



(12) 发明专利申请

(10) 申请公布号 CN 114303060 A

(43) 申请公布日 2022.04.08

(21) 申请号 202080041952.1

H·舒斯特 C·瓦格纳

(22) 申请日 2020.06.04

S·优素福 A·阿尔珀特

(30) 优先权数据

102019129341.3 2019.10.30 DE

62/858,167 2019.06.06 US

(74) 专利代理机构 北京坤瑞律师事务所 11494

代理人 封新琴

(85) PCT国际申请进入国家阶段日

2021.12.06

(51) Int.Cl.

G01N 33/50 (2006.01)

A61K 35/17 (2006.01)

G01N 33/569 (2006.01)

G01N 15/14 (2006.01)

(86) PCT国际申请的申请数据

PCT/EP2020/065567 2020.06.04

(87) PCT国际申请的公布数据

W02020/245326 EN 2020.12.10

(71) 申请人 伊玛提克斯生物技术有限公司

地址 德国图宾根市

申请人 伊玛提克斯美国公司

权利要求书6页 说明书45页

序列表40页 附图11页

(72) 发明人 S·邦克 D·毛勒 G·席姆麦克

(54) 发明名称

使用序列相似肽进行反向选择分选

(57) 摘要

本发明涉及一种选择细胞或病毒的方法,所述细胞或病毒在其表面表达与目的蛋白抗原(PAI)特异性结合的抗原结合蛋白,同时应用使用相似蛋白抗原(SPA)的反向选择。此外,本发明提出了一种确定编码抗原结合蛋白或其抗原结合部分的核酸之序列的方法,以及一种产生表达编码抗原结合蛋白或其抗原结合部分的核酸之细胞的方法。本发明还涉及一种用选定细胞群治疗受试者的方法。

1. 一种选择细胞或病毒的方法,所述细胞或病毒在其表面表达与目的蛋白抗原 (PAI) 特异性和/或选择性结合的抗原结合蛋白,所述方法包括以下步骤:

(i) 提出细胞群或病毒群;

(ii) 使步骤 (i) 的细胞群或病毒群与包含PAI和可检测标记A的第一抗原复合体 (第1AC) 或包含可检测标记A的PAI接触;

(iii) 使步骤 (i) 的细胞群或病毒群与包含相似蛋白抗原 (SPA) 的至少第二抗原复合体 (第2AC) 接触,其中SPA的氨基酸序列与PAI的氨基酸序列至少有1个氨基酸不同,其中第2AC包含可检测标记B;或与SPA和可检测标记B接触;以及

(iv) 选择至少一种与第一AC特异性和/或选择性结合的细胞或病毒,其中,可检测标记A和可检测标记B彼此系可检测地不同。

2. 根据权利要求1所述的方法,其中选定的

(i) 细胞为免疫细胞,优选为T细胞,优选为CD4或CD8 T细胞;或B细胞;或表达异源抗原结合蛋白的哺乳动物或酵母细胞;或

(ii) 选定的病毒为噬菌体。

3. 根据权利要求1或2任一项所述的方法,其中所述抗原结合蛋白选自包含T细胞受体 (TCR) 或其抗原结合片段、B细胞受体 (BCR) 或其抗原结合片段以及嵌合抗原受体 (CAR) 或其抗原结合片段的组。

4. 根据权利要求1至3任一项所述的方法,其中

(a) 细胞群包括:

(i) 免疫细胞,优选为肿瘤浸润淋巴细胞 (TIL)、T细胞受体文库、健康受试者的外周血、患病受试者的外周血或其免疫细胞富集级分;或

(ii) 真核细胞,优选为表达异源抗原结合蛋白文库的哺乳动物细胞或酵母细胞;或

(b) 病毒群包含表达异源抗原结合蛋白文库的病毒。

5. 根据权利要求4所述的方法,其中所述免疫细胞富集级分富含干细胞、T细胞,优选为CD8 T细胞或CD4 T细胞、B细胞、浆细胞。

6. 根据权利要求1至5任一项所述的方法,其中,目的蛋白抗原 (PAI) 为肿瘤相关抗原 (TAA)、病毒蛋白或细菌蛋白。

7. 根据权利要求4至6所述的方法,其中,患病受试者患有选自由免疫疾病、肿瘤性疾病、病毒引起的疾病或细菌引起的疾病组成组的疾病。

8. 根据权利要求4至7任一项所述的方法,其中,透过可检测标记一个或多个免疫细胞特异性表面标志物来选择免疫细胞富集级分。

9. 根据权利要求1至8任一项所述的方法,其包含在生长因子和/或分化因子存在的情况下孵育细胞群的进一步步骤,优选为从细胞因子组成的组中选择。

10. 根据权利要求1至9任一项所述的方法,其中,所述AC为抗原提呈细胞,或包含颗粒、PAI和可检测标记A或SPA和可检测标记B的复合体。

11. 根据权利要求1至10任一项所述的方法,其包括一个或多个以下进一步的步骤

(a) 使步骤 (i) 的细胞群与包含PAI和可检测标记C的第三抗原复合体 (第3AC) 接触,所述可检测标记C系可检测地不同于与细胞群接触的其他AC的一个或多个或所有其他可检测标记,优选为可检测地不同于至少可检测标记A,优选为可检测地不同于至少可检测标记A

和可检测标记D(如果存在可检测标记D);和/或

(b) 使步骤(i)的细胞群与包含PAI和可检测标记D的第四抗原复合体(第4AC)接触,所述可检测标记D系可检测地不同于与细胞群接触的其他AC的一个或多个或所有其他可检测标记,优选为可检测地不同于至少可检测标记A,优选为可检测地不同于至少可检测标记A和可检测标记C;和/或

(c) 使步骤(i)的细胞群与包含SPA和可检测标记E的第五抗原复合体(第5AC)接触,所述可检测标记E系可检测地不同于与细胞群接触的其他AC的一个或多个或所有其他可检测标记,优选为可检测地不同于至少可检测标记B,优选为可检测地不同于至少可检测标记B和可检测标记F(如果存在可检测标记F);和/或

(d) 使步骤(i)的细胞群与包含SPA和可检测标记F的第六抗原复合体(第6AC)接触,所述可检测标记F系可检测地不同于与细胞群接触的其他AC的一个或多个或所有其他可检测标记,优选为可检测地不同于至少可检测标记B,优选为可检测地不同于至少可检测标记B和可检测标记E;和/或

(e) 使步骤(i)的细胞群与一种或多种另外抗原复合体(AC)接触,其中每个AC均包含SPA,该SPA之至少一个氨基酸序列与第2AC的SPA氨基酸序列不同,

并且其中每个另外AC包含一个或更多标记,其中,一个或更多标记与第2AC的一个或更多标记相同或可检测地不同。

12. 根据权利要求1至11任一项所述的方法,其中

(i) 第1AC包含至少一种另外的可检测标记,第2AC包含至少一种另外的可检测标记,两者相同或不同;和/或

(ii) 所述一个或更多另外AC包含至少一个另外的可检测标记;其中选择所述至少一个另外的标记,以便可区分第1AC和第2AC以及一个或多个另外AC。

13. 根据权利要求1至11任一项所述的方法,其中,可检测标记独立选自萤光标记,优选为选自由黄嘌呤类、吡啶类、恶嗪类、菁类、苯乙烯基染料、香豆素类、卟啉类、金属-配体络合物类、萤光蛋白质类、纳米晶体、花类和酞菁类组成的组。

14. 根据权利要求1至11任一项所述的方法,其中第1AC为MHC-I或MHC-II与PAI的复合体,其中PAI为靶肽(TP),优选为肿瘤特异性靶肽和/或第2AC为MHC-I或MHC-II与SPA的复合体,其中SPA为靶相似肽(TSP),其中TSP与TP的氨基酸序列至少有1个氨基酸不同。

15. 根据权利要求14所述的方法,其中透过以下一项或多项标准选择所述至少一个TSP的氨基酸序列:

(a) TSP提呈于健康组织;

(b) 衍生自HLA类型来源;和

(c) 与各自HLA结合。

16. 根据权利要求14或15所述的方法,其中,TSP的氨基酸序列长度为8至16个氨基酸,其中:

(1) TSP的氨基酸序列的下列位置与TP的氨基酸序列不同

$X_1-X_2-X_3-X_4-X_5-X_6-X_7-X_8$

(i) 位置 X_1 、 X_2 和 X_3 ,其中位置 X_4 至 X_8 与TP相同或相似;

(ii) 位置 X_4 、 X_5 和 X_6 ,其中位置 X_1 至 X_3 以及位置 X_7 和 X_8 与TP相同或相似;或

(iii) 位置 X_7 和 X_8 ,其中位置 X_1 至 X_6 與TP相同或相似;或
如果TP的長度為8個氨基酸;或

(2) TSP的氨基酸序列的下列位置與TP的氨基酸序列不同

$X_1-X_2-X_3-X_4-X_5-X_6-X_7-X_8-X_9$

(i) 位置 X_1 、 X_2 和 X_3 ,其中位置 X_4 至 X_9 與TP相同或相似;

(ii) 位置 X_4 、 X_5 和 X_6 ,其中位置 X_1 至 X_3 以及位置 X_7 至 X_9 與TP相同或相似;或

(iii) 位置 X_4 、 X_5 、 X_6 和 X_7 ,並且其中位置 X_1 至 X_3 以及位置 X_8 至 X_9 與TP相同或相似;或

(iv) 位置 X_7 、 X_8 和 X_9 ,其中位置 X_1 至 X_6 與TP相同或相似;或

如果TP的長度為8-9個氨基酸;或

(3) TSP的氨基酸序列的下列位置与TP的氨基酸序列不同

$X_1-X_2-X_3-X_4-X_5-X_6-X_7-X_8-X_9-X_{10}$

(i) 位置 X_1 、 X_2 和 X_3 ,其中位置 X_4 至 X_{10} 与TP相同或相似;

(ii) 位置 X_4 、 X_5 、 X_6 和 X_7 ,其中位置 X_1 至 X_3 以及位置 X_8 至 X_{10} 与TP相同或相似;或

(iii) 位置 X_4 、 X_5 和 X_6 ,其中位置 X_1 至 X_3 以及位置 X_7 至 X_{10} 与TP相同或相似;或

(iv) 位置 X_8 、 X_9 和 X_{10} ,其中位置 X_1 至 X_7 与TP相同或相似;或

如果TP的长度为8-10个氨基酸;或

(4) TSP的氨基酸序列的下列位置与TP的氨基酸序列不同

$X_1-X_2-X_3-X_4-X_5-X_6-X_7-X_8-X_9-X_{10}-X_{11}$

(i) 位置 X_1 、 X_2 和 X_3 ,其中位置 X_4 至 X_{11} 与TP相同或相似;

(ii) 位置 X_4 、 X_5 、 X_6 和 X_7 ,其中位置 X_1 至 X_3 以及位置 X_8 至 X_{11} 与TP相同或相似;或

(iii) 位置 X_4 、 X_5 和 X_6 ,其中位置 X_1 至 X_3 以及位置 X_7 至 X_{11} 与TP相同或相似;或

(iv) 位置 X_8 、 X_9 、 X_{10} 和 X_{11} ,其中位置 X_1 至 X_7 与TP相同或相似;或

(v) 位置 X_9 、 X_{10} 和 X_{11} ,其中位置 X_1 至 X_8 与TP相同或相似;

如果TP的长度为8-11个氨基酸;或

(5) TSP的氨基酸序列的下列位置与TP的氨基酸序列不同

$X_1-X_2-X_3-X_4-X_5-X_6-X_7-X_8-X_9-X_{10}-X_{11}-X_{12}$

(i) 位置 X_1 、 X_2 和 X_3 ,其中位置 X_4 至 X_{12} 与TP相同或相似;

(ii) 位置 X_4 、 X_5 、 X_6 和 X_7 ,其中位置 X_1 至 X_3 以及位置 X_8 至 X_{12} 与TP相同或相似;或

(iii) 位置 X_4 、 X_5 和 X_6 ,其中位置 X_1 至 X_3 以及位置 X_7 至 X_{12} 与TP相同或相似;或

(iv) 位置 X_8 、 X_9 、 X_{10} 、 X_{11} 和 X_{12} ,其中位置 X_1 至 X_7 与TP相同或相似;或

(v) 位置 X_9 、 X_{10} 、 X_{11} 和 X_{12} ,其中位置 X_1 至 X_8 与TP相同或相似;

如果TP的长度为8-12个氨基酸。

17. 根据权利要求16所述的方法,其中,TSP的氨基酸序列长度与TP不同。

18. 根据权利要求1至17任一项所述的方法,其中,SPA或TSP的氨基酸序列与PAI或TP具有小于95%的氨基酸同一性、小于90%的氨基酸同一性或小于85%的氨基酸同一性。

19. 根据权利要求1至18任一项所述的方法,其中,使步骤(i)的细胞群与各包含不同相似蛋白抗原(SPA)的不多于10个抗原复合体(AC)接触,使用不多于9个不同SPA、不多于8个不同SPA、不多于7个不同SPA、不多于6个不同SPA、不多于5个不同SPA、不多于4个不同SPA、不多于3个不同SPA、不多于2个不同SPA或不多于1个SPA。

20. 根据权利要求19所述的方法,其中所述SPA为TSP。

21. 根据权利要求20所述的方法,其中,不同TSP的数量系介于1-10个、介于2-8个、介于3-5个或介于1-3个,优选为使用3个TSP。

22. 根据权利要求1至21任一项所述的方法,其中,权利要求1的步骤(ii)和(iii)和/或权利要求10的步骤(a)、(b)、(c)和/或(d)以任何顺序或组合连续或同时进行。

23. 根据权利要求1至22任一项所述的方法,其中步骤(iv)包括:

a) 正向选择(选择)结合于第1AC、第1和第3AC或第1、第3和第4AC的细胞;和/或

b) 负向选择(排除)结合于第2AC、第2和第5AC或第2、第5和第6AC的细胞。

24. 根据权利要求1至22任一项所述的方法,其中所述标记透过流式细胞术分析检测,优选为MACS分析、FACS分析或制备型分选分析。

25. 根据权利要求1至24任一项所述的方法,其中,步骤i)细胞群中所含的细胞为T细胞和/或B细胞且经表型分析。

26. 根据权利要求25所述的方法,其中,T细胞的表型分析包括确定一个或多个T细胞标志物。

27. 根据权利要求1至25任一项所述的方法,其中,B细胞的表型分析包括确定一个或多个B细胞标志物。

28. 根据权利要求1至27任一项所述的方法,其进一步包括以下步骤:

使步骤(i)的细胞群与包含无关蛋白抗原(IPA)的无关抗原复合体(IAC)接触,其中当IPA的氨基酸序列与PAI的氨基酸序列比对时,IPA的氨基酸序列的两个或两个以下氨基酸位置与PAI相同,其中IAC包含可检测地不同于可检测标记A的可检测标记G。

29. 根据权利要求28所述的方法,其中透过以下一项或多项标准选择至少一个IPA的氨基酸序列:

(a) IPA提呈于健康组织;

(b) 衍生自HLA类型来源;

(c) 与各自HLA结合。

30. 根据权利要求28至29所述的方法,其中,IPA或IPA中所含的至少一种蛋白质或肽的氨基酸序列与PAI具有小于30%的氨基酸同一性、小于20%的氨基酸同一性或小于10%的氨基酸同一性。

31. 根据权利要求28至30所述的方法,其中,所述IAC为MHC-I或MHC-II与无关肽(IP)的复合物,其中当与TP的氨基酸序列比对时,IP的氨基酸序列的一个氨基酸位置或无氨基酸位置与TP相同。

32. 一种确定编码抗原结合蛋白或其抗原结合部分的核酸序列的方法,所述方法包括以下步骤:

(i) 分离来自根据权利要求1至26选定细胞的编码抗原结合蛋白或其抗原结合部分的核酸;和

(ii) 确定核酸的序列。

33. 一种产生表达编码抗原结合蛋白或其抗原结合部分的核酸的细胞的方法,所述方法包括以下步骤:

(i) 提供来自本发明第一方面的方法中选定细胞的编码抗原结合蛋白或其抗原结合部

分的核酸序列；

(ii) 任选地在表达控制元素控制下产生包含步骤(i)中所提供核酸序列的核酸载体；和

(iii) 将步骤(ii)的核酸载体引入宿主细胞。

34. 根据权利要求32或33所述的方法，其中，抗原结合蛋白为TCR或其片段、BCR或其片段或抗体或其片段。

35. 根据权利要求34所述的方法，其中，TCR、BCR或抗体的氨基酸序列包含六个CDR序列。

36. 一种治疗有需要的受试者的方法，其包括以下步骤：

(i) 提出受试者的细胞群；

(ii) 使步骤(i)的细胞群与包含PAI和可检测标记A的第一抗原复合体(第1AC)或包含可检测标记A的PAI接触；

(iii) 使步骤(i)的细胞群与包含SPA的至少第二抗原复合体(第2AC)接触，其中SPA的氨基酸序列与PAI的氨基酸序列至少有1个氨基酸不同，其中第2AC包含可检测标记B；或与SPA和可检测标记B接触；以及

(iv) 选择至少一种与第一AC特异性结合的细胞，

其中，可检测标记A和可检测标记B彼此系可检测地不同

(v) 透过培养增加至少一个所选细胞数目；和

(vi) 将培养的细胞重新引入受试者体内。

37. 根据权利要求36所述的方法，其中，有需要的受试者患有选自由免疫疾病或肿瘤性疾病、病毒引起的疾病或细菌引起的疾病组成组的疾病，或处于患上所述疾病的风险。

38. 一种在其表面表达与目的蛋白抗原(PAI)特异性和/或选择性结合的抗原结合蛋白的选定免疫细胞在治疗免疫疾病或肿瘤性疾病、病毒引起的疾病或细菌引起的疾病中的用途，其包括以下步骤：

(i) 提出包含免疫细胞的细胞群；

(ii) 使步骤(i)的细胞群与包含PAI和可检测标记A的第一抗原复合体(第1AC)或包含可检测标记A的PAI接触；

(iii) 使步骤(i)的细胞群与包含SPA的至少第二抗原复合体(第2AC)接触，其中SPA的氨基酸序列与PAI的氨基酸序列至少有1个氨基酸不同，其中第2AC包含可检测标记B；或与SPA和可检测标记B接触；以及

(iv) 选择至少一种与第一AC特异性结合的细胞，

其中，可检测标记A和可检测标记B彼此系可检测地不同；肿瘤性疾病优选为癌症。

39. 一种选择免疫细胞的方法，所述免疫细胞在其表面表达与目的蛋白抗原(PAI)特异性结合的抗原结合蛋白，所述方法包括以下步骤：

(i) 提出包含免疫细胞的细胞群；

(ii) 使步骤(i)的细胞群与包含PAI和可检测标记A的第一抗原复合体(第1AC)或包含可检测标记A的PAI接触；

(iii) 使步骤(i)的细胞群与包含无关蛋白抗原(IPA)的至少第二抗原复合体(第2AC)接触，其中当IPA的氨基酸序列与PAI的氨基酸序列比对时，IPA的氨基酸序列的两个或两个

以下氨基酸位置与PAI相同,其中IAC包含可检测标记G;或与IPA和可检测标记G接触;以及
(iv) 选择至少一种与第一AC特异性结合的细胞,
其中,可检测标记A和可检测标记G彼此系可检测地不同。

使用序列相似肽进行反向选择分选

[0001] 本发明涉及一种选择细胞或病毒的方法,所述细胞或病毒在其表面表达与目的蛋白抗原 (PAI) 特异性结合的抗原结合蛋白,同时应用使用相似蛋白抗原 (SPA) 的反向选择。此外,本发明提出了一种确定编码抗原结合蛋白或其抗原结合部分的核酸之序列的方法,以及一种产生表达编码抗原结合蛋白或其抗原结合部分的核酸之细胞的方法。本发明还涉及一种用选定细胞群治疗受试者的方法。

[0002] 发明背景

[0003] 过继细胞转移 (ACT) 领域已成为治疗癌症、病毒感染和其他免疫调节性疾病的最有希望的创新方法之一。为了支持T细胞受体 (TCR) 修饰T细胞在临床上更广泛地应用,重要的是可在优选地临床前层级上正确识别和减轻风险。迄今为止,给予TCR修饰T细胞所观察到的毒性与标准ACT期间观察到的毒性相似,可大致分为三个主要类别:淋巴耗竭制剂方案引起的毒性、细胞因子相关毒性和免疫相关毒性。免疫相关毒性可分为两个子类别:即“肿瘤外/靶点”效应和“肿瘤外/脱靶”效应。最佳的基因改造T细胞治疗靶抗原是一种仅存在于肿瘤细胞上而不存在于健康细胞中的抗原;但是,大多数情况下,选定肿瘤靶抗原为过度表达或异常表达的蛋白,可能以不同程度存在于正常细胞中 (Johnson LA et al., Gene therapy with human and mouse T-cell receptors mediate cancer regression and targets normal tissues expressing cognate antigen, Blood 2009;114:535-46)。

[0004] “肿瘤外/靶点”毒性

[0005] 因此,基因改造T细胞治疗可能触发针对正常细胞 (即使是以低水平表达靶抗原的细胞) 的强细胞免疫反应。此类型的毒性称为“肿瘤外/靶点”,是由于例如改造T细胞无法区分正常细胞和表达靶抗原的癌细胞所致。靶向作用于Melan A (MLA; 也称为“T细胞识别黑色素瘤抗原1” (MART-1)) 与明显的“肿瘤外/靶点”副作用相关 (Johnson LA et al., Gene therapy with and mouse T-cell receptors mediates cancer regression and targets normal tissue expressing cognate antigens, Blood 2009, 114:535-46; van den Berg JH et al., Case report of a Fatal Serious Adverse event upon Administration of T-cells transduced with a MART-1-specific T-cell Receptor, Mol. Ther. 2015;23:1541-50)。具体而言,已发表的一份病例报告描述了转移性黑色素瘤患者接受MART-1特异性TCR转导T细胞给药后3天发生致命严重不良事件。从血液、支气管肺泡灌洗液、腹水、肿瘤部位和心脏组织中回收输注的T细胞,尽管未观察到修饰过T细胞对3-D跳动心肌细胞产生交叉反应,但作者无法排除与同种异体MHC肽复合体发生交叉反应的可能性。另外,发现多器官衰竭由靶点细胞因子释放所致。可透过选择显示足够低肿瘤外表达的靶抗原来避免肿瘤外/靶点毒性,从而在应用对肿瘤治疗上有效的剂量后使毒性可接受。

[0006] “肿瘤外/脱靶”毒性

[0007] 由于大多数肿瘤抗原均源自自身蛋白 (肿瘤相关抗原),因此,胸腺选择可有效防止高亲和力肿瘤特异性T细胞的分离。但是,TCR亲和力可透过互补决定区 (CDR) 内特定区域突变而大大增强。尽管这种方法因TCR简并性可用于提高修饰T细胞疗效,但存在下列风险:TCR可能透过交叉反应识别提呈于正常组织上的其他相关肽抗原。之前发表的结果显示:两

名患者发生致命的毒性,这两名患者输注了改造以表达靶向作用于黑色素瘤相关抗原A3 (MAGE-A3)之TCR的T细胞,该TCR与肌肉蛋白Titin中的肽发生交叉反应,但临床前研究中未预测到交叉反应性(Linette,GP et al.,Cardiovascular toxicity and titin cross-reactivity of affinity enhanced T-cells in myeloma and melanoma,Blood 2013; 122:863-71;Cameron,BJ et al.,Identification of a Titin-derived HLA-A1-presented peptide as a cross-reactive target for engineered MAGE-A3 directed T-cells,Sci.Transl.Med.2013;5:197-103)。这些患者证明,TCR改造T细胞可能具有严重且难以预测的脱靶和器官特异性毒性,并突出表明需要改进方法来定义改造TCR的特异性。因此,临床前研究中建议使用诸如肽扫描和使用更复杂细胞结构之类的策略,从而降低未来临床研究中脱靶毒性的风险。因此,开发和提出低肿瘤外/脱靶毒性的TCR仍为未满足的医疗需求。本发明提出了快速鉴定抗原结合分子(特别是与其靶抗原特异性和选择性结合的TCR)的方法,从而提供提高安全性以及降低对序列相似靶抗原(特别是健康组织上的序列相似肽)的交叉反应性的方法。本发明的快速(优选为一步)选择方法在鉴定表达具有所需抗肿瘤活性的TCR的患者来源T细胞中特别有用。

发明摘要

[0008] 本发明的第一方面涉及一种选择细胞或病毒的方法,所述细胞或病毒在其表面表达与目的蛋白抗原(PAI)特异性和/或选择性结合的抗原结合蛋白,所述方法包括以下步骤:

[0009] (i) 提出包含细胞的细胞群或病毒群;

[0010] (ii) 使步骤(i)的细胞群或病毒群与包含PAI和可检测标记A的第一抗原复合物(第1AC)或包含可检测标记A的PAI接触;

[0011] (iii) 使步骤(i)的细胞群或病毒群与包含相似蛋白抗原(SPA)的至少第二抗原复合物(第2 AC)接触,其中SPA的氨基酸序列与PAI的氨基酸序列至少有1个氨基酸不同,其中第2 AC包含可检测标记B;或与SPA和可检测标记B接触;以及

[0012] (iv) 选择至少一种与第一AC特异性和/或选择性结合的细胞或病毒,

[0013] 其中,可检测标记A和可检测标记B彼此系可检测地不同。

[0014] 本发明的第二方面进一步涉及一种确定编码抗原结合蛋白或其抗原结合部分的核酸序列的方法,所述方法包括以下步骤:

[0015] (i) 分离来自本发明第一方面的方法中选定细胞的编码抗原结合蛋白或其抗原结合部分的核酸;和

[0016] (ii) 确定核酸的序列。

[0017] 本发明的第三方面涉及一种产生表达编码抗原结合蛋白或其抗原结合部分的核酸的细胞的方法,所述方法包括以下步骤:

[0018] (i) 提供来自本发明第一方面的方法中选定细胞的编码抗原结合蛋白或其抗原结合部分的核酸序列;

[0019] (ii) 任选地在表达控制元素控制下产生包含步骤(i)中所提供核酸序列的核酸载体;和

[0020] (iii) 将步骤(ii)的核酸载体引入宿主细胞。

[0021] 本发明的第四方面涉及一种用于治疗有需要的受试者的方法,所述方法包括以下步骤:

[0022] (i) 提出包含免疫细胞的受试者细胞群;

[0023] (ii) 使步骤(i)的细胞群与包含PAI和可检测标记A的第一抗原复合体(第1 AC)接触;

[0024] (iii) 使步骤(i)的细胞群与包含SPA的至少第二抗原复合体(第2 AC)接触,其中SPA的氨基酸序列与PAI的氨基酸序列至少有1个氨基酸不同,其中第2 AC包含可检测标记B;以及

[0025] (iv) 选择至少一种与第一AC特异性结合的细胞,

[0026] 其中,可检测标记A和可检测标记B彼此系可检测地不同

[0027] (v) 透过培养增加步骤(iv)中选择的至少一个所选细胞数目;和

[0028] (vi) 将培养的细胞重新引入受试者体内。

[0029] 本发明的第五方面涉及一种选择免疫细胞的方法,所述免疫细胞在其表面表达与目的蛋白抗原(PAI)特异性结合的抗原结合蛋白,所述方法包括以下步骤:

[0030] (i) 提出包含免疫细胞的细胞群;

[0031] (ii) 使步骤(i)的细胞群与包含PAI和可检测标记A的第一抗原复合体(第1 AC)或包含可检测标记A的PAI接触;

[0032] (iii) 使步骤(i)的细胞群与包含无关蛋白抗原(IPA)的至少第二抗原复合体(第2 AC)接触,其中当IPA的氨基酸序列与PAI的氨基酸序列比对时,IPA的氨基酸序列的两个或两个以下氨基酸位置与PAI相同,其中IAC包含可检测标记G;或与IPA和可检测标记G接触;以及

[0033] (iv) 选择至少一种与第一AC特异性结合的细胞,

[0034] 其中,可检测标记A和可检测标记G彼此系可检测地不同。

[0035] 附图列表

[0036] 下文描述了本说明书中包括的附图内容。在此背景中,也请参考上面和/或下面本发明的详细描述。

[0037] 图1显示了本发明的两个示例性应用的示意图。该图的上部路径表示将门控策略应用于经过单独T细胞培养步骤的经引发T细胞时的使用,图下部分表示门控策略在直接分选方法中的使用,其中使用门控策略从天然细胞库获得的异质T细胞群富含靶标特异性T细胞。在两个实施例中,正向分选细胞部分代表在其表面表达与目的蛋白抗原特异性和/或选择性结合的抗原结合蛋白的免疫细胞。图中使用的缩写:SPA:相似蛋白抗原,PAI:目的蛋白抗原,APC:抗原提呈细胞。

[0038] 图2显示了非扩增靶标特异性T细胞的示例性门控策略。为了提高测试样本中低频靶标特异性T细胞的频率,已透过荧光染料-四聚体特异性磁珠分离法富集这些细胞。随后,细胞进行表面标志物染色并透过流式细胞仪评估。单独的2D彩色四聚体组合用于染色靶标特异性和相似肽特异性T细胞。在此实施例中,1.65%的CD8⁺ T细胞与靶肽四聚体结合,且那些靶标特异性CD8⁺ T细胞中的29.4%也与相似肽四聚体结合(Target⁺/SIM⁺),该相似肽四聚体包含3种不同的相似肽-HLA。透过将相似肽四聚体纳入分选程序,可排除高比例的交叉反应性T细胞(Target⁺/SIM⁺)。

[0039] 图3显示了经引发T细胞群的示例性门控策略。单个T细胞培养物用靶肽HLA包被人工提呈细胞反复刺激,从而增强低频靶标特异性CD8⁺ T细胞。培养4周后,对那些经引发T细胞培养物进行表面标志物染色,并以单个2D-彩色四聚体组合进行靶标-HLA和3种相似肽-HLA染色。上图显示了与靶肽四聚体和相似肽四聚体两者(Target⁺/SIM⁺)结合的单克隆富集T细胞群。下图显示了仅与靶肽四聚体结合但不与相似肽四聚体结合的单克隆富集T细胞群。透过将相似肽四聚体纳入染色程序,可从分选中排除交叉反应性T细胞(Target⁺/SIM⁺)。

