区段,加上上游loxP-ub-潮霉素-loxP盒。用于MAID 1626靶向载体的接合核酸序列(SEQ ID NO:46至SEQ ID NO:48)在表5中有所描述。

[0370] 随后,制备了一系列使用相同小鼠5'臂的人靶向载体,该小鼠5'臂包含来自带交替loxP-新霉素-loxP和loxP-潮霉素-loxP(或用于MAID 1979的frt-潮霉素-frt)选择盒的小鼠BAC克隆RP23-117i19的小鼠TCR $\alpha$ V的内源性TCR $\alpha$ 序列5'。这些特定构建体在美国专利No.9,113,616中有所描述,并示于图4A中,其中每个插入的接合序列均包括在表5和序列表中。最终TCR $\alpha$ 基因座含有5' loxP-ub-新霉素-loxP盒,加上可操作地连接至小鼠TCR $\alpha$ 恒定基因和增强子的总共54个人TCR $\alpha$ V(45个功能性)和61个人TCR $\alpha$ J基因区段。用于MAID 1771靶向载体的接合核酸序列(SEQ ID NO:57和SEQ ID NO:58)在表5中有所描述。

[0371] 在任何渐进性人源化步骤中,通过用Cre或Flp重组酶进行缺失来除去选择盒。另外,可将人TCR $\delta$ 基因座重新引入TCR $\alpha$ 序列中。

[0372] 实施例2.2:TCR $\beta$ 可变基因座的人源化

[0373] 采用在图4B中汇总并在美国专利No.9,113,616中详细描述的进行式人源化策略,用对应于67V、2D和14J人TCR $\beta$ 区段的0.6兆碱基DNA替换在小鼠TCR $\beta$ 基因座处对应于33V、2D和14J小鼠区段的0.6兆碱基DNA。用于TCR $\beta$ 基因座的进行式人源化策略的各种靶向载体的接合核酸序列汇总在表6中,并包括在序列表中。

[0374] 表6:用于各种TCR $\beta$ 基因座靶向载体的接合核酸序列

[0375]

MAID 号	SEQ ID NO	说明
1625	59	TCR $\beta$ 可变基因座上游小鼠序列的 3'端（上游小鼠胰蛋白酶原基因附近）和 <i>frit</i> -Ub-Neo- <i>frit</i> 盒的 5'端之间的接合核酸序列。
	60	<i>frit</i> -Ub-Neo- <i>frit</i> 盒的 3'端和人 TCRV $\beta$ 18-TCRV $\beta$ 29-1 插入的 5'端之间的接合核酸序列。
	61	人 TCRV $\beta$ 18-TCRV $\beta$ 29-1 插入的 3'端和小鼠 TCRV $\beta$ 区段下游小鼠序列的 5'端（下游小鼠胰蛋白酶原基因附近）之间的接合核酸序列。
1715	62	下游小鼠胰蛋白酶原基因的 3'与人 TCRD $\beta$ 1-TCRJ $\beta$ 1-1-TCRJ $\beta$ 1-6 插入的 5'端之间的接合核酸序列，包括 <i>I</i> ceI 位点。
	63	人 TCRD $\beta$ 1-TCRJ $\beta$ 1-1-TCRJ $\beta$ 1-6 插入的 3'端和 <i>loxP</i> -Ub-Hyg- <i>loxP</i> 盒的 5'端之间的接合核酸序列。
	64	<i>loxP</i> -Ub-Hyg- <i>loxP</i> 盒的 3'端和小鼠 C $\beta$ 1 基因附近小鼠序列的 5'端之间的接合核酸序列。
	65	小鼠 C $\beta$ 1 基因附近小鼠序列的 3'端和人 TCRD $\beta$ 2-TCRJ $\beta$ 2-1-TCRJ $\beta$ 2-7 插入的 5'端之间的接合核酸序列，包括 <i>NotI</i> 位点。
	66	人 TCRD $\beta$ 2-TCRJ $\beta$ 2-1-TCRJ $\beta$ 2-7 插入的 3'端和 TCR $\beta$ 可变基因座下游小鼠序列的 5'端（C $\beta$ 2 小鼠序列附近）之间的接合核酸序列。
1791	67	TCR $\beta$ 可变基因座上游小鼠序列的 3'端（上游小鼠胰蛋白酶原基因附近）和 <i>frit</i> -Ub-Hyg- <i>frit</i> 盒的 5'端之间的接合核酸序列。
	68	<i>frit</i> -Ub-Hyg- <i>frit</i> 盒的 3'端和人 TCRV $\beta$ 6-5-TCRV $\beta$ 17 插入的 5'端之间的接合核酸序列。
1792	69	TCR $\beta$ 可变基因座上游小鼠序列的 3'端（上游小鼠胰蛋白酶原基因附近）和 <i>frit</i> -Ub-Neo- <i>frit</i> 盒的 5'端之间的接合核酸序列。
	70	<i>frit</i> -Ub-Hyg- <i>frit</i> 盒的 3'端和人 TCRV $\beta$ 1-TCRV $\beta$ 12-2 插入的 5'端之间的接合核酸序列。
6192	71	小鼠 C $\beta$ 2 基因附近小鼠序列的 3'端和人 TCRBV30 外显子 2 序列的 5'端之间的接合核酸序列。
	72	人 TCRBV30 外显子 1 序列的 3'端和 TCR $\beta$ 基因座下游小鼠序列的 5'端之间的接合核酸序列。

[0376] 人TCR $\beta$ 可变区区段按IMGT数据库中的方式编号。每一接合区处的至少100bp (距每端约50bp) 包括在序列表中。

[0377] 具体地讲,通过同源重组对来自小鼠BAC克隆RP23-153p19 (英杰生命技术有限公司 (Invitrogen)) 的DNA进行修饰,然后将其用作靶向载体,以用PGK-neo盒,然后用loxP位点替换内源性小鼠TCR $\beta$ 基因座中正好3'胰蛋白酶原 (TRY) 基因簇上游17kb区域 (包括TCRBV30)。通过同源重组对来自小鼠BAC克隆RP23-461h15 (英杰生命技术有限公司 (Invitrogen)) 的DNA进行修饰,然后将其用作靶向载体,以用Ub-潮霉素盒,然后用loxP位点替换内源性小鼠TCR $\beta$ 基因座中5'胰蛋白酶原 (TRY) 基因簇下游8355bp区域 (包括TCRBV2和TCRBV3)。通过本领域已知的核型分析和筛选方法 (例如,TAQMAN<sup>TM</sup>) 确认携带双靶向染色体的ES细胞 (即,用这两种靶向载体靶向的单个内源性小鼠TCR $\beta$ 基因座)。用CRE重组酶处理经修饰的ES细胞,从而介导5'和3'loxP位点之间区域 (由从TCRBV2到TCRBV30内源性小鼠TCR $\beta$ 基因座组成) 的缺失,并且只留下单个loxP位点、潮霉素盒以及小鼠TCRBD、TCRBJ、恒定区序列和增强子序列。一个小鼠TCRV $\beta$ 留在胰蛋白酶原基因簇的5'上游,一个小鼠TCRV $\beta$ 留在小鼠E $\beta$ 的下游,如图4B中所示。



[0378] 用于TCR $\beta$ 的第一人靶向载体具有来自CTD2559j2BAC克隆(英杰生命技术有限公司(Invitrogen))的125,781bp人DNA,该克隆包含前14个连续人TCRBV基因区段(TRBV18-TRBV29-1);用于MAID 1625靶向载体的接合核酸序列(SEQ ID NO:59至SEQ ID NO:61)在表6中有所描述。

[0379] 为了用人TCR $\beta$ D和J区段替换小鼠TCR $\beta$ D和J区段,来自小鼠BAC克隆RP23-302p18(英杰生命技术有限公司(Invitrogen))和来自人BAC克隆RP11-701D14(英杰生命技术有限公司(Invitrogen))的DNA通过同源重组进行修饰,然后用作靶向载体(MAID 1715)到包含上述TCRBV小基因座(即MAID 1625)的ES细胞中。这种修饰用约25425bp含有人TCRBD1-J1、loxP Ub-潮霉素-loxP盒、小鼠恒定区1、人TCRBD2-J2的序列替换内源性小鼠TCR $\beta$ 基因座中约18540bp的区域(从3'胰蛋白酶原基因的多聚腺苷酸下游100bp到包括小鼠TCRBD1-J1、小鼠恒定区1和小鼠TCRBD2-J2的D2簇中J区段下游100bp)。通过本领域已知的核型分析和筛选方法(例如,TAQMAN™)确认携带双靶向染色体的ES细胞(即,用这两种靶向载体靶向的单个内源性小鼠TCR $\beta$ 基因座)。用CRE重组酶处理经修饰的ES细胞,从而介导潮霉素盒的缺失,只留下D1J簇中人J区段下游的单个loxP位点。用于MAID 1715靶向载体的接合核酸序列(SEQ ID NO:62至SEQ ID NO:66)在表6中有所描述。

[0380] 随后,制备了一系列使用相同小鼠5'臂的人靶向载体,该小鼠5'臂包含来自带交替选择盒的小鼠BAC克隆RP23-461h15的上游小鼠胰蛋白酶原基因周围的内源性TCR $\beta$ 序列。这些特定构建体在美国专利No.9,113,616中有所描述,并示于图4B中,其中每个插入的接合序列均包括在表6和序列表中。

[0381] 最后,生成含有可操作地连接至小鼠TCR $\beta$ 恒定基因和增强子的总共66个人TCRBV(47个功能性)和人TCR $\beta$ D和J区段(MAID 1792)的人TCR $\beta$ 小基因座。用于MAID 1792靶向载体的接合核酸序列(SEQ ID NO:69和SEQ ID NO:70)在表6中有所描述。

[0382] 小鼠TCRBV31位于TCRBC2(第二TCR $\beta$ 恒定区序列)的3'约9.4kb处,并且与其他TCRBV区段处于相反的取向。等同人V区段是TCRBV30,其位于人TCR $\beta$ 基因座中相似的位置处。为了将TCRBV31人源化,通过细菌同源重组对含有小鼠TCRBV31的小鼠BAC克隆进行修饰以制备LTVEC MAID 6192。从外显子1中的起始密码子开始、内含子、3' UTR和TCRBV31的重组信号序列(RSS)的整个编码区用同源人TCRBV30序列替换。图4B示出了位于hTCRBV30基因的外显子1和外显子2之间的内含子中的选择盒。

[0383] 用于MAID 6192靶向载体的接合核酸序列(SEQ ID NO:71和SEQ ID NO:72)在表6中有所描述。将MAID 6192DNA电穿孔到MAID 1792ES细胞中,并筛选这些细胞的小鼠TCRB31等位基因丢失和人TCRB30等位基因获得。

[0384] 使用类似的工程改造策略任选地使其余5'小鼠TCRBV区段缺失。

[0385] 在任何上述步骤中,通过用Cre或Flp重组酶进行缺失来除去选择盒。

[0386] 使对于人源化TCR $\alpha$ 可变基因座是纯合的小鼠与对于人源化TCR $\beta$ 可变基因座是纯合的小鼠交配以形成包含人源化TCR $\alpha$ 和TCR $\beta$ 可变基因座的子代。将子代繁育为相对于人源化TCR $\alpha$ 和人源化TCR $\beta$ 基因座是纯合的子代。

[0387] 经证实,包含人源化TCR $\alpha$ 和TCR $\beta$ 可变基因座的小鼠经历正常T细胞发育并包含表达来源于多个可变基因区段的可变结构域的T细胞受体。

[0388] 实施例3:T细胞辅助受体基因座的人源化

[0389] 在美国专利申请公布No.20140245466中详细描述了CD4和CD8基因座(CD8 $\alpha$ 和CD8 $\beta$ 基因座两者)的人源化,该专利申请全文以引用方式并入本文。

### [0390] 实施例3.1:CD4基因座的人源化

[0391] 具体地讲,通过采用VELOCIGENE<sup>®</sup>技术由人和小鼠细菌人工染色体(BAC)DNA构建独特的靶向载体,在单一步骤中将小鼠CD4基因座人源化(参见例如,美国专利No.6,586,251和Valenzuela等人,2003年,出处同上)。为了产生靶向载体,如美国专利申请公布No.20140245466中详细描述那样实施使用细菌人造染色体(BAC)DNA的一系列细菌同源重组(BHR)以及其他工程改造步骤。

[0392] 将人CD4靶向载体用NotI线性化,并电穿孔到F1H4小鼠ES细胞中。采用等位基因修饰测定法(Valenzuela等人,检测是否存在新霉素盒和人CD4基因),通过基因分型鉴定携带人源化CD4基因座以及小鼠CD4基因的一个拷贝的所靶向ES细胞。

[0393] 由于将人源化CD4靶向载体成功引入ES细胞中所得的最终人源化CD4基因座示于图5A。在人内含子3-lox-neo盒接合区(盒的5'端)上的序列如SEQ ID No:75所示,在lox-neo盒-人内含子3接合区(盒的3'端)上的序列如SEQ ID NO:76所示;两个序列也列于表7中。人源化CD4片段的完整核酸序列(包括在图5A中示出的pgk-neo盒)如SEQ ID NO:77所示。pgk-neo盒跨越SEQ ID No:77的残基307-2176,两个lox位点位于残基267-300和2182-2215处,而人序列跨越残基1-234和2222-18263。完整人源化CD4蛋白的氨基酸序列如SEQ ID No:78所示,其中人序列跨越氨基酸27-319(如SEQ ID NO:79所示)。

[0394] 表7. 嵌合CD4靶向载体的接合区序列

	接合区	序列	SEQ ID NO
[0395]	5'小鼠/人接合区	AGGGGAAACCCGCAAAGGATGGGACATAGGGAGACAGCTGT TAACATCTGAAACATGACCTTCTTTTCTGTGCAGCACAACCTCC TAGCTGTCACTCAAGGG(AAGAAAGTGGTGCTGGGCAAAAAA GGGGATACAGTGGAAGTACCTGTACAGCTTCCCAGAAGAA GAGCATACAATTCCACTGGAAAACTCCAACCAGAT)	73
	3'人/小鼠接合区	(CTGGTCACCTGGATGAAGTGAGGGAGGGCCCTCTGGGTTTGGG GCTGGTTTGAAGTGAAGATGAGACATCCATGAGCCAGCCTGGGGCTGG CTTCACTGAAGATC)ATCTATGTCGGGTGCGGAGAAAGAGGT AATGAAATGGCACATGCTATGTACAACTCTATTGCTGAGCAG CACCCAGTCCTGAGCTGGCTCTGAATTGAGGGTGAATTCACA CATTCTCCCCAACATCTATAATCTGG	74
[0396]	人/5' lox位点	(TATGGAGTGAAAGCCTTTGGTGTCTGAGATCTGGTCTTAGT TAAACTCTGGGATC)GGCGCGCCGAATTCCTGCAGCCCGGGCTCG AGATAACTTCGTATAATGTATGCTATACGAAGTTATATGCATCCGGG TAGGGGAGGCGCTTTTCCC	75
	3' lox位点/人	AGTATTGTTTTGCCAAGTTCTAATTCATCAGACCTCGACCTGCAGC CCTAGATAACTTCGTATAATGTATGCTATACGAAGTTATCCTAGG(CC AGAGGGCTTGGGTTGACAGAACTCAGTGGCATTCTTATCCAG AGTTTCTCTACACC)	76

[0397] 人序列在括号中,并且含有限制性内切酶位点(PI-Sce I)的序列为粗体。选择盒序列为斜体。

[0398] 通过将表达Cre重组酶的质粒电穿孔到含有人源化CD4基因座的ES细胞中除去

Floxed新霉素抗性盒。

[0399] 通过基因分型(检测是否不存在新霉素盒,是否存在人CD4基因的一个拷贝和小鼠CD4基因的一个拷贝)鉴定携带不含抗性标记的人源化CD4基因座的所靶向ES细胞。

[0400] 上述所靶向的ES细胞用作供体ES细胞,并通过VELOCIMOUSE<sup>®</sup>方法被引入8细胞期小鼠胚胎中(参见例如,美国专利No.7,294,754和Poueymirou et al.2007年,出处同上)。采用等位基因修饰测定法(Valenzuela等人,出处同上)(检测是否存在独特的人CD4基因序列),通过基因分型鉴定独立携带嵌合CD4基因的VELOCIMICE<sup>®</sup>(完全来源于供体ES细胞的F0小鼠)。使用抗人CD4抗体检测T细胞表面上的人源化CD4蛋白的表达。将本文所述的对于人源化CD4蛋白是杂合的小鼠繁育为纯合小鼠。

[0401] 实施例3.2:CD8基因座的人源化

[0402] CD8 $\alpha$ 和CD8 $\beta$ 基因共同位于基因组中,例如,在小鼠6号染色体上,这些基因位于彼此远离约37kb的位置。由于紧密连接,通过首先引入一个基因(例如CD8 $\beta$ ),然后引入第二个基因(例如CD8 $\alpha$ )而进行顺序靶向。美国专利申请公布No.20140245466中描述了具体人源化的详细步骤,该专利申请以引用方式并入本文。

[0403] 简言之,通过采用VELOCIGENE<sup>®</sup>技术,由小鼠细菌人工染色体(BAC)DNA构建独特的靶向载体,在单一步骤中将小鼠CD8 $\beta$ 基因座人源化。通过BHR对来自BAC RP23-431M6的DNA进行修饰,产生大靶向载体(LTVEC)MAID 1737,以包含同源人序列对编码CD8胞外域的小鼠外显子2-3(从内含子1中的5'接合区到内含子3中的3'接合区)的替换(图5B)。在内含子3中的3'接合区处插入loxP-Ub-Hyg盒。在所得载体的各个接合区处的核苷酸序列列于表8中并示于序列表中。人源化CD8 $\beta$ 蛋白的完整氨基酸序列如SEQ ID No:83所示;其中人序列横跨第15-165位氨基酸(如SEQ ID No:84所示)。

[0404] 表8.嵌合CD8 $\beta$ 靶向载体的接合区序列

接合区	序列	SEQ ID NO
内含子1中的小鼠/人	TGTTTGCCTGTGACATGA <sup>ACTCATTGTGACACAAA</sup> CCACTGTGCTAGGGGGGATCC <sup>ACTAGTAACGGC</sup> CGCCAGTGTGCTGGAATTCGCCC(TCGCAAGGG CCAGGCATATAAGTACACAATAAACAAATGGCAG CTCTCTCC)	80
内含子3中的人/5' lox位点	(CCCCTCCTTCCTTCCCCAGGCACTTTCCAAGTGTC AACTCTAGAGCCTAT) <sup>CGCGGCCGCACCGGTATA</sup> ACTTCGTATAATGTATGCTATACGAAGTTAT	81
内含子3中的3' lox位点/小鼠	ATAACTTCGTATAATGTATGCTATACGAAGTTAT <sup>GTCTG</sup> ACGTAGCCTATTTCTCTAGATCCAAAATGATGACA ACAAAAGGTACCTTGTG	82

[0406] 人序列在括号中,lox位点为斜体,并且限制性内切酶位点、多克隆位点和载体衍生序列为粗体。

[0407] 将靶向载体电穿孔到F1H4小鼠ES细胞中。采用等位基因修饰测定法(Valenzuela等人,检测是否存在人CD8 $\beta$ 基因),通过基因分型鉴定携带人源化CD8 $\beta$ 基因座的所靶向ES细胞。

[0408] 通过采用VELOCIGENE<sup>®</sup>技术,由小鼠细菌人工染色体(BAC)DNA构建独特的靶向载体,在单一步骤中也将小鼠CD8 $\alpha$ 基因座人源化。通过BHR对来自BAC RP23-431M6的DNA

进行修饰,产生大靶向载体 (LTVEC) MAID 1738,以包含同源人序列 (从人外显子2中的5' 接合区到内含子3中的3' 接合区 (图5A)) 对编码CD8 $\alpha$ 胞外域的小鼠外显子1-2 (从小鼠外显子1中Ala密码子27处的5' 接合区到小鼠内含子2中的3' 接合区) 的替换。这保留了外显子1起始处的小鼠前导序列。在人/小鼠序列的3' 接合区处插入lox2372-Ub-Neo盒。在所得载体的各个接合区处的核苷酸序列列于表9中并示于序列表中。人源化CD8 $\alpha$ 多肽的完整氨基酸序列如SEQ ID No:88所示;其中人序列横跨第28-179位氨基酸 (如SEQ ID NO:89所示)。

[0409] 表9. 嵌合CD8 $\alpha$ 靶向载体的接合区序列

	接合区	序列	SEQ ID NO
[0410]	外显子 1 (小鼠) 和外显子 2 (人) 处的小鼠/人	TGAACCTGCTGCTGCTGGGTGAGTCGATTATCCTGGGGAGTGGAGAAGCT(AGGCCGAGCCAGTTCGGGTGTCGCCGCTGGATCGGACCTGGAACCTGGG)	85
	人/5' lox 2372 位点	(ATGCCAGGGACAGCCCTGATACTGTAGGTAGAGTCAAGGGCTGTCCAAGT) <b>ACCGGTATAACTTCGTATAAGGTATCCTATACGAAGTTAT</b>	86
	3' lox 2372 位点/小鼠	<b>ATAACTTCGTATAAGGTATCCTATACGAAGTTATCTCGACCTGATCTTGGAGGGAGACCTGGACCGGGAGACGTGCTGGGGGCAGGGTT</b>	87

[0411] 人序列在括号中,lox位点为斜体,并且限制性内切酶位点、多克隆位点和载体衍生序列为粗体。

[0412] 将上述人源化CD8 $\alpha$ 靶向载体电穿孔到包含人源化CD8b基因座的小鼠ES细胞中,以产生包含人源化CD8b和CD8a基因座的经修饰ES细胞 (图5B)。采用等位基因修饰测定法 (Valenzuela等人),通过基因分型鉴定携带人源化CD8a和CD8b基因座的所靶向ES细胞。

[0413] 上述所靶向的ES细胞用作供体ES细胞,并通过**VELOCIMOUSE**<sup>®</sup>方法被引入8细胞期小鼠胚胎中 (参见例如,美国专利No.7,294,754和Poueymirou等人,出处同上)。采用等位基因修饰测定法 (Valenzuela等人,出处同上) (检测是否存在独特的人CD8b和CD8a基因序列),通过基因分型鉴定携带嵌合CD8b基因和嵌合CD8a基因的**VELOCIMICE**<sup>®</sup> (完全来源于供体ES细胞的F0小鼠)。

[0414] 可通过本领域技术人员已知的方法除去CD8 $\alpha$ 和CD8 $\beta$ 基因座中的选择盒。将如本文所述对于人源化CD8 $\alpha$ 和CD8 $\beta$ 基因座是杂合的小鼠繁育为纯合小鼠。使用抗人CD8抗体检测T细胞表面上的人源化CD8 $\alpha$ 和CD8 $\beta$ 的表达。

#### [0415] 实施例4: 包含人源化细胞免疫系统组分的小鼠的产生

[0416] 为了产生包含人源化细胞免疫系统组分的小鼠,可将对于各种组分 (例如MHC I、MHC II $\alpha$ 和 $\beta$ 、TCR $\alpha$ 和 $\beta$ 、CD4、CD8 $\alpha$ 和 $\beta$ 以及 $\beta$ 2M) 的人源化是纯合的小鼠以任何组合一起交配,产生具有人源化T细胞免疫应答不同组分的小鼠。例如,可使包含人源化MHC I的小鼠与包含人源化 $\beta$ 2M的小鼠交配,以产生表达人源化MHC I/ $\beta$ 2M的小鼠。采用本领域已知的方法,使对于各种组分 (例如MHC I、MHC II $\alpha$ 和 $\beta$ 、TCR $\alpha$ 和 $\beta$ 、CD4、CD8 $\alpha$ 和 $\beta$ 以及 $\beta$ 2M) 的人源化是纯合的小鼠一起交配,获得包含所有九种人源化的小鼠 (“TM I/II B C4/8”小鼠)。使用本领域已知的方法将小鼠繁育为纯合小鼠。或者,可将包含每种人源化基因的靶向载体通过顺序靶向引入相同ES细胞中以获得包含所有九种人源化的ES细胞,并通过以上实施例1-3中所述

的VELOCIMOUSE<sup>®</sup>方法将所得的ES细胞引入8细胞期小鼠胚胎中。

[0417] 实施例5:包含人源化细胞免疫系统组分的小鼠的表征

[0418] 对于人源化MHC I、MHC II $\alpha$ 和 $\beta$ 、TCR $\alpha$ 和 $\beta$ 、CD4、CD8 $\alpha$ 和 $\beta$ 是纯合的以及对于人源化 $\beta$ 2M是纯合的小鼠进行了表征。具体地讲,从小鼠中摘下脾和胸腺,并获得单细胞悬浮液。将悬浮液在4℃下以1200rpm离心5分钟,让细胞沉淀,并用4mL ACK裂解缓冲液(GIBCO)裂解每个组织的细胞,以裂解红血球。将细胞过滤通过细胞过滤网,离心沉淀,重悬于培养基中并计数。

[0419] 使用荧光染料偶联抗体:抗小鼠CD3 (17A2,BD)、抗小鼠CD19 (1D3,BD)、抗小鼠F4/80 (BM8,Biolegend)、抗小鼠CD8 $\alpha$  (53-6.7,BD)、抗小鼠CD4 (RM4-5,eBioscience)、抗人CD8 $\alpha$  (SK1,BD)和抗人CD4 (RPA-T4,BD),通过FACS分析图6A-6C和图9A-9C中示出的CD19、CD3、CD4和CD8 $\alpha$ 的细胞表面表达。使用荧光染料偶联抗体:抗小鼠CD19 (6D5,Biolegend)、抗小鼠F4/80 (BM8,Biolegend)、抗小鼠H2Db (KH95,Biolegend)、抗人HLA-A2 (BB7.2,BD)、抗人HLA-DR (G46-6,BD)、抗人B2-微球蛋白 (2M2,Biolegend)和抗小鼠I<sup>A</sup>I<sup>E</sup> (M5/114.15.2,eBioscience),通过FACS分析图7A-7F和图10A-10F中小鼠H2Db、人HLA分子(HLA-A2,B2m和HLA-DR)和小鼠MHC I<sup>A</sup>I<sup>E</sup>分子的细胞表面表达。使用荧光染料偶联抗体:抗小鼠CD3 (17A2,Biolegend)、抗小鼠CD4 (GK1.5,eBiosciences)、抗小鼠CD8 $\alpha$  (53-6.7,BD 2)、抗小鼠CD8 $\beta$  (H35-17.2,eBioscience)、抗人CD4 (OKT4,eBioscience)、抗人CD8 $\alpha$  (RPA-T8,BD 6)、抗人CD8 $\beta$  (2ST8.5H7,BD),通过FACS分析图7G和图10G中小鼠和人CD4和CD8的细胞表面表达。使用抗FoxP3 (FJK-16s,eBioscience)和抗CD25 (PC61,Biolegend),通过FACS分析图8或图11中示出的FoxP3和CD25的细胞表面表达。使用抗CD44 (IM7,BD)和抗CD62L (MEL-14,Biolegend)分析图9D-9E中示出的CD44和CD62L的细胞表面表达。

