



# (12) 发明专利申请

(10) 申请公布号 CN 116615556 A

(43) 申请公布日 2023. 08. 18

(21) 申请号 202180078244.X

(22) 申请日 2021.11.19

(30) 优先权数据

63/116,571 2020.11.20 US

63/241,486 2021.09.07 US

(85) PCT国际申请进入国家阶段日

2023.05.19

(86) PCT国际申请的申请数据

PCT/US2021/060197 2021.11.19

(87) PCT国际申请的公布数据

W02022/109339 EN 2022.05.27

(71) 申请人 贝克顿迪金森公司

地址 美国新泽西州

申请人 伊缪德克斯私人有限责任公司

(72) 发明人 辛西娅·萨科夫斯基

凯瑟琳·拉扎鲁克

玛格丽特·纳卡莫托

戴文·约瑟夫·延森

基文·达尔·雅各布森

(74) 专利代理机构 北京安信方达知识产权代理有限公司 11262

专利代理人 刘晓杰 武晶晶

(51) Int.Cl.

G12Q 1/6806 (2006.01)

权利要求书18页 说明书122页

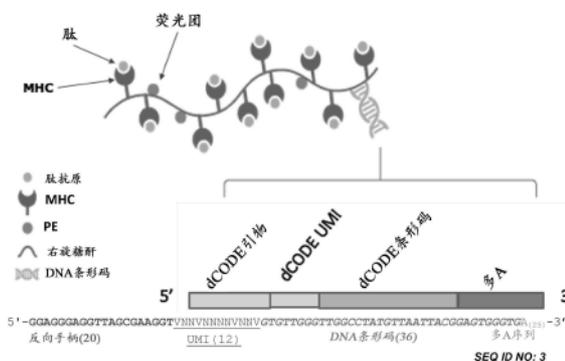
序列表3页 附图30页

(54) 发明名称

DEXTRAMER在单细胞分析中的使用

(57) 摘要

本文的公开内容包括适于在单细胞分析中使用dextramer的系统、方法、组合物和试剂盒。在一些实施方案中,使用一种或更多种允许产生蛋白(例如,抗体)和dextramer的单独文库的引物。



1. 一种方法,所述方法包括:

使多于一种受体检测构建体与多于一个细胞接触以形成与所述受体检测构建体关联的第一多于一个细胞,其中所述多于一个细胞包含多于一种细胞组分靶、和核酸靶的拷贝,其中所述多于一个细胞中的一个或更多个细胞包含受体结合试剂能够特异性结合的受体,并且其中所述多于一种受体检测构建体中的每一种包含两种或更多种受体结合试剂和受体结合试剂特异性寡核苷酸,所述受体结合试剂特异性寡核苷酸包含用于所述受体结合试剂的独特受体标识符序列;

使多于一种细胞组分结合试剂与和所述受体检测构建体关联的多于一个细胞接触以形成第二多于一个细胞,其中所述多于一种细胞组分结合试剂中的每一种包含细胞组分结合试剂特异性寡核苷酸,所述细胞组分结合试剂特异性寡核苷酸包含用于所述细胞组分结合试剂的独特标识符序列,并且其中所述细胞组分结合试剂能够与所述多于一种细胞组分靶中的至少一种特异性结合;

用多于一种寡核苷酸条形码对所述细胞组分结合试剂特异性寡核苷酸进行条形码化以产生多于一种条形码化细胞组分结合试剂特异性寡核苷酸,所述多于一种条形码化细胞组分结合试剂特异性寡核苷酸各自包含与所述独特标识符序列的至少一部分互补的序列;

用所述多于一种寡核苷酸条形码对所述受体结合试剂特异性寡核苷酸进行条形码化以产生多于一种条形码化受体结合试剂特异性寡核苷酸,所述多于一种条形码化受体结合试剂特异性寡核苷酸各自包含与所述独特受体标识符序列的至少一部分互补的序列;

用所述多于一种寡核苷酸条形码对所述多于一个细胞的所述核酸靶的拷贝进行条形码化以产生多于一种条形码化核酸分子,所述多于一种条形码化核酸分子各自包含与所述核酸靶的至少一部分互补的序列;

产生包含多于一种核酸靶文库成员、多于一种细胞组分靶文库成员和多于一种受体文库成员的测序文库,其中产生所述测序文库包括:

将测序衔接子附接至所述多于一种条形码化核酸分子或其产物,以产生所述多于一种核酸靶文库成员;以及

将测序衔接子附接至所述多于一种条形码化细胞组分结合试剂特异性寡核苷酸或其产物,以产生所述多于一种细胞组分靶文库成员;以及

将测序衔接子附接至所述多于一种条形码化受体结合试剂特异性寡核苷酸或其产物,以产生所述多于一种受体文库成员;以及

获得包含核酸靶文库成员的多于一个测序读段、细胞组分靶文库成员的多于一个测序读段和受体文库成员的多于一个测序读段的测序数据。

2. 一种方法,所述方法包括:

使多于一种受体检测构建体与多于一个细胞接触以形成与所述受体检测构建体关联的第一多于一个细胞,其中所述多于一个细胞包含多于一种细胞组分靶,其中所述多于一个细胞中的一个或更多个细胞包含受体结合试剂能够特异性结合的受体,并且其中所述多于一种受体检测构建体中的每一种包含两种或更多种受体结合试剂和受体结合试剂特异性寡核苷酸,所述受体结合试剂特异性寡核苷酸包含用于所述受体结合试剂的独特受体标识符序列;

使多于一种细胞组分结合试剂与和所述受体检测构建体关联的多于一个细胞接触以

形成第二多于一个细胞,其中所述多于一种细胞组分结合试剂中的每一种包含细胞组分结合试剂特异性寡核苷酸,所述细胞组分结合试剂特异性寡核苷酸包含用于所述细胞组分结合试剂的独特标识符序列,并且其中所述细胞组分结合试剂能够与所述多于一种细胞组分靶中的至少一种特异性结合;

用多于一种寡核苷酸条形码对所述细胞组分结合试剂特异性寡核苷酸进行条形码化以产生多于一种条形码化细胞组分结合试剂特异性寡核苷酸,所述多于一种条形码化细胞组分结合试剂特异性寡核苷酸各自包含与所述独特标识符序列的至少一部分互补的序列;

用所述多于一种寡核苷酸条形码对所述受体结合试剂特异性寡核苷酸进行条形码化以产生多于一种条形码化受体结合试剂特异性寡核苷酸,所述多于一种条形码化受体结合试剂特异性寡核苷酸各自包含与所述独特受体标识符序列的至少一部分互补的序列;

产生包含多于一种细胞组分靶文库成员和多于一种受体文库成员的测序文库,其中产生所述测序文库包括:

将测序衔接子附接至所述多于一种条形码化细胞组分结合试剂特异性寡核苷酸或其产物,以产生所述多于一种细胞组分靶文库成员;以及

将测序衔接子附接至所述多于一种条形码化受体结合试剂特异性寡核苷酸或其产物,以产生所述多于一种受体文库成员;以及

获得包含细胞组分靶文库成员的多于一个测序读段和受体文库成员的多于一个测序读段的测序数据。

3. 根据权利要求1-2中任一项所述的方法,其中用所述多于一种寡核苷酸条形码对所述多于一个细胞的所述核酸靶的拷贝进行条形码化包括:

使所述多于一个细胞的所述核酸靶的拷贝与所述多于一种寡核苷酸条形码接触,其中所述多于一种寡核苷酸条形码中的每一种寡核苷酸条形码包含第一通用序列、第一分子标记和能够与所述核酸靶杂交的靶结合区;以及

使与所述核酸靶的拷贝杂交的所述多于一种寡核苷酸条形码延伸以产生多于一种条形码化核酸分子,所述多于一种条形码化核酸分子各自包含与所述核酸靶的至少一部分互补的序列。

4. 根据权利要求1-3中任一项所述的方法,其中用所述多于一种寡核苷酸条形码对所述细胞组分结合试剂特异性寡核苷酸进行条形码化包括:

使所述细胞组分结合试剂特异性寡核苷酸与所述多于一种寡核苷酸条形码接触,其中所述多于一种寡核苷酸条形码中的每一种寡核苷酸条形码包含第一通用序列、第一分子标记和能够与所述细胞组分结合试剂特异性寡核苷酸杂交的靶结合区;以及

使与所述细胞组分结合试剂特异性寡核苷酸杂交的所述多于一种寡核苷酸条形码延伸以产生多于一种条形码化细胞组分结合试剂特异性寡核苷酸,所述多于一种条形码化细胞组分结合试剂特异性寡核苷酸各自包含与所述独特标识符序列的至少一部分互补的序列。

5. 根据权利要求1-4中任一项所述的方法,其中用所述多于一种寡核苷酸条形码对所述受体结合试剂特异性寡核苷酸进行条形码化包括:

使受体结合试剂特异性寡核苷酸与所述多于一种寡核苷酸条形码接触,其中所述多于一种寡核苷酸条形码中的每一种寡核苷酸条形码包含第一通用序列、第一分子标记和能够

与所述受体结合试剂特异性寡核苷酸杂交的靶结合区;以及

使与所述受体结合试剂特异性寡核苷酸杂交的所述多于一种寡核苷酸条形码延伸以产生多于一种条形码化受体结合试剂特异性寡核苷酸,所述多于一种条形码化受体结合试剂特异性寡核苷酸各自包含与所述独特受体标识符序列的至少一部分互补的序列。

6. 根据权利要求1-5中任一项所述的方法,其中所述靶结合区包含捕获序列,任选地(i)所述细胞组分结合试剂特异性寡核苷酸包含与配置为捕获所述细胞组分结合试剂特异性寡核苷酸的捕获序列互补的序列,和/或(ii)所述受体结合试剂特异性寡核苷酸包含与配置为捕获所述受体结合试剂特异性寡核苷酸的捕获序列互补的序列。

7. 根据权利要求1-6中任一项所述的方法,其中:

(i) 所述多于一种条形码化核酸分子中的每一种条形码化核酸分子包含第一通用序列和第一分子标记;

(ii) 所述多于一种条形码化细胞组分结合试剂特异性寡核苷酸中的每一种条形码化细胞组分结合试剂特异性寡核苷酸包含第一通用序列和第一分子标记;和/或

(iii) 所述多于一种条形码化受体结合试剂特异性寡核苷酸中的每一种条形码化受体结合试剂特异性寡核苷酸包含第一通用序列和第一分子标记。

8. 根据权利要求1-7中任一项所述的方法,其中所述第二多于一个细胞包含一个或更多个单细胞,所述方法包括在(i)使所述多于一个细胞的所述核酸靶的拷贝与所述多于一种寡核苷酸条形码接触,(ii)使所述细胞组分结合试剂特异性寡核苷酸与所述多于一种寡核苷酸条形码接触,和/或(iii)使所述受体结合试剂特异性寡核苷酸与所述多于一种寡核苷酸条形码接触之前:

将所述第二多于一个细胞分区至多于一个分区,其中所述多于一个分区中的分区包含来自所述第二多于一个细胞的单细胞;以及

在包含所述单细胞的分区中,(i)使所述多于一个细胞的所述核酸靶的拷贝与所述多于一种寡核苷酸条形码接触,(ii)使所述细胞组分结合试剂特异性寡核苷酸与所述多于一种寡核苷酸条形码接触,和/或(iii)使所述受体结合试剂特异性寡核苷酸与所述多于一种寡核苷酸条形码接触,任选地所述分区是孔或液滴。

9. 根据权利要求1-8中任一项所述的方法,其中所述多于一种寡核苷酸条形码与固体支持物关联,所述方法包括将所述固体支持物与所述分区中的所述单细胞关联,并且其中所述多于一个分区中的分区包含单个固体支持物,任选地所述方法包括在所述分区步骤之后且在所述接触步骤之前裂解所述单细胞,任选地裂解所述单细胞包括加热、与去污剂接触、改变pH或其任何组合。

10. 根据权利要求1-9中任一项所述的方法,其中所述多于一个细胞包括T细胞、B细胞、肿瘤细胞、髓样细胞、血细胞、正常细胞、胎儿细胞、母体细胞或其混合物。

11. 根据权利要求1-10中任一项所述的方法,其中所述多于一种寡核苷酸条形码中的至少10种包含不同的第一分子标记序列。

12. 根据权利要求1-11中任一项所述的方法,其中所述多于一种寡核苷酸条形码各自包含细胞标记,任选地所述多于一种寡核苷酸条形码的每一种细胞标记包含至少6个核苷酸,任选地所述多于一种寡核苷酸条形码中与同一固体支持物关联的寡核苷酸条形码包含相同的细胞标记,并且任选地所述多于一种寡核苷酸条形码中与不同固体支持物关联的寡

核苷酸条形码包含不同的细胞标记。

13. 根据权利要求9-12中任一项所述的方法,其中所述固体支持物包括合成颗粒、平坦表面或其组合;并且任选地所述多于一种寡核苷酸条形码中的至少一种寡核苷酸条形码被固定或部分固定在所述合成颗粒上,或者所述多于一种寡核苷酸条形码中的至少一种寡核苷酸条形码被包封或部分包封在所述合成颗粒中。

14. 根据权利要求1-13中任一项所述的方法,其中所述靶结合区包含多(dT)区、随机序列、靶特异性序列或其组合。

15. 根据权利要求1-14中任一项所述的方法,其中产生所述测序文库包括:

使随机引物与所述多于一种条形码化核酸分子接触,其中所述随机引物中的每一种包含第四通用序列或其互补物;以及

使与所述多于一种条形码化核酸分子杂交的所述随机引物延伸以产生第一多于一种延伸产物。

16. 根据权利要求15所述的方法,所述方法包括:

使用能够与所述第一通用序列或其互补物杂交的引物和能够与所述第四通用序列或其互补物杂交的引物扩增所述第一多于一种延伸产物,从而产生第一多于一种条形码化扩增子,

其中所述多于一种核酸靶文库成员包括所述第一多于一种条形码化扩增子或其产物,任选地,扩增所述第一多于一种延伸产物包括将测序引物的结合位点和/或测序衔接子、其互补序列和/或其部分的序列添加到所述第一多于一种延伸产物,以及

还任选地,所述方法包括基于与所述第一多于一种条形码化扩增子或其产物关联的具有不同序列的第一分子标记的数目来确定所述一个或更多个单细胞中的每一个中所述核酸靶的拷贝数。

17. 根据权利要求16所述的方法,其中产生所述测序文库包括:

使用能够与所述第一通用序列或其互补物杂交的引物和能够与所述第四通用序列或其互补物杂交的引物扩增所述第一多于一种条形码化扩增子,从而产生第二多于一种条形码化扩增子;

其中所述多于一种核酸靶文库成员包括所述第二多于一种条形码化扩增子或其产物,任选地,扩增所述第一多于一种条形码化扩增子包括将测序引物的结合位点和/或测序衔接子、其互补序列和/或其部分的序列添加到所述第一多于一种条形码化扩增子,以及

还任选地,所述方法包括基于与所述第二多于一种条形码化扩增子或其产物关联的具有不同序列的第一分子标记的数目来确定所述一个或更多个单细胞中的每一个中所述核酸靶的拷贝数。

18. 根据权利要求1-17中任一项所述的方法,其中产生所述测序文库包括:

使用所述多于一种条形码化核酸分子作为模板合成第三多于一种条形码化扩增子以产生第三多于一种条形码化扩增子,

其中所述多于一种核酸靶文库成员包括所述第三多于一种条形码化扩增子或其产物,任选地,合成所述第三多于一种条形码化扩增子包括使用能够与所述第一通用序列或其互补物杂交的引物和靶特异性引物进行PCR扩增,

还任选地,合成所述第三多于一种条形码化扩增子包括将测序引物的结合位点和/或

测序衔接子、其互补序列和/或其部分的序列添加到所述条形码化核酸分子,以及

任选地,所述方法包括基于与所述第三多于一种条形码化扩增子或其产物关联的具有不同序列的第一分子标记的数目来确定所述一个或更多个单细胞中的每一个中所述核酸靶的拷贝数。

19. 根据权利要求18所述的方法,其中所述靶特异性引物与免疫受体特异性杂交,任选地所述靶特异性引物与免疫受体的恒定区、免疫受体的可变区、免疫受体的多样性区、免疫受体的可变区和多样性区的连接部或其组合特异性杂交;并且任选地所述免疫受体是T细胞受体(TCR)和/或B细胞受体(BCR);任选地所述TCR包含TCR $\alpha$ 链、TCR $\beta$ 链、TCR $\gamma$ 链、TCR $\delta$ 链或其任何组合;并且还任选地,所述BCR包含BCR重链和/或BCR轻链。

20. 根据权利要求1-19中任一项所述的方法,其中所述多于一种条形码化核酸分子或其产物的多于一个测序读段中的每一个包含(1)分子标记序列,和/或(2)所述核酸靶的子序列,所述方法还包括基于所述核酸靶文库成员的多于一个测序读段来确定所述一个或更多个单细胞中的每一个中所述核酸靶的拷贝数;并且任选地确定所述一个或更多个单细胞中的每一个中所述核酸靶的拷贝数包括基于与所述一种或更多种核酸靶文库成员或其产物关联的具有不同序列的第一分子标记、其互补物或其组合的数目来确定所述一个或更多个单细胞中的每一个中所述核酸靶的拷贝数。

21. 根据权利要求1-20中任一项所述的方法,其中所述核酸靶包括核酸分子,任选地所述核酸分子包括核糖核酸(RNA)、信使RNA(mRNA)、微RNA、小干扰RNA(siRNA)、RNA降解产物、包含多(A)尾的RNA或其任何组合,并且还任选地所述mRNA编码免疫受体。

22. 根据权利要求1-21中任一项所述的方法,其中所述细胞组分结合试剂特异性寡核苷酸包含第三通用序列。

23. 根据权利要求1-22中任一项所述的方法,其中产生所述测序文库包括:

使用能够与所述第一通用序列或其互补物杂交的引物和能够与所述第三通用序列或其互补物杂交的引物,扩增所述多于一种条形码化细胞组分结合试剂特异性寡核苷酸或其产物,以产生多于一种扩增的条形码化细胞组分结合试剂特异性寡核苷酸,

其中所述多于一种细胞组分靶文库成员包括所述多于一种扩增的条形码化细胞组分结合试剂特异性寡核苷酸或其产物,

任选地,扩增所述多于一种条形码化细胞组分结合试剂特异性寡核苷酸包括将测序引物的结合位点和/或测序衔接子、其互补序列和/或其部分的序列添加到所述多于一种条形码化细胞组分结合试剂特异性寡核苷酸,以及

还任选地,所述方法包括基于所述细胞组分靶文库成员的多于一个测序读段来确定所述一个或更多个单细胞中所述多于一种细胞组分靶中的至少一种细胞组分靶的拷贝数。

24. 根据权利要求1-23中任一项所述的方法,其中所述细胞组分靶文库成员的多于一个测序读段中的每一个包含(1)分子标记序列,和/或(2)所述独特标识符序列的至少一部分。

25. 根据权利要求1-24中任一项所述的方法,其中所述细胞组分结合试剂特异性寡核苷酸包含第二分子标记,任选地所述多于一种细胞组分结合试剂特异性寡核苷酸中的至少10种包含不同的第二分子标记序列,任选地:(i)至少两种细胞组分结合试剂特异性寡核苷酸的所述第二分子标记序列是不同的,并且其中所述至少两种细胞组分结合试剂特异性寡

核苷酸的所述独特标识符序列是相同的;和/或(ii)至少两种细胞组分结合试剂特异性寡核苷酸的所述第二分子标记序列是不同的,并且其中所述至少两种细胞组分结合试剂特异性寡核苷酸的所述独特标识符序列是不同的。

26. 根据权利要求3-25中任一项所述的方法,其中:(i)所述测序数据中与用于所述细胞组分结合试剂的所述独特标识符序列关联的独特第一分子标记序列的数目指示所述一个或多个单细胞中所述至少一种细胞组分靶的拷贝数,所述细胞组分结合试剂能够与所述至少一种细胞组分靶特异性结合;和/或(ii)所述测序数据中与用于所述细胞组分结合试剂的所述独特标识符序列关联的独特第二分子标记序列的数目指示所述一个或多个单细胞中所述至少一种细胞组分靶的拷贝数,所述细胞组分结合试剂能够与所述至少一种细胞组分靶特异性结合。

27. 根据权利要求1-26中任一项所述的方法,所述方法包括在使所述多于一种细胞组分结合试剂与和所述受体检测构建体关联的所述第一多于一个细胞接触之后,去除所述多于一种细胞组分结合试剂中未与和所述受体检测构建体关联的所述第一多于一个细胞接触的一种或更多种细胞组分结合试剂;并且任选地去除未与和所述受体检测构建体关联的所述第一多于一个细胞接触的所述一种或更多种细胞组分结合试剂包括:去除未与所述多于一种细胞组分靶中相应的至少一种细胞组分靶接触的所述一种或更多种细胞组分结合试剂。

28. 根据权利要求1-27中任一项所述的方法,其中所述靶结合区包含多(dT)区,并且其中所述细胞组分结合试剂特异性寡核苷酸包含多(dA)区;并且任选地所述细胞组分结合试剂特异性寡核苷酸包含与多(dA)区相邻的对齐序列;并且还任选地其中

(a)所述对齐序列包含鸟嘌呤、胞嘧啶、胸腺嘧啶、尿嘧啶或其组合;

(b)所述对齐序列包含多(dT)序列、多(dG)序列、多(dC)序列、多(dU)序列或其组合;和/或

(c)所述对齐序列位于所述多(dA)区的5'。

29. 根据权利要求1-28中任一项所述的方法,其中所述细胞组分靶包括糖类、脂质、蛋白、细胞外蛋白、细胞表面蛋白、细胞标志物、B细胞受体、T细胞受体、主要组织相容性复合体、肿瘤抗原、受体、细胞内蛋白或其任何组合,任选地所述细胞组分结合试剂包含抗体或其片段,任选所述细胞组分靶在细胞表面上。

30. 根据权利要求1-29中任一项所述的方法,其中使所述多于一种受体检测构建体与所述多于一个细胞接触以形成所述第一多于一个细胞包括去除未与一种或更多种受体检测构建体结合的一个或多个细胞;任选地去除未与一种或更多种受体检测构建体结合的一个或多个细胞包括选择与一种或更多种受体检测构建体结合的细胞;并且还任选地所述受体检测构建体包含一种或更多种另外的标记,其中选择与一种或更多种受体检测构建体结合的细胞包括流式细胞术和/或基于所述一种或更多种另外的标记的存在选择细胞,任选地所述一种或更多种另外的标记包括荧光标记,还任选地去除未与和所述受体检测构建体关联的所述第一多于一个细胞接触的一种或更多种受体结合试剂包括:去除未与所述受体接触的一种或更多种受体结合试剂。

31. 根据权利要求1-30中任一项所述的方法,其中所述受体结合试剂特异性寡核苷酸包含第二通用序列。

32. 根据权利要求1-31中任一项所述的方法,其中产生所述测序文库包括:

使用能够与所述第一通用序列或其互补物杂交的引物和能够与所述第二通用序列或其互补物杂交的引物,扩增所述多于一种条形码化受体结合试剂特异性寡核苷酸或其产物,以产生多于一种扩增的条形码化受体结合试剂特异性寡核苷酸,

其中所述多于一种受体文库成员包括所述多于一种扩增的条形码化受体结合试剂特异性寡核苷酸或其产物,

任选地,扩增所述多于一种条形码化受体结合试剂特异性寡核苷酸包括将测序引物的结合位点和/或测序衔接子、其互补序列和/或其部分的序列添加到所述多于一种条形码化受体结合试剂特异性寡核苷酸。

33. 根据权利要求1-32中任一项所述的方法,其中所述受体文库成员的多于一个测序读段中的每一个包含(1)分子标记序列,(2)所述独特受体标识符序列的至少一部分或(3)细胞标记序列,并且任选地每种独特细胞标记序列指示所述第二多于一个细胞中的单细胞。

34. 根据权利要求1-33中任一项所述的方法,所述方法还包括(i)基于所述受体文库成员的多于一个测序读段来确定所述一个或更多个单细胞中所述受体的拷贝数;(ii)基于所述受体文库成员的多于一个测序读段来确定与所述一个或更多个单细胞结合的所述受体结合试剂的身份;和/或(iii)基于所述受体文库成员的多于一个测序读段来确定所述一个或更多个单细胞中所述受体的身份。

35. 根据权利要求1-34中任一项所述的方法,其中所述受体结合试剂特异性寡核苷酸包含第三分子标记,任选地所述多于一种受体结合试剂特异性寡核苷酸中的至少10种包含不同的第三分子标记序列,任选地至少两种受体结合试剂特异性寡核苷酸的第三分子标记序列是不同的,并且其中所述至少两种受体结合试剂特异性寡核苷酸的独特受体标识符序列是相同的和/或是不同的。

36. 根据权利要求3-35中任一项所述的方法,其中:(i)所述测序数据中与用于所述受体结合试剂的所述独特受体标识符序列关联的独特第一分子标记序列的数目指示所述一个或更多个单细胞中所述受体的拷贝数,所述受体结合试剂能够与所述至少一种受体靶特异性结合;和/或(ii)所述测序数据中与用于所述受体结合试剂的所述独特受体标识符序列关联的独特第三分子标记序列的数目指示所述一个或更多个单细胞中所述受体的拷贝数,所述受体结合试剂能够与所述至少一种受体特异性结合。

37. 根据权利要求1-36中任一项所述的方法,其中与相同受体结合试剂关联的受体结合试剂特异性寡核苷酸(a)包含相同的独特受体标识符序列,并且其中与不同受体结合试剂关联的受体结合试剂特异性寡核苷酸包含不同的独特受体标识符序列;(b)包括DNA、RNA、锁核酸(LNA)、肽核酸(PNA)、LNA/PNA嵌合体、LNA/DNA嵌合体、PNA/DNA嵌合体或其任何组合;或(c)长度为约10个核苷酸至约100个核苷酸。

38. 根据权利要求1-37中任一项所述的方法,所述受体包括分化簇(CD)分子、免疫受体或两者;并且任选地所述免疫受体是T细胞受体(TCR)。

39. 根据权利要求1-38中任一项所述的方法,其中所述靶结合区包含多(dT)区,并且其中所述受体结合试剂特异性寡核苷酸包含多(dA)区;任选地所述受体结合试剂特异性寡核苷酸包含与所述多(dA)区相邻的对齐序列;并且还任选地

(a)所述对齐序列包含鸟嘌呤、胞嘧啶、胸腺嘧啶、尿嘧啶或其组合；

(b)所述对齐序列包含多(dT)序列、多(dG)序列、多(dC)序列、多(dU)序列或其组合；  
和/或

(c)所述对齐序列位于所述多(dA)区的5'。

40. 根据权利要求1-39中任一项所述的方法，其中所述多于一种受体检测构建体包括一种或更多种MHC多聚体，并且其中所述两种或更多种受体结合试剂包含两种或更多种MHC-肽复合物；任选地(a) MHC多聚体包含(a-b-P)<sub>n</sub>，其中n>1，其中多肽a和多肽b一起形成能够结合肽P的功能性MHC蛋白，并且(a-b-P)是当肽P与所述功能性MHC蛋白结合时形成的MHC-肽复合物；和/或任选地(b) MHC多聚体的MHC-肽复合物与一个或更多个多聚化结构域关联。

41. 根据权利要求40所述的方法，其中所述多于一种受体检测构建体包含2种MHC多聚体、3种MHC多聚体、4种MHC多聚体、5种MHC多聚体、6种MHC多聚体、7种MHC多聚体、8种MHC多聚体、9种MHC多聚体、10种MHC多聚体、11种MHC多聚体、12种MHC多聚体、13种MHC多聚体、14种MHC多聚体、15种MHC多聚体、16种MHC多聚体、17种MHC多聚体、18种MHC多聚体、19种MHC多聚体或20种MHC多聚体，任选地所述MHC多聚体的每种MHC-肽复合物的个体抗原肽P是相同或不同的。

42. 根据权利要求40-41中任一项所述的方法，其中包含(a-b-P)<sub>n</sub>的所述一种或更多种MHC多聚体的n的值是1<n≤1000。

43. 根据权利要求40-42中任一项所述的方法，其中所述MHC蛋白是MHC I类、MHC II类、MR1、CD1或MHC样分子，并且任选地所述抗原肽P选自由与MHC I类、MHC II类、MR1、CD1或MHC样分子结合的8-mer、9-mer、10-mer、11-mer和12-mer肽组成的组。

44. 根据权利要求40-43中任一项所述的方法，其中所述一个或更多个多聚化结构域包含一个或更多个多聚化结构域连接体分子。

45. 根据权利要求40-44中任一项所述的方法，其中多聚体MHC的每种MHC蛋白或MHC-肽复合物中的一个或更多个包含一个或更多个MHC连接体分子，任选所述一个或更多个MHC连接体分子是生物素。

46. 根据权利要求44-45中任一项所述的方法，其中所述一个或更多个多聚化结构域连接体分子包括一个或更多个链霉抗生物素蛋白和/或一个或更多个抗生物素蛋白或其任何衍生物，任选地所述一个或更多个链霉抗生物素蛋白包括一个或更多个四聚体链霉抗生物素蛋白变体或一个或更多个单体链霉抗生物素蛋白变体。

47. 根据权利要求40-46中任一项所述的方法，其中MHC多聚体的每种MHC蛋白或MHC-肽复合物中的一个或更多个(a)通过链霉抗生物素蛋白-生物素连接或抗生物素蛋白-生物素连接与一个或更多个多聚化结构域关联，(b)通过接头部分与一个或更多个多聚化结构域关联，(c)通过天然二聚化和/或蛋白-蛋白相互作用与一个或更多个多聚化结构域关联，任选地所述天然二聚化和/或蛋白-蛋白相互作用选自由以下组成的组：亮氨酸拉链例如AP-1的亮氨酸拉链结构域、Fos/Jun相互作用、基于酸性/碱性卷曲螺旋结构的相互作用(例如螺旋)、抗体/抗原相互作用、多核苷酸-多核苷酸相互作用例如DNA/DNA、DNA/PNA、DNA/RNA、PNA/PNA、LNA/DNA、合成分子-合成分子相互作用和蛋白-小分子相互作用、IgG二聚体蛋白、IgM多价蛋白、螯合物/金属离子结合螯合物、strep免疫球蛋白、抗体(单克隆、多克隆和重

组的)、抗体片段及其衍生物、hexa-his(金属螯合部分)、hexa-hat GST(谷胱甘肽S-转移酶)谷胱甘肽亲和力、钙调蛋白结合肽(CBP)、Strep-标签、纤维素结合结构域、麦芽糖结合蛋白、S-肽标签、壳多糖结合标签、免疫反应性表位、表位标签、E2标签、HA表位标签、Myc表位、FLAG表位、AU1表位和AU5表位、Glu-Glu表位、KT3表位、IRS表位、Btag表位、蛋白激酶-C表位、VSV表位、介导与包括糖类、脂质和蛋白的多种化合物结合的凝集素,例如Con A(直生刀豆(*Cannvalia ensiformis*))或WGA(麦胚凝集蛋白)和四联凝集素或蛋白A或蛋白G(抗体亲和力)。

48. 根据权利要求40-47中任一项所述的方法,其中所述一个或更多个多聚化结构域包含(i)一个或更多个支架;(ii)一个或更多个载体;(iii)至少一个支架和至少一个载体;和/或(iv)一个或更多个任选取代的有机分子。

49. 根据权利要求40-48中任一项所述的方法,其中所述一个或更多个多聚化结构域包含(a)一种或更多种生物细胞和/或细胞样结构;(b)一种或更多种膜,任选地所述一种或更多种膜包括脂质体或胶束;和/或(c)一种或更多种聚合物,任选地一种或更多种合成聚合物。

50. 根据权利要求49所述的方法,其中所述一种或更多种聚合物选自由多糖组成的组,任选地所述多糖包含一个或更多个右旋糖酐部分,任选地所述一个或更多个右旋糖酐部分包含一个或更多个具有约1,000Da至1,000,000Da的分子量的右旋糖酐,并且是线性的和/或支链的。

51. 根据权利要求50所述的方法,其中所述一个或更多个右旋糖酐部分(i)与一个或更多个MHC肽复合物共价衔接;(ii)与一个或更多个MHC肽复合物非共价衔接;(iii)被修饰;(iv)包含一个或更多个氨基-右旋糖酐,任选地所述一个或更多个氨基-右旋糖酐用二乙烯基砜修饰;(v)包含一个或更多个具有约1,000Da至1,000,000Da的分子量的右旋糖酐;和/或(vi)是线性的和/或支链的,或其任何组合。

52. 根据权利要求40-51中任一项所述的方法,其中所述一个或更多个多聚化结构域包含以下中的一种或更多种:

(a) 一种或更多种选自由IgG结构域、卷曲螺旋多肽结构、DNA双链体、核酸双链体、PNA-PNA、PNA-DNA和DNA-RNA组成的组的实体;

(b) 抗体,任选地所述抗体选自由以下组成的组:多克隆抗体、单克隆抗体、IgA、IgG、IgM、IgD、IgE、IgG1、IgG2、IgG3、IgG4、IgA1、IgA2、IgM1、IgM2、人源化抗体、人源化单克隆抗体、嵌合抗体、小鼠抗体、大鼠抗体、兔抗体、人类抗体、骆驼抗体、绵羊抗体、工程化人类抗体、表位聚焦抗体、激动剂抗体、拮抗剂抗体、中和抗体、天然存在的抗体、分离的抗体、单价抗体、双特异性抗体、三特异性抗体、多特异性抗体、异源缀合抗体、免疫缀合物、免疫脂质体、标记的抗体、抗体片段、结构域抗体、纳米抗体、微抗体、大抗体、双体和融合抗体;

(c) 一种或更多种有机小支架分子或有机小分子,任选地所述一种或更多种有机小分子包括一种或更多种类固醇、一种或更多种肽、一种或更多种芳族有机分子,还任选地所述一种或更多种芳族有机分子包含一种或更多种双环结构、一种或更多种多环结构或一种或更多种单环结构;任选地所述一种或更多种单环结构包括一种或更多种任选地官能化或取代的苯环;

(d) (i) 能够聚合的单体分子;(ii) 生物聚合物,诸如一种或更多种蛋白;(iii) 小分子

支架；(iv)超分子结构，诸如一种或更多种纳米团簇；和/或(v)蛋白复合物；

(e)一种或更多种珠，任选地所述一种或更多种珠选自以下组成的组：带有亲电基团例如二乙烯基砜活化的多糖的珠、已经用甲苯磺酰基活化的酯官能化的聚苯乙烯珠、用甲苯磺酰基活化的酯官能化的磁性聚苯乙烯珠，以及其中MHC复合体已经通过所述MHC复合体中包含的亲核体与珠的亲电体的反应而共价固定至这些珠的珠。

53. 根据权利要求40-52中任一项所述的方法，其中所述多聚化结构域包含一种或更多种选自以下组成的组的化合物：琼脂糖珠、琼脂糖凝胶珠、树脂珠、玻璃珠、多孔玻璃珠、涂覆疏水聚合物的玻璃颗粒、壳聚糖涂覆珠、SH珠、胶乳珠、球形胶乳珠、等位基因型珠、SPA珠、基于PEG的树脂、PEG涂覆珠、PEG包封珠、聚苯乙烯珠、磁性聚苯乙烯珠、谷胱甘肽琼脂糖珠、磁珠、顺磁珠、蛋白A和/或蛋白G琼脂糖凝胶珠、活化羧酸珠、宏观珠、微观珠、不溶性树脂珠、基于二氧化硅的树脂、纤维素树脂、交联琼脂糖珠、聚苯乙烯珠、交联聚丙烯酰胺树脂、具有铁芯的珠、金属珠、dynabead、用NHS活化的聚甲基丙烯酸甲酯珠、链霉抗生物素蛋白-琼脂糖珠、聚丙烯、聚乙烯、右旋糖酐、尼龙、淀粉酶、天然纤维素和改性纤维素、硝化纤维素、聚丙烯酰胺、辉长岩、磁铁矿、聚合物、寡聚物、非重复部分、聚乙二醇(PEG)、单甲氧基-PEG、单-(C<sub>1</sub>-C<sub>10</sub>)烷氧基-PEG、芳氧基-PEG、聚(N-乙基吡咯烷酮)PEG、三氟乙磺酰氯单甲氧基PEG、PEG丙醛、双琥珀酰亚胺基碳酸PEG、二乙烯基苯交联聚苯乙烯珠、丙二醇均聚物、聚环氧丙烷/环氧乙烷共聚物、聚氧乙基化多元醇(例如，甘油)、聚乙烯醇、右旋糖酐、氨基右旋糖酐、基于糖类的聚合物、交联右旋糖酐珠、多糖珠、聚氨基甲酸酯珠、二乙烯基砜活化多糖、已经用甲苯磺酰基活化的酯官能化的聚苯乙烯珠、用甲苯磺酰基活化的酯官能化的磁性聚苯乙烯珠、链霉抗生物素蛋白珠、链霉抗生物素蛋白单体涂覆的珠、链霉抗生物素蛋白四聚体涂覆的珠、链霉抗生物素蛋白涂覆的Compel磁珠、抗生物素蛋白涂覆的珠、dextramer涂覆的珠、二乙烯基砜活化的右旋糖酐、羧酸修饰的珠、胺修饰的珠、抗体涂覆的珠、纤维素珠、接枝共聚珠、聚丙烯酰胺珠、任选地与N-N'-双丙烯酰基乙二胺交联的二甲基丙烯酰胺珠、中空纤维膜、荧光珠、胶原-琼脂糖珠、明胶珠、胶原-明胶珠、胶原-纤连蛋白-明胶珠、胶原珠、壳聚糖珠、胶原-壳聚糖珠、基于蛋白的珠、水凝胶珠、半纤维素、烷基纤维素、羟烷基纤维素、羧甲基纤维素、磺基乙基纤维素、淀粉、木聚糖、支链淀粉、软骨素、透明质酸、肝素、瓜尔胶、黄原胶、甘露聚糖、半乳甘露聚糖、壳多糖和壳聚糖。

54. 根据权利要求40-53中任一项所述的方法，其中所述一个或更多个多聚化结构域包含以下中的一种或更多种：(a)二聚化结构域、三聚化结构域、四聚化结构域、五聚化结构域或六聚化结构域；(b)衔接一个或更多个支架的聚合物结构，任选地所述聚合物结构包含多糖，并且还任选地所述多糖包含一个或更多个右旋糖酐部分；(c)聚酰胺、聚乙二醇、多糖、琼脂糖凝胶或其任何组合；以及(d)羧甲基右旋糖酐、右旋糖酐聚醛、羧甲基右旋糖酐内酯、环糊精或其任何组合。

55. 根据权利要求40-54中任一项所述的方法，其中所述一种或更多种MHC多聚体还包含一个或更多个选自以下组成的组的支架、载体和/或接头：链霉抗生物素蛋白(SA)和抗生物素蛋白及其衍生物、生物素、免疫球蛋白、抗体(单克隆、多克隆和重组的)、抗体片段及其衍生物、AP-1的亮氨酸拉链结构域(jun和fos)、hexa-his(金属螯合部分)、hexa-hat GST(谷胱甘肽S-转移酶)谷胱甘肽亲和力、钙调蛋白结合肽(CBP)、Strep-标签、纤维素结合结构域、麦芽糖结合蛋白、S-肽标签、壳多糖结合标签、免疫反应性表位、表位标签、E2标签、HA

表位标签、Myc表位、FLAG表位、AU1和AU5表位、Glu-Glu表位、KT3表位、IRS表位、Btag表位、蛋白激酶-C表位、VSV表位、介导与包括糖类、脂质和蛋白的多种化合物结合的凝集素,例如Con A(直生刀豆)或WGA(麦胚凝集蛋白)和四联凝集素或蛋白A或蛋白G(抗体亲和力)。

56. 根据权利要求40-55中任一项所述的方法,其中所述MHC多聚体包含以下中的一种或更多种:(a)由多聚化结构域连接部分连接的多于一个相同或不同的多聚化结构域,(b)连接至第二多聚化结构域的第一多聚化结构域,(c)一个或更多个标记,任选地所述一个或更多个标记包括所述受体结合试剂特异性寡核苷酸。

57. 根据权利要求56-56中任一项所述的方法,其中所述一个或更多个标记包括所述受体结合试剂特异性寡核苷酸和一种或更多种另外的标记,任选地所述一种或更多种另外的标记选自包括以下的组:肽标记、荧光团标记、重金属标记、同位素标记、放射标记、放射性核素、稳定同位素、同位素链和单原子、化学发光标记、生物发光标记、放射性标记、酶标记、DNA荧光染色剂、镧系元素、离子载体、与特定离子结合的螯合化合物或其任何组合。

58. 根据权利要求56-57中任一项所述的方法,其中所述一个或更多个标记:

(a) 包括共价附接的标记和/或非共价附接的标记;

(b) 与以下附接:(i)所述MHC多肽a;(ii)所述MHC多肽b;(iii)所述肽P;(iv)所述一个或更多个多聚化结构域;和/或(v)  $(a-b-P)_n$ ;

(c) 通过链霉抗生物素蛋白-生物素连接附接至  $(a-b-P)_n$ ;

(d) 是荧光团标记,任选地所述荧光团标记选自包括以下的组:异硫氰酸荧光素、罗丹明、藻红蛋白、藻蓝蛋白、别藻蓝蛋白、邻苯二甲醛、荧光胺;2-(4'-马来酰亚胺基苯胺基)萘-6-磺酸钠盐;5-(((2-碘乙酰基)氨基)乙基)氨基)萘-1-磺酸;苾-1-丁酸;AlexaFluor 350(7-氨基-6-磺酸-4-甲基香豆素-3-乙酸);AMCA(7-氨基-4-甲基香豆素-3-乙酸);7-羟基-4-甲基香豆素-3-乙酸;Marina Blue(6,8-二氟-7-羟基-4-甲基香豆素-3-乙酸);7-二甲基氨基-香豆素-4-乙酸;荧光胺-N-丁胺加合物;7-羟基-香豆素-3-羧酸;CascadeBlue(苾-三磺酸乙酰叠氮化物);Cascade Yellow;Pacific Blue(6,8二氟-7-羟基香豆素-3-羧酸);7-二乙基氨基-香豆素-3-羧酸;N-(((4-叠氮苯甲酰基)氨基)乙基)-4-氨基-3,6-二硫基-1,8-萘酰亚胺二钾盐;Alexa Fluor430;3-花十二酸;8-羟基苾-1,3,6-三磺酸三钠盐;12-(N-(7-硝基苯-2-氧杂-1,3-二唑-4-基)氨基)十二酸;N,N'-二甲基-N-(碘乙酰基)-N'-(7-硝基苯-2-氧杂-1,3-二唑-4-基)乙二胺;Oregon Green 488(二氟羧基荧光素);5-碘乙酰胺荧光素;碘化丙锭-DNA加合物;羧基荧光素、Fluor染料、Pacific Blue<sup>TM</sup>、Pacific Orange<sup>TM</sup>、Cascade Yellow<sup>TM</sup>; **AlexaFluor®** (AF), AF405、AF488、AF500、AF514、AF532、AF546、AF555、AF568、AF594、AF610、AF633、AF635、AF647、AF680、AF700、AF710、AF750、AF800;基于量子点的染料, **QDot®** 纳米晶体(Invitrogen, Molecular Probs)、**Qdot®**525、**Qdot®**565、**Qdot®**585、**Qdot®**605、**Qdot®**655、**Qdot®**705、**Qdot®**800;DyLight<sup>TM</sup>染料(Pierce) (DL), DL549、DL649、DL680、DL800;荧光素(Flu)或其任何衍生物,诸如FITC;Cy-染料、Cy2、Cy3、Cy3.5、Cy5、Cy5.5、Cy7;荧光蛋白,RPE、PerCp、APC、绿色荧光蛋白;GFP和GFP衍生的突变体蛋白;BFP、CFP、YFP、DsRed、T1、Dimer2、mRFP1、MBanana、mOrange、dTomato、tdTomato、mTangerine、mStrawberry、mCherry;串联染料,RPE-Cy5、RPE-Cy5.5、RPE-Cy7、RPE-**AlexaFluor®** 串联缀合物;RPE-Alexa610、RPE-TxRed、APC-Alexa600、APC-Alexa610、

APC-Alexa750、APC-Cy5和APC-Cy5.5;和/或

(e)能够吸收光,任选地吸收发色团和/或染料。

59.根据权利要求40-58中任一项所述的方法,其中:

(i)P被化学修饰;

(ii)P被聚乙二醇化、磷酸化和/或糖基化;

(iii)所述肽P的氨基酸残基中的一个被另一个氨基酸取代;

(iv)a和b两者都是全长肽;

(v)a是全长肽;

(vi)b是全长肽;

(vii)a被截短;

(viii)b被截短;

(ix)a和b两者都被截短;

(x)a被共价连接至b;

(xi)a被共价连接至P;

(xii)b被共价连接至P;

(xiii)a、b和P都被共价连接;

(xiv)a被非共价连接至b;

(xv)a被非共价连接至P;

(xvi)b被非共价连接至P;

(xvii)a、b和P都被非共价连接;

(xviii)a不被包含在(a-b-P)复合物中;

(xix)b不被包含在(a-b-P)复合物中;和/或

(xx)P不被包含在(a-b-P)复合物中。

60.根据权利要求40-59中任一项所述的方法,所述方法还包括:(i)基于所述受体文库成员的多于一个测序读段,将所述一个或更多个单细胞中的T细胞受体(TCR)与所述肽P关联;和/或(ii)测量对于肽P特异性的T细胞的存在、频率、数目、活性和/或状态,从而检测抗原特异性T细胞应答。

61.根据权利要求1-60中任一项所述的方法,所述方法包括将(i)条形码化核酸分子、(ii)条形码化受体结合试剂特异性寡核苷酸和(iii)条形码化细胞组分结合试剂特异性寡核苷酸中的一种或更多种与(i)条形码化核酸分子、(ii)条形码化受体结合试剂特异性寡核苷酸和(iii)条形码化细胞组分结合试剂特异性寡核苷酸中的一种或更多种物理分离,任选地物理分离包括使用一种或更多种尺寸选择试剂。

62.根据权利要求1-61中任一项所述的方法,其中所述多于一种条形码化受体结合试剂特异性寡核苷酸和所述多于一种条形码化细胞组分结合试剂特异性寡核苷酸被分别扩增。

63.根据权利要求31-62中任一项所述的方法,其中所述第二通用序列和所述第三通用序列是不同的,任选地所述第二通用序列与所述第三通用序列小于约85%相同,任选地所述第二通用序列包含与SEQ ID NO:1(GGAGGGAGGTTAGCGAAGGT)至少约85%相同的序列。

64.根据权利要求32-63中任一项所述的方法,其中扩增所述多于一种条形码化受体结

合试剂特异性寡核苷酸或其产物不包括扩增所述多于一种条形码化细胞组分结合试剂特异性寡核苷酸或其产物。

65. 根据权利要求1-64中任一项所述的方法,其中产生所述测序文库包括产生包含(i)核酸靶文库成员、(ii)细胞组分靶文库成员和/或(iii)受体文库成员的测序混合物,任选地产生测序混合物包括以预定比例混合(i)核酸靶文库成员、(ii)细胞组分靶文库成员和/或(iii)受体文库成员,还任选地在产生所述测序混合物之前,所述(i)核酸靶文库成员、(ii)细胞组分靶文库成员和/或(iii)受体文库成员在物理上彼此分离。

66. 根据权利要求65所述的方法,其中所述测序混合物包含预定比例的(i)核酸靶文库成员、(ii)细胞组分靶文库成员和/或(iii)受体文库成员;任选地细胞组分靶文库成员与受体文库成员的所述预定比例为约1:1至10000:1,任选地细胞组分靶文库成员与受体文库成员的所述预定比例被配置为实现为约1:1至10000:1的细胞组分靶文库成员的测序读段与受体文库成员的测序读段的比例。

67. 根据权利要求1-66中任一项所述的方法,其中细胞组分靶文库成员的测序读段与受体文库成员的测序读段的比例为约1:1至10000:1。

68. 根据权利要求1-67中任一项所述的方法,其中细胞组分靶文库成员的测序读段与受体文库成员的测序读段的比例与其中包括以下的方法相比为至少约2倍低:(i)所述多于一种条形码化受体结合试剂特异性寡核苷酸和所述多于一种条形码化细胞组分结合试剂特异性寡核苷酸被一起扩增;(ii)所述第二通用序列和所述第三通用序列相同;和/或(iii)所述测序混合物不包含细胞组分靶文库成员与受体文库成员的预定比例。

69. 一种试剂盒,包含:

多于一种受体检测构建体,其中受体检测构建体包含两种或更多种受体结合试剂,其中受体结合试剂能够与受体特异性结合,并且其中所述受体检测构建体中的每一种包含受体结合试剂特异性寡核苷酸,所述受体结合试剂特异性寡核苷酸包含用于所述受体结合试剂的独特受体标识符序列;以及

多于一种细胞组分结合试剂,其中所述多于一种细胞组分结合试剂中的每一种包含细胞组分结合试剂特异性寡核苷酸,所述细胞组分结合试剂特异性寡核苷酸包含用于所述细胞组分结合试剂的独特标识符序列,并且其中所述细胞组分结合试剂能够与细胞组分靶特异性结合,

其中所述受体结合试剂特异性寡核苷酸包含第二通用序列,其中所述细胞组分结合试剂特异性寡核苷酸包含第三通用序列,其中所述第二通用序列和所述第三通用序列是不同的,任选地所述第二通用序列与所述第三通用序列小于约85%相同。

70. 根据权利要求69所述的试剂盒,所述试剂盒包含多于一种寡核苷酸条形码,其中所述多于一种寡核苷酸条形码中的每一种包含第一通用序列、分子标记和靶结合区,并且其中所述多于一种寡核苷酸条形码中的至少10种包含不同的分子标记序列;并且任选地所述多于一种寡核苷酸条形码与固体支持物关联,任选地所述靶结合区包含基因特异性序列、寡(dT)序列、随机多聚体或其任何组合。

71. 根据权利要求69所述的试剂盒,所述试剂盒包含逆转录酶、缺乏5'至3'外切核酸酶活性和3'至5'外切核酸酶活性中的至少一种的DNA聚合酶、缓冲液、盒、用于逆转录反应和/或扩增反应的一种或更多种试剂,或其组合,任选地所述逆转录酶包括病毒逆转录酶,并且

还任选地所述病毒逆转录酶是鼠白血病病毒(MLV)逆转录酶或Moloney白血病病毒(MMLV)逆转录酶,任选地所述DNA聚合酶包括Klenow片段。

72. 根据权利要求70-71中任一项所述的方法,其中所述多于一种寡核苷酸条形码各自包含细胞标记,任选地所述多于一种寡核苷酸条形码的每一种细胞标记包含至少6个核苷酸,任选地所述多于一种寡核苷酸条形码中与同一固体支持物关联的寡核苷酸条形码包含相同的细胞标记,任选地所述多于一种寡核苷酸条形码中与不同固体支持物关联的寡核苷酸条形码包含不同的细胞标记。

73. 根据权利要求70-72中任一项所述的试剂盒,其中所述固体支持物包括合成颗粒、平坦表面或其组合,任选地所述多于一种寡核苷酸条形码中的至少一种寡核苷酸条形码被固定或部分固定在所述合成颗粒上,或者所述多于一种寡核苷酸条形码中的至少一种寡核苷酸条形码被包封或部分包封在所述合成颗粒中。

74. 根据权利要求69-73中任一项所述的试剂盒,所述试剂盒还包含一种或更多种引物,所述一种或更多种引物包含所述第一通用序列、所述第二通用序列和/或所述第三通用序列,任选地所述第二通用序列包含与SEQ ID NO:1 (GGAGGGAGGTTAGCGAAGGT)至少约85%相同的序列。

75. 根据权利要求69-74中任一项所述的试剂盒,其中所述细胞组分靶:

(a) 包括蛋白靶,任选地所述细胞组分结合试剂包含抗体或其片段;

(b) 包括糖类、脂质、蛋白、细胞外蛋白、细胞表面蛋白、细胞标志物、B细胞受体、T细胞受体、主要组织相容性复合体、肿瘤抗原、受体、细胞内蛋白或其任何组合;和/或

(c) 在细胞表面上。

76. 根据权利要求69-75中任一项所述的试剂盒,其中所述受体结合试剂特异性寡核苷酸包含第三分子标记,和/或所述细胞组分结合试剂特异性寡核苷酸包含第二分子标记。

77. 根据权利要求69-76中任一项所述的试剂盒,所述受体包括(a) 分化簇(CD)分子;或(b) 免疫受体,任选地所述免疫受体是T细胞受体(TCR)。

78. 根据权利要求69-77中任一项所述的试剂盒,其中所述靶结合区包含多(dT)区,并且其中所述受体结合试剂特异性寡核苷酸和/或所述细胞组分结合试剂特异性寡核苷酸包含多(dA)区。

79. 根据权利要求69-78中任一项所述的试剂盒,其中所述多于一种受体检测构建体包括一种或更多种MHC多聚体,并且其中所述两种或更多种受体结合试剂包含两种或更多种MHC-肽复合物,任选地MHC多聚体包含 $(a-b-P)_n$ ,其中 $n>1$ ,其中多肽a和多肽b一起形成能够结合肽P的功能性MHC蛋白,并且 $(a-b-P)$ 是当肽P与所述功能性MHC蛋白结合时形成的MHC-肽复合物。

80. 根据权利要求79所述的试剂盒,其中MHC多聚体的每种MHC-肽复合物与一个或更多个多聚化结构域关联。

81. 根据权利要求79-80中任一项所述的试剂盒,其中所述多于一种受体检测构建体包括(a) 2种MHC多聚体至20种MHC多聚体,任选地所述MHC多聚体的每种MHC-肽复合物的个体抗原肽P是相同或不同的;和/或(b) 一种或更多种阴性对照MHC多聚体;任选地所述一种或更多种阴性对照MHC多聚体其中每种MHC多聚体包含阴性对照肽P,并且还任选地所述阴性对照肽P是无义肽。

82. 根据权利要求79-81中任一项所述的试剂盒,其中包含(a-b-P)<sub>n</sub>的所述一种或更多种MHC多聚体的n的值是1<n≤1000,任选地n的值在2-1000之间。

83. 根据权利要求79-82中任一项所述的试剂盒,其中所述MHC蛋白是MHC I类、MHC II类、MR1、CD1或MHC样分子,并且任选地所述抗原肽P选自由与MHC I类、MHC II类、MR1、CD1或MHC样分子结合的8-mer、9-mer、10-mer、11-mer和12-mer肽组成的组。

84. 根据权利要求79-83中任一项所述的试剂盒,其中所述一个或更多个多聚化结构域包含(a)一个或更多个多聚化结构域连接体分子;(b)一个或更多个MHC连接体分子;(c)一个或更多个链霉抗生物素蛋白和/或一个或更多个抗生物素蛋白或其任何衍生物,以及任选地一个或更多个链霉抗生物素蛋白包括一个或更多个四聚体链霉抗生物素蛋白变体或一个或更多个单体链霉抗生物素蛋白变体。

85. 根据权利要求79-84中任一项所述的试剂盒,其中MHC多聚体的每种MHC蛋白或MHC-肽复合物中的一个或更多个通过链霉抗生物素蛋白-生物素连接或抗生物素蛋白-生物素连接与一个或更多个多聚化结构域关联,或通过接头部分与一个或更多个多聚化结构域关联,任选地所述一个或更多个MHC连接体分子是生物素。

86. 根据权利要求79-85中任一项所述的试剂盒,其中MHC多聚体的每种MHC蛋白或MHC-肽复合物中的一个或更多个通过天然二聚化和/或蛋白-蛋白相互作用与一个或更多个多聚化结构域关联,任选地所述天然二聚化和/或蛋白-蛋白相互作用选自由以下组成的组:亮氨酸拉链例如AP-1的亮氨酸拉链结构域、Fos/Jun相互作用、基于酸性/碱性卷曲螺旋结构的相互作用(例如螺旋)、抗体/抗原相互作用、多核苷酸-多核苷酸相互作用例如DNA/DNA、DNA/PNA、DNA/RNA、PNA/PNA、LNA/DNA、合成分子-合成分子相互作用和蛋白-小分子相互作用、IgG二聚体蛋白、IgM多价蛋白、螯合物/金属离子结合螯合物、strep免疫球蛋白、抗体(单克隆、多克隆和重组的)、抗体片段及其衍生物、hexa-his(金属螯合部分)、hexa-hat GST(谷胱甘肽S-转移酶)谷胱甘肽亲和力、钙调蛋白结合肽(CBP)、Strep-标签、纤维素结合结构域、麦芽糖结合蛋白、S-肽标签、壳多糖结合标签、免疫反应性表位、表位标签、E2标签、HA表位标签、Myc表位、FLAG表位、AU1表位和AU5表位、Glu-Glu表位、KT3表位、IRS表位、Btag表位、蛋白激酶-C表位、VSV表位、介导与包括糖类、脂质和蛋白的多种化合物结合的凝集素,例如Con A(直生刀豆(*Cannvalia ensiformis*))或WGA(麦胚凝集蛋白)和四联凝集素或蛋白A或蛋白G(抗体亲和力)。

87. 根据权利要求79-86中任一项所述的试剂盒,其中所述一个或更多个多聚化结构域包含(i)一个或更多个支架;(ii)一个或更多个载体;(iii)至少一个支架和至少一个载体;和/或(iv)一个或更多个任选取代的有机分子。

88. 根据权利要求79-87中任一项所述的试剂盒,其中所述一个或更多个多聚化结构域包含(a)一种或更多种生物细胞和/或细胞样结构;(b)一种或更多种膜,任选地所述一种或更多种膜包括脂质体或胶束;(c)一种或更多种聚合物,任选地一种或更多种合成聚合物,并且还任选地所述一种或更多种合成聚合物包括PNA、聚酰胺、PEG或其任何组合。

89. 根据权利要求88所述的试剂盒,其中所述一种或更多种聚合物选自由多糖组成的组,并且任选地所述多糖包含一个或更多个右旋糖酐部分。

90. 根据权利要求89所述的试剂盒,其中所述一个或更多个右旋糖酐部分(i)与一个或更多个MHC肽复合物共价衔接;(ii)与一个或更多个MHC肽复合物非共价衔接;(iii)被修

饰；(iv) 包含一个或更多个氨基-右旋糖酐，任选地所述一个或更多个氨基-右旋糖酐用二乙烯基砜修饰；(v) 包含一个或更多个具有约1,000Da至1,000,000Da的分子量的右旋糖酐；和/或(vi) 是线性的和/或支链的，或其任何组合。

91. 根据权利要求79-90中任一项所述的试剂盒，其中所述一个或更多个多聚化结构域包含以下中的一种或更多种：

(a) 一种或更多种选自IgG结构域、卷曲螺旋多肽结构、DNA双链体、核酸双链体、PNA-PNA、PNA-DNA和DNA-RNA组成的组的实体；

(b) 抗体，任选地所述抗体选自以下组成的组：多克隆抗体、单克隆抗体、IgA、IgG、IgM、IgD、IgE、IgG1、IgG2、IgG3、IgG4、IgA1、IgA2、IgM1、IgM2、人源化抗体、人源化单克隆抗体、嵌合抗体、小鼠抗体、大鼠抗体、兔抗体、人类抗体、骆驼抗体、绵羊抗体、工程化人类抗体、表位聚焦抗体、激动剂抗体、拮抗剂抗体、中和抗体、天然存在的抗体、分离的抗体、单价抗体、双特异性抗体、三特异性抗体、多特异性抗体、异源缀合抗体、免疫缀合物、免疫脂质体、标记的抗体、抗体片段、结构域抗体、纳米抗体、微抗体、大抗体、双体和融合抗体；以及

(c) 一种或更多种有机小支架分子或有机小分子，任选地所述一种或更多种有机小分子包括一种或更多种类固醇、一种或更多种肽、一种或更多种芳族有机分子，还任选地所述一种或更多种芳族有机分子包含一种或更多种双环结构、一种或更多种多环结构或一种或更多种单环结构；任选地所述一种或更多种单环结构包括一种或更多种任选地官能化或取代的苯环。

92. 根据权利要求79-91中任一项所述的试剂盒，其中所述一个或更多个多聚化结构域包含一种或更多种：(i) 能够聚合的单体分子；(ii) 生物聚合物，诸如一种或更多种蛋白；(iii) 小分子支架；(iv) 超分子结构，诸如一种或更多种纳米团簇；和/或(v) 蛋白复合物。

93. 根据权利要求79-92中任一项所述的试剂盒，其中所述多聚化结构域包含一种或更多种选自以下组成的组的化合物：琼脂糖珠、琼脂糖凝胶珠、树脂珠、玻璃珠、多孔玻璃珠、涂覆疏水聚合物的玻璃颗粒、壳聚糖涂覆珠、SH珠、胶乳珠、球形胶乳珠、等位基因型珠、SPA珠、基于PEG的树脂、PEG涂覆珠、PEG包封珠、聚苯乙烯珠、磁性聚苯乙烯珠、谷胱甘肽琼脂糖珠、磁珠、顺磁珠、蛋白A和/或蛋白G琼脂糖凝胶珠、活化羧酸珠、宏观珠、微观珠、不溶性树脂珠、基于二氧化硅的树脂、纤维素树脂、交联琼脂糖珠、聚苯乙烯珠、交联聚丙烯酰胺树脂、具有铁芯的珠、金属珠、dynabead、用NHS活化的聚甲基丙烯酸甲酯珠、链霉抗生物素蛋白-琼脂糖珠、聚丙烯、聚乙烯、右旋糖酐、尼龙、淀粉酶、天然纤维素和改性纤维素、硝化纤维素、聚丙烯酰胺、辉长岩、磁铁矿、聚合物、寡聚物、非重复部分、聚乙二醇(PEG)、单甲氧基-PEG、单-(C<sub>1</sub>-C<sub>10</sub>) 烷氧基-PEG、芳氧基-PEG、聚(N-乙基吡咯烷酮)PEG、三氟乙磺酰氯单甲氧基PEG、PEG丙醛、双琥珀酰亚胺基碳酸PEG、二乙烯基苯交联聚苯乙烯珠、丙二醇均聚物、聚环氧丙烷/环氧乙烷共聚物、聚氧乙基化多元醇(例如，甘油)、聚乙烯醇、右旋糖酐、氨基右旋糖酐、基于糖类的聚合物、交联右旋糖酐珠、多糖珠、聚氨基甲酸酯珠、二乙烯基砜活化多糖、已经用甲苯磺酰基活化的酯官能化的聚苯乙烯珠、用甲苯磺酰基活化的酯官能化的磁性聚苯乙烯珠、链霉抗生物素蛋白珠、链霉抗生物素蛋白单体涂覆的珠、链霉抗生物素蛋白四聚体涂覆的珠、链霉抗生物素蛋白涂覆的Compel磁珠、抗生物素蛋白涂覆的珠、dextramer涂覆的珠、二乙烯基砜活化的右旋糖酐、羧酸修饰的珠、胺修饰的珠、抗体涂覆的珠、纤维素珠、接枝共聚珠、聚丙烯酰胺珠、任选地与N-N'-双丙烯酰基乙二胺交联的二甲基

丙烯酸胺珠、中空纤维膜、荧光珠、胶原-琼脂糖珠、明胶珠、胶原-明胶珠、胶原-纤连蛋白-明胶珠、胶原珠、壳聚糖珠、胶原-壳聚糖珠、基于蛋白的珠、水凝胶珠、半纤维素、烷基纤维素、羟烷基纤维素、羧甲基纤维素、磺基乙基纤维素、淀粉、木聚糖、支链淀粉、软骨素、透明质酸、肝素、瓜尔胶、黄原胶、甘露聚糖、半乳甘露聚糖、壳多糖和壳聚糖。

94. 根据权利要求79-93中任一项所述的试剂盒,其中所述一个或更多多个多聚化结构域包含以下中的一种或更多种:(a)二聚化结构域、三聚化结构域、四聚化结构域、五聚化结构域或六聚化结构域;(b) 附接一个或更多个支架的聚合物结构,任选地所述聚合物结构包含多糖;并且任选地所述多糖包含一个或更多个右旋糖酐部分;(c) 聚酰胺、聚乙二醇、多糖、琼脂糖凝胶或其任何组合;以及(d) 羧甲基右旋糖酐、右旋糖酐聚醛、羧甲基右旋糖酐内酯、环糊精或其任何组合。

95. 根据权利要求79-94中任一项所述的试剂盒,其中所述一种或更多种MHC多聚体还包含一个或更多个选自由以下组成的组的支架、载体和/或接头:链霉抗生物素蛋白(SA)和抗生物素蛋白及其衍生物、生物素、免疫球蛋白、抗体(单克隆、多克隆和重组的)、抗体片段及其衍生物、AP-1的亮氨酸拉链结构域(jun和fos)、hexa-his(金属螯合部分)、hexa-hat GST(谷胱甘肽S-转移酶)谷胱甘肽亲和力、钙调蛋白结合肽(CBP)、Strep-标签、纤维素结合结构域、麦芽糖结合蛋白、S-肽标签、壳多糖结合标签、免疫反应性表位、表位标签、E2标签、HA表位标签、Myc表位、FLAG表位、AU1和AU5表位、Glu-Glu表位、KT3表位、IRS表位、Btag表位、蛋白激酶-C表位、VSV表位、介导与包括糖类、脂质和蛋白的多种化合物结合的凝集素,例如Con A(直生刀豆)或WGA(麦胚凝集蛋白)和四联凝集素或蛋白A或蛋白G(抗体亲和力)。

96. 根据权利要求79-95中任一项所述的试剂盒,其中所述MHC多聚体包含(a)由多聚化结构域连接部分连接的多于一个相同或不同的多聚化结构域;(b) 连接至第二多聚化结构域的第一多聚化结构域;(c) 一个或更多个标记,或其任何组合。

97. 根据权利要求96所述的试剂盒,其中所述一个或更多个标记包括

(a) 所述受体结合试剂特异性寡核苷酸;

(b) 包含所述受体结合试剂特异性寡核苷酸和一种或更多种另外的标记,任选地所述一种或更多种另外的标记选自包括以下的组:肽标记、荧光团标记、重金属标记、同位素标记、放射标记、放射性核素、稳定同位素、同位素链和单原子、化学发光标记、生物发光标记、放射性标记、酶标记、DNA荧光染色剂、镧系元素、离子载体、与特定离子结合的螯合化合物或其任何组合;和/或

(c) 共价附接的标记和/或非共价附接的标记。

98. 根据权利要求96-97中任一项所述的试剂盒,其中所述一个或更多个标记与以下附接:(i) 所述MHC多肽a;(ii) 所述MHC多肽b;(iii) 所述肽P;(iv) 所述一个或更多多个多聚化结构域;和/或(v)  $(a-b-P)_n$ ;并且任选地一个或更多个标记通过链霉抗生物素蛋白-生物素连接附接至 $(a-b-P)_n$ 。

99. 根据权利要求79-98中任一项所述的试剂盒,其中:

(i) P被化学修饰;

(ii) P被聚乙二醇化、磷酸化和/或糖基化;

(iii) 所述肽P的氨基酸残基中的一个被另一个氨基酸取代;(iv) a和b两者都是全长肽;

- (v) a是全长肽；
- (vi) b是全长肽；
- (vii) a被截短；
- (viii) b被截短；
- (ix) a和b两者都被截短；
- (x) a被共价连接至b；
- (xi) a被共价连接至P；
- (xii) b被共价连接至P；
- (xiii) a、b和P都被共价连接；
- (xiv) a被非共价连接至b；
- (xv) a被非共价连接至P；
- (xvi) b被非共价连接至P；
- (xvii) a、b和P都被非共价连接；
- (xviii) a不被包含在(a-b-P)复合物中；
- (xix) b不被包含在(a-b-P)复合物中；和/或
- (xx) P不被包含在(a-b-P)复合物中。

## DEXTRAMER在单细胞分析中的使用

[0001] 相关申请

[0002] 本申请根据35 U.S.C. §119 (e) 要求2020年11月20日提交的美国临时申请第63/116,571号的优先权以及2021年9月7日提交的美国临时申请第63/241,486号的优先权。这些申请的内容特此明确地通过引用以其整体并入。

[0003] 对序列表的引用

[0004] 本申请连同电子格式的序列表一起提交。序列表被提供为题为68EB\_317310\_W0, 创建于2021年11月3日, 大小是4.0千字节的文件。电子格式的序列表的信息通过引用以其整体并入本文。

[0005] 背景

[0006] 领域

[0007] 本公开内容总体上涉及分子生物学领域, 例如使用分子条形码化(molecular barcoding) 确定基因表达和进行免疫谱分析(immune profiling)。

[0008] 对相关技术的描述

[0009] 适应性免疫系统通过免疫细胞和抗原呈递细胞(例如, 树突状细胞、B细胞、单核细胞和巨噬细胞) 或靶细胞(例如病毒感染的细胞、细菌感染的细胞或癌细胞) 之间的特异性相互作用来指导。免疫学的一个重要领域涉及了解免疫细胞和靶细胞之间的分子相互作用。

[0010] 特别是对于T淋巴细胞(T细胞), 这种相互作用是通过克隆型T细胞受体(TCR) 和主要组织相容性复合体(MHC) I类或II类(在人类中称为人类白细胞抗原(HLA)) 之间的结合来介导的。MHC分子携带肽负荷(peptide cargo) ——抗原肽表位, 并且该肽对T细胞识别具有决定性作用。根据病原体的类型是细胞内的还是细胞外的, 抗原肽分别与MHC I类或MHC II类结合。这两类MHC复合体被不同的T细胞亚群(识别MHC I类的细胞毒性CD8+T细胞和识别MHC II类的CD4+辅助细胞) 识别。通常, TCR对MHC-肽复合物的识别导致T细胞活化、克隆扩增和T细胞分化为效应T细胞、记忆T细胞和调节性T细胞。

[0011] MHC复合体作为抗原肽受体发挥作用, 在细胞内收集肽并将其运输到细胞表面, 在那里MHC-肽复合物可以被T淋巴细胞识别。存在两类经典的MHC复合体, MHC I类和II类。这两种分子之间最重要的差异在于它们获得它们关联的肽的蛋白来源。MHC I类分子呈递源自胞质溶胶中降解的内源抗原的肽, 并因此能够展示病毒蛋白和源自癌性细胞的独特蛋白的片段。几乎所有有核细胞都在其表面表达MHC I类, 尽管不同细胞类型之间的表达水平不同。MHC II类分子结合源自外源抗原的肽。外源蛋白通过胞吞作用或吞噬作用进入细胞, 并且这些蛋白被酸化的细胞内囊泡中的蛋白酶降解, 然后被MHC II类分子呈递。MHC II类分子仅在专职抗原呈递细胞如B细胞和巨噬细胞上表达。

[0012] MHC I类和II类分子的三维结构非常相似, 但存在重要差异。MHC I类分子由两条多肽链组成, 一条跨膜的重链 $\alpha$ 和一条轻链 $\beta$ 2微球蛋白( $\beta$ 2m)。重链在称为主要组织相容性复合体(MHC) 的基因中编码, 并且其细胞外部分包含三个结构域,  $\alpha$ 1、 $\alpha$ 2和 $\alpha$ 3。 $\beta$ 2m链不在MHC基因中编码, 并由单结构域组成, 其与重链的 $\alpha$ 3结构域一起形成非常类似于免疫球蛋白结

构的折叠结构。 $\alpha 1$ 和 $\alpha 2$ 结构域配对形成肽结合槽(cleft),其由位于八条 $\beta$ 链片层上的两个分段 $\alpha$ 螺旋组成。在人类以及小鼠中存在三种不同类型的MHC I类分子。HLA-A、HLA-B、HLA-C可见于人类,而小鼠中的MHC I类分子被命名为H-2K、H-2D和H-2L。

[0013] MHC基因的显著特征是它们的多态性,其是由每个基因处的多种等位基因实现的。MHC基因的多基因性和多态性反映在肽结合槽中,使得不同的MHC复合体结合不同组的肽。肽结合槽中的可变氨基酸形成口袋,结合的肽的氨基酸侧链可以包埋在口袋中。这允许MHC的特定变体与一些肽的结合比与其他肽的结合更好。

[0014] 由于肽-MHC-T细胞受体三元复合物的短半衰期(通常在10秒和25秒之间),难以用标记的MHC-肽复合物标记特定的T细胞。为了解决这个问题,已经开发了MHC多聚体。这些MHC多聚体是包含多于一个拷贝的MHC-肽复合物的复合物,为这些复合物提供了与单体MHC-肽复合物相比增加的亲和力和相互作用半衰期。多于一个拷贝的MHC-肽复合物共价或非共价地与多聚化结构域附接。这样的MHC多聚体的已知实例包括具有IgG-多聚化结构域的MHC-二聚体、与链霉抗生物素蛋白四聚体蛋白复合的MHC-四聚体(US 5,635,363)、具有自组装卷曲螺旋结构域的MHC五聚体(US20040209295)、具有8-12个附接至streptactin的MHC streptamer和具有附接至右旋糖酐聚合物的更大数目(通常超过10个)的MHC-肽复合物的MHC dextramer。

[0015] 对T细胞识别的理解随着1996年的发现经历了巨大的技术突破:将单个肽-MHC分子多聚化成四聚体将允许肽-MHC分子和TCR之间有足够的结合强度(亲和力),以通过附接至MHC-多聚体的荧光标记来确定这种相互作用。荧光标记的MHC多聚体(I类和II类分子两者的)现在广泛用于检测T细胞和确定T细胞特异性。MHC多聚体关联的荧光可以通过例如流式细胞术或显微术来确定,或者可以基于该荧光标记通过例如流式细胞术或基于珠的分选来选择T细胞。已经开发了MHC多聚体技术,例如,通过提供具有柔性骨架的MHC多聚体,即MHC dextramer技术,实现对低亲和力T细胞的检测(例如,WO 2002/072631),并且更好地匹配T细胞识别中的巨大多样性,目的是能够检测单个样品中的多种不同T细胞特异性。抗原特异性T细胞的多重检测可以通过组合编码的MHC多聚体使用组合荧光标记方法来实现,该方法允许检测单个样品中许多不同的T细胞群体,并且最近通过使用MHC多聚体的核苷酸标记来实现。目前技术允许通过每个细胞与区室中的条形码化试剂珠共定位时将细胞特异性寡核苷酸条形码附接至来自个体细胞的多(A)mRNA分子和AbSeq Ab-Oligo,从而以大规模并行的方式(例如>10000个细胞)对单细胞进行基因表达谱分析和蛋白谱分析。对用于产生用于测序的dCODE Dextramer文库、细胞组分结合试剂特异性寡核苷酸(例如AbSeq、蛋白谱分析)文库和/或mRNA单细胞文库的方法、组合物、系统和试剂盒存在需求。

[0016] 概述

[0017] 本文的公开内容包括方法。在一些实施方案中,方法包括使多于一种受体检测构建体与多于一个细胞接触以形成与受体检测构建体关联的第一多于一个细胞,其中多于一个细胞包含多于一种细胞组分靶、和核酸靶的拷贝,其中多于一个细胞中的一个或更多个细胞包含受体结合试剂能够特异性结合的受体,并且其中多于一种受体检测构建体中的每一种包含两种或更多种受体结合试剂和受体结合试剂特异性寡核苷酸,所述受体结合试剂特异性寡核苷酸包含用于受体结合试剂的独特受体标识符序列。方法可以包括:使多于一种细胞组分结合试剂与和受体检测构建体关联的多于一个细胞接触以形成第二多于一个

细胞,其中多于一种细胞组分结合试剂中的每一种包含细胞组分结合试剂特异性寡核苷酸,所述细胞组分结合试剂特异性寡核苷酸包含用于细胞组分结合试剂的独特标识符序列,并且其中细胞组分结合试剂能够与多于一种细胞组分靶中的至少一种特异性结合。方法可以包括:用多于一种寡核苷酸条形码对细胞组分结合试剂特异性寡核苷酸进行条形码化以产生多于一种条形码化细胞组分结合试剂特异性寡核苷酸,所述多于一种条形码化细胞组分结合试剂特异性寡核苷酸各自包含与独特标识符序列的至少一部分互补的序列。方法可以包括:用多于一种寡核苷酸条形码对受体结合试剂特异性寡核苷酸进行条形码化以产生多于一种条形码化受体结合试剂特异性寡核苷酸,所述多于一种条形码化受体结合试剂特异性寡核苷酸各自包含与独特受体标识符序列的至少一部分互补的序列。方法可以包括:用多于一种寡核苷酸条形码对多于一个细胞的核酸靶的拷贝进行条形码化以产生多于一种条形码化核酸分子,所述多于一种条形码化核酸分子各自包含与核酸靶的至少一部分互补的序列。方法可以包括:产生包含多于一种核酸靶文库成员、多于一种细胞组分靶文库成员和多于一种受体文库成员的测序文库,其中产生测序文库包括:将测序衔接子附接至多于一种条形码化核酸分子或其产物,以产生多于一种核酸靶文库成员;以及将测序衔接子附接至多于一种条形码化细胞组分结合试剂特异性寡核苷酸或其产物,以产生多于一种细胞组分靶文库成员;以及将测序衔接子附接至多于一种条形码化受体结合试剂特异性寡核苷酸或其产物,以产生多于一种受体文库成员。方法可以包括:获得包含核酸靶文库成员的多于一个测序读段、细胞组分靶文库成员的多于一个测序读段和受体文库成员的多于一个测序读段的测序数据。

[0018] 本文的公开内容包括方法。在一些实施方案中,方法包括:使多于一种受体检测构建体与多于一个细胞接触以形成与受体检测构建体关联的第一多于一个细胞,其中多于一个细胞包含多于一种细胞组分靶,其中多于一个细胞中的一个或多个细胞包含受体结合试剂能够特异性结合的受体,并且其中多于一种受体检测构建体中的每一种包含两种或更多种受体结合试剂和受体结合试剂特异性寡核苷酸,所述受体结合试剂特异性寡核苷酸包含用于受体结合试剂的独特受体标识符序列。方法可以包括:使多于一种细胞组分结合试剂与和受体检测构建体关联的多于一个细胞接触以形成第二多于一个细胞,其中多于一种细胞组分结合试剂中的每一种包含细胞组分结合试剂特异性寡核苷酸,所述细胞组分结合试剂特异性寡核苷酸包含用于细胞组分结合试剂的独特标识符序列,并且其中细胞组分结合试剂能够与多于一种细胞组分靶中的至少一种特异性结合。方法可以包括:用多于一种寡核苷酸条形码对细胞组分结合试剂特异性寡核苷酸进行条形码化以产生多于一种条形码化细胞组分结合试剂特异性寡核苷酸,所述多于一种条形码化细胞组分结合试剂特异性寡核苷酸各自包含与独特标识符序列的至少一部分互补的序列。方法可以包括:用多于一种寡核苷酸条形码对受体结合试剂特异性寡核苷酸进行条形码化以产生多于一种条形码化受体结合试剂特异性寡核苷酸,所述多于一种条形码化受体结合试剂特异性寡核苷酸各自包含与独特受体标识符序列的至少一部分互补的序列。方法可以包括:产生包含多于一种细胞组分靶文库成员和多于一种受体文库成员的测序文库,其中产生测序文库包括:将测序衔接子附接至多于一种条形码化细胞组分结合试剂特异性寡核苷酸或其产物,以产生多于一种细胞组分靶文库成员;以及将测序衔接子附接至多于一种条形码化受体结合试剂特异性寡核苷酸或其产物,以产生多于一种受体文库成员。方法可以包括:获得包含细胞组

分靶文库成员的多于一个测序读段和受体文库成员的多于一个测序读段的测序数据。

[0019] 在一些实施方案中,用多于一种寡核苷酸条形码对多于一个细胞的核酸靶的拷贝进行条形码化包括:使多于一个细胞的核酸靶的拷贝与多于一种寡核苷酸条形码接触,其中多于一种寡核苷酸条形码中的每一种寡核苷酸条形码包含第一通用序列、第一分子标记和能够与核酸靶杂交的靶结合区;以及使与核酸靶的拷贝杂交的多于一种寡核苷酸条形码延伸以产生多于一种条形码化核酸分子,所述多于一种条形码化核酸分子各自包含与核酸靶的至少一部分互补的序列。在一些实施方案中,用多于一种寡核苷酸条形码对细胞组分结合试剂特异性寡核苷酸进行条形码化包括:使细胞组分结合试剂特异性寡核苷酸与多于一种寡核苷酸条形码接触,其中多于一种寡核苷酸条形码中的每一种寡核苷酸条形码包含第一通用序列、第一分子标记和能够与细胞组分结合试剂特异性寡核苷酸杂交的靶结合区;以及使与细胞组分结合试剂特异性寡核苷酸杂交的多于一种寡核苷酸条形码延伸以产生多于一种条形码化细胞组分结合试剂特异性寡核苷酸,所述多于一种条形码化细胞组分结合试剂特异性寡核苷酸各自包含与独特标识符序列的至少一部分互补的序列。在一些实施方案中,用多于一种寡核苷酸条形码对受体结合试剂特异性寡核苷酸进行条形码化包括:使受体结合试剂特异性寡核苷酸与多于一种寡核苷酸条形码接触,其中多于一种寡核苷酸条形码中的每一种寡核苷酸条形码包含第一通用序列、第一分子标记和能够与受体结合试剂特异性寡核苷酸杂交的靶结合区;以及使与受体结合试剂特异性寡核苷酸杂交的多于一种寡核苷酸条形码延伸以产生多于一种条形码化受体结合试剂特异性寡核苷酸,所述多于一种条形码化受体结合试剂特异性寡核苷酸各自包含与独特受体标识符序列的至少一部分互补的序列。

[0020] 在一些实施方案中,靶结合区包含捕获序列。在一些实施方案中,(i) 细胞组分结合试剂特异性寡核苷酸包含与配置为捕获细胞组分结合试剂特异性寡核苷酸的捕获序列互补的序列,和/或(ii) 受体结合试剂特异性寡核苷酸包含与配置为捕获受体结合试剂特异性寡核苷酸的捕获序列互补的序列。在一些实施方案中,靶结合区包含多(dT)区、随机序列、靶特异性序列或其组合。