[0040] 图4显示了仅使用靶肽四聚体分选的T细胞中的TCR可以与靶相似肽发生交叉反应。在mRNA电穿孔入健康供体T细胞后,评估使用靶肽四聚体鉴定的TCR对10种靶相似肽的交叉反应性。评估与肽载入T2细胞共培养后的IFN γ 分泌作为衡量反应性的指标。本实施例中的所有TCR与靶肽(阳性对照)反应,而与无关肽载入T2细胞、未载入T2细胞或仅效应细胞的对照不产生反应。但是,图4A和图4B中的TCR也显示了对相似肽(即,图4A中的TCR对相似肽1和10,图4B中的TCR对相似肽9和10)的反应性。仅图4C中的TCR未显示出交叉反应性,因此被选择进行进一步表征。

[0041] 图5显示了从仅与靶肽四聚体(TCR PAI⁺/SPA⁻)结合的T细胞中分离的TCR以及对于对照肽(“对照肽”)具有特异性的对照TCR和无TCR对照(“无肽”)的功能性评估。为此,TCR-mRNA被电穿孔至NFAT萤光素酶Jurkat报告细胞中,并在与肽/靶相似肽(“SIM 1、SIM 2和SIM 3”)载入T2抗原提呈细胞共培养后评估其激活情况。仅当与靶肽载入T2细胞共培养时,衍生自PAI⁺/SPA⁻分选T细胞的TCR才触发激活。对照TCR在存在对照肽的情况下显示反应性,而无TCR mRNA电穿孔的Jurkat细胞不与肽载入T2细胞反应。本实施例表明,与靶肽四聚体结合的TCR也在功能层级上显示出对那些肽的反应性。

[0042] 图6、7和8显示了基于XPRESIDENT质谱数据来自实施例4的靶相似肽1(TSP1)(图6)、TSP2(图7)和不相关肽(IP;图8)的肽提呈特征。上面部分:根据技术重复测量值的中位元相对MS信号强度绘制为单一HLA-A*02正常组织的彩色点,在其上检测到了肽。正常样本根据来源器官分组。箱须图表示多个样本的归一化信号强度,并在对数空间中定义。箱显示中位元值、第25和第75百分位数。须延伸到最低数据点仍在下四分位数的1.5四分位数范围(IQR)内,最高数据点仍在上四分位数的1.5 IQR内。下面部分:每个器官的肽检测频率显示为条形图。图表下面的数位表示每个器官分析的总样本数中检测到肽的样本数(所有器官的正常样本N \geq 628)。如果在一个样本上检测到肽,但因技术原因无法量化,则该样本纳入检测频率图中,但图表上面部分不显示任何点。脂肪:脂肪组织;肾上腺:膀胱;膀胱;血管;食管;胆囊;大肠;小肠;中枢神经;周围神经;甲状旁腺;腹膜;垂体;骨骼肌。

[0043] 选定序列列表

[0044] SEQ ID NO:1 X₁-X₂-X₃-X₄-X₅-X₆-X₇-X₈,其中X₁-X₈是长度为8个氨基酸的靶肽中氨基酸位置,每种情况下的X为任何氨基酸;

[0045] SEQ ID NO:2 X₁-X₂-X₃-X₄-X₅-X₆-X₇-X₈-X₉,其中X₁-X₉是长度为9个氨基酸的靶肽中氨基酸位置,每种情况下的X为任何氨基酸;

[0046] SEQ ID NO:3 X₁-X₂-X₃-X₄-X₅-X₆-X₇-X₈-X₉-X₁₀,其中X₁-X₁₀是长度为10个氨基酸的靶肽中氨基酸位置,每种情况下的X为任何氨基酸;

[0047] SEQ ID NO:4 X₁-X₂-X₃-X₄-X₅-X₆-X₇-X₈-X₉-X₁₀-X₁₁,其中X₁-X₁₁是长度为11个氨基酸的靶肽中氨基酸位置,每种情况下的X为任何氨基酸;

[0048] SEQ ID NO:5 $X_1-X_2-X_3-X_4-X_5-X_6-X_7-X_8-X_9-X_{10}-X_{11}-X_{12}$, 其中 X_1-X_{12} 是长度为12个氨基酸的靶肽中氨基酸位置, 每种情况下的X为任何氨基酸;

[0049] 本发明的详细说明

[0050] 在下面详细说明本发明之前, 应当理解, 本发明并不限于本文所述的特定方法、方案和试剂, 因为它们可能有变化。也应当理解, 本文所用的术语之目的仅为了描述特定实施方案, 而不是为了限制本发明的范围, 本发明的范围仅由所附权利要求限定。除非另有定义, 本文使用的所有技术和科学术语都与本领域普通技术人员通常的理解具有相同的含义。

[0051] 本说明书的全文中引用了一些文件。无论是上文还是下文, 本文引用的每份文件(包括所有专利、专利申请、科学出版物、制造商说明书、说明书等)均透过引用整体并入本文。本文中的任何内容均不应解释为承认本发明无权凭借在先发明而早于这种公开内容。本文引用的一些文件表征为“透过引用并入”。如果在这种引入的参考文献的定义或教导与本说明书中引用的定义或教导之间发生冲突的情况下, 以本说明书的文本为准。

[0052] 在下文中, 将描述本发明的要素。这些要素与特定实施方案一起列出, 但是, 应当理解, 它们可以以任何方式和任何数目组合, 以产生额外的实施方案。各种描述的实施例和优选实施方案不应该被解释为将本发明限制于仅明确描述的实施方案。这种描述应该被理解为支持并包含将明确描述的实施方案与任何数目的所公开和/或优选要素组合的实施方案。此外, 除非上下文另有指出, 否则本申请中所有描述要素的任何排列和组合都应视为由本申请的描述公开。

[0053] 定义

[0054] 为了实施本发明, 除非另有说明, 否则采用本领域文献中解释的化学、生物化学和重组DNA技术的常规方法(例如, 参见Molecular Cloning: A Laboratory Manual, 2nd Edition, J.Sambrook et al. eds., Cold Spring Harbor Laboratory Press, Cold Spring Harbor 1989)。

[0055] 在下文中, 提供了本说明书中经常来表征本发明的术语一些定义。这些术语在本说明书其余部分每种使用情况下将具有分别定义的含义和优选含义。

[0056] 除非内容中另外明确指出, 否则, 本说明书和所附权利要求书中所使用的单数形式“一个”、“一种”和“该”包括复数的指示物。

[0057] 在本发明背景中, 术语“氨基酸”系指包含取代或未取代氨基、取代或未取代羧基、以及一个或多个侧链或基团或任一个这些基团的类似物的任何单体单位。示例性侧链包括, 例如: 硫醇、硒基、磺醯基、烷基、芳基、醯基、酮基、叠氨基、羟基、胍、氰基、卤素、醯胍、烯基、炔基、醚、硼酸盐(borate)、硼酸酯(boronate)、磷酸盐、膦醯基、膦基、杂环基、烯酮、亚胺、醛、酯、硫代酸、羟胺或这些基团的任何组合。其他代表性氨基酸包括但不限于: 包含光活化交联剂的氨基酸、金属结合氨基酸、自旋标记氨基酸、荧光氨基酸、含金属氨基酸、具有新型官能基的氨基酸、与其他分子共价或非共价相互作用的氨基酸、光笼化和/或可光异构化氨基酸、放射性氨基酸、包含生物素或生物素类似物的氨基酸、糖基化氨基酸、其他碳水化合物修饰氨基酸、包含聚乙二醇或聚醚的氨基酸、重原子取代氨基酸、化学可裂解和/或光降解氨基酸、碳连接含糖氨基酸、氧化还原活性氨基酸、含氨基硫代酸的氨基酸以及包含一个或多个有毒部分的氨基酸。本文所用的术语“氨基酸”包括以下二十种天然或基因编码

的 α -氨基酸:丙氨酸(Ala或A)、精氨酸(Arg或R)、天冬酰胺(Asn或N)、天冬氨酸(Asp或D)、半胱氨酸(Cys或C)、谷氨酰胺(Gln或Q)、谷氨酸(Glu或E)、甘氨酸(Gly或G)、组氨酸(His或H)、异亮氨酸(Ile或I)、亮氨酸(Leu或L)、赖氨酸(Lys或K)、蛋氨酸(Met或M)、苯丙氨酸(Phe或F)、脯氨酸(Pro或P)、丝氨酸(Ser或S)、苏氨酸(Thr或T)、色氨酸(Trp或W)、酪氨酸(Tyr或Y)和缬氨酸(Val或V)。在未指定“X”个残基数目的情况下,则将其解释为“任何氨基酸”。这二十种天然氨基酸的结构参见例如:Stryer et al.,*Biochemistry*,5th ed.,Freeman and Company (2002)。其他氨基酸,例如:硒代半胱氨酸和吡咯赖氨酸,也可以进行基因编码(Stadtman(1996)“Selenocysteine,”*Annu Rev Biochem.*65:83-100and Ibba et al.(2002)“Genetic code:introducing pyrrolysine,”*Curr Biol.*12(13):R464-R466)。术语“氨基酸”还包括非天然氨基酸、修饰氨基酸(例如,含有修饰侧链和/或骨架)和氨基酸类似物。例如,参见Zhang et al.(2004)“Selective incorporation of 5-hydroxytryptophan into proteins in mammalian cells,”*Proc.Natl.Acad.Sci.U.S.A.*101(24):8882-8887,Anderson et al.(2004)“An expanded genetic code with a functional quadruplet codon”*Proc.Natl.Acad.Sci.U.S.A.*101(20):7566-7571,Ikeda et al.(2003)“Synthesis of a novel histidine analogue and its efficient incorporation into a protein in vivo,”*Protein Eng.Des.Sel.*16(9):699-706,Chin et al.(2003)“An Expanded Eukaryotic Genetic Code,”*Science* 301(5635):964-967,James et al.(2001)“Kinetic characterization of ribonuclease S mutants containing photoisomerizable phenylazophenylalanine residues,”*Protein Eng.Des.Sel.*14(12):983-991,Kohrer et al.(2001)“Import of amber and ochre suppressor tRNAs into mammalian cells:A general approach to site-specific insertion of amino acid analogues into proteins,”*Proc.Natl.Acad.Sci.U.S.A.*98(25):14310-14315,Bacher et al.(2001)“Selection and Characterization of *Escherichia coli* Variants Capable of Growth on an Otherwise Toxic Tryptophan Analogue,”*J.Bacteriol.*183(18):5414-5425,Hamano-Takaku et al.(2000)“A Mutant *Escherichia coli* Tyrosyl-tRNA Synthetase Utilizes the Unnatural Amino Acid Azatyrosine More Efficiently than Tyrosine,”*J.Biol.Chem.*275(51):40324-40328,and Budisa et al.(2001)“Proteins with β -(thienopyrrolyl)alanines as alternative chromophores and pharmaceutically active amino acids,”*Protein Sci.*10(7):1281-1292。氨基酸可以合并为肽、多肽或蛋白质。本说明书中使用的术语“肽”系指透过肽键连接的氨基酸短聚合物。它具有与蛋白质相同的化学(肽)键,但长度通常较短。最短的肽为二肽,由透过单个肽键连接的两个氨基酸组成。也可以为三肽、四肽、五肽等。通常,肽的长度最多为8、10、12、15、18或20个氨基酸。除非为环状肽,否则肽有一个氨基端和一个羧基端。

[0058] 本发明背景下的术语“病毒”系指小的专性细胞内寄生体,根据定义,其含有被保护性蛋白外壳(即,衣壳)包围的RNA或DNA基因组。病毒的基因组可以由DNA或RNA组成,该DNA或RNA可以为线性或环状单链(ss)或双链(ds)。整个基因组可以占据一个核酸分子(单分体基因组)或几个核酸片段(多分体基因组)。所述病毒可为双链DNA病毒,优选为肌病毒科、长尾病毒科、短尾病毒科、疱疹病毒科、腺病毒科、杆状病毒科、乳头状瘤病毒科、多分体DNA病毒科、多瘤病毒科、痘病毒科;单链DNA病毒,优选为指环病毒科、丝杆病毒科、细小病

毒科;双链RNA病毒,优选为呼肠孤病毒科;单链RNA病毒,优选为冠状病毒科、微RNA病毒科、杯状病毒科、披膜病毒科、黄病毒科、星状病毒科、动脉炎病毒科、戊型肝炎病毒科;负义单链RNA病毒,优选为沙粒病毒科、丝状病毒科、副粘病毒科、弹状病毒科、布尼亚病毒科、正粘病毒科、玻那病毒科;单链RNA逆转录病毒,优选为逆转录病毒科;或双链RNA逆转录病毒,优选为花椰菜花叶病毒科、嗜肝DNA病毒科。

[0059] 本发明背景下的术语“噬菌体”系指在细菌和古生菌中感染并复制的病毒。噬菌体依赖于在宿主生物体(通常为细菌)中复制并将其基因组注入宿主生物体的细胞质,所述基因组包含封装去氧核糖核酸(DNA)或核糖核酸(RNA)的蛋白质。在生物技术中所使用噬菌体的突出实例有:噬菌体T4 λ (T4 λ)噬菌体、T7噬菌体、fd丝状噬菌体,特别为丝状M13噬菌体,所有这些噬菌体均具有某些优点和缺点。

[0060] 本发明背景下的术语“病毒群”系指在编码其表面上表达的抗原结合蛋白的基因资讯方面不同的大量病毒。因此,病毒群可以表达异源抗原结合蛋白的文库。

[0061] 本发明背景下的术语“噬菌体展示”或“噬菌体文库”系指用于高通量筛选蛋白质相互作用的系统。简言之,编码目的蛋白的基因被插入噬菌体外壳蛋白基因,使噬菌体在其表面“展示”(即,显示)蛋白,同时保持编码目的蛋白的基因在其DNA或RNA中。这可使基因型和表型联系。随后,可对噬菌体表面上展示的蛋白质进行其他蛋白质、肽或DNA序列筛选,以研究它们在所展示分子与待筛选分子之间的相互作用。此类分子可以为抗体或其片段、TCR或其片段、BCR或其片段或CAR或其片段。TCR的片段可包含 α 可变结构域和 β 可变结构域。下文也对本发明背景下的片段作了详细定义。

[0062] 本发明背景下的术语“细胞”系指含有细胞核和细胞胞器并且可见于原生动物、真菌、植物和动物的真核细胞。动物可包含哺乳动物细胞。哺乳动物细胞尤其包括人细胞、啮齿类细胞,例如小鼠或大鼠细胞、猴细胞、猪细胞或狗细胞。真菌细胞尤其包括酵母细胞。生物技术(例如:酵母表面展示技术)中使用的典型酵母细胞为酿酒酵母细胞。

[0063] 本发明背景下的术语“酵母表面展示”或“酵母展示”或“酵母文库”系指使用酵母细胞的蛋白工程技术,所述酵母细胞表达重组目的蛋白并将这些蛋白掺入其细胞壁。这可对蛋白质,特别是抗体或其片段、TCR或其片段、BCR或其片段或CAR或其片段进行分离和工程化处理。详细来说,在酵母表面展示技术中,选择单位为酵母细胞,其装饰有成千上万份目的蛋白质,并携带编码该蛋白的质粒。质粒可在酿酒酵母和大肠杆菌之间穿梭,在酿酒酵母中进行展示和分选,而在大肠杆菌中进行DNA制备和分子生物学。在Wittrup小组首创的酵母展示形式(Chao et al.,2006)中,每个¹⁰F_n3变体都表达为与细胞壁Aga2p中存在的天然酵母蛋白发生基因融合。Aga2p为天然酵母的一个结构域,即,凝集素交配因子;通常,将其克隆到编码¹⁰F_n3变体的序列上游。另外,紧接编码¹⁰F_n3变体的序列下游为工程化的表位标签(例如c-myc和V5)。诱导后,交配因子分泌信号肽引导融合蛋白从细胞中输出;被其结合伴侣Aga1p捕获至酵母细胞壁表面,并与之形成两个二硫键。结果得到培养物,其每个酵母细胞展示单个¹⁰F_n3变体的10,000至100,000个拷贝。平均而言,变体越热稳定,其在酵母表面上的分子数量就越大(Hackel et al.,2010)。

[0064] 本发明背景下的术语“免疫细胞”系指免疫系统的细胞。免疫系统包括不同的细胞类型,例如前体细胞,其包含最终分化为B和T淋巴细胞和天然杀伤(NK)细胞的淋巴样干细胞,最终分化为粒细胞和单核细胞的骨髓母细胞,以及完全分化的白细胞。分化白细胞为胸

腺、脾、骨髓或淋巴结来源的细胞,可分为粒细胞、B淋巴细胞、T淋巴细胞和单核细胞、巨噬细胞、肥大细胞和树突状细胞等主要类别。粒细胞进一步分为中性粒细胞、嗜酸性粒细胞和嗜碱性粒细胞,它们在血液循环中吞噬细菌、病毒或真菌。B淋巴细胞为浆细胞和B记忆细胞的前体。T细胞群包括调节性T细胞、记忆T细胞、T辅助细胞和细胞毒性T细胞。T辅助细胞激活浆细胞和天然杀伤细胞,但调节性T细胞抑制B和其他T细胞的功能,从而减慢免疫反应。T记忆细胞的寿命很长,并对特定抗原拥有记忆,而细胞毒性T细胞透过与肿瘤抗原或被攻击细胞的抗原相互作用来识别并杀灭肿瘤细胞或被病毒攻击的细胞。下表1列出了由各个T细胞特定表面标志物描述的T细胞及其表面表型的实例(根据Dong and Martinez, Nature Reviews Immunology, 2010一文):

	细胞:	表面标志物:
[0065]	细胞毒性 T 细胞:	α β TCR, CD3, CD8
	调节性 T 细胞:	α β TCR, CD3, CD4
	调节性 T 细胞 (天然和可诱导):	α β TCR, CD3, CD4, CD25, CTLA4, GITR
	天然杀伤细胞:	NK1.1, SLAMF1, SLAMF6, TGF β , V α 24, J α 18
	T 辅助细胞: TH1 细胞	α β TCR, CD3, CD4, IL-12R, IFN γ R, CXCR3;
	[0066]	TH2 细胞
TH9 细胞		α β TCR, CD3, CD4;
TH17 细胞		α β TCR, CD3, CD4, IL-23R, CCR6, IL-1R, CD161;
TH22 细胞		α β TCR, CD3, CD4, CCR10;
TIL		肿瘤浸润淋巴细胞
T 记忆细胞		CCR7 ^{hi} , CD44, CD62L ^{hi} , TCR, CD3, IL-7R (CD127), IL-15R

[0067] 表1:常见的T细胞表面标志物(非穷尽列举)。

[0068] 本发明背景下的术语“肿瘤浸润淋巴细胞”(TIL)系指向肿瘤迁移并通常可见于肿瘤基质或肿瘤本身的T细胞和B细胞。TIL通常包含可用于ACT或自体细胞治疗的白细胞细胞群。此类治疗已显示出有希望的结果,例如,在各种临床试验的转移性黑素瘤患者中显示出有希望的结果(Guo et al.; “Recent updates on cancer immunotherapy”; Precision Clinical Medicine, 1(2), 2018-65-74)。ACT背景下的TIL为从手术切除肿瘤或从肿瘤片段中分离的单细胞悬液离体扩增。TIL用高剂量细胞因子(例如, IL-2)扩增。然后在“快速扩增方案”(REP)中进一步扩增表现出最佳肿瘤反应性的所选TIL系,所述REP方案应用通常两周时间的抗CD3激活。最终REP后TIL被输回患者体内。此过程还可涉及初步化疗方案以消耗内源性淋巴细胞,从而为过继转移TIL提供包围肿瘤部位的足够通路。

[0069] 本发明背景下的术语“富含免疫细胞级分”系指衍生自天然存在细胞群(例如,血液)的细胞群,其中免疫细胞的相对丰度与天然存在细胞混合物的丰度相比增加。1ml健康人类受试者的血液包含例如:470万至610万(男性)、420万至540万(女性)个红细胞,4,000

至11,000个白细胞以及200,000至500,000个血小板。因此,在血液中,免疫细胞仅占血细胞总数的0.06%至0.25%。因此,血液中富含免疫细胞级分可包含大于0.25%、更优选大于10%、再更优选大于50%、再优选大于80%、最优选大于90%的免疫细胞。富含免疫细胞级分可以富含一种或多种免疫细胞亚型。例如,富含免疫细胞级分可富含淋巴样干细胞、T细胞、B细胞、浆细胞或组合。通常,透过使用一种或多种与目的免疫细胞表面标志物特异性结合的荧光标记抗体来选择免疫细胞富集级分中的免疫细胞。在上表1中指出了用于在T细胞组中选择T细胞或子级分的合适表面标志物。例如,可以透过使用与CD8特异性结合的抗体或透过使用与CD8和CD3特异性结合的抗体来选择细胞毒性T细胞。

[0070] 本发明背景下的术语“细胞群”系指多个细胞,可为同质或异质,即,具有不同特征的细胞混合物。血液是不同细胞混合物的细胞群的一个实例。均质细胞群可透过选择特定亚型或透过克隆扩增获得。

[0071] 本发明背景下的术语“抗原结合蛋白”系指一种多肽或两种或更多种多肽的复合体,其包含与抗原特异性结合的互补位(或称为“抗原结合位点”)。抗原结合蛋白的实例有单链抗体、单链TCR、嵌合抗原受体(CAR),抗原结合复合体的实例有抗体、B细胞受体(BCR)或TCR。

[0072] 本发明背景下的术语“嵌合抗原受体”(CAR;也称为嵌合免疫受体、嵌合T细胞受体、人工T细胞受体)系指改造的受体,其将任意特异性转移至免疫效应细胞上,优选为T细胞。细胞在基因上配有CAR(一种复合膜受体分子,可提供靶向特异性和T细胞激活)。CAR的最常见形式为衍生自单克隆抗体、与CD3跨膜结构域和胞内结构域融合的单链可变片段(scFv)融合物。CAR透过细胞外部分中抗体衍生结合结构域使T细胞靶向作用于所需的细胞靶标,当遇到靶标时,T细胞透过细胞内部分信号传导结构域激活。逆转录或慢病毒载体通常促进这些受体的编码序列转移至合适的细胞中,特别是T细胞中。受体称为嵌合体的原因是它们由不同来源的部分组成。

[0073] 本发明背景下的术语“表位”系指抗原的功能性表位。功能性表位包含有助于抗原结合蛋白互补位与抗原之间的非共价相互作用的那些残基,通常为氨基酸或多糖。非共价相互作用包含静电力、凡得瓦尔力、氢键和疏水相互作用。功能性表位为构成抗原结合蛋白结构表位的残基亚组。结构表位包含抗原结合蛋白覆盖的所有残基,即,抗原结合蛋白的印迹。通常情况下,与抗体结合的抗原的功能性表位包含4至10个氨基酸。类似情况下,MHC呈现的肽的功能性表位通常包含4至8个氨基酸。

[0074] 本发明背景下的术语“表达”系指人体组织中存在蛋白质或肽,特别是PAI或SPA。术语蛋白质或肽表达表示,蛋白质或肽在蛋白质生物合成过程中,在细胞核糖体机器中从其核酸序列被翻译为其氨基酸序列。所表达的蛋白质可位于细胞内或细胞外,例如,细胞表面。表达蛋白质的人体组织可以为健康组织或患病组织。

[0075] 本发明背景下的术语“目的蛋白抗原”(PAI)系指蛋白质或蛋白质的一部分或蛋白质复合体,其包含由抗原结合蛋白互补位特异性结合的表位。PAI通常为天然存在的蛋白质,可以为任何长度。PAI优选为包含至少25个氨基酸。如果PAI相应地由TCR特异性结合,则PAI优选长度为8至12个氨基酸。PAI可以为肿瘤相关靶抗原(TAA)、病毒蛋白或细菌蛋白。PAI通常为肿瘤相关抗原(TAA),其系例如肿瘤治疗中的特异性靶标。

[0076] 本发明背景下的术语“人源化小鼠”系指携带人类基因、细胞、组织和/或器官(发

挥其生物学功能,例如,就其生物学功能而言是完整的)的基因修饰小鼠。通常情况下,免疫缺陷小鼠用作人类细胞或组织的受体,它们因缺乏宿主免疫力,可以相对容易地接受异源细胞或组织。人源化小鼠的实例有:裸鼠、严重联合免疫缺陷(SCID)小鼠、NCG小鼠、NOG(NOD/Shi-scid/IL-2R γ null)小鼠或NSG(NOD scid γ)小鼠。在其各自小鼠基因座中接受人类版本基因的小鼠称为“敲入”小鼠。B细胞和T细胞可从人源化小鼠中分离,并可用于本发明的方法。

[0077] 本发明背景下的术语“T细胞受体文库”系指包含大量不同T细胞受体(TCR)蛋白或其片段的文库,其中各个TCR蛋白或其片段不同。

[0078] 本发明背景下的“病毒抗原肽”是病毒蛋白的较短片段,其由主要组织相容性复合体(MHC)分子提呈于抗原提呈细胞(通常为患病细胞)表面上。病毒抗原肽来源于病毒,即,细胞通常被所述病毒感染。本发明背景下的病毒抗原肽可以为选自人类免疫缺陷病毒(HIV)抗原肽、人巨细胞病毒(HCMV)抗原肽、巨细胞病毒(CMV)抗原肽、人乳头瘤病毒(HPV)抗原肽、乙型肝炎病毒(HBV)抗原肽、丙型肝炎病毒(HCV)抗原肽、爱泼斯坦-巴尔病毒(EBV)抗原肽、流感抗原肽,优选为HIV、HBV、流感和HCMV抗原肽组成的组中的抗原肽。病毒抗原肽可用于本文所述方法,以及包括,例如,本文所述的实施方案包括下表2中所述的病毒抗原肽。一方面,在本文所述的方法和实施方案中使用的病毒抗原肽包括至少一种病毒抗原肽,其包含选自SEQ ID NO:6至SEQ ID NO:8的氨基酸序列中的的氨基酸序列或由其组成。

[0079] 表2:病毒抗原肽列表

[0080]

SEQ ID NO:	氨基酸序列	病毒	MHC
6	SLYNTVATL	HIV	HLA-A*02:01
7	GILGFVFTL	甲型流感	HLA-A*02:01
8	NLVPMVATV	HCMV	HLA-A*02:01

[0081] 本发明背景下的“细菌抗原肽”是细菌蛋白的较短片段,其由MHC分子提呈于抗原提呈细胞(通常为患病细胞)表面上。细菌抗原肽来源于细菌,即,细胞通常被细菌感染。此类细菌抗原肽已经在例如结核分枝杆菌感染的背景下发现。因此,本发明背景下的细菌抗原肽可为结核分枝杆菌抗原肽。

[0082] 本发明背景下的术语“肿瘤相关抗原”(TAA)系指衍生自所有蛋白质类别的自体细胞抗原,例如,由肿瘤细胞优先或专有表达的酶、受体、转录因子等。TAA可以大致分类为异常表达自身抗原、突变自身抗原和肿瘤特异性抗原。正常组织中也发现了由肿瘤细胞优先表达的TAA。但是,它们的表达与正常组织的表达不同之处在于:它们在肿瘤中的表达程度、与它们的正常对应物相比其蛋白质结构的变化或它们在肿瘤细胞内的异常亚细胞定位。可用于本文所述方法和实施方案中的TAA肽包括,例如,美国专利公开号20160187351、美国专利公开号20170165335、美国专利公开号20170035807、美国专利公开号20160280759、美国专利公开号20160287687、美国专利公开号20160346371、美国专利公开号20160368965、美国专利公开号20170022251、美国专利公开号20170002055、美国专利公开号20170029486、美国专利公开号20170037089、美国专利公开号20170136108、美国专利公开号20170101473、美国专利公开号20170096461、美国专利公开号20170165337、美国专利公开号20170189505、美国专利公开号20170173132、美国专利公开号20170296640、美国专利公开号20170253633、美国专利公开号20170260249、美国专利公开号20180051080和美国专利

公开号20180164315所述的TAA肽,本文所述的这些专利公开内容和序列表透过引用整体并入本文。此外,在本发明背景下的TAA为MHC-I类分子或MHC-II类分子(优选为MHC-I类分子)的特异性配体。

[0083] 一方面,透过本发明方法选择的抗原结合蛋白选择性地识别提呈以上列出一个或多个专利和出版物中所述TAA肽的细胞。另一方面,可用于本文所述方法和实施方案的TAA肽包括至少一个TAA,其由选自SEQ ID NO:9至164的氨基酸序列的氨基酸序列组成。一方面,透过本发明方法选择的抗原结合蛋白选择性结合提呈TAA肽/MHC复合体的细胞,其中TAA肽包含SEQ ID NO:1至164的氨基酸序列或由其组成。表3列出了TAA的其他实例。

[0084] 表3:TAA列表

SEQ ID NO :	氨基酸序列	SEQ ID NO :	氨基酸序列	SEQ ID NO :	氨基酸序列
9	YLYDSETKN A	61	VLDGKVAVV	113	KMSELQTYV
10	HLMDQPLSV	62	GLLGKVTSV	114	ALLEQTGDMSL
11	GLLKKINSV	63	KMISAIPTL	115	VIIKGLEEITV
12	FLVDGSSAL	64	GLLETTGLLA T	116	KQFEGTVEI

[0086]

13	FLFDGSANLV	65	TLNTLDINL	117	KLQEEIPVL
14	FLYKIIDEL	66	VIKGLEEI	118	GLAEFQENV
15	FILDSAETTL	67	YLEDGFAYV	119	NVAEIVIHI
16	SVDVSPPKV	68	KIWEELSVLE V	120	ALAGIVTNV
17	VADKIHSV	69	LLIPFTIFM	121	NLLIDDKGTIKL
18	IVDDLINL	70	ISLDEVAVSL	122	VLMQDSRLYL
19	GLLEELVTV	71	KISDFGLATV	123	KVLEHVVRV
20	TLDGAAVNQ V	72	KLIGNIHGNE V	124	LLWGNLPEI
21	SVLEKEIYSI	73	ILLSVLHQL	125	SLMEKNQSL
22	LLDPKTIFL	74	LDSEALLTL	126	KLLAVIHEL
23	YLMDDFSSL	75	VLQENSSDY QSNL	127	ALGDKFLLRV
24	KVWSDVTPL	76	HLLGEGAFA QV	128	FLMKNSDLYGA
25	LLWGHPRVA LA	77	SLVENIHVL	129	KLIDHQGLYL
26	KIWEELSVLE V	78	SLSEKSPEV	130	GPGIFPPPPQP
27	LLIPFTIFM	79	AMFPDTIPRV	131	ALNESLVEC
28	FLIENLLAA	80	FLIENLLAA	132	GLAALAVHL
29	LLWGHPRVA LA	81	FTAEFLEKV	133	LLLEAVWHL
30	FLLEREQLL	82	ALYGNVQQV	134	SIIYLPTL
31	SLAETIFIV	83	LFQSRIAGV	135	TLHDQVHLL
32	TLLEGISRA	84	ILAEPIYIRV	136	SLLMWITQC

