



(12) 发明专利申请

(10) 申请公布号 CN 114502579 A

(43) 申请公布日 2022. 05. 13

(21) 申请号 202080064806.0

(22) 申请日 2020.07.13

(30) 优先权数据

62/874,125 2019.07.15 US

62/975,924 2020.02.13 US

63/024,341 2020.05.13 US

63/034,157 2020.06.03 US

63/039,346 2020.06.15 US

(85) PCT国际申请进入国家阶段日

2022.03.15

(86) PCT国际申请的申请数据

PCT/US2020/041824 2020.07.13

(87) PCT国际申请的公布数据

W02021/011482 EN 2021.01.21

(71) 申请人 荷兰商新基因治疗公司

地址 荷兰阿姆斯特丹

(72) 发明人 安东尼俄斯·尼可拉斯·马瑞亚·

舒马切

卡斯汀·黎奈曼 汤玛士·奎曼

盖文·M·班斗

裘斯·F·C·加迪欧

杰罗恩·W·J·凡·赫斯特

拉奎尔·戈梅兹-鄂兰

黛博拉·苏菲·舒利科玛

(74) 专利代理机构 北京路浩知识产权代理有限公司

公司 11002

专利代理师 刘翠娥 刘成春

(51) Int.Cl.

C07K 14/725 (2006.01)

C12Q 1/6881 (2018.01)

C12N 5/0783 (2010.01)

C12N 15/10 (2006.01)

C40B 20/00 (2006.01)

C40B 30/04 (2006.01)

G16B 25/20 (2019.01)

G16B 30/00 (2019.01)

G16B 40/00 (2019.01)

A61K 39/00 (2006.01)

A61P 35/00 (2006.01)

A61P 37/02 (2006.01)

权利要求书13页 说明书124页

序列表9页 附图82页

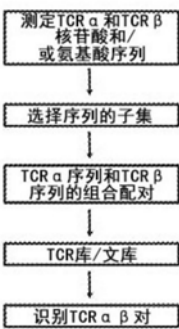
(54) 发明名称

分离TCR基因的方法

(57) 摘要

本发明提供了恢复T细胞受体 (TCR) 库的方法。在一些实施方案中,从非活样本恢复TCR库。在一些实施方案中,创建了TCR α β 对的文库。在一些实施方案中,所提供的方法用于癌症免疫疗法或诊断目的。本文的一些实施方案中描述的是从核酸组合文库中识别编码T细胞受体 α (TCR α) -链和TCR β -链的核苷酸序列的方法。

恢复用于文库生成的TCR库的方法的图示



1. 一种从不同的T细胞群恢复T细胞受体 (TCR) 库的方法,所述方法包括:
测定个体样本中的TCR- α 和TCR- β 的核苷酸序列或氨基酸序列;
从总库中选择TCR α -链和TCR β -链的序列的一个或多个子集;
通过所选的TCR α -链和TCR β -链的序列的组合配对创建TCR库,从而创建TCR $\alpha\beta$ 对的文库;以及
从所创建的TCR库中识别具有期望特征的至少一个TCR $\alpha\beta$ 对。
2. 根据权利要求1所述的方法,其中基于以下至少一个标准来选择来自所述总库的所述TCR α -链和TCR β -链的序列的一个或多个子集:
基于所述T细胞群内的频率,
基于与第二T细胞群相比的相对富集,
基于给定TCR链的DNA拷贝数和RNA拷贝数的相对差异,
基于所述TCR链的生物学特性,其中所述特性选自以下中的至少一项:(预测的)抗原特异性、(预测的)HLA-限制性、抗原-亲和性、共受体依赖性、亲代T细胞谱系(例如CD4或CD8 T细胞)或TCR序列基序,
基于基因表达的空间模式,其中空间基因表达模式衍生自组织中的起始区域或其他基因的共表达模式中的至少一种,
基于在多个样本中共现或以相似频率出现,例如在多个肿瘤病灶中出现,
分配到多个组以分别恢复TCR库的特定部分,
基于在不同实施方案中限定的多个标准的组合。
3. 根据权利要求2所述的方法,其中基于所述T细胞群内的频率的选择是基于TCR序列的所述频率的数据,所述数据用于为TCR α -链和TCR β -链创建单独的秩次序(rank order),或者为TCR α -链和TCR β -链创建组合的秩次序。
4. 根据权利要求1-3中任一项所述的方法,所述方法进一步包括测定频率阈值,所述频率阈值是基于TCR库恢复的期望深度来限定的并且用于基于频率选择TCR α -链和TCR β -链的集合。
5. 根据权利要求1-4中任一项所述的方法,其中测定TCR- α 和TCR- β 的序列是通过以下至少一项来实现的:
多重PCR;
通过靶富集进行TCR序列恢复;
通过5' RACE和PCR进行TCR序列恢复;
通过空间测序进行TCR序列恢复;
通过RNA-seq进行TCR序列恢复;
以及唯一分子标识符(UMI)的使用。
6. 根据权利要求1-5中任一项所述的方法,其中将恢复得到的TCR-链序列定义为CDR3核苷酸序列连同足够的5'-和3'-核苷酸序列信息,以基于核苷酸序列比对选择至少一个TCR V-片段家族和至少一个TCR J-片段家族来组装完整的TCR链序列。
7. 根据权利要求1-6中任一项所述的方法,其中通过所选的TCR α -链和TCR β -链的序列的组合配对来创建TCR库,从而创建TCR $\alpha\beta$ 对的文库是通过以下至少一项来实现的:
使用TCR链序列合成TCR α -链和TCR β -链的DNA或RNA片段的单独的文库,所述TCR α -链和

TCR β -链的DNA或RNA片段随后被连接成一个DNA或RNA片段,其中恰好连接了一个TCR α -链和一个TCR β -链,

TCR α -链和TCR β -链的组合是通过直接合成其中恰好连接了一个TCR α -链和一个TCR β -链的DNA或RNA片段产生的,

TCR α -链和TCR β -链的组合是通过用单独的TCR α -基因和TCR β -基因的集合来修饰细胞池,使得细胞将表达至少一个TCR α -链和至少一个TCR β -链,而在细胞内产生的,和/或

在单链TCR构建体中使TCR α -链和TCR β -链的组合连接,其中TCR α 和TCR β 可变链片段被融合,并且其中所述单链TCR构建体可以被融合到(i)单独的跨膜结构域或(ii)另外含有细胞内信号传导结构域,包括但不限于单独的与CD28信号传导结构域组合的CD3 ϵ 或CD3 ζ 信号传导结构域。

8. 根据权利要求1-7中任一项所述的方法,其中从所述创建的TCR库中识别具有期望特征的至少一个TCR $\alpha\beta$ 对通过以下至少一项来实现:

用生成的TCR $\alpha\beta$ 对的文库修饰的报告细胞池被呈递至少一种目的抗原的抗原呈递细胞刺激,并且基于至少一种活化标志物分离抗原反应性报告细胞以用于TCR分离;

用适于追踪细胞增殖的荧光染料标记用生成的TCR $\alpha\beta$ 对的文库修饰的报告细胞池,所述报告细胞池由表达至少一种目的抗原的抗原呈递细胞刺激,并且基于增殖来分离抗原反应性报告细胞以用于TCR分离;

将用生成的TCR $\alpha\beta$ 对的文库修饰的报告细胞池分成至少两个样本;样本被表达至少一种目的或非目的抗原的抗原呈递细胞刺激;刺激后,将两个报告细胞群孵育一段时间,随后通过TCR分离来分析两个报告细胞群;比较从两个样本中获得的TCR $\alpha\beta$ 对用于识别暴露于至少一种抗原的样本中具有较高丰度的TCR基因;

用生成的TCR $\alpha\beta$ 对的文库修饰的报告细胞池被呈递至少一种目的抗原的抗原呈递细胞刺激,并且基于例如报告TCR触发的NFAT-GFP或NFAT-YFP的至少一种报道基因来分离抗原反应性报告细胞以用于TCR分离;

用生成的TCR $\alpha\beta$ 对的文库修饰的报告细胞池被呈递至少一种目的抗原的抗原呈递细胞刺激,并且基于抗原特异性报告细胞的选择分离抗原反应性报告细胞以用于TCR分离,所述抗原特异性报告细胞的选择基于选择性存活,包括但不限于在TCR信号传导时获得的抗生素抗性,例如通过NFAT-嘌呤霉素转基因的使用;

将用所生成的TCR $\alpha\beta$ 对的文库修饰的报告细胞池暴露于携带目的抗原的一种或多种MHC复合物;将与MHC复合物结合的报告细胞分离以用于TCR分离;

用生成的TCR $\alpha\beta$ 对的文库修饰的报告细胞池被表达至少一种目的抗原的抗原呈递细胞刺激;随后,使用基于单细胞的液滴PCR或微流体方法识别目的TCR $\alpha\beta$ 对,以将TCR分离与至少一种活化标志物的转录物水平的检测相结合;因此,检测到T细胞池中的单个报告细胞,所述单个报告细胞中TCR $\alpha\beta$ 转录物与增加水平的活化标志物共表达。

9. 根据权利要求1-8中任一项所述的方法,其中所述个体样本包含非活起始物料。

10. 根据权利要求1-9中任一项所述的方法,其中恢复所识别的TCR库的限定部分。

11. 根据权利要求1-10中任一项所述的方法,其中恢复抗原特异性TCR序列。

12. 根据权利要求1-11中任一项所述的方法,其中恢复I类和/或II类限制性TCR序列。

13. 根据权利要求1-12中任一项所述的方法,其中恢复以下至少一项:

新生抗原特异性TCR序列

病毒特异性TCR序列

共有肿瘤抗原特异性TCR序列

自身抗原特异性TCR序列。

14. 根据权利要求12所述的方法, 所述方法进一步包括施用表达所述新生抗原特异性TCR序列的T细胞作为癌症疗法的步骤。

15. 根据权利要求1-13中任一项所述的方法, 其中所述方法用于诊断。

16. 根据权利要求15所述的方法, 其中所述诊断为从感染或自身免疫的病理部位恢复TCR库。

17. 根据权利要求1-15中任一项所述的方法, 其中所述方法用于BCR/抗体库的恢复。

18. 根据权利要求1-16中任一项所述的方法, 所述方法进一步包括从个体中分离包含所述TCR- α 和TCR- β 的核苷酸序列的核酸。

19. 根据权利要求8所述的方法, 其中所述活化标志物选自: CD25、CD69、CD62L、CD137、IFN- γ 、IL-2、TNF- α 、GM-CSF、OX40。

20. 根据权利要求1-18中任一项所述的方法, 其中DNA和RNA分离源自T细胞群, 所述T细胞群是不同细胞类型的混合物或组织样本(例如血液或肿瘤组织)的一部分。

21. 根据权利要求1-19中任一项所述的方法, 其中所述个体样本包括从体液中分离的细胞。

22. 根据权利要求20所述的方法, 其中所述细胞是肿瘤特异性T细胞或肿瘤浸润淋巴细胞。

23. 根据权利要求20-21所述的方法, 其中所述体液选自血液、尿液、血清、浆膜液、血浆、淋巴液、脑脊液、唾液、痰液、粘膜分泌物、阴道液、腹水、胸膜液、心包液、腹膜液和腹腔液。

24. 根据权利要求1所述的方法, 所述方法进一步包括使用所述TCR $\alpha\beta$ 链序列来治疗患有癌症、免疫病症、自身免疫性疾病或感染性疾病的个体。

25. 根据权利要求1-7中任一项所述的方法, 其中从所创建的TCR库中识别具有期望特征的至少一个TCR $\alpha\beta$ 对是通过以下至少一项或它们的任意组合来实现的:

基于至少一个活化标志物的识别或选择;

基于响应抗原的增殖的识别或选择;

基于在抗原刺激的细胞中与未刺激的细胞相比对具有更高丰度的TCR基因的识别的识别或选择;

基于TCR触发的报道基因激活的识别或选择;

基于选择性存活的识别或选择, 包括但不限于通过TCR信号传导获得的抗生素抗性;

基于与一种或多种MHC复合物结合的识别或选择;

使用基于单细胞的液滴PCR或微流体的识别或选择。

26. 根据权利要求8所述的方法, 其中通过从大量抗原反应性T细胞中分离DNA或RNA以产生TCR $\alpha\beta$ 特异性PCR产物来实现TCR分离, 所述产物通过DNA测序或RNA测序来分析以确定抗原反应性T细胞的TCR $\alpha\beta$ 基因序列, 或通过基于单细胞的液滴PCR或微流体方法来分析在所分析的单个T细胞中表达的所述TCR $\alpha\beta$ 基因序列。

27. 根据权利要求8所述的方法, 其中所述报告细胞是T细胞。

28. 根据权利要求24所述的方法, 其中使用基于单细胞的液滴PCR或微流体的识别或选择进一步包括活化相关基因的共表达的确定。

29. 一种创建多个T细胞文库的方法, 所述方法包括: 根据权利要求1所述的方法恢复T细胞受体 (TCR) 库; 从总库中选择TCR α -链和TCR β -链的序列到多个组中以分别恢复TCR库的特定部分, 其中创建复杂度较小的或恢复TCR库的特定部分的多个T细胞文库。

30. 根据权利要求29所述的方法, 其中TCR α -链和TCR β -链的序列的选择基于频率范围。

31. 一种从核酸的组合文库中识别核苷酸序列的方法, 所述方法包括:

提供包含多种变体核酸的组合文库, 所述多种变体核酸中的每一个均包含至少600bp的连续部分, 其中所述连续部分包含两个以上变体核苷酸子序列的组合, 其中所述两个以上变体核苷酸子序列的第一变体核苷酸子序列定义所述连续部分的第一端, 并且所述两个以上变体核苷酸子序列的第二变体核苷酸子序列定义所述连续部分的与所述第一端相对的第二端;

将所述文库导入经构造以表达由所述多种变体核酸的成员编码的一种或多种多肽的细胞群中;

基于依赖于所述两个以上变体核苷酸子序列的组合的至少一种功能特性选择所述细胞群的亚群, 其中所述亚群包括多个细胞;

从所述亚群中分离所述多种变体核酸的子集;

测定所述子集的个体成员的所述连续部分的核苷酸序列; 以及

基于所述核苷酸序列识别所述两个以上变体核苷酸子序列的至少一种组合。

32. 根据权利要求1所述的方法, 其中所述一种或多种多肽包括:

T细胞受体 α (TCR α) - 链和TCR β -链;

嵌合抗原受体 (CAR);

开关受体; 或

抗体或其抗原结合片段的一条或多条链。

33. 根据权利要求31或32所述的方法, 其中所述第一变体核苷酸子序列编码TCR α 变体氨基酸序列, 并且所述第二变体核苷酸子序列编码TCR β 变体氨基酸序列。

34. 根据权利要求31至33中任一项所述的方法, 其中所述两种以上变体核苷酸子序列编码以下中的一种或多种: TCR V区、TCR互补决定区3 (CDR3)、TCR J-片段和TCR恒定区。

35. 根据权利要求31或32所述的方法, 其中所述两个以上变体核苷酸子序列各自编码CAR的抗原结合结构域、铰链结构域、跨膜结构域或一个或多个细胞内信号传导结构域。

36. 根据权利要求31-35中任一项所述的方法, 其中所述连续部分的长度为600bp至15000bp。

37. 根据权利要求31-36中任一项所述的方法, 其中提供包括生成所述文库。

38. 根据权利要求37所述的方法, 其中生成所述文库包括:

识别编码所述一种或多种多肽的两组以上变体氨基酸序列的两组以上变体核苷酸子序列, 其中所述至少一种功能特性取决于来自两组以上变体氨基酸序列中每一组的变体氨基酸序列的组合; 以及

通过组合来自两组以上变体核苷酸子序列中每一组的变体核苷酸子序列来组装所述

连续部分,从而产生所述多种变体核酸的成员。

39. 根据权利要求31-38中任一项所述的方法,其中使所述文库中所述多种变体核酸的连续部分之间的编辑距离最大化。

40. 根据权利要求31-39中任一项所述的方法,其中所述文库包含所述两种以上变体核苷酸子序列的至少100种不同组合。

41. 根据权利要求31-40中任一项所述的方法,其中导入包括通过病毒转导、基于转座子的基因递送或核酸酶介导的位点特异性整合导入所述文库。

42. 根据权利要求41所述的方法,其中导入包括以5以下的感染复数(MOI)对所述细胞群进行病毒转导。

43. 根据权利要求31-42中任一项所述的方法,所述方法包括基于所述文库中两种以上变体核苷酸子序列的若干不同组合来调节所述细胞群的大小。

44. 根据权利要求31-43中任一项所述的方法,其中所述细胞群包含永生化T细胞或原代T细胞。

45. 根据权利要求44所述的方法,其中所述永生化T细胞或原代T细胞为人T细胞。

46. 根据权利要求31-45中任一项所述的方法,所述方法进一步包括使用标志物在表达至少一种所述多种变体核酸的所述细胞群中选择或筛选细胞。

47. 根据权利要求46所述的方法,其中所述标志物是细胞毒素抗性标志物和/或细胞表面标志物。

48. 根据权利要求31-47中任一项所述的方法,其中所述细胞群的细胞被遗传修饰。

49. 根据权利要求31-47中任一项所述的方法,其中所述群的细胞不包含赋予所述细胞所述至少一种功能特性的内源多肽。

50. 根据权利要求1至49中任一项所述的方法,其中所述细胞经遗传修饰以消除或减少CD4、CD8和CD28中的一种或多种的表达。

51. 根据权利要求48或49所述的方法,其中用CD4和/或CD8重组所述细胞并被用于筛选I类和/或II类限制性TCR序列。

52. 根据权利要求50或51所述的方法,其中所述细胞为T细胞。

53. 根据权利要求31-52中任一项所述的方法,其中选择包括基于可检测标志物的表达选择所述亚群,其中所述表达取决于所述一种或多种多肽的所述至少一种功能特性。

54. 根据权利要求53所述的方法,其中所述可检测标志物包括细胞表面标志物、细胞因子标志物、细胞增殖标志物、转录报告物、信号转导报告物和/或细胞毒性报告物。

55. 根据权利要求54所述的方法,其中

所述细胞表面标志物包括:CD25、CD69、CD62L、CD137、OX40中的一种或多种;

所述细胞因子标志物包括:IFN- γ 、IL-2、TNF- α 、IL-8、GM-CSF中的一种或多种;

所述转录报告物包括:NF- κ B、NFAT、AP-1中的一种或多种;

所述信号转导报告物包括:ZAP70, ERK1/2中的一种或多种;以及

所述细胞毒性报告物包括:CD107A、CD107B中的一种或多种。

56. 根据权利要求31-55中任一项所述的方法,其中选择包括使所述细胞群与以下所述中的一种或多种接触:

第二细胞群;

所述一种或多种多肽的配体；
所述一种或多种多肽的激动剂或拮抗剂；以及
小分子，

其中所述接触诱导的所述亚群的变化取决于所述一种或多种多肽的所述至少一种功能特性。

57. 根据权利要求56所述的方法，其中选择进一步包括：

检测所述变化的存在与否，和/或所述变化的大小；以及
基于所述检测选择所述亚群。

58. 根据权利要求56或57所述的方法，其中所述第二细胞群包含抗原呈递细胞。

59. 根据权利要求58所述的方法，其中所述抗原呈递细胞包括B-细胞和/或树突细胞。

60. 根据权利要求56-59中任一项所述的方法，其中所述第二细胞群包含原代细胞或永生化细胞。

61. 根据权利要求56-60中任一项所述的方法，其中所述文库的变体核苷酸序列源自表达包含由所述变体核苷酸子序列编码的氨基酸的变体多肽的细胞，其中所述细胞来源于个体，并且其中所述第二细胞群源自所述个体。

62. 根据权利要求31-61中任一项所述的方法，其中选择包括基于高于或低于阈值的所述至少一种功能特性的量度选择所述细胞群的第一亚群。

63. 根据权利要求62所述的方法，其中基于对所述细胞群的未选亚群中的所述功能特性的量度来确定所述阈值。

64. 根据权利要求62所述的方法，其中选择进一步包括基于高于或低于第二阈值的所述至少一种功能特性的第二量度来选择所述细胞群的第二亚群，其中所述第一亚群和第二亚群不重叠。

65. 根据权利要求64所述的方法，其中识别所述至少一个组合包括比较所述第一亚群和第二亚群之间的所述至少一个组合的丰度。

66. 根据权利要求31-64中任一项所述的方法，其中识别所述至少一种组合包括测量所述至少一种组合在所述亚群中相对于对照细胞群的富集。

67. 根据权利要求66所述的方法，其中所述细胞群包括所述对照细胞群。

68. 根据权利要求67所述的方法，其中所述细胞亚群和对照细胞群不重叠，其中不重叠表示两个群中的所述细胞具有不同的活化状态，但可以携带相同的变体核酸。

69. 根据权利要求66所述的方法，其中基于不同于所述至少一种功能特性的第二功能特性来选择所述对照细胞群。

70. 根据权利要求31-69中任一项所述的方法，其中所述亚群包含多个细胞。

71. 根据权利要求31-70中任一项所述的方法，其中所述分离不包括基于所述至少一种功能特性来分离所述亚群的单个克隆。

72. 根据权利要求31-71中任一项所述的方法，其中所述亚群包含至少1000个细胞。

73. 根据权利要求31-72中任一项所述的方法，其中测定包括通过生成连续部分的至少600bp的测序读段来对所述个体成员进行测序。

74. 根据权利要求73所述的方法，其中所述测序读段长度在600bp与15000bp之间。

75. 根据权利要求31-74中任一项所述的方法，其中测定包括至少扩增所述个体成员的

所述连续部分。

76. 根据权利要求75所述的方法, 其中扩增包括使用与不变核苷酸子序列杂交的扩增引物, 其中所述多种变体核酸中的每一个均包含所述不变核苷酸子序列, 并且其中所述不变核苷酸子序列编码所述一种或多种多肽的不变氨基酸序列。

77. 根据权利要求76所述的方法, 其中所述一种或多种多肽包含TCR α -链和TCR β -链, 并且其中所述不变氨基酸序列包含TCR α 恒定区。

78. 根据权利要求76所述的方法, 其中所述一种或多种多肽包含TCR α -链和TCR β -链, 并且其中所述不变氨基酸序列包含TCR β 恒定区。

79. 根据权利要求31-78中任一项所述的方法, 其中所述分离包括使用CRISPR/Cas9介导的来自所述亚群的基因组DNA的靶向片段化。

80. 根据权利要求31-78中任一项所述的方法, 其中所述测定包括在所述个体成员的任何扩增之前或不存在所述个体成员的任何扩增的情况下, 对于所述连续部分的每个核苷酸序列获得至少10的平均覆盖率。

81. 根据权利要求31-80中任一项所述的方法, 其中所述测定包括对于所述连续部分的每个核苷酸序列, 获得至少25、至少50、至少100、至少200、至少300、至少400、至少500或至少1000的平均覆盖率。

82. 一种从核酸组合文库中识别编码T细胞受体 α (TCR α)-链和TCR β -链的核苷酸序列的方法, 其包括:

提供包含多种变体核酸的文库, 所述多种变体核酸中的每一个均包含至少600bp的连续部分, 其中所述连续部分包含:

以下的组合:

第一变体核苷酸子序列, 其编码TCR α 变体氨基酸序列并限定所述连续部分的第一端, 和

第二变体核苷酸子序列, 其编码TCR β 变体氨基酸序列并限定与所述第一端相对的所述连续部分的第二端;

将所述文库导入经构造以表达由所述多种变体核酸的成员编码的TCR α -链和TCR β -链的永生化T细胞群中;

基于响应于使永生化T细胞与表达抗原的永生化B细胞接触的高于阈值水平的T细胞活化标志物的表达选择所述永生化T细胞群的亚群, 其中所述亚群包括多个T细胞;

从所述亚群中分离所述多种变体核酸的子集;

测定所述子集的个体成员的所述连续部分的核苷酸序列; 和

基于所述子集的核苷酸序列中的至少一种组合相对于对照的富集, 识别所述第一变体核苷酸子序列和第二变体核苷酸子序列的至少一种组合。

83. 根据权利要求82所述的方法, 其进一步包括:

基于响应于使所述永生化T细胞与所述永生化B细胞接触的低于第二阈值水平的T细胞活化标志物的表达选择所述永生化T细胞群的第二亚群, 其中所述第二亚群包括多个第二T细胞, 并且其中所述亚群和第二亚群不重叠;

从所述第二亚群中分离所述多种变体核酸的第二子集; 和

测定所述第二子集的个体成员的所述连续部分的第二核苷酸序列,

其中基于所述子集中的所述至少一种组合相对于所述第二子集的所述第二核苷酸序列中的所述至少一种组合的富集来识别所述至少一种组合。

84. 一种从核酸组合文库中识别编码嵌合抗原受体 (CAR) 铰链结构域、跨膜结构域和/或细胞内信号传导结构域的核苷酸序列的方法, 其包括:

提供包含多种变体核酸的文库, 所述多种变体核酸中的每一个均包含至少600bp的连续部分, 其中所述连续部分包含以下两种以上的组合:

编码CAR铰链结构域的第一变体核苷酸子序列;

编码CAR跨膜结构域的第二变体核苷酸子序列; 和

编码CAR细胞内信号传导结构域的第三变体核苷酸子序列,

其中所述第一、第二或第三变体核苷酸子序列中的一个限定了所述连续部分的第一端, 并且其中所述第一、第二或第三变体核苷酸子序列中的另一个限定了与所述第一端相对的所述连续部分的第二端;

将所述文库导入经构造以表达由所述多种变体核酸的成员编码的CAR的细胞群中, 其中所述细胞群包含永生化T细胞群或原代人T细胞群;

响应于使所述细胞与表达对所述CAR的抗原结合结构域特异的抗原的抗原呈递细胞接触, 基于高于阈值水平的细胞增殖选择所述细胞群的亚群, 其中所述亚群包括多个细胞;

从所述亚群中分离所述多种变体核酸的子集;

测定所述子集的个体成员的所述连续部分的核苷酸序列; 和

基于所述子集的所述核苷酸序列中的至少一种组合相对于对照的富集, 识别所述第一、第二和第三变体核苷酸子序列的至少一种组合。

85. 根据权利要求84所述的方法, 其中存在多于一个的CAR细胞内信号传导结构域。

86. 根据权利要求85所述的方法, 其中存在至少两个CAR细胞内信号传导结构域。

87. 根据权利要求84所述的方法, 其中所述至少两个CAR细胞内信号传导结构域是以下中的至少两项: CD3 ϵ 、CD3 ζ ITAM1、CD3 ζ ITAM12、CD3 ζ ITAM 123、CD3 ζ 与CD3 δ 、CD3 ϵ 和CD3 γ 的任意ITAM、CD8 α 、CD28、ICOS、4-1BB (CD137)、OX40 (CD134)、CD27和CD2。

88. 根据权利要求84所述的方法, 其包括:

基于响应于所述细胞与所述抗原呈递细胞的接触的低于第二阈值水平的细胞增殖选择所述细胞群的第二亚群, 其中所述第二亚群包括多个第二细胞, 并且其中所述亚群和第二亚群不重叠;

从所述第二亚群分离所述多种变体核酸的第二子集;

测定所述第二子集的个体成员的所述连续部分的第二核苷酸序列, 和

其中基于所述子集中所述至少一种组合相对于所述第二子集的所述第二核苷酸序列中的所述至少一种组合的富集, 识别所述至少一种组合。

89. 一种从核酸组合文库中识别核苷酸序列的方法, 其包括:

提供包含多种变体核酸的组合文库, 所述多种变体核酸中的每一个均包含至少600bp的连续部分;

将所述文库导入经构造以表达由所述多种变体核酸的成员编码的一种或多种多肽的细胞群中;

基于依赖于所述至少600bp的连续部分的至少一个功能特性来选择所述细胞群的亚

群,其中所述亚群包括多个细胞;

从所述亚群中分离所述多种变体核酸的子集;

测定所述子集的个体成员的所述连续部分的核苷酸序列;和

基于所述核苷酸序列识别所述至少600bp的连续部分。

90.根据权利要求89所述的方法,其中所述至少600bp的连续部分中的可变性分布在600个碱基对中。

91.一种从核酸文库中识别编码抗原特异性T细胞受体 α (TCR α) -链和TCR β -链对的核苷酸序列的方法,其包括:

将文库导入能够表达由多种变体核酸的成员编码的TCR α -链和TCR β -链的细胞群中,

基于响应于抗原的高于阈值水平的标志物的表达来选择所述细胞群的亚群,其中所述亚群包括多个细胞,

从所述亚群中分离所述多种变体核酸的子集,

测定所述变体核酸的核苷酸序列,和

基于所述子集内的所述核苷酸序列相对于对照的富集来识别至少一种变体核苷酸序列。

92.根据权利要求91或90所述的方法,其进一步包括提供包含编码TCR α 链和TCR β 链的所述多种变体核酸的文库。

93.一种从样本中识别编码T细胞受体 α (TCR α) -链和TCR β -链的核苷酸序列的方法,其包括:

对样本中的TCR- α 链和TCR- β 链进行测序,

选择并对TCR α -链和TCR β -链序列进行组合配对以创建TCR $\alpha\beta$ 对的文库,

将所述TCR $\alpha\beta$ 对的文库导入报告细胞池,

用呈递至少一种目的抗原的抗原呈递细胞刺激用所述TCR $\alpha\beta$ 对的文库修饰的所述报告细胞,

确定对所述至少一种目的抗原特异的TCR $\alpha\beta$ 对,和

将所述TCR $\alpha\beta$ 对导入细胞并选择含有所述TCR $\alpha\beta$ 对的细胞。

94.根据权利要求14所述的方法,其进一步包括施用表达所述抗原特异性TCR序列的T细胞以诊断或治疗感染或自身免疫的步骤。

95.根据权利要求14所述的方法,其中所述T细胞可以是自体的或同种异体的。

96.根据权利要求8所述的方法,其中所述活化标志物是CD69,并且其中分离了两个细胞群,一个细胞群具有高表达的CD69并且另一个细胞群具有低表达的CD69。

97.一种核苷酸文库,其包含根据权利要求1-30和95、96中任一项所述恢复得到的T细胞受体库。

98.一种核苷酸构建体,其包含根据权利要求1-96中任一项所述识别得到的核苷酸序列。

99.一种细胞,其包含根据权利要求98所述的核苷酸构建体。

100.一种从核酸文库中识别编码抗原特异性T细胞受体 α (TCR α) -链和TCR β -链对的核苷酸序列的方法,所述方法包括:

将所述核酸文库导入能够表达TCR α -链和TCR β -链的细胞群中,以制备细胞文库;

基于响应于抗原的高于第一阈值水平的标志物的表达来选择所述细胞文库的第一群；
和

从所述文库的所述第一群中分离变体核酸的第一群。

101. 根据权利要求100所述的方法，其进一步包括：

测定变体核酸的所述第一群的至少一个核苷酸序列或核酸同一性；和

基于所述子集内的所述核苷酸序列相对于对照的富集来识别至少一种变体核苷酸序列。

102. 根据权利要求100或101所述的方法，其中所述阈值水平基于以下至少一项：

d) 基于标志物的表达的所述总的细胞池的百分比的恢复；或者

e) 来自所述总的细胞池的最少数量的细胞的恢复；或者

f) 基于磁性探针与响应抗原表达的至少一种标志物的结合，由磁体保留的细胞的恢复。

103. 根据权利要求66-69、82、84、91或101所述的方法，其中所述对照是低于第二阈值的第二细胞群。

104. 根据权利要求66-69、82、84、91或101所述的方法，其中所述对照是以下中的一种或多种：

参考细胞群，

导入所述细胞群的所述核酸组合文库，

基于低于第二阈值的表达标志物从与第一群相同的细胞群中分选出来的细胞群，

从表达相关TCR文库的报告T细胞与不呈递任何外源性抗原的抗原呈递细胞如B细胞的共培养物中获得的至少一种细胞群。

105. 根据权利要求104所述的方法，其中所述对照（或底部样本）选自与顶部样本相同的细胞群，但具有低活化标志物表达，或其中所述底部样本获自表达所述相关TCR文库的报告T细胞和未被工程化以表达外源抗原的B细胞的共培养物。

106. 根据权利要求100-105中任一项所述的方法，其进一步包括向所述细胞群添加抗原。

107. 根据权利要求100-106中任一项所述的方法，其中所述分离第一群和/或所述对照通过以下中的至少一项来实现：a) 磁珠富集，b) 流式细胞术分选或c) 二者。

108. 一种从核酸文库中识别编码T细胞受体 α (TCR α) -链和TCR β -链的核苷酸序列的方法，所述方法包括：

将所述核酸文库导入能够表达TCR α -链和TCR β -链的细胞群中以制备细胞文库；和

基于所述子集内的所述核苷酸序列相对于对照的富集，测定变体核酸的第一群的至少一个核苷酸序列或至少一种核酸的同一性。

109. 根据权利要求108所述的方法，其中所述至少一种核酸从第一细胞群分离。

110. 根据权利要求109所述的方法，其中所述第一细胞群基于响应于抗原的高于第一阈值水平的标志物的表达来选择。

111. 一种从核酸文库中识别核苷酸序列的方法，其包括：

将所述核酸文库导入细胞群以形成细胞文库；

使所述细胞文库与第一细胞群接触；

通过磁珠富集基于至少一种标志物的表达来选择所述细胞文库的亚群；和
基于所述亚群内的所述核苷酸序列相对于对照的统计学上显著的富集或耗竭来识别至少一种核苷酸序列。

112. 根据权利要求111所述的方法，其中所述细胞亚群中的至少一些被构造为表达由所述核酸文库的成员编码的一种或多种多肽。

113. 根据权利要求111所述的方法，其中标志物表达与来自所述文库的导入核酸相关联，其中相关联表示导入的核酸改变标志物表达。

114. 根据权利要求111所述的方法，其中选择基于高于阈值水平的标志物表达。

115. 根据权利要求111所述的方法，其中选择基于高于阈值水平的至少两个标志物的表达。

116. 根据权利要求8、82、91、93或100中任一项所述的方法，其中识别或刺激或提供抗原包括以下中的一种或多种：

- a) 选择若干抗原；
- b) 创建抗原池，其中每个抗原正好存在于两个抗原池中；
- c) 评估表达至少一种T细胞受体的报告细胞对每个抗原池的反应性；和
- d) 通过评估对恰好两个抗原池的反应性，确定所述至少一种T细胞受体是否对任何选定的抗原具有反应性。

117. 根据权利要求116所述的方法，其中通过成对富集分析来检测对恰好两个抗原池的反应性。

118. 根据权利要求111所述的方法，其中所述文库是TCR文库。

119. 根据权利要求111所述的方法，其中使用活化标志物。

120. 根据权利要求111所述的方法，其中使用顶部-底部比较来评估反应性。

121. 根据前述权利要求中任一项所述的方法涉及反应性评估或进一步包括反应性评估，其中使用顶部-底部比较来评估反应性。

122. 根据前述权利要求中任一项所述的方法涉及文库，其中基于以下中的至少一项来对所述文库中的所述多种变体核酸的所述核苷酸序列进行优化：

为宿主细胞引入优选的密码子使用，

mRNA结构稳定性的优化，

重复序列的避免，

均聚物的长拉伸的避免，和

给定变体核酸序列内局部GC含量的较大差异的避免。

123. 根据权利要求91所述的方法，其中所述抗原通过抗原呈递细胞呈递。

124. 根据权利要求91所述的方法，其中所述文库是组合文库。

125. 根据前述权利要求中任一项所述的方法，其中所述抗原由细胞提供。

126. 根据前述权利要求中任一项所述的方法，其中过程涉及高度的抗原多样性和/或复杂性。

127. 前述方法中任一项涉及文库的方法，其中所述文库是组合文库。

128. 如上所述的方法，其中所述组合文库是TCR文库。

129. 一种细胞集合，所述集合包括：

至少两种T细胞的组,其中每种T细胞均被构造为表达至少一个TCR α 和TCR β 对,其中所述TCR α 和TCR β 均来自一个个体,其中所述T细胞不表达内源性TCR,并且其中所述组被构造用于一个或多个T细胞活化标志物的激活;和

至少两种B细胞的组,其中所述至少两种B细胞中的每一种均被构造为表达至少一种外源性新生抗原(或抗原),使得存在至少两种能够被产生的外源性新生抗原(或抗原),并且其中所述至少两种外源性新生抗原(或抗原)与所述个体中的外源性新生抗原(或抗原)相同。

130. 根据权利要求129所述的组合物,其中所述至少两种B细胞的组包括:

至少一种产生所述外源性新生抗原(或抗原)的第一B细胞;和

至少一种产生所述第二外源性新生抗原(或抗原)的第二B细胞。

131. 一种表达TCR的细胞的文库,所述表达TCR的细胞的文库包含:至少三种T细胞的组,

其中所述T细胞中的至少两种被构造为表达至少两个TCR α 和TCR β 对(至少两个TCR对),

其中所述至少两个TCR对来自一个个体,

其中所述至少三种T细胞不表达内源性TCR,

其中所述至少三种T细胞被构造用于在结合由B细胞呈递的抗原(或新生抗原)时一个或多个T细胞活化标志物的激活,

其中反映在若干个TCR细胞中的每个TCR对的基因组拷贝的数量使得可以读取样本中的每个TCR,并且

其中至少一种所述TCR在包含所述文库的整个组合物中分布不均。

132. 根据权利要求131所述的表达TCR的细胞的文库,其中至少一种T细胞的分布通过与B细胞呈递的抗原结合而改变。

133. 根据权利要求131所述的表达TCR的细胞的文库,其中所述至少两个TCR对近似均匀地存在于所述表达TCR的细胞的文库中。

134. 一种治疗个体的方法,所述方法包括:

识别患有肿瘤的个体;

提供至少两种T细胞的组,每种T细胞均被构造为表达至少一种不同的TCR α 和TCR β 对,其中所述TCR α 和所述TCR β 均来自所述个体,

提供至少两种B细胞的组,其中所述B细胞组经构造以表达至少两种外源性新生抗原,并且其中所述至少两种外源性新生抗原与在所述个体中发现的那些新生抗原相同;

将所述至少两种T细胞的组与所述至少两种B细胞的组组合,并基于所述至少两种T细胞经由所述至少两个外源性新生抗原的激活来选择至少两个TCR对的组合;和

向所述个体施用所述至少两个TCR对的组合,从而治疗所述肿瘤。

135. 根据权利要求134所述的方法,其中治疗减小了所述肿瘤的大小。

136. 一种治疗个体的方法,所述方法包括:

识别患有肿瘤的个体;

提供至少两种T细胞的组,每种T细胞均被构造为表达至少一种不同的TCR α 和TCR β 对,其中所述TCR α 和所述TCR β 均来自所述个体,

提供至少两种抗原呈递细胞的组,其中所述一组抗原呈递细胞源自所述个体,被构造

为表达至少两种外源性新生抗原,并且其中所述至少两种外源性新生抗原与在所述个体中发现的那些新生抗原相同;

将所述至少两种T细胞的组与所述至少两种抗原呈递细胞的组组合,并基于所述至少两种T细胞经由所述至少两个外源性新生抗原的激活来选择至少两个TCR对的组合;和向所述个体施用所述至少两个TCR对的组合,从而治疗所述肿瘤。

137. 一种药物组合物,其包含:

第一TCR对,其与个体的肿瘤中的第一抗原(或新生抗原)结合;和

第二TCR对,其与所述个体的肿瘤中的第二抗原(或新生抗原)结合。

138. 根据权利要求137所述的药物组合物,其中所述第一TCR对是MHC I类限制性的,并且其中所述第二TCR对是MHC II类限制性的。

139. 一种药物组合物,其包含:

第一TCR对,其与第一抗原结合并且是MHC I类限制性的;和

第二TCR对,其与第二抗原结合并且是MHC II类限制性的。

140. 根据权利要求137-139中任一项所述的药物组合物,其进一步包含第三TCR对。

141. 根据权利要求137-140中任一项所述的药物组合物,其中所述第一TCR对与来自肿瘤的新生抗原结合,其中所述第二TCR对与来自所述肿瘤的新生抗原结合,并且其中所述第一TCR对和第二TCR对二者均存在于所述肿瘤的宿主中。

142. 一种细胞集合,所述集合包括:

至少两种T细胞的组,其中每种T细胞均被构造为表达至少一个TCR α 和TCR β 对,其中所述对来自一个个体,其中所述T细胞不表达内源性TCR,并且其中所述组被构造用于激活一个或多个T细胞活化标志物;和

至少两种抗原呈递细胞(APC)的组,其中所述至少两种APC中的每一种均被构造为表达至少一种外源性新生抗原(或抗原),使得存在至少两种能够被产生的外源性新生抗原(或抗原),并且其中所述至少两种外源性新生抗原(或抗原)与所述个体中的外源性新生抗原(或抗原)相同。

143. 如上述方法中任一种所述的方法,其中,通过与抗原(例如新生抗原)结合来选择或识别所述TCR对和/或表达所述TCR对的所述T细胞,其中所述抗原由B细胞或抗原呈递细胞表达。

144. 如上述方法中任一种所述的方法,其中所述抗原或新生抗原来自一个个体的肿瘤,并且其中所述TCR对的TCR α 和TCR β 也各自来自所述个体。

145. 如上述方法中任一项所述的方法,其中存在至少2个、3个、4个、5个、6个、7个、8个、9个、10个、50个、100个、500个、1000个、10000个、100000个或100万个TCR对(或包括这些对的细胞),并且存在至少2个、3个、4个、5个、6个、7个、8个、9个、10个、50个、100个、500个、1000个、10000个、100000个或100万个抗原。

分离TCR基因的方法

[0001] 相关申请的交叉引用

[0002] 本申请要求于2019年7月15日提交的第62/874125号美国临时专利申请、2020年2月13日提交的第62/975924号美国临时专利申请、2020年5月13日提交的第63/024341号美国临时专利申请、2020年6月3日提交的第63/034157号美国临时专利申请和2020年6月15日提交的第63/039346号美国临时专利申请的优先权,它们的全部内容通过引用并入本文。

[0003] 关于序列表

[0004] 本申请与电子格式的序列表一起提交。序列表以名为“NTBV001WOSEQLIST.txt”的文件形式提供,该文件创建于2020年7月13日,大小为21,765字节。电子序列表中的信息通过引用整体并入本文。

技术领域

[0005] 本发明技术大致涉及分离T细胞受体 (TCR) 基因序列以恢复TCR库。还提供了治疗用组合物和方法。

背景技术

[0006] 基于PCR的技术和TCR批量链测序已被用于检测TCR $\alpha\beta$ 对 (Kobayashi等人,Nat Med 2013年,linnemann等人,Nat Med 2013年,Tran等人,Science 2015年,Tran等人,N Engl J Med 2016年,Tsuji等人,Cancer Immunol Res 2018年)。

发明内容

[0007] 在一些实施方案中,本文中描述了从核酸组合文库中识别编码T细胞受体 α (TCR α)-链和TCR β -链的核苷酸序列的方法。该方法包括 (I) 提供包含编码TCR α -链和TCR β -链的多种变体核酸的文库, (II) 将该文库导入能够表达由该多种变体核酸的成员编码的TCR α -链和TCR β -链的细胞群中, (III) 基于响应于抗原的高于阈值水平的标志物的表达来选择该细胞群的亚群,其中所述亚群包括多个细胞, (IV) 从所述亚群中分离所述多种变体核酸的子集 (subset), (V) 测定该变体核酸的核苷酸序列,和 (VI) 基于所述子集内的核苷酸序列相对于对照的富集来识别至少一个变体核苷酸序列。

[0008] 其他一些实施方案涉及从不同的T细胞群恢复T细胞受体 (TCR) 库的方法,所述方法包括: I) 测定个体样本中的TCR- α 和TCR- β 的核苷酸序列或氨基酸序列; II) 从总库中选择TCR α -链和TCR β -链的序列的一个或多个子集; III) 通过所选的TCR α -链和TCR β -链的序列的组合配对创建TCR库,从而创建TCR $\alpha\beta$ 对的文库; 以及IV) 从所创建的TCR库中识别具有期望特征的至少一个TCR $\alpha\beta$ 对。

[0009] 在一些实施方案中,基于以下至少一个标准从总库中选择TCR α -链和TCR β -链的序列的一个或多个子集: a) 基于T细胞群内的频率; b) 基于与第二T细胞群相比的相对富集; c) 基于给定TCR链的DNA拷贝数和RNA拷贝数的相对差异; d) 基于TCR链的生物学特性,其中所述特性选自以下中的至少一项: (预测的) 抗原特异性、序列基序、(预测的) HLA-限制性、亲

和性、共受体依赖性或者亲代T细胞谱系(例如CD4或CD8 T细胞);e)基于基因表达的空间模式,其中空间基因表达模式衍生自组织中的起始区域或其他基因的共表达模式中的至少一种;f)基于在多个样本中共现(co-occurrence)或以相似频率出现,例如在多个肿瘤病灶中出现;g)分配到多个组以分别恢复TCR库的特定部分;h)基于在不同实施方案中定义的多个标准的组合。在一些实施方案中,基于T细胞群内的频率的选择是基于TCR序列的频率的数据,该数据用于为TCR α -链和TCR β -链创建单独的秩次序(separate rank order),或者为TCR α -链和TCR β -链创建组合的秩次序。在一些实施方案中,该方法进一步包括测定频率阈值,该频率阈值是基于TCR库恢复的期望深度来定义的并且用于基于频率来选择TCR α -链和TCR β -链的集合(collection)。在一些实施方案中,测定TCR- α 和TCR- β 的序列是通过以下至少一项来实现的:a)多重PCR;b)通过靶富集进行TCR序列恢复;c)通过5' RACE和PCR进行TCR序列恢复;d)通过空间测序进行TCR序列恢复;或e)通过RNA-seq数据进行TCR序列恢复。在一些实施方案中,将恢复得到的TCR-链序列定义为CDR3核苷酸序列连同足够的5'-和3'-核苷酸序列信息,以基于核苷酸序列比对选择至少一个TCR V-片段家族和至少一个TCR J-片段家族来组装完整的TCR链序列。在一些实施方案中,核苷酸序列比对是基于65%序列同一性、70%序列同一性、75%序列同一性、80%序列同一性、85%序列同一性、90%序列同一性、95%序列同一性、96%序列同一性、97%序列同一性、98%序列同一性、99%序列同一性、100%序列同一性,以及其间的任何数值或范围。在一些实施方案中,最佳序列比对基于根据本领域技术人员已知的读段映射算法最小化距离测度(distance measure)。在一些实施方案中,寻求最佳比对。

[0010] 在一些实施方案中,步骤III是通过以下中的至少一项来实现的:i)使用TCR链序列合成TCR α -链和TCR β -链的DNA或RNA片段的单独的文库,所述TCR α -链和TCR β -链的DNA或RNA片段随后被连接成一个DNA或RNA片段(任选地,其中恰好连接了一个TCR α -链和一个TCR β -链),ii)通过直接合成其中恰好连接了一个TCR α -链和一个TCR β -链的DNA或RNA片段来产生TCR α -链和TCR β -链的组合,或iii)通过用以DNA-或RNA载体的形式编码的TCR α -基因和TCR β -基因的单独集合(separate collection)来修饰细胞池,使得细胞将表达一个TCR α -链和一个TCR β -链,而在细胞内产生TCR α -链和TCR β -链的组合;(iv)在单链TCR构建体中使TCR α -链和TCR β -链的组合连接(linked),该构建体含有TCR α -链和TCR β -链片段以及单独的或与CD28信号传导结构域的组合的CD3 ζ 或CD3 ϵ 信号传导结构域。在一些实施方案中,直接合成其中恰好连接了一个TCR α -链和一个TCR β -链的DNA或RNA片段消除了进行计算机模拟随机配对(in silico random pairing)和以一个DNA片段的形式合成所有组合的需要。在一些实施方案中,步骤IV通过以下中的至少一项来实现:i)用生成的TCR $\alpha\beta$ 对的文库修饰的报告细胞池或T细胞池被呈递至少一种目的抗原(antigen of interest)的抗原呈递细胞刺激,并且基于至少一种活化标志物分离抗原反应性报告细胞或T细胞以用于TCR分离;ii)用适于追踪细胞增殖的荧光染料标记用生成的TCR $\alpha\beta$ 对的文库修饰的报告细胞池或T细胞池,所述报告细胞池或T细胞池由表达至少一种目的抗原的抗原呈递细胞刺激,并且基于增殖来分离抗原反应性报告细胞或T细胞以用于TCR分离;iii)将用生成的TCR $\alpha\beta$ 对的文库修饰的报告细胞池或T细胞池分成至少两个样本;分别通过表达至少一种目的抗原或不表达目的抗原的抗原呈递细胞刺激样本;刺激后,将报告细胞群或T细胞群均孵育一段时间,随后均通过TCR分离来分析报告细胞群或T细胞群;从两个样本中获得的TCR $\alpha\beta$ 对的比较

将识别样本在暴露于至少一种抗原的样本中具有较高丰度的TCR基因;iv) 用生成的TCR $\alpha\beta$ 对的文库修饰的报告细胞池或T细胞池被呈递至少一种目的抗原的抗原呈递细胞刺激,并且基于报告TCR触发的至少一种报道基因例如NFAT-GFP或NFAT-YFP来分离抗原反应性报告细胞或T细胞;v) 用生成的TCR $\alpha\beta$ 对的文库修饰的报告细胞池或T细胞池被呈递至少一种目的抗原的抗原呈递细胞刺激,并且基于在TCR信号传导时获得的抗生素抗性基于抗原特异性报告细胞或T细胞的选择,例如通过NFAT-嘌呤霉素转基因的使用,分离抗原反应性报告细胞或T细胞以用于TCR分离;vi) 将用所生成的TCR $\alpha\beta$ 对的文库修饰的报告细胞池或T细胞池暴露于携带目的抗原的一种或多种MHC复合物;将与MHC复合物结合的报告细胞或T细胞分离以用于TCR分离;vii) 用生成的TCR $\alpha\beta$ 对的文库修饰的报告细胞池或T细胞池被表达至少一种目的抗原的抗原呈递细胞刺激;随后,使用基于单细胞的液滴PCR或微流体方法识别目的TCR $\alpha\beta$ 对,以将TCR分离与至少一种活化标志物的转录物水平的检测相结合;因此,检测到报告细胞或T细胞池内的单个报告细胞或T细胞,所述单个报告细胞或T细胞中TCR $\alpha\beta$ 转录物与增加水平的活化标志物共表达(co-expressed)。在一些实施方案中,步骤i) - vi) 中的TCR分离可以通过以下方式来实现:(i) 从大量抗原反应性报告细胞或T细胞中分离DNA或RNA以产生TCR $\alpha\beta$ 特异性PCR产物,所述产物通过DNA测序来分析以确定抗原反应性报告细胞或T细胞的TCR $\alpha\beta$ 基因序列,或(ii) 基于单细胞的液滴PCR或微流体方法来分析在所分析的单个T细胞中表达的TCR $\alpha\beta$ 基因序列。在一些实施方案中,报告细胞是T细胞。在一些实施方案中,报告细胞监测TCR接合(engagement)。在一些实施方案中,报告细胞池可以是TCR α /b-null。

[0011] 在一些实施方案中,个体的样本包括非活起始物料(non-viable starting material)。在一些实施方案中,恢复所识别的TCR库的限定部分。在一些实施方案中,通过基于包括但不限于以下的标准仅选择部分或全部检测到的TCR链来进行限定的或选择性的恢复:a) 基于T细胞群内的频率,b) 基于与第二T细胞群相比的相对富集,c) 基于给定TCR链的DNA拷贝数和RNA拷贝数的相对差异,d) 基于TCR链的生物学特性,其中所述特性选自下述中的至少一项:(预测的) 抗原特异性、(预测的) HLA-限制性、亲和性、共受体依赖性、亲代T细胞谱系(例如CD4或CD8 T细胞) 或TCR序列基序,e) 基于基因表达的空间模式,其中空间基因表达模式衍生自组织中的起始区域或其他基因的共表达模式中的至少一种,f) 基于在多个样本中共现(co-occurrence) 或以相似频率出现,例如在多个肿瘤病灶中出现,g) 分配到多个组以分别恢复TCR库的特定部分,基于在不同实施方案中限定的多个标准的组合。在一些实施方案中,TCR序列的选择性恢复指含有某些V基因片段的TCR序列的恢复。

[0012] 在一些实施方案中,恢复抗原特异性的TCR序列。在一些实施方案中,恢复治疗性的TCR序列。在一些实施方案中,恢复肿瘤反应性的TCR序列。在一些实施方案中,恢复新生抗原特异性的TCR序列。在一些实施方案中,本文所述的方法进一步包括施用表达新生抗原特异性的TCR序列的T细胞作为癌症疗法的步骤。在一些实施方案中,本文所述的方法用于诊断。在一些实施方案中,诊断是从感染或自身免疫的病理部位恢复TCR库。在一些实施方案中,本文所述的方法用于BCR/抗体库的恢复。在一些实施方案中,本文所述的方法进一步包括从个体中分离包含所述TCR- α 和TCR- β 的核苷酸序列的核酸。在一些实施方案中,活化标志物是CD4或CD8 T细胞活化标志物。可以使用任何CD4或CD8 T细胞活化标志物。在本文所述的方法的一些实施方案中,活化标志物选自:CD69、CD137、IFN- γ 、IL-2、TNF- α 、GM-

CSF。在一些实施方案中,在本文所述的方法中,从T细胞群中分离DNA和RNA,T细胞群是不同细胞类型的混合物或组织样本(例如血液或肿瘤组织)的一部分。在一些实施方案中,个体的样本包括从体液中分离的细胞。在一些实施方案中,细胞是肿瘤特异性T细胞。在一些实施方案中,体液选自血液、尿液、血清、浆膜液、血浆、淋巴液、脑脊液、唾液、痰液、粘膜分泌物、阴道液、腹水、胸膜液、心包液、腹膜液和腹腔液。在一些实施方案中,本文所述的方法进一步包括使用TCR $\alpha\beta$ 链序列来治疗患有癌症、免疫病症、自身免疫性疾病或感染性疾病的个体。在一些实施方案中,本文所述方法的步骤IV通过以下中的至少一项来实现:(i)基于至少一个活化标志物的识别或选择;(ii)基于响应抗原的增殖的识别或选择;(iii)基于在抗原刺激的细胞中与未刺激的细胞相比对具有更高丰度的TCR基因的识别的识别或选择;(iv)基于由TCR触发引起的报道基因激活的识别或选择;(v)基于选择性存活的识别或选择,包括但不限于通过TCR信号传导获得的抗生素抗性;(vi)基于与一种或多种MHC复合物结合的识别或选择;(vii)使用基于单细胞的液滴PCR或微流体的识别或选择;或它们的任意组合。在一些实施方案中,步骤(vii)进一步包括确定活化相关基因的共表达。

[0013] 在一些实施方案中,本文公开了创建多个T细胞文库的方法,该方法包括:(a)根据本文所述的方法恢复T细胞受体(TCR)库;(b)从总库中选择TCR α -链和TCR β -链的序列到多个组中以分别恢复TCR库的特定部分,其中创建复杂度较小的或恢复TCR库的特定部分的多个T细胞文库。在一些实施方案中,基于频率范围选择TCR α -链和TCR β -链序列。

[0014] 在一些实施方案中,本文描述了创建多个T细胞文库的方法,该方法包括:(a)根据本文所述的方法恢复T细胞受体(TCR)库;(b)从总库中选择TCR α -链和TCR β -链的序列到多个组中以分别恢复TCR库的特定部分,其中创建复杂度较小的或恢复TCR库的特定部分的多个T细胞文库。

[0015] 在一些实施方案中,基于频率范围选择TCR α -链和TCR β -链序列。

[0016] 在一些实施方案中,本文描述了从不同的T细胞群恢复T细胞受体(TCR)库的方法,该方法包括:I)测定个体样本中的TCR α -和TCR β -核苷酸或氨基酸序列;II)从总库中选择TCR α -和TCR β -链序列的一个或多个子集(subset);III)通过所选TCR α -和TCR β -链的序列的组合配对来创建TCR库,从而创建TCR $\alpha\beta$ 对的库;以及IV)从所创建的TCR库中识别具有期望特征的至少一个TCR $\alpha\beta$ 对。

[0017] 在一些实施方案中,提供了从核酸组合文库中识别核苷酸序列的方法。该方法包括提供包含多种变体核酸的组合文库,所述多种变体核酸中的每一个均包含至少600bp的连续部分。该方法进一步包括将文库导入经构造以表达由多种变体核酸的成员编码的一种或多种多肽的细胞群中。该方法进一步包括基于依赖于至少600bp的连续部分的至少一个功能特性选择细胞群的亚群,其中该亚群包括多个细胞。该方法进一步包括从该亚群中分离多种变体核酸的子集。该方法进一步包括确定该子集的个体成员的连续部分的核苷酸序列。该方法进一步包括基于核苷酸序列识别至少600bp的连续部分。在一些实施方案中,该方法也可以是其中至少600bp的连续部分分布在600个碱基对中的方法。

[0018] 在一些实施方案中,该方法可以包括下面描述的步骤(1)-(7)中的一步或多步。步骤(1)获得样本。样本可以是患有传染病、自身免疫性疾病或癌症的患者的组织、血液或体液。样本可以是具有活力的(viable)或非活的(non-viable)。步骤(2)对样本中的TCR α -链和TCR β -链进行测序。步骤(3)选择并对TCR α -链和TCR β -链的序列进行组合配对以创建TCR α

β 对的文库。步骤(4)将TCR $\alpha\beta$ 对的文库导入报告细胞池,例如Jurkat报告T细胞。步骤(5)用呈递至少一种目的抗原的抗原呈递细胞刺激用TCR $\alpha\beta$ 对的文库修饰的报告细胞。所述至少一种目的抗原可以是自体的或同种异体的。步骤(6)确定对所述至少一种目的抗原特异的TCR $\alpha\beta$ 对。步骤(7)将TCR $\alpha\beta$ 对导入细胞并选择含有TCR $\alpha\beta$ 对的细胞。在一些实施方案中,该方法可以包括上述步骤(1)-(7)中的一步或多步。任何步骤均可以酌情省略、重复或由本文提供的其他实施方案替代。也可以添加额外的中间步骤。

[0019] 一些实施方案涉及包含根据上述实施方案中任一项恢复的T细胞受体库的核苷酸文库。在一些实施方案中,包含根据上述实施方案中任一项识别的核苷酸序列的核苷酸构建体。在一些实施方案中,细胞包含本文所述的核苷酸构建体。

[0020] 在一些实施方案中,提供了一种从不同的T细胞群恢复T细胞受体(TCR)库的方法。该方法包括测定个体样本中的TCR- α 和TCR- β 的核苷酸序列或氨基酸序列;从总库中选择TCR α -链和TCR β -链的序列的一个或多个子集;通过所选的TCR α -链和TCR β -链的序列的组合配对创建TCR库,从而创建TCR $\alpha\beta$ 对的文库;以及从所创建的TCR库中识别具有期望特征的至少一个TCR $\alpha\beta$ 对。在一些实施方案中,提供了一种创建多个T细胞文库的方法。该方法包括根据上述方法恢复T细胞受体(TCR)库,从总库中选择TCR α -链和TCR β -链的序列到多个组中以分别恢复TCR库的特定部分,其中创建复杂度较小的或恢复TCR库的特定部分的多个T细胞库。

[0021] 在一些实施方案中,提供了从核酸组合文库中识别核苷酸序列的方法。该方法包括提供包含多种变体核酸的组合文库,所述多种变体核酸中的每一个均包含至少600bp的连续部分,其中所述连续部分包含两个以上变体核苷酸子序列的组合,其中所述两个以上变体核苷酸子序列的第一变体核苷酸子序列定义所述连续部分的第一端,并且所述两个以上变体核苷酸子序列的第二变体核苷酸子序列定义所述连续部分的与所述第一端相对的第二端;将文库导入经构造以表达由多种变体核酸的成员编码的一种或多种多肽的细胞群中;基于依赖于两个以上变体核苷酸子序列的组合的至少一种功能特性选择所述细胞群的亚群,其中所述亚群包括多个细胞;从所述亚群中分离所述多种变体核酸的子集;确定所述子集的个体成员的连续部分的核苷酸序列;以及基于所述核苷酸序列识别所述两个以上变体核苷酸子序列的至少一种组合。

[0022] 在一些实施方案中,提供了从核酸组合文库中识别编码T细胞受体 α (TCR α)-链和TCR β -链的核苷酸序列的方法。该方法包括:提供包含多种变体核酸的文库,所述多种变体核酸中的每一种均包含至少600bp的连续部分,其中所述连续部分包含:编码TCR α 变体氨基酸序列并定义所述连续部分的第一端的第一变体核苷酸子序列和编码TCR β 变体氨基酸序列并定义与所述第一端相对的所述连续部分的第二端的第二变体核苷酸子序列的组合。该方法可以进一步包括将文库导入经构造以表达由多种变体核酸的成员编码的TCR α -链和TCR β -链的永生化T细胞群中。该方法可以进一步包括基于响应于使永生化T细胞与表达抗原的永生化B细胞接触的高于阈值水平的T细胞活化标志物的表达来选择永生化T细胞群的亚群,其中所述亚群包括多个T细胞。该方法可以进一步包括从所述亚群中分离所述多种变体核酸的子集。该方法可以进一步包括确定所述子集的个体成员的连续部分的核苷酸序列;和基于所述子集的核苷酸序列中的至少一种组合相对于对照的富集来识别所述第一变体核苷酸子序列和第二变体核苷酸子序列的至少一种组合。

[0023] 在一些实施方案中,提供了从核酸组合文库中识别编码嵌合抗原受体(CAR) 铰链结构域、跨膜结构域和/或细胞内信号传导结构域的核苷酸序列的方法。该方法可以包括:提供包含多种变体核酸的文库,所述多种变体核酸中的每一个均包含至少600bp的连续部分,其中所述连续部分包含以下两种以上的组合:编码CAR铰链结构域的第一变体核苷酸子序列;编码CAR跨膜区的第二变体核苷酸子序列;和编码CAR细胞内信号传导结构域的第三变体核苷酸子序列。第一、第二或第三变体核苷酸子序列中的一个限定了连续部分的第一端,并且其中第一、第二或第三变体核苷酸子序列中的另一个限定了与第一端相对的连续部分的第二端。所述方法可以进一步包括将所述文库导入经构造以表达由所述多种变体核酸的成员编码的CAR的细胞群中。细胞群包括永生化T细胞群或原代人T细胞群。该方法可以进一步包括基于响应于使细胞与表达对CAR的抗原结合结构域特异的抗原的抗原呈递细胞接触的高于阈值水平的细胞增殖来选择细胞群的亚群,其中所述亚群包括多个细胞。该方法可以进一步包括从所述亚群中分离所述多种变体核酸的子集。所述方法可以进一步包括确定所述子集的个体成员的连续部分的核苷酸序列;以及基于所述子集的核苷酸序列中的至少一种组合相对于对照的富集来识别所述第一、第二和第三变体核苷酸子序列的至少一种组合。

[0024] 在一些实施方案中,提供了从核酸组合文库中识别核苷酸序列的方法。该方法可以包括:提供包含多种变体核酸的组合文库,所述多种变体核酸中的每一个均包含至少600bp的连续部分;将所述文库导入经构造以表达由所述多种变体核酸的成员编码的一种或多种多肽的细胞群中;基于依赖于至少600bp的连续部分的至少一个功能特性选择所述细胞群的亚群,其中所述亚群包括多个细胞;从所述亚群中分离所述多种变体核酸的子集;确定所述子集的个体成员的连续部分的核苷酸序列;以及基于核苷酸序列识别至少600bp的连续部分。

[0025] 在一些实施方案中,提供了从核酸文库中识别编码抗原特异性T细胞受体 α (TCR α)-链和TCR β -链对的核苷酸序列的方法。该方法包括将文库导入能够表达由多种变体核酸的成员编码的TCR α -链和TCR β -链的细胞群中,基于响应于抗原的高于阈值水平的标志物的表达来选择细胞群的亚群,其中所述亚群包括多个细胞。该方法可以进一步包括从所述亚群中分离所述多种变体核酸的子集。该方法可以进一步包括确定所述变体核酸的核苷酸序列,和基于子集内的核苷酸序列相对于对照的富集来识别至少一种变体核苷酸序列。

[0026] 在一些实施方案中,提供了从样本中识别编码T细胞受体 α (TCR α)-链和TCR β -链的核苷酸序列的方法。该方法可以包括对样本中的TCR- α 链和TCR- β 链进行测序,选择并对TCR α 和TCR β 链序列进行组合配对以创建TCR $\alpha\beta$ 对的文库,将TCR $\alpha\beta$ 对的文库导入报告细胞池中,用呈递至少一种目的抗原(其中所述抗原可以来自TCR- α 链和TCR- β 链所来自的同一宿主)的抗原呈递细胞刺激用TCR $\alpha\beta$ 对的文库修饰的报告细胞,确定对所述至少一种目的抗原特异的TCR $\alpha\beta$ 对,将所述TCR $\alpha\beta$ 对导入细胞并选择含有所述TCR $\alpha\beta$ 对的细胞。

[0027] 在一些实施方案中,提供了包含根据上述方法中的任意一种恢复的T细胞受体库的核苷酸文库。

[0028] 在一些实施方案中,提供了包含根据本文任意一种方法识别的核苷酸序列的核苷酸构建体。

[0029] 在一些实施方案中,提供了包含根据本文所提供的任意核苷酸的核苷酸构建体的

细胞。

[0030] 在一些实施方案中,提供了从核酸文库中识别编码抗原特异性T细胞受体 α (TCR α)-链和TCR β -链对的核苷酸序列的方法。该方法可以包括:将核酸文库导入能够表达TCR α -链和TCR β -链的细胞群中以制备细胞文库;基于响应于抗原的高于第一阈值水平的标志物的表达来选择所述细胞文库的第一群;以及从所述文库的第一群中分离变体核酸的第一群。在一些实施方案中,抗原可以是一种或多种,并且抗原以及TCR α 序列和TCR β 序列都可以在一个个体中发现。

[0031] 在一些实施方案中,提供了从核酸文库中识别编码T细胞受体 α (TCR α)-链和TCR β -链的核苷酸序列的方法。该方法可以包括:将核酸文库导入能够表达TCR α -链和TCR β -链的细胞群中以形成细胞文库;以及基于子集内的核苷酸序列相对于对照的富集来确定变体核酸的第一群的至少一个核苷酸序列或核酸同一性。

[0032] 在一些实施方案中,提供了从核酸文库中识别核苷酸序列的方法。该方法可以包括将核酸文库导入细胞群以形成细胞文库;使所述细胞文库与第一细胞群接触;基于至少一种标志物的表达通过磁珠富集来选择所述细胞文库的亚群;以及基于所述亚群内的核苷酸序列相对于对照的统计学上显著的富集或耗竭来识别至少一种核苷酸序列。

[0033] 在一些实施方案中,提供了细胞的集合。所述集合包括至少两种T细胞的组,其中每种T细胞均被构造为表达至少一个TCR α 和TCR β 对,其中所述TCR α 和TCR β 均来自一个个体,其中所述T细胞不表达内源性TCR,并且其中所述组被构造用于一个或多个T细胞活化标志物的激活;以及至少两种B细胞的组,其中所述至少两种B细胞中的每一种均被构造为表达至少一种外源性新生抗原(或抗原),使得存在至少两种能够被产生的外源性新生抗原(或抗原),并且其中所述至少两种外源性新生抗原(或抗原)与所述个体中的外源性新生抗原(或抗原)相同。

[0034] 在一些实施方案中,提供了表达TCR的细胞的文库。所述表达TCR的细胞的文库包括:至少三种T细胞的组,其中所述T细胞中的至少两种被构造为表达至少两个TCR α 和TCR β 对(至少两个TCR对),其中所述至少两个TCR对来自一个个体,其中所述至少三种T细胞不表达内源性TCR,其中所述至少三种T细胞被构造用于在结合由B细胞呈递的抗原(或新生抗原)时一个或多个T细胞活化标志物的激活,其中反映在若干个TCR细胞中的每个TCR对的基因组拷贝的数量使得可以读取样本中的每个TCR,且其中至少一种所述TCR在包含所述文库的整个组合物中分布不均。

[0035] 在一些实施方案中,提供了一种治疗个体的方法。该方法包括识别患有肿瘤的个体;提供至少两种T细胞的组,每种T细胞均被构造为表达至少一个不同的TCR α 和TCR β 对,其中所述TCR α 和所述TCR β 各自来自所述个体,提供至少两种B细胞的组,其中所述B细胞组经构造以表达至少两种外源性新生抗原,并且其中所述至少两种外源性新生抗原与在所述个体中发现的那些新生抗原相同;将所述至少两种T细胞的组与所述至少两种B细胞的组组合,并基于所述至少两种T细胞经由所述至少两种外源性新生抗原的激活来选择至少两个TCR对的组合;以及向所述个体施用所述至少两个TCR对的组合,从而治疗所述肿瘤。

[0036] 在一些实施方案中,提供了一种治疗个体的方法。该方法包括:识别患有肿瘤的个体;提供至少两种T细胞的组,每种T细胞被构造为表达至少一种不同的TCR α 和TCR β 对,其中所述TCR α 和TCR β 各自来自所述个体;提供至少两种抗原呈递细胞的组,其中所述抗原呈递

细胞的组源自所述个体,并被构造为表达至少两种外源性新生抗原,并且其中所述至少两种外源性新生抗原与在所述个体中发现的那些新生抗原相同;将所述至少两种T细胞的组与所述至少两种抗原呈递细胞的组组合,并基于所述至少两种T细胞经由所述至少两个外源性新生抗原的激活来选择至少两个TCR对的组合;以及向所述个体施用所述至少两个TCR对的组合,从而治疗所述肿瘤。

[0037] 在一些实施方案中,提供了一种药物组合物。该药物组合物可以包含:第一TCR对,其与个体的肿瘤中的第一抗原(或新生抗原)结合;和第二TCR对,其与所述个体的肿瘤中的第二抗原(或新生抗原)结合。

[0038] 在一些实施方案中,提供了一种药物组合物。该药物组合物可以包含:第一TCR对,其与第一抗原结合并且是MHC I类限制性的;和第二TCR对,其与第二抗原结合并且是MHC II类限制性的。

[0039] 在一些实施方案中,提供了一种细胞集合。该集合可以包括:至少两种T细胞的组,其中每种T细胞均被构造为表达至少一个TCR α 和TCR β 对,其中所述对来自一个个体,其中所述T细胞不表达内源性TCR,并且其中所述组被构造用于激活一个或多个T细胞活化标志物;和至少两种抗原呈递细胞(APC)的组,其中所述至少两种APC中的每一种均被构造为表达至少一种外源性新生抗原(或抗原),使得存在至少两种能够被产生的外源性新生抗原(或抗原),并且其中所述至少两种外源性新生抗原(或抗原)与所述个体中的外源性新生抗原(或抗原)相同。

附图说明

[0040] 图1示出了恢复TCR库以用于文库生成的方法。

[0041] 图2示出了恢复TCR库和识别用于癌症治疗的抗原特异性TCR α 序列和TCR β 序列的方法。

[0042] 图3示出了TCR表达盒的示例性设计。

[0043] 图4示出了用于筛选组合TCR文库的一些实施方案。

[0044] 图5示出了本文公开的用于治疗癌症患者的方法的示例性应用。

[0045] 图6示出了TCR β -P2A-TCR α -T2A-嘌呤霉素抗性盒的示意性实例。

[0046] 图7示出了组装TCR表达盒的示例性策略。

[0047] 图8示出了根据本发明的实施方案的筛选方法的示意性实例。

[0048] 图9示出了根据本发明的实施方案的筛选方法的流程图。

[0049] 图10A-图10J描述了从非活肿瘤标本中恢复TCR库以识别新生抗原特异性TCR序列。

[0050] 图10A为筛选过程的一些实施方案的示意图。

[0051] 图10B为描述TCR克隆型(clonotype)数量的图。

[0052] 图10C为描绘10,000个TCR $\alpha \times \beta$ 组合中的每一个的概率密度的图。

[0053] 图10D为描绘杀稻瘟菌素(Blasticidin)选择后所得细胞群的图。

[0054] 图10E为使用各种指示标志物的一系列FACS结果。

[0055] 图10F为描绘各种PCR产物的凝胶。

[0056] 图10G为描绘了在存在(x轴)和不存在(y轴)通过B细胞表达的TMG的情况下筛选的

经Rlog转换的平均读段计数的一系列图,其代表在10B)-10F)中描绘的pt1肿瘤样本以及对于以相同方式处理的另外三个MMRp-CRC样本(pt2、pt3和pt4)。新生抗原反应性TCR先导(lead)描绘为被圆圈包围的较大黑点。

[0057] 图10H为描绘步骤10H的图。

[0058] 图10I为描绘测量为相对于阳性对照(用PMA/离子霉素处理)的CD69阳性的活化的图。示出了由对照(PARVA-P70R)以及AKAP8L-R191W肽激活的情况。

[0059] 图10J为描绘测量为相对于阳性对照(用PMA/离子霉素处理)的CD69阳性的活化的图。示出了由对照(ETS1-R70W)以及TP53-R282W肽激活的情况。

[0060] 图10K为描绘测量为相对于阳性对照(用PMA/离子霉素处理)的CD69阳性的活化的图。示出了由对照B细胞(不表达TMG)以及由表达MG91(HspA9-p.k654RfsX42)或TMG3的B细胞激活的情况。

[0061] 图10L是描绘测量为相对于阳性对照(用PMA/离子霉素处理)的CD69阳性的活化的图。示出了由对照B细胞(不表达TMG)以及由表达MG132(ITPR3-p.L2379M)或TMG4的B细胞激活的情况。

[0062] 图11A为描述TCR克隆型数量的条形图。

[0063] 图11B为TCR表达质粒的示意图。

[0064] 图11C是描绘每个患者文库的10,000个 $\alpha \times \beta$ 组合中的每一个的概率密度的曲线图。

[0065] 图11D提供了每个TCR的读段的数量范围、平均覆盖率以及落在中位数+/- $2\log$ -单位($\text{median} \pm 2\log\text{-unit}$)范围内的TCR百分比。

[0066] 图12A为示出克隆型数量的条形图。

[0067] 图12B是描绘两个患者库的10,000个 $\alpha \times \beta$ 组合中的每一个的概率密度的图。

[0068] 图12C是提供10,000个 $\alpha \times \beta$ 组合中的每一个的覆盖率的表。

[0069] 图13A为提供了在24个抗原特异性未知的TCR中的6个具有已知抗原特异性的TCR的稀释度的表。

[0070] 图13B为示出了选择后的TCR+Jurkat报告细胞的频率的FACS结果。

[0071] 图13C为描述基于CD69活化标志物的顶部和底部T细胞群分选策略的FACS图。

[0072] 图13D为描绘各种PCR产物的凝胶。

[0073] 图13E为比较具有已知抗原特异性的TCR和具有未知抗原特异性的TCR的富集频率的图。

[0074] 图14A为50 $\times$ 50和100 $\times$ 100的设计流程图,其中掺入了(spiked in)特征性TCR链。

[0075] 图14B为描绘基于CD69活化标志物的顶部和底部T细胞群的分选策略的FACS图的集合。

[0076] 图14C为描绘各种PCR产物的凝胶。

[0077] 图14D为比较顶部样本与底部样本中的TCR代表(representation)相对于这些样本中的平均表达的倍数变化的图。

[0078] 图14E为提供了最富集的TCR序列的秩次序,以及此类TCR的碱基平均值、Log(2)倍数变化和调整后的p值的表。

[0079] 图14F为比较不存在抗原时的TCR反应性与存在抗原时的TCR反应性的图。

- [0080] 图14G为显示抗原特异性TCR和具有未知特异性的TCR在筛选中富集的概率的图。
- [0081] 图15A为描绘TCR盒的组合组装的示意图。
- [0082] 图15B为描绘TCR文库中的TCR α -链、TCR β -链和TCR $\alpha\beta$ 组合的概率密度的一系列图。
- [0083] 图15C和图15D为描绘组成更高复杂度TCR文库的替代策略的图。
- [0084] 图15E描绘了存在于通过合成四个100 $\times$ 100文库并以1:1:1:1等摩尔比混合这些文库而创建的两个200 $\times$ 200TCR文库中的TCR组合的概率密度。
- [0085] 图15F描绘了使用200 $\times$ 200文库筛选方法对pt2的TCR识别。
- [0086] 图15G示出了在100 $\times$ 100和200 $\times$ 00文库筛选中100 $\times$ 100筛选中识别的pt2 TCR的统计行为的表。
- [0087] 图16A为示出CD69+细胞的百分比与用于脉冲APC的CMV肽的量的关系的图。
- [0088] 图16B为描绘CD69+细胞的百分比与细胞接种密度的条形图。
- [0089] 图16C为描绘CD69+细胞的百分比与效应物与靶标的比值和培养皿 (culture vessel) 的关系的条形图。
- [0090] 图16D为描绘了相对于铺板的效应细胞数量和所用抗原性肽的量的CD69+细胞的百分比一系列条形图。
- [0091] 图16E为示出各种TCR的富集程度的条形图。
- [0092] 图17a-图17c用CDK4和CMV转导的Jurkat TCR KO细胞进行的肽滴定分析。(17a)用CDK4-8或CDK4-17 TCR或者用CMV-1或CMV-2TCR逆转录病毒转导Jurkat TCR KO细胞。转导的TCR由小鼠恒定区和人可变区组成。然后,对mTCR β^+ CD8 $^+$ 群进行分选(sorted)。上部图(top panel)显示分选前的细胞(cell),下部图(bottom panel)显示分选后的细胞。(17b:CDK4, 17c:CMV)与加载了分级浓度的同源肽(cognate peptide)的JY细胞(E:T比例为1:1)共培养20小时后,通过CD69上调来评估的TCR转导后或非转导的Jurkat细胞的激活。图表图例中描述了EC50值。使用PMA/离子霉素刺激作为阳性对照,并且使用不含JY细胞的培养物和加载有最高浓度(1 μ g/ml)的不相关肽的JY细胞作为阴性对照(在每个图的右侧描绘)。分析以一式三份进行(含CDK4突变肽的非转导的Jurkat细胞一式二份)。数据显示为平均值 $\pm$ 标准差(s.d.) (误差条仅显示在肽滴定曲线中)。n=1。NT,未转导的。
- [0093] 图18a-图18e。具有不同转导效率的CMV-1转导的Jurkat TCR KO细胞的杀稻瘟菌素选择(blasticidin selection)。(18a)用不同体积的CMV-1-杀稻瘟菌素逆转录病毒上清液转导Jurkat TCR KO细胞,以实现不同的转导效率。对未转导的和TCR转导后的Jurkat细胞进行CD8和mTCR β 染色。(18b,18c,18d,18e)将具有不同CMV-1转导效率的Jurkat TCR KO细胞以0.25 $\times$ 10 ^{6} 个细胞/毫升的浓度铺板,并用不同浓度的杀稻瘟菌素进行选择。四天后,通过去除杀稻瘟菌素(称为“在第4天去除杀稻瘟菌素”)或通过添加相应浓度的新的杀稻瘟菌素(称为“在第4天添加新的杀稻瘟菌素”),以0.25 $\times$ 10 ^{6} 个细胞/毫升的密度对细胞进行重新铺板。(18b,18d)分别在开始杀稻瘟菌素选择后6或7天,总活细胞和mTCR β^+ CD8 $^+$ 细胞的倍数扩增。(18c,18e)分别在开始杀稻瘟菌素选择后6天或7天,mTCR β^+ CD8 $^+$ 细胞的百分比。n=1。NT,未转导的。
- [0094] 图19a-图19d。扩大规模(Upscaling) Jurkat细胞-APC共培养物。(a、b、c、d-左)每个实验中效应Jurkat细胞的CD8和mTCR β 表达。或者用(19a)的由其中四个对TMG2.1中包含的抗原具有特异性的16个TCR组成的TCR文库,或者用(19b、19c、19d)的CDK 4-17TCR进行

Jurkat细胞转导。(19b、19c) 双阴性细胞群是由于其低CD20表达而不可能被排除的靶细胞。(19a、19b、19c、19d-右) 与表达TMG2.1的EBV LCL共培养20小时(E:T比例为1:1, 2.5×10^6 细胞/毫升细胞密度)。将未转导的EBV LCL用作阴性对照。绘制了总活细胞和CD69⁺效应细胞的百分比。绿色文本表示在1100rpm的转速下旋转1分钟(短时间旋转)的条件。使用的细胞总数: 0.5×10^6 个细胞(96W)、 $\leq 2 \times 10^6$ 个细胞(15/50ml锥形圆底管(Falcon)中分别为 $1.3/2 \times 10^6$ 个细胞)、 170×10^6 个细胞(GMP袋)。以一式三份(19c)或一式两份(19d)进行96W共培养。数据显示为平均值±标准差。 $n=1$ 。96W, 96孔圆形底板。GMP袋, MACS GMP细胞分化袋-500。*来自少量活细胞的数据。

[0095] 图20a-图20e。T细胞活化标志物CD69、CD25和CD62L的纵向分析。在存在(CDK4-8:圆形, CMV-1:三角形)或不存在(CDK4-8:矩形, CMV-1:倒三角形)表达TMG2.1的EBV LCL的情况下以1:1的E:T比例将用CDK4-8或CMV-1TCR转导后的Jurkat TCR KO细胞(>80% mTCR β +CD8⁺细胞)培养16、20、24、28和32小时, 然后对效应细胞的CD69、CD25和CD62L表达进行多色流式细胞术分析(multi-color flow cytometric analysis)。使用与EBV LCL TMG2.1共培养的未转导Jurkat TCR KO细胞作为阴性对照(菱形)。每个代表性点图来自CDK4-8转导的Jurkat细胞与效应细胞的EBV LCL TMG2.1的(20a) CD69上调、(20b) CD25上调以及(20d) CD62L下调的20小时共培养。CD69与(20c) CD25或(20e) CD62L的共表达分析。以一式三份(显示为平均值±标准差)进行共培养, $n=1$ 。NT, 未转导的。

[0096] 图21a-图21d。评估含有嘌呤霉素抗性基因的基于NFAT报告慢病毒载体在Jurkat TCR KO细胞中的功效。(21a) 带有嘌呤霉素抗性盒的四种基于NFAT的报告慢病毒质粒的示意图。(21b) 使用PEI作为转染试剂, 将4种载体转染到HEK293T细胞中。共转染另外一个质粒(pmaxGFP)作为转染效率的量度。点图示出了转染后三天未转染的和转染后的HEK293T细胞的GFP表达。(21c) 慢病毒上清液与聚凝胺联合使用以转导Jurkat TCR KO细胞。用PMA/离子霉素刺激Jurkat细胞24小时, 然后用1 μ g/ml的嘌呤霉素筛选3天。点图显示了存活的(DAPI⁻)和CD69⁺未转导的和转导后的Jurkat TCR KO细胞的百分比。(21d) 将转导和嘌呤霉素选择后的细胞扩增20天, 以进行24小时的第二轮PMA/离子霉素刺激, 随后进行4天的不同嘌呤霉素浓度的嘌呤霉素选择。绘制了存活的(DAPI⁻)和CD69⁺细胞的百分比。请注意, NT细胞与(c)中的NT细胞不同。 $n=1$ 。NT, 未转导的。

[0097] 图22a-图22c。评估基于EGFP NFAT的报告慢病毒载体在Jurkat TCR KO细胞中的功效。(22a) 具有EGFP基因的四四种基于NFAT的报告慢病毒质粒的示意图。(22b) 以PEI或FuGENE为转染试剂, 将4种NFAT载体转染到HEK293T细胞中。显示了转染后三天GFP⁺HEK293T细胞的百分比。(22c) 借助作为转导试剂的聚凝胺, 将来自PEI转染(transfection)的慢病毒上清液转导至Jurkat TCR KO细胞中。用PMA/离子霉素刺激NFAT转导后的Jurkat TCR KO细胞24小时, 并且每24小时测量一次GFP表达(显示为GFP⁺细胞的百分比和MFI), 持续3天。通过CD69上调来评估Jurkat细胞的活化。 $n=1$ 。MFI, 平均荧光强度。EGFP, 增强型绿色荧光蛋白。NT, 未转导的。

[0098] 图23。评估基于EGFP NFAT的报告慢病毒载体在原代T细胞中的功效。用图23中使用的NFAT4 $\times$ 慢病毒载体或含有不同最小启动子(minP)的NFAT4 $\times$ 慢病毒载体(NFAT4 $\times$ 新)转导原代T细胞。用PMA/离子霉素刺激NFAT转导后的原代T细胞24小时, 并且每24小时测量一次GFP表达(显示为GFP⁺细胞的百分比和MFI), 持续3天。通过CD69表达来评估Jurkat细胞

的活化。 $n=2$ 个生物学上独立的重复(显示为平均值 $\pm$ s.d)。* $P\leq 0.05$, ** $P\leq 0.01$, ns:不显著(双因素方差分析后进行Sidak多重比较检验)。MFI, 平均荧光强度。NT, 未转导的。

[0099] 图24a-图24d。基于NFAT的报告质粒在Jurkat TCR KO细胞中的非病毒递送。(24a)两个NFAT质粒的示意图。NFAT0 $\times$ 不含有任何NFAT结合位点, 并且用于评估单独的minP的背景信号。(b、c、d)将NFAT0 $\times$ 和NFAT4 $\times$ 电穿孔导入CDK4-17转导后的Jurkat TCR KO细胞($\sim 90\%$ mTCR β^+ CD8 $^+$ 细胞)中。(24b)电穿孔后20小时、2天和6天评估了在存在或不存在质粒DNA的情况下未经电穿孔或经电穿孔的Jurkat细胞的细胞存活力。(24c)NFAT转染后的Jurkat细胞的E2-Crimson和GFP单阳性和双阳性群体百分比的纵向分析。在第6天开始使用0.5 μ g/毫升的嘌呤霉素进行持续7天的抗生素选择。在开始选择后四天更新嘌呤霉素。(24d)嘌呤霉素选择后七天细胞活力(DAPI细胞)以及GFP和E2-Crimson表达的点图。 $n=1$ 。minP, 最小启动子。EGFP, 增强型绿色荧光蛋白。

[0100] 图25a、图25b。未转导的Jurkat TCR KO细胞的杀稻瘟菌素选择。将未转导的Jurkat TCR KO细胞以 0.25×10^6 个细胞/毫升的浓度铺板, 并用不同浓度的杀稻瘟菌素选择6天。四天后, 通过去除杀稻瘟菌素(称为‘在第4天去除杀稻瘟菌素’)或添加相应浓度的新的杀稻瘟菌素(称为‘在第4天添加新的杀稻瘟菌素’), 以 0.25×10^6 个细胞/毫升的浓度重新将细胞铺板。(25a)杀稻瘟菌素选择后总活细胞的倍数扩增。(25b)示出在存在或不存在4 μ g/毫升的杀稻瘟菌素的情况下培养的未转导的细胞的FSC-A/SSC-A和CD8/mTCR β 表达的流式细胞术图。 $n=1$ 。NT, 未转导的。

[0101] 图26a、26b。4 μ g/毫升的杀稻瘟菌素选择后的具有不同转导效率的CMV-1转导的Jurkat TCR KO细胞的mTCR β MFI。(26a, 26b)将具有不同CMV-1TCR逆转录病毒转导效率的Jurkat TCR KO细胞以 0.25×10^6 个细胞/毫升的浓度铺板, 并用4 μ g/毫升的杀稻瘟菌素筛选。四天后, 通过去除杀稻瘟菌素(称为: 在第4天去除杀稻瘟菌素)或添加新的4 μ g/毫升的杀稻瘟菌素(称为: 在第4天添加新的杀稻瘟菌素), 以 0.25×10^6 个细胞/毫升的浓度重新将细胞铺板。(26a)用杀稻瘟菌素选择6/7天的mTCR β^+ CD8 $^+$ 细胞的mTCR β MFI。(26b)示出在开始杀稻瘟菌素选择7天后的CD8和mTCR β 染色的代表性点图。 $n=1$ 。MFI, 平均荧光强度。

[0102] 图27。使用两种不同的抗人CD69单克隆抗体, 克隆FN50和克隆CH/4进行T细胞活化的测量。实验设置描述于图19b的图例中。同时对细胞进行CD69克隆FN50(APC)和CD69克隆CH/4(PE)染色。 $n=1$ 。96W, 96孔圆形底板。*来自少量活细胞的数据。

[0103] 图28。描述了新生抗原特异性的TCR分离平台的一些实施方案。

[0104] 图29。在仍旧保持TCR覆盖率的同时, 通过更高效地处理大量细胞来增强TCR分离平台的可扩展性。

[0105] 图30。测定杀稻瘟菌素的功效和毒性的方法。

[0106] 图31。描绘了TCR平台分离平台的各种实施方案, 其可以包括所有步骤、或每个加框的步骤(例如, 1和/或2、和/或3、和/或4)、或多个连续编号的步骤(例如, 1-7)中的任意一个。

[0107] 图32。描述了SEQ ID NO:1:表示为TCR β -P2A-TCR α -T2A-嘌呤霉素抗性表达盒的TCR基因的氨基酸序列; 以及SEQ ID NO:4:表示为TCR β -P2A-TCR α -T2A-杀稻瘟菌素的抗性表达盒的TCR基因的氨基酸序列。

[0108] 图33。SEQ ID NO:2:CD8 α -P2A-CD8 β 转基因的氨基酸序列。

- [0109] 图34.SEQ ID NO:3:HLA-A*02:01-IRES-FusionRed的示例核苷酸序列。
- [0110] 图35A示出了筛选设计示意图。使用来自卵巢癌 (OVC) 或结直肠癌 (CRC) 样本的5个特征性TCR和95个非特征性TCR创建100×100设计的组合TCR文库。
- [0111] 图35B示出了使用CD69标记物通过FACS激活T细胞的细胞分选结果。
- [0112] 图35C示出了所得的PCR产物,其中PCR具有有限数量的循环以从分选后的TCR转导的Jurkat T细胞中扩增TCRβ-P2A-TCRα盒的部分,大小为约1.5kb。
- [0113] 图35D示出了来自图35C的筛选数据的TCR富集分析。
- [0114] 图35E示出了五种特征性的抗原反应性TCR的特征。
- [0115] 图36A示出了筛选设计的示意图。使用来自卵巢癌 (OVC) 或结直肠癌 (CRC) 样本的5个特征性TCR和95个非特征性TCR创建100×100设计的组合TCR文库。
- [0116] 图36B示出了筛选的分选策略 (sorting strategy)。
- [0117] 图36C示出了TCR表达盒的提取 (retrieval)。
- [0118] 图36D示出了来自图36C的筛选数据的TCR富集分析。
- [0119] 图36E示出了图36D中前7种最显著富集的TCR的特征。
- [0120] 图37A示出了6×6组合TMG编码设计的示意图。
- [0121] 图37B示出了所有TCRα×β组合的秩次序作为pt2 TCR文库筛选的重复次数的函数的分析。
- [0122] 图37C示出了基于2或3次重复的CRC TCR文库筛选的统计分析汇总表。
- [0123] 图37D示出了用于成对TCR富集分析的pt4样本表。
- [0124] 图37E示出了成对TCR富集分析结果。
- [0125] 图38示出了筛选与验证数据之间的TCR激活与TCR背景激活的相关性。
- [0126] 图39示出了TCR代表 (representation) 的度量。该图显示,被它们的同源抗原刺激的五种特征性TCR富集在顶部样本中 (最浅的灰色阴影)。来自与表达TMG的B细胞共培养物的底部样本显示较低的rlog值。
- [0127] 图40描绘了表示为TCRβ-P2ATCRα-T2A-嘌呤霉素抗性表达盒的TCR基因的SEQ ID NO:1的图。
- [0128] 图41描绘了SEQ ID NO:2的图,一种CD8α-P2A-CD8β转基因。
- [0129] 图42描绘了SEQ ID NO:3:K562-HLA-A*02:01-IRESFusionRed的图。
- [0130] 图43描绘了SEQ ID NO:4的图,一种表示为TCRβ-P2A-TCRα-T2A杀稻瘟菌素抗性表达盒的TCR基因。

具体实施方式

[0131] 本发明提供了用于从不同的T细胞群恢复T细胞受体 (TCR) 库的方法和组合物。该方法的一些实施方案提供了从非活材料 (包括人组织标本) 中识别和分离抗原特异性TCR。一些实施方案以图31的一些或全部为依据。在一些实施方案中,该方法可以包括图31和本文中概述的步骤 (1) - (7) 中的一步或多步。例如,该方法可包括 (1) 获得样本。样本可以是患有传染病、自身免疫性疾病或癌症的患者的组织、血液或体液。样本可以是具有活力的或非活的。步骤 (2) 对样本中的TCRα-链和TCRβ-链进行测序。步骤 (3) 选择并对TCRα-链和TCRβ-链的序列进行组合配对以创建TCRαβ对的文库。步骤 (4) 将TCRαβ对的文库导入报告细胞池,

例如Jurkat报告T细胞。步骤(5)用呈递至少一种目的抗原的抗原呈递细胞刺激用TCR $\alpha\beta$ 对的文库修饰的报告细胞。所述至少一种目的抗原可以是自体的或同种异体的。步骤(6)确定对所述至少一种目的抗原特异的TCR $\alpha\beta$ 对。步骤(7)将TCR $\alpha\beta$ 对导入细胞并选择含有TCR $\alpha\beta$ 对的细胞。在一些实施方案中,该方法可以包括上述步骤(1)-(7)中的一步或多步。任何步骤均可以酌情省略、重复或由本文提供的其他实施方案替代。也可以添加额外的中间步骤。

[0132] 在一些实施方案中,对于本文的任何方法,TCR对和/或表达TCR对的T细胞通过与抗原(如新生抗原)结合来选择或识别,其中所述抗原由B细胞或抗原呈递细胞表达。在一些实施方案中,对于本文的任何一种方法,抗原或新生抗原来自一个个体的肿瘤,并且TCR对的TCR α 和TCR β 也各自来自该个体。在一些实施方案中,存在至少2个、3个、4个、5个、6个、7个、8个、9个、10个、50个、100个、500个、1000个、10000个、100000个或100万个TCR对(或包含这些对的细胞),并且存在至少2个、3个、4个、5个、6个、7个、8个、9个、10个、50个、100个、500个、1000个、10000个、100000个或100万个抗原。

