(19)中华人民共和国国家知识产权局

(12)发明专利

(10)授权公告号 CN 105705659 B
(45)授权公告日 2019.11.29

(21)申请号 201480059505.3

(22)申请日 2014.08.28

(65)同一申请的已公布的文献号
申请公布号 CN 105705659 A

(43)申请公布日 2016.06.22

(30)优先权数据
61/871,232 2013.08.28 US
61/952,036 2014.03.12 US
62/012,237 2014.06.13 US

(85)PCT国际申请进入国家阶段日
2016.04.28

(86)PCT国际申请的申请数据
PCT/US2014/053301 2014.08.28

(87)PCT国际申请的公布数据
WO2015/031691 EN 2015.03.05

(73)专利权人 贝克顿迪金森公司
地址 美国新泽西州

(72)发明人 克里斯蒂娜·范
斯蒂芬·P·A·福多尔
格伦·K·弗
杰弗里·理查德·菲舍
朱莉·威廉米

(74)专利代理机构 北京安信方达知识产权代理有限公司 11262
代理人 郑霞

(51)Int.Cl.
C12Q 1/6874(2018.01)

审查员 夏向东

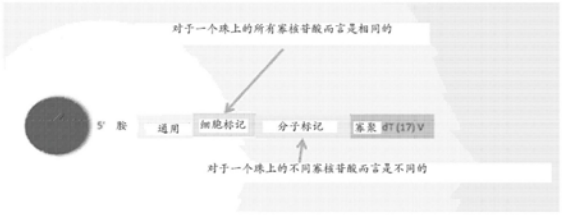
权利要求书2页 说明书182页 附图136页

(54)发明名称

大规模平行单细胞分析

(57)摘要

本披露提供了用于单细胞的多重核酸分析的方法、组合物和试剂盒。这些方法、组合物和系统可以用于大规模平行单细胞测序。这些方法、组合物和系统可以用于同时分析数以千计的细胞。这些数以千计的细胞可以包括细胞的混合群体(例如,不同类型或亚型、不同大小的细胞)。



CN 105705659 B

1.一种方法,包括:

(a)将细胞随机分配到基质上的单独的孔中,这样使得多个所述孔含有单细胞,并且其中所述基质包括至少1,000个孔;

(b)将单珠与含有所述单细胞的单孔进行关联,其中所述单珠包含多个寡核苷酸,所述多个寡核苷酸中的每一个包含细胞标记、分子标记以及靶结合区域;

(c)随机地标记来自所述单细胞的靶分子;并且

(d)对所述靶分子的数目进行估计。

2.如权利要求1所述的方法,其中所述随机地标记靶分子包括将来自所述单细胞的靶分子杂交或连接至所述多个寡核苷酸中的寡核苷酸的所述靶结合区域以产生标记的靶核酸,和任选地扩增所述标记的靶核酸。

3.如权利要求1所述的方法,其中所述细胞包括选自下组的细胞,该组由以下各项组成:稀有细胞、肿瘤细胞、来自人的细胞、来自组织的细胞、来自肿瘤的细胞、被病毒多核苷酸感染的细胞及其任何组合。

4.如权利要求2所述的方法,其中所述靶核酸是mRNA。

5.如权利要求1至4中任一项所述的方法,其中所述关联发生在裂解所述细胞之前。

6.如权利要求2所述的方法,其中在所述杂交或连接之前裂解所述细胞。

7.如权利要求4所述的方法,进一步包括从所述单孔中移出所述单珠。

8.如权利要求7所述的方法,其中所述移出包括用磁体移出。

9.如权利要求1所述的方法,其中所述多个寡核苷酸各自包括通用标记。

10.如权利要求9所述的方法,其中所述多个寡核苷酸中的单独的寡核苷酸具有不同的分子标记。

11.如权利要求9所述的方法,其中所述多个寡核苷酸中的单独的寡核苷酸具有相同的细胞标记。

12.如权利要求10至11中任一项所述的方法,其中所述多个孔的不同孔中的珠上的所述多个寡核苷酸中的单独的寡核苷酸具有不同的细胞标记。

13.如权利要求9至11中任一项所述的方法,其中所述靶结合区域包括选自下组的序列,该组由以下各项组成:寡聚-dT序列、基因特异性序列以及随机多聚体序列。

14.如权利要求12所述的方法,其中所述靶结合区域包含选自由以下组成的组的序列:寡聚-dT序列、基因特异性序列和随机多聚体序列。

15.如权利要求1所述的方法,其中所述随机地标记靶分子包括逆转录所述靶分子,从而产生经标记的靶cDNA。

16.如权利要求15所述的方法,其中所述逆转录是在聚池的珠上完成的。

17.如权利要求15或16中任一项所述的方法,其中所述经标记的靶cDNA附接至所述单珠。

18.如权利要求15至16中任一项所述的方法,进一步包括在所述经标记的靶cDNA上进行第二链合成,从而产生经标记的双链靶多核苷酸。

19.如权利要求18所述的方法,进一步包括扩增所述经标记的双链靶多核苷酸,从而产生经标记的靶扩增子。

20.如权利要求19所述的方法,其中所述扩增是用通用引物和随机多聚体引物进行。

21.如权利要求19所述的方法，其中所述扩增是用通用引物和基因特异性引物进行。

22.如权利要求19所述的方法，其中所述扩增包括执行巢式PCR反应。

23.如权利要求19所述的方法，其中所述扩增包括扩增一个或更多个所述细胞的全转录组。

24.如权利要求19至23中任一项所述的方法，其中所述估计包括确定所述经标记的靶扩增子的至少一部分的序列。

25.如权利要求19至23中任一项所述的方法，其中所述估计包括确定所述经标记的靶扩增子的至少一部分的和所述寡核苷酸的至少一部分的序列。

26.如权利要求1至3中任一项所述的方法，其中所述估计包括对与不同靶多核苷酸相关联的唯一性寡核苷酸数目/细胞进行计数。

大规模平行单细胞分析

[0001] 交叉引用

[0002] 本申请要求于2014年6月13日提交的美国临时申请号62/012,237、于2014年3月12日提交的美国临时申请号61/952,036以及于2013年8月28日提交的美国临时申请号61/871,232的权益，将这些申请中的每一篇通过引用以其全部内容结合在此。

技术领域

[0003] 本申请涉及但不限于大规模平行单细胞分析。

[0004] 背景

[0005] 多细胞团(如组织和肿瘤)可以包括异质细胞环境。这些复杂的细胞环境通常可以展示出多种表型，这些表型可以指示多种基因型。从多细胞复杂性中提炼到单细胞可变性是理解多细胞异质性的重要方面。这种理解在开发用于对抗具有多种抗性基因型的疾病的治疗方案中可以是重要的。

[0006] 发明概述

[0007] 提供的一个方面是一种方法，该方法包括获得包含多个细胞的样品；使用以下项对来自该多个细胞的第一细胞的和来自该多个细胞的第二细胞的两个或更多个多核苷酸分子、其互补体或来自其的反应产物的至少一部分进行标记：对该第一细胞具有特异性的第一相同细胞标记和对该第二细胞具有特异性的第二相同细胞标记；以及对该两个或更多个多核苷酸分子、其互补体或来自其的反应产物各自具有特异性的分子标记，其中来自该第一细胞的该两个或更多个多核苷酸分子、其互补体或来自其的反应产物的每个分子标记相对于彼此都是唯一的，并且其中来自第二细胞的该两个或更多个多核苷酸分子、其互补体或来自其的反应产物的每个分子标记相对于彼此都是唯一的。在一些实施例中，该方法进一步包括对两个或更多个多核苷酸分子、其互补体或来自其的反应产物的至少一个部分测序。在一些实施例中，该方法进一步包括分析来自该测序的序列数据，以对这些细胞的特定细胞中的这些多核苷酸的多个个体分子进行鉴定。在一些实施例中，这些细胞是癌细胞。在一些实施例中，这些细胞感染病毒多核苷酸。在一些实施例中，这些细胞是细菌或真菌。在一些实施例中，该测序包括以至少100个碱基的读取长度进行测序。在一些实施例中，该测序包括以至少500个碱基的读取长度进行测序。在一些实施例中，这些多核苷酸分子是mRNA或微小RNA，并且其互补体及其反应产物是这些mRNA或微小RNA的互补体和来自其的反应产物。在一些实施例中，这些分子标记在珠上。在一些实施例中，对单个细胞具有特异性的标记是在珠上。在一些实施例中，对单个细胞具有特异性的标记和这些分子标记是在珠上。在一些实施例中，该方法至少部分地在乳液中进行。在一些实施例中，该方法至少部分地在阵列的孔或微孔中进行。在一些实施例中，检测到与疾病或病症相关的多核苷酸的存在。在一些实施例中，该疾病或病症是癌症。在一些实施例中，检测到微小RNA、其互补体或来自其的反应产物的至少一部分。在一些实施例中，该疾病或病症是病毒感染。在一些实施例中，该病毒感染来自包膜病毒。在一些实施例中，该病毒感染来自无包膜病毒。在一些实施例中，该病毒含有双链的病毒DNA。在一些实施例中，该病毒含有单链的病毒DNA。在一些

实施例中，该病毒选自下组，该组由以下各项组成：痘病毒、疱疹病毒、水痘带状疱疹病毒、巨细胞病毒、EB病毒、嗜肝DNA病毒、乳多孔病毒、多瘤病毒及其任何组合。在一些实施例中，第一细胞来自不患有疾病或病症的人并且第二细胞来自患有该疾病或病症的人。在一些实施例中，这些人是不同的。在一些实施例中，这些人是相同的，但是细胞是在不同时间点取的。在一些实施例中，第一细胞来自患有该疾病或病症的人并且第二细胞来自同一人。在一些实施例中，该样品中的细胞包括来自组织或器官的细胞。在一些实施例中，该样品中的细胞包括来自以下项的细胞：胸腺、白细胞、红细胞、肝细胞、脾细胞、肺细胞、心脏细胞、脑细胞、皮肤细胞、胰腺细胞、胃细胞、来自口腔的细胞、来自鼻腔的细胞、结肠细胞、小肠细胞、肾细胞、来自腺体的细胞、脑细胞、神经细胞、神经胶质细胞、眼细胞、生殖器官细胞、膀胱细胞、配子细胞、人类细胞、胎儿细胞、羊膜细胞或其任何组合。

[0008] 提供的一个方面是一种固体支持物，该固体支持物包括多个寡核苷酸，其各自包括细胞标记和分子标记，其中该多个寡核苷酸的每个细胞标记都是相同的，并且该多个寡核苷酸的每个分子标记是不同的；并且其中该固体支持物是珠，该细胞标记对该固体支持物具有特异性，该固体支持物当被放置在三维笛卡尔坐标系的中心时其寡核苷酸延伸到八个八分圆的至少七个中，或其任何组合。在一些实施例中，该多个寡核苷酸进一步包括以下各项中的至少一种：样品标记；通用标记；以及靶核酸结合区。在一些实施例中，该固体支持物包括该靶核酸结合区，其中该靶核酸结合区包括选自下组的序列，该组由以下各项组成：基因特异性序列、寡聚-dT序列、随机多聚体及其任何组合。在一些实施例中，该固体支持物进一步包括靶核酸或其互补体。在一些实施例中，该固体支持物包括多个靶核酸或其互补体，该多个靶核酸或其互补体包括生物体转录组的从约0.01％至约100％的转录物或其互补体，或生物体基因组的从约0.01％至约100％的基因或其互补体。在一些实施例中，该多个寡核苷酸的细胞标记包括利用第一标记连接序列与第二随机序列连接的第一随机序列；并且该多个寡核苷酸的分子标记包括随机序列。在一些实施例中，该固体支持物选自下组，该组由以下各项组成：聚二甲基硅氧烷(PDMS)固体支持物、聚苯乙烯固体支持物、玻璃固体支持物、聚丙烯固体支持物、琼脂糖固体支持物、明胶固体支持物、磁性固体支持物、普朗尼克(pluronic)固体支持物及其任何组合。在一些实施例中，该多个寡核苷酸包括包含接头官能团的接头，并且该固体支持物包括固体支持物官能团；其中该固体支持物官能团和接头官能团彼此连接。在一些实施例中，该接头官能团和该固体支持物官能团单独地选自下组，该组由以下各项组成：C6、生物素、链霉亲和素、一种或多种伯胺、一种或多种醛、一种或多种酮及其任何组合。在一些实施例中，该多个寡核苷酸的分子标记包括至少15个核苷酸。

[0009] 提供的一个方面是一种试剂盒，该试剂盒包括在此描述的固体支持物中的任一种和使用说明书。在一些实施例中，该试剂盒进一步包括孔。在一些实施例中，该孔被包含在阵列中。在一些实施例中，该孔是微孔。在一些实施例中，该试剂盒进一步包括缓冲液。在一些实施例中，该试剂盒被包含在包装中。在一些实施例中，该包装是盒子。在一些实施例中，该包装或盒子的体积是2立方英尺或更小。在一些实施例中，该包装或盒子的体积是1立方英尺或更小。

[0010] 提供的一个方面是一种乳液，该乳液包括在此描述的固体支持物中的任一种。

[0011] 提供的一个方面是一种组合物，该组合物包括孔和在此描述的固体支持物中的任一种。

[0012] 提供的一个方面是一种组合物，该组合物包括细胞和在此描述的固体支持物中的任一种。

[0013] 在一些实施例中，该乳液或组合物进一步包括细胞。在一些实施例中，该细胞是单细胞。在一些实施例中，该孔是微孔。在一些实施例中，该微孔的体积范围是从约1,000μm^{3}至约120,000μm^3。

[0014] 提供的一个方面是一种方法，该方法包括使样品与在此披露的任何固体支持物接触，使来自该样品的靶核酸与该多个寡核苷酸中的寡核苷酸杂交。在一些实施例中，该方法进一步包括扩增靶核酸或其互补体。在一些实施例中，该方法进一步包括对靶核酸或其互补体测序，其中该测序包括对与靶核酸或其互补体结合的寡核苷酸的分子标记进行测序。在一些实施例中，该方法进一步包括确定靶核酸或其互补体的量，其中该确定包括定量靶核酸或其互补体的水平；对包含相同分子标记的多个序列计数；或其组合。在一些实施例中，该方法不包括比对任何相同的分子标记或任何相同的细胞标记。在一些实施例中，该扩增包括逆转录靶核酸。在一些实施例中，该扩增利用选自下组的方法，该组由以下各项组成：PCR、巢式PCR、定量PCR、实时PCR、数字PCR及其任何组合。在一些实施例中，该扩增在固体支持物上直接进行；在从该固体支持物转录的模板上进行；或其组合。在一些实施例中，该样品包括细胞。在一些实施例中，该细胞是单细胞。在一些实施例中，该接触发生在孔中。在一些实施例中，该孔是微孔并且被包含在微孔阵列中。

[0015] 提供的一个方面是一种装置，该装置包括多个微孔，其中该多个微孔中的每个微孔的体积范围是从约1,000μm^{3}至约120,000μm^3。在一些实施例中，该多个微孔中的每个微孔的体积是约20,000μm^3。在一些实施例中，该多个微孔包括从约96至约200,000个微孔。在一些实施例中，这些微孔被包含在材料层中。在一些实施例中，至少约10％的微孔进一步包括细胞。在一些实施例中，该装置进一步包括在此描述的固体支持物中的任一种。

[0016] 提供的一个方面是一种设备，该设备包括在此描述的装置中的任一种和液体处理器。在一些实施例中，液体处理器在约一秒内将液体递送至该多个微孔。在一些实施例中，液体处理器将液体从单输入端口递送至该多个微孔。在一些实施例中，该设备进一步包括磁体。在一些实施例中，该设备进一步包括以下项中的至少一种：入口、出口、泵、阀、通气孔、储器、样品收集室、温度控制设备或其任何组合。在一些实施例中，该设备包括样品收集室，其中样品收集室是从该设备可移除的。在一些实施例中，该设备进一步包括光学成像仪。在一些实施例中，光学成像仪产生用于控制液体处理器的输出信号。在一些实施例中，该设备进一步包括被配置成进行寡核苷酸的聚合酶链式反应(PCR)扩增的热循环机构。

[0017] 提供的一个方面是一种产生临床诊断测试结果的方法，该方法包括用以下项产生该临床诊断测试结果：在此描述的任何装置或设备；在此描述的任何固体支持物；在此描述的任何方法；或其任何组合。在一些实施例中，经由通信媒体传输临床诊断测试结果。

[0018] 提供的一个方面是一种制造在此描述的固体支持物中的任一种的方法，该方法包括将以下项附接到固体支持物：第一多核苷酸，其包含细胞标记的第一部分，和第一接头；并且接触第二多核苷酸，其包含细胞标记的第二部分、与第一接头互补的序列和分子标记。在一些实施例中，第三多核苷酸进一步包括靶核酸结合区。

[0019] 在一些实施例中，乳液、微孔或孔仅含有一个细胞。在一些实施例中，从1至2,000,000个乳液、微孔或孔各自仅含有一个细胞。在一些实施例中，该方法包括将至多一个细胞

分配到每一乳液、微孔或孔中。在一些实施例中，将单一固体支持物和单细胞分配到乳液、微孔或孔中。在一些实施例中，从1至2,000,000个乳液、微孔或孔各自已经向其中分配了一个细胞和一个固体支持物。在一些实施例中，该方法包括分配至多一个固体支持物/乳液、微孔或孔。在一些实施例中，该方法包括向从1至2,000,000种微孔、乳液或孔中的每者分配一个固体支持物和一个细胞。在一些实施例中，细胞分配是随机或非随机的。在一些实施例中，细胞分配是随机的。在一些实施例中，通过细胞分选仪分配细胞。在一些实施例中，通过使一个或多个孔、微孔或乳液与经稀释细胞的稀释溶液接触来分配细胞，这样使得至多一个细胞被分配到该一个或多个孔、微孔或乳液中。

[0020] 在一些实施例中，这些靶特异性区、该多个寡核苷酸的靶特异性区或该两个或更多个多核苷酸分子的靶特异性区包括与靶面板中的两个或更多个靶互补的序列。在一些实施例中，该靶面板中的两个或更多个靶是生物标志物。在一些实施例中，这些生物标志物是疾病或病症的生物标志物。在一些实施例中，该疾病或病症是癌症、感染、病毒感染、炎性疾病、神经退行性疾病、真菌疾病、细菌感染或其任何组合。在一些实施例中，该面板包括从：2-50,000、2-40,000、2-30,000、2-20,000、2-10,000、2-9000、2-8,000、2-7,000、2-6,000、2-5,000、2-1,000、2-800、2-700、2-600、2-500、2-400、2-300、2-200、2-100、2-75、2-50、2-40、2-30、2-20、2-10或2-5种生物标志物。

[0021] 通过引用结合

[0022] 本说明书中提到的所有出版物、专利和专利申请都通过引用结合在此，如同每一单独的出版物、专利或专利申请具体且单独地指明通过引用结合在此。

[0023] 附图简要说明

[0024] 本发明的新颖特征在所附权利要求书中具体阐述。通过参考对说明性实施例进行阐述的以下详细说明，将获得对本发明的特征和优点的更好理解，在这些实施例中利用了本发明的原理，并且在这些附图中：

[0025] 图1描绘了与示例性寡核苷酸缀合的示例性固体支持物。

[0026] 图2A-C描绘了用于使用分离-聚池合成(split-pool synthesis)来合成偶联寡核苷酸的珠的示例性工作流程。

[0027] 图3描绘了偶联寡核苷酸的示例性珠。

[0028] 图4说明了微孔阵列的示例性实施例。

[0029] 图5描绘了固体支持物在微孔阵列中的示例性分布。

[0030] 图6A-C示出了向微孔阵列上示例性分配细胞。图6A示出了K562细胞(大细胞尺寸)的分布。图6B示出了拉莫斯(Ramos)细胞(小细胞尺寸)的分布。图6C示出了向微孔阵列上分配拉莫斯细胞和偶联寡核苷酸的珠，用实线箭头指向拉莫斯细胞并且虚线箭头指向偶联寡核苷酸的珠。

[0031] 图7示出了微孔体积、固体支持物体积以及从裂解获得的生物材料的量的示例性统计。

[0032] 图8A-C说明了珠帽密封的示例性实施例。图8A-B示出了微阵列孔的图像，其中细胞和寡核苷酸珠被分配到微阵列孔的孔中并且较大的交联葡聚糖珠用于密封这些孔。点线箭头指向细胞，虚线箭头指向偶联寡核苷酸的珠并且实线箭头指向交联葡聚糖珠。图8C描绘了沉积在孔内的细胞和寡核苷酸珠(例如，寡聚珠)的示意图，其中交联葡聚糖珠用于密

封该孔。

[0033] 图9描绘了比较从微孔和管扩增的GAPDH和RPL19的扩增效率的条形图。灰条表示来自微孔的数据。白条表示来自管的数据。

[0034] 图10描绘了比较三个不同基因直接在固体支持物上的扩增特异性的琼脂糖凝胶。

[0035] 图11A-I示出了测序结果的图形表示。

[0036] 图12A-C示出了分别对于仅K562样品、仅拉莫斯样品以及K562+拉莫斯混合物样品的测序结果直方图。

[0037] 图12D-E示出了分别针对仅拉莫斯样品和仅K562样品，表3中列出的基因的拷贝数的图。

[0038] 图12F-I示出了单个基因的拷贝数。

[0039] 图12J-M示出了对于具有100个唯一性条码组合的珠，唯一性分子数目/基因(y-轴)的图。

[0040] 图12N-O示出了两种珠的放大图，其描绘这两种细胞类型的基因表达谱的总体模式。

[0041] 图12P示出了基于768个珠的基因表达谱的主成分分析的结果的散点图，其中来自K562+拉莫斯混合物样品的每个珠具有>30个分子。

[0042] 图12Q-R示出了分别针对K562样细胞(基于图12P的第一主成分的左边的珠)和拉莫斯样细胞(基于图12P的第一主成分的右边的珠)，拷贝数/扩增子/珠的直方图。

[0043] 图12S-T示出了分别针对K562样细胞(基于图12P的第一主成分的左边的珠)和拉莫斯样细胞(基于图12P的第一主成分的右边的珠)而言的单独基因拷贝数/珠或/单细胞。

[0044] 图13A描绘了小鼠细胞和拉莫斯细胞的总体基因表达模式。

[0045] 图13B-C示出了分别基于高密度样品和低密度样品的基因表达谱的主成分分析的结果的散点图。

[0046] 图13D-E描绘了分别针对来自高密度样品的拉莫斯样细胞和小鼠样细胞的以下图：读数/条码(bc)组合(y-轴)比对以分子总数目/bc组合分选的唯一性条码组合(x-轴)。

[0047] 图13F-G描绘了分别针对来自高密度样品的拉莫斯样细胞和小鼠样细胞的以下图：分子数目/条码(bc)组合(y-轴)比对以分子总数目/bc组合分选的唯一性条码组合(x-轴)。

[0048] 图13H-I描绘了分别针对来自低密度样品的拉莫斯样细胞和小鼠样细胞的以下图：读数/条码(bc)组合(y-轴)比对以分子总数目/条码组合分选的唯一性条码组合(x-轴)。

[0049] 图13J-K描绘了分别针对来自低密度样品的拉莫斯样细胞和小鼠样细胞的以下图：分子数目/条码组合(y-轴)比对以分子总数目/条码组合分选的唯一性条码组合(x-轴)。

[0050] 图14示出了描绘在X-轴上的基因和读数数目的log10的图。

[0051] 图15A示出了每个三部分细胞标记(例如，细胞条码)的检测到的基因的分布图。图15B示出了每珠(表达基因面板)的检测到的唯一性分子的分布图。

[0052] 图16描绘了基于与细胞条码相关的基因的细胞簇。

[0053] 图17A-D示出了单核细胞特异性标志物的分析。图17E示出了图16中描绘的细胞

簇。

[0054] 图18A-B示出了T细胞特异性标志物的分析。图18C示出了图16中描绘的细胞簇。

[0055] 图19A-B示出了CD8+T细胞特异性标志物的分析。图19C示出了图16中描绘的细胞簇。

[0056] 图20A示出了CD4+T细胞特异性标志物的分析。图20B示出了图16中描绘的细胞簇。

[0057] 图21A-D示出了自然杀伤细胞(NK)细胞特异性标志物的分析。图21E示出了图16中描绘的细胞簇。

[0058] 图22A-E示出了B细胞特异性标志物的分析。图22F示出了图16中描绘的细胞簇。

[0059] 图23A-F示出了Toll样受体的分析。Toll样受体主要由单核细胞和一些B细胞表达。图23G示出图16中描绘的细胞簇。

[0060] 图24描绘了基因比对读数数目log10的图。

[0061] 图25A-D示出了对于两个基因而言,分子条码比对读数数目或读数数目log10的图。

[0062] 图26A示出了表达的面板中的基因数目/细胞条码比对唯一性细胞条码数目/单细胞的图。图26B示出了检测的唯一性分子数目/珠比对细胞数目频率/负载给定数目分子的唯一性细胞条码的直方图。图26C示出了检测的唯一性GAPDH分子数目/珠比对细胞数目频率/负载给定数目分子的唯一性细胞条码的直方图。

[0063] 图27示出了856个细胞的散点图。

[0064] 图28示出了前100(就检测的分子的总数而言)的表达的热图。

[0065] 图29示出了实例12的工作流程。

[0066] 图30示出了实例13的工作流程。

[0067] 图31A-C。单细胞在含有两种不同细胞类型的受控混合物中的聚簇。A.通过12个基因表达的主成分分析的,K562和拉莫斯细胞的1:1混合物的聚簇。双标图示出了两个不同簇,其中一个簇表达拉莫斯特异性基因并且另一个簇表达K562特异性基因。B.使用具有111个基因的面板的在来自健康个体的原代B细胞的背景下含有小百分比的拉莫斯细胞的混合物的主成分分析。每个数据点的颜色指示在整个基因面板中检测到的唯一性转录物分子的总数。1198个细胞中的一组18个细胞(画圈的)展示出不同的基因表达谱并且转录水平要高得多。C.示出了每个基因在图31B的样品的前100个细胞中的表达水平的热图,利用在基因面板中检测到的转录物分子的总数进行排名。经由就相关性而言的分层聚簇使基因排序。前18个细胞(由水平红条指示)优先表达一组已知与滤泡性淋巴瘤相关的基因,如通过竖直红条所指示。

[0068] 图31D。原代B细胞的PCA分析,具有拉莫斯细胞掺入。每个数据点(单细胞)的颜色指示每个细胞携带的针对特定基因的转录物分子的数目的对数。前7行:优先由可能是拉莫斯细胞的18细胞亚群表达的基因。第一行的基因(从左到右)包括GAPDH、TCL1A、MKI67以及BCL6。第二行的基因(从左到右)包括MYC、CCND3、CD81以及GNAI2。第三行的基因(从左到右)包括IGBP1、CD20、BLNK以及DOCK8。第四行的基因(从左到右)包括IRF4、CD22、IGHM以及AURKB。第五行的基因(从左到右)包括CD38、CD10、LEF1以及AICDA。第六行的基因(从左到右)包括CD40、CD27、IL4R以及PRKCD。第七行的基因(从左到右)包括RGS1、MCL1、CD79a以及HLA-DRA。最后一行:优选由原代B细胞亚群表达但是未特异性富集在那18个细胞中的基因。

最后一行的基因(从左到右)包括IL6、Cd23a、CCR7以及CXCR5。

[0069] 图32GAPDH的表达。颜色指示每细胞观察到的唯一性转录物分子数目的自然对数。

[0070] 图33A-F示出了单核细胞相关基因的主成分分析(PCA)。图33A示出了CD16的PCA。图33B示出了CCRvarA的PCA。图33C示出了CD14的PCA。图33D示出了S100A12的PCA。图33E示出了CD209的PCA。图33F示出了IFNGR1的PCA。

[0071] 图34A-B示出了泛T细胞标志物(CD3)的主成分分析(PCA)。图34A示出了CD3D的PCA并且图34B示出了CD3E的PCA。

[0072] 图35A-E示出了CD8T细胞相关基因的主成分分析(PCA)。图35A示出了CD8A的PCA。图35B示出了EOMES的PCA。图35C示出了CD8B的PCA。图35D示出了PRF1的PCA。图35E示出了RUNX3的PCA。

[0073] 图36A-C示出了CD4T细胞相关基因的主成分分析(PCA)。图36A示出了CD4的PCA。图36B示出了CCR7的PCA。图36C示出了CD62L的PCA。

[0074] 图37A-F示出了B细胞相关基因的主成分分析(PCA)。图37A示出了CD20的PCA。图37B示出了IGHD的PCA。图37C示出了PAX5的PCA。图37D示出了TCL1A的PCA。图37E示出了IGHM的PCA。图37F示出了CD24的PCA。

[0075] 图38A-C示出了自然杀伤细胞相关基因的主成分分析(PCA)。图38A示出了KIR2DS5的PCA。图38B示出了CD16的PCA。图38C示出了CD62L的PCA。

[0076] 图39通过由塞托细胞(CytoSeq Cells)测定的81个基因的PCA分析在人类PBMC样品(632个细胞)中同时鉴定主要细胞类型,其中高度相关的细胞表达谱由类似颜色编码。

[0077] 图40A-B PBMC样品的单细胞基因表达谱的相关分析。40A.示出了样品中的跨632个细胞的两两相关系数的矩阵。将这些细胞排序,这样使得具有高度相关的基因表达谱的那些被分组在一起。40B.示出了每个细胞对每个基因的表达的热图。将这些细胞(列)按与以上相关矩阵相同的方式排序。将这些基因(行)排序,这样使得跨细胞具有高度相似的表达模式的基因被分组在一起。可以通过细胞共表达的该组基因鉴定每个细胞簇的细胞类型。在每个主要细胞簇内,就基因表达而言存在实质程度的异质性。

[0078] 图41数据表示来自同一供体的PBMC样品的重复实验的731个细胞的数据。将具有类似基因表达谱(基于使用相关系数的分层聚簇)的细胞用类似颜色绘出。

[0079] 图42示出了证明基因之间的基因表达谱的相关性的热图。

[0080] 图43CytoSeq的描述。A.CytoSeq的实验程序。B.附接至珠的寡核苷酸的结构。

[0081] 图44剖析CD3+T细胞的亚群。A.供体1未刺激样品的PCA揭示了细胞的两个主要分支。每个细胞内的特定基因的表达水平(唯一性转录物分子的log)用颜色指示。辅助T细胞相关细胞因子和效应基因富集在较下支的细胞中,而细胞毒性T细胞相关基因富集在较上支的细胞中。此处示出了代表性基因。第一行示出了辅助T细胞相关基因并且包括(从左到右)CD4、SELL和CCR7。第二行示出了细胞毒性T细胞相关基因并且包括(从左到右)CD8A、NKG2D和EOMES。B.供体1抗-CD3/抗-CD28刺激样品的PCA,示出了指定基因的表达向表示辅助T细胞和细胞毒性T细胞的两个主要分支之一的富集。这些基因以低量存在于未刺激样品中。前两行示出了已知与活化的T细胞相关的基因并且在第一行中包括(从左到右)IRF4、CD69和MYC并且在第二行中包括(从左到右)GAPDH、TNF和IFNG。第三行示出了已知与活化的辅助T细胞相关的基因并且包括(从左到右)IL2、LTA和CD40LG。第四行示出了已知与活化的

细胞毒性T细胞相关的基因并且包括(从左到右)CCL4、CCL3和GZMB。C.促成了以下基因的总体表达水平的细胞数目,这些基因当在汇总数据中比较刺激与未刺激样品时展示出大的倍数变化。对于若干细胞因子(红色箭头)而言,来自仅小数目的细胞的贡献负责整个群体中的大的总体基因表达变化。

[0082] 图45.在两个供体中已经经历用抗-CD28/抗-CD3珠刺激的T细胞样品和相应的未刺激样品的PCA绘图,重点是以下基因的表达,这些基因在未刺激样品中清楚显示出在辅助或细胞毒性亚群中的优先表达。每个数据点(单细胞)的颜色指示每细胞针对指定基因的log(唯一性转录物分子的数目)。对于每个供体中的每对刺激和未刺激图,将颜色范围调节成相同的。A.已知与辅助T细胞和细胞毒性T细胞均相关的基因。B.已知与细胞毒性T细胞相关的基因。C.已知与辅助T细胞相关的基因。

[0083] 图46A-D在两个供体中已经经历用抗-CD28/抗-CD3珠刺激的T细胞样品和相应的未刺激样品的PCA绘图,重点是以下基因的表达,这些基因表达于刺激的样品中但以低或不可测水平表达于未刺激的样品中。每个数据点(单细胞)的颜色指示每细胞针对指定基因的log(唯一性转录物分子的数目)。对于每个供体中的每对刺激和未刺激图,将颜色范围调节成相同的。46A和46D.活化后由两个分支的细胞表达的基因。46B.活化后优先表达于较上支细胞中的基因。已知这些基因与活化的细胞毒性T细胞相关。46C.活化后优先表达于较下支细胞中的基因。已知这些基因与活化的辅助T细胞相关。

[0084] 图47来自供体1的未刺激的CD3+T细胞的数据聚簇示出了CD4和CD8细胞的分离以及一组表达颗粒酶K和颗粒酶A但几乎不表达CD8的细胞。顶部:示出了每对细胞之间的相关性的热图。高度相关的细胞被分组在一起。底部:示出了每个细胞的每个基因的表达水平的热图。经由双向分层聚簇使细胞和基因排序。

[0085] 图48.类似于图47,但示出了来自供体1的抗-CD3/抗-CD28刺激的CD3+T细胞样品的数据。顶部:示出了每对细胞之间的相关性的热图。高度相关的细胞被分组在一起。底部:示出了每个细胞的每个基因的表达水平的热图。经由双向分层聚簇使细胞和基因排序。

[0086] 图49A-C在供体1中,与未刺激样品相比,在抗CD28/抗CD3刺激的样品中观察到不同细胞因子的大的总体倍数变化。A-B:这些细胞因子的大的倍数变化在很大程度上仅由几个单细胞(用正方形或圆圈封闭的点)促成。许多这些细胞因子由相同的小数目的细胞促成。C:针对辅助T细胞的Th2和Th17亚群而言,这些细胞因子的共表达模式与特征细胞因子组合相一致。

[0087] 图50A-B.剖析CD8+T细胞的亚群。A.CytoSeq数据的聚簇定义了CD8+细胞的两个主要组-一个组表达中心记忆/初始细胞具有的基因,并且另一个组表达效应记忆细胞/效应细胞具有的基因。此处示出了供体2的未刺激的样品的数据。顶部:示出了每对细胞之间的相关性的热图。底部:示出了每个细胞中的每个基因的表达水平的热图。经由双向分层聚簇使细胞和基因排序。B.通过用CMV肽池刺激之后γ干扰素(IFNG)在来自两个供体的CD8+T细胞中的表达来鉴定稀有抗原特异性T细胞。在2D主成分空间上绘出每个细胞。表达IFNG的细胞(画圈的)通常属于以下者之列,它们在面板中具有最多总检测转录物(由颜色指示)。在供体2中,最高的表达细胞(正方形)不产生IFNG但表达细胞因子IL6和IL1B。紧挨着每个圆圈的数字指示以递减顺序针对该细胞检测到的总唯一性转录物分子数目的排名。

[0088] 图51.类似于图50A,除了此处的数据表示供体2CMV刺激样品的数据之外,即,

CytoSeq数据的聚簇定义了CD8+细胞的两个主要组-一个组表达中心记忆/初始细胞具有的基因，并且另一个组表达效应记忆细胞/效应细胞具有的基因。此处示出了供体2的未刺激的样品的数据。顶部：示出了每对细胞之间的相关性的热图。底部：示出了每个细胞中的每个基因的表达水平的热图。经由双向分层聚簇使细胞和基因排序。

[0089] 图52.在主成分空间中绘出数据。颜色指示特定基因的log(检测到的唯一性转录物分子的数目)。A.用CMV肽池刺激后，显得由更大比例的细胞表达的基因。B.富集在一个分支的细胞中的基因。还已知这些基因与初始和中心记忆CD8+T细胞相关。C.富集在另一分支的细胞中的基因。已知这些基因与效应和效应记忆CD8+T细胞相关。D.表达颗粒酶K的细胞在PC空间上占据初始/中心记忆细胞与效应/效应记忆细胞之间的区域。E.表达HLA-DRA的细胞构成特殊亚群。F.表达于两个分支的细胞中的基因。

[0090] 图53.与图50B相同，除了数据表示未刺激的对照的数据之外。供体1的样品中的细胞都不表达IFNG，而供体2的样品中的一个细胞表达IFNG，但是跨整个基因面板的总体表达较低(排名1069)。将颜色标度调节到与刺激的样品的对应图的颜色标度相匹配。

[0091] 图54.示出了在供体1和2的CMV刺激的CD8+T细胞中表达γ干扰素(IFNG)的细胞中的基因面板的异质表达的热图。还示出了携带在供体2中检测到的最多总转录物的细胞。这一特定细胞不表达IFNG但强烈表达IL6、IL1B和CCL4。基于相关性，通过双向分层聚簇为这些细胞和基因排序。细胞ID是指在基因面板的检测到的转录物的总数中的排名，并且被指示在图50的PCA绘图中。

[0092] 图55.扩增方案。第一PCR使用基因特异性引物和针对通用亿明达(Illumina)测序引物1序列的引物扩增附接至珠的分子。第二PCR使用侧翼于亿明达测序引物2序列的巢式基因特异性引物和针对通用亿明达测序引物1序列的引物扩增第一PCR产物。第三PCR添加P5和P7以及样品索引，以便使PCR产物进入亿明达测序文库。150bp x 2测序揭示了读数1上的细胞标记和分子标记、读数2上的基因以及索引1读数上的样品索引。

[0093] 图56描绘了用于分析来自样品的分子的工作流程的示意图。

[0094] 图57描绘了用于分析来自样品的分子的工作流程的示意图。

[0095] 图58A-B描绘了PCR产物的琼脂糖凝胶。

[0096] 图59描绘了多个基因的测序读数的绘图。

[0097] 图60A-D描绘了分别针对Lys、Phe、Thr和Dap掺入型对照而言，观察到的读数/检测的标记(RPLD)的绘图。图60E描绘了读数比对输入的绘图。

[0098] 图61描绘了对于不同基因而言，观察到的读数/检测的标记(RPLD)的绘图。

[0099] 图62描绘了对于不同基因而言，观察到的读数/检测的标记(RPLD)的绘图。

[0100] 图63描绘了针对不同基因而言的总读数(标记)比对rpld的绘图。

[0101] 图64描绘了针对未检测的基因而言的RPKM的绘图。

[0102] 图65描绘了用于合成分子条码的示意图。

[0103] 图66A-C描绘了用于合成分子条码的示意图。

[0104] 图67示出了用于随机地标记核酸的工作流程的示意图。

[0105] 图68是用于随机地标记核酸的工作流程的示意图。

[0106] 图69说明了机械固定器，微孔阵列基底可以被夹持在其中，由此形成反应室或孔，可以向其中用移液管移取样品和试剂，用于进行多重、单细胞随机标记/分子索引实验。上：

示出了固定器的上下部分和用于与微孔阵列基底形成防漏密封的弹性垫圈的分解图。下：固定器的分解侧视图。

[0107] 图70说明了机械固定器，当微孔阵列基底被夹持在该固定器内时，其产生两个反应室或孔。

[0108] 图71说明了用于与图69和70中说明的机械固定器一起使用的弹性(例如，聚二甲基硅氧烷)垫圈的两个实例。这些弹性垫圈与微孔阵列基底一起提供防漏密封，以在微孔阵列周围产生试剂孔。这些垫圈可以含有一个(上)、两个(下)或更多个开口，用于产生试剂孔。

[0109] 图72描绘了微孔阵列被包装在其内的柱体的一个实施例。左：柱体的分解图，说明了(自下而上)微孔阵列基底、限定流动池或阵列室的垫圈、试剂和/或废物储器部件(用于限定隔区含有预加载测定试剂或储存用过的试剂)以及盖(用于密封试剂和废物储器并限定入口和出口)。右：柱体设计的一个实施例的装配视图，说明了用于使外磁体与微孔阵列非常接近的浮雕。

[0110] 图73描绘了被设计成包括板上的测定试剂与包装的微孔阵列的柱体的一个实施例。

[0111] 图74提供了用于进行多重、单细胞随机标记/分子索引测定的仪器系统的示意性说明。该仪器系统可以提供多种控制和分析能力，并且可以被包装为单独的模块或包装为完全集成的系统。微孔阵列可以与作为系统的固定部件或可移除的流动池整合，或可以被包装在可移除的柱体内，该柱体进一步包括预加载的测定试剂储器和其他功能性。

[0112] 图75说明了有待通过用于进行多重、单细胞随机标记/分子索引测定的自动化系统进行的工艺步骤的一个实施例。

[0113] 图76说明了用于为当前披露的测定系统提供仪器控制和数据分析能力的计算机系统或处理器的一个实施例。

[0114] 图77示出了一个框图，说明了可以与本披露的测定系统的示例性实施例结合使用的计算机系统架构的一个实例。

[0115] 图78描绘了一个图表，示出了具有多个计算机系统、手机、个人数据助理和附网存储(NAS)的网络，其可以与本披露的测定系统的示例性实施例一起使用。

[0116] 图79描绘了多处理器计算机系统的框图，该系统可以与本披露的测定系统的示例性实施例一起使用。

[0117] 图80描绘了测试样品的分析和经由通信媒体从该测试样品获得的测试结果的通信的图表。

[0118] 详细描述

[0119] 在此披露了用于分析多个样品中的分子的方法、试剂盒和组合物。通常，这些方法、试剂盒和组合物包括(a)用分子条码随机地标记两个或更多个样品中的分子，以产生经标记的分子；并且(b)检测这些经标记的分子。这些分子条码可以包括一个或多个靶特异性区、标记区、样品索引区、通用PCR区、适配子、接头或其组合。这些经标记的分子可以包括a)分子区；b)样品索引区；以及c)标记区。分子区可以包括分子条码最初所附接的分子的至少一部分。分子区可以包括分子条码最初所附接的分子的片段。样品索引区可以用于确定分子区的来源。样品索引区可以用于确定分子区来源于哪个样品。样品索引区可以用于区分

来自两个或更多个不同样品的分子区。标记区可以用于为来源于同一来源的相同分子区赋予唯一性身份。标记区可以用于为来源于同一样品的相同分子区赋予唯一性身份。

[0120] 用于分析多个样品中的分子的方法可以包括:a)通过以下产生多个样品标签化核酸:i)使包括多个核酸的第一样品与多个第一样品标签接触,以产生多个第一样品标签化核酸;并且ii)使包括多个核酸的第二样品与多个第二样品标签接触,以产生多个第二样品标签化核酸,其中该多个第二样品标签不同于这些第一样品标签;b)使该多个样品标签化核酸与多个分子鉴定物标记接触,以产生多个经标记的核酸;并且c)检测这些经标记的核酸的至少一部分,由此确定在多个样品中的多个核酸的计数。该多个样品可以包括单细胞。

[0121] 可替代地,用于分析多个样品中的分子的方法可以包括:a)产生多个经标记的核酸,包括i)使第一样品与第一多个样品标签接触,其中该第一多个样品标签包括相同的核酸序列;ii)使该第一样品与第一多个分子鉴定物标记接触,这些分子鉴定物标记可以包括不同的核酸序列,其中使该第一样品与该第一多个样品标签或该第一多个分子鉴定物标记接触同时或顺序发生,以产生多个第一经标记的核酸;iii)使第二样品与第二多个样品标签接触,其中该第二多个样品标签可以包括相同的核酸序列;iv)使该第二样品与第二多个分子鉴定物标记接触,这些分子鉴定物标记可以包括不同的核酸序列,其中使该第二样品与该第二多个样品标签或该第二多个分子鉴定物标记接触同时或顺序发生,以产生多个第二经标记的核酸,其中该多个经标记的核酸可以包括该多个第一经标记的核酸和这些第二经标记的核酸;并且b)确定多个不同的经标记的核酸,由此确定在多个样品中的多个核酸的计数。

[0122] 用于分析多个样品中的分子的方法可以包括:a)使多个样品与多个样品标签和多个分子鉴定物标记接触,这些样品可以包括两个或更多个不同的核酸,以产生多个经标记的核酸,其中:i)该多个经标记的核酸可以包括附接至两个或更多个样品标签和两个或更多个分子鉴定物标记的两个或更多个核酸;ii)附接至来自该多个样品的第一样品的核酸的样品标签不同于附接至来自该多个样品的第二样品的核酸分子的样品标签;并且iii)在同一样品中的两个或更多个相同的核酸附接至两个或更多个不同的分子鉴定物标记;并且b)检测这些经标记的核酸的至少一部分,由此确定该多个样品中的两个或更多个不同核酸的计数。

[0123] 图56描绘了用于定量样品中的RNA分子的示例性工作流程。如图56的步骤1所示,通过使一组分子鉴定物标记(115)与RNA分子的聚A尾区域随机杂交,这些RNA分子(110)可以被逆转录,以产生cDNA分子(105)。这些分子鉴定物标记(115)可以包括寡聚dT区域(120)、标记区(125)以及通用PCR区(130)。该组分子鉴定物标记可以包含960个不同类型的标记区。如图56的步骤2所示,可以将这些经标记的cDNA分子(170)纯化,以除去过量的分子鉴定物标记(115)。纯化可以包括Ampure珠纯化。如图56的步骤3所示,可以扩增这些经标记的cDNA分子(170),以产生经标记的扩增子(180)。扩增可以包括多重PCR扩增。扩增可以包括在单一反应体积中用96种多重引物进行的多重PCR扩增。扩增可以包括定制引物(135)和通用引物(140)。定制引物(135)可以与经标记的cDNA分子(170)的cDNA(105)部分内的区域杂交。通用引物(140)可以与经标记的cDNA分子(170)的通用PCR区(130)杂交。如步骤4所示,可以通过巢式PCR进一步扩增这些经标记的扩增子(180)。巢式PCR可以包括在单一反应体积中用96种多重引物进行的多重PCR。巢式PCR可以包括定制引物(145)和通用引物

(140)。定制引物(135)可以与经标记的扩增子(180)的cDNA(105)部分内的区域杂交。通用引物(140)可以与经标记的扩增子(180)的通用PCR区(130)杂交。如步骤5所示,一个或多个适配子(150,155)可以附接至经标记的扩增子(180),以产生经适配子标记的扩增子(190)。该一个或多个适配子可以经由连接附接至经标记的扩增子(180)。如步骤6所示,该一个或多个适配子(150,155)可以用于在经适配子标记的扩增子(190)上进行一个或多个另外的测定。该一个或多个适配子(150,155)可以与一个或多个引物(160,165)杂交。该一个或多个引物(160,165)可以是PCR扩增引物。该一个或多个引物(160,165)可以是测序引物。该一个或多个适配子(150,155)可以用于进一步扩增经适配子标记的扩增子。该一个或多个适配子(150,155)可以用于对经适配子标记的扩增子测序。

[0124] 图57描绘了用于分析来自两个或更多个样品的核酸的工作流程的示例性示意图。如图57所示,用于分析来自两个或更多个样品的核酸的方法可以包括选择两个或更多个基因供分析并且基于所选基因设计定制引物(210)。该方法可以进一步包括补充一个或多个包含核酸(例如,RNA)的样品与一个或多个掺入型对照(220)。可以用分子条码(或样品标签或分子鉴定物标记)和定制引物通过多重RT-PCR扩增该样品中的核酸(230),以产生经标记的扩增子。可以进一步用一个或多个测序适配子处理经标记的扩增子,以产生经适配子标记的扩增子(240)。可以分析经适配子标记的扩增子(250)。如图57所示,分析经标记的扩增子(250)可以包括以下项中的一种或多种:(1)检测通用PCR引物序列、聚A和/或分子条码(或样品标签、分子鉴定物标记);(2)对未附接至适配子和/或条码(例如,分子条码、样品标签、分子鉴定物标记)的经适配子标记的扩增子(例如,96个基因和掺入型对照)的末端进行图谱读数;以及(3)计数和/或概述不同的经适配子标记的扩增子的数目。

[0125] 图67示出了用于用分子条码(1220)随机地标记核酸的工作流程的示意图。如图67的步骤1所示,可以用一组分子条码(1220)随机地标记RNA分子。分子条码(1220)可以包括靶结合区(1221)、标记区(1222)、样品索引区(1223)以及通用PCR区(1224)。在一些情况下,靶结合区包括与RNA分子中的聚A序列杂交的寡聚dT序列。标记区(1222)可以包括可以用于区分两个或更多个不同分子条码的唯一性序列。当分子条码与RNA分子杂交时,标记区可以用于为相同的RNA分子赋予唯一性身份。对于一组分子条码而言,样品索引区(1223)可以是相同的。样品索引区(1223)可以用于区分来自不同样品的经标记的核酸。通用PCR区(1224)可以用作用于扩增经标记的分子的引物结合位点。一旦将RNA分子用分子条码标记,这些RNA分子便可以被逆转录,以产生含有RNA分子的cDNA拷贝(1210)和分子条码(1220)的经标记的cDNA分子(1230)。

[0126] 如图67的步骤2所示,可以通过Ampure珠纯化除去过量的寡核苷酸(例如,分子条码)。如图67的步骤3所示,可以通过多重PCR扩增经标记的cDNA分子。可以通过在单一反应体积中使用第一组正向引物(图67中的F1,1235)和通用引物(1240)进行经标记的cDNA分子的多重PCR,以产生经标记的扩增子(1245)。如图67的步骤4所示,可以使用巢式引物通过多重PCR进行扩增经标记的扩增子。可以通过在单一反应体积中使用第二组正向引物(图67中的F2,1250)和通用引物(1240)进行经标记的扩增子的巢式引物扩增,以产生经标记的巢式PCR扩增子。在一些情况下,F2引物(1250)含有适配子(1251)和靶结合区(1252)。F2引物的靶结合区(1252)可以与经标记的扩增子杂交并且可以引发经标记的扩增子的扩增。巢式PCR扩增子的适配子(1251)和通用PCR区(1224)可以用在对经标记的巢式PCR扩增子的测序

中。可以通过MiSeq对扩增子测序。可替代地，可以通过HiSeq对扩增子测序。

[0127] 图68示出了用于随机地标记核酸的工作流程的示意图。如图68的步骤1所示，可以用一组分子条码(1320)随机地标记RNA分子(1305)。分子条码可以包括靶结合区(1321)、标记区(1322)和通用PCR区(1323)。一旦分子条码附接至RNA分子，这些RNA分子(1305)便可以被逆转录，以产生包含RNA分子的cDNA拷贝(1310)和分子条码(1320)的经标记的cDNA分子(1325)。如图68的步骤2所示，可以通过Ampure珠纯化来纯化经标记的cDNA分子，以除去过量的寡核苷酸(例如，分子条码)。如图68的步骤3所示，可以通过多重PCR扩增经标记的扩增子。可以通过在单一反应体积中使用第一组正向引物(图68中的F1，1330)和通用引物(1335)进行经标记的cDNA分子的多重PCR，以产生经标记的扩增子(1360)。如图67的步骤4所示，可以使用巢式引物通过多重PCR进行扩增经标记的扩增子。可以通过在单一反应体积中使用第二组正向引物(图68中的F2，1340)和样品索引引物(1350)进行经标记的扩增子的巢式引物扩增，以产生经标记的巢式PCR扩增子。在一些情况下，F2引物(1340)含有适配子(1341)和靶结合区(1342)。F2引物的靶结合区(1342)可以与经标记的扩增子杂交并且可以引发经标记的扩增子的扩增。样品索引引物(1350)可以包括通用引物区(1351)、样品索引区(1352)和适配子区(1353)。如图68的步骤4所示，样品索引引物的通用引物区(1351)可以与经标记的扩增子的通用PCR区杂交。样品索引引物的样品索引区(1352)可以用于区分两个或更多个样品。适配子区(1341，1353)可以用于对经标记的巢式PCR扩增子测序。可以通过MiSeq对扩增子测序。可替代地，可以通过HiSeq对扩增子测序。

[0128] 在此进一步披露了产生一个或多个文库的方法。该一个或多个文库可以包括多个经标记的分子。该一个或多个文库可以包括多个经标记的扩增子。该一个或多个文库可以包括多个富集的分子或其衍生物(例如，经标记的分子、经标记的扩增子)。通常，产生一个或多个文库的方法包括(a)随机地标记来自两个或更多个样品的多个分子，以产生多个经标记的分子，其中这些经标记的分子包括分子区、样品索引区和标记区；并且(b)从该多个经标记的分子产生一个或多个文库，其中(i)该一个或多个文库包括两个或更多个不同的经标记的分子，(ii)该两个或更多个不同的经标记的分子的区别在于分子区、样品索引区、标记区或其组合。

[0129] 用于产生一个或多个文库的方法可以包括：a)通过以下产生多个样品标签化核酸：i)使包括多个核酸的第一样品与多个第一样品标签接触，以产生多个第一样品标签化核酸；并且ii)使包括多个核酸的第二样品与多个第二样品标签接触，以产生多个第二样品标签化核酸，其中该多个第一样品标签不同于这些第二样品标签；并且b)使该多个样品标签化核酸与多个分子鉴定物标记接触，以产生多个经标记的核酸，由此产生经标记的核酸文库。

[0130] 与样品接触可以是随机或非随机的。例如，样品与样品标签的接触可以是随机或非随机接触。在一些实施例中，使样品与样品标签随机接触。在一些实施例中，使样品与样品标签非随机接触。与多个核酸接触可以是随机或非随机的。例如，多个核酸与样品标签的接触可以是随机或非随机接触。在一些实施例中，使该多个核酸与样品标签随机接触。在一些实施例中，使该多个核酸与样品标签非随机接触。

[0131] 在此进一步披露了产生一组或多组经标记的珠的方法。产生一组或多组经标记的珠的方法可以包括将一个或多个核酸附接至一个或多个珠，由此产生一组或多组的经标记

的珠。该一个或多个核酸可以包括一个或多个分子条码。该一个或多个核酸可以包括一个或多个样品标签。该一个或多个核酸可以包括一个或多个分子鉴定物标记。该一个或多个核酸可以包括a)引物区;b)样品索引区;以及c)接头或适配子区。该一个或多个核酸可以包括a)引物区;b)标记区;以及c)接头或适配子区。该一个或多个核酸可以包括a)样品索引区;以及b)标记区。该一个或多个核酸可以进一步包括引物区。该一个或多个核酸可以进一步包括靶特异性区。该一个或多个核酸可以进一步包括接头区域。该一个或多个核酸可以进一步包括适配子区。该一个或多个核酸可以进一步包括样品索引区。该一个或多个核酸可以进一步包括标记区。

[0132] 在此进一步披露了用于选择一个或多个定制引物的方法。选择用于分析多个样品中的分子的定制引物的方法可以包括:a)第一遍,其中选择的引物可以包括:i)不多于三个连续鸟嘌呤、不多于三个连续胞嘧啶、不多于四个连续腺嘌呤以及不多于四个连续胸腺嘧啶;ii)作为鸟嘌呤或胞嘧啶的至少3、4、5或6个核苷酸;以及iii)并不易于形成发夹结构的序列;b)第二遍,包括:i)第一轮选择具有所有转录物的高覆盖的多个序列;以及ii)随后的一轮或多轮,选择以下序列,该序列具有剩余转录物的最高覆盖和不多于4的与其他选择的序列的互补分数;并且c)添加序列至挑选的组,直至覆盖饱和或定制引物的总数小于或等于约96。

[0133] 在此进一步披露了用于分析来自两个或更多个样品的两个或更多个分子的试剂盒。该试剂盒可以包括(a)第一容器,该容器包括第一组分子条码,其中(i)第一组分子条码的分子条码包括样品索引区和标记区;(ii)第一组分子条码的两个或更多个条码的样品索引区是相同的;并且(iii)第一组分子条码的两个或更多个条码的标记区是不同的;以及(b)第二容器,该容器包括第二组分子条码,其中(i)第二组分子条码的分子条码包括样品索引区和标记区;(ii)第二组分子条码的两个或更多个条码的样品索引区是相同的;(iii)第二组分子条码的两个或更多个条码的标记区是不同的;(iv)第二组分子条码的条码的样品索引区不同于第一组分子条码的条码的样品索引区;并且(v)第二组分子条码的两个或更多个条码的标记区与第一组分子条码的两个或更多个条码的标记区是相同的。

[0134] 可替代地,该试剂盒包括:a)多个珠,其中该多个珠的一个或多个珠可以包括多个核酸中的至少一个,其中多个核酸中的至少一个可以包括:i)至少一个引物序列,其中该多个核酸中的至少一个的引物序列对于该多个珠而言是相同的;ii)珠特异性序列,其中该多个核酸中的任一个核酸的珠特异性序列是相同的,并且其中珠特异性序列对于该多个珠中的任一个珠而言是不同的;以及iii)随机序列,其中随机序列对于该多个核酸中的任一个核酸而言是不同的;b)引物,该引物可以包括与引物序列互补的序列;以及c)一种或多种适于核酸扩增的扩增剂。

[0135] 可替代地,该试剂盒包括:a)第一容器,该容器包括第一组样品标签,其中(i)第一组样品标签的样品标签包括样品索引区;并且(ii)第一组样品标签的样品标签的样品索引区是至少约80%相同的;以及b)第二容器,该容器包括第一组分子鉴定物标记,其中(i)第一组分子鉴定物标记的分子鉴定物标记包括标记区;并且(ii)第一组分子鉴定物标记的总分子鉴定物标记的至少约30%的标记区是不同的

[0136] 在更详细地描述本发明方法、试剂盒和组合物之前,应理解的是本发明不限于描述的具体方法、试剂盒或组合物,因为这些当然可以改变。还应理解的是,在此使用的术语

仅是出于描述具体实施例的目的，并不旨在是限制性的，因为本发明的范围将仅由所附权利要求书限制。提出实例以便为本领域普通技术人员提供如何制造和使用本发明的完整披露和描述，并不旨在限制诸位发明人考虑作为他们的发明的范围，它们也不旨在表示以下实验是进行的所有或仅有的实验。就使用的数字(例如，量、温度等)而言，已努力确保其精确度，但仍应考虑到有一些实验误差和偏差。除非另外指明，份数是重量份，分子量是重量平均分子量，温度是摄氏度，并且压力是或接近大气压。

[0137] 提供了用于随机标记多个样品或复杂的核酸制剂中的核酸的方法、试剂盒和组合物。这些方法、试剂盒和组合物可用于阐明细胞应答、分化或信号转导的机制以及进行多种多样的临床测量。通过阅读如将在下文中更全面地描述的本发明的方法、试剂盒和组合物，本发明的这些和其他目的、优点和特征对于本领域普通技术人员而言将变得显而易见。

[0138] 在此披露的方法包括将一个或多个分子条码、样品标签和/或分子鉴定物标记附接至来自两个或更多个样品的两个或更多个分子。这些分子条码、样品标签和/或分子鉴定物标记可以包括一个或多个寡核苷酸。在一些情况下，将分子条码、样品标签和/或分子鉴定物标记附接至分子包括随机地标记这些分子。用于随机地标记分子的方法可见于例如美国序列号12/969,581和13/327,526。通常，随机标记方法包括将多个标签和标记寡核苷酸随机附接至一个或多个分子。相对于有待标记的一个或多个分子，过量地提供分子条码、样品标签和/或分子鉴定物标记。在随机标记中，每个有待标记的单个分子具有附接至该多个分子条码、样品标签和/或分子鉴定物标记的单独概率。每个有待标记的单个分子附接至特定分子条码、样品标签和/或分子鉴定物标记的概率可以与任何其他有待标记的单个分子大致相同。因此，在一些情况下，假设样品中的任何分子遇到任何标签和标记的概率是相等的，这一假设可用在数学计算中，以估算该样品中的分子数目。在一些情形下，可以例如通过选择具有将提高或降低那些分子条码、样品标签和/或分子鉴定物标记与单个分子的结合效率的不同特性的标签和标记来控制附接概率。这些标签和标记在数目上也可变化，以改变在随机标记过程中特定分子条码、样品标签和/或分子鉴定物标记遇到结合配偶体的概率。例如，一种标记在标记池中大量存在，由此提高大量存在的标记遇到至少一个结合配偶体的几率。

[0139] 在此披露的方法可以进一步包括组合两个或更多个样品。在此披露的方法可以进一步包括组合来自两个或更多个样品的一个或多个分子。例如，在此披露的方法包括组合第一样品和第二样品。可以在进行一个或多个随机标记程序之后组合该两个或更多个样品。可以在将一组或多组分子条码附接至来自该两个或更多样品的两个或更多个分子之后组合该两个或更多个样品。可以在将一组或多组样品标签附接至来自该两个或更多样品的两个或更多个分子之后组合该两个或更多个样品。可以在将一组或多组分子鉴定物标记附接至来自该两个或更多样品的两个或更多个分子之后组合该两个或更多个样品。例如，在与该多个分子鉴定物标记接触之前组合第一和第二样品。

[0140] 可替代地，可以在进行一个或多个随机标记程序之前组合该两个或更多个样品。可以在将一组或多组分子条码附接至来自该两个或更多样品的两个或更多个分子之前组合该两个或更多个样品。可以在将一组或多组样品标签附接至来自该两个或更多样品的两个或更多个分子之前组合该两个或更多个样品。可以在将一组或多组分子鉴定物标记附接至来自该两个或更多样品的两个或更多个分子之前组合该两个或更多个样品。

[0141] 可以在对来自该两个或更多个样品的两个或更多个分子或其衍生物(例如,经标记的分子、扩增子)进行一个或多个测定之后组合该两个或更多个样品。该一个或多个测定可以包括一个或多个扩增反应。该一个或多个测定可以包括一个或多个富集测定。该一个或多个测定可以包括一个或多个检测测定。例如,在检测经标记的核酸之后组合第一和第二样品。

[0142] 可以在对来自该两个或更多个样品的两个或更多个分子或其衍生物(例如,经标记的分子、扩增子)进行一个或多个测定之前组合该两个或更多个样品。该一个或多个测定可以包括一个或多个扩增反应。该一个或多个测定可以包括一个或多个富集测定。该一个或多个测定可以包括一个或多个检测测定。例如,在检测经标记的核酸之前组合第一和第二样品。

[0143] 支持物

[0144] 本披露包括用于对单细胞进行多重序列分析的组合物和方法。本披露的方法和组合物提供了固体支持物的用途。在一些情况下,在此披露的方法、试剂盒和组合物包括支持物。

[0145] 术语“支持物”、“固体支持物”“半固体支持物”以及“基底”可互换使用,并且是指具有一个或多个刚性或半刚性表面的材料或一组材料。支持物可以指任何表面,其可从溶液转移到溶液中或形成用于进行基于寡核苷酸的测定的结构。支持物或基底可以是固体支持物。可替代地,支持物是非固体支持物。支持物可以指不溶性、半溶性或不溶性材料。当支持物包括接头、支架、结构单元或附接至其上的其他反应性部分时,它可以被称为“官能化的”,而当固体支持物缺少附接至其上的这样一个反应性部分时,它可以被称为“非官能化的”。支持物可以在溶液中不受约束地利用,如以微量滴定孔形式;以流通形式,如在柱中;或在试纸条(dipstick)中。

[0146] 支持物或基底可以包括膜、纸、塑料、涂覆的表面、平表面、玻璃、载玻片、芯片或其任何组合。在许多实施例中,支持物的至少一个表面大体可以是平的,但是在一些实施例中,对于不同的化合物,可能希望的是用例如孔、凸起区域、针、蚀刻沟槽等在物理上分开合成区域。根据其他实施例,一个或多个固体支持物可以采用树脂、凝胶、微球或其他几何构型的形式。可替代地,一个或多个固体支持物包括二氧化硅芯片、微粒、纳米粒子、平板以及阵列。固体支持物可以包括珠(例如,硅胶、可控孔度玻璃、磁珠、Dynabead、王(Wang)树脂;梅里菲尔德(Merrifield)树脂、交联葡聚糖/琼脂糖珠、纤维素珠、聚苯乙烯珠等),毛细管,平支持物如玻璃纤维过滤器、玻璃表面、金属表面(钢、金、银、铝、硅以及铜)、玻璃支持物、塑料支持物、硅支持物、芯片、过滤器、膜、微孔板、载玻片或类似物,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),晶片,梳状物(comb),针或针头(例如,适于组合合成或分析的针阵列)或平表面(如晶片(例如,硅晶片)、带有具有或不具有滤底的凹陷的晶片)的凹陷或纳升孔的阵列中的珠。

[0147] 适用于聚合物(包括蛋白质)阵列合成的方法和技术已经描述于美国专利公开号20050074787,WO 00/58516,美国专利号5,143,854、5,242,974、5,252,743、5,324,633、5,384,261、5,405,783、5,424,186、5,451,683、5,482,867、5,491,074、5,527,681、5,550,215、5,571,639、5,578,832、5,593,839、5,599,695、5,624,711、5,631,734、5,795,716、5,831,070、5,837,832、5,856,101、5,858,659、5,936,324、5,968,740、5,974,164、5,981,

185、5,981,956、6,025,601、6,033,860、6,040,193、6,090,555、6,136,269、6,269,846及6,428,752,PCT公开号WO 99/36760和WO 01/58593中,出于所有目的将其全部通过引用以其全部内容结合在此。描述了具体实施例中的合成技术的专利包括美国专利号5,412,087、6,147,205、6,262,216、6,310,189、5,889,165以及5,959,098。核酸阵列描述于许多上述专利中,但是许多相同技术可以应用于多肽阵列。另外的示例性基底披露于美国专利号5,744,305以及美国专利公开号20090149340和20080038559中。

[0148] 将经标记的核酸附接至支持物可以包括胺-硫醇交联、马来酰亚胺交联、N-羟基琥珀酰亚胺或N-羟基硫代琥珀酰亚胺、泽农(Zenon)或位点点击(SiteClick)。将经标记的核酸附接至支持物可以包括将生物素附接至该多个经标记的核酸并用链霉亲和素涂覆该一个或多个珠。

[0149] 在一些情况下,固体支持物可以包括分子支架。示例性分子支架可以包括抗体、抗原、亲和试剂、多肽、核酸、细胞器以及类似物。分子支架可以被连接在一起(例如,固体支持物可以包括多个连接的分子支架)。分子支架可以通过氨基酸接头、核酸接头、小分子连接(例如,生物素和亲和素)和/或基质连接(例如,PEG或甘油)连接在一起。连接可以是非共价的。连接可以是共价的。在一些情况下,分子支架可以不被连接。多个单独的分子支架可以用于本披露的方法中。

[0150] 在一些情况下,支架可以包括纳米粒子。纳米粒子可以是镍、金、银、碳、铜、硅酸盐、铂钴、氧化锌、二氧化硅结晶和/或银纳米粒子。可替代地或另外地,纳米粒子可以是包埋在多孔锰氧化物中的金纳米粒子。纳米粒子可以是铁纳米粒子。纳米粒子可以是镶有碳纳米粒子的纳米四脚体(nanotetrapod)。

[0151] 支持物可以包括聚合物。支持物可以包括基质。基质可以进一步包括一个或多个珠。聚合物可以包括PEG、甘油、多糖或其组合。聚合物可以是塑料、橡胶、尼龙、硅酮、氯丁橡胶和/或聚苯乙烯。聚合物可以是天然聚合物。天然聚合物的实例包括但不限于虫胶、琥珀、羊毛、蚕丝、纤维素以及天然橡胶。聚合物可以是合成聚合物。合成聚合物的实例包括但不限于合成橡胶、酚醛树脂(或Bakelite)、氯丁橡胶、尼龙、聚氯乙烯(PVC或乙烯树脂(vinyl))、聚苯乙烯、聚乙烯、聚丙烯、聚丙烯腈、PVB以及硅酮。

[0152] 支持物可以是半固体支持物。支持物可以包括凝胶(例如,水凝胶)。术语“水凝胶”、“凝胶”等在此可互换使用,并且是指这样一种材料,该材料不是易于流动的液体也不是固体而是凝胶,该凝胶由按重量计从0.5%或更多并且优选少于40%的成凝胶溶质材料和从95%或更少并且优选超过55%的水组成。可以通过使用优选作为合成溶质(但可以是天然溶质,例如用于形成明胶)的溶质形成本发明的凝胶,所述溶质形成互连单元,所述互连单元结合至水、陷捕水、吸收水和/或以其他方式保水,并且由此与水结合产生凝胶,其中水包括结合水和非结合水。该凝胶可以是本发明的水凝胶贴片的基本结构,所述水凝胶贴片除成凝胶溶质材料和水之外还包括另外的组分(如酶和盐),将在此进一步描述这些组分。凝胶可以是聚合物凝胶。

[0153] 固体支持物可以包括结构化纳米结构。例如,结构化纳米结构可以包括捕获容器(例如,微型蜂巢),其可以包括用于捕获细胞和/或细胞内容物的寡核苷酸。在一些情况下,结构化纳米结构可以不需要添加外源性试剂。

[0154] 在一些情况下,支持物包括珠。珠可以包括任何类型的实心或空心球体、球、承座、

圆柱体或其他类似配置，其由塑料、陶瓷、金属或聚合材料构成，其上可以固定核酸(例如，共价地或非共价地)。珠可以包括一根或多根尼龙线。珠的形状可以是球形的。珠的形状可以是非球形的。珠可以是未抛光的，或如果被抛光，则抛光珠可以在处理(例如，用烷化剂)之前被粗糙化。珠可以包括可以是球形的(例如，微球)或具有不规则形状的离散颗粒。珠可以包括多种材料，包括但不限于顺磁性材料、陶瓷、塑料、玻璃、聚苯乙烯、甲基苯乙烯、丙烯酸聚合物、钛、胶乳、琼脂糖、纤维素、尼龙以及类似物。珠可以附接至或被包埋在一个或多个支持物中。珠可以附接至凝胶或水凝胶。珠可以被包埋在凝胶或水凝胶中。珠可以附接至基质。珠可以被包埋在基质中。珠可以附接至聚合物。珠可以被包埋在聚合物中。可以使用珠上存在的用作位置地址的寡核苷酸鉴定该珠在支持物(例如，凝胶、基质、支架或聚合物)内的空间位置。珠的实例包括但不限于链霉亲和素珠、琼脂糖珠、磁珠、Dynabeads®、MACS®微珠、缀合抗体的珠(例如，抗免疫球蛋白微珠)、缀合A蛋白的珠、缀合G蛋白的珠、缀合A/G蛋白的珠、缀合L蛋白的珠、缀合寡聚dT的珠、二氧化硅珠、二氧化硅样珠、抗生物素微珠、抗荧光染料微珠以及BcMag™羧基封端的磁珠。珠的直径可以是约5μm、10μm、20μm、25μm、30μm、35μm、40μm、45μm或50μm。珠可以指可以为生物颗粒和大分子(如DNA和RNA)的固定提供增加的表面积的任何三维结构。

[0155] 支持物可以是多孔的。支持物可以是可渗透性的或半渗透性的。支持物可以是固体的。支持物可以是半固体的。支持物可以是具延展性的。支持物可以是具弹性的。在一些情况下，支持物可以被模制成一定形状。例如，可以将支持物放置在物体上并且该支持物可以呈该物体的形状。在一些情况下，将支持物放置在器官上并且呈该器官的形状。在一些情况下，通过3D打印产生支持物。

[0156] 支持物(例如，珠、纳米粒子)的直径可以是至少约0.1、0.5、1、2、3、4、5、6、7、8、9、10、20、30、100、500、1000或2000或更大微米。固体支持物(例如，珠)的直径可以是至多约0.1、0.5、1、2、3、4、5、6、7、8、9、10、20、30、100、500、1000或2000或更大微米。珠的直径可以是约20微米。

[0157] 在一些情况下，固体支持物包括树状聚合物。树状聚合物可以比珠小。树状聚合物可以是亚细胞的。树状聚合物的直径可以小于1、0.9、0.8、0.7、0.6、0.5、0.4、0.3、0.2或0.1微米。树状聚合物的直径可以小于0.09、0.08、0.07、0.06、0.05、0.04、0.03或0.01微米。树状聚合物可以包括三个主要部分：核心、内壳和外壳。树状聚合物可以被合成为在这些部分的每者中具有不同的功能性。树状聚合物的这些部分的不同功能性可以控制化合物的特性，如溶解度、热稳定性和附接，用于特定应用。树状聚合物可以合成地处理。树状聚合物可以通过发散性合成来合成。发散性合成可以包括从多功能核心组装树状聚合物，将该核心通过一系列反应向外延伸。发散合成可以包括一系列迈克尔(Michael)反应。可替代地，树状聚合物可以通过汇集性合成来合成。汇集性合成可以包括从在球体表面结束的小分子构建树状聚合物，并且反应可以向内进行并且最终附接至核心。树状聚合物还可以通过点击化学制备。点击化学可以包括狄尔斯-阿尔德(Diels-Alder)反应、硫醇-炔反应、叠氮化物-炔烃反应或其组合。树状聚合物的实例包括但不限于聚(酰胺基胺)(PAMAM)树状聚合物、PEG-核心树状聚合物、含磷树状聚合物、聚丙烯亚胺树状聚合物以及聚赖氨酸树状聚合物。树状聚合物可以是手性树状聚合物。可替代地，树状聚合物可以是非手性树状聚合物。

[0158] 固体支持物可以包括树状聚合物的一部分。树状聚合物的该部分可以包括树状突

(dendron)。树状突在聚焦点处可以包括具有多个末端基团和单一反应功能的单分散性楔形树状聚合物区段。固体支持物可以包括聚酯树状突。树状突的实例包括但不限于聚酯-8-羟基-1-乙炔双-MPA树状突、聚酯-16-羟基-1-乙炔双-MPA树状突、聚酯-32-羟基-1-乙炔双-MPA树状突、聚酯-8-羟基-1-羧基双-MPA树状突、聚酯-16-羟基-1-羧基双-MPA树状突以及聚酯-32-羟基-1-羧基双-MPA树状突。

[0159] 固体支持物可以包括超支化聚合物。超支化聚合物可以包括具有树状聚合物样特性的多分散性树状大分子。通常,以单一合成聚合步骤制备超支化聚合物。超支化聚合物可以基于2,2-双(羟基甲基)丙酸(双-MPA)单体。超支化聚合物的实例包括但不限于超支化双-MPA聚酯-16-羟基、超支化双-MPA聚酯-32-羟基和超支化双-MPA聚酯-64-羟基。

[0160] 固体支持物可以是阵列或微阵列。固体支持物可以包括离散区域。固体支持物可以是可寻址的阵列。在一些情况下,该阵列包括固定在固体表面上的多个探针。该多个探针允许经标记的分子和/或经标记的扩增子与固体表面的杂交。该多个探针包括与经标记的分子和/或标记的扩增子的至少一部分互补的序列。在一些情况下,该多个探针包括与样品标签、分子鉴定物标记、核酸或其组合的至少一部分互补的序列。在其他情况下,该多个探针包括与通过样品标签或分子鉴定物标记附接至核酸而形成的连接处互补的序列。

[0161] 阵列可以包括一个或多个探针。这些探针可以处于多种形式。阵列可以包括以下探针,该探针包括与靶核酸的至少一部分互补的序列和与样品标签或分子鉴定物标记的唯一性鉴定物区互补的序列,其中该样品标签或分子鉴定物标记包括寡核苷酸。与靶核酸的至少一部分互补的序列可以附接至该阵列。与唯一性鉴定物区互补的序列可以附接至该阵列。阵列可以包括第一探针和第二探针,该第一探针包括与靶核酸的至少一部分互补的序列,该第二探针与唯一性鉴定物区互补。经随机标记的核酸可以按不同方式与阵列杂交。例如,经随机标记的核酸的唯一性鉴定物区和靶核酸的连接处可以与阵列上的探针杂交。经随机标记的核酸的区域中可以存在空位,其可以与阵列上的探针杂交。经随机标记的核酸的不同区域可以与阵列上的两个或更多个探针杂交。因此,阵列探针可以处于许多不同形式。阵列探针可以包括与唯一性鉴定物区互补的序列、与靶核酸互补的序列或其组合。经随机标记的核酸与阵列的杂交可以通过多种方式进行。例如,经随机标记的核酸的两个或更多个核苷酸可以与阵列上的一个或多个探针杂交。与探针杂交的经随机标记的核酸的两个或更多个核苷酸可以是连续的核苷酸、非连续的核苷酸或其组合。可以通过本领域已知的任何方法检测与探针杂交的经随机标记的核酸。例如,可以直接检测经随机标记的核酸。直接检测经随机标记的核酸可以包括检测荧光团、半抗原或可检测标记。可以间接检测经随机标记的分子。经随机标记的核酸的间接检测可以包括连接法或其他酶法或非酶法。

[0162] 该阵列可以处于多种形式。例如,该阵列可以处于16-、32-、48-、64-、80-、96-、112-、128-、144-、160-、176-、192-、208-、224-、240-、256-、272-、288-、304-、320-、336-、352-、368-、384-或400-形式。可替代地,该阵列处于8x60K、4x180K、2x400K、1x1M形式。在其他情况下,该阵列处于8x15K、4x44K、2x105K、1x244K形式。

[0163] 该阵列可以包括单一阵列。单一阵列可以在单一基底上。可替代地,该阵列在多个基底上。该阵列可以包括多种形式。该阵列可以包括多个阵列。该多个阵列可以包括两个或更多个阵列。例如,该多个阵列可以包括至少约2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、25、30、35、40、45、50、55、60、65、70、75、80、85、90、95或100个阵列。在一些情

况下，该多个阵列中的至少两个阵列是相同的。可替代地，该多个阵列中的至少两个阵列是不同的。

[0164] 在一些情况下，该阵列包括对称的隔室区。例如，该阵列包括0.5x 0.5毫米(mm)、1x 1mm、1.5x 1.5mm、2x 2mm、2.5x 2.5mm、3x 3mm、3.5x 3.5mm、4x 4mm、4.5x 4.5mm、5x 5mm、5.5x 5.5mm、6x 6mm、6.5x 6.5mm、7x 7mm、7.5x 7.5mm、8x 8mm、8.5x8.5mm、9x 9mm、9.5x 9.5mm、10x 10mm、10.5x 10.5mm、11x 11mm、11.5x 11.5mm、12x 12mm、12.5x 12.5mm、13x 13mm、13.5x13.5mm、14x 14mm、14.5x 14.5mm、15x 15mm、15.5x 15.5mm、16x 16mm、16.5x 16.5mm、17x 17mm、17.5x 17.5mm、18x 18mm、18.5x 18.5mm、19x 19mm、19.5x 19.5mm或20x 20mm的隔室区。在一些情况下，该阵列包括6.5x 6.5mm的隔室区。可替代地，该阵列包括不对称的隔室区。例如，该阵列包括6.5x 0.5mm、6.5x 1mm、6.5x 1.5mm、6.5x 2mm、6.5x 2.5mm、6.5x 3mm、6.5x 3.5mm、6.5x 4mm、6.5x 4.5mm、6.5x 5mm、6.5x 5.5mm、6.5x 6mm、6.5x 6.5mm、6.5x 7mm、6.5x 7.5mm、6.5x 8mm、6.5x 8.5mm、6.5x 9mm、6.5x 9.5mm、6.5x 10mm、6.5x 10.5mm、6.5x 11mm、6.5x 11.5mm、6.5x 12mm、6.5x 12.5mm、6.5x 13mm、6.5x 13.5mm、6.5x 14mm、6.5x14.5mm、6.5x 15mm、6.5x 15.5mm、6.5x 16mm、6.5x 16.5mm、6.5x 17mm、6.5x 17.5mm、6.5x 18mm、6.5x 18.5mm、6.5x 19mm、6.5x 19.5mm或6.5x 20mm的隔室区。

[0165] 该阵列可以包括至少约1微米(μm)、2μm、3μm、4μm、5μm、6μm、7μm、8μm、9μm、10μm、15μm、20μm、25μm、30μm、35μm、40μm、45μm、50μm、55μm、60μm、65μm、70μm、75μm、80μm、85μm、90μm、95μm、100μm、125μm、150μm、175μm、200μm、225μm、250μm、275μm、300μm、325μm、350μm、375μm、400μm、425μm、450μm、475μm或500μm的点。在一些情况下，该阵列包括70μm的点。

[0166] 该阵列可以包括至少约1μm、2μm、3μm、4μm、5μm、6μm、7μm、8μm、9μm、10μm、15μm、20μm、25μm、30μm、35μm、40μm、45μm、50μm、55μm、60μm、65μm、70μm、75μm、80μm、85μm、90μm、95μm、100μm、125μm、150μm、175μm、200μm、225μm、250μm、275μm、300μm、325μm、350μm、375μm、400μm、425μm、450μm、475μm、500μm、525μm、550μm、575μm、600μm、625μm、650μm、675μm、700μm、725μm、750μm、775μm、800μm、825μm、850μm、875μm、900μm、925μm、950μm、975μm、1000μm的特征螺距。在一些情况下，该阵列包括161μm的特征螺距。

[0167] 该阵列可以包括一个或多个探针。在一些情况下，该阵列包括至少约5、10、15、20、25、30、40、50、60、70、80、90或100个探针。可替代地，该阵列包括至少约200、300、400、500、600、700、800、900、1000、1100、1200、1300、1400、1500、1600、1700、1800、1900、2000、2100、2200、2300、2400、2500、2600、2700、2800、2900或3000个探针。该阵列可以包括至少约3500、4000、4500、5000、5500、6000、6500、7000、7500、8000、8500、9000、9500或10000个探针。在一些情况下，该阵列包括至少约960个探针。可替代地，该阵列包括至少约2780个探针。这些探针可以对该多个寡核苷酸标签具有特异性。这些探针可以对该多个寡核苷酸标签的至少一部分具有特异性。这些探针可以对该多个寡核苷酸标签的总数的至少约5％、10％、20％、30％、40％、50％、60％、70％、80％、90％、95％、97％或100％具有特异性。可替代地，这些探针对该多个寡核苷酸标签的不同寡核苷酸标签的总数的至少约5％、10％、20％、30％、40％、50％、60％、70％、80％、90％、95％、97％或100％具有特异性。这些探针可以是寡核苷酸。这些寡核苷酸可以为至少约1、2、5、10、20、30、40、50、60、70、80、90、100个核苷酸长。在其他情况下，这些探针是非特异性探针。例如，这些探针可以对附接至经标记的分子的可检

测标记具有特异性。探针可以是链霉亲和素。

[0168] 阵列可以是打印的阵列。在一些情况下，打印的阵列包括附接至基底的一个或多个寡核苷酸。例如，打印的阵列包括附接至环氧硅烷基底的5'胺修饰的寡核苷酸。

[0169] 可替代地，该阵列包括具有一个或多个孔的载玻片。该载玻片可以包括至少约2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、25、30、35、40、45、50、55、60、65、70、75、80、85、90、95或100个孔。可替代地，该载玻片包括至少约125、150、175、200、250、300、350、400、450、500、650、700、750、800、850、900、950或1000个孔。在一些情况下，该载玻片包括16个孔。可替代地，该载玻片包括96个孔。在其他情况下，该载玻片包括至少约80、160、240、320、400、480、560、640、720、800、880或960个孔。

[0170] 在一些情况下，固体支持物是昂飞(Affymetrix)3K标签阵列、阿瑞杰特(Arrayjet)非接触打印阵列或应用微阵列公司(Applied Microarrays Inc)(AMI)阵列。可替代地，支持物包括接触式打印机、击打式打印机、点阵式打印机或针式打印机。

[0171] 固体支持物可以包括使用在微孔中自组装的珠。例如，固体支持物包括亿明达的珠阵列技术(BeadArray Technology)。可替代地，固体支持物包括雅培分子公司(Abbott Molecular)的珠阵列技术和应用微阵列公司的FlexiPlexTM系统。

[0172] 在其他情况下，固体支持物是平板。平板的实例包括但不限于MSD多阵列平板、MSD Multi-Spot®平板、微孔板、ProteOn微孔板、阿尔法平板(AlphaPlate)、DELFIA平板、艾搜平板(IsoPlate)以及鲁马平板(LumaPlate)。

[0173] 该方法可以进一步包括将多个经标记的核酸中的至少一个附接至支持物。支持物可以包括多个珠。支持物可以包括阵列。支持物可以包括载玻片。

[0174] 该载玻片可以包括一个或多个孔。可以在载玻片上蚀刻一个或多个孔。该一个或多个孔可以包括至少960个孔。该载玻片可以包括一个或多个探针。可以将该一个或多个探针打印到该载玻片上。该一个或多个孔可以进一步包括一个或多个探针。可以将该一个或多个探针打印到该一个或多个孔内。一个或多个探针可以包括960个核酸。

[0175] 在此披露的方法和试剂盒可以进一步包括将多个第一样品标签、多个第二样品标签、多个分子鉴定物标记或其任何组合分配进微孔板中。在此披露的方法和试剂盒可以进一步包括将一个或多个珠分配进微孔板中。在此披露的方法和试剂盒可以进一步包括将多个样品分配进微孔板的多个孔中。该多个样品中的一个或多个可以包括多个细胞。该多个样品中的一个或多个可以包括多个核酸。该方法可以进一步包括将一个或更少的细胞分配至多个孔中。可以在微孔板中裂解该多个细胞。该方法可以进一步包括在微孔板中合成cDNA。合成cDNA可以包括mRNA的逆转录。微孔板可以包括在PDMS上通过软光刻制造的、在硅晶片上蚀刻的、在载玻片上蚀刻的、在载玻片上光刻胶图案化或其组合的微孔板。微孔可以包括在微毛细管板上的洞。微孔板可以包括油包水乳剂。微孔板可以包括至少一个或多个孔。微孔板可以包括至少约6个孔、12个孔、48个孔、96个孔、384个孔、960个孔或1000个孔。

[0176] 这些方法和试剂盒可以进一步包括芯片。微孔板可以附接至芯片。芯片可以包括至少约6个孔、12个孔、48个孔、96个孔、384个孔、960个孔、1000个孔、2000个孔、3000个孔、4000个孔、5000个孔、6000个孔、7000个孔、8000个孔、9000个孔、10,000个孔、20,000个孔、30,000个孔、40,000个孔、50,000个孔、60,000个孔、70,000个孔、80,000个孔、90,000个孔、100,000个孔、200,000个孔、500,000个孔或一百万个孔。孔可以包括至少约300微米^2、400

微米^2、500微米^2、600微米^2、700微米^2、800微米^2、900微米^2、1000微米^2、1100微米^2、1200微米^2、1300微米^2、1400微米^2、1500微米^{2}的面积。该方法可以进一步包括在芯片上分配约10,000与30,000个之间的样品。

[0177] 官能化的表面和寡核苷酸

[0178] 珠可以包括官能化的表面。官能化的表面可以指固体支持物的包含官能团的表面。官能团可以是能够与另一官能团形成附接的基团。例如,官能团可以是生物素,其可以与另一官能团链霉亲和素形成附接。示例性官能团可以包括但不限于醛、酮、羧基基团、氨基基团、生物素、链霉亲和素、核酸、小分子(例如,用于点击化学)、同双功能和异双功能试剂(例如,N-琥珀酰亚胺基(4-碘代乙酰基)氨基苯甲酸酯(SIAB)、二顺丁烯二酰亚胺、二硫代-双-硝基苯甲酸(DTNB)、N-琥珀酰亚胺基-S-乙酰基-硫代乙酸酯(SATA)、N-琥珀酰亚胺基-3-(2-吡啶基二硫代)丙酸酯(SPDP)、琥珀酰亚胺基4-(N-马来酰亚胺基甲基)-环己烷-1-甲酸酯(SMCC)和6-肼基烟酰胺(hydrazinonicotimide)(HYNIC))以及抗体。在一些情况下,官能团是羧基基团(例如,COOH)。

[0179] 寡核苷酸(例如,核酸)可以附接至官能化的固体支持物。固体支持物或类似结构上的固定的寡核苷酸可以用作核酸探针,并且可以进行杂交测定,其中可以在复杂的生物样品中检测特异性靶核酸。

[0180] 可以将固体支持物(例如,珠)官能化,用于固定寡核苷酸。寡核苷酸可以通过固体支持物与该寡核苷酸之间形成的共价酰胺键缀合至该固体支持物。

[0181] 支持物可以缀合至至少约10、20、30、40、50、60、70、80、90、100、200、300、400、500、600、700、800、900或1000或更多个寡核苷酸。支持物可以缀合至至少约100000、200000、300000、400000、500000、600000、700000、800000、900000、1000000、2000000、3000000、4000000、5000000、6000000、7000000、8000000、9000000或10000000、100000000、500000000、1000000000或更多个寡核苷酸。支持物可以缀合至至少约100000、200000、300000、400000、500000、600000、700000、800000、900000、1000000、2000000、3000000、4000000、5000000、6000000、7000000、8000000、9000000或10000000、100000000、500000000、1000000000或更多个寡核苷酸。支持物可以缀合至至少一百万个寡核苷酸。支持物可以缀合至至少一千万个寡核苷酸。支持物可以缀合至至少两千五百万个寡核苷酸。支持物可以缀合至至少五千万个寡核苷酸。支持物可以缀合至至少一亿个寡核苷酸。支持物可以缀合至至少2.5亿个寡核苷酸。支持物可以缀合至至少5亿个寡核苷酸。支持物可以缀合至至少7.5亿个寡核苷酸。支持物可以缀合至至少约十亿、二十亿、三十亿、四十亿、五十亿、六十亿、七十亿、八十亿、九十亿、100亿、110亿、120亿、130亿、140亿或150亿个寡核苷酸。支持物可以缀合至至少十亿个寡核苷酸。支持物可以缀合至至少五十亿个寡核苷酸。

[0182] 寡核苷酸可以经由接头附接至支持物(例如,珠、聚合物、凝胶)。缀合可以包括共价或非共价附接。缀合可以在珠与核酸之间引入可变间隔物。支持物与寡核苷酸之间的接头是可切割的(例如,光可切割连接、酸不稳定接头、热敏接头以及可酶促切割的接头)。

[0183] 用于将分子缀合至支持物的交联剂可以包括能够与固体支持物表面上存在的功能团并与分子中存在的官能团反应的试剂。能够具有这样的反应性的试剂可以包括醛、酮、羧基基团、氨基基团、生物素、链霉亲和素、核酸、小分子(例如,用于点击化学)、同双功能和异双功能试剂(例如,N-琥珀酰亚胺基(4-碘代乙酰基)氨基苯甲酸酯(SIAB)、二顺丁烯二酰

亚胺、二硫代-双-硝基苯甲酸(DTNB)、N-琥珀酰亚胺基-S-乙酰基-硫代乙酸酯(SATA)、N-琥珀酰亚胺基-3-(2-吡啶基二硫代)丙酸酯(SPDP)、琥珀酰亚胺基4-(N-马来酰亚胺基甲基)-环己烷-1-甲酸酯(SMCC)和6-肼基烟酰胺(HYNIC))。

[0184] 珠可以被羧基官能团官能化并且寡核苷酸可以被氨基官能团官能化。

[0185] 支持物可以是光滑的。可替代地或另外地,支持物可以包括凹、脊或孔。支持物可以包括微孔阵列。微孔阵列可以被有助于寡核苷酸的附接的官能团官能化。微孔阵列上的官能团对于微孔阵列上的不同位置而言可以是不同的。微孔阵列上的官能团对于微孔阵列上的所有区域而言可以是相同的。

[0186] 测定系统部件

[0187] 微孔阵列

[0188] 如上所述的,微孔阵列被用来将单细胞和珠(一个珠/细胞)截留在具有限定体积的小反应室内。每个珠包括用于随机标记细胞mRNA分子的整个互补体并对其进行数字计数的寡核苷酸探针文库,当细胞裂解后的这些探针被释放。在本披露的一个实施例中,微孔阵列是测定系统的消耗性部件。在其他实施例中,微孔阵列可以是可重复使用的。在任一情况下,它们可以被配置成被用作用于手动进行测定的独立式装置,或者它们可以被配置成包括提供测定程序的完全或部分自动的仪器的可移除或固定部件。

[0189] 阵列的微孔可以被制造成多种形状和尺寸,对其进行选择以便优化在每个孔中俘获单细胞和珠的效率。适当的孔几何形状包括但不限于圆柱形、圆锥形,半球形、矩形或多面体(例如,由若干平面组成的三维几何体,例如六角柱、八角柱、倒三棱锥、倒四棱锥、倒五棱锥、倒六棱椎或倒截棱锥)。微孔可以包括组合了这些几何形状中的两种或更多种的形状。例如,在一个实施例中,它可以是部分圆柱形的,其余部分具有倒锥形的形状。在另一个实施例中,它可以包括两个并列圆柱体,一个的直径大于另一个,这两个圆柱体通过延伸圆柱体的全长(深度)的竖直通道(即,平行于圆柱轴)连接。通常,每个微孔的开口端(或口)被定位在微孔阵列的上表面上,但是在一些实施例中,开口可以被定位阵列的下表面上。通常,微孔的封闭端(或底部)是平的,但曲面(例如,凸出或凹陷)也是可能的。通常,基于有待俘获在微孔中的细胞和/或珠的类型确定微孔的形状(和尺寸)。

[0190] 可以就孔的直径和深度而言来表征微孔尺寸。如在此使用的,微孔的直径是指可以内切于微孔几何体的平面横截面内的最大圆。在本披露的一个实施例中,微孔的直径范围可以是有待俘获在微孔内的细胞和/或珠的直径的从约0.1倍至约5倍。在其他实施例中,微孔直径是有待俘获在微孔内的细胞和/或珠的直径的至少0.1倍、至少0.5倍、至少1倍、至少2倍、至少3倍、至少4倍或至少5倍。在仍其他实施例中,微孔直径是有待俘获在微孔内的细胞和/或珠的直径的至多5倍、至多4倍、至多3倍、至多2倍、至多1倍、至多0.5倍或至多0.1倍。在一个实施例中,微孔直径是有待俘获在微孔内的细胞和/或珠的直径的约2.5倍。本领域的普通技术人员应理解的是,微孔直径可以落入由任何这些值所限制的任何范围内(例如,有待俘获在微孔内的细胞和/或珠的直径的从约0.2倍至约3.5倍)。可替代地,可以就绝对尺寸而言指定微孔的直径。在本披露的一个实施例中,微孔的直径范围可以是从约5至约50微米。在其他实施例中,微孔直径是至少5微米、至少10微米、至少15微米、至少20微米、至少25微米、至少30微米、至少35微米、至少40微米、至少45微米或至少50微米。在仍其他实施例中,微孔直径是至多50微米、至多45微米、至多40微米、至多35微米、至多30微米、至多25

微米、至多20微米、至多15微米、至多10微米或至多5微米。在一个实施例中，微孔直径是约30微米。本领域的普通技术人员应理解的是，微孔直径可以落入由任何这些值所限制的任何范围内(例如，从约28微米至约34微米)。

[0191] 对微孔深度进行选择，以便优化细胞和珠俘获效率同时还提供包含在孔内的测定缓冲液和其他试剂的有效交换。在本披露的一个实施例中，微孔的深度范围可以是有待俘获在微孔内的细胞和/或珠的直径的从约0.1倍至约5倍。在其他实施例中，微孔深度是有待俘获在微孔内的细胞和/或珠的直径的至少0.1倍、至少0.5倍、至少1倍、至少2倍、至少3倍、至少4倍或至少5倍。在仍其他实施例中，微孔深度是有待俘获在微孔内的细胞和/或珠的直径的至多5倍、至多4倍、至多3倍、至多2倍、至多1倍、至多0.5倍或至多0.1倍。在一个实施例中，微孔深度是有待俘获在微孔内的细胞和/或珠的直径的约2.5倍。本领域的普通技术人员应理解的是，微孔深度可以落入由任何这些值所限制的任何范围内(例如，有待俘获在微孔内的细胞和/或珠的直径的从约0.2倍至约3.5倍)。可替代地，可以就绝对尺寸而言指定微孔的深度。在本披露的一个实施例中，微孔的深度范围可以是从约10至约60微米。在其他实施例中，微孔深度是至少10微米、至少20微米、至少25微米、至少30微米、至少35微米、至少40微米、至少50微米或至少60微米。在仍其他实施例中，微孔深度是至多60微米、至多50微米、至多40微米、至多35微米、至多30微米、至多25微米、至多20微米或至多10微米。在一个实施例中，微孔深度是约30微米。本领域的普通技术人员应理解的是，微孔深度可以落入由任何这些值所限制的任何范围内(例如，从约24微米至约36微米)。

[0192] 微孔阵列的孔被安排在一维、二维或三维阵列中，其中三维阵列可以例如通过堆积一系列的两个或更多个二维阵列(即，通过堆积包含微孔阵列的两个或更多个基底)来实现。对孔之间的模式和间隔进行选择，以便优化在每个孔中俘获单细胞和珠的效率，并且最大化阵列的每单位面积的孔数目。可以根据多种随机或非随机模式来散布孔，例如，它们可以在阵列基底的表面上完全随机地分布，或者它们可以被安排成方形网格、矩形网格或六角网格。在本披露的一个实施例中，孔之间的中心距(或间隔)可以从约15微米至约75微米变化。在其他实施例中，孔之间的间隔是至少15微米、至少20微米、至少25微米、至少30微米、至少35微米、至少40微米、至少45微米、至少50微米、至少55微米、至少60微米、至少65微米、至少70微米或至少75微米。在仍其他实施例中，微孔间隔是至多75微米、至多70微米、至多65微米、至多60微米、至多55微米、至多50微米、至多45微米、至多40微米、至多35微米、至多30微米、至多25微米、至多20微米或至多15微米。在一个实施例中，微孔间隔是约55微米。本领域的普通技术人员应理解的是，微孔间隔可以落入由任何这些值所限制的任何范围内(例如，从约18微米至约72微米)。

[0193] 微孔阵列在微孔之间可以包括表面特征，其被设计成帮助引导细胞和珠进入孔和/或防止它们沉降在孔之间的表面上。适合的表面特征的实例包括但不限于环绕孔和/或跨过孔之间的表面的隆起的、呈脊形的或尖形的表面特征。

[0194] 微孔阵列中的孔的总数目由孔的模式和间隔以及阵列的总尺寸决定。在本披露的一个实施例中，阵列中的微孔数目的范围可以是从约96至约5,000,000或更高。在其他实施例中，阵列中的微孔数目是至少96、至少384、至少1,536、至少5,000、至少10,000、至少25,000、至少50,000、至少75,000、至少100,000、至少500,000、至少1,000,000或至少5,000,000。在仍其他实施例中，阵列中的微孔数目是至多5,000,000、至多1,000,000、至多75,

000、至多50,000、至多25,000、至多10,000、至多5,000、至多1,536、至多384或至多96个孔。在一个实施例中,阵列中的微孔数目是约96。在另一个实施例中,微孔数目是约150,000。本领域的普通技术人员应理解的是,阵列中的微孔数目可以落入由任何这些值所限制的任何范围内(例如,从约100至325,000)。

[0195] 可以使用本领域的普通技术人员已知的多种制造技术中的任一种来制造微孔阵列。可以使用的制造方法的实例包括但不限于体微机械加工技术,如光刻法和湿化学蚀刻、等离子蚀刻或深反应离子蚀刻;微模塑和微压花;激光微机械加工;3D打印或使用可固化材料的其他直接写入制造工艺;以及类似技术。

[0196] 微孔阵列可以由本领域的普通技术人员已知的多种基底材料中的任一种制造而来,其中材料的选择典型地取决于制造技术的选择,反之亦然。适合的材料的实例包括但不限于硅、熔融二氧化硅、玻璃、聚合物(例如,琼脂糖、明胶、水凝胶、聚二甲基硅氧烷(PDMS;弹性体)、聚甲基丙烯酸甲酯(PMMA)、聚碳酸酯(PC)、聚丙烯(PP)、聚乙烯(PE)、高密度聚乙烯(HDPE)、聚酰亚胺、环烯烃聚合物(COP)、环烯烃共聚物(COC)、聚对苯二甲酸乙二酯(PET)和环氧树脂)、金属或金属膜(例如,铝、不锈钢、铜、镍、铬和钛)以及类似物。典型地,亲水材料对于制造微孔阵列而言是令人希望的(以增强可湿性并最小化细胞和其他生物材料的非特异性结合),但是可以(例如,通过氧等离子体处理,或聚氧乙烯表层的接枝)可处理或涂覆的疏水材料也是可以使用的。使用多孔亲水材料来制造微孔阵列可以是令人希望的,以便有助于装置中截留的气泡的毛细管芯吸/通气。在一些实施例中,用光学胶制造微孔阵列。在一些实施例中,用经等离子体或电晕处理的材料制造微孔阵列。使用经等离子体或电晕处理的材料可以使材料亲水。在一些实施例中,经等离子体或电晕处理的材料(如亲水材料)可以比未处理的材料更稳定。在一些实施例中,微孔阵列由单一材料制造而来。在其他实施例中,微孔阵列可以包括两种或更多种已经结合在一起或机械连接的不同材料。

[0197] 多种表面处理和表面改性技术可以用于改变微孔阵列表面的特性。实例包括但不限于氧等离子体处理以使得疏水材料表面更亲水、使用湿或干蚀刻技术以使玻璃和硅表面光滑(或粗糙)、将聚氧乙烯或或其他聚合物层吸附和/或接枝到基底表面以使它们更亲水并且更不易于生物分子和细胞的非特异性吸附、使用硅烷反应以将化学反应性官能团接枝到另外的惰性硅和玻璃表面等。光脱保护技术可以用于选择性地活化阵列结构中的特定位置处的化学反应性官能团,例如,选择性添加或活化微孔的内壁上的化学反应性官能团(如伯胺或羧基基团)可以用于将寡核苷酸探针、肽、蛋白质或其他生物分子共价偶联至微孔壁。通常,利用的表面处理或表面改性的选择取决于所希望的表面特性的类型和制成微孔阵列的材料的类型两者。

[0198] 在一些实施例中,可能有利的是在例如细胞裂解步骤过程中将微孔的开口密封,以防止相邻微孔之间的靶核酸进行交叉杂交。可以使用帽(如固体支持物或珠)来密封微孔,其中珠的直径大于微孔的直径。例如,用作帽的珠可以比微孔的直径大至少约10%、20%、30%、40%、50%、60%、70%、80%或90%。可替代地,帽可以比微孔的直径大至多约10%、20%、30%、40%、50%、60%、70%、80%或90%。

[0199] 用作帽的珠可以包括交联葡聚糖珠(例如,交联葡聚糖(Sephadex))。交联葡聚糖的范围可以是从约10微米至约80微米。珠帽的交联葡聚糖可以是从20微米至约50微米。帽可以包括例如无机纳米孔膜(例如,氧化铝)、透析膜、载玻片、盖玻片和/或亲水塑料薄膜

(例如,用与裂解缓冲液水合的琼脂糖薄膜涂覆的薄膜)。

[0200] 在一些实施例中,帽可以允许缓冲液穿过进出微孔,同时阻止大分子(例如,核酸)从孔中迁移出来。通过帽可以阻挡至少约1、2、3、4、5、6、7、8、9、10、12、13、14、15、16、17、18、19或20或更多个核苷酸的大分子迁移进入或离开微孔。通过帽可以阻挡至多约1、2、3、4、5、6、7、8、9、10、12、13、14、15、16、17、18、19或20或更多个核苷酸的大分子迁移进入或离开微孔。

[0201] 在一些实施例中,密封的微孔阵列可以包括微孔顶部的单层珠。在一些实施例中,密封的微孔阵列可以包括微孔顶部的多层珠。密封的微孔阵列可以包括约1、2、3、4、5或6或更多层珠。

[0202] 机械固定器

[0203] 当手动进行多重、单细胞随机标记/分子索引测定时,方便的是将微孔阵列安装在机械固定器中,以便产生反应室并有助于将细胞悬浮液和测定试剂移液管移取或分发到阵列上(图69和70)。在图69说明的实例中,固定器接受在1mm厚的基底上制造的微孔阵列,并且提供呈硅酮垫圈形式的机械支持物,以便将测定试剂限制在16mm宽x 35mm长x大约4mm深的反应室中,由此允许使用800微升至1毫升的细胞悬浮液和珠悬浮液(包含基于珠的寡核苷酸标记)进行测定。

[0204] 固定器由刚性的机器加工的顶板和底板(例如,铝)以及用于产生室壁或孔的可压缩(例如,硅酮、二甲基硅氧烷)垫圈组成。设计特征包括:(i)用于根据需要将显微镜物镜旋转进出适当位置的倒角孔径边缘和间隙(用于在不同放大倍数下查看微孔阵列)。(ii)硅酮垫圈的可控压缩,以确保与微孔阵列基底统一地可重复的形成防漏密封。(iii)用于方便操作的系留紧固件。(iv)用于牢固且可重复地定位阵列的定位夹紧机构。(v)用于在漂洗步骤过程中移除阵列的方便拆卸。

[0205] 顶板和底板可以使用多种材料(例如,铝、阳极氧化铝、不锈钢、特氟龙、聚甲基丙烯酸甲酯(PMMA)、聚碳酸酯(PC)或类似的刚性聚合物材料)使用多种技术中的任一种(例如,常规机器加工、CNC机器加工、注塑、3D打印等)来制造。

[0206] 硅酮(聚二甲基硅氧烷;PDMS)垫圈可以被配置成产生多个室(参见图71),以便平行地运行对照和实验(或重复实验,或多个独立实验)。垫圈使用特氟龙模具由PDMS或类似的弹性体材料模塑而来,所述模具包括针对垂直垫圈的拔模角度以提供良好的释放特征。可替代地,模具可以由铝或其他材料(例如,黑迭尔林(delrin)、聚醚酰亚胺(ultem)等)机器加工而来,并且如果必要用特氟龙涂覆以提供良好的释放特征。垫圈模具设计是颠倒的,即,这样使得模制件的顶表面(即与在铸造过程中用于覆盖模具的、载玻片或硅晶片界面结合处的表面)在使用过程中变为用于与微孔阵列基底产生密封的表面,由此避免在产生光滑垫圈表面(以确保与阵列基底形成防漏密封)中模具表面粗糙度和表面污染的潜在问题,并且通过在铸造过程中使用微孔阵列基底作为基部还提供了基底材料的灵活选择和预组装的选择权。垫圈模具设计还可以包括孔区边界处的聚力脊(force focusing ridge),即模具中的一个或多个中央台面(mesa)(其形成一个或多个孔)在成为该一个或多个孔的周界的位置处具有凸起的脊,这样使得在填充后放置在模具顶部的盖停留在精确位置处的小接触面积上,在所述精确位置处良好的边缘轮廓对于在使用过程中在垫圈与基底之间形成防漏密封而言是关键的。

[0207] 仪器系统

[0208] 本披露还包括用于支持多重、单细胞随机标记/分子索引测定的自动化的仪器系统和消耗品。此类系统可以包括并入与流动池整合的微孔阵列的消耗性柱体，连同用于提供控制和分析功能性所必需的仪器装备，所述功能性是如(i)射流控制、(ii)温度控制、(iii)细胞和/或珠分配和收集机构、(iv)细胞裂解机构、(v)成像能力以及(vi)图像处理。在一些实施例中，系统的输入包括细胞样品并且输出包括包含具有附接的寡核苷酸的珠的珠悬浮液，所述寡核苷酸包含样品标签、细胞标签和分子索引标签。在其他实施例中，系统可以包括另外的功能性，如用于进行PCR扩增的热循环能力，在这种情况下系统的输入包括细胞样品并且输出包括由寡核苷酸的扩增产生的寡核苷酸文库，所述寡核苷酸包含最初附接至珠上的样品标签、细胞标签和分子索引标签。在仍其他实施例中，系统还可以包括测序能力，需要或不需要寡核苷酸扩增，在这种情况下，系统的输入是细胞样品并且输出包括数据集，所述数据集进一步包括与感兴趣的靶序列相关的所有样品标签、细胞标签和分子索引标签的序列。

[0209] 微孔阵列流动池

[0210] 在自动化测定系统的许多实施例中，微孔阵列基底被包装在流动池内，所述流动池提供与流体处理系统的其余部分的方便界面结合并且有助于被递送至微孔阵列的流体(例如，细胞和珠悬浮液、裂解缓冲液、漂洗缓冲液等)的交换。设计特征可以包括：(i)一个或多个用于引入细胞样品、珠悬浮液和/或其他测定试剂的入口，(ii)一个或多个被设计成提供均匀填充和有效流体交换同时最小化同涡或死区的微孔阵列室，以及(iii)一个或多个用于将流体递送到样品收集点和/或废物储器的出口。在一些实施例中，流动池的设计可以包括与多个微孔阵列界面结合的多个微孔阵列室，这样使得可以平行处理一个或多个细胞样品。在一些实施例中，流动池的设计可以进一步包括用于跨阵列室的宽度产生匀流速度剖面图(即“活塞流”)的特征，以提供细胞和珠到微孔的更均匀递送，例如，通过使用作为“流动扩散器(flow diffuser)”的位于室入口附近和微孔阵列上游的多孔隔板，或通过将每个阵列室分成若干分段，这些分段共同覆盖相同的总阵列区，但是通过它们分开的入口液流平行地流动。在一些实施例中，流动池可以围绕或并入多于一个微孔阵列基底。在一些实施例中，整合的微孔阵列/流动池组件可以构成系统的固定部件。在一些实施例中，微孔阵列/流动池可以是从仪器可移除的。

[0211] 通常，将优化流体通道和一个或多个阵列室在流动池设计中的尺寸，以便(i)提供细胞和珠到微孔阵列的均匀递送，和(ii)最小化样品和试剂消耗。在一些实施例中，流体通道的宽度将在50微米与20mm之间。在其他实施例中，流体通道的宽度可以是至少50微米、至少100微米、至少200微米、至少300微米、至少400微米、至少500微米、至少750微米、至少1mm、至少2.5mm、至少5mm、至少10mm或至少20mm。在仍其他实施例中，流体通道的宽度可以是至多20mm、至多10mm、至多5mm、至多2.5mm、至多1mm、至多750微米、至多500微米、至多400微米、至多300微米、至多200微米、至多100微米或至多50微米。在一个实施例中，流体通道的宽度是约2mm。本领域的普通技术人员应理解的是，流体通道的宽度可以落入由任何这些值所限制的任何范围内(例如，从约250微米至约3mm)。

[0212] 在一些实施例中，流体通道的深度将在50微米与10mm之间。在其他实施例中，流体通道的深度可以是至少50微米、至少100微米、至少200微米、至少300微米、至少400微米、至

少500微米、至少750微米、至少1mm、至少1.25mm、至少1.5mm、至少1.75mm、至少2mm、至少2.5mm、至少3mm、至少3.5mm、至少4mm、至少4.5mm、至少5mm、至少5.5mm、至少6mm、至少6.5mm、至少7mm、至少7.5mm、至少8mm、至少8.5mm、至少9mm或至少9.5mm。在其他实施例中，流体通道的深度可以是至多10mm、至多9.5mm、至多9mm、至多8.5mm、至多8mm、至多7.5mm、至多7mm、至多6.5mm、至多6mm、至多5.5mm、至多5mm、至多4.5mm、至多4mm、至多3.5mm、至多3mm、至多2mm、至多1.75mm、至多1.5mm、至多1.25mm、至多1mm、至多750微米、至多500微米、至多400微米、至多300微米、至多200微米、至多100微米或至多50微米。在一个实施例中，流体通道的深度是约1mm。本领域的普通技术人员应理解的是，流体通道的深度可以落入由任何这些值所限制的任何范围内(例如，从约800微米至约1mm)。

[0213] 流动池可以使用本领域的普通技术人员已知的多种技术和材料制造而来。通常，流动池被制造为单独零件并且随后被机械地夹固或永久地结合到微孔阵列基底上。适合的制造技术的实例包括常规机器加工、CNC机器加工、注塑、3D打印、对齐并层压一层或多层激光切割或冲切聚合物薄膜或多种微制造技术中的任一种，如光刻法和湿化学蚀刻、干蚀刻、深反应离子蚀刻或激光微机械加工。一旦流动池零件已经被制造出，则它可以机械地附接至微孔阵列基底，例如通过将它抵着微孔阵列基底夹紧(使用或未使用垫圈)，或者可以使用本领域的普通技术人员已知的多种技术中的任一种(取决于使用的材料的选择)使它直接结合到微孔阵列基底上，例如通过使用阳极键合、热键合、超声波焊接或多种粘合剂或粘附膜中的任一种，包括基于环氧树脂的、基于丙烯酸的、基于硅酮的、可UV固化的、基于聚氨酯的或基于氰基丙烯酸酯的粘合剂。

[0214] 流动池可以使用本领域的普通技术人员已知的多种材料制造而来。适合的材料的实例包括但不限于硅、熔融二氧化硅、玻璃、多种聚合物中的任一种(例如，聚二甲基硅氧烷(PDMS;弹性体)、聚甲基丙烯酸甲酯(PMMA)、聚碳酸酯(PC)、聚丙烯(PP)、聚乙烯(PE)、高密度聚乙烯(HDPE)、聚酰亚胺、环烯烃聚合物(COP)、环烯烃共聚物(COC)、聚对苯二甲酸乙二酯(PET)、环氧树脂)、金属(例如，铝、不锈钢、铜、镍、铬和钛)或这些材料的组合。

[0215] 柱体

[0216] 在自动化测定系统的许多实施例中，微孔阵列(具有或不具有附接流动池)将被包装在消耗性柱体内，所述柱体与仪器系统界面结合并且可以并入另外的功能性。柱体的设计特征可以包括(i)一个或多个入口，用于与仪器形成流体连接和/或将细胞样品、珠悬浮液和/或其他测定试剂手动引入柱体中;(ii)一个或多个旁路通道，即用于自计量细胞样品和珠悬浮液，以避免满溢和/或回流;(iii)一个或多个整合的微孔阵列/流动池组件、或一个或多个室，其内定位有一个或多个微阵列基底;(iv)整合的微型泵或其他流体致动机构，用于控制流体流经装置;(v)整合的微型阀，用于将预加载试剂隔区化和/或控制流体流动通过装置;(vi)用于为俘获空气提供逃逸路径的通气孔;(vii)一个或多个样品和试剂废物储器;(viii)一个或多个出口，用于与仪器形成流体连接和/或提供经处理的样品收集点;(ix)机械接口特征，用于相对于仪器系统重复性定位可移除的消耗性柱体，并且用于提供进入，这样使得外磁体可以与微孔阵列非常接近;(x)整合的温度控制部件和/或热接口，用于提供与仪器系统的良好热接触;以及(xi)光学接口特征，例如透明窗，用于对微孔阵列进行光学探询。在一些实施例中，柱体被设计成平行地处理多于一个样品。在装置的一些实施例中，柱体可以进一步包括适于与独立式PCR热循环仪和/或测序仪界面结合的一个或多个

可移除样品收集室。在装置的一些实施例中，柱体自身适于与独立式PCR热循环仪和/或测序仪界面结合。

[0217] 在装置的一些实施例中，柱体可以进一步包括部件，所述部件被设计成产生防止大分子扩散(或增加其路径长度和扩散时间)的物理和/或化学屏障，以便最小化微孔之间的交叉污染。此类屏障的实例包括但不限于用于将细胞和珠递送到微孔阵列的一套蛇形通道，在裂解或孵育步骤过程中经过按压与微孔阵列基底表面接触的可伸缩压板或可变形膜，使用较大的珠(例如如上所述的交联葡萄糖珠)来堵塞微孔的开口，或在裂解或孵育步骤过程中使不互混的疏水流体从柱体内的储器中释放出来以有效地分离并隔区化阵列中的每个微孔。任何或所有这些屏障或不具有此类屏障的实施例可以与增加微孔中的和邻近微孔的溶液的粘度组合，例如通过添加溶液组分如甘油或聚乙二醇。

[0218] 通常，将优化流体通道和一个或多个阵列室在柱体设计中的尺寸，以便(i)提供细胞和珠到微孔阵列的均匀递送，和(ii)最小化样品和试剂消耗。在一些实施例中，流体通道的宽度将在50微米与20mm之间。在其他实施例中，流体通道的宽度可以是至少50微米、至少100微米、至少200微米、至少300微米、至少400微米、至少500微米、至少750微米、至少1mm、至少2.5mm、至少5mm、至少10mm或至少20mm。在仍其他实施例中，流体通道的宽度可以是至多20mm、至多10mm、至多5mm、至多2.5mm、至多1mm、至多750微米、至多500微米、至多400微米、至多300微米、至多200微米、至多100微米或至多50微米。在一个实施例中，流体通道的宽度是约2mm。本领域的普通技术人员应理解的是，流体通道的宽度可以落入由任何这些值所限制的任何范围内(例如，从约250微米至约3mm)。

[0219] 在一些实施例中，流体通道在柱体设计中的深度将在50微米与10mm之间。在其他实施例中，流体通道的深度可以是至少50微米、至少100微米、至少200微米、至少300微米、至少400微米、至少500微米、至少750微米、至少1mm、至少1.25mm、至少1.5mm、至少1.75mm、至少2mm、至少2.5mm、至少3mm、至少3.5mm、至少4mm、至少4.5mm、至少5mm、至少5.5mm、至少6mm、至少6.5mm、至少7mm、至少7.5mm、至少8mm、至少8.5mm、至少9mm或至少9.5mm。在仍其他实施例中，流体通道的深度可以是至多10mm、至多9.5mm、至多9mm、至多8.5mm、至多8mm、至多7.5mm、至多7mm、至多6.5mm、至多6mm、至多5.5mm、至多5mm、至多4.5mm、至多4mm、至多3.5mm、至多3mm、至多2mm、至多1.75mm、至多1.5mm、至多1.25mm、至多1mm、至多750微米、至多500微米、至多400微米、至多300微米、至多200微米、至多100微米或至多50微米。在一个实施例中，流体通道的深度是约1mm。本领域的普通技术人员应理解的是，流体通道的深度可以落入由任何这些值所限制的任何范围内(例如，从约800微米至约1mm)。

[0220] 柱体可以使用本领域的普通技术人员已知的多种技术和材料制造而来。通常，使用多种机械组装或结合技术中的任一种将柱体制造为一系列单独的零部件(图72)并且随后组装(图72和73)。适合的制造技术的实例包括但不限于常规机器加工、CNC机器加工、注塑、热成形以及3D打印。一旦柱体部件已经被制造出，可以使用螺钉、夹子等将它们进行机械组装，或使用多种技术中的任一种(取决于使用的材料的选择)使它们永久结合，例如通过使用热或超声波键合/焊接或多种粘合剂或粘附膜中的任一种，包括基于环氧树脂的、基于丙烯酸的、基于硅酮的、可UV固化的、基于聚氨酯的或基于氰基丙烯酸酯的粘合剂。

[0221] 柱体部件可以使用多种适合的材料中的任一种制造而来，包括但不限于硅、熔融二氧化硅、玻璃、多种聚合物中的任一种(例如，聚二甲基硅氧烷(PDMS;弹性体)、聚甲基丙

烯酸甲酯(PMMA)、聚碳酸酯(PC)、聚丙烯(PP)、聚乙烯(PE)、高密度聚乙烯(HDPE)、聚酰亚胺、环烯烃聚合物(COP)、环烯烃共聚物(COC)、聚对苯二甲酸乙二酯(PET)、环氧树脂)或金属(例如,铝、不锈钢、铜、镍、铬和钛)。

[0222] 如上所述的,柱体的入口和出口特征可以被设计成提供与仪器的方便且防漏的流体连接,或可以用作开放储器,用于将样品和试剂手动地用移液管移取进出柱体。入口和出口连接器的方便机械设计的实例包括但不限于螺纹连接器、模锻连接器、鲁尔锁紧连接器、鲁尔滑锁(Luer slip)或"滑动型(slip tip)"连接器、压入配合连接器以及类似物。在一些实施例中,柱体的入口和出口可以进一步包括帽、弹簧加载盖或塞、相变材料或聚合物膜,当柱体被定位在该仪器中时,其被打开或刺破,并且用于防止在储存过程中柱体内表面的污染和/或当柱体被从仪器上移除时防止流体溢出。如上所指出的,在一些实施例中,柱体的一个或多个出口可以进一步包括可移除的样品收集室,其适于与独立式PCR热循环仪和/或测序仪界面结合。

[0223] 如上所指出的,在一些实施例中,柱体可以包括整合的微型泵或其他流体致动机构,用于控制流体流经装置。适合的微型泵或流体致动机构的实例包括但不限于机电或气动致动的微型注射器或柱塞机构、化学推进剂、气动地或通过外部活塞致动的膜隔膜泵、气动致动的试剂小袋或气囊、或电渗泵。

[0224] 如上所述的,在一些实施例中,柱体可以包括微型阀,用于将预加载试剂隔区化和/或控制流体流经装置。适合的微型阀的实例包括但不限于使用可以被熔化或溶解的蜡或聚合物塞或可以被刺破的聚合物膜制造的一次性"阀";使用可变形膜构造的夹管阀和使用可变形膜片构造的气动、液压、磁、电磁或机电(螺线管)致动的止回阀以及微型门阀。

[0225] 如上所指出的,在一些实施例中,柱体可以包括用于为俘获空气提供逃逸路径的通气孔。可以根据本领域的普通技术人员已知的多种技术构造通气孔,例如使用聚二甲基硅氧烷(PDMS)或其他疏水材料的多孔塞,所述材料允许毛细管芯吸空气但阻止被水渗透。通气孔还可以通过疏水阻隔材料被构造为孔,这样使得在操作过程中在所用压力下不会发生孔壁的润湿。

[0226] 通常,柱体的机械接口特征提供了柱体相对于仪器系统的易于移除但高度精确且可重复的定位。适合的机械接口特征包括但不限于定位销、定位导件、机械止动件以及类似物。在一些实施例中,机械设计特征将包括凸凹特征,用于使外部设备(例如,磁体或光学部件)与微孔阵列室非常接近(图72)。

[0227] 在一些实施例中,柱体还包括温度控制部件或热接口特征,用于与外部温度控制模块配合。适合的温度控制元件的实例包括但不限于电阻加热元件、微型红外发射光源、珀耳帖(Peltier)加热或冷却装置、散热器、热敏电阻、热电偶以及类似物。热接口特征将典型地由作为良好热导体的材料(例如,铜、金、银、铝等)制成并且将典型地包括一个或多个能够与外部加热块或冷却块产生良好热接触的平表面。

[0228] 在许多实施例中,柱体将包括光学接口特征,用于对微孔阵列进行光学成像或光谱学探询。典型地,柱体将包括光学透明窗,例如,微孔基底本身或与微孔阵列相对的流动池的或微阵列室的侧面,所述透明窗由满足用于探测微孔阵列的成像或光谱技术的频谱要求的材料制成。适合的光学窗材料的实例包括但不限于玻璃、熔融二氧化硅、聚甲基丙烯酸甲酯(PMMA)、聚碳酸酯(PC)、环烯烃聚合物(COP)或环烯烃共聚物(COC)。典型地,柱体将包

括第二光学透明或半透明窗或区域，其可以用于以横向、反射或倾斜照明方向为微孔阵列照明。

[0229] 仪器

[0230] 本披露还包括用于在多重、单细胞随机标记/分子索引测定自动化中使用的仪器。如上所指出的，这些仪器可以提供控制和分析功能性，如(i)射流控制、(ii)温度控制、(iii)细胞和/或珠分配和收集机构、(iv)细胞裂解机构、(v)磁场控制、(vi)成像能力以及(vii)图像处理。在一些实施例中，仪器系统可以包括一个或多个模块(其一个可能的实施例示意性地说明于图74中)，其中每个模块为系统提供一个或多个特定功能特征集。在其他实施例中，仪器系统可以被包装，这样使得所有系统功能性位于同一包装内。图75提供了包括在自动化系统的一个实施例中的工艺步骤的示意性说明。如上所指出的，在一些实施例中，系统可以包括另外的功能单元，作为该系统的集成部件或作为模块化部件，其将系统的功能能力扩展到包括PCR扩增(或其他类型的寡核苷酸扩增技术)和寡核苷酸测序。

[0231] 通常，仪器系统将提供射流能力，用于将样品和/或试剂递送到与该系统连接的一个或多个测定柱体内的一个或多个微阵列室或流动池中。测定试剂和缓冲液可以被存储在瓶、试剂和缓冲液药筒或其他与柱体入口连接的适合容器中。系统还可以包括废物储器，其所处的形式为瓶、废物药筒或其他用于收集一个或多个测定柱体的下游流体的适合废物容器。控制流体流动通过系统将典型地通过使用泵(或其他流体致动机构)和阀进行。适合的泵的实例包括但不限于注射泵、可编程注射泵、蠕动泵、隔膜泵以及类似物。在一些实施例中，可以借助在试剂和缓冲液容器的一个或多个入口处或在一个或多个测定柱体的入口处施加正气动压来控制流体流动通过系统。在一些实施例中，可以借助在废物储器的一个或多个出口处或在一个或多个测定柱体的出口处抽真空来控制流体流动通过系统。适合的阀的实例包括但不限于止回阀、机电两通或三通阀、气动两通和三通阀以及类似物。在一些实施例中，可以在测定洗涤/漂洗步骤过程中应用脉动流，以有助于一个或多个微孔阵列流动池或室内的流体的完全且有效的交换。

[0232] 如上所指出的，在一些实施例中，仪器系统可以包括用于进一步有助于在微孔阵列上均匀分配细胞和珠的机构。此类机构的实例包括但不限于振动、振荡、漩涡、再循环流、低频搅拌(例如，使用摇板或通过脉冲弹性(例如，硅酮)膜，所述膜形成室壁或在流体通道附近)或高频搅拌(例如，通过使用压电换能器)。在一些实施例中，将这些机构中的一种或多种与流动池或阵列室的内壁上的物理结构或特征组合应用，例如，夹层/顶帽结构、V状物(chevron)或脊阵列，以有助于在阵列室内混合细胞或珠和/或帮助阻止其聚池化。流动池或阵列室的上表面或下表面上的流动增强肋状物可以用来控制流动速度谱并且减小跨微孔开口的剪切力(即，为了防止在试剂交换和漂洗步骤过程中将细胞或珠从微孔中拉出)。

[0233] 在一些实施例中，仪器系统可以包括机械细胞裂解能力，作为使用洗涤剂或其他试剂的替代方案。使用高频压电换能器进行声处理是适合技术的一个实例。

[0234] 在一些实施例中，仪器系统将包括温度控制功能性，用于有助于测定结果的准确度和再现性的目的，例如，冷却微孔阵列流动池或室对于最小化微孔之间的分子扩散可以是有利的。可以被并入仪器系统中的温度控制部件的实例包括但不限于电阻加热元件、红外光源、珀耳帖加热或冷却装置、散热器、热敏电阻、热电偶以及类似物。在该系统的一些实施例中，温度控制器可以提供在规定的时间间隔内的可编程温度变化。

[0235] 如在本披露的其他地方所指出的，所披露的方法的许多实施例都利用磁场来在完成测定后将珠从微孔中移出。在一些实施例中，仪器系统可以进一步包括使用磁场来将珠输送进出微孔阵列流动池或室。用于提供磁场控制的适合手段的实例包括但不限于相对于柱体在一个或多个固定位置中使用电磁体，或必要时使用被机械重新定位的永磁体。在仪器系统的一些实施例中，施加的一个或多个磁场的强度将通过改变施加至一个或多个电磁体的电流量而改变。在仪器系统的一些实施例中，施加的磁场的强度被通过使用例如步进电机驱动的线性致动器、伺服电机驱动的线性致动器或凸轮轴机构改变相对于一个或多个微阵列室的位置的一个或多个永磁体的位置而改变。在仪器系统的一些实施例中，使用脉冲磁场可能是有利的，例如，以防止磁珠簇集。在一些实施例中，非常接近阵列或室的磁体可以在至少两个位置之间相对于微孔阵列被移动一次或多次。磁体的运动可以用于搅动微孔内的珠，以有助于将珠从微孔中移出，或以收集希望的位置处的磁珠。

[0236] 如上所指出的，在许多实施例中，仪器系统将包括光学成像和/或其他光谱能力。这样的功能性可以有用于例如检查一个或多个微孔阵列，以确定该阵列是否已经被细胞和/或珠均匀且最佳地填充。可以应用多种成像模式中的任一种，包括但不限于明场、暗场以及荧光/发光成像。成像模式的选择将影响微孔阵列、流动池和柱体室的设计，因为流动池或阵列室的阵列基底和/或相对壁需要在感兴趣的光谱范围内必须是透明或半透明的。在一些实施例中，可以在单一图像中将每个微孔阵列整个成像。在一些实施例中，可以将一系列图像“平铺”，以产生整个阵列的高分辨率图像。在一些实施例中，代表阵列的子分段的单一图像可以用于整体上评价阵列的特性(例如，细胞或珠分布)。在一些实施例中，可以进行双波长激发和发射(或多波长激发和/或发射)成像。多种光源中的任一种都可以用于提供成像和/或激发光，包括但不限于钨灯、卤钨灯、弧光灯、激光器、发光二极管(LED)或激光二极管。多种图像传感器中的任一种都可以用于成像目的，包括但不限于光电二极管阵列、电荷耦合装置(CCD)摄像机或CMOS图像传感器。光学系统将典型地包括多种光学部件，用于使穿过系统的光束转向、成形、过滤和/或聚焦。适合的光学部件的实例包括但不限于透镜、反射镜、棱镜、衍射光栅、着色玻璃滤光片、窄带干涉滤光片、宽带干涉滤光片、分色反射镜、光学纤维、光学波导以及类似物。在一些实施例中，仪器系统可以使用光学透明的微阵列基底作为波导，用于将激发光递送到微孔阵列。成像模式的选择还可以允许与随机标记/分子索引测定平行地使用有待运行的其他类型的测定，例如，与明场成像平行地使用台盼蓝活细胞/死细胞测定，与荧光成像平行地使用基于荧光的活细胞/死细胞测定等。相关微孔中的单个细胞的生存力数据和与每个珠相关的细胞标签的相关性可以在分析来自多重、单细胞测定的数据方面提供另外的辨别标准。可替代地，呈多个细胞的统计形式的生存力数据可以用于增强测定的分析能力和质量保证。

[0237] 在一些实施例中，系统可以包括用于探测微孔阵列的非成像和/或非光学能力。用于检测俘获的气泡、确定阵列中的细胞和/或珠分布等的非成像和/或非光学技术的实例包括但不限于测量光散射、紫外/可见/红外吸收测量(例如，使用掺入染料的经染色的细胞和/或珠)、相干拉曼(raman)散射以及电导测量(例如，使用与微孔阵列配准的微制造电极阵列)。

[0238] 系统处理器和软件

[0239] 通常，被设成支持多重、单细胞随机标记/分子索引测定的自动化的仪器系统将包

括处理器或计算机连同软件，以提供(i)仪器控制功能性、(ii)图像处理和分析能力以及(iii)数据存储、分析和显示功能性。

[0240] 在许多实施例中，仪器系统将包括计算机(或处理器)和计算机可读介质，所述介质包括用于提供用户界面连同所有系统功能的手动、半自动化或全自动化控制的代码，即射流系统、温度控制系统、细胞和/或珠分配功能、磁珠操纵功能以及成像系统的控制。由仪器控制软件提供的流体控制功能的实例包括但不限于流体体积流量，流体流速，样品和珠添加、试剂添加的时机和持续时间以及漂洗步骤。由仪器控制软件提供的温度控制功能的实例包括但不限于指定一个或多个温度设定点以及温度变化的时机、持续时间和斜坡率的控制。由仪器控制软件提供的细胞和/或珠分配功能的实例包括但不限于搅拌参数的控制，如幅度、频率和持续时间。由仪器控制软件提供的磁场功能的实例包括但不限于施加的一个或多个磁场的时机和持续时间，并且在电磁体的情况下，还有磁场强度。由仪器控制软件提供的成像系统控制功能的实例包括但不限于自动对焦能力，照明和/或激发光暴露时间和强度的控制，图像采集速率、暴露时间的控制以及数据存储选项。

[0241] 在仪器系统的一些实施例中，该系统将进一步包括计算机可读介质，其包括用于提供图像处理和分析能力的代码。由软件提供的图像处理和分析能力的实例包括但不限于手动、半自动化或全自动化图像曝光调整(例如，白平衡、反差调整、信号平均和其他降噪能力等)，自动化目标鉴定(即用于鉴定图像中的细胞和珠)，自动化统计分析(即用于确定微孔阵列的每单位面积鉴定出的细胞和/或珠的数目，或用于鉴定含有多于一个细胞或多于一个珠的孔)以及手动测量能力(例如，用于测量目标之间的距离等)。在一些实施例中，仪器控制和图像处理/分析软件将被写为独立的软件模块。在一些实施例中，仪器控制和图像处理/分析软件将被并入集成包装中。在一些实施例中，系统软件可以提供整合的实时图像分析和仪器控制，这样使得细胞和珠样品加载步骤可以被延长或重复，直到达到最佳细胞/珠分布。

[0242] 在仪器系统的一些实施例中，该系统将包括计算机可读介质，其包括用于提供序列数据分析的代码。可以由数据分析软件提供的测序数据分析功能性的实例包括但不限于(i)算法，用于基于由通过运行测定法产生的寡核苷酸文库的测序提供的数据来确定读数数目/基因/细胞和唯一性转录物分子数目/基因/细胞；(ii)测序数据的统计分析，例如主成分分析，用于预测转录物分子数目/基因/细胞的测定的置信区间等；(iii)序列比对能力，用于将基因序列数据与已知参考序列进行比对；(iv)样品条码、细胞条码和分子条码的解码/多路解编；以及(v)分子标记的自动化聚簇，以补偿扩增或测序误差。

[0243] 通常，包括在当前披露的仪器系统(如图76所说明的)中的计算机或处理器可以被进一步理解为逻辑设备，其可以读取来自介质511和/或网络端口505的指令，其可以任选地与具有固定介质512的服务器509连接。系统500(如图76所示的)可以包括CPU 501、磁盘驱动器503、任选的输入装置(如键盘515和/或鼠标516)以及任选的监测器507。可以通过从指定的通信介质传至在本地或远程位置的服务器来实现数据通信。通信介质可以包括发送和/或接收数据的任何工具。例如，通信介质可以是网络连接、无线连接或因特网连接。这样的连接可以提供经由万维网的通信。可以设想的是，与本披露有关的数据可以通过这样的网络或连接进行传输，以由接收方522接收和/或审阅，如图76所说明的。

[0244] 图77是一个框图，说明了可以与本披露的示例性实施例结合使用的计算机系统

100的第一示例性架构。如图77所描绘的，示例性计算机系统可以包括用于处理指令的处理器102。处理器的非限制性实例包括：英特尔XeonTM处理器、AMD OpteronTM处理器、三星32位RISC ARM 1176JZ(F)-S v1.0TM处理器、ARM Cortex-A8三星S5PC100TM处理器、ARM Cortex-A8苹果A4TM处理器、迈威尔(Marvell)PXA 930TM处理器或功能等效的处理器。多个执行线程可以用于并行处理。在一些实施例中，也可以使用多处理器或具有多个核心的处理器，无论是以单个计算机系统、成群地还是通过网络跨系统分布的，所述网络包括多个计算机、手机和/或个人数据助理装置。

[0245] 如图77所说明的，可以将高速缓存104连接至或并入处理器102中，以便为最近已经被或频繁地被处理器102使用的指令或数据提供高速存储器。通过处理器总线108将处理器102与北桥106连接。北桥106通过存储总线112与随机存取内存(RAM)110连接并且管理处理器102对RAM 110的访问。北桥106还通过芯片集总线116与南桥114连接。南桥114进而与外设总线118连接。外设总线可以是例如PCI、PCI-X、PCI Express或其他外设总线。北桥和南桥通常被称为处理器芯片集并且管理处理器、RAM与外设总线118上的外设部件之间的数据传输。在一些替代性架构中，可以将北桥的功能性并入处理器中而取代使用单独的北桥芯片。

[0246] 在一些实施例中，系统100可以包括附接至外设总线118的加速器卡122。加速器可以包括现场可编程门阵列(FPGA)或其他用于加速某些处理的硬件。例如，加速器可以用于自适应性数据重构或用于评价扩展集处理中所用的代数表达式。

[0247] 软件和数据被存储在外部存储器124中并且可以被加载到RAM110和/或缓存104中，用于为处理器所用。系统100包括用于管理系统资源的操作系统；操作系统的非限制性实例包括：Linux、Windows™、MACOS™、BlackBerry OS™、iOS™和其他功能等效的操作系统，以及在操作系统之上运行的用于根据本发明的示例性实施例管理数据存储和优化的应用软件。