[0087]

33	ILQDGQFLV	85	FLLEREQLL	137	FLLDKPQDLSI
34	VIFEGEPMYL	86	LLPLELSLA	138	YLLDMPLWYL
35	SLFESLEYL	87	SLAETIFIV	139	GLLDCPIFL
36	SLLNQPKAV	88	AILNVDEKNQ V	140	VLIEYNFSI
37	GLAEFQENV	89	RLFEEVLGV	141	TLYNPERTITV
38	KLLAVIHEL	90	YLDEVAFML	142	AVPPPPSSV
39	TLHDQVHLL	91	KLIDEDEPLF L	143	KLQEELNKV
40	TLYNPERTIT V	92	KLFEKSTGL	144	KLMDPGSLPPL
41	KLQEKIQEL	93	SLLEVNEASS V	145	ALIVSLPYL
42	SVLEKEIYSI	94	GVYDGREHT V	146	FLLDGSANV
43	RVIDDSL VVG V	95	GLYPVTLVG V	147	ALDPSGNQLI
44	VLFGELPAL	96	ALLSSVAEA	148	ILIKHLVKV
45	GLVDIMVHL	97	TLLEGISRA	149	VLLDTILQL
46	FLNAIETAL	98	SLIEESEEL	150	HLIAEIHTA
47	ALLQALMEL	99	ALYVQAPTV	151	SMNGGVFAV
48	ALSSQAEV	100	KLIYKDLVSV	152	MLAEKLLQA
49	SLITGQDLLS V	101	ILQDGQFLV	153	YMLDIFHEV
50	QLIEKNWLL	102	SLLDYEVSI	154	ALWLPTDSATV
51	LLDPKTIFL	103	LLGDSSFFL	155	GLASRILDA
52	RLHDENILL	104	VIFEGEPMYL	156	ALSVLRLAL

[0088]	53	GLPSATTTV	105	ALSYILPYL	157	SYVKVLHHL
	54	GLLPSAESIKL	106	FLFVDPELV	158	VYLPKIPSW
	55	KTASINQNV	107	SEWGSPhAA VP	159	NYEDHFPLL
	56	SLLQHLIGL	108	ALSELERVL	160	VYIAELEKI
	57	YLMDDFSSL	109	SLFESLEYL	161	VHFEDTGKTLF
	58	LMYPYIYHV	110	KVLEYVIKV	162	VLSPFILTL
	59	KVWSDVTPL	111	VLLNEILEQV	163	HLLEGSVGV
	60	LLWGHPRVA LA	112	SLLNQPKAV		

[0089] 此外,在本发明背景下的TAA抗原肽为MHC-I类分子或MHC-II类分子(优选为MHC-I类分子)的特异性配体。

[0090] 本发明背景下的术语“肿瘤特异性抗原”系指只在肿瘤细胞上表达的抗原。它们包括因突变而产生的新抗原,例如:肿瘤细胞中的点突变或移码突变。肿瘤特异性抗原的实例有p53或BCR-ABL。

[0091] 本发明背景下的术语“MHC”是短语“主要组织相容性复合体”的缩写。MHC是一组细胞表面受体,其在脊椎动物中对于针对改变的天然或外来蛋白质建立获得性免疫起重要作用,进而决定了组织内的组织相容性。MHC分子的主要功能是与来自改变蛋白质或病原体的抗原结合,并将它们展示于细胞表面上以供适当的T细胞识别。人MHC也称为HLA(人白细胞抗原)复合体或HLA。MHC基因家族分为三个亚组:I类、II类和III类。肽和MHC I类的复合体由负载适当TCR的CD8阳性T细胞进行识别,而肽和MHC II类分子的复合体由负载适当TCR的CD4阳性辅助T细胞进行识别。由于CD8依赖型和CD4依赖型这两种反应共同并协同地促进抗肿瘤作用,因此,确定和表征肿瘤相关抗原和相应TCR在开发癌症免疫治疗(如:疫苗和细胞治疗)中非常重要。

[0092] 本发明背景下的术语“MHC-I”系指MHC I类分子或MHC-I。MHC I分子由一条 α 链(也称为MHC I重链)和一条 β 链(其构成 β 2微球蛋白分子)组成。 α 链包含三个 α 结构域,即, α 1结构域、 α 2结构域和 α 3结构域。 α 1和 α 2结构域主要有助于形成肽袋,以产生肽配体MHC (pMHC)复合体。MHC-I通常与源自胞质抗原蛋白的肽结合,所述肽在泛素化后被蛋白酶体降解,随后透过抗原加工相关特异性转运蛋白(TAP)从胞质溶质被转运至内质网(ER)。MCH I通常与长度为8-12个氨基酸的肽结合。

[0093] 本发明背景下的术语“MHC-II”系指MHC II类分子或MHC-II。MHC-II分子由 α 和 β 链组成,其中 α 链包含两个 α 结构域,即 α 1结构域和 α 2结构域, β 链包含两个 β 结构域,即 β 1结构域和 β 2结构域。MHC II通常在ER中折叠,与称为不变链的蛋白质复合,然后被转运至晚期内体区室,在此,不变链被组织蛋白酶切断,并且一段短的片段仍与MHC II的肽结合槽保持结合,称为II类相关不变链肽(CLIP)。然后,这种占位肽通常与更高亲和力的肽交换,后者衍

生自胞吞区室中的蛋白水解降解蛋白。MHC-II通常与长度为10-30个氨基酸的肽或与长度为13-25个氨基酸的肽结合。

[0094] 本发明背景下的术语“HLA”系指不同人类之间氨基酸序列不同的分子。但是,HLA可以透过国际公认的HLA命名法(IMGT命名法)来识别。HLA-A基因位于染色体6的短臂上并编码HLA-A的较大 α -链成分。HLA-A α -链的变异是HLA功能的关键。这种变异促进了人群中的遗传多样性。由于每种HLA对某些结构的肽均具有不同的亲和力,因此,更多种类的HLA意味着在细胞表面上“提呈”更多种类的抗原。每个人最多可以表达两种类型的HLA-A,各来自父母双方。有些人将从父母双方遗传相同的HLA-A,这降低了其个体HLA多样性。但是,大多数人都遗传两种不同拷贝的HLA-A。所有HLA组都具有相同的模式。换句话说,每个人只能表达编码当前1740种活性蛋白的2432种已知HLA-A等位基因中的一种或两种。HLA-A*02表示特定的HLA等位基因,其中字母A表示等位基因属于哪个HLA基因,而首码“*02首码”表示A2血清型。在MHC I类依赖性免疫反应中,肽不仅必须能与肿瘤细胞表达的某些MHC-I类分子结合,而且它们之后还必须能被负载特异性TCR的T细胞识别。

[0095] 本发明背景下的术语“靶肽”(TP)系指目的蛋白抗原(PAI)较短的肽、部分或片段。靶肽的氨基酸序列通常包含长度为8-12个氨基酸、长度为8-11个氨基酸或长度为8-10个氨基酸。优选为,靶肽氨基酸序列通常包含长度为8-11个氨基酸。靶肽可以与MHC-I分子或MHC-II分子结合。靶肽是否与MHC-I或MHC-II分子结合取决于靶肽的天然来源,即,靶肽是否在细胞质中合成、在蛋白酶体中加工或被内吞作用吸收并随后加工。另外,是否与MHC-I或MHC-II分子的结合槽结合取决于靶肽的长度。在一实施例中,长度为8-12、8-11或8-10个氨基酸的靶肽通常与MHC-I结合。在另一实施例中,靶肽的氨基酸序列可包含长度为13-23个氨基酸,优选为长度为13-18个氨基酸。长度为13-25或13-18个氨基酸的靶肽通常与MHC-II结合。

[0096] 本发明背景下的术语“抗原复合体”(AC)系指包含抗原的复合体,该抗原例如,直接或间接透过MHC或其肽结合部分附着至载体或可溶性多聚化MHC或其肽结合部分的表面。此类载体可以为细胞或合成材料。如果抗原附着至细胞,该细胞可以为抗原提呈细胞(APC),优选为人APC。载体的合成材料可以为珠或颗粒,优选为微珠、微粒或纳米颗粒。此类珠可以为磁珠或顺磁珠。珠或微粒通常由聚合物制成,可以用偶联残基对的第一成员共价或非共价包覆。偶联残基对中的第二成员与MHC或其肽结合部分共价或非共价偶联。优选的第一和第二偶联残基对包含链霉亲和素及生物素成员。技术人员了解其他偶联残基对。因此,在一优选实施方案中,可以用链霉亲和素包覆载体,这可固定包含生物素部分的MHC分子或其肽结合部分。相反,生物素包覆的载体可固定包含链霉亲和素部分的MHC分子或其肽结合部分。可溶性多聚化MHC或其肽结合部分可包含两个或更多MHC,其中每个MHC被共价或非共价(优选为共价)偶联至偶联残基对的第三成员和偶联残基的第四成员,其中第四成员具有至少两个第三成员的结合位点,优选为3、4、5、6、7或8个结合位点,特别优选为4个结合位点。生物素是偶联残基对的优选第三成员,链霉亲和素是偶联残基对的优选第四成员。链霉亲和素具有四个生物素的结合位点。因此,如果包含生物素的MHC肽复合体与链霉亲和素接触,将形成可溶性四聚体,其中四个肽载MHC(或其肽结合片段)与链霉亲和素非共价结合。因此,在一优选实施方案中,可溶性多聚化MHC或其肽结合片段为包含四个MHC肽复合体的复合体,其中每个MHC肽复合体共价附着至一个生物素,所述生物素又与链霉亲和素非共

价结合。

[0097] 术语“偶联残基对”系指以高亲和力彼此特异性和非共价结合的两个实体。优选情况为, K_d 小于 10^{-10} mol/L, 更优选为小于 10^{-11} mol/L, 更优选为小于 10^{-12} mol/L, 再更优选为小于 10^{-13} mol/L。优选情况为, 结合对其中至少一个成员的分子量低于 500g/mol。此类分子可共价附着于 MHC 或其肽结合片段的一条链, 而不会干扰 MHC 与 TCR 相互作用的能力。优选偶联残基对为生物素-链霉亲和素以及生物素-亲和素残基对。或者, 偶联残基对其中一个成员可以为与 MHC 一条链融合的蛋白质。实例包括几丁质结合蛋白 (CBP)、麦芽糖结合蛋白 (MBP)、链球菌标签谷胱甘肽 S-转移酶 (GST)、poly (His) 标签、V5 标签、Myc 标签、HA 标签、Spot 标签、T7 标签和 NE 标签。该对的另一成员由各自的蛋白质标签确定, 即, 几丁质、麦芽糖、生物素、谷胱甘肽、金属基质 (例如, Ni 基质) 或与 V5、Myc、HA、Spot、T7 或 NE 标签特异性结合的抗体。

[0098] 本发明背景下的术语“相似蛋白抗原” (SPA) 系指蛋白质或蛋白质的一部分或蛋白质复合物, 其包含由抗原结合蛋白互补位结合的表位。SPA 的氨基酸序列由 PAI 确定。SPA 的氨基酸序列与给定的 PAI (即, 目的 PAI) 的氨基酸序列至少有一个氨基酸不同。其目的是鉴定同时与 PAI 以及 SPA 结合的抗原结合蛋白, 即, 未表现出对 PAI 的预期特异性和/或选择性的抗原结合蛋白。此类抗原结合蛋白可引起肿瘤外/脱靶毒性。对于给定的 PAI 和抗原结合蛋白组合, SPA 为以下三类之一:

[0099] (1) 相似氨基酸序列, 同一表位:

[0100] 如果相对于 PAI 不同的 SPA 氨基酸不属于由给定的 PAI 特异性抗原结合蛋白结合的表位, 则该抗原结合蛋白将以相同的亲和力与 PAI 和 SPA 两者结合。具有这种性质的抗原结合蛋白将透过本发明的方法进行反向选择。

[0101] (2) 相似氨基酸序列, 相似表位:

[0102] 如果相对于 PAI 不同的至少一个 SPA 氨基酸属于由给定的 PAI 特异性抗原结合蛋白结合的表位, 则该抗原结合蛋白将以不同的亲和力与 PAI 和 SPA 结合。可以透过本发明的方法选择对 SPA 比对 PAI 显示明显更低结合力的抗原结合蛋白。在此方面, 明显更低的结合力表示: 在 PAI 和 SPA 浓度相同的情况下, 与 PAI 和 SPA 结合之间的差异为至少 2 倍、至少 3 倍、至少 4 倍、至少 5 倍、至少 6 倍、至少 7 倍、至少 8 倍、至少 9 倍、至少 10 倍、至少 15 倍、至少 20 倍、至少 30 倍、至少 40 倍、至少 50 倍、至少 70 倍、至少 100 倍、至少 200 倍, 优选为至少 50 倍, 更优选为至少 100 倍。

[0103] (3) 相似氨基酸序列, 不同表位:

[0104] 如果发散 (diverging) 氨基酸位于属于给定 PAI 特异性抗原结合蛋白结合的表位的位置, 则抗原结合蛋白可能根本不与 SPA 结合。具有这种性质的抗原结合蛋白将透过本发明的方法进行选择。

[0105] SPA 的氨基酸序列通常基于天然存在蛋白质的氨基酸序列, 因为此类蛋白质可在肿瘤患者的健康组织上表达。优选情况为, SPA 为天然存在蛋白质或其片段。特别是, SPA 与 PAI 存在于同一物种中。因此, 所期望的是, 本发明方法中包括的 SPA 具有可鉴定和反向选择类别 (1) 和 (2) 中抗原结合蛋白的氨基酸序列。如果 SPA 的氨基酸序列与 PAI 的氨基酸序列密切相关, 则 SPA 仅可能反向选择不合适的抗原结合蛋白。因此, 优选情况为, 本发明方法中使用的 SPA 的氨基酸序列与 PAI 的氨基酸序列具有至少 50%、至少 60%、至少 70%、至少 80%、

至少90%或至少95%的相似性。因此,在一优选实施方案中,SPA与PAI的氨基酸序列的不同为1-20个氨基酸,更优选为2-10个氨基酸。优选情况为本发明方法中使用的SPA,特别是靶相似肽(TSP)在健康组织上表达,优选为每细胞大于10个拷贝,优选为每细胞大于20个拷贝,优选为每细胞大于50个拷贝,优选为每细胞大于100个拷贝。表达的相对强度可透过多种本领域已知的方法来测定,包括用荧光标记的抗原结合蛋白对健康和患病细胞进行FACS分析或进行质谱分析。也可以使用RNA测序方法进行基因表达分析。选择本发明方法中使用的SPA另一标准为其在原发正常组织上的提呈频率。该频率是指SPA有多常提呈于正常组织(即,健康组织),有别于定义特定健康组织(例如,细胞)中SPA数量的拷贝数。拷贝数与该频率是针对给定PAI选择SPA的重要标准。与PAI的相似性越高并且在正常组织上的提呈频率和每细胞拷贝数(CpC)越高,SPA的相关性就越高。

[0106] 本发明背景下的术语“靶相似肽”(TSP)系指SPA的较短肽、部分或片段。TSP氨基酸序列通常包含长度为8至16个氨基酸。TSP通常由MHC提呈。与SPA相似,TSP与TP的氨基酸序列具有至少50%、至少60%、至少70%、至少80%、至少90%的相似性。TSP通常长度为8至16个氨基酸,例如8至12个氨基酸,优选为8至11个氨基酸。特别地,当TSP与MHC-I结合时,其长度为8至11或8-12个氨基酸。在另一实施例中,当TSP与MHC-II结合时通常长度为13至25个氨基酸。TSP与TP可能有一个或多个氨基酸不同,只要它们满足上面列出的相似性得分即可,并可按如下所述确定。因此,如果长度为8个氨基酸的TSP的氨基酸序列与给定的TP对齐,则可包含与TP相应氨基酸相似的1至8个氨基酸。其他氨基酸可能与TP相同或不同。因此,长度为9个氨基酸的TSP可包含与TP相应氨基酸相似的1至9个氨基酸;长度为10个氨基酸的TSP可包含与TP相应氨基酸相似的1至10个氨基酸;长度为11个氨基酸的TSP可包含与TP相应氨基酸相似的1至11个氨基酸;长度为12个氨基酸的TSP可包含与TP相应氨基酸相似的1至12个氨基酸。如果TSP与MHC II结合,则其长度通常为13至25个氨基酸,因此,长度为13个氨基酸的TSP可包含与TP相应氨基酸相似的1至13个氨基酸;长度为14个氨基酸的TSP可包含与TP相应氨基酸相似的1至14个氨基酸;长度为15个氨基酸的TSP可包含与TP相应氨基酸相似的1至15个氨基酸;长度为16个氨基酸的TSP可包含与TP相应氨基酸相似的1至16个氨基酸;长度为17个氨基酸的TSP可包含与TP相应氨基酸相似的1至17个氨基酸;长度为18个氨基酸的TSP可包含与TP相应氨基酸相似的1至18个氨基酸;长度为19个氨基酸的TSP可包含与TP相应氨基酸相似的1至19个氨基酸;长度为20个氨基酸的TSP可包含与TP相应氨基酸相似的1至20个氨基酸。选择本发明方法中使用的TSP另一标准为其在原发正常组织上的提呈频率。该提呈频率是指TSP有多常提呈于正常组织(即,健康组织),有别于定义给定健康组织不同样本上TSP数量的拷贝数,例如,如果12个脂肪组织样本中有6个样本检测到某个TSP,则其提呈频率为50%。给定TSP的提呈频率可透过本领域已知的方法来确定,包括实施例4中使用的MS分析(参见图6至图8,其显示了三种不同TSP的提呈频率)。因此,优选情况为,本发明的方法中使用TSP,其在至少一种健康组织中提呈频率为至少10%,优选为在至少一种健康组织中提呈频率为至少20%,更优选为在至少一种健康组织中提呈频率为至少30%。优选情况为,所选TSP存在于至少一种,优选为至少两种,更优选为至少三种健康组织中。这些组织优选地选自分别在图6和图7中就其TSP1和TSP2提呈进行分析的那些组织。

[0107] 拷贝数与该频率是针对给定TP选择TSP的重要标准。与TP的相似性越高并且在正常组织上的提呈频率和CpC越高,TSP的相关性就越高。

[0108] 本发明背景下的术语“无关抗原复合体” (IAC) 系指包含无关蛋白抗原 (IPA) 的AC。此类AC可以为:例如,APC或可溶的多聚化MHC肽复合体或固定于载体上的MHC肽复合体。无关蛋白抗原定义如下。

[0109] 本发明背景下的术语“无关蛋白抗原” (IPA) 系指未被所选TCR结合的蛋白抗原。针对TCR与其各自靶肽的结合进行筛选。TCR与其靶肽结合后,会触发所需免疫反应或T细胞介导免疫反应。无关肽不会触发此类所需免疫反应,因为在筛选过程中TCR不会与无关肽结合。IPA可以由管家基因编码的蛋白质。通常情况下,IPA与TP的氨基酸序列具有至少小于50%、至少小于40%、至少小于30%、至少小于20%、至少小于10%的相似性。

[0110] 本发明背景下的术语“无关肽” (IP) 系指IPA的较短肽、部分或片段。IP氨基酸序列通常包含长度为8至16个氨基酸。此类IP通常与MHC-I结合。在一些实施例中,当IP与MHC-II结合时,IP的长度可包含13-25个氨基酸。当IP与MHC-II结合时,其长度还可包含13-18个氨基酸。IP可由管家基因编码。

[0111] 本发明背景下的“管家基因”通常为组成性基因,是维持基本细胞功能所必需的,并在正常和病理生理条件下表达于生物体的所有细胞中。尽管在大多数非病理情况下某些管家基因以相对恒定的速率表达,但其他管家基因的表达可能根据实验条件而有所不同。管家基因在基因组中占活跃基因的大部分,其表达对生存显然至关重要。为了满足各种组织中的代谢要求,可对管家基因表达水准进行微调。管家基因的实例列举如下(非穷尽列举):转录因子、翻译因子、阻遏分子、RNA剪接分子、RNA结合蛋白、核糖体蛋白、线粒体核糖体蛋白、RNA聚合酶、蛋白质加工基因、热休克蛋白、组蛋白、细胞周期、细胞凋亡、致癌基因、DNA修复、DNA复制、代谢参与基因(例如,参与碳水化合物代谢、柠檬酸循环、脂质代谢、氨基酸代谢的基因)、NADH脱氢酶、细胞色素C氧化酶、ATP酶、溶酶体、蛋白酶体、核糖核酸酶、硫氧还原酶、受体、通道、转运蛋白、HLA/免疫球蛋白/细胞识别、激酶、细胞骨架、生长因子、肿瘤坏死因子 α 。与IPA相似,IP与TP的氨基酸序列具有至少小于50%、至少小于40%、至少小于30%、至少小于20%、至少小于10%,优选至少小于30%、至少小于20%、至少小于10%的相似性。

[0112] 本发明背景下的术语“氨基酸序列同一性”系指序列同一性的百分比,透过在比较视窗中比较两个最佳比对序列而确定,其中比较视窗中的序列部分与参考序列(不包含添加或删除)相比,可包含添加或删除(即,缺口),从而实现两个序列的最佳比对。该百分比的计算方法为:透过确定两个序列中出现相同核酸碱基或氨基酸残基的位置数以产生匹配位置数,将匹配位置数除以比较视窗中的位置总数,然后所得结果乘以100便可得出序列同一性的百分比。

[0113] 在两个或两个以上多肽或核酸序列背景下的术语“同一的”系指两个或多个序列或子序列相同,即包含相同序列的氨基酸或核酸。当在比较视窗或指定区域上进行比较和比对以获得最大对应性时(透过使用下列序列比较演算法之一或人工比对和目视检查测得),如果序列具有指定百分比的相同氨基酸残基(例如,在指定区域内具有至少70%、至少75%、至少80%、至少81%、至少82%、至少83%、至少84%、至少85%、至少86%、至少87%、至少88%、至少89%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%或至少99%的同一性),则序列相互之间“基本同一”。这些定义也指测试序列的补体。因此,关于多肽和多核苷酸序列的比较,在说明书通篇中用术语“至少

80%序列同一性”。此表达优选指与相应参考多肽或相应参考多核苷酸具有至少80%、至少81%、至少82%、至少83%、至少84%、至少85%、至少86%、至少87%、至少88%、至少89%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%或至少99%的序列同一性。

[0114] 本发明背景下的术语“序列比较”系指其中一个序列充当参考序列与测试序列进行比较的程序。当使用序列比较演算法时,将测试和参考序列输入电脑,必要时指定子序列座标,并指定序列演算法程序参数。通常使用默认程序参数,也可指定其他参数。然后,序列比较演算法根据程序参数算得测试序列相对于参考序列的序列同一性或相似性百分比。在比较两个序列并且未指定比较中的参考序列以供计算序列同一性百分比的情况下,若无特别另有说明,则参考两个待比较序列中较长的一个来计算序列同一性。如果指出了参考序列,则序列同一性基于SEQ ID所示的参考序列全长来确定(若无特别另有说明)。

[0115] 在序列比对中,术语“比较窗口”系指序列连续位置的区段,该区段与具有相同位置数的序列连续位置的参考区段相比。优选情况是,分别与SPA和TSP比对时,使用PAI的整个长度优选为TP作为比较窗口。如果TP是一个10个氨基酸长的MHC 1提呈肽,则在10个氨基酸的比较窗口中确定相似性。在这种情况下,有一个氨基酸错配的9个氨基酸长的SPA与给定TP具有80%的同一性。

[0116] 用于比较的“序列比对方法”为本领域所熟知。用于比较的序列最佳比对可透过,例如,Smith和Waterman的局部同源性演算法(Adv. Appl. Math. 2:482, 1970)、Needleman和Wunsch的同源性比对演算法(J. Mol. Biol. 48:443, 1970)、Pearson和Lipman的搜索相似性方法(Proc. Natl. Acad. Sci. USA 85:2444, 1988)、这些演算法的电脑化实现(例如, Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, Wis. 中的GAP、BESTFIT、FASTA和TFASTA)或手动比对和目视检查(例如,参见 Ausubel et al., Current Protocols in Molecular Biology (1995 supplement))进行。适用于确定百分比序列同一性和序列相似性的演算法有BLAST和BLAST 2.0演算法,描述分别参见Altschul et al. (Nuc. Acids Res. 25:3389-402, 1977)和Altschul et al. (J. Mol. Biol. 215:403-10, 1990)。进行BLAST分析的软体可透过国家生物技术资讯中心(<http://www.ncbi.nlm.nih.gov/>)公开获得。该演算法涉及首先透过识别查询序列中长度W的短词来识别高分序列对(HSP),这些短词在与数据库序列中相同长度的词比对时匹配或满足某个正值阈值分数T。T称为邻近词分数阈值(Altschul et al., 同上)。这些最初邻近命中词充当激活搜索以查找包含它们的较长HSP的种子。命中词沿着每个序列在两个方向延伸,直至累积比对分数可增加。对于核苷酸序列,使用参数M(匹配残基对的奖励分数;始终>0)和N(错配残基罚分;始终<0)来计算累积分数。对于氨基酸序列,使用得分矩阵来计算累积分数。在以下情况下,停止命中词在每个方向上延伸:累积比对分数比其最大实现值减少X;由于一个或多个负得分残基比对的累积,累积分数变为零或更低;或到达任一序列的末尾。BLAST演算法参数W、T和X确定比对的灵敏度和速度。BLASTN程序(用于核苷酸序列)使用词长(W) 11、期望值(E) 10、M=5、N=-4以及两条链的比较作为预设值。对于氨基酸序列, BLASTP程序使用词长3、期望值(E) 10以及BLOSUM62评分矩阵(参见Henikoff and Henikoff, Proc. Natl. Acad. Sci. USA 89:10915, 1989)比对(B) 50、期望值(E) 10、M=5、N=-4以及两条链的比较作为预设值。

[0117] 两个氨基酸序列的术语“相似性”考虑了给定位置上两个氨基酸的相关性(例如,参见下表4)。例如,TP和TSP中两个氨基酸序列的相似性可以使用BLAST演算法确定,其对两个氨基酸之间的相似性进行统计学分析(参见,例如Karlin and Altschul, Proc.Natl.Acad.Sci.USA 90:5873-87,1993)。此类比对的优选设置为:词长3、期望值(E)10、使用BLOSUM62或PMBEC评分矩阵(Kim et al.2009 BMC Bioinformatics),优选为确定相似时使用PMBEC评分矩阵。这些矩阵例如基于氨基酸之间的进化性或功能相似性来定量氨基酸相似性,其与根据理化参数的相似性有密切相关。对于给定TP序列中氨基酸的每次取代,可使用这些矩阵来计算得分(十进位值),从而表明TP序列中的氨基酸与TSP序列中被取代氨基酸的相似性。可透过加总TP序列中单次取代的效果(得分)来考虑多次取代。根据定义,TSP可实现的最大分数由未取代TP序列提供,而导致TSP的任何取代都会在得分矩阵中罚分,最终导致TSP的分数较低。但是,此最大分数取决于TP的长度和氨基酸序列(即,不同TP序列的最大分数将不同)。通常情况下,较长的氨基酸序列会导致较高的分数。但是,TP的分数取决于分配给其组成的氨基酸分数。为了能够在不考虑不同TP序列最大分数差异的情况下计算和比较TSP与TP的相似性,将TSP与TP的相似性计算结果之各十进位值进行转换为百分比分数,因此,TP序列的最大分数始终为100%。

[0118] BLAST演算法提供的另一种相似性度量指标为最小总和概率(P(N)),其表示氨基酸序列之间按机率发生匹配的概率。例如,如果测试氨基酸与参考氨基酸相比最小总和概率小于约0.2,通常小于约0.01,更通常小于约0.001,则认为氨基酸与参考序列相似。优选为半保守氨基酸取代,尤其是保守氨基酸取代,其中氨基酸被化学相关氨基酸取代。典型的取代为脂肪族氨基酸之中、具有脂肪族羟基侧链的氨基酸之中、具有酸性残基的氨基酸之中、酰胺衍生物之中、具有碱性残基的氨基酸之中或具有芳香族残基的氨基酸之中。下表4中列出了典型的半保守和保守取代。

氨基酸	保守取代	半保守取代
A	G; S; TN; V; C	
C	A; V; L	M; I; F; G
D	E; N; Q	A; S; T; K; R; H
E	D; Q; N	A; S; T; K; R;
F	W; Y; L; M; H	I; V; A
G	A	S; N; T; D; E; N; Q
H	Y; F; K; R	L; M; A
I	V; L; M; A	F; Y; W; G
K	R; H	D; E; N; Q; S; T; A

[0119]

L	M; I; V; A	F; Y; W; H; C
M	L; I; V; A	F; Y; W; C;
N	Q	D; E; S; T; A; G; K; R
P	V; I	L; A; M; W; Y; S; T; C; F
Q	N	D; E; A; S; T; L; M; K; R
[0120] R	K; H	N; Q; S; T; D; E; A
S	A; T; G; N	D; E; R; K
T	A; S; G; N; V D; E; R; K; I	
V	A; L; I M; T; C; N	
W	F; Y; H	L; M; I; V; C
Y	F; W; H	L; M; I; V; C

[0121] 表4:氨基酸以及分别的保守和半保守取代。

[0122] 如果新的半胱氨酸仍以游离硫醇形式存在,则从A、F、H、I、L、M、P、V、W或Y变为C为半保守的。此外,本领域技术人员将理解,空间需求位置上的甘氨酸不应被取代,并且不应将P引入具有 α -螺旋或 β -折叠结构的蛋白质部分中。

[0123] 本发明背景下的术语“可检测标记”系指透过因标记所发挥的特性或特定特征来标记不同分子或细胞而让此不同分子得以被选择的分子。适于标记的分子为蛋白质、DNA或RNA或合成材料(例如珠或其他合适材料)。有关蛋白质,标记策略可导致不同分子(包括生物素、报告酶、荧光团、磁标记和放射性同位素)共价连接至靶蛋白质或肽或核苷酸序列。可以使用短DNA或RNA条码“标签”来标记单细胞,以鉴定测序实验中源自同一细胞的读数。标记可以为荧光标记,例如:黄嘌呤类、吡啶类、恶嗪类、菁类、苯乙烯基染料、香豆素类、卟啉类、金属-配体络合物类、荧光蛋白质类、纳米晶体、花类和酞菁类、藻红蛋白(SA-PE)、链霉亲和素-藻蓝蛋白(SA-APC)或链霉亲和素亮紫421(SA-BV421);RNA条码或DNA条码或放射性标记。放射性标记通常为一个分子,其中一个或多个原子被放射性对应物(即,放射性同位素)取代。蛋白质、肽、DNA或RNA可以进行放射性标记。磁标记可包含磁珠或磁纳米颗粒,其可以用特定表面抗原的抗体包覆。磁标记可用于磁激活细胞分选(MACS)。