[0133] 在一些实施方案中,采用的和/或由上述方法产生的任何组合物也被预期作为文库和/或试剂盒和/或组合物和/或用于它们在医学应用和/或筛选系统中的应用。在一些实施方案中,组合物可以通过本文提供的任何方法选择或产生的TCR α -链和TCR β -链配对。在一些实施方案中,组合物可以是本文提供的任何方法中涉及的任何组分或来自本文提供的任何方法的产品。

[0134] 还提供了各种细胞组合物、文库、细胞和蛋白质相关的治疗剂。在一些实施方案中,共培养物至少包含细胞群的第一类的细胞和第二类的细胞。在一些实施方案中,共培养物中的第一类的细胞和第二类的细胞可以接触并诱导表型变化。在一些实施方案中,共培养物被保持在培养皿中。在一些实施方案中,共培养物被保持在培养袋(culture bag)中。在一些实施方案中,第一类的细胞和第二类的细胞是不同细胞类型的两个或多个群。在一些实施方案中,细胞群恰好是不同细胞类型的两个群。在一些实施方案中,本文提供的任何方法的任何细胞集合或中间体或所得细胞群被认为是特定的组合物、文库、治疗剂等。在一些实施方案中,共培养物包含T细胞和B细胞。

[0135] 在一些实施方案中,共培养物中的第一类细胞是T细胞。在一些实施方案中,第一类包括人T细胞。在一些实施方案中,第一类包括Jurkat T细胞。在一些实施方案中,第一类包括经工程改造以表达人CD8a和CD8b且缺乏内源性TCR表达的Jurkat T细胞。在一些实施方案中,第一类包括表达一种或多种变体核酸分子的Jurkat T细胞。在一些实施方案中,第一类包括表达一种或多种变体TCR的Jurkat T细胞。在一些实施方案中,由于低密度预培养,Jurkat T细胞表达低背景水平的活化标志物,包括但不限于CD69。

[0136] 在一些实施方案中,共培养物中的第二类的细胞包括能够呈递抗原的细胞。在一些实施方案中,其他类型的细胞包括能够呈递抗原的人细胞。在一些实施方案中,其他类型的细胞包括人肿瘤细胞。在一些实施方案中,另一种类型的细胞包括人B细胞。在一些实施方案中,另一种类型的细胞包括人自体B细胞。在一些实施方案中,其他类型的细胞包括人自体永生B细胞。在一些实施方案中,B细胞群经工程改造以表达外源抗原。在一些实施方案中,B细胞群经工程改造以表达多种外源抗原。在一些实施方案中,B细胞群经工程改造以表达多种外源性新生抗原(exogenous neo-antigen)。在一些实施方案中,B细胞群经工程改造以表达多种小基因(minigene)形式的多种外源性新生抗原。在一些实施方案中,B细胞

群经工程改造以表达单一TMG (single TMG) 形式的多种外源性新生抗原。在一些实施方案中, B细胞群经工程改造以表达多种TMG (multiple TMG) 形式的多种外源性新生抗原。在一些实施方案中, B细胞群中的单个细胞仅表达单个外源小基因或TMG。在一些实施方案中, B细胞群中的单个细胞可以表达多个外源小基因或TMG。

[0137] 在一些实施方案中, 提供了一种组合物。该组合物包含: i) 通过一种或多种T细胞活化标志物测定为活化的第一T细胞群; 和ii) 另一选择的第二T细胞群作为参考群 (reference population), 其表达相同的TCR文库或在相同的TCR文库的质粒池 (plasmid pool) 中, 其中相对于第二T细胞群, 一种或多种TCR在第一T细胞群中富集。

[0138] 在一些实施方案中, 提供了一种细胞集合, 所述集合包括被构造为表达至少一种TCR α 和TCR β 对的一组T细胞, 其中所述对来自一个个体, 其中所述T细胞不表达内源性TCR, 并且其中所述一组T细胞被构造为用于一种或多种T细胞活化标志物的激活; 和一组B细胞, 其中所述一组B细胞被构造为表达至少一种外源性新生抗原, 并且其中所述至少一种外源性新生抗原来自所述个体的肿瘤。

[0139] 在一些实施方案中, 所述集合中存在至少两个TCR对和至少两个外源性新生抗原。在一些实施方案中, 组合物中存在至少2个、3个、4个、5个、6个、7个、8个、9个、10个、50个、100个、500个、1000个、10000个、100000个或100万个TCR对 (或包括这些对的细胞)。在一些实施方案中, 所述集合中存在至少2个、3个、4个、5个、6个、7个、8个、9个、10个、50个、100个、500个、1000个、10000个、100000个或100万个抗原 (或表达这些抗原的B细胞)。在一些实施方案中, 组合物中存在至少2个、3个、4个、5个、6个、7个、8个、9个、10个、50个、100个、500个、1000个、10000个、100000个或100万个TCR对 (或包含这些对的细胞), 并且所述集合中存在至少2个、3个、4个、5个、6个、7个、8个、9个、10个、50个、100个、500个、1000个、10000个、100000个或100万个抗原。在一些实施方案中, 每对或其部分的序列是不同的。在一些实施方案中, 每个TCR对是不同的。

[0140] 在一些实施方案中, 两种以上细胞群存在于共培养物中的比例在1000:1至1:1000的范围内。

[0141] 在一些实施方案中, 设定两种以上细胞群存在于共培养物中的比率以允许诱导接触诱导的表型 (contact-induced phenotype)。

[0142] 在一些实施方案中, B细胞群和T细胞群在共培养物中存在的比例允许T细胞活化。

[0143] 在一些实施方案中, B细胞群和T细胞群在共培养物中存在的比例允许表达特定TCR的T细胞亚类的T细胞活化。在一些实施方案中, 两种以上细胞群存在于共培养物中的比例在1000:1至1:1000的范围内。在一些实施方案中, 该比例足以导致TCR细胞活化。

[0144] 在一些实施方案中, 共培养物或组合物包含a) 经工程改造以表达人CD8a、CD8b和TCR变体文库且缺乏内源性TCR表达的Jurkat T细胞; 和b) 永生化的自体人B细胞, 其中所述自体人B细胞以单个小基因或单个TMG的形式表达多种抗原。在一些实施方案中, 将共培养物保持在培养皿或培养袋中, 并且B细胞群和T细胞群以允许T细胞活化的比例存在。抗原特异性T细胞活化仅通过TCR文库中的特定TCR介导。

[0145] 在一些实施方案中, T细胞 (表达TCR对) 与B细胞 (提供新生抗原) 的数量和/或比率为至少2比2, 例如, 至少10比10、至少1000 (TCR对) 比10 (新生抗原)、至少10000 (TCR对) 比10 (新生抗原)、至少10000 (TCR对) 比100 (新生抗原)、至少10000 (TCR对) 比1000 (新生抗原)、

至少100000 (TCR对) 比10 (新生抗原)、100000 (TCR对) 比100 (新生抗原)、100000 (TCR对) 比1000 (新生抗原)、100000 (TCR对) 比10000 (新生抗原)、至少50比50、至少100比100、至少1000000 (TCR) 比10000 (抗原), 包括在任意两个前述比率之间定义的任何范围。在一些实施方案中, 单个TCR对和/或抗原存在于每个细胞 (T或B) 中, 使得上述每个数字也可以代表细胞数。

[0146] 在一些实施方案中, 组合物包含B细胞和T细胞的共培养物, 其中存在由T细胞表达的至少两种不同的TCR对, 并且其中存在由B细胞表达的至少两种不同的抗原。

[0147] 在一些实施方案中, T细胞中存在多种TCR, 并且B细胞中存在多种抗原。在一些实施方案中, 组合物被构造为可以诱导由至少一种TCR和一种抗原 (或至少两种以上TCR和/或两种以上抗原) 介导的T细胞活化。在一些实施方案中, 组合物中存在低背景水平的CD69。在一些实施方案中, 组合物包括自体APC (或自体永生生化B细胞)。在一些实施方案中, 对T细胞进行工程改造以表达TCR文库 (和/或缺乏内源性TCR表达)。在一些实施方案中, T细胞被构造为使得它们被剥夺天然TCR表达, 但能够通过外源TCR对激活TCR。

[0148] 在一些实施方案中, 提供了一种细胞集合。所述集合包括: 至少两种T细胞的组。在一些实施方案中, 每T细胞均被构造为表达至少一个TCR α 和TCR β 对。在一些实施方案中, TCR α 和TCR β 均来自一个个体, 所述T细胞不表达内源性TCR, 并且所述组被构造用于一个或多个T细胞活化标志物的激活。所述集合进一步包括一组至少两种B细胞的组。所述至少两种B细胞中的每一个均被构造为表达至少一种外源性新生抗原 (或抗原), 使得存在至少两种能够被产生的外源性新生抗原 (或抗原), 并且所述至少两种外源性新生抗原 (或抗原) 与所述个体中的外源性新生抗原 (或抗原) 相同。

[0149] 在一些实施方案中, 在所述细胞组合物中, 所述至少两种B细胞的组包括: 至少一个产生外源性新生抗原 (或抗原) 的第一B细胞; 和至少一个产生第二外源性新生抗原 (或抗原) 的第二B细胞。

[0150] 在一些实施方案中, 提供了TCR (或TCR表达细胞) 文库。所述文库包括: 至少三种T细胞的组, 其中所述T细胞中的至少两个被构造为表达至少两个TCR α 和TCR β 对 (至少两个TCR对), 其中所述至少两个TCR对来自一个个体, 其中所述至少三种T细胞不表达内源性TCR, 其中所述至少三种T细胞被构造用于在结合由B细胞呈递的抗原 (或新生抗原) 时一个或多个T细胞活化标志物的激活, 其中反映在若干个TCR细胞中的每个TCR对的基因组拷贝的数量使得可以读取样本中的每个TCR, 并且其中至少一种所述TCR在整个包含所述文库的组合物中分布不均。

[0151] 在一些实施方案中, 至少一种T细胞的分布通过与B细胞呈递的抗原结合而改变。在一些实施方案中, 至少两个TCR对近似均匀地存在于文库中。

[0152] 在一些实施方案中, 提供了一种细胞集合。所述集合包括: 至少两种T细胞的组, 其中每个T细胞均被构造为表达至少一个TCR α 和TCR β 对, 其中所述对来自一个个体, 其中所述T细胞不表达内源性TCR, 并且其中所述组被构造用于一个或多个T细胞活化标志物的激活。所述细胞集合可以进一步包括至少两种抗原呈递细胞 (APC) 的组, 其中所述至少两个APC中的每一个均被构造为表达至少一种外源性新生抗原 (或抗原), 使得存在至少两种能够被产生的外源性新生抗原 (或抗原), 并且其中所述至少两种外源性新生抗原 (或抗原) 与所述个体中的外源性新生抗原 (或抗原) 相同。

[0153] 在一些实施方案中, 套装(set)或试剂盒包括第一T细胞群和第二T细胞群。在一些实施方案中, 第一T细胞群由共有(share)可测量的特定表型的T细胞组成。在一些实施方案中, 第一T细胞群包括共有T细胞活化表型的T细胞。在一些实施方案中, 第一T细胞群可以是选定的T细胞群。

[0154] 在一些实施方案中, 所述第一T细胞群包括共有一定表达水平的一种或多种标志物的T细胞。在一些实施方案中, 第一T细胞群包括共有一种或多种T细胞活化标志物表达的T细胞。

[0155] 在一些实施方案中, 所述T细胞表达变体核酸分子文库。在一些实施方案中, 所述T细胞表达TCR文库。

[0156] 在一些实施方案中, 第二T细胞群可以是参考群。参考群可以是选定或未选定的T细胞群, 其表达的TCR文库与选定的第一T细胞群所表达的相同。在一些实施方案中, 参考物(reference)是由第一T细胞群表达的相同TCR文库的质粒池。

[0157] 在一些实施方案中, 每种TCR的量使得可以读取样本中的每种TCR, 并且其中存在至少一个在整个组合中不均匀分布的TCR对。在一些实施方案中, 反映在若干个TCR细胞中的每个TCR对的基因组拷贝的数量使得可以读取样本中的每个TCR。在一些实施方案中, 至少一个TCR在包含文库的整个组合中表达不平均。在一些实施方案中, 大多数TCR在选定的T细胞群和参考细胞群中大致同等地表达(represented)。在一些实施方案中, 存在于TCR文库中的所有TCR的90%以上在第一T细胞群和第二T细胞群(例如, 参考群体)中均有表达。在一些实施方案中, 存在于TCR文库中的所有TCR的99%以上在第一T细胞群和第二T细胞群中均有表达。

[0158] 在一些实施方案中, 一种或多种TCR在第一群中相对于第二群富集。在一些实施方案中, 一种或多种TCR在第一群中相对于第二群在统计学上显著富集。

[0159] 在一些实施方案中, 所述组合包含由筛选方法(例如, 顶部-底部比较(top-bottom comparison))得到的前(top)1%、2%、3%、4%、5%、6%、7%、8%、9%或至少10%的TCR(或表达这些TCR的细胞)。

[0160] 在一些实施方案中, 所有TCR的99%以上存在于第一(选定的)T细胞群中。在一些实施方案中, 组合中的大多数TCR大致均匀地分布在第一T细胞群和第二T细胞群(例如, 这种情况下的参考细胞)中。

[0161] 在一些实施方案中, 识别、分离和/或提供具有期望特征的至少一个TCR $\alpha\beta$ 对。在一些实施方案中, 具有期望特征的至少一个TCR $\alpha\beta$ 对源自肿瘤浸润淋巴细胞(tumor-infiltrating lymphocyte, TIL)。在一些实施方案中, 具有期望特征的至少一个TCR $\alpha\beta$ 对源自肿瘤浸润淋巴细胞(TIL), 并用于同一个体中的癌症治疗。在一些实施方案中, 具有期望特征的至少一个TCR $\alpha\beta$ 对源自外周血。在一些实施方案中, 所述期望特征是对抗原的特异性。在一些实施方案中, 期望特征是新生抗原的识别。在一些实施方案中, 期望特征是病毒抗原的识别。在一些实施方案中, 期望特征是由肿瘤细胞表达的共有抗原的识别。在一些实施方案中, 期望特征是对MHC I类或MHC II类的限制。在一些实施方案中, 期望特征是对抗原的亲合力。在一些实施方案中, 期望特征是对抗原缺乏反应性。在一些实施方案中, 需要多个特征。在一些实施方案中, 该TCR对被构造用于这些特征中的任意一个或多个。

[0162] 在一些实施方案中, 使用和/或制备和/或调节具有期望特征的至少一个TCR $\alpha\beta$ 对

用于治疗。在一些实施方案中,使用和/或制备和/或调节至少一个TCR $\alpha\beta$ 对用于治疗。在一些实施方案中,使用或构造至少一个TCR $\alpha\beta$ 对用于癌症治疗。在一些实施方案中,使用和/或制备和/或调节至少一个TCR $\alpha\beta$ 对用于感染性疾病的治疗。在一些实施方案中,使用和/或制备和/或调节至少一个TCR $\alpha\beta$ 对用于自身免疫性疾病的治疗。在一些实施方案中,使用和/或制备和/或调节至少一个TCR $\alpha\beta$ 对以改造重组蛋白用于治疗。在一些实施方案中,施用重组蛋白用于治疗。

[0163] 在一些实施方案中,使用至少一个TCR $\alpha\beta$ 对来改造T细胞以用于治疗。在一些实施方案中,使用至少两个TCR $\alpha\beta$ 对来改造T细胞以用于治疗。在一些实施方案中,使用两个以上的TCR $\alpha\beta$ 对来改造T细胞以用于治疗。在一些实施方案中,使用五个TCR $\alpha\beta$ 对来改造T细胞以用于治疗。在一些实施方案中,使用十个TCR $\alpha\beta$ 对来改造T细胞以用于治疗。在一些实施方案中,使用二十个TCR $\alpha\beta$ 对来改造T细胞以用于治疗。在一些实施方案中,施用改造的T细胞以进行治疗。在一些实施方案中,使用病毒将TCR $\alpha\beta$ 对引入T细胞。在一些实施方案中,病毒是慢病毒。在一些实施方案中,病毒是逆转录病毒。在一些实施方案中,病毒是腺病毒。在一些实施方案中,病毒介导TCR整合到T细胞的基因组中。

[0164] 在一些实施方案中,病毒导致TCR的瞬时表达。在一些实施方案中,病毒携带TCR DNA作为通过同源定向修复 (Homology-directed-repair,HDR) 在T细胞中进行基因组双链断裂的修复模板。

[0165] 在一些实施方案中,使用非病毒基因递送方法 (non-viral gene delivery method) 将TCR $\alpha\beta$ 对引入T细胞。在一些实施方案中,非病毒基因递送方法基于电穿孔。在一些实施方案中,非病毒基因递送方法基于可向T细胞引入细胞膜的暂时穿孔以将组分递送到T细胞中的其他方法。在一些实施方案中,非病毒基因递送方法涉及转座酶。在一些实施方案中,非病毒基因递送方法涉及核酸酶。

[0166] 在一些实施方案中,核酸酶是CRISPR/Cas9复合物。

[0167] 在一些实施方案中,用TCR修饰被工程改造的T细胞,并进一步进行遗传修饰以控制其表型和反应性。

[0168] 在一些实施方案中,将表达对不同抗原具有特异性的不同TCR $\alpha\beta$ 对的被工程化改造的T细胞组合成细胞组合物施用。在一些实施方案中,该组合允许靶向多种抗原,其可能比单一疗法更有效。在一些实施方案中,该组合允许MHC-I类和MHC-II类限制性T细胞一起使用,它们可协同用于实体癌 (solid cancer) 的治疗。在一些实施方案中,该组合允许一起靶向主干和分支肿瘤突变。在一些实施方案中,该组合基于利用相等比例的每种工程改造的T细胞群。在一些实施方案中,该组合基于利用不同细胞数量的每种工程改造的T细胞群。

[0169] 在一些实施方案中,该组合物包含基于使用多于一种TCR基因的工程改造的T细胞产品。在一些实施方案中,工程改造的T细胞包括以下至少一种:它对于接受它的患者是自体的;它是TIL衍生的;它使用至少一种I类TCR和一种II类TCR;以及它采用相等的比率。

[0170] 本文提供的一些实施方案规避了恢复TCR α -链和TCR β -链的天然组合的需要,并且可以应用于非活细胞材料和非活组织样本。通过生成组合TCR $\alpha\beta$ 文库,本发明的一些实施方案允许从储存或存档的样本中识别抗原特异性TCR $\alpha\beta$ 对。例如,本发明的实施方案可以解决由非存活T细胞的细胞膜完整性的丧失引起的与来自不同T细胞的TCR α 和TCR β mRNA转录物的混合相关的问题。在此类混合中,原始TCR $\alpha\beta$ 对的信息丢失了。本文提供的方法的一些实

施方案解决了先前描述的TCR文库筛选技术由于恢复的TCR文库偏向高频率的TCR序列而引起的低灵敏度。本文提供的方法的一些实施方案消除了在下游应用中包括恢复的TCR链的完整库的需要,允许例如将TCR发现集中于具有期望性质的TCR链。与单细胞方法例如液滴PCR和微流体装置不同,从本文描述的T细胞中恢复特定TCR $\alpha\beta$ 对的方法不需要限制可扩展性的特定仪器和活细胞材料。与从先验(a priori)不允许恢复天然TCR $\alpha\beta$ 对的T细胞群中恢复TCR α 序列和TCR β 序列的集合的批量PCR方法不同,本文描述的方法采用一种设计来恢复所识别的TCR库的限定部分以恢复目的TCR $\alpha\beta$ 对。

[0171] 本文描述的方法可用于治疗和诊断目的和/或组合物和/或药物。重组TCR基因可用于产生对免疫疗法(例如,包括癌症免疫疗法)具有期望的特异性的T细胞。对于癌症免疫治疗应用,可通过选择TCR基因,通过TCR基因转移产生抗原特异性T细胞,从而产生具有期望的特异性的T细胞。在一些实施方案中,该方法基于抗原特异性可以通过编码TCR $\alpha\beta$ 对的基因的转移而在T细胞之间转移这一观察。可以通过利用 γ -逆转录病毒载体或慢病毒载体、基于转座子的基因递送平台、mRNA递送(例如通过电穿孔或纳米颗粒)或包括CRISPR/Cas9的基因组工程工具,将目的TCR基因引入人T细胞的基因组中。后者通过将新的TCR链位点特异性整合到内源TCR位点,能够同时敲除内源TCR链。

[0172] 在一些实施方案中,所得的选定分子对用于治疗个体和/或用于个体的药物,以治疗本文提供的任何病症。

[0173] 在一些实施方案中,提供了治疗个体的方法。该方法包括识别患有肿瘤的个体;提供至少两种T细胞的组,每个T细胞均被构造为表达至少一个不同的TCR α 和TCR β 对,其中所述TCR α 和所述TCR β 各自来自所述个体,提供至少两种B细胞的组,其中所述B细胞组经构造以表达至少两种外源性新生抗原,并且其中所述至少两种外源性新生抗原与在所述个体中发现的那些新生抗原相同;将所述至少两种T细胞的组与所述至少两种B细胞的组组合,并基于所述至少两种T细胞经由所述至少两种外源性新生抗原的激活来选择至少两个TCR对的组合;以及向所述个体施用所述至少两个TCR对的组合,从而治疗所述肿瘤。在一些实施方案中,在本方法中可以使用本文提供的任何T细胞与B细胞的数量比。在一些实施方案中,所述治疗减小肿瘤的大小。在一些实施方案中,所述治疗使肿瘤的大小减小至少1%、2%、3%、4%、5%、10%、20%、30%、40%、50%、60%、70%、80%、90%、95%、99%或100%,包括在任意两个前述值之间限定的任何范围。

[0174] 在一些实施方案中,提供了治疗个体的方法。该方法包括识别患有肿瘤的个体;提供至少两种T细胞的组,每个T细胞均被构造为表达至少一种不同的TCR α 和TCR β 对,其中所述TCR α 和TCR β 各自来自所述个体;提供至少两种抗原呈递细胞的组,其中所述抗原呈递细胞的组源自所述个体,被构造为表达至少两种外源性新生抗原,并且其中所述至少两种外源性新生抗原与在所述个体中发现的那些新生抗原相同;将所述至少两种T细胞的组与所述至少两种抗原呈递细胞的组组合,并基于所述至少两种T细胞经由所述至少两种外源性新生抗原的激活来选择至少两个TCR对的组合;和向所述个体施用所述至少两个TCR对的组合,从而治疗所述肿瘤。在一些实施方案中,可以采用两个以上TCR对,例如,2个、3个、4个、5个、6个、7个以上TCR对。在一些实施方案中,通过细胞疗法施用TCR对。在一些实施方案中,所述对具有不同于其他对的序列。

[0175] 在一些实施方案中,本文通过任何方法提供的任何选定的TCR对或对的组合可用

于本文提供的治疗方法中。

[0176] 在一些实施方案中,提供了一种药物组合物。在一些实施方案中,该药物组合物包含第一TCR对,其与个体的肿瘤中的第一抗原(或新生抗原)结合;和第二TCR对,其与所述个体的肿瘤中的第二抗原(或新生抗原)结合。在一些实施方案中,可以采用两个以上的TCR对,例如,2个、3个、4个、5个、6个、7个以上的TCR对。在一些实施方案中,通过细胞疗法施用TCR对。在一些实施方案中,所述对具有不同于其他对的序列。在一些实施方案中,第一TCR对是MHC I类限制性的,并且其中所述第二TCR对是MHC II类限制性的。

[0177] 在一些实施方案中,提供了一种药物组合物。它可以包括第一TCR对,其与第一抗原结合并且是MHC I类限制性的;和第二TCR对,其与第二抗原结合并且是MHC II类限制性的。

[0178] 在一些实施方案中,组合物可以进一步包含第三TCR对。

[0179] 在一些实施方案中,所述第一TCR对与来自肿瘤的新生抗原结合,其中所述第二TCR对与来自所述肿瘤的新生抗原结合,并且其中所述第一TCR对和第二TCR对二者均存在于所述肿瘤的宿主中。

[0180] 用于癌症治疗用途的重组TCR基因可从不同来源获得。首先,可以从活的患者标本如血液或肿瘤组织中检测和分离对肿瘤抗原具有特异性的T细胞。本领域描述的技术包括分离MHC-多聚体结合T细胞,以及通过流式细胞术或磁珠选择在抗原刺激后分离表达某些表型的标志物或分泌某些细胞因子的T细胞。随后,分离的抗原特异性T细胞可用于通过基于单细胞PCR的技术、TCR批量测序或基于微流体的PCR技术来确定表达的TCR基因的序列。第二,异体CTL系统或动物模型(例如HLA转基因和/或人TCR转基因小鼠模型)为肿瘤抗原特异性T细胞/TCR提供了替代来源。第三,治疗性TCR基因可以通过噬菌体-、酵母-或T细胞-展示系统从以重组TCR文库形式表达的体外突变的TCR链中选择。

[0181] 用于癌症治疗的T细胞(或TCR)可根据期望的治疗标准进行选择;首先,用于癌症治疗的TCR基因理想地识别在重要组织中低表达或无表达的肿瘤特异性抗原。第二,TCR应以高灵敏度识别其抗原,例如少量抗原应触发经TCR修饰的T细胞针对肿瘤细胞的效应功能,例如溶细胞活性。第三,TCR应与在重要组织中表达的其他抗原无交叉反应性。

[0182] 通过TCR基因转移可靶向不同的肿瘤抗原,包括细胞谱系特异性抗原(例如MART-1)、过表达抗原(如WT-1)、癌症/睾丸(C/T)抗原(例如NY-ESO-1、MAGE-A4、MAGE-A10)、病毒抗原(例如HPV E6、E7)和突变蛋白(新生抗原)。值得注意的是,新生抗原特异性TCR序列可能特别适用于癌症的治疗。例如,在免疫检查点阻断治疗和过继性T细胞治疗后,新生抗原特异性T细胞与晚期转移性癌症消退相关。由于肿瘤中的绝大多数基因组突变是乘客突变(passenger mutation),并且仅在一小部分肿瘤中发现,并且由于MHC限制,个体患者之间T细胞可识别的肿瘤新生抗原库有很大差异。因此,利用TCR基因转移产生用于治疗的新生抗原特异性T细胞通常需要针对每个患者或肿瘤的一个或多个新的新生抗原特异性TCR序列。鉴于用于TCR基因转移的任何TCR的安全性要求,商业上可扩展的方法理想地依赖于自体组织,因为可以假定直接来自患者的新生抗原特异性TCR是安全的。此外,例如优选使用非活组织,例如存档的肿瘤样本,以实现商业上可扩展的过程,因为其避免了处理存活患者细胞以进行TCR分离的需要。本文描述的方法解决了在每位患者的基础上以高灵敏度识别相关新生抗原特异性TCR的重大需求。此外,如本文所述方法中描述的新生抗原特异性TCR基因

转移可使未受益于其它疗法例如免疫检查点阻断的患者受益。

[0183] 在一些实施方案中,提供了从核酸组合文库中识别编码T细胞受体 α (TCR α)-链和TCR β -链的核苷酸序列的方法。所述方法包括:a) 提供包含多种变体核酸的文库,所述多种变体核酸中的每一个均包含至少600bp的连续部分,其中所述连续部分包含以下的组合:1) 第一变体核苷酸子序列,其编码TCR α 变体氨基酸序列并限定所述连续部分的第一端,和2) 第二变体核苷酸子序列,其编码TCR β 变体氨基酸序列并限定与所述第一端相对的所述连续部分的第二端。所述方法还包括将所述文库导入经构造以表达由所述多种变体核酸的成员编码的TCR α -链和TCR β -链的永生化T细胞群中,并基于响应于使永生化T细胞与表达抗原的永生化B细胞接触的高于阈值水平的T细胞活化标志物的表达选择所述永生化T细胞群的亚群,其中所述亚群包括多个T细胞和/或从所述亚群中分离所述多种变体核酸的子集;和/或测定所述子集的个体成员的所述连续部分的核苷酸序列;和/或基于所述子集的核苷酸序列中的至少一种组合相对于对照的富集来识别所述第一变体核苷酸子序列和第二变体核苷酸子序列的至少一种组合。

[0184] 在一些实施方案中,提供了一种从核酸组合文库中识别编码嵌合抗原受体 (CAR) 铰链结构域、跨膜结构域和/或细胞内信号传导结构域的核苷酸序列的方法。该方法包括:提供包含多种变体核酸的文库,所述多种变体核酸中的每一个均包含至少600bp的连续部分,其中所述连续部分包含以下两种以上的组合:1) 编码CAR铰链结构域的第一变体核苷酸子序列;2) 编码CAR跨膜结构域的第二变体核苷酸子序列;和3) 编码CAR细胞内信号传导结构域的第三变体核苷酸子序列。第一、第二或第三变体核苷酸子序列中的一个限定了所述连续部分的第一端,并且第一、第二或第三变体核苷酸子序列中的另一个限定了与所述第一端相对的所述连续部分的第二端。所述方法进一步包括将所述文库导入经构造以表达由所述多种变体核酸的成员编码的CAR的细胞群中,其中所述细胞群包含永生化T细胞群或原代人T细胞群。该方法可以进一步包括基于响应于使细胞与表达对所述CAR的抗原结合结构域特异的抗原的抗原呈递细胞接触的、高于阈值水平的细胞增殖选择所述细胞群的亚群,其中所述亚群包括多个细胞。所述方法可以进一步包括从所述亚群中分离所述多种变体核酸的子集,和/或测定所述子集的个体成员的所述连续部分的核苷酸序列;和/或基于子集的所述核苷酸序列中的至少一种组合相对于对照的富集,识别所述第一、第二和第三变体核苷酸子序列的至少一种组合。

[0185] 定义

[0186] 在整个说明书中,词语“包括 (comprise)”或例如“包含 (comprises)”或“含有 (comprising)”的变体将被理解为暗示包括所述的要素、整数或步骤,或元素组、整数组或步骤组,但不排除任何其他要素、整数或步骤,或元素组、整数组或步骤组。

[0187] 提供以下术语和方法的解释以更好地描述本发明,并指导本领域普通技术人员实践本发明。除非上下文另有明确说明,否则单数形式“一个 (a)”、“一个 (an)”和“该 (the)”指一个或一个以上。例如,术语“包括一个核酸分子”包括单个或多个核酸分子,并且被认为等同于短语“包括至少一个核酸分子”。除非上下文另有明确说明,否则术语“或 (or)”指所述替代要素的单个要素或两个以上要素的组合。如本文所用,“包含 (comprise)”意指“包括 (include)”。因此,“包含A或B”意指“包括A、B或A和B”,而不排除附加要素。除非另有说明,否则当本定义可能与其他可能的定义不同时,以本文提供的定义为准。

[0188] 除非另有说明,否则本文使用的所有技术术语和科学术语具有与本发明所属领域的普通技术人员通常理解的含义相同的含义。本文提及的所有HUGO基因命名委员会(HGNC)标识符(ID)均通过引用整体并入本文。尽管与本文所述的方法和材料相似或等效的方法和材料可用于本发明的实践或测试,但下文描述了合适的方法和材料。材料、方法和实施例仅是说明性的,而不是限制性的。

[0189] “T细胞受体”或“TCR”表示在T细胞或T淋巴细胞表面发现的分子,其识别作为肽与主要组织相容性复合体(major histocompatibility complex,MHC)分子结合的抗原。MHC分子包括I类、II类和III类。I类和II类MHC分子在免疫应答中均起关键作用。MHC I类分子在所有有核细胞和血小板——实质上在除红细胞以外的所有细胞中均有表达。它为杀伤性T细胞提供表位,也被称为细胞毒性T淋巴细胞(cytotoxic T lymphocyte,CTL)。MHC II类可由所有细胞类型有条件地表达,但通常仅发生于“专职”抗原呈递细胞(APC):巨噬细胞、B细胞,尤其是树突状细胞(DC)。APC摄取抗原性蛋白质,进行抗原加工,并返回其分子部分(称为表位的部分),然后将其展示在与MHC II类分子偶联的APC表面(抗原呈递)。在细胞表面,该表位可被T细胞受体(TCR)等免疫结构识别。在一些实施方案中,TCR包含两条多肽链,TCR α 和TCR β (分别通过TRA和TRB编码)。在一些实施方案中,TCR包括TCR γ 链和TCR δ 链(分别由TRG和TRD编码)。在一些实施方案中,TCR包含细胞外可变区和细胞外恒定区。在一些实施方案中,TCR α 链和TCR β 链的可变结构域包含三个高变互补决定区(CDR),表示为CDR1、CDR2和CDR3。在一些实施方案中,CDR3是主要的抗原识别区。在一些实施方案中,TCR α 链基因包含V和J,而TCR β 链基因包含有助于TCR多样性的V、D和J基因片段。

[0190] 如本文所用,术语“TCR库(TCR repertoire)”指样本或文库中的TCR链的集合。一个集合可以包含至少两个以上不同的TCR链变体。在一些实施方案中,“TCR库”指样本或文库中所有TCR链的集合。在一些实施方案中,“TCR库”是指样本或文库中TCR链的子集(subset)或选择(selection)的集合(collection)。在一些实施方案中,“TCR库”指样本或文库中TCR $\alpha\beta$ 对的集合。在一些实施方案中,“TCR库”指样本或文库中TCR $\alpha\beta$ 对的子集或选择的集合。

[0191] 例如,TCR链的子集或选择可以基于TCR链的频率。在一些实施方案中,“TCR链的频率”指在编码(部分)所有TCR链氨基酸序列的所有核酸的总数中编码(部分)特定TCR链氨基酸序列的核酸分子(RNA和/或DNA)的绝对数量。在一些实施方案中,编码(部分)特定TCR链氨基酸序列的核酸分子的绝对数量可以使用“唯一分子标识符”(UMI)(如例如在Kivioja等人,Nat Meth 2011和Islam等人,Nat Meth 2013中描述的原理)基于唯一分子的计数来确定。在一些实施方案中,TCR链频率可以表示为百分比。所有TCR链的总数可以仅包括编码TCR α -链的核酸分子,仅包括编码TCR β -链的核酸分子或包括编码TCR α -链和TCR β -链二者的核酸分子。在一些实施方案中,频率表示为样本中TCR链具有的频率等于或高于、等于、高于或低于0.001%、0.01%、0.1%、1%、2%、3%、4%、5%、6%、7%、8%、9%、10%、20%、30%、40%、50%、60%、70%、80%、90%、100%或它们之间的任何数量或范围。使用编码(部分)TCR链氨基酸序列的核酸分子的绝对数量或相应的百分比,可以获得TCR链的秩次序(rank order)。在一些实施方案中,频率被表示为TCR链位于样本的秩次序的前1000、前900、前800、前700、前600、前500、前450、前400、前350、前300、前250、前200、前150、前140、前130、前120、前110、前100、前90、前80、前70、前60、前50、前40、前30、前20、前10、前5或它

们之间的任何数量或范围。在一些实施方案中，“TCR链的频率”指TCR链相对于样本中全部TCR链的频率。在一些实施方案中，“TCR链的频率”指TCR链相对于样本中少于全部的TCR链或相对于样本中TCR链的子集(subset)的频率。

[0192] 如本文所用,术语“频率阈值”指待包含于TCR链的子集或选择中的给定的TCR链在样本中出现的最小频率。在一些实施方案中,TCR链的频率阈值包括位于样本的秩次序的前1000、前900、前800、前700、前600、前500、前450、前400、前350、前300、前250、前200、前150、前140、前130、前120、前110、前100、前90、前80、前70、前60、前50、前40、前30、前20、前10、前5或任何之间的数字或范围。在一些实施方案中,频率阈值表示为包括样本中等于和高于、等于和低于、等于、高于或低于0.001%、0.01%、0.1%、1%、2%、3%、4%、5%、6%、7%、8%、9%、10%、20%、30%、40%、50%、60%、70%、80%、90%、100%或它们之间的任何数量或范围的所有TCR链。在一些实施方案中,“频率阈值”指相对于样本中全部TCR链的阈值。在一些实施方案中,“频率阈值”指相对于样本中少于全部的TCR链或相对于样本中TCR链的子集的阈值。

[0193] 如本文所用,术语“相对富集”指一个样本中的TCR链比另一个样本中的TCR链具有更高的丰度或频率。可以比较的样本包括,例如,肿瘤样本和血液、不同的肿瘤样本、肿瘤样本和非肿瘤样本、来自同一肿瘤的不同区域的样本、来自肿瘤核心和肿瘤边界(boundary)或边缘(margin)的样本、来自具有不同活化或分化状态的T细胞的样本等。

[0194] 如本文所用,根据本领域技术人员使用“开关受体(switch receptor)”,并且包括:开关受体包括但不限于用于将通常与T细胞抑制或凋亡相关的细胞外信号转化为T细胞激活信号的受体分子。这可以通过将结合抑制性配体或凋亡诱导配体(例如但不限于TGFR2、FAS或TIGIT)的胞外结构域(ECD)与来自T细胞激活受体(例如CD3 ϵ 、CD28、IL2RB)的胞内信号传导结构域(intracellular signaling domain,ISD)融合来实现。本领域技术人员可以理解,融合受体分子也可以设计成通过将结合T细胞激活配体的ECD与来自T细胞抑制受体(例如PD-1或CTLA-4)的ISD结合来抑制T细胞功能。在一些实施方案中,开关受体可以包含但不限于与一个、两个或甚至更多个信号传导结构域融合的ECD。在一些实施方案中,开关受体分子包括但不限于包含以下组分的受体分子:除了ECD和ISD之外还包含不同跨膜结构域(TM)或任何其他新的组分,包括但不限于ECD和TM和/或TM和ISD的不同结构域之间的接头或间隔序列。

[0195] 如本文所用,根据本领域技术人员使用“单链TCR”。该术语进一步包括但不限于用接头共价连接TCR α 和TCR β 可变链片段。单链TCR包括但不限于用与TCR β 恒定区融合的接头共价连接TCR α 和TCR β 可变链片段,并与TCR α 恒定区反式共表达。在一些实施方案中,单链TCR包括但不限于用与TCR α 恒定区融合的接头共价连接TCR α 和TCR β 可变链片段,并与TCR β 恒定区反式共表达。单链TCR的一些实施方案包括但不限于用接头共价连接TCR α 和TCR β 可变链片段,并单独或与CD28信号传导结构域组合地与CD3 ϵ 或CD3 ζ 信号传导结构域融合。

[0196] 如本文所用,术语“基因表达的空间模式”指基因在特定区域或空间中的表达。在一些实施方案中,“基因表达的空间模式”指基因在例如肿瘤的组织内的表达,即肿瘤内的表达。在一些实施方案中,“基因表达的空间模式”指在以一种或多种表型标志物的表达或不表达为特征的区域或空间中基因表达的富集。在一些实施方案中,表型标志物可以是与表型相关的任何标志物,包括但不限于一种或多种表面标志物或其片段、一种或多种蛋白

质或其片段、一种或多种RNA如微RNA (microRNA)、干扰小RNA (siRNA) 或任何其他RNA。在其他实施方案中,“基因表达的空间模式”指一个基因的表达或不表达与至少一个其他基因的表达或不表达的组。

[0197] 如本文所用,术语“共表达模式”包括一种或多种基因在相同细胞或相同组织样本中的表达。在一些实施方案中,术语“共表达模式”指在同一细胞或同一组织样本中不存在一种或多种基因的表达。

[0198] 术语“癌症”指经历了丧失分化、生长速率增加、侵入周围组织的特征性间变并能够转移的恶性赘生物(neoplasm)。术语“癌症”应理解为包括以个体体内细胞不受控制的生长为特征的疾病。在一些实施方案中,术语“癌症”和“肿瘤”可互换使用。在一些实施方案中,术语“肿瘤”是指良性或非恶性生长。

[0199] 术语“文库(library)”指TCR链的集合。在一些实施方案中,文库包括组合形成TCR $\alpha\beta$ 对的TCR链的集合。在一些实施方案中,文库包括组合形成TCR $\alpha\beta$ 对的TCR链的子集或选择的集合。

[0200] 如本文所用,术语“新生抗原(neo-antigen)”指源自肿瘤特异性基因组突变的抗原。例如,新生抗原可由非同义单核苷酸突变导致的肿瘤样本中突变蛋白的表达产生,或由突变诱导的移码(frame-shift)导致的替代开放阅读框架的表达产生。因此,新生抗原可能与病理状况有关。在一些实施方案中,“突变的蛋白质”指在与典型氨基酸序列的相同位置处包含至少一个不同的氨基酸的蛋白质。在一些实施方案中,突变的蛋白质包括相对于典型氨基酸序列的插入、缺失、取代、由移码产生的氨基酸的包含物或它们的任何组合。

[0201] 术语“治疗(treatment)”包括其在本领域中的普通含义,并且包括减轻疾病的至少一种症状或其他实施方案,或降低疾病严重性等。治疗不需要实现完全治愈,或根除疾病的每个症状或表现,以构成可行的治疗。如在相关领域中所公知的,用作治疗剂的组合物可以降低给定疾病状态的严重性,但不需要消除疾病的每一种表现就被认为是有用的。减少疾病的影响(例如,通过减少其症状的数量或严重性,或通过增加另一种治疗的有效性,或通过产生另一种有益效果),或减少个体中疾病发生或恶化的可能性,就足够了。

[0202] “抗体”表示包括至少一个特异性识别和结合抗原表位的轻链或重链免疫球蛋白可变区的多肽。在一些实施方案中,抗体由重链和轻链组成,每一条链都具有可变区,称为重链(V_H)可变区和轻链(V_L)可变区。 V_H 区和 V_L 区共同负责与抗体识别的抗原结合。术语抗体包括完整的免疫球蛋白及其变体和部分,如Fab' 片段、F(ab')₂片段和任何其他来源于完整免疫球蛋白的分子。

[0203] 如本文所用,短语“B细胞受体(BCR)/抗体库(B-cell receptor(BCR)/antibody repertoire)”指样本中BCR链或抗体链的集合。在一些实施方案中,“BCR/抗体库”指样本中所有BCR链或抗体链的集合。在一些实施方案中,“BCR/抗体库”指样本中BCR链或抗体链的子集或选择的集合。

[0204] 如本文所用,术语“新鲜冷冻的(fresh-frozen)”或“速冻的(snap-frozen)”指在收集后的短时间内冷冻组织或细胞样本。在一些实施方案中,在冷冻前不对组织或细胞样本进行保存。术语“新鲜冷冻的”或“速冻的”可以互换使用。

[0205] 如本文所用,术语“TCR分离(TCR isolation)”包括评估TCR α 链和TCR β 链的哪些特定组合介导所需的功能。术语“TCR分离”可以指单链TCR分子的分离。根据所需功能和TCR盒

的设计,TCR分离方法可能会有所不同。

[0206] 术语“活化标志物(activation marker)”包括如本领域技术人员所理解的该术语的全部范围,并进一步表示响应于外部刺激在细胞内差异调节的一个或多个基因。用作活化标志物的基因可以是细胞基因组的天然部分或通过本领域技术人员已知的基因工程工具引入(例如病毒基因递送)。值得注意的是,差异调节可以描述在RNA水平上检测到的基因表达的增加或减少。在某些情况下,转录水平的这种变化可导致蛋白质水平的可检测变化。通过非限制性实例,T细胞中与T细胞受体触发相关的活化标志物可以包括CD69、CD137、IFN- γ 、IL-2、TNF- α 、GM-CSF、OX40以及人工报道基因如NFAT-GFP或NFAT-嘌呤霉素抗性基因。

[0207] 如本文所用,术语“TCR文库”指编码TCR的质粒的多克隆集合或含有这些质粒的细胞。一个集合可以包含至少两个以上不同的TCR链变体。“TCR文库”可以包括编码TCR的质粒的子集或选择的集合。“TCR文库”可以指可以从编码TCR的质粒的集合中表达的所有TCR的集合。在一些实施方案中,“TCR文库”指可从编码TCR的质粒的集合中表达的TCR的子集或选择的集合。在一些实施方案中,“TCR文库”是指编码TCR的质粒的多克隆集合中的TCR $\alpha\beta$ 对的集合。在一些实施方案中,“TCR文库”指编码TCR的质粒的多克隆集合中的TCR $\alpha\beta$ 对的子集或选择的集合。

[0208] 各种实施方案的一般描述

[0209] 本文所述的一些实施方案涉及从核酸组合文库中识别编码T细胞受体 α (TCR α)-链和TCR β -链的核苷酸序列的方法。在一些实施方案中,所述方法包括(I)提供包含编码TCR α -链和TCR β -链的多种变体核酸的文库,(II)将该文库导入能够表达由该多种变体核酸的成员编码的TCR α -链和TCR β -链的细胞群中,(III)基于响应于抗原的高于阈值水平的标志物的表达来选择该细胞群的亚群,其中所述亚群包含多个细胞,(IV)从所述亚群中分离所述多种变体核酸的子集,(V)测定该变体核酸的核苷酸序列,和(VI)基于所述子集内的核苷酸序列相对于对照的富集来识别至少一个变体核苷酸序列。在一些实施方案中,富集可以是基于使用适当分析软件的统计学富集。在一些实施方案中,R DESeq2套装软件(package)可用于确定富集的显著性。在一些实施方案中,显著性的p值阈值可以限定为0.2、0.1、0.05、0.01、0.001、0.0001、0或这些值之间的任意值。

[0210] 在一些实施方案中,描述了一种从不同的T细胞群恢复T细胞受体(TCR)文库的方法。所述方法可以包括(I)测定个体样本中的TCR- α 和TCR- β 核苷酸序列,(II)基于至少一个标准从总库中选择TCR α -链和TCR β -链的序列的一个或多个子集;(III)通过所选的TCR α -和TCR β -链的序列的组合配对来创建TCR库,从而创建TCR $\alpha\beta$ 对的库;以及IV)从所创建的TCR库中识别具有期望特征的至少一个TCR $\alpha\beta$ 对。图1-图3中提供了该方法的各種实施方案。

[0211] 在一些实施方案中,TCR链选择最初将基于如本文进一步描述的频率阈值。在一些实施方案中,选择将基于一个以上的标准。例如,可以基于所分析的肿瘤类型选择一种以上的选择标准。在一些实施方案中,TCR链选择基于筛选效率的阈值。在一些实施方案中,基于可以有效筛选多少个组合TCR $\alpha\beta$ 链来选择选择标准。作为一个实例,如果可以有效地筛选 1×10^6 个TCR $\alpha\beta$ 对,则可以选择多达1000个TCR α 链和1000个TCR β 链。在一些实施方案中,可以选择不相等数量的TCR α 链和TCR β 链,例如以补偿相当一部分T细胞携带两个框内TCR α 重排(in-frame TCR α rearrangement)的事实。在一些实施方案中,TCR α 与TCR β 的比率可以是例

如100万:1至1:100万。在一些实施方案中,比率是这些范围之间的任何比率,包括例如100000:1、10000:1、1000:1、100:1、10:1、1:1、1:10、1:100、1:1000、1:10000、1:100000。

[0212] 在一些实施方案中,在通过本文所述方法生成的文库中识别的TCR链序列可用于治疗或诊断已自其分离出该TCR链序列的患者。在一些实施方案中,在通过本文所述方法生成的文库中识别的TCR链序列可用于治疗或诊断除了自其分离出该TCR链序列的患者之外的患者。例如,从一名患者中分离的TCR链序列可识别由另一名患者共有的肿瘤抗原。

[0213] 在一些实施方案中,通过本文所述的方法生成大量文库。在一些实施方案中,筛选大量文库允许预测TCR特征,因此允许例如TCR链的特定选择。

[0214] 在一些实施方案中,TCR文库可以包括i) 组合TCR文库;ii) 表达这种文库的细胞;和/或iii) TCR扩增子测序文库。

[0215] 在一些实施方案中,TCR文库是以组合方式精确表达一个 α 和一个 β TCR链的质粒的多克隆集合。文库的性质使得与给定 β 链配对的给定 α 链的频率与该 β 链的总体表示(overall representation)成比例。相反,与给定 α 链配对的给定 β 链的频率与该 α 链的总体表示(overall representation)成比例(由此可以得出结论,可以控制文库中单个链的频率)。在中位频率(median frequency) $\pm 1\log_2$ 单位范围内的单个组合的频率百分比为25%、50%、60%、70%、80%、90%、95%、86%、97%、98%、99%、100%或介于上述任意两个值之间的任何值。

[0216] 在一些实施方案中,文库可以涉及表达相关核酸序列的细胞。表达可包括稳定的或瞬时的方法,并可由DNA/RNA(或其衍生物)赋予。表达这种文库的多克隆池中的细胞数目可以是池中存在的TCR变体数目的20倍、30倍、40倍、50倍、75倍、100倍、125倍、150倍、200倍、250倍、300倍、400倍、500倍、750倍、1000倍、5000倍、10000倍、100000倍、1000000倍。

[0217] 在一些实施方案中,TCR扩增子测序文库是代表给定样本中TCR频率的DNA分子的集合。扩增子包含在给定细胞中表达的关于 α 链和 β 链的信息,并且需要一段(a stretch of)超过600个连续核苷酸的序列来确定V和J区的同一性,以及 α 链和 β 链的CDR3序列。

[0218] 在一些实施方案中,文库是约1.5kb的变体文库。

[0219] 本文所述方法中使用的不同的T细胞群可包括任何谱系的T细胞或其混合物。在一些实施方案中,不同的T细胞群包含CD4或CD8 T细胞。在一些实施方案中,不同的T细胞群包含初始T细胞(naïve T cell)。在一些实施方案中,不同的T细胞群包含效应T细胞。任何类型的效应T细胞均可以在不同的T细胞群中找到,包括Th₁、Th₂、Th₁₇和细胞毒性T淋巴细胞(CTL)。在一些实施方案中,不同的T细胞群包含调节性T细胞(regulatory T cell, T_{reg})。在一些实施方案中,不同的T细胞群包含记忆性T细胞。任何记忆亚型均可以在不同的T细胞群中找到,包括中枢记忆性T细胞(T_{CM}细胞)、效应记忆性T细胞(T_{EM}细胞和T_{EMRA}细胞)、组织驻留记忆T细胞(tissue resident memory T cell, T_{RM})、记忆性干细胞样T细胞(memory stem cell T cell, T_{SCM})和虚拟记忆T细胞。在一些实施方案中,虚拟记忆T细胞包括CD4阳性T细胞。在一些实施方案中,虚拟记忆T细胞包括CD8阳性记忆T细胞。在一些实施方案中,不同的T细胞群包含调节性T细胞。在一些实施方案中,不同的T细胞群包含功能障碍T细胞(dysfunctional T cell)。功能障碍T细胞的特征在于(1)高水平的抑制性受体,(2)在分泌趋化因子CXCL13的同时丧失经典的效应子功能(effector function)(例如IFN- γ 、IL-2和TNF- α),以及(3)与细胞毒性相关的转录谱的高表达水平,包括高水平的颗粒酶B。在一些实

施方案中,不同的T细胞群包含 $\alpha\beta$ T细胞。在一些实施方案中,不同的T细胞群包含 $\gamma\delta$ T细胞。不同的T细胞群可包括自然杀伤T细胞(NKT)和粘膜相关恒定T细胞(mucosal associated invariant T cell, MAIT)。T细胞群可以是不同细胞类型的混合物的一部分,也可以是组织样本的一部分,例如血液或肿瘤组织。不同的T细胞群可以包含不同谱系的T细胞的混合物,也可以包含T细胞和非T细胞的混合物。

[0220] 在一些实施方案中,在个体的样本中测定TCR- α 和TCR- β 的核苷酸序列。利用从样本中获得的DNA或RNA可以测定TCR- α 和TCR- β 的核苷酸序列。在一些实施方案中,测定TCR- α 和TCR- β 的核苷酸序列包括使用多重PCR。在一些实施方案中,测定TCR- α 和TCR- β 的核苷酸序列包括通过靶富集进行TCR序列恢复。例如,TCR基因捕获可用于靶富集(Linnemann et al, Nat Med 2013)。在一些实施方案中,TCR序列恢复包括利用5' RACE和PCR进行恢复。在一些实施方案中,TCR序列恢复包括利用空间测序。

[0221] 在一些实施方案中,从活细胞中分离DNA或RNA。在一些实施方案中,从保存的细胞或保存的组织样本中分离DNA或RNA。保存的细胞和保存的组织样本可以是活的或非活的。保存的组织样本可以包含活细胞或非活细胞,或者活细胞和非活细胞二者的组合。可从通过任何保存方法保存的样本(sample)或样品(specimen)中分离DNA或RNA,包括速冻细胞或组织以及固定的或福尔马林固定的/石蜡包埋的(FFPE)样本。细胞和组织样本的保存方法以及DNA和RNA分离方法是本领域技术人员已知的。在一些实施方案中,样本是肿瘤样本。在一些实施方案中,肿瘤样本是FFPE样本。在一些实施方案中,肿瘤样本是速冻样本。在一些实施方案中,T细胞群是不同细胞类型的混合物的一部分或组织样本或体液如血液、尿液、引流淋巴结(drainage lymph node)或肿瘤组织的一部分。在一些实施方案中,样本是非活肿瘤样本。在一些实施方案中,非活肿瘤样本是速冻样本或FFPE样本。

[0222] 在一些实施方案中,基于至少一个标准从总库中选择TCR α -链和TCR β -链的序列的一个或多个子集:a)基于在T细胞群内的频率,b)基于与第二T细胞群相比的相对富集,c)基于TCR链的生物学特性,其中所述特性选自以下至少一项:(预测的)抗原特异性、(预测的)HLA-限制性、亲和性、共受体依赖性或非代T细胞谱系(例如CD4或CD8 T细胞),d)基于基因表达的空间模式,其中空间基因表达模式衍生自组织中的起始区域或其他基因的表达模式中的至少一种,例如包括共表达,e)基于在多个样本中共现或以相似频率出现,例如在多个肿瘤病灶中出现,f)选择到多个组以分别恢复TCR库的特定部分,g)基于在不同实施方案中限定的多个标准的组合。

[0223] 在一些实施方案中,选择标准可以用于排除而不是包括(包括,例如,上面段落的选项b或c)。这可以应用于本文提供的任何实施方案。因此,它们可以以不向个体施用或提供的方式而应用,或在TCR集合中排除它们的包含项(inclusion)。

[0224] 在一些实施方案中,基于给定TCR链的DNA拷贝数和RNA拷贝数的相对差异,从总库中选择TCR α -链和TCR β -链的序列。例如,基于给定TCR链的基因组DNA和RNA的定量,可以获得RNA衍生拷贝数和基因组拷贝数之间的比率。在一些实施方案中,选择其中RNA拷贝数远高于基因组DNA拷贝数从而导致比值大于1的TCR链。在一些实施方案中,可以对所得到的任何给定TCR相对于样本中的所有其他TCR的比率进行排序和选择。例如,与基于较低比值的较低秩(lower rank)的TCR相比,可以选择基于较大比值的较高秩的TCR,从而选择具有较大RNA拷贝数的TCR链。在一些实施方案中,与基于较大比值具有较高秩的TCR相比,可以选

择基于较低比值具有较低秩的TCR,从而选择具有较低RNA拷贝数的TCR链。可以将秩次序调整为RNA衍生的拷贝数与基因组拷贝数之间的比率任何数值。

[0225] 在一些实施方案中,选择任意数目的TCR α 链和任意数量的TCR β 链,每个链最多包括1000个TCR α 链和TCR β 链。在一些实施方案中,分别选择超过1000个TCR α 链和TCR β 链。在一些实施方案中,选择多个TCR α 链和TCR β 链以产生约 1×10^6 个TCR $\alpha\beta$ 对。在一些实施方案中,选择多个TCR α 链和TCR β 链以产生多于 1×10^6 个TCR $\alpha\beta$ 对。

[0226] 在一些实施方案中,基于T细胞群内的频率从总库中选择TCR α -链和TCR β -链序列的一个或多个子集。在一些实施方案中,关于TCR序列的频率的数据被用于为TCR α -链和TCR β -链创建单独的秩次序。例如,对于含有T细胞的样本,可使用多重PCR、靶富集或5'-RACE和PCR测定编码(部分)不同TCR链氨基酸序列的核酸分子的绝对数量。在一些实施方案中,DNA用于本文所述的方法中。在一些实施方案中,RNA用于本文所述的方法中。得到的TCR链序列集合被分为TCR α -序列集合和TCR β -链序列集合。其中TCR片段在氨基酸序列水平上框外连接,和/或其中引入了终止密码子,和/或其中存在移码突变,和/或其中存在缺陷剪接位点的任何无效的TCR链序列,都被从集合中去除。在一些实施方案中,使用“唯一分子标识符”(UMI)基于唯一分子的计数确定编码(部分)特定TCR链氨基酸序列的核酸分子的绝对数量,并按降序排序以获得TCR α -链和TCR β -链的秩次序。在一些实施方案中,使用编码特定TCR链的核酸分子的绝对数量(或分别在总TCR α -链和TCR β -链中的相应百分比)以降序对每个集合进行分类,以获得TCR α -链和TCR β -链的秩次序。

[0227] 在本文提供的任何实施方案中,可以收集或产生RNA,并且可以对RNA进行测序来代替DNA测序。

[0228] 在一些实施方案中,TCR序列的频率数据被用于为TCR α -链和TCR β -链创建组合的秩次序。例如,对于含有T细胞的样本,可使用多重PCR、靶富集或5'-RACE和PCR来测定编码(部分)不同的TCR链氨基酸序列的核酸分子的绝对数量。在一些实施方案中,DNA用于本文所述的方法中。在一些实施方案中,RNA用于本文所述的方法中。其中TCR片段在氨基酸序列水平上框外连接,和/或其中引入终止密码子,和/或其中存在移码突变,和/或其中存在缺陷剪接位点的任何无效的TCR链序列,都被从集合中去除。使用编码特定的TCR链核酸分子的绝对数量或TCR链总集合中的相应百分比(可能通过使用唯一分子标识符”(UMI))按降序对剩余的TCR链序列进行排序,以获得TCR链的秩。

[0229] 在一些实施方案中,基于TCR库恢复的期望深度来限定频率阈值。例如,对于含有T细胞的样本,可使用多重PCR、靶富集或5'-RACE和PCR并可使用“唯一分子标识符”(UMI)来测定编码(部分)不同TCR链氨基酸序列的核酸分子的绝对数量。在一些实施方案中,DNA用于本文所述的方法中。在一些实施方案中,RNA用于本文所述的方法中。得到的TCR链序列集合被分为TCR α -链序列集合和TCR β -链序列集合。其中TCR片段在氨基酸序列水平上框外连接,和/或其中引入了终止密码子,和/或其中存在移码突变,和/或其中存在缺陷剪接位点的任何无效的TCR链序列,都被从集合中去除。使用编码特定TCR链的核酸分子的绝对数量(或分别在总TCR α -链或TCR β -链中的相应百分比)以降序对每个集合进行分类,以获得TCR α -链和TCR β -链的秩次序。如果意图恢复样本中的十个频率最高的TCR,则可能只选择前十个频率最高的TCR α -链和TCR β -链。相反,如果希望恢复样本中大部分的TCR库,可以选择超过10个频率最高的TCR α -链和TCR β -链。较低的频率阈值可以导致更大的深度,但也可以在

选定的TCR链池和由此产生的TCR文库中产生更高的多样性。重要的是,也可以从如在所公开的实施方案之一中所描述的组合的秩次序中选择TCR α -链或TCR β -链。

[0230] 在一些实施方案中,较低的频率阈值用于基于频率选择TCR α -链和TCR β -链的集合。在一些实施方案中,不需要选择相等数量的TCR α -链和TCR β -链。例如,对于含有T细胞的样本,可使用多重PCR、靶富集或5' -RACE和PCR并可能使用“唯一分子标识符”(UMI)来测定编码(部分)不同TCR链氨基酸序列的核酸分子的绝对数量。在一些实施方案中,DNA用于本文所述的方法中。在一些实施方案中,RNA用于本文所述的方法中。得到的TCR链序列集合被分为TCR α -链序列集合和TCR β -链序列集合。其中TCR片段在氨基酸序列水平上框外连接,和/或其中引入了终止密码子,和/或其中存在移码突变,和/或其中存在缺陷剪接位点的任何无效的TCR链序列,都被从集合中去除。使用编码特定TCR链的核酸分子的绝对数量(或分别在总TCR α -链或TCR β -链中的相应百分比)以降序对每个集合进行分类,以获得TCR α -链和TCR β -链的秩次序。得到的秩次序可能包含不同数目的TCR α -链或TCR β -链,从而阻止选择相等数量的TCR α -链和TCR β -链,或者可能期望从一个类别中选择比另一个类别更多的TCR链,例如因为细胞中两个TCR α 基因座都倾向于经历有效重排(productive rearrangement)。此外,如果选择高于或低于某一频率(以绝对数量或百分比表示)的所有TCR链,这可能导致分别选择不同数量的TCR α -链和TCR β -链。重要的是,也可以如在前述实施方案之一中描述的从组合的秩次序中选择TCR α -链或TCR β -链。

[0231] 在一些实施方案中,基于定量频率数据选择前100个最丰富(most abundant)的TCR α -链和TCR β -链。在一些实施方案中,基于定量频率数据选择多于前100个最丰富的TCR α -链和TCR β -链。在一些实施方案中,基于频率数据来选择前100个、前200个、前300个、前400个、前500个、前600个、前700个、前800个、前900个、前1000个,或它们之间的任何数量或范围内的最丰富的TCR α -链和TCR β -链。在一些实施方案中,基于频率数据选择多于前1000个最丰富的TCR α -链和TCR β -链。在一些实施方案中,基于频率数据选择前5%、前10%、前20%、前30%、前40%、前50%、前60%、前70%、前80%、前90%、前100%或它们之间的任何数量或范围内的TCR α -链和TCR β -链。在一些实施方案中,选定的链用作组装TCR $\alpha\beta$ 对的集合的构建块(building block)。

[0232] 在一些实施方案中,基于与第二T细胞群相比的相对富集从总库中选择TCR α -链和TCR β -链的序列的一个或多个子集。在一些实施方案中,当与另一样本相比时,基于给定样本中最高倍数富集(highest fold enrichment)来选择前100个TCR α -链和TCR β -链。在一些实施方案中,基于相对富集来选择前100个最丰富的TCR α -链和TCR β -链。在一些实施方案中,最高倍数富集是相对于另一个样本而言的。在一些实施方案中,基于相对富集选择多于前100个最丰富的TCR α -链和TCR β -链。在一些实施方案中,基于相对富集来选择前100个、前200个、前300个、前400个、前500个、前600个、前700个、前800个、前900个、前1000个或它们之间的任何数量或范围内的最丰富的TCR α -链和TCR β -链。在一些实施方案中,基于相对富集选择多于前1000个最丰富的TCR α -链和TCR β -链。在一些实施方案中,基于相对富集选择前5%、前10%、前20%、前30%、前40%、前50%、前60%、前70%、前80%、前90%、前100%或它们之间的任何数量或范围内的TCR α -链和TCR β -链。在一些实施方案中,来自肿瘤病灶的TCR链是基于与它们在血液中的相应频率相比的相对富集来选择的。在一些实施方案中,来自肿瘤病灶的TCR链是基于与第二肿瘤病灶相比的相对富集来选择的。在一些实施方案

中,对多个样本并行进行TCR链的定量。例如,可以并行分析来自同一个体的多个肿瘤病灶、匹配的肿瘤病灶(matched tumor lesion)和血液样本,或者较大肿瘤病灶的多个离散采样切片。通过分析来自同一个体的多个样本或匹配样本,可以确定TCR链的生物学相关性。例如,与在外周血中也以高频率出现或在单个肿瘤病灶中出现的TCR链相比,与血液相比在肿瘤中频率富集或在多个肿瘤病灶中出现的TCR链更可能与肿瘤抗原的识别相关。在一些实施方案中,基于与肿瘤边界(boundary)或肿瘤边缘(margin)(其可包括肿瘤周围的正常组织)相比肿瘤核心中的TCR链频率来选择TCR链。在任何描述的实施方案中,相对频率差可以用于基于相对频率的倍数差来创建秩次序。相对频率的倍数差可以是 10^{-6} 和 10^6 之间的任何数字。在一些实施方案中,可以优先选择或排除仅在至少两个比较样本之一中发现的TCR链用于生成TCR文库。在一些实施方案中,排序前100的TCR α -链和TCR β -链用于生成TCR库。在一些实施方案中,多于排序前100的TCR α -链和TCR β -链用于生成TCR库。在一些实施方案中,比较不同样本中的TCR链库,以靶向选择具有更高新生抗原特异性的可能性的TCR链。在一些实施方案中,基于相对富集对TCR链序列进行排序,随后根据基于频率的秩次序进行选择,例如,如上所述。任何排序的标准和任何标准的组合均可用于TCR链选择。

[0233] 在本文提供的任何实施方案中,也可以使用复合度量(composite metrics)。也就是说,可以通过两个或多个方面的组合来进行排序,例如也可以通过频率和肿瘤富集来进行排序。在一些实施方案中,TCR在肿瘤中既是高频率的又是在肿瘤中富集的。

[0234] 在一些实施方案中,基于TCR链的生物学特性或序列特征,从总库中选择TCR α -链和TCR β -链序列的一个或多个子集。在一些实施方案中,TCR链的生物学特性或序列特征选自以下的至少一种:(预测的)抗原特异性、(预测的)HLA-限制性、亲和性、共受体依赖性 or 亲代T细胞谱系(例如CD4或CD8 T细胞)。在一些实施方案中,通过基于计算机模拟算法(in silico algorithm-based)的预测获得关于生物特性的信息。

[0235] 在一些实施方案中,使用算法来识别样本中的TCR聚类簇(cluster)。例如,可以使用关于TCR聚类簇的信息进行TCR链的靶向选择,以便随后生成TCR文库。在一些实施方案中,使用关于TCR聚类簇的信息选择具有限定的属性的聚类簇。在一些实施方案中,使用关于TCR聚类簇的信息将聚类簇与公共TCR数据库进行比较。在一些实施方案中,使用关于TCR聚类簇的信息去除具有高不相关的概率的TCR特异性的聚类簇或TCR链。在一些实施方案中,基于聚类簇与公共TCR数据库的比较来去除聚类簇。高不相关的概率的示例性TCR特异性包括,例如,识别源自流感、CMV、EBV和其他病毒和细菌感染原(bacterial infectious agent)的病毒表位。生成的已从中去除不相关的TCR链的TCR链集可随后用于TCR文库生成。在一些实施方案中,关于TCR聚类簇的信息用于优先包括具有相关氨基酸序列的TCR链的聚类簇。

[0236] 在一些实施方案中,基于TCR链的氨基酸序列识别TCR的特性。在一些实施方案中,基于TCR $\alpha\beta$ 复合体(complex)的结构特征识别TCR的特性。示例性特性包括,例如,HLA-限制性、抗原特异性、共受体依赖性、亲代T细胞谱系、TCR聚类簇间的共有特性等。

[0237] 在一些实施方案中,基于空间信息从总库中选择TCR α -链和TCR β -链序列的一个或多个子集。在一些实施方案中,基于基因或蛋白质表达的空间模式从总库中选择TCR α -链和TCR β -链序列的一个或多个子集,其中空间基因或蛋白质表达模式衍生自以下至少一种:组织中的起始区域或其他基因或蛋白质的共表达模式(为了避免疑问,包括可能不存在共表

达)。在一些实施方案中,基于肿瘤内定位选择TCR链。在一些实施方案中,基于在显示某些表型标志物的过表达(或不表达)的空间中的富集来选择TCR链。表型标志物可以是与表型相关的任何标志物,包括但不限于一种或多种表面标志物或其片段、一种或多种蛋白质或其片段、一种或多种RNA如microRNA、siRNA或任何其他RNA。

[0238] 在一些实施方案中,使用空间测序方法过滤TCR链。空间测序能够从细胞中恢复转录组或遗传信息,包括TCR序列信息以及细胞在组织中的位置。例如,回收临近细胞组并标记,以标记它们在组织内的起点(spot of origin),从而将转录组或遗传信息与空间维度联系起来。从组织中相同或邻近空间位置恢复的细胞形成空间聚类簇。基于某些信息,可以优先选择来自某些空间聚类簇的TCR链。可以使用的信息包括例如解剖信息。例如,位于肿瘤中心或三级淋巴结构中的细胞簇可能比位于肿瘤边界处的簇更受关注。可使用的信息包括,例如,转录组和/或蛋白表达信息。例如,PD-1和CD39高表达的空间聚类簇比此类标志物低表达的聚类簇更可能富集新生抗原特异性TCR链。因此,可以选择PD-1和CD39高表达的聚类簇来过滤TCR链。在一些实施方案中,例如,基于与肿瘤边界相比在肿瘤中心的过表达来选择聚类簇。表1所示为示例性选择参数。

[0239] 表1. 标志物选择的示例性参数(包括其任何组合)

	基因标志物	空间聚类中表达的相关性
[0240]	PD-1, Tim-3, LAG-3, CD109, CD200, GITR, TNFRSF18, CD39, CD103, IFN- γ , IL-2, TNF- α , Ki67, Granzyme B, CXCL13 CD8, CD4, CTLA-4, FoxP3, RBPJ, ZBED2, ETV1, ID3, MAF,	过表达可以标记可能的目的聚类簇
	PRDM1, EOMES, ITM2A, KIR2DL4, PTMS, FAM3C, AKAP5, CD7, PHLDA1, ENTPD1, CCL3, CCL4, TNFRSF9, ZBED2, XCL2, HMGB1P17, ILDR1, C10orf113, BTNL8, DLGAP5, MIR4298, CDC25C, KIF20A, E2F2, TSHZ2	
[0241]	CD45RA, KLRG1, IL7R, TCF7, CCR7	过表达可以标记关注较低的聚类簇

[0242] 可以使用选定的TCR链的组生成TCR库。在一些实施方案中,使用空间分辨

(spatially resolved)的RNA-或DNA-测序。在一些实施方案中,例如,与T细胞表型相关的另外的转录物一起批量(bulk)恢复TCR链群。与T细胞表型相关的示例性转录物包括CTLA-4、PD-1、CD103、CD39、FoxP3、IFN- γ 、IL-2、CXCL13等。在一些实施方案中,使用从多个空间聚类簇恢复的解剖位置和T细胞特异性转录组来识别关注的聚类簇(cluster of interest)。

[0243] 在一些实施方案中,基于在多个样本中共现或以相似频率出现来从总库中选择TCR α -链和TCR β -链序列的一个或多个子集。在一些实施方案中,基于在多个肿瘤病灶中的出现(occurrence)来从总库中选择TCR α -链和TCR β -链序列的一个或多个子集。在一些实施方案中,使用来自多个肿瘤病灶的TCR链频率信息来过滤关注的TCR链。在一些实施方案中,用于过滤关注的TCR链的特定信息包括排除仅发生在一个肿瘤病灶中的TCR链。在一些实施方案中,用于过滤关注的TCR链的特定信息包括选择性包含在所有测试的肿瘤病灶中出现的TCR链。

[0244] 在一些实施方案中,基于选择到多个组中来从总库中选择TCR α -链和TCR β -链序列的一个或多个子集,以分别恢复TCR库的特定部分。在一些实施方案中,将频率高于限定的阈值的所有TCR α -链和TCR β -链序列一起选入一组。例如,可以将包括总TCR α -链和TCR β -链序列的某个百分比的,例如高于和/或等于10%、9%、8%、7%、6%、5%、4%、3%、2%、1%、0.5%、0.1%、0.05%或0.01%或该范围之间的任何其他数值的所有TCR α -链和TCR β -链序列,选入一组。通过类似的原理,例如,可以将排序位置高达并等于(up and equal to)10(前10)、100(前100)、1000(前1000)或10000(前10000)的所有TCR α -链和TCR β -链序列选入一组。在一些实施方案中,将低于限定的阈值或在两个限定的阈值之间的所有TCR α -链和TCR β -链序列选入一组。例如,可以将低于1%或在1%和0.1%之间的所有TCR α -链和TCR β -链序列选入一组。通过类似的原理,例如,可以将排序位置(rank position)低于10或排序位置在10和100之间的所有TCR α -链和TCR β -链序列选入一组。在一些实施方案中,将频率高于限定的阈值的所有TCR α -链和TCR β -链序列选入一组,并将低于限定的阈值的所有TCR α -链和TCR β -链序列选入另一组。例如,将高于和等于1%的所有TCR α -链和TCR β -链序列选入一组,将低于1%的所有TCR α -链和TCR β -链序列选入另一组。通过类似的原理,例如,将排序位置高于或等于10的所有TCR α -链和TCR β -链序列选入一组,并将排序位置低于10的所有TCR α -链和TCR β -链序列选入另一组。

[0245] 在一些实施方案中,筛选了大量的TCR,而没有产生相当复杂的TCR文库(TCR library)。例如,可以生成多个子文库(sub-library)。在一些实施方案中,TCR文库的复杂度小于由包括的所有TCR α -链和TCR β -链序列的随机配对产生的复杂度。在一些实施方案中,所生成的TCR文库不包含所有可能的TCR α 和TCR β 组合。在一些实施方案中,将TCR链的集合分离到单独的池中以创建一个或多个复杂度比通过随机配对TCR α -链和TCR β -链序列所获得的复杂度更低的文库。在一些实施方案中,基于排序阈值(ranking threshold)来池化(pool)TCR链。在一些实施方案中,秩次序(rank order)中某个位置内的所有TCR形成池用于生成TCR文库。例如,排序前50的TCR α -链和TCR β -链序列可以包含在池1中;排序前25至前75的TCR α -链和TCR β -链序列可包含在池2中;等等。作为另一个实例,排序前50至前100的TCR α -链和TCR β -链序列可以包含在池2中;等等。可以使用任何排序标准,包括针对TCR α -链和TCR β -链序列的不同阈值。在一些实施方案中,排序基于频率。在一些实施方案中,排序基

于与参考样本相比的相对富集度。在一些实施方案中,基于空间信息来池化(pool) TCR链。例如,来自给定空间聚类簇的所有TCR α -链和TCR β -链可以形成特定的池。在一些实施方案中,基于TCR的特性来池化(pool) TCR链。任何TCR特性均可用于池化(pool) TCR。例如,具有定义的序列特征或预测属性的所有TCR可以形成一个特定的池。预测属性的实例包括,例如,共受体依赖性、起源T细胞谱系(originating T cell lineage)、HLA-限制性、特异性等。

[0246] 在一些实施方案中,基于如在不同实施方案中定义多个标准的组合,从总库中选择TCR α -链和TCR β -链序列的一个或多个子集。

[0247] 在一些实施方案中,通过所选TCR α -链和TCR β -链序列的组合配对来创建TCR库(TCR repertoire)。在一些实施方案中,组合配对包括所有选择的TCR α -链和TCR β -链序列的随机配对。可以通过选定的TCR α -链和TCR β -链序列的组合配对来创建TCR $\alpha\beta$ 链的文库。在一些实施方案中,使用选定的TCR链序列来合成TCR α -链和TCR β -链DNA或RNA片段的文库。使用本领域技术人员已知的克隆策略(例如,包括但不限于Gibson分子组装和Golden Gate组装),可以通过精确连接一个TCR α -链和一个TCR β -链DNA或RNA片段来产生人工TCR基因。在一些实施方案中,通过直接合成其中恰好连接了一个TCR α -链和一个TCR β -链的DNA或RNA片段来产生TCR α -链和TCR β -链的组合。在一些实施方案中,通过用单独的TCR α -基因和TCR β -基因的集合来修饰细胞池,使得细胞将表达约一个TCR α -链和一个TCR β -链,而在细胞内产生TCR α -链和TCR β -链的组合。

[0248] 在一些实施方案中,通过所选的TCR α -链和TCR β -链的序列的组合配对来创建TCR库,从而创建TCR $\alpha\beta$ 对的文库是通过以下的至少一种来实现的:a) 使用TCR链序列合成TCR α -链和TCR β -链的DNA或RNA片段的单独文库,所述TCR α -链和TCR β -链的DNA或RNA片段随后被连接成一个DNA或RNA片段,其中恰好连接了一个TCR α -链和一个TCR β -链,b) 通过直接合成其中恰好连接了一个TCR α -链和一个TCR β -链的DNA或RNA片段来产生TCR α -链和TCR β -链的组合,c) 通过用单独的TCR α -和TCR β -基因的集合来修饰细胞池,这样细胞将表达约一个TCR α -链和一个TCR β -链,而在细胞内产生TCR α -链和TCR β -链的组合,和/或d) TCR α -链和TCR β -链的组合被连接在单链TCR构建体中,其中TCR α 和TCR β 可变链片段两者被融合,并且其中所述单链TCR构建体可以被融合到(i) 单独的跨膜结构域或(ii) 另外含有细胞内信号传导结构域,包括但不限于单独的或与CD28信号传导结构域组合的CD3 ϵ 或CD3 ζ 信号传导结构域。

[0249] 对于本文的任何实施方案,在一些实施方案中,恢复I类和/或II类限制性TCR序列。

[0250] 对于本文提供的任何实施方案,在一些实施方案中,恢复以下至少一项:新生抗原特异性TCR序列、病毒特异性TCR序列、共有肿瘤抗原特异性TCR序列和/或自身抗原特异性TCR序列。

[0251] 在一些实施方案中,活化标志物可以选自:CD25、CD69、CD62L、CD137、IFN- γ 、IL-2、TNF- α 、GM-CSF、OX40。

[0252] 在一些实施方案中,TCR库代表样本中的所有TCR- α 链和TCR- β 链序列。在一些实施方案中,TCR库代表从样本中恢复的所有TCR- α 链和TCR- β 链序列。在一些实施方案中,从样本中存在的TCR序列的总库中选择TCR库作为TCR- α 链和TCR- β 链序列的子集。在一些实施方案中,从从样本中恢复的TCR序列的总库中选择TCR库作为TCR- α 链和TCR- β 链序列的子集。

[0253] 在一些实施方案中,该方法包括从所创建的TCR库中识别至少一个TCR $\alpha\beta$ 对。在一些实施方案中,TCR $\alpha\beta$ 对表示新生成的组合。在一些实施方案中,TCR $\alpha\beta$ 对表示不是新生成的组合。在一些实施方案中,用生成的TCR $\alpha\beta$ 对的文库修饰报告细胞池或T细胞池。在一些实施方案中,修饰后的报告细胞池或修饰后T细胞池可以被装载 (loaded) 有至少一种目的抗原的抗原呈递细胞刺激。可使用任何针对报告细胞或T细胞的刺激试验 (stimulation assay)。对报告细胞或T细胞的刺激试验是本领域技术人员已知的。在一些实施方案中,基于至少一种活化标志物来分离抗原反应性报告细胞或T细胞。例如,可以使用任何CD4或CD8 T细胞活化标志物。在一些实施方案中,可以使用任何CD标志物。例如,在一些实施方案中,活化标志物可以包括标志物如CD69、CD137、IFN- γ 、IL-2、TNF- α 、GM-CSF。在一些实施方案中,基于增殖来分离抗原反应性报告细胞或T细胞。在一些实施方案中,基于通过报告细胞或T细胞活化依赖性表达抗性基因而获得的对抗生素的抗性选择来分离抗原反应性报告细胞或T细胞。可以使用任何报告细胞或T细胞分离的方法,包括但不限于例如本领域技术人员已知的磁珠富集或流式细胞术。在一些实施方案中,可能不需要分离报告细胞或T细胞。例如,报告细胞或T细胞增殖或使用抗生素选择可消除选择的需要。在一些实施方案中,从大量抗原反应性报告细胞或T细胞中获得RNA。从大量抗原反应性报告细胞或T细胞中获得的RNA可用于产生TCR $\alpha\beta$ 特异性cDNA。在一些实施方案中,通过DNA测序分析TCR $\alpha\beta$ 特异性cDNA以确定抗原反应性报告细胞或T细胞的TCR $\alpha\beta$ 基因序列。在一些实施方案中,从大量 (bulk) 抗原反应性报告细胞或T细胞中获得DNA以产生TCR α/β 特异性PCR产物,通过DNA测序分析该产物以确定抗原反应性报告细胞或T细胞的TCR $\alpha\beta$ 基因序列。在一些实施方案中,定义的TCR $\alpha\beta$ 对可以与基于分子核酸的标识符 (“条形码”) 相关联,该标识符可以通过对从RNA或DNA产生的特异性PCR产物进行测序来检测。

[0254] 在一些实施方案中,使用基于单细胞的方法来确定TCR $\alpha\beta$ 对。例如,基于单细胞的方法包括液滴PCR。使用基于单细胞的方法,可以分析抗原反应性T细胞的TCR $\alpha\beta$ 基因序列。在一些实施方案中,通过一种或多种活化标志物来识别抗原反应性T细胞。可以使用任何CD标志物,例如包括CD4或CD8 T细胞活化标志物。在一些实施方案中,活化标志物包括例如CD69、CD137、IFN- γ 、IL-2、TNF- α 、GM-CSF。在一些实施方案中,抗原反应性T细胞由其转录谱来识别。

[0255] 在一些实施方案中,通过大量T细胞中TCR $\alpha\beta$ 基因插入物 (insertion) 的基因组PCR来测定TCR $\alpha\beta$ 对。在一些实施方案中,对产生的PCR产物进行DNA测序分析。

[0256] 在一些实施方案中,使用报道基因来识别TCR转导的报告细胞或T细胞的活化。报道基因可以报告TCR触发。示例性的报道基因包括例如NFAT-GFP或NFAT-YFP。在一些实施方案中,基于通过抗性基因的T细胞活化依赖性表达而获得的对抗生素的抗性的选择来分离抗原反应性报告细胞或T细胞。示例性的报道基因包括例如NFAT-嘌呤霉素抗性或NFAT-潮霉素抗性。在一些实施方案中,使用报道基因的组合。

[0257] 在一些实施方案中,通过与携带目的抗原的MHC复合物结合来识别抗原反应性细胞。

[0258] 在上述任一实施方案中,从所创建的TCR库中识别至少一个TCR $\alpha\beta$ 对。TCR $\alpha\beta$ 对的期望特征可以包括抗原特异性、TCR亲和力、TCR共受体依赖性、HLA-限制性、TCR交叉反应性、TCR抗肿瘤反应性或它们的任意组合。

[0259] 在上述任一实施方案中,可将恢复的TCR链序列定义为包含CDR3核苷酸序列以及足够的5'-和3'-核苷酸序列信息,以基于核苷酸序列比对选择至少一个TCR V-片段家族和至少一个TCR J-片段家族,从而组装完整的TCR链序列。在一些实施方案中,以2位(digit)或4位(digit)分辨率识别J基因。在一些实施方案中,基于65%序列同一性、70%序列同一性、75%序列同一性、80%序列同一性、85%序列同一性、90%序列同一性、95%序列同一性、96%序列同一性、97%序列同一性、98%序列同一性、99%序列同一性、100%序列同一性,以及其间的任何数值或范围进行核苷酸序列比对。在一些实施方案中,获得足够的序列信息以从所创建的具有期望特征的TCR库中识别TCR- α 链和TCR- β 链。

[0260] 在一些实施方案中,由CDR3核苷酸序列定义恢复的TCR链。在一些实施方案中,由CDR3氨基酸序列定义恢复的TCR链。

[0261] 在一些实施方案中,由足够的5'-和3'-核苷酸序列信息定义恢复的TCR链,以选择至少一个TCR V-片段家族和至少一个TCR J-片段家族。在一些实施方案中,由足够的氨基酸序列信息定义恢复的TCR链,以选择至少一个TCR V-片段家族和至少一个TCR J-片段家族。在一些实施方案中,恢复的TCR链被定义为足够的核苷酸或氨基酸序列信息以明确地识别所产生的TCR文库中的TCR $\alpha\beta$ 对。在一些实施方案中,恢复的TCR链被定义为明确地识别所创建的TCR文库中的TCR $\alpha\beta$ 对的唯一分子标识符,例如基于核苷酸的条形码。

[0262] 在一些实施方案中,基于核苷酸序列比对来定义TCR链序列。在一些实施方案中,基于氨基酸序列比对来定义TCR链序列。通过核苷酸或氨基酸序列比对,可以组装出完整的TCR链序列。

[0263] 在一些实施方案中,可以使用来自个体的包含如上所述的非活起始材料的样本。举例来说,非活起始材料可以包括非活细胞或非活组织样本。可以通过本领域已知的任何方法来保存非活起始材料。

[0264] 在上述任一实施方案中,可以恢复所识别的TCR库的限定部分。在一些实施方案中,所识别的TCR库的限定部分包括恢复TCR库的选定部分而不是完整的TCR库。可通过上文所述的任何标准来限定所选的TCR序列,例如T细胞群内的限定的频率、与第二T细胞群相比的相对富集、TCR链的生物学特性、基因表达的空间模式、在多个样本中以相似频率出现或共现、选择进入TCR链的多个组或池,或它们的任意组合。在一些实施方案中,由给定的抗原特异性限定选定的TCR库。在一些实施方案中,抗原特异性包括对新生抗原的特异性。在一些实施方案中,抗原特异性包括预测的抗原特异性。

[0265] 在一些实施方案中,恢复抗原特异性TCR序列。在一些实施方案中,恢复新生抗原特异性TCR序列。例如,新生抗原可以是在肿瘤中发现的被抗原特异性T细胞识别的突变蛋白。因此,可以存在针对肿瘤的抗原特异性T细胞,其TCR序列对肿瘤或其肿瘤抗原具有特异性。

[0266] 在前述任一实施方案中,表达抗原特异性TCR序列的T细胞可用于诊断或治疗感染或自身免疫疾病。

[0267] 在前述任一实施方案中,可以施用表达新生抗原特异性TCR序列的T细胞作为癌症疗法。例如,可以使用新生抗原特异性T细胞靶向表达新生抗原的肿瘤。在一些实施方案中,通过将新生抗原特异性TCR链引入T细胞来产生新生抗原特异性T细胞。在一些实施方案中,表达新生抗原特异性TCR序列的T细胞可以是自体的或同种异体的。

[0268] 在前述任一实施方案中,该方法可用于诊断。例如,针对某一组织抗原的抗原特异性TCR的存在可能指示自身免疫性疾病。作为其他实例,针对某些病原体的抗原特异性TCR的存在可以指示感染性疾病。

[0269] 在一些实施方案中,诊断是从感染的病理部位恢复TCR库。在一些实施方案中,诊断是从自身免疫部位恢复TCR库。例如,感染或自身免疫位点的细胞或组织可能表达某些T细胞识别的特定抗原。通过测定能在感染或自身免疫位点检测特定抗原的T细胞的TCR序列,可恢复与感染或自身免疫位点相关或对其特异的TCR库。为了识别针对自身免疫介导的自身抗原(self-antigen)或病原体的TCR序列,可以测试从选定的TCR α -链和TCR β -链产生的组合TCR $\alpha\beta$ 文库针对选定的自身抗原或病原体衍生抗原组的反应性。

[0270] 在前述任一实施方案中,该方法可用于恢复BCR/抗体库。例如,可以恢复表达BCR受体或产生特定抗原特异性抗体的B细胞。因此,可通过应用上述任何方法来测定恢复的B细胞的BCR/抗体库,以对免疫球蛋白重链和轻链进行恢复、选择和组合配对,从而产生抗体库。可以从所创建的抗体库中选择具有目的特性的抗体。

[0271] 在上述任一实施方案中,该方法可以包括从患者样本中分离包含TCR- α 和TCR- β 核酸序列的核酸。核酸可以是DNA或RNA。可以从个体的任何组织或细胞中分离核酸,包括但不限于血液、皮肤、肝脏、骨髓、活检材料等。在一些实施方案中,个体是人。在一些实施方案中,个体是哺乳动物。在一些实施方案中,个体是动物。

[0272] 在本文的任一实施方案中,来自个体的样本可以包含从体液中分离的细胞。在一些实施方案中,所述细胞是肿瘤特异性T细胞或肿瘤浸润的淋巴细胞。在一些实施方案中,体液选自血液、尿液、血清、浆膜液、血浆、淋巴液、脑脊液、唾液、痰液、粘膜分泌物、阴道液、腹水、胸膜液、心包液、腹膜液和腹腔液。

[0273] 在一些实施方案中,基于至少一个标准从总库中选择TCR α -链和TCR β -链的序列的一个或多个子集:基于在T细胞群内的频率、基于与第二T细胞群相比的相对富集、基于给定TCR链的DNA拷贝数和RNA拷贝数的相对差异、基于TCR链的生物学特性,其中所述特性选自下述至少一项:(预测的)抗原特异性、(预测的)HLA-限制性、亲和性、共受体依赖性,亲代T细胞谱系(例如CD4或CD8 T细胞)或TCR序列基序,基于基因表达的空间模式,其中空间基因表达模式衍生自组织中的起始区域或其他基因的共表达模式中的至少一种、基于在多个样本中共现以相似频率出现,例如在多个肿瘤病灶中出现、基于分配到多个组以分别恢复TCR库的特定部分、基于在不同实施方案中限定的多个标准的组合。

[0274] 在一些实施方案中,测定TCR- α 和TCR- β 的序列是通过下列至少一种方法实现的:多重PCR;通过靶富集进行TCR序列恢复;通过5' RACE和PCR进行TCR序列恢复;通过空间测序进行TCR序列恢复;通过RNA-seq进行TCR序列恢复;以及唯一分子标识符(UMI)的使用。

[0275] 在一些实施方案中,步骤III是通过以下至少一项来实现的:使用TCR链序列来合成TCR α -链和TCR β -链的DNA或RNA片段的文库,所述TCR α -链和TCR β -链的DNA或RNA片段被连接成一个DNA或RNA片段(任选地,其中恰好连接了一个TCR α -链和一个TCR β -链),通过直接合成其中恰好连接了一个TCR α -链和一个TCR β -链的DNA或RNA片段来产生TCR α -链和TCR β -链的组合,或通过TCR α -和TCR β -基因的单独集合来修饰细胞池,使得细胞将表达一个TCR α -链和一个TCR β -链,而在细胞内产生TCR α -链和TCR β -链的组合,在单链TCR构建体中使TCR α -链和TCR β -链的组合连接,该构建体含有TCR链片段以及单独的或与CD28信号传导结构域组

合的CD3 ζ 或CD3 ϵ 信号传导结构域。

[0276] 在一些实施方案中,步骤IV通过以下至少一项来实现:用生成的TCR $\alpha\beta$ 对的文库修饰的报告T细胞池被呈递至少一种目的抗原的抗原呈递细胞刺激,并且基于至少一种活化标志物分离抗原反应性报告细胞以用于TCR分离;用适于追踪细胞增殖的荧光染料标记用生成的TCR $\alpha\beta$ 对的文库修饰的报告细胞池,所述报告细胞池由表达至少一种目的抗原的抗原呈递细胞刺激,并且基于增殖来分离抗原反应性报告细胞以用于TCR分离;将用生成的TCR $\alpha\beta$ 对的文库修饰的报告细胞池分成至少两个样本;用表达至少一种目的或非目的抗原的抗原呈递细胞刺激样本;刺激后,将两个报告细胞群孵育一段时间,随后通过TCR分离分析两个报告细胞群;从两个样本中获得的TCR $\alpha\beta$ 对的比较将识别在暴露于至少一种抗原的样本中具有更高丰度的TCR基因;用生成的TCR $\alpha\beta$ 对的文库修饰的报告细胞池被呈递至少一种目的抗原的抗原呈递细胞刺激,并且基于例如报告TCR触发的NFAT-GFP或NFAT-YFP的至少一种报道基因,分离抗原反应性报告细胞以用于TCR分离;用生成的TCR $\alpha\beta$ 对的文库修饰的报告细胞池被呈递至少一种目的抗原的抗原呈递细胞刺激,并且基于抗原特异性报告细胞的选择分离抗原反应性报告细胞以用于TCR分离,所述抗原特异性报告细胞的选择基于在TCR信号传导时获得的选择性存活,包括但不限于获得的抗生素抗性,例如通过使用NFAT-嘌呤霉素转基因;将用所生成的TCR $\alpha\beta$ 对的文库修饰的报告细胞池暴露于携带目的抗原的一种或多种MHC复合物;将与MHC复合物结合的报告细胞分离以用于TCR分离;用生成的TCR $\alpha\beta$ 对的文库修饰的报告细胞池被表达至少一种目的抗原的抗原呈递细胞刺激;随后,使用基于单细胞的液滴PCR或微流体方法识别目的TCR $\alpha\beta$ 对,以将TCR分离与至少一种活化标志物的转录物水平的检测相结合;因此,检测到T细胞池内的单个报告细胞,其中TCR $\alpha\beta$ 转录物与增加水平的活化标志物共表达。

[0277] 在一些实施方案中,活化标志物选自:CD4或CD8 T细胞活化标志物、CD69、CD137、IFN- γ 、IL-2、TNF- α 、GM-CSF、OX40。

[0278] 在一些实施方案中,活化标志物为CD69,分离出两个细胞群用于进一步分析,一个细胞群具有高表达的CD69,而另一个细胞群具有低表达的CD69。

[0279] 在一些实施方案中,步骤IV通过以下至少一项来实现:基于至少一种活化标志物的识别或选择;基于响应抗原的增殖的识别或选择;基于在抗原刺激的细胞中与未刺激的细胞相比对具有更高丰度的TCR基因的识别的识别或选择;基于通过TCR触发的报道基因激活的识别或选择;基于选择性存活的识别或选择,包括但不限于在TCR信号传导时获得的抗生素抗性;基于与一种或多种MHC复合物结合的识别或选择;使用基于单细胞的液滴PCR或微流体的识别或选择;或它们的任意组合。

