



(12)发明专利申请

(10)申请公布号 CN 108603233 A

(43)申请公布日 2018.09.28

(21)申请号 201680078293.2

(74)专利代理机构 北京安信方达知识产权代理有限公司 11262

(22)申请日 2016.11.01

代理人 刘晓杰

(30)优先权数据

62/250,422 2015.11.03 US

(51)Int.Cl.

G12Q 1/6886(2018.01)

(85)PCT国际申请进入国家阶段日

G01N 33/574(2006.01)

2018.07.02

G01N 33/53(2006.01)

(86)PCT国际申请的申请数据

PCT/US2016/059877 2016.11.01

(87)PCT国际申请的公布数据

W02017/079139 EN 2017.05.11

(71)申请人 艾匹克科学公司

地址 美国加利福尼亚州

(72)发明人 瑞安·迪塔莫尔 D·马里努茨

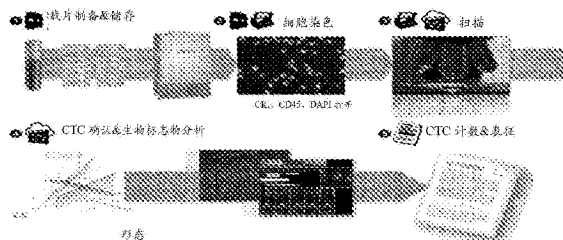
权利要求书1页 说明书15页 附图9页

(54)发明名称

转移性疾病中循环肿瘤细胞(CTC)的单细胞基因组图谱分析以表征疾病异质性

(57)摘要

本披露提供了一种检测癌症患者中疾病的异质性的方法,该方法包括(a)进行直接分析,其包括针对获自该患者的血液样本中的有核细胞的免疫荧光染色和形态表征以确认和计数循环肿瘤细胞(CTC);(b)从该样本中分离该CTC;(c)单独表征基因组参数以生成该CTC中每一种的基因组图谱,以及(c)基于该图谱确定该癌症患者中疾病的异质性。在一些实施方式中,该癌症是前列腺癌。在一些实施方式中,该前列腺癌是激素难治性的。



1. 一种检测癌症患者中疾病的异质性的方法,该方法包括:
  - (a) 进行直接分析,其包括针对获自该患者的血液样本中的有核细胞的免疫荧光染色和形态表征以确认和计数循环肿瘤细胞 (CTC);
  - (b) 从所述样本中分离该CTC;
  - (c) 单独表征基因组参数以生成该CTC中每一种的基因组图谱,以及 (c) 基于所述图谱确定该癌症患者中疾病的异质性。
2. 如权利要求1所述的方法,其中所述癌症是前列腺癌。
3. 如权利要求2所述的方法,其中所述前列腺癌是激素难治性的。
4. 如权利要求1-3中任一项所述的方法,其中有核细胞的该免疫荧光染色包括广谱细胞角蛋白、分化簇 (CD) 45和二脒基-2-苯基吲哚 (DAPI)。
5. 如权利要求1-4中任一项所述的方法,其中所述基因组参数包括拷贝数变异 (CNV) 标签。
6. 如权利要求5所述的方法,其中所述拷贝数变异 (CNV) 标签包括基因扩增或删除。
7. 如权利要求6所述的方法,其中所述CNV标签包括与雄激素非依赖性细胞生长相关的基因。
8. 如权利要求6所述的方法,其中所述删除包括磷酸酶和张力蛋白同源物基因 (PTEN) 的缺失。
9. 如权利要求6所述的方法,其中所述基因扩增包括AR基因的扩增。
10. 如权利要求1-4中任一项所述的方法,其中所述基因组参数包括基因组不稳定性。
11. 如权利要求10所述的方法,其中所述基因组不稳定性通过测量大规模转变 (LST) 来表征。
12. 如权利要求10所述的方法,其中所述基因组不稳定性通过测量基因组改变百分比 (PGA) 来表征。

## 转移性疾病中循环肿瘤细胞 (CTC) 的单细胞基因组图谱分析 以表征疾病异质性

[0001] 本申请要求2015年11月3日提交的美国临时申请号62/250,422 (其全部内容通过引用并入本文) 的权益。

### 技术领域

[0002] 本发明一般涉及癌症诊断领域, 并且更具体地涉及用于循环肿瘤细胞 (CTC) 的单细胞基因组图谱分析以表征疾病异质性的方法。

### 背景技术

[0003] 在连续的癌症疗法之后, 出现多个癌细胞亚群, 每个亚群具有可能赋予耐药性或易感性的不同遗传畸变。组织活检可能无法检测到这些亚群, 但血液的液体活检可以帮助识别这些重要的肿瘤细胞, 并表征患者的肿瘤如何随着时间的推移已演变。单细胞基因组图谱分析是一项用于研究癌症的进化和多样性以及了解罕见细胞在肿瘤进展中的作用的强大新工具。克隆多样性注定将在侵袭、转移和抵抗疗法的进化中发挥重要作用。

[0004] 在美国 (US), 前列腺癌是最常诊断的实体器官恶性肿瘤, 并且仍然是美国男性中癌症死亡的第二主因。仅在2014年, 预计的前列腺癌发病率为233,000例, 其中29,480位男性死亡, 这使得转移性前列腺癌疗法确实成为未满足的医疗需求。Siegel等人, 2014. CA Cancer J Clin. 2014;64 (1) :9-29。来自欧洲的流行病学研究显示可比数据, 其中2012年估计发病率为416 700例新病例, 占男性癌症诊断的22.8%。总共预计有92 200例特定的前列腺癌死亡, 使其成为男性最有可能致死的三种癌症之一, 其中死亡率为9.5%。