[0021] 在一些实施方案中,(i) 多于一种条形码化核酸分子中的每一种条形码化核酸分子包含第一通用序列和第一分子标记;(ii) 多于一种条形码化细胞组分结合试剂特异性寡核苷酸中的每一种条形码化细胞组分结合试剂特异性寡核苷酸包含第一通用序列和第一分子标记;和/或(iii) 多于一种条形码化受体结合试剂特异性寡核苷酸中的每一种条形码化受体结合试剂特异性寡核苷酸包含第一通用序列和第一分子标记。

[0022] 在一些实施方案中,第二多于一个细胞包含一个或更多个单细胞。方法可以包括:在(i) 使多于一个细胞的核酸靶的拷贝与多于一种寡核苷酸条形码接触,(ii) 使细胞组分结合试剂特异性寡核苷酸与多于一种寡核苷酸条形码接触,和/或(iii) 使受体结合试剂特异性寡核苷酸与多于一种寡核苷酸条形码接触之前:将第二多于一个细胞分区至多于一个分区,其中多于一个分区中的分区包含来自第二多于一个细胞的单细胞;以及在包含单细胞的分区中,(i) 使多于一个细胞的核酸靶的拷贝与多于一种寡核苷酸条形码接触,(ii) 使细胞组分结合试剂特异性寡核苷酸与多于一种寡核苷酸条形码接触,和/或(iii) 使受体结合试剂特异性寡核苷酸与多于一种寡核苷酸条形码接触。在一些实施方案中,分区是孔或液滴。在一些实施方案中,多于一种寡核苷酸条形码与固体支持物关联,方法包括将固体支

持物与分区中的单细胞关联,并且其中多于一个分区中的分区包含单个固体支持物。方法可以包括:在分区步骤之后且在接触步骤之前裂解单细胞。在一些实施方案中,裂解单细胞包括加热、与去污剂接触、改变pH或其任何组合。在一些实施方案中,多于一个细胞包括T细胞、B细胞、肿瘤细胞、髓样细胞、血细胞、正常细胞、胎儿细胞、母体细胞或其混合物。

[0023] 在一些实施方案中,多于一种寡核苷酸条形码中的至少10种包含不同的第一分子标记序列。在一些实施方案中,多于一种寡核苷酸条形码各自包含细胞标记。在一些实施方案中,多于一种寡核苷酸条形码的每一种细胞标记包含至少6个核苷酸。在一些实施方案中,多于一种寡核苷酸条形码中与同一固体支持物关联的寡核苷酸条形码包含相同的细胞标记。在一些实施方案中,多于一种寡核苷酸条形码中与不同固体支持物关联的寡核苷酸条形码包含不同的细胞标记。在一些实施方案中,固体支持物包括合成颗粒、平坦表面或其组合。在一些实施方案中,多于一种寡核苷酸条形码中的至少一种寡核苷酸条形码被固定或部分固定在合成颗粒上,或者多于一种寡核苷酸条形码中的至少一种寡核苷酸条形码被包封或部分包封在合成颗粒中。在一些实施方案中,合成颗粒是可破坏的(例如,可破坏的水凝胶颗粒)。在一些实施方案中,合成颗粒包括珠(例如,珠是琼脂糖凝胶珠(sepharose bead)、链霉抗生物素蛋白珠、琼脂糖珠(agarose bead)、磁珠、缀合珠、蛋白A缀合珠、蛋白G缀合珠、蛋白A/G缀合珠、蛋白L缀合珠、寡(dT)缀合珠、二氧化硅珠、二氧化硅样珠、抗生物素微珠、抗荧光染料微珠,或其任何组合)。在一些实施方案中,合成颗粒包含选自自由以下组成的组的材料:聚二甲基硅氧烷(PDMS)、聚苯乙烯、玻璃、聚丙烯、琼脂糖、明胶、水凝胶、顺磁物质、陶瓷、塑料、玻璃、甲基苯乙烯、丙烯酸聚合物、钛、胶乳、琼脂糖凝胶、纤维素、尼龙、硅酮,以及它们的任何组合。在一些实施方案中,多于一种寡核苷酸条形码中的每一种寡核苷酸条形码包含接头官能团。在一些实施方案中,合成颗粒包含固体支持物官能团。在一些实施方案中,支持物官能团和接头官能团彼此关联。在一些实施方案中,接头官能团和支持物官能团单独地选自自由C6、生物素、链霉抗生物素蛋白、一种或更多种伯胺、一种或更多种醛、一种或更多种酮及其任何组合组成的组。

[0024] 在一些实施方案中,产生测序文库包括:使随机引物与多于一种条形码化核酸分子接触,其中随机引物中的每一种包含第四通用序列或其互补物;以及使与多于一种条形码化核酸分子杂交的随机引物延伸以产生第一多于一种延伸产物。方法可以包括:使用能够与第一通用序列或其互补物杂交的引物和能够与第四通用序列或其互补物杂交的引物扩增第一多于一种延伸产物,从而产生第一多于一种条形码化扩增子,其中多于一种核酸靶文库成员包括第一多于一种条形码化扩增子或其产物。在一些实施方案中,扩增第一多于一种延伸产物包括将测序引物的结合位点和/或测序衔接子、其互补序列和/或其部分的序列添加到第一多于一种延伸产物。方法可以包括:基于与第一多于一种条形码化扩增子或其产物关联的具有不同序列的第一分子标记的数目来确定一个或更多个单细胞中的每一个中核酸靶的拷贝数。

[0025] 在一些实施方案中,产生测序文库包括:使用能够与第一通用序列或其互补物杂交的引物和能够与第四通用序列或其互补物杂交的引物扩增第一多于一种条形码化扩增子,从而产生第二多于一种条形码化扩增子;其中多于一种核酸靶文库成员包括第二多于一种条形码化扩增子或其产物。在一些实施方案中,扩增第一多于一种条形码化扩增子包括将测序引物的结合位点和/或测序衔接子、其互补序列和/或其部分的序列添加到第一多

于一种条形码化扩增子。方法可以包括：基于与第二多于一种条形码化扩增子或其产物关联的具有不同序列的第一分子标记的数目来确定一个或多个单细胞中的每一个中核酸靶的拷贝数。在一些实施方案中，第一多于一种条形码化扩增子和/或第二多于一种条形码化扩增子包含全转录组扩增(WTA)产物。

[0026] 在一些实施方案中，产生测序文库包括：使用多于一种条形码化核酸分子作为模板合成第三多于一种条形码化扩增子以产生第三多于一种条形码化扩增子，其中多于一种核酸靶文库成员包括第三多于一种条形码化扩增子或其产物。在一些实施方案中，合成第三多于一种条形码化扩增子包括使用能够与第一通用序列或其互补物杂交的引物和靶特异性引物进行PCR扩增。在一些实施方案中，合成第三多于一种条形码化扩增子包括将测序引物的结合位点和/或测序衔接子、其互补序列和/或其部分的序列添加到条形码化核酸分子。方法可以包括：基于与第三多于一种条形码化扩增子或其产物关联的具有不同序列的第一分子标记的数目来确定一个或多个单细胞中的每一个中核酸靶的拷贝数。在一些实施方案中，靶特异性引物与免疫受体特异性杂交。在一些实施方案中，靶特异性引物与免疫受体的恒定区、免疫受体的可变区、免疫受体的多样性区、免疫受体的可变区和多样性区的连接部(junction)或其组合特异性杂交。在一些实施方案中，免疫受体是T细胞受体(TCR)和/或B细胞受体(BCR)。在一些实施方案中，TCR包含TCR $\alpha$ 链、TCR $\beta$ 链、TCR $\gamma$ 链、TCR $\delta$ 链或其任何组合。在一些实施方案中，BCR包含BCR重链和/或BCR轻链。

[0027] 在一些实施方案中，多于一种条形码化核酸分子或其产物的多于一个测序读段的每一个包含(1)分子标记序列，和/或(2)核酸靶的子序列。方法可以包括：基于核酸靶文库成员的多于一个测序读段来确定一个或多个单细胞中的每一个中核酸靶的拷贝数。在一些实施方案中，确定一个或多个单细胞中的每一个中核酸靶的拷贝数包括基于与一种或更多种核酸靶文库成员或其产物关联的具有不同序列的第一分子标记、其互补物或其组合的数目来确定一个或多个单细胞中的每一个中核酸靶的拷贝数。在一些实施方案中，多于一种条形码化核酸分子包括条形码化脱氧核糖核酸(DNA)分子、条形码化核糖核酸(RNA)分子或其组合。在一些实施方案中，核酸靶包括核酸分子(例如，核糖核酸(RNA)、信使RNA(mRNA)、微RNA、小干扰RNA(siRNA)、RNA降解产物、包含多(A)尾的RNA或其任何组合)。在一些实施方案中，mRNA编码免疫受体。

[0028] 在一些实施方案中，细胞组分结合试剂特异性寡核苷酸包含第三通用序列。在一些实施方案中，产生测序文库包括：使用能够与第一通用序列或其互补物杂交的引物和能够与第三通用序列或其互补物杂交的引物，扩增多于一种条形码化细胞组分结合试剂特异性寡核苷酸或其产物，以产生多于一种扩增的条形码化细胞组分结合试剂特异性寡核苷酸，其中多于一种细胞组分靶文库成员包括多于一种扩增的条形码化细胞组分结合试剂特异性寡核苷酸或其产物。在一些实施方案中，扩增多于一种条形码化细胞组分结合试剂特异性寡核苷酸包括将测序引物的结合位点和/或测序衔接子、其互补序列和/或其部分的序列添加到多于一种条形码化细胞组分结合试剂特异性寡核苷酸。方法可以包括：基于细胞组分靶文库成员的多于一个测序读段来确定一个或多个单细胞中多于一种细胞组分靶中的至少一种细胞组分靶的拷贝数。

[0029] 在一些实施方案中，细胞组分靶文库成员的多于一个测序读段中的每一个包含(1)分子标记序列，和/或(2)独特标识符序列的至少一部分。在一些实施方案中，细胞组分

结合试剂特异性寡核苷酸包含第二分子标记。在一些实施方案中,多于一种细胞组分结合试剂特异性寡核苷酸中的至少10种包含不同的第二分子标记序列。在一些实施方案中,至少两种细胞组分结合试剂特异性寡核苷酸的第二分子标记序列是不同的,并且其中所述至少两种细胞组分结合试剂特异性寡核苷酸的独特标识符序列是相同的。在一些实施方案中,至少两种细胞组分结合试剂特异性寡核苷酸的第二分子标记序列是不同的,并且其中所述至少两种细胞组分结合试剂特异性寡核苷酸的独特标识符序列是不同的。在一些实施方案中,测序数据中与用于细胞组分结合试剂的独特标识符序列关联的独特第一分子标记序列的数目指示一个或更多个单细胞中至少一种细胞组分靶的拷贝数,所述细胞组分结合试剂能够与至少一种细胞组分靶特异性结合。在一些实施方案中,测序数据中与用于细胞组分结合试剂的独特标识符序列关联的独特第二分子标记序列的数目指示一个或更多个单细胞中至少一种细胞组分靶的拷贝数,所述细胞组分结合试剂能够与至少一种细胞组分靶特异性结合。

[0030] 方法可以包括:在使多于一种细胞组分结合试剂与和受体检测构建体关联的第一多于一个细胞接触之后,去除多于一种细胞组分结合试剂中未与和受体检测构建体关联的第一多于一个细胞接触的一种或更多种细胞组分结合试剂。在一些实施方案中,去除未与和受体检测构建体关联的第一多于一个细胞接触的一种或更多种细胞组分结合试剂包括:去除未与多于一种细胞组分靶中相应的至少一种细胞组分靶接触的一种或更多种细胞组分结合试剂。

[0031] 在一些实施方案中,靶结合区包含多(dT)区,并且其中细胞组分结合试剂特异性寡核苷酸包含多(dA)区。在一些实施方案中,细胞组分结合试剂特异性寡核苷酸包含与多(dA)区相邻的对齐序列(alignment sequence)。在一些实施方案中,(a)对齐序列包含鸟嘌呤、胞嘧啶、胸腺嘧啶、尿嘧啶或其组合;(b)对齐序列包含多(dT)序列、多(dG)序列、多(dC)序列、多(dU)序列或其组合;和/或(c)对齐序列位于多(dA)区的5'。在一些实施方案中,细胞组分结合试剂特异性寡核苷酸通过接头与细胞组分结合试剂关联。在一些实施方案中,细胞组分结合试剂特异性寡核苷酸被配置为能够从细胞组分结合试剂脱离。方法可以包括:使细胞组分结合试剂特异性寡核苷酸与细胞组分结合试剂解离。在一些实施方案中,接头包含碳链。在一些实施方案中,碳链包含2-30个碳,并且还任选地碳链包含12个碳。在一些实施方案中,接头包含5'氨基修饰物C12(5AmMC12)或其衍生物。在一些实施方案中,细胞组分靶包括蛋白靶。在一些实施方案中,细胞组分结合试剂包含抗体或其片段。在一些实施方案中,细胞组分靶包括糖类、脂质、蛋白、细胞外蛋白、细胞表面蛋白、细胞标志物、B细胞受体、T细胞受体、主要组织相容性复合体、肿瘤抗原、受体、细胞内蛋白或其任何组合。在一些实施方案中,细胞组分结合试剂包含抗体或其片段。在一些实施方案中,细胞组分靶在细胞表面上。在一些实施方案中,使多于一种受体检测构建体与多于一个细胞接触以形成第一多于一个细胞包括去除未与一种或更多种受体检测构建体结合的一个或更多个细胞。在一些实施方案中,去除未与一种或更多种受体检测构建体结合的一个或更多个细胞包括选择与一种或更多种受体检测构建体结合的细胞。

[0032] 在一些实施方案中,受体检测构建体包含一种或更多种另外的标记,其中选择与一种或更多种受体检测构建体结合的细胞包括流式细胞术和/或基于一种或更多种另外的标记的存在选择细胞。在一些实施方案中,一种或更多种另外的标记包括荧光标记。在一些

实施方案中,去除未与和受体检测构建体关联的第一多于一个细胞接触的一种或更多种受体结合试剂包括:去除未与受体接触的一种或更多种受体结合试剂。在一些实施方案中,受体结合试剂特异性寡核苷酸包含第二通用序列。在一些实施方案中,产生测序文库包括:使用能够与第一通用序列或其互补物杂交的引物和能够与第二通用序列或其互补物杂交的引物,扩增多于一种条形码化受体结合试剂特异性寡核苷酸或其产物,以产生多于一种扩增的条形码化受体结合试剂特异性寡核苷酸,其中多于一种受体文库成员包括多于一种扩增的条形码化受体结合试剂特异性寡核苷酸或其产物。在一些实施方案中,扩增多于一种条形码化受体结合试剂特异性寡核苷酸包括将测序引物的结合位点和/或测序衔接子、其互补序列和/或其部分的序列添加到多于一种条形码化受体结合试剂特异性寡核苷酸。

[0033] 在一些实施方案中,受体文库成员的多于一个测序读段中的每一个包含(1)分子标记序列,和/或(2)独特受体标识符序列的至少一部分。在一些实施方案中,受体文库成员的多于一个测序读段中的每一个包含细胞标记序列。在一些实施方案中,每种独特细胞标记序列指示第二多于一个细胞中的单细胞。方法可以包括:基于受体文库成员的多于一个测序读段来确定一个或更多个单细胞中受体的拷贝数。方法可以包括:基于受体文库成员的多于一个测序读段来确定与一个或更多个单细胞结合的受体结合试剂的身份。方法可以包括:基于受体文库成员的多于一个测序读段来确定一个或更多个单细胞中受体的身份。在一些实施方案中,受体结合试剂特异性寡核苷酸包含第三分子标记。在一些实施方案中,多于一种受体结合试剂特异性寡核苷酸中的至少10种包含不同的第三分子标记序列。在一些实施方案中,至少两种受体结合试剂特异性寡核苷酸的第三分子标记序列是不同的,并且其中所述至少两种受体结合试剂特异性寡核苷酸的独特受体标识符序列是相同的。在一些实施方案中,至少两种受体结合试剂特异性寡核苷酸的第三分子标记序列是不同的,并且其中所述至少两种受体结合试剂特异性寡核苷酸的独特受体标识符序列是不同的。在一些实施方案中,测序数据中与用于受体结合试剂的独特受体标识符序列关联的独特第一分子标记序列的数目指示一个或更多个单细胞中受体的拷贝数,所述受体结合试剂能够与至少一种受体靶特异性结合。在一些实施方案中,测序数据中与用于受体结合试剂的独特受体标识符序列关联的独特第三分子标记序列的数目指示一个或更多个单细胞中受体的拷贝数,所述受体结合试剂能够与至少一种受体特异性结合。在一些实施方案中,与相同受体结合试剂关联的受体结合试剂特异性寡核苷酸包含相同的独特受体标识符序列,并且其中与不同受体结合试剂关联的受体结合试剂特异性寡核苷酸包含不同的独特受体标识符序列。

[0034] 在一些实施方案中,受体结合试剂特异性寡核苷酸包括DNA、RNA、锁核酸(locked nucleic acid,LNA)、肽核酸(PNA)、LNA/PNA嵌合体、LNA/DNA嵌合体、PNA/DNA嵌合体或其任何组合。在一些实施方案中,受体结合试剂特异性寡核苷酸的长度为约10个核苷酸至约100个核苷酸。在一些实施方案中,受体包括分化簇(CD)分子。在一些实施方案中,受体包括免疫受体(例如,T细胞受体(TCR))。在一些实施方案中,独特受体标识符序列的长度为约3个核苷酸至约100个核苷酸。在一些实施方案中,靶结合区包含多(dT)区,并且其中受体结合试剂特异性寡核苷酸包含多(dA)区。在一些实施方案中,受体结合试剂特异性寡核苷酸包含与多(dA)区相邻的对齐序列。在一些实施方案中,(a)对齐序列包含鸟嘌呤、胞嘧啶、胸腺嘧啶、尿嘧啶或其组合;(b)对齐序列包含多(dT)序列、多(dG)序列、多(dC)序列、多(dU)序

列或其组合；和/或(c)对齐序列位于多(dA)区的5'。在一些实施方案中,受体结合试剂特异性寡核苷酸通过接头与受体结合试剂关联。在一些实施方案中,受体结合试剂特异性寡核苷酸被配置为能够从受体结合试剂脱离。方法可以包括:使受体结合试剂特异性寡核苷酸与受体结合试剂解离。在一些实施方案中,接头包含碳链。在一些实施方案中,碳链包含2-30个碳,并且还任选地碳链包含12个碳。在一些实施方案中,接头包含5'氨基修饰物C12(5AmMC12)或其衍生物。

[0035] 多于一种受体检测构建体可以包括一种或更多种MHC多聚体,并且其中所述两种或更多种受体结合试剂包含两种或更多种MHC-肽复合物。MHC多聚体可以包含 $(a-b-P)_n$ ,其中 $n>1$ ,其中多肽a和多肽b一起形成能够结合肽P的功能性MHC蛋白,并且 $(a-b-P)$ 是当肽P与功能性MHC蛋白结合时形成的MHC-肽复合物。在一些实施方案中,MHC多聚体的每种MHC-肽复合物与一个或更多个多聚化结构域关联。在一些实施方案中,多于一种受体检测构建体包含2种MHC多聚体、3种MHC多聚体、4种MHC多聚体、5种MHC多聚体、6种MHC多聚体、7种MHC多聚体、8种MHC多聚体、9种MHC多聚体、10种MHC多聚体、11种MHC多聚体、12种MHC多聚体、13种MHC多聚体、14种MHC多聚体、15种MHC多聚体、16种MHC多聚体、17种MHC多聚体、18种MHC多聚体、19种MHC多聚体或20种MHC多聚体。在一些实施方案中,所述MHC多聚体的每种MHC-肽复合物的个体抗原肽P是相同或不同的。在一些实施方案中,多于一种受体检测构建体包括一种或更多种阴性对照MHC多聚体。在一些实施方案中,一种或更多种阴性对照MHC多聚体,其中每种MHC多聚体包含阴性对照肽P。在一些实施方案中,所述阴性对照肽P是无义肽。在一些实施方案中,所述一种或更多种阴性对照MHC多聚体是空MHC多聚体。在一些实施方案中,多于一种受体检测构建体包含一种或更多种阳性对照MHC多聚体,其中每种MHC多聚体包含阳性对照肽P。在一些实施方案中,包含 $(a-b-P)_n$ 的所述一种或更多种MHC多聚体的n的值是 $1<n\leq 1000$ 。在一些实施方案中,n的值在2-3、3-4、4-5、5-6、6-7、7-8、8-9、9-10、10-11、11-12、12-13、13-14、14-15、15-16、16-17、17-18、18-19、19-20、20-21、21-22、22-23、23-24、24-25、25-26、26-27、27-28、28-29、29-30、30-35、35-40、40-45、45-50、50-55、55-60、60-65、65-70、75-80、80-85、85-90、90-95、95-100、100-110、110-120、120-130、130-140、140-150、150-160、160-170、170-180、180-190、190-200、200-225、225-250、250-275、275-300、300-325、325-350、350-375、375-400、400-450、450-500、500-550、550-600、600-650、650-700、700-750、750-800、800-850、850-900、900-950或950-1000之间。在一些实施方案中,所述MHC蛋白是MHC I类。在一些实施方案中,所述MHC蛋白是MHC I类、MHC II类、MR1、CD1或MHC样分子,并且抗原肽P选自自由与MHC I类、MHC II类、MR1、CD1或MHC样分子结合的8-mer、9-mer、10-mer、11-mer和12-mer肽组成的组。

[0036] 在一些实施方案中,多聚体MHC的每种MHC蛋白或MHC-肽复合物中的一个或更多个和一个或更多个多聚化结构域之间的关联是共价关联和/或非共价关联。在一些实施方案中,一个或更多个多聚化结构域包含一个或更多个多聚化结构域连接体分子(connector molecule)。在一些实施方案中,多聚体MHC的每种MHC蛋白或MHC-肽复合物中的一个或更多个包含一个或更多个MHC连接体分子。在一些实施方案中,一个或更多个多聚化结构域连接体分子包括一个或更多个链霉抗生物素蛋白和/或一个或更多个抗生物素蛋白,或其任何衍生物。在一些实施方案中,一个或更多个链霉抗生物素蛋白包括一个或更多个四聚体链霉抗生物素蛋白变体或一个或更多个单体链霉抗生物素蛋白变体。一个或更多个MHC连接

体分子可以是生物素。在一些实施方案中,MHC多聚体的每种MHC蛋白或MHC-肽复合物中的一个或更多个通过链霉抗生物素蛋白-生物素连接或抗生物素蛋白-生物素连接与一个或更多个多聚化结构域关联。在一些实施方案中,MHC多聚体的每种MHC蛋白或MHC-肽复合物中的一个或更多个通过接头部分与一个或更多个多聚化结构域关联。在一些实施方案中,MHC多聚体的每种MHC蛋白或MHC-肽复合物中的一个或更多个通过天然二聚化和/或蛋白-蛋白相互作用与一个或更多个多聚化结构域关联。在一些实施方案中,天然二聚化和/或蛋白-蛋白相互作用选自由以下组成的组:亮氨酸拉链例如AP-1的亮氨酸拉链结构域、Fos/Jun相互作用、基于酸性/碱性卷曲螺旋结构的相互作用(例如螺旋)、抗体/抗原相互作用、多核苷酸-多核苷酸相互作用例如DNA/DNA、DNA/PNA、DNA/RNA、PNA/PNA、LNA/DNA、合成分子-合成分子相互作用和蛋白-小分子相互作用、IgG二聚体蛋白、IgM多价蛋白、螯合物/金属离子结合螯合物、strep免疫球蛋白、抗体(单克隆、多克隆和重组的)、抗体片段及其衍生物、hexa-his(金属螯合部分)、hexa-hat GST(谷胱甘肽S-转移酶)谷胱甘肽亲和力、钙调蛋白结合肽(CBP)、Strep-标签、纤维素结合结构域、麦芽糖结合蛋白、S-肽标签、壳多糖结合标签、免疫反应性表位、表位标签、E2标签、HA表位标签、Myc表位、FLAG表位、AU1表位和AU5表位、Glu-Glu表位、KT3表位、IRS表位、Btag表位、蛋白激酶-C表位、VSV表位、介导与包括糖类、脂质和蛋白的多种化合物结合的凝集素(lectin),例如Con A(直生刀豆(*Cannvalia ensiformis*))或WGA(麦胚凝集蛋白(wheat germ agglutinin))和四联凝集素(tetranectin)或蛋白A或蛋白G(抗体亲和力)。

[0037] 在一些实施方案中,一个或更多个多聚化结构域包含(i)一个或更多个支架;(ii)一个或更多个载体;(iii)至少一个支架和至少一个载体;和/或(iv)一个或更多个任选取代的有机分子。在一些实施方案中,任选取代的有机分子包含一种或更多种官能化的环状结构。在一些实施方案中,官能化的环状结构包含苯环。在一些实施方案中,任选取代的有机分子包括支架分子,所述支架分子包含至少三个反应性基团,或至少三个适于非共价附接的位点。在一些实施方案中,一个或更多个多聚化结构域包含一种或更多种生物细胞和/或细胞样结构。在一些实施方案中,一个或更多个多聚化结构域包含一种或更多种膜。在一些实施方案中,一种或更多种膜包括脂质体或胶束。在一些实施方案中,一个或更多个多聚化结构域包含一种或更多种聚合物(例如,一种或更多种合成聚合物)。在一些实施方案中,一种或更多种聚合物包括多糖。

[0038] 多糖可以包含一个或更多个右旋糖酐部分。在一些实施方案中,一个或更多个右旋糖酐部分(i)与一个或更多个MHC肽复合物共价附接;(ii)与一个或更多个MHC肽复合物非共价附接;和/或(iii)被修饰。一个或更多个右旋糖酐部分可以包含一个或更多个氨基-右旋糖酐。一个或更多个氨基-右旋糖酐可以用二乙烯基砜修饰。在一些实施方案中,一个或更多个右旋糖酐部分包含一个或更多个具有约1,000至50,000Da的分子量的右旋糖酐。在一些实施方案中,一个或更多个右旋糖酐部分包含一个或更多个具有约50,000至150,000Da(例如,约50,000至60,000、约60,000至70,000、约70,000至80,000、约80,000至90,000、约90,000至100,000、约100,000至110,000、约110,000至120,000、约120,000至130,000、约130,000至140,000或约140,000至150,000Da)的分子量的右旋糖酐。在一些实施方案中,一个或更多个右旋糖酐部分包含一个或更多个右具有约150,000Da至270,000Da的分子量的右旋糖酐。一个或更多个右旋糖酐部分可以是线性的和/或支链的。

[0039] 在一些实施方案中,一种或更多种合成聚合物包括PNA、聚酰胺、PEG或其任何组合。在一些实施方案中,一个或更多个多聚化结构域包含一种或更多种选自以下组成的组的实体: IgG结构域、卷曲螺旋多肽结构、DNA双链体、核酸双链体、PNA-PNA、PNA-DNA和DNA-RNA。在一些实施方案中,一个或更多个多聚化结构域包含抗体。在一些实施方案中,抗体选自以下组成的组: 多克隆抗体、单克隆抗体、IgA、IgG、IgM、IgD、IgE、IgG1、IgG2、IgG3、IgG4、IgA1、IgA2、IgM1、IgM2、人源化抗体、人源化单克隆抗体、嵌合抗体、小鼠抗体、大鼠抗体、兔抗体、人类抗体、骆驼抗体、绵羊抗体、工程化人类抗体、表位聚焦抗体(epitope-focused antibody)、激动剂抗体、拮抗剂抗体、中和抗体、天然存在的抗体、分离的抗体、单价抗体、双特异性抗体、三特异性抗体、多特异性抗体、异源缀合抗体、免疫缀合物、免疫脂质体、标记的抗体、抗体片段、结构域抗体、纳米抗体(nanobody)、微抗体(minibody)、大抗体(maxibody)、双体(diabody)和融合抗体。在一些实施方案中,一个或更多个多聚化结构域包含一种或更多种有机小支架分子或有机小分子。在一些实施方案中,一种或更多种有机小分子包括一种或更多种类固醇、一种或更多种肽和/或一种或更多种芳族有机分子。在一些实施方案中,一种或更多种芳族有机分子包含一种或更多种双环结构、一种或更多种多环结构或一种或更多种单环结构; 任选地所述一种或更多种单环结构包括一种或更多种任选地官能化或取代的苯环。在一些实施方案中,一个或更多个多聚化结构域包含一种或更多种: (i) 能够聚合的单体分子; (ii) 生物聚合物, 诸如一种或更多种蛋白; (iii) 小分子支架; (iv) 超分子结构, 诸如一种或更多种纳米团簇; 和/或(v) 蛋白复合物。

[0040] 一个或更多个多聚化结构域可以包含一种或更多种珠。在一些实施方案中,一种或更多种珠包括一种或更多种带有亲电基团例如二乙烯基砜活化的多糖的珠、已经用甲苯磺酰基活化的酯官能化的聚苯乙烯珠、用甲苯磺酰基活化的酯官能化的磁性聚苯乙烯珠, 以及其中MHC复合体已经通过MHC复合体中包含的亲核体与珠的亲电体的反应而共价固定至这些珠的珠。在一些实施方案中,一种或更多种珠选自以下组成的组: 琼脂糖凝胶珠、聚丙烯酰胺葡聚糖(sephacryl)珠、聚苯乙烯珠、琼脂糖珠、多糖珠、聚氨基甲酸酯(polycarbamate)珠和可以悬浮在水性缓冲液中的任何其他种类的珠。在一些实施方案中,多聚化结构域包含一种或更多种选自以下的化合物: 琼脂糖珠、琼脂糖凝胶珠、树脂珠、玻璃珠、多孔玻璃珠、涂覆疏水聚合物的玻璃颗粒、壳聚糖涂覆珠、SH珠、胶乳珠、球形胶乳珠、等位基因型珠(allele-type beads)、SPA珠、基于PEG的树脂、PEG涂覆珠、PEG包封珠、聚苯乙烯珠、磁性聚苯乙烯珠、谷胱甘肽琼脂糖珠、磁珠、顺磁珠、蛋白A和/或蛋白G琼脂糖凝胶珠、活化羧酸珠、宏观珠(macroscopic beads)、微观珠(microscopic beads)、不溶性树脂珠、基于二氧化硅的树脂、纤维素树脂、交联琼脂糖珠、聚苯乙烯珠、交联聚丙烯酰胺树脂、具有铁芯的珠、金属珠、dynabead、用NHS活化的聚甲基丙烯酸甲酯珠、链霉抗生物素蛋白-琼脂糖珠、聚丙烯、聚乙烯、右旋糖酐、尼龙、淀粉酶、天然纤维素和改性纤维素、硝化纤维素、聚丙烯酰胺、辉长岩、磁铁矿、聚合物、寡聚物、非重复部分、聚乙二醇(PEG)、单甲氧基-PEG、单-(C<sub>1</sub>-C<sub>10</sub>)烷氧基-PEG、芳氧基-PEG、聚(N-乙基吡咯烷酮)PEG、三氟乙磺酰氯单甲氧基PEG(tresyl monomethoxy PEG)、PEG丙醛、双琥珀酰亚胺基碳酸PEG、二乙烯基苯交联聚苯乙烯珠、丙二醇均聚物、聚环氧丙烷/环氧乙烷共聚物、聚氧乙基化多元醇(例如,甘油)、聚乙烯醇、右旋糖酐、氨基右旋糖酐、基于糖类的聚合物、交联右旋糖酐珠、多糖珠、聚

氨基甲酸酯珠、二乙烯基砷活化多糖、已经用甲苯磺酰基活化的酯官能化的聚苯乙烯珠、用甲苯磺酰基活化的酯官能化的磁性聚苯乙烯珠、链霉抗生物素蛋白珠、链霉抗生物素蛋白单体涂覆的珠、链霉抗生物素蛋白四聚体涂覆的珠、链霉抗生物素蛋白涂覆的Compel磁珠、抗生物素蛋白涂覆的珠、dextramer涂覆的珠、二乙烯基砷活化的右旋糖酐、羧酸修饰的珠、胺修饰的珠、抗体涂覆的珠、纤维素珠、接枝共聚珠、聚丙烯酰胺珠、任选地与N-N' -双丙烯酰基乙二胺交联的二甲基丙烯酰胺珠、中空纤维膜、荧光珠、胶原-琼脂糖珠、明胶珠、胶原-明胶珠、胶原-纤连蛋白-明胶珠、胶原珠、壳聚糖珠、胶原-壳聚糖珠、基于蛋白的珠、水凝胶珠、半纤维素、烷基纤维素、羟烷基纤维素、羧甲基纤维素、磺基乙基纤维素、淀粉、木聚糖、支链淀粉、软骨素、透明质酸、肝素、瓜尔胶、黄原胶、甘露聚糖、半乳甘露聚糖、壳多糖和壳聚糖。

[0041] 一个或更多个多聚化结构域可以包含二聚化结构域、三聚化结构域、四聚化结构域、五聚化结构域或六聚化结构域。在一些实施方案中，一个或更多个多聚化结构域包含附接一个或更多个支架的聚合物结构。在一些实施方案中，聚合物结构包含多糖。在一些实施方案中，多糖包含一个或更多个右旋糖酐部分。在一些实施方案中，一个或更多个多聚化结构域包含聚酰胺、聚乙二醇、多糖、琼脂糖凝胶或其任何组合。在一些实施方案中，一个或更多个多聚化结构域包含羧甲基右旋糖酐、右旋糖酐聚醛、羧甲基右旋糖酐内酯、环糊精或其任何组合。在一些实施方案中，一个或更多个多聚化结构域具有小于1,000Da、1,000Da至小于10,000Da、10,000Da至小于100,000Da、100,000Da至小于1,000,000Da、超过1,000,000Da的分子量。在一些实施方案中，所述一种或更多种MHC多聚体还包含一个或更多个选自由以下组成的组的支架、载体和/或接头：链霉抗生物素蛋白(SA)和抗生物素蛋白及其衍生物、生物素、免疫球蛋白、抗体(单克隆、多克隆和重组的)、抗体片段及其衍生物、AP-1的亮氨酸拉链结构域(jun和fos)、hexa-his(金属螯合部分)、hexa-hat GST(谷胱甘肽S-转移酶)谷胱甘肽亲和力、钙调蛋白结合肽(CBP)、Strep-标签、纤维素结合结构域、麦芽糖结合蛋白、S-肽标签、壳多糖结合标签、免疫反应性表位、表位标签、E2标签、HA表位标签、Myc表位、FLAG表位、AU1和AU5表位、Glu-Glu表位、KT3表位、IRS表位、Btag表位、蛋白激酶-C表位、VSV表位、介导与包括糖类、脂质和蛋白的多种化合物结合的凝集素，例如Con A(直生刀豆)或WGA(麦胚凝集蛋白)和四联凝集素或蛋白A或蛋白G(抗体亲和力)。在一些实施方案中，所述MHC多聚体包含由多聚化结构域连接部分连接的多于一个相同或不同的多聚化结构域。在一些实施方案中，所述MHC多聚体包含连接至第二多聚化结构域的第一多聚化结构域。

[0042] 一种或更多种MHC多聚体可以包含一个或更多个标记。一个或更多个标记可以包括受体结合试剂特异性寡核苷酸。在一些实施方案中，一个或更多个标记包括受体结合试剂特异性寡核苷酸和一种或更多种另外的标记。在一些实施方案中，一种或更多种另外的标记是肽标记、荧光团标记、重金属标记、同位素标记、放射标记、放射性核素、稳定同位素、同位素链和单原子、化学发光标记、生物发光标记、放射性标记、酶标记、DNA荧光染色剂、镧系元素、离子载体、与特定离子结合的螯合化合物或其任何组合。在一些实施方案中，一个或更多个标记包括共价附接的标记和/或非共价附接的标记。在一些实施方案中，一个或更多个标记与以下附接：(i)MHC多肽a；(ii)MHC多肽b；(iii)肽P；(iv)一个或更多个多聚化结构域；和/或(v)  $(a-b-P)_n$ 。在一些实施方案中，一个或更多个标记通过链霉抗生物素蛋白-生物素连接附接至  $(a-b-P)_n$ 。

[0043] 一种或更多种另外的标记可以是荧光团标记。在一些实施方案中,荧光团标记是异硫氰酸荧光素、罗丹明、藻红蛋白、藻蓝蛋白、别藻蓝蛋白、邻苯二甲醛、荧光胺;2-(4'-马来酰亚胺基苯胺基)萘-6-磺酸钠盐;5-(((2-碘乙酰基)氨基)乙基)氨基)萘-1-磺酸;苾-1-丁酸;AlexaFluor 350(7-氨基-6-磺酸-4-甲基香豆素-3-乙酸);AMCA(7-氨基-4-甲基香豆素-3-乙酸);7-羟基-4-甲基香豆素-3-乙酸;Marina Blue(6,8-二氟-7-羟基-4-甲基香豆素-3-乙酸);7-二甲基氨基-香豆素-4-乙酸;荧光胺-N-丁胺加合物;7-羟基-香豆素-3-羧酸;CascadeBlue(苾-三磺酸乙酰叠氮化物);Cascade Yellow;Pacific Blue(6,8-二氟-7-羟基香豆素-3-羧酸);7-二乙基氨基-香豆素-3-羧酸;N-(((4-叠氮苯甲酰基)氨基)乙基)-4-氨基-3,6-二硫基-1,8-萘酰亚胺二钾盐;Alexa Fluor 430;3-苾十二酸;8-羟基苾-1,3,6-三磺酸三钠盐;12-(N-(7-硝基苯-2-氧杂-1,3-二唑-4-基)氨基)十二酸;N,N'-二甲基-N-(碘乙酰基)-N'-(7-硝基苯-2-氧杂-1,3-二唑-4-基)乙二胺;Oregon Green 488(二氟羧基荧光素);5-碘乙酰胺荧光素;碘化丙锭-DNA加合物;羧基荧光素、Fluor染料、Pacific Blue<sup>TM</sup>、Pacific Orange<sup>TM</sup>、Cascade Yellow<sup>TM</sup>; **AlexaFluor®** (AF), AF405、AF488、AF500、AF514、AF532、AF546、AF555、AF568、AF594、AF610、AF633、AF635、AF647、AF680、AF700、AF710、AF750、AF800;基于量子点的染料, **QDot®** 纳米晶体 (Invitrogen, MolecularProbs)、**Qdot®**525、**Qdot®**565、**Qdot®**585、**Qdot®**605、**Qdot®**655、**Qdot®**705、**Qdot®**800;DyLight<sup>TM</sup>染料(Pierce) (DL), DL549、DL649、DL680、DL800;荧光素(Flu)或其任何衍生物,诸如FITC;Cy-染料、Cy2、Cy3、Cy3.5、Cy5、Cy5.5、Cy7;荧光蛋白,RPE、PerCp、APC、绿色荧光蛋白;GFP和GFP衍生的突变体蛋白,BFP、CFP、YFP、DsRed、T1、Dimer2、mRFP1、MBanana、mOrange、dTomato、tdTomato、mTangerine、mStrawberry、mCherry;串联染料,RPE-Cy5、RPE-Cy5.5、RPE-Cy7、**RPE-AlexaFluor®** 串联缀合物;RPE-Alexa610、RPE-TxRed、APC-Alexa600、APC-Alexa610、APC-Alexa750、APC-Cy5、APC-Cy5.5或其任组合。

[0044] 在一些实施方案中,一种或更多种另外的标记能够吸收光(例如,发色团和/或染料)。在一些实施方案中,一种或更多种另外的标记能够在激发后发射光,任选地是一种或更多种荧光染料,还任选地一种或更多种荧光染料选自 **AlexaFluor®** (AF) 家族,其包括 **AF®**350、AF405、AF430、AF488、AF500、AF514、AF532、AF546、AF555、AF568、AF594、AF610、AF633、AF635、AF647、AF680、AF700、AF710、AF750和AF800;选自基于量子点(**Qdot®**)的染料家族,其包括 **Qdot®**525、**Qdot®**565、**Qdot®**585、**Qdot®**605、**Qdot®**655、**Qdot®**705、**Qdot®**800;选自DyLight<sup>TM</sup>染料(DL)家族,其包括DL549、DL649、DL680、DL800;选自发荧光小染料家族,其包括FITC、Pacific Blue<sup>TM</sup>、Pacific Orange<sup>TM</sup>、Cascade Yellow<sup>TM</sup>、Marina Blue<sup>TM</sup>、DSred、DSred-2、7-AAD、T0-Pro-3;选自Cy-染料家族,其包括Cy2、Cy3、Cy3.5、Cy5、Cy5.5、Cy7;选自藻胆蛋白(Phycobili Protein)家族,其包括R-藻红蛋白(RPE)、PerCP、别藻蓝蛋白(APC)、B-藻红蛋白、C-藻蓝蛋白;选自荧光蛋白家族,其包括(E)GFP和GFP((增强的)绿色荧光蛋白)衍生的突变体蛋白,BFP、CFP、YFP、DsRed、T1、Dimer2、mRFP1、MBanana、mOrange、dTomato、tdTomato、mTangerine;选自具有PRE的串联染料家族,其包括RPE-Cy5、RPE-Cy5.5、RPE-Cy7、**RPE-AlexaFluor®** 串联缀合物;RPE-Alexa610、RPE-TxRed;选自具有APC的串联染料家族,其包括APC-Alexa600、APC-Alexa610、APC-Alexa750、APC-Cy5、APC-

Cy5.5;选自钙染料家族,其包括Indo-1-Ca<sup>2+</sup>、Indo-2-Ca<sup>2+</sup>。

[0045] 除了a和b之外,MHC多聚体还可以包含一种或更多种另外的多肽。在一些实施方案中,MHC-肽复合物的一种多肽是重链多肽。在一些实施方案中,MHC-肽复合物的一种多肽是b2M多肽。在一些实施方案中,(i)P被化学修饰;(ii)P被聚乙二醇化、磷酸化和/或糖基化;(iii)肽P的氨基酸残基中的一个被另一个氨基酸取代;(iv)a和b两者都是全长肽;(v)a是全长肽;(vi)b是全长肽;(vii)a被截短;(viii)b被截短;(ix)a和b两者都被截短;(x)a被共价连接至b;(xi)a被共价连接至P;(xii)b被共价连接至P;(xiii)a、b和P都被共价连接;(xiv)a被非共价连接至b;(xv)a被非共价连接至P;(xvi)b被非共价连接至P;(xvii)a、b和P都被非共价连接;(xviii)a不被包含在(a-b-P)复合物中;(xix)b不被包含在(a-b-P)复合物中;和/或(xx)P不被包含在(a-b-P)复合物中。在一些实施方案中,MHC多聚体包含一种或更多种增加稳定性的组分(例如,HEG和/或TEG)。

[0046] 方法可以包括:基于受体文库成员的多于一个测序读段,将一个或更多个单细胞中的T细胞受体(TCR)与肽P关联。方法可以包括:测量对于肽P特异性的T细胞的存在、频率、数目、活性和/或状态,从而检测抗原特异性T细胞应答。

[0047] 在一些实施方案中,第一通用序列、第二通用序列、第三通用序列和/或第四通用序列是相同的。在一些实施方案中,第一通用序列、第二通用序列、第三通用序列和/或第四通用序列是不同的。在一些实施方案中,第一通用序列、第二通用序列、第三通用序列和/或第四通用序列包含测序引物的结合位点和/或测序衔接子、其互补序列和/或其部分。在一些实施方案中,测序衔接子包括P5序列、P7序列、其互补序列和/或其部分。在一些实施方案中,测序引物包括读段1测序引物、读段2测序引物、其互补序列和/或其部分。

[0048] 方法可以包括:使(i)条形码化核酸分子、(ii)条形码化受体结合试剂特异性寡核苷酸和(iii)条形码化细胞组分结合试剂特异性寡核苷酸中的一种或更多种与(i)条形码化核酸分子、(ii)条形码化受体结合试剂特异性寡核苷酸和(iii)条形码化细胞组分结合试剂特异性寡核苷酸中的一种或更多种物理分离。在一些实施方案中,物理分离包括使用一种或更多种尺寸选择试剂。在一些实施方案中,多于一种条形码化受体结合试剂特异性寡核苷酸和多于一种条形码化细胞组分结合试剂特异性寡核苷酸被分别扩增。

[0049] 第二通用序列和第三通用序列可以是不同的。在一些实施方案中,第二通用序列与第三通用序列小于约85%相同。在一些实施方案中,第二通用序列包含与SEQ ID NO:1 (GGAGGGAGGTTAGCGAAGGT)至少约85%相同的序列。在一些实施方案中,扩增多于一种条形码化受体结合试剂特异性寡核苷酸或其产物不包括扩增多于一种条形码化细胞组分结合试剂特异性寡核苷酸或其产物。在一些实施方案中,产生测序文库包括产生包含(i)核酸靶文库成员、(ii)细胞组分靶文库成员和/或(iii)受体文库成员的测序混合物。在一些实施方案中,产生测序混合物包括以预定比例混合(i)核酸靶文库成员、(ii)细胞组分靶文库成员和/或(iii)受体文库成员。在一些实施方案中,在产生测序混合物之前,(i)核酸靶文库成员、(ii)细胞组分靶文库成员和/或(iii)受体文库成员在物理上彼此分离。在一些实施方案中,测序混合物包含预定比例的(i)核酸靶文库成员、(ii)细胞组分靶文库成员和/或(iii)受体文库成员。

[0050] 在一些实施方案中,细胞组分靶文库成员与受体文库成员的预定比例是约1:1、1.1:1、1.2:1、1.3:1、1.4:1、1.5:1、1.6:1、1.7:1、1.8:1、1.9:1、2:1、2.5:1、3:1、4:1、5:1、

6:1、7:1、8:1、9:1、10:1、11:1、12:1、13:1、14:1、15:1、16:1、17:1、18:1、19:1、20:1、21:1、22:1、23:1、24:1、25:1、26:1、27:1、28:1、29:1、30:1、31:1、32:1、33:1、34:1、35:1、36:1、37:1、38:1、39:1、40:1、41:1、42:1、43:1、44:1、45:1、46:1、47:1、48:1、49:1、50:1、51:1、52:1、53:1、54:1、55:1、56:1、57:1、58:1、59:1、60:1、61:1、62:1、63:1、64:1、65:1、66:1、67:1、68:1、69:1、70:1、71:1、72:1、73:1、74:1、75:1、76:1、77:1、78:1、79:1、80:1、81:1、82:1、83:1、84:1、85:1、86:1、87:1、88:1、89:1、90:1、91:1、92:1、93:1、94:1、95:1、96:1、97:1、98:1、99:1、100:1、200:1、300:1、400:1、500:1、600:1、700:1、800:1、900:1、1000:1、2000:1、3000:1、4000:1、5000:1、6000:1、7000:1、8000:1、9000:1或10000:1。在一些实施方案中,细胞组分靶文库成员与受体文库成员的预定比例被配置为实现是以下的细胞组分靶文库成员的测序读段与受体文库成员的测序读段的比例:约1:1、1.1:1、1.2:1、1.3:1、1.4:1、1.5:1、1.6:1、1.7:1、1.8:1、1.9:1、2:1、2.5:1、3:1、4:1、5:1、6:1、7:1、8:1、9:1、10:1、11:1、12:1、13:1、14:1、15:1、16:1、17:1、18:1、19:1、20:1、21:1、22:1、23:1、24:1、25:1、26:1、27:1、28:1、29:1、30:1、31:1、32:1、33:1、34:1、35:1、36:1、37:1、38:1、39:1、40:1、41:1、42:1、43:1、44:1、45:1、46:1、47:1、48:1、49:1、50:1、51:1、52:1、53:1、54:1、55:1、56:1、57:1、58:1、59:1、60:1、61:1、62:1、63:1、64:1、65:1、66:1、67:1、68:1、69:1、70:1、71:1、72:1、73:1、74:1、75:1、76:1、77:1、78:1、79:1、80:1、81:1、82:1、83:1、84:1、85:1、86:1、87:1、88:1、89:1、90:1、91:1、92:1、93:1、94:1、95:1、96:1、97:1、98:1、99:1、100:1、200:1、300:1、400:1、500:1、600:1、700:1、800:1、900:1、1000:1、2000:1、3000:1、4000:1、5000:1、6000:1、7000:1、8000:1、9000:1或10000:1。在一些实施方案中,细胞组分靶文库成员的测序读段与受体文库成员的测序读段的比例是约1:1、1.1:1、1.2:1、1.3:1、1.4:1、1.5:1、1.6:1、1.7:1、1.8:1、1.9:1、2:1、2.5:1、3:1、4:1、5:1、6:1、7:1、8:1、9:1、10:1、11:1、12:1、13:1、14:1、15:1、16:1、17:1、18:1、19:1、20:1、21:1、22:1、23:1、24:1、25:1、26:1、27:1、28:1、29:1、30:1、31:1、32:1、33:1、34:1、35:1、36:1、37:1、38:1、39:1、40:1、41:1、42:1、43:1、44:1、45:1、46:1、47:1、48:1、49:1、50:1、51:1、52:1、53:1、54:1、55:1、56:1、57:1、58:1、59:1、60:1、61:1、62:1、63:1、64:1、65:1、66:1、67:1、68:1、69:1、70:1、71:1、72:1、73:1、74:1、75:1、76:1、77:1、78:1、79:1、80:1、81:1、82:1、83:1、84:1、85:1、86:1、87:1、88:1、89:1、90:1、91:1、92:1、93:1、94:1、95:1、96:1、97:1、98:1、99:1、100:1、200:1、300:1、400:1、500:1、600:1、700:1、800:1、900:1、1000:1、2000:1、3000:1、4000:1、5000:1、6000:1、7000:1、8000:1、9000:1或10000:1。在一些实施方案中,细胞组分靶文库成员的测序读段与受体文库成员的测序读段的比例与其中包括以下的方法相比为至少约2倍低:(i) 多于一种条形码化受体结合试剂特异性寡核苷酸和多于一种条形码化细胞组分结合试剂特异性寡核苷酸被一起扩增;(ii) 第二通用序列和第三通用序列相同;和/或(iii) 测序混合物不包含细胞组分靶文库成员与受体文库成员的预定比例。

[0051] 本文的公开内容包括试剂盒。在一些实施方案中,试剂盒包含:多于一种受体检测构建体,其中受体检测构建体包含两种或更多种受体结合试剂,其中受体结合试剂能够与受体特异性结合,并且其中受体检测构建体中的每一种包含受体结合试剂特异性寡核苷酸,所述受体结合试剂特异性寡核苷酸包含用于受体结合试剂的独特受体标识符序列。试剂盒可以包含:多于一种细胞组分结合试剂,其中多于一种细胞组分结合试剂中的每一种包含细胞组分结合试剂特异性寡核苷酸,所述细胞组分结合试剂特异性寡核苷酸包含用于

细胞组分结合试剂的独特标识符序列,并且其中细胞组分结合试剂能够与细胞组分靶特异性结合,其中受体结合试剂特异性寡核苷酸包含第二通用序列,其中细胞组分结合试剂特异性寡核苷酸包含第三通用序列,其中第二通用序列和第三通用序列是不同的。在一些实施方案中,第二通用序列与第三通用序列小于约85%相同。

[0052] 试剂盒可以包含:多于一种寡核苷酸条形码,其中多于一种寡核苷酸条形码中的每一种包含第一通用序列、分子标记和靶结合区,并且其中多于一种寡核苷酸条形码中的至少10种包含不同的分子标记序列。在一些实施方案中,多于一种寡核苷酸条形码与固体支持物关联。在一些实施方案中,靶结合区包含基因特异性序列、寡(dT)序列、随机多聚体或其任何组合。试剂盒可以包含:逆转录酶。在一些实施方案中,逆转录酶包括病毒逆转录酶,并且还任选地病毒逆转录酶是鼠白血病病毒(MLV)逆转录酶和/或Moloney鼠白血病病毒(MMLV)逆转录酶。试剂盒可以包含:缺乏5'至3'外切核酸酶活性和3'至5'外切核酸酶活性中的至少一种的DNA聚合酶(例如,包括Klenow片段)。试剂盒可以包含:缓冲液、盒或两者。试剂盒可以包含:一种或更多种用于逆转录反应和/或扩增反应的试剂。在一些实施方案中,多于一种寡核苷酸条形码各自包含细胞标记。在一些实施方案中,多于一种寡核苷酸条形码的每一种细胞标记包含至少6个核苷酸。在一些实施方案中,多于一种寡核苷酸条形码中与同一固体支持物关联的寡核苷酸条形码包含相同的细胞标记。在一些实施方案中,多于一种寡核苷酸条形码中与不同固体支持物关联的寡核苷酸条形码包含不同的细胞标记。

[0053] 固体支持物可以包括合成颗粒、平坦表面,或其组合。在一些实施方案中,多于一种寡核苷酸条形码中的至少一种寡核苷酸条形码被固定或部分固定在合成颗粒上,或者多于一种寡核苷酸条形码中的至少一种寡核苷酸条形码被包封或部分包封在合成颗粒中。在一些实施方案中,合成颗粒是可破坏的(例如,可破坏的水凝胶颗粒)。在一些实施方案中,合成颗粒包括珠。珠可以是琼脂糖凝胶珠、链霉抗生物素蛋白珠、琼脂糖珠、磁珠、缀合珠、蛋白A缀合珠、蛋白G缀合珠、蛋白A/G缀合珠、蛋白L缀合珠、寡(dT)缀合珠、二氧化硅珠、二氧化硅样珠、抗生物素微珠、抗荧光染料微珠或其任何组合。在一些实施方案中,合成颗粒包含选自以下组成的组的材料:聚二甲基硅氧烷(PDMS)、聚苯乙烯、玻璃、聚丙烯、琼脂糖、明胶、水凝胶、顺磁物质、陶瓷、塑料、玻璃、甲基苯乙烯、丙烯酸聚合物、钛、胶乳、琼脂糖凝胶、纤维素、尼龙、硅酮,以及其任何组合。在一些实施方案中,多于一种寡核苷酸条形码中的每一种寡核苷酸条形码包含接头官能团。在一些实施方案中,合成颗粒包含固体支持物官能团。在一些实施方案中,支持物官能团和接头官能团彼此关联,并且任选地接头官能团和支持物官能团单独地选自以下组成的组:C6、生物素、链霉抗生物素蛋白、一种或更多种伯胺、一种或更多种醛、一种或更多种酮及其任何组合。

[0054] 试剂盒可以包含:一种或更多种包含第一通用序列、第二通用序列和/或第三通用序列的引物。在一些实施方案中,第二通用序列包含与SEQ ID NO:1 (GGAGGGAGGTTAGCGAAGGT)至少约85%相同的序列。在一些实施方案中,靶结合区包含多(dT)区,并且其中细胞组分结合试剂特异性寡核苷酸包含多(dA)区。在一些实施方案中,细胞组分结合试剂特异性寡核苷酸包含与多(dA)区相邻的对齐序列。在一些实施方案中,(a)对齐序列包含鸟嘌呤、胞嘧啶、胸腺嘧啶、尿嘧啶或其组合;(b)对齐序列包含多(dT)序列、多(dG)序列、多(dC)序列、多(dU)序列或其组合;和/或(c)对齐序列位于多(dA)区的5'。在

一些实施方案中,细胞组分结合试剂特异性寡核苷酸通过接头与细胞组分结合试剂关联。在一些实施方案中,细胞组分结合试剂特异性寡核苷酸被配置为能够从细胞组分结合试剂脱离。在一些实施方案中,接头包含碳链。在一些实施方案中,碳链包含2-30个碳,例如12个碳。在一些实施方案中,接头包含5' 氨基修饰物C12 (5AmMC12) 或其衍生物。

[0055] 在一些实施方案中,细胞组分靶包括蛋白靶。在一些实施方案中,细胞组分结合试剂包含抗体或其片段。在一些实施方案中,细胞组分靶包括糖类、脂质、蛋白、细胞外蛋白、细胞表面蛋白、细胞标志物、B细胞受体、T细胞受体、主要组织相容性复合体、肿瘤抗原、受体、细胞内蛋白或其任何组合。在一些实施方案中,细胞组分结合试剂包含抗体或其片段。在一些实施方案中,细胞组分靶在细胞表面上。

[0056] 在一些实施方案中,受体结合试剂特异性寡核苷酸包含第三分子标记,和/或细胞组分结合试剂特异性寡核苷酸包含第二分子标记。在一些实施方案中,受体结合试剂特异性寡核苷酸的长度为约10个核苷酸至约100个核苷酸。在一些实施方案中,受体包括分化簇(CD)分子。在一些实施方案中,受体包括免疫受体(例如,T细胞受体(TCR))。在一些实施方案中,独特受体标识符序列的长度为约3个核苷酸至约100个核苷酸。在一些实施方案中,靶结合区包含多(dT)区,并且其中受体结合试剂特异性寡核苷酸包含多(dA)区。在一些实施方案中,受体结合试剂特异性寡核苷酸包含与多(dA)区相邻的对齐序列。在一些实施方案中,(a)对齐序列包含鸟嘌呤、胞嘧啶、胸腺嘧啶、尿嘧啶或其组合;(b)对齐序列包含多(dT)序列、多(dG)序列、多(dC)序列、多(dU)序列或其组合;和/或(c)对齐序列位于多(dA)区的5'。在一些实施方案中,受体结合试剂特异性寡核苷酸通过接头与受体结合试剂关联。在一些实施方案中,受体结合试剂特异性寡核苷酸被配置为能够从受体结合试剂脱离。在一些实施方案中,接头包含碳链。在一些实施方案中,碳链包含2-30个碳,例如12个碳。在一些实施方案中,接头包含5' 氨基修饰物C12 (5AmMC12) 或其衍生物。

[0057] 在一些实施方案中,多于一种受体检测构建体包括一种或更多种MHC多聚体,并且其中两种或更多种受体结合试剂包含两种或更多种MHC-肽复合物。在一些实施方案中,MHC多聚体包含 $(a-b-P)_n$ ,其中 $n>1$ ,其中多肽a和多肽b一起形成能够结合肽P的功能性MHC蛋白,并且 $(a-b-P)$ 是当肽P与功能性MHC蛋白结合时形成的MHC-肽复合物。在一些实施方案中,MHC多聚体的每种MHC-肽复合物与一个或更多个多聚化结构域关联。在一些实施方案中,多于一种受体检测构建体包含2种MHC多聚体、3种MHC多聚体、4种MHC多聚体、5种MHC多聚体、6种MHC多聚体、7种MHC多聚体、8种MHC多聚体、9种MHC多聚体、10种MHC多聚体、11种MHC多聚体、12种MHC多聚体、13种MHC多聚体、14种MHC多聚体、15种MHC多聚体、16种MHC多聚体、17种MHC多聚体、18种MHC多聚体、19种MHC多聚体或20种MHC多聚体。在一些实施方案中,所述MHC多聚体的每种MHC-肽复合物的个体抗原肽P是相同或不同的。在一些实施方案中,多于一种受体检测构建体包括一种或更多种阴性对照MHC多聚体。在一些实施方案中,一种或更多种阴性对照MHC多聚体,其中每种MHC多聚体包含阴性对照肽P。在一些实施方案中,所述阴性对照肽P是无义肽。在一些实施方案中,所述一种或更多种阴性对照MHC多聚体是空MHC多聚体。在一些实施方案中,多于一种受体检测构建体包含一种或更多种阳性对照MHC多聚体,其中每种MHC多聚体包含阳性对照肽P。在一些实施方案中,包含 $(a-b-P)_n$ 的所述一种或更多种MHC多聚体的n的值是 $1<n\leq 1000$ 。在一些实施方案中,n的值在2-3、3-4、4-5、5-6、6-7、7-8、8-9、9-10、10-11、11-12、12-13、13-14、14-15、15-16、16-17、17-18、18-19、

19-20、20-21、21-22、22-23、23-24、24-25、25-26、26-27、27-28、28-29、29-30、30-35、35-40、40-45、45-50、50-55、55-60、60-65、65-70、75-80、80-85、85-90、90-95、95-100、100-110、110-120、120-130、130-140、140-150、150-160、160-170、170-180、180-190、190-200、200-225、225-250、250-275、275-300、300-325、325-350、350-375、375-400、400-450、450-500、500-550、550-600、600-650、650-700、700-750、750-800、800-850、850-900、900-950或950-1000之间。在一些实施方案中,所述MHC蛋白是MHC I类、MHC II类、MR1、CD1或MHC样分子。在一些实施方案中,所述MHC蛋白是MHC I类,并且抗原肽P选自由与MHC I类、MHC II类、MR1、CD1或MHC样分子结合的8-mer、9-mer、10-mer、11-mer和12-mer肽组成的组。

[0058] 在一些实施方案中,多聚体MHC的每种MHC蛋白或MHC-肽复合物中的一个或更多个和一个或更多个多聚化结构域之间的关联是共价关联和/或非共价关联。在一些实施方案中,一个或更多个多聚化结构域包含一个或更多个多聚化结构域连接体分子。在一些实施方案中,多聚体MHC的每种MHC蛋白或MHC-肽复合物中的一个或更多个包含一个或更多个MHC连接体分子。在一些实施方案中,一个或更多个多聚化结构域连接体分子包括一个或更多个链霉抗生物素蛋白和/或一个或更多个抗生物素蛋白,或其任何衍生物。在一些实施方案中,一个或更多个链霉抗生物素蛋白包括一个或更多个四聚体链霉抗生物素蛋白变体或一个或更多个单体链霉抗生物素蛋白变体。一个或更多个MHC连接体分子可以是生物素。在一些实施方案中,MHC多聚体的每种MHC蛋白或MHC-肽复合物中的一个或更多个通过链霉抗生物素蛋白-生物素连接或抗生物素蛋白-生物素连接与一个或更多个多聚化结构域关联。在一些实施方案中,MHC多聚体的每种MHC蛋白或MHC-肽复合物中的一个或更多个通过接头部分与一个或更多个多聚化结构域关联。在一些实施方案中,MHC多聚体的每种MHC蛋白或MHC-肽复合物中的一个或更多个通过天然二聚化和/或蛋白-蛋白相互作用与一个或更多个多聚化结构域关联。在一些实施方案中,天然二聚化和/或蛋白-蛋白相互作用选自由以下组成的组:亮氨酸拉链例如AP-1的亮氨酸拉链结构域、Fos/Jun相互作用、基于酸性/碱性卷曲螺旋结构的相互作用(例如螺旋)、抗体/抗原相互作用、多核苷酸-多核苷酸相互作用例如DNA/DNA、DNA/PNA、DNA/RNA、PNA/PNA、LNA/DNA、合成分子-合成分子相互作用和蛋白-小分子相互作用、IgG二聚体蛋白、IgM多价蛋白、螯合物/金属离子结合螯合物、strep免疫球蛋白、抗体(单克隆、多克隆和重组的)、抗体片段及其衍生物、hexa-his(金属螯合部分)、hexa-hat GST(谷胱甘肽S-转移酶)谷胱甘肽亲和力、钙调蛋白结合肽(CBP)、Strep-标签、纤维素结合结构域、麦芽糖结合蛋白、S-肽标签、壳多糖结合标签、免疫反应性表位、表位标签、E2标签、HA表位标签、Myc表位、FLAG表位、AU1表位和AU5表位、Glu-Glu表位、KT3表位、IRS表位、Btag表位、蛋白激酶-C表位、VSV表位、介导与包括糖类、脂质和蛋白的多种化合物结合的凝集素,例如Con A(直生刀豆(*Cannalia ensiformis*))或WGA(麦胚凝集蛋白)和四联凝集素或蛋白A或蛋白G(抗体亲和力)。

[0059] 在一些实施方案中,一个或更多个多聚化结构域包含(i)一个或更多个支架;(ii)一个或更多个载体;(iii)至少一个支架和至少一个载体;和/或(iv)一个或更多个任选取代的有机分子。在一些实施方案中,任选取代的有机分子包含一种或更多种官能化的环状结构。在一些实施方案中,官能化的环状结构包含苯环。在一些实施方案中,任选取代的有机分子包括支架分子,所述支架分子包含至少三个反应性基团,或至少三个适于非共价衔接的位点。在一些实施方案中,一个或更多个多聚化结构域包含一种或更多种生物细胞和/

或细胞样结构。在一些实施方案中，一个或更多个多聚化结构域包含一种或更多种膜。在一些实施方案中，一种或更多种膜包括脂质体或胶束。在一些实施方案中，一个或更多个多聚化结构域包含一种或更多种聚合物。在一些实施方案中，一种或更多种合成聚合物。在一些实施方案中，一种或更多种聚合物选自自由多糖组成的组。

[0060] 在一些实施方案中，多糖包含一个或更多个右旋糖酐部分。在一些实施方案中，一个或更多个右旋糖酐部分(i)与一个或更多个MHC肽复合物共价附接；(ii)与一个或更多个MHC肽复合物非共价附接；和/或(iii)被修饰。在一些实施方案中，一个或更多个右旋糖酐部分包含一个或更多个氨基-右旋糖酐。在一些实施方案中，一个或更多个氨基-右旋糖酐用二乙烯基砜修饰。在一些实施方案中，一个或更多个右旋糖酐部分包含一个或更多个具有约1,000Da至50,000Da的分子量的右旋糖酐。在一些实施方案中，一个或更多个右旋糖酐部分包含一个或更多个具有约50,000至150,000Da的分子量的右旋糖酐。在一些实施方案中，一个或更多个右旋糖酐部分包含一个或更多个具有约150,000-270,000Da的分子量的右旋糖酐。在一些实施方案中，一个或更多个右旋糖酐部分是线性的和/或支链的。

[0061] 在一些实施方案中，一种或更多种合成聚合物包括PNA、聚酰胺、PEG或其任何组合。在一些实施方案中，一个或更多个多聚化结构域包含一种或更多种选自自由以下组成的组的实体：IgG结构域、卷曲螺旋多肽结构、DNA双链体、核酸双链体、PNA-PNA、PNA-DNA和DNA-RNA。在一些实施方案中，一个或更多个多聚化结构域包含抗体。在一些实施方案中，抗体选自自由以下组成的组：多克隆抗体、单克隆抗体、IgA、IgG、IgM、IgD、IgE、IgG1、IgG2、IgG3、IgG4、IgA1、IgA2、IgM1、IgM2、人源化抗体、人源化单克隆抗体、嵌合抗体、小鼠抗体、大鼠抗体、兔抗体、人类抗体、骆驼抗体、绵羊抗体、工程化人类抗体、表位聚焦抗体、激动剂抗体、拮抗剂抗体、中和抗体、天然存在的抗体、分离的抗体、单价抗体、双特异性抗体、三特异性抗体、多特异性抗体、异源缀合抗体、免疫缀合物、免疫脂质体、标记的抗体、抗体片段、结构域抗体、纳米抗体、微抗体、大抗体、双体和融合抗体。在一些实施方案中，一个或更多个多聚化结构域包含一种或更多种有机小支架分子或有机小分子。在一些实施方案中，一种或更多种有机小分子包括一种或更多种类固醇、一种或更多种肽和/或一种或更多种芳族有机分子。在一些实施方案中，一种或更多种芳族有机分子包含一种或更多种双环结构、一种或更多种多环结构或一种或更多种单环结构；任选地所述一种或更多种单环结构包括一种或更多种任选地官能化或取代的苯环。在一些实施方案中，一个或更多个多聚化结构域包含一种或更多种：(i)能够聚合的单体分子；(ii)生物聚合物，诸如一种或更多种蛋白；(iii)小分子支架；(iv)超分子结构，诸如一种或更多种纳米团簇；和/或(v)蛋白复合物。

[0062] 在一些实施方案中，一个或更多个多聚化结构域包含一种或更多种珠。在一些实施方案中，一种或更多种珠选自自由以下组成的组：带有亲电基团例如二乙烯基砜活化的多糖的珠、已经用甲苯磺酰基活化的酯官能化的聚苯乙烯珠、用甲苯磺酰基活化的酯官能化的磁性聚苯乙烯珠，以及其中MHC复合体已经通过MHC复合体中包含的亲核体与珠的亲电体的反应而共价固定至这些珠的珠。在一些实施方案中，一种或更多种珠选自自由以下组成的组：琼脂糖凝胶珠、聚丙烯酰胺葡聚糖珠、聚苯乙烯珠、琼脂糖珠、多糖珠、聚氨基甲酸酯珠和可以悬浮在水性缓冲液中的任何其他种类的珠。在一些实施方案中，多聚化结构域包含一种或更多种选自自由以下组成的组的化合物：琼脂糖珠、琼脂糖凝胶珠、树脂珠、玻璃珠、多孔玻璃珠、涂覆疏水聚合物的玻璃颗粒、壳聚糖涂覆珠、SH珠、胶乳珠、球形胶乳珠、等位基

因型珠、SPA珠、基于PEG的树脂、PEG涂覆珠、PEG包封珠、聚苯乙烯珠、磁性聚苯乙烯珠、谷胱甘肽琼脂糖珠、磁珠、顺磁珠、蛋白A和/或蛋白G琼脂糖凝胶珠、活化羧酸珠、宏观珠、微观珠、不溶性树脂珠、基于二氧化硅的树脂、纤维素树脂、交联琼脂糖珠、聚苯乙烯珠、交联聚丙烯酰胺树脂、具有铁芯的珠、金属珠、dynabead、用NHS活化的聚甲基丙烯酸甲酯珠、链霉抗生物素蛋白-琼脂糖珠、聚丙烯、聚乙烯、右旋糖酐、尼龙、淀粉酶、天然纤维素和改性纤维素、硝化纤维素、聚丙烯酰胺、辉长岩、磁铁矿、聚合物、寡聚物、非重复部分、聚乙二醇(PEG)、单甲氧基-PEG、单-(C<sub>1</sub>-C<sub>10</sub>)烷氧基-PEG、芳氧基-PEG、聚(N-乙基吡咯烷酮)PEG、三氟乙磺酰氯单甲氧基PEG、PEG丙醛、双琥珀酰亚胺基碳酸PEG、二乙烯基苯交联聚苯乙烯珠、丙二醇均聚物、聚环氧丙烷/环氧乙烷共聚物、聚氧乙基化多元醇(例如,甘油)、聚乙烯醇、右旋糖酐、氨基右旋糖酐、基于糖类的聚合物、交联右旋糖酐珠、多糖珠、聚氨基甲酸酯珠、二乙烯基砜活化多糖、已经用甲苯磺酰基活化的酯官能化的聚苯乙烯珠、用甲苯磺酰基活化的酯官能化的磁性聚苯乙烯珠、链霉抗生物素蛋白珠、链霉抗生物素蛋白单体涂覆的珠、链霉抗生物素蛋白四聚体涂覆的珠、链霉抗生物素蛋白涂覆的Compel磁珠、抗生物素蛋白涂覆的珠、dextramer涂覆的珠、二乙烯基砜活化的右旋糖酐、羧酸修饰的珠、胺修饰的珠、抗体涂覆的珠、纤维素珠、接枝共聚珠、聚丙烯酰胺珠、任选地与N-N'-双丙烯酰基乙二胺交联的二甲基丙烯酰胺珠、中空纤维膜、荧光珠、胶原-琼脂糖珠、明胶珠、胶原-明胶珠、胶原-纤连蛋白-明胶珠、胶原珠、壳聚糖珠、胶原-壳聚糖珠、基于蛋白的珠、水凝胶珠、半纤维素、烷基纤维素、羟烷基纤维素、羧甲基纤维素、磺基乙基纤维素、淀粉、木聚糖、支链淀粉、软骨素、透明质酸、肝素、瓜尔胶、黄原胶、甘露聚糖、半乳甘露聚糖、壳多糖和壳聚糖。在一些实施方案中,一个或多个多聚化结构域包含二聚化结构域、三聚化结构域、四聚化结构域、五聚化结构域或六聚化结构域。

[0063] 在一些实施方案中,一个或多个多聚化结构域包含附接一个或多个支架的聚合物结构。在一些实施方案中,聚合物结构包含多糖。在一些实施方案中,多糖包含一个或多个右旋糖酐部分。在一些实施方案中,一个或多个多聚化结构域包含聚酰胺、聚乙二醇、多糖、琼脂糖凝胶或其任何组合。在一些实施方案中,一个或多个多聚化结构域包含羧甲基右旋糖酐、右旋糖酐聚醛、羧甲基右旋糖酐内酯、环糊精或其任何组合。在一些实施方案中,一个或多个多聚化结构域具有小于1,000Da、1,000Da至小于10,000Da、10,000Da至小于100,000Da、100,000Da至小于1,000,000Da、超过1,000,000Da的分子量。在一些实施方案中,所述一种或更多种MHC多聚体还包含一个或多个选自由以下组成的组的支架、载体和/或接头:链霉抗生物素蛋白(SA)和抗生物素蛋白及其衍生物、生物素、免疫球蛋白、抗体(单克隆、多克隆和重组的)、抗体片段及其衍生物、AP-1的亮氨酸拉链结构域(jun和fos)、hexa-his(金属螯合部分)、hexa-hat GST(谷胱甘肽S-转移酶)谷胱甘肽亲和力、钙调蛋白结合肽(CBP)、Strep-标签、纤维素结合结构域、麦芽糖结合蛋白、S-肽标签、壳多糖结合标签、免疫反应性表位、表位标签、E2标签、HA表位标签、Myc表位、FLAG表位、AU1和AU5表位、Glu-Glu表位、KT3表位、IRS表位、Btag表位、蛋白激酶-C表位、VSV表位、介导与包括糖类、脂质和蛋白的多种化合物结合的凝集素,例如Con A(直生刀豆)或WGA(麦胚凝集蛋白)和四联凝集素或蛋白A或蛋白G(抗体亲和力)。在一些实施方案中,所述MHC多聚体包含由多聚化结构域连接部分连接的多于一个相同或不同的多聚化结构域。在一些实施方案中,所述MHC多聚体包含连接至第二多聚化结构域的第一多聚化结构域。

[0064] 一种或更多种MHC多聚体可以包含一个或更多个标记。在一些实施方案中,一个或更多个标记包括受体结合试剂特异性寡核苷酸。在一些实施方案中,一个或更多个标记包括受体结合试剂特异性寡核苷酸和一种或更多种另外的标记。在一些实施方案中,一种或更多种另外的标记包括肽标记、荧光团标记、重金属标记、同位素标记、放射标记、放射性核素、稳定同位素、同位素链和单原子、化学发光标记、生物发光标记、放射性标记、酶标记、DNA荧光染色剂、镧系元素、离子载体、与特定离子结合的螯合化合物或其任何组合。在一些实施方案中,一个或更多个标记包括共价附接的标记和/或非共价附接的标记。在一些实施方案中,一个或更多个标记与以下附接:(i)MHC多肽a;(ii)MHC多肽b;(iii)肽P;(iv)一个或更多个多聚化结构域;和/或(v)  $(a-b-P)_n$ 。在一些实施方案中,一个或更多个标记通过链霉抗生物素蛋白-生物素连接附接至  $(a-b-P)_n$ 。

[0065] 在一些实施方案中,一种或更多种另外的标记是荧光团标记。在一些实施方案中,荧光团标记选自包括以下的组:异硫氰酸荧光素、罗丹明、藻红蛋白、藻蓝蛋白、别藻蓝蛋白、邻苯二甲醛、荧光胺;2-(4'-马来酰亚胺基苯胺基)萘-6-磺酸钠盐;5-(((2-碘乙酰基)氨基)乙基)氨基)萘-1-磺酸;苾-1-丁酸;AlexaFluor 350(7-氨基-6-磺酸-4-甲基香豆素-3-乙酸);AMCA(7-氨基-4-甲基香豆素-3-乙酸);7-羟基-4-甲基香豆素-3-乙酸;Marina Blue(6,8-二氟-7-羟基-4-甲基香豆素-3-乙酸);7-二甲基氨基-香豆素-4-乙酸;荧光胺-N-丁胺加合物;7-羟基-香豆素-3-羧酸;CascadeBlue(苾-三磺酸乙酰叠氮化物);Cascade Yellow;Pacific Blue(6,8二氟-7-羟基香豆素-3-羧酸);7-二乙基氨基-香豆素-3-羧酸;N-(((4-叠氮苯甲酰基)氨基)乙基)-4-氨基-3,6-二硫基-1,8-萘酰亚胺二钾盐;Alexa Fluor 430;3-苾十二酸;8-羟基苾-1,3,6-三磺酸三钠盐;12-(N-(7-硝基苯-2-氧杂-1,3-二唑-4-基)氨基)十二酸;N,N'-二甲基-N-(碘乙酰基)-N'-(7-硝基苯-2-氧杂-1,3-二唑-4-基)乙二胺;Oregon Green 488(二氟羧基荧光素);5-碘乙酰胺荧光素;碘化丙锭-DNA加合物;羧基荧光素、Fluor染料、Pacific Blue<sup>TM</sup>、Pacific Orange<sup>TM</sup>、Cascade Yellow<sup>TM</sup>;AlexaFluor® (AF), AF405、AF488、AF500、AF514、AF532、AF546、AF555、AF568、AF594、AF610、AF633、AF635、AF647、AF680、AF700、AF710、AF750、AF800;基于量子点的染料,QDot®纳米晶体(Invitrogen,MolecularProbs)、Qdot®525、Qdot®565、Qdot®585、Qdot®605、Qdot®655、Qdot®705、Qdot®800;DyLight<sup>TM</sup>染料(Pierce)(DL),DL549、DL649、DL680、DL800;荧光素(Flu)或其任何衍生物,诸如FITC;Cy-染料、Cy2、Cy3、Cy3.5、Cy5、Cy5.5、Cy7;荧光蛋白,RPE、PerCp、APC、绿色荧光蛋白;GFP和GFP衍生的突变体蛋白,BFP、CFP、YFP、DsRed、T1、Dimer2、mRFP1、MBanana、mOrange、dTomato、tdTomato、mTangerine、mStrawberry、mCherry;串联染料,RPE-Cy5、RPE-Cy5.5、RPE-Cy7、RPE-AlexaFluor®串联缀合物;RPE-Alexa610、RPE-TxRed、APC-Alexa600、APC-Alexa610、APC-Alexa750、APC-Cy5和APC-Cy5.5。

[0066] 在一些实施方案中,一种或更多种另外的标记能够吸收光(例如,发色团和/或染料)。在一些实施方案中,一种或更多种另外的标记能够在激发后发射光,任选地是一种或更多种荧光染料,还任选地一种或更多种荧光染料选自AlexaFluor® (AF)家族,其包括AF®350、AF405、AF430、AF488、AF500、AF514、AF532、AF546、AF555、AF568、AF594、AF610、

AF633、AF635、AF647、AF680、AF700、AF710、AF750和AF800；选自基于量子点(Qdot®)的染料家族，其包括Qdot®525、Qdot®565、Qdot®585、Qdot®605、Qdot®655、Qdot®705、Qdot®800；选自DyLight™染料(DL)家族，其包括DL549、DL649、DL680、DL800；选自发荧光小染料家族，其包括FITC、Pacific Blue™、Pacific Orange™、Cascade Yellow™、Marina Blue™、DSred、DSred-2、7-AAD、T0-Pro-3；选自Cy-染料家族，其包括Cy2、Cy3、Cy3.5、Cy5、Cy5.5、Cy7；选自藻胆蛋白家族，其包括R-藻红蛋白(RPE)、PerCP、别藻蓝蛋白(APC)、B-藻红蛋白、C-藻蓝蛋白；选自荧光蛋白家族，其包括(E)GFP和GFP((增强的)绿色荧光蛋白)衍生的突变体蛋白，BFP、CFP、YFP、DsRed、T1、Dimer2、mRFP1、MBanana、mOrange、dTomato、tdTomato、mTangerine；选自具有PRE的串联染料家族，其包括RPE-Cy5、RPE-Cy5.5、RPE-Cy7、RPE-AlexaFluor®串联缀合物；RPE-Alexa610、RPE-TxRed；选自具有APC的串联染料家族，其包括APC-Alexa600、APC-Alexa610、APC-Alexa750、APC-Cy5、APC-Cy5.5；选自钙染料家族，其包括Indo-1-Ca2+、Indo-2-Ca2+。

[0067] 在一些实施方案中，除了a和b之外，MHC多聚体还包含一种或更多种另外的多肽。在一些实施方案中，MHC-肽复合物的一种多肽是重链多肽。在一些实施方案中，MHC-肽复合物的一种多肽是b2M多肽。在一些实施方案中，(i) P被化学修饰；(ii) P被聚乙二醇化、磷酸化和/或糖基化；(iii) 肽P的氨基酸残基中的一个被另一个氨基酸取代；(iv) a和b两者都是全长肽；(v) a是全长肽；(vi) b是全长肽；(vii) a被截短；(viii) b被截短；(ix) a和b两者都被截短；(x) a被共价连接至b；(xi) a被共价连接至P；(xii) b被P共价连接至P；(xiii) a、b和P都被共价连接；(xiv) a被非共价连接至b；(xv) a被非共价连接至P；(xvi) b被非共价连接至P；(xvii) a、b和P都被非共价连接；(xviii) a不被包含在(a-b-P)复合物中；(xix) b不被包含在(a-b-P)复合物中；和/或(xx) P不被包含在(a-b-P)复合物中。在一些实施方案中，MHC多聚体包含一种或更多种增加稳定性的组分(例如，HEG和/或TEG)。