[0248] 在这个实例中，系统100还包括与外设总线连接的网络接口卡(NIC)120和121，用于为外部存储器(如附网存储(NAS))和其他可以用于分布式并行处理的计算机系统提供网络接口。

[0249] 图78是一个图表，示出了网络200，其具有多个计算机系统202a和202b、多个手机和个人数据助理202c以及附网存储(NAS)204a和204b。在示例性实施例中，系统212a、212b和212c可以管理数据存储并针对存储在附网存储(NAS)214a和214b中的数据优化数据访问。数学模型可以用于数据并且使用跨计算机系统212a和212b、和手机以及个人数据助理系统212c的分布式并行处理进行评价。计算机系统212a和212b和手机以及个人数据助理系统212c还可以为存储在附网存储(NAS)214a和214b中的数据的自适应性数据重构提供并行处理。图78仅说明了一个实例，并且多种多样的其他计算机架构和系统也可以与本发明的不同实施例结合使用。例如，刀片式服务器可以用来提供并行处理。处理器刀片可以通过底板进行连接，以提供并行处理。存储器也可以通过单独的网络接口与底板连接或作为附网存储(NAS)。

[0250] 在一些示例性实施例中，处理器可以保持分开的存储空间并且通过网络接口、底板或其他连接器传输数据，用于通过其他处理器进行并行处理。在其他实施例中，一些或所有处理器可以使用共享的虚拟地址存储空间。

[0251] 图79是根据一个示例性实施例的多处理器计算机系统300的框图，所述系统使用共享的虚拟地址存储空间。该系统包括多个可以访问共享的存储子系统304的处理器302a-f。该系统在存储子系统304中并入了多个可程编硬件存储算法处理器(MAP)306a-f。MAP 306a-f各自可以包括存储器308a-f以及一个或多个现场可编程门阵列(FPGA)310a-f。MAP提供了可配置性功能单元并且具体算法或算法部分可以被提供给FPGAs 310a-f，用于与对应的处理器密切配合进行处理。例如，MAP可以用于相对于数据模型来评价代数表达式以及用于在示例性实施例中进行自适应型数据重构。在这个实例中，每个MAP都可被所有处理器全球访问，用于这些目的。在一种配置中，每个MAP都可以使用直接内存访问(DMA)来访问相关的存储器308a-f，从而允许它独立于对应的微处理器302a-f并且与其不同步地执行任务。在这种配置中，MAP可以将结果直接给至另一MAP，用于流水操作和并行执行算法。

[0252] 以上计算机架构和系统仅仅是实例，并且多种多样的其他计算机、手机和个人数据助理架构和系统可以与示例性实施例结合使用，包括使用一般处理器、协同处理器、FPGA和其他可编程逻辑装置、片上系统(SOC)、专用集成电路(ASIC)以及其他处理和逻辑元件的任何组合的系统。在一些实施例中，全部或部分的计算机系统可以在软件或硬件中实施。任何种类的数据存储介质都可以与示例性实施例结合使用，包括随机存取内存、硬盘驱动器、闪速存储器、磁带驱动器、磁盘阵列、附网存储(NAS)以及其他局部或分布式数据存储装置和系统。

[0253] 在示例性实施例中，可以使用在任何以上或其他计算机架构和系统上执行的软件模块来实施本披露的计算机子系统。在其他实施例中，可以部分地或完全地在以下项中实现该系统的功能：固件、可编程逻辑装置(如在图79中提及的现场可编程门阵列(FPGA))、片上系统(SOC)、专用集成电路(ASIC)或其他处理和逻辑元件。例如，可以通过使用硬件加速器卡(如图77中所说明的加速器卡122)通过硬件加速来实施集处理器和优化器。

[0254] 寡核苷酸(例如，分子条码)

[0255] 在此披露的方法和试剂盒可以包括一种或多种寡核苷酸或其用途。寡核苷酸可以附接至在此披露的固体支持物。寡核苷酸与固体支持物的附接可以通过该固体支持物和该寡核苷酸上的官能团进行。寡核苷酸可以被称为分子条码。寡核苷酸可以被称为标记(例如，分子标记、细胞标记)或标签(例如，样品标签)。

[0256] 寡核苷酸可以包括通用标记。通用标记对于样品中的所有寡核苷酸而言可以是相同的。通用标记对于一组寡核苷酸中的寡核苷酸而言可以是相同的。通用标记对于两组或更多组寡核苷酸而言可以是相同的。通用标记可以包括可与测序引物杂交的核酸序列。测序引物可以用于对包含通用标记的寡核苷酸测序。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联系的测序引物。通用标记可以包括可与PCR引物杂交的核酸序列。通用标记可以包括可与测序引物和PCR引物杂交的核酸序列。可与测序和/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用来引发寡核苷酸的转录的序列。通用标记可以包括可用来延伸寡核苷酸或寡核苷酸内的区域的序列。通用标记的长度可以是至少约1、2、3、4、5、10、15、20、25、30、35、40、45、50或更多个核苷酸。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多约1、2、3、4、5、10、15、20、25、30、35、40、45、50或更多个核苷酸。

[0257] 寡核苷酸可以包括细胞标记。细胞标记可以包括可为寡核苷酸所接触的细胞提供

信息(例如,确定哪个核酸来源于哪个细胞)的核酸序列。同一固体支持物上的至少60%、70%、80%、85%、90%、95%、97%、99%或100%的寡核苷酸可以包括相同的细胞标记。同一固体支持物上的至少60%的寡核苷酸可以包括相同的细胞标记。同一固体支持物上的至少95%的寡核苷酸可以包括相同的细胞标记。同一固体支持物上的所有寡核苷酸可以包括相同的细胞标记。第一固体支持物上的寡核苷酸的细胞标记可以不同于第二固体支持物上的寡核苷酸的细胞标记。

[0258] 细胞标记的长度可以是至少约1、2、3、4、5、10、15、20、25、30、35、40、45、50或更多个核苷酸。细胞标记的长度可以是至多约300、200、100、90、80、70、60、50、40、30、20、15、12、10、9、8、7、6、5、4或更少个核苷酸。细胞标记可以包括约5至约200个之间的核苷酸。细胞标记可以包括约10至约150个之间的核苷酸。细胞标记的长度可以包括约20至约125个之间的核苷酸。

[0259] 寡核苷酸可以包括分子标记。分子标记可以包括可为与该寡核苷酸杂交的特定核酸种类提供识别信息的核酸序列。与同一固体支持物缀合的寡核苷酸可以包括不同的分子标记。以此方式,分子标记可以区分与不同寡核苷酸杂交的靶核酸(例如,基因)的类型。分子标记的长度可以是至少约1、2、3、4、5、10、15、20、25、30、35、40、45、50或更多个核苷酸。分子标记的长度可以是至多约300、200、100、90、80、70、60、50、40、30、20、15、12、10、9、8、7、6、5、4或更少个核苷酸。

[0260] 寡核苷酸可以包括样品标记(例如,样品索引)。样品标记可以包括可提供有关靶核酸起源于哪里的信息的核酸序列。例如,在不同实验中所用的不同固体支持物上的样品标记可以是不同的。样品标记的长度可以是至少约1、2、3、4、5、10、15、20、25、30、35、40、45、50或更多个核苷酸。样品标记的长度可以是至多约300、200、100、90、80、70、60、50、40、30、20、15、12、10、9、8、7、6、5、4或更少个核苷酸。

[0261] 寡核苷酸可以包括通用标记、细胞标记、分子标记以及样品标记或其任何组合。在组合时,样品标记可以用于区分样品之间的靶核酸,细胞标记可以用于区分来自样品中的不同细胞的靶核酸,分子标记可以用于区分细胞中的不同靶核酸(例如,相同标靶核酸的不同拷贝),并且通用标记可以用于扩增靶核酸并对其测序。

[0262] 通用标记、分子标记、细胞标记、接头标记和/或样品标记可以包括核苷酸的随机序列。核苷酸的随机序列可以是计算机产生的。核苷酸的随机序列可以没有与其相关的模式。通用标记、分子标记、细胞标记、接头标记和/或样品标记可以包括核苷酸的非随机序列(例如,这些核苷酸包括一定模式)。通用标记、分子标记、细胞标记、接头标记和/或样品标记的序列可以是可商购的序列。通用标记、分子标记、细胞标记、接头标记和/或样品标记的序列可以是/包括随机物序列。随机物序列可以指由给定长度的该随机物的所有可能序列构成的寡核苷酸序列。可替代地或另外地,通用标记、分子标记、细胞标记、接头标记和/或样品标记可以包括核苷酸的预定序列。

[0263] 图1示出了本披露的包含通用标记、细胞标记和分子标记的示例性寡核苷酸。

[0264] 图3示出了示例性的偶联寡核苷酸的固体支持物,其包括与寡核苷酸(312)偶联的固体支持物(301)。寡核苷酸(312)包括化学基团(5'胺,302)、通用标记(303)、细胞标记(311)、分子标记(分子BC(Molecular BC),311)以及靶结合区(寡聚dT,310)。在这个示意图中,细胞标记(311)包括第一细胞标记(CL部分1,304)、第一接头(接头1,305)、第二细胞标

记(CL部分2,306)、第二接头(接头2,307)、第三细胞标记(CL部分3,308)。细胞标记(311)对于固体支持物上的每个寡核苷酸而言是共有的。两个或更多个珠的细胞标记(311)可以是不同的。两个或更多个珠的细胞标记(311)的区别可以在于细胞标记(例如,CL部分1(304)、CL部分2(306)、CL部分3(308))。两个或更多个珠的细胞标记(311)的区别可以在于第一细胞标记(304)、第二细胞标记(306)、第三细胞标记(308)或其组合。细胞标记(311)的第一和第二接头(303,305)对于两个或更多个偶联寡核苷酸的固体支持物而言可以是相同的。通用标记(303)对于两个或更多个偶联寡核苷酸的固体支持物而言可以是相同的。通用标记(303)对于同一固体支持物上的两个或更多个寡核苷酸而言可以是相同的。分子标记(311)对于固体支持物上的至少两个或更多个寡核苷酸而言可以是不同的。固体支持物可以包括100或更多个寡核苷酸。固体支持物可以包括1000或更多个寡核苷酸。固体支持物可以包括10000或更多个寡核苷酸。固体支持物可以包括100000或更多个寡核苷酸。

[0265] 除通用标记、细胞标记和分子标记之外,寡核苷酸可以包括靶结合区。靶结合区可以包括可与靶核酸(例如,有待分析的细胞核酸)结合的核酸序列。靶结合区可以是基因特异性序列。例如,靶结合区可以包括可附接(例如,杂交)至特定靶核酸的特定位置的核酸序列。靶结合区可以包括非特异性靶核酸序列。非特异性靶核酸序列可以指独立于靶核酸的特定序列可与多个靶核酸结合的序列。例如,靶结合区可以包括随机多聚体序列或寡聚dT序列(例如,一段可与mRNA上的聚腺苷酸化尾杂交的胸苷核苷酸)。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。靶结合区的长度可以是至少约5、10、15、20、25、30、35、40、45、50或更多个核苷酸。靶结合区的长度可以是至多约5、10、15、20、25、30、35、40、45、50或更多个核苷酸。

[0266] 寡核苷酸可以包括多个标记。例如,寡核苷酸可以包括至少约1、2、3、4、5、6、7或8或更多个通用标记。寡核苷酸可以包括至多约1、2、3、4、5、6、7或8或更多个通用标记。寡核苷酸可以包括至少约1、2、3、4、5、6、7或8或更多个细胞标记。寡核苷酸可以包括至多约1、2、3、4、5、6、7或8或更多个细胞标记。寡核苷酸可以包括至少约1、2、3、4、5、6、7或8或更多个分子标记。寡核苷酸可以包括至多约1、2、3、4、5、6、7或8或更多个分子标记。寡核苷酸可以包括至少约1、2、3、4、5、6、7或8或更多个样品标记。寡核苷酸可以包括至多约1、2、3、4、5、6、7或8或更多个样品标记。寡核苷酸可以包括至少约1、2、3、4、5、6、7或8或更多个靶结合区。寡核苷酸可以包括至多约1、2、3、4、5、6、7或8或更多个靶结合区。

[0267] 当寡核苷酸包括多于一种类型的标记(例如,多于一种细胞标记或多于一种分子标记)时,这些标记可以穿插着接头标记序列。接头标记序列的长度可以是至少约5、10、15、20、25、30、35、40、45、50或更多个核苷酸。接头标记序列的长度可以是至多约5、10、15、20、25、30、35、40、45、50或更多个核苷酸。在一些情况下,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寡核苷酸的合成,如图2A所图示的。

[0268] 缀合至固体支持物的寡核苷酸的数目可以是细胞中靶核酸的数目的1、2、3、4、5、6、7、8、9或10倍。在一些情况下,至少10%、20%、30%、40%、50%、60%、70%、80%、90%或100%的寡核苷酸被靶核酸结合。在一些情况下,至多10%、20%、30%、40%、50%、60%、70%、80%、90%或100%的寡核苷酸被靶核酸结合。在一些情况下,至少1、2、3、4、5、6、7、8、9、10、20、30、40、50、60、70、80、90或100或更多个不同靶核酸被固体支持物上的寡核苷酸捕

获。在一些情况下，至多1、2、3、4、5、6、7、8、9、10、20、30、40、50、60、70、80、90或100或更多个不同靶核酸被固体支持物上的寡核苷酸捕获。

[0269] 聚合物可以包括另外的固体支持物。例如，聚合物可以点缀着珠。珠在空间上可以定位在聚合物的不同区域。包含本披露的寡核苷酸的珠或支持物可以在空间上编址。珠或支持物可以包括对应于聚合物的空间地址的条码。例如，多个珠或支持物中的每个珠或支持物都可以包括条码，该条码对应于聚合物上的位置，例如阵列或多个微孔的特定微孔上的位置。可以将空间地址解码，以确定珠或支持物所定位的位置。例如，空间地址(如条码)可以通过使寡核苷酸与该条码杂交或通过对该条码测序来解码。可替代地，珠或支持物可以带有其他类型的条码，如图形特征、化学基团、颜色、荧光或其任何组合，用于解码空间地址的目的。

[0270] 在此披露的方法和试剂盒可以包括一组或多组分子条码。一个或多个分子条码可以包括样品索引区和标记区。一组分子条码中的两个或更多个分子条码可以包括相同的样品索引区和两个或更多个不同的标记区。两组或更多组分子条码中的两个或更多个分子条码可以包括两个或更多个不同的样品索引区。来自一组分子条码的两个或更多个分子条码可以包括不同的标记区。来自两组或更多组分子条码的两个或更多个分子条码可以包括相同的标记区。来自两组或更多组分子条码的分子条码的区别可以在于其样品索引区。基于其标记区，来自两组或更多组分子条码的分子条码可以是相似的。

[0271] 分子条码可以进一步包括靶特异性区、适配子区、通用PCR区、靶特异性区或其任何组合。分子条码可以包括通用PCR区和靶特异性区。分子条码可以包括一个或多个二级结构。分子条码可以包括发夹结构。分子条码可以包括靶特异性区和可切割茎。

[0272] 在此披露的方法和试剂盒可以包括一组或多组样品标签。一个或多个样品标签可以包括样品索引区。一个或多个样品标签可以包括样品索引区。一组样品标签中的两个或更多个样品标签可以包括相同的样品索引区。两组或更多组样品标签中的两个或更多个样品标签可以包括两个或更多个不同的样品索引区。

[0273] 样品标签可以进一步包括靶特异性区、适配子区、通用PCR区、靶特异性区或其任何组合。样品标签可以包括通用PCR区和靶特异性区。样品标签可以包括一个或多个二级结构。样品标签可以包括发夹结构。样品标签可以包括靶特异性区和可切割茎。

[0274] 在此披露的方法和试剂盒可以包括一组或多组分子鉴定物标记。一个或多个分子鉴定物标记可以包括标记区。一个或多个分子鉴定物标记可以包括标记区。一组分子鉴定物标记中的两个或更多个分子鉴定物标记可以包括两个或更多个不同的标记区。两组或更多组分子鉴定物标记中的两个或更多个分子鉴定物标记可以包括两个或更多个相同的标记区。分子鉴定物标记可以进一步包括靶特异性区、适配子区、通用PCR区、靶特异性区或其任何组合。分子鉴定物标记可以包括通用PCR区和靶特异性区。分子鉴定物标记可以包括一个或多个二级结构。分子鉴定物标记可以包括发夹结构。分子鉴定物标记可以包括靶特异性区和可切割茎。

[0275] 分子条码、样品标签或分子鉴定物标记可以包括至少约1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、30、40、50、60、70、80、90、100、200、300、400、500、600、700、800、900或1000个核苷酸或碱基对。在另一个实例中，样品标签或分子鉴定物标记包括至少约1500、2,000；2500、3,000；3500、4,000；4500、5,000；5500、6,000；6500、7,000；7500、

8,000;8500、9,000;9500或10,000个核苷酸或碱基对。

[0276] 分子条码、样品标签或分子鉴定物标记可以是多聚体，例如随机多聚体。多聚体序列可以是例如非随机或随机二聚体、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。标签可以由一组单核苷酸随机产生。标签可以通过随机并入单核苷酸组装而来。

[0277] 分子条码、样品标签或分子鉴定物标记也可以进行非随机组装，以产生非随机产生的不同标签的文库，但是该文库包括足够实践所述方法的数目的不同标签。

[0278] 在一些实施例中，分子条码、样品标签或分子鉴定物标记可以包括靶核酸的消减。消减可以是例如靶核酸的一端或两端的酶消化。消减可以与添加的分子条码、样品标签(例如，样品索引区、样品标记)、细胞标记以及分子鉴定物标记(例如，分子标记)的添加结合使用。添加的标签和该消减的组合可以含有与特定起始分子相关的信息。通过向分子条码、样品标签或分子鉴定物标记中添加随机消减，使得当检测允许确定添加的寡核苷酸和该随机消减两者时，所添加的标签的较小多样性对于对靶核酸的数目计数而言可能是必要的。

[0279] 分子条码、样品标签或分子鉴定物标记可以包括靶特异性区。靶特异性区可以包括与分子互补的序列。在一些情况下，分子是mRNA分子并且靶特异性区包括与该mRNA分子的聚A尾互补的寡聚dT序列。靶特异性区也可以充当DNA和/或RNA合成的引物。例如，靶特异性区的寡聚dT序列可以充当mRNA分子的cDNA拷贝的第一链合成的引物。可替代地，靶特异性区包括与分子的任何部分互补的序列。在其他情况下，靶特异性区包括可与分子杂交或连接的随机序列。靶特异性区可以允许样品标签或分子鉴定物标记附接至所述分子。样品标签或分子鉴定物标记的附接可以通过在此披露的任何方法(例如，杂交、连接)进行。在一些情况下，靶特异性区包括被一种或多种限制酶识别的序列。靶特异性区可以包括至少约1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、30、40、50、60、70、80、90、100、200、300、400、500、600、700、800、900或1000个核苷酸或碱基对。在另一个实例中，靶特异性区包括至少约1500、2000、2500、3000、3500、4000、4500、5000、5500、6000、6500、7000、7500、8000、8500、9000、9500或10000个核苷酸或碱基对。优选地，靶特异性区包括至少约5-10、10-15、10-20、10-30、15-30或20-30个核苷酸或碱基对。

[0280] 在一些情况下，靶特异性区对特定基因或基因产物具有特异性。例如，靶特异性区包括与p53基因或基因产物的区域互补的序列。因此，样品标签和分子鉴定物标记只能附接至包含p53特异性序列的分子。可替代地，靶特异性区对多个不同基因或基因产物具有特异性。例如，靶特异性区包括寡聚dT序列。因此，样品标签和分子鉴定物标记可以附接至包含聚A序列的任何分子。在另一个实例中，靶特异性区包括与多个不同基因或基因产物互补的随机序列。因此，样品标签或分子鉴定物标记可以附接至具有与靶特异性区互补的序列的任何分子。在其他情况下，靶特异性区包括限制性位点突出端(例如，EcoRI粘端突出端)。样品标签或分子鉴定物标记可以连接至包含与限制性位点突出端互补的序列的任何分子。

[0281] 在一些情况下，靶特异性区对特定微小RNA或微小RNA产物具有特异性。例如，靶特异性区包括与特定微小RNA或微小RNA产物的区域互补的序列。靶特异性区包括与特定面板的微小RNA或特定面板的微小RNA产物的区域互补的序列。因此，样品标签和分子鉴定物标记只能附接至包含微小RNA特异性序列的分子。可替代地，靶特异性区对多个不同微小RNA或微小RNA产物具有特异性。例如，靶特异性区包括与包含在两个或更多个微小RNA中的区

域互补的序列，如含有共有序列的微小RNA面板。因此，样品标签和分子鉴定物标记可以附接至包含共有微小RNA序列的任何分子。在另一个实例中，靶特异性区包括与多个不同微小RNA或微小RNA产物互补的随机序列。因此，样品标签或分子鉴定物标记可以附接至具有与靶特异性区互补的序列的任何微小RNA分子。在其他情况下，靶特异性区包括限制性位点突出端(例如，EcoRI粘端突出端)。样品标签或分子鉴定物标记可以连接至包含与限制性位点突出端互补的序列的任何微小RNA分子。

[0282] 在此披露的分子条码或分子鉴定物标记通常包括标记区。标记区可以用于唯一性地确定目标种类的存在，由此用鉴定物对每个种类进行标记，所述鉴定物可以用于区分两个在其他方面相同或几乎相同的靶标。该多个样品标签和分子鉴定物标记的标记区可以包括以下项的集合：不同的半导体纳米晶体，金属化合物，肽，寡核苷酸，抗体，小分子，同位素，与其相关或包埋在其中的具有不同形状、颜色、条码或衍射图的颗粒或结构，数字串，蛋白质或核酸的随机片段，不同的同位素或其任何组合。标记区可以包括简并序列。标记区可以包括至少约1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、30、40、50、60、70、80、90、100、200、300、400、500、600、700、800、900或1000个核苷酸或碱基对。在另一个实例中，标记区包括至少约1500；2,000；2500、3,000；3500、4,000；4500、5,000；5500、6,000；6500、7,000；7500、8,000；8500、9,000；9500或10,000个核苷酸或碱基对。优选地，标记区包括至少约10-30、15-40或20-50个核苷酸或碱基对。

[0283] 在一些情况下，分子条码、样品标签或分子鉴定物标记包括通用引物结合位点。通用引物结合位点允许将通用引物附接至经标记的分子和/或经标记的扩增子。通用引物在本领域是熟知的，并且包括但不限于-47F(M13F)、αMF、AOX3’、AOX5’、BGH_r、CMV_-30、CMV_-50、CVM_f、LACrmt、λgt10F、λ(lambda)gt 10R、λgt11F、λgt11R、M13rev、M13正向(-20)、M13反向、male、p10SEQP_pQE、pA_-120、pet_4、pGAP正向、pGL_RVpr3、pGLpr2_R、pKLAC1_4、pQE_FS、pQE_RS、puc_U1、puc_U2、revers_A、seq_IRES_tam、seq_IRES_zpet、seq_ori、seq_PCR、seq_pIRES-、seq_pIRES+、seq_pSecTag、seq_pSecTag+、seq_retro+PSI、SP6、T3-prom、T7-prom以及T7-term_Inv。通用引物与通用引物结合位点的附接可以用于对经标记的分子和/或经标记的扩增子进行扩增、检测和/或测序。通用引物结合位点可以包括至少约1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、30、40、50、60、70、80、90、100、200、300、400、500、600、700、800、900或1000个核苷酸或碱基对。在另一个实例中，通用引物结合位点包括至少约1500；2,000；2500、3,000；3500、4,000；4500、5,000；5500、6,000；6500、7,000；7500、8,000；8500、9,000；9500或10,000个核苷酸或碱基对。优选地，通用引物结合位点包括10-30个核苷酸或碱基对。

[0284] 分子条码、样品标签或分子鉴定物标记可以包括适配子区。适配子区可以允许一个或多个探针的杂交。适配子区可以允许一个或多个HCR探针的杂交。

[0285] 分子条码、样品标签或分子鉴定物标记可以包括一个或多个可检测标记。

[0286] 分子条码、样品标签和分子鉴定物标记可以充当杂交链式反应(HCR)的引发物。样品标签或分子鉴定物标记的适配子区可以充当HCR的引发物。通用引物结合位点可以充当HCR的引发物。

[0287] 在一些情况下，分子条码、样品标签或分子鉴定物标记是单链的。在其他情况下，分子条码、样品标签或分子鉴定物标记是双链的。分子条码、样品标签或分子鉴定物标记可

以是直链的。可替代地，分子条码、样品标签或分子鉴定物标记包括二级结构。如在此使用的，“二级结构”包括三级、四级等...结构。在一些情况下，二级结构是发夹、茎环结构、内环、凸环、分支结构或假结体、多茎环结构、四叶型结构或任何三维结构。在一些情况下，二级结构是发夹。发夹可以包括突出端序列。发夹的突出端序列可以充当聚合酶链式反应和/或逆转录反应的引物。突出端序列包括与样品标签或分子鉴定物标记所附接的分子互补的序列，并且突出端序列与该分子杂交。突出端序列可以与该分子连接并且充当聚合酶链式反应和/或逆转录反应的模板。在一些实施例中，分子条码、样品标签或分子鉴定物标记包括核酸和/或合成的核酸和/或修饰的核酸。

[0288] 在一些情况下，该多个分子条码、样品标签(例如，样品索引区、样品标记)、细胞标记以及分子鉴定物标记(例如，分子标记)包括至少约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或100个不同的分子条码、样品标签(例如，样品索引区、样品标记)、细胞标记以及分子鉴定物标记(例如，分子标记)。在其他情况下，该多个分子条码、样品标签(例如，样品索引区、样品标记)、细胞标记以及分子鉴定物标记(例如，分子标记)包括至少约200；300；400；500；600；700；800；900；1,000；2,000；3,000；4,000；5,000；6,000；7,000；8,000；9,000；或10000个不同的分子条码、样品标签(例如，样品索引区、样品标记)、细胞标记以及分子鉴定物标记(例如，分子标记)。可替代地；该多个分子条码、样品标签(例如，样品索引区、样品标记)、细胞标记以及分子鉴定物标记(例如，分子标记)包括至少约20,000；30,000；40,000；50,000；60,000；70,000；80,000；90,000；或100,000个不同的分子条码、样品标签(例如，样品索引区、样品标记)、细胞标记以及分子鉴定物标记(例如，分子标记)。

[0289] 该多个分子条码、样品标签(例如，样品索引区、样品标记)、细胞标记以及分子鉴定物标记(例如，分子标记)中的分子条码、样品标签(例如，样品索引区、样品标记)、细胞标记以及分子鉴定物标记(例如，分子标记)的数目通常超过有待被标记的分子的数目。在一些情况下，该多个分子条码、样品标签(例如，样品索引区、样品标记)、细胞标记以及分子鉴定物标记(例如，分子标记)中的分子条码、样品标签(例如，样品索引区、样品标记)、细胞标记以及分子鉴定物标记(例如，分子标记)的数目是比有待被标记的分子的数目多至少约2、3、4、5、6、7、8、9、10、15、20、30、40、50、60、70、80、90或100倍。

[0290] 该多个分子条码、样品标签(例如，样品索引区、样品标记)、细胞标记以及分子鉴定物标记(例如，分子标记)中的不同分子条码、样品标签(例如，样品索引区、样品标记)、细胞标记以及分子鉴定物标记(例如，分子标记)的数目通常超过有待被标记的不同分子的数目。在一些情况下，该多个分子条码、样品标签(例如，样品索引区、样品标记)、细胞标记以及分子鉴定物标记(例如，分子标记)中的不同分子条码、样品标签(例如，样品索引区、样品标记)、细胞标记以及分子鉴定物标记(例如，分子标记)的数目是比有待被标记的不同分子的数目多至少约1、1.5、2、2.5、3、3.5、4、4.5、5、6、7、8、9、10、15、20、30、40、50、60、70、80、90或100倍。

[0291] 在一些情况下，进行分子的随机标记包括多个分子条码、样品标签(例如，样品索引区、样品标记)、细胞标记以及分子鉴定物标记(例如，分子标记)，其中该多个分子条码、样品标签(例如，样品索引区、样品标记)、细胞标记以及分子鉴定物标记(例如，分子标记)中的不同分子条码、样品标签(例如，样品索引区、样品标记)、细胞标记以及分子鉴定物标

记(例如,分子标记)的浓度是相同的。在这样的情况下,该多个分子条码、样品标签(例如,样品索引区、样品标记)、细胞标记以及分子鉴定物标记(例如,分子标记)包括相等数目的每种不同的分子条码、样品标签或分子鉴定物标记。

[0292] 在一些情况下,进行分子的随机标记包括多个分子条码、样品标签(例如,样品索引区、样品标记)、细胞标记以及分子鉴定物标记(例如,分子标记),其中该多个分子条码、样品标签(例如,样品索引区、样品标记)、细胞标记以及分子鉴定物标记(例如,分子标记)中的不同分子条码、样品标签(例如,样品索引区、样品标记)、细胞标记以及分子鉴定物标记(例如,分子标记)的浓度是不同的。在这样的情况下,该多个分子条码、样品标签(例如,样品索引区、样品标记)、细胞标记以及分子鉴定物标记(例如,分子标记)包括不同数目的每种不同的分子条码、样品标签或分子鉴定物标记。

[0293] 在一些情况下,在该多个分子条码、样品标签(例如,样品索引区、样品标记)、细胞标记以及分子鉴定物标记(例如,分子标记)中,一些分子条码、样品标签(例如,样品索引区、样品标记)、细胞标记以及分子鉴定物标记(例如,分子标记)以高于其他分子条码、样品标签(例如,样品索引区、样品标记)、细胞标记以及分子鉴定物标记(例如,分子标记)的浓度存在。在一些情况下,用不同浓度的分子条码、样品标签(例如,样品索引区、样品标记)、细胞标记以及分子鉴定物标记(例如,分子标记)进行随机标记在没有增加所用的不同标记的数目的情况下扩展样品测量动态范围。例如,考虑用全部处于相等浓度下的10种不同的分子条码、样品标签(例如,样品索引区、样品标记)、细胞标记以及分子鉴定物标记(例如,分子标记)随机标记3种核酸样品分子。我们期望观察到3种不同的标记。现在考虑30种核酸分子而非3种核酸分子,我们期望观察到所有10种标记。相比之下,如果我们仍然使用10种不同的随机标记并且将标记的相对比变为1:2:3:4…10,则在用3种核酸分子的情况下,我们期望观察到1-3种之间的标记,但是在用30种分子的情况下,我们将仅期望观察到大约5种标记,从而扩展了用相同数目的随机标记进行测量的范围。

[0294] 该多个分子条码、样品标签(例如,样品索引区、样品标记)、细胞标记以及分子鉴定物标记(例如,分子标记)中的不同分子条码、样品标签(例如,样品索引区、样品标记)、细胞标记以及分子鉴定物标记(例如,分子标记)的相对比可以是1:X,其中X是至少约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或100。可替代地,该多个分子条码、样品标签(例如,样品索引区、样品标记)、细胞标记以及分子鉴定物标记(例如,分子标记)中的“n”种不同分子条码、样品标签(例如,样品索引区、样品标记)、细胞标记以及分子鉴定物标记(例如,分子标记)的相对比是1:A:B:C:…Zn,其中A、B、C…Zn是至少约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或100。

[0295] 在一些情况下,该多个分子条码、样品标签(例如,样品索引区、样品标记)、细胞标记以及分子鉴定物标记(例如,分子标记)中的两种或更多种不同分子条码、样品标签(例如,样品索引区、样品标记)、细胞标记以及分子鉴定物标记(例如,分子标记)的浓度是相同的。对于“n”种不同的分子条码、样品标签(例如,样品索引区、样品标记)、细胞标记以及分子鉴定物标记(例如,分子标记)而言,至少2、3、4…n种不同的分子条码、样品标签(例如,样品索引区、样品标记)、细胞标记以及分子鉴定物标记(例如,分子标记)的浓度是相同的。可替代地,该多个分子条码、样品标签(例如,样品索引区、样品标记)、细胞标记以及分子鉴定

物标记(例如,分子标记)中的两种或更多种不同分子条码、样品标签(例如,样品索引区、样品标记)、细胞标记以及分子鉴定物标记(例如,分子标记)的浓度是不同的。对于“n”种不同的分子条码、样品标签(例如,样品索引区、样品标记)、细胞标记以及分子鉴定物标记(例如,分子标记)而言,至少2、3、4…n种不同的分子条码、样品标签(例如,样品索引区、样品标记)、细胞标记以及分子鉴定物标记(例如,分子标记)的浓度是不同的。在一些情况下,对于“n”种不同的分子条码、样品标签(例如,样品索引区、样品标记)、细胞标记以及分子鉴定物标记(例如,分子标记)而言,至少2、3、4…n种不同的分子条码、样品标签(例如,样品索引区、样品标记)、细胞标记以及分子鉴定物标记(例如,分子标记)的浓度差异是至少约0.1、0.2、0.3、0.4、0.5、0.6、0.7、0.8、0.9、1、1.25、1.5、1.75、2、2.25、2.5、2.75、3、4、5、6、7、8、9、10、15、20、30、40、50、60、70、80、90、100、200、300、400、500、600、700、800、900或1000倍。

[0296] 在一些情况下,在该多个分子条码、样品标签(例如,样品索引区、样品标记)、细胞标记以及分子鉴定物标记(例如,分子标记)中,至少约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%的不同分子条码、样品标签(例如,样品索引区、样品标记)、细胞标记以及分子鉴定物标记(例如,分子标记)具有相同的浓度。可替代地,在该多个分子条码、样品标签(例如,样品索引区、样品标记)、细胞标记以及分子鉴定物标记(例如,分子标记)中,至少约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%的不同分子条码、样品标签(例如,样品索引区、样品标记)、细胞标记以及分子鉴定物标记(例如,分子标记)具有不同同的浓度。

[0297] 如图65所示,分子条码(1004)可以是分开合成的。分子条码(1004)可以包括通用PCR区(1001)、一个或多个鉴定物区(1002)和靶特异性区。该一个或多个鉴定物区可以包括样品索引区、标记区或其组合。该一个或多个鉴定物区可以是相邻的。该一个或多个鉴定物区可以是不相邻的。可以将单独分子条码合并,以产生包含多个不同鉴定物区的多个分子条码(1005)。样品标签可以按如图65所描绘的类似方式合成,其中该一个或多个鉴定物区包括样品索引区。分子鉴定物标记可以按如图65所描绘的类似方式合成,其中该一个或多个鉴定物区包括标记区。

[0298] 可以将靶特异性区连接至鉴定物区,以产生包含靶特异性区的分子条码。可以将5’和3’外切核酸酶添加至反应中,以除去未连接产物。分子条码可以包括通用引物结合位点、标记区和靶特异性区并且可以对5’和3’外切核酸酶具有抗性。如在此使用的,术语“通用引物结合位点”和“通用PCR区”可以可互换地使用并且是指可用于引发扩增反应的序列。可以将3’磷酸基团从连接的鉴定物区除去,以产生没有3’磷酸基团的分子条码。可以通过酶法除去3’磷酸基团。例如,T4多核苷酸激酶可以用于除去3’磷酸基团。

[0299] 合成分子条码的另一种方法描绘于图66A中。如图66A所示,可以通过连接两个或更多个寡核苷酸片段(1121和1127)来合成分子条码(1128)。一个寡核苷酸片段(1121)可以包括通用引物结合位点(1122)、鉴定物区(1123)和第一夹板区(1123)。另一个寡核苷酸片段(1128)可以包括第二夹板区(1125)和靶特异性区(1126)。连接酶(例如,T4DNA连接酶)可以用于连接这两个寡核苷酸片段(1121和1127),以产生分子条码(1128)。第一夹板区(1124)和第二夹板区(1125)的双链连接可以产生具有桥接夹板区(1129)的分子条码

(1128)。

[0300] 通过连接两个寡核苷酸片段合成分子条码的替代方法描绘于图66B中。如图66B所示,通过连接两个寡核苷酸片段(1150和1159)合成分子条码(1158)。一个寡核苷酸片段(1150)可以包括通用引物结合位点(1151)、一个或多个鉴定物区(1152)和连接序列(1153)。另一个寡核苷酸片段(1159)可以包括与第一寡核苷酸片段(1150)的连接序列(1153)互补的连接序列(1154)、靶特异性区(1155)的互补体和标记(1156)。寡核苷酸片段(1159)也可以包括防止寡核苷酸片段延伸的3'磷酸。如图66B的步骤1所示,两个寡核苷酸片段的连接序列(1153和1154)可以退火并且聚合酶可以用于延伸第一寡核苷酸片段(1150)的3'端,以产生分子条码(1158)。分子条码(1158)可以包括通用引物结合位点(1151)、一个或多个鉴定物区(1152)、连接序列(1153)以及靶特异性序列(1157)。分子条码(1158)的靶特异性序列(1157)可以是第二寡核苷酸片段(1159)的靶特异性区(1155)的互补体的互补体。可以将包含标记(1156)的寡核苷酸片段从分子条码(1158)中除去。例如,标记(1156)可以包括生物素并且可以经由链霉亲和素捕获来除去包含生物素标记(1156)的寡核苷酸片段(1159)。在另一个实例中,标记(1156)可以包括5'磷酸并且可以经由外切核酸酶(例如,λ外切核酸酶)来除去包含5'磷酸(1156)的寡核苷酸片段(1159)。

[0301] 如图66C所描绘,包含通用引物结合位点(1171)、一个或多个鉴定物区(1172)、第一连接序列(1173)的第一寡核苷酸片段(1170)与包含第二连接序列(1174)和靶序列的RNA互补体(1175)的第二寡核苷酸片段(1176)退火。步骤1可以包括使第一和第二连接序列(1173和1174)退火,随后逆转录靶序列的RNA互补体(1175),以产生包含通用引物结合位点(1171)、一个或多个鉴定物区(1172)、第一连接序列(1173)和靶特异性区(1178)的分子条码(1177)。可以通过RNA酶处理选择性降解包含靶序列的RNA互补体的寡核苷酸片段。

[0302] 可以优化分子条码、样品标签(例如,样品索引区、样品标记)、细胞标记以及分子鉴定物标记(例如,分子标记)的序列,以最小化分子条码、样品标签(例如,样品索引区、样品标记)、细胞标记以及分子鉴定物标记(例如,分子标记)的二聚化。可以扩增分子条码、样品标签或分子鉴定物标记二聚体并且导致扩增子的形成,所述扩增子在其每端上包含两个通用引物结合位点以及靶特异性区和唯一性鉴定物区。因为分子条码、样品标签(例如,样品索引区、样品标记)、细胞标记以及分子鉴定物标记(例如,分子标记)的浓度远大于DNA模板的数目,这些分子条码、样品标签或分子鉴定物标记二聚体在扩增反应中可以胜过经标记的DNA分子。未扩增的DNA导致假阴性,而扩增的分子条码、样品标签或分子鉴定物标记二聚体导致较高的假阳性。因此,可以优化分子条码、样品标签(例如,样品索引区、样品标记)、细胞标记以及分子鉴定物标记(例如,分子标记),以最小化分子条码、样品标签或分子鉴定物标记二聚体形成。可替代地,丢弃二聚化的分子条码、样品标签(例如,样品索引区、样品标记)、细胞标记以及分子鉴定物标记(例如,分子标记),由此消除分子条码、样品标签或分子鉴定物标记二聚体形成。

[0303] 可替代地,可以通过向分子条码、样品标签或分子鉴定物标记序列中掺入一种或多种修饰来消除或减少分子条码、样品标签或分子鉴定物标记二聚体形成。包含通用引物结合位点、唯一性鉴定物区和靶特异性区(包含尿嘧啶和3'磷酸基团)的分子条码、样品标签或分子鉴定物标记与靶核酸退火。靶核酸可以是被限制性内切核酸酶消化的片段。限制性核酸内切酶可以对识别位点进行识别。PCR扩增可以包括一个或多个正向引物以及一个

或多个反向引物。PCR扩增可以包括巢式PCR，其使用对分子条码、样品标签或分子鉴定物标记的通用引物结合位点具有特异性的正向引物和对分子条码、样品标签或分子鉴定物标记的靶特异性区具有特异性的正向引物以及对靶核酸具有特异性的反向引物。可以使用Pfu DNA聚合酶扩增靶核酸，该聚合酶无法扩增包含一个或多个尿嘧啶的模板。因此，任何二聚化的分子条码、样品标签(例如，样品索引区、样品标记)、细胞标记以及分子鉴定物标记(例如，分子标记)都无法通过Pfu DNA聚合酶扩增。

[0304] 合成寡核苷酸的方法(例如，分子条码)

[0305] 可以合成寡核苷酸。可以通过例如将寡核苷酸上的5'氨基偶联(例如，通过1-乙基-3-(3-二甲基氨基丙基)碳二亚胺)至官能化的固体支持物的羧基来合成寡核苷酸。

[0306] 可以通过多次清洗从该反应混合物中去除未偶联的寡核苷酸。这些固体支持物可以分成多个孔(例如，96个孔)。每个固体支持物可以分成不同的孔。可以使用分离/聚池合成方法进行寡核苷酸合成。该分离/聚池方法可以利用一种包含反应性部分(例如，待合成的寡核苷酸)的固体支持物的聚池。可以将该聚池分成固体支持物的单独聚池。可以使每个聚池经受第一反应，该第一反应可导致对每个聚池中固体支持物的不同修饰(例如，将不同的核酸序列添加至寡核苷酸)。在该反应之后，这些固体支持物的聚池可以被合并、混合、及再次分离。可以使每个分离聚池经受第二反应或随机化，这对于每个聚池再次是不同的。该过程可以持续进行直到目标化合物的文库形成。

[0307] 使用分离/聚池合成，添加至寡核苷酸的核酸序列可以通过引物延伸结合(例如，克列诺(Klenow)延伸)。有待添加至寡核苷酸的核酸序列可以称为引物片段。对于每个单独聚池的每个引物片段可以包含不同的序列(例如，在细胞标记、分子标记、样品标记、或它们的任何组合中)。引物片段可以包含可与寡核苷酸的接头标记序列杂交的序列(例如，偶联至固体支持物的寡核苷酸)。引物片段可以进一步包括第二细胞标记和第二接头标记序列。可以将引物延伸用于将第二细胞标记序列和第二接头标记序列引入到偶联至固体支持物的寡核苷酸上(参见图2B)。在引物延伸合并新序列之后，可以将这些固体支持物组合。可以将组合的固体支持物加热至使酶变性。可以将组合的固体支持物加热至破坏杂交。可以将组合的固体支持物再次分成多个孔。可以重复该过程以将另外的序列添加至缀合固体支持物的寡核苷酸。

[0308] 该分离/聚池过程可以导致产生至少约1000、10000、100000、500000、或1000000或更多个不同的寡核苷酸。该过程可以导致产生至多约1000、10000、100000、500000、或1000000或更多个不同的寡核苷酸。

[0309] 分离聚池合成可以包括化学合成。可在单个反应中在固体支持物上使用DMT化学，然后聚池到合成反应中来合成不同的寡核苷酸。可以将该分离/聚池过程重复1、2、3、4、5、6、7、8、9、10或更多次。可以将该分离/聚池过程重复11、12、13、14、15、16、17、18、19、20、25、30、35、40、45、50、55、60、65、70、75、80、85、90、95、100或更多次。可以将该分离/聚池过程重复2次或更多次。可以将该分离/聚池过程重复3次或更多次。可以将该分离/聚池过程重复5次或更多次。可以将该分离/聚池过程重复10次或更多次。

[0310] 在此进一步披露的是产生一个或多个组的经标记的珠(例如，缀合寡核苷酸的珠)的方法。产生一个或多个组的经标记的珠的方法可以包括将一个或多个核酸附接至一个或多个珠，由此产生一个或多个组的经标记的珠。该一个或多个核酸可以包括一个或多个分

子条码。该一个或多个核酸可以包括一个或多个样品标签(例如,样品标记,样品索引区)。该一个或多个核酸可以包括一个或多个细胞标记。该一个或多个核酸可以包括一个或多个分子鉴定物标记(例如,分子标记)。该一个或多个核酸可以包括a)引物区;b)样品索引区;以及c)接头区或适配子区。该一个或多个核酸可以包括a)引物区;b)标记区(例如,分子标记);以及c)接头区或适配子区。该一个或多个核酸可以包括a)样品索引区(例如,样品标签);以及b)标记区(例如,分子标记)。该一个或多个核酸可以包括a)样品索引区;以及b)细胞标记。该一个或多个核酸可以包括a)细胞标记;以及b)分子标记。该一个或多个核酸可以包括a)样品索引区;b)细胞标记;以及c)分子标记。该一个或多个核酸可以进一步包括引物区。该一个或多个核酸可以进一步包括靶特异性区。该一个或多个核酸可以进一步包括接头区。该一个或多个核酸可以进一步包括适配子区。该一个或多个核酸可以进一步包括样品索引区。该一个或多个核酸可以进一步包括标记区。

[0311] 可替代地,该方法包括:a)使多个第一核酸沉积进多个孔中,其中该多个孔的两个或更多个不同的孔可以包括该多个核酸的两个或更多个不同核酸;b)使该多个孔的一个或多个孔与一个或较少的珠接触,以产生多个单标记珠,其中多个第一经标记的珠的单标记珠包括附接至该多个第一核酸的核酸的珠;c)从该多个孔汇聚该多个第一经标记的珠以产生第一经标记的珠的聚池;d)将第一经标记的珠的该聚池分配至随后的多个孔,其中该随后多个孔中的两个或更多个孔包括多个随后的核酸中的两个或更多个不同核酸;以及e)将该多个随后的核酸的一个或多个核酸附接至一个或多个第一经标记的珠,以产生多个经唯一性标记的珠。

[0312] 文库

[0313] 在此披露的是产生分子文库的方法。该方法可以包括:(a)随机地对来自两个或更多个样品中的两种或更多种分子进行标记,以产生经标记的分子,其中这些经标记的分子包括(i)基于或衍生自这两种或更多种分子的分子区,(ii)用于区分来自两个或更多个样品的两种或更多种分子的样品索引区;以及(iii)用于区分来自单一样品的两种或更多种分子的标记区。随机进行标记可以包括使用一个或多个组的分子条码。随机进行标记可以包括使用一个或多个组的样品标签。随机进行标记可以包括使用一个或多个组的分子鉴定物标记。

[0314] 随机地标记两种或更多种分子可以包括使两个或更多个样品与多个样品标签和多个分子特异性标记接触,以产生多个经标记的核酸。该接触可以是随机的。该方法可以进一步包括对一个或多个经标记的分子进行扩增,从而产生文库的经标记的分子的富集群体。该方法可以进一步包括对来自两个或更多个样品的两种或更多种分子进行一个或多个测定。该方法可以进一步包括进行一个或多个拉下实验(pull-down assay)。

[0315] 产生经标记的核酸文库的方法可以进一步包括将一个或多个对照添加至两个或更多个样品。可以随机地标记一个或多个对照以产生经标记的对照。该一个或多个对照可以用于测量产生经标记的分子的效率。

[0316] 可以将在此披露的文库用于多种应用中。例如,可以将该文库用于测序应用。可以将该文库储存和多次使用,以生成用于分析的样品。一些应用包括,例如,基因分型多态性、研究RNA加工、以及选择克隆代表进行测序。

[0317] 样品制备和应用

[0318] 可以将在此披露的寡核苷酸(例如,分子条码、样品标签、分子标记、细胞标记)用于多种方法中。这些寡核苷酸可以在用于核酸分析的方法中。核酸分析可以包括,但不限于,基因分型、基因表达、拷贝数变异、和分子计数。

[0319] 本披露提供了用于多重核酸分析的方法。该方法可以包括(a)使来自细胞的一个或多个寡核苷酸与附连到支持物上的一个或多个寡核苷酸接触,其中附连到该支持物上的一个或多个寡核苷酸包括(i)细胞标记区,该区域包括通过非随机序列连接的两个或更多个随机物序列;以及(ii)分子标记区。以及(b)对来自该细胞的一个或多个寡核苷酸进行一个或多个测定。

[0320] 进一步在此披露的是产生单细胞核酸文库的方法。该方法可以包括(a)使来自细胞的一个或多个寡核苷酸与附连到支持物上的一个或多个寡核苷酸接触,其中附连到该支持物上的一个或多个寡核苷酸包括(i)细胞标记区,该区域包括通过非随机序列连接的两个或更多个随机物序列;以及(ii)分子标记区。以及(b)对来自该细胞的一个或多个寡核苷酸进行一个或多个测定。

[0321] 在一些情况下,该方法包括将一个或多个细胞添加至微孔阵列上。可以通过计数来确定有待添加的细胞数目。可以使用缓冲液(例如,磷酸盐缓冲液、HEPES、Tris)洗去过量或未结合的细胞。可以被微孔阵列的这些孔捕获的细胞的数目可与细胞的大小相关。例如,取决于微孔的设计,较大的细胞可以比较小的细胞更容易被捕获,如图6所示。可以将不同微孔(例如,不同大小)用于捕获不同细胞类型。

[0322] 此处描述的方法允许添加序列,这些序列可以是用于测序或其他分子分析的核酸。这些方法可以允许检测核酸变体、突变体、多态性、倒位、缺失、回复突变和其他在RNA或DNA分子群体中发现的定性事件。例如,这些方法可以允许鉴定靶标频率(例如,基因表达或等位基因的分布)。例如,这些方法还允许鉴定突变或在基因组或转录组中的SNP,例如来自患病的或非患病的受试者。这些方法还允许确定来自受试者的生物试样中污染与传染的存在与否,如外来生物或病毒,如细菌或真菌。

[0323] 可以通过任何方法将细胞添加至微孔中。在一些实施例中,将细胞作为稀释的细胞样品添加至微孔中。在一些实施例中,将细胞添加至微孔中并且允许通过重力沉降在微孔中。在一些实施例中,将细胞添加至微孔中并且使用离心将这些细胞沉降在微孔中。在一些实施例中,通过将一个或多个细胞注入一个或多个微孔中将细胞添加至微孔中。例如,通过将单细胞注入微孔中而将该单细胞添加至微孔中。细胞的注入可以通过使用任何装置或方法,例如通过使用显微操纵器。在一些实施例中,可以使用磁体将细胞添加至微孔中。例如,细胞可以在它们的表面上涂布有磁性颗粒(例如磁性微粒或磁性纳米颗粒)并且使用磁体或磁场添加至微孔中。

[0324] 可以将包含细胞的微孔阵列与缀合寡核苷酸的固体支持物(例如,珠)接触。可以去除未捕获的缀合寡核苷酸的固体支持物(例如,用缓冲液洗去)。图5描绘了具有捕获的固体支持物的微孔阵列。微孔可以包括至少一个固体支持物。微孔可以包括至少两个固体支持物。微孔可以包括至多一个固体支持物。微孔可以包括至多两个固体支持物。微孔可以包括至少约1、2、3、4、5、6、7、8、9、或10或更多个固体支持物。微孔可以包括至多约1、2、3、4、5、6、7、8、9、或10或更多个固体支持物。微孔阵列的一些微孔可以包括1个固体支持物并且微孔阵列的一些微孔可以包括两个或更多个固体支持物,如图5中所示。对于本披露的任何方

法，可无需覆盖该微孔。换言之，在该方法的过程中可无需密封该微孔。当微孔未被覆盖(例如，密封)时，这些孔以是间隔开的，这样使得一个微孔的内容物不会扩散至另一个微孔。

[0325] 可替代地，或另外地，可以在与缀合寡核苷酸的支持物接触之前捕获和/或纯化细胞。用于捕获和/或纯化细胞的方法可以包括使用抗体、分子支架、和/或珠。可以通过流式细胞术纯化细胞。可以将可商购的试剂盒用于捕获或纯化细胞。例如，可以将Dynabeads(R)用于分离细胞。可以将磁性分离用于纯化细胞。可以通过离心纯化细胞。

[0326] 可以通过产生包括细胞和支持物的悬浮液而将细胞与缀合寡核苷酸的支持物接触。悬浮液可以包括凝胶。可以在与缀合寡核苷酸的支持物接触之前将细胞固定在支持物上或在溶液中。可替代地，可以将细胞添加至包括缀合寡核苷酸的支持物的悬浮液。例如，可以将细胞添加至内嵌有缀合寡核苷酸的支持物的水凝胶。

[0327] 可以将单细胞与偶联单寡核苷酸的固体支持物接触。可以将单细胞与缀合多个寡核苷酸的固体支持物接触。多个细胞与缀合单寡核苷酸的固体支持物相互作用。多个细胞与缀合多个寡核苷酸的固体支持物相互作用。这些缀合寡核苷酸的固体支持物可以是细胞类型特异性的。可替代地，缀合寡核苷酸的支持物可以与两个或更多个不同细胞类型相互作用。

[0328] 裂解

[0329] 可以将微孔中的细胞裂解。可以通过机械裂解、热裂解、光学裂解、和/或化学裂解进行裂解。化学裂解可以包括使用消化酶类，如蛋白酶K、胃蛋白酶、以及胰蛋白酶。可以通过将裂解缓冲液添加到微孔中进行裂解。裂解缓冲液可以包括Tris HCl。裂解缓冲液可以包含至少约0.01、0.05、0.1、0.5、或1M或更多Tris HCl。裂解缓冲液可以包含至多约0.01、0.05、0.1、0.5、或1M或更多Tris HCL。裂解缓冲液可以包含约0.1M Tris HCl。裂解缓冲液的pH可以至少是约1、2、3、4、5、6、7、8、9、或10或更多。裂解缓冲液的pH可以至多是约1、2、3、4、5、6、7、8、9、或10或更多。在一些情况下，裂解缓冲液的pH是约7.5。裂解缓冲液可以包含盐(例如，LiCl)。裂解缓冲液中盐的浓度可以至少是约0.1、0.5、或1M或更多。裂解缓冲液中盐的浓度可以至多是约0.1、0.5、或1M或更多。在一些情况下，裂解缓冲液中盐的浓度是约0.5M。裂解缓冲液可以包含洗涤剂(例如，SDS、十二烷基硫酸锂、曲通X、吐温、NP-40)。裂解缓冲液中洗涤剂的浓度可以至少是约0.0001％、0.0005％、0.001％、0.005％、0.01％、0.05％、0.1％、0.5％、1％、2％、3％、4％、5％、6％、或7％或更多。裂解缓冲液中洗涤剂的浓度可以至多是约0.0001％、0.0005％、0.001％、0.005％、0.01％、0.05％、0.1％、0.5％、1％、2％、3％、4％、5％、6％、或7％或更多。在一些情况下，裂解缓冲液中洗涤剂的浓度可以是约1％十二烷基硫酸锂(Li dodecyl sulfate)。该裂解方法中所用时间可以依赖于所用洗涤剂的量。在一些情况下，所用洗涤剂越多，裂解所需时间越短。裂解缓冲液可以包含螯合剂(例如，EDTA、EGTA)。裂解缓冲液中螯合剂的浓度可以至少是约1mM、5mM、10mM、15mM、20mM、25mM、或30mM或更多。裂解缓冲液中螯合剂的浓度可以至多是约1mM、5mM、10mM、15mM、20mM、25mM、或30mM或更多。在一些情况下，裂解缓冲液中螯合剂的浓度是约10mM。裂解缓冲液可以包含还原剂(例如，β-巯基乙醇、DTT)。裂解缓冲液中还原剂的浓度可以至少是约1mM、5mM、10mM、15mM、或20mM或更多。裂解缓冲液中还原剂的浓度可以至多是约1mM、5mM、10mM、15mM、或20mM或更多。在一些情况下，裂解缓冲液中还原剂的浓度是约5mM。在一些情况下，裂解缓冲液可以包含约0.1M TrisHCl、约pH 7.5、约0.5M LiCl、约1％十二烷基硫酸

锂、约10mM EDTA和约5mM DTT。

[0330] 可以在约4、10、15、20、25、或30C的温度下进行裂解。裂解可以进行约1、5、10、15、或20或更多分钟。裂解的细胞可以包含至少约100000、200000、300000、400000、500000、600000、或700000或更多个靶核酸分子。裂解的细胞可以包含至多约100000、200000、300000、400000、500000、600000、或700000或更多个靶核酸分子。图7示出关于可以从裂解中获得的靶核酸(即mRNA)的浓度的示例性统计。密封

[0331] 可以在裂解过程中密封微孔阵列的微孔。密封可以用于防止相邻微孔之间靶核酸的交叉杂交。可以使用如在图8A和8B中所示的帽来密封微孔。帽可以是固体支持物。帽可以包括珠。珠的直径可以大于微孔的直径。例如,帽可以至少比微孔的直径大约10%、20%、30%、40%、50%、60%、70%、80%或90%。例如,帽可以至多比微孔的直径大约10%、20%、30%、40%、50%、60%、70%、80%或90%。

[0332] 帽可以包括交联葡聚糖珠(例如,Sephadex)。交联葡聚糖可以在从约10微米至约80微米的范围内。帽的交联葡聚糖可以是从20微米至约50微米。帽可以包括例如anopore无机膜(例如,氧化铝)、透析膜、载玻片、盖玻片、和/或亲水性塑料膜(例如膜,该膜包覆与裂解缓冲液水合的琼脂糖的薄膜)。

[0333] 帽可以允许缓冲液穿过进入微孔,但可以阻止大分子(例如,核酸)从孔中迁移出来。通过帽可以阻挡至少约1、2、3、4、5、6、7、8、9、10、12、13、14、15、16、17、18、19、或20或更多个核苷酸的大分子迁移进入或离开微孔。通过帽可以阻挡至多约1、2、3、4、5、6、7、8、9、10、12、13、14、15、16、17、18、19、或20或更多个核苷酸的大分子迁移进入或离开微孔。

[0334] 密封的微孔阵列可以包括微孔顶部的单层的珠。密封的微孔阵列可以包括微孔顶部的多层的珠。密封的微孔阵列可以包含约1、2、3、4、5、或6或更多层的珠。

[0335] 将一个珠或多个珠沉积到固体支持物(例如,微孔阵列)上可以是随机或非随机的。例如,使珠与微孔阵列接触可以是随机或非随机接触。在一些实施例中,使珠与微孔阵列随机接触。在一些实施例中,使珠与微孔阵列非随机接触。将多个珠沉积到微孔阵列上可以是随机或非随机的。例如,使多个珠与微孔阵列接触可以是随机或非随机接触。在一些实施例中,使多个珠与微孔阵列随机接触。在一些实施例中,使多个珠与微孔阵列非随机接触。

[0336] 对分子的随机标记

[0337] 其中样品标签或分子鉴定物标记是寡核苷酸,寡核苷酸附接至核酸可以通过多种方法发生,包括,但不限于,寡核苷酸杂交至核酸。在一些情况下,寡核苷酸包含靶特异性区。靶特异性区可以包括与待标记的分子的至少一部分互补的序列。靶特异性区可以与分子杂交,从而产生经标记的核酸。寡核苷酸杂交至核酸之后可以是核酸延伸反应。核酸延伸反应可以是逆转录。

[0338] 使多个核酸与样品标签附接(可替代地称为接触)可以包括使样品标签杂交至该多个核酸中的一个或多个。使多个核酸与样品标签接触可以包括进行核酸延伸反应。核酸延伸反应可以是逆转录反应。

[0339] 使多个核酸与分子鉴定物标记接触可以包括使分子鉴定物标记杂交至该多个核酸中的一个或多个。使多个核酸与分子鉴定物标记接触可以包括进行核酸延伸反应。核酸延伸反应可以包括逆转录。

[0340] 使多个核酸与分子鉴定物标记接触可以包括使样品标签杂交至该多个核酸中的一个或多个。使多个核酸与分子鉴定物标记接触可以包括使分子鉴定物标记杂交至样品标签。

[0341] 使多个核酸与样品标签接触可以包括使分子鉴定物标记杂交至该多个核酸中的一个或多个。使多个核酸与样品标签接触可以包括将样品标签杂交至分子鉴定物标记。

[0342] 样品标签和/或分子鉴定物标记附接至核酸可以通过连接发生。使多个核酸与样品标签接触可以包括将标签接触连接至该多个核酸中的任一个。使多个核酸与分子鉴定物标记接触可以包括使分子鉴定物标记连接至该多个核酸中的一个或多个。使多个核酸与样品标签接触可以包括使分子鉴定物标记连接至一个或多个核酸。使多个核酸与分子鉴定物标记接触可以包括使样品标签连接至这些核酸中的一个或多个。连接技术包括平端连接和粘端连接。连接反应可包括DNA连接酶，例如DNA连接酶I、DNA连接酶III、DNA连接酶IV和T4DNA连接酶。连接反应可包括RNA连接酶，例如T4RNA连接酶I和T4RNA连接酶II。

[0343] 连接方法在例如萨拉布鲁克(Sambrook)等人.(2001)和新英格兰生物实验室(New England biolab)目录(这两篇文献均通过引用并入本文而用于所有目的)中进行了描述。方法包括使用T4DNA连接酶，其催化在具有平端和粘端的双链体DNA或RNA中并列的5’磷酸和3’羟基末端之间形成磷酸二酯键；使用Taq DNA连接酶，其催化在与互补靶DNA杂交的两个相邻寡核苷酸中并列的5’磷酸和3’羟基末端之间形成磷酸二酯键；使用大肠杆菌DNA连接酶，其催化在含有粘性末端的双链体DNA中并列的5’-磷酸和3’-羟基末端之间形成磷酸二酯键；以及使用T4RNA连接酶，其催化通过形成3’→5’磷酸二酯键使末端为5’磷酰基的核酸供体与末端为3’羟基的核酸受体连接，底物包括单链RNA和DNA以及二核苷焦磷酸；或者本领域中描述的任何其他方法。在将适配子与一个或两个末端连接之前，可采用一种或多种酶例如内切核酸酶处理片段化的DNA，以利于通过产生适于连接的末端而进行的连接。

[0344] 在一些情况下，寡核苷酸的两个末端均附接至分子。例如，寡核苷酸的两个末端可与分子的一个或多个末端杂交和/或连接。在一些情况下，寡核苷酸的两个末端与分子的两个末端的附接导致环形的经标记的核酸的形成。寡核苷酸的两个末端也可附接至分子的同一末端。例如，寡核苷酸的5’末端与分子的3’末端连接，并且寡核苷酸的3’末端与分子的3’末端杂交，从而产生在一个末端具有发夹结构的经标记的核酸。在一些情况下，寡核苷酸附接至分子的中部。

[0345] 在一些情况下，寡核苷酸附接至核酸包括将一个或多个寡核苷酸接头附接至多个核酸。方法可以进一步包括将一个或多个寡核苷酸接头附接至样品标签化的核酸。方法可以进一步包括将一个或多个寡核苷酸接头附接至经标记的核酸。将一个或多个寡核苷酸接头附接至核酸、样品标签或分子鉴定物标记可以包括将一个或多个寡核苷酸接头连接至核酸、样品标签或分子鉴定物标记。一个或多个接头可以包括至少约1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、25、30、35、40、45、50、55、60、70、80、90、100个核苷酸。在一些情况下，接头可以包括至少约1000个核苷酸。

[0346] 在一些情况下，分子条码与分子的附接包括使用一个或多个适配子。如在此所使用，术语“适配子”和“适配子区”可以互换地使用。适配子可以包含允许适配子与分子附接的靶特异性区，以及允许分子条码与适配子附接的寡核苷酸特异性区域。适配子可以进一步包括通用引物。适配子可以进一步包括通用PCR区。可通过包括但不限于杂交和/或连接

的方法将适配子附接至分子和/或分子条码。

[0347] 用于连接适配子与核酸片段的方法是众所周知的。适配子可以是双链、单链或部分单链的。在一些方面，适配子由具有互补性(例如，约10至30个碱基，或约15至40个碱基的完全互补性)区域的两个寡核苷酸形成，从而使得当两个寡核苷酸杂交在一起时，它们形成双链区域。任选地，这两个寡核苷酸中的一个或两个可具有与另一寡核苷酸不互补的区域，因而在该适配子的一个或两个末端形成单链突出端。单链突出端可以突出约1个至约8个碱基，或约2个至约4个碱基。突出端可与通过采用限制性内切酶切割产生的突出端互补，以利于“粘端”连接。适配子可包含其他特征，例如引物结合位点和限制位点。在一些方面，限制位点可针对IIS型限制酶或在其识别序列以外进行切割的另一种酶，例如EcoP151(参见，穆克(Mucke)等人，分子生物学杂志(J Mol Biol)2001,312(4):687-698，以及US 5,710,000，这些文献通过引用整体并入本文)。

[0348] 在一些情况下，随机计数多个样品中核酸的拷贝数目包括检测适配子、适配子的互补体、适配子的反向互补体或它们的一部分，以确定不同的经标记的核酸的数目。检测适配子、适配子的互补体、适配子的反向互补体或它们的一部分可以包括对适配子、适配子的互补体、适配子的反向互补体或它们的一部分进行测序。

[0349] 分子条码可附接至分子的任何区域。例如，分子条码可附接至多核苷酸(例如，DNA、RNA)的5’末端或3’末端。例如，分子条码的靶特异性区包含与分子的5’区中的序列互补的序列。分子条码的靶特异性区也可包含与分子的3’区中的序列互补的序列。在一些情况下，分子条码附接至基因或基因产物内的区域。例如，基因组DNA被片段化，且样品标签或分子鉴定物标记附接至该片段化的DNA。在其他情况下，RNA分子被选择性地剪接，并且分子条码附接至该选择性剪接的变体。在另一个实例中，多核苷酸被消化，并且分子条码附接至被消化的多核苷酸。在另一个实例中，分子条码的靶特异性区包含与分子内的序列互补的序列。

[0350] 包含发夹的分子条码、样品标签(例如，样品索引)、细胞标记、或分子鉴定物标记(例如，分子标记)可以充当用于杂交链反应(HCR)的探针，并且因此可以称之为HCR探针。HCR探针可以包括分子条码，该分子条码包含发夹结构。HCR探针可以包括样品标签，该样品标签包含发夹结构。HCR探针可以包括分子鉴定物标记，该分子鉴定物标记包含发夹结构。本文进一步公开了基于随机标记的杂交链反应(HCR)方法，该方法包括：采用HCR探针随机地标记一个或多个核酸分子，其中HCR探针包括分子条码，该分子条码包含发夹，并且该一个或多个核酸分子充当杂交链反应的引发物。本文进一步公开了基于随机标记的杂交链反应(HCR)方法，该方法包括：采用HCR探针随机地标记一个或多个核酸分子，其中HCR探针包括样品标签，该样品标签包含发夹，并且该一个或多个核酸分子充当杂交链反应的引发物。本文进一步公开了基于随机标记的杂交链反应(HCR)方法，该方法包括：采用HCR探针随机地标记一个或多个核酸分子，其中HCR探针包括分子鉴定物标记，该分子鉴定物标记包含发夹，并且该一个或多个核酸分子充当杂交链反应的引发物。

[0351] HCR探针可以包括带有突出端区域的发夹。发夹的突出端区域可以包含靶特异性区。突出端区域可以包含寡dT序列。包含一个或多个核酸分子的样品可在进行随机标记之前使用一种或多种限制性核酸酶进行处理。突出端区域可以包含限制性内切酶识别序列。包含一个或多个核酸分子的样品可在进行随机标记之前与一个或多个适配子接触，以产生

适配子-核酸分子杂合体。突出端区域和茎部可与一个或多个适配子互补。HCR探针可以包括带有环的发夹。HCR探针的环可以包括标记区和/或样品索引区。

[0352] 第一HCR探针与核酸分子的杂交可导致经标记的核酸的形成，其中使第一HCR探针线性化，以产生第一线性化HCR探针。经标记的核酸的第一线性化HCR探针可充当第二HCR探针与经标记的核酸杂交的引发物，以产生具有两个线性化HCR探针的经标记的核酸。第二线性化HCR探针可充当另一杂交反应的引发物。这一过程可重复多次，以产生具有多个线性化HCR探针的经标记的核酸。HCR探针上的可检测标记可以使得能够检测经标记的核酸。可检测标记可以是任何类型的标记(例如，荧光团、生色团、小分子、纳米粒子、半抗原、酶、抗体、磁体)。可检测标记可以包含单个标记的片段。当可检测标记紧密靠近时，它们可以产生可检测的信号。当HCR探针为发夹时，可检测标记可能由于距离太远而无法产生可检测的信号。当HCR探针被线性化并且多个线性化的HCR探针杂交在一起时，可检测标记可充分靠近，以产生可检测的信号。例如，HCR探针可以包含两个作为可检测标记的茫部分。可替代地，可检测标记可以是纳米粒子。基于随机标记的HCR方法可使多个发夹HCR探针能够附接至经标记的核酸，这可导致信号放大。基于随机标记的HCR可提高核酸分子的检测、分析和/或定量的灵敏度。基于随机标记的HCR可提高一个或多个核酸分子的检测、分析和/或定量的准确度。

[0353] 裂解后，细胞的靶核酸可以杂交至缀合至固体支持物的寡核苷酸。靶核酸可以杂交至寡核苷酸的靶结合区。核酸可以杂交至寡核苷酸的任何区域。

[0354] 在一些情况下，不是所有寡核苷酸都可以结合靶核酸。这是因为在一些情况下，寡核苷酸的数目大于靶核酸的数目。缀合至固体支持物的寡核苷酸的数目可以是细胞中靶核酸的数目的1、2、3、4、5、6、7、8、9、或10倍。至少10％、20％、30％、40％、50％、60％、70％、80％、90％或100％的寡核苷酸可以与靶核酸结合。至多10％、20％、30％、40％、50％、60％、70％、80％、90％或100％的寡核苷酸可以与靶核酸结合。在一些情况下，至少1、2、3、4、5、6、7、8、9、10、20、30、40、50、60、70、80、90或100或更多个不同靶核酸可以被固体支持物上的寡核苷酸捕获。在一些情况下，至多1、2、3、4、5、6、7、8、9、10、20、30、40、50、60、70、80、90或100或更多个不同靶核酸可以被固体支持物上的寡核苷酸捕获。

[0355] 在一些情况下，靶核酸的拷贝数的至少约40％、50％、60％、70％、80％、90％、95％、96％、97％、98％、99％或100％结合至固体支持物上的寡核苷酸。在一些情况下，靶核酸的拷贝数的至多约40％、50％、60％、70％、80％、90％、95％、96％、97％、98％、99％或100％结合至固体支持物上的寡核苷酸。

[0356] 收回

[0357] 裂解后，可以将固体支持物收回。可以通过使用磁体进行固体支持物的收回。可以通过熔融微孔阵列和/或超声处理进行固体支持物的收回。固体支持物的收回可以包括离心。固体支持物的收回可以包括尺寸排阻。在一些情况下，从微孔回收至少约50％、60％、70％、80％、90％、95％或100％的固体支持物。在一些情况下，从微孔回收至多约50％、60％、70％、80％、90％、95％或100％的固体支持物。

[0358] 逆转录

[0359] 本文公开的方法可进一步包括对标记的RNA分子进行逆转录，以产生经标记的cDNA分子。在一些情况下，寡核苷酸的至少一部分充当逆转录反应的引物。寡核苷酸的寡dT

部分可以充当cDNA分子的第一链合成的引物。

[0360] 在一些情况下，标记的cDNA分子可用作用于新的随机标记反应的分子。标记的cDNA可以具有在逆转录之前附接至RNA的第一标签或标签组，以及附接至cDNA分子的第二标签或标签组。这些多重标记反应可用于，例如，确定在第一标签和第二标签的附接之间发生的事件(例如任选的扩增反应或逆转录反应)的效率。

[0361] 在另一个实例中，寡核苷酸附接至RNA分子的5’末端，以产生标记的RNA分子。标记的RNA分子的逆转录可通过加入逆转录引物而进行。在一些情况下，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引物的长度为12-18个核苷酸，并结合至在哺乳动物mRNA的3’末端的内源性聚(A)+尾部。随机六核苷酸引物可在多个互补位点结合至mRNA。靶特异性寡核苷酸引物通常选择性地引发感兴趣的mRNA。

[0362] 在一些情况下，该方法包括对标记的RNA分子反复进行逆转录，以产生多个经标记的cDNA分子。本文公开的方法可包括进行至少约1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19或20次逆转录反应。该方法可包括进行至少约25、30、35、40、45、50、55、60、65、70、75、80、85、90、95或100次逆转录反应。

[0363] 核酸合成(例如，cDNA合成)可以在收回的固体支持物上进行。核酸合成可以在管中和/或在转子上进行以保持固体支持物悬浮。可以将所得合成的核酸用于随后的核酸扩增和/或测序技术中。核酸合成可以包括在附接至固体支持物上的寡核苷酸的RNA上产生cDNA拷贝。产生cDNA拷贝可以包括使用具有RT活性的逆转录酶(RT)或DNA聚合酶。这可以导致单链cDNA分子的产生。在核酸合成之后，可以将未使用的寡核苷酸从固体支持物中去除。可以通过外切核酸酶处理(例如，通过ExoI)进行寡核苷酸的去除。

[0364] 在一些实施例中，可以使用化学切割将核酸从固体支持物中去除。例如，可以将存在于核酸中的化学基团或经修饰的碱基用于促进将其从固体支持物中去除。例如，酶可以用于从固体支持物中去除核酸。例如，可以通过限制性内切核酸酶消化将核酸从固体支持物中去除。例如，用尿嘧啶-d-糖基化酶(UDG)处理含有dUTP或ddUTP的核酸可以用于从固体支持物中去除核酸。例如，可以使用进行核苷酸切除(例如，碱基切除修复酶(例如，脱嘌呤/脱嘧啶(AP)核酸内切酶))的酶将核酸从固体支持物中去除。在一些实施例中，可以使用光分裂(photocleavable)基团以及光将核酸从固体支持物中去除。在一些实施例中，可切割性接头可以用于从固体支持物中去除核酸。例如，可切割性接头可以包括以下各项中的至少一种：生物素/亲和素、生物素/链霉抗生物素蛋白、生物素/中性链亲和素、Ig蛋白A、光不稳定性接头、酸或碱不稳定性接头基团、或适体。

[0365] 在一些实施例中，核酸不进行扩增。在一些实施例中，在核酸测序前核酸不进行扩增。在一些实施例中，未附接至固体支持物的核酸可以不经预先扩增直接测序。在一些实施例中，当未附接至固体支持物时，核酸可以无需进行扩增而直接测序，例如，附接至固体支持物的核酸可以在附接至固体支持物的同时直接测序。在一些实施例中，已经从固体支持物中去除的核酸可以直接测序。例如，已经从固体支持物中去除的核酸可以无需进行扩增而直接测序。任何有利于无需扩增进行测序的测序平台可以用于进行该测序。

[0366] 扩增

[0367] 在核酸已经合成(例如，逆转录的)之后，可以将其扩增。扩增可以多路方式进行，其中多个靶核酸序列同时进行扩增。扩增可以将测序适配子添加至核酸。扩增可以通过聚

合酶链式反应(PCR)来进行。PCR可指用于通过DNA的互补链的同时引物延伸使特异性DNA序列体外扩增的反应。PCR可涵盖该反应的派生形式，包括但不限于，RT-PCR、实时PCR、巢式PCR、定量PCR、多重PCR、数字PCR，和组装PCR。

[0368] 该方法可以进一步包括进行一个或多个扩增反应以产生经标记的核酸扩增子。可以在检测经标记的核酸之前扩增这些经标记的核酸。本发明进一步包括在进行一个或多个扩增反应之前组合第一和第二样品。

[0369] 扩增反应可以包括扩增至少一部分样品标签。扩增反应可以包括扩增至少一部分标记。扩增反应可以包括扩增至少一部分样品标签、标记、核酸、或它们的组合。扩增反应可以包括扩增多个核酸的至少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％。该方法可以进一步包括进行一个或多个cDNA合成反应，以产生样品标签化的核酸或经分子鉴定物标记的核酸的一个或多个cDNA拷贝。

[0370] 对经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。对经标记的核酸的扩增可以包括对经标记的核酸的指数式扩增。对经标记的核酸的扩增可以包括对经标记的核酸的线性扩增。

[0371] 在一些情况下，对经标记的核酸的扩增包括非基于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(circle-to-circle amplification)。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合成和转录的多重循环以扩增DNA或RNA靶标(WO 89/01050;WO 88/10315;以及美国专利号5,130,238;5,409,818;5,466,586;5,514,545;5,554,517;5,888,779;6,063,603;以及6,197,554)、连接酶链式反应(LCR)、Qβ复制酶(Qβ)方法，如描述于美国专利号4,786,600，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、一种扩增方法，其中将引物杂交至核酸序列并且将所得双链体在延伸反应和扩增之前切割，使用缺乏5’外切核酸酶活性的核酸聚合酶的链置换扩增(美国专利号6,214,587)，滚环扩增、以及分支延伸扩增(RAM)(美国专利号5,942,391)。

[0372] 对经标记的核酸的扩增可包括基于杂交链反应(HCR)的方法(德克斯(Dirks)和皮尔斯(Pierce)，PNAS，2004;张(Zhang)等人，分析化学(Anal Chem)，2012)。基于HCR的方法可包括基于DNA的HCR。基于HCR的方法可包括一个或多个标记的探针。该一个或多个标记的探针可以包括本文所公开的一个或多个样品标签或分子鉴定物标记、或其互补体。

[0373] 在一些情况下，本文公开的方法进一步包括对经标记的核酸(例如，标记的RNA、标记的DNA、标记的cDNA)进行聚合酶链反应，以产生经标记的扩增子。经标记的扩增子可以是双链分子。该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包含样品标签或分子鉴定物标记。可替代地，经标记的扩增子是单链分子。该单链分子可包括DNA、RNA或它们的组合。本发明的核酸可包括合成的或改变的核酸。

[0374] 可通过诸如PCR、HD-PCR、下一代PCR、数字RTA或其任意组合的方法进行聚合酶链反应。另外的PCR方法包括但不限于，等位基因特异性PCR、Alu PCR、组装PCR、不对称PCR、微滴PCR、乳液PCR、依赖解旋酶的扩增HDA、热启动PCR、反向PCR、指数后线性(LATE)-PCR

(linear-after-the-exponential(LATE)-PCR)、长PCR、多重PCR、巢式PCR、半巢式PCR、定量PCR、RT-PCR、实时PCR、单细胞PCR以及递降PCR(touchdown PCR)或它们的组合。

[0375] 多重PCR反应可以包括巢式PRC反应。该方法可以包括引物对,其中第一引物退火至多个核酸中任一个的从该多个核酸中任一个的3'端的至少约300至400个核苷酸处,并且第二引物退火至多个核酸中任一个的从该多个核酸中任一个的3'端的至少约200至300个核苷酸处,其中该第一引物和第二引物朝向该多个核酸中任一个的3'端产生互补DNA合成。

[0376] 在一些情况下,进行聚合酶链反应包括将第一靶特异性引物与经标记的核酸退火。可替代地或另外地,进行聚合酶链反应进一步包括将通用引物与样品标签或分子鉴定物标记的通用引物结合位点区域退火,其中该样品标签或分子鉴定物标记在经标记的核酸或经标记的扩增子上。本文公开的方法可进一步包括将第二靶特异性引物与经标记的核酸和/或经标记的扩增子退火。

[0377] 在一些情况下,该方法包括反复扩增经标记的核酸以产生多个经标记的扩增子。本文公开的方法可包括进行至少约1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19或20次扩增反应。可替代地,该方法包括进行至少约25、30、35、40、45、50、55、60、65、70、75、80、85、90、95或100次扩增反应。

[0378] 其他适合的扩增方法包括连接酶链式反应(LCR)(例如,吴(Wu)和华莱士(Wallace),基因组学(Genomics),4:560(1989),兰德格伦(Landegren)等人,科学(Science),241:1077(1988)以及巴林杰(Barringer)等人基因(Gene)89:117(1990)),转录扩增(郭(Kwoh)等人,美国科学院院刊(Proc.Natl.Acad.Sci.USA),86:1173(1989)以及WO88/10315),自持序列复制(瓜特利(Guatelli)等人,美国科学院院刊(Proc.Nat.Acad.Sci.87,1874(1990)以及WO 90/06995)、靶多核苷酸序列的选择性扩增(美国专利号6,410,276)、共有序列引物聚合酶链反应(CP-PCR)(美国专利号4,437,975)、任意引物聚合酶链反应(AP-PCR)(美国专利号5,413,909、5,861,245)、滚环扩增(RCA)(例如,福爱尔(Fire)和徐(Xu),PNAS 92:4641(1995)以及刘(Liu)等人,美国化学会志(J.Am.Chem.Soc.118:1587(1996))以及美国专利号5,648,245)、链置换扩增(参见见蓝斯肯(Lasken)和艾格霍姆(Egholm),生物技术趋势(Trends Biotechnol.)2003 21(12):531-5;巴克(Barker)等人,基因组研究(Genome Res.)2004年5月;14(5):901-7;迪恩(Dean)等人,美国科学院院报(Proc Natl Acad Sci U S A.)2002;99(8):5261-6;沃克(Walker)等人,1992,核酸研究(Nucleic Acids Res.)20(7):1691-6,1992以及帕兹(Paez)等人,核酸研究(Nucleic Acids Res.)2004;32(9):e71),Qbeta复制酶(描述于PCT专利申请号PCT/US87/00880中)以及基于核酸的序列扩增(NABSA)(参见美国专利号5,409,818、5,554,517和6,063,603,其中每个专利均通过引用并入本文)。其他可使用的扩增方法描述于美国专利号6,582,938、5,242,794、5,494,810、4,988,617和美国公开号20030143599,其中每个专利通过引用并入本文。DNA还可通过多重基因座特异性PCR或使用适配子连接和单引物PCR(参见金斯勒(Kinzler)和沃格尔斯坦(Vogelstein),NAR(1989)17:3645-53)进行扩增。也可使用其他可用的扩增方法,如平衡PCR(Makrigiorgos等人,(2002),自然生物技术(Nat Biotechnol),第20卷,第936-9页)。

[0379] 分子倒置探针("MIP")也可用于对选定的靶标进行扩增。可生成MIP以使环前探针(pre-circle probe)的末端与位于待扩增区域侧翼的区域互补。可通过延伸探针末端而使

缺口闭合，从而使靶标的互补体在末端连接以形成闭环之前并入MIP。该闭环可通过测序或杂交来扩增和检测，如先前在Hardenbol等人，基因组研究(Genome Res.)15:269-275(2005)以及在美国专利号6,858,412)中所公开的。

[0380] 扩增可进一步包括将一个或多个对照核酸添加至一个或多个包括多个核酸的样品中。扩增可进一步包括将一个或多个对照核酸添加至多个核酸中。对照核酸可以包括对照标记。

[0381] 扩增可以包括使用一种或多种非天然核苷酸。非天然核苷酸可以包括光不稳定和/或可触发的核苷酸。非天然核苷酸的实例包括，但不限于，肽核酸(PNA)、吗啉代和锁核酸(LNA)、以及二醇核酸(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多个循环中。添加非天然核苷酸也可以用于鉴别扩增反应中特定循环或时间点的产物。

[0382] 进行一个或多个扩增反应可以包括使用一个或多个引物。该一个或多个引物可以包括一个或多个寡核苷酸。该一个或多个寡核苷酸可以包括至少约7-9个核苷酸。该一个或多个寡核苷酸可以包括少于12-15个核苷酸。该一个或多个引物可以退火至多个经标记的核酸的至少一部分。该一个或多个引物可以退火至多个经标记的核酸的3’端和/或5’端。该一个或多个引物可以退火至多个经标记的核酸的内部区。内部区可以是从该多个经标记的核酸的3’端的至少约50、100、150、200、220、230、240、250、260、270、280、290、300、310、320、330、340、350、360、370、380、390、400、410、420、430、440、450、460、470、480、490、500、510、520、530、540、550、560、570、580、590、600、650、700、750、800、850、900或1000个核苷酸。该一个或多个引物可以包括固定的引物面板。该一个或多个引物可以包括至少一个或多个定制引物。该一个或多个引物可以包括至少一个或多个对照引物。该一个或多个引物可以包括至少一个或多个管家基因引物。该一个或多个寡核苷酸可以包括选自下组的序列，该组由表23中的序列组成。该一个或多个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。该一个或多个定制引物可以退火至第一样品标签、第二样品标签、分子鉴定物标记、核酸或它们的产物。该一个或多个引物可以包括通用引物和定制引物。该定制引物可以被设计成扩增一个或多个靶核酸。这些靶核酸可以包括一个或多个样品中总核酸的亚群。这些靶核酸可以包括一个或多个样品中总的经标记的核酸的亚群。该一个或多个引物可以包括至少96个或更多个定制引物。该一个或多个引物可以包括至少960个或更多个定制引物。该一个或多个引物可以包括至少9600个或更多个定制引物。该一个或多个定制引物可以退火至两个或更多个不同的经标记的核酸。该两个或更多个不同的经标记的核酸可以对应于一个或多个基因。

[0383] 在此披露的是选择定制引物的方法，该方法包括：a)第一遍，其中选择的引物可以包括：i)不多于三个连续的鸟嘌呤，不多于三个连续的胞嘧啶，不多于四个连续的腺嘌呤，以及不多于四个连续的胸腺嘧啶；ii)为鸟嘌呤或胞嘧啶的至少3、4、5、或6个核苷酸；以及iii)不易于形成发夹结构的序列；b)第二遍，包括：i)第一轮，选择具有所有转录物的高覆盖的多个序列；以及ii)一个或多个随后的轮，选择具有剩余转录物的最高覆盖和不大于4的与其他选择的序列的互补分数的序列；以及c)添加序列至挑选的组，直至覆盖饱和或客户引物的总数小于或等于约96。

[0384] 选择定制引物的方法可以进一步包括选择至少一个共同引物，其基于一个或多个

mRNA转录物、包括结构RNA的非编码转录物、转录假基因、由基因组注释过程提供的模型mRNA、对应于基因组重叠群的序列、或它们的任何组合。

[0385] 选择定制引物的方法可以进一步包括引物选择方法，该方法富集一个或多个亚组的核酸。该一个或多个亚组可以包括低丰度mRNA。

[0386] 选择定制引物的方法可以进一步包括计算算法。在该方法中使用的引物可以使用Primer 3(一种基于用户定义的输入序列建议引物序列的计算机程序)进行设计。也可使用其他的引物设计，或者可在没有计算机程序辅助下用眼睛选择引物。在该程序中有许多可用的选项以使引物设计适合于大多数应用。Primer 3可以考虑许多因素，包括但不限于，寡核苷酸的解链温度、长度、GC含量、3’末端稳定性、估计的二级结构、与不期望的序列退火或对其进行扩增的可能性(例如散在重复序列)以及在同一引物的两个拷贝之间形成引物二聚体的可能性。在设计引物对时，Primer 3可以考虑产物大小和解链温度、在引物对中的两个引物之间形成引物二聚体的可能性、引物的解链温度之间的差异以及引物相对于将要避免的特定目标区域的位置。

[0387] 在此披露的方法、组合物和试剂盒可以包括披露于表23-24中的一个或多个引物。

[0388] 测序

[0389] 在一些方面，确定不同的经标记的核酸的数目可以包括确定经标记的核酸或其任意产物(例如，经标记的扩增子、经标记的cDNA分子)的序列。在一些情形下，扩增的靶核酸可经受测序。确定经标记的核酸或其任意产物的序列可以包括进行测序反应以确定样品标签、分子鉴定物标记的至少一部分的、经标记的核酸的至少一部分的、其互补体的、其反向互补体的或它们的任意组合的序列。在一些情况下，仅对样品标签或样品标签的一部分进行测序。在一些情况下，仅对分子鉴定物标记或分子鉴定物标记的一部分进行测序。

[0390] 确定经标记的核酸或其任意产物的序列可以通过诸如以下的测序方法进行：Helioscope™单分子测序、纳米孔DNA测序、Lynx Therapeutics的大规模平行签名测序(MPSS)、454焦磷酸测序、单分子实时(RNAP)测序、Illumina(Solexa)测序、SOLiD测序、Ion Torrent™、离子半导体测序、单分子SMRT(™)测序、聚合酶克隆测序(Polony sequencing)、DNA纳米球测序以及VisiGen生物技术(及VisiGen Biotechnologies)的方法。可替代地，确定经标记的核酸或其任意产物的序列可以使用测序平台，包括但不限于：Illumina提供的基因组分析仪(Genome Analyzer)IIx、HiSeq和MiSeq，单分子实时(SMRT™)技术，例如由太平洋生物科学公司(Pacific Biosciences)(加利福尼亚(California))提供的PacBio RS系统和Solexa测序仪，真正单分子测序(tSMS™)技术，例如由Helicos Inc(剑桥(Cambridge)，马萨诸塞州(MA))提供的HeliScope™测序仪。

[0391] 在一些实施例中，经标记的核酸包括代表来自生物体基因组基因的约0.01％至生物体基因组基因的约100％的核酸。例如，可以使用包括多个多聚体的靶标互补区域，通过从样品中捕获含有互补序列的基因，对约0.01％的生物体基因组基因至约100％的生物体基因组基因进行测序。在一些实施例中，经标记的核酸包括代表来自约0.01％的生物体转录组转录物至约100％的生物体转录组转录物的核酸。例如，可以使用包括聚-T尾的靶标互补区域，通过从样品中捕获mRNA，对约0.501％的生物体转录组转录物至约100％的生物体转录组转录物进行测序。

[0392] 在一些情况下，确定经标记的核酸或其任意产物的序列包括配对末端测序、纳米

孔测序、高通量测序、鸟枪法测序、染料终止剂测序、多重引物DNA测序、引物步移、桑格双脱氧测序法、Maxim-Gilbert测序、焦磷酸测序、真正的单分子测序或它们的任意组合。可替代地，可以通过电子显微镜分析法或化学敏感场效应晶体管(chemFET)阵列来确定经标记的核酸或其任意产物的序列。

[0393] 可以使用多种测序方法进行确定核酸的序列(例如，经扩增的核酸、经标记的核酸、经标记的核酸的cDNA拷贝等)，这些方法包括但不限于：杂交测序(SBH)、连接法测序(SBL)、量化增量荧光核苷酸附加测序(quantitative incremental fluorescent nucleotide addition sequencing)(QIFNAS)、分段连接与断裂、荧光共振能量转移(FRET)、分子信标、TaqMan报告探针消化、焦磷酸测序、荧光原位测序(FISSEQ)、FISSEQ珠粒、摆动测序(wobble sequencing)、多重测序、聚合集群(polymerized colony)(POLONY)测序；纳米格滚环测序(nanogrid rolling circle sequencing)(ROLONY)、等位基因特异性寡核苷酸连接检验(allele-specific oligo ligation assay)(例如，寡核苷酸连接检验(OLA)、使用连接的线性探针和滚环扩增(RCA)读出、连接的持锁探针的单模板分子(single template molecule)OLA、和/或使用连接的环形持锁探针和滚环扩增(RCA)读出的单模板分子OLA)等。也可以使用高通量测序方法，如使用平台(如Roche 454、Illumina Solexa、ABI-SOLiD、ION Torrents、完整基因(Complete Genomics)、太平洋生物科学(Pacific Bioscience)、螺旋(Helicos)、Polonator平台)的循环阵列测序。测序可以包括MiSeq测序。测序可以包括HiSeq测序。测序可以读取细胞标记、分子标记和/或原始寡核苷酸上的基因。

[0394] 在另一个实例中，确定经标记的核酸或其任意产物的序列包括RNA-Seq或微小RNA测序。可替代的，确定经标记的核酸或其任意产物的序列包括蛋白质测序技术，例如埃德曼降解、多肽质量指纹分析、质谱法或蛋白酶消化。

[0395] 在某些实施例中，测序反应可发生在固相或半固相支持物上、在凝胶中、在乳液中、在表面上、在珠粒上、在液滴中、在连续流中、在稀释液中或在一个或多个物理上分离的体积中。

[0396] 测序可包括对经标记的核酸的至少约10、20、30、40、50、60、70、80、90、100个或更多个核苷酸或碱基对进行测序。在一些情况下，测序包括对经标记的核酸的至少约200、300、400、500、600、700、800、900、1000个或更多个核苷酸或碱基对进行测序。在其他情况下，测序包括对经标记的核酸的至少约1500；2,000；3,000；4,000；5,000；6,000；7,000；8,000；9,000；或10,000个或更多个核苷酸或碱基对进行测序。

[0397] 测序可以包括至少约200、300、400、500、600、700、800、900、1000个或更多个测序读数/运行。在一些情况下，测序包括测序至少约1500；2,000；3,000；4,000；5,000；6,000；7,000；8,000；9,000；或10,000个或更多个测序读数/运行。测序可以包括小于或等于约1,600,000,000个测序读数/运行。测序可以包括小于或等于约200,000,000个读数/运行。

[0398] 确定不同的经标记的核酸的数目可以包括一个或多个阵列。

[0399] 确定不同的经标记的核酸的数目可以包括使经标记的核酸与一个或多个探针接触。

[0400] 如在此描述的探针可以包含与经标记的核酸或经标记的扩增子的至少一部分互补的序列。该多个探针可以排列于固体支持物上的非连续区域中，其中固体支持物上的非

连续区域包含具有相同或接近相同的序列的探针。在一些情况下，固体支持物上的两个或更多个非连续区域包含两种不同的探针，这些探针包含与寡核苷酸标签的两个不同的唯一性鉴定物区域的序列互补的序列。

[0401] 在一些情况下，该多个探针与阵列杂交。该多个探针能够使经标记的分子与阵列杂交。该多个探针可包含与随机标记寡dT互补的序列。可替代地，或另外地，该多个探针包含与该分子互补的序列。

[0402] 确定不同的经标记的核酸的数目可以包括使经标记的核酸与多个探针的阵列接触。确定不同的经标记的核酸的数目可以包括使经标记的核酸与多个探针的载玻片接触。

[0403] 确定不同的经标记的核酸的数目可以包括经标记的探针杂交、靶特异性扩增、靶特异性测序、用特异性针对靶小核苷酸多态性的经标记的核苷酸进行测序、用特异性针对限制性内切酶消化模式的经标记的核苷酸进行测序、用特异性针对突变的经标记的核苷酸进行测序、或它们的组合。

[0404] 确定不同的经标记的核酸的数目可以包括序列特异性标记的流式细胞仪分选。确定不同的经标记的核酸的数目可以包括检测附接至珠的经标记的核酸。检测附接至珠的经标记的核酸可以包括荧光检测。

[0405] 确定不同的经标记的核酸的数目可以包括通过靶特异性探针和经标记的核酸或靶特异性的经标记的标记探针之间的荧光共振能量转移(FRET)，对多个经标记的核酸进行计数。

[0406] 经标记的核酸的检测

[0407] 本文公开的方法可以进一步包括对经标记的核酸和/或经标记的扩增子的检测。对经标记的核酸和/或经标记的扩增子的检测可以包括经标记的核酸与表面例如固体支持物的杂交。该方法可以进一步包括靶序列与核酸结合蛋白的免疫沉淀反应。经标记的核酸和/或经标记的扩增子的检测可使得能够或辅助确定不同的经标记的核酸的数目。

[0408] 在一些情况下，该方法进一步包括将经标记的核酸和/或经标记的扩增子与可检测标记接触，以产生与可检测标记偶联的经标记的核酸。本文公开的方法可以进一步包括检测与可检测标记偶联的经标记的核酸。对经标记的核酸或其任意产物(例如，经标记的扩增子、与可检测标记偶联的经标记的核酸)的检测可以包括对样品标签或分子鉴定物标记、分子、可检测标记、样品标签或分子鉴定物标记的互补体、分子的互补体或它们的任意组合的至少一部分的检测。

[0409] 对经标记的核酸或其任意产物的检测可以包括乳剂或液滴。例如，经标记的核酸或其任意产物可以在乳剂或液滴中。液滴可以是小体积的第一液体，该第一液体被不混溶的第二液体包封，例如乳剂的连续相(和/或被较大的液滴)。液滴的体积，和/或乳剂中液滴的平均体积，可以例如小于约1微升(或在约1微升和1纳升之间或在约1微升和1皮升之间)、小于约1纳升(或在约1纳升和1皮升之间)、或小于约1皮升(或在约1皮升和1飞升之间)等等。液滴(或乳剂的液滴)可以具有小于约1000、100、或10微米、或约1000至10微米等的直径(或平均直径)。液滴可以是球形或非球形的。可以产生具有平均直径为约、小于约、或大于约0.001、0.01、0.05、0.1、1、5、10、20、30、40、50、60、70、80、100、120、130、140、150、160、180、200、300、400、或500微米的液滴。液滴可以具有为约0.001至约500、约0.01至约500、约0.1至约500、约0.1至约100、约0.01至约100、或约1至约100微米的平均直径。液滴可以是简

单液滴或化合物液滴。如在此所使用，术语乳剂可指不互溶液体的混合物(例如油和水)。油相和/或油包水乳剂允许水性液滴中反应混合物的区室化。这些乳剂可以包括连续油相中的水性液滴。在此所提供的乳剂可以是水包油乳剂，其中这些液滴是连续水相中的油滴。当将乳剂或液滴用于分离(例如，空间上分离)单细胞时，可以不使用固体支持物。因此，有待标签化与分析的核酸可以不被结合至固体支持物，并且在此类情况下；当被标签化时，细胞标记可以对应于存在于乳剂或液滴中的单细胞或细胞群体。因此，乳剂或液滴可以有效地用单细胞或多个细胞分离标签化或标记化步骤并且细胞标记可以用于鉴定来自单细胞或多个细胞的核酸。在一些实施例中，可以将液滴应用于微孔中，例如，类似于将珠应用于微孔阵列中。

[0410] 可替代地，对经标记的核酸或其任意产物的检测包括一种或多种溶液。在其他情况下，对经标记的核酸的检测包括一个或多个容器。

[0411] 对经标记的核酸或其任意产物(例如，经标记的扩增子、与可检测标记偶联的经标记的核酸)的检测可以包括检测每个经标记的核酸或其产物。例如，本文公开的方法包括对每个经标记的核酸的至少一部分进行测序，从而检测每个经标记的核酸。

[0412] 在一些情况下，对经标记的核酸和/或经标记的扩增子的检测包括电泳、光谱法、显微镜检查、化学发光法、冷发光法、荧光法、免疫荧光法、比色法或电化学发光法。例如，该方法包括对荧光染料的检测。对经标记的核酸或其任意产物的检测可以包括比色法。例如，比色法包括使用比色计或比色读取器。比色计和比色读取器的非限制性列表包括森索维逊氏比色阵列成像读取器(Sensovation's Colorimetric Array Imaging Reader，CLAIR)、ESEQuant侧流免疫测定读取器(Lateral Flow Immunoassay Reader)、SpectraMax 340PC 38、SpectraMax Plus 384、SpectraMax 190、VersaMax，VMax和EMax。

[0413] 单独使用或与用于检测经标记的核酸和/或扩增子的其他方法组合使用的另外的方法可以包括使用阵列检测器、荧光读取器、非荧光检测器、CR读取器、发光计或扫描仪。在一些情况下，检测经标记的核酸和/或经标记的扩增子包括使用阵列检测器。阵列检测器的实例包括但不限于，二极管阵列检测器、光电二极管阵列检测器、HLPC光电二极管阵列检测器、阵列检测器、锗阵列检测器、CMOS及CCD阵列检测器、门控线性CCD阵列检测器、InGaAs光电二极管阵列系统和TE冷却CCD系统。该阵列检测器可以是微阵列检测器。微阵列检测器的非限制性实例包括微电极阵列检测器、光学DNA微阵列检测平台、DNA微阵列检测器、RNA微阵列检测器和蛋白质微阵列检测器。

[0414] 在一些情况下，荧光读取器用于检测经标记的核酸和/或经标记的扩增子。该荧光读取器可以读取在生物芯片上、载玻片上或微孔板中的1、2、3、4、5种或更多种颜色的荧光微阵列或其他结构。在一些情况下，该荧光读取器是森索维逊荧光阵列成像读取器(Sensovation Fluorescence Array imaging Reader，FLAIR)。可替代地，该荧光读取器是荧光酶标仪，例如Gemini XPS荧光酶标仪、Gemini Em荧光酶标仪、基于Finstruments® Fluoroskan滤波器的荧光酶标仪、PHERAstar酶标仪、FIUOstar酶标仪、POLARstar Omega酶标仪、FLUOstar OPTIMA多模式酶标仪和POLARstar OPTIMA多模式酶标仪。荧光读取器的另外的实例包括PharosFXTM和PharosFX Plus系统。

[0415] 在一些情况下，对经标记的核酸和/或经标记的扩增子的检测包括使用酶标仪。在一些情况下，该酶标仪是xMarkTM微孔板吸光度分光光度计、iMark微孔板吸光度读取器、

EnSpire®多模式酶标仪、EnVision多标记酶标仪、VICTOR X多标记酶标仪、FlexStation、SpectraMax Paradigm、SpectraMax M5e、SpectraMax M5、SpectraMax M4、SpectraMax M3、SpectraMax M2-M2e、FilterMax F系列、Fluoroskan Ascent FL微孔板荧光计和发光计、Fluoroskan Ascent微孔板荧光计、Luminoskan Ascent微孔板发光计、Multiskan EX微孔板光度计、Muliskan FC微孔板光度计和Muliskan GO微孔板光度计。在一些情况下，酶标仪检测吸光度、荧光、发光、时间分辨荧光、光散射或其任意组合。在一些实施例中，酶标仪检测动态光散射。在一些情况下，酶标仪可检测静态光散射。在一些情况下，对经标记的核酸和/或经标记的扩增子的检测包括使用微孔板成像器。在一些情况下，微孔板成像器包括ViewLux uHTS微孔板成像器和BioRad微孔板成像系统。

[0416] 对经标记的核酸和/或其产物的检测可以包括使用发光计。发光计的实例包括但不限于，SpectraMax L、GloMax®-96微孔板发光计、GloMax®-20/20单管式发光计、带有InstinctTM软件的GloMax®-Multi+、GloMax®-MultiJr单管式多模式读取器、LUMIstar OPTIMA、LEADER HC+发光计、LEADER 450i发光计和LEADER 50i发光计。

[0417] 在一些情况下，对经标记的核酸和/或经标记的扩增子的检测包括使用扫描仪。扫描仪包括平板扫描仪，例如由佳能(Cannon)、爱普生(Epson)、HP、富士通(Fujitsu)和施乐(Xerox)提供的扫描仪。平板扫描仪的另外的实例包括FMBIO®荧光成像扫描仪(例如，FMBIO®II、III和III Plus系统)。扫描仪可以包括微孔板扫描仪，例如Arrayit ArrayPixTM微阵列微孔板扫描仪。在一些情况下，扫描仪是由伯乐公司(Bio-rad)提供的个性化分子成像仪(Personal Molecular Imager)TM(PMI)系统。

[0418] 对经标记的核酸的检测可以包括使用测量带电粒子的质荷比的分析技术，例如质谱分析。在一些实施例中，结合色谱分离技术测量带电粒子的质荷比。在一些实施例中，测序反应与带电粒子质荷比的测量结合使用。在一些实施例中，标签包括同位素。在一些实施例中，同位素的类型或比例在标签库中得以控制或操纵。

[0419] 对经标记的核酸或其任意产物的检测包括使用小粒子和/或光散射。例如，扩增的分子(例如，经标记的扩增子)附接至半抗原或直接附接至小粒子并与阵列杂交。小粒子的大小可以在纳米至微米的范围内。当光在粒子表面发生散射时，可以对粒子进行检测。

[0420] 在小粒子有色的情况下可以使用比色测定法，或者可以使用比色检测系统对半抗原进行染色。在一些情形下，平板扫描仪可以用于检测从粒子散射的光或者有色材料的显影。本文公开的方法可以进一步包括使用光吸收材料。光吸收材料可以用于阻断不需要的光散射或反射。光吸收材料可以是食用色素或其他材料。在一些情形下，对经标记的核酸或其任意产物的检测包括使经标记的核酸与离轴白光接触。

[0421] 在一些实施例中，可以同时检测来自样品中的两种或更多种不同类型的生物材料。例如，可以同时检测选自下组的来自样品中的两种或更多种不同类型的生物材料，该组由以下各项组成：DNA、RNA(例如，microRNA、mRNA等)、核苷酸、蛋白、以及糖类。例如，可以使用在此描述的方法同时检测来自样品的DNA和RNA。

[0422] 数据分析

[0423] 可以将测序数据用于计数细胞中靶核酸分子的数目。例如，细胞中靶核酸的多个拷贝可以结合至固体支持物上不同的寡核苷酸。当对多个靶核酸进行扩增以及测序时，它

们可以包括不同的分子标记。对同一个靶核酸的分子标记的数目可以指示细胞中靶核酸的拷贝的数目。当确定细胞中靶核酸的浓度时，确定靶核酸的拷贝数目可以用于去除扩增偏倚。

[0424] 测序数据可以用于对受试者基因分型。通过将靶核酸与不同细胞标记进行比较，可以确定靶核酸的拷贝数变化和/或浓度。通过将靶核酸的浓度与不同细胞标记进行比较，测序数据可以用来确定细胞基因型异质性。例如，样品的第一细胞可以包括高浓度的靶核酸，而样品的第二细胞可以不包括靶核酸，或可以包括低浓度的靶核酸，从而指示细胞样品的异质性。

[0425] 确定细胞基因型异质性可用于诊断、预后、以及确定疾病的疗程。例如，如果样品的第一细胞包括靶核酸，但样品的第二细胞不包括该靶核酸，但包括第二靶核酸，那么疗程可包括靶向第一基因型的试剂(例如，药物)和靶向第二基因型的试剂(例如，药物)。

[0426] 在一些实施例中，某些序列类型可以关联至DNA或RNA图谱。例如，T细胞受体和/或B细胞受体序列可以连接至转录图谱、微小RNA图谱、或样品(如单细胞)的基因组突变图谱。在一些实施例中，某些序列类型可以关联至抗原性或蛋白表达谱。例如，通过将抗体结合至表面(例如，包含蛋白(例如，包括T细胞受体和/或B细胞受体序列的抗体的靶蛋白)的表面)，T细胞受体和/或B细胞受体序列可以关联至抗原性或蛋白表达谱。

[0427] 在一些实施例中，存在或不存在序列(例如，病毒序列)可以关联至DNA或RNA图谱。例如，存在或不存在序列(例如，病毒序列)可以关联至转录图谱、微小RNA图谱、或样品(如单个细胞)的基因组突变图谱。

[0428] 试剂盒

[0429] 本披露提供了用于实施本披露方法的试剂盒。试剂盒可以包括以下各项中的一项或多项：微孔阵列、寡核苷酸、以及固体支持物。试剂盒可以包括用于重构和/或稀释寡核苷酸和/或固体支持物的试剂。试剂盒可以包括用于将寡核苷酸缀合至固体支持物的试剂。试剂盒可以进一步包括一种或多种另外的试剂，其中这些另外的试剂可以选自：洗涤缓冲液；对照试剂、用于扩增(例如，进行cDNA合成以及PCR)靶核酸的扩增试剂、以及用于将寡核苷酸缀合至固体支持物的缀合试剂。主题试剂盒的多个组分可以在分开的容器中，或可以组合在单一容器中。

[0430] 试剂盒可以包括用于使用试剂盒的组分以实践这些主题方法的说明书。用于实践主题方法的说明书可以记录在适合的记录媒体上。例如，说明书可以印刷在基质(如纸或塑料等)上。如此，说明书可以作为包装插入物存在于试剂盒中、存在于试剂盒或其组分的容器的标签中(即，与包装或分包装相关)等。在一些实施例中，说明书可以作为电子存储数据文件存在于合适的计算机可读存储介质(例如，CD-ROM、磁盘、闪存盘等)中。在一些实施例中，实际说明书可以不存在于试剂盒中，但是提供了从远程源(例如，经由互联网)获得说明书的手段。例如，试剂盒可以包括网址，可以在该网址上查看说明书和/或可以从该网址下载说明书。如同说明书一样，用于获得说明书的手段也记录在适合的基质上。

[0431] 进一步在此披露的是用于分析来自两个或更多个样品的两种或更多种分子的试剂盒。在此披露的试剂盒可以包括多个珠、引物以及足够处理至少约384个样品的扩增剂。这些样品中的任何一个可以包括单细胞。核酸扩增可以导致在样品中测量约1、2、5、10、15、20、25、30、35、40、45、50、55、60、65、70、75、80、85、90、95、100、200、300、400、500、600、700、

800、900或1000个目标核酸。核酸扩增可以导致在样品中测量约1000个目标核酸。核酸扩增可以导致在样品中测量约100个目标核酸。核酸扩增可以导致在单细胞中测量约5%、10%、15%、20%、25%、30%、35%、40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、95%、100%的总核酸。核酸扩增可以导致在单细胞中对所有核酸序列的总体测量。核酸扩增可以导致通过测序在单细胞中测量目标核酸序列。核酸扩增可以导致通过阵列在单细胞中测量目标核酸序列。

[0432] 扩增剂可以包括固定的引物面板。扩增剂可以包括至少一对定制引物。扩增剂可以包括至少一对对照引物。扩增剂可以包括至少一对管家基因引物。合适的扩增剂可以包括PCR预混合液。试剂盒可以进一步包括用于引物设计和优化的说明书。试剂盒可以进一步包括微孔板，其中该微孔板可以包括至少一个孔，其中分布有不多于一个珠。试剂盒可以进一步包括一个或多个另外的容器。该一个或多个另外的容器可以包括一个或多个另外的多个样品标签。在一个或多个另外的容器中的一个或多个另外的多个样品标签不同于在第一容器中的第一多个样品标签。该一个或多个另外的容器可以包括一个或多个另外的分子鉴定物标记。该一个或多个另外的容器的一个或多个另外的分子鉴定物标记与第二容器的一个或多个另外的分子鉴定物标记相同。

[0433] 在此披露的方法和试剂盒可以包括使用一个或多个移液器吸头和/或容器(例如，管、小瓶、多孔板、微孔板、埃彭道夫管、载玻片、珠)。在一些情况下，这些移液器吸头是低结合性移液器吸头。可替代地，或另外地，这些容器可以是低结合性容器。低结合性移液器吸头和低结合性容器可具有降低的与基于硅树脂的吸头和非低结合性容器相关的浸出和/或后续样品降解。与非低结合性移液器吸头和容器相比，低结合性移液器吸头和低结合性容器可具有降低的样品结合。低结合性吸头的实例包括但不限于，Corning® DeckWorksTM低结合性吸头和Avant Premium低结合性带刻度吸头。低结合性容器的非限制性列表包括Corning® Costar®低结合性微量离心管和Cosmobrand低结合性PCR管以及微量离心管。

[0434] 任何在此披露的试剂盒可以进一步包括软件。例如，试剂盒可以包括用于分析序列(如条码或靶序列)的软件。例如，试剂盒可以包括用于分析序列(如条码或靶序列)的软件，以计数唯一性靶分子(例如，来自单细胞的唯一性靶分子)。例如，试剂盒可以包括用于分析序列(如条码或靶序列)的软件，以计数唯一性靶分子(例如，来自基因(例如，来自单细胞的基因)的唯一性靶分子)。

[0435] 微孔和微孔阵列

[0436] 在一些情况下，本披露的方法提供用于使包括缀合的寡核苷酸的固体支持物与细胞接触。接触步骤可以在表面上进行。示例性表面可以包括微孔、管、烧瓶、以及芯片。在一些情况下，表面包括微孔。在一些情况下，微孔是微孔阵列的一部分。

[0437] 微孔阵列的微孔可以具有使每个微孔能够包含至少1、2、3、4、5、6、7、8、9、或10个或更多个细胞的尺寸与形状。微孔可以具有使每个微孔能够包含至多1、2、3、4、5、6、7、8、9、或10个或更多个细胞的尺寸与形状。微孔阵列的微孔可以具有使每个微孔能够包含至少1、2、3、4、5、6、7、8、9、或10个或更多个固体支持物的尺寸与形状。微孔可以具有使每个微孔能够包含至多1、2、3、4、5、6、7、8、9、或10个或更多个固体支持物的尺寸与形状。微孔可以包括至多一个细胞以及一个固体支持物。微孔可以包括至多一个细胞以及两个固体支持物。微孔可以包括至少一个细胞以及至多一个固体支持物。微孔可以包括至少一个细胞以及至多

两个固体支持物。

[0438] 微孔阵列中的微孔可以是水平布置的。微孔可以是垂直布置的。微孔可以是等间距或接近等间距布置的。微孔阵列可以具有与一个或多个微孔相关的标志物。例如，微孔阵列的微孔可以被分成多个组，每个组由规定数目的微孔构成。这些组可以提供在基质的主要表面上。可以提供标志物，以使得各组的位置可以被确定。标志物可以通过肉眼检测出。标志物可以是需要光学器件可见的标志物(例如，荧光标记、发射标志、UV标志)。

[0439] 微孔阵列可以包含至少约96、384、1000、5000、10000、15000、100000、150000、500000、1000000、或5000000个或更多个微孔。微孔阵列可以包含至多约96、384、1000、5000、10000、15000、100000、150000、500000、1000000、或5000000个或更多个微孔。

[0440] 微孔的形状可以是圆柱形的。微孔的形状可以是非圆柱形的，如由多个面构成的多面体(例如，平行六面体、六角柱、或八角柱)、倒锥形、倒金字塔形(倒三棱锥、倒四角锥、倒五棱锥、倒六角锥、或具有七个或更多个角的倒多角锥)。微孔可以包括组合这些形状中的两种或更多种的形状。例如，它可以是部分圆柱形的，其余部分具有倒锥形的形状。微孔的形状可以是切去倒锥形或倒金字塔形顶部的一部分的形状。微孔的口可以在微孔的顶部或微孔的底部。微孔的底部可以是平的，但曲面(例如，凸出或凹陷)也是可能的。可以考虑有待存储于微孔中的细胞类型和/或固体底物(例如，形状、尺寸)以确定微孔的形状和尺寸。

[0441] 微孔的直径可以指最大的圆，该最大的圆可以内切在微孔的平面形状中。微孔的直径可以是有待包含在微孔中的细胞和/或固体支持物的直径的至少约0.1、0.5、1、2、或3倍或更多倍。微孔的直径可以是有待包含在微孔中的细胞和/或固体支持物的直径的至多约0.1、0.5、1、2、或3倍或更多倍。微孔的直径可以是固体支持物的直径的至少约10％、20％、30％、40％、或50％或更多。微孔的直径可以是固体支持物的直径的至多约10％、20％、30％、40％、或50％或更多。微孔的直径可以是至少约5、10、15、20、25、30、35、40、45、或50微米或更多微米。微孔的直径可以是至多约5、10、15、20、25、30、35、40、45、或50微米或更多微米。微孔的直径为约25微米。在一些情况下，微孔的直径为约30微米。在一些情况下，微孔的直径为约28微米。

[0442] 微孔体积与固体支持物体积之间的差异可以至少是约1x $10^{(-14)}m^3$、1.5x $10^{(-14)}m^3$、1.7x $10^{(-14)}m^3$、2.0x $10^{(-14)}m^3$、2.5x $10^{(-14)}m^3$、或3.0x $10^{(-14)}m^3$或更多。微孔体积与固体支持物体积之间的差异可以至多是约1x $10^{(-14)}m^3$、1.5x $10^{(-14)}m^3$、1.7x $10^{(-14)}m^3$、2.0x $10^{(-14)}m^3$、2.5x $10^{(-14)}m^3$、或3.0x $10^{(-14)}m^3$或更多。微孔体积与固体支持物体积之间的差异可以至少是约1x $10^{(-11)}$L、1.5x $10^{(-11)}$L、1.7x $10^{(-11)}$L、2.0x $10^{(-11)}$L、2.5x $10^{(-11)}$L、或3.0x $10^{(-11)}$L或更多。微孔体积与固体支持物体积之间的差异可以至多是约1x $10^{(-11)}$L、1.5x $10^{(-11)}$L、1.7x $10^{(-11)}$L、2.0x $10^{(-11)}$L、2.5x $10^{(-11)}$L、或3.0x $10^{(-11)}$L或更多。图7示出关于微孔、固体支持物的体积，以及微孔与固体支持物体积之间差异的示例性统计。

[0443] 微孔的深度可以是有待包含在微孔中的细胞和/或固体支持物的直径的至少约0.1、0.5、1、2、3、4、或5倍或更多倍。微孔的深度可以是有待包含在微孔中的细胞和/或固体支持物的直径的至多约0.1、0.5、1、2、3、4、或5倍或更多倍。微孔的深度可以是固体支持物的深度的至少约10％、20％、30％、40％、或50％或更多。微孔的深度可以是固体支持物的深度的至多约10％、20％、30％、40％、或50％或更多。微孔的深度可以是至少约5、10、15、20、

25、30、35、40、45、或50微米或更多微米。微孔的深度可以是至多约5、10、15、20、25、30、35、40、45、或50微米或更多微米。微孔的深度可以是约30微米。微孔的深度可以是约28微米。微孔可以是平的或基本上平的。

[0444] 微孔阵列可以包括这些孔之间的间距。这些孔之间的间距可以是至少约5、10、25、20、25、30、35、40、45、或50微米或更多微米。这些孔之间的间距可以是至多约5、10、25、20、25、30、35、40、45、或50微米或更多微米。这些孔之间的间距可以是约15微米。这些孔之间的间距可以是约25微米。

[0445] 可以在微孔内壁的任何位置处凹下和隆起的高度方面存在差异。通过在孔内壁的一部分上形成凹下和隆起，该孔已经处理光滑，可以将官能度添加至孔中。可以通过蚀刻使得微孔内壁光滑。可以适当地选择蚀刻装置中的真空度、蚀刻气体的类型、蚀刻步骤等。例如，可以通过湿蚀刻或通过将热氧化步骤与氧化膜蚀刻组合使微孔内壁变得光滑。可以将微孔的内壁官能化(例如，用寡核苷酸、反应基、官能团进行官能化)。

[0446] 微孔阵列可以由以下各项制成：硅、金属(例如，铝、不锈钢、铜、镍、铬、以及钛)、PDMS(弹性体)、玻璃、聚丙烯、琼脂糖、明胶、普流尼克(pluoronic)(例如，普流尼克F127)、塑料(例如，天然亲水性的塑料(如PMMA))、塑料(例如，PP、COP、COC)以及疏水性但是可以处理成为亲水性的弹性体(例如，PDMS)、水凝胶(例如，聚丙烯酰胺、藻酸盐)或树脂(例如，聚酰亚胺、聚乙烯、氯乙烯、聚丙烯、聚碳酸酯、丙烯酸以及聚对苯二甲酸乙二酯)。微孔阵列可以由疏水性材料制成。微孔阵列可以由疏水性但被涂覆(例如，通过氧等离子体处理)成为亲水性的材料制成。微孔阵列可以由亲水性但被涂覆成为疏水性的材料制成。

[0447] 可以组装微孔阵列。微孔阵列组装可以包括获得具有图式的硅晶片(例如，用SU8光刻胶制成的形成图式化柱)，并且将其用PDMS材料孵育，以经由软光刻(例如，在80C下持续几个小时)生成孔的阵列。例如，未固化的PDMS可以是液体。未固化的PDMS可以填充柱之间的间隙。当通过加热将PDMS固化时，它可以变为固体，由此生成孔的阵列。可以将光学胶粘剂(例如，NOA81/NOA63)施用至PDMS材料(例如，使用UV光)，以生成柱的阵列(例如，多个阵列)。施用可以进行持续至少约1秒、2秒、3秒、4秒、5秒、6秒、7秒、8秒、9秒、10秒或1分钟、2分钟、3分钟、4分钟、或5或更多分钟。可以将包含琼脂糖的层施用至光学胶粘剂。琼脂糖层可以是至少约1％、2％、3％、4％、5％、6％、7％、8％、9％、10％或更多琼脂糖。琼脂糖层可以是至多约1％、2％、3％、4％、5％、6％、7％、8％、9％、10％或更多琼脂糖。琼脂糖层可以是约5％琼脂糖。琼脂糖层可以设置在聚胶膜(Gelbond film)、或任何亲水性底物上，琼脂糖可附着至该底物。对光学表面上琼脂糖层的孵育可以是至少约1、2、3、4、5、6、7、8、9、10分钟或更多分钟。对光学表面上琼脂糖层的孵育可以是至多约1、2、3、4、5、6、7、8、9、10分钟或更多分钟。

[0448] 在一些情况下，本披露的方法可以使用不包括微孔的表面。表面可以是玻璃、塑料、金属。可以用固体支持物、细胞外基质、聚合物涂覆表面。表面可以不包括孔。表面可以包括空间排布成限制分子扩散的固体支持物。本披露用于捕获细胞和/或细胞内含物的方法可以发生在平表面上。本披露用于捕获细胞和/或细胞内含物的方法可以发生在悬浮液中。

[0449] 细胞和样品

[0450] 本披露的细胞可以是来自动物(例如，人类、大鼠、猪、马、母牛、狗、小鼠)的细胞。

在一些情况下，该细胞是人类细胞。细胞可以是人类胎儿细胞。人类胎儿细胞可以获得自怀有胎儿的母亲。细胞可以是来自怀孕母亲的细胞。细胞可以是来自脊椎动物、无脊椎动物、真菌、古细菌、或细菌的细胞。细胞可以来自多细胞组织(例如，器官(例如，脑、肝、肺、肾、前列腺、卵巢、脾、淋巴结、甲状腺、胰脏、心、骨骼肌、肠、喉、食道、以及胃)、胚泡)。细胞可以是来自细胞培养物的细胞。细胞可以是HeLa细胞、K562细胞、Ramos细胞、杂交瘤细胞、干细胞、未分化细胞、分化细胞、循环细胞、CHO细胞、3T3细胞、以及类似物。

[0451] 在一些情况下，该细胞是癌性细胞。癌细胞的非限制性实例可以包括前列腺癌细胞、乳癌细胞、结肠癌细胞、肺癌细胞、脑癌细胞、以及卵巢癌细胞。在一些情况下，该细胞来自癌(例如，循环肿瘤细胞)。癌的非限制性实例可以包括腺瘤(adenoma)、腺癌(adenocarcinoma)、鳞状细胞癌、基底细胞癌、小细胞癌、大细胞未分化癌、软骨肉瘤、以及纤维肉瘤。

[0452] 在一些情况下，该细胞是稀有细胞。稀有细胞可以是循环肿瘤细胞(CTC)、循环上皮细胞(CEC)、循环干细胞(CSC)、干细胞、未分化干细胞、癌干细胞、骨髓细胞、祖细胞、泡沫细胞、胎儿细胞、间充质细胞、循环内皮细胞、循环子宫内膜细胞、滋养细胞、免疫系统细胞(宿主或移植物)、结缔组织细胞、细菌、真菌、或病原体(例如，细菌或原生动物)、微粒、细胞片段、蛋白质和核酸、细胞器、其他细胞成分(例如，线粒体和核)、以及病毒。

[0453] 在一些情况下，该细胞来自肿瘤。在一些情况下，该肿瘤是良性或恶性的。该肿瘤细胞可以包括转移细胞。在一些情况下，该细胞来自包括多个不同细胞类型(例如，不同基因型)的实体组织。

[0454] 该细胞可以包括病毒、细菌、真菌和寄生物。病毒可以包括但不限于：DNA或RNA动物病毒(例如，微小核糖核酸病毒科(例如，脊髓灰质炎病毒))、呼肠孤病毒科(例如，轮状病毒)、披膜病毒科(例如，脑炎病毒、黄热病病毒、风疹病毒)、正粘病毒科(例如，流感病毒)、副粘病毒科(例如，呼吸道合胞体病毒、麻疹病毒、腮腺炎病毒、副流感病毒)、弹状病毒科(例如，狂犬病病毒)、冠状病毒科、布尼亚病毒科、黄病毒科、丝状病毒科、砂粒病毒科、本雅病毒科以及逆转录病毒科(例如，人类嗜T细胞病毒(HTLV)、人体免疫缺损病毒(HIV)、乳多空病毒科(例如，乳头瘤病毒)、腺病毒科(例如，腺病毒)、疱疹病毒科(例如，单纯性疱疹病毒)、以及痘病毒科(例如，天花病毒))。

[0455] 可以用于本披露方法中的示例性细菌可以包括Actinomedurae、衣氏放线菌、炭疽杆菌、蜡样芽胞杆菌、肉毒梭菌、艰难梭菌、产气荚膜梭菌、破伤风梭菌、棒状杆菌、粪肠球菌、单核细胞增生李斯特菌、诺卡氏菌属、痤疮丙酸杆菌、金黄色葡萄球菌、表皮葡萄球菌(Staphylococcus epiderm)、变异链球菌、肺炎链球菌以及类似物。革兰氏阴性细菌包括但不局限于：猫阿菲波菌(Afipia felis)、类杆菌(Bacteriodes)、杆状巴尔通体、百日咳杆菌(Bortadella pertussis)、伯氏疏螺旋体、回归热疏螺旋体、布鲁氏菌(Brucella)、肉芽肿鞘杆菌、弯曲杆菌、大肠杆菌、土拉热弗朗西丝菌、阴道加德纳菌、埃及嗜血杆菌(Haemophilius aegyptius)、杜克雷嗜血杆菌(Haemophilius ducreyi)、流感嗜血杆菌(Haemophilius influenziae)、幽门螺杆菌、嗜肺军团菌、问号钩端螺旋体、脑膜炎奈瑟菌、牙龈卟啉单胞菌、斯氏普罗维登斯菌(Providencia sturti)、铜绿假单胞菌、肠炎沙门菌(Salmonella enteridis)、伤寒沙门菌、粘质沙雷菌、波伊德志贺氏菌、念珠状链杆菌、化脓链球菌、梅毒螺旋体、霍乱弧菌、小肠结肠炎耶尔森菌、鼠疫耶尔森菌、以及类似物。其他细

菌可以包括鸟分枝杆菌(Myobacterium avium)、麻风分枝杆菌(Myobacterium leprae)、结核分枝杆菌(Myobacterium tuberculosis)、Bartonella henseiae、鹦鹉热衣原体、沙眼衣原体、伯氏考克斯体、肺炎支原体、螨立克次体、普氏立克次氏体、立氏立克次体、姜虫病立克次氏体、斑疹伤寒立克次氏体、尿素分解脲原体、肺炎双球菌、查菲埃立克体(Ehrlichia chafensis)、屎肠球菌、脑膜炎球菌、以及类似物。

[0456] 用于本披露的方法中的示例性真菌可以包括但不限于:曲霉属、假丝酵母属(Candidae)、白色念珠菌、粗球孢子菌、隐球菌(Cryptococci)及其组合。

[0457] 用于本披露的方法中的示例性寄生物可以包括但不限于:结肠小袋纤毛虫、微小隐孢子虫、卡晏环孢子虫(Cyclospora cayatanensis)、脑炎微孢子虫属(Encephalitozoa)、溶组织内阿米巴(Entamoeba histolytica)、比氏肠孢虫(Enterocytozoon bieneusi)、兰伯贾第虫、利什曼原虫(Leishmaniae)、疟原虫(Plasmodii)、鼠弓形体、锥体虫(Trypanosomae)、梯形阿米巴(trapezoidal amoeba)、蠕虫(worms)(例如,蠕虫(helminthes))、尤其是寄生蠕虫(包括但不限于线虫纲(蛔虫(例如,鞭虫、钩虫、蛲虫、蛔虫(ascarids)、丝状虫(filarids)等)))、绦虫纲(例如,绦虫)。

[0458] 本披露的样品可以是来自动物(例如,人类、大鼠、猪、马、母牛、狗、小鼠)的样品。在一些情况下,样品是人类样品。样品可以是人类胎儿样品。人类胎儿样品可以获得自怀有胎儿的母亲。样品可以是来自怀孕母亲的样品。样品可以是来自脊椎动物、无脊椎动物、真菌、古细菌、或细菌的样品。样品可以来自多细胞组织(例如,器官(例如,脑、肝、肺、肾、前列腺、卵巢、脾、淋巴结、甲状腺、胰脏、心、骨骼肌、肠、喉、食道、以及胃)、胚泡)。样品可以是来自细胞培养物的细胞。

[0459] 样品可以包括多个细胞。样品可以包括多个相同类型的细胞。样品可以包括多个不同类型的细胞。样品可以包括处在细胞周期和/或分化路径中相同点处的多个细胞。样品可以包括处在细胞周期和/或分化路径中不同点处的多个细胞。样品可以包括多个样品。

[0460] 该多个样品可以包括至少5、10、20、30、40、50、60、70、80、90或100或更多个样品。该多个样品可以包括至少约100、200、300、400、500、600、700、800、900或1000或更多个样品。该多个样品可以包括至少约1000、2000、3000、4000、5000、6000、7000、8000个样品、9000、或10,000个样品,或100,000个样品,或1,000,000个或更多个样品。该多个样品可以包括至少约10,000个样品。

[0461] 第一样品中的一个或多个核酸可以不同于第二样品中的一个或多个核酸。第一样品中的一个或多个核酸可以不同于多个样品中的一个或多个核酸。该一个或多个核酸可以包括至少为约1个核苷酸、2个核苷酸、5个核苷酸、10个核苷酸、20个核苷酸、50个核苷酸、100个核苷酸、200个核苷酸、300个核苷酸、500个核苷酸、1000个核苷酸、2000个核苷酸、3000个核苷酸、4000个核苷酸、5000个核苷酸、10,000个核苷酸、100,000个核苷酸、1,000,000个核苷酸的长度。

[0462] 第一样品可以包括一个或多个细胞并且第二样品可以包括一个或多个细胞。第一样品的一个或多个细胞可以具有与第二样品的一个或多个细胞相同的细胞类型。第一样品的一个或多个细胞可以与该多个样品的一个或多个不同细胞具有不同的细胞类型。所述细胞类型可以是软骨细胞、破骨细胞、脂细胞、成肌细胞、干细胞、内皮细胞或平滑肌细胞。所述细胞类型可以是免疫细胞类型。所述免疫细胞类型可以是T细胞、B细胞、血小板、树突细

胞、中性粒细胞、巨噬细胞或单核细胞。

[0463] 该多个样品可以包括一种或多种恶性细胞。该一种或多种恶性细胞可以衍生自肿瘤、肉瘤或白血病。

[0464] 该多个样品可以包括至少一种体液。该体液可以包括血液、尿液、淋巴液、唾液。该多个样品可以包括至少一种血液样品。

[0465] 该多个样品可以包括来自一种或多种生物组织的至少一种细胞。所述一种或多种生物组织可以是骨骼、心脏、胸腺、动脉、血管、肺、肌肉、胃、肠、肝、胰腺、脾、肾、胆囊、甲状腺、肾上腺、乳腺、卵巢、前列腺、睾丸、皮肤、脂肪、眼或脑。

[0466] 该生物组织可以包括受感染的组织、患病组织、恶性组织、钙化组织或健康组织。

[0467] 该多个样品可以来自一个或多个来源。该多个样品可以来自两个或更多个来源。该多个样品可以来自一个或多个受试者。该多个样品可以来自两个或更多个受试者。该多个样品可以来自同一个受试者。该一个或更多个受试者可以来自同一物种。该一个或更多个受试者可以来自不同物种。该一个或更多个受试者可以是健康的。该一个或更多个受试者可以受疾病、障碍或病症的影响。该多个样品可以包括选自下组来源的细胞:哺乳动物、细菌、病毒、真菌或植物。该一个或多个样品可以来自人类、马、母牛、鸡、猪、大鼠、小鼠、猴、兔、豚鼠、绵羊、山羊、狗、猫、鸟、鱼、蛙以及果蝇。

[0468] 可以并行获得该多个样品。可以同时获得该多个样品。可以顺序获得该多个样品。可以在获得一个或多个不同样品为期多年、100年、10年、5年、4年、3年、2年或1年的过程中获得多个样品。可以在获得一个或多个不同样品的约1年之内获得一个或多个样品。可以在获得一个或多个不同样品的12个月、11个月、10个月、9个月、8个月、7个月、6个月、4个月、3个月、2个月或1个月之内获得一个或多个样品。可以在获得一个或多个不同样品的30天、28天、26天、24天、21天、20天、18天、17天、16天、15天、14天、13天、12天、11天、10天、9天、8天、7天、6天、5天、4天、3天、2天或1天之内获得一个或多个样品。可以在获得一个或多个不同样品的约24小时、22小时、20小时、18小时、16小时、14小时、12小时、10小时、8小时、6小时、4小时、2小时或1小时之内获得一个或多个样品。可以在获得一个或多个不同样品的约60秒、45秒、30秒、20秒、10秒、5秒、2秒或1秒之内获得一个或多个样品。可以在获得一个或多个不同样品的小于1秒之内获得一个或多个样品。

[0469] 靶分子

[0470] 本文公开的方法和试剂盒可用于对分子进行随机标记。此类分子包括但不限于多核苷酸和多肽。如本文所使用的,术语“多核苷酸”和“核酸分子”是指任何长度的核苷酸的聚合形式,该核苷酸可以是核糖核苷酸、脱氧核糖核苷酸、锁定核酸(LNA)或肽核酸(PNA),其包含嘌呤和嘧啶碱基,或其他天然的、化学或生物化学修饰的、非天然的或衍生的核苷酸碱基。“多核苷酸”或“核酸分子”可由单个核苷酸或碱基对组成。可替代地,“多核苷酸”或“核酸分子”包含两个或更多个核苷酸或碱基对。例如,“多核苷酸”或“核酸分子”包含至少约2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、30、40、50、60、70、80、90、100、200、300、400、500、600、700、800、900或1000个核苷酸或碱基对。在另一个实例中,该多核苷酸包含至少约1500、2000、2500、3000、3500、4000、4500、5000、5500、6000、6500、7000、7500、8000、8500、9000、9500或10000个核苷酸或碱基对。多核苷酸的主链可包含糖和磷酸基团(正如通常可在RNA或DNA中所见的)或者经修饰或取代的糖或磷酸基团。多核苷酸可包含经

修饰的核苷酸，如甲基化的核苷酸和核苷酸类似物。核苷酸的序列可以被非核苷酸成分中断。因此，术语“核苷”、“核苷酸”、“脱氧核苷”和“脱氧核苷酸”通常包括类似物，例如本文所述的那些类似物。这些类似物是那些与天然存在的核苷或核苷酸具有一些共同的结构特征的分子，从而使得当其并入核酸或寡聚核苷序列时，它们能够与溶液中的天然存在的核酸序列杂交。通常情况下，这些类似物通过对碱基、核糖或磷酸二酯部分进行替换和/或修饰从天然存在的核苷和核苷酸衍生得到。可以定制改变，以根据需要使杂合体形成稳定或不稳定化，或者增强与互补核酸序列杂交的特异性。在一些情形下，这些分子是DNA、RNA或DNA-RNA杂合体。这些分子可以是单链或双链的。在一些情形下，这些分子是RNA分子，如mRNA、rRNA、tRNA、ncRNA、lncRNA、siRNA、microRNA或miRNA。这些RNA分子可以是聚腺苷酸化的。可替代地，这些mRNA分子不是聚腺苷酸化的。可替代地，这些分子是DNA分子。这些DNA分子可以是基因组DNA。这些DNA分子可包含外显子、内含子、非翻译区或其任意组合。在一些情况下，这些分子是分子的面板。

[0471] 本文公开的方法和试剂盒可用于随机地标记单独出现的相同或几乎相同的分子和/或不同分子。在一些情形下，本文所公开的方法和试剂盒可用于随机地标记相同或几乎相同的分子(例如，分子包含相同或几乎相同的序列)。例如，待标记的分子包含至少约80％、85％、90％、95％、96％、97％、98％、99％或100％的序列一致性。几乎相同的分子可相差少于约100、90，80、70、60、50、40、30、25、20、25、10、9、8、7、6、5、4、3、2或1个核苷酸或碱基对。多个样品中一个或多个样品中的多个核酸可以包括两个或更多个相同序列。多个样品中一个或多个中的至少约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％的总核酸可以包括相同序列。多个样品中一个或多个样品中的多个核酸可以包括至少两个不同序列。多个样品中一个或多个中的至少约5％、10％、15％、20％、25％、30％、35％、40％、45％、50％、55％、60％、65％、70％、75％、80％、81％、82％、83％、84％、85％、86％、87％、88％、89％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％、100％的总核酸可以包括至少两个不同序列。在一些情形下，待标记的分子是彼此的变体。例如，待标记的分子可包含单核苷酸多态性或其他类型的突变。在另一个实例中，待标记的分子是剪接变体。在一些情形下，随机地标记至少一个分子。在其他情形下，随机地标记至少2、3、4、5、6、7、8、9或10个相同或几乎相同的分子。可替代地，随机地标记至少20、30、40、50、60、70、80、90、100、200、300、400、500、600、700、800、900或1000个相同或几乎相同的分子。在其他情况下，随机地标记至少1500；2,000；2500；3,000；3500；4,000；4500；5,000；6,000；7,000；8,000；9,000；或10000个相同或几乎相同的分子。在其他情况下；随机地标记至少15,000；20,000；25,000；30,000；35,000；40,000；45,000；50,000；60,000；70,000；80,000；90,000；或100,000个相同或几乎相同的分子。