[0124] 本发明背景下的术语“可检测地不同”系指其中存在的两个标记仅在它们所发出的信号可不同的情况。例如,两个细胞可以用荧光标记进行标记,因此,不可透过标记本身特性(即,荧光)来区分。但是,附着于一个细胞上的荧光标记可能发出红色信号,而附着与第二个细胞上的荧光信号可能发出绿色信号。因此,示例性细胞的两个标记系可检测地不同。

[0125] 本发明背景下的术语“流式细胞术分析”系指一种分选技术,其包括测量样本中特定细胞群或细胞亚群的化学和物理性质。样本通常为悬浮液,并进行调整导致一次一个细胞的流动通过检测单元(通常为激发荧光团的雷射光束和光检测器)。检测到的信号,例如,由细胞流过而散射的光是细胞(即其成分)的特征。可透过此技术在短时间内分析多个细胞。流式细胞术的常规应用为细胞计数、细胞分选、细胞特征及功能确定、疾病(例如,癌症)

诊断、生物标志物检测或微生物检测。受欢迎的流式细胞仪技术为荧光激活细胞分选 (FACS)。FACS技术利用了荧光染料标签或标记来标记靶细胞的能力,从而可基于特定细胞群的个别标记谱对细胞进行分选。

[0126] 术语“磁激活细胞分选”(MACS)系指利用与特定细胞表面抗原相应抗体缀合的功能性微颗粒或纳米颗粒的分选技术。在施加磁场梯度的情况下,可以相对于所使用抗原为正向或负向方式分离磁靶向细胞。技术人员充分了解不同种类的分选分析。

[0127] 本发明背景下的术语“特异性结合”系指当靶标包含特异性和非特异性结合位点时,抗原结合蛋白或其片段(例如,抗体或其片段或TCR或其片段)与其靶标特异性结合位点结合。但是,有时不可避免的是蛋白质与紧密相关蛋白质结合,则与靶标实际结合可能具有特异性,但相对于预期靶标结合,蛋白质视为具有非特异性。如果在RT下表面等离子共振 (SPR)测得本发明的抗原结合蛋白或其片段以 10^{-5} M或更小的Kd与抗原结合,则其视为与给定抗原特异性结合。结合部分对特异性结合靶标的解离常数(Kd)比抗原结合蛋白或其片段的结合部分对不特异性结合靶标(例如,相似蛋白抗原(SPA),优选为靶相似肽(TSP))的解离常数(Kd)低至少2倍、至少5倍、至少7倍、10倍,优选为至少20倍,更优选为至少50倍,再优选为至少100倍、200倍、500倍或1000倍。

[0128] 通常情况下,如果与给定TP特异性结合的本发明抗原结合蛋白为TCR或其片段,则其Kd的范围为 3×10^{-5} 至 1×10^{-7} 、 2×10^{-5} 至 5×10^{-7} 、 1×10^{-5} 至 1×10^{-6} 或 5×10^{-6} 至 1×10^{-6} 。在这种情况下,优选情况为,抗原结合蛋白同时具有比抗原结合蛋白的结合部分对不特异性结合靶标(例如,TSP)的Kd低至少2倍、至少5倍、至少7倍、10倍,优选为至少20倍,更优选为至少50倍,再优选为至少100倍、200倍、500倍或1000倍的对TP的Kd。因此,例如,选定TCR可以以 1×10^{-6} 的KD与TP结合,以 1×10^{-5} 的KD与TSP结合。

[0129] 通常情况下,如果与给定TP特异性结合的本发明抗原结合蛋白为亲和力成熟的TCR或其片段或双特异性形式的可溶性分子或其片段,例如TCER®分子或其片段,则KD范围为 9×10^{-9} 至 1×10^{-12} 、 8×10^{-9} 至 5×10^{-12} 、 7×10^{-9} 至 1×10^{-11} 、 6×10^{-9} 至 2×10^{-11} 、 5×10^{-9} 至 5×10^{-11} 、 4×10^{-9} 至 8×10^{-11} 、 3×10^{-9} 至 1×10^{-10} 。在这种情况下,优选情况为,抗原结合蛋白同时具有比抗原结合蛋白的结合部分对不特异性结合靶标(例如,TSP)的Kd低至少2倍、至少5倍、至少7倍、10倍,优选为至少20倍,更优选为至少50倍,再优选为至少100倍、200倍、500倍或1000倍的对TP的Kd。本文中称为“TCER®”分子或“TCER®”的双特异性形式的分子通常包含与T细胞表面分子特异性结合的第一多肽和与MHC肽复合体特异性结合的第二多肽。

[0130] 通常情况下,如果本发明的抗原结合蛋白为抗体或其片段或与给定PAI特异性结合的B细胞,则KD范围为 9×10^{-9} 至 1×10^{-12} 、 8×10^{-9} 至 5×10^{-12} 、 7×10^{-9} 至 1×10^{-11} 、 6×10^{-9} 至 2×10^{-11} 、 5×10^{-9} 至 5×10^{-11} 、 4×10^{-9} 至 8×10^{-11} 、 3×10^{-9} 至 1×10^{-10} 。在这种情况下,优选情况为,抗原结合蛋白同时具有比抗原结合蛋白的结合部分对不特异性结合靶标(例如,SPA)的Kd低至少2倍、至少5倍、至少7倍、10倍,优选为至少20倍,更优选为至少50倍,再优选为至少100倍、200倍、500倍或1000倍的对PAI的Kd。

[0131] 在某些情况下,特别是TCR的背景下,如果本发明的抗原结合蛋白(特别是在TCR的背景下)与给定TP特异性结合可透过使用功能测定法来确定,例如,TCR激活测定法(如IFN γ 释放测定法)。因此,特异性结合可以透过反应来表征,例如针对TP检测到的信号大于30%、大于40%、大于50%、大于60%、大于70%、大于80%、大于90%、大于100%此类测定

法中针对TSP获得的反应(即,信号)。

[0132] 本发明背景下的术语“选择性结合”系指抗原结合蛋白(例如TCR或抗体)选择性地识别或结合至优选仅一个特异性表位的特征,优选为与另一个表位、肽或蛋白没有或基本上没有结合(无交叉反应性)。透过流式细胞术评估表位结合的阈值可透过使用非四聚体染色对照物进行评估。可根据非四聚体染色对照物设置门控,方式为该门控中出现<0.01%的细胞。该门控可应用于来自使用目的四聚体染色的同一供体的样本。出现于该门控中的细胞视为与目的表位选择性结合。

[0133] 本发明背景下的术语“T细胞受体”(TCR)系指免疫球蛋白超家族的异二聚体细胞表面蛋白,其与参与介导信号转导的CD3复合体的不变蛋白质相关。TCR以 $\alpha\beta$ 和 $\gamma\delta$ 形式存在,其结构相似,但解剖位置非常不同,也可能功能也非常不同。天然异二聚体 $\alpha\beta$ TCR和 $\gamma\delta$ TCR的细胞外部分各自均包含两个多肽,每个多肽都有一个膜近端恒定结构域和一个膜远端可变结构域。每个恒定和可变结构域都包括一个链内二硫键。可变结构域包含与抗体互补决定区(CDR)类似的高度多态性环。使用TCR基因疗法可克服目前的一些障碍。它可让受试者(患者)自己的T细胞在短时间内具有所需的特异性并产生足够数量的T细胞,从而避免其耗尽。TCR将被转导至有效的T细胞(例如中央记忆T细胞或具有干细胞特性的T细胞),这可确保转移后产生更好的持久性、保留性和功能。TCR工程化T细胞将被注入透过化学疗法或照射致使淋巴细胞减少的癌症患者,从而获得有效的植入,但抑制免疫抑制。天然 α - β 异二聚体TCR具有 α 链和 β 链。每个 α 链都包括可变、连接和恒定区域, β 链通常还包括可变区和连接区之间的短多样性区域,但是该多样性区域常被视为连接区域的一部分。TCR α 和 β 链的恒定区或C区分别称为TRAC和TRBC(Lefranc, (2001), Curr Protoc Immunol附录1:附录10)。每个可变区(本文称为 α 可变结构域和 β 可变结构域)都包括嵌入在框架序列中的三个互补决定区(CDR),其中一个称为CDR3的高度可变区。 α 可变结构域CDR称为CDRa1、CDRa2、CDRa3,而 β 可变结构域CDR称为CDRb1、CDRb2、CDRb3。存在几种类型的 α 链可变(V α)区和几种类型的 β 链可变(V β)区,由其框架、CDR1和CDR2序列及部分定义的CDR3序列区分。V α 类型透过唯一的TRAV号在IMGT命名中提及,V β 类型透过唯一的TRBV号在IMGT命名中提及(Folch and Lefranc, (2000), Exp Clin Immunogenet 17(1):42-54; Scaviner and Lefranc, (2000), Exp Clin Immunogenet 17(2):83-96; LeFranc and LeFranc, (2001), "T cell Receptor Factsbook", Academic Press)。关于免疫球蛋白抗体和TCR基因的更多资讯,参见国际免疫基因学资讯系统, Lefranc M-P等人(Nucleic Acids Res. 2015 Jan; 43(Database issue):D413-22; and <http://www.imgt.org/>)。常规的TCR抗原结合位点通常包括六个CDR,包含来自 α 和 β 链可变区的CDR组,其中CDR1和CDR3序列与MHC蛋白结合的肽抗原的识别和结合有关,而CDR2序列与MHC蛋白的识别和结合有关。与抗体类似,TCR包含的框架区为插入CDR之间的氨基酸序列,即,指不同TCR之间相对保守的TCR α 和 β 链可变区的那些部分。每个TCR的 α 和 β 链各有四个FR,本文中分别称为FR1-a、FR2-a、FR3-a、FR4-a和FR1-b、FR2-b、FR3-b、FR4-b。因此, α 链可变结构域可称为(FR1-a) - (CDRa1) - (FR2-1) - (CDRa2) - (FR3-a) - (CDRa3) - (FR4-a), β 链可变结构域可称为(FR1-b) - (CDRb1) - (FR2-b) - (CDRb2) - (FR3-b) - (CDRb3) - (FR4-b)。

[0134] “病毒或细菌引起的疾病”也可称为病毒或细菌感染。在本发明的背景下,导致疾病的病毒可选自例如人类免疫缺陷病毒(HIV)、人巨细胞病毒(HCMV)、巨细胞病毒(CMV)、人

乳头瘤病毒 (HPV)、乙型肝炎病毒 (HBV)、丙型肝炎病毒 (HCV)、人乳头瘤病毒感染 (HPV)、爱泼斯坦-巴尔病毒 (EBV)、流感病毒 (优选为人类免疫缺陷病毒 (HIV)) 构成的组。在本发明的背景下,引起疾病的细菌可为结核分枝杆菌。因此,这种细菌引起的疾病为结核病。本领域技术人员会理解,当抗原结合蛋白靶向作用于病毒抗原肽(例如HIV肽)时,则抗原结合蛋白可用于治疗HIV。因此,靶向作用于病毒或细菌抗原肽TA-C的抗原结合蛋白适合用于治疗所述抗原性病毒或细菌抗原肽的来源病毒或细菌。

[0135] 本发明背景下的术语“免疫疾病”系指免疫系统触发的疾病。术语“疾病”系指异常状况,特别是诸如生病或伤害之类的异常医疗状况,其中组织、器官或个体不再能够有效地执行其功能。相反,如果不存在异常状况并且组织、器官或个体无病理发现,则在本文中系指健康组织、器官或个体。在健康组织中,无细胞随机迁移,细胞在表征组织的结构上相互粘附,并帮助实现其功能。健康组织中不存在转移。通常情况下,疾病与表明该疾病存在的特定症状或体征相关,但这也不一定如此。因此,存在此类症状或体征可能提示有患病组织、器官或个体。这些症状或体征的改变可能提示此类疾病进展。疾病进展通常以此类症状或体征增加或减少为特征,其可能提示疾病“恶化”或“改善”。疾病“恶化”以组织、器官或生物体有效履行其功能的能力下降为特征,而疾病“改善”通常以组织、器官或器官有效履行其功能的能力增加为特征。有“患病风险”的组织、器官或个人处于健康状态,但显示有疾病发生的可能性。通常情况下,发生疾病的风险与此类疾病的早期或较弱体征或症状相关。在此类情况下,仍然可透过治疗来预防疾病发生。疾病的实例包括但不限于感染性疾病、外伤性疾病、炎症性疾病、皮肤病、内分泌疾病、肠道疾病、神经系统疾病、关节疾病、遗传疾病、自身免疫疾病和各种类型的癌症。本文所定义的健康组织通常包含健康细胞或由其组成。

[0136] 本发明背景下的术语“肿瘤性疾病”系指以细胞异常生长为特征的疾病,也称为肿瘤。肿瘤性疾病是导致肿瘤生长的疾病。恶性肿瘤为癌性疾病,可以缓慢或快速生长,并有转移或扩散至多个组织和器官的风险。“肿瘤”意指一组异常的细胞或组织,其透过快速、不受控制的细胞增殖而生长,并在引发新的生长的刺激因素停止后继续生长。肿瘤表现为与正常组织部分或完全缺乏结构组织和功能协调,通常形成明显的组织块,可为良性,也可为恶性。肿瘤性疾病可导致癌症,其中示例性癌症类型疾病包括但不限于基底细胞癌、膀胱癌、骨癌、脑瘤、乳腺癌、伯基特淋巴瘤、宫颈癌、结肠癌、皮肤T细胞淋巴瘤、食道癌、视网膜母细胞瘤、胃癌、胃肠道间质瘤、神经胶质瘤、霍奇金淋巴瘤、卡波西肉瘤、白血病、淋巴瘤、黑色素瘤、口咽癌、卵巢癌、胰腺癌、胸膜肺母细胞瘤、前列腺癌、喉癌、甲状腺癌和尿道癌。

[0137] 本发明的背景下的术语“治疗”系指对有需要(即,患有疾病或病症)的受试者的治疗性用途,意指逆转、减轻、抑制此类疾病、病症或状况的一种或多种症状进展。因此,治疗不仅指使疾病完全治愈的治疗,而且还指减缓疾病进展和/或延长受试者生存期的治疗。

[0138] 本发明背景下的术语“免疫细胞特异性表面标志物”系指细胞表面抗原,其用作帮助鉴定和分类免疫细胞的标志。表征不同T细胞亚型的此类标志物实例见上表1。大多数免疫细胞特异性表面标志物为细胞质膜内的分子或抗原。这些分子不仅作为标志物,而且还有关键的功能作用。

[0139] 术语“生长因子”或“分化因子”在本发明背景下可互换使用,系指能够刺激细胞生长、细胞增殖和细胞分化并调节多个细胞过程的分子。生长因子通常为蛋白质或类固醇激素。主要分子的实例列举如下(非穷尽列举):生长因子,例如:集落刺激因子(CSF)、巨噬细

胞集落刺激因子 (M-CSF)、粒细胞集落刺激因子 (G-CSF) 和粒细胞巨噬细胞集落刺激因子 (GM-CSF); 表皮生长因子 (EGF); 促红细胞生成素 (EPO); 成纤维细胞生长因子 (FGF); 胎牛生长激素 (FBS); 肝细胞生长因子 (HGF); 胰岛素; 胰岛素样生长因子 (IGF); 白介素; 神经调节蛋白; 神经营养因子; T细胞生长因子 (TCGF); 转化生长因子 (TGF); 肿瘤坏死因子 α (TNF α); 血管内皮生长因子 (VEGF)。

[0140] 在本发明背景中, 术语“核酸”系指去氧核糖核苷酸或核糖核苷酸碱基或两者的单链或双链寡聚物或聚合物。核苷酸单体由核碱基、五碳糖 (诸如, 但不限于核糖或2'-去氧核糖) 和一至三个磷酸基团组成。通常情况是, 核酸透过各核苷酸单体之间的磷酸二酯键形成。在本发明背景中, 术语核酸包括但不限于核糖核酸 (RNA) 和去氧核糖核酸 (DNA) 分子, 但也包括包含其他连接的合成形式核酸 (例如, 如Nielsen等人所描述的肽核酸 (Science 254:1497-1500, 1991)。通常情况是, 核酸为单链或双链分子, 并由天然存在核苷酸组成。核酸单链的描述也定义了 (至少部分地定义了) 互补链的序列。核酸可以为单链, 也可为双链, 或者可包含双链和单链序列的部分。典型的双链核酸分子可具有3'或5'突出端, 因此不需要或假设在其整个长度上完全为双链。核酸可透过生物学、生物化学或化学合成方法或本领域已知的任何方法获得, 包括但不限于扩增方法和RNA逆转录方法。术语核酸包括染色体或染色体片段、载体 (例如, 表达载体)、表达盒、裸露DNA或RNA聚合物、引物、探针、cDNA、基因组DNA、重组DNA、cRNA、mRNA、tRNA、微RNA (miRNA) 或小干扰RNA (siRNA)。核酸可以为: 例如, 单链、双链或三链的, 并且不限于任何特定长度。除非另有说明, 否则除了明确指出的任何序列外, 特定核酸序列包含或编码互补序列。

[0141] 在本发明背景中, 术语“载体”系指编码目的蛋白质的多核苷酸或包含多肽和编码目的蛋白质的多核苷酸的混合物, 其能够被引入细胞或能够将其中包含的蛋白质和/或核酸引入细胞。载体的实例包括但不限于质粒、粘粒、噬菌体、病毒或人工染色体。载体用于将目的基因产物 (例如, 外源或异源DNA) 引入宿主细胞。载体可能含有“复制子”多核苷酸序列, 其有助于在宿主细胞中自主复制载体。外源DNA定义为异源DNA, 其为宿主细胞中非天然存在的DNA, 例如, 其复制载体分子、编码可选择的或可筛选的标志物或编码转基因。载体一旦进入宿主细胞, 可以独立于宿主染色体DNA或与宿主染色体DNA同时复制, 并且可以产生数个拷贝的载体及其插入的DNA。此外, 载体还可能含有必需的要素, 其允许将插入的DNA转录为mRNA分子或以其他方式引起插入的DNA复制为多拷贝的RNA。载体可进一步包括调节目的基因表达的“表达控制序列”。表达控制序列通常为多肽或多核苷酸, 诸如激活子、增强子、沉寂子、绝缘子或抑制子。在包含编码一种或多种目的基因产物的多种多核苷酸的载体中, 所述表达可透过一种或多种表达控制序列一起或单独控制。更具体地, 载体上包含的每种多核苷酸可透过单独的表达控制序列控制, 或者载体上包含的所有多核苷酸可透过单一表达控制序列控制。由单一表达控制序列控制的单一载体上包含的多核苷酸可形成开放阅读框。一些表达载体额外含有与插入DNA相邻的序列元件, 其增加所表达mRNA的半衰期和/或允许mRNA翻译为蛋白质分子。因此可以快速合成由插入DNA编码的许多mRNA和多肽分子。此类载体可包含调节元件, 诸如激活子、增强子、终止子等, 以在受试者给药后引起或引导所述多肽的表达。用于动物细胞表达载体的激活子和增强子实例包括SV40的早期激活子和增强子 (Mizukami T. et al. 1987)、Moloney小鼠白血病病毒的LTR激活子和增强子 (Kuwana Y et al. 1987)、免疫球蛋白H链的激活子 (Mason JO et al. 1985) 和增强子 (Gillies SD

et al.1983)等。只要可以插入和表达编码人抗体C区的基因,即可使用任何动物细胞的表达载体。合适的载体实例包括pAGE107(Miyaji H et al.1990)、pAGE103(Mizukami T et al.1987)、pHSG274(Brady G et al.1984)、pKCR(O'Hare K et al.1981)、pSG1βd2-4-(Miyaji H et al.1990)等。质粒的其他实例包括包含复制起点的复制质粒或整合质粒(例如pUC、pcDNA、pBR)。

[0142] 本发明背景下的术语“抗原结合部分”或“抗原结合片段”系指分子,特别是氨基酸链,其长度较短,但保留了亲本蛋白的结合特异性和/或选择性,因为其仍然包含一个或多个负责亲本蛋白结合特异性和/或选择性的必需氨基酸序列。如果如下测定之“抗原结合部分”或“抗原结合片段”对亲本蛋白之靶标的Kd比亲本蛋白Kd至少高10倍或以下、高5倍或以下、高3倍或以下、高2倍或以下或相等,则认为其保留了结合特异性。例如,TCR的抗原结合片段为α和β链的可变结构域,抗体的抗原结合部分为可变轻链和重链。TCR、BCR或抗体的抗原结合部分是位于α和β或轻链和重链相应可变区的CDR。因此,可透过测量TCR对其靶肽或抗体对其抗原的结合亲和力来评估抗原结合蛋白(例如,抗体、TCR、BCR或其免疫功能部分或片段)的结合和/或特异性。本文使用的术语“片段”系指天然存在的片段(例如:剪接变体或肽片段)以及人工构建的片段,特别系指透过基因技术手段获得的片段。

[0143] 本发明背景下的术语“Kd”(测量单位为“mol/L”,有时缩写为“M”)系指结合部分(例如抗体或其片段)和靶分子(例如抗原或其表位)之间特定相互作用的解离平衡常数。亲和力可透过本领域已知的常见方法测量,包括但不限于基于表面等离子共振(SPR)的测定法(例如BIAcore测定法);生物层干涉法(BLI)、酶联免疫吸附测定法(ELISA);以及竞争性测定法(例如,放射免疫测定法(RIA))。低亲和力抗体通常与抗原缓慢结合并易于解离,而高亲和力抗体通常与抗原更快地结合并通常保持结合时间更长。测量结合亲和力的多种方法为本领域已知,其中任何一种方法均可用于本发明之目的。本公开书全文的各种抗原结合蛋白的Kd均在室温(即20℃)下透过SPR测量。

[0144] 术语“B细胞受体”(BCR)系指提呈于B细胞表面上具有抗原样结构的受体。B细胞在其第一次遇到与其受体结合的抗原(其“同源抗原”)时被激活,该细胞进行增殖并分化以生成抗体分泌血浆B细胞和记忆B细胞群。BCR透过生化信号传导和透过抗原提呈细胞从免疫突触中物理获得抗原来控制B细胞激活,在与抗原相互作用时具有两种关键功能。一种功能为信号转导,参与受体寡聚化的变化。第二种功能为介导内在化,以进行抗原的后续加工以及将肽提呈至辅助T细胞。识别抗原的BCR部分由三个不同的遗传区域(称为V、D和J)组成,它们在免疫系统独特的组合过程中在基因层级上被剪接和重组。形成I型跨膜受体蛋白的免疫球蛋白分子通常位于B淋巴细胞的外表面。从结构上来说,BCR包含透过二硫键连接的一种同种型(IgD、IgM、IgA、IgG或IgE)膜结合免疫球蛋白分子和信号转导部分:Ig-α/Ig-β(CD79)异二聚体。二聚体的各成员均跨过质膜,并有一条负载基于免疫受体酪氨酸激活基序(ITAM)的细胞质尾。

[0145] 本发明背景下的术语“抗体”系指分泌免疫球蛋白,其缺乏跨膜区域,因而可被释放至血流和体腔中。人源抗体根据其拥有的重链分组为不同的同种型。有五种类型的人Ig重链,用希腊字母表示为:α、γ、δ、ε和μ。存在的重链类型界定抗体类别,即,这些链分别存在于IgA、IgD、IgE、IgG和IgM抗体中,各抗体分别具有不同的作用,并指导针对不同类型抗原的相应免疫反应。不同的重链在大小和组成上各不相同;并可包含大约450个氨基酸

(Janeway et al. (2001) Immunobiology, Garland Science)。IgA存在于粘膜区域,例如,肠道、呼吸道、泌尿生殖道以及唾液、眼泪和母乳中,可防止病原体定殖(Underdown&Schiff (1986) Annu. Rev. Immunol. 4:389-417)。IgD主要充当尚未暴露于抗原的B细胞上的抗原受体,并参与激活嗜碱性粒细胞和masT细胞,以产生抗微生物因子(Geisberger et al. (2006) Immunology 118:429-437;Chen et al. (2009) Nat. Immunol. 10:889-898)。IgE透过与过敏原结合而参与过敏反应,从而触发masT细胞和嗜碱性粒细胞释放组胺。IgE还参与保护免于寄生虫侵害(Pier et al. (2004) Immunology, Infection, and Immunity, ASM Press)。IgG提供对抗入侵病原体的大多数基于抗体的免疫力,是唯一能够穿过胎盘让胎儿获得被动免疫力的抗体同种型(Pier et al. (2004) Immunology, Infection, and Immunity, ASM Press)。在人体中,有四种不同的IgG亚类(IgG1、2、3和4),以其在血清中的丰度顺序命名,其中IgG1丰度最高(~66%),其次为IgG2(~23%)、IgG3(~7%)和IgG(~4%)。不同IgG类别的生物学特征由各自铰链区的结构决定。IgM在B细胞表面以单体形式表达,并以亲和力非常高的五聚体形式分泌。在产生足量IgG之前,IgM在早期B细胞介导的(体液)免疫中参与病原体的消除(Geisberger et al. (2006) Immunology 118:429-437)。抗体不仅以单体存在,而且还已知可形成两个Ig单位的二聚体(例如IgA)、四个Ig单位的四聚体(例如,硬骨鱼IgM)或五个Ig单位的五聚体(例如,哺乳动物IgM)。抗体通常由包含两条相同重链和两条相同轻链的四个多肽链制成,它们透过二硫键连接并且类似于“Y”形大分子。每条链均包含许多免疫球蛋白结构域,其中一些为恒定结构域而其他为可变结构域。免疫球蛋白结构域由排列于两片的7至9条反向平行链的2层三明治结构组成。通常情况下,抗体重链包含四个Ig结构域,其中三个为恒定结构域(CH结构域:CH1、CH2、CH3),另一个为可变结构域(VH)。轻链通常包含一个恒定Ig结构域(CL)和一个可变Ig结构域(VL)。VH和VL区可进一步细分为高可变区,称为互补决定区(CDR),其间穿插称为框架区(FR)的更为保守区域。每个VH和VL由三个CDR和四个FR组成,按以下顺序从氨基端至羧基端排列:FR1、CDR1、FR2、CDR2、FR3、CDR3、FR4。重链和轻链的可变区包含与抗原相互作用的结合结构域。抗体的恒定区可介导免疫球蛋白与宿主组织或因子结合,所述宿主组织或因子包括免疫系统的各种细胞(例如,效应细胞)和经典补体系统的第一组分(C1q)。本文所用的术语抗体还包含嵌合抗体、人源化抗体或人抗体。

[0146] 本文所用的抗体、TCR或BCR或CAR的术语“抗原结合片段”(或“结合部分”或“片段”)系指保留与抗原特异性结合能力的抗体TCR、BCR或CAR的一个或多个片段。研究显示,可透过全长抗体、TCR、BCR或CAR的片段来执行抗体、TCR、BCR或CAR的抗原结合功能。术语“抗体、BCR或CAR的抗原结合部分”所包含的抗体结合片段实例包括(i) Fab片段,由VL、VH、CL和CH结构域组成的单价片段;(ii) F(ab')₂片段,包含在铰链区透过二硫键连接两个Fab片段的二价片段;(iii) 由VH和CH结构域组成的Fd片段;(iv) 由抗体单臂的VL和VH结构域组成的Fv片段;(v) 由VH结构域组成的dAb片段(Ward et al., (1989) Nature 341:544-546);(vi) 分离的互补决定区(CDR);以及(vii) 可任选透过合成连接子连接的两个或更多个分离CDR的组合。此外,尽管Fv片段的两个结构域VL和VH由独立的基因编码,但是它们可使用重组方法透过合成连接子连接,从而使它们成为一条蛋白链,其中VL区和VH区配对以形成单价分子(称为单链Fv(scFv);参见例如,Bird et al. (1988) Science 242:423-426;和Huston et al. (1988) Proc. Natl. Acad. Sci. USA 85:5879-5883)。此类单链抗体预期也包

含于抗体、BCR或CAR的术语“抗原结合片段”内。术语“TCR的抗原结合部分”包含至少TCR α 和 β 链的CDR1和CDR3,优选为 α 和 β 链的CDR1、CDR2和CDR3。尽管这些CDR优选为包含于其天然框架区背景中,但是它们也可包含于另一种蛋白质(所谓的蛋白质支架)中,所述蛋白质以位于 α 和/或 β 链的相似方式使它们放在一起。TCR的抗原结合部分优选为包含 α 和 β 链的可变结构域。抗体、TCR、BCR或CAR的抗原结合片段可包含于单体、二聚体、三聚体、四聚体或多聚体蛋白质复合体,从而赋予此类复合体一种或多种不同抗原结合特异性。此外,尽管Fv片段的两个结构域VL和VH由独立的基因编码,但是它们可使用重组方法透过合成连接子连接,从而使它们成为一条蛋白链,其中VL区和VH区配对以形成单价分子(称为单链Fv(scFv);参见例如,Bird et al.(1988) Science 242:423-426;和Huston et al.(1988) Proc.Natl.Acad.Sci.USA 85:5879-5883)。此类单链抗体预期也包含于术语“抗体的抗原结合片段”内。其中抗体的抗原结合片段用于产生单价、二价或多价结合分子的其他形式为本领域已知,称为双抗体、四抗体和纳米抗体。类似情况为,scFV的单链TCR在一条透过连接子连接的蛋白链上包含 α 和 β 链的可变结构域。

[0147] 本说明书中所使用的术语“受试者”系指在本文中可互换使用的“个体”、“受试者”或“患者”,并且指可以从本发明中受益的任何哺乳动物。特别是,“个体”为人。受试者可为健康受试者。

[0148] 本发明背景下的术语“有需要的受试者”系指患有或处于患病(例如,增生性疾病或病症、由病毒引起的疾病或由细菌引起的疾病)风险的受试者。此类增生性疾病或病症(例如,癌症)涉及细胞的失调和/或不适当增殖。增生性病症或疾病可为,例如,以肿瘤疾病的癌症或肿瘤细胞有TAA的表达(更特别是以TAA)为特征的所述肿瘤疾病。

[0149] 因此,特别优选的癌症为TA阳性癌症,特别是TAA阳性癌症。

[0150] 本发明的权利要求书和说明书的常用术语的缩写词:

[0151]	抗原复合体	AC
	B细胞受体	BCR
	嵌合抗原受体	CAR
	互补决定区	CDR
	荧光激活细胞分选	FACS
	人白细胞抗原	HLA
	无关抗原复合体	IAC
	无关肽	IP
	无关蛋白抗原	IPA
	磁激活细胞分选	MACS
	主要组织相容性复合体I/II	MHC I/II
	目的蛋白抗原	PAI
	相似蛋白抗原	SPA
	T细胞受体	TCR
	靶肽	TP
	靶相似肽	TSP
	肿瘤相关抗原	TAA

[0152] 表5:常用术语的缩写词

[0153] 实施方案

[0154] 下文将更详细地定义本发明的不同方面。如此定义的各方面可与任何其他一个或多个方面组合,除非明确指出不可。特别是,提示为优选或有利的任何特征可与提示为优选或有利的任何其他一个或多个特征组合。

[0155] 免疫治疗构成令人兴奋且迅速演变的领域,且证明基因修饰的T细胞受体(TCR)可用于产生所需特异性的T淋巴细胞群,为抗原特异性T细胞治疗提供了新的机会。总体而言,TCR修饰的T细胞具有透过T细胞的正常生物学靶向作用于广泛的自身和非自身靶标的能力。但是,“肿瘤外/靶点”或“肿瘤外/脱靶”效应可导致免疫相关毒性的极大不良作用。透过在鉴定TCR阶段已纳入的相似蛋白抗原(相似肽),发明人能够在表征过程之前排除一部分交叉反应性T细胞,从而提高整个TCR发现程序的效率。为此目的,染色组中纳入1D标记或2D标记相似肽多聚体。内部数据库可让发明人鉴定正常人组织上存在的高度相关靶序列相似肽。如果假定用于临床用途的TCR识别到此类靶序列相似肽,则代表安全性风险。因此,发明人开发出了一种内部搜索演算法,将靶相似肽的公共和内部基因组数据库搜索与来自内部数据库的健康组织上实际MS检测到的肽结果结合。发明人在TCR鉴定期间早期使用了一组靶相似肽,使得能够早期取消选择(de-selection)交叉反应性TCR。为此目的,针对靶肽和靶相似肽都产生了荧光染料(链霉亲和素)标记的肽主要组织相容性复合体(pMHC)四聚体,可透过至少一种不同的荧光染料区分。将对靶标和相似肽均呈阳性的细胞排除出下游TCR鉴定的T细胞分选。

[0156] 这一令人意外的发现尤其提供了以下与本领域相比的优点:(i)降低所选TCR与健康组织上相似肽的交叉反应性;(ii)提高所选TCR的安全性特征;(iii)因早期选择了靶标特异性TCR,因而可有效、快速鉴定和表征TCR;(v)分选过程中排除相似肽结合实现特异性TCR选择;(vi)TCR发挥降低的脱靶和肿瘤外细胞毒性作用;以及(vii)提供了改进的特异性、选择性和安全TCR。

[0157] 本发明的第一方面涉及一种选择细胞或病毒的方法,所述细胞或病毒在其表面表达与目的蛋白抗原(PAI)特异性和/或选择性结合的抗原结合蛋白,所述方法包括以下步骤:

[0158] (i)提出细胞群或病毒群;

[0159] (ii)使步骤(i)的细胞群或病毒群与包含PAI和可检测标记A的第一抗原复合体(第1 AC)或包含可检测标记A的PAI接触;

[0160] (iii)使步骤(i)的细胞群或病毒群与包含相似蛋白抗原(SPA)的至少第二抗原复合体(第2 AC)接触,其中SPA的氨基酸序列与PAI的氨基酸序列至少有1个氨基酸不同,其中第2 AC包含可检测标记B;或与SPA和可检测标记B接触;以及

[0161] (iv)选择至少一种与第一AC特异性和/或选择性结合的细胞或病毒,

[0162] 其中,可检测标记A和可检测标记B彼此系可检测地不同。

[0163] 在本发明第一方面的一实施方案中,细胞基于反向选择原理进行选择:优选表达于患病组织上的肿瘤抗原(TP)的MHC提呈短肽用可检测标记物标记,表达于健康组织上的含相似序列的肽(TSP)用可检测标记物标记。这种方法中使用的标记系可检测地不同。在让细胞(优选为T细胞)与TP和TSP接触时,所述细胞与TP、TSP其中之一结合或与TP和TSP两者

结合或两者均不结合,且基于正向选择标准或负向选择标准进行选择。在传统分选方法中,可鉴定具有相应可检测细胞信号的以下细胞:可透过检测TP标记(即,细胞与MHC提呈肽结合情况下的目的肽)的信号检测到一种细胞。另一种细胞可透过检测TP标记(即,免疫细胞与MHC提呈肽结合情况下的目的肽)和透过检测TSP标记来发出信号。第三种细胞只能透过检测TSP标记发出信号。正向选择的仅是TP标记可检测到的细胞,因为如果所述细胞为T细胞,那么该细胞便是含有能够与MHC提呈肽结合之目的TCR的细胞。反向选择(或负向选择)描述如下:由两个标记(即TP和TSP标记)检测到的细胞被负向选择,因为这些细胞除了与TP结合外还与TSP结合。仅透过TSP标记检测到的细胞也被负向选择,因为它们仅与健康组织上表达的TSP结合,并且通常应避免与TSP结合以减少脱靶效应。通常情况下,给定细胞与给定PAI(优选为TP)和一个或多个SPA(优选为TSP)的结合并非全部结合或完全不结合。因此,选择也可基于PAI(优选为TP)与SPA(优选为TSP)结合的相对差异。在PAI和SPA浓度相同的情况下,若细胞与PAI(优选为TP)的结合力强至少2倍、至少3倍、至少4倍、至少5倍、至少6倍、至少7倍、至少8倍、至少9倍、至少10倍、至少15倍、至少20倍、至少30倍、至少40倍、至少50倍、至少70倍、至少100倍、至少200倍,则视它们为特异性结合。优选情况为,与PAI(优选为TP)的结合比本发明方法中使用的任意一种或多种SPA(优选为TSP)强至少10倍。再优选为,结合强至少20倍,更优选为至少50倍。给定细胞(例如,T细胞或其TCR或B细胞)的结合强度或亲和力可透过多种测定法确定,通常表示为TCR的解离常数(Kd)。但是,就本发明方法之目的而言,无需精确测定所述细胞与给定PAI(优选为TP)和一个或多个SPA(优选为TSP)的Kd。测定相对结合强度即已足够,例如,可透过细胞FACS分析测定,其中将给定TP的荧光信号与一个或多个TSP的荧光信号进行比较。在此类测定中,必须考虑PAI(优选为TP)和一种或多种SPA(优选为TSP)的相对摩尔量。如果TP和TSP以相同摩尔量被添加入细胞中,并发现所使用各标记物的荧光强度存在差异,则当等摩尔量的SPA(优选为TSP)显示归因于PAI(优选为TP)的荧光至少强于SPA(优选为TSP)10倍时,细胞与PAI(优选为TP)特异性结合。本领域技术人员理解,在据此透过调整门控选择细胞时,可考虑步骤(ii)和(iii)中与细胞群接触的PAI(优选为TP)和SPA(优选为TSP)的摩尔比变化。

[0164] 在本发明第一方面的一实施方案中,选定细胞为表达异源抗原结合蛋白的哺乳动物细胞或表达异源抗原结合蛋白的酵母细胞。哺乳动物细胞可为任何哺乳动物细胞,例如,人细胞、小鼠细胞(优选为人源化小鼠细胞)、大鼠细胞、猪细胞、猴细胞或犬细胞。通常情况下,哺乳动物细胞可为任何抗原提呈细胞(APC)。优选情况为,哺乳动物细胞为人细胞。特别地,将哺乳动物细胞改造为表达异源抗原结合蛋白,例如,TCR或其片段、或BCR或其片段、或CAR或其片段、或抗体或其片段。如果所选择细胞为表达异源抗原结合蛋白的酵母细胞,则优选此类酵母细胞为酿酒酵母细胞。特别地,将酵母细胞改造为表达异源抗原结合蛋白,例如,TCR或其片段、或BCR或其片段、或CAR或其片段、或抗体或其片段。

[0165] 在本发明第一方面的另一实施方案中,所述方法选择了病毒。所选病毒可以为选自由双链DNA病毒,优选为肌病毒科、长尾病毒科、短尾病毒科、疱疹病毒科、腺病毒科、杆状病毒科、乳头状瘤病毒科、多分体DNA病毒科、多瘤病毒科、痘病毒科;单链DNA病毒,优选为指环病毒科、丝杆病毒科、细小病毒科;双链RNA病毒,优选为呼肠孤病毒科;单链RNA病毒,优选为冠状病毒科、微RNA病毒科、杯状病毒科、披膜病毒科、黄病毒科、星状病毒科、动脉炎病毒科、戊型肝炎病毒科;负义单链RNA病毒,优选为沙粒病毒科、丝状病毒科、副粘病毒科、

弹状病毒科、布尼亚病毒科、正粘病毒科、玻那病毒科；单链RNA逆转录病毒，优选为逆转录病毒科；或双链RNA逆转录病毒，优选为花椰菜花叶病毒科、嗜肝DNA病毒科组成的组中任何的病毒。更优选为，选定的病毒为噬菌体。噬菌体优选为选自由噬菌体T4 λ (T4 λ) 噬菌体、T7噬菌体、fd丝状噬菌体(优选为丝状噬菌体M13)组成的组。所选病毒(例如噬菌体)可与珠(例如适合于顺序磁分选的磁珠)结合。在此实施方案中，优选情况为，标记(例如标记A和标记B)为条码标记，优选为RNA条码或DNA条码，如前文所述。

[0166] 在本发明第一方面的另一实施方案中，用于选择细胞的方法包括步骤(i)中提供细胞群。优选情况为，细胞群包含真核细胞。更优选的情况为，真核细胞为表达异源抗原结合蛋白文库的哺乳动物细胞或表达异源抗原结合蛋白文库的酵母细胞。因此，本发明第一方面的方法可用于酵母展示。

[0167] 在本发明第一方面的另一实施方案中，用于选择病毒的方法包括步骤(i)中提供病毒群。优选情况为，病毒群包含表达异源抗原结合蛋白文库的病毒。在一优选实施方案中，病毒群包含噬菌体，因此，本发明第一方面的方法可用于噬菌体展示。

[0168] 本发明第一方面的另一实施方案涉及一种选择免疫细胞的方法，所述细胞在其表面表达与目的蛋白抗原(PAI)特异性和/或选择性结合的抗原结合蛋白，所述方法包括以下步骤：

[0169] (i) 提出包含免疫细胞的细胞群；

[0170] (ii) 使步骤(i)的细胞群与包含PAI和可检测标记A的第一抗原复合体(第1AC)或包含可检测标记A的PAI接触；

[0171] (iii) 使步骤(i)的细胞群与包含相似蛋白抗原(SPA)的至少第二抗原复合体(第2AC)接触，其中SPA的氨基酸序列与PAI的氨基酸序列至少有1个氨基酸不同，其中第2AC包含可检测标记B；或与SPA和可检测标记B接触；以及

[0172] (iv) 选择至少一种与第一AC特异性和/或选择性结合的免疫细胞，

[0173] 其中，可检测标记A和可检测标记B彼此系可检测地不同。

[0174] 在本发明第一方面的一实施方案中，免疫细胞基于反向选择原理进行选择：优选表达于患病组织上的肿瘤抗原(TP)的MHC提呈短肽用可检测标记物标记，表达于健康组织上的含相似序列的肽(TSP)用可检测标记物标记。这种方法中使用的标记系可检测地不同。在让免疫细胞(优选为T细胞)与TP和TSP接触时，所述免疫细胞与TP、TSP其中之一结合或与TP和TSP两者结合或两者均不结合，且基于正向选择标准或负向选择标准进行选择。在传统分选方法中，可鉴定具有相应可检测细胞信号的以下细胞：可透过检测TP标记(即，免疫细胞与MHC提呈肽结合情况下的目的肽)的信号检测到一种细胞。另一种细胞可透过检测TP标记(即，免疫细胞与MHC提呈肽结合情况下的目的肽)和透过检测TSP标记来发出信号。第三种细胞只能透过检测TSP标记发出信号。正向选择的仅是TP标记可检测到的细胞，因为如果所述免疫细胞为T细胞，那么该细胞便是含有能够与MHC提呈肽结合之目的TCR的细胞。反向选择(或负向选择)描述如下：由两个标记(即TP和TSP标记)检测到的细胞被负向选择，因为这些细胞除了与TP结合外还与TSP结合。仅透过TSP标记检测到的细胞也被负向选择，因为它们仅与健康组织上表达的TSP结合，并且通常应避免与TSP结合以减少脱靶效应。通常情况下，给定免疫细胞与给定PAI(优选为TP)和一个或多个SPA(优选为TSP)的结合并非全部结合或完全不结合。因此，选择也可基于PAI(优选为TP)与SPA(优选为TSP)结合的相对差

异。在PAI和SPA浓度相同的情况下,若免疫细胞与PAI(优选为TP)的结合力强至少2倍、至少3倍、至少4倍、至少5倍、至少6倍、至少7倍、至少8倍、至少9倍、至少10倍、至少15倍、至少20倍、至少30倍、至少40倍、至少50倍、至少70倍、至少100倍、至少200倍,则视它们为特异性结合。优选情况为,与PAI(优选为TP)的结合比本发明方法中使用的任意一种或多种SPA(优选为TSP)强至少10倍。再优选为,结合强至少20倍,更优选为至少50倍。给定免疫细胞(例如,T细胞或其TCR)的结合强度或亲和力可透过多种测定法确定,通常表示为TCR的解离常数(Kd)。但是,就本发明方法之目的而言,无需精确测定所述免疫细胞与给定PAI(优选为TP)和一个或多个SPA(优选为TSP)的Kd。测定相对结合强度即已足够,例如,可透过免疫细胞FACS分析测定,其中将给定TP的荧光信号与一个或多个TSP的荧光信号进行比较。在此类测定中,必须考虑PAI(优选为TP)和一种或多种SPA(优选为TSP)的相对摩尔量。如果TP和TSP以相同摩尔量被添加入免疫细胞中,并发现所使用各标记物的荧光强度存在差异,则当等摩尔量的SPA(优选为TSP)显示归因于PAI(优选为TP)的荧光至少强于SPA(优选为TSP)10倍时,免疫细胞与PAI(优选为TP)特异性结合。本领域技术人员理解,在据此透过调整门控选择细胞时,可考虑步骤(ii)和(iii)中与细胞群接触的PAI(优选为TP)和SPA(优选为TSP)的摩尔比变化。

[0175] 在一实施方案中,让步骤(i)的细胞群与包含SPA(优选为TSP)的至少第二抗原复合体(第2 AC)、包含SPA(优选为TSP)的至少第三抗原复合体(第3 AC)、包含SPA(优选为TSP)的至少第四抗原复合体(第4 AC)、包含SPA(优选为TSP)的至少第五抗原复合体(第5 AC)、包含SPA(优选为TSP)的至少第六抗原复合体(第6 AC)、包含SPA(优选为TSP)的至少第七抗原复合体(第7 AC)、包含SPA(优选为TSP)的至少第八抗原复合体(第8 AC)、包含SPA(优选为TSP)的至少第九抗原复合体(第9 AC)、包含SPA(优选为TSP)的至少第十抗原复合体(第10 AC)接触。因此,优选情况为使用不超过十个SPA(优选为TSP)、使用不超过九个SPA(优选为TSP)、使用不超过八个SPA(优选为TSP)、使用不超过七个SPA(优选为TSP)、使用不超过六个SPA(优选为TSP)、使用不超过五个SPA(优选为TSP)、使用不超过四个SPA(优选为TSP)、使用不超过三个SPA(优选为TSP)、使用不超过两个SPA(优选为TSP)或使用不超过一个SPA(优选为一个TSP)。因此,优选情况为,在本发明的方法中,优选使用1-10个不同TSP,优选为2-8个不同TSP,更优选为3-5个不同TSP。在一再优选实施方案中,使用三个TSP。

[0176] 在一实施方案中,本发明第一方面的方法选择了免疫细胞。优选情况为,此免疫细胞为T细胞或B细胞。更优选的情况为,T细胞为CD4 T细胞。在一再优选实施方案中,T细胞为CD8 T细胞。信号传导结构域优选包含CD3。在本发明第一方面的另一实施方案中,免疫细胞在其表面表达抗原结合蛋白。优选情况为,如果要选择的免疫细胞为T细胞,则抗原结合蛋白为TCR或其抗原结合片段。优选情况为,如果要选择的免疫细胞为B细胞,则抗原结合蛋白为BCR或其抗原结合片段。此类抗原识别或抗原结合位点优选为单链可变片段(scFv),并优选靶向作用于系为TAA的PAI。共刺激结构域优选包含CD28或4-1BB。信号传导结构域优选包含CD3。还优选为,如果要选择的免疫细胞为T细胞,则抗原结合蛋白为CAR或其抗原结合片段。

[0177] 在本发明第一方面的另一实施方案中,用于选择免疫细胞的方法包括步骤(i)中提供包含免疫细胞的细胞群。包含免疫细胞的细胞群源自健康受试者的外周血。在另一实施方案中,包含免疫细胞的细胞群源自患病受试者的外周血。优选情况为,包含免疫细胞的

细胞群源自健康或患病受试者的外周血中富含免疫细胞的级分。优选情况为,富含免疫细胞的级分富含干细胞、T细胞、B细胞或浆细胞。再优选情况为,免疫富集级分富含CD4 T细胞和/或CD8 T细胞。在另一实施方案中,包含免疫细胞的细胞群可源自肿瘤浸润淋巴细胞(TIL)或TCR文库。优选情况为,TCR文库包含大量不同的T细胞受体(TCR)蛋白或其片段,其中每个TCR蛋白或其片段均不同。

[0178] 在本发明第一方面的另一实施方案中,透过可检测标记一个或多个免疫细胞特异性表面标志物来选择免疫细胞富集级分中的一个或多个免疫细胞。优选情况为,免疫细胞表面标志物选自由CD3、CD8、CD4和CD19组成的组。

[0179] 在本发明第一方面的另一实施方案中,步骤(i)中的细胞群可在生长因子和/或细胞因子存在的情况下在进一步的步骤中进行孵育。优选情况为,细胞因子为白介素。更优选的情况为,白介素选自由IL-1、IL-2、IL-7、IL-10、IL-12、IL-15、IL-17、IL-21和IL-23组成的组。最优选的情况为,步骤(i)的细胞群用IL-2、IL-7、IL-15和/或IL-21孵育。

[0180] 在另一实施方案中,目的蛋白抗原(PAI)为肿瘤相关抗原(TAA)、病毒蛋白或细菌蛋白。优选情况为,如果PAI为靶肽,即PAI的较短片段,则靶肽为病毒抗原肽或细菌抗原肽。在本发明第一方面的另一实施方案中,患病受试者患有选自由免疫疾病、肿瘤性疾病、病毒引起的疾病或细菌引起的疾病组成组中的疾病。优选情况为,肿瘤性疾病为癌症。优选情况为,病毒引起的疾病为病毒感染;细菌引起的疾病为细菌感染。优选情况为,病毒感染由选自由HIV、HCMV、CMV、HPV、HBV、HCV、HPV、EBV、流感病毒组成组中的病毒引起。更优选情况为,病毒感染由HIV引起。优选情况为,细菌感染由结核分枝杆菌引起。此类疾病为结核病。

[0181] 在第一方面的另一优选实施方案中,所述方法包括在步骤(ii)中使步骤(i)的细胞群与包含PAI和可检测标记A的第一抗原复合物(第1 AC)或与PAI和可检测标记A接触。包含PAI和标记A的第一AC优选为抗原提呈细胞。在另一优选实施方案中,第一AC为包含颗粒、PAI和可检测标记A的复合物。更优选情况为,所述颗粒为纳米珠或微珠。还优选的情况为,MHC分子连接至纳米珠或微珠。在另一优选实施方案中,第一AC由PAI和可检测标记A组成。在另一实施方案中,第一AC为包含颗粒、SPA和可检测标记B的复合物。更优选情况为,所述颗粒为纳米珠或微珠。还优选的情况为,MHC分子连接至纳米珠或微珠。在另一优选实施方案中,第一AC由SPA和可检测标记B组成。

[0182] 或者,PAI可包含可检测标记。如果PAI为氨基酸链且标记与此氨基酸链共价连接,则为优选实施方案。实例有荧光标记或荧光蛋白(如GFP或EGFP)。在后一种情况下,优选为荧光蛋白透过肽键与PAI连接。

[0183] 在另一优选实施方案中,本发明第一方面的方法进一步包括以下步骤中的一个或多个步骤:使细胞群与包含另外的PAI和另外的可检测标记的另外AC以及包括另外的SPA和另外的可检测标记的另外AC接触。使用包含另外的PAI和另外的可检测标记C的第三AC和包含另外的SPA和另外的可检测标记D的第四AC反映了所谓的2D多聚体多工(2DMM)方法,其可用高灵敏度(0.0001%)以高度节省细胞的方式进行特定的罕见细胞检测。两种针对每种特异性用不同荧光染料标记的多聚体(肽MHC)可透过使用9种不同的荧光染料鉴定一个样本中最多36种不同的特异性。在一优选实施方案中,使步骤(i)的细胞群与包含PAI和可检测标记C的第三抗原复合物(第3 AC)接触,所述可检测标记C系可检测地不同于与细胞群接触的其他AC的一个或多个或所有其他可检测标记。优选情况为,标记可检测地不同于至少可

检测标记A,优选为可检测地不同于至少可检测标记A和可检测标记D(如果存在可检测标记D)。如果将步骤(i)的细胞群与包含PAI和可检测标记D的第四抗原复合体(第4 AC)接触,则存在可检测标记D,所述可检测标记D系可检测地不同于与细胞群接触的其他AC的一个或多个或所有其他可检测标记。可检测标记物D优选为可检测地不同于至少可检测标记A。还优选为,可检测标记D系可检测地不同于至少可检测标记A和可检测标记C。在一实施方案中,使步骤(i)的细胞群与包含SPA和可检测标记E的第五抗原复合体(第5 AC)接触,所述可检测标记E系可检测地不同于与细胞群接触的其他AC的一个或多个或所有其他可检测标记。优选情况为,标记物E可检测地不同于至少可检测标记物B。还优选为,可检测标记物E系可检测地不同于至少可检测标记B和可检测标记物F(6)如果存在可检测标记F)。如果将步骤(i)的细胞群进一步与包含SPA和可检测标记F的第六抗原复合体(第6 AC)接触,则存在可检测标记F,所述可检测标记F系可检测地不同于与细胞群接触的其他AC的一个或多个或所有其他可检测标记。优选情况为,标记物F可检测地不同于至少可检测标记物B。还优选为,可检测标记物F系可检测地不同于至少可检测标记B和可检测标记物E。

[0184] 在另一实施方案中,使步骤(i)的细胞群与包含PAI(优选为TP)和可检测标记A的第一抗原复合体(第1 AC)接触;与包含相似蛋白抗原(SPA)(优选为TSP)的至少第二抗原复合体(第2 AC)接触,其中SPA(优选为TSP)的氨基酸序列与PAI(优选为TP)的氨基酸序列至少有1个氨基酸不同,其中,第2 AC包含可检测地不同于标记A的可检测标记B;与一至十个,即一、二、三、四、五、六、七、八、九或十个,优选为二至四个,最优选为两个另外抗原复合体(AC)接触,其中每个AC包含不同的SPA(优选为不同的TSP),该SPA之至少一个氨基酸序列与第2 AC的SPA氨基酸序列不同,并且其中每个另外AC包含一个或多个标记,其中一个或多个标记系可检测地不同于第2 AC的一个或多个标记。优选情况为,第1 AC包含另外的第二标记,其与第2 AC的一个或多个标记以及另外AC的一个或多个标记系可检测地不同。还优选为细胞群为T细胞群。进一步优选为,与第1 AC特异性和/或选择性结合的所选免疫细胞为T细胞。还优选为,在所选T细胞表面上与PAI特异性和/或选择性结合的抗原结合蛋白为TCR。

[0185] 在另一实施方案中,使步骤(i)的细胞群与包含PAI(优选为TP)和可检测标记A的第一抗原复合体(第1 AC)接触;与包含相似蛋白抗原(SPA)(优选为TSP)的至少第二抗原复合体(第2 AC)接触,其中SPA(优选为TSP)的氨基酸序列与PAI(优选为TP)的氨基酸序列至少有1个氨基酸不同,其中,第2 AC包含可检测标记B;与一至十个,即一、二、三、四、五、六、七、八、九或十个,优选为二至四个,最优选为两个另外抗原复合体(AC)接触,其中每个AC包含不同的SPA(优选为不同的TSP),该SPA之至少一个氨基酸序列与第2 AC的SPA氨基酸序列不同,并且其中每个另外AC包含一个或多个标记,其中一个或多个标记与第2 AC的一个或多个标记相同。还优选为细胞群为T细胞群。进一步优选为,与第1 AC特异性和/或选择性结合的所选免疫细胞为T细胞。还优选为,在所选T细胞表面上与PAI特异性和/或选择性结合的抗原结合蛋白为TCR。

[0186] 在另一实施方案中,使步骤(i)的细胞群与包含PAI(优选为TP)和可检测标记A的第一抗原复合体(第1 AC)接触;与包含相似蛋白抗原(SPA)(优选为TSP)的至少第二抗原复合体(第2 AC)接触,其中SPA(优选为TSP)的氨基酸序列与PAI(优选为TP)的氨基酸序列至少有1个氨基酸不同,其中,第2 AC包含可检测地不同于标记A的可检测标记B,且第1 AC包含至少一种另外的可检测标记,第2 AC包含至少一种另外的可检测标记,两者相同。优选为

细胞群为T细胞群。进一步优选为,与第1 AC特异性和/或选择性结合的所选免疫细胞为T细胞。还优选为,在所选T细胞表面上与PAI特异性和/或选择性结合的抗原结合蛋白为TCR。

[0187] 在另一实施方案中,使步骤(i)的细胞群与包含PAI(优选为TP)和可检测标记A的第一抗原复合体(第1 AC)接触;与包含SPA(优选为TSP)的至少第二抗原复合体(第2 AC)接触,其中SPA(优选为TSP)的氨基酸序列与PAI(优选为TP)的氨基酸序列至少有1个氨基酸不同,其中,第2 AC包含可检测地不同于标记A的可检测标记B,且第1 AC包含至少一种另外的可检测标记,第2 AC包含至少一种另外的可检测标记,两者不同。优选为细胞群为T细胞群。进一步优选为,与第1 AC特异性和/或选择性结合的所选免疫细胞为T细胞。还优选为,在所选T细胞表面上与PAI特异性和/或选择性结合的抗原结合蛋白为TCR。

[0188] 在另一实施方案中,使步骤(i)的细胞群与包含PAI(优选为TP)和可检测标记A的第一抗原复合体(第1 AC)接触;与包含相似蛋白抗原(SPA)(优选为TSP)的至少第二抗原复合体(第2 AC)接触,其中SPA(优选为TSP)的氨基酸序列与PAI(优选为TP)的氨基酸序列至少有1个氨基酸不同,其中,第2 AC包含可检测地不同于标记A的可检测标记B,且第1 AC包含至少一种另外的可检测标记,第2 AC包含至少一种另外的可检测标记,两者相同,且使步骤(i)的细胞群与一种或多种另外抗原复合体(AC)接触,其中每个AC均包含SPA,该SPA之至少一个氨基酸序列与第2 AC的SPA氨基酸序列不同,一个或更多另外AC包含至少一种另外的可检测标记;其中选择至少一个另外的标记,以便可区分第1 AC和第2 AC以及一个或多个另外AC。优选为细胞群为T细胞群。进一步优选为,与第1 AC特异性和/或选择性结合的所选免疫细胞为T细胞。还优选为,在所选T细胞表面上与PAI特异性和/或选择性结合的抗原结合蛋白为TCR。

[0189] 在另一实施方案中,使步骤(i)的细胞群与包含PAI(优选为TP)和可检测标记A的第一抗原复合体(第1 AC)接触;与包含相似蛋白抗原(SPA)(优选为TSP)的至少第二抗原复合体(第2 AC)接触,其中SPA(优选为TSP)的氨基酸序列与PAI(优选为TP)的氨基酸序列至少有1个氨基酸不同,其中,第2 AC包含可检测地不同于标记A的可检测标记B,且第1 AC包含至少一种另外的可检测标记,第2 AC包含至少一种另外的可检测标记,两者不同,且使步骤(i)的细胞群与一种或多种另外抗原复合体(AC)接触,其中每个AC均包含SPA,该SPA之至少一个氨基酸序列与第2 AC的SPA氨基酸序列不同,并且其中一个或更多另外AC包含至少一种另外的可检测标记;其中选择至少一个另外的标记,以便可区分第1 AC和第2 AC以及一个或多个另外AC。优选为细胞群为T细胞群。进一步优选为,与第1 AC特异性和/或选择性结合的所选免疫细胞为T细胞。还优选为,在所选T细胞表面上与PAI特异性和/或选择性结合的抗原结合蛋白为TCR。

[0190] 在另一实施方案中,使步骤(i)的细胞群与包含PAI(优选为TP)和可检测标记A的第一抗原复合体(第1 AC)接触;与包含相似蛋白抗原(SPA)(优选为TSP)的至少第二抗原复合体(第2 AC)接触,其中SPA(优选为TSP)的氨基酸序列与PAI(优选为TP)的氨基酸序列至少有1个氨基酸不同,其中,第2 AC包含可检测地不同于标记A的可检测标记B,且第1 AC包含至少一种另外的可检测标记,第2 AC包含至少一种另外的可检测标记,两者相同,且使步骤(i)的细胞群与一至十个,即一、二、三、四、五、六、七、八、九或十个,优选为二至四个,最优选为两个另外抗原复合体(AC)接触,其中每个AC均包含不同的SPA,优选为该SPA之至少一个氨基酸序列与第2 AC的SPA氨基酸序列不同者,并且其中一个或更多另外AC包含至少

一种另外的可检测标记;其中选择至少一个另外的标记,以便可区分第1 AC和第2 AC以及一个或多个另外AC。优选为细胞群为T细胞群。进一步优选为,与第1 AC特异性和/或选择性结合的所选免疫细胞为T细胞。还优选为,在所选T细胞表面上与PAI特异性和/或选择性结合的抗原结合蛋白为TCR。

[0191] 在上述每个实施方案中,优选为每个SPA(特别是每个TSP)与PAI的氨基酸序列(特别是TP的氨基酸序列)具有至少50%、至少60%、至少70%、至少80%、至少90%的相似性。