[0068] 附图简述

[0069] 图1图示了非限制性示例性条形码。

[0070] 图2示出条形码化和数字计数的非限制性示例性工作流程。

[0071] 图3是示出用于从多于一个靶产生在3'末端条形码化的靶的索引文库的非限制性示例性过程的示意图。

[0072] 图4A-图4B描绘了在不使用独特引物(图4A)和使用独特引物(图4B)的情况下产生的AbSeq和Dextramer测序文库的非限制性示例性示意图。

[0073] 图5描绘了本文提供的组合物和方法的非限制性示例性示意工作流程。

[0074] 图6描绘了包含dCODE®和AbSeq™文库PCR扩增子两者(约170bp)的非限制性示例性生物分析仪迹线。

[0075] 图7描绘了包含dCODE PCR2产物(约190bp)的非限制性示例性生物分析仪迹线。

[0076] 图8描绘了包含索引PCR3 dCODE文库产物(约285bp)的非限制性示例性生物分析仪迹线。

[0077] 图9A-图9F描绘了Immudex dCODE和BD Rhapsody系统单细胞测序的非限制性示例性实验工作流程和分子机制。图9A描绘了与BD Rhapsody系统相容的非限制性示例性Immudex dCODE Dextramer设计。图9B描绘了使用dCODE Dextramer的非限制性的示例性BD

Rhapsody系统靶向工作流程。图9C描绘了从盒(cartridge)回收的具有杂交的mRNA和dCODE的非限制性示例性珠。图9D-图9F描绘了非限制性示例性实验设计:用3种Dextramer染色的人类PBMC。

[0078] 图10A-图10B示出了与检测不同于阴性对照的特定细胞群体上的dextramer相关的非限制性示例性数据。

[0079] 图11A-图11B示出了与EBV+抗原特异性T细胞中基因表达分析相关的非限制性示例性数据。

[0080] 图12A-图12D描绘了非限制性示例性生物分析仪迹线。图12A描绘了Ri0dCODE®和AbSeq™文库PCR1扩增子两者(峰166bp)的生物分析仪迹线。图12B描绘了Ri0dCODE®文库PCR2扩增子(峰约190bp)的生物分析仪迹线。图12C-图12D描绘了Abseq(图12C;峰265bp)和Ri0 dCODEdextramer(图12D;峰约290bp)的最终文库(索引PCR)制备物的生物分析仪迹线。

[0081] 图13A-图13B描绘了dCODE Dextramer®(图13A)和BD® AbSeq(图13B;4-plex,高表达者(high expressor))的非限制性示例性饱和曲线。

[0082] 详述

[0083] 以下详述中参考了形成本文的一部分的附图。在附图中,除非上下文另外指示,否则相似的符号通常标识相似的组成部分。在详述、附图和权利要求书中描述的说明性实施方案不意味着是限制性的。在不脱离本文呈现的主题的精神或范围的情况下,可以利用其他实施方案,并且可以做出其他改变。将容易理解的是,如本文一般描述的以及附图中图示的本公开内容的方面能够以各种不同的配置来布置、替换、组合、分离和设计,所有这些都都在本文中明确设想并且构成本公开内容的一部分。

[0084] 本文提及的所有专利、公开的专利申请、其他出版物和来自GenBank的序列以及其他数据库关于相关技术通过引用以其整体并入。

[0085] 对少量核酸(例如信使核糖核苷酸(mRNA)分子)进行定量对于确定例如在不同发育阶段或在不同环境条件下在细胞中表达的基因是临床上重要的。然而,确定核酸分子(例如,mRNA分子)的绝对数目也可以是非常具有挑战性的,尤其是当分子数目非常小时。确定样品中分子的绝对数目的一种方法是数字聚合酶链式反应(PCR)。理想地,PCR在每个循环中产生相同拷贝的分子。然而,PCR可具有缺点,使得每个分子以随机概率复制,且此概率根据PCR循环和基因序列而变化,这导致扩增偏倚和不准确的基因表达测量。具有独特分子标记(molecular labels,也称为分子索引(molecular indexes,MI))的随机条形码可以用于计数分子数目和校正扩增偏倚。诸如Precise™测定(Cellular Research, Inc. (Palo Alto, CA))和Rhapsody™测定(Becton, Dickinson and Company (Franklin Lakes, NJ))的随机条形码化可以通过使用分子标记(ML)在逆转录(RT)期间标记mRNA来纠正由PCR和文库制备步骤引起的偏倚。

[0086] Precise™测定可以利用具有在多(T)寡核苷酸上的大量(例如6561种至65536种)独特分子标记序列的随机条形码的非耗尽性池(non-depleting pool),以在RT步骤期间与样品中的所有多(A)-mRNA杂交。随机条形码可以包含通用PCR引发位点。在RT期间,靶基因分子与随机条形码随机反应。每一种靶分子可以与随机条形码杂交,导致产生随机条形码化的互补核糖核苷酸(cDNA)分子。在标记后,可以将来自微孔板微孔的随机条形码化cDNA

分子汇集到单个管中用于PCR扩增和测序。可以分析原始测序数据以产生读段的数目、具有独特分子标记序列的随机条形码的数目以及mRNA分子的数目。

[0087] 本文的公开内容包括方法。在一些实施方案中,方法包括使多于一种受体检测构建体与多于一个细胞接触以形成与受体检测构建体关联的第一多于一个细胞,其中多于一个细胞包含多于一种细胞组分靶、和核酸靶的拷贝,其中多于一个细胞中的一个或多个细胞包含受体结合试剂能够特异性结合的受体,并且其中多于一种受体检测构建体中的每一种包含两种或更多种受体结合试剂和受体结合试剂特异性寡核苷酸,所述受体结合试剂特异性寡核苷酸包含用于受体结合试剂的独特受体标识符序列。方法可以包括:使多于一种细胞组分结合试剂与和受体检测构建体关联的多于一个细胞接触以形成第二多于一个细胞,其中多于一种细胞组分结合试剂中的每一种包含细胞组分结合试剂特异性寡核苷酸,所述细胞组分结合试剂特异性寡核苷酸包含用于细胞组分结合试剂的独特标识符序列,并且其中细胞组分结合试剂能够与多于一种细胞组分靶中的至少一种特异性结合。方法可以包括:用多于一种寡核苷酸条形码对细胞组分结合试剂特异性寡核苷酸进行条形码化以产生多于一种条形码化细胞组分结合试剂特异性寡核苷酸,所述多于一种条形码化细胞组分结合试剂特异性寡核苷酸各自包含与独特标识符序列的至少一部分互补的序列。方法可以包括:用多于一种寡核苷酸条形码对受体结合试剂特异性寡核苷酸进行条形码化以产生多于一种条形码化受体结合试剂特异性寡核苷酸,所述多于一种条形码化受体结合试剂特异性寡核苷酸各自包含与独特受体标识符序列的至少一部分互补的序列。方法可以包括:用多于一种寡核苷酸条形码对核酸靶的拷贝进行条形码化以产生多于一种条形码化核酸分子,所述多于一种条形码化核酸分子各自包含与核酸靶的至少一部分互补的序列。方法可以包括:产生包含多于一种核酸靶文库成员、多于一种细胞组分靶文库成员和多于一种受体文库成员的测序文库,其中产生测序文库包括:将测序衔接子附接至多于一种条形码化核酸分子或其产物,以产生多于一种核酸靶文库成员;以及将测序衔接子附接至多于一种条形码化细胞组分结合试剂特异性寡核苷酸或其产物,以产生多于一种细胞组分靶文库成员;以及将测序衔接子附接至多于一种条形码化受体结合试剂特异性寡核苷酸或其产物,以产生多于一种受体文库成员。方法可以包括:获得包含核酸靶文库成员的多于一个测序读段、细胞组分靶文库成员的多于一个测序读段和受体文库成员的多于一个测序读段的测序数据。

[0088] 本文的公开内容包括方法。在一些实施方案中,方法包括:使多于一种受体检测构建体与多于一个细胞接触以形成与受体检测构建体关联的第一多于一个细胞,其中多于一个细胞包含多于一种细胞组分靶,其中多于一个细胞中的一个或多个细胞包含受体结合试剂能够特异性结合的受体,并且其中多于一种受体检测构建体中的每一种包含两种或更多种受体结合试剂和受体结合试剂特异性寡核苷酸,所述受体结合试剂特异性寡核苷酸包含用于受体结合试剂的独特受体标识符序列。方法可以包括:使多于一种细胞组分结合试剂与和受体检测构建体关联的多于一个细胞接触以形成第二多于一个细胞,其中多于一种细胞组分结合试剂中的每一种包含细胞组分结合试剂特异性寡核苷酸,所述细胞组分结合试剂特异性寡核苷酸包含用于细胞组分结合试剂的独特标识符序列,并且其中细胞组分结合试剂能够与多于一种细胞组分靶中的至少一种特异性结合。方法可以包括:用多于一种寡核苷酸条形码对细胞组分结合试剂特异性寡核苷酸进行条形码化以产生多于一种条形

码化细胞组分结合试剂特异性寡核苷酸,所述多于一种条形码化细胞组分结合试剂特异性寡核苷酸各自包含与独特标识符序列的至少一部分互补的序列。方法可以包括:用多于一种寡核苷酸条形码对受体结合试剂特异性寡核苷酸进行条形码化以产生多于一种条形码化受体结合试剂特异性寡核苷酸,所述多于一种条形码化受体结合试剂特异性寡核苷酸各自包含与独特受体标识符序列的至少一部分互补的序列。方法可以包括:产生包含多于一种细胞组分靶文库成员和多于一种受体文库成员的测序文库,其中产生测序文库包括:将测序衔接子附接至多于一种条形码化细胞组分结合试剂特异性寡核苷酸或其产物,以产生多于一种细胞组分靶文库成员;以及将测序衔接子附接至多于一种条形码化受体结合试剂特异性寡核苷酸或其产物,以产生多于一种受体文库成员。方法可以包括:获得包含细胞组分靶文库成员的多于一个测序读段和受体文库成员的多于一个测序读段的测序数据。

[0089] 本文的公开内容包括试剂盒。在一些实施方案中,试剂盒包含:多于一种受体检测构建体,其中受体检测构建体包含两种或更多种受体结合试剂,其中受体结合试剂能够与受体特异性结合,并且其中受体检测构建体中的每一种包含受体结合试剂特异性寡核苷酸,所述受体结合试剂特异性寡核苷酸包含用于受体结合试剂的独特受体标识符序列。试剂盒可以包含:多于一种细胞组分结合试剂,其中多于一种细胞组分结合试剂中的每一种包含细胞组分结合试剂特异性寡核苷酸,所述细胞组分结合试剂特异性寡核苷酸包含用于细胞组分结合试剂的独特标识符序列,并且其中细胞组分结合试剂能够与细胞组分靶特异性结合,其中受体结合试剂特异性寡核苷酸包含第二通用序列,其中细胞组分结合试剂特异性寡核苷酸包含第三通用序列,其中第二通用序列和第三通用序列是不同的。在一些实施方案中,第二通用序列与第三通用序列小于约85%相同。

#### [0090] 定义

[0091] 除非另外定义,否则本文使用的技术术语和科学术语具有与本公开内容所属领域的普通技术人员通常所理解的含义。参见,例如,Singleton等人,Dictionary of Microbiology and Molecular Biology,第2版,J.Wiley&Sons (New York,NY 1994); Sambrook等人,Molecular Cloning,A Laboratory Manual,Cold Spring Harbor Press (Cold Spring Harbor,NY 1989)。为了本公开内容的目的,下文定义了以下术语。

[0092] 如本文使用的,术语“衔接子”可以意指促进关联的核酸的扩增或测序的序列。关联的核酸可以包括靶核酸。关联的核酸可以包括空间标记、靶标记、样品标记、索引标记或条形码序列(例如,分子标记)中的一种或更多种。衔接子可以是线性的。衔接子可以是预腺苷酸化的衔接子。衔接子可以是双链或单链的。一种或更多种衔接子可以位于核酸的5'端或3'端。当衔接子在5'端和3'端包含已知序列时,已知序列可以是相同或不同的序列。位于多核苷酸的5'端和/或3'端的衔接子可以能够与固定在表面上的一种或更多种寡核苷酸杂交。在一些实施方案中,衔接子可以包含通用序列。通用序列可以是两种或更多种核酸分子共有的核苷酸序列的区域。两种或更多种核酸分子也可以具有不同序列的区域。因此,例如,5'衔接子可以包含相同和/或通用核酸序列,并且3'衔接子可以包含相同和/或通用序列。可以存在于多于一个核酸分子的不同成员中的通用序列可以允许使用与通用序列互补的单种通用引物复制或扩增多于一个不同序列。类似地,可以存在于核酸分子的集合中的不同成员中的至少一种、两种(例如,一对)或更多种通用序列可以允许使用与通用序列互补的至少一种、两种(例如,一对)或更多种单一通用引物复制或扩增多于一个不同序列。因

此,通用引物包含可与此类通用序列杂交的序列。可以修饰具有靶核酸序列的分子以将通用衔接子(例如,非靶核酸序列)附接到不同靶核酸序列的一个末端或两个末端。与靶核酸衔接的一种或更多种通用引物可以提供通用引物杂交的位点。与靶核酸衔接的一种或更多种通用引物可以彼此相同或不同。

[0093] 如本文使用的,术语“关联”或“与...关联”可以意指两个或更多个物质可以被鉴定为在某个时间点共定位。关联可以意指,两个或更多个物质在或曾经在相似的容器内。关联可以是信息学关联。例如,关于两个或更多个物质的数字信息可以被存储并且可以用于确定一种或更多种物质在某个时间点共定位。关联也可以是物理关联。在一些实施方案中,两个或更多个关联的物质彼此“拴系”、“衔接”或“固定”或与共同的固体或半固体表面“拴系”、“衔接”或“固定”。关联可以指用于将标记与固体或半固体支持物(诸如珠)衔接的共价或非共价方式。关联可以是靶与标记之间的共价键。关联可以包括两个分子(诸如靶分子和标记)之间的杂交。

[0094] 如本文使用的,术语“互补”可以指两个核苷酸之间精确配对的能力。例如,如果核酸在给定位置处的核苷酸能够与另一个核酸的核苷酸形成氢键,则这两个核酸被认为在该位置处是彼此互补的。两个单链核酸分子之间的互补性可以是“部分的”,其中只有一些核苷酸结合,或者当单链分子之间存在全部互补性时它可以是完全的。如果第一核苷酸序列与第二核苷酸序列互补,则第一核苷酸序列可以被称为第二序列的“互补物”。如果第一核苷酸序列和与第二序列相反的序列(即,核苷酸顺序相反)互补,则第一核苷酸序列可以被称为第二序列的“反向互补物”。如本文使用的,“互补”序列可以指序列的“互补物”或“反向互补物”。从本公开内容理解,如果一个分子可以与另一个分子杂交,则其可以与其所杂交的分子互补或部分互补。

[0095] 如本文使用的,术语“数字计数”可以指用于估计样品中靶分子数目的方法。数字计数可以包括确定已经与样品中的靶关联的独特标记的数目的步骤。这种方法(其本质上可以是随机的)将计数分子的问题从相同分子的定位和鉴定之一转化为有关检测一组预定义标记的一系列是/否数字问题。

[0096] 如本文使用的,术语“一种标记(label)”或“多于一个标记(labels)”可以指与样品中的靶关联的核酸代码。标记可以是例如核酸标记。标记可以是完全或部分可扩增的标记。标记可以是完全或部分可测序的标记。标记可以是可鉴定为有区别的天然核酸的一部分。标记可以是已知的序列。标记可以包括核酸序列的连接处,例如天然和非天然序列的连接处。如本文使用的,术语“标记”可以与术语“索引”、“标签”或“标记-标签”互换使用。标记可以传达信息。例如,在多种实施方案中,可以使用标记来确定样品的身份、样品的来源、细胞的身份和/或靶。

[0097] 如本文使用的,术语“非耗尽性储库(non-depleting reservoir)”可以指由许多不同标记组成的条形码(例如,随机条形码)的池。非耗尽性储库可以包括大量不同的条形码,使得当非耗尽性储库与靶池关联时,每一种靶可能与独特条形码关联。每一种标记的靶分子的独特性可以通过随机选择的统计来确定,并且取决于与标记的多样性相比在集合中相同的靶分子的拷贝数。所得的标记的靶分子的集合的大小可以通过条形码化处理的随机性质来确定,并且然后对检测到的条形码的数目的分析允许计算原始集合或样品中存在的靶分子的数目。当存在的靶分子的拷贝数与独特条形码的数目的比率低时,标记的靶分子

是高度独特的(即,多于一个靶分子被给定标记标记的概率非常低)。

[0098] 如本文使用的,术语“核酸”是指多核苷酸序列或其片段。核酸可以包括核苷酸。核酸对于细胞可以是外源的或内源的。核酸可以存在于无细胞环境中。核酸可以是基因或其片段。核酸可以是DNA。核酸可以是RNA。核酸可以包括一种或更多种类似物(例如,改变的主链、糖或核酸碱基)。类似物的一些非限制性实例包括:5-溴尿嘧啶、肽核酸、非天然核酸(xeno nucleic acid)、吗啉代核酸(morpholinos)、锁核酸、二醇核酸、苏糖核酸、二脱氧核苷酸、虫草菌素、7-脱氮-GTP、荧光团(例如,罗丹明或与糖连接的荧光素)、含硫醇的核苷酸、生物素连接的核苷酸、荧光碱基类似物、CpG岛、甲基-7-鸟苷、甲基化的核苷酸、肌苷、硫代尿苷、假尿苷、二氢尿苷、癸苷(queuosine)以及怀俄苷(wyosine)。“核酸”、“多核苷酸”、“靶多核苷酸”和“靶核酸”可以互换使用。

[0099] 核酸可以包括一种或更多种修饰(例如,碱基修饰、主链修饰),以为核酸提供新的或增强的特征(例如,改进的稳定性)。核酸可以包含核酸亲和标签。核苷可以是碱基-糖组合。核苷的碱基部分可以是杂环碱基。此类杂环碱基的两个最常见的类别是嘌呤和嘧啶。核苷酸可以是还包括与核苷的糖部分共价连接的磷酸基团的核苷。对于包括呋喃戊糖的那些核苷,磷酸基团可以连接到糖的2'、3'或5'羟基部分。在形成核酸时,磷酸基团可以将相邻的核苷彼此共价连接以形成线性聚合化合物。继而,此线性聚合化合物的各端可以进一步连接而形成环状化合物;然而,线性化合物通常是合适的。此外,线性化合物可以具有内部核苷酸碱基互补性,并且因此可以按产生完全或部分双链化合物的方式折叠。在核酸中,磷酸基团通常可以称为形成核酸的核苷间主链。连键(linkage)或主链可以是3'到5'磷酸二酯连键。

[0100] 核酸可以包括修饰的主链和/或修饰的核苷间连键。修饰的主链可以包括那些在主链中保留磷原子的主链和那些在主链中没有磷原子的主链。合适的其中含磷原子的修饰的核酸主链可以包括,例如,硫代磷酸酯、手性硫代磷酸酯、二硫代磷酸酯、磷酸三酯、氨基烷基磷酸三酯、甲基和其他烷基磷酸酯诸如3'-亚烷基磷酸酯、5'-亚烷基磷酸酯、手性磷酸酯、亚磷酸酯、磷酰胺酯(phosphoramidate)(包括3'-氨基磷酰胺酯和氨基烷基磷酰胺酯、磷酸二酰胺酯(phosphorodiamidates)、硫代磷酰胺酯(thionophosphoramidates)、硫代烷基磷酸酯、硫代烷基磷酸三酯、硒代磷酸酯和硼磷酸酯,具有正常3'-5'连接、2'-5'连接的类似物以及具有反向极性的类似物(其中一个或多个核苷酸间连接是3'至3'、5'至5'或2'至2'连接)。

[0101] 核酸可以包括由短链烷基或环烷基核苷间连键,混合杂原子,和烷基或环烷基核苷间连键,或者一个或多个短链杂原子的或杂环的核苷间连键形成的多核苷酸主链。这些可以包括具有吗啉代(morpholino)连键的那些(部分由核苷的糖部分形成);硅氧烷主链;硫化物、亚砷和砷主链;甲乙酰基(formacetyl)和硫代甲乙酰基主链;亚甲基甲乙酰基和硫代甲乙酰基主链;核糖乙酰基主链;含烯烃的主链;氨基磺酸酯主链;亚甲基亚氨基和亚甲基胍基主链;磺酸酯和磺酰胺主链;酰胺主链;和具有混合的N、O、S和CH<sub>2</sub>组分部分的其他的那些。

[0102] 核酸可以包括核酸模拟物。术语“模拟物”可以意图包括其中仅呋喃糖环或呋喃糖环和核苷酸间连键二者被非呋喃糖基团替代的多核苷酸,仅呋喃糖环的替代也可以称为糖替代物(surrogate)。可以维持杂环碱基部分或修饰的杂环碱基部分,以与适当的靶核酸杂

交。一种这样的核酸可以是肽核酸(PNA)。在PNA中,多核苷酸的糖主链可以被含酰胺的主链替代,特别是被氨基乙基甘氨酸主链替代。核苷酸可以被保留,并且直接或间接与主链的酰胺部分的氮杂氮原子结合。PNA化合物中的主链可以包含两个或更多个连接的氨基乙基甘氨酸单元,这使得PNA具有含酰胺的主链。杂环碱基部分可以直接或间接与主链的酰胺部分的氮杂氮原子结合。

[0103] 核酸可以包括吗啉代主链结构。例如,核酸可以包含替代核糖环的6元吗啉代环。在这些实施方案的一些中,磷酸二酰胺酯或其他非磷酸二酯核苷间连键可以替代磷酸二酯连键。

[0104] 核酸可以包括具有附接至吗啉代环的杂环碱基的连接的吗啉代单元(例如,吗啉代核酸)。连接基团可以连接吗啉代核酸中的吗啉代单体单元。基于非离子吗啉代的寡聚化合物与细胞蛋白可以具有较少的不期望的相互作用。基于吗啉代的多核苷酸可以是核酸的非离子模拟物。吗啉代类别内的各种化合物可以使用不同的连接基团来连接。另外类别的多核苷酸模拟物可以称为环己烯基核酸(CeNA)。核酸分子中通常存在的呋喃糖环可以被环己烯基环替代。使用亚磷酰胺化学可以制备CeNA DMT保护的亚磷酰胺单体并用于寡聚化合物合成。将CeNA单体掺入核酸链可以增加DNA/RNA杂合体的稳定性。CeNA寡腺苷酸酯可以与核酸互补物形成复合物,具有与天然复合物相似稳定性。另外的修饰可以包括锁核酸(LNA),其中2'-羟基基团与糖环的4'碳原子连接,从而形成2'-C,4'-C-氧亚甲基连键,从而形成双环糖部分。连接可以是亚甲基(-CH<sub>2</sub>-),桥接2'氧原子和4'碳原子的基团,其中n是1或2。LNA和LNA类似物可以显示与互补核酸的非常高的双链体热稳定性(T<sub>m</sub>=+3°C至+10°C)、对3'-外切核酸酶降解的稳定性和良好的溶解性。

[0105] 核酸可以包括核碱基(通常简称为“碱基”)修饰或取代。如本文使用的,“未修饰的”或“天然的”核碱基可以包括嘌呤碱基(例如,腺嘌呤(A)和鸟嘌呤(G)),以及嘧啶碱基(例如,胸腺嘧啶(T)、胞嘧啶(C)和尿嘧啶(U))。经修饰的核碱基可以包括其他合成以及天然的核碱基,诸如5-甲基胞嘧啶(5-me-C)、5-羟甲基胞嘧啶、黄嘌呤、次黄嘌呤、2-氨基腺嘌呤、腺嘌呤和鸟嘌呤的6-甲基衍生物和其他烷基衍生物、腺嘌呤和鸟嘌呤的2-丙基衍生物和其他烷基衍生物,2-硫代尿嘧啶、2-硫代胸腺嘧啶和2-硫代胞嘧啶、5-卤素尿嘧啶(5-halouracil)和胞嘧啶、5-丙炔基(-C≡C-CH<sub>3</sub>)尿嘧啶和胞嘧啶以及嘧啶碱基的其他炔基衍生物,6-偶氮尿嘧啶、胞嘧啶和胸腺嘧啶、5-尿嘧啶(假尿嘧啶)、4-硫代尿嘧啶,8-卤素、8-氨基、8-硫代、8-硫代烷基、8-羟基和其他8-取代的腺嘌呤和鸟嘌呤,5-卤素特别是5-溴、5-三氟甲基和其他5-取代的尿嘧啶和胞嘧啶,7-甲基鸟嘌呤和7-甲基腺嘌呤、2-F-腺嘌呤、2-氨基腺嘌呤、8-氮杂鸟嘌呤和8-氮杂腺嘌呤、7-脱氮鸟嘌呤和7-脱氮腺嘌呤和3-脱氮鸟嘌呤和3-脱氮腺嘌呤。修饰的核碱基可以包括三环嘧啶,诸如吩噻嗪胞苷(1H-嘧啶并(5,4-b)(1,4)苯并噻嗪-2(3H)-酮)、吩噻嗪胞苷(1H-嘧啶并(5,4-b)(1,4)苯并噻嗪-2(3H)-酮),G-钳(G-clamps)诸如取代的吩噻嗪胞苷(例如,9-(2-氨基乙氧基)-H-嘧啶并(5,4-(b)(1,4)苯并噻嗪-2(3H)-酮)、吩噻嗪胞苷(1H-嘧啶并(5,4-b)(1,4)苯并噻嗪-2(3H)-酮),G-钳诸如取代的吩噻嗪胞苷(例如,9-(2-氨基乙氧基)-H-嘧啶并(5,4-(b)(1,4)苯并噻嗪-2(3H)-酮)、咪唑胞苷(2H-嘧啶并(4,5-b)吡啶-2-酮)、吡啶并吡啶胞苷(H-吡啶并(3',2':4,5)吡咯并[2,3-d]嘧啶-2-酮)。

[0106] 如本文使用的,术语“样品”可以指包含靶的组合物。用于通过所公开的方法、装置

和系统进行分析的合适样品包括细胞、组织、器官或生物体。

[0107] 如本文使用的,术语“采样装置”或“装置”可以指可以取样品的切片和/或将所述切片放置在基底上的装置。采样装置可以指例如荧光激活细胞分选(FACS)机、细胞分选机、活检针、活检装置、组织切片装置、微流体装置、刀片格栅和/或超薄切片机。

[0108] 如本文使用的,术语“固体支持物”可以指可以附接多于一个条形码(例如,随机条形码)的离散固体或半固体表面。固体支持物可以包括任何类型的实心的、多孔的或空心的球体、球、承座(bearing)、圆柱体或由塑料、陶瓷、金属或聚合材料(例如,水凝胶)构成的其他类似配置,其上可以固定核酸(例如,共价地或非共价地)。固体支持物可以包括可以是球形的(例如,微球)或具有非球形或不规则形状的离散颗粒,所述形状诸如立方体、长方形、锥形、圆柱形、圆锥形、椭圆形或圆盘形等。珠的形状可以是非球形的。以阵列间隔开的多于一个固体支持物可以不包括基底。固体支持物可以与术语“珠”互换使用。

[0109] 如本文使用的,术语“随机条形码”可以指本公开内容的包含标记的多核苷酸序列。随机条形码可以是可用于随机条形码化的多核苷酸序列。随机条形码可以用于对样品中的靶定量。随机条形码可以用于控制标记与靶关联后可能发生的错误。例如,随机条形码可用于评估扩增或测序错误。与靶关联的随机条形码可以称为随机条形码-靶或随机条形码-标签-靶。

[0110] 如本文使用的,术语“基因特异性随机条形码”可以指包含标记和基因特异性的靶结合区的多核苷酸序列。随机条形码可以是可用于随机条形码化的多核苷酸序列。随机条形码可以用于对样品中的靶定量。随机条形码可以用于控制标记与靶关联后可能发生的错误。例如,随机条形码可用于评估扩增或测序错误。与靶关联的随机条形码可以称为随机条形码-靶或随机条形码-标签-靶。

[0111] 如本文使用的,术语“随机条形码化”可以指核酸的随机标记(例如,条形码化)。随机条形码化可以利用递归泊松策略来关联并对与靶关联的标记进行定量。如本文使用的,术语“随机条形码化”可以与“随机进行标记”互换使用。

[0112] 如本文使用的,术语“靶”可以指可与条形码(例如,随机条形码)关联的组合物。用于通过所公开的方法、装置和系统进行分析的示例性合适的靶包括寡核苷酸、DNA、RNA、mRNA、微RNA、tRNA等。靶可以是单链的或双链的。在一些实施方案中,靶可以是蛋白、肽或多肽。在一些实施方案中,靶是脂质。如本文使用的,“靶”可以与“物质(species)”互换使用。

[0113] 如本文使用的,术语“逆转录酶”可以指具有逆转录酶活性(即,催化从RNA模板合成DNA)的一组酶。通常,这样的酶包括但不限于逆转录病毒逆转录酶、逆转录转座子逆转录酶、逆转录质粒(retroplasmid)逆转录酶、逆转录子逆转录酶、细菌逆转录酶、II组内含子衍生的逆转录酶,及它们的突变体、变体或衍生物。非逆转录病毒逆转录酶包括非LTR逆转录转座子逆转录酶、逆转录质粒逆转录酶、逆转录子逆转录酶和II组内含子逆转录酶。II组内含子逆转录酶的实例包括乳酸乳球菌(*Lactococcus lactis*)LI.LtrB内含子逆转录酶、细长嗜热聚球藻(*Thermosynechococcus elongatus*)TeI4c内含子逆转录酶或嗜热脂肪地芽孢杆菌(*Geobacillus stearothermophilus*)GsI-IIC内含子逆转录酶。其他类别的逆转录酶可以包括许多类型的非逆转录病毒逆转录酶(即,尤其是逆转录子、II组内含子、以及多样性产生型逆转录元件)。

[0114] 术语“通用衔接子引物”、“通用引物衔接子”或“通用衔接子序列”可互换地使用,

以指可以用于与条形码(例如,随机条形码)杂交以产生基因特异性条形码的核苷酸序列。通用衔接子序列可以例如是在本公开内容的方法中使用的遍及所有条形码通用的已知序列。例如,当使用本文公开的方法标记多于一个靶时,每一种靶特异性序列可以连接到相同的通用衔接子序列。在一些实施方案中,多于一个通用衔接子序列可以用于本文公开的方法中。例如,当使用本文公开的方法标记多于一个靶时,至少两种靶特异性序列连接到不同的通用衔接子序列。通用衔接子引物及其互补物可以被包括在两种寡核苷酸中,其中的一种寡核苷酸包含靶特异性序列且另一种寡核苷酸包含条形码。例如,通用衔接子序列可以是包含靶特异性序列的寡核苷酸的一部分以产生与靶核酸互补的核苷酸序列。包含条形码和通用衔接子序列的互补序列的第二寡核苷酸可与核苷酸序列杂交并产生靶特异性条形码(例如,靶特异性随机条形码)。在一些实施方案中,通用衔接子引物具有与本公开内容的方法中使用的通用PCR引物不同的序列。

[0115] “8mer”是由8个氨基酸组成的肽。“9mer”是由9个氨基酸组成的肽。“10mer”是由10个氨基酸组成的肽。“11mer”是由11个氨基酸组成的肽。“12mer”是由12个氨基酸组成的肽。

[0116] “氨基酸残基”可以是由肽键或不同于肽键的键连接的天然或非天然的氨基酸残基。氨基酸残基可以是D-构型或L-构型。氨基酸残基包含氨基末端部分(NH<sub>2</sub>)和羧基末端部分(COOH),由包含碳原子或碳原子链的中间部分隔开,至少其中之一包含至少一个侧链或官能团。NH<sub>2</sub>是指存在于氨基酸或肽的氨基末端的氨基基团,并且COOH是指存在于氨基酸或肽的羧基末端的羧基基团。通用术语氨基酸包括本领域技术人员已知的天然和非天然氨基酸。此外,非天然氨基酸残基包括但不限于修饰的氨基酸残基、L-氨基酸残基和立体异构体D-氨基酸残基。

[0117] 锚定氨基酸(anchor amino acid):锚定氨基酸在本文可与锚定残基互换使用,并且是具有氨基酸侧链的抗原肽的氨基酸,所述氨基酸侧链结合到排列在MHC分子的肽结合沟内的口袋中,从而将肽锚定到MHC分子上。负责肽与MHC分子主要锚定的锚定残基被称为一级锚定氨基酸。有助于抗原肽与MHC分子结合但程度低于一级锚定氨基酸的氨基酸被称为二级锚定氨基酸。

[0118] 锚定基序:结合某种MHC分子的抗原肽中锚定残基的模式。结合不同MHC分子的肽具有由肽序列中锚定残基的模式定义的不同锚定基序。

[0119] 锚定残基:锚定残基在本文可与锚定氨基酸互换使用。

[0120] 锚定位置:抗原肽序列中锚定氨基酸的位置。对于MHC II,锚定位置定义在9-mer核心基序中。

[0121] 抗原呈递细胞:如本文使用的抗原呈递细胞(APC)是在其表面展示与MHC复合的外来抗原的细胞。

[0122] 抗原肽、抗原肽P:可与P、结合肽、肽表位P或简单表位互换使用。结合或能够结合到MHC分子的结合沟中的任何肽分子。

[0123] 抗原多肽:在生物体中表达的含有一种或更多种抗原肽的多肽或蛋白。

[0124] 适配体:如本文使用的术语适配体被定义为结合特定靶分子的寡核苷酸或肽分子。适配体通常通过从大的随机序列池中选择它们来产生,但是也存在天然适配体。适配体可以分为DNA适配体、RNA适配体和肽适配体。

[0125] 抗生物素蛋白:如本文使用的抗生物素蛋白是在鸟类、爬行动物和两栖动物的蛋

清和组织中发现的糖蛋白。它包含四个相同的亚基,具有67,000-68,000道尔顿的组合质量。每个亚基由128个氨基酸组成,并结合一个生物素分子。生物活性分子:生物活性分子是自身具有生物活性/效应或当施用至生物系统时能够诱导生物活性/效应的分子。生物活性分子包括佐剂、免疫靶(例如抗原)、酶、受体活性调节剂、受体配体、免疫增强剂、药物、毒素、细胞毒性分子、共受体、一般的蛋白和肽、糖部分、脂质基团、包括siRNA的核酸、纳米颗粒、小分子。

[0126] 生物素:如本文使用的生物素也称为维生素H或维生素B<sub>7</sub>。生物素的化学式是C<sub>10</sub>H<sub>16</sub>N<sub>2</sub>O<sub>3</sub>S。

[0127] 双特异性捕获分子:对至少两个不同抗原具有结合特异性的分子。该分子也可以是三特异性或多特异性的。

[0128] 载体(carrier):如本文使用的载体可以是与MHC肽复合物直接或间接关联的任何类型的分子。在本公开内容中,载体通常指这样的官能化聚合物(例如右旋糖酐),其能够与MHC-肽复合物反应,从而将MHC-肽复合物共价附接至载体;或者能够与支架分子(例如链霉抗生物素蛋白)反应,从而将链霉抗生物素蛋白共价附接至载体,然后链霉抗生物素蛋白可以结合MHC-肽复合物。载体和支架在本文可互换使用,其中支架通常指多聚化结构域的较小分子,而载体通常指较大的分子和/或细胞样结构。

[0129] 卷曲螺旋多肽:可与卷曲螺旋肽和卷曲螺旋结构互换使用。如本文使用的术语卷曲螺旋多肽是蛋白中的结构基序,其中2-7个 $\alpha$ -螺旋像绳股一样卷曲在一起。

[0130] 右旋糖酐:如本文使用的术语右旋糖酐是一种复杂的支链多糖,由许多葡萄糖分子连接成不同长度的链组成。直链由葡萄糖分子之间的 $\alpha 1 \rightarrow 6$ 糖苷连键组成,而支链从 $\alpha 1 \rightarrow 3$ 连键开始(并且在一些情况下,也有 $\alpha 1 \rightarrow 2$ 和 $\alpha 1 \rightarrow 4$ 连键)。

[0131] 折叠:蛋白三级结构的体外或体内折叠。

[0132] 免疫监测:本文提供的免疫监测是指在感染性疾病的诊断和治疗中的免疫状态测试。它还指在疫苗接种程序之前、期间和之后的免疫状态测试。

[0133] 免疫监测过程:一系列一种或更多种免疫监测分析。

[0134] 标记(label):本文中标记可与标记分子互换使用。如本文描述的标记是可在测定中检测的可标识物质,并且可以与分子附接,产生标记的分子。然后可以研究标记的分子的行为。

[0135] 标记(labeling):本文中标记(labeling)意指将标记(label)附接至分子。

[0136] 接头分子:接头分子和接头在本文可互换使用。接头分子是共价或非共价连接两个或更多个分子,从而产生由包括接头分子的所有分子组成的更大的复合物的分子。

[0137] 免疫谱分析(profiling):如本文使用的免疫谱分析定义了对个体抗原特异性T细胞库的谱分析。

[0138] 标志物:标志物在本文可与标志物分子互换使用。标志物是与属于实体或与实体关联的分子特异性地共价或非共价关联的分子。

[0139] MHC I在本文与MHC I类可互换使用,并表示主要组织相容性复合体I类。MHC II在本文与MHC II类可互换使用,并表示主要组织相容性复合体II类。

[0140] MHC分子:如本文各处使用的MHC分子被定义为如本文定义的任何MHC I类分子或MHC II类分子,包括MHC I类分子。如本文各处使用的“MHC I类分子”可与MHC I分子互换使

用,并定义为包含1-3个亚基的分子,包括MHC I重链、与MHC IB2微球蛋白链组合的MHC I重链、通过柔性接头与MHC IB2微球蛋白链组合的MHC I重链、与抗原肽组合的MHC I重链、通过接头与抗原肽组合的MHC I重链、与抗原肽组合的MHC I重链/MHC IB2微球蛋白二聚体、以及通过与重链或 $\beta$ 2微球蛋白的柔性接头与抗原肽组合的MHC I重链/MHC IB2微球蛋白二聚体。MHC I分子链可以通过单个或一组天然氨基酸的取代,或通过插入或缺失来改变,以增强或削弱归于所述分子的功能。MHC复合物:MHC复合物在本文可与MHC-肽复合物互换使用,并且定义了与抗原肽组合的任何MHC I和/或MHC II分子,除非指明MHC复合物是空的,即不与抗原肽复合。

[0141] MHC I类样分子(包括非经典MHC I类分子)包括CD1d、HLA E、HLA G、HLA F、HLA H、MIC A、MIC B、ULBP-1、ULBP-2和ULBP-3。

[0142] “不含肽的MHC I类分子”在本文中“不含肽的MHC I分子”可互换使用,并且如本文中各处使用的意指如上定义的不含有肽的MHC I类分子。不含肽的MHC I类分子也称为“空”MHC分子。

[0143] MHC分子可适当地为脊椎动物MHC分子,诸如人类、小鼠、大鼠、猪、牛科动物或禽类MHC分子。这样的来自不同物种的MHC复合物具有不同的名称。例如在人类中,MHC复合物被称为HLA。本领域技术人员将很容易得知来自各种物种的MHC复合物的名称。

[0144] 通常,术语“MHC分子”旨在包括所有等位基因。举例来说,在人类中,例如HLA A、HLA B、HLA C、HLA D、HLA E、HLA F、HLA G、HLA H、HLA DR、HLA DQ和HLA DP等位基因是感兴趣的,应被包括;并且在小鼠系统中,H-2等位基因是感兴趣的,应被包括。同样,在大鼠系统中,RT1-等位基因;在猪系统中,SLA-等位基因;在牛科动物系统中,BoLA;在禽类系统中,例如鸡-B等位基因,是感兴趣的,应被包括。

[0145] “MHC复合物”和“MHC构建体”在本文可互换使用。

[0146] 如本文使用的术语“MHC复合物”和“MHC多聚体”意指这样的复合物及其多聚体,其能够执行归于所述复合物或多聚体的至少一种功能。这些术语包括经典和非经典MHC复合物两者。与MHC复合物相关的“经典”和“非经典”的含义是本领域技术人员熟知的。非经典MHC复合物是MHC样复合物的亚组。术语“MHC复合物”包括MHC I类分子、MHC II类分子以及MHC样分子(I类和II类两者),包括非经典MHC I类和II类分子的亚组。

[0147] MHC多聚体:本文的术语“MHC多聚体”、“MHC-多聚体”、“MHC聚体”和“MHC'聚体”可互换使用,以表示包含通过共价键或非共价键保持在一起的多于一个MHC-肽复合物的复合物。

[0148] 多聚化结构域:多聚化结构域是可以附接一个或更多个MHC或MHC-肽复合物的分子、分子复合物或固体支持物。多聚化结构域由一个或更多个载体和/或一个或更多个支架组成,并且还可以包含一个或更多个连接载体与支架、载体与载体、支架与支架的接头。多聚化结构域还可以包含一个或更多个可用于将MHC复合物和/或其他分子附接至多聚化结构域的接头。因此,多聚化结构域包含IgG、链霉抗生物素蛋白、抗生物素蛋白、streptactin、胶束、细胞、聚合物、右旋糖酐、多糖、珠和其他类型的固体支持物,以及带有反应性基团或带有能够结合MHC复合物和其他分子的化学基序的有机小分子;诸如本文在其他地方详细示出的。

[0149] 如本文各处使用的“一个或更多个”旨在包括一个和多于一个。这适用于例如MHC

肽复合物和结合实体。当多于一个MHC肽复合物附接到多聚化结构域(诸如支架或载体分子)时,MHC肽复合物的数目只需要受到多聚化结构域容量的限制。

[0150] 支架:支架通常是带有能够与MHC-肽复合物上的反应性基团反应的反应性基团的有机分子。特别地,有环状结构(例如官能化环烷烃或官能化芳环结构)的有机小分子被称为支架。支架和载体在本文可互换使用,其中支架通常指多聚化结构域的较小分子,而载体通常指较大的分子和/或细胞样结构。

[0151] 染色:通过使标记的分子与细胞表面或细胞内限定的蛋白或其他结构结合来对细胞进行特异性或非特异性标记。这些细胞是悬浮物或是组织的一部分。标记的分子可以是MHC多聚体、抗体或能够结合细胞表面特定结构的类似分子。

[0152] 链霉抗生物素蛋白:如本文使用的链霉抗生物素蛋白是从细菌亲和素链霉菌(*Streptomyces avidinii*)中纯化的四聚体蛋白。链霉抗生物素蛋白因其对生物素极强的亲和力而广泛应用于分子生物学。

[0153] 条形码

[0154] 条形码化,诸如随机条形码化,已经在例如,Fu等人,Proc Natl Acad Sci U.S.A.,2011May 31,108(22):9026-31;US2011/0160078;Fan等人,Science,2015,347(6222):1258367;US2015/0299784;和WO2015/031691中描述;这些中的每一项的内容,包括任何支持或补充信息或材料,通过引用以其整体并入本文。在一些实施方案中,本文公开的条形码可以是随机条形码,所述随机条形码是可以用于对靶进行随机标记(例如,条形码化、加标签)的多核苷酸序列。如果随机条形码的不同条形码序列的数目与待标记的任何靶的出现数目的比例可以是以下或可以是约以下:1:1、2:1、3:1、4:1、5:1、6:1、7:1、8:1、9:1、10:1、11:1、12:1、13:1、14:1、15:1、16:1、17:1、18:1、19:1、20:1、30:1、40:1、50:1、60:1、70:1、80:1、90:1、100:1或这些值中的任何两个值之间的数字或范围,则条形码可以称为随机条形码。靶可以是包括具有相同或几乎相同序列的mRNA分子的mRNA物质。如果随机条形码的不同条形码序列的数目与待标记的任何靶的出现数目的比例是至少以下或是至多以下:1:1、2:1、3:1、4:1、5:1、6:1、7:1、8:1、9:1、10:1、11:1、12:1、13:1、14:1、15:1、16:1、17:1、18:1、19:1、20:1、30:1、40:1、50:1、60:1、70:1、80:1、90:1或100:1,则条形码可以称为随机条形码。随机条形码的条形码序列可以称为分子标记。

[0155] 条形码(例如,随机条形码)可以包括一种或更多种标记。示例性标记可以包括通用标记、细胞标记、条形码序列(例如,分子标记)、样品标记、板标记、空间标记和/或前空间标记(pre-spatial label)。图1图示了具有空间标记的示例性条形码104。条形码104可以包含可以将条形码与固体支持物105连接的5' 胺。条形码可以包含通用标记、维度标记、空间标记、细胞标记和/或分子标记。条形码中不同标记(包括但不限于通用标记、维度标记、空间标记、细胞标记和分子标记)的顺序可以变化。例如,如图1中所示,通用标记可以是最5' 侧的标记(5' -most label),且分子标记可以是最3' 侧的标记(3' -most label)。空间标记、维度标记和细胞标记可以处于任何顺序。在一些实施方案中,通用标记、空间标记、维度标记、细胞标记和分子标记处于任何顺序。条形码可以包含靶结合区。靶结合区可以与样品中的靶(例如,靶核酸、RNA、mRNA、DNA)相互作用。例如,靶结合区可以包含可以与mRNA的多(A)尾相互作用的寡(dT)序列。在一些情况下,条形码的标记(例如,通用标记、维度标记、空间标记、细胞标记和条形码序列)可以由1个、2个、3个、4个、5个、6个、7个、8个、9个、10个、11

个、12个、13个、14个、15个、16个、17个、18个、19个或20个或更多个核苷酸隔开。

[0156] 标记(例如,细胞标记)可以包含一组独特的定义长度的核酸子序列,例如每一种七个核苷酸(相当于一些汉明错误校正代码中使用的比特数目),其可被设计成提供错误校正能力。可以设计包含七个核苷酸序列的错误校正子序列组,使得所述组中的序列的任何成对组合展现出定义的“遗传距离”(或错配碱基数),例如一组错误校正子序列可被设计成展现三个核苷酸的遗传距离。在这种情况下,对于标记的靶核酸分子的序列数据组中的错误校正序列的审查(在下文更详细地描述)可允许人们检测或校正扩增错误或测序错误。在一些实施方案中,用于产生错误校正代码的核酸子序列的长度可以变化,例如,它们的长度可以是以下或可以是约以下:1个、2个、3个、4个、5个、6个、7个、8个、9个、10个、15个、20个、30个、31个、40个、50个核苷酸或这些值中的任何两个值之间的数字或范围的核苷酸。在一些实施方案中,其他长度的核酸子序列可以用来产生错误校正代码。

[0157] 条形码可以包含靶结合区。靶结合区可以与样品中的靶相互作用。靶可以是以下或包括以下:核糖核酸(RNA)、信使RNA(mRNA)、微RNA、小干扰RNA(siRNA)、RNA降解产物、各自含有多(A)尾的RNA或其任何组合。在一些实施方案中,多于一个靶可以包括脱氧核糖核酸(DNA)。

[0158] 在一些实施方案中,靶结合区可以包括可以与mRNA的多(A)尾相互作用的寡(dT)序列。条形码的一种或更多种标记(例如,通用标记、维度标记、空间标记、细胞标记和条形码序列(例如,分子标记))可以通过间隔区(spacer)与条形码的另一种或两种剩余标记隔开。间隔区可以是例如1个、2个、3个、4个、5个、6个、7个、8个、9个、10个、11个、12个、13个、14个、15个、16个、17个、18个、19个或20个或更多个核苷酸。在一些实施方案中,条形码的标记中没有一个标记被间隔区隔开。

#### [0159] 通用标记

[0160] 条形码可以包含一种或更多种通用标记。在一些实施方案中,一种或更多种通用标记对于附接至给定固体支持物的条形码组中的所有条形码可以是相同的。在一些实施方案中,一种或更多种通用标记对于附接至多于一个珠的所有条形码可以是相同的。在一些实施方案中,通用标记可以包括能够与测序引物杂交的核酸序列。测序引物可以用于对包括通用标记的条形码进行测序。测序引物(例如,通用测序引物)可以包括与高通量测序平台相关的测序引物。在一些实施方案中,通用标记可以包括能够与PCR引物杂交的核酸序列。在一些实施方案中,通用标记可以包括能够与测序引物和PCR引物杂交的核酸序列。能够与测序引物或PCR引物杂交的通用标记的核酸序列可以被称为引物结合位点。通用标记可以包括可以用于引发条形码转录的序列。通用标记可以包括可以用于延伸条形码或条形码内的区域的序列。通用标记的长度可以是以下或可以是约以下:1个、2个、3个、4个、5个、10个、15个、20个、25个、30个、35个、40个、45个、50个核苷酸或这些值中的任何两个值之间的数字或范围的核苷酸。例如,通用标记可以包括至少约10个核苷酸。通用标记的长度可以是至少以下或可以是至多以下:1个、2个、3个、4个、5个、10个、15个、20个、25个、30个、35个、40个、45个、50个、100个、200个或300个核苷酸。在一些实施方案中,可裂解接头或修饰的核苷酸可以是通用标记序列的一部分,以使条形码能够从支持物上被裂解下来。

#### [0161] 维度标记

[0162] 条形码可以包含一种或更多种维度标记。在一些实施方案中,维度标记可以包括

提供关于标记(例如,随机标记)发生的维度的信息的核酸序列。例如,维度标记可以提供关于靶被条形码化的时间的信息。维度标记可以与样品中条形码化(例如,随机条形码化)的时间关联。维度标记可以在标记的时间被激活。不同的维度标记可以在不同的时间被激活。维度标记提供关于靶、靶的组和/或样品被条形码化的顺序的信息。例如,在细胞周期的G0期可以将细胞的群体条形码化。在细胞周期的G1期,可以用条形码(例如,随机条形码)对细胞再次进行脉冲处理。在细胞周期的S期,可以用条形码对细胞再次进行脉冲处理,等等。每次脉冲(例如,细胞周期的每个时期)时的条形码可以包含不同的维度标记。以这种方式,维度标记提供关于哪些靶在细胞周期的哪个时期被标记的信息。维度标记可以探询许多不同的生物学时间。示例性的生物学时间可以包括但不限于细胞周期、转录(例如,转录起始)和转录物降解。在另一种实例中,样品(例如,细胞、细胞的群体)可以在用药物和/或疗法治疗之前和/或之后标记。不同靶的拷贝数的变化可以指示样品对药物和/或疗法的响应。

[0163] 维度标记可以是可激活的。可激活的维度标记可以在特定时间点被激活。可激活的标记可以被例如组成性地激活(例如,不关闭)。可激活的维度标记可以被例如可逆地激活(例如,可激活的维度标记可以被打开和关闭)。维度标记可以被例如可逆地激活至少1次、2次、3次、4次、5次、6次、7次、8次、9次、10次或更多次。维度标记可以被可逆地激活例如至少1次、2次、3次、4次、5次、6次、7次、8次、9次、10次或更多次。在一些实施方案中,可以用荧光、光、化学事件(例如,裂解,连接另一种分子,添加修饰(例如,聚乙二醇化、类泛素化(sumoylate)、乙酰化、甲基化、去乙酰化、去甲基化)、光化学事件(例如,光罩(photocaging))以及引入非天然的核苷酸将维度标记激活。

[0164] 在一些实施方案中,维度标记对于附接至给定固体支持物(例如,珠)的所有条形码(例如,随机条形码)可以是相同的,但对于不同的固体支持物(例如,珠)是不同的。在一些实施方案中,同一固体支持物上至少60%、70%、80%、85%、90%、95%、97%、99%或100%的条形码可以包含相同的维度标记。在一些实施方案中,同一固体支持物上至少60%或至少95%的条形码可以包含相同的维度标记。

[0165] 多于一个固体支持物(例如,珠)中可以呈现多达 $10^6$ 种或更多种独特维度标记序列。维度标记的长度可以是以下或可以是约以下:1个、2个、3个、4个、5个、10个、15个、20个、25个、30个、35个、40个、45个、50个核苷酸,或这些值中的任何两个值之间的数字或范围的核苷酸。维度标记的长度可以是至少以下或可以是至多以下:1个、2个、3个、4个、5个、10个、15个、20个、25个、30个、35个、40个、45个、50个、100个、200个或300个核苷酸。维度标记可以包含长度在约5个至约200个之间的核苷酸、约10个至约150个之间的核苷酸或约20个至约125个之间的核苷酸。

#### [0166] 空间标记

[0167] 条形码可以包含一种或更多种空间标记。在一些实施方案中,空间标记可以包含提供关于与条形码关联的靶分子的空间取向的信息的核酸序列。空间标记可以与样品中的坐标关联。坐标可以是固定的坐标。例如,坐标可以相对于基底固定。空间标记可以参考二维或三维网格。坐标可以相对于界标(landmark)固定。界标可在空间中被鉴定。界标可以是可被成像的结构。界标可以是生物结构,例如解剖学界标。界标可以是细胞界标,例如细胞器。界标可以是非天然界标,诸如具有可鉴定标识(identifiable identifier)(诸如色码、条形码、磁特性(magnetic property)、荧光、放射性或独特尺寸或形状)的结构。空间标记

可以与物理分区(例如,孔、容器或液滴)关联。在一些实施方案中,将多于一个空间标记一起用于编码空间中的一个或更多个位置。

[0168] 空间标记对于附接至给定固体支持物(例如,珠)的所有条形码可以是相同的,但对于不同的固体支持物(例如,珠)是不同的。在一些实施方案中,同一固体支持物上包含相同空间标记的条形码的百分比可以是以下或可以是约以下:60%、70%、80%、85%、90%、95%、97%、99%、100%或这些值中的任何两个值之间的数字或范围。在一些实施方案中,同一固体支持物上包含相同空间标记的条形码的百分比可以是至少或至多60%、70%、80%、85%、90%、95%、97%、99%或100%。在一些实施方案中,同一固体支持物上至少60%或至少95%的条形码可以包含相同的空间标记。

[0169] 多于一个固体支持物(例如,珠)中可以呈现多达 $10^6$ 种或更多种独特空间标记序列。空间标记的长度可以是以下或可以是约以下:1个、2个、3个、4个、5个、10个、15个、20个、25个、30个、35个、40个、45个、50个核苷酸或这些值中的任何两个值之间的数字或范围的核苷酸。空间标记的长度可以是至少以下或至多以下:1个、2个、3个、4个、5个、10个、15个、20个、25个、30个、35个、40个、45个、50个、100个、200个或300个核苷酸。空间标记可以包含长度在约5个至约200个之间的核苷酸、约10个至约150个的核苷酸或约20个至约125个之间的核苷酸。

#### [0170] 细胞标记

[0171] 条形码(例如,随机条形码)可以包含一种或更多种细胞标记。在一些实施方案中,细胞标记可以包含提供用于确定哪种靶核酸来源于哪种细胞的信息的核酸序列。在一些实施方案中,细胞标记对于附接至给定固体支持物(例如,珠)的所有条形码是相同的,但对于不同的固体支持物(例如,珠)是不同的。在一些实施方案中,同一固体支持物上包含相同细胞标记的条形码的百分比可以是以下或可以是约以下:60%、70%、80%、85%、90%、95%、97%、99%、100%或这些值中的任何两个值之间的数字或范围。在一些实施方案中,同一固体支持物上包含相同细胞标记的条形码的百分比可以是以下或可以是约以下:60%、70%、80%、85%、90%、95%、97%、99%或100%。例如,同一固体支持物上至少60%或至少95%的条形码可以包含相同的细胞标记。

[0172] 多于一个固体支持物(例如,珠)中可以呈现多达 $10^6$ 种或更多种独特细胞标记序列。细胞标记的长度可以是以下或可以是约以下:1个、2个、3个、4个、5个、10个、15个、20个、25个、30个、35个、40个、45个、50个核苷酸或这些值中的任何两个值之间的数字或范围的核苷酸。细胞标记的长度可以是至少以下或可以是至多以下:1个、2个、3个、4个、5个、10个、15个、20个、25个、30个、35个、40个、45个、50个、100个、200个或300个核苷酸。例如,细胞标记可以包含长度在约5个至约200个之间的核苷酸、约10个至约150个之间的核苷酸或约20个至约125个之间的核苷酸。

#### [0173] 条形码序列

[0174] 条形码可以包含一种或更多种条形码序列。在一些实施方案中,条形码序列可以包含为与条形码杂交的特定类型的靶核酸物质提供鉴定信息的核酸序列。条形码序列可以包含为与条形码(例如,靶结合区)杂交的靶核酸物质的特定出现提供计数器(例如,提供粗略估计)的核酸序列。

[0175] 在一些实施方案中,将一组相异的(diverse)条形码序列附接至给定固体支持物

(例如,珠)。在一些实施方案中,可以有以下或可以有约以下的独特分子标记序列: $10^2$ 种、 $10^3$ 种、 $10^4$ 种、 $10^5$ 种、 $10^6$ 种、 $10^7$ 种、 $10^8$ 种、 $10^9$ 种或这些值中的任何两个值之间的数字或范围。例如,多于一种条形码可以包括约6561种具有不同序列的条形码序列。作为另一种实例,多于一种条形码可以包括约65536种具有不同序列的条形码序列。在一些实施方案中,可以有至少以下或可以有至多以下的独特条形码序列: $10^2$ 种、 $10^3$ 种、 $10^4$ 种、 $10^5$ 种、 $10^6$ 种、 $10^7$ 种、 $10^8$ 种或 $10^9$ 种。独特分子标记序列可以附接至给定固体支持物(例如,珠)。在一些实施方案中,独特分子标记序列被颗粒(例如水凝胶珠)部分或全部包含。

[0176] 在不同实施方式中,条形码的长度可以是不同的。例如,条形码的长度可以是以下或可以是约以下:1个、2个、3个、4个、5个、10个、15个、20个、25个、30个、35个、40个、45个、50个核苷酸或这些值中的任何两个值之间的数字或范围的核苷酸。作为另一种实例,条形码的长度可以是至少以下或可以是至多以下:1个、2个、3个、4个、5个、10个、15个、20个、25个、30个、35个、40个、45个、50个、100个、200个或300个核苷酸。

#### [0177] 分子标记

[0178] 条形码(例如,随机条形码)可以包含一种或更多种分子标记。分子标记可以包含条形码序列。在一些实施方案中,分子标记可以包含为与条形码杂交的特定类型的靶核酸物质提供鉴定信息的核酸序列。分子标记可以包含为与条形码(例如,靶结合区)杂交的靶核酸物质的特定出现提供计数器的核酸序列。

[0179] 在一些实施方案中,将一组相异的分子标记附接至给定固体支持物(例如,珠)。在一些实施方案中,可以有以下或可以有约以下的独特分子标记序列: $10^2$ 种、 $10^3$ 种、 $10^4$ 种、 $10^5$ 种、 $10^6$ 种、 $10^7$ 种、 $10^8$ 种、 $10^9$ 种或这些值中的任何两个值之间的数字或范围。例如,多于一种条形码可以包括约6561种具有不同序列的分子标记。作为另一种实例,多于一种条形码可以包括约65536种具有不同序列的分子标记。在一些实施方案中,可以有至少以下或可以有至多以下的独特分子标记序列: $10^2$ 种、 $10^3$ 种、 $10^4$ 种、 $10^5$ 种、 $10^6$ 种、 $10^7$ 种、 $10^8$ 种或 $10^9$ 种。具有独特分子标记序列的条形码可以附接至给定固体支持物(例如,珠)。

[0180] 对于使用多于一个随机条形码进行的条形码化(例如随机条形码化),不同分子标记序列的数量与任何靶的出现次数的比例可以是以下或可以是约以下:1:1、2:1、3:1、4:1、5:1、6:1、7:1、8:1、9:1、10:1、11:1、12:1、13:1、14:1、15:1、16:1、17:1、18:1、19:1、20:1、30:1、40:1、50:1、60:1、70:1、80:1、90:1、100:1,或这些值中的任何两个值之间的数字或范围。靶可以是包括具有相同或几乎相同序列的mRNA分子的mRNA物质。在一些实施方案中,不同分子标记序列的数量与任何靶的出现次数的比例是至少以下或是至多以下:1:1、2:1、3:1、4:1、5:1、6:1、7:1、8:1、9:1、10:1、11:1、12:1、13:1、14:1、15:1、16:1、17:1、18:1、19:1、20:1、30:1、40:1、50:1、60:1、70:1、80:1、90:1或100:1。

[0181] 分子标记的长度可以是以下或可以是约以下:1个、2个、3个、4个、5个、10个、15个、20个、25个、30个、35个、40个、45个、50个或这些值中的任何两个值之间的数字或范围的核苷酸。分子标记的长度可以是至少以下或可以是至多以下:1个、2个、3个、4个、5个、10个、15个、20个、25个、30个、35个、40个、45个、50个、100个、200个或300个核苷酸。

#### [0182] 靶结合区

[0183] 条形码可以包含一个或更多个靶结合区,诸如捕获探针。在一些实施方案中,靶结合区可以与感兴趣的靶杂交。在一些实施方案中,靶结合区可以包含与靶(例如,靶核酸、靶

分子,例如待分析的细胞核酸)特异性杂交(例如,与特定基因序列特异性杂交)的核酸序列。在一些实施方案中,靶结合区可以包含可以附接(例如,杂交)至特定靶核酸的特定位置的核酸序列。在一些实施方案中,靶结合区可以包含能够与限制性酶位点突出端(例如,EcoRI粘性末端突出端)特异性杂交的核酸序列。然后条形码可以连接至包含与限制性位点突出端互补的序列的任何核酸分子。

[0184] 在一些实施方案中,靶结合区可以包含非特异性靶核酸序列。非特异性靶核酸序列可以指可以独立于靶核酸的特定序列结合多于一个靶核酸的序列。例如,靶结合区可以包含随机多聚体序列、多(dA)序列、多(dT)序列、多(dG)序列、多(dC)序列或其组合。例如,靶结合区可以是与mRNA分子上的多(A)尾杂交的寡(dT)序列。随机多聚体序列可以是,例如,随机二聚体、三聚体、四聚体、五聚体、六聚体、七聚体、八聚体、九聚体、十聚体或任何长度的更高多聚体序列。在一些实施方案中,对于附接至给定珠的所有条形码,靶结合区是相同的。在一些实施方案中,对于附接至给定珠的多于一种条形码,靶结合区可以包括两种或更多种不同的靶结合序列。靶结合区的长度可以是以下或可以是约以下:5个、10个、15个、20个、25个、30个、35个、40个、45个、50个核苷酸或这些值中的任何两个值之间的数字或范围的核苷酸。靶结合区的长度可以是至多约5个、10个、15个、20个、25个、30个、35个、40个、45个、50个或更多个核苷酸。例如,可以使用逆转录酶诸如Moloney鼠白血病病毒(MMLV)逆转录酶逆转录mRNA分子,以产生具有多(dC)尾的cDNA分子。条形码可以包括具有多(dG)尾的靶结合区。在条形码的多(dG)尾和cDNA分子的多(dC)尾之间碱基配对后,逆转录酶将模板链从细胞RNA分子转换到条形码,并继续向条形码的5'末端复制。通过这样做,得到的cDNA分子在该cDNA分子3'末端上含有条形码序列(诸如分子标记)。

[0185] 在一些实施方案中,靶结合区可以包含寡(dT),所述寡(dT)可以与包含聚腺苷酸化末端的mRNA杂交。靶结合区可以是基因特异性的。例如,可以将靶结合区配置为与靶的特定区域杂交。靶结合区的长度可以是以下或可以是约以下:1个、2个、3个、4个、5个、6个、7个、8个、9个、10个、11个、12个、13个、14个、15个、16个、17个、18个、19个、20个、21个、22个、23个、24个、25个、26个、27个、28个、29个、30个核苷酸或这些值中的任何两个值之间的数字或范围的核苷酸。靶结合区的长度可以是至少以下或可以是至多以下:1个、2个、3个、4个、5个、6个、7个、8个、9个、10个、11个、12个、13个、14个、15个、16个、17个、18个、19个、20个、21个、22个、23个、24个、25个、26个、27个、28个、29个或30个核苷酸。靶结合区的长度可以是约5-30个核苷酸。当条形码包含基因特异性靶结合区时,条形码在本文中可称为基因特异性条形码。

[0186] 定向特性(Orientation Property)

[0187] 随机条形码(例如,随机条形码)可以包含一种或更多种可以用于定向(例如,比对)条形码的定向特性。条形码可以包含用于等电聚焦的部分。不同的条形码可以包含不同的等电聚焦点。当将这些条形码引入样品中时,样品可以经历等电聚焦,以便于将条形码定向成已知的方式。以这种方式,定向特性可以用于开发样品中条形码的已知的映射。示例性定向特性可以包括电泳迁移率(例如,基于条形码的尺寸)、等电点、自旋、电导率和/或自组装。例如,具有自组装的定向特性的条形码激活时可以自组装成特定的定向(例如,核酸纳米结构)。

[0188] 亲和特性(Affinity Property)

[0189] 条形码(例如,随机条形码)可以包含一种或更多种亲和特性。例如,空间标记可以包含亲和特性。亲和特性可以包括可以促进条形码与另一种实体(例如,细胞受体)结合的化学部分和/或生物部分。例如,亲和特性可以包括抗体,例如,对样品上的特定部分(例如,受体)特异性的抗体。在一些实施方案中,抗体可以将条形码引导至特定细胞类型或分子。在特定细胞类型或分子处和/或特定细胞类型或分子附近的靶可以被标记(例如,被随机标记)。在一些实施方案中,亲和特性可以提供空间标记的核苷酸序列之外的空间信息,因为抗体可以将条形码引导至特定位置。抗体可以是治疗性抗体,例如单克隆抗体或多克隆抗体。抗体可以是人源化的或嵌合的。抗体可以是裸抗体或融合抗体。

[0190] 抗体可以是全长(即,天然存在的或通过正常免疫球蛋白基因片段重组过程形成的)免疫球蛋白分子(例如,IgG抗体)或免疫球蛋白分子的免疫活性(即,特异性结合性)部分(如抗体片段)。

[0191] 抗体片段可以是例如抗体的一部分,诸如F(ab')<sub>2</sub>、Fab'、Fab、Fv、sFv等。在一些实施方案中,抗体片段可以与由全长抗体识别的相同抗原结合。抗体片段可以包括由抗体的可变区组成的分离的片段,诸如由重链和轻链的可变区组成的“Fv”片段和其中轻链和重链可变区通过肽接头连接的重组单链多肽分子(“scFv蛋白”)。示例性抗体可以包括但不限于癌细胞抗体、病毒抗体、与细胞表面受体(CD8、CD34、CD45)结合的抗体和治疗性抗体。

[0192] 通用衔接子引物

[0193] 条形码可以包含一种或更多种通用衔接子引物。例如,基因特异性条形码(诸如基因特异性随机条形码)可以包含通用衔接子引物。通用衔接子引物可以指遍及所有条形码的通用的核苷酸序列。通用衔接子引物可以用于构建基因特异性条形码。通用衔接子引物的长度可以是以下或可以是约以下:1个、2个、3个、4个、5个、6个、7个、8个、9个、10个、11个、12个、13个、14个、15个、16个、17个、18个、19个、20个、21个、22个、23个、24个、25个、26个、27个、28个、29个、30个或这些值中的任何两个值之间的数字或范围的核苷酸。通用衔接子引物的长度可以是至少以下或可以是至多以下:1个、2个、3个、4个、5个、6个、7个、8个、9个、10个、11个、12个、13个、14个、15个、16个、17个、18个、19个、20个、21个、22个、23个、24个、25个、26个、27个、28个、29个或30个核苷酸。通用衔接子引物的长度可以是5-30个核苷酸。

[0194] 接头

[0195] 当条形码包含多于一个类型的标记(例如,多于一种细胞标记或多于一种条形码序列,诸如一种分子标记)时,标记之间可以散布有接头标记序列。接头标记序列的长度可以是至少以下、至少约以下、至多以下或至多约以下:5个、10个、15个、20个、25个、30个、35个、40个、45个、50个或更多个核苷酸。在一些实施方案中,接头标记序列的长度是12个核苷酸。接头标记序列可以用于促进条形码的合成。接头标记可以包括错误校正(例如,汉明)码。

[0196] 固体支持物

[0197] 在一些实施方案中,本文公开的条形码(诸如随机条形码)可以与固体支持物关联。固体支持物可以是例如合成颗粒。在一些实施方案中,固体支持物上的多于一种条形码(例如,第一多于一种条形码)的一些或所有条形码序列(诸如,随机条形码(例如,第一条条形码序列)的分子标记)相差至少一个核苷酸。同一固体支持物上的条形码的细胞标记可以是相同的。不同的固体支持物上的条形码的细胞标记可以相差至少一个核苷酸。例如,第一固

体支持物上的第一多于一种条形码的第一细胞标记可以具有相同的序列,且第二固体支持物上的第二多于一种条形码的第二细胞标记可以具有相同的序列。第一固体支持物上的第一多于一种条形码的第一细胞标记和第二固体支持物上的第二多于一种条形码的第二细胞标记可以相差至少一个核苷酸。细胞标记可以是例如约5-20个核苷酸长。条形码序列可以是例如约5-20个核苷酸长。合成颗粒可以是例如珠。

[0198] 珠可以是例如硅胶珠、可控孔径玻璃珠、磁珠、Dynabead、Sephadex/琼脂糖凝胶珠、纤维素珠、聚苯乙烯珠或其任何组合。珠可以包括材料诸如聚二甲基硅氧烷(PDMS)、聚苯乙烯、玻璃、聚丙烯、琼脂糖、明胶、水凝胶、顺磁物质、陶瓷、塑料、玻璃、甲基苯乙烯、丙烯酸聚合物、钛、胶乳、琼脂糖凝胶、纤维素、尼龙、硅酮或其任何组合。

[0199] 在一些实施方案中,珠可以用条形码或随机条形码官能化的聚合物珠(例如可变形珠或凝胶珠)(诸如来自10X Genomics(San Francisco,CA)的凝胶珠)。在一些实施方式中,凝胶珠可以包括基于聚合物的凝胶。凝胶珠可以例如通过将一种或更多种聚合物前体包封到液滴中来产生。在将聚合物前体暴露于促进剂(例如,四甲基乙二胺(TEMED))后,可以产生凝胶珠。

[0200] 在一些实施方案中,颗粒可以是可破坏的(例如,可溶解的、可降解的)。例如,聚合物珠可以例如在期望的条件下溶解、熔化或降解。所期望的条件可以包括环境条件。所期望的条件可以导致聚合物珠以受控方式溶解、熔化或降解。凝胶珠可以由于化学刺激、物理刺激、生物刺激、热刺激、磁刺激、电刺激、光刺激或其任何组合而溶解、熔化或降解。

[0201] 例如,分析物和/或试剂(诸如寡核苷酸条形码)可以偶联/固定至凝胶珠的内表面(例如,经由寡核苷酸条形码和/或用于产生寡核苷酸条形码的材料的扩散而可及的内部)和/或凝胶珠的外表面或本文描述的任何其他微胶囊。偶联/固定可以经由任何形式的化学键合(例如,共价键、离子键)或物理现象(例如,范德华力、偶极-偶极相互作用等)。在一些实施方案中,本文描述的试剂与凝胶珠或任何其他微胶囊的偶联/固定可以是可逆的,诸如,例如经由不稳定型部分(例如,经由化学交联物,包括本文描述的化学交联物)。在施加刺激后,不稳定型部分可以被裂解并释放所固定的试剂。在一些实施方案中,不稳定型部分是二硫键。例如,在经由二硫键将寡核苷酸条形码固定至凝胶珠的情况下,使二硫键暴露于还原剂可以裂解二硫键并从珠释放寡核苷酸条形码。不稳定型部分可以作为凝胶珠或微胶囊的一部分、作为将试剂或分析物与凝胶珠或微胶囊连接的化学接头的一部分和/或作为试剂或分析物的一部分被包括。在一些实施方案中,多于一种条形码的至少一种条形码可以被固定在颗粒上、被部分固定在颗粒上、被包封在颗粒中、被部分包封在颗粒中或其任何组合。

[0202] 在一些实施方案中,凝胶珠可以包括宽范围的不同的聚合物,包括但不限于:聚合物、热敏聚合物、光敏聚合物、磁性聚合物、pH敏感聚合物、盐敏感聚合物、化学敏感聚合物、聚电解质、多糖、肽、蛋白和/或塑料。聚合物可以包括但不限于以下材料:诸如聚(N-异丙基丙烯酰胺)(PNIPAAm)、聚(苯乙烯磺酸酯)(PSS)、聚(烯丙基胺)(PAAm)、聚(丙烯酸)(PAA)、聚(乙烯亚胺)(PEI)、聚(双烯丙基二甲基-氯化铵)(PDADMAC)、聚(吡咯)(poly(pyrolle), PPy)、聚(乙基吡咯烷酮)(PVPON)、聚(乙基吡啶)(PVP)、聚(甲基丙烯酸)(PMAA)、聚(甲基丙烯酸甲酯)(PMMA)、聚苯乙烯(PS)、聚(四氢呋喃)(PTHF)、聚(邻苯二甲醛)(PPA)、聚(己基紫精)(PHV)、聚(L-赖氨酸)(PLL)、聚(L-精氨酸)(PARG)、聚(乳酸-共-羟基乙酸)(PLGA)。

[0203] 许多化学刺激可以用于触发珠的破坏、溶解或降解。这些化学改变的实例可以包括但不限于pH介导的珠壁改变、经由交联键的化学裂解使珠壁崩解、珠壁的触发解聚和珠壁转换反应。批量(bulk)改变也可以用于触发珠的破坏。

[0204] 通过各种刺激对微胶囊的批量或物理改变在设计胶囊以释放试剂方面也提供了许多优点。在宏观尺度上发生批量或物理改变,其中珠破裂是由刺激引起的机械-物理力的结果。这些过程可以包括但不限于压力引起的破裂、珠壁熔化或珠壁的孔隙率的改变。

[0205] 生物刺激也可以用于触发珠的破坏、溶解或降解。通常,生物触发物类似于化学触发物,但是许多实例使用生物分子或生命系统中常见的分子,诸如酶、肽、糖、脂肪酸、核酸等。例如,珠可以包含具有对特定蛋白酶的裂解敏感的肽交联的聚合物。更特别地,一种实例可以包括包含GFLGK肽交联的微胶囊。在添加生物触发物(诸如蛋白酶组织蛋白酶B)后,壳壁的肽交联被裂解并且珠的内容物被释放。在其他情况下,蛋白酶可以是热激活的。在另一种实例中,珠包括包含纤维素的壳壁。壳聚糖水解酶的添加用作纤维素键裂解、壳壁解聚及其内部内容物释放的生物触发物。

[0206] 还可以在施加热刺激后诱导珠释放其内容物。温度的改变可以引起珠的各种改变。热量的变化可以引起珠熔化,使得珠壁崩解。在其他情况下,热量可以增加珠的内部组分的内部压力,使得珠破裂或爆炸。在又其他的情况下,热量可以使珠转化成收缩的脱水状态。热量还可以作用于珠壁内的热敏聚合物,从而引起珠的破坏。

[0207] 将磁性纳米颗粒包括在微胶囊的珠壁中可以允许珠的触发破裂以及将珠引导成阵列。本公开内容的装置可以包括用于任一目的的磁珠。在一个实例中,将 $\text{Fe}_3\text{O}_4$ 纳米颗粒掺入含聚电解质的珠中在存在振荡磁场刺激的情况下触发破裂。

[0208] 珠可以由于电刺激的结果被破坏、溶解或降解。与先前部分中描述的磁性颗粒类似,电敏珠可以允许珠的触发破裂以及其他功能,诸如电场中的对齐、电导率或氧化还原反应。在一种实例中,含电敏材料的珠在电场中对齐,从而可以控制内部试剂的释放。在其他实例中,电场可以在珠壁本身内引起氧化还原反应,这可以增加孔隙率。

[0209] 也可以使用光刺激来破坏珠。许多光触发物是可能的,并且可以包括使用各种分子(诸如能够吸收特定波长范围的光子的纳米颗粒和发色团)的系统。例如,金属氧化物涂层可以用作胶囊触发物。涂覆有 $\text{SiO}_2$ 的聚电解质胶囊的UV照射可以导致珠壁的崩解。在又另一种实例中,可以将可光切换材料(诸如偶氮苯基团)掺入珠壁中。在施加UV或可见光后,诸如这些的化学物质在吸收光子后经历可逆的顺式至反式异构化。在此方面,掺入光子开关(photon switch)产生在施加光触发物后可以崩解或变得更多孔的珠壁。

[0210] 例如,在图2中图示的条形码化(例如,随机条形码化)的非限制性实例中,在框208处将细胞(诸如单细胞)引入微孔阵列的多于一个微孔上之后,在框212处可以将珠引入微孔阵列的多于一个微孔上。每个微孔可以包含一个珠。珠可以包含多于一个条形码。条形码可以包含衔接至珠的5' 胺区域。条形码可以包含通用标记、条形码序列(例如,分子标记)、靶结合区或其任何组合。

[0211] 本文公开的条形码可以与固体支持物(例如,珠)关联(例如,衔接)。与固体支持物关联的条形码可各自包含选自以下组的条形码序列,该组包括至少100种或1000种具有独特序列的条形码序列。在一些实施方案中,与固体支持物关联的不同条形码可以包含具有不同序列的条形码。在一些实施方案中,与固体支持物关联的条形码的一定百分比包含相

同的细胞标记。例如,所述百分比可以是以下或可以是约以下:60%、70%、80%、85%、90%、95%、97%、99%、100%,或这些值中的任何两个值之间的数字或范围。作为另一个实例,所述百分比可以是至少以下或可以是至多以下:60%、70%、80%、85%、90%、95%、97%、99%或100%。在一些实施方案中,与固体支持物关联的条形码可以具有相同的细胞标记。与不同固体支持物关联的条形码可以具有选自以下组的不同的细胞标记,该组包括至少100种或1000种具有独特序列的细胞标记。

[0212] 本文公开的条形码可以与固体支持物(例如,珠)关联(例如,附接)。在一些实施方案中,可以用包括与多于一种条形码关联的多于一个合成颗粒的固体支持物将样品中的多于一个靶条形码化。在一些实施方案中,固体支持物可以包括与多于一种条形码关联的多于一个合成颗粒。不同固体支持物上的多于一种条形码的空间标记可以相差至少一个核苷酸。固体支持物可以例如在二维或三维包括多于一种条形码。合成颗粒可以是珠。珠可以是硅胶珠、可控孔径玻璃珠、磁珠、Dynabead、Sephadex/琼脂糖凝胶珠、纤维素珠、聚苯乙烯珠或其任何组合。固体支持物可以包括聚合物、基质、水凝胶、针阵列装置、抗体或其任何组合。在一些实施方案中,固体支持物可以自由浮动。在一些实施方案中,固体支持物可以包埋到半固体或固体阵列中。条形码可以不与固体支持物关联。条形码可以是单独的核苷酸。条形码可以与基底关联。

[0213] 如本文使用的,术语“拴系的”、“附接的”和“固定的”可以互换使用,并且可以指用于将条形码附接至固体支持物的共价或非共价方式。可以将各种不同的固体支持物中的任何一种用作固体支持物,以用于附接预先合成的条形码或用于条形码的原位固相合成。

[0214] 在一些实施方案中,固体支持物是珠。珠可以包括一种或更多种类型的实心的、多孔的或空心的球体、球、承座、圆柱体或可以固定核酸(例如,共价地或非共价地)的其他类似配置。珠可以由例如塑料、陶瓷、金属、聚合物材料或其任何组合构成。珠可以是或包括球形的(例如,微球)或具有非球形或不规则形状的离散颗粒,所述形状诸如立方体、长方形、锥形、圆柱形、圆锥形、椭圆形或圆盘形等。在一些实施方案中,珠的形状可以是非球形的。

[0215] 珠可以包括各种材料,包括但不限于顺磁性材料(例如,镁、钼、锂和钽)、超顺磁性材料(例如,铁氧体( $\text{Fe}_3\text{O}_4$ ;磁铁矿)纳米颗粒)、铁磁材料(例如,铁、镍、钴,它们的一些合金,以及一些稀土金属化合物)、陶瓷、塑料、玻璃、聚苯乙烯、二氧化硅、甲基苯乙烯、丙烯酸聚合物、钛、胶乳、琼脂糖凝胶、琼脂糖、水凝胶、聚合物、纤维素、尼龙,或其任何组合。

[0216] 在一些实施方案中,珠(例如,标记所附接的珠)是水凝胶珠。在一些实施方案中,珠包括水凝胶。

[0217] 本文公开的一些实施方案包括一个或更多个颗粒(例如,珠)。每个颗粒可以包含多于一个寡核苷酸(例如,条形码)。多于一个寡核苷酸中的每一个可以包含条形码序列(例如,分子标记序列)、细胞标记和靶结合区(例如,寡(dT)序列、基因特异性序列、随机多聚体或其组合)。多于一个寡核苷酸的每一个的细胞标记序列可以是相同的。不同颗粒上的寡核苷酸的细胞标记序列可以是不同的,使得可以鉴定不同颗粒上的寡核苷酸。在不同实施方式中,不同细胞标记序列的数目可以是不同的。在一些实施方案中,细胞标记序列的数目可以是以下或可以是约以下:10、100、200、300、400、500、600、700、800、900、1000、2000、3000、4000、5000、6000、7000、8000、9000、10000、20000、30000、40000、50000、60000、70000、80000、90000、100000、 $10^6$ 、 $10^7$ 、 $10^8$ 、 $10^9$ 、这些值中的任何两个值之间的数字或范围或更多。

在一些实施方案中,细胞标记序列的数目可以是至少以下或可以是至多以下:10、100、200、300、400、500、600、700、800、900、1000、2000、3000、4000、5000、6000、7000、8000、9000、10000、20000、30000、40000、50000、60000、70000、80000、90000、100000、 $10^6$ 、 $10^7$ 、 $10^8$ 或 $10^9$ 。在一些实施方案中,多于一个颗粒中的不超过1、2、3、4、5、6、7、8、9、10、20、30、40、50、60、70、80、90、100、200、300、400、500、600、700、800、900、1000或更多个颗粒包括具有相同细胞序列的寡核苷酸。在一些实施方案中,包括具有相同细胞序列的寡核苷酸的多于一个颗粒可以是至多0.1%、0.2%、0.3%、0.4%、0.5%、0.6%、0.7%、0.8%、0.9%、1%、2%、3%、4%、5%、6%、7%、8%、9%、10%或更多。在一些实施方案中,多于一个颗粒全都不具有相同的细胞标记序列。