[0280] 在一些实施方案中,报告细胞是T细胞。

[0281] 在一些实施方案中,使用基于单细胞的液滴PCR或微流体进行识别或选择;或其任何组合还包括确定活化相关基因的共表达。

[0282] 在一些实施方案中,本文描述了创建多个T细胞文库的方法,所述方法包括:(a)根据本文所述方法恢复T细胞受体(TCR)库;(b)从总库中选择TCR α -链和TCR β -链的序列到多个组中,以分别恢复TCR库的特定部分,其中创建复杂度较小的或恢复TCR库的特定部分的多个T细胞文库。

[0283] 在一些实施方案中,基于频率范围选择TCR α -链和TCR β -链的序列。

[0284] 在一些实施方案中,基于门控(gating)对细胞进行选择或分类(sorted)。在一些实施方案中,基于样本中最高的0.1%、0.5%、1%、5%、10%、20%、30%、40%、50%、60%、70%、80%、90%、99.9%或介于它们之间的任何数量或范围内的活的单细胞来对细胞进行分类。在一些实施方案中,基于样本中最低的0.1%、0.5%、1%、5%、10%、20%、30%、40%、50%、60%、70%、80%、90%、99.9%或介于它们之间的任何数量或范围内的活的单细胞来对细胞进行分类。

[0285] 在一些实施方案中,将核酸文库导入细胞群中。在一些实施方案中,细胞群是二倍体。在一些实施方案中,细胞群具有任何倍性。在一些实施方案中,第一细胞群选自文库细胞群。在一些实施方案中,测量第一细胞群中核酸序列的富集和/或耗竭来识别目的核酸序列。在一些实施方案中,通过将群与参考物(reference)进行比较来测量第一群中核酸序列的富集和/或耗竭。在一些实施方案中,所述参考物可以是第二细胞群或核酸序列文库。在一些实施方案中,通过与多于一个的参考物进行比较来测量第一细胞群中核酸序列的富集和/或耗竭。

[0286] 在一些实施方案中,基于流式细胞术分选来分离第一群和/或第二群。在一些实施方案中,基于检测表型的变化进行流式细胞术分选。在一些实施方案中,通过使文库细胞群与另一细胞群接触来诱导表型的变化。

[0287] 在一些实施方案中,通过荧光标记的探针与细胞的结合来检测表型的改变。在一些实施方案中,基于阈值进行流式细胞术分选。在一些实施方案中,阈值基于通过流式细胞术分选从总细胞的部分中恢复具有最高荧光信号的细胞的百分比的意图。在一些实施方案中,基于荧光信号分离前0.1%、0.5%、1%、5%、10%、20%、30%、40%、50%、60%、70%、80%、90%、99.9%的细胞。在一些实施方案中,阈值基于通过流式细胞术分选从总细胞的部分中恢复具有低荧光信号的细胞的百分比的意图。在一些实施方案中,基于荧光信号分离后(bottom)0.1%、0.5%、1%、5%、10%、20%、30%、40%、50%、60%、70%、80%、90%、99.9%的细胞。在一些实施方案中,使用多个阈值将样本分成第一群体和第二群体。

[0288] 在一些实施方案中,阈值基于通过基于荧光信号强度的流式细胞术从总细胞池中恢复最小数量的细胞的意图,例如,如果要从 10×10^6 个细胞中恢复至少 1×10^6 个细胞,则基于荧光信号分离前10%或后10%的细胞。

[0289] 在一些实施方案中,阈值基于第二细胞群的荧光信号强度。在一些实施方案中,阈值基于高于参考群中的荧光信号强度。在一些实施方案中,阈值基于低于参考群中的荧光信号强度。在一些实施方案中,参考群中的荧光信号强度基于具有二级标志物(secondary marker)表达的群的亚群。

[0290] 在一些实施方案中,基于珠富集分离第一群和/或第二群。在一些实施方案中,基于表型的改变进行磁珠富集。在一些实施方案中,通过使文库细胞群与另一细胞群接触来诱导表型的变化。

[0291] 在一些实施方案中,使文库细胞群与另一(another)细胞群(其可称为“其它(the other)细胞群”)接触是使其与经基因工程改造而改变表型的另一细胞群接触。在一些实施方案中,所述另一细胞群是经基因工程改造的细胞的多克隆池。在一些实施方案中,对所述另一细胞群进行基因工程改造以表达变体分子。在一些实施方案中,对所述另一细胞群进行基因工程改造以表达一种或多种抗原。在一些实施方案中,对所述另一细胞群进行基因

工程改造以表达小基因形式的一种或多种抗原。在一些实施方案中,对所述另一细胞群进行基因工程改造以表达串联小基因形式的一种或多种抗原。

[0292] 在一些实施方案中,所述另一细胞群是能够呈递抗原的细胞(抗原呈递细胞(Antigen-Presenting Cell,APC))。在一些实施方案中,所述另一细胞群是树突细胞。在一些实施方案中,所述另一细胞群是单核细胞。在一些实施方案中,所述另一细胞群是用MHC I类和/或II类等位基因(allele)工程改造的细胞。在一些实施方案中,所述另一细胞群可以是B细胞。在一些实施方案中,所述另一细胞群可以是自体细胞。在一些实施方案中,所述另一细胞群可以是自体B细胞。在一些实施方案中,所述另一细胞群可以是永生化的自体B细胞。在一些实施方案中,所述另一细胞群可以通过EBV感染永生化的自体B细胞。

[0293] 在一些实施方案中,使文库细胞群与所述另一细胞群接触引发在属于二者中任一细胞群的细胞上表达的因子之间的特异性相互作用。在一些实施方案中,因子之间的相互作用是受体-配体相互作用(receptor-ligand interaction)。在一些实施方案中,受体-配体相互作用中的受体是T细胞受体(TCR)。在一些实施方案中,受体-配体相互作用中的配体是存在于主要组织相容性复合体(major histocompatibility complex,MHC)上的抗原。在一些实施方案中,文库细胞群和所述另一细胞群之间的受体-配体相互作用引发可检测的表型变化。

[0294] 在一些实施方案中,在文库细胞群中表达的变体的集合是质粒文库,每个质粒表达单个TCR α 链和单个TCR β 链的组合。在一些实施方案中,通过组合连接TCR α 链和TCR β 链的集合来构建TCR文库。在一些实施方案中,TCR文库中可以存在TCR α 和TCR β 的所有组合。在一些实施方案中,使用多个复杂度较低的文库来创建复杂度较高的文库。在一些实施方案中,组合后的复杂度较高的文库的复杂度小于具有存在于所述组合后的复杂度较高的文库中的TCR α 和TCR β 的所有组合的文库。在一些实施方案中,配对信息或配对可能性(likelihood)的信息用于设计复杂度较低的多个文库,以使这些TCR α -TCR β 对出现在TCR文库中的机会最大化。在一些实施方案中,较低复杂度的文库可以包含一个或多个TCR α 链和一个或多个TCR β 链的所有组合。在一些实施方案中,通过首先产生编码TCR α 链和TCR β 链的单核苷酸分子的多种变体,然后混合编码TCR α 链和TCR β 链的分子的两种以上不同变体来构建TCR文库。

[0295] 在一些实施方案中,通过探针与细胞的结合来检测表型的变化。在一些实施方案中,在结合探针之前不需要用任何固定剂处理细胞。在一些实施方案中,探针允许将磁珠耦合到细胞。在一些实施方案中,磁珠富集允许从文库细胞群中分离第一和/或第二细胞群。在一些实施方案中,第一细胞群或第二细胞群是由磁体保留的细胞。在一些实施方案中,第一细胞群或第二细胞群是未被磁体保留的细胞。在一些实施方案中,多种探针的结合用于通过顺序磁珠富集(sequential magnetic bead enrichment)来分离第一群和/或第二群。在一些实施方案中,探针与CD62L结合。在一些实施方案中,探针与CD69结合。在一些实施方案中,至少一个核苷酸序列在第一细胞群中在统计学上显著的富集或耗竭。在一些实施方案中,相对于参考物确定统计学上显著的富集或耗竭。在一些实施方案中,至少一个核苷酸序列在第一细胞群中相对于第二细胞群在统计学上富集。在一些实施方案中,流式细胞术分选和磁珠富集相结合。

[0296] 在一些实施方案中,提供了从核酸文库中识别核苷酸序列的方法。该方法包括将

所述文库导入细胞群；使细胞文库与第二细胞群接触，通过磁珠富集基于至少一种标志物的表达来选择所述细胞文库的第一群，基于选定的第一群内的核苷酸序列相对于对照的统计学上显著的富集或耗竭来识别至少一种核苷酸序列。在一些实施方案中，富集或耗竭的实体是包含在核酸文库中的核苷酸序列。在一些实施方案中，第一细胞群中的至少一些被构造为表达由核酸文库成员编码的一种或多种多肽。在一些实施方案中，标志物表达与从文库中导入的核酸相关联。在一些实施方案中，“相关联(linked)”表示导入的核酸改变标志物表达。

[0297] 在一些实施方案中，提供了从核酸文库中识别核苷酸序列的方法。该方法包括：将核酸文库导入细胞群以形成细胞文库；使所述细胞文库与第一细胞群接触；通过磁珠富集基于至少一种标志物的表达来选择所述细胞文库的亚群；以及基于所述亚群内的核苷酸序列相对于对照的统计学上显著的富集或耗竭来识别至少一种核苷酸序列。在一些实施方案中，至少一些细胞亚群被构造为表达由核酸文库成员编码的一种或多种多肽。

[0298] 在一些实施方案中，选择是基于高于第一阈值水平的标志物的表达。在一些实施方案中，标志物适用于磁珠富集，这可能意味着但不限于所述标志物可通过细胞外表达被磁珠标记(例如，其必须是可到达细胞外的)。在一些实施方案中，核苷酸序列编码表达的多肽。

[0299] 在一些实施方案中，导入细胞群的核酸文库导致变体分子的表达。在一些实施方案中，这样的变体分子是T细胞受体序列。在一些实施方案中，这样的变体分子是开关受体。在一些实施方案中，这样的变体分子是CAR分子。

[0300] 在一些实施方案中，提供了从(任选的)核酸组合文库中识别编码T细胞受体 α (TCR α)-链和TCR β -链的核苷酸序列的方法。所述方法包括：将所述核酸文库导入能够表达TCR α -链和TCR β -链的细胞群中以制备细胞文库；并且基于子集内的核苷酸序列相对于对照的富集来确定变体核酸的第一群的至少一个核苷酸序列或核酸的同一性。在一些实施方案中，从第一细胞群中分离所述至少一个核酸。在一些实施方案中，基于响应于抗原的高于第一阈值水平的标志物的表达来选择第一细胞群。

[0301] 在一些实施方案中，本文提供的任何方法进一步包括施用表达抗原特异性TCR序列的T细胞以诊断或治疗感染或自身免疫的步骤。

[0302] 在一些实施方案中，对于本文提供的任何方法或组合物而言，T细胞可以是自体的或同种异体的。

[0303] 在一些实施方案中，对于本文提供的任何方法或组合物而言，活化标志物是CD69。

[0304] 在一些实施方案中，对于本文提供的任何方法或组合物而言，分离两个细胞群，一个细胞群具有高表达的CD69，而另一个细胞群具有低表达的CD69。

[0305] 在一些实施方案中，对于本文提供的任何方法或组合物而言，提供了包含根据本文提供的任何一种方法恢复的T细胞受体库的核苷酸文库。

[0306] 在一些实施方案中，对于本文提供的任何方法或组合物而言，提供了包含根据本文提供的任何方法所识别的核苷酸序列的核苷酸构建体。

[0307] 在一些实施方案中，对于本文提供的任何方法或组合物而言，提供了包含根据上述的核苷酸构建体的细胞。

[0308] 在一些实施方案中，提供了从样本中识别编码T细胞受体 α (TCR α)-链和TCR β -链的

核苷酸序列的方法。该方法包括：a) 对样本中的TCR- α 链和TCR β -链进行测序，b) 选择并对TCR α -链和TCR β -链的序列进行组合配对以创建TCR $\alpha\beta$ 对的文库，c) 将TCR $\alpha\beta$ 对的文库导入报告细胞池，d) 用呈递至少一种目的抗原的抗原呈递细胞刺激用TCR $\alpha\beta$ 对的文库修饰的所述报告细胞（在一些实施方案中，这可以通过本文所述的确切的两池方法来完成），e) 确定对所述至少一种目的抗原特异的TCR $\alpha\beta$ 对，和f) 将所述TCR $\alpha\beta$ 对导入细胞并选择含有所述TCR $\alpha\beta$ 对的细胞。

[0309] 在一些实施方案中，提供了从核酸组合文库中识别编码抗原特异性T细胞受体 α (TCR α) -链和TCR β -链对的核苷酸序列的方法。该方法包括：a) 将文库导入能够表达由多种变体核酸的成员编码的TCR α -链和TCR β -链的细胞群中，b) 基于响应于抗原的高于阈值水平的标志物的表达来选择细胞群的亚群（其可以任选地在抗原呈递细胞中），其中所述亚群包括多个细胞，c) 从所述亚群分离所述多种变体核酸的子集，d) 测定所述变体核酸的核苷酸序列，和e) 基于所述子集内的核苷酸序列相对于对照的富集来识别至少一种变体核苷酸序列。在一些实施方案中，该方法进一步包括提供包含编码TCR α 链和TCR β 链的多种变体核酸的文库。

[0310] 在一些实施方案中，被分选的 (sorted) 细胞的百分比是基于对照培养物和具有新生抗原表达的细胞培养物中具有标志物表达的T细胞百分比的比较。任何标志物都可用于细胞分选。在一些实施方案中，用于分选细胞的标志物包括例如CD4、CD8和CD69。

[0311] 图35A-图35E中示出了其他实施方案，它们示出了从通过基因合成产生的TCR文库中恢复抗原特异性TCR。这些实施方案示出了通过增加E:T的比例来增加TCR文库筛选平台的灵敏度的想法。这些条项 (item) 也体现在下面的一些实例中。

[0312] 本文使用的“步骤”的数字仅表示在该特定部分中讨论的实施方案，字母表示过程本身的步骤。因此，例如，“步骤35”表示标题为35的实施方案，而“a”表示该实施方案的步骤“a”。术语“步骤 (step)”表示过程的一部分，并不一定要求任何一个步骤在另一个步骤开始之前完成。

[0313] 步骤35A) 筛选设计示意图。使用来自卵巢癌 (OVC) 或结直肠癌 (CRC) 样本的5个特征性TCR和95个非特征性TCR创建 100×100 设计的组合TCR文库。该文库由Twist Bioscience使用人类V、CDR3和J片段组装而成，而恒定区 (C) 来自鼠源。该文库用于Jurkat报告T细胞的逆转录病毒转导。将多克隆报告T细胞与经工程改造以TMG形式呈递同源抗原的抗原呈递细胞 (APC) 共培养。对于本实施例，表达TMG抗原的EBV-LCL细胞和未被工程改造以呈现特定抗原的EBV-LCL用于共培养。

[0314] 步骤35B) 筛选的分选 (sorting) 策略。表达如35A) 中所述产生的 100×100 设计TCR文库的Jurkat报告T细胞与35A) 中所述的APC以1:1/1:2和1:3的比例共培养21小时。共培养后，使用基于B细胞上的CD20表达的磁珠选择来耗尽APC。B细胞耗竭后，使用CD69标志物，通过FACS对细胞进行T细胞活化分选。在一些实施方案中，可以通过任何方法对文库进行分选。分选策略包括（从左到右）顺序门控以选择淋巴细胞、门控以选择单细胞、门控以排除CD20⁺细胞，以及分别捕获表达高CD69和低CD69的细胞的两个分选门（“顶部 (top)”和“底部 (bottom)”）。结果示于图35B) 的图中。

[0315] 步骤35C) TCR表达盒的提取 (retrieval)。在一些实施方案中，可以通过多种技术中的任何一种来提取相关的TCR表达盒。在一些实施方案中，来自图35B) 的顶部和底部样本

的TCR表达盒可使用10C)中描述的PCR策略,随后进行条形码PCR来提取。还包括质粒TCR文库的对照PCR。使用安捷伦(Agilent) TapeStation分析第二轮PCR后的PCR产物。结果示于图35C中。

[0316] 步骤35D) 筛选数据的TCR富集分析。在一些实施方案中,可以以多种方式对来自步骤35C)的PCR反应产物池进行分析。例如,来自35C)的PCR产物池用于文库制备,并使用纳米孔技术进行测序。使用DESeq2 R包恢复TCR α 链和 β 链身份(identity),并识别差异表达的TCR组合。B细胞存在(x轴)和不存在(y轴)TMG表达的筛选的平均Rlog转换读段计数代表效应物与靶(E:T)比例为1:1、1:2和1:3的情况。五种特征性抗原反应性TCR在图35D中被描绘为较大的灰点。

[0317] 步骤35E) 五种特征性抗原反应性TCR的特征。使用DESeq2 R套装软件确定了35D)数据中最具差异性的TCR $\alpha \times \beta$ 链组合的排序以及伴随的p值。差异表征分析(Differential representation analysis)是本领域技术人员已知的,并且基于线性模型,假设富集的TCR被定义为相对于不存在抗原时“顶部”和“底部”样本,在存在(presented)抗原时在“顶部”样本中富集而在“底部”样本中耗竭。这些特性列于图35E中。

[0318] 图36中示出了其他实施方案,其示出了从通过基因合成生成的TCR文库中抗原特异性TCR的恢复。这些实施方案示出了使用基于珠的分选(bead-based sorting)而不是基于流的分选(flow-based sorting)作为分离用于TCR文库筛选的细胞的方式的思想。基于珠的细胞分选的优势在于其比基于FACS的细胞分选具有更高的可扩展性。

[0319] 步骤36A) 筛选设计示意图。使用来自卵巢癌(OVC)或结直肠癌(CRC)样本的5个特征性TCR和95个非特征性TCR创建 100×100 设计的组合TCR文库。该文库由Twist Bioscience使用人V、CDR3和J片段组装,而恒定区(C)为鼠源。该文库用于Jurkat报告T细胞的逆转录病毒转导。将多克隆报告T细胞与经工程改造以TMG形式呈递特异性抗原的抗原呈递细胞(APC)共培养。

[0320] 步骤36B) 筛选的分选(sorting)策略。表达如36A)中所述产生的 100×100 设计TCR文库的Jurkat报告T细胞与36A)中所述的APC以1:1的比例共培养21小时。共培养后,使用来自Miltenyi的AutoMACS进行连续细胞分离。首先,使用死细胞去除试剂盒去除死细胞。对于活细胞,使用基于B细胞上CD20表达的磁珠选择来耗尽APC。B细胞耗竭后,使用CD62L,一种在非活化的Jurkat细胞上表达的标志物的表达来分离CD20细胞。然后用抗CD69-生物素标记的抗体分别对CD62L-和CD62L+染色,然后用抗生物素微珠分离细胞。用于从中提取TCR盒的最终组分是代表“顶部”组分的CD20-、CD62L-、CD69+细胞和代表“底部”组分的CD20-、CD62L+、CD69-细胞。细胞分离过程的示意图示于图36B中。

[0321] 步骤36C) TCR表达盒的提取(retrieval)。在一些实施方案中,可以通过多种技术中的任何一种来提取相关的TCR表达盒。在一些实施方案中,使用在10C)中描述的PCR策略,随后进行条形码PCR来提取来自图36B)的顶部和底部样本的TCR表达盒。还包括质粒TCR文库的对照PCR。使用安捷伦(Agilent) TapeStation分析第二轮PCR后的PCR产物。结果示于图36C中。

[0322] 步骤36D) 筛选数据的TCR富集分析。在一些实施方案中,可以以多种方式对来自步骤36C)的PCR反应产物池进行分析。例如,来自36C)的PCR产物池用于文库制备,并使用纳米孔技术进行测序。使用DESeq2 R套装软件通过与库中存在的链进行比对,恢复TCR α 链和 β 链

的身份(identity),并识别差异表达的TCR组合。对于每个TCR,以灰色表示在B细胞存在(x轴)和不存在(y轴)TMG表达时筛选的平均Rlog转换读段计数,以黑点表示五种特征性抗原反应性TCR。

[0323] 步骤36E) 前7种最显著富集的TCR的特征。使用DESeq2 R套装软件识别来自36D) 中的数据差异表达的TCR $\alpha \times \beta$ 链的组合。差异表征分析(differential representation analysis)是本领域技术人员已知的,并且基于线性模型,假设富集的TCR被定义为相对于未表达TMG的“顶部”和“底部”样本而言,在表达TMG的“顶部”样本中富集,在表达TMG的“底部”样本中耗竭。前7个最显著命中的 α 链和 β 链,以及它们的表征(representation),它们的log2转换的倍数变化和差异表征的显著性列于图36E)。

[0324] 图37示出了其他实施方案,其示出了用于识别来自结直肠癌(CRC)患者2和患者4(pt2/pt4)的新生抗原反应性TCR的遗传筛查(genetic screen)的另外的分析。这些实施方案示出了使用组合串联小基因(tandemminigene, TMG) 编码来针对大量小基因有效筛选TCR文库的想法。

[0325] 步骤37A) 6 $\times$ 6组合TMG编码设计示意图。假定一个需要针对表达36个TMG的APC筛选TCR文库的筛选条件,可以创建APC池,每个APC池都表达6个TMG的独特组合。例如,C1池由表达TMG1、TMG2、TMG3、TMG4、TMG5和TMG6的APC组成。R1池由表达TMG1、TMG7、TMG 13、TMG19、TMG25和TMG 31的APC组成。可以对每个APC池执行单独的TCR库筛选。根据筛选方法中TCR识别的两个池的组合,被识别出的TMG可以确定为在两个池中均有代表(represented)的TMG。

[0326] 步骤37B) 作为pt2 TCR文库筛选的重复次数的函数的所有TCR $\alpha \times \beta$ 组合的秩次序的分析。使用全部3次重复或3次重复中的2次对来自图10G) 的pt2 TCR文库筛选数据进行分析。针对这两种情况,使用DESeq2 R套装软件识别了差异表达的TCR组合。富集显著性排序基于Wald统计(使用DESeq2计算),给最高排序(排序1) 分配最高的Wald统计值。基于基于2或3次重复的分析的排序分别表示在y轴和x轴上。图10G) 中识别的新生抗原反应性TCR先导(lead) 表示为较大的黑点。

[0327] 步骤37C) 基于CRC TCR库筛选的2或3次重复的统计分析汇总表。将图37B) 的分析应用于图10G) 中进行的所有患者TCR文库筛选。表中列出了图37B) 中识别的所有新生抗原反应性TCR先导及其Bonferroni校正p值,以了解相对于底部样本,顶部样本中TCR富集的统计学显著性。该p值表示基于3次筛选重复试验或3次筛选重复试验中的2次的任意组合的分析。最后一列中的复选标记表示,无论筛选分析中是否包括2次或3次筛选重复,也无论选择了3个重复中的2个组合中的哪一个,都会选择相同的排序靠前的TCR先导。图37B) 中示出了一个实例,其中从pt4中选择用于进一步验证的7个排序靠前的新生抗原反应性TCR(见图10G) 与基于仅使用2个重复的分析的7个排序靠前的TCR相同。

[0328] 步骤37D) 用于成对TCR富集分析的pt4样本表。纳入了六份样本(每份样本均由顶部和底部样本表示),用于成对TCR富集分析。样本包括表达TCR文库的报告T细胞与分别表达TMG1、TMG2、TMG3或TMG4的B细胞系的共培养物(分别为样本1-4),以及以1:1:1:1的比例与这些B细胞系混合的池的共培养物(样本5)。对于样本6,与未被工程改造以表达任何外源性抗原的B细胞进行共培养。

[0329] 步骤37E) 成对TCR富集分析结果。使用DESeq2分析图19D) 中所有可能的样本对的顶部样本相对于底部样本的TCR富集。差异表征分析是本领域技术人员已知的,并且基于线

性模型,假设富集的TCR被定义为相对于不表达TMG时的“顶部”和“底部”样本,在表达TMG时在“顶部”样本中富集,在“底部”样本中耗竭。Bonferroni校正p值按递增顺序排序和绘制。浅灰色点代表 $\alpha 43.\beta 16$ TCR分析的p值,深灰色点代表 $\alpha 9.\beta 14$ TCR分析的p值。顶部异常值(outliers)与图10G)中确定的TCR相同。TCR $\alpha 9.\beta 14$ 的反应性可归因于由TMG4表达的抗原,因为这是样本4和5之间共有的TMG效应。TCR $\alpha 43.\beta 16$ 的反应性可归因于TMG3,因为这是样本3和样本5之间的共有TMG。

[0330] 图38中示出了其他实施方案,其示出了用于识别来自结直肠癌(CRC)患者2的新生抗原反应性TCR的遗传筛查的另外的分析。这些实施方案示出了平台的灵敏度,并支持可从遗传筛查数据确定TCR特征(如TCR敏感性)的观点。

[0331] 步骤38A) 筛选和验证数据之间的TCR激活和TCR背景激活的相关性。基于图10G)的数据,鉴定了遗传筛查方法中TCR反应性不同的16个CRC pt2 TCR。首先,使用DESeq2 R套装软件计算经Rlog转换的读段计数,从顶部样本的Rlog值中减去底部样本的Rlog值,并表示通过B细胞在存在(x轴) TMG表达和不存在(y轴) TMG表达的情况下进行的共培养(图38)。在独立验证实验中,这些TCR在报告T细胞中表达,并与表达相关TMG(TM2)的EBV-B细胞共培养。通过FACS分析,使用CD69标志物测定T细胞活化。为了将验证数据与TCR文库筛选数据进行比较,将筛选中TCR的倍数激活定义为来自与表达TM2的EBV-B细胞共培养的顶部样本和底部样本的Rlog值之间的差异(y轴;中间图)。验证实验中的倍数激活定义为CD69+细胞与表达TM2的细胞与非工程改造的EBV-B细胞共培养后的比率(x轴;中间图)。筛选背景被定义为来自与非工程改造的EBV-B细胞共培养的给定TCR在顶部(top)与底部(bottom)的富集(y轴;右图)。验证实验的背景定义为CD69+细胞与非工程改造的EBV-B细胞共培养后的百分比(x轴;右图)。

[0332] 在一些实施方案中,从TCR $\alpha\beta$ 文库中恢复抗原反应性TCR可以通过基于对抗原的响应分离一个或多个亚群来实现。在一些实施方案中,该方法需要一个或多个以下步骤:i) 报告T细胞的基因工程以允许表达TCR $\alpha\beta$ 文库的TCR;ii) 进行这些细胞与表达至少一种抗原的抗原呈递细胞的共培养;iii) 基于T细胞活化标志物的细胞分离成a) 表达一种或多种T细胞活化标志物的“顶部”群;和b) 缺乏(或低)表达一种或多种T细胞活化标志物的“底部”群;iv) 使用对基因组DNA的PCR和随后的深度测序从顶部和底部样本进行TCR识别;和v) 相对于底部样本在顶部样本中富集的至少一种抗原反应性TCR的识别。T细胞活化标志物的表达可以是区分活化T细胞的标志物的相对高表达(例如CD69),或区分非活化T细胞的标志物的相对低表达(例如CD62L)。

[0333] 在一些实施方案中,顶部(top)是前1、5、10、20、30、40或50%,底部(bottom)是后1、5、10、20、30、40或50%,包括顶部和底部的任意两个标注值之间的任意对范围。在一些实施方案中,在本文提供的任何实施方案中,采用顶部/底部法(top/bottom approach)(其中,使用顶部群和底部群两者)。

[0334] 在一些实施方案中,提供了从核酸文库中识别编码抗原特异性T细胞受体 α (TCR α)-链和TCR β -链对的核苷酸序列的方法。该方法包括a) 将核酸文库导入能够表达TCR α -链和TCR β -链的细胞群中以制备细胞文库;b) 基于响应于抗原的高于第一阈值水平的标志物的表达来选择所述细胞文库的第一群;和c) 从所述文库的第一群中分离变体核酸的第一群。在一些实施方案中,所述方法进一步包括a) 确定变体核酸的所述第一群的至少一个核

核苷酸序列或核酸同一性；和b) 基于所述子集内的所述核苷酸序列相对于对照的富集来识别至少一种变体核苷酸序列。在一些实施方案中，所述阈值水平基于以下至少一项：

[0335] a) 基于标志物的表达的所述总的细胞池的百分比的恢复；或者

[0336] b) 来自所述总的细胞池的最少数量的细胞的恢复；或者

[0337] c) 基于磁性探针与响应抗原表达的至少一种标志物的结合，由磁体保留的细胞的恢复。

[0338] 在一些实施方案中，对照是低于第二阈值的第二细胞群。在一些实施方案中，所述对照是以下的一种或多种：参考细胞群、导入细胞群的核酸组合文库、基于低于第二阈值的表达标志物从与第一群相同的细胞群分选出来的细胞群、和/或从表达相关TCR文库的报告T细胞与未被工程改造以表达外源抗原的B细胞的共培养物中获得的至少一种细胞群。在一些实施方案中，底部(或对照)样本选自与顶部样本相同的细胞群，但具有低活化标志物表达，或其中底部样本获自表达相关TCR文库的报告T细胞和未被工程改造以表达外源抗原的B细胞的共培养物。在一些实施方案中，顶部部分和底部部分之间没有重叠(overlap)。在一些实施方案中，所述方法进一步包括将抗原加入细胞群中。在一些实施方案中，分离第一群和/或对照通过以下至少一项来实现：a) 磁珠富集，b) 流式细胞术分选或c) 二者。在一些实施方案中，对照是以下的一种或多种：参考细胞群、导入细胞群的核酸组合文库、基于低于第二阈值的表达标志物从与第一群相同的细胞群中分选的细胞群、或从表达相关TCR文库的报告T细胞与不呈递任何外源性抗原的抗原呈递细胞如B细胞的共培养物中获得的至少一种细胞群。

[0339] 在一些实施方案中，建立了顶部-底部法(top-bottom approach)，使得抗原反应性TCR在抗原刺激后将变为活化标志物阳性，因此相对于底部群，此类TCR将在顶部群中富集。如实施例24所述，通过各种附图说明了顶部-底部法。

[0340] 在一些实施方案中，底部样本(或对照)可以是任何参考细胞群或TCR质粒的参考文库。底部样本可以从与顶部样本相同的细胞群中分选出来，但是具有低活化标志物表达。底部样本可从表达相关TCR文库的报告T细胞和未被工程改造以表达外源抗原的B细胞的共培养物中获得。底部样本可以是TCR质粒文库，用于创建从中分选顶部样本的报告T细胞。在一些实施方案中，在差异TCR表征分析期间，可以将顶部和底部样本中的TCR表征(representation)与任何其他另外的样本中的TCR表征(representation)进行比较。在一些实施方案中，此类另外的样本可以是质粒TCR文库。在一些实施方案中，此类另外的样本可以衍生自表达相关TRC文库的报告T细胞和未被工程改造以表达外源抗原的B细胞的共培养物。

[0341] 在一些实施方案中，该方法是一种从不同的T细胞群恢复T细胞受体(TCR)库的方法。该方法可以包括测定个体样本中配对的TCR α 链和TCR β 链的核苷酸序列或氨基酸序列；从总库中选择TCR $\alpha\beta$ 对序列；通过创建所选的TCR α/β 对的文库来创建TCR库；以及从所创建的TCR文库中识别具有期望特征的至少一个TCR α/β 对。

[0342] 在一些实施方案中，TCR α/β 对可以包括TCR序列基序。

[0343] 在一些实施方案中，本文提供了TCR $\alpha\beta$ 对的文库。在一些实施方案中，可以通过本文提供的任何方法来创建这些文库。在一些实施方案中，文库来自非活样本。

[0344] 在一些实施方案中，本文提供了将TCR $\alpha\beta$ 对文库引入报告细胞池的细胞群和/或文

库。因此,在一些实施方案中,提供了包括选择的和组合配对的TCR α -链和TCR β -链序列 (TCR $\alpha\beta$ 对的文库)的报告细胞池。

[0345] 在一些实施方案中,文库可以是经刺激的文库(stimulated library)。在一些实施方案中,文库可以包括用TCR $\alpha\beta$ 对文库修饰的报告细胞,并且可以进一步包括呈递至少一种目的抗原的抗原呈递细胞。在一些实施方案中,所述至少一种目的抗原可以是自体的或同种异体的。

[0346] 在一些实施方案中,本文提供了涉及来自样本或非活样本的特征的一个或多个TCR $\alpha\beta$ 对的文库。

[0347] 关于筛选的其他实施方案

[0348] 描述了本发明的其他实施方案。本发明的任何实施方案都可以与本文提供的任何其他实施方案相结合。本文提供了通过遗传功能增益/功能丧失筛选(也称为“遗传变体文库筛选”)经由高通量核酸测序来识别变体序列文库中的目的序列的方法(本文也称为“筛选方法”)。筛选方法可用于筛选变体核苷酸序列文库,其中仅通过识别至少600bp的变体核苷酸序列才能明确区分单个核酸序列。

[0349] 在一些实施方案中,任何合适的方法均可采用识别涉及池化抗原(pooled antigen)的方法。例如,在一些实施方案中,识别或刺激包括:a)选择若干抗原;b)创建抗原池,其中每个抗原正好存在于两个抗原池中;c)评估表达至少一种T细胞受体的报告细胞对每个抗原池的反应性;和d)通过评估对恰好两个抗原池的反应性,确定所述至少一种T细胞受体是否对任何选定的抗原具有反应性。在一些实施方案中,通过成对富集分析检测针对恰好两个抗原池的反应性。在一些实施方案中,文库是TCR文库。在一些实施方案中,采用活化标志物。在一些实施方案中,采用顶部-底部比较(如本文所述)来评估反应性。在一些实施方案中,虽然其他方法可以使用针对单个池的反应性来创建独特的反应性模式,但是该过程可以使用成对分析来通过特异性分析重复来增加信号强度。

[0350] 本发明的筛选方法可以是高通量方法。多克隆基因文库筛选可允许在单个实验中筛选大量(例如,几万个)蛋白质变体,而不需要产生和分析单个克隆(individual clone)。此外,与单个变异基因的合成相比,生成变异基因文库的成本可以更低且耗时更少。筛选方法的一些实施方案允许变体的功能选择,并且可以基于一种或多种功能特性选择蛋白质变体。本发明的筛选方法可以是灵敏的筛选方法。虽然目的变体是根据功能表型选择的,但目的变体的序列同一性是基于DNA测序方法识别的,这种方法甚至可以以高灵敏度来检测罕见的变体(rare variant)。在一些实施方案中,本文提供的实施方案的灵敏度可以允许人们以期望水平进行区分,包括1:1000、1:10000、1:100000、1:1000000或甚至更低。在一些实施方案中,更高的灵敏度是可能的,条件是:(i)分析足够数量的细胞和(ii)可以产生足够的序列读段。在一些实施方案中,要考虑的一个因素是分用(demultiplexing)。然而,这可以通过以下方式解决,例如:i)不进行多路复用(multiplexing)或ii)以丢弃更多读段为代价来提高条形码阈值(例如,更多的测序)。

[0351] 筛选方法的一些实施方案是对蛋白质变体进行高通量遗传变体文库筛选的方法,其中变体之间的区分基于>200个氨基酸的延伸和基于多克隆群分析。在一些实施方案中,本筛选方法包括筛选方案和生物信息学过程,其克服了一些NGS测序读段数中的高错误率,例如由Oxford Nanopore平台产生的那些错误率。

[0352] 本发明的筛选方法可以包括识别任何合适的变体蛋白,其与大小无关并且不受序列多样性位置的限制。在一些实施方案中,该方法包括从大的TCR集合中选择目的T细胞受体(TCR)序列,其中不同的TCR α 链和TCR β 链的配对是未知的或不明确的,并且包括确定编码TCR α 和TCR β 可变区的变体基因盒的全序列。在一些实施方案中,本发明的筛选方法允许基于由TCR介导的报告细胞的功能性应答(例如CD69上调),以高通量(例如,通过避免产生克隆)筛选TCR文库。

[0353] 在一些实施方案中,该方法包括从分子设计方面有很大差异的大量CAR变体的集合中识别具有增强性质的嵌合抗原受体(CAR)序列,例如通过从几个不同信号传导结构域的池中选择的多达三个不同信号传导结构域的组合组装。在一些实施方案中,本筛选方法允许通过筛选在整个CAR分子中具有突变多样性的变体来全面增强CAR。

[0354] 对于本文提供的任何实施方案,在适当的情况下,可以存在一个以上的CAR细胞内信号传导结构域。在一些实施方案中,存在至少两个CAR细胞内信号传导结构域。在一些实施方案中,所述至少两个CAR细胞内信号传导结构域是以下中的至少两个:CD3 ϵ 、CD3 ζ ITAM1、CD3 ζ ITAM12、CD3 ζ ITAM 123、CD3 ζ 与CD3 δ 、CD3 ϵ 和CD3 γ 的任意ITAM、CD8 α 、CD28、ICOS、4-1BB(CD137)、OX40(CD134)、CD27和CD2。

[0355] 在一些实施方案中,本发明的筛选方法包括分析多克隆报告细胞群而不衍生单细胞克隆。在一些实施方案中,筛选方法可用于筛选含有>10000个变体的文库以识别目的组合(例如,在100倍的覆盖率(coverage)下),而不产生单细胞克隆。

[0356] 在一些实施方案中,覆盖量(amount of coverage)取决于许多因素,包括筛选的主要焦点(富集筛选(较低的覆盖率)或耗竭筛选(较高的覆盖率))、文库内个体变体的表征的分布、选择期间的细胞损失等。一般来说,可以使用50-10000的范围,包括100-2000或例如100-400倍的富集筛选(enrichment screen)。

[0357] 在一些实施方案中,本发明的筛选方法提供了分子先导识别(molecule lead identification)和增强的遗传筛查方法,用于在整个完整蛋白质中具有突变多样性的较大蛋白质。与其他可用的方法相比,该方法在一些实施方案中可在识别分子先导方面提供高通量(降低成本和时间线(timeline))和高灵敏度。

[0358] 一般而言,筛选方法可以包括(1)产生包含至少两个可以(或仅可以)通过测定它们的总核苷酸序列的至少600bp而被明确地识别的变体核苷酸序列的变体核苷酸序列文库;(2)将变体核苷酸序列文库导入报告细胞;(3)基于至少一种功能特性选择报告细胞;(4)从选择的报告细胞中分离变体核苷酸序列;(5)测定分离的变体核苷酸序列的总核苷酸序列的至少600bp和(6)选择至少一种目的变体核苷酸序列。

[0359] 对于本文提供的任何实施方案(提供对600bp延伸(stretch)或氨基酸等效长度的参考物),文库的序列可变性可以存在于总计超过600bp的延伸中。在一些实施方案中,文库将含有或由或基本上由或包含长于1500bp的扩增子组成。在一些实施方案中,文库将包含/由或基本上由至少30、40、50、60、70、80、90、95、98、99或100%的大于1500bp的扩增子组成。

[0360] 关于图8,提供了本发明的筛选方法的一些实施方案。该方法可以包括提供810组合文库811,组合文库811包括多种变体核酸(此处例示为812a、812b、812c)。文库中的每种变体核酸可以包含长度至少为600bp的连续部分814。

[0361] 如本文所用,关于生物聚合物(例如,核酸或多肽)的“连续的(contiguous)”,是指

没有中间序列(例如,没有中间核苷酸或核苷酸序列的核苷酸序列,没有中间氨基酸或氨基酸序列的氨基酸序列)的生物聚合物的各个构建块(例如,核苷酸或氨基酸)的序列。连续部分814可以包含两个以上变体核苷酸子序列。

[0362] 如本文所用,“变体核苷酸子序列”可以包括定义核酸和/或由核酸编码的多肽的功能活性单位的核苷酸序列家族中的任何一个。在一些情况下,变体核苷酸子序列可赋予和/或促成核酸和/或由核酸编码的多肽的不同功能活性(例如,结合亲和力、特异性)。在一些情况下,核苷酸序列家族通过以下方式赋予和/或促成不同的功能活性:变体核苷酸子序列在核酸中的位置;与共有序列(consensus sequence)具有序列相似性;和/或具有可变区和不变区,其中不变区被相同核苷酸序列家族的其它成员共有。在一些实施方案中,TCR文库包括具有对应于TCR α -链或其一个或多个功能域(例如,TCR α V区、TCR α 互补决定区3(CDR3)、TCR α J-片段、TCR α 恒定区)的变体核苷酸子序列的变体核酸,以及具有对应于TCR β -链或其功能域(例如,TCR β V区、TCR β 互补决定区3(CDR3)、TCR β J-片段、TCR β 恒定区)的变体核苷酸子序列的变体核酸。在一些实施方案中,CAR文库包括具有对应于一个或多个CAR功能域(例如,抗原结合结构域、铰链结构域、跨膜结构域和细胞内信号传导结构域,其可以包括2-3个信号模块)的变体核苷酸子序列的变体核酸。

[0363] 进一步参考图8,文库的核酸812a可以包含文库中多个可能的变体(816a,816b)之一的变体核苷酸子序列816a。核酸812a还可以在不同位置包含另一个多个可能的变体(818a,818b)之一的变体核苷酸子序列818a。文库中的另一个核酸812b可以包含来自第一组合816a/818a的变体核苷酸子序列的不同组合816a/818b。然而,第三核酸812c可以包含来自第一组合816a/818a或第二组合816a/818b的变体核苷酸子序列的不同组合816b/818b。此外,连续部分的一端(例如,5'端或3'端)可由一个变体核苷酸子序列定义,而连续部分的另一端(例如,分别为3'端或5'端)可由另一个变体核苷酸子序列定义。

[0364] 连续部分可由下式表示:5'-A*-X-B*-3',其中A*和B*表示不同家族的变体核苷酸子序列,X可以不存在(在这种情况下,连续部分为5'-A*-B*-3'),或者如果存在,可以是任何长度的任何核苷酸序列。在一些实施方案中,A*可以是变体核苷酸子序列家族的任何成员(例如,A1、A2、A3等)。在一些实施方案中,B*可以是变体核苷酸子序列家族的任何成员(例如,B1、B2、B3等)。在一些实施方案中,X可以包括变体核苷酸子序列的一个或多个家族的成员(例如,C、D等)。

[0365] 筛选方法可进一步包括将文库820导入细胞群822,细胞群822可以表达由多种变体核酸的成员826a、826b编码的一种或多种基因产物(例如,多肽)824a、824b。

[0366] 筛选方法可包括基于至少一种功能特性832(例如,所表达的多肽与配体的结合)选择细胞群的亚群830,其中所述功能特性取决于引入细胞的核酸成员826a、826b中的变体核苷酸子序列的组合。细胞亚群可以包括多个细胞。在一些实施方案中,细胞亚群可以包括多种变体核酸的多个不同成员,其中不同成员通过具有不同的变体核苷酸子序列的组合而彼此不同。

[0367] 在一些实施方案中,筛选方法可以包括例如通过从细胞中提取基因组DNA来从细胞亚群中分离840(多种变体核酸的子集842)。

[0368] 在一些实施方案中,筛选方法可以进一步包括测定多种变体核酸的子集的个体成员的连续部分的核苷酸序列850,例如,通过对子集842中的变体核酸的至少连续部分进行

高通量测序。

[0369] 在一些实施方案中,所述方法可进一步包括识别860存在于亚群的细胞中的变体核苷酸子序列816a/818a的组合。在一些实施方案中,识别包括分析在细胞亚群中发现的两种以上组合中的变体核苷酸子序列的组合与预定阈值水平相比,或与该特定组合在对照细胞亚群中的丰度相比是否富集。在一些实施方案中,富集的组合被鉴定为赋予细胞选择细胞亚群所依据的功能特性。

[0370] 参照图9,提供了本筛选方法的一些实施方案。方法900可以包括910提供包含多种变体核酸的组合文库,所述多种变体核酸中的每一个具有至少600bp的连续部分,其中所述连续部分包括两个以上变体核苷酸子序列的组合,其中所述两个以上变体核苷酸子序列的第一变体核苷酸子序列定义所述连续部分的第一端,并且所述两个以上变体核苷酸子序列的第二变体核苷酸子序列定义所述连续部分的与所述第一端相对的第二端。920可以将组合文库导入经构造以表达由所述多种变体核酸的成员编码的一种或多种多肽的细胞群中。然后,该方法可以包括930基于依赖于所述两个以上变体核苷酸子序列的至少一种功能特性来选择所述细胞群的亚群,其中所述亚群包括多个细胞。该方法可以包括940从所述亚群中分离所述多种变体核酸的子集(例如,一个或多个),和950测定该子集的各个成员的连续部分的核苷酸序列。然后,960基于确定的核苷酸序列,可以识别两个以上变体核苷酸子序列的一个或多个组合。

[0371] 在一些实施方案中,对于每个筛选,都存在一个“参考物”和一个“顶部”样本。“顶部”样本包含显示最高CD69表达的分选细胞,“参考物”样本包含显示低CD69表达的分选细胞。在一些实施方案中,可以基于任何其他活化标志物来分离样本,包括但不限于CD25、CD62L、CD137、IFN- γ 、IL-2、TNF- α 、GM-CSF、合成启动子报告标志物或增殖标志物。为此,可以针对库中复杂度为Y的每种变体以至少100-200倍的覆盖率分离变体。每个样本的变体/细胞总数可以 $>100-200 \times Y$ 。

[0372] 当一个核酸的变体核苷酸序列编码的氨基酸序列不同于另一个核酸中相应的变体核苷酸序列编码的氨基酸序列时,可以认为这一核酸与组合文库中的另一核酸“不同(different)”或“有区别(distinct)”。在一些实施方案中,一种核酸可以被认为与组合文库中的另一种核酸“不同(different)”或“有区别(distinct)”,是基于它们编码相同氨基酸序列时它们的核苷酸序列的差异。

[0373] 组合文库可以包括任何合适数量的不同(或“变体”)核酸。在一些实施方案中,组合文库包括约100种以上,例如约200种以上、约300种以上、约400种以上、约500种以上、约600种以上、约700种以上、约800种以上、约900种以上、约1000种以上、约2000种以上、约3000种以上、约4000种以上、约5000种以上、约7500种以上、约10000种以上、约20000种以上、约50000种以上、约 1×10^5 种以上、约 2×10^5 种以上、约 5×10^5 种以上、约 1×10^6 种以上、约 1×10^7 种以上、约 1×10^8 种以上、约 1×10^9 种以上、约 1×10^{10} 种以上,包括约 1×10^{11} 种以上不同的核酸,或在由上述任何两个值限定的范围内的若干不同的核酸。在一些实施方案中,组合文库包括约100种至约200种、约200种至约500种、约500至约1000种、约1000种至约5000种、约5000种至约 1×10^4 种、约 1×10^4 种至约 2×10^4 种、约 2×10^4 种至约 5×10^4 种、约 5×10^4 种至约 1×10^5 种、约 1×10^5 种至约 1×10^6 种、约 1×10^6 种至约 1×10^7 种、约 1×10^7 种至约 1×10^8 种、约 1×10^8 种至约 1×10^9 种、约 1×10^9 种至约 1×10^{10} 种、约 1×10^{10} 种至约 1×10^{11} 种

或更多不同的核酸。在一些实施方案中,文库包括至少100种,例如至少200种、至少300种、至少400种、至少500种、至少600种、至少700种、至少800种、至少900种、至少1000种、至少2000种、至少3000种、至少4000种、至少5000种、至少7500种、至少10000种、至少20000种、至少50000种、至少 1×10^5 种、至少 2×10^5 种、至少 5×10^5 种、至少 1×10^6 种、至少 1×10^7 种、至少 1×10^8 种、至少 1×10^9 种、至少 1×10^{10} 种,包括至少 1×10^{11} 种的两个变体核苷酸子序列的不同组合。

[0374] 在一些实施方案中,一种或多种多肽(在一些实施方案中由组合文库的变体核酸编码)包括:T细胞受体 α (TCR α)-链和TCR β -链;嵌合抗原受体(CAR);开关受体;或抗体或其抗原结合片段的一条或多条链。在一些实施方案中,第一变体核苷酸子序列编码TCR α 变体氨基酸序列,第二变体核苷酸子序列编码TCR β 变体氨基酸序列。在一些实施方案中,所述两种以上变体核苷酸子序列编码以下一种或多种:TCR V区、TCR互补决定区3(CDR3)、TCR J-片段和TCR恒定区。在一些实施方案中,两种以上变体核苷酸子序列中的每一种均编码CAR的抗原结合结构域、铰链结构域、跨膜结构域或一个或多个细胞内信号传导结构域。

[0375] 在一些实施方案中,使文库中多种变体核酸的相邻部分之间的编辑距离最大化。在一些实施方案中,例如通过控制密码子使用,使组合文库的任何两种变体核酸之间的编辑距离最大化。在一些实施方案中,其可以包括密码子优化。在一些实施方案中,本文提供的涉及文库的任何方法可以包括基于以下至少一项来优化文库中的核苷酸序列(例如,多种变体核酸中的核苷酸序列):为宿主细胞引入优选的密码子使用、mRNA结构稳定性的优化、重复序列的避免、均聚物的长延伸的避免和给定变体核酸序列内局部GC含量的较大差异的避免。

[0376] 在一些实施方案中,基于以下至少一项来优化文库中多种变体核酸的核苷酸序列:1)本文提供的任何方法,其中细胞群的细胞被遗传修饰,或2)本文提供的任何方法,其中细胞用CD4和/或CD8重组并用于筛选I类和/或II类限制性TCR序列。

[0377] 在一些实施方案中,使用的细胞是T细胞。

[0378] 在一些实施方案中,细胞亚群和对照细胞群不重叠。在一些实施方案中,不重叠表示两个群中的细胞具有不同的活化状态,但可以携带相同的变体核酸。

[0379] 在一些实施方案中,所述一种或多种多肽包含TCR α -链和TCR β -链,并且其中所述不变氨基酸(invariant amino acid)序列包含TCR β 恒定区。

[0380] 在一些实施方案中,确定包括获得连续部分的每个核苷酸序列的至少25、至少50、100、至少200、至少300、至少400、至少500或至少1000的平均覆盖率(average coverage)。

[0381] 在一些实施方案中,涉及反应性评价或可进一步包括反应性评价的任何方法均可采用顶部-底部比较(top-bottom comparison)来评价反应性。

[0382] 在一些实施方案中,本文提供的涉及文库的任何方法可包括基于以下至少一项来优化文库中的核苷酸序列(例如,多种变体核酸的核苷酸序列):为宿主细胞引入优选的密码子使用、mRNA结构稳定性的优化、重复序列的避免、均聚物的长延伸的避免和给定变体核酸序列内局部GC含量的较大差异的避免。

[0383] 在一些实施方案中,通过抗原呈递细胞来呈递抗原。

[0384] 在一些实施方案中,所述文库是组合文库。

[0385] 在一些实施方案中,由细胞提供抗原。

[0386] 在一些实施方案中,该过程涉及高度的抗原多样性和/或复杂性。

[0387] 在涉及库的一些实施方案中,所述文库是组合文库。在一些实施方案中,组合文库是TCR文库。

[0388] 连续部分可具有至少600bp的任何合适长度。在一些实施方案中,在基于功能特性选择细胞亚群之后,连续部分的长度取决于用于对连续部分进行测序的测序平台的读段长度(例如,准确的读段长度)。在一些实施方案中,所述连续部分具有至少600bp的长度,例如,至少700bp、至少800bp、至少900bp、至少1000bp、至少1100bp、至少1200bp、至少1300bp、至少1400bp、至少1500bp、至少1750bp、至少2000bp、至少2500bp、至少3000bp、至少4000bp、至少5000bp、至少6000bp、至少7000bp、至少8000bp、至少9000bp、至少10000bp、至少11000bp、至少12000bp、至少13000bp、至少14000bp或至少15000bp,或在上述任意两个值限定的范围内的长度。在一些实施方案中,连续部分的长度为约600bp至约15000bp,例如约800bp至约12000bp,约1000bp至约10000bp,约1000bp至约8000bp,约1000bp至约6000bp,约1000bp至约5000bp,包括约1000bp至约4000bp。

[0389] 组合文库中的变体核酸可以具有任何合适数量的变体核苷酸序列。在一些实施方案中,组合文库中的所有核酸具有相同数量的变体核苷酸序列。在一些实施方案中,组合文库中的变体核酸具有2个、3个、4个、5个、6个、7个、8个、9个、10个、12个、15个、20个以上变体核苷酸序列,每个变体核苷酸序列可以具有可以组合方式组装以组装变体核酸的连续部分的变体。

[0390] 组合文库可以包括每个变体核苷酸序列的任何合适数量的变体。在一些实施方案中,组合文库中变体核苷酸序列的变体的数量是2以上、3以上、4以上、5以上、6以上、7以上、8以上、9以上、10以上、12以上、14以上、16以上、18以上、20以上、25以上、30以上、35以上、40以上、45以上、50以上、75以上,100以上、125以上、150以上、175以上、200以上、250以上、300以上、350以上、400以上、450以上、500以上、600以上、700以上、800以上、900以上、1000以上、1500以上、2500以上、3000以上、4000以上、5000以上、7500以上、10000以上,或在上述任何两个值所限定的范围内的变体的数量。在一些实施方案中,组合文库中变体核苷酸序列的变体数量为2至10、10至20、20至30、30至40、40至50、50至100、100至200、200至500、500至1000、1000至2000、2000至5000或5000至10000。在一些实施方案中,文库内的个体变体的频率分布使得>80%的那些变体具有在从中值频率/8开始到中值频率*8结束的范围内的频率。

[0391] 在一些实施方案中,对于本文提供的任何与筛选和/或文库相关方法,通过与抗原(如新生抗原)结合来选择或识别TCR对和/或表达TCR对的T细胞,其中所述抗原由B细胞或抗原呈递细胞表达。

[0392] 在一些实施方案中,对于本文提供的任何与筛选和/或文库相关方法,所述抗原或新生抗原来自一个个体的肿瘤,并且TCR对的TCR α 和TCR β 也各自来自所述个体(意味着单个个体同时具有抗原序列以及TCR α 和TCR β 序列)。

[0393] 在一些实施方案中,对于本文提供的任何与筛选和/或文库相关方法,存在至少2个、3个、4个、5个、6个、7个、8个、9个、10个、50个、100个、500个、1000个、10000个、100000个或100万个TCR对(或包含这些对的细胞),并且存在至少2个、3个、4个、5个、6个、7个、8个、9个、10个、50个、100个、500个、1000个、10000个、100000个或100万个抗原。

[0394] 在一些实施方案中,对于本文提供的任何与筛选和/或文库相关方法,a)通过与抗原(例如新生抗原)结合来选择或识别TCR对和/或表达TCR对的T细胞,其中所述抗原由B细胞或抗原呈递细胞表达,b)所述抗原或新生抗原来自一个个体的肿瘤,并且TCR对的TCR α 和TCR β 也各自来自所述个体(意味着单个个体同时具有抗原序列以及TCR α 序列和TCR β 序列),和c)存在至少2个、3个、4个、5个、6个、7个、8个、9个、10个、50个、100个、500个、1000个、10000个、100000个或100万个TCR对(或包含这些对的细胞)并且存在至少2个、3个、4个、5个、6个、7个、8个、9个、10个、50个、100个、500个、1000个、10000个、100000个或100万个抗原。

[0395] 组合文库可以是任何合适的生物分子(例如,蛋白质、核酸、核蛋白等)的文库。在一些实施方案中,合适的生物分子是其中序列(例如,蛋白质和/或核酸序列)可以在对应于至少600个核苷酸的序列单元(例如,氨基酸或核苷酸)的连续部分上变化的生物分子。合适的组合文库包括但不限于TCR、CAR、抗体、RNA引导的核酸酶等的组合文库。

[0396] 在一些实施方案中,组合文库包括来自不同T细胞群的T细胞受体(TCR)库。在一些实施方案中,组合文库的多个核酸包括编码一个或多个TCR α 功能域(例如,TCR α V区、TCR α 互补决定区3(CDR3)、TCR α J-片段、TCR α 恒定区)和一个或多个TCR β 功能域(例如,TCR β V区、TCR β 互补决定区3(CDR3)、TCR β J-片段、TCR β 恒定区)的变体核苷酸序列。在一些实施方案中,组合文库TCR的核酸的连续部分具有600bp至2000bp之间,例如800bp至1900bp之间、1000bp至1900bp之间、1200bp至1900bp之间、1400bp至1900bp之间、1500bp至约1900bp之间、1600bp至1900bp之间、1700bp至1900bp,或约1800bp的长度。

[0397] 在一些实施方案中,组合文库的多个核酸包括编码一个或多个嵌合抗原受体(CAR)功能域(例如,抗原结合结构域、铰链结构域、跨膜结构域和细胞内信号传导结构域,细胞内信号传导结构域可以包括2-3个信号传导模块)的变体核苷酸序列。在一些实施方案中,组合文库CAR的核酸的连续部分具有600bp至2000bp之间,例如800bp至1900bp之间、1000bp至1800bp之间、1200bp至1800bp之间、1400bp至1700bp之间或约1500bp的长度。

[0398] 在一些实施方案中,组合文库包括抗体重链和轻链序列的库。在一些实施方案中,组合文库的多个核酸包括编码一个或多个抗体重链功能域(例如,重链可变区(包括一个或多个CDR、框架区)和/或重链恒定区(包括CH1、CH2、CH3和铰链区中的一个或多个)的变体核苷酸序列。在一些实施方案中,组合文库的多个核酸包括编码一个或多个抗体轻链功能域(例如,轻链可变区(包括一个或多个CDR,框架区)和/或轻链恒定区)的变体核苷酸序列。

[0399] 在一些实施方案中,组合文库的核酸包括含有变体核酸的合适载体。在一些实施方案中,组合文库的核酸包含合适的调控序列和/或非编码序列。合适的非编码序列包括但不限于启动子、信号肽、剪接位点、终止密码子和poly(A)信号序列。在一些实施方案中,核酸包含选择标志物(例如,抗生素抗性基因、荧光分子或细胞表面标志物)。

[0400] 在一些实施方案中,筛选方法包括生成组合文库。可以用任何合适的选项来制备组合文库。在一些实施方案中,生成组合文库包括识别编码一种或多种多肽的两组以上变体氨基酸序列的两组以上变体核苷酸子序列,其中所述至少一种功能特性取决于来自两组以上变体氨基酸序列中每一组的变体氨基酸序列的组合;以及通过组合来自两组以上变体核苷酸子序列中的每一组的变体核苷酸子序列来组装连续部分,从而产生多种变体核酸的成员。

[0401] 在一些实施方案中,生成组合文库包括识别编码待由报告细胞表达的多肽的变体氨基酸序列的变体核苷酸子序列的多个变体。当多肽包括两个以上变体核苷酸子序列时,可以通过组合来自每个变体核苷酸子序列的变体来组装连续部分。在一些实施方案中,产生组合文库包括计算机模拟(in silico)设计文库的核酸,例如,通过控制密码子使用使编辑距离最大化。任何合适的算法都可以用来使编辑距离最大化。可以使用基于例如计算机生成的设计的任何合适的选项来合成组合文库。在一些实施方案中,组合文库包含来自不同T细胞群的TCR库。

[0402] 可以使用将组合文库引入细胞的任何合适的选项。合适的选项包括但不限于病毒转导、转座子介导的基因递送、转化、电穿孔、核酸酶介导的位点特异性整合(nuclease mediated site-specific integration)(如CRISPR/Cas9、TALEN)。在一些实施方案中,将组合文库引入细胞包括病毒转导、基于转座子的基因递送或核酸酶介导的位点特异性整合。

[0403] 可以将组合文库引入任何合适的细胞,例如被构造为表达由文库的核酸编码的多肽的报告细胞。合适的细胞包括但不限于哺乳动物细胞、昆虫细胞、酵母和细菌。在一些实施方案中,合适的载体包括病毒、酵母、细菌和噬菌体。尽管为了简单起见,本发明通篇使用术语“细胞”,但本文预期本文中“细胞”的所有此类内容不仅包括各种形式的T细胞(例如永生T细胞)、酵母和细菌,而且还可更一般地与任何载体(包括病毒和噬菌体)一起使用。因此,本文所用的围绕“细胞”(指可引入组合文库的细胞)的内容可包括真核细胞、原核细胞,并表示病毒和噬菌体也可用作载体的选择。细胞可以是细胞系、永生细胞或原代细胞。在一些实施方案中,细胞是人类细胞,或衍生自人类细胞。在一些实施方案中,细胞群包括永生T细胞或原代T细胞。在一些实施方案中,永生T细胞或原代T细胞是人类T细胞。在一些实施方案中,将组合文库引入永生T细胞或原代T细胞(例如,通过病毒转导)。在一些实施方案中,细胞不表现出或几乎不表现出功能特性,基于该功能特性将选择细胞以识别目的变体核苷酸子序列的组合。在一些实施方案中,细胞不表现出或几乎不表现出由文库的核酸编码的多肽介导并依赖于变体核苷酸子序列的组的功能特性。在一些实施方案中,细胞群的细胞是经工程改造的,例如,经遗传修饰的。在一些实施方案中,对细胞进行工程改造,例如遗传修饰,以减少或消除细胞对功能特性的内源表达或背景表达。在一些实施方案中,对细胞进行工程改造,例如遗传修饰,以增强细胞在用组合文库导入时表现出功能特性的能力。在一些实施方案中,对细胞进行工程改造,例如遗传修饰,以促进培养物中群的生长和/或维持。在一些实施方案中细胞群不包含赋予细胞至少一种功能特性的内源性多肽。在一些实施方案中,对细胞进行遗传修饰以引入或增强或消除或减少CD4、CD8和CD28中的一种或多种的表达。在一些实施方案中,经遗传修饰的细胞是T细胞。

[0404] 在一些实施方案中,导入组合文库的细胞群的每个细胞平均包括多种核酸中的一种核酸。在一些实施方案中,用组合文库以10以下,例如7以下、5以下、3以下、2以下,包括1以下的感染复数(multiplicity of infection, MOI)转导细胞群。在一些实施方案中,导入包括以5以下的感染复数(MOI)对细胞群进行病毒转导。在一些实施方案中,使用核酸酶介导的位点特异性整合(例如,CRISPR/Cas9,TALEN)将一种或两种核酸准确导入细胞群的每个细胞中。

[0405] 引入组合文库的细胞群的大小可以包括任何合适数量的细胞。在一些实施方案

中,细胞的数量取决于文库的大小、变体核酸在文库中的相对表征(relative representation)、群体中每种变体核酸的期望表征水平(也称为“覆盖”)中的一项或多项、所进行的筛选的类型(例如,是进行“富集”筛选(主要目标是识别富集的变体)还是进行“耗竭”筛选(主要目标是识别耗竭的变体)、文库内个体变体的表征(representation)和错误率以及基于与细胞损失(cell loss)相关的至少一种功能特性选择总细胞群的亚群所需的过程步骤。此类过程步骤的实例包括但不限于通过流式细胞术选择成功转导的细胞和/或选择由表达的多肽介导的功能特性。在一些实施方案中,筛选方法包括基于文库中两种以上变体核苷酸子序列的许多不同组合来调节细胞群的大小。

[0406] 在一些实施方案中,筛选方法包括基于包含在核酸中的选择标志物,识别已经通过接受文库的核酸而成功修饰的细胞。在一些实施方案中,筛选方法包括使用标志物在表达多种变体核酸中的至少一种的细胞群中选择或筛选细胞。在一些实施方案中,所述标记是细胞毒素抗性标志物和/或细胞表面标志物。可以使用任何合适的方法选择成功修饰的细胞,这取决于所使用的选择性标志物。在一些实施方案中,基于抗生素抗性选择细胞,例如但不限于对嘌呤霉素或杀稻瘟菌素的抗性。在一些实施方案中,基于可检测的标志物表达选择细胞,例如但不限于,通过可用于用流式细胞术分选的细胞表面标志物或荧光分子。在一些实施方案中,基于细胞表面标志物选择细胞,例如适于基于磁珠的富集。

[0407] 选择细胞群的亚群可以基于由变体核酸编码的多肽的任何合适的功能特性。合适的功能特性包括但不限于配体结合(例如,抗原结合)、响应刺激的信号转导(例如,响应抗原结合)。信号转导可以包括但不限于磷酸化、易位、信号传导结构域相互作用或转录变化。可以使用任何合适的选项进行基于依赖于变体核苷酸子序列的组合的功能特性选择细胞群的亚群。在一些实施方案中,使用但不限于标志物的表达或响应于刺激的细胞增殖来测量合适的功能输出。

[0408] 在一些实施方案中,选择包括基于可检测的标志物的表达来选择亚群,其中所述表达取决于一种或多种多肽的至少一种功能特性。在一些实施方案中,可检测的标志物包括细胞表面标志物、细胞因子标志物、细胞增殖标志物、转录报告物、信号转导报告物和/或细胞毒性报告物。在一些实施方案中,细胞表面标志物包括以下一种或多种:CD69、CD62L、CD137;细胞因子标志物包括以下一种或多种:IFN- γ 、IL-2、TNF- α 、GM-CSF;转录报告物包含以下一种或多种:NF- κ B、NFAT、AP-1;信号转导报告物包含以下一种或多种:ZAP70,ERK1/2;并且细胞毒性报告物包含以下一种或多种:CD107A、CD107B、颗粒酶B。

[0409] 在一些实施方案中,选择细胞群的亚群包括选择表现出功能特性(例如,在功能测定中阳性应答)的细胞。在一些实施方案中,选择细胞群的亚群包括选择不表现出功能特性(例如,在功能测定中阴性应答)的细胞。在一些实施方案中,筛选方法包括选择表现出功能特性的细胞群的亚群,和选择不表现出功能特性的细胞群的另一亚群。在一些实施方案中,选择细胞群的亚群包括基于每个亚群的细胞所表现的功能特性的水平或程度的层级(stratification)来选择多个细胞亚群。

[0410] 在一些实施方案中,选择亚群包括使细胞群与以下一种或多种接触:第二细胞群;一种或多种多肽的配体;一种或多种多肽的激动剂或拮抗剂;以及小分子,其中由接触诱导的亚群的变化取决于一种或多种多肽的至少一种功能特性。在一些实施方案中,选择亚群包括检测变化的存在与否,和/或变化的大小;以及基于检测选择亚群。在一些实施方案中,

第二细胞群包含抗原呈递细胞。在一些实施方案中,抗原呈递细胞包括B细胞和/或树突细胞。在一些实施方案中,第二细胞群包括原代细胞或永生化细胞。在一些实施方案中,文库的变体核苷酸子序列源自表达包含由所述变体核苷酸子序列编码的氨基酸的变体多肽的细胞,其中所述细胞获自个体,且其中所述第二细胞群源自所述个体。

[0411] 在一些实施方案中,选择包括基于高于或低于阈值的至少一种功能特性的量度来选择细胞群的第一亚群。在一些实施方案中,基于细胞群的未选亚群中的功能特性的量度来确定阈值。在一些实施方案中,选择进一步包括基于高于或低于第二阈值的至少一种功能特性的第二度量来选择细胞群的第二亚群,其中所述第一亚群和所述第二亚群不重叠。在一些实施方案中,识别至少一个组合包括比较第一亚群和第二亚群之间的至少一个组合的丰度。

[0412] 在一些实施方案中,当组合文库是包含TCR α 和TCR- β 变体的TCR文库时,细胞亚群的选择是基于细胞通过一种或多种标志物的表达变化对抗原呈递作出反应的能力。合适的标志物包括但不限于CD69、CD62L、CD137、IFN- γ 、IL-2、TNF- α 和GM-CSF。在一些实施方案中,标志物是启动子活性报告物,包括但不限于NF- κ B、NFAT和AP-1。

[0413] 在一些实施方案中,抗原由抗原呈递细胞呈递,包括但不限于B细胞(例如永生化B细胞)和树突细胞。在一些实施方案中,抗原的身份未知。在一些实施方案中,抗原是新生抗原。

[0414] 在一些实施方案中,当组合文库是含有CAR功能域变体的CAR文库时,细胞亚群的选择是基于细胞通过细胞增殖的变化和/或标志物表达的变化对抗原呈递作出反应的能力。合适的标志物包括但不限于CD69、CD62L、CD137、IFN- γ 、IL-2、TNF- α 和GM-CSF。在一些实施方案中,标志物是信号转导报告物,包括但不限于ZAP70磷酸化和ERK1/2磷酸化。在一些实施方案中,标志物是细胞毒性报告物,例如但不限于CD107A和CD107B。在一些实施方案中,标志物是启动子活性报告物,例如但不限于NF- κ B、NFAT和AP-1。

[0415] 在一些实施方案中,亚群包括多个细胞。在一些实施方案中,分离不包括基于至少一种功能特性分离亚群的单个克隆。在一些实施方案中,亚群包含至少1000个细胞(例如,100个变体的10倍覆盖率)。

[0416] 在一些实施方案中,基于依赖于变体核苷酸子序列的组的功能特性选择的亚群包括约1000个以上细胞,例如约2000个以上细胞、约3000个以上细胞、约4000个以上细胞、约5000个以上细胞、约7500个以上细胞、约10000个以上细胞、约20000个以上细胞、约50000个以上细胞、约 1×10^5 个以上细胞、约 2×10^5 个以上细胞、约 5×10^5 个以上细胞、约 1×10^6 个以上细胞、约 1×10^7 个以上细胞、约 1×10^8 个以上细胞、约 1×10^9 个以上细胞、约 1×10^{10} 个以上细胞、约 1×10^{11} 个以上细胞,包括约 1×10^{12} 个以上细胞,或由上述任何两个值限定的范围内的若干个细胞。在一些实施方案中,基于依赖于变体核苷酸子序列的组的功能特性选择的亚群包括约1000至约 1×10^{12} 个细胞、例如约2000至约 1×10^{12} 个细胞、约3000至约 1×10^{10} 个细胞、约5000至约 1×10^9 个细胞、约5000至约 1×10^9 个细胞,包括约 1×10^4 至约 1×10^9 个细胞。

[0417] 在一些实施方案中,TCR对的功能是与抗原结合。

[0418] 在一些实施方案中,亚群是初始群体的一部分,例如,小于 $10^{-14}\%$ 、 $10^{-13}\%$ 、 $10^{-12}\%$ 、 $10^{-11}\%$ 、 $10^{-10}\%$ 、 $10^{-9}\%$ 、 $10^{-8}\%$ 、0.0000001%、0.00001%、0.0001%、0.001%、

0.01%、0.1%、1%、2%、3%、4%、5%、6%、7%、8%、9%、10%、20%、30%、40%、50%、60%、70%、80%或90%的原始群体(包括任意两个前述值之间限定的任何范围)。

[0419] 在一些实施方案中,基于依赖于变体核苷酸子序列的组的功能特性选择的亚群的大小足够大,使得文库中的变体核酸在亚群中得到充分代表。在一些实施方案中,亚群的大小提供文库中的变体核酸的变体的总数的以下数值的倍数覆盖率(fold coverage):约10以上,例如约20以上、约30以上、约40以上、约50以上、约60以上、约70以上、约80以上、约90以上、约100以上、约120以上、约140以上、约160以上、约180以上、约200以上、约250以上、约300以上、约400以上、约500以上,包括约1000以上,或在由任何两个前述值限定的范围内的倍数覆盖率。在一些实施方案中,亚群的大小提供文库中的变体核酸的变体总数的以下数值的倍数覆盖率:约10至约1000,例如约20至约1000、约30至约750、约40至约500、约50至约500、约50至约400、约60至约300、约70至约250,包括约80至约200的倍数覆盖率。

[0420] 可使用任何合适的方法从亚群中分离变体核酸。在一些实施方案中,分离变体核酸包括使用任何合适的方法从亚群中提取基因组DNA。在一些实施方案中,分离变体核酸包括从亚群中提取大量基因组DNA。在一些实施方案中,分离变体核酸不包括从亚群中分离个体克隆和从个体克隆中分离基因组DNA。在一些实施方案中,分离变体核酸不包括从亚群中分离单个克隆和扩增单个克隆,从而从扩增的克隆群体中分离基因组DNA。在一些实施方案中,分离变体核酸包括从亚群中提取大量基因组DNA,而不从亚群中分离或扩增单个克隆。在一些实施方案中,分离变体核酸包括使用任何合适的方法从亚群中提取RNA。在一些实施方案中,分离变体核酸包括从亚群中提取大量RNA。在一些实施方案中,分离变体核酸包括从亚群中提取mRNA。在一些实施方案中,分离变体核酸不包括从亚群中分离个体克隆和从个体克隆中分离RNA。在一些实施方案中,分离变体核酸不包括通过单细胞PCR方法分析单细胞。在一些实施方案中,分离变体核酸不包括从亚群中分离单个克隆和扩增单个克隆,从而从扩增的克隆群体中分离RNA。在一些实施方案中,分离变体核酸包括从亚群中提取RNA,而不从亚群中分离或扩增单个克隆。术语“RNA”是一个属术语,包括例如天然和人工形式的RNA。尽管本说明书经常概述关于“DNA”的选项,但是应当理解,所有此类选项和实施方案都可以替代地用于RNA。

[0421] 在一些实施方案中,从亚群中提取的基因组DNA的量足以为文库的变体核酸中的每个变体提供充分的覆盖率。在一些实施方案中,从亚群中提取的基因组DNA的量足以提供文库中的变体核酸的变体总数的以下数值的倍数覆盖率:约10以上,例如约20以上、约30以上、约40以上、约50以上、约60以上、约70以上、约80以上、约90以上、约100以上、约120以上、约140以上、约160以上、约180以上、约200以上、约250以上、约300以上、约400以上、约500以上、包括约1000以上,或由前述任何两个值所限定的范围内的倍数覆盖率。在一些实施方案中,从亚群中提取的基因组DNA的量足以提供文库中的变体核酸的变体总数的以下数值的倍数覆盖率:约10至约1000,例如约20至约1000、约30至约750、约40至约500、约50至约500、约50至约400、约60至约300、约70至约250,包括约80至约200的倍数覆盖率。在一些实施方案中,测定包括在个体成员的任何扩增之前或不存在个体成员的任何扩增的情况下,对于连续部分的每个核苷酸序列获得至少10的平均覆盖率。

[0422] 在一些实施方案中,从亚群中分离变体核酸包括扩增提取的变体核酸。可以扩增变体核酸的任何合适的部分。在一些实施方案中,基本上仅扩增变体核酸的连续部分。在一

些实施方案中,扩增编码整个多肽的变体核酸部分。在一些实施方案中,扩增产物(或扩增子)的大小为至少600bp,例如,至少700bp、至少800bp、至少900bp、至少1000bp、至少1100bp、至少1200bp、至少1300bp、至少1400bp、至少1500bp、至少1750bp、至少2000bp、至少2500bp、至少3000bp、至少4000bp、至少5000bp、至少6000bp、至少7000bp、至少8000bp、至少9000bp、至少10000bp、至少11000bp、至少12000bp、至少13000bp、至少14000bp或至少15000bp,或在任何两个前述值限定的范围内的长度。在一些实施方案中,扩增产物(或扩增子)的大小为约600bp至约15000bp,例如约800bp至约12000bp、约1000bp至约10000bp、约1000bp至约8000bp、约1000bp至约6000bp、约1000bp至约5000bp、约1000bp至约4000bp、约1000bp至约3000bp,包括约1000bp至约2000bp。

[0423] 在一些实施方案中,测定包括至少扩增多个核酸的个体成员的连续部分。在一些实施方案中,扩增包括使用与不变核苷酸子序列杂交的扩增引物,其中多种变体核酸中的每一个均包含不变核苷酸子序列,并且其中不变核苷酸子序列编码一种或多种多肽的不变氨基酸序列。在一些实施方案中,扩增包括使用与变体核苷酸序列之外的核苷酸子序列杂交的扩增引物,所述变体核苷酸序列包括但不限于基因载体的非编码核苷酸序列。在一些实施方案中,一种或多种多肽包含TCR α -链和TCR β -链,并且其中所述不变氨基酸序列包含TCR α 恒定区。

[0424] 在一些实施方案中,进行扩增以向扩增得到的扩增子文库中的组合文库的变体核酸中的每种变体提供足够的覆盖率。在一些实施方案中,扩增分离的变体核酸以提供所得扩增子文库中组合文库的变体核酸的变体总数的以下数值的倍数覆盖率:约1000以上,例如约2000以上、约3000以上、约4000以上、约5000以上、约6000以上、约7000以上、约8000以上、约9000以上、约10000以上、约12000以上、约14000以上、约16000以上、约18000以上、约20000以上、约25000以上、约30000以上,约40000以上,约50000以上,包括约100000以上的倍数覆盖率,或在由任何两个前述值限定的范围内的倍数覆盖率。在一些实施方案中,扩增分离得到的变体核酸以提供所得扩增子文库中组合文库的变体核酸的变体总数的以下数值的倍数覆盖率:约1000至约100000,例如约2000至约100000、约3000至约75000、约4000至约50000、约5000至约50000、约5000至约40000、约6000至约30000、约7000至约25000,包括约8000至约20000的倍数覆盖率。在一些实施方案中,测定包括对于连续部分的每个核苷酸序列获得至少1000的平均覆盖率。

[0425] 在一些实施方案中,以减少所得扩增子文库中的扩增偏差(amplification bias)的方式进行扩增。在一些实施方案中,通过减少扩增循环的数量来减少扩增偏差。在一些实施方案中,通过具有减少扩增循环的数量的足够大的细胞亚群来减少扩增偏差。在一些实施方案中,使用唯一分子标识符(UMI)来减少由于扩增偏差引起的测序数据中的偏差。

[0426] 在一些实施方案中,从亚群中分离变体核酸不包括扩增分离得到的变体核酸。

[0427] 在一些实施方案中,从亚群中分离变体核酸包括使用基于CRISPR的选择性文库制备。可以使用任何适合基于CRISPR的选择性文库制备的选项。在一些实施方案中,分离包括使用来自亚群的基因组DNA的CRISPR/Cas9介导的靶向片段化。在一些实施方案中,从亚群中分离变体核酸包括将从亚群中提取的基因组DNA去磷酸化,在目的序列侧翼的位置(例如变体核酸的连续部分)引入CRISPR/Cas9介导的双链断裂,并在双链断裂处连接接头(例如测序接头)。然后可以使用任何合适的方法对接头连接的序列进行测序。在一些实施方案

中,测定子集的个体成员的连续部分的核苷酸序列包括对所述子集的每个个体成员进行条形码编码(barcoding)。

[0428] 可以使用任何合适的选项来确定子集的个体成员的连续部分的核苷酸序列。在一些实施方案中,确定连续部分的核苷酸序列涉及对连续部分进行测序。可以使用任何合适的测序平台。合适的测序平台包括但不限于桑格测序、焦磷酸测序、单分子测序、离子半导体测序、合成测序、联合探针锚定聚合测序(combinatorial probe anchor synthesis sequencing)、连接测序、单分子实时(SMRT)测序和/或纳米孔测序。在一些实施方案中,测定连续部分的核苷酸序列涉及使用允许长测序读段数的测序平台。在一些实施方案中,测定包括通过产生连续部分的至少600bp的测序读段数对个体成员进行测序。在一些实施方案中,测序读段数的长度在600bp和15000bp之间。在一些实施方案中,测定连续部分的核苷酸序列包括产生或获得至少600bp,例如至少700bp、至少800bp、至少900bp、至少1000bp、至少1100bp、至少1200bp、至少1300bp、至少1400bp、至少1500bp、至少1750bp、至少2000bp、至少2500bp、至少3000bp、至少4000bp、至少5000bp、至少6000bp、至少7000bp、至少8000bp、至少9000bp、至少10000bp、至少11000bp、至少12000bp、至少13000bp、至少14000bp、或至少15000bp,或由任意前述两个值限定的范围内的连续序列读段长度。在一些实施方案中,测序读段数为连续部分的约600bp至约1000bp、约1000bp至约2000bp、约2000bp至约3000bp、约3000bp至约4000bp、约4000bp至约5000bp、约5000bp至约7000bp、约7000bp至约10000bp,和/或约10000bp至约15000bp。

[0429] 可以使用任何合适的选项来完成基于核苷酸序列识别两种以上变体核苷酸子序列的至少一种组合。在一些实施方案中,通过确定所述组合在基于功能特性选择的亚群中富集或耗竭来识别目的组合。组合的相对丰度可以基于从亚群测定的核苷酸序列中的组合丰度与参考丰度水平的任何适当比较。合适的丰度参考水平包括但不限于基于缺乏功能特性、不同水平的应答或不同类型的应答而选择的另一个细胞亚群中的组合的丰度;或未被选择的亚群中组合的丰度。在一些实施方案中,通过确定与负向选择的亚群相比在正向选择的亚群中组合是富集或耗竭来识别目的组合。在一些实施方案中,通过确定组合在基于功能特性而选择的亚群中与组合在组合文库中的丰度相比的富集或耗竭来识别目的组合。

[0430] 在一些实施方案中,识别至少一种组合包括测量至少一种组合在亚群中相对于对照细胞群的富集。在一些实施方案中,细胞群包括对照细胞群,并且其中细胞亚群和对照细胞群不重叠。在一些实施方案中,基于不同于至少一种功能特性的第二功能特性来选择对照细胞群。

[0431] 在一些实施方案中,筛选方法包括:提供包含多种变体核酸的文库,所述多种变体核酸中的每一个包含至少600bp的连续部分,其中所述连续部分包含:第一变体核苷酸子序列,其编码TCR α 变体氨基酸序列并限定所述连续部分的第一端,和第二变体核苷酸子序列,其编码TCR β 变体氨基酸序列并限定与所述第一端相对的所述连续部分的第二端的组合;将所述文库导入经构造以表达由所述多种变体核酸的成员编码的TCR α -链和TCR β -链的永生化T细胞群中;基于响应于使永生化T细胞与表达抗原的永生化B细胞接触的高于阈值水平的T细胞活化标志物的表达选择所述永生化T细胞群的亚群,其中所述亚群包括多个T细胞;从所述亚群中分离所述多种变体核酸的子集;测定所述子集的个体成员的所述连续部分的核苷酸序列;和基于所述子集的核苷酸序列中的至少一种组合相对于对照的富集,识别所

述第一变体核苷酸子序列和第二变体核苷酸子序列的至少一种组合。在一些实施方案中,所述方法进一步包括:基于响应于使所述永生化T细胞与所述永生化B细胞接触的低于第二阈值水平的T细胞活化标志物的表达选择所述永生化T细胞群的第二亚群,其中所述第二亚群包括多个第二T细胞,并且其中所述亚群和第二亚群不重叠;从所述第二亚群中分离所述多种变体核酸的第二子集;和测定所述第二子集的个体成员的所述连续部分的第二核苷酸序列,其中基于所述子集中的所述至少一种组合相对于所述第二子集的所述第二核苷酸序列中的所述至少一种组合的富集来识别所述至少一种组合。

[0432] 在一些实施方案中,所述筛选方法包括:提供包含多种变体核酸的文库,所述多种变体核酸中的每一种均包含至少600bp的连续部分,其中所述连续部分包含以下两种以上的组合:编码CAR铰链结构域的第一变体核苷酸子序列;编码CAR跨膜结构域的第二变体核苷酸子序列;以及编码CAR细胞内信号传导结构域的第三变体核苷酸子序列,其中所述第一、第二或第三变体核苷酸子序列中的一个限定了所述连续部分的第一端,并且其中所述第一、第二或第三变体核苷酸子序列中的另一个限定了与所述第一端相对的所述连续部分的第二端;将所述文库导入经构造以表达由所述多种变体核酸的成员编码的CAR的细胞群中,其中所述细胞群包含永生化T细胞群或原代人T细胞群;响应于使所述细胞与表达对所述CAR的抗原结合结构域特异的抗原的抗原呈递细胞接触,基于高于阈值水平的细胞增殖选择所述细胞群的亚群,其中所述亚群包括多个细胞;从所述亚群中分离所述多种变体核酸的子集;测定所述子集的个体成员的所述连续部分的核苷酸序列;和基于所述子集的所述核苷酸序列中的至少一种组合相对于对照的富集,识别所述第一、第二和第三变体核苷酸子序列的至少一种组合。在一些实施方案中,所述方法进一步包括:基于响应于所述细胞与所述抗原呈递细胞的接触的低于第二阈值水平的细胞增殖选择所述细胞群的第二亚群,其中所述第二亚群包括多个第二细胞,并且其中所述亚群和第二亚群不重叠;从所述第二亚群分离所述多种变体核酸的第二亚群;测定所述第二子集的个体成员的所述连续部分的第二核苷酸序列,和其中基于所述子集中所述至少一种组合相对于所述第二子集的所述第二核苷酸序列中的所述至少一种组合的富集,识别所述至少一种组合。

[0433] 在本发明内容中,各种实施方案通常被描述为涉及包含两个以上变体核苷酸子序列的连续部分。此外,这些实施方案通常涉及第一变体核苷酸子序列和第二变体核苷酸,并且基于依赖于两个以上变体核苷酸子序列的至少一个功能特性来选择细胞群的亚群。然而,对于涉及两个以上变体核苷酸的所有此类实施方案而言,明确考虑的另一个实施方案是涉及单个功能序列(即,单个变体核苷酸)的实施方案。在此类实施方案中,600bp序列将是可编码例如具有单一功能的单一蛋白质优于单个核酸序列的序列。因此,对于本文发明描述的涉及“两个以上变体核苷酸子序列”的所有实施方案,还设想将所述方法应用于其中再次只有“变体核苷酸子序列”的情况,其中序列本身是600bp以上。例如,在一些实施方案中,提供了从核酸组合文库中识别核苷酸序列的方法。该方法包括提供包含多种变体核酸的组合文库,所述多种变体核酸中的每一个包含均至少600bp的连续部分。所述方法进一步包括将所述文库导入经构造以表达由所述多种变体核酸的成员编码的一种或多种多肽的细胞群中。该方法进一步包括基于依赖于至少600bp的连续部分的至少一个功能特性选择细胞群的亚群,其中该亚群包括多个细胞。所述方法进一步包括从所述亚群中分离所述多种变体核酸的子集。所述方法进一步包括测定该亚群的个体成员的连续部分

的核苷酸序列。所述方法进一步包括基于核苷酸序列识别至少600bp的连续部分。在一些实施方案中,所述方法也可以是其中至少600bp的连续部分分布在600个碱基对中的方法。

[0434] 提供了本筛选方法的进一步的实施方案。在一些实施方案中,本发明的筛选方法可用于任何蛋白质变体筛选,其中 (a) 蛋白质序列意图在200个氨基酸以上的连续区域中变化; (b) 蛋白质变体可在可暴露于选择性压力的报告细胞(包括酵母和细菌)或另一种功能性载体(例如病毒、噬菌体)中表达;和 (c) 可以在选择性压力(例如抗原结合、对抗原的反应的基因表达等)后的至少一种功能特性上选择表达目的蛋白变体的报告细胞。

[0435] 生成含有至少两个变体核苷酸子序列的变体核苷酸序列的文库,该变体核苷酸子序列只能通过测定其总核苷酸序列的至少600bp来明确识别

[0436] 在一些实施方案中,通过任何合适的计算机模拟设计和蛋白质工程方法产生变体核苷酸序列的文库。每种变体都可以在基因载体中编码,这将导致变体核苷酸序列在细胞内的表达。根据要进行的精确筛选,变体核苷酸序列可与适当的启动子、信号肽、剪接供体/受体、终止密码子和poly(A)信号序列组合。表达构建体还可以包含选择标志物,例如抗生素抗性基因、荧光分子或细胞表面标志物。

[0437] 在一些实施方案中,为了增强以高置信度识别变体的能力,可以使用计算机算法来控制密码子使用,从而使任何两种变体之间的编辑距离(其可以是给定核苷酸序列转化为文库中任何其它变体核苷酸序列所涉及的核苷酸变化的数目)最大化并增强区分变体核苷酸序列的能力。变体序列文库可通过任何合适的方法产生,例如但不限于DNA合成。

[0438] 在一些实施方案中,本发明的筛选方法可用于筛选变体只能通过测定200个以上的氨基酸来确定地识别任何合适的蛋白质。实例包括但不限于TCR、CAR、抗体、RNA指导的核酸酶、合成开关受体和蛋白质定向进化(directed protein evolution)。

[0439] 在一些实施方案中,文库是包含TCR α 变体和TCR β 变体的TCR文库。TCR文库可以是任何合适的文库,其中任何给定的TCR α 链和TCR β 链可以与多于一个互补二聚配偶体(complimentary dimerization partner)一起出现,或者二聚配偶体是未知的。在这两种情况下,TCR变体只能通过对TCR α 和TCR β 可变序列进行测序来明确识别。

[0440] 在一些实施方案中,文库是包含CAR变体的CAR库。CAR文库可以是任何合适的文库,其中需要对200个以上的氨基酸进行测序以可靠地识别任何给定的CAR变体。实例包括具有高度多样性的抗原结合结构域的任何CAR文库或使用一些或所有CAR蛋白结构域的任何其他组合构建物,例如其中铰链结构域、跨膜结构域和两个信号传导结构域的变体被组合以产生具有500个以上变体的文库的文库。

[0441] 在一些实施方案中,所述文库是含有抗体变体的抗体文库。变体核苷酸序列可以包括编码抗体重链和轻链的核酸,所述抗体重链和轻链在一个表达构建体中配对,所述表达构建体只能通过对重链和轻链序列进行测序来明确识别。在一些实施方案中,可以将抗体变体的文库导入报告细胞(或另一种功能载体)中,并基于至少一种功能特性(例如抗原结合)来选择抗体变体的文库。

[0442] 将变体核苷酸序列的文库导入报告细胞

[0443] 可以通过任何合适的方法将变体核苷酸序列的文库导入报告细胞。在一些实施方案中,使用逆转录病毒或慢病毒基因递送将文库导入报告细胞(或,如本文所述的任何载体)。可以使用任何合适的方法将文库导入报告细胞。报告细胞可以是任何合适的细胞(或

载体)。在一些实施方案中,可以基于几个标准选择报告细胞:(1)其表现出依赖于所引入的变体核苷酸序列和响应于外部选择压力的可测量的功能增益或功能损失的能力;(2)用于测量对选择压力的响应的标志物分子的最小背景表达,和(3)培养物中报告细胞的增殖和细胞存活。合适的报告细胞包括但不限于永生化的Jurkat T细胞或用于筛选的原代人T细胞,例如免疫受体(TCR/CAR)文库。本文还提及了其他载体的选项。

[0444] 在一些实施方案中,平均每个报告细胞仅用一个变体核苷酸序列转导,因此病毒转导是以低MOI进行的。在一些实施方案中,为了维持多克隆报告细胞池中每个变体核苷酸序列的表征(representation)水平(下文:覆盖率),待转导的报告细胞的数量与文库变体的数量直接相关。

[0445] 在一些实施方案中,在用文库修饰后,可以例如基于抗生素抗性(例如,嘌呤霉素或杀稻瘟菌素)选择成功修饰的报告细胞。此类选择可以通过在低MOI转导后从群体中消除未转导的细胞来减少总细胞数。在一些实施方案中,选择的强度取决于变体序列文库的多样性。在一些实施方案中,将文库引入更大的报告细胞群中,其中文库具有更高的多样性以维持群体中的文库覆盖率。在一些实施方案中,使用足够数量的多克隆报告细胞以在选择后维持特定水平的覆盖率。

[0446] 在一些实施方案中,可以通过转座子介导的基因递送将文库导入报告细胞。在一些实施方案中,可以通过DNA-核酸酶介导的位点特异性整合(例如,使用CRISPR/Cas9和TALEN)将文库导入报告细胞。

[0447] 在一些实施方案中,可基于通过修饰的报告细胞的细胞表面标志物的基于珠的富集来选择修饰的报告细胞。在一些实施方案中,可以通过流式细胞术分选细胞表面标记或荧光分子来选择修饰的报告细胞。

[0448] 在一些实施方案中,对报告细胞进行遗传修饰,以(1)增强其响应于外部选择压力而表现出功能增益或功能丧失的能力;(2)减少用于测量对选择压力的响应的标志物分子的背景表达;和/或(3)增强在细胞培养物中保持报告细胞群的能力。

[0449] 基于至少一种功能特性选择报告细胞

[0450] 在一些实施方案中,向报告细胞的多克隆群体施加选择性压力,以测量报告细胞依赖于表达的蛋白变体的功能增益或功能丧失。例如,可以用抗原表达细胞刺激报告细胞。随后,可以基于合适的标志物来分离对刺激作出反应的报告细胞。例如,可以使用对TCR转导的Jurka T细胞响应抗原表达细胞的CD69上调对报告细胞进行流式细胞术分选。可以以足够的覆盖率对有响应和无响应的群(population)进行分离,并分别进行分析。因此,可以通过测定阳性群体中的富集和阴性群体中的耗竭来测量给定变体核苷酸序列的相对倍数富集(relative fold enrichment)。

[0451] 在一些实施方案中,以下列方式中的一种或多种向报告细胞施加选择性压力:用受体激动剂或拮抗剂刺激;暴露于小分子;暴露于一种以上的同时刺激。在一些实施方案中,基于蛋白标志物上调或下调来分离报告细胞,例如用于标志物阳性和标志物阴性的报告细胞的流式细胞术或基于珠的分选。在一些实施方案中,基于耐药性/敏感性来分离报告细胞,其导致报告细胞的选择性存活或细胞死亡。在一些实施方案中,基于多种标志物分子来分离报告细胞。

[0452] 在一些实施方案中,仅分离一个群,而不是分离阳性和阴性响应的报告细胞。可以

通过与未暴露于选择性压力的多克隆报告细胞进行比较来建立给定的变体核苷酸序列的倍数富集。

[0453] 可以使用上述选择报告细胞的方法中的任何一种或它们的组合。从选定的报告细胞中分离变体核苷酸序列

[0454] 在一些实施方案中,为了在大规模水平上(on a bulk level)分析分离的多克隆报告细胞群,使用任何合适的方法来分离基因组DNA(gDNA)。随后,通过PCR扩增变体核苷酸序列。可以针对完整的变体核苷酸序列或仅针对其中变体表现出突变多样性的区域进行此类扩增。可以在能够提供足够的读段长度和测序读段总数的平台上制备PCR扩增子用于NGS分析,例如但不限于Oxford Nanopore技术(Oxford Nanopore technology)。在一些实施方案中,可以使用任何合适的选项(例如使用唯一分子标识符(UMI)和最少数量的PCR循环)来改进PCR方案以避免定义的变体核苷酸序列的任何有偏向扩增(biased amplification)。