[0420] 使用BD Fortessa进行所有流式细胞分析。使用FlowJo分析数据。

[0421] 图6A-6C、图7A-7G和图8中描述了在胸腺中的表达。对照小鼠和人源化TM I/II B C4/8小鼠中胸腺细胞和CD3<sup>+</sup>细胞的绝对数目以及胸腺T细胞的总体发育是相当的(数据未示出)。图6A表明,具有人源化细胞免疫系统的小鼠(TM I/II B C4/8)胸腺中B细胞和T细胞的比例与存在于对照小鼠中的比例近似。将TM I/II B C4/8小鼠的胸腺中F4/80细胞的频率和数目与对照小鼠进行比较(图6B,数据未示出)。另外,使人源化CD4和CD8对于所有九种细胞免疫基因人源化的小鼠(TM I/II B C4/8)的胸腺细胞上表达,这类似于非人源化对照小鼠中小鼠CD4和CD8的表达(图6C)。使人源化 $\beta$ 2M在人源化TM I/II B C4/8小鼠B细胞和巨噬细胞的表面上表达,而对照小鼠的B细胞和巨噬细胞中缺少它的表达(图7A和7B)。类似地,人源化MHC I和II存在于人源化TM I/II B C4/8小鼠的B细胞和巨噬细胞两者的表面上(图7C和7D),而小鼠MHC I类和II类分子检测不到(图7E和7F)。使人源化CD4、CD8 $\alpha$ 和CD8 $\beta$ 在获自人源化TM I/II B C4/8小鼠的CD3<sup>+</sup>胸腺细胞表面上表达,而对照小鼠的CD3<sup>+</sup>胸腺细胞中缺少它们的表达(图7G)。人源化TM I/II B CD4/8表达调节性T细胞(Treg)(图8)、NK细胞(CD335<sup>+</sup>CD3<sup>-</sup>)和单核细胞(CD11b<sup>+</sup>)(数据未示出)。

[0422] 在图9A-9D和10A-10G中示出了在脾脏中的表达。对于细胞免疫系统组分人源化的小鼠(TM I/II B CD4/8)的脾脏包括相当多绝对数目的CD3<sup>+</sup>细胞,以及几乎正常比例的B和T细胞(图9A,数据未示出)。将TM I/II B C4/8小鼠的脾脏中F4/80细胞的频率和数目与对照小鼠进行比较(图9B,数据未示出)。对于细胞免疫系统组分人源化的小鼠(TM I/II B

CD4/8) 在CD3<sup>+</sup>脾细胞上表达人源化CD4和CD8 $\alpha$  (图9C)。人源化TM I/II/BCd4/8小鼠包括效应记忆 (CD44<sup>+</sup>CD62L<sup>-</sup>) CD4<sup>+</sup>和CD8<sup>+</sup>T细胞和中枢记忆 (CD44<sup>+</sup>CD62L<sup>+</sup>) CD8<sup>+</sup>T细胞 (图9D和9E)。

[0423] 如图10A和10B所示,使人源化 $\beta$ 2M在人源化TM I/II B C4/B小鼠的脾脏中B细胞和巨噬细胞的表面上表达,而在对照小鼠脾脏的B细胞和巨噬细胞中缺少它的表达和小鼠MHC分子的表达。类似地,人源化MHC I和II存在于人源化TM I/II B C4/B小鼠的脾脏中B细胞和巨噬细胞两者的表面上 (图10C和图10D),而小鼠MHC I类和II类分子检测不到 (图10E和图10F)。使人源化CD4、CD8 $\alpha$ 和CD8 $\beta$ 在获自人源化TM I/II BC4/8小鼠的CD3<sup>+</sup>脾细胞的表面上表达,而对照小鼠的CD3<sup>+</sup>脾细胞中缺少它们的表达 (图10G)。与对照小鼠相比,TM I/II B C4/8小鼠具有接近正常的脾调节性T细胞表达 (图11),并表达脾NK细胞 (CD335<sup>+</sup>CD3<sup>-</sup>) 和单核细胞 (CD11b<sup>+</sup>)。

#### [0424] 实施例6:对递呈到T细胞和用人肽活化T细胞的评估

[0425] 为了确定包含人源化细胞免疫系统组分的小鼠是否表现出人源化T细胞免疫应答,测试了来自对于细胞免疫系统组分人源化的小鼠 (TM I/II BCD4/8) 的脾细胞递呈MAGE-A3并对MAGE-A3 (由人HLA-A2特异性递呈的肽) 作出应答的能力。

[0426] 合成由人HLA-A2特异性递呈的肽MAGE-A3 (切尔泰克生物科学公司 (Celtek Biosciences)), 将其稀释于PBS中,并与完全弗氏佐剂 (CFA; Chondrex公司) 等体积混合,以使得200 $\mu$ g的MAGE-A3包含在200 $\mu$ l乳液中。将50 $\mu$ l乳液注射到每只动物上的4个斑点中。其中两个斑点各自在后胁腹中,另两个斑点各自靠近对于人源化MHC I、MHC II $\alpha$ 和 $\beta$ 、TCR $\alpha$ 和 $\beta$ 、CD4、CD8 $\alpha$ 和 $\beta$ 以及 $\beta$ 2M纯合的小鼠 (TM I/II B CD4/8) 或表达内源性MHC I、MHC II $\alpha$ 和 $\beta$ 、TCR $\alpha$ 和 $\beta$ 、CD4、CD8 $\alpha$ 和 $\beta$ 以及 $\beta$ 2M的对照小鼠的每个肩部。

[0427] 获得来自免疫小鼠的脾悬浮液并加以分离。在ACK裂解缓冲液 (生命科技公司 (Life Technologies)) 中将红血球裂解,并将脾细胞悬浮于RPMI完全培养基中。在不存在或在存在10 $\mu$ g/mL或1 $\mu$ g/mL稀释的MAGE-A3肽的情况下,在酶联免疫斑点测定法中测试包被5 $\mu$ g/mL小鼠IFN- $\gamma$ 捕获抗体 (BD生物科学事业部 (BD Biosciences)) 的PVDF板 (密理博公司 (Millipore)) 每孔2 $\times$ 10<sup>5</sup>个分离的脾细胞。与肽温育16-20小时后,将板洗涤并与生物素酰化检测抗体 (BD生物科学事业部 (BD Biosciences)) 一起温育,然后洗涤,用链霉亲和素-HRP (MabTech) 处理,接着洗涤,用TMB底物 (Mabtech) 显色,并通过AID酶联免疫斑点读数器计数。

[0428] 虽然每种基因型仅示出一只小鼠,但对于每种基因型测试了几只小鼠,并且所有样品重复三次,标准偏差以误差棒表示。如图12所示,在用HLA-A2特异性肽MAGE-A3处理后,只有来自对于人源化MHC I、MHC II $\alpha$ 和 $\beta$ 、TCR $\alpha$ 和 $\beta$ 、CD4、CD8 $\alpha$ 和 $\beta$ 以及 $\beta$ 2M中每一者是纯合的小鼠 (TM I/II B CD4/8) 的样品通过分泌IFN- $\gamma$ 作出应答,这表明在由人源化HLA-A2递呈MAGE-A3之后,来自这些小鼠的T细胞被活化。

#### [0429] 实施例7:使用LCMV感染模型评估T细胞功能

[0430] 为了确定包含人源化细胞免疫系统组分的小鼠是否对感染表现出正常反应,测试人源化小鼠清除淋巴细胞性脉络丛脑膜炎病毒 (LCMV) 的能力。LCMV是一种小鼠热带病毒,是否感染取决于病毒菌株。暴露于Armstrong菌株可导致急性感染,在这种感染中小鼠可迅速产生T细胞应答对抗病毒并在约一周内清除感染。在另一方面,Clone 13病毒不能被清除,因此T细胞渐渐“耗尽” (表达与T细胞耗尽相关的标记,例如PD1、Lag3、Tim3),形成慢性

感染。已经证实,用Armstrong菌株感染CD8缺失或MHC I类缺陷型小鼠使得高病毒滴度得以维持(J.Virol.68:8056-63(1994)(《病毒学杂志》,第68卷,第8056-8063页,1994年))。因此,由于病毒感染取决于T细胞活性,LCMV是测试T细胞功能的理想模型。

[0431] 为了确定包含人源化细胞免疫系统组分(例如MHC I、MHC II $\alpha$ 和 $\beta$ 、TCR $\alpha$ 和 $\beta$ 、CD4、CD8 $\alpha$ 和 $\beta$ 以及 $\beta$ 2M)的小鼠是否表现出正常的T细胞功能,在第0天用 $2 \times 10^5$ ffu的Armstrong病毒株i.p.感染对照和人源化小鼠(TM I/II B C4/8)两者。在第3、6、9和12天,摘下器官并测量病毒滴度。如图13A所示,对照和人源化小鼠都能够清除Armstrong感染。

[0432] 在第0天也用 $4.5 \times 10^5$ ffu克隆13病毒i.v.感染对照和人源化小鼠两者。在第21天摘下器官并测量病毒滴度。如图13B所示,这两种小鼠品系都能够形成慢性LCMV感染。还测量了人源化小鼠表达PD1、Lag3和Tim3(T细胞耗尽的标记)的能力。采集未感染小鼠和感染3周后的受感染人源化小鼠的血液,并采用流式细胞术用PE-Cy7偶联的抗PD1抗体(BIOLEGEND)、PerCpCy5.5偶联的Lag3抗体(BIOLEGEND)和PE偶联的Tim3抗体(安迪生物科技公司(R&D Systems))染色。图13C中的数据是对指定受体染色呈阳性的细胞的定量结果。人源化小鼠(TM I/II B C4/8)和对照小鼠B6都在用慢性LCMV克隆13菌株感染3周后表达所有三种T细胞耗尽标记。

[0433] 为了评估对于细胞免疫系统组分人源化的小鼠中的记忆T细胞应答,用 $2 \times 10^5$ ffu的Armstrong菌株感染5只对照小鼠和4只人源化小鼠,并在第17天用 $4.5 \times 10^5$ ffu的克隆13菌株超严重感染(人源化小鼠和对照小鼠各2只进行模拟感染,作为另外的对照)。在初次感染后第31天,摘下器官,分析病毒滴度。如图14所示,已发生急性LCMV感染的5/5对照小鼠和3/4人源小鼠随后免受慢性LCMV感染,这表明这些动物中记忆T细胞应答完整。

[0434] 为了分析细胞应答的本质,在第0天用 $2 \times 10^5$ ffu的Armstrong病毒株感染对照和人源化小鼠。在第10天(图15A-15B)或在感染后指定时间点(图15C-15D),使用三种已知活化人CD8<sup>+</sup>T细胞的HLA-A2限制性肽(GPC10-18、N69-77或Z49-58)(参见Botten et al.(2007)J.Virol.81:2307-17,or gp33,an immunodominant LCMV peptide recognized by mice on a H-2D<sup>b</sup> background(Botten等人,2007年,《病毒学杂志》,第81卷,第2307-2317页))或由小鼠在H-2D<sup>b</sup>背景上识别的免疫显性LCMV肽gp33分析细胞应答的特异性。具体地讲,从摘下的脾脏中分离CD8<sup>+</sup>T细胞并用肽脉冲处理。通过酶联免疫斑点测定法(图15A-15B)或通过对细胞内IFN $\gamma$ 染色(图15C-15D)测量产生干扰素- $\gamma$ (IFN $\gamma$ )的Cd8<sup>+</sup>细胞。

[0435] 由gp33肽特异性活化从对照动物中分离的CD8<sup>+</sup>T细胞(图15A),而由HLA-A2限制性肽活化从人源化动物中分离的CD8<sup>+</sup>T细胞(图15B)。如通过在用肽刺激时表达IFN $\gamma$ 的能力所监测,CD8<sup>+</sup>T细胞活化的时间进程表明,在对照和人源化小鼠两者中CD8<sup>+</sup>T细胞在感染后前两周内扩增,并在清除病毒之后检测不到(图15C-15D)。虽然看起来对gp33肽的应答在对照动物中更强,但应当注意,gp33是已知的免疫显性LCMV表位,而尚未鉴定出免疫显性HLA-A2限制性LCMV表位。总之,包含人源化或基本上人源化的T细胞免疫系统的动物能够处理LCMV表达的蛋白质,在人源化MHC分子上递呈它们,并通过人源化T细胞受体活化T细胞。

[0436] 等同方案

[0437] 本领域的技术人员将认识到,或能够使用不超过常规实验的手段确定本文所述发明的具体实施例的许多等同方案。这些等同方案旨在包括在以下权利要求中。

[0438] 在整个本申请中引用的所有非专利文献、专利申请和专利的全部内容全文以引用

方式并入本文。



- [0001] 序列表
- [0002] <110> 瑞泽恩制药公司
- [0003] <120> 非人动物中的人源化T细胞介导的免疫应答
- [0004] <130> 10145W001
- [0005] <150> 62/143,687
- [0006] <151> 2015-04-06
- [0007] <150> 62/158,804
- [0008] <151> 2015-05-08
- [0009] <150> 62/186,935
- [0010] <151> 2015-06-30
- [0011] <160> 91
- [0012] <170> PatentIn版本3.5
- [0013] <210> 1
- [0014] <211> 19
- [0015] <212> DNA
- [0016] <213> 人工序列
- [0017] <220>
- [0018] <223> 合成
- [0019] <400> 1
- [0020] cagaacgcca ggctgtaac 19
- [0021] <210> 2
- [0022] <211> 20
- [0023] <212> DNA
- [0024] <213> 人工序列
- [0025] <220>
- [0026] <223> 合成
- [0027] <400> 2
- [0028] ggagagcagg gtcagtcaac 20
- [0029] <210> 3
- [0030] <211> 24
- [0031] <212> DNA
- [0032] <213> 人工序列
- [0033] <220>
- [0034] <223> 合成
- [0035] <400> 3
- [0036] caccgccact cacagtcct taca 24
- [0037] <210> 4
- [0038] <211> 22
- [0039] <212> DNA
- [0040] <213> 人工序列
- [0041] <220>

---

[0042]	<223> 合成
[0043]	<400> 4
[0044]	gtgggcacca tcttcatcat tc 22
[0045]	<210> 5
[0046]	<211> 22
[0047]	<212> DNA
[0048]	<213> 人工序列
[0049]	<220>
[0050]	<223> 合成
[0051]	<400> 5
[0052]	cttcctttcc aggggtgtgac tc 22
[0053]	<210> 6
[0054]	<211> 23
[0055]	<212> DNA
[0056]	<213> 人工序列
[0057]	<220>
[0058]	<223> 合成
[0059]	<400> 6
[0060]	aggcctgcga tcaggtggca cct 23
[0061]	<210> 7
[0062]	<211> 19
[0063]	<212> DNA
[0064]	<213> 人工序列
[0065]	<220>
[0066]	<223> 合成
[0067]	<400> 7
[0068]	ggtggagagg ctattcggc 19
[0069]	<210> 8
[0070]	<211> 17
[0071]	<212> DNA
[0072]	<213> 人工序列
[0073]	<220>
[0074]	<223> 合成
[0075]	<400> 8
[0076]	gaacacggcg gcatcag 17
[0077]	<210> 9
[0078]	<211> 23
[0079]	<212> DNA
[0080]	<213> 人工序列
[0081]	<220>
[0082]	<223> 合成
[0083]	<400> 9

[0084] tgggcacaaac agacaatcgg ctg 23  
[0085] <210> 10  
[0086] <211> 140  
[0087] <212> DNA  
[0088] <213> 人工序列  
[0089] <220>  
[0090] <223> 合成  
[0091] <400> 10  
[0092] tttgtaaaca aagtctaccc agagacagat gacagacttc agctccaatg ctgattgggtt 60  
[0093] cctcacttgg gaccaaccct accggtataa cttcgtataa ggtatcctat acgaagttat 120  
[0094] atgcatggcc tccgcgccgg 140  
[0095] <210> 11  
[0096] <211> 140  
[0097] <212> DNA  
[0098] <213> 人工序列  
[0099] <220>  
[0100] <223> 合成  
[0101] <400> 11  
[0102] cgacctgcag ccggcgcgcc ataacttcgt ataaggtatc ctatacgaag ttatctcgag 60  
[0103] cacaggcatt tgggtgggca gggatggacg gtgactggga caatcgggat ggaagagcat 120  
[0104] agaatgggag ttaggaaga 140  
[0105] <210> 12  
[0106] <211> 17  
[0107] <212> DNA  
[0108] <213> 人工序列  
[0109] <220>  
[0110] <223> 合成  
[0111] <400> 12  
[0112] cgaggagccc cgtaca 17  
[0113] <210> 13  
[0114] <211> 20  
[0115] <212> DNA  
[0116] <213> 人工序列  
[0117] <220>  
[0118] <223> 合成  
[0119] <400> 13  
[0120] aagcgcacga actccttggt 20  
[0121] <210> 14  
[0122] <211> 17  
[0123] <212> DNA  
[0124] <213> 人工序列  
[0125] <220>

[0126]	<223> 合成
[0127]	<400> 14
[0128]	ctctgtcggc tatgtgg 17
[0129]	<210> 15
[0130]	<211> 22
[0131]	<212> DNA
[0132]	<213> 人工序列
[0133]	<220>
[0134]	<223> 合成
[0135]	<400> 15
[0136]	ggactcccag aatctcctga ga 22
[0137]	<210> 16
[0138]	<211> 25
[0139]	<212> DNA
[0140]	<213> 人工序列
[0141]	<220>
[0142]	<223> 合成
[0143]	<400> 16
[0144]	gagtcataaa ccatcactgt gaaga 25
[0145]	<210> 17
[0146]	<211> 16
[0147]	<212> DNA
[0148]	<213> 人工序列
[0149]	<220>
[0150]	<223> 合成
[0151]	<400> 17
[0152]	tggtgggttg ctggaa 16
[0153]	<210> 18
[0154]	<211> 90
[0155]	<212> DNA
[0156]	<213> 人工序列
[0157]	<220>
[0158]	<223> 合成
[0159]	<400> 18
[0160]	ctgtttcttc cctaactccc attctatgct cttccatccc gaccgcggcc caatctctct 60
[0161]	ccactacttc ctgcctacat gtatgtaggt 90
[0162]	<210> 19
[0163]	<211> 80
[0164]	<212> DNA
[0165]	<213> 人工序列
[0166]	<220>
[0167]	<223> 合成

- [0168] <400> 19  
[0169] caaggtttcc tcctatgatg cttgtgtgaa actcggggcc ggccagcatt taacagtaca 60  
[0170] gggatgggag cacagctcac 80  
[0171] <210> 20  
[0172] <211> 80  
[0173] <212> DNA  
[0174] <213> 人工序列  
[0175] <220>  
[0176] <223> 合成  
[0177] <400> 20  
[0178] gaaagcagtc ttcccagcct tcacactcag aggtacaaat cccattttc atattagcga 60  
[0179] ttttaattta ttctagcctc 80  
[0180] <210> 21  
[0181] <211> 80  
[0182] <212> DNA  
[0183] <213> 人工序列  
[0184] <220>  
[0185] <223> 合成  
[0186] <400> 21  
[0187] tcttcctaa ctcccattct atgctcttcc atcccaccg cggeccaatc tctctccact 60  
[0188] acttctgcc tacatgtatg 80  
[0189] <210> 22  
[0190] <211> 100  
[0191] <212> DNA  
[0192] <213> 人工序列  
[0193] <220>  
[0194] <223> 合成  
[0195] <400> 22  
[0196] gaggttcctcc atcacttcac tgggtagcac agctgtaact gtcagcctg tcttggtctg 60  
[0197] cagggtgtgg gcgttgccgg tggggccggt taaggttcca 100  
[0198] <210> 23  
[0199] <211> 100  
[0200] <212> DNA  
[0201] <213> 人工序列  
[0202] <220>  
[0203] <223> 合成  
[0204] <400> 23  
[0205] tcccacatcc tattttaatt tgctccatgt tctcatctcc atcagcacag ctgagataa 60  
[0206] cttcgataa tgtatgctat acgaagtat atgcatggcc 100  
[0207] <210> 24  
[0208] <211> 100  
[0209] <212> DNA

[0210]	<213> 人工序列
[0211]	<220>
[0212]	<223> 合成
[0213]	<400> 24
[0214]	atacgaagtt atgctagtaa ctataacggt cctaaggtag cgagtggctt acaggtaggt 60
[0215]	gcgtgaagct tctacaagca cagttgcccc ctgggaagca 100
[0216]	<210> 25
[0217]	<211> 17
[0218]	<212> DNA
[0219]	<213> 人工序列
[0220]	<220>
[0221]	<223> 合成
[0222]	<400> 25
[0223]	tgcgccgat cttagcc 17
[0224]	<210> 26
[0225]	<211> 18
[0226]	<212> DNA
[0227]	<213> 人工序列
[0228]	<220>
[0229]	<223> 合成
[0230]	<400> 26
[0231]	ttgaccgatt ccttgcg 18
[0232]	<210> 27
[0233]	<211> 21
[0234]	<212> DNA
[0235]	<213> 人工序列
[0236]	<220>
[0237]	<223> 合成
[0238]	<400> 27
[0239]	acgagcgggt tcggcccatt c 21
[0240]	<210> 28
[0241]	<211> 18
[0242]	<212> DNA
[0243]	<213> 人工序列
[0244]	<220>
[0245]	<223> 合成
[0246]	<400> 28
[0247]	ccccacagca cgtttcct 18
[0248]	<210> 29
[0249]	<211> 23
[0250]	<212> DNA
[0251]	<213> 人工序列

[0252] <220>  
[0253] <223> 合成  
[0254] <400> 29  
[0255] cgtccattg aagaaatgac act 23  
[0256] <210> 30  
[0257] <211> 15  
[0258] <212> DNA  
[0259] <213> 人工序列  
[0260] <220>  
[0261] <223> 合成  
[0262] <400> 30  
[0263] tggcagccta agagg 15  
[0264] <210> 31  
[0265] <211> 18  
[0266] <212> DNA  
[0267] <213> 人工序列  
[0268] <220>  
[0269] <223> 合成  
[0270] <400> 31  
[0271] cccacagea cgtttcct 18  
[0272] <210> 32  
[0273] <211> 17  
[0274] <212> DNA  
[0275] <213> 人工序列  
[0276] <220>  
[0277] <223> 合成  
[0278] <400> 32  
[0279] acccgctcgg tccatt 17  
[0280] <210> 33  
[0281] <211> 18  
[0282] <212> DNA  
[0283] <213> 人工序列  
[0284] <220>  
[0285] <223> 合成  
[0286] <400> 33  
[0287] agcctaagag ggagtgtc 18  
[0288] <210> 34  
[0289] <211> 20  
[0290] <212> DNA  
[0291] <213> 人工序列  
[0292] <220>  
[0293] <223> 合成

[0294] <400> 34  
[0295] agaccctggt gatgctggaa 20  
[0296] <210> 35  
[0297] <211> 18  
[0298] <212> DNA  
[0299] <213> 人工序列  
[0300] <220>  
[0301] <223> 合成  
[0302] <400> 35  
[0303] cgcttggtg ctccactt 18  
[0304] <210> 36  
[0305] <211> 18  
[0306] <212> DNA  
[0307] <213> 人工序列  
[0308] <220>  
[0309] <223> 合成  
[0310] <400> 36  
[0311] tcgaagtgga gaggttta 18  
[0312] <210> 37  
[0313] <211> 21  
[0314] <212> DNA  
[0315] <213> 人工序列  
[0316] <220>  
[0317] <223> 合成  
[0318] <400> 37  
[0319] tggaatggag tgagcagctt t 21  
[0320] <210> 38  
[0321] <211> 20  
[0322] <212> DNA  
[0323] <213> 人工序列  
[0324] <220>  
[0325] <223> 合成  
[0326] <400> 38  
[0327] gcacggtccc cttcttagtg 20  
[0328] <210> 39  
[0329] <211> 18  
[0330] <212> DNA  
[0331] <213> 人工序列  
[0332] <220>  
[0333] <223> 合成  
[0334] <400> 39  
[0335] tgacttccta aatttctc 18



---

[0336]	<210> 40
[0337]	<211> 21
[0338]	<212> DNA
[0339]	<213> 人工序列
[0340]	<220>
[0341]	<223> 合成
[0342]	<400> 40
[0343]	ctggcggcctt gaagaatttg g 21
[0344]	<210> 41
[0345]	<211> 24
[0346]	<212> DNA
[0347]	<213> 人工序列
[0348]	<220>
[0349]	<223> 合成
[0350]	<400> 41
[0351]	catgatttcc aggttggtt tgtc 24
[0352]	<210> 42
[0353]	<211> 26
[0354]	<212> DNA
[0355]	<213> 人工序列
[0356]	<220>
[0357]	<223> 合成
[0358]	<400> 42
[0359]	cgatttgcca gctttgaggc tcaagg 26
[0360]	<210> 43
[0361]	<211> 21
[0362]	<212> DNA
[0363]	<213> 人工序列
[0364]	<220>
[0365]	<223> 合成
[0366]	<400> 43
[0367]	cctcacttgg gaccaaccct a 21
[0368]	<210> 44
[0369]	<211> 20
[0370]	<212> DNA
[0371]	<213> 人工序列
[0372]	<220>
[0373]	<223> 合成
[0374]	<400> 44
[0375]	ttgtcccagt caccgtccat 20
[0376]	<210> 45
[0377]	<211> 24