[0192] 在另一实施方案中,提供了可检测标记。本领域技术人员熟知如何标记目的蛋白抗原、靶肽、相似蛋白抗原或目的靶相似肽。上面指定为A-F的可检测标记独立地选自自由磁标记、荧光标记、RNA条码、DNA条码或放射性标记组成的组。优选情况为,磁标记可包含磁珠或磁纳米颗粒,其可以用针对特定表面抗原的抗体包覆。磁标记可用于磁激活细胞分选(MACS)。优选情况为,可检测标记为选自自由黄嘌呤类、吡啶类、恶嗪类、菁类、苯乙烯基染料、香豆素类、卟啉类、金属-配体络合物类、荧光蛋白质类、纳米晶体、花类和酞菁类组成组中的荧光标记。还优选使用藻红蛋白(SA-PE)、链霉亲和素-藻蓝蛋白(SA-APC)或链霉亲和素亮紫421(SA-BV421)作为可检测标记A-F的荧光标记。在另一优选实施方案中,第1 AC为MHC-I或MHC-II与PAI的复合体,且PAI为靶肽(TP)。优选TP为TAA。另外或可替代情况为,第2 AC为MHC-I或MHC-II与SPA的复合体,且SPA为靶相似肽(TSP)。在另一优选实施方案中,第1 AC和第2 AC为可溶性多聚化MHC-肽复合体。

[0193] PAI和SPA的功能性差异和相似点:

[0194] 如上所述,PAI优选为表达于患病组织,而SPA相反,其优选为表达于健康人组织,因此基于健康组织上的表达选择SPA。发明人开发了一种内部高通量技术平台(XPRESIDENT),其包括大型免疫肽组数据库,由之前发现提呈于健康组织中的肽组成。SPA优选来自MHC,优选HLA类型来源,即,SPA能够与各MHC分子(优选为HLA分子)结合。这在免疫细胞为T细胞且SPA在与HLA分子结合下提呈至T细胞从而让T细胞可识别HLA提呈SPA的情况下是必需的。因此,优选选择已知在与PAI相同的HLA同种异型上提呈的SPA。优选情况为,本发明方法中使用的SPA在健康组织的细胞上表达,表达量为每细胞大于10个拷贝,优选为每细胞大于20个拷贝,优选为每细胞大于50个拷贝,再优选为每细胞大于100个拷贝。为了避免脱靶/肿瘤外毒性,需要反向选择能够与此类相对丰富的SPA结合且同时与PAI结合的T细胞。

[0195] TSP的数量:

[0196] T细胞的TCR识别给定TP内的氨基酸亚组,即,TCR的表位。因此,如果使用太多不同的TSP,则可能不存在主要或排他地与TP结合但不与TSP结合的TCR。因此,优选情况为,使用的TSP不超过10个、使用的TSP不超过9个、使用的TSP不超过8个、使用的TSP不超过7个、使用的TSP不超过6个、使用的TSP不超过5个、使用的TSP不超过4个、使用的TSP不超过3个、使用的TSP不超过2个或使用的TSP不超过1个。因此,优选情况为,在本发明的方法中,使用1-10个不同TSP、2-8个不同TSP、3-5个不同TSP或1-3个不同TSP。在一优选实施方案中,使用三个TSP。TSP为SPA的片段,并根据与上述SPA列出的相同标准选择。同样,选择在健康组织中强烈表达的TSP。因此,本发明方法中纳入的TSP为具有如上定义的高序列相似性者,即,优选每个SPA(特别是每个TSP)与PAI的氨基酸序列(特别是TP的氨基酸序列)具有至少50%、至少60%、至少70%、至少80%、至少90%的相似性,并在健康组织上表达最高。

[0197] TP和TSP的长度:

[0198] 在一实施方案中,TP的长度为8-11个氨基酸。TP也可包含12个氨基酸。在一实施方案中,TP的长度为13-25个氨基酸。在另一实施方案中,TP的长度为13-18个氨基酸。通常情况下,选择与给定TP具有相同长度的TSP。但是,可替代的情况为,TSP的长度可比TP长或短1至3个氨基酸。在TP为MHC提呈的实施方案中,选择一个或多个TSP的长度以便它们也可以是MHC提呈的。例如,如果TP的长度为8个氨基酸,则优选TSP的长度为7个或少于7个氨基酸或8个或8个以上的氨基酸。更优选为,使用包含不同长度氨基酸序列的TSP混合物。在TP与MHC-I结合的一实施方案中,TP的长度为8-12个氨基酸。在TP与MHC-I结合的另一实施方案中,TP的长度为8-11个氨基酸。在TP与MHC-I结合的一实施方案中,TP的长度为8-10个氨基酸。在TP与MHC-II结合的一实施方案中,TP的长度为13-23个氨基酸。在TP与MHC-II结合的一优选实施方案中,TP的长度为13-18个氨基酸。

[0199] TP和TSP的结构差异/相似性:

[0200] 在另一实施方案中,基于与TP的高序列相似性(相似性BLAST搜索),从健康组织提呈的HLA结合肽的XPRESIDENT数据库选择了TSP。XPRESIDENT数据库包含健康或患病组织上不同HLA同种异型提呈的肽。优选情况为,TSP和TP由相同的HLA同种异型提呈。提呈TSP和TP的HLA同种异型可选自由HLA-A、HLA-B、HLA-C、HLA-E、HLA-F、HLA-G、HLA-H、HLA-J、HLA-K、HLA-L组成的组。优选情况为,HLA-A蛋白选自由HLA-A1、HLA-A2、HLA-A3和HLA-A11组成的组。优选的HLA-A等位基因为HLA-A*02:01、HLA-A*01:01、HLA-A*03:01或HLA-A*24:02。优选的HLA-B等位基因为HLA-B*07:02、HLA-B*08:01、HLA-B*15:01、HLA-B*35:01或HLA-B*44:05。

[0201] 一般而言,对于上面列出的大多数HLA同种异型,给定MHC提呈肽的第二个氨基酸(从N端开始计数)和C端氨基酸不包含在与该肽特异性结合的TCR所识别的该肽表位中。

[0202] 在另一实施方案中,TSP的氨基酸序列长度为8至16个氨基酸,TP的长度为8个氨基酸,并且其中TSP的氨基酸序列的下列位置与TP的氨基酸序列不同:

[0203] $X_1-X_2-X_3-X_4-X_5-X_6-X_7-X_8$ (SEQ ID NO:1)

[0204] (i) 位置 X_1 、 X_2 和 X_3 ,并且其中位置 X_4 至 X_8 与TP相同或相似,优选为相同;或

[0205] (ii) 位置 X_4 、 X_5 和 X_6 ,并且其中位置 X_1 至 X_3 以及 X_7 和 X_8 与TP相同或相似,优选为相同;或

[0206] (iii) 位置 X_7 和 X_8 ,并且其中位置 X_1 至 X_6 与TP相同或相似,优选为相同。

[0207] 在一优选实施方案中,位置 X_1 、 X_2 和 X_3 在TSP中突变,其中位置 X_4 至 X_8 与TP相比相同。在另一优选实施方案中,位置 X_4 、 X_5 和 X_6 在TSP中突变,且位置 X_1 至 X_3 和 X_7 和 X_8 与TP相同。在另一优选实施方案中,位置 X_7 和 X_8 在TSP中突变,且位置 X_1 至 X_6 与TP相同。在另一优选实施方案中,位置 X_7 和 X_8 在TSP中突变,且位置 X_1 至 X_6 与TP相同。

[0208] 在另一优选实施方案中,在(i)至(iii)中的上述TSP混合物,即,具有不同突变模式的TSP用于本发明第一方面的方法中。在另一优选实施方案中,使用了(i)至(iii)中的上述TSP混合物(即,具有不同突变模式的TSP),也使用了不同长度的氨基酸序列。使用此类TSP混合物可快速有效地正向选择与TP结合的免疫细胞,并负向选择与一种或多种TSP结合的免疫细胞。

[0209] 在另一优选实施方案中,TSP的氨基酸序列长度为8至16个氨基酸,TP的长度为9个

氨基酸, TSP的氨基酸序列的下列位置与TP的氨基酸序列不同

[0210] $X_1-X_2-X_3-X_4-X_5-X_6-X_7-X_8-X_9$ (SEQ ID NO:2)

[0211] (i) 位置 X_1 、 X_2 和 X_3 , 并且其中位置 X_4 至 X_9 与TP相同或相似, 优选为相同;

[0212] (ii) 位置 X_4 、 X_5 和 X_6 , 并且其中位置 X_1 至 X_3 以及位置 X_7 至 X_9 与TP相同或相似, 优选为相同;或

[0213] (iii) 位置 X_4 、 X_5 、 X_6 和 X_7 , 并且其中位置 X_1 至 X_3 以及位置 X_8 至 X_9 与TP相同;或(iv) 位置 X_7 、 X_8 和 X_9 , 并且其中位置 X_1 至 X_6 与TP相同或相似, 优选为相同。

[0214] 在一优选实施方案中, 位置 X_1 、 X_2 和 X_3 在TSP中突变, 其中位置 X_4 至 X_9 与TP相比相同。在另一优选实施方案中, 位置 X_4 、 X_5 、 X_6 和 X_7 在TSP中突变, 其中位置 X_1 至 X_3 和位置 X_7 和 X_9 与TP相同。在另一优选实施方案中, 位置 X_7 至 X_9 在TSP中突变, 且位置 X_1 至 X_6 与TP相同。在另一优选实施方案中, 位置 X_7 至 X_9 在TSP中突变, 且位置 X_1 至 X_6 与TP相同。在另一优选实施方案中, 在(i)至(iv)中的上述TSP混合物, 即, 具有不同突变模式的TSP用于本发明第一方面的方法中。在另一优选实施方案中, 使用了(i)至(iv)中的上述TSP混合物(即, 具有不同突变模式的TSP), 也使用了不同长度的氨基酸序列。

[0215] 在另一优选实施方案中, TSP的氨基酸序列长度为8至16个氨基酸, TP的长度为10个氨基酸, TSP的氨基酸序列的下列位置与TP的氨基酸序列不同

[0216] $X_1-X_2-X_3-X_4-X_5-X_6-X_7-X_8-X_9-X_{10}$ (SEQ ID NO:3)

[0217] (i) 位置 X_1 、 X_2 和 X_3 , 其中位置 X_4 至 X_{10} 与TP相同或相似, 优选为相同;

[0218] (ii) 位置 X_4 、 X_5 、 X_6 和 X_7 , 其中位置 X_1 至 X_3 以及位置 X_8 至 X_{10} 与TP相同或相似, 优选为相同;或

[0219] (iii) 位置 X_4 、 X_5 和 X_6 , 并且其中位置 X_1 至 X_3 以及位置 X_7 至 X_{10} 与TP相同或相似, 优选为相同;或

[0220] (iv) 位置 X_8 、 X_9 和 X_{10} , 并且其中位置 X_1 至 X_7 与TP相同或相似, 优选为相同。

[0221] 在一优选实施方案中, 位置 X_1 、 X_2 和 X_3 在TSP中突变, 其中位置 X_4 至 X_{10} 与TP相比相同。在另一优选实施方案中, 位置 X_4 至 X_7 在TSP中突变, 其中位置 X_1 至 X_3 和位置 X_8 和 X_{10} 与TP相同。在另一优选实施方案中, 位置 X_4 至 X_6 在TSP中突变, 其中位置 X_1 至 X_3 和位置 X_7 至 X_{10} 与TP相同。在另一优选实施方案中, 位置 X_8 至 X_{10} 在TSP中突变, 且位置 X_1 至 X_7 与TP相同。在另一优选实施方案中, 在(i)至(iv)中的上述TSP混合物, 即, 具有不同突变模式的TSP用于本发明第一方面的方法中。在另一优选实施方案中, 使用了(i)至(iv)中的上述TSP混合物(即, 具有不同突变模式的TSP), 也使用了不同长度的氨基酸序列。

[0222] 在另一优选实施方案中, TSP的氨基酸序列长度为8至16个氨基酸, TP的长度为11个氨基酸, TSP的氨基酸序列的下列位置与TP的氨基酸序列不同

[0223] $X_1-X_2-X_3-X_4-X_5-X_6-X_7-X_8-X_9-X_{10}-X_{11}$ (SEQ ID NO:4)

[0224] (i) 位置 X_1 、 X_2 和 X_3 , 其中位置 X_4 至 X_{11} 与TP相同或相似, 优选为相同;

[0225] (ii) 位置 X_4 、 X_5 、 X_6 和 X_7 , 其中位置 X_1 至 X_3 以及位置 X_8 至 X_{11} 与TP相同或相似, 优选为相同;或

[0226] (iii) 位置 X_4 、 X_5 和 X_6 , 并且其中位置 X_1 至 X_3 以及位置 X_7 至 X_{11} 与TP相同或相似, 优选为相同;或

[0227] (iv) 位置 X_8 、 X_9 、 X_{10} 和 X_{11} , 其中位置 X_1 至 X_7 与TP相同或相似, 优选为相同;或

[0228] (v) 位置 X_9 、 X_{10} 和 X_{11} ,其中位置 X_1 至 X_8 与TP相同或相似,优选为相同。

[0229] 在一优选实施方案中,位置 X_1 、 X_2 和 X_3 在TSP中突变,其中位置 X_4 至 X_{11} 与TP相比相同。在另一优选实施方案中,位置 X_4 至 X_7 在TSP中突变,其中位置 X_1 至 X_3 和位置 X_8 和 X_{11} 与TP相同。在另一优选实施方案中,位置 X_4 至 X_6 在TSP中突变,其中位置 X_1 至 X_3 和位置 X_7 至 X_{11} 与TP相同。在另一优选实施方案中,位置 X_8 至 X_{11} 在TSP中突变,且位置 X_1 至 X_7 与TP相同。在另一优选实施方案中,位置 X_9 至 X_{11} 在TSP中突变,且位置 X_1 至 X_8 与TP相同。

[0230] 在另一优选实施方案中,在(i)至(iv)中的上述TSP混合物,即,具有不同突变模式的TSP用于本发明第一方面的方法中。在另一优选实施方案中,使用了(i)至(iv)中的上述TSP混合物(即,具有不同突变模式的TSP),也使用了不同长度的氨基酸序列。

[0231] 在另一优选实施方案中,TSP的氨基酸序列长度为8-16个氨基酸,TP的长度为12个氨基酸,TSP的氨基酸序列的下列位置与TP的氨基酸序列不同

[0232] X_1 - X_2 - X_3 - X_4 - X_5 - X_6 - X_7 - X_8 - X_9 - X_{10} - X_{11} - X_{12} (SEQ ID NO:5)

[0233] (i) 位置 X_1 、 X_2 和 X_3 ,其中位置 X_4 至 X_{12} 与TP相同或相似;

[0234] (ii) 位置 X_4 、 X_5 、 X_6 和 X_7 ,其中位置 X_1 至 X_3 以及位置 X_8 至 X_{12} 与TP相同或相似;或

[0235] (iii) 位置 X_4 、 X_5 和 X_6 ,其中位置 X_1 至 X_3 以及位置 X_7 至 X_{12} 与TP相同或相似;或

[0236] (iv) 位置 X_8 、 X_9 、 X_{10} 、 X_{11} 和 X_{12} ,其中位置 X_1 至 X_7 与TP相同或相似;或

[0237] (v) 位置 X_9 、 X_{10} 、 X_{11} 和 X_{12} ,其中位置 X_1 至 X_8 与TP相同或相似。

[0238] 在一优选实施方案中,位置 X_1 、 X_2 和 X_3 在TSP中突变,其中位置 X_4 至 X_{12} 与TP相比相同。在另一优选实施方案中,位置 X_4 至 X_7 在TSP中突变,其中位置 X_1 至 X_3 和位置 X_8 和 X_{12} 与TP相同。在另一优选实施方案中,位置 X_4 至 X_6 在TSP中突变,其中位置 X_1 至 X_3 和位置 X_7 至 X_{12} 与TP相同。在另一优选实施方案中,位置 X_8 至 X_{12} 在TSP中突变,且位置 X_1 至 X_7 与TP相同。在另一优选实施方案中,位置 X_9 至 X_{12} 在TSP中突变,且位置 X_1 至 X_8 与TP相同。

[0239] 在另一优选实施方案中,在(i)至(iv)中的上述TSP混合物,即,具有不同突变模式的TSP用于本发明第一方面的方法中。在另一优选实施方案中,使用了(i)至(iv)中的上述TSP混合物(即,具有不同突变模式的TSP),也使用了不同长度的氨基酸序列。

[0240] 在本发明第一方面的另一实施方案中,SPA或SPA中包含的至少一种蛋白质或肽的氨基酸序列与PAI的氨基酸序列具有至少50%、至少60%、至少70%、至少80%或至少90%或至少95%的相似性。在另一实施方案中,SPA或SPA中所含的至少一种蛋白质或肽的氨基酸序列与PAI具有小于等于90%、小于等于89%、小于等于88%、小于等于87%或小于等于86%的氨基酸同一性。在另一实施方案中,SPA或SPA中所含的至少一种蛋白质或肽的氨基酸序列与PAI具有小于等于85%、小于等于84%、小于等于83%、小于等于82%、小于等于81%或小于等于80%的氨基酸同一性。

[0241] 在本发明第一方面的另一实施方案中,TSP的氨基酸序列与TP具有小于等于96%、小于等于95%、小于等于94%、小于等于93%、小于等于92%或小于等于91%的氨基酸同一性。在另一实施方案中,TSP的氨基酸序列与TP具有小于90%、小于等于89%、小于等于88%、小于等于87%或小于等于86%的氨基酸同一性。在另一实施方案中,TSP的氨基酸序列与TP具有小于等于85%、小于等于84%、小于等于83%、小于等于82%、小于等于81%或小于等于80%的氨基酸同一性。

[0242] 在另一实施方案中,TSP在健康组织上的绝对表达与本发明方法中纳入的TSP的最

低序列同一性相关。如果TSP在健康组织上高表达,由于亲和力效应,仅以低亲和力与TSP结合的TCR可能仍与健康组织中表达的TSP结合。因此,如果给定的TSP在健康组织上拷贝数低,则仅其显示与TP高度相似的情况下才纳入在本发明的方法中。相应地,如果给定的TSP在健康组织上拷贝数高,虽然其与TP相似性低,也纳入在本发明的方法中。例如,如果TSP的每个健康细胞的拷贝数低于10,则在与TP具有至少90%的序列相似性时纳入。如果TSP的拷贝数介于1至25之间,则在它们与TP具有至少85%的序列相似性时纳入此类TSP。如果TSP的拷贝数介于25至100/细胞之间,则在它们与TP具有至少80%的序列相似性时纳入此类TSP。如果TSP的拷贝数介于100至250/细胞之间,则在它们与TP具有至少75%的序列相似性时纳入此类TSP。如果TSP的拷贝数大于250/细胞,则在它们与TP具有至少50%的序列相似性时纳入此类TSP。

[0243] 在本发明第一方面所述方法的另一实施方案中,所述方法的步骤(ii)和(iii)连续或同时进行。在本发明第一方面的另一实施方案中,上述步骤(a)、(b)、(c)和(d)连续或同时进行。步骤(a)、(b)、(c)和(d)是否组合进行取决于可检测标记所标记的AC数量的使用情况。

[0244] 在本发明第一方面所述方法的另一实施方案中,步骤(iv)包括正向选择(选择)结合于第1 AC、第1和第3 AC或第1、第3和第4 AC的细胞。在另一实施方案中,步骤(iv)包括负向选择(排除)结合于第2 AC、第2和第5 AC或第2、第5和第6 AC的细胞。在另一优选实施方案中,步骤(iv)包括选择结合于第1 AC、第1和第3 AC或第1、第3和第4 AC的细胞以及排除结合于第2 AC、第2和第5 AC或第2、第5和第6 AC的细胞。

[0245] 在本发明第一方面所述方法的另一实施方案中,可检测标记透过流式细胞术分析检测。在一优选实施方案中,可检测标记透过FACS分析检测。在另一优选实施方案中,可检测标记透过制备型分选分析检测。

[0246] 在本发明第一方面所述方法的另一实施方案中,本发明第一方面所述方法的步骤i)细胞群中所含的细胞为T细胞且经表型分析。在另一实施方案中,步骤i)细胞群中所含的细胞为B细胞且经表型分析。

[0247] 在本发明第一方面所述方法的另一实施方案中,T细胞的表型分析包括确定一个或多个T细胞标志物。T细胞标志物优选选自由CD3、CD4、CD8、CD11a、CD14、CD19、CD25、CD27、CD28、CD44、CD45RA、CD45R0、CD57、CD62L、CD69、CD122、CD127、CD137、CD197(CCR7)、IFN γ 、IL-2、TNF α 、IL7R和端粒长度组成的组。在另一优选实施方案中,T细胞标志物为CD45RA、CD45R0、CD197、CD25、CD27、CD57、CD95、CD127和CD62L。特别优选为CD69和CD137用于T细胞表型分析。在另一优选实施方案中,B细胞的表型分析包括确定一个或多个B细胞标志物。B细胞标志物优选选自由CD19、CD27、CD45R、CD21、CD40、CD20、CD38和CD83组成的组。

[0248] 在本发明第一方面所述方法的另一实施方案中,所述方法进一步包括的步骤有:使步骤(i)的细胞群与包含无关蛋白抗原(IPA)的无关抗原复合体(IAC)接触,其中当IPA的氨基酸序列与PAI的氨基酸序列比对时,IPA的氨基酸序列的两个或两个以下氨基酸位置与PAI相同,并且其中IAC包含可检测地不同于可检测标记A的可检测标记G。优选情况为,无关蛋白抗原为管家基因的基因产物。管家基因产物在正常和病理生理条件下在生物体的所有细胞中表达,这使其适合作为参考基因,因为在不同或变化的细胞条件下其通常不会上调或下调。一般而言,天然免疫细胞群不应包含任何与管家基因或其衍生肽结合的免疫细胞。

因此,在本发明的方法中纳入IPC可鉴定与AC非特异性结合的T细胞,这些T细胞也是不受欢迎的。

[0249] 在另一实施方案中,透过以下一项或多项标准选择至少一个IP的氨基酸序列:IP提呈于健康组织;IP衍生自HLA类型来源;或与各自HLA结合。优选情况为,IP或IPA中所含的至少一种蛋白质或肽的氨基酸序列与PAI具有小于50%、小于40%、小于30%、小于29%、小于28%、小于27%、小于26%、小于25%、小于24%、小于23%、小于22%、小于21%、小于20%、小于19%、小于18%、小于17%、小于16%、小于15%、小于10%、小于5%的氨基酸同一性。在另一实施方案中,IAC为MHC-I或MHC-II与IP的复合体。优选情况为,当与TP的氨基酸序列比对时,IP的氨基酸序列的一个氨基酸位置或无氨基酸位置与TP相同。优选情况为,IP由管家基因编码。

[0250] 本发明的第二方面进一步涉及一种确定编码抗原结合蛋白或其抗原结合部分的核酸序列的方法,所述方法包括以下步骤:

[0251] (i) 分离来自本发明第一方面的方法中选定细胞的编码抗原结合蛋白或其抗原结合部分的核酸;和

[0252] (ii) 确定核酸的序列。

[0253] 在一优选实施方案中,透过本领域熟知的方法(例如,有机萃取、固相萃取),使用包含含有亚氨基二乙酸基团的苯乙烯-二乙烯基苯共聚物树脂,从所选的免疫细胞中分离核酸。在另一实施方案中,分离的核酸为DNA或RNA。在另一实施方案中,优选在分离后扩增核酸。优选情况为,透过聚合酶链反应(PCR)进行扩增。更优选为,用cDNA末端快速扩增PCR(RACE PCR)来扩增核酸。在另一实施方案中,透过逆转录从RNA范本中合成DNA,从而产生互补DNA(cDNA)。逆转录酶(RT)使用RNA范本和与RNA 3'末端互补的短引物来指导第一链cDNA的合成,其可直接用作PCR的范本。在另一实施方案中,分离核酸的序列可透过本领域已知的方法来确定,例如:下一代测序法,例如,透过同时鉴定DNA碱基(因为每个碱基都发出独特的荧光信号)并将其添加入核酸链中所进行的Illumina(Solexa)测序;基于焦磷酸测序法的Roche 454测序,这是一种在核苷酸透过聚合酶掺入到新的DNA链后同样使用荧光来检测焦磷酸盐释放的一种技术;或Ion Torrent:质子/PGM测序法测量质子从DNA聚合酶掺入单个碱基的直接释放。

[0254] 本发明的第三方面涉及一种产生表达编码抗原结合蛋白或其抗原结合部分的核酸的细胞的方法,所述方法包括以下步骤:

[0255] (i) 提供来自本发明第一方面的方法中选定细胞的编码抗原结合蛋白或其抗原结合部分的核酸序列;

[0256] (ii) 任选地在表达控制元素控制下产生包含步骤(i)中所提供核酸序列的核酸载体;和

[0257] (iii) 将步骤(ii)的核酸载体引入宿主细胞。

[0258] 在一实施方案中,抗原结合蛋白或其抗原结合部分被克隆至载体中。

[0259] 在一实施方案中,抗原结合蛋白为TCR或其抗原结合片段、BCR或其抗原结合片段或抗体或其抗原结合片段。在另一实施方案中,抗原结合蛋白为TCR或其部分由至少 α 和 β 链的可变结构域组成。优选情况为,TCR或其抗原结合部分的序列被插入合适的载体中。在另一实施方案中,TCR、BCR或抗体的氨基酸序列包含六个CDR。在另一实施方案中,鉴定TCR的

可变 α 和/或 β 结构域的两个或三个CDR被插入框架或另一TCR或抗体中。优选情况为,将CDR的基因序列(其由TCR的可变 α 结构域组成)被克隆入包含框架区的合适载体中。表达载体可包含编码轻链或重链或 α 和 β 链两者(或其可变结构域)的核酸——即所谓的“串联型”,它们也可被单独载体中所包含的核酸编码。优选使用串联类型人源化抗体表达载体(Shitara K et al. J Immunol Methods.1994 Jan.3;167(1-2):271-8)。串联型人源化抗体表达载体的实例包括,例如:pKANTEX93(WO 97/10354)和pEE18。

[0260] 在另一实施方案中,步骤(ii)的载体被引入宿主细胞。在一实施方案中,此类重组宿主细胞可用于产生至少一种本发明的抗原结合蛋白或其部分。优选情况为,宿主细胞用编码抗原结合蛋白或其抗原结合部分的核酸和/或载体转化、转导或转染。用编码抗原结合蛋白或抗原结合蛋白部分的核酸对宿主细胞进行转导或转染使用本领域熟知的方法(例如US20190216852中所述的方法)进行。在另一实施方案中,包含抗原结合蛋白或其抗原结合部分的宿主细胞可以为真核细胞,例如植物、动物、真菌或藻类,也可以为原核细胞,例如细菌或原生动物。宿主细胞可以是培养细胞或原代细胞,即直接从生物体(例如人)分离。宿主细胞可以是贴壁T细胞或悬浮细胞(即,在悬浮液中生长的细胞)。为了产生抗原结合蛋白或部分抗原结合蛋白(例如,重组TCR或其片段)之目的,宿主细胞优选为哺乳动物细胞。最优选情况为,宿主细胞为人细胞。虽然宿主细胞可以是任何细胞类型,可以来自任何类型的组织,可以是任何发育阶段的细胞,但是宿主细胞优选为外周血白细胞(PBL)或外周血单核细胞(PBMC)或B细胞。更优选为,宿主细胞为T细胞。T细胞可以是任何T细胞,诸如培养的T细胞,例如,原代T细胞或来自培养T细胞系的T细胞,例如Jurkat、SupT1等,或获得自哺乳动物的T细胞,优选为来自人类患者的T细胞或T细胞前体。如果获得自哺乳动物,T细胞可从许多来源获得,包括但不限于血液、骨髓、淋巴结、胸腺或其他组织或液体。优选情况为,T细胞为人T细胞。更优选为,T细胞是分离自人的T细胞。T细胞可以是任何类型的T细胞并且可以是任何发育阶段的,包括但不限于CD4阳性和/或CD8阳性、CD4阳性辅助T细胞,例如,Th1和Th2细胞、CD8-阳性T细胞(例如,细胞毒性T细胞)、肿瘤浸润细胞(TIL)、记忆T细胞、幼稚T细胞。优选情况为,T细胞为CD8阳性T细胞或CD4阳性T细胞。在另一实施方案中,宿主细胞可以为用于重组表达的任何细胞。优选情况为,宿主细胞为中国仓鼠卵巢(CHO)细胞。

[0261] 本发明的第四方面涉及一种用于治疗有需要的受试者的方法,所述方法包括以下步骤:

[0262] (i) 提出包含免疫细胞的受试者细胞群;

[0263] (ii) 使步骤(i)的细胞群与包含PAI和可检测标记A的第一抗原复合体(第1 AC)或包含可检测标记A的PAI接触;

[0264] (iii) 使步骤(i)的细胞群与包含SPA的至少第二抗原复合体(第2 AC)接触,其中SPA的氨基酸序列与PAI的氨基酸序列至少有1个氨基酸不同,其中第2 AC包含可检测标记B;以及

[0265] (iv) 选择至少一种与第一AC特异性结合的细胞,

[0266] 其中,可检测标记A和可检测标记B彼此系可检测地不同

[0267] (v) 透过培养增加至少一个所选细胞数目;和

[0268] (vi) 将培养的细胞重新引入受试者体内。

[0269] 这种方法为ACT法。优选情况为,在基因改造和功能改善之后,所选的细胞被转移

入受试者体内。优选情况为,细胞源自在其经基因改造后将转入的受试者,即,所述细胞的供体与改造细胞的受体相同。受试者是指上文所定义的有需要的受试者。优选情况为,有需要的受试者患有选自由免疫疾病或肿瘤性疾病、病毒引起的疾病或细菌引起的疾病组成组的疾病,或处于患上所述疾病的风险。优选肿瘤性疾病为癌症。优选病毒引起的疾病为HIV。优选细菌引起的疾病为结核病。

[0270] 本发明的第五方面涉及一种选择免疫细胞的方法,所述免疫细胞在其表面表达与目的蛋白抗原(PAI)特异性结合的抗原结合蛋白,所述方法包括以下步骤:

[0271] (i) 提出包含免疫细胞的细胞群;

[0272] (ii) 使步骤(i)的细胞群与包含PAI和可检测标记A的第一抗原复合体(第1 AC)或包含可检测标记A的PAI接触;