[0218] 在每个颗粒上的多于一个寡核苷酸可以包含不同的条形码序列(例如,分子标记)。在一些实施方案中,条形码序列的数目可以是以下或可以是约以下:10、100、200、300、400、500、600、700、800、900、1000、2000、3000、4000、5000、6000、7000、8000、9000、10000、20000、30000、40000、50000、60000、70000、80000、90000、100000、 $10^6$ 、 $10^7$ 、 $10^8$ 、 $10^9$ ,或这些值中的任何两个值之间的数字或范围。在一些实施方案中,条形码序列的数目可以是至少以下或可以是至多以下:10、100、200、300、400、500、600、700、800、900、1000、2000、3000、4000、5000、6000、7000、8000、9000、10000、20000、30000、40000、50000、60000、70000、80000、90000、100000、 $10^6$ 、 $10^7$ 、 $10^8$ 或 $10^9$ 。例如,多于一个寡核苷酸中的至少100种包含不同的条形码序列。作为另一种实例,在单个颗粒中,多于一个寡核苷酸中的至少100种、500种、1000种、5000种、10000种、15000种、20000种、50000种、这些值中的任何两个值之间的数字或范围或更多种包含不同的条形码序列。一些实施方案提供了多于一个包含条形码的颗粒。在一些实施方案中,待标记的靶和不同条形码序列的出现(或拷贝或数目)的比例可以是至少1:1、1:2、1:3、1:4、1:5、1:6、1:7、1:8、1:9、1:10、1:11、1:12、1:13、1:14、1:15、1:16、1:17、1:18、1:19、1:20、1:30、1:40、1:50、1:60、1:70、1:80、1:90或更高。在一些实施方案中,多于一个寡核苷酸的每一个还包含样品标记、通用标记或二者。颗粒可以是例如纳米颗粒或微米颗粒。

[0219] 珠的尺寸可以不同。例如,珠的直径范围可以为从0.1微米至50微米。在一些实施方案中,珠的直径可以是以下或可以是约以下:0.1微米、0.5微米、1微米、2微米、3微米、4微米、5微米、6微米、7微米、8微米、9微米、10微米、20微米、30微米、40微米、50微米或这些值中的任何两个值之间的数字或范围。

[0220] 珠的直径可以与基底的孔的直径相关。在一些实施方案中,珠的直径可以比孔的直径长或短以下或者约以下:10%、20%、30%、40%、50%、60%、70%、80%、90%、100%或这些值中的任何两个值之间的数字或范围。珠的直径可以与细胞(例如,被基底的孔捕获的单细胞)的直径相关。在一些实施方案中,珠的直径可以比孔的直径长或短至少或至多10%、20%、30%、40%、50%、60%、70%、80%、90%、100%、150%、200%、250%或300%。珠的直径可以与细胞(例如,被基底的孔捕获的单细胞)的直径相关。在一些实施方案中,珠的直径可以比细胞的直径长或短以下或约以下:10%、20%、30%、40%、50%、60%、70%、80%、90%、100%、150%、200%、250%、300%或这些值中的任何两个值之间的数字或范围。

[0221] 珠可以附接至基底和/或包埋到基底中。珠可以附接至凝胶、水凝胶、聚合物和/或

基质和/或包埋到凝胶、水凝胶、聚合物和/或基质中。珠在基底(例如凝胶、基质、支架或聚合物)中的空间位置可以使用珠上的条形码上存在的空间标记来鉴定,该空间标记可以用作位置地址。

[0222] 珠的实例可以包括但不限于链霉抗生物素蛋白珠、琼脂糖珠、磁珠、Dynabeads®、MACS®微珠、抗体缀合的珠(例如,抗免疫球蛋白微珠)、蛋白A缀合的珠、蛋白G缀合的珠、蛋白A/G缀合的珠、蛋白L缀合的珠、寡(dT)缀合的珠、二氧化硅珠、二氧化硅样珠、抗生物素微珠、抗荧光染料微珠和BcMag™羧基封端磁珠。

[0223] 珠可以与量子点或荧光染料关联(例如,用量子点或荧光染料浸渍),以使其在一个荧光光学通道或多于一个光学通道中发荧光。珠可以与氧化铁或氧化铬关联,使其成为顺磁性或铁磁性。珠可以是可鉴定的。例如,可以使用照相机对珠成像。珠可以具有与珠关联的可检测代码。例如,珠可以包含条形码。珠可以改变尺寸,例如由于在有机溶液或无机溶液中溶胀。珠可以是疏水性或亲水性的。珠可以是生物相容的。

[0224] 固体支持物(例如,珠)可以被可视化。固体支持物可以包含可视化标签(例如,荧光染料)。固体支持物(例如,珠)可以蚀刻有标识符(例如,数字)。标识符可以通过对珠成像来可视化。

[0225] 固体支持物可以包括可溶性、半溶性或不溶性材料。当固体支持物包括接头、支架、构建模块(building block)或其他与其附接的反应性部分时,固体支持物可以被称为“官能化的”,而当固体支持物缺少这种与其附接的反应性部分时,固体支持物可以被称为“非官能化的”。固体支持物可以以溶液中游离,诸如在微量滴定孔中的形式;以流通形式,诸如在柱中;或以浸量尺(dipstick)使用。

[0226] 固体支持物可以包括膜、纸(paper)、塑料、涂覆表面、平坦表面、玻璃、载玻片、芯片或其任何组合。固体支持物可以采取树脂、凝胶、微球或其他几何配置的形式。固体支持物可以包括二氧化硅芯片、微米颗粒、纳米颗粒、板、阵列、毛细管、平坦支持物诸如玻璃纤维过滤器、玻璃表面、金属表面(钢、金、银、铝、硅和铜)、玻璃支持物、塑料支持物、硅支持物、芯片、过滤器、膜、微孔板、载玻片、塑料材料包括多孔板或膜(例如,由聚乙烯、聚丙烯、聚酰胺、聚偏二氟乙烯形成),和/或晶片、梳、针或针头(例如,适于组合合成或分析的针阵列)或珠,平坦表面诸如晶片(例如,硅晶片)的凹陷或纳升孔阵列,具有凹陷的晶片(具有或不具有过滤器底部)。

[0227] 固体支持物可以包括聚合物基质(例如,凝胶、水凝胶)。聚合物基质可以能够渗透细胞内空间(例如,细胞器周围)。聚合物基质可以能够被泵送到整个循环系统。

[0228] 基底和微孔阵列

[0229] 如本文使用的,基底可以指固体支持物类型。基底可以指可以包含本公开内容的条形码或随机条形码的固体支持物。基底可以例如包括多于一个微孔。基底可以例如是包括两个或更多个微孔的孔阵列。在一些实施方案中,微孔可以包括定义体积的小的反应室。微孔可以捕获一个或更多个细胞,或者捕获仅一个细胞。在一些实施方案中,微孔可以捕获一个或更多个固体支持物。在一些实施方案中,微孔可以仅捕获一个固体支持物。在一些实施方案中,微孔捕获单细胞和单个固体支持物(例如,珠)。微孔可以包含本公开内容的条形码试剂。

[0230] 条形码化的方法

[0231] 本公开内容提供了用于估计身体样品(例如,组织、器官、肿瘤、细胞)中不同位置处的不同靶的数目的方法。该方法可以包括将条形码(例如,随机条形码)紧密接近样品放置,裂解样品,将不同的靶与条形码关联,对靶进行扩增和/或对靶进行数字计数。该方法还可以包括对从条形码上的空间标记获得的信息进行分析和/或将所述信息可视化。在一些实施方案中,该方法包括使样品中的多于一个靶可视化。将多于一个靶映射到样品的映射图上可以包括产生样品的二维映射图或三维映射图。可以在将样品中的多于一个靶条形码化(例如,随机条形码化)之前或之后产生二维映射图和三维映射图。将样品中的多于一个靶可视化可以包括将多于一个靶映射到样品的映射图上。将多于一个靶映射到样品的映射图上可以包括产生样品的二维映射图或三维映射图。可以在对样品中的多于一个靶进行条形码化之前或之后产生二维映射图和三维映射图。在一些实施方案中,可以在裂解样品之前或之后产生二维映射图和三维映射图。在产生二维映射图或三维映射图之前或之后裂解样品可以包括加热样品、使样品与去污剂接触、改变样品的pH或其任何组合。

[0232] 在一些实施方案中,将多于一种靶条形码化包括将多于一种条形码与多于一种靶杂交以产生条形码化靶(例如,随机条形码化靶)。将多于一种靶条形码化可以包括产生条形码化靶的索引文库。产生条形码化靶的索引文库可以用包含多于一种条形码(例如,随机条形码)的固体支持物来进行。

#### [0233] 使样品和条形码接触

[0234] 本公开内容提供了用于使样品(例如,细胞)与本公开内容的基底接触的方法。可以使包括例如细胞、器官或组织薄切片的样品与条形码(例如,随机条形码)接触。细胞可以例如通过重力流来接触,其中可以使细胞沉淀并且产生单层。样品可以是组织薄切片。可以将薄切片放置在基底上。样品可以是一维的(例如,形成平坦表面)。可以使样品(例如,细胞)分散遍及基底,例如,通过在基底上生长/培养细胞。

[0235] 当条形码紧密接近靶时,靶可以与条形码杂交。条形码可以按不可耗尽的比例接触,使得每一种不同的靶可以与本公开内容的不同条形码关联。为了确保靶与条形码之间的有效关联,可以将靶与条形码交联。

#### [0236] 细胞裂解

[0237] 在细胞和条形码的分配之后,可以将细胞裂解以释放靶分子。细胞裂解可以通过各种手段中的任何一种来完成,例如通过化学或生化手段,通过渗透冲击,或通过热裂解、机械裂解或光学裂解的手段。可以通过添加包含去污剂(例如,SDS、十二烷基硫酸锂、Triton X-100、Tween-20或NP-40)、有机溶剂(例如,甲醇或丙酮)或消化酶(例如,蛋白酶K、胃蛋白酶或胰蛋白酶)或其任何组合的细胞裂解缓冲液来裂解细胞。为了增加靶与条形码的关联,可以通过例如降低裂解物的温度和/或增加裂解物的粘度来改变靶分子的扩散速率。

[0238] 在一些实施方案中,可以使用滤纸来裂解样品。可以在滤纸上部用裂解缓冲液浸泡滤纸。可以将滤纸用压力施加至样品,这可以促进样品的裂解以及样品的靶与基底的杂交。

[0239] 在一些实施方案中,裂解可以通过机械裂解、热裂解、光学裂解和/或化学裂解来进行。化学裂解可以包括使用消化酶,诸如蛋白酶K、胃蛋白酶和胰蛋白酶。裂解可以通过将裂解缓冲液添加至基底来进行。裂解缓冲液可以包含Tris HCl。裂解缓冲液可以包含至少

约0.01M、0.05M、0.1M、0.5M或1M或更多的Tris HCl。裂解缓冲液可以包含至多约0.01M、0.05M、0.1M、0.5M或1M或更多的Tris HCl。裂解缓冲液可以包含约0.1M Tris HCl。裂解缓冲液的pH可以是至少以下、至少约以下、至多以下或至多约以下：1、2、3、4、5、6、7、8、9或10。在一些实施方案中，裂解缓冲液的pH是约7.5。裂解缓冲液可以包含盐(例如，LiCl)。裂解缓冲液中的盐的浓度可以是至少以下、至少约以下、至多以下或至多约以下：0.1M、0.5M或1M。在一些实施方案中，裂解缓冲液中的盐的浓度是约0.5M。裂解缓冲液可以包含去污剂(例如，SDS、十二烷基硫酸锂、triton X、tween、NP-40)。裂解缓冲液中的去污剂浓度可以是至少以下、至少约以下、至多以下或至多约以下：0.0001%、0.0005%、0.001%、0.005%、0.01%、0.05%、0.1%、0.5%、1%、2%、3%、4%、5%、6%或7%。在一些实施方案中，裂解缓冲液中的去污剂浓度是约1%的十二烷基硫酸锂。裂解方法中使用的的时间可以取决于所使用的去污剂的量。在一些实施方案中，使用的去污剂越多，裂解所需的时间越少。裂解缓冲液可以包含螯合剂(例如，EDTA、EGTA)。裂解缓冲液中的螯合剂浓度可以是至少约1mM、5mM、10mM、15mM、20mM、25mM或30mM或更高。裂解缓冲液中的螯合剂浓度可以是至多约1mM、5mM、10mM、15mM、20mM、25mM或30mM或更高。在一些实施方案中，裂解缓冲液中的螯合剂浓度是约10mM。裂解缓冲液可以包含还原剂(例如， $\beta$ -巯基乙醇、DTT)。裂解缓冲液中还原剂的浓度可为至少以下、至少约以下、至多以下或至多约以下：1mM、5mM、10mM、15mM或20mM。在一些实施方案中，裂解缓冲液中的还原剂浓度是约5mM。在一些实施方案中，裂解缓冲液可以包含约0.1M Tris HCl，约pH 7.5，约0.5M LiCl，约1%十二烷基硫酸锂，约10mM EDTA和约5mM DTT。

[0240] 裂解可以在约4°C、10°C、15°C、20°C、25°C或30°C的温度进行。裂解可以进行约1分钟、5分钟、10分钟、15分钟或20分钟或更多分钟。裂解的细胞可以包含至少以下、至少约以下、至多以下或至多约以下：100000、200000、300000、400000、500000、600000或700000个靶核酸分子。

[0241] 将条形码附接至靶核酸分子

[0242] 在细胞裂解和核酸分子从细胞释放之后，核酸分子可以与共定位的固体支持物的条形码随机关联。关联可以包括使条形码的靶识别区与靶核酸分子的互补部分杂交(例如，条形码的寡(dT)可以与靶的多(A)尾相互作用)。可以选择用于杂交的测定条件(例如，缓冲液pH、离子强度、温度等)以促进形成特定的稳定的杂交体。在一些实施方案中，可以将从裂解的细胞释放的核酸分子与基底上的多于一个探针关联(例如，与基底上的探针杂交)。当探针包含寡(dT)时，可以将mRNA分子与探针杂交并且逆转录。寡核苷酸的寡(dT)部分可以充当用于cDNA分子的第一链合成的引物。例如，在图2中图示的条形码化的非限制性实例中，在框216处，mRNA分子可以与珠上的条形码杂交。例如，单链的核苷酸片段可以与条形码的靶结合区杂交。

[0243] 附接还可以包括将条形码的靶识别区与靶核酸分子的一部分连接。例如，靶结合区可以包含可以能够与限制性位点突出端(例如，EcoRI粘性末端突出端)特异性杂交的核酸序列。测定程序还可以包括用限制性酶(例如，EcoRI)处理靶核酸以产生限制性位点突出端。然后条形码可以连接至包含与限制性位点突出端互补的序列的任何核酸分子。连接酶(例如，T4 DNA连接酶)可以用于连接两个片段。

[0244] 例如，在图2中图示的条形码化的非限制性实例中，在框220处，随后可以将来自多

于一个细胞(或多于一个样品)的标记的靶(例如,靶-条形码分子)汇集至例如管中。标记的靶可以通过例如回收(retrieving)条形码和/或附接靶-条形码分子的珠来汇集。

[0245] 可以通过使用磁珠和外部施加的磁场来实现附接的靶-条形码分子的基于固体支持物的集合的回收。汇集靶-条形码分子后,所有进一步的处理可以在单个反应容器中进行。进一步的处理可以包括,例如,逆转录反应、扩增反应、裂解反应、解离反应和/或核酸延伸反应。进一步的处理反应可以在微孔内进行,即,不先汇集来自多于一个细胞的标记的靶核酸分子。

#### [0246] 逆转录或核酸延伸

[0247] 本公开内容提供了使用逆转录(例如,在图2的框224处)或核酸延伸来产生靶-条形码缀合物的方法。靶-条形码缀合物可以包含条形码以及靶核酸的全部或一部分的互补序列(即,条形码化的cDNA分子,诸如随机条形码化的cDNA分子)。关联的RNA分子的逆转录可以通过添加逆转录引物连同逆转录酶而发生。逆转录引物可以是寡(dT)引物、随机六核苷酸引物或靶特异性寡核苷酸引物。寡(dT)引物的长度可以是12-18个核苷酸或可以是约12-18个核苷酸,并且与哺乳动物mRNA的3'端的内源多(A)尾结合。随机六核苷酸引物可以在各个互补位点处与mRNA结合。靶特异性寡核苷酸引物通常选择性地引发感兴趣的mRNA。

[0248] 在一些实施方案中,mRNA分子向标记的RNA分子的逆转录可以通过添加逆转录引物而发生。在一些实施方案中,逆转录引物是寡(dT)引物、随机六核苷酸引物或靶特异性寡核苷酸引物。通常,寡(dT)引物的长度是12-18个核苷酸,并且与哺乳动物mRNA的3'端的内源多(A)尾结合。随机六核苷酸引物可以在各个互补位点处与mRNA结合。靶特异性寡核苷酸引物通常选择性地引发感兴趣的mRNA。

[0249] 在一些实施方案中,靶是cDNA分子。例如,可以使用逆转录酶诸如Moloney鼠白血病病毒(MMLV)逆转录酶逆转录mRNA分子,以产生具有多(dC)尾的cDNA分子。条形码可以包括具有多(dG)尾的靶结合区。在条形码的多(dG)尾和cDNA分子的多(dC)尾之间碱基配对后,逆转录酶将模板链从细胞RNA分子转换到条形码,并继续向条形码的5'末端复制。通过这样做,得到的cDNA分子在该cDNA分子3'末端上含有条形码序列(诸如分子标记)。

[0250] 逆转录可以重复地发生以产生多于一个标记的cDNA分子。本文公开的方法可以包括进行至少、至少约、至多或至多约1次、2次、3次、4次、5次、6次、7次、8次、9次、10次、11次、12次、13次、14次、15次、16次、17次、18次、19次、20次、25次、30次、35次、40次、45次、50次、55次、60次、65次、70次、75次、80次、85次、90次、95次或100次逆转录反应。

#### [0251] 扩增

[0252] 可以进行一个或更多个核酸扩增反应(例如,在图2的框228处)以产生标记的靶核酸分子的多于一个拷贝。扩增可以以多重化方式进行,其中多于一个靶核酸序列同时进行扩增。扩增反应可以用于向核酸分子添加测序衔接子。扩增反应可以包括扩增样品标记(如果存在)的至少一部分。扩增反应可以包括扩增细胞标记和/或条形码序列(例如,分子标记)的至少一部分。扩增反应可以包括扩增样品标签、细胞标记、空间标记、条形码序列(例如,分子标记)、靶核酸或其组合的至少一部分。扩增反应可以包括扩增多于一个核酸的0.5%、1%、2%、3%、4%、5%、6%、7%、8%、9%、10%、15%、20%、25%、30%、35%、40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、95%、97%、100%或这些值中的任何两个值之间的范围或数字。该方法还可以包括进行一个或更多个cDNA合成反应以产生

包含样品标记、细胞标记、空间标记和/或条形码序列(例如,分子标记)的靶-条形码分子的一个或更多个cDNA拷贝。

[0253] 在一些实施方案中,扩增可以使用聚合酶链式反应(PCR)来进行。如本文使用的,PCR可以指用于通过DNA的互补链的引物同时延伸使特定DNA序列体外扩增的反应。如本文使用的,PCR可以涵盖反应的衍生形式,包括但不限于,RT-PCR、实时PCR、巢式PCR、定量PCR、多重化PCR、数字PCR以及组装PCR。

[0254] 标记的核酸的扩增可以包括非基于PCR的方法。非基于PCR的方法的实例包括但不限于多重置换扩增(MDA)、转录介导的扩增(TMA)、基于核酸序列的扩增(NASBA)、链置换扩增(SDA)、实时SDA、滚环扩增或环到环扩增。其他非基于PCR的扩增方法包括DNA依赖性RNA聚合酶驱动的RNA转录扩增或RNA指导的DNA合成和转录的多于一个循环以扩增DNA或RNA靶、连接酶链式反应(LCR)、和QB复制酶(QB)方法、回文探针的使用、链置换扩增、使用限制性内切核酸酶的寡核苷酸驱动的扩增、使引物与核酸序列杂交并且将所得双链体在延伸反应和扩增之前裂解的扩增方法、使用缺乏5'外切核酸酶活性的核酸聚合酶的链置换扩增、滚环扩增和分支延伸扩增(RAM)。在一些实施方案中,扩增不产生环化转录物。

[0255] 在一些实施方案中,本文公开的方法还包括对标记的核酸(例如,标记的RNA、标记的DNA、标记的cDNA)进行聚合酶链式反应以产生标记的扩增子(例如,随机标记的扩增子)。标记的扩增子可以是双链分子。双链分子可包括双链RNA分子、双链DNA分子或者与DNA分子杂交的RNA分子。双链分子的一条或两条链可以包含样品标记、空间标记、细胞标记和/或条形码序列(例如,分子标记)。标记的扩增子可以是单链分子。单链分子可以包括DNA、RNA或其组合。本公开内容的核酸可以包括合成的或改变的核酸。

[0256] 扩增可以包括使用一种或更多种非天然核苷酸。非天然核苷酸可以包括光不稳定或可触发的核苷酸。非天然核苷酸的实例可以包括但不限于肽核酸(PNA)、吗啉代核酸和锁核酸(LNA)以及乙二醇核酸(GNA)与苏糖核酸(TNA)。可以将非天然核苷酸添加至扩增反应的一个或更多个循环中。添加非天然核苷酸可以用于鉴定扩增反应中特定循环或时间点的产物。

[0257] 进行一个或更多个扩增反应可以包括使用一种或更多种引物。一种或更多种引物可以包含以下或包含约以下:1个、2个、3个、4个、5个、6个、7个、8个、9个、10个、11个、12个、13个、14个或15个或更多个核苷酸。一种或更多种引物可以包含至少以下、至少约以下、至多以下或至多约以下:1个、2个、3个、4个、5个、6个、7个、8个、9个、10个、11个、12个、13个、14个或15个或更多个核苷酸。一种或更多种引物可以包含少于12-15个核苷酸。一种或更多种引物可以退火至多于一个标记的靶(例如,随机标记的靶)的至少一部分。一种或更多种引物可以退火至多于一个标记的靶的3'端或5'端。一种或更多种引物可以退火至多于一个标记的靶的内部区域。内部区域可以与多于一个标记的靶的3'末端距离至少以下、至少约以下、至多以下或至多约以下:50个、100个、150个、200个、220个、230个、240个、250个、260个、270个、280个、290个、300个、310个、320个、330个、340个、350个、360个、370个、380个、390个、400个、410个、420个、430个、440个、450个、460个、470个、480个、490个、500个、510个、520个、530个、540个、550个、560个、570个、580个、590个、600个、650个、700个、750个、800个、850个、900个或1000个核苷酸。一种或更多种引物可以包括一组固定的引物。一种或更多种引物可以包括至少一种或更多种定制引物。一种或更多种引物可以包括至少一种或更

多种对照引物。一种或更多种引物可以包括至少一种或更多种基因特异性引物。

[0258] 一种或更多种引物可以包括通用引物。通用引物可以退火至通用引物结合位点。一种或更多种定制引物可以退火至第一样品标记、第二样品标记、空间标记、细胞标记、条形码序列(例如,分子标记)、靶或其任何组合。一种或更多种引物可以包括通用引物和定制引物。定制引物可以被设计成扩增一种或更多种靶。靶可以包括一个或更多个样品中总核酸的子集。靶可以包括一个或更多个样品中总标记靶的子集。一种或更多种引物可以包括至少96种或更多种定制引物、至少960种或更多种定制引物、或至少9600种或更多种定制引物。一种或更多种定制引物可以退火至两种或更多种不同的标记的核酸。两种或更多种不同的标记的核酸可以对应于一种或更多种基因。

[0259] 可以在本公开内容的方法中使用任何扩增方案。例如,在一种方案中,第一轮PCR可以使用基因特异性引物和针对通用Illumina测序引物1序列的引物来扩增附接至珠的分子。第二轮PCR可以使用侧翼为Illumina测序引物2序列的巢式基因特异性引物和针对通用Illumina测序引物1序列的引物扩增第一PCR产物。第三轮PCR添加P5和P7以及样品索引,以使PCR产物变成Illumina测序文库。使用150bp×2测序的测序可以揭示读段1上的细胞标记和条形码序列(例如,分子标记)、读段2上的基因以及索引1读段上的样品索引。

[0260] 在一些实施方案中,可以使用化学裂解将核酸从基底去除。例如,存在于核酸中的化学基团或经修饰的碱基可以用于促进将核酸从固体支持物去除。例如,酶可以用于将核酸从基底去除。例如,通过限制性内切核酸酶消化可以将核酸从基底去除。例如,用尿嘧啶-d-糖苷酶(UDG)处理含dUTP或ddUTP的核酸可以用于将核酸从基底去除。例如,可以使用进行核苷酸切除的酶(诸如,碱基切除修复酶,诸如无嘌呤/无嘧啶(apurinic/aprimidinic, AP)内切核酸酶)将核酸从基底去除。在一些实施方案中,可以使用可光裂解基团以及光将核酸从基底去除。在一些实施方案中,可以使用可裂解接头将核酸从基底去除。例如,可裂解接头可以包括以下中的至少一种:生物素/抗生物素蛋白、生物素/链霉抗生物素蛋白、生物素/中性抗生物素蛋白、Ig蛋白A、光不稳定型接头、酸或碱不稳定型接头基团或适配体。

[0261] 当探针是基因特异性时,可以将分子与探针杂交,并且逆转录和/或扩增。在一些实施方案中,在核酸已经合成(例如,逆转录)之后,核酸可以被扩增。扩增可以以多重方式进行,其中多种靶核酸序列同时扩增。扩增可以将测序衔接子添加至核酸。

[0262] 在一些实施方案中,可以例如用桥接扩增在基底上进行扩增。cDNA可以加同聚物尾,以便产生相容末端,用于使用基底上的寡(dT)探针进行桥接扩增。在桥接扩增中,与模板核酸的3'末端互补的引物可以是共价地附接至固体颗粒的每对引物中的第一引物。当包含模板核酸的样品与颗粒接触并进行单个热循环时,可以将模板分子退火至第一引物,并且第一引物通过添加核苷酸而向前延长以形成双链体分子,所述双链体分子由模板分子和与模板互补的新形成的DNA链构成。在下一循环的加热步骤中,双链体分子可以变性,从颗粒释放模板分子并且留下通过第一引物附接至颗粒的互补DNA链。在随后的退火和延长步骤的退火阶段中,互补链可以与第二引物杂交,第二引物在从第一引物去除的位置处与互补链的区段互补。这种杂交可导致互补链在第一引物和第二引物之间形成桥,通过共价键连接第一引物并通过杂交连接第二引物。在延长阶段,通过在同一反应混合物中添加核苷酸,第二引物可以在反向方向上延长,从而将桥转化为双链桥。然后开始下一个循环,并且双链桥可以变性以产生两个单链核酸分子,每个单链核酸分子具有的一个末端分别经由第

一引物和第二引物附接至颗粒表面,其中每个单链核酸分子的另一个末端是未附接的。在这第二个循环的退火和延长步骤中,每条链可以与同一颗粒上先前未使用的另外的互补引物杂交,以形成新的单链桥。现在杂交的两个先前未使用的引物延长从而将两个新的桥转换成双链桥。

[0263] 扩增反应可以包括扩增多于一种核酸的至少以下、至少约以下、至多以下或至多约以下:1%、2%、3%、4%、5%、6%、7%、8%、9%、10%、15%、20%、25%、30%、35%、40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、95%、97%或100%。

[0264] 标记的核酸的扩增可以包括基于PCR的方法或非基于PCR的方法。标记的核酸的扩增可以包括对标记的核酸的指数式扩增。标记的核酸的扩增可以包括对标记的核酸的线性扩增。扩增可以通过PCR进行。PCR可以涵盖反应的衍生形式,包括但不限于,RT-PCR、实时PCR、巢式PCR、定量PCR、多重化PCR、数字PCR、抑制PCR、半抑制PCR以及组装PCR。

[0265] 在一些实施方案中,标记的核酸的扩增包括非基于PCR的方法,例如本文描述的任何基于非PCR的方法(例如,MDA、TMA、NASBA、SDA、实时SDA、滚环扩增或环到环扩增、DNA依赖性RNA聚合酶驱动的RNA转录扩增或RNA指导的DNA合成和转录的多于一个循环以扩增DNA或RNA靶、LCR、QB方法、回文探针的使用、链置换扩增、使用限制性内切核酸酶的寡核苷酸驱动的扩增、使引物与核酸序列杂交并且将所得双链体在延伸反应和扩增之前裂解的扩增方法、使用缺乏5'外切核酸酶活性的核酸聚合酶的链置换扩增、滚环扩增和RAM)。

[0266] 在一些实施方案中,本文公开的方法还包括对扩增的扩增子(例如,靶)进行巢式聚合酶链式反应。扩增子可以是双链分子。双链(ds)分子可以包括dsRNA分子、dsDNA分子或与DNA分子杂交的RNA分子。双链分子的一条或两条链可以包含样品标签或分子标识符标记。可选地,扩增子可以是单链分子。单链分子可以包括DNA、RNA或其组合。本发明的核酸可以包括合成的或改变的核酸。

[0267] 在一些实施方案中,该方法包括反复扩增标记的核酸以产生多于一个扩增子。本文公开的方法可以包括进行至少、至少约、至多或至多约1次、2次、3次、4次、5次、6次、7次、8次、9次、10次、11次、12次、13次、14次、15次、16次、17次、18次、19次、20次、25次、30次、35次、40次、45次、50次、55次、60次、65次、70次、75次、80次、85次、90次、95次或100次扩增反应。

[0268] 扩增还可以包括将一种或更多种对照核酸添加至一个或多个包含多于一个核酸的样品中。扩增还可以包括将一种或更多种对照核酸添加至多于一个核酸。对照核酸可以包含对照标记。

[0269] 扩增可以包括使用一种或更多种非天然核苷酸。非天然核苷酸可以包括光不稳定型和/或可触发的核苷酸。非天然核苷酸的实例包括但不限于PNA、吗啉代核酸和LNA,以及GNA和TNA。可以将非天然核苷酸添加至扩增反应的一个或多个循环中。添加非天然核苷酸可以用于鉴定扩增反应中特定循环或时间点的产物。

[0270] 进行一个或多个扩增反应可以包括使用一种或更多种引物。一种或更多种引物可以包括一种或更多种寡核苷酸。一种或更多种寡核苷酸可以包含至少约7-9个核苷酸。一种或更多种寡核苷酸可以包含少于12-15个核苷酸。一种或更多种引物可以退火至多于一个标记的核酸的至少一部分。一种或更多种引物可以退火至多于一个标记的核酸的3'端和/或5'端。一种或更多种引物可以退火至多于一个标记的核酸的内部区域。内部区域可以与多于一个标记的核酸的3'末端距离至少约50个、100个、150个、200个、220个、230个、240

个、250个、260个、270个、280个、290个、300个、310个、320个、330个、340个、350个、360个、370个、380个、390个、400个、410个、420个、430个、440个、450个、460个、470个、480个、490个、500个、510个、520个、530个、540个、550个、560个、570个、580个、590个、600个、650个、700个、750个、800个、850个、900个或1000个核苷酸。一种或更多种引物可以包括一组固定的引物。一种或更多种引物可以包括至少一种或更多种定制引物、至少一种或更多种对照引物、至少一种或更多种管家基因引物或其组合。一种或更多种引物可以包括通用引物。通用引物可以退火至通用引物结合位点。一种或更多种定制引物可以退火至第一样品标签、第二样品标签、分子标识符标记、核酸或其产物。一种或更多种引物可以包括通用引物和定制引物。定制引物可以被设计成扩增一种或更多种靶核酸。靶核酸可以包括一个或更多个样品中总核酸的子集。在一些实施方案中，引物是与本公开内容的阵列附接的探针。

[0271] 在一些实施方案中，将样品中的多于一个靶条形码化（例如，随机条形码化）还包括产生条形码化靶（例如，随机条形码化靶）或靶的条形码化片段的索引文库。不同的条形码的条形码序列（例如，不同的随机条形码的分子标记）可以彼此不同。产生条形码化靶的索引文库包括从样品中的多于一个靶产生多于一个索引多核苷酸。例如，对于包括第一索引靶和第二索引靶的条形码化靶的索引文库，第一索引多核苷酸的标记区与第二索引多核苷酸的标记区可以相差以下、相差约以下、相差至少以下或相差至多以下：1个、2个、3个、4个、5个、6个、7个、8个、9个、10个、20个、30个、40个、50个核苷酸或这些值中的任何两个值之间的数字或范围的核苷酸。在一些实施方案中，产生条形码化靶的索引文库包括使多于一个靶（例如mRNA分子）与包含多（T）区和标记区的多于一个寡核苷酸接触；以及使用逆转录酶进行第一链合成以产生单链标记的cDNA分子（每一种包含cDNA区和标记区），其中多于一个靶包括至少两种不同序列的mRNA分子，且多于一个寡核苷酸包括至少两种不同序列的寡核苷酸。产生条形码化靶的索引文库还可以包括扩增单链标记的cDNA分子以产生双链标记的cDNA分子；以及对双链标记的cDNA分子进行巢式PCR以产生标记的扩增子。在一些实施方案中，该方法可以包括产生衔接子标记的扩增子。

[0272] 条形码化（例如，随机条形码化）可以包括使用核酸条形码或标签以标记个体核酸（例如，DNA或RNA）分子。在一些实施方案中，其涉及在从mRNA产生cDNA分子时将DNA条形码或标签添加至cDNA分子。可以进行巢式PCR以使PCR扩增偏倚最小化。可以添加用于测序（例如下一代测序（NGS））使用的衔接子。例如在图2的框232处，可以使用测序结果来确定靶的一个或更多个拷贝的细胞标记、分子标记和核苷酸片段的序列。

[0273] 图3是示出了产生条形码化靶（例如，随机条形码化靶）的索引文库，诸如条形码化的mRNA或其片段的索引文库的非限制性示例性过程的示意图。如步骤1中示出的，逆转录过程可以用独特分子标记序列、细胞标记序列和通用PCR位点对每个mRNA分子进行编码。具体地，通过将一组条形码（例如，随机条形码）310与RNA分子302的多（A）尾区308杂交（例如，随机杂交），可以将RNA分子302逆转录以产生标记的cDNA分子304（包括cDNA区306）。条形码310中的每一个可以包括靶结合区，例如多（dT）区312、标记区314（例如，条形码序列或分子）和通用PCR区316。

[0274] 在一些实施方案中，细胞标记序列可以包含3个至20个核苷酸。在一些实施方案中，分子标记序列可以包含3个至20个核苷酸。在一些实施方案中，多于一个随机条形码中的每一个还包括通用标记和细胞标记中的一种或更多种，其中通用标记对于固体支持物上

的多于一个随机条形码是相同的,并且细胞标记对于固体支持物上的多于一个随机条形码是相同的。在一些实施方案中,通用标记可以包含3个至20个核苷酸。在一些实施方案中,细胞标记包含3个至20个核苷酸。

[0275] 在一些实施方案中,标记区314可以包含条形码序列或分子标记318和细胞标记320。在一些实施方案中,标记区314可以包括通用标记、维度标记和细胞标记中的一种或更多种。条形码序列或分子标记318的长度可以是以下、可以是约以下、可以是至少以下或可以是至多以下:1个、2个、3个、4个、5个、6个、7个、8个、9个、10个、20个、30个、40个、50个、60个、70个、80个、90个、100个或这些值中的任何两个值之间的数字或范围的核苷酸。细胞标记320的长度可以是以下、可以是约以下、可以是至少以下或可以是至多以下:1个、2个、3个、4个、5个、6个、7个、8个、9个、10个、20个、30个、40个、50个、60个、70个、80个、90个、100个或这些值中的任何两个值之间的数字或范围的核苷酸。通用标记的长度可以是以下、可以是约以下、可以是至少以下或可以是至多以下:1个、2个、3个、4个、5个、6个、7个、8个、9个、10个、20个、30个、40个、50个、60个、70个、80个、90个、100个或这些值中的任何两个值之间的数字或范围的核苷酸。通用标记对于固体支持物上的多于一个随机条形码可以是相同的,并且细胞标记对于固体支持物上的多于一个随机条形码是相同的。维度标记的长度可以是以下、可以是约以下、可以是至少以下或可以是至多以下:1个、2个、3个、4个、5个、6个、7个、8个、9个、10个、20个、30个、40个、50个、60个、70个、80个、90个、100个或这些值中的任何两个值之间的数字或范围的核苷酸。

[0276] 在一些实施方案中,标记区314可以包括以下、可以包括约以下、可以包括至少以下或可以包括至多以下:1种、2种、3种、4种、5种、6种、7种、8种、9种、10种、20种、30种、40种、50种、60种、70种、80种、90种、100种、200种、300种、400种、500种、600种、700种、800种、900种、1000种或这些值中的任何两个值之间的数字或范围的不同标记,诸如条形码序列或分子标记318和细胞标记320。每一种标记的长度可以是以下、可以是约以下、可以是至少以下或可以是至多以下:1个、2个、3个、4个、5个、6个、7个、8个、9个、10个、20个、30个、40个、50个、60个、70个、80个、90个、100个核苷酸或这些值中的任何两个值之间的数字或范围的核苷酸。一组条形码或随机条形码310可以含有以下、可以含有约以下、可以含有至少以下或可以含有至多以下:10种、20种、40种、50种、70种、80种、90种、 $10^2$ 种、 $10^3$ 种、 $10^4$ 种、 $10^5$ 种、 $10^6$ 种、 $10^7$ 种、 $10^8$ 种、 $10^9$ 种、 $10^{10}$ 种、 $10^{11}$ 种、 $10^{12}$ 种、 $10^{13}$ 种、 $10^{14}$ 种、 $10^{15}$ 种、 $10^{20}$ 种或这些值中的任何两个值之间的数字或范围的条形码或随机条形码310。并且条形码或随机条形码310的组可以例如,各自包含独特标记区314。标记的cDNA分子304可以进行纯化以去除过量的条形码或随机条形码310。纯化可以包括Ampure珠纯化。

[0277] 如步骤2中示出的,来自步骤1中的逆转录过程的产物可以汇集至1支管中,并且用第1PCR引物池和第1通用PCR引物进行PCR扩增。因为独特标记区314,汇集是可能的。特别地,可以将标记的cDNA分子304扩增以产生巢式PCR标记的扩增子322。扩增可以包括多重PCR扩增。扩增可以包括以单一反应体积用96种多重引物进行的多重PCR扩增。在一些实施方案中,在单一反应体积中,多重PCR扩增可以利用、利用约、利用至少或利用至多10、20、40、50、70、80、90、 $10^2$ 、 $10^3$ 、 $10^4$ 、 $10^5$ 、 $10^6$ 、 $10^7$ 、 $10^8$ 、 $10^9$ 、 $10^{10}$ 、 $10^{11}$ 、 $10^{12}$ 、 $10^{13}$ 、 $10^{14}$ 、 $10^{15}$ 、 $10^{20}$ 个或这些值中的任何两个值之间的数字或范围的多重引物。扩增可以包括使用包括靶向特定基因的定制引物326A-C的第1PCR引物池324和通用引物328。定制引物326可以与标记的

cDNA分子304的cDNA部分306' 内的区域杂交。通用引物328可以与标记的cDNA分子304的通用PCR区域316杂交。

[0278] 如图3的步骤3中示出的,来自步骤2中的PCR扩增的产物可以用巢式PCR引物池和第2通用PCR引物扩增。巢式PCR可以使PCR扩增偏倚最小化。特别地,巢式PCR标记的扩增子322可通过巢式PCR进行进一步扩增。巢式PCR可以包括在单个反应体积中用巢式PCR引物332a-c的巢式PCR引物池330和第2通用PCR引物328' 进行的多重PCR。巢式PCR引物池328可以包含以下、可以包含约以下、可以包含至少以下或可以包含至多以下:1种、2种、3种、4种、5种、6种、7种、8种、9种、10种、20种、30种、40种、50种、60种、70种、80种、90种、100种、200种、300种、400种、500种、600种、700种、800种、900种、1000种或这些值中的任何两个值之间的数字或范围的不同的巢式PCR引物332。巢式PCR引物330可以包含衔接子334,并与标记的扩增子322的cDNA部分306' 内的区域杂交。通用引物328' 可以包含衔接子336,并与标记的扩增子322的通用PCR区域316杂交。由此,步骤3产生衔接子标记的扩增子338。在一些实施方案中,巢式PCR引物332和第2通用PCR引物328' 可以不包含衔接子334和衔接子336。而是,衔接子334和衔接子336可以连接至巢式PCR的产物以产生衔接子标记的扩增子338。

[0279] 如步骤4中示出的,可以使用文库扩增引物将来自步骤3的PCR产物进行PCR扩增用于测序。特别地,可以使用衔接子334和衔接子336对衔接子标记的扩增子338进行一个或更多个另外的测定。衔接子334和衔接子336可以与引物340和引物342杂交。一种或更多种引物340和引物342可以是PCR扩增引物。一种或更多种引物340和引物342可以是测序引物。一种或更多种衔接子334和衔接子336可以用于衔接子标记的扩增子338的进一步扩增。一种或更多种衔接子334和衔接子336可以用于对衔接子标记的扩增子338测序。引物342可以包含板索引344,使得使用同一组条形码或随机条形码310产生的扩增子可以使用下一代测序(NGS)在一个测序反应中测序。

[0280] 利用灵敏的受体结合试剂特异性寡核苷酸和细胞组分结合试剂特异性寡核苷酸对单细胞群体进行多组学表征

[0281] 在一些实施方案中,提供了用于产生dCODE Dextramer文库、细胞组分结合试剂特异性寡核苷酸(例如AbSeq、蛋白谱分析)文库和/或mRNA单细胞文库的方法、组合物、系统和试剂盒。在一些实施方案中,提供了用于Immudex dextramer技术与BD Rhapsody系统的相容性的独特引物设计。在一些实施方案中,所公开的方法和组合物允许高表达的AbSeq抗体与低表达的dextramer的单独的下一代测序文库制备,这使用户能够减少检测dextramer分子的测序成本。

[0282] Immudex dextramer已经用于使用目前可用的平台来对TCR/MHC/肽复合物进行谱分析,其中用户使用与用于蛋白文库相同的引物来扩增dextramer信号。本文的公开内容包括使得能够在BD Rhapsody平台上使用dextramer的方法、组合物和试剂盒。在一些实施方案中,所公开的方法和组合物采用新的引物来扩增dextramer条形码文库。与上述目前可用的方法不同,本文提供的方法可以采用单独的引物来制备蛋白文库("AbSeq")和dextramer文库。不受任何特定理论的约束,当使用相同的扩增手柄("引物")扩增所有产物时,文库中产物的比例是固定的。使用单独的引物允许用户调整产物的比例,这可以显著降低获得相同的分辨率的测序成本。这可以应用于调整dextramer:蛋白的比例,如本文公开的。图4A-图4B描绘了在不使用独特引物(图4A)和使用独特引物(图4B)的情况下产生的AbSeq和

Dextramer测序文库的非限制性示例性示意图。当dextramer和蛋白条形码一起扩增(例如,使用相同的引物)时,每种标志物的相对比例无法调整。例如,如果蛋白A的表达为dextramer A的10,000倍高,则条形码的比例将为~10,000:1。这导致高测序成本,因为必须对蛋白A条形码的所有10,000个拷贝进行测序才能观察到dextramer A条形码的单个拷贝。使用独立于AbSeq文库的用于dextramer文库的独特引物允许用户产生用于测序的两种独立的扩增产物(AbSeq文库和dextramer文库)。这意味着可以在测序输入期间调整相对比例,这可以允许用户观察到低表达的dextramer条形码,而无需对蛋白条形码的所有分子进行测序。在本文提供的一些实施方案中,使用mRNA靶引物、dCODE引物和AbSeq引物对捕获珠结合的cDNA进行PCR1扩增(例如,如图5中示出的)。dCODE dextramer寡核苷酸(例如,受体结合试剂特异性寡核苷酸)和AbSeq蛋白谱分析寡核苷酸(例如,细胞组分结合试剂特异性寡核苷酸)可以包含不同的扩增手柄(例如,分别是第二通用序列与第三通用序列)。BD AbSeq和dCODE® PCR1产物可以与mRNA靶向PCR1产物分离(例如,通过双侧尺寸选择)。dCODE®文库可以与AbSeq文库从PCR1共纯化。dCODE®和BD Rhapsody mRNA靶向PCR1产物可以进行PCR2扩增。dCODE、AbSeq和mRNA文库可以用文库索引引物进行单独的索引PCR。在索引PCR之后,可以组合dCODE®、BD Rhapsody mRNA和BD AbSeq文库以用于测序(例如,以各索引文库的预定比例)。

[0283] 本文的公开内容包括用于在单细胞分析中使用dextramer的方法、组合物、试剂盒和系统。在一些实施方案中,引物被配置为允许dextramer与单细胞分析系统(例如,BD Rhapsody™系统)一起使用,实现对高表达的AbSeq™抗体与低表达的dextramer的单独的下一代测序文库制备,这使得用户能够减少检测dextramer分子的测序成本。

[0284] dextramer,例如来自Immudex的dextramer,已经被用于对TCR/MHC/肽复合物进行谱分析。Immudex的dCODE Dextramers™及其使用已经在WO 2015/185067、WO 2015/188839和WO 2002/072631中描述,其中每一项都通过引用以其整体特此并入。本文公开的方法、系统、试剂盒和组合物能够实现使用dextramer进行例如BD Rhapsody™系统的单细胞分析。在一些实施方案中,一种或更多种引物被配置为扩增dextramer条形码文库。在一些实施方案中,相同的引物可以用于扩增dextramer信号,例如,用于蛋白文库。在一些实施方案中,单独的引物分别用于扩增蛋白文库(“AbSeq”)和dextramer文库。在实施例2中示出如何在单细胞分析中使用dextramer的非限制性实例,其说明了在BD Rhapsody™单细胞分析系统上利用灵敏的dCODE Dextramer®和BD® AbSeq Ab-Oligo在单细胞水平对T细胞群体的多组学表征。实施例1描述了在BD Rhapsody™单细胞分析系统中使用dextramer(例如,dCODE Dextramer)的染色过程的非限制性实例。实施例1描述了用于在BD Rhapsody™单细胞分析系统中使用dextramer(例如,dCODE Dextramer)制备靶向mRNA文库的程序的非限制性示例。

[0285] 本文的公开内容包括方法。在一些实施方案中,方法包括使多于一种受体检测构建体与多于一个细胞接触以形成与受体检测构建体关联的第一多于一个细胞,其中多于一个细胞包含多于一种细胞组分靶、和核酸靶的拷贝,其中多于一个细胞中的一个或多个细胞包含受体结合试剂能够特异性结合的受体,并且其中多于一种受体检测构建体中的每一种包含两种或更多种受体结合试剂和受体结合试剂特异性寡核苷酸,所述受体结合试剂

特异性寡核苷酸包含用于受体结合试剂的独特受体标识符序列。方法可以包括：使多于一种细胞组分结合试剂与和受体检测构建体关联的多于一个细胞接触以形成第二多于一个细胞，其中多于一种细胞组分结合试剂中的每一种包含细胞组分结合试剂特异性寡核苷酸，所述细胞组分结合试剂特异性寡核苷酸包含用于细胞组分结合试剂的独特标识符序列，并且其中细胞组分结合试剂能够与多于一种细胞组分靶中的至少一种特异性结合。方法可以包括：用多于一种寡核苷酸条形码对细胞组分结合试剂特异性寡核苷酸进行条形码化以产生多于一种条形码化细胞组分结合试剂特异性寡核苷酸，所述多于一种条形码化细胞组分结合试剂特异性寡核苷酸各自包含与独特标识符序列的至少一部分互补的序列。方法可以包括：用多于一种寡核苷酸条形码对受体结合试剂特异性寡核苷酸进行条形码化以产生多于一种条形码化受体结合试剂特异性寡核苷酸，所述多于一种条形码化受体结合试剂特异性寡核苷酸各自包含与独特受体标识符序列的至少一部分互补的序列。方法可以包括：用多于一种寡核苷酸条形码对多于一个细胞的核酸靶的拷贝进行条形码化以产生多于一种条形码化核酸分子，所述多于一种条形码化核酸分子各自包含与核酸靶的至少一部分互补的序列。方法可以包括：产生包含多于一种核酸靶文库成员、多于一种细胞组分靶文库成员和多于一种受体文库成员的测序文库，其中产生测序文库包括：将测序衔接子附接至多于一种条形码化核酸分子或其产物，以产生多于一种核酸靶文库成员；以及将测序衔接子附接至多于一种条形码化细胞组分结合试剂特异性寡核苷酸或其产物，以产生多于一种细胞组分靶文库成员；以及将测序衔接子附接至多于一种条形码化受体结合试剂特异性寡核苷酸或其产物，以产生多于一种受体文库成员。方法可以包括：获得包含核酸靶文库成员的多于一个测序读段、细胞组分靶文库成员的多于一个测序读段和受体文库成员的多于一个测序读段的测序数据。

[0286] 本文公开的方法可以包括：使多于一种受体检测构建体与多于一个细胞接触以形成与受体检测构建体关联的第一多于一个细胞，其中多于一个细胞包含多于一种细胞组分靶，其中多于一个细胞中的一个或更多个细胞包含受体结合试剂能够特异性结合的受体，并且其中多于一种受体检测构建体中的每一种包含两种或更多种受体结合试剂和受体结合试剂特异性寡核苷酸，所述受体结合试剂特异性寡核苷酸包含用于受体结合试剂的独特受体标识符序列。方法可以包括：使多于一种细胞组分结合试剂与和受体检测构建体关联的多于一个细胞接触以形成第二多于一个细胞，其中多于一种细胞组分结合试剂中的每一种包含细胞组分结合试剂特异性寡核苷酸，所述细胞组分结合试剂特异性寡核苷酸包含用于细胞组分结合试剂的独特标识符序列，并且其中细胞组分结合试剂能够与多于一种细胞组分靶中的至少一种特异性结合。方法可以包括：用多于一种寡核苷酸条形码对细胞组分结合试剂特异性寡核苷酸进行条形码化以产生多于一种条形码化细胞组分结合试剂特异性寡核苷酸，所述多于一种条形码化细胞组分结合试剂特异性寡核苷酸各自包含与独特标识符序列的至少一部分互补的序列。方法可以包括：用多于一种寡核苷酸条形码对受体结合试剂特异性寡核苷酸进行条形码化以产生多于一种条形码化受体结合试剂特异性寡核苷酸，所述多于一种条形码化受体结合试剂特异性寡核苷酸各自包含与独特受体标识符序列的至少一部分互补的序列。方法可以包括：产生包含多于一种细胞组分靶文库成员和多于一种受体文库成员的测序文库，其中产生测序文库包括：将测序衔接子附接至多于一种条形码化细胞组分结合试剂特异性寡核苷酸或其产物，以产生多于一种细胞组分靶文库成

员;以及将测序衔接子附接至多于一种条形码化受体结合试剂特异性寡核苷酸或其产物,以产生多于一种受体文库成员。方法可以包括:获得包含细胞组分靶文库成员的多于一个测序读段和受体文库成员的多于一个测序读段的测序数据。

[0287] 用多于一种寡核苷酸条形码对多于一个细胞的核酸靶的拷贝进行条形码化可以包括:使多于一个细胞的核酸靶的拷贝与多于一种寡核苷酸条形码接触,其中多于一种寡核苷酸条形码中的每一种寡核苷酸条形码包含第一通用序列、第一分子标记和能够与核酸靶杂交的靶结合区;以及使与核酸靶的拷贝杂交的多于一种寡核苷酸条形码延伸,以产生多于一种条形码化核酸分子,所述多于一种条形码化核酸分子各自包含与核酸靶的至少一部分互补的序列。

[0288] 用多于一种寡核苷酸条形码对细胞组分结合试剂特异性寡核苷酸进行条形码化可以包括:使细胞组分结合试剂特异性寡核苷酸与多于一种寡核苷酸条形码接触,其中多于一种寡核苷酸条形码中的每一种寡核苷酸条形码包含第一通用序列、第一分子标记和能够与细胞组分结合试剂特异性寡核苷酸杂交的靶结合区;以及使与细胞组分结合试剂特异性寡核苷酸杂交的多于一种寡核苷酸条形码延伸以产生多于一种条形码化细胞组分结合试剂特异性寡核苷酸,所述多于一种条形码化细胞组分结合试剂特异性寡核苷酸各自包含与独特标识符序列的至少一部分互补的序列。

[0289] 用多于一种寡核苷酸条形码对受体结合试剂特异性寡核苷酸进行条形码化可以包括:使受体结合试剂特异性寡核苷酸条形码与多于一种寡核苷酸条形码接触,其中多于一种寡核苷酸条形码中的每一种寡核苷酸条形码包含第一通用序列、第一分子标记和能够与受体结合试剂特异性寡核苷酸杂交的靶结合区;以及使与受体结合试剂特异性寡核苷酸杂交的多于一种寡核苷酸条形码延伸以产生多于一种条形码化受体结合试剂特异性寡核苷酸,所述多于一种条形码化受体结合试剂特异性寡核苷酸各自包含与独特受体标识符序列的至少一部分互补的序列。靶结合区可以包含捕获序列。在一些实施方案中,(i)细胞组分结合试剂特异性寡核苷酸包含与配置为捕获细胞组分结合试剂特异性寡核苷酸的捕获序列互补的序列和/或(ii)受体结合试剂特异性寡核苷酸包含与配置为捕获受体结合试剂特异性寡核苷酸的捕获序列互补的序列。在一些实施方案中,(i)多于一种条形码化核酸分子中的每一种条形码化核酸分子包含第一通用序列和第一分子标记;(ii)多于一种条形码化细胞组分结合试剂特异性寡核苷酸中的每一种条形码化细胞组分结合试剂特异性寡核苷酸包含第一通用序列和第一分子标记;和/或(iii)多于一种条形码化受体结合试剂特异性寡核苷酸中的每一种条形码化受体结合试剂特异性寡核苷酸包含第一通用序列和第一分子标记。

[0290] 第二多于一个细胞可以包含一个或更多个单细胞。方法可以包括:在(i)使多于一个细胞的核酸靶的拷贝与多于一种寡核苷酸条形码接触,(ii)使细胞组分结合试剂特异性寡核苷酸与多于一种寡核苷酸条形码接触,和/或(iii)使受体结合试剂特异性寡核苷酸与多于一种寡核苷酸条形码接触之前:将第二多于一个细胞分区至多于一个分区,其中多于一个分区中的分区包含来自第二多于一个细胞的单细胞;以及在包含单细胞的分区中,(i)使多于一个细胞的核酸靶的拷贝与多于一种寡核苷酸条形码接触,(ii)使细胞组分结合试剂特异性寡核苷酸与多于一种寡核苷酸条形码接触,和/或(iii)使受体结合试剂特异性寡核苷酸与多于一种寡核苷酸条形码接触。分区可以是孔或液滴。多于一种寡核苷酸条形码

可以与固体支持物关联,方法包括将固体支持物与分区中的单细胞关联,并且其中多于一个分区中的分区包含单个固体支持物。方法可以包括:在分区步骤之后且在接触步骤之前裂解单细胞。裂解单细胞可以包括加热、与去污剂接触、改变pH或其任何组合。多于一个细胞可以包括T细胞、B细胞、肿瘤细胞、髓样细胞、血细胞、正常细胞、胎儿细胞、母体细胞或其混合物。多于一种寡核苷酸条形码中的至少10种可以包含不同的第一分子标记序列。多于一种寡核苷酸条形码各自可以包含细胞标记。多于一种寡核苷酸条形码的每一种细胞标记可以包含至少6个核苷酸。多于一种寡核苷酸条形码中与同一固体支持物关联的寡核苷酸条形码可以包含相同的细胞标记。多于一种寡核苷酸条形码中与不同固体支持物关联的寡核苷酸条形码可以包含不同的细胞标记。靶结合区可以包含多(dT)区、随机序列、靶特异性序列或其组合。

[0291] 固体支持物可以包括合成颗粒、平坦表面,或其组合。多于一种寡核苷酸条形码中的至少一种寡核苷酸条形码可以固定或部分固定在合成颗粒上,或者多于一种寡核苷酸条形码中的至少一种寡核苷酸条形码可以包封或部分包封在合成颗粒中。合成颗粒可以是可破坏的(例如,可破坏的水凝胶颗粒)。合成颗粒可以包括珠(例如,珠可以是琼脂糖凝胶珠、链霉抗生物素蛋白珠、琼脂糖珠、磁珠、缀合珠、蛋白A缀合珠、蛋白G缀合珠、蛋白A/G缀合珠、蛋白L缀合珠、寡(dT)缀合珠、二氧化硅珠、二氧化硅样珠、抗生物素微珠、抗荧光染料微珠或其任何组合)。合成颗粒可以包含选自由以下组成的组的材料:聚二甲基硅氧烷(PDMS)、聚苯乙烯、玻璃、聚丙烯、琼脂糖、明胶、水凝胶、顺磁物质、陶瓷、塑料、玻璃、甲基苯乙烯、丙烯酸聚合物、钛、胶乳、琼脂糖凝胶、纤维素、尼龙、硅酮及其任何组合。多于一种寡核苷酸条形码中的每一种寡核苷酸条形码可以包含接头官能团。合成颗粒可以包含固体支持物官能团。支持物官能团和接头官能团可以彼此关联。接头官能团和支持物官能团可以单独地选自由以下组成的组:C6、生物素、链霉抗生物素蛋白、一种或更多种伯胺、一种或更多种醛、一种或更多种酮及其任何组合。

[0292] 产生测序文库可以包括:使随机引物与多于一种条形码化核酸分子接触,其中随机引物中的每一种包含第四通用序列或其互补物;以及使与多于一种条形码化核酸分子杂交的随机引物延伸以产生第一多于一种延伸产物。方法可以包括:使用能够与第一通用序列或其互补物杂交的引物和能够与第四通用序列或其互补物杂交的引物扩增第一多于一种延伸产物,从而产生第一多于一种条形码化扩增子,其中多于一种核酸靶文库成员包括第一多于一种条形码化扩增子或其产物。扩增第一多于一种延伸产物可以包括将测序引物的结合位点和/或测序衔接子、其互补序列和/或其部分的序列添加到第一多于一种延伸产物。方法可以包括:基于与第一多于一种条形码化扩增子或其产物关联的具有不同序列的第一分子标记的数目来确定一个或更多个单细胞中的每一个中核酸靶的拷贝数。

[0293] 产生测序文库可以包括:使用能够与第一通用序列或其互补物杂交的引物和能够与第四通用序列或其互补物杂交的引物扩增第一多于一种条形码化扩增子,从而产生第二多于一种条形码化扩增子;其中多于一种核酸靶文库成员包括第二多于一种条形码化扩增子或其产物。扩增第一多于一种条形码化扩增子可以包括将测序引物的结合位点和/或测序衔接子、其互补序列和/或其部分的序列添加到第一多于一种条形码化扩增子。方法可以包括:基于与第二多于一种条形码化扩增子或其产物关联的具有不同序列的第一分子标记的数目来确定一个或更多个单细胞中的每一个中核酸靶的拷贝数。第一多于一种条形码化

扩增子和/或第二多于一种条形码化扩增子可以包含全转录组扩增(WTA)产物。

[0294] 产生测序文库可以包括:使用多于一种条形码化核酸分子作为模板合成第三多于一种条形码化扩增子以产生第三多于一种条形码化扩增子,其中多于一种核酸靶文库成员包括第三多于一种条形码化扩增子或其产物。合成第三多于一种条形码化扩增子可以包括使用能够与第一通用序列或其互补物杂交的引物和靶特异性引物进行PCR扩增。合成第三多于一种条形码化扩增子可以包括将测序引物的结合位点和/或测序衔接子、其互补序列和/或其部分的序列添加到条形码化核酸分子。方法可以包括:基于与第三多于一种条形码化扩增子或其产物关联的具有不同序列的第一分子标记的数目来确定一个或多个单细胞中的每一个中核酸靶的拷贝数。在一些实施方案中,靶特异性引物与免疫受体特异性杂交。在一些实施方案中,靶特异性引物与免疫受体的恒定区、免疫受体的可变区、免疫受体的多样性区、免疫受体的可变区和多样性区的连接部或其组合特异性杂交。免疫受体可以是T细胞受体(TCR)和/或B细胞受体(BCR)。TCR可以包含TCR $\alpha$ 链、TCR $\beta$ 链、TCR $\gamma$ 链、TCR $\delta$ 链或其任何组合。BCR可以包含BCR重链和/或BCR轻链。多于一种条形码化核酸分子或其产物的多于一个测序读段的每一个可以包含(1)分子标记序列,和/或(2)核酸靶的子序列。方法可以包括:基于核酸靶文库成员的多于一个测序读段来确定一个或多个单细胞中的每一个中核酸靶的拷贝数。确定一个或多个单细胞中的每一个中核酸靶的拷贝数可以包括基于与一种或更多种核酸靶文库成员或其产物关联的具有不同序列的第一分子标记、其互补物或其组合的数目来确定一个或多个单细胞中的每一个中核酸靶的拷贝数。多于一种条形码化核酸分子可以包括条形码化脱氧核糖核酸(DNA)分子、条形码化核糖核酸(RNA)分子或其组合。核酸靶可以包括核酸分子(例如,核糖核酸(RNA)、信使RNA(mRNA)、微RNA、小干扰RNA(siRNA)、RNA降解产物、包含多(A)尾的RNA或其任何组合)。在一些实施方案中,mRNA编码免疫受体。

[0295] 细胞组分结合试剂特异性寡核苷酸可以包含第三通用序列。产生测序文库可以包括:使用能够与第一通用序列或其互补物杂交的引物和能够与第三通用序列或其互补物杂交的引物,扩增多于一种条形码化细胞组分结合试剂特异性寡核苷酸或其产物,以产生多于一种扩增的条形码化细胞组分结合试剂特异性寡核苷酸,其中多于一种细胞组分靶文库成员包括多于一种扩增的条形码化细胞组分结合试剂特异性寡核苷酸或其产物。扩增多于一种条形码化细胞组分结合试剂特异性寡核苷酸可以包括将测序引物的结合位点和/或测序衔接子、其互补序列和/或其部分的序列添加到多于一种条形码化细胞组分结合试剂特异性寡核苷酸。方法可以包括:基于细胞组分靶文库成员的多于一个测序读段来确定一个或多个单细胞中多于一种细胞组分靶中的至少一种细胞组分靶的拷贝数。

[0296] 细胞组分靶文库成员的多于一个测序读段中的每一个可以包含(1)分子标记序列,和/或(2)独特标识符序列的至少一部分。细胞组分结合试剂特异性寡核苷酸可以包含第二分子标记。多于一种细胞组分结合试剂特异性寡核苷酸中的至少10个可以包含不同的第二分子标记序列。在一些实施方案中,至少两种细胞组分结合试剂特异性寡核苷酸的第二分子标记序列是不同的,并且其中所述至少两种细胞组分结合试剂特异性寡核苷酸的独特标识符序列是相同的。在一些实施方案中,至少两种细胞组分结合试剂特异性寡核苷酸的第二分子标记序列是不同的,并且其中所述至少两种细胞组分结合试剂特异性寡核苷酸的独特标识符序列是不同的。在一些实施方案中,测序数据中与用于细胞组分结合试剂的

独特标识符序列关联的独特第一分子标记序列的数目指示一个或多个单个细胞中至少一种细胞组分靶的拷贝数,所述细胞组分结合试剂能够与至少一种细胞组分靶特异性结合。在一些实施方案中,测序数据中与用于细胞组分结合试剂的独特标识符序列关联的独特第二分子标记序列的数目指示一个或多个单个细胞中至少一种细胞组分靶的拷贝数,所述细胞组分结合试剂能够与至少一种细胞组分靶特异性结合。

[0297] 方法可以包括:在使多于一种细胞组分结合试剂与和受体检测构建体关联的第一多于一个细胞接触之后,去除多于一种细胞组分结合试剂中未与和受体检测构建体关联的第一多于一个细胞接触的一种或更多种细胞组分结合试剂。在一些实施方案中,去除未与和受体检测构建体关联的第一多于一个细胞接触的一种或更多种细胞组分结合试剂可以包括:去除未与多于一种细胞组分靶中相应的至少一种细胞组分靶接触的一种或更多种细胞组分结合试剂。

[0298] 靶结合区可以包含多(dT)区,并且细胞组分结合试剂特异性寡核苷酸可以包含多(dA)区。细胞组分结合试剂特异性寡核苷酸可以包含与多(dA)区相邻的对齐序列。在一些实施方案中,(a)对齐序列包含鸟嘌呤、胞嘧啶、胸腺嘧啶、尿嘧啶或其组合;(b)对齐序列包含多(dT)序列、多(dG)序列、多(dC)序列、多(dU)序列或其组合;和/或(c)对齐序列位于多(dA)区的5'。细胞组分结合试剂特异性寡核苷酸可以通过接头与细胞组分结合试剂关联。细胞组分结合试剂特异性寡核苷酸可以被配置为能够从细胞组分结合试剂脱离。方法可以包括:使细胞组分结合试剂特异性寡核苷酸与细胞组分结合试剂解离。接头可以包含碳链。碳链可以包含2-30个碳,并且还任选地碳链可以包含12个碳。接头可以包含5'氨基修饰物C12(5AmMC12)或其衍生物。细胞组分靶可以包括蛋白靶。细胞组分结合试剂可以包含抗体或其片段。细胞组分靶可以包括糖类、脂质、蛋白、细胞外蛋白、细胞表面蛋白、细胞标志物、B细胞受体、T细胞受体、主要组织相容性复合体、肿瘤抗原、受体、细胞内蛋白或其任何组合。细胞组分结合试剂可以包含抗体或其片段。细胞组分靶可以在细胞表面上。

[0299] 使多于一种受体检测构建体与多于一个细胞接触以形成第一多于一个细胞可以包括去除未与一种或更多种受体检测构建体结合的一个或多个细胞。去除未与一种或更多种受体检测构建体结合的一个或多个细胞可以包括选择与一种或更多种受体检测构建体结合的细胞。受体检测构建体可以包含一种或更多种另外的标记,其中选择与一种或更多种受体检测构建体结合的细胞包括流式细胞术和/或基于一种或更多种另外的标记的存在选择细胞。一种或更多种另外的标记可以包括荧光标记。去除未与和受体检测构建体关联的第一多于一个细胞接触的一种或更多种受体结合试剂可以包括:去除未与受体接触的一种或更多种受体结合试剂。

[0300] 受体结合试剂特异性寡核苷酸可以包含第二通用序列。产生测序文库可以包括:使用能够与第一通用序列或其互补物杂交的引物和能够与第二通用序列或其互补物杂交的引物,扩增多于一种条形码化受体结合试剂特异性寡核苷酸或其产物,以产生多于一种扩增的条形码化受体结合试剂特异性寡核苷酸,其中多于一种受体文库成员包括多于一种扩增的条形码化受体结合试剂特异性寡核苷酸或其产物。扩增多于一种条形码化受体结合试剂特异性寡核苷酸可以包括将测序引物的结合位点和/或测序衔接子、其互补序列和/或其部分的序列添加到多于一种条形码化受体结合试剂特异性寡核苷酸。

[0301] 受体文库成员的多于一个测序读段中的每一个可以包含(1)分子标记序列,和/或

(2) 独特受体标识符序列的至少一部分。受体文库成员的多于一个测序读段中的每一个可以包含细胞标记序列。在一些实施方案中,每种独特细胞标记序列指示第二多于一个细胞中的单细胞。方法可以包括:基于受体文库成员的多于一个测序读段来确定一个或更多个单细胞中受体的拷贝数。方法可以包括:基于受体文库成员的多于一个测序读段来确定与一个或更多个单细胞结合的受体结合试剂的身份。方法可以包括:基于受体文库成员的多于一个测序读段来确定一个或更多个单细胞中受体的身份。

[0302] 受体结合试剂特异性寡核苷酸可以包含第三分子标记。多于一种受体结合试剂特异性寡核苷酸中的至少10种可以包含不同的第三分子标记序列。在一些实施方案中,至少两种受体结合试剂特异性寡核苷酸的第三分子标记序列是不同的,并且其中所述至少两种受体结合试剂特异性寡核苷酸的独特受体标识符序列是相同的。在一些实施方案中,至少两种受体结合试剂特异性寡核苷酸的第三分子标记序列是不同的,并且其中所述至少两种受体结合试剂特异性寡核苷酸的独特受体标识符序列是不同的。在一些实施方案中,测序数据中与用于受体结合试剂的独特受体标识符序列关联的独特第一分子标记序列的数目指示一个或更多个单细胞中受体的拷贝数,所述受体结合试剂能够与至少一种受体靶特异性结合。在一些实施方案中,测序数据中与用于受体结合试剂的独特受体标识符序列关联的独特第三分子标记序列的数目指示一个或更多个单细胞中受体的拷贝数,所述受体结合试剂能够与至少一种受体特异性结合。与相同受体结合试剂关联的受体结合试剂特异性寡核苷酸可以包含相同的独特受体标识符序列,并且与不同受体结合试剂关联的受体结合试剂特异性寡核苷酸可以包含不同的独特受体标识符序列。

[0303] 受体结合试剂特异性寡核苷酸可以包括DNA、RNA、锁核酸(LNA)、肽核酸(PNA)、LNA/PNA嵌合体、LNA/DNA嵌合体、PNA/DNA嵌合体或其任何组合。受体结合试剂特异性寡核苷酸的长度可以是以下、可以是约以下、可以是至少以下或可以是至多以下:1个、2个、3个、4个、5个、6个、7个、8个、9个、10个、20个、30个、40个、50个、60个、70个、80个、90个、100个或这些值中的任何两个值之间的数字或范围的核苷酸。受体可以包括分化簇(CD)分子。受体可以包括免疫受体(例如T细胞受体(TCR))。独特受体标识符序列的长度可以是约3个核苷酸至约100个核苷酸。靶结合区可以包含多(dT)区,并且受体结合试剂特异性寡核苷酸可以包含多(dA)区。受体结合试剂特异性寡核苷酸可以包含与多(dA)区相邻的对齐序列。在一些实施方案中,(a)对齐序列包含鸟嘌呤、胞嘧啶、胸腺嘧啶、尿嘧啶或其组合;(b)对齐序列包含多(dT)序列、多(dG)序列、多(dC)序列、多(dU)序列或其组合;和/或(c)对齐序列位于多(dA)区的5'。受体结合试剂特异性寡核苷酸可以通过接头与受体结合试剂关联。受体结合试剂特异性寡核苷酸可以被配置为能够从受体结合试剂脱离。方法可以包括:使受体结合试剂特异性寡核苷酸与受体结合试剂解离。接头可以包含碳链。碳链可以包含2-30个碳,并且还任选地碳链可以包含12个碳。接头可以包含5'氨基修饰物C12(5AmMC12)或其衍生物。

[0304] 多于一种受体检测构建体可以包括一种或更多种MHC多聚体,并且其中所述两种或更多种受体结合试剂包含两种或更多种MHC-肽复合物。MHC多聚体可以包含(a-b-P)<sub>n</sub>,其中n>1,其中多肽a和多肽b一起形成能够结合肽P的功能性MHC蛋白,并且(a-b-P)是当肽P与功能性MHC蛋白结合时形成的MHC-肽复合物。MHC多聚体的每种MHC-肽复合物可以与一个或更多个多聚化结构域关联。多于一种受体检测构建体可以包含至少以下或至多以下:2种

MHC多聚体、3种MHC多聚体、4种MHC多聚体、5种MHC多聚体、6种MHC多聚体、7种MHC多聚体、8种MHC多聚体、9种MHC多聚体、10种MHC多聚体、11种MHC多聚体、12种MHC多聚体、13种MHC多聚体、14种MHC多聚体、15种MHC多聚体、16种MHC多聚体、17种MHC多聚体、18种MHC多聚体、19种MHC多聚体或20种MHC多聚体。所述MHC多聚体的每种MHC-肽复合物的个体抗原肽P可以是相同或不同的。

[0305] 多于一种受体检测构建体可以包括一种或更多种阴性对照MHC多聚体。一种或更多种阴性对照MHC多聚体,其中每种MHC多聚体可以包含阴性对照肽P。所述阴性对照肽P可以是无义肽。所述一种或更多种阴性对照MHC多聚体可以是空MHC多聚体。多于一种受体检测构建体可以包含一种或更多种阳性对照MHC多聚体,其中每种MHC多聚体可以包含阳性对照肽P。

[0306] 包含 $(a-b-P)_n$ 的所述一种或更多种MHC多聚体的n的值可以是 $1 < n \leq 1000$ 。n的值可以在2-3、3-4、4-5、5-6、6-7、7-8、8-9、9-10、10-11、11-12、12-13、13-14、14-15、15-16、16-17、17-18、18-19、19-20、20-21、21-22、22-23、23-24、24-25、25-26、26-27、27-28、28-29、29-30、30-35、35-40、40-45、45-50、50-55、55-60、60-65、65-70、75-80、80-85、85-90、90-95、95-100、100-110、110-120、120-130、130-140、140-150、150-160、160-170、170-180、180-190、190-200、200-225、225-250、250-275、275-300、300-325、325-350、350-375、375-400、400-450、450-500、500-550、550-600、600-650、650-700、700-750、750-800、800-850、850-900、900-950、或950-1000或这些值中的任何两个值之间的数字或范围之间。所述MHC蛋白可以是MHC I类。所述MHC蛋白可以是MHC I类,并且抗原肽P可以是与MHC I类结合的8-mer、9-mer、10-mer、11-mer或12-mer肽。

[0307] 多聚体MHC的每种MHC蛋白或MHC-肽复合物中的一个或更多个和一个或更多个多聚化结构域之间的关联可以是共价关联和/或非共价关联。一个或更多个多聚化结构域可以包含一个或更多个多聚化结构域连接体分子。多聚体MHC的每种MHC蛋白或MHC-肽复合物中的一个或更多个可以包含一个或更多个MHC连接体分子。一个或更多个多聚化结构域连接体分子可以包括一个或更多个链霉抗生物素蛋白和/或一个或更多个抗生物素蛋白,或其任何衍生物。一个或更多个链霉抗生物素蛋白可以包括一个或更多个四聚体链霉抗生物素蛋白变体或一个或更多个单体链霉抗生物素蛋白变体。一个或更多个MHC连接体分子可以是生物素。MHC多聚体的每种MHC蛋白或MHC-肽复合物中的一个或更多个可以通过链霉抗生物素蛋白-生物素连接或抗生物素蛋白-生物素连接与一个或更多个多聚化结构域关联。MHC多聚体的每种MHC蛋白或MHC-肽复合物中的一个或更多个可以通过接头部分与一个或更多个多聚化结构域关联。MHC多聚体的每种MHC蛋白或MHC-肽复合物中的一个或更多个可以通过天然二聚化和/或蛋白-蛋白相互作用与一个或更多个多聚化结构域关联。天然二聚化和/或蛋白-蛋白相互作用可以选自由以下组成的组:亮氨酸拉链例如AP-1的亮氨酸拉链结构域、Fos/Jun相互作用、基于酸性/碱性卷曲螺旋结构的相互作用(例如螺旋)、抗体/抗原相互作用、多核苷酸-多核苷酸相互作用例如DNA/DNA、DNA/PNA、DNA/RNA、PNA/PNA、LNA/DNA、合成分子-合成分子相互作用和蛋白-小分子相互作用、IgG二聚体蛋白、IgM多价蛋白、螯合物/金属离子结合螯合物、strep免疫球蛋白、抗体(单克隆、多克隆和重组的)、抗体片段及其衍生物、hexa-his(金属螯合部分)、hexa-hat GST(谷胱甘肽S-转移酶)谷胱甘肽亲和力和钙调蛋白结合肽(CBP)、Strep-标签、纤维素结合结构域、麦芽糖结合蛋白、S-肽标签、

壳多糖结合标签、免疫反应性表位、表位标签、E2标签、HA表位标签、Myc表位、FLAG表位、AU1表位和AU5表位、Glu-Glu表位、KT3表位、IRS表位、Btag表位、蛋白激酶-C表位、VSV表位、介导与包括糖类、脂质和蛋白的多种化合物结合的凝集素,例如Con A(直生刀豆(*Cannalia ensiformis*))或WGA(麦胚凝集蛋白)和四联凝集素或蛋白A或蛋白G(抗体亲和力)。

[0308] 一个或更多个多聚化结构域可以包含(i)一个或更多个支架;(ii)一个或更多个载体;(iii)至少一个支架和至少一个载体;和/或(iv)一个或更多个任选取代的有机分子。任选取代的有机分子可以包含一种或更多种官能化的环状结构。官能化的环状结构可以包含苯环。任选取代的有机分子可以包括支架分子,所述支架分子包含至少三个反应性基团,或至少三个适于非共价附接的位点。一个或更多个多聚化结构域可以包含一种或更多种生物细胞和/或细胞样结构。一个或更多个多聚化结构域可以包含一种或更多种膜。一种或更多种膜可以包括脂质体或胶束。一个或更多个多聚化结构域可以包含一种或更多种聚合物(例如,一种或更多种合成聚合物)。

[0309] 一种或更多种聚合物可以选自由多糖组成的组。多糖可以包含一个或更多个右旋糖酐部分。一个或更多个右旋糖酐部分可以(i)与一个或更多个MHC肽复合物共价附接;(ii)与一个或更多个MHC肽复合物非共价附接;和/或(iii)被修饰。一个或更多个右旋糖酐部分可以包含一个或更多个氨基-右旋糖酐。一个或更多个氨基-右旋糖酐可以用二乙烯基砜修饰。一个或更多个右旋糖酐部分可以包含一个或更多个具有约1,000至50,000Da(例如,约1,000至5,000、约5,000至10,000、约10,000至15,000、约15,000至20,000、约20,000至25,000、约25,000至30,000、约30,000至35,000、约35,000至40,000、约40,000至45,000或约45,000至50,000Da,或这些值中的任何两个值之间的数字或范围)的分子量的右旋糖酐。一个或更多个右旋糖酐部分可以包含一个或更多个具有为约50,000至150,000Da(例如,约50,000至60,000、约60,000至70,000、约70,000至80,000、约80,000至90,000、约90,000至100,000、约100,000至110,000、约110,000至120,000、约120,000至130,000、约130,000至140,000或约140,000至150,000Da)的分子量的右旋糖酐。一个或更多个右旋糖酐部分可以包含一个或更多个具有约150,000-270,000Da(例如,约150,000至160,000、约160,000至170,000、约170,000至180,000、约180,000至190,000、约190,000至200,000、约200,000至210,000、约210,000至220,000、约220,000至230,000、约230,000至240,000、约240,000至250,000、约250,000至260,000、约260,000至270,000、约270,000至280,000、约280,000至290,000、约290,000至300,000、约300,000至310,000、约310,000至320,000、约320,000至330,000、约330,000至340,000、约340,000至350,000、约350,000至360,000、约360,000至370,000、约370,000至380,000、约380,000至390,000、约390,000至400,000、约400,000至410,000、约410,000至420,000、约420,000至430,000、约430,000至440,000、约440,000至450,000、约450,000至460,000、约460,000至470,000、约470,000至480,000、约480,000至490,000、约490,000至500,000、约500,000至550,000、约550,000至600,000、约600,000至650,000、约650,000至700,000、约700,000至750,000、约750,000至800,000、约800,000至850,000、约850,000至900,000、约900,000至950,000、或约950,000至1,000,000Da,或这些值中的任何两个值之间的数字或范围)的分子量的右旋糖酐。一个或更多个右旋糖酐部分可以是线性的和/或支链的。

[0310] 一种或更多种合成聚合物可以包括PNA、聚酰胺、PEG或其任何组合。一个或更多个

多聚化结构域可以包含一种或更多种选自由以下组成的组的实体：IgG结构域、卷曲螺旋多肽结构、DNA双链体、核酸双链体、PNA-PNA、PNA-DNA和DNA-RNA。一个或更多个多聚化结构域可以包含抗体。抗体可以选自由以下组成的组：多克隆抗体、单克隆抗体、IgA、IgG、IgM、IgD、IgE、IgG1、IgG2、IgG3、IgG4、IgA1、IgA2、IgM1、IgM2、人源化抗体、人源化单克隆抗体、嵌合抗体、小鼠抗体、大鼠抗体、兔抗体、人类抗体、骆驼抗体、绵羊抗体、工程化人类抗体、表位聚焦抗体、激动剂抗体、拮抗剂抗体、中和抗体、天然存在的抗体、分离的抗体、单价抗体、双特异性抗体、三特异性抗体、多特异性抗体、异源缀合抗体、免疫缀合物、免疫脂质体、标记的抗体、抗体片段、结构域抗体、纳米抗体、微抗体、大抗体、双体和融合抗体。一个或更多个多聚化结构域可以包含一种或更多种有机小支架分子或有机小分子。一种或更多种有机小分子可以包括一种或更多种类固醇、一种或更多种肽和/或一种或更多种芳族有机分子。一种或更多种芳族有机分子可以包含一种或更多种双环结构、一种或更多种多环结构或一种或更多种单环结构；任选所述一种或更多种单环结构可以包括一种或更多种任选官能化或取代的苯环。一个或更多个多聚化结构域可以包含一种或更多种：(i)能够聚合的单体分子；(ii)生物聚合物，诸如一种或更多种蛋白；(iii)小分子支架；(iv)超分子结构，诸如一种或更多种纳米团簇；和/或(v)蛋白复合物。一个或更多个多聚化结构域可以包含一种或更多种珠。一种或更多种珠可以选自由以下组成的组：带有亲电基团例如二乙烯基砜活化的多糖的珠、已经用甲苯磺酰基活化的酯官能化的聚苯乙烯珠、用甲苯磺酰基活化的酯官能化的磁性聚苯乙烯珠，以及其中MHC复合体已经通过MHC复合体中包含的亲核体与珠的亲电体的反应而共价固定至这些珠的珠。一种或更多种珠可以选自由以下组成的组：琼脂糖凝胶珠、聚丙烯酰胺葡聚糖珠、聚苯乙烯珠、琼脂糖珠、多糖珠、聚氨基甲酸酯珠和可以悬浮在水性缓冲液中的任何其他种类的珠。