[0378] <212> DNA  
[0379] <213> 人工序列  
[0380] <220>  
[0381] <223> 合成  
[0382] <400> 45  
[0383] tgcatctcga gcacaggcat ttgg 24  
[0384] <210> 46  
[0385] <211> 100  
[0386] <212> DNA  
[0387] <213> 人工序列  
[0388] <220>  
[0389] <223> 合成  
[0390] <400> 46  
[0391] atggagtagt cagaacacac tcttcagaag ggactcctga tttcaaaggg ggtaccgggc 60  
[0392] cccccctcga ggtcgacata acttcgtata gcatacatta 100  
[0393] <210> 47  
[0394] <211> 108  
[0395] <212> DNA  
[0396] <213> 人工序列  
[0397] <220>  
[0398] <223> 合成  
[0399] <400> 47  
[0400] ggccatgcat ataacttcgt atagcataca ttatacgaag ttataccggt gcgatcgcg 60  
[0401] gcttcctct tctaaccact aattcaaaaa ggattgtaag taatgttt 108  
[0402] <210> 48  
[0403] <211> 145  
[0404] <212> DNA  
[0405] <213> 人工序列  
[0406] <220>  
[0407] <223> 合成  
[0408] <400> 48  
[0409] agacagaccc ctaaacacct ccaaattaaa agcggcaaag agataagggt ggagctccac 60  
[0410] cgcggtggcg gccgccaccg cggaggagct cgaggtttcc ggtacttaac aacagagcac 120  
[0411] agatttagtg gtgagggact ctctc 145  
[0412] <210> 49  
[0413] <211> 100  
[0414] <212> DNA  
[0415] <213> 人工序列  
[0416] <220>  
[0417] <223> 合成  
[0418] <400> 49  
[0419] atggagtagt cagaacacac tcttcagaag ggactcctga tttcaaaggg ggtaccgggc 60

[0420] ccccccctcga ggtcgacata acttcgtata gcatacatta 100  
[0421] <210> 50  
[0422] <211> 109  
[0423] <212> DNA  
[0424] <213> 人工序列  
[0425] <220>  
[0426] <223> 合成  
[0427] <400> 50  
[0428] ggccatgcat ataacttcgt atagcataca ttatacgaag ttataccggt gcgatcgctc 60  
[0429] aagcatgcaa gggtaacata tgttatgaga ttatatcttc tttatctca 109  
[0430] <210> 51  
[0431] <211> 100  
[0432] <212> DNA  
[0433] <213> 人工序列  
[0434] <220>  
[0435] <223> 合成  
[0436] <400> 51  
[0437] atggagtagt cagaacacac tcttcagaag ggactcctga tttcaaaggg ggtaccggg 60  
[0438] ccccccctcg agaagttcct attccgaagt tcctattctc 100  
[0439] <210> 52  
[0440] <211> 108  
[0441] <212> DNA  
[0442] <213> 人工序列  
[0443] <220>  
[0444] <223> 合成  
[0445] <400> 52  
[0446] gttcctattc cgaagttcct attctctaga aagtatagga acttcctagg gcgatcgctc 60  
[0447] ctctccaggc tcgaattagt attacagttg aggcacgttg tcctcccg 108  
[0448] <210> 53  
[0449] <211> 100  
[0450] <212> DNA  
[0451] <213> 人工序列  
[0452] <220>  
[0453] <223> 合成  
[0454] <400> 53  
[0455] atggagtagt cagaacacac tcttcagaag ggactcctga tttcaaaggg ggtaccgggc 60  
[0456] ccccccctcga ggtcgacata acttcgtata gcatacatta 100  
[0457] <210> 54  
[0458] <211> 108  
[0459] <212> DNA  
[0460] <213> 人工序列  
[0461] <220>

- [0462] <223> 合成  
[0463] <400> 54  
[0464] ggccatgcat ataacttcgt atagcataca ttatacgaag ttataccggt gcgatcgccg 60  
[0465] cctccatttc cttcatagga aacatgaagt gaatggggct gtgtgtgt 108  
[0466] <210> 55  
[0467] <211> 100  
[0468] <212> DNA  
[0469] <213> 人工序列  
[0470] <220>  
[0471] <223> 合成  
[0472] <400> 55  
[0473] atggagtagt cagaacacac ttttcagaag ggactcctga tttcaaaggg ggtaccgggc 60  
[0474] cccccctcga ggtcgacata acttcgtata gcatacatta 100  
[0475] <210> 56  
[0476] <211> 108  
[0477] <212> DNA  
[0478] <213> 人工序列  
[0479] <220>  
[0480] <223> 合成  
[0481] <400> 56  
[0482] ggccatgcat ataacttcgt atagcataca ttatacgaag ttataccggt gcgatcgctg 60  
[0483] ggagcacgtt ccattattat aacaactttc tgaacacaag agggcagt 108  
[0484] <210> 57  
[0485] <211> 100  
[0486] <212> DNA  
[0487] <213> 人工序列  
[0488] <220>  
[0489] <223> 合成  
[0490] <400> 57  
[0491] atggagtagt cagaacacac ttttcagaag ggactcctga tttcaaaggg ggtaccgggc 60  
[0492] cccccctcga ggtcgacata acttcgtata gcatacatta 100  
[0493] <210> 58  
[0494] <211> 108  
[0495] <212> DNA  
[0496] <213> 人工序列  
[0497] <220>  
[0498] <223> 合成  
[0499] <400> 58  
[0500] ggccatgcat ataacttcgt atagcataca ttatacgaag ttataccggt gcgatcgctt 60  
[0501] taaggtgagg aggcaggcaa taccctctt ccaccgcatt ctcaatcc 108  
[0502] <210> 59  
[0503] <211> 100

- [0504] <212> DNA  
[0505] <213> 人工序列  
[0506] <220>  
[0507] <223> 合成  
[0508] <400> 59  
[0509] gggggggtgg ggtggaggag gaggtacag catctcctct ccttcctctc tggtaccgaa 60  
[0510] gttcctattc cgaagttcct attctctaga aagtatagga 100  
[0511] <210> 60  
[0512] <211> 108  
[0513] <212> DNA  
[0514] <213> 人工序列  
[0515] <220>  
[0516] <223> 合成  
[0517] <400> 60  
[0518] gaagttccta ttctctagaa agtataggaa ctctctaggg ttccaccggt gcgatcgct 60  
[0519] gaatatacta aaaaccactt aattatata ttgaaagggt ggatgtta 108  
[0520] <210> 61  
[0521] <211> 108  
[0522] <212> DNA  
[0523] <213> 人工序列  
[0524] <220>  
[0525] <223> 合成  
[0526] <400> 61  
[0527] ctctctccta cccagtcct ctcacacgag cctgaaggcc ctgccaaggt ggcgcgctt 60  
[0528] tcaaatgtt gttgagttca aagtgggcaa cagaaaagg ggtgtgag 108  
[0529] <210> 62  
[0530] <211> 130  
[0531] <212> DNA  
[0532] <213> 人工序列  
[0533] <220>  
[0534] <223> 合成  
[0535] <400> 62  
[0536] aataaatagt aaatttctgt agaatcataa tgaggtctag acccccgggc tcgataacta 60  
[0537] taacggtcct aaggtagcga aatggcgct aatcaagccc agctcttcat gctgcatttt 120  
[0538] tatcttcttt 130  
[0539] <210> 63  
[0540] <211> 100  
[0541] <212> DNA  
[0542] <213> 人工序列  
[0543] <220>  
[0544] <223> 合成  
[0545] <400> 63

[0546] ttgactcggg ggtgcctggg ttgactgca atgacagtt gctgggaagg accggtataa 60  
[0547] cttcgtataa tgtatgctat acgaagtat atgcatggcc 100  
[0548] <210> 64  
[0549] <211> 100  
[0550] <212> DNA  
[0551] <213> 人工序列  
[0552] <220>  
[0553] <223> 合成  
[0554] <400> 64  
[0555] ccggcgcgcc ataacttcgt ataatgtatg ctatacgaag ttatgtcgac ataaggtaa 60  
[0556] acagagtcgt cccttcccat ctggaaccct ctacctttct 100  
[0557] <210> 65  
[0558] <211> 107  
[0559] <212> DNA  
[0560] <213> 人工序列  
[0561] <220>  
[0562] <223> 合成  
[0563] <400> 65  
[0564] gttgatgaat cataaaagaa gagatattca agaaaaggat ggccacactg cggccgcaga 60  
[0565] ggtattcaag gaaaatgcag actcttcacg taagagggat gaggggc 107  
[0566] <210> 66  
[0567] <211> 100  
[0568] <212> DNA  
[0569] <213> 人工序列  
[0570] <220>  
[0571] <223> 合成  
[0572] <400> 66  
[0573] tccccggagt cggagggtgg accggagctg gaggagctgc cgcggtggcg gccgatgcca 60  
[0574] ttctattacc tctttctccg caccgcacat agataaagct 100  
[0575] <210> 67  
[0576] <211> 100  
[0577] <212> DNA  
[0578] <213> 人工序列  
[0579] <220>  
[0580] <223> 合成  
[0581] <400> 67  
[0582] gggggggtgg ggtggaggag gagggtacag catctcctct ccttcctctc tggtaccgaa 60  
[0583] gttcctattc cgaagtccct attctctaga aagtatagga 100  
[0584] <210> 68  
[0585] <211> 108  
[0586] <212> DNA  
[0587] <213> 人工序列

[0588] <220>  
[0589] <223> 合成  
[0590] <400> 68  
[0591] gaagttccta ttctctagaa agtataggaa cttcctaggg tttcaccggt gcgatcgca 60  
[0592] agcaattaac tgccccctggc ccagttgcct cctctgataa tgcattgt 108  
[0593] <210> 69  
[0594] <211> 100  
[0595] <212> DNA  
[0596] <213> 人工序列  
[0597] <220>  
[0598] <223> 合成  
[0599] <400> 69  
[0600] gggggggtgg ggtggaggag gaggtacag catctcctct ccttcctctc tggtagcgaa 60  
[0601] gttcctattc cgaagttcct attctctaga aagtatagga 100  
[0602] <210> 70  
[0603] <211> 108  
[0604] <212> DNA  
[0605] <213> 人工序列  
[0606] <220>  
[0607] <223> 合成  
[0608] <400> 70  
[0609] gaagttccta ttctctagaa agtataggaa cttcctaggg tttcaccggt gcgatcgct 60  
[0610] tatctagtag acttaattaa ggatcgatcc ggcgcgcaa tagtcatg 108  
[0611] <210> 71  
[0612] <211> 100  
[0613] <212> DNA  
[0614] <213> 人工序列  
[0615] <220>  
[0616] <223> 合成  
[0617] <400> 71  
[0618] gttttccaga cttcaacttg actatcagcc agaaattcag tggcaaacc ccaccagtc 60  
[0619] cctaagtga ggccccctggg gagtatggtt agggctcagg 100  
[0620] <210> 72  
[0621] <211> 100  
[0622] <212> DNA  
[0623] <213> 人工序列  
[0624] <220>  
[0625] <223> 合成  
[0626] <400> 72  
[0627] caccaccaa agaaagtgcc caggagaagg gcaaggagag agcagagcat agttcaagat 60  
[0628] ggtctttgtc taggcttgtc tactctgcac ttgtacttcc 100  
[0629] <210> 73

[0630] <211> 202  
[0631] <212> DNA  
[0632] <213> 人工序列  
[0633] <220>  
[0634] <223> 合成  
[0635] <400> 73  
[0636] aggggaaacc cgcaaaggat gggacatagg gagacagctg ttaacatctg aaacatgacc 60  
[0637] ttcttttctg tgcagcaciaa ctctctagctg tctactcaagg gaagaaagtg gtgctgggca 120  
[0638] aaaaagggga tacagtggaa ctgacctgta cagcttccca gaagaagagc atacaattcc 180  
[0639] actggaaaaa ctccaaccag at 202  
[0640] <210> 74  
[0641] <211> 240  
[0642] <212> DNA  
[0643] <213> 人工序列  
[0644] <220>  
[0645] <223> 合成  
[0646] <400> 74  
[0647] ctggtcacct ggatgaagtg agggagggcc ctctgggttt ggggctgggt ttgaactgag 60  
[0648] acatccatga gccagcctgg ggctggcttc actgaagatc atctatgtcg ggtgcggaga 120  
[0649] aagaggtaat gaaatggcac atgctatgta caaactctat tgctgagcag caccagtc 180  
[0650] tgagctggct ctgaattgag ggtgaaatc acacattctc cccaacatc tataatctgg 240  
[0651] <210> 75  
[0652] <211> 151  
[0653] <212> DNA  
[0654] <213> 人工序列  
[0655] <220>  
[0656] <223> 合成  
[0657] <400> 75  
[0658] tatggagtga aagcctttgg tgtctgagat ctggtcttag ttaaactctg ggatcggcgc 60  
[0659] gccgaattcc tgcagcccg gctcgagata acttcgtata atgtatgcta tacgaagtta 120  
[0660] tatgcatccg gtaggggag gcgcttttcc c 151  
[0661] <210> 76  
[0662] <211> 151  
[0663] <212> DNA  
[0664] <213> 人工序列  
[0665] <220>  
[0666] <223> 合成  
[0667] <400> 76  
[0668] agtattgttt tgccaagttc taattccatc agacctgac ctgcagccct agataacttc 60  
[0669] gtataatgta tgctatacga agttatccta ggccagaggg cttgggttga cagaaactca 120  
[0670] gtggcattct tatccagagt ttctctacac c 151  
[0671] <210> 77



[0672]	<211> 18263
[0673]	<212> DNA
[0674]	<213> 人工序列
[0675]	<220>
[0676]	<223> 合成
[0677]	<220>
[0678]	<221> 尚未归类的特征
[0679]	<222> (2957) .. (2957)
[0680]	<223> n = A、T、C或G
[0681]	<220>
[0682]	<221> 尚未归类的特征
[0683]	<222> (3193) .. (3193)
[0684]	<223> n = A、T、C或G
[0685]	<400> 77
[0686]	aagaaagtgg tgctgggcaa aaaaggggat acagtggaac tgacctgtac agcttcccag 60
[0687]	aagaagagca tacaattcca ctggaaaaac tccaaccaga taaagattct gggaaatcag 120
[0688]	ggctccttct taactaaagg tagggttgcc tggctcccca tccaggagg aaaacacact 180
[0689]	atggagtga agccttttgt gtctgagatc tggctctagt taaactctgg gatcggcgcg 240
[0690]	ccgaattcct gcagcccggg ctcgagataa cttcgtataa tgtatgctat acgaagtatt 300
[0691]	atgcatccgg gtaggggagg cgcttttccc aaggcagtct ggagcatgcg ctttagcagc 360
[0692]	cccgtgggc acttgcgct acacaagtgg cctctggcct cgcacacatt ccacatccac 420
[0693]	cggtaggcgc caaccggctc cgcttcttgg tggcccttc gcgccacctt ctactcctcc 480
[0694]	cctagtcagg aagttcccc ccgcccgcga gctcgcgtcg tgcaggacgt gacaaatgga 540
[0695]	agtagcacgt ctactagtc tcgtgcagat ggacagcacc gctgagcaat ggaagcgggt 600
[0696]	aggcctttgg ggcagcgcc aatagcagct ttgctccttc gctttctggg ctacagaggt 660
[0697]	gggaaggggt gggctccggg gcgggctcag gggcgggctc aggggcgggg cgggcgcccc 720
[0698]	aagtcctcc ggaggcccg cattctgcac gcttcaaaag cgcacgtctg ccgcgctgtt 780
[0699]	ctctcttcc tcatctccg gcctttcgac ctgcagccaa ttgttgacaa ttaatcatcg 840
[0700]	gcatagtata tcggcatagt ataatacgac aaggtgagga actaaaccat gggatcgcc 900
[0701]	attgaacaag atggattgca cgcaggttct ccggccgctt ggggtggagag gctattcggc 960
[0702]	tatgactggg cacaacagac aatcggtgc tctgatgcc ccgtgttccg gctgtcagcg 1020
[0703]	cagggcgcc cggttctttt tgtcaagacc gacctgtccg gtgccctgaa tgaactgcag 1080
[0704]	gacgaggcag cgcggtatc gtggtggcc acgacggcg ttccttgcc agctgtgctc 1140
[0705]	gacgttgta ctgaagcgg aagggaactg ctgctattgg gcgaagtgc ggggcaggat 1200
[0706]	ctcctgtcat ctacattgc tcctgccgag aaagtatcca tcatggctga tgcaatgcgg 1260
[0707]	cggctgcata cgcttgatcc ggctacctgc ccattcgacc accaagcgaa acatcgcac 1320
[0708]	gagcgagcac gtactcgat ggaagccgg cttgtcgatc aggatgatct ggacgaagag 1380
[0709]	catcaggggc tcgcgccag cgaactgttc gccaggtca aggcgcgcac gcccgacggc 1440
[0710]	gatgatctcg tcgtgacca tggcgatgcc tggcttgcga atatcatggt ggaaaatggc 1500
[0711]	cgctttctg gattcatcga ctgtggccgg ctgggtgtgg cggaccgcta tcaggacata 1560
[0712]	gcgttggcta cccgtgatat tgctgaagag cttggcggcg aatgggctga ccgcttctc 1620
[0713]	gtgctttacg gtatcgccgc tcccgattcg cagcgcatcg ctttctatcg ctttcttgac 1680

[0714]	gagttcttct gaggggatcc gctgtaagtc tgcagaaatt gatgatctat taaacaataa	1740
[0715]	agatgtccac taaaatggaa gtttttcttg tcatactttg ttaagaaggg tgagaacaga	1800
[0716]	gtacctacat tttgaatgga aggattggag ctacgggggt gggggtgggg tgggattaga	1860
[0717]	taaatgcctg ctctttactg aaggctcttt actattgctt tatgataatg tttcatagtt	1920
[0718]	ggatatcata atttaacaa gcaaaaccaa attaagggcc agctcattcc tcccactcat	1980
[0719]	gatctataga tctatagatc tctcgtggga tcattgtttt tctcttgatt cccactttgt	2040
[0720]	ggttctaagt actgtggttt ccaaagtgtg cagtttcata gcctgaagaa cgagatcagc	2100
[0721]	agcctctgtt ccacatacac ttcatctca gtattgtttt gccagttct aattccatca	2160
[0722]	gacctcgacc tgcagcccta gataactcg tataatgtat gctatacgaa gttatcctag	2220
[0723]	gccagagggc ttgggttgac agaaactcag tggcattctt atccagagtt tctctacacc	2280
[0724]	aactgctggt ggcccaggga aagggtggtat gtgaatttca atattttaat atttaatat	2340
[0725]	catgaactta ttttagtgag ttttagaaca atcactatca cttaaaaccc gtgatttctt	2400
[0726]	gagtattgtt gctacagacc tatgtagata atactttgca cagtgactca tatgtataat	2460
[0727]	cctagcactg tgggaggctg aggccggagg attgcttgag tccaggagtt caagaccagc	2520
[0728]	ctgaacaaca tagtgagact ctgtctctat gaaaaaaaaat atatataat ttttttggga	2580
[0729]	gacaaggtct agttctatca cccaggtcc agtgcagtgg tgtgatctcg gctcactgca	2640
[0730]	atctccacct cccaggtca agtcatcatc ccacctcagc ctcccaagta gctgggacta	2700
[0731]	caggcatgca ccacatgcc aggctaattt ttgtattttt tatagagaca gggtttcacc	2760
[0732]	atgttggcca ggctggtctc gaactcatga gctcaagtga tccactcacc ttggcctctc	2820
[0733]	agagtgtctg aattacaggt gtgtgtcact atgcctagcc aaaaaaatt tttttaatta	2880
[0734]	aaaaaaaaa ggccggtgt agtggctcac acctgtaatc cagaactttg ggagtttgag	2940
[0735]	gtgggcagat caccgnggt caggagtca agaccagtct ggccaacatg gtgaaacccg	3000
[0736]	gtcttacta aaaatacaaa aattagccag gtgtgggggt gcagtcctgt acttcagct	3060
[0737]	actcaggagg ctgaggcagg agactcgctt gaacctggga ggcaaaggct gcagttagct	3120
[0738]	gagattgcac cactgcactc cagcctgggt gacagagcaa gacttcatct caaaaaaaaa	3180
[0739]	aaaaaagctg canatttatt attattatta ttagtttatt tattttattt tttgagacag	3240
[0740]	agtctcgttc tgctgcccag gctggagtgc ggtggcgtga tcttggtca ttgcaacctc	3300
[0741]	cacctcccgg gttcaagtga ttctctgcc tcagcctccc gagtagctgg gactacaggc	3360
[0742]	gtatgccacc atgcctggtt aattttttgt acttttagta gagacagagt ttacgggtgt	3420
[0743]	tagccaggct ggtcttgatc tctgacctc gtgatttacc ctcttggtcc tcccaaagt	3480
[0744]	ctgggattac aggcgtgagt cactgtgccc ggcccagaat catttttttc tacttttttt	3540
[0745]	tttttgaggc aaactctcga tctgttccc aggctggagt gcagtgggca tgatcttggc	3600
[0746]	tactgcaag ctctgcctcc caggttcaag caattctcct gcctcagcct cctgagtagc	3660
[0747]	tgggactaca ggcgtgtgcc accatgcccg gctaatttgc gtatttttag tagagaccgg	3720
[0748]	ttttcatcat attggccagg ctggtcttga actcctgacc tcaagtgatt ctcccactt	3780
[0749]	agcctcccaa agtgcctggga ttacaggcat gagctactgc acttggcctt ttctcctggt	3840
[0750]	tttaaaacta ttatatgtc attacaaaat atttgggtcaa tgaagaaaag aatatggaag	3900
[0751]	aaaatcaaat gcatgcatac ttctatcact cagagatata ctctgctaac attttgattg	3960
[0752]	attttcttcc aatctttttt ttttttttc tttttgagac agggctctcac tctgctgccc	4020
[0753]	aggctggagt acagtggcat gaccacaaca catcacagcc tcaagtgatc ttcccacttc	4080
[0754]	agccttccca gtagctggga ctacagggtc acgccaccat gtacacctaa ttttttactt	4140
[0755]	ttttagagaga tgagacttca ccatgttgct caggctggtc ttgaattcct aggtcgaagt	4200

[0756]	gatcttcccg ctttggcctc ccaaagtgt gggattatag gtatgagcca ctgcatgtgg	4260
[0757]	cctattttct tccactgttg ttccggcgtg agaattatt atacataatt acgtaaata	4320
[0758]	tatcatactg tatatacctt ttttctact ctttcttaa gttatatcat aatgagacta	4380
[0759]	ccaattatta gacttttttt cttttttttg agacggagtc tcggtctgtc acctaggctg	4440
[0760]	gagtgcattg gcgcgatctc agctcgtgtc aacctctgcc tcccagggtc aagcaattct	4500
[0761]	gcctcagcct cccgagtagc tgggactaca gacacgtgcc accatgcccc gctaactttt	4560
[0762]	ttattttttt attagagaca gggttccacc atgctagcag gatgggtctca atctctcgac	4620
[0763]	ttcgtgatca gcccggcttg gcctcccaaa gtgctgggag tacagggtgt agccaccgca	4680
[0764]	ctcggcctag actaactatt taaagtaatc tggcaatgtt taacgaatac aaaactctaa	4740
[0765]	aacccttgga cctaataata gctatttttg aaagtctact tgacagaaat aaaattgtga	4800
[0766]	atattctttt ttgttgtttt tttagacag agtctcattt ggacgcctag gctggagtgc	4860
[0767]	agtggcatga tctcggttaa ctgcaacctc cacctcctgg gttcaagtga ttctcctgcc	4920
[0768]	tcagcctcct gagcagctgg gattacaggt gtgcaccacc atgtctggct aatttttgca	4980
[0769]	tttttagtag atggggtttc accatgttga ccagggtggt ctggaacttc taccctcaag	5040
[0770]	tgatctaccc accttggcct cccaaagtgc tgggattaca ggtgtgagcc accacgcctg	5100
[0771]	accagtgaac acttaataat atctatggaa aggtgttatt ataagaattg cttgtggggc	5160
[0772]	cgggcgtggt ggctcacgcc tgtaatccca gcactttggg aggctgtggc aggcggatca	5220
[0773]	cgaggtcagg agatcaagat catcctggct aacacggtga aaccccgctc ctactaaaa	5280
[0774]	tacaaaaaaa ttagccaggc gtggtggcgg gcacttgtaa tcccagctat ccaggaggct	5340
[0775]	gaggcaggag aattgcgtga acccaggagg cggaggtcgc agtgagctga gaccgtgcca	5400
[0776]	ttgcactcca gcctgagtga cagagtgaga ctccatcaca aaaaataaat aaataaata	5460
[0777]	ataaaatata aataagtaaa taaaggtcag gagtgggtgc tcacgcctgt aatcccagca	5520
[0778]	ctttgggagg ccgaggtgga cagatcatga ggtcatgaga tcaagaccat cctggctaac	5580
[0779]	acagtgaaac cctgcctcta ctaaaaatac aaaaagtcac ccagggtgtg tggcacacac	5640
[0780]	ctatagtccc agctacttgg gaggtgagg caggagaatc acttgaaccc aggaggcaga	5700
[0781]	ggttgcatgt agctgagatc gcgccactac actccagcct aggcgacaga gcaagactct	5760
[0782]	gtctcaaaat aaataaataa ataaataaat aaataaataa ataaataaaa taaaaagcac	5820
[0783]	acacacacac acacacacac acacacaatg caaaagaccc accctactac aactaacatt	5880
[0784]	atatttaatg gtgaaaaact gaattctttc tcctaagtg caggaataag acaaatgtgt	5940
[0785]	ctgctcttac tactcttatt caacataata ctgcaatccc ttgccagtgc aataaggcaa	6000
[0786]	gaaaaatgaa ataaaaggaa aactgatcag aaagaaagaa ataaaactgt tcctatttgt	6060
[0787]	ggatgacatg attacataga aaatctcaaa gaatctgtaa gaaacttctt agaattaata	6120
[0788]	aatgaattca tcaaggttgc agaataaag ataaacataa aaaatctatt gtatttctat	6180
[0789]	atattagcaa ggaacatgtg tacacagaaa ttaaaactac aataccattt ataattgtct	6240
[0790]	aaaaaggcca ggcattgttg ctcacacctg taattcctgc actttgggag gccaaagtgt	6300
[0791]	gaagattgct taagcccagg agttcaagac cagcccgggc aacatagtga gaccttgtct	6360
[0792]	ctacaaaaag taaaaaatta gctgagcatg gccgggtgca gtggctcact cctgtaaccc	6420
[0793]	caacactttg ggaggctgag gcgggcggat catgaggtca ggagatcgag accatcctgg	6480
[0794]	ctaacacggt gaaacctgt ctctactaaa aacacaaaaa attagctgga tgtgtgtgga	6540
[0795]	ggcgcctgta gaccagcta ctcgggaagc tgaggcagga gaatggcgtg aacctgggag	6600
[0796]	gcggagcttg cagtgagctg agattgtgcc actgcactcc agcctgggtg acacagttag	6660
[0797]	actacgtctc aaaaaaaaaa aaaaaaaaaa tagctgagca ttatggtgta tgccgttagt	6720