[0472] 在其他情形下，本文公开的方法和试剂盒可用于随机地标记不同的分子。例如，待标记的分子包括小于75％、70％、65％、60％、55％、50％、45％、40％、35％、30％、25％、20％、15％、10％、5％、4％、3％、2％、1％序列一致性。不同的分子可相差至少约1、2、3、4、5、6、7、8、9、10、15、20、25、30、40、50、60、70、80、90、100个或更多个核苷酸或碱基对。在一些情形下，随机地标记至少一个分子。在其他情形下，随机地标记至少2、3、4、5、6、7、8、9或10个不同的分子。可替代地，随机地标记至少20、30、40、50、60、70、80、90、100、200、300、400、

500、600、700、800、900或1000个不同的分子。在其他情况下，随机地标记至少1500；2,000；2500；3,000；3500；4,000；4500；5,000；6,000；7,000；8,000；9,000；或10000个不同的分子。在其他情况下；随机地标记至少15,000；20,000；25,000；30,000；35,000；40,000；45,000；50,000；60,000；70,000；80,000；90,000；或100,000不同的分子。

[0473] 待标记的不同分子能以不同浓度或量存在于样品中。例如，样品中一种分子的浓度或量大于另一种分子的浓度或量。在一些情形下，样品中至少一种分子的浓度或量是该样品中至少另一种分子的浓度或量的至少约1.5、2、3、4、5、6、7、8、9、10、11、12、13、14、15、20、25、30、35、40、45、50、60、70、80、90或100倍或更多倍。在一些情形下，样品中至少一种分子的浓度或量是该样品中至少另一种分子的浓度或量的至少约1000倍或更多倍。在另一个实例中，样品中一种分子的浓度或量小于另一种分子的浓度或量。样品中至少一种分子的浓度或量可以比该样品中至少另一种分子的浓度或量低至少约1.5、2、3、4、5、6、7、8、9、10、11、12、13、14、15、20、25、30、35、40、45、50、60、70、80、90或100倍或更多倍。样品中至少一种分子的浓度或量可以比该样品中至少另一种分子的浓度或量低至少约1000倍或更多倍。

[0474] 在一些情形下，待标记的分子处于一个或多个样品中。待标记的分子可处于两个或更多个样品中。该两个或更个样品可包含不同量或浓度的待标记的分子。在一些情形下，一个样品中一种分子的浓度或量可大于不同样品中相同分子的浓度或量。例如，血液样品可能含有与尿样品相比更高的量的特定分子。可替代地，将一个样品分成两个或更多个子样品。子样品可含有不同量或浓度的同一种分子。一个样品中至少一种分子的浓度或量可以是另一样品中相同分子的浓度或量的至少约1.5、2、3、4、5、6、7、8、9、10、11、12、13、14、15、20、25、30、35、40、45、50、60、70、80、90或100倍或更多倍。可替代地，一个样品中一种分子的浓度或量可小于不同样品中相同分子的浓度或量。例如，心脏组织样品可能含有与肺组织样品相比更高的量的特定分子。一个样品中至少一种分子的浓度或量可以比另一样品中相同分子的浓度或量低至少约1.5、2、3、4、5、6、7、8、9、10、11、12、13、14、15、20、25、30、35、40、45、50、60、70、80、90或100倍或更多倍。在一些情形下，两个或更多个不同样品中分子的不同浓度或量被称为样品偏差。

[0475] 在此披露的方法和试剂盒可以用于分析来自两个或更多个样品的两种或更多种分子。两种或更多种分子可以包括两种或更多种多肽。该方法可以包括确定两种或更多种经标记的多肽的身份。确定两种或更多种经标记的多肽的身份可以包括质谱法。该方法可以进一步包括将第一样品的经标记的多肽与第二样品的经标记的多肽组合。可以在确定不同的经标记的多肽的数目之前组合经标记的多肽。该方法可以进一步包括将第一样品标签化多肽和第二样品标签化多肽组合。可以在与该多个分子鉴定物标记接触之前组合第一样品标签化多肽和第二样品标签化多肽。确定不同的经标记的多肽的数目可以包括检测经标记的多肽的至少一部分。检测经标记的多肽的至少一部分可以包括检测样品标签、分子鉴定物标记、多肽、或它们的组合的至少一部分。

[0476] 如本文所使用的，术语“多肽”是指包含至少一个肽的分子。在一些情形下，该多肽由单肽组成。可替代地，该多肽包含两个或更多个肽。例如，该多肽包含至少约2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、30、40、50、60、70、80、90、100、200、300、400、500、600、700、800、900或1000个肽。多肽的实例包括但不限于，氨基酸、蛋白质、肽、激素、寡糖、脂质、糖脂、磷脂、抗体、酶、激酶、受体、转录因子和配体。

[0477] 受试者

[0478] 在此披露的方法和试剂盒可以包括使用来自一个或多个受试者的细胞或样品。受试者可以是人类或非人类受试者。受试者可以是活的。受试者可以是死的。受试者可以是由护理员(例如,医学专业人士)照看的人类。受试者可以疑似患有疾病。受试者可以患有疾病。受试者可以具有疾病的症状。受试者可以是提供一个或多个样品的受试者。受试者可以是哺乳动物、爬行动物、两栖动物、和/或鸟。受试者可以是非人类灵长动物。

[0479] 酶

[0480] 在此披露的方法和试剂盒可以包括一种或多种酶。酶的实例包括但不限于连接酶、逆转录酶、聚合酶和限制性核酸酶。在一些情形下,寡核苷酸标签与分子的附接包括使用一种或多种连接酶。连接酶的实例包括但不限于DNA连接酶(如DNA连接酶I、DNA连接酶III、DNA连接酶IV以及T4DNA连接酶)和RNA连接酶(例如T4RNA连接酶I和T4RNA连接酶II)。

[0481] 在此披露的方法和试剂盒可以进一步包括使用一种或多种逆转录酶。在一些情形下,该逆转录酶是HIV-1逆转录酶、M-MLV逆转录酶、AMV逆转录酶和端粒酶逆转录酶。在一些情形下,该逆转录酶是M-MLV逆转录酶。

[0482] 在一些情形下,本文公开的方法和试剂盒包括使用一种或多种聚合酶。聚合酶的实例包括但不限于DNA聚合酶和RNA聚合酶。在一些情形下,该DNA聚合酶是DNA聚合酶I、DNA聚合酶II、DNA聚合酶III全酶和DNA聚合酶IV。市售的DNA聚合酶包括但不限于:Bst 2.0DNA聚合酶、Bst 2.0WarmStart™DNA聚合酶、Bst DNA聚合酶、硫化叶菌DNA聚合酶IV、Taq DNA聚合酶、9°N™m DNA聚合酶、Deep VentR™(exo-)DNA聚合酶、Deep VentR™DNA聚合酶、Hemo KlenTaq™、LongAmp®Taq DNA聚合酶、OneTaq®DNA聚合酶、Phusion®DNA聚合酶、Q5™高保真度DNA聚合酶、Therminator™γDNA聚合酶、Therminator™DNA聚合酶、Therminator™II DNA聚合酶、Therminator™III DNA聚合酶、VentR®DNA聚合酶、VentR®(exo-)DNA聚合酶、Bsu DNA聚合酶、phi29DNA聚合酶、T4DNA聚合酶、T7DNA聚合酶、末端转移酶、Titanium®Taq聚合酶、KAPA Taq DNA聚合酶以及KAPA Taq热启动DNA聚合酶。

[0483] 可替代地,该聚合酶是RNA聚合酶,例如RNA聚合酶I、RNA聚合酶II、RNA聚合酶III、大肠杆菌聚(A)聚合酶、phi6RNA聚合酶(RdRP)、Poly(U)聚合酶、SP6RNA聚合酶和T7RNA聚合酶。

[0484] 在一些情况下,本文公开的方法和试剂盒包含一种或多种限制酶。该限制酶包括I型、II型、III型和IV型限制酶。在一些情况下,I型酶为在远离其识别序列的位置随机切割DNA的复合的、多亚基、组合性的限制性以及修饰性酶。通常,II型酶在靠近其识别序列或在其识别序列内的确定的位置切割DNA。它们可产生离散的限制性片段和独特的凝胶带型图样。III型酶也是大的、组合性的限制性以及修饰性酶。它们通常在其识别序列以外进行切割,并且可能需要在同一DNA分子内在相反方向上的两段此类序列来完成切割;它们很少进行完全消化。在一些情况下,IV型酶识别经修饰的、通常为甲基化的DNA,并且可由大肠杆菌的McrBC和Mrr系统来例示。

[0485] 另外的试剂

[0486] 在此披露的方法和试剂盒可以包括使用一种或多种试剂。该试剂的实例包括但不限于PCR试剂、连接试剂、逆转录试剂、酶试剂、杂交试剂、样品制备试剂以及用于核酸纯化

和/或分离的试剂。

[0487] 在此披露的方法和试剂盒可以包括使用一种或多种缓冲液。缓冲液的实例包括但不限于洗涤缓冲液、连接缓冲液、杂交缓冲液、扩增缓冲液和逆转录缓冲液。在一些情况下，杂交缓冲液是市售的缓冲液，如TMAC Hyb溶液、SSPE杂交溶液和ECONOTM杂交缓冲液。本文公开的缓冲液可包含一种或多种洗涤剂。

[0488] 在此披露的方法和试剂盒可以包括使用一种或多种载体。载体可提高或改善本文公开的一种或多种反应(例如，连接反应、逆转录、扩增、杂交)的效率。载体可减少或防止分子或其任意产物(例如，经标记的分子、经标记的cDNA分子、经标记的扩增子)的非特异性损失。例如，所述载体可减少经标记的分子通过吸附至表面而引起的非特异性损失。所述载体可降低分子、经标记的分子或其任意产物对表面或基质(例如，容器、埃彭道夫管(eppendorf tube)、移液管吸头)的亲和力。可替代地，所述载体可增加分子或其任意产物对表面或基质(例如，珠粒、阵列、玻璃、载玻片、芯片)的亲和力。载体可以保护分子或其任意产物不发生降解。例如，载体可以保护RNA分子或其任意产物不被核糖核酸酶降解。可替代地，载体可以保护DNA分子或其任意产物不被DNase降解。载体的实例包括但不限于核酸分子(例如DNA和/或RNA)或多肽。DNA载体的实例包括质粒、载体(vector)、聚腺苷酸化的DNA和DNA寡核苷酸。RNA载体的实例包括聚腺苷酸化的RNA、噬菌体RNA、MS2噬菌体RNA、大肠杆菌RNA、酵母RNA、酵母tRNA、哺乳动物RNA、哺乳动物tRNA、短聚腺苷酸化的合成核糖核苷酸和RNA寡核苷酸。RNA载体可以是聚腺苷酸化的RNA。可替代地，RNA载体可以是非聚腺苷酸化的RNA。在一些情况下，该载体来自细菌、酵母或病毒。例如，该载体可以是来源于细菌、酵母或病毒的核酸分子或多肽。例如，该载体是来自枯草芽孢杆菌的蛋白质。在另一个实例中，该载体是来自大肠杆菌的核酸分子。或者，该载体是来自哺乳动物(例如，人、小鼠、山羊、大鼠、牛、绵羊、猪、狗或兔)、禽类、两栖动物或爬行动物的核酸分子或肽。

[0489] 在此披露的方法和试剂盒可以包括使用一种或多种对照剂。对照剂可以包括对照寡核苷酸、钝化酶、非特异性竞争剂。可替代地，对照剂包括清晰杂交、清晰探针对照、核酸模板、掺入型对照、PCR扩增对照。PCR扩增对照可以是阳性对照。在其他情况下，PCR扩增对照是阴性对照。核酸模板对照可具有已知的浓度。对照剂可包含一个或多个标记。

[0490] 掺入型对照可以是加入到反应或样品中的模板。例如，可将掺入型模板加入到扩增反应中。可在第一个扩增循环后的任何时间将掺入型模板加入到扩增反应中。在一些情形下，在第2个、第3个、第4个、第5个、第6个、第7个、第8个、第9个、第10个、第11个、第12个、第13个、第14个、第15个、第20个、第25个、第30个、第35个、第40个、第45个或第50个扩增循环之后将掺入型模板加入到扩增反应中。可在最后一个扩增循环之前的任何时间将掺入型模板加入到扩增反应中。掺入型模板可包含一个或多个核苷酸或核酸碱基对。掺入型模板可包含DNA、RNA或其任意组合。掺入型模板可包含一个或多个标记。

[0491] 可检测标记

[0492] 本文公开的方法、试剂盒和组合物可以进一步包含可检测标记。术语“可检测标记”、“标签”或“标记”可以互换地使用并且是指附接至分子(例如，核苷酸、核苷酸聚合物、或核酸结合因子、分子条码)的任何化学部分。化学部分可以共价地附接至分子。化学部分可以非共价地附接至分子。分子条码、样品标签以及分子鉴定物标记可以进一步包括可检测标记、标签或标记。优选地，该标记是可检测的，并使核苷酸或核苷酸聚合物对于本发明

的实施者是可检测的。可与本文所公开的方法结合使用的可检测标记包括，例如，荧光标记、化学发光标记、淬灭剂、放射性标记、生物素、茋部分、金或它们的组合。可检测标记的非限制性实例包括发光分子、荧光染料、荧光淬灭剂、有色分子、放射性同位素或闪烁体。

[0493] 在一些情况下，本文公开的方法进一步包括将一个或多个可检测标记附接至分子条码、分子鉴定物标记、样品标签、经标记的核酸或其任意产物(例如，经标记的扩增子)。该方法可以包括将两个或更多种可检测标记附接至分子条码、分子鉴定物标记、样品标签或经标记的核酸。可替代地，该方法包括将至少约3、4、5、6、7、8、9、或10种可检测标记附接至分子条码、分子鉴定物标记、样品标签或经标记的核酸。在一些情况下，可检测标记是CyTM标记。该CyTM标记是Cy3标记。可替代地，或另外地，可检测标记是生物素。在一些实施例中，将可检测标记附接至探针，该探针结合至分子条码、分子鉴定物标记、样品标签或经标记的核酸。这可以例如在将核酸或经标记的核酸与阵列杂交之后发生。在一个实例中，核酸或经标记的核酸与其在阵列上的配偶体结合。在结合之后，可结合该经标记的核酸的探针与阵列上的分子结合。可以使用多个探针和标记重复这一过程，以降低由于标记与阵列的非特异结合或分子与阵列的非特异结合而产生信号的可能性。

[0494] 供体-受体对可以用作可检测标记。可以将供体或受体附接至结合核酸的探针。该探针可以是，例如，可以与核酸或经标记的核酸结合的核酸探针。可以加入相应的供体或受体以产生信号。

[0495] 在一些情形下，可检测标记是自由染料(Freedom dye)、阿力克萨荧光®染料(AlexaFluor® dye)、CyTM染料、荧光素染料或LICOR IRDyes®。在一些情况下，自由染料(Freedom dye)是荧光素(6-FAMTM，6-羧基荧光素)、MAX(NHS酯)、TYETM 563、TEX615、TYETM 665、TYE 705。可检测标记可以是阿力克萨荧光染料。阿力克萨荧光®染料的实例包括阿力克萨荧光®488(NHS酯)、阿力克萨荧光®532(NHS酯)、阿力克萨荧光®546(NHS酯)、阿力克萨荧光® 594(NHS酯)、阿力克萨荧光®647(NHS酯)、阿力克萨荧光®660(NHS酯)或阿力克萨荧光®750(NHS酯)。可替代地，可检测标记为CyTM染料。CyTM染料的实例包括但不限于Cy2、Cy3、Cy3B、Cy3.5、Cy5、Cy5.5和Cy7。在一些情况下，可检测标记是荧光素染料。荧光素染料的非限制性实例包括6-FAMTM(叠氮化物)、6-FAMTM(NHS酯)、荧光素dT、JOE(NHS酯)、TETTM和HEXTM。在一些情况下，可检测标记是LI-COR IRDyes®例如5’ IRDye®700、5’ IRDye®800、或IRDye®800CW(NHS酯)。在一些情况下，可检测标记是TYETM 563。可替代地，可检测标记是Cy3。

[0496] 可检测标记可以是罗丹明(Rhodamine)染料。罗丹明染料的实例包括但不限于，罗丹明GreenTM-X(NHS酯)、TAMRATM、TAMRATM(NHS酯)、罗丹明RedTM-X(NHS酯)、ROXTM(NHS酯)和5’TAMRATM(叠氮化物)。在其他情况下，可检测标记是维尔瑞德(WellRED)染料。维尔瑞德染料包括但不限于，维尔瑞德D4染料、维尔瑞德D3染料和维尔瑞德D2染料。在一些情况下，可检测标记是德克萨斯宏®-X(TexasRed®-X)(NHS酯)、光环®640(Lightcycler® 640)(NHS酯)或Dy 750(NHS酯)。

[0497] 在一些情况下，可检测标记包括接头分子。接头分子的实例包括但不限于，生物素、亲和素、链霉亲和素、HRP、蛋白A、蛋白G、抗体或其片段、Grb2、多组氨酸、Ni2+、FLAG标

签、myc标签。可替代地，可检测标记包括重金属、电子供体/受体、吖啶酯、染料和量热底物。在其他情况下、可检测标记包括酶，如碱性磷酸酶、过氧化物酶和萤光素酶。

[0498] 质量的变化可以被认为是一种可检测标记，正如表面等离子体共振检测的情况。本领域技术人员将容易确认可在本发明的操作中使用的本文未提及的有用的可检测标记。

[0499] 在一些情况下，可检测标记与引物一起使用。例如，通用引物采用可检测标记进行标记(例如，Cy3标记的通用引物、荧光团标记的通用引物)。可替代地，靶特异性引物采用可检测标记进行标记(例如，TYE563标记的靶标特异性引物)。在其他情况下，可检测标记与样品标签或分子鉴定物标记一起使用。例如，寡核苷酸标签采用可检测标记进行标记(例如，生物素标记的寡核苷酸标签)。在其他情况下，可检测标记与核酸模板分子一起使用。可检测标记可以用于检测经标记的分子或经标记的扩增子。可替代地，可检测标记用于检测核酸模板分子。

[0500] 在一些情况下，可检测标记附接至引物、分子条码、样品标签、分子鉴定物标记、经标记的分子、经标记的扩增子、探针、HCR探针和/或未经标记的核酸。用于将可检测标记附接至引物、寡核苷酸标签、经标记的分子、经标记的扩增子和/或未经标记的核酸的方法包括但不限于化学标记法和酶标记法。在一些情况下，可检测标记通过化学标记法进行附接。在一些实施例中，化学标记技术包含化学反应性基团。反应性基团的非限制性实例包括胺反应性琥珀酰亚胺酯(例如NHS-荧光素或NHS-罗丹明)、胺反应性异硫氰酸酯衍生物(包括FITC)和巯基反应性马来酰亚胺活化荧光剂(例如荧光素-5-马来酰亚胺)。在一些实施例中，任何上述反应性染料与另一分子的反应导致在荧光团与接头和/或试剂之间形成稳定的共价键。在一些实施例中，反应性基团为异硫氰酸酯。在一些实施例中，通过赖氨酸侧链的伯胺将标记附接至试剂。在一些实施例中，化学标记法包括NHS-酯化学法。

[0501] 可替代地，可检测标记通过酶标记法进行附接。酶标记法可以包括但不限于，生物素受体肽/生物素连接酶(AP/Bir A)、酰基载体蛋白质/磷酸泛酰巯基乙胺转移酶(ACP/PPTase)、人O6-烷基鸟嘌呤转移酶(hAGT)、Q-标签/转谷氨酰胺酶(TGase)、醛标签/甲酰甘氨酸生成酶、突变的原核脱卤素酶(HaloTagTM)和法尼基化基序/蛋白质法尼基转移酶(PFTase)方法。亲和标记法可以包括但不限于，使用二氢叶酸还原酶(DHFR)和FK506结合蛋白12的Phe36Val突变体(FKBP12(F36V))的非共价方法以及金属-螯合方法。

[0502] 交联试剂可用于将可检测标记附接至引物、寡核苷酸标签、经标记的分子、经标记的扩增子和/或未经标记的核酸。在一些情况下，交联试剂是戊二醛。戊二醛可以与胺基团反应，以通过几种途径形成交联。例如，在还原条件下，戊二醛两端的醛与胺偶合从而形成仲胺连接。

[0503] 在一些情况下，可检测标记与引物、寡核苷酸标签、经标记的分子、经标记的扩增子和/或未经标记的核酸的附接包括高碘酸盐活化以及后续的还原胺化。在一些情况下，磺基-SMCC或其他异源双功能交联剂用于将可检测标记与引物、寡核苷酸标签、经标记的分子、经标记的扩增子和/或未经标记的核酸偶联。例如，磺基-SMCC用于将酶与药物偶联。在一些实施例中，将酶在一个步骤中进行激活和纯化，然后在第二步骤中与药物偶联。在一些实施例中，将交联的方向性限定为一个特定的方向(例如，从酶上的胺到抗体上的巯基)。

[0504] 疾病/病症

[0505] 本文公开了用于对受试者的疾病或病症的状态或结果进行诊断、监测和/或预后

的方法、试剂盒和组合物。通常，该方法包括：(a)对来自两个或更多个样品的两种或更多种分子进行随机标记，以生成两种或更多种经标记的核酸；(b)检测和/或定量该两种或更多种经标记的核酸；并且(c)基于对该两种或更多种经标记的核酸的检测和/或定量，对受试者的疾病或病症的状态或结果进行诊断、监测和/或预后。该方法可以进一步包括确定治疗方案。该两个或更多个样品可以包括来自患有疾病或病症的受试者的一个或多个样品。该两个或更多个样品可以包括来自健康受试者的一个或多个样品。该两个或更多个样品可以包括来自对照的一个或多个样品。

[0506] 监测疾病或病症可以进一步包括对治疗方案进行监控。对治疗方案进行监控可以包括确定治疗方案的疗效。在一些情况下，对治疗方案进行监控包括施用、终止、添加或改变治疗方案。改变治疗方案可以包括增加或减少治疗方案的剂量、给药频率或给药方式。治疗方案可以包含一种或多种治疗药物。该治疗药物可以是抗癌药、抗病毒药、抗菌药、抗病原体药物或它们的任意组合。

[0507] 癌症

[0508] 在一些情况下，该疾病或病症是癌症。待进行随机标记的分子可以来自癌细胞或组织。在一些情况下，该癌症是肉瘤、癌、淋巴瘤或白血病。肉瘤是骨、软骨、脂肪、肌肉、血管或其他结缔组织或支持性组织的癌症。肉瘤包括但不限于，骨癌、纤维肉瘤、软骨肉瘤、尤因氏肉瘤、恶性血管内皮瘤、恶性神经鞘瘤、双侧前庭神经鞘瘤、骨肉瘤、软组织肉瘤(例如，软组织腺泡状肉瘤、血管肉瘤、叶状囊性肉瘤、皮肤纤维肉瘤、硬纤维瘤、上皮样肉瘤、骨外骨肉瘤、纤维肉瘤、血管外皮细胞瘤、血管肉瘤、卡波西氏肉瘤、平滑肌肉瘤、脂肪肉瘤、淋巴管肉瘤、淋巴肉瘤、恶性纤维组织细胞瘤、神经纤维肉瘤、横纹肌肉瘤和滑膜肉瘤)。

[0509] 上皮癌是开始于上皮细胞(即覆盖身体表面、产生激素并构成腺体的细胞)的癌症。作为非限制性实例，上皮癌癌包括乳腺癌、胰腺癌、肺癌、结肠癌、结直肠癌、直肠癌、肾癌、膀胱癌、胃癌、前列腺癌、肝癌、卵巢癌、脑癌、阴道癌、外阴癌、子宫癌、口癌、阴茎癌、睾丸癌、食道癌、皮肤癌、输卵管癌、头颈癌、胃肠间质癌、腺癌、皮肤或眼内黑色素瘤、肛区癌症、小肠癌、内分泌系统癌症、甲状腺癌、甲状旁腺癌、肾上腺癌、尿道癌、肾盂癌、输尿管癌、子宫内膜癌、宫颈癌、垂体癌、中枢神经系统(CNS)肿瘤、原发性中枢神经系统淋巴瘤、脑干神经胶质瘤和脊椎轴肿瘤。在一些情况下，该癌症是皮肤癌，例如基底细胞癌、鳞状细胞癌、黑色素瘤、非黑色素瘤或光化性(日光性)角化病。

[0510] 在一些情况下，该癌症是肺癌。肺癌可开始于气道(气道将气管分岔以供应肺(支气管)或肺的小气囊(肺泡))。肺癌包括非小细胞肺癌(NSCLC)、小细胞肺癌和间皮瘤。NSCLC的实例包括鳞状细胞癌、腺癌和大细胞癌。间皮瘤可以是肺和胸腔的内衬(胸膜)或腹腔的内衬(腹膜)的癌性肿瘤。间皮瘤可能是由于石棉暴露引发的。该癌症可以是脑癌，例如胶质母细胞瘤。

[0511] 可替代地，该癌症可以是中枢神经系统(CNS)肿瘤。CNS肿瘤可以被归类为神经胶质瘤或非神经胶质瘤。神经胶质瘤可以是恶性神经胶质瘤、高分级神经胶质瘤、弥散性内源性脑桥胶质瘤。神经胶质瘤的实例包括星形细胞瘤、少突神经胶质细胞瘤(或者少突神经胶质细胞瘤和星形细胞瘤成分的混合)和室管膜瘤。星形细胞瘤包括但不限于，低度星形细胞瘤、间变性星形细胞瘤、多形性胶质母细胞瘤、纤维状细胞性星形细胞瘤、多形性黄色星形细胞瘤和室管膜下巨细胞星形细胞瘤。少突神经胶质细胞瘤包括低度少突神经胶质细胞瘤

(或寡星状细胞瘤)和间变性少突神经胶质细胞瘤。非神经胶质瘤包括脑膜瘤、垂体腺瘤、原发性中枢神经系统淋巴瘤和髓母细胞瘤。在一些情况下,该癌症是脑膜瘤。

[0512] 白血病可以是急性淋巴细胞性白血病、急性粒细胞性白血病、慢性淋巴细胞性白血病或慢性粒细胞性白血病。其他类型的白血病包括多毛细胞白血病、慢性粒-单核细胞性白血病和幼年粒-单核细胞性白血病。

[0513] 淋巴瘤是淋巴细胞的癌症,并且可从B淋巴细胞或T淋巴细胞发展而成。淋巴瘤的两种主要类型是霍奇金淋巴瘤(以前称为霍奇金病)和非霍奇金淋巴瘤。霍奇金淋巴瘤的特点是里-施细胞(Reed-Sternberg cell)的存在。非霍奇金淋巴瘤是所有那些不属于霍奇金淋巴瘤的淋巴瘤。非霍奇金淋巴瘤可以是缓慢进展淋巴瘤和侵袭性淋巴瘤。非霍奇金淋巴瘤包括但不限于,弥漫性大B细胞淋巴瘤、滤泡性淋巴瘤、粘膜相关淋巴组织淋巴瘤(MALT)、小细胞淋巴细胞性淋巴瘤、套细胞淋巴瘤、伯基特淋巴瘤、纵隔大B细胞淋巴瘤、瓦尔登斯特伦巨球蛋白血症(Waldenströmmacroglobulinemia)、结节性边缘区B细胞淋巴瘤(NMZL)、脾边缘区淋巴瘤(SMZL)、结外边缘区B细胞淋巴瘤、血管内大B细胞淋巴瘤、原发性渗出性淋巴瘤和淋巴瘤样肉芽肿病。

[0514] 病原性感染

[0515] 在一些情况下,该疾病或病症是病原性感染。待进行随机标记的分子可以来自病原体。该病原体可以是病毒、细菌、真菌或原生动物。在一些情况下,该病原体可以是原生动物,例如棘阿米巴属(Acanthamoeba)(例如,阿斯特罗尼棘阿米巴(A.astronyxis)、卡氏棘阿米巴(A.castellanii)、卡伯特森氏棘阿米巴(A.culbertsoni)、A.hatchetti、多食棘阿米巴(A.polyphaga)、A.rhysodes、A.healyi、A.divionensis)、Brachiola(例如,B.connori、B.vesicularum)、隐孢子虫属(Cryptosporidium)(例如,微小隐孢子虫(C.parvum))、环孢子虫属(Cyclospora)(例如,卡晏环孢子虫(C.cayetanensis))、脑炎微孢子虫属(Encephalitozoon)(例如,兔脑炎微孢子虫(E.cuniculi)、海伦脑炎微孢子虫(E.hellem)、肠脑炎微孢子虫(E.intestinalis))、内阿米巴属(Entamoeba)(例如,溶组织内阿米巴(E.histolytica))、肠孢虫属((Enterocytozoon)(例如,比氏肠孢虫(E.bieneusi))、贾第虫属(Giardia)(例如,兰伯贾第虫(G.lamblia))、等孢子球虫属(Isospora)(例如,贝氏等孢子球虫(I.belli))、微孢子虫属(Microsporidium)(例如,非洲微孢子虫(M.africanum)、M.ceylonensis)、耐格里原虫属(Naegleria)(例如,福氏耐格里原虫(N.fowleri))、微粒子虫属(Nosema)(例如,海藻小孢子虫(N.algerae)、眼微粒子虫(N.ocularum))、具褶孢虫属(Pleistophora)、气管普孢虫属(Trachipleistophora)(例如,害人气管普孢虫(T.anthropophthera)、人气管普孢虫(T.hominis))和角膜微孢子虫属(Vittaforma)(例如,角膜微孢子虫(V.corneae))。该病原体可以是真菌,如假丝酵母属(Candida)、曲霉属(Aspergillus)、隐球酵母属(Cryptococcus)、组织胞浆菌属(Histoplasma)、肺囊虫属(Pneumocystis)和葡萄穗霉属(Stachybotrys)。

[0516] 该病原体可以是细菌。细菌的实例包括但不限于,博德特氏菌属(Bordetella)、疏螺旋体属(Borrelia)、布鲁氏菌属(Brucella)、弯曲杆菌属(Campylobacter)、衣原体属(Chlamydia)、嗜衣体属(Chlamydophila)、梭菌属(Clostridium)、棒杆菌属(Corynebacterium)、肠球菌属(Enterococcus)、埃希氏菌属(Escherichia)、弗朗西斯氏菌属(Francisella)、嗜血菌属(Haemophilus)、螺杆菌属(Helicobacter)、军团菌属

(Legionella)、钩端螺旋体属(Leptospira)、利斯特氏菌属(Listeria)、分枝杆菌属(Mycobacterium)、支原体属(Mycoplasma)、奈瑟氏球菌属(Neisseria)、假单细胞菌属(Pseudomonas)、立克次氏体属(Rickettsia)、沙门氏菌属(Salmonella)、志贺氏菌属(Shigella)、葡萄球菌属(Staphylococcus)、链球菌属(Streptococcus)、密螺旋体属(Treponema)、弧菌属(Vibrio)或耶尔森氏菌属(Yersinia)。

[0517] 该病毒可以是逆转录病毒。逆转录病毒的实例包括但不限于单链RNA-RT(ssRNA-RT)病毒和双链DNA-RT(dsDNA-RT)病毒。ssRNA-RT病毒的非限制性实例包括逆转录病毒、α逆转录病毒、β逆转录病毒、γ逆转录病毒、δ逆转录病毒、ε逆转录病毒、慢病毒、泡沫病毒(spuma virus)、转座病毒(metavirirus)和假病毒。dsDNA-RT病毒的非限制性实例包括嗜肝DNA病毒(hepadenovirus)和花椰菜花叶病毒。该病毒可以是DNA病毒。该病毒可以是RNA病毒。DNA病毒可以是双链DNA(dsDNA)病毒。在一些情况下,dsDNA病毒是腺病毒、疱疹病毒或痘病毒。腺病毒的实例包括但不限于腺病毒和犬传染性肝炎病毒。疱疹病毒的实例包括但不限于,单纯疱疹病毒、水痘带状疱疹病毒、巨细胞病毒和EB病毒。痘病毒的非限制性列表包括天花病毒、牛痘病毒、绵羊痘病毒、猴痘病毒和痘苗病毒。DNA病毒可以是单链DNA(ssDNA)病毒。ssDNA病毒可以是细小病毒。细小病毒的实例包括但不限于,细小病毒B19、犬细小病毒、小鼠细小病毒、猪细小病毒、猫传染性粒细胞缺乏症病毒和貂肠炎病毒。

[0518] 该病毒可以是RNA病毒。RNA病毒可以是双链RNA(dsRNA)病毒、正义单链RNA病毒((+)ssRNA病毒)或反义单链RNA病毒((-)ssRNA病毒)。dsRNA病毒的非限制性目录包括呼肠孤病毒、正呼肠孤病毒、质型多角体病毒、轮状病毒、蓝舌病病毒和植物呼肠孤病毒。(+)ssRNA病毒的实例包括但不限于小核糖核酸病毒和披膜病毒。小核糖核酸病毒的实例包括但不限于,肠病毒、鼻病毒、肝病毒属、心病毒属、口蹄疫病毒、脊髓灰质炎病毒、小RNA病毒、马鼻病毒、嵴病毒、捷申病毒和柯萨奇病毒。在一些情况下,披膜病毒是风疹病毒、辛德毕斯病毒、东方马脑炎病毒、西方马脑炎病毒、委内瑞拉马脑炎病毒、罗斯河病毒、阿尼昂-尼昂病毒、基孔肯雅病(Chikungunya)或塞姆利基森林病毒。(-)ssRNA病毒的非限制性列表包括正粘病毒和棒状病毒。正粘病毒的实例包括但不限于甲型流感病毒、乙型流感病毒、丙型流感病毒、传染性鲑鱼贫血病毒和索哥托病毒。弹状病毒的实例包括但不限于质型弹状病毒、迪路弹状病毒(dichorhabdovirus)、短暂热病毒、狂犬病病毒、粒外弹状病毒和水泡病毒。

[0519] 胎儿异常

[0520] 在一些情况下,该疾病或病症是妊娠。本文公开的方法和试剂盒可以包括诊断怀孕的受试者中的胎儿状况。本文公开的方法和试剂盒可以包括鉴别胎儿突变或遗传异常。待进行随机标记的分子可以来自胎儿细胞或组织。可替代地,或另外地,待标记的分子可以来自怀孕的受试者。

[0521] 本文公开的方法和试剂盒可以用于诊断、预测或监测常染色体三体性(例如,13、15、16、18、21或22号染色体三体性)。在一些情况下,三体性可与增加的流产几率相关联(例如,15、16或22号染色体三体性)。在其他情况下,所检测的三体性是活产三体性,这可能表明婴儿将带有出生缺陷(例如,13号染色体三体性(帕塔综合征)、18号染色体三体(爱德华兹综合征)和21号染色体三体(唐氏综合征))出生。所述异常也可能是性染色体异常(例如,XXY(克氏综合征)、XYY(雅各布斯氏综合征)或XXX(X三体性))。待标记的一种或多种分子可以位于如下一条或多条染色体上:13号染色体、18号染色体、21号染色体、X染色体或Y染色

体。例如，所述分子位于21号染色体和/或18号染色体上，和/或13号染色体上。

[0522] 可以根据本文公开的方法和试剂盒确定的胎儿的另外的状况包括：一条或多条染色体的单体性(X染色体单体性，也被称为特纳氏综合征)、一条或多条染色体的三体性(13、18、21和X染色体)、一条或多条染色体的四体性和五体性(在人类中最常见于性染色体，例如XXXX、XXYY、XXXY、XYYY、XXXXX、XXXXY、XXXYY、XYYYY和XXYYY)、单倍性、三倍性(每种染色体有三条，例如人类中的69条染色体)、四倍性(每种染色体有四条，例如人类中的92条染色体)、五倍性和多倍性。

[0523] 在此进一步披露的是法医分析的方法，该方法包括以上描述的任何方法。法医科学家可使用在犯罪现场发现的各种样品(例如，血、精液、皮肤、唾液、毛发)中的核酸，以鉴别个体(如犯罪者)在案发现场的存在。该过程正式地命名为DNA分析，但也可以称为“基因指纹分析”。例如，DNA分析包括测量和比较不同样品和人中重复DNA的变量段的长度，如短串联重复以及小卫星。对于将来自一个人的DNA样品与在犯罪现场发现的样品中DNA进行匹配，这种方法通常是极为可靠的技术。然而，如果现场被来自若干人的DNA所污染，那么鉴别可能很复杂。在这种情况下，以及在其他法医学应用中，从单细胞或少量细胞中获得绝对定量的核酸可以是有利的。

[0524] 在一些情况下，该疾病或病症是免疫障碍。免疫障碍可以是炎性疾病、自身免疫性障碍、肠道易激综合症或溃疡性结肠炎。自身免疫性疾病的实例包括克罗恩病(Chrohn's disease)、狼疮、以及格雷夫斯氏病(Graves'disease)。

[0525] 在一些情况下，该疾病或障碍是神经疾病或障碍。神经疾病或障碍可以是获得性癫痫失语症(Acquired Epileptiform Aphasia)、急性播散性脑脊髓炎(Acute Disseminated Encephalomyelitis)、肾上腺脑白质营养不良(Adrenoleukodystrophy)、胼胝体发育不良(Agenesis of the corpus callosum)、失认症(Agnosia)、艾卡尔迪综合征(Aicardi syndrome)、亚历山大症(Alexander disease)、阿尔珀斯病(Alpers'disease)、交替性偏瘫(Alternating hemiplegia)、阿尔茨海默病(Alzheimer's disease)、肌萎缩性脊髓侧索硬化症(Amyotrophic lateral sclerosis)(见运动神经元病(Motor Neuron Disease))、先天无脑畸形(Anencephaly)、天使人症候群(Angelman syndrome)、血管瘤病(Angiomatosis)、缺氧症(Anoxia)、失语症(Aphasia)、失用症(Apraxia)、蛛网膜囊肿(Arachnoid cysts)、蛛网膜炎(Arachnoiditis)、阿-基氏脑畸形(Arnold-Chiari malformation)、动静脉畸形(Arteriovenous malformation)、阿斯伯格综合症(Asperger's syndrome)、共济失调毛细血管扩张(Ataxia Telangiectasia)、注意缺陷多动障碍(Attention Deficit Hyperactivity Disorder)、孤独症(Autism)、言语听觉处理失调(Auditory processing disorder)、自主神经功能障碍(Autonomic Dysfunction)、背痛(Back Pain)、登氏症(Batten disease)、白塞氏病(Behcet's disease)、贝尔氏麻痹(Bell's palsy)、良性特发性眼睑痉挛(Benign Essential Blepharospasm)、良性局限性肌萎缩症(Benign Focal Amyotrophy)、良性颅内高压症(Benign Intracranial Hypertension)、双侧额顶区多小脑回(Bilateral frontoparietal polymicrogyria)、皮层下动脉硬化性脑病(Binswanger's disease)、睑痉挛(Blepharospasm)、布洛克-苏兹贝格综合征(Bloch-Sulzberger syndrome)、臂丛神经损伤(Brachial plexus injury)、脑脓肿(Brain abscess)、脑损伤(Brain damage)、脑外伤(Brain injury)、脑肿瘤(Brain

tumor)、脊髓半侧横断综合征(Brown-Sequard syndrome)、卡纳万病(Canavan disease)、腕隧道综合症(Carpal tunnel syndrome)(CTS)、灼痛(Causalgia)、中枢性疼痛综合症(Central pain syndrome)、脑桥中央髓鞘溶解(Central pontine myelinolysis)、中央核肌病(Centronuclear myopathy)、头部障碍(Cephalic disorder)、脑动脉瘤(Cerebral aneurysm)、脑动脉硬化(Cerebral arteriosclerosis)、大脑萎缩(Cerebral atrophy)、脑性巨人症(Cerebral gigantism)、大脑性麻痹(Cerebral palsy)、腓骨肌萎缩症(Charcot-Marie-Tooth disease)、小脑扁桃体下疝畸形(Chiari malformation)、舞蹈病(Chorea)、慢性脱髓鞘性神经炎(Chronic inflammatory demyelinating polyneuropathy)(CIDP)、慢性疼痛(Chronic pain)、慢性区域性疼痛综合征(Chronic regional pain syndrome)、科芬-劳里综合征(Coffin Lowry syndrome)、昏迷(Coma)(包括持续性植物状态(Persistent Vegetative State))、复合物I缺陷综合征(Complex I deficiency syndrome)、复合物I缺陷综合征(Complex I deficiency syndrome)、复合物II缺陷综合征(Complex II deficiency syndrome)、复合物III缺陷综合征(Complex III deficiency syndrome)、复合物IV/COX缺陷综合征(Complex IV/COX deficiency syndrome)、复合物V缺陷综合征(Complex V deficiency syndrome)、先天性双侧面瘫(Congenital facial diplegia)、皮质基底核退化症(Corticobasal degeneration)、颅动脉炎(Cranial arteritis)、颅缝早闭(Craniosynostosis)、克雅二氏症(Creutzfeldt-Jakob disease)、累积创伤失调(Cumulative trauma disorders)、库兴氏综合征(Cushing's syndrome)、巨细胞性包涵体病(Cytomegalic inclusion body disease)(CIBD)、巨细胞病毒感染(Cytomegalovirus Infection)、第四脑室闭锁综合征(Dandy-Walker syndrome)、道森病(Dawson disease)、复合物I的线粒体NADH脱氢酶组分缺陷(Deficiency of mitochondrial NADH dehydrogenase component of Complex I)、德氏综合征(De Morsier's syndrome)、下臂丛麻痹(Dejerine-Klumpke palsy)、德热里纳-索塔斯病(Dejerine-Sottas disease)、睡眠相位后移症候群(Delayed sleep phase syndrome)、痴呆(Dementia)、皮肌炎(Dermatomyositis)、神经性运用障碍(Neurological Dyspraxia)、糖尿病性神经病(Diabetic neuropathy)、弥漫性硬皮病(Diffuse sclerosis)、家族性自主神经异常(Dysautonomia)、计算障碍(Dyscalculia)、书写障碍(Dysgraphia)、阅读障碍(Dyslexia)、肌张力障碍(Dystonia)、早期婴儿型癫痫性脑病(Early infantile epileptic encephalopathy)、空蝶鞍综合征(Empty sella syndrome)、脑炎(Encephalitis)、脑膨出(Encephalocele)、脑三叉神经血管瘤病(Encephalotrigeminal angiomatosis)、大便失禁(Encopresis)、癫痫(Epilepsy)、厄尔布氏麻痹(Erb's palsy)、红斑性肢痛症(Erythromelalgia)、特发性震颤(Essential tremor)、法布莱氏病(Fabry's disease)、法尔氏综合征(Fahr's syndrome)、昏厥(Fainting)、家族性痉挛性麻痹(Familial spastic paralysis)、高热惊厥(Febrile seizures)、费希尔氏综合征(Fisher syndrome)、弗里德里奇共济失调(Friedreich's ataxia)、FART综合征(FART Syndrome)、高歇尔病(Gaucher's disease)、格斯特曼综合症(Gerstmann's syndrome)、巨细胞性动脉炎(Giant cell arteritis)、巨细胞性包涵体病(Giant cell inclusion disease)、球形细胞样脑白质病(Globoid cell Leukodystrophy)、灰质异位症(Gray matter heterotopia)、格林-巴利综合征(Guillain-Barre syndrome)、HTLV-1型相关脊髓病

(HTLV-1associated myelopathy)、蛋白球色素退变综合征(Hallervorden-Spatz disease)、头部损伤(Head injury)、头痛(Headache)、面部单侧痉挛(Hemifacial Spasm)、遗传性痉挛性下半身麻痹(Hereditary Spastic Paraplegia)、多神经炎型遗传性共济失调(Heredopathia atactica polyneuritiformis)、耳带状疱疹(Herpes zoster oticus)、带状疱疹(Herpes zoster)、Hirayama syndrome、前脑无裂畸形(Holoprosencephaly)、亨廷顿氏病(Huntington's disease)、积水性无脑(Hydranencephaly)、脑积水(Hydrocephalus)、皮质醇增多症(Hypercortisolism)、组织缺氧(Hypoxia)、免疫介导的脑脊髓炎(Immune-Mediated encephalomyelitis)、包涵体肌炎(Inclusion body myositis)、色素失调症(Incontinentia pigmenti)、婴儿植烷酸贮积病(Infantile phytanic acid storage disease)、婴儿雷夫叙姆病(Infantile Refsum disease)、婴儿痉挛症(Infantile spasms)、炎症性肌病(Inflammatory myopathy)、颅内囊肿(Intracranial cyst)、颅内压增高(Intracranial hypertension)、朱伯特综合症(Joubert syndrome)、卡恩斯塞尔综合征(Kearns-Sayre syndrome)、肯尼迪氏症(Kennedy disease)、金斯伯恩综合征(Kinsbourne syndrome)、先天性短颈综合征(Klippel Feil syndrome)、克拉伯病(Krabbe disease)、库福-瑞科波综合征(Kufor-Rakeb syndrome)、库格尔贝格-韦兰德病(Kugelberg-Welander disease)、苦鲁病(Kuru)、拉福拉病(Lafora disease)、兰-爱二氏肌无力综合征(Lambert-Eaton myasthenic syndrome)、获得性癫痫失语综合征(Landau-Kleffner syndrome)、延髓外侧(瓦伦伯格氏)综合征(Lateral medullary(Wallenberg)syndrome)、学习障碍(Learning disabilities)、利氏病(Leigh's disease)、勒-格二氏综合征(Lennox-Gastaut syndrome)、自毁容貌综合征(Lesch-Nyhan syndrome)、脑白质营养不良(Leukodystrophy)、路易体痴呆症(Lewy body dementia)、无脑回畸形(Lissencephaly)、闭锁综合征(Locked-In syndrome)、肌萎缩性侧索硬化症(Lou Gehrig's disease)、腰椎间盘突出症(Lumbar disc disease)、莱姆病-神经系统后遗症(Lyme disease-Neurological Sequelae)、马查多-约瑟夫病(脊髓小脑性共济失调3型)(Machado-Joseph disease(Spinocerebellar ataxia type 3))、巨脑征(Macrencephaly)、枫糖尿症(Maple Syrup Urine Disease)、巨脑畸形(Megalencephaly)、梅-罗综合征(Melkersson-Rosenthal syndrome)、美尼尔氏综合症(Menieres disease)、脑膜炎(Meningitis)、门克斯病(Menkes disease)、异染性脑白质营养不良(Metachromatic leukodystrophy)、头小畸型(Microcephaly)、偏头痛(Migraine)、米勒-费希尔综合征(Miller Fisher syndrome)、短暂性脑缺血发作(Mini-Strokes)、线粒体病(Mitochondrial disease)、线粒体功能障碍(Mitochondrial dysfunction)、线粒体肌病(Mitochondrial Myopathies)、线粒体呼吸链复合体I缺陷(Mitochondrial Respiratory Chain Complex I Deficiency)、默比乌斯综合征(Mobius syndrome)、单肢肌萎缩(Monomelic amyotrophy)、运动神经元病(Motor Neuron Disease)、运动技能障碍(Motor skills disorder)、烟雾病(Moyamoya disease)、黏多醣贮积症(Mucopolysaccharidoses)、多发梗塞性痴呆(Multi-Infarct Dementia)、多灶性运动神经病(Multifocal motor neuropathy)、多发性硬化症(Multiple sclerosis)、伴有姿势性低血压的多系统萎缩症(Multiple system atrophy with postural hypotension)、肌肉萎缩症(Muscular dystrophy)、肌痛性脑脊髓炎(Myalgic encephalomyelitis)、重症肌无力

(Myasthenia gravis)、Myelinoclastic弥漫性硬化(Myelinoclastic diffuse sclerosis)、乳儿肌阵挛性脑病(Myoclonic Encephalopathy of infants)、肌阵挛(Myoclonus)、肌病(Myopathy)、肌管性肌病(Myotubular myopathy)、先天性肌强直(Myotonia congenita)、NADH-辅酶Q还原酶缺乏症(NADH-coenzyme Q reductase deficiency)、NADH:Q(1)氧化还原酶缺乏症(NADH:Q(1)oxidoreductase deficiency)、发作性嗜睡病(Narcolepsy)、多发性神经纤维瘤(Neurofibromatosis)、神经阻滞剂恶性综合征(Neuroleptic malignant syndrome)、AIDS的神经系统表现(Neurological manifestations of AIDS)、狼疮的神经系统后遗症(Neurological sequelae of lupus)、神经性肌强直(Neuromyotonia)、神经元蜡样脂褐质沉积症(Neuronal ceroid lipofuscinosis)、神经元移行异常(Neuronal migration disorders)、尼曼匹克症(Niemann-Pick disease)、非24小时睡醒周期综合征(Non 24-hour sleep-wake syndrome)、非言语型学习障碍(Nonverbal learning disorder)、奥沙利文-麦克劳德综合征(O'Sullivan-McLeod syndrome)、枕神经痛(Occipital Neuralgia)、隐性椎管闭合不全综合征(Occult Spinal Dysraphism Sequence)、大田原综合征(Ohtahara syndrome)、橄榄体桥脑小脑萎缩(Olivopontocerebellar atrophy)、眼阵挛-肌阵挛综合征(Opsoclonus myoclonus syndrome)、视神经炎(Optic neuritis)、起立性低血压(Orthostatic Hypotension)、过度使用症候群(Overuse syndrome)、氧化磷酸化障碍(oxidative phosphorylation disorders)、视觉重复(Palinopsia)、感觉异常(Paresthesia)、帕金森氏病(Parkinson's disease)、先天强直性肌痉挛(Paramyotonia Congenita)、副肿瘤病(Paraneoplastic diseases)、阵发性发作(Paroxysmal attacks)、帕瑞-罗姆伯格氏综合征(也称为罗姆伯格氏综合征)(Parry-Romberg syndrome(Rombergs Syndrome))、佩-梅二氏病(Pelizaeus-Merzbacher disease)、周期性瘫痪(Periodic Paralyses)、周围神经病(Peripheral neuropathy)、持续性植物状态(Persistent Vegetative State)、流行性神经障碍(Pervasive neurological disorders)、旋光性喷嚏反射(Photic sneeze reflex)、植烷酸贮积病(Phytanic Acid Storage disease)、皮克氏病(Pick's disease)、神经受压(Pinched Nerve)、垂体肿瘤(Pituitary Tumors)、PMG、小儿麻痹症(Polio)、多小脑回(Polymicrogyria)、多发性肌炎(Polymyositis)、孔洞脑(Porencephaly)、小儿麻痹后遗症(Post-Polio syndrome)、带状疱疹后遗神经痛(Postherpetic Neuralgia(PHN))、感染后脑脊髓炎(Postinfectious Encephalomyelitis)、体位性低血压(Postural Hypotension)、普拉德-威利综合征(Prader-Willi syndrome)、原发性脊髓侧索硬化(Primary Lateral Sclerosis)、朊病毒病(Prion diseases)、进行性面部偏侧萎缩(也称为罗姆伯格氏综合征)(Progressive Hemifacial Atrophy,Rombergs_Syndrome)、进行性多灶性白质脑病(Progressive multifocal leukoencephalopathy)、进行性硬化性灰白质萎缩(Progressive Sclerosing Poliodystrophy)、进行性核上性麻痹(Progressive Supranuclear Palsy)、假性脑瘤(Pseudotumor cerebri)、雷-享二氏综合征(I型和II型)(Ramsay-Hunt syndrome)、拉斯马森脑炎(Rasmussen's encephalitis)、反射交感性营养不良综合征(Reflex sympathetic dystrophy syndrome)、雷夫叙姆病(Refsum disease)、重复性运动损伤(Repetitive motion disorders)、重复性压力损伤(Repetitive stress injury)、多动腿综合征(Restless legs syndrome)、逆转录病毒相关的脊髓病

(Retrovirus-associated myelopathy)、雷特氏症(Rett syndrome)、莱以氏综合症(Reye's syndrome)、罗姆伯格氏综合征(Rombergs_Syndrome)、狂犬病(Rabies)、圣维特斯舞蹈病(Saint Vitus dance)、山德霍夫氏病(Sandhoff disease)、精神分裂症(Schytsophrenia)、谢耳德氏病(Schilder's disease)、脑裂畸形(Schizencephaly)、感知综合机能障碍(Sensory Integration Dysfunction)、视隔发育不良综合征(Septo-optic dysplasia)、摇晃婴儿综合症(Shaken baby syndrome)、带状疱疹(Shingles)、立位低血压综合征(Shy-Drager syndrome)、干燥综合征(Sjogren's syndrome)、睡眠呼吸暂停(Sleep apnea)、昏睡症(Sleeping sickness)、餍足(Snatiation)、索托氏症(Sotos syndrome)、痉挛状态(Spasticity)、脊柱裂(Spina bifida)、脊髓受伤(Spinal cord injury)、脊髓肿瘤(Spinal cord tumors)、脊髓性肌萎缩(Spinal muscular atrophy)、椎管狭窄(Spinal stenosis)、斯蒂尔-理查森-奥尔谢夫斯基综合征(Steele-Richardson-Olszewski syndrome)(参见进行性核上性麻痹(Progressive Supranuclear Palsy))、脊髓小脑性共济失调(Spinocerebellar ataxia)、僵人综合征(Stiff-person syndrome)、中风(Stroke)、斯特格-韦伯氏综合征(Sturge-Weber syndrome)、亚急性硬化性全脑炎(Subacute sclerosing panencephalitis)、皮层下动脉硬化性脑病(Subcortical arteriosclerotic encephalopathy)、浅表性铁质沉着(Superficial siderosis)、西登哈姆氏舞蹈病(Sydenham's chorea)、晕厥(Syncope)、共感觉(Synesthesia)、脊髓空洞症(Syringomyelia)、迟发性运动障碍(Tardive dyskinesia)、戴萨克斯症(Tay-Sachs disease)、颞动脉炎(Temporal arteritis)、骨髓栓系综合征(Tethered spinal cord syndrome)、肌强直性白内障(Thomsen disease)、胸廓出口综合征(Thoracic outlet syndrome)、三叉神经痛(Tic Douloureux)、托德麻痹(Todd's paralysis)、图雷特氏综合症(Tourette syndrome)、短暂性脑缺血发作(Transient ischemic attack)、传染性海绵状脑病(Transmissible spongiform encephalopathies)、横贯性脊髓炎(Transverse myelitis)、外伤性脑损伤(Traumatic brain injury)、震颤(Tremor)、三叉神经痛(Trigeminal neuralgia)、热带痉挛性轻截瘫(Tropical spastic paraparesis)、嗜眠病(Trypanosomiasis)、结节性硬化症(Tuberous sclerosis)、血管炎(Vasculitis)(包括颞动脉炎(temporal arteritis))、视网膜小脑及脊髓血管瘤症(Von Hippel-Lindau disease)(VHL)、维恩梅里斯症(Viliuisk Encephalomyelitis)(VE)、瓦伦贝格综合征(Wallenberg's syndrome)、韦德尼希-霍夫曼病(Werdnig-Hoffman disease)、韦斯特综合症(West syndrome)、颈椎过度屈伸损伤(Whiplash)、威廉斯综合征(Williams syndrome)、威尔森氏症(Wilson's disease)、X-连锁隐性遗传性脊髓延髓型肌萎缩(X-Linked Spinal and Bulbar Muscular Atrophy)、或赵苇格氏症(Zellweger syndrome)。

[0526] 定义

[0527] 在提供了一系列值时,应理解每个中间值,到下限的第十个单位(除非上下文清晰地另外指示),还具体地披露了该范围的上限与下限之间的值。陈述范围内的任何陈述值或中间值以及该规定范围内的任何其他陈述值或中间值之间每个较小的范围涵盖在本发明之内。这些较小范围的上限和下限可以独立地包括在所述范围内或排除在其外,并且每个范围(其中两个极限之一、没有一个、或都包括在所述较小的范围内)也包括在本发明内,服从所述范围中任何特别排除的限值。在所陈述的范围包括一个或两个限值时,排除了那些

被包括的限值的任一个或两者的范围也被包括在本发明之内。

[0528] 除非另外定义，在此所使用的全部技术术语和科学术语具有与本发明所属领域之内的普通技术人员通常所理解的相同的意思。虽然类似或等效于在此描述的那些方法和材料的任何方法和材料也可以用于本发明的实践或测试，但现在描述一些潜在以及优选的方法和材料。在此提及的所有公开是通过引用结合在此，以结合所引用的这些公开来披露和描述这些方法和/或材料。应理解的是，本披露取代结合的出版物的任何矛盾的披露。

[0529] 如将对于本领域技术人员清楚的是，在阅读本披露时，在此描述和展示的单独实施例的每一个具有离散的组成部分和特征，这些组成部分和特征可以在不偏离本发明的范围或精神的情况下易于与任何其他一些实施例的特征分离或组合。可以按照所叙述的事件的顺序或按照逻辑上可行的任何其他顺序来进行任何叙述的方法。

[0530] 必须注意，除非上下文另外明确指示，否则如在此和所附权利要求中所使用，单数形式“一个/一种(a/an)”和“该”包括复数指示物。因而，例如，提及“一个/种细胞”时包括多个/种此类细胞，并且提及“所述/该肽”时包括提及一个或多个肽和本领域技术已知的其等价物(例如，多肽)，等等。

[0531] 如在此所使用，术语“标记”可以指唯一性寡核苷酸序列，该序列可允许鉴定出相应的核酸碱基和/或核酸序列。在一些实施例中，核酸碱基和/或核酸序列可以位于更大多核苷酸序列的特定位置处(例如，附接至珠上的多核苷酸)。

[0532] 如在此所使用，术语“杂交”可以指这样一个过程，其中两条单链多核苷酸非共价地结合，以形成稳定的双链多核苷酸。术语“杂交”还可以指三链杂交。所得(通常)双链多核苷酸是“杂合体”或“双链体”。

[0533] 如在此所使用，“核苷”可以包括天然核苷，如2'-脱氧以及2'-羟基形式。关于核苷的“类似物”可以包括合成的核苷，这些合成的核苷包含经修饰的碱基部分和/或经修饰的糖部分等。类似物能够杂交。类似物可以包括合成的核苷，这些合成的核苷被设计为增强结合特性、降低复杂度、增加特异性等。示例性类型的类似物可以包括寡核苷酸磷酰胺酯(在此称为“酰胺化物”)、肽核酸(在此称为“PNA”)、寡糖-2'-O-烷基核糖核苷酸、含多核苷酸的C-5丙炔基嘧啶、以及锁核酸(LNA)。

[0534] 如在此所使用，术语“核酸分子”、“核酸序列”、“核酸片段、“寡核苷酸”、“寡核苷酸片段”和“多核苷酸”可以互换地使用并且可以旨在包括但不限于，可以具有不同长度的核苷酸的聚合物形式，脱氧核糖核苷酸或核糖核苷酸或其类似物。核酸分子可以包括单链DNA(ssDNA)、双链DNA(dsDNA)、单链RNA(ssRNA)和双链RNA(dsRNA)。不同核酸分子可以具有不同的三维结构，并且可以执行不同的功能。核酸分子的非限制性实例可以包括基因、基因片段、基因组间隙(genomic gap)、外显子、内含子，基因间DNA(包括，但不限于，异染色质DNA)、信使RNA(mRNA)、转移RNA、核糖体RNA、核酶、小干扰RNA(siRNA)、miRNA、核仁小RNA(snoRNA)、cDNA、重组多核苷酸、分支多核苷酸、质粒、载体、序列的分离的DNA、序列的分离的RNA、核酸探针以及引物。

[0535] 寡核苷酸可以指由磷酸二酯键或其类似物连接的天然或修饰核苷单体的线性聚合物。“寡核苷酸片段”是指被切割为两个或更多个更小寡核苷酸序列的寡核苷酸序列。寡核苷酸可以是天然的或合成的。寡核苷酸可以包括脱氧核糖核苷、核糖核苷、及其非天然类似物(例如其异头形式、肽核酸(PNA))等。寡核苷酸能够通过单体与单体相互作用的正常模

式(如沃森-克里克(Watson-Crick)型的碱基配对、碱基堆积、霍氏(Hoogsteen)或反霍氏(reversed Hoogsteen)类型的碱基配对等)特异性结合至靶基因组。寡核苷酸与术语“多核苷酸”可以在此处可互换地使用。

[0536] 除非另外指出,每当寡核苷酸由一系列字母(如“ATGCCTG”)表示时,应理解的是,核苷酸从左往右是以5’至3’的顺序,并且“A”表示脱氧腺苷,“C”表示脱氧胞苷,“G”表示脱氧鸟苷,“T”表示脱氧胸苷,以及“U”表示核糖核苷,尿苷。

[0537] 寡核苷酸可以包括一个或多个非标准核苷酸、一个或多个核苷酸类似物和/或修饰的核苷酸。修饰的核苷酸的实例可以包括但不限于:二氨基嘌呤、5-氟尿嘧啶、5-溴尿嘧啶、5-氯尿嘧啶、5-碘尿嘧啶、次黄嘌呤、黄嘌呤、4-乙酰基胞嘧啶、5-(羧基羟基甲基)尿嘧啶、5-羧基甲基氨基甲基-2-硫代尿苷、5-羧基甲基氨基甲基尿嘧啶、二氢尿嘧啶、β-D-半乳糖基辫苷、肌苷、N6-异戊烯基腺嘌呤、1-甲基鸟嘌呤、1-甲基肌苷、2,2-二甲基鸟嘌呤、2-甲基腺嘌呤、2-甲基鸟嘌呤、3-甲基胞嘧啶、5-甲基胞嘧啶、N6-腺嘌呤、7-甲基鸟嘌呤、5-甲基氨基甲基尿嘧啶、5-甲氧基氨基甲基-2-硫代尿嘧啶、β-D-甘露糖基辫苷、5’-甲氧基羧基甲基尿嘧啶、5-甲氧基尿嘧啶、2-甲基硫代-D46-异戊烯基腺嘌呤、尿嘧啶-5-氧基乙酸(v)、怀丁苷(wybutoxosine)、假尿嘧啶、辫苷(queosine)、2-硫代胞嘧啶、5-甲基-2-硫代尿嘧啶、2-硫代尿嘧啶、4-硫代尿嘧啶、5-甲基尿嘧啶、尿嘧啶-5-氧基乙酸甲基酯、尿嘧啶-5-氧基乙酸(v)、5-甲基-2-硫代尿嘧啶、3-(3-氨基-3-N-2-羧基丙基)尿嘧啶、(acp3)w、寡核苷酸磷酰胺酯(在此称为“酰胺化物”)、肽核酸(在此称为“PNA”)、寡-2’-O-烷基核糖核苷酸、含多核苷酸的C-5丙炔基嘧啶、锁核酸(LNA)、2,6-二氨基嘌呤以及类似物。核酸分子还可以在碱基部分(例如,在通常能够与互补核苷酸形成氢键的一个或多个原子处和/或在通常不能够与互补核苷酸形成氢键的一个或多个原子处)、糖部分或磷酸骨架处进行修饰。

[0538] 如在此所使用,“样品”可以指单细胞或许多细胞。核酸分子可以获得自一个或多个样品。样品可以包括单细胞类型或两种或更多种细胞类型的组合。样品可以包括能执行类似功能的细胞(如例如在组织中发现的那些)的集合。样品可以包括一种或多种组织。组织的实例可以包括但不限于:上皮组织(例如,皮肤,腺、肠、皮肤以及器官(如肝、肺、肾)的内衬)、内皮(例如,血管和淋巴管的内衬)、间皮(例如,胸膜腔、腹膜腔以及心包腔的内衬)、间充质(例如,填充器官(包括脂肪、肌肉、骨骼、软骨以及肌腱细胞)之间空间的细胞)、血细胞(例如,红细胞、粒细胞、中性粒细胞、嗜酸性粒细胞、嗜碱性粒细胞、单核细胞、T淋巴细胞(又称为T细胞)、B淋巴细胞(又称为B细胞)、浆细胞、巨核细胞等)、神经细胞、生殖细胞(例如,精子、卵母细胞)、羊水细胞、胎盘、干细胞等。样品可以获得自培养物中一个或多个单细胞、宏基因组样品、胚胎干细胞、诱导性多能干细胞,癌样品、组织切片和活检或其任意组合。

[0539] 如在此所使用,术语“有机体”可以包括但不限于:人类、非人类灵长动物、母牛、马、绵羊、山羊、猪、狗、猫、兔、小鼠、大鼠、沙鼠、蛙、蟾蜍、鱼(例如,斑马鱼(Danio rerio))、线虫(例如,秀丽隐杆线虫(C.elegans))及其任何转基因种类。术语“有机体”还可以包括但不限于:酵母菌(例如,酿酒酵母)细胞、酵母菌四分体、酵母菌落、细菌、细菌菌落、病毒粒子、病毒体、病毒样颗粒和/或其培养物等。

[0540] 如在此所使用,术语“附接”、“缀合”、以及“偶联”可以互换地使用并且可以指共价相互作用和非共价相互作用。

[0541] 尽管在此已经显示并说明了本发明的优选实施例，但是本领域的普通技术人员将清楚的是仅作为举例而提供了此类实施例。本领域的普通技术人员现在将会想到众多变体、变化、以及替代，而不背离本发明。应该理解的是，在此说明的本发明的实施例的不同替代方案可以用于实施本发明。预期的是以下权利要求书限定了本发明的范围以及由此覆盖在这些权利要求和它们的等效物的范围内的方法和结构。

[0542] 实例

[0543] 实例1:酶法分离-聚池合成

[0544] 在此实例中，使用酶法分离-聚池合成来生产偶联寡核苷酸的珠。如在图2A中所示，将一组寡核苷酸添加至第一板的每个孔中。在一组寡核苷酸中的寡核苷酸包括5'胺、通用序列、细胞标记和接头。对于每组寡核苷酸而言，该5'胺、通用序列和接头是相同的。该通用序列和接头彼此不同。然而，对于每组寡核苷酸而言，该细胞标记是不同的。因此，每个孔具有不同的细胞标记。

[0545] 在酶法分离-聚池合成的步骤1中，偶联寡核苷酸的珠是通过将单个珠添加至每个孔并且进行1-乙基-3-(3-二甲基氨基丙基)碳二亚胺(EDC)偶联反应合成的。从步骤1获得的这些寡核苷酸珠包括偶联多个寡核苷酸的珠。寡核苷酸包括5'-胺、通用序列，细胞标记1、和接头1(参见图2A)。在同一珠上的寡核苷酸是相同的。然而，在第一珠上的寡核甘酸不同于在第二珠上的寡核苷酸。

[0546] 在酶法分离-聚池合成的步骤2中，进行多次洗涤以去除未偶联的寡核苷酸。一旦去除这些未偶联的寡核苷酸，将偶联寡核苷酸的珠聚池(参见图2A)。从步骤2获得的偶联寡核苷酸的珠包括偶联多个单链寡核苷酸的珠。该单链寡核苷酸包括5'胺、通用序列、细胞标记1和接头1。在珠上的每个寡核苷酸是相同的。然而，每个珠包括不同的寡核苷酸。偶联至不同珠上的寡核苷酸的区别在于细胞标记1序列。

[0547] 如在图2B中所示，将一组寡核苷酸添加至第二板的每个孔。在一组寡核苷酸中的寡核苷酸包括第一接头、细胞标记、和第二接头。对于每组寡核苷酸而言，该第一接头和该第二接头是相同的。该第一接头和该第二接头彼此不同。然而，对于每组寡核苷酸而言，该细胞标记是不同的。因此，每个孔具有不同的细胞标记。

[0548] 在酶法分离-聚池合成的步骤3中，将在步骤2中聚池的偶联寡核苷酸的珠分离到第二板的孔中。因为在第二板的孔中的寡核苷酸的第一接头与偶联到珠上的寡核苷酸的接头互补，进行使用克列诺(Klenow)大片段的引物延伸以将来自第二板的寡核苷酸偶联至来自步骤2的偶联寡核苷酸的珠上。从步骤3获得的偶联寡核苷酸的珠包括偶联多个双链寡核苷酸的珠。该双链寡核苷酸包括5'胺、通用序列、细胞标记1，接头1、细胞标记2、和接头2(参见图2B)。

[0549] 在酶法分离-聚池合成的步骤4中，进行多次洗涤以去除未偶联的寡核苷酸和克列诺大片段酶。将第二板进行加热以使双链寡核苷酸变性，并且将偶联寡核苷酸的珠聚池(参见图2B)。从步骤4获得的偶联寡核苷酸的珠包括偶联多个单链寡核苷酸的珠。该单链寡核苷酸包括5'胺、通用序列、细胞标记1，接头1、细胞标记2、和接头2。在珠上的每个寡核苷酸是相同的。然而，每个珠包括不同的寡核苷酸。偶联至不同珠上的寡核苷酸的区别在于组合的细胞标记序列。例如，第一珠可以包括以下寡核苷酸，这些寡核苷酸包含:细胞标记A的第一细胞标记和细胞标记C的第二细胞标记;并且第二珠可以包括以下寡核苷酸，这些寡核苷

酸包含:细胞标记C的第一细胞标记和细胞标记D的第二细胞标记。因此,该第一珠和该第二珠可以包括相同的细胞标记(在这种情况下,为细胞标记C),然而,该第一珠和该第二珠的组合的细胞标记序列是不同的(例如,对于第一珠,组合的细胞标记序列是细胞标记A+细胞标记C;对于第二珠,组合的细胞标记序列是细胞标记C+细胞标签A)。在其他情况下,两个珠可以包括含有不同细胞标记的寡核苷酸。例如,第一珠可包括含有细胞标记A和细胞标记B的寡核苷酸,并且第二珠可包括含有细胞标记C和细胞标记D的寡核苷酸。在这种情况下,第一珠的两种细胞标记不同于第二珠的两种细胞标记。

[0550] 如在图2C中所示,将一组寡核苷酸添加至第三板的每个孔。在一组寡核苷酸中的寡核苷酸包括接头、细胞标记、分子标记、和寡dT。对于每组寡核苷酸而言,该接头和寡dT序列是相同的。然而,对于每组寡核苷酸而言,该细胞标记是不同的。因此,每个孔具有不同的细胞标记。此外,对于在一组内的寡核苷酸,该分子标记是不同的。因此,单个孔包含具有相同细胞标记、但不同分子标记的多个寡核苷酸。来自不同孔的寡核苷酸可以包含相同的分子标记。

[0551] 在酶法分离-聚池合成的步骤5中,将在步骤4中聚池的偶联寡核苷酸的珠分离到第三板的孔中。因为在第二板的孔中的寡核苷酸的接头与偶联到珠上的寡核苷酸的第二接头互补,进行使用克列诺(Klenow)大片段的引物延伸以将来自第三板的寡核苷酸偶联至来自步骤4的偶联寡核苷酸的珠上。从步骤5获得的偶联寡核苷酸的珠包括偶联多个双链寡核苷酸的珠。该双链寡核苷酸包括5'胺、通用序列、细胞标记1、接头1、细胞标记2、接头2、细胞标记3、分子标记和寡dT(参见图2C)。

[0552] 在酶法分离-聚池合成的步骤6中,进行多次洗涤以去除未偶联的寡核苷酸和克列诺大片段酶。将第三板进行加热以使双链寡核苷酸变性,并且将偶联寡核苷酸的珠聚池(参见图2C)。从步骤4获得的偶联寡核苷酸的珠包括偶联多个单链寡核苷酸的珠。该单链寡核苷酸包括5'胺、通用序列、细胞标记1、接头1、细胞标记2、接头2、细胞标记3、分子标记和寡dT。在单个珠上的多个单链寡核苷酸可以通过分子标记来区分。在单个珠上的多个寡核苷酸的细胞标记部分是相同的。每个珠包括不同的寡核苷酸。偶联至不同珠上的寡核苷酸的区别在于细胞标记序列。来自不同珠的寡核苷酸上的分子标记可以是相同的。来自不同珠的寡核苷酸上的分子标记可以是不同的。两个或更多个珠的差别可以在于组合的细胞标记序列。例如,第一珠可以包括含有细胞标记A、细胞标记B和细胞标记C的寡核苷酸,并且第二珠可以包括含有细胞标记B、细胞标记D和细胞标记A的寡核苷酸。在这种情况下,该第一珠和第二珠均含有细胞标记B,然而两种其他细胞标记是不同的。因此,两个或更多个珠可以包括区别在于至少一种细胞标记的寡核苷酸。两个或更多个珠可以包括区别在于至少两种细胞标记的寡核苷酸。两个或更多个珠可以包括区别在于至少三种细胞标记的寡核苷酸。然而,珠可以包括含有两种或更多种相同细胞标记的寡核苷酸。例如,珠可以包括含有细胞标记A、细胞标记A和细胞标记D的寡核苷酸。珠可以包括含有至少三种相同的细胞标记的寡核苷酸。例如,珠可以包括含有细胞标记A、细胞标记A和细胞标记A的寡核苷酸。珠可以包括含有三种不同的细胞标记的寡核苷酸。例如,珠可以包括含有细胞标记A、细胞标记D和细胞标签E的寡核苷酸。珠可以包括含有至少两种不同分子标记的至少两种寡核苷酸。例如,珠可以包括含有分子标记A的第一寡核苷酸以及含有分子标记D的第二寡核苷酸。然而,珠可以包括含有第一分子标记的寡核苷酸的多个拷贝。因此,珠可以包括含有相同分子标记的

至少两个寡核苷酸。例如，珠可以包括含有分子标记A的第一寡核苷酸以及含有分子标记A的第二寡核苷酸。在珠上的至少30％的寡核苷酸可以包括不同的分子标记。在珠上的至少40％的寡核苷酸可以包括不同的分子标记。在珠上的至少50％的寡核苷酸可以包括不同的分子标记。在珠上的至少60％的寡核苷酸可以包括不同的分子标记。在珠上的少于30％的寡核苷酸可以包括相同的分子标记。在珠上的少于20％的寡核苷酸可以包括相同的分子标记。在珠上的少于15％的寡核苷酸可以包括相同的分子标记。在珠上的少于10％的寡核苷酸可以包括相同的分子标记。在珠上的少于5％的寡核苷酸可以包括相同的分子标记。

[0553] 酶法分离-聚池合成技术可以在多个板或具有更多数目的孔的板上进行，以产生更多数目的偶联寡核苷酸的珠。使用的三个单独的细胞标记部分可增加在这些珠上的总体细胞标记部分的多样性。对于每个细胞标记部分具有96个不同的序列选择，可以产生884,736个不同的细胞标记组合。

[0554] 实例2：在管和微孔中的扩增的比较

[0555] 本披露提供了一种用于捕获细胞的方法。约5,000个拉莫斯细胞(Ramos cell)被捕获在包含直径为约30微米的微孔的微孔阵列上。一些细胞未被捕获。对于该实验的对照是捕获于管中的等量细胞。将在管中的细胞和在微孔阵列中的细胞两者均进行裂解。允许核酸杂交至一种缀合的珠。进行GAPDH和RPL19基因的实时PCR。

[0556] 图9示出实时PCR扩增的结果。来自微孔的产量大于来自管中的核酸的产量，指示在微孔中该核酸与该寡核苷酸的杂交比在管中的更有效(对应地，比较灰色条和白色条)。

[0557] 实例3：第二合成链的扩增和珠上的合成的比较

[0558] 如实例1所述获得并裂解细胞。对RPL19、TUBB、和GAPDH使用通用引物，脱离固体支持物从第二链扩增，或直接在固体支持物上扩增。图10示出与未直接从固体支持物进行的扩增相比，直接在固体支持物上的扩增(图10)产生更少的脱靶扩增。GAPDH和TUBB扩增产生正确大小的产物，不管何种方法(在图10中每三个重复的左泳道对应于管形式中的固体支持物加裂解物，每三个重复的中间泳道对应于来自微孔的固体支持物，并且每三个重复的右泳道对应于固体支持物加纯化的核酸)。RPL19产物具有最低限度的脱靶扩增产物，但当纯化的核酸与固体支持物一起使用时仅产生一种强的产物。这些实验表明，与使用脱离固体支持物合成的第二链进行的扩增相比，直接在珠上的扩增产生更少的脱靶扩增产物。

[0559] 实例4：靶核酸的多重分析

[0560] 如实例1所述获得并裂解细胞。靶核酸与包含寡核苷酸的固体支持物杂交。多个拷贝的靶核酸与包含寡dT序列的靶结合区杂交。使用逆转录酶将该多个拷贝的靶核酸进行逆转录。逆转录结合寡核苷酸的多个特征，该靶核酸的拷贝与该多个特征杂交(例如，分子标记、细胞标记、以及通用标记)。使用PCR扩增该靶核酸的多个拷贝。对经扩增的靶核酸拷贝进行测序。对经测序的靶核酸进行计数以确定在细胞中该靶核酸的拷贝数。计数是通过针对靶核酸的相同序列读数的每一个，对不同分子标记的数目进行计数来进行。以此方式，可以减小扩增偏差。

[0561] 实例5：评估分离-聚池合成用以产生以下珠的功效，这些珠具有一个细胞标记组合的克隆拷贝

[0562] 在此实例中，评价了分离-聚池合成用以产生以下珠的功效，这些珠具有一个细胞标记组合的克隆拷贝。通过如实例1中描述的酶法分离-聚池合成方法，合成偶联寡核苷酸

的珠。从拉莫斯细胞纯化250ng的总RNA，其相当于来自25,000个细胞的RNA。将该总RNA与35,000个偶联寡核苷酸的珠接触，导致mRNA与这些偶联寡核苷酸的珠杂交。在与这些偶联寡核苷酸的珠杂交的mRNA上进行cDNA合成。使用包含18、175和1750个珠的样品用于进一步分析。对与来自18-、175-和1750-珠样品的珠结合的cDNA，使用GAPDH-特异性引物和IGJ-特异性引物进行PCR扩增反应。对附接至珠上的cDNA分子进行测序。图11A-I示出测序结果的图形表示。对于图11A-C，分别对于18-珠、175-珠和1750-珠样品，读数数目/珠是绘制于y轴上，并且唯一性条码(例如，细胞标记组合)是绘制于x轴上。对于图11D-F，分别对于18-珠、175-珠和1750-珠样品，唯一性条码数目/珠是绘制于y轴上，并且唯一性条码(例如，细胞标记组合)是绘制于x轴上。对于图11G-I，分别对于18-珠、175-珠和1750-珠样品，唯一性分子数目/珠是绘制于y轴上，并且唯一性条码是绘制于x轴上。图11G-I的这些结果是通过分子的总数分选的。对于不同样品，唯一性分子数目中值/珠示于表1中。测序结果的数值示于表2中。

[0563]

表 1

	唯一性分子的中值数目		
	18 珠	175 珠	1750 珠
IGJ	78	85	40
GAPDH	22	45	25

[0564]

表 2

预期的珠#	17.5	175	1750
读数的总数目	58321	60308	133043
在恒定 1 中的 >=8 匹配	56385	57615	123349
在恒定 2 中的 >=8 匹配	54117	55187	115126
在恒定 1 和 2 中的 >=8 匹配	54114	55185	115107
在所有 3 个亚条码中的完全匹配	38585	46066	95217
在基因 (40 bp) 中的完全匹配	29968	33775	72260

[0565]

唯一性条码组合的总数目	239	407	1654
有用读数%	51.388%	56.00%	54.31%
唯一性条码组合 >20 读数的数目	5	26	288

[0566] 实例6.使用偶联寡核苷酸的珠进行单细胞RNA标记

[0567] 在此实例中，评价了使用寡核苷偶联的珠进行单细胞RNA标记的功效。制备三种细胞样品如下：

[0568]

	样品1:仅K562	样品2：仅拉莫斯（Ramos）	样品3：拉莫斯＋K562混合物
微孔数目	约10000	约10000	约10000
拉莫斯细胞数目	0	5000	3750
K562细胞数目	1000	0	2500

[0569] 将样品的细胞悬浮液添加到微孔的顶部，并且允许细胞沉降到微孔阵列的孔中。将未由微孔阵列捕获的细胞在磷酸盐缓冲盐水(PBS)浴中洗去。将偶联寡核苷酸的珠(如通过实例1中所述的酶法分离-聚池合成方法制备的)添加至微孔阵列。偶联寡核苷酸的珠包括具有多个寡核苷酸的磁珠。在珠上的每个寡核苷酸包括5'胺、通用序列、细胞标记1、接头1、细胞标记2、接头2、细胞标记3、分子标记以及寡dT。对于在相同珠上的每个寡核苷酸，这些寡核苷酸的序列除了分子标记外是相同的。对于在不同珠上的寡核苷酸，细胞标记1、2、和3组合是不同的。在微孔阵列中，添加约5-6个珠/孔。在一些情况下，对于每10个孔，50个珠可以沉积在该阵列上，其中0-2个珠落到每个孔中。允许珠沉降至孔中并且在PBS浴中洗去未捕获的珠。磁体放置该微孔阵列下方。将细胞通过添加冷裂解缓冲液裂解。将该阵列以及磁体放置在冷铝块上持续5分钟。来自裂解的细胞的mRNA与偶联到珠上的寡核苷酸杂交。将该阵列用过量的裂解缓冲液进行洗涤以除去未结合的mRNA。通过将磁体放置在微孔阵列顶部，从这些孔收回珠。将这些收回的珠进行洗涤。使用Superscript III，在50℃在旋转器上，在珠上进行cDNA合成持续50分钟。通过在旋转器上在37℃用ExoI处理30分钟，去除来自珠上的寡核苷酸的非延伸寡dT。对cDNA进行基因特异性PCR扩增。被选择用于该基因特异性PCR的基因是细胞类型特异性的，并且示于表3中。对PCR扩增产物进行测序。测序统计示于表4中。图12A-C示出分别对于仅K562样品、仅拉莫斯样品、以及K562+拉莫斯混合物样品的测序结果直方图。对于图12A-C，唯一性分子/条码是绘制于y轴上，并且以读数/bc分选的唯一性bc组合索引绘制于x轴上。

[0570]

表 3		
数目	基因	细胞类型
1	CD74	拉莫斯-特异性
2	CD79a	拉莫斯-特异性
3	IGJ	拉莫斯-特异性
4	TCL1A	拉莫斯-特异性
5	SEPT9	拉莫斯-特异性
6	CD27	拉莫斯-特异性
7	CD41	K562-特异性
8	GYPA	K562-特异性
9	GATA1	K562-特异性
10	GATA2	K562-特异性
11	HBG1	K562-特异性
12	GAPDH	共同的

[0571]

表 4			
	样品 1：仅 K562	样品 2:仅拉莫斯（Ramos）	样品 3：拉莫斯＋K562 混合物
拉莫斯细胞数目	0	5000	3750
K562 细胞数目	1000	0	2500
读数的总数目	717718	1329189	2399025
在恒定 1 中的 >＝8 匹配	657911	1201081	2026726
在恒定 2 中的 >＝8 匹配	581581	1071364	1513466

[0572]

在恒定1和2中的 >=8匹配	581508	1071153	1513102
在所有3个亚条码中的完全匹配	481564	862348	1248073
在基因(40 bp)中的完全匹配	283463	575713	1004338
有用读数%	39.50%	43.31%	41.86%
唯一性条码组合的总数目	8501	29647	28783
唯一性条码组合 > 30分子的数目	145	1072	768
捕获效率	0.145	0.2144	0.12288

[0573] 对于单细胞类型样品(例如,仅K562样品、仅拉莫斯样品),使用单细胞标记以确定拷贝数。图12D-E示出分别针对仅拉莫斯样品和仅K562样品,表3中列出的基因的拷贝数的图。对于图12D-E,分子数目/条码(bc)组合是绘制于y轴上,并且以分子总数目/bc组合分选的唯一性条码组合是绘制于x轴上。在图12D-E中示出的结果是基于来自以下珠的测序数据,这些珠具有>30的唯一性分子总数。这些结果证明,每珠每扩增子的分子比例匹配对该细胞类型的预期。对于仅K562细胞样品,分子数目的偏斜更严重并且它显现HBG1(在这种细胞类型中是高度丰富的)具有可变的拷贝数。然而,GAPDH拷贝数显现是恒定的,即使分子总数/珠是偏斜的。单独基因的拷贝数示于图12F-I中。对于图12F-G,该拷贝数表示为分别针对仅拉莫斯细胞和仅K562细胞而言的每珠或单细胞的拷贝。对于图12H-I,该拷贝数表示为分别针对仅拉莫斯细胞和仅K562细胞而言的每珠或单细胞的相对丰度。

[0574] 使用单细胞标记确定在K562+拉莫斯混合物样品中单细胞的细胞类型。对来自具有最丰富分子的100个唯一性条码组合的测序结果进行分析,以评价单细胞标记用以确定K562+拉莫斯混合物样品中的单细胞的细胞类型的功效。图12J-M示出对于具有100个唯一性条码组合的珠,唯一性分子数目/基因(y-轴)的图。在x-轴上的数字是指基因(参见表3)。图12J-M清楚地描绘了K562细胞和拉莫斯细胞的一般基因表达模式。图12N-O示出了描绘这两种细胞类型的基因表达谱的一般模式的两种珠的放大图。图12N示出K562样细胞的基因表达谱的一般模式,并且图12O示出拉莫斯样细胞的基因表达谱的一般模式。图12P示出基于768个珠的基因表达谱的主成分分析的结果的散点图,其中来自K562+拉莫斯混合物样品具有>30个分子/珠。组分1,其绘制于x轴上,将两个细胞类型分开。组分2,其绘制于y轴上,将具有高HBG1拷贝数和低HBG1拷贝数的K562细胞分开。在散点图上的每个点表示一个唯一性条码组合,它相当于一个珠或一个细胞。基于该主成分分析,409个珠对应于K562细胞,并

且347个珠对应于拉莫斯细胞。针对K562-样细胞类型和拉莫斯样细胞类型，确定来自表3的基因拷贝数。图12Q-R示出分别针对K562-样细胞(基于图12P的第一主组分的左边的珠)和拉莫斯样细胞(基于图12P的第一主组分的右边的珠)，每珠每扩增子的拷贝数的直方图。对于图12Q-R，每个bc组合对应的数目是在y-轴上，并且以分子总数/bc组合分选的唯一性条码组合是在x-轴上。图12S-T示出分别针对K562-样细胞(基于图12P的第一主组分的左边的珠)和拉莫斯样细胞(基于图12P的第一主组分的右边的珠)而言的每珠或单细胞的单独基因拷贝数。表5示出了对于单细胞和混合物样品而言的平均拷贝数/珠。

[0575]

表 5

	单细胞类型样品		K562 + 拉莫斯混合物样品	
基因	仅 K562	仅拉莫斯	K562 样	拉莫斯样
CD74	0.00	39.95	0.10	7.50
CD79a	0.02	30.97	0.84	18.88
IGJ	0.03	42.43	0.81	27.76
TCL1A	0.01	31.78	0.71	19.44
SEPT9	0.88	3.89	1.35	1.52
CD27	0.00	5.31	0.03	1.30
CD41	0.61	0.00	0.47	0.01
GYPA	1.92	0.00	0.73	0.02
GATA2	1.38	0.00	0.60	0.04
GATA1	0.94	0.00	1.04	0.04
HBG1	201.09	0.00	72.27	1.37

[0576]

GAPDH	51.77	39.13	44.94	13.53
GAPDH 读数冗余	2.04	1.47	7.67	7.22

[0577] 实例7.评价珠之间的交互作用

[0578] 在此实例中，对珠之间的交互作用进行评价。将包含小鼠EL4细胞和拉莫斯细胞的样品如下进行制备：

[0579]

	高密度	低密度
微孔数目	约10000	约10000
小鼠EL4细胞数目	2500	1500

拉莫斯细胞数目	3750	1500

[0580] 将样品的细胞悬浮液添加到微孔的顶部，并且允许细胞沉降到微孔阵列的孔中。将未由微孔阵列捕获的细胞在磷酸盐缓冲盐水(PBS)浴中洗去。将偶联寡核苷酸的珠(如通过实例1中所述的酶法分离-聚池合成方法制备的)添加至微孔阵列。偶联寡核苷酸的珠包括具有多个寡核苷酸的磁珠。在珠上的每个寡核苷酸包括5'胺、通用序列、细胞标记1、接头1、细胞标记2、接头2、细胞标记3、分子标记以及寡dT。对于在相同珠上的每个寡核苷酸，这些寡核苷酸的序列除了分子标记外是相同的。对于在不同珠上的寡核苷酸，细胞标记1、2、和3组合是不同的。在微孔阵列中，添加约5-6个珠/孔。允许珠沉降至孔中并且在PBS浴中洗去未捕获的珠。磁体放置该微孔阵列下方。将细胞通过添加冷裂解缓冲液裂解。将该阵列以及磁体放置在冷铝块上持续5分钟。来自裂解的细胞的mRNA与偶联到珠上的寡核苷酸杂交。将该阵列用过量的裂解缓冲液进行洗涤以除去未结合的mRNA。通过将磁体放置在微孔阵列顶部，从这些孔收回珠。将这些收回的珠进行洗涤。使用Superscript III，在50℃在旋转器上，在珠上进行cDNA合成持续50分钟。通过在旋转器上在37℃用ExoI处理30分钟，去除来自珠上的寡核苷酸的非延伸寡dT。对cDNA进行基因特异性PCR扩增。被选择用于该基因特异性PCR的基因是细胞类型特异性的，并且示于表6中。

[0581]

表 6

数目	基因	细胞类型
1	HS_CD74	人

[0582]

2	HS_CD79a	人
3	HS_IGJ	人
4	HS_TCL1A	人
5	HS_SEPT9	人
6	HS_CD27	人
7	HS_GAPDH	人
8	MM_B2M	小鼠
9	MM_ACTM	小鼠
10	MM_HPRT	小鼠
11	MM_SHDA	小鼠

[0583] 对PCR扩增产物进行测序。测序统计示于表7中。

[0584]

表 7		
	低密度	高密度
拉莫斯细胞数目	15000	3750
小鼠细胞数目	1500	2500
读数的总数目	2391780	4038217
在恒定 1 中的 >=8 匹配	2162945	3651643
在恒定 2 中的 >=8 匹配	1981835	3356493
在恒定 1 和 2 中的 >=8 匹配	1981626	3355787
在所有 3 个亚条码中的完全匹配	1645994	2790879
在基因 (40 bp) 中的完全匹配	1083013	2171930
有用读数%	45%	54%
唯一性条码组合的总数目	16695	36595
唯一性条码组合 >30 分子的数目	80	281
捕获效率	0.03	0.04

[0585] 对于高密度和低密度样品，确定针对100个具有最丰富的分子的唯一性条码组合的基因表达谱。基于测序结果产生基因表达谱。图13A示出对于来自高密度样品的100个唯一性条码组合中的35个的基因表达谱图。对于图13A，唯一性分子数目是在y-轴上，并且基因参考号是在x-轴上(关于对应于基因参考号的基因，参见表6)。图13A清楚地描绘了小鼠细胞和拉莫斯细胞的一般基因表达模式。图13B-13C示出分别基于高密度样品和低密度样品的基因表达谱的主成分分析的结果的散点图。组分1，其绘制于x轴上，将两个细胞类型分开。组分2，其绘制于y轴上，指示拉莫斯细胞群体内的基因表达可变性。在散点图上的每个点表示一个唯一性条码组合，它相当于一个珠或一个细胞。基于该高密度样品的主成分分析，144个珠对应于小鼠细胞并且132个珠对应于拉莫斯细胞。基于该低密度样品的主成分分析，52个珠对应于小鼠细胞并且27个珠对应于拉莫斯细胞。

[0586] 一旦确定细胞类型，则通过检测来自不同细胞类型中的表6的基因来对这些珠之间的交互作用进行评估。图13D-E描绘了针对分别来自高密度样品的拉莫斯样细胞和小鼠样细胞的以下图：读数/条码(bc)组合(y-轴)比对以分子总数目/bc组合分选的唯一性条码组合(x-轴)。图13F-G描绘了针对分别来自高密度样品的拉莫斯样细胞和小鼠样细胞的以下图：分子数目/bc组合的(y-轴)比对以分子总数目/bc组合分选的唯一性条码组合(x-轴)。图13H-I描绘了针对分别来自低密度样品的拉莫斯样细胞和小鼠样细胞的以下图：读数/条码(bc)组合(y-轴)比对以分子总数目/bc组合分选的唯一性条码组合(x-轴)。图13J-

K描绘了针对分别来自低密度样品的拉莫斯样细胞和小鼠样细胞的以下图:分子数目/bc组合(y-轴)比对以分子总数目/bc组合分选的唯一性条码组合(x-轴)。表8示出了针对低密度和高密度样品而言的每唯一性分子的平均倍数覆盖或读数冗余。

[0587]

表 8

	低密度		高密度	
基因	拉莫斯样细胞	小鼠样细胞	拉莫斯样细胞	小鼠样细胞
HS_CD74	29.75	3.17	23.75	2.15
HS_CD79a	47.2	4.09	42.30	2.67
HS_IGJ	29.65	1.39	30.23	2.4

[0588]

HS_TCL1A	45.74	2.26	39.00	4.13
HS_SEPT9	11.85	1.00	12.75	1.18
HS_CD27	37.99	1.00	32.12	1.10
HS_GAPDH	19.97	1.55	17.37	2.57
MM_B2M	1.21	31.98	3.05	31.48
MM_ACTM	1.05	29.08	1.90	28.38
MM_HPRT	1.02	39.96	1.03	43.65
MM_SHDA	1.00	39.60	1.02	29.60

[0589] 表8中的结果显示,在拉莫斯细胞中,与小鼠基因相比,对于人基因的每唯一性分子的平均倍数覆盖更高,并且反之亦然。

[0590] 作为对照,将小鼠和人细胞的混合物在一个管中裂解,转至用珠进行cDNA合成,并且对cDNA进行测序。如所预期,观察到大量的唯一性条码(bc)组合,并且大多数珠仅具有总计一至两个拷贝。

[0591] 这些结果证明在珠之间存在最小交互作用,并且该交互作用可以经生物信息学鉴定。

[0592] 实例8.产生单细胞核酸文库

[0593] 可以将在此披露的缀合寡核苷酸的支持物用于产生单细胞核酸文库。在此实例中,单细胞核酸文库是通过将细胞样品添加至具有缀合寡核苷酸的支持物的表面(例如,网格)来产生的。缀合寡核苷酸的支持物包括缀合至珠的多个寡核苷酸。寡核苷酸包括(a)细胞标记区,该区包括通过接头连接的至少两个不同区域;以及(b)分子标记区。在珠上的两个或更多个寡核苷酸包括相同的细胞标记区。在珠上的两个或更多个寡核苷酸包括两个或更多个不同的分子标记区。在两个或更多个不同珠上的两个或更多个寡核苷酸包括两个或

更多个不同的细胞标记区。因此，每个与缀合寡核苷酸的支持物相关联的细胞具有不同的细胞标记区。在细胞样品中的细胞的浓度是足够稀释的，以使得一个或更少细胞能够与在该表面上的一个缀合寡核苷酸的支持物相关联。使用裂解缓冲液将细胞裂解。来自细胞的mRNA与缀合寡核苷酸的支持物的寡核苷酸杂交。由此，所有来自细胞的mRNA均标记有包括相同细胞标记区的寡核苷酸。来自细胞的两个或更多个mRNA标记有包括两个或更多个不同分子标记区的两个或更多个寡核苷酸。将磁体施加到该表面上以从该表面纯化缀合寡核苷酸的固体支持物。可以单独地从该表面纯化缀合寡核苷酸的固体支持物。将与缀合寡核苷酸的固体支持物上的寡核苷酸杂交的mRNA进行逆转录以产生经标记的cDNA。经标记的cDNA包括该mRNA的反向互补物和该mRNA所杂交的寡核苷酸的拷贝。将经标记的cDNA通过PCR扩增以产生经标记的扩增子。可以通过限制性酶消化将经标记的cDNA和/或经标记的扩增子从珠去除。从经标记的扩增子产生来自该单细胞的核酸文库。

[0594] 可替代地，一起纯化缀合寡核苷酸的固体支持物。mRNA的逆转录可以在组合的缀合寡核苷酸的固体支持物上进行。因为来自不同细胞的mRNA标记有包括不同细胞标记区的寡核苷酸，这些细胞标记区可以用于确定经标记的cDNA或经标记的扩增子来源于哪个细胞。由此，可以产生来自多个细胞的核酸文库，其中作为经标记的扩增子的来源的细胞的身份可以通过细胞标记区来确定。

[0595] 单细胞核酸文库还可以通过在裂解细胞之前使这些细胞与试剂接触来产生。该试剂可以是抗原、药物、细胞、毒素、等等。因此，可以产生特化的单细胞核酸文库。可以使用对核酸文库的分析来产生单细胞药物表达谱。还可以从这些核酸文库确定单细胞水平上的信号转导通路。这些核酸文库还可以用于确定抗原对特异性细胞类型的作用。

[0596] 实例9.单细胞表达谱分析

[0597] 可以使用在此披露的缀合寡核苷酸的支持物来确定单细胞的表达谱。在此实例中，使包括细胞混合物的细胞样品与多个抗体接触。使用流式细胞术纯化细胞亚群。将该细胞亚群添加至微孔阵列。将多个缀合寡核苷酸的支持物添加至该微孔阵列。缀合寡核苷酸的支持物包括偶联至纳米粒子的多个寡核苷酸。寡核苷酸包括(a)细胞标记区，该区包括通过两个预先确定的序列连接的三个不同序列;以及(b)分子标记区。在纳米粒子上的两个或更多个寡核苷酸包括相同的细胞标记区。在纳米粒子上的两个或更多个寡核苷酸包括两个或更多个不同的分子标记区。在两个或更多个不同纳米粒子上的两个或更多个寡核苷酸包括两个或更多个不同的细胞标记区。因此，每个与缀合寡核苷酸的支持物相关联的细胞具有不同的细胞标记区。

[0598] 将磁体施加至微孔阵列，并且将与缀合寡核苷酸的支持物不相关联的细胞洗去。将包括裂解缓冲液的海绵放置在该微孔阵列的顶部，由此裂解细胞。

[0599] 来自裂解的细胞的mRNA与珠上的寡核苷酸杂交。将mRNA进行逆转录以产生经标记的cDNA。经标记的cDNA包括该mRNA的反向互补物和该mRNA所杂交的寡核苷酸的拷贝。将经标记的cDNA通过PCR扩增以产生经标记的扩增子。对经标记的扩增子进行测序。因为来自细胞的每个mRNA标记有相同的细胞标记，并且来自不同细胞的mRNA标记有不同的细胞标记，使用经标记的扩增子的序列信息来产生单细胞表达谱。

[0600] 实例10:通过单细胞测序进行免疫分型

[0601] 从受试者收集血液样品，并且从该血液样品分离外周血单核细胞(PMBC)。将PMBC

培养在RPMI1640培养基中，并且放置于培养箱中过夜。将PMBC在PBS中洗涤多次以去除血清。将约7000个PMBC沉积到具有32,400个孔的微孔阵列上。由此，在微孔阵列上的大多数孔不包含细胞，并且微孔阵列上的一些孔仅包含1个细胞。将缀合寡核苷酸的珠施加至该微孔阵列。每一个缀合寡核苷酸的珠包含附接至珠的约10亿个寡核苷酸。附接至珠的每个寡核苷酸包括5'胺、通用序列、三部分细胞标记(例如，通过两个接头连接的三个细胞标记区段)、分子标记、以及寡dT。每个珠包含唯一性三-部分细胞标记，该标记是三个细胞标记区段的唯一性组合的结果。在单一珠上的全部寡核苷酸包含相同的三部分细胞标记。来自不同珠的寡核苷酸包含不同的三部分细胞标记。每个孔包含1个或更少的缀合寡核苷酸的珠。将细胞裂解试剂施加至该微孔阵列，导致细胞的裂解。来自细胞的聚腺苷酸化分子(例如，mRNA)与来自缀合寡核苷酸的珠的寡核苷酸的寡dT序列杂交。在旋转器上，将与来自缀合寡核苷酸的珠的寡核苷酸杂交的聚腺苷酸化分子用SuperScript II在42℃进行逆转录90分钟。将来自缀合寡核苷酸的珠的寡核苷酸用作用于第一链cDNA合成的引物。在cDNA合成中掺入SMART寡聚物，以使得superscript II可以在其到达末端时将SMART寡聚物序列的互补物添加至cDNA的3'端。cDNA合成反应产生与未延伸寡核苷酸(例如，未附接至来自细胞的聚腺苷酸化分子的寡核苷酸)和延伸寡核苷酸(例如，附接至聚腺苷酸化分子的并且包括聚腺苷酸化分子/cDNA杂交体的寡核苷酸)缀合的珠。

[0602] 将珠组合，并且对包括聚腺苷酸化分子/cDNA杂交体的寡核苷酸进行扩增。进行多重PCR来从珠上的cDNA扩增一面板的98个基因(参见表9)。用于多重PCR的引物包括第一基因特异性引物(其被设计为位于距离mRNA的3'端大约500个碱基对处)和巢式基因特异性引物(其被设计为位于距离mRNA的3'端约300个碱基对处)。用于多重PCR的引物被设计为需要在该面板中引物的最后6个碱基的并非显著的互补性。如果检测到多重PCR引物的互补性，则手动更换引物。多重PCR反应包含以下步骤:1)15个循环的第一基因特异性PCR(KAPA多重混合物，每种引物50nM-第一基因特异性引物和通用引物，该通用引物与缀合寡核苷酸的珠的通用序列互补)、Ampure净化(0.7x珠与模板比率)、15个循环的巢式基因特异性PCR(KAPA多重混合物，每种引物50nM-巢式基因特异性引物和通用引物，该通用引物与缀合寡核苷酸的珠的通用序列互补)、Ampure净化(0.7x珠与模板比率)、8个循环的最终PCR以添加全长亿明达(Illumina)适配子(KAPA HiFi ReadyMix)、以及Ampure净化(1x珠与模板比率)。

[0603]

表 9.基因面板（Gene Panel）

细胞类型	基因	细胞类型	基因	细胞类型	基因	细胞类型	基因
B 细胞	PAX5	单核细胞	CD14	初始的	CD62L（SELL）	Th17	IL17A
	CD19	经典单核细胞	S100A12		CD45RA		IL17F
	CD20		CCR2	初始的 Th	THPOK/ZBTB7B		IL21
	BCMA/TNFRSF17		SELL/CD62L（L-选择素）	初始的 Tc	RUNX3		IL22
	BAFF	非经典单核细胞	CD16/FCGR3B	记忆细胞	CD45RO/PTPRC		CCL20
	TCL1A		CX3CR1		CD44		IL23R
	TACI		ITGAL	中心记忆 CD8+/	CCR7		RORa/RORA
初始	IGHD	常规	CD1b		TXK		RORgamat/RORC

[0604]

B	IGHM	髓样DC	FOXQ1	CD4+细胞	MBD2	滤泡T辅助细胞	OX40L/TNFSF4/CD252
B记忆细胞	CD27		CD209/DC-SIGN		BCL6		CXCR5
	CD38		CD1e	效应记忆CD8+/CD4+细胞	BLIMP1		SLAM/SLAMF1
	CD24		CCL17	Th1	CXCR3		ICOS
	AICDA		DTNA		IFNGR1		SAP/SH2D1A
	CD95	浆细胞样树突细胞（稀有）	CLEC4C/CD303		IL12RB2	活化的T细胞	CD69
过渡型B细胞	CD10	稀有髓样树突细胞（0.02%）	CD141/TM		IFNγ	活化的T细胞和B细胞	CD30
调节性B细胞（Breg）	IL10	NKT		Th2	IL33R/IL1RL1	Toll样受体（先天的）	TLR1
血浆	RASD1		PLZF/ZBTB16		IL4R		TLR2
	AMPD		SLAMF1		CCR4		TLR3

[0605]

	1						
	SDC1（CD138）	T细胞	CD3（CD3D）		CRTH2/PTGDR2		TLR4
NK	OSBPL5		CD3（CD3E）		IL4		TLR5
	CD56/NCAM1	细胞毒性T细胞	CD8（CD8A）		IL5		TLR6
	IGFBP7		CD8（CD8B）	调节性T细胞（Treg）	CD25		TLR7
	KIR2DS5		PRF1（穿孔素）		FOXP1		TLR8
	KIR2DS2		EOMES		TGFβ		TLR9
	RAB4B	辅助性T细胞	CD4		IL10		TLR10

[0606] 对这些扩增产物进行测序。使用Bowtie2,将具有150bp的序列读数与表9中列出的98个基因的完整mRNA序列进行比对。序列比对的结果(参见表10)证明该多重PCR反应产生高特异性产物。图14示出了描绘在X-轴上的基因和读数数目的log10的图。98个基因中的16个基因是不存在的。这些基因的不存在可归因于以下事实:一些基因靶向可不存在于此血液样品中的稀有细胞。总的来说,检测到来自该98基因面板的约84%的基因。

[0607]

表 10	
总计	6357075
比对 0 次	703616
比对恰好 1 次	5584201
比对 >1 次	69258
比对恰 1 次的%	88%

[0608] 表11示出了总体测序统计的结果。对于读数1,总读数1匹配标准要求与三部分细胞标记(例如细胞条码)的完美匹配以及与接头的至多1个错配。

[0609]

表 11	
总读数数目	6357075
总读数 1 匹配标准	4384245
读数 2 也比对	3943667
读数 2 比对%	89.95%
唯一性细胞 bc 数目	31129
读数计数/唯一性 bc > 100	3228
读数计数/唯一性 bc > 50	3721
有用读数%	62.04%

[0610] 图15A示出了每个三部分细胞标记(例如,细胞条码)的检测到的基因的分布图。图15B示出了每珠(表达该基因面板)的检测到的唯一性分子的分布图。

[0611] 进行细胞聚簇分析以确定测序结果是否可以用于分析基于细胞条码的细胞群体。使用SPADE(由诺兰实验室(Nolan lab)开发的用于CyTOF数据的最小生成树算法)来基于17个基因的存在/不存在对细胞聚簇。对于认为存在的基因,对于该基因的平均测序冗余必须大于5倍。在序列过滤之后,存在与大于20个唯一性分子相关联的约500个唯一性细胞条码(例如细胞标记)。每个唯一性细胞条码对应于单细胞。基于与唯一性细胞条码相关联的基因,将细胞聚簇为多个细胞类型。表12示出了可以用于明确鉴定细胞类型的基因列表。由此,与CD20、IGHM、TCL1A和CD24相关联的细胞条码被称作B-细胞,而与CD8A、CD3D、CD3E、CD4和CD62L相关联的细胞条码被称作T-细胞。将来自表9的剩余基因映射到细胞簇。图16描绘了基于与细胞条码相关联的基因的细胞簇。簇的大小与被指定到该簇的细胞数目是成比例的。示于图16中的结果证明,细胞和分子条码化的组合可以用于唯一地标记来自单细胞的分子的拷贝,这可以使得能够通过单细胞测序进行免疫分型。除了基于表12中列出的基因将PMBC聚簇到主要细胞类型中之外,该98基因面板还可以用于鉴定主要细胞类型的亚型的簇。表13示出了通过单细胞测序检测的每个主要细胞类型的频率。如在表13中所示,除了CD8+T细胞,每个细胞类型的百分比对应于正常细胞百分比范围。在PMBC样品中观察到略微更高百分比的CD8+T细胞。使用基于图16的细胞簇,使用来自该98基因面板的另外的基因的表达谱来进一步分析这些细胞簇。

[0612]

表 12	
主要细胞类型	基因
B 细胞	CD20、IGHM、TCL1A、CD24
T 细胞	CD8A、CD3D、CD3E、CD4、CD62L
NKT 细胞	ZBTB16
树突细胞	CD209
天然杀伤细胞	KIR2DS5、KIR2DS2、CD16
单核细胞	CD16、CD14、CCR2、S100A12、CD62L

[0613]

表 13			
细胞类型	细胞#	百分比	正常范围
单核细胞	67	13.3%	10%-30%
NK	85	16.9%	高达 15%
B	47	9.3%	高达 15%
CD8	210	41.7%	5%-30%
CD4	94	18.7%	25%-60%
总指定簇	503	100.0%	

[0614] 图17A-D示出单核细胞特异性标志物的分析。图17E示出了图16中描绘的细胞簇。图17A示出CD14的细胞表达谱，该CD14是单核细胞特异性标志物。“暖色”(例如红色)代表高基因表达，并且“冷色”(例如蓝色)代表低基因表达。如在图17A中所示，CD14高表达于单核细胞群体中并且在其他细胞类型中具有低至不表达。已知存在于单核细胞和NK两者中的CD16的细胞表达谱示于图17B中。如在图17B中所示，这些单核细胞和NK细胞簇具有CD16的高表达，而其他细胞类型具有低至不表达。已知CCR2和S100A12高表达于单核细胞中。CCR2和S100A12单核细胞-特异性表达也分别在图17C和D中所示的细胞表达谱中得以证实。然而，CCR2和S100A12的表达分成单核细胞的两个分支。其他细胞类型展现出CCR2和S100A12的低至不表达。

[0615] 图18A-B示出T细胞特异性标志物的分析。图18C示出了图16中描绘的细胞簇。图18A示出了作为CD3分子的链的CD3D的细胞表达谱。CD3是全T细胞标志物。图18A示出CD3D高

表达于CD8+T细胞的两个分支中，并且中度表达于CD8+T细胞的第三分支中。然而，CD3D不是高表达于CD4+T细胞中。同样，其他细胞类型具有CD3D的低至不表达。图18B示出了作为CD3分子的链的CD3E的细胞表达谱。图18B示出CD3D高表达于CD4+T细胞中。CD8+T细胞的不同分支展现出CD3D的高至中度表达。在其他细胞类型中观察到CD3D的少至不表达。

[0616] 图19A-B示出CD8+T细胞特异性标志物的分析。图19C示出了图16中描绘的细胞簇。图19A示出了作为CD8分子的链的CD8A的细胞表达谱。如图19A中所示，CD8+T细胞的不同分支具有不同水平的CD8A表达，其中一些分支具有CD8A的高表达，其他分支具有中度表达，并且一个分支展现出低至不表达。在CD16+NK细胞的一个分支中观察到高CD8A表达。已经在文献中报道高达80％的NK细胞表达CD8。在其他细胞类型中观察到CD8A的少至不表达。图19B示出了作为CD8模型的链的CD8B的细胞表达谱。如图19B中所示，CD8+T细胞的不同分支具有不同水平的CD8B表达，其中一个分支具有CD8B的高表达，一些分支具有中度表达，并且两个分支展现出低至不表达。还在CD16+NK细胞的一个分支中观察到高CD8B表达。在其他细胞类型中观察到CD8B的很少至不表达。

[0617] 图20A示出CD4+T细胞特异性标志物的分析。图20B示出图16中描绘的细胞簇。图20A示出CD4的表达谱。在CD4+T细胞簇中的细胞的一个亚群中观察到CD4的中度表达，并且在单核细胞簇的一个分支中观察到CD4的高表达。先前已经在文献中记载单核细胞也表达CD4。在CD8+T-细胞的分支和在NK细胞中观察到CD4的中度至低表达。在其他细胞类型中观察到CD4的低至不表达。

[0618] 图21A-D示出自然杀伤细胞(NK)细胞特异性标志物的分析。图21E示出了图16中描绘的细胞簇。图21A示出KIR2DS2的表达谱。所有的细胞类型展现出KIR2DS2的很少至不表达。图21B示出KIR2DS5的表达谱。已知杀伤细胞免疫球蛋白受体(KIR)表达于NK细胞和T细胞亚群中。在NK细胞的2个分支中观察到KIR2DS5的高表达，并且在NK细胞的一个分支中观察到KIR2DS5的中度至低表达。在CD8+T细胞的2个分支中观察到KIR2DS5的中度至高表达。在所有其他细胞类型中观察到KIR2DS5的低至不表达。已知OSBPL5和IGFBP7高表达于NK细胞中。图21C示出了OSBPL5的表达谱。OSBPL5高表达于NK细胞的一个分支中。在B细胞的分支中观察到OSBPL5的中度至低表达。在所有其他细胞类型中观察到OSBPL5的低至不表达。图21D示出了IGFPBP7的表达谱。在NK细胞的两个分支和单核细胞的一个分支中观察到IGFPBP7的高表达。在B细胞的一个分支中观察到IGFPBP7的中度表达。在所有其他细胞类型中观察到IGFPBP7的低至不表达。

[0619] 图22A-E示出B细胞特异性标志物的分析。图22F示出了图16中描绘的细胞簇。图22A示出IGHM CH4的表达谱。IGHM CH4高表达于B细胞的一个分支中，并且中度表达于B细胞的第二分支中。在其他细胞类型中观察到IGHM CH4的低至不表达。图22B示出PAX5的表达谱。PAX5高表达于B细胞的一个分支中。在所有其他细胞类型中观察到PAX5的低至不表达。图22C示出CD20的表达谱。CD20高表达于B细胞的一个分支中。在所有其他细胞类型中观察到CD20的低至不表达。图22D示出TCL1A的表达谱。在所有其他细胞类型中观察到TCL1A的低至不表达。图22E示出IGHD CH2的表达谱。IGHD CH2高表达于B细胞的一个分支中。在所有其他细胞类型中观察到IGHD CH2的低至不表达。

[0620] 图23A-F示出Toll样受体的分析。Toll样受体主要由单核细胞和一些B细胞表达。图23G示出图16中描绘的细胞簇。图23A示出TLR1的表达谱。单核细胞的一个分支展现出

TLR1的高表达，并且单核细胞的两个分支展现出TLR1的中度表达。在所有其他细胞类型中观察到TLR1的低至不表达。图23B示出TLR4的表达谱。单核细胞的一个分支展现出TLR4的高表达。在单核细胞的两个分支和NK细胞的一个分支中观察到了TLR4的中度表达。在所有其他细胞类型中观察到TLR4的低至不表达。图23C示出TLR7的表达谱。在单核细胞的一个分支中观察到了TLR7的高表达，并且在NK细胞的一个分支中观察到了TLR7的中度表达。在所有其他细胞类型中观察到TLR7的低至不表达。图23D示出TLR2的表达谱。在B细胞的一个分支中观察到TLR2的高表达。在所有其他细胞类型中观察到TLR2的低至不表达。图23E示出TLR3的表达谱。在B细胞的一个分支中观察到TLR3的高表达。在所有其他细胞类型中观察到TLR3的低至不表达。图23F示出了TLR8的表达谱。在单核细胞的三个分支中观察到TLR8的高表达。在单核细胞的两个分支和NK细胞的一个分支中观察到了TLR8的中度至低表达。在所有其他细胞类型中观察TLR8的低至不表达。

[0621] 这些结果证明大规模平行单细胞测序可成功地鉴定PMBC中的主要细胞类型。测序结果还确定了在FAC中用于鉴定细胞类型的一些细胞标志物不具有高mRNA表达(例如，NK细胞的CD56，B细胞的CD19)。此外，在基因面板中的许多基因跨多种细胞类型表达。这些表达谱可以用于主要细胞类型内的亚型细胞(例如，活化的细胞对比静息细胞等)。实例11：鉴定群体中的稀有细胞

[0622] 在这个实验中，使用大规模平行单细胞测序来从癌细胞和非癌细胞的混合物中鉴定癌细胞。将拉莫斯(伯基特淋巴瘤)细胞添加至从健康个体分离的CD19+B细胞的群体中。在混合群体中的拉莫斯细胞的浓度是约4％-5％。将约7000个正常B细胞和300个拉莫斯细胞沉积在具有25,200个孔的微孔阵列上。由此，在微孔阵列上的大多数孔不包含细胞，并且微孔阵列上的一些孔仅包含1个细胞。将缀合寡核苷酸的珠施加至该微孔阵列。每一个缀合寡核苷酸的珠包含附接至珠的约10亿个寡核苷酸。附接至珠的每个寡核苷酸包括5'胺、通用序列、三部分细胞标记(例如，通过两个接头连接的三个细胞标记区段)、分子标记、以及寡dT。每个珠包含唯一性三-部分细胞标记，该标记是三个细胞标记区段的唯一性组合的结果。在单一珠上的全部寡核苷酸包含相同的三部分细胞标记。来自不同珠的寡核苷酸包含不同的三部分细胞标记。每个孔包含1个或更少的缀合寡核苷酸的珠。将细胞裂解试剂施加至该微孔阵列，导致细胞的裂解。来自细胞的聚腺苷酸化分子(例如，mRNA)与来自缀合寡核苷酸的珠的寡核苷酸的寡dT序列杂交。在旋转器上，将与来自缀合寡核苷酸的珠的寡核苷酸杂交的聚腺苷酸化分子用SuperScript II在42℃进行逆转录90分钟。将来自缀合寡核苷酸的珠的寡核苷酸用作用于第一链cDNA合成的引物。在cDNA合成中掺入SMART寡聚物，以使得superscript II可以在其到达末端时将SMART寡聚物序列的互补物添加至cDNA的3'端。cDNA合成反应产生与未延伸寡核苷酸(例如，未附接至来自细胞的聚腺苷酸化分子的寡核苷酸)和延伸寡核苷酸(例如，附接至聚腺苷酸化分子的并且包括聚腺苷酸化分子/cDNA杂交体的寡核苷酸)缀合的珠。

[0623] 将珠组合，并且对包括聚腺苷酸化分子/cDNA杂交体的寡核苷酸进行扩增。进行多重PCR来从珠上的cDNA扩增一面板的111个基因。这111个基因代表B细胞的不同亚群的标志物。用于多重PCR的引物包括第一基因特异性引物(其被设计为位于距离mRNA的3'端大约500个碱基对处)和巢式基因特异性引物(其被设计为位于距离mRNA的3'端约300个碱基对处)。用于多重PCR的引物被设计为需要在该面板中引物的最后6个碱基的并非显著的互补

性。如果检测到多重PCR引物的互补性，则手动更换引物。多重PCR反应包含以下步骤：1)15个循环的第一基因特异性PCR(KAPA多重混合物，每种引物50nM-第一基因特异性引物和通用引物，该通用引物与缀合寡核苷酸的珠的通用序列互补)、Ampure净化(0.7x珠与模板比率)、15个循环的巢式基因特异性PCR(KAPA多重混合物，每种引物50nM-巢式基因特异性引物和通用引物，该通用引物与缀合寡核苷酸的珠的通用序列互补)、Ampure净化(0.7x珠与模板比率)、8个循环的最终PCR以添加全长亿明达(Illumina)适配子(KAPA HiFi ReadyMix)、以及Ampure净化(1x珠与模板比率)。

[0624] 对这些扩增产物进行测序。使用Bowtie2，将包含150bp的序列读数与这111个基因(表17)的完整mRNA序列进行比对。序列比对的结果(参见表14)证明该多重PCR反应产生高特异性产物。图24描绘了基因比对读数数目的log10的图。111个基因中的24个是不存在的。这些基因中的至少两个即RAG1和RAG2涉及VDJ重组并且应仅在前B细胞中存在，这两个基因应该是不存在的。少数不存在的基因是特异性针对血浆细胞，非常稀有地保存在冷冻细胞中。

[0625]

表 17.