[0455] 在一些实施方案中,作为防止变异核苷酸序列的偏向扩增的手段,分离来自所选报告细胞的gDNA并进行基于CRISPR的选择性文库制备。可以使基因组DNA去磷酸化,并且可以在目的序列侧翼的序列中引入CRISPR/Cas9介导的双链断裂。在此处理之后,与双链断裂直接相邻的核苷酸可以保持磷酸化,这可以允许在随后的文库制备步骤中进行磷酸化依赖的接头连接。根据前间隔序列邻近基序(Protospacer adjacent motif,PAM)的方向和所用的前间隔序列,可以使用合适的测序选项(例如,Oxford Nanopore技术)对插入物进行特异性测序。

[0456] 测定分离的变体核苷酸序列的总核苷酸序列的至少600 bp

[0457] 在一些实施方案中,使用合适的NGS平台对PCR扩增子进行测序。可以使用任何合适的测序平台对扩增子进行测序,这取决于遗传变体文库的特性。合适的测序平台包括但不限于Oxford Nanopore技术。

[0458] 至少一种目的变体核苷酸序列的选择

[0459] 在一些实施方案中,通过测定如通过两种细胞群中的变体读段计数所确定的正向选择的(标志物分子阳性)报告细胞群中的变体富集和负向选择的(标志物分子阴性)报告细胞群中的变体耗竭的相对富集来选择目的变体核苷酸序列,例如TCR和CAR。在一些实施方案中,基于以下一项或多项来选择目的变体核苷酸序列:相对富集;在不同的选择压力,例如不同的抗原剂量下发生的相对富集;在基于不同标志物分子分离的多个细胞群中的相对富集;基于标志物分子的组合分离的细胞群中的相对富集;在多个细胞群中发生的相对富集,这种细胞群由不同类型的报告细胞组成;相对于原始基因变体文库在分离的报告细胞群中富集变体核苷酸序列,例如TCR或CAR。

[0460] 在一些实施方案中,以下方案或其子部分中的任何一个可以是本文提供的实施方案的一部分或与本文提供的实施方案组合。方案编号为1-145,如下所示:

[0461] 1. 一种从不同的T细胞群恢复T细胞受体(TCR)库的方法,所述方法包括:

[0462] 测定个体样本中的TCR- α 和TCR- β 的核苷酸序列或氨基酸序列;

[0463] 从总库中选择TCR α -链和TCR β -链的序列的一个或多个子集;

[0464] 通过所选的TCR α -链和TCR β -链的序列的组合配对创建TCR库,从而创建TCR $\alpha\beta$ 对的文库;以及

[0465] 从所创建的所述TCR库中识别具有期望特征的至少一个TCR $\alpha\beta$ 对。

[0466] 2. 根据方案1所述的方法, 其中基于以下至少一个标准来选择来自所述总库的所述TCR α -链和TCR β -链的序列的一个或多个子集:

[0467] 基于所述T细胞群内的频率,

[0468] 基于与第二T细胞群相比的相对富集,

[0469] 基于给定TCR链的DNA拷贝数和RNA拷贝数的相对差异,

[0470] 基于所述TCR链的生物学特性, 其中所述特性选自以下中的至少一项: (预测的) 抗原特异性、(预测的) HLA-限制性、抗原-亲和性、共受体依赖性、亲代T细胞谱系 (例如CD4或CD8 T细胞) 或TCR序列基序,

[0471] 基于基因表达的空间模式, 其中空间基因表达模式衍生自组织中的起始区域或其他基因的共表达模式中的至少一种,

[0472] 基于在多个样本中共现或以相似频率出现, 例如在多个肿瘤病灶中出现,

[0473] 分配到多个组以分别恢复TCR库的特定部分,

[0474] 基于在不同实施方案中限定的多个标准的组合。

[0475] 3. 根据方案2所述的方法, 其中基于所述T细胞群内的频率的选择是基于TCR序列的所述频率的数据, 所述数据用于为TCR α -链和TCR β -链创建单独的秩次序, 或者为TCR α -链和TCR β -链创建组合的秩次序。

[0476] 4. 根据方案1-3中任一项所述的方法, 所述方法进一步包括测定频率阈值, 所述频率阈值是基于TCR库恢复的期望深度来限定的并且用于基于频率选择TCR α -链和TCR β -链的集合。

[0477] 5. 根据方案1-4中任一项所述的方法, 其中测定TCR- α 和TCR- β 的序列是通过以下至少一项来实现的:

[0478] 多重PCR;

[0479] 通过靶富集进行TCR序列恢复;

[0480] 通过5' RACE和PCR进行TCR序列恢复;

[0481] 通过空间测序进行TCR序列恢复;

[0482] 通过RNA-seq进行TCR序列恢复;

[0483] 以及唯一分子标识符 (UMI) 的使用。

[0484] 6. 根据方案1-5中任一项所述的方法, 其中将恢复得到的TCR-链序列定义为CDR3核苷酸序列连同足够的5' -和3' -核苷酸序列信息, 以基于核苷酸序列比对选择至少一个TCR V-片段家族和至少一个TCR J-片段家族来组装完整的TCR链序列。

[0485] 7. 根据方案1-6中任一项所述的方法, 其中通过所选的TCR α -链和TCR β -链的序列的组合配对来创建TCR库, 从而创建TCR $\alpha\beta$ 对的文库是通过以下至少一项来实现的:

[0486] 使用TCR链序列合成TCR α -链和TCR β -链的DNA或RNA片段的单独的文库, 所述TCR α -链和TCR β -链的DNA或RNA片段随后被连接成一个DNA或RNA片段, 其中恰好连接了一个TCR α -链和一个TCR β -链,

[0487] TCR α -链和TCR β -链的组合是通过直接合成其中恰好连接了一个TCR α -链和一个TCR β -链的DNA或RNA片段产生的,

[0488] TCR α -链和TCR β -链的组合是通过用单独的TCR α -基因和TCR β -基因的集合来修饰细胞池, 使得细胞将表达至少一个TCR α -链和至少一个TCR β -链, 而在细胞内产生的, 和/或

[0489] 在单链TCR构建体中使TCR α -链和TCR β -链的组合连接,其中TCR α 和TCR β 可变链片段都被融合,并且其中所述单链TCR构建体可以被融合到(i)单独的跨膜结构域或(ii)另外含有细胞内信号传导结构域,包括但不限于单独的或与CD28信号传导结构域组合的CD3 ϵ 或CD3 ζ 信号传导结构域。

[0490] 8.根据方案1-7中任一项所述的方法,其中从所述创建的TCR库中识别具有期望特征的至少一个TCR $\alpha\beta$ 对通过以下至少一项来实现:

[0491] 用生成的TCR $\alpha\beta$ 对的文库修饰的报告细胞池被呈递至少一种目的抗原的抗原呈递细胞刺激,并且基于至少一种活化标志物分离抗原反应性报告细胞以用于TCR分离;

[0492] 用适于追踪细胞增殖的荧光染料标记用生成的TCR $\alpha\beta$ 对的文库修饰的报告细胞池,所述报告细胞池由表达至少一种目的抗原的抗原呈递细胞刺激,并且基于增殖来分离抗原反应性报告细胞以用于TCR分离;

[0493] 将用生成的TCR $\alpha\beta$ 对的文库修饰的报告细胞池分成至少两个样本;样本被表达至少一种目的或非目的抗原的抗原呈递细胞刺激;刺激后,将两个报告细胞群孵育一段时间,随后通过TCR分离来分析两个报告细胞群;比较从两个样本中获得的TCR $\alpha\beta$ 对用于识别暴露于至少一种抗原的样本中具有较高丰度的TCR基因;

[0494] 用生成的TCR $\alpha\beta$ 对的文库修饰的报告细胞池被呈递至少一种目的抗原的抗原呈递细胞刺激,并且基于例如报告TCR触发的NFAT-GFP或NFAT-YFP的至少一种报道基因来分离抗原反应性报告细胞以用于TCR分离;

[0495] 用生成的TCR $\alpha\beta$ 对的文库修饰的报告细胞池被呈递至少一种目的抗原的抗原呈递细胞刺激,并且基于抗原特异性报告细胞的选择分离抗原反应性报告细胞以用于TCR分离,所述抗原特异性报告细胞的选择基于选择性存活,包括但不限于在TCR信号传导时获得的抗生素抗性,例如通过NFAT-嘌呤霉素转基因的使用;

[0496] 将用所生成的TCR $\alpha\beta$ 对的文库修饰的报告细胞池暴露于携带目的抗原的一种或多种MHC复合物;将与MHC复合物结合的报告细胞分离以用于TCR分离;

[0497] 用生成的TCR $\alpha\beta$ 对的文库修饰的报告细胞池被表达至少一种目的抗原的抗原呈递细胞刺激;随后,使用基于单细胞的液滴PCR或微流体方法识别目的TCR $\alpha\beta$ 对,以将TCR分离与至少一种活化标记的转录物水平的检测相结合;因此,检测到T细胞池中的单个报告细胞,所述单个细胞中TCR $\alpha\beta$ 转录物与增加水平的活化标志物共表达。

[0498] 9.根据方案1-8中任一项所述的方法,其中所述个体样本包含非活起始物料。

[0499] 10.根据方案1-9中任一项所述的方法,其中恢复所识别的TCR库的限定部分。

[0500] 11.根据方案1-10中任一项所述的方法,其中恢复抗原特异性TCR序列。

[0501] 12.根据方案1-11中任一项所述的方法,其中恢复I类和/或II类限制性TCR序列。

[0502] 13.根据方案1-12中任一项所述的方法,其中恢复以下至少一项:

[0503] 新生抗原特异性TCR序列

[0504] 病毒特异性TCR序列

[0505] 共有肿瘤抗原特异性TCR序列

[0506] 自身抗原特异性TCR序列。

[0507] 14.根据方案12所述的方法,所述方法进一步包括施用表达所述新生抗原特异性TCR序列的T细胞作为癌症疗法的步骤。

- [0508] 15. 根据方案1-13中任一项所述的方法, 其中所述方法用于诊断。
- [0509] 16. 根据方案15所述的方法, 其中所述诊断为从感染或自身免疫的病理部位恢复TCR库。
- [0510] 17. 根据方案1-15中任一项所述的方法, 其中所述方法用于BCR/抗体库的恢复。
- [0511] 18. 根据方案1-16中任一项所述的方法, 所述方法进一步包括从个体中分离包含所述TCR- α 和TCR- β 的核苷酸序列的核酸。
- [0512] 19. 根据方案8所述的方法, 其中所述活化标志物选自: CD25、CD69、CD62L、CD137、IFN- γ 、IL-2、TNF- α 、GM-CSF、OX40。
- [0513] 20. 根据方案1-18中任一项所述的方法, 其中DNA和RNA分离源自T细胞群, 所述T细胞群是不同细胞类型的混合物或组织样本(例如血液或肿瘤组织)的一部分。
- [0514] 21. 根据方案1-19中任一项所述的方法, 其中所述个体样本包括从体液中分离的细胞。
- [0515] 22. 根据方案20所述的方法, 其中所述细胞是肿瘤特异性T细胞或肿瘤浸润淋巴细胞。
- [0516] 23. 根据方案20-21所述的方法, 其中所述体液选自血液、尿液、血清、浆膜液、血浆、淋巴液、脑脊液、唾液、痰液、粘膜分泌物、阴道液、腹水、胸膜液、心包液、腹膜液和腹腔液。
- [0517] 24. 根据方案1所述的方法, 所述方法进一步包括使用所述TCR $\alpha\beta$ 链序列来治疗患有癌症、免疫病症、自身免疫性疾病或感染性疾病的个体。
- [0518] 25. 根据方案1-7中任一项所述的方法, 其中从所创建的TCR库中识别具有期望特征的至少一个TCR $\alpha\beta$ 对是通过以下至少一项或它们的任意组合来实现的:
- [0519] 基于至少一个活化标志物的识别或选择;
- [0520] 基于响应抗原的增殖的识别或选择;
- [0521] 基于在抗原刺激的细胞中与未刺激的细胞相比对具有更高丰度的TCR基因的识别的识别或选择;
- [0522] 基于TCR触发的报道基因激活的识别或选择;
- [0523] 基于选择性存活的识别或选择, 包括但不限于通过TCR信号传导获得的抗生素抗性;
- [0524] 基于与一种或多种MHC复合物结合的识别或选择;
- [0525] 使用基于单细胞的液滴PCR或微流体的识别或选择。
- [0526] 26. 根据方案8所述的方法, 其中通过从大量抗原反应性T细胞中分离DNA或RNA以产生TCR $\alpha\beta$ 特异性PCR产物来实现TCR分离, 所述产物通过DNA测序或RNA测序来分析以确定抗原反应性T细胞的TCR $\alpha\beta$ 基因序列, 或通过基于单细胞的液滴PCR或微流体方法来分析在所分析的单个T细胞中表达的所述TCR $\alpha\beta$ 基因序列。
- [0527] 27. 根据方案8所述的方法, 其中所述报告细胞是T细胞。
- [0528] 28. 根据方案24所述的方法, 其中使用基于单细胞的液滴PCR或微流体的识别或选择进一步包括活化相关基因的共表达的确定。
- [0529] 29. 一种创建多个T细胞文库的方法, 所述方法包括: 根据方案1所述的方法恢复T细胞受体(TCR)库; 从总库中选择TCR α -和TCR β -链序列到多个组中以分别恢复TCR库的特定

部分,其中创建复杂度较小的或恢复TCR库的特定部分的多个T细胞文库。

[0530] 30.根据方案29所述的方法,其中TCR α -链和TCR β -链的序列的选择基于频率范围。

[0531] 31.一种从核酸的组合文库中识别核苷酸序列的方法,所述方法包括:

[0532] 提供包含多种变体核酸的组合文库,所述多种变体核酸中的每一个均包含至少600bp的连续部分,其中所述连续部分包含两个以上变体核苷酸子序列的组合,其中所述两个以上变体核苷酸子序列的第一变体核苷酸子序列定义所述连续部分的第一端,并且所述两个以上变体核苷酸子序列的第二变体核苷酸子序列定义所述连续部分的与所述第一端相对的第二端;

[0533] 将所述文库导入经构造以表达由所述多种变体核酸的成员编码的一种或多种多肽的细胞群中;

[0534] 基于依赖于所述两种以上变体核苷酸子序列的组合的至少一种功能特性选择所述细胞群的亚群,其中所述亚群包括多个细胞;

[0535] 从所述亚群中分离所述多种变体核酸的子集;

[0536] 测定所述子集的个体成员的所述连续部分的核苷酸序列;以及

[0537] 基于所述核苷酸序列识别所述两个以上变体核苷酸子序列的至少一种组合。

[0538] 32.根据方案1所述的方法,其中所述一种或多种多肽包括:

[0539] T细胞受体 α (TCR α) -链和TCR β -链;

[0540] 嵌合抗原受体 (CAR) ;

[0541] 开关受体;或

[0542] 抗体或其抗原结合片段的一条或多条链。

[0543] 33.根据方案31或32所述的方法,其中所述第一变体核苷酸子序列编码TCR α 变体氨基酸序列,并且所述第二变体核苷酸子序列编码TCR β 变体氨基酸序列。

[0544] 34.根据方案31至33中任一项所述的方法,其中所述两种以上变体核苷酸子序列编码以下中的一种或多种:TCR V区、TCR互补决定区3 (CDR3)、TCR J-片段和TCR恒定区。

[0545] 35.根据方案31或32所述的方法,其中所述两个以上变体核苷酸子序列各自编码CAR的抗原结合结构域、铰链结构域、跨膜结构域或一个或多个细胞内信号传导结构域。

[0546] 36.根据方案31-35中任一项所述的方法,其中所述连续部分的长度为600bp至15000bp。

[0547] 37.根据方案31-36中任一项所述的方法,其中提供包括生成所述文库。

[0548] 38.根据方案37所述的方法,其中生成所述文库包括:

[0549] 识别编码所述一种或多种多肽的两组以上变体氨基酸序列的两组以上变体核苷酸子序列,其中所述至少一种功能特性取决于来自两组以上变体氨基酸序列中每一组的变体氨基酸序列的组合;以及

[0550] 通过组合来自两组以上变体核苷酸子序列中每一组的变体核苷酸子序列来组装所述连续部分,从而产生所述多种变体核酸的成员。

[0551] 39.根据方案31-38中任一项所述的方法,其中使所述文库中所述多种变体核酸的连续部分之间的编辑距离最大化。

[0552] 40.根据方案31-39中任一项所述的方法,其中所述文库包含所述两种以上变体核苷酸子序列的至少100种不同组合。

- [0553] 41. 根据方案31-40中任一项所述的方法,其中导入包括通过病毒转导、基于转座子的基因递送或核酸酶介导的位点特异性整合导入所述文库。
- [0554] 42. 根据方案41所述的方法,其中导入包括以5以下的感染复数(MOI)对所述细胞群进行病毒转导。
- [0555] 43. 根据方案31-42中任一项所述的方法,所述方法包括基于所述文库中两种以上变体核苷酸子序列的若干不同组合来调节所述细胞群的大小。
- [0556] 44. 根据方案31-43中任一项所述的方法,其中所述细胞群包含永生化T细胞或原代T细胞。
- [0557] 45. 根据方案44所述的方法,其中所述永生化T细胞或原代T细胞为人T细胞。
- [0558] 46. 根据方案31-45中任一项所述的方法,所述方法进一步包括使用标志物在表达至少一种所述多种变体核酸的所述细胞群中选择或筛选细胞。
- [0559] 47. 根据方案46所述的方法,其中所述标志物是细胞毒素抗性标志物和/或细胞表面标志物。
- [0560] 48. 根据方案31-47中任一项所述的方法,其中所述细胞群的细胞被遗传修饰。
- [0561] 49. 根据方案31-47中任一项所述的方法,其中所述群的细胞不包含赋予所述细胞所述至少一种功能特性的内源多肽。
- [0562] 50. 根据方案1至49中任一项所述的方法,其中所述细胞经遗传修饰以消除或减少CD4、CD8和CD28中的一种或多种的表达。
- [0563] 51. 根据方案48或49所述的方法,其中用CD4和/或CD8重建所述细胞并被用于筛选I类和/或II类限制性TCR序列。
- [0564] 52. 根据方案50或51所述的方法,其中所述细胞为T细胞。
- [0565] 53. 根据方案31-52中任一项所述的方法,其中选择包括基于可检测标志物的表达选择所述亚群,其中所述表达取决于所述一种或多种多肽的所述至少一种功能特性。
- [0566] 54. 根据方案53所述的方法,其中所述可检测标志物包括细胞表面标志物、细胞因子标志物、细胞增殖标志物、转录报告物、信号转导报告物和/或细胞毒性报告物。
- [0567] 55. 根据方案54所述的方法,其中
- [0568] 所述细胞表面标志物包括:CD25、CD69、CD62L、CD137、OX40中的一种或多种;
- [0569] 所述细胞因子标志物包括:IFN- γ 、IL-2、TNF- α 、IL-8、GM-CSF中的一种或多种;
- [0570] 所述转录报告物包括:NF- κ B、NFAT、AP-1中的一种或多种;
- [0571] 所述信号转导报告物包括:ZAP70、ERK1/2中的一种或多种;以及
- [0572] 所述细胞毒性报告物包括:CD107A、CD107B中的一种或多种。
- [0573] 56. 根据方案31-55中任一项所述的方法,其中选择包括使所述细胞群与以下所述中的一种或多种接触:
- [0574] 第二细胞群;
- [0575] 所述一种或多种多肽的配体;
- [0576] 所述一种或多种多肽的激动剂或拮抗剂;以及
- [0577] 小分子,
- [0578] 其中所述接触诱导的所述亚群的变化取决于所述一种或多种多肽的所述至少一种功能特性。

- [0579] 57. 根据方案56所述的方法, 其中选择进一步包括:
- [0580] 检测所述变化的存在与否, 和/或所述变化的大小; 以及
- [0581] 基于所述检测选择所述亚群。
- [0582] 58. 根据方案56或57所述的方法, 其中所述第二细胞群包含抗原呈递细胞。
- [0583] 59. 根据方案58所述的方法, 其中所述抗原呈递细胞包括B-细胞和/或树突细胞。
- [0584] 60. 根据方案56-59中任一项所述的方法, 其中所述第二细胞群包含原代细胞或永生化细胞。
- [0585] 61. 根据方案56-60中任一项所述的方法, 其中所述文库的变体核苷酸序列源自表达包含由所述变体核苷酸子序列编码的氨基酸的变体多肽的细胞, 其中所述细胞来源于个体, 并且其中所述第二细胞群源自所述个体。
- [0586] 62. 根据方案31-61中任一项所述的方法, 其中选择包括基于高于或低于阈值的所述至少一种功能特性的量度选择所述细胞群的第一亚群。
- [0587] 63. 根据方案62所述的方法, 其中基于对所述细胞群的未选亚群中的所述功能特性的量度来确定所述阈值。
- [0588] 64. 根据方案62所述的方法, 其中选择进一步包括基于高于或低于第二阈值的所述至少一种功能特性的第二量度来选择所述细胞群的第二亚群, 其中所述第一亚群和第二亚群不重叠。
- [0589] 65. 根据方案64所述的方法, 其中识别所述至少一个组合包括比较所述第一亚群和第二亚群之间的所述至少一个组合的丰度。
- [0590] 66. 根据方案31-64中任一项所述的方法, 其中识别所述至少一种组合包括测量所述至少一种组合在所述亚群中相对于对照细胞群的富集。
- [0591] 67. 根据方案66所述的方法, 其中所述细胞群包括所述对照细胞群。
- [0592] 68. 根据方案67所述的方法, 其中所述细胞亚群和对照细胞群不重叠, 其中不重叠表示两个群中的所述细胞具有不同的活化状态, 但可以携带相同的变体核酸。
- [0593] 69. 根据方案66所述的方法, 其中基于不同于所述至少一种功能特性的第二功能特性来选择所述对照细胞群。
- [0594] 70. 根据方案31-69中任一项所述的方法, 其中所述亚群包含多个细胞。
- [0595] 71. 根据方案31-70中任一项所述的方法, 其中所述分离不包括基于所述至少一种功能特性来分离所述亚群的单个克隆。
- [0596] 72. 根据方案31-71中任一项所述的方法, 其中所述亚群包含至少1000个细胞。
- [0597] 73. 根据方案31-72中任一项所述的方法, 其中测定包括通过生成连续部分的至少600bp的测序读段数来对所述个体成员进行测序。
- [0598] 74. 根据方案73所述的方法, 其中所述测序读段数长度在600bp与15000bp之间。
- [0599] 75. 根据方案31-74中任一项所述的方法, 其中测定包括至少扩增所述个体成员的所述连续部分。
- [0600] 76. 根据方案75所述的方法, 其中扩增包括使用与不变核苷酸子序列杂交的扩增引物, 其中所述多种变体核酸中的每一个均包含所述不变核苷酸子序列, 并且其中所述不变核苷酸子序列编码所述一种或多种多肽的不变氨基酸序列。
- [0601] 77. 根据方案76所述的方法, 其中所述一种或多种多肽包含TCR α -链和TCR β -链, 并

且其中所述不变氨基酸序列包含TCR α 恒定区。

[0602] 78. 根据方案76所述的方法, 其中所述一种或多种多肽包含TCR α -链和TCR β -链, 并且其中所述不变氨基酸序列包含TCR β 恒定区。

[0603] 79. 根据方案31-78中任一项所述的方法, 其中所述分离包括使用CRISPR/Cas9介导的来自所述亚群的基因组DNA的靶向片段化。

[0604] 80. 根据方案31-78中任一项所述的方法, 其中所述测定包括在所述个体成员的任何扩增之前或不存在所述个体成员的任何扩增的情况下, 对于所述连续部分的每个核苷酸序列获得至少10的平均覆盖率。

[0605] 81. 根据方案31-80中任一项所述的方法, 其中所述测定包括对于所述连续部分的每个核苷酸序列, 获得至少25、至少50、至少100、至少200、至少300、至少400、至少500或至少1000的平均覆盖率。

[0606] 82. 一种从核酸组合文库中识别编码T细胞受体 α (TCR α) -链和TCR β -链的核苷酸序列的方法, 其包括:

[0607] 提供包含多种变体核酸的文库, 所述多种变体核酸中的每一个均包含至少600bp的连续部分, 其中所述连续部分包含:

[0608] 以下的组合:

[0609] 第一变体核苷酸子序列, 其编码TCR α 变体氨基酸序列并限定所述连续部分的第一端, 和

[0610] 第二变体核苷酸子序列, 其编码TCR β 变体氨基酸序列并限定与所述第一端相对的所述连续部分的第二端;

[0611] 将所述文库导入经构造以表达由所述多种变体核酸的成员编码的TCR α -链和TCR β -链的永生化的T细胞群中;

[0612] 基于响应于使永生化的T细胞与表达抗原的永生化的B细胞接触的高于阈值水平的T细胞活化标志物的表达选择所述永生化的T细胞群的亚群, 其中所述亚群包括多个T细胞;

[0613] 从所述亚群中分离所述多种变体核酸的子集;

[0614] 测定所述子集的个体成员的所述连续部分的核苷酸序列; 和

[0615] 基于所述子集的核苷酸序列中的至少一种组合相对于对照的富集, 识别所述第一变体核苷酸子序列和第二变体核苷酸子序列的至少一种组合。

[0616] 83. 根据方案82所述的方法, 其进一步包括:

[0617] 基于响应于使所述永生化的T细胞与所述永生化的B细胞接触的低于第二阈值水平的T细胞活化标志物的表达选择所述永生化的T细胞群的第二亚群, 其中所述第二亚群包括多个第二T细胞, 并且其中所述亚群和第二亚群不重叠;

[0618] 从所述第二亚群中分离所述多种变体核酸的第二子集; 和

[0619] 测定所述第二子集的个体成员的所述连续部分的第二核苷酸序列,

[0620] 其中基于所述子集中的所述至少一种组合相对于所述第二子集的所述第二核苷酸序列中的所述至少一种组合的富集来识别所述至少一种组合。

[0621] 84. 一种从核酸组合文库中识别编码嵌合抗原受体 (CAR) 铰链结构域、跨膜结构域和/或细胞内信号传导结构域的核苷酸序列的方法, 其包括:

[0622] 提供包含多种变体核酸的文库, 所述多种变体核酸中的每一个均包含至少600bp

的连续部分,其中所述连续部分包含以下两种以上的组合:

[0623] 编码CAR铰链结构域的第一变体核苷酸子序列;

[0624] 编码CAR跨膜结构域的第二变体核苷酸子序列;和

[0625] 编码CAR细胞内信号传导结构域的第三变体核苷酸子序列,

[0626] 其中所述第一、第二或第三变体核苷酸子序列中的一个限定了所述连续部分的第一端,并且其中所述第一、第二或第三变体核苷酸子序列中的另一个限定了与所述第一端相对的所述连续部分的第二端;

[0627] 将所述文库导入经构造以表达由所述多种变体核酸的成员编码的CAR的细胞群中,其中所述细胞群包含永生化T细胞群或原代人T细胞群;

[0628] 响应于使所述细胞与表达对所述CAR的抗原结合结构域特异的抗原的抗原呈递细胞接触,基于高于阈值水平的细胞增殖选择所述细胞群的亚群,其中所述亚群包括多个细胞;

[0629] 从所述亚群中分离所述多种变体核酸的子集;

[0630] 测定所述子集的个体成员的所述连续部分的核苷酸序列;和

[0631] 基于所述子集的所述核苷酸序列中的至少一种组合相对于对照的富集,识别所述第一、第二和第三变体核苷酸子序列的至少一种组合。

[0632] 85.根据方案84所述的方法,其中存在多于一个的CAR细胞内信号传导结构域。

[0633] 86.根据方案85所述的方法,其中存在至少两个CAR细胞内信号传导结构域。

[0634] 87.根据方案84所述的方法,其中所述至少两个CAR细胞内信号传导结构域是以下中的至少两个:CD3 ϵ 、CD3 ζ ITAM1、CD3 ζ ITAM12、CD3 ζ ITAM 123、CD3 ζ 与CD3 δ 、CD3 ϵ 和CD3 γ 的任意ITAM、CD8 α 、CD28、ICOS、4-1BB (CD137)、OX40 (CD134)、CD27和CD2。

[0635] 88.根据方案84所述的方法,其包括:

[0636] 基于响应于所述细胞与所述抗原呈递细胞的接触的低于第二阈值水平的细胞增殖选择所述细胞群的第二亚群,其中所述第二亚群包括多个第二细胞,并且其中所述亚群和第二亚群不重叠;

[0637] 从所述第二亚群分离所述多种变体核酸的第二子集;

[0638] 测定所述第二子集的个体成员的所述连续部分的第二核苷酸序列,和

[0639] 其中基于所述子集中所述至少一种组合相对于所述第二子集的所述第二核苷酸序列中的所述至少一种组合的富集,识别所述至少一种组合。

[0640] 89.一种从核酸组合文库中识别核苷酸序列的方法,其包括:

[0641] 提供包含多种变体核酸的组合文库,所述多种变体核酸中的每一个均包含至少600bp的连续部分;

[0642] 将所述文库导入经构造以表达由所述多种变体核酸的成员编码的一种或多种多肽的细胞群中;

[0643] 基于依赖于所述至少600bp的连续部分的至少一个功能特性来选择所述细胞群的亚群,其中所述亚群包括多个细胞;

[0644] 从所述亚群中分离所述多种变体核酸的子集;

[0645] 测定所述子集的个体成员的所述连续部分的核苷酸序列;和

[0646] 基于所述核苷酸序列识别所述至少600bp的连续部分。

[0647] 90. 根据方案89所述的方法, 其中所述至少600bp的连续部分中的可变性分布在600个碱基对中。

[0648] 91. 一种从核酸文库中识别编码抗原特异性T细胞受体 α (TCR α) -链和TCR β -链对的核苷酸序列的方法, 其包括:

[0649] 将文库导入能够表达由多种变体核酸的成员编码的TCR α -链和TCR β -链的细胞群中,

[0650] 基于响应于抗原的高于阈值水平的标志物的表达来选择所述细胞群的亚群, 其中所述亚群包括多个细胞,

[0651] 从所述亚群中分离所述多种变体核酸的子集,

[0652] 测定所述变体核酸的核苷酸序列, 和

[0653] 基于所述子集内的所述核苷酸序列相对于对照的富集来识别至少一种变体核苷酸序列。

[0654] 92. 根据方案91或90所述的方法, 其进一步包括提供包含编码TCR α 链和TCR β 链的所述多种变体核酸的文库。

[0655] 93. 一种从样本中识别编码T细胞受体 α (TCR α) -链和TCR β -链的核苷酸序列的方法, 其包括:

[0656] 对样本中的TCR- α 链和TCR- β 链进行测序,

[0657] 选择并对TCR α -链和TCR β -链序列进行组合配对以创建TCR $\alpha\beta$ 对的文库,

[0658] 将所述TCR $\alpha\beta$ 对的文库导入报告细胞池,

[0659] 用呈递至少一种目的抗原的抗原呈递细胞刺激用所述TCR $\alpha\beta$ 对的文库修饰的所述报告细胞,

[0660] 确定对所述至少一种目的抗原特异的TCR $\alpha\beta$ 对, 和

[0661] 将所述TCR $\alpha\beta$ 对导入细胞并选择含有所述TCR $\alpha\beta$ 对的细胞。

[0662] 94. 根据方案14所述的方法, 其进一步包括施用表达所述抗原特异性TCR序列的T细胞以诊断或治疗感染或自身免疫的步骤。

[0663] 95. 根据方案14所述的方法, 其中所述T细胞可以是自体的或同种异体的。

[0664] 96. 根据方案8所述的方法, 其中所述活化标志物是CD69, 并且其中分离了两个细胞群, 一个细胞群具有高表达的CD69并且另一个细胞群具有低表达的CD69。

[0665] 97. 一种核苷酸文库, 其包含根据方案1-30和95、96中任一项所述恢复得到的T细胞受体库。

[0666] 98. 一种核苷酸构建体, 其包含根据方案1-96中任一项所述识别得到的核苷酸序列。

[0667] 99. 一种细胞, 其包含根据方案98所述的核苷酸构建体。

[0668] 100. 一种从核酸文库中识别编码抗原特异性T细胞受体 α (TCR α) -链和TCR β -链对的核苷酸序列的方法, 所述方法包括:

[0669] 将所述核酸文库导入能够表达TCR α -链和TCR β -链的细胞群中, 以制备细胞文库;

[0670] 基于响应于抗原的高于第一阈值水平的标志物的表达来选择所述细胞文库的第一群; 和

[0671] 从所述文库的所述第一群中分离变体核酸的第一群。

- [0672] 101. 根据方案100所述的方法, 其进一步包括:
- [0673] 测定变体核酸的所述第一群的至少一个核苷酸序列或核酸同一性; 和
- [0674] 基于所述子集内的所述核苷酸序列相对于对照的富集来识别至少一种变体核苷酸序列。
- [0675] 102. 根据方案100或101所述的方法, 其中所述阈值水平基于以下至少一项:
- [0676] d) 基于标志物的表达的所述总的细胞池的百分比的恢复; 或者
- [0677] e) 来自所述总的细胞池的最少数量的细胞的恢复; 或者
- [0678] f) 基于磁性探针与响应抗原表达的至少一种标志物的结合, 由磁体保留的细胞的恢复。
- [0679] 103. 根据方案66-69、82、84、91或101所述的方法, 其中所述对照是低于第二阈值的第二细胞群。
- [0680] 104. 根据方案66-69、82、84、91或101所述的方法, 其中所述对照是以下中的一种或多种:
- [0681] 参考细胞群,
- [0682] 导入所述细胞群的所述核酸组合文库,
- [0683] 基于低于第二阈值的表达标志物从与第一群相同的细胞群中分选出来的细胞群,
- [0684] 从表达相关TCR文库的报告T细胞与不呈递任何外源性抗原的抗原呈递细胞如B细胞的共培养物中获得的至少一种细胞群。
- [0685] 105. 根据方案104所述的方法, 其中所述对照 (或底部样本) 选自与顶部样本相同的细胞群, 但具有低活化标志物表达, 或其中所述底部样本获自表达所述相关TCR文库的报告T细胞和未被工程化以表达外源抗原的B细胞的共培养物。
- [0686] 106. 根据方案100-105中任一项所述的方法, 其进一步包括向所述细胞群添加抗原。
- [0687] 107. 根据方案100-106中任一项所述的方法, 其中所述分离第一群和/或所述对照通过以下中的至少一项来实现: a) 磁珠富集, b) 流式细胞术分选或c) 二者。
- [0688] 108. 一种从核酸文库中识别编码T细胞受体 α (TCR α) -链和TCR β -链的核苷酸序列的方法, 所述方法包括:
- [0689] 将所述核酸文库导入能够表达TCR α -链和TCR β -链的细胞群中以制备细胞文库; 和
- [0690] 基于所述子集内的所述核苷酸序列相对于对照的富集, 测定变体核酸的第一群的至少一个核苷酸序列或至少一种核酸的同一性。
- [0691] 109. 根据方案108所述的方法, 其中所述至少一种核酸从第一细胞群分离。
- [0692] 110. 根据方案109所述的方法, 其中所述第一细胞群基于响应于抗原的高于第一阈值水平的标志物的表达来选择。
- [0693] 111. 一种从核酸文库中识别核苷酸序列的方法, 其包括:
- [0694] 将所述核酸文库导入细胞群以形成细胞文库;
- [0695] 使所述细胞文库与第一细胞群接触;
- [0696] 通过磁珠富集基于至少一种标志物的表达来选择所述细胞文库的亚群; 和
- [0697] 基于所述亚群内的所述核苷酸序列相对于对照的统计学上显著的富集或耗竭来识别至少一种核苷酸序列。

- [0698] 112. 根据方案111所述的方法, 其中所述细胞亚群中的至少一些被构造为表达由所述核酸文库的成员编码的一种或多种多肽。
- [0699] 113. 根据方案111所述的方法, 其中标志物表达与来自所述文库的导入核酸相关联, 其中相关联表示导入的核酸改变标志物表达。
- [0700] 114. 根据方案111所述的方法, 其中选择基于高于阈值水平的标志物表达。
- [0701] 115. 根据方案111所述的方法, 其中选择基于高于阈值水平的至少两个标志物的表达。
- [0702] 116. 根据方案8、82、91、93或100中任一项所述的方法, 其中识别或刺激或提供抗原包括以下中的一种或多种:
- [0703] a) 选择若干抗原;
- [0704] b) 创建抗原池, 其中每个抗原正好存在于两个抗原池中;
- [0705] c) 评估表达至少一种T细胞受体的报告细胞对每个抗原池的反应性; 和
- [0706] d) 通过评估对恰好两个抗原池的反应性, 确定所述至少一种T细胞受体是否对任何选定的抗原具有反应性。
- [0707] 117. 根据方案116所述的方法, 其中通过成对富集分析来检测对恰好两个抗原池的反应性。
- [0708] 118. 根据方案111所述的方法, 其中所述文库是TCR文库。
- [0709] 119. 根据方案111所述的方法, 其中使用活化标志物。
- [0710] 120. 根据方案111所述的方法, 其中使用顶部-底部比较来评估反应性。
- [0711] 121. 根据前述方案中任一项所述的方法涉及反应性评估或进一步包括反应性评估, 其中使用顶部-底部比较来评估反应性。
- [0712] 122. 根据前述方案中任一项所述的方法涉及文库, 其中基于以下中的至少一项来对所述文库中的所述多种变体核酸的所述核苷酸序列进行优化:
- [0713] 为宿主细胞引入优选的密码子使用,
- [0714] mRNA结构稳定性的优化,
- [0715] 重复序列的避免,
- [0716] 均聚物的长拉伸的避免, 和
- [0717] 给定变体核酸序列内局部GC含量的较大差异的避免。
- [0718] 123. 根据方案91所述的方法, 其中所述抗原通过抗原呈递细胞呈递。
- [0719] 124. 根据方案91所述的方法, 其中所述文库是组合文库。
- [0720] 125. 根据前述方案中任一项所述的方法, 其中所述抗原由细胞提供。
- [0721] 126. 根据前述方案中任一项所述的方法, 其中过程涉及高度的抗原多样性和/或复杂性。
- [0722] 127. 前述方法中任一项涉及文库的方法, 其中所述文库是组合文库。
- [0723] 128. 如上所述的方法, 其中所述组合文库是TCR文库。
- [0724] 129. 一种细胞集合, 所述集合包括:
- [0725] 至少两种T细胞的组, 其中每种T细胞均被构造为表达至少一个TCR α 和TCR β 对, 其中所述TCR α 和TCR β 各自来自一个个体, 其中所述T细胞不表达内源性TCR, 并且其中所述组被构造用于一个或多个T细胞活化标志物的激活; 和

[0726] 至少两种B细胞的组,其中所述至少两种B细胞中的每一种均被构造为表达至少一种外源性新生抗原(或抗原),使得存在至少两种能够被产生的外源性新生抗原(或抗原),并且其中所述至少两种外源性新生抗原(或抗原)与所述个体中的外源性新生抗原(或抗原)相同。

[0727] 130.根据方案129所述的组合物,其中所述至少两种B细胞的组包括:

[0728] 至少一种产生所述外源性新生抗原(或抗原)的第一B细胞;和

[0729] 至少一种产生所述第二外源性新生抗原(或抗原)的第二B细胞。

[0730] 131.一种表达TCR的细胞的文库,所述表达TCR的细胞的文库包含:至少三个种细胞的组,

[0731] 其中所述T细胞中的至少两种被构造为表达至少两个TCR α 和TCR β 对(至少两个TCR对),

[0732] 其中所述至少两个TCR对来自一个个体,

[0733] 其中所述至少三种T细胞不表达内源性TCR,

[0734] 其中所述至少三种T细胞被构造用于在结合由B细胞呈递的抗原(或新生抗原)时一个或多个T细胞活化标志物的激活,

[0735] 其中反映在若干个TCR细胞中的每个TCR对的基因组拷贝的数量使得可以读取样本中的每个TCR,并且

[0736] 其中至少一种所述TCR在包含所述文库的整个组合物中分布不均。

[0737] 132.根据方案131所述的表达TCR的细胞的文库,其中至少一种T细胞的分布通过与B细胞呈递的抗原结合而改变。

[0738] 133.根据方案131所述的表达TCR的细胞的文库,其中所述至少两个TCR对近似均匀地存在于所述表达TCR的细胞的文库中。

[0739] 134.一种治疗个体的方法,所述方法包括:

[0740] 识别患有肿瘤的个体;

[0741] 提供至少两种T细胞的组,每种T细胞被构造为表达至少一种不同的TCR α 和TCR β 对,其中所述TCR α 和所述TCR β 各自来自所述个体,

[0742] 提供至少两种B细胞的组,其中所述B细胞的组经构造以表达至少两种外源性新生抗原,并且其中所述至少两种外源性新生抗原与在所述个体中发现的那些新生抗原相同;

[0743] 将所述至少两种T细胞的组与所述至少两种B细胞的组组合,并基于所述至少两种T细胞经由所述至少两个外源性新生抗原的激活来选择至少两个TCR对的组合;和

[0744] 向所述个体施用所述至少两个TCR对的组合,从而治疗所述肿瘤。

[0745] 135.根据方案134所述的方法,其中治疗减小了所述肿瘤的大小。

[0746] 136.一种治疗个体的方法,所述方法包括:

[0747] 识别患有肿瘤的个体;

[0748] 提供至少两种T细胞的组,每种T细胞均被构造为表达至少一种不同的TCR α 和TCR β 对,其中所述TCR α 和所述TCR β 各自来自所述个体,

[0749] 提供至少两种抗原呈递细胞的组,其中所述抗原呈递细胞的组源自所述个体,被构造为表达至少两种外源性新生抗原,并且其中所述至少两种外源性新生抗原与在所述个体中发现的那些新生抗原相同;

[0750] 将所述至少两种T细胞的组与所述至少两种抗原呈递细胞的组组合,并基于所述至少两种T细胞经由所述至少两个外源性新生抗原的激活来选择至少两个TCR对的组合;和

[0751] 向所述个体施用所述至少两个TCR对的组合,从而治疗所述肿瘤。

[0752] 137.一种药物组合物,其包含:

[0753] 第一TCR对,其与个体的肿瘤中的第一抗原(或新生抗原)结合;和

[0754] 第二TCR对,其与所述个体的肿瘤中的第二抗原(或新生抗原)结合。

[0755] 138.根据方案137所述的药物组合物,其中所述第一TCR对是MHC I类限制性的,并且其中所述第二TCR对是MHC II类限制性的。

[0756] 139.一种药物组合物,其包含:

[0757] 第一TCR对,其与第一抗原结合并且是MHC I类限制性的;和

[0758] 第二TCR对,其与第二抗原结合并且是MHC II类限制性的。

[0759] 140.根据方案137-139中任一项所述的药物组合物,其进一步包含第三TCR对。

[0760] 141.根据方案137-140中任一项所述的药物组合物,其中所述第一TCR对与来自肿瘤的新生抗原结合,其中所述第二TCR对与来自所述肿瘤的新生抗原结合,并且其中所述第一TCR对和第二TCR对二者均存在于所述肿瘤的宿主中。

[0761] 142.一种细胞集合,所述集合包括:

[0762] 至少两种T细胞的组,其中每种T细胞均被构造为表达至少一个TCR α 和TCR β 对,其中所述对来自一个个体,其中所述T细胞不表达内源性TCR,并且其中所述组被构造用于激活一个或多个T细胞活化标志物;和

[0763] 至少两种抗原呈递细胞(APC)的组,其中所述至少两种APC中的每一种均被构造为表达至少一种外源性新生抗原(或抗原),使得存在至少两种能够被产生的外源性新生抗原(或抗原),并且其中所述至少两种外源性新生抗原(或抗原)与所述个体中的外源性新生抗原(或抗原)相同。

[0764] 143.如上述方法中任一种所述的方法,其中,通过与抗原(例如新生抗原)结合来选择或识别所述TCR对和/或表达所述TCR对的所述T细胞,其中所述抗原由B细胞或抗原呈递细胞表达。

[0765] 144.如上述方法中任一种所述的方法,其中所述抗原或新生抗原来自一个个体中的肿瘤,并且其中所述TCR对的TCR α 和TCR β 也各自来自所述个体。

[0766] 145.如上述方法中任一项所述的方法,其中存在至少2个、3个、4个、5个、6个、7个、8个、9个、10个、50个、100个、500个、1000个、10000个、100000个或100万个TCR对(或包括这些对的细胞),并且存在至少2个、3个、4个、5个、6个、7个、8个、9个、10个、50个、100个、500个、1000个、10000个、100000个或100万个抗原。

[0767] 附加实施方案

[0768] 各种方法的各种实施方案也呈现在图10A-10J中。在一些实施方案中,图10A-10J中的图示出了,从错配型结直肠癌(mismatch-proficient colorectal cancer,MMRp-CRC)的肿瘤中分离新生抗原反应性TCR的方法,作为其他肿瘤的示例,该方法是可能的。这些肿瘤通常具有较低的突变负荷(mutational burden)。有充分证据表明,突变负荷与免疫肿瘤治疗(immuno-oncology therapy)(例如抗PD1/PD-L1检查点抑制疗法)的应答有很好的相关性。实际上,在接受免疫检查点抑制剂治疗的MMRp-CRC队列中观察到了不良反应率(poor

response rate)。

[0769] 与免疫检查点抑制剂疗法类似,T细胞疗法是通过免疫系统对肿瘤细胞的识别来驱动的。鉴于MMRp-CRC中的低突变负荷,预计会发现T细胞对肿瘤新生抗原的反应性有限。出乎意料地,在筛选的所有四份MMRp-CRC患者样本中发现(如下文所述和以下实施例13所示)识别肿瘤新生抗原的TCR,从而能够对该患者组使用TCR T细胞疗法(否则治疗选择将会是有限的)。

[0770] 在一些实施方案中,该方法可以涉及以下概述为过程10A-10J概述的一个或多个步骤(以及一些所述实施方案的附图)。任何步骤均可酌情被本文提供的其他实施方案重复或替代。也可以添加其他的中间步骤。

[0771] 步骤10A) 筛选过程示意图。在一些实施方案中,错配型结直肠癌(例如,MMRp-CRC)患者样本接受TCR识别平台,从常规非存活肿瘤活检获得遗传信息开始。从肿瘤浸润淋巴细胞(TIL)中提取大量TCR测序信息,并用于组装 α 链和 β 链表达盒的组合配对文库。这些表达盒在Jurkat报告T细胞中表达,并针对以串联小基因(TMG)形式将肿瘤新生抗原表达为小基因的自体B细胞进行筛选。通过回收活化的报告T细胞并分离它们的TCR,可以识别新生抗原反应性TCR。

[0772] 步骤10B) 人MMRp-CRC样本中浸润淋巴细胞的批量TCR测序。在一些实施方案中,可以对来自10A的产物进行TCR的批量测序。在一些实施方案中,通过Milaboratory对人MMRp-CRC肿瘤样本pt1进行批量TCR测序。经过比对和TCR识别后,克隆型根据它们的CDR3氨基酸序列及它们的V和J同一性收缩。唯一克隆型(unique clonotype)的数量代表 α 链和 β 链二者。结果示于图10B的图中。

[0773] 步骤10C) TCR文库的质量控制。在一些实施方案中,可以采用可选的质量控制检查。在一些实施方案中,选择来自10A)中的肿瘤样本的100条最普遍的 α 链和 β 链并用于创建组合文库。该文库由Twist Bioscience使用人类V片段、CDR3片段和J片段组装而成,而恒定区(C)来自鼠源。使用可变TCR α 链和 β 链结构域两侧的引物对来扩增两条链,并使用纳米孔测序来揭示两条链的身份(unveil the identity of both chains)。示出了10000个 $\alpha \times \beta$ 组合中的每一个的表征(representation)。结果示于图10C中的图中,其中概率密度表示在y轴上。

[0774] 步骤10D) Jurkat报告T细胞中的文库表达。在一些实施方案中,上述文库可在报告T细胞例如Jurkat中表达。将来自10C的文库)转染到Phoenix细胞中用于病毒生产,并使用所得病毒上清液进行CD8+TCR-KO Jurkat报告T细胞的逆转录病毒转导。使用杀稻瘟菌素对细胞进行选择,并使用针对小鼠TCR- β 恒定区的抗体检测TCR表达的阳性。结果示于图10D的图中。

[0775] 步骤10E) 筛选的分选(sorting)策略。在一些实施方案中,可以通过任何方法对文库进行筛选。以1:1的比例将来自10D的Jurkat报告T细胞)与以多个串联小基因(TMG)的形式表达pt1变体的B细胞共培养21小时。然后使用CD69标志物,通过FACS对细胞进行T细胞活化分选。分选策略包括(从左到右)顺序门控以选择淋巴细胞、门控以选择单细胞、门控以排除CD20⁺细胞,以及分别捕获表达高CD69和低CD69的细胞的两个分选门(“上(top)”和“下(bottom)”)。结果示于图10E的图中。

[0776] 步骤10F) TCR表达盒的提取。在一些实施方案中,可以通过多种技术中的任何一种

来提取相关的TCR表达盒。在一些实施方案中,使用在10C)中描述的PCR策略,随后进行条形码PCR来提取来自10E)的顶部和底部样本的TCR表达盒(三次重复,并且包括在Jurkat报告T细胞与不表达TMG的B细胞的共培养物上的三个对照筛选)。还包括质粒TCR文库的对照PCR。使用安捷伦TapeStation(左侧)分析PCR产物。按1:1的比例将PCR产物混合,并在TapeStation(右侧)上分析。结果示于图10F中。

[0777] 步骤10G) 筛选分析。在一些实施方案中,可以用多种方法对来自步骤F的PCR产物池进行分析。例如,使用来自10F)的PCR产物池进行文库制备并测序。使用DESeq2 R套装软件恢复TCR α 链和 β 链身份并识别差异表达的TCR组合。B细胞存在(x轴)和不存在(y轴)TMG表达时筛选的平均Rlog转换的读段计数表示了10B)-10F)中所述的pt1肿瘤样本,以及以相同方式处理的另外三个MMRp-CRC样本(pt2、pt3和pt4)。新生抗原反应性TCR先导在图10G中描绘为圆圈状较大的黑点。

[0778] 步骤10H) 相关TMG的解卷积(deconvolution)。在一些实施方案中,可以以多种方式对相关TMG解卷积。例如,使用表达单个TMG构建体的B细胞重新筛选(单次复制)10G)中识别的新生抗原反应性TCR。作为实例,表示pt1样本的多路分解筛选(demultiplexing screen)。pt1 TCR先导识别pt1-TMG1,而非pt1-TMG2。结果示于图10H的图中。

[0779] 步骤10I) pt1 TCR先导抗原的识别。可通过多种方式来识别TCR先导抗原(例如pt1)。例如,通过向B细胞加载pt1-TMG1中所代表的单个小基因的肽来识别pt1 TCR先导识别的新生抗原。使用表达pt1-TMG1的B细胞作为阳性对照。将APC与Jurkat报告T细胞共培养,后者表达在10G)中识别的pt1新生抗原反应性TCR先导。活化被测量为相对于阳性对照(用PMA/离子霉素处理)的CD69阳性。对照(PARVA-P70R)以及AKAP8L-R191W肽的激活作用示于图10I中。

[0780] 步骤10J) pt3 TCR先导抗原的识别。可通过多种方式来识别TCR先导抗原(例如pt3)。例如,通过向B细胞加载pt3-TMG1中所代表的单个小基因的肽来识别pt3 TCR先导识别的新生抗原。使用表达pt3-TMG1的B细胞作为阳性对照。将APC与Jurkat报告T细胞共培养,后者表达在10G)中识别的pt3新生抗原反应性TCR先导。活化被测量为相对于阳性对照(使用PMA/离子霉素治疗)的CD69阳性。对照(ETS1-R70W)以及TP53-R282W肽的激活作用示于图10J中。

[0781] 步骤10K) 第一pt4 TCR先导抗原的识别。可以通过多种方式识别第一TCR先导抗原(例如pt4)。例如,通过B细胞中的一个微小基因的表达(编码HSPA9-p.K654RfsX42的MG91)来识别第一pt4 TCR先导识别的新生抗原。使用表达pt4-TMG3的B细胞作为阳性对照。将APC与Jurkat报告T细胞共培养,后者表达10G)中识别的第一pt4新生抗原反应性TCR先导。活化被测量为相对于阳性对照(使用PMA/离子霉素治疗)的CD69阳性。对照(不表达MG或TMG的B细胞)以及MG91/TMG3样本的激活作用示于图10K中。

[0782] 步骤10L) 第二pt4 TCR先导抗原的识别。第二TCR先导抗原(如pt4)可通过多种方式识别。例如,通过B细胞中的一个微小基因(编码ITPR3-p.L2379M的MG132)的表达来识别第二pt4 TCR先导识别的新生抗原。使用表达pt4-TMG4的B细胞作为阳性对照。将APC与Jurkat报告T细胞共培养,后者表达在10G)中识别的第二pt4新生抗原反应性TCR先导。活化被测量为相对于阳性对照(使用PMA/离子霉素处理)的CD69阳性。对照(不表达MG或TMG的B细胞)以及MG132/TMG4样本的激活作用示于图10L中。

[0783] 图12A-12C中描述了一些附加实施方案：从黑素瘤中回收TCR库用于生成TCR $\alpha\beta$ 文库。图12A描述了人黑素瘤样本中浸润淋巴细胞的批量TCR测序。通过Milaboratory (莫斯科/俄罗斯) 对两份人肿瘤样本 (pt5和pt6) 进行批量TCR测序。经过比对和TCR识别后, 克隆型根据它们的CDR3氨基酸序列及它们的V和J同一性收缩。唯一克隆型的数量表示每个肿瘤样本的 α 链和 β 链二者。在图12B中, 描述了100 $\times$ 100个文库的质量控制。针对来自A) 的每个样本选择100个最普遍的 α 链和 β 链, 并用于创建组合文库。合成B) 中的TCR- β 区域和TCR- α 区域, 并使用由Twist Bioscience执行的组合克隆方法将其插入图11B所示的逆转录病毒构建体中。使用可变TCR α 链和 β 链结构域两侧的引物对来扩增两条链, 并使用纳米孔测序来识别两条链的身份。对于每个患者文库, 示出了10000个 $\alpha \times \beta$ 组合中的每一个的表征 (representation)。图12C) 中提供了患者文库的TCR表征 (representation) 的特征。对于12B) 中的每个患者文库, 示出了每个TCR的读段的数量范围、平均覆盖率以及在中位数 $\pm 2\log$ 单位范围内的TCR百分比。

[0784] 图13A-13E中提供了一些附加实施方案 (作为一些实施方案的步骤, 图13A-13E示出了从通过质粒人工混合生成的TCR文库中恢复抗原特异性TCR的情况。在步骤13A) 中, 通过人工混合TCR质粒生成文库。通过人工混合质粒生成TCR文库。将各表达一种特征性TCR的6个质粒混合到各表达一种非特征性卵巢癌 (OVC) TCR的11个质粒和各表达一种非特征性结肠癌 (CRC) TCR的13个质粒的池中。在特征性TCR中, 将CMV1和CDK4-8 TCR质粒以1:10.000的摩尔比混合到此池中, 并将CMV2、CDK4-17、GCN1L1和NY-ES0 1G4 TCR质粒以1:100.0000的摩尔比混合到此池中。频率示于图13A中。步骤13B) 中Jurkat报告T细胞中的文库表达。在一些实施方案中, 来自步骤13A) 的TCR的混合物可以在报告T细胞如Jurkat中表达。将来自步骤13A) 的TCR质粒池转染到Phoenix细胞中用于病毒生产, 并将所得病毒上清液用于TCR-KO Jurkat报告T细胞的逆转录病毒转导。使用嘌呤霉素选择细胞, 并使用针对鼠TCR- β 恒定区的抗体检测阳性TCR表达。结果示于图13B的图中。在步骤13C) 中, 提供了用于筛选的分选策略 (sorting strategy)。在一些实施方案中, 可以通过任何方法对文库进行分选。将来自步骤13B) 的Jurkat报告T细胞与以多个串联小基因 (TMG) 的形式表达表征的TCR的每种同源抗原的B细胞以1:1的比例共培养21小时。然后使用CD69标志物, 通过FACS对细胞进行分选以激活T细胞。分选策略包括 (从左到右) 顺序门控以选择淋巴细胞、门控以选择单细胞、门控以排除CD20+细胞, 以及分别捕获表达高CD69和低CD69的细胞的两个分选门 (“顶部 (top)” 和 “底部 (bottom)”)。结果示于图13C的图中。

[0785] 在步骤13D) 中, 提供关于TCR表达盒的信息。在一些实施方案中, 可以通过多种技术中的任何一种来提取相关的TCR表达盒。在一些实施方案中, 使用在10C) 中描述的PCR策略, 随后进行条形码PCR来提取来自步骤13C) 的顶部和底部样本的TCR表达盒 (用表达TMG的B细胞进行筛选的一个重复和用不表达外源性抗原的B细胞进行筛选的一个重复)。还包括质粒TCR文库的对照PCR。使用Agilent TapeStation分析第二轮PCR后的PCR产物。结果示于图13D中。

[0786] 在步骤13E) 中, 提供了对筛选数据的分析。在一些实施方案中, 可以以多种方式对来自步骤13D) 的PCR产物池进行分析。例如, 使用来自13D) 的PCR产物池进行文库制备, 并使用纳米孔技术对所述PCR产物池进行分析进行测序。使用DESeq2 R套装软件恢复TCR α 链和 β 链身份 (identity), 并识别差异表达的TCR $\alpha \times \beta$ 链组合。以相对于TCR测量频率 (x轴) 的方式

表示顶部样本与底部样本中归一化读段计数之间的 $\log_2$ 转换比(y轴)。在图13E中,特征性(抗原特异性)TCR用灰色表示,而非特征性(无关的)TCR用黑色表示。

[0787] 图14中示出了附加实施方案,其示出了从通过基因合成生成的TCR文库中恢复抗原特异性TCR。在步骤14A(图14A)中,提供了筛选设计的示意图。使用来自卵巢癌(OVC)或结肠直肠癌(CRC)样本的5个特征性TCR和45个或95个非特征性TCR,分别构建 50×50 和 100×100 设计的组合TCR文库。该文库由Twist Bioscience使用人V、CDR3和J片段组装而成,而恒定区I来自鼠源。该文库用于Jurkat报告T细胞的逆转录病毒转导。将多克隆报告T细胞与经工程改造以各种方式呈递抗原的抗原呈递细胞(APC)共培养。APC包括装载肽的JY细胞、各表达不同微小基因的EBV-LCL细胞系的混合物、表达TMG的EBV-LCL细胞以及未被工程改造以表达特定抗原的EBV-LCL。

[0788] 在步骤14B中,提供了用于筛选的分选策略。将表达如14A)中所述产生的 50×50 设计的TCR文库的Jurkat报告T细胞与14A)中提及的APC以1:1的比例共培养21小时。然后使用CD69标志物,通过FACS对细胞进行T细胞活化分选。在一些实施方案中,可以通过任何方法对库进行分选。分选策略包括(从左到右)顺序门控以选择淋巴细胞、门控以选择单细胞、门控以排除CD20+细胞,以及分别捕获表达高CD69和低CD69的细胞的两个分选门(“顶部”和“底部”)。结果示于图14B的图中。

[0789] 步骤14C(图14C)显示了TCR表达盒的提取。在一些实施方案中,可以通过多种技术中的任何一种来提取相关的TCR表达盒。在一些实施方案中,使用在10C)中描述的PCR策略,随后进行条形码PCR来提取来自图14B)的顶部和底部样本的TCR表达盒。还包括质粒TCR文库的对照PCR。使用Agilent TapeStation分析第二轮PCR后的PCR产物。结果示于图14C中。

[0790] 步骤14D)示出了对 50×50 筛选数据的分析。在一些实施方案中,可以以多种方式对来自步骤14C)的PCR产物池进行分析。例如,使用来自14C)的PCR产物池进行文库制备,并使用纳米孔技术对所述PCR产物池进行测序。恢复了TCR α 和 β 链的身份(identity),并将顶部和底部样本(y轴)中TCR表征(representation)之间的倍数变化表示为每个TCR的TCR平均表征(representation)(x轴)的函数。

[0791] 步骤14E)示出了前10种最显著富集的TCR的特征。使用DESeq2 R套装软件识别来自14D)中数据的差异表达的TCR $\alpha \times \beta$ 链组合。差异表征分析是本领域技术人员已知的,并且基于假设富集的TCR被定义为相对于不存在抗原时的“顶部”和“底部”样本,在存在抗原时“顶部”样本中富集,而在存在抗原时的“底部”样本中耗竭的线性模型。前10个最显著命中的 α 和 β 链,以及它们的表征,它们的 $\log_2$ 转换的倍数变化和差异表征的显著性列于图14E。

[0792] 步骤14F)示出了对 100×100 筛选数据的分析。 100×100 文库的筛选类似于14) — 14E)中所述的 50×50 文库的筛选。识别TCR α 链和 β 链后,使用DESeq2 R套装软件识别差异表达的TCR组合。差异表征分析是本领域技术人员已知的。图14F显示了在B细胞存在(x轴)和不存在(y轴)TMG表达的情况下 100×100 文库筛选的平均Rlog转化读段计数。5个掺入(spiked-in)的特征性TCR被描绘为带圆圈的较大黑点。

[0793] 步骤14G示出了特征性TCR的秩次序。 100×100 文库中所代表的所有TCR组合的富集显著性的秩次序,其中按照14E)中所述进行统计分析)。使用DESeq2 R套装软件计算Wald统计量,并表示为Wald统计量递减的有序图(概率度量;y轴)。掺入特征的TCR以灰色阴影表示。插图:前20个统计学上最显著富集的TCR的放大图。

[0794] 图15A-15D描述了使用基因合成创建TCR库的一些实施方案。在步骤15A) (图15A) 中,提供了TCR文库的示意性布局。含有TCR β 链和 α 链二者以及嘌呤霉素选择标志物的逆转录病毒构建体被用作构建文库的支架。TCR α 链或 β 链(V-CDR3-J)的可变区被合成为每个具有100个寡核苷酸池的池。随后,在组合克隆反应中使用 α 链和 β 链池,以获得总复杂度为10000的文库。TCR链合成和组合克隆二者均由Twist Bioscience (旧金山/美国) 进行。使用该构建体的逆转录病毒转导最终导致单一转录物的表达,由于在2A位点处的肽切割,这导致TCR β 链和 α 链,以及产生的嘌呤霉素抗性标志物的翻译。

[0795] 在步骤15B) 中,提供了100个 α 链和100个 β 链的组合TCR文库的质量控制。使用可变TCR α 链和 β 链域两侧的引物对来扩增两条链,并使用纳米孔测序对用于识别两条链的身份。示出了100条 α 链(左图)、100条 β 链(中图)或10000个 $\alpha \times \beta$ 组合的每一个的表征(representation)。

[0796] 在步骤15C) 中,从多个复杂度较低的文库创建复杂度较高的文库。此示意图描述了从多个复杂度较低的文库创建一个更复杂的文库的想法。在一些实施方案中,复杂度较低的文库不包含任何 α 链和 β 链的可能重叠的组合。在一些其他实施方案中,复杂度较低的文库确实包含 α 链和 β 链的可能重叠的组合。在本例中,通过以等摩尔比混合四个 100×100 文库,创建了200个 α 链和200个 β 链的组合文库(200×200 文库)。在本实例中,使四个子文库生成为以下选项的组合文库:i) TCR α 1-100号和TCR β 1-100号;ii) TCR α 101-200号和TCR β 1-100号;iii) TCR α 1-100号和TCR β 101-200号;以及iv) TCR α 101-200号和TCR β 101-200号。在一些实施方案中,文库的复杂度可以在 50×50 到 2000×2000 之间变化。

[0797] 步骤15D) (如图15D所示) 通过基于配对信息或配对可能性的设计来减少TCR文库。该示意图描述了在不影响或仅有限地影响文库中存在的抗原特异性TCR链对的数量情况下降低TCR库的复杂度的想法。例如,对于 100×100 库,可以通过 10×10 设计的10个组合子文库的等摩尔混合来使10000的复杂度降低。这导致TCR库的复杂度降低了10倍。关于TCR α 链和 β 链配对的信息,或 α 链和 β 链之间的配对可能性的信息,可以被包括在文库的设计中,以这样的方式,实验识别的或以其他方式已知的TCR α - β 对的所有 α 链和 β 链,或可能配对的 α 链和 β 链,在单个组合子文库中被表示。在一些实施方案中,该原理可以被应用于复杂度更高或更低的复合文库。在一些实施方案中,组合子文库的复杂性可以更高或更低。在一些实施方案中,组合子文库可以包含重叠的TCR链组合。在一些实施方案中,复合文库可以具有100 (10×10) -90000 (300×300) 的范围。在一些实施方案中,子文库的大小范围可以是1 (1×1) -100 (50×50)。

[0798] 图16A-16E中示出了其他的实施方案,提供了用于从TCR库中识别TCR α/β 对的共培养条件。

[0799] 图16A) 描述了使用CD69作为活化的Jurkat细胞的选择标志物的结果。用CMV#1TCR转导表达hCD8的Jurkat细胞,转导效率为21%。在 370°C 下,用不同量的CMV_{pp65}肽(范围为0-10 μg /毫升,如所示)或1 μg /毫升的MART-1不相关肽(IRR)对JY细胞进行脉冲处理1小时。1小时后洗涤细胞,在 370°C 下在U形底96孔板中将 1×10^5 个细胞与 1×10^5 个转导的Jurkat细胞/孔共培养20小时。收获细胞,用抗人CD69(克隆:FN50)染色,并通过FACS进行分析。

[0800] 图16B) 描述了依赖于接种密度的CD69背景表达。以不同密度($0.25 \times 10^6/\text{ml}$ 、 $0.5 \times 10^6/\text{ml}$ 和 $1 \times 10^6/\text{ml}$)培养表达hCD8的Jurkat细胞和CMV#1TCR 2天。收获细胞,用抗人CD69

(克隆:FN50)染色,并通过FACS进行分析。

[0801] 图16C描述了活化的Jurkat细胞在不同培养容器中以及在各种效应物与靶标的比率下的CD69表达。将表达hCD8的Jurkat细胞和TCR文库(由4个特征性抗原特异性TCR的 α 链和 β 链组成的 4×4 组合文库)与表达同源小基因的靶B细胞(TM2.1)共培养,同时保持 2.5×10^5 个效应细胞/ 0.32 cm^2 培养面积。使用效应细胞与靶细胞的1:1或1:2的比例,在U形底96孔板中培养细胞。或者,在T25培养瓶中以细胞总量的1/3或1/2培养细胞,否则将用于该表面积,同时保持效应细胞与靶细胞的1:1的比例。此外,在T75培养瓶中以效应细胞与靶细胞的1:2和3:1的比例培养细胞。孵育20小时后,收获细胞,用抗人CD69(克隆:FN50)染色,并通过FACS进行分析。

[0802] 图16D描述了在各种共培养密度下的CD69表达。在U形底96孔板中将表达hCD8和CDK4-17或CDK4-8 TCR的Jurkat细胞与用指示量的CDK423-32(24L)肽或用不相关的MART-126-35(27L)肽脉冲的JY细胞以1:1的比例共培养。效应细胞的接种密度为每孔 125×10^3 个、 250×10^3 个或 500×10^3 个。20小时后,收获细胞,用抗人CD69(克隆:FN50)染色,并通过FACS进行分析。

[0803] 图16E描述了在遗传筛查中使用CD69作为T细胞活化的标志物,以识别反应性TCR。将等比表达未知特异性TCR的Jurkat hCD8+细胞池与以指定频率表达CDK4-17、CDK4-8、CMV#1、CMV#2或GCN1L1 TCR的Jurkat hCD8+细胞混合。将这种细胞混合物与表达所有同源小基因的靶B细胞共培养。20小时后,收获细胞,用抗人CD69(克隆:FN50)染色,并分选为分别表达高CD69水平和低CD69水平的“顶部”和“底部”细胞群。这些样本中的每一个均含有约10%的分选细胞总量。从这些细胞中提取基因组DNA,并通过PCR扩增TCRB可变区。对样本进行Illumina测序,并将结果读段数映射至参考TCRB序列。以顶部样本相对于底部样本的归一化读段数计算倍数富集。

[0804] 图17-图27示出了其他实施方案。

[0805] 图17和图27描述了CD69用于检测抗原激活的T细胞的用途。CD69表达模式的知识允许在T细胞受体(TCR)文库筛选中检测和选择活化的T细胞。

[0806] 图18、图25a和图25b以及图26描述了杀稻瘟菌素用于选择用TCR基因转导的Jurkat报告T细胞,从而提供了在引入TCR文库后对报告T细胞的有效选择的用途。

[0807] 图19描述了刺激细胞培养袋中的Jurkat报告T细胞,从而在TCR文库筛选过程中刺激大量用TCR文库转导的报告T细胞。使用大量细胞可以通过保持足够的覆盖率提高TCR文库筛选的敏感性。

[0808] 图20描述了用不同的TCR转导的Jurkat报告T细胞上的CD69、CD25和CD62L表达的纵向分析。了解表达标志物的纵向表达从而选择单个(或多个)活化标志物来特异性检测抗原激活的T细胞,并进行两步选择程序,例如通过磁珠富集。

[0809] 图21至图24描述了使用不同的NFAT报告系统来检测抗原激活的T细胞的,从而了解NFAT报道基因盒的设计、递送类型和它们的基因组插入位点控制报道基因的与抗原无关的背景表达的功能性和水平。数据表明,不同的病毒递送载体、克隆报告T细胞群、不同的报告T细胞或不同的报告系统可能导致更优化的报告功能。

[0810] 在一些实施方案中,该方法可以包括下面(以及在图31中)描述的步骤(1)-(7)。步骤(1)获取样本。样本可以是患有传染病、自身免疫性疾病或癌症的患者的组织、血液或体

液。样本可以是活的或非活的。步骤(2)对样本中的TCR- α 链和TCR- β 链进行测序。步骤(3)选择并对TCR α -链和TCR β -链序列进行组合配对以创建TCR $\alpha\beta$ 对的文库。步骤(4)将TCR $\alpha\beta$ 对的文库导入报告细胞池,例如Jurkat报告T细胞。步骤(5)用呈递至少一种目的抗原的抗原呈递细胞刺激用TCR $\alpha\beta$ 对的文库修饰的报告细胞。所述至少一种目的抗原可以是自体的或同种异体的。步骤(6)确定对所述至少一种目的抗原特异的TCR $\alpha\beta$ 对。步骤(7)将TCR $\alpha\beta$ 对导入细胞并选择含有TCR $\alpha\beta$ 对的细胞。

[0811] 在一些实施方案中,该方法可以包括上述步骤(1)-(7)中的一步或多步。任何步骤均可酌情省略、重复或由本文提供的其他实施方案替代。也可以添加额外的中间步骤。例如,一些实施方案包括步骤(2)和(3)。其他实施例包括步骤(5)和(6)。还有一些包括步骤(7)。一些实施方案包括步骤(1)-(7),并且进一步包括向患者施用含有TCR $\alpha\beta$ 对的细胞用于治疗。在一些实施方案中,抗原呈递细胞可以通过将新生抗原文库引入B细胞而获得。新生抗原文库可以是自体的或同种异体的。一些实施方案涉及通过选择TCR链子集来创建TCR库。一些实施方案涉及包含来自新生抗原文库的任何新生抗原的B细胞。一些实施方案涉及基于富集/耗竭的遗传筛查的应用。一些实施方案涉及用大尺寸扩增子进行遗传筛查。一些实施方案涉及用于检测TCR修饰的T细胞的方法。一些实施方案结合了前述实施方案中的任何一个或多个。

[0812] 一些实施方案涉及包含根据上述实施方案中任一个恢复的T细胞受体库的核苷酸文库。

[0813] 在一些实施方案中,包含根据上述实施方案中任一个识别的核苷酸序列的核苷酸构建体。在一些实施方案中,细胞包含本文所述的核苷酸构建体。

[0814] 一些实施方案以图28为根据。在一些实施方案中,通过应用可扩展和微创的遗传筛查方法来实现新生抗原特异性TCR识别。将少量非活的存档肿瘤组织用作肿瘤内TCR序列的来源,而不是TIL。接下来,使用逆转录病毒基因转移将已识别的肿瘤内TCR文库导入永生化的T淋巴细胞细胞系,称为Jurkat报告T细胞系。富集表达TCR文库的Jurkat细胞(效应细胞,简称E),随后筛选它们对表达突变组的患者衍生的APC(靶细胞,简称T)的反应性。流式细胞术分选后,根据参与细胞增殖和下游信号转导的早期T细胞活化标志物CD69(作为一个例子)的表达选择Jurkat细胞³²。在一些实施方案中,可以基于任何其它活化标志物来分离样本,包括但不限于CD25、CD62L、CD137、IFN- γ 、IL-2、TNF- α 、GM-CSF、合成启动子报告标志物或增殖标志物。作为最后一步,通过下一代测序(next generation sequencing)来识别新生抗原特异性TCR。目前的TCR分离平台提供了对10000个TCR文库的筛选。为了能够从该文库中以高灵敏度识别新生抗原特异性TCR并确保TCR在不同处理步骤中不丢失,必须在筛选过程中多次表征(represented)每个唯一的TCR,以保持TCR的覆盖范围。因此,必须筛选大量TCR转导的Jurkat细胞和APC³³。

[0815] 在将逆转录病毒TCR文库转导至Jurkat TCR KO细胞后,有效且高通量的筛选程序有助于成功富集经TCR转导的Jurkat细胞(>80% mTCR β^+ CD8 $^+$ 细胞),且不会引起毒性和损失TCR覆盖率。在一些实施方案中,96孔圆形底板可用于APC-Jurkat共培养。在一些实施方案中,GMP袋可用于APC-Jurkat共培养。在一些实施方案中,共培养可以在封闭的系统中进行。在一些实施方案中,可在GMP袋中建立与168 $\times$ 106 Jurkat细胞和APC的共培养物。在一些实施方案中,共培养可以进行16小时、20小时、24小时、48小时或32小时。在一些实施方案中,

读段数可以关于CD69、CD25或CD62L。在一些实施方案中,读段数可以是CD69和CD25的组合。在一些实施方案中,读段数可以是CD69和CD62L的组合。一些实施方案涉及与CD69组合的CD25⁺CD62L⁻Jurkat细胞的选择。在一些实施方案中,可以使用GMP袋。

[0816] 抗生素选择是富集TCR转导的Jurkat T细胞的一个有吸引力的策略。在一些实施方案中,可以使用杀稻瘟菌素进行选择。一些实施方案以图30为根据。在一些实施方案中,在选择的第4天,将细胞以其起始密度重新铺板在含有或不含相应浓度的杀稻瘟菌素的培养基中。在一些实施方案中,杀稻瘟菌素的浓度为4μg/毫升。一些实施方案是7天选择,起始细胞密度为 0.25×10^6 个细胞/ml,杀稻瘟菌素浓度为4μg/毫升,在第4天去除抗生素。在一些实施方案中,可以以0.5e6个细胞/ml铺板并添加6μg/毫升的杀稻瘟菌素,并在第4天以0.5e6/ml重新铺板细胞,不添加杀稻瘟菌素。

[0817] 在一些实施方案中,报告系统可以是AP-1或NFκB信号传导通路。

[0818] 一些实施方案以图29为根据。在一些实施方案中,一个目标是扩大文库中TCR的数量,以允许对大于10000个TCR进行高灵敏度筛选。这突出了通过优化各种流程步骤来提高TCR发现平台的可扩展性的需要,以允许更有效地处理大量细胞,同时仍然保持TCR覆盖率。在一些实施方案中,使用了四种已知的HLA-A*02:01限制性TCR—CDK4 TCR克隆8和17(简称CDK4-8和17)以及CMV TCR克隆1和2(简称CMV-1和2)。这两种CDK4 TCR对突变的细胞周期蛋白依赖性激酶4(CDK4_{R24C})肽具有特异性。在多个黑素瘤患者中识别出了这种突变衍生的新生抗原表位³⁴。此外,两个CMV TCR靶向人巨细胞病毒(CMV)成分pp65编码的肽³⁵。对于CDK4和CMV表位,在研究中使用了具有潜在不同亲和力的两个不同的TCR克隆。这允许评价TCR对同源肽的亲和力在筛选过程中的作用。

[0819] 其他实施方案-实施例19中的结果

[0820] TCR基因疗法涉及工程改造自体T细胞以表达对癌症抗原具有所需的特异性的TCR。一类癌症抗原是非同义体细胞突变衍生的新生抗原(nonsynonymous somatic mutation-derived neo-antigen),它们仅在恶性细胞上表达,因此是TCR基因疗法的一个有吸引力的靶点。然而,大多数新生抗原对给定患者的肿瘤是独特的,靶向它们需要个性化的方法。为了克服这一挑战,提供了一种完全个性化的新生抗原特异性TCR基因疗法,方法是结合遗传筛查方法,从肿瘤活检中分离的患者TCR基因文库中识别此类TCR。目前的TCR分离平台允许筛选10000个TCR,这涉及处理大量TCR文库转导的报告Jurkat T细胞和表达新生抗原的抗原呈递细胞(APC),以最大限度地提高筛选灵敏度。然而,对大量细胞的处理可能会影响平台的可扩展性。因此,本研究旨在通过考察处理大量细胞的替代方法,提高筛选平台的可扩展性,同时保持其灵敏度。首先,研究表明,杀稻瘟菌素选择可有效富集表达TCR的Jurkat细胞,且毒性最低。接下来,将Jurkat细胞-APC共培养物从已建立的96孔板形式放大至更高通量的MACS GMP细胞分化袋设置,导致Jurkat细胞的类似的活化。此外,研究了不同方法,以用更具可扩展性的基于珠的选择来代替目前在筛选过程中使用的流式细胞术分选。因此,除CD69外,还检测了另外两种T细胞活化标志物CD25和CD62L的表达谱,发现它们是与CD69联合进行两步基于珠的富集的潜在和合适候选者。此外,评估了基于NFAT(活化T细胞的核因子家族)的报告系统,该系统将通过利用报道基因(如抗生素抗性盒或适用于基于珠的选择的细胞表面标记)来规避流式细胞术分选的使用。然而,NFAT-报告系统表现出了高水平的背景信号。

[0821] 在肿瘤微环境中,细胞毒性T淋巴细胞(CD8⁺T细胞)主要负责肿瘤消退¹。T细胞表达独特的T细胞受体(TCR),这些T细胞受体是由α链和β链组成的异二聚体。每个α链和β链由一个恒定区和可变区组成²。可变区赋予给定T细胞对抗原呈递细胞(APC)在其主要组织相容性复合体(MHC)上呈递的同源肽的特异性和亲和力。MHC分子在人类中也被称为人类白细胞抗原(HLA)^{2,3}。CD8/CD4和CD28是T细胞共受体的实例,它们稳定TCR-HLA复合物,并与CD3ζ一起启动涉及蛋白酪氨酸磷酸化和细胞质钙释放的下游信号传导。这种下游信号传导诱导转录因子NFAT(活化T细胞的核因子家族)、AP-1和NF-κB的核易位(nuclear translocation),以及随后的T细胞活化特异性基因的转录^{4,5}。活化的CD8⁺T细胞能够通过产生多种炎性细胞因子如白细胞介素-2(IL-2)、IFN-γ和TNF-α来杀伤表达病毒、细菌或癌症抗原的靶细胞。分泌型IL-2与T细胞上的IL-2受体结合,通过刺激更多IL-2的产生和增强T细胞的增殖形成正反馈环(positive feedback loop)^{5,6}。

[0822] 迄今为止,癌症免疫疗法已被证明是治疗晚期肿瘤的最有效方法之一。许多免疫疗法利用了细胞毒性T细胞的杀伤特性^{7,8}。一个实例是阻断免疫检查点分子CTLA-4和PD-1,它们抑制肿瘤微环境中CD8⁺T细胞的细胞毒性活性^{9,10}。检查点阻断是一种可扩展的常规给药疗法,在非同义突变率高¹¹的癌症中最有效例如黑素瘤^{12,13}和非小细胞肺癌^{14,15}。此外,自体肿瘤浸润淋巴细胞(TIL)的过继性细胞转移,在体外扩增并补充IL-2,已显示出治疗黑素瘤^{16,17}和子宫颈癌^{18,19}的潜力。然而,肿瘤抗原特异性TIL通常是终末分化的T细胞,并在体外扩增过程中可能是寿命短暂的和丢失的^{7,20}。

[0823] 上述免疫疗法均未识别出参与肿瘤消退的肿瘤抗原反应性TCR。这些TCR的检测对于用TCR对肿瘤抗原进行基因工程改造T细胞是有用的。此类疗法被称为TCR基因疗法,并且呈现出优于其他免疫治疗策略的优势,因为它允许产生大量具有所需抗原特异性的“更合适的”T细胞²¹。肿瘤抗原可以是非自身抗原或自身抗原^{7,21}。自身抗原一直是癌症疫苗试验的主要焦点,但可能由于对自身抗原的中枢耐受性,这些试验一直无效²²。然而,一些针对异常表达自身抗原的TCR基因治疗试验已显示出临床疗效。例如,在滑膜细胞肉瘤和黑素瘤患者中靶向癌症种系NY-ESO-1表位已显示出治疗潜力²³。然而,靶向肿瘤相关的自身抗原也经常导致严重的靶向毒性(on-target toxicities)^{24,25},这是需要针对非自身抗原进行免疫疗法的基础。一种这样的类型是突变衍生的新生抗原^{7,8,21}。

[0824] 新生抗原来自非同义体细胞突变并导致产生健康组织中不存在的新多肽。这使得新生抗原成为免疫疗法的有用靶标,因为它们在健康组织中的完全不存在将防止靶向毒性。此外,新生抗原特异性T细胞的发现不受针对高亲和力自身抗原反应性T细胞的中枢耐受性的影响。新生抗原主要由乘客基因(passenger gene)中的突变产生,这些突变不会赋予恶性细胞任何存活优势。这些突变通常对每个患者都是独特的,因此靶向新生抗原需要涉及基因组测序方法的个性化治疗^{8,26}。

[0825] 随着最近开发的全外显子组测序(whole-exome sequencing, WES)和RNA测序,很明显在TIL中发现了新生抗原反应性T细胞,并且它们可以介导肿瘤消退^{27,28}。例如,WES与高度特异性和敏感的肽-MHC(pMHC)多聚体偶联²⁹已导致从黑素瘤患者的TIL材料中识别出新生抗原特异性T细胞^{13,30}。尽管这种方法是可扩展的,但它依赖于有限的HLA等位基因覆盖,并且需要算法来预测pMHC结合。这些限制可能导致某些新生抗原反应性TCR被忽视。为了规避pMHC多聚体的限制,已有研究使用质谱来识别与MHC结合的新生抗原。将这种方法与WES

和转录组测序作为参考物相结合,导致在鼠肿瘤细胞系中发现了新生抗原³¹。然而,这种设置是低吞吐量的,并且经常导致假阴性(false negative)⁸。尽管可以通过上述方法同时识别多种新生抗原特异性TCR,但仍需要一个更具可扩展性和灵敏度的平台来发现此类TCR。

[0826] 通过应用可扩展且微创的基因筛查方法,可以实现新生抗原特异性TCR识别。将少量非活的存档肿瘤组织用作肿瘤内TCR序列,而不是TIL的来源。接下来,使用逆转录病毒基因转移将已识别的肿瘤内TCR文库导入永生化T淋巴细胞细胞系,称为Jurkat报告T细胞系。富集表达TCR文库的Jurkat细胞(效应细胞,简称E),随后筛选它们对表达突变组的患者衍生APC(靶细胞,简称T细胞)的反应性。流式细胞术分选后,根据参与细胞增殖和下游信号转导的早期T细胞活化标志物CD69的表达选择Jurkat细胞³²。作为最后一步,通过下一代测序识别新生抗原特异性TCR。在一些实施方案中,目前的TCR分离平台提供了10000个TCR的文库筛选。与上述一致,实施例19提供了进一步的结果和证据来支持这一方法。

[0827] 可以通过多种方式实现TCR的选择。可使用合适的分析工具测定变体富集,包括但不限于DESeq2 R套装软件。变体富集可以包括对比报告T细胞与表达TMG的B细胞接触的顶底对(top-bottom pair),以及报告T细胞与未被工程改造以表达TMG的B细胞接触的顶底对。变体富集可包括根据DESeq2 Wald检验统计量按降序排列TCR组合。变体富集可能包括根据经过Bonferroni调整的p值确定排序较高的TCR组合的统计学显著性。至少一种TCR组合的选择可以基于调整后的p值和其他统计度量。本实例中的程序可作为单次重复执行,或作为一式两份、三份或三次以上重复来执行,以提高TCR反应性的灵敏度。对于来自与表达TMG的APC的共培养物的样本,以及对于来自含有未被工程改造以表达TMG的APC的共培养物的样本,重复次数可能会有所不同。这些选项中的任何一项都可以与本文提供的任何方法相结合。

[0828] 本发明的实施方案在以下实施例中进一步说明,这些实施例仅出于说明目的而给出,而不旨在以任何方式限制所要求保护的的主题的范围。

[0829] 实施例

[0830] 实施例1

[0831] 本实施例描述了从非活肿瘤样本中恢复TCR库以识别新生抗原特异性TCR序列。

[0832] 从新鲜冷冻的或固定的或福尔马林固定的/石蜡包埋(FFPE)的肿瘤样本中分离DNA或RNA,并用于进行批量TCR α -链和TCR β -链测序。基于使用“唯一分子标识符”(UMI)的独特分子的计数,确定编码(部分)特定TCR链氨基酸序列的核酸分子的绝对数量。或者,UMI不包括在TCR α -链和TCR β -链测序中,TCR链的频率是根据下一代测序读段计数而不是UMI计数来测量的。例如,通过应用例如肿瘤内TCR链丰度的标准,从已识别的TCR序列的总集合中选择一组定义的TCR α -链和TCR β -链的集合。随后,通过DNA或RNA合成分别产生选定的TCR α -链和TCR β -链的DNA或RNA片段。RNA片段可以通过标准技术转化为cDNA。通过将所有选定的TCR α -链和TCR β -链组合配对到编码TCR $\alpha\beta$ 基因的单一表达构建体中,重新创建了肿瘤内TCR $\alpha\beta$ 对的原始库的确定部分。单一表达构建体可用于表达单个TCR α 链和TCR β 链的给定组合。或者,TCR α 链和TCR β 链可以由单独的表达构建体表达。可以使用任何合适的表达载体,包括病毒载体。为了稳定表达,使用逆转录病毒或慢病毒载体或颗粒。得到的TCR $\alpha\beta$ 基因的文库在报告T细胞池中表达。新生抗原刺激使表达文库的T细胞被激活,并基于T细胞活化标志物富集新生抗原特异性T细胞。随后,通过在受抗原刺激的样本中的相对于未受抗原刺激的样本

的富集来识别被表达的新生抗原特异性TCR $\alpha\beta$ 基因。使用已被识别的TCR基因或TCR基因组来设计新生抗原特异性T细胞以用于癌症治疗。

[0833] 实施例2

[0834] 此实施例描述了为生成TCR $\alpha\beta$ 文库而恢复TCR库。

[0835] 从新鲜冷冻的或固定的或福尔马林固定的/石蜡包埋 (FFPE) 的样本中分离DNA或RNA,并用于进行批量TCR α -链和TCR β -链测序。基于使用“唯一分子标识符”(UMI)的唯一分子计数,确定编码(部分)特定TCR链氨基酸序列的核酸分子的绝对数量。例如,通过应用例如TCR链丰度的标准,从已识别的TCR序列的总集合中选择一组定义的TCR α -链和TCR β -链的集合。随后,通过DNA或RNA合成分别产生选定的TCR α -链和TCR β -链的DNA或RNA片段。RNA片段可以通过标准技术转化为cDNA。通过将所有选定的TCR α -链和TCR β -链组合配对到TCR $\alpha\beta$ 基因中,重新创建了TCR $\alpha\beta$ 对的原始库的限定部分。成对的TCR α 链和TCR β 链代表了一个TCR $\alpha\beta$ 文库,包括一个选定的TCR库。

[0836] 实施例3

[0837] 本实施例描述了利用恢复的TCR库的文库通过免疫疗法治疗癌症患者。

[0838] 通过实施例1和实施例2中概述的方法恢复TCR库并生成TCR $\alpha\beta$ 文库。TCR $\alpha\beta$ 基因的文库在报告T细胞池中表达。新生抗原特异性T细胞由抗原刺激活化,并在T细胞活化的基础上进行分离。随后,识别被表达的新生抗原特异性TCR $\alpha\beta$ 基因。通过在T细胞中表达TCR基因,被识别的TCR基因或一组TCR基因被用于工程改造新生抗原特异性T细胞用于癌症治疗。将工程改造的新生抗原特异性T细胞输注至癌症患者体内,作为治疗癌症的免疫疗法。癌症患者可能是其TCR $\alpha\beta$ 库被测序的患者或其癌症携带或表达相同新生抗原的患者。用于治疗的细胞可以是自体的或同种异体的。

[0839] 实施例4

[0840] 本实施例描述了从感染或自身免疫部位恢复TCR库。

[0841] 从新鲜冷冻的或固定的或福尔马林固定的/石蜡包埋 (FFPE) 的获自感染或自身免疫部位的样本中分离DNA或RNA,并用于进行批量TCR α -链和TCR β -链测序。例如,通过应用例如TCR链丰度的标准,从已识别的TCR序列的总集合中选择一组定义的TCR α -链和TCR β -链的集合。随后,通过DNA或RNA合成分别产生选定的TCR α -链和TCR β -链的DNA或RNA片段。RNA片段可以通过标准技术转化为cDNA。通过将所有选定的TCR α -链和TCR β -链组合配对到TCR $\alpha\beta$ 基因中,重新创建了TCR $\alpha\beta$ 对原始库的确定部分。通过测定能在感染或自身免疫位点检测特定抗原的T细胞的TCR序列,可恢复与感染或自身免疫位点相关或对其特异的TCR库。

[0842] 所得到的TCR $\alpha\beta$ 基因文库在报告T细胞池中表达。抗原特异性T细胞由抗原刺激激活,并在T细胞激活的基础上进行分离。随后,识别表达的抗原特异性TCR $\alpha\beta$ 基因。使用已识别的TCR基因或TCR基因组来诊断或治疗感染或自身免疫。

[0843] 实施例5

[0844] 本实施例描述了从通过人工混合TCR质粒产生的TCR文库中恢复抗原特异性TCR。

[0845] 以TCR β -P2A-TCR α -T2A-嘌呤霉素抗性的格式生成TCR表达盒(图6示出了TCR β -P2A-TCR α -T2A-嘌呤霉素抗性盒的示意性实例;以这种格式表示的TCR的实例给出于SEQ ID NO:1,图32,图40中)。通过将一种新生抗原特异性TCR与多种非相关TCR以多种比例混合,生成了多个TCR文库。生成的TCR文库包含一个HLA-A*02:01限制性新生抗原特异性TCR,频率

为1:10、1:100、1:1000和1:10000。为了创建含有频率为1:10的新生抗原特异性TCR的TCR文库,将逆转录病毒表达载体中编码新生抗原特异性TCR的一个质粒拷贝与其他九种逆转录病毒载体中编码不同的其他特异性TCR的每种载体的一个质粒拷贝混合。因此,将逆转录病毒表达载体中编码新生抗原特异性TCR的一个质粒拷贝与其他九种逆转录病毒载体中编码不同的其他特异性TCR的每种的1,11,111和1111个质粒拷贝混合,分别以1:10、1:100、1:1000和1:10000的频率获得含有新生抗原特异性TCR的TCR文库。

[0846] 在一些实施方案中,可以使用编码SEQ ID NO:1的任何一种或多种核酸。在一些实施方案中,可以通过在5个特征性TCR中以1:16、1:100、1:2500和1:10000的频率混合来改变上述情况。最后,可以使用24个而不是9个其他TCR质粒进行混合。

[0847] 通过本领域技术人员已知的方法将每个TCR文库分别转染到兼向性病毒产生细胞(amphotropic virus producer cell)如Phoenix-Ampho(ATCC CRL-3213)中。所得逆转录病毒病毒颗粒用于转导Jurkat报告T细胞系。Jurkat报告T细胞系缺乏内源性TCR表达(例如在Mezzadra等人Nature 2017中描述的),并且在用CD8 α -P2A-CD8 β 转基因(SEQ ID NO:2)转导后使用本领域技术人员已知的方法被修饰以表达人CD8 α 和CD8 β 。使用低MOI用单个TCR文库转导Jurkat报告T细胞,以便将TCR转导的Jurkat T细胞的频率限制为总T细胞的25-30%。通过在转导后向细胞培养基中添加嘌呤霉素来正向选择经修饰以表达TCR β -P2A-TCR α -T2A-嘌呤霉素抗性转基因的Jurkat T细胞。遗传修饰细胞的抗生素选择是本领域普通技术人员已知的。

[0848] 在一些实施方案中,可以使用编码SEQ ID NO:2(图33)的任何一种或多种核酸。图41描述了CD8 α -P2A-CD8 β 转基因SEQ ID NO:2的图。

[0849] 本领域普通技术人员将进一步理解,本文所述的方法可以包括不使用嘌呤霉素的选择,例如通过FACS或基于珠的选择对TCR转导的细胞进行分选。因此,TCR盒可能缺少嘌呤霉素选择基因。