[0798] cccagctact ggggaggctg aggtgggagg attgcttgag ccctaggagg gcaaggctgc 6780  
 [0799] agtgagccat gatcacacca ctgctttcca gcctcggttag gagagcaaga ccctatctca 6840  
 [0800] aaaaaaaaaa aaaaaaaaaa agaaaagaaa agaaaagaaa agaaaaagaa agagagaaag 6900  
 [0801] aaatacttag gtgtaaatct aaaaaacatg cgtagggcca ggtgcagtgg ctcatgcctg 6960  
 [0802] taatcccagc actttgggaa gttgaggctg gcgcatcact tgaagtcggg agtttgagac 7020  
 [0803] cagcctggcc aacatgggtga aaccccgctct ctactaaaaa tgcaaaaatt aggcaggtgt 7080  
 [0804] tgtggcgcac gcctgatccc agctactttg gaggtgagg caggagaatt gcttcaaccc 7140  
 [0805] gggaggcaga ggttgacgtg agccaagact gttccactgc actccagcct gggcaacaga 7200  
 [0806] gtaagagtct gtctcccgaa aaaaaaaaaa agaaaaaaga aagcattgaa ttgtatgcta 7260  
 [0807] aaaactacac gatgctgatt aaagaagtca aagaagatct aaatatatgg agagacatgc 7320  
 [0808] tgtactcatg gattgatgga ttggaagact caacataaga cagatatcaa ttttcccaa 7380  
 [0809] attaatatac aagtttaatc caattcctat aaaaatacca gcaagatttt ttgtagatat 7440  
 [0810] aaacaagttg gccaggtgta gtggttaca cctgtaatcc tagcactttg ggaggctgag 7500  
 [0811] gtgggaagat cgcttgagcc caggtgttca cgactgcagt gagctatgat tgtgtcactg 7560  
 [0812] cattccagct ggcactccag cctaagtac aaaggagac cctgtctcaa aaacaaaaac 7620  
 [0813] aaaacaaaaa taattttgct ctgcaaaatc cctattaaga agaagaaaag aggtctgggca 7680  
 [0814] cagtggctca ccgctgtaat cccagcacgt tgggaggctg aggcaggctg atcacttcag 7740  
 [0815] cccagaagtt tgagatcagc ctgggcaaca tgaggaaacc ccgtctctac caaaaaaaaa 7800  
 [0816] aaaaggtaca tacacacaca cacacacaca cacacacata cacaagtata tacacatata 7860  
 [0817] tatacacata caggtgaata gatgtatata catctattta ttgtgaatat acatctatac 7920  
 [0818] acacacgtgt gtgtacacat atatttataa tttattttta tttattttatt tatttttgag 7980  
 [0819] acagagtctt gctctgtcac ccaggctggg tgcacctgta ttcccaacga cacaggaggc 8040  
 [0820] tgagggtgga gaatcactga gccaggagg cagaggtgc agtgagccaa gatgttgcc 8100  
 [0821] ggttgccctg gcaacagagc gagaccctat atcaaaaaag aagaataata agaaaagaca 8160  
 [0822] gtttacagaa tataagaaaa tatattcaca atccacatac ttagcaaagg actggtatct 8220  
 [0823] agaatatgat aaacaactct caaaactcaa aaccaaaaaa atgaacaatt caattagaaa 8280  
 [0824] acaggccgaa aaggacatac agttggcaaa taagcacatg aaaagttgtt caacatcatt 8340  
 [0825] aatcattagg gatattgaca ttaaaaccac aataggctat cactaaacct atcagaatgg 8400  
 [0826] ctaaatacaa aattggaaca ccaccaaatg ctgatgagga tgtggagaaa ctgggtcatt 8460  
 [0827] cttccaatat tgggtggagg ctaaaatggc aaagccactc tggaaaacag tttgatagtt 8520  
 [0828] tcttataaaa caaaacatgc ggccgggcgc ggtagctcac gcctgtaatc ccagcacttt 8580  
 [0829] gggaggccga ggcgggtgga tcacgaggtc aggagatcga gaccatcctg gctaacacgg 8640  
 [0830] tgaaaccctg tctctactaa aaatacaaaa aattagccgg gcgtggtggc gggcgctgt 8700  
 [0831] agtcccagct actcgggagg ctgaggcagg agaattgtgt gaaccggga ggcggagctt 8760  
 [0832] gcagtgagcc gagatcgcgc cattgcactc caacctggga gacggaggga gactccgtct 8820  
 [0833] caaaaaaaca aaacaaca aacaaaaaac atgcaacaat ccagcaatat tgcacccta 8880  
 [0834] ggcatttata ctagagcaat gaagacttat gccacacaa aaagctgcac acaaatgttc 8940  
 [0835] atagcagctt tattcatggt agccaacaat tagaaacaat ctagatgtcc ttcaactggt 9000  
 [0836] gaatgattac atccatacca cgaaatactt ttacagcaata aaaagatga atcatagtac 9060  
 [0837] acaccacaac ctggatgaat ctccaggga ttatgctgag tgaaaaaag ccaatctcaa 9120  
 [0838] aagtaatat actgtattaa tccatttata taacattctt aaaataacta attatagaaa 9180  
 [0839] tggagaacag atgagtgatt gccaggggtt aaggggtca gggatgggga ggggaagggg 9240

[0840]	tatggctaca	aaaagcaaca	accttatggc	gccgaaaatg	ttctgtattc	tgattgtgtc	9300
[0841]	aatgtgagca	tactggttga	gatatagtgc	tacagttttg	caagttatta	ccatcagagt	9360
[0842]	aaactggata	gagggcacat	aggatttctc	tgtattactt	cttacaactg	caagtgaatc	9420
[0843]	tacaattatc	tcaaaaataat	aagtttagtt	taatgctagg	cgtggtggct	cacatctgta	9480
[0844]	atctcagctc	tttgggaggc	tgagacgggt	ggatggcttg	agtccaggag	ttcgagacca	9540
[0845]	gcctggccaa	catggcaaaa	cgggtctcta	ctaaaaatac	aaaaattagc	tgggcgtggt	9600
[0846]	ggcaagtgcc	tgtagtccca	gctactcggg	aggctgaggc	aggagaattg	cttgaacccg	9660
[0847]	ggaggtggag	gttgacagta	gccgagatca	cgcactaca	ctgtagcttg	ggcgacagag	9720
[0848]	tgaggctctt	tctcaaaaaa	aaaaaaaaaa	aaaaaaaaagc	aggcaggcag	ggccaggaaa	9780
[0849]	gcgtataatt	tttgtagtgc	aaatgactaa	cctaaaaagt	gaagattggc	caggcgcagt	9840
[0850]	ggctcacgcc	tgtaatccca	gcactttggg	aggccaaggc	gggtggatca	cgaggtcagg	9900
[0851]	agattgagcc	actctggcta	acacagtga	accccgcttc	tactaaaata	caaaaaatta	9960
[0852]	gctgggcgtg	gtggcaccgc	cctgtagttg	cagctacttg	ggaggctgag	gcaggagaat	10020
[0853]	cacttgaacc	caggaggcga	agttgcagcg	agccgagatc	acactactgc	actccagcct	10080
[0854]	gggtgacaaa	gtgagattct	gtctcaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaagt	10140
[0855]	aagtttacct	ttttttttaa	atttttcttc	ttttccttcc	ctactttgtg	agataatttt	10200
[0856]	cttcttttta	aaaagccaag	agcttacttc	tgtaagtaaa	gattatctta	agacaactta	10260
[0857]	gaaatgtata	ttattagtat	tttctatttc	attgtaagtt	atttgtaa	attgggtttg	10320
[0858]	gtgctaacct	agaattccat	caaattaatt	gtccccta	atatggccat	tatcattttg	10380
[0859]	tctaacattg	tatcctatta	acaatgctgt	aagtattatt	tttgtagcta	aattatggtt	10440
[0860]	tgcattttta	aattattggt	ttaaggataa	agttccagaa	atgaaattaa	ggatatgaac	10500
[0861]	tttttgagca	catcttgtca	gactgagta	gtattattta	aaacttttgg	ggggggcaat	10560
[0862]	tttataattg	aaaaatata	cattgtttta	atttgcat	ctttcactgc	ctatgagatt	10620
[0863]	aaaacaatgc	actactttcc	aaaaattctt	aagtcttttg	tgttgatgct	tttgttctgt	10680
[0864]	ttctatggat	ctcatcttcc	ttcagaacag	ctccccctcc	caacttctcg	atttctaaca	10740
[0865]	ataacagtat	caccctcctt	gttctcccaa	tttctgaaac	acagagtc	gtttttttct	10800
[0866]	ctgcttcaat	ccctggtttc	ctatcgtcat	caattatgac	ctttccttgc	tttgaaagt	10860
[0867]	ttttggcccg	ggcatgatgt	ctgccaccta	ttgtaatcct	agcactttgg	gaggctgagg	10920
[0868]	cggtggtgat	acttgacctg	aggatttcga	gaccagcctg	ggcaacagg	cgaaacctcg	10980
[0869]	tctctacaaa	aaatacaaaa	gttagtcggg	agtgggtgca	catgcttgta	gtcccagtta	11040
[0870]	cttggggggc	tgaggtggca	ggatctcttg	agcccacgag	gtagatgttg	cagtgagccg	11100
[0871]	tgattgcgcc	actgcacccc	agcctaggtg	acagagttag	accctgtctc	aaaaaaaaaa	11160
[0872]	aaatgttcta	gtttcttctt	cttctttgtt	cccatgggaa	tgccaccatc	accagccaag	11220
[0873]	gtcacatac	ctcccacctg	gattacagt	agcttccagg	taatttggtc	tgctactagt	11280
[0874]	ctcgctact	tggatttccc	ttccccctgc	tgcagcattg	ccttccaaag	ccatgctttg	11340
[0875]	cacatgccac	atcctagccc	attagactaa	gcctagaagc	ctctgcagga	cgttcacctt	11400
[0876]	ctcagcgcca	ctgctcagtt	tcccagtgga	aacctctgca	cccaggaggt	ttccccacag	11460
[0877]	cttgctgtg	ctgcctctct	ggagcttttc	tcccttctctg	taatgtctt	gtgtctctct	11520
[0878]	gtctctagtc	cattgcctat	acctcttttt	tttttttttt	ttgagatgga	gtctctctct	11580
[0879]	ctcatccagg	ctggagtga	gtggcgcat	ctcggtcac	tgcaacctt	gtctctctgg	11640
[0880]	ttcaagggat	tctcctgcct	cagcctcccg	agtaactggg	attacaggcg	tgcaccacca	11700
[0881]	ttcctggcta	atttttgtat	tttttagtaa	gactgggttt	caccatgttg	gccaggctgg	11760

[0882]	tcttgaactc	ctgccctcag	gtgatccacc	tgcctcggcc	tcccagagt	ctgggattac	11820
[0883]	aggcgtgagc	caccgcacct	gccacaggcc	catacctctt	ttaagtcttc	attcaatacc	11880
[0884]	agttgtccca	tgaatttgtc	ccagactcac	tcatatgctt	agacctttca	tattatcttg	11940
[0885]	ccatagcttt	ttcaaagtat	gggacagcat	ggacaagcag	gccatggttt	tcttttgaag	12000
[0886]	agaagcaagg	aggcagagtt	attttaggag	gagggttata	catttcattt	tgaaccaatt	12060
[0887]	gcgtttgggg	tgatggcagg	atattaacat	aaacttattt	cttgaccat	tggaaatgtg	12120
[0888]	tgcctagaac	tgaggagaga	ggtcagggt	ggcagtaaca	acttgccac	aatctgcaga	12180
[0889]	gctgactggg	gatgaggtg	aatttagaat	gtctgtagaa	acggggaaga	gaaccaaaga	12240
[0890]	cagagtctgg	gacaacacct	aaatgtagat	gtcagagcaa	gagttcaaga	cgaagaaaaa	12300
[0891]	cgaatcatat	ttagaaatgg	aggggaggaa	caaaagaggc	ggagcaaagt	ggggcagaac	12360
[0892]	cagagtaggc	cacgctttta	agaagtttg	taaaggaact	gtgaaaggaa	tgtagttgaa	12420
[0893]	tttcagggt	agctggggaa	ttaaagcagt	gtgtagatcc	agggcaaaca	gcaagtaggg	12480
[0894]	caggaaccac	tgaaggaaca	aataaagggg	gaggttgggt	ccaggttgct	ttgagtaggg	12540
[0895]	aagttttttt	aaaaagtgtg	aaactgaagg	tgtgggtgg	attgggtgcc	tgccgtgctc	12600
[0896]	tgaggaagct	tggggcaact	gtgtgctgag	gctgtgaggt	tgtctggaag	gggctcctgg	12660
[0897]	acagtaagag	ctgagcagtg	gggaagagga	ctgtgtggtc	tggaagagga	gagaaaggag	12720
[0898]	agtgagtga	tgaactggta	tccaggtcc	cacaccaagg	cagaaaggag	gagaggacct	12780
[0899]	gggcatctca	gggaggcaga	ggcagtagca	agcagggtga	gaggctttag	tcttagccac	12840
[0900]	ctttgcccc	ttcctccaaa	tatacattct	aagtaaaaac	aaaacaaaac	agaactgttt	12900
[0901]	gctatgtaaa	tttagcttct	aaagccctgt	tctacagaga	ttttggagct	tccactgcac	12960
[0902]	ccagaaaatg	cacagctaaa	gagaaaactt	cccttggtga	tggttattag	attttacaag	13020
[0903]	aagaggccaa	aggagacaca	tacttatgcc	agaagaactt	tccagagata	gcattgcata	13080
[0904]	gcgaaatagc	ctgaattatt	tttatttttt	aaaacatttt	ttcttttctt	ttttcttttc	13140
[0905]	tttttctttt	tttttttttt	tttttgagac	agagtctcac	tctgtcacc	aggctggagt	13200
[0906]	gcagtggcgt	gatcttggct	caactgcaat	tccacctccc	gggttcaagc	cattctctct	13260
[0907]	cctcagcctc	ccaagtagct	gggattacag	gcatgcgtca	ctatgctctg	gctaattttt	13320
[0908]	tttttctttt	tttttttggt	attttttagta	gagatggggt	ttcaccatgt	tgccagggt	13380
[0909]	ggtcttgaac	tctgacctc	aagtgatcca	ccgccttgcc	ctcccaaagt	gctgggattt	13440
[0910]	caggcgtgag	ccaccgcacc	cggccaaaaa	tttcttttct	ttaagatgag	gcctcactct	13500
[0911]	gttggccagg	ctggagtgea	gtgttacaat	catagctcac	tgtaactttg	aactcctggg	13560
[0912]	ctcaagtgat	cctcctgctt	cagcctctca	agtagctggg	attacaggca	tgtgccacca	13620
[0913]	caccagctta	atttttttta	aaataatttt	tttttagagac	gagggtctcg	attggctgcc	13680
[0914]	taggttggtc	ccagactcct	gacgggtgct	attttaatcc	tagctccacc	acttacggga	13740
[0915]	gtcaaaaattc	aaaagataga	aaagggcata	taggctgggt	gcagtggctc	acacctgcaa	13800
[0916]	tcccagcaat	ttgggaggct	gaggtgggct	ggttgcttga	ggtcaggagt	tcgagatcag	13860
[0917]	cctgggcaac	atggcaaaac	ttgtatctac	taaaaatata	aaaattagcc	agatgtgggt	13920
[0918]	gtgtacacct	gtaatcccag	ctactccgaa	ggctgaggca	agagaatccc	ttgaactcag	13980
[0919]	gaggcagagg	ttacaatgag	cagagatcga	acactcgact	ccataaaaac	aaacaaacaa	14040
[0920]	aaaaagaaag	caggctgggt	gtggtggctc	acgcctgtaa	ccccagcact	tcgggaggcc	14100
[0921]	aaggcgagcg	gatcacctga	ggttgggcat	tcgagaccag	cctgaccaac	aaggagaaac	14160
[0922]	cctgtctcta	ctgaaaatac	aaaattagcc	gggcttggtt	gcgcatgcct	gtaatctcag	14220
[0923]	ctactcggga	ggcagaggca	agataattgc	ttgaacccgg	gaggcggagg	ttgcggtgag	14280

[0924]	ccaagatcat gccattgcac tccaacctgg gcaacaatag cgaaactcca tctcaaaaaa	14340
[0925]	aaaaaagcaa agggcatata gtgaaaagct ttcttcttac acatgagtat tcaattcctc	14400
[0926]	ttcctagagg caaccaaggt tatttttgtg tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg	14460
[0927]	tgttttggga cagtctcact ctctcaccaa ggctggaatg cagtgggtgcg atctcactgc	14520
[0928]	aaactctgcc tcccagtcct aagcagtcct gtgcctcagc ctcccagttt ttttttcttt	14580
[0929]	taaatgggggt ctatttctgt cgcccagggt ggagtgcagt ggcatgatca tagctcactg	14640
[0930]	cagcctcgac ctctgggtc aggttatcct cccacctcag cctccggcat agctggggct	14700
[0931]	actggcatgc accaccacac tcagttaatt ttttttcttt tttagacag agtctcactc	14760
[0932]	tgtcacctag actggagtgc agtggtgcca tctcatttgt ttcactgcaa cctttgactt	14820
[0933]	ctgggctcaa gtgattctcc cacctcagcc tccaaggcgg gctaattaaa aaaaattttt	14880
[0934]	tttttttttt ttttagagat ggggtttcgc catgttggcc aggctgatct cgaactcctg	14940
[0935]	ggcacaacaa atctttccac ctcgatcttt caaagagctg ggatgagaga tttccacat	15000
[0936]	gcctggcctc attttctttt ttaatttttt tttagacatt atagctcttt ttaatggcct	15060
[0937]	cattttctta tgtttaattc gagaattatt cttttcatat acaaagaata tattttctcc	15120
[0938]	acctttaaaa acaaatagta gactgtttaa catctcgctt tattcagtta gtgatgtttc	15180
[0939]	ttagatacgg gtccaaatta gtacacaaag cacttctca ttcctctctt acggctgcct	15240
[0940]	agcagtccac tgaatgggtg agctatgac tatttaacct attctttatt gatggacatt	15300
[0941]	tggttttgta tatacatttg taattctgta tagattacaa atcaccatcc aaagaaattg	15360
[0942]	tactggttta ttctctaca atgtgtgaga gttgggtaat tacttaatct caatatgtga	15420
[0943]	gagtttaggc agttacctaa tctctctgag tctcagttt tctatctgca aaataaacia	15480
[0944]	aacagtgttg acagtatcta tttctcgga ttattgtgga gattactgag atgatgcctg	15540
[0945]	taaagtattt ggcatgtagg agttgggtgt ctccaaataa ggatatgatt ttatttgtat	15600
[0946]	ttgtgagcta ctgtcccagc caggtaaatg gatatgatga gacctcctg ccagaccggg	15660
[0947]	tttctctgat tagaacgagg agcagatgtt gcaggaaatt agcaactgat atcagaagag	15720
[0948]	ccgtgggcat tctcttgcca gaggtgccct gtctccaggg cgcctcagtc ccccccata	15780
[0949]	tgtctctgct tcccaggctc atccaagctg aatgatcgcg ctgactcaag aagaagcctt	15840
[0950]	tgggaccaag gaaactttcc cctgatcatc aagaatctta agatagaaga ctgagatact	15900
[0951]	tacatctgtg aagtggagga ccagaaggag gaggtgcaat tgctagtgtt cggtatgtgag	15960
[0952]	tggggcagggt ggggatgagg atacctcctg cctgggtccc tccccacta ctcccaccc	16020
[0953]	tgcaccaa at ccagcctgag ctggtgatac cgcagcagcc ccaagaggac caggctgtca	16080
[0954]	aactggcctc caaatgtctt aaaacccttc ttgatcaggt gagggatgct ggtgggcgga	16140
[0955]	ggagggaaga ggccttgga aaaggaaaga aaagggaagg aggcaaggga aggagggaga	16200
[0956]	gagactgggg aagagaggat gaggggagag gaggaagaa gagagagagg aggggagagg	16260
[0957]	gaaaccctat cttggctggg ggtgcgcagc tgggtgctgg gaggaaggag atgttgggac	16320
[0958]	ggcgataatg gagagatgtt gttggtttcc tgttgtctgc ctttctcctt ggggatggta	16380
[0959]	tgtgtgtgac acagctggcc tttccctcca cagtactgc caactctgac acccactgc	16440
[0960]	ttcaggggca gagcctgacc ctgaccttg agagcccccc tggtagtagc ccctcagtgc	16500
[0961]	aatgtaggag tccaaggggt aaaaacatac agggggggaa gacctctcc gtgtctcagc	16560
[0962]	tggagctcca ggatagtggc acctggacat gactgtctt gcagaaccag aagaaggtgg	16620
[0963]	agttcaaat agacatctg gtgctaggta aggaagccc ctcttcgcgc agtctctcc	16680
[0964]	ctgccccagg ggctgacagc ccctccctct gctctgactg ccctgtttct ggttctgggtg	16740
[0965]	ctgggagggtc aggagtggag aagactaggt ccctagagc tgaggcctgt cttgaaggac	16800

[0966]	tcactggggc cctcatcctc agggggctga ttggcagcca cccctcagtg tggtggacat	16860
[0967]	ggagaaagga aaggctgggg aaggtaagga tgctagaggc ccgagtctcc tttggaggcc	16920
[0968]	ccaaaggagg aatgtcaggg agcttacttt ctttgttgcc tcagctccac acccctacca	16980
[0969]	agttggcaaa tccacttact cagggacact aacaccagta agccaaccct gatgatgttc	17040
[0970]	tatgtgttac ctctggacct ctaagccagg ccactgtggg gagaccaagg tcctacccca	17100
[0971]	gatcctgtcc cctgggtgct tatgtgactt aaggtagaca taaggtagtg tgccagttaa	17160
[0972]	gtgcatgtac gctgattgaa atcctgggtc tgccacaacc atgtgacctt gggtagttaa	17220
[0973]	ctaaacctct ctgcaccttg gtttcagcct ctgtgaaatg gggatgatgt taactgccat	17280
[0974]	agtactacc tcgtattaag ttgaggactg atatacgtaa ggcactgaaa atggtgcctg	17340
[0975]	gcacagagta agccctagtt aagtgttcgc tgttatittg tgaagggtga tgaatacgcc	17400
[0976]	tctaaggagt ggaggccaaa tggcttctgt ggtccaggaa tcctaaggac agcaaggatc	17460
[0977]	ccctgtggct gggctgctct gtgatggctt ccgggaggag ggaggtggcc tgctgtagga	17520
[0978]	aaatgctggg tggaagaagg gagagaaggc tggagaggtg ggaaggaact gaagtatctg	17580
[0979]	aagtgacaag gtgggtgtct ggactcgtcg ggtcccttc catctccctg ctgcctccac	17640
[0980]	atgccaaccc cactcgtgca ccctcatctt cctatctcct caccagggt ctctcccttc	17700
[0981]	ccacctccag ctttccagaa ggcctccagc atagtctata agaaagagg ggaacagggtg	17760
[0982]	gagttctect tcccactcgc ctttacagtt gaaaagctga cgggcagtgg cgagctgtgg	17820
[0983]	tggcaggcgg agagggttc ctctccaag tcttgatca ctttgacct gaagaacaag	17880
[0984]	gaagtgtctg taaaacgggt taccaggac cctaagctcc agatgggcaa gaagctccc	17940
[0985]	ctccacctca ccctgcccc ggccttgctt cagtatgtg gctctggaaa cctcacctg	18000
[0986]	gccctgaag cgaaaacagg aaagttgcat caggaagtga acctggtggt gatgagaggt	18060
[0987]	gaggggccag gccaggagg ggtgggcagg ggaaggagtt ggagggcct ggcccagggc	18120
[0988]	tccctctgag gcaagccagg cccaagagg ggatgcctag gccctggtca cctggatgaa	18180
[0989]	gtgaggagg gccctctggg tttgggctg gttttgaact gagacatcca tgagccagcc	18240
[0990]	tgggctggc ttactgaag atc	18263
[0991]	<210>	78
[0992]	<211>	458
[0993]	<212>	PRT
[0994]	<213>	人工序列
[0995]	<220>	
[0996]	<223>	合成
[0997]	<400>	78
[0998]	Met Cys Arg Ala Ile Ser Leu Arg Arg Leu Leu Leu Leu Leu Gln	
[0999]	1 5 10 15	
[1000]	Leu Ser Gln Leu Leu Ala Val Thr Gln Gly Lys Lys Val Val Leu Gly	
[1001]	20 25 30	
[1002]	Lys Lys Gly Asp Thr Val Glu Leu Thr Cys Thr Ala Ser Gln Lys Lys	
[1003]	35 40 45	
[1004]	Ser Ile Gln Phe His Trp Lys Asn Ser Asn Gln Ile Lys Ile Leu Gly	
[1005]	50 55 60	
[1006]	Asn Gln Gly Ser Phe Leu Thr Lys Gly Pro Ser Lys Leu Asn Asp Arg	
[1007]	65 70 75 80	



[1008]	Ala Asp Ser Arg Arg Ser Leu Trp Asp Gln Gly Asn Phe Pro Leu Ile
[1009]	85 90 95
[1010]	Ile Lys Asn Leu Lys Ile Glu Asp Ser Asp Thr Tyr Ile Cys Glu Val
[1011]	100 105 110
[1012]	Glu Asp Gln Lys Glu Glu Val Gln Leu Leu Val Phe Gly Leu Thr Ala
[1013]	115 120 125
[1014]	Asn Ser Asp Thr His Leu Leu Gln Gly Gln Ser Leu Thr Leu Thr Leu
[1015]	130 135 140
[1016]	Glu Ser Pro Pro Gly Ser Ser Pro Ser Val Gln Cys Arg Ser Pro Arg
[1017]	145 150 155 160
[1018]	Gly Lys Asn Ile Gln Gly Gly Lys Thr Leu Ser Val Ser Gln Leu Glu
[1019]	165 170 175
[1020]	Leu Gln Asp Ser Gly Thr Trp Thr Cys Thr Val Leu Gln Asn Gln Lys
[1021]	180 185 190
[1022]	Lys Val Glu Phe Lys Ile Asp Ile Val Val Leu Ala Phe Gln Lys Ala
[1023]	195 200 205
[1024]	Ser Ser Ile Val Tyr Lys Lys Glu Gly Glu Gln Val Glu Phe Ser Phe
[1025]	210 215 220
[1026]	Pro Leu Ala Phe Thr Val Glu Lys Leu Thr Gly Ser Gly Glu Leu Trp
[1027]	225 230 235 240
[1028]	Trp Gln Ala Glu Arg Ala Ser Ser Ser Lys Ser Trp Ile Thr Phe Asp
[1029]	245 250 255
[1030]	Leu Lys Asn Lys Glu Val Ser Val Lys Arg Val Thr Gln Asp Pro Lys
[1031]	260 265 270
[1032]	Leu Gln Met Gly Lys Lys Leu Pro Leu His Leu Thr Leu Pro Gln Ala
[1033]	275 280 285
[1034]	Leu Pro Gln Tyr Ala Gly Ser Gly Asn Leu Thr Leu Ala Leu Glu Ala
[1035]	290 295 300
[1036]	Lys Thr Gly Lys Leu His Gln Glu Val Asn Leu Val Val Met Arg Val
[1037]	305 310 315 320
[1038]	Ala Gln Leu Asn Asn Thr Leu Thr Cys Glu Val Met Gly Pro Thr Ser
[1039]	325 330 335
[1040]	Pro Lys Met Arg Leu Thr Leu Lys Gln Glu Asn Gln Glu Ala Arg Val
[1041]	340 345 350
[1042]	Ser Glu Glu Gln Lys Val Val Gln Val Val Ala Pro Glu Thr Gly Leu
[1043]	355 360 365
[1044]	Trp Gln Cys Leu Leu Ser Glu Gly Asp Lys Val Lys Met Asp Ser Arg
[1045]	370 375 380
[1046]	Ile Gln Val Leu Ser Arg Gly Val Asn Gln Thr Val Phe Leu Ala Cys
[1047]	385 390 395 400
[1048]	Val Leu Gly Gly Ser Phe Gly Phe Leu Gly Phe Leu Gly Leu Cys Ile
[1049]	405 410 415