[0273] (iii) 使步骤(i)的细胞群与包含无关蛋白抗原(IPA)的至少第二抗原复合体(第2 AC)接触,其中当IPA的氨基酸序列与PAI的氨基酸序列比对时,IPA的氨基酸序列的两个或两个以下氨基酸位置与PAI相同,其中IAC包含可检测标记G;或与IPA和可检测标记G接触;以及

[0274] (iv) 选择至少一种与第一AC特异性结合的细胞,

[0275] 其中,可检测标记A和可检测标记G彼此系可检测地不同。

[0276] 根据本发明第五方面的选择过程按本发明第一方面的上述方法实施。

实施例

[0277] 实施例1:有和没有事先靶标特异性扩增的靶肽特异性T细胞的直接分选。图2和图3显示了两种不同方法的比较,这两种方法可分选出靶标特异性T细胞,同时免除识别靶标和相似肽的交叉反应性T细胞。方法1(图2)无需事先扩增靶标特异性T细胞。透过磁珠分离法来分离并富集PBMC中的T细胞群。用荧光染料缀合的靶肽和相似肽四聚体对T细胞群进行染色。随后,透过使用靶向作用于靶四聚体的其中一种荧光染料缀合物的磁珠分离法,可进一步富集那些细胞以获得靶标特异性T细胞。然后对靶肽特异性T细胞群的表面标志物(例如,CD4和CD8)以及生存力标志物进行染色,以排除死的细胞。透过使用流式细胞术分选方法,可以分选出靶标特异性T细胞的所需表面标志物表达,同时免除靶+相似肽特异性T细胞,如图2所示。方法2(图3)利用靶肽HLA包被的人工抗原提呈细胞刺激以扩增低频靶标特异性T细胞。此时,在单个容器中培养富集CD8T细胞,从而让单克隆或寡克隆靶肽特异性T细胞群生长。在用人工抗原提呈细胞反复刺激后,用表面标志物以及靶肽和相似肽四聚体对单个单克隆或寡克隆细胞群进行染色(靶肽和相似肽四聚体用2种不同的荧光染料标记,各采用2D染色方法),这可以区分靶标特异性和交叉反应性单或寡克隆T细胞群,如图3所示。

[0278] 实施例2:衍生自仅用靶肽多聚体分选的T细胞的T细胞受体功能评估。为了评估透过用靶多聚体分选而鉴定的TCR的功能性和特异性,T细胞受体mRNA透过体外转录产生,随后用于透过电穿孔转染健康供体的CD8阳性T细胞。电穿孔后18小时,然后将20,000个转染T细胞与用靶肽、不同靶序列相似肽、无关肽或空载T2细胞载入的T2细胞以1:1的比例共培养。共培养开始后24小时收集上清液,并透过ELISA技术分析所分泌的IFN- γ 。细胞因子分泌证明了各T细胞的抗原识别和活性,如图4所示。尽管图4A、B和C中的所有TCR均能识别靶标(阳性对照),但图4A和图4B中的TCR对在正常组织上表达的靶序列相似肽也具有交叉反

应性,因此被排除出进一步分析,只有“干净的”TCR(图4C)才值得被选择用于进一步表征。(“靶标”=靶肽;TP;“SIM 1-10”=靶相似肽;TSP 1-10)。

[0279] 本实施例中使用的肽:

[0280] TP和SIM 1-SIM 10全部为9元体。

[0281] • TP和SIM 1在第2、5、8和9氨基酸位置不同,其中SIM 1在第2位置有异亮氨酸残基,在第5位置有苏氨酸残基,在第8位置有亮氨酸残基,在第9位置有缬氨酸残基。

[0282] • TP和SIM 2在第3、4和7位置不同,其中SIM 2在第3位置有异亮氨酸残基,在第4位置有谷氨酸残基,在第7位置有谷氨酰胺残基。

[0283] • TP和SIM 3在第2、7、8和9位置不同,其中SIM 3在第3位置有异亮氨酸残基,在第7和第8位置有谷氨酸残基,在第9位置有异亮氨酸残基。

[0284] • TP和SIM 4在第4、5和8位置不同,其中SIM 4在第4位置有赖氨酸残基,在第5位置有天冬酰胺残基,在第8位置有酪氨酸残基。

[0285] • TP和SIM 5在第4、7和8位置不同,其中SIM 5在第4位置有天冬酰胺残基,在第7位置有脯氨酸残基,在第8位置有酪氨酸残基。

[0286] • TP和SIM 6在第6、7和8位置不同,其中SIM 6在第6位置有缬氨酸残基,在第7和第8位置有亮氨酸残基。

[0287] • TP和SIM 7在第5、6和8位置不同,其中SIM 7在第5位置有赖氨酸残基,在第6位置有谷氨酰胺残基,在第8位置有甲硫氨酸残基。

[0288] • TP和SIM 8在第3、5、7和9位置不同,其中SIM 8在第3位置有丝氨酸残基,在第5位置有谷氨酸残基,在第7和第9位置有缬氨酸残基。

[0289] • TP和SIM 9在第2、4、5和9位置不同,其中SIM 9在第2位置有缬氨酸残基,在第4位置有甘氨酸残基,在第5位置有丙氨酸残基,在第9位置有缬氨酸残基。

[0290] • TP和SIM 10在第1、4和6位置不同,其中SIM 10在第1位置有缬氨酸残基,在第4位置有组氨酸残基,在第9位置有谷氨酰胺残基。

[0291] • TP和对照肽在第4-9位置不同。

[0292] 实施例3:衍生自靶肽以及靶肽和相似肽特异性T细胞的T细胞受体功能评估。为了该目的,T细胞受体mRNA使用体外转录产生,随后用于透过电穿孔转染NFAT-萤光素酶Jurkat细胞。转染的Jurkat细胞开始在其表面上暂态表达新引入的TCR。电穿孔后18小时,然后将50,000个Jurkat细胞与载入靶肽或用于分选的相似肽以及对照肽或无肽的T2细胞以1:1的比例与共孵育。当TCR与其同源肽-HLA特异性结合后,信号传导使NFAT激活,继而表达萤光素酶。过夜孵育后,添加萤光素酶底物,当T细胞被激活时可检测到发光信号。图5显示了当用靶肽载入T2细胞刺激时,衍生自靶四聚体结合T细胞的TCR导致功能激活。(“靶标”=靶肽;TP;“SIM 1-3”=靶相似肽;TSP 1-3)。

[0293] 本实施例中使用的肽:

[0294] TP和SIM 1-SIM 3全部为9元体。

[0295] • TP和SIM 1在第4、6和7氨基酸位置不同,其中SIM 1在第4位置有谷氨酸残基,在第6位置有亮氨酸残基,在第7位置有异亮氨酸残基。

[0296] • TP和SIM 2在第2、7和8位置不同,其中SIM 2在第2位置有甲硫氨酸残基,在第7位置有谷氨酸残基,在第8位置有赖氨酸残基。

[0297] • TP和SIM 3在第1、5和6位置不同,其中SIM 3在第1位置有苯丙氨酸残基,在第5位置有甘氨酸残基,在第6位置有丝氨酸残基。

[0298] • TP和对照肽在第1位置和第4-9位置不同。

[0299] 总体而言,使用BLASTP、BLOSUM62评分矩阵、词长为3和期望值(E)为10时,SIM 1与TP具有77%的相似性,SIM 2与TP具有77%的相似性,SIM 3与TP具有75%的相似性。

[0300] 实施例4:给定靶肽的相关和无关靶相似肽。

[0301] 肽作为TSP与给定TP的相关性主要由其与TP的相似性决定,且还可由其提呈频率以及在原代正常组织上的定量提呈水准(每个细胞的拷贝数(CpC))决定。与TP的相似性越高并且在正常组织上的提呈频率和CpC越高,TSP的相关性就越高。表6显示了TP、两个相应TSP以及IP的示例性序列。每个肽都描述了与TP相同氨基酸(aa)的数量,与TP序列相比基于pmbec位置评分矩阵的相似性以及正常组织上的CpC范围。图6、7和8额外显示了两个TSP以及IP的肽提呈谱。TSP(TSP1)与TP有4个相同的氨基酸,但与和靶肽相比具有5个相同氨基酸的TSP2比较,与靶标具有更高的总体相似性。根据两个TSP与TP的相似性及其在正常组织中的提呈情况,视为两个TSP具有相关性(图6和7)。所描述的IP显示在正常组织上提呈频率甚至更高(图8),通常也以每个细胞较高的拷贝数提呈。但是,序列相似性以及相同氨基酸数量相当低(相似性17%,相同氨基酸0个)。

[0302] 表6:

	氨基酸序列	与 TP 相同的 aa 数量	与 TP 的相似性 (PMBEC 得分)	正常组织的 CpC 范围
TP	VLLHHQIGL (SEQ ID NO: 164)	9	100%	n.a.
[0303] TSP1	ALMYHTITL (SEQ ID NO: 165)	4	63%	5-60
TSP2	LLLAHIAL (SEQ ID NO: 166)	5	55 %	15-35
IP	AIVDKVPSV (SEQ ID NO: 167)	0	17%	55-600

<212> PRT
<213> 人工序列
<220>
<223> TAA
<400> 16
Ser Val Asp Val Ser Pro Pro Lys Val
1 5
<210> 17
<211> 8
<212> PRT
<213> 人工序列
<220>
<223> TAA
<400> 17
Val Ala Asp Lys Ile His Ser Val
1 5
<210> 18
<211> 9
<212> PRT
<213> 人工序列
<220>
<223> TAA
<400> 18
Ile Val Asp Asp Leu Thr Ile Asn Leu
1 5
<210> 19
<211> 9
<212> PRT
<213> 人工序列
<220>
<223> TAA
<400> 19
Gly Leu Leu Glu Glu Leu Val Thr Val
1 5
<210> 20
<211> 10
<212> PRT
<213> 人工序列
<220>

<223> TAA
 <400> 20
 Thr Leu Asp Gly Ala Ala Val Asn Gln Val
 1 5 10
 <210> 21
 <211> 10
 <212> PRT
 <213> 人工序列
 <220>
 <223> TAA
 <400> 21
 Ser Val Leu Glu Lys Glu Ile Tyr Ser Ile
 1 5 10
 <210> 22
 <211> 9
 <212> PRT
 <213> 人工序列
 <220>
 <223> TAA
 <400> 22
 Leu Leu Asp Pro Lys Thr Ile Phe Leu
 1 5
 <210> 23
 <211> 9
 <212> PRT
 <213> 人工序列
 <220>
 <223> TAA
 <400> 23
 Tyr Leu Met Asp Asp Phe Ser Ser Leu
 1 5
 <210> 24
 <211> 9
 <212> PRT
 <213> 人工序列
 <220>
 <223> TAA
 <400> 24
 Lys Val Trp Ser Asp Val Thr Pro Leu

<212> PRT
<213> 人工序列
<220>
<223> TAA
<400> 29
Leu Leu Trp Gly His Pro Arg Val Ala Leu Ala
1 5 10
<210> 30
<211> 9
<212> PRT
<213> 人工序列
<220>
<223> TAA
<400> 30
Phe Leu Leu Glu Arg Glu Gln Leu Leu
1 5
<210> 31
<211> 9
<212> PRT
<213> 人工序列
<220>
<223> TAA
<400> 31
Ser Leu Ala Glu Thr Ile Phe Ile Val
1 5
<210> 32
<211> 9
<212> PRT
<213> 人工序列
<220>
<223> TAA
<400> 32
Thr Leu Leu Glu Gly Ile Ser Arg Ala
1 5
<210> 33
<211> 9
<212> PRT
<213> 人工序列
<220>

<223> TAA
<400> 33
Ile Leu Gln Asp Gly Gln Phe Leu Val
1 5
<210> 34
<211> 10
<212> PRT
<213> 人工序列
<220>
<223> TAA
<400> 34
Val Ile Phe Glu Gly Glu Pro Met Tyr Leu
1 5 10
<210> 35
<211> 9
<212> PRT
<213> 人工序列
<220>
<223> TAA
<400> 35
Ser Leu Phe Glu Ser Leu Glu Tyr Leu
1 5
<210> 36
<211> 9
<212> PRT
<213> 人工序列
<220>
<223> TAA
<400> 36
Ser Leu Leu Asn Gln Pro Lys Ala Val
1 5
<210> 37
<211> 9
<212> PRT
<213> 人工序列
<220>
<223> TAA
<400> 37
Gly Leu Ala Glu Phe Gln Glu Asn Val

<212> PRT
<213> 人工序列
<220>
<223> TAA
<400> 42
Ser Val Leu Glu Lys Glu Ile Tyr Ser Ile
1 5 10
<210> 43
<211> 11
<212> PRT
<213> 人工序列
<220>
<223> TAA
<400> 43
Arg Val Ile Asp Asp Ser Leu Val Val Gly Val
1 5 10
<210> 44
<211> 9
<212> PRT
<213> 人工序列
<220>
<223> TAA
<400> 44
Val Leu Phe Gly Glu Leu Pro Ala Leu
1 5
<210> 45
<211> 9
<212> PRT
<213> 人工序列
<220>
<223> TAA
<400> 45
Gly Leu Val Asp Ile Met Val His Leu
1 5
<210> 46
<211> 9
<212> PRT
<213> 人工序列
<220>

<223> TAA
<400> 46
Phe Leu Asn Ala Ile Glu Thr Ala Leu
1 5
<210> 47
<211> 9
<212> PRT
<213> 人工序列
<220>
<223> TAA
<400> 47
Ala Leu Leu Gln Ala Leu Met Glu Leu
1 5
<210> 48
<211> 9
<212> PRT
<213> 人工序列
<220>
<223> TAA
<400> 48
Ala Leu Ser Ser Ser Gln Ala Glu Val
1 5
<210> 49
<211> 11
<212> PRT
<213> 人工序列
<220>
<223> TAA
<400> 49
Ser Leu Ile Thr Gly Gln Asp Leu Leu Ser Val
1 5 10
<210> 50
<211> 9
<212> PRT
<213> 人工序列
<220>
<223> TAA
<400> 50
Gln Leu Ile Glu Lys Asn Trp Leu Leu

1 5
 <210> 51
 <211> 9
 <212> PRT
 <213> 人工序列
 <220>
 <223> TAA
 <400> 51
 Leu Leu Asp Pro Lys Thr Ile Phe Leu
 1 5
 <210> 52
 <211> 9
 <212> PRT
 <213> 人工序列
 <220>
 <223> TAA
 <400> 52
 Arg Leu His Asp Glu Asn Ile Leu Leu
 1 5
 <210> 53
 <211> 9
 <212> PRT
 <213> 人工序列
 <220>
 <223> TAA
 <400> 53
 Gly Leu Pro Ser Ala Thr Thr Thr Val
 1 5
 <210> 54
 <211> 11
 <212> PRT
 <213> 人工序列
 <220>
 <223> TAA
 <400> 54
 Gly Leu Leu Pro Ser Ala Glu Ser Ile Lys Leu
 1 5 10
 <210> 55
 <211> 9

<212> PRT
<213> 人工序列
<220>
<223> TAA
<400> 55
Lys Thr Ala Ser Ile Asn Gln Asn Val
1 5
<210> 56
<211> 9
<212> PRT
<213> 人工序列
<220>
<223> PRAME-004
<400> 56
Ser Leu Leu Gln His Leu Ile Gly Leu
1 5
<210> 57
<211> 9
<212> PRT
<213> 人工序列
<220>
<223> TAA
<400> 57
Ser Leu Leu Gln His Leu Ile Gly Leu
1 5
<210> 58
<211> 9
<212> PRT
<213> 人工序列
<220>
<223> TAA
<400> 58
Ser Leu Leu Gln His Leu Ile Gly Leu
1 5
<210> 59
<211> 9
<212> PRT
<213> 人工序列
<220>

<223> TAA
<400> 59
Lys Val Trp Ser Asp Val Thr Pro Leu
1 5
<210> 60
<211> 11
<212> PRT
<213> 人工序列
<220>
<223> TAA
<400> 60
Leu Leu Trp Gly His Pro Arg Val Ala Leu Ala
1 5 10
<210> 61
<211> 9
<212> PRT
<213> 人工序列
<220>
<223> TAA
<400> 61
Val Leu Asp Gly Lys Val Ala Val Val
1 5
<210> 62
<211> 9
<212> PRT
<213> 人工序列
<220>
<223> TAA
<400> 62
Gly Leu Leu Gly Lys Val Thr Ser Val
1 5
<210> 63
<211> 9
<212> PRT
<213> 人工序列
<220>
<223> TAA
<400> 63
Lys Met Ile Ser Ala Ile Pro Thr Leu

<212> PRT
<213> 人工序列
<220>
<223> TAA
<400> 68
Lys Ile Trp Glu Glu Leu Ser Val Leu Glu Val
1 5 10
<210> 69
<211> 9
<212> PRT
<213> 人工序列
<220>
<223> TAA
<400> 69
Leu Leu Ile Pro Phe Thr Ile Phe Met
1 5
<210> 70
<211> 10
<212> PRT
<213> 人工序列
<220>
<223> TAA
<400> 70
Ile Ser Leu Asp Glu Val Ala Val Ser Leu
1 5 10
<210> 71
<211> 10
<212> PRT
<213> 人工序列
<220>
<223> TAA
<400> 71
Lys Ile Ser Asp Phe Gly Leu Ala Thr Val
1 5 10
<210> 72
<211> 11
<212> PRT
<213> 人工序列
<220>

<223> TAA
<400> 72
Lys Leu Ile Gly Asn Ile His Gly Asn Glu Val
1 5 10
<210> 73
<211> 9
<212> PRT
<213> 人工序列
<220>
<223> TAA
<400> 73
Ile Leu Leu Ser Val Leu His Gln Leu
1 5
<210> 74
<211> 9
<212> PRT
<213> 人工序列
<220>
<223> TAA
<400> 74
Leu Asp Ser Glu Ala Leu Leu Thr Leu
1 5
<210> 75
<211> 13
<212> PRT
<213> 人工序列
<220>
<223> TAA
<400> 75
Val Leu Gln Glu Asn Ser Ser Asp Tyr Gln Ser Asn Leu
1 5 10
<210> 76
<211> 11
<212> PRT
<213> 人工序列
<220>
<223> TAA
<400> 76
His Leu Leu Gly Glu Gly Ala Phe Ala Gln Val

<212> PRT
<213> 人工序列
<220>
<223> TAA
<400> 81
Phe Thr Ala Glu Phe Leu Glu Lys Val
1 5
<210> 82
<211> 9
<212> PRT
<213> 人工序列
<220>
<223> TAA
<400> 82
Ala Leu Tyr Gly Asn Val Gln Gln Val
1 5
<210> 83
<211> 9
<212> PRT
<213> 人工序列
<220>
<223> TAA
<400> 83
Leu Phe Gln Ser Arg Ile Ala Gly Val
1 5
<210> 84
<211> 11
<212> PRT
<213> 人工序列
<220>
<223> TAA
<400> 84
Ile Leu Ala Glu Glu Pro Ile Tyr Ile Arg Val
1 5 10
<210> 85
<211> 9
<212> PRT
<213> 人工序列
<220>

<212> PRT
<213> 人工序列
<220>
<223> TAA
<400> 94
Gly Val Tyr Asp Gly Arg Glu His Thr Val
1 5 10
<210> 95
<211> 10
<212> PRT
<213> 人工序列
<220>
<223> TAA
<400> 95
Gly Leu Tyr Pro Val Thr Leu Val Gly Val
1 5 10
<210> 96
<211> 9
<212> PRT
<213> 人工序列
<220>
<223> TAA
<400> 96
Ala Leu Leu Ser Ser Val Ala Glu Ala
1 5
<210> 97
<211> 9
<212> PRT
<213> 人工序列
<220>
<223> TAA
<400> 97
Thr Leu Leu Glu Gly Ile Ser Arg Ala
1 5
<210> 98
<211> 9
<212> PRT
<213> 人工序列
<220>

<223> TAA
<400> 98
Ser Leu Ile Glu Glu Ser Glu Glu Leu
1 5
<210> 99
<211> 9
<212> PRT
<213> 人工序列
<220>
<223> TAA
<400> 99
Ala Leu Tyr Val Gln Ala Pro Thr Val
1 5
<210> 100
<211> 10
<212> PRT
<213> 人工序列
<220>
<223> TAA
<400> 100
Lys Leu Ile Tyr Lys Asp Leu Val Ser Val
1 5 10
<210> 101
<211> 9
<212> PRT
<213> 人工序列
<220>
<223> TAA
<400> 101
Ile Leu Gln Asp Gly Gln Phe Leu Val
1 5
<210> 102
<211> 9
<212> PRT
<213> 人工序列
<220>
<223> TAA
<400> 102
Ser Leu Leu Asp Tyr Glu Val Ser Ile

<212> PRT
<213> 人工序列
<220>
<223> TAA
<400> 107
Ser Glu Trp Gly Ser Pro His Ala Ala Val Pro
1 5 10
<210> 108
<211> 9
<212> PRT
<213> 人工序列
<220>
<223> TAA
<400> 108
Ala Leu Ser Glu Leu Glu Arg Val Leu
1 5
<210> 109
<211> 9
<212> PRT
<213> 人工序列
<220>
<223> TAA
<400> 109
Ser Leu Phe Glu Ser Leu Glu Tyr Leu
1 5
<210> 110
<211> 9
<212> PRT
<213> 人工序列
<220>
<223> TAA
<400> 110
Lys Val Leu Glu Tyr Val Ile Lys Val
1 5
<210> 111
<211> 10
<212> PRT
<213> 人工序列
<220>

<223> TAA
<400> 111
Val Leu Leu Asn Glu Ile Leu Glu Gln Val
1 5 10
<210> 112
<211> 9
<212> PRT
<213> 人工序列
<220>
<223> TAA
<400> 112
Ser Leu Leu Asn Gln Pro Lys Ala Val
1 5
<210> 113
<211> 9
<212> PRT
<213> 人工序列
<220>
<223> TAA
<400> 113
Lys Met Ser Glu Leu Gln Thr Tyr Val
1 5
<210> 114
<211> 11
<212> PRT
<213> 人工序列
<220>
<223> TAA
<400> 114
Ala Leu Leu Glu Gln Thr Gly Asp Met Ser Leu
1 5 10
<210> 115
<211> 11
<212> PRT
<213> 人工序列
<220>
<223> TAA
<400> 115
Val Ile Ile Lys Gly Leu Glu Glu Ile Thr Val

1	5	10
<210> 116		
<211> 9		
<212> PRT		
<213> 人工序列		
<220>		
<223> TAA		
<400> 116		
Lys Gln Phe Glu Gly Thr Val Glu Ile		
1	5	
<210> 117		
<211> 9		
<212> PRT		
<213> 人工序列		
<220>		
<223> TAA		
<400> 117		
Lys Leu Gln Glu Glu Ile Pro Val Leu		
1	5	
<210> 118		
<211> 9		
<212> PRT		
<213> 人工序列		
<220>		
<223> TAA		
<400> 118		
Gly Leu Ala Glu Phe Gln Glu Asn Val		
1	5	
<210> 119		
<211> 9		
<212> PRT		
<213> 人工序列		
<220>		
<223> TAA		
<400> 119		
Asn Val Ala Glu Ile Val Ile His Ile		
1	5	
<210> 120		
<211> 9		

<212> PRT
<213> 人工序列
<220>
<223> TAA
<400> 120
Ala Leu Ala Gly Ile Val Thr Asn Val
1 5
<210> 121
<211> 12
<212> PRT
<213> 人工序列
<220>
<223> TAA
<400> 121
Asn Leu Leu Ile Asp Asp Lys Gly Thr Ile Lys Leu
1 5 10
<210> 122
<211> 10
<212> PRT
<213> 人工序列
<220>
<223> TAA
<400> 122
Val Leu Met Gln Asp Ser Arg Leu Tyr Leu
1 5 10
<210> 123
<211> 9
<212> PRT
<213> 人工序列
<220>
<223> MAG-003
<400> 123
Lys Val Leu Glu His Val Val Arg Val
1 5
<210> 124
<211> 9
<212> PRT
<213> 人工序列
<220>

1	5	10
<210> 129		
<211> 10		
<212> PRT		
<213> 人工序列		
<220>		
<223> TAA		
<400> 129		
Lys Leu Ile Asp His Gln Gly Leu Tyr Leu		
1	5	10
<210> 130		
<211> 12		
<212> PRT		
<213> 人工序列		
<220>		
<223> TAA		
<400> 130		
Gly Pro Gly Ile Phe Pro Pro Pro Pro Pro Gln Pro		
1	5	10
<210> 131		
<211> 9		
<212> PRT		
<213> 人工序列		
<220>		
<223> TAA		
<400> 131		
Ala Leu Asn Glu Ser Leu Val Glu Cys		
1	5	
<210> 132		
<211> 9		
<212> PRT		
<213> 人工序列		
<220>		
<223> TAA		
<400> 132		
Gly Leu Ala Ala Leu Ala Val His Leu		
1	5	
<210> 133		
<211> 9		

<212> PRT
<213> 人工序列
<220>
<223> TAA
<400> 133
Leu Leu Leu Glu Ala Val Trp His Leu
1 5
<210> 134
<211> 9
<212> PRT
<213> 人工序列
<220>
<223> TAA
<400> 134
Ser Ile Ile Glu Tyr Leu Pro Thr Leu
1 5
<210> 135
<211> 9
<212> PRT
<213> 人工序列
<220>
<223> TAA
<400> 135
Thr Leu His Asp Gln Val His Leu Leu
1 5
<210> 136
<211> 9
<212> PRT
<213> 人工序列
<220>
<223> TAA
<400> 136
Ser Leu Leu Met Trp Ile Thr Gln Cys
1 5
<210> 137
<211> 11
<212> PRT
<213> 人工序列
<220>

<223> TAA
<400> 137
Phe Leu Leu Asp Lys Pro Gln Asp Leu Ser Ile
1 5 10
<210> 138
<211> 10
<212> PRT
<213> 人工序列
<220>
<223> TAA
<400> 138
Tyr Leu Leu Asp Met Pro Leu Trp Tyr Leu
1 5 10
<210> 139
<211> 9
<212> PRT
<213> 人工序列
<220>
<223> TAA
<400> 139
Gly Leu Leu Asp Cys Pro Ile Phe Leu
1 5
<210> 140
<211> 9
<212> PRT
<213> 人工序列
<220>
<223> TAA
<400> 140
Val Leu Ile Glu Tyr Asn Phe Ser Ile
1 5
<210> 141
<211> 11
<212> PRT
<213> 人工序列
<220>
<223> TAA
<400> 141
Thr Leu Tyr Asn Pro Glu Arg Thr Ile Thr Val

<212> PRT
<213> 人工序列
<220>
<223> TAA
<400> 146
Phe Leu Leu Asp Gly Ser Ala Asn Val
1 5
<210> 147
<211> 10
<212> PRT
<213> 人工序列
<220>
<223> TAA
<400> 147
Ala Leu Asp Pro Ser Gly Asn Gln Leu Ile
1 5 10
<210> 148
<211> 9
<212> PRT
<213> 人工序列
<220>
<223> TAA
<400> 148
Ile Leu Ile Lys His Leu Val Lys Val
1 5
<210> 149
<211> 9
<212> PRT
<213> 人工序列
<220>
<223> TAA
<400> 149
Val Leu Leu Asp Thr Ile Leu Gln Leu
1 5
<210> 150
<211> 9
<212> PRT
<213> 人工序列
<220>

1	5	10
<210> 155		
<211> 9		
<212> PRT		
<213> 人工序列		
<220>		
<223> TAA		
<400> 155		
Gly Leu Ala Ser Arg Ile Leu Asp Ala		
1	5	
<210> 156		
<211> 9		
<212> PRT		
<213> 人工序列		
<220>		
<223> TAA		
<400> 156		
Ala Leu Ser Val Leu Arg Leu Ala Leu		
1	5	
<210> 157		
<211> 9		
<212> PRT		
<213> 人工序列		
<220>		
<223> TAA		
<400> 157		
Ser Tyr Val Lys Val Leu His His Leu		
1	5	
<210> 158		
<211> 9		
<212> PRT		
<213> 人工序列		
<220>		
<223> TAA		
<400> 158		
Val Tyr Leu Pro Lys Ile Pro Ser Trp		
1	5	
<210> 159		
<211> 9		

<212> PRT

<213> 人工序列

<220>

<223> TAA

<400> 159

Asn Tyr Glu Asp His Phe Pro Leu Leu

1 5

<210> 160

<211> 9

<212> PRT

<213> 人工序列

<220>

<223> TAA

<400> 160

Val Tyr Ile Ala Glu Leu Glu Lys Ile

1 5

<210> 161

<211> 12

<212> PRT

<213> 人工序列

<220>

<223> TAA

<400> 161

Val His Phe Glu Asp Thr Gly Lys Thr Leu Leu Phe

1 5 10

<210> 162

<211> 9

<212> PRT

<213> 人工序列

<220>

<223> TAA

<400> 162

Val Leu Ser Pro Phe Ile Leu Thr Leu

1 5

<210> 163

<211> 9

<212> PRT

<213> 人工序列

<220>

<223> TAA
<400> 163
His Leu Leu Glu Gly Ser Val Gly Val
1 5
<210> 164
<211> 9
<212> PRT
<213> 人工序列
<220>
<223> 目标肽
<400> 164
Val Leu Leu His His Gln Ile Gly Leu
1 5
<210> 165
<211> 9
<212> PRT
<213> 人工序列
<220>
<223> 目标相似肽 1
<400> 165
Ala Leu Met Tyr His Thr Ile Thr Leu
1 5
<210> 166
<211> 9
<212> PRT
<213> 人工序列
<220>
<223> 目标相似肽 2
<400> 166
Leu Leu Leu Ala His Ile Ile Ala Leu
1 5
<210> 167
<211> 9
<212> PRT
<213> 人工序列
<220>
<223> 无关肽
<400> 167
Ala Ile Val Asp Lys Val Pro Ser Val