[0311] 多聚化结构域可以包含一种或更多种选自由以下组成的组的化合物：琼脂糖珠、琼脂糖凝胶珠、树脂珠、玻璃珠、多孔玻璃珠、涂覆疏水聚合物的玻璃颗粒、壳聚糖涂覆珠、SH珠、胶乳珠、球形胶乳珠、等位基因型珠、SPA珠、基于PEG的树脂、PEG涂覆珠、PEG包封珠、聚苯乙烯珠、磁性聚苯乙烯珠、谷胱甘肽琼脂糖珠、磁珠、顺磁珠、蛋白A和/或蛋白G琼脂糖凝胶珠、活化羧酸珠、宏观珠、微观珠、不溶性树脂珠、基于二氧化硅的树脂、纤维素树脂、交联琼脂糖珠、聚苯乙烯珠、交联聚丙烯酰胺树脂、具有铁芯的珠、金属珠、dynabead、用NHS活化的聚甲基丙烯酸甲酯珠、链霉抗生物素蛋白-琼脂糖珠、聚丙烯、聚乙烯、右旋糖酐、尼龙、淀粉酶、天然纤维素和改性纤维素、硝化纤维素、聚丙烯酰胺、辉长岩、磁铁矿、聚合物、寡聚物、非重复部分、聚乙二醇(PEG)、单甲氧基-PEG、单-(C<sub>1</sub>-C<sub>10</sub>)烷氧基-PEG、芳氧基-PEG、聚(N-乙基吡咯烷酮)PEG、三氟乙磺酰氯单甲氧基PEG、PEG丙醛、双琥珀酰亚胺基碳酸PEG、二乙烯基苯交联聚苯乙烯珠、丙二醇均聚物、聚环氧丙烷/环氧乙烷共聚物、聚氧乙基化多元醇(例如，甘油)、聚乙烯醇、右旋糖酐、氨基右旋糖酐、基于糖类的聚合物、交联右旋糖酐珠、多糖珠、聚氨基甲酸酯珠、二乙烯基砜活化多糖、已经用甲苯磺酰基活化的酯官能化的聚苯乙烯珠、用甲苯磺酰基活化的酯官能化的磁性聚苯乙烯珠、链霉抗生物素蛋白珠、链霉抗生物素蛋白单体涂覆的珠、链霉抗生物素蛋白四聚体涂覆的珠、链霉抗生物素蛋白涂覆的Compel磁珠、抗生物素蛋白涂覆的珠、dextramer涂覆的珠、二乙烯基砜活化的右旋糖酐、羧酸修饰的珠、胺修饰的珠、抗体涂覆的珠、纤维素珠、接枝共聚珠、聚丙烯酰胺珠、任选地与N-N'-双丙烯酰基乙二胺交联的二甲基丙烯酰胺珠、中空纤维膜、荧光珠、胶原-琼脂糖

珠、明胶珠、胶原-明胶珠、胶原-纤连蛋白-明胶珠、胶原珠、壳聚糖珠、胶原-壳聚糖珠、基于蛋白的珠、水凝胶珠、半纤维素、烷基纤维素、羟烷基纤维素、羧甲基纤维素、磺基乙基纤维素、淀粉、木聚糖、支链淀粉、软骨素、透明质酸、肝素、瓜尔胶、黄原胶、甘露聚糖、半乳甘露聚糖、壳多糖和壳聚糖。

[0312] 一个或更多个多聚化结构域可以包含二聚化结构域、三聚化结构域、四聚化结构域、五聚化结构域或六聚化结构域。一个或更多个多聚化结构域可以包含能够附接一个或更多个支架的聚合物结构。聚合物结构可以包含多糖。多糖可以包含一个或更多个右旋糖酐部分。一个或更多个多聚化结构域可以包含聚酰胺、聚乙二醇、多糖、琼脂糖凝胶或其任何组合。一个或更多个多聚化结构域可以包含羧甲基右旋糖酐、右旋糖酐聚醛、羧甲基右旋糖酐内酯、环糊精或其任何组合。在一些实施方案中，一个或更多个多聚化结构域具有小于1,000Da、1,000Da至小于10,000Da、10,000Da至小于100,000Da、100,000Da至小于1,000,000Da、超过1,000,000Da或这些值中的任何两个值之间的数字或范围的分子量。所述一种或更多种MHC多聚体还可以包含一个或更多个选自由以下组成的组的支架、载体和/或接头：链霉抗生物素蛋白(SA)和抗生物素蛋白及其衍生物、生物素、免疫球蛋白、抗体(单克隆、多克隆和重组的)、抗体片段及其衍生物、AP-1的亮氨酸拉链结构域(jun和fos)、hexahis(金属螯合部分)、hexa-hat GST(谷胱甘肽S-转移酶)谷胱甘肽亲和力、钙调蛋白结合肽(CBP)、Strep-标签、纤维素结合结构域、麦芽糖结合蛋白、S-肽标签、壳多糖结合标签、免疫反应性表位、表位标签、E2标签、HA表位标签、Myc表位、FLAG表位、AU1和AU5表位、Glu-Glu表位、KT3表位、IRS表位、Btag表位、蛋白激酶-C表位、VSV表位、介导与包括糖类、脂质和蛋白的多种化合物结合的凝集素，例如Con A(直生刀豆)或WGA(麦胚凝集蛋白)和四联凝集素或蛋白A或蛋白G(抗体亲和力)。所述MHC多聚体可以包含由多聚化结构域连接部分连接的多于一个相同或不同的多聚化结构域。所述MHC多聚体可以包含连接至第二多聚化结构域的第一多聚化结构域。

[0313] 所述一种或更多种MHC多聚体可以包含一个或更多个标记。一个或更多个标记可以包括受体结合试剂特异性寡核苷酸。一个或更多个标记可以包括受体结合试剂特异性寡核苷酸和一种或更多种另外的标记。一种或更多种另外的标记可以选自包括以下的组：肽标记、荧光团标记、重金属标记、同位素标记、放射标记、放射性核素、稳定同位素、同位素链和单原子、化学发光标记、生物发光标记、放射性标记、酶标记、DNA荧光染色剂、镧系元素、离子载体、与特定离子结合的螯合化合物或其任何组合。一个或更多个标记可以包括共价附接的标记和/或非共价附接的标记。可以将一个或更多个标记与以下附接：(i)MHC多肽a；(ii)MHC多肽b；(iii)肽P；(iv)一个或更多个多聚化结构域；和/或(v)  $(a-b-P)_n$ 。一个或更多个标记可以通过链霉抗生物素蛋白-生物素连接附接至  $(a-b-P)_n$ 。一种或更多种另外的标记可以是荧光团标记。荧光团标记可以选自包括以下的组：异硫氰酸荧光素、罗丹明、藻红蛋白、藻蓝蛋白、别藻蓝蛋白、邻苯二甲醛、荧光胺；2-(4'-马来酰亚胺基苯胺基)萘-6-磺酸钠盐；5-(((2-碘乙酰基)氨基)乙基)氨基)萘-1-磺酸；苝-1-丁酸；AlexaFluor 350(7-氨基-6-磺酸-4-甲基香豆素-3-乙酸)；AMCA(7-氨基-4-甲基香豆素-3-乙酸)；7-羟基-4-甲基香豆素-3-乙酸；Marina Blue(6,8-二氟-7-羟基-4-甲基香豆素-3-乙酸)；7-二甲基氨基-香豆素-4-乙酸；荧光胺-N-丁胺加合物；7-羟基-香豆素-3-羧酸；CascadeBlue(苝-三磺酸乙酰叠氮化物)；Cascade Yellow；Pacific Blue(6,8二氟-7-羟基香豆素-3-羧酸；7-二

乙基氨基-香豆素-3-羧酸;N-(((4-叠氮苯甲酰基)氨基)乙基)-4-氨基-3,6-二巯基-1,8-萘酰亚胺二钾盐;Alexa Fluor430;3-苾十二酸;8-羟基苾-1,3,6-三磺酸三钠盐;12-(N-(7-硝基苯-2-氧杂-1,3-二唑-4-基)氨基)十二酸;N,N'-二甲基-N-(碘乙酰基)-N'- (7-硝基苯-2-氧杂-1,3-二唑-4-基)乙二胺;Oregon Green 488(二氟羧基荧光素);5-碘乙酰胺荧光素;碘化丙锭-DNA加合物;羧基荧光素、Fluor染料、Pacific Blue<sup>TM</sup>、Pacific Orange<sup>TM</sup>、Cascade Yellow<sup>TM</sup>; **AlexaFluor®** (AF), AF405、AF488、AF500、AF514、AF532、AF546、AF555、AF568、AF594、AF610、AF633、AF635、AF647、AF680、AF700、AF710、AF750、AF800;基于量子点的染料, **QDot®**纳米晶体(Invitrogen, Molecular Probs)、**Qdot®**525、**Qdot®**565、**Qdot®**585、**Qdot®**605、**Qdot®**655、**Qdot®**705、**Qdot®**800;DyLight<sup>TM</sup>染料(Pierce) (DL), DL549、DL649、DL680、DL800;荧光素(Flu)或其任何衍生物,诸如FITC;Cy-染料、Cy2、Cy3、Cy3.5、Cy5、Cy5.5、Cy7;荧光蛋白,RPE、PerCp、APC、绿色荧光蛋白;GFP和GFP衍生的突变体蛋白,BFP、CFP、YFP、DsRed、T1、Dimer2、mRFP1、MBanana、mOrange、dTomato、tdTomato、mTangerine、mStrawberry、mCherry;串联染料,RPE-Cy5、RPE-Cy5.5、RPE-Cy7、**RPE-AlexaFluor®** 串联缀合物;RPE-Alexa610、RPE-TxRed、APC-Alexa600、APC-Alexa610、APC-Alexa750、APC-Cy5和APC-Cy5.5。

[0314] 一种或更多种另外的标记可以能够吸收光(例如,发色团和/或染料)。在一些实施方案中,一种或更多种另外的标记能够在激发后发射光,任选地是一种或更多种荧光染料,还任选地一种或更多种荧光染料选自**AlexaFluor®** (AF) 家族,其包括**AF®**350、AF405、AF430、AF488、AF500、AF514、AF532、AF546、AF555、AF568、AF594、AF610、AF633、AF635、AF647、AF680、AF700、AF710、AF750和AF800;选自基于量子点(**Qdot®**)的染料家族,其包括**Qdot®**525、**Qdot®**565、**Qdot®**585、**Qdot®**605、**Qdot®**655、**Qdot®**705、**Qdot®**800;选自DyLight<sup>TM</sup>染料(DL)家族,其包括DL549、DL649、DL680、DL800;选自发荧光小染料家族,其包括FITC、Pacific Blue<sup>TM</sup>、Pacific Orange<sup>TM</sup>、Cascade Yellow<sup>TM</sup>、Marina Blue<sup>TM</sup>、DSred、DSred-2、7-AAD、T0-Pro-3;选自Cy-染料家族,其包括Cy2、Cy3、Cy3.5、Cy5、Cy5.5、Cy7;选自藻胆蛋白家族,其包括R-藻红蛋白(RPE)、PerCP、别藻蓝蛋白(APC)、B-藻红蛋白、C-藻蓝蛋白;选自荧光蛋白家族,其包括(E)GFP和GFP((增强的)绿色荧光蛋白)衍生的突变体蛋白,BFP、CFP、YFP、DsRed、T1、Dimer2、mRFP1、MBanana、mOrange、dTomato、tdTomato、mTangerine;选自具有PRE的串联染料家族,其包括RPE-Cy5、RPE-Cy5.5、RPE-Cy7、**RPE-AlexaFluor®** 串联缀合物;RPE-Alexa610、RPE-TxRed;选自具有APC的串联染料家族,其包括APC-Alexa600、APC-Alexa610、APC-Alexa750、APC-Cy5、APC-Cy5.5;选自钙染料家族,其包括Indo-1-Ca<sup>2+</sup>、Indo-2-Ca<sup>2+</sup>。

[0315] 除了a和b之外,MHC多聚体还可以包含一种或更多种另外的多肽。MHC-肽复合物的一种多肽可以是重链多肽。MHC-肽复合物的一种多肽可以是b2M多肽。在一些实施方案中,(i)P被化学修饰;(ii)P被聚乙二醇化、磷酸化和/或糖基化;(iii)肽P的氨基酸残基中的一个被另一个氨基酸取代;(iv)a和b两者都是全长肽;(v)a是全长肽;(vi)b是全长肽;(vii)a被截短;(viii)b被截短;(ix)a和b两者都被截短;(x)a被共价连接至b;(xi)a被共价连接至P;(xii)b被P共价连接至P;(xiii)a、b和P都被共价连接;(xiv)a被非共价连接至b;(xv)a被

非共价连接至P; (xvi) b被非共价连接至P; (xvii) a、b和P都被非共价连接; (xviii) a不被包含在(a-b-P)复合物中; (xix) b不被包含在(a-b-P)复合物中; 和/或(xx) P不被包含在(a-b-P)复合物中。MHC多聚体可以包含一种或更多种增加稳定性的组分(例如, HEG和/或TEG)。

[0316] 方法可以包括: 基于受体文库成员的多于一个测序读段, 将一个或更多个单细胞中的T细胞受体(TCR)受体与肽P关联。方法可以包括: 测量对于肽P特异性的T细胞的存在、频率、数目、活性和/或状态, 从而检测抗原特异性T细胞应答。

[0317] 第一通用序列、第二通用序列、第三通用序列和/或第四通用序列可以是相同的。第一通用序列、第二通用序列、第三通用序列和/或第四通用序列可以是不同的。第一通用序列、第二通用序列、第三通用序列和/或第四通用序列可以包含测序引物的结合位点和/或测序衔接子、其互补序列和/或其部分。测序衔接子可以包括P5序列、P7序列、其互补序列和/或其部分。测序引物可以包括读段1测序引物、读段2测序引物、其互补序列和/或其部分。

[0318] 方法可以包括: 使(i)条形码化核酸分子、(ii)条形码化受体结合试剂特异性寡核苷酸和(iii)条形码化细胞组分结合试剂特异性寡核苷酸中的一种或更多种与(i)条形码化核酸分子、(ii)条形码化受体结合试剂特异性寡核苷酸和(iii)条形码化细胞组分结合试剂特异性寡核苷酸中的一种或更多种物理分离。在一些实施方案中, 物理分离可以包括使用一种或更多种尺寸选择试剂。

[0319] 多于一种条形码化受体结合试剂特异性寡核苷酸和多于一种条形码化细胞组分结合试剂特异性寡核苷酸可以分别扩增。第二通用序列和第三通用序列可以是不同的。第二通用序列可以与第三通用序列小于约85%(例如, 0.00000001%、0.0000001%、0.000001%、0.00001%、0.0001%、0.001%、0.01%、0.1%、1%、2%、3%、4%、5%、6%、7%、8%、9%、10%、11%、12%、13%、14%、15%、16%、17%、18%、19%、20%、21%、22%、23%、24%、25%、26%、27%、28%、29%、30%、31%、32%、33%、34%、35%、36%、37%、38%、39%、40%、41%、42%、43%、44%、45%、46%、47%、48%、49%、50%、51%、52%、53%、54%、55%、56%、57%、58%、59%、60%、61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、81%、82%、83%、84%或这些值中的任何两个值之间的数字或范围)相同。第二通用序列可以包含与SEQ ID NO:1(GGAGGGAGGTTAGCGAAGGT)至少约85%相同的序列。在一些实施方案中, 扩增多于一种条形码化受体结合试剂特异性寡核苷酸或其产物不包括扩增多于一种条形码化细胞组分结合试剂特异性寡核苷酸或其产物。

[0320] 在一些实施方案中, 产生测序文库可以包括产生包含(i)核酸靶文库成员、(ii)细胞组分靶文库成员和/或(iii)受体文库成员的测序混合物。在一些实施方案中, 产生测序混合物可以包括以预定比例混合(i)核酸靶文库成员、(ii)细胞组分靶文库成员和/或(iii)受体文库成员。在产生测序混合物之前, (i)核酸靶文库成员、(ii)细胞组分靶文库成员和/或(iii)受体文库成员可以在物理上彼此分离。测序混合物可以包含预定比例的(i)核酸靶文库成员、(ii)细胞组分靶文库成员和/或(iii)受体文库成员。细胞组分靶文库成员与受体文库成员的预定比例可以是约1:1、1.1:1、1.2:1、1.3:1、1.4:1、1.5:1、1.6:1、1.7:1、1.8:1、1.9:1、2:1、2.5:1、3:1、4:1、5:1、6:1、7:1、8:1、9:1、10:1、11:1、12:1、13:1、14:1、15:1、16:1、17:1、18:1、19:1、20:1、21:1、22:1、23:1、24:1、25:1、26:1、27:1、28:1、

29:1、30:1、31:1、32:1、33:1、34:1、35:1、36:1、37:1、38:1、39:1、40:1、41:1、42:1、43:1、44:1、45:1、46:1、47:1、48:1、49:1、50:1、51:1、52:1、53:1、54:1、55:1、56:1、57:1、58:1、59:1、60:1、61:1、62:1、63:1、64:1、65:1、66:1、67:1、68:1、69:1、70:1、71:1、72:1、73:1、74:1、75:1、76:1、77:1、78:1、79:1、80:1、81:1、82:1、83:1、84:1、85:1、86:1、87:1、88:1、89:1、90:1、91:1、92:1、93:1、94:1、95:1、96:1、97:1、98:1、99:1、100:1、200:1、300:1、400:1、500:1、600:1、700:1、800:1、900:1、1000:1、2000:1、3000:1、4000:1、5000:1、6000:1、7000:1、8000:1、9000:1、10000:1或所述值中的任何两个值之间的数字或范围。细胞组分靶文库成员与受体文库成员的预定比例可以被配置为实现可以是以下的细胞组分靶文库成员的测序读段与受体文库成员的测序读段的比例:约1:1、1.1:1、1.2:1、1.3:1、1.4:1、1.5:1、1.6:1、1.7:1、1.8:1、1.9:1、2:1、2.5:1、3:1、4:1、5:1、6:1、7:1、8:1、9:1、10:1、11:1、12:1、13:1、14:1、15:1、16:1、17:1、18:1、19:1、20:1、21:1、22:1、23:1、24:1、25:1、26:1、27:1、28:1、29:1、30:1、31:1、32:1、33:1、34:1、35:1、36:1、37:1、38:1、39:1、40:1、41:1、42:1、43:1、44:1、45:1、46:1、47:1、48:1、49:1、50:1、51:1、52:1、53:1、54:1、55:1、56:1、57:1、58:1、59:1、60:1、61:1、62:1、63:1、64:1、65:1、66:1、67:1、68:1、69:1、70:1、71:1、72:1、73:1、74:1、75:1、76:1、77:1、78:1、79:1、80:1、81:1、82:1、83:1、84:1、85:1、86:1、87:1、88:1、89:1、90:1、91:1、92:1、93:1、94:1、95:1、96:1、97:1、98:1、99:1、100:1、200:1、300:1、400:1、500:1、600:1、700:1、800:1、900:1、1000:1、2000:1、3000:1、4000:1、5000:1、6000:1、7000:1、8000:1、9000:1、或10000:1、所述值中的任何两个值之间的数字或范围。细胞组分靶文库成员的测序读段与受体文库成员的测序读段的比例可以是约1:1、1.1:1、1.2:1、1.3:1、1.4:1、1.5:1、1.6:1、1.7:1、1.8:1、1.9:1、2:1、2.5:1、3:1、4:1、5:1、6:1、7:1、8:1、9:1、10:1、11:1、12:1、13:1、14:1、15:1、16:1、17:1、18:1、19:1、20:1、21:1、22:1、23:1、24:1、25:1、26:1、27:1、28:1、29:1、30:1、31:1、32:1、33:1、34:1、35:1、36:1、37:1、38:1、39:1、40:1、41:1、42:1、43:1、44:1、45:1、46:1、47:1、48:1、49:1、50:1、51:1、52:1、53:1、54:1、55:1、56:1、57:1、58:1、59:1、60:1、61:1、62:1、63:1、64:1、65:1、66:1、67:1、68:1、69:1、70:1、71:1、72:1、73:1、74:1、75:1、76:1、77:1、78:1、79:1、80:1、81:1、82:1、83:1、84:1、85:1、86:1、87:1、88:1、89:1、90:1、91:1、92:1、93:1、94:1、95:1、96:1、97:1、98:1、99:1、100:1、200:1、300:1、400:1、500:1、600:1、700:1、800:1、900:1、1000:1、2000:1、3000:1、4000:1、5000:1、6000:1、7000:1、8000:1、9000:1、10000:1或所述值中的任何两个值之间的数字或范围。

[0321] 细胞组分靶文库成员的测序读段与受体文库成员的测序读段的比例与其中包括以下的方法相比可以为至少约2倍低(例如,2倍、3倍、4倍、5倍、6倍、7倍、8倍、9倍、10倍、20倍、30倍、40倍、50倍、60倍、70倍、80倍、90倍、100倍或这些值中的任何两个值之间的数字或范围):(i)多于一种条形码化受体结合试剂特异性寡核苷酸和多于一种条形码化细胞组分结合试剂特异性寡核苷酸被一起扩增;(ii)第二通用序列和第三通用序列相同;和/或(iii)测序混合物不包含细胞组分靶文库成员与受体文库成员的预定比例。

#### [0322] 蛋白谱分析和样品索引

[0323] 核酸靶可以包括核酸分子,诸如例如,核糖核酸(RNA)、信使RNA(mRNA)、微RNA、小干扰RNA(siRNA)、RNA降解产物、包含多(A)尾的RNA或其任何组合。核酸靶可以包括样品索引寡核苷酸。样品索引寡核苷酸可以包含样品索引序列。本文提供的多于一种样品索引组

合物中的至少两种样品索引组合物的样品索引序列可以包括不同的序列。核酸靶可以包括细胞组分结合试剂特异性寡核苷酸。细胞组分结合试剂特异性寡核苷酸可以包含用于细胞组分结合试剂的独特标识符序列。在本文提供的方法和组合物的一些实施方案中,核酸靶是结合试剂寡核苷酸(例如,抗体寡核苷酸(“AbOligo”或“AbO”)、结合试剂寡核苷酸、细胞组分结合试剂特异性寡核苷酸、样品索引寡核苷酸)。本文公开的一些实施方案提供了多于一种组合物,所述多于一种组合物各自包含与寡核苷酸(例如,结合试剂寡核苷酸)缀合的细胞组分结合试剂(诸如蛋白结合试剂),其中寡核苷酸包含用于与其缀合的细胞组分结合试剂的独特标识符。细胞组分结合试剂(诸如条形码化抗体)及其用途(诸如对细胞进行样品索引)已在US2018/0088112和US2018/0346970中描述;这些中的每一项的内容通过引用以其整体并入本文。

[0324] 用于同时确定蛋白表达和基因表达以及用于进行样品索引的系统、方法、组合物和试剂盒也在US2020/0232032中描述,其内容通过引用以其整体并入本文。在一些这样的实施方案中,与细胞组分结合试剂(例如,抗体)关联的寡核苷酸包含以下中的一种或更多种:独特分子标记序列、引物衔接子、抗体特异性条形码序列、对齐序列和/或多(A)序列。在一些实施方案中,寡核苷酸通过接头(例如,5AmMC12)与细胞组分结合试剂关联。

#### [0325] 免疫组库谱分析

[0326] 本文的公开内容包括用于将具有分子标记(或分子索引)的条形码(例如,随机条形码)附接到被条形码化或被标记的核酸靶(例如,脱氧核糖核酸分子和核糖核酸分子)的5'末端的系统、方法、组合物和试剂盒。本文公开的基于5'的转录物计数方法可以互补或补充例如基于3'的转录物计数方法(例如,Rhapsody™测定(Becton,Dickinson and Company, Franklin Lakes,NJ)、Chromium™单细胞3'解决方案(10X Genomics,San Francisco,CA))。条形码化的核酸靶可用于序列鉴定、转录物计数、选择性剪接分析、突变筛选和/或以高通量方式进行全长测序。对5'末端(相对于被标记的靶核酸靶的5')转录物计数可以揭示在核酸分子的5'末端或较靠近核酸分子的5'末端的选择性剪接同种型和变体(包括但不限于剪接变体、单核苷酸多态性(SNP)、插入、缺失、取代)。在一些实施方案中,方法可以涉及分子内杂交。使用5'条形码化和/或3'条形码化确定核酸靶(例如,免疫受体的V(D)J区)的序列的方法在US2020/0109437中描述;该文献的内容通过引用以其整体并入本文。用于在核酸靶的5'末端上进行分子条形码化的系统、方法、组合物和试剂盒已在US2019/0338278中描述,其内容通过引用以其整体并入本文。在一些实施方案中,本文提供的用于基于5'的基因表达谱分析的系统、方法、组合物和试剂盒可以与基于随机引发和延伸(RPE)的全转录组分析方法和组合物一起使用,这些方法和组合物已经在US2020/0149037中描述,其内容通过引用以其整体并入本文。

#### [0327] 抗原肽的设计和产生

[0328] 本文公开的抗原肽可以在所公开的方法中用作为本文公开的MHC多聚体的一部分。用于本文提供的抗原肽的设计和产生的特征和原理将在下文中更详细地描述。

[0329] MHC I类蛋白通常在其肽结合沟中结合八聚体、九聚体、十聚体或十一聚体(8-mer、9-mer、10-mer、11-mer)肽,在一些情况下上至12mer肽。个体MHC I类等位基因对给定范围内的肽长度有个体偏好。MHC II类蛋白通常结合总长度为13-18个氨基酸,包含含有重要氨基酸锚定残基的9'-mer核心基序的肽。然而,与大多数MHC I类分子不同,对总长度没

有严格限定。

[0330] 对于一些MHC等位基因,最佳肽长度和对所谓锚定位置上的特定氨基酸残基的偏好是已知的。

[0331] 为了鉴定源自给定MHC等位基因针对的特定蛋白的高亲和力结合肽,有必要系统地研究蛋白的氨基酸序列来鉴定推定的高亲和力结合肽。虽然给定的肽是结合物,但它不一定是功能性T细胞表位。功能性需要通过功能分析来确认,例如,如本文其他地方描述的ELISPOT、CTL杀伤测定或流式细胞术测定。

[0332] 在一些实施方案中,抗原肽可以通过计算预测(例如使用NetMHC)或通过选择特定的8-mer、9-mer、10-mer、11-mer、12-mer氨基酸序列来产生。对于一些MHC分子,肽的结合亲和力可以在目前可用的数据库中预测。

[0333] 结合肽P的设计

[0334] 设计结合肽P的第一步是获得感兴趣的蛋白或抗原多肽的氨基酸序列。

[0335] 在许多情况下,必须从中鉴定抗原肽的蛋白的氨基酸序列是已知的。然而,当只有基因组DNA序列是已知的时,即基因的阅读框和转录方向是未知的时,DNA序列需要以两个方向翻译所有三个阅读框,导致给定基因组的总共六个氨基酸序列。

[0336] 可以如下描述地从这些氨基酸序列鉴定结合肽。在具有内含子/外显子基因结构的生物体中,必须相应地修改本方法,以鉴定由部分源自两个独立内含子的氨基酸序列组合衍生的肽序列基序。cDNA序列可以被翻译成实际的氨基酸序列以允许肽鉴定。在蛋白序列已知的情况下,这些可以直接用于预测肽表位。

[0337] 结合肽序列可以通过全面方法(产生潜在的任何MHC等位基因的结合肽序列)或定向方法(鉴定具有某些优选特征(诸如对MHC蛋白的亲合力、对MHC蛋白的特异性、通过细胞中的蛋白水解形成的可能性和其他重要特征)的结合肽的子集)从任何蛋白序列预测。

[0338] MHC I类结合肽序列的设计

[0339] 许多参数影响个体结合肽P的设计,以及在特定应用中使用的一组结合肽的选择。

[0340] MHC-肽复合物的重要特征是物理和化学(例如蛋白水解)稳定性。必须考虑这些关于抗原肽P、MHC-肽复合物和MHC多聚体的产生以及它们在给定应用中的使用的参数的相关性。例如,MHC-肽复合物在测定缓冲液(例如PBS)、血液或体内的稳定性对于特定应用可能是非常重要的。

[0341] 在MHC-肽复合物与TCR的相互作用中,必须考虑许多另外的特征,包括对TCR的结合亲和力和特异性、串扰程度、与其他TCR的不期望的结合或相互作用。最后,必须考虑关于MHC-肽复合物、MHC多聚体或抗原肽与其所被应用于的样品或个体的相互作用的许多参数。这些参数包括免疫原性、变应原性,以及与“错误的”T细胞的不期望的相互作用导致的副作用,包括与例如自身免疫性疾病的串扰以及与抗原特异性T细胞以外的其他细胞的不期望的相互作用。

[0342] 对于一些应用,例如集中在一种抗原上的个体免疫应答的免疫谱分析,有利的是该抗原的所有可能的结合肽都包括在应用中(即下面描述的用于结合肽设计的“全面方法”)。对于其他应用,对于应用中包含的每种HLA-等位基因,包含几种或仅一种结合肽可能就足够了(即“定向方法”,由此可以仅包含最有效的结合肽)。个性化诊断、治疗和疫苗通常将介于这两种极端之间,因为在例如靶向给定个体的疫苗中只需要包含几种或仅一种结合

肽,但是特异性结合肽可能必须从由全面方法设计的结合肽中挑选,并通过使用涉及所有可能结合肽的免疫谱分析研究来鉴定。免疫谱分析的原理在本文的其他地方描述。

[0343] a) 全面方法

[0344] MHC I类结合肽P的预测使用全面方法如下进行。实际的蛋白序列被分成8-mer、9-mer、10-mer和11-mer肽序列。这是如下进行的:从氨基酸位置1起始鉴定第一个8-mer;然后将起始位置移动一个氨基酸,鉴定第二个8-mer;然后将起始位置移动一个氨基酸,鉴定第三个8-mer。对于每一轮肽鉴定,该程序通过将起始位置移动一个氨基酸来继续。产生的肽将是氨基酸位置1-8、2-9、3-10等。该过程可以手动或通过软件程序(诸如W0 2009/106073的图2中公开的)的方式进行。然后以相同的方式分别对9-mer、10-mer、11-mer和12-mer重复该程序。

[0345] b) 定向方法

[0346] 定向方法从全面方法中产生的结合肽中鉴定结合肽P的优选子集。这个优选的子集在给定的背景中特别有价值。选择抗原肽(P)子集的一种方法是使用共有序列来选择一组能够结合个体MHC等位基因并将适合“一般(average)”个体的相关结合肽。这样的共有序列通常只考虑结合肽对MHC蛋白的亲合力;换句话说,在设计结合肽具有形成稳定的MHC-肽复合物的高概率,但是不确定该MHC-肽复合物在群体中是否高度相关,并且更不确定该MHC-肽复合物在给定个体中是否高度相关的情况下,鉴定结合肽的子集。

[0347] 对于I类MHC等位基因,结合肽的共有序列通常由式X1-X2-X3-X4-...-Xn给出,其中n等于8、9、10或11,并且其中X代表20种天然存在的氨基酸中的一种,任选地如本申请中其他地方描述地被修饰。X1-Xn可以被进一步定义。因此,共有序列中的某些位置比其他位置更可能有助于与给定的MHC分子结合。

[0348] MHC I的抗原肽结合是通过抗原肽的特定氨基酸侧链与MHC分子的肽结合沟内的离散口袋的相互作用来完成的。肽结合沟由MHC I重链的a1和a2结构域形成,并包含六个称为A、B、C、D、E、F的口袋。对于人类HLA分子,将抗原肽与MHC I关联的主要结合能由抗原肽的位置2和C-末端处的氨基酸与MHC I分子的B和F结合口袋的相互作用提供。抗原肽的负责肽与MHC分子的主要锚定的氨基酸在下文中称为一级锚定氨基酸并且它们形成的基序称为一级锚定基序。抗原肽的其他氨基酸侧链也可能有助于抗原肽与MHC分子的锚定,但程度较低。这样的氨基酸通常被称为二级锚定氨基酸,并形成二级锚定基序。

[0349] 不同的HLA等位基因具有排列在肽结合沟的不同口袋中的不同氨基酸,使得不同的等位基因能够结合具有特定锚定氨基酸基序的独特抗原肽组库。因此,对于选择的共有序列,某些位置是所谓的锚定位置,并且这些位置的有用氨基酸的选择仅限于那些能够适合HLA分子中相应结合口袋的氨基酸。例如,对于结合HLA-A\*02的肽,X2和X9分别是对接(docking)到HLA分子的B和F口袋中的一级锚定位置,并且结合肽中这两个位置处的有用氨基酸优选地限于X2为亮氨酸或甲硫氨酸,以及X9位置处为缬氨酸或亮氨酸。不同的是,结合HLA-B\*08的肽的一级锚定位置是X3、X5和X9,并且这些位置处相应的优选氨基酸是X3位置处为赖氨酸,X5位置处为赖氨酸或精氨酸,以及X9位置处为亮氨酸。

[0350] 不同的HLA等位基因可以分组成簇或超型(supertyping),其中超型的等位基因共有肽结合口袋相似性,即它们能够识别同一类型的抗原肽一级锚定基序。因此,抗原肽可以根据其结合给定HLA分子或给定HLA超型的能力,基于其氨基酸序列,例如一级锚定基序的身

份来选择。

[0351] 表1列出了特别令人感兴趣的抗原肽一级锚定基序,其中结合在口袋B和口袋F中的有用氨基酸的实例示出为单字母代码。

[0352] 表1:HLA-I超型家族及其抗原肽锚定基序

[0353]

| 超型      | 锚定基序    |                        |           |                  | HLA 等位基因的实例  |
|---------|---------|------------------------|-----------|------------------|--|
|         | B 口袋特异性 | 示例 aa<br>B 口袋          | F 口袋特异性   | 示例 aa<br>F 口袋    |  |
| A01     | 小的和脂肪族的 | A, T, S, V, L, I, M, Q | 芳族和大的疏水性的 | F, W, Y, L, I, M | A*0101, A*2601, A*2602, A*2603, A*3002, A*3003, A*3004, A*3201 |
| A01/A03 | 小的和脂肪族的 | A, T, S, V, L, I, M, Q | 芳族和碱性的    | Y, R, K          | A*3001, A*3201, A*7401   |

|         |            |                                 |                 |                                    |  |
|---------|------------|---------------------------------|-----------------|------------------------------------|--|
| A01/A24 | 小的、脂肪族和芳族的 | A, S, T, V, L, I, M, Q, F, W, Y | 芳族和大的疏水性的       | F, W, Y, L, I, M                   | A*2902   |
| A02     | 小的和脂肪族的    | A, T, S, V, L, I, M, Q          | 脂肪族和小的疏水性的      | L, I, V, M, Q, A                   | A*0201, A*0202, A*0203, A*0204, A*0205, A*0206, A*0207, A*0214, A*0217, A*6802, A*6901   |
| A03     | 小的和脂肪族的    | A, T, S, V, L, I, M, Q          | 碱性的             | R, H, K                            | A*0301, A*1101, A*3101, A*3301, A*3303, A*6601, A*6801, A*7401   |
| A24     | 芳族和脂肪族的    | F, W, Y, L, I, V, M, Q          | 芳族、脂肪族和疏水性的     | F, W, Y, L, I, V, M, Q, A          | A*2301, A*2402   |
| B07     | 脯氨酸        | P                               | 芳族、脂肪族和疏水性的     | F, W, Y, L, I, V, M, Q, A          | B*0702, B*0703, B*0705, B*1508, B*3501, B*3503, B*4201, B*5101, B*5102, B*5103, B*5301, B*5401, B*5501, B*5502, B*5601, B*6701, B*7801         |
| B08     | 未定义        |                                 | 芳族、脂肪族和疏水性的     | F, W, Y, L, I, V, M, Q, A          | B*0801, B*0802   |
| B27     | 碱性的        | R, H, K                         | 芳族、脂肪族、碱性和疏水性性的 | F, W, Y, L, I, V, M, Q, A, R, H, K | B*1402, B*1503, B*1509, B*1510, B*1518, B*2702, B*2703, B*2704, B*2705, B*2706, B*2707, B*2709, B*3801, B*3901, B*3902, B*3909, B*4801, B*7301 |
| B44     | 酸性的        | D, E                            | 芳族、脂肪族和疏水性的     | F, W, Y, L, I, V, M, Q, A          | B*1801, B*3701, B*4001, B*4002, B*4006, B*4402, B*4403, B*4501   |
| B58     | 小的         | A, S, T                         | 芳族、脂肪族和疏水性的     | F, W, Y, L, I, V, M, Q, A          | B*1516, B*1517, B*5701, B*5702, B*5801, B*5802   |
| B62     | 脂肪族的       | L, I, V, M, Q                   | 芳族、脂肪族和疏水性的     | F, W, Y, L, I, V, M, Q, A          | B*1501, B*1502, B*1512, B*1513, B*4501, B*4601, B*5201   |

[0354]

[0355] 能够结合给定MHC分子的抗原肽P不是必须具有与MHC分子的两个主要锚定口袋都相容的一级锚定氨基酸残基,而是可以具有一个或不具有适于结合所讨论的MHC分子的一级锚定氨基酸。然而,具有针对给定MHC等位基因的优选一级锚定基序增加了抗原肽对该给

定等位基因的亲和力,并从而增加了制备稳定和有用的MHC-肽分子的可能性。

[0356] 因此,在一些实施方案中,可以基于抗原肽在一级锚定位置和/或二级锚定位置处具有的氨基酸,根据抗原肽结合给定HLA或其他MHC分子的能力来鉴定和选择抗原肽。

[0357] 可用的软件程序使用神经网络或已建立的结合偏好来预测特定结合肽与特定MHC I类等位基因的相互作用。用于预测和选择有用抗原肽的另一个有用参数是细胞内蛋白水解机制在体内产生所讨论的结合肽的概率。例如,对于给定的抗原,可以考虑内体(endosomal)蛋白酶、胞质蛋白酶和膜结合蛋白酶活性的组合作用以及TAP1和TAP2转运蛋白特异性。然而,蛋白水解活性在个体间差异很大,并且对于个性化诊断、治疗或疫苗接种,忽略这些一般的蛋白水解数据可能是合意的。

[0358] 在一些实施方案中,抗原肽P的鉴定包括预测肽P与一种或更多种MHC I类等位基因的理论结合亲和力及结合亲和力阈值(nM)(或简称为“亲和力阈值”)。在一些实施方案中,抗原肽P的鉴定包括预测肽P与MHC I类分子的理论结合亲和力,其中亲和力阈值为1000nM(结合物)、500nM(弱结合物)或50nM(强结合物)。

[0359] 使用上述原理,可以鉴定能够结合一种或更多种类型的MHC分子并产生稳定的MHC-肽复合物的个体肽或肽的子集。然后可以在功能测定(诸如干扰素 $\gamma$ 释放测定(例如ELISPOT)、细胞毒性测定(例如CTL杀伤测定))中或使用如本文其他地方描述的其他方法测试鉴定的肽的生物学相关性。作为替代或补充,鉴定的抗原肽结合选择的MHC分子的能力可以在结合测定(如Biacore测量、竞争测定或本领域技术人员已知的其他可用于测量肽与MHC分子结合的测定)中确定。

[0360] 具有氨基酸取代的抗原肽P

[0361] 可以修饰感兴趣的抗原肽,例如,抗原肽P可以具有一个或更多个氨基酸取代,诸如1个、2个、3个、4个、5个、6个、7个或8个氨基酸取代。一个或更多个氨基酸取代可以在氨基酸锚定基序内,在氨基酸锚定基序外,或两者都有。在一些实施方案中,一个或更多个氨基酸取代在9-mer核心基序内。在一些实施方案中,一个或更多个氨基酸取代在9-mer核心基序外。

[0362] 在一些实施方案中,这些氨基酸取代包括用“等效氨基酸残基”取代。“等效氨基酸残基”是指能够替代多肽中另一个氨基酸残基而不实质上改变多肽结构和/或功能性的氨基酸残基。因此,等效氨基酸具有相似的特性,诸如侧链的体积(bulkiness)、侧链极性(极性或非极性)、疏水性(疏水性或亲水性)、pH(酸性、中性或碱性)和碳分子的侧链组织(芳族/脂肪族)。因此,“等效氨基酸残基”可以被视为“保守氨基酸取代”。

[0363] 在一些实施方案中,等效氨基酸的分类指以下类别:1)HRK、2)DENQ、3)C、4)STPAG、5)MILV和6)FYW。在一些实施方案中,在如本文应用的术语“等效氨基酸取代”的含义内,下文指示的氨基酸组内的一种氨基酸可以被另一种氨基酸取代:

[0364] 具有极性侧链的氨基酸(Asp、Glu、Lys、Arg、His、Asn、Gln、Ser、Thr、Tyr和Cys)

[0365] 具有非极性侧链的氨基酸(Gly、Ala、Val、Leu、Ile、Phe、Trp、Pro和Met)

[0366] 具有脂肪族侧链的氨基酸(Gly、Ala、Val、Leu、Ile)

[0367] 具有环状侧链的氨基酸(Phe、Tyr、Trp、His、Pro)

[0368] 具有芳族侧链的氨基酸(Phe、Tyr、Trp)

[0369] 具有酸性侧链的氨基酸(Asp、Glu)

[0370] 具有碱性侧链的氨基酸(Lys、Arg、His)

[0371] 具有酰胺侧链的氨基酸(Asn、Gln)

[0372] 具有羟基侧链的氨基酸(Ser、Thr)

[0373] 具有含硫侧链(Cys、Met)的氨基酸

[0374] 中性、弱疏水性氨基酸(Pro、Ala、Gly、Ser、Thr)

[0375] 亲水性、酸性氨基酸(Gln、Asn、Glu、Asp)和

[0376] 疏水性氨基酸(Leu、Ile、Val)

[0377] 维恩图是根据氨基酸的特性对其进行分组的另一种方法(Livingstone&Barton, CABIOS, 9, 745-756, 1993)。在一些实施方案中,一种或更多种氨基酸可以被同一维恩图组中的另一种取代。

[0378] 在一些实施方案中,这些氨基酸取代包括用“非等效氨基酸残基”取代。非等效氨基酸残基是根据上述分组具有与它们所取代的氨基酸的特性不同的特性的氨基酸残基。

[0379] 在一些实施方案中,修饰的抗原肽P包含选自上文表1中包括的HLA基序组的锚定基序;诸如包含根据上文表1的氨基酸位置2和/或9的一级锚定氨基酸残基。在一些实施方案中,修饰的抗原肽P包含用选自自由以下组成的组的氨基酸残基取代位置2的氨基酸残基:i)丙氨酸、苏氨酸、丝氨酸、缬氨酸、亮氨酸、异亮氨酸、甲硫氨酸、谷氨酰胺、苯丙氨酸、色氨酸和酪氨酸;ii)丙氨酸、苏氨酸、丝氨酸、缬氨酸、亮氨酸、异亮氨酸、甲硫氨酸和谷氨酰胺;iii)精氨酸、组氨酸和赖氨酸;iv)天冬氨酸和谷氨酸;或v)丙氨酸、苏氨酸和丝氨酸。

[0380] 在一些实施方案中,抗原肽P或修饰的抗原肽P包含用选自自由以下组成的组的氨基酸残基取代位置9或10的氨基酸残基:i)苯丙氨酸、色氨酸、酪氨酸、亮氨酸、异亮氨酸、缬氨酸、谷氨酰胺、丙氨酸、精氨酸、组氨酸、赖氨酸和甲硫氨酸;ii)苯丙氨酸、色氨酸、酪氨酸、亮氨酸、异亮氨酸、缬氨酸、谷氨酰胺、丙氨酸和甲硫氨酸;iii)亮氨酸、异亮氨酸、缬氨酸、谷氨酰胺、丙氨酸和甲硫氨酸;iv)苯丙氨酸、色氨酸、酪氨酸、亮氨酸、异亮氨酸和甲硫氨酸;v)谷氨酰胺和丙氨酸,以及vi)酪氨酸、精氨酸和赖氨酸。

[0381] 在一些实施方案中,氨基酸取代增加了肽对MHC分子的亲和力,并从而增加了MHC-肽复合物的稳定性。

[0382] 在一些实施方案中,氨基酸取代降低了肽对MHC分子的亲和力,并从而降低了MHC-肽复合物的稳定性。在一些实施方案中,氨基酸取代增加了一种或更多种T细胞受体对含有修饰抗原肽的MHC-肽复合物的总体亲和力。在一些实施方案中,氨基酸取代降低了一种或更多种T细胞受体对含有修饰抗原肽的MHC-肽复合物的总体亲和力。

[0383] 抗原肽P片段

[0384] 在一些实施方案中,一种或更多种抗原肽包含一种或更多种抗原肽的片段或由一种或更多种抗原肽的片段组成,所述片段诸如所述一种或更多种抗原肽P的1个、2个、3个、4个、5个、6个、7个、8个、9个、10个、11个或12个氨基酸组成的片段。

[0385] 其他肽修饰

[0386] 除了通过全面方法和/或定向方法设计的结合肽之外,同源肽和已经在氨基酸侧链或主链中修饰的肽可以用作结合肽。

[0387] 在一些实施方案中,本文公开的抗原肽被一种或更多种类型的翻译后修饰(诸如本文其他地方公开的一种或更多种翻译后修饰)修饰。抗原肽中的一个或多个氨基酸上

可以发生相同或不同类型的翻译后修饰。因此,在一些实施方案中,任一氨基酸可以用相同或不同类型的修饰修饰一次、两次或三次。此外,所述相同和/或不同的修饰可以存在于本文公开的结合肽的1个、2个、3个、4个、5个、6个、7个、8个、9个、10个、11个或12个氨基酸上。

[0388] 同源肽

[0389] 同源MHC肽序列可以来自多于一个强烈同源的等位基因的存在,来自小的插入、缺失、倒位或取代。如果它们与由全面方法衍生的肽充分同源,即与由全面方法衍生的一种或两种结合肽具有大于例如多于90%、多于80%、或多于70%、或多于60%的氨基酸序列同一性,则它们可以是良好的候选物。锚定残基的同一性通常是最重要的。

[0390] MHC结合肽可以是分开的或组合的表位来源,即由源自两种不同肽片段和/或蛋白的肽片段的连接形成。这样的肽可以是DNA水平上的遗传重组的结果,或者是由于蛋白周转期间的蛋白复合物分解期间的肽片段关联而产生。这样的肽也可以是蛋白合成期间错误反应的结果,即由某种混合RNA处理引起的。如果较长的肽的一部分形成环,只留下结合在沟中的肽的末端部分,也可以观察到组合肽表位类型。

[0391] 不常见的、人工的和化学修饰的氨基酸

[0392] 具有不常见氨基酸(诸如硒代半胱氨酸和吡咯赖氨酸)的肽也可以结合在MHC沟中。例如具有异构体D-型的人工氨基酸还可以构成可以结合在MHC分子的结合沟中的异构体D-肽。结合肽也可以包含被化学修饰的氨基酸或与反应性基团连接的氨基酸,反应性基团可以被激活以诱导肽的变化或破坏肽。示例翻译后修饰如下所示。然而,也可以对氨基酸侧链或肽主链进行化学修饰。

[0393] 任何修饰都可以单独或组合地存在于肽的任何位置处,例如位置1、2、3、4、5、6等,上至n。

[0394] 表2:肽的翻译后修饰

| 蛋白一级结构和翻译后修饰   |  |
|----------------|--|
| N-末端           | 乙酰化、甲酰化、焦谷氨酸、甲基化、糖化(Gly)、豆蔻酰化、氨基甲酰化  |
| C-末端           | 酰胺化、糖基磷脂酰肌醇(GPI)、O-甲基化、糖基磷脂酰肌醇化、泛素化、类泛素化(Sumoylation)  |
| 赖氨酸            | 甲基化、乙酰化、酰化、羟基化、泛素化、类泛素化、锁链素形成、ADP-核糖基化、脱氨和氧化成醛   |
| 半胱氨酸           | 二硫键、异戊烯化(prenylation)、棕榈酰化   |
| [0395] 丝氨酸/苏氨酸 | 磷酸化、糖基化  |
| 酪氨酸            | 磷酸化、硫酸化、卟啉环连接、黄素连接 GFP 辅基(prosthetic group)(Thr-Tyr-Gly 序列)形成、赖氨酸酪氨酸醌(LTQ)形成、多巴醌(Topaquinone, TPQ)形成 |
| 天冬酰胺           | 脱酰胺、糖基化  |
| 天冬氨酸           | 琥珀酰亚胺形成  |
| 谷氨酰胺           | 转谷氨酰胺化   |
| 谷氨酸            | 羧化、甲基化、聚谷氨酰化、聚甘氨酰化   |
| 精氨酸            | 瓜氨酸化、甲基化   |
| 脯氨酸            | 羟基化  |

[0396] 翻译后修饰的肽

[0397] 抗原肽P的氨基酸也可以根据所讨论的氨基酸以各种方式修饰,或者修饰可以影

响肽的氨基末端或羧基末端(参见紧接在上面的表2)。这样的肽修饰作为亲本蛋白翻译后加工的结果天然发生。下面给出了主要的翻译后修饰的非穷尽描述,分为下面列出的三种主要类型(a、b、c)。

[0398] a) 将化学部分添加至结合肽P的修饰:

[0399] 乙酰化,添加乙酰基基团,通常在蛋白的N-末端处。烷基化,添加烷基基团(例如甲基、乙基)。

[0400] 甲基化,添加甲基基团,通常在赖氨酸或精氨酸残基处,是烷基化的一种。脱甲基包括去除甲基基团。

[0401] C-末端酰胺化。

[0402] 生物素化,用生物素附加物(appendage)甲酰化对保守赖氨酸残基的酰化。

[0403] 依赖于维生素K的 $\gamma$ -羧化。

[0404] 谷氨酰化,谷氨酸残基通过微管蛋白多聚谷氨酰化酶与微管蛋白和一些其他蛋白共价连接。糖基化(glycosylation),将糖基基团添加至天冬酰胺、羟赖氨酸、丝氨酸或苏氨酸,产生糖蛋白。不同于糖化(glycation),糖化被认为是糖的非酶促衔接。

[0405] 甘氨酸化,将一个到多于40个甘氨酸残基共价连接至微管蛋白C-末端尾。

[0406] 可以共价衔接血红素部分。

[0407] 羟基化,是将一个或多个羟基基团(-OH)引入到化合物(或基团(radical))中从而氧化它的任何化学过程。待羟基化的主要残基是脯氨酸。羟基化发生在C<sup>Y</sup>原子,形成羟脯氨酸(Hyp)。在某一情况下,脯氨酸可以替代地在其C<sup>P</sup>原子上被羟基化。赖氨酸也可以在其C<sup>6</sup>原子上被羟基化,形成羟赖氨酸(Hyl)。碘化。

[0408] 异戊二烯化,添加类异戊二烯基团(例如法尼醇和香叶基香叶醇(geranylgeraniol))。

[0409] 脂酰化,衔接脂酸官能团,如异戊烯化、GPI锚形成、豆蔻酰化、法尼基化、类香叶基化(geranylation)。

[0410] 可以共价衔接核苷酸或其衍生物,如ADP-核糖基化和黄素衔接。

[0411] 氧化,赖氨酸可以被氧化成醛。

[0412] 聚乙二醇化,将聚乙二醇基团添加至蛋白。典型的活性氨基酸包括赖氨酸、半胱氨酸、组氨酸、精氨酸、天冬氨酸、谷氨酸、丝氨酸、苏氨酸、酪氨酸。也可以使用N-末端氨基基团和C-末端羧酸。

[0413] 可以共价衔接磷脂酰肌醇。

[0414] 磷酸泛酰巯基乙胺化(Phosphopantetheinylation),添加来自辅酶A的4'-磷酸泛酰巯基乙胺基部分,如脂肪酸、聚酮、非核糖体肽和亮氨酸生物合成。

[0415] 磷酸化,添加磷酸基团,通常向丝氨酸、酪氨酸、苏氨酸或组氨酸添加。

[0416] 焦谷氨酸形成是N-末端谷氨酰胺自我攻击,导致环状焦谷氨酸基团的形成的结果。

[0417] 通过脯氨酰异构酶的脯氨酸外消旋化。

[0418] tRNA介导的氨基酸添加,诸如精氨酸化。

[0419] 硫酸化,将硫酸基团添加至酪氨酸。

[0420] 硒酰化(硒蛋白中硒的共翻译掺入)。

- [0421] b) 添加蛋白或肽的修饰:
- [0422] ISG化, 与ISG15蛋白(干扰素刺激基因15)共价连接。
- [0423] 类泛素化, 与SUMO蛋白(小泛素相关修饰物)共价连接。
- [0424] 泛素化, 与蛋白泛素共价连接。
- [0425] c) 将一种或更多种氨基酸转化为不同氨基酸的修饰:
- [0426] 瓜氨酸化或脱亚胺化, 将精氨酸转化为瓜氨酸。
- [0427] 脱酰胺, 将谷氨酰胺转化为谷氨酸或将天冬酰胺转化为天冬氨酸。
- [0428] 肽修饰可以作为单个氨基酸或多于一个氨基酸(即单独或组合)的修饰发生。修饰可以存在于肽内的任何位置, 即肽P的整个长度的位置1、2、3、4、5等处。
- [0429] 结合肽的来源
- [0430] A) 来自天然来源
- [0431] 结合肽可以通过天然蛋白或由体外翻译mRNA衍生的蛋白的酶促消化或蛋白水解从天然来源获得。结合肽也可以从MHC结合沟洗脱。
- [0432] B) 来自重组来源
- [0433] 1) 作为单体或多聚体肽
- [0434] 可选地, 肽可以作为单体抗原肽或作为多聚(多联)抗原肽的形式由转染的细胞重组产生。任选地, 在与MHC蛋白结合之前, 多聚体抗原肽被裂解形成单体抗原肽。
- [0435] 2) 作为更大重组蛋白的一部分
- [0436] 结合肽也可以构成更大重组蛋白的一部分, 例如, 由以下组成:
- [0437] 2a) 对于MHC I类结合肽;
- [0438] 肽-接头-b2m, b2hi是全长的或截短的; 肽-接头-MHC I类重链, 重链是全长的或截短的。最重要的是, 截短的I类重链将由细胞外部分, 即a1、a2和a3结构域组成。重链片段也可以仅包含a1和a2结构域, 或单独包含a1结构域, 或与设计结构域或蛋白片段附接的任何片段或全长b2hi或重链。
- [0439] C) 来自化学合成
- [0440] MHC结合肽也可以根据标准方案通过固相或液相合成进行化学合成。
- [0441] 源自一种抗原的抗原肽的综合集合可以通过固相合成方案的修改来制备。
- [0442] 通过在每次氨基酸偶联之后添加部分裂解步骤, 修改了用于在固体支持物上合成全长抗原的方案。因此, 合成的起点是已附接可裂解接头的固体支持物。然后添加第一氨基酸X1(对应于抗原的C-末端)并进行偶联反应。固体支持物现在带有分子“接头-X1”。洗涤后, 一部分(例如10%)可裂解接头现在被裂解, 以将X1释放到溶液中。将上清液转移到收集容器中。添加带有可裂解接头的另外的固体支持物, 例如对应于固体支持物初始量的10%。
- [0443] 然后添加第二氨基酸X2并与X1或可裂解接头偶联, 以在固体支持物上形成分子“接头-X2”和“接头-X1-X2”。洗涤后, 一部分(例如10%)可裂解接头被裂解, 以将X2和X1-X2释放到溶液中。将上清液收集到收集容器中, 因此收集容器现在包含X1、X2和X1-X2。添加带有可裂解接头的另外的固体支持物, 例如对应于固体支持物初始量的10%。
- [0444] 然后添加第三氨基酸X3并与X2或可裂解接头偶联, 以在固体支持物上形成分子“接头-X3”、“接头-X2-X3”和“接头-X1-X2-X3”。洗涤后, 一部分(例如10%)可裂解接头被裂解, 以将X3、X2-X3和X1-X2-X3释放到溶液中。将上清液收集到收集容器中, 因此收集容器现

在包含X1、X2、X3、X1-X2、X2-X3和X1-X2-X3。添加带有可裂解接头的另外的固体支持物,例如对应于固体支持物初始量的10%。这种逐步偶联和接头的部分裂解持续到到达抗原的N-末端。收集容器现在将包含大量不同长度和序列的肽。在采用10%部分裂解的本实例中,大部分肽将是对应于I类抗原肽的8'-mer、9'-mer、10'-mer和10-mer。例如,对于100个氨基酸的抗原,8'-mer将由序列X1-X2-X3-X4-X5-X6-X7-X8、X2-X3-X4-X5-X6-X7-X8-X9、...、X93-X94-X95-X96-X97-X98-X99-X100组成。

[0445] 任选地,在许多偶联和裂解步骤之后或在每个偶联和裂解步骤之后,固体支持物上经使用的(失活的)接头可以再生,以保持高比例的接头可用于合成。抗原肽的集合可用于作用于例如在ELISPOT测定中刺激CTL的APC展示的池,或者抗原肽可以与一种或更多种MHC等位基因混合,以形成大量不同的MHC-肽复合物,这些复合物可以例如用于形成大量不同的MHC多聚体,这些MHC多聚体可以例如在流式细胞术实验中使用。

[0446] 选择用于产生MHC单体和MHC多聚体的MHC等位基因

[0447] 已知人类中有超过600种MHC等位基因(1类和2类);对于其中的许多,肽结合特征是已知的。WO 2009/106073的图3呈现了HLA I类等位基因的列表。不同HLA等位基因的频率差异很大,在不同的种族群体之间也是如此——如针对前30个HLA类等位基因图示的(例如WO 2009/106073中的图4)。因此,仔细选择与希望研究的群体相对应的MHC等位基因是最重要的。

[0448] MHC蛋白可以选自由以下组成的HLA等位基因的组:A\*0201、C\*0701、A\*0101、A\*0301、C\*0702、C\*0401、B\*4402、B\*0702、B\*0801、C\*0501、C\*0304、C\*0602、A\*1101、B\*4001、A\*2402、B\*3501、C\*0303、B\*5101、C\*1203、B\*1501、A\*2902、A\*2601、A\*3201、C\*0802、A\*2501、B\*5701、B\*1402、C\*0202、B\*1801、B 403、C\*0401、C\*0701、C\*0602、A\*0201、A\*2301、C\*0202、A\*0301、C\*0702、B\*5301、B\*0702、C\*1601、B\*1503、B\*5801、A\*6802、C\*1701、B 501、B\*4201、A\*3001、B\*3501、A\*0101、C\*0304、A\*3002、B\*0801、A\*3402、A\*7401、A\*3303、C\*1801、A\*2902、B 403、B\*4901、A\*0201、C\*0401、A\*2402、C\*0702、C\*0701、C\*0304、A\*0301、B\*0702、B\*3501、C\*0602、C\*0501、A\*0101、A\*1101、B\*5101、C\*1601、B 403、C\*0102、A\*2902、C\*0802、B\*1801、A\*3101、B\*5201、B\*1402、C\*0202、C\*1203、A\*2601、A\*6801、B\*0801、A\*3002、B 402、A\*1101、A\*2402、C\*0702、C\*0102、A\*3303、C\*0801、C\*0304、A\*0201、B 001、C\*0401、B\*5801、B 601、B\*5101、C\*0302、B\*3802、A\*0207、B\*1501、A\*0206、C\*0303、B\*1502、A\*0203、B 403、C\*1402、B\*3501、C\*0602、B\*5401、B\*1301、B\*4002、B\*5502和A\*2601。在一些实施方案中,MHC蛋白选自由以下组成的HLA等位基因的组:HLA-A\*A0101、A0201、A0301、A1101、A2402、A2501、A2601、A2902、A3101、A3201、A6801、B0702、B0801、B1503、B1801、B3501、B4002、B4402、B4501和B5101。

[0449] 将肽加载到MHC多聚体中

[0450] 根据肽和MHC的来源以及根据应用,可以以多种方式将肽加载到MHC聚体(MHCmer)MHC I类中。

[0451] 抗原肽可以在不同时间和以不同形式添加到其他肽链,如下

[0452] a) 在MHC复合物折叠期间加载抗原肽:

[0453] a. 以游离肽形式添加抗原肽,MHC I类分子在体外组装期间最常通过折叠反应中的个体组分加载肽,即由纯化的重组重链a与纯化的重组b2微球蛋白和肽或肽混合物组成。

[0454] b. 抗原肽是重组蛋白构建体的一部分

[0455] 可选地,待折叠到结合沟中的肽可以与例如重链或其片段一起由具有重链-柔性接头-肽结构的基因构建体编码。这种重组分子然后在体外与p2微球蛋白一起折叠。

[0456] b) 抗原肽通过交换反应替代另一抗原肽:a. “溶液中的”交换反应

[0457] 期望的肽的加载也可以通过体外交换反应进行,其中已经就位于结合沟中的肽被另一肽种类交换。

[0458] b. “原位”交换反应

[0459] 肽交换反应也可以在亲本分子被附接至其他分子、结构、表面、人工膜或天然膜和纳米颗粒时发生。

[0460] c. 辅助交换反应。

[0461] 该方法可以通过用包含亚稳定氨基酸类似物的肽制备亲本构建体来改善,所述亚稳定氨基酸类似物通过光诱导或化学诱导而分裂,从而使亲本结构自由,使期望的肽可以进入结合沟中。

[0462] d. 通过体内加载展示

[0463] 用期望的肽加载表达于细胞表面的MHC I类分子可以通过交换反应进行。可选地,细胞可以用肽本身或母体蛋白转染,母体蛋白然后被加工,导致体内类似情况,其中在经转染的细胞表达MHC的天然进程期间,肽结合在沟中。对于专职抗原呈递细胞例如树突状细胞、巨噬细胞、Langerhans细胞的情况,蛋白和肽可以通过吞噬作用被细胞自身摄取,并且然后以天然方式结合至MHC复合物,并在细胞表面以正确的MHC背景表达。

[0464] MHC多聚体

[0465] 在一些实施方案中,MHC多聚体在50,000Da和1,000,000Da之间,诸如50,000至980,000Da(例如,50,000至960,000、50,000至940,000、50,000至920,000、50,000至900,000、50,000至880,000、50,000至860,000、50,000至840,000、50,000至820,000、50,000至800,000、50,000至780,000、50,000至760,000、50,000至740,000、50,000至720,000、50,000至700,000、50,000至680,000、50,000至660,000、50,000至640,000、50,000至620,000、50,000至600,000、50,000至580,000、50,000至560,000、50,000至540,000、50,000至520,000、50,000至500,000、50,000至480,000、50,000至460,000、50,000至440,000、50,000至420,000、50,000至400,000、50,000至380,000、50,000至360,000、50,000至340,000、50,000至320,000、50,000至300,000、50,000至280,000、50,000至260,000、50,000至240,000、50,000至220,000、50,000至200,000、50,000至180,000、50,000至160,000、50,000至140,000、50,000至120,000、50,000至100,000、50,000至80,000、50,000至60,000、100,000至980,000、100,000至960,000、100,000至940,000、100,000至920,000、100,000至900,000、100,000至880,000、100,000至860,000、100,000至840,000、100,000至820,000、100,000至800,000、100,000至780,000、100,000至760,000、100,000至740,000、100,000至720,000、100,000至700,000、100,000至680,000、100,000至660,000、100,000至640,000、100,000至620,000、100,000至600,000、100,000至580,000、100,000至560,000、100,000至540,000、100,000至520,000、100,000至500,000、100,000至480,000、100,000至460,000、100,000至440,000、100,000至420,000、100,000至400,000、100,000至380,000、100,000至360,000、100,000至340,000、100,000至320,000、100,000至300,000、100,000至280,000、100,000至

260,000、100,000至240,000、100,000至220,000、100,000至200,000、100,000至180,000、100,000至160,000、100,000至140,000、100,000至120,000、150,000至980,000、150,000至960,000、150,000至940,000、150,000至920,000、150,000至900,000、150,000至880,000、150,000至860,000、150,000至840,000、150,000至820,000、150,000至800,000、150,000至780,000、150,000至760,000、150,000至740,000、150,000至720,000、150,000至700,000、150,000至680,000、150,000至660,000、150,000至640,000、150,000至620,000、150,000至600,000、150,000至580,000、150,000至560,000、150,000至540,000、150,000至520,000、150,000至500,000、150,000至480,000、150,000至460,000、150,000至440,000、150,000至420,000、150,000至400,000、150,000至380,000、150,000至360,000、150,000至340,000、150,000至320,000、150,000至300,000、150,000至280,000、150,000至260,000、150,000至240,000、150,000至220,000、150,000至200,000、150,000至180,000、150,000至160,000Da)。在一些实施方案中,MHC多聚体在1,000,000Da和3,000,000Da之间,例如1,000,000至2,800,000Da、1,000,000至2,600,000Da、1,000,000至2,400,000Da、1,000,000至2,200,000Da、1,000,000至2,000,000Da、1,000,000至1,800,000Da、1,000,000至1,600,000Da、1,000,000至1,400,000Da。

[0466] 以上说明了如何设计和产生MHC多聚体的关键组分,即MHC-肽复合物。在下文中描述如何产生本文公开的MHC单体或MHC多聚体产物。

[0467] 每种多聚体的MHC复合物的数目

[0468] 可能的MHC单体和多聚体的非穷尽列表说明可能性,“n”表示包含在本文公开的多聚体中的MHC复合物的数目:

[0469] a) n=1,单体;

[0470] b) n=2,二聚体,多聚化可以例如基于IgG支架、

[0471] 具有两个MHC的链霉抗生物素蛋白、卷曲螺旋二聚化例如Fos.Jun二聚化;

[0472] c) n=3,三聚体,多聚化可以例如基于具有三个MHC的作为支架的链霉抗生物素蛋白、TNF $\alpha$ -MHC杂合体、三重DNA-MHC缀合物或其他三聚体结构;

[0473] d) n=4,四聚体,多聚化可以例如基于所有四个结合位点都被MHC分子占据的链霉抗生物素蛋白或基于二聚体IgA;e) n=5,五聚体,多聚化例如可以围绕五聚体卷曲-螺旋结构发生;

[0474] f) n=6,六聚体;

[0475] g) n=7,七聚体;

[0476] h) n=8-12,八聚体-十二聚体,多聚化可以例如使用Streptactin;

[0477] i) n=10,十聚体,多聚化可以例如使用IgM;

[0478] j)  $1 < n < 100$ ,可以例如使用Dextramer作为多聚化结构域聚合物,诸如多肽、多糖和右旋糖酐;

[0479] k)  $1 < n < 1000$ ,多聚化可以例如利用树突状细胞(DC)、抗原呈递细胞(APC)、胶束、脂质体、珠、表面例如微量滴定板、管、微阵列装置、微流体系统;

[0480] l)  $1 < n$ ,n为数十亿或数万亿或更高,多聚化可以例如发生在珠和表面例如微量滴定板、管、微阵列装置、微流体系统上。

[0481] 在一些实施方案中,本文公开的组包含MHC多聚体(a-b-P)<sub>n</sub>,其中n>1,包含两种或

更多种MHC蛋白,每种都与抗原肽P复合以形成MHC-肽复合物。在优选实施方案中,MHC蛋白是I类MHC蛋白。

[0482] 在一些实施方案中,本文公开的组包含MHC多聚体(a-b-P)<sub>n</sub>,其中n的值选自以下组成的组:1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、50、55、60、65、70、75、80、85、90、95、100、110、120、130、140、150、160、170、180、190、200、250、300、350、400、450、500、550、600、650、700、750、800、850、900、950和1000。

[0483] 在一些实施方案中,本文公开的组包含MHC多聚体(a-b-P)<sub>n</sub>,其中n的值是1<n≤1000,诸如在2-3、3-4、4-5、5-6、6-7、7-8、8-9、9-10、10-11、11-12、12-13、13-14、14-15、15-16、16-17、17-18、18-19、19-20、20-21、21-22、22-23、23-24、24-25、25-26、26-27、27-28、28-29、29-30、30-35、35-40、40-45、45-50、50-55、55-60、60-65、65-70、75-80、80-85、85-90、90-95、95-100、100-110、110-120、120-130、130-140、140-150、150-160、160-170、170-180、180-190、190-200、200-225、225-250、250-275、275-300、300-325、325-350、350-375、375-400、400-450、450-500、500-550、550-600、600-650、650-700、700-750、750-800、800-850、850-900、900-950、950-1000之间。

[0484] 在一些实施方案中,本文公开的组包含MHC多聚体(a-b-P)<sub>n</sub>,其中n的值为>1、诸如2、诸如>2、诸如≥2、诸如3、诸如>3、诸如≥3、诸如4、诸如>4、诸如≥4、诸如5、诸如>5、诸如≥5、诸如6、诸如>6、诸如≥6、诸如7、诸如>7、诸如≥7、诸如8、诸如>8、诸如≥8、诸如9、诸如>9、诸如≥9、诸如10、诸如>10、诸如≥10。

[0485] 因此,MHC多聚体包括MHC-二聚体、MHC-三聚体、MHC-四聚体、MHC-五聚体、MHC-六聚体和MHC n-聚体,以及包含两个或更多个MHC-肽复合物的有机分子、细胞、膜、聚合物和颗粒。示例的基于有机分子的多聚体包括官能化的环状结构,诸如苯环,其中例如苯环被官能化并共价连接至例如三个MHC复合物;示例的基于细胞的MHC多聚体包括树突状细胞和抗原呈递细胞(APC);示例的基于膜的MHC多聚体包括在其膜中带有MHC-肽复合物的脂质体和胶束;示例的基于聚合物的MHC多聚体包括MHC-Dextramer(许多MHC-肽复合物共价或非共价连接至其的右旋糖酐),并且示例颗粒包括表面上固定有MHC复合物的珠或其他固体支持物。显然,可以使用任何种类的多聚化结构域,包括任何种类的细胞、聚合物、蛋白或其他分子结构,或者颗粒和固体支持物。

[0486] MHC复合物的三种组分中的任何一种都可以是下述任何一种来源。该列表是非穷尽性的。完整列表将涵盖所有脊索动物物种。来源意指该序列与特定物种的天然存在的序列相同或高度同源。

[0487] 来源列表:人类、小鼠、灵长类动物(包括黑猩猩、大猩猩、猩猩)、猴(包括猕猴)、猪(Porcine (Swine/Pig))、牛科动物(牛/羚羊)、马科动物(马)、驼类动物(骆驼)、反刍动物(鹿)、犬科动物(狗)、猫科动物(猫)、鸟(包括鸡、火鸡)、鱼、爬行动物和两栖动物。

[0488] 在一些实施方案中,本文公开的MHC是HLA-A型的MHC I类复合物。在一些实施方案中,MHC是HLA-B型的MHC I类复合物。在一些实施方案中,MHC是HLA-C型的MHC I类复合物。

[0489] MHC多聚体的产生

[0490] 产生各种类型的MHC多聚体的不同方法在例如美国专利第5,635,363号、W002/072631和W099/42597、US20040209295中描述,并且在本文的其他地方描述。简言之,MHC多

聚体可以如下产生：首先表达和纯化MHC蛋白的个体蛋白组分，并且然后将MHC蛋白组分和肽组合来形成MHC-肽复合物。然后适当数目的MHC-肽复合物通过共价键或非共价键与多聚化结构域连接在一起。这可以通过多聚化结构域的反应性基团（例如右旋糖酐聚合物上的乙烯基磺官能团）和MHC蛋白上的反应性基团（例如蛋白表面的氨基基团）之间的化学反应，或者通过MHC蛋白的一部分（例如生物素化的肽组分）和多聚化结构域（例如链霉抗生物素蛋白四聚体蛋白上针对生物素的四个结合位点）之间的非共价相互作用来实现。作为替代，MHC多聚体可以通过与MHC蛋白的一种组分融合的氨基酸螺旋的非共价关联形成，以形成五聚体MHC多聚体，由构成多聚化结构域的卷曲螺旋结构中的五个螺旋保持在一起。

[0491] 用于MHC和多聚化结构域的共价偶联的适当化学反应包括通过亲电试剂活化的亲核取代（例如，酰化诸如酰胺形成、吡唑啉酮（pyrazolone）形成、异噁唑酮形成；烷基化；乙烯基化；二硫化物形成），碳-杂多键的加成（例如通过膦酸与醛或酮的反应的烯炔形成；芳基化；通过与烷基硼酸酯或烯醇醚反应使芳烃/杂芳烃烷基化），使用亲核试剂活化的亲核取代（例如缩合；脂肪族卤化物或甲苯磺酸用烯醇醚或烯胺的烷基化），以及环加成。

[0492] 能够在多聚化结构域和MHC-肽复合物之间提供非共价相互作用的适当分子包括以下分子对和分子：链霉抗生物素蛋白/生物素、抗生物素蛋白/生物素、抗体/抗原、DNA/DNA、DNA/RNA、PNA/PNA、LNA/DNA、亮氨酸拉链例如Fos/Jun、IgG二聚体蛋白、IgM多价蛋白、酸性/碱性卷曲螺旋、螯合物/金属离子结合螯合物、链霉抗生物素蛋白（SA）和抗生物素蛋白及其衍生物、生物素、免疫球蛋白、抗体（单克隆、多克隆和重组的）、抗体片段及其衍生物、AP-1的亮氨酸拉链结构域（jun和fos）、hexa-his（金属螯合部分）、hexa-hat GST（谷胱甘肽S-转移酶）谷胱甘肽亲和力、

[0493] 钙调蛋白结合肽（CBP）、Strep-标签、纤维素结合结构域、麦芽糖结合蛋白、S-肽标签、壳多糖结合标签、免疫反应性表位、表位标签、E2标签、HA表位标签、Myc表位、FLAG表位、AU1和AU5表位、Glu-Glu表位、KT3表位、IRS表位、Btag表位、蛋白激酶-C表位、VSV表位、介导与包括糖类、脂质和蛋白的多种化合物结合的凝集素，例如Con A（直生刀豆）或WGA（麦胚凝集蛋白）和四联凝集素或蛋白A或蛋白G（抗体亲和力）。还包括这样的结合实体的组合。特别地，当MHC复合物被加标签时，结合实体可以是“抗标签（anti-tag）”。“抗标签”意指与标签结合的抗体和能够与这样的标签结合的任何其他分子。

[0494] MHC蛋白链的产生

[0495] MHC I类重链（HC）和 $\beta 2m$ 微球蛋白（ $\beta 2m$ ）可以从各种来源获得。

[0496] a) 通过从天然表达所讨论的MHC I类或 $\beta 2m$ 分子的真核细胞中纯化的天然来源。

[0497] b) 所述分子可以通过重组方法获得，例如使用

[0498] a. 从天然表达所讨论的MHC或 $\beta 2m$ 分子的细胞获得的mRNA的体外翻译；

[0499] b. 通过经HC和/或 $\beta 2m$ 基因转染的哺乳动物、酵母、细菌或其他来源的细胞的表达和纯化。这后一种方法通常是选择的方法。用于转染/转化的遗传物质可以是：

[0500] i. 从细胞、组织或生物体中分离的天然来源的

[0501] ii. 合成来源的，即与天然DNA序列相同的合成基因，或者可以对其进行修饰以引入分子变化或容易重组表达。

[0502] 遗传物质可以编码 $\beta 2m$ 的全部或仅编码 $\beta 2m$ 的片段，MHC I类重链的全部或仅MHC I类重链的片段。特别令人感兴趣的是由以下组成的MHC I类重链片段：完整的链减去膜内结

构域,仅由细胞外a1和a2 I类重链结构域组成的链,或任何提及的包含修饰或添加的设计结构域或序列的β2m和重链片段。

#### [0503] 修饰的MHC I复合物

[0504] 以上述任何方式修饰的MHC I复合物可以结合TCR。

[0505] 修饰包括天然或非天然氨基酸或任何其他有机分子的突变(取代、缺失或插入)。突变不限于增加MHC复合体稳定性的突变,并且可以引入MHC复合物中的任何位置。一种特别令人感兴趣的实例是引入MHC I重链a3亚基中的突变。a3亚基与T细胞表面的CD8分子相互作用。为了使MHC多聚体与非特异性T细胞表面的CD8分子的结合最小化,可以使参与和CD8相互作用的a3结构域中的氨基酸突变。这样的突变能够导致MHC与CD8分子的结合被改变或消除。另一种特别令人感兴趣的实例是在负责结合CD4分子的MHC II分子的p2结构域区域中的突变。

[0506] 另一种实施方案是化学修饰的MHC复合物,其中化学修饰可以被引入复合物中的任何位置,例如其中肽结合槽中的肽具有附接的二硝基苯基基团的MHC复合物。

[0507] 修饰的MHC复合物也可以是MHC I或MHC II融合蛋白,其中融合蛋白不必比天然蛋白更稳定。特别令人感兴趣的是与编码能够被Bir A酶生物素化的氨基酸序列的基因融合的MHC复合物(Schatz,P.J.(1993),*Biotechnology* 11(10):1138-1143)。该生物素化序列可以与MHC I分子的p2m或重链的COOH末端或MHC II的a链或b链的COOH末端融合。类似地,能够被酶促或化学修饰的其他序列可以融合到MHC复合物的NH2末端或COOH末端。

#### [0508] 对空MHC复合物和MHC-肽复合物进行稳定

[0509] 自然界中经典的MHC复合物嵌入在膜中。一些实施方案包括包含可溶性形式的MHC的多聚体,其中膜锚定的MHC复合物的跨膜结构域和胞质结构域被去除。去除分子的膜锚定部分可能影响MHC复合物的稳定性。在产生和使用MHC多聚体时,MHC复合物的稳定性是重要的参数。

[0510] MHC I复合物具有单个膜锚定的重链,该重链包含完整的肽结合沟,并且当与p2m复合时以可溶性形式稳定。长期稳定性依赖于在肽结合沟中结合肽。在肽结合沟中没有肽的情况下,重链和p2m倾向于解离。类似地,对肽结合沟中的结合具有高亲和力的肽通常将使MHC复合物的可溶性形式稳定,而对肽结合沟具有低亲和力的肽通常将具有较小的稳定作用。

[0511] 在自然界中,MHC I分子由与p2m组合的重链和通常8-11个氨基酸的肽组成。在本文中,MHC I分子还包括由以下组成的分子:重链和p2m组成的分子(空MHC);或与肽组合的重链、或与肽组合的包含a1和a2亚基的截短的重链;或与全长或截短的p2m链组合的全长或截短的重链。这些MHC I分子可以在大肠杆菌(*E. coli*)中以重组蛋白产生,纯化和在体外重折叠(Garboczi等人,(1992),*Proc.Natl.Acad.Sci.* 89,3429-33)。可选地,可以使用昆虫细胞系统或哺乳动物细胞系统。为了产生稳定的MHC I复合物,并从而产生可靠的MHC I多聚体,可以遵循几种策略。针对MHC I复合物的稳定策略在下文中描述。

#### [0512] 针对MHCI复合物的稳定策略

[0513] 产生共价蛋白融合体

[0514] MHC I分子可以通过在MHC I复合物的个体组分之间引入一个或多个接头来稳定。这可以由以下组成的复合物:通过接头与p2m融合的重链和可溶性肽;通过接头与p2m

融合的重链;通过接头与肽共价连接的重链/p2m二聚体(肽连接至重链或p2m),并且其中在重链和p2m之间可以有接头或可以没有接头;通过接头与肽融合的重链;或者通过接头与肽融合的重链的a1和a2亚基。在所有这些示例蛋白融合中,每个重链、p2m和肽都可以被截短。

[0515] 接头可以是柔性接头,例如由甘氨酸和丝氨酸构成,并且是例如5-20个残基之间的长度。接头也可以是刚性的,具有确定的结构,例如由氨基酸如谷氨酸、丙氨酸、赖氨酸和亮氨酸构成,产生例如更刚性的结构。

[0516] 在重链-p2m融合蛋白中,p2m的COOH末端可以与重链的NH2末端共价连接,或者p2m的NH2末端可以与重链的COOH末端连接。融合蛋白还可以包含插入到重链中的p2m结构域或截短的p2m结构域,以形成“重链(第一部分)-p2m-重链(后一部分)”形式的融合蛋白。同样,融合蛋白可以包含插入到p2m链中的重链结构域或截短的重链,以形成“P2m(第一部分)-重链-p2m(后一部分)”形式的融合蛋白。

[0517] 在肽-p2m融合蛋白中,肽的COOH末端优选地与p2m的NH2末端连接,但肽也可以通过其NH2末端与p2m的COOH末端连接。在重链-肽融合蛋白中,优选地将重链的NH2末端与肽的COOH末端融合,但是融合也可以在重链的COOH末端和肽的NH2末端之间。在重链-p2m肽融合蛋白中,重链的NH2末端可以与p2m的COOH末端融合,并且p2m的NH2末端可以与肽的COOH末端融合。

[0518] 通过与非天然组分结合进行非共价稳定

[0519] 非天然组分与MHC I复合物的非共价结合可以导致稳定性增加。非天然组分可以与重链和p2m两者结合,并以这种方式促进复合物的组装,和/或稳定所形成的复合物。可选地,非天然组分可以与p2m或重链结合,并以这种方式稳定多肽的正确构象,并以这种方式增加重链对p2m和/或肽的亲合力,或增加p2m对肽的亲合力。

[0520] 这里,非天然组分意指抗体、肽、适配体或任何其他具有结合MHC复合物肽链段(stretches)能力的分子。抗体在这里应被理解为截短或全长(同种型IgG、IgM、IgA、IgE的)抗体、Fab、scFv或双Fab片段或双体。

[0521] 特别令人感兴趣的实例是通过与重链以及p2m相互作用结合MHC I分子的抗体。抗体可以是双特异性抗体,其一臂与MHC复合物的重链结合,并且另一臂与MHC复合物的p2m结合。可选地,抗体可以是单特异性的,并结合在重链和p2m之间的连接处。