[0005] 尽管使用化学或手术去势对前列腺癌的荷尔蒙疗法已经证明是成功的, 但大多数患者最终还是会展至转移并显示出对进一步激素操纵的抵抗性的疾病阶段。这被称为转移性去势抵抗性前列腺癌 (mCRPC)。然而, 尽管如此命名, 还是有证据表明在mCRPC中可以持续存在雄激素受体 (AR) 介导的信号传导和基因表达, 即使在面对去势的雄激素水平时也是如此。这可能部分归因于与雄激素合成的酶的上调, AR的过度表达或伴随着各种甾体配体的混杂识别的突变AR的出现。在20-30%的mCRPC中发现的雄激素受体 (AR) -基因扩增的进展被认为是激素剥夺疗法的结果, 并且是治疗失败的主要原因。对mCRPC患者的治疗仍然是一个重大的临床挑战。研究进一步阐明了PI3K-AKT-mTOR和雄激素受体 (AR) 信号轴 (signaling axis) 之间的直接联系, 揭示了激素抵抗进展过程中这些途径之间的动态相互作用。PTEN是人前列腺癌中最常被删除/突变的肿瘤抑制基因之一。作为PI3K/AKT/mTOR途径的脂质磷酸酶和负调节因子, PTEN控制着许多细胞过程, 包括存活、生长、增殖、代谢、迁移和细胞结构。PTEN缺失可以用作前列腺癌的诊断和预后生物标志物, 以及预测患者对新兴的PI3K/AKT/mTOR抑制剂的响应。

[0006] 在2004年之前, 没有任何治疗证明可以提高mCRPC男性患者的存活率。用米托蒽醌, 用泼尼松或氢化可的松对患者的治疗仅仅是为了减轻疼痛并改善生活质量, 但在总生存期 (OS) 方面没有益处。2004年, 两项主要的3期临床试验TAX 327和SWOG (西南肿瘤组织)

9916的结果确立了泰索帝®(多西他赛)作为mCRPC患者的主要化学治疗选择。已经针对mCRPC研究了利用雄激素受体(AR)靶向疗法、化学疗法、联合疗法和免疫疗法的额外激素治疗,并且最近的结果为这个难以治疗的患者群提供了额外的选择。随着仅在过去5年中测试和批准用于治疗转移性去势抵抗性前列腺癌(mCRPC)患者的新型药剂的指数增长的出现,已经出现了关于这些药剂的最佳排序或组合的问题。在最佳排序方法方面帮助指导临床医生的指南有几条,而且大多数可就症状的存在或缺乏、表现状态以及疾病负担做出评估,以帮助确定这些药剂的最佳排序。Mohler等人,2014,J Natl Compr Canc Netw.2013;11(12):1471-1479;Cookson等人,2013,J Urol.2013;190(2):429-438。目前,批准的治疗方法包括紫杉烷类细胞毒性剂,如泰索帝®(多西他赛)和Jevtana®(卡巴他赛),以及抗雄激素激素治疗药物,如Zytiga®(arbiteron,阻止雄激素产生)或Xtandi®(恩杂鲁胺,一种雄激素受体(AR)抑制剂)。

[0007] 临床医生面临的挑战是决定施用这些疗法的最佳顺序以向患者提供最大的益处。然而,治疗失败仍然是基于患者对各种疗法的异质响应以及来自每种药剂的交叉抗性的重大挑战。Mezynski等人,Ann Oncol.2012;23(11):2943-2947;Noonan等人,Ann Oncol.2013;24(7):1802-1807;Pezaro等人,Eur Urol.2014,66(3):459-465。此外,患者可能会错过从被证明可提供总体生存获益的每种药物中获得实质性益处的治疗窗。因此,确认最有可能从目标疗法中受益的目标人群的更好方法仍然是一个重要目标。

[0008] 循环肿瘤细胞(CTC)代表癌症诊断中的显著进步,而其非侵入性测量甚至使其更具吸引力。Cristofanilli等人,N Engl J Med 2004,351:781-91。从原发性肿瘤或其转移部位释放的CTC具有关于该肿瘤的生物学的重要信息。历史上,血流中CTC的水平极低,再加上其表型未知,这严重阻碍了它们的检测并限制了它们的临床效用。为了利用它们的信息,最近已经出现了各种用于CTC的检测、分离和表征的技术。CTC有可能提供一种在治疗过程中实时评估进展性癌症的非侵入性手段,并进一步地,通过监测治疗响应中出现的表型生理和遗传变化来帮助指导治疗。在大多数晚期前列腺癌患者中,原发性肿瘤已被移除,并且预期CTC由转移酶脱落的细胞组成,从而提供“液体活检”。虽然CTC通常定义为EpCAM/细胞角蛋白阳性(CK+)细胞,CD45-,但在形态学上截然不同,最近的证据表明存在其他CTC候选群体,包含EpCAM/细胞角蛋白阴性(CK-)细胞或比传统CTC更小的细胞。这些关于CTC群体异质性的研究结果表明,无富集CTC平台有利于基于阳性选择技术,该技术根据大小、密度或EpCAM阳性分离CTC,容易忽视重要CTC亚群。

[0009] CRPC对患有这种晚期形式的前列腺癌的患者和管控这些患者的临床医生都提出了严重的挑战。临床医生通常面临的问题是提供对导致疾病恶化的机制的全面诊断和评估,以努力指导适当和个体化的治疗。通过确认适当的治疗和预后指标,靶向疗法的潜在临床益处得以增加,并且临床医生能够更好地控制CRPC,改善患者的生活质量,并增强临床效果。需要了解个别CTC中的亚克隆CNV驱动基因改变的频率和基因组不稳定性,结合细胞表型,以使得能够更准确地观察异源疾病,预测治疗响应并确认新型的抗性机制。本发明解决了这种需要并提供了相关的优点。

[0010] 发明概述

[0011] 本发明提供了一种检测癌症患者中疾病的异质性的方法,该方法包括(a)进行直接分析,其包括针对获自该患者的血液样本中的有核细胞的免疫荧光染色和形态表征以确