CD19	AURKB	FOXP1	CCND3	TLR1	FOXP3	CXCL12	GNAI2
CD27	CD81	MCL1	IL12A	TLR2	LAG3	CCL3	RGS1
CD138	CD80	IFNB1	IFNG	TLR3	CD73	CCL14	CD5
CD38	CD23a	BLNK	TNFA	TLR4	CD70	CCL20	CD22
CD24	CD44	CD40LG	IL2	TLR5	CCR7	CCL18	PIK3CD
CD10	LEF1	IGBP1	IL4	TLR6	CD45RA	TCL1A	DOCK8

[0626]

CD95	CXCR5	IRF4	IL6	TLR7	PDCD1	TACI	CD11b
CD21	PRKCB	CD79a	BAFF	TLR8	MYC	AICDA	FCGR2B
CXCR3	PRKCD	LTA	IGHE	TLR9	CD25	FCRL4	CD72
CD40	CD20	HDAC5	IGHD	TLR10	FCAMR	BCL2	BCL11B
CD69	CD30	RAG1	IGHM	GAPDH	CCND2	FASLG	CD86
CD1c	CD30L	RAG2	IGHA	CD9	MKI67	BCL6	TBX21
IL10	BAFFR	CD1d	IGHG1	CD11c	IL21R	IGHG2	PRDM1
IL4R	CMRF-35H	TGFB1	IGHG4	IL6R	HLA-DRA	IGHG3	

[0627]

表 14	
总计	5711013
比对 0 次	504775
比对恰好 1 次	5203308
比对 >1 次	2930
比对恰 1 次的%	91.6%

[0628] 表15示出了总体测序统计的结果。对于读数1，总读数1匹配标准要求与三部分细胞标记(例如细胞条码)的完美匹配以及与接头的至多1个错配。

[0629]

表 15	
总读数数目	5711013
总读数 1 匹配标准	3795915
读数 2 也比对	3495392
读数 2 比对%	92%
唯一性细胞 bc 数目	40764
读数计数/唯一性 bc > 100	3313
读数计数/唯一性 bc > 50	4154
有用读数%	61%

[0630] 图25A-D示出对于两个基因而言，分子条码对读数数目或读数数目的log10的图。图25A示出对于CD79而言，分子条码(通过丰度分选)对读数数目的图。图25B示出对于CD79而言，分子条码(通过丰度分选)对读数数目的log10的图。图25C示出对于GAPDH而言，分子条码(通过丰度分选)对读数数目的图。图25D示出对于GAPDH而言，分子条码(通过丰度分选)对读数数目的log10的图。

[0631] 保留856个细胞用于分析。图26A示出了在该面板中表达的基因数目/细胞条码对比唯一性细胞条码数目/单细胞的图。图26B示出检测的唯一性分子数目/珠对比细胞数目频率/负载给定数目分子的唯一性细胞条码的直方图。细胞的一个小亚群显示明显更高数量的mRNA分子和从该111基因面板表达的基因的数目(参见图26A-B中的圆圈区)。图26C示出检测的唯一性GAPDH分子数目/珠对比细胞数目频率/负载给定数目分子的唯一性细胞条码的直方图。

[0632] 使用主成分分析(PCA)产生细胞的散点图。图27示出了856个细胞的散点图。PCA鉴定出与多数细胞相比具有不同基因表达谱的细胞的小亚群。细胞的该亚群包括18个细胞，

是分析的全部细胞的约2%。该百分比类似于添加到该群体中的拉莫斯细胞的百分比。

[0633] 拉莫斯细胞衍生自滤泡性B细胞并强烈表达B细胞分化标志物CD20、CD22、CD19、CD10和BCL6。拉莫斯细胞还表达IgM并且过表达c-myc。图28示出前100(就检测的分子的总数而言)的表达的热图。表达更高水平的mRNA的细胞的亚群(18个细胞)还强烈表达作为拉莫斯细胞的已知标志物(例如,CD10、Bcl-6、CD22、C-my、和IgM)的基因。

[0634] 这些结果证明大规模平行单细胞测序成功地鉴定出细胞悬浮液中的异常细胞类型的小亚群(低至2%)。大规模平行单细胞测序可用于癌症诊断(例如,活组织检查/循环肿瘤细胞)。因为癌细胞具有更大的尺寸,并且负载更多的mRNA,它们可以易于从正常细胞分化。

[0635] 实例12:使用RESOLVE进行gDNA靶标的大规模平行单细胞全基因组和多重扩增

[0636] 图29示出此实例的工作流程。如在图29所示,将细胞悬浮液施加到微孔阵列(2901)。细胞悬浮液中的细胞的数目小于微孔阵列中的孔的数目,这样使得将该细胞悬浮液施加到该微孔阵列导致微孔阵列中的一个孔包含1个或更少细胞。将缀合寡核苷酸的珠(2905)施加至该微孔阵列。缀合寡核苷酸的珠(2905)包含附接到寡核苷酸的珠(2910),该寡核苷酸包括5'胺(2915)、通用引物序列(2920)、细胞标记(2925)、分子标记(2930)和随机物(2935)。缀合寡核苷酸的珠包含约10亿寡核苷酸。寡聚核苷酸包含5'胺、通用引物序列、细胞标记、分子标记、以及随机物。在单一珠上的每个寡核苷酸包含相同的细胞标记。然而,在单一珠上的两个或更多个寡核苷酸可以包含两个或更多个不同的分子标记。珠可以包含具有相同分子标记的寡核苷酸的多个拷贝。

[0637] 在将缀合寡核苷酸的珠添加至该微孔阵列之后,将细胞裂解缓冲液施加至阵列表面。如在图29中所示,来自细胞的基因组DNA(2945)与缀合寡核苷酸的珠(2940)的随机物序列(2935)杂交。将中和缓冲液添加至阵列表面。将DNA聚合酶(例如,Phi29)和dNTP添加至阵列表面。随机物序列(2935)充当用于扩增基因组DNA的引物,由此产生缀合gDNA的珠(2555)。缀合gDNA的珠(2955)包括含有5'胺(2915)、通用引物序列(2920)、细胞标记(2925)、分子标记(2925)、随机物(2935)和基因组DNA拷贝(2955)的寡核苷酸。原始基因组DNA(2945)与随机物(2935)基因组DNA拷贝(2955)杂交。对于单一珠,有多个不同的基因组DNA分子附接至寡核苷酸。

[0638] 如在图29中所示,将来自孔的缀合gDNA的珠(2950)组合到埃彭道夫管(2960)中。将含有随机物的gDNA MDA混合物上的基因组DNA、dNTP和DNA聚合酶(例如,Phi29)添加至包含组合的缀合gDNA的珠的埃彭道夫管中。将经标记的基因组DNA进一步扩增以产生溶液中的经标记的扩增子(2965)。经标记的扩增子(2965)包括通用引物序列(2920)、细胞标记(2925)、分子标记(2930)、随机物(2935)、和基因组DNA拷贝(2955)。将经标记的扩增子剪切成约1kb或更小的较小碎片。可替代地,经标记的扩增子可以通过Tagmentation(Nextera)片段化。对经标记的扩增子进行剪切或片段化产生标记的片段(2980)和未标记的片段(2985)。标记的片段(2980)包含通用引物序列(2920)、细胞标记(2925)、分子标记(2930)、随机物(2935)、和基因组DNA拷贝(2955)片段。将适配子(2970,7975)添加到这些片段中。该通用引物序列可以用来通过杂交下拉(hybridization pulldown)或PCR(使用该通用引物序列和针对适配子(2970,2975)之一的引物)选择经标记的片段(2980)。

[0639] 可以对经标记的片段测序。可以使用包含细胞标记、分子标记和基因组片段的序

列的序列读数来鉴定来自细胞悬浮液的细胞群体。可以使用主成分分析基于已知细胞标志物来产生细胞散点图。可替代地，或另外地，可以使用SPADE来产生细胞簇图。可以使用计算机软件程序产生一个列表，该列表包括细胞标记和分子标记以及与该细胞标记相关联的基因组片段。

[0640] 实例13：大规模平行测序以鉴定异质群体中的细胞

[0641] 对于此实例的实验工作流程示于图30中。如在图30中所示，细胞的混合群体被随机分散在微孔阵列上。在此实例中，细胞的混合群体包括拉莫斯细胞和K562细胞的混合物。细胞悬浮液包括低浓度的细胞，以使得阵列中的每个微孔包含1个或更少的细胞。将细胞施加到该微孔阵列之后，将多个缀合寡核苷酸的珠随机分散到该微孔阵列上。寡核苷酸珠包含多个含有5'胺、通用引物序列，细胞标记、分子标记、以及寡dT的寡核苷酸。来自单一珠的多个寡核苷酸的细胞标记是相同的。单一珠可以包括含有相同分子标记的多个寡核苷酸。此外，单一珠可以包括含有不同分子标记的多个寡核苷酸。与第一珠缀合的寡核苷酸的细胞标记不同于与第二珠缀合的寡核苷酸的细胞标记。因此，细胞标记可以用于区分两个或更多个缀合寡核苷酸的珠。将细胞裂解，并且将来自单细胞的RNA分子附接至相同孔中的缀合寡核苷酸的珠。图30示出了RNA的聚A序列附接至寡核苷酸的寡dT序列。在来自单独细胞的RNA分子附接至相同孔中的缀合寡核苷酸的珠之后，将珠组合为单一样品。在该单一样品中的珠上进行cDNA合成反应。图30示出了cDNA合成的产物包括附接至寡核苷酸的珠，该寡核苷酸包含5'胺、通用引物序列、细胞标记、分子标记、寡dT和RNA分子拷贝。为了简单起见，在图30中仅描绘一个寡核苷酸，然而，在此实例中，每个缀合寡核苷酸的珠包括约10亿个寡核苷酸。如在图30中所示，使用与通用引物杂交序列杂交的通用引物以及与RNA分子拷贝杂交的基因特异性引物，对在单一样品中的珠进行多重PCR。将基因特异性引物设计为结合至来自表16中所示的基因面板(gene panel)的拉莫斯特异性基因或K562特异性基因。作为对照，还将GAPDH基因特异性引物用于该多重PCR反应中。最后，使用新一代测序来对扩增的产物进行测序。测序读数包括关于细胞标记、分子标记和基因的信息。使用主成分分析，基于关于细胞标记、分子标记和基因的序列信息构建细胞的散点图。类似于如何使用FAC分选细胞和如何将基于表面标志物的散点图用于划分细胞，使用细胞标记以鉴定来自单细胞的基因，并且使用分子标记确定该基因的量。然后使用该组合信息使基因表达谱与单独细胞相关。如在图31A中所示，使用细胞和分子标记进行单细胞大规模平行测序能够成功地鉴定混合的细胞群体中的两个细胞群体(K562和拉莫斯细胞)。

[0642]

表 16

基因	细胞	基因	细胞
CD74	拉莫斯特异性	CD41	K562 特异性
CD79a	拉莫斯特异性	GYPA	K562 特异性
IGJ	拉莫斯特异性	GATA2	K562 特异性
TCL1A	拉莫斯特异性	GATA1	K562 特异性
SEPT9	拉莫斯特异性	HBG1	K562 特异性
CD27	拉莫斯特异性	GAPDH	共同的

[0643] 实例14:大规模平行单细胞测序与主成分分析

[0644] 在此实例中,平行地,将来自单独细胞的mRNA分子用缀合寡核苷酸的珠进行随机地标记。PBMC分离自血液,并且在-80℃在RPMI1640加FBS和DMSO中冷冻。将PMBC解冻并且用PBS洗涤三次。将包括细胞类型(总计4000个细胞)混合物的PBMC样品随机施加至琼脂糖微孔阵列。该琼脂糖微孔阵列包含37,500个细胞。经由围绕微孔阵列的PDMS垫圈,将150,000个缀合寡核苷酸的珠的混合物随机施加至微孔阵列。缀合寡核苷酸的珠描绘在图1中。为了简单起见,仅示出附接至珠的一个寡核苷酸,然而,缀合寡核苷酸的珠包含约10亿个寡核苷酸。

[0645] 通过将微孔阵列放置于冷块上持续10分钟并且通过将裂解缓冲液施加到阵列表面,使细胞裂解。一旦在孔内的细胞裂解,将来自单细胞的mRNA分子经由寡dT序列附接至缀合寡核苷酸的珠上。将磁体施加到该阵列,并且将该阵列用洗涤缓冲液洗涤两次。

[0646] 将具有附接的mRNA分子的珠组合到埃彭道夫管中。将附接至珠的mRNA分子进行逆转录以产生cDNA。以下cDNA合成混合物的制备如下:

[0647]

组分	体积(uL)
水	8
dNTP(10mM)	2
5x第一链缓冲液	4
MgCl2	2.4

SuperRase In	1
SMART寡聚物(50uM)	0.4
0.1M DTT	1
100x BSA	0.2
SSII	1
总计	20

[0648] 将cDNA合成混合物添加至包含具有附接的mRNA分子的珠的埃彭道夫管中。在旋转器上,将该埃彭道夫管在40℃孵育90分钟。在珠上发生cDNA合成反应。90分钟后,将磁体施加到该管,并且将cDNA混合物去除并且用以下ExoI反应混合物替换:

[0649]

组分	体积(uL)
ExoI缓冲液	2
水	17
ExoI	1

[0650] 在旋转器上,将管在37℃孵育30分钟。然后将管转移至热循环仪在80℃持续15分钟。在将管在80℃孵育持续15分钟之后,将70微升的TE+吐温20添加到该管。将磁体施加到该管并去除缓冲液。然后将这些珠重悬浮于50微升的TE+吐温20中。

[0651] 通过实时PCR,使用以下扩增混合物扩增附接至珠的cDNA:

[0652]

组分	体积(uL)
2x iTaq混合物	10
GAPDH ILMN(10uM)	0.6
ILR2(10uM)	0.6
珠	2
水	6.8
总计	20

[0653] 对经标记的cDNA扩增子进行测序以检测细胞标记、分子索引、和基因。将测序读数与细胞标记、然后基因、并且最后分子标记进行比对。与4个或更多个基因相关联的或与10个或更多个唯一性转录物分子相关联的细胞标记被指定到细胞,其中对每个唯一性转录物分子进行不止一次的测序。使用对于所有检测的来自表9的基因的主成分分析来鉴定对数据中的变化具有最大贡献的基因集合。在该主成分分析中使用632个单细胞。基于测序结果,98个基因中有81个被检测。

[0654] 图32示出对于GAPDH表达的主成分分析图。如在图32中所示,基于主成分空间的定位观察到两个细胞簇。

[0655] 图33A-F示出对于单核细胞相关基因的主成分分析(PCA)。图33A示出了CD16的PCA。图33B示出了CCRvarA的PCA。图33C示出了CD14的PCA。图33D示出了S100A12的PCA。图33E示出了CD209的PCA。图33F示出了IFNGR1的PCA。

[0656] 图34A-B示出对于泛T细胞标志物(CD3)的主成分分析(PCA)。图34A示出了CD3D的PCA并且图34B示出了CD3E的PCA。

[0657] 图35A-E示出对于CD8T细胞相关基因的主成分分析(PCA)。图35A示出了CD8A的PCA。图35B示出了EOMES的PCA。图35C示出了CD8B的PCA。图35D示出了PRF1的PCA。图35E示出了RUNX3的PCA。

[0658] 图36A-C示出对于CD4T细胞相关基因的主成分分析(PCA)。图36A示出了CD4的PCA。图36B示出了CCR7的PCA。图36C示出了CD62L的PCA。

[0659] 图37A-F示出对于B细胞相关基因的主成分分析(PCA)。图37A示出了CD20的PCA。图37B示出了IGHD的PCA。图37C示出了IGHD的PCA。图37D示出了TCL1A的PCA。图37E示出了IGHM的PCA。图37F示出了CD24的PCA。

[0660] 图38A-C示出对于自然杀伤细胞相关基因的主成分分析(PCA)。图38A示出了KIR2DS5的PCA。图38B示出了CD16的PCA。图38C示出了CD62L的PCA。

[0661] 基于主成分分析,单核细胞和淋巴细胞在PC1上形成两个不同簇。B、T、和NK细胞形成另一个簇,留作沿着PC2的簇中的连续体。图39示出GAPDH表达的PCA分析,其中具有对细胞类型和细胞亚型的注释。图40描绘了一个热图,示出了在细胞之间的基因表达谱方面的相关性。沿着从左上角开始的对角线,这些细胞是单核细胞、初始CD4T细胞、初始CD8T细胞、细胞毒性CD8T细胞、NK细胞、以及B细胞。图41示出另一个形式的热图,证明了基因表达与细胞类型之间的相关性。图42示出热图,证明了基因之间的基因表达谱方面的相关性。

[0662] 实例15:通过数字基因表达细胞计数法揭露细胞异质性

[0663] 呈现一种基因表达细胞计数法的途径,该途径将新一代测序与单细胞的随机条码化组合。数千个细胞随机沉积到约150,000个微孔的阵列上。添加承载细胞-和转录物-条码捕获探针的珠的文库,以使得每个细胞被分配到具有唯一性细胞条码的珠旁边。在细胞裂解之后,将mRNA与珠杂交,并且聚池以用于逆转录、扩增、以及测序。当对带有条码的转录物进行计数并且指定给源细胞时,对每个细胞的数字基因表达谱进行重构。我们应用该技术以将人类造血系统解析到细胞亚群,并且表征免疫细胞对于体外刺激的异质性应答。此外,该方法的高灵敏度通过稀有细胞(例如抗原-特异性T细胞、以及在高背景的正常细胞中的肿瘤细胞)的检测得以证实。

[0664] 介绍

[0665] 理解在细胞的大集合中的细胞多样性和功能需要测量特异性基因或由单独细胞表达的蛋白。流式细胞术被良好建立以用于测量单细胞的蛋白质表达,但mRNA表达测量是典型地在整体样品中进行,模糊了单独的细胞贡献。虽然使用微量滴定板或商业的微流芯片进行单细胞mRNA表达测量近来已经被报道(1-5),这些途径是极低通量的并且难以放大规模。由于这些限制,目前的大多数研究被限制在探询的细胞数目和探索的条件数目两方面中。

[0666] 此处,我们已经开发了一种高度可放大规模的途径,该途径使得能够跨任意数目的基因对数千个单细胞进行常规数字基因表达谱分析。使用小规模的工程化和组合化学以大规模平行方式来用唯一性细胞条码标记细胞中的所有mRNA分子。此外,在细胞内的每个转录物拷贝标记有分子条码,从而允许绝对的数字基因表达测量(6)。将来自所有细胞的标记的mRNA分子进行聚池、扩增、并且测序。使用每个序列上的细胞条码和分子条码对每个细胞的数字基因表达谱进行重构。这一高度可放大规模的技术使得能够进行基因表达细胞计数法,我们称之为CytoSeq。我们已经将该技术应用至造血系统的多参数遗传分类,并且证

实了它用于研究细胞异质性和检测群体中的稀有细胞的用途。

[0667] 结果

[0668] CytoSeq

[0669] 在图43A中概括了该程序。首先将细胞悬浮液加载到具有高达150,000个微孔的微制作表面上。每个30微米直径的微孔具有约20皮升的体积。调节细胞数目以使得在10个或更多个孔中仅约1个接收一个细胞。细胞通过重力沉降在这些孔内。

[0670] 将磁珠加载到该微孔阵列至饱和,以使得珠部分地位于孔内的每个细胞的顶部、或邻近处。珠的大小被选择为使得每个微孔可以仅容纳一个珠。每个磁珠负载约十亿个具有图43B中描画的结构的寡核苷酸模板。每个寡核苷酸显示通用引发位点、随后为细胞标记、分子标记、以及寡(dT)的捕获序列。在每个珠上的所有寡核苷酸具有相同的细胞标记但包含各种分子标记。我们已经发明了组合的分离-聚池方法以合成具有近一百万多样性的珠。将两个单细胞用相同的细胞标记进行标记的概率是低的(在10^{-4}量级上),因为仅约10%的孔被单细胞占据。类似地,在单一珠上的分子标记的多样性是在10^{4}量级上,并且相同细胞中的相同基因的两个转录物分子被相同分子标记进行标签化的可能性也是低的。

[0671] 将裂解缓冲液施加到微孔阵列的表面上,并且扩散到这些微孔中。从细胞释放的具有聚(dA)尾的mRNA分子与珠上的寡核苷酸的3'端上的寡(dT)杂交。因为细胞邻近于珠,在裂解缓冲液的高盐条件和mRNA的高局部浓度(数十纳摩尔)下,mRNA分子被捕获在珠上。

[0672] 裂解和杂交后,使用磁体将所有珠从微孔阵列收集到管中。从此往后,在单一管中进行所有反应。使用常规方案(方法)在珠上进行cDNA合成。将源自每个细胞的cDNA分子共价附接至其相应的珠,每个在5'端上标记有细胞标记和分子标记。进行巢式多重PCR扩增感兴趣的基因(图55)。因为来自每个细胞的mRNA已经作为cDNA拷贝到珠上,可以对这些珠进行反复扩增和分析,例如,针对基因的不同集合。

[0673] 对扩增子的测序揭示了细胞标记、分子标记、以及基因一致性(图55)。计算分析将基于细胞标记的读数分类,并且将具有相同分子标记和基因序列的读数缩小到单一条目中以抑制任何扩增偏倚。使用分子标记使我们能够测量分子绝对数目/细胞/基因,并且因此允许直接跨可以经历不同测序深度的生物样品比较细胞表达水平。

[0674] 在控制的细胞混合物中相异的细胞类型的鉴定

[0675] 为了测量该方法的将两个细胞类型分开的能力,将K562和拉莫斯细胞的约1:1的混合物加载到具有10,000个孔的微孔阵列上。使用约6000个细胞捕获约1000个细胞。选择具有12个基因的面板,并且从珠扩增这些基因。该面板由5个特异性针对K562(骨髓性白血病)细胞的基因、6个特异性针对拉莫斯(滤泡性淋巴瘤)细胞的基因、以及管家基因GAPDH组成(表18)。在约1000个细胞被捕获在各自具有单一珠的10,000-孔阵列上的情况下,仅10%的珠应负载mRNA,并且理论上应该在测序数据中仅观察到1000个唯一性细胞标记的最大值。实际上,在数据过滤后(参见过滤标准方法),我们发现与显著数目的读数相关联的768个细胞标记。作为比较,我们在具有相似数目的细胞和珠的微量离心管中进行整体细胞裂解和mRNA捕获,并且观察了大量的细胞标记,其中大多数仅一个读数与每个细胞标记相关联。这证明了微孔阵列在限制来自单细胞的mRNA与相同孔中的珠杂交方面是有效的。

[0676] 使用主成分分析(PCA),对768个单细胞的每一个的基因表达谱进行聚簇(图31A)。第一主成分(PC)清楚地将单细胞基于细胞类型分为两个主要簇。贡献于第一主成分的正面

的基因是对于拉莫斯特异性的那些基因，贡献于相同主成分的负面的基因是对于K562特异性的那些基因。这种基于细胞特异性表达将细胞成功聚簇到组中显示孔间污染(如果有的话)是可忽略不计的。该第二主成分突出显示了在K562细胞内的胎血红蛋白(HBG1)的高度可变性，其已经在先前观察到(7)。

[0677]

表 18

基因	外侧引物	具有共同 5'侧翼序列的巢式引物
CD41	CCCCTGGAAGAA GATGATGA	CAGACGTGTGCTCTTCCGATCTTTCT CCAACAAGTTGCCTCC
GYPD	GAGGAAATGAAG CCAAACACA	CAGACGTGTGCTCTTCCGATCTAATC GTGACCTTAAAGGCCC
GATA1	TTAGCCACCTCAT GCCTTTC	CAGACGTGTGCTCTTCCGATCTCTAC TGTGGTGGCTCCGCT
GATA2	GGAGGAGGATTGT GCTGATG	CAGACGTGTGCTCTTCCGATCTGTGT CCGCATAAGAAAAAGAATC
HBG1	GCAAGAAGGTGC TGACTTCC	CAGACGTGTGCTCTTCCGATCTCTGC ATGTGGATCCTGAGAA
CD27	CTGCAGTCCCATC CTCTTGT	CAGACGTGTGCTCTTCCGATCTGATG AGGTGGAGAGTGGGAA
IGJ	GGACATAACAGAC TTGGAAGCA	CAGACGTGTGCTCTTCCGATCTCAAT CCATTTTGTAACTGAACCTT
TCL1A	AAGCCTCTGGGTC AGTGGT	CAGACGTGTGCTCTTCCGATCTTGG AAAAGGGATAGAGGTTGG
CD74	TAGACAGATCCCC GTTCCTG	CAGACGTGTGCTCTTCCGATCTACA GGGAGAAGGGATAACCC
SEPT9	CAGCATCCCAGCC TTGAG	CAGACGTGTGCTCTTCCGATCTCCTC AATGGCCTTTTGCTAC
CD79a	CCTCTAAACTGCC CCACCTC	CAGACGTGTGCTCTTCCGATCTCCTT AATCGCTGCCTCTAGG
GAPD H	CACATGGCCUCCA AGGAGUAA	CAGACGTGTGCTCTTCCGATCTCAG CAAGAGCACAAGAGGAA

[0678] 在另一个实验中，我们将拉莫斯(伯基特淋巴瘤)细胞以几个百分比加入来自健康个体的原初B细胞中。具有111个基因的面板(表22)被设计为代表不同状态的B细胞。对1198个单细胞进行分析。发现构成18个单细胞(约群体的1.5％)的一小组群体具有与其余相比不同的基因表达模式(图31B)。已知由该组优先表达的基因与伯基特淋巴瘤相关联，例如MYC和IgM、以及由滤泡性B细胞(其是伯基特淋巴瘤起源的B细胞亚群)特异性表达的B细胞

分化标志物(CD10、CD20、CD22、BCL6)(图31C和31D)。此外,这组细胞负载更高水平的CCND3和GAPDH、以及总体更高的mRNA含量,如通过基于分析分子索引检测到的唯一性mRNA分子的总数目确定的(图31B)。这个发现是与以下事实一致,淋巴瘤细胞在物理上比健康个体中的原初B细胞更大,并且它们快速增生并产生更大量的转录物。

[0679] 人PBMC中的多个细胞类型的同时鉴定

[0680] 虽然控制的实验涉及两个不同细胞类型的人工混合物,但大多数天然存在的生物样品包含具有众多细胞类型和状态的多种群体,在基因表达谱方面具有更细微的差别。一个突出实例是血液。我们进行一个实验,其中我们的目的在于通过测量对于每个主要细胞类型具有特异性的具有98个基因的面板(表19)的表达谱,同时地鉴定人外周血单核细胞(PBMC)中的所有的主要细胞类型,包括单核细胞、NK细胞、和不同的T细胞亚群和B细胞亚群。不像被多数限制于表面蛋白标志物的传统免疫分型,我们包括了编码细胞因子、转录因子、和除了表面蛋白外还有的具有各种细胞功能的胞内蛋白的基因。我们用PCA使用81个存在的基因来分析632个单一PBMC的数字基因表达谱(图32-39)。第一主成分清楚地将单核细胞和淋巴细胞分为两个正交簇,如在一个簇中CD16a、CD14、S100A12、和CCR2的表达以及在另一个簇中淋巴细胞相关基因的表达所证实的。淋巴细胞的不同亚型处于沿着第二主成分的连续体中,其中B细胞(表达IgM、IgD、TCL1A、CD20、CD24、PAX5)在一端,初始T细胞(表达CD4、CCR7、CD62L)在中间,并且细胞毒性T细胞(表达CD8A、CD8B、EOMES、PRF1)在另一端。表达杀伤细胞样免疫球蛋白受体CD16a和穿孔素(PRF1)的自然杀伤细胞位于单核细胞和细胞毒性T细胞之间的空间中。我们还观察到,细胞代谢指示物GAPDH以最高水平表达于单核细胞中并且在B细胞中表达最低,推测大多数B细胞处于静息。跨细胞的基因表达谱的相关分析重申了用PCA的观察结果,并且揭示了在每个主要细胞类型内的细胞的另外更小的亚群(图40A-B)。具有731个细胞的相同PBMC样品的重复实验产生了大为相似的分离以及细胞类型频率(图41)。

[0681]

表 22

基因	外侧引物	具有共同 5'侧翼序列的巢式引物
CD19	GCAGGGTCCCA GTCCTATG	CAGACGTGTGCTCTTCCGATCTCCAAT CATGAGGAAGATGCA

[0682]

CD27	TCCAGGAGGATTACCGAAAA	CAGACGTGTGCTCTTCCGATCTCCATCCAAGGGAGAGTGAGA
CD138	AATGGCAAAGGAAGGTGGAT	CAGACGTGTGCTCTTCCGATCTGCAGACACCTTGGACATCCT
CD38	AGATCTGAGCCAGTCGCTGT	CAGACGTGTGCTCTTCCGATCTTGGTGCAGAGCTGAAGATTTT
CD24	AAAAGTGGGCTTGATTCTGC	CAGACGTGTGCTCTTCCGATCTTTTTGTTCGCATGGTCACAC
CD10	ATATTCCTTTGGGCCTCTGC	CAGACGTGTGCTCTTCCGATCTTCAAGTTTGGGTCTGTGCTG
CD95	CCCCCGAAAATGTTCAATAA	CAGACGTGTGCTCTTCCGATCTTGCTCTTGTCATACCCCCA
CD21	TAGCTTCCTCCTCTGGTGGT	CAGACGTGTGCTCTTCCGATCTTTTGCCTTTCCATAATCACTCA
CXCR3	CTGGCTCTCCCCAATATCCT	CAGACGTGTGCTCTTCCGATCTGCTCTGAGGACTGCACCATT
CD40	GTGGTGTTGGGGTATGGTTT	CAGACGTGTGCTCTTCCGATCTATACACAGATGCCCATTGCA
CD69	AGACAGGTCCTTTTCGATGG	CAGACGTGTGCTCTTCCGATCTTGTGCAATATGTGATGTGGC
CD1c	TTGAGACAGGCACATACAGCTT	CAGACGTGTGCTCTTCCGATCTTTGCTTCCTCAATCTGTCCA
IL10	CCCCAACCACTTCATTCTTG	CAGACGTGTGCTCTTCCGATCTTTCAATTCCTCTGGGAATGTT
IL4R	TGCCTAGAGGTGCTCATTCA	CAGACGTGTGCTCTTCCGATCTGTTGATGCTGGAGGCAGAAT
IL21R	AGCCTGGGTCACAGATCAAG	CAGACGTGTGCTCTTCCGATCTAGGTAGGAGGGTGGATGGAG
IL6R	CCAGCACCAGGGAGTTTCTA	CAGACGTGTGCTCTTCCGATCTAGGAAAGGATTGGAACAGCA
CXCL12	GGGTTTCAGGTTCCAATCAG	CAGACGTGTGCTCTTCCGATCTTTTGTAACTTTTTGCAAGGCA
CCL3	GTGAGGAGTGGGTCCAGAAA	CAGACGTGTGCTCTTCCGATCTAGTGGGGAGGAGCAGGAG
CCL14	CCATTCCCTTCTTCCTCCTC	CAGACGTGTGCTCTTCCGATCTTACCTACAAGATCCCGCGTC

[0683]

CCL20	TTGGACATAGCCCAAGAACA	CAGACGTGTGCTCTTCCGATCTTGTGCCTCACTGGACTTGTC
CCL18	ACCTGAAGCTGAATGCCTGA	CAGACGTGTGCTCTTCCGATCTCTGGAGGCCACCTCTTCTAA
TCL1A	GGTAAACACGCCTGCAAAC	CAGACGTGTGCTCTTCCGATCTCAGGACTCAGAAGCCTCTGG
TACI	CAACAAAGCACAGTGTTAAATGAA	CAGACGTGTGCTCTTCCGATCTTGTGTCAGCTACTGCGGAAA
AICDA	TGAGCAGATCCACAGGAAAA	CAGACGTGTGCTCTTCCGATCTGAAATGGAGTCTCAAAGCTTCA
FCRL4	TCCCAACTACGCTGATTTGA	CAGACGTGTGCTCTTCCGATCTGACCAAAAGGAATGTGTGGG
BCL2	TGCAAGAGTGACAGTGGATTG	CAGACGTGTGCTCTTCCGATCTTCAACCAAGGTTTGCTTTTGT
FASLG	AGAGGCTGAAAGAGGCCAAT	CAGACGTGTGCTCTTCCGATCTAATATGGGTTGCATTTGGTCA
BCL6	AAATCTGCAGAAGGAAAAATGTG	CAGACGTGTGCTCTTCCGATCTAGTTTTCAATGATGGGCGAG
AURKB	GCTCAAGGGAGAGCTGAAGA	CAGACGTGTGCTCTTCCGATCTGACTACCTGCCCCCAGAGAT
CD81	GTGGCGTGTATGAGTGGAGA	CAGACGTGTGCTCTTCCGATCTCACTCGCCCAGAGACTCAG
CD80	GCACATCTCATGGCAGCTAA	CAGACGTGTGCTCTTCCGATCTGCTTCACAAACCTTGCTCCT
CD23a	ACATTTTCTGCCACCCAAAC	CAGACGTGTGCTCTTCCGATCTAACAGCACCCTCTCCAGATG
CD44	GCCTGGTAGAATTGGCTTTTC	CAGACGTGTGCTCTTCCGATCTTTTTGTAGCCAACATTCATTCAA
LEF1	CAATTGGCAGCCCTATTTCA	CAGACGTGTGCTCTTCCGATCTGTTCAGCAGACTGGTTTGCA
CXCR5	CCGTGAGGATGTCACTCAGA	CAGACGTGTGCTCTTCCGATCTACGAGGAAGCCCTAAGACGT
PRKCB	TTGAGCCTGGGGTGTAAGAC	CAGACGTGTGCTCTTCCGATCTGTCTTCCAGGATTCACGGTG

[0684]

PRKCD	GAGCACCTCCTGGAAGATTG	CAGACGTGTGCTCTTCCGATCTTAAGCACCAGTGGGACTGTG
CD20	TAGGAGCAGGCCTGAGAAAA	CAGACGTGTGCTCTTCCGATCTGATTCCTCTCCAAACCCATG
CD30	TGTTTTGGGGAAAGTTGGAG	CAGACGTGTGCTCTTCCGATCTCTGTTTGCCCAGTGTTTGTG
CD30L	TGCAACCCAACTGTGTGTTA	CAGACGTGTGCTCTTCCGATCTTTTCACCAACTGTTCTCTGAGC
BAFFR	GCCCTGAGCAACAATAGCAG	CAGACGTGTGCTCTTCCGATCTTTCAGCTCTTCACTCCAGCA
CMRF-35H	AGGAAAAGATGTGGCTCACG	CAGACGTGTGCTCTTCCGATCTGGAGTTGGGGAGAACTGTCA
PRDM1	TCGAATAATCCAGGGAAACC	CAGACGTGTGCTCTTCCGATCTACCAAAGCATCACGTTGACA
HLA-DRA	GGCTTTACAAAGCTGGCAAT	CAGACGTGTGCTCTTCCGATCTTATGCCTCTTCGATTGCTCC
GNAI2	CCTTGAGTGTGTCTGCGTGT	CAGACGTGTGCTCTTCCGATCTCCACAGAATTGGGTTCCAAG
RGS1	AACTGGGAAGGCCAGGTAAC	CAGACGTGTGCTCTTCCGATCTTGTTTTCAAATTGCCATTGC
CD5	CTTTCTCCACGCCATTTGAT	CAGACGTGTGCTCTTCCGATCTACTAGGATATGGGGTGGGCT
CD22	GGGATCTGCTCGTCATCATT	CAGACGTGTGCTCTTCCGATCTGTTTCTGCCTCTGAGGGAAA
PIK3CD	GCGTGCGCGTTATTTATTTA	CAGACGTGTGCTCTTCCGATCTTGTCTGGGGAAGGCAAGTTA
DOCK8	GCAGTCAGCCAGAAATCACA	CAGACGTGTGCTCTTCCGATCTTTTTCTCCTCTCTGGGACCA
CD11b	TGAAAAGTCTCCCTTTCCAGA	CAGACGTGTGCTCTTCCGATCTCCTTCAGACAGATTCCAGGC
FCGR2B	GGAGAGGAGAGATGGGGATT	CAGACGTGTGCTCTTCCGATCTGAGTGAGTGCCCCTTTTCTT
CD72	CTCATGCCAACAAGAACCTG	CAGACGTGTGCTCTTCCGATCTTGACCCACACCTGACACTTC
BCL11B	TCGTGGAACACAGGCAAAC	CAGACGTGTGCTCTTCCGATCTTTGCATTTGTACTGGCAAGG

[0685]

CD86	TCAAGGCAACCAGAGGAAAC	CAGACGTGTGCTCTTCCGATCTACTAAGGGATGGGGCAGTCT
TBX21	ACCTTTTCGTTGGCATGTGT	CAGACGTGTGCTCTTCCGATCTTCAGGGAAAGGACTCACCTG
FOXP1	ATGCTGAAGGCATTTCTTGG	CAGACGTGTGCTCTTCCGATCTCTGTGAGCATGGTGCTTCAT
MCL1	GAGGGGAGTGGTGGGTTTAT	CAGACGTGTGCTCTTCCGATCTCAAAAGGGAAAGGGAGGATT
IFNB1	AGGGGAAAACTCATGAGCAG	CAGACGTGTGCTCTTCCGATCTTCACTGTGCCTGGACCATAG
BLNK	TTGGGCAGAAAGAAAAATGG	CAGACGTGTGCTCTTCCGATCTCAAAAGATTCCACCAGACTGAA
CD40LG	CCTCCCCCAGTCTCTCTTCT	CAGACGTGTGCTCTTCCGATCTGAGTCAGGCCGTTGCTAGTC
IGBP1	GGCTGATCTTCCCACAACAC	CAGACGTGTGCTCTTCCGATCTACGAGGGCAAAGATGCTAAA
IRF4	ATTCCCGTGTTGCTTCAAAC	CAGACGTGTGCTCTTCCGATCTAGAACTGCCAGCAGGTAGGA
CD79a	CACTTCCCTGGGACATTCTC	CAGACGTGTGCTCTTCCGATCTCTCACTCTTCTCCAGGCCAG
LTA	TGATGTCTGTCTGGCTGAGG	CAGACGTGTGCTCTTCCGATCTCCACACACAGAGGAAGAGCA
HDAC5	CCAGCCTGTAGGAAACCAAA	CAGACGTGTGCTCTTCCGATCTCTCCTTCTATCTCCAGGGCC
RAG1	GGATGCAGGTGGTTTTTGAT	CAGACGTGTGCTCTTCCGATCTCATTGTACCCATTTTACATTTTCTT
RAG2	CAAACCTTAAACACCCAGAAGC	CAGACGTGTGCTCTTCCGATCTATAACAATTCGGCAGTTGGC
CD1d	GAACCAGTTTCCTCCTGTGC	CAGACGTGTGCTCTTCCGATCTAAGATGTGGAGGCTGTTGCT
TGFB1	GACTGCGGATCTCTGTGTCA	CAGACGTGTGCTCTTCCGATCTTCTGCACTATTCCTTTGCCC
CD9	TCAGTATGATCTTGTGCTGTGCT	CAGACGTGTGCTCTTCCGATCTTACCCATGAAGATTGGTGGG
CD11c	CACAGCATGAGAGGCTCTGT	CAGACGTGTGCTCTTCCGATCTTCTCAGTTCCGATTTCCCAG

[0686]

FOXP3	TCAGGATCTGAGGTCCCAAC	CAGACGTGTGCTCTTCCGATCTTCACCTGTGTATCTCACGCA
LAG3	AGAGCTGTCTAGCCCAGGTG	CAGACGTGTGCTCTTCCGATCTTGGTGTCCTTTCTCTGCTCC
CD73	CTTAACGTGGGAGTGGAACC	CAGACGTGTGCTCTTCCGATCTGTGTGCAAATGGCAGCTAGA
CD70	TCTCAGCTTCCACCAAGGTT	CAGACGTGTGCTCTTCCGATCTTCACTGGGACACTTTTGCCT
CCR7	CAGGGGAGAGTGTGGTGTTT	CAGACGTGTGCTCTTCCGATCTGACATGCACTCAGCTCTTGG
CD45RA	TGCATAGTTCCCATGTTAAATCC	CAGACGTGTGCTCTTCCGATCTTACCAGGAATGGATGTCGCT
PDCD1	ACATCCTACGGTCCCAAGGT	CAGACGTGTGCTCTTCCGATCTGCAGAAGTGCAGGCACCTA
MYC	TGCATGATCAAATGCAACCT	CAGACGTGTGCTCTTCCGATCTTTGGACTTTGGGCATAAAAGA
CD25	AAATCACGGCAGTTTTCAGC	CAGACGTGTGCTCTTCCGATCTCTCATCTGTGCACTCTCCCC
FCAMR	GTGGGAAGAGAAGCTGATGC	CAGACGTGTGCTCTTCCGATCTTCAAGCATTATCCACGTCCA
CCND2	TGTGATGCCATATCAAGTCCA	CAGACGTGTGCTCTTCCGATCTTCAGTGTATGCGAAAAGGTTTTT
MKI67	AGCCTCTCTTGGGCTTTCTT	CAGACGTGTGCTCTTCCGATCTGTTTTCCCTGCCTGGAACTT
CCND3	CTTTGCTGCTGAAGGCTCAT	CAGACGTGTGCTCTTCCGATCTACAAGTGGTGGTAACCCTGG
IL12A	TGCTTCCTAAAAAGCGAGGT	CAGACGTGTGCTCTTCCGATCTGAACTAGGGAGGGGGAAAGA
IFNG	GCAGCCAACCTAAGCAAGAT	CAGACGTGTGCTCTTCCGATCTATCCAGTTACTGCCGGTTTG
TNFA	GAATGCTGCAGGACTTGAGA	CAGACGTGTGCTCTTCCGATCTACTTCCTTGAGACACGGAGC
IL2	ACCCAGGGACTTAATCAGCA	CAGACGTGTGCTCTTCCGATCTGCTGATGAGACAGCAACCATT
IL4	GACATCTTTGCTGCCTCCA	CAGACGTGTGCTCTTCCGATCTATGAGAAGGACACTCGCTGC

[0687]

IL6	TTAAGGAGTTCC TGCAGTCCA	CAGACGTGTGCTCTTCCGATCTTCCAC TGGGCACAGAACTTA
BAFF	TCCTTCGCTTTG CTTGTCTT	CAGACGTGTGCTCTTCCGATCTAGGTG GAAAAATAGATGCCAGTC
IGHE	CCCGGAAGTCTA TGCGTTT	CAGACGTGTGCTCTTCCGATCTAGGAC ATCTCGGTGCAGTG
IGHD	TGTGTGAGGTGT CTGGCTTC	CAGACGTGTGCTCTTCCGATCTAGGAG CACCACGTTCTGG
IGHM	CCCGGAGAAGT ATGTGACCA	CAGACGTGTGCTCTTCCGATCTGTACT TCGCCCACAGCATC
IGHA	CTGAACGAGCT GGTGACG	CAGACGTGTGCTCTTCCGATCTAGTAC CTGACTTGGGCATCC
IGHG1	CAAGGGCCCAT CGGTCTT	CAGACGTGTGCTCTTCCGATCTTTGTG ACAAAACTCACACATGC
IGHG4	CAAGGGCCCAT CGGTCTT	CAGACGTGTGCTCTTCCGATCTCAAAT ATGGTCCCCCATGC
IGHG2	CAAGGGCCCAT CGGTCTT	CAGACGTGTGCTCTTCCGATCTGCAAA TGTTGTGTCGAGTGC
IGHG3	CAAGGGCCCAT CGGTCTT	CAGACGTGTGCTCTTCCGATCTACCCC ACTTGGTGACACAAC
TLR1	CCATTCCGCAGT ACTCCATT	CAGACGTGTGCTCTTCCGATCTAAGGA AAAGAGCAAACGTGG
TLR2	TTGGTTGACTTC ATGGATGC	CAGACGTGTGCTCTTCCGATCTGGAAA CAGCACAAATGAACTTAA
TLR3	CATCATGCAGTT CAACAAGC	CAGACGTGTGCTCTTCCGATCTATGCA CTCTGTTTGCGAAGA
TLR4	GGGTGTGTTTCC ATGTCTCA	CAGACGTGTGCTCTTCCGATCTTTGAA AGTGTGTGTGTCCGC
TLR5	TCAGGCTGTTGC ATGAAGAA	CAGACGTGTGCTCTTCCGATCTGTATG CCCTTGCTGGACCTA
TLR6	ATGCGCAGTAAA AACTCGTG	CAGACGTGTGCTCTTCCGATCTTACAG TTCCACGCTGAGCTG
TLR7	GCCTGTACTTTC AGCTGGGTA	CAGACGTGTGCTCTTCCGATCTAAGGT GTTTGTGCCATTTGG
TLR8	GGTGAGCTCTGA TTGCTTCA	CAGACGTGTGCTCTTCCGATCTTATCA GGAGGCAGGGATCAC

[0688]

TLR9	GACCGGGTCAG TGGTCTCT	CAGACGTGTGCTCTTCCGATCTGGTGA TCCTGAGCCCTGAC
TLR10	TGCAGTGAGCT GAGATCGAG	CAGACGTGTGCTCTTCCGATCTATGGA AAACATCCTCATGGC
GAPDH	CACATGGCCUCC AAGGAGUAA	CAGACGTGTGCTCTTCCGATCTCAGCA AGAGCACAAGAGGAA

[0689]

表 19		
基因	外侧引物	具有共同 5'侧翼序列的巢式引物
CD19	GCAGGGTCCCAG TCCTATG	CAGACGTGTGCTCTTCCGATCTCCAAT CATGAGGAAGATGCA
CD20	TAGGAGCAGGCC TGAGAAAA	CAGACGTGTGCTCTTCCGATCTGATTC CTCTCCAAACCCATG
BAFF	TCCTTCGCTTTG CTTGTCTT	CAGACGTGTGCTCTTCCGATCTAGGT GGAAAAATAGATGCCAGTC
TCL1A	GGTAAACACGCC TGCAAAC	CAGACGTGTGCTCTTCCGATCTCAGG ACTCAGAAGCCTCTGG
TACI	CAACAAAGCAC AGTGTTAAATGA A	CAGACGTGTGCTCTTCCGATCTTGTGT CAGCTACTGCGGAAA
IGHD	TGTGTGAGGTGT CTGGCTTC	CAGACGTGTGCTCTTCCGATCTAGGA GCACCACGTTCTGG
IGHM	CCCGGAGAAGTA TGTGACCA	CAGACGTGTGCTCTTCCGATCTGTACT TCGCCCACAGCATC
CD27	TCCAGGAGGATT ACCGAAAA	CAGACGTGTGCTCTTCCGATCTCCATC CAAGGGAGAGTGAGA
CD38	AGATCTGAGCCA GTCGCTGT	CAGACGTGTGCTCTTCCGATCTTGGT GCAGAGCTGAAGATTTT
CD24	AAAAGTGGGCTT GATTCTGC	CAGACGTGTGCTCTTCCGATCTTTTTG TTCGCATGGTCACAC
AICDA	TGAGCAGATCCA CAGGAAAA	CAGACGTGTGCTCTTCCGATCTGAAA TGGAGTCTCAAAGCTTCA
CD95	CCCCCGAAAATG TTCAATAA	CAGACGTGTGCTCTTCCGATCTTGCTC TTGTCATACCCCCA
CD10	ATATTCCTTTGGG	CAGACGTGTGCTCTTCCGATCTTCAA

[0690]

	CCTCTGC	GTTTGGGTCTGTGCTG
IL10	CCCCAACCACTTCATTCTTG	CAGACGTGTGCTCTTCCGATCTTTCAATTCCTCTGGGAATGTT
CD138	AATGGCAAAGGAAGGTGGAT	CAGACGTGTGCTCTTCCGATCTGCAGACACCTTGGACATCCT
CD45RA	TGCATAGTTCCCATGTTAAATCC	CAGACGTGTGCTCTTCCGATCTTACCAGGAATGGATGTCGCT
BCL6	AAATCTGCAGAAGGAAAAATGTG	CAGACGTGTGCTCTTCCGATCTAGTTTTCAATGATGGGCGAG
PRDM1	TCGAATAATCCAGGGAAACC	CAGACGTGTGCTCTTCCGATCTACCAAAGCATCACGTTGACA
CXCR3	CTGGCTCTCCCCAATATCCT	CAGACGTGTGCTCTTCCGATCTGCTCTGAGGACTGCACCATT
IFNG	GCAGCCAACCTAAGCAAGAT	CAGACGTGTGCTCTTCCGATCTATCCAGTTACTGCCGGTTTG
IL4R	TGCCTAGAGGTGCTCATTCA	CAGACGTGTGCTCTTCCGATCTGTTGATGCTGGAGGCAGAAT
IL4	GACATCTTTGCTGCCTCCA	CAGACGTGTGCTCTTCCGATCTATGAGAAGGACACTCGCTGC
CCL20	TTGGACATAGCCCAAGAACA	CAGACGTGTGCTCTTCCGATCTTGTGCCTCACTGGACTTGTC
CD25	AAATCACGGCAGTTTTCAGC	CAGACGTGTGCTCTTCCGATCTCTCATCTGTGCACTCTCCCC
FOXP1	ATGCTGAAGGCATTTCTTGG	CAGACGTGTGCTCTTCCGATCTCTGTGAGCATGGTGCTTCAT
TGFB1	GACTGCGGATCTCTGTGTCA	CAGACGTGTGCTCTTCCGATCTTCTGCACTATTCCTTTGCCC
CXCR5	CCGTGAGGATGTCACTCAGA	CAGACGTGTGCTCTTCCGATCTACGAGGAAGCCCTAAGACGT
CD69	AGACAGGTCCTTTTCGATGG	CAGACGTGTGCTCTTCCGATCTTGTGCAATATGTGATGTGGC
CD30	TGTTTTGGGGAAAGTTGGAG	CAGACGTGTGCTCTTCCGATCTCTGTTTGCCCAGTGTTTGTG
PDCD1	ACATCCTACGGTCCCAAGGT	CAGACGTGTGCTCTTCCGATCTGCAGAAGTGCAGGCACCTA
LAG3	AGAGCTGTCTAG	CAGACGTGTGCTCTTCCGATCTTGGT

[0691]

	CCCAGGTG	GTCCTTTCTCTGCTCC
PAX5	TGACGTGTGTTG CTTTTGTG	CAGACGTGTGCTCTTCCGATCTACTTG GGAGAAAACAGGGGT
TNFRSF 17	GCTTTCCACTCC CAGCTATG	CAGACGTGTGCTCTTCCGATCTTGCTT TGAGTGCTACGGAGA
RASD1	GGGGGAGGGAT GTGAAGTTA	CAGACGTGTGCTCTTCCGATCTATCTT GTCTGTGATTCCGGG
AMPD1	ACAGATGACCCA ATGCAATTC	CAGACGTGTGCTCTTCCGATCTGAGC ACCTGTGATATGTGCG
OSBPL5	AGACCGATGCAC AGTCTTCC	CAGACGTGTGCTCTTCCGATCTCTTCA CGTCTGGCCTCAGTC
CD56	GGAGCACTCAA GTGTGACGA	CAGACGTGTGCTCTTCCGATCTTTTTC TATGGAGCCTTCCGA
IGFBP7	CATCCAATTCCC AAGGACAG	CAGACGTGTGCTCTTCCGATCTGGTG AAGGTGCCGAGCTATA
KIR2DS 5	GCTCTTCCTCAA ACCACGAA	CAGACGTGTGCTCTTCCGATCTCACA CTCCTTTGCTTAGCCC
KIR2DS 2	TCCTCACACCAC GAATCTGA	CAGACGTGTGCTCTTCCGATCTCACTC CTTTGCTTAGCCCAC
RAB4B	CCAGCTCACCTG TTCTCCAG	CAGACGTGTGCTCTTCCGATCTGAATC CCGTACCTGCTGCT
CD14	CTAAAGGACTGC CAGCCAAG	CAGACGTGTGCTCTTCCGATCTATAAC CTGACACTGGACGGG
S100A12	CACATTCCTGTG CATTGAGG	CAGACGTGTGCTCTTCCGATCTATACT CAGTTCGGAAGGGGC
CCR2 变体 B	GAAGGAGGGAG ACATGAGCA	CAGACGTGTGCTCTTCCGATCTACTG GTCCTTAGCCCCATCT
CD62L	TCAGTTGGCTGA CTTCCACA	CAGACGTGTGCTCTTCCGATCTTTAGT TTGGGGGTTTTGCTG
CD16	TCTTGGCCAGGG TAGTAAGAA	CAGACGTGTGCTCTTCCGATCTGTCA GTTCCAATGAGGTGGG
CX3CR1	CGTCCAGACCTT GTTCACAC	CAGACGTGTGCTCTTCCGATCTCCAC AAATAGTGCTCGCTTTC
CD1b	TAGAGGGCCAG GACATCATC	CAGACGTGTGCTCTTCCGATCTTTGCT CCTTTTGCTATGCCT

[0692]

FOXQ1	TGCTATTGACCG ATGCTTCA	CAGACGTGTGCTCTTCCGATCTGCAA CGGGCTACAGCTTTAT
CD209	GCTCTTGTTCTT GCCGTTTT	CAGACGTGTGCTCTTCCGATCTGAGT CCCTCAGTGGAGCAAG
CD1e	CACAAGCACATT CATCTCTTCC	CAGACGTGTGCTCTTCCGATCTATTCA GGGCCAGCTTCATAA
CCL17	TACTTCAAGGGA GCCATTCC	CAGACGTGTGCTCTTCCGATCTTTTGT AACTGTGCAGGGCAG
DTNA	AGCAACGTGGA GTCAGTCTGT	CAGACGTGTGCTCTTCCGATCTCTCAC CTTCTCTTGCCTTGG
CLEC4C	TTATTTTCTGGG GCTGTCAGA	CAGACGTGTGCTCTTCCGATCTCATTC TGGCACTCAGGTGAA
ZBTB16	TGATCAAGCACC TGAGAACG	CAGACGTGTGCTCTTCCGATCTTACCA GTGCACCATCTGCAC
SLAMF1	TGCAAAACCCAG AAGCTAAAA	CAGACGTGTGCTCTTCCGATCTGTTCT GTGCAAATGGCATTC
CD3D	AGAGCTGTGTGG AGCTGGAT	CAGACGTGTGCTCTTCCGATCTGGAG TCTTCTGCTTTGCTGG
CD3E	GCCCTCTTGCCA GGATATTT	CAGACGTGTGCTCTTCCGATCTGCATG TAAGTTGTCCCCCAT
CD8A	CTGGCCTCTGCT CAACTAGC	CAGACGTGTGCTCTTCCGATCTATGGT ACAAGCAATGCCTGC
CD8B	CAGCCTCAAGGG GAAGGTAT	CAGACGTGTGCTCTTCCGATCTTGCTT AACCCATGGATCCTG
PRF1	CCCTGCAGTCAC AGCTACAC	CAGACGTGTGCTCTTCCGATCTTCAG GGCTGGTCTTTTAGGA
EOMES	TGGGATAATGTA AAACTGGTGCT	CAGACGTGTGCTCTTCCGATCTCATCC CCATGATATTTGGGA
CD4	AGCTAGCCTGAG AGGGAACC	CAGACGTGTGCTCTTCCGATCTTCCTC CAGACCATTCAGGAC
THPOK	GGCTCTGCCTTG CACTATTT	CAGACGTGTGCTCTTCCGATCTCTCTT CCTCCCTTCCATGC
RUNX3	TAAGGCCCAAAG TGGGTACA	CAGACGTGTGCTCTTCCGATCTTAGG AAGCACGAGGAAAGGA
CD45RO	ACCCTCTCTCCC TCCCTTTC	CAGACGTGTGCTCTTCCGATCTTAGTT GGCTATGCTGGCATG

[0693]

CD44	GCCTGGTAGAATTGGCTTTTC	CAGACGTGTGCTCTTCCGATCTTTTTGTAGCCAACATTCATTCAA
CCR7	CAGGGGAGAGTGTGGTGTTT	CAGACGTGTGCTCTTCCGATCTACTCAGCTCTTGGCTCCACT
TXK	ACATCAAGCTCCATTGTTTCG	CAGACGTGTGCTCTTCCGATCTTTTGCCTGCACTCTTTGTAGG
MBD2	GCCTGGCACGTAATAGCTTG	CAGACGTGTGCTCTTCCGATCTAGGAAAGAAATGCCCTTGGT
IFNGR1	GAGGATGTGTGGCATTTTCA	CAGACGTGTGCTCTTCCGATCTGGTTCCTAGGTGAGCAGGTG
IL12RB2	AGCAGGCTGTACACAGCAGA	CAGACGTGTGCTCTTCCGATCTGACACTAGGCACATTGGCTG
IL33R	ACTGTGCCCTCATCCAGAAC	CAGACGTGTGCTCTTCCGATCTAACGACGCCAAGGTGATACT
CCR4	TGGTGAAATGCAGAGTCAATG	CAGACGTGTGCTCTTCCGATCTTCAGGAGGAAGGCTTACACC
CRTH2	TGAATTTTGCTTGGTGGATG	CAGACGTGTGCTCTTCCGATCTTGTCAGTGGAAGAAGCAGATG
IL5	CAGTGAGAATGAGGGCCAAG	CAGACGTGTGCTCTTCCGATCTGAATGAGGGCCAAGAAAGAG
IL17A	AAAATGAAACCCTCCCCAAA	CAGACGTGTGCTCTTCCGATCTTCCTTTGGAGATTAAGGCCC
IL17F	CTGCATCAATGCTCAAGGAA	CAGACGTGTGCTCTTCCGATCTCCAAGGCTGCTCTGTTTCTT
IL21	AAATCAAGCTCCCAAGGTCA	CAGACGTGTGCTCTTCCGATCTTGTGAATGACTTGGTCCCTG
IL22	ATGCCCCAAAGCGATTTTT	CAGACGTGTGCTCTTCCGATCTCAAAGGAAACCAATGCCACT
IL23R	TCCCTCATTGAAAGATGCAA	CAGACGTGTGCTCTTCCGATCTTAGAATCATTAGGCCAGGCG
RORA	TGCAAGCCATTTATGGGAAT	CAGACGTGTGCTCTTCCGATCTCCTTGGGTTTTCTTTTCAATTC
RORC	ATTTCCATGGTGCTCCAGTC	CAGACGTGTGCTCTTCCGATCTAGAGAAGCAGAAGTCGCTCG
OX40L	CTGCTGGCCCTGTACCTG	CAGACGTGTGCTCTTCCGATCTCTCCACCCTGGCCAAGAT

[0694]

ICOS	TTCAGCTGACTTGGACAACCT	CAGACGTGTGCTCTTCCGATCTGGACAACCTGACTGGCTTTG
SH2D1A	GGGTGTTGGTGAACTTGGTT	CAGACGTGTGCTCTTCCGATCTTTTAATATGGATGCCGTGGG
CCR2 变体 A	GTTGCCCAGTGTGTTTCTGA	CAGACGTGTGCTCTTCCGATCTAACCAGGCAACTTGGGAACT
TLR1	CCATTCCGCAGTACTCCATT	CAGACGTGTGCTCTTCCGATCTAAGGAAAAGAGCAAACGTGG
TLR2	TTGGTTGACTTCATGGATGC	CAGACGTGTGCTCTTCCGATCTGGAAACAGCACAAATGAACTTAA
TLR3	CATCATGCAGTTCAACAAGC	CAGACGTGTGCTCTTCCGATCTATGCACTCTGTTTGCGAAGA
TLR4	GGGTGTGTTTCCATGTCTCA	CAGACGTGTGCTCTTCCGATCTTTGAAAGTGTGTGTGTCCGC
TLR5	TCAGGCTGTTGCATGAAGAA	CAGACGTGTGCTCTTCCGATCTGTATGCCCTTGCTGGACCTA
TLR6	ATGCGCAGTAAAAACTCGTG	CAGACGTGTGCTCTTCCGATCTTACAGTTCCACGCTGAGCTG
TLR7	GCCTGTACTTTCAGCTGGGTA	CAGACGTGTGCTCTTCCGATCTAAGGTGTTTGTGCCATTTGG
TLR8	GGTGAGCTCTGATTGCTTCA	CAGACGTGTGCTCTTCCGATCTTATCAGGAGGCAGGGATCAC
TLR9	GACCGGGTCAGTGGTCTCT	CAGACGTGTGCTCTTCCGATCTGGTGATCCTGAGCCCTGAC
TLR10	TGCAGTGAGCTGAGATCGAG	CAGACGTGTGCTCTTCCGATCTATGGAAAACATCCTCATGGC
GAPDH	CACATGGCCUCCAAGGAGUAA	CAGACGTGTGCTCTTCCGATCTCAGCAAGAGCACAAGAGGAA

[0695] 研究人类T细胞对体外刺激的应答的多样性

[0696] 当检查整体样品的基因表达模式时，由样品的细胞组成和在每个细胞类型或亚型中的每个基因的表达水平两者促成观察到的模式。这两个作用不能通过整体分析来揭露，但仅可用大规模单细胞分析来揭露。为了说明，我们使用我们的平台来研究人类T细胞对体外刺激的应答的可变性。

[0697] 我们通过从血液供体进行阴性选择来纯化CD3+T细胞，并且用抗-CD28/抗-CD3珠刺激它们持续6小时，并且采用经刺激的细胞和未经刺激的细胞的单独等分试样进行实验。我们设计了具有93个基因的面板(表20)，该面板涵盖由不同的T细胞亚群表达的表面蛋白、细胞因子、趋化因子、和效应分子。将总计3517个和1478个单细胞分别针对经刺激和未经刺

激的样品进行分析。

[0698]

表 20

基因	外侧引物	具有共同 5'侧翼序列的巢式引物
GAPDH	GACTTCAACAGCG ACACCCA	CAGACGTGTGCTCTTCCGATCTGCC CTCAACGACCACTTTGT
CD3D	GAAAACGCATCCT GGACCCA	CAGACGTGTGCTCTTCCGATCTTGA TGTCATTGCCACTCTGCT
CD3E	AAGTTGTCCCCCAT CCCAAA	CAGACGTGTGCTCTTCCGATCTCTG GGGATGGACTGGGTAAAT
CD8A	ACTGCTGTCCCAA ACATGCA	CAGACGTGTGCTCTTCCGATCTATG CCTGCCCATTGGAGAGAA
CD8B	CCACCATCTTTGCA GGTTGC	CAGACGTGTGCTCTTCCGATCTGCT GTCCAGTTCCCAGAAGG
CD4	CTGGGAGAGGGGG TAGCTAG	CAGACGTGTGCTCTTCCGATCTACC ACTTCCCTCAGTCCCAA
FOXP3	ACAGAAGCAGCGT CAGTACC	CAGACGTGTGCTCTTCCGATCTGGG TCTCTTGAGTCCCGTG
CCR7	GGGGAGAGTGTGG TGTTTCC	CAGACGTGTGCTCTTCCGATCTCTC TTGGCTCCACTGGGATG
CD5	ATCAATGGTCCAA GCCGCAT	CAGACGTGTGCTCTTCCGATCTAGG TCACAGATCTTCCCCCG
IL32	CTTTCCAGTCCTAC GGAGCC	CAGACGTGTGCTCTTCCGATCTTGC TCTGAACCCCAATCCTC
CD28	ACCATCACAGGCA TGTTCCT	CAGACGTGTGCTCTTCCGATCTTGT AGATGACCTGGCTTGCC
SELL	GCATCTCATGAGTG CCAAGC	CAGACGTGTGCTCTTCCGATCTCCT GCCCCCAGACCTTTTATC
CD27	TGCAGAGCCTTGT CGTTACA	CAGACGTGTGCTCTTCCGATCTCGT GACAGAGTGCCTTTTCG

[0699]

GZMB	AGGTGAAGATGACAGTGCAGG	CAGACGTGTGCTCTTCCGATCTAGGCCCTCTTGTGTGTAACA
GZMA	GGAACCATGTGCCAAGTTGC	CAGACGTGTGCTCTTCCGATCTCCTTTGTTGTGCGAGGGTGT
GZMH	AGTGTTGCTGACAGTGCAGA	CAGACGTGTGCTCTTCCGATCTCCAAAGAAGACACAGACCGGT
GZMK	TTGCCACAAAGCCTGGAATC	CAGACGTGTGCTCTTCCGATCTAAAGCAACCTTGTCCCGCCT
PRF1	GGAGTCCAGCGAATGACGTC	CAGACGTGTGCTCTTCCGATCTCATGGCCACGTTGTCATTGT
NKG2D	CAACACCCAGGGGATCAGTG	CAGACGTGTGCTCTTCCGATCTCCACCCTCCACAGGAAATTG
LAG3	AGCTGTACCAGGGGGAGAG	CAGACGTGTGCTCTTCCGATCTCTTTGGAGAAGACAGTGGCGA
CD160	GGAAGACAGCCAGATCCAGTG	CAGACGTGTGCTCTTCCGATCTTTGTGCAGACCAAGAGCACC
CD244	GGGCTGAGAATGAGGCAGTT	CAGACGTGTGCTCTTCCGATCTGGAAAGCGACAAGGGTGAAC
EOMES	ACTTAACAGCTGCAGGGGC	CAGACGTGTGCTCTTCCGATCTACTAACTTGAACCGTGTTTAAGG
TBX21	TTATAACCATCAGCCCGCCA	CAGACGTGTGCTCTTCCGATCTAGAAAAGGGGCTGGAAAGGG
PRDM1	ACCAAAGCATCACGTTGACAT	CAGACGTGTGCTCTTCCGATCTACATGTGAATGTTGAGCCCA
IRF4	CTCTTCAGCATCCCCCGTAC	CAGACGTGTGCTCTTCCGATCTGCCCCCAAATGAAAGCTTGA
ZNF683	GGAGAGCGTCCATTCCAGTG	CAGACGTGTGCTCTTCCGATCTATCCACCTGAAGCTGCACC
ZBED2	AATGTACCAGCCAGTCAGCG	CAGACGTGTGCTCTTCCGATCTGGTTTTGGTGGAGCTGACGA
CD30	TTTACTCATCGGGCAGCCAC	CAGACGTGTGCTCTTCCGATCTTGTTTGCCCAGTGTTTGTGC
CD69	GCTGTAGACAGGTCCTTTTCG	CAGACGTGTGCTCTTCCGATCTAGTGTTGGAAAATGTGCAATATGTG
HLA-DRA	GGGTCTGGTGGGCATCATTA	CAGACGTGTGCTCTTCCGATCTGCCTCTTCGATTGCTCCGTA

[0700]

CD38	AGGTCAATGCCAGAGACGGA	CAGACGTGTGCTCTTCCGATCTATCAGCATACCTTTATTGTGATCTATC
TNFRSF9	TGGCATGTGAGTCATTGCTC	CAGACGTGTGCTCTTCCGATCTTTTTGATGTGAGGGGCGGAT
MKI67	TACTTTTTCGCCTCCCAGGG	CAGACGTGTGCTCTTCCGATCTTCCTGCCCCACCAAGATCAT
BIRC5	TGCCACGGCCTTTCCTTAAA	CAGACGTGTGCTCTTCCGATCTTTGTCTAAGTGCAACCGCCT
FOSL1	CTCCTGACAGAAGGTGCCAC	CAGACGTGTGCTCTTCCGATCTGGTGATTGGACCAGGCCATT
MCL1	GACTGGCTACGTAGTTCGGG	CAGACGTGTGCTCTTCCGATCTTTTGCTTAGAAGGATGGCGC
MYC	AGCTACGGAACTCTTGTGCG	CAGACGTGTGCTCTTCCGATCTCAACCTTGGCTGAGTCTTGA
TYMS	TCAGTCTTTAGGGGTTGGGC	CAGACGTGTGCTCTTCCGATCTATGTGCATTTCAATCCCACGTAC
CDCA7	CCAGTCTAGTTTCTGGGCAGG	CAGACGTGTGCTCTTCCGATCTATGTAAACCATTGCTGTGCCATT
UHRF1	CCAGTTCTTCCTGACACCGG	CAGACGTGTGCTCTTCCGATCTCCAAAGTTTGCAGCCTATACC
SAP30	ACCAACCAGACCAGGACTTA	CAGACGTGTGCTCTTCCGATCTTCACTAGGAGACGTGGAATTG
CX3CR1	CACCCGTCCAGACCTTGTT	CAGACGTGTGCTCTTCCGATCTTGTTTTCCTCTTAACGTTAGACCAC
BCL2	TGCAAGAGTGACAGTGGATTG	CAGACGTGTGCTCTTCCGATCTGCTGATATTCTGCAACACTGTACA
BCL6	TGTCCTCACGGTGCCTTTT	CAGACGTGTGCTCTTCCGATCTGTAGGCAGACACAGGGACTT
FASLG	CCTCAAGGGGGACTGTCTTTC	CAGACGTGTGCTCTTCCGATCTGCATATCCTGAGCCATCGGT
FAS	ATTGCTGGTAGAGACCCCCA	CAGACGTGTGCTCTTCCGATCTCCCCCATTTCCCCGATGT
CCL4	CCCAGCCAGCTGTGGTATTC	CAGACGTGTGCTCTTCCGATCTTGGAACTGAACTGAGCTGCT
IFNG	CTAGGCAGCCAACCTAAGCA	CAGACGTGTGCTCTTCCGATCTCCTGCAATCTGAGCCAGTGC

[0701]

TNF	AGTGGACCTTAGGCCTTCCT	CAGACGTGTGCTCTTCCGATCTGGCTCAGACATGTTTTCCGTG
IL2	TCACTTAAGACCCAGGGACTT	CAGACGTGTGCTCTTCCGATCTAAGCATCATCTCAACACTGACTT
IL4	ACCATGAGAAGGACACTCGC	CAGACGTGTGCTCTTCCGATCTCGGGCTTGAATTCCTGTCCT
IL6	CGGCAAATGTAGCATGGGC	CAGACGTGTGCTCTTCCGATCTGGAAAGTGGCTATGCAGTTTG
IL1A	GGCATCCTCCACAATAGCAGA	CAGACGTGTGCTCTTCCGATCTGCATTTTGGTCCAAGTTGTGC
IL1B	CTTAAAGCCCGCCTGACAGA	CAGACGTGTGCTCTTCCGATCTACATTCTGATGAGCAACCGC
IL3	ACAGACGACTTTGAGCCTCG	CAGACGTGTGCTCTTCCGATCTATTTCACCTTTTCCTGCGGC
IL13	GGAGCCAAGGGTTCAGAGAC	CAGACGTGTGCTCTTCCGATCTTGCTACCTCACTGGGGTCCT
IL31	GGCCATCTCTTCCTTTCGGA	CAGACGTGTGCTCTTCCGATCTGTGTGGGAACTCTGCCGTG
IL24	CTCACCCCATCATCCCTTTCC	CAGACGTGTGCTCTTCCGATCTGCCCAGTGAGACTGTGTTGT
IL26	TACTGACGGCATGTTAGGTG	CAGACGTGTGCTCTTCCGATCTTGTGTGTGGAGTGGGATGTG
LTA	AGGCAGGGAGGGGACTATTT	CAGACGTGTGCTCTTCCGATCTGGAGAAACAGAGACAGGCCC
IL5	GCAGTGAGAATGAGGGCCA	CAGACGTGTGCTCTTCCGATCTAGGCATACTGACACTTTGCC
CSF2	AGCCAGTCCAGGAGTGAGAC	CAGACGTGTGCTCTTCCGATCTGGCCACACTGACCCTGATAC
IL21	CCCAAGGTCAAGATCGCCAC	CAGACGTGTGCTCTTCCGATCTCTGCCAGCTCCAGAAGATGT
IL22	TGGGAAGCCAAACTCCATCAT	CAGACGTGTGCTCTTCCGATCTGGAAACCAATGCCACTTTTGT
IL17A	GCCTTCAAGACTGAACACCGA	CAGACGTGTGCTCTTCCGATCTGCCCCTCAGAGATCAACAGAC
IL17F	TTGGAGAAGGTGCTGGTGAC	CAGACGTGTGCTCTTCCGATCTCTTACCCAGTGCTCTGCAAC

[0702]

TGFB1	TATTCCTTTGCCCGGCATCA	CAGACGTGTGCTCTTCCGATCTACCTTGGGCACTGTTGAAGT
CCL20	ACTTGCACATCATGGAGGGT	CAGACGTGTGCTCTTCCGATCTTCCATAAGCTATTTTGGTTTAGTGC
IL12A	GGTCCCTCCAAACCGTTGTC	CAGACGTGTGCTCTTCCGATCTGAACTAGGGAGGGGGAAAGAAG
CXCL12	TGGGAGTTGATCGCCTTTCC	CAGACGTGTGCTCTTCCGATCTCTCATTCTGAAGGAGCCCCAT
CCL3	TGGACTGGTTGTTGCCAAAC	CAGACGTGTGCTCTTCCGATCTCTCTGAGAGTTCCCCTGTCC
CCL14	TTCCTCCTCATCACCATCGC	CAGACGTGTGCTCTTCCGATCTCTTACCACCCCTCAGAGTGC
CCL18	GAAGCTGAATGCCTGAGGGG	CAGACGTGTGCTCTTCCGATCTGTCCCATCTGCTATGCCCA
CCL17	GAGTGCTGCCTGGAGTACTT	CAGACGTGTGCTCTTCCGATCTCTCACCCCAGACTCCTGACT
IL12B	GCTATGGTGAGCCGTGATTG	CAGACGTGTGCTCTTCCGATCTTCCTCACCCCCACCTCTCTA
CXCR3	GACCTCAGAGGCCTCCTACT	CAGACGTGTGCTCTTCCGATCTCCAATATCCTCGCTCCCGG
IL33R	TTCAGGACTCCCTCCAGCAT	CAGACGTGTGCTCTTCCGATCTAGGTACCAAATGCCTGTGCC
IL4R	TGAACTTCAGGGAGGGTGGT	CAGACGTGTGCTCTTCCGATCTTCCTCGTATGCATGGAACCC
CCR4	CCAAAGGGAAGAGTGCAGGG	CAGACGTGTGCTCTTCCGATCTATTCTGTATAACACTCATATCTTTGCC
IL23R	AGAATCATTAGGCCAGGCGTG	CAGACGTGTGCTCTTCCGATCTCTGGCCAATATGCTGAAACCC
IL21R	ATTTGAGGCTGCAGTGAGCT	CAGACGTGTGCTCTTCCGATCTAGACAAGAGCTGGCTCACCT
CXCR5	CCTCCCCAGCCTTTGATCAG	CAGACGTGTGCTCTTCCGATCTTCCTCGCAAGCTGGGTAATC
IL6R	CCAGCACCAGGGAGTTTCTA	CAGACGTGTGCTCTTCCGATCTACAGCATGTCACAAGGCTGT
CXCL13	AGGCAGATGGAACTTGAGCC	CAGACGTGTGCTCTTCCGATCTGCATTCGAAGATCCCCAGACTT

[0703]

LIF	TCCCCATCGTCCTCCTTGTC	CAGACGTGTGCTCTTCCGATCTTTGCCGGCTCTCCAGAGTA
PTPRCv1（CD45RA）	GTTCCCATGTTAAATCCCATTCAT	CAGACGTGTGCTCTTCCGATCTTACCAGGAATGGATGTCGCTAATCA
PTPRCv2（CD45RO）	ACCCTCTCTCCCTCCCTTTC	CAGACGTGTGCTCTTCCGATCTTAGTTGGCTATGCTGGCATG
IL10	CCCCAACCACTTCATTCTTG	CAGACGTGTGCTCTTCCGATCTTTCAATTCCTCTGGGAATGTT
CD40LG	CCTCCCCCAGTCTCTCTTCT	CAGACGTGTGCTCTTCCGATCTGAGTCAGGCCGTTGCTAGTC

[0704] 在未刺激的样品中，PCA分析揭示出两个主要的细胞亚群。更近对在每个亚群中富集的基因的观察显示一个亚群代表表达CD8A、CD8B、NKG2D、GZMA、GZMH、GZMK、和EOMES的CD8+细胞，并且另一个亚群代表表达CD4、CCR7和SELL的CD4+细胞(图44A和图45)。

[0705] 在经刺激的样品中，细胞的两个分支立即清楚示于PCA图上(图44B和图46A-D)。该第一主成分代表在IFNG、TNF、CD69、和GAPDH的表达的不同水平方面单独细胞对刺激剂的应答程度。CCL3、CCL4、和GZMB(它们是与细胞毒性T细胞相关联的细胞因子和效应分子)以及与疲乏细胞相关联的标志物LAG3的表达被定位至上分支中的细胞。IL2、LTA、CD40LG、和CCL20(它们是与辅助T细胞相关联的细胞因子)的表达被定位至下分支。其他已经已知在活化的T细胞中上调的基因(包括ZBED2、IL4R、PRDM1、TBX21、MYC、FOSL1、CSF2、TNFRSF9、BCL2和FASLG)以不同程度表达于小数量的细胞中(图46A-D)。这些细胞因子、效应分子、和转录因子中的大多数不被未经刺激的样品中的细胞表达或以非常低的水平表达。虽然在此短暂的刺激期内应答的多数细胞推定为记忆细胞，但我们观察到小群体细胞产生较低水平的IL2并且没有其他细胞因子和效应分子，并且可以代表初始细胞(图44B，箭头)。

[0706] 为了完全理解应答中的异质性，我们基于成对相关系数对细胞聚簇。虽然CD4和CD8细胞的两个主要组是明显的，但就表达的活化的基因的组合和水平而言，在每个集合内存在显著差异性(图47和图48)。

[0707] 我们观察到存在几种细胞因子，即IL4、IL5、IL13、IL17F、IL22、LIF、IL3、以及IL21，与未经刺激的相比，在经刺激的样品中它们作为整体被上调几百倍或更多倍，但仅由样品中的少数细胞促成(图44C)。这些细胞因子的亚类由相同的细胞表达(图49A-C)。例如，相同的单细胞贡献于IL17F和IL22(它们是Th17细胞的特征)的计数的大多数。另外7个细胞表达IL4、IL5、IL13(它们是Th2细胞的特征)的各种组合，并且表达它们的各种组合。这种观察结果突出了大规模单细胞分析的重要性，尤其是当对总体表达变化的贡献是源自稀有子群时。

[0708] 我们用来自第二血液供体的T细胞重复了相同的刺激实验，并且分析了分别在经刺激和未经刺激的样品中的669个和595个单细胞的谱图。虽然在该个体中活化的整体水平较低(就表达变化而言的幅度较小)(可能指示对于刺激的个体间可变性)，但我们观察到PCA分析中的相同趋势以及在单独细胞对刺激的应答方面的异质性(图48)。

[0709] 鉴定稀有抗原特异性T细胞

[0710] 我们证明了我们的平台用以使用CD8+T细胞群体中的抗原特异性细胞的模型来鉴定稀有细胞的实用性。我们将相同的两个血液供体的新鲜血液暴露于CMV pp65肽池，所述供体对巨细胞病毒(CMV)呈血清反应阳性。每个供体的单独的未处理的血液等分试样用作阴性对照。我们随后分离CD8+T细胞，并且在我们的平台上分析经刺激和未经刺激的细胞的应答。我们分别从供体2的CMV刺激和未刺激的以及供体1的CMV刺激和未刺激的样品中的2274、2337、581、和253个细胞获得数据。

[0711] 除了供体1的阴性对照产生相对小数目的在聚簇分析中形成明显簇的细胞，所有其余样品显示细胞的两个主要组(图50A、51和52)。在一个组中的细胞表达初始细胞和中心记忆细胞相关标志物SELL、CCR7、和CD27，而在另一个组中的细胞表达效应记忆细胞(CCL4、CX3CR1、CXCR3)和效应细胞相关基因(EOMES、GZMA、GZMB、GZMH、TBX21、ZNF683)。存在着占据两个分支之间的空间并且表达颗粒酶K(GZMK)的不同小细胞亚群，连同表达HLA-DRA的细胞的另一个亚群。不同类型的颗粒酶的差异表达已经在先前报道过(8)。我们的结果重述了在先前使用CD8+T细胞的CyTOF实验中观察到的那些(9)。

[0712] 虽然相当大比例的细胞似乎通过表达CD69和MYC对暴露于抗原有应答(图52)，我们发现仅几种细胞表达IFNG，即活化的抗原特异性细胞的特征细胞因子。大多数表达IFNG的细胞也处于负载基因面板中的最多的总检测转录物分子(指示活性细胞状态)的那些细胞之列，并且属于效应记忆/效应细胞簇(图50B和53)。我们分别鉴定出供体1中的581个细胞中的5个(0.86％)以及供体2中的2274个细胞中的2个(0.09％)基于IFNG表达和总体转录水平可能是CMV特异性的。在那些细胞中，就表达的效应分子(例如颗粒酶)和细胞因子(例如IFNG、IL2、CCL3、CCL4、TNF、CSF2、IL4)的组合和水平而言，存在大量的异质性(图54)。感兴趣的是，在供体2中表达最多转录物的单细胞表达IL6和IL1B两者但不表达IFNG。

[0713] 讨论

[0714] 在此实例中，我们呈现高度可扩大规模的mRNA细胞计数法，该方法使用递归泊松策略以分离单细胞、唯一性地编码细胞内容物、并且编码单独分子，以用于定量分析。我们已经显示，我们可以在包含数千个细胞的样品中同时对属于每个细胞的转录物分子进行鉴定和计数。此外，我们已经证明使用这种技术以在天然存在的异质系统中基于细胞表达谱表征单独细胞，并且在大的背景群体中检测稀有细胞。

[0715] CytoSeq的通量和简单性呈现了优于现有途径的大进步，所述现有途径涉及微量滴定板或微流芯片以用于对单细胞的基因表达的基于测序的测量。因为实验程序简单，并且每个细胞的试剂消耗低(在纳升范围内)，它能够使技术人员易于跨多种条件对大量的细胞进行单细胞分析。仅在此研究中，我们跨12个实验对总计约14,600个单细胞进行基因表达谱分析，这如果通过现有途径进行的话是高成本和耗时的。通过CytoSeq测量的细胞数目可以简单通过以下方式进一步扩大规模：增加微孔阵列的大小以及具有条码的珠的文库大小，这是易于通过组合合成实现的。此外，对细胞大小的均匀性没有限制，由此允许在不经

任何预分选的情况下对包含具有多种细胞大小和形状的细胞(例如在此实例中示出的PBMC)的复杂样品进行直接分析。

[0716] CytoSeq数据类似于流式细胞术(FC)的数据,但是具有重要差异。首先,CytoSeq提供了在研究的基因产物的数目和类型方面的更多的多功能性。不同于流式细胞术主要被限制于少数的表面蛋白并且需要最佳结合抗体,CytoSeq允许经由核酸扩增技术测量任何转录的mRNA。最佳引物设计和测定条件使我们能够对于任意选择的具有100个或更多个基因的面板经由多重PCR常规地实现约88%映射率(表21)。另外,在样品中的每个单细胞的完整转录组还可以经由珠结合的cDNA的通用扩增来测量,但技术人员必须注意常用通用扩增技术(7)的相对低的效率以及跨上千个细胞测量整个转录组所需要的高测序深度。

[0717] 其次,相比于依赖抗体结合动力学的流式细胞术,CytoSeq通过分子索引提供了基因表达水平的数字型绝对读数。它具有对单一稀有细胞事件的更高的灵敏度和特异性,因为该检测是通过共表达对稀有细胞具有特异性的许多基因来实现的。因此,与需要一定数量的事件以形成可靠的用于门控的簇的流式细胞术相比,它消耗更少量的样品。

[0718] 我们的数据展示了单细胞对比整体分析的重要性。例如,我们显示了在作为整体的上千个细胞的样品中最高表达的基因是仅由一个或几个细胞贡献的情况。最重要的是,我们的实验展示了在单细胞基因表达研究中检查大量的细胞和大量的基因两者的重要性,这是一个在现有途径中极其受限制的能力。这种用于跨生物样品中的上千个单细胞对表达进行常规测量的工具的可获得性可以有助于加速对复杂生物系统的理解并且促进在临床诊断学中的新颖应用,例如循环肿瘤细胞分析和免疫应答监测。我们设想,我们的大规模平行单细胞条码化方案还可以用于测量基因组,连同同时测量基因组和转录组,以用于研究在例如癌症生物学和神经科学领域中的单细胞基因组不稳定性。

[0719]

表 21

实验	读数的总数目	具有与面板中的基因恰好1个匹配的读数的数目	与面板中的一个基因比对的读数%	具有与细胞条码准确匹配以及与一个基因对准的读数的数目	在基因和条码比对后的读数%	满足过滤标准的唯一性细胞条码的数目	与那些细胞条码相关联的读数的数目

[0720]

K562 + 拉莫斯	2399025	2154454	90%	1175715	49%	768	859470
原初 B + 拉莫斯	5711013	5203308	91%	3495392	61%	1198	2868577
PBMC	1270214	1105687	87%	803151	63%	632	670576
PBMC 重复	3927672	3468538	88%	2459367	63%	731	1920956
供体 1，抗 CD3/抗 CD28 刺激的	3529898	3249998	92%	2122416	60%	3517	1466000
供体 1，抗 CD3/抗 CD28 阴性对照	1557996	1292211	83%	939094	60%	1478	719351
供体 2，抗 CD3/抗 CD28 刺激的	606865	552877	91%	403943	67%	669	246234
供体 2，抗 CD3/抗 CD28 阴性对照	332951	283723	85%	205762	62%	595	86866
CMV 刺激的供体 1	1064648	958410	90%	697057	65%	581	401629
供体 1 CMV 阴性对照	619957	547259	88%	406801	66%	253	192605
CMV 刺激的供体 2	1902977	1692734	89%	1229667	65%	2274	688296
供体 2 CMV 阴性对照	1671419	1346637	81%	977344	58%	2337	715453

[0721] 珠文库的合成

[0722] 珠是由赛卢拉研究公司(Cellular Research,Inc.)使用分离-聚池组合途径制造的。简言之,将用羧基官能化的二十微米磁珠分配到包含寡聚物的96个管中,这些寡聚物具有5'胺、随后为通用序列、对于不同管而言不同的细胞标记的第一部分、以及接头序列。通过碳二亚胺化学将这些寡聚物共价偶联至这些珠上。将珠聚池并分离到包含寡聚物的第二组96个管中,这些寡聚物具有在5'端上的第二接头序列、随后为对于不同管而言不同的细胞标记的第二部分、以及与第一接头互补的序列。在经由第一接头与溶液中的寡聚物杂交后,珠上的寡聚物由DNA聚合酶延伸。将珠聚池并分离到包含寡聚物的第三组96个管中,这些寡聚物具有在5'端上的寡(dA)、随后为充当分子标记的随机物序列、细胞标记的第三部分、以及与第二接头互补的序列。在经由第二接头与溶液中的寡聚物杂交后,珠上的寡聚物由DNA聚合酶延伸。最终珠文库的大小为96x96x 96(884,736)个细胞标记。

[0723] 制作微孔阵列

[0724] 使用标准光刻法制作微孔阵列。将柱形阵列图案化于硅晶片上的光刻胶上。将PDMS倾倒到该晶片上以产生微孔的阵列。该晶片的复制品是使用PDMS微孔阵列作为模板用NOA63光学胶粘剂制成的。在每个实验之前,将琼脂糖(5%,IX-A型,西格玛公司)微孔阵列从NOA63复制品浇注。

[0725] 样品制备

[0726] 将K562和拉莫斯细胞在具有10%FBS和1x抗生素-抗真菌剂的RPM1-1640中培养。来自健康供体的原初B细胞是购自乐天生物科学(Sanguine Biosciences)。来自健康供体的PBMC是分离自肝素钠管中的新鲜全血,其使用淋巴细胞分离剂溶液(干细胞公司(StemCell))采集自斯坦福血液中心。

[0727] T细胞刺激

[0728] 两个CMV血清反应阳性血液供体的肝素化全血是获得自斯坦福血液中心。对于CMV刺激,将1ml的全血用稀释于PBS(美天旎生物技术(Miltenyi Biotec))中终浓度为1.81μg/ml的CMV pp65肽库在37℃进行刺激持续6小时。将每个供体的全血的单独等分部分与作为阴性对照的PBS孵育。将CD8+T细胞使用RosetteSep混合液(StemCell)进行分离,并且随后沉积到微孔阵列上。对于抗-CD3/抗-CD28刺激,使用RosetteSep T细胞富集混合液将来自相同两个供体的T细胞从全血中分离,并且重悬浮于具有10%FBS和1x抗生素-抗真菌剂的RPMI-1640中。以约1:1珠对细胞比率,将来自每个供体的细胞的一个等分部分与免疫磁珠人类T-活化剂(Dynabeads Human T-Activator)CD3/CD28(生命技术公司(Life Technologies))一起在37℃孵育6小时。将来自每个供体的细胞的分开的等分试样放置于培养箱中,没有刺激并且用作阴性对照。

[0729] 单细胞捕获

[0730] 将单细胞悬浮液用移液管移取到微孔阵列上,密度为约1个细胞/10个微孔。在洗涤以去除未捕获的细胞之后,将磁珠以密度约5个珠/孔加载以饱和该微孔阵列。在洗涤以去除过量珠之后,将冷裂解缓冲液(0.1M Tris-HCl pH 7.5、0.5M LiCl、1%LiSDS、10mM EDTA、5mM DTT)用移液管移取在该微孔阵列的表面上。在磁体片上孵育10分钟之后,将这些珠从该微孔阵列中收回。将珠收集在微量离心管中,并且用A洗涤缓冲液(0.1M Tris-HCl、0.5M LiCl、1mM EDTA)洗涤两次,并且用B洗涤缓冲液(20mM Tris-HCl pH 7.5、50mM KCl、3mM MgCl2)洗涤一次。从此往后,在单一管中进行所有反应。

[0731] cDNA合成

[0732] 在杂交箱中，将洗涤的珠重悬浮于在置于旋转器上的微量离心管中的40μL Rt混合物(1x第一链缓冲液、1μL SuperRase抑制剂、1μL SuperScript II或SuperScript III、3mM另外的MgCl2、1mM dNTP、0.2ug/μL BSA)中，在50℃持续50分钟(当使用SuperScript III用于使用K562和拉莫斯细胞的早期实验时)或者在42℃持续90分钟(当使用Superscript II用于其余实验时)。将珠用在20μL的1x ExoI缓冲液中的1μL ExoI(NEB)在37℃处理30分钟，并且在80℃处理15分钟。

[0733] 多重PCR和测序

[0734] 每个基因面板包含通过Primer3设计的两组基因特异性引物。编写定制MATLAB脚本以选择PCR引物，使得在组内的引物之间存在最少的3'端互补性。每个面板中的引物列于表21中。扩增方案示于图55中。用KAPA快速多重试剂盒(KAPA Fast Multiplex Kit)，对珠进行PCR，其中使用50nM的在第一引物组中的每个基因特异性引物和400nM通用引物，体积为100μL或200μL，采用以下循环方案：95℃ 3分钟；15个循环：95℃ 15秒，60℃ 60秒，72℃ 90秒；72℃ 5分钟。将磁珠回收，并且将PCR产物用0.7x Ampure XP进行纯化。将一半的纯化产物用于下一轮巢式PCR，该下一轮巢式PCR采用第二引物组，使用相同的KAPA试剂盒和循环方案。在用0.7x Ampure XP净化后，将1/10th的产物输入最终PCR反应，借此附加全长亿明达适配子(1x KAPA HiFi Ready Mix、200nM的P5、200nM的P7。95℃ 5分钟；8个循环：98℃ 15秒，60℃ 30秒，72℃ 30秒；72℃ 5min)。

[0735] 数据分析

[0736] 在亿明达MiSeq仪器上，使用150x 2bp化学物进行文库测序，深度中值为160万个读数/样品。测序揭示了每个读数的细胞标记、分子标记、和基因(图55)。每个读数的基因的比对是用比对软件'bowtie'(参考)完成的。使用定制MATLAB脚本对每个读数的细胞标记和分子标记进行分析。将读数首先通过细胞标记、然后通过基因标记和分子标记来分组。为了计算唯一性分子数目/基因/细胞，对具有相同细胞标记和基因比对的读数的分子标记进行聚簇。大于1个碱基的编辑距离被认为是唯一性簇，并且由此为唯一性转录物分子。对于每个样品构建包含每个细胞的数字基因表达信息的表-在该表中的每行代表唯一性细胞标记，每列代表基因，并且该表中的每个条目代表基因内唯一性分子计数/细胞标记。对该表进行过滤使得进行仅一次(即冗余度＝1)测序的唯一性分子被去除。随后，将唯一性分子总数小于10的细胞或者面板中的共表达4个或更少基因的细胞去除。然后将过滤的表用于聚簇分析。采用MATLAB中的内置函数对于经对数转换的转录物计数(其中添加1的假计数)进行主成分分析和分层聚簇。

[0737] 在实例15中引用参考文献，将所有这些文献通过引用以其全文结合：

[0738] 1.A.K.沙勒克(Shalek)等人，单一细胞转录组学揭示免疫细胞中表达和剪接中的二型性(Single-cell transcriptomics reveals bimodality in expression and splicing in immune cells).自然(Nature)498,236(2013年1月13日)。

[0739] 2.S.C.本多尔(Bendall)等人，在人类造血连续体中差异性免疫和药物反应的单一细胞大量细胞计数法(Single-cell mass cytometry of differential immune and drug responses across a human hematopoietic continuum).科学(Science)332,687(2011年5月6日)。

[0740] 3.A.R.吴(Wu)等人,单一细胞RNA-测序方法的定量评估(Quantitative assessment of single-cell RNA-sequencing methods).自然方法(Nature methods)11,41(2014年1月)。

[0741] 4.B.特罗伊特莱因(Treutlein)等人,使用单一细胞RNA-seq重构远端肺上皮的谱系等级(Reconstructing lineage hierarchies of the distal lung epithelium using single-cell RNA-seq).自然(Nature)509,371(2014年5月15日)。

[0742] 5.S.伊斯兰(Islam)等人,通过高度多重RNA-seq对单一细胞转录景观的表征(Characterization of the single-cell transcriptional landscape by highly multiplex RNA-seq).基因组研究(Genome research)21,1160(2011年7月)。

[0743] 6.G.K.付(Fu),J.胡(Hu),P.H.王(Wang),S.P.福多尔(Fodor),通过随机附接不同标记对单独的DNA分子计数(Counting individual DNA molecules by the stochastic attachment of diverse labels).美国国家科学院院刊(Proceedings of the National Academy of Sciences of the United States of America)108,9026(2011年5月31日)。

[0744] 7.G.K.付(Fu),J.威廉米(Wilhelmy),D.斯特恩(Stern),H.C.范(Fan),S.P.福多尔(Fodor),细胞mRNA的数字编码使得能够通过单一分子计数进行精确且绝对的基因表达测量(Digital encoding of cellular mRNAs enabling precise and absolute gene expression measurement by single-molecule counting).分析化学(Analytical chemistry)86,2867(2014年3月18日)。

[0745] 8.K.布拉特克(Bratke),M.屈佩尔(Kuepper),B.巴德(Bade),J.C.菲尔绍(Virchow),Jr.,W.吕特曼(Luttmann),在自然杀伤细胞中以及在外周血中CD8+T细胞分化期间人类颗粒酶A、B、和K的差异性表达(Differential expression of human granzymes A,B,and K in natural killer cells and during CD8+T cell differentiation in peripheral blood).欧洲免疫学杂志(European journal of immunology)35,2608(2005年9月)。

[0746] 9.E.W.纽厄尔(Newell),N.西加尔(Sigal),S.C.本多尔(Bendall),G.P.诺兰(Nolan),M.M.戴维斯(Davis),通过飞行时间的细胞计数法示出CD8+T细胞表型的连续体内的组合细胞因子表达和病毒特异性细胞位置(Cytometry by time-of-flight shows combinatorial cytokine expression and virus-specific cell niches within a continuum of CD8+T cell phenotypes).免疫(Immunity)36,142(2012年1月27日)。

[0747] 实例16:单细胞定量方案的开发

[0748] 图56描绘了用于定量样品中的RNA分子的一般工作流程。在此实例中,在样品中的RNA分子的总数目等于单细胞中的RNA分子的总数目。如在图56的步骤1中所示,通过一组分子鉴定物标记(115)随机杂交到RNA分子的聚A尾区域,RNA分子(110)被逆转录而产生cDNA分子(105)。分子鉴定物标记(115)包括寡dT区(120)、标记区(125)、以及通用PCR区(130)。该组分子鉴定物标记包含960个不同类型的标记区。

[0749] 部分I.RNA分子的逆转录和标记

[0750] 通过混合以下项来制备RNA样品:

[0751]

基因	RNA分子数目

Lys(掺入型对照)	456
Phe(掺入型对照)	912
Thr(掺入型对照)	1824
Dap(掺入型对照)	6840
Kan(掺入型对照)	7352
淋巴细胞系RNA	10pg(1细胞当量)
MS2载体(没有聚A)	6x 10^{11}

[0752] 通过在埃彭道夫管中制备如下的标记混合物来标记RNA分子：

[0753]

	量 (μL)
RNA 样品	2
ms2 RNA 1μg/μL	1
10 mM dNTP	1

[0754]

960 dT 寡聚物池(4 号组)10 μM	0.4
水	9.1

[0755] 注释：dT寡聚物池(4号组)是指分子鉴定物标记的集合。

[0756] 通过在65℃孵育5分钟，分子鉴定物标记与RNA分子杂交。将标记混合物储存在冰上，持续至少1分钟。

[0757] 通过添加如下所述的逆转录混合物对标记的RNA分子进行逆转录：

[0758]

	量(μL)
5X第一链缓冲液	4
0.1M DTT	1
superase-in 20u/μL	0.5
superscript III RT	1

[0759] 一旦将逆转录混合物添加到包含标记混合物反应的埃彭道夫管中，则通过将样品孵育于37℃持续5分钟、随后在50℃孵育持续30分钟、并且最后在75℃孵育持续15分钟进行逆转录反应。经标记的RNA分子的逆转录产生了经标记的cDNA分子(170)。

[0760] 一旦RNA分子被逆转录和标记，则通过Ampure珠纯化将过量的寡聚物从样品去除(图1的步骤2)。通过以下方式进行Ampure珠纯化：将20μL的ampure珠添加至含有经逆转录的且标记的RNA分子的埃彭道夫管中，并且将管在室温孵育5分钟，将珠用70％乙醇洗涤以去除过量的寡聚物。一旦通过乙醇洗涤去除过量的寡聚物，则将20μL的10mM Tris添加到包含珠结合的经标记cDNA分子的管中。

[0761] 如在图56的步骤3中所示，通过多重PCR扩增经标记的cDNA分子(170)。经标记的cDNA分子的定制扩增是通过使用定制正向引物(F1，图56中的135)和通用PCR引物(140)进行的。表23列出96个不同的定制正向引物，这些引物用于扩增96个不同的基因以产生单一反应体积中的经标记的扩增子(180)。

[0762] 为了优化多重PCR反应，制备了3个多重PCR反应混合物。多重PCR反应1如下制备：

[0763]

	反应 1
	量（μL）
10 X 钛	5
10 mM dNTP	1.5
水	35.5
1 μM 各 F1 引物池	5
PCR004 10μM	1
纯化的 cDNA	1
钛聚合酶	1

[0764]

[0765] 多重PCR反应的反应条件是1个循环的94℃持续2分钟；随后25个循环：94℃持续30秒，57℃持续60秒，以及68℃持续1分钟；然后1个循环的68℃持续7分钟，以及1个在4℃的维持循环。

[0766] 多重PCR反应2和3如下制备：

[0767]

	反应 2	反应 3
	量（μL）	量（μL）
2 X 凯杰（Qiagen）多重混合物	25	25
1 μM 各 F1 引物池	5	5
PCR004 10 μM	1	1
Q 溶液		5
水	18	13
纯化的 cDNA	1	1

[0768] 对于反应2和3的多重PCR反应条件是1个循环的95℃持续15分钟；随后25个循环：94℃持续30秒，57℃持续90秒，以及72℃持续1分钟；然后1个循环的68℃持续7分钟，以及1个在4℃的维持循环。

[0769] F1引物池包含以下引物：

[0770]

F1 PCR 引物	序列
100611KanF2	CTGCCTCGGTGAGTTTTCTC
Lys_L_269	CTTCCCGTTACGGTTTTGAC

[0771]

phe_L_177	AAAACCGGATTAGGCCATTA
thr_L_332	TCTCGTCATGACCGAAAAAG
dap_L_276	CAACGCCTACAAAAGCCAGT

[0772] 通过巢式PCR选择性地扩增Kan、Phe和Dap对照基因。巢式PCR扩增反应如下制备：

[0773]

多重 PCR Rxn # →	1	2	3	1	2	3	1	2	3
PCR Rxn # →	1	2	3	4	5	6	7	8	9
	μL	μL	μL	μL	μL	μL	μL	μL	μL
10x Taq	2.5	2.5	2.5	2.5	2.5	2.5	2.5	2.5	2.5
10 mM dNTP	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5
水	22.25	22.25	22.25	22.25	22.25	22.25	22.25	22.25	22.25
Cy3 PCR004 10μM	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5
KanF3_5P 5 μM	1	1	1						
Phe_L_215 5 μM				1	1	1			
Dap_L_290 5 μM							1	1	1
多重 PCR Rxn	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5
USB taq	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25

[0774] 注释：用于PCR反应1、4、和7的多重PCR反应是多重PCR反应#1。用于PCR反应2、5、和8的多重PCR反应是多重PCR反应#2。用于PCR反应3、6、和9的多重PCR反应是多重PCR反应#3。

[0775] 用于巢式PCR的引物披露如下：

[0776]

巢式PCR引物	序列
KanF4_5P	/5Phos/GTGGCAAAGCAAAAGTTCAA
Phe_L_215	TGAGAAAGCGTTTGATGATGTA
Dap_L_290	GCCAGTTTATCCCGTCAAAG

[0777] 对于反应1-9的PCR扩增反应条件是1个循环的94℃持续2分钟；30个循环：94℃持续20秒，55℃持续20秒，以及72℃持续20秒；然后1个循环的72℃持续4分钟，以及1个在4℃的维持循环。

[0778] 将PCR扩增反应1-9的4μL PCR产物在琼脂糖凝胶上运行。如在图58A中所示，反应1-3显示Kan对照基因的存在，反应4-6显示该Phe对照基因的存在，并且反应7-9显示Dap对照基因的存在。

[0779] 将来自PCR反应1-9的PCR产物进行制备用于杂交到应用微阵列公司(Applied Microarray Inc.,AMI)阵列上。杂交混合物如下制备：

[0780]

	μL
PCR产物	20
洗涤A(6X SSPE+0.01％曲通X-100)	55
Cy3寡聚物(760pM)	1

[0781] 将分别与包含来自PCR反应1-9的PCR产物的混合物相对应的杂交混合物1-9在95℃变性5分钟，并且然后放置于4℃。将杂交混合物转移至AMI阵列片上，并且在37℃孵育过夜。

[0782] 在过夜杂交之后，对该AMI阵列片进行洗涤并且然后扫描。理论和实际的测量和百分比准确性描绘如下：

[0783]

杂交混合物#→	1	2	3	4	5	6	7	8	9
多重PCR条件（Rxn #）	钛(1)	凯杰(Qiagen)(2)	凯杰+Q(3)	钛(1)	凯杰(2)	凯杰+Q(3)	钛(1)	凯杰(2)	凯杰+Q(3)
理论测量	3676（生物分析仪）			912			6840		
实际测量	1826	1740	2116	299	235	251	1165	1077	172
%检测	49.7	47.3	57.6	32.8	25.8	27.5	17	15.7	2.5

[0784] 注释：理论测量是基于Kan、Phe和Dap对照基因的100％的检测。

[0785] 通过Ampure纯化来对来自反应2的PCR产物进行纯化。Ampure纯化是如下进行的：

[0786]

	μL
来自X01样品2的F1PCR产物	30
Ampure珠	30

[0787] 将Ampure纯化反应在室温下孵育5分钟，并且然后在70％乙醇中洗涤。将纯化的PCR产物从在30μL水中的珠进行洗脱。PCR产物的浓度是6ng/μL，如通过Nanodrop光谱仪确定的。

[0788] 部分II:文库制备方案

[0789] 使用从X01样品2纯化的PCR产物(参见实例1)来制备DNA文库。混合以下引物来创造F2引物池:

[0790]

F2PCR引物	序列
Lys_L_269	CTTCCCGTTACGGTTTTGAC
phe_L_177	AAAACCGGATTAGGCCATTA
thr_L_332	TCTCGTCATGACCGAAAAAG
dap_L_276	CAACGCCTACAAAAGCCAGT

[0791] 混合以下项来制备F2引物混合物

[0792]

F2 引物混合物	μL
水	750
F2 引物池 1 uM 每个/100 uM 总的	100

[0793] 将F2引物混合物孵育于95℃持续3分钟,并且然后储存在冰上。将以下连接混合物添加到F2引物混合物以产生F2引物连接混合物:

[0794]

连接混合物	μL
10 X DNA 连接酶缓冲液 NEB	100
T4 PNK USB	50

[0795] 将F2引物连接混合物在37℃孵育1小时,随后在65℃孵育20分钟。将F2PCR引物经历乙醇沉淀,并且通过Nanodrop分光光度计确定引物池的浓度。重悬浮F2引物池以产生以下最终浓度:1uM每个/100uM总的。

[0796] 如在图56的步骤4中所示,通过多重PCR扩增经标记的扩增子(180)。使用96个不同的定制正向引物(F2,图56中的145)和通用PCR引物(140)在单一反应体积中扩增经标记的扩增子(来自实例1的X01样品2)。表24列出96个不同的定制正向引物。

[0797] 多重PCR反应如下制备:

[0798]

多重 PCR 混合物	μL
2 X 凯杰（Qiagen）多重混合物	25
1 μM 各 F2 引物池激酶	5
PCR004 5'P 10 μM	1
水	18
纯化的第一 PCR X01 样品 2	1

[0799] 多重PCR条件是1个循环的95℃持续15分钟;随后18个循环:94℃持续30秒,57℃持续90秒,以及72℃持续1分钟;然后1个循环的68℃持续7分钟,以及1个在4℃的维持循环。将多重扩增子通过Ampure纯化进行纯化,并且用50μL的水进行洗脱。通过Nanodrop分光光度计,扩增子的浓度被确定为30ng/μL。将5μL的扩增子在琼脂糖凝胶上运行(图58B)。

[0800] 如图56的步骤5中所示,适配子(150,155)连接到经标记的扩增子(180)上以产生经适配子标记的扩增子(190)。经适配子标记的扩增子如下产生:

[0801]

适配子混合物	μL
10X T4连接酶USB	10
水	60
纯化的巢式PCR产物	10
退火的聚池的96个ABC适配子50μM	10
T4DNA连接酶(3μl neb hc,7μl usb)	10

[0802] 将适配子混合物在16℃孵育4小时。该经适配子标记的扩增子通过Ampure纯化进行纯化,并且在20μL的10mM Tris中洗脱。

[0803] 将纯化的经适配子标记的扩增子进行间隙修复,并且进行PCR扩增,如下:

[0804]

填充物和 PCR 混合物	μL
10x thermoPol 缓冲液	5
10 mM dNTP	1.5
水	32
CR P1 10 μM	3
CR IDX D1 10 μM	3
纯化的经适配子标记的扩增子	5
Vent exo- 2 u/μL	0.5

[0805] PCR条件是1个循环的72℃持续2分钟，随后94℃持续1分钟；12循环：94℃持续15秒，60℃持续15秒，以及72℃持续30秒；1个循环的72℃持续4分钟，以及1个在4℃的维持循环。将PCR产物通过Ampure纯化进行纯化，并且用30μL的TE进行洗脱。纯化的PCR产物的浓度是22ng/μL(83nM)，如通过Nanodrop光谱法确定的。将5μL的PCR纯化产物在1％琼脂糖凝胶上运行(图58B)

[0806] 部分III.对经适配子标记的扩增子文库进行测序

[0807] 使用MiSeq测序仪对经适配子标记的扩增子文库进行测序。

[0808] 下文示出序列图谱概述：

[0809]

	需要完全匹配	允许 1 bp 错配
总读数对	7,724,955	
具有通用引物和聚 A 匹配的 RNA 的#	2,499,444 (32%)	4,716,378 (61%)
映射到靶标 (96 个基因) 的 RNA 的#	2,373,700	4,489,485

[0810] 如以上序列图谱概述中所示，许多读数由于严格的聚A匹配标准而丢失。图59示出跨所有检测的基因的读数和计数。

[0811] 还使用测序读数定量特异性基因。图61-62描绘了对于不同基因而言，观察的读数/检测的标记(RPLD)的图。常规rpkm值也示于图61-62描绘的图中。图59总结了针对不同基因而言的RPLD和RPKM的比较。

[0812] 图63描绘了针对不同基因而言的总读数(标记)对rpld的图。

[0813] 在图4、7和8中表示的数据还以数字形式示于表25中。

[0814] 图64描绘了针对未检测的基因而言的RPKM的图。

[0815] 在经适配子标记的扩增子文库中的掺入型对照的量是通过MiSeq测序来确定。在下表中示出了来自对掺入型对照进行MiSeq测序的结果。

[0816]

掺入型对照 (Spike-in Control)	输入 N(mfg)	读数	标记 (K)
Dap	6840	1,920,503	893
Phe	912	470,738	859
Thr	1824	410,664	847
Lys	456	282,174	847
Kan	7352	24	23

[0817] 在上表中，输入N是指掺入型对照的初始数目；读数是指读数对的总数目；并且标

记(K)是指通过测序检测的不同标记的数目。图60A-D描绘分别针对Lys、Phe、Thr、和Dap掺入型对照而言，观察的读数/检测的标记(RPLD)的图。图60E描绘了读数比对输入的绘图。

[0818]

表 23

名称	序列
NM_144646.3F1	TTGACTTTGCCTTGGAGAGC
NR_015342.1F1	TTTTTCTTACAGTGTCTTGGCATA
NM_000193.2F1	CGTGACCCTAAGCGAGGAG
NM_001777.3F1	TTTGCAGTGATTTGAAGACCA
NM_000600.3F1	GGCATTCCTTCTTCTGGTCA
NM_021127.2F1	CTGGGCTATATACAGTCCTCAAA
NM_004318.3F1	GGGGTGATTATGACCAGTTGA
NM_002467.4F1	TGCATGATCAAATGCAACCT
NM_001773.2F1	TCTTCCGAAAAATCCTCTTCC

[0819]

NM_001770.5F1	CTGGGGTCCCAGTCCTATG
NM_001718.4F1	TGTACTGGGAAGGCAATTTCA
NR_023920.1F1	GAGCCGCTGGGGTTACTC
NM_000267.3F1	CAGTTAGTTGCTGCACATGGA
NM_000633.2F1	TTGCATTTCTTTTGGGGAAG
NM_000314.4F1	GTCATGCATGCAGATGGAAG
NM_021151.3F1	GCTGCAGTGAGCTGTGATGT
NM_002415.1F1	GTTCCTCTCCGAGCTCACC
NM_004985.3F1	TCCGAAAGTTTCCAATTCCA
NM_005375.2F1	TTGTTTGGGAGACTCTGCATT
NM_000555.3F1	GACCCCACTTGGACTGGTAG
NM_001668.3F1	GTGATCTTGATTGCGGCTTT
NM_025237.2F1	GGGGGAAAAACTACAAGTGC
NM_021117.3F1	TGATTCCTTTTCCTGCCTGT
NM_016316.2F1	AAAAACCTCCAGGCCAGACT
NM_021975.3F1	AATCAAAATAACGCCCCAGA
NM_004333.4F1	TTGCTAAAAATTGGCAGAGC
NM_001621.4F1	TTGTTAAGTGCCAAACAAAGGA
NM_005239.5F1	AAGCTGGGAAGAGCAAAGC
NM_000485.2F1	AGGACAGAGGGTGGTCGTC
NM_004048.2F1	TGAGTGCTGTCTCCATGTTTG
NM_001657.2F1	CCTCACAGCTGTTGCTGTTATT
NM_012238.4F1	AAAACACCCAGCTAGGACCA
NM_002055.4F1	AACTGAGGCACGAGCAAAGT
NM_002392.4F1	GCTTTATGGGTGGATGCTGA
NM_001625.3F1	ATAATATCGCCAGCCTCAGC
NM_002110.3F1	TCCAGAGTGTGCTGGATGAC
NM_002943.3F1	TGCAAGCCATTTATGGGAAT
NM_000059.3F1	TGGAATGAGGTCTCTTAGTACAGTT
NM_018136.4F1	TCCCAGAAACACCTGTAAGGA
NM_003467.2F1	TGTCTAGGCAGGACCTGTGG

[0820]

NM_004958.3F1	AGTGATGCTGCGACTCACAC
NM_006139.3F1	GGCTCAGAAAGTCTCTCTTTCC
NM_002693.2F1	CTCCCAAACTCAGGCTTTCA
NM_001080432.2F1	AAAGCGCTGGGATTACAGG
NM_005954.2F1	CGTCCAGTTGCTTGGAGAAG
NM_024865.2F1	AATAACCTTGGCTGCCGTCT
NM_001905.2F1	GGGAATTCTCAGTGCCAACT
NM_002046.4F1	GCATCCTGGGCTACACTGAG
NM_002253.2F1	TGCTGGGAACAATGACTATAAGA
NM_002356.5F1	GCCTAAAACACTTTGGGTGGT
NM_000189.4F1	GGGTGCCCACAAAATAGAGA
NM_000546.5F1	GAGACTGGGTCTCGCTTTGT
NM_152860.1F1	TGGGGAAGGCTTTCTCTAGG
NM_016231.4F1	TTCAACTTGAGTGATCTGAGCTG
NM_000518.4F1	TATGGGCAACCCTAAGGTGA
NM_000905.3F1	CGCTGCGACACTACATCAAC
NM_005038.2F1	TGGAGTCTTGCTCTGTCACC
NM_000041.2F1	ACGAGGTGAAGGAGCAGGT
NM_005957.4F1	CGATGCCTTTGGGTAGAGAG
NR_002785.2F1	ACTGATCGTCCAAGGACTGG
NM_000321.2F1	AAAAAGAAATCTGGTCTTGTTAGAAAA
NM_152756.3F1	TTGAAAAGTGGTAAGGAATTGTGA
NM_000610.3F1	CACCAAGAATTGATTTTGTAGCC
NR_033314.1F1	AAAAATGGGGGAAAATGGTG
NM_017460.5F1	CATGGTTGAAACCCCATCTC
NR_002196.1F1	TTCAAAGCCTCCACGACTCT
NM_000591.3F1	GCTGGAACAGGTGCCTAAAG
NM_000106.5F1	CCCTAAGGGAACGACACTCA
NM_138712.3F1	ACCTGCTACAAGCCCTGGA
NM_004304.4F1	GGATCCCTAAGACCGTGGAG
NM_000754.3F1	CCACCTCAGAGGCTCCAA

[0821]

NM_000492.3F1	TGCTGTATTTTAAAAGAATGATTATGA
NM_000444.4F1	GTAGCTGGGACGCTGGTTTA
NM_002463.1F1	ATTCCCTTCCCCCTACAAGA
NM_000552.3F1	CCTGAGTGCAACGACATCAC
NM_005430.3F1	GGGGGAACCAGCAGAAAT
NM_003150.3F1	GACCTAGGGCGAGGGTTC
NM_000388.3F1	AATTCCTGAAGCCAGATCCA
NM_007294.3F1	AAAATGTTTATTGTTGTAGCTCTGG
NM_005933.3F1	TTTCAAGAGCTCAACAGATGACA
NM_002343.3F1	GACTGCCCGGACAAGTTTT
NM_000376.2F1	GAGAAGGTGCCCCAAAATG
NM_002462.3F1	AGCCACTGGACTGACGACTT
NM_021005.3F1	GGAGGACTAGTGAGGGAGGTG
NM_012343.3F1	GGCAAGTGATGTGGCAATTA
NM_001741.2F1	GTTGGAGCACCTGGAAAGAA
NM_014417.4F1	ATGCCTGCCTCACCTTCAT
NM_014009.3F1	ACAGGGGCACTGTCAACAC
NM_006908.4F1	AAAAATCATGTGTTGCAGCTTT
NM_005228.3F1	TGCTTTCACAACATTTGCAG
NM_013994.2F1	AATGTTTCCTTGTGCCTGCT
NM_000639.1F1	ATATCCTGAGCCATCGGTGA
NM_002701.4F1	TTTTGGTACCCCAGGCTATG
NM_000268.3F1	ACCCCGTGGCATTACATAAC
NM_003140.1F1	CTTCCAGGAGGCACAGAAAT
NM_000551.3F1	CTAACCTGGGCGACAGAGTG

[0822]

表 24	
名称	序列
NM_144646.3F2	ATATTTGGACATAACAGACTTGGAA
NR_015342.1F2	TGCTGACTTTTAAAATAAGTGATTCG

[0823]

NM_000193.2F2	GCGGCAGAGTAGCCCTAAC
NM_001777.3F2	TGGGCTATTTCTATTGCTGCT
NM_000600.3F2	AATGGAAAGTGGCTATGCAG
NM_021127.2F2	GGTTGTAGTCACTTTAGATGGAAAA
NM_004318.3F2	TTTGTTTGACTTTGAGCACCA
NM_002467.4F2	AATGTTTCTCTGTAAATATTGCCATT
NM_001773.2F2	CACCCCCATATGGTCATAGC
NM_001770.5F2	AGCACCAGGTGATCCTCAG
NM_001718.4F2	TGTTTTGCTGTAACATTGAAGGA
NR_023920.1F2	TAATGCCACAGTGGGGATG
NM_000267.3F2	GGGCCTAAACTTTGGCAGTT
NM_000633.2F2	TTTTACCTTCCATGGCTCTTTT
NM_000314.4F2	GCCTTACTCTGATTCAGCCTCTT
NM_021151.3F2	CGTAACAAAATTCATTGTGGTGT
NM_002415.1F2	AGAACCGCTCCTACAGCAAG
NM_004985.3F2	GTGCTTTCTTTTGTGGGACA
NM_005375.2F2	GGGAGTTCTGCATTTGATCC
NM_000555.3F2	TGGGTCAGAGGACTTCAAGG
NM_001668.3F2	AGGGTTCTGATCACATTGCAC
NM_025237.2F2	CTGCAGGACTGGTCGTTTTT
NM_021117.3F2	AGGGCAGGGTAGAGAGGGTA
NM_016316.2F2	TTCTTCCATGCGGAGAAATC
NM_021975.3F2	CATGGCTGAAGGAAACCAGT
NM_004333.4F2	TTGCCAGCTATCACATGTCC
NM_001621.4F2	TCTTTTCCTGTACCAGGTTTTTC
NM_005239.5F2	TGACTGGGAACATCTTGCTG
NM_000485.2F2	TGGCACCTGTACCCTTCTTC
NM_004048.2F2	TTCAATCTCTTGCACTCAAAGC
NM_001657.2F2	TGGAGTCACTGCCAAGTCAT
NM_012238.4F2	TTTGCATGATGTTTGTGTGC
NM_002055.4F2	GCACCCACTCTGCTTTGACT

[0824]

NM_002392.4F2	ACCATGTAGCCAGCTTTCAA
NM_001625.3F2	GCAACTGGGCATGAGTACCT
NM_002110.3F2	CCACACCCCCTTCCTACTC
NM_002943.3F2	AGTCTGCTTATTTCCAGCTGTTT
NM_000059.3F2	TCCTGTTCAAAAGTCAGGATGA
NM_018136.4F2	AAATCACAAATCCCCTGCAA
NM_003467.2F2	CTGAACATTCCAGAGCGTGT
NM_004958.3F2	CAGTGGGACCACCCTCACT
NM_006139.3F2	TCTGTAGATGACCTGGCTTGC
NM_002693.2F2	TCAGAACCAAGATGCCAACA
NM_001080432.2F2	CATGACCCAGCCTATGGTTT
NM_005954.2F2	ACCTCCTGCAAGAAGAGCTG
NM_024865.2F2	TTGGGAGGCTTTGCTTATTTT
NM_001905.2F2	CTGGGAAACACTCCTTGCAT
NM_002046.4F2	CAACGAATTTGGCTACAGCA
NM_002253.2F2	CAAAGGTCATAATGCTTTCAGC
NM_002356.5F2	TTTGACGTATCTTTTCATCCAA
NM_000189.4F2	TGTTGTTGGTTTCCAAAAAGG
NM_000546.5F2	GCCAACTTTTGCATGTTTTG
NM_152860.1F2	CCCAAGCTGATCTGGTGGT
NM_016231.4F2	TGCTGTGAAAGAAACAAACATTG
NM_000518.4F2	GCACGTGGATCCTGAGAACT
NM_000905.3F2	CCAGCCCAGAGACACTGATT
NM_005038.2F2	CACGCCCAGCTAATTTTTGT
NM_000041.2F2	CCTGGTGGAAGACATGCAG
NM_005957.4F2	TCACACCTGTAATCCCAGCA
NR_002785.2F2	CAGAGCTCCGCCTCATTAGT
NM_000321.2F2	TCCATTTCATCATTGTTTCTGC
NM_152756.3F2	TGGTGTTTGTAGGTCACTGAACA
NM_000610.3F2	AACATGGTCCATTCACCTTTATG
NR_033314.1F2	AGAGCGAGACTCCGTCTCAA

[0825]

NM_017460.5F2	AGTGAGCTGAGATTGCACCA
NR_002196.1F2	AGACGGCCTTGAGTCTCAGT
NM_000591.3F2	GGGAATCCCTTCCTGGTC
NM_000106.5F2	CTTCCTGCCTTTCTCAGCAG
NM_138712.3F2	TGCAGGTGATCAAGAAGACG
NM_004304.4F2	GGTTTTGAGCATGGGTTCAT
NM_000754.3F2	CCAGCCCACTCCTATGGAT
NM_000492.3F2	AAACTGGGACAGGGGAGAAC
NM_000444.4F2	TTTGGGTAGGTGACCTGCTT
NM_002463.1F2	TCACTGAACGAATGAGTGCTG
NM_000552.3F2	ACGATGTGCAGGACCAGTG
NM_005430.3F2	AATTTGCACTGAAACGTGGA
NM_003150.3F2	CTGTTGTGGCCCATTAAAGAA
NM_000388.3F2	TTCCCTCCAGCAGTGGTATT
NM_007294.3F2	CACCAGGAAGGAAGCTGTTG
NM_005933.3F2	TTTCCTTGTGTTCTTCCAAGC
NM_002343.3F2	TCGCAGGCATTACTAATCTGAA
NM_000376.2F2	CTCTGGCTGGCTAACTGGAA
NM_002462.3F2	AGAGCCCCACCCTCAGAT
NM_021005.3F2	TGTGCAGAGTTCTCCATCTGA
NM_012343.3F2	TGCCTGTTACAAATATCAAGGAA
NM_001741.2F2	TTTCCCTTCTTGCATCCTTC
NM_014417.4F2	TGTGACCACTGGCATTCATT
NM_014009.3F2	CTCACACACACGGCCTGTTA
NM_006908.4F2	CACTTGACCAATACTGACCCTCT
NM_005228.3F2	GTGTGTGCCCTGTAACCTGA
NM_013994.2F2	CCACTTCCCACTTGCAGTCT
NM_000639.1F2	TGTGTGTGTGTGTGTGTGTGT
NM_002701.4F2	TCTCCCATGCATTCAAACTG
NM_000268.3F2	TCTAAGTGTTCCTCACTGACAGG
NM_003140.1F2	TACTCTGCAGCGAAGTGCAA

[0826]

NM_000551.3F2	CCAAGATCACACCATTGCAC
NM_144646.3F2	ATATTTGGACATAACAGACTTGGAA

[0827]

表 25

基因	读数数目	标记数目	每 kb 的读数/百万 (RPKM)
APOE	1585	408	0.2
APRT	11280	56	103.4
AREG	147	102	0.0
ASPM	4683	53	4.4
B2M	209362	698	3891.7
BBC3	8	1	0.0
BCL2	3627	27	33.2
BDNF	12778	116	0.3
CD19	38	5	43.1
CD44	6789	47	8.1
COMT	2828	16	10.7
CTPS1	3998	15	25.4
CXCR4	10547	54	19.2
CYP3A4	80982	267	0.1
DCX	28	24	0.0
ETS2	6	5	0.0
FASLG	3182	565	0.2
FTO	8877	58	11.3
GAPDH	227129	661	3870.8
HCK	294	2	2.4
HK2	593	2	12.7
IGJ	119449	454	438.9
KDR	2	2	0.0
KRAS	64	31	6.8
LTF	126	90	0.0

[0828]

MARCKS	1563	12	36.9
MIF	17775	89	760.4
MLL	72	9	2.6
MTHFR	4854	282	3.9
MX1	100701	285	119.0
MX2	2145	13	45.2
MYB	18361	100	2.8
MYC	6859	27	130.5
NF1	4	1	3.7
NNT	15673	78	14.1
PMAIP1	50604	244	126.9
POLG	5163	46	7.4
POU5F1	1924	12	1.0
PPID	27354	303	39.0
PTEN	20884	109	51.6
RAC1	12454	67	44.8
RB1	1420	14	46.1
RELA	3893	26	17.9
RICTOR	898	5	5.2
RORA	954	7	0.1
SOST	1	1	0.0
SP7	1	1	0.0
STAT3	706	28	16.9
TP53	900	34	14.6
VHL	11576	106	0.0

[0829] 尽管在此已经显示并说明了本发明的优选实施例，但是本领域的普通技术人员将清楚的是仅作为举例而提供了此类实施例。本领域的普通技术人员现在将会想到众多变体、变化、以及替代，而不背离本发明。应该理解的是，在此说明的本发明的实施例的不同替代方案可以用于实施本发明。

[0830] 仅提供在此讨论的这些公开在本申请的提交日期之前的披露内容。在此不应被视为承认本发明因现有发明而无权先于这些公开。另外，所提供的公开日期可能与实际公开

日期不同,实际公开日期可能需要独立地确定。

[0831] 实施方案

[0832] 在此披露的是用于分析两个或更多个样品中的分子的方法。该方法可以包括:a)通过以下产生多个样品标签化核酸:i)使包括多个核酸的第一样品与多个第一样品标签接触,以产生多个第一样品标签化核酸;并且ii)使包括多个核酸的第二样品与多个第二样品标签接触,以产生多个第二样品标签化核酸,其中该多个第二样品标签不同于这些第一样品标签;b)使该多个样品标签化核酸与多个分子鉴定物标记接触,以产生多个经标记的核酸;以及c)检测这些经标记的核酸中的至少一个,由此确定在多个样品中的多个核酸的计数。该多个样品中的一个或多个可以包括单细胞或细胞裂解物。该多个样品中的一个或多个可以由单细胞组成。样品标签可以包括细胞标记,该细胞标记鉴定经标记的核酸从中起源的细胞。该多个由单细胞组成的样品可以来自一个或多个来源。样品标签可以包括鉴定单细胞的来源的样品索引区。分子鉴定物标记可以称为分子标记。该多个样品中的一个或多个可以包括少于1,000,000个细胞。该多个样品中的一个或多个可以包括少于100,000个细胞。该多个样品中的一个或多个可以包括少于10,000个细胞。该多个样品中的一个或多个可以包括少于1,000个细胞。该多个样品中的一个或多个可以包括少于100个细胞。该多个样品中的一个或多个可以包括细胞裂解物。

[0833] 可替代地,用于分析多个样品中的分子的方法可以包括:a)产生多个经标记的核酸包括:i)使第一样品与第一多个样品标签接触,其中该第一多个样品标签包括相同核酸序列;ii)使第一样品与第一多个包括不同核酸序列的分子鉴定物标记接触,由此产生多个第一经标记的核酸;iii)使第二样品与第二多个样品标签接触,其中该第二多个样品标签包括相同核酸序列;iv)使第二样品与第二多个包括不同核酸序列的分子鉴定物标记接触,由此产生多个第二经标记的核酸,其中该多个经标记的核酸包括该多个第一经标记的核酸和第二经标记的核酸;并且b)确定多个不同的经标记的核酸,由此确定在多个样品中的多个核酸的计数。样品标签可以包括细胞标记,该细胞标记鉴定经标记的核酸从中起源的细胞。样品标签可以包括鉴定单细胞的来源的样品索引区。分子鉴定物标记可以称为分子标记。

[0834] 可替代地,用于分析多个样品中的分子的方法可以包括:a)使多个包括两个或更多个不同核酸的样品与多个样品标签和多个分子鉴定物标记接触,以产生多个经标记的核酸,其中:i)该多个经标记的核酸包括附接至两个或更多个样品标签和两个或更多个分子鉴定物标记的两个或更多个核酸;ii)附接至来自该多个样品的第一样品的核酸的样品标签不同于附接至来自该多个样品的第二样品的核酸分子的样品标签;并且iii)在同一样品中的两个或更多个相同的核酸附接至两个或更多个不同的分子鉴定物标记;并且b)检测这些经标记的核酸的至少一部分,由此确定该多个样品中的两个或更多个不同核酸的计数。样品标签可以包括细胞标记,该细胞标记鉴定经标记的核酸从中起源的细胞。样品标签可以包括鉴定单细胞的来源的样品索引区。分子鉴定物标记可以称为分子标记。

[0835] 在此进一步披露的是用于分析多个样品中的分子的方法,该方法包括:a)使来自多个样品的第一样品第一多个分子与第一组分子条码接触,以产生第一多个经标记的分子,其中该第一多个分子条码的分子条码包括标记区和样品索引区;b)使来自该多个样品的第二样品的第二多个分子与第二组分子条码接触,以产生第二多个经标记的分子,其中

该第二多个分子条码的分子条码包括标记区和样品索引区，并且其中该第一多个分子条码和该第二多个分子条码至少区别在于这些分子条码的样品索引区；以及c)检测该第一多个经标记的分子的两个或更多个分子的至少一部分，和该第二多个经标记的分子的两个或更多个分子的至少一部分，由此确定在该多个样品中的两个或更多个分子的计数。该第一多个分子可以包括核酸分子。该第二多个分子可以包括核酸分子。标记区可以称为分子标记。分子条码可以进一步包括细胞标记。在其中该多个样品的样品由单细胞组成的实例中，该样品索引区可以称为细胞标记。

[0836] 在此披露的是选择定制引物的方法，该方法包括：a)第一遍，其中选择的引物包括：i)不多于三个连续的鸟嘌呤，不多于三个连续的胞嘧啶，不多于四个连续的腺嘌呤，以及不多于四个连续的胸腺嘧啶；ii)作为鸟嘌呤或胞嘧啶的至少3、4、5或6个核苷酸；以及iii)并不易于形成发夹结构的序列；b)第二遍，包括：i)第一轮选择具有所有转录物的高覆盖的多个序列；以及ii)随后的一轮或多轮，选择以下序列，该序列具有剩余转录物的最高覆盖和不多于4的与其他选择的序列的互补分数；并且c)添加序列至挑选的组，直至覆盖饱和或定制引物的总数小于或等于约96。

[0837] 在此进一步披露的是用于产生经标记的分子文库的方法，该方法包括：a)通过以下产生多个样品标签化核酸：i)使包括多个核酸的第一样品与多个第一样品标签接触，以产生多个第一样品标签化核酸；并且ii)使包括多个核酸的第二样品与多个第二样品标签接触，以产生多个第二样品标签化核酸，其中该多个第一样品标签不同于这些第二样品标签；并且b)使该多个样品标签化核酸与多个分子鉴定物标记接触，以产生多个经标记的核酸，由此产生经标记的核酸文库。

[0838] 在此披露的是用于分析多个样品中的分子的试剂盒。该试剂盒可以包括：a)两个或更多个组的分子条码，其中一个或多个分子条码的组中的分子条码包括样品索引区和标记区，其中(i)一组分子条码中的分子条码的样品索引区是相同的；以及(ii)第一组分子条码的样品索引区不同于第二组分子条码的样品索引区；以及b)多个珠。可以将两个或更多个组的分子条码附接至该多个珠。可以将两个或更多个组的分子条码缀合至该珠。标记区可以称为分子标记。分子条码可以进一步包括细胞标记。在其中该多个样品中的样品由单细胞组成的实例中，该样品索引区可以称为细胞标记。

[0839] 用于分析多个样品中的分子的试剂盒可以包括：a)包括第一多个分子条码的第一容器，其中(i)分子条码包括样品索引区和标记区；(ii)该第一多个分子条码的分子条码的总数中的至少约80％的样品索引区是相同的；以及(iii)该第一多个分子条码中的两个或更多个条码的标记区是不同的；以及(b)包括第二多个分子条码的第二容器，其中(i)分子条码包括样品索引区和标记区；(ii)该第一多个分子条码的分子条码的总数中的至少约80％的样品索引区是相同的；以及(iii)该第一多个分子条码中的两个或更多个条码的标记区是不同的；其中该第一多个分子条码的样品索引区不同于该第二多个分子条码的样品索引区。标记区可以称为分子标记。分子条码可以进一步包括细胞标记。在其中该多个样品的样品由单细胞组成的实例中，该样品索引区可以称为细胞标记。

[0840] 可替代地，用于分析多个样品中的分子的试剂盒包括：a)包括第一多个样品标签的第一容器，其中该多个样品标签包括相同核酸序列；以及b)包括第一多个分子鉴定物标记的第二容器，其中该多个分子鉴定物标记包括两个或更多个不同核酸序列。标记区可以

称为分子标记。在其中该多个样品中的样品由单细胞组成的实例中，该样品标签可以称为细胞标记。该试剂盒可以进一步包括：包括第一多个细胞标记的第三容器，其中该多个细胞标记包括两个或更多个不同核酸序列。

[0841] 在此披露的试剂盒和方法可以包括一个或多个组的分子条码。在此披露的试剂盒和方法可以包括一个或多个分子条码。分子条码可以包括样品索引区、分子标记区、细胞标记区、或它们的组合。一组分子条码中的至少两个分子条码可以包括两个或更多个不同标记区。两个或更多个组的分子条码中的两个或更多个分子条码的标记区可以是相同的。两个或更多个组的分子条码可以包括：包括相同标记区的分子条码。在其中该多个样品中的样品由单细胞组成的实例中，该样品标签可以称为细胞标记。

[0842] 在此披露的分子条码可以包括样品索引区。两个或更多个组的分子条码中的分子条码的样品索引区可以是不同的。样品索引区可以包括一个或多个核苷酸。两个或更多个不同组的分子条码的样品索引区的两个或更多个序列可以小于约90％、85％、80％、75％、70％、65％、60％、55％、50％、45％、40％、35％、30％、25％、20％、15％、10％、或5％同源。两个或更多个不同组的分子条码的样品索引区的两个或更多个序列可以小于约80％同源。两个或更多个不同组的分子条码的样品索引区的两个或更多个序列可以小于约60％同源。两个或更多个不同组的分子条码的样品索引区的两个或更多个序列可以小于约40％同源。两个或更多个不同组的分子条码的样品索引区的两个或更多个序列可以小于约20％同源。

[0843] 在此披露的分子条码可以包括细胞标记。两个或更多个组的分子条码中的分子条码的细胞标记可以是不同的。细胞标记可以包括一个或多个核苷酸。两个或更多个不同组的分子条码的细胞标记的两个或更多个序列可以小于约90％、85％、80％、75％、70％、65％、60％、55％、50％、45％、40％、35％、30％、25％、20％、15％、10％、或5％同源。两个或更多个不同组的分子条码的细胞标记的两个或更多个序列可以小于约80％同源。两个或更多个不同组的分子条码的细胞标记的两个或更多个序列可以小于约60％同源。两个或更多个不同组的分子条码的细胞标记的两个或更多个序列可以小于约40％同源。两个或更多个不同组的分子条码的细胞标记的两个或更多个序列可以小于约20％同源。

[0844] 在此披露的分子条码可以进一步包括通用PCR区。分子条码可以进一步包括靶特异性区。分子条码可以包括一个或多个核苷酸。标记区可以包括一个或多个核苷酸。样品索引区可以包括一个或多个核苷酸。通用PCR区可以包括一个或多个核苷酸。靶特异性区可以包括一个或多个核苷酸。

[0845] 在此披露的试剂盒和方法可以包括一个或多个组的样品标签。在此披露的试剂盒和方法可以包括一个或多个样品标签。样品标签可以包括样品索引区。第一样品标签组的样品标签的样品索引区可以不同于第二样品标签组的样品标签的样品索引区。样品索引区可以包括一个或多个核苷酸。两个或更多个不同组的样品标签的样品索引区的两个或更多个序列可以小于约90％、85％、80％、75％、70％、65％、60％、55％、50％、45％、40％、35％、30％、25％、20％、15％、10％、或5％同源。两个或更多个不同组的样品标签的样品索引区的两个或更多个序列可以小于约80％同源。两个或更多个不同组的样品标签的样品索引区的两个或更多个序列可以小于约60％同源。两个或更多个不同组的样品标签的样品索引区的两个或更多个序列可以小于约40％同源。两个或更多个不同组的样品标签的样品索引区的两个或更多个序列可以小于约20％同源。

[0846] 在此披露的试剂盒和方法可以包括一个或多个组的分子鉴定物标记。在此披露的试剂盒和方法可以包括一个或多个分子鉴定物标记。分子鉴定物标记可以包括标记区。一组分子鉴定物标记中的两个或更多个分子鉴定物标记的标记区可以是不同的。标记区可以包括一个或多个核苷酸。一组分子鉴定物标记中的两个或更多个分子鉴定物标记的标记区的序列可以小于约90％、85％、80％、75％、70％、65％、60％、55％、50％、45％、40％、35％、30％、25％、20％、15％、10％、或5％同源。一组分子鉴定物标记中的两个或更多个分子鉴定物标记的标记区的序列可以小于约80％同源。一组分子鉴定物标记中的两个或更多个分子鉴定物标记的标记区的序列可以小于约60％同源。一组分子鉴定物标记中的两个或更多个分子鉴定物标记的标记区的序列可以小于约40％同源。一组分子鉴定物标记中的两个或更多个分子鉴定物标记的标记区的序列可以小于约20％同源。标记区可以称为细胞标记区。

[0847] 在此披露的试剂盒和方法可以进一步包括一个或多个引物。该一个或多个引物可以包括与通用PCR区至少部分互补的序列。该一个或多个引物可以包括与通用PCR区至少约50％互补的序列。该一个或多个引物可以包括与通用PCR区至少约80％互补的序列。

[0848] 在此披露的试剂盒和方法可以进一步包括一种或多种扩增剂。扩增剂可以包括固定的引物面板。扩增剂可以包括一个或多个定制引物。扩增剂可以包括一个或多个对照引物。扩增剂可以包括一个或多个管家基因引物。扩增剂可以包括一种或多种PCR试剂。该一种或多种PCR试剂可以包括聚合酶、脱氧核糖核苷三磷酸(dNTP)、缓冲液、或它们的组合。

[0849] 在此披露的试剂盒和方法可以进一步包括一个或多个珠。可以将分子条码附接至该一个或多个珠。可以将样品标签附接至该一个或多个珠。可以将分子鉴定物标记附接至该一个或多个珠。

[0850] 在此进一步披露的是用于产生一个或多个组的珠的方法。该方法可以包括:a)使多个第一核酸沉积进多个孔中,其中该多个孔的两个或更多个不同的孔可以包括该多个核酸的两个或更多个不同核酸;b)使该多个孔的一个或多个孔与一个或较少的珠接触,以产生多个单标记珠,其中多个第一经标记的珠的单标记珠包括附接至该多个第一核酸的核酸的珠;c)从该多个孔汇聚该多个第一经标记的珠以产生第一经标记的珠的聚池;d)将第一经标记的珠的该聚池分配至随后的多个孔,其中该随后多个孔中的两个或更多个孔包括多个随后的核酸中的两个或更多个不同核酸;以及e)将该多个随后的核酸的一个或多个核酸附接至一个或多个第一经标记的珠,以产生多个经唯一性标记的珠。

[0851] 在此披露的方法和试剂盒可以用于分析多个核酸。在此披露的方法和试剂盒可以用于分析少于约100,000,000个核酸。在此披露的方法和试剂盒可以用于分析少于约10,000,000个核酸。在此披露的方法和试剂盒可以用于分析少于约1,000,000个核酸。在此进一步披露的是分析多个蛋白的方法。该方法可以包括:a)通过以下产生多个样品标签化多肽:i)使包括多个多肽的第一样品与多个第一样品标签接触,以产生多个第一样品标签化多肽;并且ii)使包括多个多肽的第二样品与多个第二样品标签接触,以产生多个第二样品标签化多肽,其中该多个第一样品标签不同于该多个第二样品标签;b)使该多个样品标签化多肽与多个分子鉴定物标记接触,以产生多个经标记的多肽;并且c)检测这些经标记的多肽的至少一部分,由此确定在该多个样品中的该多个多肽的计数。

[0852] 分析多个样品中的多肽的方法可以进一步包括确定一个或多个经标记的多肽的身份。确定该一个或多个经标记的多肽的身份可以包括质谱法。该方法可以进一步包括将

第一样品的经标记的多肽与第二样品的经标记的多肽组合。可以在确定不同的经标记的多肽的数目之前组合经标记的多肽。该方法可以进一步包括将第一样品标签化多肽和第二样品标签化多肽组合。可以在与该多个分子鉴定物标记接触之前组合第一样品标签化多肽和第二样品标签化多肽。确定不同的经标记的多肽的数目可以包括检测标签化的经标记的多肽的至少一部分。检测标签化的经标记的多肽的至少一部分标记可以包括检测样品标签、分子特异性标签、多肽、或它们的组合的至少一部分。

[0853] 在此披露的方法可以包括使多个样品与多个样品标签和多个分子鉴定物标记接触。使该多个样品与该多个样品标签和该多个分子鉴定物标记接触可以同时发生。使该多个样品与该多个样品标签和该多个分子鉴定物标记接触可以并行发生。使该多个样品与该多个样品标签和该多个分子鉴定物标记接触可以顺序发生。使该多个样品与该多个样品标签接触可以发生在使该多个样品与该多个分子鉴定物标记接触之前。使该多个样品与该多个样品标签接触可以发生在使该多个样品与该多个分子鉴定物标记接触之后。

[0854] 在此披露的方法可以包括使第一样品与第一多个样品标签和第一多个分子鉴定物标记接触。使第一样品与第一多个样品标签和第一多个分子鉴定物标记接触可以同时发生。使第一样品与第一多个样品标签和第一多个分子鉴定物标记接触可以并行发生。使第一样品与第一多个样品标签和第一多个分子鉴定物标记接触可以顺序发生。使第一样品与第一多个样品标签接触可以发生在使第一样品与第一多个分子鉴定物标记接触之前。使第一样品与第一多个样品标签接触可以发生在使第一样品与第一多个分子鉴定物标记接触之后。

[0855] 在此披露的方法可以包括使第二样品与第二多个样品标签和第二多个分子鉴定物标记接触。使第二样品与第二多个样品标签和第二多个分子鉴定物标记接触可以同时发生。使第二样品与第二多个样品标签和第二多个分子鉴定物标记接触可以并行发生。使第二样品与第二多个样品标签和第二多个分子鉴定物标记接触可以顺序发生。使第二样品与第二多个样品标签接触可以发生在使第二样品与第二多个分子鉴定物标记接触之前。使第二样品与第二多个样品标签接触可以发生在使第二样品与第二多个分子鉴定物标记接触之后。

[0856] 在此披露的方法和试剂盒可以进一步包括组合两个或更多个样品。在此披露的试剂盒和方法可以进一步包括组合第一样品和第二样品。可以在与该多个分子鉴定物标记接触之前组合第一和第二样品。可以在检测经标记的核酸之前组合第一和第二样品。可以在随机地标记两个或更多个样品中的两个或更多个分子之前组合这两个或更多个样品。可以在随机地标记两个或更多个样品中的两个或更多个分子之后组合这两个或更多个样品。可以在检测两个或更多个样品中的两个或更多个分子之前组合这两个或更多个样品。可以在检测两个或更多个样品中的两个或更多个分子之后组合这两个或更多个样品。可以在分析两个或更多个样品中的两个或更多个分子之前组合这两个或更多个样品。可以在分析两个或更多个样品中的两个或更多个分子之后组合这两个或更多个样品。可以在对两个或更多个样品中的两个或更多个分子进行一个或多个测定之前组合这两个或更多个样品。可以在对两个或更多个样品中的两个或更多个分子进行一个或多个测定之后组合这两个或更多个样品。

[0857] 在此披露的方法和试剂盒可以包括对样品中的两个或更多个分子进行一个或多

个测定。该一个或多个测定可以包括一个或多个扩增反应。在此披露的方法和试剂盒可以进一步包括进行一个或多个扩增反应以产生经标记的核酸扩增子。可以在检测经标记的核酸之前扩增这些经标记的核酸。本发明进一步包括在进行一个或多个扩增反应之前组合第一和第二样品。

[0858] 扩增反应可以包括扩增至少一部分样品标签。扩增反应可以包括扩增至少一部分标记。扩增反应可以包括扩增至少一部分样品标签、标记、核酸、或它们的组合。扩增反应可以包括扩增多个核酸的核酸总数的至少约1%、5%、10%、15%、20%、25%、30%、35%、40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、95%、或100%。扩增反应可以包括扩增多个核酸的核酸总数的至少约1%。扩增反应可以包括扩增多个核酸的核酸总数的至少约5%。扩增反应可以包括扩增多个经标记的核酸的经标记的核酸的总数的至少约1%、5%、10%、15%、20%、25%、30%、35%、40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、95%、或100%。扩增反应可以包括扩增多个经标记的核酸的经标记的核酸的至少约1%。扩增反应可以包括扩增多个经标记的核酸的经标记的核酸的总数的至少约5%的总数。扩增反应可以包括扩增多个经标记的核酸的经标记的核酸的总数的至少约10%。扩增反应可以包括扩增多个核酸的核酸的总数的少于约95%、90%、80%、70%、60%或50%。扩增反应可以包括扩增多个核酸的核酸的总数的少于约50%。扩增反应可以包括扩增多个核酸的核酸的总数的少于约20%。扩增反应可以包括扩增多个核酸的核酸的总数的少于约10%。扩增反应可以包括扩增多个经标记的核酸的经标记的核酸的总数的少于约95%、90%、80%、70%、60%或50%。扩增反应可以包括扩增多个经标记的核酸的经标记的核酸的总数的少于约40%。扩增反应可以包括扩增多个经标记的核酸的经标记的核酸的总数的少于约25%。扩增反应可以包括扩增多个经标记的核酸的经标记的核酸的总数的少于约10%。

[0859] 该一个或多个扩增反应可以导致样品中约1、2、5、10、15、20、25、30、35、40、45、50、55、60、65、70、75、80、85、90、95、100、200、300、400、500、600、700、800、900或1000个标靶核酸的扩增。该一个或多个扩增反应可以导致样品中约2000个目标核酸的扩增。该一个或多个扩增反应可以导致样品中约1000个目标核酸的扩增。该一个或多个扩增反应可以导致约2000个目标分子的扩增。该一个或多个扩增反应可以导致样品中约100个目标核酸的扩增。

[0860] 扩增反应可以包括一个或多个聚合酶链式反应(PCR)。该一个或多个多重聚合酶链反应可以包括多重PCR、巢式PCR、绝对PCR、HD-PCR、下一代PCR、数字RTA、或它们的任何组合。该一个或多个聚合酶链式反应可以包括多重PCR。该一个或多个聚合酶链式反应可以包括巢式PCR。

[0861] 进行一个或多个扩增反应可以包括使用一个或多个引物。该一个或多个引物可以包括一个或多个寡核苷酸。该一个或多个寡核苷酸可以包括至少约7-9个核苷酸。该一个或多个寡核苷酸可以包括少于12-15个核苷酸。该一个或多个引物可以退火至多个经标记的核酸的至少一部分。该一个或多个引物可以退火至多个经标记的核酸的3’端和/或5’端。该一个或多个引物可以退火至多个经标记的核酸的内部区。内部区可以是从该多个经标记的核酸的3’端的至少约50、100、150、200、220、230、240、250、260、270、280、290、300、310、320、330、340、350、360、370、380、390、400、410、420、430、440、450、460、470、480、490、500、510、520、530、540、550、560、570、580、590、600、650、700、750、800、850、900或1000个核苷酸。内

部区可以是从该多个经标记的核酸的3’端的至少约2000个核苷酸。该一个或多个引物可以包括固定的引物面板。该一个或多个引物可以包括至少一个或多个定制引物。该一个或多个引物可以包括至少一个或多个对照引物。该一个或多个引物可以包括至少一个或多个管家基因引物。该一个或多个寡核苷酸可以包括选自下组的序列，该组由表1中的序列组成。该一个或多个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。通用引物可以退火至通用PCR区。该一个或多个定制引物可以退火至样品标签的至少一部分。该一个或多个定制引物可以退火至分子鉴定物标记的至少一部分。该一个或多个定制引物可以退火至分子条码的至少一部分。该一个或多个定制引物可以退火至第一样品标签、第二样品标签、分子鉴定物标记、核酸或它们的产物。该一个或多个引物可以包括通用引物和定制引物。该一个或多个引物可以包括至少约96个或更多个定制引物。该一个或多个引物可以包括至少约960个或更多个定制引物。该一个或多个引物可以包括至少约9600个或更多个定制引物。该一个或多个定制引物可以退火至两个或更多个不同的经标记的核酸。该两个或更多个不同的经标记的核酸可以对应于一个或多个基因。

[0862] 多重PCR反应可以包括巢式PCR反应。巢式PCR反应可以包括：包括第一引物和第二引物的一对引物。第一引物可以退火至该多个核酸中的一个或多个核酸的区。该一个或多个核酸的区可以是从该一个或多个核酸的3’端至少约300至400个核苷酸。第二引物可以退火至该多个核酸中的一个或多个核酸的区。该一个或多个核酸的区可以是从该一个或多个核酸的3’端至少200至300个核苷酸。

[0863] 在此披露的方法和试剂盒可以进一步包括进行一个或多个cDNA合成反应以产生这些分子或它们的衍生物(例如经标记的分子)的一个或多个cDNA拷贝。该一个或多个cDNA合成反应可以包括一个或多个逆转录反应。

[0864] 在此披露的方法和试剂盒可以包括一个或多个样品。在此披露的方法和试剂盒可以包括多个样品。该多个样品可以包括至少约2、3、4、5、10、20、30、40、50、60、70、80、90或100或更多个样品。该多个样品可以包括至少约100、200、300、400、500、600、700、800、900或1000或更多个样品。该多个样品可以包括至少约1000、2000、3000、4000、5000、6000、7000、8000个样品、9000、或10,000个样品，或100,000个样品，或1,000,000个或更多个样品。该多个样品可以包括至少约10,000个样品。该多个样品可以包括至少约2个样品。该多个样品可以包括至少约5个样品。该多个样品可以包括至少约10个样品。该多个样品可以包括至少约50个样品。该多个样品可以包括至少约100个样品。

[0865] 在此披露的方法和试剂盒可以包括一个或多个样品，该一个或多个样品包括一个或多个细胞。在此披露的方法和试剂盒可以包括两个或更多个样品，该两个或更多个包括一个或多个细胞。第一样品可以包括一个或多个细胞。第二样品可以包括一个或多个细胞。第一样品的一个或多个细胞可以具有与第二样品的一个或多个细胞相同的细胞类型。

[0866] 在此披露的方法和试剂盒可以包括多个样品。该多个样品可以来自一个或多个受试者。该多个样品可以来自两个或更多个受试者。该多个样品可以来自同一个受试者。该两个或更多个受试者可以来自同一物种。该两个或更多个受试者可以来自不同物种。该多个样品可以来自一个或多个来源。该多个样品可以来自两个或更多个来源。该多个样品可以来自同一个受试者。该两个或更多个来源可以来自同一物种。该两个或更多个来源可以来自不同物种。

[0867] 可以并行获得该多个样品。可以顺序获得该多个样品。可以经两个或更多个时段获得该多个样品。该两个或更多个时段可以相距一个或多个小时。该两个或更多个时段可以相距一日或多日。该两个或更多个时段可以相距一周或多周。该两个或更多个时段可以相距一个或多个月。该两个或更多个时段可以相距一年或多年。

[0868] 该多个样品可以来自一种或多种体液、组织、细胞、器官、或肌肉。该多个样品可以包括一个或多个血液样品。

[0869] 在此披露的方法和试剂盒可以包括一个或多个样品，该一个或多个样品包括一个或多个核酸。两个或更多个样品可以包括一个或多个核酸。两个或更多个样品可以包括两个或更多个核酸。第一样品的一个或多个核酸可以不同于第二样品的一个或多个核酸。第一样品中的核酸可以与第二样品中的核酸至少约50％相同。第一样品中的核酸可以与第二样品中的核酸至少约70％相同。第一样品中的核酸可以与第二样品中的核酸至少约80％相同。

[0870] 该一个或多个样品中的多个核酸可以包括两个或更多个相同序列。该一个或多个样品中的至少约5％、10％、15％、20％、25％、30％、35％、40％、45％、50％、55％、60％、65％、70％、75％、80％、85％、90％、95％、100％的总核酸可以包括相同序列。一个或多个样品中的多个核酸可以包括至少两个不同序列。该一个或多个样品中的至少约5％、10％、15％、20％、25％、30％、35％、40％、45％、50％、55％、60％、65％、70％、75％、80％、85％、90％、95％、100％的总核酸可以包括不同序列。