[0522] 另一特别令人感兴趣的实例是仅当重链正确折叠时结合重链的抗体。这里正确折叠是这样的构象,其中MHC复合物能够以受限的T细胞能够识别MHC-肽复合物并被活化的方式结合和呈递肽。这种类型的抗体可以是与克隆W6/32(来自Dako,Denmark的M0736)产生的抗体类似的抗体,其识别人类和一些猴的含有p2m、重链和肽的完整MHC复合物的构象表位。

[0523] 产生修饰的蛋白或蛋白组分

[0524] 改进MHC I复合物稳定性的一种方法是增加结合肽对MHC复合物的亲合力。这可以通过肽中相关位置处的氨基酸突变/取代,通过化学修饰肽中相关位置处的氨基酸或通过合成肽中相关位置处的非天然氨基酸来实现。可选地,突变、化学修饰、天然或非天然氨基酸的插入、或缺失可以被引入肽结合槽,即在容纳负责将肽锚定至肽结合槽的肽侧链的结合口袋中。此外,可以将反应性基团引入抗原肽中;在肽结合之前、期间或之后,反应性基团可以与肽结合槽的氨基酸残基反应,从而将肽共价连接至结合口袋。

[0525] 突变/取代、化学修饰、天然或非天然氨基酸的插入、或缺失也可以引入重链和/或

p2m的肽结合槽外的位置处。通过实例,已经表明在人类b2Gh的位置nn处用YY取代XX增强了MHC I类分子复合物的生化稳定性,并因此可以导致亚优势肽表位的更有效的抗原呈递。

[0526] 一些实施方案包括通过突变、化学修饰、氨基酸交换或缺失去除重链中的“不想要的半胱氨酸残基”。

[0527] “不想要的半胱氨酸残基”在这里应被理解为不参与最终MHC I分子正确折叠的半胱氨酸。不直接参与正确折叠的MHC I分子的形成的半胱氨酸的存在可以导致分子内二硫桥的形成,从而在体外重折叠期间导致不正确折叠的MHC复合物。对MHC I复合物进行共价稳定的另一种方法是在MHC复合物的两个亚基之间共价附接接头。这可以是肽和重链之间或重链和β2微球蛋白之间的接头。

[0528] 对MHCI复合物进行其他稳定

[0529] 用可溶性添加剂进行稳定

[0530] 蛋白在水性溶液中的稳定性取决于溶液的组成。盐、去污剂、有机溶剂、聚合物等的添加可以影响稳定性。可以添加盐、去污剂、有机溶剂、聚合物和任何其他可溶性添加剂来增加MHC复合物的稳定性。特别令人感兴趣的是增加MHC分子表面张力而不结合MHC分子的添加剂。实例是蔗糖、甘露糖、甘氨酸、甜菜碱、丙氨酸、谷氨酰胺、谷氨酸和硫酸铵。甘油、甘露醇和山梨糖醇也包括在该组中,尽管它们能够结合极性区域。

[0531] 另一组特别令人感兴趣的添加剂能够增加MHC分子的表面张力,并同时与蛋白中的带电基团相互作用。实例是MgSO<sub>4</sub>、NaCl、聚乙二醇、2-甲基-2,4-戊二醇和硫酸胍。

[0532] MHC复合物的正确形成依赖于肽在肽结合槽中的结合;结合的肽表现为使复合物稳定在其正确的构象。添加摩尔过量的肽将迫使平衡趋向正确折叠的MHC-肽复合物。同样,过量的p2m也被预期驱动折叠过程趋向正确折叠的MHC I复合物的方向。因此,与肽结合槽中结合的肽相同的肽和/或p2m被包括在内作为稳定可溶性添加剂。

[0533] 用于稳定MHC II分子的特别令人感兴趣的其他添加剂是BSA、胎儿和小牛血清或血清中具有蛋白稳定作用的个体蛋白组分。

[0534] 所有上述可溶性添加剂都可以被添加至任何含有MHC复合物的溶液中,以增加分子的稳定性。所述添加剂可以在重折叠过程期间添加,所述添加剂可以添加至形成的MHC复合物、添加至可溶性MHC单体、添加至包含一个或多个MHC复合物的MHC多聚体溶液或添加至在用MHC多聚体分析MHC特异性T细胞期间使用的溶液。

[0535] 用于稳定MHC分子的特别令人感兴趣的其他添加剂是BSA、胎儿和小牛血清或血清中具有蛋白稳定作用的个体蛋白组分。

[0536] 化学修饰的MHC I复合物

[0537] 有许多氨基酸对化学交联剂具有反应性。在下文中,描述了有利于MHC I复合物的交联或修饰的化学反应。

[0538] 链和肽两者的N-末端的氨基基团以及赖氨酸侧链的氨基基团是亲核的,并且可以用于许多化学反应,包括通过亲电试剂活化的亲核取代(例如,酰化诸如酰胺形成、吡啶酮形成、异噁唑酮形成;烷基化;乙烯基化;二硫化物形成),碳-杂多键的加成(例如通过膦酸盐与醛或酮反应的烯炔形成;芳基化;通过与烷基硼酸酯或烯醇醚反应使芳炔/杂芳炔烷基化),使用亲核试剂活化的亲核取代(例如缩合;脂肪族卤化物或甲苯磺酸与用烯醇醚或烯胺的烷基化),以及环加成。可用于与氨基基团反应的示例试剂是活化的羧酸,诸如NHS-

酯、四氟酚酯和五氟酚酯、酸酐、酰氯 (acid chloride) 和酰氟, 以形成稳定的酰胺键。同样, 磺酰基氯可以与这些氨基基团反应形成稳定的磺酰胺。异氰酸酯也可以与氨基基团反应形成稳定的脲, 并且异硫氰酸酯可以用来引入硫脲连接。

[0539] 醛类, 诸如甲醛和戊二醛将与氨基基团反应形成席夫碱, 然后可以进一步还原成仲胺。

[0540] 精氨酸侧链上的胍基基团将与相同类型的试剂进行类似的反应。另一种非常有用的氨基酸是半胱氨酸。侧链上的硫醇容易被马来酰亚胺、乙烯基砜和卤化物烷基化形成稳定的硫醚, 并且与其他硫醇反应将产生二硫化物。

[0541] 链和肽两者的C-末端以及谷氨酸和天冬氨酸侧链上的羧酸也可用于引入交联。它们需要用诸如碳二亚胺的试剂活化, 并且然后可以与氨基基团反应生成稳定的酰胺。

[0542] 因此, 可以采用大量的化学物质来形成共价交联。关键点在于化学试剂是双官能的, 能够与两个氨基酸残基反应。

[0543] 它们可以是同双官能的, 具有两个相同的反应性部分, 诸如戊二醛; 或者可以是异双官能的, 具有两个不同的反应性部分, 诸如GMBS (马来酰亚胺丁酰氧基-琥珀酰亚胺酯)。

[0544] 可选地, 可以使用两种或更多种试剂; 即GMBS可用于在a链上引入马来酰亚胺, 并且亚氨基硫烷可用于在b链上引入硫醇; 然后马来酰胺和硫醇可以在两条链之间形成硫醚连接。

[0545] 在本文公开的方法和组合物的一些实施方案中, 一些类型的交联是特别有用的。折叠的MHC复合物可以与具有大量 (多达数百个) 乙烯基砜的右旋糖酐反应。它们可以与a链和b链两者上的赖氨酸残基以及与从结合位点突出的肽上的赖氨酸残基反应, 有效地交联整个MHC复合物。这样的交联确实是有利的反应, 因为当第一个赖氨酸残基与右旋糖酐反应时, MHC复合物变为锚定至右旋糖酐, 有利于MHC复合物和右旋糖酐多聚化结构域之间的进一步反应。这种右旋糖酐化学的另一个很大的优点是它可以与荧光染料标记组合; 即右旋糖酐与一个或若干个MHC复合物和一个或多个荧光蛋白 (诸如APC) 反应。

[0546] 另一种有价值的方法是将上述分子生物学工具与化学交联剂组合。例如, 一个或多个赖氨酸残基可以插入到a链中, 与b链中的谷氨酸并置, 在引入氨基基团和羧酸之后, 通过添加碳二亚胺进行反应。这样的反应在水中通常不是很有效, 除非在这种情况下, 基团很好地定向于反应。这意味着避免了过度反应, 否则所述过度反应可以导致MHC复合物的变性或改变构象。

[0547] 同样, 右旋糖酐多聚化结构域可以与适当修饰的MHC复合物交联; 即使可以使MHC复合物的一条或两条链富集赖氨酸残基, 增加对乙烯基砜右旋糖酐的反应性。赖氨酸可以在与肽结合槽相对的位置处插入, 使MHC复合物有利于T细胞识别。

[0548] 另一种有价值的化学工具是使用延长的和柔性的交联剂。延长的接头将允许两条链相互作用, 而连接它们的接头产生的应变很少或没有, 同时在复合物解离时保持链彼此靠近。过量的肽还将有利于解离的MHC复合物的重新形成。

[0549] 多聚化结构域

[0550] 许多MHC复合物与多聚化结构域关联形成MHC多聚体。多聚化结构域的尺寸跨越宽范围, 从基于有机小分子支架的多聚化结构域到基于细胞结构或固体支持物的大多聚体。因此, 多聚化结构域可以基于不同类型的载体或支架, 并且同样, MHC复合物与多聚化结构

域的附接可以包括共价或非共价接头。不同种类的多聚化结构域的特征在下文描述。

[0551] 多聚化结构域的分子量

[0552] 在一些实施方案中,多聚化结构域小于1,000Da(小分子支架)。实例包括短肽(例如包含10个氨基酸)和各种小分子支架(例如芳环结构)。

[0553] 在一些实施方案中,多聚化结构域在1,000Da和10,000Da之间(小分子支架、小肽、小聚合物)。实例包括脂肪族化合物和芳族化合物两者的多环结构、包含例如10-100个氨基酸的肽和其他聚合物诸如右旋糖酐、聚乙二醇和聚脲。在一些实施方案中,多聚化结构域在10,000Da和100,000Da之间(小分子支架、聚合物,例如右旋糖酐、链霉抗生物素蛋白、IgG、五聚体结构)。实例包括蛋白和大的多肽、小分子支架诸如类固醇、右旋糖酐、二聚体链霉抗生物素蛋白和诸如在五聚体中使用的多亚基蛋白。在一些实施方案中,多聚化结构域在100,000Da和1,000,000Da之间(小分子支架、聚合物,例如右旋糖酐、链霉抗生物素蛋白、IgG、五聚体结构)。典型的实例包括较大的聚合物,诸如右旋糖酐(例如在Dextramer中使用的)和链霉抗生物素蛋白四聚体。在一些实施方案中,多聚化结构域大于1,000,000Da(小分子支架、聚合物,例如右旋糖酐、链霉抗生物素蛋白、IgG、五聚体结构、细胞、脂质体、人工脂质双层、聚苯乙烯珠和其他珠)。具有该尺寸的大多数实例包括细胞或基于细胞的结构,诸如胶束和脂质体,以及珠和其他固体支持物。

[0554] 如本文所公开的,多聚化结构域可以包含载体分子、支架或两者的组合。

[0555] 多聚化结构域的类型

[0556] 原则上,任何种类的载体或支架都可以用作多聚化结构域,包括任何种类的细胞、聚合物、蛋白或其他分子结构,或颗粒和固体支持物。下面列出了多聚化结构域的不同类型和具体实例。

[0557] 细胞.细胞可以用作载体。细胞可以是活的且有丝分裂活跃的,由于辐射或化学处理而是活的且有丝分裂不活跃的,或者细胞可以是死的。MHC表达可以是天然的(即未受刺激)或可以由例如IFN- $\gamma$ 诱导/刺激的。特别令人感兴趣的是天然抗原呈递细胞(APC),诸如树突状细胞、巨噬细胞、Kupfer细胞、Langerhans细胞、B细胞和任何MHC表达细胞,所述MHC表达细胞是天然表达的、被转染的或是杂交瘤。细胞样结构.细胞样载体包括在其膜中带有MHC-肽复合物的基于膜的结构,诸如胶束、脂质体和其他膜结构,以及噬菌体,诸如丝状噬菌体。

[0558] 固体支持物.固体支持物包括珠、颗粒物质和其他表面。一些实施方案包括带有亲电基团(例如二乙烯基砷活化的多糖)的珠(磁珠或非磁性珠)、已经用甲苯磺酰基活化的酯官能化的聚苯乙烯珠、用甲苯磺酰基活化的酯官能化的磁性聚苯乙烯珠,并且其中MHC复合物可以通过MHC复合物中包含的亲核体与珠的亲电试剂的反应共价固定至这些珠。珠可以由琼脂糖凝胶、聚丙烯酰胺葡聚糖、聚苯乙烯、琼脂糖、多糖、聚氨基甲酸酯制成,或者可以是能够悬浮在水性缓冲液中的任何其他种类的珠。

[0559] 一些实施方案包括表面,即表面上带有固定的MHC复合物的固体支持物和颗粒。特别令人感兴趣的是用于微量滴定板或其他板形式的孔、试剂管、载玻片或用于微阵列分析的其他支持物、微流体室或装置的管道或通道、Biacore芯片和珠。

[0560] 分子.多聚化结构域也可以是通过非共价键保持在一起的分子或分子的复合物。构成多聚化结构域分子可以是有机小分子或大的聚合物,并且可以是柔性的线性分子或

刚性的球状结构,诸如蛋白。用于多聚化结构域的不同种类的分子在下文描述。

[0561] 有机小分子.这里的有机小分子包括类固醇、肽、线性或环状结构、和芳族或脂肪族结构以及许多其他结构。典型的有机小支架是官能化的苯环,即用许多反应性基团(诸如胺)官能化的苯环,许多MHC分子可以共价连接至所述反应性基团。然而,构成连接MHC复合物和多聚化结构域的接头的反应性基团的类型,以及支架结构的类型,可以从化学结构的长列表中选择。下面列出了支架结构的非全面列表。

[0562] 典型支架包括芳族结构、苯二氮草类(benzodiazepines)、乙内酰胺类(hydantoins)、哌嗪类、吡啶类、咪唑类、噻唑类、类固醇、二酮哌嗪类、吗啉类、莨菪烷类(tropanes)、香豆素类、喹啉类、吡咯类、噁唑类、氨基酸前体、环状或芳族环结构以及许多其他结构。典型的载体包括线性的和支链的聚合物,诸如肽、多糖、核酸以及许多其他的聚合物。因此,基于有机小分子或聚合物分子的多聚化结构域包括大量不同的结构,包括致密的小分子,线性结构,聚合物,多肽,聚脲,聚氨基甲酸酯,环状结构,天然化合物衍生物, $\alpha$ -肽, $\beta$ -肽, $\gamma$ -肽,和 $\omega$ -肽、单取代肽、二取代肽和三取代肽、L-型肽和D-型肽、环己烷主链修饰的 $\beta$ -肽和环戊烷主链修饰的 $\beta$ -肽、插烯多肽(vinylogous polypeptides)、糖多肽、聚酰胺、插烯磺酰胺肽、多磺酰胺缀合肽(即具有辅基)、聚酯、多糖诸如右旋糖酐和氨基右旋糖酐、聚氨基甲酸酯、聚碳酸酯、聚脲、聚肽基膦酸酯、氮杂肽(azatide)、类肽(寡聚N-取代的甘氨酸)、聚醚、乙氧基甲缩醛低聚物、聚硫醚、聚乙二醇(PEG)、聚乙烯、聚二硫化物、聚亚芳基硫化物、多核苷酸、PNA、LNA、吗啉代核酸、寡聚吡咯啉酮(oligo pyrrolinone)、聚脲、聚亚胺、聚乙烯亚胺、聚乙酸酯、聚苯乙烯、聚乙炔、聚乙烯、脂质、磷脂、糖脂、多环(脂肪族)、多环(芳族)、多杂环、蛋白聚糖、聚硅氧烷、聚异氰化物(Polyisocyanides)、聚异氰酸酯、聚甲基丙烯酸酯,单官能、双官能、三官能和寡官能(oligofunctional)开链烃,单官能、双官能、三官能和寡官能非芳族碳环、单环、双环、三环和多环烃、桥联多环烃,单官能、双官能、三官能和寡官能非芳族杂环、单环、双环、三环和多环杂环、桥联多环杂环,单官能、双官能、三官能和寡官能芳族碳环、单环、双环、三环和多环芳族碳环,单官能、双环、三环和多环芳族碳环、单官能、双官能、三官能和寡官能芳族杂环、单环、双环、三环和多环杂环,螯合物、富勒烯以及以上结构的任何组合,以及许多其他结构。

[0563] 生物聚合物.这里的生物分子包括肽、蛋白(包括抗体、卷曲螺旋、链霉抗生物素蛋白以及许多其他蛋白)、核酸诸如DNA和RNA、以及多糖诸如右旋糖酐。生物聚合物可以与MHC复合物反应(例如,化学偶联至例如蛋白的氨基基团的许多MHC复合物),或者可以通过例如载体DNA分子和许多各自偶联至MHC复合物的DNA寡核苷酸之间的DNA双链体形成而连接。另一种基于生物聚合物的多聚化结构域是基于链霉抗生物素蛋白的四聚体,其中链霉抗生物素蛋白结合多达四个生物素化的MHC复合物,如以上描述的。

[0564] 自组装多聚体结构.存在几种商业MHC多聚体的实例,其中多聚体通过自组装形成。因此,五聚体(Pentamer)通过形成以表观平面结构将5个MHC复合物保持在一起的卷曲螺旋结构而形成。以类似的方式,Streptamer基于Streptactin蛋白,Streptactin蛋白寡聚形成包含几个MHC复合物的MHC多聚体。

[0565] 在下文中,描述了基于分子多聚化结构域制备MHC多聚体的替代方法。所述方法涉及一种或更多种上述类型的多聚化结构域。

[0566] MHC dextramer可以通过链霉抗生物素蛋白-生物素相互作用将MHC复合物偶联至

右旋糖酐来制备。原则上,生物素-链霉抗生物素蛋白可以被任何二聚化结构域替代,其中二聚化结构域的一半与MHC-肽复合物偶联,并且另一半与右旋糖酐偶联。例如,酸性螺旋(卷曲螺旋二聚体的一半)与MHC偶联或融合,并且碱性螺旋(卷曲螺旋二聚体的另一半)与右旋糖酐偶联。混合两者导致通过形成酸性/碱性卷曲螺旋结构,MHC与右旋糖酐结合。

[0567] 通过利用抗体与精心选择的抗原(该抗原是天然存在的或作为标签被添加到MHC分子的不参与肽结合的部分)结合的能力,抗体可以用作支架。例如,IgG和IgE将能够结合两个MHC分子,具有五聚体结构的IgM将能够结合10个MHC分子。抗体可以是全长的或截短的;标准抗体片段包括Fab2片段。

[0568] 参与卷曲螺旋结构的肽可以通过产生稳定的二聚体、三聚体、四聚体和五聚体相互作用来充当支架。这里的实例是Fos-Jun异二聚体卷曲螺旋,大肠杆菌同源三聚体卷曲螺旋结构域Lpp-56,在水中在生理pH形成离散、稳定的 $\alpha$ -螺旋五聚体的工程化Trp-拉链蛋白。

[0569] 合适的支架、载体和/或接头的另外的实例是链霉抗生物素蛋白(SA)和抗生物素蛋白及其衍生物、生物素、免疫球蛋白、抗体(单克隆、多克隆和重组的)、抗体片段及其衍生物、AP-1的亮氨酸拉链结构域(jun和fos)、hexa-his(金属螯合部分)、hexa-hat GST(谷胱甘肽S-转移酶)、谷胱甘肽、钙调蛋白结合肽(CBP)、Strep-标签、纤维素结合结构域、麦芽糖结合蛋白、S-肽标签、壳多糖结合标签、免疫反应性表位、表位标签、E2标签、HA表位标签、Myc表位、FLAG表位、AU1和AU5表位、Glu-Glu表位、KT3表位、IRS表位、Btag表位、蛋白激酶-C表位、VSV表位、介导与包括糖类、脂质和蛋白的多种化合物结合的凝集素,例如Con A(直生刀豆)或WGA(麦胚凝集蛋白)和四联凝集蛋白或蛋白A或蛋白G(抗体亲和力)。还包括这样的结合实体的组合。非限制性实例是链霉抗生物素蛋白-生物素和jun-fos。特别地,当MHC分子被加标签时,结合实体可以是“抗标签”。“抗标签”意指与标签结合的抗体或能够与这样的标签结合的任何其他分子。

[0570] 除了偶联或结合到多聚化结构域之外,MHC复合物可以通过其他方式进行多聚化。因此,可以在MHC的多聚化期间形成多聚化结构域。一种这样的方法是用二聚化结构域延伸结合的抗原肽。抗原肽的一端延伸有二聚化结构域A(例如酸性螺旋,卷曲螺旋二聚体的一半),并且另一端延伸有二聚化结构域B(例如碱性螺旋,卷曲螺旋二聚体的另一半)。当MHC复合物被加载这些延伸肽/与这些延伸肽混合时,将形成以下多聚体结构:A-MHC-BA-MHC-BA-MHC-B等。混合物中的抗原肽可以是相同的,也可以是具有可比延伸二聚化结构域的肽的混合物。可选地,一种肽的两端延伸有相同的二聚化结构域A,并且另一种肽(相同的氨基酸序列或不同的氨基酸序列)的两端延伸有二聚化结构域B。当MHC和肽混合时,形成以下结构:A-MHC-AB-MHC-BA-MHC-AB-MHC-B等。通过肽延伸的MHC复合物的多聚化限于MHC II分子,因为MHC I分子的肽结合沟通常在两端是封闭的,从而限制了可以嵌入沟中的肽的尺寸,并因此阻止肽延伸出沟。

[0571] 另一种适用于MHC复合物的多聚化方法是基于MHC复合物的N-末端和/或C-末端的延伸。例如,MHC复合物的N-末端延伸有二聚化结构域A,并且C-末端延伸有二聚化结构域B。当MHC复合物被一起孵育时,它们彼此配对并形成如下的多聚体:A-MHC-BA-MHC-BA-MHC-BA-MHC-B等。可选地,一种MHC复合物的N-末端和C-末端两者都延伸有二聚化结构域A,并且另一种MHC复合物制备物(相同或不同的MHC)的N-末端和C-末端延伸有二聚化结构域B。当这两种类型的MHC复合物被一起孵育时,将形成多聚体:A-MHC-AB-MHC-BA-MHC-AB-MHC-B

等。

[0572] 在所有上述实施例中,延伸可以与肽/MHC复合物化学偶联,或者通过基因融合作为延伸引入。

[0573] 二聚化结构域AB可以是能够彼此结合的任何分子对,诸如酸性/碱性卷曲螺旋、抗体-抗原、DNA-DNA、PNA-PNA、DNA-PNA、DNA-RNA、LNA-DNA、亮氨酸拉链例如Fos/Jun、链霉抗生物素蛋白-生物素和本文其他地方描述的其他分子对。

[0574] 接头分子

[0575] 许多MHC复合物与多聚化结构域关联形成MHC多聚体。MHC复合物与多聚化结构域的附接可以包括共价接头或非共价接头,并且可以包括小的反应性基团以及大的蛋白-蛋白相互作用。

[0576] 多聚化结构域和MHC复合物的偶联包括实体X(与多聚化结构域附接或为多聚化结构域的一部分)和实体Y(与MHC复合物附接或为MHC复合物的一部分)的关联。因此,连接多聚化结构域和MHC复合物的接头包括XY部分。

[0577] 共价接头.XY连接可以是共价的,在该情况下,X和Y是反应性基团。在这种情况下,X可以是亲核基团(诸如-NH<sub>2</sub>、-OH、-SH、-NH-NH<sub>2</sub>),并且Y可以是亲电基团(诸如CHO、COOH、CO),它们反应形成共价键XY;或者Y可以是亲核基团,并且X可以是亲电基团,它们反应形成共价键XY。还存在其他可能性,例如,反应性基团中的任一个可以是能够与另一个反应性基团反应的基团。许多反应性基团X和Y,以及X和Y反应后形成的键在WO 2009/106073的图5示出。

[0578] X和Y可以是天然包含在多聚化结构域和/或MHC复合物中的反应性基团,或者X和Y可以是被人工添加的反应性基团。因此,含有反应性基团的接头可以连接至多聚化结构域和MHC复合物中的任一个;随后引入的反应性基团可用于共价连接多聚化结构域和MHC复合物。

[0579] MHC复合物的示例天然反应性基团包括包含-NH<sub>2</sub>、-OH、-SH和-NH-的氨基酸侧链。多聚化结构域的示例天然反应性基团包括多糖(诸如右旋糖酐)的羟基,但当多肽用作多聚化结构域时,还包括多肽的包含-NH<sub>2</sub>、-OH、-SH和-NH-的氨基酸侧链。在一些MHC多聚体中,MHC复合物的一种多肽(即b2M,重链或抗原肽)通过蛋白融合连接至多聚化结构域。因此,在融合蛋白的翻译期间,酰基基团(反应性基团X或Y)和氨基基团(反应性基团Y或X)反应形成酰胺键。多聚化结构域和MHC复合物之间的键是共价的并且由天然反应性基团之间的反应产生的示例MHC多聚体包括MHC-五聚体(在US20040209295中描述)和MHC-二聚体,其中多聚化结构域和MHC复合物之间的连接在这两种情况下均在融合蛋白的翻译期间产生。

[0580] 示例人工反应性基团包括通过包含反应性基团的接头分子的关联而被附接至多聚化结构域或MHC复合物的反应性基团。通过右旋糖酐羟基与二乙烯基砷的反应来活化右旋糖酐,引入了可以与例如MHC复合物的胺反应的反应性乙烯基基团,以形成现在连接多聚化结构域(右旋糖酐聚合物)和MHC复合物的胺。另一种右旋糖酐多聚化结构域的活化包括导致右旋糖酐被马来酰亚胺基团修饰的多步反应,如US 6,387,622中描述的。在这种方法中,MHC复合物的氨基基团被转化为能够与活化的右旋糖酐的马来酰亚胺基团反应的-SH基团。因此,在后一实例中,多聚化结构域的反应性基团(马来酰亚胺)和MHC复合物的反应性基团(硫醇)两者均被人工引入。

[0581] 有时使用活化试剂是为了使反应性基团更具反应性。例如,酸(诸如谷氨酸或天冬氨酸)可以通过添加例如碳二亚胺和NHS或硝基苯酚,或通过把酸部分转化为甲苯磺酰基活化的酯而转化为活化的酯。活化的酯与亲核体诸如-IMH<sub>2</sub>、-SH、-OH等有效地反应。

[0582] 在一些实施方案中,根据多聚化结构域包含的反应性基团的性质,与MHC复合物形成共价键的多聚化结构域(包括有机小支架分子、蛋白、蛋白复合物、聚合物、珠、脂质体、胶束、细胞)可以分成独立的组。一个组包括带有亲核基团(例如-NH<sub>2</sub>、-OH、-SH、-CN、-NH-NH<sub>2</sub>)的多聚化结构域,通过多糖、包含例如赖氨酸、丝氨酸和半胱氨酸的多肽示例;另一组多聚化结构域带有亲电基团(例如-COOH、-CHO、-CO、NHS-酯、甲苯磺酰基活化的酯和其他活化的酯、酸酐),通过例如包含例如谷氨酸和天冬氨酸的多肽或乙烯基砜活化的右旋糖酐示例;还有一组多聚化结构域带有自由基或共轭双键。

[0583] 因此,适用于本公开内容的多聚化结构域包括含有W0 2009/106073的图5中示出的任何反应性基团的多聚化结构域,或者可以与其他反应性基团反应形成W0 2009/106073的图5中示出的键的多聚化结构域。

[0584] 同样,根据MHC复合物内包含的反应性基团的性质,MHC复合物可以分成独立的组。一个组包括带有亲核基团(例如-NH<sub>2</sub>、-OH、-SH、-CN、-NH-NH<sub>2</sub>)例如赖氨酸、丝氨酸和半胱氨酸的MHC;另一组MHC带有亲电基团(例如-COOH、-CHO、-CO、NHS-酯、甲苯磺酰基活化的酯和其他活化酯、酸酐),通过例如谷氨酸和天冬氨酸示例;还有一组MHC带有自由基或共轭双键。

[0585] MHC复合物的反应性基团由MHC-肽复合物的氨基酸携带(并且可以由MHC-肽复合物的任何肽包括抗原肽所包含),或者可选地,MHC复合物的反应性基团通过包含适当反应性基团的分子的共价或非共价衔接引入。

[0586] 在这方面优选的反应性基团包括-CSO<sub>2</sub>OH、苯基氯化物、-SH、-SS、醛、羟基、异氰酸酯、硫醇、胺、酯、硫酯、羧酸、三键、双键、醚、酰氯、磷酸、咪唑、卤化芳环、其任何前体或任何受保护的反应性基团,以及许多其他基团。反应性基团的示例配对和所得的形成的键在W0 2009/106073的图5中示出。

[0587] 可以采用的反应包括酰化(形成酰胺、吡啶啉酮、异噁唑酮、嘧啶、香豆素、喹啉酮(quinolinon)、邻苯二甲酰肼(phthalhydrazide)、二酮哌嗪、苯二氮革酮(benzodiazepinone)和乙内酰脲)、烷基化、乙烯基化、二硫化物形成、Wittig反应、Horner-Wittig-Emmans反应、芳基化(形成双芳基或乙烯基芳烃)、缩合反应、环加成((2+4),(3+2))、碳-碳多键的加成、多键的环加成、碳-杂多键的加成、亲核芳族取代、过渡金属催化的反应,并且可以包括形成醚、硫醚、仲胺、叔胺、β-羟基醚、β-羟基硫醚、β-羟基胺、β-氨基醚、酰胺、硫代酰胺、肟、磺酰胺、二官能化合物和三官能化合物、取代的芳族化合物、乙烯基取代的芳族化合物、炔取代的芳族化合物、双芳基化合物、肼、羟基胺醚、取代的环烯烃、取代的环二烯、取代的1,2,3三唑、取代的环烯烃,β-羟基酮、β-羟基醛、乙烯基酮、乙烯基醛、取代的烯烃、取代的炔烃、取代的胺,以及许多其他的反应。

[0588] MHC dextramer可以通过MHC复合物与右旋糖酐主链的共价偶联来制备,例如通过MHC复合物与右旋糖酐主链的化学偶联。如果MHC复合物是MHC I,则MHC复合物可以通过重链或β<sub>2</sub>微球蛋白偶联,或者如果MHC复合物是MHC II,则可以通过α链或p链偶联。MHC复合物可以作为包含重链/β<sub>2</sub>微球蛋白或α链/p链或任一组合的折叠复合物与肽结合槽中的肽偶

联在一起。可选地,任一蛋白链可以与右旋糖酐偶联,并且然后在体外与MHC复合物的未与右旋糖酐偶联的另一条链一起折叠,并与肽一起折叠。MHC复合物与右旋糖酐多聚化结构域的直接偶联可以通过氨基基团或硫化物基团。任一基团可以是MHC复合物的天然组分,或者化学附接至MHC复合物。可选地,半胱氨酸可以被引入MHC复合物的任一链的基因中。

[0589] 另一种将MHC复合物共价连接到右旋糖酐多聚化结构域的方法是使用抗原肽作为MHC和右旋糖酐之间的接头。将一端包含抗原肽接头与右旋糖酐偶联。这里的抗原肽意指能够结合在MHC复合物的肽结合槽中的肽。例如,10个或更多个抗原肽可以偶联至一个右旋糖酐分子。当MHC复合物被添加至这样的肽-右旋糖酐构建体时,MHC复合物将与抗原肽结合,并因此MHC-肽复合物被展示在右旋糖酐多聚化结构域周围。抗原肽可以彼此相同或不同。类似地,只要MHC复合物能够结合右旋糖酐多聚化结构域上的一个或更多个肽,MHC复合物可以彼此相同或不同。

[0590] 非共价接头.连接多聚化结构域和MHC复合物的接头包括XY部分。以上描述了不同种类的共价连接XY。然而,XY连接也可以是非共价的。

[0591] 非共价XY连接可以包括天然二聚化配对,诸如抗原-抗体对、DNA-DNA相互作用,或者可以包括小分子和蛋白之间的天然相互作用,例如生物素和链霉抗生物素蛋白之间的相互作用。人工XY实例包括XY对,诸如 $\text{HiS}_6$ 标签(X)与Ni-NTA(Y)的相互作用和PNA-PNA相互作用。

[0592] 蛋白-蛋白相互作用.非共价接头可以包括通过非共价相互作用保持在一起的两个或更多个多肽或蛋白的复合物。属于该组的示例多肽和蛋白包括Fos/Jun、酸性/碱性卷曲螺旋结构、抗体/抗原(其中抗原是肽)和许多其他多肽和蛋白。

[0593] 一些实施方案包括多肽和/或蛋白之间的非共价相互作用,由US 20040209295中描述的五聚体结构代表。一些实施方案包括使用对MHC的与肽结合沟相对的表面具有亲和力的抗体。因此,抗MHC抗体将通过它的两个结合位点结合两个MHC复合物,并以这种方式产生二价MHC多聚体。此外,抗体可以通过结合相互作用稳定MHC复合物。这对MHC II类复合物特别有意义,因为这些MHC II类复合物不如I类MHC复合物稳定。

[0594] 多核苷酸-多核苷酸相互作用.非共价接头可以包括非共价相互作用的核苷酸。示例相互作用包括PNA/PNA、DNA/DNA、RNA/RNA、LNA/DNA和任何其他核酸双链体结构,以及这样的天然和非天然多核苷酸的任何组合,诸如DNA/PNA、RNA/DNA和PNA/LNA。

[0595] 蛋白-小分子相互作用.非共价接头可以包括大分子(例如蛋白、多核苷酸)和大分子的小分子配体。相互作用可以是天然的(即,可见于自然界,诸如链霉抗生物素蛋白/生物素相互作用)或非天然的(例如His-标签肽/Ni-NTA相互作用)。示例相互作用包括链霉抗生物素蛋白/生物素和抗生物素抗体/生物素。

[0596] 组合——非共价连接分子.蛋白、多核苷酸、有机小分子和其他分子的其他组合可用于将MHC连接至多聚化结构域。这些其他组合包括蛋白-DNA相互作用(例如DNA结合蛋白,诸如基因调节蛋白CRP与其DNA识别序列相互作用),RNA适配体-蛋白相互作用(例如对生长激素特异性的RNA适配体与生长激素相互作用)。

[0597] 合成分子-合成分子相互作用.非共价接头可以包括通过非共价相互作用保持在一起的两个或更多个有机分子的复合物。示例相互作用是与同一金属离子结合的两个螯合物分子(例如EDTA- $\text{Ni}^{++}$ -NTA),或与NTA- $\text{Ni}^{++}$ 结合的短的聚组氨酸肽(例如, $\text{His}_6$ )。

[0598] 在另一种优选实施方案中,多聚化结构域是珠。珠通过不可裂解的接头或可裂解的接头被共价或非共价地涂覆MHC多聚体或单一MHC复合物。例如,珠可以被涂覆链霉抗生物素蛋白单体,链霉抗生物素蛋白单体继而与生物素化的MHC复合物关联;或者珠可以被涂覆链霉抗生物素蛋白四聚体,每个四聚体与0个、1个、2个、3个或4个生物素化的MHC复合物关联;或者珠可以被涂覆MHC-dextramer,其中例如MHC-dextramer的反应性基团(例如二乙烯基砷活化的右旋糖酐主链)已经与珠上的亲核基团反应,以在dextramer的右旋糖酐和珠之间形成共价连接。

[0599] 在一些实施方案中,上述MHC多聚体(例如,其中多聚化结构域是珠)还包含柔性或刚性且水溶性的接头,接头允许固定的MHC复合物与细胞(诸如对MHC复合物具有亲和力的T细胞)有效地相互作用。在一些实施方案中,接头是可裂解的,允许MHC复合物从珠释放。如果T细胞已经通过与MHC复合物结合被固定,T细胞可以通过裂解这种可裂解的接头非常温和地释放。非限制性可裂解接头在WO 2009/106073的图6中示出。在一些实施方案中,接头有利地在生理条件被裂解,允许分离细胞的完整性。

[0600] 接头分子的另外的非限制性实例包括钙调蛋白结合肽(CBP)、6xHIS、蛋白A、蛋白G、生物素、抗生物素蛋白、链霉抗生物素蛋白、Strep-标签、纤维素结合结构域、麦芽糖结合蛋白、S-肽标签、壳多糖结合标签、免疫反应性表位、表位标签、加GST标签的蛋白、E2标签、HA表位标签、Myc表位、FLAG表位、AU1和AU5表位、Glu-Glu表位、KT3表位、IRS表位、Btag表位、蛋白激酶-C表位、VSV表位。本文件中其他地方描述的二聚化结构域和多聚化结构域的列表定义了多聚化结构域和MHC复合物之间的可选的非共价接头。

[0601] 上述二聚化结构域和多聚化结构域代表特异性结合相互作用。另一种类型的非共价相互作用包括非特异性吸附,例如蛋白到表面上的非特异性吸附。例如,蛋白到玻璃珠上的非共价吸附代表了这类XY相互作用。同样,MHC复合物(包括全长多肽链,包括跨膜部分)与例如树突状细胞的细胞膜的相互作用是非共价的、主要非特异性的XY相互作用的实例。

[0602] 在一些上述实施方案中,若干多聚化结构域(例如与生物素化MHC复合物结合的链霉抗生物素蛋白四聚体)被连接至另一多聚化结构域(例如珠)。出于本公开内容的目的,我们将较小和较大的多聚化结构域两者以及组合的多聚化结构域都称为多聚化结构域。

[0603] MHC多聚体的另外的特征

[0604] 另外的组分被可以偶联至载体,或者作为不偶联至载体的单独组分添加。

[0605] 生物活性分子与MHC多聚体的附接

[0606] MHC复合物与特定T细胞受体的接合导致T细胞中的信号传导级联。然而,T细胞通常通过进入凋亡来对单一信号刺激做出应答。T细胞需要第二信号才能被活化并开始发展成特定的活化状态,例如成为活性细胞毒性T细胞、辅助性T细胞或调节性T细胞。

[0607] 应当理解,本文公开的MHC多聚体可以还包含一个或更多个另外的取代基。术语“一或更多(one or more)”、“多于一(a plurality)”、“一(a)”、“一(an)”和“该(the)”的定义也适用于此。可以将这样的生物活性分子附接至构建体从而影响构建体的,例如关于结合特性、作用、MHC分子特异性、溶解度、稳定性或可检测性的特征。例如,可以在MHC复合物之间提供间距,可以插入荧光共振能量转移(FRET)供体/受体对中的一个或两个发色团,可以附接官能团,或者可以附接具有生物活性的基团。

[0608] MHC多聚体可以与各种分子共价或非共价关联,所述各种分子:具有佐剂作用;作

为免疫靶,例如抗原;具有生物活性,例如酶、受体活性调节物、受体配体、免疫增强剂、药物、毒素、共受体、一般的蛋白和肽;糖部分;脂质基团;包括siRNA的核酸;纳米颗粒;小分子。在下文中,这些分子统称为生物活性分子。这样的分子可以使用与本文其他地方描述的将MHC复合物附接至多聚化结构域的原理相同的原理附接至MHC多聚体。简言之,附接可以通过生物活性分子上的反应性基团和多聚化结构域的反应性基团之间、和/或生物活性分子上的反应性基团和MHC-肽复合物的反应性基团之间的化学反应来完成。可选地,通过多聚化结构域的部分和生物活性分子的部分之间、或MHC-肽复合物的部分和生物活性分子的部分之间的非共价相互作用来实现附接。在生物分子与多聚化结构域的共价和非共价附接两者中,接头分子可以将生物分子与多聚化结构域连接起来。接头分子可以共价或非共价地与两种分子附接。接头分子的实例在本文的其他地方描述。一些MHC聚体结构比其他MHC聚体结构更允许这些种类的修饰。

[0609] 生物活性分子可以重复附接,通过Toll或其他受体帮助固有免疫系统识别和对其进行刺激。

[0610] 特别地,生物活性分子可以选自:(i)蛋白,诸如MHC I类样蛋白,如MIC A、MIC B、CD1d、HLA E、HLA F、HLA G、HLA H、ULBP-1、ULBP-2和ULBP-3;(ii)共刺激分子,诸如CD2、CD3、CD4、CD5、CD8、CD9、CD27、CD28、CD30、CD69、CD134(OX40)、CD137(4-1BB)、CD147、CDw150(SLAM)、CD152(CTLA-4)、CD153(CD30L)、CD40L(CD154)、NKG2D、ICOS、HVEM、HLA II类、PD-1、Fas(CD95)、T细胞和/或NK细胞上表达的FasL、CD40、CD48、CD58、CD70、CD72、B7.1(CD80)、B7.2(CD86)、B7RP-1、B7-H3、PD-L1、PD-L2、CD134L、CD137L、ICOSL、APC细胞和/或肿瘤细胞上表达的LIGHT;(iii)细胞调节分子,诸如在NK细胞上表达的CD16、NKp30、NKp44、NKp46、NKrd0、2B4、KIR、LIR、CD94/NKG2A、CD94/NKG2C、IFN- $\alpha$ 、IFN- $\beta$ 、IFN- $\gamma$ 、IL-1、IL-2、IL-3、IL-4、IL-6、IL-7、IL-8、IL-10、IL-11、IL-12、IL-15、CSF(集落刺激因子)、维生素D3、IL-2毒素、环孢菌素、FK-506、雷帕霉素、TGF- $\beta$ 、克霉唑、尼群地平(nitrendipine)和卡律布德蝎毒素(charybdotoxin);(iv)辅助分子诸如LFA-1、CD11a/18、CD54(ICAM-1)、CD106(VCAM)和CD49a,b,c,d,e,f/CD29(VLA-4);(v)黏附分子,诸如ICAM-1、ICAM-2、GlyCAM-1、CD34、抗-LFA-1、抗CD44、抗 $\beta$ 7、趋化因子、CXCR4、CCR5、抗选择素L、抗选择素E和抗选择素P;(vi)毒性分子,选自毒素、酶、抗体、放射性同位素、化学发光物质、生物发光物质、聚合物、金属颗粒和半抗原,诸如环磷酰胺、甲氨蝶呤、硫唑嘌呤、咪唑立宾(mizoribine)、15-脱氧精胍菌素(15-deoxuspergualin)、新霉素、星形孢菌素、genestein、除莠霉素A、假单胞菌外毒素A、皂草素、利妥昔单抗(Rituxan)、蓖麻毒素、吉妥单抗奥佐米星、志贺毒素、重金属如无机和有机汞、和FN18-CRM9、放射性同位素诸如掺入的碘、钴、硒、氙和磷的同位素,和半抗原诸如DNP和地高辛;和前述任何组合,以及如果相关的话,针对前述的抗体(单克隆、多克隆和重组的)。也可以使用抗体衍生物或其片段。

[0611] 包括MHC多聚体组的方法

[0612] 一些实施方案涉及确定本文定义的组中提供的MHC多聚体与样品中MHC肽识别细胞的结合,诸如怀疑包含MHC肽识别细胞的样品。

[0613] 如本文各处使用的,术语“MHC-肽识别细胞”意图意指能够识别并结合MHC多聚体的细胞。上面给出了“MHC多聚体”的意图含义。MHC-肽识别细胞也可以称为MHC肽识别细胞克隆、靶细胞、靶MHC肽识别细胞、靶MHC分子识别细胞、MHC分子受体、MHC受体、MHC肽特异性

受体或肽特异性细胞。术语“MHC-肽识别细胞”意图包括所有识别并结合MHC分子的正常、异常和缺陷细胞的所有亚群。实际上,正是MHC-肽识别细胞上的受体与MHC分子结合。

[0614] 在疾病和各种状况中,肽由MHC多聚体展示,由免疫系统识别,并产生靶向这样的MHC多聚体的细胞(MHC-肽识别细胞)。因此,这样的MHC蛋白识别细胞的存在是展示MHC蛋白识别细胞识别的肽的MHC多聚体存在的直接指示。展示的肽是指示性的,并可能涉及各种疾病和状况。

[0615] 本文公开的MHC多聚体具有许多用途,并且是例如在治疗、诊断、预后、监测、分层和确定疾病或状况状态的领域中有价值且强大的工具。因此,MHC多聚体可以应用于涉及MHC-肽识别细胞的检测的各种方法和许多应用中,包括诸如流式细胞术、免疫组织化学(IHC)和ELISA样分析的分析。

#### [0616] 检测原理

[0617] 本文公开的诊断程序、免疫监测和一些治疗方法都涉及抗原特异性T细胞的鉴定和/或计数和/或分离。抗原特异性T细胞的鉴定和计数可以通过多种方式进行,并且目前采用几种测定来提供该信息。

[0618] 在下文中,描述了如何使用如本文所述的MHC多聚体和/或抗原肽在各种方法和测定中检测特异性T细胞受体(TCR),并从而检测抗原特异性T细胞。在一些实施方案中,检测包括检测样品中抗原特异性TCR/T细胞的存在,从样品中检测和分离具有抗原特异性TCR的细胞或实体,以及检测和富集样品中具有抗原特异性TCR的细胞或实体。

[0619] 样品可以是生物样品,包括固体组织、固体组织切片,和流体样品诸如但不限于全血、血清、血浆、鼻分泌物、痰、尿液、汗液、唾液、透皮渗出物、咽渗出物、支气管肺泡灌洗液、气管抽吸物、脑脊液、滑液、关节液、玻璃体液、阴道或尿道分泌物、精液等。在本文中,分解的细胞组织,诸如例如毛发、皮肤、滑膜组织、组织活检和指甲刮拭物(scrapings)也被认为是生物样品。

[0620] 本文公开的许多测定和方法对于测定血液样品中的T细胞特别有用。血液样品包括但不限于全血样品或经过处理以去除红细胞和血小板的血液(例如,通过Ficoll密度离心或本领域技术人员已知的其他这样的方法),并直接使用剩余的PBMC样品,其包含感兴趣的T细胞以及B细胞、巨噬细胞和树突状细胞。还包括以其他方式处理的血液样品,例如通过选择或取消选择血液中的细胞或实体来分离各种血细胞亚群。

[0621] 为了能够通过MHC多聚体检测特异性T细胞,可以使用标记(例如,包含独特受体标识符序列的受体结合试剂特异性寡核苷酸)。

#### [0622] 标记分子

[0623] 标记分子是可以在特定分析中被检测的分子,即标记分子提供了可通过所用方法检测到的信号。在一些实施方案中,标记分子的量可以被量化。在一些实施方案中,一种或更多种MHC多聚体包含一个或更多个标记。在一些实施方案中,一个或更多个标记包括受体结合试剂特异性寡核苷酸。在一些实施方案中,一个或更多个标记包括受体结合试剂特异性寡核苷酸和一种或更多种另外的标记,任选地所述一种或更多种另外的标记选自包括以下的组:肽标记、荧光团标记、重金属标记、同位素标记、放射标记、放射性核素、稳定同位素、同位素链和单原子、化学发光标记、生物发光标记、放射性标记、酶标记、DNA荧光染色剂、镧系元素、离子载体、与特定离子结合的螯合化合物或其任何组合。

[0624] 标记分子可以是适合于直接或间接检测的任何标记分子。术语“直接”意指标记分子本身可以被检测,而不需要二级分子,即“一级”标记分子。术语“间接”意指标记分子可以通过使用一种或更多种“二级”分子来检测,即通过检测二级分子与一级分子的结合来进行检测。

[0625] 可以将标记分子附接至多聚化结构域。在一些实施方案中,将标记分子附接至MHC分子。

[0626] 标记分子还可以通过合适的接头附接。适于与标记分子附接的接头是本领域技术人员将容易地知道的,并如本文其他地方关于MHC分子至多聚化结构域的附接描述的。

[0627] 这样的合适的标记化合物的实例是聚合物、核酸、寡核苷酸、肽、荧光标记、磷光标记、酶标记、化学发光标记、生物发光标记、半抗原、抗体、染料、纳米颗粒标记、元素、金属颗粒、重金属标记、同位素标记、放射性同位素、稳定同位素、同位素链和单原子。

[0628] 标记可以是有机或无机分子或颗粒。

[0629] 有机分子标记包括核糖核酸(例如RNA、DNA或非天然DNA、RNA和XNA(例如PNA、LNA、GNA、TNA)和单核苷酸、肽和其他聚酰胺(例如包含b-氨基酸残基的肽)、脂质、糖类、氨基酸和许多其他分子。

[0630] 无机分子标记包括元素(例如镧、铈、镨、钕、钐、钆、铽、镱、镱、铟和其他已知的元素)。元素可以通过螯合物与接头偶联,螯合物与离子配位(与离子非共价相互作用),其中螯合物然后与接头连接(在诸如钆的情况下,其中元素可以以离子形式存在),或者元素可以包含在胶束中。对于一些应用,稀有元素是有利的,而对于一些应用,重金属特别有利。

[0631] 分子标记可以具有在1Da和几百万Da之间的分子量。在一些情况下,非常低的分子量可能是有利的,诸如分子量为1-10Da、11-50Da、50-250Da或251-500Da。这可以是例如当使用质谱来检测元素标记(例如钆,Gd)的身份时的情况。在其他情况下,可以优选低分子量,例如501-2000Da、2001-5000Da或5001-10000Da。这可以是使用例如肽标记时的情况,其中肽标记包含约10-40个氨基酸残基。在又其他情况下,高分子量的分子标记是实用的,并且分子标记的分子量可以是10001-50000Da、50001-200000Da或200000-1000000Da。这可以是例如在其中使用核糖核酸标记的情况下的情况,其中编码区(也称为条形码区或条形码序列)具有显著长度(例如10-20个核苷酸),并且其中具有各自10-20个核苷酸的侧翼引物结合区加上具有不同实际用途的其他序列是可行的。所得的寡核苷酸标记在这些情况下可以是30-1000nt长,对应于约10000-600000Da的分子量。最后,多分子结构,例如在通过结合模板DNA中的特定区域将许多不同的荧光蛋白排列成阵列的情况下,其中总体标记因此包含与许多蛋白结合的长寡核苷酸,并且标记的总分子量因此可以是50000-200000Da、200001-100000或1000001-10000000Da。

[0632] 在一些实施方案中,标记是荧光团和发射或吸收辐射的其他分子。荧光团和发射或吸收辐射的其他分子可以是有机或无机性质的,并且可以是例如小分子以及大蛋白。在一些实施方案中,如果所有荧光团和发射或吸收辐射的其他分子都在相同的窄的最佳发射波长范围内,诸如具有1-10nm、11-30nm、31-100nm、101-200nm、201-300nm、301-400nm、401-500nm、501-600nm、601-700nm、701-800nm、800-900nm、901-1200nm、1201-1500nm或大于1500nm范围内的最佳波长,是特别有利的。例如,如果仪器具有可检测的窄波长范围,则所有标签都落在该检测范围内是有利的。另一方面,如果用于检测由标记发射的辐射的仪器

具有宽跨度的可检测波长,则期望实验中使用的不同标记落在上述范围中的数个范围内,因为这将导致不同标记的发射之间几乎没有重叠,并因此更准确地检测实验中不同标记的相对丰度。发射的辐射可以是磷光、发光、荧光等等。

[0633] 标记化合物可以适当地选自荧光标记诸如5-(和6)-羧基荧光素、5-或6-羧基荧光素、6-(荧光素)-5-(和6)-甲酰氨基己酸、异硫氰酸荧光素(FITC)、罗丹明、四甲基罗丹明、和染料诸如Cy2、Cy3和Cy5、任选取代的香豆素包括AMCA、PerCP、藻胆蛋白包括R-藻红蛋白(RPE)和别藻红蛋白(APC)、Texas Red、Princeton Red、绿色荧光蛋白(GFP)及其类似物,以及R-藻红蛋白或别藻红蛋白与例如Cy5或Texas Red的缀合物,以及基于半导体纳米晶体的无机荧光标记(如量子点和Qdot™纳米晶体),以及基于镧系元素的时间解析荧光标记如Eu<sup>3+</sup>和Sm<sup>3+</sup>,选自半抗原诸如DNP、生物素和地高辛,选自酶标记诸如辣根过氧化物酶(HRP)、碱性磷酸酶(AP)、β-半乳糖苷酶(GAL)、葡萄糖-6-磷酸脱氢酶、β-N-乙酰氨基葡萄糖苷酶、b-葡萄糖醛酸酶、转化酶、黄嘌呤氧化酶、萤火虫萤光素酶和葡萄糖氧化酶(GO),选自发光标记诸如鲁米诺(luminol)、异鲁米诺、吡啶酯、1,2-二氧杂环丁烷和吡啶并哒嗪,选自放射性标记诸如掺入的碘、钴、硒、氙和磷的同位素,以及选自单原子诸如锌(Zn)、铁(Fe)、镁(Mg),任何镧系元素(Ln)包括La、Ce、Pr、Nd、Pm、Sm、Eu、Gd、Tb、Dy、Ho、Er、Tm、Yb和Lu;钪(Sc)和钇(Y)。

[0634] 放射性标记在标记MHC多聚体所携带的肽的方面可能特别令人感兴趣。

[0635] 基于标记分子的特定特性,存在不同的标记和检测原理。不同类型的标记的实例是放射性辐射的发射(放射性核素、同位素)、光的吸收(例如染料、发色团)、激发后的光发射(来自荧光染料的荧光)、NMR(来自顺磁性分子的核磁共振)和光的反射(来自例如诸如各种尺寸和形状的金珠、塑料珠或玻璃珠/颗粒的散射)。可选地,标记分子可以具有酶促活性,通过酶促活性,它们催化标记分子附近环境中的化学物质之间的反应,产生信号,这包括光的产生(化学发光)、发色团染料的沉淀或可被另一层检测分子检测的沉淀。酶促产物可以沉积在酶的位置,或者在基于细胞的分析系统中,与细胞膜反应或扩散到酶促产物所附着的细胞中。标记分子的实例和相关检测原理在表3中示出。

[0636] 表3:标记分子的实例和相关检测原理

| 标记物质                              | 作用  | 测定原理   |
|-----------------------------------|---|--|
| 荧光染料                              | 具有特定光谱的光的发射   | 光度学、显微术、光谱术<br>PMT、摄影胶片、CCD (彩色捕获器件或电荷耦合器件)。             |
| 放射性核素                             | 辐照, $\alpha$ 、 $\beta$ 或 $\gamma$ 射线  | 闪烁计数、GM 管、摄影胶片、磷光成像屏的激发                                  |
| 酶;<br>HRP (辣根过氧化物酶), 一般过氧化物酶      | 使用鲁米诺为氧受体催化 $H_2O_2$ 还原, 产生氧化的鲁米诺+光<br>使用可溶性染料或含有半抗原(诸如生物素残基)的分子作为氧受体催化 $H_2O_2$ 还原, 产生沉淀。半抗原可以被检测分子识别。 | 光度学、显微术、光谱术<br>PMT、摄影胶片、CCD(彩色捕获器件或电荷耦合器件),<br>二级标记连接的抗体 |
| 颗粒; 金、聚苯乙烯珠、花粉和其他颗粒               | 关联实体的散射、反射和透明度的变化   | 显微术、细胞术、电子显微术<br>PMT、光检测装置、流式细胞仪散射                       |
| [0637] AP (碱性磷酸酶)                 | 催化不可检测的分子向沉淀的可检测分子的化学转化, 诸如染料或半抗原   | 光度学、显微术、光谱术<br>二级标记连接抗体                                  |
| 离子载体或与特定离子(例如 $Ca^{2+}$ )结合的螯合化合物 | 当结合时吸收和发射光谱的变化。强度变化   | 光度学, 细胞术, 光谱术  |
| 镧系元素                              | 荧光  | 光度学, 细胞术, 光谱术  |
|                                   | 磷光  |  |
|                                   | 顺磁性的  | NMR (核磁共振)   |
| DNA 荧光染色剂                         | 碘化丙锭<br>Hoechst 染色剂<br>DAPI<br>AMC<br>DraQ5™<br>吡啶橙<br>7-AAD  | 光度学, 细胞术, 光谱术  |
| 寡核苷酸标签/标识符                        | 独特序列  | PCR 扩增、测序  |

[0638] 光度学应被理解为任何可用于检测源自发射一个或多个波长或光谱的光的光源的光的强度、分析光的波长光谱和/或测量光的累积的方法。

[0639] 标记分子可用于标记MHC多聚体以及与MHC多聚体一起使用的其他试剂, 例如抗体、适配体或能够结合另一蛋白、糖、DNA或其他分子中的特定结构的其他蛋白或分子。在下文中, 能够结合另一分子中特定结构的分子被称为标志物。可以将标记分子通过共价连接附接至给定的MHC多聚体或任何其他蛋白标志物, 如本文其他地方针对MHC多聚体至多聚化结构域的附接描述的。可以直接在标记分子中的反应性基团和标志物分子中的反应性基团之间进行附接, 或者可以通过与标记分子和标志物共价附接的接头进行附接, 两者如本文其他地方描述的。当对MHC多聚体进行标记时, 可以将标记附接至MHC复合物(重链、p2m或肽)或附接至多聚化结构域。

[0640] 具体地, 可以将一个或多个标记分子附接至载体分子, 或可以将一个或多个

标记分子附接至一个或更多个支架,或可以将一个或更多个标记化合物附接至一个或更多个MHC复合物,或可以将一个或更多个标记化合物附接至载体分子和/或一个或更多个支架和/或一个或更多个MHC复合物,或可以将一个或更多个标记化合物附接至MHC分子所携带的肽。

[0641] 标志物上的单个标记分子并不总是产生足够的信号强度。信号强度可以通过将单一标记分子组装成包含两个或更多个标记分子残基的大多标记化合物来改进。多标记化合物的产生可以通过标记分子与主要结构分子的共价或非共价关联来实现。这样的结构的实例是合成聚合物或天然聚合物(例如dextramer)、蛋白(例如链霉抗生物素蛋白)或聚合物。多标记化合物中的标记分子可以都是同一类型的,或者可以是不同标记分子的混合物。

[0642] 在一些应用中,作为组合步骤或以单独步骤应用不同的MHC复合物可以是有利的。这样的不同的MHC多聚体可以被不同地标记(即通过用不同的标记化合物标记),使得能够对不同的靶MHC肽识别细胞可视化。因此,如果存在具有不同标记化合物的数种不同的MHC多聚体,如果每种MHC多聚体呈递不同的肽,则有可能同时鉴定多于一种特异性受体。

[0643] 检测原理可以应用于流式细胞术、固定细胞术(stationary cytometry)和基于批次的分析。大多数基于批次的方法可以根据检测的目的使用任何标记物质。流式细胞术主要利用荧光,而固定细胞术主要利用光吸收,例如来自酶促活性的染料或发色团沉积。在以下部分中,将例示涉及流式细胞术的荧光检测的原理,并且将在固定细胞术的背景中例示涉及发色团检测的原理。然而,标记分子可以应用于本公开内容中描述的任何分析。

[0644] 在流式细胞术中特别使用的标记分子

[0645] 在流式细胞术中,典型的标记通过其荧光来检测。最常见的,阳性检测是基于来自单荧光染料的光的存在,但是在其他技术中,信号是通过发射光的波长偏移来检测的;如在基于FRET的技术中,其中激发的荧光染料将其能量转移到相邻结合的荧光染料,该相邻结合的荧光染料发射光,或者当使用Ca<sup>2+</sup>螯合荧光探针时,在与钙结合后发射(和吸收)光谱改变。

[0646] 流式细胞术中使用的一些非限制性标记分子在表4-表5中示出,并在下文中描述。

[0647] 荧光标记可以选自以下一种或更多种:(i) Fluor染料, Pacific Blue<sup>TM</sup>、Pacific Orange<sup>TM</sup>、Cascade Yellow<sup>TM</sup>; (ii) AlexaFluor® (AF), 例如AF405、AF488、AF500、AF514、AF532、AF546、AF555、AF568、AF594、AF610、AF633、AF635、AF647、AF680、AF700、AF710、AF750、AF800; (iii) 基于量子点的染料, QDot® 纳米晶体(Invitrogen, Molecular Probes), 例如Qdot®525、Qdot®565、Qdot®585、Qdot®605、Qdot®655、Qdot®705、Qdot®800; (iv) DyLight<sup>TM</sup>染料(Pierce) (DL), 例如DL549、DL649、DL680、DL800; (v) 荧光素(Flu)或其任何衍生物, 例如FITC; (vi) Cy-染料, 例如Cy2、Cy3、Cy3.5、Cy5、Cy5.5、Cy7; (vii) 荧光蛋白, 例如RPE、PerCp、APC、绿色荧光蛋白、GFP和GFP衍生的突变体蛋白; BFP、CFP、YFP、DsRed、T1、Dimer2、mRFP1、MBanana、mOrange、dTomato、tdTomato、Tangerine、Strawberry、mCherry; (viii) 串联染料, 例如RPE-Cy5、RPE-Cy5.5、RPE-Cy7、RPE-AlexaFluor® 串联缀合物、RPE-Alexa610、RPE-TxRed、APC-Alexa600、APC-Alexa610、APC-Alexa750、APC-Cy5、APC-Cy5.5; (ix) 离子载体; 离子螯合荧光探针, 例如当结合特定离子时改变波长的探针, 诸如当与特定离子(诸如钙)结合时改变强度的钙探针; (x) 同一标志物上的荧光染料组合。因

此,标志物不是通过单一荧光染料来识别的,而是通过作为荧光染料的特定组合的识别代码以及相互相关的强度比来识别的。

[0648] 表4: 优选荧光染料的实例

| 荧光团/荧光染料  | 激发 nm | 发射 nm |
|---|-------|-------|
| 2-(4'-马来酰亚胺基苯胺基)萘-6-磺酸钠盐                        | 322   | 417   |
| 5-(((2-碘乙酰基)氨基)乙基)氨基)萘-1-磺酸                     | 336   | 490   |
| 苝-1-丁酸  | 340   | 376   |
| AlexaFluor 350 (7-氨基-6-磺酸-4-甲基香豆素-3-乙酸)         | 346   | 442   |
| AMCA (7-氨基-4-甲基香豆素-3-乙酸)                        | 353   | 442   |
| 7-羟基-4-甲基香豆素-3-乙酸                               | 360   | 455   |
| Marina Blue (6,8-二氟-7-羟基-4-甲基香豆素-3-乙酸)          | 362   | 459   |
| 7-二甲基氨基-香豆素-4-乙酸                                | 370   | 459   |
| 荧光胺-N-丁胺加合物                                     | 380   | 464   |
| 7-羟基-香豆素-3-羧酸                                   | 386   | 448   |
| CascadeBlue (苝-三磺酸乙酰叠氮化物)                       | 396   | 410   |
| Cascade Yellow                                  | 409   | 558   |
| [0649] Pacific Blue (6,8 二氟-7-羟基香豆素-3-羧酸)       | 416   | 451   |
| 7-二乙基氨基-香豆素-3-羧酸                                | 420   | 468   |
| N-(((4-叠氮苯甲酰基)氨基)乙基)-4-氨基-3,6-二硫基-1,8-萘酰亚胺二钾盐   | 426   | 534   |
| Alexa Fluor 430                                 | 434   | 539   |
| 3-苝十二酸  | 440   | 448   |
| 8-羟基苝-1,3,6-三磺酸三钠盐                              | 454   | 511   |
| 12-(N-(7-硝基苯-2-氧杂-1,3-二唑-4-基)氨基)十二酸             | 467   | 536   |
| N,N'-二甲基-N-(碘乙酰基)-N'-(7-硝基苯-2-氧杂-1,3-二唑-4-基)乙二胺 | 478   | 541   |
| Oregon Green 488 (二氟羧基荧光素)                      | 488   | 518   |
| 5-碘乙酰胺荧光素                                       | 492   | 515   |
| 碘化丙锭-DNA 加合物                                    | 493   | 636   |
| 羧基荧光素   | 495   | 519   |

[0650] 优选荧光染料家族的实例在表5中示出。

[0651] 表5: 荧光染料家族示例荧光染料

| 荧光染料家族                 | 示例荧光染料   |
|------------------------|--|
| [0652] AlexaFluor®(AF) | AF®350、AF405、AF430、AF488、AF500、AF514、AF532、AF546、AF555、AF568、AF594、AF610、AF633、AF635、AF647、AF680、AF700、AF710、AF750、AF800 |

|        |                      |   |
|--------|----------------------|---|
|        | 基于量子点(Qdot®)的染料      | Qdot®525、Qdot®565、Qdot®585、Qdot®605、Qdot®655、Qdot®705、Qdot®800  |
|        | DyLight™ 染料(DL)小荧光染料 | DL549、DL649、DL680、DL800<br>FITC、Pacific Blue™、Pacific Orange™、Cascade Yellow™、Marina Blue™、DSred、DSred-2、7-AAD、TO-Pro-3               |
|        | Cy-染料                | Cy2、Cy3、Cy3.5、Cy5、Cy5.5、Cy7   |
|        | 藻胆蛋白:                | R-藻红蛋白(RPE)、PerCP、别藻蓝蛋白(APC)、B-藻红蛋白、C-藻蓝蛋白  |
| [0653] | 荧光蛋白                 | (E)GFP 和 GFP ((增强型)绿色荧光蛋白)衍生的突变体蛋白; BFP、CFP、YFP、DsRed、T1、Dimer2、mRFP1、MBanana、mOrange、dTomato、tdTomato、mTangerine、mStrawberry、mCherry |
|        | 具有 RPE 的串联染料         | RPE-Cy5、RPE-Cy5.5、RPE-Cy7、RPE-AlexaFluor® 串联缀合物; RPE-Alexa610、RPE-TxRed   |
|        | 具有 APC 的串联染料         | APC-Alexa600、APC-Alexa610、APC-Alexa750、APC-Cy5、APC-Cy5.5  |
|        | 钙染料                  | Indo-1-Ca2+、Indo-2-Ca2+   |

## [0654] 肽标记

[0655] 标记可以是包含一段连续氨基酸残基的肽标记。这是可以确定其身份的‘编码区’。

[0656] 肽标记可以包含限定数目的连续氨基酸或由限定数目的连续氨基酸组成。因此，在一些实施方案中，肽标记包含2个或更多个连续氨基酸，诸如2-3、3-4、4-5、5-6、6-7、7-8、8-9、9-10、10-11、11-12、12-13、13-14、14-15、15-16、16-17、17-18、18-19、19-20、20-21、21-22、22-23、23-24、24-25、25-26、26-27、27-28、28-29、29-30、30-31、31-32、32-33、33-34、34-35、35-36、36-37、37-38、38-39、39-40、40-45、45-50、50-55、55-60、60-65、65-70、70-75、75-80、80-85、85-90、90-95、95-100、100-110、110-120、120-130、130-140、140-150、150-160、160-170、170-180、180-190、190-200、200-225、225-250、250-275、275-300、300-350、350-400、400-450、450-500、500-600、600-700、700-800、800-900、900-1000、1000-1500、1500-2000，或超过2000个连续氨基酸。

[0657] 肽标记可以包含一段连续的氨基酸残基(编码区)和蛋白酶裂解位点。蛋白酶裂解位点可以优选地位于将标记连接至MHC多聚体的接头附近。

[0658] 当MHC多聚体靠近蛋白酶时，肽标记被裂解，并且编码区从MHC多聚体释放。可以对样品细胞进行沉淀，并且可以通过质谱分析上清液，来确定所释放的标记的身份和量。

[0659] 能够裂解肽标记的蛋白酶可以被涂覆在样品细胞的表面上，例如通过添加抗体-蛋白酶缀合物，其中抗体识别特定的细胞表面结构。

[0660] 肽标记可以包含天然(或标准)氨基酸、非天然存在的氨基酸(非蛋白源性或非标准氨基酸)或两者。在一些实施方案中，肽标记包含标准和非标准氨基酸。天然氨基酸是天然存在于自然界中并天然掺入多肽中(蛋白源性)的天然存在的氨基酸。它们由以下组成：20种遗传编码的氨基酸Ala、Arg、Asn、Asp、Cys、Gln、Glu、Gly、His、Ile、Leu、Lys、Met、Phe、Pro、Ser、Tyr、Thr、Trp、Val；以及2种通过独特的合成机制掺入蛋白中的氨基酸：Sec(硒代半胱氨酸，或U)和Py1(吡咯赖氨酸，O)。这些氨基酸都是L-立体异构体。除了22种天然或标准氨基酸之外，还有许多其他非天然存在的氨基酸(非蛋白源性或非标准氨基酸)。它们或

者不存在于蛋白中,或者不是由标准细胞机制直接和单独产生的。非标准氨基酸通常通过对标准氨基酸的修饰诸如翻译后修饰来形成。本文公开的任何氨基酸可以是L-构型或D-构型。标准和/或非标准氨基酸可以通过肽键连接形成线性肽链。

[0661] 如本领域已知的,术语肽还包括由化学或酶催化反应引入的翻译后修饰。同样,功能等效者可以包括化学修饰,诸如泛素化、标记(例如,用放射性核素、各种酶等标记)、聚乙二醇化(用聚乙二醇衍生)、或通过插入(或通过经化学合成取代为)通常不存在于人类蛋白中的一个氨基酸(或更多个氨基酸)。

[0662] 蛋白翻译后修饰 (PTM) 通过共价添加官能团或蛋白、蛋白水解裂解调节性亚基或降解整个蛋白来增加蛋白组的功能多样性。这些修饰包括磷酸化、糖基化、泛素化、亚硝基化、甲基化、乙酰化、脂化 (C-末端糖基磷脂酰肌醇 (GPI) 锚定、N-末端豆蔻酰化、S-豆蔻酰化、S-异戊烯化)、酰胺化和蛋白水解,并影响正常细胞生物学和发病机制的几乎所有方面。

[0663] 受体结合试剂特异性寡核苷酸

[0664] 如本文公开的MHC单体或MHC多聚体可以包含至少一个核酸标记,诸如核苷酸标记,例如寡核苷酸标记(例如,受体结合试剂特异性寡核苷酸)。在一些实施方案中,标记是寡核苷酸,诸如DNA寡核苷酸(DNA标记)。术语核酸标记、核酸分子、核苷酸标记、寡核苷酸标记、DNA分子、DNA标记、DNA标签、受体结合试剂特异性寡核苷酸、DNA寡核苷酸和核酸组分在本文可互换使用。在一些实施方案中,核酸标记包含以下组分中的一种或更多种:条形码区域、5' 第一引物区域(正向)、3' 第二引物区域(反向)、随机核苷酸区域、连接体分子、增加稳定性的组分、任何上述组分之间的短核苷酸接头、用于测序的衔接子和/或退火区域。

[0665] 核酸标记可以包含至少一个条形码区域(例如,条形码序列、独特受体标识符序列)。条形码区域包含连续核酸的序列。本文公开的核酸标记可以包含许多连续的核酸。核酸可以是任何类型的核酸或其修饰物,天然存在的或合成制备的(人工核酸)。核酸标记可以包含DNA或由DNA组成,和/或包含RNA或由RNA组成。

[0666] 在一些实施方案中,核酸标记包含以下或由以下组成:人工核酸或非天然核酸(Xeno nucleic acid, XNA)。化学家已经设计和合成了人工核酸类似物,并且人工核酸类似物包括PNA、吗啉代核酸和LNA,以及GNA、TNA、HNA和CeNA。它们中的每一种都通过分子主链的变化区别于天然存在的DNA或RNA。在一些实施方案中,核酸标记包含以下或由以下组成:XNA、PNA、LNA、TNA、GNA、HNA和CeNA中的一种或更多种。在一些实施方案中,至少一种核酸分子包含以下或由以下组成:DNA、RNA和/或人工核苷酸(诸如PLA或LNA)。优选DNA,但也可以包含其他核苷酸以例如增加稳定性。

[0667] 寡核苷酸可以是天然寡核苷酸,诸如DNA或RNA,或者是PNA、LNA或另一种类型的非天然寡核苷酸。寡核苷酸可以在碱基实体、糖实体或连接个体核苷酸的接头上被修饰的。核酸分子的长度也可以变化,例如具有范围20-100个核苷酸,诸如30-100个、30-80个或30-50个核苷酸内的长度。在一些实施方案中,至少一种核酸分子是长度为1个、2个、3个、4个、5个、6个、7个、8个、9个、10个、11个、12个、13个、14个、15个、16个、17个、18个、19个、20个、21个、22个、23个、24个、25个、26个、27个、28个、29个、30个、31-35个、36-50个、51-100个或超过100个核苷酸的寡核苷酸。

[0668] 在一些实施方案中,核酸标记包含1个至1,000,000个核苷酸,诸如1个、2个、3个、4个、5个、6个、7个、8个、9个、10个、11个、12个、13个、14个、15个、16个、17个、18个、19个、20

个、21个、22个、23个、24个、25个、26个、27个、28个、29个、30个、31个、32个、33个、34个、35个、36个、37个、38个、39个、40个核苷酸；例如1-3个、3-5个、5-10个、10-15个、15-20个、20-25个、25-30个、30-35个、35-40个、40-50个、50-60个、60-70个、70-80个、80-90个、90-100个、100-110个、110-120个、120-130个、130-140个、140-150个、150-175个、175-200个、200-250个、250-300个、300-400个、400-500个、500-600个、600-700个、700-800个、800-900个、900-1000个、1000-1500个、1500-2000个、2000-3000个、3000-4000个、4000-5000个、5000-7500个、7500-10,000个、10,000-100,000个、100,000-1,000,000个核苷酸。

[0669] 核酸标记可以包含许多连续的核酸。核酸的序列用作可以被识别(诸如扩增和/或测序)的代码。在一些实施方案中,核酸标记包含被设计成通过例如PCR扩增的中心核酸链段(条形码区域)。

[0670] 核酸标记的可识别连续核酸或可识别序列在本文被称为“条形码”、“条形码区域”、“核酸条形码”、“独特序列”、“独特核苷酸序列”、“独特受体标识符序列”和“编码序列”(可互换使用)。条形码区域包含构成核酸序列的多个连续核酸。

[0671] 在一些实施方案中,核酸条形码是范围为10个至超过50个核苷酸的独特寡核苷酸序列。在该实施方案中,条形码在3'末端和5'末端具有共有的扩增序列,并在中间具有独特序列。这种独特序列可以通过测序揭示,并可以作为用于给定MHC多聚体的特定条形码。独特序列,即条形码,包含一系列核苷酸,这些核苷酸一起形成可以基于其组成进行特异性识别的序列(一系列核苷酸)。该序列组成使得仅基于每种条形码的独特序列,条形码#1能够与条形码#2、#3、#4等,多达超过100,000种条形码相区分。完整的核苷酸条形码也可以包含一系列彼此连接的独特核苷酸序列的组合。这一系列独特序列将一起分配条形码。

[0672] 在一些实施方案中,每种独特核苷酸序列(条形码)包含1个、2个、3个、4个、5个、6个、7个、8个、9个、10个、11个、12个、13个、14个、15个、16个、17个、18个、19个、20个、21个、22个、23个、24个、25个、26个、27个、28个、29个、30个、31-35个、36-50个、51-100个或超过100个核苷酸。

[0673] 在一些实施方案中,标记是寡核苷酸,和/或独特序列具有1个、2个、3个、4个、5个、6个、7个、8个、9个、10个、11个、12个、13个、14个、15个、16个、17个、18个、19个、20个、21个、22个、23个、24个、25个、26个、27个、28个、29个、30个、31-35个、36-50个、51-100个或超过100个核苷酸的长度。在一些实施方案中,独特序列短于标记的总长度。

[0674] 在一些实施方案中,条形码区域包含以下或由以下组成:2-5个、5-10个、10-15个、15-20个、20-25个、25-30个、30-40个、40-45个、45-50个、50-60个、60-70个、70-80个、80-90个、90-100个、100-110个、110-120个、120-130个、130-140个、140-150个、150-175个、175-200个、200-250个、250-300个、300-400个、400-500个核苷酸。

[0675] 独特核苷酸序列(条形码)仅用作MHC分子与其靶之间分子相互作用的识别标签。独特核苷酸序列优选地不与任何天然存在的DNA序列相同,尽管可以存在序列相似性或同一性。

[0676] 每个核酸条形码应该与给定实验中的另外的条形码保持足够的差异,以允许给定的条形码的特定识别,使其与其他条形码可区分。

[0677] 核酸组分(例如,DNA)具有特殊的结构。因此,在实施方案中,至少一种核酸分子(标记)包含至少5'第一引物区域、中心区域(条形码区域)和3'第二引物区域。以这种方式,

中心区域(条形码区域)可以被引物组扩增。

[0678] 核酸分子与多聚化结构域的偶联也可以变化。因此,在一些实施方案中,至少一个核酸分子通过链霉抗生物素蛋白-生物素结合和/或抗生物素蛋白-生物素结合与所述多聚化结构域连接。也可以使用其他偶联部分。

[0679] 在一些实施方案中,核酸标记包含连接体分子,该连接体分子能够与多聚化结构域或MHC分子上的组分相互作用。在一些实施方案中,连接体分子是生物素或抗生物素蛋白。在一些实施方案中,接头包含链霉抗生物素蛋白,标记通过其生物素或抗生物素蛋白连接体分子与链霉抗生物素蛋白结合。在一些实施方案中,核酸标记包含随机核苷酸区域。该随机nt区域是用于检测标记污染物的潜在工具。在一些实施方案中,随机nt区域包含3-20个核苷酸,诸如3个、4个、5个、6个、7个、8个、9个、10个、11个、12个、13个、14个、15个、16个、17个、18个、19个或20个nt。

[0680] 实验中使用的不同标记可以具有相同的扩增特性并共有共同的引物区域:共同的引物区域和共有的扩增特性一起将确保在细胞相互作用和分选后存在的所有标记被同等地扩增,由此没有序列会由于测序反应而被偏倚。

[0681] 在不同标记上使用相同的引物区域,存在一种标记被另一种标记污染的固有风险——尤其是在扩增反应之后。为了能够追溯潜在的污染,可以在核酸标记中包含短的“随机核苷酸区域”。由于随机核苷酸区域对于每种标记都是独特的,因此将可以检查测序数据并查看给定标记是否存在多种读段。即随机核苷酸区域是克隆性对照区域。在一些实施方案中,随机核苷酸区域由2-20个核酸;诸如2个、3个、4个、5个、6个、7个、8个、9个、10个、11个、12个、13个、14个、15个、16个、17个、18个、19个或20个核酸组成。由6个核苷酸组成的随机核苷酸区域在本文可以称为TM6',以此类推。在一些实施方案中,核酸标记包含一种或更多种增加稳定性的组分(诸如HEG或TEG)。

[0682] 当与细胞混合时,标记优选地是稳定的:因为与细胞混合可能使标记暴露于核酸酶消化。使这种情况最小化的措施可以是在寡核苷酸标记的一个或两个末端添加六乙二醇(HEG)或TEG形式的修饰。另外,在所应用的缓冲液中的稳定性可以通过添加对寡核苷酸起保护作用的成分(例如鲑鱼DNA和EDTA)来确保。

[0683] 在一些实施方案中,核酸标记包含样品识别序列。为了能够在每个测序反应中分析多于一个样品,核酸标记可以被分配另外的识别特征,即样品识别序列。样品识别序列不是标记初始设计的一部分,而是在细胞相互作用和在PCR中通过引物分选后被分配——因此来源于同一样品的所有细胞将具有相同的样品识别序列。在一些实施方案中,样品识别序列是由2-20个核苷酸(诸如2个、3个、4个、5个、6个、7个、8个、9个、10个、11个、12个、13个、14个、15个、16个、17个、18个、19个、20个或这些中的任何两个之间的范围的核苷酸)组成的短序列。可以将样品识别序列附接至引物,诸如正向引物。

[0684] 在一些实施方案中,核酸标记是包含正向引物、条形码区域和反向引物的“1寡核苷酸系统”。在一些实施方案中,核酸标记是具有两个序列的“2寡核苷酸系统”,第一序列包含正向引物、条形码区域和退火区域;并且第二序列包含退火区域、条形码区域和反向引物。

[0685] 试剂盒

[0686] 本文公开内容包括试剂盒,例如,包含以下的试剂盒:多于一种受体检测构建体,

其中受体检测构建体包含两种或更多种受体结合试剂,其中受体结合试剂能够与受体特异性结合,并且其中受体检测构建体中的每一种包含受体结合试剂特异性寡核苷酸,所述受体结合试剂特异性寡核苷酸包含用于受体结合试剂的独特受体标识符序列。试剂盒可以包含:多于一种细胞组分结合试剂,其中多于一种细胞组分结合试剂中的每一种包含细胞组分结合试剂特异性寡核苷酸,所述细胞组分结合试剂特异性寡核苷酸包含用于细胞组分结合试剂的独特标识符序列,并且其中细胞组分结合试剂能够与细胞组分靶特异性结合。在一些实施方案中,受体结合试剂特异性寡核苷酸包含第二通用序列,细胞组分结合试剂特异性寡核苷酸包含第三通用序列,并且第二通用序列和第三通用序列是不同的。在一些实施方案中,第二通用序列与第三通用序列小于约85%(例如,0.000000001%、0.00000001%、0.0000001%、0.000001%、0.00001%、0.0001%、0.001%、0.01%、0.1%、1%、2%、3%、4%、5%、6%、7%、8%、9%、10%、11%、12%、13%、14%、15%、16%、17%、18%、19%、20%、21%、22%、23%、24%、25%、26%、27%、28%、29%、30%、31%、32%、33%、34%、35%、36%、37%、38%、39%、40%、41%、42%、43%、44%、45%、46%、47%、48%、49%、50%、51%、52%、53%、54%、55%、56%、57%、58%、59%、60%、61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、81%、82%、83%、84%或这些值中的任何两个值之间的数字或范围)相同。