认和计数循环肿瘤细胞 (CTC) ; (b) 从该样本中分离该CTC; (c) 单独表征基因组参数以生成该CTC中每一种的基因组图谱, 以及 (c) 基于该图谱确定该癌症患者中疾病的异质性。在一些实施方式中, 该癌症是前列腺癌。在一些实施方式中, 该前列腺癌是激素难治性的。

[0012] 在一些实施方式中, 有核细胞的该免疫荧光染色包括广谱细胞角蛋白、分化簇 (CD) 45、二脒基-2-苯基吡啶以及二脒基-2-苯基吡啶 (DAPI) 。

[0013] 在一些实施方式中, 该基因组参数包括拷贝数变异 (CNV) 标签。在一些实施方式中, 该CNV标签包括基因扩增或删除。在一些实施方式中, 该基因扩增包括AR基因的扩增。在一些实施方式中, 该删除包括磷酸酶和张力蛋白同源物基因 (PTEN) 的缺失。在一些实施方式中, 该CNV标签包括与雄激素非依赖性细胞生长相关的基因。

[0014] 在一些实施方式中, 该基因组参数包括基因组不稳定性。在一些实施方式中, 该基因组不稳定性通过测量大规模转变 (LST) 来表征。在一些实施方式中, 该基因组不稳定性通过测量基因组改变百分比 (PGA) 来表征。

[0015] 根据发明详述, 以及根据权利要求书, 本发明的其他特征和优点将是显而易见的。

## 附图说明

[0016] 图1A示出了对标准Epic CTC分析过程的描述。使用多参数数字病理学算法分析图像以检测CTC候选者并定量蛋白质生物标志物表达水平。CTC分类显示在基于网络的报告中, 并由经过培训的技术人员确认。图1B显示了对CTC恢复和基因组图谱分析工作流程的描述。单个细胞分离、进行全基因组扩增, 以及NGS文库制备。在Illumina NextSeq 500上执行测序。

[0017] 图2提供了执行的生物信息学分析的图。对原始FASTQ文件进行评估并进行质量过滤。读数与hg 38参考基因组 (UCSC) 对齐, 去除PCR重复, 并通过MAPQ得分30过滤。对过滤后读数>250K的样本进行拷贝数改变的分析。过滤后的对齐文件用Epic的拷贝数管道 (Epic's Copy Number Pipelines) 进一步分析。一条管道用于使用1M bp窗口估算基因组不稳定性, 而另一条用于基因特异性拷贝数测量。<sup>1</sup>LST: 至少10Mb的邻近区域之间的染色体断裂的数目。<sup>2</sup>PGA: 携带拷贝数改变 (扩增或删除) 的患者的基因组百分比。

[0018] 图3A-3C示出了单个细胞中的拷贝数变异 (CNV) 。分别从LNCaP、PC3和VcaP中分离单个细胞并通过全基因组测序对其进行拷贝数变异的分析。扩增和删除可以在重复中可再现地观察。还示出了每种细胞系的代表性图像。细胞用CK混合物、AR、CD45和DAPI染色。此处示出了每种细胞系的5个重复, 以证明可再现性。图3D中描述了来自每种细胞系的已知基因组改变。用Circos生成绘图: Krzywinski, M. 等人Circos: an Information Aesthetic for Comparative Genomics. *Genome Res* (2009) 19:1639-1645。

[0019] 图4A-4D示出了CNV (图4A和4B) 和基因组不稳定性测量 (图4C和4D) 。图4A示出了3种代表性细胞系和健康供体白细胞 (WBC) 对照中AR的log2基因组拷贝数的比较。VcaP含有AR的扩增, 而LNCaP和PC3保持有AR的2个拷贝。图4B示出了3种代表性细胞系和健康供体WBC对照中PTEN的log2基因组拷贝数的比较。证实了PC3纯合PTEN的缺失, 在具有显著z-分数的许多细胞中观察到LNCaP杂合PTEN的缺失。图4C示出了3种代表性细胞系和健康供体WBC对照中的断点# (LST) 的比较。与LNCaP (wt p53和杂合PTEN缺失) 和WBC对照相比, 在PC3 (PTEN无效, p53突变体) 和VcaP (p53突变体) 中检测到的断点数量更多。图4D示出了在3种代表性

细胞系和健康供体WBC对照中改变的基因组的%的比较。PC3示出了最高的改变百分比,揭示了可能由PTEN和p53两者的缺失导致的遗传不稳定性和多倍性。

[0020] 发明详述

[0021] 本披露部分地基于以下发现:整合的单细胞全基因组CNV分析提供

[0022] 跨越多个重复的可再现的拷贝数图谱,并确认了已知的局灶性CNV事件(包含AR扩增和PTEN缺失)的存在。本披露进一步部分地基于以下发现:

[0023] 全基因组拷贝数分析可以用于通过测量LST和PGA来可再现地表征基因组不稳定性。如本文所披露的,与野生型(LNCaP)相比,在p53突变细胞系(PC3&VCaP)中检测到基因组不稳定性最高。了解个别CTC中的亚克隆CNV驱动基因改变的频率和基因组不稳定性,结合细胞表型,可以使得能够更准确地观察异源疾病,潜在的治疗响应并确认新型的抗性机制。

[0024] 肿瘤内异质性的增加与对疗法的内在抗性和不良后果相关。已示出了CTC反映转移性患者中的异质性疾病和活性转移性肿瘤群体。本文所述的非富集CTC分析平台通过允许异质CTC群体的单细胞分辨率和准确的基因组图谱分析来实现本发明的方法。为了表征肿瘤内异质性,使用非富集CTC分析平台执行循环肿瘤细胞(CTC)的单细胞全基因组拷贝数分析。在掺入血液中以模拟患者样本的个体前列腺癌细胞中检测到治疗敏感性(例如分别针对PI3K抑制剂或AR-靶向疗法的PTEN删除或雄激素受体(AR)扩增)的标记。除了检测局灶性可操作改变之外,还通过测量大规模转变(LST)和基因组改变%(PGA)来表征基因组不稳定性。