[0850] 用表达重组HLA-A*02:01-IRES-FusionRed转基因的抗原负载K562细胞(K562-HLA-A*02:01-IRES-FusionRed;SEQ ID NO:3;图34、图42)刺激TCR转导的Jurkat T细胞。已经描述了转基因表达K562细胞的产生(例如,Hirano等人,Clin Canc Res 2006;Butler等人,Int Immunol 2010;Butler等人,Clin Canc Res 2007;Lorenz等人,Hum Gene Ther 2017),且为本领域技术人员所知。通过在37°C下用目的肽脉冲90分钟并随后洗涤获得肽负载的K562-HLA-A2细胞。本领域技术人员将理解,本文提及的肽呈递HLA-A*02:01阳性K562细胞系可被其他HLA-A*02:01阳性抗原呈递细胞替代。或者,可以使用变体(在HLA-A02*01转导的K562细胞中表达TMG)。

[0851] 以下刺激条件将分别用于每个TCR文库:

[0852] 1.4e⁷ TCR转导的Jurkat T细胞+4e⁷ K562-HLA-A2,装载1 μ g新生抗原肽

[0853] 2.4e⁷ TCR转导的Jurkat T细胞+8e⁷ K562-HLA-A2,装载1 μ g新生抗原肽

[0854] 随后,收获共培养物,并用荧光染料标记的CD4、CD8和CD69抗体对共培养物进行染色。使用BD Biosciences ARIAFusion流式细胞仪分选器,针对每种刺激条件对下列群进行分选:

[0855] 1.活的、单细胞、CD4⁺、CD8⁺、CD69^高(基于CD69荧光信号,CD69^高包括最高5-50%的活的、单细胞、CD4⁺、CD8⁺细胞)

[0856] 2. 活的、单细胞、CD4⁺、CD8⁺、CD69^低 (基于低CD69荧光信号, CD69^低包括5-50%的活的、单细胞、CD4⁺、CD8⁺细胞)

[0857] 或者, 以下刺激条件可分别用于每个TCR文库:

[0858] 1. 4e⁷ TCR转导的Jurkat T细胞+4e⁷ K562-HLA-A2, 装载1μg新生抗原肽

[0859] 2. 4e⁷ TCR转导的Jurkat T细胞+8e⁷ K562-HLA-A2, 装载1μg新生抗原肽

[0860] 3. 4e⁷ TCR转导的Jurkat T细胞+4e⁷ K562-HLA-A2, 装载1μg的无关对照肽

[0861] 4. 4e⁷ TCR转导的Jurkat T细胞+8e⁷ K562-HLA-A2, 装载1μg的无关对照肽

[0862] 随后, 收获共培养物, 并用荧光染料标记的CD4、CD8和CD69抗体对共培养物进行染色。使用BD Biosciences ARIA Fusion流式细胞仪分选器, 对每个样本的活的、单细胞、CD4⁺、CD8⁺、CD69⁺细胞进行分选。

[0863] 从分选的TCR转导的Jurkat T细胞中分离基因组DNA, 并用作PCR的模板以扩增部分TCRβ-P2A-TCRα盒。所得PCR产物的大小约为1.5kb, 可使用Oxford Nanopore MinIon测序仪进行测序, 以比较分选的T细胞群 (CD69^低和CD69^高或CD69⁺) 中新生抗原特异性TCR序列的相对丰度。

[0864] 在一些实施方案中, Oxford Nanopore MinIon可以由其他测序仪器代替, 或者可以采用其他测序策略。

[0865] 实施例6

[0866] 本实施例描述了从通过基因合成生成的TCR文库中恢复抗原特异性TCR。

[0867] 通过基因合成生成多个其中包含两个HLA-A*02:01限制性、新生抗原特异性TCR的TCR文库。为此, 合成了来源于新生抗原特异性TCR的两个TCRα链片段和两个TCRβ链片段以及来源于具有其他特异性的TCR的98个TCRα链片段和98个TCRβ链片段。或者, 可以采用5+95个TCR。随后, 使用所得片段生成含有以TCRβ-P2A-TCRα格式的TCR表达盒的TCR文库。为了生成文库, 将所有选择的TCRα和TCRβ片段混合并连接到编码TCRβ-P2A-TCRα盒的连续核酸分子上。重要的是, 每个TCR盒只能连接一个TCRβ- 和一个TCRα- 片段。通过选择和混合不同数量的TCRα和TCRβ片段, 可以创建不同复杂度的TCR文库:

[0868] 1. 来自新生抗原特异性TCR的两个TCRα链和两个TCRβ链片段将创建复杂度为4 (4个不同的TCRαβ对) 的TCR文库

[0869] 2. 来自新生抗原特异性TCR的两个TCRα链和两个TCRβ链片段以及来自具有其他特异性的TCR的八个TCRα链和八个TCRβ链片段将创建复杂度为100 (100个不同的TCRαβ对) 的TCR文库

[0870] 3. 来自新生抗原特异性TCR的两个TCRα链和两个TCRβ链片段以及来自具有其他特异性的TCR的48个TCRα链和48个TCRβ链片段将创建复杂度为2500 (2500个不同的TCRαβ对) 的TCR文库

[0871] 4. 来自新生抗原特异性TCR的两个TCRα链和两个TCRβ链片段以及来自具有其他特异性的TCR的98个TCRα链和98个TCRβ链片段将创建复杂度为10000 (10000个不同的TCRαβ对) 的TCR文库

[0872] 或者, 对于TCRα链片段和TCRβ链片段, 选项可以是: 4+0; 5+5; 5+45和5+95设计。

[0873] 所得的TCR文库将包含TCRβ-P2A-TCRα-T2A-嘌呤霉素抗性格式 (SEQ ID NO:1, 图40) 的TCR表达盒。

[0874] 通过本领域技术人员已知的方法将每个TCR文库分别转染到兼向性病毒产生细胞如Phoenix-Ampho (ATCC CRL-3213) 中。所得逆转录病毒病毒颗粒用于转导Jurkat报告T细胞系。Jurkat报告T细胞系缺乏内源性TCR表达(例如在Mezzadra等人, Nature 2017中描述的), 并且在用CD8 α -P2A-CD8 β 转基因 (SEQ ID NO:2) 转导后使用本领域技术人员已知的方法修饰以表达人CD8 α 和CD8 β 。使用低MOI用单个TCR文库转导Jurkat报告T细胞, 以便将TCR转导的Jurkat T细胞的频率限制为总T细胞的25-30%。通过在转导后向细胞培养基中添加嘌呤霉素来正向选择经修饰以表达TCR β -P2A-TCR α -T2A-嘌呤霉素抗性转基因的Jurkat T细胞。遗传修饰细胞的抗生素选择是本领域普通技术人员已知的。

[0875] 本领域普通技术人员将进一步理解, 本文所述的方法可以包括不使用嘌呤霉素的选择, 例如通过FACS或基于磁珠的选择对TCR转导的细胞进行分选。因此, TCR盒可能缺少嘌呤霉素选择基因。

[0876] 用表达重组HLA-A*02:01-IRES-FusionRed转基因 (K562-HLA-A*02:01-IRES-FusionRed; SEQ ID NO:3) 的负载抗原的K562细胞刺激TCR转导的Jurkat T细胞6小时。已经描述了转基因表达K562细胞的产生(例如, Hirano等人, Clin Canc Res 2006; Butler等人, Int Immunol 2010; Butler等人, Clin Canc Res 2007; Lorenz等人, Hum Gene Ther 2017), 且为本领域技术人员所知。通过在37°C下用目的肽脉冲90分钟并随后洗涤获得肽负载的K562-HLA-A2细胞。本领域技术人员将理解, 本文提及的肽呈递HLA-A*02:01阳性K562细胞系可被其他HLA-A*02:01阳性抗原呈递细胞替代。

[0877] 将使用以下刺激条件:

[0878] 1. 4×10^7 TCR转导的Jurkat T细胞+ 4×10^7 K562-HLA-A2, 装载1 μ g新生抗原肽

[0879] 2. 4×10^7 TCR转导的Jurkat T细胞+ 8×10^7 K562-HLA-A2, 装载1 μ g新生抗原肽

[0880] 随后, 收获共培养物, 并用荧光染料标记的CD4、CD8和CD69抗体对共培养物进行染色。使用BD Biosciences ARIA Fusion流式细胞仪分选器, 针对每种刺激条件对下列群进行分选:

[0881] 1. 活的、单细胞、CD4⁺、CD8⁺、CD69^高 (基于CD69荧光信号, CD69^高包括最高5-60%的活的、单细胞、CD4⁺、CD8⁺细胞)

[0882] 2. 活的、单细胞、CD4⁺、CD8⁺、CD69^低 (基于低CD69荧光信号, CD69^低包括5-50%的活的、单细胞、CD4⁺、CD8⁺细胞)

[0883] 或者, 以下刺激条件可分别用于每个TCR文库:

[0884] 1. 4×10^7 TCR转导的Jurkat T细胞+ 4×10^7 K562-HLA-A2, 装载1 μ g新生抗原肽

[0885] 2. 4×10^7 TCR转导的Jurkat T细胞+ 8×10^7 K562-HLA-A2, 装载1 μ g新生抗原肽

[0886] 3. 4×10^7 TCR转导的Jurkat T细胞+ 4×10^7 K562-HLA-A2, 装载1 μ g的无关对照肽

[0887] 4. 4×10^7 TCR转导的Jurkat T细胞+ 8×10^7 K562-HLA-A2, 装载1 μ g的无关对照肽

[0888] 随后, 收获共培养物, 并用荧光染料标记的CD4、CD8和CD69抗体对共培养物进行染色。使用BD Biosciences ARIA Fusion流式细胞仪分选器, 对每个样本的活的、单细胞、CD4⁺、CD8⁺、CD69⁺细胞进行分选。

[0889] 从分选的TCR转导的Jurkat T细胞中分离基因组DNA, 并用作PCR的模板以扩增部分TCR β -P2A-TCR α 盒。所得PCR产物的大小约为1.5kb, 可使用Oxford Nanopore MinIon测序仪进行测序, 以比较分选的T细胞群 (CD69^低和CD69^高或CD69⁺) 中新生抗原特异性TCR序列的

相对丰度。

[0890] 在一些实施方案中,Oxford Nanopore MinIon可以由其他测序仪器代替,或者可以采用其他测序策略。

[0891] 实施例7

[0892] 本实施例描述了从新鲜冷冻的黑素瘤病变产生的TCR文库中恢复新生抗原特异性TCR。

[0893] 从新鲜冷冻的黑素瘤样本中分离DNA和RNA,并一式两份使用:

[0894] 首先,利用DNA和/或RNA进行批量TCR α -链和TCR β -链测序。基于使用“唯一分子标识符”(UMI)的唯一分子计数,确定编码(部分)特定TCR链氨基酸序列的核酸分子的绝对数量。或者,可以使用读段计数。得到的TCR链序列的集合被分成TCR α -链序列的集合和TCR β -链序列的集合。任何无效的TCR链序列,其中TCR片段在氨基酸序列水平上框外连接,和/或其中引入终止密码子,和/或其中存在移码突变,和/或其中存在缺陷剪接位点,都被从集合中去除。使用编码特定TCR链的核酸分子的绝对数量(或分别在总TCR α -链或TCR β -链中的相应百分比)以降序对每个集合进行分类,以获得TCR α -链和TCR β -链的秩次序。前100个丰度最高的TCR α -链和TCR β -链是通过DNA合成选择和产生的片段。为了生成文库,将所有选择的TCR α 和TCR β 片段混合并连接到编码TCR β -P2A-TCR α 盒的连续核酸分子上。重要的是,每个盒只能连接一个TCR β -和TCR α -片段。得到的TCR文库将含有约10000个TCR β -P2A-TCR α -T2A-嘌呤霉素抗性格式的TCR表达盒(SEQ ID NO:1)。

[0895] 其次,使用来源于肿瘤以及健康组织的DNA和RNA,使用全外显子组测序(WES)来确定肿瘤特异性突变集,并使用RNA-seq建立表达的突变基因集。可以生成以串联阵列编码多种肿瘤衍生的突变肽的串联小基因(TMG)构建体。TMG构建体用于生成体外转录的mRNA(例如,Stevanović等人,Science 2017)。或者,TMG表达构建体可用于B细胞(APC)的病毒生产/转导。

[0896] 同时,使用来自黑素瘤患者的匹配自体血液产生永生B细胞。本领域技术人员已知人B细胞的EBV永生化(例如,Traggiai等人,Methods Mol Biol 2012)。永生化的自体B细胞用于通过用TMG-mRNA将B细胞电穿孔来产生表达抗原的B细胞。抗原呈递细胞(APC)的电穿孔已在前面描述过,并且是本领域技术人员已知的。本领域普通技术人员将理解,其他方法也将允许使自体B细胞永生化,并通过这种自体的永生B细胞诱导抗原表达,例如用肽脉冲、用编码TMG的DNA质粒转染或用编码TMG的病毒颗粒转导。

[0897] 通过本领域技术人员已知的方法将TCR文库转染到兼向性病毒产生细胞如Phoenix-Ampho(ATCC CRL-3213)中。所得逆转录病毒病毒颗粒用于转导Jurkat报告T细胞系。Jurkat报告T细胞系缺乏内源性TCR表达(例如在Mezzadra等人,Nature 2017中描述的),并且在用CD8 α -P2A-CD8 β 转基因(SEQ ID NO:2)转导后使用本领域技术人员已知的方法修饰以表达人CD8 α 和CD8 β 。使用低MOI用TCR文库转导Jurkat报告T细胞,以便将TCR转导的Jurkat T细胞的频率限制为总T细胞的25-30%。通过向细胞培养基中添加嘌呤霉素来正向选择经修饰以表达TCR β -P2A-TCR α -T2A-嘌呤霉素抗性转基因的Jurkat T细胞。遗传修饰细胞的抗生素选择是本领域普通技术人员已知的。本领域普通技术人员将进一步理解,本文所述的方法可以包括不使用嘌呤霉素的选择,例如通过FACS或基于珠的选择对TCR转导的细胞进行分选。因此,TCR盒可能缺少嘌呤霉素选择基因。

[0898] 使用以下条件通过加载抗原的B细胞刺激TCR转导的Jurkat T细胞6小时:

[0899] 1. 4×10^7 TCR转导的Jurkat T细胞+ 4×10^7 装载抗原的自体B细胞

[0900] 2. 4×10^7 TCR转导的Jurkat T细胞+ 4×10^7 自体B细胞

[0901] 随后,收获共培养物,并用荧光染料标记的CD4、CD8和CD69抗体对共培养物进行染色。使用BD Biosciences ARIAFusion流式细胞仪分选器,针对每种刺激条件对下列群体进行分选:

[0902] 1. 活的、单细胞、 $CD4^+$ 、 $CD8^+$ 、 $CD69^{\text{高}}$ (基于CD69荧光信号, $CD69^{\text{高}}$ 包括最高5-50%的活的、单细胞、 $CD4^+$ 、 $CD8^+$ 细胞)

[0903] 2. 活的、单细胞、 $CD4^+$ 、 $CD8^+$ 、 $CD69^{\text{低}}$ (基于低CD69荧光信号, $CD69^{\text{低}}$ 包括5-50%的活的、单细胞、 $CD4^+$ 、 $CD8^+$ 细胞)

[0904] 或者,以下刺激条件可分别用于每个TCR文库:

[0905] 1. 4×10^7 TCR转导的Jurkat T细胞+ 4×10^7 K562-HLA-A2,装载1 μ g新生抗原肽

[0906] 2. 4×10^7 TCR转导的Jurkat T细胞+ 8×10^7 K562-HLA-A2,装载1 μ g新生抗原肽

[0907] 3. 4×10^7 TCR转导的Jurkat T细胞+ 4×10^7 K562-HLA-A2,装载1 μ g的无关对照肽

[0908] 4. 4×10^7 TCR转导的Jurkat T细胞+ 8×10^7 K562-HLA-A2,装载1 μ g的无关对照肽

[0909] 随后,收获共培养物,并用荧光染料标记的CD4、CD8和CD69抗体对共培养物进行染色。使用BD Biosciences ARIAFusion流式细胞仪分选器,对每个样本的活的、单细胞、 $CD4^+$ 、 $CD8^+$ 、 $CD69^+$ 细胞进行分选。

[0910] 从分选的TCR转导的Jurkat T细胞中分离基因组DNA,并用作PCR的模板以扩增部分TCR β -P2A-TCR α 盒。所得PCR产物的大小约为1.5kb,可使用Oxford Nanopore MinIon测序仪进行测序,以比较分选的T细胞群($CD69^{\text{低}}$ 和 $CD69^{\text{高}}$ 或 $CD69^+$)中新生抗原特异性TCR序列的相对丰度。

[0911] 在一些实施方案中,Oxford Nanopore MinIon可以由其他测序仪器代替,或者可以采用其他测序策略。

[0912] 实施例8

[0913] 本实施例描述了使用基因合成创建TCR库。

[0914] TCR可以在如图3所示的盒设计中表达。为了合成大量这样的TCR盒,每个TCR链的高度可变部分被合成为片段,例如CDR3-J片段。其他组件可以是“现成的(off-the-shelf)”构建块。为了重建所有目的TCR链,合成了可变/独特片段,并与“现成的”构建块混合。原理如图7所示。随后,将所有组分混合在一起,组装将创建如图3所示的TCR盒。组装由特定的序列单链突出端(overhangs)控制:特定的CDR3-J片段将仅融合到某些V片段,并且由于不同片段之间的重叠,每个TCR盒将最终将采用TCR β -P2A-TCR α 格式。每一对TCR α 可以与每一个TCR β 配对。最后,将所有TCR β -P2A-TCR α 盒克隆到表达载体中。在某些情况下,可以使用V-CDR3-J(对于 α 和 β 链二者)作为可变部分,使用CB-P2A作为现成的构建块。

[0915] 或者,任何以及所有可能的TCR β -P2A-TCR α 组合完全通过基因合成产生。通过这种方式,可以组合配对所有TCR α 和TCR β 片段。

[0916] 另一种方法包括如上所述通过组合合成或通过基因合成产生TCR α 和TCR β 片段。然而,不是产生TCR β -P2A-TCR α 盒,而是将所得的TCR α 和TCR β 链的集合克隆到单独的表达载体中。用载体集合修饰细胞,以这样的方式使得每个T细胞平均表达一个如上所述的TCR α 和一

个TCR β ,类似于上述的组合配对。

[0917] 作为替代方法,可以使用修饰的TCR,例如与CD3 ϵ 或CD3 ζ 信号传导结构域单独融合或与CD3 ϵ 或CD3 ζ 和CD28信号传导结构域的组合融合的单链TCR构建体,而不仅仅是TCR α 和TCR β 。

[0918] 实施例9

[0919] 本实施例描述了从TCR库中识别TCR $\alpha\beta$ 对。

[0920] 用生成的TCR $\alpha\beta$ 对的文库修饰的T细胞池由呈递至少一种目的抗原的抗原呈递细胞刺激,并且基于至少一种用于TCR分离的活化标志物分离抗原反应性T细胞。或者,用适于追踪细胞增殖的荧光染料标记T细胞池,由表达至少一种目的抗原的抗原呈递细胞刺激,并基于增殖分离抗原反应性T细胞以进行TCR分离。活化T细胞的纯化可通过抗体标记和随后基于流式细胞术分选的分选、基于磁珠的选择或任何其他基于抗体结合的选择方法来实现。

[0921] 在另一种方法中,用生成的TCR $\alpha\beta$ 对的文库修饰的T细胞池被分成至少两个样本。样本由表达或不表达至少一种目的抗原的抗原呈递细胞刺激。刺激后,将两种T细胞群培养一段时间,然后通过TCR分离分析两种T细胞群。比较从两个样本中获得的TCR $\alpha\beta$ 对将识别暴露于至少一种抗原的样本中具有更高丰度的TCR基因。增殖检测可以基于荧光染料(如CFSE或Cell Tracer Violet)对稀释度的检测。通过流式细胞术根据稀释的荧光信号分选增殖细胞。

[0922] 在另一种方法中,用生成的TCR $\alpha\beta$ 对的文库修饰的T细胞池被呈递至少一种目的抗原的抗原呈递细胞刺激,并且基于至少一种报告TCR触发的报道基因,例如NFAT-GFP或NFAT-YFP来分离抗原反应性T细胞。

[0923] 在另一种方法中,用生成的TCR $\alpha\beta$ 对的文库修饰的T细胞池被呈递至少一种目的抗原的抗原呈递细胞刺激,并且使用基于TCR信号传递时获得的抗生素抗性的抗原特异性T细胞的选择,例如通过使用NFAT-嘌呤霉素转基因,分离抗原反应性T细胞用于TCR分离。或者,将修饰的T细胞池暴露于携带目的抗原的一种或多种MHC复合物。将与MHC复合物结合的T细胞分离用于TCR分离。可通过流式细胞术分选或磁珠富集进行基于MHC复合物结合的分选。

[0924] 上述任何方法中的TCR分离可以通过以下方法实现:(i)从大量抗原反应性T细胞中分离DNA或RNA以产生TCR $\alpha\beta$ 特异性PCR产物,所述产物通过DNA测序或RNA测序来分析以确定抗原反应性T细胞的TCR $\alpha\beta$ 基因序列,或(ii)基于单细胞的液滴PCR或微流体方法来分析在所分析的单个T细胞中表达的所述TCR $\alpha\beta$ 基因序列。以这种方式,检测TCR $\alpha\beta$ 转录物与升高水平的活化标志物共表达的T细胞池中的单个T细胞。

[0925] 在进一步的方法中,用产生的TCR $\alpha\beta$ 对的文库修饰的T细胞池被表达至少一种目的抗原的抗原呈递细胞刺激。随后,使用基于单细胞的液滴PCR或微流体方法识别目的TCR $\alpha\beta$ 对,以将TCR分离与至少一种活化标志物的转录物水平的检测相结合。由此,检测其中TCR $\alpha\beta$ 转录物与一种或多种活化标志物的增加的水平共表达的T细胞池内的单个T细胞。

[0926] 实施例10

[0927] 本实施例描述了从编码变体TCR α 和TCR β 多肽的核酸序列的组合文库中识别功能性TCR α /TCR β 组合的筛选方法。

[0928] 为了从非活肿瘤中识别新生抗原特异性TCR,开发了通过肿瘤衍生的TCR α 链和TCR

β链的组合组装产生TCR文库的方法。通过对从肿瘤组织中分离的DNA或RNA进行TCR批量链测序来识别这些TCR链。TCRαβ对被编码为约1.5kb的转基因并导入报告细胞。通过用抗原表达细胞刺激,可以在遗传变体文库筛选中选择表达抗原反应性TCRαβ组合的报告细胞。鉴于组合组装,每个TCR变体只能通过确定TCRα和TCRβ可变序列来明确识别。因此,在遗传筛查过程中分离的报告细胞中编码的转基因作为约1.5kb的PCR扩增子被回收,使用Oxford Nanopore测序法进行全长测序,并使用生物信息学分析管道(bioinformatic analysis pipeline)进行分析。

[0929] 此过程可以识别TCR先导(lead),这些TCR先导可进一步评价在癌症治疗中的潜在用途。

[0930] 1.生成变体核苷酸序列文库

[0931] 如本文的一个或多个实施例中所述选择TCRα链和TCRβ链,并组合配对,从而产生TCR变体的总集合。由于组合配对,任何给定的TCR变体只能通过确定TCRα和TCRβ可变序列来明确地识别。这可能涉及对约1.5kb的PCR扩增子进行测序。

[0932] 2.将变体核苷酸序列文库导入报告细胞

[0933] 将步骤1)中描述的组合TCR文库转染到产生病毒的Phoenix-ampho细胞中,以产生编码文库中存在的所有TCR变体的逆转录病毒。该病毒随后用于转导缺乏内源性TCR表达的Jurkat T细胞。一种杀稻瘟菌素选择标志物在TCR文库中共同编码,其允许对表达TCR池的转导的报告T细胞进行抗生素选择。

[0934] 3.基于至少一种功能特性的报告细胞的选择

[0935] 以低密度接种表达多种TCR的报告T细胞的多克隆混合物,随后与表达潜在新生抗原的永生B细胞共培养。报告T细胞的数量使得所有TCR变体的平均覆盖率至少为100。随后,收获共培养物,并基于T细胞活化标志物CD69的高表达或低表达进行FACS分选。收获并对这些“顶部”和“底部”群体分别进行进一步分析。其他活化标志物可用于FACS-分选,其中CD62L、CD137、IFN-γ、IL-2、TNF-α和GM-CSF可替代CD69或与CD69组合以选择活化的报告T细胞。此外,各种启动子活性报道基因可用于选择活化的细胞。

[0936] 4.从选定的报告细胞中分离变体核苷酸序列

[0937] 分离基因组DNA并对编码TCR的逆转录病毒插入物进行PCR扩增。PCR引物位于逆转录病毒载体中和TCRα链的恒定区,产生约1500个碱基的扩增子。使用足够的基因组DNA以至少100的平均覆盖率代表所有TCR变体。使扩增最小化,以防止特定TCR变体扩增中的偏差,但对于文库中所代表的每种TCR,应产生至少10000的平均覆盖率。对顶部和底部样本均进行基因组DNA的TCR扩增。

[0938] 5.测定分离的变体核苷酸序列的总核苷酸序列的至少600bp

[0939] 进一步处理扩增的TCR序列用于Oxford Nanopore测序。在第一个PCR反应中,使用了加尾引物(tailed primer)。这些引物含有一个新的结合位点,可用于使用经过快速连接化学修饰的带条形码的外部引物进行第二次PCR。顶部和底部样本上的PCR使用不同的条形码。在所有PCR步骤中,使扩增最小化,以防止特定TCR变体扩增中的偏差。然而,使用足够的PCR产物以至少10000的平均覆盖率来代表所有TCR变体。

[0940] 在上述方案的替代方案中,进一步处理扩增的TCR序列用于Oxford Nanopore测序。在第一个PCR反应中,用未加尾的引物扩增TCR盒。在随后的一轮PCR中,使用了加尾引

物。这些引物含有一个新的结合位点,可用于使用经过快速连接化学修饰的带条形码的外部引物进行第三次PCR。顶部和底部样本上的PCR使用不同的条形码。在所有PCR步骤中,使扩增最小化,以防止特定TCR变体扩增中的偏差。然而,使用足够的PCR产物以至少10000的平均覆盖率来代表所有TCR变体。本领域普通技术人员将认识到,可以采用替代的PCR扩增策略。

[0941] 将来自顶部和底部样本的带有条形码PCR产物按等摩尔比池化 (pooled), 并且在最后一步中, 将快速1D测序衔接子连接到该池上以得到准备测序的文库制备物 (library preparation)。将该文库加载到Oxford Nanopore R9.4.1流通池上 (flow cell), 并对文库中编码的每个TCR进行测序, 直至至少100个读段 (read) 的平均覆盖率。

[0942] 对于生物信息学分析, 使用Oxford Nanopore guppy工具包。使用guppy_basecaller从原始数据中提取序列读段 (sequence read)。使用guppy_barcode对样本进行分用 (demultiplex)。或者, 对于基于GridIon的测序, 使用MinKnow软件包获得分用的序列读段。使用guppy_aligner将序列读段与由单个 α 和 β 链序列组成的参考物进行比对, 并从得到的bam比对文件中提取每个读段的 α 链和 β 链身份 (identity)。在最后一步中, 计算每种TCR的出现频率, 并用于进一步分析。

[0943] 6. 至少一种目的变体核苷酸序列的选择

[0944] 为了识别目的变体核苷酸序列, 通过测定正向选择的 (标志物分子阳性) 报告细胞群中的变体富集和负向选择的 (标志物分子阴性) 报告细胞群中的变体耗竭 (通过测定两种细胞群中的变体读段计数) 基于相对富集来选择TCR。

[0945] 实施例11

[0946] 本实施例描述了从编码变体CAR蛋白结构域的核酸序列的组合文库中识别功能性CAR变体的筛选方法。

[0947] 为了识别具有最佳功能特性的CAR设计, 可以以组合方式组装CAR分子结构域。CAR分子由 (i) 一个抗原结合结构域, (ii) 一个铰链结构域, (iii) 一个跨膜结构域和 (iv) 一个细胞内信号传导结构域 (通常由2-3个信号模块组成) 组成, 形成一个约1.5kb的合成分子。可以将CAR变体文库导入报告细胞中, 并通过用抗原表达细胞刺激, 可在遗传筛查中选择表达CAR变体导致所需活化表型的报告细胞。鉴于几个分子结构域的组合装配, 每个变体只能通过确定所有可变分子部分的序列来明确识别。因此, 在遗传筛查过程中分离的报告细胞中编码的转基因作为约1.5kb的PCR扩增子被回收, 使用Oxford Nanopore测序法进行全长测序, 并使用定制的 (customized) 生物信息分析pipeline进行分析。

[0948] 该过程可以识别可进一步评价潜在治疗用途的CAR先导, 例如在癌症中。

[0949] 1. 变体核苷酸序列文库的生成

[0950] CAR变体文库是通过组合组装以下几个CAR蛋白结构域生成的: 2个铰链结构域、12个跨膜结构域和13个信号传导结构域 (每个变体中包含3个信号传导结构域), 产生具有超过50000个蛋白变体的文库。为了明确地识别任何给定的CAR变体, 可以对1.3kb的PCR扩增子进行测序。

[0951] 2. 将变体核苷酸序列文库导入报告细胞

[0952] 将步骤1中描述的CAR变体文库转染到产生病毒的Phoenix-ampho或293T细胞中, 以分别产生编码文库中存在的所有CAR变体的逆转录病毒或慢病毒。该病毒随后用于转导

永生化的Jurkat T细胞或体外激活的原代人T细胞。细胞表面标志物和/或抗生素选择标志物在CAR变体文库内共编码,这允许选择表达CAR变体池的转导的报告T细胞。

[0953] 3. 基于至少一种功能特性的报告细胞的选择

[0954] 用细胞增殖染料标记表达CAR变体文库的报告T细胞的多克隆混合物,以低密度接种,并将所述多克隆混合物与表达CAR抗原结合结构域的同源配体的抗原呈递细胞共培养。报告T细胞的使用量使得所有CAR变体的平均覆盖率至少为100。随后,收获共培养物并基于已分裂至少一次或未分裂的T细胞进行流式细胞术分选。收获并进一步分析这些各自的“顶部”和“底部”群。其他活化标志物,例如CD69、CD137、IFN- γ 、IL-2、TNF- α 和GM-CSF可以单独或组合用于基于流式细胞术的应答和非应答T细胞的分选。此外,各种转录因子活性报道基因(NF- κ B,NFAT,AP-1)、信号转导报道基因(ZAP70,ERK1/2磷酸化)或细胞毒性报道基因(CD107A表达)可用于选择应答和非应答T细胞。

[0955] 4. 从选定的报告细胞中分离变体核苷酸序列

[0956] 从选择的表达CAR变体的报告T细胞中,分离基因组DNA,并对编码CAR的逆转录病毒或慢病毒插入物进行PCR扩增。PCR引物与CAR插入物的不变区域结合,产生约1300个碱基的平均扩增子大小。

[0957] 或者,可以应用类似的基于唯一分子标识符(UMI)的方法。使用足够的基因组DNA以用至少为100的平均覆盖率来代表所有CAR变体。使PCR循环次数最小化,以防止特定CAR变体扩增中的偏差,但对于文库中所代表的每种CAR,应产生至少10000的平均覆盖率。对应答和非应答报告T细胞(顶部和底部样本)进行基因组DNA的CAR变体扩增。

[0958] 5. 测定分离的变体核苷酸序列的总核苷酸序列的至少600bp

[0959] 扩增的CAR序列经进一步处理后用于Oxford Nanopore测序。在第一个PCR反应中,使用了加尾引物。这些引物含有一个新的结合位点,可用于使用经过快速连接化学修饰的带条形码的外部引物进行第二次PCR。顶部和底部样本上的PCR使用不同的条形码。在所有PCR步骤中,使扩增最小化,以防止特定CAR变体扩增中的偏差。然而,使用足够的PCR产物以至少10000的平均覆盖率来代表所有CAR变体。

[0960] 将来自顶部和底部样本的带有条形码的PCR产物按等摩尔比池化,并且在最后一步中,将快速1D测序衔接子连接到该池上以得到准备测序的文库制备物。将该文库加载到Oxford Nanopore R9.4.1流通池上(flow cell),并对文库中编码的每一个CAR进行测序,直至至少100个读段的平均覆盖率。

[0961] 对于生物信息学分析,使用Oxford Nanopore guppy工具包。使用guppy_basecaller从原始数据中提取序列读段。使用guppy_barcode对样本进行分用。使用guppy_aligner将序列读段与由单个CAR变体序列组成的参考物进行比对,并从得到的bam比对文件中提取每个读段的CAR变体身份(identity)。在最后一步中,计算每种CAR变体的出现频率,并用于进一步分析。作为规避PCR扩增偏差的一种选择,可以应用基于UMI的CAR变体计数。

[0962] 6. 至少一种目的变体核苷酸序列的选择

[0963] 为了识别目的变体核苷酸序列,通过测定正向选择的(标志物分子阳性)报告细胞群中的变体富集和负向选择的(标志物分子阴性)报告细胞群中的变体耗竭(通过测定两种细胞群中的变体读段计数)基于相对富集来选择CAR。

[0964] 实施例12

[0965] 本实施例描述了通过如实施例10和实施例11所述的本筛选方法筛选得到的细胞数的计算。

[0966] 对于大小为100个组合的组合文库,通过在选择阳性应答者时恢复10000个细胞,并恢复10000个阴性应答者,可实现100倍的覆盖率(100xcoverage)。因此,选择前的细胞群可大于20000。

[0967] 对于大小为1000个组合的组合文库,通过在选择阳性应答者时恢复100000个细胞,并恢复100000个阴性应答者,可实现100倍的覆盖率(100xcoverage)。因此,选择前的细胞群可大于200000。

[0968] 对于大小为 1×10^9 个组合的组合文库,通过在选择阳性应答者时恢复 1×10^{11} 细胞,并恢复 1×10^{11} 阴性应答者,可实现100倍的覆盖率(100xcoverage)。因此,选择前的细胞群可大于 2×10^{11} 。

[0969] 实施例13

[0970] 本实施例描述了从错配型结直肠癌(MMRp-CRC)肿瘤中恢复新生抗原特异性T细胞受体序列(图10A)。

[0971] 从4份新鲜冷冻的MMRp-CRC肿瘤样本中分离出DNA和RNA,并以两种方式使用:

[0972] 首先,使用DNA和/或RNA进行批量TCR α -链和TCR β -链测序(由MiLaboratory进行;莫斯科/俄罗斯)。所得到的TCR链序列的集合被分成TCR α -链序列集合和TCR β -链序列集合,导致每个样本约10000-30000个TCR α -链和TCR β -链的集合(图10B和11A)。

[0973] 图10B和图11A描述了人肿瘤样本中浸润淋巴细胞的批量TCR测序。通过Milaboratory对四份人MMRp-CRC肿瘤样本进行了批量TCR测序。在比对和TCR识别后,克隆型根据其CDR3氨基酸序列及其V和J同一性(identity)收缩(collapsed)。独特克隆型的数量表示每个肿瘤样本的 α 链和 β 链二者。

[0974] 任何无效的TCR链序列,其中TCR片段(也称为TCR基因元件)在氨基酸序列水平上框外连接,和/或其中引入终止密码子,和/或其中存在移码突变,和/或其中存在缺陷剪接位点,都被从集合中去除。使用编码特定TCR链的核酸分子的读段计数或唯一分子标识符(UMI)计数(或分别在总TCR α -链或TCR β -链中的相应百分比)以降序对每个集合进行分类,以获得TCR α -链和TCR β -链的秩次序。

[0975] 选择前100个丰度最高的TCR α -链和TCR β -链用于TCR文库生成(由Twist Bioscience进行;旧金山/美国)。简而言之,所选的TCR α -链和TCR β -链是通过DNA合成产生的片段。为了生成文库,将所有选择的TCR α 和TCR β 片段组合连接到编码TCR β -P2A-TCR α 盒的连续核酸分子上(图11BTCR表达质粒的示意图。可以使用含有TCR β 链和TCR α 链二者以及杀稻瘟菌素选择标志物的逆转录病毒构建体作为构建文库的支架)。使用这种构建体的逆转录病毒转导最终导致单一转录物的表达,最终导致单一转录物的表达,由于在2A位点处的肽切割,这导致TCR β 链和 α 链,以及产生的杀稻瘟菌素抗性标志物的翻译)。每个盒可以以TCR β -P2A-TCR α -T2A-杀稻瘟菌素抗性的格式连接一个TCR β -和TCR α -片段(图11B)(SEQ ID NO:4)(图43)。产生的TCR文库通常包含10000个TCR变体,在非常窄的频率范围内没有丢失(图10C;图11C和11D)。100 $\times$ 100文库质控。对于来自11A)的每个样本,选择100个最普遍的 α 链和 β 链,并用于创建组合文库。以由Twist Bioscience进行的组合克隆方法,合成了图

11B) 中的TCR- β 和TCR- α 区域,并将其插入图11B)中的载体中。使用可变TCR α 链和 β 链结构域两侧的引物对来扩增两条链,并使用纳米孔测序来识别两条链的身份(identity)。

[0976] 对于每个患者文库,示出了10000个 $\alpha \times \beta$ 组合中每一个的表征(representation)(图11D)。患者文库的TCR表征(representation)的特征。对于11C)中的每个患者文库,示出了每个TCR的读段的数量范围、平均覆盖率以及落在中位数 $\pm 2 \log$ 单位范围内的TCR百分比。其次,使用来源于肿瘤以及健康组织的DNA和RNA,使用全外显子组测序(WES)和RNA测序来确定肿瘤特异性突变集,以建立表达的突变基因集。识别和选择肿瘤特异性突变的管道可以提供选择。基于WES和RNA-seq识别这种突变的方法已在前面描述过,并且是本领域技术人员已知的。随后,产生以串联阵列编码多种肿瘤衍生的突变肽串联小基因(TMG)构建体。TMG构建体将12个单独的肿瘤突变编码为25mer多肽小基因,这些基因被连接并包括LAMP-1信号转导和跨膜结构域(例如在Gros等人,Nat Med 2016中描述的;pt4除外)以及使用2A元件与LAMP-1胞质结构域融合的嘌呤霉素抗性标志物。对于Pt4,使用编码33个或34个不同的串联小基因,不含LAMP-1信号转导和跨膜结构域,但含有嘌呤霉素抗性标志物的串联小基因设计。同时,使用来自MMRp-CRC患者的匹配自体血液来产生永生化B细胞。人B细胞的EBV永生化是本领域技术人员已知的(例如,Traggiai等人,Methods Mol Biol 2012)。使用永生化的自体B细胞通过用编码TMG的病毒颗粒进行逆转录病毒转导来产生表达抗原的B细胞。文献中已描述了用除TMG之外的其他基因进行原代人B细胞逆转录病毒转导的方案(如Kwakkenbos等人,Nat Med 2010),这些方案是本领域技术人员已知的。可以基于对嘌呤霉素的抗性来选择TMG转导的B细胞。嘌呤霉素选择的方法是本领域技术人员已知的。使用Fugene转染试剂和本领域技术人员已知的方案将TCR文库转染到Phoenix-Ampho病毒产生细胞(ATCC CRL-3213)中。产生的逆转录病毒病毒颗粒用于转导Jurkat报告T细胞系。Jurkat报告T细胞系缺乏内源性TCR表达(例如Mezzadra等人,Nature 2017中描述了这种基因敲除的产生),并且使用本领域技术人员已知的方法在用CD8 α -P2A-CD8 β 转基因(SEQ ID NO:2)转导后被修饰以表达人CD8 α 和CD8 β 。用TCR文库转导Jurkat报告T细胞,得到占总活T细胞的10-60%的经TCR修饰的T细胞(图10D)。在TCR文库中使用鼠TCR恒定结构域序列允许通过使用鼠TCR β 恒定结构域特异性抗体的流式细胞术检测TCR修饰的Jurkat T细胞。通过向细胞培养基中添加杀稻瘟菌素,将经修饰以表达TCR β -P2A-TCR α -T2A-杀稻瘟菌素抗性转基因的Jurkat T细胞正向选择至高纯度(图10D)。遗传修饰细胞的抗生素选择是本领域普通技术人员已知的。

[0977] 接下来,通过与TMG转导的B细胞共培养来刺激TCR转导的Jurkat T细胞。在共培养前3天,以低密度(0.1×10^6 个细胞/ml)接种Jurkat报告T细胞。以相等(1:1)的比例将表达不同TMG的B细胞混合。在总体积为72ml的培养基中,将 90×10^6 个池化的B细胞(或缺乏TMG表达的对照B细胞)与 90×10^6 个Jurkat报告T细胞(比例为1:1)混合。将200 μ l(每孔 0.5×10^6 个细胞)的该混合物分布在约360孔经TC处理的U形底96孔板上。将板以1000rpm离心1分钟,并在37 $^{\circ}$ C下孵育20-22小时。

[0978] 随后,收获共培养物,并用荧光染料标记的CD20和CD69抗体对共培养物进行染色。接下来,使用含有4%甲醛的固定缓冲液固定细胞。使用BD Biosciences ARIA Fusion流式细胞仪分选器,针对每种刺激条件对下列群体进行分选(图10E):

[0979] 淋巴细胞、单细胞、CD20 $^{-}$ 、CD69 $^{\text{高}}$ (基于CD69荧光信号,CD69 $^{\text{高}}$ 包括最高的10-15%的

单细胞、CD20⁺细胞)

[0980] 淋巴细胞、单细胞、CD20⁺、CD69^低(基于低CD69荧光信号,CD69^低包括10-15%的单细胞、CD20⁺细胞)

[0981] 从分选的TCR转导的Jurkat T细胞中分离基因组DNA,并使用所述基因组DNA作为模板,使用本领域技术人员已知的PCR方法,以有限循环次数的多轮PCR扩增TCRβ-P2A-TCRα盒的一部分。所得PCR产物的大小约为1.5kb(图10F),并使用Oxford Nanopore MinIon或GridIon测序仪进行测序。至此,整个筛选程序在最小覆盖率为100倍(100x)的情况下进行,在TMG表达的情况下进行三次重复,在没有TMG表达的情况下进行三次重复。使用DESeq2 R套装软件恢复TCRα链和β链身份(identity)并识别差异表达的TCR组合。在B细胞中存在和不存在TMG表达的情况下,平均Rlog转换读段计数可用于描述新生抗原反应性TCR先导。新生抗原反应性TCR先导被描绘为被圆圈包围的较大黑点(图10G)。

[0982] 为了对TCR先导识别的TMG进行解卷积,使用在单次复制中表达单个TMG构建体的B细胞(而不是表达TMG的B细胞的混合物)筛选TCR文库(图10H)。随后,为了确定由TCR先导识别的确切新生抗原,使用了装载有在TMG中编码的单个25mer肽的B细胞(图10I和10J)。使用表达在TMG中编码的单个小基因的B细胞测定由TCR先导识别的确切新生抗原(图10K和10L)。

[0983] 综上所述,本实施例表明所描述的平台能够成功地从新鲜冷冻的肿瘤材料中识别出新生抗原特异性TCR。

[0984] 实施例14

[0985] 本实施例描述了从黑色素瘤肿瘤样本中恢复TCR序列用于产生患者特异性TCRαβ文库(图12A-12C)。

[0986] 从两份新鲜冷冻的黑色素瘤肿瘤样本中分离DNA和RNA,并以以下两种方式使用:

[0987] 首先,使用DNA和/或RNA进行批量TCRα-链和TCRβ-链测序(由Milaboratory进行;莫斯科/俄罗斯)。所得到的TCR链序列的集合被分成TCRα-链序列的集合和TCRβ-链序列的集合,导致每个样本约5000-10000个TCRα-链和TCRβ-链的集合(图12A)。通过Milaboratory对两份人黑色素瘤肿瘤样本进行了批量TCR测序。在比对和TCR识别后,克隆型根据其CDR3氨基酸序列及其V和J同一性(identity)收缩(collapsed)。独特克隆型的数量表示每个肿瘤样本的α链和β链二者。

[0988] 任何无效的TCR链序列,其中TCR片段(也称为TCR基因元件)在氨基酸序列水平上框外连接,和/或其中引入终止密码子,和/或其中存在移码突变,和/或其中存在缺陷剪接位点,都被从集合中去除。使用编码特定TCR链的核酸分子的读段计数或唯一分子标识符(UMI)计数(或分别在总TCRα-链或TCRβ-链中的相应百分比)以降序对每个集合进行分类,以获得TCRα-链和TCRβ-链的秩次序。选择前100个丰度最高的TCRα-链和TCRβ-链用于TCR文库生成(由Twist Bioscience执行;旧金山/美国)。简而言之,所选的TCRα-链和TCRβ-链是通过DNA合成产生的片段。为了生成文库,将所有选择的TCRα和TCRβ片段组合连接到编码TCRβ-P2A-TCRα盒的连续核酸分子上(图11B TCR表达质粒的示意图。可以使用含有TCRβ链和TCRα链二者以及杀稻瘟菌素选择标志物的逆转录病毒构建体作为构建文库的支架)。使用这种构建体的逆转录病毒转导最终导致单一转录物的表达,由于在2A位点处的肽切割,这导致TCRβ链和α链,以及产生的杀稻瘟菌素抗性标志物的翻译)。每个盒可以以TCRβ-P2A-

TCR α -T2A-杀稻瘟菌素抗性的格式连接一个TCR β -和TCR α -片段(图11B)(SEQ ID NO:4)。产生的TCR文库通常将包含10000个TCR变体,在非常窄的频率范围内没有丢失(图12B)。100 $\times$ 100文库质控。对于来自12A)的每个样本,可以选择100个最普遍的 α 链和 β 链,并用于创建组合文库。可以通过Twist Bioscience进行的组合克隆方法合成12A)中的TCR- β 和TCR- α 区域,并将其插入11B)中的载体中。可以使用可变TCR α 链和 β 链结构域两侧的引物对来扩增两条链,并使用纳米孔测序来识别两条链的身份。对于每个患者文库,示出了10000个 $\alpha\times\beta$ 组合中每一个的表征(representation)。12C)患者文库的TCR表征(representation)的特征。对于12B)中的每个患者文库,示出了每个TCR的读段的数量范围、平均覆盖率以及落在中位数 $\pm 2\log$ 单位范围内的TCR百分比。

[0989] 综上所述,本实施例表明组合TCR文库可以基于批量TCR测序来合成和克隆。

[0990] 实施例15

[0991] 本实施例描述了从通过混合TCR质粒生成的TCR文库中恢复抗原特异性TCR。

[0992] 通过将各自编码单个特征性TCR的6个质粒与各自编码单个非特征性TCR的24个质粒混合生成TCR文库(图13A)。TCR表达质粒设计如图15A所示。使用Fugene转染试剂和本领域技术人员已知的方案,将该TCR文库转染到Phoenix-Ampho病毒产生细胞(ATCC CRL-3213)中。产生的逆转录病毒病毒颗粒用于转导Jurkat报告T细胞系。Jurkat报告T细胞系缺乏内源性TCR表达(例如在Mezzadra等人,Nature 2017中描述了这种基因敲除的产生),并且使用本领域技术人员已知的方法在用CD8 α -P2A-CD8 β 转基因(SEQ ID NO:2)转导后被修饰以表达人CD8 α 和CD8 β 。用TCR文库转导Jurkat报告T细胞,得到占总活T细胞的80%的经TCR修饰的T细胞。在TCR文库中使用鼠TCR恒定结构域序列(图15A)允许通过使用鼠TCR β 恒定结构域特异性抗体的流式细胞术检测TCR修饰的Jurkat T细胞。通过向细胞培养基中添加嘌呤霉素,正向选择经修饰以表达TCR β -P2A-TCR α -T2A-嘌呤霉素抗性转基因的Jurkat T细胞至高纯度(图13B)。遗传修饰细胞的抗生素选择是本领域普通技术人员已知的。

[0993] 接下来,通过与TMG转导的B细胞共培养来刺激TCR转导的Jurkat T细胞。在共培养前3天,以低密度(0.1×10^6 个细胞/ml)接种Jurkat报告T细胞。在总体积为72ml的培养基中,将表达TMG的 90×10^6 个池化B细胞(或缺乏TMG表达的对照B细胞)与 90×10^6 个Jurkat报告T细胞(比例为1:1)混合。将200 μ l(每孔 0.5×10^6 个细胞)的该混合物分布在约360孔经TC处理的U形底96孔板上。将板以1000rpm离心1分钟,并在37 $^{\circ}$ C下孵育20-22小时。

[0994] 随后,收获共培养物,并用荧光染料标记的CD20和CD69抗体对共培养物进行染色。然后使用含有4%甲醛的固定缓冲液固定细胞。使用BD Biosciences ARIAFusion流式细胞仪分选器,针对每种刺激条件对下列群体进行了分选(图13C):

[0995] 1. 淋巴细胞、单细胞、CD20 $^{-}$ 、CD69 $^{\text{高}}$ (基于CD69荧光信号,CD69 $^{\text{高}}$ 包括最高10-15%的单细胞、CD20 $^{-}$ 细胞)

[0996] 2. 淋巴细胞、单细胞、CD20 $^{-}$ 、CD69 $^{\text{低}}$ (基于低CD69荧光信号,CD69 $^{\text{低}}$ 包括10-15%的单细胞、CD20 $^{-}$ 细胞)

[0997] 从分选的TCR转导的Jurkat T细胞中分离基因组DNA,并使用所述基因组DNA作为模板,使用本领域技术人员已知的PCR方法,以有限循环次数的多轮PCR扩增TCR β -P2A-TCR α 盒的一部分。所得PCR产物的大小约为1.5kb(图13D),并且可以使用Oxford Nanopore MinIon或GridIon测序仪使用本领域技术人员已知的技术进行测序。使用本领域技术人员

已知的比对技术恢复TCR身份(identity),并且对于每一个TCR,将顶部样本相对于底部样本的归一化读段计数的log2转换的倍数富集表示为log10转换的TCR频率的函数(图13E TMG表达的情况下的单次重复:在没有TMG表达的情况下的单次重复)。

[0998] 综上所述,本实施例表明可从TCR质粒的混合物中分离抗原反应性TCR。

[0999] 实施例16

[1000] 本实施例描述了从通过基因合成产生的TCR文库中恢复抗原特异性TCR。

[1001] 选择五种已知抗原反应性的特征性TCR,以及45或95种非特征性TCR(分别用于50×50和100×100文库;图14A)用于生成TCR文库(由Twist Bioscience进行;旧金山/美国)。简而言之,来自这些TCR的TCR α -链和TCR β -链是通过DNA合成作为片段合成的。为了生成文库,将所有选择的TCR α 和TCR β 片段组合连接到编码TCR β -P2A-TCR α 盒(SEQ ID 1)的连续核酸分子上。使用含有TCR β 链和TCR α 链二者以及嘌呤霉素选择标志物的逆转录病毒构建体作为构建文库的支架。使用这种构建体的逆转录病毒转导最终最终导致单一转录物的表达,由于在2A位点处的肽切割,这导致TCR β 链和 α 链,以及产生的嘌呤霉素抗性标志物的翻译。每个盒以TCR β -P2A-TCR α -T2A-嘌呤霉素抗性的格式(SEQ ID NO:1)连接一个TCR β -和TCR α -片段。通过由Twist Bioscience进行的组合克隆方法合成了50或100个TCR中的每一个的TCR- β 和TCR- α 区域,并将其插入该载体中。

[1002] 使用Fugene转染试剂和本领域技术人员已知的方案将50×50文库转染到Phoenix-Ampho病毒产生细胞(ATCC CRL-3213)中。产生的逆转录病毒病毒颗粒用于转导Jurkat报告T细胞系。在用CD8 α -P2A-CD8 β 转基因(SEQ ID NO:2)转导后,使用本领域技术人员已知的方法修饰Jurkat报告T细胞以表达人CD8 α 和CD8 β 。用TCR文库转导Jurkat报告T细胞,得到占总活T细胞的15%的经TCR修饰的T细胞(基于转导后4天的染色——嘌呤霉素选择后和检测当天的纯度>60%)。在TCR文库(SEQ ID NO:1)中使用鼠TCR恒定结构域序列允许通过使用鼠TCR β 恒定结构域特异性抗体的流式细胞术检测TCR修饰的Jurkat T细胞。通过向细胞培养基中添加嘌呤霉素,正向选择经修饰以表达TCR β -P2A-TCR α -T2A-嘌呤霉素抗性转基因的Jurkat T细胞至高纯度。遗传修饰细胞的抗生素选择是本领域普通技术人员已知的。

[1003] 接下来,用经工程改造以各种方式呈递抗原的APC刺激TCR转导的Jurkat T细胞。APC包括装载肽的JY细胞、各表达不同小基因的EBV-LCL细胞系的混合物、表达TMG的EBV-LCL细胞以及未被工程改造以表达特定抗原的EBV-LCL。在共培养前3天,以低密度(0.1×10^6 个细胞/ml)接种Jurkat报告T细胞。将APC与Jurkat报告T细胞以1:1的比例以 2.5×10^6 个细胞/ml的浓度混合。将200 μ L(每孔 0.5×10^6 个细胞)分布在约40孔经TC处理的U形底96孔板上。将板以1000rpm离心1分钟,并在37℃下孵育20-22小时。

[1004] 随后,收获共培养物,并用荧光染料标记的CD8、CD20和CD69抗体对共培养物进行染色。使用BD Biosciences ARIAFusion流式细胞仪分选器,针对每种刺激条件对下列群进行分选(图14B):

[1005] 1.淋巴细胞、单细胞、活细胞、CD20⁻、CD8⁺、CD69^高(基于CD69荧光信号,CD69^高包括最高10-15%的单细胞、活细胞、CD20⁻、CD8⁺细胞)

[1006] 2.淋巴细胞、单细胞、活细胞、CD20⁻、CD8⁺、CD69^低(基于低CD69荧光信号,CD69^低包括10-15%的单细胞、活细胞、CD20⁻、CD8⁺细胞)

[1007] 从分选的TCR转导的Jurkat T细胞中分离基因组DNA,并使用所述基因组DNA作为模板,使用本领域技术人员已知的PCR方法,以有限循环次数的多轮PCR扩增TCR β -P2A-TCR α 盒的一部分。所得PCR产物的大小约为1.5kb(图14C),并使用Oxford Nanopore MinIon测序仪使用本领域技术人员已知的技术进行测序。使用本领域技术人员已知的比对技术恢复TCR身份(identity),对于每一个TCR,将顶部样本相对于底部样本的归一化读段计数的倍数富集(y轴)表示为平均TCR表征(representation)的函数(x轴;图14D),14D;在外源性抗原存在的情况下,每个条件的单次重复;在缺乏外源抗原表达情况下,每个条件的单次重复)。使用所描述的DESeq2 R套装软件识别在相对于“底部”的“顶部”样本中富集的TCR组合,使用假设富集的TCR被定义为相对于不存在抗原的“顶部”和“底部”样本,在存在抗原时在“顶部”样本中富集,在“底部”样本中耗竭的线性模型。示出了TCR α 链和 β 链身份,以及关键的统计度量(图14E)。

[1008] 以类似于50 $\times$ 50筛选的方式进行100 $\times$ 100文库筛选,不同之处在于将Jurkat报告T细胞工程改造为缺乏内源性TCR表达并外源性表达人CD8 α 和CD8 β 。与50 $\times$ 50文库筛选相比,在表达TMG的B细胞的情况下进行了三次重复筛选,在未被工程改造以表达TMG的B细胞的情况下进行了一次重复。用TCR文库转导Jurkat报告T细胞,得到占总活T细胞的14%的经TCR修饰的T细胞。使用DESeq2 R软件包计算Rlog转换的读段计数,并从顶部样本的所有重复的每个TCR的平均Rlog值中减去底部样本的所有重复的每个TCR的平均Rlog值,并示出了存在(x轴)和不存在(y轴)由B细胞表达的TMG的情况下进行的共培养(图14F)。5个掺入特征的TCR被描绘为被圆圈包围的较大黑点。Wald统计量被用作排序概率测量图的度量(图14G)。

[1009] 综上所述,本实施例表明可以筛选50 $\times$ 50和100 $\times$ 100设计的组合文库以识别抗原反应性TCR。

[1010] 实施例17

[1011] 本实施例描述了使用基因合成来创建TCR库。

[1012] 选择五种已知抗原反应性的特征性TCR以及95种非特征性TCR用于生成TCR文库(由Twist Bioscience进行;旧金山/美国)。简而言之,所选的TCR α -链和TCR β -链是通过DNA合成产生的片段。为了生成文库,将所有选择的TCR α 和TCR β 片段组合连接到编码TCR β -P2A-TCR α 盒的连续核酸分子上(SEQ ID NO:1;可以使用含有TCR β 链和TCR α 链二者以及嘌呤霉素选择标志物的逆转录病毒构建体作构建文库的支架。使用这种构建体的逆转录病毒转导最终导致单一转录物的表达,由于在2A位点处的肽切割,这导致TCR β 链和 α 链,以及产生的嘌呤霉素抗性标志物的翻译。每个盒以TCR β -P2A-TCR α -T2A-嘌呤霉素抗性的格式(SEQ ID NO:1,图15A)连接一个TCR β -和TCR α -片段。可以使用可变TCR α 链和 β 链结构域两侧的引物对来扩增两条链,使用纳米孔测序来识别两条链的身份。TCR链的识别是本领域技术人员已知的。对于100 $\times$ 100文库,示出了10000个 $\alpha\times\beta$ 组合中每一个的表征(representation)。得到的TCR文库包含所有10000个TCR变体,在非常窄的频率范围内没有损失(图15B)。

[1013] 设想了组成组合文库的替代方式,其中可以通过多个重叠或非重叠的组合子文库来创建更高复杂度的文库,这些子文库共同地但不是单独地表示需要存在于复合文库中的所有TCR组合(图15C)。或者,可以通过组成一个更大的由更小的组合子文库组成的文库,降低文库的复杂度,使得不是复合文库中的所有 α 链和所有 β 链的所有可能组合都被表示。使

用配对信息或配对可能性的信息,本领域技术人员可以以这样的方式设计组合子文库,使得配对的链或可能构成一对的链都包含在一个组合子文库中(图15D)。

[1014] 图15E描述了存在于通过合成四个 100×100 文库并以1:1:1:1的比例混合而产生的两个 200×200 TCR文库中的TCR组合的识别。每种可能的TCR组合(x轴)的出现表示为其密度(y轴)。中位数 $\pm \log^2$ 单位范围内的TCR组合的数量分别为92%和88%。对于两个患者文库(图15E)测试了通过以1:1:1:1的比例混合四个 100×100 库(如图15C所示)来创建 200×200 库的原理。这表明可以通过混合多个复杂度较低的组合TCR文库来创建复杂度较高的组合TCR文库。

[1015] 图15F示出了在 200×200 文库筛选中存在和不存在新生抗原时的pt2TCR反应性。 200×200 文库是通过组合连接193个最常表达的TCR α 和TCR β 链而创建的,如使用来自图10A)-10H)的批量TCR测序数据所测得的)。此外,还包括7条先前已确定特征的TCR α 和TCR β 链。 200×200 文库的筛选类似于10A)-10H)中所述的 100×100 文库筛选,不同之处在于gDNA分离后的覆盖率在57-133的范围内,并且在表达TMG的情况下进行了两次重复,在没有TMG的情况下进行了两次重复。另一个不同之处在于,在进行FACS分选之前,使用抗CD20或抗-Ly6G微珠将B细胞耗尽。TCR α 链和 β 链识别后,使用DESeq2 R套装软件识别了差异表达的TCR组合。差异表征分析是本领域技术人员已知的。图15F描述了 200×200 文库筛选中存在(x轴)和不存在(y轴)由B细胞表达的TMG时的平均Rlog转化读段计数。六个掺入特征的TCR被描述为较大的深灰色圆点。一个特征性TCR未被表示,因为它仅限于在pt2 EBV-LCL中不表达的HLA-等位基因。使用 100×100 文库筛选在10A)-10H)中识别的TCR用较大的浅灰点表示。i)在 200×200 文库筛选中识别出的和ii)在 100×100 库中未识别出的其他TCR先导被表示为较大的中间阴影化的灰点。

[1016] 这表明在 200×200 筛选方法中识别出了被掺杂(spiked)到文库中的六个TCR。此外,这表明可以使用 200×200 文库筛选识别出 100×100 文库中没有表征的新TCR先导。因此,与使用 100×100 TCR文库的筛选相比,在饱和文库筛选中可从 200×200 或更复杂的TCR文库中识别出更多的新生抗原反应性TCR。

[1017] 图15G表示六种特征性TCR的统计度量,以及在10A)-10H)中使用 100×100 文库筛选在 100×100 和 200×200 文库筛选中识别的TCR。在10A)-10H)中使用 100×100 文库筛选识别的TCR的表征(representation),如DESeq2碱基平均值度量(BaseMean metric)所测量的,在 200×200 筛选中比在 100×100 筛选中低6倍以上。这可能是未能识别这些TCR的原因,并强调了TCR文库中TCR变体的同等表征(representation)的重要性。

[1018] 综上所述,本实施例显示了如何使用基因合成来创建TCR文库,以及如何基于组合多个组合子文库的思想来增加或降低文库的复杂度。

[1019] 实施例18

[1020] 本实施例描述了优化共培养条件以从TCR库中识别TCR α/β 对。

[1021] 在本实施例中,调节共培养条件并用于从高度稀释的抗原反应性TCR的TCR库中识别TCR α/β 对。为了测试CD69作为T细胞活化标志物用于筛选目的的适用性,将表达hCD8和CMV#1TCR的Jurkat T细胞与负载不同量CMV肽的JY细胞共培养。APC的肽负荷是本领域技术人员已知的。通过FACS测得的CD69阳性随着抗原性肽的浓度而增加(图16A)。

[1022] 为了测试接种密度是否影响Jurkat T细胞的CD69染色背景,以不同密度接种细

胞。如通过FACS测量的,低密度接种降低了CD69背景表达(图16B)。为了测试最佳共培养板,将TCR+Jurkat报告T细胞与表达相关抗原的B细胞在T75或T25烧瓶或U形底96孔板中共培养。当在U形底96孔板中进行共培养时,Jurkat T细胞的活化最为显著(图16C)。为了测试共培养密度对T细胞活化的影响,以不同密度共培养TCR+Jurkat T细胞。当共培养物中接种250000或125000个效应细胞时,Jurkat T细胞活化最高效(图16D)。为了在遗传筛查设置中测试图16A-16D中确定的共培养条件,通过混合Jurkat T细胞系建立了Jurkat报告T细胞的多克隆池,所述Jurkat T细胞系各自表达五种特征性TCR之一或二十四种非特征性(非相关性)TCR之一。如此进行混合,使得表达特征性TCR的细胞以1:10000至1:1000000的频率存在。在共培养前3天以低密度(100000个细胞/ml)接种细胞,并在U形底96孔板中以250000个细胞的接种密度共培养。

[1023] 20小时后,收获共培养物,并用荧光染料标记的CD20和CD69抗体对共培养物进行染色。接下来,使用含有4%甲醛的固定缓冲液固定细胞。使用BD Biosciences ARIA Fusion流式细胞仪分选器,针对每种刺激条件对下列群体进行分选:

[1024] 1. 淋巴细胞、单细胞、CD20⁻、CD69^高(基于CD69荧光信号,CD69^高包括最高10-15%的单细胞、CD20⁻细胞)

[1025] 2. 淋巴细胞、单细胞、CD20⁻、CD69^低(基于低CD69荧光信号,CD69^低包括10-15%的单细胞、CD20⁻细胞)

[1026] 从分选的TCR转导的Jurkat T细胞中分离基因组DNA,并使用所述基因组DNA作为模板,使用本领域技术人员已知的PCR方法,以有限循环次数的多轮PCR扩增TCRB盒。所得PCR产物的大小约为0.5kb,并且可以使用本领域技术人员已知的技术,使用Illumina测序仪进行测序。使用本领域技术人员已知的比对技术恢复TCR身份,对于每一个TCR,表示顶部样本相对于底部样本的归一化读段计数的倍数富集(图16E)。

[1027] 综上所述,本实施例表明可以使用所述共培养条件的遗传筛查可用于识别频率为1:1000000以上的抗原反应性TCR。

[1028] 实施例19

[1029] 为了能够从这种文库中以高灵敏度识别新生抗原特异性TCR,并确保TCR在不同的处理步骤中不丢失,在筛选过程中必须多次表征每个唯一的TCR,以保持TCR的覆盖率。因此,必须筛选大量经TCR转导的Jurkat细胞和APC³³。此外,一个目标是提高文库中的TCR的数量,以便能够对超过10000个TCR进行高灵敏度筛选。这突出表明,通过优化各种流程步骤来提高TCR发现平台的可扩展性的必要性,以允许在仍然保持TCR覆盖的同时更有效地处理大量细胞。因此,本研究的目的是进一步优化这些流程步骤,以提高可扩展性,同时保持TCR分离平台的灵敏度。

[1030] 在这些研究中,使用了四种已知的HLA-A*02:01限制性TCR-CDK4 TCR克隆8和17(简称CDK4-8和17)以及CMV TCR克隆1和2(简称CMV-1和2)。这两种CDK4 TCR对突变的细胞周期蛋白依赖性激酶4(CDK4_{R24C})肽具有特异性。在多个黑素瘤患者中识别出这种突变衍生的新生抗原表位³⁴。此外,两个CMV TCR靶向人巨细胞病毒(CMV)成分pp65编码的肽³⁵。对于CDK4和CMV表位,在研究中使用了具有潜在不同亲和力的两个不同的TCR克隆。这可能允许评价同源肽的TCR亲和力在筛选过程中的作用。

[1031] 作为处理大量细胞的第一步,评价了逆转录病毒转导后杀稻瘟菌素选择作为富集

表达TCR的Jurkat细胞的方法。观察到杀稻瘟菌素选择导致转导的Jurkat细胞高效富集,且毒性最小。此外,研究发现,将共培养形式从圆形底96孔(96W)板扩大到GMP袋,可实现至少 170×10^6 个效应物和靶细胞的高通量处理,同时维持效应细胞的可比活化。

[1032] 接下来,探索了用基于珠的选择步骤代替TCR文库筛选过程中的流式细胞术分选步骤的不同方法。除CD69外,在纵向共培养试验中还评价了T细胞活化标志物CD25和CD62L。CD25作为IL-2受体 α 链发挥作用,已知在活化后由T细胞表达³²,而CD62L在初始T细胞上表达,并在T细胞活化时下调,使效应T细胞能够重新进入血流³⁶。CD25和CD62L表现出有希望的表达谱(expression profiles),因此是与CD69组合的两步基于珠选择的候选物。

[1033] 最后,评价了基于NFAT的报告系统的功效。已经建立了包含多个人类NFAT结合位点、后跟一个最小启动子和一个报道基因的载体,以研究T细胞的抗原特异性应答³⁷⁻⁴⁰。最小启动子不含任何NFAT结合位点,其唯一目的是在转录因子与NFAT结合位点结合时启动报道基因的转录。用这些构建体转导的Jurkat细胞仅在T细胞活化时表达报道基因。可以引入适用于基于珠的选择的细胞表面标志物或抗生素抗性基因作为报道基因以规避流式细胞术选择步骤。然而,结果表明,使用这种报告系统导致了高背景,因此,这种报告系统不太可能被纳入筛选平台。

[1034] 结果

[1035] 肽滴定分析表明,CDK4 TCR克隆8和17之间以及CMV TCR克隆1和2之间的亲和力没有差异或差异极小。

[1036] 四种模型TCR(CDK4 TCR克隆8和17以及CMV TCR克隆1和2)对其同源肽的亲和力在将它们用于研究之前进行了表征。这可通过在共培养分析中进行肽滴定来实现。

[1037] 通过逆转录病毒基因转移将4种TCR导入Jurkat TCR敲除(KO)CD8⁺细胞系。转导后的Jurkat TCR KO细胞为35-45% $mTCR\beta^+CD8^+$ 并显示CD8⁻群(图17a,上图)。为了避免在共培养分析过程中校正转导效率,对经TCR转导的Jurkat细胞进行了分选,以获得约90%的 $mTCR\beta^+CD8^+$ 细胞群(图17a,底图)。

[1038] 使用CD69作为读出(readout),将CDK4-8和17-转导的Jurkat TCR KO细胞与载有分级浓度的CDK4突变肽的表达HLA-A*02:01的JY细胞共培养20小时,显示CDK4-17 TCR对其同源肽的亲和力略高于CDK4-8TCR(图17b,CDK4-8的EC50:43ng/ml,CDK4-17的EC50:23ng/ml)。在CMV TCR的情况下,当将表达这些TCR的Jurkat TCR KO细胞与载有分级浓度的CMVpp65肽的表达HLA-A*02:01的JY细胞共培养时,CMV-1和2TCR对CMVpp65肽表现出相当的亲和力(图17c,CMV-1的EC50:2.2ng/ml,CMV-2的EC50:1.2ng/ml)。

[1039] 因此,所述数据表明CDK4-8和17对CDK4突变肽具有相似(或最小不同)的亲和力,并且CMV-1和2对CMVpp65肽具有相当(或最小差异)的亲和力。然而,这些不同的TCR克隆作为研究TCR亲和力对筛选过程影响的工具价值有限。因此,在继续测试TCR文库之前,CDK4 TCR中的一种和CMV TCR中的一种被标准地纳入研究。

[1040] 利用杀稻瘟菌素高效筛选出具有不同的TCR转导效率的Jurkat细胞。

[1041] 在将逆转录病毒TCR文库转导至Jurkat TCR KO细胞中之后,有效且高通量的筛选程序对于成功富集经TCR转导的Jurkat细胞(>80% $mTCR\beta^+CD8^+$ 细胞)而不引起毒性和损失TCR覆盖率非常有用。抗生素选择是富集TCR转导的Jurkat T细胞的一个有吸引力的策略,并对其进行了评价。

[1042] 为此目的,最初通过使用CDK4-17 TCR评价杀稻瘟菌素选择。由于TCR文库的转导效率各不相同(通常为10-56% $mTCR\beta^+CD8^+$ 细胞),因此使用具有不同CDK4-17转导效率的Jurkat TCR KO细胞(10%和30% $mTCR\beta^+CD8^+$ 细胞)评价杀稻瘟菌素选择。与使用较低的抗生素浓度(2和4 $\mu g/ml$)和较高的起始细胞密度(0.5和 1×10^6 个细胞/ ml)相比,以 0.25×10^6 个细胞/ ml 的密度铺板并且用6 $\mu g/ml$ 杀稻瘟菌素选择7天的细胞导致最高的倍数扩增。此外,上述选择条件导致约90%的 $mTCR\beta^+CD8^+$ 细胞(数据未示出)。

[1043] 为了验证CDK4-17的最佳杀稻瘟菌素选择条件是否也可以应用于其他TCR,检测了CMV TCR中的一种,CMV-1。此外,如果在更早的时间点去除杀稻瘟菌素允许选择继续并改善倍数扩增,则进行评价。为了实现这一点,在选择的第4天,将细胞以其起始密度重新铺板在含有或不含相应浓度的杀稻瘟菌素的培养基中(分别称为“第4天去除杀稻瘟菌素”和“第4天添加新的杀稻瘟菌素”)。

[1044] 为了获得不同的转导效率,使用不同体积的逆转录病毒上清液对Jurkat TCR KO细胞进行转导(6、20和60% $mTCR\beta^+CD8^+$ 细胞,图18a)。

[1045] 将经TCR转导的细胞以 0.25×10^6 个细胞/ ml 的浓度铺板,并用不同浓度的杀稻瘟菌素(0、4 $\mu g/ml$ 、5 $\mu g/ml$ 和6 $\mu g/ml$)进行筛选。选择后6天,在第4天去除或不去除杀稻瘟菌素的细胞之间在倍数扩增方面没有明显差异。与CDK4-17 TCR的数据相反,用较低的杀稻瘟菌素浓度(4 $\mu g/ml$ 和5 $\mu g/ml$)选择的差异转导的Jurkat细胞似乎比用6 $\mu g/ml$ 这种抗生素选择的细胞具有稍高的倍数扩增(图18b,图25a示出了未转导细胞的倍数扩增)。尽管在倍数扩增方面存在这些差异,但不同的杀稻瘟菌素浓度始终导致80-90%的 $mTCR\beta^+CD8^+$ 细胞的选择效率(图18c,图25b示出了在4 $\mu g/ml$ 条件下NT细胞的百分比是由于存在非常少的存活细胞,其中一些出现在 $mTCR\beta^+CD8^+$ 门中)。最后,在第4天去除杀稻瘟菌素对倍数扩增和选择效率没有任何影响(图18b,c)。基于该数据,仅使用4 $\mu g/ml$ 杀稻瘟菌素选择的Jurkat细胞在第7天继续进行测量,因为该抗生素浓度导致高倍数扩增和80-90%的 $mTCR\beta^+CD8^+$ 细胞。此外,从三种试验浓度推断使用4 $\mu g/ml$ 的杀稻瘟菌素的毒性最小。

[1046] 选择后7天,就总活细胞和 $mTCR\beta^+CD8^+$ 细胞两方面而言,在第4天去除了杀稻瘟菌素的经TCR转导的Jurkat细胞显示出比在第4天添加了杀稻瘟菌素的细胞略高的倍数扩增(图18d)。然而,在第4天去除杀稻瘟菌素对选择效率没有影响,这导致在所有试验条件下 $mTCR\beta^+CD8^+$ 细胞>90%(图18e)。

[1047] 值得注意的是,根据第4天是否去除或添加了杀稻瘟菌素,观察到了CD8和 $mTCR\beta$ 双阳性的Jurkat细胞的 $mTCR\beta$ MFI的差异。在第4天,添加了杀稻瘟菌素的细胞显示出比去除了杀稻瘟菌素的细胞更高的MFI。这种差异在选择后第7天比第6天更明显(图26)。

[1048] 综上所述,在3种不同的杀稻瘟菌素浓度(4 $\mu g/ml$ 、5 $\mu g/ml$ 、6 $\mu g/ml$)下,可以高效地筛选出CMV-1转导效率不同的Jurkat TCR KO细胞。就倍数扩增而言,较低的杀稻瘟菌素浓度似乎毒性较低,尽管三种试验抗生素浓度之间的差异很小。此外,研究表明,在更早的时间点去除杀稻瘟菌素允许选择继续进行,并导致比在完整的7天选择期间将细胞暴露于抗生素略好的倍数扩增。

[1049] 在GMP袋中进行的Jurkat细胞-APC共培养导致CD69活化标志物的表达与在圆底96W板中进行的共培养相当。

[1050] 提高筛选平台的可扩展性还需要优化已建立的圆底96W板共培养形式,使其能够

在保持TCR的覆盖率的同时更有效地处理大量效应物和靶细胞。先前进行的研究表明,在比96W板孔更大表面积 of 平底细胞培养容器中进行共培养会导致Jurkat细胞的CD69表达下降。此外,改变E:T比例和此类共培养物的起始细胞密度并未导致效应细胞的CD69上调增强(increased CD69 upregulation)(数据未示出)。这些数据表明,使用圆底培养容器,如圆底96W板,可实现更有效的效应物-靶接触。因此,决定确定一个更大规模的共培养系统,该系统将增强效应物-靶的接触,并允许在单个培养容器中处理较高的细胞数量。

[1051] 为了比较不同的共培养系统,使用其中40-50%为TCR阳性的CD8⁺Jurkat细胞作为效应细胞,以获得明确的CD69阳性群和阴性群(图19a、19b、19c、19d-左)。作为靶细胞,使用表达HLA-A*02:01的EBV永生B细胞(EBV LCL),该细胞已被工程改造以表达由表达TCR的Jurkat细胞(EBV LCL TMG2.1)识别的抗原。

[1052] 首先通过在置于支架上的15和50ml Falcon管中进行共培养来检查增强效应物-靶接触。由于Falcon管未进行细胞培养处理,因此使用少量细胞($\leq 2 \times 10^6$ 个细胞,根据96W板的孔与15/50ml Falcon管之间的直径差异按比例放大)来评价共培养20小时后的细胞活力和效应细胞活化的情况。活细胞和CD69⁺细胞的百分比在96W板中与在Falcon管中的共培养物之间相当(图19a)。然而,将细胞数量增加到 38×10^6 导致细胞活力的急剧下降(图19a)。

[1053] 值得注意的是,由于CD69上调是Jurkat细胞活化的唯一量度,因此验证了抗人CD69抗体(克隆FN50)是否提供了CD69表达的准确表征(representation)。与针对不同表位的抗人CD69抗体(克隆CH/4)共染色显示两个不同抗体克隆之间的染色水平相当(图27)。

[1054] 接下来,检查在Falcon管共培养中的细胞损失是否是由于与大气接触的培养基表面积小,妨碍了有效的气体交换。为了对其进行评估,将Falcon管离心旋转(spun down),并从含有 38×10^6 个细胞的试管中取出14ml培养基,得到的总体积为1ml。然而,在与 38×10^6 个细胞的Falcon管共培养物中观察到细胞活力急剧下降(图19b)。此外,还进一步测试了250ml试管中的共培养物,因为这些容器的直径更宽,可以使得进行更有效的气体交换(因为没有合适的离心机转子,没有进行离心旋转)。尽管250ml管共培养导致与96W板共培养相似的CD69上调,但细胞死亡增加(图19b)。

[1055] 由于细胞死亡的增加,因此得出结论认为在(Falcon)管中进行大量细胞的共培养是不可行的。对细胞活力的这种影响可能是由于气体交换的效率低下和(Falcon)管中放置的大量细胞所致。因此,在MACS GMP细胞分化袋-500(简称GMP袋),一种可以容纳大量细胞并允许气体交换穿过袋的整个表面区域的培养容器中评估了细胞活力和效应物-靶接触。

[1056] 当前的TCR文库筛选平台允许筛选10000个TCR,并且需要 170×10^6 个效应物和靶细胞的共培养,以保持TCR的覆盖率。因此,在GMP袋中建立了 170×10^6 个Jurkat细胞和APC的共培养物。

[1057] 在替代方案中,筛选平台可以包括 80×10^6 个效应细胞和 80×10^6 个靶细胞,因此值可以是 160×10^6 个。或者,可以与 90×10^6 个细胞建立共培养物,因为由于使用多通道,可能损失~10%,并且可能希望确保至少保留 80×10^6 个细胞。

[1058] 为了增强效应物-靶的接触并允许有效的气体交换,将GMP袋水平放入圆形滤器(round colander)中(图19c)。共培养20小时后,在GMP袋和96W板的共培养物之间观察到相当的细胞活力和CD69上调。然而,在GMP袋条件下CD69 MFI略有降低(图19c)。因此,评估了

离心旋转 (spinning down) GMP袋是否会增加CD69 MFI。为了能够使GMP袋离心旋转,将它们垂直放置在离心机转子中,并移动到滤器中,而不干扰 (disturbing) 细胞。未离心旋转的GMP袋也被置于直立位置。CD69+ Jurkat细胞的百分比和MFI在GMP袋和96W平板中的共培养物之间是相当的。此外,没有观察到依赖于效应物和靶细胞在20小时共培养之前是否离心旋转的CD69表达的主要差异 (图19d)。然而,在GMP袋中共培养20小时后,细胞活力下降。令人惊讶的是,不同GMP袋之间的细胞损失量不同 (图19d)。

[1059] 总之,数据表明,在GMP袋中的共培养物导致效应细胞的活化,与已经建立的96W平板共培养形式类似。然而,为了检查图19d中细胞死亡增加的原因,需要进一步研究—例如,放置袋子的不同方式。

[1060] CD25和CD62L显示出与CD69相结合的基于珠的两步选择的有希望的表达谱。

[1061] 迄今为止,CD69已被用作活化标志物,用于在20小时共培养和流式细胞术分选后,在TCR文库筛选平台中选择活化的Jurkat细胞。为了评估更具可扩展性的基于珠的选择方法,除了CD69之外,还探索了两种辅助T细胞活化标志物CD25和CD62L的表达谱。为此,在与表达同源抗原的靶细胞共培养后,对TCR转导的Jurkat细胞的CD69、CD25和CD62L表达进行了纵向分析。

[1062] 共培养开始后16-32小时, Jurkat细胞的CD69上调似乎保持稳定 (图20a)。相比之下,CD25⁺和CD69⁺CD25⁺效应细胞的表达谱在24小时时间点后显示轻微上调 (图20b、20c)。进一步观察到,在共培养开始后的16至20小时时间点之间,CD62L下调略有增加,并保持稳定,直至32小时时间点。然而,当使用CD62L作为活化标志物时,背景信号较高,因为在开始共培养后分析的各个时间点处,经TCR转导的未刺激的和未转导的Jurkat细胞的CD62L阴性在20-40%之间的 (图20d)。CD69与CD62L共表达分析发现,非特异性下调CD62L表达的效应细胞未上调CD69。此外,在共培养开始后16-32小时,可以观察到CD69上调但CD62L表达缺乏的活化Jurkat细胞群 (图20e)。

[1063] CD25和CD62L在该初步纵向分析中显示出有希望的表达谱,因为它们的表达与CD69上调相关,这表明它们可能与CD69联合用于基于珠的两步选择过程。

[1064] 慢病毒NFAT报告系统在Jurkat和原代T细胞中产生高背景信号。

[1065] 在TCR文库筛选过程中,使用内源性细胞表面活化标志物来选择反应性TCR的另一种方法是在Jurkat细胞中建立一个基于NFAT的报告系统。建立该报告系统可能允许对活化的Jurkat细胞进行抗生素选择,这将规避流式细胞术分选。或者,抗生素抗性盒可以被任何适合于基于珠富集的细胞表面标志物替代。

[1066] 设计了四种不同的基于NFAT的慢病毒报告载体。自身失活 (SIN) 型慢病毒质粒含有截短的无活性的5' LTR启动子,因此整合进入基因组中后,嘌呤霉素抗性报道基因的转录应仅依赖于NFAT转录因子与NFAT结合位点的结合⁴¹。其中三个构建体含有两个、四个或六个相同的NFAT结合位点,后跟一个最小的IL2启动子。第四个载体由包含至少三个不同的NFAT结合位点的完全IL2启动子组成 (图21a)⁶。研究了不同的基于NFAT的报告质粒,以确定最敏感的报告质粒。

[1067] 将基于NFAT的慢病毒载体与另一种pmaxGFP质粒共转染入HEK293T细胞。转染后三天产生病毒的细胞的瞬时GFP表达被用作转染效率的量度。HEK293T细胞为100%GFP⁺,表明慢病毒转染成功 (图21b)。然后,将慢病毒上清液转导至Jurkat TCR KO细胞中,随后将其暴

露于非生理性PMA/离子霉素刺激24小时,以评估报告系统的功效。研究表明,刺激后72小时,NFAT激活达到最大值³⁸。因此,在PMA/离子霉素刺激后,进行3天的嘌呤霉素选择。然而,NFAT转导的Jurkat细胞的细胞活力急剧降低并且在嘌呤霉素暴露的转导和非转导细胞之间是相当的(图21c,上部图)。活化细胞>99%CD69⁺,意味着刺激是有效的(图21c,下部图)。

[1068] 将NFAT转导的嘌呤霉素选择的细胞扩增20天,并进行第二轮刺激和选择。注意,本实验中的未转导的Jurkat细胞与图21c中的未转导的细胞不同。PMA/离子霉素刺激,随后用不同浓度的嘌呤霉素进行4天选择,在刺激的和未刺激的NFAT转导的细胞之间在细胞活力方面没有产生主要差异(图21d)。令人惊讶的是,转导细胞的CD69表达低于未转导细胞(图21d)。一种解释可能是,NFAT转导的Jurkat细胞未从第一轮非生理性PMA/离子霉素刺激中完全恢复。

[1069] 综上所述,该数据表明,在T细胞刺激后,嘌呤霉素耐受的NFAT转导的Jurkat细胞没有富集。尽管pmaxGFP质粒的转染是有效的,但慢病毒构建体的转染和转导可能不成功(图21b)。其他可能的解释可能是抗生素选择欠佳或报告系统泄漏。为了解决上述的一些场景,使用含有EGFP报道基因的NFAT报告构建体来进行以下实验(图22a)。在另一种选择中,报道基因可以是任何适合基于珠的选择/抗生素抗性基因的细胞表面标志物。

[1070] 通过使用EGFP NFAT质粒,可以通过读取HEK293T细胞的CMV驱动GFP表达来直接评估慢病毒转染效率。这表明PEI是比FuGENE更有效的慢病毒转染试剂(图22b)。为了研究报告系统的效力,用PMA/离子霉素刺激NFAT转导的Jurkat TCR KO细胞(PEI转染)24小时,并每24小时读取GFP表达,持续3天。对于任何基于NFAT的载体,在GFP⁺细胞的百分比和MFI方面,在刺激和未刺激的Jurkat细胞之间没有观察到明显的差异(图22c)。CD69在激活时的上调揭示了PMA/离子霉素刺激是有效的(图22c)。总之,该数据表明慢病毒NFAT报告系统在Jurkat细胞中具有非常高的背景信号。

[1071] 接下来评估在NFAT转导的原代T细胞中是否也观察到类似的背景信号。为此,使用了两种基于NFAT的慢病毒载体——NFAT4×和新NFAT4×。与NFAT4×相反,新NFAT4×不含最小IL2启动子,而是含有一个复杂性较低的启动子,称为minP。Jutz等人先前已描述过minP³⁸。用NFAT慢病毒上清液转导来自两个健康供体的原代T细胞,然后用PMA/离子霉素刺激24小时。每24小时测量一次GFP表达,持续3天。与Jurkat细胞的数据相反,在刺激条件下观察到的GFP⁺细胞的百分比和MFI略高于未刺激条件下的GFP⁺细胞,尽管这些差异不是非常显著(图23)。此外,使用具有不同最小启动子的NFAT4×构建体并未在转导的T细胞中产生更特异的GFP表达。尽管CD69在PMA/离子霉素刺激时没有上调,但在受刺激和未受刺激的NFAT转导细胞之间GFP表达的变化表明刺激是有效的(图23)。

[1072] 总之,数据表明基于NFAT的报告载体的慢病毒递送在Jurkat和原代细胞中均产生了高背景信号。

[1073] 将NFAT报告构建体非病毒递送到Jurkat细胞中导致高背景信号。

[1074] 接下来,通过设计NFAT0×和NFAT4×(图24a)两种不同的构建体,研究了非病毒NFAT报道基因(reporter)递送系统在Jurkat细胞中的效力。NFAT0×不含任何NFAT结合位点,因此可作为对照评估最小启动子的背景信号。两个质粒均含有一个EGFP报道基因和组成型表达的E2-Crimson和嘌呤霉素抗性基因。通过用两个NFAT质粒使Jurkat细胞电穿孔,构建体将被随机线性化,并整合到基因组中的任意位置。因此,预期转染效率较低。然而,通

过进行嘌呤霉素选择,在基因组中稳定整合质粒的Jurkat细胞将被富集。将通过刺激选定的Jurkat细胞进一步评估NFAT背景信号。

[1075] 用NFAT构建体转染的表达CDK4-17的Jurkat TCR KO细胞没有表现出细胞活力随时间的显著降低(图24b)。令人惊讶的是,电穿孔后两天,观察到GFP⁺和GFP⁺E2-Crimson⁺细胞增加,这对于NFAT0×转染的Jurkat细胞更为明显。转染后6天这两种细胞群的减少意味着双阳性和GFP单阳性表达是暂时的(图24c)。转染后6天开始嘌呤霉素选择,持续7天。在选择期间,E2-Crimson和GFP单阳性和双阳性Jurkat细胞的百分比增加(图24c)。在嘌呤霉素处理后7天,由于没有存活的非电穿孔的Jurkat细胞,选择似乎结束了(图24d)。然而,GFP⁺和GFP⁺E2-Crimson⁺细胞的百分比与NFAT转染的Jurkat细胞的E2-Crimson表达相当。这表明NFAT报告质粒的非病毒递送也导致高背景信号。

[1076] 关于实施例19的讨论

[1077] 通过关注新生抗原特异性TCR的识别来治疗癌症的完全个性化的TCR基因疗法是非常有用的。这种方法依赖于遗传筛查,并且因此需要对大量表达TCR的Jurkat细胞和APC进行处理。为了能够治疗许多患者,筛选平台应允许有效处理效应物和靶细胞,而不影响筛选敏感性。在此,介绍了为提高新生抗原反应性TCR分离平台的可扩展性而开展的研究。

[1078] 对TCR转导的Jurkat TCR KO细胞的杀稻瘟菌素选择。

[1079] 将TCR文库逆转录病毒基因转移至Jurkat TCR KO细胞后,为了富集表达TCR的Jurkat细胞而不引起毒性和损失TCR覆盖率,有效的选择方法是有用的。为此,已证明抗生素杀稻瘟菌素可导致CMV-1和CDK4-17 TCR转导的Jurkat TCR KO细胞的有效选择(图18,数据未示出)。尽管所测试的不同杀稻瘟菌素浓度和杀稻瘟菌素选择方法之间的差异不是主要的,但是选择CMV-1转导的细胞7天(起始密度为 0.25×10^6 个细胞/ml和起始4 μ g/ml的杀稻瘟菌素,在培养的第4天去除杀稻瘟菌素)导致富集>90%的mTCR β^+ CD8⁺Jurkat细胞和最高倍数的扩增(图18)。

[1080] 然而,这种选择方法导致约70%的mTCR β^+ CD8⁺Jurkat细胞被10000个TCR的文库转导。此外,在共培养试验之前,当Jurkat细胞扩增时,mTCR β^+ CD8⁺细胞的百分比进一步降低(数据未示出)。因此,为了富集用患者来源的TCR文库转导的Jurkat细胞,规范地应用6 μ g/ml的杀稻瘟菌素选择并在第4天添加新的杀稻瘟菌素,以获得>70%的mTCR β^+ CD8⁺Jurkat细胞群(数据未示出)。此外,应用 0.5×10^6 个细胞/ml的起始细胞密度,以减少所需培养瓶的数量。

[1081] 使用两种已知的TCR进行杀稻瘟菌素选择研究,而患者筛选涉及10000种非特征性TCR的转导。因此,对表达TCR文库的Jurkat细胞进行抗生素选择可能需要比富集表达单一TCR的细胞更苛刻的处理。在Spindler等人的研究中也进行了类似的观察,其中使用嘌呤霉素选择表达TCR文库的Jurkat细胞14天。选择后CD3⁺TCR $\alpha\beta^+$ 细胞从~10%富集到~50%⁴²。

[1082] 如果在第4天添加杀稻瘟菌素导致具有较低mTCRBMFI的表达TCR的细胞的损失,或者如果观察到的MFI差异仅仅是由于抗生素处理时间延长导致Jurkat细胞表型改变的结果,则这一问题仍有待解决(图26)。此外,检测低转导效率(<10% mTCR β^+ CD8⁺)是否会导致TCR覆盖率下降也是有用的。这些问题可以通过将不同TCR的文库转导至Jurkat细胞并对富集杀稻瘟菌素的细胞的TCR进行测序来解决。

[1083] 总之,为TCR转导的Jurkat细胞寻找最优化和最有效的富集选择方法的过程表明,

使用已知TCR的研究是有用的。然而,实验结果有待进一步验证,在筛选未知TCR文库时可能会涉及额外的调整。

[1084] 在GMP袋中的Jurkat细胞-APC共培养物。

[1085] 为了进一步提高筛选平台的可扩展性,Jurkat细胞-APC共培养形式涉及从已建立的96W圆底形式升级。

[1086] 我们的理解是,本研究首次检验了在GMP袋中的新共培养形式的功效,其导致与在96W板中的共培养类似的CD69表达(图19c、19d)。GMP袋可以以模拟96W板的圆底孔的形式折叠(folded),从而增强效应物与靶的接触。此外,GMP袋允许在其整个表面区域上进行有效的气体交换,以防止细胞损失。然而,如图19d所示,与96W平板共培养相比,GMP袋中的共培养导致细胞死亡增加。细胞活力的下降可能是由于短时间旋转步骤所需的GMP袋的不同放置所致。将细胞收集在GMP袋的底部,GMP袋底部可能是培养容器中唯一不允许进行有效气体交换的部分。注意,在图19c中,GMP袋没有离心旋转,并且将细胞收集在袋的一侧,这可能是细胞活力不受影响的原因。尽管重复进行GMP袋共培养对于进一步检查图19h中的细胞死亡原因是有用的,但这些观察结果强调了建立将柔性GMP袋放入培养箱的一致方法的必要性。例如,可以设计具有适当的形状同时还能保持CO₂渗透性的铸件(cast)。

[1087] 使用GMP袋进行Jurkat细胞-APC共培养的另一个好处是,可以在保持相似的细胞密度的同时添加更多的细胞。这可以允许筛选超过10000个TCR。此外,GMP袋是一个封闭系统,这将大大降低交叉污染的可能性。

[1088] 值得注意的是,仅基于T细胞活化标志物CD69的表达评估了不同共培养系统的功效。此实验设置未考虑GMP袋共培养可能对TCR覆盖率产生的影响。为了研究这一点,可以将表达TCR文库的Jurkat细胞与APC在GMP袋和96W板中平行共培养。随后将对活化的Jurkat细胞进行基于CD69的流式细胞术分选和下一代测序来比较特异性TCR的富集。

[1089] 其他实验将评估在GMP袋中进行共培养对TCR分离过程的影响。然而,初步数据显示了一个有前景的共培养系统,该系统应在至少10000个TCR的文库筛选中提供对效应物和靶细胞的有效处理。

[1090] 用基于珠的选择取代基于流式细胞术的分选。

[1091] 目前的一些TCR筛选平台涉及在共培养后对表达CD69的Jurkat T细胞进行流式细胞术分选,以识别新生抗原特异性TCR。然而,基于珠选择活化的Jurkat细胞将提高TCR筛选平台的可扩展性。因此,决定评价用于基于珠的选择的另外的或替代活化标志物,其优选在激活的和非激活的Jurkat群之间给出明确的分离(clear separation)。