[1050]	Leu Cys Cys Val Arg Cys Arg His Gln Gln Arg Gln Ala Ala Arg Met
[1051]	420 425 430
[1052]	Ser Gln Ile Lys Arg Leu Leu Ser Glu Lys Lys Thr Cys Gln Cys Pro
[1053]	435 440 445
[1054]	His Arg Met Gln Lys Ser His Asn Leu Ile
[1055]	450 455
[1056]	<210> 79
[1057]	<211> 293
[1058]	<212> PRT
[1059]	<213> 人工序列
[1060]	<220>
[1061]	<223> 合成
[1062]	<400> 79
[1063]	Lys Lys Val Val Leu Gly Lys Lys Gly Asp Thr Val Glu Leu Thr Cys
[1064]	1 5 10 15
[1065]	Thr Ala Ser Gln Lys Lys Ser Ile Gln Phe His Trp Lys Asn Ser Asn
[1066]	20 25 30
[1067]	Gln Ile Lys Ile Leu Gly Asn Gln Gly Ser Phe Leu Thr Lys Gly Pro
[1068]	35 40 45
[1069]	Ser Lys Leu Asn Asp Arg Ala Asp Ser Arg Arg Ser Leu Trp Asp Gln
[1070]	50 55 60
[1071]	Gly Asn Phe Pro Leu Ile Ile Lys Asn Leu Lys Ile Glu Asp Ser Asp
[1072]	65 70 75 80
[1073]	Thr Tyr Ile Cys Glu Val Glu Asp Gln Lys Glu Glu Val Gln Leu Leu
[1074]	85 90 95
[1075]	Val Phe Gly Leu Thr Ala Asn Ser Asp Thr His Leu Leu Gln Gly Gln
[1076]	100 105 110
[1077]	Ser Leu Thr Leu Thr Leu Glu Ser Pro Pro Gly Ser Ser Pro Ser Val
[1078]	115 120 125
[1079]	Gln Cys Arg Ser Pro Arg Gly Lys Asn Ile Gln Gly Gly Lys Thr Leu
[1080]	130 135 140
[1081]	Ser Val Ser Gln Leu Glu Leu Gln Asp Ser Gly Thr Trp Thr Cys Thr
[1082]	145 150 155 160
[1083]	Val Leu Gln Asn Gln Lys Lys Val Glu Phe Lys Ile Asp Ile Val Val
[1084]	165 170 175
[1085]	Leu Ala Phe Gln Lys Ala Ser Ser Ile Val Tyr Lys Lys Glu Gly Glu
[1086]	180 185 190
[1087]	Gln Val Glu Phe Ser Phe Pro Leu Ala Phe Thr Val Glu Lys Leu Thr
[1088]	195 200 205
[1089]	Gly Ser Gly Glu Leu Trp Trp Gln Ala Glu Arg Ala Ser Ser Ser Lys
[1090]	210 215 220
[1091]	Ser Trp Ile Thr Phe Asp Leu Lys Asn Lys Glu Val Ser Val Lys Arg

[1092]	225	230	235	240
[1093]	Val Thr Gln Asp Pro Lys Leu Gln Met Gly Lys Lys Leu Pro Leu His			
[1094]		245	250	255
[1095]	Leu Thr Leu Pro Gln Ala Leu Pro Gln Tyr Ala Gly Ser Gly Asn Leu			
[1096]		260	265	270
[1097]	Thr Leu Ala Leu Glu Ala Lys Thr Gly Lys Leu His Gln Glu Val Asn			
[1098]		275	280	285
[1099]	Leu Val Val Met Arg			
[1100]	290			
[1101]	<210> 80			
[1102]	<211> 142			
[1103]	<212> DNA			
[1104]	<213> 人工序列			
[1105]	<220>			
[1106]	<223> 合成			
[1107]	<400> 80			
[1108]	tgtttgctg tgacatgaac tcattgtgac acaaaccact gtgctagggg ggatccacta 60			
[1109]	gtaacggccg ccagtgtgct ggaattcgcc ctgcgaaggg ccaggcatat aagtacacaa 120			
[1110]	taaacaaatg gcagctctct cc 142			
[1111]	<210> 81			
[1112]	<211> 99			
[1113]	<212> DNA			
[1114]	<213> 人工序列			
[1115]	<220>			
[1116]	<223> 合成			
[1117]	<400> 81			
[1118]	cccctccttc cttccccagg cactttccaa gtgtcaactc tagagcctat cgcggccgca 60			
[1119]	ccggtataac ttcgtataat gtatgctata cgaagttat 99			
[1120]	<210> 82			
[1121]	<211> 90			
[1122]	<212> DNA			
[1123]	<213> 人工序列			
[1124]	<220>			
[1125]	<223> 合成			
[1126]	<400> 82			
[1127]	ataacttcgt ataatgtatg ctatacgaag ttatgtcgac gtagcctatt tctctagatc 60			
[1128]	caaaatgatg acaacaaaag gtaccttgtg 90			
[1129]	<210> 83			
[1130]	<211> 210			
[1131]	<212> PRT			
[1132]	<213> 人工序列			
[1133]	<220>			

[1134]	<223> 合成
[1135]	<400> 83
[1136]	Met Gln Pro Trp Leu Trp Leu Val Phe Ser Met Lys Leu Ala Val Leu
[1137]	1 5 10 15
[1138]	His Gly Asn Ser Val Leu Gln Gln Thr Pro Ala Tyr Ile Lys Val Gln
[1139]	20 25 30
[1140]	Thr Asn Lys Met Val Met Leu Ser Cys Glu Ala Lys Ile Ser Leu Ser
[1141]	35 40 45
[1142]	Asn Met Arg Ile Tyr Trp Leu Arg Gln Arg Gln Ala Pro Ser Ser Asp
[1143]	50 55 60
[1144]	Ser His His Glu Phe Leu Ala Leu Trp Asp Ser Ala Lys Gly Thr Ile
[1145]	65 70 75 80
[1146]	His Gly Glu Glu Val Glu Gln Glu Lys Ile Ala Val Phe Arg Asp Ala
[1147]	85 90 95
[1148]	Ser Arg Phe Ile Leu Asn Leu Thr Ser Val Lys Pro Glu Asp Ser Gly
[1149]	100 105 110
[1150]	Ile Tyr Phe Cys Met Ile Val Gly Ser Pro Glu Leu Thr Phe Gly Lys
[1151]	115 120 125
[1152]	Gly Thr Gln Leu Ser Val Val Asp Phe Leu Pro Thr Thr Ala Gln Pro
[1153]	130 135 140
[1154]	Thr Lys Lys Ser Thr Leu Lys Lys Arg Val Cys Arg Leu Pro Arg Pro
[1155]	145 150 155 160
[1156]	Glu Thr Gln Lys Gly Leu Thr Cys Ser Leu Thr Thr Leu Ser Leu Leu
[1157]	165 170 175
[1158]	Val Val Cys Ile Leu Leu Leu Leu Ala Phe Leu Gly Val Ala Val Tyr
[1159]	180 185 190
[1160]	Phe Tyr Cys Val Arg Arg Arg Ala Arg Ile His Phe Met Lys Gln Phe
[1161]	195 200 205
[1162]	His Lys
[1163]	210
[1164]	<210> 84
[1165]	<211> 151
[1166]	<212> PRT
[1167]	<213> 人工序列
[1168]	<220>
[1169]	<223> 合成
[1170]	<400> 84
[1171]	Val Leu His Gly Asn Ser Val Leu Gln Gln Thr Pro Ala Tyr Ile Lys
[1172]	1 5 10 15
[1173]	Val Gln Thr Asn Lys Met Val Met Leu Ser Cys Glu Ala Lys Ile Ser
[1174]	20 25 30
[1175]	Leu Ser Asn Met Arg Ile Tyr Trp Leu Arg Gln Arg Gln Ala Pro Ser

[1176]	35	40	45
[1177]	Ser Asp Ser His His Glu Phe Leu Ala Leu Trp Asp Ser Ala Lys Gly		
[1178]	50	55	60
[1179]	Thr Ile His Gly Glu Glu Val Glu Gln Glu Lys Ile Ala Val Phe Arg		
[1180]	65	70	75
[1181]	Asp Ala Ser Arg Phe Ile Leu Asn Leu Thr Ser Val Lys Pro Glu Asp		
[1182]	85	90	95
[1183]	Ser Gly Ile Tyr Phe Cys Met Ile Val Gly Ser Pro Glu Leu Thr Phe		
[1184]	100	105	110
[1185]	Gly Lys Gly Thr Gln Leu Ser Val Val Asp Phe Leu Pro Thr Thr Ala		
[1186]	115	120	125
[1187]	Gln Pro Thr Lys Lys Ser Thr Leu Lys Lys Arg Val Cys Arg Leu Pro		
[1188]	130	135	140
[1189]	Arg Pro Glu Thr Gln Lys Gly		
[1190]	145	150	
[1191]	<210> 85		
[1192]	<211> 100		
[1193]	<212> DNA		
[1194]	<213> 人工序列		
[1195]	<220>		
[1196]	<223> 合成		
[1197]	<400> 85		
[1198]	tgaacctgct gctgctgggt gagtcgatta tcctggggag tggagaagct aggccgagcc 60		
[1199]	agttccgggt gtcgccgtg gatcggacct ggaacctggg 100		
[1200]	<210> 86		
[1201]	<211> 90		
[1202]	<212> DNA		
[1203]	<213> 人工序列		
[1204]	<220>		
[1205]	<223> 合成		
[1206]	<400> 86		
[1207]	atgccaggga cagccctgat actgtaggta gagtcaaggg ctgtccaagt accggtataa 60		
[1208]	cttcgtataa ggtatcctat acgaagttat 90		
[1209]	<210> 87		
[1210]	<211> 89		
[1211]	<212> DNA		
[1212]	<213> 人工序列		
[1213]	<220>		
[1214]	<223> 合成		
[1215]	<400> 87		
[1216]	ataacttcgt ataaggtatc ctatacgaag ttatctcgac ctgatcttgg agggagacct 60		
[1217]	ggaccgggag acgtgctggg ggcagggtt 89		

[1218]	<210> 88
[1219]	<211> 243
[1220]	<212> PRT
[1221]	<213> 人工序列
[1222]	<220>
[1223]	<223> 合成
[1224]	<400> 88
[1225]	Met Ala Ser Pro Leu Thr Arg Phe Leu Ser Leu Asn Leu Leu Leu Leu
[1226]	1 5 10 15
[1227]	Gly Glu Ser Ile Ile Leu Gly Ser Gly Glu Ala Arg Pro Ser Gln Phe
[1228]	20 25 30
[1229]	Arg Val Ser Pro Leu Asp Arg Thr Trp Asn Leu Gly Glu Thr Val Glu
[1230]	35 40 45
[1231]	Leu Lys Cys Gln Val Leu Leu Ser Asn Pro Thr Ser Gly Cys Ser Trp
[1232]	50 55 60
[1233]	Leu Phe Gln Pro Arg Gly Ala Ala Ala Ser Pro Thr Phe Leu Leu Tyr
[1234]	65 70 75 80
[1235]	Leu Ser Gln Asn Lys Pro Lys Ala Ala Glu Gly Leu Asp Thr Gln Arg
[1236]	85 90 95
[1237]	Phe Ser Gly Lys Arg Leu Gly Asp Thr Phe Val Leu Thr Leu Ser Asp
[1238]	100 105 110
[1239]	Phe Arg Arg Glu Asn Glu Gly Tyr Tyr Phe Cys Ser Ala Leu Ser Asn
[1240]	115 120 125
[1241]	Ser Ile Met Tyr Phe Ser His Phe Val Pro Val Phe Leu Pro Ala Lys
[1242]	130 135 140
[1243]	Pro Thr Thr Thr Pro Ala Pro Arg Pro Pro Thr Pro Ala Pro Thr Ile
[1244]	145 150 155 160
[1245]	Ala Ser Gln Pro Leu Ser Leu Arg Pro Glu Ala Cys Arg Pro Ala Ala
[1246]	165 170 175
[1247]	Gly Gly Ala Val Lys Gly Thr Gly Leu Asp Phe Ala Cys Asp Ile Tyr
[1248]	180 185 190
[1249]	Ile Trp Ala Pro Leu Ala Gly Ile Cys Val Ala Leu Leu Leu Ser Leu
[1250]	195 200 205
[1251]	Ile Ile Thr Leu Ile Cys Tyr His Arg Ser Arg Lys Arg Val Cys Lys
[1252]	210 215 220
[1253]	Cys Pro Arg Pro Leu Val Arg Gln Glu Gly Lys Pro Arg Pro Ser Glu
[1254]	225 230 235 240
[1255]	Lys Ile Val
[1256]	<210> 89
[1257]	<211> 152
[1258]	<212> PRT
[1259]	<213> 人工序列

[1260]	<220>
[1261]	<223> 合成
[1262]	<400> 89
[1263]	Arg Pro Ser Gln Phe Arg Val Ser Pro Leu Asp Arg Thr Trp Asn Leu
[1264]	1 5 10 15
[1265]	Gly Glu Thr Val Glu Leu Lys Cys Gln Val Leu Leu Ser Asn Pro Thr
[1266]	20 25 30
[1267]	Ser Gly Cys Ser Trp Leu Phe Gln Pro Arg Gly Ala Ala Ala Ser Pro
[1268]	35 40 45
[1269]	Thr Phe Leu Leu Tyr Leu Ser Gln Asn Lys Pro Lys Ala Ala Glu Gly
[1270]	50 55 60
[1271]	Leu Asp Thr Gln Arg Phe Ser Gly Lys Arg Leu Gly Asp Thr Phe Val
[1272]	65 70 75 80
[1273]	Leu Thr Leu Ser Asp Phe Arg Arg Glu Asn Glu Gly Tyr Tyr Phe Cys
[1274]	85 90 95
[1275]	Ser Ala Leu Ser Asn Ser Ile Met Tyr Phe Ser His Phe Val Pro Val
[1276]	100 105 110
[1277]	Phe Leu Pro Ala Lys Pro Thr Thr Thr Pro Ala Pro Arg Pro Pro Thr
[1278]	115 120 125
[1279]	Pro Ala Pro Thr Ile Ala Ser Gln Pro Leu Ser Leu Arg Pro Glu Ala
[1280]	130 135 140
[1281]	Cys Arg Pro Ala Ala Gly Gly Ala
[1282]	145 150
[1283]	<210> 90
[1284]	<211> 149
[1285]	<212> DNA
[1286]	<213> 人工序列
[1287]	<220>
[1288]	<223> 5'端嵌合人/小鼠MHC I基因座的序列
[1289]	小鼠/人序列的接合区
[1290]	<400> 90
[1291]	agtgtcgccg cggacgctgg atataaagtc caccgagccc gcagaactca gaagtcgcga 60
[1292]	atcgccgaca ggtgcgatgg ccgtcatggc gccccgaacc ctcgtcctgc tactctcggg 120
[1293]	ggctctggcc ctgaccaga cctgggcg 149
[1294]	<210> 91
[1295]	<211> 159
[1296]	<212> DNA
[1297]	<213> 人工序列
[1298]	<220>
[1299]	<223> 3'端嵌合人/小鼠MHC I基因座的序列
[1300]	人/小鼠序列的接合区
[1301]	<400> 91

---

[1302] ggtgggtgcct tctggacagg agcagagata cacctgccat gtgcagcatg agggtttgcc 60  
[1303] caagcccctc accctgagat ggggtaagga gagtgtgggt gcagagctgg ggtcaggga 120  
[1304] agctggagct ttctgcagac cctgagctgc tcagggtg 159