[0025] 本发明提供了一种检测癌症患者中疾病异质性的方法,该方法包括(a)进行直接分析,其包括对获自该患者的血液样本中的有核细胞的免疫荧光染色和形态表征以确认和计数循环肿瘤细胞(CTC);(b)从该样本中分离该CTC;(c)单独地表征基因组参数以生成该CTC中每种的基因组图谱,以及(c)基于该图谱确定该癌症患者中疾病的异质性。在一些实施方式中,该癌症是前列腺癌。在一些实施方式中,该前列腺癌是激素难治性的。

[0026] 在一些实施方式中,有核细胞的该免疫荧光染色包括广谱细胞角蛋白、分化簇(CD)45、二脒基-2-苯基吲哚(DAPI)和雄激素受体(AR)。

[0027] 在一些实施方式中,该基因组参数包括拷贝数变异(CNV)标签。在一些实施方式中,该CNV标签包括基因扩增或删除。在一些实施方式中,该基因扩增包括AR基因的扩增。在一些实施方式中,该删除包括磷酸酶和张力蛋白同源物基因(PTEN)的缺失。在一些实施方式中,该CNV标签包括与雄激素非依赖性细胞生长相关的基因。

[0028] 在一些实施方式中,该基因组参数包括基因组不稳定性。在一些实施方式中,该基因组不稳定性通过测量大规模转变(LST)来表征。在一些实施方式中,该基因组不稳定性通过测量基因组改变百分比(PGA)来表征。

[0029] 在一些实施方式中,基于该图谱确定该癌症患者中疾病的异质性确认疾病的新机制。

[0030] 在一些实施方式中,基于该图谱确定该癌症患者中疾病的异质性预测对治疗的积极响应。

[0031] 在一些实施方式中,基于该图谱确定该癌症患者中疾病的异质性预测对治疗的抗性。

[0032] 必须注意,如在本说明书和所附权利要求中所使用的,除非内容另外明确指明,否

则单数形式“一种”、“一个”和“该”包含复数指代。因此,例如,对“生物标志物”的提及包含两种或更多种生物标志物等等的混合物。

[0033] 术语“约”,特别是关于给定的量时,意在涵盖正或负百分之五的偏差。

[0034] 如在包含所附权利要求书的本申请中所使用的,除非内容另外明确指明,否则单数形式“一种”、“一个”和“该”包含复数指代,并且可与“至少一个”和“一个或多个”可互换使用。

[0035] 如本文中所用,术语“包括”、“包含”、“含有”及其任何变型旨在涵盖非排他性囊括,以便包括、包含或含有元素或元素列表的过程、方法、方法限定的产品或物质的组合物不仅包含那些元素,而且可以包含这样的过程、方法、方法限定的产品或物质的组合物中没有明确列出的或者并非其固有的其他元素。

[0036] 如本文中所用,在液体活检样本的上下文中使用的术语“提供”意在涵盖获得该样本的任何和全部手段。该术语在实践所要求保护的方法的情况下涵盖导致该样本的存在的所有直接和间接手段。

[0037] 如本文中所用,术语“患者”优选指人类,但也涵盖其他哺乳动物。注意,如本文中所用,术语“生物体”、“个体”、“对象”或“患者”被用作同义词并且可互换使用。

[0038] 如本文所述的组合物和方法中所使用的,术语“癌症”是指或描述了哺乳动物的生理状况,其典型特征在于细胞生长不受调节。在一个实施方式中,该癌症是上皮癌。在一个实施方式中,该癌症是前列腺癌。在本文所述的方法和组合物的各种实施方式中,该癌症可以包含但不限于乳腺癌、肺癌、前列腺癌、结肠直肠癌、脑癌、食道癌、胃癌、膀胱癌、胰腺癌、宫颈癌、头颈癌、卵巢癌、黑素瘤和多药耐药性癌症;或其亚型和分期(phase)。在另一个替代实施方式中,该癌症是“早期”癌症。在又一个实施方式中,该癌症是“晚期”癌症。如本文中所用,术语“肿瘤”是指恶性或者良性的所有肿瘤细胞生长和增殖,以及所有的癌前细胞和组织癌细胞和组织。该癌症可以是淋巴增殖性癌症,例如前体B淋巴母细胞性白血病/淋巴母细胞性淋巴瘤、滤泡性B细胞非霍奇金淋巴瘤、霍奇金淋巴瘤前体T细胞淋巴母细胞性白血病/淋巴母细胞性淋巴瘤、未成熟T细胞的赘生物、外周胸腺后T细胞的赘生物、T细胞幼淋巴细胞白血病、外周T细胞淋巴瘤、未明确的间变性大细胞淋巴瘤、成人T细胞白血病/淋巴瘤、慢性淋巴细胞白血病、套细胞淋巴瘤、滤泡性淋巴瘤、边缘区淋巴瘤、多毛细胞白血病、弥漫性大B细胞淋巴瘤、伯基特淋巴瘤、淋巴浆细胞性淋巴瘤、前体T淋巴母细胞性白血病/淋巴母细胞性淋巴瘤、T细胞幼淋巴细胞性白血病、血管免疫母细胞性淋巴瘤或结节性淋巴细胞为主的霍奇金淋巴瘤。

[0039] 如本文中所用,术语“循环肿瘤细胞”或“CTC”意在涵盖存在于生物样本中且与癌症有关的任何罕见细胞。可以以单个细胞或以CTC簇的形式存在的CTC通常是从患者的循环中发现的浓度非常低的实体瘤中脱落的上皮细胞。