[1092] 使用另外两种T细胞活化标志物CD25和CD62L进行纵向共培养分析,结果显示CD25表达的异质性大于活化的Jurkat细胞的CD69表达的异质性(图20a、20b)。此外,CD62L下调并非仅对活化的Jurkat细胞具有特异性,因为一些未转导和未刺激的TCR转导的效应细胞呈CD62L阴性(图20d)。因此,这两种活化标志物对于基于单个标记珠的选择方法来说不是有吸引力的候选物。然而,CD25和CD62L可用于基于珠的连续选择过程(sequential bead-based selection process),分别包括CD62L⁺或CD25⁺效应细胞的耗竭或富集,然后通过CD69正向选择捕获活化细胞。值得注意的是,在24小时时间点之后,CD25⁺和CD69⁺CD25⁺群体略微升高。然而,考虑到CD25已被证明是一种晚期活化标志物,这并不奇怪⁴³。这也表明,对于分析具有这两种活化标志物的效应细胞而言,较晚的时间点最为理想。据报道,CD62L脱

落 (shedding) 在T细胞激活后约4小时达到峰值,因此是早期活化标志物^{44,45}。一项研究显示,在4小时峰值之后,CD62L经历了动态表达谱 (dynamic expression profile)⁴⁴。然而,在此,在共培养20小时后观察到CD62L非常稳定的下调,其遵循与CD69上调类似的模式 (图20a、20d)。这将使我们能够在Jurkat细胞活化后16-32小时用CD69和CD62L进行共表达分析 (图20e)。然而,这种纵向分析需要通过将TCR转导的Jurkat TCR KO细胞与不表达同源肽的EBV LCL共培养进行重复实验 (repetition)。这将允许更准确地估计三种活化标志物的背景信号。

[1093] 最后,检查了慢病毒NFAT-报告系统允许抗生素选择或基于珠富集活化的Jurkat细胞的功效。然而,NFAT载体的慢病毒递送在Jurkat和原代细胞中均产生高背景信号 (图22和图24)。采用这种报告系统的研究主要依赖于具有最小背景信号的克隆群体^{37,38}。将具有嘌呤霉素抗性报道基因的NFAT转导的Jurkat细胞暴露于两轮抗生素选择中 (图21)。然而,尽管假设在初始抗生素治疗后获得了半克隆群体 (semi-clonal population),但是在第二轮选择后,受刺激和未受刺激的Jurkat细胞之间没有明显差异。另一种解释可能是次优的非生理性PMA/离子霉素刺激,先于第二次嘌呤霉素选择,产生<30%的CD69⁺NFAT转导的Jurkat细胞 (图21d)。此外,受刺激的和未受刺激的EGFP NFAT转导的Jurkat细胞之间没有明显差异 (图22)。因此,由于缺乏GFP的差异表达,选择克隆是不可行的。结果表明,报告系统渗漏 (leakiness) 不是由最小IL2启动子引起的,因为另一个最小启动子minP显示了相似水平的背景 (图23)。因此得出结论,渗漏可能是慢病毒递送系统的一个特征。

[1094] SIN NFAT逆转录病毒载体已被成功使用和建立,用于研究抗原特异性T细胞应答^{37,38,40}。NFAT质粒的非病毒递送也导致高背景信号,因为在非刺激的Jurkat细胞中报道基因EGFP的表达升高 (图24)。然而,需要活化T细胞,以进一步评估NFAT转染的Jurkat细胞的渗漏性。此外,可对E2-Crimson单阳性群体进行流式细胞术分选,以随后在T细胞刺激试验中评价该群体的功效和GFP渗漏性。未来研究的一个领域可能是研究AP-1和NF- κ B报告系统的敏感性,因为这些转录因子利用不同的信号通路,这些通路可能对背景信号有不同的影响³⁸。例如,NF- κ B报告系统已成功应用于Jurkat细胞中嵌合抗原受体文库⁴⁰的筛选。

[1095] 本研究旨在提高新生抗原特异性TCR分离平台的可扩展性。这些实验非常有用,因为对报道Jurkat T细胞和APC的更有效处理最终将允许对更多的癌症患者的治疗。首先,筛选逆转录病毒TCR转导的Jurkat TCR KO细胞的杀稻瘟菌素选择被确定为最高效和毒性最小的效应细胞富集过程。此外,将Jurkat细胞-APC共培养物从96W的形式扩大到更具可扩展性的GMP袋设置,导致了相当水平的Jurkat细胞活化。最后,已经奠定了基础,可以在TCR发现平台中用更具可扩展性的基于珠的表达新生抗原反应性TCR的Jurkat细胞的选择来替代流式细胞术分选。

[1096] 材料和方法—实施例19

[1097] 细胞系

[1098] 人Jurkat细胞系 (克隆E6-1,TIB-152) 购自ATCC,人EBV永生化HLA-A2-阳性B细胞系,称为JY,购自ECACC。Jurkat和JY细胞系二者均在RPMI-1640培养基中培养,HEPES (Gibco) 补充有热灭活的10%胎牛血清 (FBS,Gibco)、100U/ml青霉素 (Gibco) 和100 μ g/ml链霉素 (Gibco)。人胚肾293T (HEK293T,CRL-3216) 和Phoenix Ampho细胞系购自ATCC,在补充有10%FBS、100U/ml青霉素 (Gibco) 和100 μ g/ml链霉素的DMEM培养基 (Gibco) 中培养。将所

有细胞系均保持在37℃和5%CO₂的条件下。

[1099] 人健康供体中T细胞和B细胞的分离

[1100] 在知情同意后,从Sanquin Blood Supply (荷兰) 获得了健康人类供体的血块黄层(Buffycoats)。使用Ficoll Paque Plus (西格玛-奥德里奇 (Sigma-Aldrich)) 从血块黄层中分离外周血单核细胞(PBMCs)。

[1101] 使用MojoSort人泛B细胞分离试剂盒(BioLegend) 进行负向选择,分离HLA-A2-阳性B细胞,随后使用人γ 疱疹病毒4 (HHV-4, ATCC VR-1492) 进行EBV永生。EBV永生B细胞(EBV LCLs) 在RPMI-1640培养基中培养,HEPES (Gibco) 补充有20%FBS、100U/ml青霉素(Gibco) 和100μg/ml链霉素(Gibco)。

[1102] 使用CD3/CD28 Dynabeads (欧洲生命技术公司 (Life Technologies Europe)) 分离并活化T细胞,并在RPMI-1640培养基中培养,HEPES (Gibco) 补充有10%人血清(西格玛-奥德里奇)、100U/ml青霉素(Gibco)、100μg/ml链霉素(Gibco) 和细胞因子(5ng/ml人IL-15和100IU/ml人IL-2,派普泰克公司 (Peprotech))。将所有原代细胞保持在37℃和5%CO₂的条件下。

[1103] 流式细胞术和分选

[1104] 通过向PBS (Gibco) 中添加2%FBS制成FACS缓冲液。为了进行流式细胞术分析,将细胞在4℃黑暗条件下染色20min,并用FACS缓冲液洗涤一次。LIVE/DEAD Fixable Near-IR (欧洲生命技术公司,1:1000稀释) 或DAPI (西格玛-奥德里奇,1:1000稀释) 被用作活/死染色剂(live/dead stain)。在4℃下用Cytotfix将产生病毒的细胞(BD生物科学 (BD Biosciences)) 固定30分钟,并在测量前用PBS洗涤一次。在BD LSR Fortessa光谱仪上测量样本,并使用FlowJo v10.6.1软件进行分析。

[1105] 为了分选,用FBS涂覆收集管。以每15ml试管中0.5ml抗体溶液对20×10⁶个细胞进行染色。在4℃黑暗条件下进行染色20分钟,随后用10ml FACS缓冲液洗涤细胞。然后,在FACS缓冲液中将细胞浓度调节至15-20×10⁶个细胞/ml,并使样本通过一个35μm的细胞过滤器(cell strainer)。在分选前立即加入DAPI (西格玛-奥德里奇,1:1000稀释) 作为活/死染色剂。在FACS Aria Fusion上对样本进行分选。

[1106] 在FACS缓冲液中稀释用于流式细胞术和分选的抗体。使用了以下抗体和稀释液:抗人CD8 APC (克隆SK1, BD生物科学,1:300稀释液) 和抗人CD8 APC-R700 (克隆RPA-T8, BD生物科学,1:600稀释液)、抗小鼠TCRβPE (克隆H57-597, BD生物科学,1:150稀释液)、抗人CD69PE/APC/BV510 (克隆FN50, BD生物科学,分别为1:100、1:200和1:300的稀释液) 和与CD69 PE (克隆CH/4, 赛默飞世尔科技 (ThermoFisher Scientific), 1:100), 抗人CD20 FITC (克隆2H7, BD生物科学,1:25稀释液) 和抗人CD20 PE-Cy7 (克隆L27, BD生物科学,1:200稀释液)、抗人CD25BV711 (克隆2A3, BD生物科学,1:200稀释液)、抗人CD62L APC (克隆DREG-56, BD生物科学,1:25稀释液)、抗人CD3 PerCPy5.5 (克隆SK7, BD生物科学,1:25稀释液)。注意:除非另有说明,否则抗人CD69克隆FN50抗体用于CD69读出(readout)。

[1107] 逆转录病毒转染和转导

[1108] 为了进行逆转录病毒转染,将~1.8×10⁶个Phoenix Amphi细胞接种在10cm培养皿中,并在24小时后用10μg pMP71-TCR质粒转染,使用FuGENE-6 (普洛麦格公司 (Promega)) 作为转染试剂。转导的TCR由小鼠恒定区和人可变区组成,以防止转导的TCR与Jurkat细胞

系的内源性TCR二聚化。因此,不同TCR的转导效率可以通过用针对小鼠TCRB (mTCRB) 区域的单一抗体染色来测量。转染48小时后收集病毒上清液,并通过0.45 μ m过滤器。为了进行逆转录病毒转导,用0.5ml/孔的RetroNectin (RN, 宝生物 (Takara)) 包被未经TC处理的24W板,并在4℃下孵育过夜或在室温 (RT) 下孵育2小时。在RT下用0.5ml/孔的封闭缓冲液 (补充有5% FBS的PBS) 封闭平板30分钟,并用PBS洗涤一次。将0.5ml培养基中的 2×10^5 个Jurkat细胞接种到RN包被的24W板中,并与0.5ml病毒上清液混合。将板在880 $\times$ g下离心旋转90分钟 (加速3, 减速0)。3~4天后用流式细胞仪检测转导效率。

[1109] 慢病毒转染和转导

[1110] 为了进行慢病毒转染,将 $\sim 3 \times 10^6$ 个HEK293T细胞接种在10cm培养皿中,24小时后使用第三代包装构建体 (packaging construct) (pRSV-Rev和pCgpV包装载体以及pCMV-VSV-G包膜载体 (envelope vector), Cell Biolabs) 和转染试剂聚乙烯亚胺 (PEI, 1mg/ml, Polysciences) 或Lipofectamine 3000 (赛默飞世尔科技) 用基于10 μ g pSMPUW-NFAT-IL-2的报告构建体转染。对于某些慢病毒转染,将pmaxGFP (来自SE细胞系4D-核转染X试剂盒L, Lonza) 共转染以测量转染效率。转染48小时、72小时后收集病毒上清液,并通过0.45 μ m过滤器。转导前48小时,在补充有IL-2和IL-15的培养基存在的情况下,用CD3/CD28 Dynabeads激活并分离原代T细胞 (参见“从人类健康供体中分离T细胞和B细胞”)。如“逆转录病毒转染和转导”部分所述,通过旋转转导RN包被的 (RN-coated) 板进行原代T细胞的慢病毒转导 (24W板中每孔 2.5×10^5 个原代T细胞)。为了进行Jurkat细胞的慢病毒转导,将 2×10^5 个细胞重新悬浮在含有聚凝胺 (西格玛-奥尔德里奇) 的1ml病毒上清液中。将细胞铺在24W板上,并在3~4天后用流式细胞术测定转导效率。转导后的原代T细胞每3-4天接受一次含有细胞因子的新鲜培养基。

[1111] Jurkat TCR敲除CD8⁺细胞系的建立

[1112] 用pMP71-CD8 α -P2A-CD8 β 逆转录病毒转导Jurkat E6-1细胞系以表达CD8。此外,如Scheper等人⁴⁶的研究所述,通过敲除TCR α 链和 β 链来消除内源性TCR表达。对CD8高表达的单细胞克隆进行分选和扩增。

[1113] Jurkat细胞的电穿孔

[1114] 为了实现pMKRQ-NFAT质粒向Jurkat细胞的非病毒递送 (图24), 遵循来自Lonza (货号V4XC-1024) 的用于Jurkat克隆E6.1 ATCC的Amaxa4D-Nucleofector方案。

[1115] 共培养测定

[1116] CDK4突变体 (ALDPHSGHFV) 和CMVpp65 (NLVPMVATV) 肽均从Pepscan订购。在37℃和5%CO₂的条件下培养1小时后,将肽加载到密度为 1×10^6 个细胞/ml的JY细胞上。接下来,用培养基洗涤JY细胞两次。用编码一系列 (a string of) 25mer多肽的逆转录病毒载体转导EBV LCLs。这些多肽包含目的表位 (例如,CDK4突变体和CMVpp65肽), 并被称为串联小基因 (TMG) 2.1。将不含同源肽或负载不相关肽的JY细胞和未转导的EBV LCLs用作阴性对照。

[1117] 在共培养前24小时,以低密度 ($0.25/0.5 \times 10^6$ 个细胞/ml) 接种效应细胞和靶细胞。除非另有说明,否则使用200000个效应细胞和靶细胞 (总体积200 μ l; E:T比例为1:1) 在96W圆底板中共培养20小时;在20小时培养之前,以1100rpm的速度将板向下旋转1分钟。将抗人CD20抗体纳入检测组,以排除靶细胞。

[1118] 图19c、19d中使用的MACS GMP细胞分化袋-500购自美天旎生物技术公司

(Miltenyi Biotec)。

[1119] 试剂

[1120] 使用了以下附加试剂:50ng/ml的佛波醇12-肉豆蔻酸13-乙酸酯(PMA,西格玛-奥尔德里奇)、1 μ M离子霉素钙盐(西格玛-奥尔德里奇)、嘌呤霉素二盐酸盐(Gibco)、杀稻瘟菌素S HCl(Gibco)。

[1121] 倍数扩增

[1122] 将图18中的倍数扩增值用作毒性的量度,并且通过将选择后的总细胞计数除以选择开始时的总细胞计数来计算。获得的值针对第4天的重新铺板进行了校正。

[1123] 统计分析

[1124] 使用GraphPad Prism 8进行EC50值计算和统计分析。在图23中,进行了双因素ANOVA,随后进行Sidak多重比较检验,P值 ≤ 0.05 被认为具有统计学意义(*P ≤ 0.05 ,**P ≤ 0.01)。

[1125] 参考文献

[1126] 1.Shankaran V,Ikeda H,Bruce AT,et al. IFN γ , and lymphocytes prevent primary tumour development and shape tumour immunogenicity. Nature. 2001;410(6832):1107-1111. doi:10.1038/35074122

[1127] 2.Davis MM,Bjorkman PJ. T-cell antigen receptor genes and T-cell recognition. Nature. 1988;334(6181):395-402. doi:10.1038/334395a0

[1128] 3.Neefjes J, Jongsma MLM, Paul P, Bakke O. Towards a systems understanding of MHC class I and MHC class II antigen presentation. Nat Rev Immunol. 2011;11(12):823-836. doi:10.1038/nri3084

[1129] 4.Weiss A, Littman DR. Signal transduction by lymphocyte antigen receptors. Cell. 1994;76(2):263-274. doi:10.1016/0092-8674(94)90334-4

[1130] 5.Crabtree GR, Clipstone NA. Signal Transmission Between the Plasma Membrane and Nucleus of T Lymphocytes. Annu Rev Biochem. 1994;63(1):1045-1083. doi:10.1146/annurev.bi.63.070194.005145

[1131] 6.Rao A, Luo C, Hogan PG. Transcription factors of the NFAT family: regulation and function. Annu Rev Immunol. 1997;15(1):707-747. doi:10.1146/annurev.immunol.15.1.707

[1132] 7.Tran E, Robbins PF, Rosenberg SA. Final common pathway' of human cancer immunotherapy: Targeting random somatic mutations. Nat Immunol. 2017;18(3):255-262. doi:10.1038/ni.3682

[1133] 8.Schumacher TN, Scheper W, Kvistborg P. Cancer Neoantigens. Annu Rev Immunol. 2019;37(1):173-200. doi:10.1146/annurev-immunol-042617-053402

[1134] 9.Leach DR, Krummel MF, Allison JP. Enhancement of antitumor immunity by CTLA-4 blockade. Science (80-). 1996;271(5256):1734-1736. doi:10.1126/science.271.5256.1734

[1135] 10.Sharma P, Allison JP. The future of immune checkpoint therapy. Science (80-). 2015;348(6230):56-61. doi:10.1126/science.aaa8172

- [1136] 11.Castle JC,Uduman M,Pabla S,Stein RB,Buell JS.Mutation-derived neoantigens for cancer immunotherapy.Front Immunol.2019;10(AUG).doi:10.3389/fimmu.2019.01856
- [1137] 12.Hodi FS,O' Day SJ,McDermott DF,et al.Improved survival with ipilimumab in patients with metastatic melanoma.N Engl J Med.2010;363(8):711-723.doi:10.1056/NEJMoal003466
- [1138] 13.Van Rooij N, Van Buuren MM, Philips D, et al. Tumor exome analysis reveals neoantigen-specific T-cell reactivity in an ipilimumab-responsive melanoma.J Clin Oncol.2013;31(32).doi:10.1200/JCO.2012.47.7521
- [1139] 14.Rizvi NA,Mazières J,Planchard D,et al.Activity and safety of nivolumab,an anti-PD-1 immune checkpoint inhibitor,for patients with advanced,refractory squamous non-small-cell lung cancer(CheckMate 063):A phase 2,single-arm trial.Lancet Oncol.2015;16(3):257-265.doi:10.1016/S1470-2045(15)70054-9
- [1140] 15.Borghaei H,Paz-Ares L,Horn L,et al.Nivolumab versus docetaxel in advanced nonsquamous non-small-cell lung cancer.N Engl J Med.2015;373(17):1627-1639.doi:10.1056/NEJMoal507643
- [1141] 16.Yee C,Thompson JA,Byrd D,et al.Adoptive T cell therapy using antigen-specific CD8+T cell clones for the treatment of patients with metastatic melanoma:In vivo persistence,migration,and antitumor effect of transferred T cells.Proc Natl Acad Sci U S A.2002;99(25):16168-16173.doi:10.1073/pnas.242600099
- [1142] 17.Rosenberg SA,Yang JC,Sherry RM,et al.Durable complete responses in heavily pretreated patients with metastatic melanoma using T-cell transfer immunotherapy.Clin Cancer Res.2011;17(13):4550-4557.doi:10.1158/1078-0432.CCR-11-0116
- [1143] 18. **Stevanović** S,Draper LM,Langhan MM,et al.Complete regression of metastatic cervical cancer after treatment with human papillomavirus-targeted tumor-infiltrating T cells.J Clin Oncol.2015;33(14):1543-1550.doi:10.1200/JCO.2014.58.9093
- [1144] 19.Stevanovic S,Helman SR,Wunderlich JR,et al.A Phase II Study of Tumor-infiltrating Lymphocyte Therapy for Human Papillomavirus-associated Epithelial Cancers.Clin Cancer Res.2019;25(5):1486-1493.doi:10.1158/1078-0432.CCR-18-2722
- [1145] 20.Papaioannou NE,Beniata O V.,Vitsos P,Tsitsilonis O,Samara P.Harnessing the immune system to improve cancer therapy.Ann Transl Med.2016;4(14).doi:10.21037/atm.2016.04.01
- [1146] 21.Zhang J,Wang L.The emerging world of TCR-T cell trials against cancer:A systematic review.Technol Cancer Res Treat.2019;18.doi:10.1177/

1533033819831068

[1147] 22.Rosenberg SA,Yang JC,Restifo NP.Cancer immunotherapy:Moving beyond current vaccines.Nat Med.2004;10(9):909-915.doi:10.1038/nm1100

[1148] 23.Robbins PF,Kassim SH,Tran TLN,et al.A pilot trial using lymphocytes genetically engineered with an NY-ESO-1-reactive T-cell receptor: Long-term follow-up and correlates with response.Clin Cancer Res.2015;21(5):1019-1027.doi:10.1158/1078-0432.CCR-14-2708

[1149] 24.Parkhurst MR,Yang JC,Langan RC,et al.T cells targeting carcinoembryonic antigen can mediate regression of metastatic colorectal cancer but induce severe transient colitis.Mol Ther.2011;19(3):620-626.doi:10.1038/mt.2010.272

[1150] 25.Morgan RA,Chinnasamy N,Abate-Daga D,et al.Cancer regression and neurological toxicity following anti-MAGE-A3 TCR gene therapy.J Immunother.2013;36(2):133-151.doi:10.1097/CJI.0b013e3182829903

[1151] 26.Schumacher TN,Schreiber RD.Neoantigens in cancer immunotherapy.Science (80-).2015;348(6230):69-74.doi:10.1126/science.aaa4971

[1152] 27.Tran E,Turcotte S,Gros A,et al.Cancer immunotherapy based on mutation-specific CD4+T cells in a patient with epithelial cancer.Science (80-).2014;344(6184):641-645.doi:10.1126/science.1251102

[1153] 28.Tran E,Robbins PF,Lu Y-C,et al.T-Cell Transfer Therapy Targeting Mutant KRAS in Cancer.N Engl J Med.2016;375(23):2255-2262.doi:10.1056/NEJMoa1609279

[1154] 29.Rodenko B,Toebe M,Hadrup SR,et al.Generation of peptide-MHC class I complexes through UV-mediated ligand exchange.Nat Protoc.2006;1(3):1120-1132.doi:10.1038/nprot.2006.121

[1155] 30.Cohen CJ,Gartner JJ,Horovitz-Fried M,et al.Isolation of neoantigen-specific T cells from tumor and peripheral lymphocytes.J Clin Invest.2015;125(10):3981-3991.doi:10.1172/JCI82416

[1156] 31.Yadav M,Jhunjhunwala S,Phung QT,et al.Predicting immunogenic tumour mutations by combining mass spectrometry and exome sequencing.Nature.2014;515(7528):572-576.doi:10.1038/nature14001

[1157] 32.López-Cabrera M,Santis AG,Fernández-Ruiz E,et al.Molecular cloning,expression,and chromosomal localization of the human earliest lymphocyte activation antigen AIM/CD69,a new member of the C-Type animal lectin superfamily of signal-transmitting receptors.J Exp Med.1993;178(2):537-548.doi:10.1084/jem.178.2.537

[1158] 33.Krishnamani V,Stamnes MA,Piper RC.MALTA:A calculator for estimating the coverage with shRNA,CRISPR,and cDNA libraries.SoftwareX.2019;9:154-160.doi:10.1016/j.softx.2019.01.006

- [1159] 34. Wölfel T, Hauer M, Schneider J, et al. A p16INK4a-insensitive CDK4 mutant targeted by cytolytic T lymphocytes in a human melanoma. *Science* (80-) .1995;269(5228):1281-1284. doi:10.1126/science.7652577
- [1160] 35. Wills MR, Carmichael AJ, Mynard K, et al. The human cytotoxic T-lymphocyte (CTL) response to cytomegalovirus is dominated by structural protein pp65: frequency, specificity, and T-cell receptor usage of pp65-specific CTL. *J Virol*.1996;70(11):7569-7579. doi:10.1128/jvi.70.11.7569-7579.1996
- [1161] 36. Galkina E, Tanousis K, Preece G, et al. L-Selectin Shedding Does Not Regulate Constitutive T Cell Trafficking but Controls the Migration Pathways of Antigen-activated T Lymphocytes. *J Exp Med*.2003;198(9):1323-1335. doi:10.1084/jem.20030485
- [1162] 37. Hooijberg E, Bakker AQ, Ruizendaal JJ, Spits H. NFAT-controlled expression of GFP permits visualization and isolation of antigen-stimulated primary human T cells. *Blood*.2000;96(2):459-466. doi:10.1182/blood.v96.2.459.014k50_459_466
- [1163] 38. Jutz S, Leitner J, Schmetterer K, et al. Assessment of costimulation and coinhibition in a triple parameter T cell reporter line: Simultaneous measurement of NF- κ B, NFAT and AP-1. *J Immunol Methods*.2016;430:10-20. doi:10.1016/j.jim.2016.01.007
- [1164] 39. Morimoto S, Fujiki F, Kondo K, et al. Establishment of a novel platform cell line for efficient and precise evaluation of T cell receptor functional avidity. *Oncotarget*.2018;9(75):34132-34141. doi:10.18632/oncotarget.26139
- [1165] 40. Rydzek J, Nerreter T, Peng H, et al. Chimeric Antigen Receptor Library Screening Using a Novel NF- κ B/NFAT Reporter Cell Platform. *Mol Ther*.2019;27(2):287-299. doi:10.1016/j.ymthe.2018.11.015
- [1166] 41. Zufferey R, Dull T, Mandel RJ, et al. Zufferey, R., Dull, T., Mandel, R.J., Bukovsky, A., Quiroz, D., Naldini, L., & Trono, D. (1998). Self-inactivating lentivirus vector for safe and efficient in vivo gene delivery. *Journal of Virology*, 72(12), 9873-80. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/9811723> 0A<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC110499>
- [1167] 42. Spindler MJ, Nelson AL, Wagner EK, et al. Massively parallel interrogation and mining of natively paired human TCR $\alpha\beta$ repertoires. *Nat Biotechnol*. Published online 2020. doi:10.1038/s41587-020-0438-y
- [1168] 43. Reddy M, Eirikis E, Davis C, Davis HM, Prabhakar U. Comparative analysis of lymphocyte activation marker expression and cytokine secretion profile in stimulated human peripheral blood mononuclear cell cultures: An in vitro model to monitor cellular immune function. *J Immunol Methods*.2004;293(1-

2):127-142.doi:10.1016/j.jim.2004.07.006

[1169] 44.Chao CC,Jensen R,Dailey MO.Mechanisms of L-selectin regulation by activated T cells.J Immunol.1997;159(4):1686-1694.Accessed May 15, 2020.<http://www.ncbi.nlm.nih.gov/pubmed/9257829>

[1170] 45.Yang S,Liu F,Wang QJ,Rosenberg SA,Morgan RA.The Shedding of CD62L (L-Selectin)Regulates the Acquisition of Lytic Activity in Human Tumor Reactive T Lymphocytes.Teague RM,ed.PLoS One.2011;6(7):e22560.doi:10.1371/journal.pone.0022560

[1171] 46.Scheper W,Kelderman S,Fanchi LF,et al.Low and variable tumor reactivity of the intratumoral TCR repertoire in human cancers.Nat Med.2019; 25(1):89-94.doi:10.1038/s41591-018-0266-5

[1172] 实施例20

[1173] 本实施例描述了使用不同的效应物与靶(E:T)的比例从TCR文库中恢复抗原特异性TCR。

[1174] 选择五种已知抗原反应性的特征性TCR,以及95种非特征性TCR(对于100×100文库;图35A)用于生成TCR文库(由Twist Bioscience进行;旧金山/美国)。简而言之,来自这些TCR的TCR α -链和TCR β -链是通过DNA合成作为片段合成的。为了生成文库,将所有选择的TCR α 和TCR β 片段组合连接到编码TCR β -P2A-TCR α 盒(SEQ ID NO:1)的连续核酸分子上。使用含有TCR β 链和TCR α 链的逆转录病毒构建体以及嘌呤霉素选择标志物作为构建文库的支架。使用这种构建体的逆转录病毒转导最终导致单一转录物的表达,由于在2A位点处的肽切割,这导致TCR β 链和 α 链,以及产生的嘌呤霉素抗性标志物的翻译。每个盒以TCR β -P2A-TCR α -T2A-嘌呤霉素抗性的格式(SEQ ID NO:1)连接一个TCR β -和TCR α -片段。通过由Twist Bioscience进行的组合克隆方法合成了100个TCR中的每一个的TCR- β 和TCR- α 区域,并将其插入该载体中。

[1175] 使用Fugene转染试剂和本领域技术人员已知的方案将100×100文库转染到Phoenix-Ampho病毒产生细胞(ATCC CRL-3213)中。产生的逆转录病毒颗粒用于转导Jurkat报告T细胞系。Jurkat报告T细胞缺乏内源性TCR表达(例如在Mezzadra等人Nature 2017中描述了这种基因敲除的产生),并且在用CD8 α -P2A-CD8 β 转基因(SEQ ID NO:2)转导后,使用本领域技术人员已知的方法修饰Jurkat报告T细胞以表达人CD8 α 和CD8 β 。用TCR文库转导Jurkat报告T细胞,得到占总活T细胞25%的经TCR修饰的T细胞(基于转导后4天的染色—嘌呤霉素选择后和检测当天的纯度>80%)。在TCR文库(SEQ ID NO:1)中使用鼠TCR恒定结构域序列允许通过使用鼠TCR β 恒定结构域特异性抗体的流式细胞术检测TCR修饰的Jurkat T细胞。通过向细胞培养基中添加嘌呤霉素,正向选择经修饰以表达TCR β -P2A-TCR α -T2A-嘌呤霉素抗性转基因的Jurkat T细胞至高纯度。遗传修饰细胞的抗生素选择是本领域普通技术人员已知的。

[1176] 接下来,用经工程改造以呈递抗原的APC刺激TCR转导的Jurkat T细胞。将表达TMG的EBV-LCL细胞以10%的比例掺入经过工程改造以表达特定抗原的EBV-LCL。在共培养前3天,以低密度(0.1×10^6 个细胞/ml)接种Jurkat报告T细胞。以1:1、1:2和1:3的比例将浓度为 2.5×10^6 个细胞/ml的APC与Jurkat报告T细胞混合。每个E:T比例进行单次重复(存在TMG

表达),并包括单次阴性对照重复(不存在TMG表达)。将200 μ L(每孔 0.5×10^6 个细胞)分别分布在经过TC处理的U形底96孔板的 ~ 360 、540或720个孔中。将板以1000rpm离心1分钟,并在37 $^{\circ}$ C下孵育20-22小时。

[1177] 随后,收获共培养物,并用购自Miltenyi的抗CD20微珠标记细胞,并转移至放置在磁体上的LS-柱。使用磁珠选择,CD20⁺细胞被保留在柱中,而CD20⁻细胞通过柱并被收集到阴性部分中。然后用荧光染料标记的针对CD20和CD69的抗体对这些CD20⁻细胞进行染色。接下来,使用含有4%甲醛的固定缓冲液固定细胞。使用BD生物科学ARIAFusion流式细胞仪分选器,针对每种刺激条件对下列群体进行分选(图35B):

[1178] 1.淋巴细胞、单细胞、CD20⁻、CD69^高(基于CD69荧光信号,CD69^高包括最高10-15%的单细胞、CD20⁻细胞)

[1179] 2.淋巴细胞、单细胞、CD20⁻、CD69^低(基于低CD69荧光信号,CD69^低包括10-15%的单细胞、CD20⁻细胞)

[1180] 从分选的TCR转导的Jurkat T细胞中分离基因组DNA,并使用所述基因组DNA作为模板,使用本领域技术人员已知的PCR方法,以有限循环次数的多轮PCR扩增TCR β -P2A-TCR α 盒的一部分。所得PCR产物的大小约为1.5kb(图35C),并且可以使用Oxford Nanopore MinIon或GridIon测序仪使用本领域技术人员已知的技术进行测序。使用本领域技术人员已知的比对技术恢复TCR身份,并且使用DESeq2 R套装软件识别差异表达的TCR组合。使用DESeq2 R套装软件计算Rlog转换的读段计数,从顶部样本的Rlog值中减去底部样本的Rlog值,并表示存在(x轴)和不存在(y轴)由B细胞表达的TMG的情况下进行的共培养(图35D)。5个特征性TCR用较大的灰点表示。TCR身份以及关键统计指标代表五个特征性TCR(图35E)。

[1181] 综上所述,本实施例表明提高靶细胞与效应细胞的比值可以提高TCR文库筛选平台的灵敏度。

[1182] 实施例21

[1183] 本实施例描述了使用基于珠的细胞分选策略从通过基因合成产生的TCR文库中恢复抗原特异性TCR。

[1184] 选择五种已知抗原反应性的特征性TCR,以及95种非特征性TCR(对于100 $\times$ 100文库;图36A)用于生成TCR文库(由Twist Bioscience进行;旧金山/美国)。简而言之,来自这些TCR的TCR α -链和TCR β -链是通过DNA合成作为片段合成的。为了生成文库,将所有选择的TCR α 和TCR β 片段组合连接到编码TCR β -P2A-TCR α 盒(SEQ ID NO:1)的连续核酸分子上。可以使用含有TCR β 链和TCR α 链以及嘌呤霉素选择标志物的逆转录病毒构建体作构建文库的支架。使用这种构建体的逆转录病毒转导最终导致单一转录物的表达,由于在2A位点处的肽切割,这导致TCR β 链和 α 链,以及产生的嘌呤霉素抗性标志物的翻译。每个盒以TCR β -P2A-TCR α -T2A-嘌呤霉素抗性的格式(SEQ ID NO:1)连接一个TCR β -和TCR α -片段。通过由Twist Bioscience进行的组合克隆方法合成了50或100个TCR中的每一个的TCR- β 和TCR- α 区域,并将其插入该载体中。

[1185] 使用Fugene转染试剂和本领域技术人员已知的方案将100 $\times$ 100文库转染到Phoenix-Ampho病毒产生细胞(ATCC CRL-3213)中。产生的逆转录病毒病毒颗粒用于转导Jurkat报告T细胞系。Jurkat报告T细胞缺乏内源性TCR表达(例如在Mezzadra等人,Nature 2017中描述了这种基因敲除的产生),并且在用CD8 α -P2A-CD8 β 转基因(SEQ ID NO:2)转导

后使用本领域技术人员已知的方法修饰Jurkat报告T细胞以表达人CD8 α 和CD8 β 。用TCR文库转导Jurkat报告T细胞,得到占总活T细胞25%的经TCR修饰的T细胞(基于转导后4天的染色—嘌呤霉素选择后和检测当天的纯度>80%)。在TCR文库(SEQ ID NO:1)中使用鼠TCR恒定结构域序列允许通过使用鼠TCR β 恒定结构域特异性抗体的流式细胞术检测TCR修饰的Jurkat T细胞。通过向细胞培养基中添加嘌呤霉素,正向选择经修饰以表达TCR β -P2A-TCR α -T2A-嘌呤霉素抗性转基因的Jurkat T细胞至高纯度。遗传修饰细胞的抗生素选择是本领域普通技术人员已知的。

[1186] 接下来,用经过或未经过工程改造以呈递源自肿瘤的抗原的APC刺激TCR转导的Jurkat T细胞(各1次重复)。在共培养前3天,以低密度(0.1×10^6 细胞/ml)接种Jurkat报告T细胞。将浓度为 2.5×10^6 个细胞/ml的APC与Jurkat报告T细胞以1:1的比例混合。将200 μ L(每孔 0.5×10^6 个细胞)分布在经过TC处理的U形底96孔板的约360个孔中。将板以1000rpm离心1分钟,并在37 $^{\circ}$ C下孵育20-22小时。

[1187] 随后,收获共培养物,并使用死细胞去除试剂盒去除死细胞。基于多步骤分离策略分离活细胞(图36B),用抗CD20微珠标记并转移到AutoMACS(Miltenyi, Bergisch Gladbach, 德国)上。使用磁珠选择,CD20+细胞被保留到柱中,而CD20-细胞通过柱并被收集到阴性部分中。然后用CD62L微珠标记这些CD20-细胞,以使用autoMACS将未激活的Jurkat细胞与激活的Jurkat细胞分离。在这种分离之后,用CD69-生物素抗体标记阴性部分(CD62L-细胞)和阳性部分(CD62L+细胞)二者,并使用抗生物素微珠分离CD69-和CD69+部分。收集以下群用于进一步分析(图36B):

[1188] 1. CD20 $^{-}$ 、CD62L $^{+}$ 、CD69 $^{-}$ (“底部”部分 $\rightarrow$ 未激活的Jurkat T细胞)

[1189] 2. CD20 $^{-}$ 、CD62L $^{-}$ 、CD69 $^{+}$ (“顶部”部分 $\rightarrow$ 激活的Jurkat T细胞)

[1190] 从基于珠的分选的TCR转导的Jurkat T细胞中分离基因组DNA,并使用所述基因组DNA作为模板,使用本领域技术人员已知的PCR方法,以有限循环次数的多轮PCR扩增TCR β -P2A-TCR α 盒的一部分。所得PCR产物的大小约为1.5kb(图36C),并且可以使用本领域技术人员已知的技术使用Oxford Nanopore MinIon或GridIon测序仪进行测序。使用本领域技术人员已知的比对技术恢复TCR身份,并且使用DESeq2 R套装软件识别差异表达的TCR组合。使用DESeq2 R套装软件计算每种TCR的Rlog转换的读段计数,从顶部样本的Rlog值中减去底部的Rlog值,并代表在存在(x轴)和不存在(y轴)通过B细胞的TMG表达的情况下进行的共培养(图36D)。TCR α 和 β 链身份,以及关键统计度量代表顶部样本中差异最大的7个TCR(图36E)。

[1191] 综上所述,本实施例表明基于珠的分选在将活化细胞与非活化细胞分离方面是成功的,并且该技术可用作使用遗传筛查方法从TCR文库中识别抗原特异性TCR的方法,作为基于FACS的细胞分离的替代方法。

[1192] 实施例22

[1193] 本实施例描述了使用成对TCR富集分析可以解析组合TMG编码。本实施例证明了使用组合TMG编码混合表达TMG的APC池的可能性,以及表达由TCR库中的TCR识别的抗原的TMG的分辨率。

[1194] 组合TMG编码允许有效地筛选大量TMG,而无需相应地缩放筛选数量,同时保持从TCR库筛选数据中识别出由TCR先导识别的TMG的潜力。组合TMG编码基于制作表达TMG的APC

池的原则,并且每个TMG精确地表征在一个池的组合中。由于池的组合对于每个TMG来说都是唯一的,因此可以在TCR文库筛选中识别由TCR识别的TMG。可使用成对TCR富集分析识别由来自TCR文库筛选的给定TCR先导识别的编码抗原的TMG,该分析通过特异性分析成对的池而不是单独的单个池来增加信号强度。

[1195] 图37A表示组合TMG设计的一个实例。12个池(C1-6和R1-6)中表达了36个TMG,每个池表达6个TMG。表达TMG的APC池可以通过多克隆APC工程改造(例如,多克隆病毒生产和APC转导或单克隆病毒生产,将病毒混合到池中,以及多克隆APC转导)或通过混合各自表达单个TMG的工程化APC细胞系来创建。例如,如果成对TCR富集分析显示APC池C1和R1都能激活给定的TCR,则该TCR正在识别由TMG1表达的抗原。这是因为TMG1是两个池中唯一代表的TMG。虽然在图37A)中,组合TMG编码基于恰好存在于两个池中的TMG,但是可以进行类似的组合TMG编码设计,其中每个TMG恰好存在于3、4、5、6、7、8、9、10个以上的池中。如果每个TMG恰好存在于3个池中,则通过识别给定TCR识别的3个池,可以推断出识别的TMG。被识别的TMG是所有3个池共有的TMG。

[1196] 在图10G)中,显示了三重筛选的结果,其中包括3个对照筛选,其中表达TCR文库的报告T细胞与未被工程改造以表达TMG的APC共培养,和3个基于与各自表达单个TMG的APC细胞系池的共培养的筛选。如上所述的组合TMG编码(其中每个TMG恰好在2个池中表征)在基于两次重复的筛选方法中涉及足够的TCR识别灵敏度。使用了DESeq2 R套装软件对来自图10G)中的数据的差异表征的TCR $\alpha \times \beta$ 链组合进行统计分析。如图37B)和37C)所示,基于3次重复中的2次,我们测试了在CRC患者筛查(图10G)中识别的TCR先导的排序或显著性是否相似。如图37B所示),当为pt2挑选7个TCR先导时,基于3个重复中的2个重复的任何组合的组合分析,将挑选相同的TCR。以类似的方式,对于图10G中识别的任何TCR先导,相同(组)的TCR将基于使用3个重复中的2个被挑选出来而进行统计分析(图37C“排序”列)。此外,对于在4个CRC患者TCR库筛选中确定的任何TCR先导,基于2或3次重复的统计分析,调整后的p值相似(图37C)。这表明2次重复足够进行TCR识别,且与使用3次重复的筛选具有相似的灵敏度。这允许使用组合TMG设计来识别已识别出(recognized)的TMG,其中每个TMG在TMG池中被精确地表征两次。

[1197] 为了测试成对TCR富集分析以解决从组合TMG编码设计的TMG识别的灵敏度,选择6个pt4样本(每个样本均由顶部和底部样本组成)进行分析(图37D)。使用DESeq2分析了6份样本的所有15种可能的成对组合,以获得顶部样本相对于底部样本的TCR差异表征。使用交互模型进行分析,其中对比了顶部样本相对于一对池底部样本的TCR富集,以及顶部样本相对于所有其他池底部样本的TCR富集。所有调整后的p值按从最低到最高p值的顺序排序,如图37E所示)。两个经调整的p值与其他值不同。这些值表示样本4和5的成对TCR富集分析中的TCR $\alpha 9\beta 14$,以及样本3和5的成对TCR富集分析中的TCR $\alpha 43\beta 16$ 。这些结果表明,在样本4和5中表征的TMG4以及在样本3和5中表征的TMG3编码由这些相应的TCR识别的抗原。

[1198] 综上所述,本实施例表明,各自表达多种TMG的APC池可用于针对大量小基因进行TCR文库筛选,并且这种组合TMG编码方法允许识别抗原反应性TCR以及表达所识别抗原的TMG。

[1199] 实施例23

[1200] 本实施例描述了可从遗传筛查数据中获得TCR特征。本实施例表明,TCR文库筛选

方法对于推导TCR特征足够灵敏。为此,将来自TCR识别平台的功能性遗传筛查数据与关于TCR诱导活化和背景活化的独立验证实验进行比较。

[1201] 评估了图10G)中识别的16个CRC pt2 TCR对于识别出的TMG (TMG2)的反应性。在遗传筛查方法中观察到的活化和在验证实验中观察到的活化之间观察到良好的相关性(图38);中间图)。此外,在遗传筛查方法中观察到的背景活化和在验证实验中观察到的背景活化之间观察到良好的相关性(图38);右图)。

[1202] 这些数据表明,可以在筛选阶段从功能性遗传筛查数据中确定TCR特征,这些特征可以在使用TCR文库筛选方法进行TCR鉴定后的独立实验中确定。例如,可以从TCR文库筛选中确定相对TCR灵敏度,以及在不存在抗原的情况下的TCR活化,如图38)中例示的。相对TCR灵敏度对于确定用于治疗癌症患者的T细胞产品工程化的最佳TCR可能很重要。在不存在抗原的情况下,TCR激活可能构成负向选择标准,因为表达此类TCR的工程改造的T细胞可能识别非癌细胞上存在的抗原。

[1203] 在TCR文库筛选阶段可包括额外的筛选,以减少筛选阶段后需要进行的TCR特性分析的数量。这些包括但不限于:i)使用表达野生型TMG的APC作为对照进行筛选,以在TCR文库筛选阶段确定TCR的突变特异性;ii)使用被工程改造为缺乏I类或II类表现(class II presentation)的基本成分(B2M或CIITA/CD74)的APC和被工程改造为表达TMG作为对照的筛选,以确定TCR的I/II类限制性;iii)使用来自不同强度启动子来表达TMG的APC进行筛选,以确定TCR灵敏度。

[1204] 总之,本实施例表明,可以从TCR文库筛选中确定TCR特征,这具有减少在TCR文库筛选阶段之后需要进行的TCR表征实验的数量的优点,并且因此可以减少从筛选到包括TCR表征所需的时间的总量。

[1205] 实施例24

[1206] 通过分离报告T细胞亚群从TCR $\alpha\beta$ 文库中恢复抗原反应性TCR(“顶部-底部法(Top-bottom approach)”)

[1207] 本实施例描述了通过基于对抗原的应答来分离一个或多个亚群从TCR $\alpha\beta$ 文库中恢复抗原反应性TCR。简而言之,该方法需要以下步骤:i)报告T细胞的基因工程化改造以允许表达TCR $\alpha\beta$ 文库的TCR;ii)进行这些细胞与表达至少一种抗原的抗原呈递细胞的共培养;iii)基于T细胞活化标志物的细胞分离成a)表达一种或多种T细胞活化标志物的“顶部”群;和b)缺乏(或具有低)表达一种或多种T细胞活化标记的“底部”群;iv)使用对基因组DNA的PCR和随后的深度测序从顶部和底部样本进行TCR鉴定;和v)识别相对于底部样本在顶部样本中富集的至少一种抗原反应性TCR。T细胞活化标志物的表达可以是区分活化T细胞的标志物的相对高表达(CD69),或区分非活化T细胞的标志物的相对低表达(CD62L)。

[1208] 顶部-底部法背后的原理是,抗原反应性TCR在抗原刺激后将变为活化标志物阳性,因此相对于底部群,此类TCR将在顶部群中富集。如下所述,通过各种附图说明了顶部-底部比较。

[1209] 在替代实施方案中,底部样本可以是任何参考细胞群或TCR质粒参考文库。底部样本可以从与顶部样本相同的细胞群中分选出来,但是具有低活化标志物表达。底部样本可从表达相关TCR文库的报告T细胞和未被工程改造以表达外源抗原的B细胞的共培养物中获得。底部样本可以是用于产生报告T细胞的TCR质粒文库,顶部样本是从该报告T细胞中分选

出来的。在一些实施方案中,在差异TCR表征分析期间,可以将顶部和底部样本中的TCR表征 (representation) 与任何其他另外的样本中的TCR表征 (representation) 进行比较。在一些实施方案中,此类另外的样本可以是质粒TCR文库。在其他实施方案中,此类另外的样本可以衍生自表达相关TRC文库的报告T细胞和未被工程化以表达外源抗原的B细胞的共培养物。

[1210] 图39) 中的rlog值表示TCR表征 (representation) 的度量。该图显示,与任何其他样本相比,受同源抗原刺激的五种特征性TCR在顶部样本(最浅的灰色阴影)中富集。来自与表达TMG的B细胞共培养物的底部样本显示出比来自与未被工程改造以表达外源抗原的B细胞共培养的顶部或底部样本更低的rlog值。基于这些结果,本领域技术人员将认识到,使用顶部-底部法识别抗原反应性TCR使用i) 来自与顶部样本相同的细胞群但具有低活化标志物表达(底部+TMG)的底部样本将比使用ii) 来自与未被工程改造以表达外源抗原的B细胞的共培养物的底部样本更灵敏。这是因为顶部+TMG和底部+TMG样本之间的差异大于顶部+TMG和顶部/底部-TMG样本之间的差异,因此可以使用前一种比较更灵敏地检测TCR反应性。底部样本(即,参考TCR表征 (representation)) 可来自相关质粒TCR文库(图39中的黑点), 其与来自与未被工程改造以表达外源抗原的B细胞的共培养物的顶部或底部样本中的TCR表征 (representation) 相关。

[1211] 该实施例的其他方面显示在图39中,其显示了对来自图14F,G) 的100×100组合文库的遗传筛查的附加分析。本示例说明了顶部-底部法(top-bottom approach), 其中底部样本中的细胞理想地被从与顶部样本中的细胞相同的群中分选(sorted) 出来,但活化标志物为阴性。在图39) 中,使用DESeq2 R套装软件对图14F,G) 的数据进行Rlog变换,并代表五个特征性TCR中的每一个。每个点代表一个独立的重复,并且根据样本类型对点进行着色。顶部样本代表CD69表达最高的前10%细胞,而底部样本代表从相同细胞群中分选出来但CD69表达较低的细胞。在图16E) 中,将顶部-底部法应用于从24个T细胞系创建的Jurkat hCD8+细胞池,每个T细胞系表达具有已知身份但具有未知特异性的不同TCR。这些细胞系以1:1的比例混合,表达CDK4-17、CDK4-8、CMV#1、CMV#2或GCN1L1 TCR的Jurkat hCD8+细胞也以1:10000-1:1000000的不同频率混合。与表达这5种TCR的同源抗原的B细胞共培养后,对细胞进行CD69(T细胞活化标志物) 染色。CD69表达最高的前10%的T细胞被FACS分选到顶部样本中,大致等量的表达低水平CD69的T细胞被FACS分选到底部样本中。使用基因组DNA PCR分析两个样本中的TCR含量,然后使用Illumina平台进行下一代测序。计算了顶部与底部样本中TCR的相对丰度,并使用此方法观察到所有五种已知特异性的TCR富集。

[1212] 在图13A) -E) 中,通过将以1:10000-1:100000频率表达六种抗原特异性TCR的质粒混合到具有已知身份但未知特异性的质粒池中来创建TCR文库。该质粒池用于产生逆转录病毒并感染Jurkat hCD8+TCR KO细胞。在该多克隆报告T细胞池和表达抗原特异性TCR同源抗原的B细胞之间进行共培养。CD69表达最高的前10%的T细胞被FACS分选到顶部样本中,大致等量的表达低水平CD69的T细胞被FACS分选到底部样本中(图13C)。使用基因组DNA PCR分析两个样本中的TCR含量(图13D), 然后使用Illumina平台进行下一代测序。计算了顶部与底部样本中TCR的相对丰度,并使用该方法观察到所有五种已知特异性的TCR富集(图13E)。

[1213] 在图10A) -J) 中,从使用结直肠癌(CRC) 样本的批量TCR测序获得的最普遍的TCRα

和TCR β 链创建组合TCR文库。该文库用于Jurkat hCD8+TCR KO细胞的逆转录病毒转导(图10D), Jurkat hCD8+TCR KO细胞与表达肿瘤特异性新生抗原的EBV-B细胞共培养。CD69表达最高的前10%的T细胞被FACS分选到顶部样本中, 大致等量的表达低水平CD69的T细胞被FACS分选到底部样本中(图10E)。使用基因组DNA PCR分析所有样本中的TCR含量(图10F), 然后使用Oxford Nanopore技术进行下一代测序。计算了顶部样本与底部样本中TCR的相对丰度。对另外三个CRC标本采用这种方法, 我们共识别出11个候选的新生抗原反应性TCR(图10G)。其中四项已在独立共培养实验中得到进一步验证(图10H-K)。

[1214] 在图36A)-E)中, 使用autoMACS (Miltenyi, Bergisch Gladbach, 德国) 基于珠的分选法分离顶部和底部样本。使用来自卵巢癌 (OVC) 或结直肠癌 (CRC) 样本的5个特征性TCR和95个非特征性TCR来创建100 $\times$ 100设计的组合TCR文库, 并在报告T细胞中表达。在与表达五种特征性TCR的同源抗原的B细胞共培养后, 使用CD20-微珠耗竭B细胞。然后使用CD62L和CD69标志物, 使用基于珠的连续分选 (sequential bead-based sorting) 获得顶部和底部样本。CD62L是非活化T细胞上表达的标志物, 而CD69是活化T细胞表达的标志物。将不结合CD62L-微珠且确实结合CD69-微珠的细胞分离到顶部样本中, 将确实结合CD62L-微珠且不结合CD69-微珠的细胞分离到底部样本中。使用基因组DNA PCR分析两个样本中的TCR含量(图36C), 然后使用Oxford Nanopore技术和筛选分析进行下一代测序(图36D,E)。五种特征性TCR在顶部样本相对于底部样本最显著富集的7种TCR之中。这表明, 在TCR文库筛选方法中, 基于珠的分选可用作分离顶部和底部样本的替代策略。

[1215] 本实施例表明, 顶部-底部法通常可应用于表达多种TCR的报告T细胞, 以允许对抗原反应性TCR进行功能性遗传筛查。顶部-底部法可以独立于创建表达多个TCR的多克隆报告T细胞的方式而应用。通过从以各种方式获得的表达TCR的报告T细胞中识别抗原反应性TCR来支持这一点: i) 通过混合各自表达单个定义的TCR的报告T细胞系; ii) 通过混合编码定义的TCR和多克隆病毒生产的质粒以及报告T细胞的转导; 和iii) 通过使用用于多克隆病毒生产和转导的报告T细胞TCR文库。

[1216] 此外, 本实施例表明顶部-底部法可应用于各种细胞分离和TCR识别技术: i) 用不同T细胞活化标志物 (CD69单独或CD62L/CD69组合) 染色; ii) 用不同细胞分选技术 (FACS分选和基于珠的分选) 进行分离; 和iii) 使用不同的下一代测序技术 (Illumina和Oxford Nanopore技术) 的分析。

[1217] 在一些实施方案中, 上述实施例中的方法 (顶部-底部法) 可以应用于创建TCR库的任何方法, 包括但不限于: i) 图15中描述的文库设计的各种方式; ii) 基于可识别在单个细胞内表达的TCRa链和TCRb链序列 (配对TCR测序) 的TCR测序技术的文库设计; iii) 基于TCRa和TCRb序列的细胞内连接和随后的克隆的文库设计。顶部-底部法可以基于通过任何T细胞活化标志物或T细胞活化标志物的组合进行的细胞分离, 所述T细胞活化标志物包括但不限于CD69、CD62L、CD137、CD25、IFN- γ 、IL-2、TNF- α 、GM-CSF。顶部-底部法可应用于任何增殖标志物, 包括但不限于CellTrace染色。顶部-底部法可应用于基于任何T细胞活化报道基因的细胞分离, 包括但不限于CD69 (启动子) -EGFRt和CD69 (启动子) -LNGFR报道基因。顶部-底部法可以应用于任何细胞分离技术, 包括但不限于基于FACS的分选、基于微流体的细胞分选和基于珠的细胞分选。顶部-底部法可以使用任何下一代测序技术进行分析, 包括但不限于使用Illumina、Oxford Nanopore或Pacific Biosciences技术进行测序。

〈110〉 荷兰商新基因治疗公司 (Neogene Therapeutics B.V.)

〈120〉 分离TCR基因的方法

<130> KHP222110026.6

<150> US 62/874125

<151> 2019-07-15

<150> US 62/975924

<151> 2020-02-13

<150> US 63/024341

<151> 2020-05-13

<150> US 63/034157

<151> 2020-06-03

<150> US 63/039346

<151> 2020-06-15

<160> 4

<170> PatentIn version 3.5

 $\langle 210 \rangle$ 1

<211> 819

<212> PRT

〈213〉 人工序列 (Artificial Sequence)

 $\langle 220 \rangle$

〈223〉合成的：表示为

TCR β -P2A-TCR α -T2A-嘌呤霉素抗性表达盒的TCR 基因的氨基酸序列

 $\langle 400 \rangle$ 1

Met Gly Thr Arg Leu Leu Cys Trp Ala Ala Leu Cys Leu Leu Gly Ala

1 5 10 15

Glu Leu Thr Glu Ala Gly Val Ala Gln Ser Pro Arg Tyr Lys Ile Ile

20 25 30

Glu Lys Arg Gln Ser Val Ala Phe Trp Cys Asn Pro Ile Ser Gly His

35 40 45

Ala Thr Leu Tyr Trp Tyr Gln Gln Ile Leu Gly Gln Gly Pro Lys Leu

50 55 60

Leu Ile Gln Phe Gln Asn Asn Gly Val Val Asp Asp Ser Gln Leu Pro

65 70 75 80

Lys Asp Arg Phe Ser Ala Glu Arg Leu Lys Gly Val Asp Ser Thr Leu

85 90 95

Lys Ile Gln Pro Ala Lys Leu Glu Asp Ser Ala Val Tyr Leu Cys Ala

100 105 110

Ser Ser Leu Tyr Pro Tyr Asp Gly Arg Ser Ser Tyr Glu Gln Tyr Phe

115	120	125
Gly Pro Gly Thr Arg Leu Thr	Val Thr Glu Asp Leu Arg Asn Val Thr	
130	135	140
Pro Pro Lys Val Ser Leu Phe Glu Pro Ser Lys Ala Glu Ile Ala Asn		
145	150	155
Lys Gln Lys Ala Thr Leu Val Cys Leu Ala Arg Gly Phe Phe Pro Asp		
165	170	175
His Val Glu Leu Ser Trp Trp Val Asn Gly Lys Glu Val His Ser Gly		
180	185	190
Val Cys Thr Asp Pro Gln Ala Tyr Lys Glu Ser Asn Tyr Ser Tyr Cys		
195	200	205
Leu Ser Ser Arg Leu Arg Val Ser Ala Thr Phe Trp His Asn Pro Arg		
210	215	220
Asn His Phe Arg Cys Gln Val Gln Phe His Gly Leu Ser Glu Glu Asp		
225	230	235
Lys Trp Pro Glu Gly Ser Pro Lys Pro Val Thr Gln Asn Ile Ser Ala		
245	250	255
Glu Ala Trp Gly Arg Ala Asp Cys Gly Ile Thr Ser Ala Ser Tyr His		
260	265	270
Gln Gly Val Leu Ser Ala Thr Ile Leu Tyr Glu Ile Leu Leu Gly Lys		
275	280	285
Ala Thr Leu Tyr Ala Val Leu Val Ser Gly Leu Val Leu Met Ala Met		
290	295	300
Val Lys Lys Lys Asn Ser Gly Ser Gly Ala Thr Asn Phe Ser Leu Leu		
305	310	315
Lys Gln Ala Gly Asp Val Glu Glu Asn Pro Gly Pro Met Thr Ser Ile		
325	330	335
Arg Ala Val Phe Ile Phe Leu Trp Leu Gln Leu Asp Leu Val Asn Gly		
340	345	350
Glu Asn Val Glu Gln His Pro Ser Thr Leu Ser Val Gln Glu Gly Asp		
355	360	365
Ser Ala Val Ile Lys Cys Thr Tyr Ser Asp Ser Ala Ser Asn Tyr Phe		
370	375	380
Pro Trp Tyr Lys Gln Glu Leu Gly Lys Gly Pro Gln Leu Ile Ile Asp		
385	390	395
Ile Arg Ser Asn Val Gly Glu Lys Lys Asp Gln Arg Ile Ala Val Thr		
405	410	415
Leu Asn Lys Thr Ala Lys His Phe Ser Leu His Ile Thr Glu Thr Gln		
420	425	430

Pro	Glu	Asp	Ser	Ala	Val	Tyr	Phe	Cys	Ala	Ala	Tyr	Asn	Gln	Gly	Gly				
																435	440	445	
Lys	Leu	Ile	Phe	Gly	Gln	Gly	Thr	Glu	Leu	Ser	Val	Lys	Pro	Asp	Ile				
																450	455	460	
Gln	Asn	Pro	Glu	Pro	Ala	Val	Tyr	Gln	Leu	Lys	Asp	Pro	Arg	Ser	Gln				
																465	470	475	480
Asp	Ser	Thr	Leu	Cys	Leu	Phe	Thr	Asp	Phe	Asp	Ser	Gln	Ile	Asn	Val				
																485	490	495	
Pro	Lys	Thr	Met	Glu	Ser	Gly	Thr	Phe	Ile	Thr	Asp	Lys	Cys	Val	Leu				
																500	505	510	
Asp	Met	Lys	Ala	Met	Asp	Ser	Lys	Ser	Asn	Gly	Ala	Ile	Ala	Trp	Ser				
																515	520	525	
Asn	Gln	Thr	Ser	Phe	Thr	Cys	Gln	Asp	Ile	Phe	Lys	Glu	Thr	Asn	Ala				
																530	535	540	
Thr	Tyr	Pro	Ser	Ser	Asp	Val	Pro	Cys	Asp	Ala	Thr	Leu	Thr	Glu	Lys				
																545	550	555	560
Ser	Phe	Glu	Thr	Asp	Met	Asn	Leu	Asn	Phe	Gln	Asn	Leu	Ser	Val	Met				
																565	570	575	
Gly	Leu	Arg	Ile	Leu	Leu	Leu	Lys	Val	Ala	Gly	Phe	Asn	Leu	Leu	Met				
																580	585	590	
Thr	Leu	Arg	Leu	Trp	Ser	Ser	Gly	Ser	Gly	Glu	Gly	Arg	Gly	Ser	Leu				
																595	600	605	
Leu	Thr	Cys	Gly	Asp	Val	Glu	Glu	Asn	Pro	Gly	Pro	Met	Thr	Glu	Tyr				
																610	615	620	
Lys	Pro	Thr	Val	Arg	Leu	Ala	Thr	Arg	Asp	Asp	Val	Pro	Arg	Ala	Val				
																625	630	635	640
Arg	Thr	Leu	Ala	Ala	Ala	Phe	Ala	Asp	Tyr	Pro	Ala	Thr	Arg	His	Thr				
																645	650	655	
Val	Asp	Pro	Asp	Arg	His	Ile	Glu	Arg	Val	Thr	Glu	Leu	Gln	Glu	Leu				
																660	665	670	
Phe	Leu	Thr	Arg	Val	Gly	Leu	Asp	Ile	Gly	Lys	Val	Trp	Val	Ala	Asp				
																675	680	685	
Asp	Gly	Ala	Ala	Val	Ala	Val	Trp	Thr	Thr	Pro	Glu	Ser	Val	Glu	Ala				
																690	695	700	
Gly	Ala	Val	Phe	Ala	Glu	Ile	Gly	Pro	Arg	Met	Ala	Glu	Leu	Ser	Gly				
																705	710	715	720
Ser	Arg	Leu	Ala	Ala	Gln	Gln	Gln	Met	Glu	Gly	Leu	Leu	Ala	Pro	His				
																725	730	735	
Arg	Pro	Lys	Glu	Pro	Ala	Trp	Phe	Leu	Ala	Thr	Val	Gly	Val	Ser	Pro				

740	745	750
Asp His Gln Gly Lys Gly Leu Gly Ser Ala Val Val Leu Pro Gly Val		
755	760	765
Glu Ala Ala Glu Arg Ala Gly Val Pro Ala Phe Leu Glu Thr Ser Ala		
770	775	780
Pro Arg Asn Leu Pro Phe Tyr Glu Arg Leu Gly Phe Thr Val Thr Ala		
785	790	795
Asp Val Glu Val Pro Glu Gly Pro Arg Thr Trp Cys Met Thr Arg Lys		
805	810	815
Pro Gly Ala		
<210> 2		
<211> 467		
<212> PRT		
<213> 人工序列 (Artificial Sequence)		
<220>		
<223> 合成的; CD8 α -P2A-CD8 β 转基因的氨基酸序列		
<400> 2		
Met Ala Leu Pro Val Thr Ala Leu Leu Leu Pro Leu Ala Leu Leu Leu		
1 5 10 15		
His Ala Ala Arg Pro Ser Gln Phe Arg Val Ser Pro Leu Asp Arg Thr		
20 25 30		
Trp Asn Leu Gly Glu Thr Val Glu Leu Lys Cys Gln Val Leu Leu Ser		
35 40 45		
Asn Pro Thr Ser Gly Cys Ser Trp Leu Phe Gln Pro Arg Gly Ala Ala		
50 55 60		
Ala Ser Pro Thr Phe Leu Leu Tyr Leu Ser Gln Asn Lys Pro Lys Ala		
65 70 75 80		
Ala Glu Gly Leu Asp Thr Gln Arg Phe Ser Gly Lys Arg Leu Gly Asp		
85 90 95		
Thr Phe Val Leu Thr Leu Ser Asp Phe Arg Arg Glu Asn Glu Gly Tyr		
100 105 110		
Tyr Phe Cys Ser Ala Leu Ser Asn Ser Ile Met Tyr Phe Ser His Phe		
115 120 125		
Val Pro Val Phe Leu Pro Ala Lys Pro Thr Thr Thr Pro Ala Pro Arg		
130 135 140		
Pro Pro Thr Pro Ala Pro Thr Ile Ala Ser Gln Pro Leu Ser Leu Arg		
145 150 155 160		
Pro Glu Ala Cys Arg Pro Ala Ala Gly Gly Ala Val His Thr Arg Gly		
165 170 175		

Leu Asp Phe Ala Cys Asp Ile Tyr Ile Trp Ala Pro Leu Ala Gly Thr		
180	185	190
Cys Gly Val Leu Leu Leu Ser Leu Val Ile Thr Leu Tyr Cys Asn His		
195	200	205
Arg Asn Arg Arg Arg Val Cys Lys Cys Pro Arg Pro Val Val Lys Ser		
210	215	220
Gly Asp Lys Pro Ser Leu Ser Ala Arg Tyr Val Gly Ser Gly Ala Thr		
225	230	235
Asn Phe Ser Leu Leu Lys Gln Ala Gly Asp Val Glu Glu Asn Pro Gly		
245	250	255
Pro Met Arg Pro Arg Leu Trp Leu Leu Leu Ala Ala Gln Leu Thr Val		
260	265	270
Leu His Gly Asn Ser Val Leu Gln Gln Thr Pro Ala Tyr Ile Lys Val		
275	280	285
Gln Thr Asn Lys Met Val Met Leu Ser Cys Glu Ala Lys Ile Ser Leu		
290	295	300
Ser Asn Met Arg Ile Tyr Trp Leu Arg Gln Arg Gln Ala Pro Ser Ser		
305	310	315
Asp Ser His His Glu Phe Leu Ala Leu Trp Asp Ser Ala Lys Gly Thr		
325	330	335
Ile His Gly Glu Glu Val Glu Gln Glu Lys Ile Ala Val Phe Arg Asp		
340	345	350
Ala Ser Arg Phe Ile Leu Asn Leu Thr Ser Val Lys Pro Glu Asp Ser		
355	360	365
Gly Ile Tyr Phe Cys Met Ile Val Gly Ser Pro Glu Leu Thr Phe Gly		
370	375	380
Lys Gly Thr Gln Leu Ser Val Val Asp Phe Leu Pro Thr Thr Ala Gln		
385	390	395
Pro Thr Lys Lys Ser Thr Leu Lys Lys Arg Val Cys Arg Leu Pro Arg		
405	410	415
Pro Glu Thr Gln Lys Gly Pro Leu Cys Ser Pro Ile Thr Leu Gly Leu		
420	425	430
Leu Val Ala Gly Val Leu Val Leu Leu Val Ser Leu Gly Val Ala Ile		
435	440	445
His Leu Cys Cys Arg Arg Arg Arg Ala Arg Leu Arg Phe Met Lys Gln		
450	455	460
Phe Tyr Lys		
465		
<210> 3		

<211> 2405

<212> DNA

<213> 人工序列 (Artificial Sequence)

<220>

<223> 合成的; K562-HLA-A*02:01-IRES-FusionRed的核苷酸序列

<400> 3

```

atggctgtta tggcccctag aacactggtg ctgctgctgt ctggtgccct ggctctgaca 60
caaacatggg ccggcagcca cagcatgcgg tactttttca ccagcgtgtc cagacctggc 120
agaggcgagc ctagattcat tgccgtgggc tacgtggacg acaccagtt cgtcagattc 180
gattccgatg ccgccagcca gcggatggaa ctagagcac cttggatcga gcaagagggc 240
cccagatatt gggacggcga gacaagaaaa gtgaaggccc acagccagac acacagagtg 300
gatctgggaa ccctgcgggg ctactacaat cagtctgagg ccggctctca caccgtgcag 360
aggatgtatg gctgtgacgt gggcagcgat tggcggttcc tgagaggcta tcaccagtac 420
gcctacgacg gcaaggacta tatcgccctg aaagaggacc tgcggtcttg gacagccgcc 480
gatatggctg ccagaccac aaagcacaag tgggaagccg ctcacgtggc cgaacagctg 540
agagcttata tggaaggcac ctgtgtggaa tggctgcgga gatacctgga aaacggcaaa 600
gagacactgc agcggacaga cgcccctaag acacacatga cacaccacgc cgtgtccgac 660
cacgaagcca cacttagatg ttgggccctg agcttctacc ccgccgagat cactctgaca 720
tggcagagag atggcgagga tcagaccag gataccgagc tgggtgaaac aagacctgcc 780
ggcgacggca ccttccagaa atgggctgct gtggtggtgc ctagcggcca agagcagaga 840
tacacctgtc acgtgcagca cgaggccctg cctaagcctc ttactactgag atgggagccc 900
agcagccagc ctacaatccc catcgtggga atcattgccg gcctggtgct gtttggcgcc 960
gtgattacag gtgcagtggg ggccgctgtt atgtggcgga gaaagagcag cgacagaaaa 1020
ggcggcagct actctcaggc cgccagctct gattctgccc agggctctga tgtgtctctg 1080
accgcctgca aagtctgagc ggccgctacg taaattccgc cccccccct ccccccccc 1140
taacgttact ggccgaagcc gcttgaata aggccggtgt gcgtttgtct atatgttatt 1200
ttccaccata ttgccgtctt ttggcaatgt gagggcccg aaacctggcc ctgtcttctt 1260
gacgagcatt ctaggggtc tttccctct cgccaaagga atgcaaggtc tgttgaatgt 1320
cgtgaaggaa gcagttcctc tggaagcttc ttgaagacaa acaacgtctg tagcgacct 1380
ttgcaggcag cggaaccccc cacctggcga caggtgcctc tgcggccaaa agccacgtgt 1440
ataagataca cctgcaaagg cggcacaacc ccagtgccac gttgtgagtt ggatagttgt 1500
ggaaagagtc aaatggctct cctcaagcgt attcaacaag gggctgaagg atgccagaa 1560
ggtaccccat tgtatgggat ctgatctggg gcctcggtgc acatgcttta catgtgttta 1620
gtcgaggtta aaaaaacgtc taggcccccc gaaccacggg gacgtggttt tcctttgaaa 1680
aacacgatga taatatggcc acaacctagg tgtccgagct gatcaaagaa aacatgcccc 1740
tgaagctgta catggaagg accgtgaaca accaccactt caagtgcacc agcgaaggcg 1800
agggcaagcc ttatgagggc acccagacca tgcggatcaa ggtggtggaa ggtggccctc 1860
tgcccttcgc ctttgatata ctggccacca gctttatgta cggcagccgg accttcatca 1920
agcaccctcc aggcattccc gatttcttca agcagagctt tcccagggc ttcacctggg 1980

```

agagagtgac cacatacgag gatggcggcg tgctgacagc cacacaggat acaagtctgc 2040
 aggacggctg cctgatctac aacgtgaaag tgcgggggcgt gaacttcccc gccaatggac 2100
 ctgtgatgca gaagaaaacc ctcggctggg aagccagcac cgagacaatg tatectgccg 2160
 acggtggact ggaagggggcc tgtgatatgg ccctgaagct tgttggcggc ggacacctga 2220
 tctgcaacct ggaaccacc tacagaagca agaagccgc caccaacctg aagatgcccg 2280
 gcgtgtacaa cgtggaccac cggctggaaa ggatcaaaga ggccgacgac gagacatacg 2340
 tggaacagca cgaagtggcc gtggccagat actctacagg cggagctggc gacggcggca 2400
 aatga 2405

<210> 4

<211> 755

<212> PRT

<213> 人工序列 (Artificial Sequence)

<220>

<223> 合成的; 表示为

TCR β -P2A-TCR α -T2A-杀稻瘟菌素抗性表达盒的TCR基因的氨基酸序列

<400> 4

Met	Gly	Pro	Gln	Leu	Leu	Gly	Tyr	Val	Val	Leu	Cys	Leu	Leu	Gly	Ala
1			5					10						15	
Gly	Pro	Leu	Glu	Ala	Gln	Val	Thr	Gln	Asn	Pro	Arg	Tyr	Leu	Ile	Thr
			20					25						30	
Val	Thr	Gly	Lys	Lys	Leu	Thr	Val	Thr	Cys	Ser	Gln	Asn	Met	Asn	His
			35					40						45	
Glu	Tyr	Met	Ser	Trp	Tyr	Arg	Gln	Asp	Pro	Gly	Leu	Gly	Leu	Arg	Gln
			50				55							60	
Ile	Tyr	Tyr	Ser	Met	Asn	Val	Glu	Val	Thr	Asp	Lys	Gly	Asp	Val	Pro
65					70					75					80
Glu	Gly	Tyr	Lys	Val	Ser	Arg	Lys	Glu	Lys	Arg	Asn	Phe	Pro	Leu	Ile
				85						90					95
Leu	Glu	Ser	Pro	Ser	Pro	Asn	Gln	Thr	Ser	Leu	Tyr	Phe	Cys	Ala	Ser
				100						105					110
Arg	Ala	Ser	Gly	Arg	Glu	Gln	Phe	Phe	Gly	Pro	Gly	Thr	Arg	Leu	Thr
				115						120					125
Val	Leu	Glu	Asp	Leu	Arg	Asn	Val	Thr	Pro	Pro	Lys	Val	Ser	Leu	Phe
				130						135					140
Glu	Pro	Ser	Lys	Ala	Glu	Ile	Ala	Asn	Lys	Gln	Lys	Ala	Thr	Leu	Val
145					150					155					160
Cys	Leu	Ala	Arg	Gly	Phe	Phe	Pro	Asp	His	Val	Glu	Leu	Ser	Trp	Trp
					165					170					175
Val	Asn	Gly	Lys	Glu	Val	His	Ser	Gly	Val	Cys	Thr	Asp	Pro	Gln	Ala

180	185	190
Tyr Lys Glu Ser Asn Tyr Ser Tyr Cys Leu Ser Ser Arg Leu Arg Val		
195	200	205
Ser Ala Thr Phe Trp His Asn Pro Arg Asn His Phe Arg Cys Gln Val		
210	215	220
Gln Phe His Gly Leu Ser Glu Glu Asp Lys Trp Pro Glu Gly Ser Pro		
225	230	235
Lys Pro Val Thr Gln Asn Ile Ser Ala Glu Ala Trp Gly Arg Ala Asp		
245	250	255
Cys Gly Ile Thr Ser Ala Ser Tyr His Gln Gly Val Leu Ser Ala Thr		
260	265	270
Ile Leu Tyr Glu Ile Leu Leu Gly Lys Ala Thr Leu Tyr Ala Val Leu		
275	280	285
Val Ser Gly Leu Val Leu Met Ala Met Val Lys Lys Lys Asn Ser Gly		
290	295	300
Ser Gly Ala Thr Asn Phe Ser Leu Leu Lys Gln Ala Gly Asp Val Glu		
305	310	315
Glu Asn Pro Gly Pro Met Lys Ser Leu Arg Val Leu Leu Val Ile Leu		
325	330	335
Trp Leu Gln Leu Ser Trp Val Trp Ser Gln Asp Gln Gln Val Lys Gln		
340	345	350
Asn Ser Pro Ser Leu Ser Val Gln Glu Gly Arg Ile Ser Ile Leu Asn		
355	360	365
Cys Asp Tyr Thr Asn Ser Met Phe Asp Tyr Phe Leu Trp Tyr Lys Lys		
370	375	380
Tyr Pro Ala Glu Gly Pro Thr Phe Leu Ile Ser Ile Ser Ser Ile Lys		
385	390	395
Asp Lys Asn Glu Asp Gly Arg Phe Thr Val Phe Leu Asn Lys Ser Ala		
405	410	415
Lys His Leu Ser Leu His Ile Val Pro Ser Gln Pro Gly Asp Ser Ala		
420	425	430
Val Tyr Phe Cys Ala Ala Phe Leu Gln Ser Asn Asp Met Arg Phe Gly		
435	440	445
Ala Gly Thr Arg Leu Thr Val Lys Pro Asp Ile Gln Asn Pro Glu Pro		
450	455	460
Ala Val Tyr Gln Leu Lys Asp Pro Arg Ser Gln Asp Ser Thr Leu Cys		
465	470	475
Leu Phe Thr Asp Phe Asp Ser Gln Ile Asn Val Pro Lys Thr Met Glu		
485	490	495

Ser Gly Thr Phe Ile Thr Asp Lys Cys Val Leu Asp Met Lys Ala Met		
500	505	510
Asp Ser Lys Ser Asn Gly Ala Ile Ala Trp Ser Asn Gln Thr Ser Phe		
515	520	525
Thr Cys Gln Asp Ile Phe Lys Glu Thr Asn Ala Thr Tyr Pro Ser Ser		
530	535	540
Asp Val Pro Cys Asp Ala Thr Leu Thr Glu Lys Ser Phe Glu Thr Asp		
545	550	555
Met Asn Leu Asn Phe Gln Asn Leu Ser Val Met Gly Leu Arg Ile Leu		
565	570	575
Leu Leu Lys Val Ala Gly Phe Asn Leu Leu Met Thr Leu Arg Leu Trp		
580	585	590
Ser Ser Gly Ser Gly Glu Gly Arg Gly Ser Leu Leu Thr Cys Gly Asp		
595	600	605
Val Glu Glu Asn Pro Gly Pro Met Lys Thr Phe Asn Ile Ser Gln Gln		
610	615	620
Asp Leu Glu Leu Val Glu Val Ala Thr Glu Lys Ile Thr Met Leu Tyr		
625	630	635
Glu Asp Asn Lys His His Val Gly Ala Ala Ile Arg Thr Lys Thr Gly		
645	650	655
Glu Ile Ile Ser Ala Val His Ile Glu Ala Tyr Ile Gly Arg Val Thr		
660	665	670
Val Cys Ala Glu Ala Ile Ala Ile Gly Ser Ala Val Ser Asn Gly Gln		
675	680	685
Lys Asp Phe Asp Thr Ile Val Ala Val Arg His Pro Tyr Ser Asp Glu		
690	695	700
Val Asp Arg Ser Ile Arg Val Val Ser Pro Cys Gly Met Cys Arg Glu		
705	710	715
Leu Ile Ser Asp Tyr Ala Pro Asp Cys Phe Val Leu Ile Glu Met Asn		
725	730	735
Gly Lys Leu Val Lys Thr Thr Ile Glu Glu Leu Ile Pro Leu Lys Tyr		
740	745	750
Thr Arg Asn		
755		

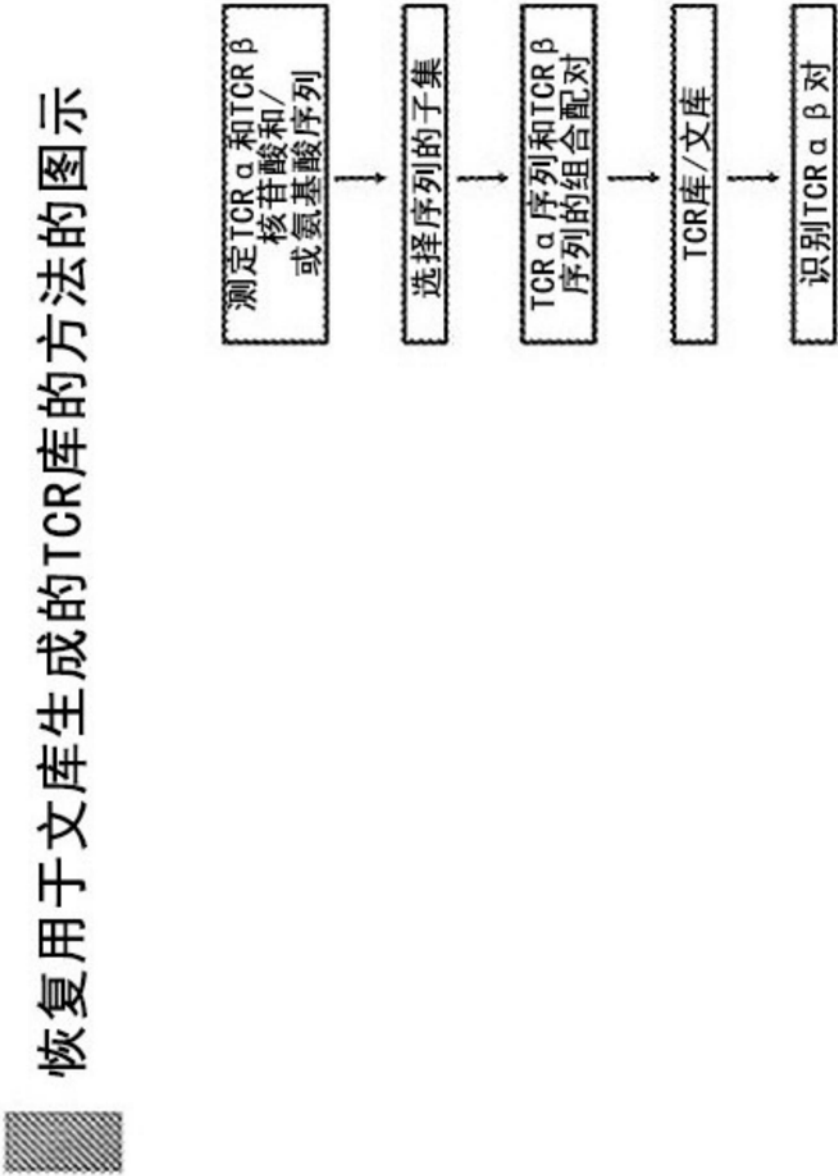


图1

恢复TCR库并识别用于癌症治疗的抗原特异性TCR α 和TCR β 序列的方法的图示

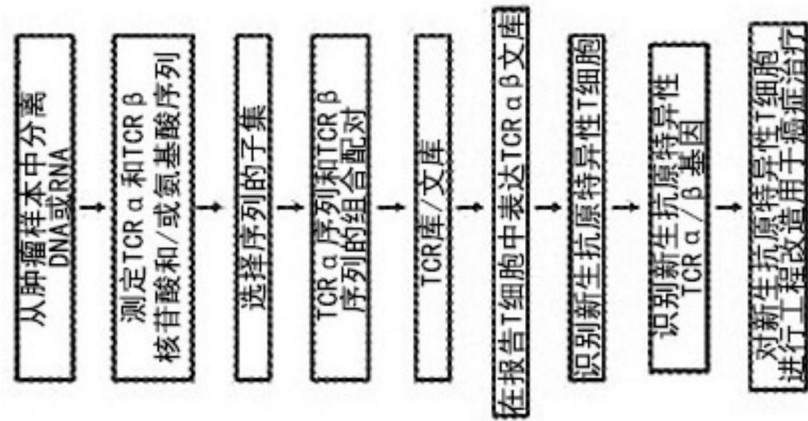


图2



TCR表达盒的示例性设计

TCRβ V-区	TCRβ CDR3	TCRβ J-片段	TCRβ C-链构建	P2A	TCRα V-区	TCRα CDR3	TCRα J-片段	TCRα C-结构域
321-357 bp	15-90 bp	27 bp	519bp	66 bp	303-351 bp	15-90 bp	30 bp	414 bp
几种人类变体	对于每个 TCR都是独一无二的	几种人类变体	对于每个TCR都是相同的	对于每个TCR盒都是相同的	几种人类变体	对于每个 TCR都是独一无二的	几种人类变体	对于每个TCR都是相同的

图3

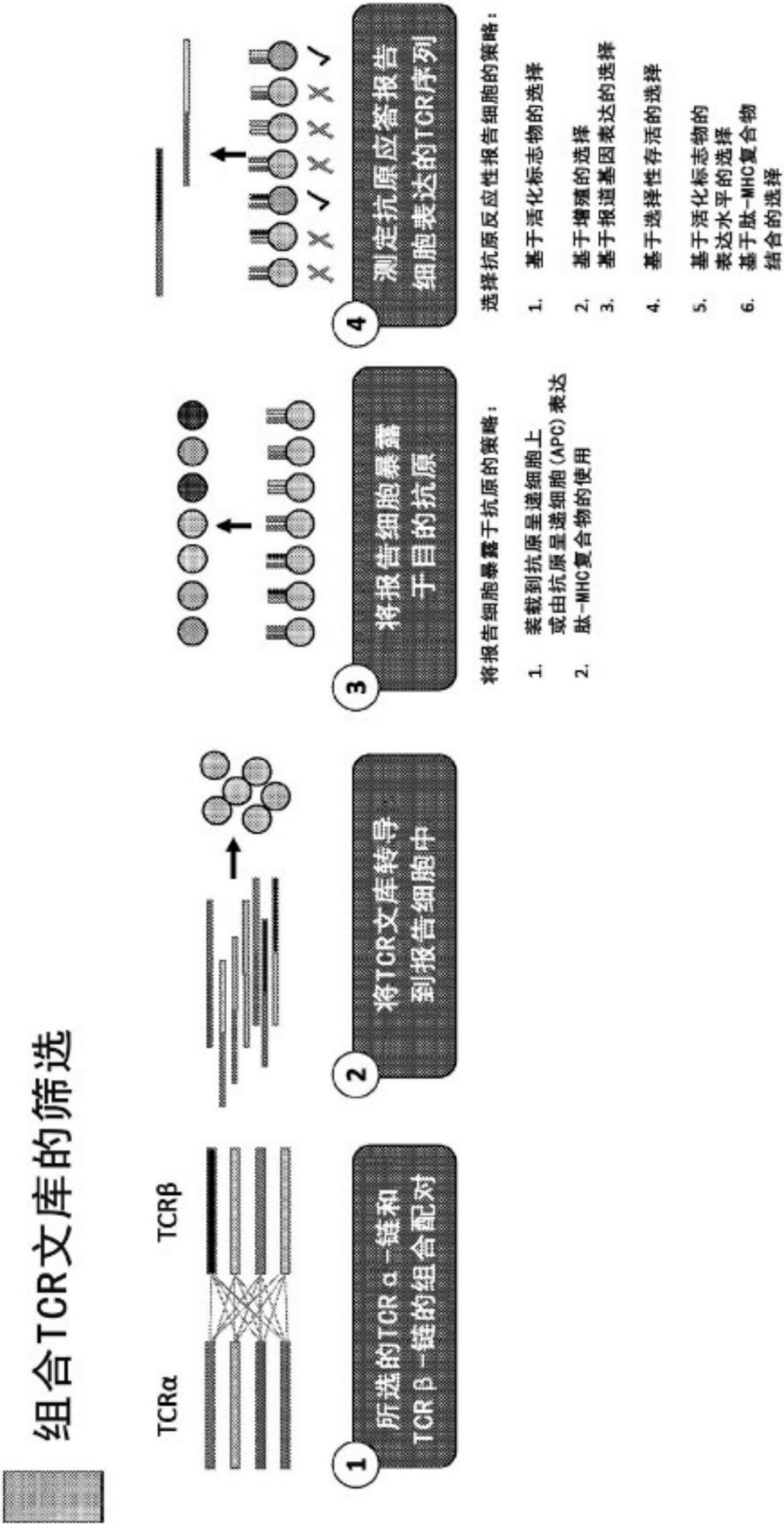


图4

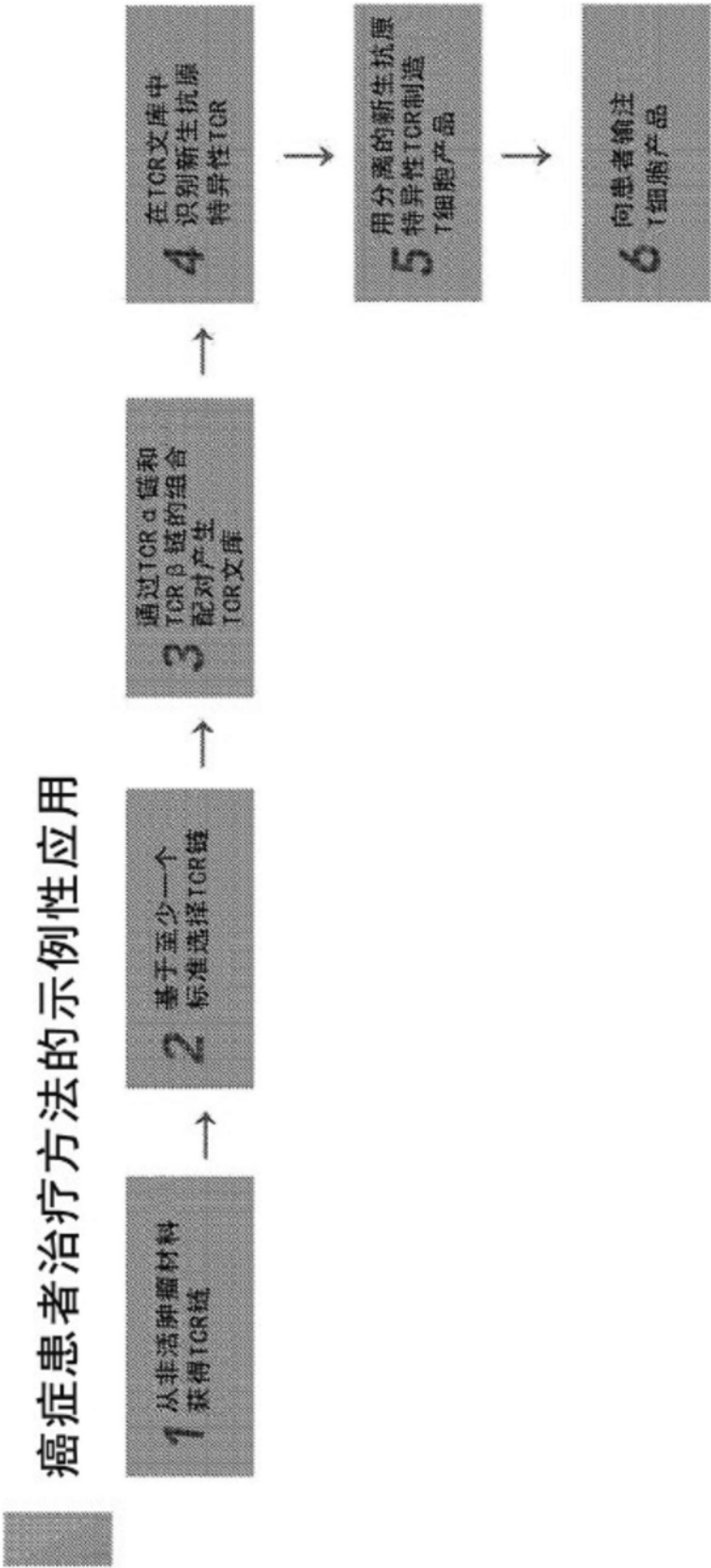


图5



表示为TCR β -P2A-TCR α -T2A-嘌呤霉素抗性
表达盒的TCR基因的示意性实例

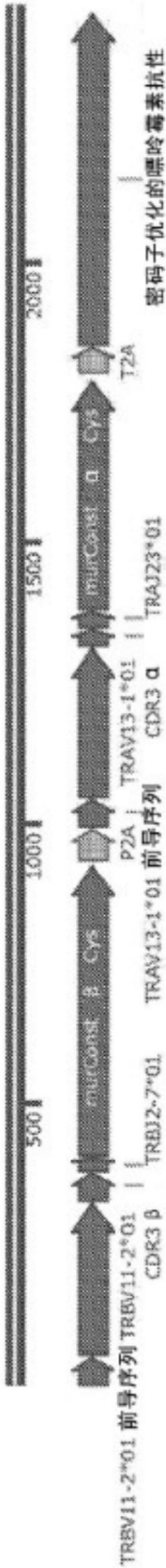
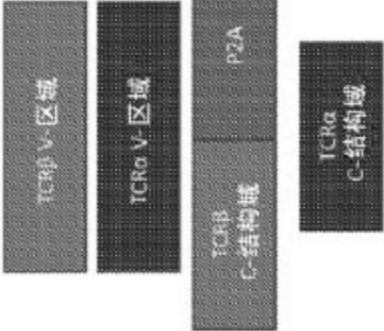