[0687] 试剂盒可以包含:多于一种寡核苷酸条形码,其中多于一种寡核苷酸条形码中的每一种包含第一通用序列、分子标记和靶结合区,并且其中多于一种寡核苷酸条形码中的至少10种包含不同的分子标记序列。多于一种寡核苷酸条形码可以与固体支持物关联。靶结合区可以包含基因特异性序列、寡(dT)序列、随机多聚体或其任何组合。试剂盒可以包含:逆转录酶。逆转录酶可以包括病毒逆转录酶(例如,鼠白血病病毒(MLV)逆转录酶或Moloney鼠白血病病毒(MMLV)逆转录酶)。试剂盒可以包含:缺乏5'至3'外切核酸酶活性和3'至5'外切核酸酶活性中的至少一种的DNA聚合酶(例如,Klenow片段)。试剂盒可以包含:缓冲液、盒或两者。试剂盒可以包含:一种或更多种用于逆转录反应和/或扩增反应的试剂。在一些实施方案中,多于一种寡核苷酸条形码各自包含细胞标记。多于一种寡核苷酸条形码的每一种细胞标记可以包含至少6个核苷酸。在一些实施方案中,多于一种寡核苷酸条形码中与同一固体支持物关联的寡核苷酸条形码可以包含相同的细胞标记。在一些实施方案中,多于一种寡核苷酸条形码中与不同固体支持物关联的寡核苷酸条形码可以包含不同的细胞标记。

[0688] 固体支持物可以包括合成颗粒、平坦表面,或其组合。多于一种寡核苷酸条形码中的至少一种寡核苷酸条形码可以固定或部分固定在合成颗粒上,或者多于一种寡核苷酸条形码中的至少一种寡核苷酸条形码可以包封或部分包封在合成颗粒中。合成颗粒可以是可破坏的(例如,可破坏的水凝胶颗粒)。合成颗粒可以包括珠。珠可以是琼脂糖凝胶珠、链霉抗生物素蛋白珠、琼脂糖珠、磁珠、缀合珠、蛋白A缀合珠、蛋白G缀合珠、蛋白A/G缀合珠、蛋白L缀合珠、寡(dT)缀合珠、二氧化硅珠、二氧化硅样珠、抗生物素微珠、抗荧光染料微珠或其任何组合。合成颗粒可以包含选自由以下组成的组的材料:聚二甲基硅氧烷(PDMS)、聚苯乙烯、玻璃、聚丙烯、琼脂糖、明胶、水凝胶、顺磁物质、陶瓷、塑料、玻璃、甲基苯乙烯、丙烯酸聚合物、钛、胶乳、琼脂糖凝胶、纤维素、尼龙、硅酮及其任何组合。多于一种寡核苷酸条形码

中的每一种寡核苷酸条形码可以包含接头官能团。合成颗粒可以包含固体支持物官能团。其中支持物官能团和接头官能团可以彼此关联,并且任选地接头官能团和支持物官能团单独地选自以下组成的组:C6、生物素、链霉抗生物素蛋白、一种或更多种伯胺、一种或更多种醛、一种或更多种酮及其任何组合。

[0689] 试剂盒可以包含:一种或更多种包含第一通用序列、第二通用序列和/或第三通用序列的引物。第二通用序列可以包含与SEQ ID NO:1 (GGAGGGAGGTTAGCGAAGGT)至少约80% (例如,80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、100%或这些值中的任何两个值之间的数字或范围)相同的序列。靶结合区可以包含多(dT)区,并且细胞组分结合试剂特异性寡核苷酸可以包含多(dA)区。细胞组分结合试剂特异性寡核苷酸可以包含与多(dA)区相邻的对齐序列。在一些实施方案中,(a)对齐序列包含鸟嘌呤、胞嘧啶、胸腺嘧啶、尿嘧啶或其组合;(b)对齐序列包含多(dT)序列、多(dG)序列、多(dC)序列、多(dU)序列或其组合;和/或(c)对齐序列位于多(dA)区的5'。细胞组分结合试剂特异性寡核苷酸可以通过接头与细胞组分结合试剂关联。细胞组分结合试剂特异性寡核苷酸可以被配置为能够从细胞组分结合试剂脱离。接头可以包含碳链。碳链可以包含2-30个碳,并且还任选地碳链可以包含12个碳。接头可以包含5'氨基修饰物C12 (5AmMC12)或其衍生物。

[0690] 细胞组分靶可以包括蛋白靶。细胞组分结合试剂可以包含抗体或其片段。细胞组分靶可以包括糖类、脂质、蛋白、细胞外蛋白、细胞表面蛋白、细胞标志物、B细胞受体、T细胞受体、主要组织相容性复合体、肿瘤抗原、受体、细胞内蛋白或其任何组合。细胞组分结合试剂可以包含抗体或其片段。细胞组分靶可以在细胞表面上。受体结合试剂特异性寡核苷酸可以包含第三分子标记,和/或细胞组分结合试剂特异性寡核苷酸可以包含第二分子标记。受体结合试剂特异性寡核苷酸的长度可以为约10个核苷酸至约100个核苷酸。受体可以包括分化簇(CD)分子。受体可以包括免疫受体(例如T细胞受体(TCR))。独特受体标识符序列的长度可以是约3个核苷酸至约100个核苷酸。靶结合区可以包含多(dT)区,并且受体结合试剂特异性寡核苷酸可以包含多(dA)区。受体结合试剂特异性寡核苷酸可以包含与多(dA)区相邻的对齐序列。在一些实施方案中,(a)对齐序列包含鸟嘌呤、胞嘧啶、胸腺嘧啶、尿嘧啶或其组合;(b)对齐序列包含多(dT)序列、多(dG)序列、多(dC)序列、多(dU)序列或其组合;和/或(c)对齐序列位于多(dA)区的5'。受体结合试剂特异性寡核苷酸可以通过接头与受体结合试剂关联。受体结合试剂特异性寡核苷酸可以被配置为能够从受体结合试剂脱离。接头可以包含碳链。碳链可以包含2-30个碳,并且还任选地碳链可以包含12个碳。接头可以包含5'氨基修饰物C12 (5AmMC12)或其衍生物。多于一种受体检测构建体可以包含本文所述的一种或更多种MHC多聚体。两种或更多种受体结合试剂包含本文所述的两种或更多种MHC-肽复合物。

## 实施例

[0691] 上文讨论的实施方案的一些方面在以下实施例中进一步详细公开,其并不是以任何方式意在限制本公开内容的范围。

[0692] 实施例1方案

[0693] dCODE™ Dextramer® (BD-Rhapsody™) 染色方案

[0694] 该方案可用于使用BD Rhapsody™单细胞分析系统,与靶向mRNA分析一起,对细胞样品中的抗原特异性T细胞进行谱分析和定量。

[0695] dCODE™ **Dextramer**® (BD-Rhapsody™相容的)试剂可以包含带有多个MHC-肽复合物的右旋糖酐聚合物主链、相应的独特DNA条形码寡核苷酸和R-藻红蛋白(PE),以用于在将样品加载到Rhapsody™系统中之前分选dCODE Dextramer阳性细胞。

[0696] 独特DNA条形码寡核苷酸可以包含BD-Rhapsody™相容的用于PCR扩增的PCR手柄序列、独特分子标识符(UMI)序列和指定MHC-肽特异体(specificity)的DNA条形码序列(参见,例如,图9A;SEQ ID NO:3)。

[0697] dCODE Dextramer (Rhapsody™)可以在含有1%牛血清白蛋白(BSA)和15mM NaN<sub>3</sub>, pH 7.2的PBS缓冲液中以 $1.6 \times 10^{-7}$ M的浓度提供。

[0698] 2μl(一次测试)可用于 $1-3 \times 10^6$ 个PBMC的染色。

[0699] 每种dCODE Dextramer可以通过其HLA-等位基因/肽/DNA条形码进行独特识别。

[0700] dCODE Dextramer (Rhapsody™)-Gold:每种25个测试(50μL)、50个测试(100μl)或150个测试(300μl)的单试剂。

[0701] dCODE Dextramer (Rhapsody™)-Explore:16种、32种、48种、64种、80种或96种dCODE Dextramer (Rhapsody™)试剂的组,每种用于10个测试(20μl)、25个测试(50μl)或50个测试(100μl)。

[0702] 试剂可以在2-8°C在黑暗中储存。

[0703] 另外的材料可以包括:细胞标记缓冲液:PBS, pH 7.4,含有1%-5%血清和0.1g/l鲑鱼精DNA(经剪切的);或BD™染色缓冲液(FBS)(目录号554656),添加了0.1g/l鲑鱼精DNA(经剪切的);洗涤缓冲液:PBS, pH 7.4,含有1%-5%血清;PBS中的100μM d-生物素;用于分选的鉴定相关细胞表面标志物(例如CD3、CD4、CD8)的抗体;用于分选活细胞的可固定存活力染色剂(Fixable Viability Stain);BD AbSeq Ab-Oligo;钙黄绿素AM(Thermo Fisher Scientific目录号1430);和/或INCYTO™一次性血细胞计数器(Cincyto目录号DHC-N01-5)。\*对于FACS分选,不使用与钙黄绿素AM具有相似光发射光谱的荧光团(FITC)和与Draq7™具有相似光发射光谱的荧光团(Cy5、Cy5-串联或APC),钙黄绿素AM和Draq7™用于通过Rhapsody™扫描仪评估存活力。

[0704] dCODE Dextramer特异性引物:dCODE PCR1引物(5'-GGAGGGAGGTTAGCGAAGGT-3'; SEQ ID NO:1)和dCODE PCR2引物(5'-CAGACGTGTGCTCTCCGATCTGGAGGGAGGTTAGCGAAGGT-3'; SEQ ID NO:2)。引物可以以10μM浓度使用。

[0705] BD方案:Doc ID:210966:Single Cell Capture and cDNA Synthesis with the BD Rhapsody™Single-Cell Analysis System(用BD Rhapsody™进行单细胞捕获和cDNA合成);Doc ID:214293:BD Rhapsody™Targeted mRNA and AbSeq Amplification Kit(使用BD Rhapsody™靶向mRNA和AbSeq扩增试剂盒制备mRNA靶向和BD™AbSeq文库)。

[0706] 制备单细胞PBMC样品

[0707] 注意,对于较大的dCODE Dextramer试剂池,应增加孵育时间。对于超过25种特异体,可以使用20-30min孵育。

[0708] (步骤1)在50μL细胞标记缓冲液中重悬 $1-3 \times 10^6$ 个PBMC。

[0709] 如果使用FVS,将细胞重悬在1000μl洗涤缓冲液(不含叠氮化物)中,并添加建议体

积的FVS。

[0710] 在RT孵育15min。(所有孵育必须避光进行)。添加2ml洗涤缓冲液,离心300-600g 5min,将细胞重悬在50μl细胞标记缓冲液中,并继续步骤2。

[0711] 制备dCODE Dextramer池并用dCODE Dextramer标记细胞

[0712] \*在用抗体染色之前,用dCODE Dextramer进行细胞标记。

[0713] (步骤2)将dCODE **Dextramer**<sup>®</sup> (Rhapsody<sup>™</sup>)以10,000×g离心1min。

[0714] (步骤3)向空流式管(**Falcon**<sup>®</sup>管,5mL圆底聚苯乙烯试管(Corning目录号352054))中添加0.2μl 100μM d-生物素/dCODE Dextramer特异体。

[0715] (步骤4)添加2μl每种dCODE Dextramer特异体并混合(用于500μL管的Latch Rack (Thermo Fisher Scientific目录号4890);8通道螺旋盖管加盖器(Thermo Fisher Scientific目录号4105MAT))。

[0716] (步骤5)将dCODE Dextramer试剂池添加至细胞样品中并混合。

[0717] (步骤6)在室温孵育10min(所有孵育必须避光进行)。(dCODE Dextramer池应立即使用,而不能存储)。在孵育时,准备2x AbSeq主混合物。

[0718] 准备2×BD AbSeq Ab-oligo标记主混合物

[0719] 该方案可以包括在每次实验前产生新鲜汇集的抗体。该方案可以包括产生具有30%过量的池,以确保用于标记的充足体积(试剂是黏性的,并容易形成气泡)。对于high-plex组,该方案可以包括使用8通道螺旋盖管加盖器(Thermo Fisher Scientific目录号4105MAT)和多通道移液器将BD AbSeq Ab-Oligo吸移到8管排管中。该方案基于使用HPBMC(外周血单个核细胞)。

[0720] (步骤7)将所有待汇集的BD AbSeq Ab-Oligo置于用于500μL管的Latch Rack中。排列管,使其可以容易地用8通道螺旋盖管加盖器开盖和重新加盖,并用多通道移液器进行等分。(用于500μL管的Latch Rack(Thermo Fisher Scientific目录号4890);8通道螺旋盖管加盖器(Thermo Fisher Scientific目录号4105MAT))。

[0721] (步骤8)在具有板转接管(plate adapter tubes)的台式离心机中,对Latch Rack中的BD AbSeq Ab-Oligo以400×g离心30秒,并放置在冰上。

[0722] (步骤9)在预扩增工作区中,如表6中示出的,将试剂吸移到在冰上的新的1.5mL LoBind管中。

[0723] 表6:混合物

| 组分   | (N=抗体数) 1个样品    | 1<N<100 加上 30%过量  |
|--|-----------------|-------------------|
| 每种 BD AbSeq Ab-Oligo                       | 2 μl            | 2 × (1+0.3) μl    |
| [0724] <b>BD 染色缓冲液(FBS)</b><br>目录号 554656) | 100μl - (2 × N) | 100μl - (2.6 × N) |
| <b>总计</b>                                  | 100 μl          | 130μl             |

[0725] (步骤10)将2X AbSeq标记主混合物吸移混合,并放回冰上。

[0726] AbSeq<sup>™</sup>标记

[0727] \*在用抗体染色之前,用dCODE Dextramer进行细胞标记。

[0728] (步骤11)使用标记缓冲液,使dCODE标记反应的体积到达100μl。

[0729] 标记缓冲液的体积:100μl-细胞体积-2×“dCODE特异体数目”(如果体积>100μl,

则继续下一步)。

[0730] (步骤12)将AbSeq 2×池添加到样品中,并吸移混合。

[0731] (步骤13)添加分选抗体缀合物。使用制造商建议的抗体体积,并进行吸移混合。

[0732] (步骤14)在4℃孵育30-60min。

[0733] 洗涤标记的细胞

[0734] (步骤15)如果在4ml Falcon管中进行染色,则添加2ml洗涤缓冲液。

[0735] (步骤16)以300-600x g离心5min,并去除上清液。

[0736] 总共重复3次。如果在96孔微量滴定板中进行标记,则每孔使用200μl洗涤缓冲液进行6次顺序洗涤。在每次洗涤之间以300-600x g离心5min,并去除上清液。

[0737] (步骤17)将细胞重悬在足够体积的洗涤缓冲液中,并将样品储存在冰上直至进行分选。如果不进行细胞分选,直接转到步骤20。

[0738] (步骤18)使用dCODE Dextramer的PE荧光按照您的分选设备的指南和操作对Dextramer阳性细胞进行FACS分选。(使用分选“产量”模式,因为如果使用“纯度”分选模式,您将损失更多阳性细胞)。

[0739] (步骤19)将经分选的细胞直接收集到含有适当体积的血清的管中。在进行分选时,将未经分选和经分选的细胞保持在4℃。(在分选管中使用最大体积的100%血清(FBS),仍有用于经分选的细胞的体积的空间;分选管中的FCS提高了经分选的细胞的存活力)。

[0740] (步骤20)以300-600g对经分选的细胞样品离心5-10min(取决于分选体积),倒转以将上清液倾析到生物危害性废物中。保持管倒置,用无绒拭纸(lint-free wiper)温和地吸干,以去除管边缘残留的上清液。立即进行步骤21。

[0741] 制备用于Rhapsody™细胞计数盒加载的细胞和cDNA合成

[0742] (步骤21)从开始就按照BD方案:“Single Cell Capture and cDNA Synthesis with the BD Rhapsody™Single-Cell Analysis System(用BD Rhapsody™单细胞分析系统进行单细胞捕获和cDNA合成)”Doc ID:210966。

[0743] 文库产生

[0744] (步骤22)使用下面的方案和方案mRNA Targeted and BD™AbSeq Library Preparation with the BD Rhapsody™Targeted mRNA and AbSeq Amplification Kit(使用BD Rhapsody™靶向mRNA和AbSeq扩增试剂盒制备mRNA靶向和BD™AbSeq文库)(Doc ID:214293)进行文库产生。

[0745] 测序要求

[0746] 对于dCODE文库的测序,按照“mRNA Targeted and BD™AbSeq Library Preparation with the BD Rhapsody™Targeted mRNA and AbSeq Amplification Kit”方案(Doc ID:214293)中关于AbSeq的要求和建议。

[0747] dCODE Dextramer® (Rhapsody™) 文库制备

[0748] 该方案描述了从cDNA捕获珠的文库制备。

[0749] 该方案提供了关于使用用于在Illumina测序仪上进行测序的BD Rhapsody™单细胞分析系统或BD Rhapsody™快速单细胞分析系统创建dCODE Dextramer®(有或没有AbSeq™和mRNA)单细胞文库的说明。

[0750] 为了创建文库,dCODE® Dextramer、BD AbSeq和BD Rhapsody mRNA靶被编码在

细胞捕获珠上,并且然后在PCR1中扩增。

[0751] 在PCR1之后,用Agencourt® AMPure® XP磁珠通过双侧尺寸选择将BD AbSeq和dCODE® PCR1产物与mRNA靶向PCR1产物分离。文库分子的尺寸选择是通过特定和连续使用DNA样品和AMPure珠之间的体积比来实现的。

[0752] mRNA和AbSeq文库的成功制备可以包括以下:(1)对dCODE®和BD Rhapsody mRNA靶向PCR1产物进行PCR2扩增;(2)用文库索引引物对dCODE、AbSeq和mRNA文库进行单独的索引PCR;和/或(3)在索引PCR之后,可以组合dCODE®、BD Rhapsody mRNA和BD AbSeq文库以进行测序。

[0753] 图5描绘了本文提供的组合物和方法的非限制性示例性示意工作流程。PCR1:通用寡核苷酸(黑色)、dCODE®(橙色)、ABSeq™(紫色)和mRNA靶(黄色)代表了PCR1中用于扩增捕获珠结合的cDNA的引物。PCR2:mRNA靶向文库和dCODE®文库的选择性扩增。PCR3:对每个文库使用不同的索引引物(与Illumina测序系统相容),能够对池中的文库进行测序。

[0754] 该方案可以包括以下材料:Agencourt AMPure XP磁珠(Beckman Coulter目录号A63880);无水乙醇,分子生物学级;不含核酸酶的水;用于1.5mL管的磁分离架;和/或Qubit™ dsDNA HS测定试剂盒(Thermo Fisher Scientific目录号Q32851)。该方案可以包括BD Rhapsody™靶向mRNA和AbSeq扩增试剂盒(目录号633774)。

[0755] dCODE Dextramer® (Rhapsody™)可以与以下BD Rhapsody™工作流程组合:mRNA Targeted and BD™AbSeq Library Preparation with the BD Rhapsody™Targeted mRNA and AbSeq Amplification Kit (Doc ID:214293)(工作流程的概述在图5中示出)。dCODE Dextramer®可以与BD™AbSeq文库制备(有或没有AbSeq™抗体)组合使用。

[0756] 对捕获的dCODE®、AbSeq寡核苷酸和靶向的mRNA进行PCR1扩增(步骤1)在预扩增工作区中,将试剂吸移到在冰上的新的1.5mL LoBind管中,使用表7(dCODE®,AbSeq™和靶向的mRNA)或表8(仅dCODE®和靶向的mRNA)中示出的反应混合物。(如果扩增多于一个样品,则使用20%过量)。

[0757] 表7:PCR1反应混合物,用于dCODE®,ABSeq™和靶mRNA

| A: 试剂盒组分名称                        | 1x (μL) | 2 x (+20%) |
|-----------------------------------|---------|------------|
| 不含核酸酶的水 <sup>a)</sup>             | 4.0     | 9.6        |
| PCR 主混合物 <sup>a)</sup>            | 100.0   | 240.0      |
| PCR1 引物组(mRNA 靶向引物) <sup>b)</sup> | 40.0    | 96.0       |
| AbSeq PCR1 引物 <sup>a)</sup>       | 12.0    | 28.8       |
| dCODE 引物 2 (10μM) <sup>c)</sup>   | 12.0    | 28.8       |
| 通用寡核苷酸 <sup>a)</sup>              | 20.0    | 48.0       |
| 珠 RT/PCR 增强剂 <sup>a)</sup>        | 12.0    | 28.8       |

[0758]

| 总体积    | 200   | 480 |
|--------|---|-----|
| [0759] | a) “BD Rhapsody™ 靶向 mRNA 和 AbSeq 扩增试剂盒(目录号 633774)” 中提供的试剂。<br>b) BD Biosciences 引物<br>c) 参见方案末尾的引物 |     |

[0760] 表8:PCR1反应混合物,仅用于dCODE®和靶mRNA

| B: 试剂盒组分名称                        | 1x (μL)   | 2 x (+20%) |
|-----------------------------------|---|------------|
| 不含核酸酶的水 <sup>a)</sup>             | 16.0  | 38.4       |
| PCR 主混合物 <sup>a)</sup>            | 100.0   | 240.0      |
| PCR1 引物组(mRNA 靶向引物) <sup>b)</sup> | 40.0  | 96.0       |
| dCODE 引物 2 (10μM) <sup>c)</sup>   | 12.0  | 28.8       |
| 通用寡核苷酸 <sup>a)</sup>              | 20.0  | 48.0       |
| 珠 RT/PCR 增强剂 <sup>a)</sup>        | 12.0  | 28.8       |
| <b>总体积</b>                        | <b>200</b>  | <b>480</b> |
| [0761]                            | a) “BD Rhapsody™ 靶向 mRNA 和 AbSeq 扩增试剂盒(目录号 633774)” 中提供的试剂。<br>b) BD Biosciences 引物<br>c) 参见方案末尾的引物 |            |

[0762] (步骤2)将混合物温和地涡旋,短暂离心,并放回冰上。

[0763] (步骤3)如下进行:完整样品,跳到步骤5;子样品,继续步骤4。

[0764] (步骤4)对外切核酸酶I处理的珠进行次级采样:(a)基于由BD™ Rhapsody扫描仪检测到的含有活细胞和珠的孔的数目或盒中靶向以供捕获的细胞的数目,确定用于靶向测序的次级采样的珠的体积;(b)将珠吸移混合以完全重悬,并将计算体积的珠悬浮液吸移到新的1.5mL LoBind管中。剩余的珠可在2℃至8℃储存≤3个月。

[0765] (步骤5)将珠重悬缓冲液(目录号650000066)中外切核酸酶I处理的珠的管放置到1.5mL磁体上持续<2分钟。去除上清液。

[0766] (步骤6)将管从磁体上取下,并将珠重悬在200μL PCR1反应混合物中。不进行涡旋。

[0767] (步骤7)确保珠完全重悬,将50μL PCR1反应混合物与珠吸移到四个0.2mL PCR管的每一个中。将任何残留混合物转移到其中一个管中。

[0768] (步骤8)使反应混合物到达扩增后工作区。

[0769] (步骤9)如表9中示出的对热循环仪进行编程。不使用快速循环模式。

[0770] 表9:PCR程序

| 步骤   | 循环                 | 温度                | 时间     |
|------|--------------------|-------------------|--------|
| 热启动  | 1                  | 95°C <sup>a</sup> | 3 min  |
| 变性   | 11-15 <sup>b</sup> | 95°C              | 30 sec |
| 退火   |                    | 60°C              | 3 min  |
| 延伸   |                    | 72°C              | 1 min  |
| 最终延伸 | 1                  | 72°C              | 5 min  |
| 保持   | 1                  | 4°C               | ∞      |

[0771] <sup>a</sup> 为了避免在变性步骤之前，由于热循环仪上的孵育时间延长而导致珠沉降，在加载样品之前，将仪器暂停在 95°C。不同的热循环仪可能有不同的暂停时间设置。然而，在某些品牌的热循环仪中，BD Biosciences 观察到暂停/取消暂停功能存在步骤跳过错误。为了确保不跳过完整的三分钟的变性，请验证您的热循环仪上的暂停/取消暂停功能是否正常工作。为了避免步骤跳过问题，可以在紧接着三分钟的 95°C 变性步骤之前添加一分钟 95°C 暂停步骤。

<sup>b</sup> 建议的 PCR 循环可能需要针对不同的细胞类型、BD™ AbSeq 组中的抗体数目和细胞数目进行优化，如表 10 中示出的。

[0772] 表10:PCR循环

| PCR1 中的细胞数 | 建议的用于静息 PBMC 的 PCR 循环 |
|------------|-----------------------|
| 500        | 15                    |
| 1,000      | 14                    |
| 2,500      | 13                    |
| 5,000      | 12                    |
| 10,000     | 11                    |

[0774] (步骤10) 通过启动热循环仪程序，将扩增后热循环仪的盖和加热块升温 (Ramp heated) 至  $\geq 95^{\circ}\text{C}$ ，并且然后暂停。在用移液器温和地混合每支管以确保使珠悬浮均匀之后，才进行热循环。

[0775] (步骤11) 对于每支 0.2mL PCR 管，温和地吸移混合，立即将管放入热循环仪并取消暂停热循环仪程序。终止点：PCR 可以过夜运行，但在 PCR 后  $\leq 24$  小时进行纯化。

[0776] (步骤12) 在 PCR 后，将管短暂离心。

[0777] (步骤13) 将四种反应混合物吸移混合并合并到新的 1.5mL LoBind 管中。在下一步中保留上清液。

[0778] (步骤14) 将 1.5mL 管放置在磁体上持续 2 分钟，并且在不扰动珠的情况下仔细地将上清液 (mRNA 靶向 PCR1 产物和 BD AbSeq PCR1 产物) 吸移到新的 1.5mL LoBind 管中。注：(任选的) 从磁体上取下含有细胞捕获珠的管，并将 200  $\mu\text{L}$  冷的珠重悬缓冲液 (目录号 650000066) 吸移到管中。吸移混合。不进行涡旋。在扩增后工作区中，将珠储存在  $2^{\circ}\text{C}$  至  $8^{\circ}\text{C}$ 。

[0779] 通过双侧尺寸选择纯化 PCR1 产物

[0780] 进行双侧 AMPure 珠纯化，以将较短的 BD AbSeq PCR1 产物 (~170bp) 与较长的 mRNA 靶向 PCR1 产物 (350-800bp) 分离。在该方案中，保留上清液 (dCODE® 和 BD AbSeq 产物) 和 AMPure 珠 (mRNA 靶向产物) 两者进行纯化。在扩增后工作区中进行纯化。

[0781] 分离 BD AbSeq PCR1 产物与 mRNA 靶向 PCR1 产物

[0782] (步骤1) 在新 5.0mL LoBind 管中，通过将 4.0mL 分子生物学级无水乙醇 (主要供应商) 与 1.0mL 无核酸酶的水 (主要供应商) 混合，制备 5mL 新鲜的 80% (v/v) 乙醇。将管涡旋 10

秒以进行混合。制备新鲜的80%乙醇,并在 $\leq 24$ 小时内使用。

[0783] (步骤2)使Agencourt AMPure XP磁珠(Beckman Coulter目录号15°C至25°C)达到室温(15°C至25°C)。高速涡旋1分钟,直到珠完全重悬。

[0784] (步骤3)将140 $\mu$ L AMPure珠与200 $\mu$ L mRNA靶向PCR1产物和BD AbSeq PCR1产物一起吸移到管中(进行PCR1的步骤14)。吸移混合10次。

[0785] (步骤4)在室温(15°C至25°C)孵育5分钟。

[0786] (步骤5)将1.5mL LoBind管放置在磁体上5分钟。

[0787] (步骤6)将管保持在磁体上,在不扰动珠(mRNA靶向PCR1产物)的情况下将340 $\mu$ L上清液(BD AbSeq PCR1产物)转移到新的1.5mL管中。

[0788] (步骤7)将上清液(步骤6)储存在室温(15°C至25°C),同时按“纯化mRNA靶向PCR1产物”纯化并洗脱mRNA靶向PCR1产物。按“纯化BD AbSeq PCR1产物”纯化BD AbSeq PCR1产物。

[0789] 纯化mRNA靶向PCR1产物

[0790] (步骤1)将管保持在磁体上,向与mRNA靶向PCR1产物结合的AMPure珠的管中温和地添加500 $\mu$ L新鲜的80%乙醇,并孵育30秒。去除上清液。

[0791] (步骤2)重复步骤1一次,共进行两次洗涤。

[0792] (步骤3)将管保持在磁体上,使用小体积移液器从管中取出并弃去残留上清液。

[0793] (步骤4)使珠在室温(15°C至25°C)空气干燥5分钟。

[0794] (步骤5)将管从磁体上取下,并将珠沉淀物重悬在30 $\mu$ L洗脱缓冲液(目录号91-1084)中。剧烈吸移混合,直到珠均匀分散。小团块不影响性能。

[0795] (步骤6)在室温(15°C至25°C)孵育2分钟,并短暂离心。

[0796] (步骤7)将管放置在磁体上,直到溶液澄清,通常 $\leq 30$ 秒。

[0797] (步骤8)将洗脱液( $\sim 30\mu$ L)吸移到新的1.5mL LoBind管中(纯化的mRNA靶向PCR1产物)。

[0798] 终止点:在于 $\leq 24$ 小时内继续之前在2°C至8°C储存,或在-25°C至-15°C储存 $\leq 6$ 个月。

[0799] 纯化BD AbSeq PCR1产物

[0800] (步骤1)将100 $\mu$ L AMPure XP珠吸移到来自“分离BD AbSeq PCR1产物与mRNA靶向PCR1产物”的步骤6的含有340 $\mu$ L BD AbSeq PCR1产物的管中。吸移混合10次。

[0801] (步骤2)在室温(15°C至25°C)孵育5分钟。

[0802] (步骤3)放置在磁体上5分钟。

[0803] (步骤4)将管保持在磁体上,取出并弃去上清液。

[0804] (步骤5)将管保持在磁体上,温和地添加500 $\mu$ L新鲜的80%乙醇,并孵育30秒。去除上清液。

[0805] (步骤6)重复步骤5一次,共进行两次洗涤。

[0806] (步骤7)将管保持在磁体上,使用小体积移液器从管中取出并弃去残留上清液。

[0807] (步骤8)使珠在室温(15°C至25°C)空气干燥5分钟。

[0808] (步骤9)将管从磁体上取下,并将珠沉淀物重悬在30 $\mu$ L洗脱缓冲液(目录号91-1084)中。剧烈吸移混合,直到珠均匀分散。小团块不影响性能。

- [0809] (步骤10) 在室温(15℃至25℃) 孵育2分钟, 并短暂离心。
- [0810] (步骤11) 将管放置在磁体上, 直到溶液澄清, 通常≤30秒。
- [0811] (步骤12) 将洗脱液(~30μL) 吸移到新的1.5mL LoBind管中(纯化的BD AbSeq PCR1产物)。
- [0812] 终止点: 在于≤24小时内继续之前在2℃至8℃储存, 或在-25℃至-15℃储存≤6个月。
- [0813] 注: dCODE®文库与来自PCR1的AbSeq文库共纯化。
- [0814] 对BD AbSeq PCR1产物进行定量
- [0815] (步骤1) 通过使用高灵敏度试剂盒(Agilent目录号5067-4626) 用Agilent生物分析仪测量BD AbSeq PCR1产物的最大峰(~170bp) 的产量。按照制造商的说明。
- [0816] (步骤2) 用洗脱缓冲液(目录号91-1084) 将BD AbSeq产物的等分试样稀释至0.1-1.1ng/μL, 然后对BD AbSeq PCR1产物进行索引PCR。
- [0817] 图6描绘了包含dCODE®和AbSeq™文库PCR扩增子两者(约170bp) 的非限制性示例性生物分析仪迹线。
- [0818] 对mRNA靶向PCR1产物进行PCR2
- [0819] (步骤1) 在预扩增工作区中, 将试剂吸移到在冰上的新的1.5mL LoBind管中, 以产生表11中示出的反应混合物。
- [0820] 在使用BD Rhapsody™ 10X PCR2定制引物和/或BD Rhapsody™ 10X PCR2补充引物之前, 将1份10X PCR引物储备液稀释到9份IDTE缓冲液中, 以制备1X引物溶液。BD Rhapsody™靶向(预先设计) 引物组以1X浓度提供, 并且不应稀释。
- [0821] 表11: PCR2反应混合物

| 组分                       | 对于 1 个文库(μL) | 对于 1 个文库+20%过量(μL) |
|--------------------------|--------------|--------------------|
| PCR 主混合物(目录号 91-1083)    | 25.0         | 30.0               |
| 通用寡核苷酸(目录号 650000074)    | 2.0          | 2.4                |
| PCR2 引物组(BD Biosciences) | 10.0         | 12.0               |
| (任选的) PCR2 组补充物          | (2.5)        | (3.0)              |
| 不含核酸酶的水(目录号 650000076)   | 至多 8.0       | 至多 9.6             |
| <b>总计</b>                | <b>45.0</b>  | <b>54.0</b>        |

- [0822] [0823] (步骤2) 将混合物温和地涡旋, 短暂离心, 并放回冰上。
- [0824] (步骤3) 使PCR2混合物到达扩增后工作区。
- [0825] (步骤4) 在新的0.2mL PCR管中, 将5.0μL纯化的mRNA靶向PCR1产物吸移到45μL PCR2反应混合物中。
- [0826] (步骤5) 温和地涡旋, 并短暂离心。
- [0827] (步骤6) 如表12中示出的对热循环仪进行编程。不使用快速循环模式。
- [0828] 表12: PCR程序

| 步骤                                | 循环              | 温度  | 时间     |
|-----------------------------------|-----------------|-----|--------|
| 热启动                               | 1               | 95℃ | 3 min  |
| 变性                                | 10 <sup>a</sup> | 95℃ | 30 sec |
| 退火                                |                 | 60℃ | 3 min  |
| 延伸                                |                 | 72℃ | 1 min  |
| 最终延伸                              | 1               | 72℃ | 5 min  |
| 保持                                | 1               | 4℃  | ∞      |
| <sup>a</sup> 循环数可能需要根据细胞数和类型进行优化。 |                 |     |        |

[0830] 终止点:PCR可以过夜运行。

[0831] 纯化mRNA靶向PCR2产物

[0832] 在扩增后工作区进行纯化。

[0833] (步骤1)使AMPure XP珠达到室温(15℃至25℃),并以高速涡旋1分钟,直到珠完全重悬。

[0834] (步骤2)对mRNA靶向PCR2产物进行短暂离心。

[0835] (步骤3)将40μL AMPure XP珠吸移到含有50μL mRNA靶向PCR2产物的管中。吸移混合10次。

[0836] (步骤4)在室温(15℃至25℃)孵育5分钟。

[0837] (步骤5)将管放置在排管磁体上3分钟。去除上清液。

[0838] (步骤6)将管保持在磁体上,温和地向管中添加200μL新鲜的80%乙醇,并孵育30秒。去除上清液。

[0839] (步骤7)重复步骤6一次,共进行两次洗涤。

[0840] (步骤8)将管保持在磁体上,使用小体积移液器从管中取出并弃去残留上清液。

[0841] (步骤9)使珠在室温(15℃至25℃)空气干燥3分钟。

[0842] (步骤10)将管从磁体上取下,并将珠沉淀物重悬在30μL洗脱缓冲液(目录号91-1084)中。吸移混合,直到珠完全重悬。

[0843] (步骤11)在室温(15℃至25℃)孵育2分钟,并短暂离心。

[0844] (步骤12)将管放置在磁体上,直到溶液澄清,通常≤30秒。

[0845] (步骤13)将全部洗脱液(~30μL)吸移到新的1.5mL LoBind管中(纯化的mRNA靶向PCR2产物)。

[0846] 终止点:在于当天继续之前在2℃至8℃储存,或在-25℃至-15℃储存≤6个月。

[0847] (步骤14)通过使用Qubit dsDNA HS测定试剂盒用Qubit™荧光计对2μL mRNA靶向PCR2产物进行定量来估计浓度。按照制造商的说明。

[0848] (步骤15)用洗脱缓冲液(目录号91-1084)将mRNA靶向PCR2产物的等分试样稀释至0.2-2.7ng/μL。

[0849] 对纯化的PCR1产物进行PCR2

[0850] dCODE和(任选的)AbSeq纯化的PCR1产物用于使用独特dCODE文库引物(下文列出)扩增dCODE®文库。

[0851] (步骤1)在预扩增工作区中,将试剂吸移到在冰上的新的1.5mL LoBind管中,以产生表13中示出的反应混合物。(如果超过一个样品,则使反应混合物上有20%过量)。

[0852] 表13:dCODE特异性PCR2反应混合物

| 试剂盒组分名称 <sup>a)</sup>  | $\mu\text{l}$ 1x | 2x +20%             |
|--|------------------|---------------------|
| 不含核酸酶的水  | 8                | 19.2                |
| PCR 主混合物   | 25               | 60.0                |
| 通用寡核苷酸   | 2                | 4.8                 |
| dCODE 特异性引物 2  | 10               | 24.0                |
| <b>总计</b>  | 45               | 108                 |
| PCR1 纯化的 AbSeq 文库(未稀释)   | 5                | 5 $\mu\text{l}$ /样品 |
| a)用于 PCR2 扩增的试剂包含在“BD Rhapsody™ 靶向 mRNA 和 AbSeq 扩增试剂盒(目录号 633774)”中。 |                  |                     |

[0854] (步骤2)将混合物温和地涡旋,短暂离心,并放回冰上。

[0855] (步骤3)使PCR2混合物到达扩增后工作区。

[0856] (步骤4)在两支单独的新的0.2mL PCR管中:(i)mRNA靶向PCR1产物:将5.0 $\mu\text{L}$ 产物吸移到45.0 $\mu\text{L}$  mRNA靶向PCR2反应混合物中;(ii)样品标签PCR1产物:将5.0 $\mu\text{L}$ 产物吸移到45.0 $\mu\text{L}$ 样品标签PCR2反应混合物中。

[0857] (步骤5)温和地涡旋,并短暂离心。

[0858] (步骤6)对于dCODE文库PCR2,如表14中示出的对热循环仪进行编程。(不使用快速循环模式)

[0859] 表14:dCODE文库PCR2

| 步骤                   | 循环  | 温度   | 时间       |
|----------------------|-----|------|----------|
| 热启动                  | 1   | 95°C | 3 min    |
| 变性                   | 10* | 95°C | 30 sec   |
| 退火                   |     | 66°C | 30 sec   |
| 延伸                   |     | 72°C | 1 min    |
| 最终延伸                 | 1   | 72°C | 5 min    |
| 保持                   | 1   | 4°C  | $\infty$ |
| *循环数可能需要根据预计的细胞数进行优化 |     |      |          |

[0862] 纯化dCODE文库

[0863] (步骤1)使AMPure XP珠达到室温(15°C至25°C),并以高速涡旋1分钟,直到珠完全重悬。

[0864] (步骤2)对PCR2产物进行短暂离心。

[0865] (步骤3)将60 $\mu\text{L}$  AMPure珠吸移到50.0 $\mu\text{L}$  PCR2产物中。

[0866] (步骤4)吸移混合10次,并在室温(15°C至25°C)孵育5分钟。

[0867] (步骤5)将管放置在排管磁体上3分钟。去除上清液。

[0868] (步骤6)将管保持在磁体上,温和地向管中添加200 $\mu\text{L}$ 新鲜的80%乙醇,并孵育30秒。去除上清液。

[0869] (步骤7)重复步骤6一次,共进行两次洗涤。

[0870] (步骤8)将管保持在磁体上,使用小体积移液器从管中取出并弃去残留上清液。

- [0871] (步骤9)使珠在室温(15℃至25℃)空气干燥3分钟。
- [0872] (步骤10)将管从磁体上取下,并将珠沉淀物重悬在30μL洗脱缓冲液(目录号91-1084)中。吸移混合,直到珠完全重悬。
- [0873] (步骤11)在室温(15℃至25℃)孵育2分钟,并短暂离心。
- [0874] (步骤12)将管放置在磁体上,直到溶液澄清,通常≤30秒。
- [0875] (步骤13)将样品的全部洗脱液(~30μL)吸移到新的1.5mL LoBind管中(纯化的dCODE PCR2产物)。
- [0876] (步骤14)通过使用Qubit dsDNA HS测定试剂盒用Qubit™荧光计对2μL PCR2产物进行定量来估计样品的浓度。
- [0877] 终止点:在于当天继续之前在2℃至8℃储存,或在-25℃至-15℃储存≤6个月。
- [0878] 对dCODE®库进行定量
- [0879] 通过使用高灵敏度试剂盒(Agilent目录号5067-4626)用Agilent生物分析仪测量dCODE®PCR2产物(扩增子应为约190bp)的产量。按照制造商的说明。
- [0880] 图7描绘了包含dCODE PCR2产物(约190bp)的非限制性示例性生物分析仪迹线。
- [0881] 进行索引PCR以制备最终文库
- [0882] 使用不同的反向索引引物(在“BD Rhapsody™靶向mRNA和AbSeq扩增试剂盒(目录号633774)”中提供),对每种制备的纯化PCR1产物(AbSeq™文库(任选的)和PCR2产物(dCODE和mRNA)进行索引PCR3。
- [0883] (步骤1)在预扩增工作区,针对每种产物制备1个文库+20%过量的最终扩增混合物(表15)。将试剂吸移到在冰上的新的1.5mL LoBind管中。
- [0884] 对于单个盒或样品,考虑对该盒或样品的所有文库使用相同的索引。如果文库将被不同地索引,则针对每种文库类型制备包含不同文库反向引物的单独索引PCR混合物。
- [0885] 表15:索引PCR混合物

| 组分                                       | 对于 1 个文库(μL) | 对于 1 个文库+20%过量(μL) |
|--|--------------|--------------------|
| PCR 主混合物(目录号 91-1083)                    | 25.0         | 30.0               |
| 文库正向引物(目录号 91-1085)                      | 2.0          | 2.4                |
| 文库反向引物 1-4 (目录号 650000080, 650000091-93) | 2.0          | 2.4                |
| 不含核酸酶的水(目录号 650000076)                   | 18.0         | 21.6               |
| <b>总计</b>                                | <b>47.0</b>  | <b>56.4</b>        |

- [0886] [0887] (步骤2)将混合物温和地涡旋,短暂离心,并放回冰上。
- [0888] (步骤3)使索引PCR混合物到达扩增后工作区。
- [0889] (步骤4)在两支单独的新的0.2mL PCR管中:(a)mRNA靶向PCR2产物:将3.0μL 0.2-2.7ng/μL产物吸移到47.0μL索引PCR混合物中;(b)BD AbSeq PCR1产物:将3.0μL 0.1-1.1ng/μL产物吸移到47.0μL索引PCR混合物中。对dCODE®PCR2纯化产物的Index PCR3也按照步骤4中对AbSeq™文库所描述的进行——使用3.0μL 0.1-1.1ng/μLdCODE®PCR2产物加入47.0μL index PCR3混合物中。
- [0890] (步骤5)温和地涡旋,并短暂离心。

[0891] (步骤6)如表16中示出的对热循环仪进行编程。不使用快速循环模式。

[0892] 表16:PCR

| 步骤   | 循环               | 温度   | 时间     |
|------|------------------|------|--------|
| 热启动  | 1                | 95°C | 3 min  |
| 变性   | 6-8 <sup>a</sup> | 95°C | 30 sec |
| 退火   |                  | 60°C | 30 sec |
| 延伸   |                  | 72°C | 30 sec |
| 最终延伸 | 1                | 72°C | 1 min  |
| 保持   | 1                | 4°C  | ∞      |

<sup>a</sup> 建议的 PCR 循环如表 17 中示出的。

[0894] 表17:PCR循环数

| mRNA 靶向文库的索引 PCR 输入浓度(ng/μL) | BD AbSeq 文库的索引 PCR 输入浓度(ng/μL) | 建议的 PCR 循环 |
|------------------------------|--------------------------------|------------|
| 1.2-2.7                      | 0.5-1.1                        | 6          |
| 0.6-1.2                      | 0.25-0.5                       | 7          |
| 0.2-0.6                      | 0.1-0.25                       | 8          |

[0896] 终止点:PCR可以过夜运行。

[0897] dCODE®最终索引文库的定量与质量控制

[0898] 通过使用高灵敏度试剂盒(Agilent目录号5067-4626)用Agilent生物分析仪测量索引PCR3产物(约290bp)的产量。按照制造商的说明。图8描绘了包含索引PCR3 dCODE文库产物(约285bp)的非限制性示例性生物分析仪迹线。

[0899] 测序要求

[0900] 对于dCODE文库的测序,按照“mRNA Targeted and BD<sup>TM</sup>AbSeq Library Preparation with the BD Rhapsody<sup>TM</sup>Targeted mRNA and AbSeq Amplification Kit”方案(Doc ID:214293)中关于AbSeq<sup>TM</sup>的要求和建议。

[0901] BD Rhapsody<sup>TM</sup>方案

[0902] mRNA Targeted and BD<sup>TM</sup>AbSeq Library Preparation with the BD Rhapsody<sup>TM</sup>Targeted mRNA and AbSeq Amplification Kit(Doc ID:214293)。

[0903] 用于dCODE<sup>TM</sup>Dextramer®特异性文库制备的扩增引物

[0904] dCODE PCR1引物(5'-GGAGGGAGGTTAGCGAAGGT-3';SEQ ID NO:1)和dCODE PCR2引物(5'-CAGACGTGTGCTCTCCGATCTGGAGGGAGGTTAGCGAAGGT-3';SEQ ID NO:2)。引物可以以10 μM浓度使用。

[0905] 实施例2

[0906] 在BD Rhapsody单细胞分析系统上利用灵敏的dCODE Dextramer和BD AbSeq Ab-Oligo在单细胞水平对T细胞群体进行多组学表征

[0907] 过继性转移的抗原特异性T细胞在治疗一些病毒相关疾病和恶性肿瘤中显示出巨大的效力。过继性T细胞疗法发展的主要驱动力是成功表征T细胞功能状态和抗原特异性的能力。然而,这受到抗原特异性T细胞检测效率低下的限制,可能是由于抗原特异性T细胞的低频率和对已知MHC-肽复合物的低结合亲和力。本实施例展示了组合两种强大技术(即来

自Immudex的先进的dCODE™ **Dextramer**®和使用BD Rhapsody™单细胞分析系统的单细胞多组学分析)的组合,使用本文提供的组合物和方法,来检测和表征数千个PBMC中的疾病特异性CD8+T细胞。所公开的组合物和方法可以鉴定超过350种mRNA以及一组超过20种**BD**® AbSeq细胞表面蛋白标志物,这些标志物可能与T细胞活化状态相关。这些数据可以用于定义来自PBMC群体的富集的CD8+Dextramer+细胞的T细胞表型以及抗原特异性。本实施例概述了所公开的组合物和方法用于高分辨率T细胞谱分析的效用,其在免疫肿瘤学、感染性疾病和自身免疫中具有较广泛的意义和效用。采用与实施例1中提供的方案类似的方案。

#### [0908] 方法

[0909] 图9A-图9F描绘了Immudex dCODE和BD Rhapsody系统单细胞测序的非限制性示例性实验工作流程和分子机制。图9A描绘了与BD Rhapsody系统相容的非限制性示例性Immudex dCODE Dextramer设计。它们包含与特定抗原表位关联的可以与靶向mRNA同时扩增和测序的dCODE特异性条形码。图9B描绘了使用dCODE Dextramer的非限制性示例性BD Rhapsody系统靶向工作流程。图9C描绘了从盒回收的具有杂交的mRNA和dCODE的非限制性示例性珠。首先用一组dCODE Dextramer对细胞进行染色,然后在BD Rhapsody系统上捕获单细胞,该系统在细胞裂解前使用微孔分离个体细胞。在裂解后,来自mRNA和dCODE Dextramer的多腺苷酸化序列(如图9C中示出的)被捕获在珠上。完成cDNA和用于测序的文库制备,然后使用SeqGeq™软件进行数据分析。图9D-图9F描绘了非限制性示例性实验设计:用3种Dextramer染色的人类PBMC。在本研究中,用两种肽抗原EBV和破伤风类毒素(Tetanus toxoid, TT)刺激hPBMC。在刺激之后,在BD Rhapsody系统上捕获之前,用dCODE EBV特异性和TT特异性Dextramer以及阴性对照Dextramer同时对细胞进行染色。使用BD Rhapsody系统以及常规FACS检查Dextramer特异体群体,其中用相同的dCODE Dextramer试剂独立标记细胞。dCODE **Dextramer**®特异体为DRB1\*0101,具有以下抗原肽:EBV/TSLYNLRRGTALA (SEQ ID NO:4), TT/KIYSYFSPVISKV (SEQ ID NO:5) 阴性对照Clip/PVSKMRMATPLLMQA (SEQ ID NO:6)。门控抗体CD3/APC和CD4/BV421用于将MHC II类抗原特异性T细胞限定为CD3+、CD4+和dCODE (PE) 阳性。

#### [0910] 结果

[0911] Dextramer可以在不同于阴性对照的特定细胞群体上检测到

[0912] 图10A-图10B示出了检测不同于阴性对照的特定细胞群体上的Dextramer相关的非限制性示例性数据。来自使用dCODE **Dextramer**®鉴定hPBMC样品中的抗原特异性T细胞。在来自使用6,000个细胞(图10A)和30,000个细胞(图10B)的两个独立实验的tSNE图中示出的EBV和TT dCODE **Dextramer**®检测。Neg-dex代表来自阴性对照的非特异性结合,其可以与EBV和TT dCODE **Dextramer**®两者区分开。

[0913] 差异性基因表达分析显示出dCODE **Dextramer**®阳性T细胞中的不同的转录谱

[0914] 图11A-图11B示出了与EBV+抗原特异性T细胞中基因表达分析相关的非限制性示例性数据。可以鉴定Dextramer EBV+群体与EBV-阴性群体之间的差异性基因表达。来自上调基因和下调基因的分子的总和能够分别在TriMap空间的相对坐标中可见,表明这些细胞具有不同的转录谱(表18-表19)。

[0915] 表18:上调基因

[0916]

| 特征            | FC   | q-Val    | 平均值测试  | 平均值对照 |
|---------------|------|----------|--------|-------|
| ADEX5001 (Ab) | 5.46 | 1.00E-17 | 435.06 | 78.88 |
| CD27          | 3.43 | 1.24E-09 | 13.24  | 3.16  |
| GZMK          | 2.76 | 1.00E-17 | 109.78 | 39.08 |
| FYB           | 1.93 | 7.81E-13 | 34.83  | 17.57 |
| VNN2          | 1.87 | 3.72E-03 | 9.96   | 4.86  |
| CD244         | 1.83 | 3.82E-02 | 4.59   | 2.05  |
| KLRK1         | 1.64 | 2.26E-05 | 12.24  | 7.09  |
| FYN           | 1.56 | 1.00E-07 | 41.43  | 26.24 |
| CXCR4         | 1.49 | 6.19E-09 | 72.09  | 48.18 |
| CD6           | 1.46 | 4.27E-06 | 18.44  | 12.28 |

[0917]

表19: 下调基因

| 特征          | FC   | q-Val    | 平均值测试 | 平均值对照  |
|-------------|------|----------|-------|--------|
| TARP_refseq | 0.28 | 8.56E-05 | 4.91  | 19.98  |
| CCR3        | 0.33 | 1.20E-03 | 1.48  | 6.49   |
| CCR2        | 0.34 | 1.64E-11 | 9.66  | 30.56  |
| CD8B        | 0.35 | 2.25E-04 | 0.99  | 4.71   |
| IL13        | 0.45 | 1.67E-02 | 4.70  | 11.79  |
| IFITM3      | 0.47 | 2.31E-12 | 11.12 | 24.76  |
| LAT2        | 0.47 | 6.03E-05 | 2.67  | 6.75   |
| CD44        | 0.49 | 1.78E-15 | 23.67 | 49.48  |
| MZB1        | 0.50 | 5.44E-06 | 4.12  | 9.28   |
| GNLY        | 0.53 | 1.29E-03 | 54.51 | 103.01 |
| HAVCR2      | 0.54 | 2.36E-09 | 11.21 | 21.76  |
| ENTPD1      | 0.55 | 4.72E-08 | 11.62 | 21.98  |
| CD33        | 0.57 | 1.84E-06 | 6.97  | 12.96  |
| CD4         | 0.57 | 1.25E-09 | 15.18 | 27.29  |
| IFITM2      | 0.58 | 2.09E-13 | 25.20 | 44.44  |
| IL2RA       | 0.62 | 5.75E-07 | 20.00 | 33.08  |

[0918]

[0919]

[0920] 结论

[0921] dCODEDextramer®试剂和BD AbSeq抗体可以用于使用BD Rhapsody系统检测供体样品中的多种抗原特异性T细胞群体。BD靶向mRNA和AbSeq抗体组能够实现对这些供体样品中的疾病特异性T细胞的深入表征。dCODEDextramer®试剂也可以与BD Rhapsody CDR 3VDJ单细胞方案一起使用来提供检测到的疾病特异性T细胞受体的可变区的序列。

[0922] dCODE Dextramer与BD Rhapsody系统相容,并且可以与靶向mRNA组一起用于鉴定抗原特异性细胞。所有用于染色细胞的dCODE Dextramer都在测序实验中检测到,并通过常规FACS进一步验证。因此,本文提供的组合物和方法即使在抗原特异性T细胞频率较低的情况下也能够鉴定具有阳性而非阴性Dextramer表达的不同细胞群体。

[0923] 实施例3

[0924] dCODE Dextramer和BD AbSeq Ab-Oligo文库制备

[0925] 本实施例概述了所公开的组合物和方法用于高分辨率T细胞谱分析的效用。采用与实施例1中提供的方案类似的方案。图12A-图12D描绘了非限制性示例性生物分析仪迹线。图12A描绘了Ri0dCODE®和AbSeq™文库PCR1扩增子两者(峰166bp)的生物分析仪迹线。在PCR1后,将文库分开。图12B描绘了Ri0dCODE®文库PCR2扩增子(峰约190bp)的生物分析仪迹线,并展示了在使用对dCODE特异性的引物的第二PCR(PCR2)后仅Ri0 dCODE Dextramer的单独扩增。单独的dCODE文库的可视化可以用作测序前的QC指标。图12C-图12D描绘了Abseq(图12C;峰265bp)和Ri0 dCODEDextramer(图12D;峰约290bp)的最终文库(索引PCR)制备物的生物分析仪迹线。由于它们现在被独立扩增,这允许用户按照它们的建议测序深度如本文描述的对这些文库进行测序。

[0926] 组合的dCODE和AbSeq文库可能需要深度测序来检测dCODE分子。图13A-图13B描绘了dCODEDextramer®(图13A)和BD® AbSeq(图13B;4-plex,高表达者)的非限制性示例性饱和曲线。饱和曲线可以取决于所使用的AbSeq和dCODE组的类型。如在图13A-图13B中观察到的,在每个细胞2000个读段的测序深度,对于dCode,预计每个细胞约400个分子被检测,但是对于AbSeq,预计每个细胞约2200个分子被检测。因此,如果这些文库没有被分开且它们必须一起测序,则用户将不得不对文库进行非常深的测序才能够检测到dCODE分子。因此,如本文描述地分离文库可能是该问题的最佳解决方案。

[0927] 术语

[0928] 在至少一些先前描述的实施方案中,一种实施方案中使用的一个或更多个要素可以互换地用于另一种实施方案中,除非这种替换在技术上不可行。本领域技术人员将理解,在不脱离所要求保护的的主题的范围的情况下,可以对上述方法和结构进行各种其他的省略、添加和修改。所有此类修改和改变都旨在落在由所附权利要求书限定的主题的范围之内。

[0929] 本领域的技术人员将理解,对于本文公开的这个和其他过程和方法,在该过程和方法中执行的功能可以以不同的顺序实现。此外,所概述的步骤和操作仅作为示例提供,并且该步骤和操作中的一些可以是任选的,组合成较少的步骤和操作,或者扩展成另外的步骤和操作,而不偏离所公开的实施方案的本质。

[0930] 关于本文中基本上任何复数和/或单数术语的使用,在对于上下文和/或应用适当的情况下,本领域技术人员可以从复数转换为单数和/或从单数转换为复数。为了清楚起见,可以在本文明确阐述各种单数/复数排列。如本说明书和所附权利要求书中使用的,单数形式“一(a)”、“一(an)”和“所述/该(the)”包括复数指代物,除非上下文另外明确指示。除非另外说明,否则在本文中对“或”的任何提及旨在涵盖“和/或”。

[0931] 本领域技术人员将理解,一般来说,本文使用的术语,并且尤其是所附权利要求(例如,所附权利要求的主体)中的术语,通常意在作为“开放式”术语(例如,术语“包括(including)”应解释为“包括但不限于(including but not limited to)”,术语“具有(having)”应解释为“具有至少(having at least)”,术语“包括(includes)”应解释为“包括但不限于(includes but is not limited to)”,等等)。本领域技术人员将进一步理解,如果意图所介绍的权利要求陈述的特定数字,则这样的意图将明确地陈述于权利要求中,并且在这种陈述不存在的情况下,不存在这种意图。例如,作为对理解的帮助,以下所附权利要求书可以包含介绍性措辞“至少一种”和“一种或更多种”的使用,以介绍权利要求陈述。然而,此类措辞的使用不应解读为意味着由不定冠词“一(a)”或“一(an)”介绍权利要求

陈述会将任何包含此类介绍的权利要求陈述的具体权利要求限制为包含仅一种此类陈述的实施方案,甚至当同一权利要求包括介绍性措辞“一种或更多种”或“至少一种”以及不定冠词诸如“一(a)”或“一(an)”时也是如此(例如,“一(a)”和/或“一(an)”应解释为意指“至少一种”或“一种或更多种”);这对于使用定冠词来介绍权利要求陈述同样适用。此外,即使明确地陈述了介绍的权利要求陈述的特定数字,本领域技术人员将认识到,此类陈述应解释为意指至少所陈述的数字(例如,仅陈述“两种陈述”而没有其他修饰词意指至少两种陈述或两种或更多种陈述)。此外,在使用类似于“A、B和C等中的至少一种”的惯例的那些情况下,通常这种句法结构意在为本领域技术人员将理解该惯例的意义(例如,“具有A、B和C中的至少一种的系统”将包括但不限于具有单独的A,具有单独的B,具有单独的C,A和B一起,A和C一起,B和C一起,和/或A、B和C一起等的系统)。在使用类似于“A、B或C等中的至少一种”的惯例的那些情况下,通常这种句法结构意在为本领域技术人员将理解该惯例的意义(例如,“具有A、B或C中的至少一种的系统”将包括但不限于具有单独的A,具有单独的B,具有单独的C,A和B一起,A和C一起,B和C一起,和/或A、B和C一起等的系统)。本领域技术人员将进一步理解,实际上,无论在说明书、权利要求书还是在附图中,呈现两个或更多个替代术语的任何分离性词语和/或措辞应被理解为考虑到包括术语之一、任一术语或两个术语的可能性。例如,短语“A或B”应被理解为包括“A”或“B”或“A和B”的可能性。

[0932] 此外,当本公开内容的特征或方面以马库什组(Markush group)描述时,本领域技术人员将意识到,本公开内容还由此以马库什组的任何个体成员或成员子组描述。

[0933] 如本领域技术人员将理解的,为了任何和所有目的,诸如在提供书面描述方面,本文公开的所有范围还包括该范围的任何和所有可能的子范围和子范围的组合。任何列出的范围都可以很容易地被识别为充分描述并使相同的范围能被分解为至少相等的一半、三分之一、四分之一、五分之一、十分之一等。作为非限制性实例,本文讨论的每个范围可以容易地分解为下三分之一、中三分之一和上三分之一等。如本领域技术人员还将理解的,所有语言,诸如“多达”、“至少”、“大于”、“小于”等包括所陈述的数字,并且指可以随后分解为如上文讨论的子范围的范围。最后,如本领域技术人员将理解的,范围包括每个个体的成员。因此,例如,具有1-3个物品的组是指具有1个、2个或3个物品的组。类似地,具有1-5个物品的组是指具有1个、2个、3个、4个或5个物品的组,等等。

[0934] 从前述内容,应当理解,本文出于说明的目的已经描述了本公开内容的各种实施方案,并且可以在不脱离本公开内容的范围和精神的情况下进行各种修改。因此,本文公开的各种实施方案并不旨在进行限制,真正的范围和精神由以下权利要求来指示。

## 序列表

- <110> 贝克顿迪金森公司  
 伊缪德克斯私人有限责任公司  
 辛西娅·萨科夫斯基  
 凯瑟琳·拉扎鲁克  
 玛格丽特·纳卡莫托  
 戴文·约瑟夫·延森  
 基文·达尔·雅各布森
- <120> DEXTRAMER在单细胞分析中的使用
- <130> 68EB-317310-WO
- <150> 63/116,571
- <151> 2020-11-20
- <150> 63/241,486
- <151> 2021-09-07
- <160> 6
- <170> PatentIn version 3.5
- <210> 1
- <211> 20
- <212> DNA
- <213> 人工序列(Artificial Sequence)
- <220>
- <223> PCR1引物
- <400> 1  
 ggagggaggt tagcgaaggt 20
- <210> 2
- <211> 42
- <212> DNA
- <213> 人工序列(Artificial Sequence)
- <220>
- <223> PCR2引物
- <400> 2  
 cagacgtgtg ctcttccgat ctggaggag gttagcgaag gt 42
- <210> 3
- <211> 93
- <212> DNA
- <213> 人工序列(Artificial Sequence)
- <220>
- <223> dCODE Dextramer寡核苷酸

- <220>  
 <221> misc\_feature  
 <222> (22) .. (23)  
 <223> n是a,c,g或t  
 <220>  
 <221> misc\_feature  
 <222> (25) .. (28)  
 <223> n是a,c,g或t  
 <220>  
 <221> misc\_feature  
 <222> (30) .. (31)  
 <223> n是a,c,g或t  
 <400> 3  
 ggagggaggt tagcgaaggt vnnvnnnnvn nvgtgttggg ttggcctatg ttaattacgg 60  
 agtgggtgaa aaaaaaaaaa aaaaaaaaaa aaa 93  
 <210> 4  
 <211> 13  
 <212> PRT  
 <213> 人工序列(Artificial Sequence)  
 <220>  
 <223> 抗原肽(EBV)  
 <400> 4  
 Thr Ser Leu Tyr Asn Leu Arg Arg Gly Thr Ala Leu Ala  
 1 5 10  
 <210> 5  
 <211> 13  
 <212> PRT  
 <213> 人工序列(Artificial Sequence)  
 <220>  
 <223> 抗原肽(破伤风类毒素(Tetanus toxoid))  
 <400> 5  
 Lys Ile Tyr Ser Tyr Phe Pro Ser Val Ile Ser Lys Val  
 1 5 10  
 <210> 6  
 <211> 15  
 <212> PRT  
 <213> 人工序列(Artificial Sequence)  
 <220>  
 <223> 抗原肽(阴性对照Clip)