[0040] 如本文中所用,“传统的CTC”是指细胞角蛋白阳性、CD45阴性的,包含DAPI核并且在形态上不同于周围白血细胞的单个CTC。

[0041] 如本文中所用,“非传统的CTC”是指至少在一个特征上不同于传统CTC的CTC。

[0042] 从广义上讲,生物样本可以是含有CTC的任何样本。样本可以包括体液,如血液;细胞制剂的可溶部分或培养有细胞的培养基的等分试样;从细胞中分离或提取的染色体、细胞器或膜;在溶液中或与底物结合的基因组DNA、RNA或cDNA;一种细胞;组织;组织印迹;指

纹;多种细胞;皮肤等。获自对象的生物样本可以是含有细胞的并且包含任何其中可以检测到CTC的材料任何样本。样本可以是,例如全血、血浆、唾液或其他含有细胞的体液或组织。

[0043] 在具体的实施方式中,该生物样本是血液样本。如本文所述,样本可以是全血,更优选是外周血或外周血细胞部分。如本领域技术人员将理解的,血液样本可以包含但不限于T细胞、单核细胞、嗜中性粒细胞、红细胞、血小板和微泡(例如外来体和外来体样囊泡)的血液的任何部分或组分。在本披露的上下文中,包含在血液样本中的血细胞涵盖任何有核细胞并且不限于全血的组分。因此,血细胞包含,例如白细胞(WBC)以及包含CTC的罕见细胞。

[0044] 本披露的样本可以各自含有通过本领域熟知的方法(如FACS、免疫组织化学)可区分的多个细胞群和细胞亚群。例如,血液样本可以含有非有核细胞群,例如红细胞群(如4-5百万/ $\mu\text{l}$ )或血小板群(150,000-400,000个细胞/ $\mu\text{l}$ ),以及有核细胞群,例如WBC群(如4,500-10,000个细胞/ $\mu\text{l}$ )、CEC群或CTC群(循环肿瘤细胞;如2-800个细胞/ $\mu\text{l}$ )。WBC可以含有例如嗜中性粒细胞(2,500-8,000个细胞/ $\mu\text{l}$ )、淋巴细胞(1,000-4,000个细胞/ $\mu\text{l}$ )、单核细胞(100-700个细胞/ $\mu\text{l}$ )、嗜酸性粒细胞(50-500个细胞/ $\mu\text{l}$ )、嗜碱性粒细胞(25-100个细胞/ $\mu\text{l}$ )等的细胞亚群。本披露的样本是非富集样本,即它们不富集有任何特定的有核细胞群或亚群。例如,非富集血液样本不富集有CTC、WBC、B细胞、T细胞、NK细胞、单核细胞等。

[0045] 在一些实施方式中,该样本是获自健康对象或基于已知的临床既定标准(包含,例如年龄、种族、家族和病史)被认为处于癌症或现有癌症转移高危的对象的血液样本。在一些实施方式中,该血液样本来自基于组织或液体活检和/或手术或临床依据被诊断患有癌症的对象。在一些实施方式中,该血液样本获自显示出癌症的临床表现和/或本领域中熟知的或者呈现特定癌症的任何已知风险因素的患者。在一些实施方式中,该癌症是膀胱癌,例如膀胱尿路上皮癌。

[0046] 如本文在生成CTC数据的情况中所使用的,术语直接分析意指与在检测之前使该样本富集CTC之后相反,在该样本中存在所有周围有核细胞的背景下检测CTC。在一些实施方式中,该方法包括提供包含CTC和至少200个周围白血细胞(WBC)的视野的显微术。

[0047] 本披露的基本方面是所披露的方法关于CTC检测的无与伦比的鲁棒性(robustness)。本文披露的关于CTC的罕见事件检测基于直接分析,即非富集有群体,其涵盖在周围的非罕见事件的背景下确认罕见事件。根据所披露的方法对该罕见事件的确认固有的将周围事件识别为非罕见事件。考虑到周围非罕见事件并确定非罕见事件的平均值(例如非罕见事件的平均细胞尺寸),这通过消除噪声允许对检测方法的校准。结果是所披露的方法的鲁棒性不能用不基于直接分析的方法实现,而是将富集的群体与罕见事件的固有扭曲的背景相比较。本文披露的直接分析方法的鲁棒性使得实现了CTC(包含本文所述的CTC亚型)的表征,其允许确认不能用其他CTC检测方法实现的表型和异质性,并且在请求保护的方法的背景下实现生物标志物的分析。

[0048] 在一些实施方式中,本文披露的方法可以进一步涵盖个体患者风险因素和成像数据,其包含本领域中已知和使用的任何形式的成像模态,例如但不限于,通过X射线计算机断层摄影术(CT)、超声波、正电子发射断层扫描(PET)、电阻抗断层扫描和磁共振(MRI)。应该理解,本领域技术人员可以基于各种已知标准选择成像模态。如本文所述,本发明的方法可



以涵盖一个或多个成像数据。在本文披露的方法中,一种或多种个体风险因素可以选自年龄、种族、家族史。应该理解,本领域技术人员可以基于各种已知的标准选择额外的个体风险因素。如本文所述,本发明的方法可以涵盖一种或多种个体风险因素。相应地,生物标记可以包含成像数据、个体风险因子和CTC数据。如本文所述,生物标记还可以包含但不限于包括如下的生物分子:核苷酸、核酸、核苷、氨基酸、糖、脂肪酸、类固醇、代谢物、肽、多肽、蛋白质、碳水化合物、脂质、激素、抗体、用作生物大分子的替代物的相关区域及其组合(如糖蛋白、核糖核蛋白、脂蛋白);以及生物分子的部分或片段。