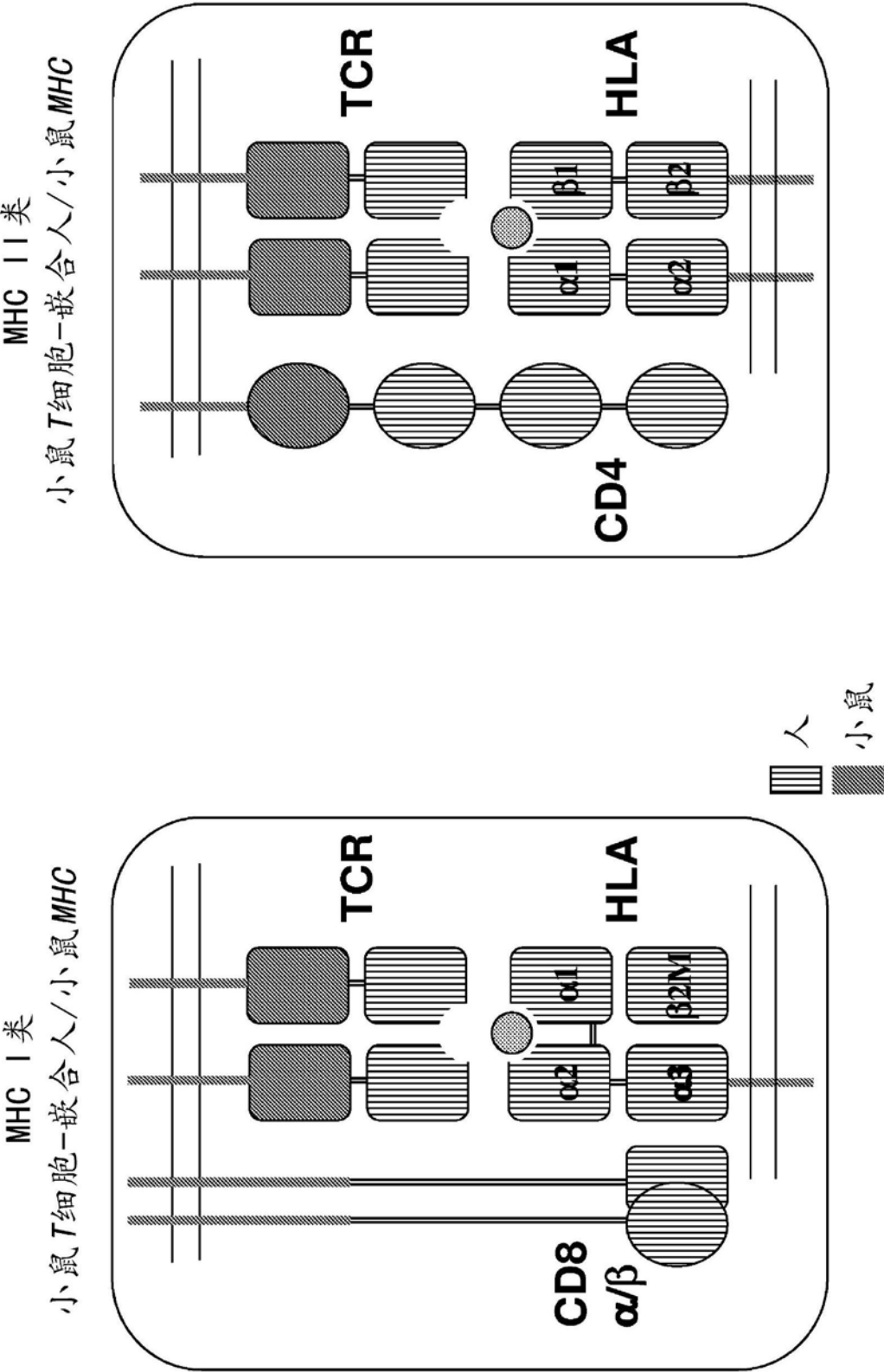


图1

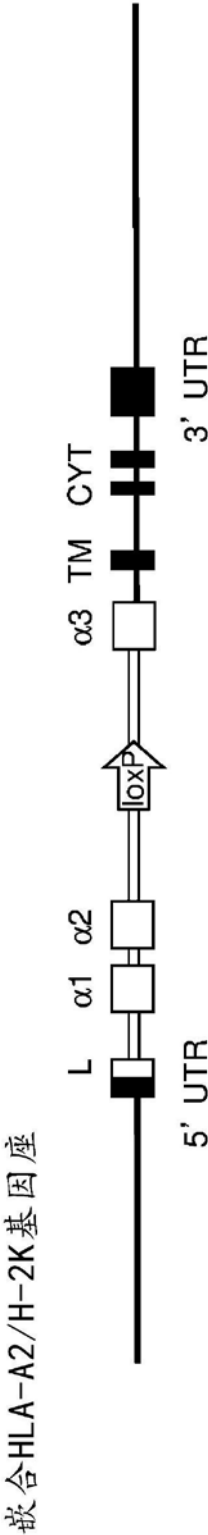


图2A

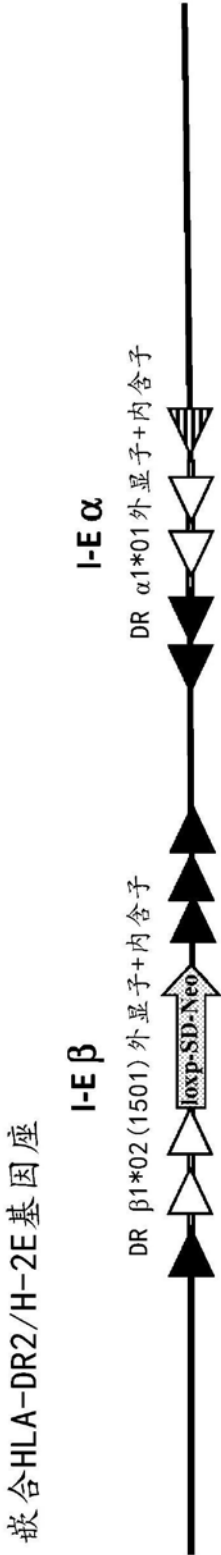


图2B

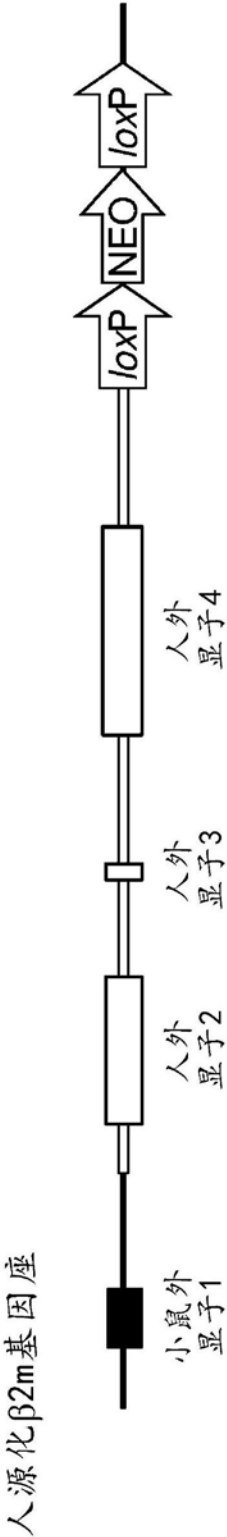


图2C

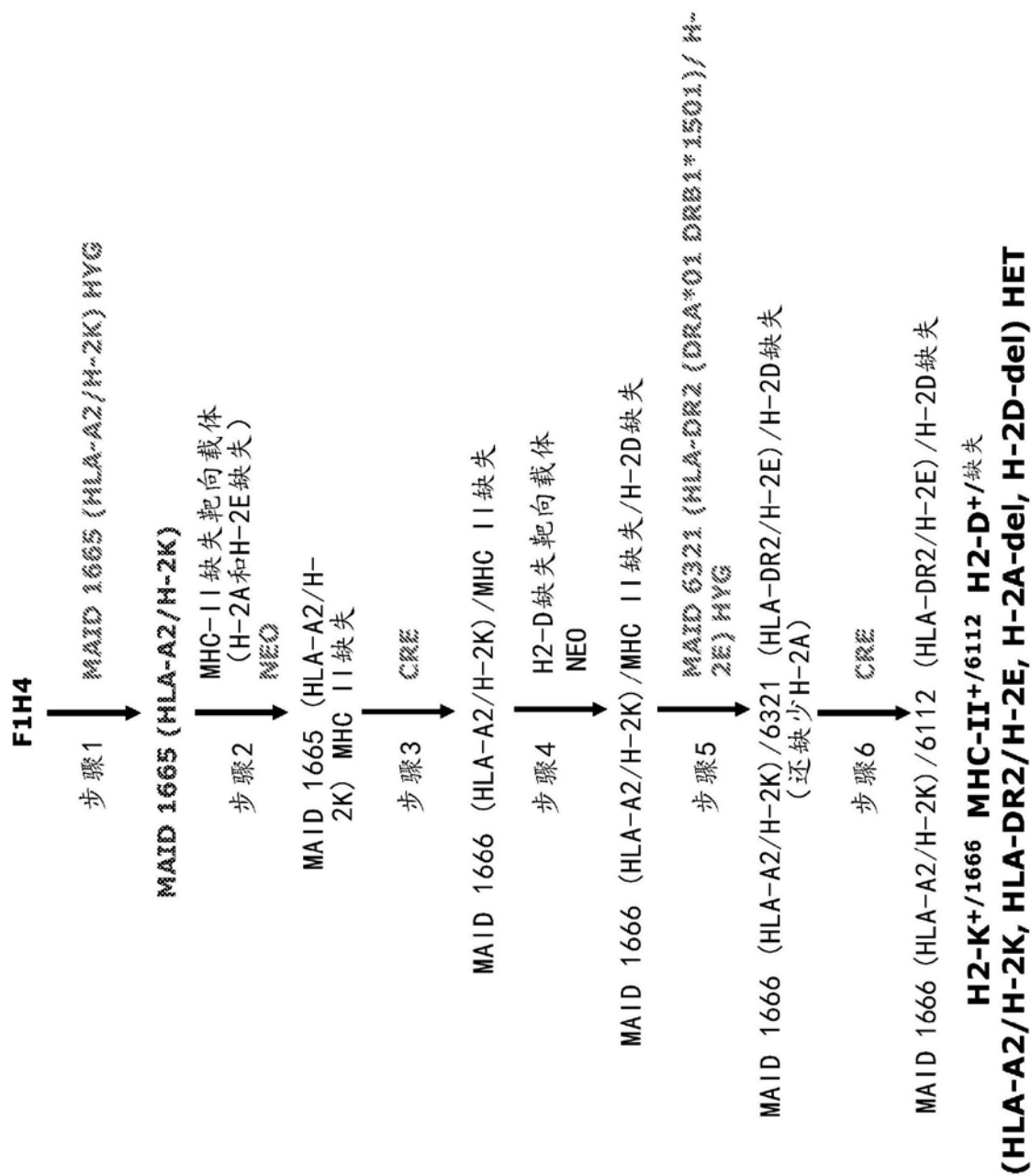


图3A

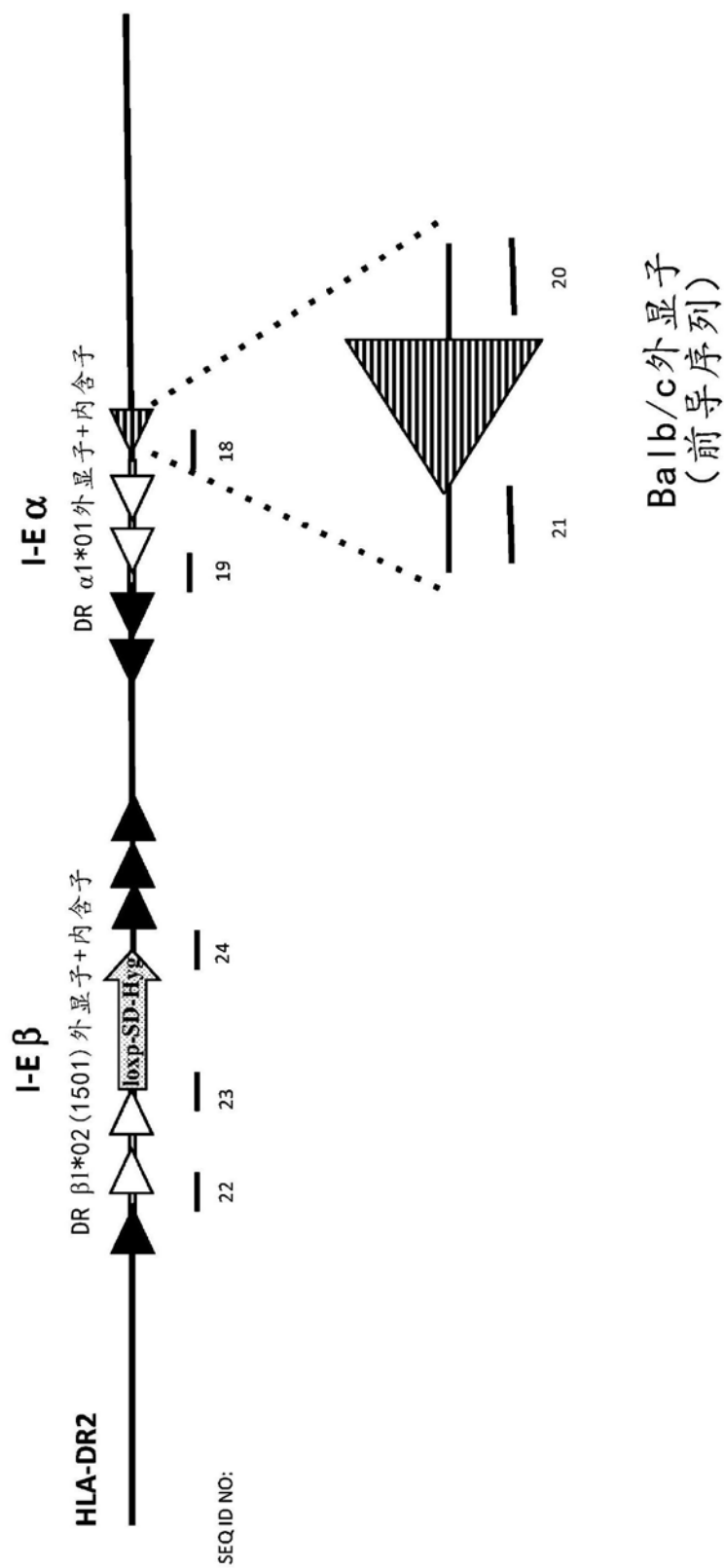


图3B

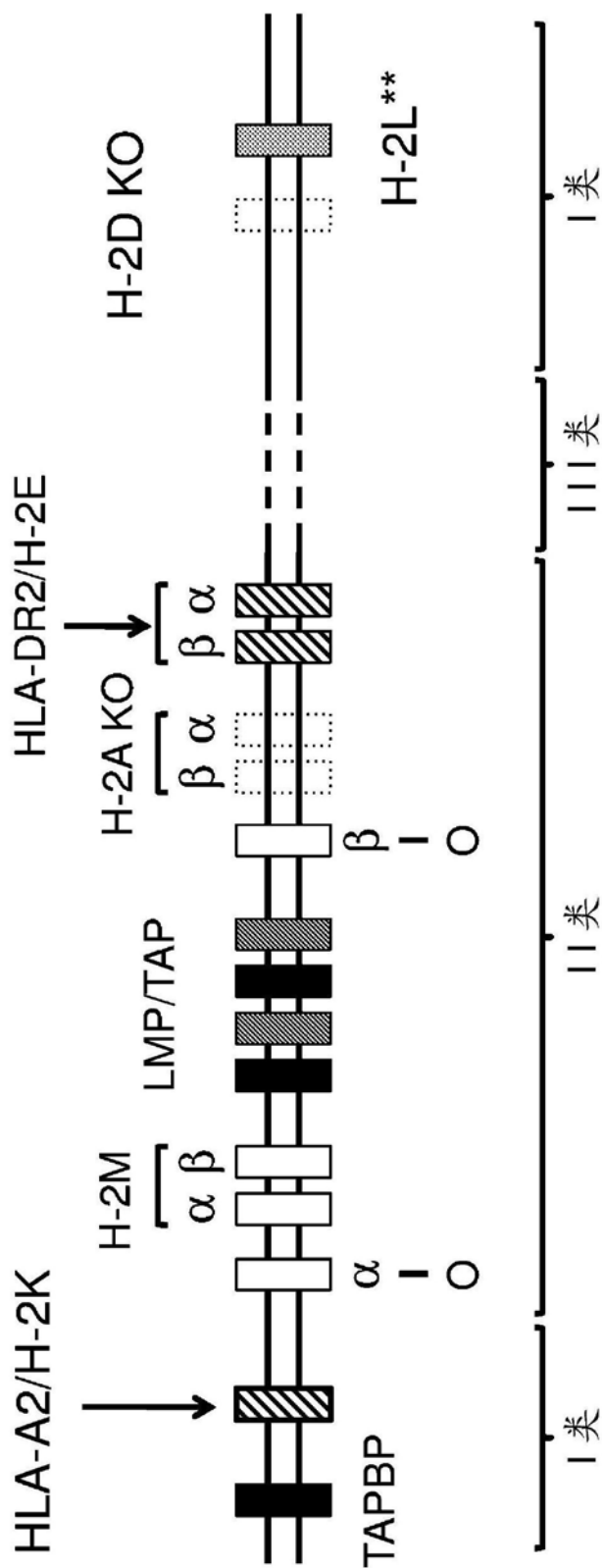


图3C

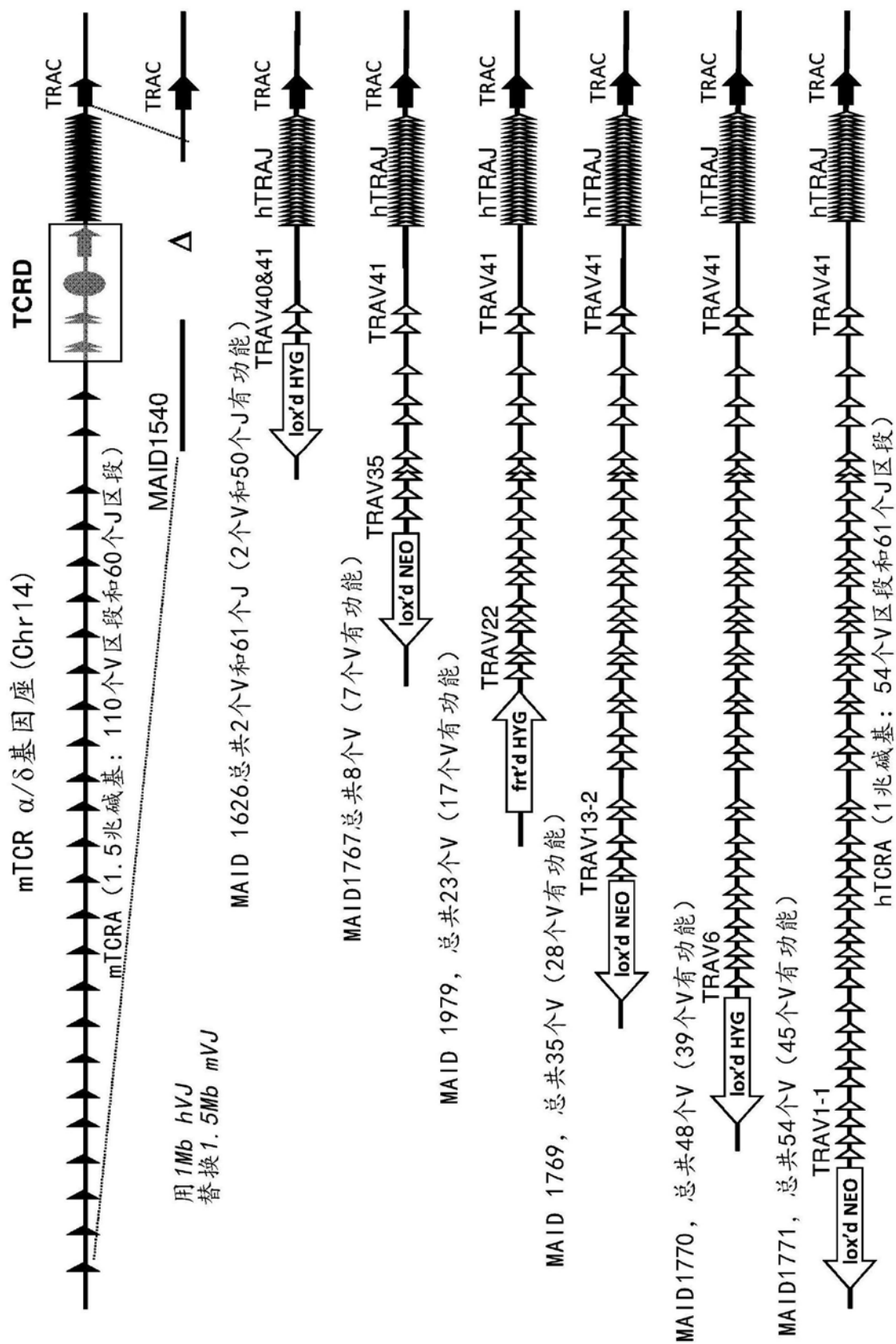


图4A



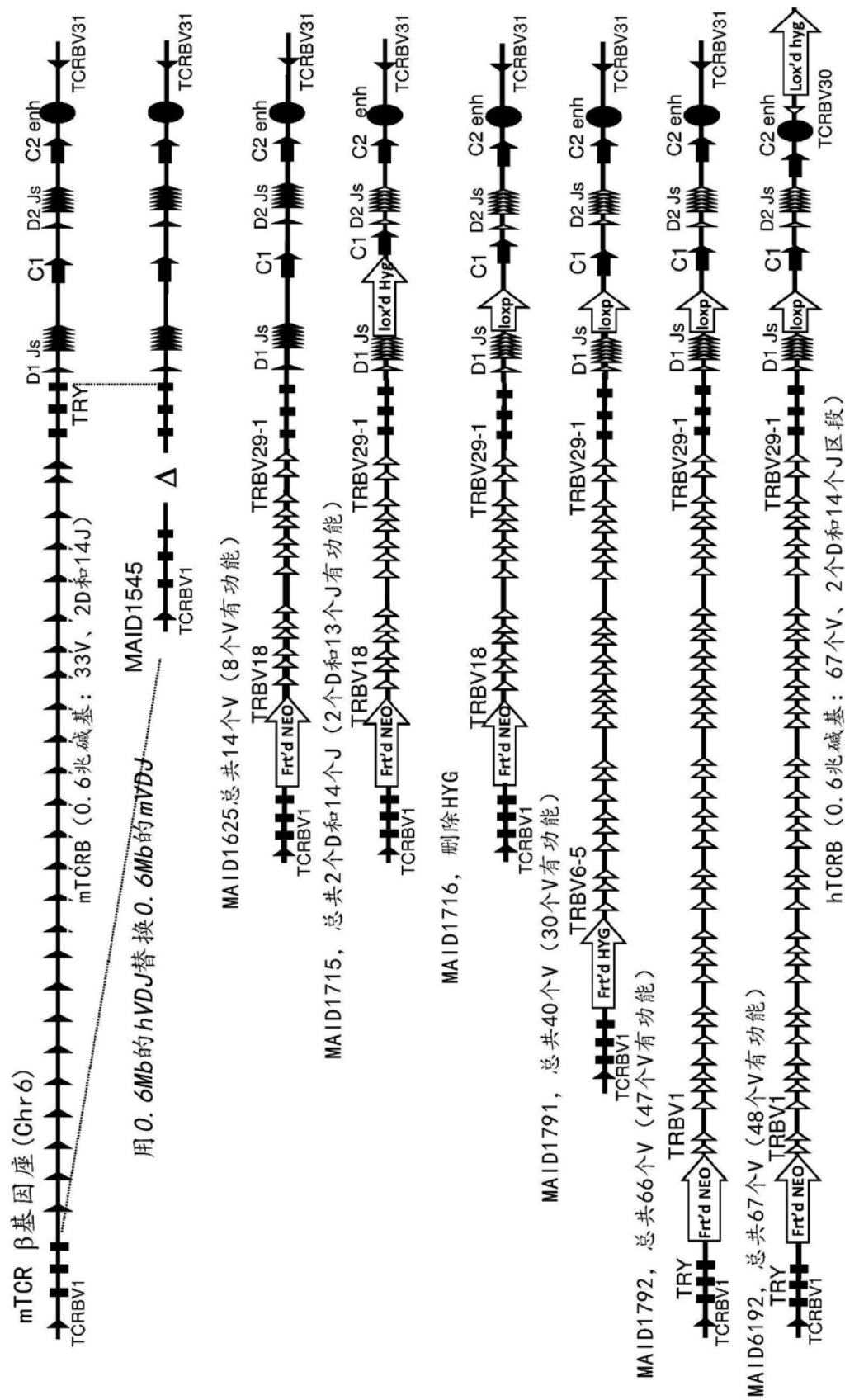


图4B

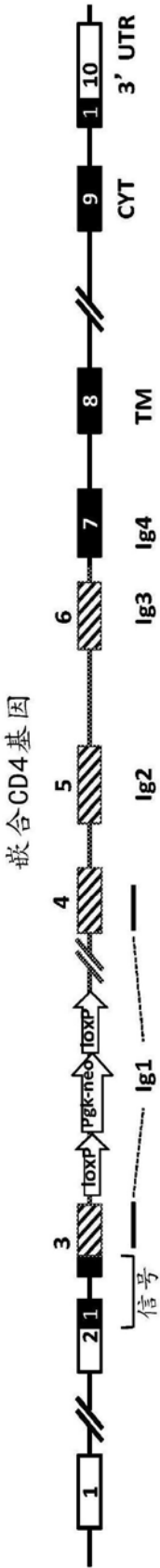


图5A

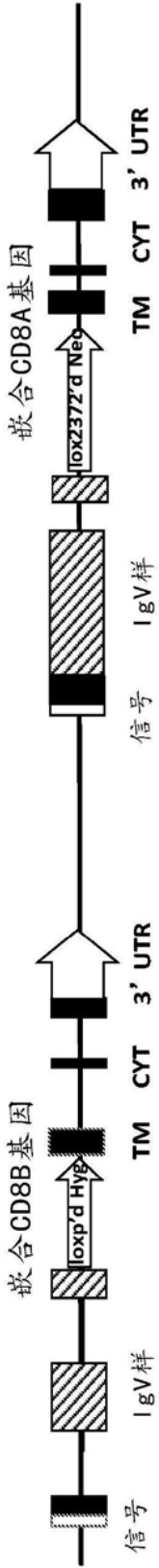


图5B

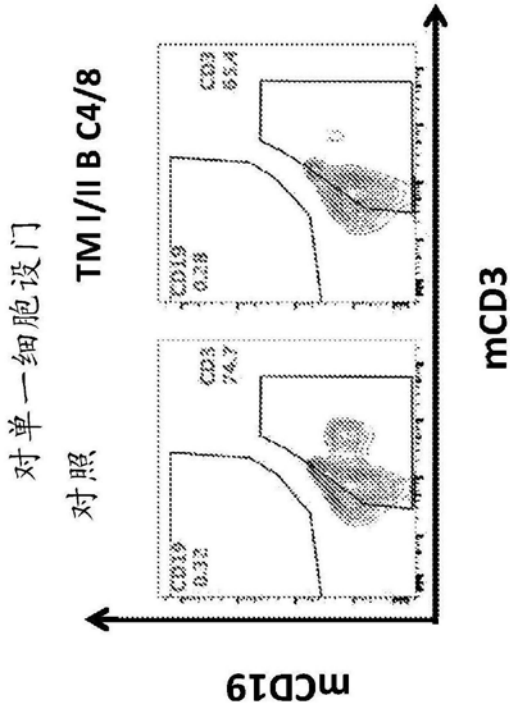


图6A

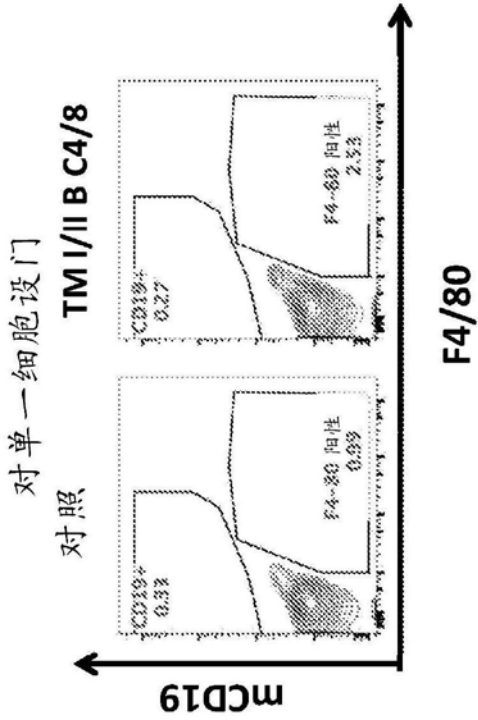


图6B

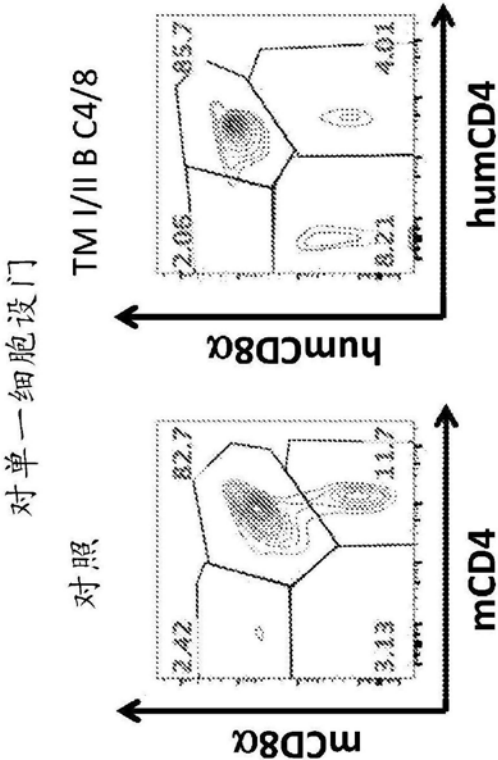


图6C

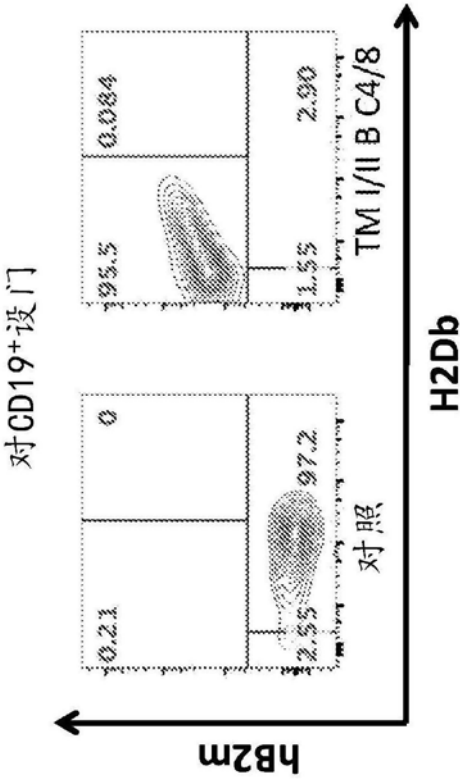


图7A