[0871] 该多个核酸可以包括RNA、DNA、cDNA、mRNA、基因组DNA、小RNA、非编码RNA、或细胞的其他核酸内容物。该多个核酸可以包括mRNA。该多个核酸可以包括RNA。该多个核酸可以包括mRNA。该多个核酸可以包括DNA。

[0872] 在此披露的方法和试剂盒可以包括一个或多个样品标签。在此披露的方法和试剂盒可以包括一个或多个的多个样品标签。这些样品标签可以包括样品索引区。第一多个样品标签的样品索引区可以不同于第二多个样品标签的样品索引区。这些样品标签可以包括一个或多个核苷酸。

[0873] 这些样品标签可以包括至少约5、10、15、20、25、30、35、40、45、50、55、60、65、70、75、80、85、90、95或100个核苷酸。这些样品标签可以包括至少约5个或更多个核苷酸。这些样品标签可以包括至少约10个或更多个核苷酸。这些样品标签可以包括少于约200个核苷酸。这些样品标签可以包括少于约100个核苷酸。这些样品标签可以包括少于约60个核苷酸。

[0874] 这些样品标签可以进一步包括通用引物结合位点。这些样品标签可以进一步包括通用PCR区。这些样品标签可以进一步包括一个或多个适配子区。这些样品标签可以进一步包括一个或多个靶特异性区。

[0875] 在此披露的方法和试剂盒可以包括一个或多个分子鉴定物标记。在此披露的方法和试剂盒可以包括一个或多个的多个分子鉴定物标记。该一个或多个的多个分子鉴定物标记可以包括两个或更多个不同分子鉴定物标记。该一个或多个的多个分子鉴定物标记可以包括50或更多个不同分子鉴定物标记。该一个或多个的多个分子鉴定物标记可以包括90或更多个不同分子鉴定物标记。该一个或多个的多个分子鉴定物标记可以包括100或更多个不同分子鉴定物标记。该一个或多个的多个分子鉴定物标记可以包括300或更多个不同分

子鉴定物标记。该一个或多个的多个分子鉴定物标记可以包括500或更多个不同分子鉴定物标记。该一个或多个的多个分子鉴定物标记可以包括960或更多个不同分子鉴定物标记。该一个或多个的多个分子鉴定物标记可以包括一个或多个分子鉴定物标记的多个拷贝。两个或更多个的多个分子鉴定物标记可以包括一个或多个相同分子鉴定物标记。两个或更多个的多个分子鉴定物标记可以包括10或更多个相同分子鉴定物标记。第一多个分子鉴定物标记的分子鉴定物标记可以与第二多个分子鉴定物标记的分子鉴定物标记至少约30％相同。第一多个分子鉴定物标记的分子鉴定物标记可以与第二多个分子鉴定物标记的分子鉴定物标记至少约50％相同。第一多个分子鉴定物标记的分子鉴定物标记可以与第二多个分子鉴定物标记的分子鉴定物标记至少约80％相同。

[0876] 分子鉴定物标记可以包括标记区(例如分子标记区、分子索引区)。第一多个的分子鉴定物标记的两个或更多个分子鉴定物标记的标记区可以是不同的。一个或多个的多个分子鉴定物标记可以包括至少约20个不同标记区。一个或多个的多个分子鉴定物标记可以包括至少约50个不同标记区。一个或多个的多个分子鉴定物标记可以包括至少约96个不同标记区。一个或多个的多个分子鉴定物标记可以包括至少约200个不同标记区。一个或多个的多个分子鉴定物标记可以包括至少约500个不同标记区。一个或多个的多个分子鉴定物标记可以包括至少约960个不同标记区。

[0877] 这些分子鉴定物标记可以包括至少约1、2、3、4、5、6、7、8、9、10或更多个核苷酸。这些分子鉴定物标记可以包括至少约20、30、40、50或更多个核苷酸。这些分子鉴定物标记可以包括至少约21个核苷酸。

[0878] 这些分子鉴定物标记可以进一步包括靶特异性区。靶特异性区可以包括寡dT序列。

[0879] 这些分子鉴定物标记可以进一步包括一个或多个染料标记。这些分子鉴定物标记可以进一步包括Cy3染料。这些分子鉴定物标记可以进一步包括Tye563染料。

[0880] 在此披露的方法和试剂盒可以包括一个或多个经标记的分子。可以通过使多个分子与多个样品标签接触来产生一个或多个经标记的分子。可以通过使多个核酸与多个样品标签接触来产生一个或多个经标记的分子。使多个核酸与多个样品标签接触可以包括将一个或多个样品标签连接至一个或多个核酸。使多个核酸与多个样品标签接触可以包括将一个或多个样品标签杂交至一个或多个核酸。使多个核酸与多个样品标签接触可以包括进行一个或多个核酸增长反应。该一个或多个核酸增长反应可以包括逆转录。

[0881] 在此披露的方法和试剂盒可以进一步包括将一个或多个寡核苷酸接头附接至该多个核酸。这些方法和试剂盒可以进一步包括将一个或多个寡核苷酸接头附接至该样品标签化核酸。这些方法和试剂盒可以进一步包括将一个或多个寡核苷酸接头附接至这些经标记的核酸。该一个或多个接头可以包括至少约10个核苷酸。

[0882] 在此披露的方法和试剂盒可以进一步包括将一个或多个经标记的核酸附接至支持物。该支持物可以包括固体支持物。该支持物可以包括珠。支持物可以包括阵列。支持物可以包括载玻片。

[0883] 将经标记的核酸附接至支持物可以包括胺-硫醇交联、马来酰亚胺交联、N-羟基琥珀酰亚胺或N-羟基硫代琥珀酰亚胺、泽农(Zenon)、位点点击(SiteClick)、或它们的组合。将经标记的核酸附接至支持物可以包括将生物素附接至一个或多个经标记的核酸。

[0884] 支持物可以包括一个或多个珠。该一个或多个珠可以包括经包覆的珠。经包覆的珠可以包覆有链亲和素。

[0885] 支持物可以包括阵列。支持物可以包括一个或多个探针。可以将经标记的核酸附接至该一个或多个探针。该一个或多个探针可以包括一个或多个寡核苷酸。可以将该一个或多个探针附接至经标记的核酸的至少一部分。附接至该一个或多个探针的经标记的核酸的部分可以包括至少一部分的样品标签、分子鉴定物标记、分子条码、核酸、或它们的组合。

[0886] 支持物可以包括载玻片。该载玻片可以包括一个或多个孔。可以在载玻片上蚀刻一个或多个孔。该一个或多个孔可以包括至少960个孔。该载玻片可以包括一个或多个探针。可以将该一个或多个探针打印到该载玻片上。该一个或多个孔可以进一步包括一个或多个探针。可以将该一个或多个探针打印到该一个或多个孔内。一个或多个探针可以包括960个核酸。这些核酸可以是不同的。这些核酸可以是相同的。

[0887] 在此披露的方法和试剂盒可以用于确定在一个或多个样品中的一个或多个分子的计数。确定一个或多个分子的计数可以包括确定不同的经标记的核酸的数目。确定不同的经标记的核酸的数目可以包括检测经标记的核酸的至少一部分。检测经标记的核酸的至少一部分可以包括检测至少一部分的样品标签、分子鉴定物标记、分子条码、核酸、或它们的组合。

[0888] 确定不同的经标记的核酸的数目可以包括测序。测序可以包括MiSeq测序。测序可以包括HiSeq测序。确定不同的经标记的核酸的数目可以包括阵列。确定不同的经标记的核酸的数目可以包括使经标记的核酸与一个或多个探针接触。

[0889] 确定不同的经标记的核酸的数目可以包括使经标记的核酸与阵列接触。阵列可以包括多个探针。确定不同的经标记的核酸的数目可以包括使经标记的核酸与多个探针的载玻片接触。

[0890] 确定不同的经标记的核酸的数目可以包括经标记的探针杂交、靶特异性扩增、靶特异性测序、用特异性针对靶小核苷酸多态性的经标记的核苷酸进行测序、用特异性针对限制性内切酶消化模式的经标记的核苷酸进行测序、用特异性针对突变的经标记的核苷酸进行测序、或它们的组合。

[0891] 确定不同的经标记的核酸的数目可以包括序列特异性标记的流式细胞仪分选。确定不同的经标记的核酸的数目可以包括检测附接至珠的经标记的核酸。检测附接至珠的经标记的核酸可以包括荧光检测。

[0892] 确定不同的经标记的核酸的数目可以包括通过靶特异性探针和经标记的核酸或靶特异性的经标记的标记探针之间的荧光共振能量转移(FRET),对多个经标记的核酸进行计数。确定不同的经标记的核酸的数目可以包括将经标记的核酸附接至支持物。

[0893] 在此披露的方法和试剂盒可以进一步包括靶序列与核酸结合蛋白的免疫沉淀反应。

[0894] 在此披露的方法和试剂盒可以进一步包括将多个样品分配进微孔板的多个孔中。该多个样品中的一个或多个可以包括多个细胞。该多个样品中的一个或多个可以包括多个核酸。在此披露的方法和试剂盒可以进一步包括将一个或更少的细胞分配至多个孔。可以在微孔板中裂解该多个细胞。在此披露的方法和试剂盒可以进一步包括在微孔板中合成cDNA。合成cDNA可以包括mRNA的逆转录。

[0895] 在此披露的方法和试剂盒可以进一步包括将多个第一样品标签、多个第二样品标签、多个分子鉴定物标记或其任何组合分配进微孔板中。

[0896] 在此披露的方法和试剂盒可以进一步包括将一个或多个珠分配进微孔板中。微孔板可以包括在PDMS上通过软光刻制造的、在硅晶片上蚀刻的、在载玻片上蚀刻的、在载玻片上光刻胶图案化或其组合的微孔板。微孔可以包括在微毛细管板上的洞。微孔板可以包括油包水乳剂。微孔板可以包括至少一个或多个孔。微孔板可以包括至少约6个孔、12个孔、48个孔、96个孔、384个孔、960个孔或1000个孔。

[0897] 在此披露的方法和试剂盒可以进一步包括芯片。微孔板可以附接至芯片。芯片可以包括至少约6个孔、12个孔、48个孔、96个孔、384个孔、960个孔、1000个孔、2000个孔、3000个孔、4000个孔、5000个孔、6000个孔、7000个孔、8000个孔、9000个孔、10,000个孔、20,000个孔、30,000个孔、40,000个孔、50,000个孔、60,000个孔、70,000个孔、80,000个孔、90,000个孔、100,000个孔、200,000个孔、500,000个孔、或一百万个孔。孔可以包括至少约300微米^2、400微米^2、500微米^2、600微米^2、700微米^2、800微米^2、900微米^2、1000微米^2、1100微米^2、1200微米^2、1300微米^2、1400微米^2、1500微米^{2}的面积。在此披露的方法和试剂盒可以进一步包括在芯片上分配约10,000和30,000个之间的样品。

[0898] 在此披露的方法和试剂盒可以进一步包括诊断对其有需要的受试者体内的病症、疾病或失调。

[0899] 在此披露的方法和试剂盒可以进一步包括预测对其有需要的受试者体内的病症、疾病或失调。在此披露的方法和试剂盒可以进一步包括确定用于对其有需要的受试者体内的病症、疾病或失调的治疗。

[0900] 该多个样品可以包括来自患有疾病或病症的受试者的一个或多个样品。该多个样品可以包括来自健康受试者的一个或多个样品。

[0901] 在此进一步披露的是法医分析的方法，该方法包括：a)随机地标记两个或更多个样品中的两个或更多个分子以产生两个或更多个经标记的分子；以及b)检测该两个或更多个经标记的分子。

[0902] 选择定制引物的方法可以进一步包括基于一个或多个核酸选择定制引物。该一个或多个核酸可以包括mRNA转录物、非编码转录物(包括结构RNA)、转录的假基因、由基因组注释过程提供的模型mRNA、对应于基因组重叠群的序列、或它们的任何组合。该一个或多个核酸可以是RNA。该一个或多个核酸可以是mRNA。该一个或多个核酸可以包括一个或多个外显子。选择定制引物的方法可以进一步包括富集一个或多个亚组的核酸。该一个或多个亚组包括低丰度mRNA。选择定制引物的方法可以进一步包括计算算法。

[0903] 在此披露的方法和试剂盒可以包括使用一个或多个对照。该一个或多个对照是掺入型对照。该一个或多个对照可以包括核酸。该一个或多个包括多个核酸的样品可以掺入有一个或多个对照核酸。该一个或多个对照核酸可以用于测量产生经标记的核酸文库的效率。

[0904] 在此披露的方法和试剂盒可以用于产生一个或多个核酸文库。该一个或多个核酸文库可以包括多个经标记的核酸或它们的衍生物(例如经标记的扩增子)。产生经标记的核酸文库的方法可以包括用两个或更多个组的分子条码随机地标记两个或更多个样品中的两个或更多个核酸，以产生多个经标记的核酸。产生经标记的核酸文库的方法可以包括使

两个或更多个样品与多个样品标签和多个分子特异性标记接触，以产生多个经标记的核酸。经标记的核酸可以包括样品索引区、标记区和核酸区。样品索引区可以用于赋予样品的或子样品的身份至核酸。样品索引区可以用于确定核酸的来源。标记区可以用于赋予独特身份至核酸，由此使得能够区别同一样品或子样品中的两个或更多个相同核酸。

[0905] 产生核酸文库的方法可以进一步包括扩增一个或多个经标记的核酸，以产生一个或多个富集的经标记的核酸。该方法可以进一步包括进行一个或多个富集的经标记的核酸的拉下实验(pull-down assay)。该方法可以进一步包括纯化该一个或多个富集的经标记的核酸。

[0906] 在此披露的试剂盒可以包括多个珠、引物和/或扩增剂。一个或多个试剂盒可以用于分析至少约10、20、30、40、50、60、70、80、90、或100或更多个样品或子样品。一个或多个试剂盒可以用于分析至少约96个样品。一个或多个试剂盒可以用于分析至少约384个样品。试剂盒可以进一步包括用于引物设计和优化的说明书。

[0907] 试剂盒可以进一步包括一个或多个微孔板。该一个或多个微孔板可以用于分配一个或多个珠。该一个或多个微孔板可以用于分配来自一个或多个样品的一个或多个分子或它们的衍生物(例如经标记的分子、经标记的扩增子)。

[0908] 试剂盒可以进一步包括一个或多个另外的容器。该一个或多个另外的容器可以包括一个或多个另外的多个样品标签。在一个或多个另外的容器中的一个或多个另外的多个样品标签可以不同于在第一容器中的第一多个样品标签。该一个或多个另外的容器可以包括一个或多个另外的多个分子鉴定物标记。该一个或多个另外的容器的一个或多个另外的多个分子鉴定物标记可以与第二容器的一个或多个另外的分子鉴定物标记至少约50％相同。该一个或多个另外的容器的一个或多个另外的多个分子鉴定物标记可以与第二容器的一个或多个另外的分子鉴定物标记至少约80％相同。该一个或多个另外的容器的一个或多个另外的多个分子鉴定物标记可以与第二容器的一个或多个另外的分子鉴定物标记至少约90％相同。

[0909] 在此进一步披露了产生一组或多组经标记的珠的方法。产生一组或多组经标记的珠的方法可以包括将一个或多个核酸附接至一个或多个珠，由此产生一组或多组的经标记的珠。该一个或多个核酸可以包括一个或多个分子条码。该一个或多个核酸可以包括一个或多个样品标签。该一个或多个核酸可以包括一个或多个分子鉴定物标记。该一个或多个核酸可以包括a)引物区；b)样品索引区；以及c)接头或适配子区。该一个或多个核酸可以包括a)引物区；b)标记区；以及c)接头或适配子区。该一个或多个核酸可以包括a)样品索引区；以及b)标记区。该一个或多个核酸可以进一步包括引物区。该一个或多个核酸可以进一步包括靶特异性区。该一个或多个核酸可以进一步包括接头区域。该一个或多个核酸可以进一步包括适配子区。该一个或多个核酸可以进一步包括样品索引区。该一个或多个核酸可以进一步包括标记区。

[0910] 用于一组经标记的珠的核酸的引物区可以是至少约70％相同。用于一组经标记的珠的核酸的引物区可以是至少约90％相同。用于一组经标记的珠的核酸的引物区可以是相同的。

[0911] 用于一组经标记的珠的核酸的样品索引区可以是至少约70％相同。用于一组经标记的珠的核酸的样品索引区可以是至少约90％相同。用于一组经标记的珠的核酸的样品索

引区可以是相同的。用于两个或更多个组的样品索引化珠的核酸的样品索引区可以是少于约40％相同。用于两个或更多个组的样品索引化珠的核酸的样品索引区可以是少于约50％相同。用于两个或更多个组的样品索引化珠的核酸的样品索引区可以是少于约60％相同。用于两个或更多个组的样品索引化珠的核酸的样品索引区可以是不同的。

[0912] 用于两个或更多个组的经标记的珠的核酸的标记区可以是至少约70％相同。用于两个或更多个组的经标记的珠的核酸的标记区可以是至少约90％相同。用于两个或更多个组的经标记的珠的核酸的标记区可以是相同的。用于一组经标记的珠的核酸的标记区可以是少于约40％相同。用于一组经标记的珠的核酸的标记区可以是少于约50％相同。用于一组经标记的珠的核酸的标记区可以是少于约60％相同。用于一组经标记的珠的两个或更多个核酸的标记区可以是不同的。

[0913] 用于一组经标记的珠的核酸的接头区或适配子区可以是至少约70％相同。用于一组经标记的珠的核酸的接头区或适配子区可以是至少约90％相同。用于一组经标记的珠的核酸的接头区或适配子区可以是相同的。

[0914] 用于两个或更多个组的靶特异性珠的核酸的靶特异性区可以是至少约70％相同。用于两个或更多个组的靶特异性珠的核酸的靶特异性区可以是至少约90％相同。用于两个或更多个组的靶特异性珠的核酸的靶特异性区可以是相同的。用于一组靶特异性珠的核酸的靶特异性区可以是少于约40％相同。用于一组靶特异性珠的核酸的靶特异性区可以是少于约50％相同。用于一组靶特异性珠的核酸的靶特异性区可以是少于约60％相同。用于一组靶特异性珠的两个或更多个核酸的靶特异性区可以是不同的。

[0915] 该一个或多个组的经标记的珠可以包括一百万或更多个经标记的珠。该一个或多个组的经标记的珠可以包括一千万或更多个经标记的珠。

[0916] 将该一个或多个核酸附接至珠可以包括共价附接。将该一个或多个核酸附接至珠可以包括缀合。将该一个或多个核酸附接至珠可以包括离子相互作用。

[0917] 这些珠可以是经包覆的珠。可以将核酸附接至一个或多个标签。这些珠可以包覆有链亲和素。可以将核酸附接至生物素。这些珠还可以包覆有抗体或核酸，并且可以经由此类表面包覆材料将核酸间接附接至这些珠。

[0918] 在一个方面中，本披露提供了一种组合物，该组合物包括：固体支持物，其中所述固体支持物包括多个寡核苷酸，其中所述多个寡核苷酸中的至少两个包括细胞标记和分子标记，其中所述多个寡核苷酸中的所述至少两个的所述细胞标记是相同的，并且其中所述多个寡核苷酸中的所述至少两个的所述分子标记是不同的。在一些实施例中，该多个寡核苷酸进一步包括样品标记。在一些实施例中，该多个寡核苷酸进一步包括靶结合区。在一些实施例中，靶结合区包括被适配为杂交至靶核酸的序列。在一些实施例中，靶结合区包括选自下组的序列，该组由以下各项组成：随机多聚体，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；基因特异性引物；和寡聚dT；或其任何组合。在一些实施例中，该多个寡核苷酸包括通用标记。在一些实施例中，该通用标记包括针对测序引物的结合位点。在一些实施例中，该多个寡核苷酸包括接头。在一些实施例中，该接头包括官能团。在一些实施例中，该接头位于所述寡核苷酸的5’。在一些实施例中，接头选自下组，该组由以下各项组成：C6、生物素、链亲和素、伯胺、醛、和酮，或它们的任何组合。在一些实施例中，该固体支持物由聚苯乙烯构成。在一些实施例中，

固体支持物是磁性的。在一些实施例中，该固体支持物选自下组，该组由以下各项组成：PDMS固体支持物、玻璃固体支持物、聚丙烯固体支持物、琼脂糖固体支持物、明胶固体支持物、磁性固体支持物、和普朗尼克(pluronic)固体支持物，或它们的任何组合。在一些实施例中，该固体支持物包括珠。在一些实施例中，该固体支持物包括约20微米的直径。在一些实施例中，该固体支持物包括从约5微米至约40微米的直径。在一些实施例中，该固体支持物包括官能团。在一些实施例中，该官能团选自下组，该组由以下各项组成：C6、生物素、链亲和素、伯胺、醛、和酮，或它们的任何组合。在一些实施例中，该细胞标记包括多个细胞标记。在一些实施例中，该多个细胞标记散布有多个接头标记序列。在一些实施例中，该多个寡核苷酸包括从10,000至10亿个寡核苷酸。

[0919] 在一个方面中，本披露提供了固体支持物，该固体支持物包括：第一寡核苷酸，该第一寡核苷酸包括：第一细胞标记，该第一细胞标记包括第一随机序列、第二随机序列、和第一接头标记序列，其中所述第一接头标记序列连接所述第一随机序列和所述第二随机序列；以及包括随机序列的第一分子标记；以及第二寡核苷酸，该第二寡核苷酸包括：第二细胞标记，该第二细胞标记包括第三随机序列、第四随机序列、和第二接头标记序列，其中所述第二接头标记序列连接所述第三随机序列和所述第四随机序列；以及包括随机序列的第二分子标记，其中所述第一细胞标记和所述第二细胞标记是相同的并且所述第一分子标记和所述第二分子标记是不同的。在一些实施例中，该第一和第二寡核苷酸进一步包括相同样品索引区。在一些实施例中，样品索引区包括随机序列。在一些实施例中，样品索引区是4-12个核苷酸长。在一些实施例中，细胞标记直接附接至所述分子标记。在一些实施例中，细胞标记和所述分子标记是通过接头标记序列附接的。在一些实施例中，所述细胞标记的随机序列是从4-12个核苷酸长。在一些实施例中，所述细胞标记的恒定序列是至少4个核苷酸长。在一些实施例中，细胞标记具有至少12个核苷酸的总长。在一些实施例中，细胞标记进一步包括一个或多个另外的随机序列。在一些实施例中，细胞标记进一步包括一个或多个另外的接头标记序列。在一些实施例中，该一个或多个另外的接头标记序列连接该一个或多个另外的随机序列。在一些实施例中，该细胞标记的随机序列是4-12个核苷酸长。

[0920] 在一个方面中，本披露提供了一种组合物，该组合物包括：固体支持物，其中所述固体支持物包括多个寡核苷酸，其中所述多个寡核苷酸中的至少两个包括：细胞标记、分子标记；和靶结合区；以及多个靶核酸，其中所述多个寡核苷酸中的所述至少两个的所述细胞标记是相同的，并且其中所述多个寡核苷酸中的所述至少两个的所述分子标记是不同的。在一些实施例中，靶结合区包括适配为杂交到所述多个靶核酸中的至少一个上的序列。在一些实施例中，靶结合区包括选自下组的序列，该组由以下各项组成：随机多聚体，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；基因特异性引物；和寡聚dT；或其任何组合。在一些实施例中，该多个寡核苷酸包括从10,000至10亿个寡核苷酸。在一些实施例中，该多个寡核苷酸包括大于所述多个靶核酸中的靶核酸的数目的寡核苷酸数目。在一些实施例中，该多个靶核酸包括相同靶核酸的多个拷贝。在一些实施例中，该多个靶核酸包括不同靶核酸的多个拷贝。在一些实施例中，该多个靶核酸结合至所述多个寡核苷酸。在一些实施例中，寡核苷酸进一步包括样品标记。在一些实施例中，该多个寡核苷酸包括通用标记。在一些实施例中，该通用标记包括针对测序引物的结合位点。在一些实施例中，该多个寡核苷酸包括接头。在一些实施例中，

该接头包括官能团。在一些实施例中，该接头位于所述寡核苷酸的5’。在一些实施例中，该官能团包括氨基。在一些实施例中，接头选自下组，该组由以下各项组成：C6、生物素、链亲和素、伯胺、醛、和酮，或它们的任何组合。在一些实施例中，该固体支持物由聚苯乙烯构成。在一些实施例中，该固体支持物是磁性的。在一些实施例中，该固体支持物选自下组，该组由以下各项组成：PDMS固体支持物、玻璃固体支持物、聚丙烯固体支持物、琼脂糖固体支持物、明胶固体支持物、磁性固体支持物、和普朗尼克(pluronic)固体支持物，或它们的任何组合。在一些实施例中，该固体支持物包括珠。在一些实施例中，该固体支持物包括约20微米的直径。在一些实施例中，该固体支持物包括从约5微米至约40微米的直径。在一些实施例中，该固体支持物包括官能团。在一些实施例中，该官能团包括羧基。在一些实施例中，该官能团选自下组，该组由以下各项组成：C6、生物素、链亲和素、伯胺、醛、和酮，或它们的任何组合。在一些实施例中，该细胞标记包括多个细胞标记。在一些实施例中，该多个细胞标记散布有多个接头标记序列。

[0921] 在一个方面中，本披露提供了一种试剂盒，该试剂盒包括：第一固体支持物，其中所述第一固体支持物包括第一多个寡核苷酸，其中所述第一多个寡核苷酸包括相同第一细胞标记，第二固体支持物，其中所述第二固体支持物包括第二多个寡核苷酸，其中所述第二多个寡核苷酸包括相同第二细胞标记；使用说明书，其中所述第一细胞标记和所述第二细胞标记是不同的。在一些实施例中，来自所述第一多个寡核苷酸和所述第二多个寡核苷酸的寡核苷酸包括分子标记。在一些实施例中，所述寡核苷酸的分子标记是不同的。在一些实施例中，所述寡核苷酸的分子标记是相同的。在一些实施例中，一些所述寡核苷酸的分子标记是不同的并且一些是相同的。在一些实施例中，来自所述第一多个寡核苷酸和所述第二多个寡核苷酸的寡核苷酸包括靶结合区。在一些实施例中，试剂盒进一步包括：微孔阵列。在一些实施例中，试剂盒进一步包括：缓冲液。在一些实施例中，缓冲液选自下组，该组由以下各项组成：复水缓冲液、稀释缓冲液、和稳定缓冲液，或它们的任何组合。

[0922] 在一个方面中，本披露提供了用于确定靶核酸的量的方法，该方法包括：使样品与固体支持物接触，其中所述固体支持物包括多个寡核苷酸，其中所述多个寡核苷酸中的至少两个包括细胞标记和分子标记，其中所述多个寡核苷酸中的所述至少两个的所述细胞标记是相同的，并且其中所述多个寡核苷酸中的所述至少两个的所述分子标记是不同的；并且将来自所述样品的所述靶核酸与所述多个寡核苷酸的寡核苷酸杂交。在一些实施例中，样品包括细胞。在一些实施例中，在所述杂交之前裂解样品。在一些实施例中，杂交包括将相同靶核酸的多个拷贝与所述多个寡核苷酸杂交。在一些实施例中，该方法进一步包括：扩增所述靶核酸。在一些实施例中，扩增包括逆转录所述靶核酸。在一些实施例中，扩增包括使用选自下组的方法进行扩增，该组由以下各项组成：PCR、定量PCR、实时PCR、和数字PCR、或它们的任何组合。在一些实施例中，在所述固体支持物上直接进行扩增。在一些实施例中，在从所述固体支持物转录的模板上进行扩增。在一些实施例中，该方法进一步包括：将所述靶核酸测序。在一些实施例中，测序包括将所述靶核酸和所述分子标记测序。在一些实施例中，该方法进一步包括：确定所述靶核酸的量。在一些实施例中，该确定包括量化所述靶核酸的水平。在一些实施例中，该确定包括对所述靶核酸的经测序的分子标记的数目进行计数。在一些实施例中，该接触发生在微孔中。在一些实施例中，该微孔是由选自下组的材料制成的，该组由以下各项组成：亲水塑料、塑料、弹性体、和水凝胶，或它们的任何组合。

在一些实施例中，该微孔包括琼脂糖。在一些实施例中，该微孔是微孔阵列的一个微孔。在一些实施例中，该微孔阵列包括至少90个微孔。在一些实施例中，该微孔阵列包括至少150,000个微孔。在一些实施例中，该微孔包括至少一个固体支持物/孔。在一些实施例中，该微孔包括至多两个固体支持物/孔。在一些实施例中，该微孔具有容纳至多两个所述固体支持物的尺寸。在一些实施例中，该微孔具有容纳至多一个所述固体支持物的尺寸。在一些实施例中，该微孔是至少25微米深。在一些实施例中，该微孔直径是至少25微米。

[0923] 在一个方面中，本披露提供了用来减小靶核酸的扩增偏倚的方法，该方法包括：使样品与固体支持物接触，其中所述固体支持物包括多个寡核苷酸，其中所述多个寡核苷酸中的至少两个包括细胞标记和分子标记，其中所述多个寡核苷酸中的所述至少两个的所述细胞标记是相同的，并且其中所述多个寡核苷酸中的所述至少两个的所述分子标记是不同的；并且将来自所述样品的靶核酸与所述多个寡核苷酸杂交；扩增所述靶核酸；将所述靶核酸测序，其中所述测序对所述靶核酸和所述靶核酸所结合的所述寡核苷酸的所述分子标记进行测序；并且确定所述靶核酸的量。在一些实施例中，杂交包括将相同靶核酸的多个拷贝与所述多个寡核苷酸杂交。在一些实施例中，该确定包括对相同靶核酸的经测序的分子标记的数目进行计数。在一些实施例中，该计数将对所述相同靶核酸的拷贝的数目计数。在一些实施例中，样品包括细胞。在一些实施例中，扩增包括逆转录所述靶核酸。在一些实施例中，扩增包括使用选自下组的方法进行扩增，该组由以下各项组成：PCR、定量PCR、实时PCR、和数字PCR、或它们的任何组合。在一些实施例中，在所述固体支持物上直接进行扩增。在一些实施例中，在从所述固体支持物转录的模板上进行扩增。

[0924] 在一个方面中，本披露提供了一种组合物，该组合物包括：微孔；池；以及固体支持物，其中所述固体支持物包括多个寡核苷酸，其中所述多个寡核苷酸中的至少两个包括细胞标记和分子标记，其中所述多个寡核苷酸中的所述至少两个的所述细胞标记是相同的，并且其中所述多个寡核苷酸中的所述至少两个的所述分子标记是不同的。在一些实施例中，所述多个寡核苷酸中的至少两个进一步包括样品标记。在一些实施例中，所述多个寡核苷酸中的至少两个进一步包括靶结合区。在一些实施例中，靶结合区包括选自下组的序列，该组由以下各项组成：随机多聚体，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；基因特异性引物；和寡聚dT；或其任何组合。在一些实施例中，该多个寡核苷酸包括通用标记。在一些实施例中，该通用标记包括针对测序引物的结合位点。在一些实施例中，该固体支持物由聚苯乙烯构成。在一些实施例中，该固体支持物是磁性的。在一些实施例中，该固体支持物选自下组，该组由以下各项组成：PDMS固体支持物、玻璃固体支持物、聚丙烯固体支持物、琼脂糖固体支持物、明胶固体支持物、磁性固体支持物、和普朗尼克(pluronic)固体支持物，或它们的任何组合。在一些实施例中，该固体支持物包括珠。在一些实施例中，该固体支持物具有约20微米的直径。在一些实施例中，该固体支持物具有从约5微米至约40微米的直径。在一些实施例中，该细胞标记包括多个细胞标记。在一些实施例中，该多个细胞标记散布有多个接头序列。在一些实施例中，该微孔是由选自下组的材料制成的，该组由以下各项组成：亲水塑料、塑料、弹性体、和水凝胶，或它们的任何组合。在一些实施例中，该微孔包括琼脂糖。在一些实施例中，该微孔是微孔阵列的微孔。在一些实施例中，该微孔包括至少一个固体支持物/孔。在一些实施例中，该微孔包括至多两个固体支持物/孔。在一些实施例中，该微孔具有容纳至少

一个所述固体支持物和至少一个所述细胞的尺寸。在一些实施例中，微孔具有容纳至多一个所述固体支持物和至少一个所述细胞的尺寸。在一些实施例中，该微孔是至少25微米深。在一些实施例中，该微孔直径是至少25微米。在一些实施例中，该微孔是平的。

[0925] 在一个方面中，本披露提供了一种装置，该装置包括：包括第一微孔阵列的第一基质；其中所述第一微孔阵列包括配置为进行多重的单细胞随机标记和分子索引测定的、第一预定空间安排中的多个第一微孔。

[0926] 在一些实施例中，该装置包括第一基质，该第一基质包括至少一个第二微孔阵列，其中所述至少一个第二微孔阵列包括在至少一个第二预定空间安排中的多个至少一个第二微孔。在一些实施例中，这些第一微孔和该至少一个第二微孔是相同的。在一些实施例中，这些第一微孔和该至少一个第二微孔是不同的。在一些实施例中，该第一预定空间安排和该至少一个第二预定空间安排是相同的。在一些实施例中，该第一预定空间安排和该至少一个第二预定空间安排是不同的。在一些实施例中，该预定空间安排包括一维的或二维的阵列模式。在一些实施例中，二维阵列模式包括方形网格、矩形网格、或六角网格。在一些实施例中，这些微孔包括圆柱几何形状、圆锥几何形状、半球几何形状、矩形几何形状、多面体几何形状，或它们的组合。在一些实施例中，这些微孔的直径是在约5微米和约50微米之间。在一些实施例中，这些微孔的深度是在约10微米和约60微米之间。在一些实施例中，两个相邻微孔之间的中心到中心的间距是在约15微米和约75微米之间。在一个实施例中，在第一或至少一个第二微孔阵列中的微孔的总数是在约96和约5,000,000之间。在一些实施例中，该第一基质包括硅、熔融二氧化硅、玻璃、聚合物、金属，或它们的组合。在一些实施例中，该第一基质包括琼脂糖或水凝胶。在一些实施例中，微孔阵列进一步包括至少一个表面特征，其中所述表面特征围绕一个或多个单独微孔或跨过单独微孔之间的空间，并且其中所述表面特征是半球形的、脊形的、或尖形的。

[0927] 在一个方面中，本披露提供了一种装置，该装置包括：包括至少一个第一微孔阵列的第一基质；以及包括顶板、底板、和垫片的机械夹具；其中当第一基质和机械夹具处于组装形式时，第一基质位于垫片和底板之间，垫片与第一基质形成防漏密封，并且顶板和垫片形成涵盖所述至少一个第一微孔阵列的至少一个第一室，这样使得细胞样品和基于珠的寡核苷酸标记可以分散进所述至少第一室，以进行多重的单细胞随机标记和分子索引测定。

[0928] 在一些实施例中，该至少一个第一微孔阵列是任何在此描述的阵列。在一些实施例中，垫片是从聚二甲基硅氧烷(PDMS)或类似弹性体材料制造的。在一些实施例中，顶板和底板是从铝、阳极氧化铝、不锈钢、特氟龙、聚甲基丙烯酸甲酯、聚碳酸酯、或类似的硬聚合物材料制造的。

[0929] 在一个方面中，本披露提供了一种装置，该装置包括：进一步包括至少一个微孔阵列的至少一个基质；和流动池；其中出于递送流体至所述微孔阵列的目的，流动池围绕或被附接至所述至少一个基质，并且包括至少一个入口和至少一个出口；并且其中该装置被配置为进行多重的单细胞随机标记和分子索引测定。

[0930] 在一些实施例中，所述至少一个基质包括至少一个如在此描述的微孔阵列。在一些实施例中，流动池进一步包括与多个微孔阵列配合的多个微阵列室，这样使得可以平行处理一个或多个样品。在一些实施例中，流动池进一步包括具孔阻隔物或流扩散器，以提供细胞和珠到至少一个微孔阵列的更均匀递送。在一些实施例中，流动池进一步包括将每一

包含微孔阵列的室分为分段的分隔物，这些分段集体地覆盖同一总阵列区域并且提供细胞和珠到至少一个微孔阵列的更均匀递送。在一些实施例中，并入该装置的流体通道的宽度是在约50微米和20mm之间。在一些实施例中，并入该装置的流体通道的深度是在约50微米和约2mm之间。在一些实施例中，流动池是从选自下组的材料制造的，该组由以下各项组成：硅、熔融二氧化硅、玻璃、聚二甲基硅氧烷(PDMS;弹性体)、聚甲基丙烯酸甲酯(PMMA)、聚碳酸酯(PC)、聚丙烯(PP)、聚乙烯(PE)、高密度聚乙烯(HDPE)、聚酰亚胺、环烯烃聚合物(COP)、环烯烃共聚物(COC)、聚对苯二甲酸乙二酯(PET)、环氧树脂、金属，或这些材料的组合。在一些实施例中，该装置包括配置为进行自动化多重的单细胞随机标记和分子索引测定的仪器系统的固定部件。在一些实施例中，该装置包括配置为进行自动化多重的单细胞随机标记和分子索引测定的仪器系统的可移除部件。

[0931] 在一个方面中，本披露提供了柱体，该柱体包括：进一步包括至少一个第一微孔阵列的至少一个第一基质；至少一个第一流动池或微孔阵列室；样品的或试剂的一个或多个储器；并且其中出于递送流体至所述至少一个第一微孔阵列的目的，该柱体进一步包括至少一个入口和至少一个出口；并且其中该柱体被配置为进行多重的单细胞随机标记和分子索引测定。

[0932] 在一些实施例中，所述至少一个第一基质包括如在此描述的至少一个第一微孔阵列。在一些实施例中，该柱体包括多个微孔阵列并且被配置为平行处理一个或多个样品。在一些实施例中，该至少一个第一流动池或微孔阵列室进一步包括具孔阻隔物或流量扩散器，以提供细胞和珠到至少一个第一微孔阵列的更均匀递送。在一些实施例中，至少一个第一流动池或微孔阵列室进一步包括将该至少一个第一流动池或微孔阵列室分为分段的分隔物，这些分段集体地覆盖同一总阵列区域并且提供细胞和珠到微孔阵列的更均匀递送。在一些实施例中，并入该柱体的流体通道的宽度是在约50微米和200微米之间。在一些实施例中，并入该柱体的流体通道的宽度是在约200微米和2mm之间。在一些实施例中，并入该柱体的流体通道的宽度是在约2mm和10mm之间。在一些实施例中，并入该柱体的流体通道的宽度是在约10mm和20mm之间。在一些实施例中，并入该柱体的流体通道的深度是在约50微米和约2mm之间。在一些实施例中，并入该柱体的流体通道的深度是在约500微米和1mm之间。在一些实施例中，并入该柱体的流体通道的深度是在约1mm和约2mm之间。在一些实施例中，该一个或多个流动池或微孔阵列室是从选自下组的材料制造的，该组由以下各项组成：硅、熔融二氧化硅、玻璃、聚二甲基硅氧烷(PDMS;弹性体)、聚甲基丙烯酸甲酯(PMMA)、聚碳酸酯(PC)、聚丙烯(PP)、聚乙烯(PE)、高密度聚乙烯(HDPE)、聚酰亚胺、环烯烃聚合物(COP)、环烯烃共聚物(COC)、聚对苯二甲酸乙二酯(PET)、环氧树脂、金属，或这些材料的组合。在一些实施例中，该装置包括配置为进行自动化多重的单细胞随机标记和分子索引测定的仪器系统的可移除的、可消耗的部件。在一些实施例中，该柱体进一步包括旁路通道或其他设计特征，以便提供分配或注射进该柱体的细胞样品或珠悬浮液的自身计量。在一些实施例中，该柱体进一步包括集成微型泵，以便控制穿过该装置的流体流动。在一些实施例中，该柱体进一步包括集成微型阀，以便分隔预加载的试剂并且以便控制穿过该装置的流体流动。在一些实施例中，该柱体进一步包括通气孔，以便提供滞留空气的逃逸路径。在一些实施例中，该柱体进一步包括设计要素，以便产生物理的或化学的屏障，这些屏障有效增加路径长度并且防止或最小化微孔之间的分子扩散，其中这些设计要素选自下组，该组由以下各项组

成:蛇形通道模式,用于细胞和珠到至少一个第一微孔阵列的递送,被压入与至少一个第一微孔阵列的表面接触的可伸缩压板或可变形膜,或不互混的疏水流体从该柱体内的储器的释放。在一些实施例中,该柱体进一步包括集成的温度控制部件或集成的热界面,以便提供与外部仪器系统的良好热接触。在一些实施例中,该柱体进一步包括光学的界面或窗口,用于至少一个第一微孔阵列的光学成像。在一些实施例中,该柱体进一步包括被配置为与独立式PCR热循环仪和/或测序仪配合的一个或多个可移除样品收集室。在一些实施例中,该柱体自身被配置为直接与独立式PCR热循环仪和/或测序仪配合。

[0933] 在一些方面中,本披露提供了一种仪器系统,该仪器系统包括:进一步包括至少一个第一微孔阵列的至少一个第一流动池或柱体;和流动控制器;其中该流动控制器控制细胞样品、基于珠的寡核苷酸标记试剂、和其他测定试剂到至少一个第一微孔阵列的递送,并且该仪器系统被配置为进行多重的单细胞随机标记和分子索引测定。

[0934] 在一些实施例中,该至少一个第一微孔阵列如在此描述。在一些实施例中,该至少一个第一流动池是该系统的固定部件。在一些实施例中,该至少一个第一流动池是该系统的可移除的、可消耗的部件。在一些实施例中,该至少一个第一柱体是该系统的可移除的、可消耗的部件。在一些实施例中,由使用者将细胞样品和基于珠的寡核苷酸试剂直接分配或注射进该柱体。在一些实施例中,将除了细胞样品的测定试剂预加载进该柱体。在一些实施例中,该仪器系统进一步包括成像系统,以便使该至少一个微孔阵列成像。在一些实施例中,该仪器系统进一步包括细胞的或珠的分配系统,用于促进细胞和珠跨该至少一个第一微孔阵列的均匀分配,其中在所述分配系统下的机制选自下组,该组由以下各项组成:振动、振荡、漩涡、再循环流、低频搅拌、或高频搅拌。在一些实施例中,该仪器系统进一步包括细胞裂解系统,其中该系统使用高频压电换能器,用于超声处理细胞。在一些实施例中,该仪器系统进一步包括温度控制器,用于维持使用者指定的温度,或用于经两个或更多个指定时间间隔,升高在两个或更多个指定温度之间升降温度。在一些实施例中,该仪器系统进一步包括磁场控制器,用于从微孔洗脱珠。在一些实施例中,该仪器系统进一步包括被编程用来提供系统功能的用户界面和控制的计算机或处理器。在一些实施例中,该仪器系统进一步包括程序代码,用于提供实时图象分析能力。在一些实施例中,偶联实时图像分析和仪器控制功能,这样使得可以延长或重复细胞的和珠的样品加载步骤,直至实现最佳的细胞的/珠的分配。在一些实施例中,该仪器系统进一步包括集成的PCR热循环仪,用于寡核苷酸标记的扩增。在一些实施例中,该仪器系统进一步包括集成的测序仪,用于寡核苷酸文库的测序,由此提供样品至答案能力。在一些实施例中,这些细胞样品包括患者样品,并且多重的单细胞随机标记和分子索引测定的结果被用于临床诊断应用。在一些实施例中,这些细胞样品包括患者样品,并且由医疗提供者使用多重的单细胞随机标记和分子索引测定的结果,以做出明智的医疗保健治疗决策。

[0935] 在一个方面中,本发明提供了被编程用来进行一个或多个以下序列数据分析功能的存在于计算机可读媒质中的软件:确定读取数目/基因/细胞,和唯一性转录物分子数目/基因/细胞;主成分分析或其他统计分析,以预测用于确定转录物分子数目/基因/细胞的置信区间;基因序列数据与已知参照序列的比对;样品条码、细胞条码、和分子条码的解码/多路解编;以及分子标记的自动聚簇,以补偿扩增的或测序的误差;其中通过进行多重的单细胞随机标记和分子索引测定产生序列数据。

[0936] 在一个方面中，本披露提供了一种组合物，该组合物包括：固体支持物，其中该固体支持物包括多个寡核苷酸，其中该多个寡核苷酸中的至少两个包括细胞标记和分子标记，其中该多个寡核苷酸中的该至少两个的细胞标记是相同的，并且其中该多个寡核苷酸中的该至少两个的分子标记是不同的。

[0937] 在一些实施例中，该多个寡核苷酸进一步包括样品标记。在一些实施例中，该多个寡核苷酸进一步包括靶结合区。在一些实施例中，该靶结合区包括被适配为杂交至靶核酸的序列。在一些实施例中，该靶核酸包括多个靶核酸，该多个靶核酸包括有机体的转录组的转录物的至少0.01％、0.02％、0.05％、0.1％、0.2％、0.3％、0.4％、0.5％、1％、2％、3％、4％、5％、6％、7％、8％、9％、10％、11％、12％、13％、14％、15％、16％、17％、18％、19％、20％、25％、30％、35％、40％、45％、50％、60％、70％、80％、90％、或100％。在一些实施例中，该靶核酸是DNA。在一些实施例中，该靶核酸是RNA。在一些实施例中，该靶核酸是mRNA。在某些实施例中，该DNA是基因组DNA。在某些实施例中，剪切基因组DNA。在一些实施例中，剪切的基因组DNA包括有机体的基因组的基因的至少0.01％、0.02％、0.05％、0.1％、0.2％、0.3％、0.4％、0.5％、1％、2％、3％、4％、5％、6％、7％、8％、9％、10％、11％、12％、13％、14％、15％、16％、17％、18％、19％、20％、25％、30％、35％、40％、45％、50％、60％、70％、80％、90％、或100％。在一些实施例中，靶结合区包括选自下组的序列，该组由以下各项组成：随机多聚体，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；基因特异性引物；和寡聚dT；或其任何组合。在一些实施例中，该多个寡核苷酸包括通用标记。在一些实施例中，该通用标记包括针对测序引物的结合位点。在一些实施例中，该多个寡核苷酸包括接头。在一些实施例中，该接头包括官能团。在一些实施例中，该接头位于寡核苷酸的5’。在一些实施例中，接头选自下组，该组由以下各项组成：C6、生物素、链亲和素、伯胺、醛、和酮，或它们的任何组合。在一些实施例中，该固体支持物由聚苯乙烯构成。在一些实施例中，该固体支持物是磁性的。在一些实施例中，该固体支持物选自下组，该组由以下各项组成：PDMS固体支持物、玻璃固体支持物、聚丙烯固体支持物、琼脂糖固体支持物、明胶固体支持物、磁性固体支持物、和普朗尼克(pluronic)固体支持物，或它们的任何组合。在一些实施例中，该固体支持物包括珠。在一些实施例中，该固体支持物包括约20微米的直径。在一些实施例中，该固体支持物包括从约5微米至约40微米的直径。在一些实施例中，该固体支持物包括官能团。在一些实施例中，该官能团选自下组，该组由以下各项组成：C6、生物素、链亲和素、伯胺、醛、和酮，或它们的任何组合。在一些实施例中，该细胞标记包括多个细胞标记。在一些实施例中，该多个细胞标记散布有多个接头标记序列。在一些实施例中，该多个寡核苷酸包括从10,000至10亿个寡核苷酸。在一些实施例中，该多个寡核苷酸包括从10,000至10亿个靶结合区。在一些实施例中，该多个寡核苷酸包括从10,000至10亿个不同靶结合区。在一些实施例中，该多个寡核苷酸包括从10,000至10亿个相同靶结合区。在一些实施例中，不同靶结合区可以杂交至有机体的转录组的转录物的至少0.01％、0.02％、0.05％、0.1％、0.2％、0.3％、0.4％、0.5％、1％、2％、3％、4％、5％、6％、7％、8％、9％、10％、11％、12％、13％、14％、15％、16％、17％、18％、19％、20％、25％、30％、35％、40％、45％、50％、60％、70％、80％、90％、或100％。在一些实施例中，不同靶结合区可以杂交至有机体的转录组的转录物的至少0.01％、0.02％、0.05％、0.1％、0.2％、0.3％、0.4％、0.5％、1％、2％、3％、4％、5％、6％、7％、8％、9％、

10％、11％、12％、13％、14％、15％、16％、17％、18％、19％、20％、25％、30％、35％、40％、45％、50％、60％、70％、80％、90％、或100％。

[0938] 在一个方面中，本披露提供了一种组合物，该组合物包括：固体支持物，其中该固体支持物包括多个寡核苷酸，其中该多个寡核苷酸中的至少两个包括细胞标记和分子标记，其中该多个寡核苷酸中的该至少两个的细胞标记是相同的，并且其中该多个寡核苷酸中的该至少两个的分子标记是不同的。

[0939] 在一些实施例中，该多个寡核苷酸进一步包括样品标记。在一些实施例中，该多个寡核苷酸进一步包括靶结合区。在一些实施例中，该靶结合区包括被适配为杂交至靶核酸的序列。在一些实施例中，该靶核酸包括多个靶核酸，该多个靶核酸包括有机体的转录组的转录物的至少0.01％、0.02％、0.05％、0.1％、0.2％、0.3％、0.4％、0.5％、1％、2％、3％、4％、5％、6％、7％、8％、9％、10％、11％、12％、13％、14％、15％、16％、17％、18％、19％、20％、25％、30％、35％、40％、45％、50％、60％、70％、80％、90％、或100％。在一些实施例中，该靶核酸包括剪切的基因组DNA，其中该剪切的基因组DNA包括有机体的基因组的基因的至少0.01％、0.02％、0.05％、0.1％、0.2％、0.3％、0.4％、0.5％、1％、2％、3％、4％、5％、6％、7％、8％、9％、10％、11％、12％、13％、14％、15％、16％、17％、18％、19％、20％、25％、30％、35％、40％、45％、50％、60％、70％、80％、90％、或100％。在一些实施例中，靶结合区包括寡聚dT。在一些实施例中，该多个寡核苷酸中的该至少两个包括第一寡核苷酸和第二寡核苷酸，其中该第一寡核苷酸包括第一细胞标记和第一分子标记，其中该第一细胞标记包括第一随机序列、第二随机序列、和第一接头标记序列，其中该第一接头标记序列连接该第一随机序列和该第二随机序列；并且该第一分子标记包括随机序列；并且该第二寡核苷酸包括第二细胞标记和第二分子标记，其中该第二细胞标记包括第三随机序列、第四随机序列、和第二接头标记序列，其中该第二接头标记序列连接该第三随机序列和该第四随机序列；并且该第二分子标记包括随机序列，并且其中该第一细胞标记和该第二细胞标记是相同的并且该第一分子标记和该第二分子标记是不同的。

[0940] 在一个方面中，本披露提供了一种试剂盒，该试剂盒包括任何在此描述的组合物和使用说明书。

[0941] 在一个方面中，本披露提供了一种方法，该方法包括：使样品与固体支持物接触，其中该固体支持物包括多个寡核苷酸，其中该多个寡核苷酸中的至少两个包括细胞标记和分子标记，其中该多个寡核苷酸中的该至少两个的细胞标记是相同的，并且其中该多个寡核苷酸中的该至少两个的分子标记是不同的；并且将来自该样品的靶核酸与该多个寡核苷酸的寡核苷酸杂交。

[0942] 在一些实施例中，该样品包括细胞。在一些实施例中，在该杂交之前裂解样品。在一些实施例中，杂交包括将相同靶核酸的多个拷贝与该多个寡核苷酸杂交。在一些实施例中，该方法进一步包括逆转录靶核酸。在一些实施例中，该方法进一步包括进行寡核苷酸扩增。在一些实施例中，扩增包括使用选自下组的方法进行扩增，该组由以下各项组成：PCR、定量PCR、实时PCR、和数字PCR、或它们的任何组合。

[0943] 在一个方面中，本披露提供了固体支持物，该固体支持物包括：第一寡核苷酸，该第一寡核苷酸包括：第一细胞标记，该第一细胞标记包括第一随机序列、第二随机序列、和第一接头标记序列，其中该第一接头标记序列连接该第一随机序列和该第二随机序列；以

及包括随机序列的第一分子标记;以及第二寡核苷酸,该第二寡核苷酸包括:第二细胞标记,该第二细胞标记包括第三随机序列、第四随机序列、和第二接头标记序列,其中该第二接头标记序列连接该第三随机序列和该第四随机序列;以及包括随机序列的第二分子标记,其中该第一细胞标记和该第二细胞标记是相同的并且该第一分子标记和该第二分子标记是不同的。在一些实施例中,该第一和第二寡核苷酸进一步包括相同样品索引区。在一些实施例中,该样品索引区包括随机序列。在一些实施例中,该样品索引区是4-12个核苷酸长。在一些实施例中,该细胞标记直接附接至该分子标记。在一些实施例中,该细胞标记和该分子标记是通过接头标记序列附接的。在一些实施例中,该细胞标记的随机序列是从4-12个核苷酸长。在一些实施例中,该细胞标记的恒定序列是至少4个核苷酸长。在一些实施例中,该细胞标记具有至少12个核苷酸的总长。在一些实施例中,该细胞标记进一步包括一个或多个另外的随机序列。在一些实施例中,该细胞标记进一步包括一个或多个另外的接头标记序列。在一些实施例中,该一个或多个另外的接头标记序列连接该一个或多个另外的随机序列。在一些实施例中,该细胞标记的随机序列是4-12个核苷酸长。

[0944] 在一个方面中,本披露提供了一种组合物,该组合物包括:固体支持物,其中该固体支持物包括多个寡核苷酸,其中该多个寡核苷酸中的至少两个包括:细胞标记、分子标记;和靶结合区;以及多个靶核酸,其中该多个寡核苷酸中的该至少两个的细胞标记是相同的,并且其中该多个寡核苷酸中的该至少两个的分子标记是不同的。

[0945] 在一些实施例中,该靶结合区包括适配为杂交到该多个靶核酸中的至少一个上的序列。在一些实施例中,靶结合区包括选自下组的序列,该组由以下各项组成:随机多聚体,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;基因特异性引物;和寡聚dT;或其任何组合。在一些实施例中,该多个寡核苷酸包括从10,000至10亿个寡核苷酸。在一些实施例中,该多个寡核苷酸包括大于该多个靶核酸的靶核酸数目的寡核苷酸数目。在一些实施例中,该多个靶核酸包括相同靶核酸的多个拷贝。在一些实施例中,该多个靶核酸包括不同靶核酸的多个拷贝。在一些实施例中,该多个靶核酸结合至该多个寡核苷酸。在一些实施例中,该寡核苷酸进一步包括样品标记。在一些实施例中,该多个寡核苷酸包括通用标记。在一些实施例中,该通用标记包括针对测序引物的结合位点。在一些实施例中,该多个寡核苷酸包括接头。在一些实施例中,该接头包括官能团。在一些实施例中,该接头位于该寡核苷酸的5'。在一些实施例中,官能团包括氨基。在一些实施例中,接头选自下组,该组由以下各项组成:C6、生物素、链亲和素、伯胺、醛、和酮,或它们的任何组合。在一些实施例中,该固体支持物由聚苯乙烯构成。在一些实施例中,固体支持物是磁性的。在一些实施例中,该固体支持物选自下组,该组由以下各项组成:PDMS固体支持物、玻璃固体支持物、聚丙烯固体支持物、琼脂糖固体支持物、明胶固体支持物、磁性固体支持物、和普朗尼克(pluronic)固体支持物,或它们的任何组合。在一些实施例中,该固体支持物包括珠。在一些实施例中,该固体支持物包括约20微米的直径。在一些实施例中,该固体支持物包括从约5微米至约40微米的直径。在一些实施例中,该固体支持物包括官能团。在一些实施例中,官能团包括羧基。在一些实施例中,该官能团选自下组,该组由以下各项组成:C6、生物素、链亲和素、伯胺、醛、和酮,或它们的任何组合。在一些实施例中,该细胞标记包括多个细胞标记。在一些实施例中,该多个细胞标记散布有多个接头标记序列。

[0946] 在一个方面中，本披露提供了一种试剂盒，该试剂盒包括：第一固体支持物，其中该第一固体支持物包括第一多个寡核苷酸，其中该第一多个寡核苷酸包括相同第一细胞标记；第二固体支持物，其中该第二固体支持物包括第二多个寡核苷酸，其中该第二多个寡核苷酸包括相同第二细胞标记；和使用说明书，其中该第一细胞标记和该第二细胞标记是不同的。

[0947] 在一些实施例中，来自该第一多个寡核苷酸和该第二多个寡核苷酸的寡核苷酸包括分子标记。在一些实施例中，这些寡核苷酸的分子标记是不同的。在一些实施例中，这些寡核苷酸的分子标记是相同的。在一些实施例中，这些寡核苷酸中的一些的分子标记是不同的并且一些是相同的。在一些实施例中，来自该第一多个寡核苷酸和该第二多个寡核苷酸的寡核苷酸包括靶结合区。在一些实施例中，该试剂盒包括微孔阵列。在一些实施例中，该试剂盒进一步包括缓冲液。在一些实施例中，缓冲液选自下组，该组由以下各项组成：复水缓冲液、稀释缓冲液、和稳定缓冲液，或它们的任何组合。

[0948] 在一个方面中，本披露提供了用于确定靶核酸的量的方法，该方法包括：使样品与固体支持物接触，其中该固体支持物包括多个寡核苷酸，其中该多个寡核苷酸中的至少两个包括细胞标记和分子标记，其中该多个寡核苷酸中的该至少两个的细胞标记是相同的，并且其中该多个寡核苷酸中的该至少两个的分子标记是不同的；并且将来自该样品的靶核酸与该多个寡核苷酸的寡核苷酸杂交。

[0949] 在一些实施例中，该样品包括细胞。在一些实施例中，在该杂交之前裂解样品。在一些实施例中，杂交包括将相同靶核酸的多个拷贝与该多个寡核苷酸杂交。在一些实施例中，该方法进一步包括扩增靶核酸。在一些实施例中，该扩增包括逆转录靶核酸。在一些实施例中，扩增包括使用选自下组的方法进行扩增，该组由以下各项组成：PCR、定量PCR、实时PCR、和数字PCR、或它们的任何组合。在一些实施例中，在固体支持物上直接进行扩增。在一些实施例中，在从固体支持物转录的模板上进行扩增。在一些实施例中，该方法进一步包括对靶核酸进行测序。在一些实施例中，该测序包括对靶核酸和分子标记进行测序。在一些实施例中，该方法进一步包括确定靶核酸的量。在一些实施例中，该确定包括定量靶核酸的水平。在一些实施例中，该确定包括对靶核酸的经测序的分子标记的数目进行计数。在一些实施例中，该接触发生在微孔中。在一些实施例中，该微孔是由选自下组的材料制成的，该组由以下各项组成：亲水塑料、塑料、弹性体、和水凝胶，或它们的任何组合。在一些实施例中，该微孔包括琼脂糖。在一些实施例中，该微孔是微孔阵列的一个微孔。在一些实施例中，该微孔阵列包括至少90个微孔。在一些实施例中，该微孔阵列包括至少150,000个微孔。在一些实施例中，该微孔包括至少一个固体支持物/孔。在一些实施例中，该微孔包括至多两个固体支持物/孔。在一些实施例中，该微孔具有容纳至多两个固体支持物的尺寸。在一些实施例中，该微孔具有容纳至多一个固体支持物的尺寸。在一些实施例中，该微孔是至少25微米深。在一些实施例中，该微孔直径是至少25微米。

[0950] 在一个方面中，本披露提供了用来减小靶核酸的扩增偏倚的方法，该方法包括：使样品与固体支持物接触，其中该固体支持物包括多个寡核苷酸，其中该多个寡核苷酸中的至少两个包括细胞标记和分子标记，其中该多个寡核苷酸中的该至少两个的细胞标记是相同的，并且其中该多个寡核苷酸中的该至少两个的分子标记是不同的；并且将来自该样品的靶核酸与该多个寡核苷酸杂交；扩增该靶核酸或其互补体。对该靶核酸或其互补体进行

测序，其中该测序对该靶核酸或其互补体和该靶核酸或其互补体所结合的寡核苷酸的分子标记进行测序。确定靶核酸的量。

[0951] 在一些实施例中，杂交包括将相同靶核酸的多个拷贝与该多个寡核苷酸杂交。在一些实施例中，该确定包括对相同靶核酸的经测序的分子标记的数目进行计数。在一些实施例中，该计数对相同靶核酸的拷贝的数目进行计数。在一些实施例中，该样品包括细胞。在一些实施例中，该扩增包括逆转录靶核酸。在一些实施例中，扩增包括使用选自下组的方法进行扩增，该组由以下各项组成：PCR、定量PCR、实时PCR、和数字PCR、或它们的任何组合。在一些实施例中，该扩增在固体支持物上直接进行。在一些实施例中，在从固体支持物转录的模板上进行扩增。

[0952] 在一个方面中，本披露提供了一种组合物，该组合物包括微孔；池；以及固体支持物，其中该固体支持物包括多个寡核苷酸，其中该多个寡核苷酸中的至少两个包括细胞标记和分子标记，其中该多个寡核苷酸中的该至少两个的细胞标记是相同的，并且其中该多个寡核苷酸中的该至少两个的分子标记是不同的。

[0953] 在一些实施例中，该多个寡核苷酸中的该至少两个进一步包括样品标记。在一些实施例中，该多个寡核苷酸中的该至少两个进一步包括靶结合区。在一些实施例中，靶结合区包括选自下组的序列，该组由以下各项组成：随机多聚体，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；基因特异性引物；和寡聚dT；或其任何组合。在一些实施例中，该多个寡核苷酸包括通用标记。在一些实施例中，该通用标记包括针对测序引物的结合位点。在一些实施例中，该固体支持物由聚苯乙烯构成。在一些实施例中，该固体支持物是磁性的。在一些实施例中，该固体支持物选自下组，该组由以下各项组成：PDMS固体支持物、玻璃固体支持物、聚丙烯固体支持物、琼脂糖固体支持物、明胶固体支持物、磁性固体支持物、和普朗尼克(pluronic)固体支持物，或它们的任何组合。在一些实施例中，该固体支持物包括珠。在一些实施例中，该固体支持物具有约20微米的直径。在一些实施例中，该固体支持物具有从约5微米至约40微米的直径。在一些实施例中，该细胞标记包括多个细胞标记。在一些实施例中，该多个细胞标记散布有多个接头序列。在一些实施例中，该微孔是由选自下组的材料制成的，该组由以下各项组成：亲水塑料、塑料、弹性体、和水凝胶，或它们的任何组合。在一些实施例中，该微孔包括琼脂糖。在一些实施例中，该微孔是微孔阵列的微孔。在一些实施例中，该微孔包括至少一个固体支持物/孔。在一些实施例中，该微孔包括至多两个固体支持物/孔。在一些实施例中，该微孔具有容纳至少一个固体支持物和至少一个细胞的尺寸。在一些实施例中，微孔具有容纳至多一个固体支持物和至少一个细胞的尺寸。在一些实施例中，该微孔是至少25微米深。在一些实施例中，该微孔直径是至少25微米。在一些实施例中，该微孔是平的。

[0954] 在一个方面中，本披露提供了一种装置，该装置包括多个微孔，其中该多个微孔包括至少两个微孔；并且其中该多个微孔的每一微孔具有范围从约1,000μm^{3}至约120,000μm^{3}的体积。在一些实施例中，该多个微孔中的每个微孔的体积是约20,000μm^3。在一些实施例中，该多个微孔包括从约1,000至约5,000,000个微孔。在一些实施例中，该多个微孔包括约100,000至约200,000个微孔。在一些实施例中，这些微孔包括在单层的材料中。在一些实施例中，至少约10％的微孔进一步包括细胞。在一些实施例中，至少约10％的微孔进一步包括固体支持物，该固体支持物包括多个寡核苷酸，其中该多个寡核苷酸中的至少两个包括细

胞标记和分子标记，其中该多个寡核苷酸中的该至少两个的细胞标记是相同的，并且其中该多个寡核苷酸中的该至少两个的分子标记是不同的。在一些实施例中，该固体支持物是磁化的。

[0955] 在一个方面中，本披露提供了一种设备，该设备包括任何在此描述的装置、和液体处理器。

[0956] 在一些实施例中，该液体处理器在约1秒内递送液体至该多个微孔。在一些实施例中，该设备从单输入口递送液体至该多个微孔。在一些实施例中，该设备进一步包括磁体。在一些实施例中，该设备进一步包括以下项中的至少一种：入口、出口、泵、阀、通气孔、储器、样品收集室、温度控制设备或其任何组合。在一些实施例中，该设备包括样品收集室，其中样品收集室是从该设备可移除的。在一些实施例中，该设备进一步包括光学成像仪。在一些实施例中，光学成像仪产生用于控制液体处理器的输出信号。在一些实施例中，该设备进一步包括被配置成进行寡核苷酸的聚合酶链式反应(PCR)扩增的热循环机构。

[0957] 在一个方面中，本披露提供了用于产生临床诊断测试结果的方法，该方法包括用任何在此描述的装置或设备产生该临床诊断测试结果。在一些实施例中，经由通信媒体传输临床诊断测试结果。

[0958] 在一个方面中，本披露提供了一种装置，该装置包括：进一步包括一个或多个微孔阵列的一个或多个基质；其中该微孔阵列被用于进行多重的单细胞随机标记和分子索引测定。

[0959] 在一些实施例中，这些基质的微孔阵列包括按一维的或二维的阵列模式安排的微孔。在一些实施例中，微孔的二维阵列模式选自下组，该组包括：方形网格、矩形网格、或六角网格。

[0960] 在一些实施例中，使用选自下组的孔几何形状制造微孔阵列的微孔，该组包括：圆柱、圆锥、半球、矩形、或多面体。在一些实施例中，使用包括两个或更多个选自下组的分量几何体的整体几何体制造微孔阵列的微孔，该组包括：圆柱、圆锥、半球、矩形、或多面体。在一些实施例中，微孔阵列中的微孔的直径是在约5微米和约50微米之间。在一些实施例中，微孔阵列中的微孔的深度是在约10微米和约60微米之间。在一些实施例中，微孔阵列中的微孔之间的中心到中心的间距是在约15微米和约75微米之间。在一个实施例中，每一微孔阵列中的微孔的总数是在约96和约5,000,000之间。在一些实施例中，该一个或多个基质是从选自下组的材料制造的，该组包括：硅、熔融二氧化硅、玻璃、聚合物、或金属。在一些实施例中，该一个或多个基质是从琼脂糖或水凝胶制造的。在一些实施例中，微孔阵列进一步包括围绕这些微孔或跨过微孔之间的空间的、并且选自下组的微孔之间的表面特征，该组包括半球形的、脊形的、或尖形的表面特征。

[0961] 在一个方面中，本披露提供了一种装置，该装置包括：进一步包括一个或多个微孔阵列的基质；以及包括顶板、底板、和垫片的机械夹具；其中当该基质位于垫片和底板之间时，垫片与基质形成防漏密封，并且顶板和垫片形成涵盖微孔阵列的一个或多个室，这样使得一个或多个细胞样品和基于珠的寡核苷酸标记可以分配进这些室，用于进行多重的单细胞随机标记和分子索引测定的目的。

[0962] 在一些实施例中，该基质包括一个或多个如在此描述的微孔阵列。在一些实施例中，垫片是从聚二甲基硅氧烷(PDMS)或类似弹性体材料制造的。在一些实施例中，顶板和底

板是从铝、阳极氧化铝、不锈钢、特氟龙、聚甲基丙烯酸甲酯、聚碳酸酯、或类似的硬聚合物材料制造的。

[0963] 在一个方面中，本披露提供了一种装置，该装置包括：进一步包括一个或多个微孔阵列的一个或多个基质；和一个或多个流动池；其中出于递送流体至这些微孔阵列的目的，该一个或多个流动池围绕或被附接至该一个或多个基质，并且包括至少一个入口和至少一个出口；并且其中该装置被用于进行多重的单细胞随机标记和分子索引测定。

[0964] 在一些实施例中，该一个或多个基质包括如在此描述的任何一个或多个微孔阵列。在一些实施例中，该一个或多个流动池中的每一个进一步包括与多个微孔阵列配合的多个微阵列室，这样使得可以平行处理一个或多个样品。在一些实施例中，该一个或多个流动池进一步包括具孔阻隔物或流量扩散器，以提供细胞和珠到微孔阵列的更均匀递送。在一些实施例中，该一个或多个流动池进一步包括将包含微孔阵列的室分为分段的分隔物，这些分段集体地覆盖同一总阵列区域并且提供细胞和珠到微孔阵列的更均匀递送。在一些实施例中，并入该装置的流体通道的宽度是在约50微米和20mm之间。在一些实施例中，并入该装置的流体通道的深度是在约50微米和约2mm之间。在一些实施例中，该一个或多个流动池是从选自下组的材料制造的，该组由以下各项组成：硅、熔融二氧化硅、玻璃、聚二甲基硅氧烷(PDMS；弹性体)、聚甲基丙烯酸甲酯(PMMA)、聚碳酸酯(PC)、聚丙烯(PP)、聚乙烯(PE)、高密度聚乙烯(HDPE)、聚酰亚胺、环烯烃聚合物(COP)、环烯烃共聚物(COC)、聚对苯二甲酸乙二酯(PET)、环氧树脂、金属，或这些材料的组合。在一些实施例中，该装置包括用于进行自动化多重的单细胞随机标记和分子索引测定的仪器系统的固定部件。在一些实施例中，该装置包括用于进行自动化多重的单细胞随机标记和分子索引测定的仪器系统的可移除部件。

[0965] 在一个方面中，本披露提供了一种柱体，该柱体包括：进一步包括一个或多个微孔阵列的一个或多个基质；一个或多个流动池或微孔阵列室；样品的或试剂的一个或多个储器；并且其中出于递送流体至微孔阵列的目的，该柱体进一步包括至少一个入口和至少一个出口；并且其中该柱体被用于进行多重的单细胞随机标记和分子索引测定。

[0966] 在一些实施例中，该一个或多个基质包括如在此描述的任何一个或多个微孔阵列。在一些实施例中，该一个或多个流动池或微孔阵列室与多个微孔阵列配合，这样使得可以平行处理一个或多个样品。在一些实施例中，该一个或多个流动池或微孔阵列室进一步包括具孔阻隔物或流量扩散器，以提供细胞和珠到微孔阵列的更均匀递送。在一些实施例中，该一个或多个流动池或微孔阵列室进一步包括将流动池或室分为分段的分隔物，这些分段集体地覆盖同一总阵列区域并且提供细胞和珠到微孔阵列的更均匀递送。在一些实施例中，并入该柱体的流体通道的宽度是在约50微米和200微米之间。在一些实施例中，并入该柱体的流体通道的宽度是在约200微米和2mm之间。在一些实施例中，并入该柱体的流体通道的宽度是在约2mm和10mm之间。在一些实施例中，并入该柱体的流体通道的宽度是在约10mm和20mm之间。在一些实施例中，并入该柱体的流体通道的深度是在约50微米和约10mm之间。在一些实施例中，并入该柱体的流体通道的深度是在约500微米和1mm之间。在一些实施例中，并入该柱体的流体通道的深度是在约1mm和约2mm之间。在一些实施例中，该一个或多个流动池或微孔阵列室是从选自下组的材料制造的，该组由以下各项组成：硅、熔融二氧化硅、玻璃、聚二甲基硅氧烷(PDMS；弹性体)、聚甲基丙烯酸甲酯(PMMA)、聚碳酸酯(PC)、聚

丙烯(PP)、聚乙烯(PE)、高密度聚乙烯(HDPE)、聚酰亚胺、环烯烃聚合物(COP)、环烯烃共聚物(COC)、聚对苯二甲酸乙二酯(PET)、环氧树脂、金属，或这些材料的组合。在一些实施例中，该装置包括用于进行自动化多重的单细胞随机标记和分子索引测定的仪器系统的可移除的、可消耗的部件。在一些实施例中，该柱体进一步包括旁路通道或其他设计特征，以便提供分配或注射进该柱体的细胞样品或珠悬浮液的自身计量。在一些实施例中，该柱体进一步包括集成微型泵，以便控制穿过该装置的流体流动。在一些实施例中，该柱体进一步包括集成微型阀，以便分隔预加载的试剂并且以便控制穿过该装置的流体流动。在一些实施例中，该柱体进一步包括通气孔，以便提供滞留空气的逃逸路径。在一些实施例中，该柱体进一步包括设计要素，以便产生物理的或化学的屏障，这些屏障有效增加路径长度并且防止或最小化微孔之间的分子扩散，其中该设计要素选自下组，该组由以下各项组成：蛇形通道的模式，用于细胞和珠到微孔阵列的递送，被压入与微孔阵列的表面接触的可伸缩压板或可变形膜，或不互混的疏水流体从该柱体内的储器的释放。在一些实施例中，该柱体进一步包括集成的温度控制部件或集成的热界面，以便提供与外部仪器系统的良好热接触。在一些实施例中，该柱体进一步包括光学的界面或窗口，用于该一个或多个微孔阵列的光学成像。在一些实施例中，该柱体进一步包括被配置为与独立式PCR热循环仪和/或测序仪配合的一个或多个可移除样品收集室。在一些实施例中，该柱体自身被配置为直接与独立式PCR热循环仪和/或测序仪配合。

[0967] 在一个方面中，本披露提供了一种仪器系统，该仪器系统包括：进一步包括一个或多个微孔阵列的一个或多个流动池或柱体；和流动控制器；其中该流动控制器控制细胞样品、基于珠的寡核苷酸标记试剂、和其他测定试剂到微孔阵列的递送，并且该仪器系统被用于进行多重的单细胞随机标记和分子索引测定。

[0968] 在一些实施例中，该一个或多个微孔阵列是任何在此描述的任何。在一些实施例中，该一个或多个流动池是该系统的固定部件。在一些实施例中，该一个或多个流动池是该系统的可移除的、可消耗的部件。在一些实施例中，该一个或多个柱体是该系统的可移除的、可消耗的部件。在一些实施例中，由使用者将细胞样品和基于珠的寡核苷酸试剂直接分配或注射进该柱体。在一些实施例中，将除了细胞样品的测定试剂预加载进该柱体。在一些实施例中，该仪器系统进一步包括成像系统，以便使该微孔阵列成像。在一些实施例中，该仪器系统进一步包括细胞的或珠的分配系统，用于促进细胞和珠跨微孔阵列的均匀分配，其中在该分配系统下的机制选自下组，该组由以下各项组成：振动、振荡、漩涡、再循环流、低频搅拌、或高频搅拌。在一些实施例中，该仪器系统进一步包括细胞裂解系统，其中该系统使用高频压电换能器，用于超声处理细胞。在一些实施例中，该仪器系统进一步包括温度控制器，用于维持使用者指定温度，或用于经两个或更多个指定时间间隔，在两个或更多个指定温度之间升降温度。在一些实施例中，该仪器系统进一步包括磁场控制器，用于从微孔洗脱珠。在一些实施例中，该仪器系统进一步包括被编程用来提供系统功能的用户界面和控制的计算机或处理器。在一些实施例中，该仪器系统进一步包括程序代码，用于提供实时图象分析能力。在一些实施例中，偶联实时图像分析和仪器控制功能，这样使得可以延长或重复细胞的和珠的样品加载步骤，直至实现最佳的细胞的/珠的分配。在一些实施例中，该仪器系统进一步包括集成的PCR热循环仪，用于寡核苷酸标记的扩增。在一些实施例中，该仪器系统进一步包括集成的测序仪，用于寡核苷酸文库的测序，由此提供样品至答案能力。

在一些实施例中，这些细胞样品包括患者样品，并且多重的单细胞随机标记和分子索引测定的结果被用于临床诊断应用。在一些实施例中，这些细胞样品包括患者样品，并且由医疗提供者使用多重的单细胞随机标记和分子索引测定的结果，以做出明智的医疗保健治疗决策。

[0969] 在一个方面中，本发明提供了被编程用来进行一个或多个以下序列数据分析的存在于计算机可读媒质中的软件：确定读取数目/基因/细胞，和唯一性转录物分子数目/基因/细胞；主成分分析或其他统计分析，以预测用于确定转录物分子数目/基因/细胞的置信区间；基因序列数据与已知参照序列的比对；样品条码、细胞条码、和分子条码的解码/多路解编；以及分子标记的自动聚簇，以补偿扩增的或测序的误差；其中通过进行多重的单细胞随机标记和分子索引测定产生序列数据。

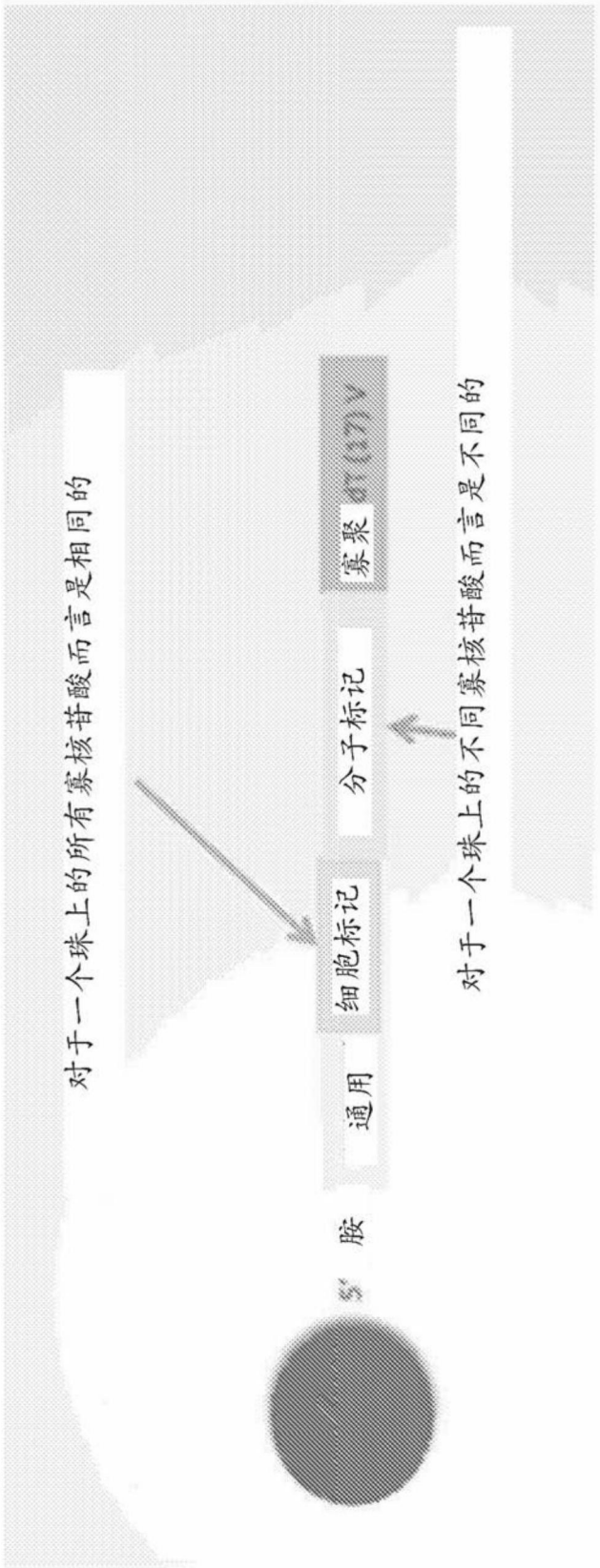

对于一个珠上的所有寡核苷酸而言是相同的
5' 胺
通用
细胞标记
分子标记
寡聚 dT(37)V
对于一个珠上的不同寡核苷酸而言是不同的

图1

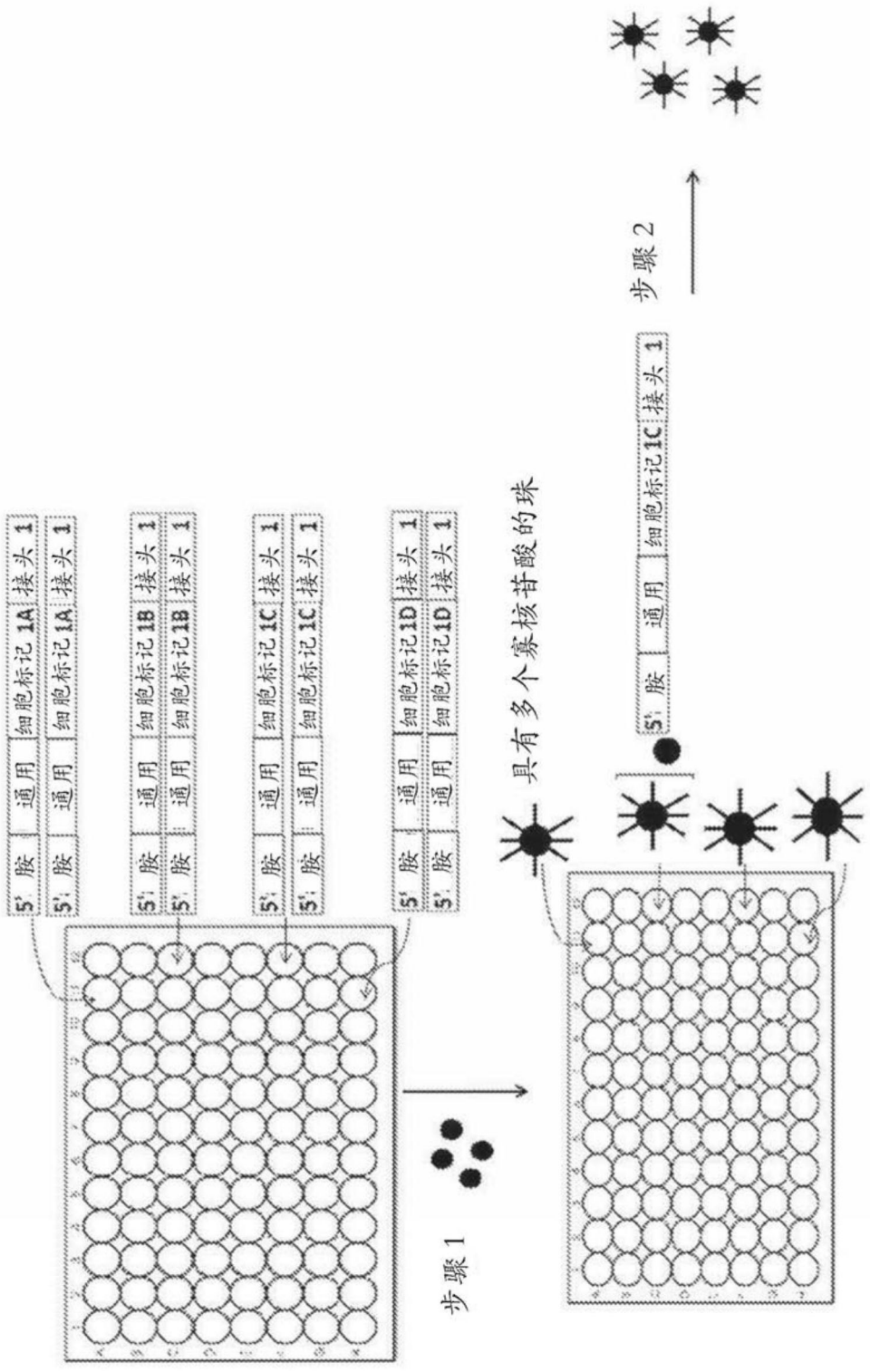


图2A

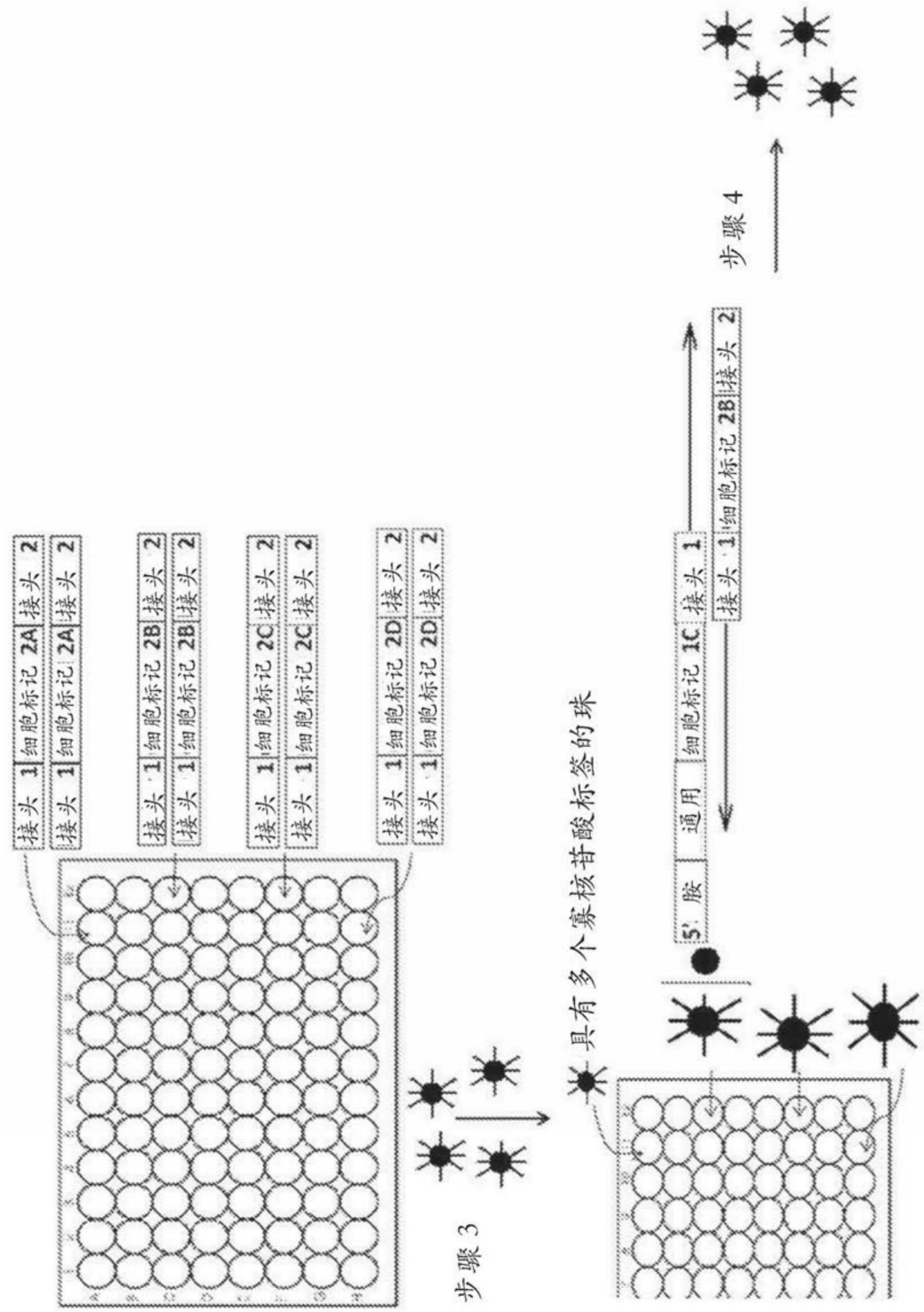

接头 1 细胞标记 2A 接头 2
接头 1 细胞标记 2A 接头 2
接头 1 细胞标记 2B 接头 2
接头 1 细胞标记 2B 接头 2
接头 1 细胞标记 2C 接头 2
接头 1 细胞标记 2C 接头 2
接头 1 细胞标记 2D 接头 2
接头 1 细胞标记 2D 接头 2
步骤 3
具有多个寡核苷酸标签的珠
5' 胺 通用 细胞标记 1C 接头 1
接头 1 细胞标记 2B 接头 2
步骤 4

图2B

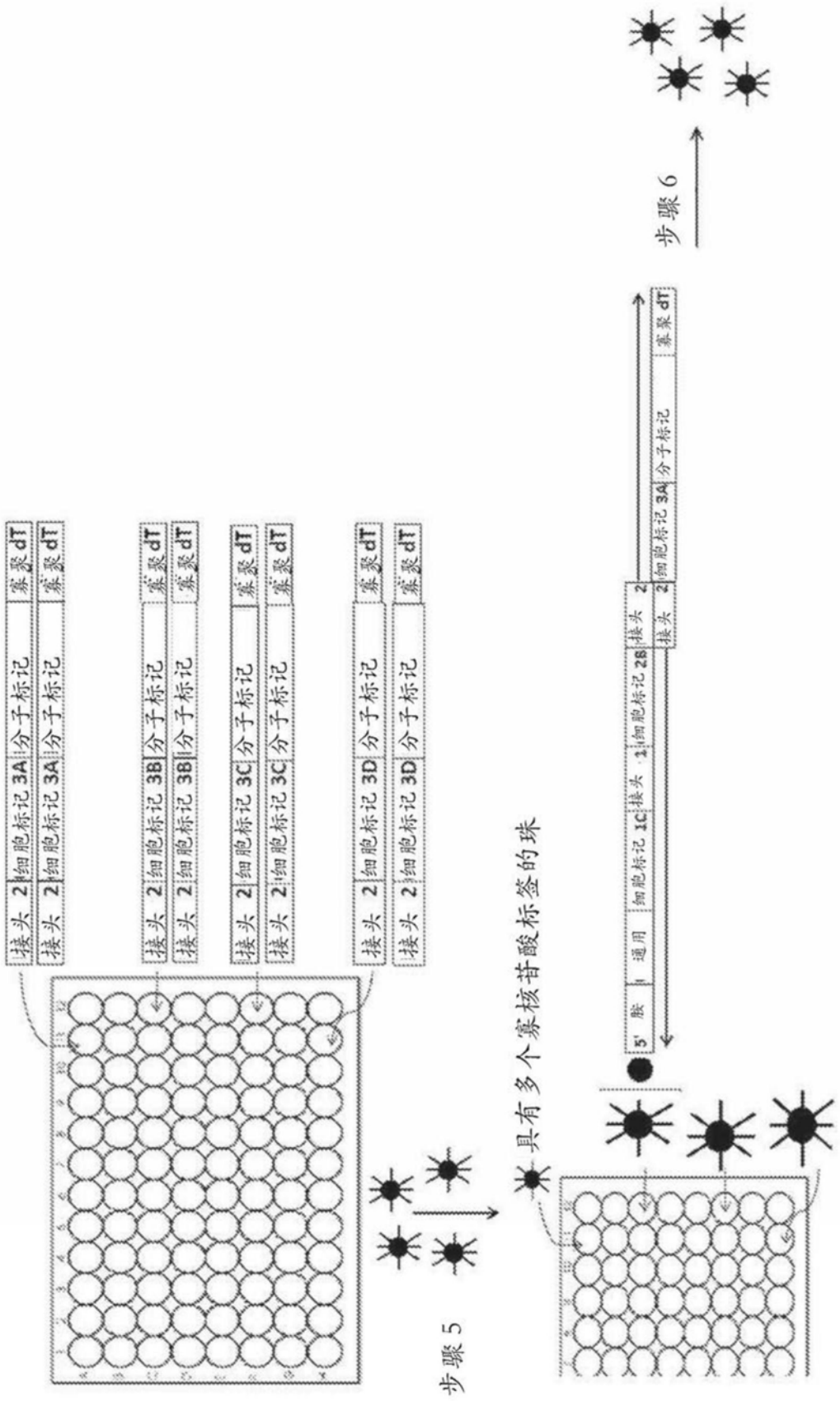

接头 2 细胞标记 3A 分子标记 寡聚dT
接头 2 细胞标记 3A 分子标记 寡聚dT
接头 2 细胞标记 3B 分子标记 寡聚dT
接头 2 细胞标记 3B 分子标记 寡聚dT
接头 2 细胞标记 3C 分子标记 寡聚dT
接头 2 细胞标记 3C 分子标记 寡聚dT
接头 2 细胞标记 3D 分子标记 寡聚dT
接头 2 细胞标记 3D 分子标记 寡聚dT
步骤 5
具有多个寡核苷酸标签的珠
5' 胺 1 通用 细胞标记 1C 接头 2 细胞标记 2B 接头 2
接头 2 细胞标记 3A 分子标记 寡聚dT
步骤 6

图2C

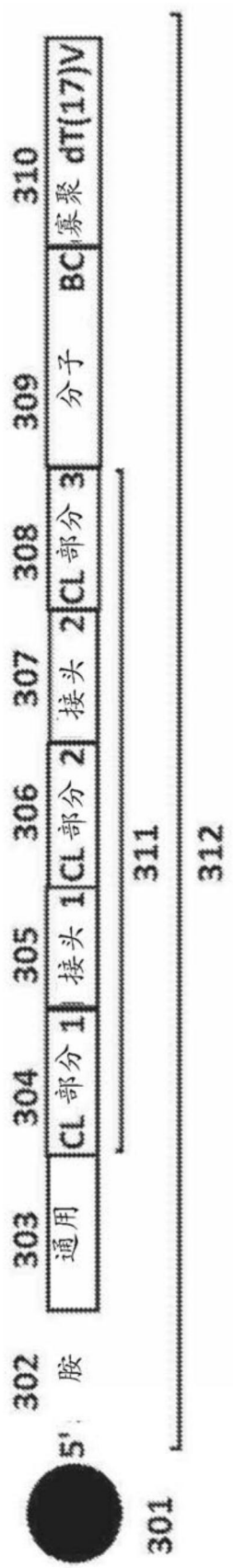

5'
胺
302
通用
303
CL 部分 1
304
接头 1
305
CL 部分 2
306
接头 2
307
CL 部分 3
308
分子
309
BC
寡聚 dT(17)V
310
311
312
301

图3

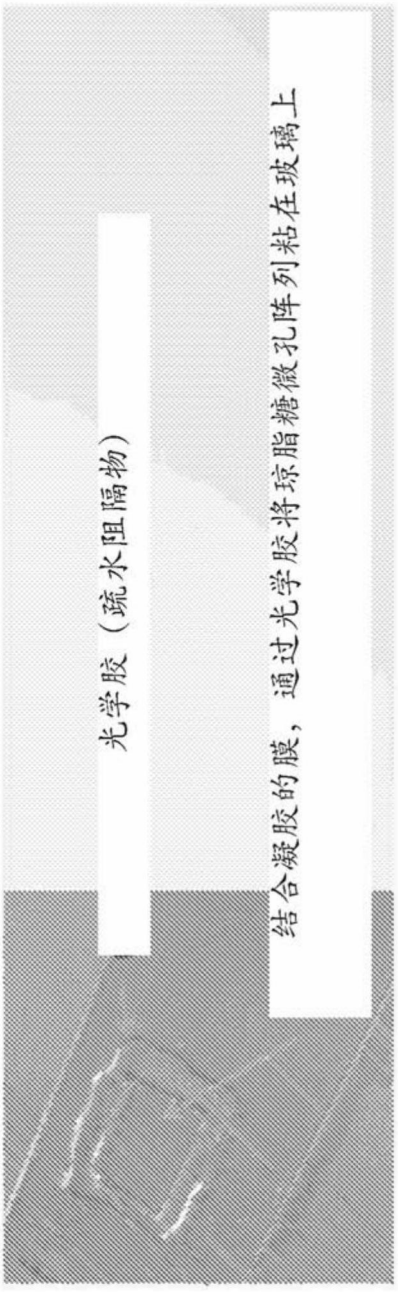


图4

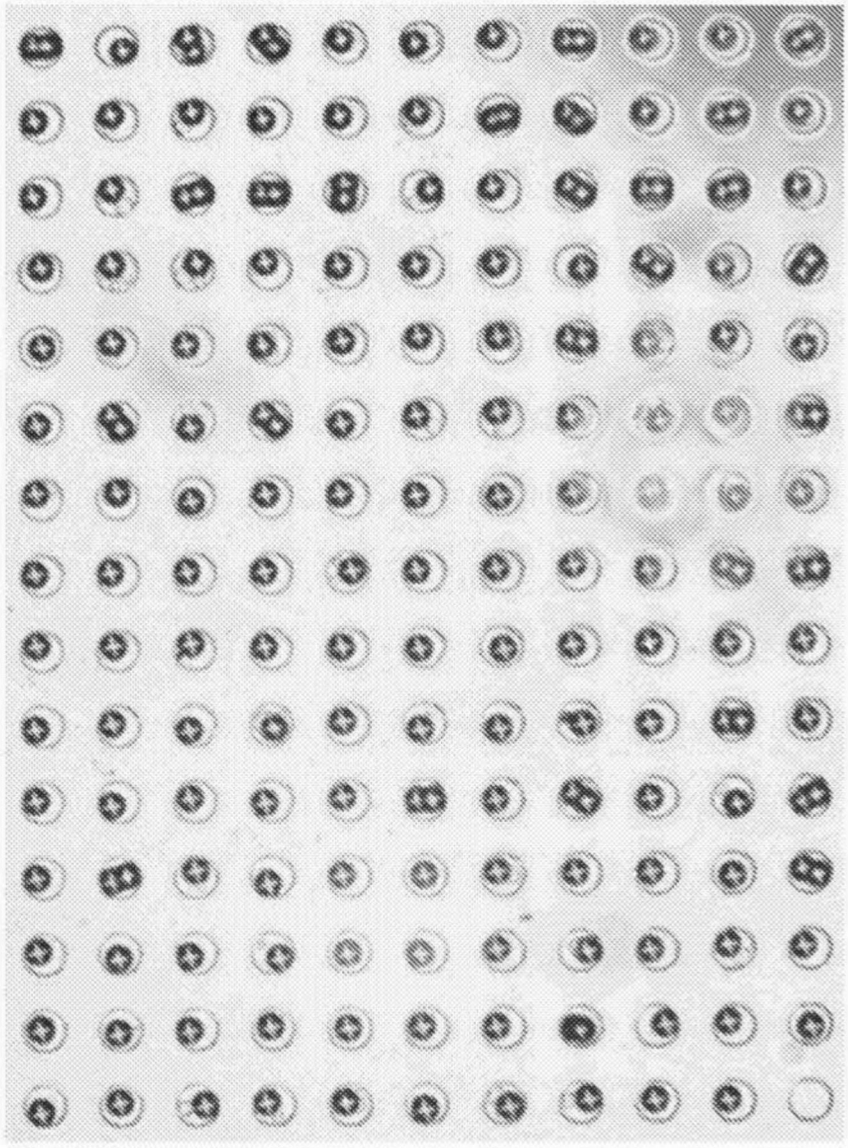


图5



图6A

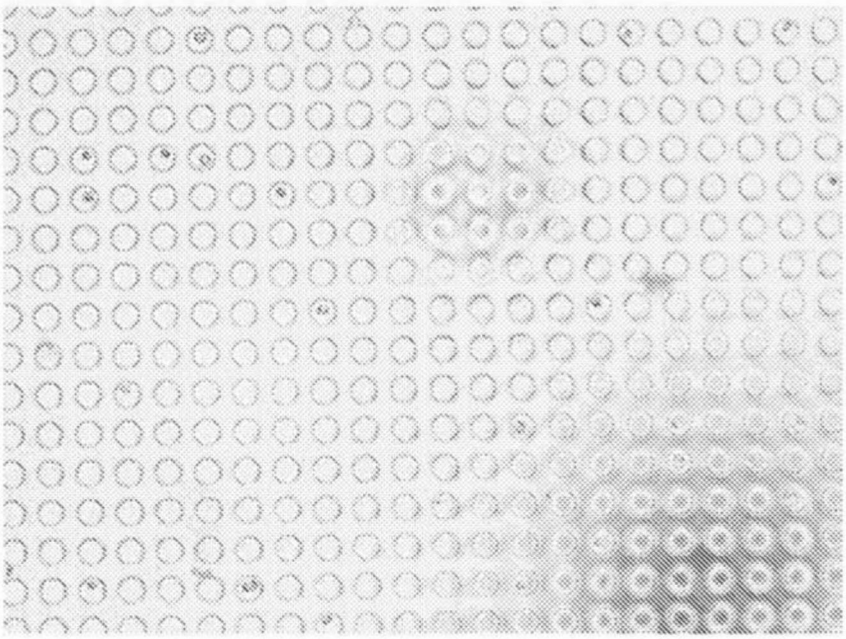


图6B

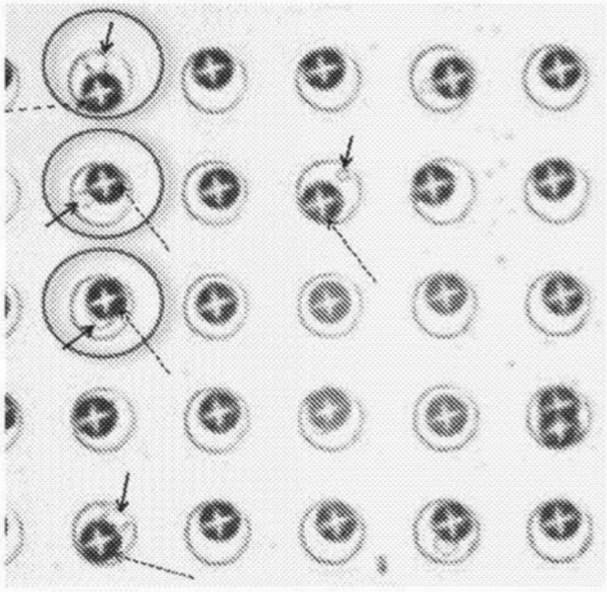


图6C

孔体积	2.12E-14	m3
珠体积	4.19E-15	m3
孔体积-珠体积	1.70E-14	m3
孔体积-珠体积	1.70E-11	L
#mRNA /细胞	2.50E+05	分子
mRNA的浓度	1.47E+16	分子/L
mRNA的浓度	2.44E-08	摩尔/L
mRNA的浓度	24.40	nM

图7

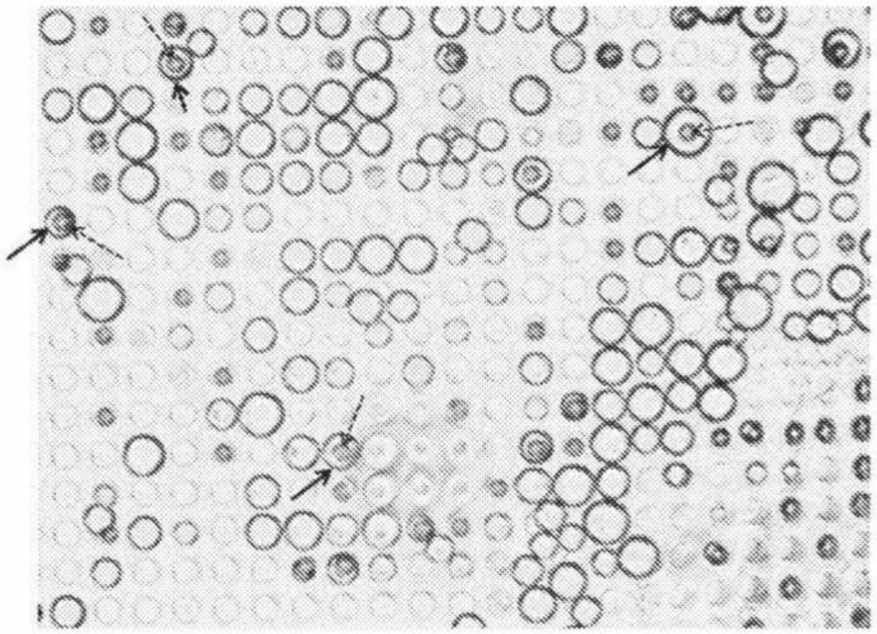


图8A

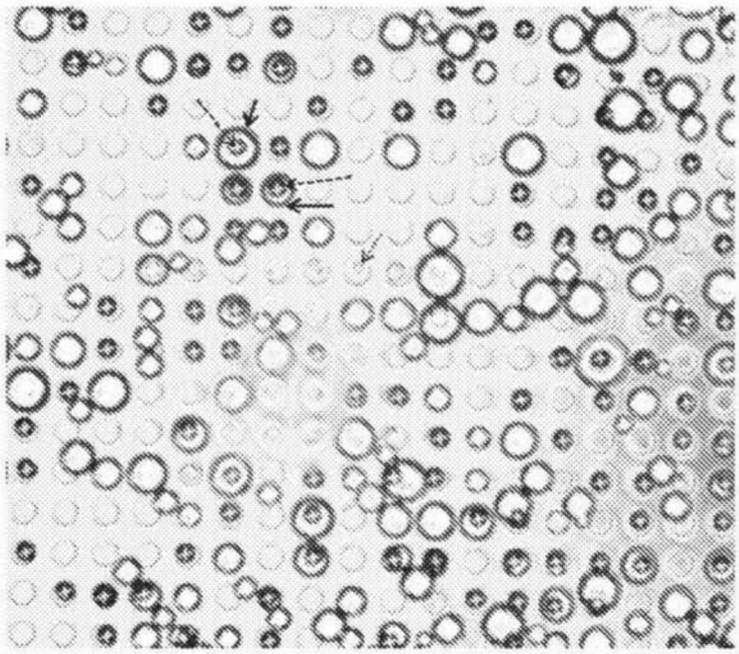


图8B

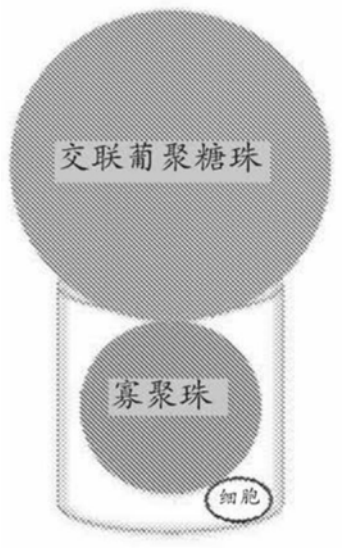


图8C

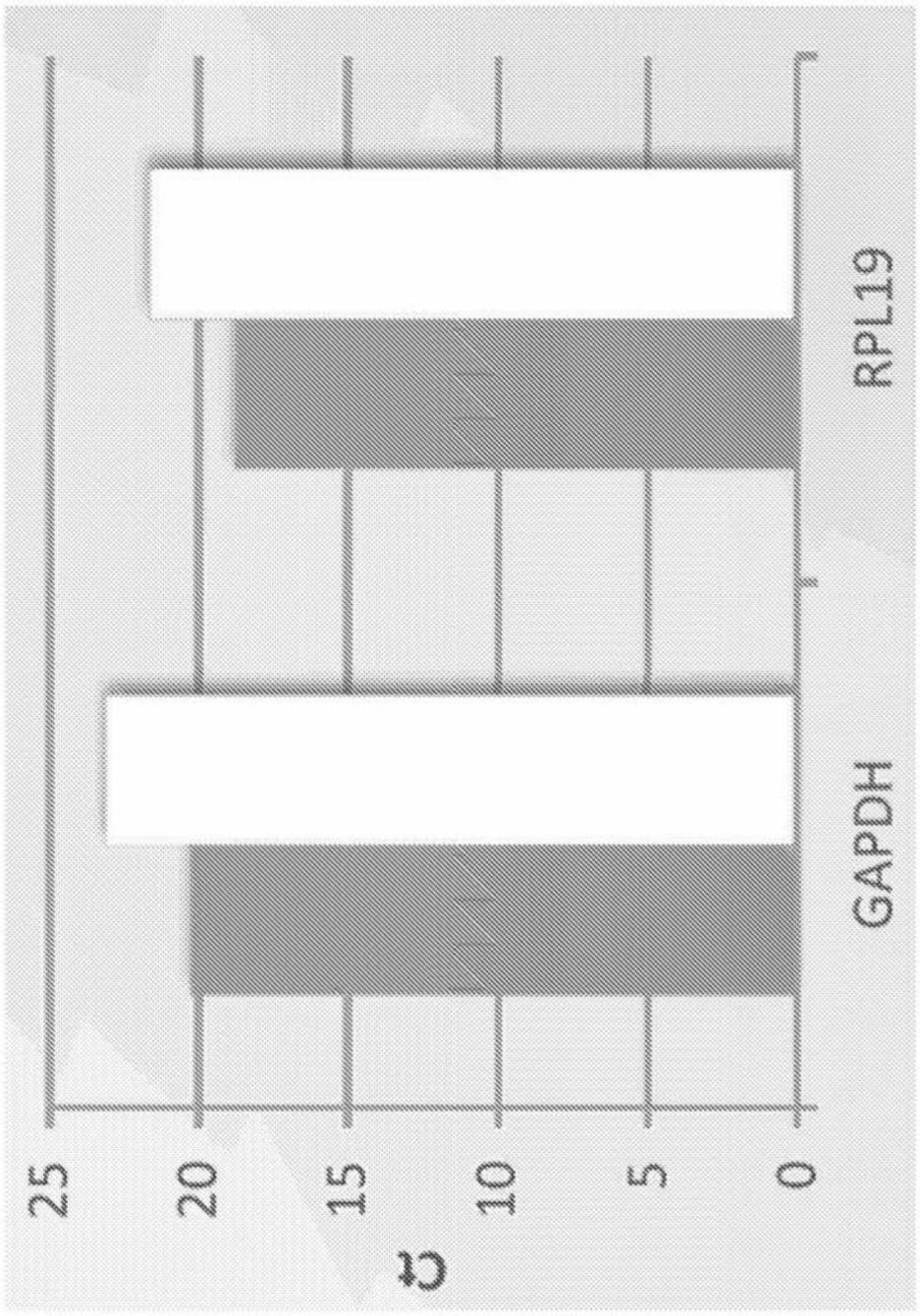


图9

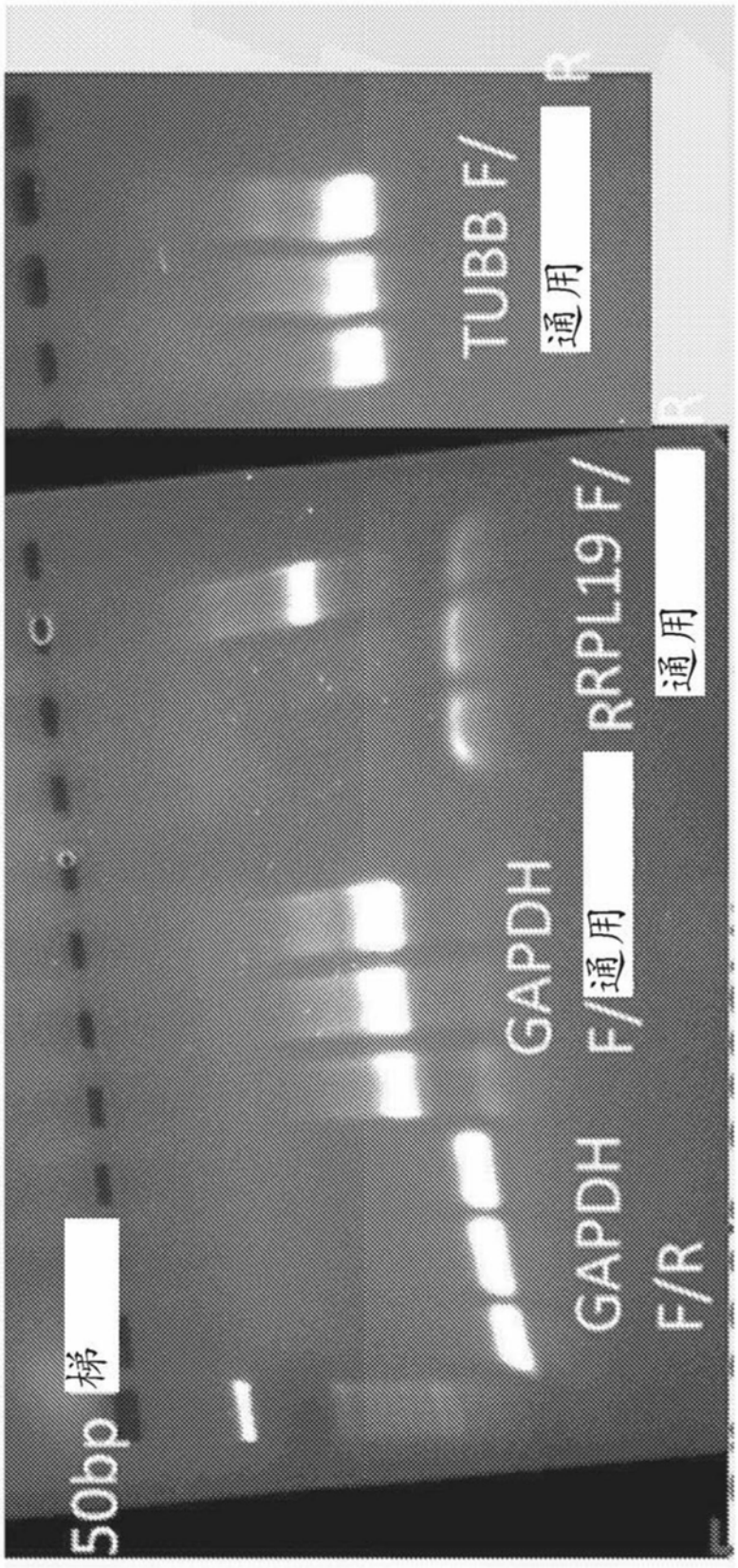


图10

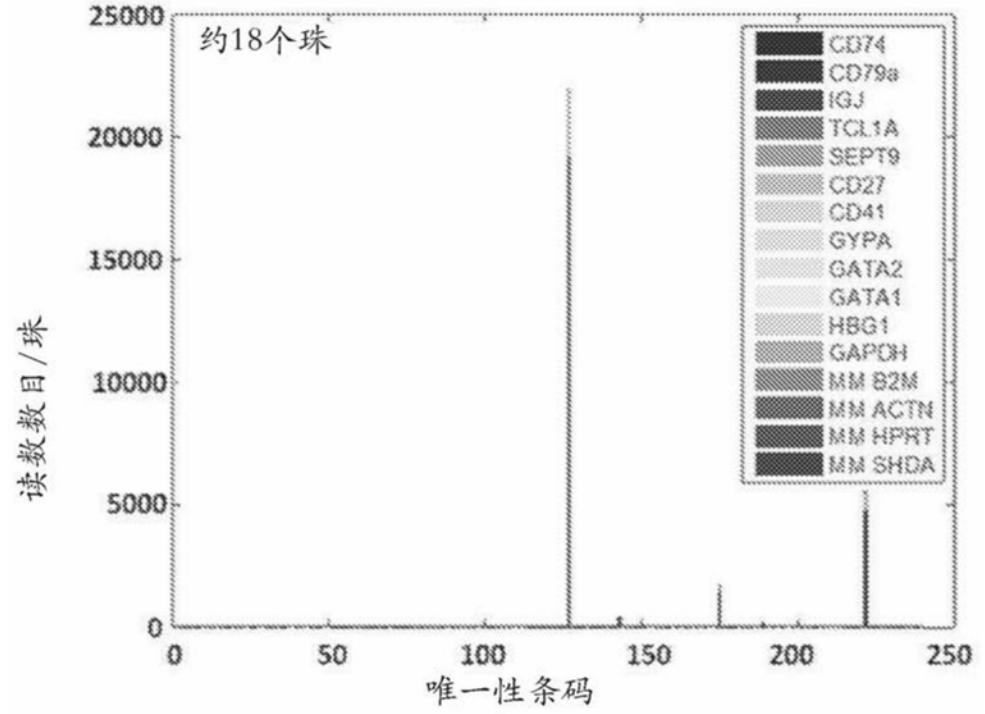

约18个珠
读数数目/珠
唯一性条码
0
5000
10000
15000
20000
25000
0
50
100
150
200
250
CD74
CD79a
IGJ
TCL1A
SEPT9
CD27
CD41
GYPA
GATA2
GATA1
HBG1
GAPDH
MM B2M
MM ACTN
MM HPRT
MM SHDA

图11A

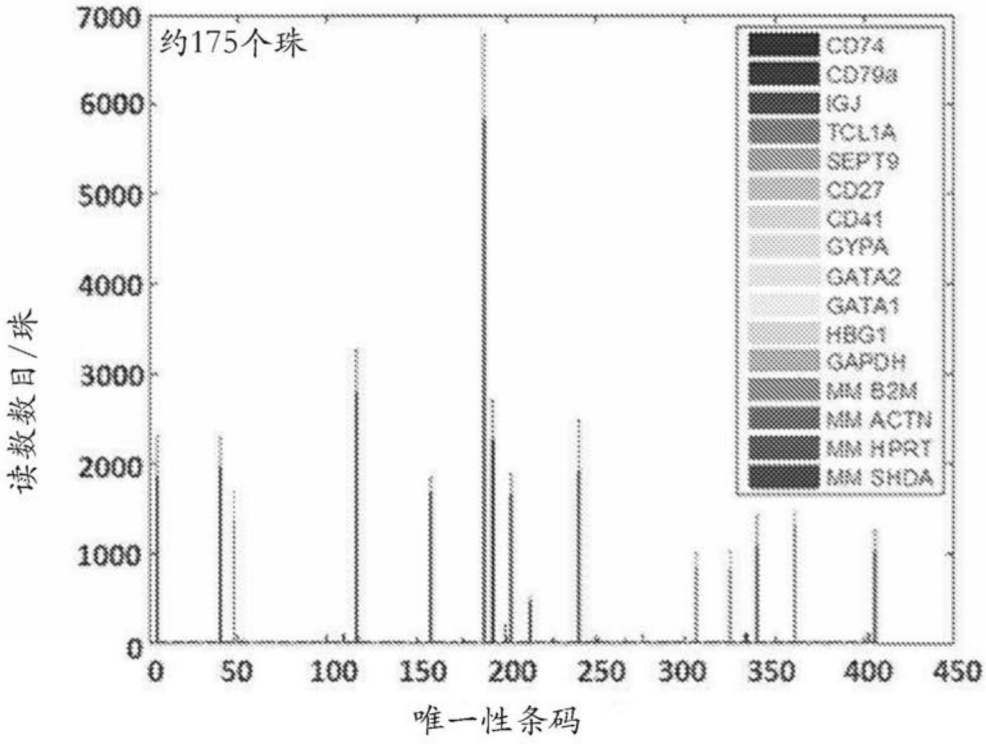

约175个珠
读数数目/珠
唯一性条码
0
1000
2000
3000
4000
5000
6000
7000
0
50
100
150
200
250
300
350
400
450
CD74
CD79a
IGJ
TCL1A
SEPT9
CD27
CD41
GYPA
GATA2
GATA1
HBG1
GAPDH
MM B2M
MM ACTN
MM HPRT
MM SHDA

图11B

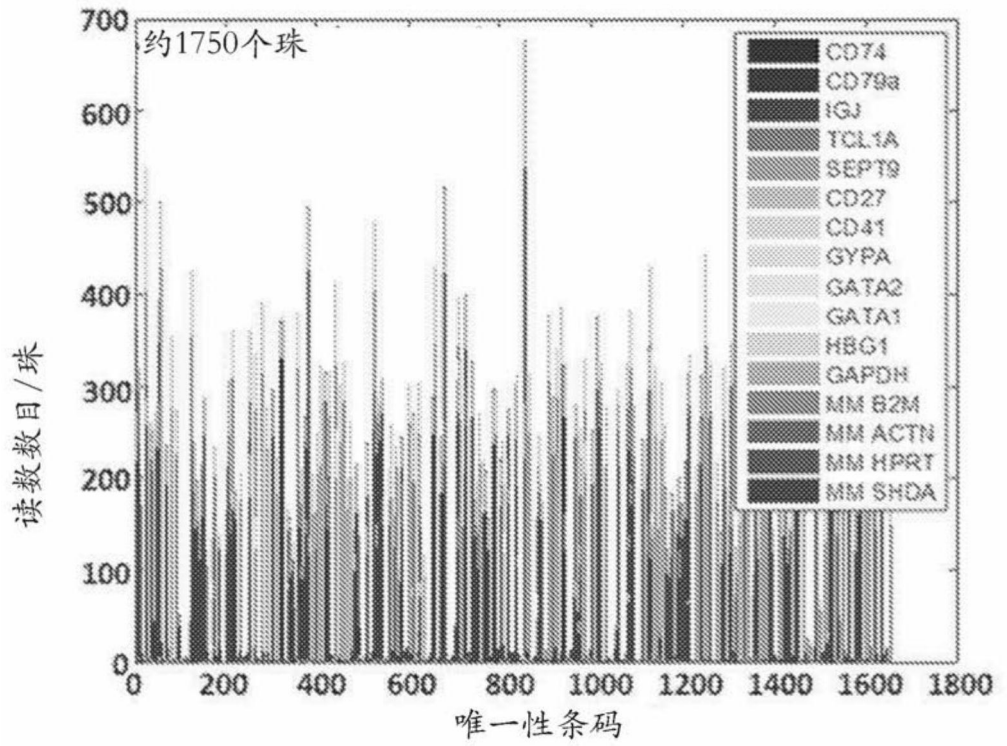

约1750个珠
读数数目/珠
唯一性条码
CD74
CD79a
IGJ
TCL1A
SEPT9
CD27
CD41
GYPA
GATA2
GATA1
HBG1
GAPDH
MM B2M
MM ACTN
MM HPRT
MM SHDA
0
100
200
300
400
500
600
700
0
200
400
600
800
1000
1200
1400
1600
1800

图11C

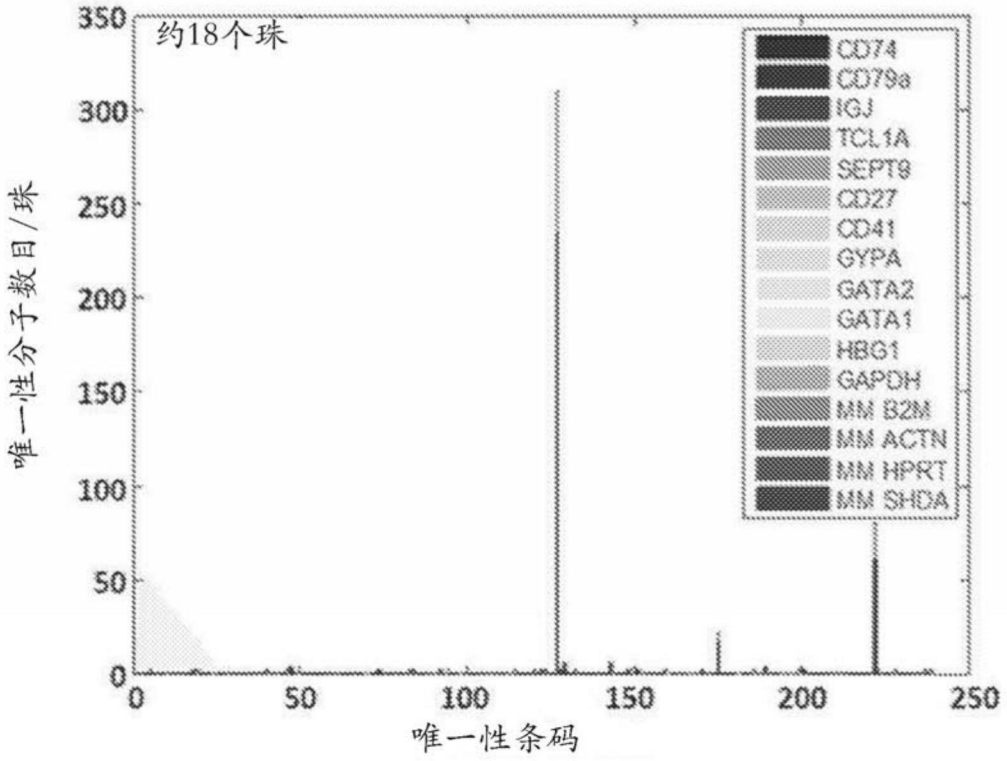

约18个珠
唯一性分子数目/珠
唯一性条码
CD74
CD79a
IGJ
TCL1A
SEPT9
CD27
CD41
GYPA
GATA2
GATA1
HBG1
GAPDH
MM B2M
MM ACTN
MM HPRT
MM SHDA
0
50
100
150
200
250
300
350
0
50
100
150
200
250

图11D

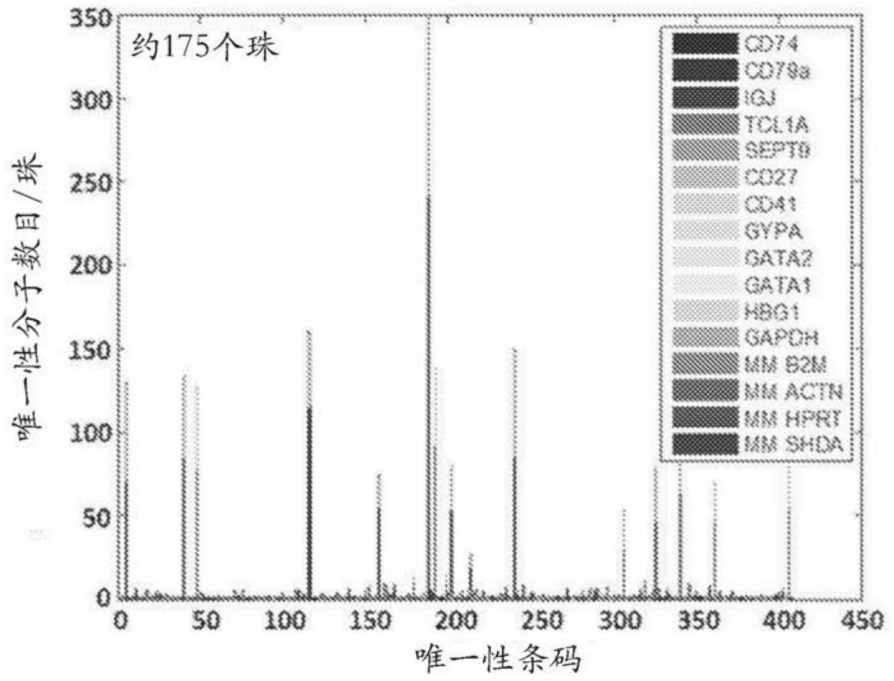

约175个珠
唯一性分子数目/珠
唯一性条码
CD74
CD79a
IGJ
TCL1A
SEPT9
CD27
CD41
GYPA
GATA2
GATA1
HBG1
GAPDH
MM B2M
MM ACTN
MM HPRT
MM SHDA

图11E

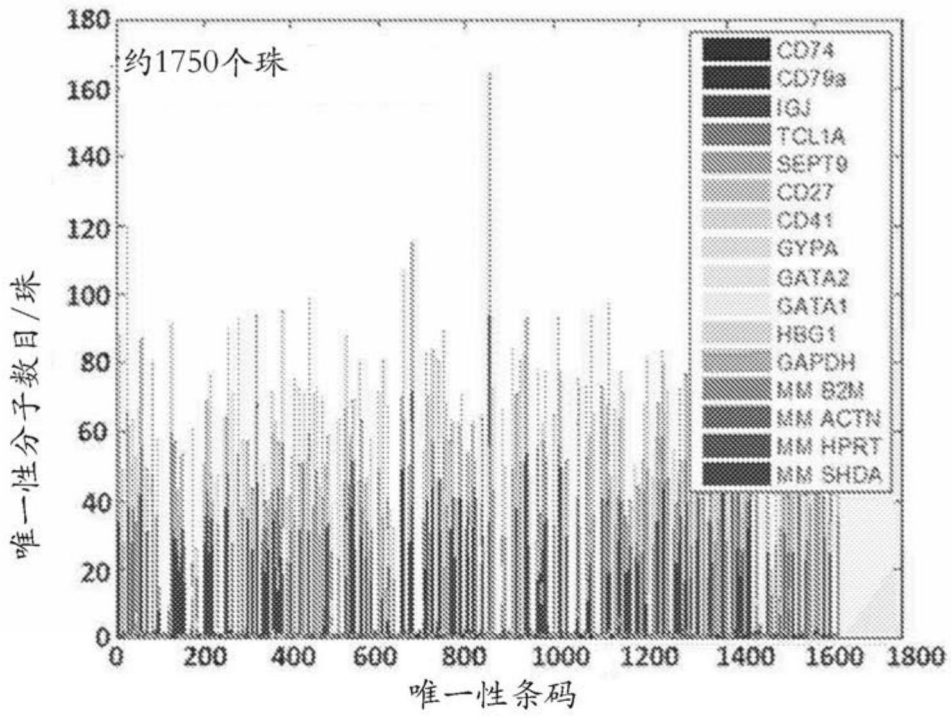

约1750个珠
唯一性分子数目/珠
唯一性条码
CD74
CD79a
IGJ
TCL1A
SEPT9
CD27
CD41
GYPA
GATA2
GATA1
HBG1
GAPDH
MM B2M
MM ACTN
MM HPRT
MM SHDA

图11F

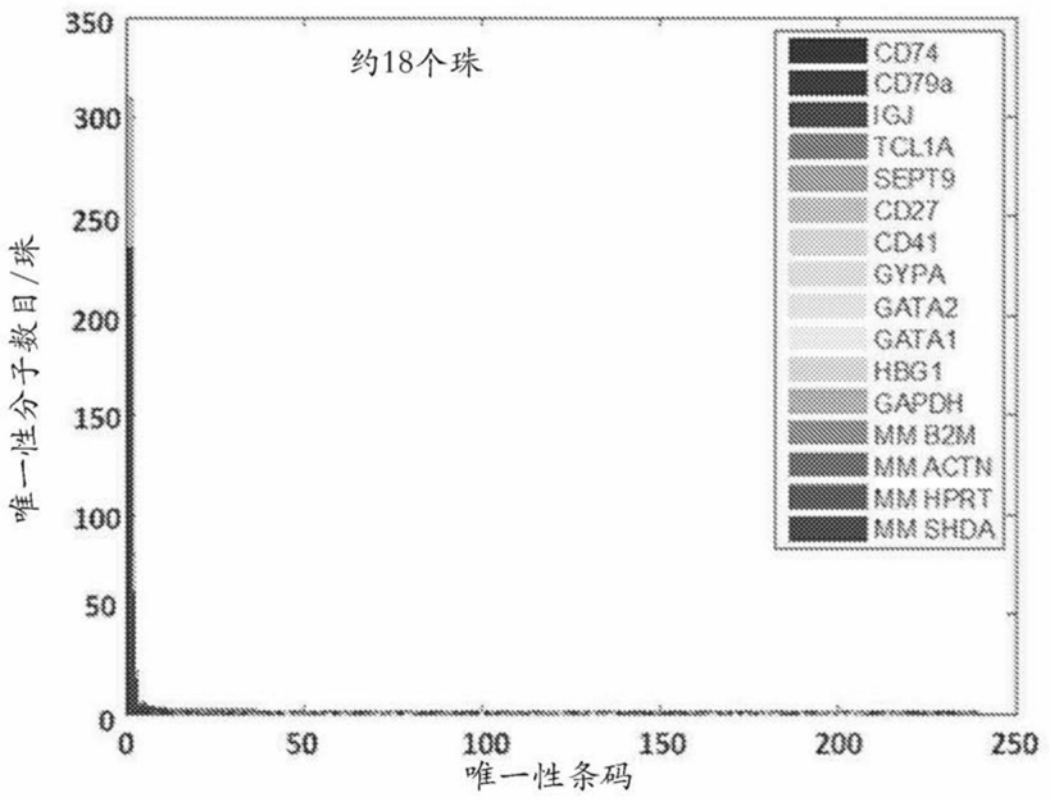

约18个珠
CD74
CD79a
IGJ
TCL1A
SEPT9
CD27
CD41
GYPA
GATA2
GATA1
HBG1
GAPDH
MM B2M
MM ACTN
MM HPRT
MM SHDA
唯一性分子数目/珠
唯一性条码
0
50
100
150
200
250
300
350

图11G

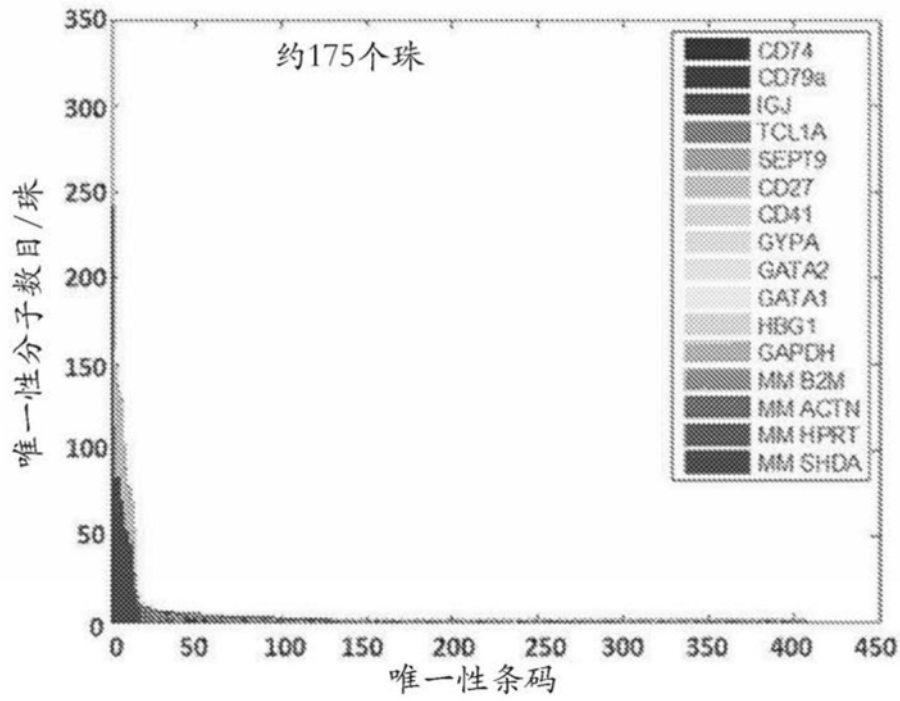

约175个珠
CD74
CD79a
IGJ
TCL1A
SEPT9
CD27
CD41
GYPA
GATA2
GATA1
HBG1
GAPDH
MM B2M
MM ACTN
MM HPRT
MM SHDA
唯一性分子数目/珠
唯一性条码
0
50
100
150
200
250
300
350
400
450

图11H

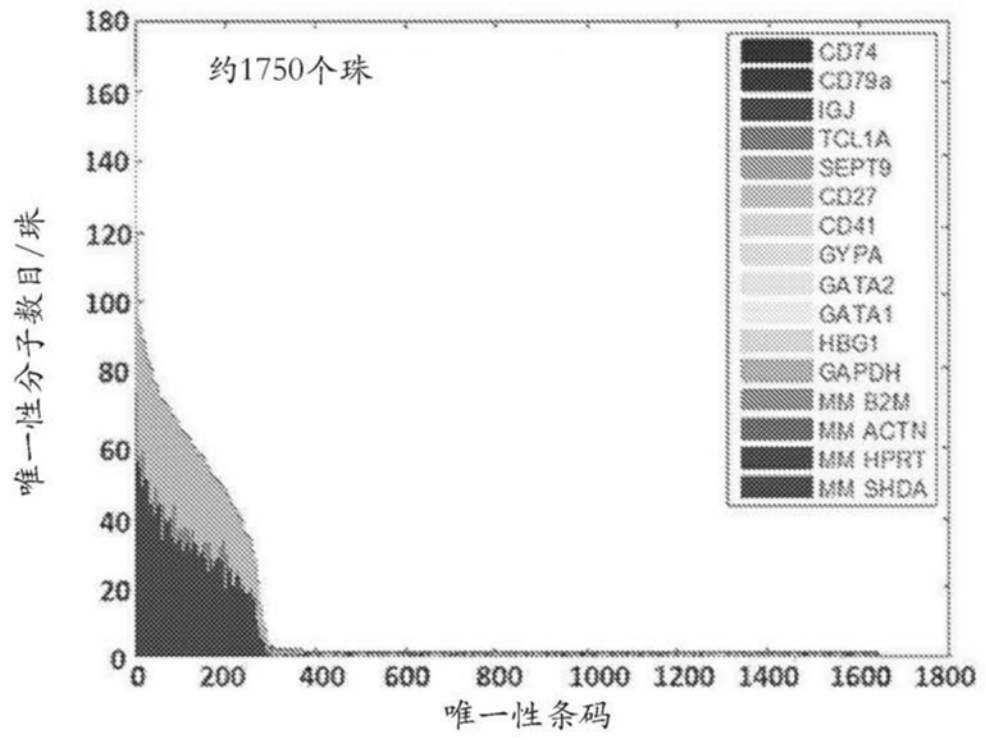

约1750个珠
CD74
CD79a
IGJ
TCL1A
SEPT9
CD27
CD41
GYPA
GATA2
GATA1
HBG1
GAPDH
MM B2M
MM ACTN
MM HPRT
MM SHDA
唯一性分子数目/珠
唯一性条码
0
20
40
60
80
100
120
140
160
180
0
200
400
600
800
1000
1200
1400
1600
1800