[0049] CTC数据可以包含形态、遗传、表观遗传特征和免疫荧光特征。如本领域技术人员将理解的,生物标记可以包含生物分子或生物分子的片段,其变化和/或检测可以与癌症,与其他可测量特征以单独或相组合地形式相关联。可以呈现为单个细胞或CTC簇的CTC通常是从实体瘤中脱落的上皮细胞,并且在对象的循环中以非常低的浓度存在。相应地,血样中CTC的检测可以被称为罕见事件检测。血细胞群中的CTC具有小于1:1,000的丰度,例如小于1:5,000、1:10,000、1:30,000、1:50,000、1:100,000、1:300,000、1:500,000或1:1,000,000的丰度。在一些实施方式中,该细胞群中的CTC具有1:50,000至1:100,000的丰度。

[0050] 本披露的样本可以通过任何手段(包含,例如通过实体组织活检或流体活检)获得(参见,例如Marrinucci D.等人,2012,Phys.Biol.9 016003)。简而言之,在具体的实施方式中,过程可以涵盖裂解并去除7.5mL血液样本中的红细胞,将剩余的有核细胞沉积在专用显微镜载片上,其中每片容纳相当于大致0.5mL的全血。血液样本可以提取自任何已知的包含血细胞或其组分的来源,例如静脉、动脉、外周、组织、髓,以及诸如此类。该样本可以使用熟知的常规临床方法(例如,用于绘制和处理全血的程序)进行处理。在一些实施方式中,将血液样本抽吸入可以含有EDTA或Streck无细胞DNA™的抗凝血性采血管(BCT)。在其他实施方式中,将血液样本抽吸入CellSave®管(Veridex)。在进一步处理之前,血液样本可以进一步储存长达12小时、24小时、36小时、48小时或60小时。

[0051] 在一些实施方式中,本披露的方法包括获得该血液样本的白细胞(WBC)计数的初始步骤。在某些实施方式中,该WBC计数可以通过使用HemoCue®WBC装置(瑞典恩厄尔霍尔姆的Hemocue)获得。在一些实施方式中,该WBC计数用来确定每个载片上有核细胞的一致加载体积的涂布所需的血液量并反算每血液体积的CTC的当量。

[0052] 在一些实施方式中,本披露的方法包括裂解该血液样本中红细胞的初始步骤。在一些实施方式中,该红细胞,例如通过向该血液样本中加入氯化铵溶液来裂解。在某些实施方式中,在红细胞裂解之后使血液样本离心,并将有核细胞重悬,例如在PBS溶液中。

[0053] 在一些实施方式中,来自样本(例如血液样本)的有核细胞作为单层沉积在平面载体上。该平面载体可以是任何材料,例如任何荧光透明的材料,任何有助于细胞附着材料,任何有助于容易去除细胞碎片的材料,任何厚度<100μm的材料。在一些实施方式中,该材料是薄膜。在一些实施方式中,该材料是载玻片。在某些实施方式中,该方法涵盖将来自该血液样本的有核细胞作为单层沉积在载玻片上的初始步骤。该载玻片可以经涂覆以允许最大程度地保留活细胞(参见,例如Marrinucci D.等人,2012,Phys.Biol.9 016003)。在一些实施方式中,将约50万、100万、150万、200万、250万、300万、350万、400万、450万或500万个有核细胞沉积在该载玻片上。在一些实施方式中,本披露的方法包括将约3百万个细胞沉积到该载玻片上。在另外的实施方式中,本披露的方法包括将约2百万至约3百万个细胞沉

积在该载玻片上。在一些实施方式中,在本披露的方法完成之后,该载玻片和固定的细胞样本可用于进一步处理或实验。

[0054] 在一些实施方式中,本披露的方法包括确认该非富集血液样本中的有核细胞的初始步骤。在一些实施方式中,该有核细胞用荧光染色剂确认。在某些实施方式中,该荧光染色剂包含核酸特异性染色剂。在某些实施方式中,该荧光染色剂是二脒基-2-苯基吲哚(DAPI)。在一些实施方式中,有核细胞的免疫荧光染色包括广谱细胞角蛋白(CK)、分化簇(CD) 45和DAPI。在本文进一步描述的一些实施方式中,CTC包括相比周围有核细胞的区别性免疫荧光染色。在一些实施方式中,CTC的该区别性不同免疫荧光染色包括DAPI (+)、CK (+)和CD45 (-)。在一些实施方式中,CTC的确认进一步包括比较广谱细胞角蛋白荧光染色与周围有核细胞的强度。在一些实施方式中,该CTC数据通过荧光扫描显微术生成以检测血液样本中有核细胞的免疫荧光染色。Marrinucci D.等人,2012,Phys.Biol.9016003)。

[0055] 在具体的实施方式中,所有有核细胞被保留,并用靶向细胞角蛋白(CK)的单克隆抗体、专门存在于上皮细胞中的中间丝、靶向常见白细胞抗原CD45的泛白细胞特异性抗体和核染色剂DAPI进行免疫荧光染色。有核血细胞可以在多个荧光通道中成像,以产生保留了细胞核轮廓和细胞质分布的细胞学细节的高质量且高分辨率的数字图像。虽然周围WBC可以用该靶向CD45的泛白细胞特异性抗体确认,但CTC可以被确认为DAPI (+)、CK (+)和CD45 (-)。在本文所述的方法中,该CTC包括相比周围有核细胞的区别性免疫荧光染色。

[0056] 在进一步的实施方式中,该CTC数据包含也被称为高清CTC(HD-CTC)的传统CTC。传统的CTC是CK阳性的、CD45阴性的,含有完整的DAPI阳性细胞核,没有可识别的凋亡改变或受破坏的外观,并且在形态上不同于周围白细胞(WBC)。DAPI (+)、CK (+)和CD45 (-)强度在如前所述的HD-CTC计数过程中可以归类为可测量的特征。Nieva等人,Phys Biol 9:016004(2012)。本文披露的方法所采用的无富集的直接分析导致高灵敏度和高特异性,同时增添了高分辨率细胞形态学以使得能够对已知为异质性的CTC群体进行详细的形态学表征。