图6

TCR表达盒的示例性策略设计



“现成的” 片段



合成的每个TCR

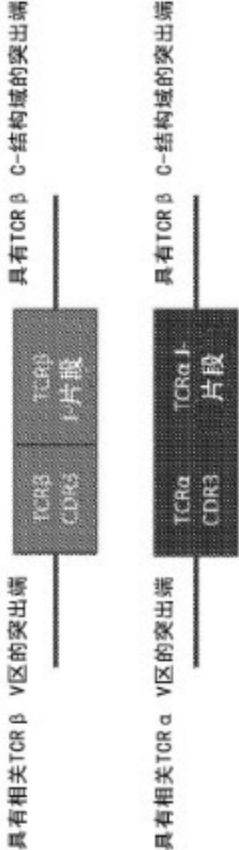


图7

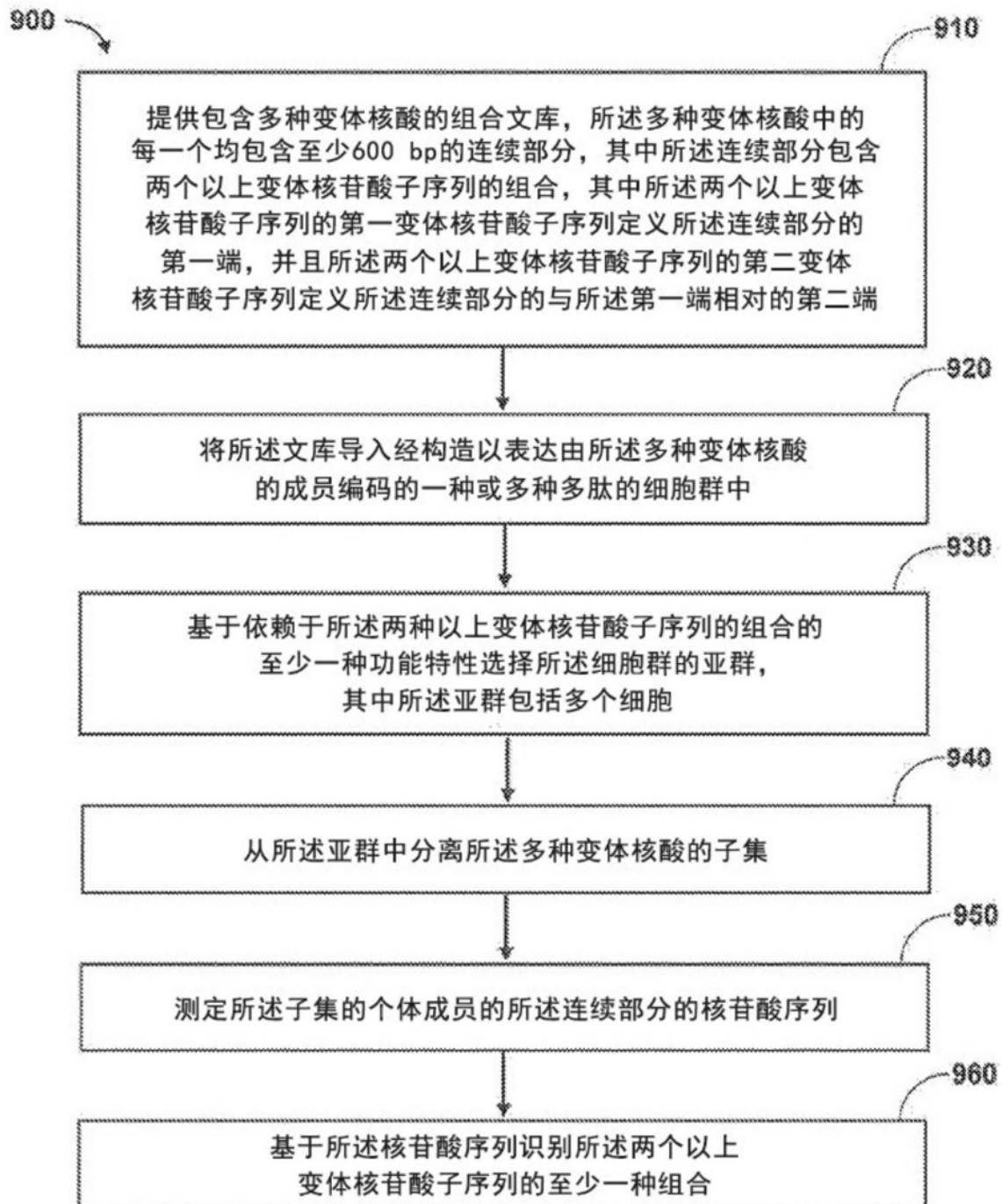


图9

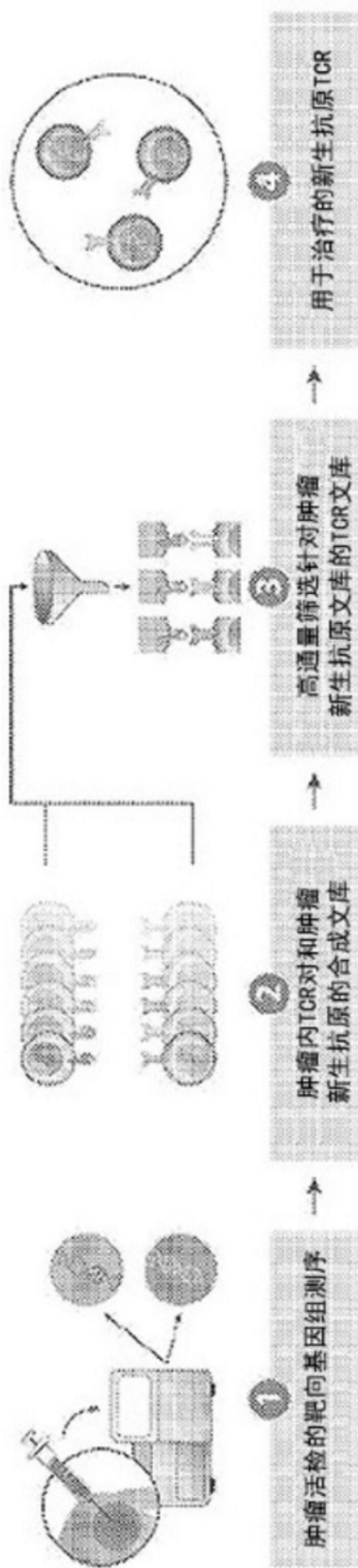


图10A

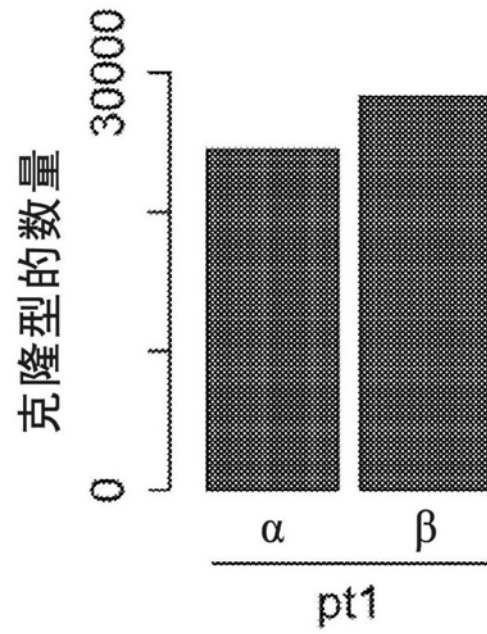


图10B

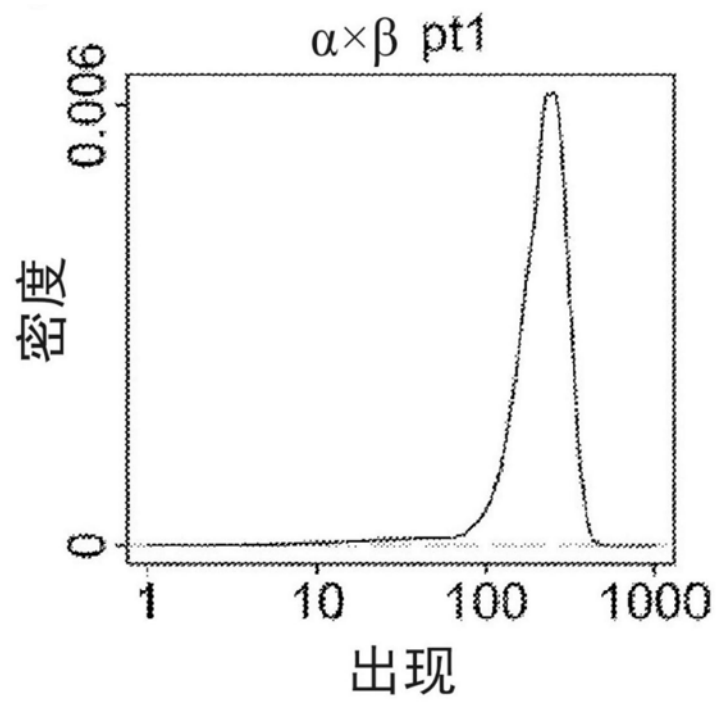


图10C

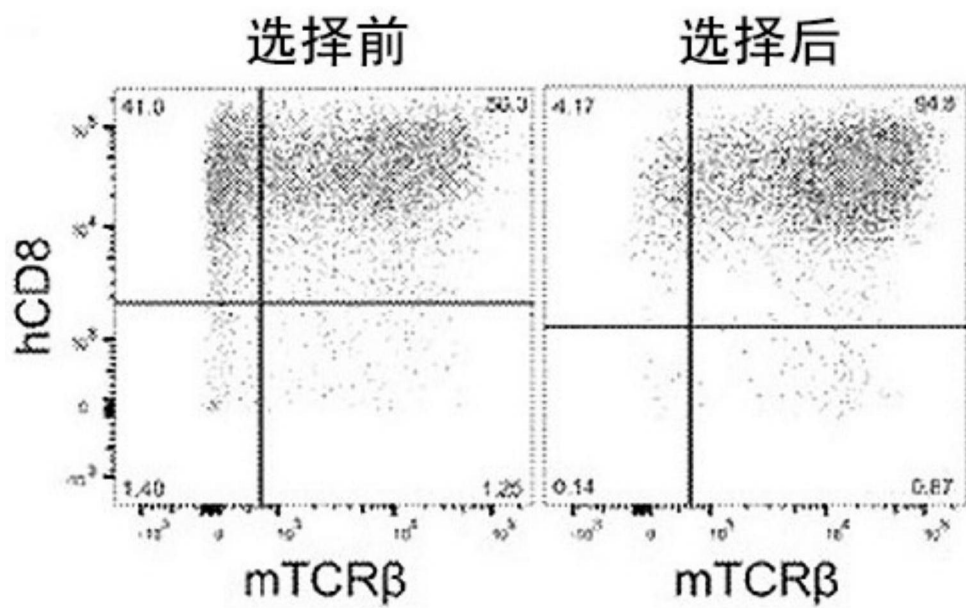


图10D

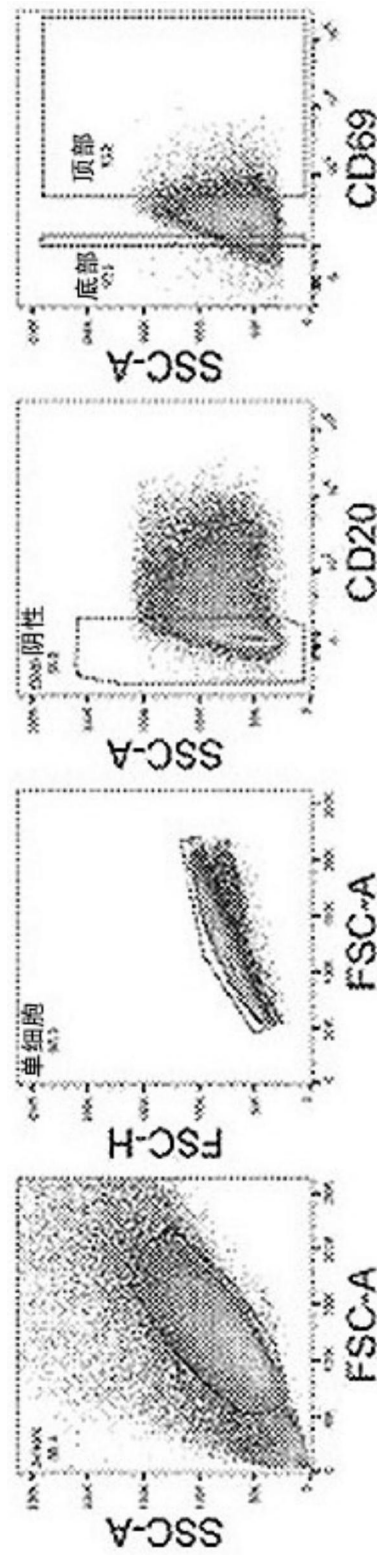


图10E

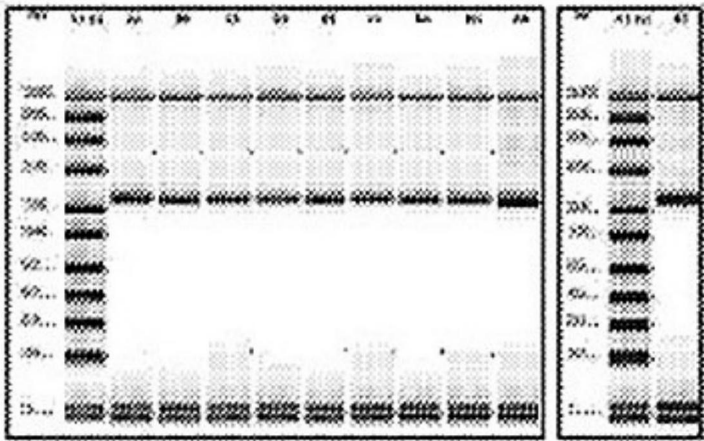


图10F

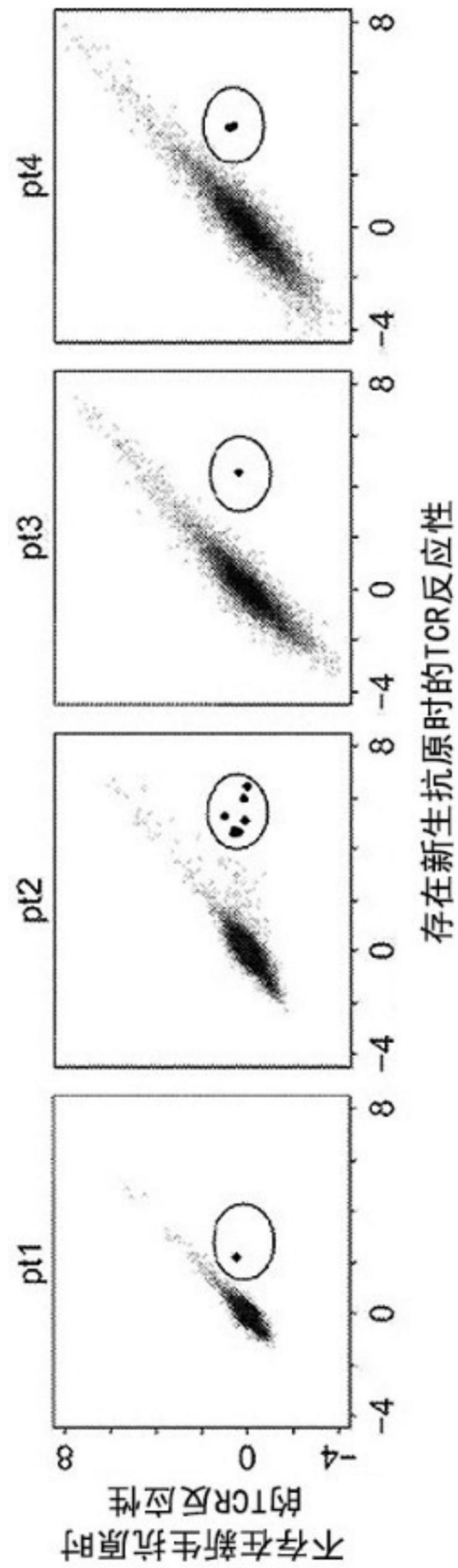


图10G

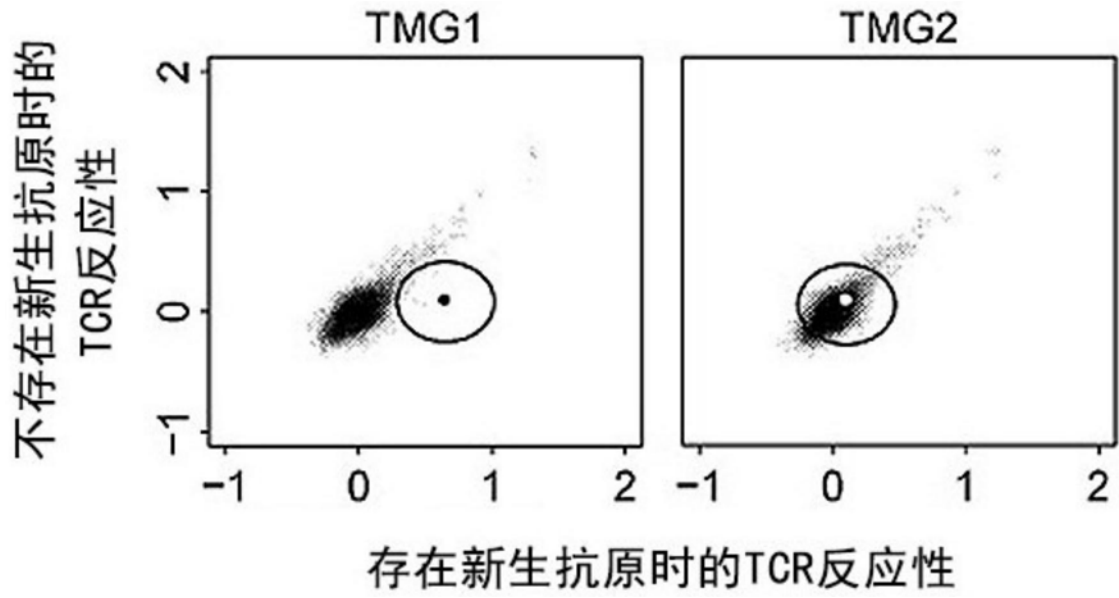


图10H

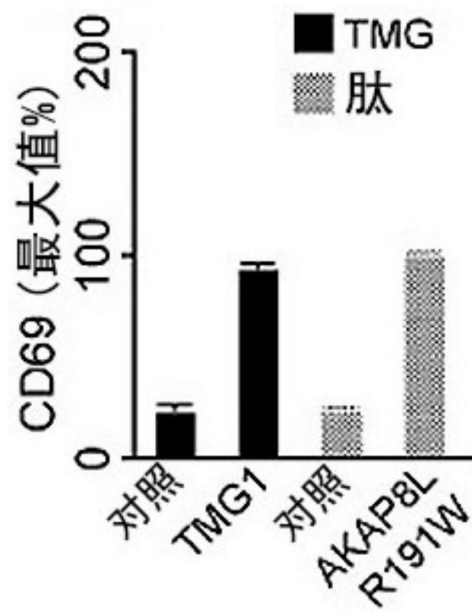


图10I

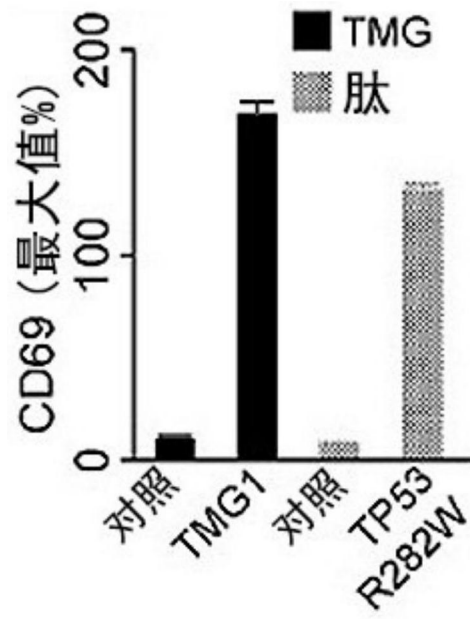


图10J

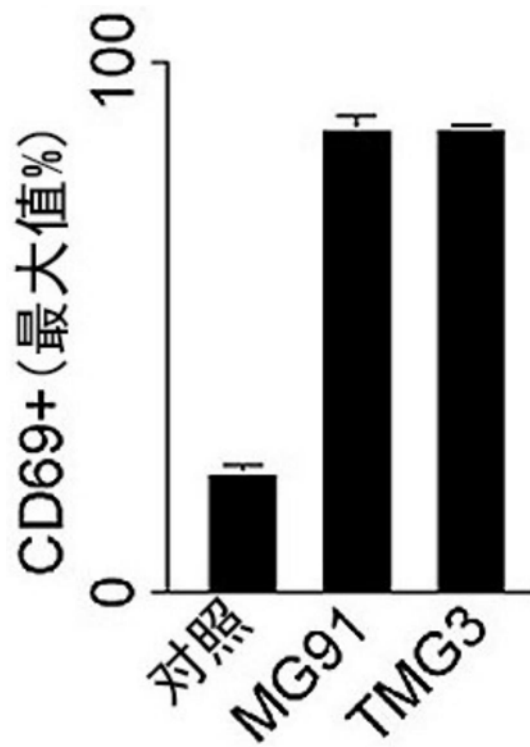


图10K

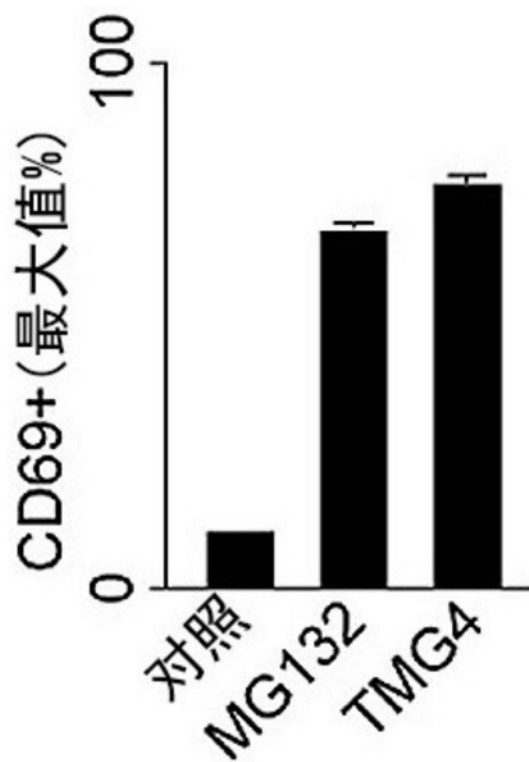


图10L

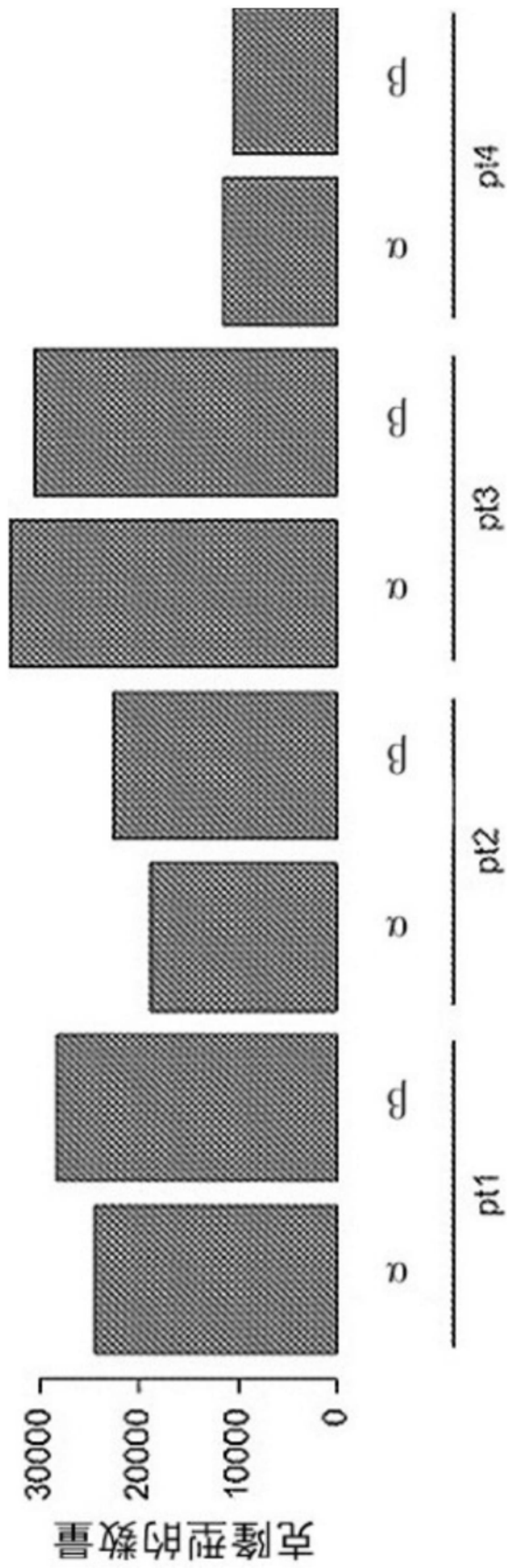


图11A



图11B

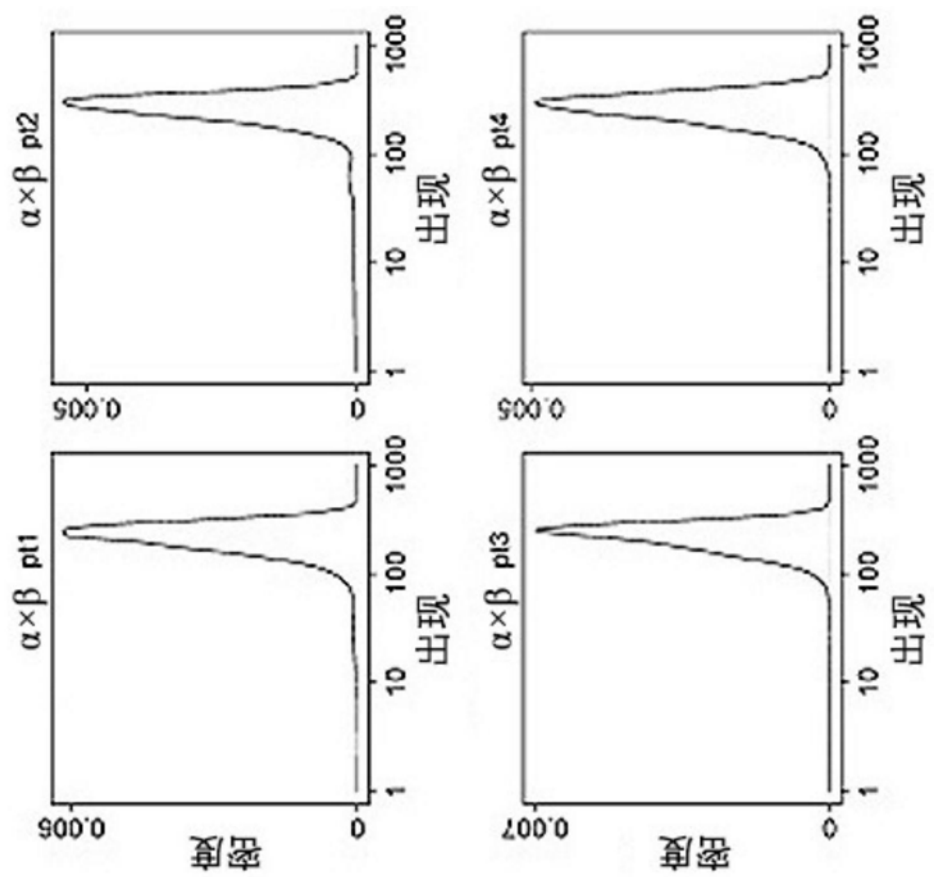


图11C

样本	读段的数量(范围)	平均覆盖率	落在中位数+/-2log单位范围内的TCR的百分比
pt1	14-498	254	97.2%
pt2	5-592	291	97.7%
pt3	49-495	240	97.8%
pt4	60-742	297	97.6%

图11D

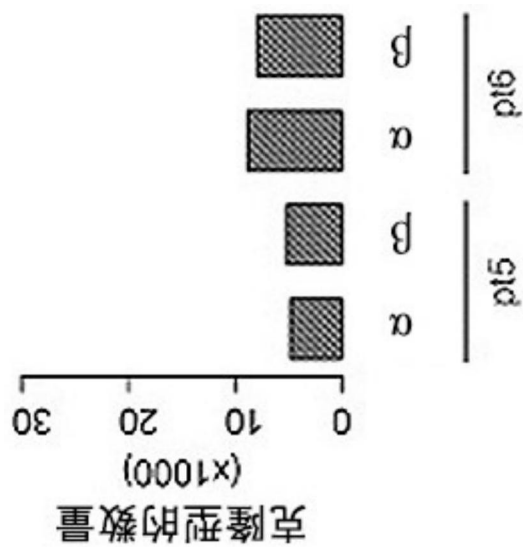


图12A

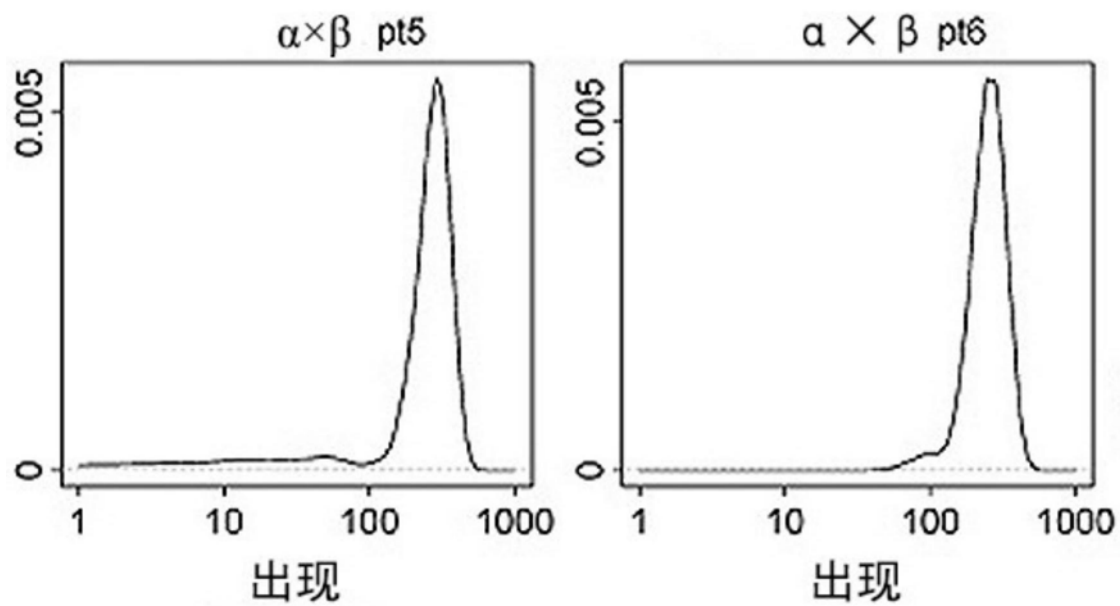


图12B

样本	读段的数量(范围)	平均覆盖率	落在中位数 $\pm 2 \log$ 单位范围内的TCR的百分比
pt5	3-611	299	97.6%
pt6	48-569	276	98.2%

图12C

TCR质粒	稀释度
CMV1	10.000
CMV2	100.000
CDK4-8	10.000
CDK4-17	100.000
GCN1L1	100.000
NY ESO 1G4	100.000

图13A

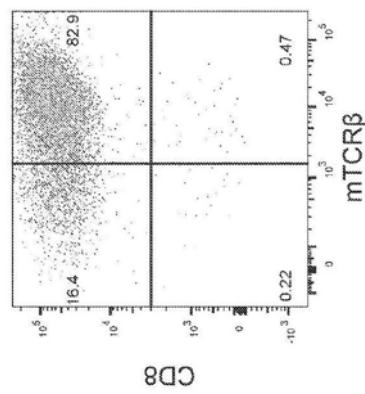


图13B

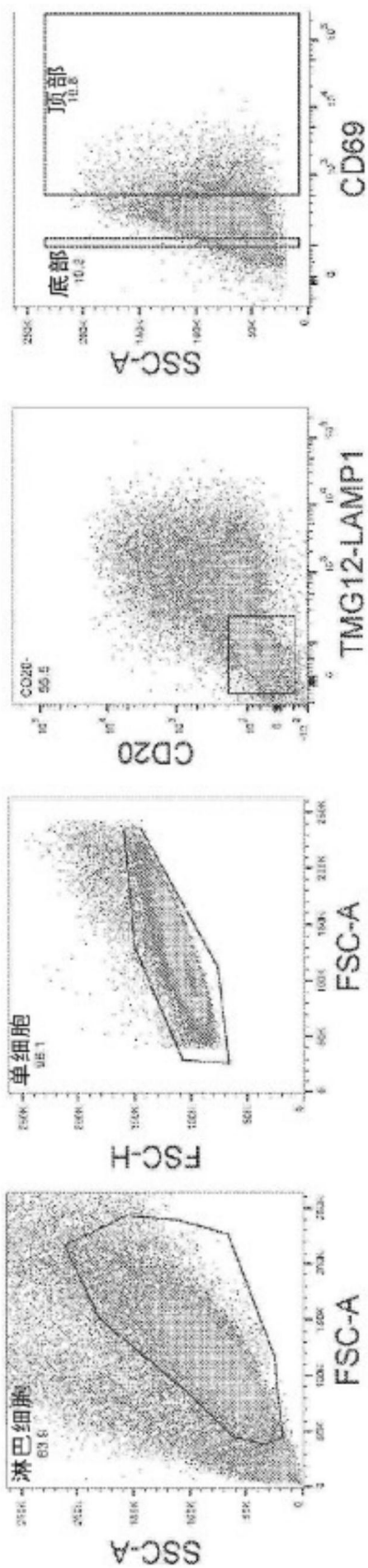


图13C

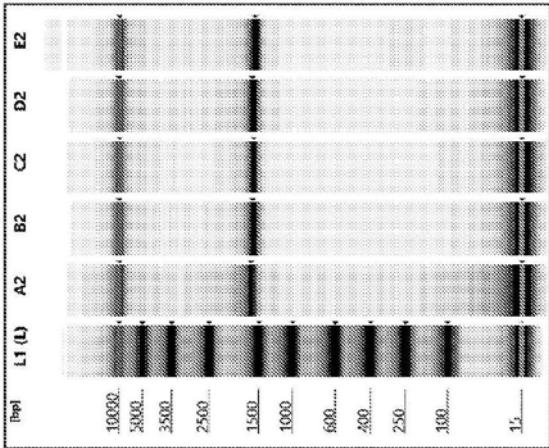


图13D

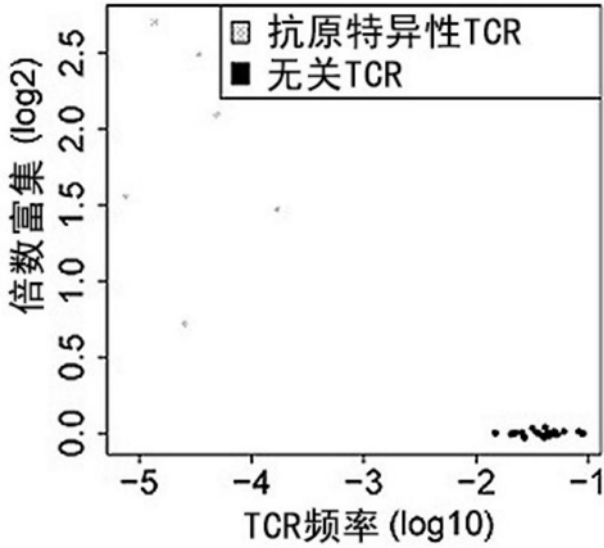


图13E

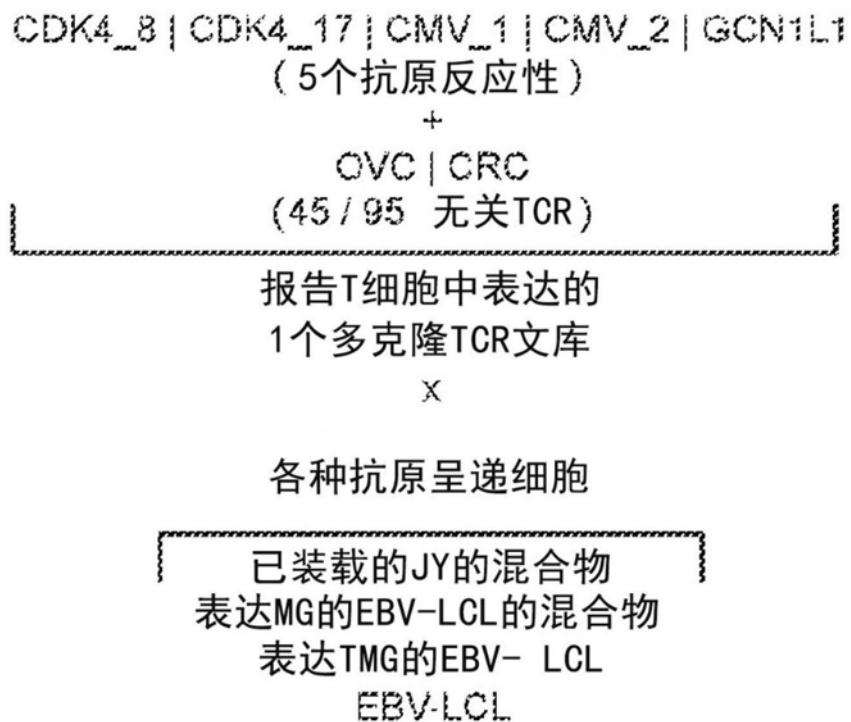


图14A

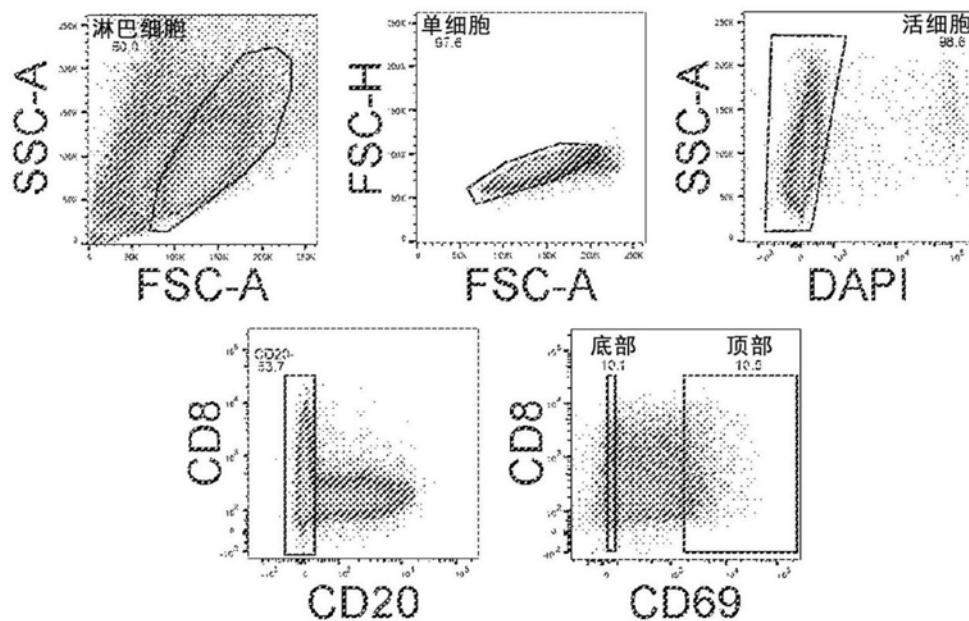


图14B

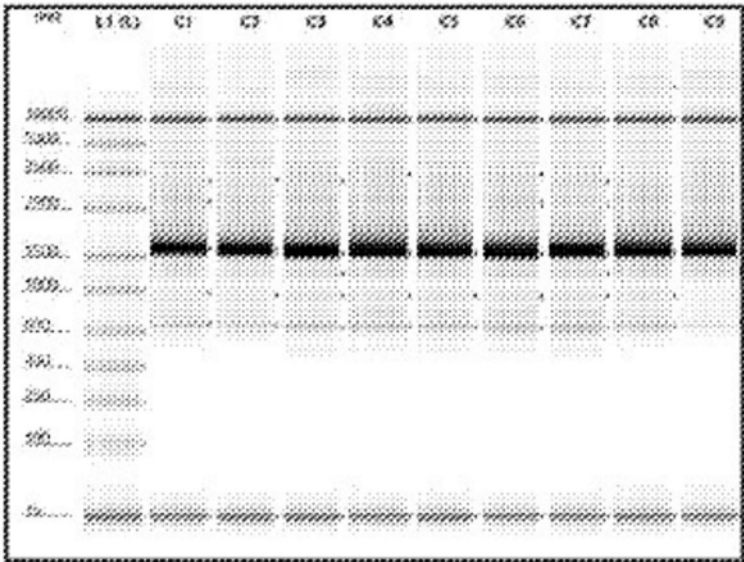


图14C

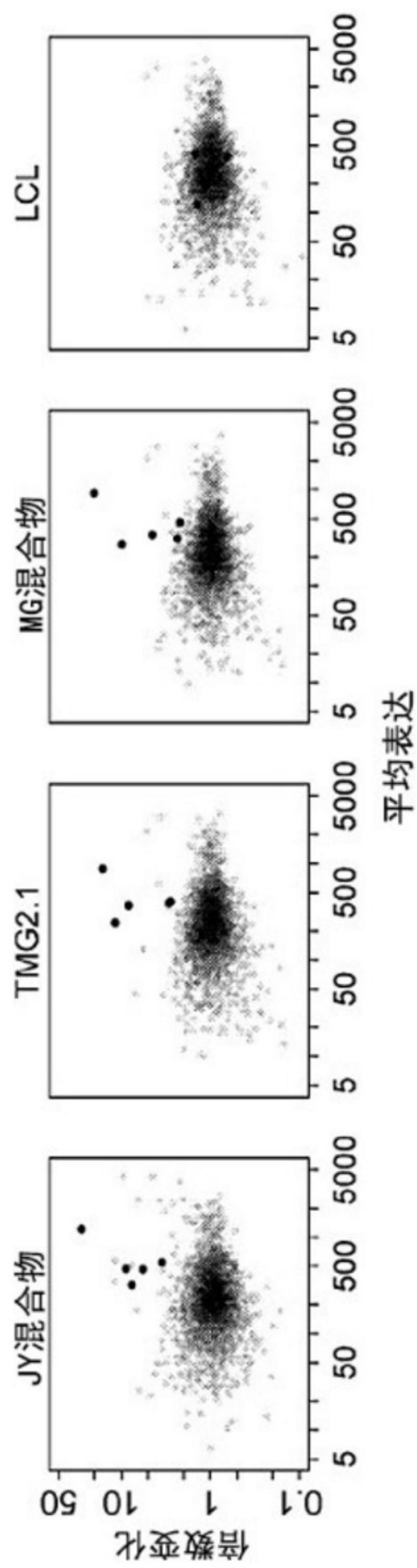


图14D

排序	α	β	碱基平均值	Log2 FC	调整后的p值
1	CDK4_17	CDK4_17	872	3.81	4.3E-42
2	CDK4_8	CDK4_8	244	2.75	1.2E-10
3	GCN1L1	GCN1L1	346	2.57	6.8E-10
4	CMV_2	CMV_2	449	2.11	7.0E-09
5	CMV_1	CMV_1	362	1.89	3.4E-01
6	CMV_1	GCN1L1	154	1.19	4.1E-01
7	OVC48.1	OVC21.2	97	1.55	8.1E-01
8	OVC21.2	CDK4_8	135	1.22	9.0E-01
9	CDK4_8	OVC21.2	52	1.63	9.6E-01
10	CMV_2	OVC48.1	99	1.34	1.0E+00

图14E

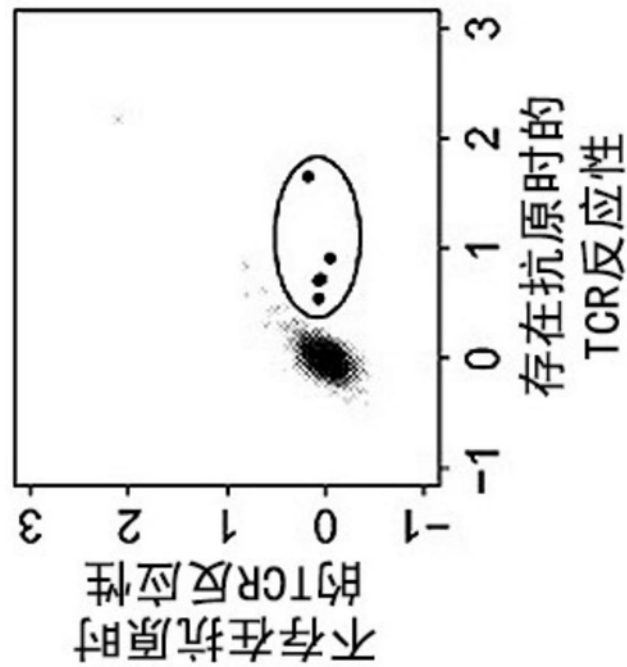


图14F

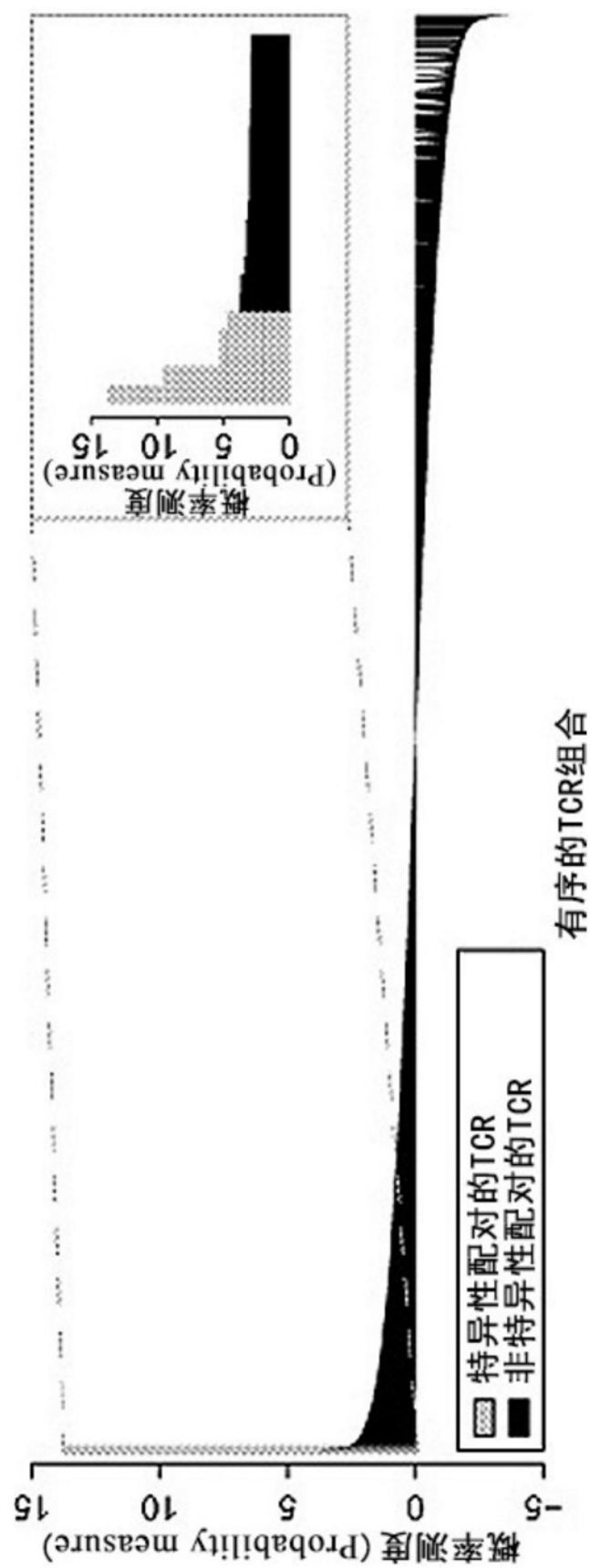


图14G

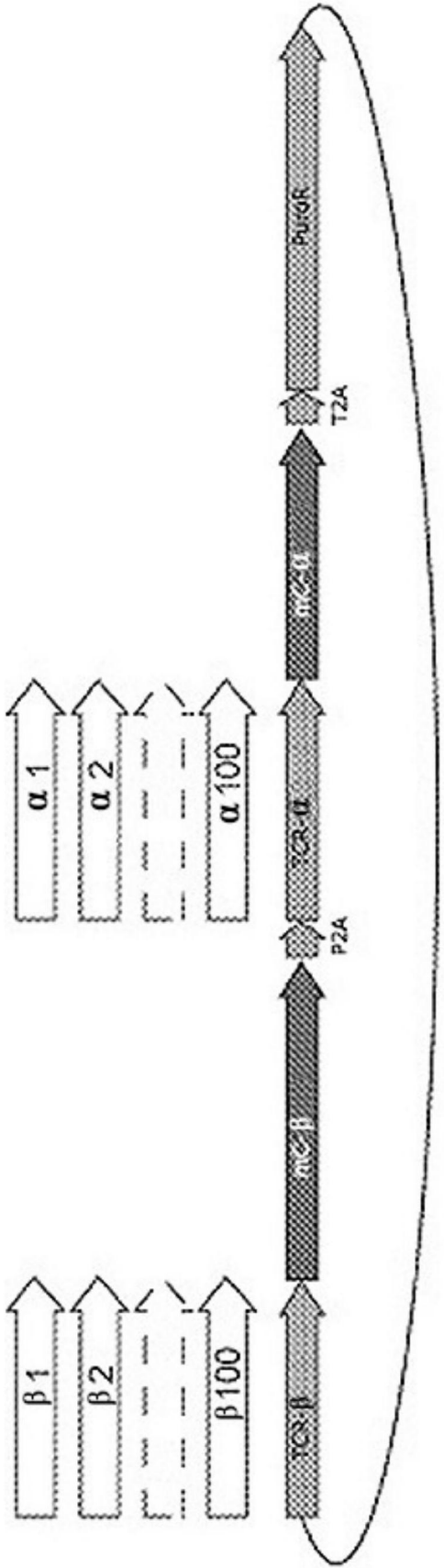


图15A

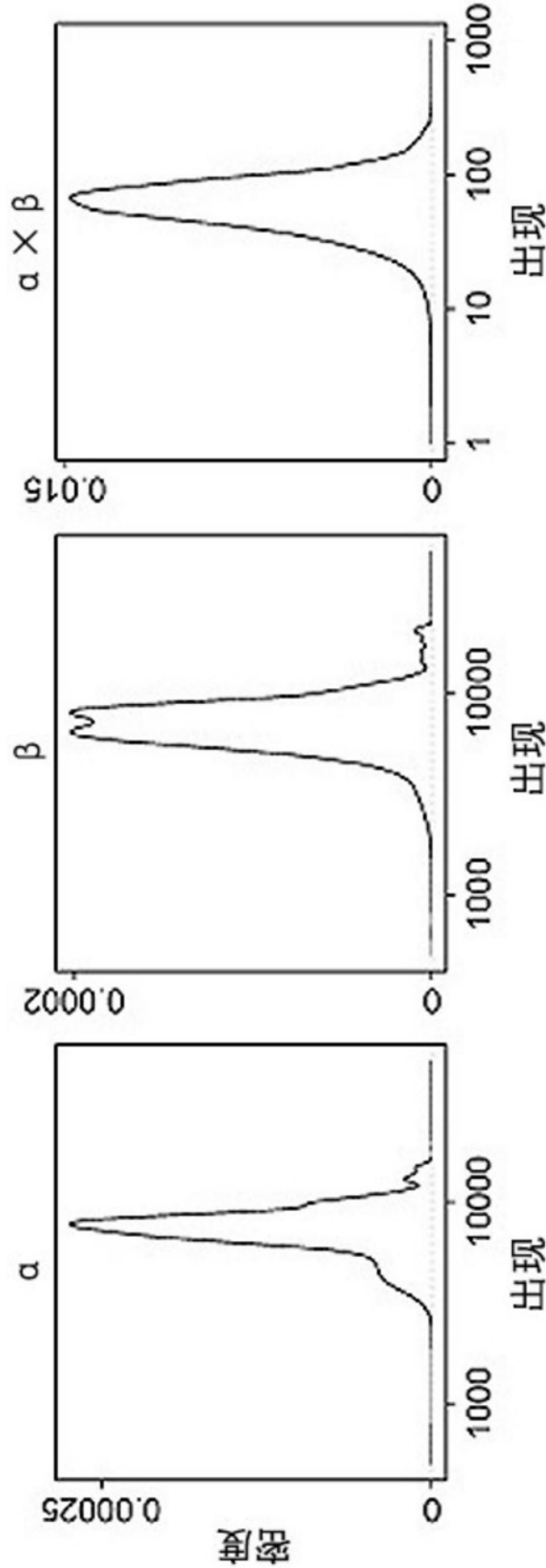


图15B

200 × 200 文库 $\approx$ 4 × (100 × 100 文库)

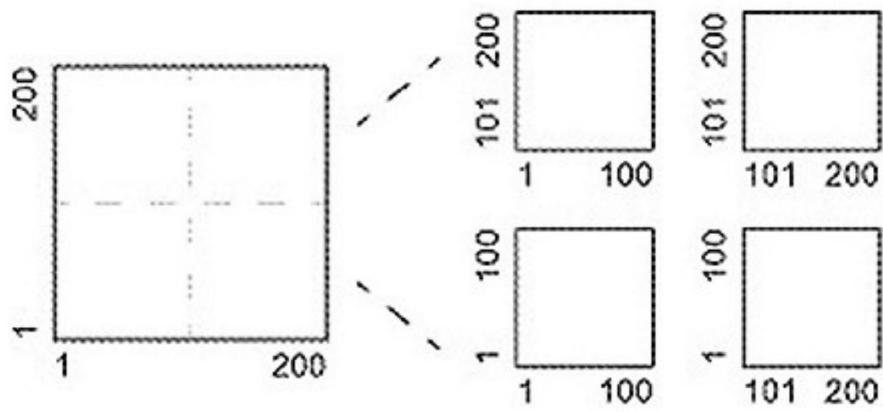


图15C

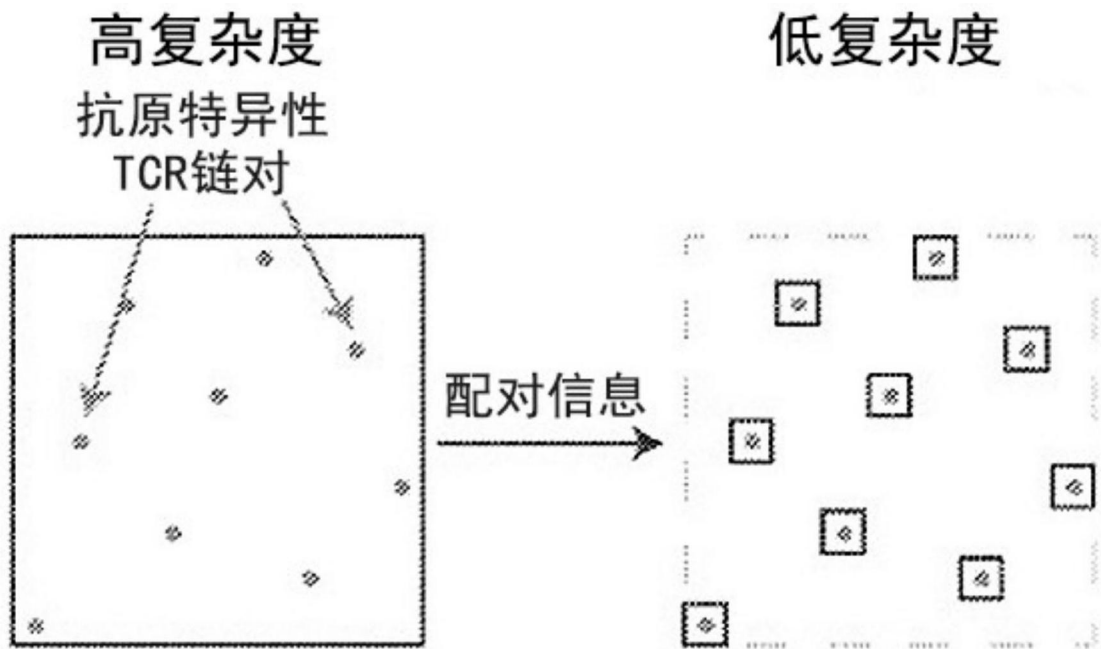


图15D

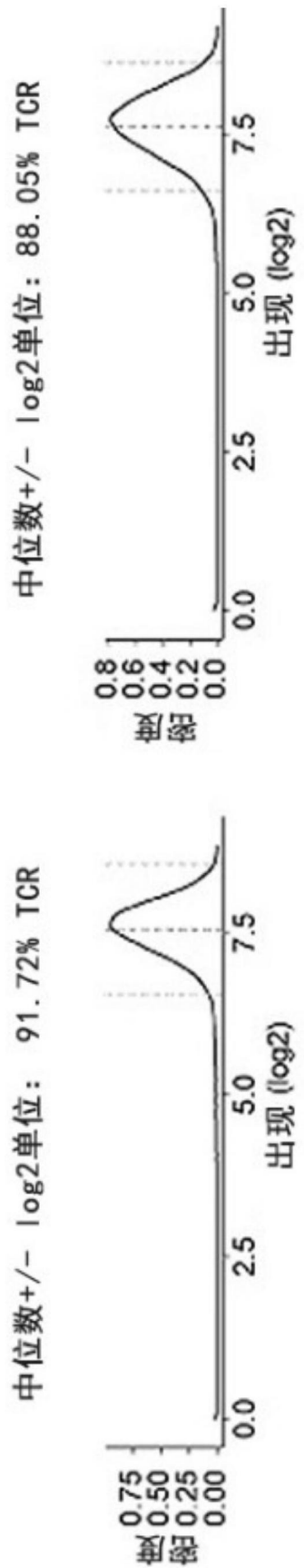


图15E

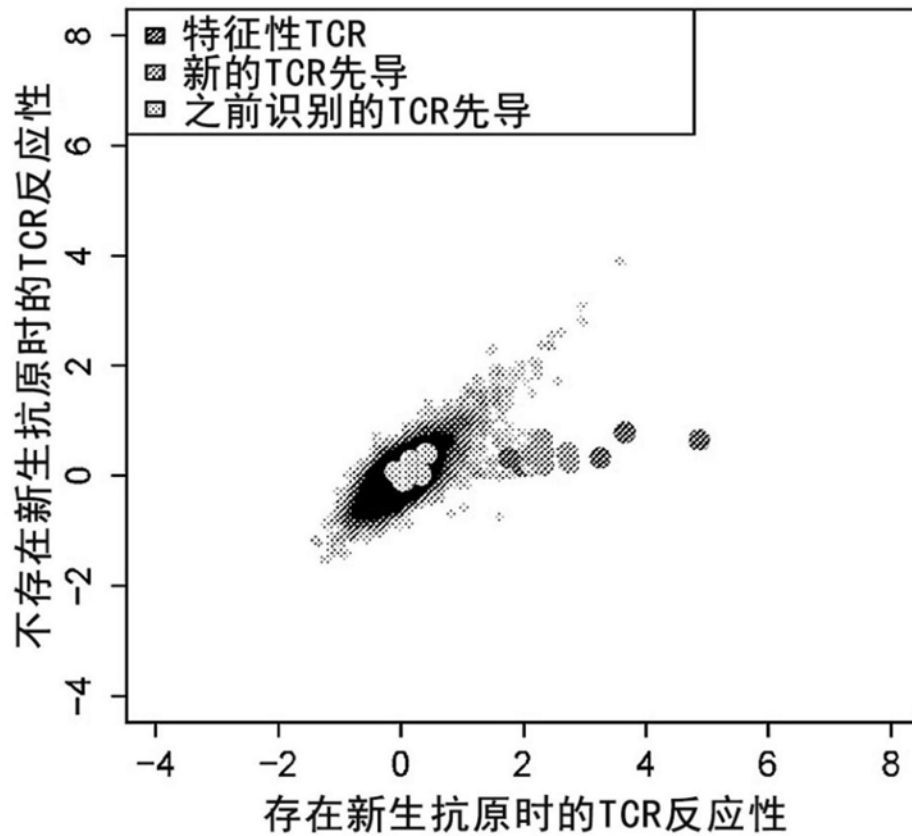


图15F

α	β	200 x 200				100 x 100			
		碱基平均值	log2FC	bonferroni	排序	碱基平均值	log2FC	bonferroni	
NYESO_1g4	NYESO_1g4	895.756	3.559	2.83E-87	1				
GCN1L1	GCN1L1	169.73	3.512	2.19E-17	4				
CDK4_17	CDK4_17	123.15	3.255	2.96E-12	7				
CMV_2	CMV_2	408.908	1.505	1.71E-08	10				
CMV_1	CMV_1	276.288	1.867	1.60E-07	12				
TCR24	TCR43	155.563	1.845	3.88E-04	22	721.624	2.929	2.60E-123	
CDK4_8	CDK4_8	160.185	1.309	1.00E+00	41				
TCR81	TCR97	27.535	0.561	1.00E+00	9693	3452.72	2.306	3.41E-243	
TCR83	TCR29	8.754	-0.738	1.00E+00	14727	515.73	2.598	2.35E-63	
TCR88	TCR35	1.494	-1.485	1.00E+00	19661	455.633	2.368	1.05E-50	
TCR76	TCR20	49.302	0.208	1.00E+00	22192	357.861	3.768	3.27E-93	
TCR53	TCR17	53.18	-0.145	1.00E+00	26478	410.276	2.599	6.56E-64	
TCR24	TCR62	51.637	0.027	1.00E+00	37425	574.068	3.978	4.80E-148	