图11I

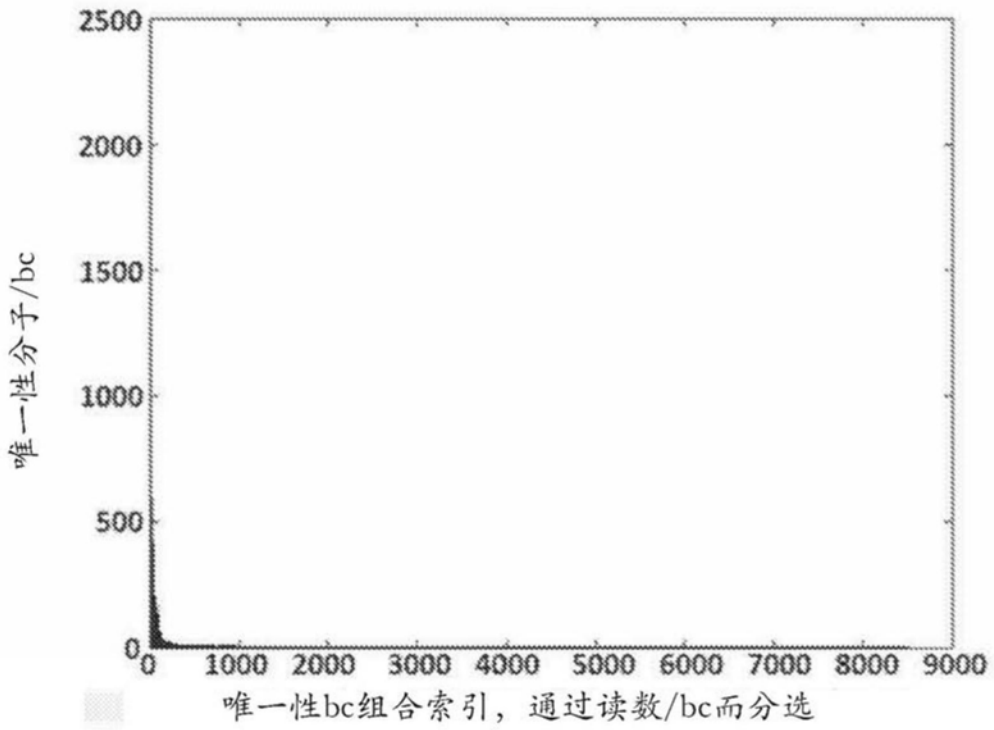

唯一性分子/bc
唯一性bc组合索引，通过读数/bc而分选
0
500
1000
1500
2000
2500
0
1000
2000
3000
4000
5000
6000
7000
8000
9000

图12A

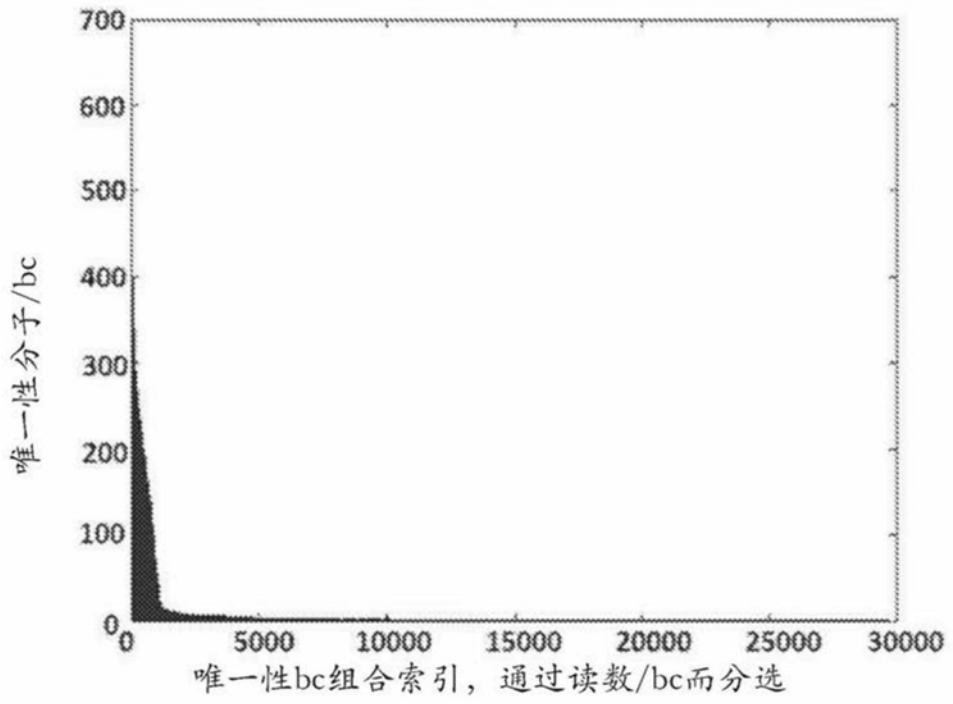

700
600
500
400
300
200
100
0
唯一性分子/bc
0
5000
10000
15000
20000
25000
30000
唯一性bc组合索引，通过读数/bc而分选

图12B

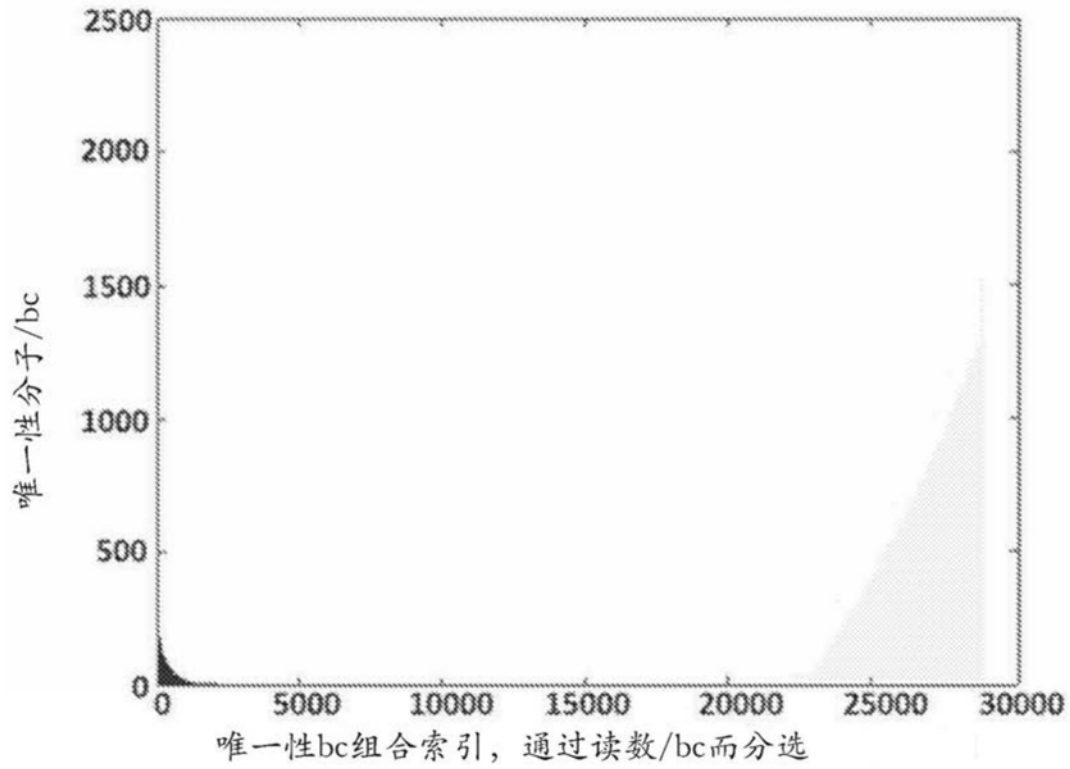

2500
2000
1500
1000
500
0
唯一性分子/bc
0
5000
10000
15000
20000
25000
30000
唯一性bc组合索引，通过读数/bc而分选

图12C

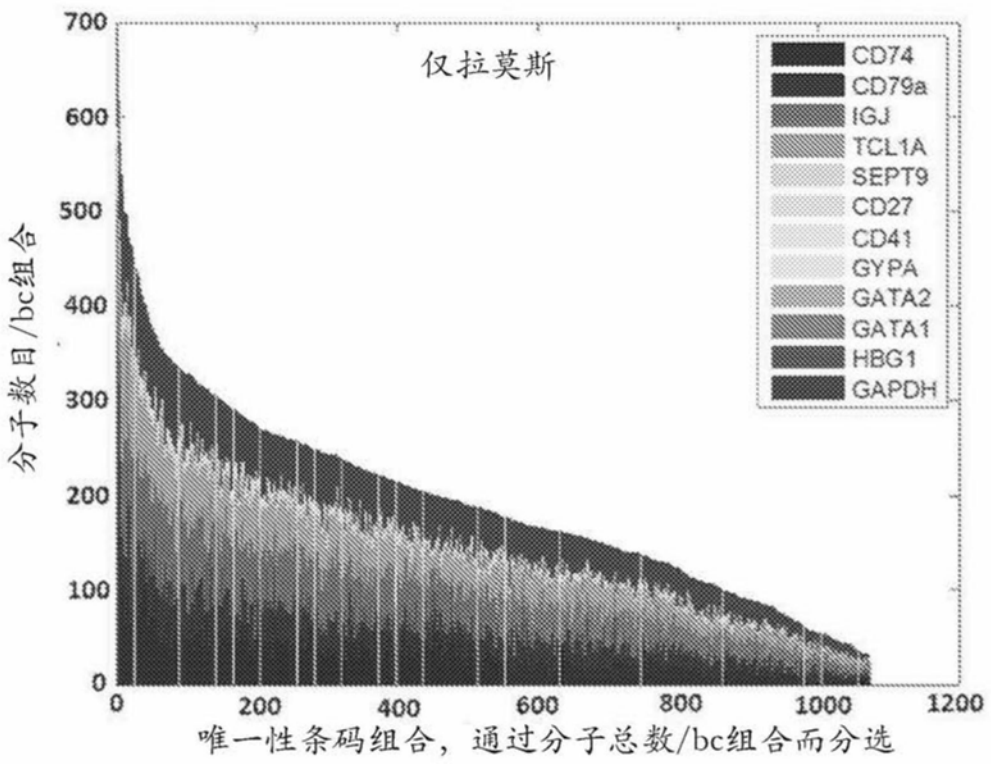

仅拉莫斯
CD74
CD79a
IGJ
TCL1A
SEPT9
CD27
CD41
GYPA
GATA2
GATA1
HBG1
GAPDH
分子数目/bc组合
0
100
200
300
400
500
600
700
0
200
400
600
800
1000
1200
唯一性条码组合，通过分子总数/bc组合而分选

图12D

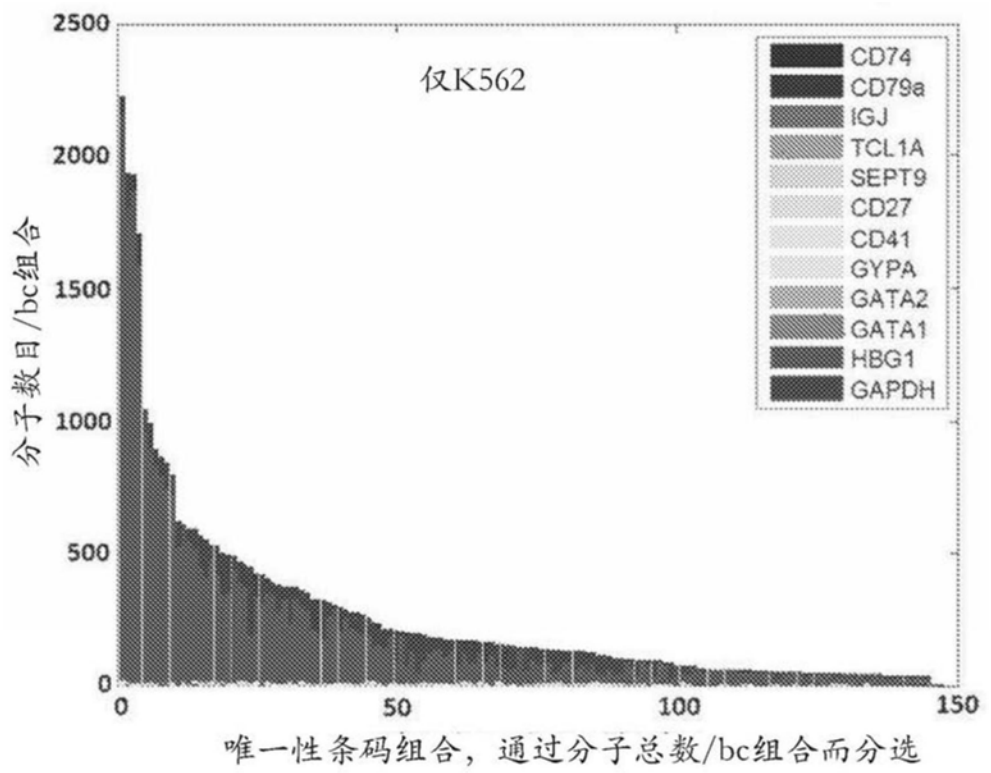

仅K562
CD74
CD79a
IGJ
TCL1A
SEPT9
CD27
CD41
GYPA
GATA2
GATA1
HBG1
GAPDH
分子数目/bc组合
0
500
1000
1500
2000
2500
0
50
100
150
唯一性条码组合，通过分子总数/bc组合而分选

图12E

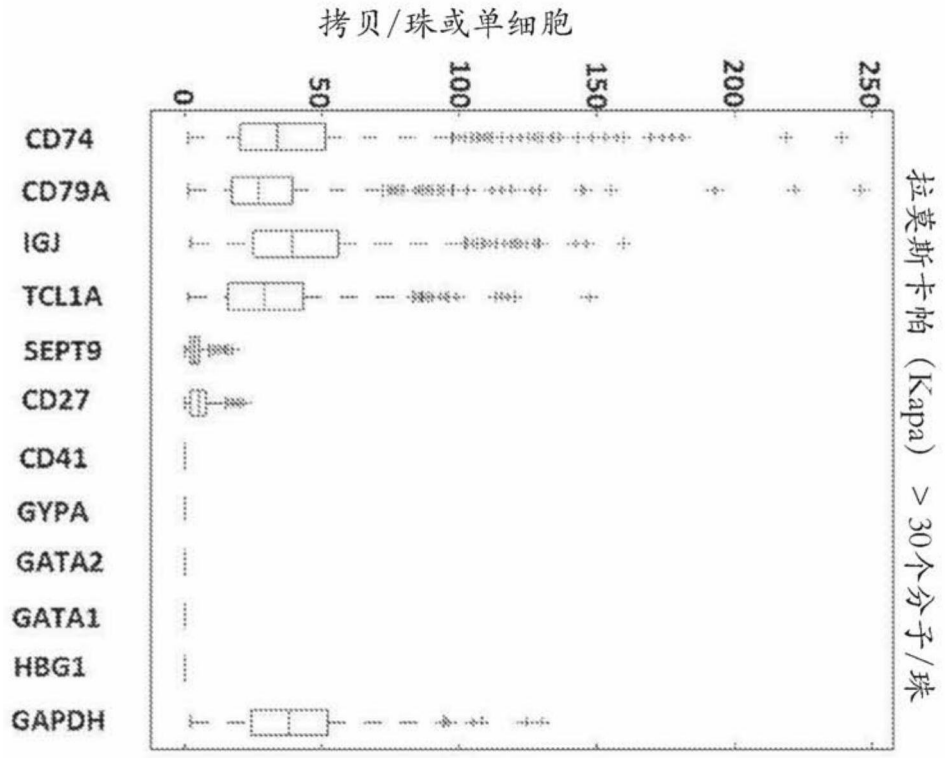

拷贝/珠或单细胞
0
50
100
150
200
250
CD74
CD79A
IGJ
TCL1A
SEPT9
CD27
CD41
GYPA
GATA2
GATA1
HBG1
GAPDH
拉莫斯卡帕（Kapa）＞30个分子/珠

图12F

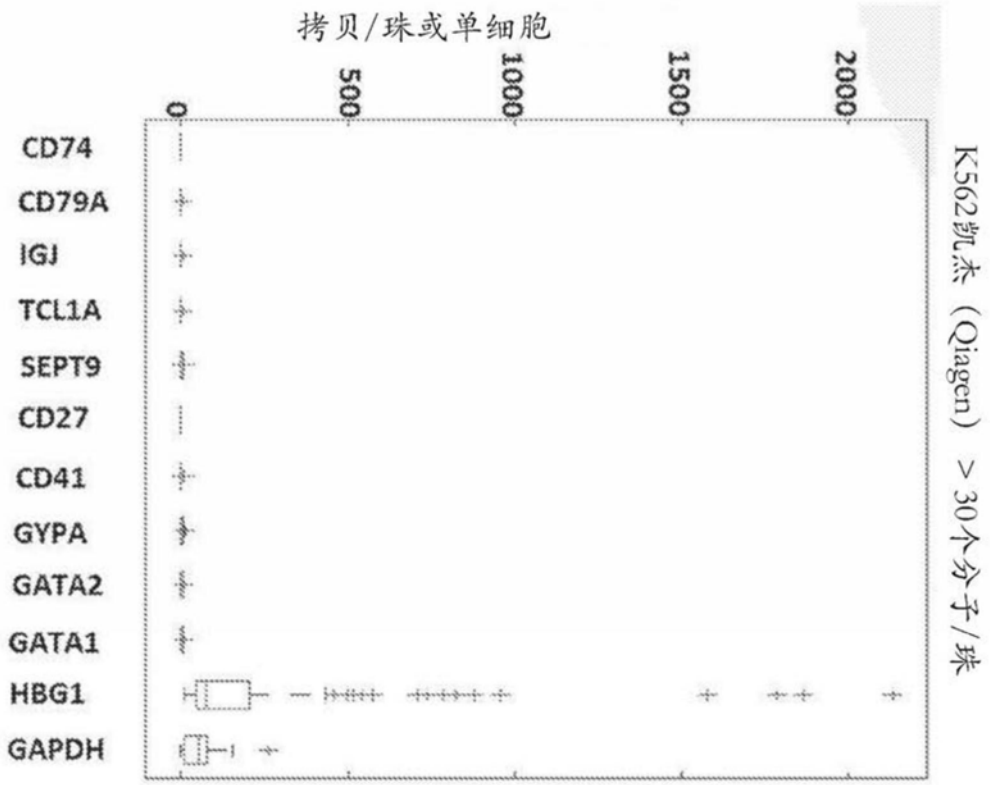

拷贝/珠或单细胞
0
500
1000
1500
2000
CD74
CD79A
IGJ
TCL1A
SEPT9
CD27
CD41
GYPA
GATA2
GATA1
HBG1
GAPDH
K562凯杰（Qiagen）＞30个分子/珠

图12G

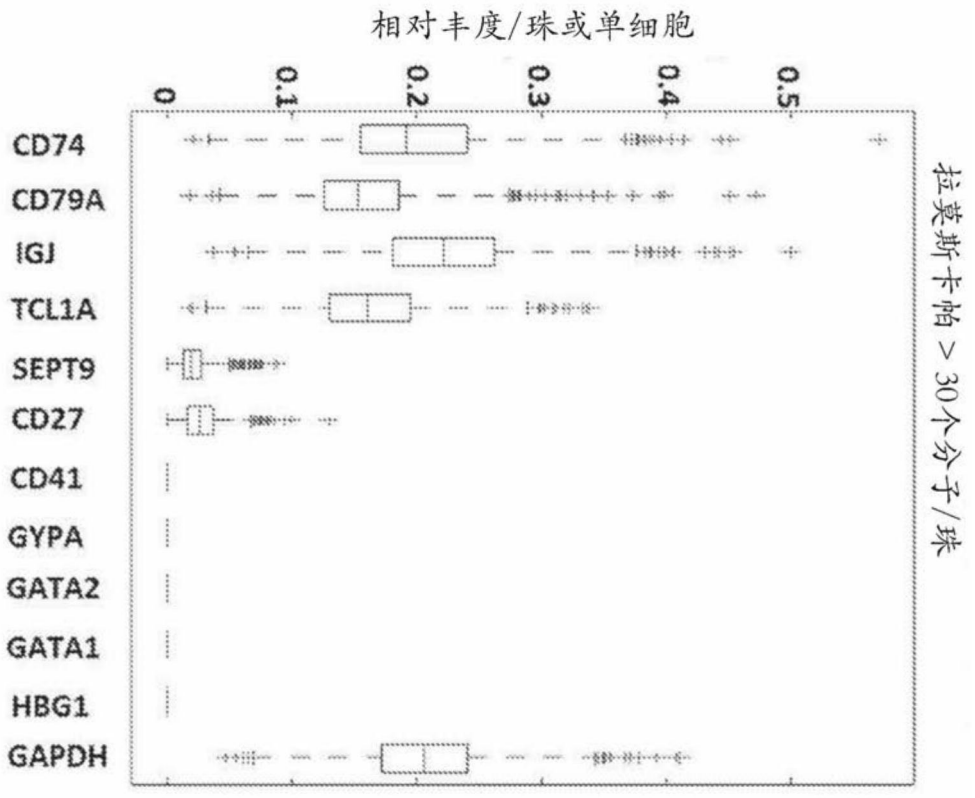

相对丰度/珠或单细胞
0
0.1
0.2
0.3
0.4
0.5
CD74
CD79A
IGJ
TCL1A
SEPT9
CD27
CD41
GYPA
GATA2
GATA1
HBG1
GAPDH
拉莫斯卡帕 > 30个分子/珠

图12H

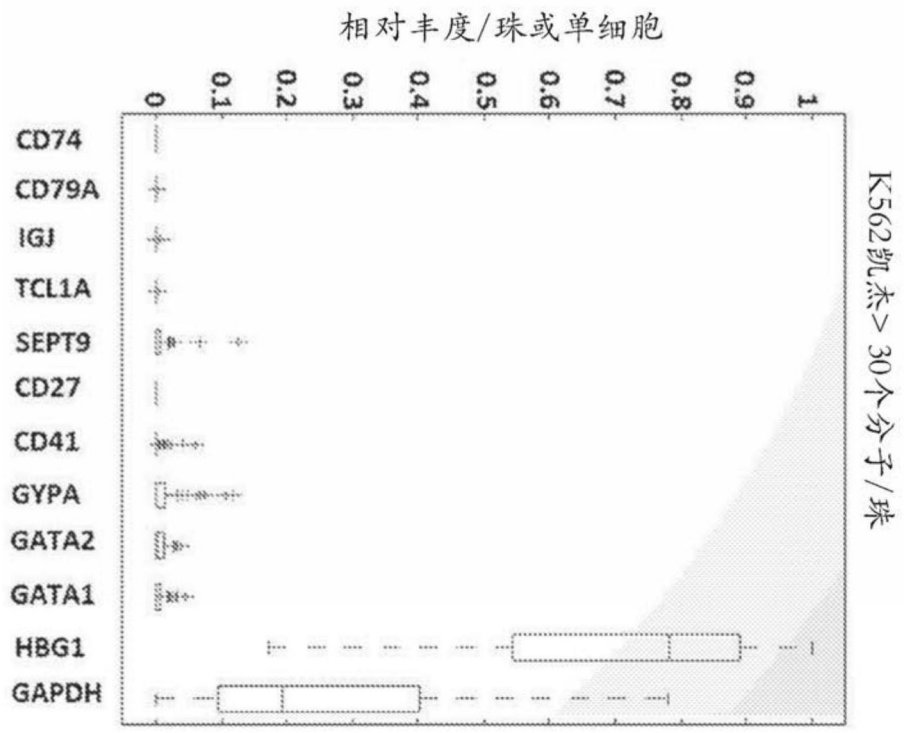

相对丰度/珠或单细胞
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
CD74
CD79A
IGJ
TCL1A
SEPT9
CD27
CD41
GYPA
GATA2
GATA1
HBG1
GAPDH
K562凯杰> 30个分子/珠

图12I

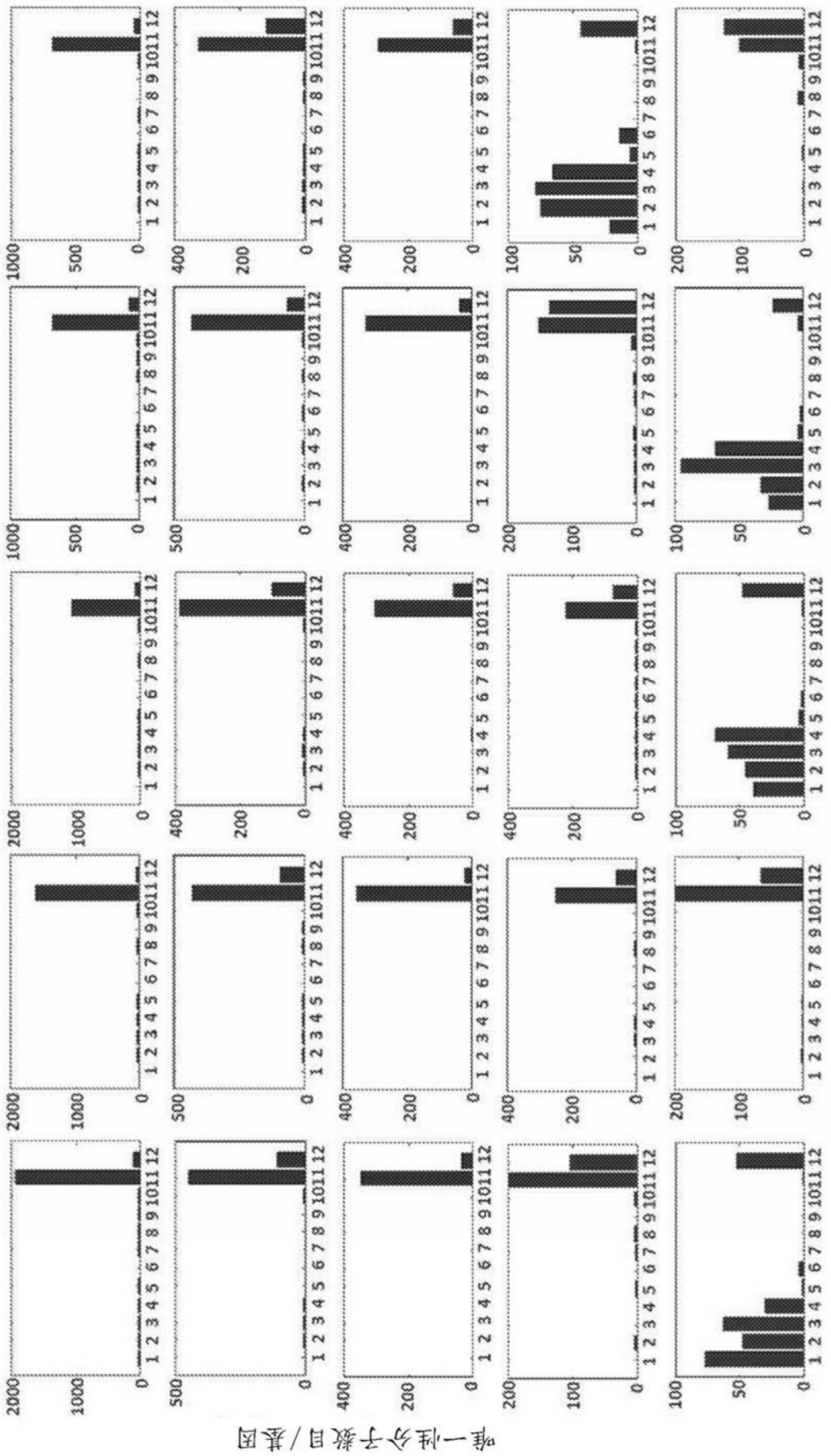


图12J

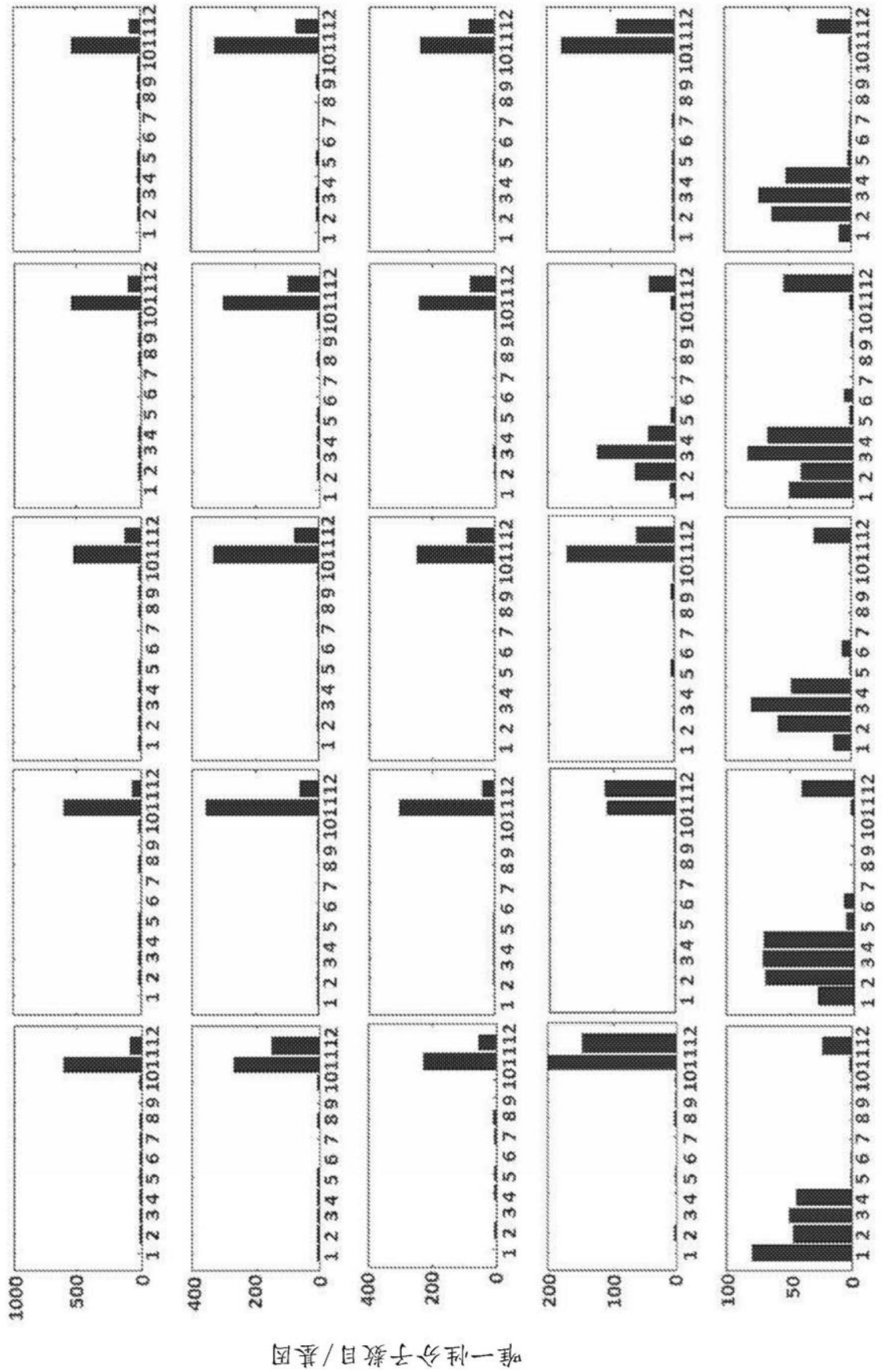


图12K

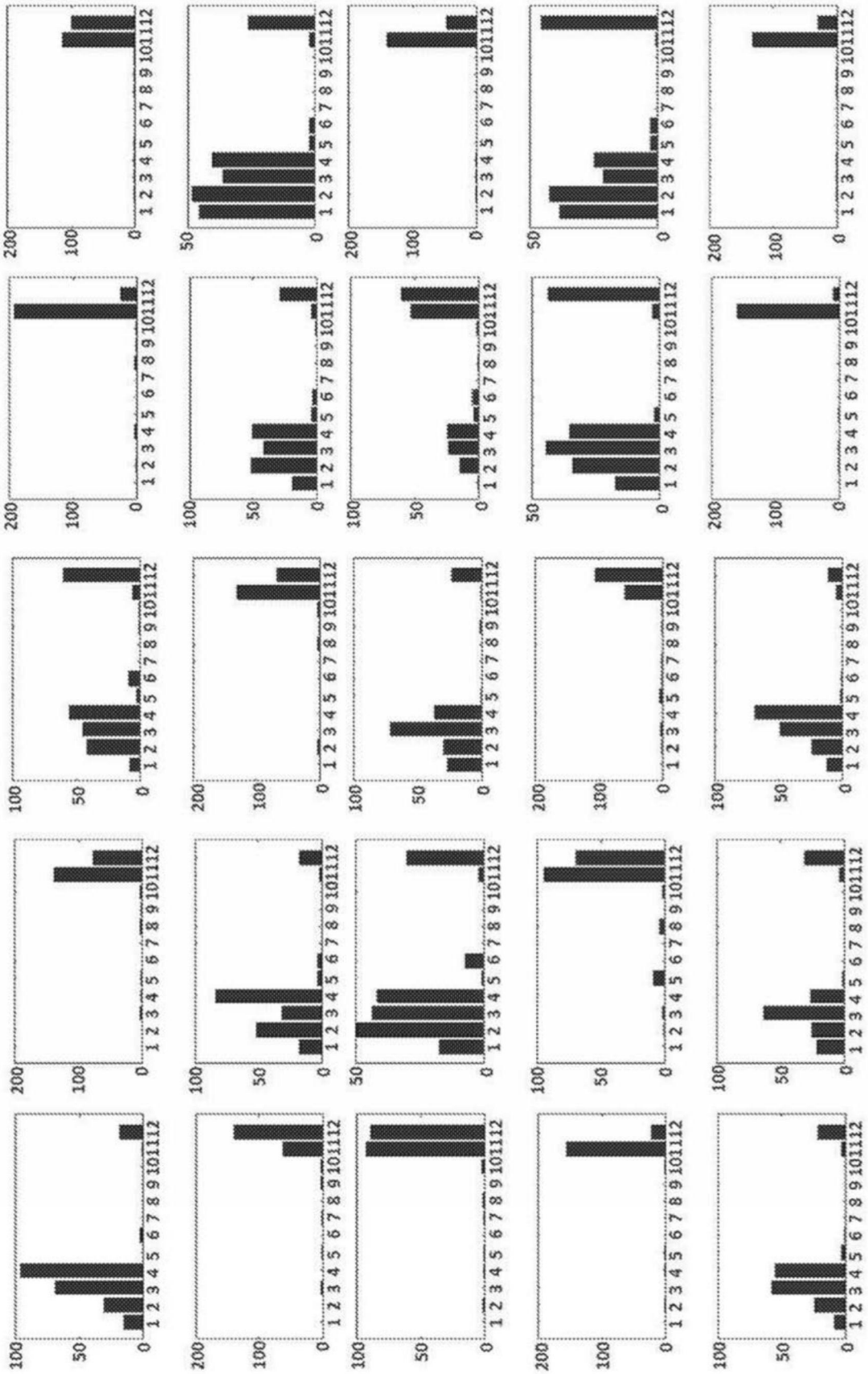


图12L

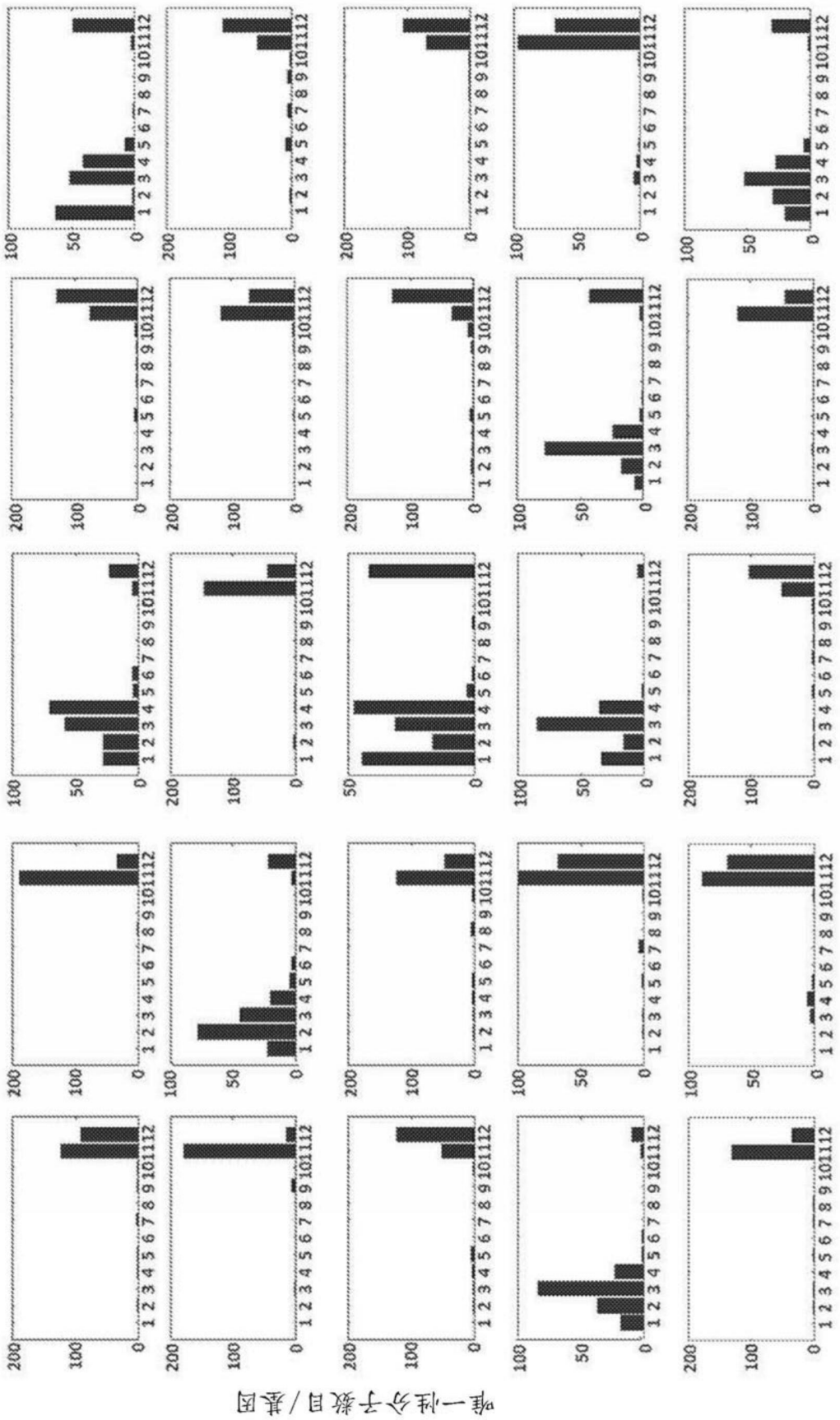


图12M

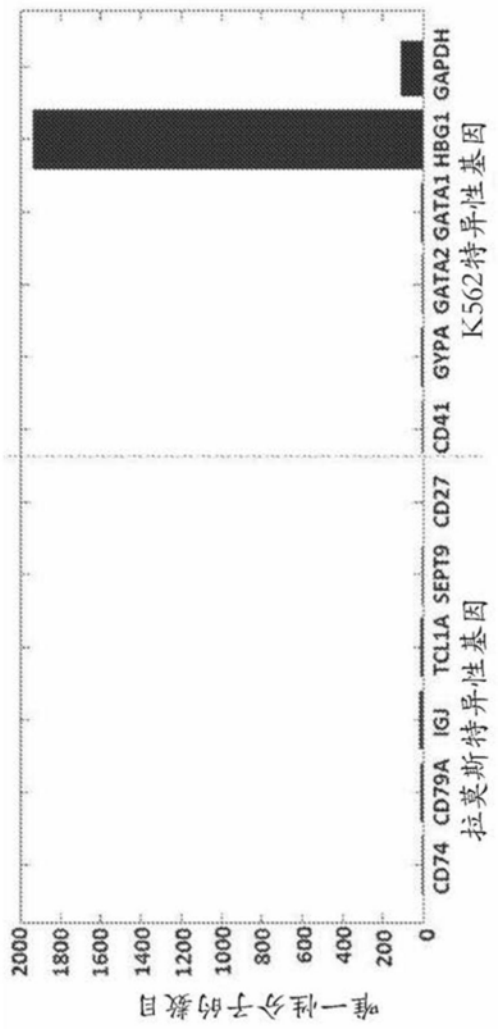

唯一性分子的数目
2000
1800
1600
1400
1200
1000
800
600
400
200
0
CD74 CD79A IGJ TCL1A SEPT9 CD27
拉莫斯特异性基因
CD41 GYPA GATA2 GATA1 HBG1 GAPDH
K562特异性基因

图12N

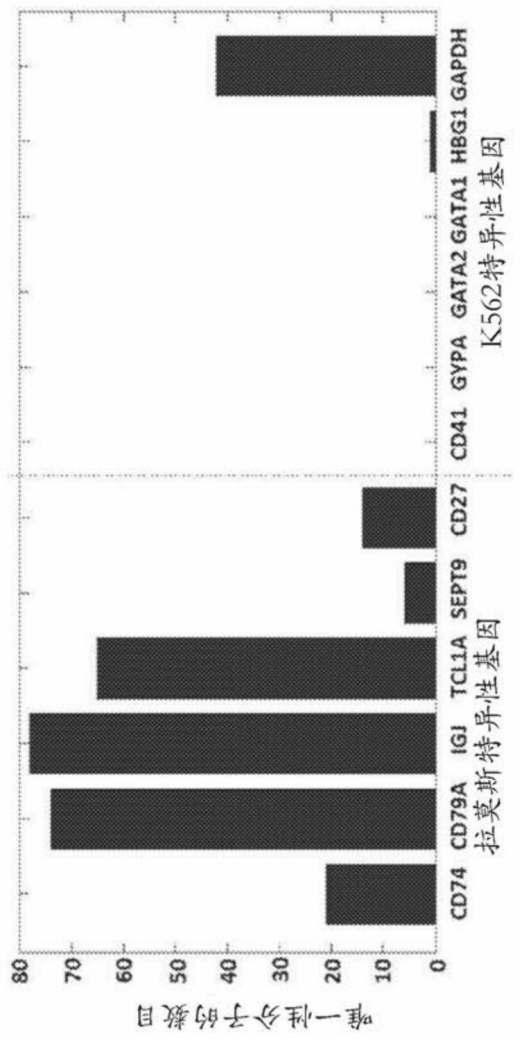

唯一性分子的数目
0
10
20
30
40
50
60
70
80
CD74 CD79A IGJ TCL1A SEPT9 CD27
拉莫斯特异性基因
CD41 GYPA GATA2 GATA1 HBG1 GAPDH
K562特异性基因

图120

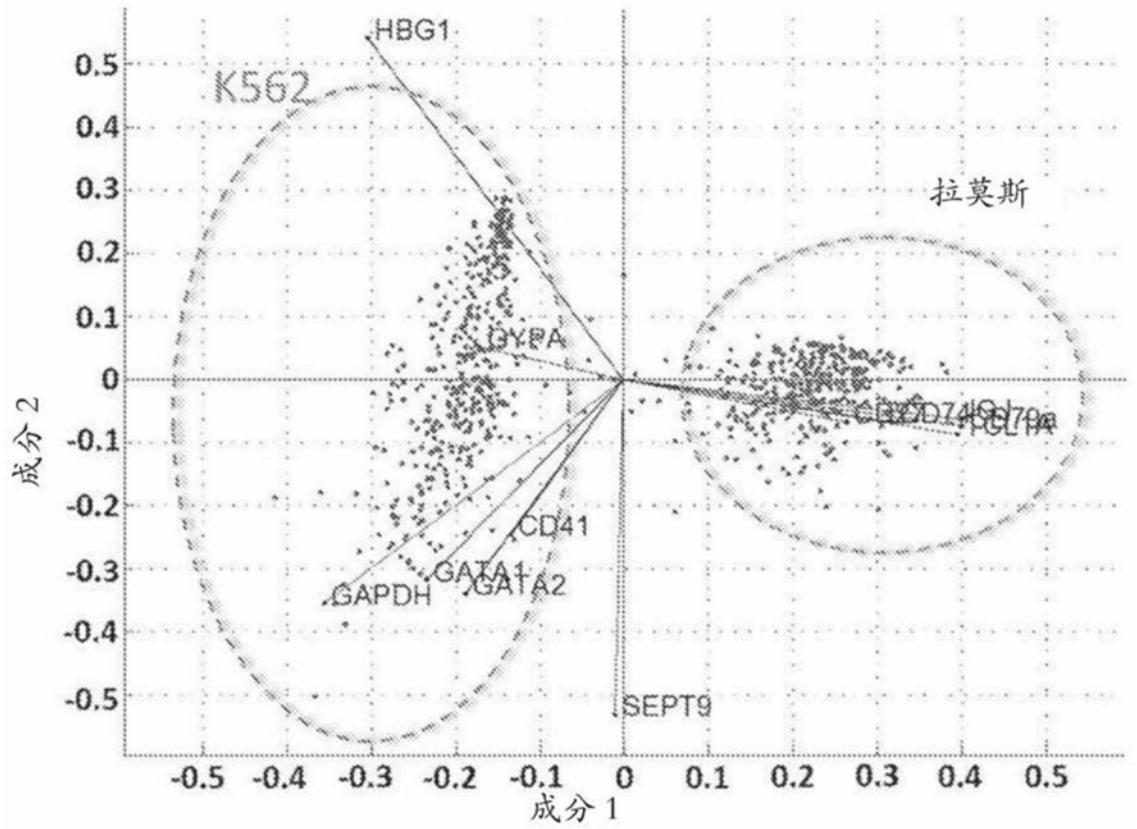

HBG1
K562
拉莫斯
GYPA
CD41
GATA1
GAPDH
GATA2
SEPT9
成分 2
成分 1

图12P

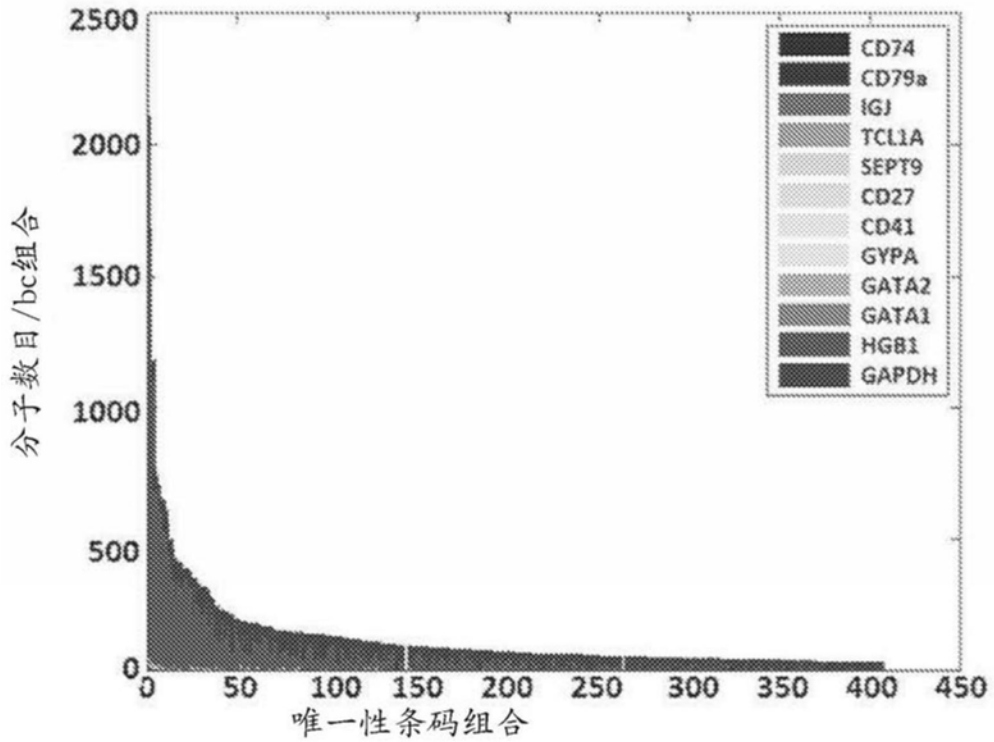

CD74
CD79a
IGJ
TCL1A
SEPT9
CD27
CD41
GYPA
GATA2
GATA1
HGB1
GAPDH
分子数目/bc组合
唯一性条码组合

图12Q

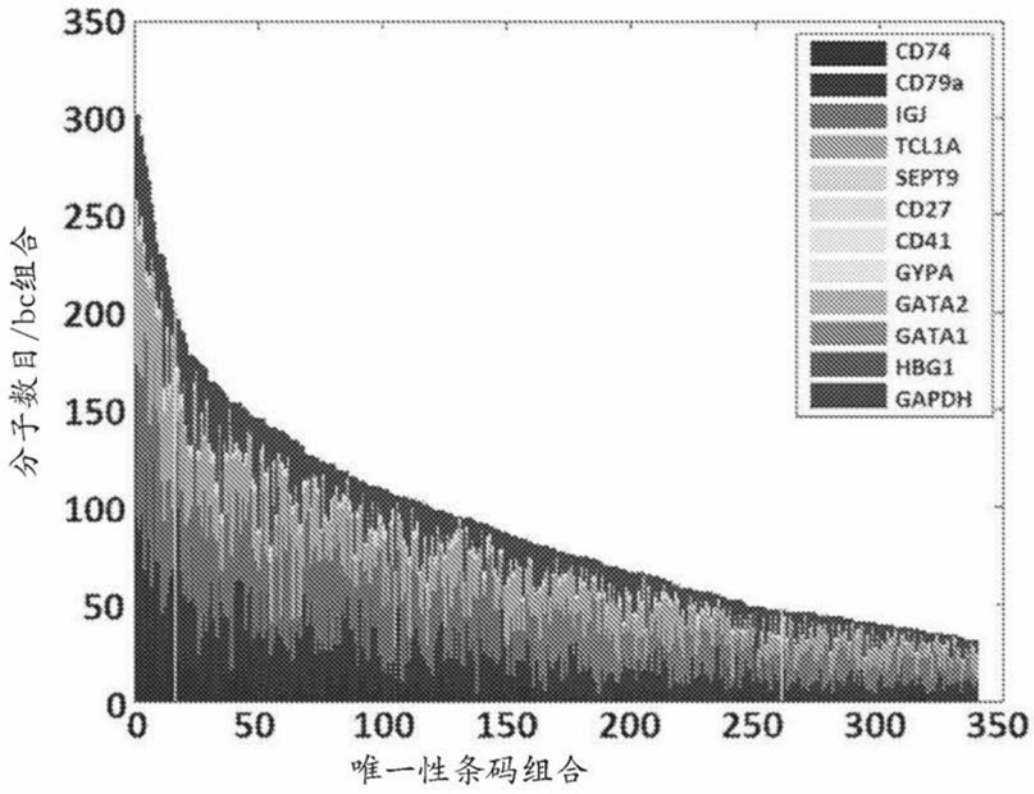

350
300
250
200
150
100
50
0
分子数目/bc组合
0
50
100
150
200
250
300
350
唯一性条码组合
CD74
CD79a
IGJ
TCL1A
SEPT9
CD27
CD41
GYPA
GATA2
GATA1
HBG1
GAPDH

图12R

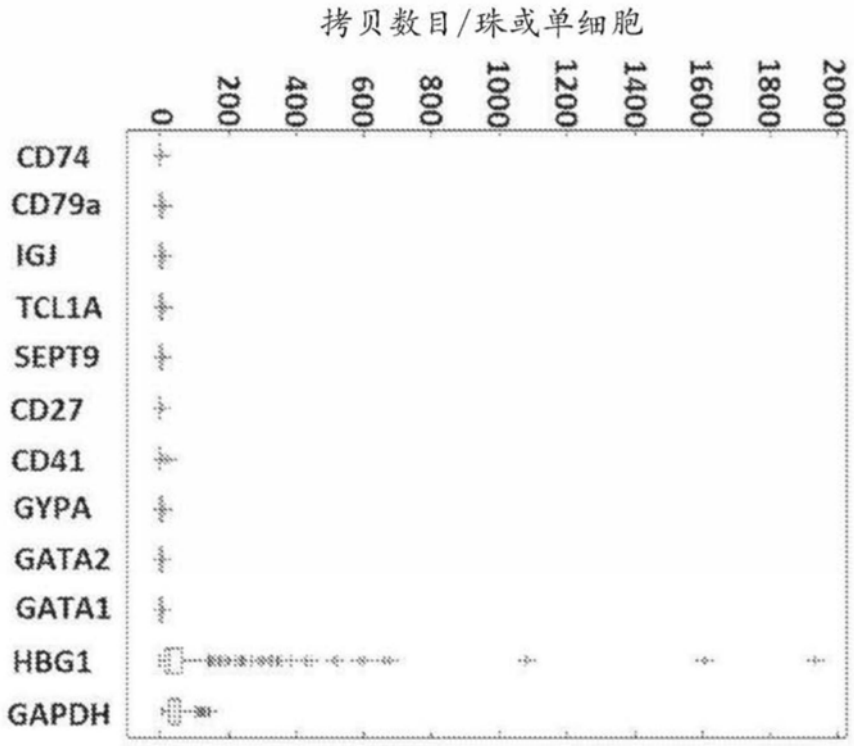

拷贝数目/珠或单细胞
0
200
400
600
800
1000
1200
1400
1600
1800
2000
CD74
CD79a
IGJ
TCL1A
SEPT9
CD27
CD41
GYPA
GATA2
GATA1
HBG1
GAPDH

图12S

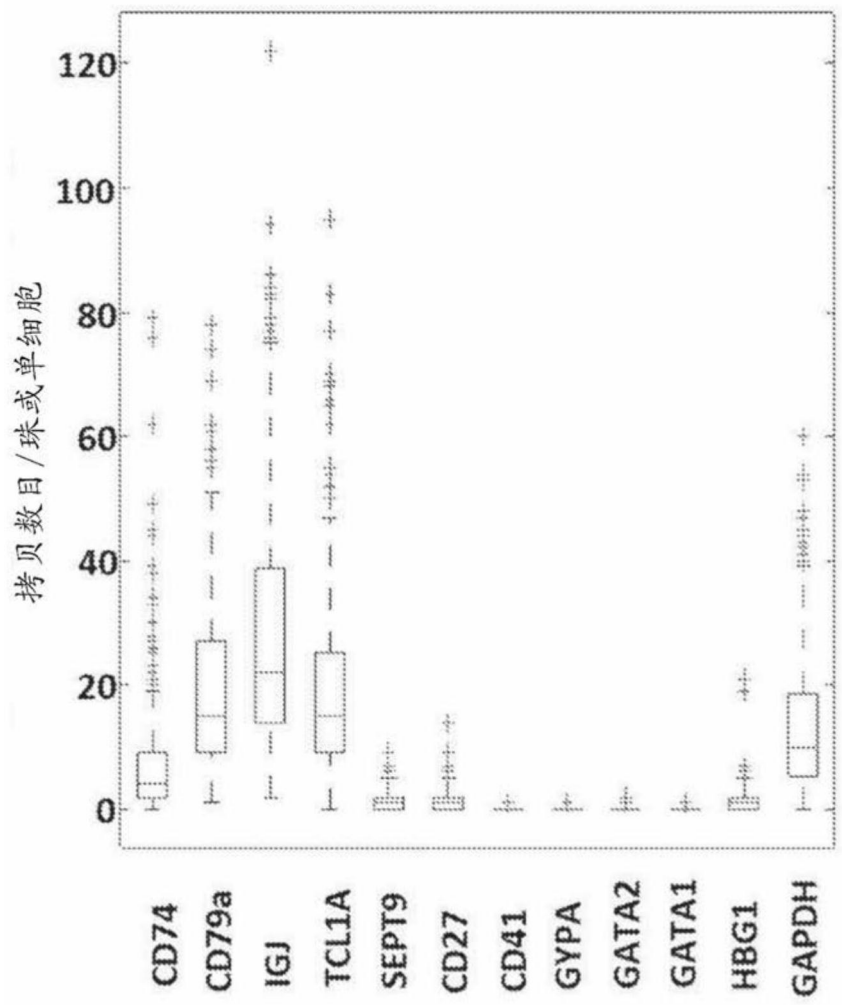

拷贝数目/珠或单细胞
120
100
80
60
40
20
0
CD74
CD79a
IGJ
TCL1A
SEPT9
CD27
CD41
GYPA
GATA2
GATA1
HBG1
GAPDH

图12T

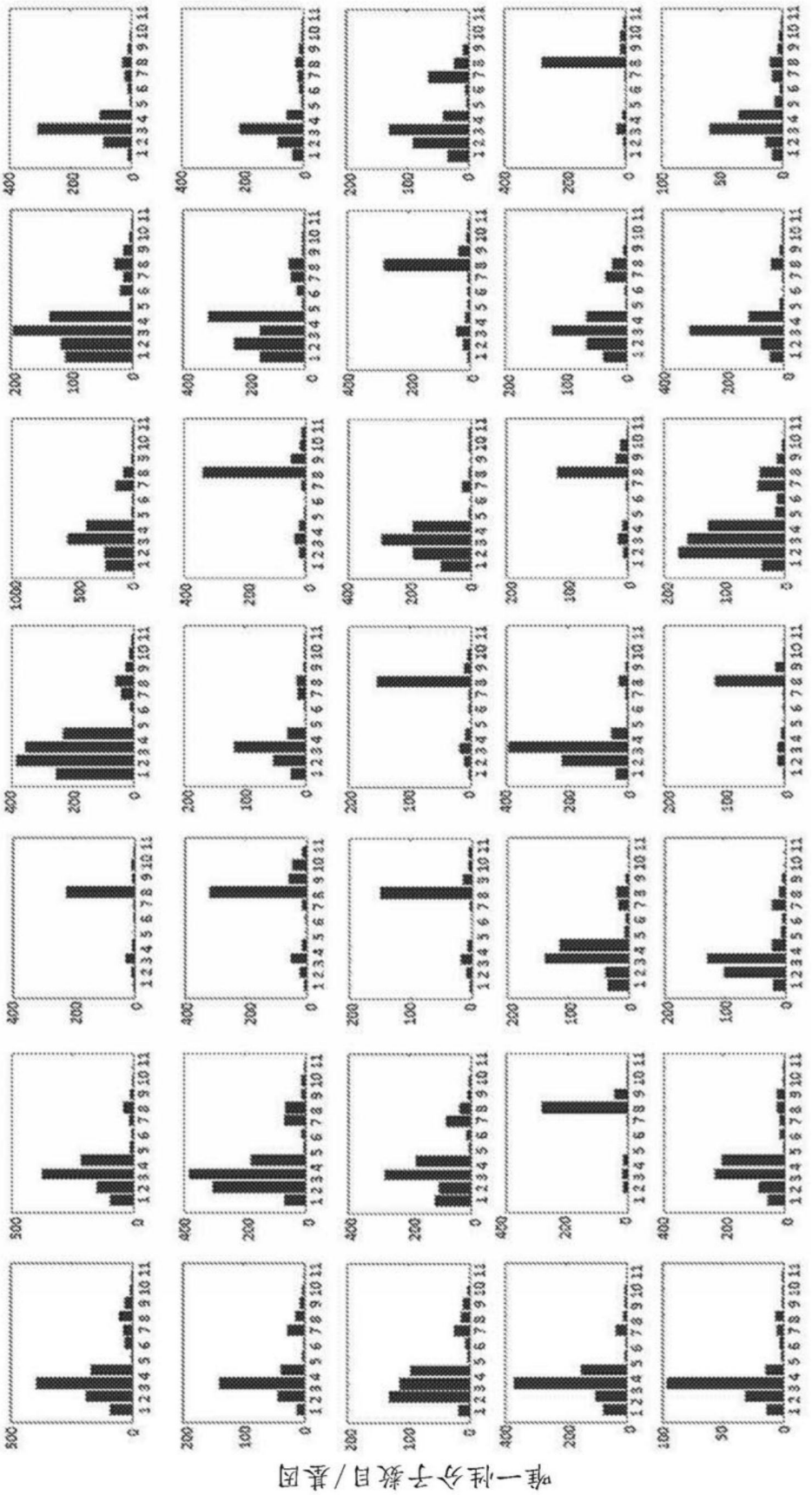


图13A

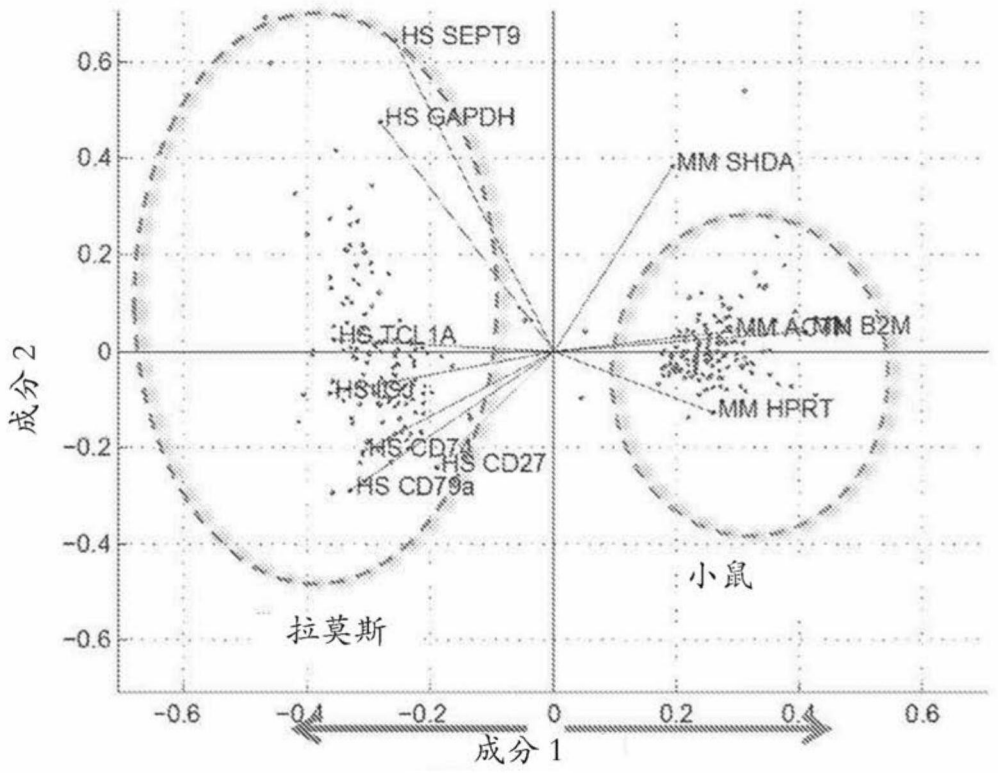

HS SEPT9
HS GAPDH
MM SHDA
HS TCL1A
MM HPRT
HS CD74
HS CD27
HS CD79a
拉莫斯
小鼠
成分 2
成分 1
0.6
0.4
0.2
0
-0.2
-0.4
-0.6

图13B

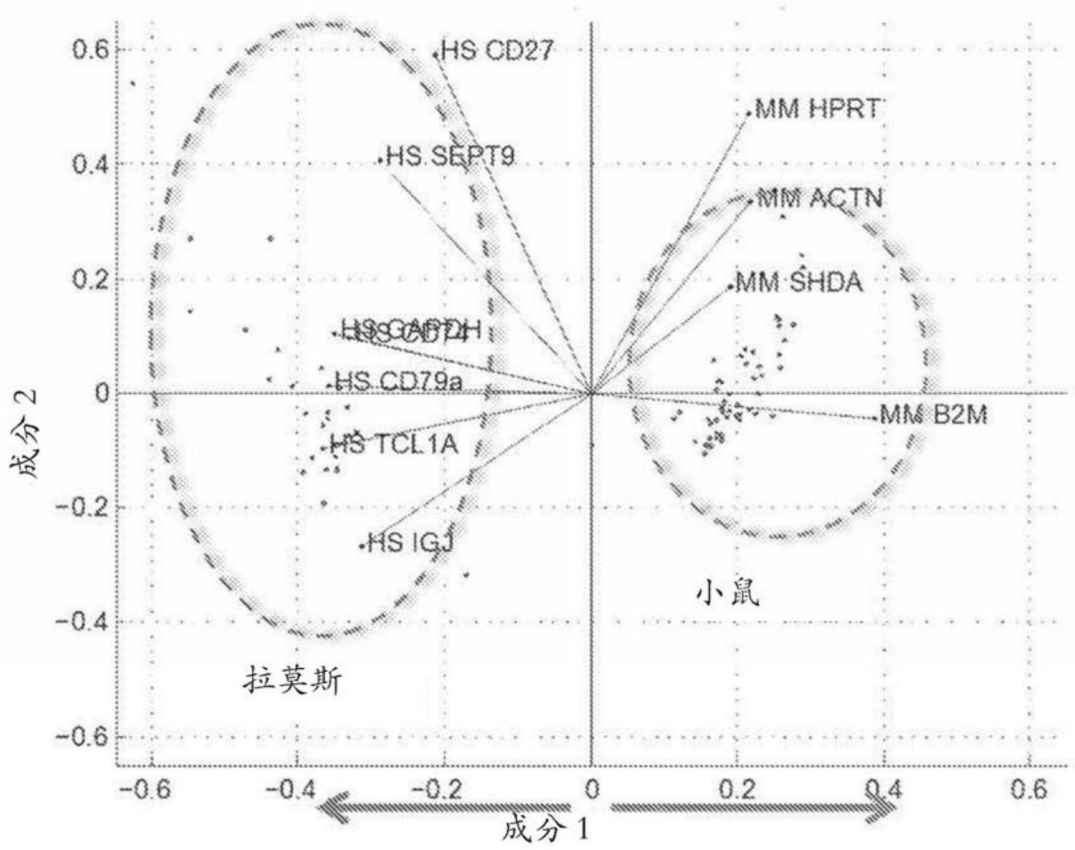

HS CD27
MM HPRT
HS SEPT9
MM ACTN
MM SHDA
HS GAPDH
HS CD74
HS CD79a
MM B2M
HS TCL1A
HS IGJ
小鼠
拉莫斯
成分 2
成分 1
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.6
-0.4
-0.2
0
0.2
0.4
0.6

图13C

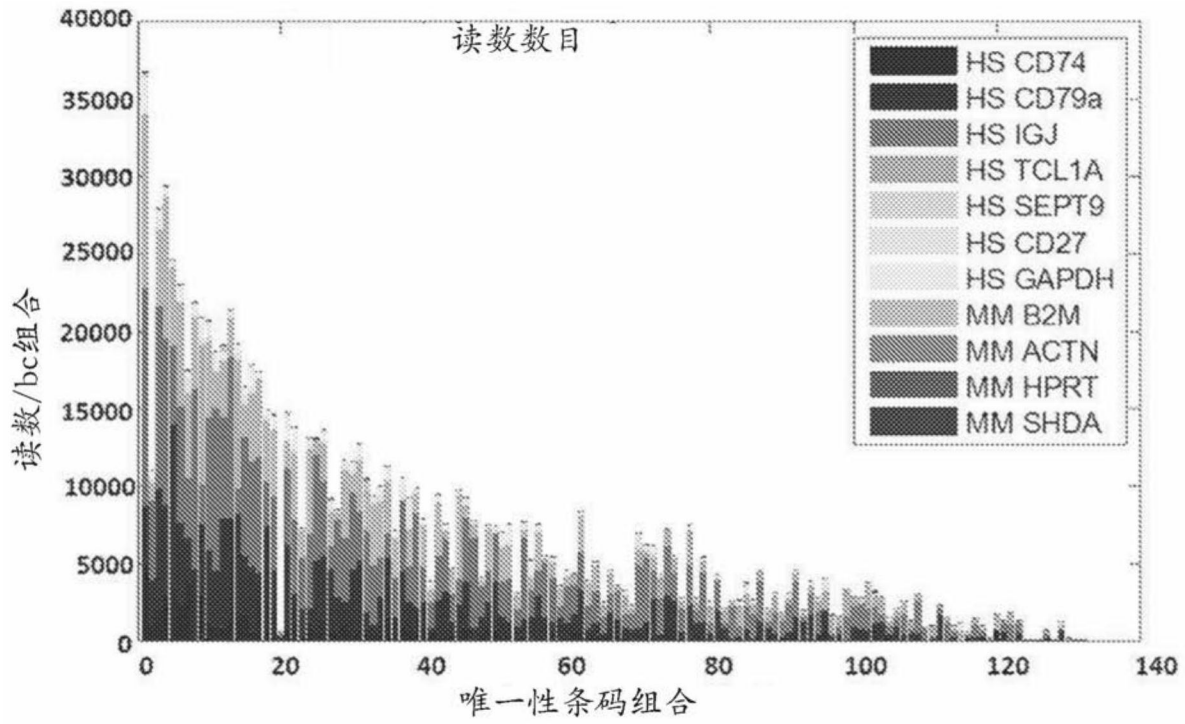

读数数目
HS CD74
HS CD79a
HS IGJ
HS TCL1A
HS SEPT9
HS CD27
HS GAPDH
MM B2M
MM ACTN
MM HPRT
MM SHDA
读数/bc组合
唯一性条码组合
40000
35000
30000
25000
20000
15000
10000
5000
0
0
20
40
60
80
100
120
140

图13D

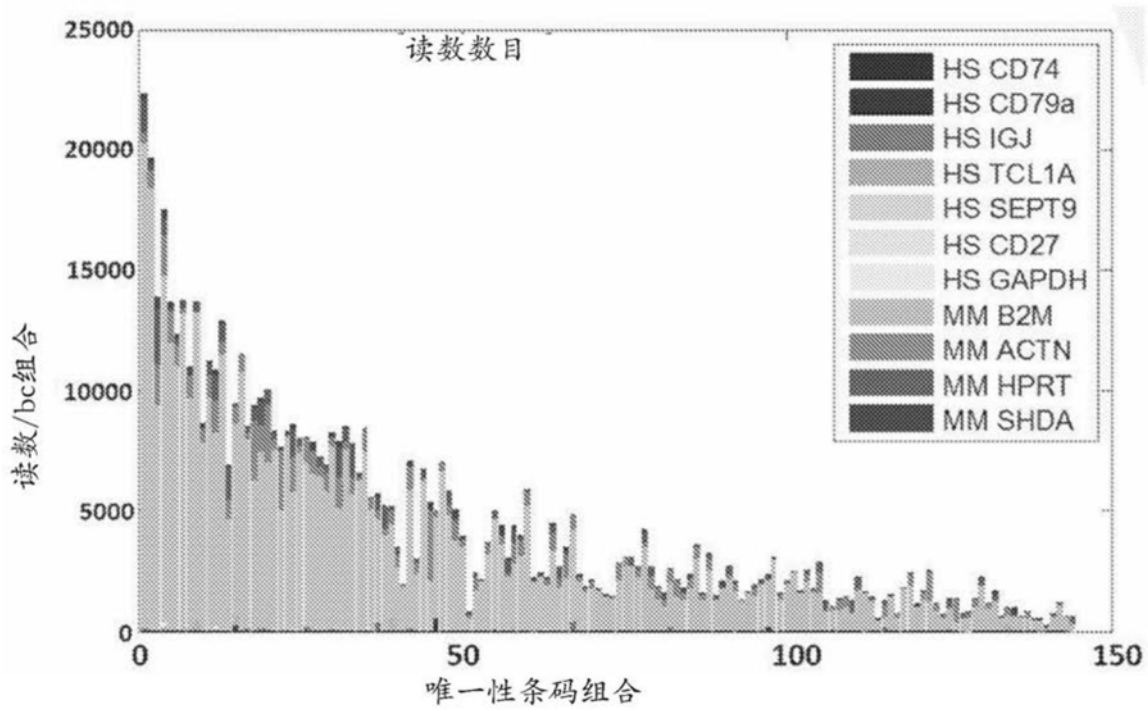

读数数目
读数/bc组合
唯一性条码组合
25000
20000
15000
10000
5000
0
0
50
100
150
HS CD74
HS CD79a
HS IGJ
HS TCL1A
HS SEPT9
HS CD27
HS GAPDH
MM B2M
MM ACTN
MM HPRT
MM SHDA

图13E

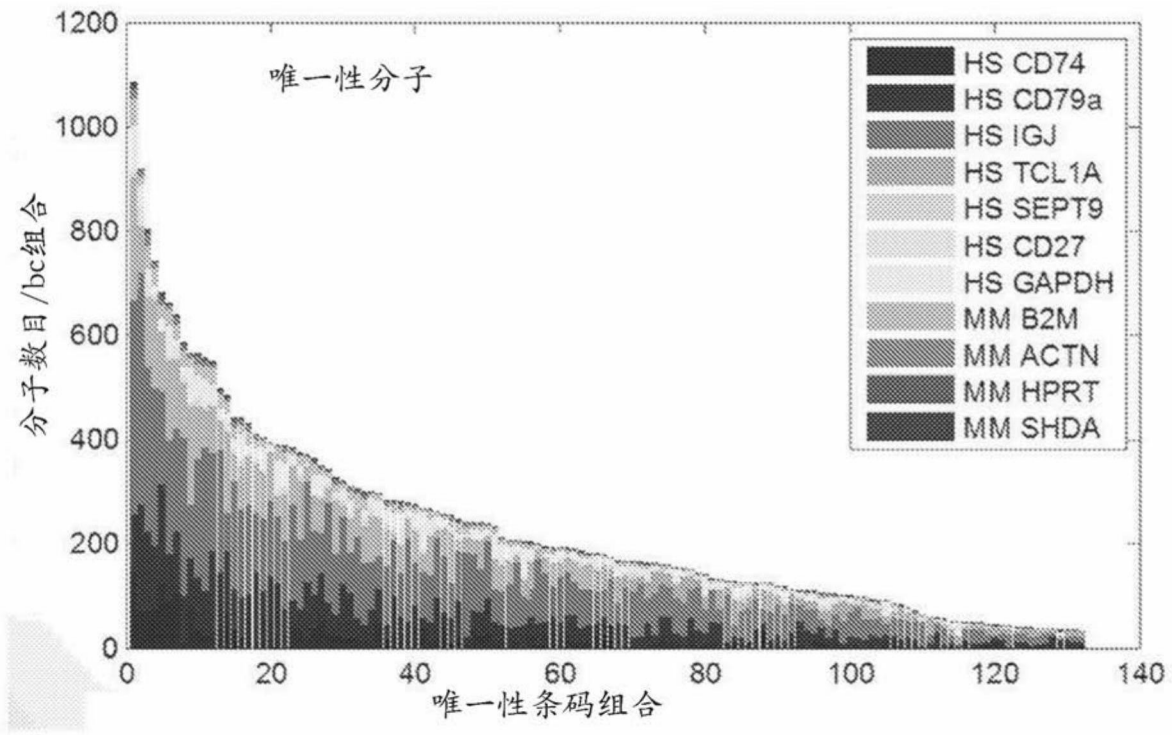

唯一性分子
分子数目/bc组合
唯一性条码组合
1200
1000
800
600
400
200
0
0
20
40
60
80
100
120
140
HS CD74
HS CD79a
HS IGJ
HS TCL1A
HS SEPT9
HS CD27
HS GAPDH
MM B2M
MM ACTN
MM HPRT
MM SHDA

图13F

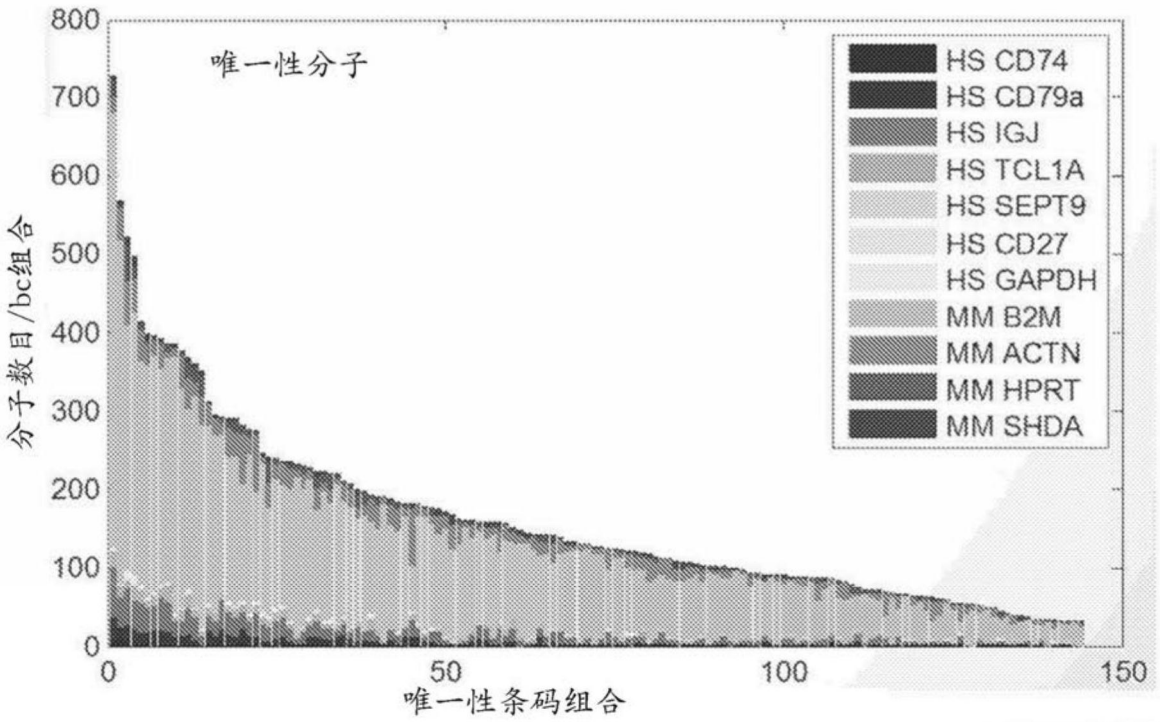

唯一性分子
HS CD74
HS CD79a
HS IGJ
HS TCL1A
HS SEPT9
HS CD27
HS GAPDH
MM B2M
MM ACTN
MM HPRT
MM SHDA
分子数目/bc组合
唯一性条码组合
800
700
600
500
400
300
200
100
0
0
50
100
150

图13G

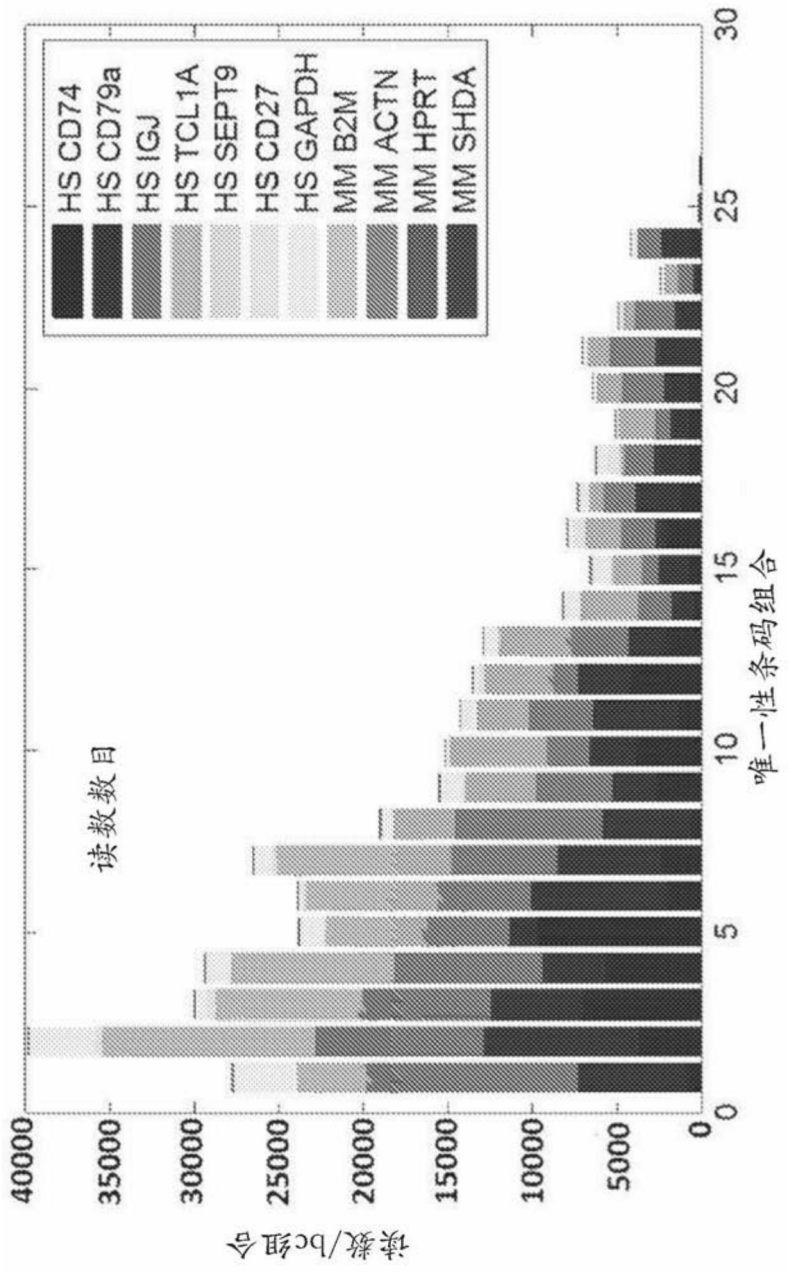

HS CD74
HS CD79a
HS IGJ
HS TCL1A
HS SEPT9
HS CD27
HS GAPDH
MM B2M
MM ACTN
MM HPRT
MM SHDA
读数数目
唯一性条码组合
读数/bc组合
0
5000
10000
15000
20000
25000
30000
35000
40000
0
5
10
15
20
25
30

图13H

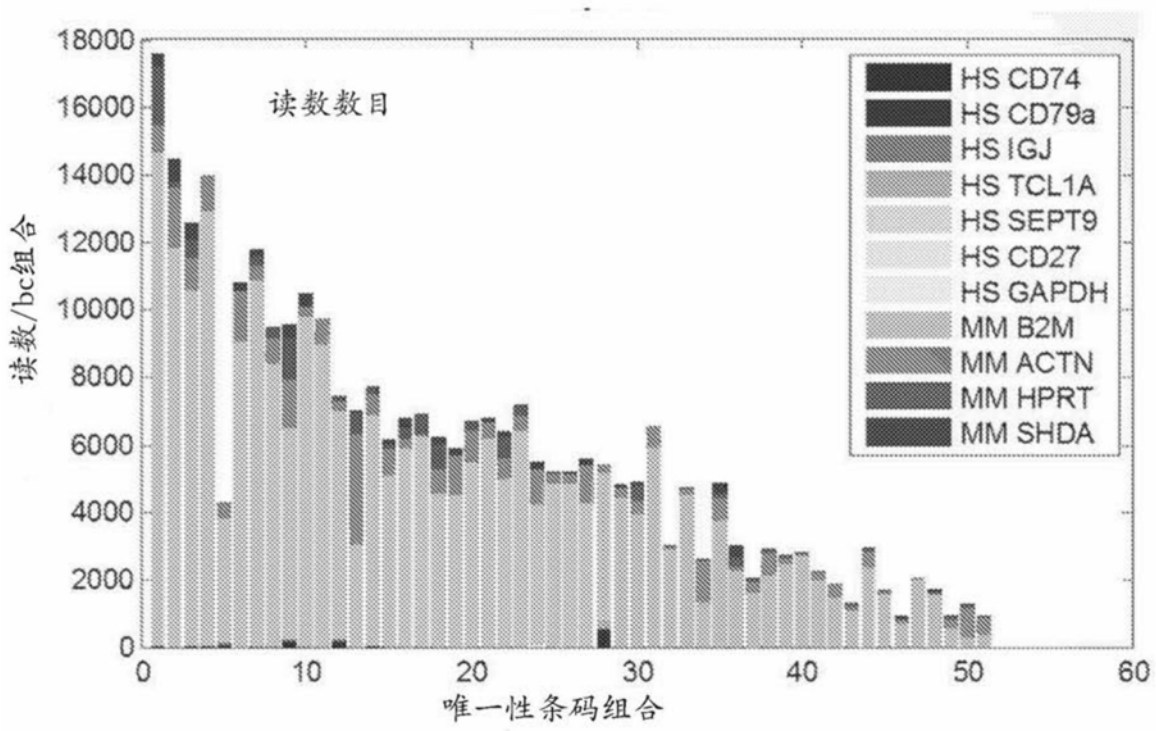

读数数目
读数/bc组合
唯一性条码组合
HS CD74
HS CD79a
HS IGJ
HS TCL1A
HS SEPT9
HS CD27
HS GAPDH
MM B2M
MM ACTN
MM HPRT
MM SHDA
0
2000
4000
6000
8000
10000
12000
14000
16000
18000
0
10
20
30
40
50
60

图13I

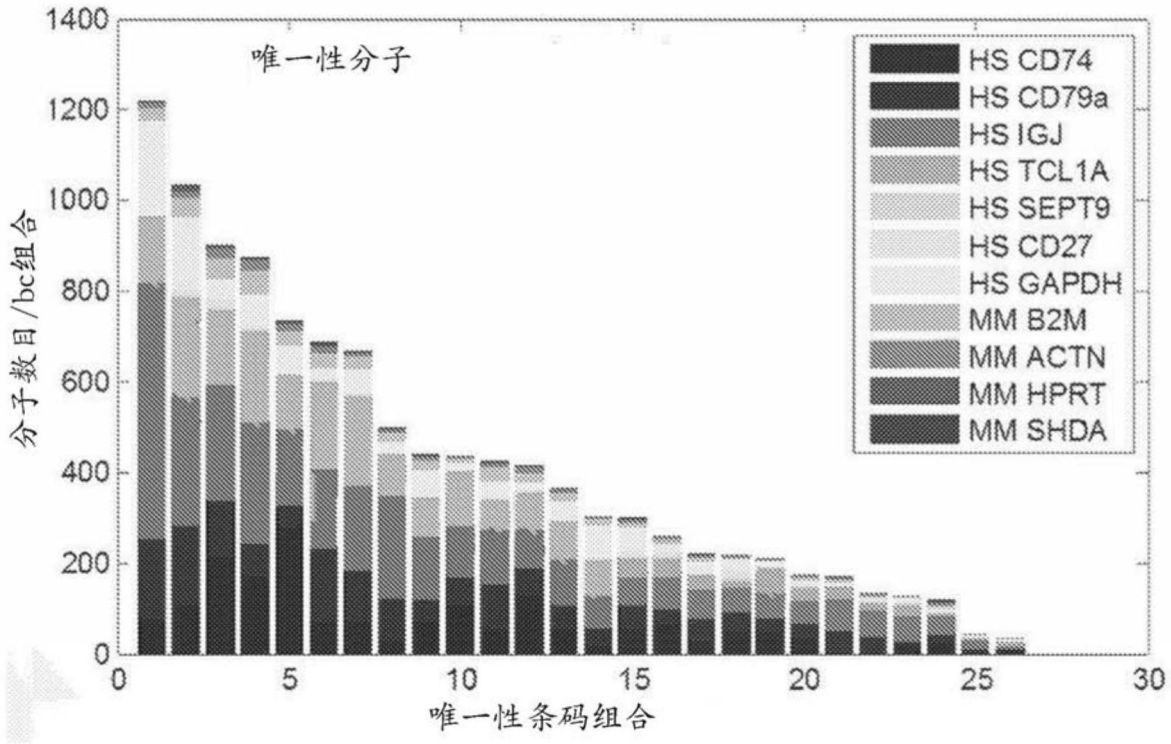

唯一性分子
分子数目/bc组合
唯一性条码组合
HS CD74
HS CD79a
HS IGJ
HS TCL1A
HS SEPT9
HS CD27
HS GAPDH
MM B2M
MM ACTN
MM HPRT
MM SHDA
0
200
400
600
800
1000
1200
1400
0
5
10
15
20
25
30

图13J

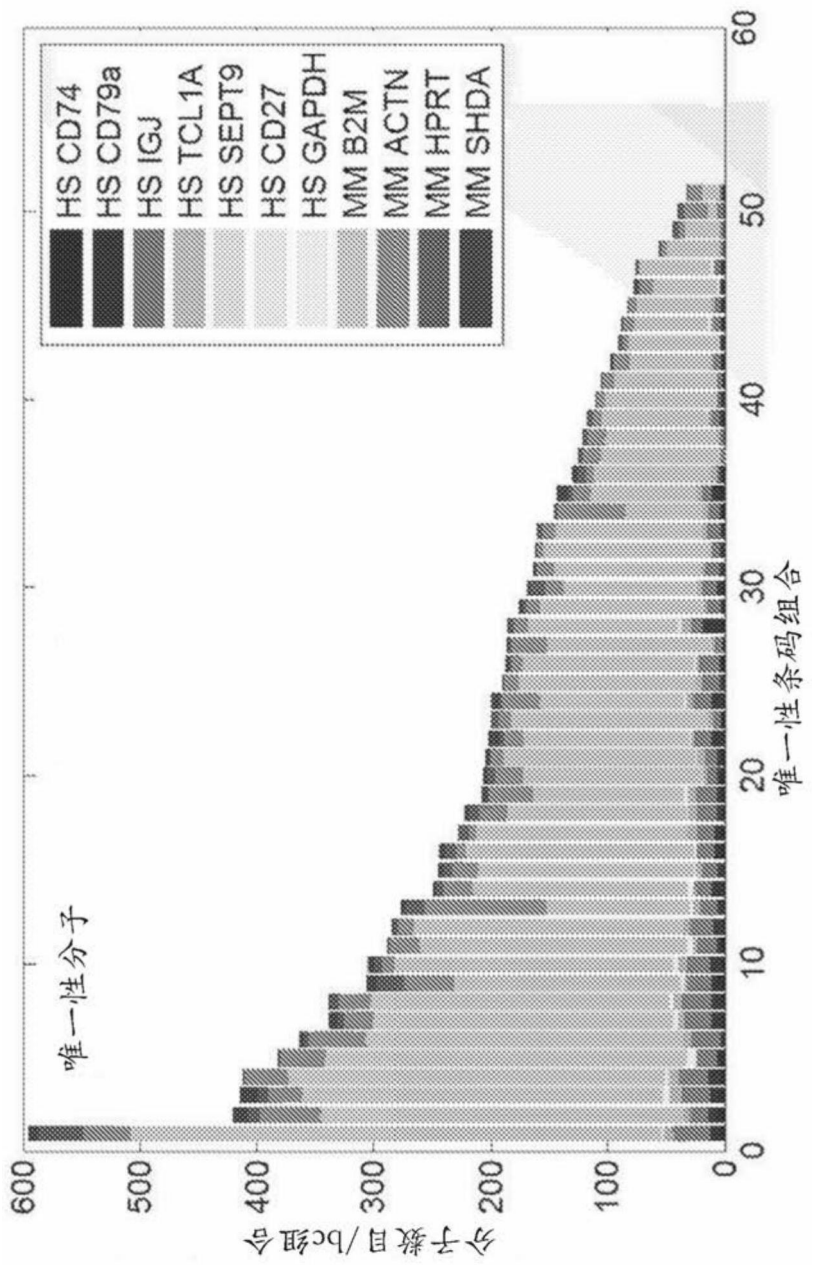

HS CD74
HS CD79a
HS IGJ
HS TCL1A
HS SEPT9
HS CD27
HS GAPDH
MM B2M
MM ACTN
MM HPRT
MM SHDA
唯一性分子
唯一性条码组合
分子数目/bc组分
0
100
200
300
400
500
600
10
20
30
40
50
60

图13K


log10(读数的数目)
0
1
2
3
4
5
6
CD30
CD69
CD19
CD20
BAFF
TCL1A
TACI
PAX5
TNFRSF17
CD27
CD38
CD24
AICDA
CD95
IGHD_CH2
IGHM_CH4
IL10
CD10
BCL6
PRDM1
IXK
MBD2
CCR7
S100A12
CCR2varA
CCR2varB
CD62L
CD11b
FOXQ1
CD209
CD1c
CCL17
DTNA
CD8A
CD8B
PRF1
EOMES
PDCD1
LAG3
CXCR5
OX40L
ICOS
SH2D1A
CD4
GAPDH
CD44
CD45RO
CD14
CD45RA
RUNX3
THPOK
OSBPL5
CD56
IGFBP7
KIR2DS5
KIR2DS2
RAB4B
ZBTB16
SLAMF1
CD16a
CX3CR1
CD138
RASD1
AMPD1
CLEC4C
CD3D
CD3E
CXCR3
IFNG
IFNGR1
IL12RB2
CCL20
IL17A
IL17F
IL21
IL22
IL23R
RORA
RORC
IL4R
IL4
IL2RB
CCR4
CRTH2
IL5
TLR1
TLR2
TLR3
TLR4
TLR5
TLR6
TLR7
TLR8
TLR9
TLR10
CD25
FOXP3
TGFB1

图14

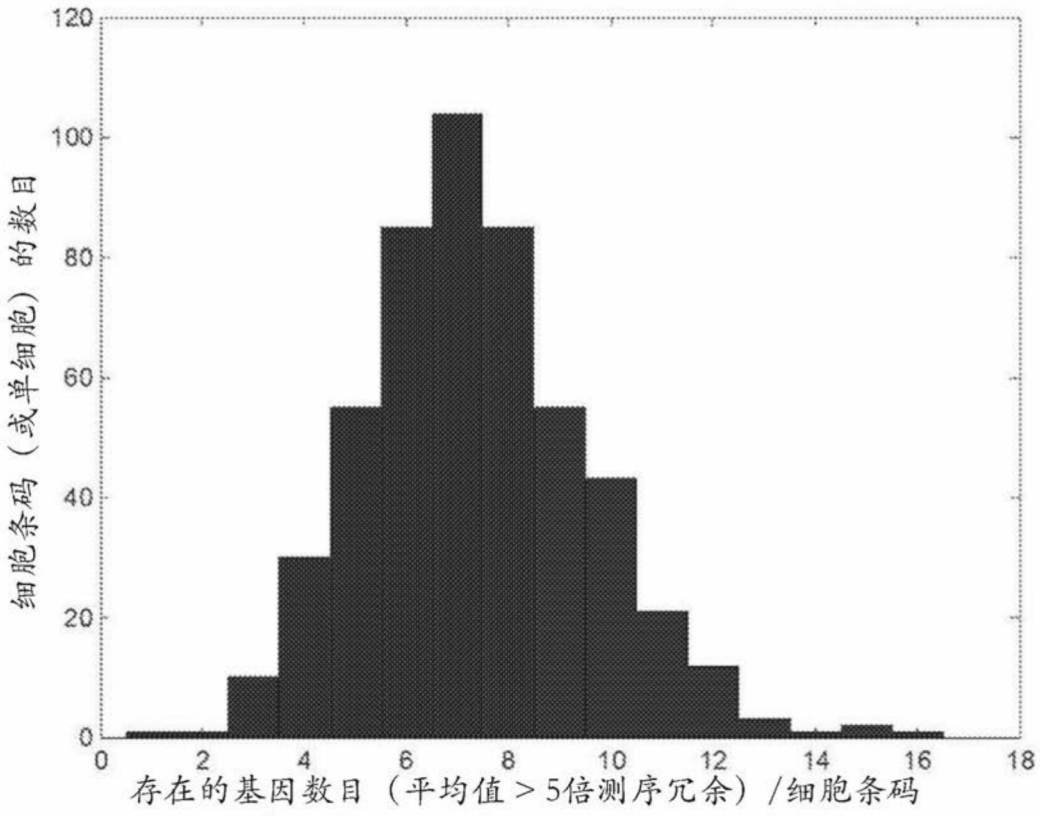

120
100
80
60
40
20
0
细胞条码（或单细胞）的数目
0
2
4
6
8
10
12
14
16
18
存在的基因数目（平均值＞5倍测序冗余）/细胞条码

图15A

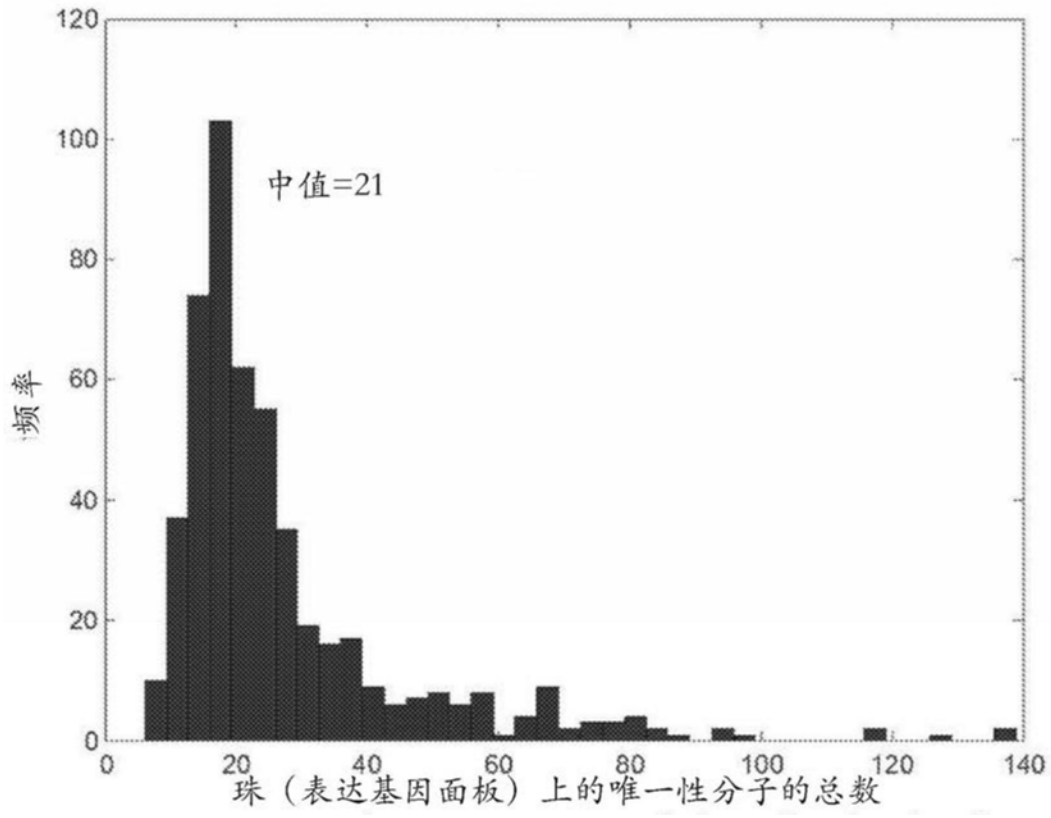

120
100
80
60
40
20
0
频率
中值=21
0
20
40
60
80
100
120
140
珠（表达基因面板）上的唯一性分子的总数

图15B

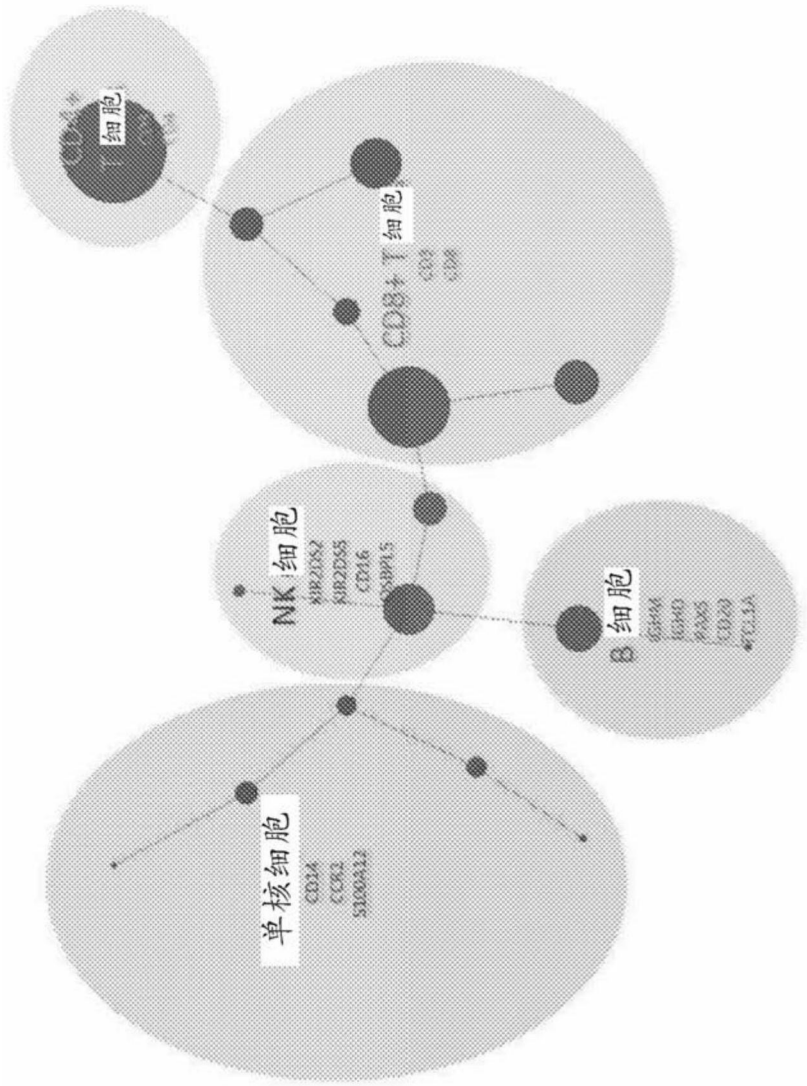

细胞
CD8+ T 细胞
CD3
CD8
NK 细胞
KIR2DL2
KIR2DS5
CD16
CXCR1S
B 细胞
CD19
CD20
单核细胞
CD14
CCR2
S100A12

图16

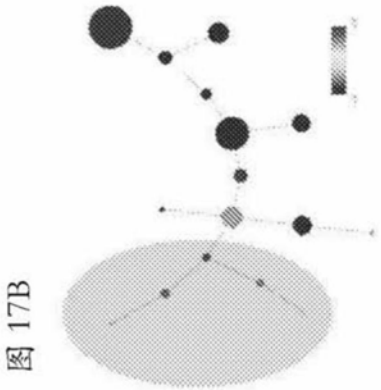


图 17B

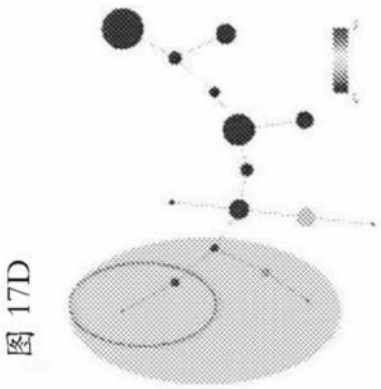


图 17D

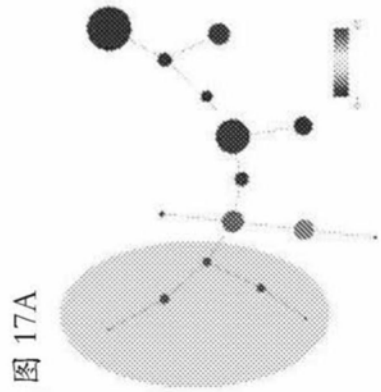


图 17A

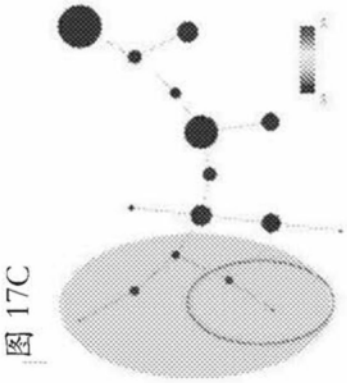


图 17C

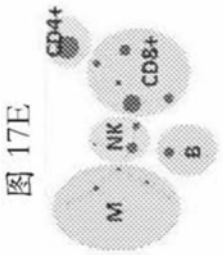


图 17E

图 18A

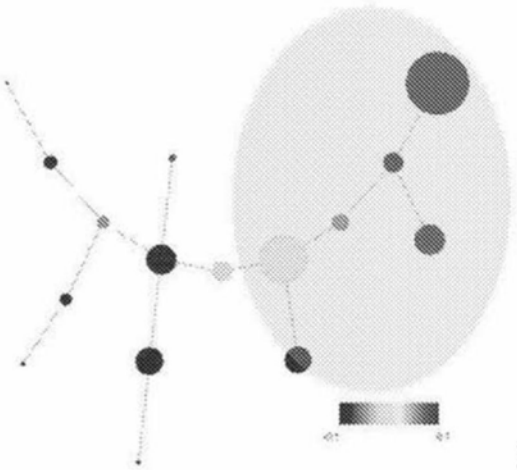


图 18B

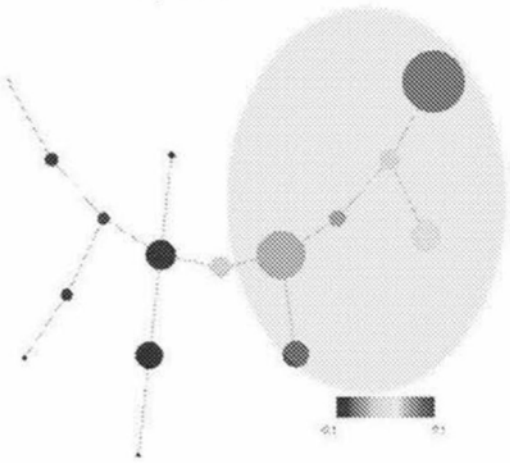


图 18C

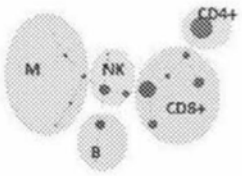


图 19A

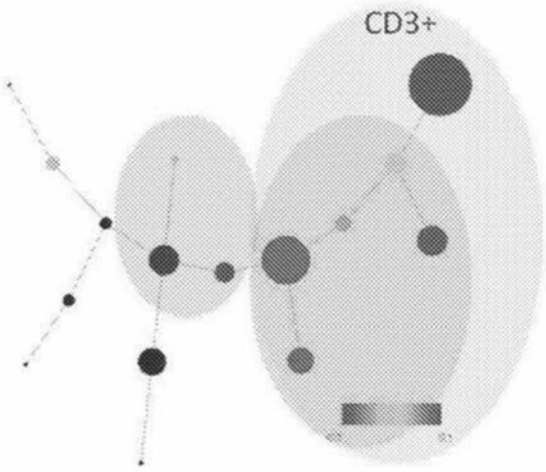


图 19B

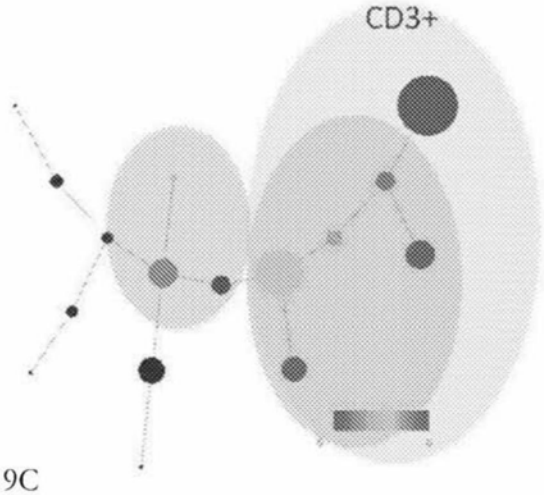


图 19C

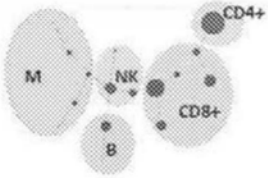


图 20A

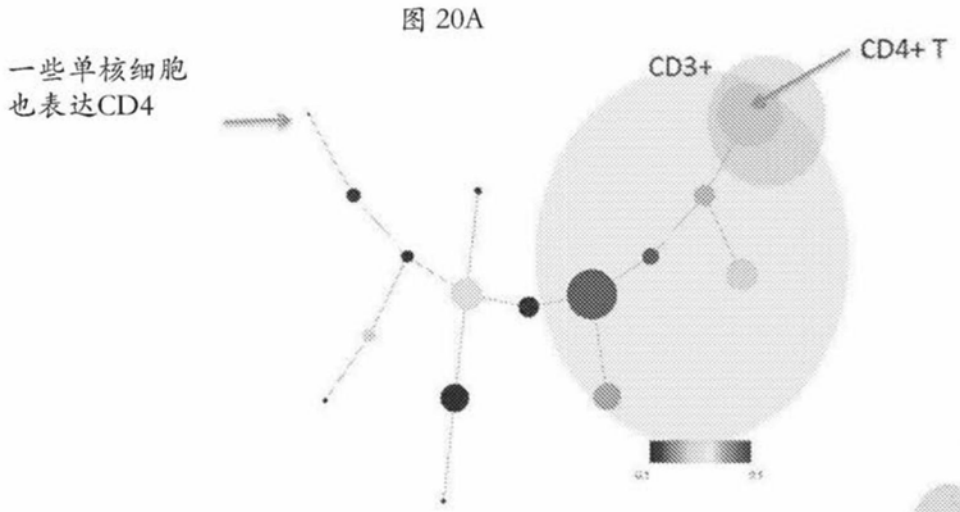


图 20B

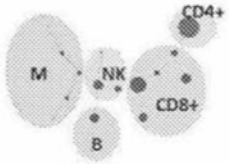


图 21A

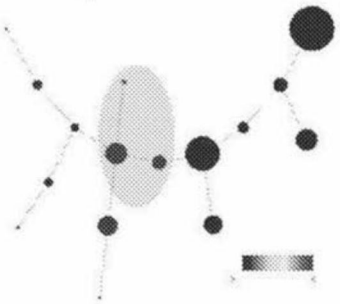


图 21B

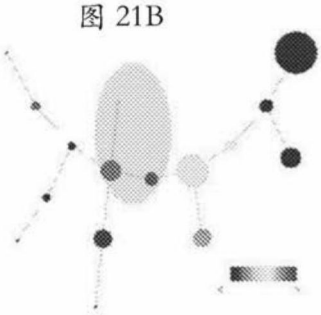


图 21C

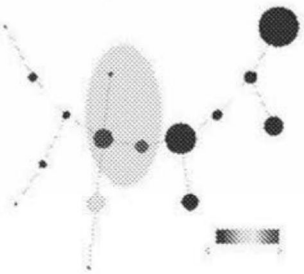


图 21D

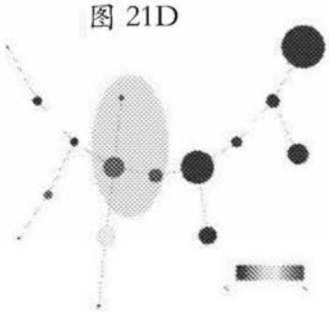


图 21E

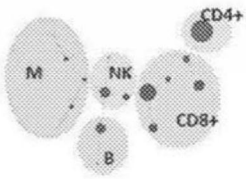


图 22A

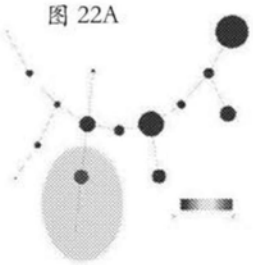


图 22B

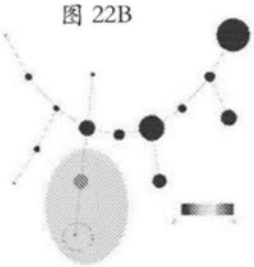


图 22C

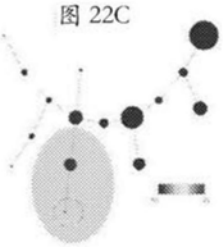


图22D

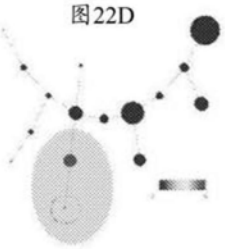


图22E

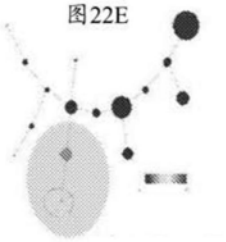


图 22F

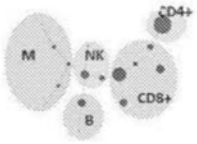


图 23A

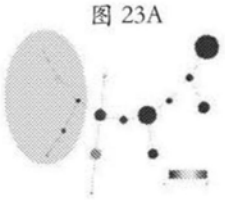


图 23B

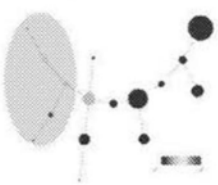


图 23C

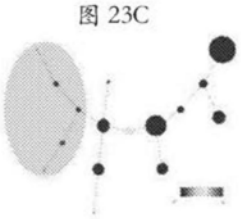


图 23D

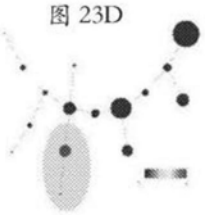


图 23E

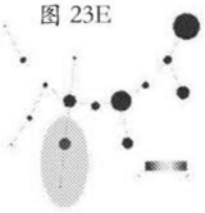


图 23F

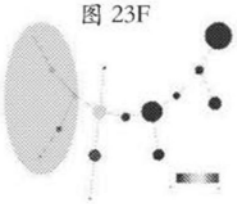
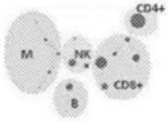


图 23G



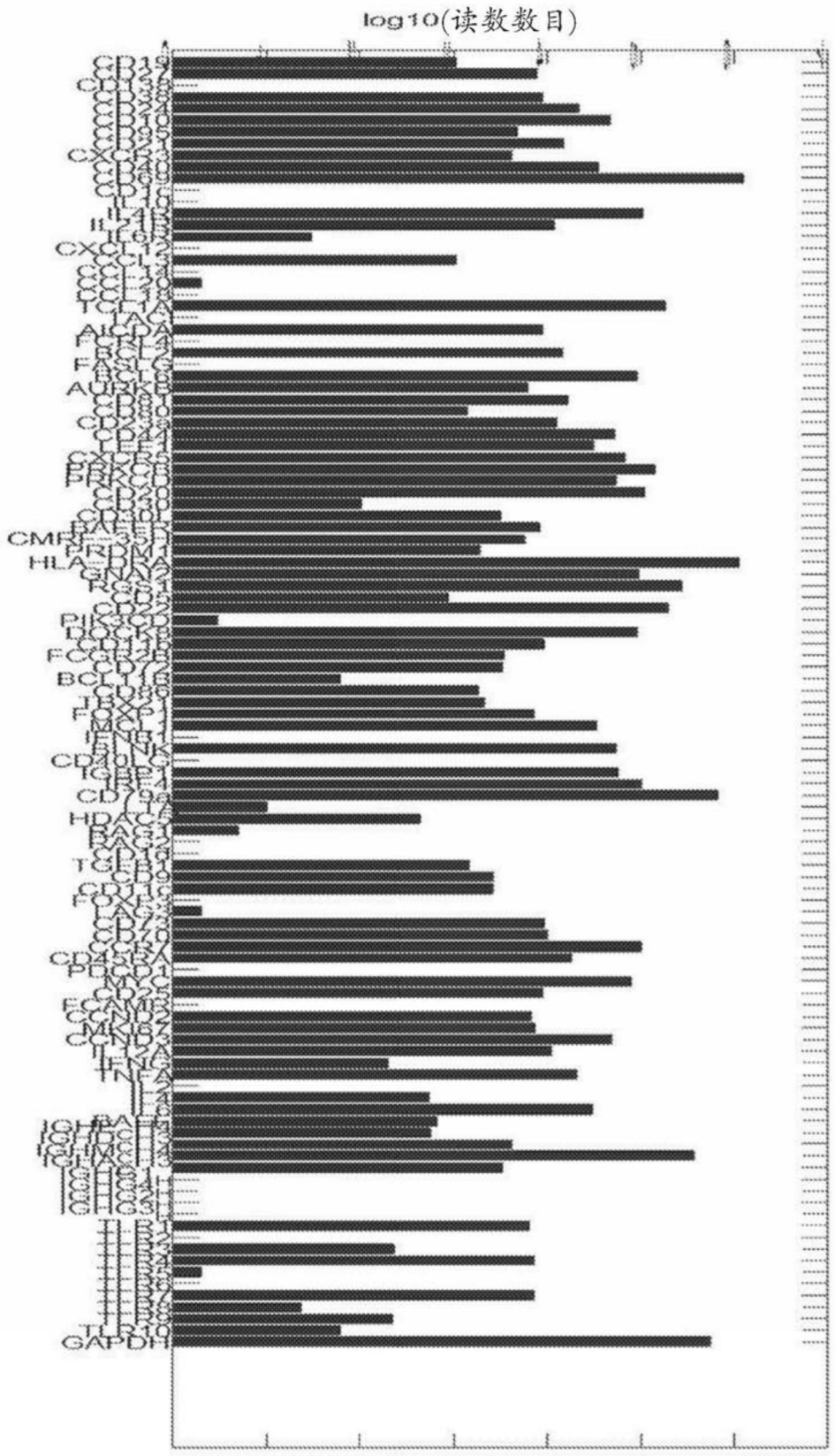


图24

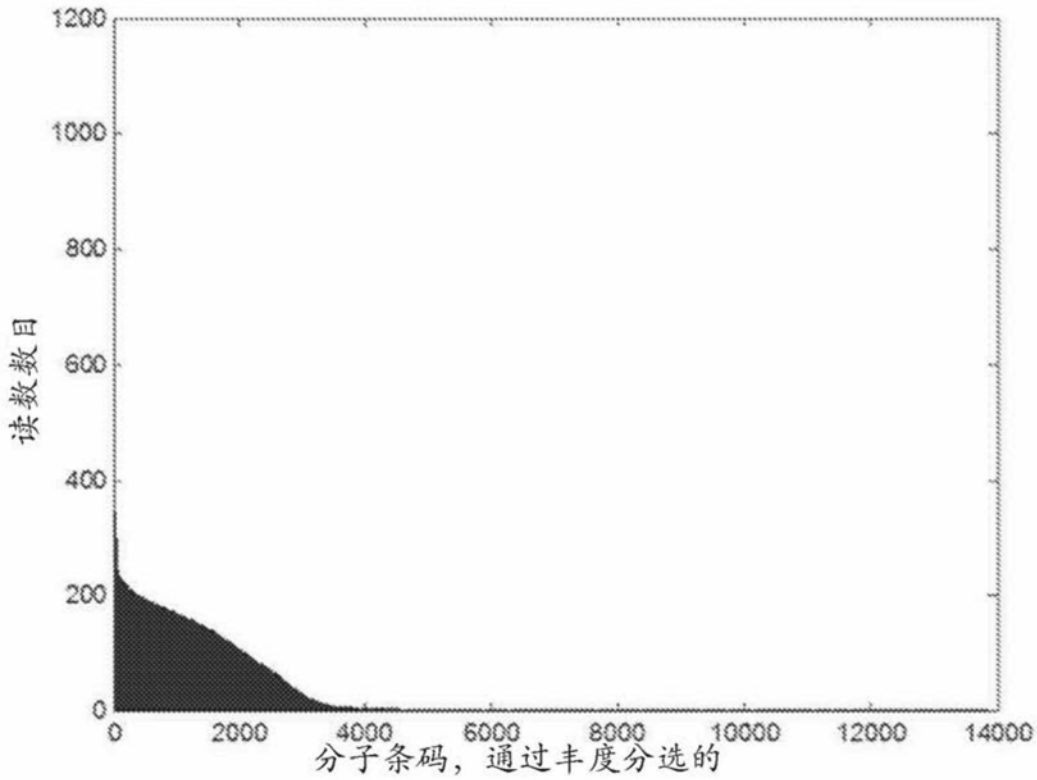

1200
1000
800
600
400
200
0
读数数目
0
2000
4000
6000
8000
10000
12000
14000
分子条码，通过丰度分选的

图25A

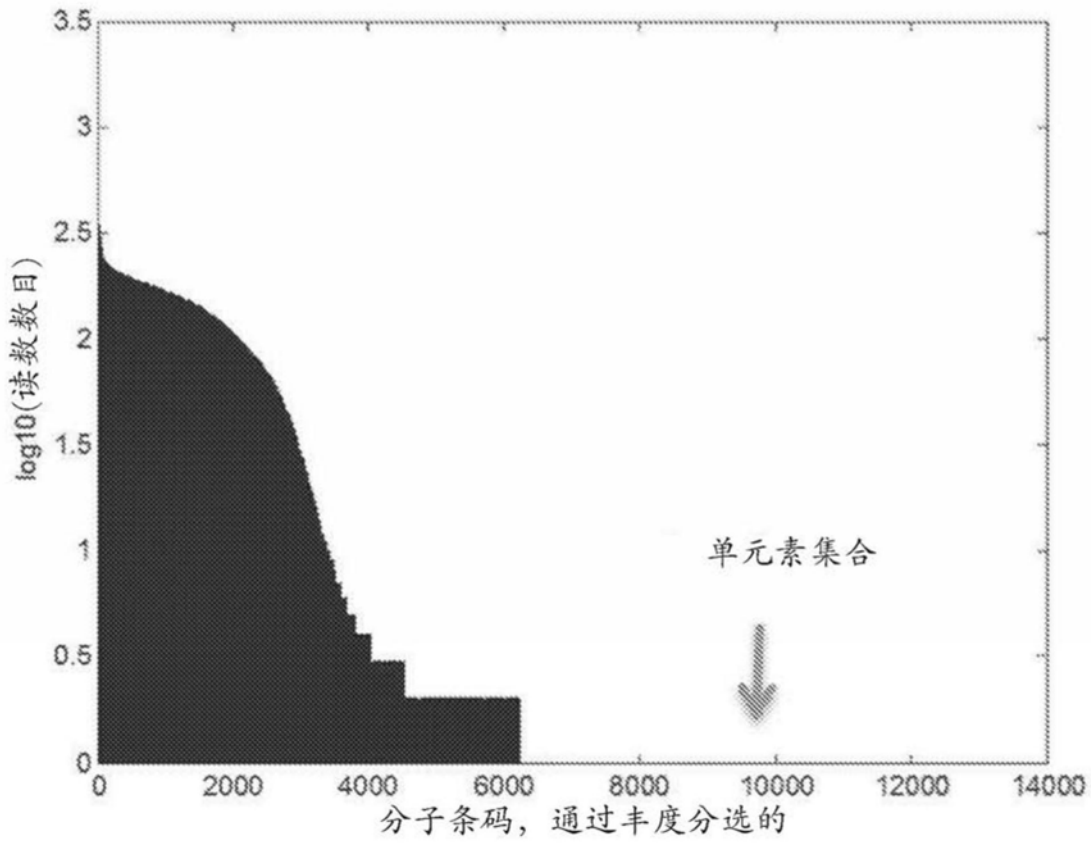

3.5
3
2.5
2
1.5
1
0.5
0
log10(读数数目)
0
2000
4000
6000
8000
10000
12000
14000
单元素集合
分子条码，通过丰度分选的

图25B

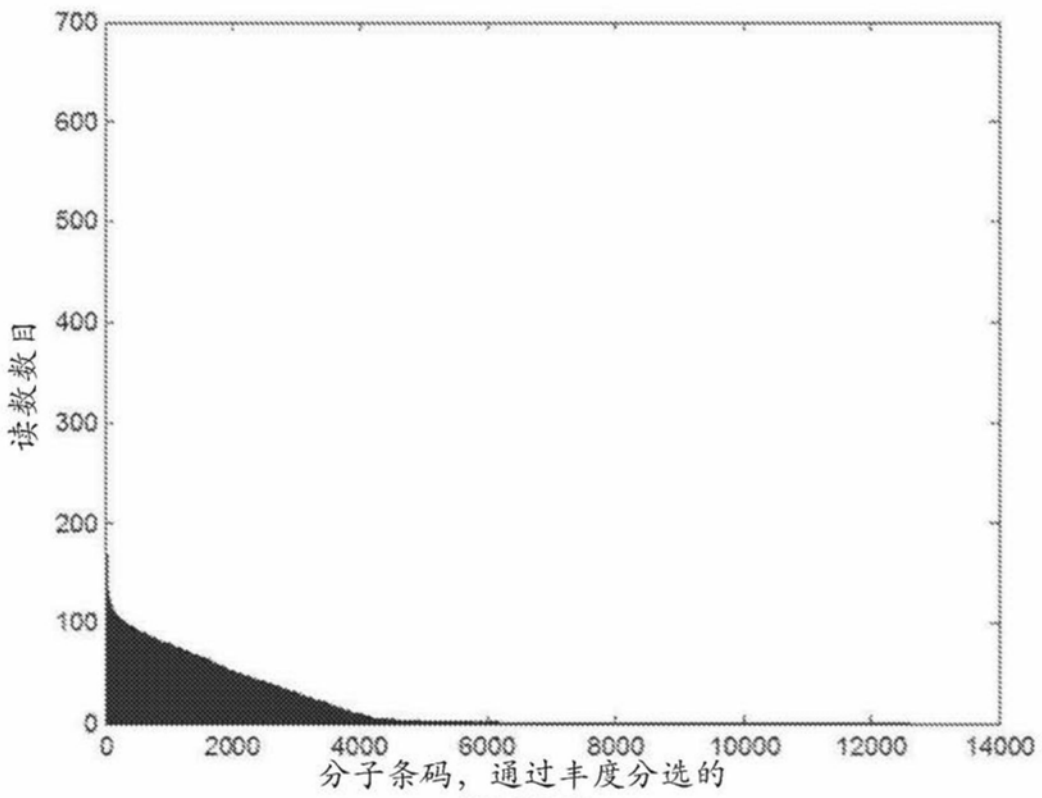

700
600
500
400
300
200
100
0
读数数目
0
2000
4000
6000
8000
10000
12000
14000
分子条码，通过丰度分选的

图25C

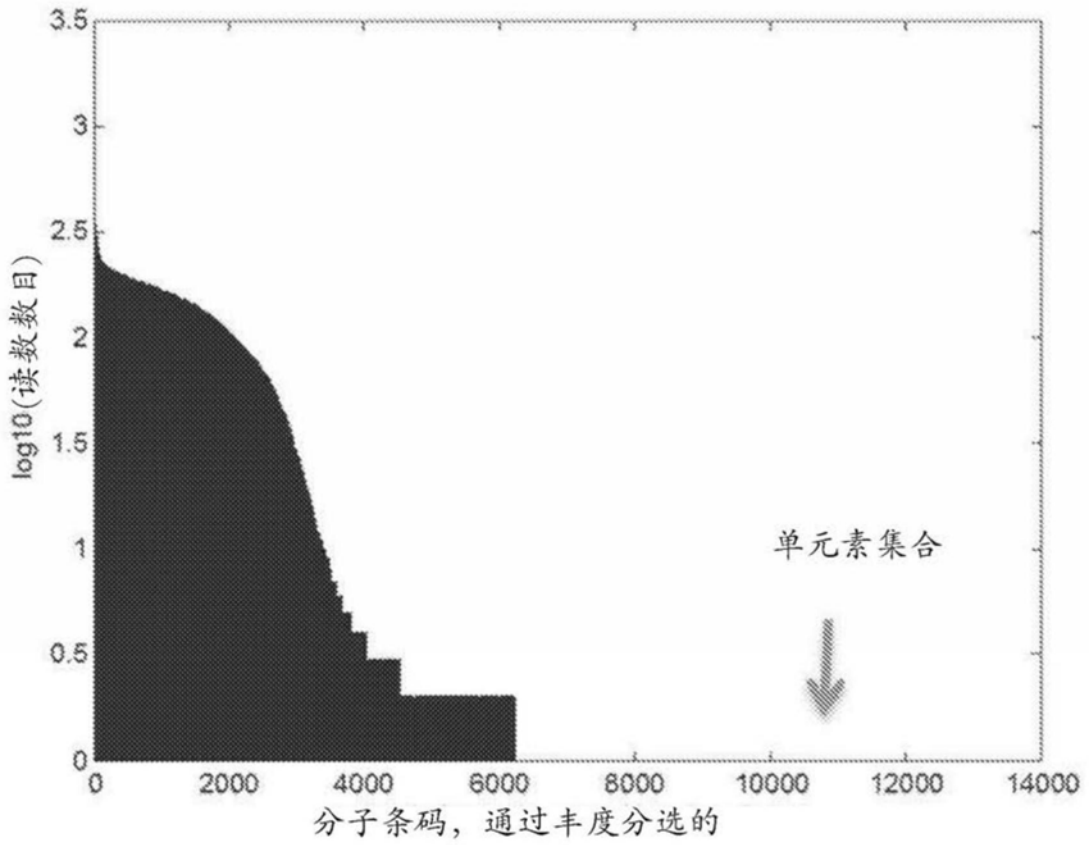

log10(读数数目)
单元素集合
分子条码，通过丰度分选的

图25D

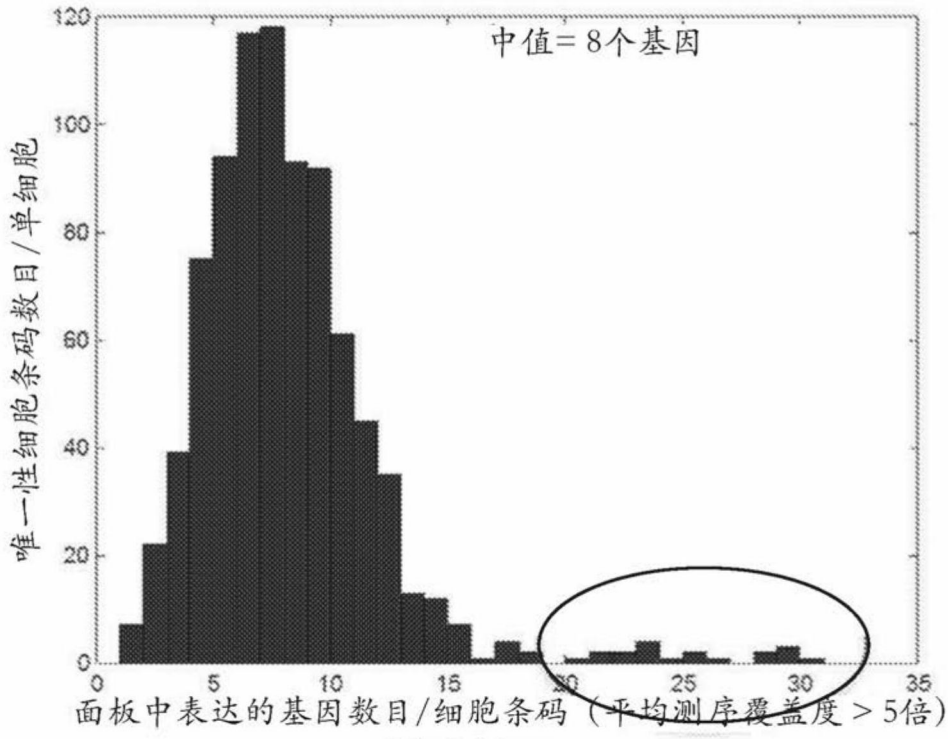

中值= 8个基因
唯一性细胞条码数目/单细胞
面板中表达的基因数目/细胞条码（平均测序覆盖度 > 5倍）

图26A

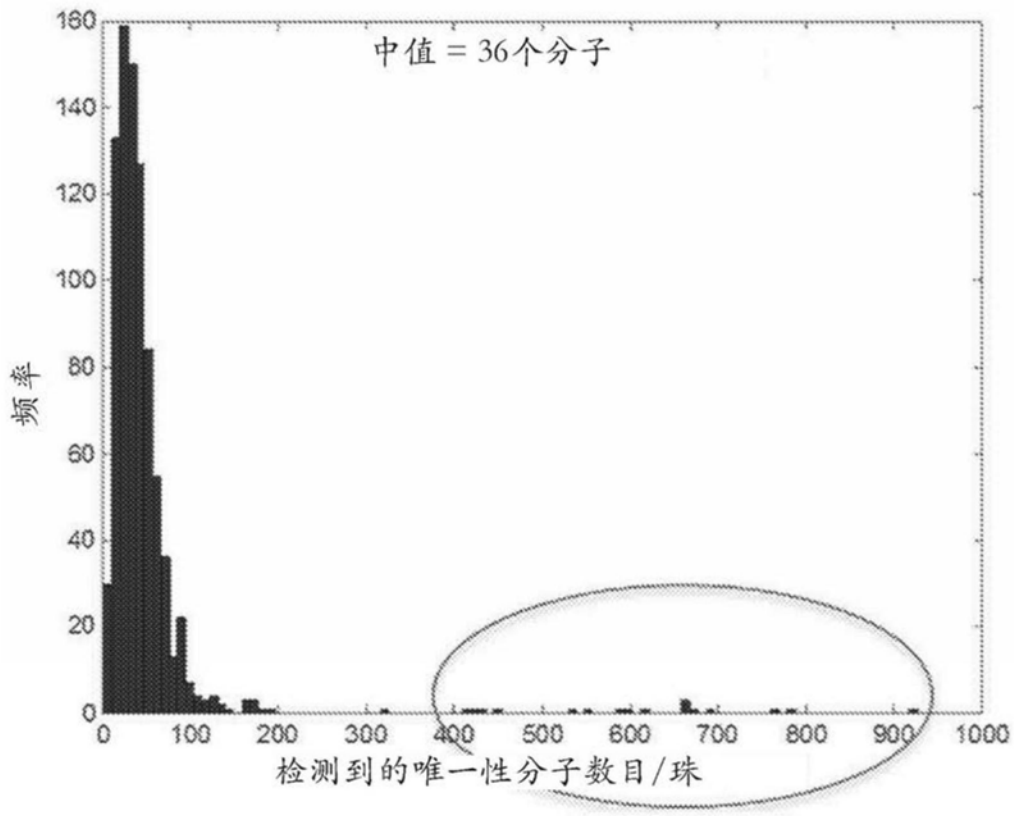

中值 = 36个分子
频率
检测到的唯一性分子数目/珠
0
20
40
60
80
100
120
140
160
0
100
200
300
400
500
600
700
800
900
1000

图26B

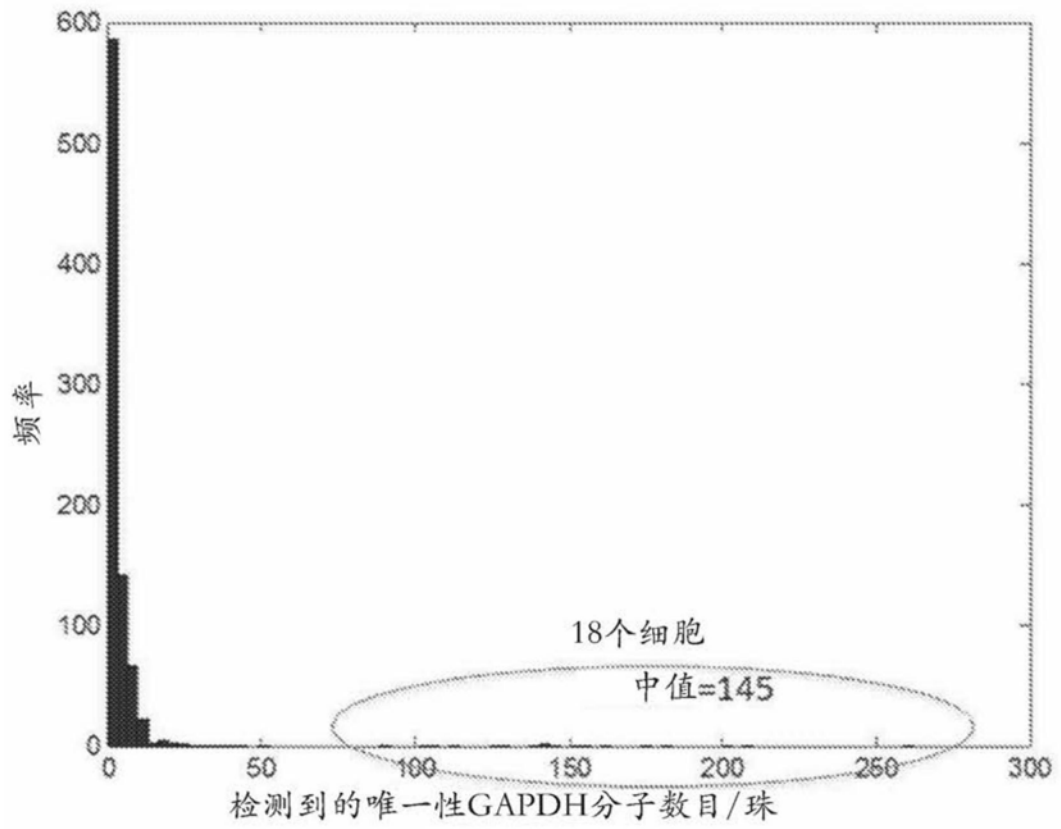

600
500
400
300
200
100
0
频率
18个细胞
中值=145
0
50
100
150
200
250
300
检测到的唯一性GAPDH分子数目/珠