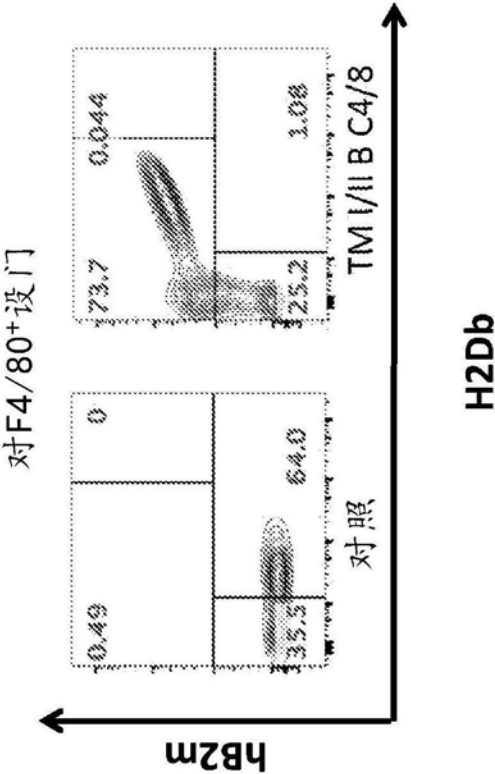


图7B

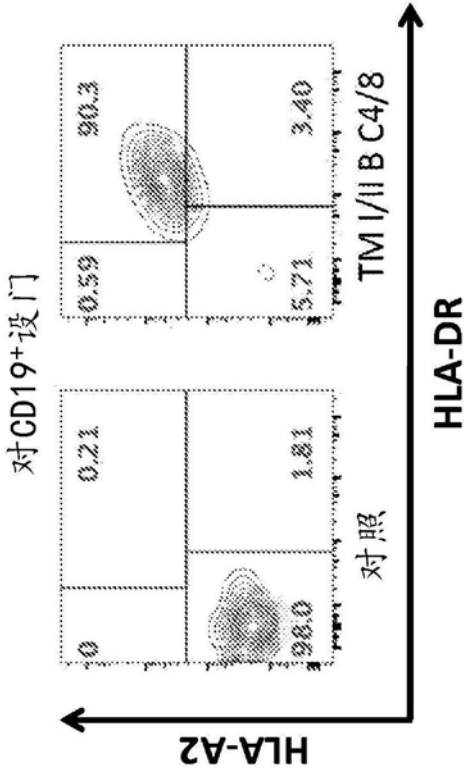


图7C

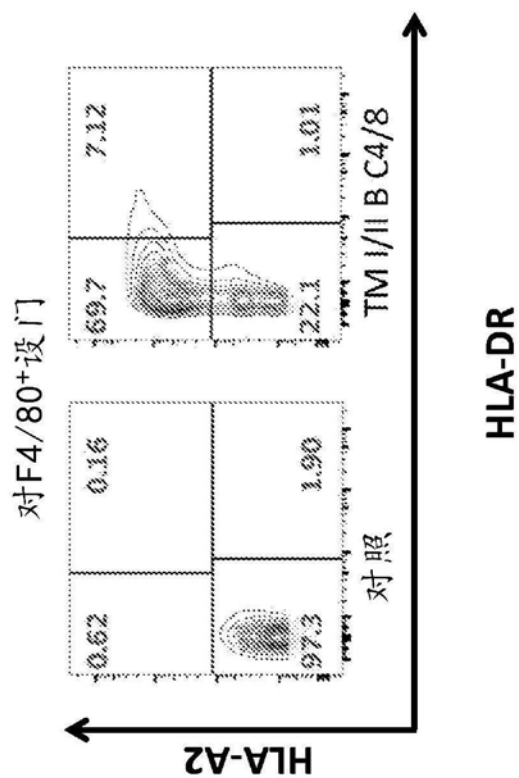


图7D

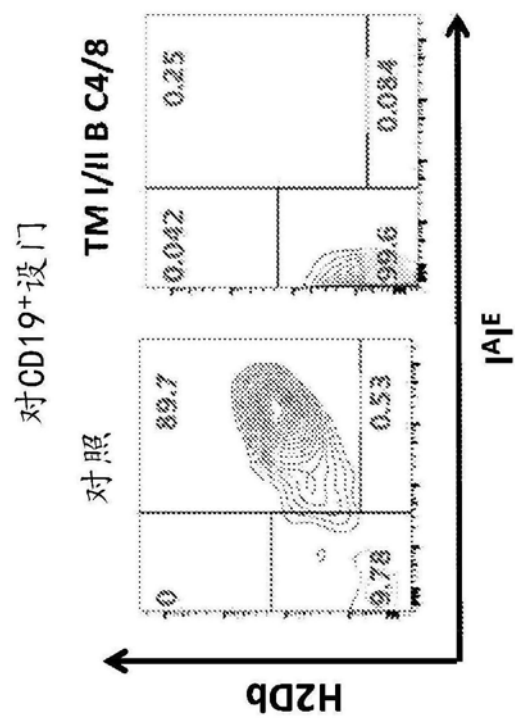


图7E

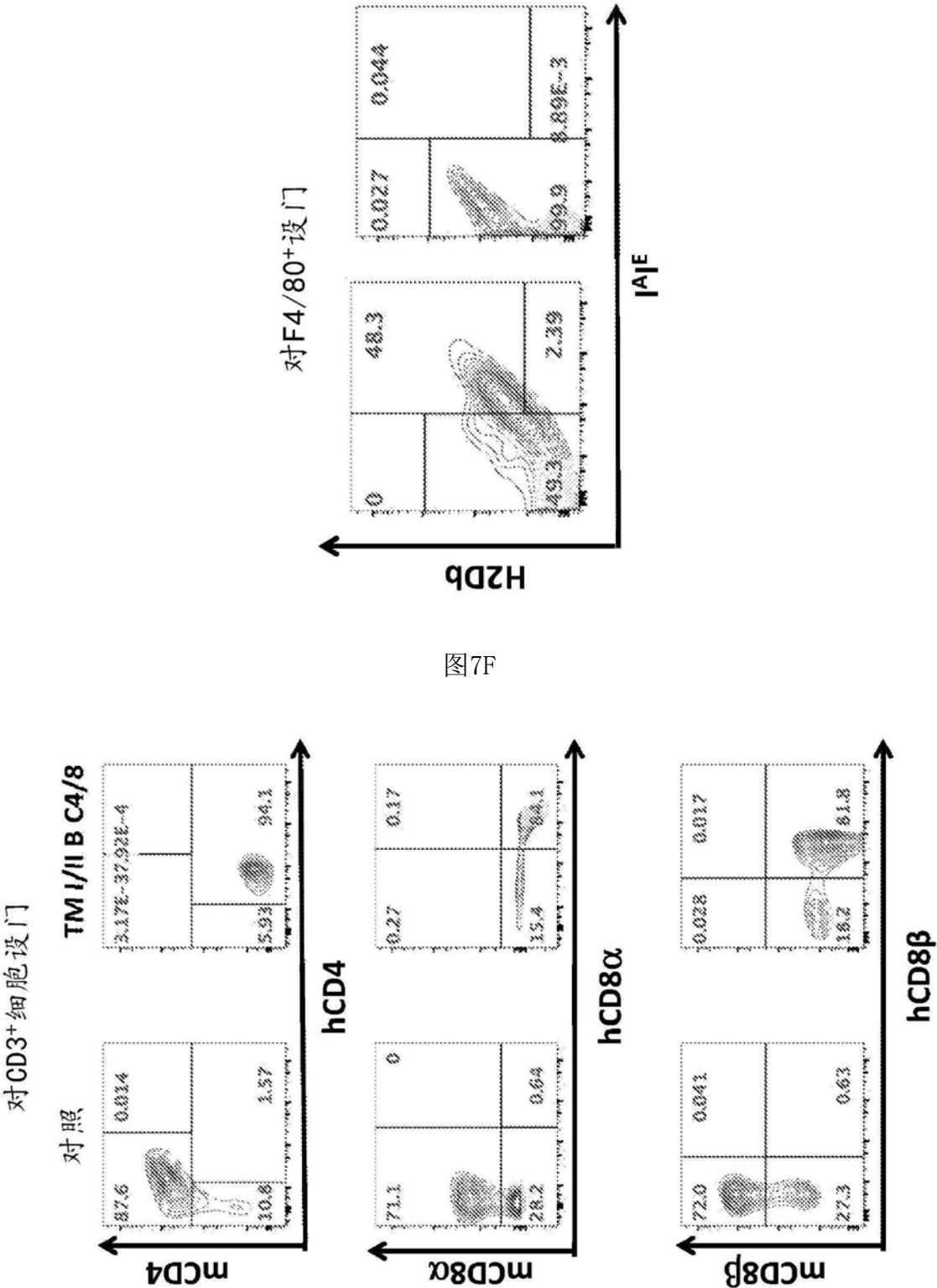


图7F

图7G



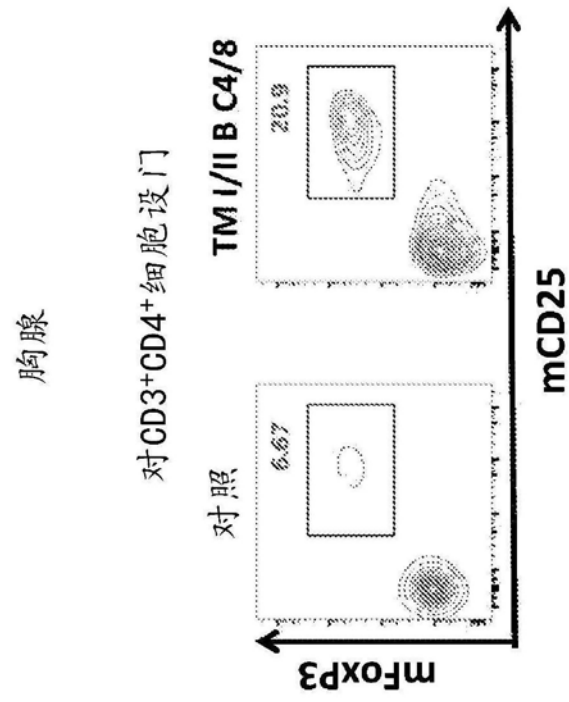


图8

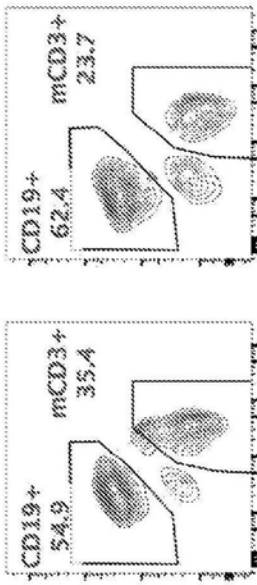
脾

对单一细胞设门

对照

TM I/II B C4/8

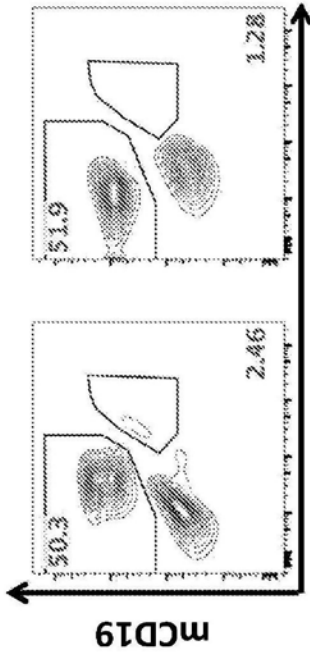
mCD19



mCD3

图9A

对单一细胞设门



F4/80

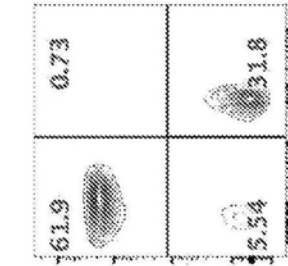
图9B

对CD3<sup>+</sup>细胞设门

对照

TM I/II B C4/8

mCD4



mCD8α

hCD4

hCD8α

图9C

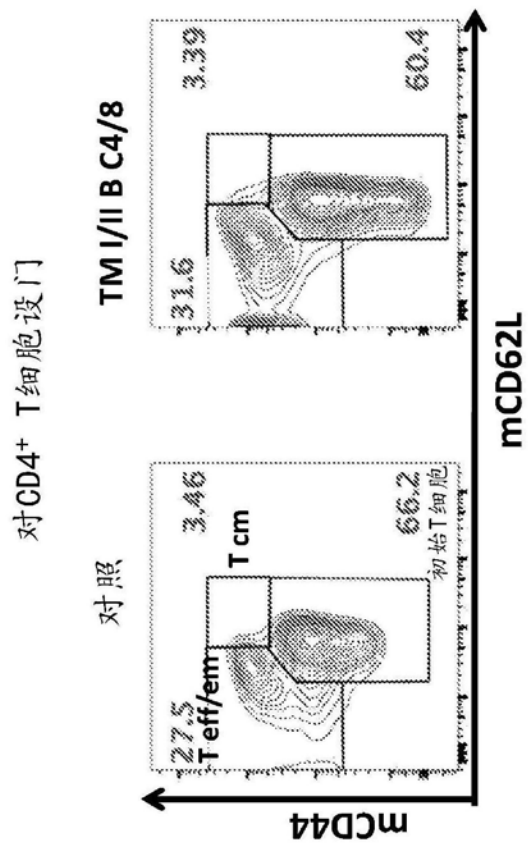


图9D

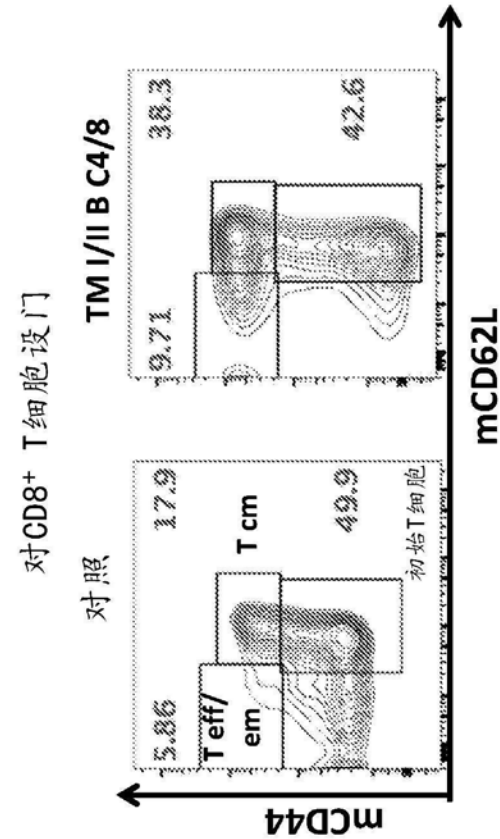


图9E

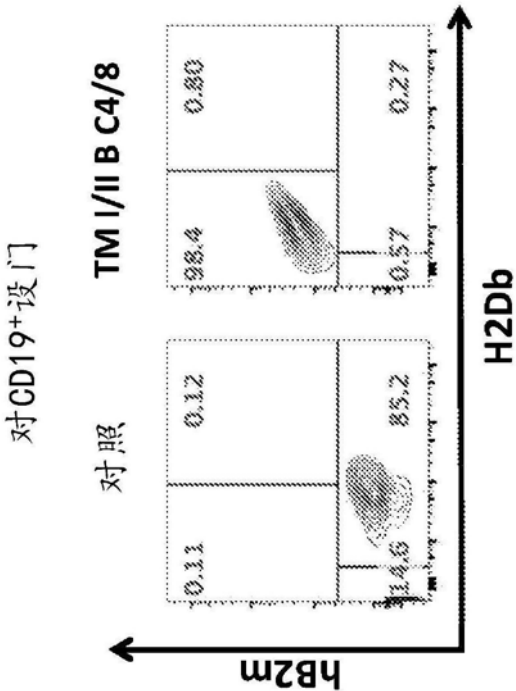


图10A

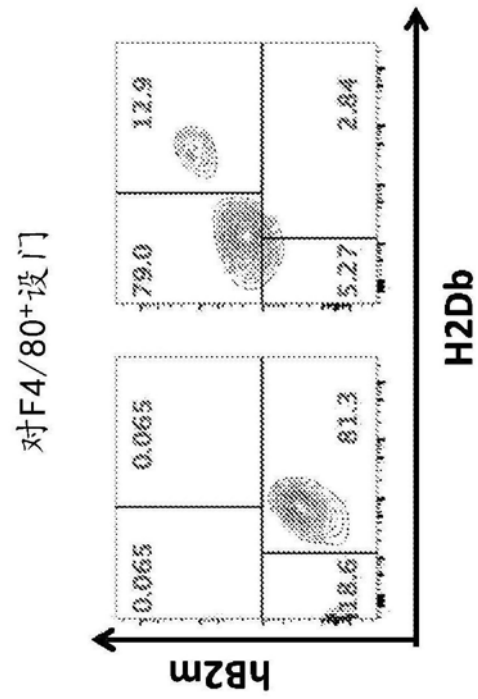


图10B