[0057] 虽然可以将CTC确认为包括DAPI (+)、CK (+)和CD45 (-)细胞,但本发明的方法可以用本领域技术人员选择用于生成CTC数据和/或确认CTC和CTC簇的任何其他生物标志物来实施。本领域技术人员知道如何选择形态特征、生物分子或生物分子的片段,其变化和/或检测可以与CTC相关联。分子生物标志物包含但不限于包括如下的生物分子:核苷酸、核酸、核苷、氨基酸、糖、脂肪酸、类固醇、代谢物、肽、多肽、蛋白质、碳水化合物、脂质、激素、抗体、用作生物大分子的替代物的相关区域及其组合(如糖蛋白、核糖核蛋白、脂蛋白)。该术语还涵盖生物分子的部分或片段,例如蛋白质或多肽的肽片段。

[0058] 本领域技术人员将会理解,可以使用许多方法来生成CTC数据,包含基于显微术的方法,包含荧光扫描显微术(参见,如Marrinucci D.等人,2012,Phys.Biol.9 016003);测序方法;质谱方法,如MS/MS、LC-MS/MS、多响应监测(MRM)或SRM和产物离子监测(PIM);还包含基于抗体的方法,如免疫荧光、免疫组织化学、免疫测定,如Western印迹、酶联免疫吸附测定(ELISA)、免疫沉淀、放射免疫测定、斑点印迹和FACS。免疫测定技术和方案通常是本领域技术人员已知的(Price和Newman,Principles and Practice of Immunoassay,第二版,Grove's Dictionaries,1997;以及Gosling,Immunoassays:A Practical Approach,Oxford University Press,2000)。可以使用各种免疫测定技术,包含竞争性和非竞争性免疫测定法(Self等人,Curr.Opin.Biotechnol.,7:60-65(1996),还参见John R.Crowther,

The ELISA Guidebook, 第一版, Humana Press 2000, ISBN 0896037282 以及 An Introduction to Radioimmunoassay and Related Techniques, Chard T 编辑, Elsevier Science 1995, ISBN 0444821198)。

[0059] 本领域已知且未具体描述的标准分子生物学技术通常遵循 Sambrook 等人, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, New York (1989), 以及 Ausubel 等人, *Current Protocols in Molecular Biology*, John Wiley and Sons, Baltimore, Md. (1989) 以及 Perbal, *A Practical Guide to Molecular Cloning*, John Wiley & Sons, New York (1988), 以及 Watson 等人, *Recombinant DNA*, Scientific American Books, New York 以及 Birren 等人 (编辑) *Genome Analysis: A Laboratory Manual Series*, 1-4 卷 Cold Spring Harbor Laboratory Press, New York (1998)。聚合酶链式反应 (PCR) 通常可以按照 PCR 方案进行: *A Guide to Methods and Applications*, Academic Press, San Diego, Calif. (1990)。如果分辨率足以确认本发明的所述生物标志物, 则任何能够确定特定样本的 DNA 拷贝数图谱的方法都可用于根据本发明的分子图谱分析。本领域技术人员知道并能够使用许多不同的平台来以足以确认本发明的所述一种或多种生物标志物的拷贝数的分辨率评估全基因组拷贝数变化。

[0060] 原位杂交测定法是熟知的并且通常描述于 Angerer 等人, *Methods Enzymol.* 152: 649-660 (1987)。在原位杂交测定中, 将细胞, 例如来自活检的细胞, 固定到固体载体, 通常是载玻片上。如果要探测 DNA, 会用热或碱使该细胞变性。然后使该细胞在中等温度下与杂交溶液接触以允许标记过的特异性探针退火。该探针优选使用放射性同位素或荧光报道分子标记。FISH (荧光原位杂交) 使用荧光探针, 其仅结合序列的那些与其有高度序列相似性的部分。

[0061] FISH 是用于检测和定位细胞中特定多核苷酸序列的细胞遗传学技术。例如, FISH 可以用于检测染色体上的 DNA 序列。FISH 也可以用于检测和定位组织样本内的特定 RNA, 例如 mRNA。在 FISH 中使用荧光探针, 其结合与其有高度序列相似性的特定的核苷酸序列。可以使用荧光显微术来查明该荧光探针是否结合以及在哪里结合。除了检测特定的核苷酸序列, 例如易位、融合、断裂、重复和其他染色体异常之外, FISH 还可以帮助定义细胞和组织内特定基因拷贝数和/或基因表达的空间-时间模式。

[0062] 核酸测序技术是适于基因表达分析的方法。这些方法的基本原理是, 样本中检测到 cDNA 序列的次数与对应于该序列的 RNA 的相对表达直接相关。这些方法有时用术语数字基因表达 (DGE) 来指代以反映所得数据的离散数字属性。应用这一原理的早期方法是基因表达系列分析 (SAGE) 和大规模平行标签测序 (MPSS)。参见, 例如 S. Brenner, 等人, *Nature Biotechnology* 18 (6): 630-634 (2000)。最近, “下一代” 测序技术的出现使得 DGE 更简单, 通量更高, 价格更实惠。因此, 更多的实验室能够利用 DGE 来筛选比以前可能的更多的个体患者样本中更多基因的表达。参见, 例如 J. Marionni, *Genome Research* 18 (9): 1509-1517 (2008); R. Morin, *Genome Research* 18 (4): 610-621 (2008); A. Mortazavi, *Nature Methods* 5 (7): 621-628 (2008); N. Cloonan, *Nature Methods* 5 (7): 613-619 (2008)。J. Marionni, *Genome Research* 18 (9): 1509-1517 (2008); R. Morin, *Genome Research* 18 (4): 610-621 (2008); A. Mortazavi, *Nature Methods* 5 (7): 621-628 (2008); N. Cloonan, *Nature Methods* 5 (7): 613-619 (2008)。