<400> 6

Pro Val Ser Lys Met Arg Met Ala Thr Pro Leu Leu Met Gln Ala

1

5

10

15

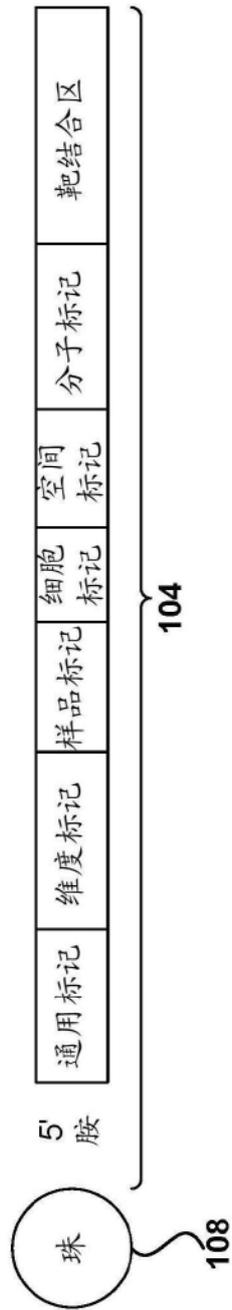


图1

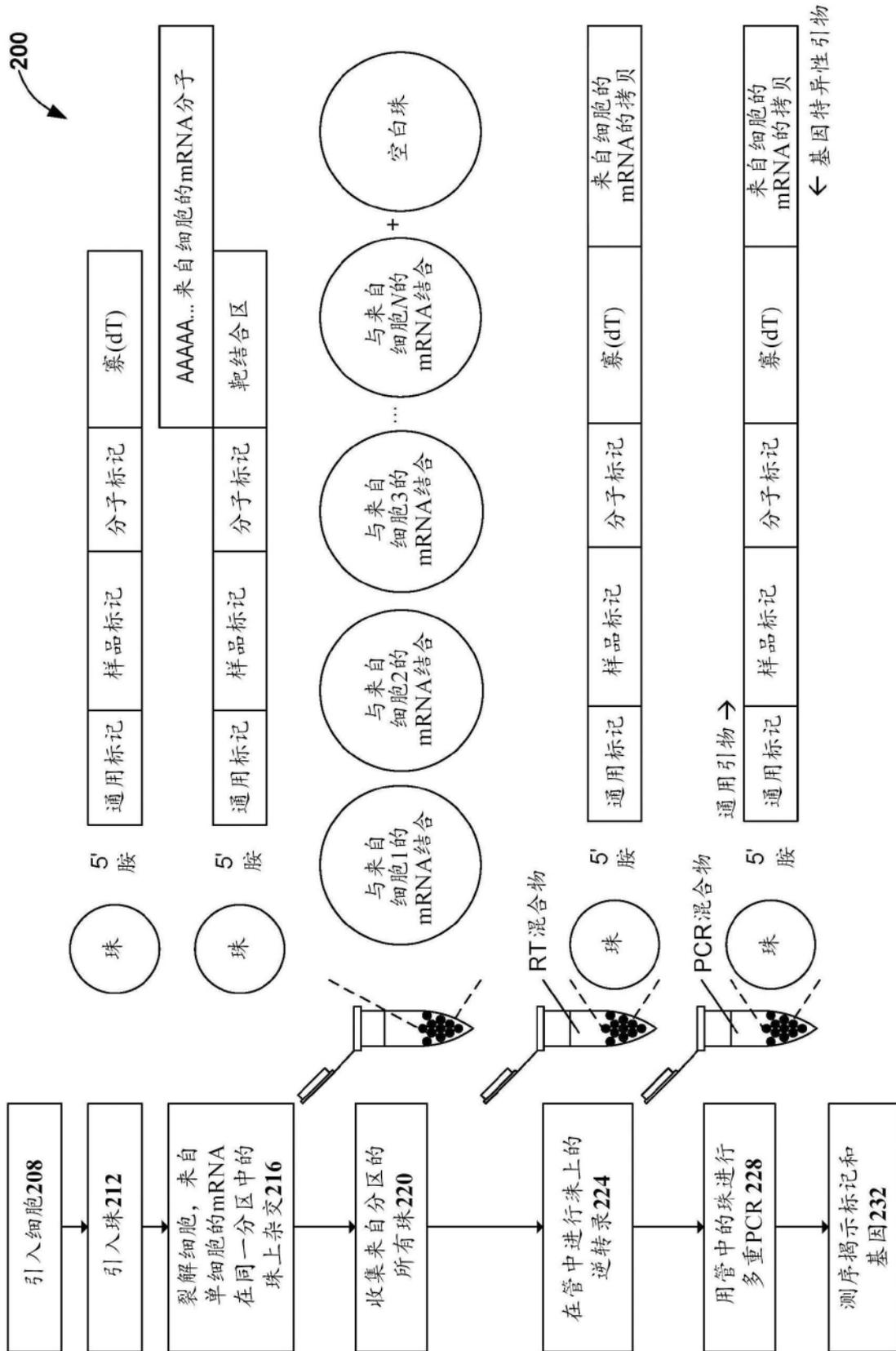


图2

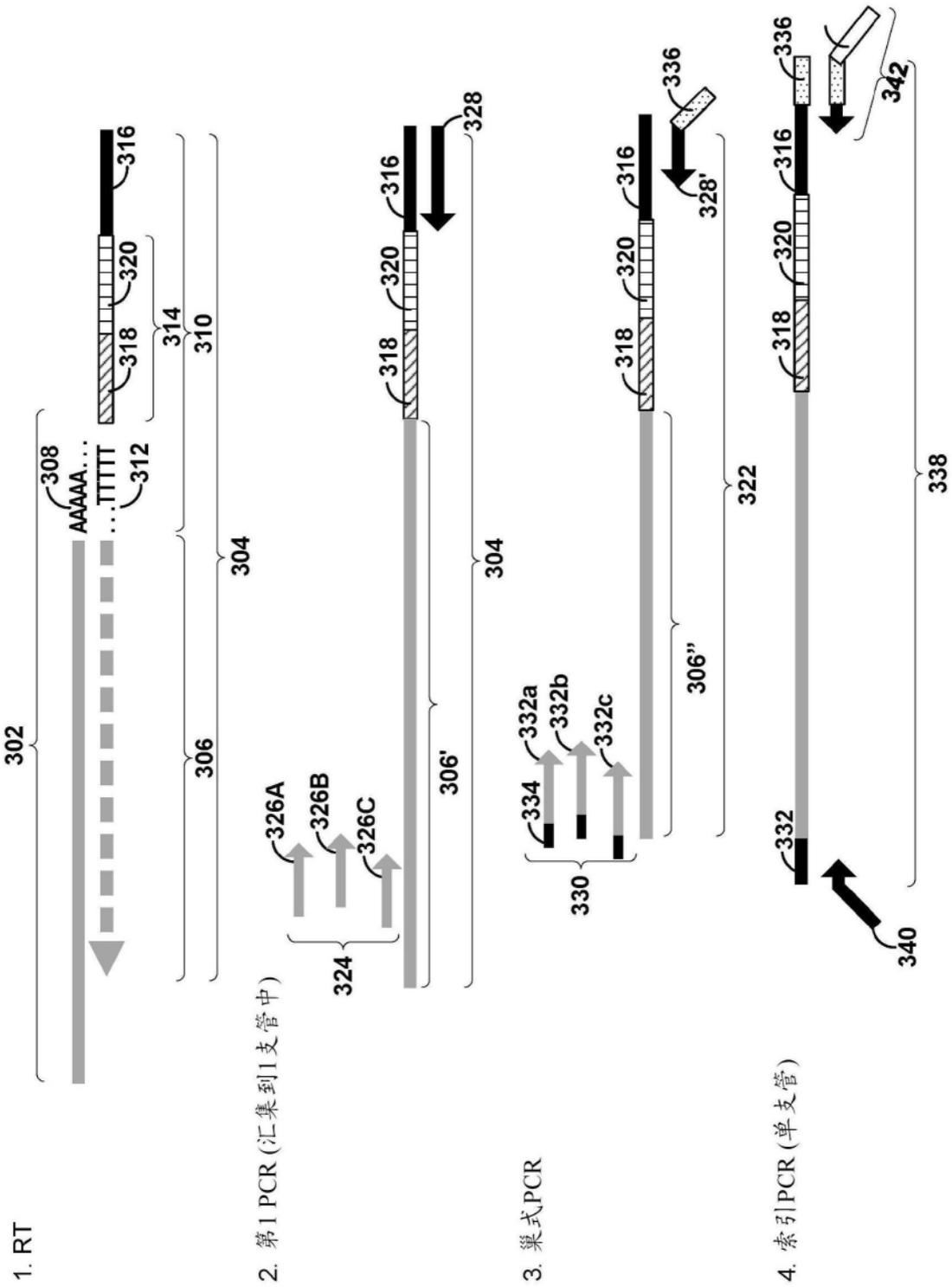


图3

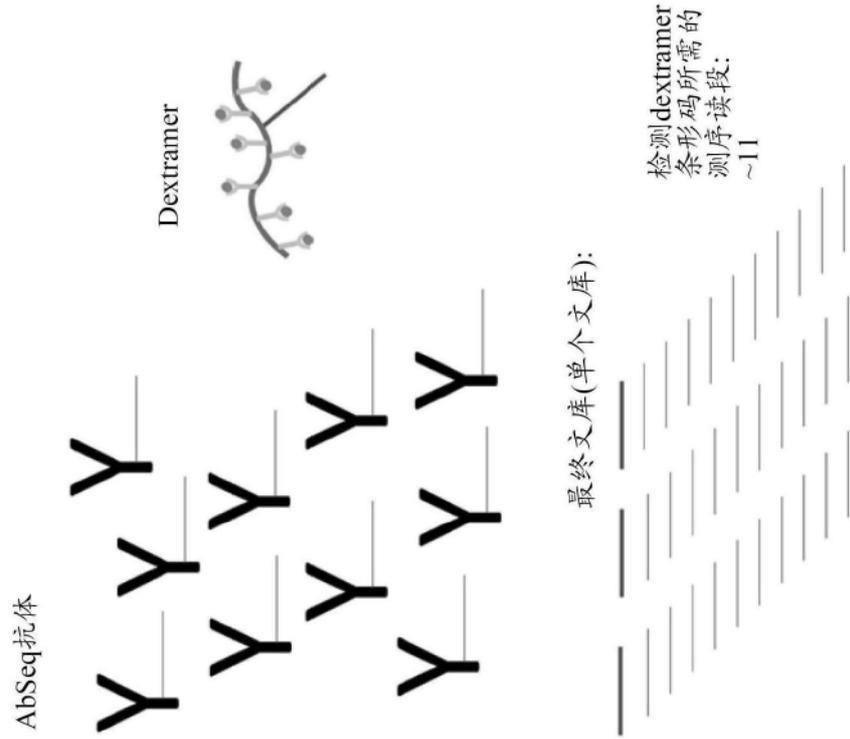


图4A

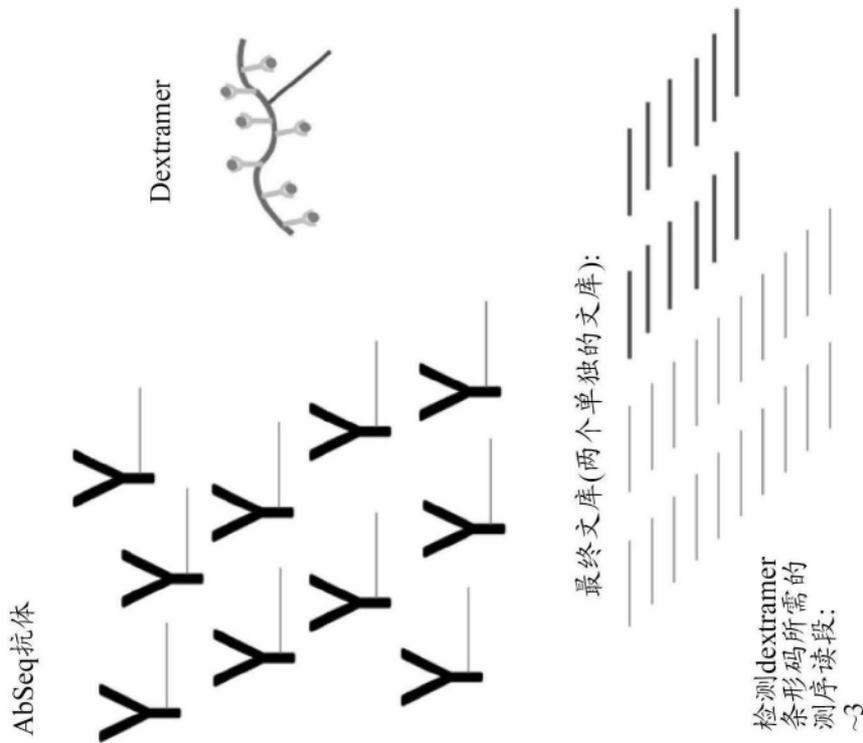


图4B

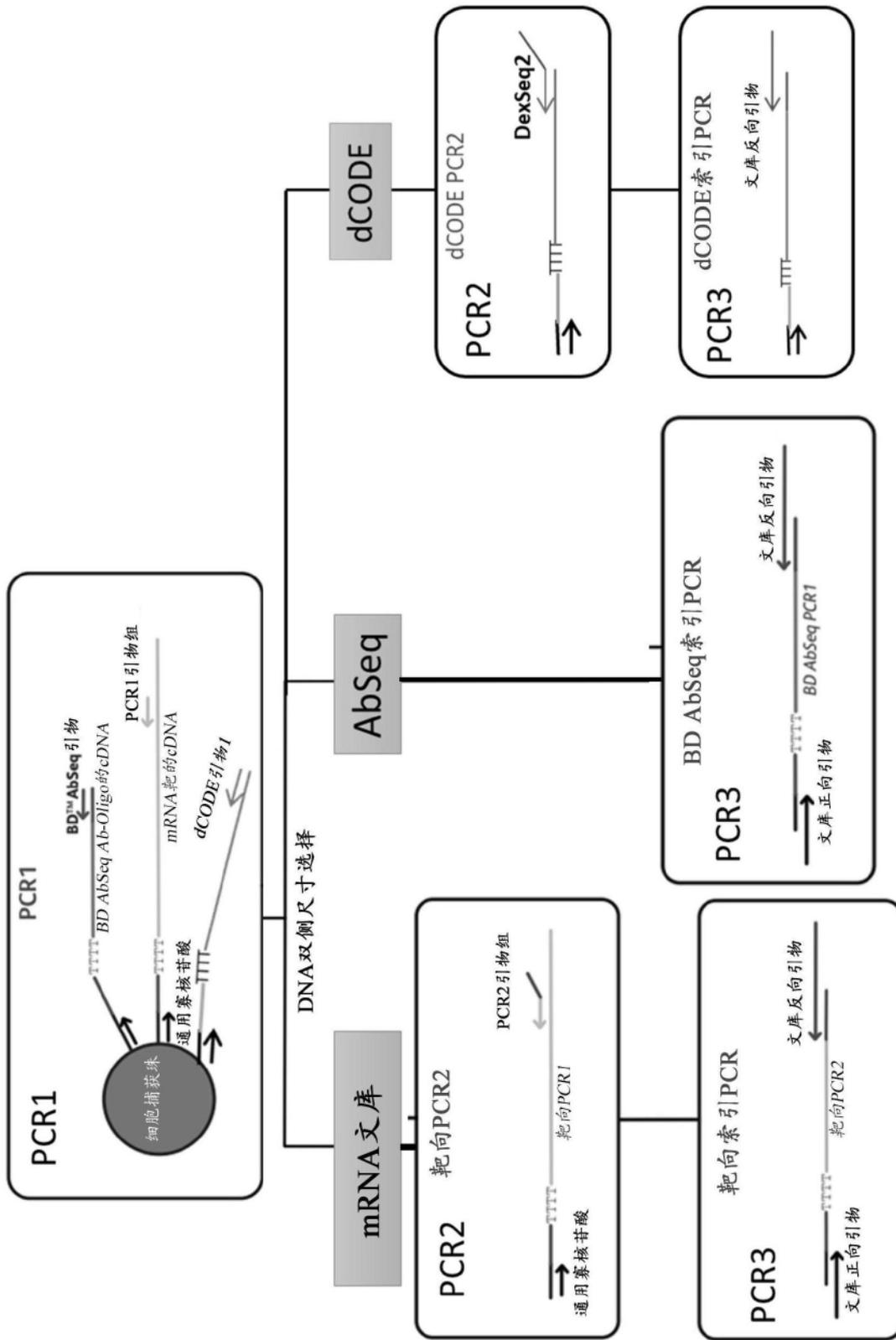


图5

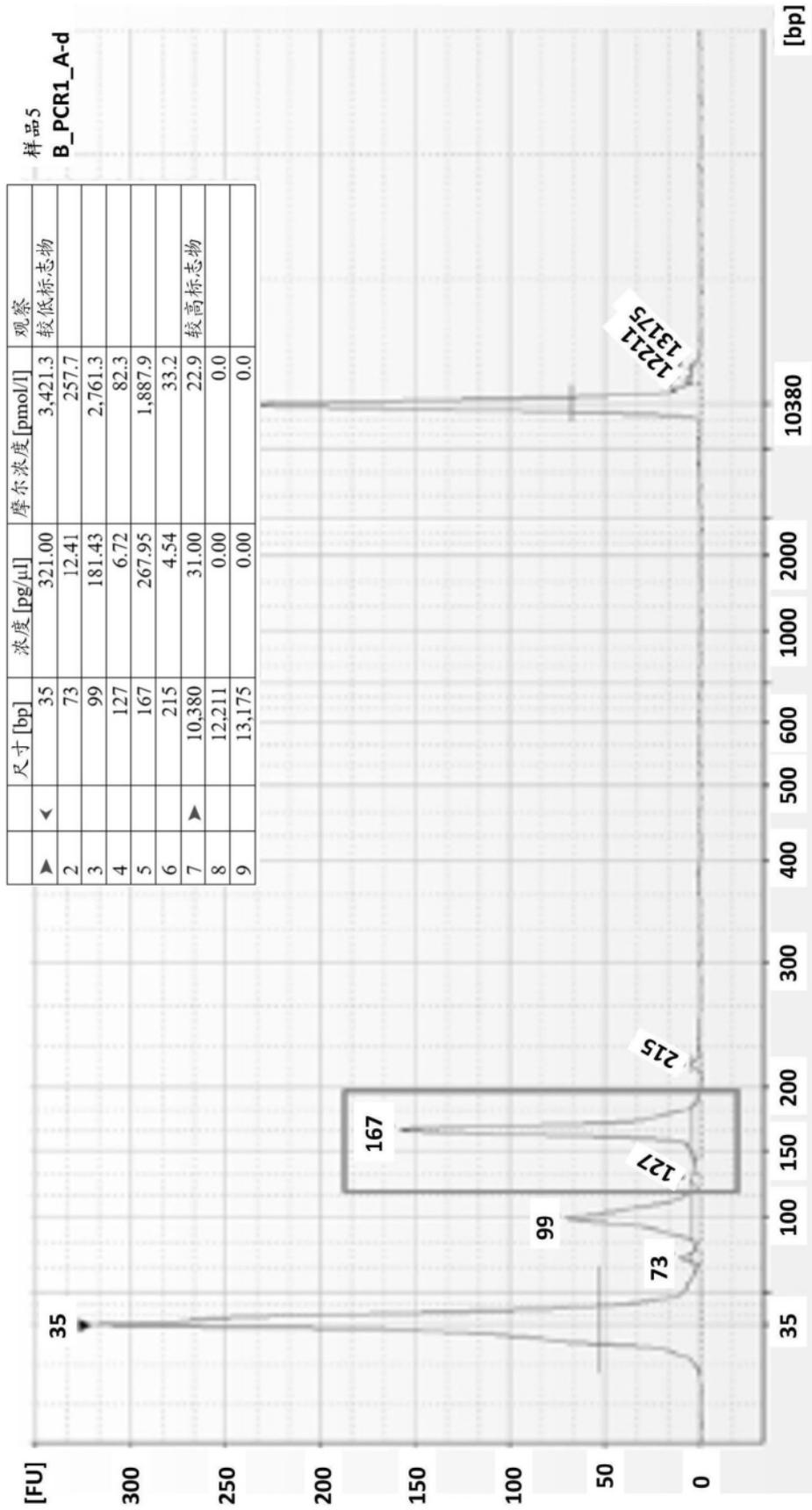


图6

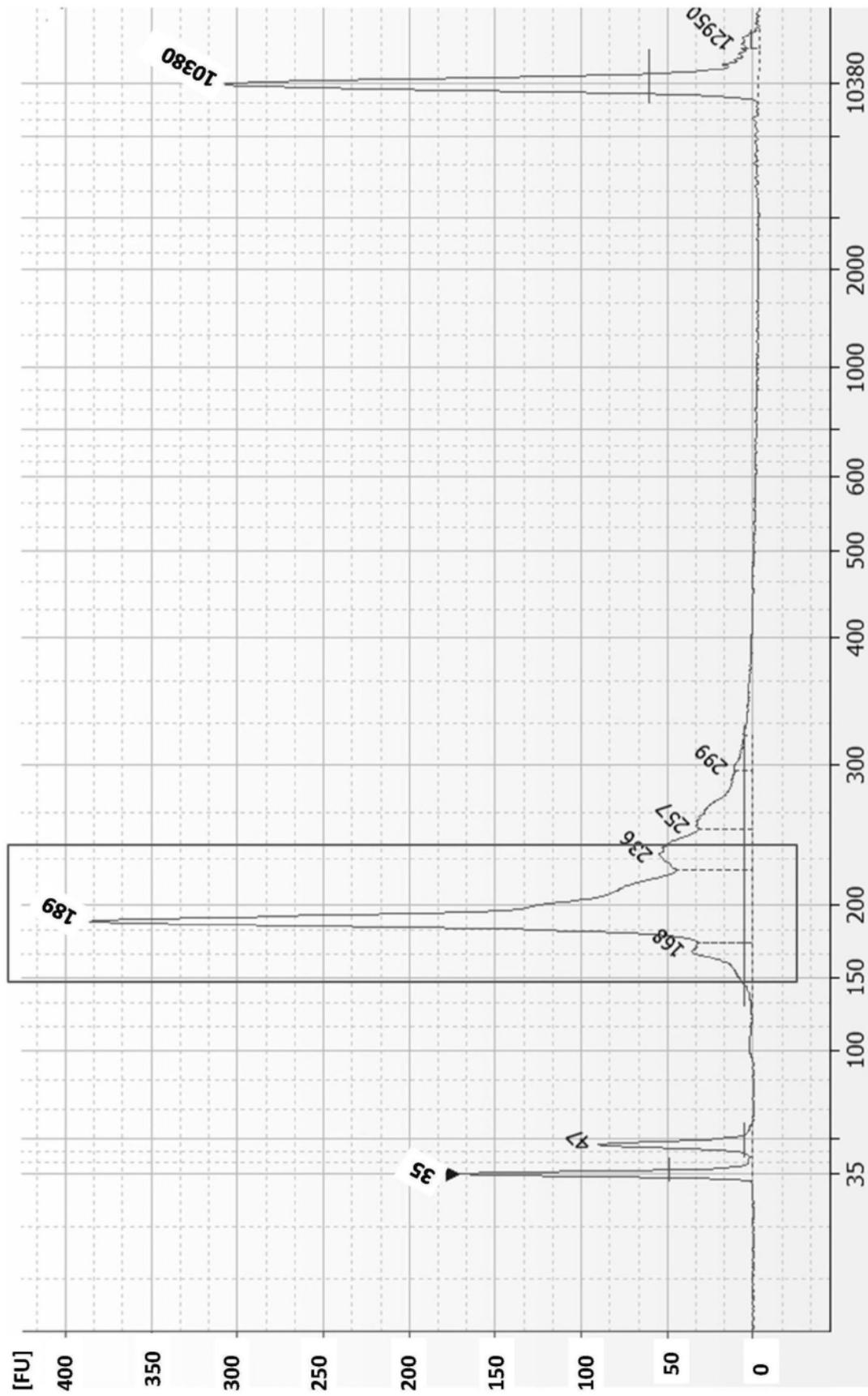


图7

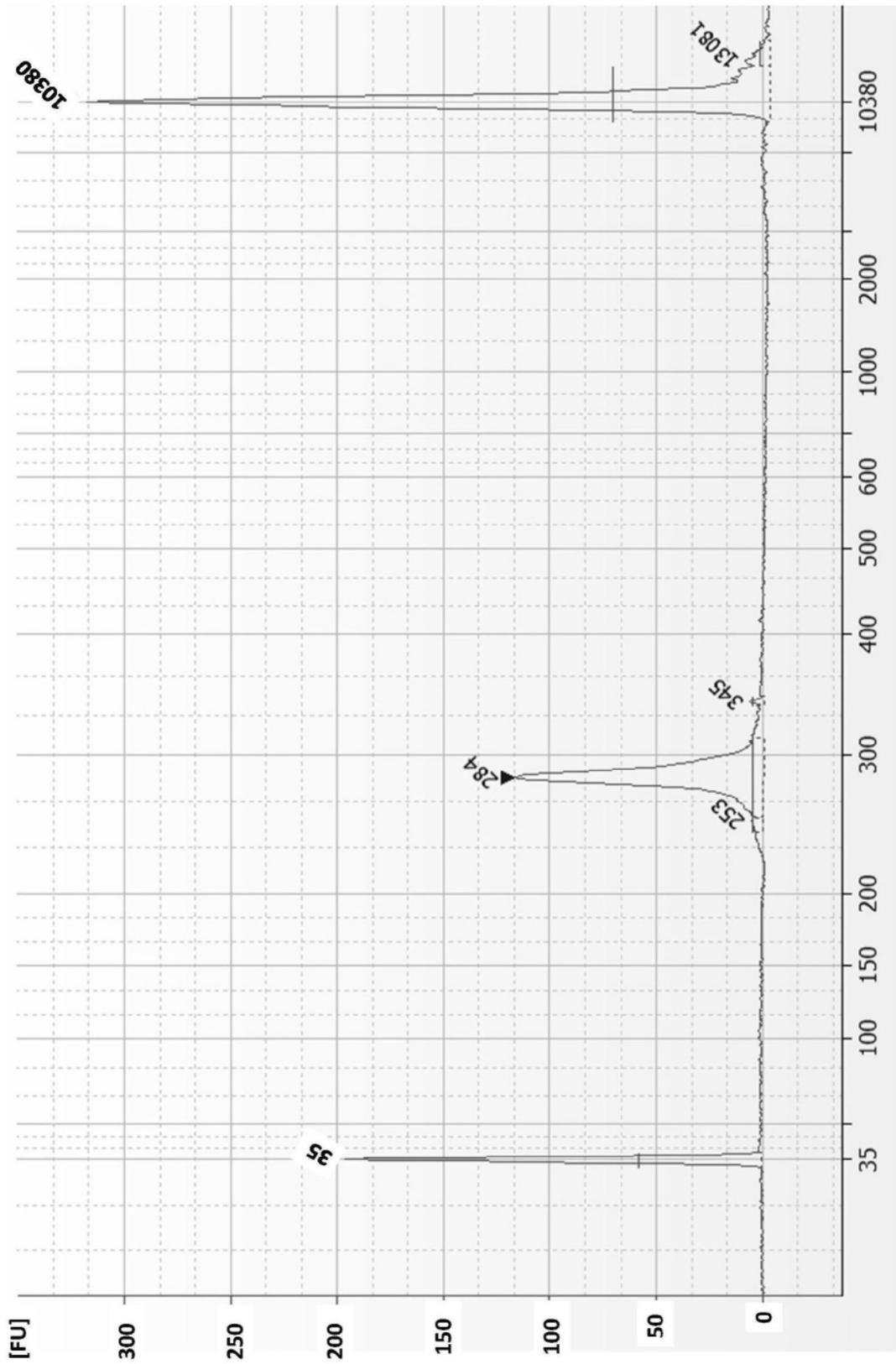


图8

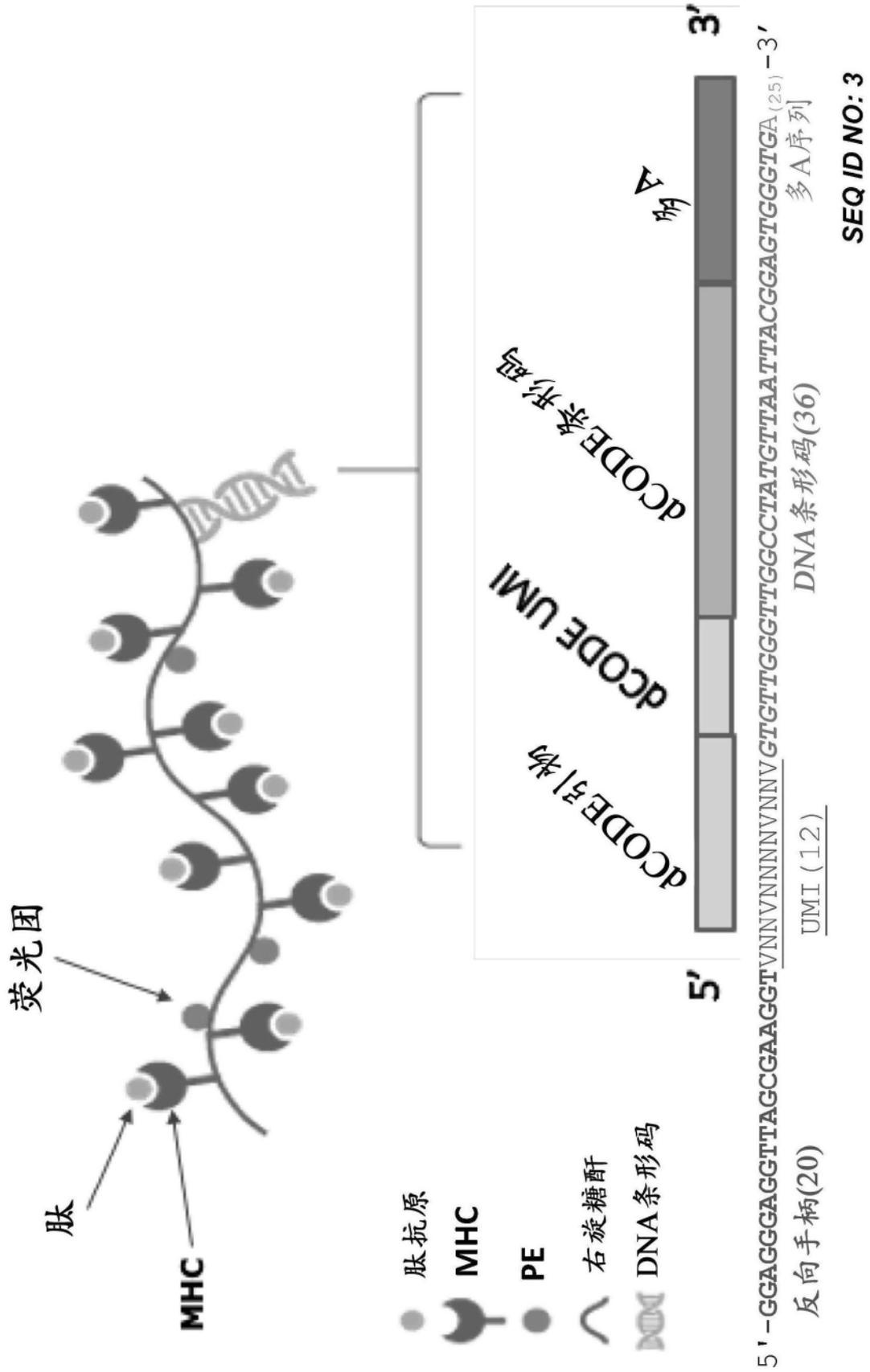


图9A

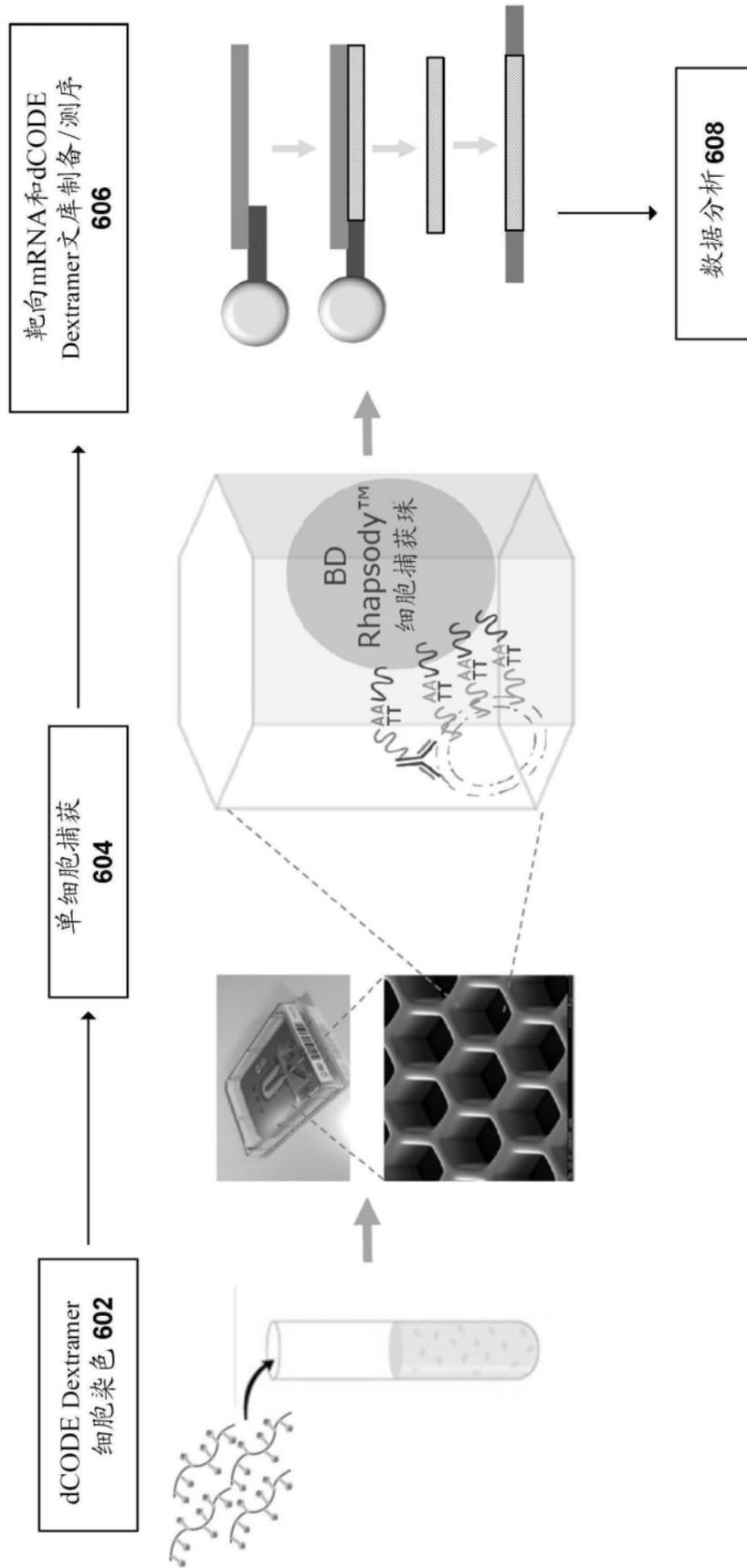


图9B

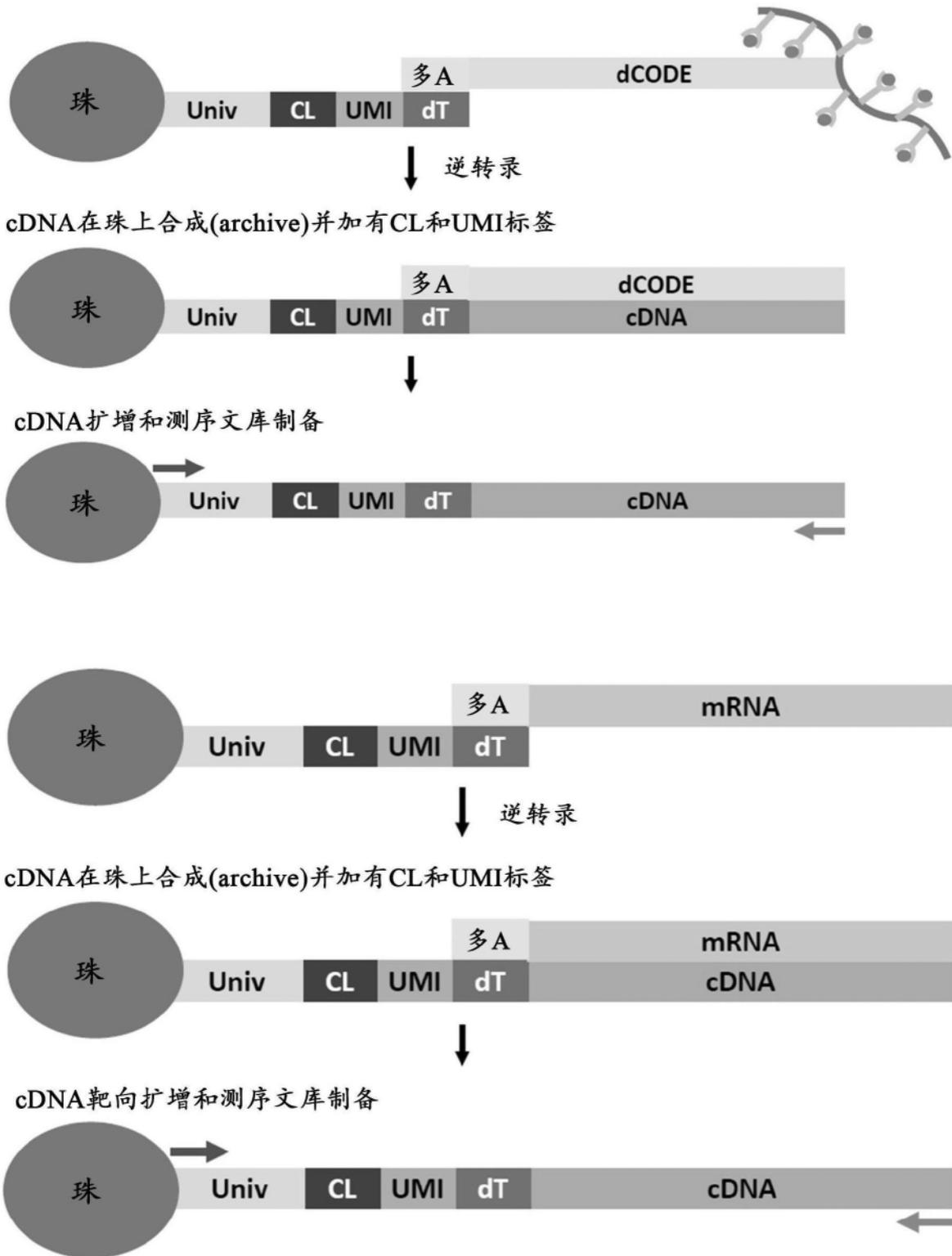


图9C

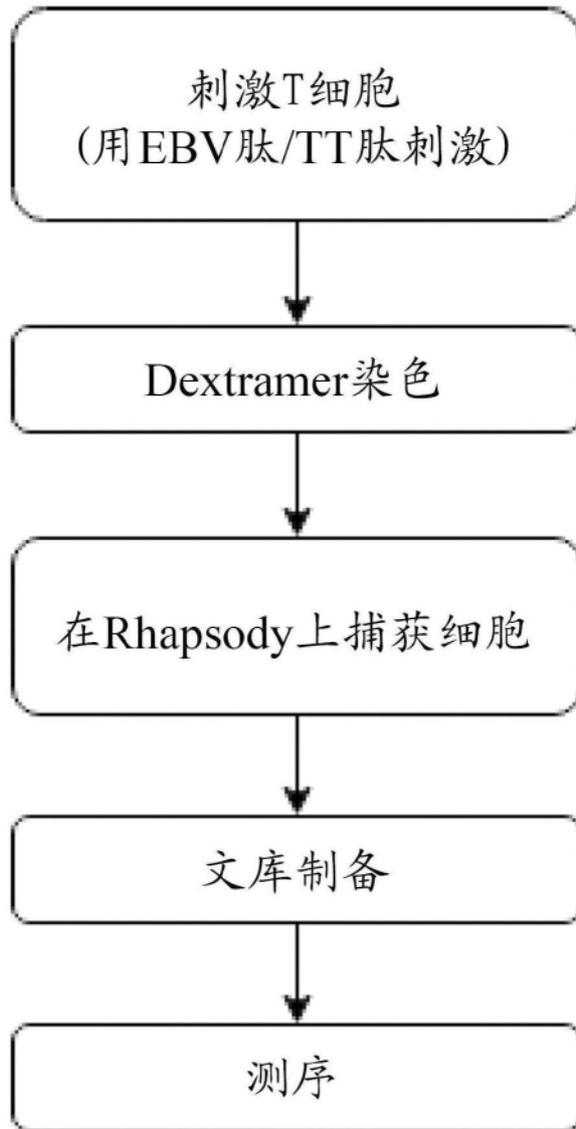


图9D

阴性对照      破伤风类毒素      EBV

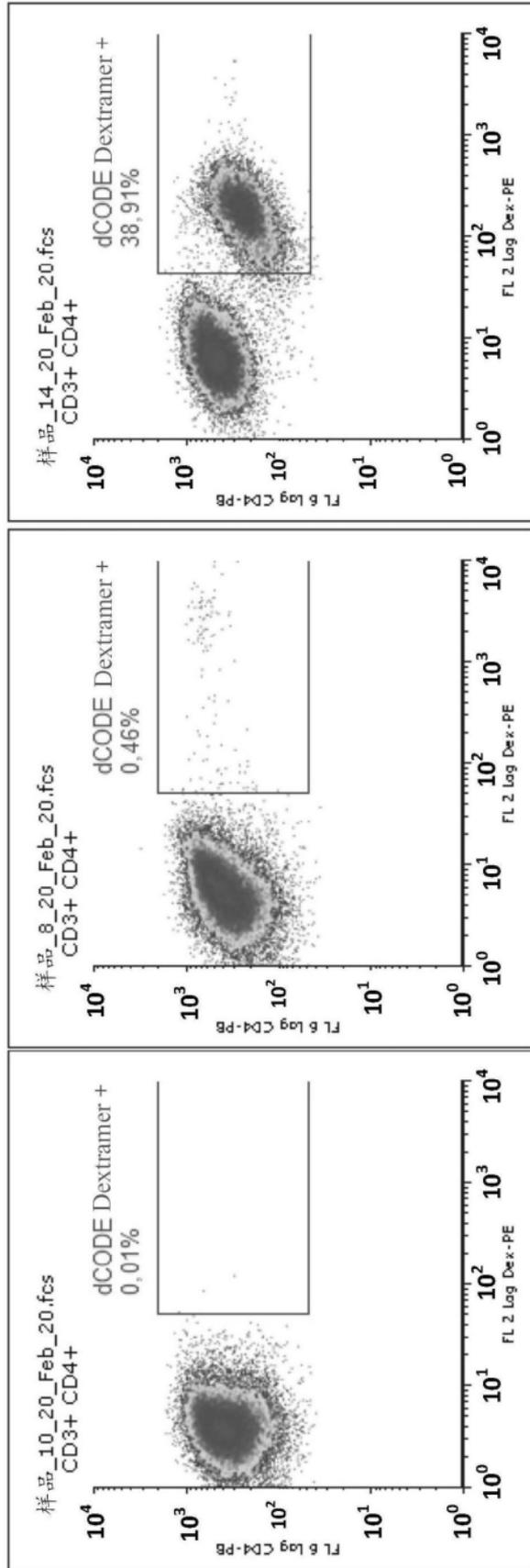


图9E

# 门控：散射, CD3/APC+, CD4/BV421+

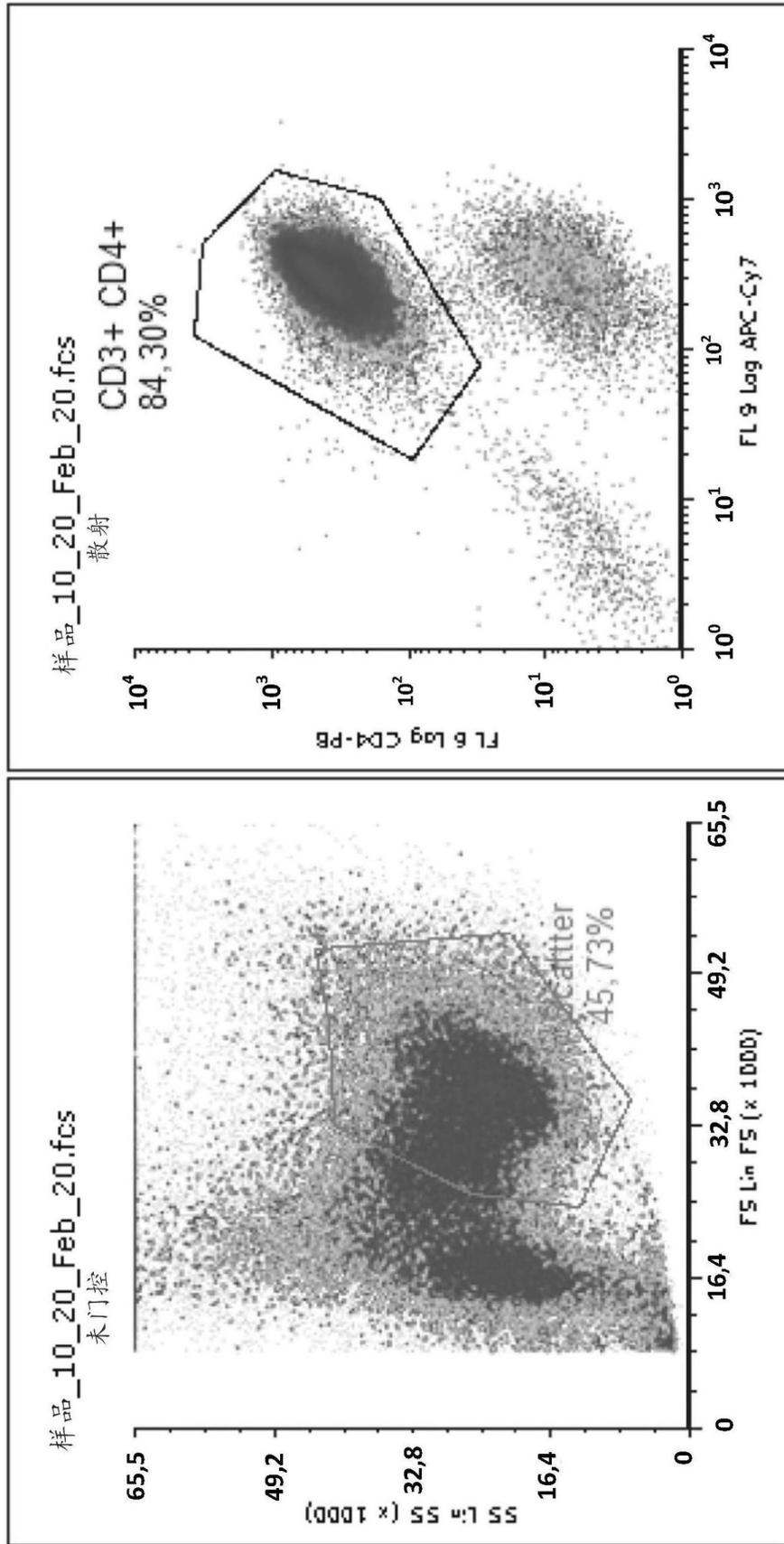


图9F

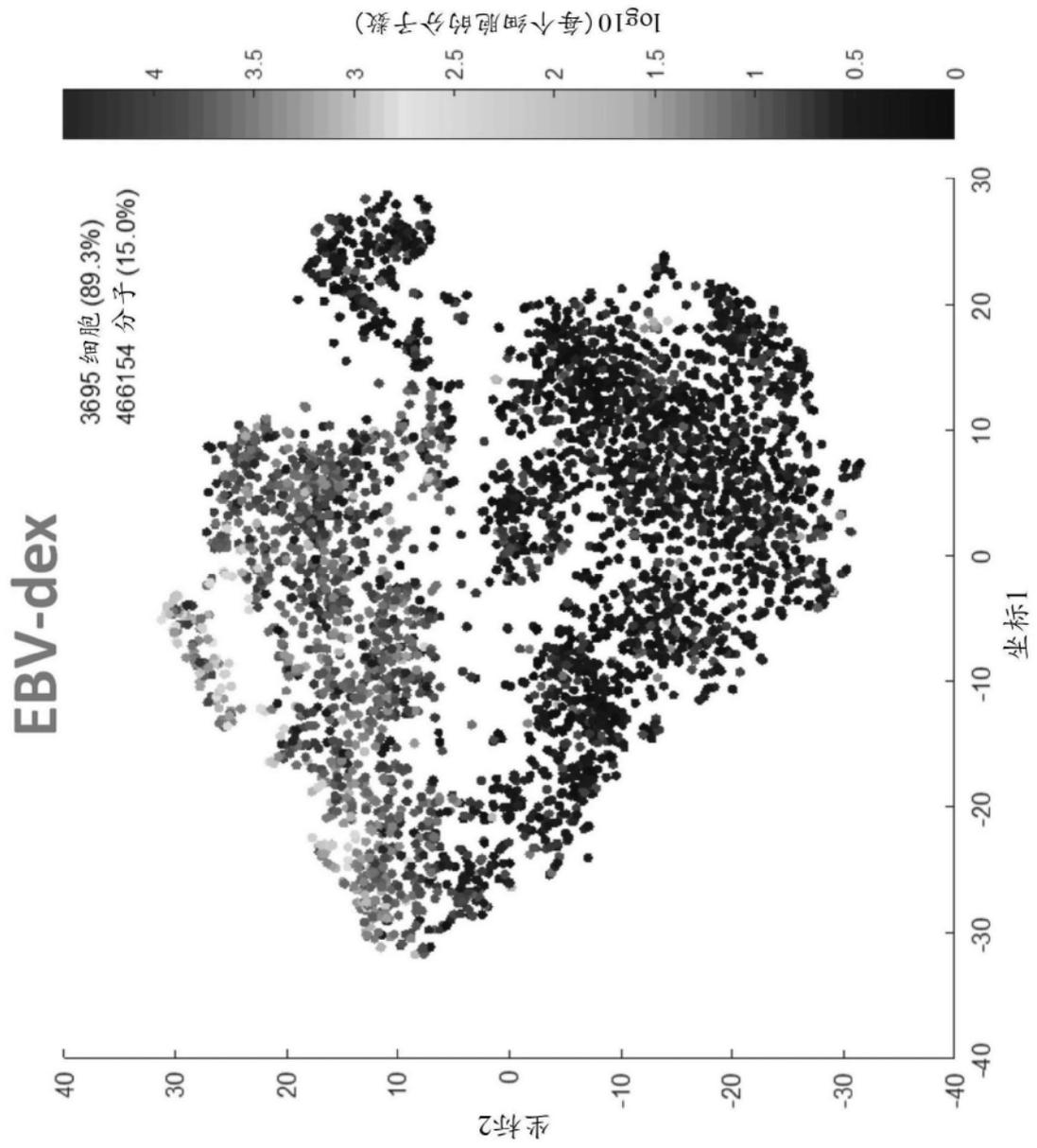


图10A

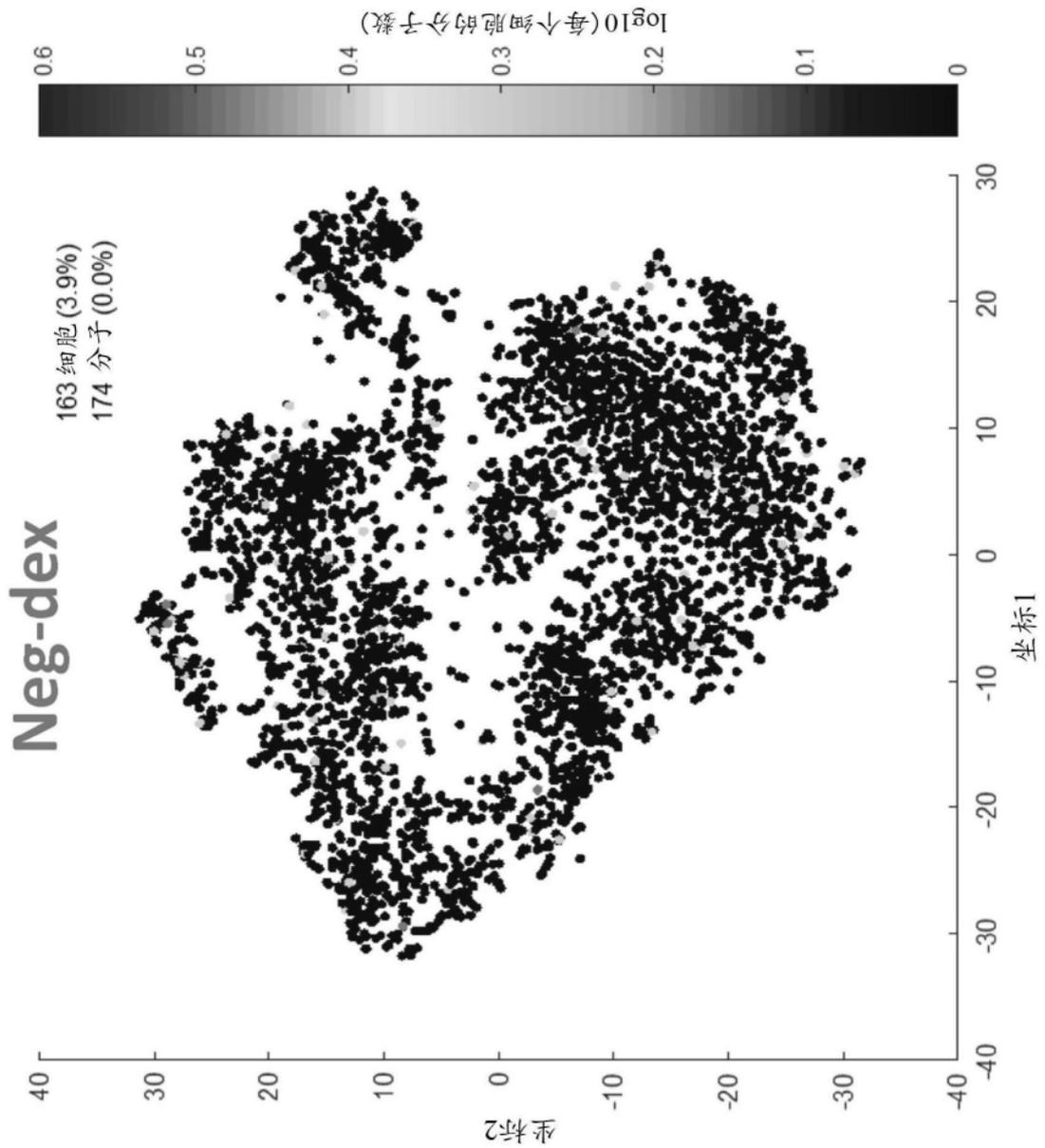


图10A(续)

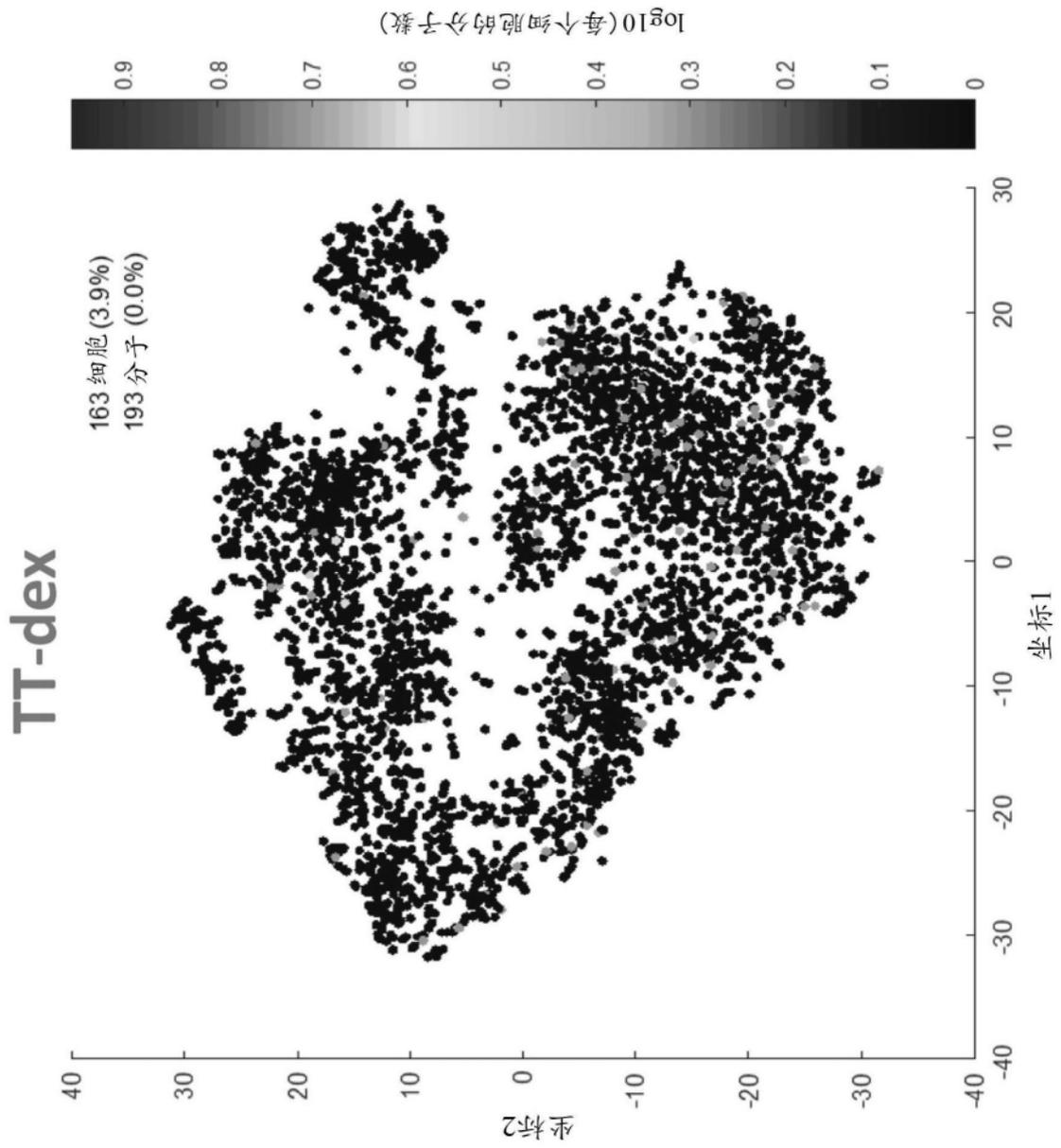


图10A(续)

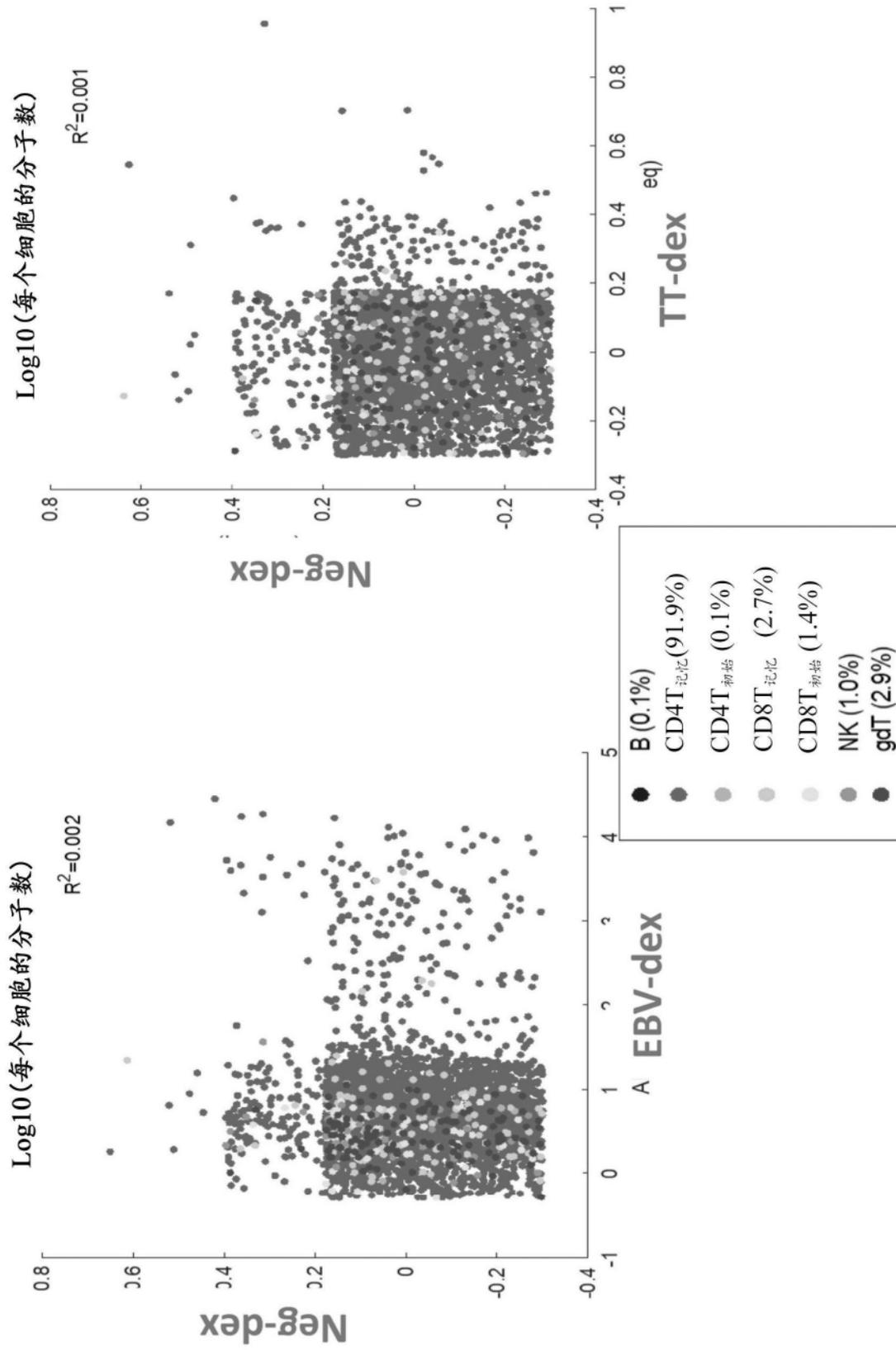


图10A(续)

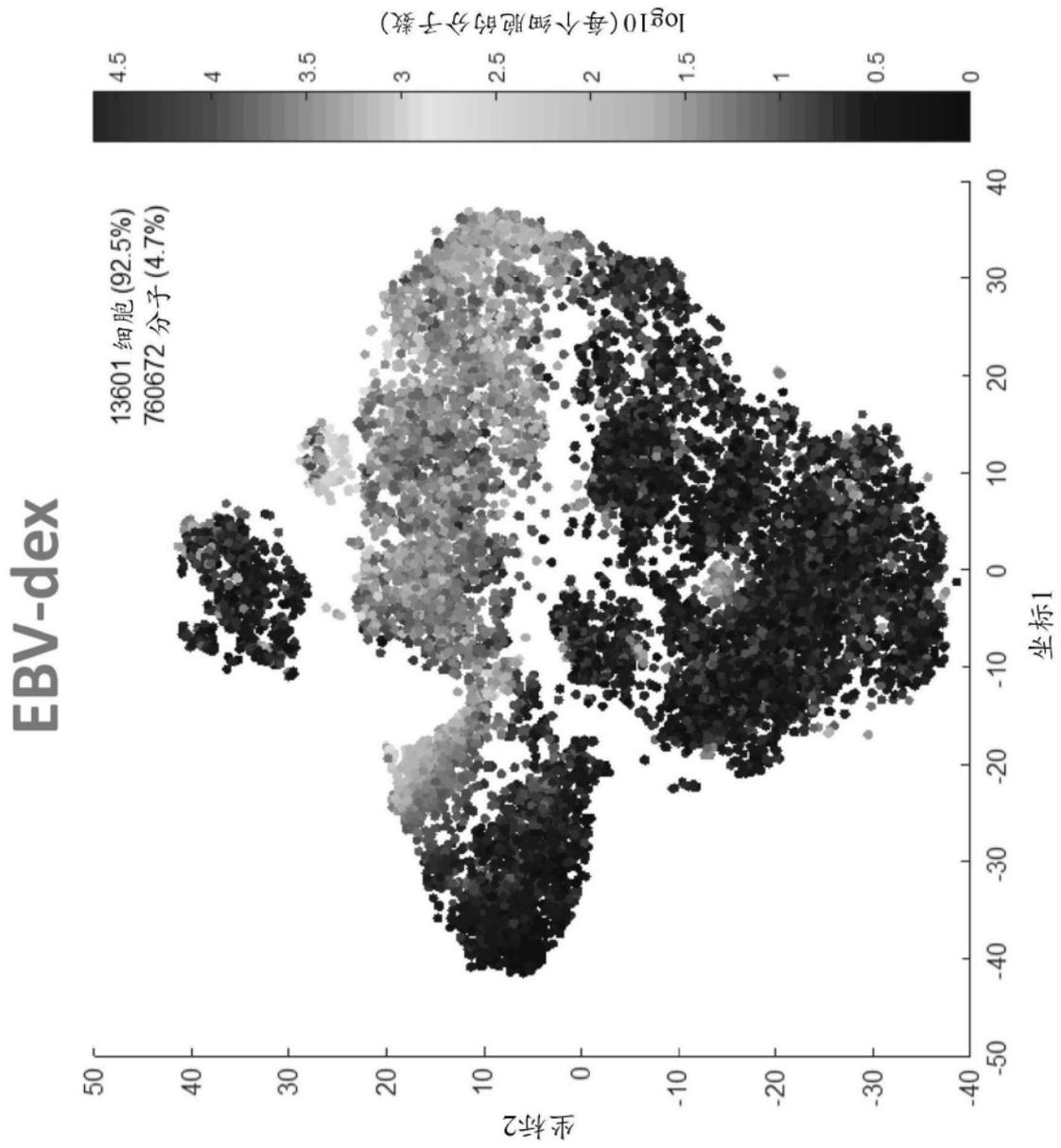


图10B

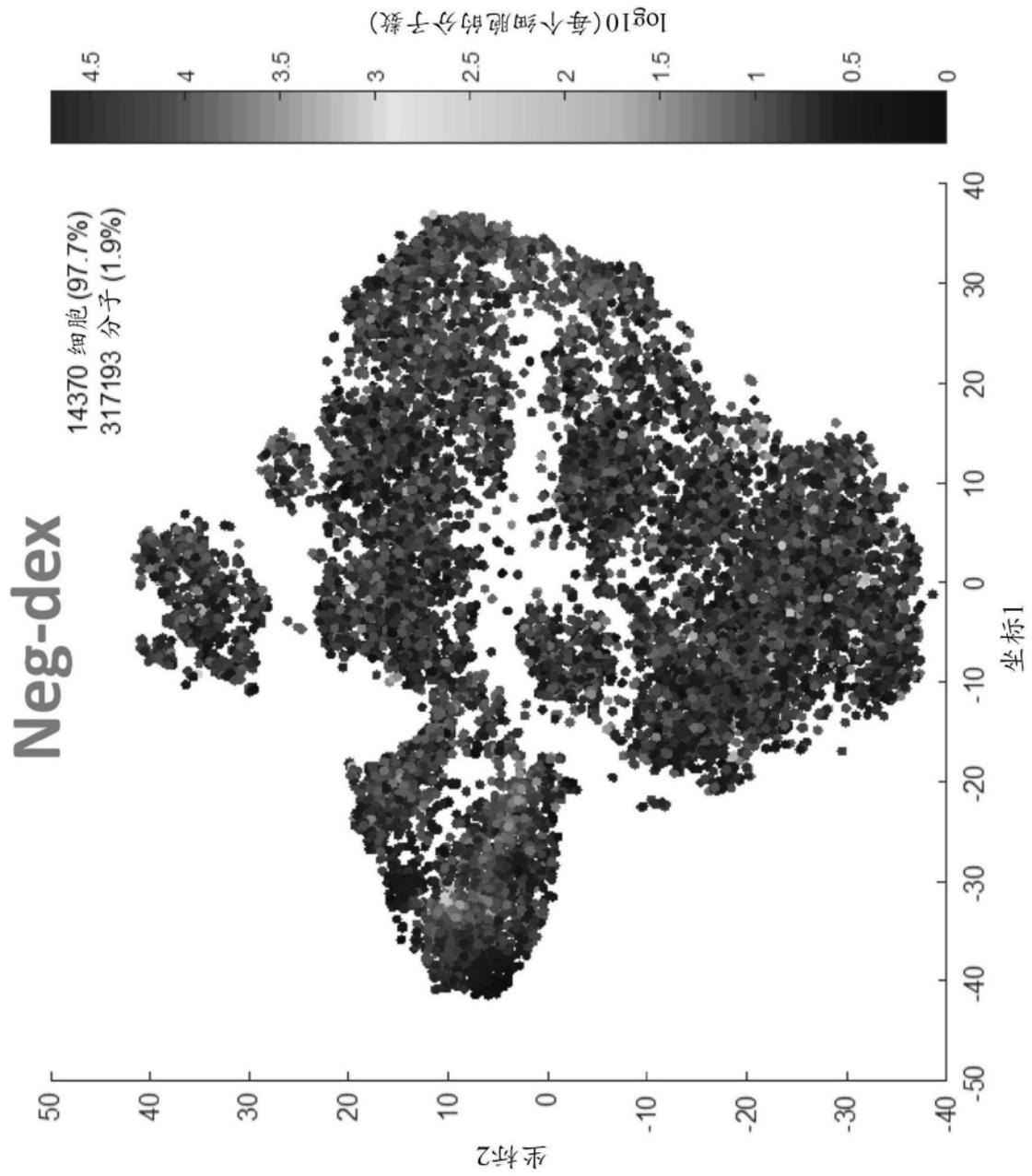


图10B(续)

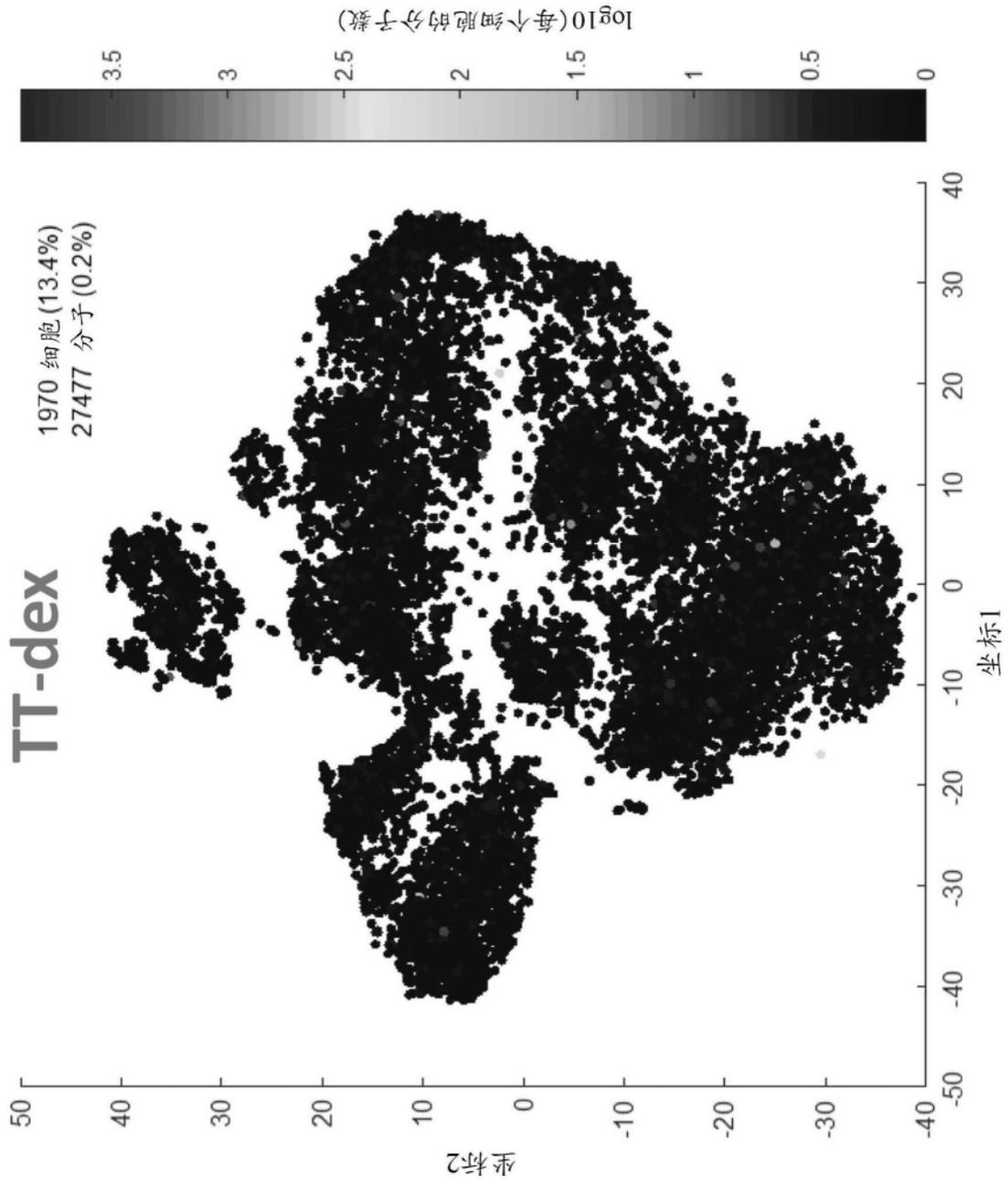


图10B(续)

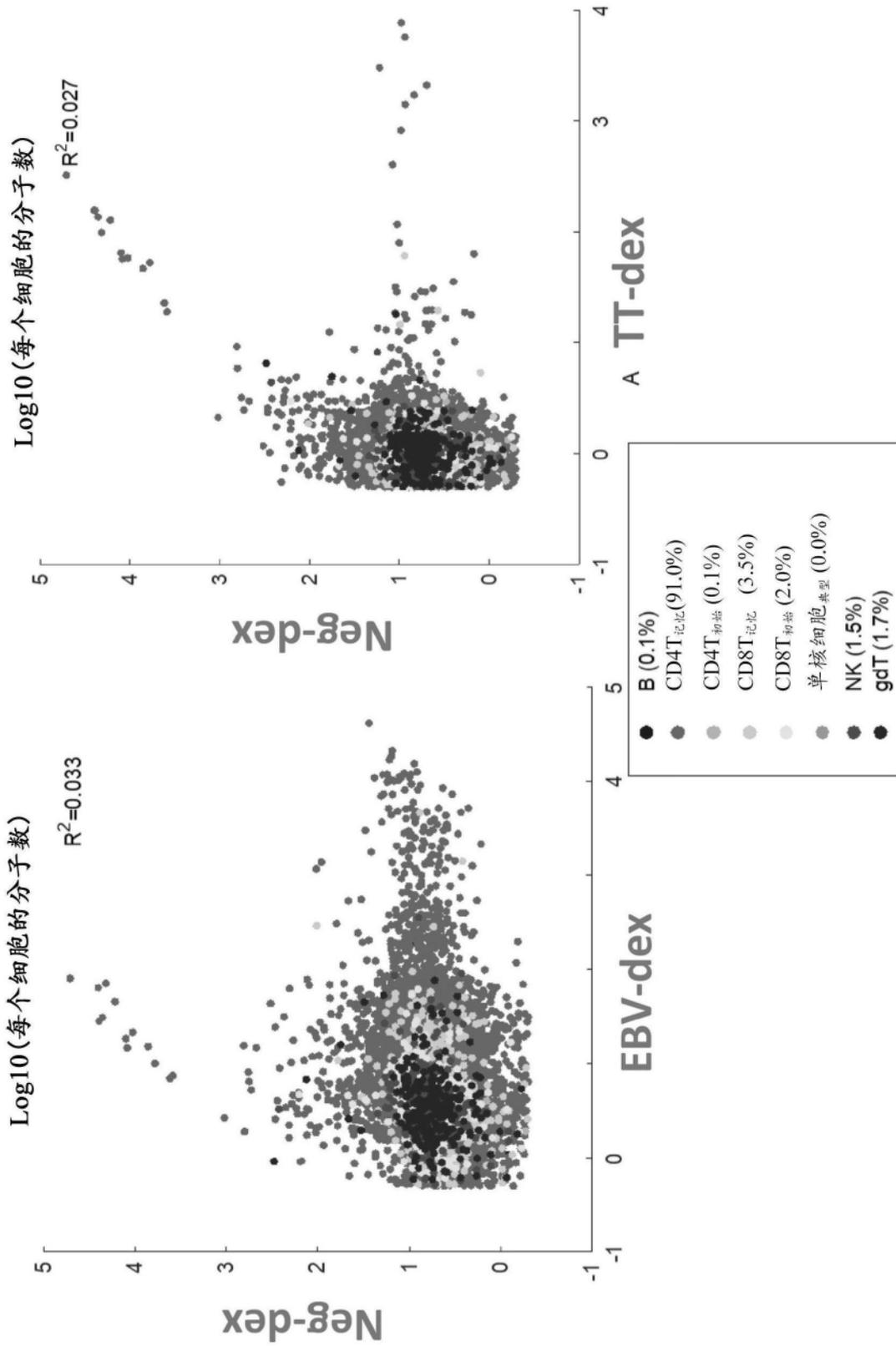
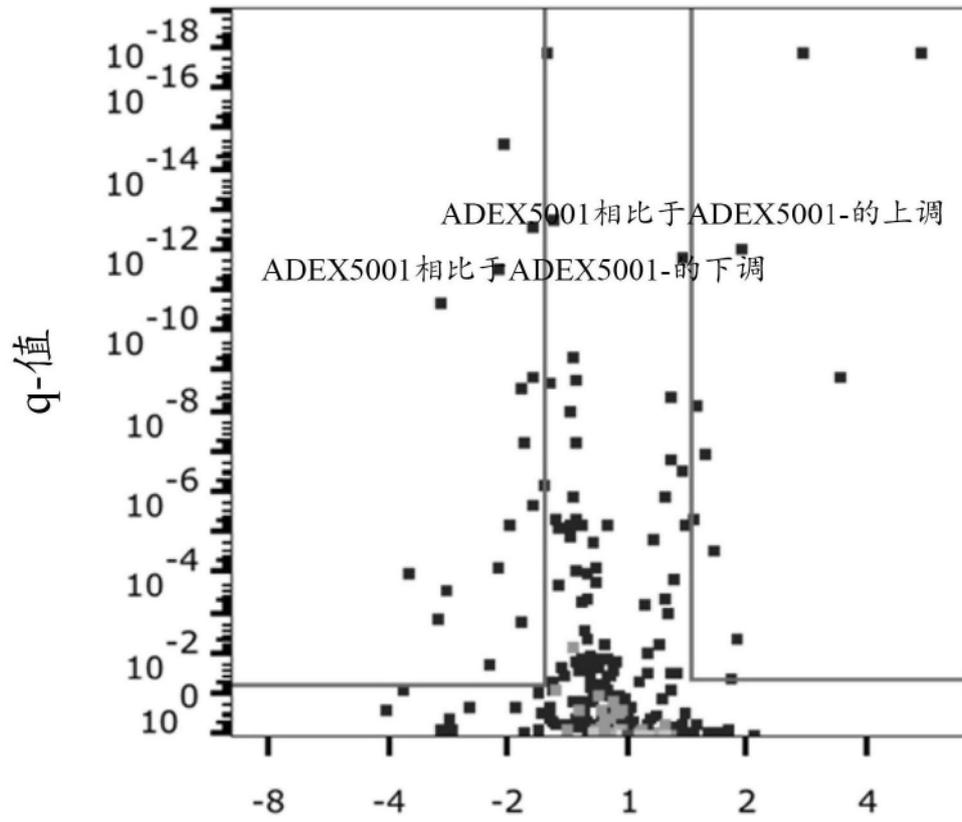


图10B(续)





EBV+细胞相比于阴性细胞的变化倍数

图11A(续)



**31.6228**  **1426.1872**  
EBV+细胞相比于阴性细胞的下调  
(总和)



**31.6228**  **1206.0295**  
EBV+细胞相比于阴性细胞的上调  
(总和)

图11B

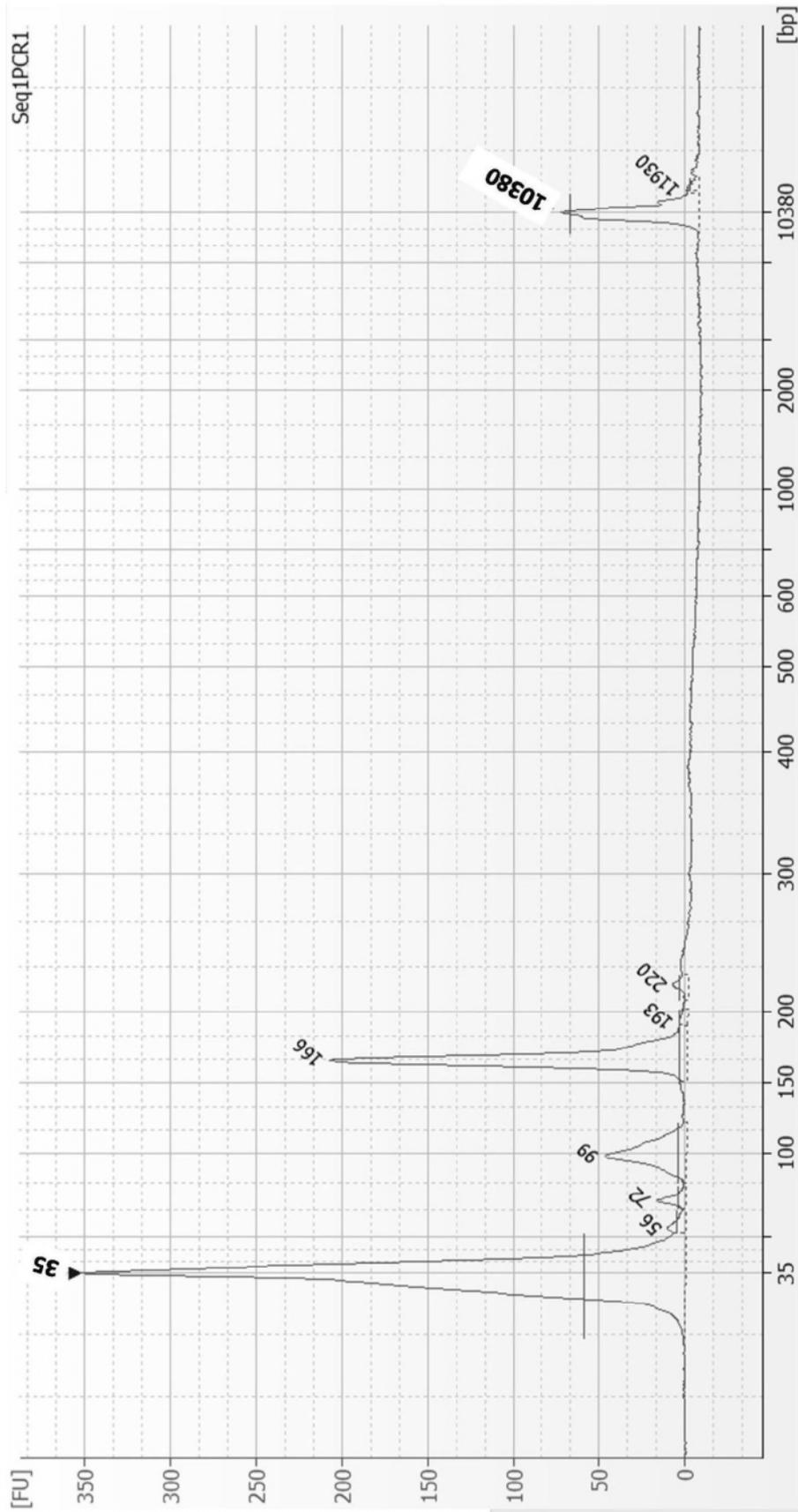


图12A

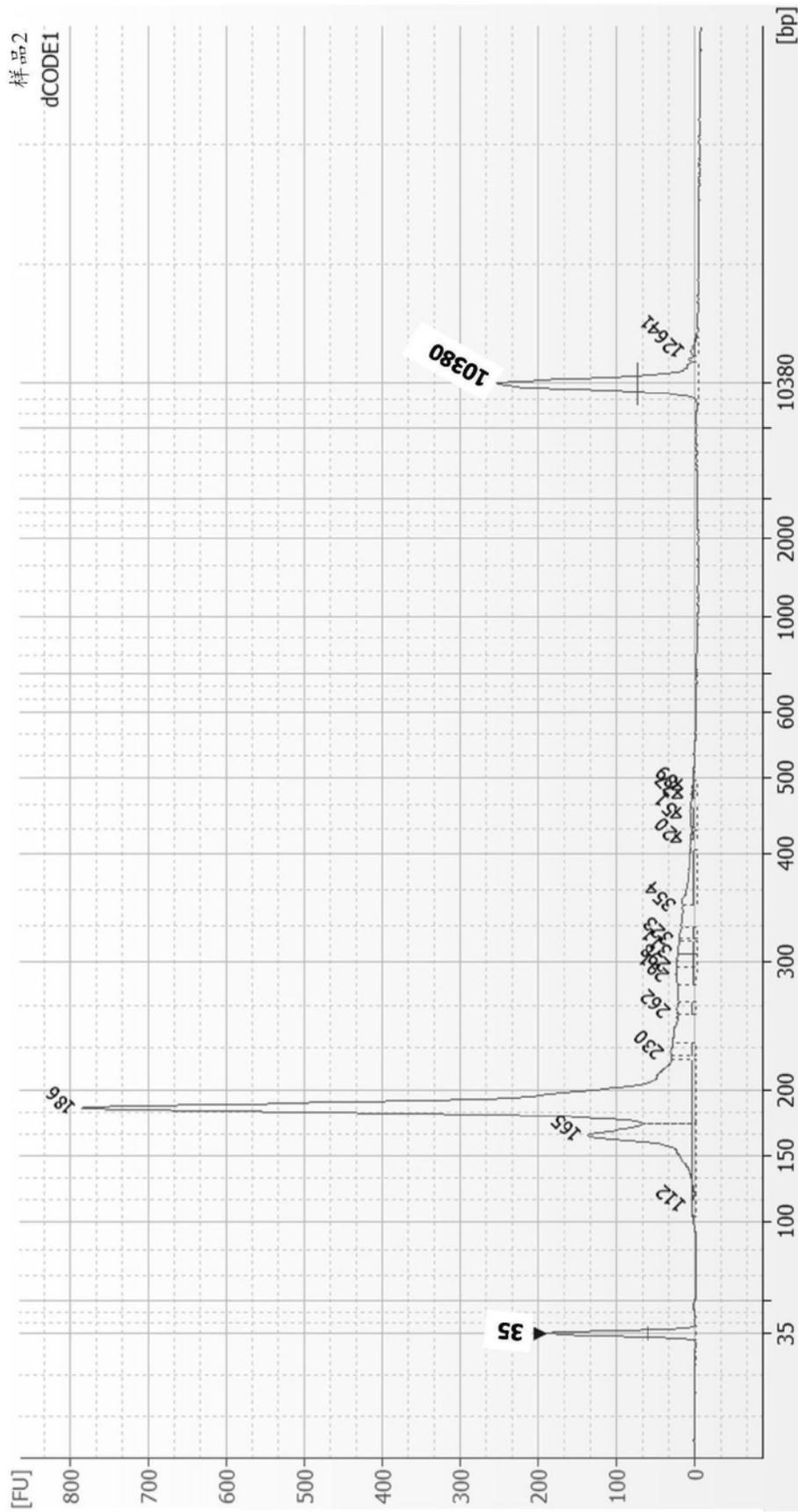


图12B

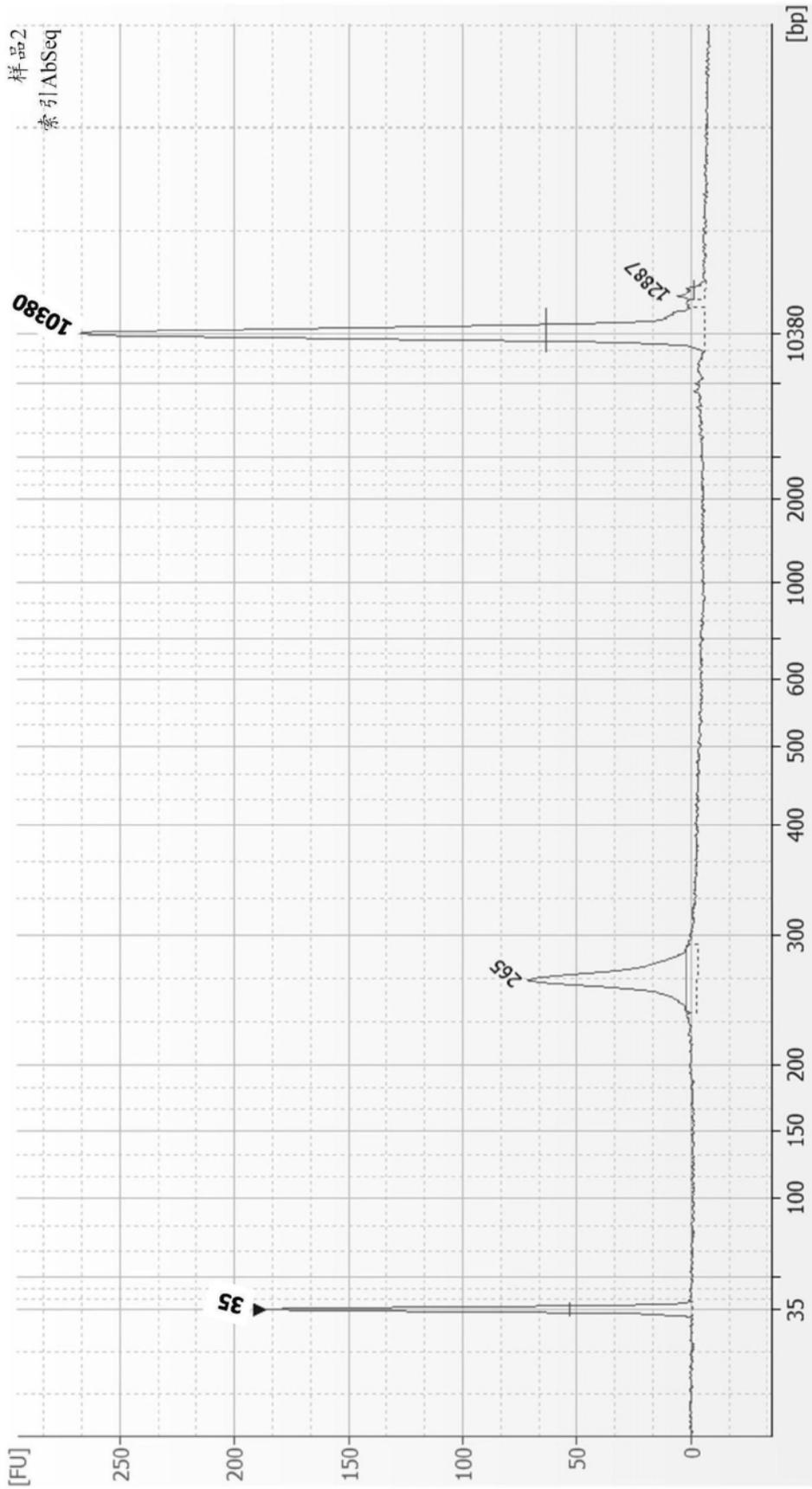


图12C

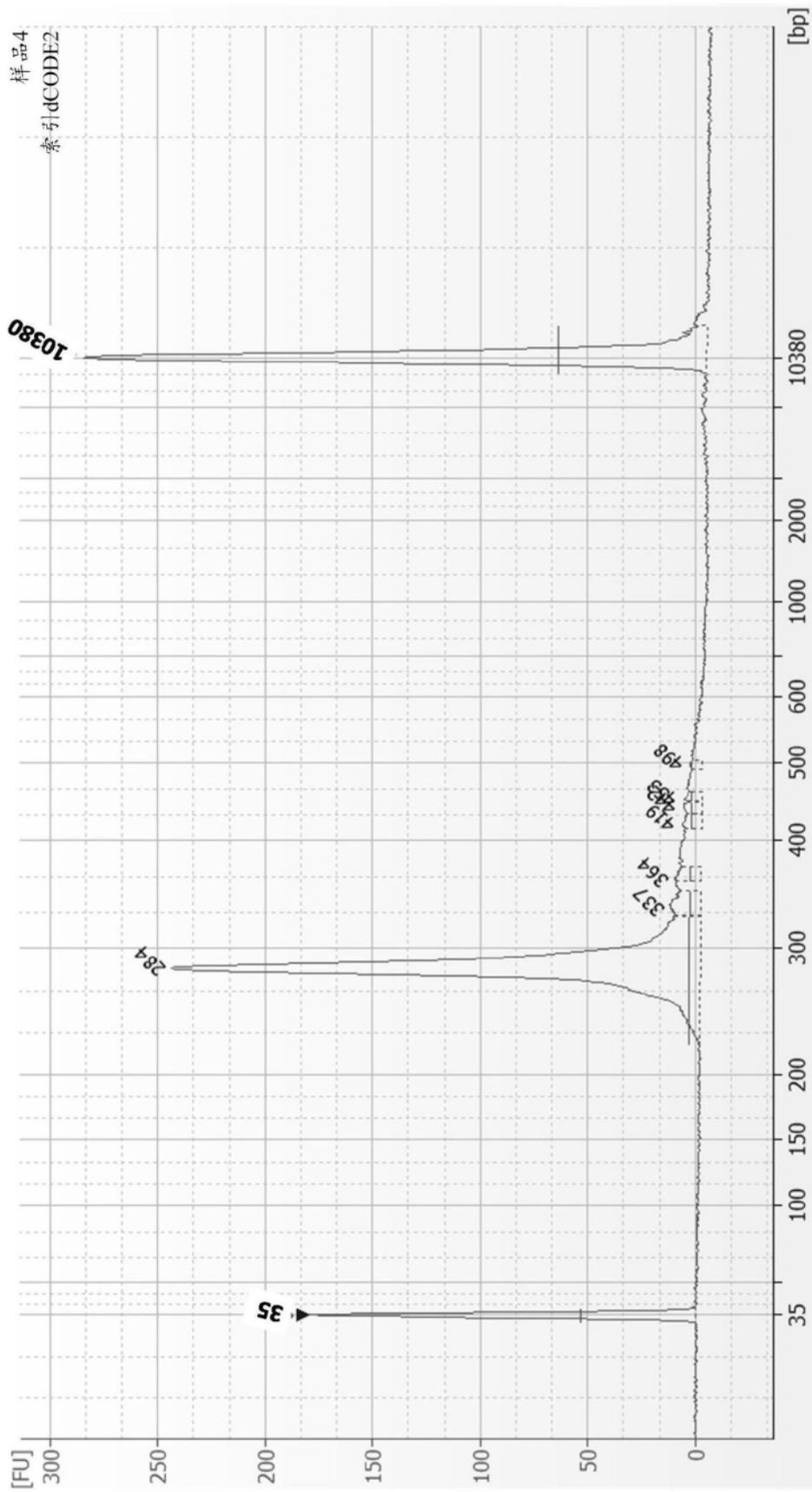


图12D

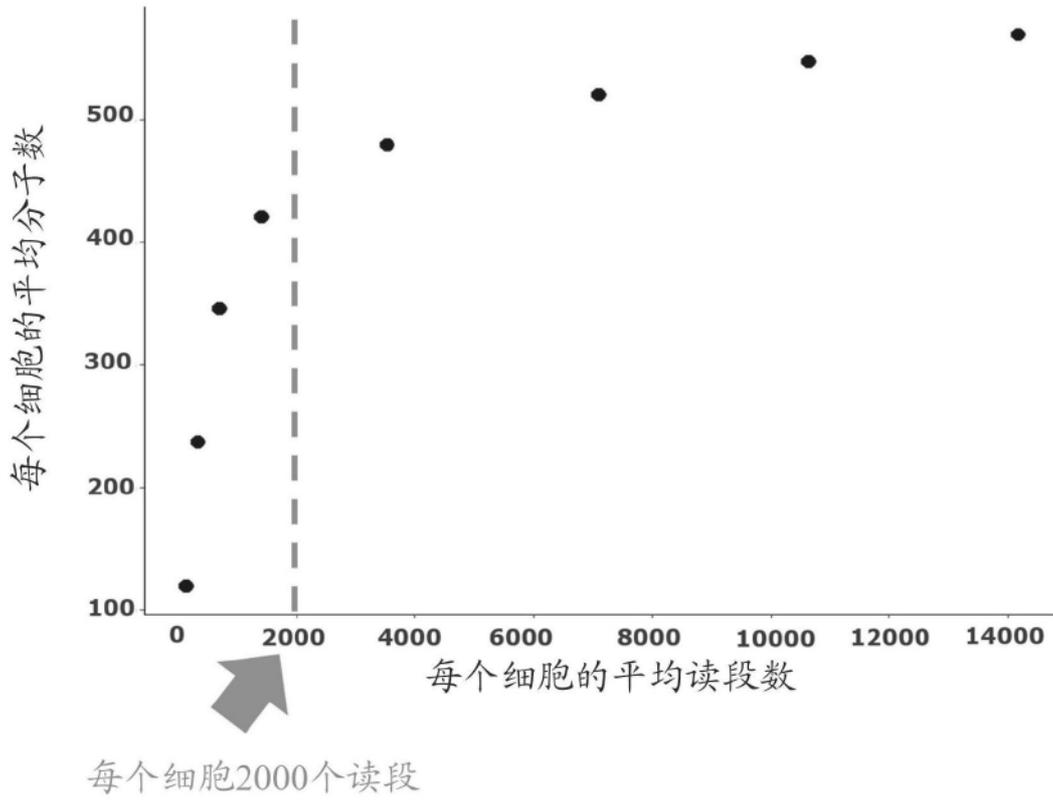


图13A

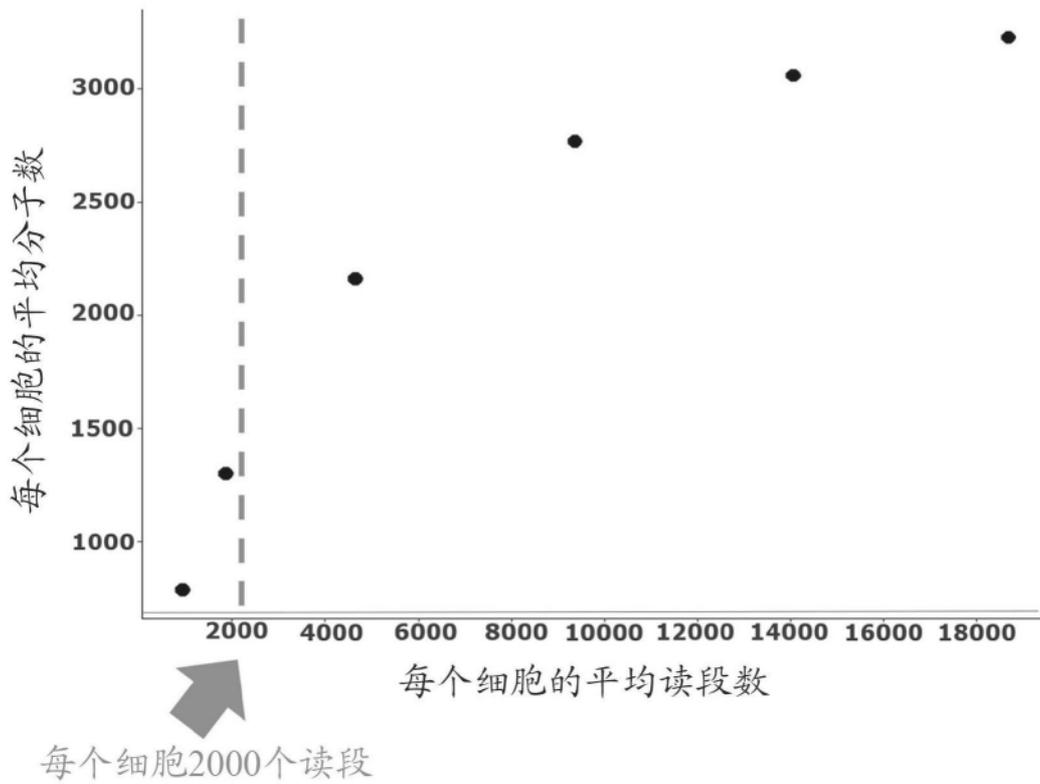


图13B