图15G

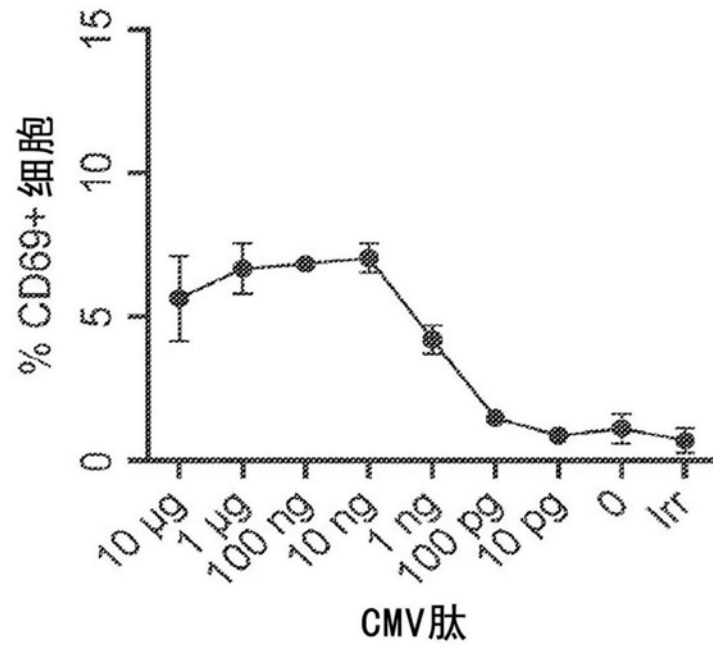


图16A

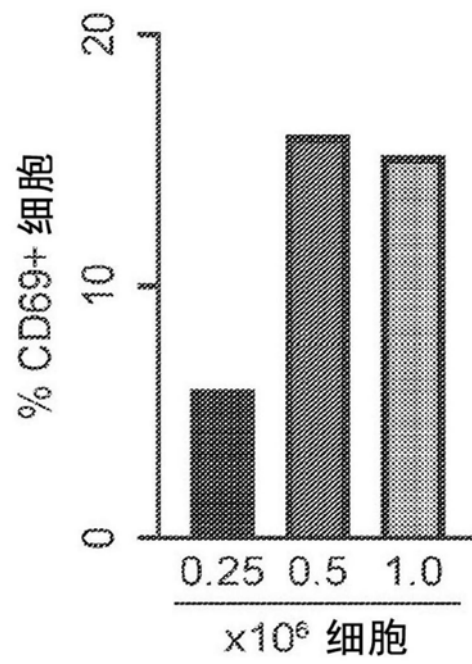


图16B

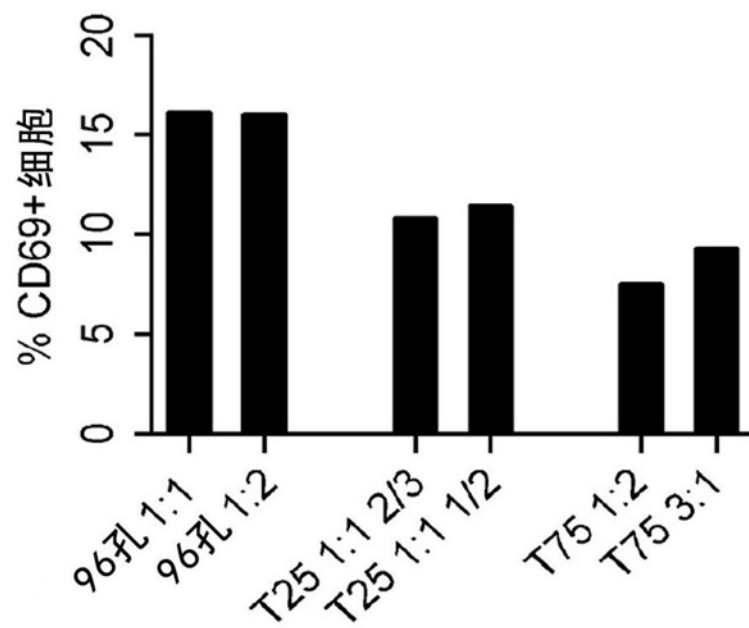


图16C

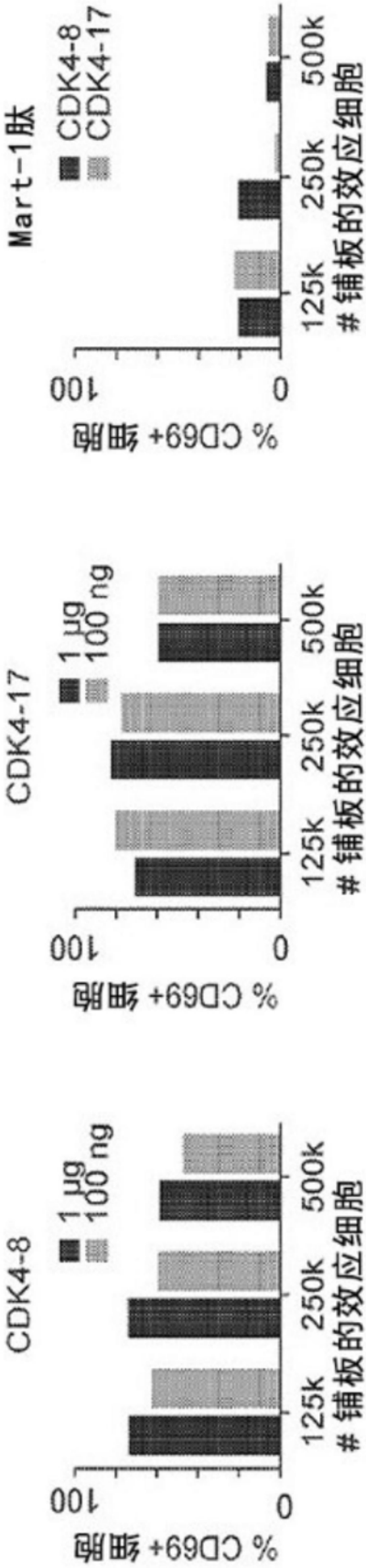


图16D

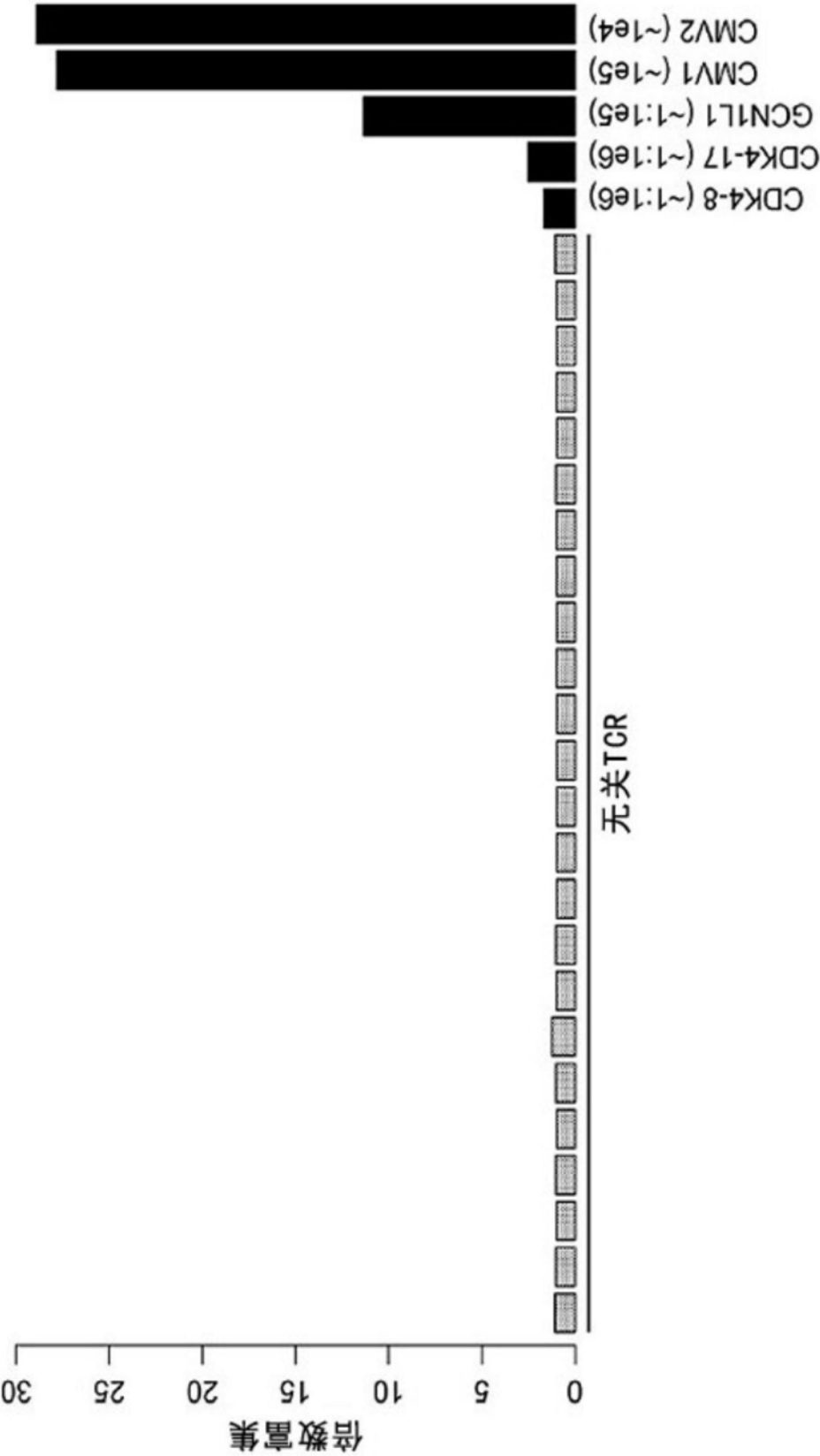


图16E

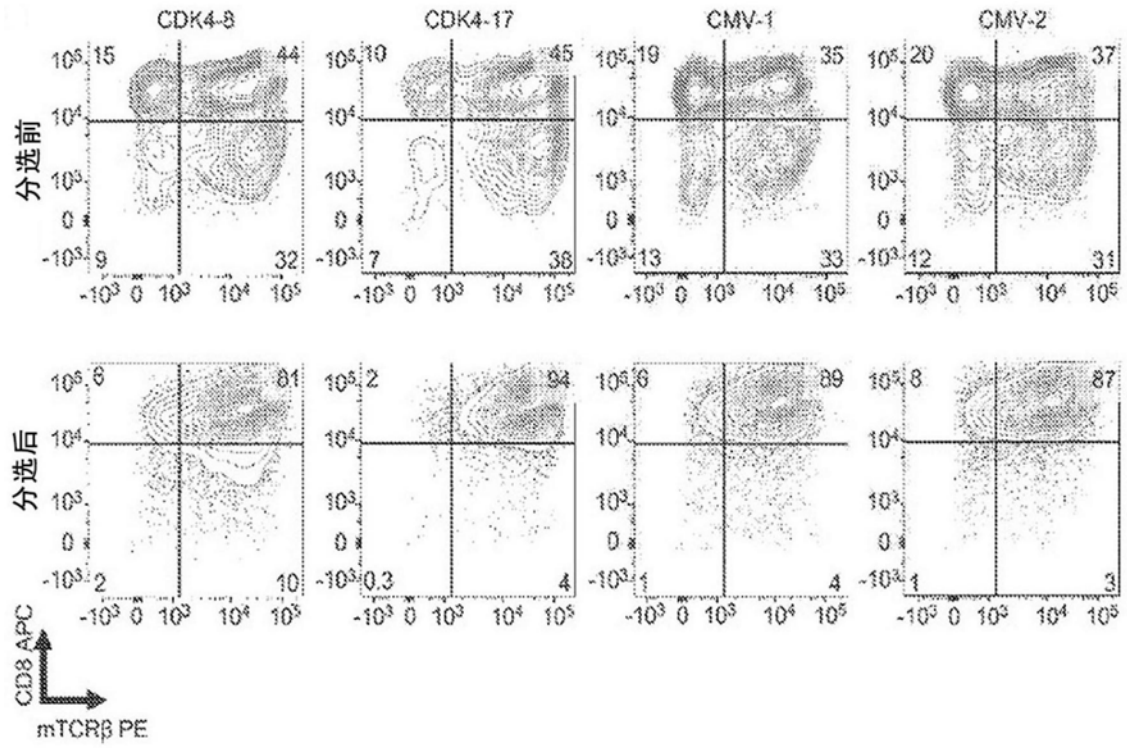


图17a

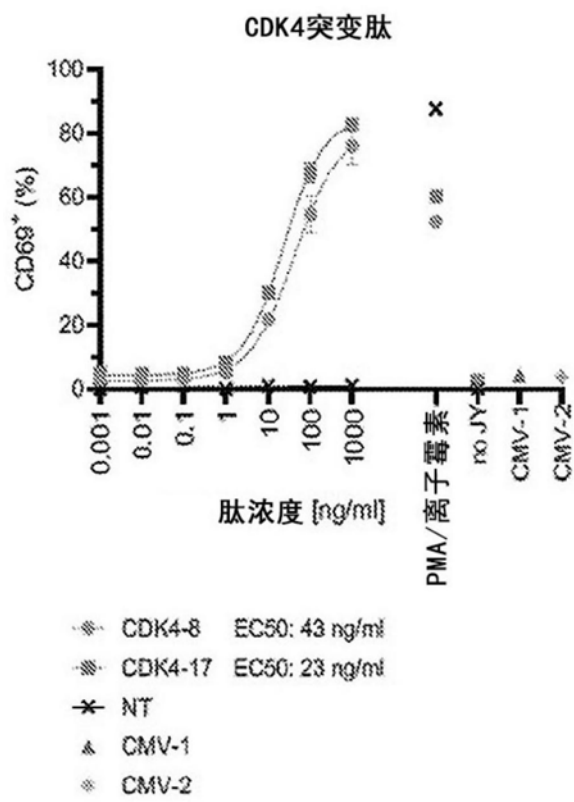


图17b

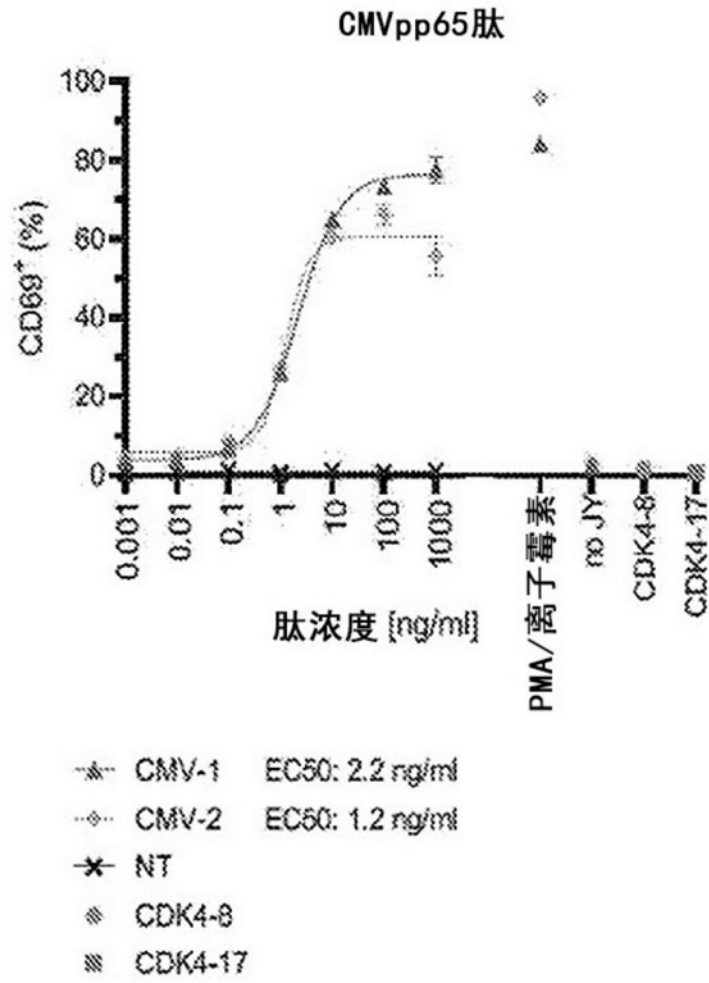


图17c

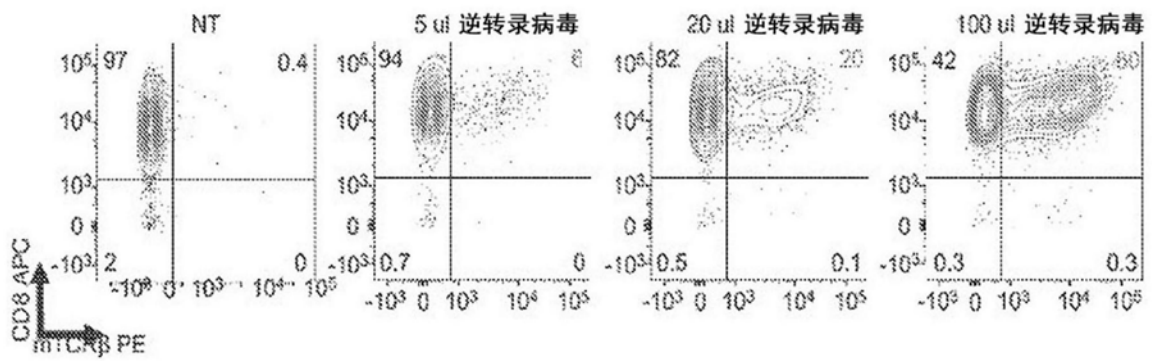


图18a

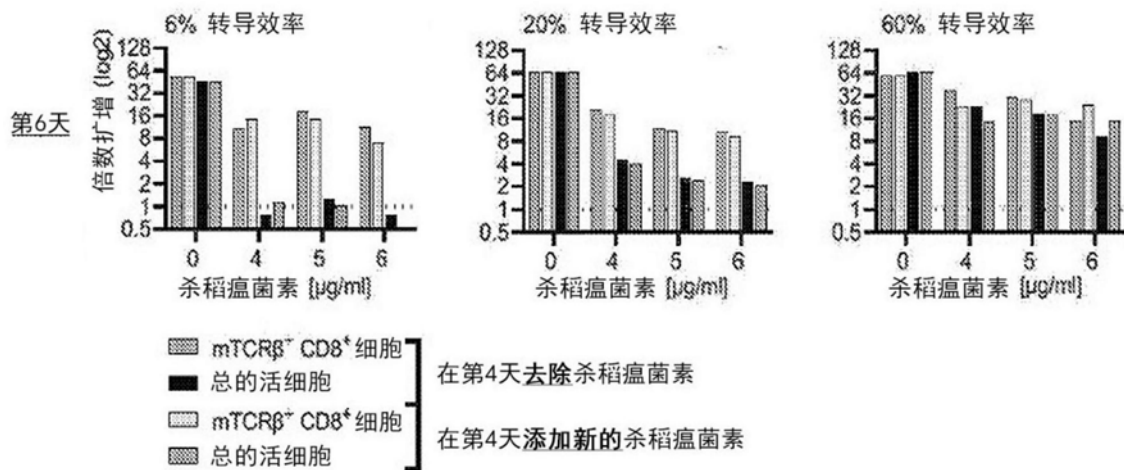


图18b

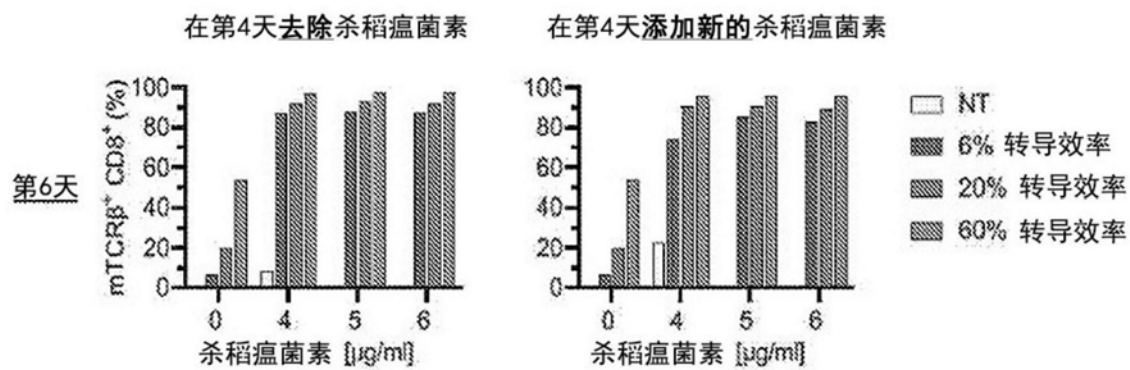


图18c

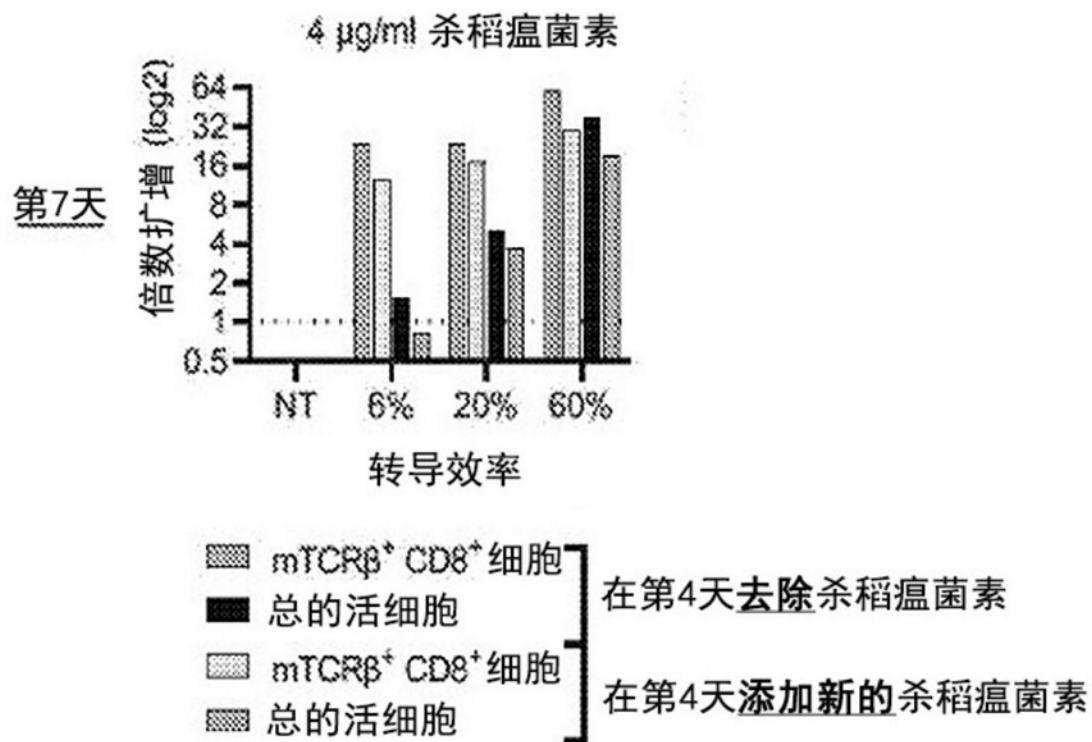


图18d

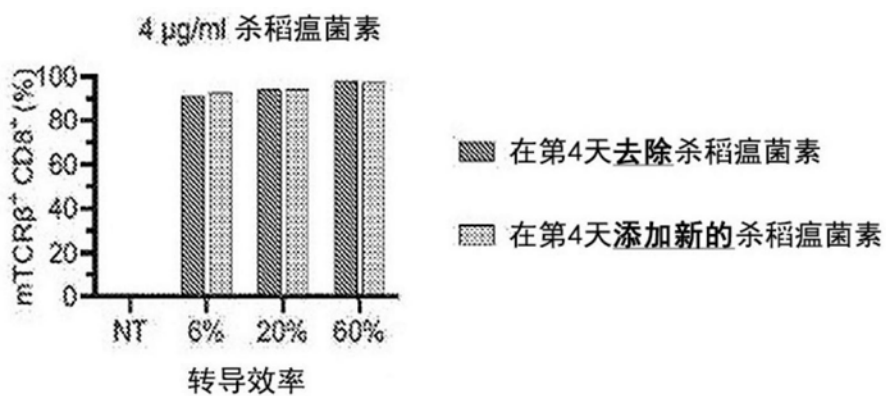


图18e

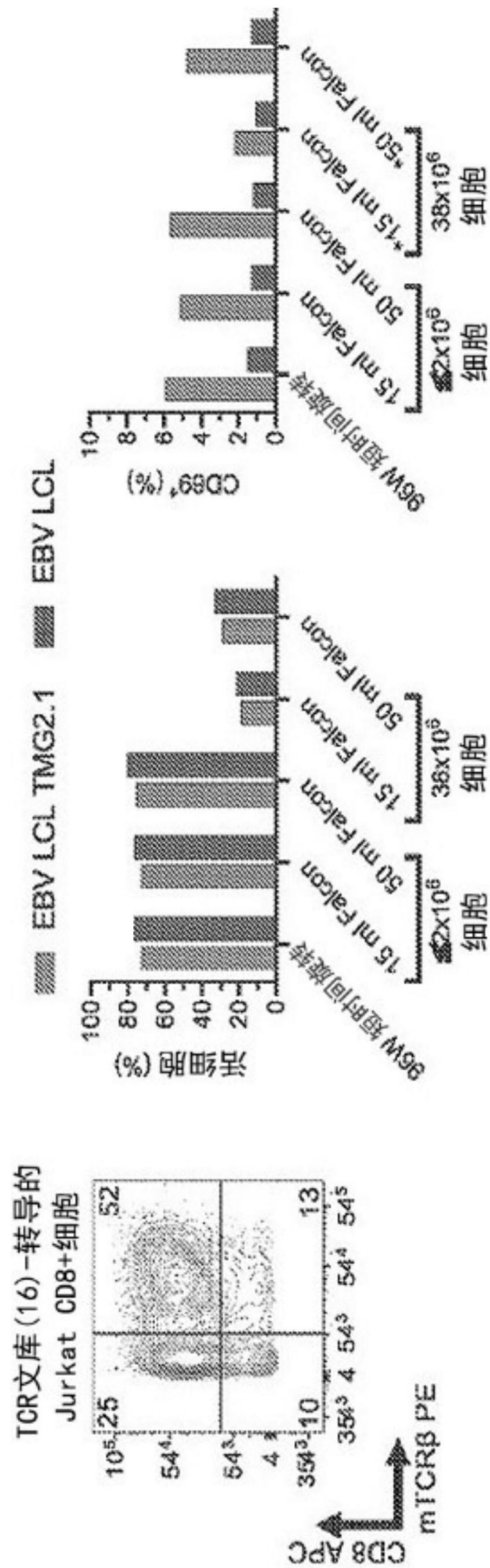


图19a

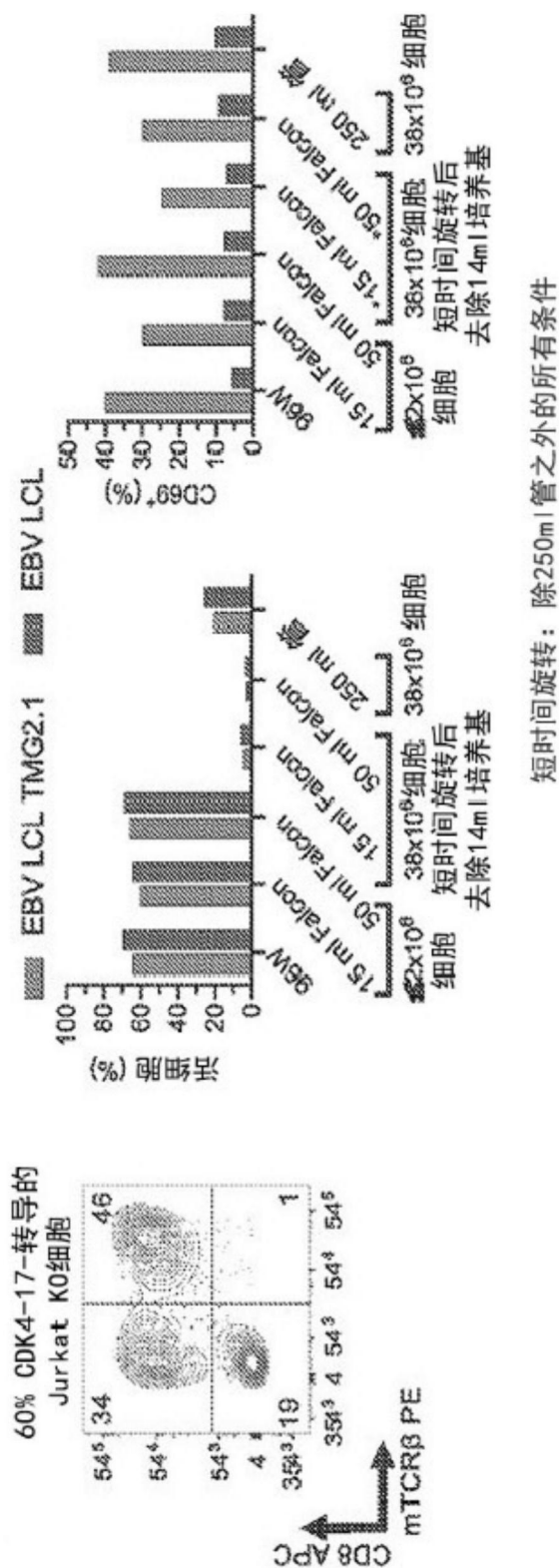


图 19b

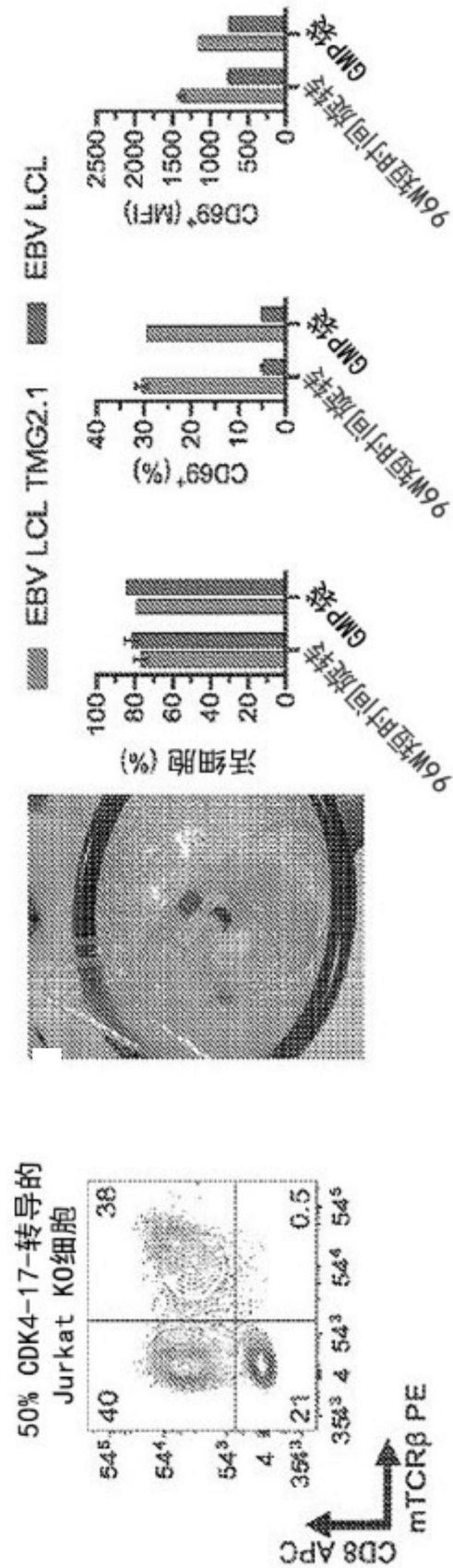


图19c

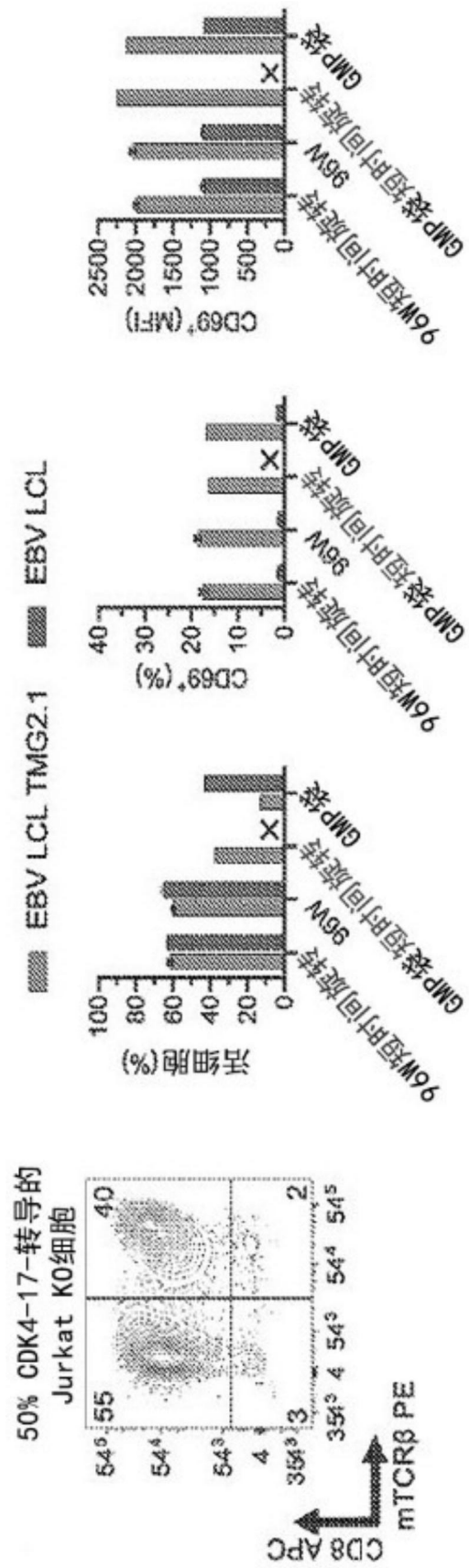


图19d

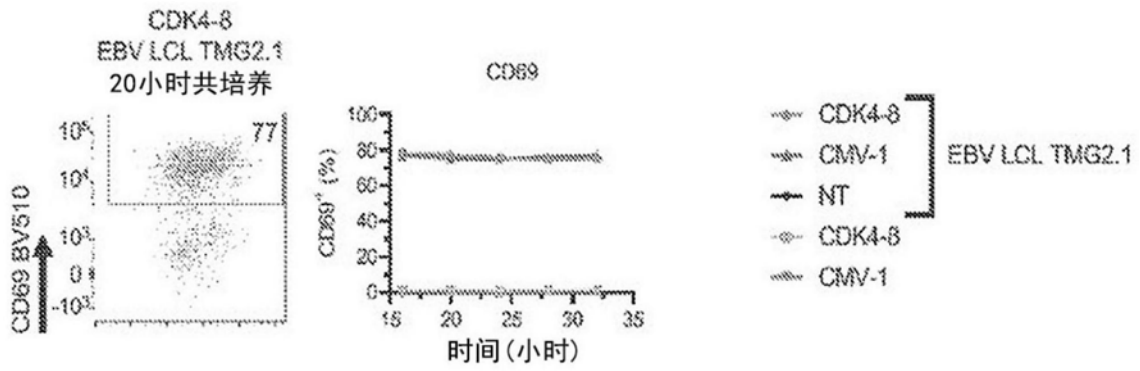


图20a

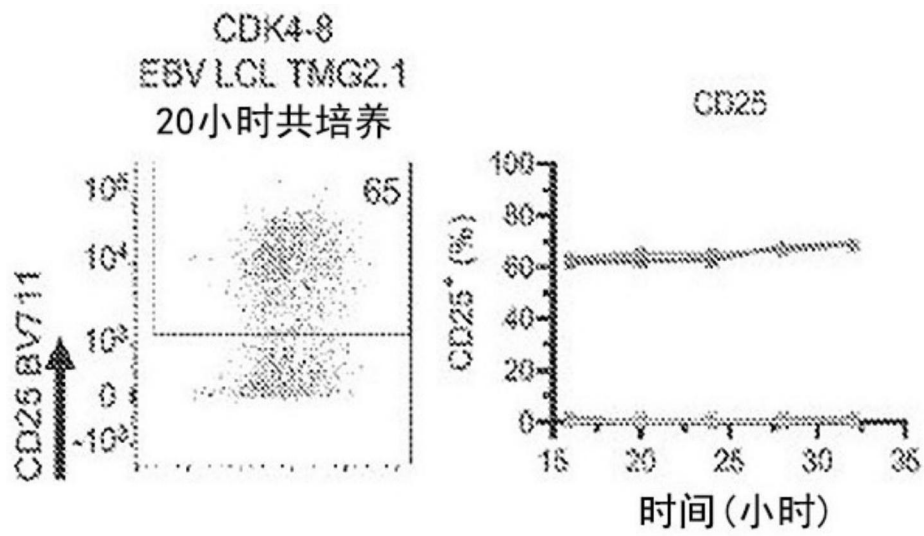


图20b

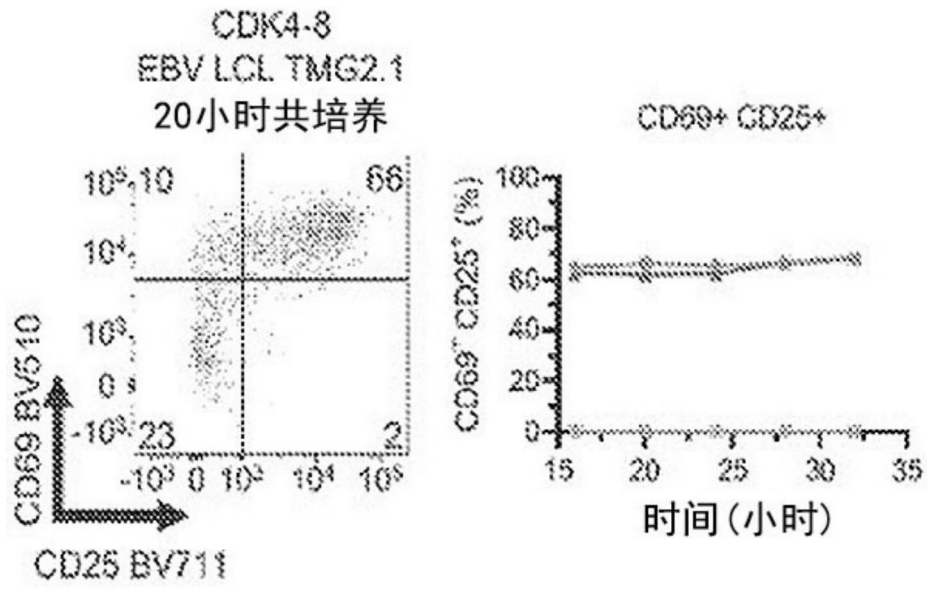


图20c

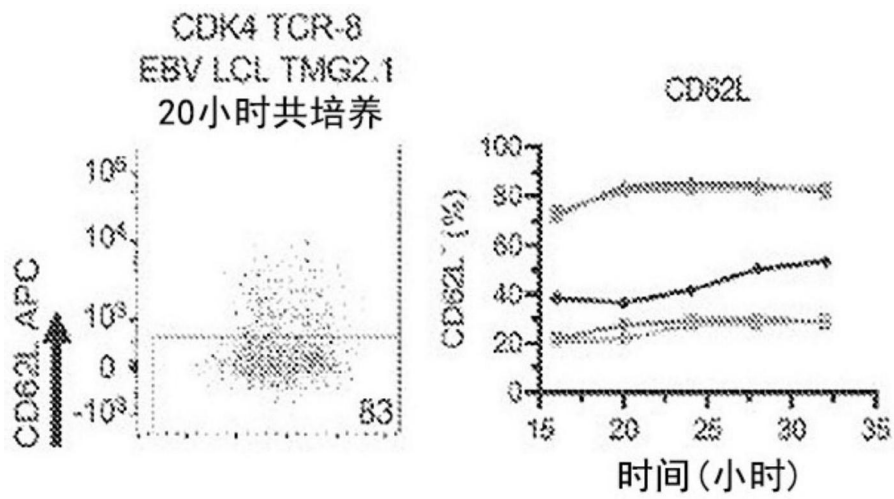


图20d

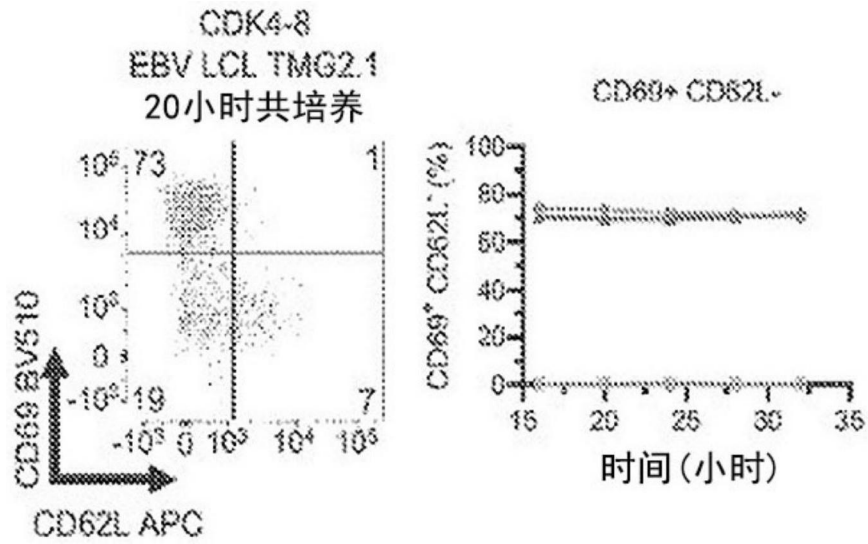


图20e

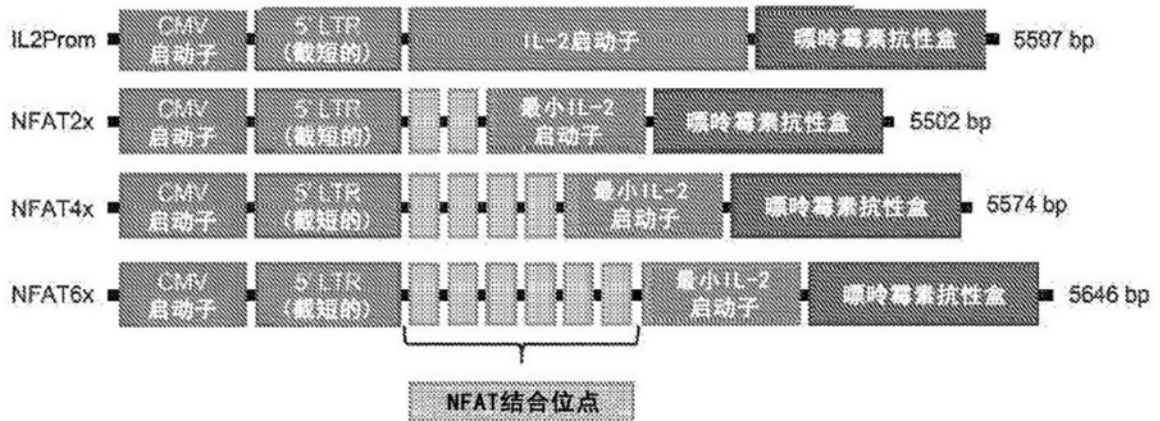


图21a

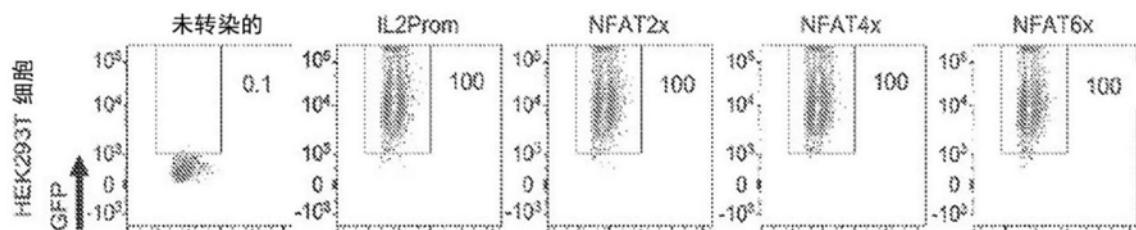


图21b

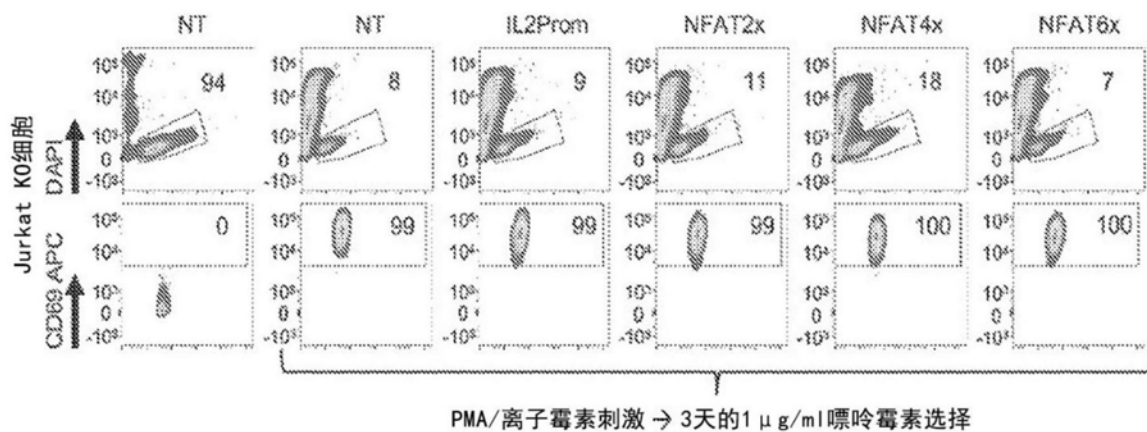


图21c

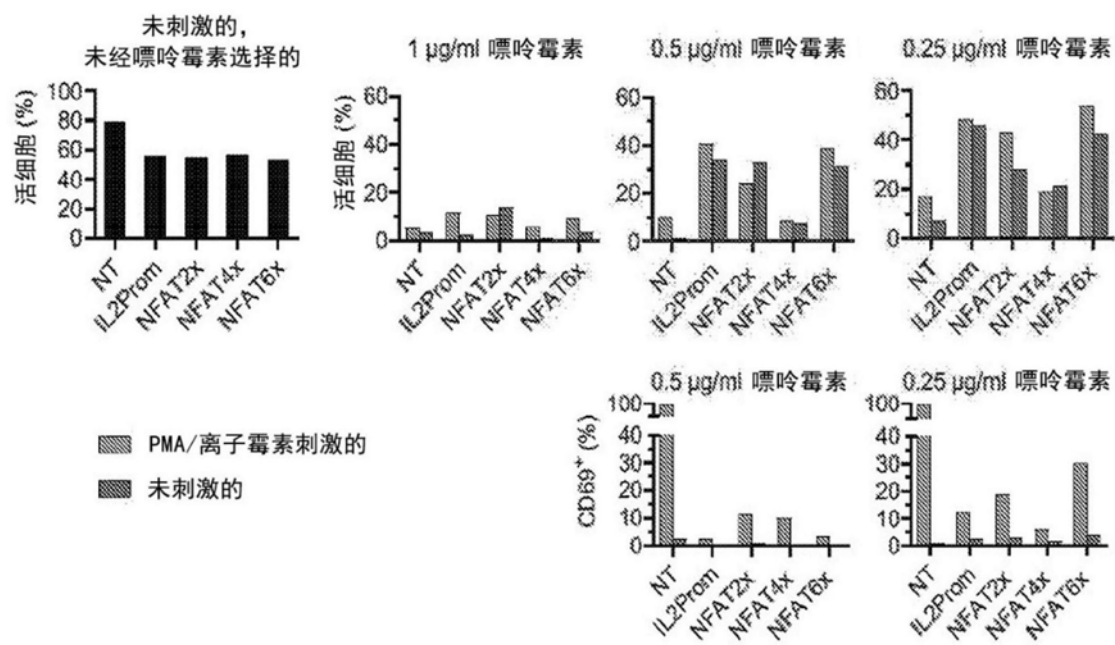


图21d

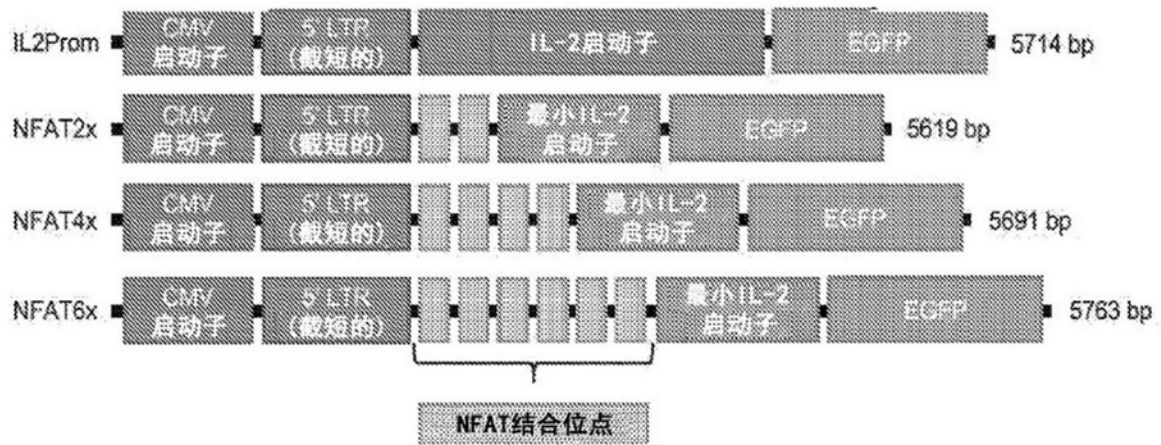


图22a

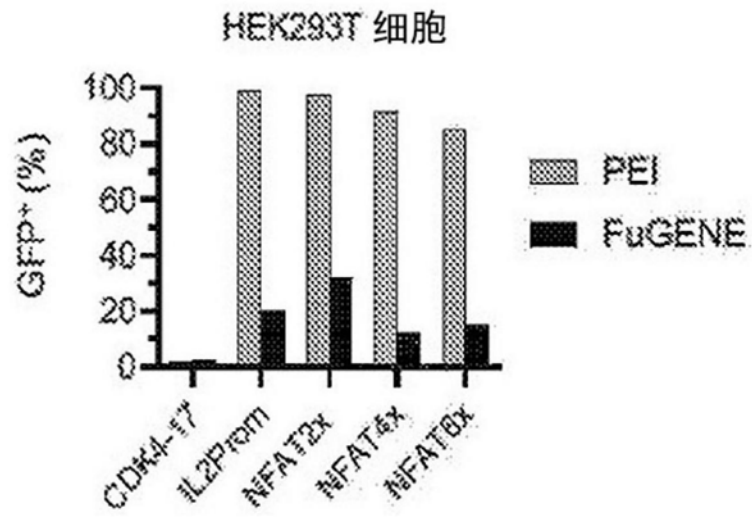


图22b

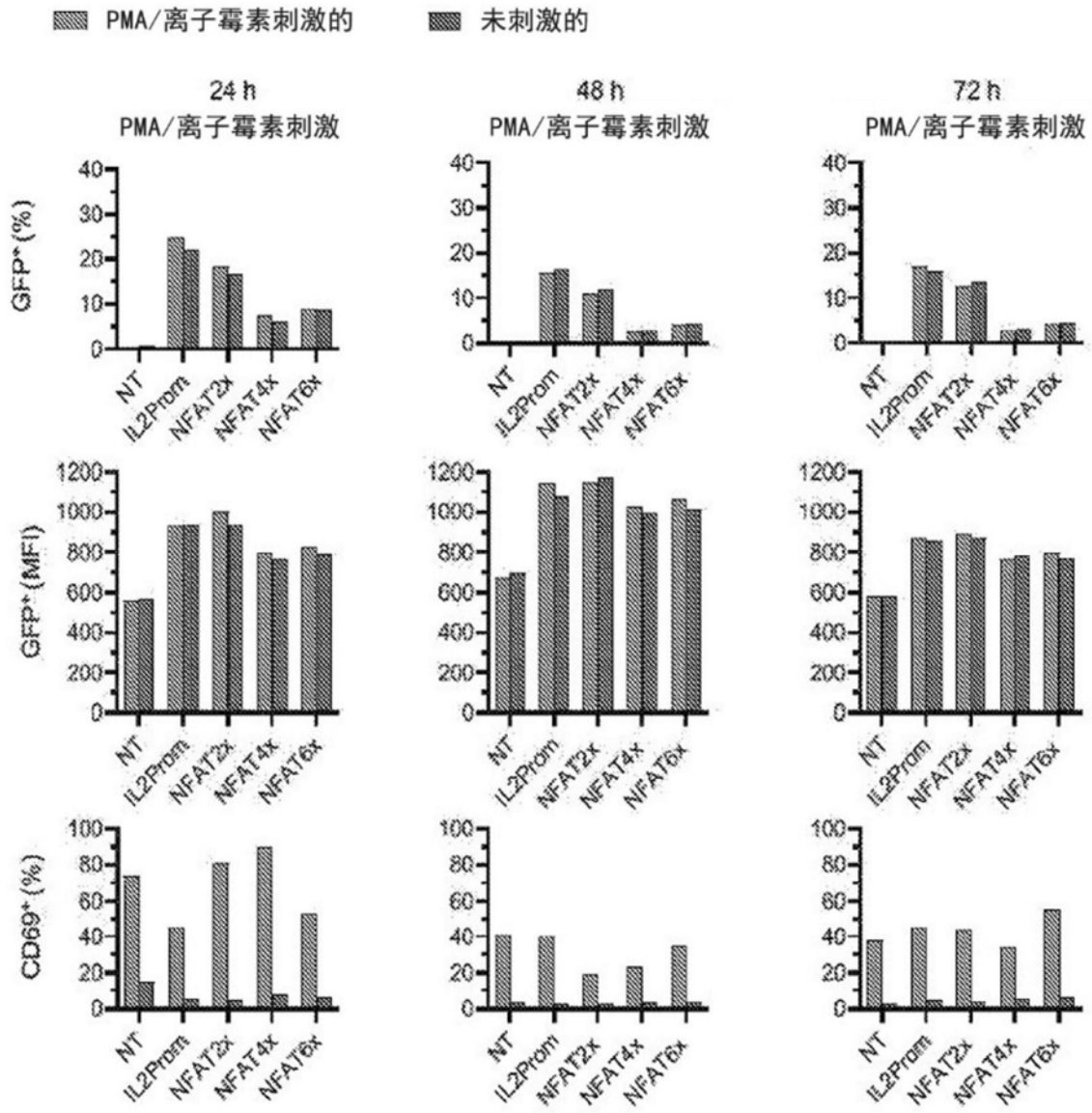


图22c

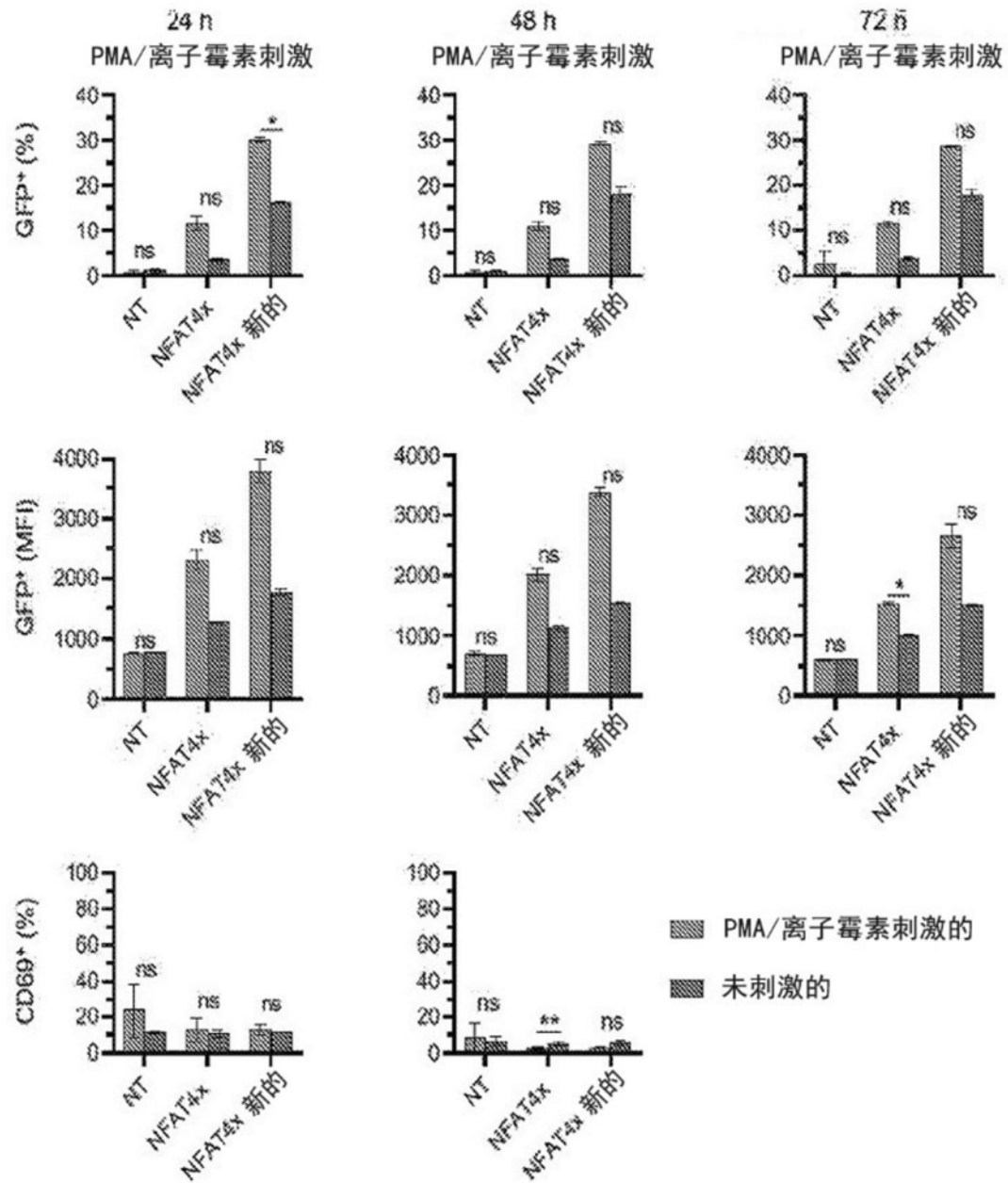


图23

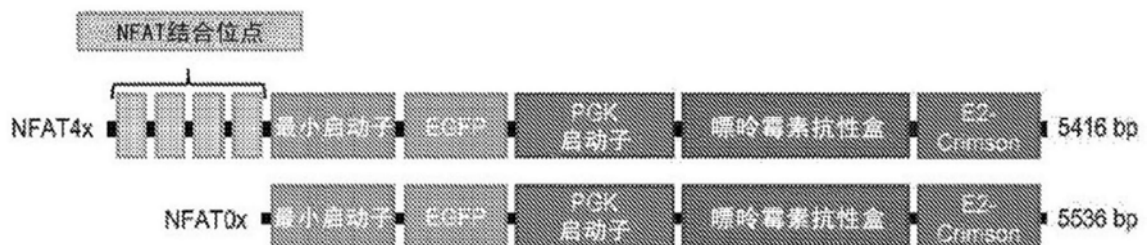


图24a

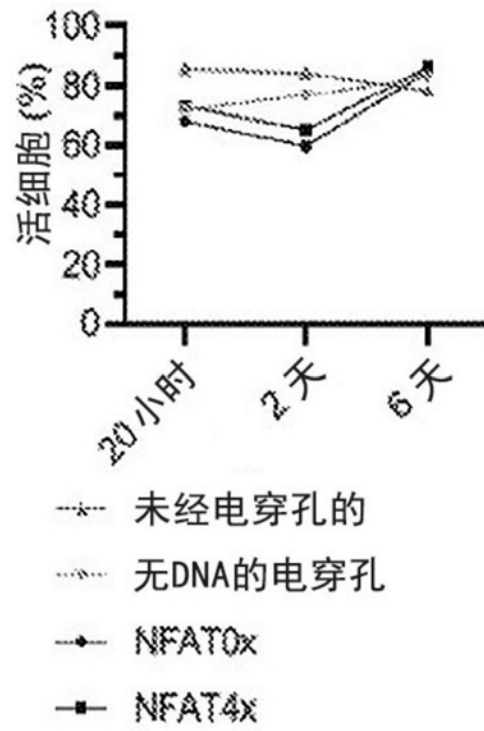


图24b

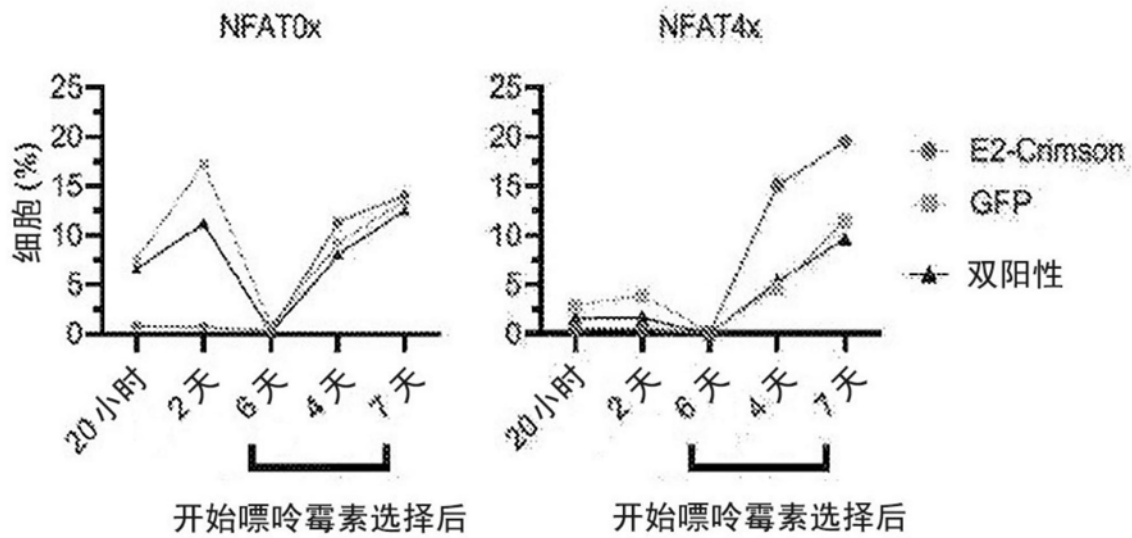


图24c

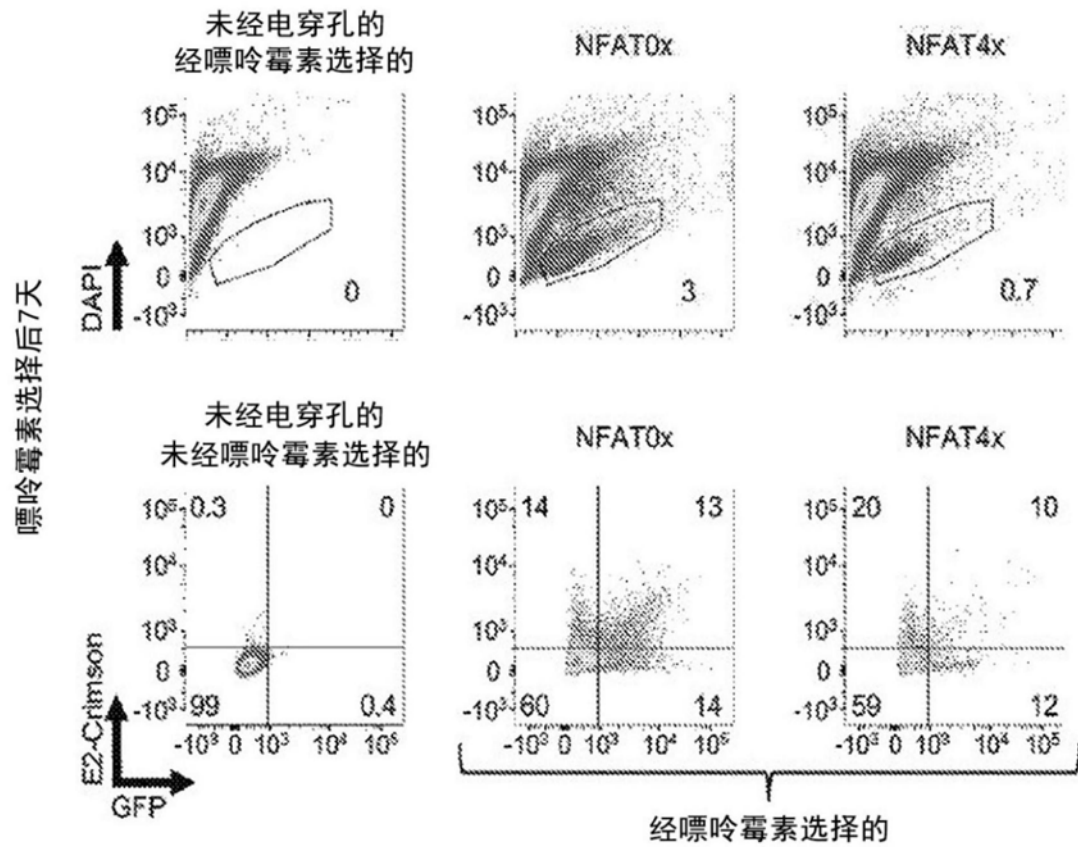


图24d

在第4天去除杀稻瘟菌素

在第4天添加新的杀稻瘟菌素

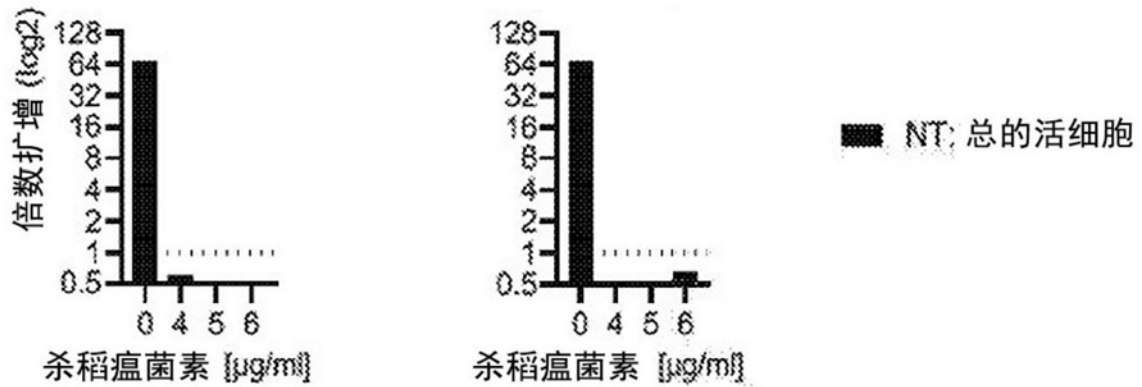


图25a

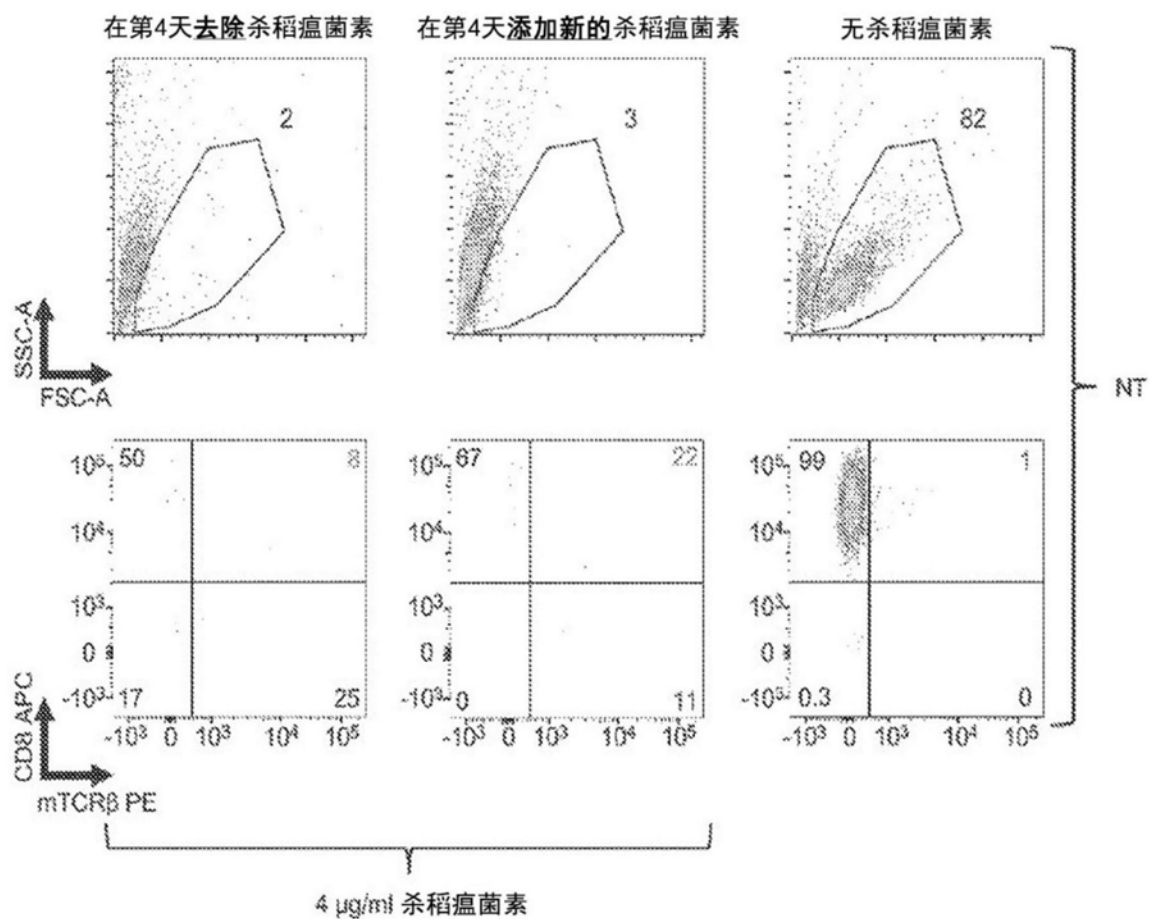


图25b

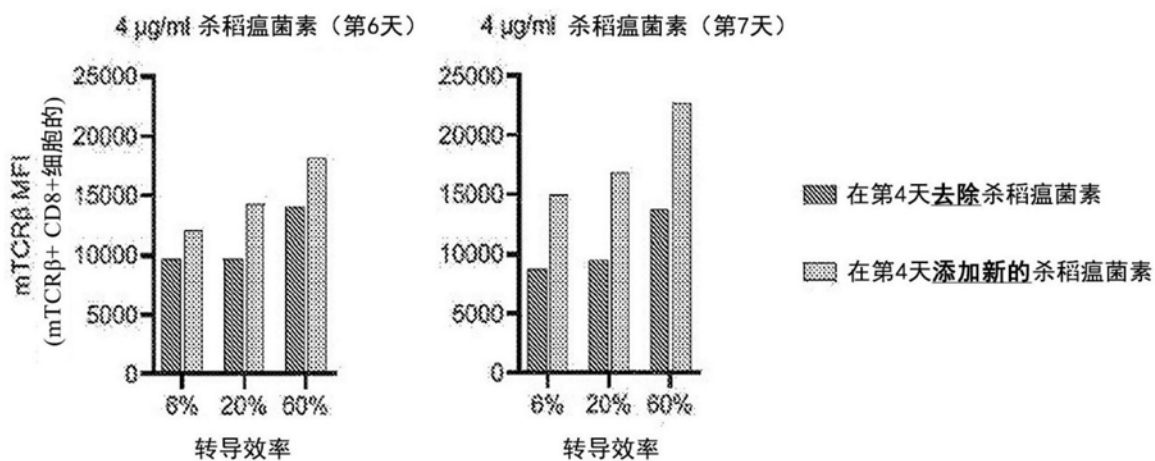


图26a

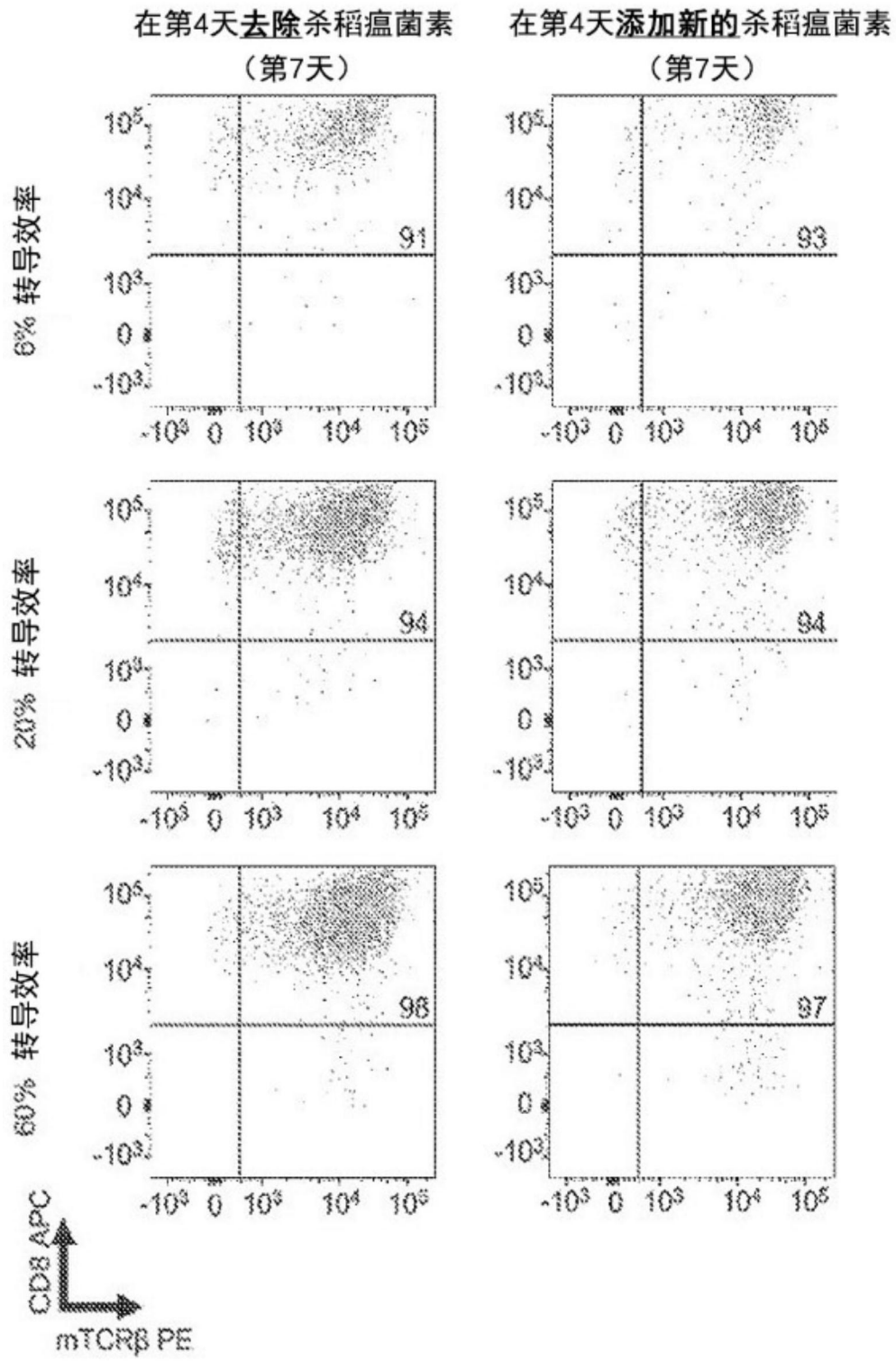


图26b

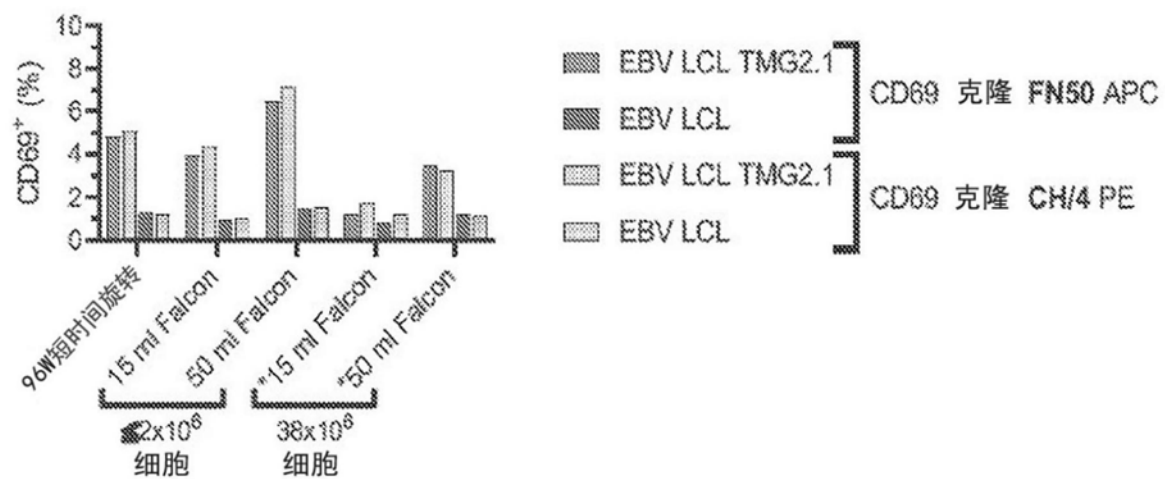


图27

新生抗原特异性TCR分离平台

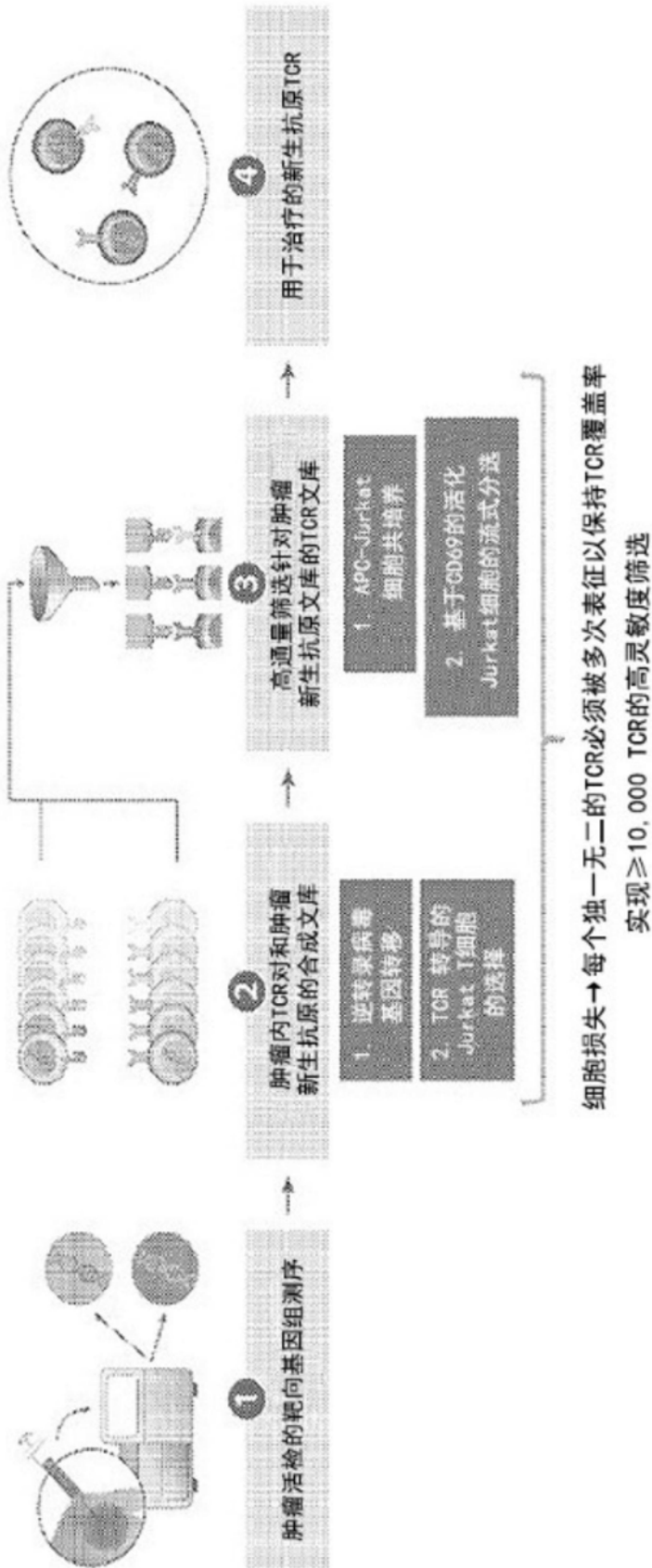


图28

研究的目的是



图29

210

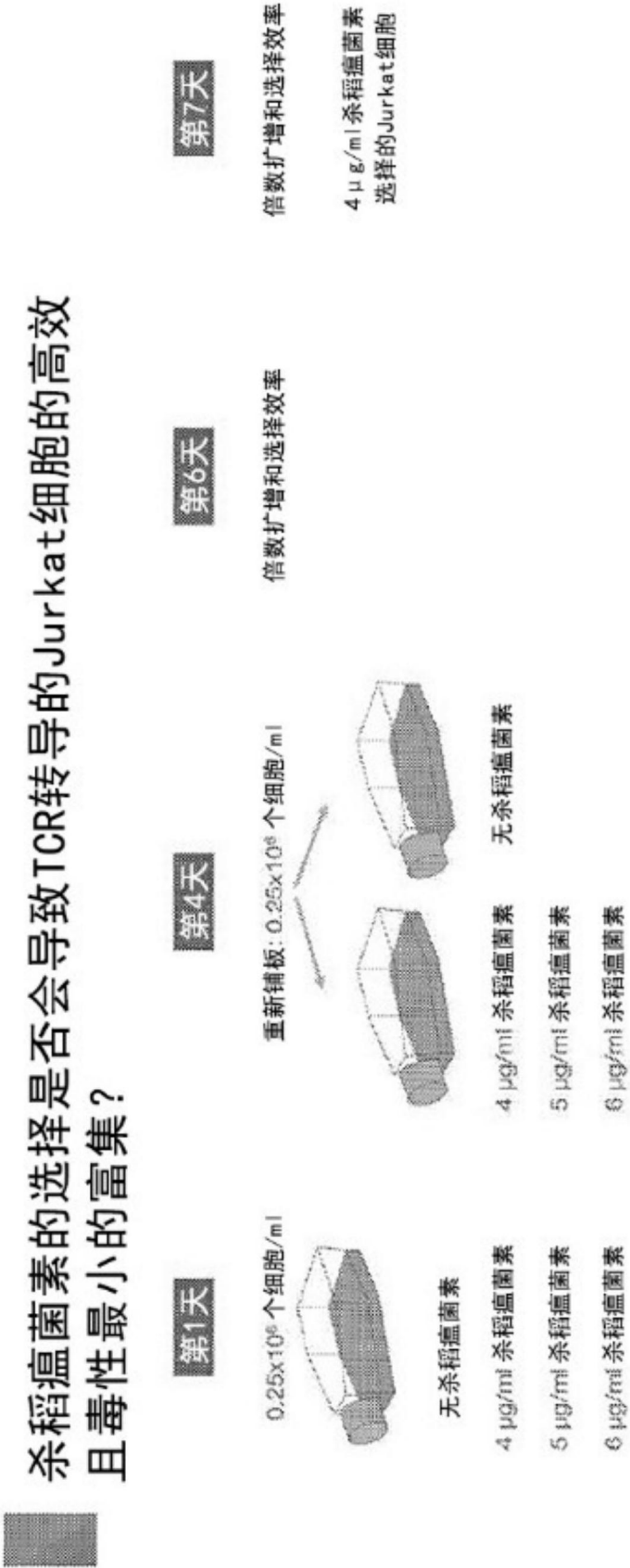


图30

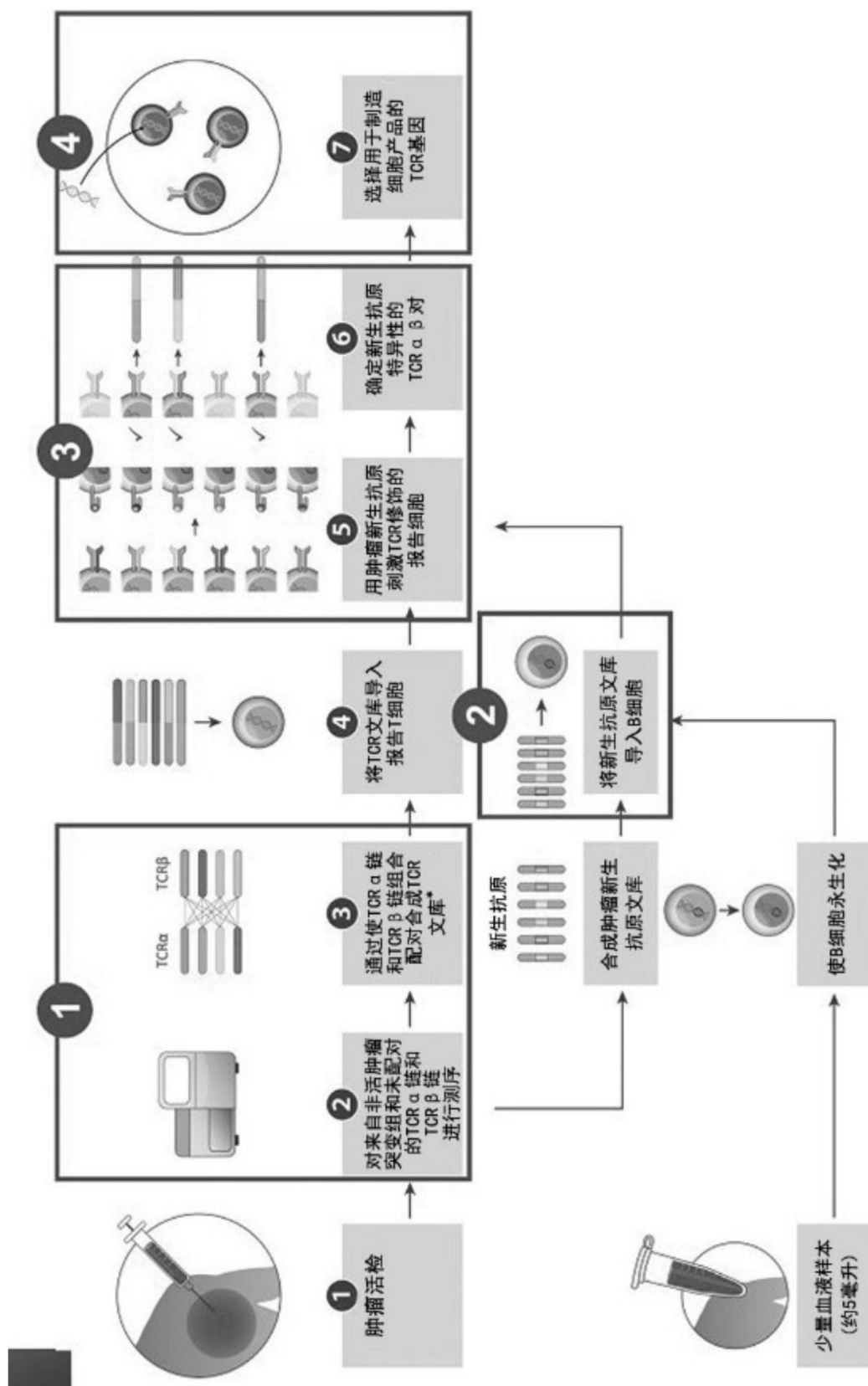


图31

SEQ ID NO:1: 表示为TCR β -P2A-TCR -T2A-嘌呤霉素抗性表达盒的
TCR基因的氨基酸序列

MGTRLLCWAALCLLGAELTEAGVAQSPRYKIIKRQSVAFWCNPISGHATLYWYQQILG
QGPKLLIQFQNNGVVDDSQLPKDRFSAERLKGVDSTLKIQPAKLEDSAVYLCASSLYPY
DGRSSYEQYFGPGTRLTVTEDLRNVTPPKVSLFEP SKAEIANKQKATLVCLARGFFPDHV
ELSWVWNGKEVHSGVCTDPQAYKESNYSYCLSSRLRVSATFWHNPRNHFRCQVQFHG
LSEEDKWPEGSPKPVTONISAEAWGRADCGITSASYHQGVLSATILYEILLGKATLYAVL
VSGLVLMAMVKKKNSGSGATNFSLLKQAGDVEENPGPMTSIRAVFIFLWLQLDLVNGE
NVEQHPSTLSVQEGDSAVIKCTYSDSASNYFPWKQELGKGPQLIIDIRSNVGEKKDQRI
AVTLNKTAKHFSLHITETQPEDSAVYFCAAYNQGGKLIFGQGTELSVKPDIQNPEPAVYQ
LKDPRSQDSTLCLFTDFDSQINVPKTMESGTFITDKCVLDMKAMDSKSNGAIAWSNQT
FTCQDIFKETNATYPSSDVPCDATLTEKSFETDMNLNFQNL SVMGLRILLKLVAGFNLL
MTLRLWSSSGSGEGRGSLTTCGDVEENPGPMTEYKPTVRLATRDDVPRAVRTLAAAFAD
YPATRIHTVDPDRIHIERVTELQELFLTRVGLDIGKVVWADDGA AVAVWTTPESEAGAV
FAEIGPRMAELSGSRLAAQQQMEGLLAPHRPKPAWFLATVGVSPDHQGKGLGSAVVL
PGVEAAERAGVPAFLET SAPRNLPFYERLGFVTADVEVPEGPRTWCMTKPGA

SEQ ID NO: 4: 表示为TCR β -P2A-TCR -T2A-杀稻瘟菌素抗性表达盒的
的TCR基因的氨基酸序列

MGPQLLGYVVLCLLGAGPLEAQVTQNPRYLITVTGKKLTVTCSQNMNHEYMSWYRQD
PGLGLRQIYYSMNVEVTDKGDVPEGYKVS RKEKRNFLILESPSPNQTSLYFCASRASGR
EQFFGPGTRLTVLEDLRNVTPPKVSLFEP SKAEIANKQKATLVCLARGFFPDHVELSWW
VNGKEVHSGVCTDPQAYKESNYSYCLSSRLRVSATFWHNPRNHFRCQVQFHGLSEEDK
WPEGSPKPVTONISAEAWGRADCGITSASYHQGVLSATILYEILLGKATLYAVLVSGLV
MAMVKKKNSGSGATNFSLLKQAGDVEENPGPMKSLRVLLVILWLQLSWVWSQDQV
KQNSPSLSVQEGRISILNCDYTNSMFDYFLWYKKYPAEGPTFLISSIKDKNEDGRFTVF
LNKSAKHL SLHIVPSQPGDSAVYFCA AFLQSNDMRFGAGTRLTVKPDIQNPEPAVYQLK
DPRSQDSTLCLFTDFDSQINVPKTMESGTFITDKCVLDMKAMDSKSNGAIAWSNQTST
CQDIFKETNATYPSSDVPCDATLTEKSFETDMNLNFQNL SVMGLRILLKLVAGFNLLMT
LRLWSSSGSGEGRGSLTTCGDVEENPGPMKTFNISQQDLELVEVATEKITMLYEDNKHVV
GAAIRTKTGEIISAVHIEAYIGRVTVCAEAIAIGSAVSNQKDFDTIVAVRHPYSDEVDRSI
RVVSPCGMCRELISDYAPDCFVLIEMNGKLVKTTIEELIPLKYTRN

图32

CD8 α -P2A-CD8 β 转基因的氨基酸序列 SEQ ID NO: 2

MALPVTALLPLALLHAARPSQFRVSPLDRTWNLGETVELKCQVLLSNPTSGCSWLFQ
PRGAAASPTFLLYLSQNKPKAAEGLDTQRFSGKRLGDTFVLTLSDFRRENEGYYFCSAL
SNSIMYFSHFVPVFLPAKPTTTPAPRPPTPAPTIASQPLSLRPEACRPAAGGAVHTRGLDF
ACDIYWAPLAGTCGVLLLSLVITLYCNHRNRRRVCKCPRPVVKSGDKPSLSARYVGSG
ATNFSLLKQAGDVEENPGPMRPRLWLLAAQLTVLHGNSVLQQTPAYIKVQTNKMVM
LSCEAKISLSNMRIYWLRQRQAPSSDSHHEFLALWDSAKGTIHGEEVEQEKIAVFRDASR
FILNLTSVKPEDSGIYFCMIVGSPELTFGKGTQLSVVDFLPTTAQPTKKSTLKKRVCRLPR
PETQKGPLCSPITLGLLVAGVLVLLVSLGVAIHLCCRERRARLRFMKQFYK

图33

SEQ ID NO3: K562-HLA- A*02:01-IRES-FusionRed的核苷酸序列实例

ATGGCTGTTATGGCCCCTAGAACACTGGTGCTGCTGCTGTCTGGTGCCCTGGCTCTG
ACACAAACATGGGCGGCAGCCACAGCATGCGGTACTTTTCACCAGCGTGTCCAG
ACCTGGCAGAGGCGAGCCTAGATTCAATTGCCGTGGGCTACGTGGACGACACCCAGT
TCGTCAGATTTCGATTCCGATGCCGCCAGCCAGCGGATGGAACCTAGAGCACCTTGG
ATCGAGCAAGAGGGGCCCCGAGTATTGGGACGGCGAGACAAGAAAAGTGAAGGCCC
ACAGCCAGACACACAGAGTGGATCTGGGAACCCTGCGGGGCTACTACAATCAGTCT
GAGGCCGGCTCTCACACCGTGCAGAGGATGTATGGCTGTGACGTGGGCAGCGATTG
GCGGTTCTGAGAGGCTATCACCAGTACGCCTACGACGGCAAGGACTATATCGCCCT
GAAAGAGGACCTGCGGTCTTGACAGCCGCCGATATGGCTGCCCAGACCACAAAGC
ACAAGTGGGAAGCCGCTCACGTGGCCGAACAGCTGAGAGCTTATCTGGAAGGCACC
TGTGTGGAATGGCTGCGGAGATACCTGGAACCGGCAAAGAGACACTGCAGCGGAC
AGACGCCCCTAAGACACACATGACACACCACGCCGTGTCCGACCACGAAGCCACAC
TTAGATGTTGGGCCCTGAGCTTCTACCCCGCCGAGATCACACTGACATGGCAGAGAG
ATGGCGAGGATCAGACCCAGGATACCGAGCTGGTGGAAACAAGACCTGCCGGCGAC
GGCACCTTCCAGAAATGGGCTGCTGTGGTGGTGCCTAGCGGCCAAGAGCAGAGATA
CACCTGTCACGTGCAGCACGAGGGCCTGCCTAAGCCTCTTACACTGAGATGGGAGC
CCAGCAGCCAGCCTACAATCCCCATCGTGGGAATCATTGCCGGCCTGGTGTGTTG
GCGCCGTGATTACAGGTGCAGTGGTGGCCGCTGTTATGTGGCGGAGAAAGAGCAGC
GACAGAAAAGGCGGCAGCTACTCTCAGGCCGCCAGCTCTGATTCTGCCAGGGCTC
TGATGTGTCTCTGACCGCCTGCAAAGTCTGAGCGGCCGCTACGTAAATTCCGCCCC
CCCCCTCCCCCCCCCTAACGTTACTGGCCGAAGCCGCTTGGAATAAGGCCGGTGTG
CGTTTGCTATATGTTATTTTCCACCATATTGCCGTCTTTTGGCAATGTGAGGGCCCCG
GAAACCTGGCCCTGTCTTCTTGACGAGCATTCCTAGGGGTCTTTCCCCTCTCGCCAA
AGGAATGCAAGGTCTGTTGAATGTCGTGAAGGAAGCAGTTCCTCTGGAAGCTTCTTG
AAGACAAACAACGTCTGTAGCGACCCTTTGCAGGCAGCGGAACCCCCCACCTGGCG
ACAGGTGCCTCTGCGGCCAAAAGCCACGTGTATAAGATACACCTGCAAAGGCGGCA
CAACCCACGTGCCACGTTGTGAGTTGGATAGTTGTGGAAAGAGTCAAATGGCTCTCC
TCAAGCGTATTCAACAAGGGGCTGAAGGATGCCCAGAAGGTACCCCAITGTATGGG
ATCTGATCTGGGGCCTCGGTGCACATGCTTTACATGTGTTTAGTCGAGGTTAAAAA
ACGTCTAGGCCCCCGAACCACGGGGACGTGGTTTTCTTTGAAAAACACGATGATA
ATATGGCCACAACCATGGTGTCCGAGCTGATCAAAGAAAACATGCCCATGAAGCTG
TACATGGAAGGGACCGTGAACAACCACCACTTCAAGTGCACCAGCGAAGGCGAGGG
CAAGCCTTATGAGGGCACCCAGACCATGCGGATCAAGGTGGTGGAAAGGTGGCCCTC
TGCCTTTCGCCTTTGATATCCTGGCCACCAGCTTTATGTACGGCAGCCGGACCTTCAT
CAAGCACCCCTCCAGGCATCCCCGATTTCTTCAAGCAGAGCTTTCCCGAGGGCTTCAC
CTGGGAGAGAGTGACCACATACGAGGATGGCGGCGTGCTGACAGCCACACAGGATA
CAAGTCTGCAGGACGGCTGCCTGATCTACAACGTGAAAGTGCGGGGCGTGAACCTC
CCCGCCAATGGACCTGTGATGCAGAAGAAAACCCCTCGGCTGGGAAGCCAGCACCGA
GACAATGTATCCTGCCGACGGTGGACTGGAAGGGGCCTGTGATATGGCCCTGAAGC

TTGTTGGCGGCGGACACCTGATCTGCAACCTGGAAACCACCTACAGAAGCAAGAAG
CCCGCCACCAACCTGAAGATGCCCCGGCGTGTAACGTGGACCACCGGCTGGAAAG
GATCAAAGAGGGCCGACGACGAGACATACGTGGAACAGCACGAAGTGGCCGTGGCC
AGATACTCTACAGGCGGAGCTGGCGACGGCGGCAATGA

图34

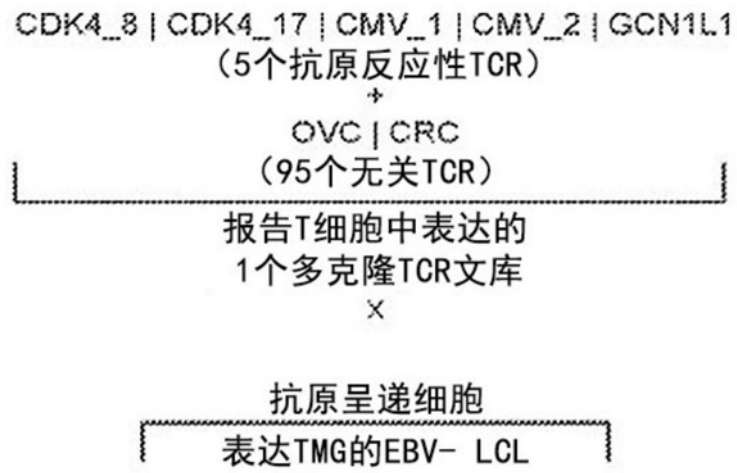


图35A

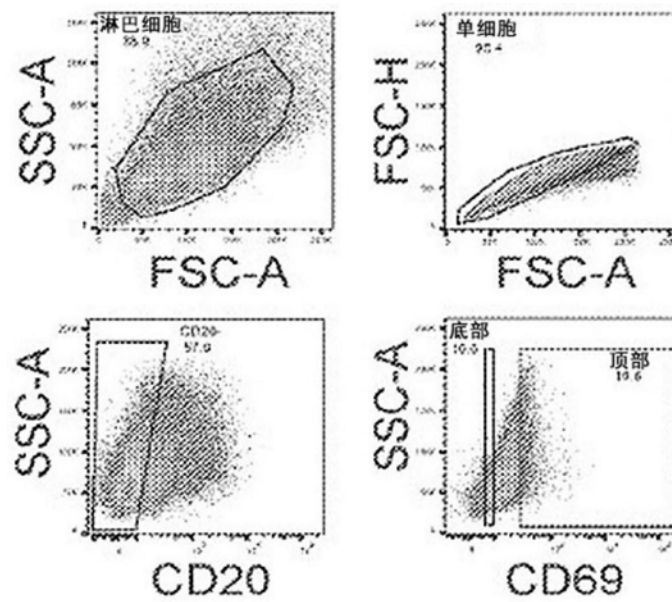


图35B

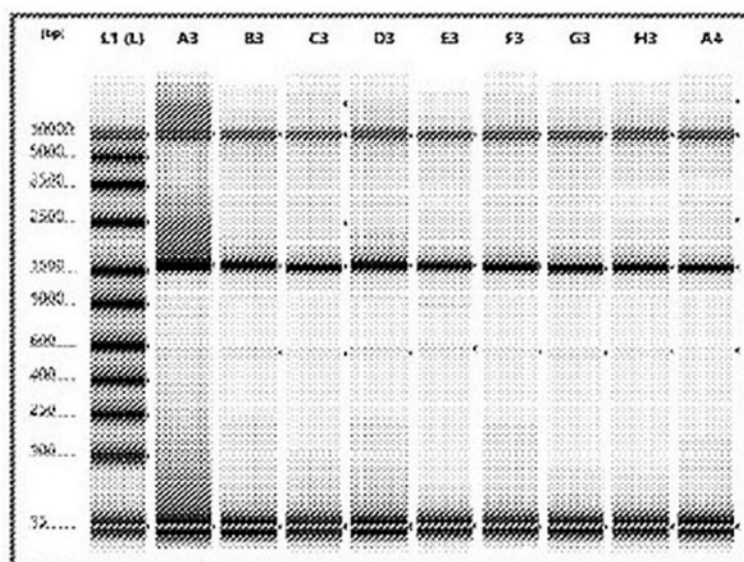


图35C

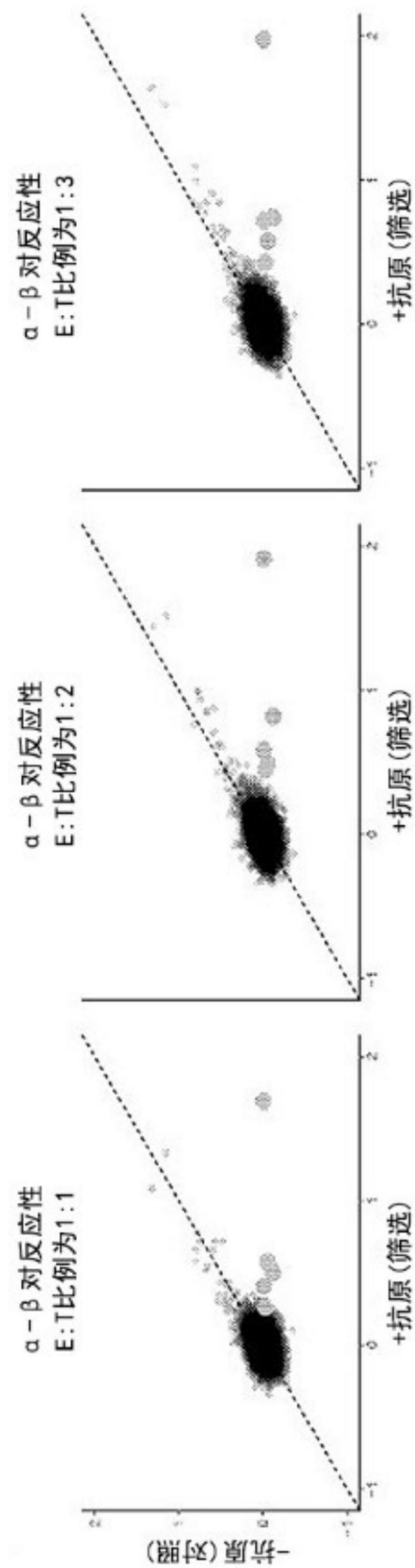


图35D

E:T 1:1				
排序	α-β 对		P值	
1	CDK4_17_TCR α	CDK4_17_TCR β	1.24e-40	
2	CDK4_8_TCR α	CDK4_8_TCR β	6.26e-08	
3	CMV_2_TCR α	CMV_2_TCR β	2.04e-03	
25	CMV_1_TCR α	CMV_1_TCR β	2.45e-02	
159	GCN1L1_TCR α	GCN1L1_TCR β	8.74e-02	

E:T 1:2				
排序	α-β 对		P值	
1	CDK4_17_TCR α	CDK4_17_TCR β	2.04e-10	
2	CDK4_8_TCR α	CDK4_8_TCR β	1.55e-08	
3	CMV_2_TCR α	CMV_2_TCR β	2.53e-05	
5	CMV_1_TCR α	CMV_1_TCR β	4.35e-04	
108	GCN1L1_TCR α	GCN1L1_TCR β	3.52e-02	

E:T 1:3				
排序	α-β 对		P值	
1	CDK4_17_TCR α	CDK4_17_TCR β	1.23e-12	
2	CMV_2_TCR α	CMV_2_TCR β	2.30e-07	
3	CDK4_8_TCR α	CDK4_8_TCR β	2.70e-06	
9	CMV_1_TCR α	CMV_1_TCR β	1.62e-03	
71	GCN1L1_TCR α	GCN1L1_TCR β	2.49e-02	

图35E



图36A

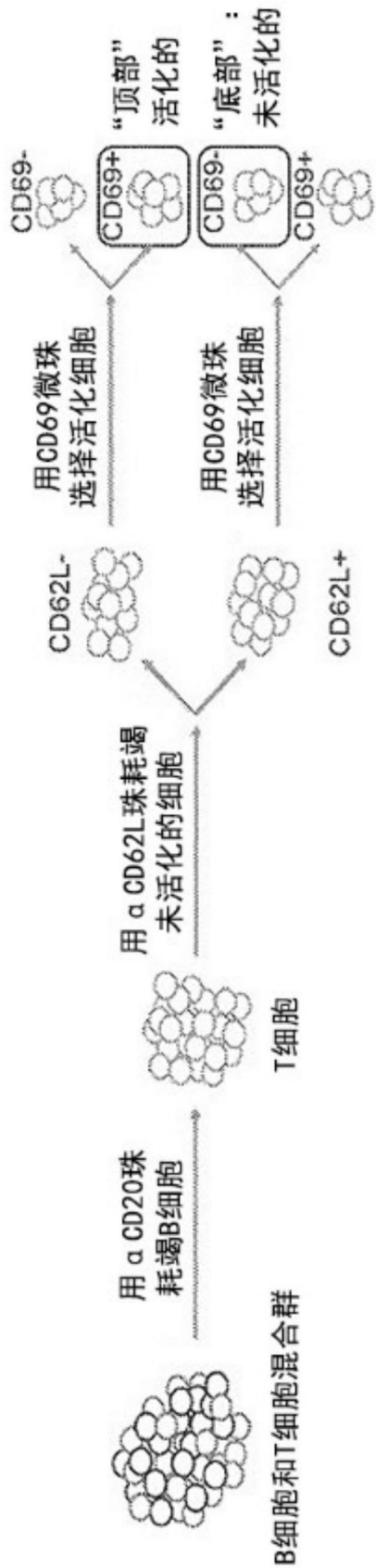


图36B

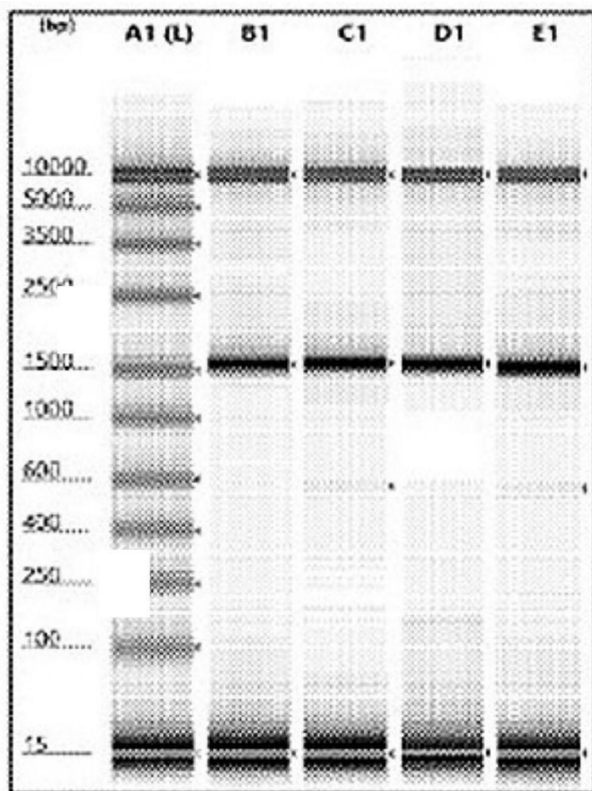


图36C

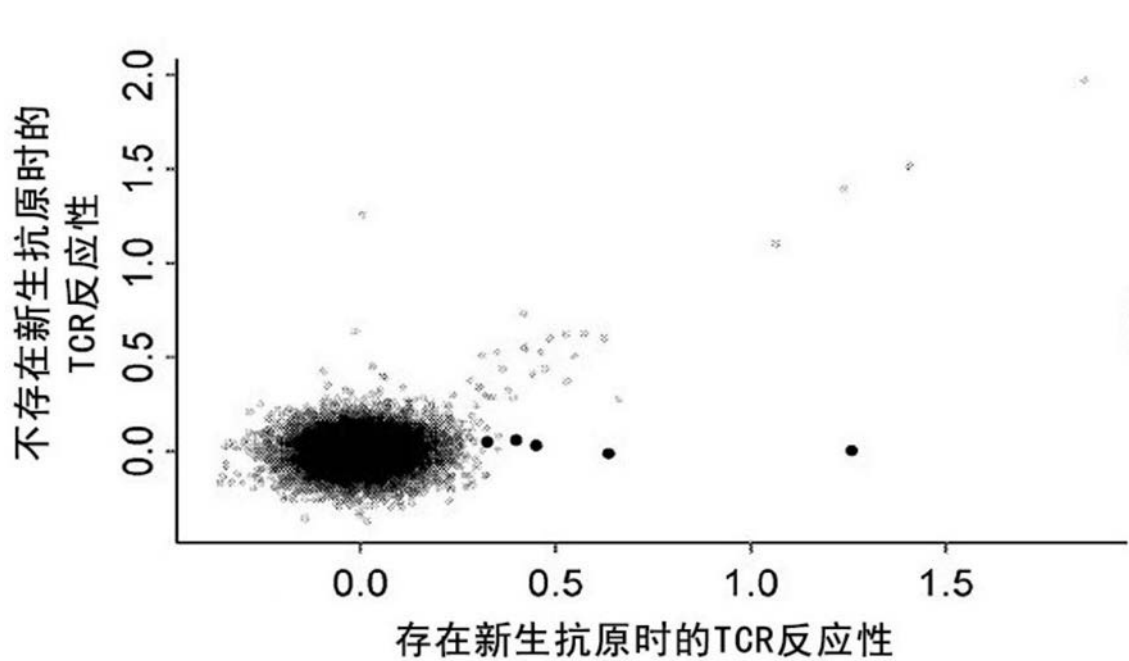


图36D

排序	α	β	碱基平均值	Log2 FC	调整后的p值
1	CDK4_17	CDK4_17	221	1.54	1.37E-18
2	GCN1L1	GCN1L1	99	1.15	4.09E-03
3	CMV_1	CMV_1	113	0.73	1.00E+00
4	OVC21.1	OVC48.1	268	0.55	1.00E+00
5	CRC11.1	OVC48.1	406	0.44	1.00E+00
6	CMV_2	CMV_2	186	0.49	1.00E+00
7	CDK4_8	CDK4_8	64	0.80	1.00E+00

图36E

	池 C1	池 C2	池 C3	池 C4	池 C5	池 C6
池 R1	TMG1	TMG7	TMG13	TMG19	TMG25	TMG31
池 R2	TMG2	TMG8	TMG14	TMG20	TMG26	TMG32
池 R3	TMG3	TMG9	TMG15	TMG21	TMG27	TMG33
池 R4	TMG4	TMG10	TMG16	TMG22	TMG28	TMG34
池 R5	TMG5	TMG11	TMG17	TMG23	TMG29	TMG35
池 R6	TMG6	TMG12	TMG18	TMG24	TMG30	TMG36

图37A

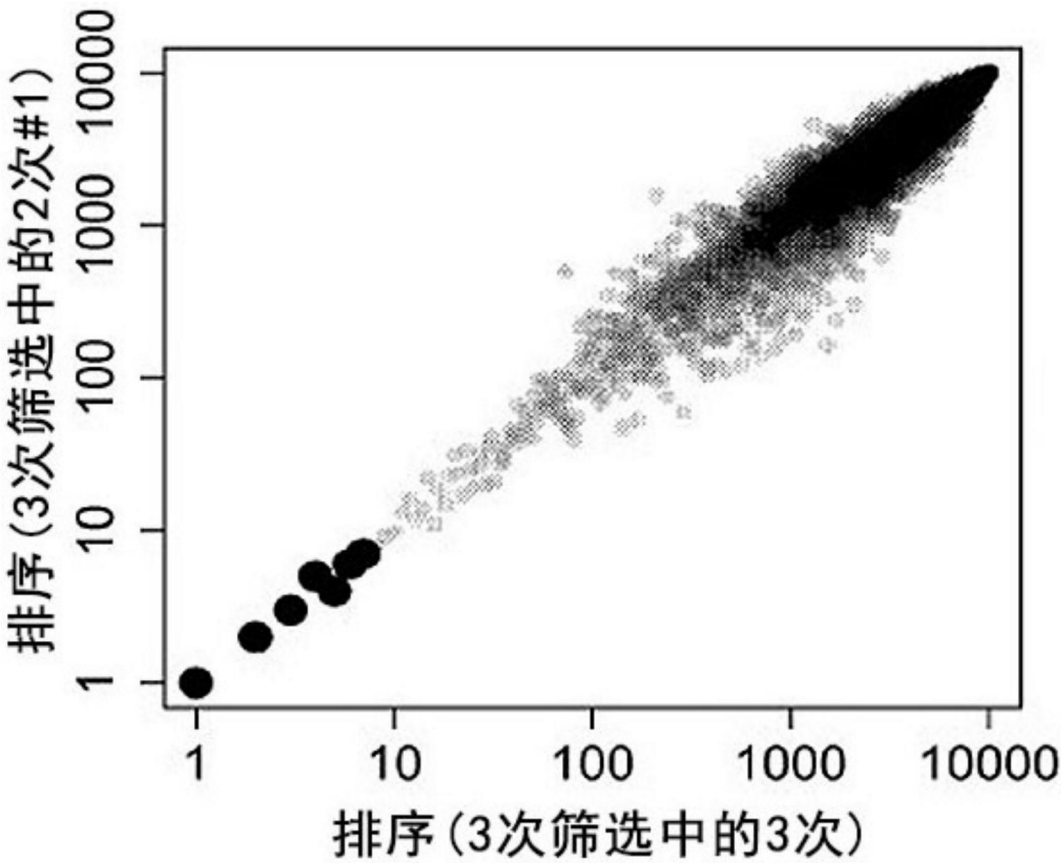


图37B

样本	TCR		3次筛选	2次筛选#1	2次筛选#2	2次筛选#3	排序
pt1	TCR7 α	.TCR5 β	2.1E-11	9.0E-15	4.0E-08	2.8E-12	✓
pt2	TCR24 α	.TCR62 β	8.8E-106	3.4E-133	6.9E-120	7.3E-117	✓
pt2	TCR76 α	.TCR20 β	1.4E-87	5.6E-93	1.5E-91	7.4E-114	✓
pt2	TCR24 α	.TCR43 β	9.7E-72	9.3E-76	4.6E-56	2.3E-66	✓
pt2	TCR63 α	.TCR17 β	1.1E-51	2.7E-56	8.9E-49	1.2E-68	✓
pt2	TCR61 α	.TCR97 β	4.9E-51	2.5E-56	5.1E-60	2.9E-49	✓
pt2	TCR83 α	.TCR29 β	3.8E-40	1.2E-38	1.3E-35	5.5E-39	✓
pt2	TCR86 α	.TCR35 β	5.5E-39	8.3E-36	3.7E-41	4.0E-33	✓
pt3	TCR55 α	.TCR24 β	1.5E-21	5.2E-30	1.5E-29	4.3E-21	✓
pt4	TCR43 α	.TCR16 β	1.8E-15	1.7E-17	5.4E-16	8.5E-11	✓
pt4	TCR9 α	.TCR14 β	1.8E-07	3.3E-06	1.3E-10	7.5E-05	✓

图37C

样本	抗原
1	TMG1
2	TMG2
3	TMG3
4	TMG4
5	混合物(TMGI-4)
6	无TMG

图37D

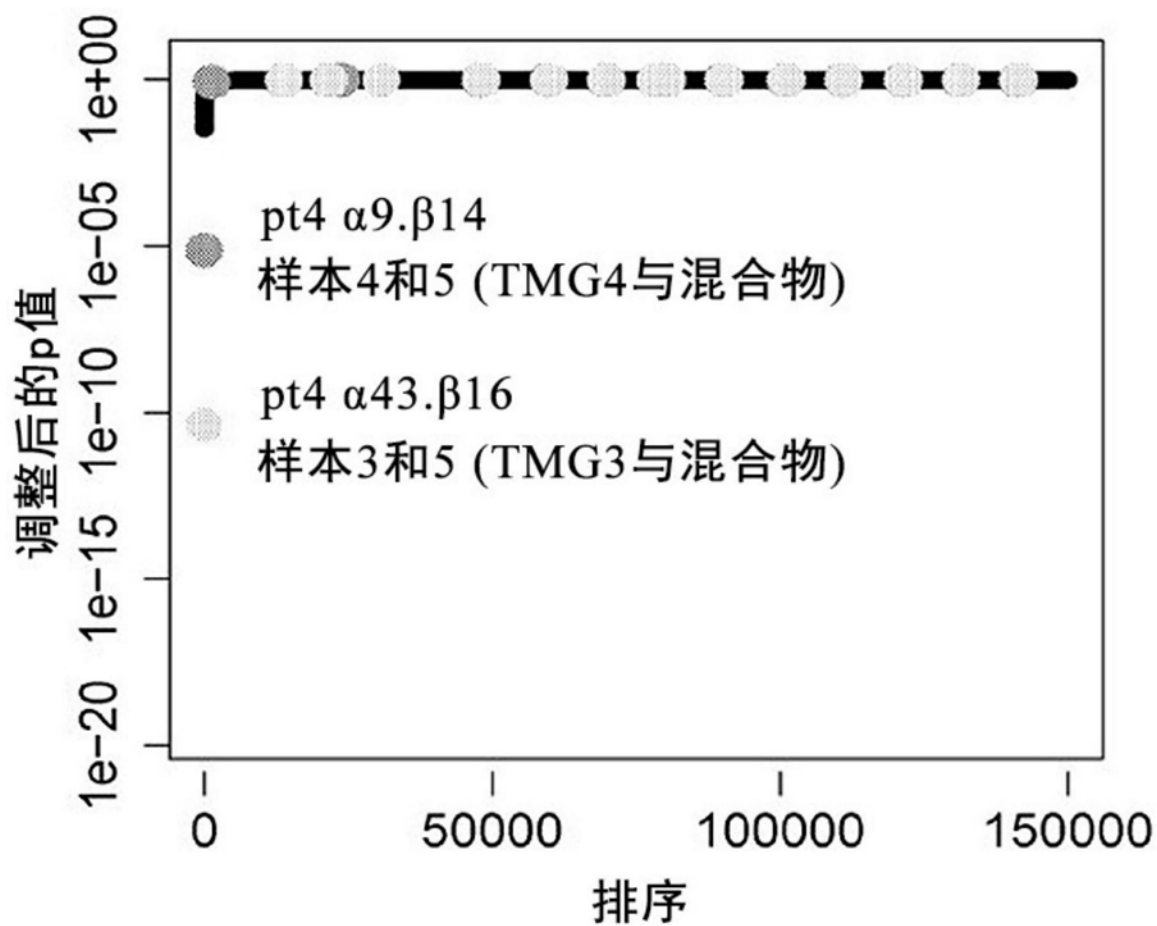


图37E

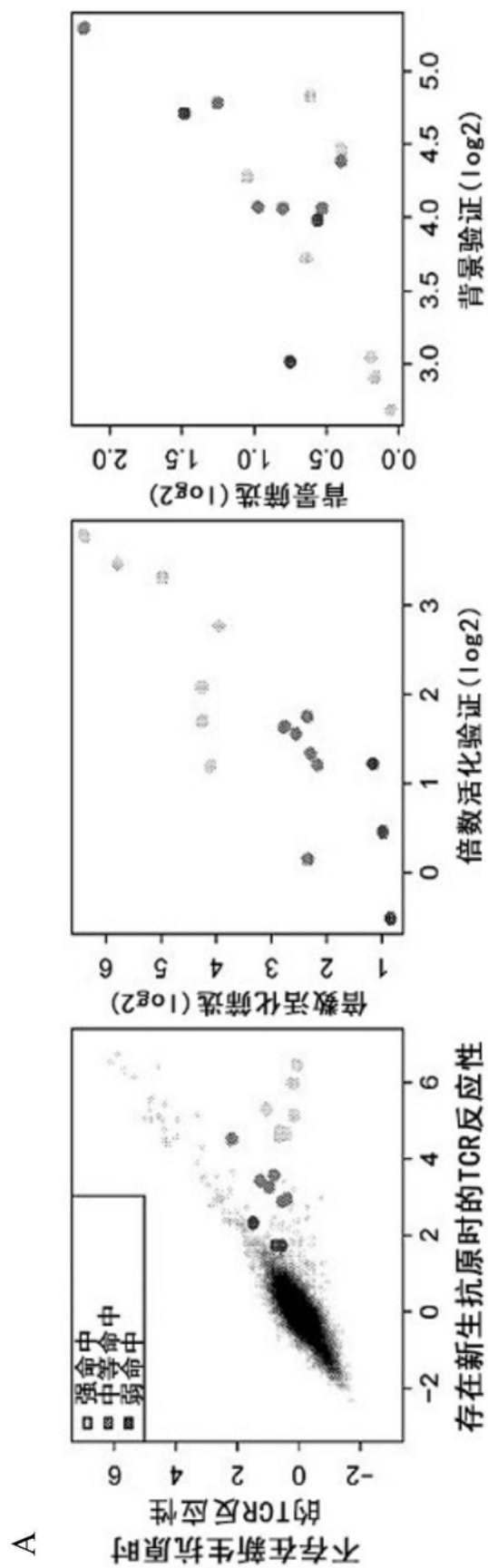


图38

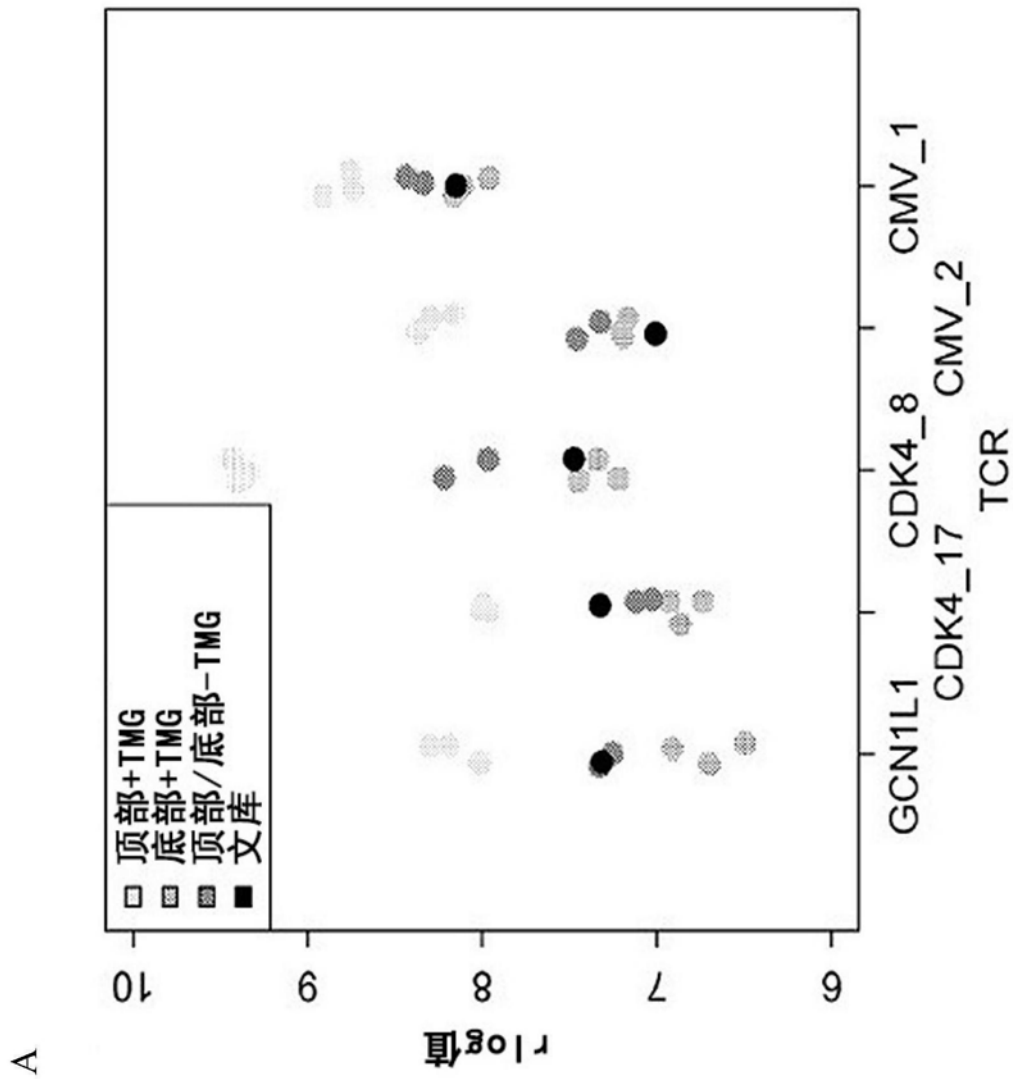


图39

SEQ ID NO: 1, 表示为TCR β -P2ATCR α -T2A-嘌呤霉素抗性表达盒的TCR基因的图解。

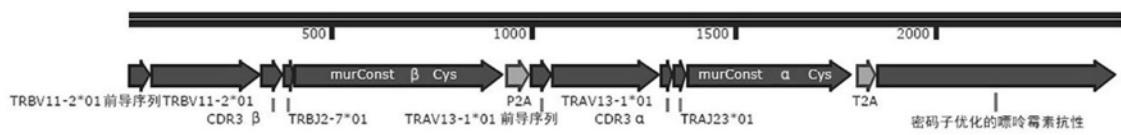


图40

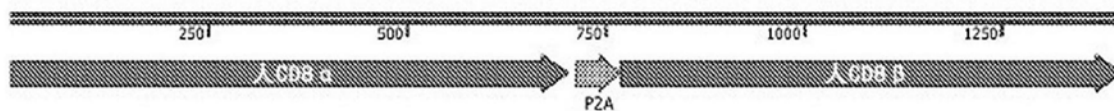
SEQ ID NO: 2, CD8 α -P2A-CD8 β 转基因的图解

图41

SEQ ID NO 3: K562-HLA-A*02:01-IRESFusionRed的图解

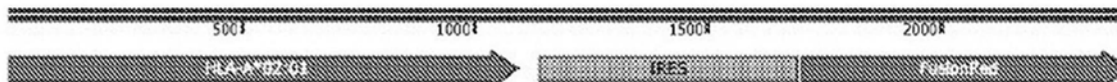


图42

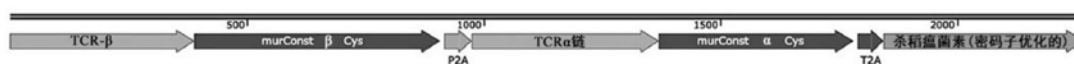
SEQ ID NO: 4, 表示为TCR β -P2A-TCR α -T2A杀稻瘟菌素抗性表达盒的TCR基因的图解。

图43