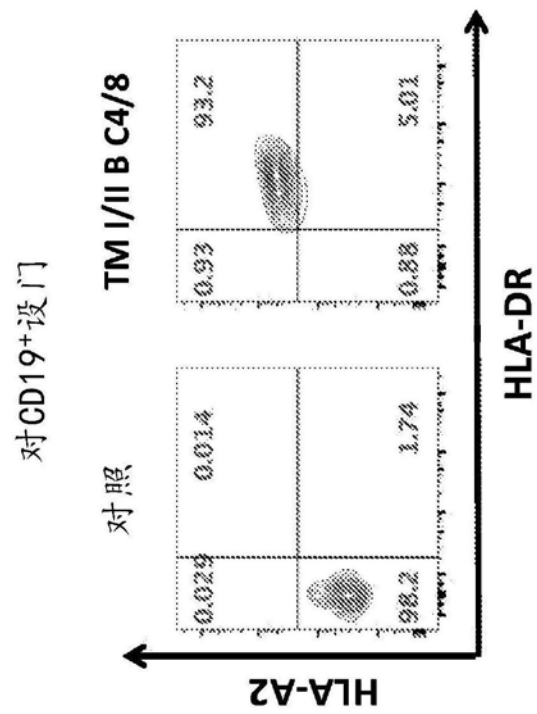


图10C

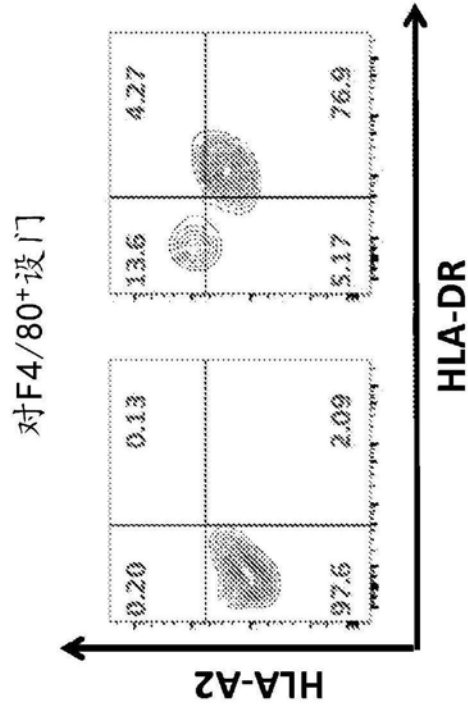


图10D

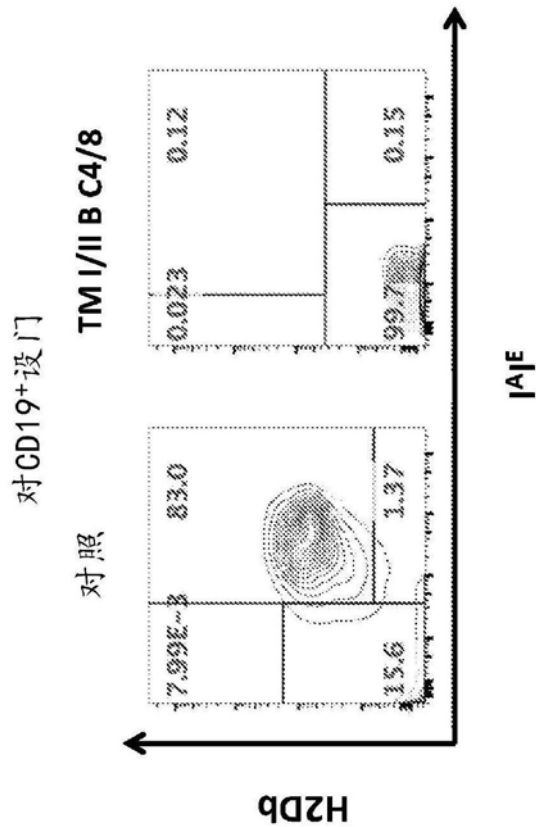


图10E

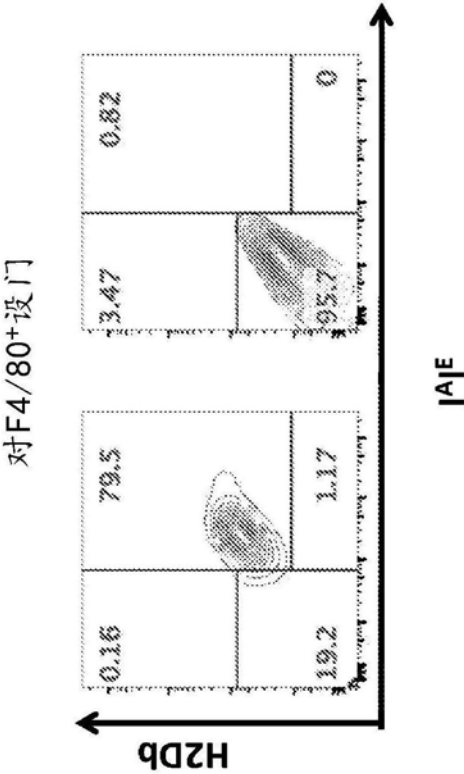


图10F

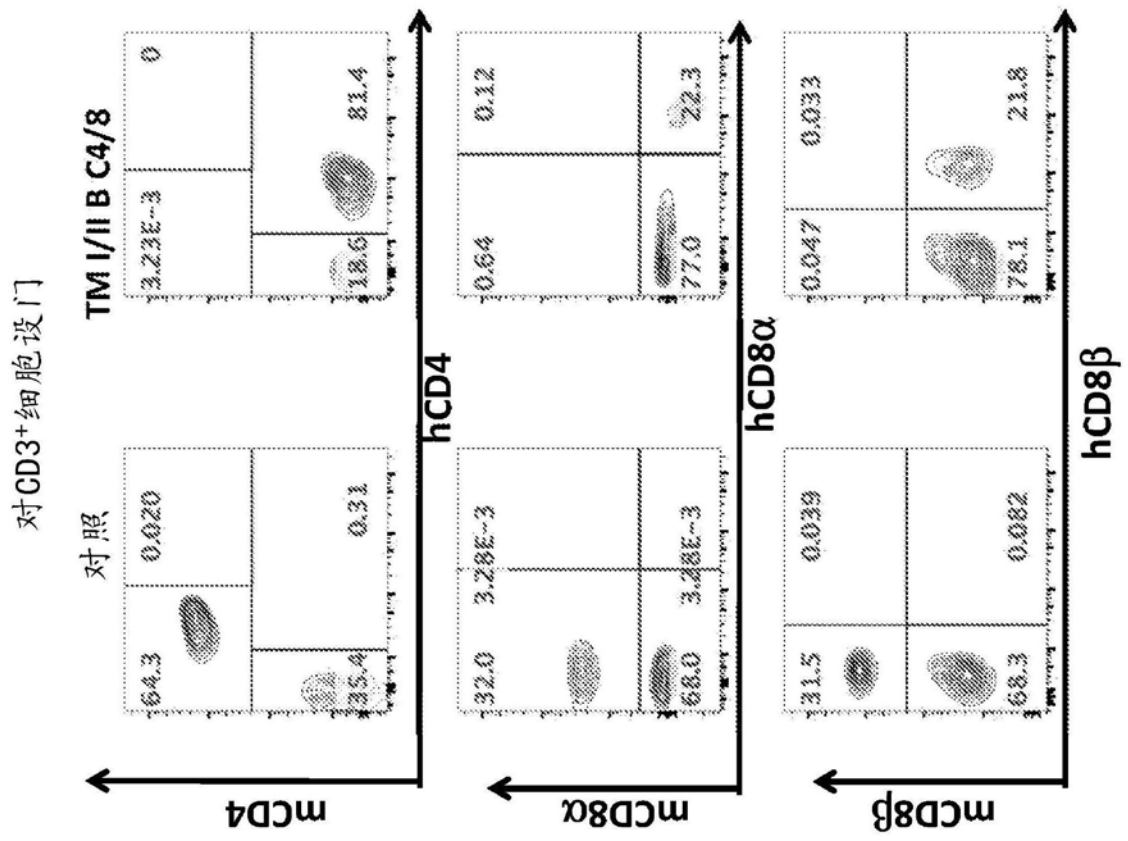


图10G



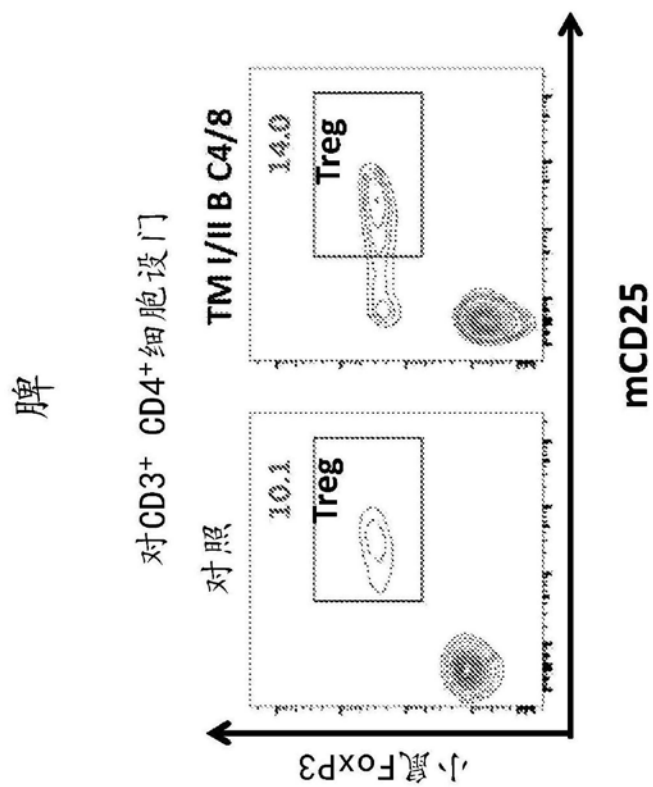


图11

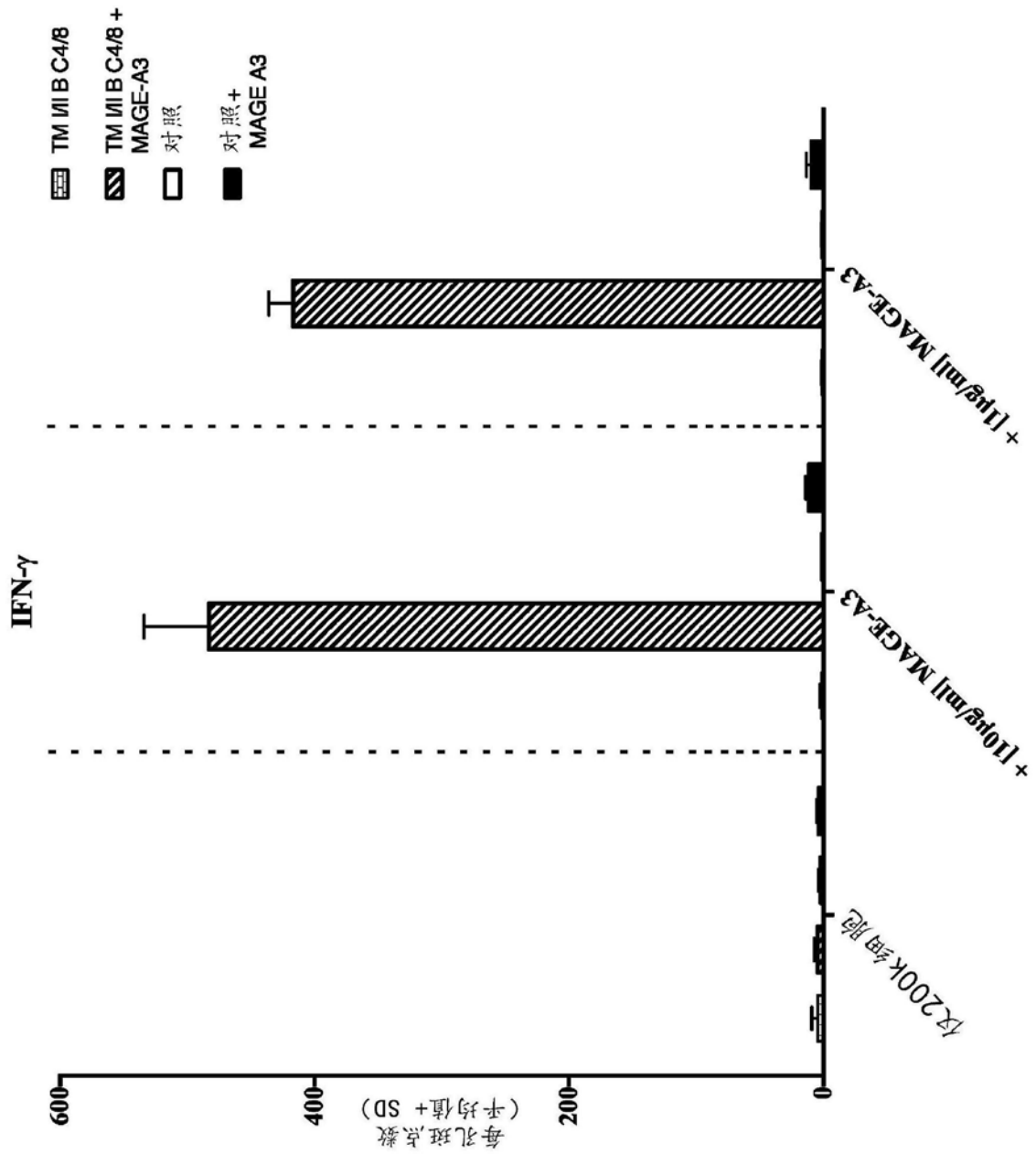


图12

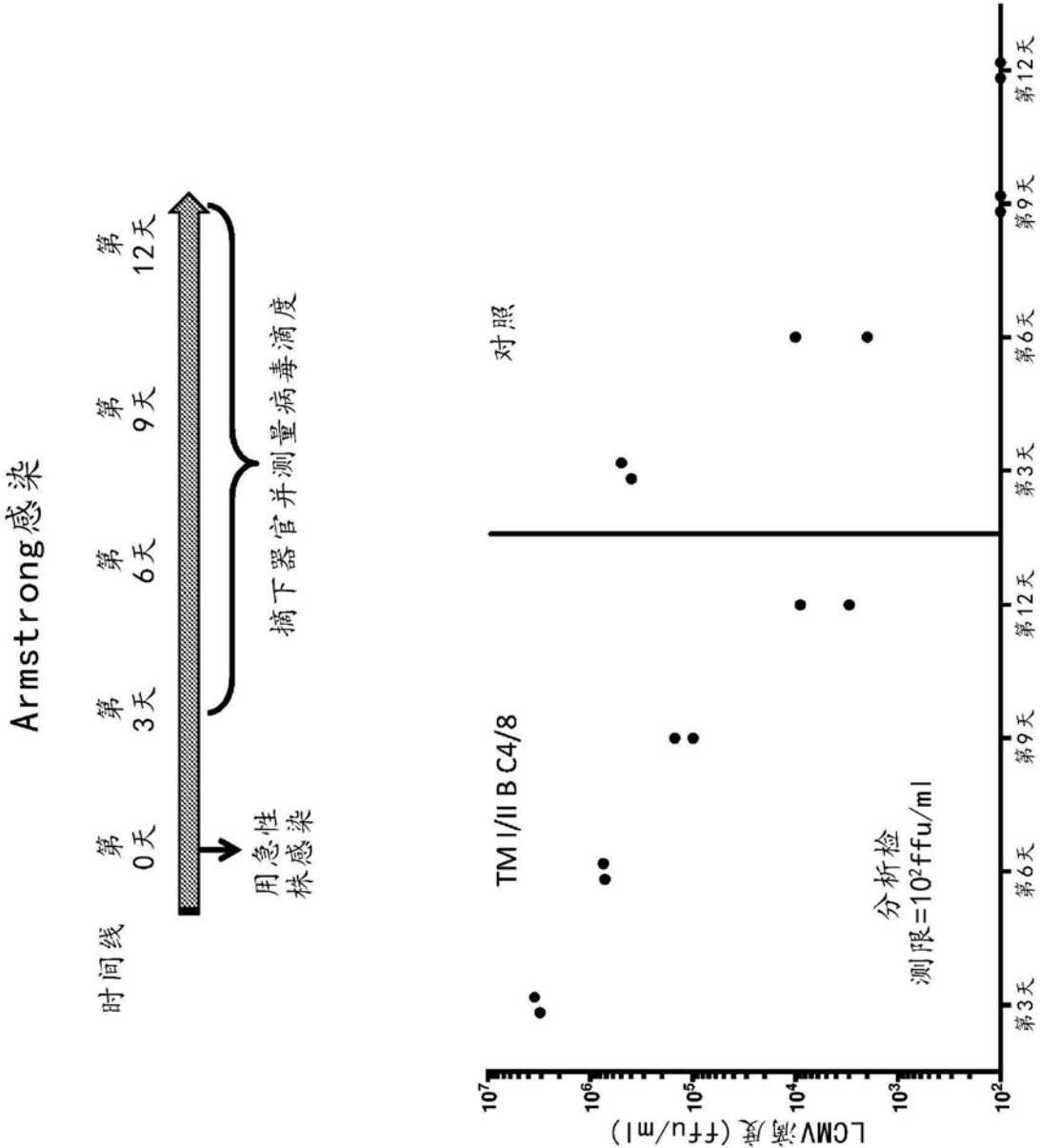


图13A

克隆13感染

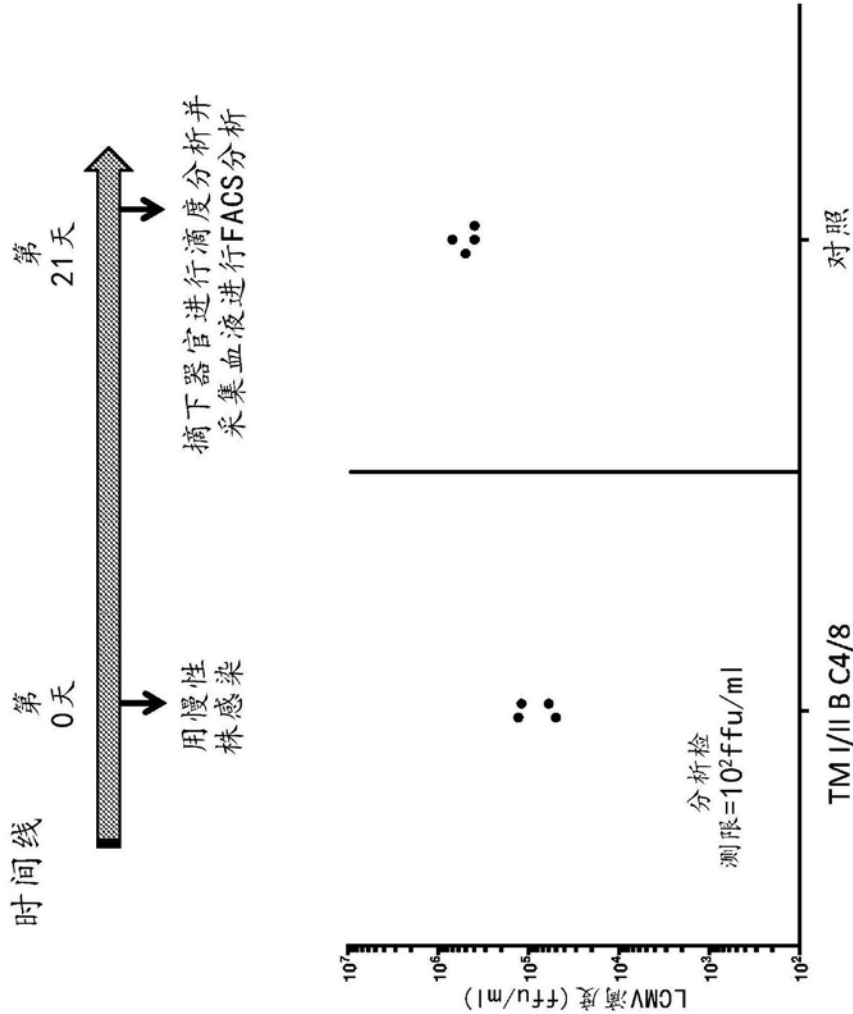


图13B

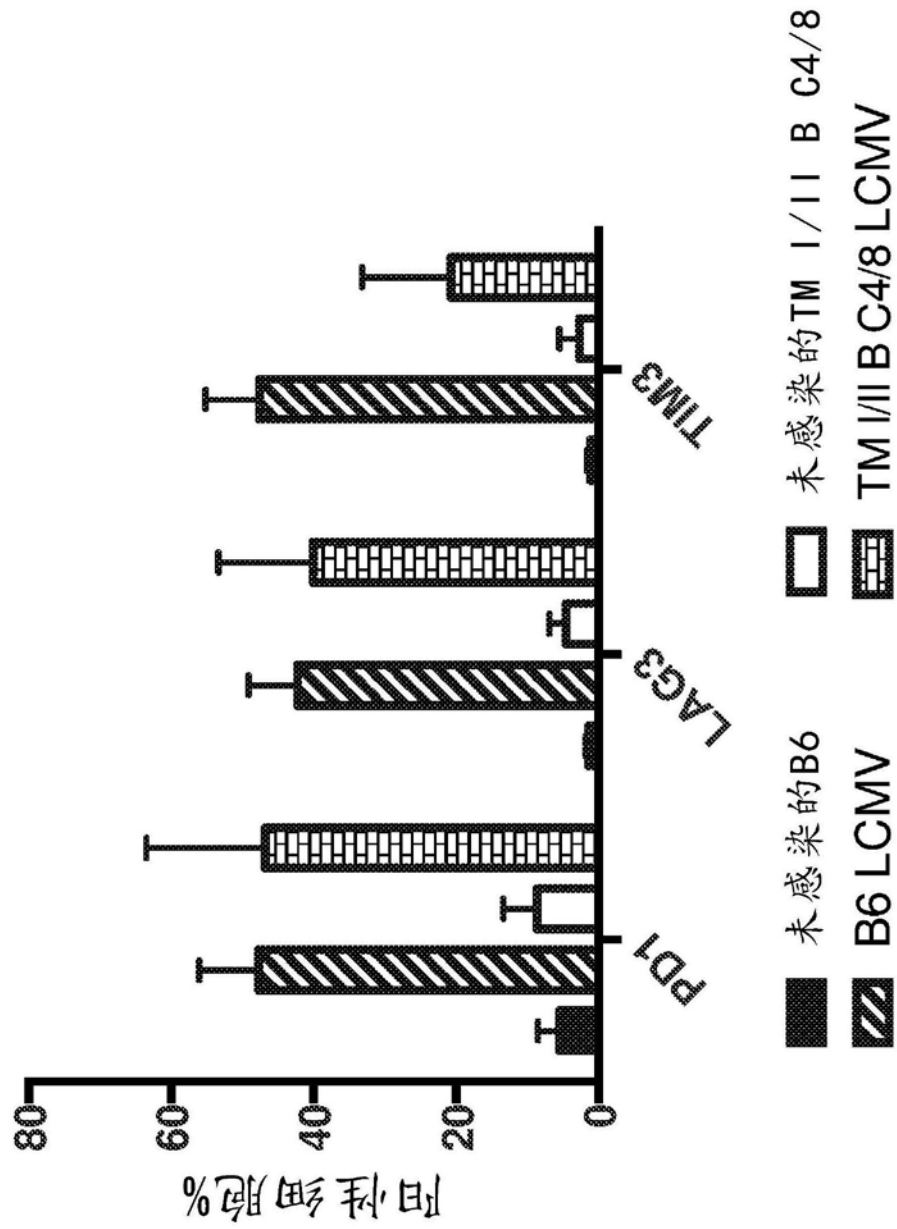


图13C

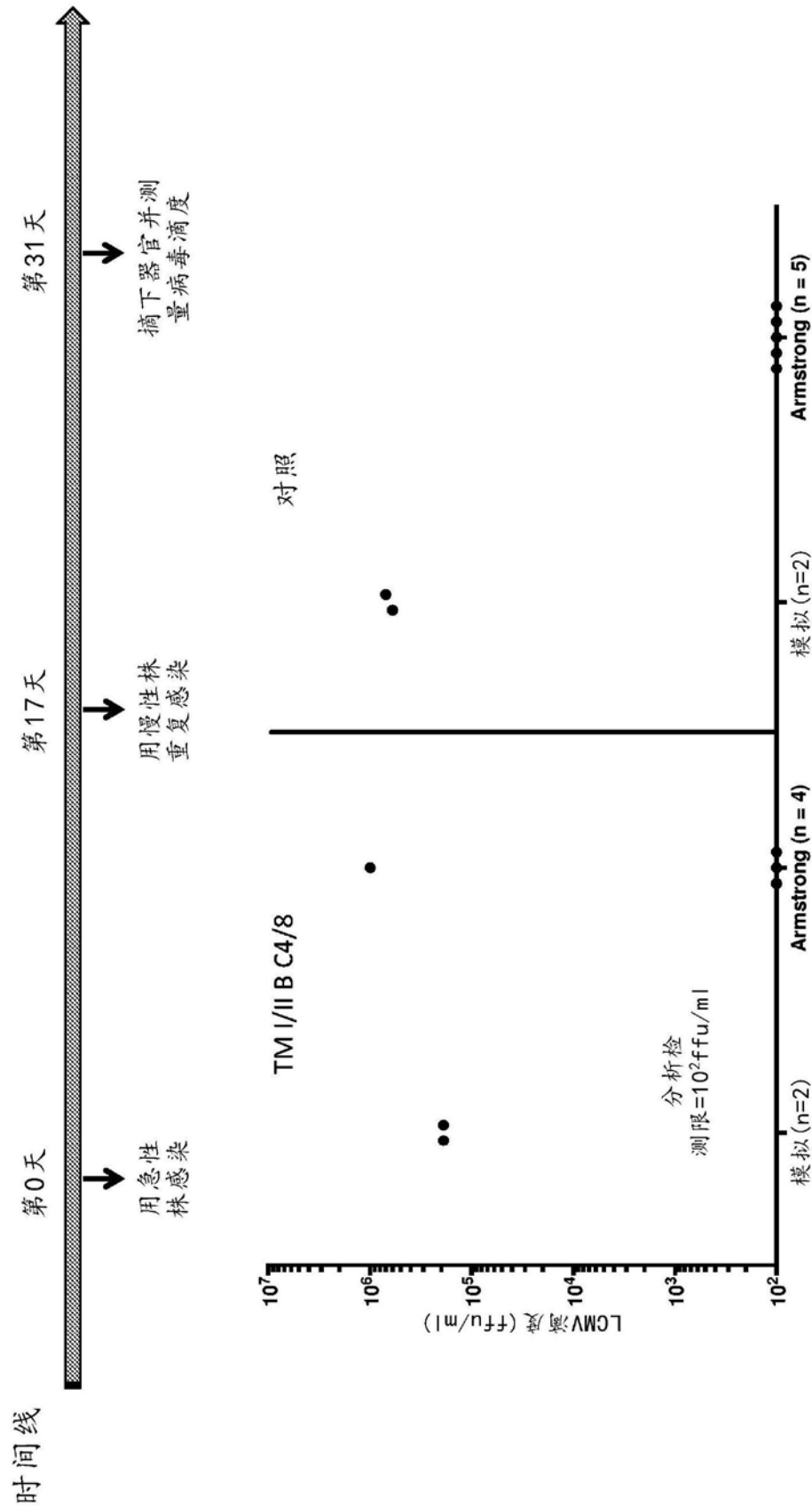


图14

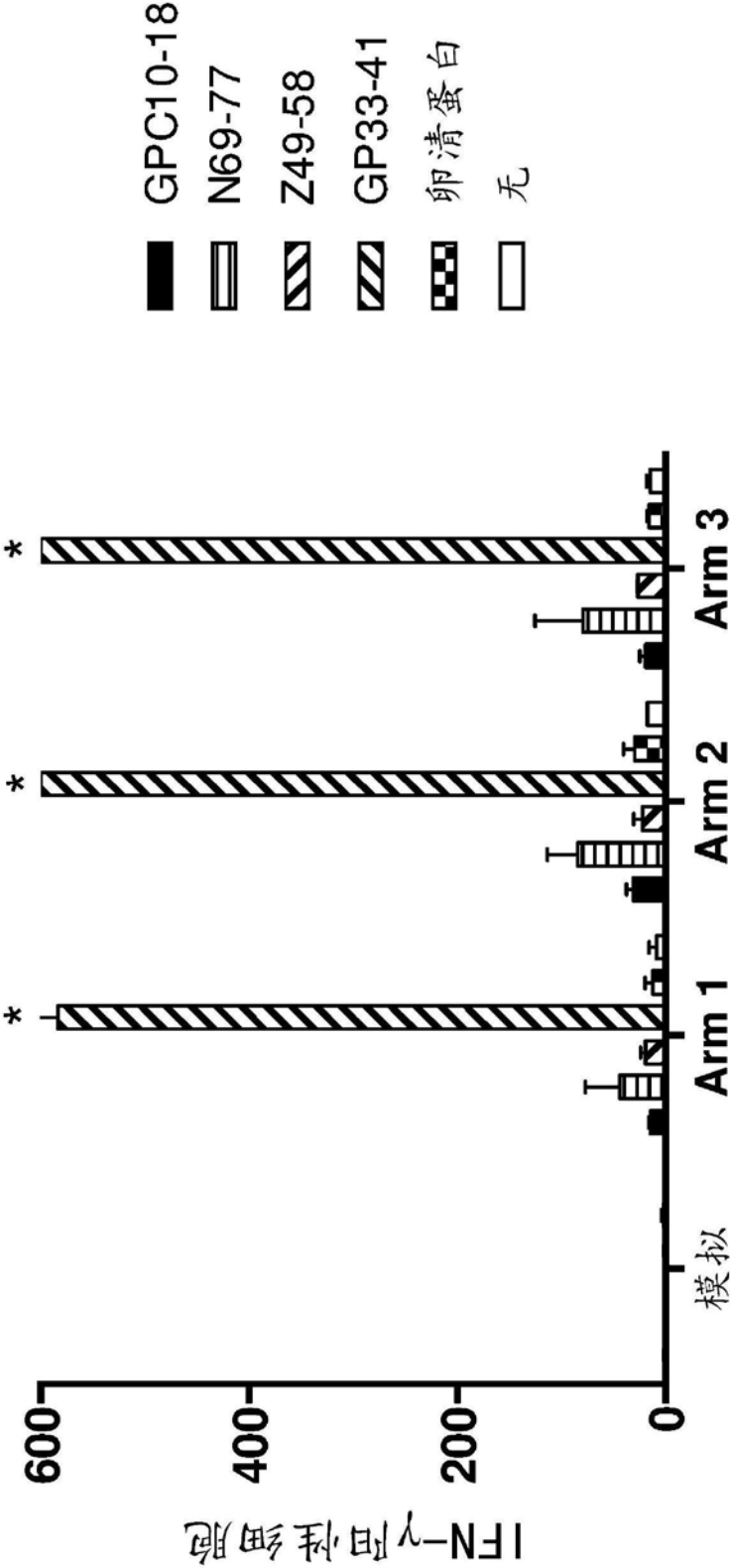


图15A

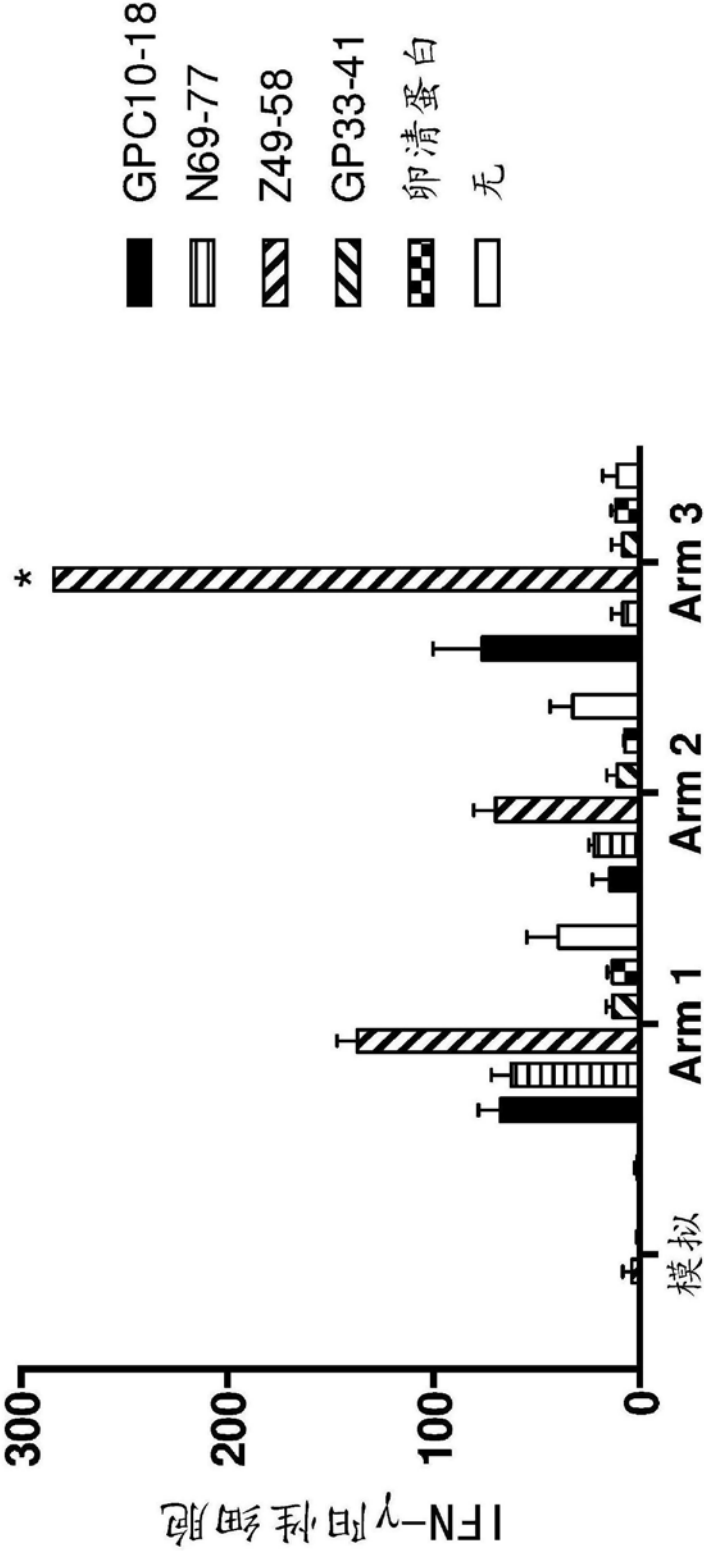


图15B



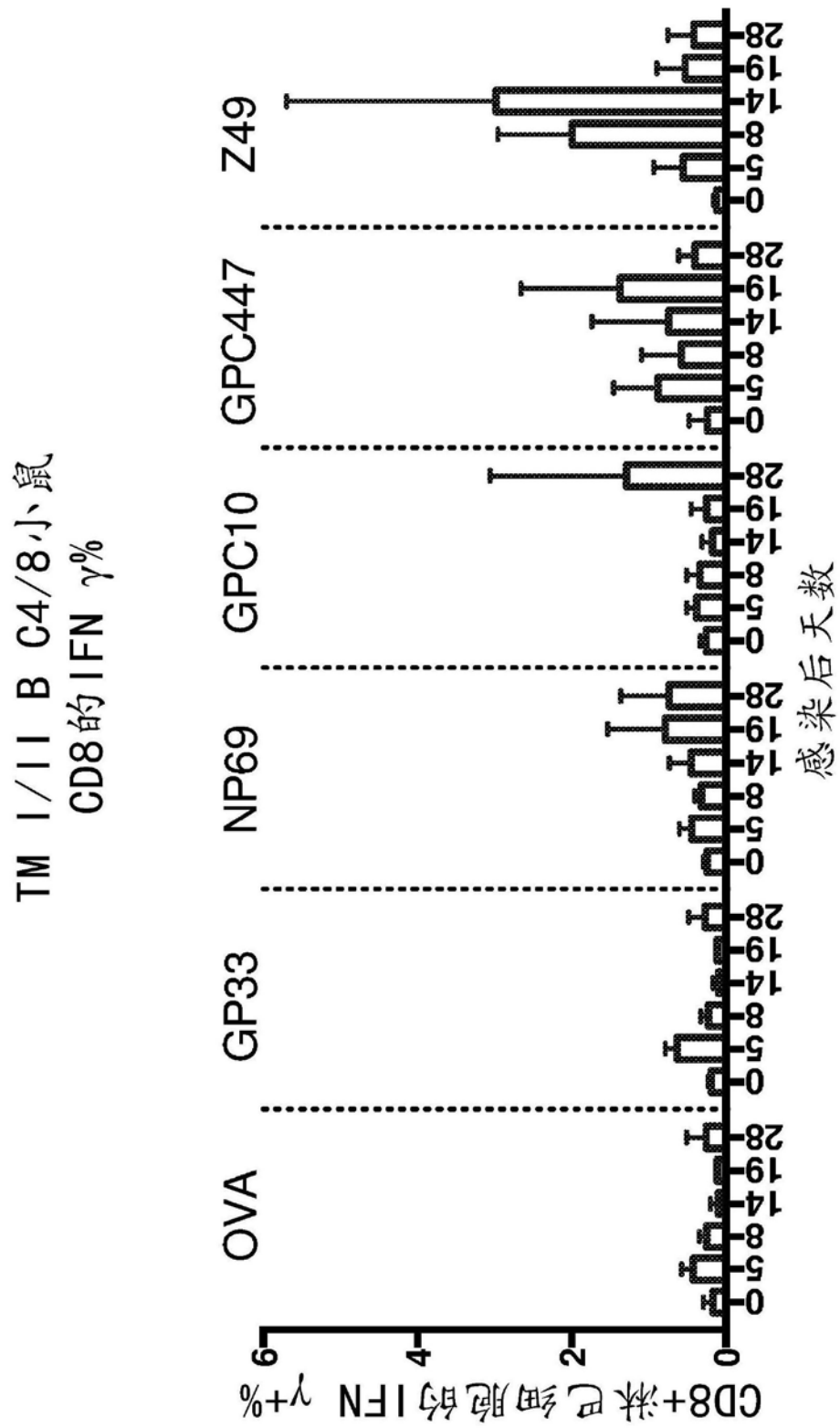


图15C

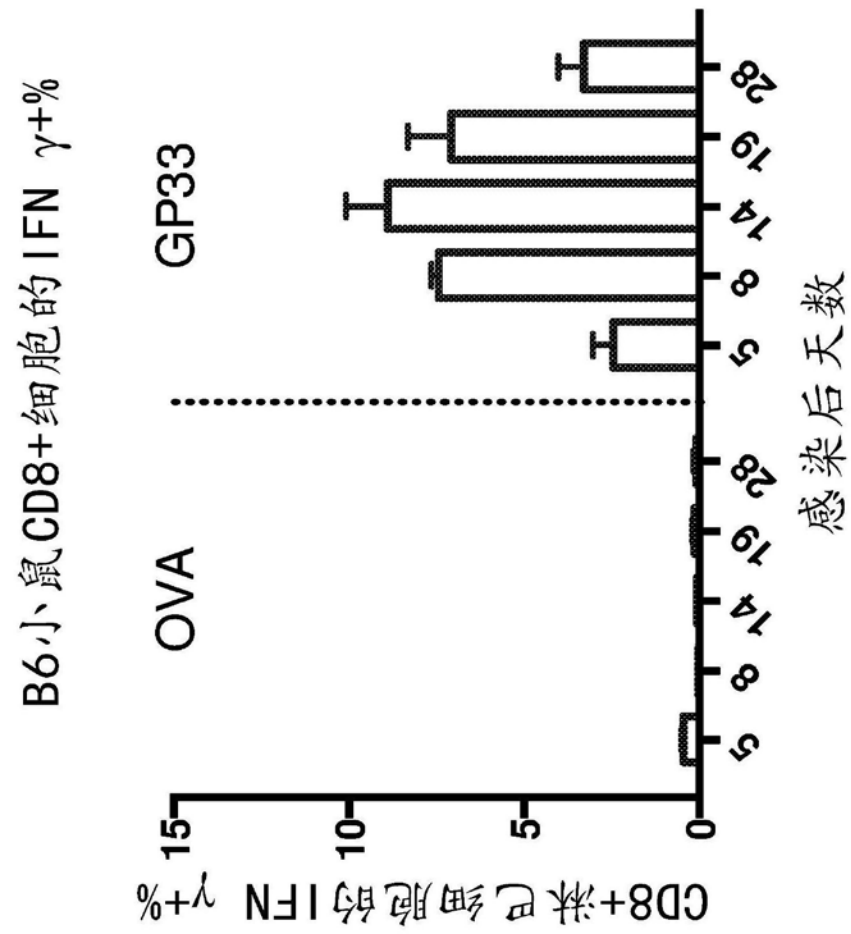


图15D