图26C

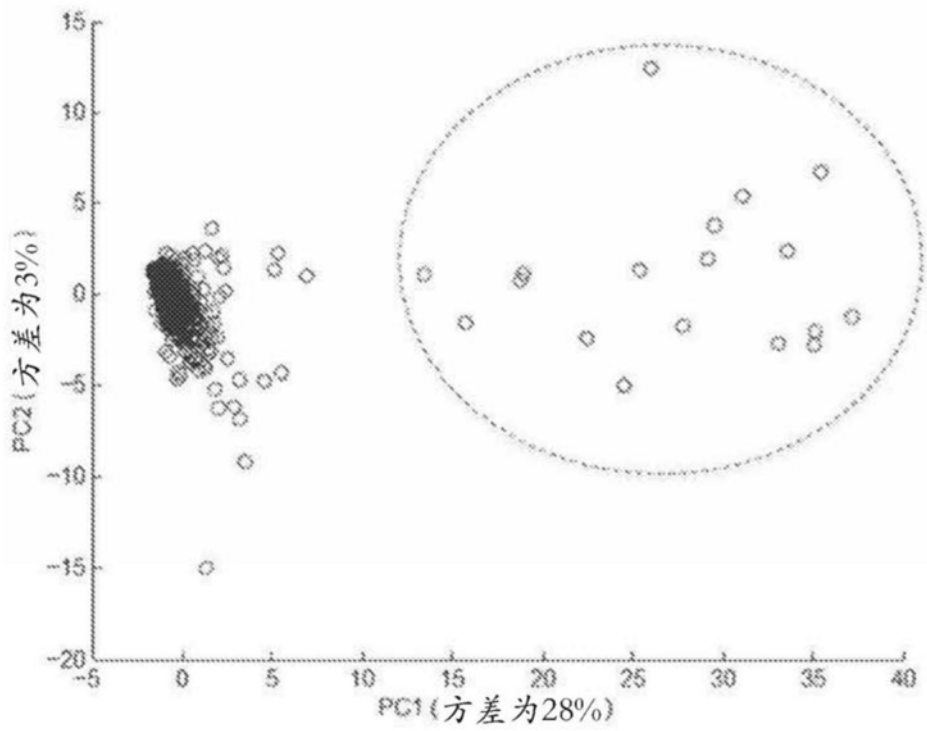

15
10
5
0
-5
-10
-15
-20
PC2(方差为3%)
-5
0
5
10
15
20
25
30
35
40
PC1(方差为28%)

图27

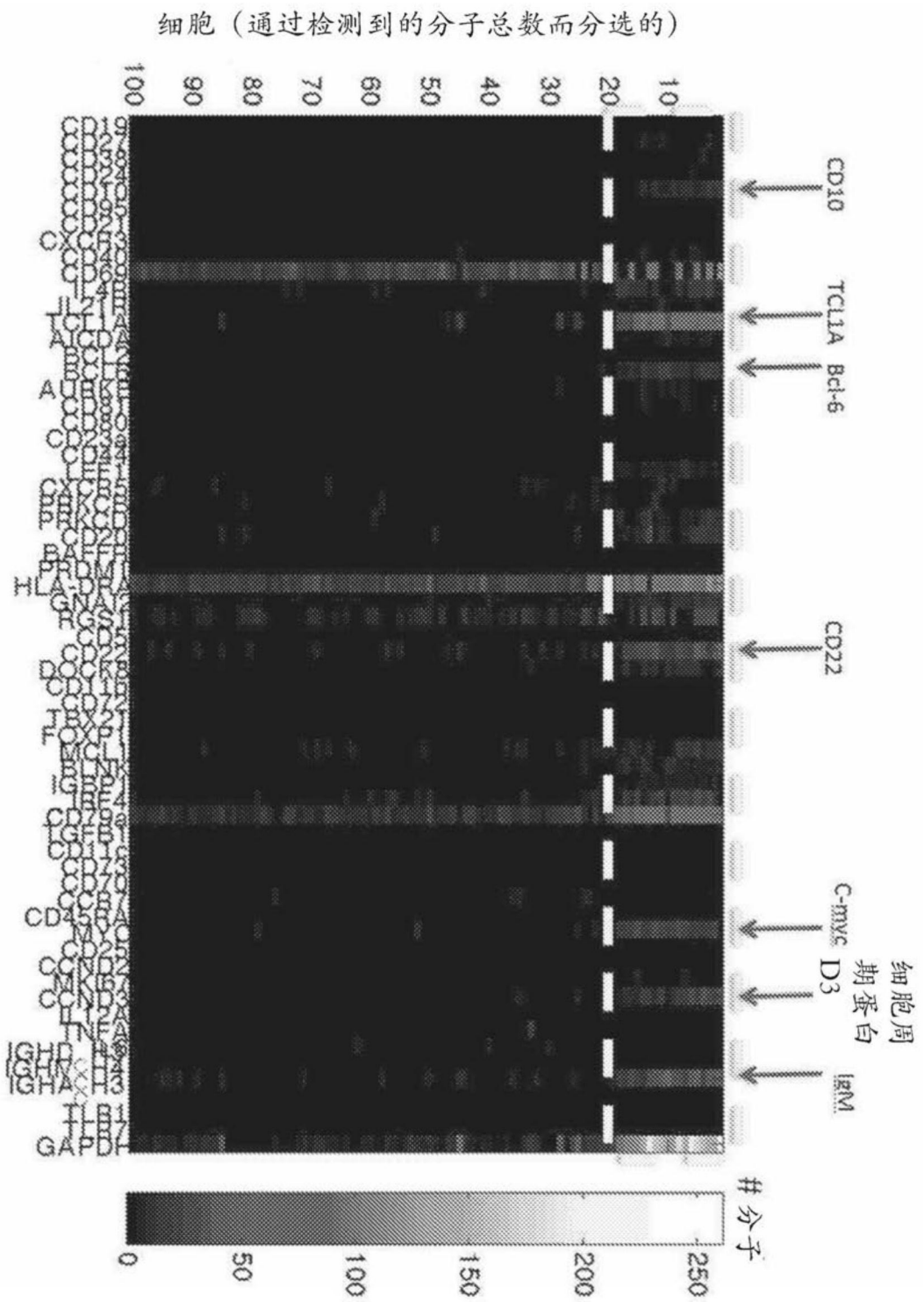

细胞（通过检测到的分子总数而分选的）
100
90
80
70
60
50
40
30
20
10
CD10
TCL1A
Bcl-6
CD22
C-myc
D3
IgM
细胞周期蛋白
分子
0
50
100
150
200
250

图28

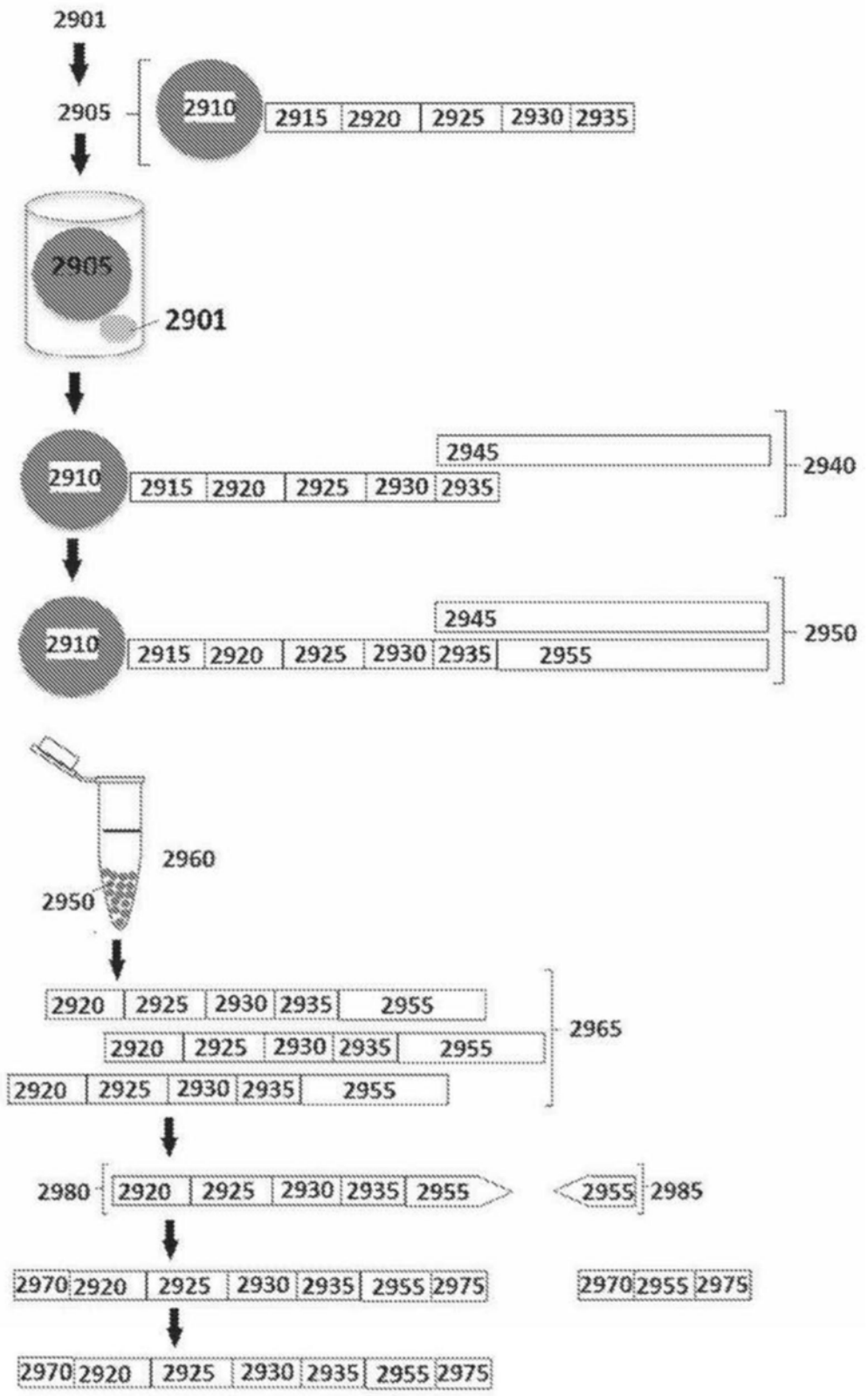

2901
2905
2910
2915 2920 2925 2930 2935
2905
2901
2910
2945
2915 2920 2925 2930 2935
2940
2910
2945
2915 2920 2925 2930 2935 2955
2950
2960
2950
2920 2925 2930 2935 2955
2920 2925 2930 2935 2955
2920 2925 2930 2935 2955
2965
2980
2920 2925 2930 2935 2955
2955
2985
2970 2920 2925 2930 2935 2955 2975
2970 2955 2975
2970 2920 2925 2930 2935 2955 2975

图29

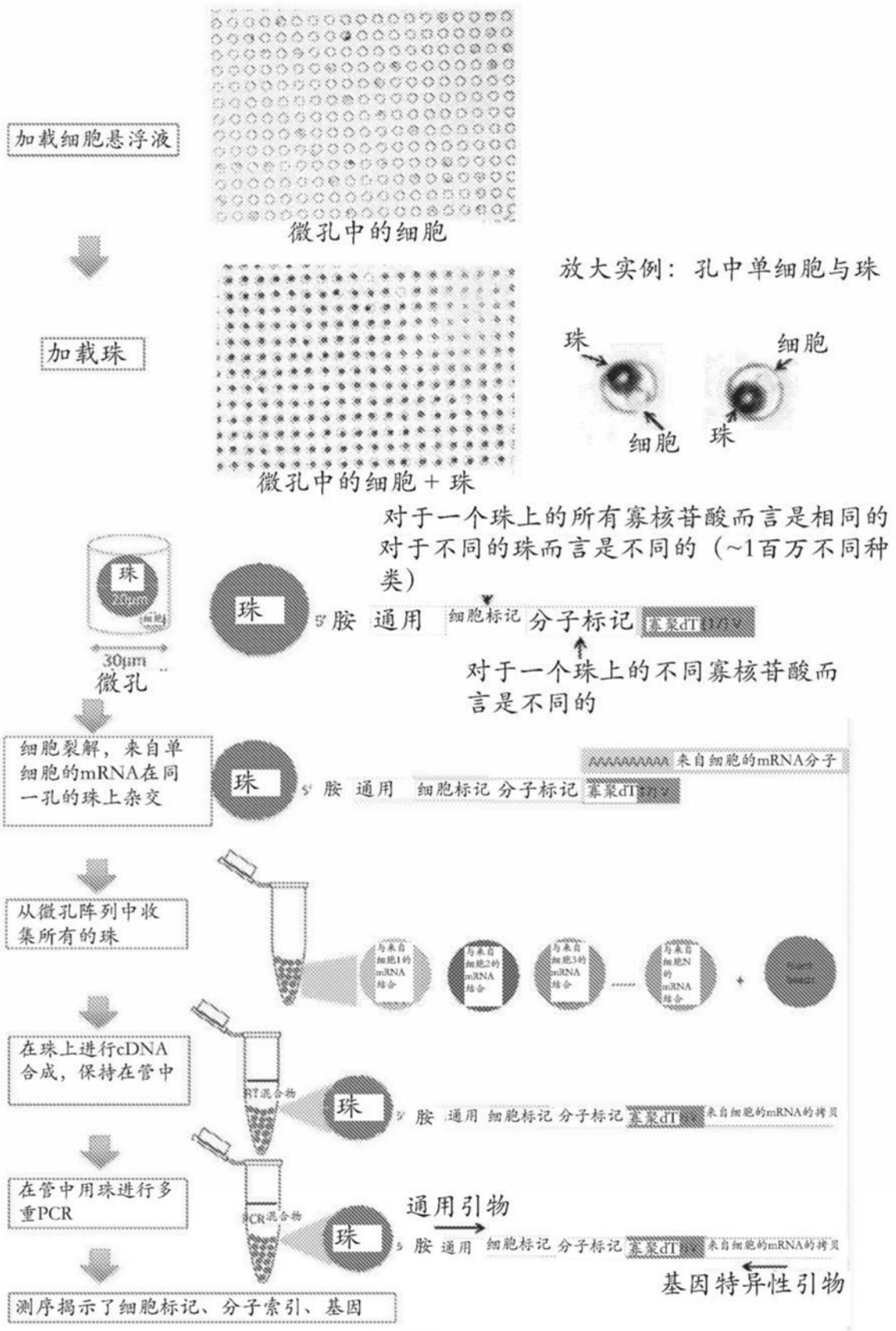

加载细胞悬浮液
微孔中的细胞
加载珠
微孔中的细胞 + 珠
放大实例：孔中单细胞与珠
珠
细胞
细胞
珠
对于一个珠上的所有寡核苷酸而言是相同的
对于不同的珠而言是不同的（~1百万不同种类）
珠
5' 胺 通用 细胞标记 分子标记 寡聚dT
对于一个珠上的不同寡核苷酸而言是不同的
30μm
微孔
细胞裂解，来自单细胞的mRNA在同一孔的珠上杂交
来自细胞的mRNA分子
从微孔阵列中收集所有的珠
在珠上进行cDNA合成，保持在管中
RT混合物
来自细胞的mRNA的拷贝
在管中用珠进行多重PCR
PCR混合物
通用引物
基因特异性引物
测序揭示了细胞标记、分子索引、基因

图30

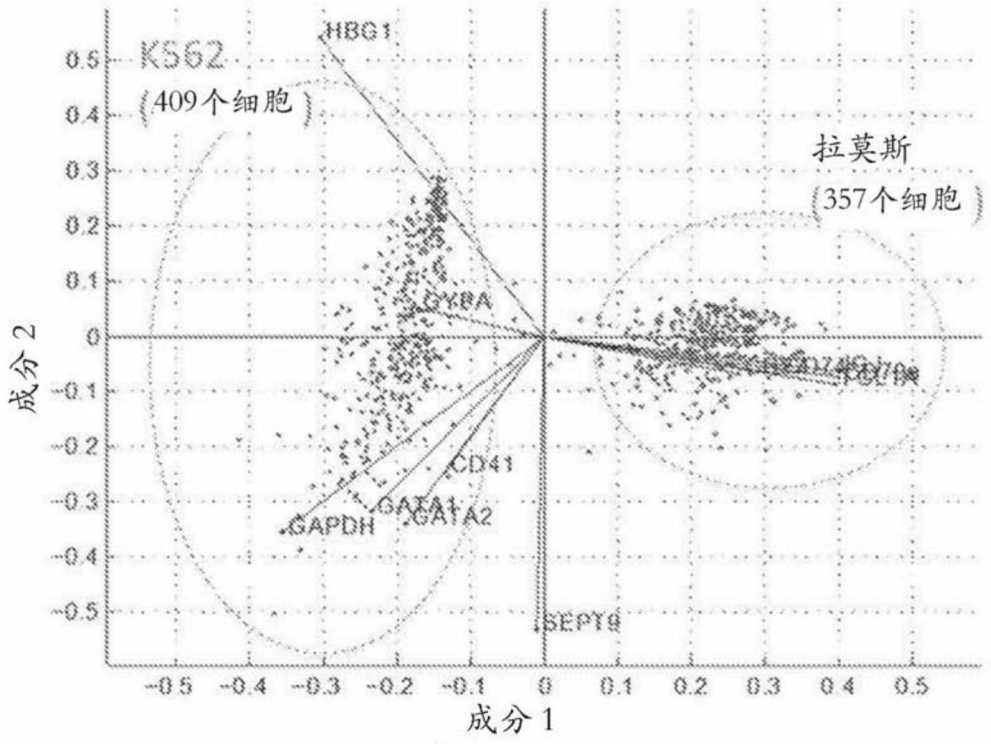

K562
(409个细胞)
拉莫斯
(357个细胞)
HBG1
GYPA
CD41
GATA1
GATA2
GAPDH
SEPT9
成分 2
成分 1

图31A

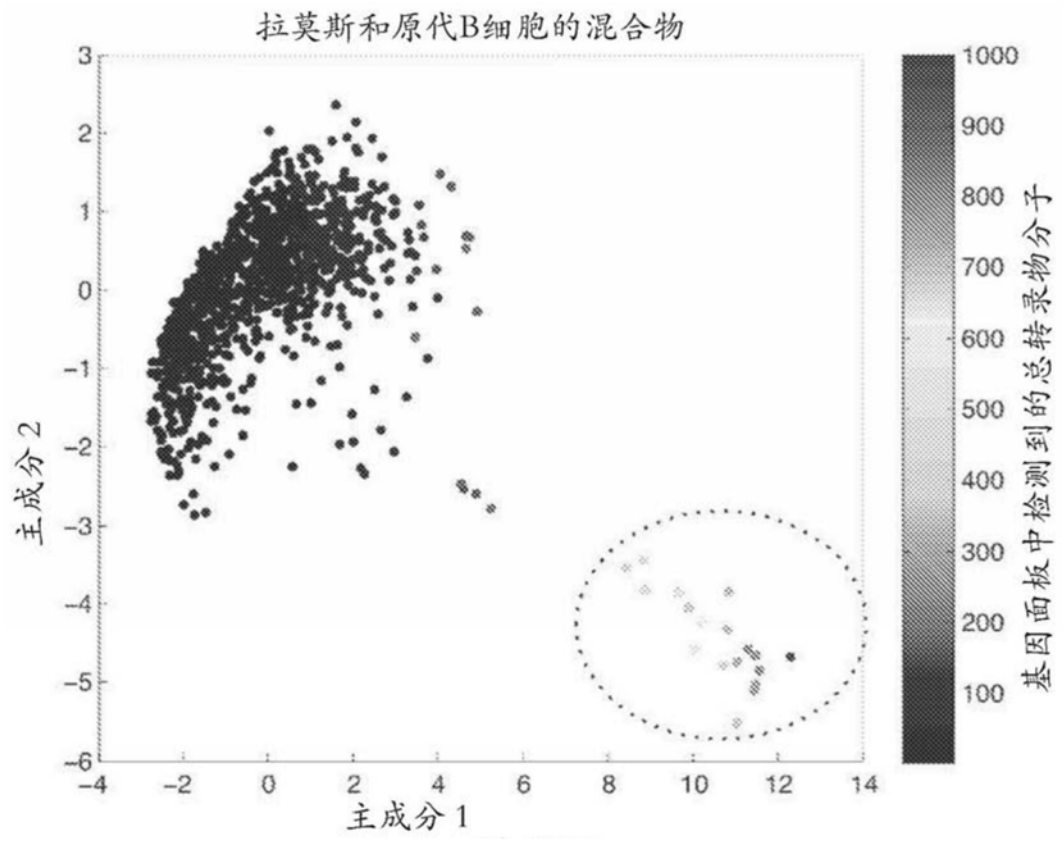

拉莫斯和原代B细胞的混合物
主成分 2
主成分 1
基因面板中检测到的总转录物分子

图31B

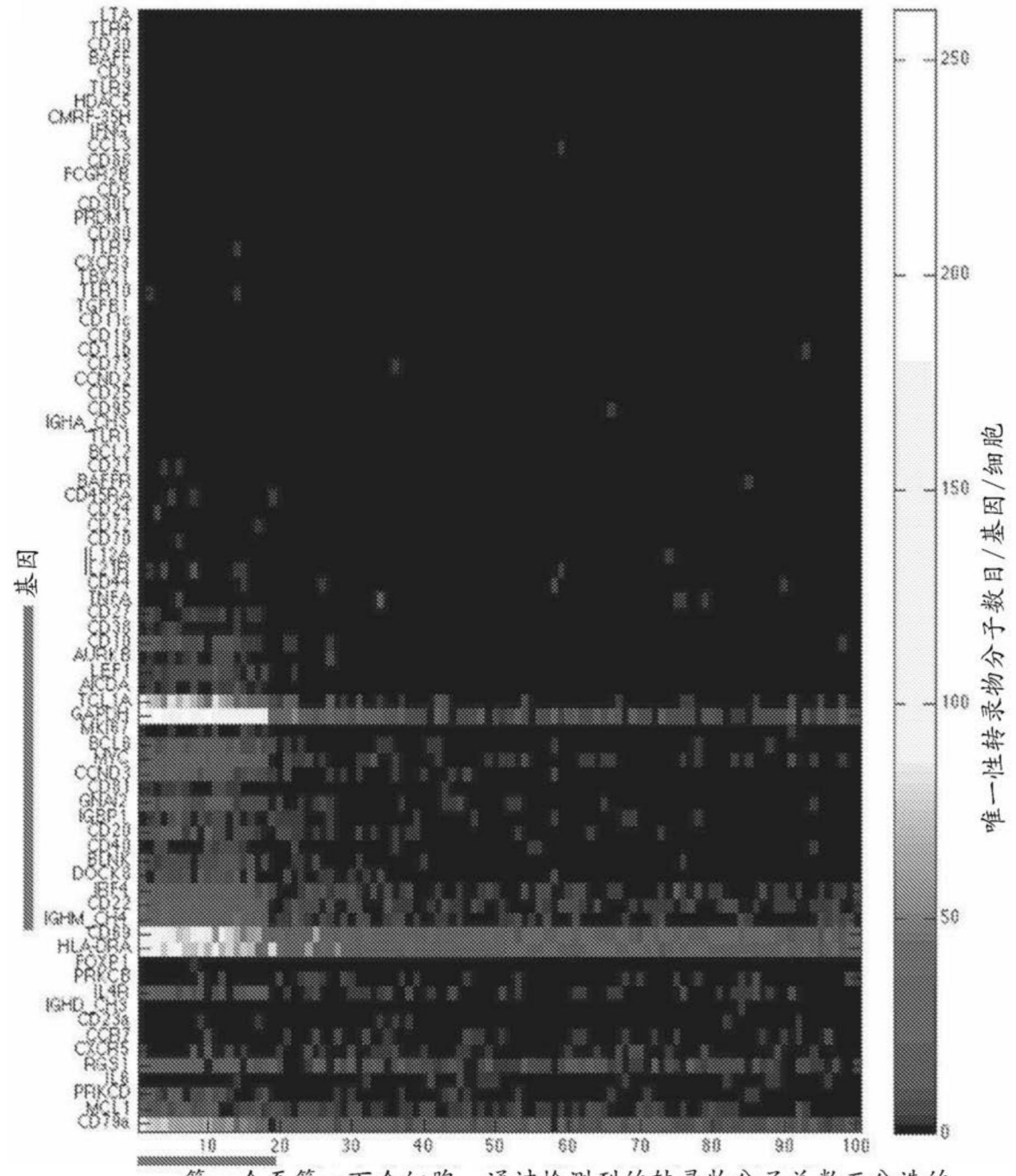

基因
唯一性转录物分子数目/基因/细胞
250
200
150
100
50
0
10
20
30
40
50
60
70
80
90
100
第一个至第一百个细胞，通过检测到的转录物分子总数而分选的

图31C

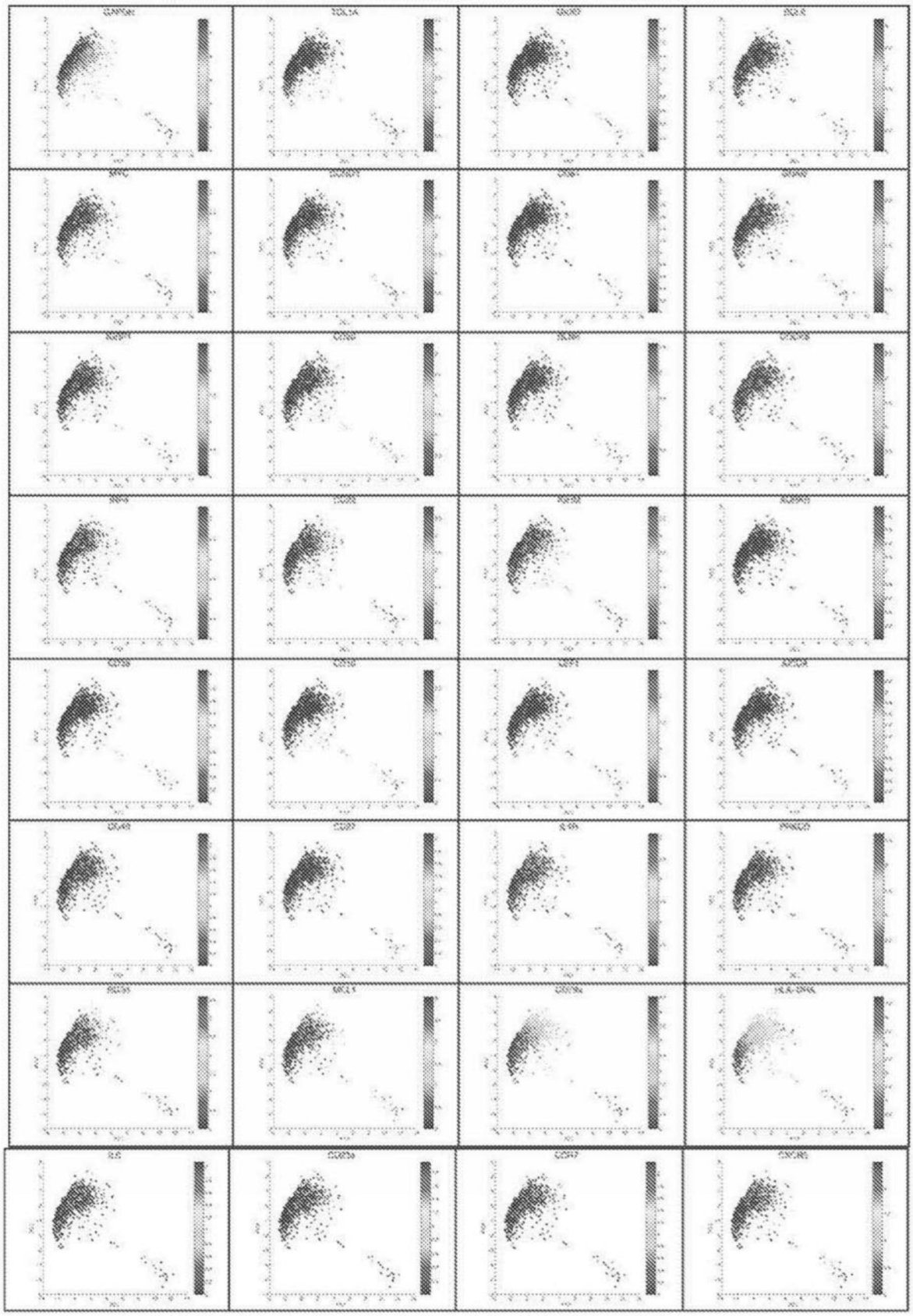


图31D

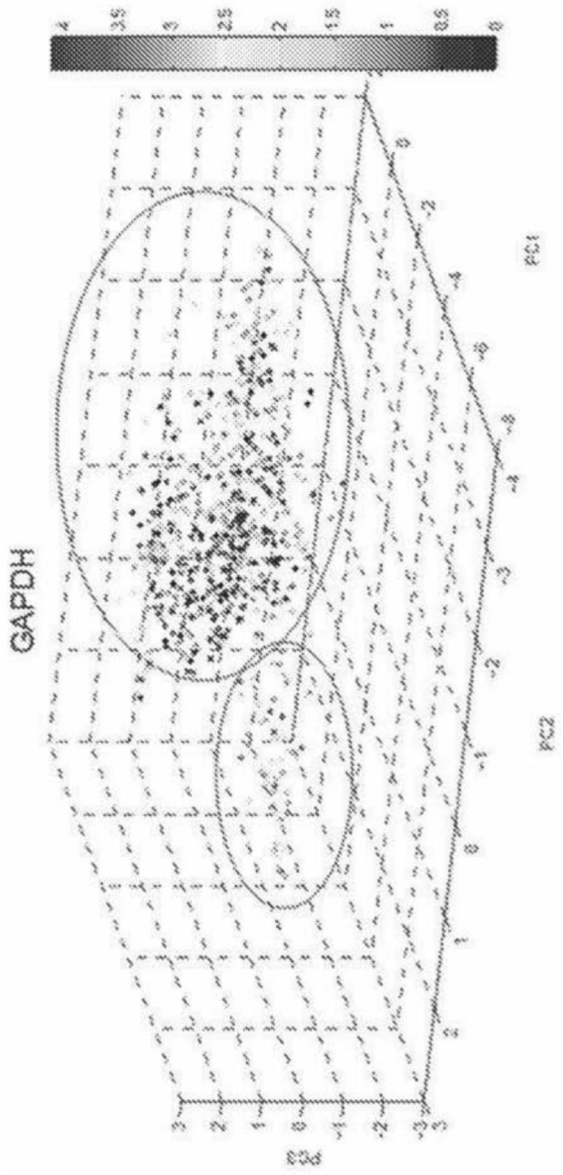


图32

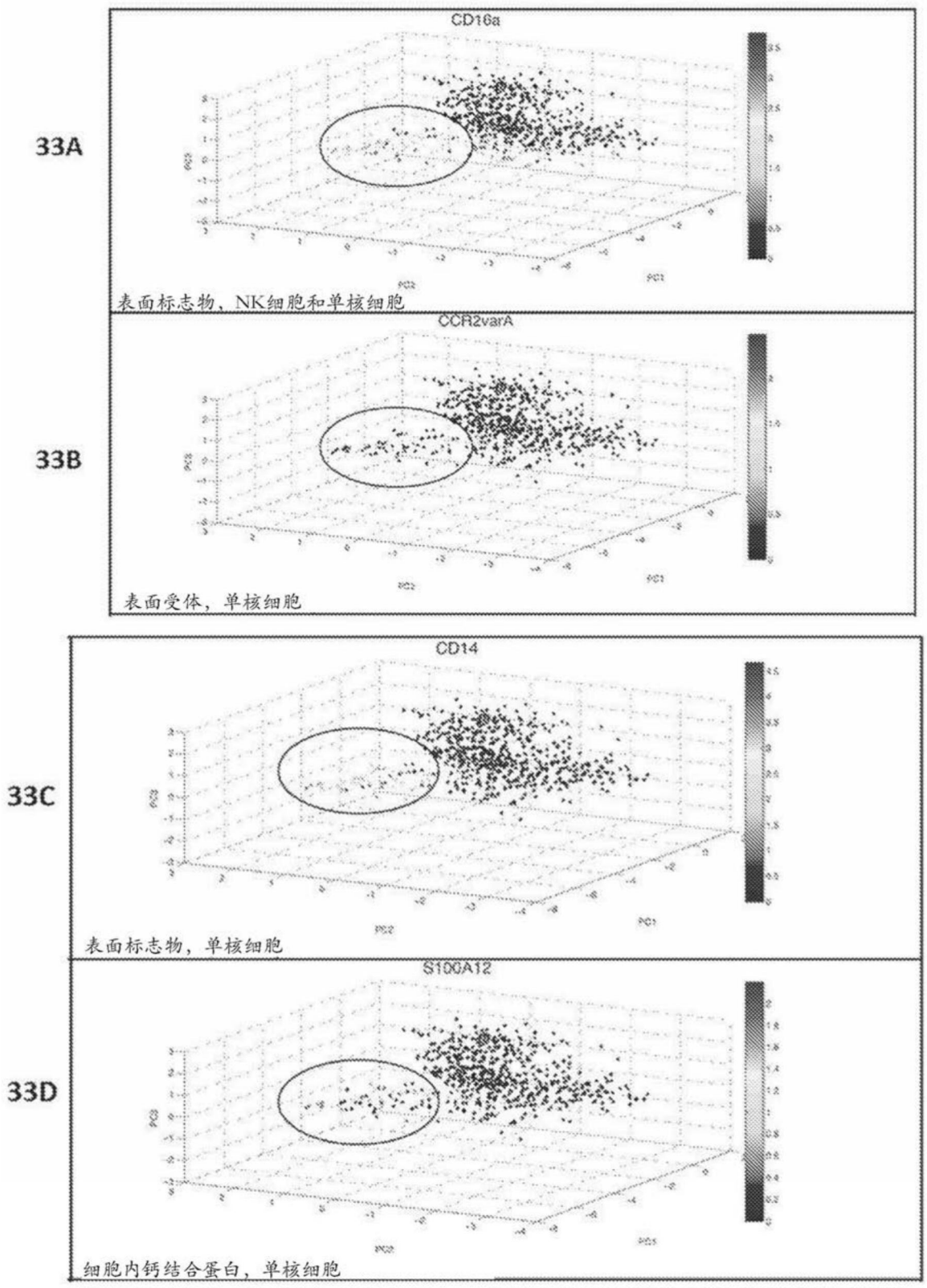

33A
CD16a
表面标志物，NK细胞和单核细胞
33B
CCR2varA
表面受体，单核细胞
33C
CD14
表面标志物，单核细胞
33D
S100A12
细胞内钙结合蛋白，单核细胞

图33A-D

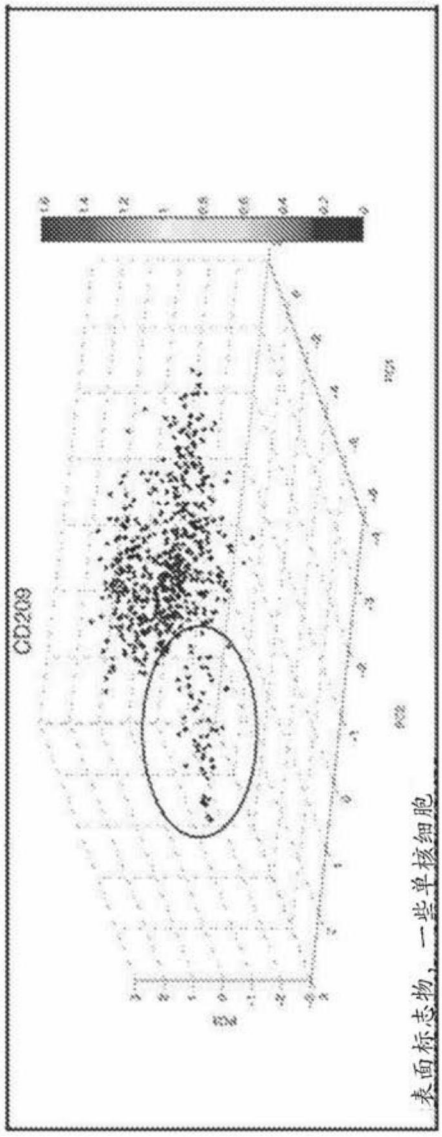


图33E

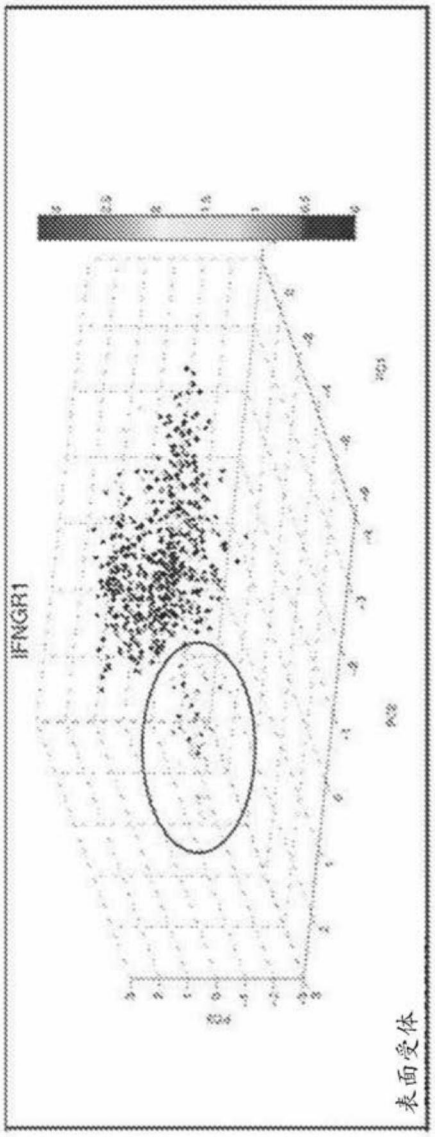


图33F

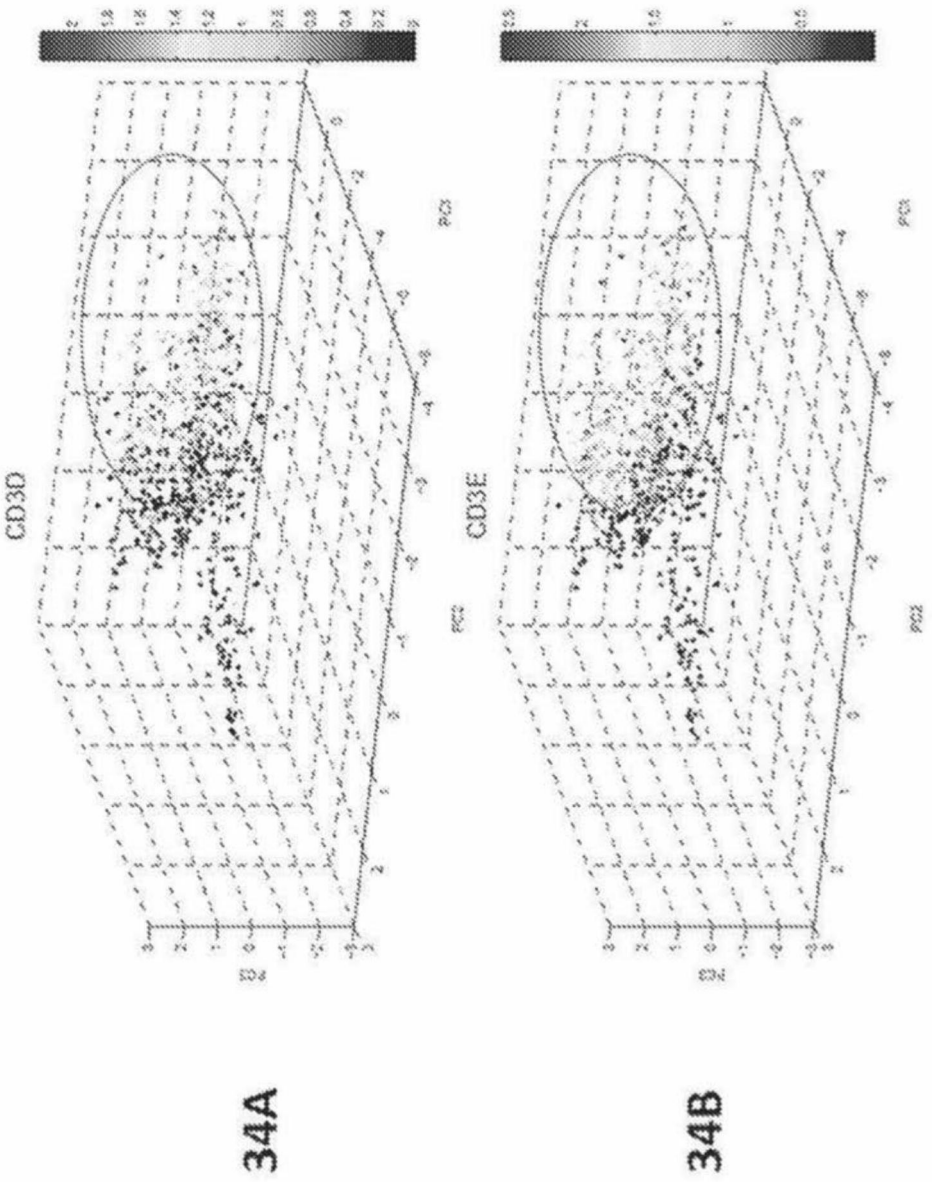


图34A-B

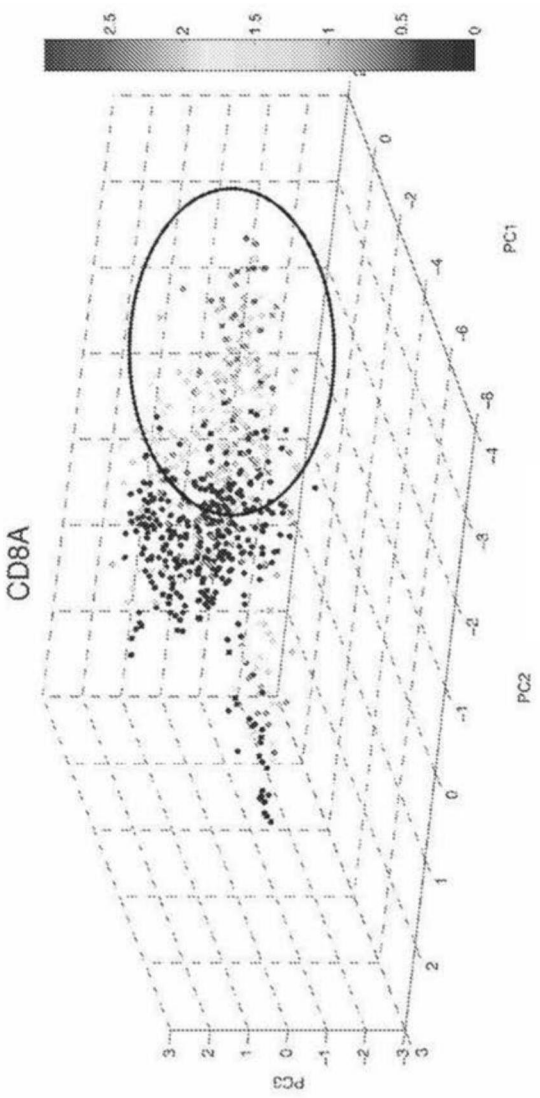


图35A

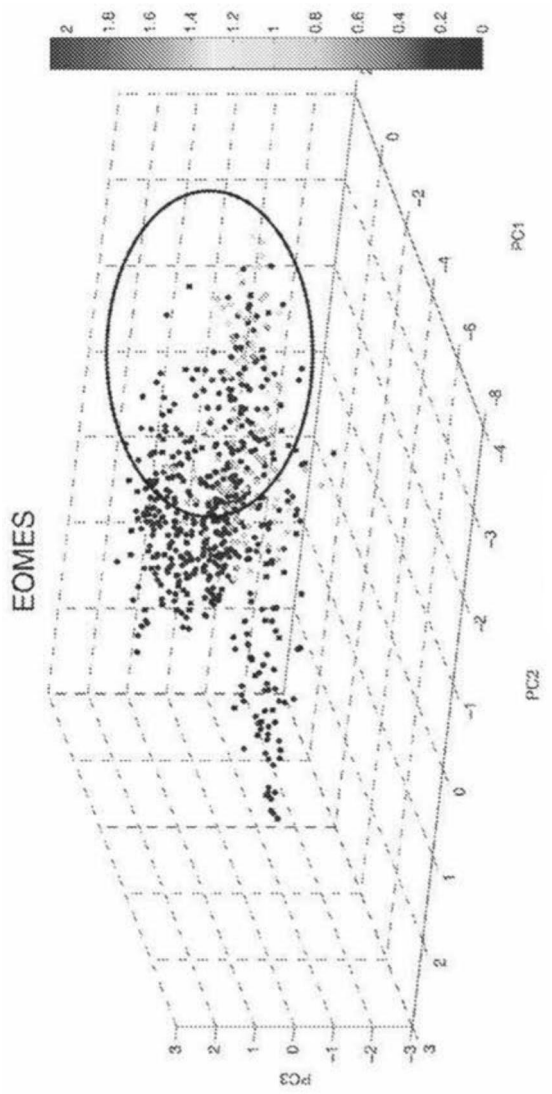


图35B

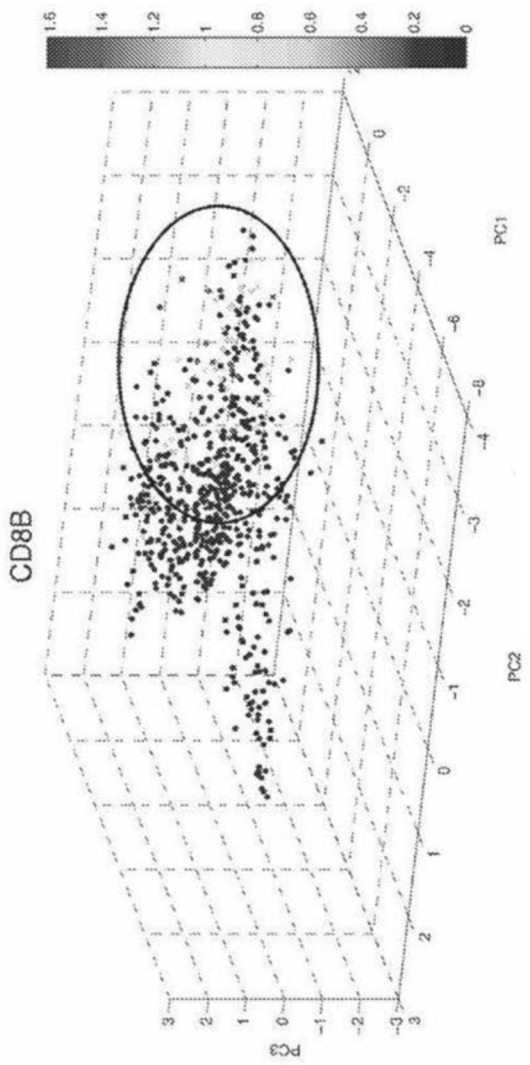


图35C

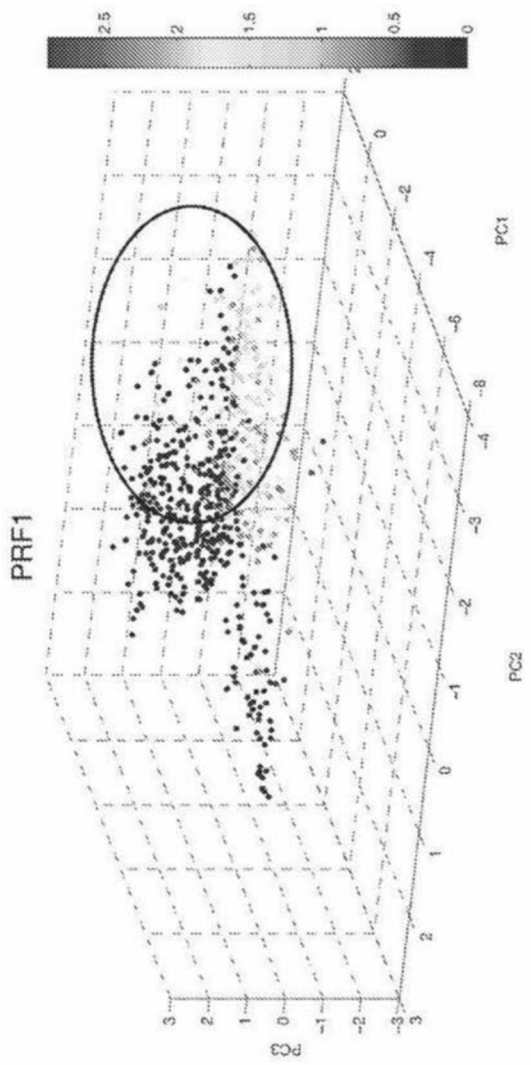


图35D

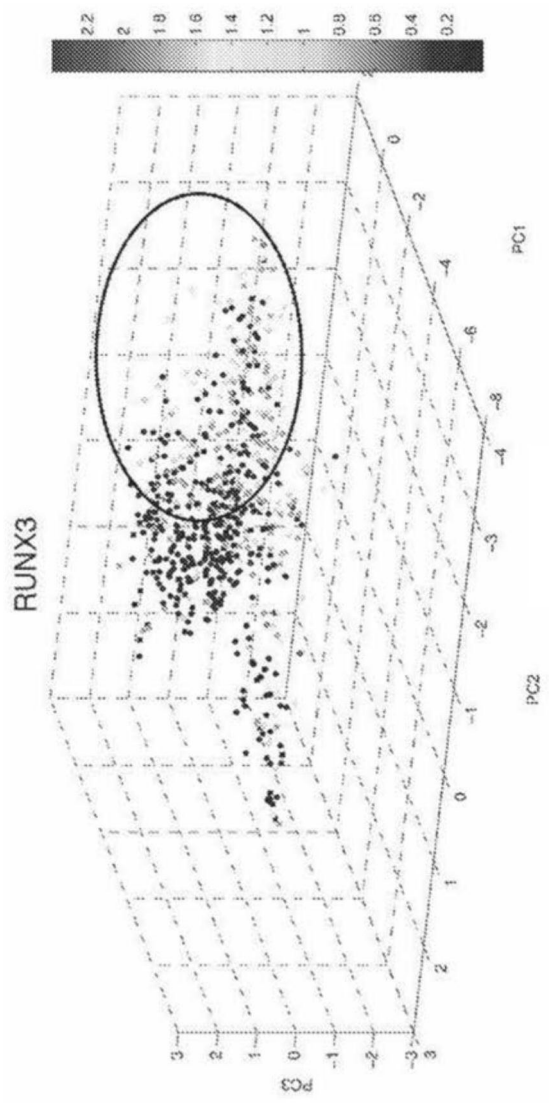


图35E

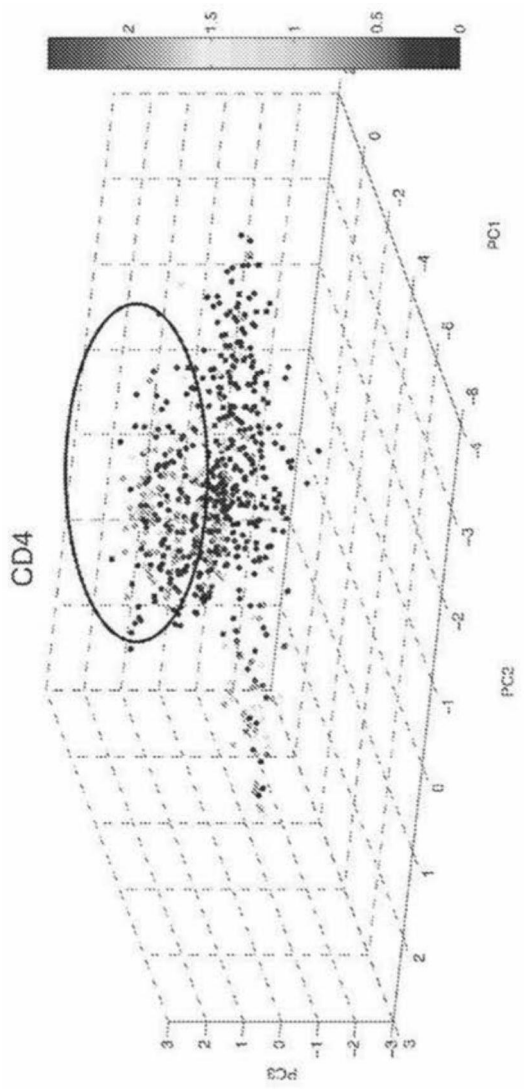


图36A

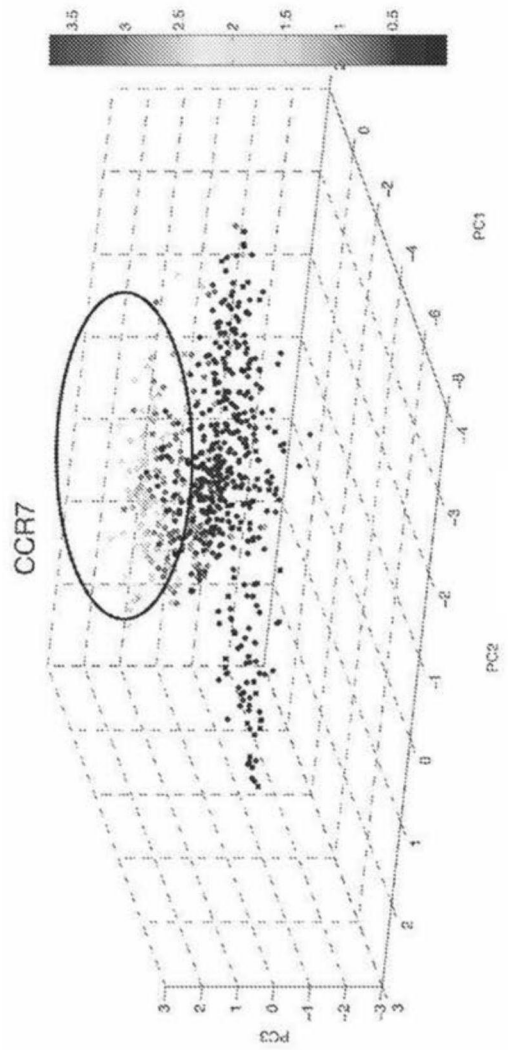


图36B

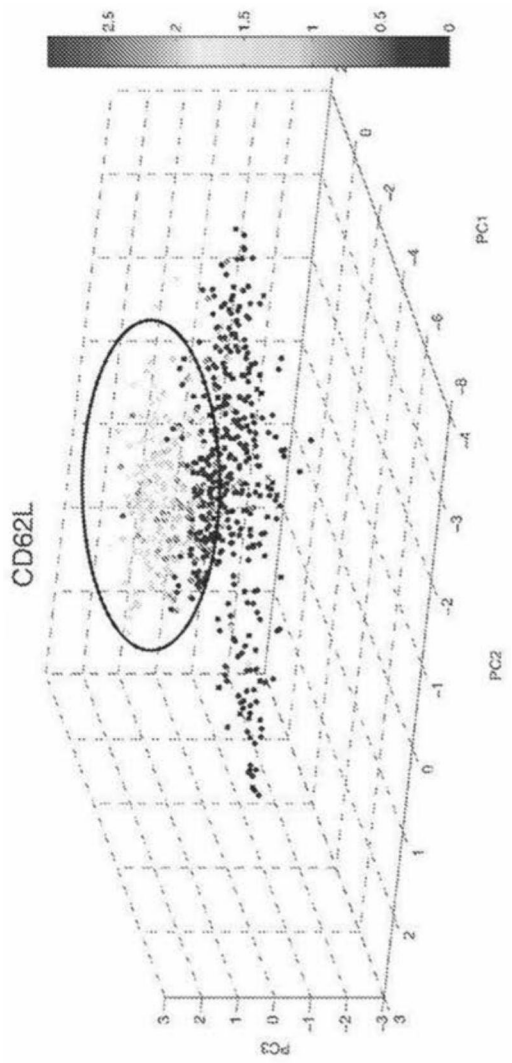


图36C

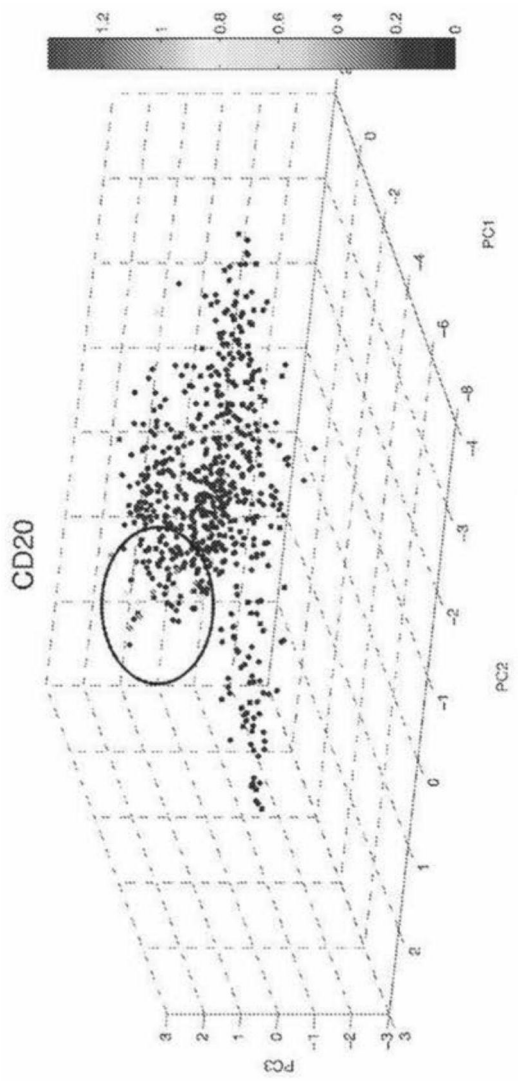


图37A

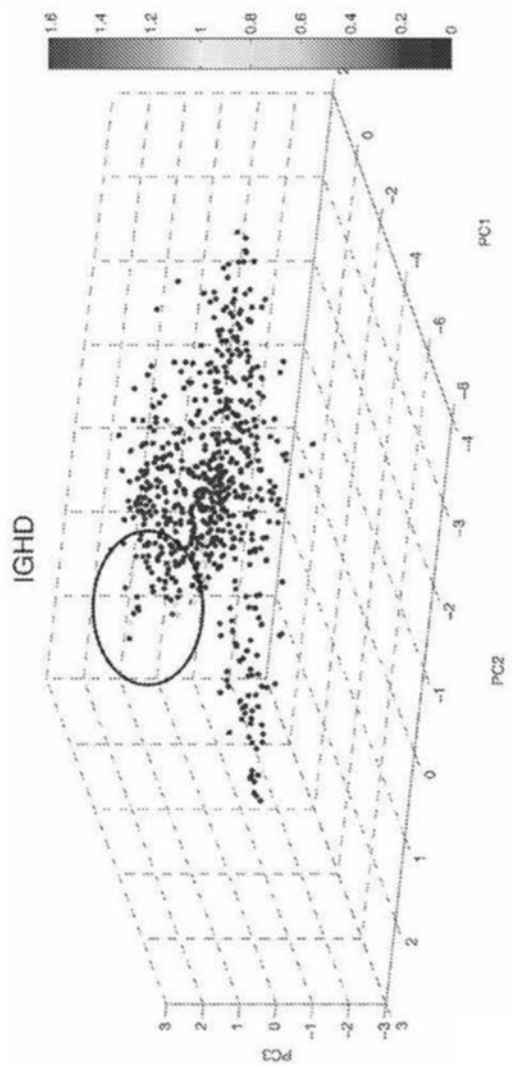


图37B

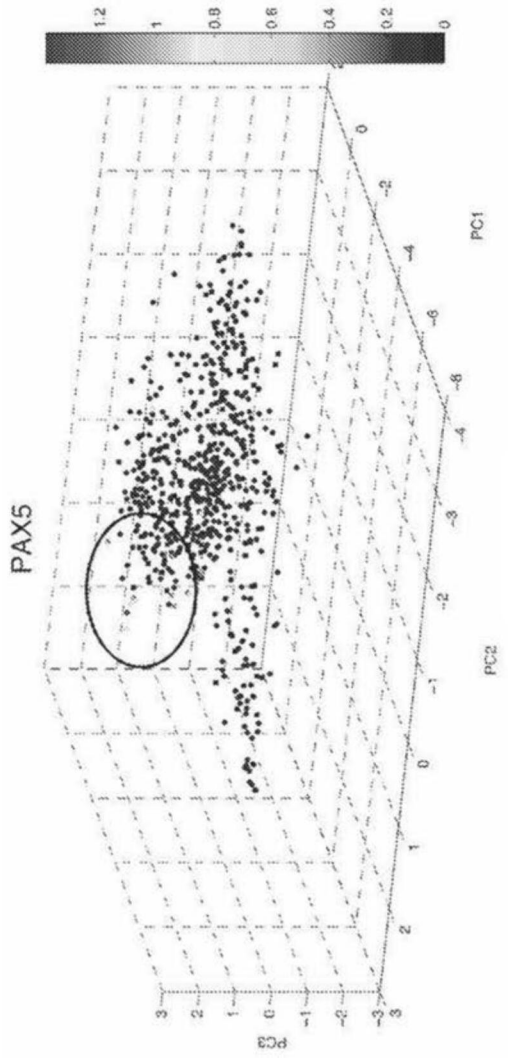


图37C

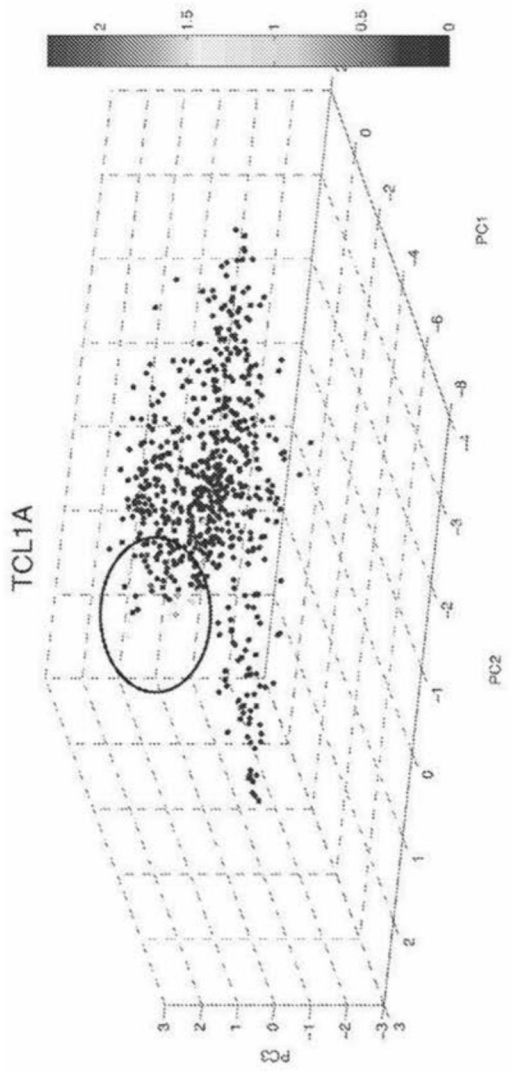


图37D

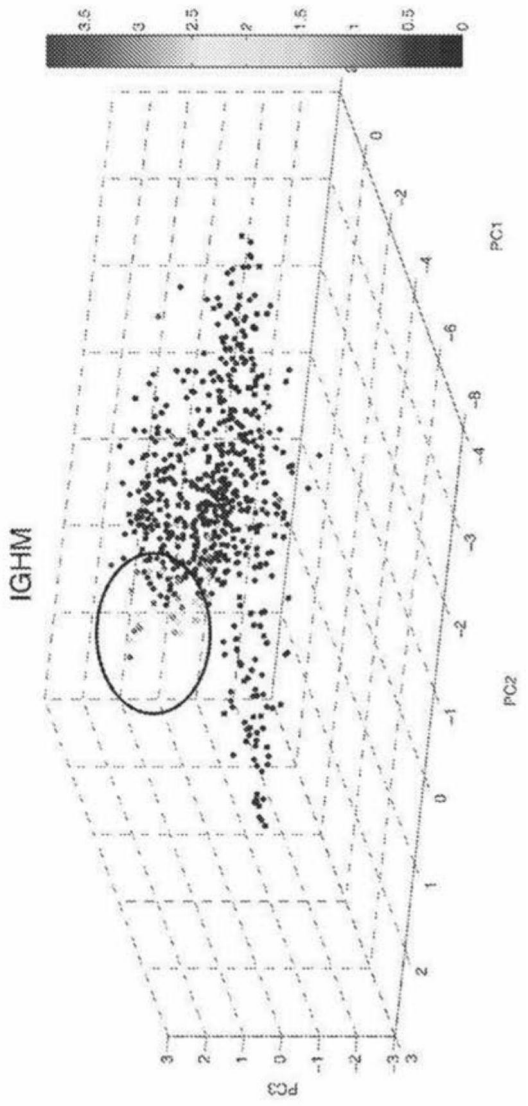


图37E

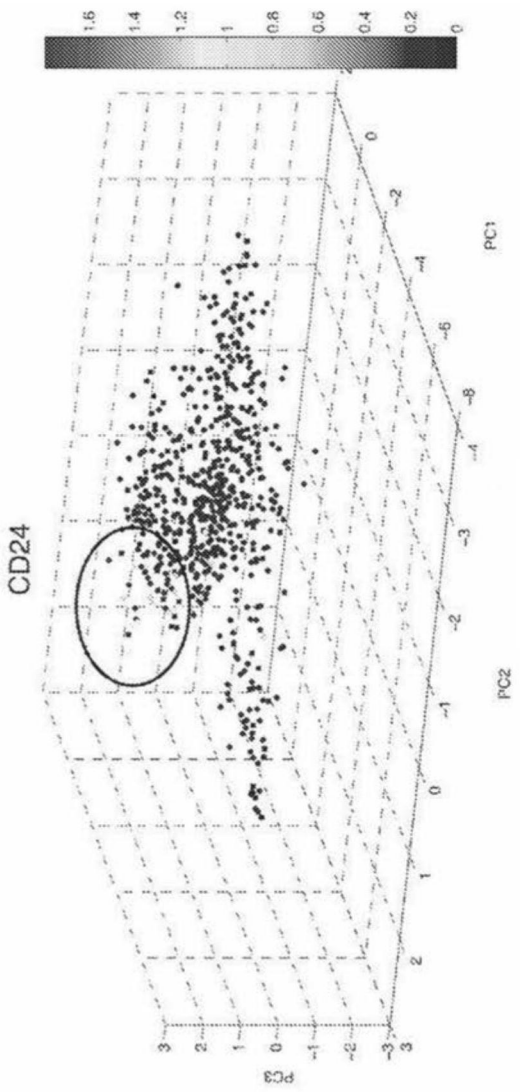


图37F

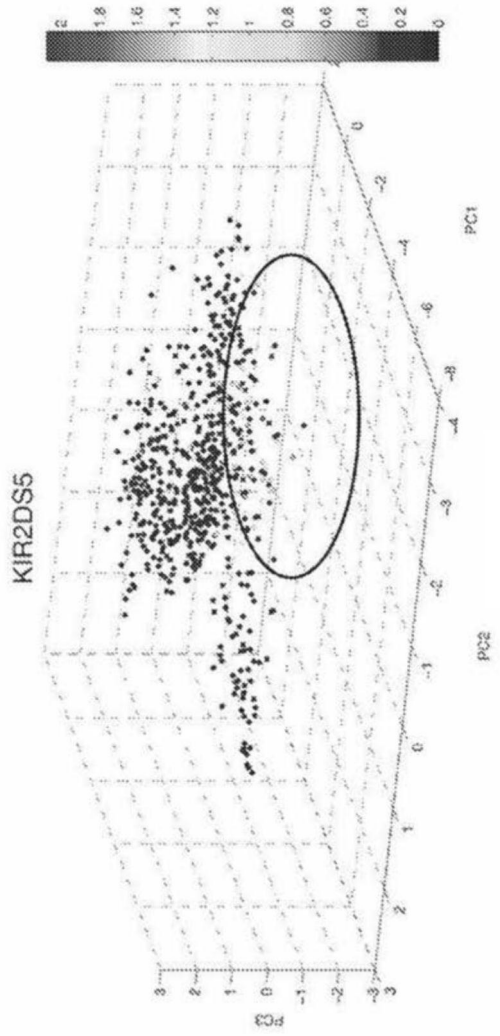


图38A

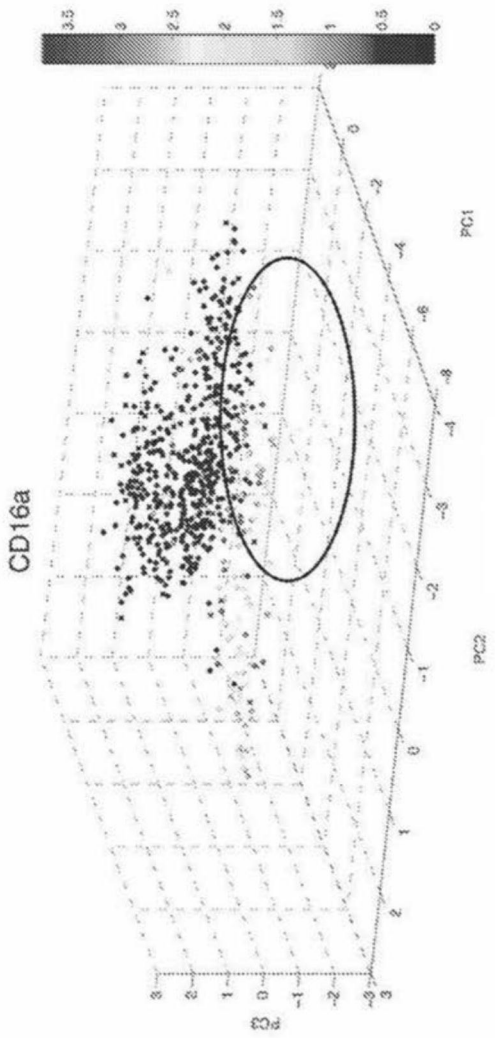


图38B

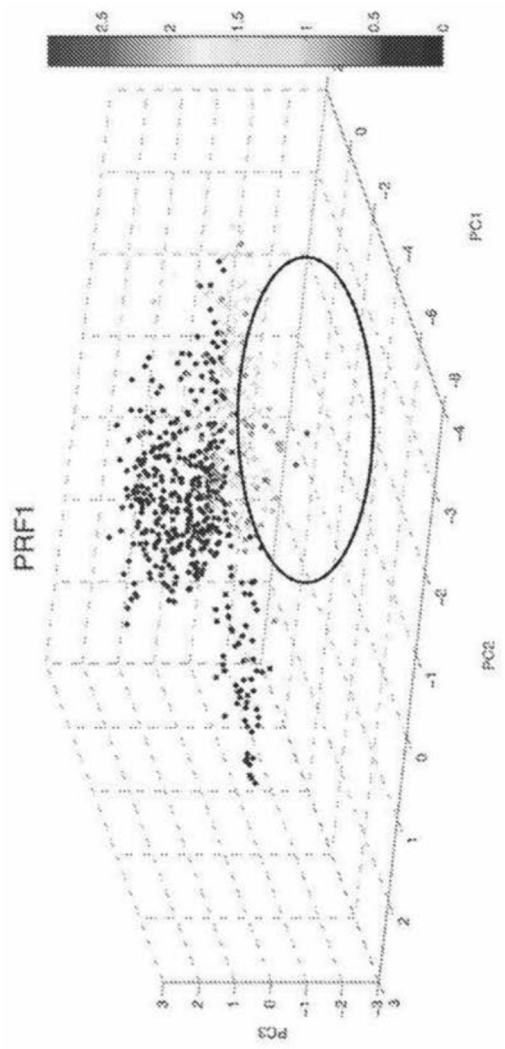


图38C

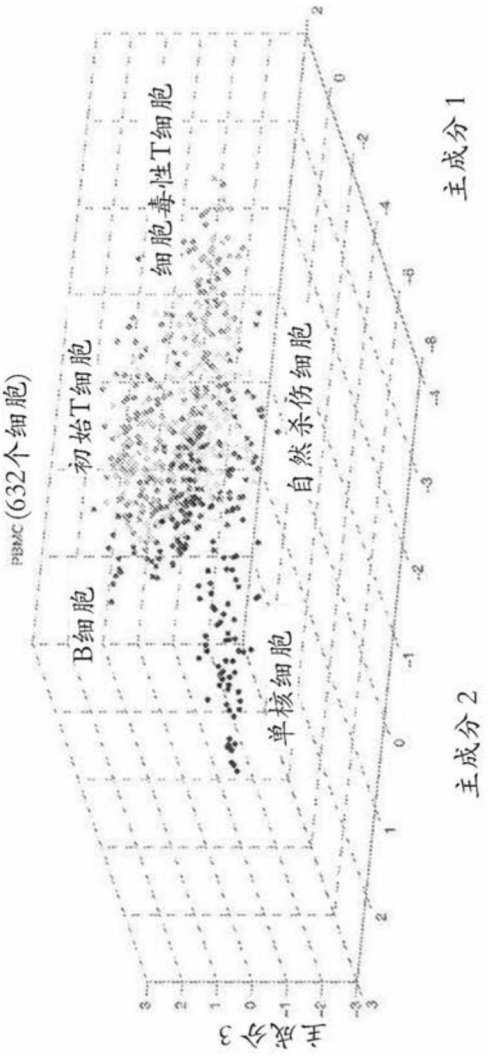


图39

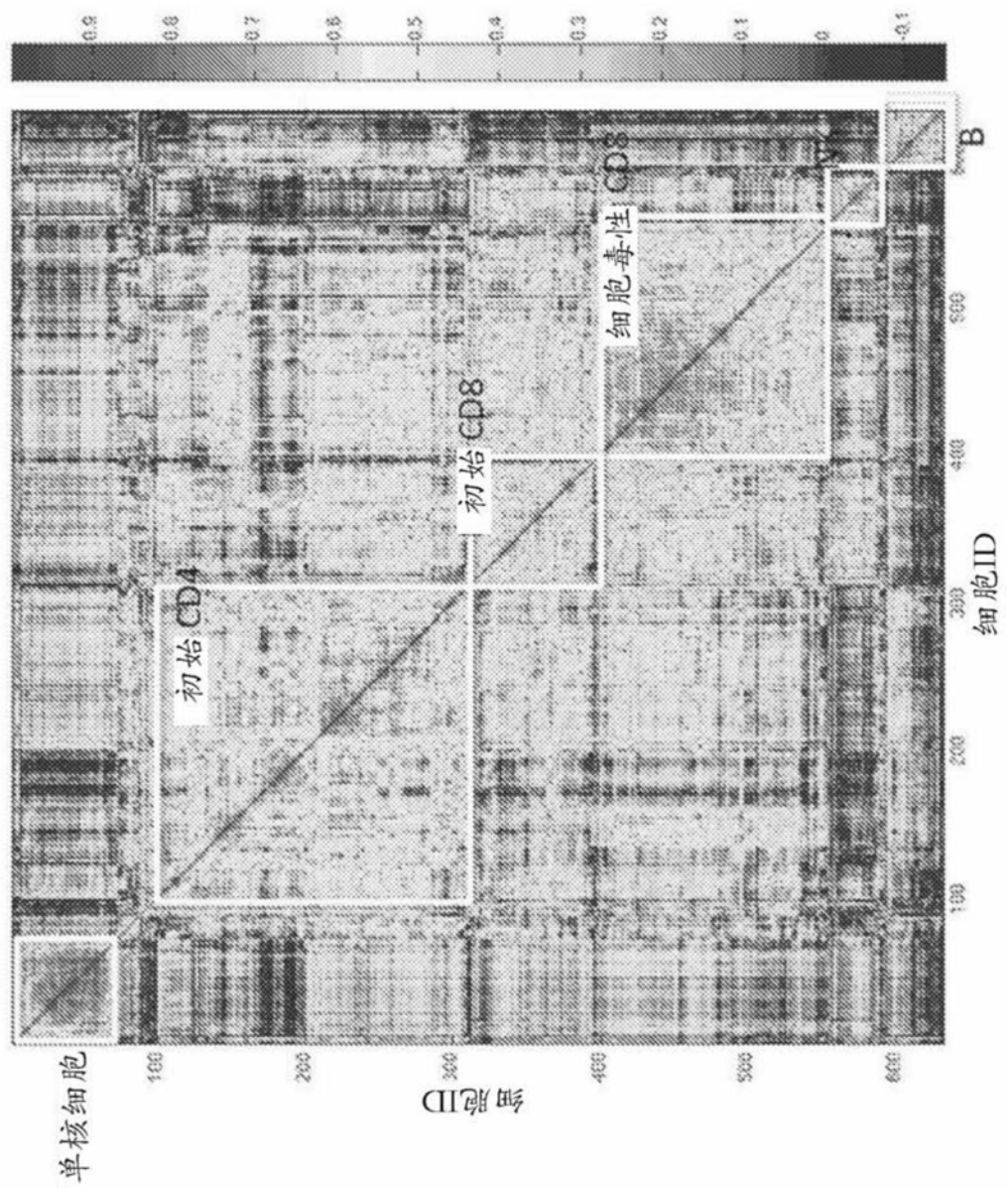


图40A

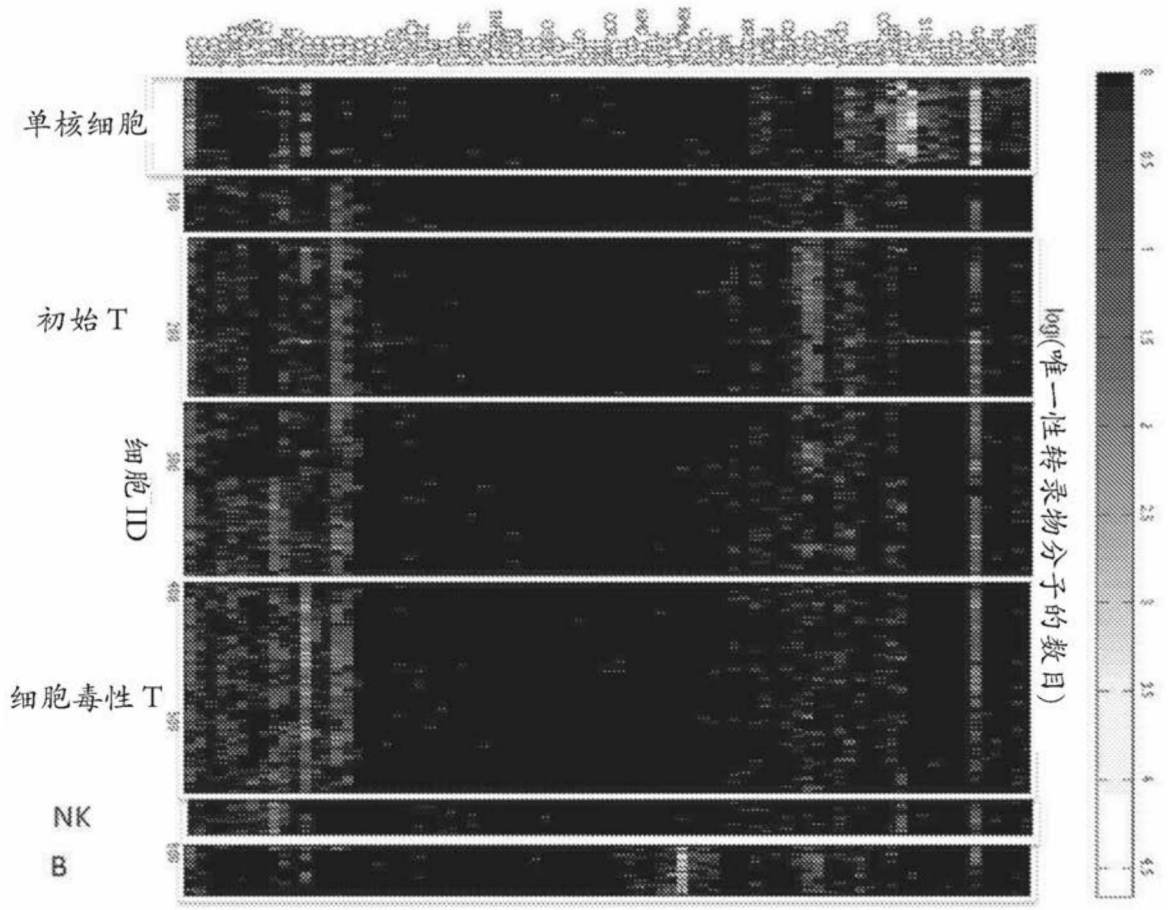

单核细胞
初始 T
细胞 ID
细胞毒性 T
NK
B
log(唯一性转录物分子的数目)

图40B

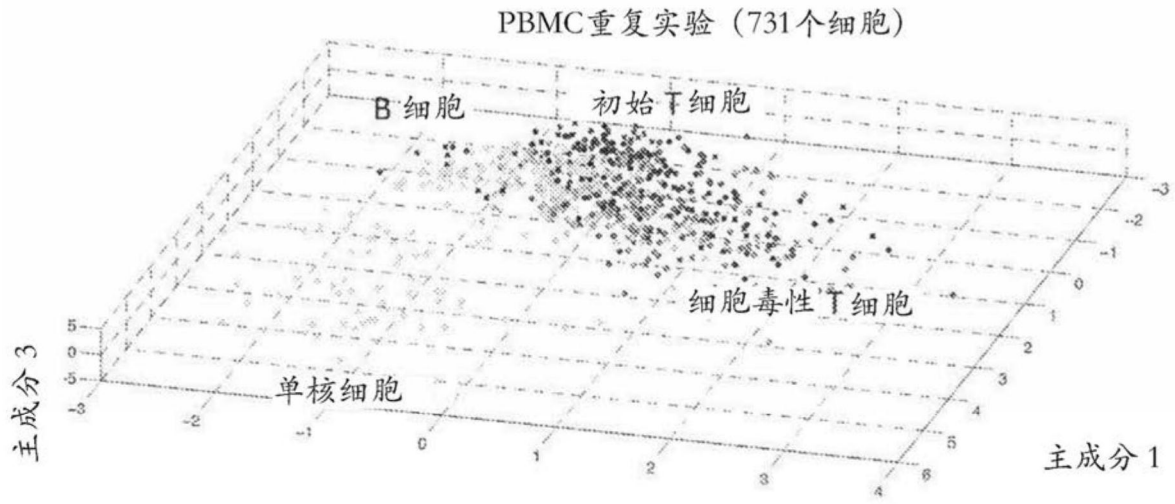

PBMC重复实验（731个细胞）
B 细胞
初始 T 细胞
细胞毒性 T 细胞
单核细胞
主成分 3
主成分 1

图41

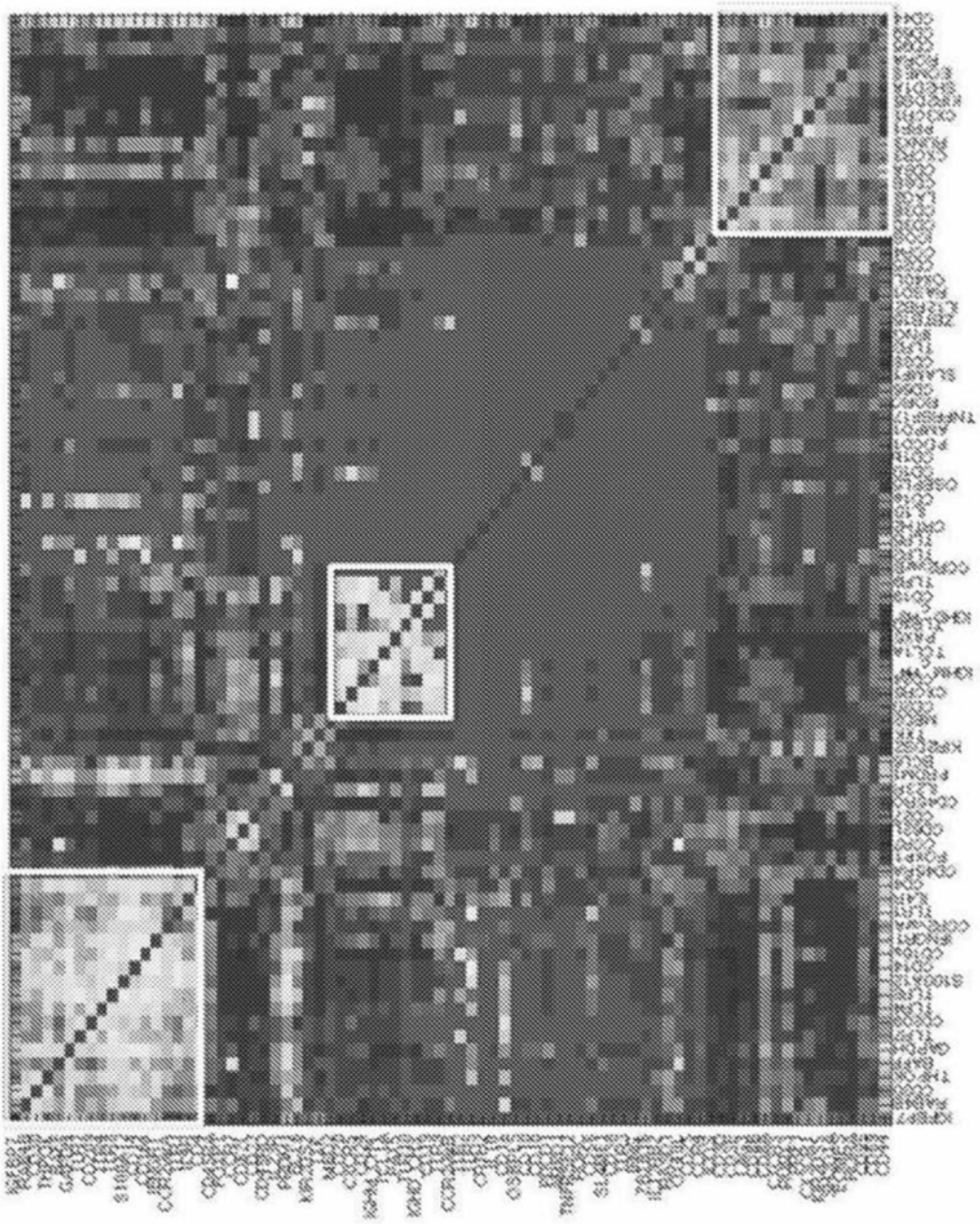


图42

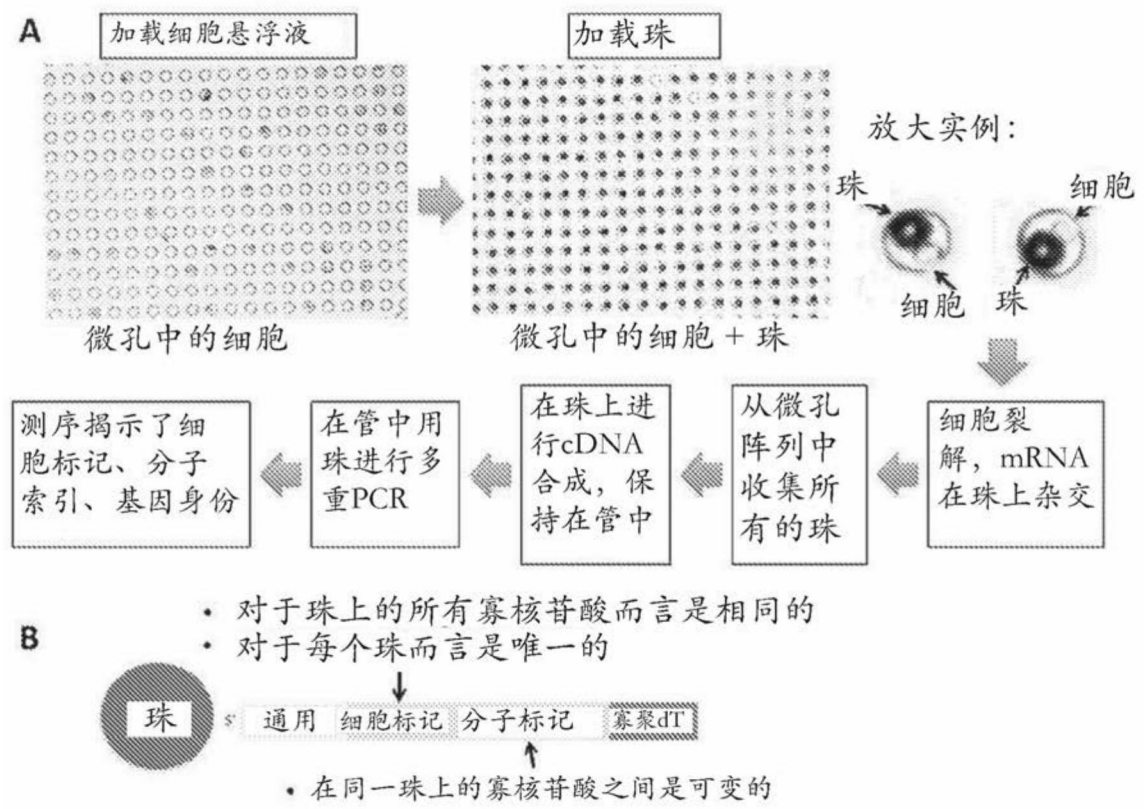

A
加载细胞悬浮液
加载珠
放大实例：
珠
细胞
细胞
珠
微孔中的细胞
微孔中的细胞 + 珠
细胞裂解，mRNA在珠上杂交
从微孔阵列中收集所有的珠
在珠上进行cDNA合成，保持在管中
在管中用珠进行多重PCR
测序揭示了细胞标记、分子索引、基因身份
B
• 对于珠上的所有寡核苷酸而言是相同的
• 对于每个珠而言是唯一的
珠
5'
通用
细胞标记
分子标记
寡聚dT
• 在同一珠上的寡核苷酸之间是可变的

图43

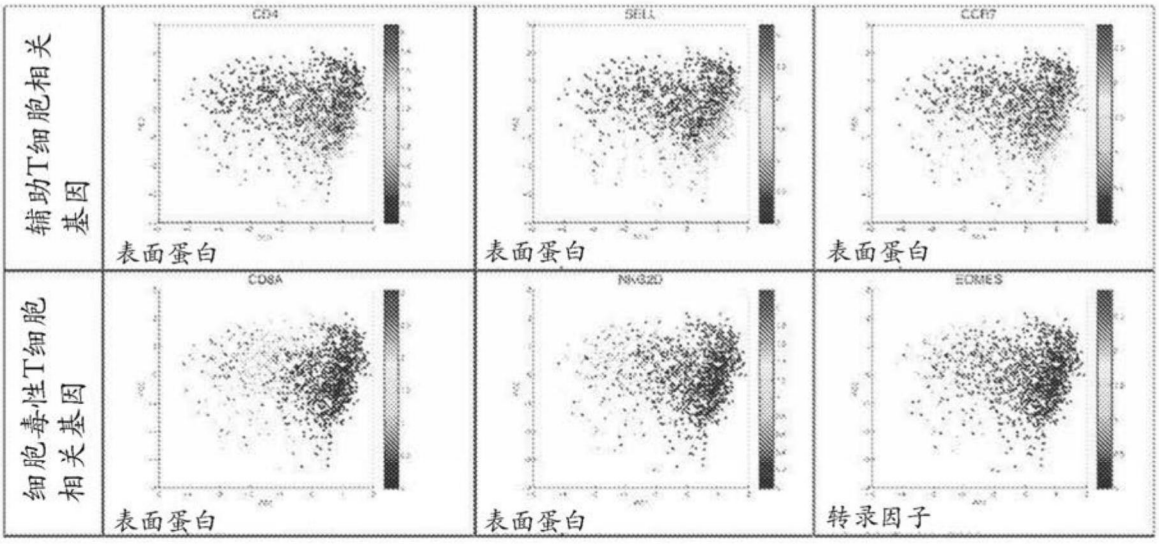

辅助T细胞相关基因
细胞毒性T细胞相关基因
CD4
SELL
CCR7
CD8A
NKG2D
EOMES
表面蛋白
表面蛋白
表面蛋白
表面蛋白
表面蛋白
转录因子

图44A

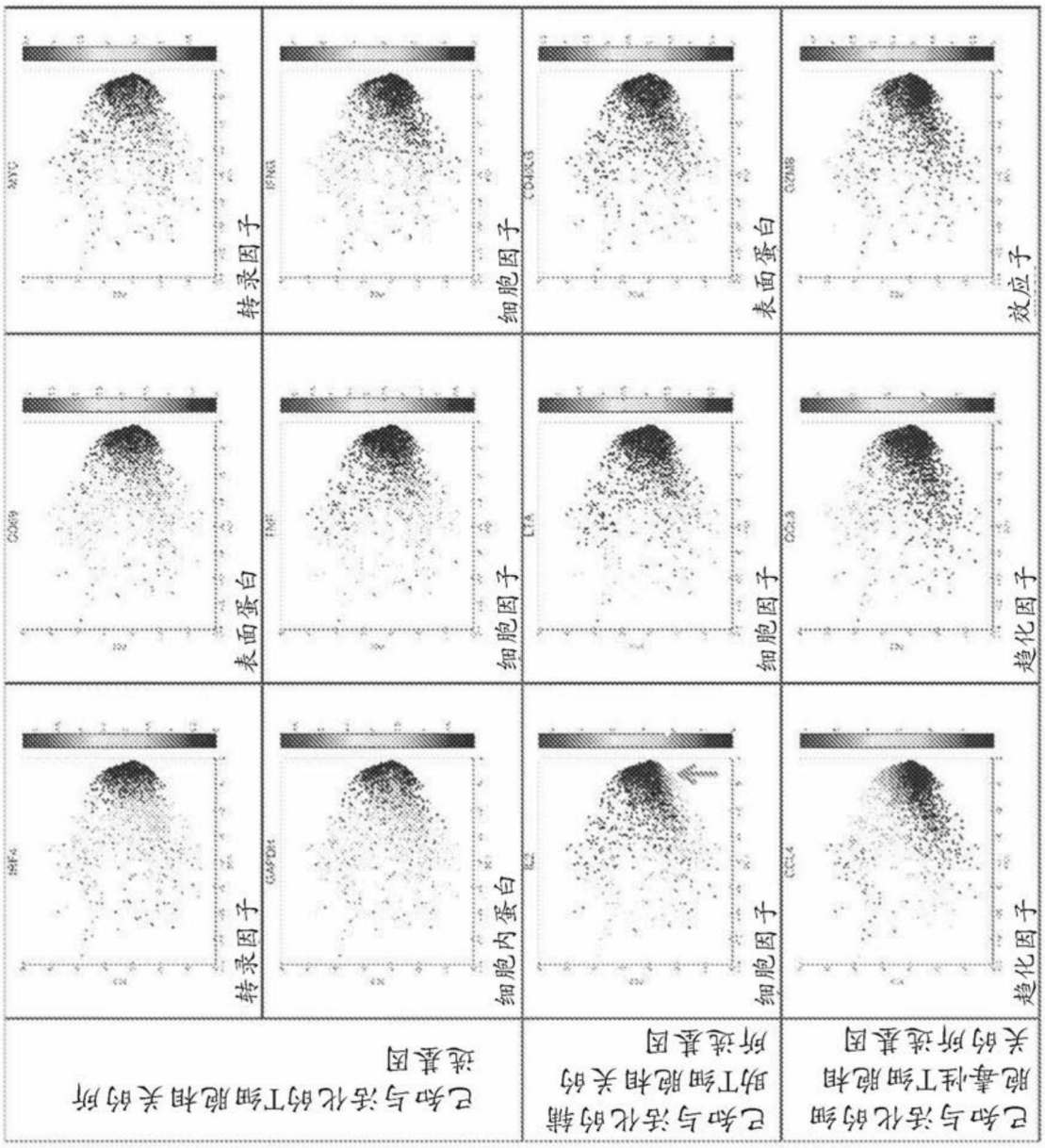

已知与活化的细胞毒性T细胞相关的所选基因
已知与活化的辅助T细胞相关的所选基因
已知与活化的T细胞相关的所选基因
趋化因子
细胞因子
细胞因子
细胞内蛋白
表面蛋白
转录因子
效应子
表面蛋白
细胞因子
细胞因子
趋化因子
转录因子

图44B

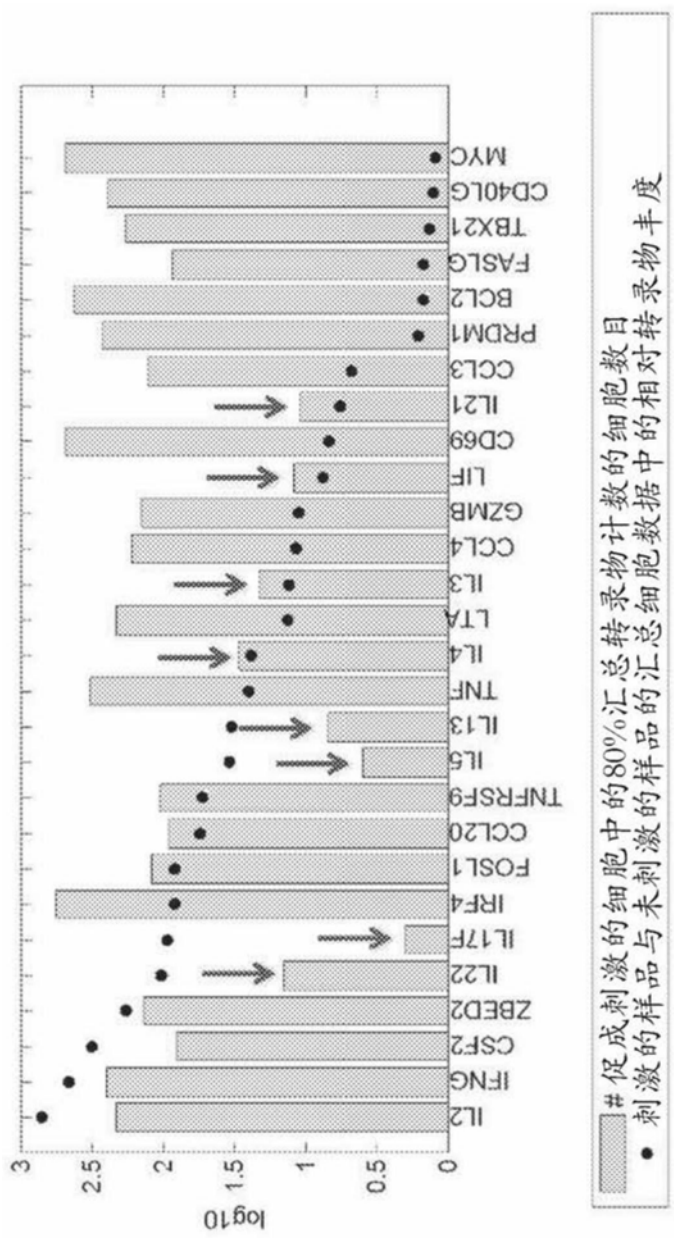

log10
3
2.5
2
1.5
1
0.5
0
IL2
IFNG
CSF2
ZBED2
IL22
IL17F
IRF4
FOSL1
CCL20
TNFRSF9
IL5
IL13
TNF
IL4
LTA
IL3
CCL4
GZMB
LIF
CD69
IL21
CCL3
PRDM1
BCL2
FASLG
TBX21
CD40LG
MYC
促成刺激的细胞中的80%汇总转录物计数的细胞数目
• 刺激的样品与未刺激的样品的汇总细胞数据中的相对转录物丰度

图44C

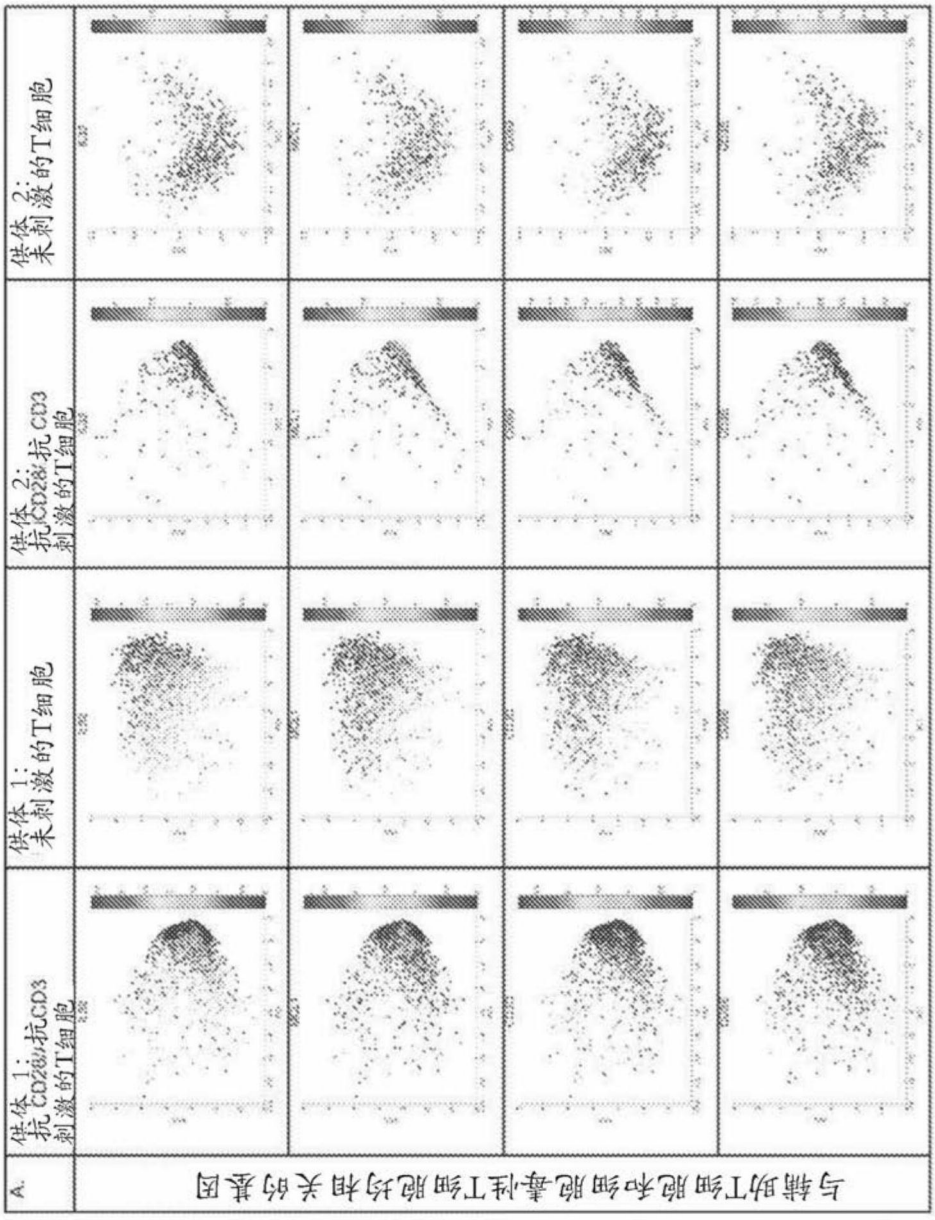

A.
与辅助T细胞和细胞毒性T细胞均相关的基因
供体 1:
抗CD28/抗CD3
刺激的T细胞
供体 1:
未刺激的T细胞
供体 2:
抗CD28/抗CD3
刺激的T细胞
供体 2:
未刺激的T细胞

图45A

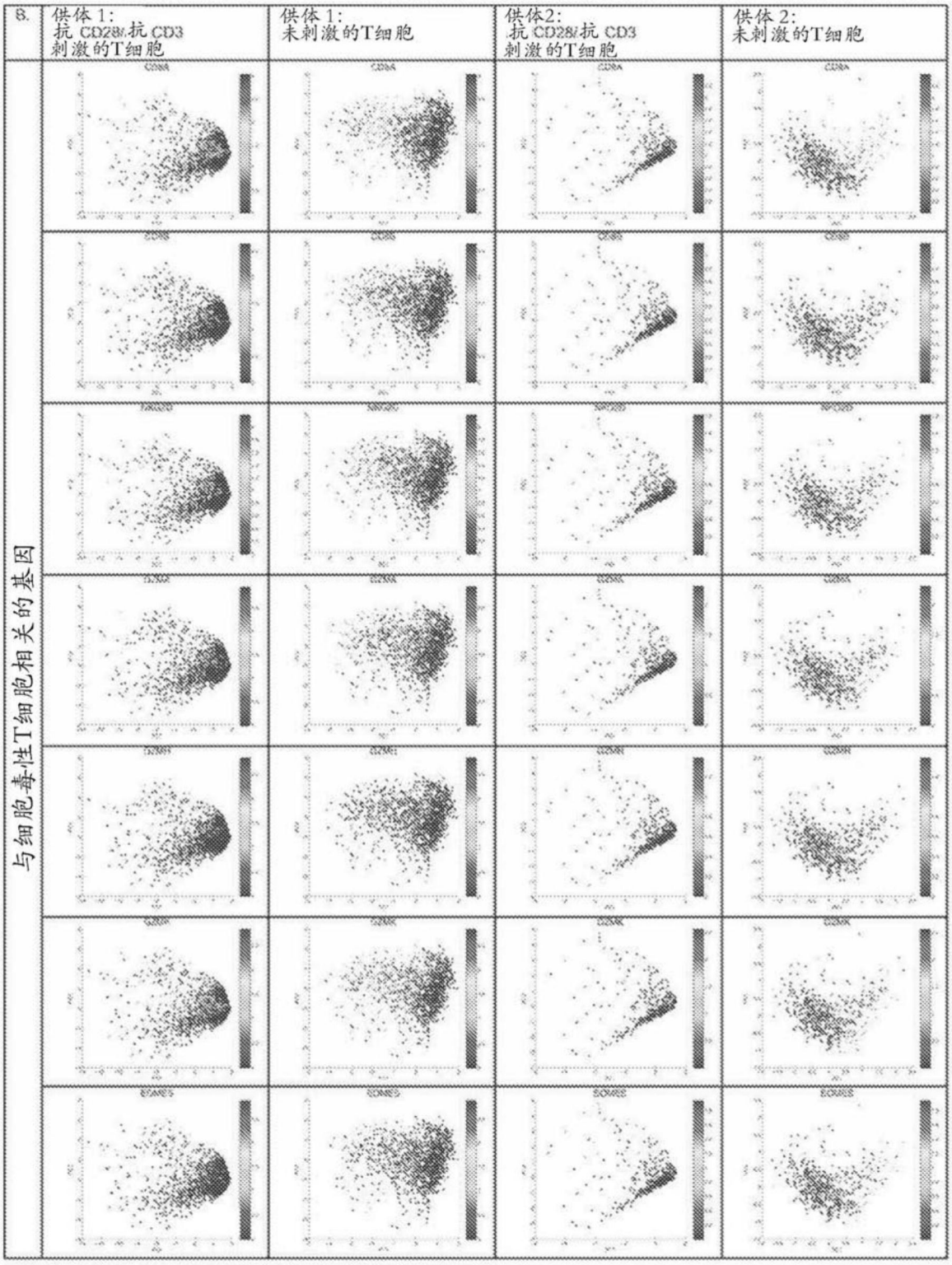

B.
供体 1:
抗 CD28/抗 CD3
刺激的T细胞
供体 1:
未刺激的T细胞
供体2:
抗 CD28/抗 CD3
刺激的T细胞
供体 2:
未刺激的T细胞
与细胞毒性T细胞相关的基因

图45B

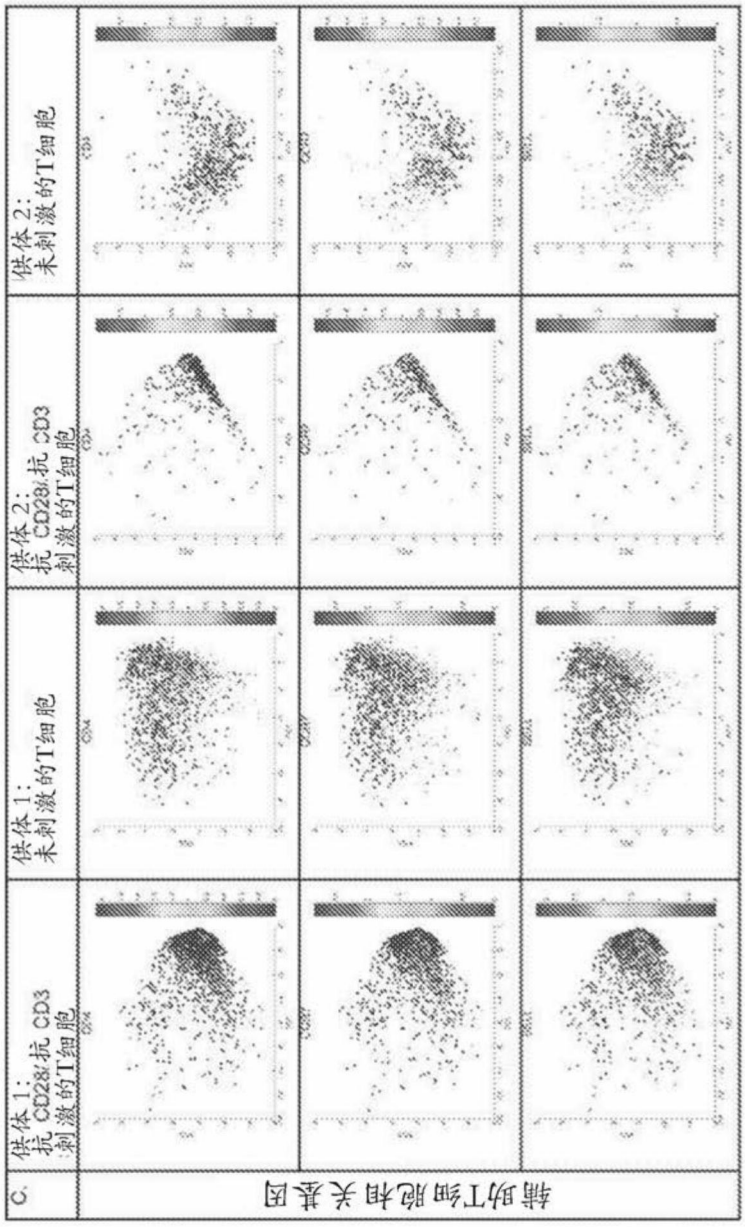

C.
辅助T细胞相关基因
供体 1:
抗 CD28/抗 CD3
刺激的T细胞
供体 1:
未刺激的T细胞
供体 2:
抗 CD28/抗 CD3
刺激的T细胞
供体 2:
未刺激的T细胞

图45C

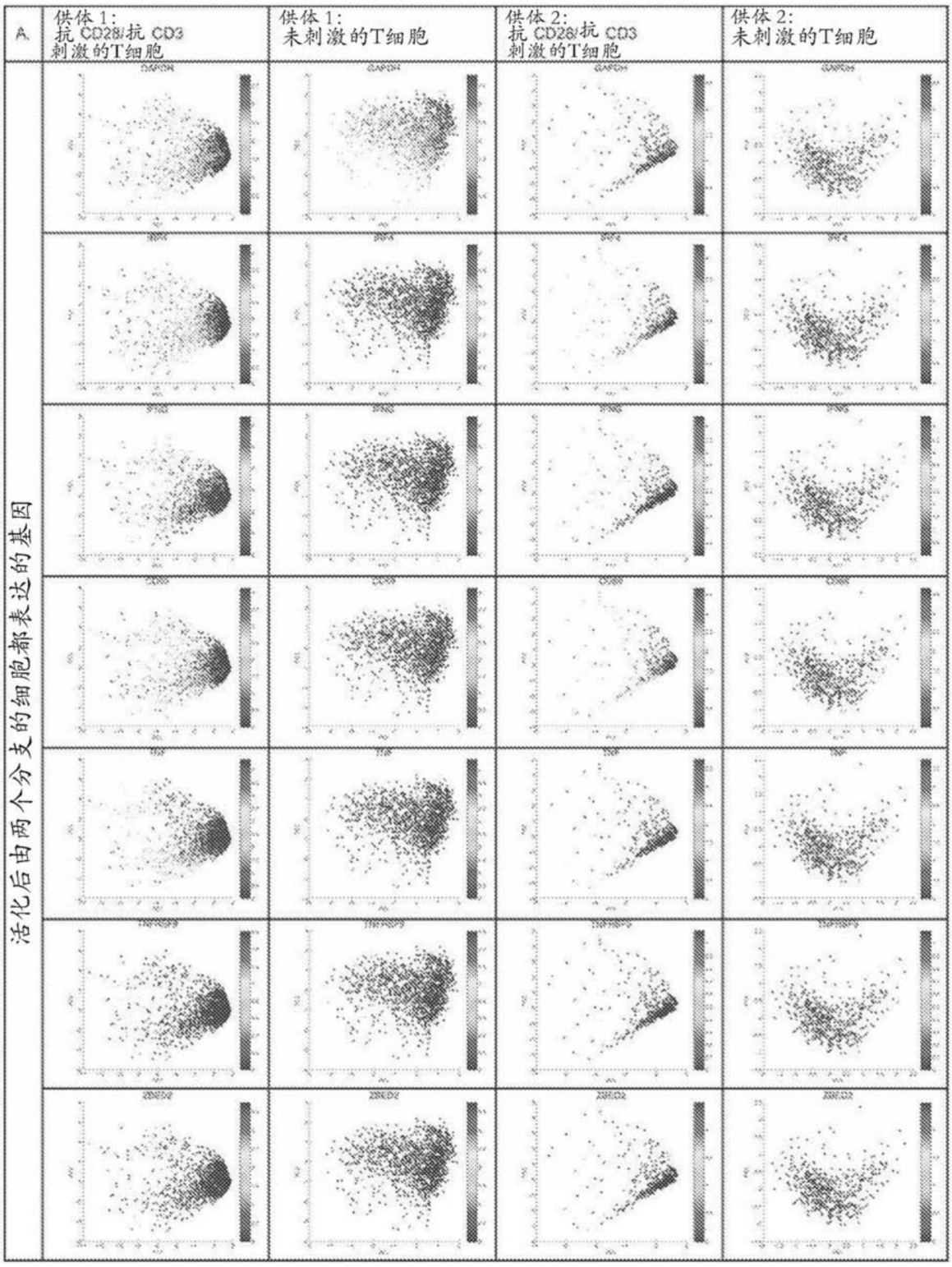

A
供体 1:
抗 CD28/抗 CD3
刺激的T细胞
供体 1:
未刺激的T细胞
供体 2:
抗 CD28/抗 CD3
刺激的T细胞
供体 2:
未刺激的T细胞
活化后由两个分支的细胞都表达的基因

图46A

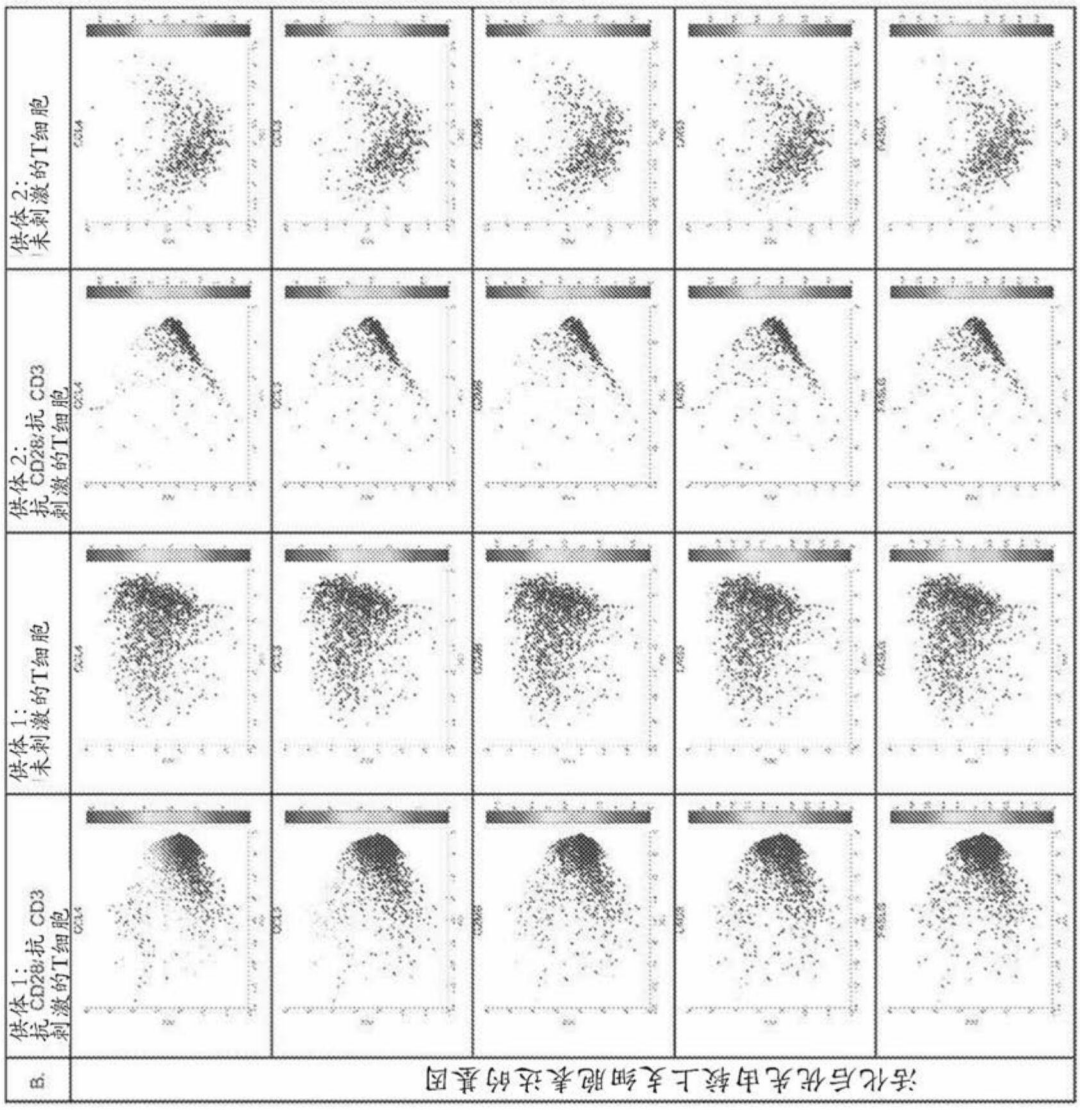

B.
活化后优先由较上支细胞表达的基因
供体 1:
抗 CD28i 抗 CD3
刺激的T细胞
供体 1:
未刺激的T细胞
供体 2:
抗 CD28i 抗 CD3
刺激的T细胞
供体 2:
未刺激的T细胞

图46B

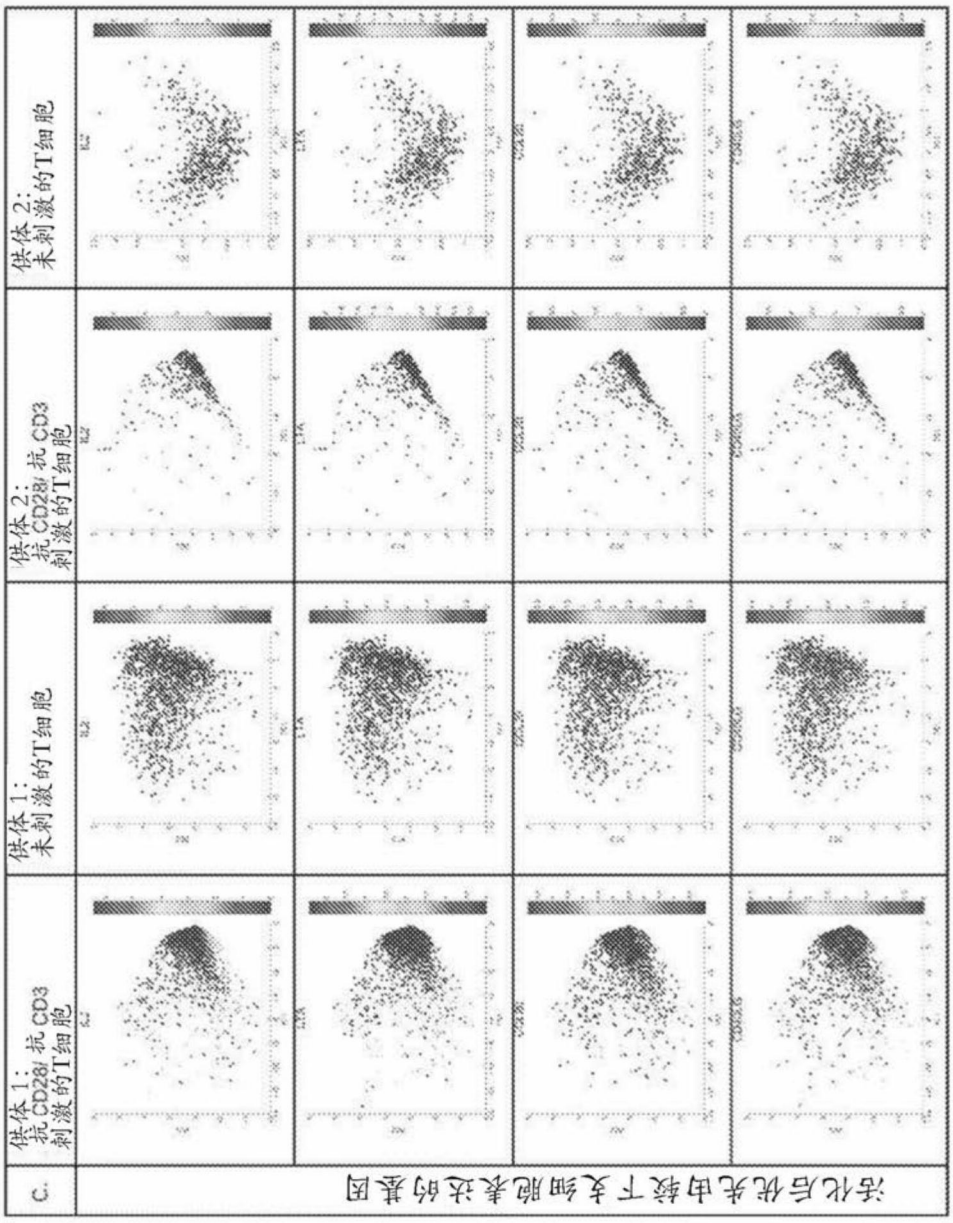

C.
活化后优先由较下支细胞表达的基因
供体 1:
抗 CD28/ 抗 CD3
刺激的T细胞
供体 1:
未刺激的T细胞
供体 2:
抗 CD28/ 抗 CD3
刺激的T细胞
供体 2:
未刺激的T细胞

图46C

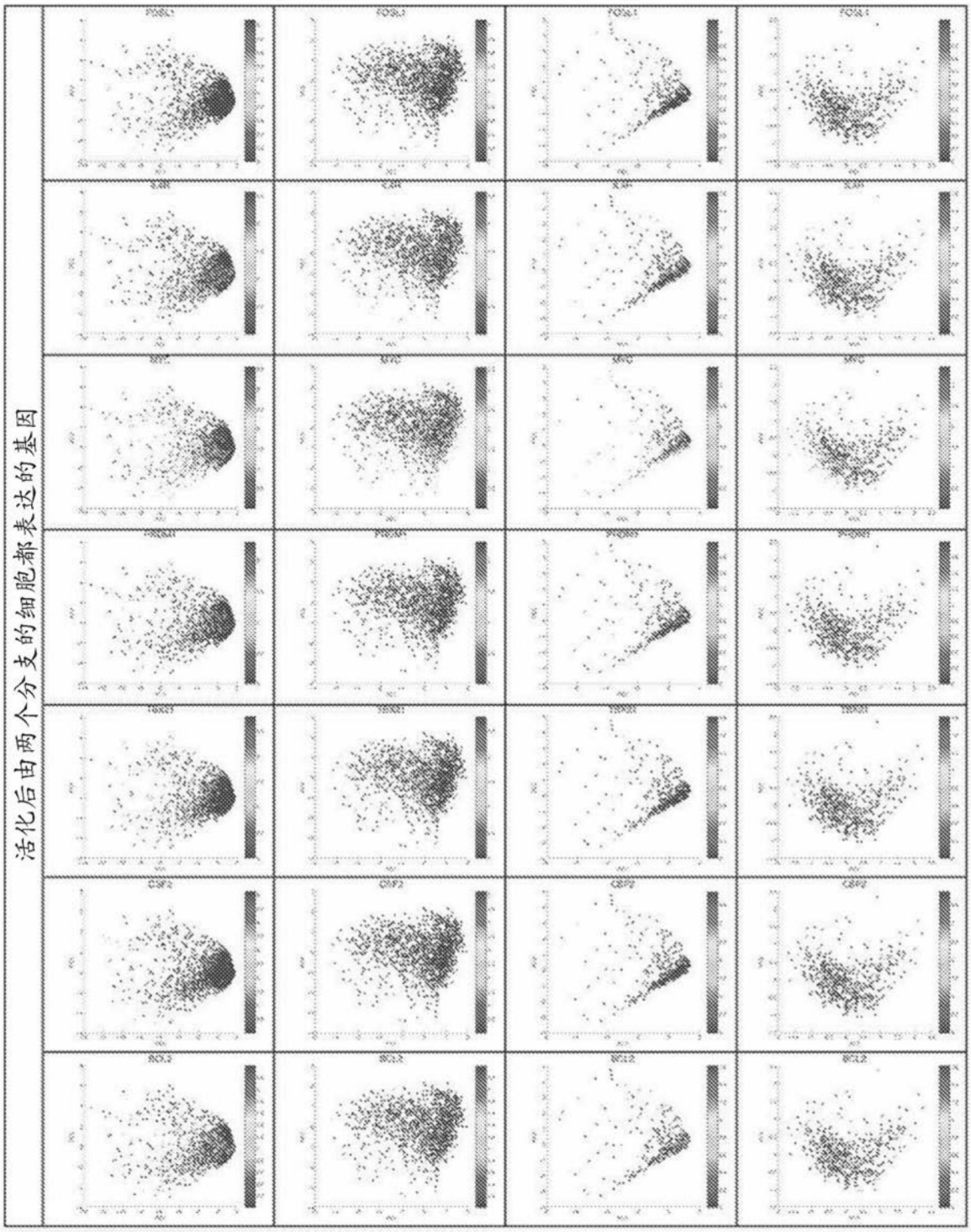


图46D

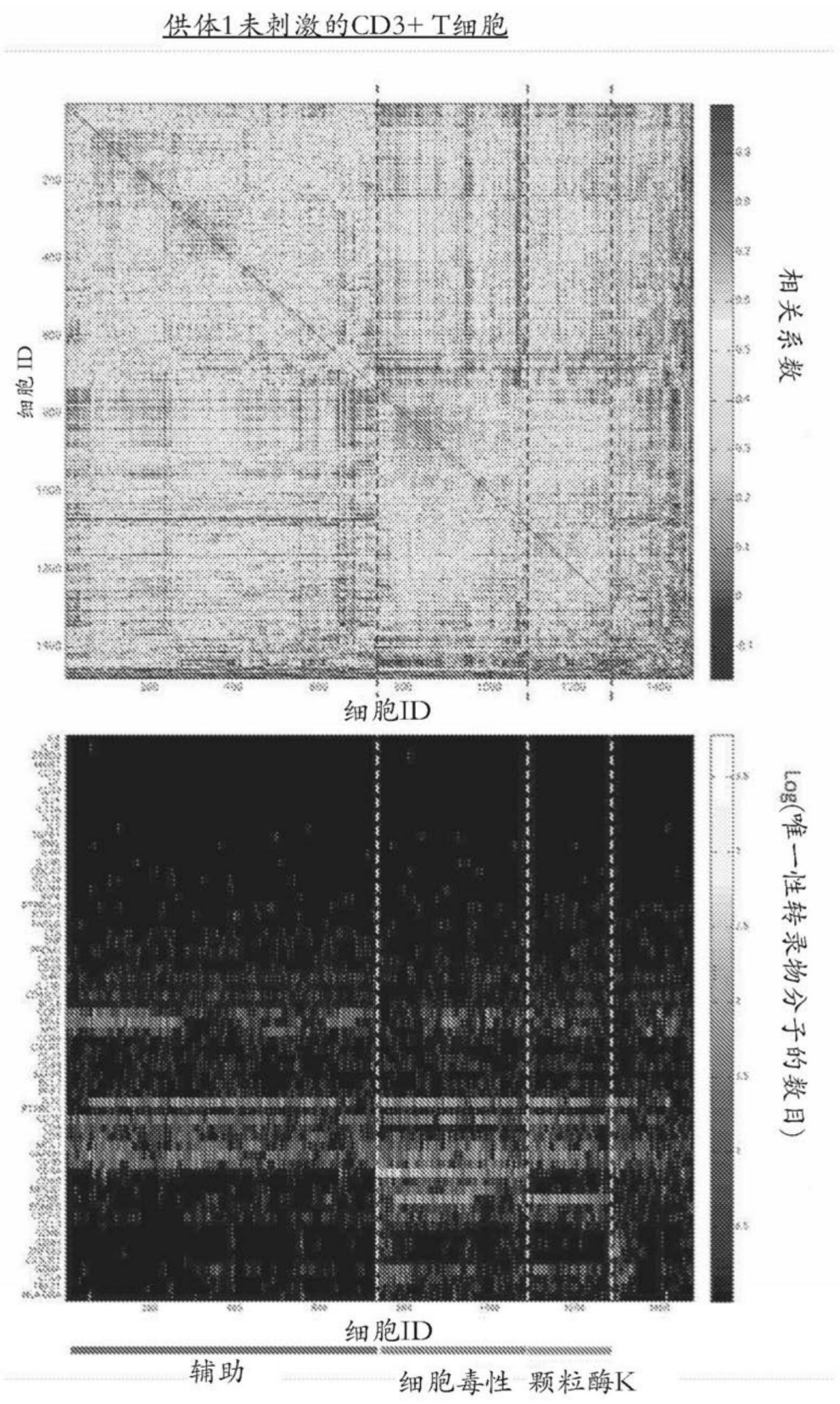

供体1未刺激的CD3+ T细胞
细胞ID
相关系数
细胞ID
Log(唯一性转录物分子的数目)
细胞ID
辅助
细胞毒性 颗粒酶K

图47

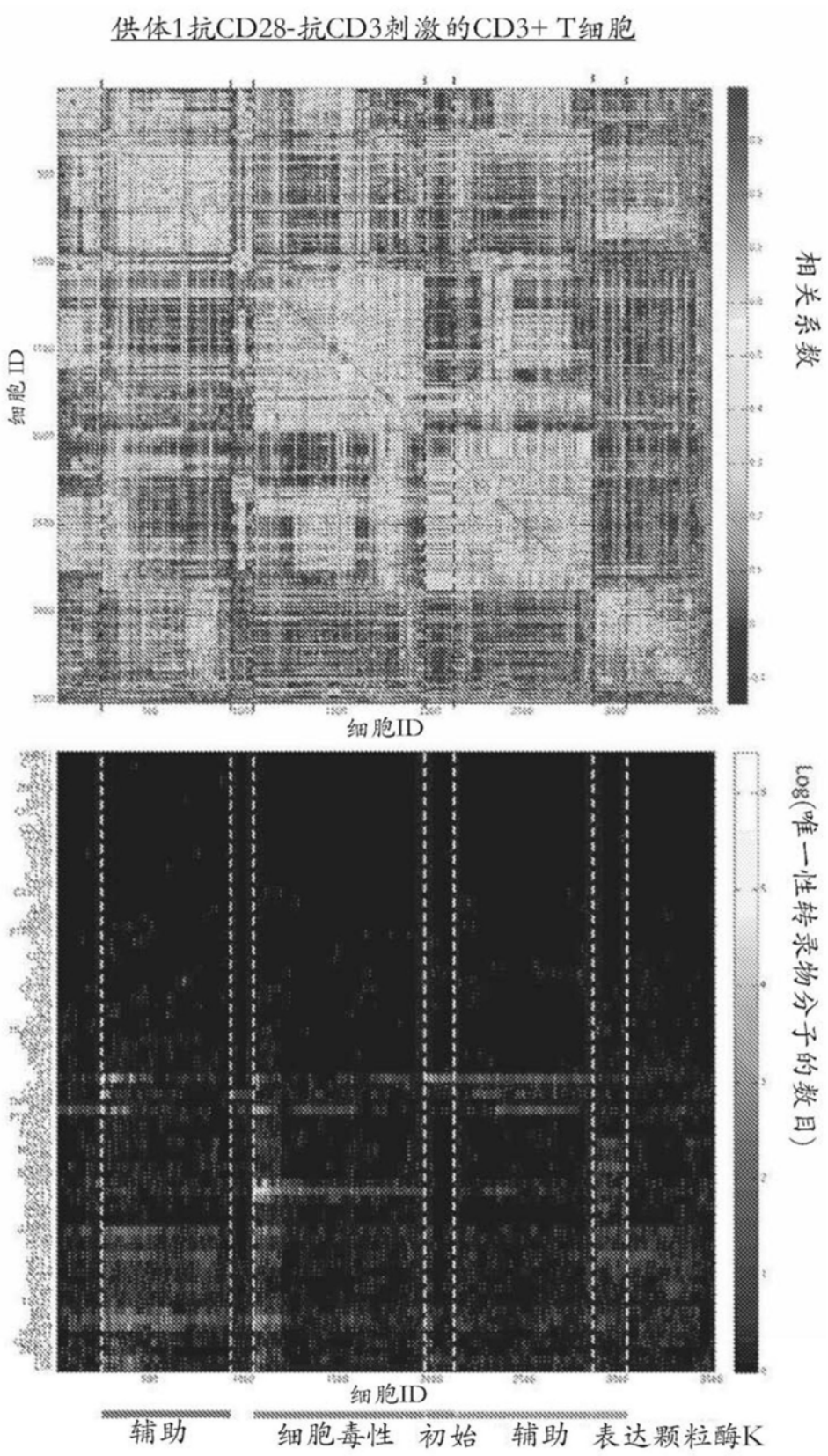

供体1抗CD28-抗CD3刺激的CD3+ T细胞
细胞ID
相关系数
细胞ID
log(唯一性转录物分子的数目)
辅助
细胞毒性
初始
辅助
表达颗粒酶K

图48

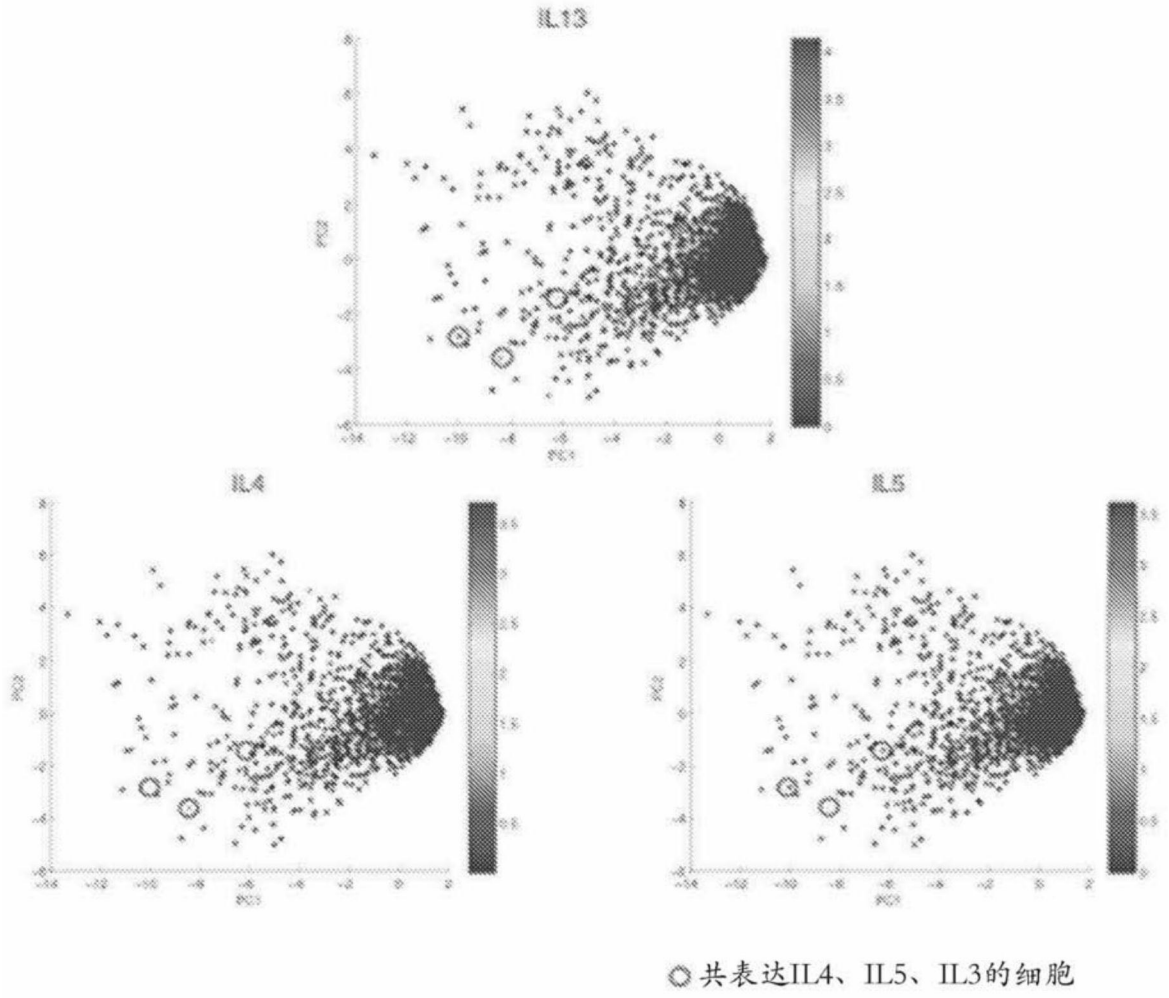

IL13
IL4
IL5
共表达IL4、IL5、IL3的细胞

图49A

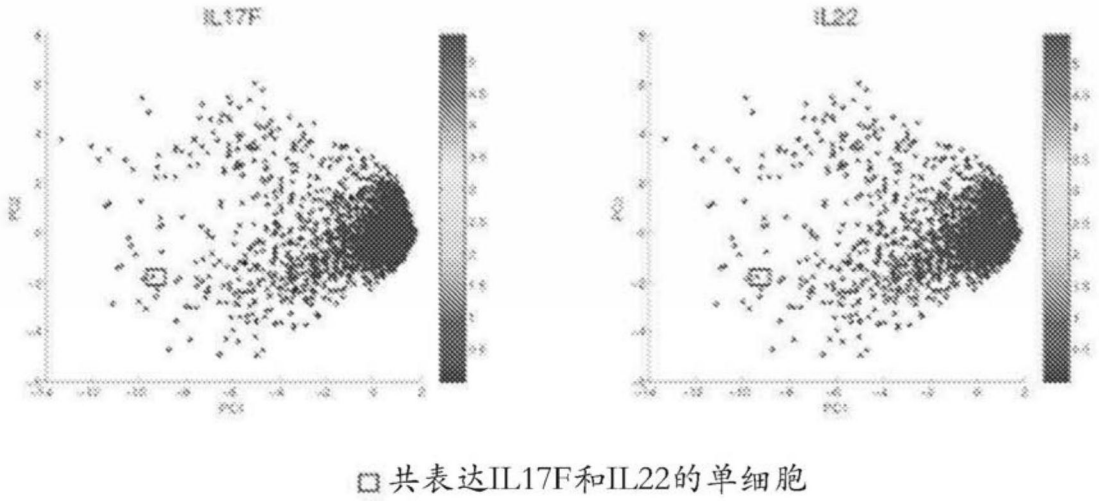

IL17F
IL22
共表达IL17F和IL22的单细胞

图49B

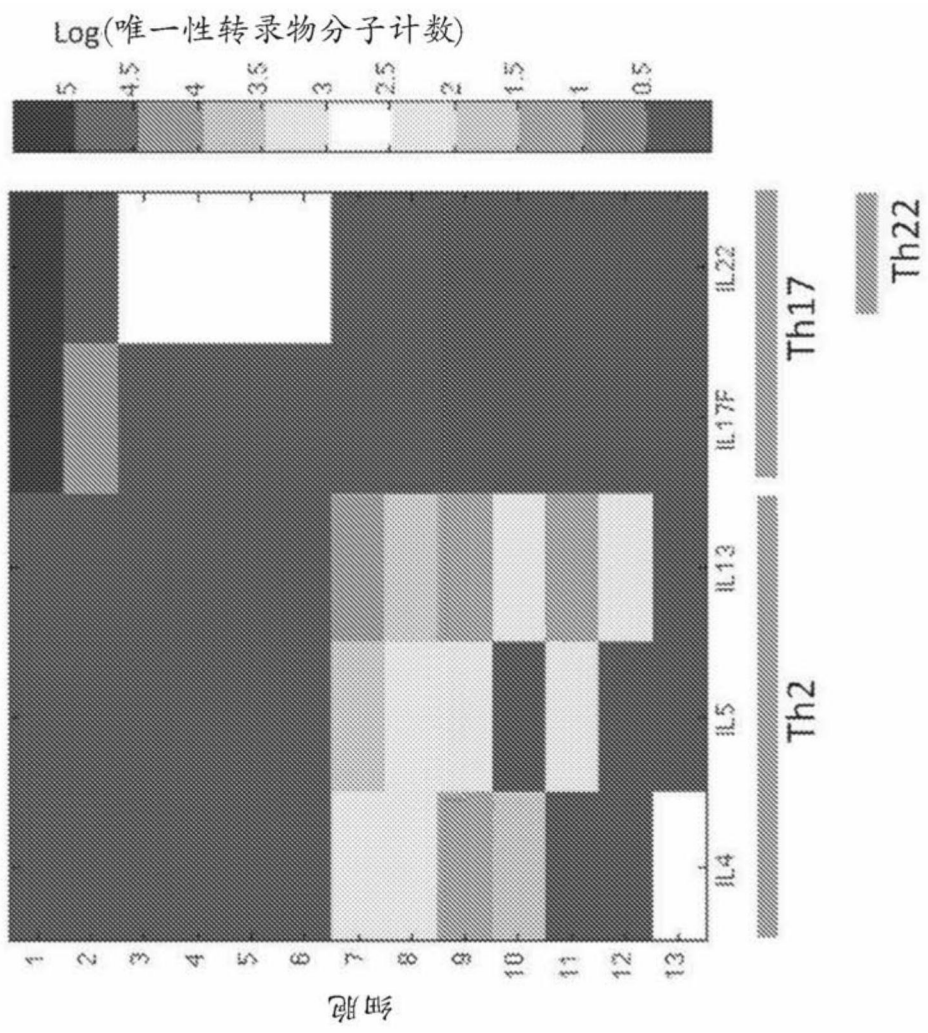

Log(唯一性转录物分子计数)
5
4.5
4
3.5
3
2.5
2
1.5
1
0.5
IL22
IL17F
IL13
IL5
IL4
Th17
Th22
Th2
1
2
3
4
5
6
7
8
9
10
11
12
13
细胞