1

5

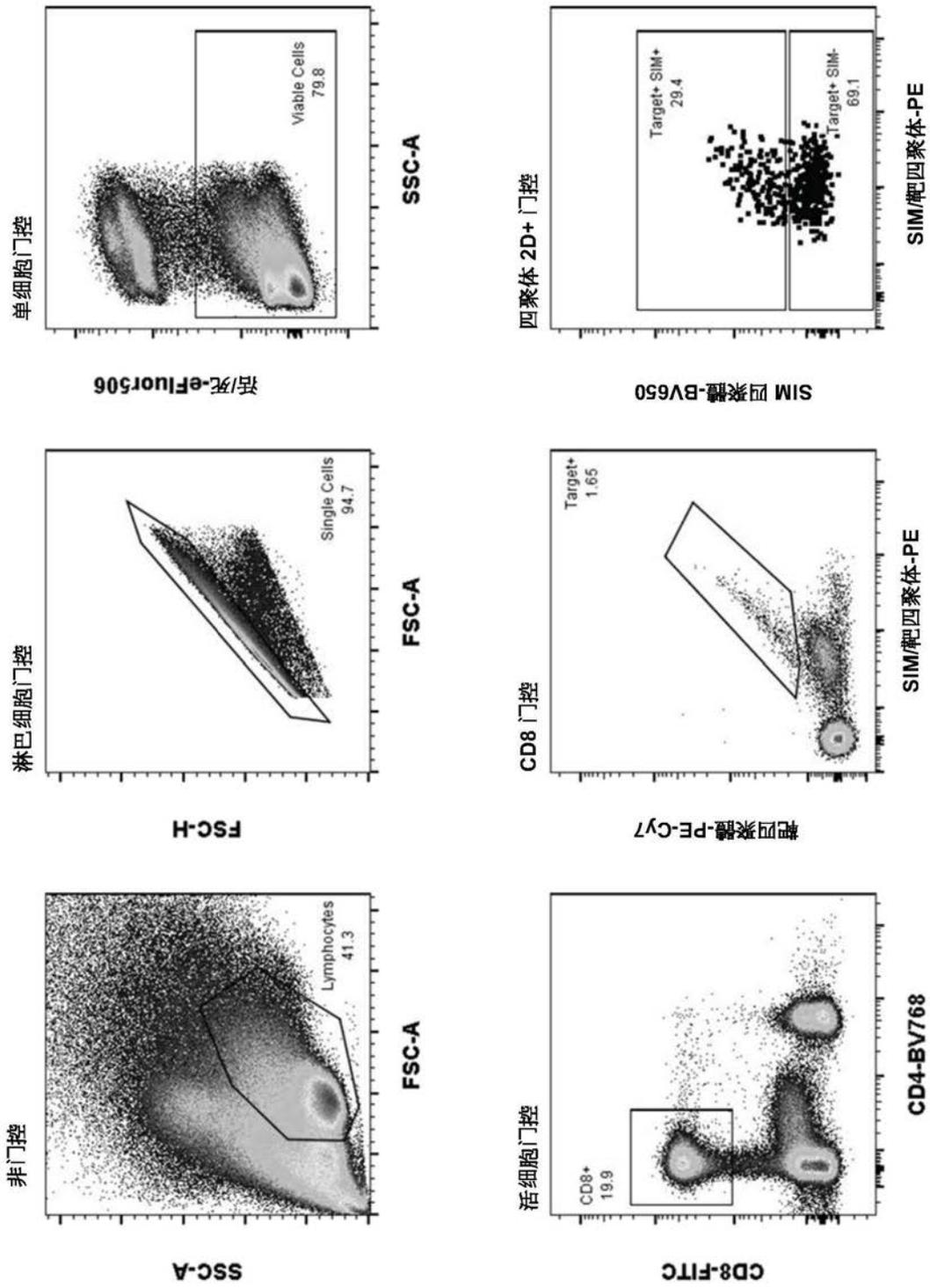


图2

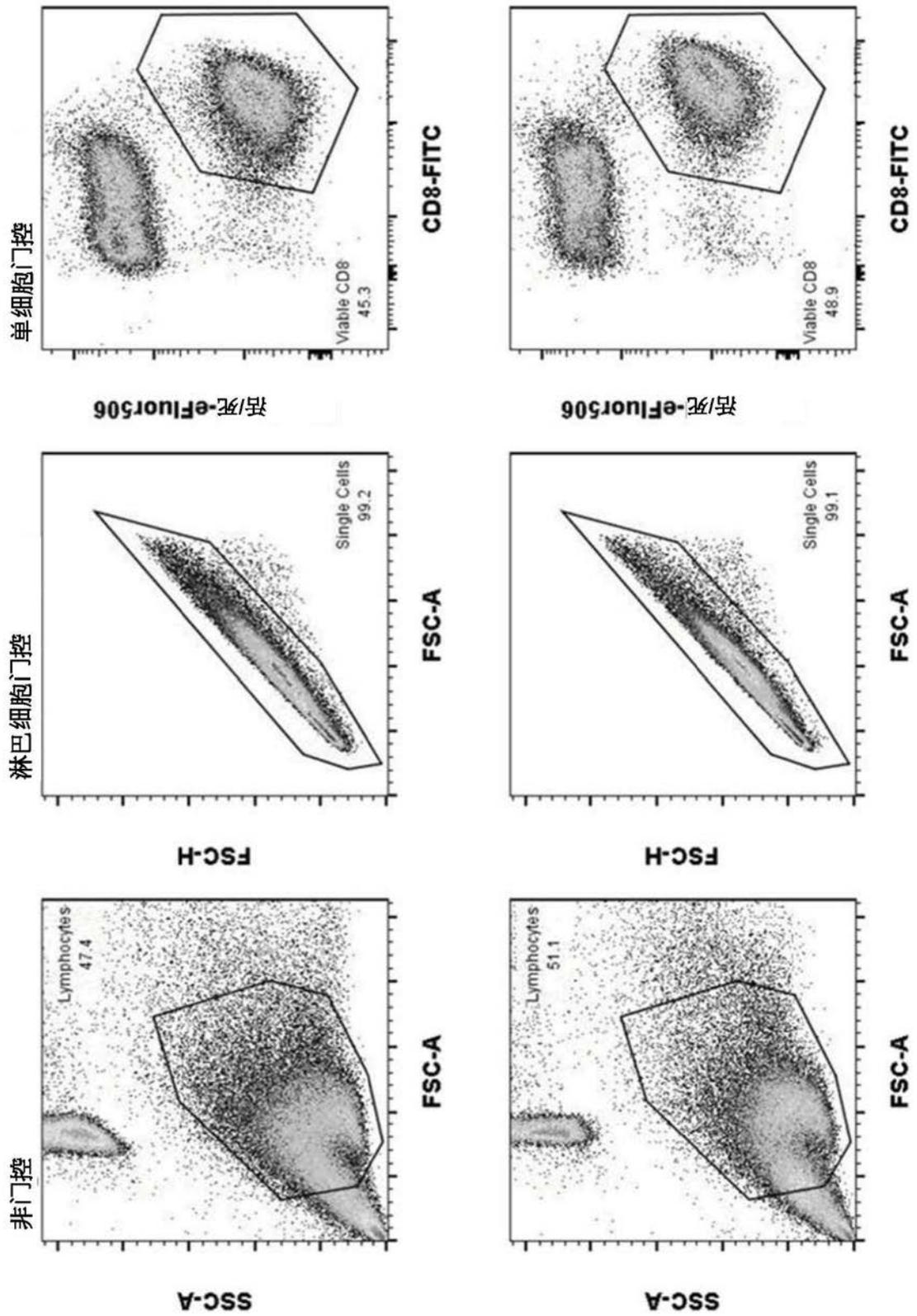


图3

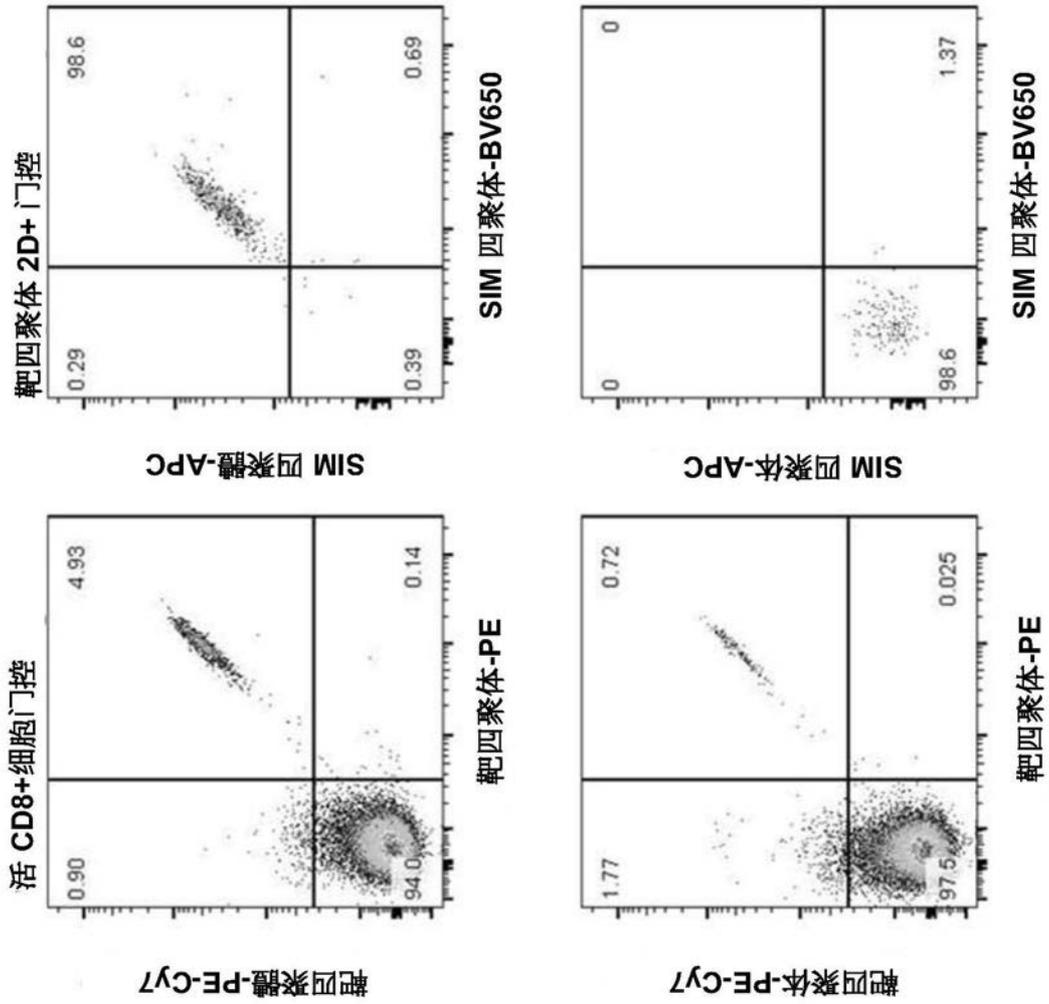


图3 (续)

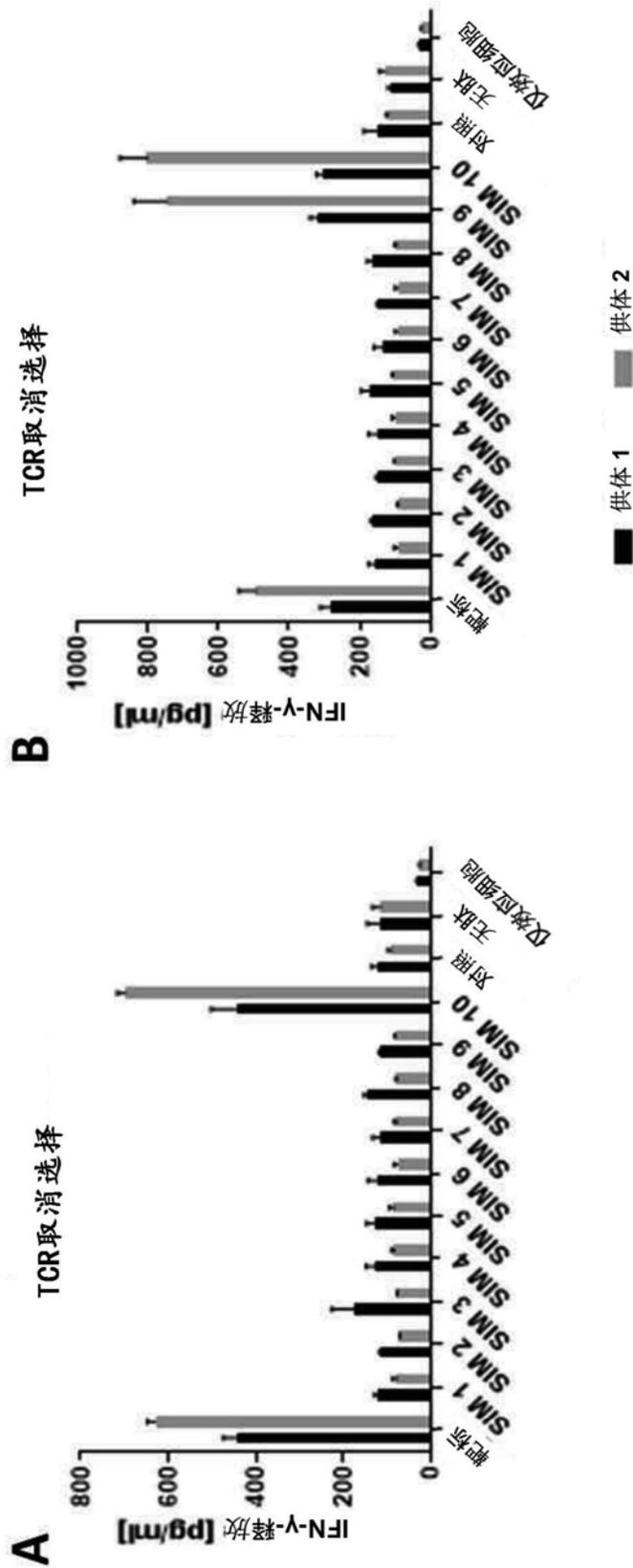


图4

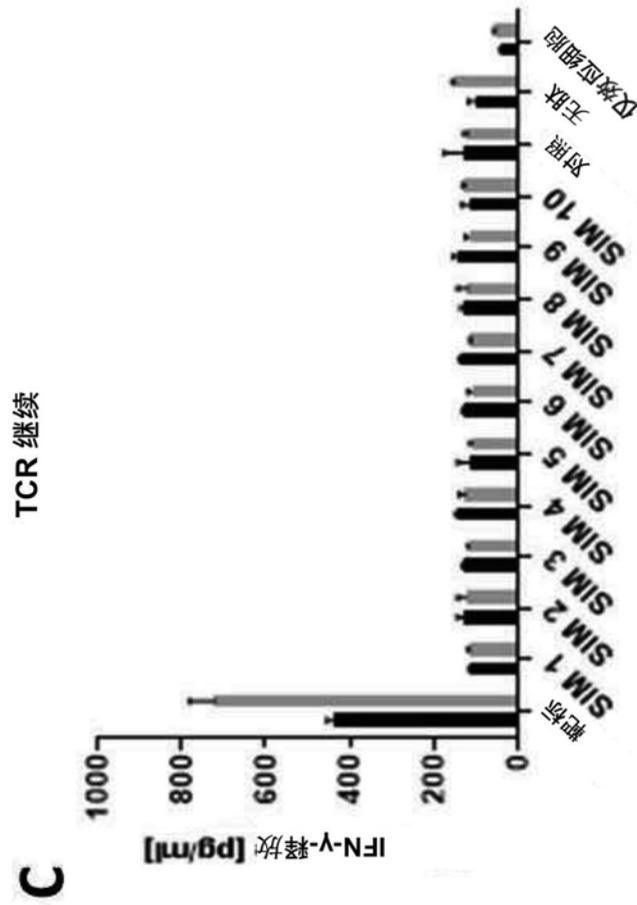


图4(续)

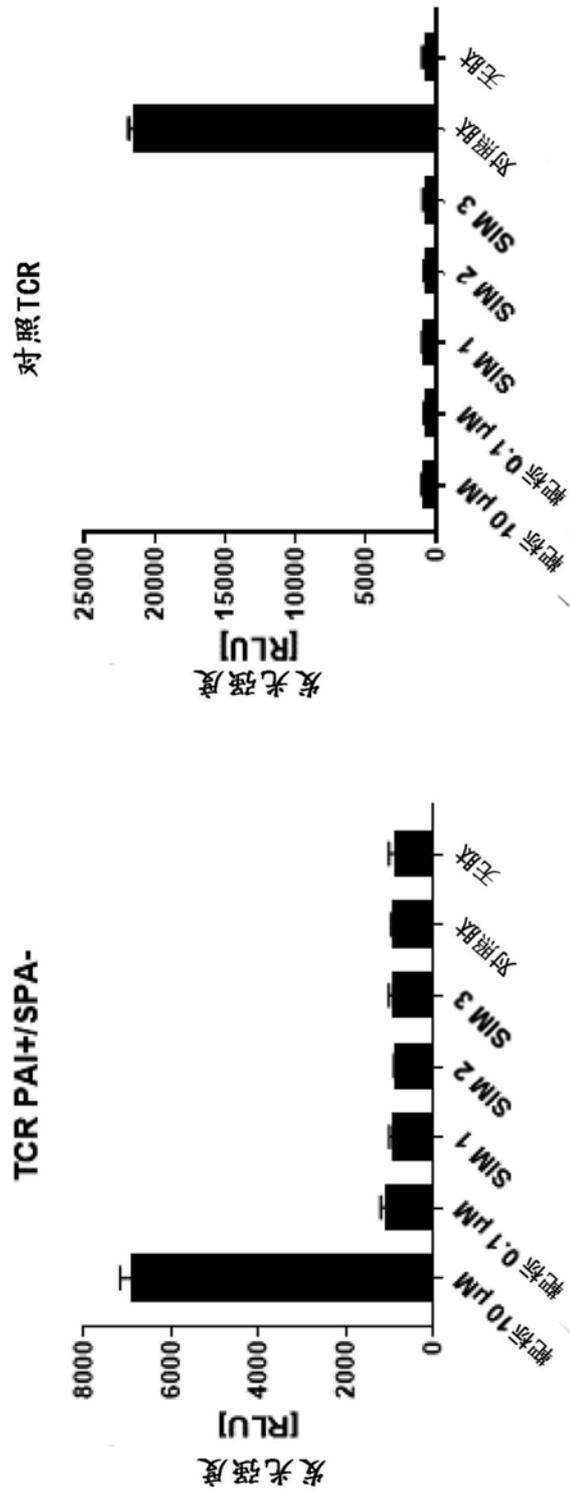


图5

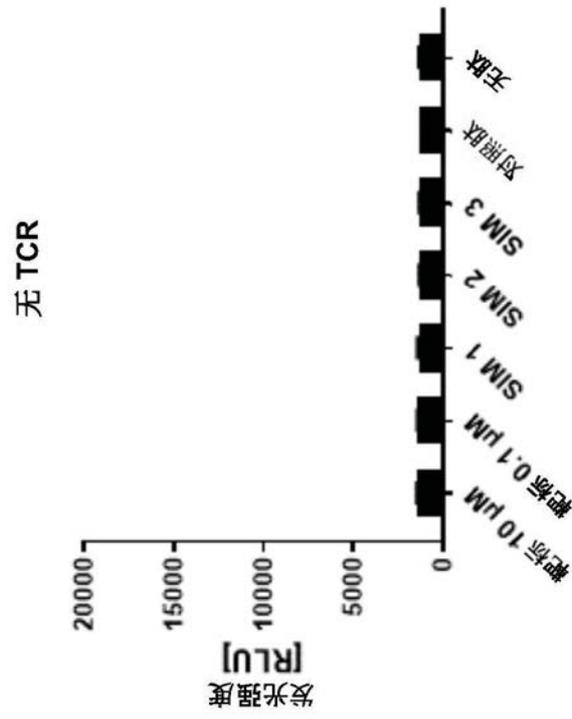


图5 (续)

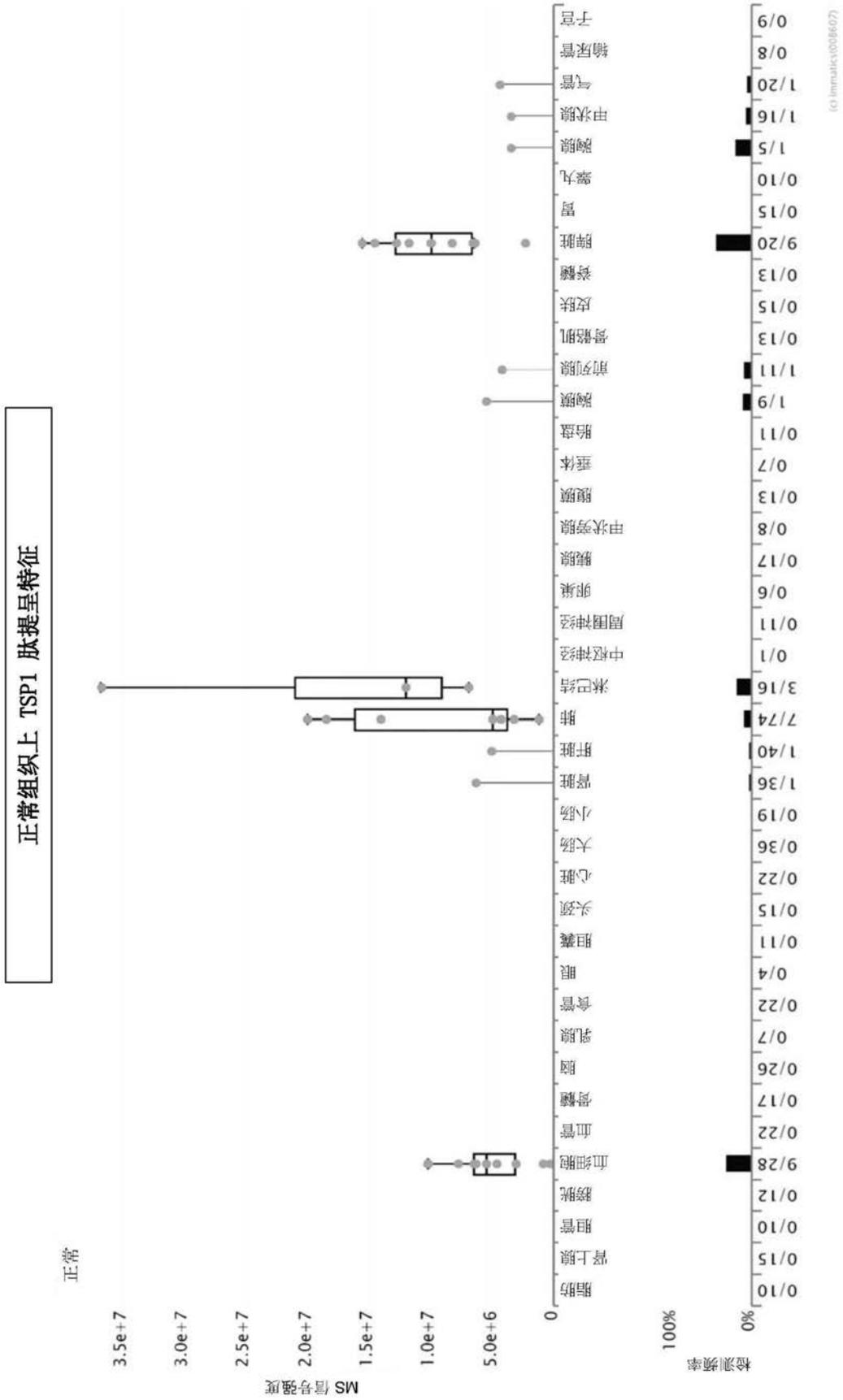


图6

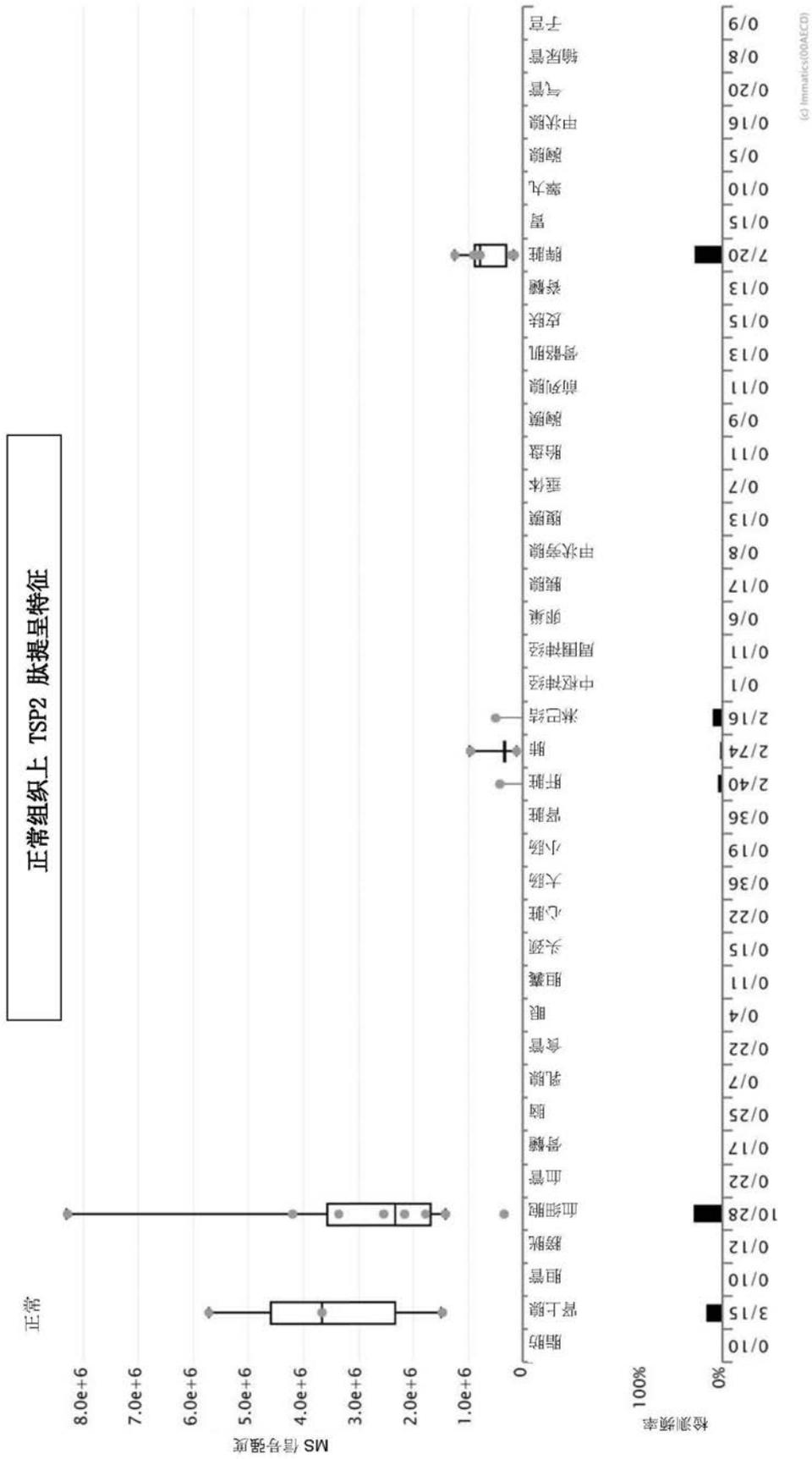


图7

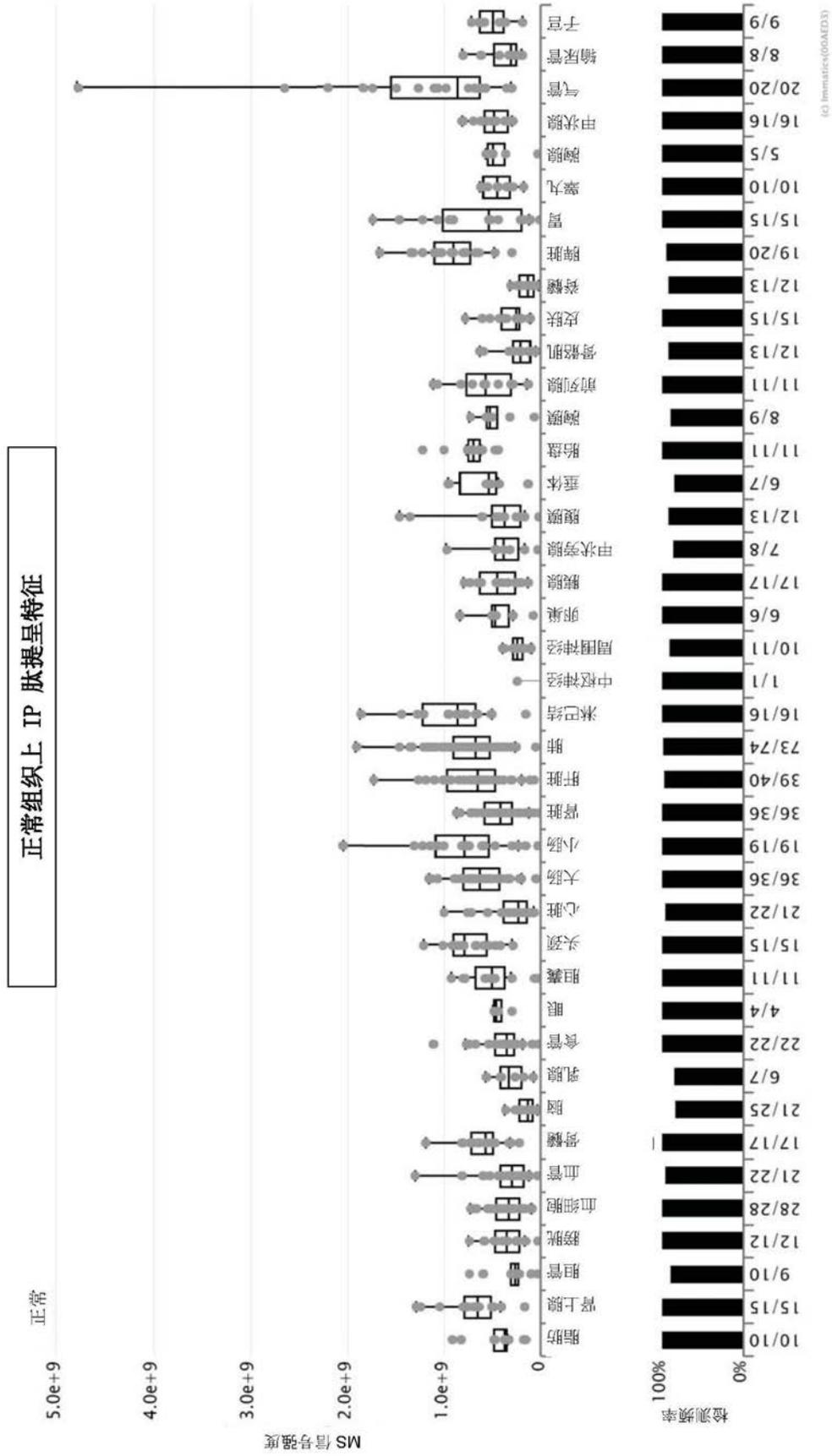


图8