图49C

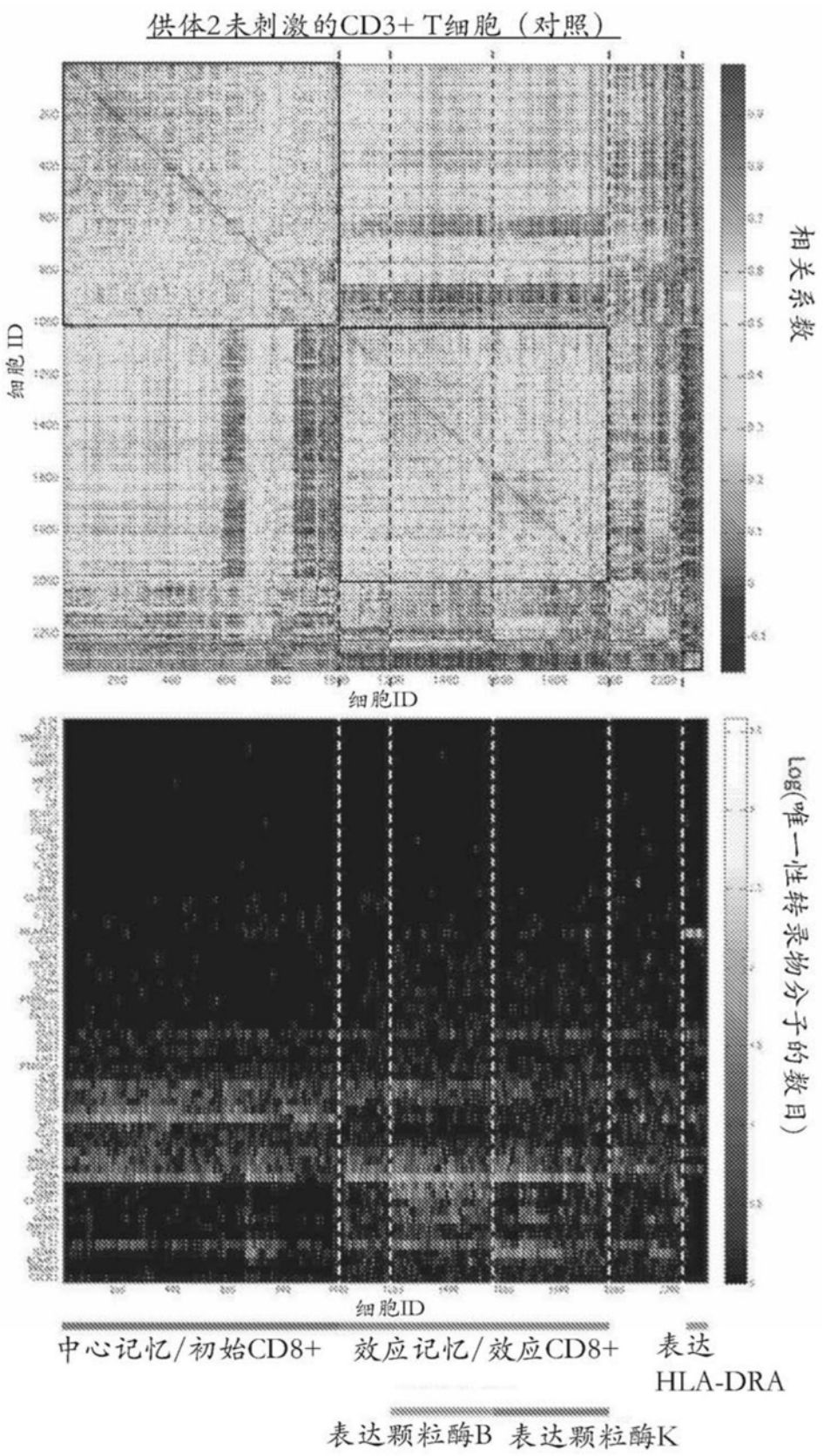

供体2未刺激的CD3+ T细胞（对照）
细胞ID
细胞ID
相关系数
细胞ID
细胞ID
Log(唯一性转录物分子的数目)
中心记忆/初始CD8+
效应记忆/效应CD8+
表达
HLA-DRA
表达颗粒酶B
表达颗粒酶K

图50A

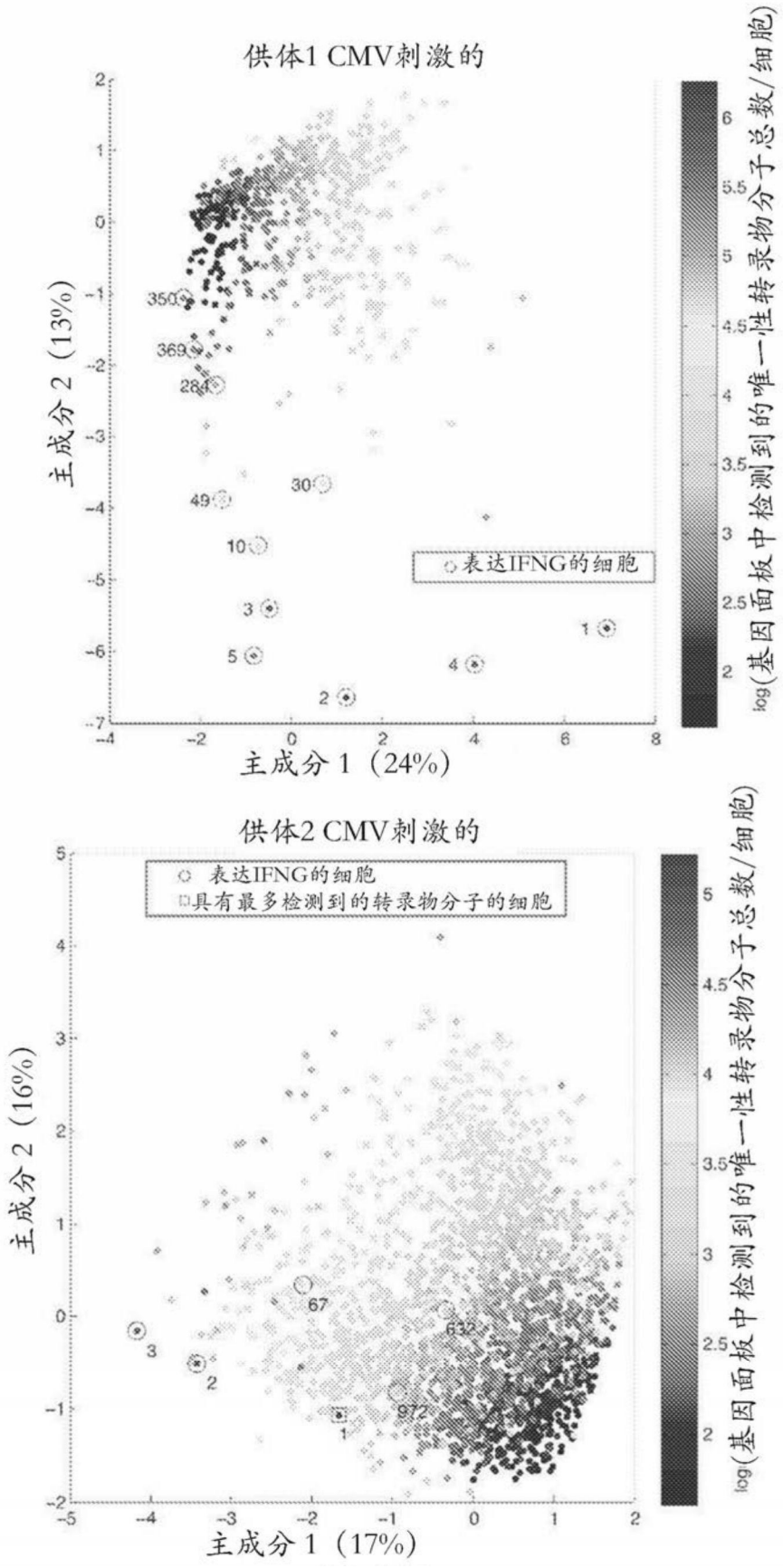

供体1 CMV刺激的
主成分 2（13%）
主成分 1（24%）
表达IFNG的细胞
log(基因面板中检测到的唯一性转录物分子总数/细胞)
供体2 CMV刺激的
表达IFNG的细胞
具有最多检测到的转录物分子的细胞
主成分 2（16%）
主成分 1（17%）
log(基因面板中检测到的唯一性转录物分子总数/细胞)

图50B

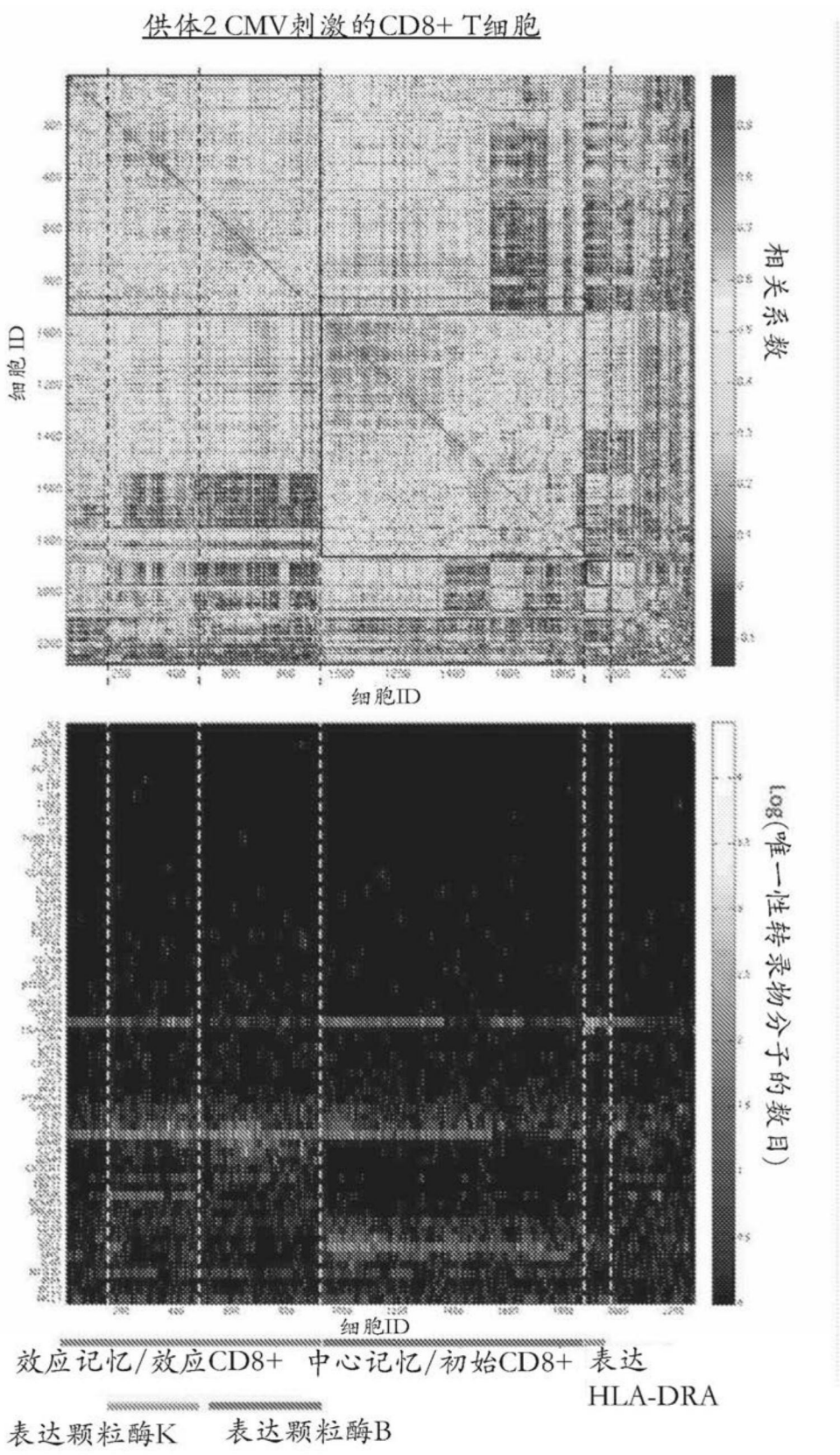

供体2 CMV刺激的CD8+ T细胞
细胞ID
相关系数
细胞ID
log(唯一性转录物分子的数目)
效应记忆/效应CD8+
中心记忆/初始CD8+
表达HLA-DRA
表达颗粒酶K
表达颗粒酶B

图51

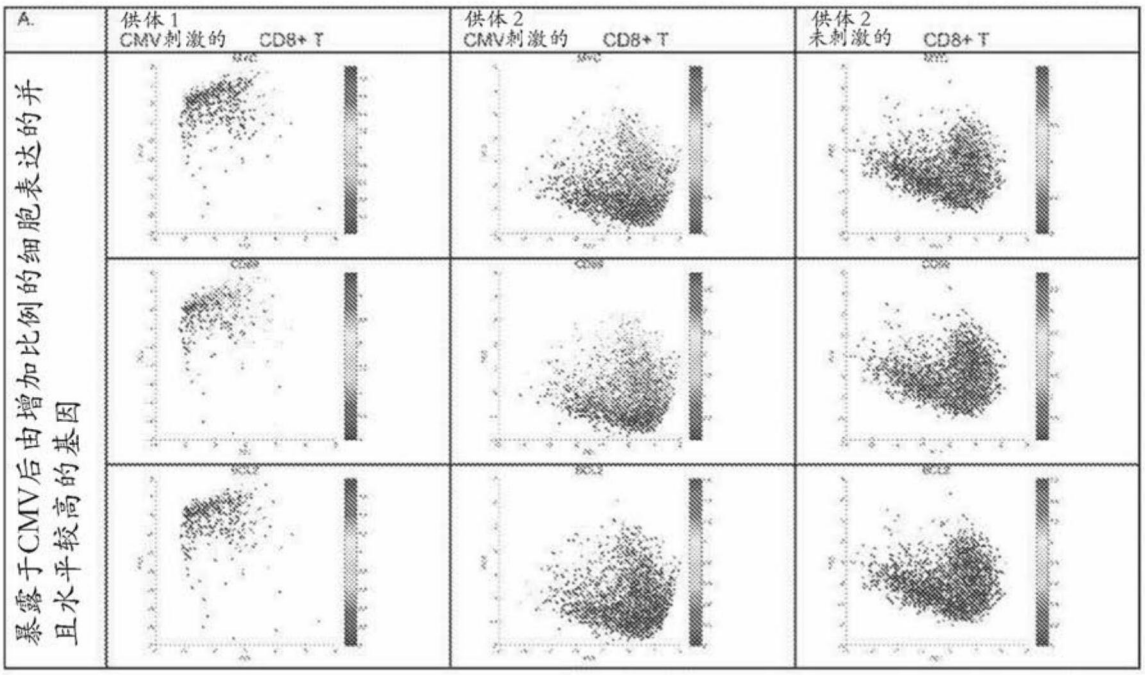

A.
供体1
CMV刺激的 CD8+ T
供体2
CMV刺激的 CD8+ T
供体2
未刺激的 CD8+ T
暴露于CMV后由增加比例的细胞表达的并且水平较高的基因

图52A

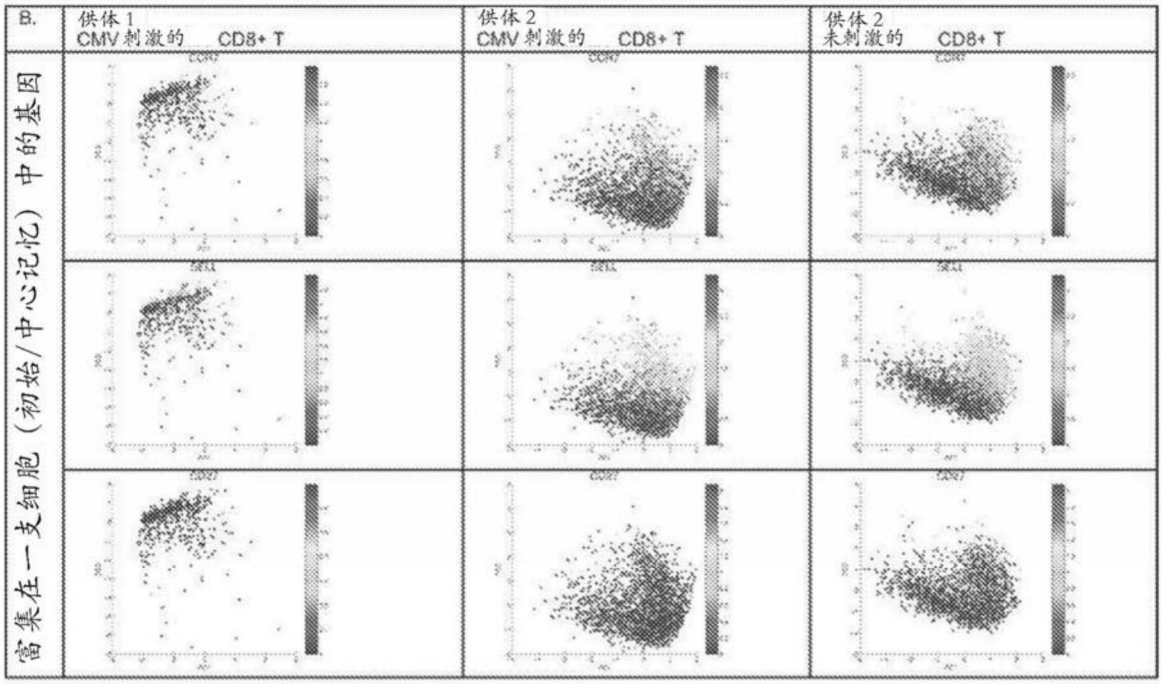

B.
供体1
CMV刺激的 CD8+ T
供体2
CMV刺激的 CD8+ T
供体2
未刺激的 CD8+ T
富集在一支细胞（初始/中心记忆）中的基因

图52B

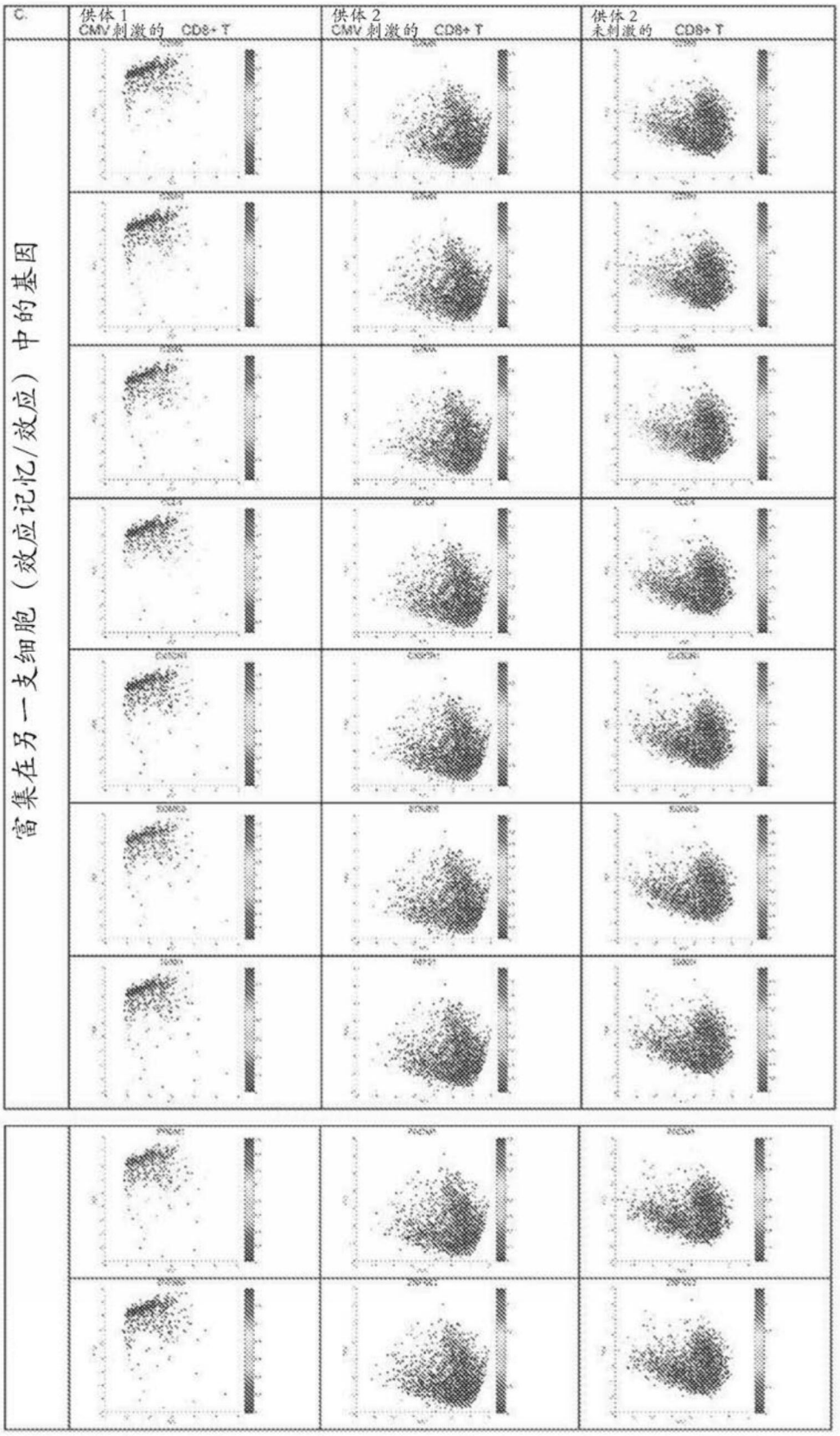

C.
供体 1
CMV 刺激的 CD8+ T
供体 2
CMV 刺激的 CD8+ T
供体 2
未刺激的 CD8+ T
富集在另一支细胞（效应记忆/效应）中的基因

图52C

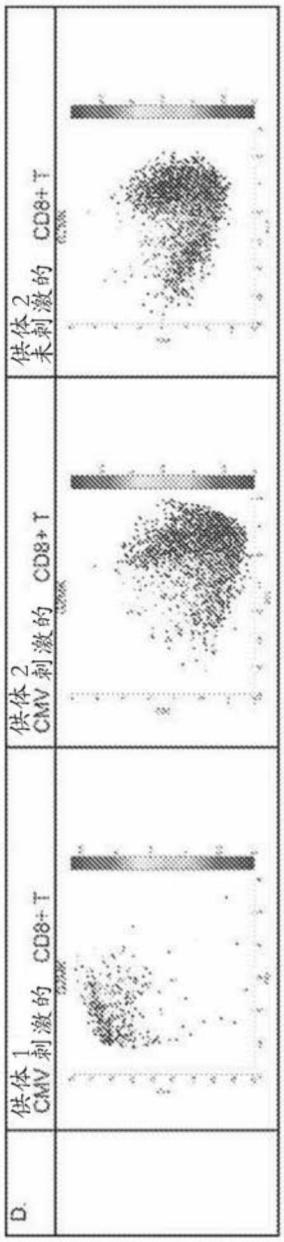


图52D

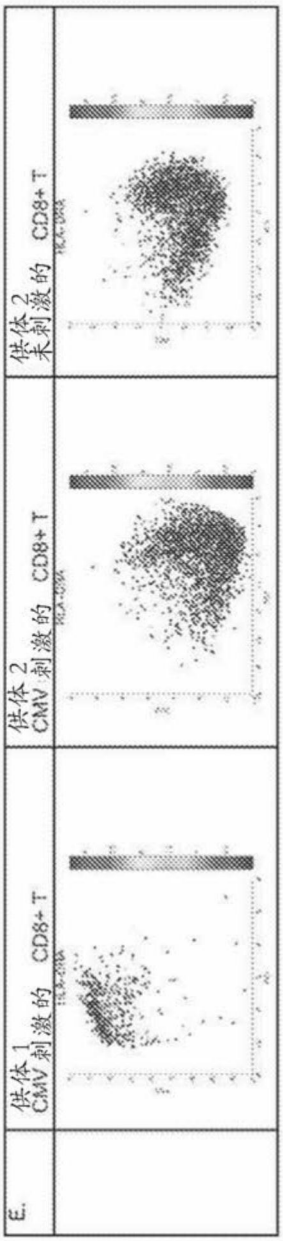


图52E

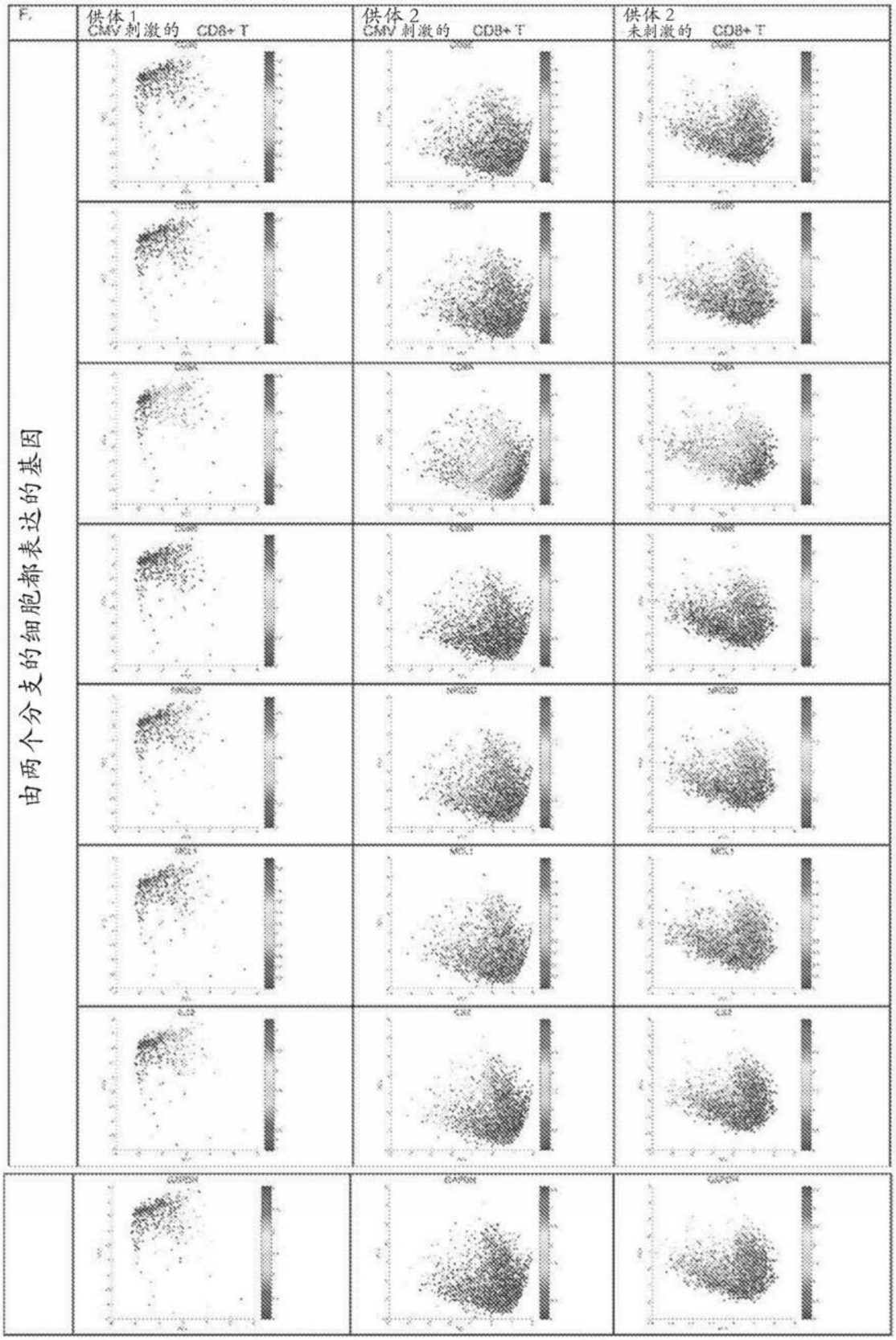

F.
供体 1
CMV 刺激的 CD8+ T
供体 2
CMV 刺激的 CD8+ T
供体 2
未刺激的 CD8+ T
由两个分支的细胞都表达的基因

图52F

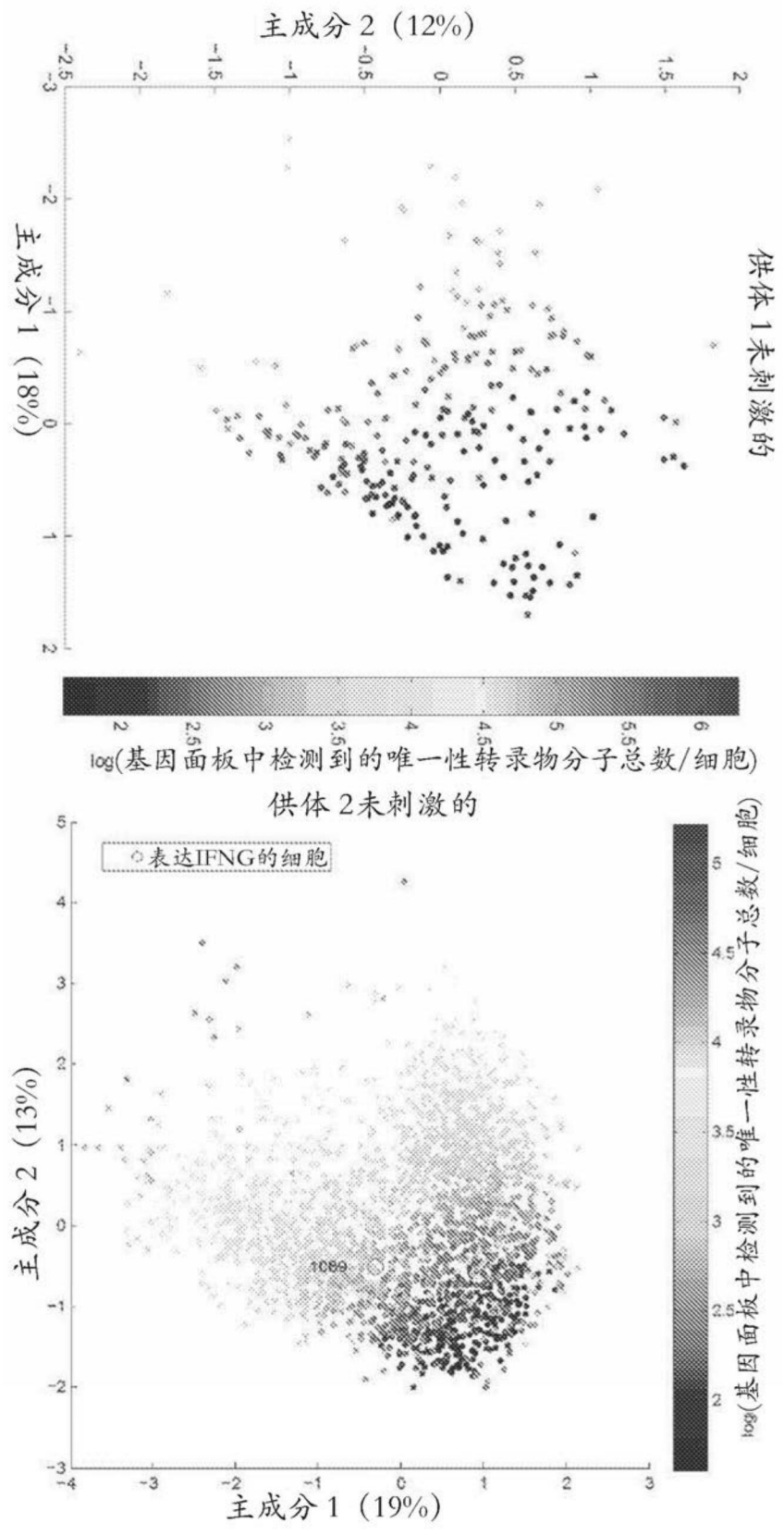

主成分 2（12%）
主成分 1（18%）
供体 1未刺激的
log(基因面板中检测到的唯一性转录物分子总数/细胞)
供体 2未刺激的
表达IFNG的细胞
1089
主成分 2（13%）
主成分 1（19%）
log(基因面板中检测到的唯一性转录物分子总数/细胞)

图53

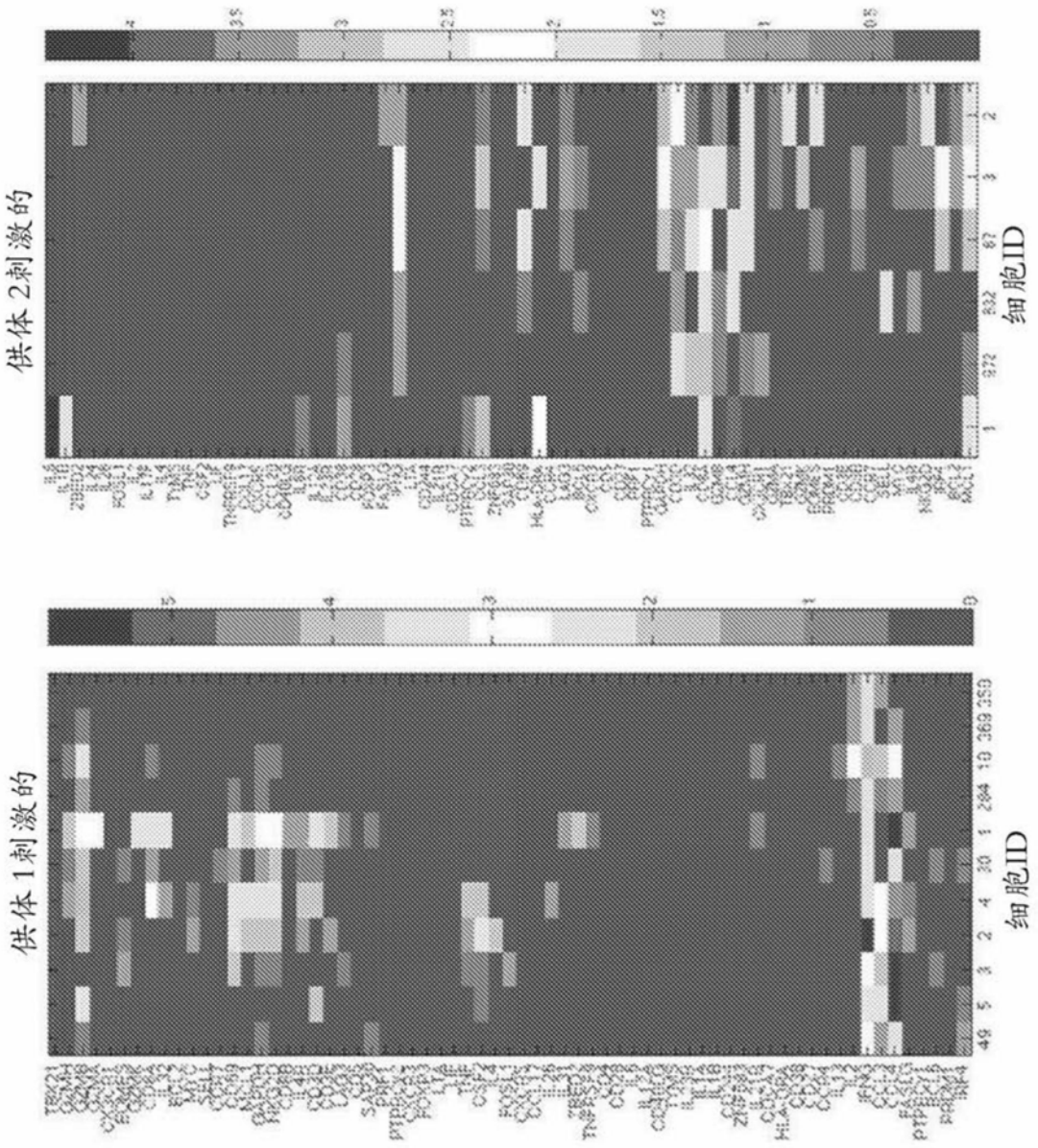


图54

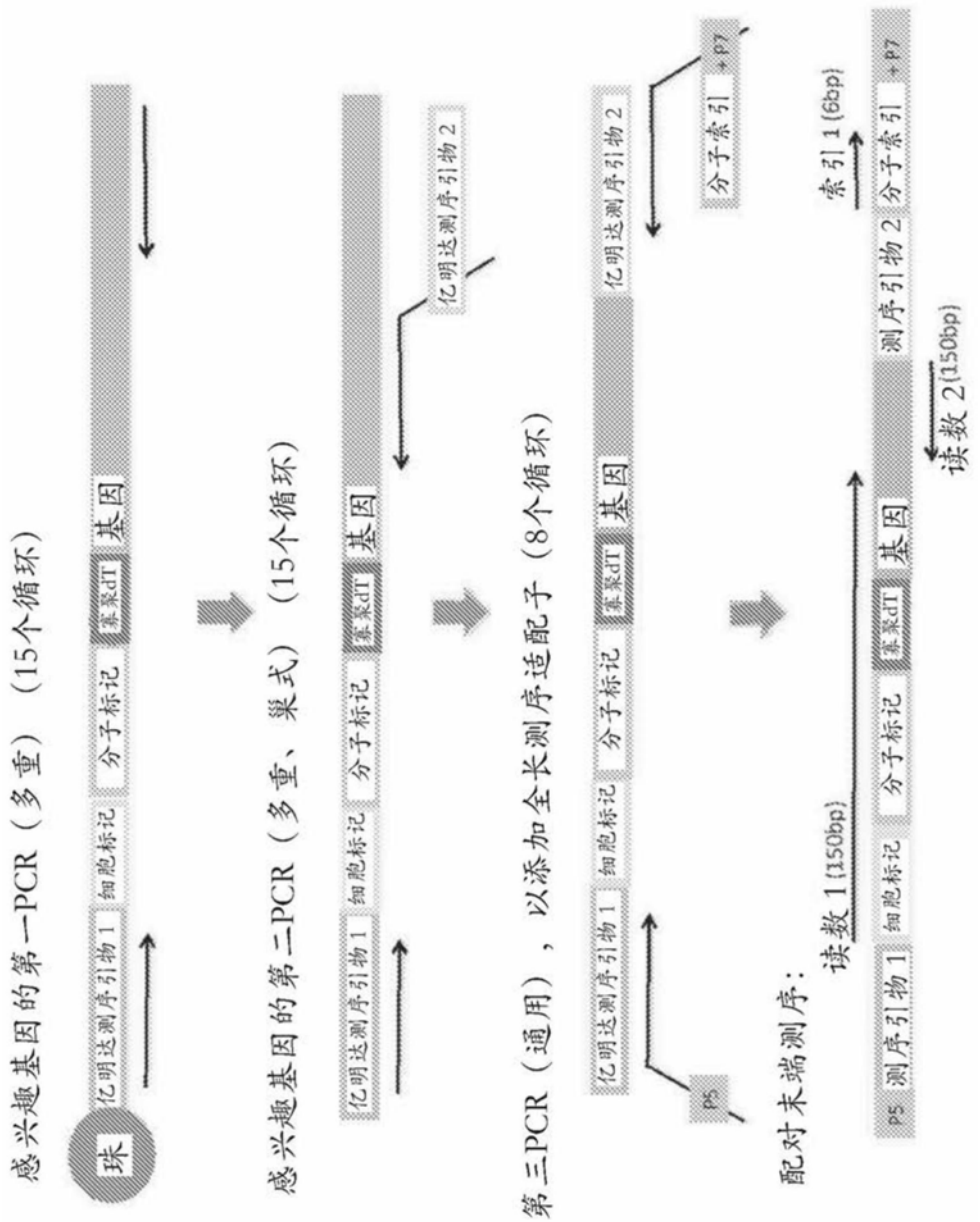

感兴趣基因的第一PCR（多重）（15个循环）
珠
亿明达测序引物1
细胞标记
分子标记
重聚dT
基因
感兴趣基因的第二PCR（多重、巢式）（15个循环）
亿明达测序引物2
第三PCR（通用），以添加全长测序适配子（8个循环）
P5
P7
分子索引
索引1(6bp)
配对末端测序：
读数1(150bp)
测序引物1
测序引物2
读数2(150bp)

图55

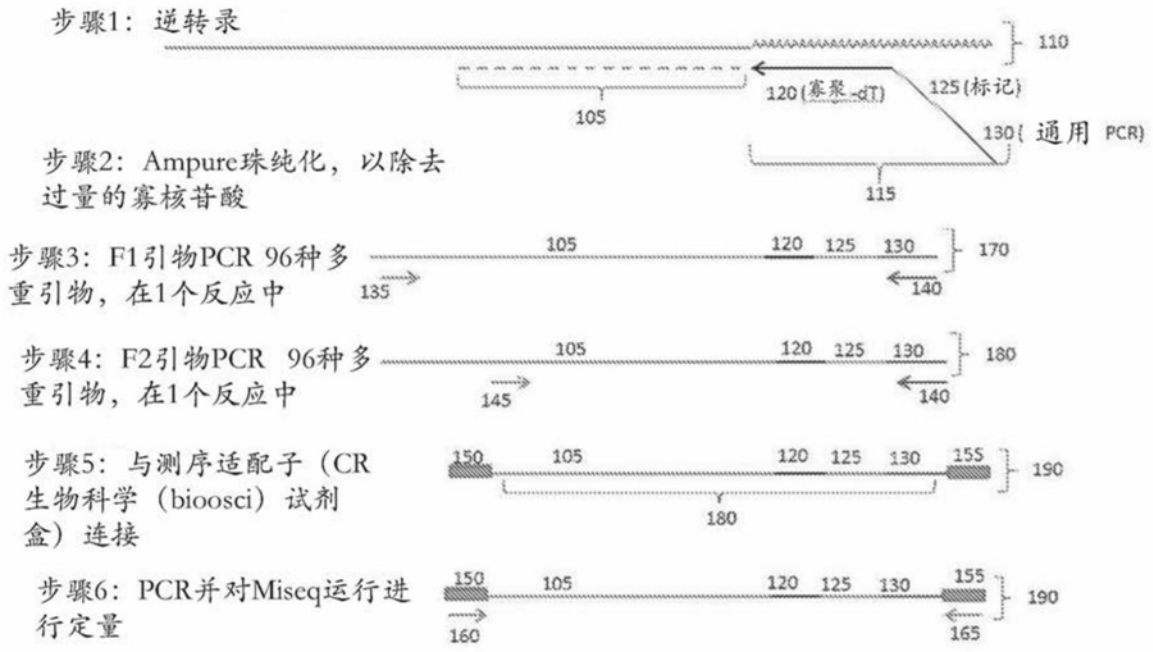

步骤1：逆转录
110
120(寡聚-dT)
125(标记)
105
130(通用 PCR)
115
步骤2：Ampure珠纯化，以除去过量的寡核苷酸
步骤3：F1引物PCR 96种多重引物，在1个反应中
105
120 125 130
170
135
140
步骤4：F2引物PCR 96种多重引物，在1个反应中
105
120 125 130
180
145
140
步骤5：与测序适配子（CR生物科学（bioosci）试剂盒）连接
150
105
120 125 130
155
190
180
步骤6：PCR并对Miseq运行进行定量
150
105
120 125 130
155
190
160
165

图56

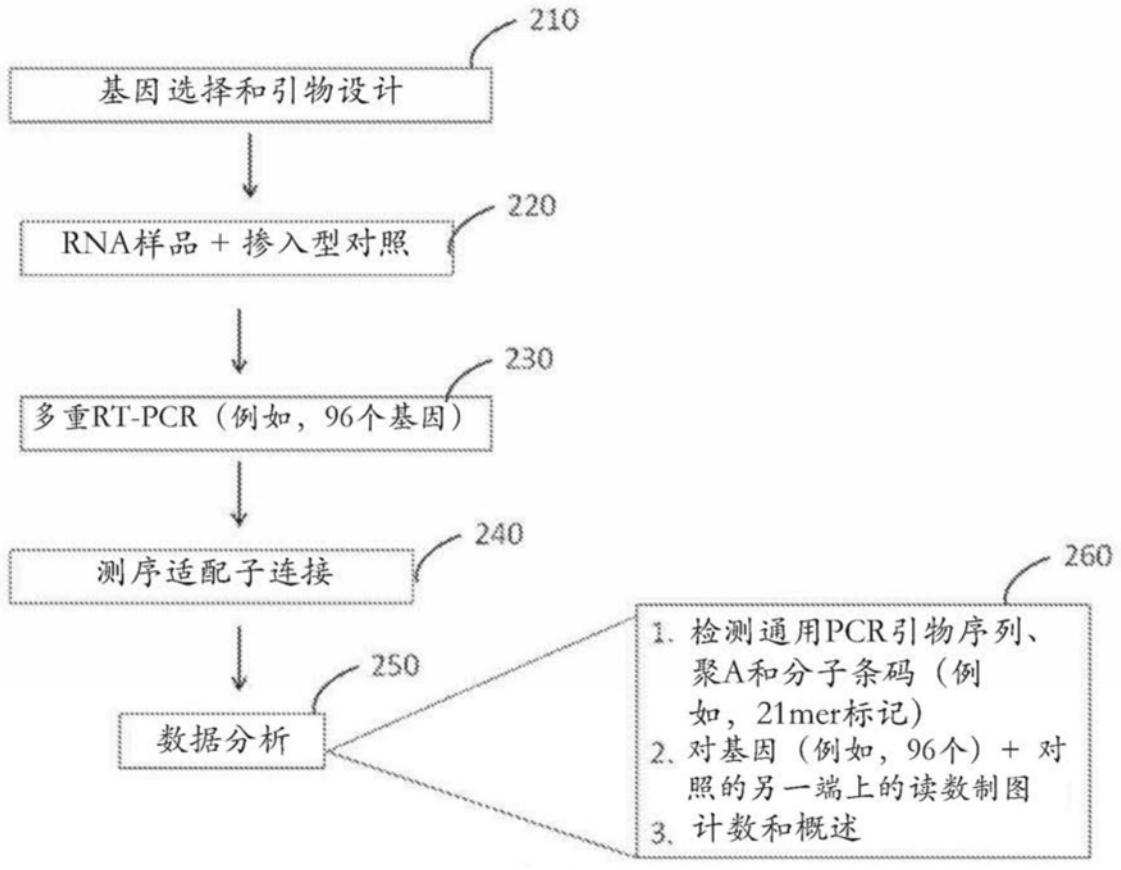

210
基因选择和引物设计
220
RNA样品 + 掺入型对照
230
多重RT-PCR（例如，96个基因）
240
测序适配子连接
250
数据分析
260
1. 检测通用PCR引物序列、聚A和分子条码（例如，21mer标记）
2. 对基因（例如，96个）+ 对照的另一端上的读数制图
3. 计数和概述

图57

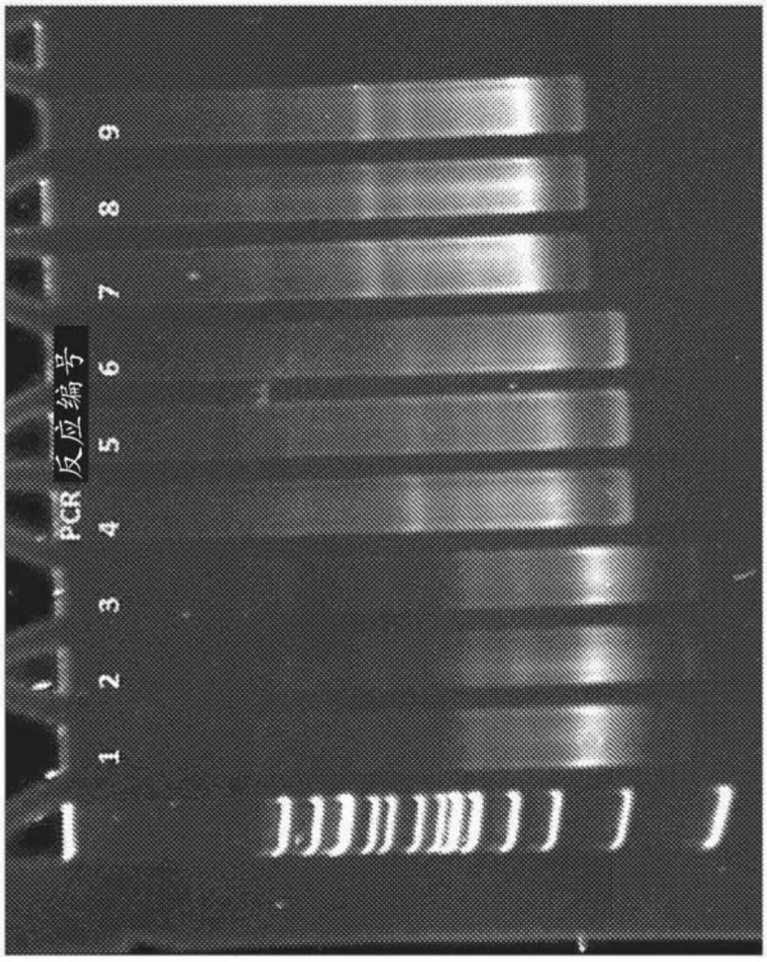


图58A

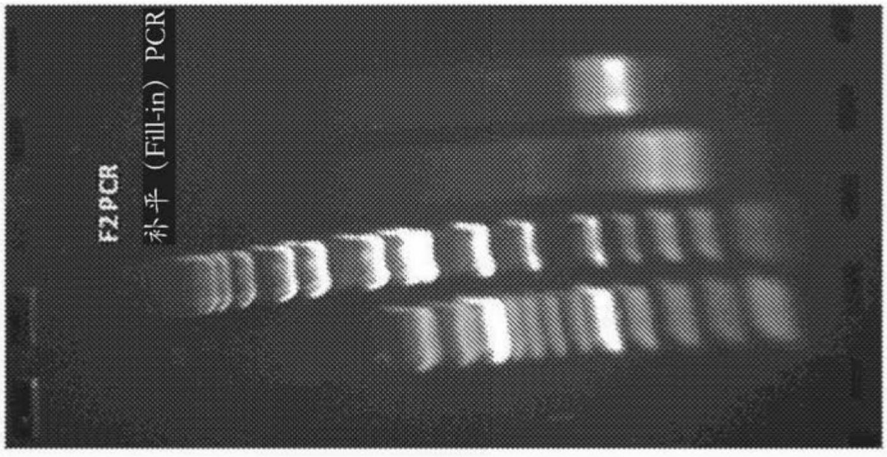


图58B

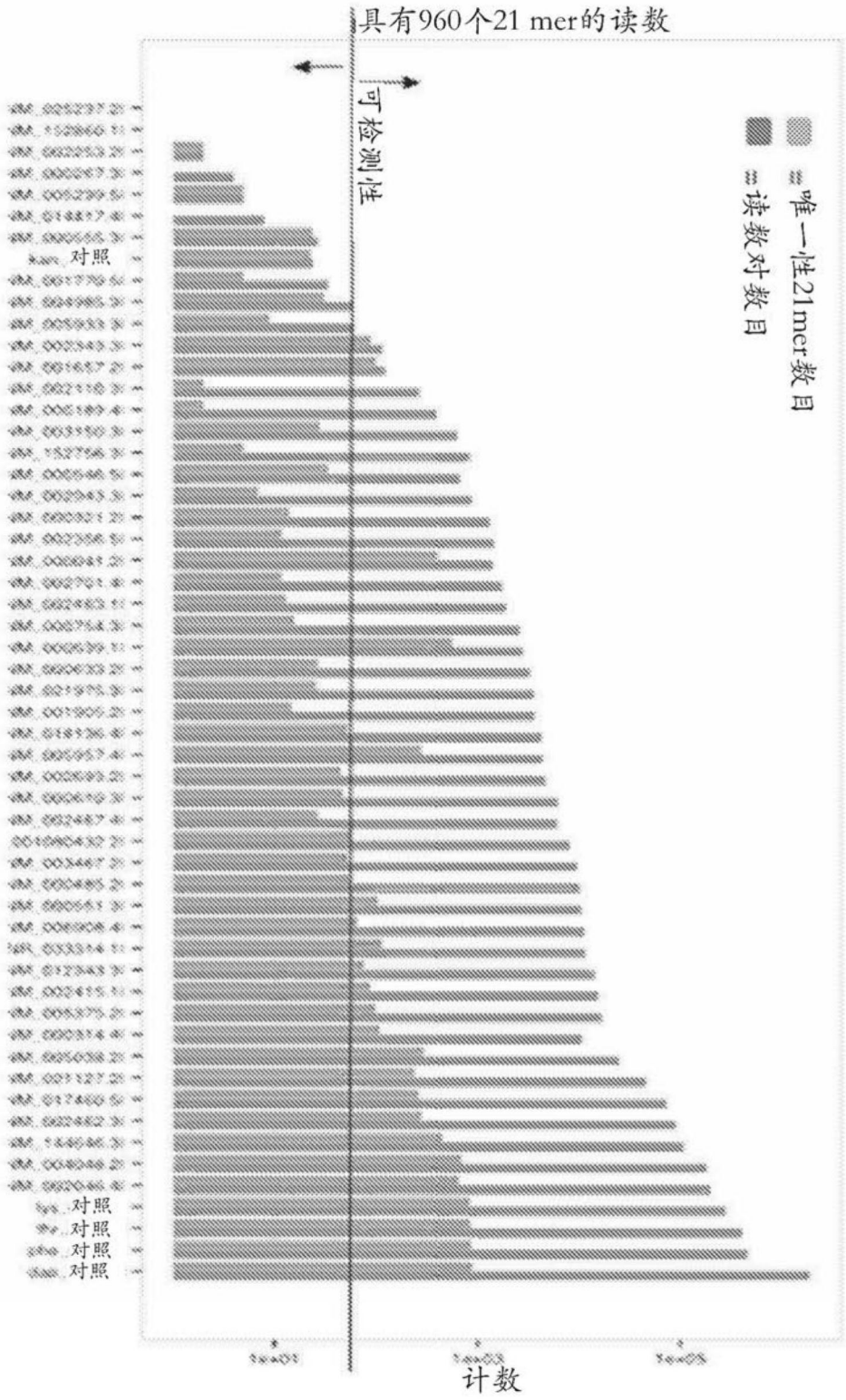


图59

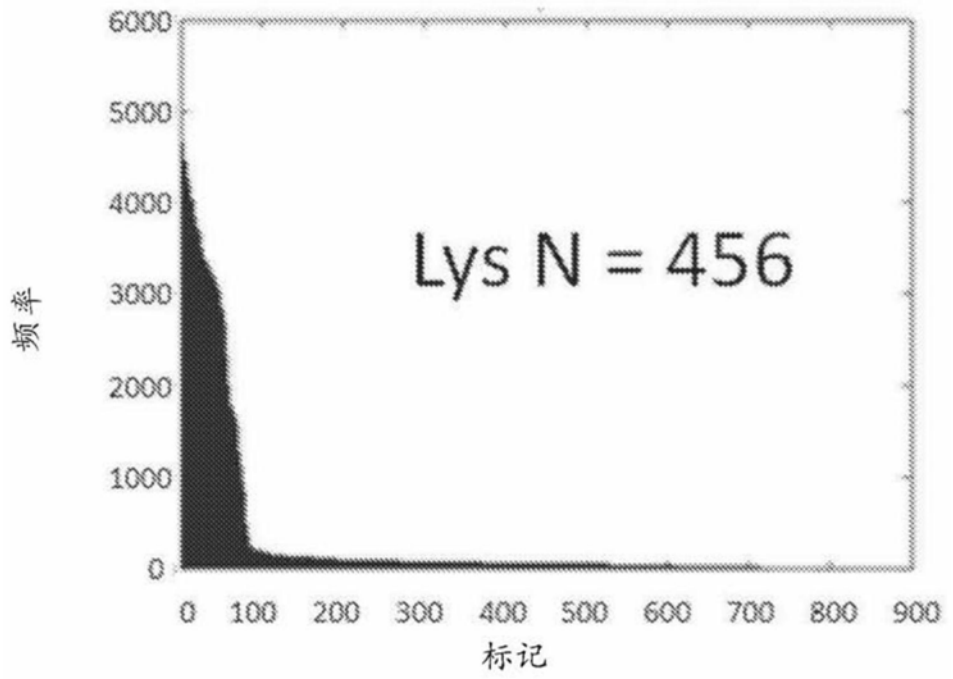

Lys N = 456
频率
6000
5000
4000
3000
2000
1000
0
0 100 200 300 400 500 600 700 800 900
标记

图60A

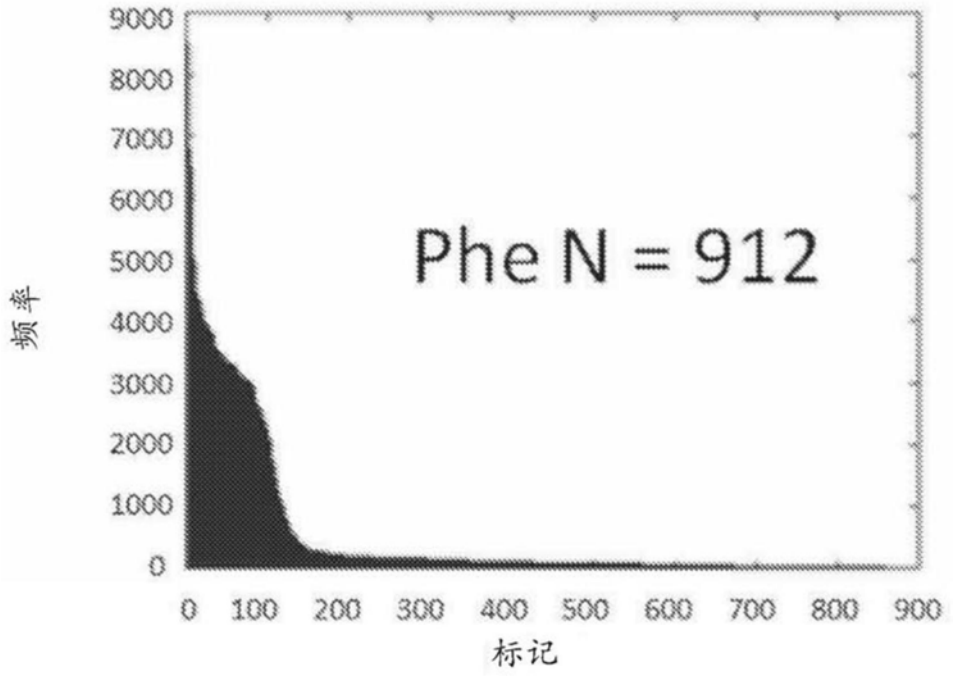

Phe N = 912
频率
9000
8000
7000
6000
5000
4000
3000
2000
1000
0
0 100 200 300 400 500 600 700 800 900
标记

图60B

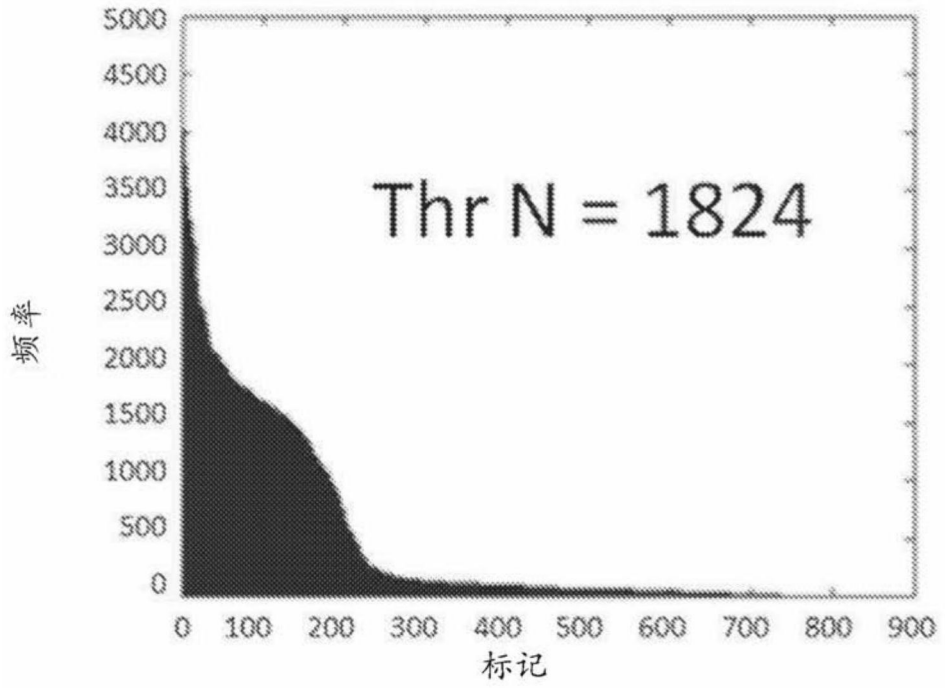

Thr N = 1824
频率
标记
5000
4500
4000
3500
3000
2500
2000
1500
1000
500
0
0 100 200 300 400 500 600 700 800 900

图60C

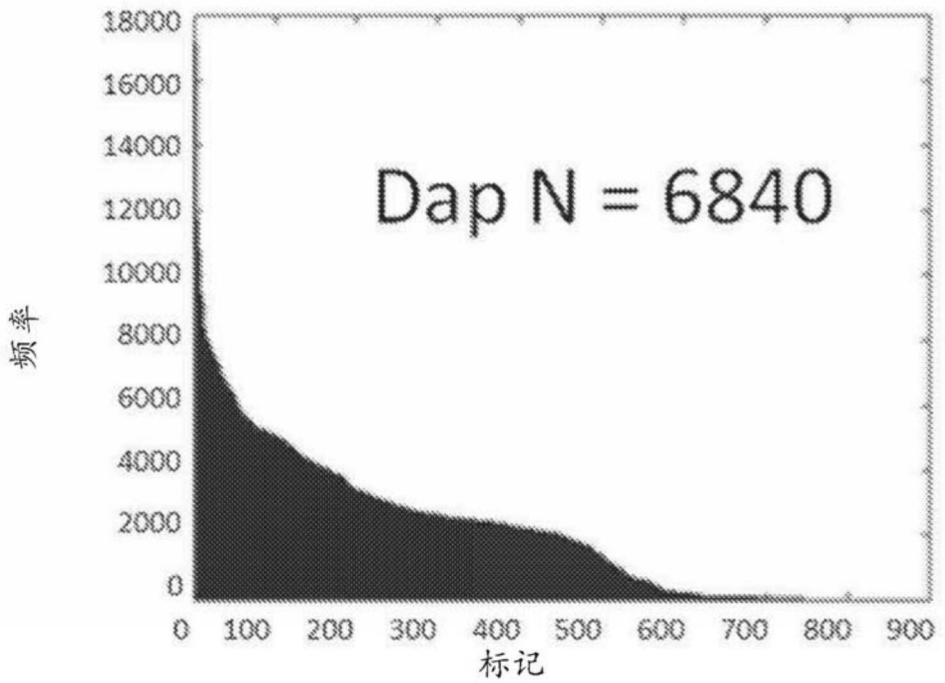

Dap N = 6840
频率
标记
18000
16000
14000
12000
10000
8000
6000
4000
2000
0
0 100 200 300 400 500 600 700 800 900

图60D

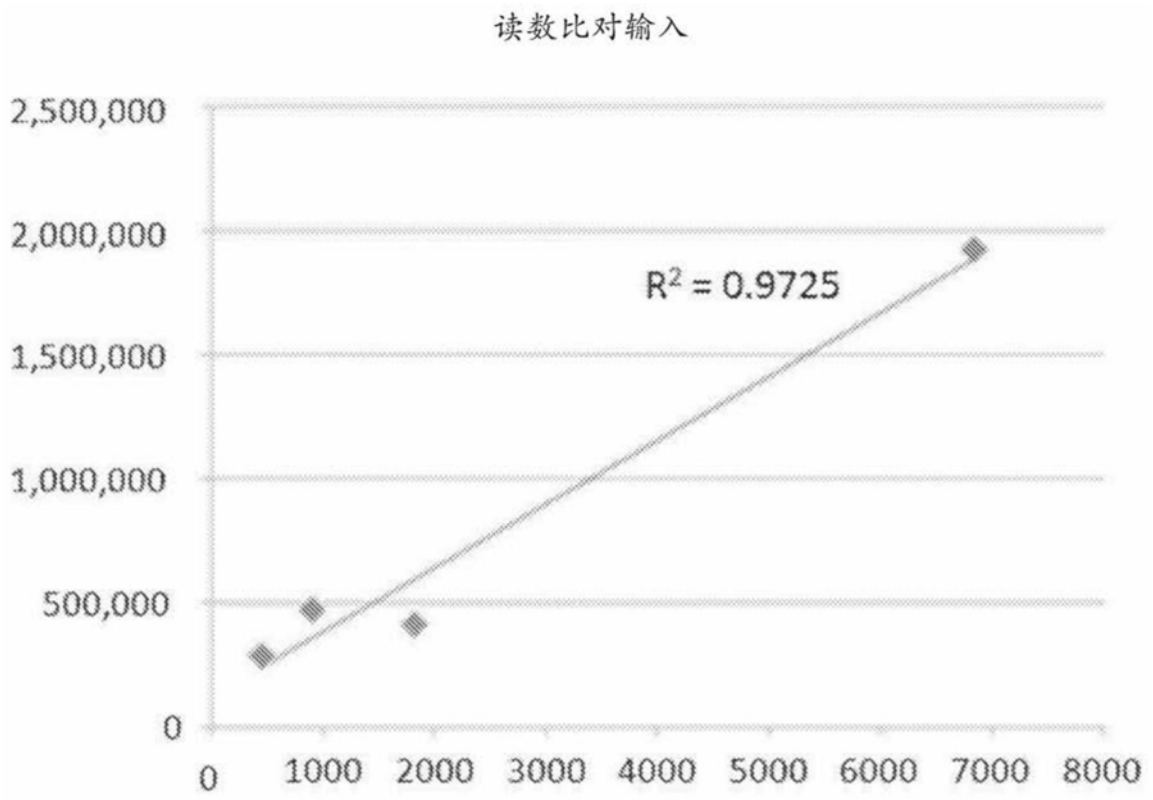

读数比对输入
2,500,000
2,000,000
1,500,000
1,000,000
500,000
0
$R^2 = 0.9725$
0
1000
2000
3000
4000
5000
6000
7000
8000

图60E

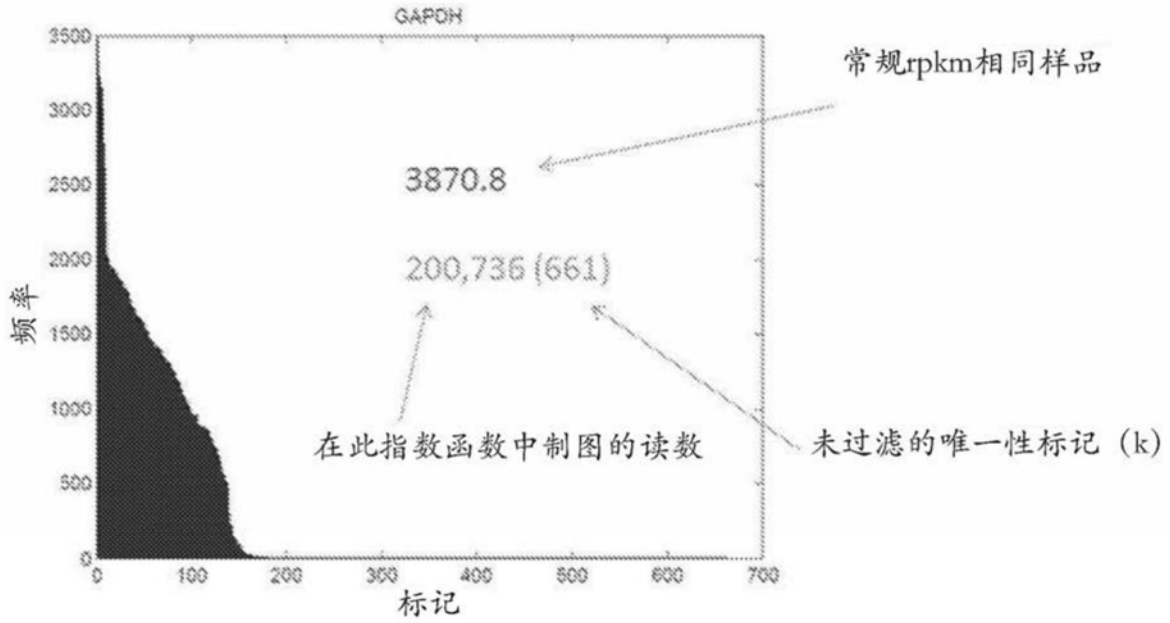

GAPDH
频率
3500
3000
2500
2000
1500
1000
500
0
0
100
200
300
400
500
600
700
标记
常规rpkm相同样品
3870.8
200,736 (661)
在此指数函数中制图的读数
未过滤的唯一性标记（k）

图61

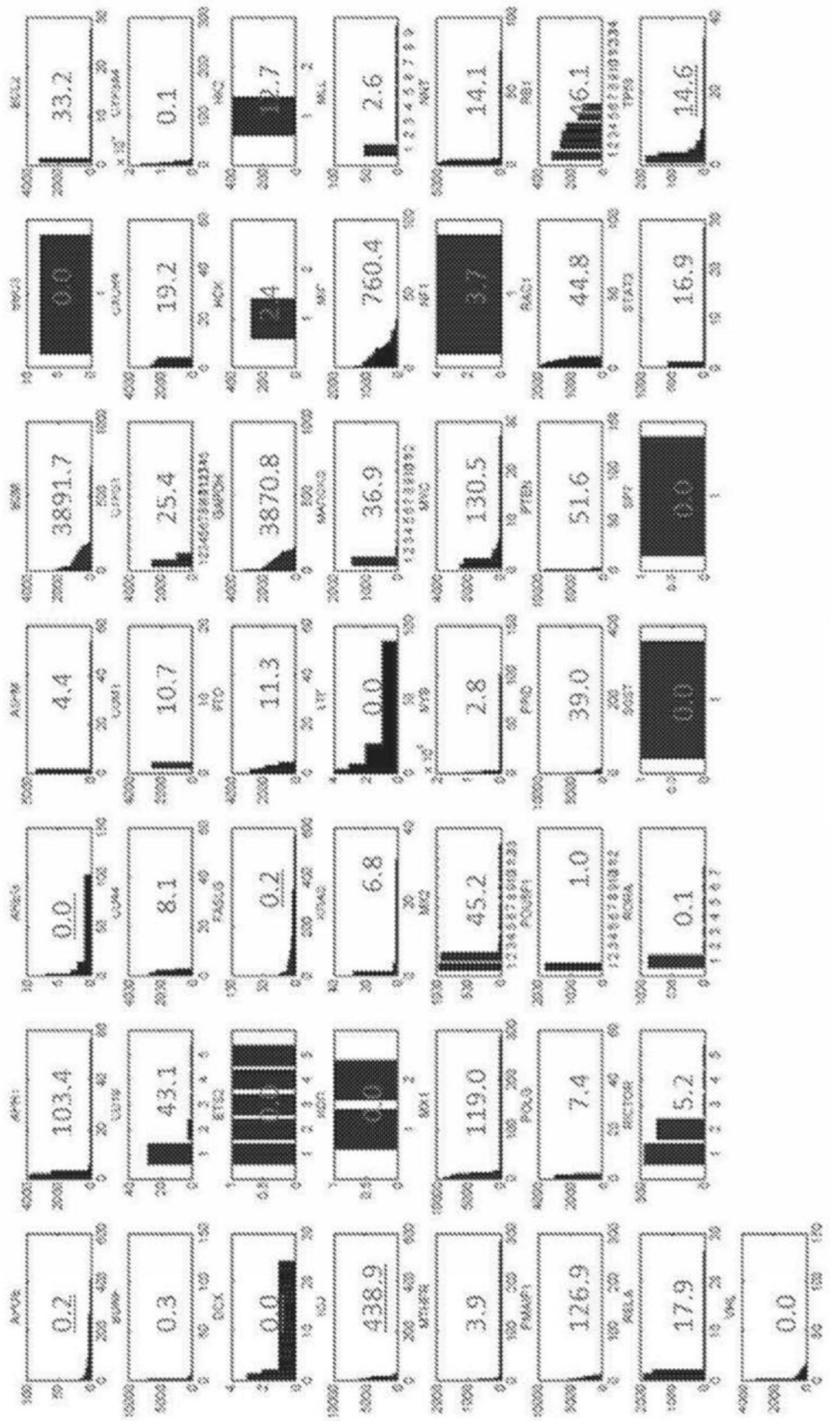

33.2
0.1
1.7
2.6
14.1
46.1
14.6
0.0
19.2
24
760.4
3.7
44.8
16.9
3891.7
25.4
3870.8
36.9
130.5
51.6
0.0
4.4
10.7
11.3
0.0
2.8
39.0
0.0
0.0
8.1
0.2
6.8
45.2
1.0
0.1
103.4
43.1
0.0
0.0
119.0
7.4
5.2
0.2
0.3
0.0
438.9
3.9
126.9
17.9
0.0

图62

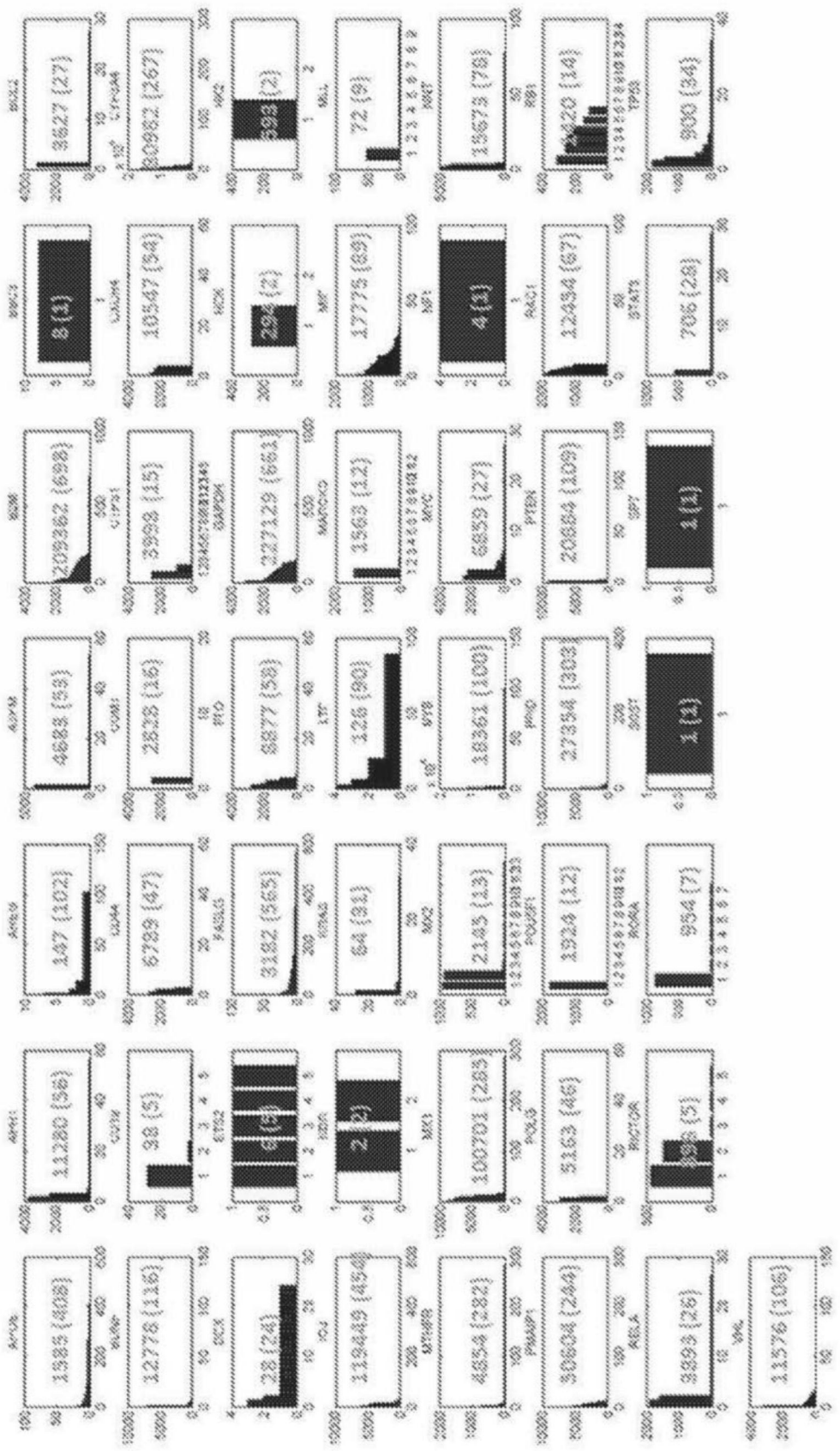


图63

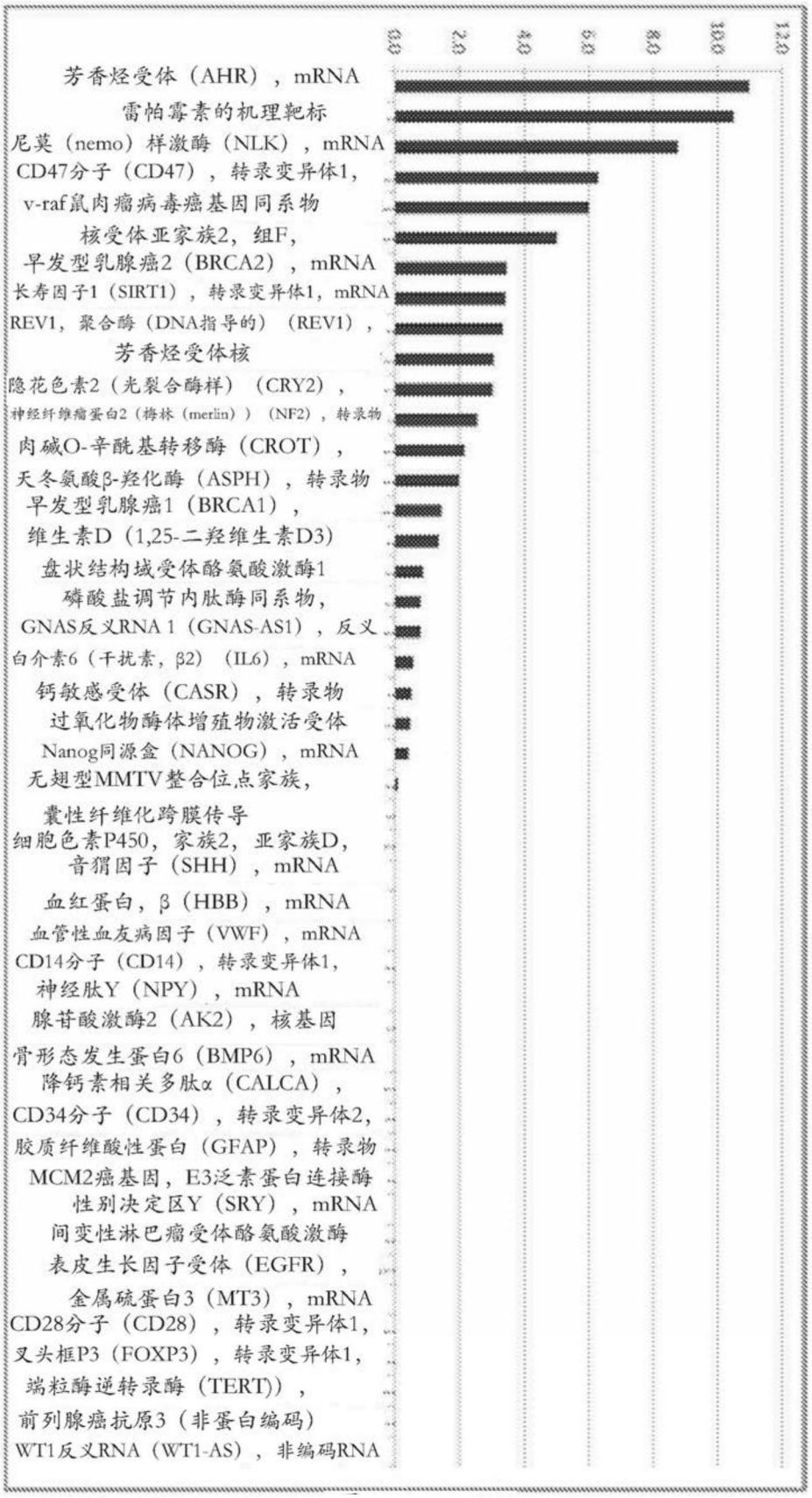

0.0
20
40
60
80
100
120
芳香烃受体（AHR），mRNA
雷帕霉素的机理靶标
尼莫（nemo）样激酶（NLK），mRNA
CD47分子（CD47），转录变异体1，
v-raf鼠肉瘤病毒癌基因同系物
核受体亚家族2，组F，
早发型乳腺癌2（BRCA2），mRNA
长寿因子1（SIRT1），转录变异体1，mRNA
REV1，聚合酶（DNA指导的）（REV1），
芳香烃受体核
隐花色素2（光裂合酶样）（CRY2），
神经纤维瘤蛋白2（梅林（merlin））（NF2），转录物
肉碱O-辛酰基转移酶（CROT），
天冬氨酸β-羟化酶（ASPH），转录物
早发型乳腺癌1（BRCA1），
维生素D（1,25-二羟维生素D3）
盘状结构域受体酪氨酸激酶1
磷酸盐调节内肽酶同系物，
GNAS反义RNA 1（GNAS-AS1），反义
白介素6（干扰素，β2）（IL6），mRNA
钙敏感受体（CASR），转录物
过氧化物酶体增殖物激活受体
Nanog同源盒（NANOG），mRNA
无翅型MMTV整合位点家族，
囊性纤维化跨膜传导
细胞色素P450，家族2，亚家族D，
音猬因子（SHH），mRNA
血红蛋白，β（HBB），mRNA
血管性血友病因子（VWF），mRNA
CD14分子（CD14），转录变异体1，
神经肽Y（NPY），mRNA
腺苷酸激酶2（AK2），核基因
骨形态发生蛋白6（BMP6），mRNA
降钙素相关多肽α（CALCA），
CD34分子（CD34），转录变异体2，
胶质纤维酸性蛋白（GFAP），转录物
MCM2癌基因，E3泛素蛋白连接酶
性别决定区Y（SRY），mRNA
间变性淋巴瘤受体酪氨酸激酶
表皮生长因子受体（EGFR），
金属硫蛋白3（MT3），mRNA
CD28分子（CD28），转录变异体1，
叉头框P3（FOXP3），转录变异体1，
端粒酶逆转录酶（TERT)），
前列腺癌抗原3（非蛋白编码）
WT1反义RNA（WT1-AS），非编码RNA

图64

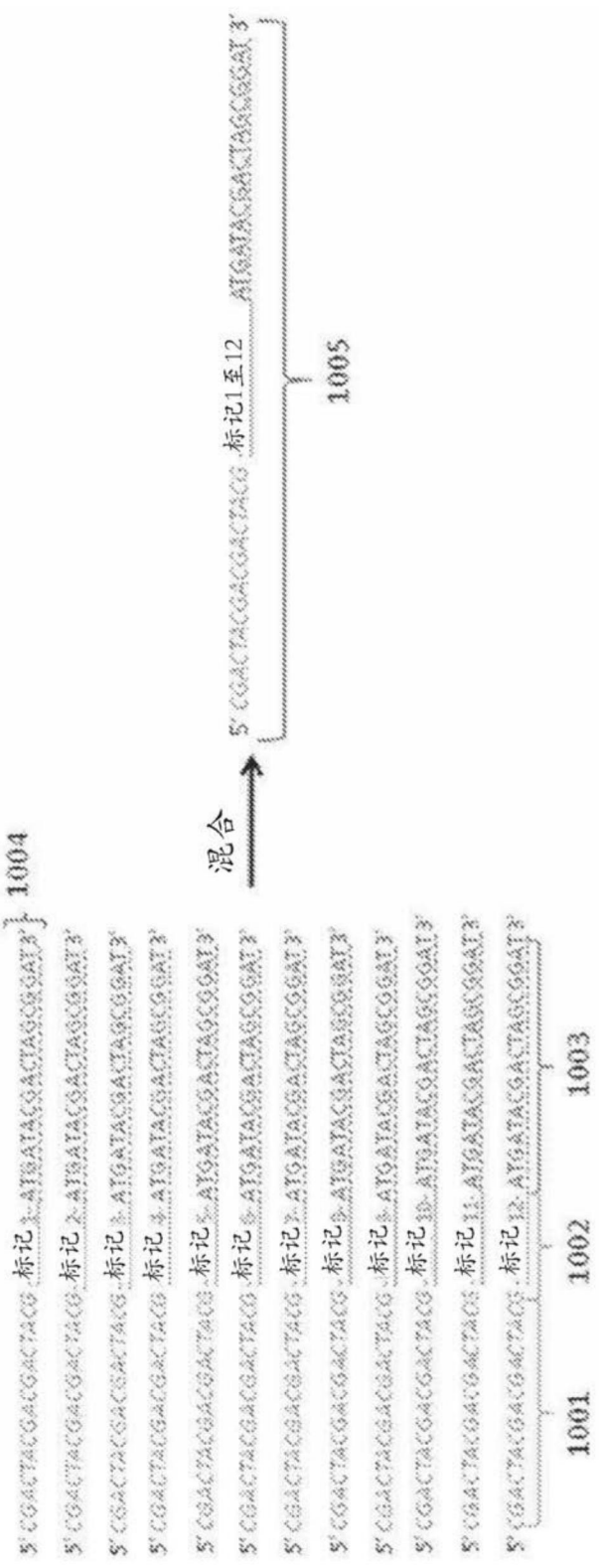


图65

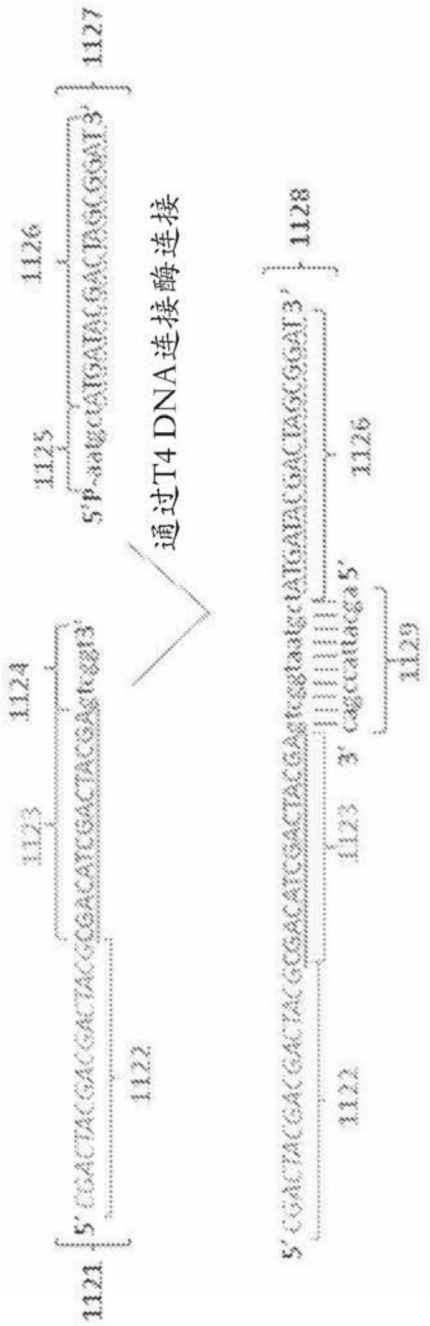


图66A

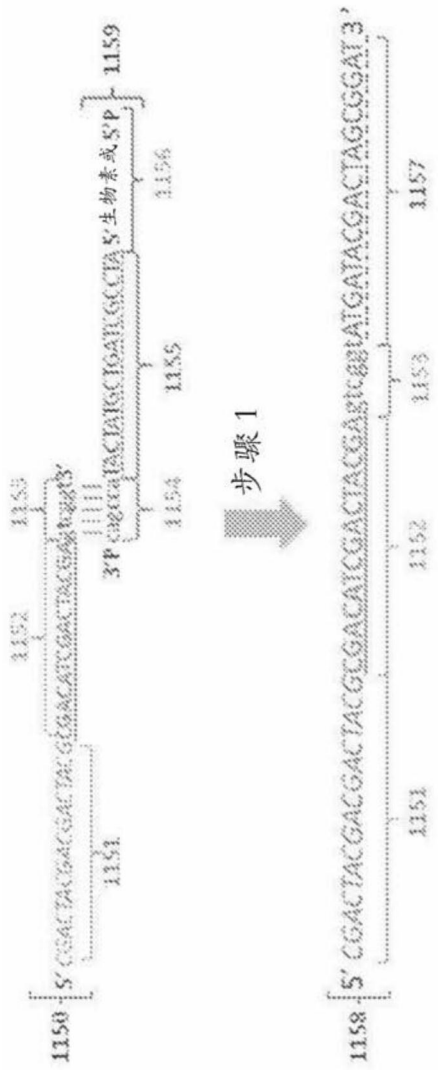

步骤 1
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
5' 生物素或 5'P

图66B

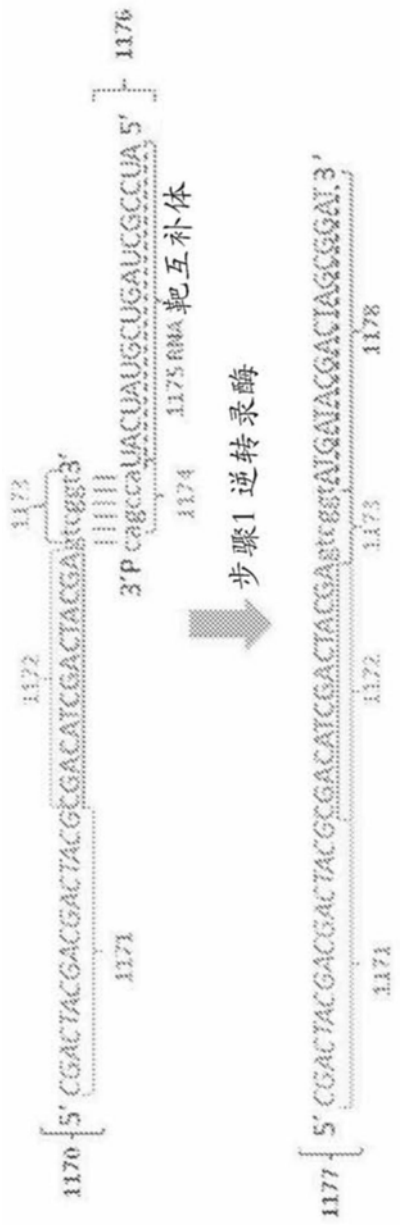


图66C

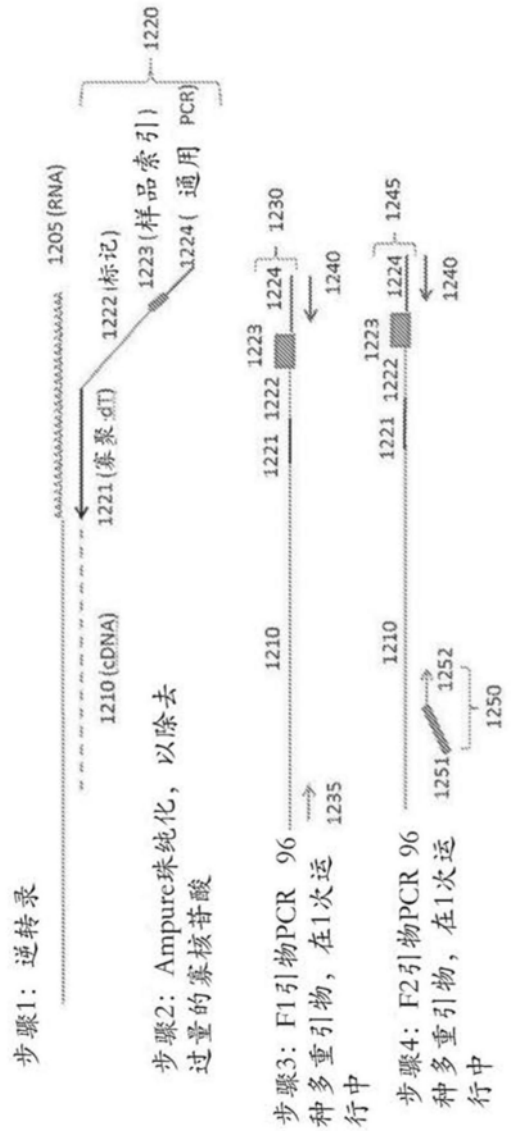

步骤1：逆转录
1205(RNA)
1210(cDNA)
1221(寡聚dT)
1222(标记)
1223(样品索引)
1224(通用 PCR)
1220
步骤2：Ampure珠纯化，以除去过量的寡核苷酸
步骤3：F1引物PCR 96种多重引物，在1次运行中
1210
1221
1222
1223
1224
1230
1240
1235
步骤4：F2引物PCR 96种多重引物，在1次运行中
1245
1251
1252
1250

图67

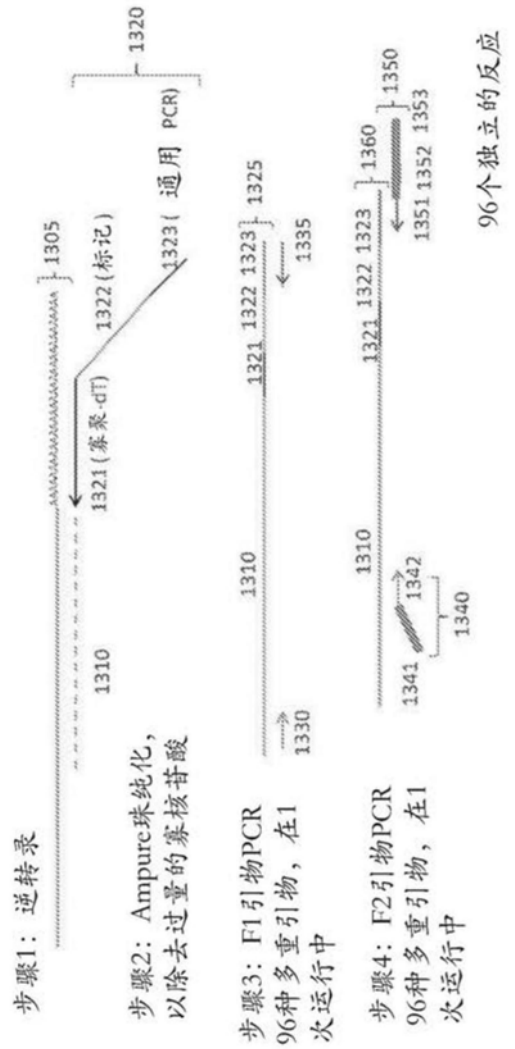

步骤1：逆转录
步骤2：Ampurc珠纯化，以除去过量的寡核苷酸
步骤3：F1引物PCR 96种多重引物，在1次运行中
步骤4：F2引物PCR 96种多重引物，在1次运行中
96个独立的反应
1305
1310
1320
1321（寡聚-dT）
1322（标记）
1323（通用 PCR）
1325
1330
1335
1340
1341
1342
1350
1351
1352
1353
1360

图68

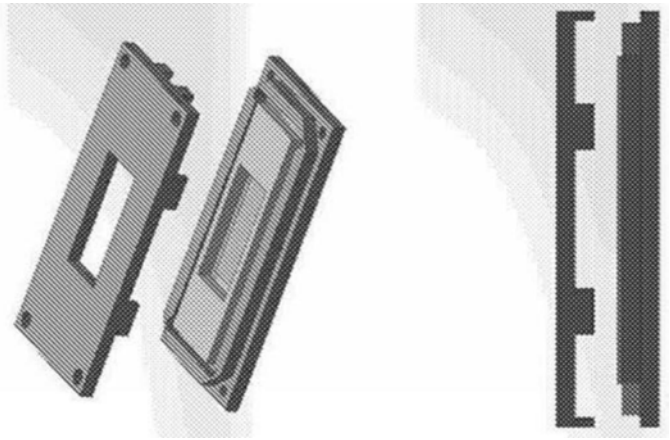


图69

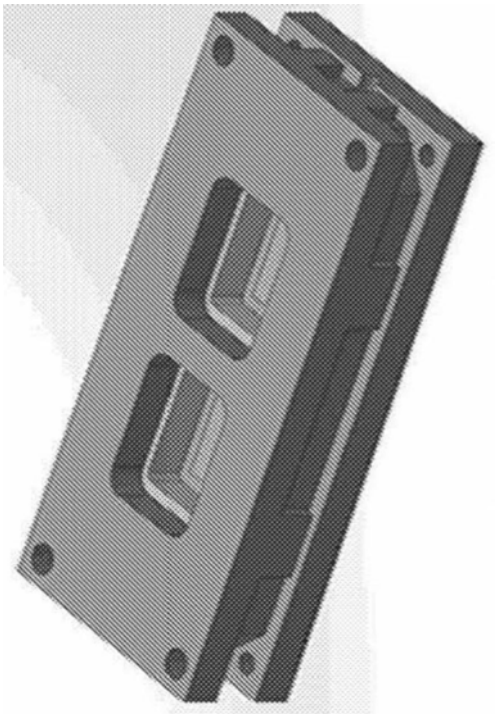


图70

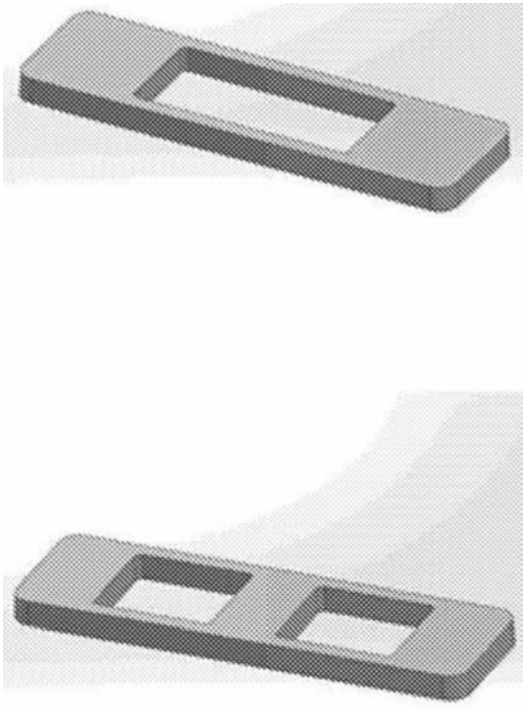


图71

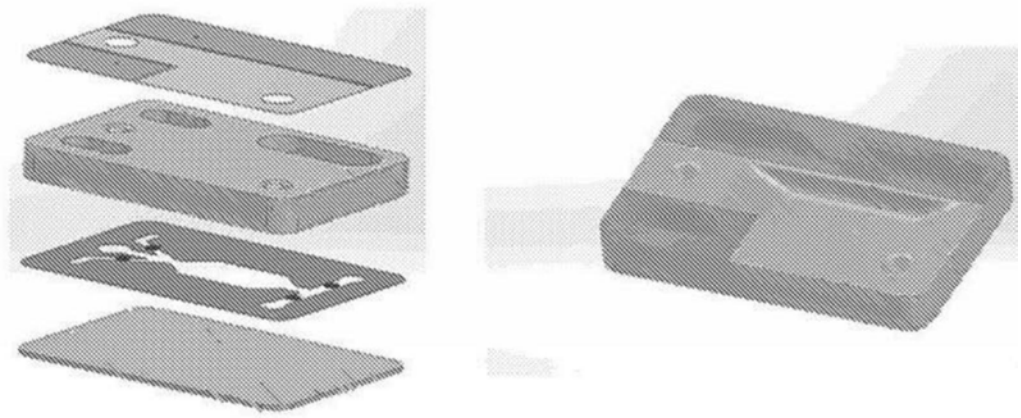


图72

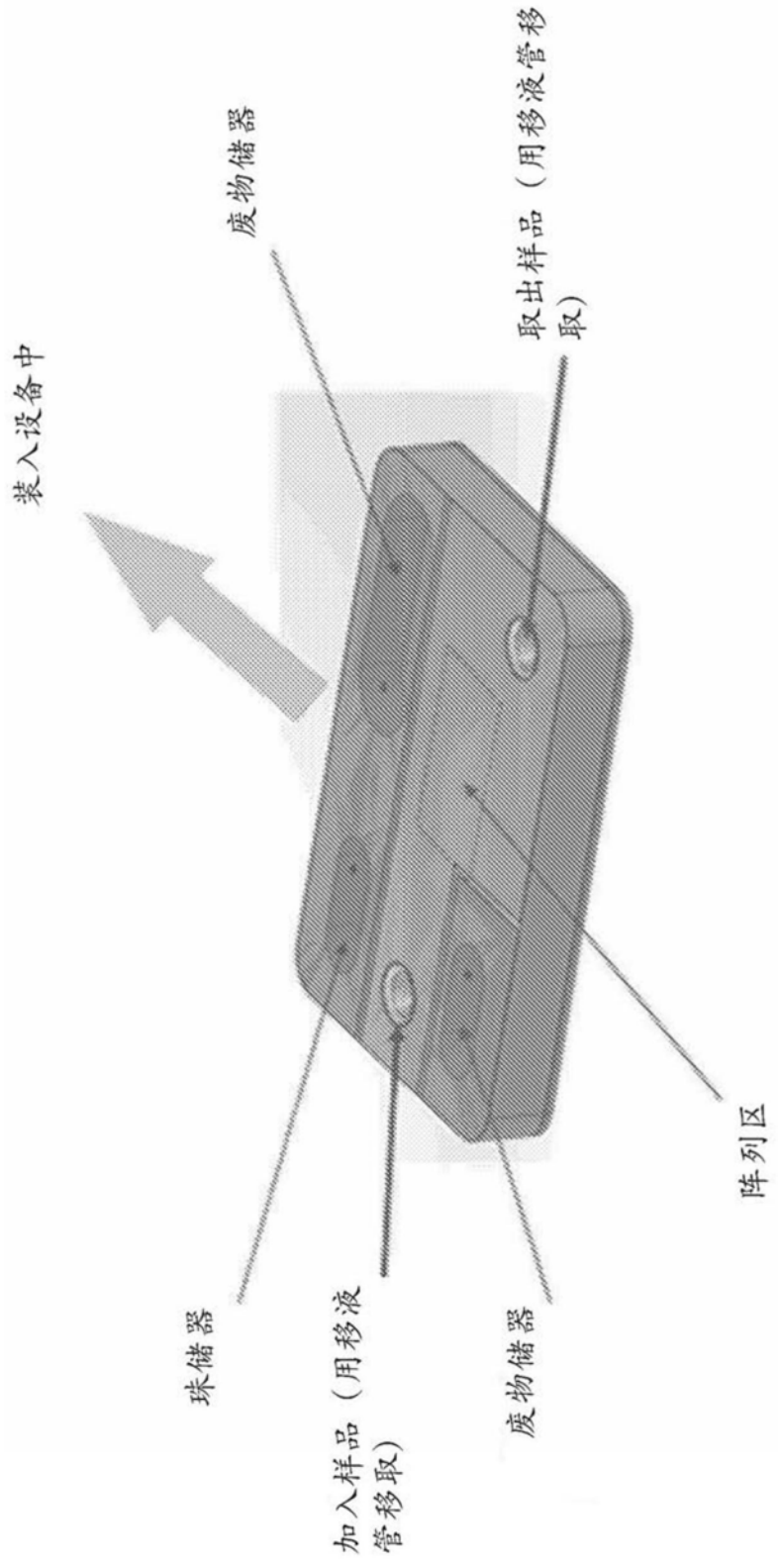


图73

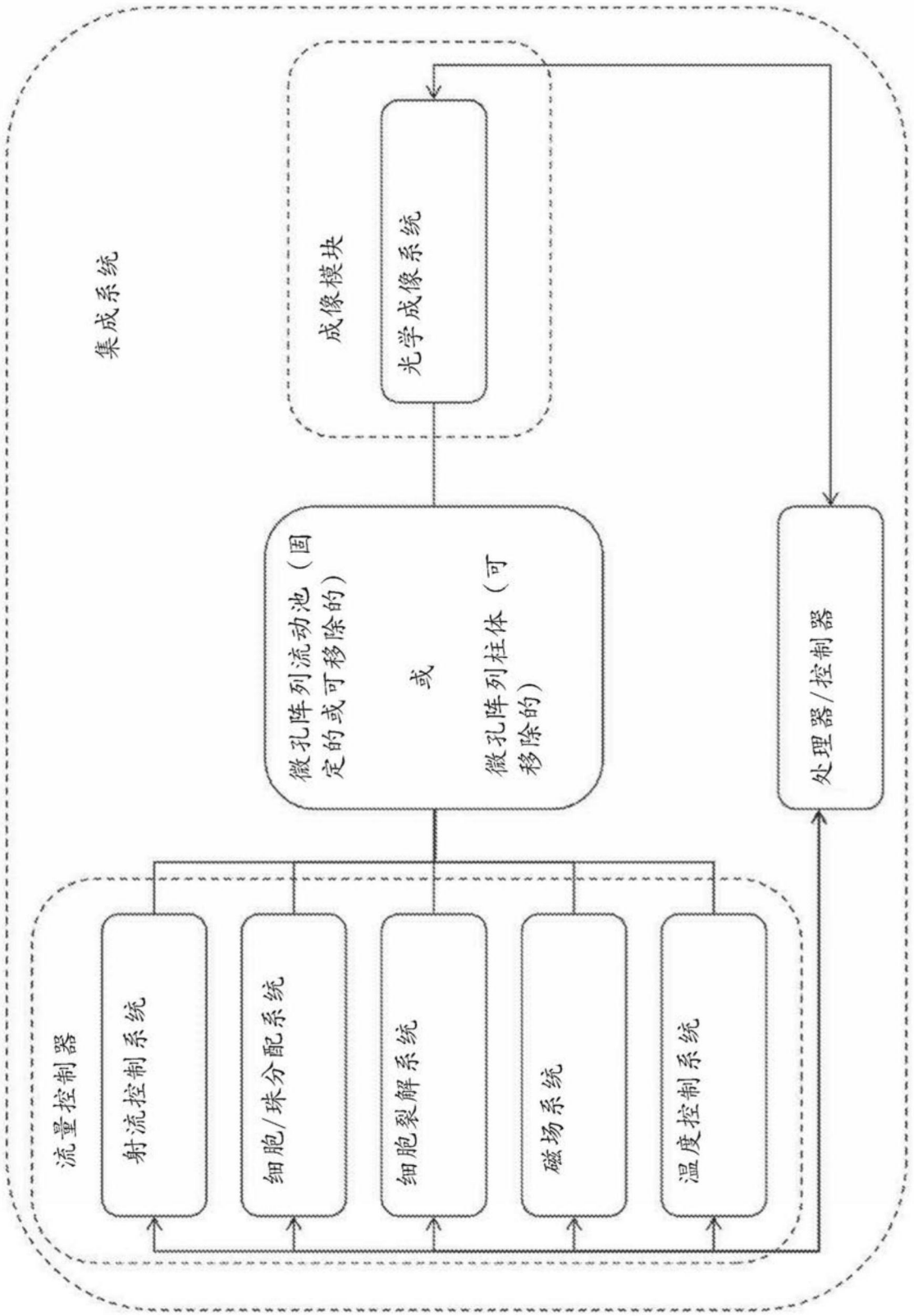

集成系统
成像模块
光学成像系统
微孔阵列流动池（固定的或可移除的）
或
微孔阵列柱体（可移除的）
处理器/控制器
流量控制器
射流控制系统
细胞/珠分配系统
细胞裂解系统
磁场系统
温度控制系统

图74

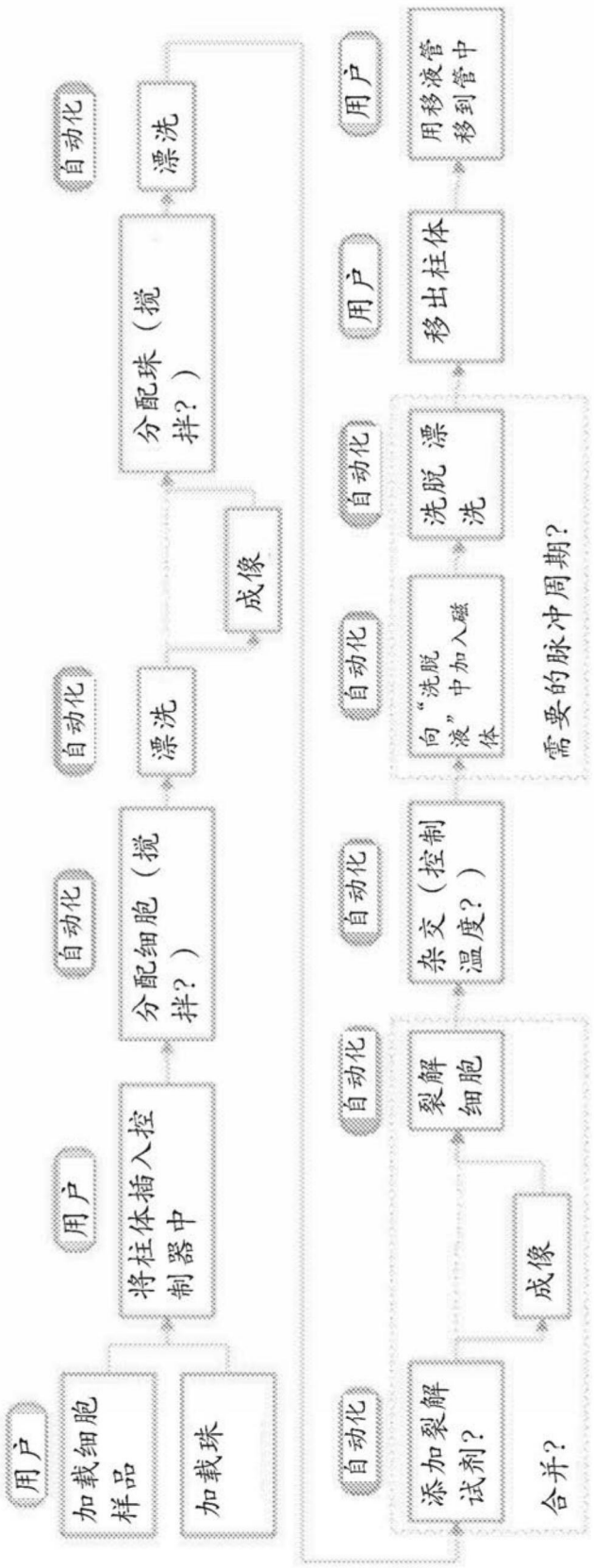

用户
加载细胞样品
加载珠
用户
将柱体插入控制器中
自动化
分配细胞（搅拌？）
自动化
漂洗
自动化
分配珠（搅拌？）
漂洗
成像
自动化
裂解细胞
自动化
杂交（控制温度？）
自动化
向“洗脱液”中加入磁体
自动化
洗脱 漂洗
用户
移出柱体
用户
用移液管移到管中
自动化
添加裂解试剂？
合并？
成像
需要的脉冲周期？

图75

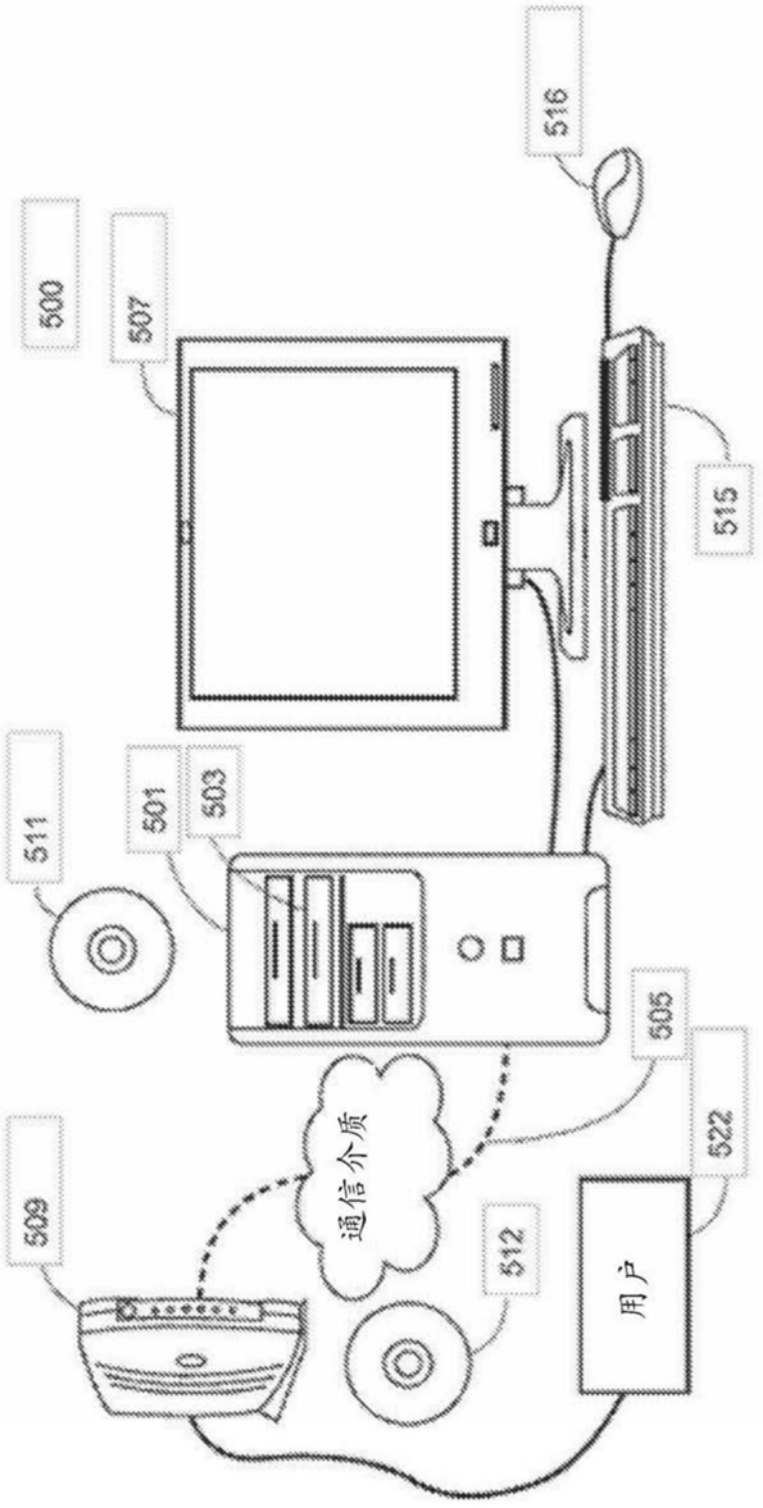


图76

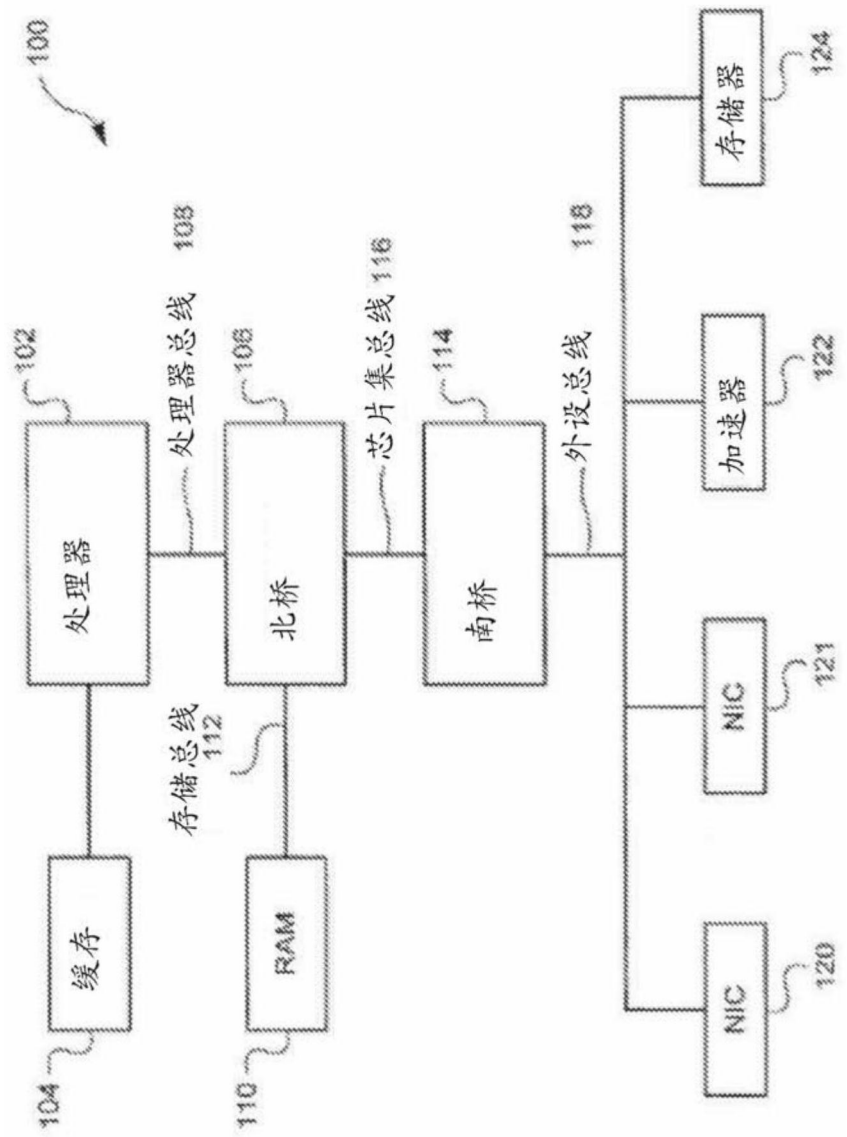

100
处理器
102
处理器总线
103
北桥
106
芯片集总线
116
南桥
114
外设总线
118
存储器
124
加速器
122
NIC
121
NIC
120
缓存
104
存储总线
112
RAM
110

图77

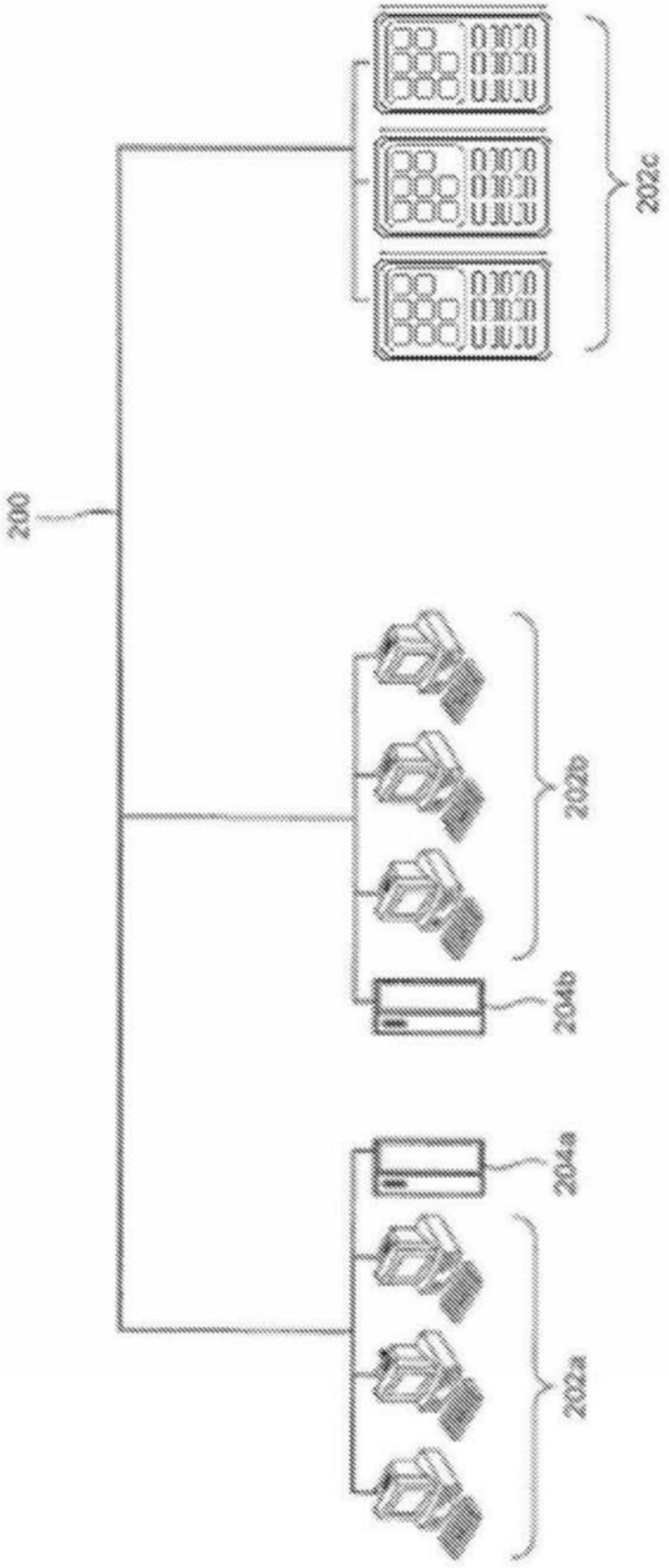


图78

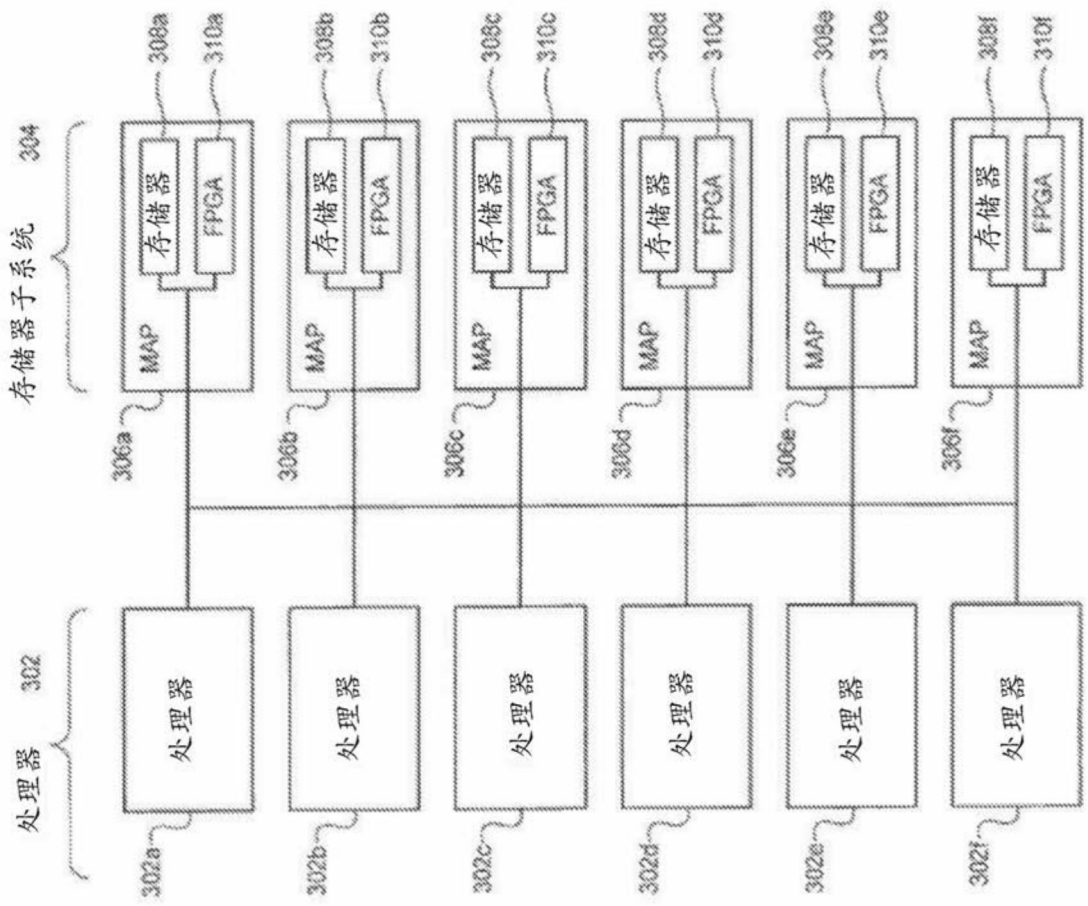

处理器 302
存储器子系统 304
处理器 302a
处理器 302b
处理器 302c
处理器 302d
处理器 302e
处理器 302f
MAP 306a
MAP 306b
MAP 306c
MAP 306d
MAP 306e
MAP 306f
存储器 308a
存储器 308b
存储器 308c
存储器 308d
存储器 308e
存储器 308f
FPGA 310a
FPGA 310b
FPGA 310c
FPGA 310d
FPGA 310e
FPGA 310f

图79

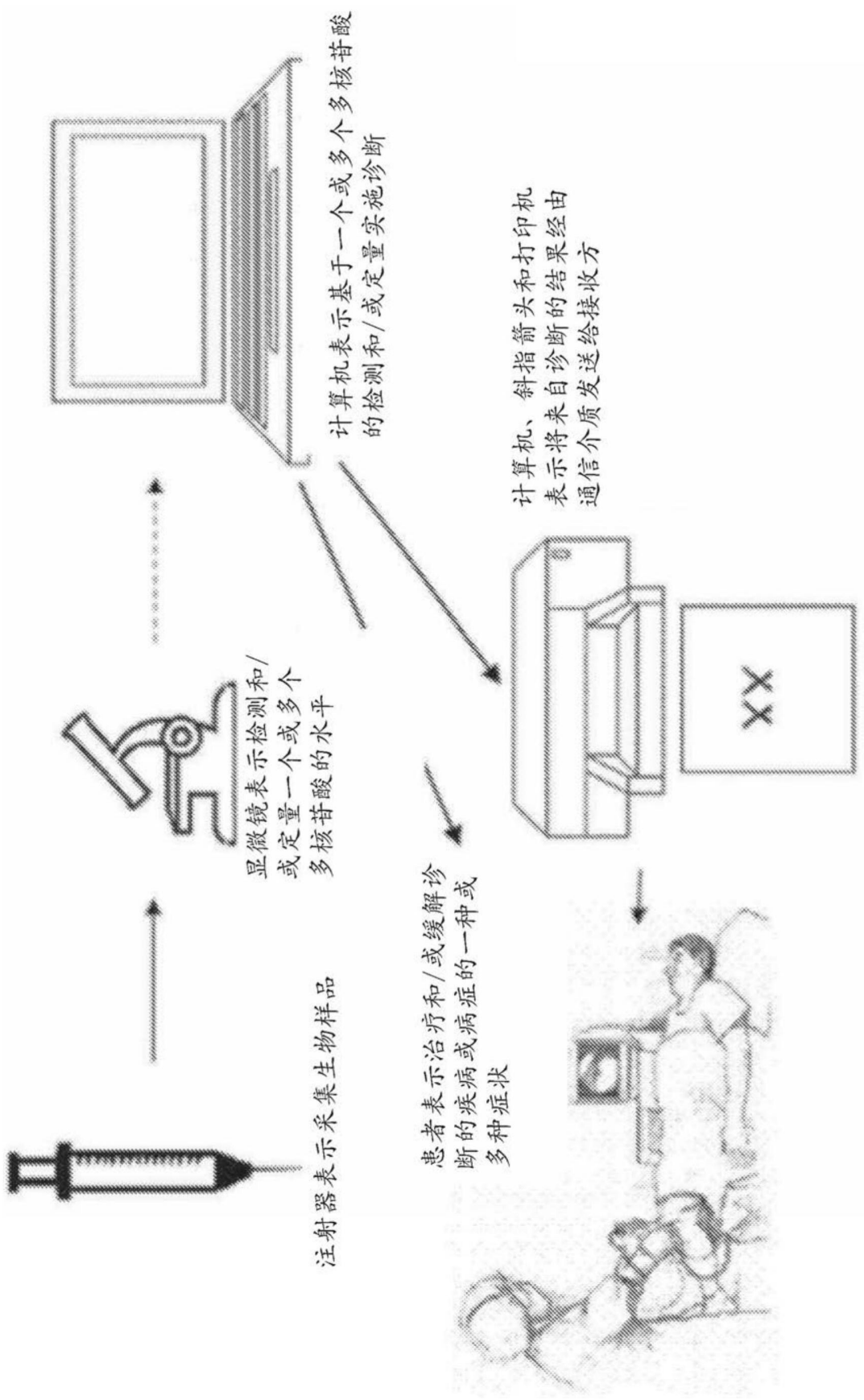

注射器表示采集生物样品
显微镜表示检测和/或定量一个或多个多核苷酸的水平
计算机表示基于一个或多个多核苷酸的检测和/或定量实施诊断
计算机、斜指箭头和打印机表示将来自诊断的结果经由通信介质发送给接收方
患者表示治疗和/或缓解诊断的疾病或病症的一种或多种症状
XX

图80