[0063] 本领域技术人员将进一步理解,可以使用本领域已知的任何类别的标志物特异性结合试剂来检测生物标志物的存在或不存在,包含,例如抗体、适体、融合蛋白(如包含蛋白质受体或蛋白质配体组分的融合蛋白)或生物标志物特异性小分子结合物。在一些实施方式中,CK或CD45的存在或不存在由抗体决定。技术人员将进一步理解,可以通过评估生物标志物的染色体基因座处的染色体拷贝数变化来测量生物标志物的存在或不存在。基因组生物标志物可以通过任何技术,例如(作为举例)比较基因组杂交(CGH)或通过细胞系(如癌细胞)的单核苷酸多态性阵列(基因分型微阵列)来确认。除了使用诸如qPCR或原位杂交之类的技术进行进一步分析之外,生物信息学方法还可以使用针对扩增和删除的适当拷贝数阈值来确认区分细胞系群并指示生物标志物的染色体畸变区域。用于检测染色体DNA拷贝数变化的核酸检测方法包含:(i)对完整组织或细胞样本的原位杂交检测,(ii)对提取自组织样本的染色体DNA的微阵列杂交检测,以及(iii)提取自组织样本的染色体DNA的聚合酶链式反应(PCR)或其他扩增法。使用核酸的任何这些形式的合成类似物(如肽核酸)的检测也可以使用。

[0064] 该生物标志物可以使用可检测标记的基于核酸的探针(例如脱氧核糖核酸(DNA)探针或蛋白质核酸(PNA)探针)或经设计/选择以与经设计的特异性染色体靶杂交的未标记引物,通过杂交检测来检测。该未标记引物,例如通过聚合酶链式反应(PCR,其中在引物结合后,聚合酶扩增靶核酸序列用于随后的检测)用于扩增检测。用于PCR或其他扩增检测的检测探针优选是荧光的,并且还更优选地是可用于“实时PCR”的检测探针。荧光标记还优选用于原位杂交,但也可以使用杂交技术中常用的其他可检测标记,例如酶促标记、显色标记和同位素标记。有用的探针标记技术描述于Molecular Cytogenetics:Protocols and Applications,Y.-S.Fan,编辑,第二章,“Labeling Fluorescence In Situ Hybridization Probes for Genomic Targets”,L.Morrison等人,21-40页,Humana Press,.COPYRGT.2002,通过引入并入本文。在通过微阵列分析检测该基因组生物标志物时,应用这些探针标记技术来标记来自患者样本的染色体DNA提取物,然后将其与微阵列杂交。

[0065] 在其他实施方式中,生物标志物蛋白质可以通过免疫学手段或其他蛋白质测定法来检测。可用于本发明以测量生物标志物水平的蛋白质测定方法可以包括(i)涉及标记的抗体或蛋白质与表达的生物标志物的结合的免疫测定方法,(ii)确定表达的生物标志物的质谱法,以及(iii)针对表达的生物标志物的基于蛋白质组的测定或“蛋白芯片”测定。有用的免疫测定方法包含使用本领域已知的任何形式进行的溶液相测定,例如但不限于ELISA形式、夹心形式、竞争性抑制形式(包括正向或反向竞争抑制测定)或荧光极化形式,以及固相测定,如免疫组织化学(称为“IHC”)。

[0066] 本披露的抗体特异性结合生物标志物。该抗体可以使用本领域已知的任何合适的方法来制备。参见,例如Coligan,Current Protocols in Immunology(1991);Harlow& Lane,Antibodies:A Laboratory Manual(1988);Goding,Monoclonal Antibodies:Principles and Practice(第二版,1986)。该抗体可以是天然的或者全部或部分合成产生的任何免疫球蛋白或其衍生物。其所有的保持特异性结合能力的衍生物也包括在该术语中。该抗体具有与免疫球蛋白结合结构域同源或大部分同源的结合结构域,并且可以来自天然来源;或者部分或全部合成产生。该抗体可以是单克隆或多克隆抗体。在一些实施方式

中,抗体是单链抗体。本领域普通技术人员将理解,抗体可以以多种形式(包含,例如人源化、部分人源化、嵌合、嵌合人源化等)中的任何形式提供。该抗体可以是抗体片段,包括但不限于Fab、Fab'、F(ab')<sub>2</sub>、scFv、Fv、dsFv双抗体,以及Fd片段。该抗体可以通过任何方式产生。例如,该抗体可以通过完整抗体的片段化来酶促或化学产生和/或其可以由编码部分抗体序列的基因重组产生。该抗体可以包括单链抗体片段。替代地或附加地,该抗体可以包括,例如通过二硫键连接在一起的多条链,以及获自这样的分子的任何功能片段,其中这样的片段保留了亲本抗体分子的特异性结合特性。由于其尺寸较小,作为整个分子的功能组分,抗体片段可以提供优于完整抗体的优点以用于某些免疫化学技术和实验应用中。

[0067] 当在本发明的方法中生成CTC数据时,可以在本文所述的方法中使用可检测标记来直接或间接检测该生物标志物。可以使用多种可检测标记,其中标记的选择依据所需的灵敏度、与该抗体缀合的便宜性、稳定性要求以及可用的仪器和处置条款。本领域技术人员熟悉基于本发明的方法中该生物标志物的测定检测来选择合适的可检测标记。合适的可检测标记包含但不限于荧光染料(例如荧光素、异硫氰酸荧光素(FITC)、Oregon Green™、罗丹明、德克萨斯红、四罗丹明异硫氰酸酯(TRITC)、Cy3、Cy5、AlexaFluor®647、AlexaFluor®555、AlexaFluor®488)、荧光标志物(如绿色荧光蛋白(GFP)、藻红蛋白等)、酶(如萤光素酶、辣根过氧化物酶、碱性磷酸酶等)、纳米颗粒、生物素、地高辛、金属,以及诸如此类。

[0068] 对于基于质谱学(mass-pectrometry)的分析,使用同位素试剂(如同位素编码的亲标签(ICAT))的差异标签或使用等压标签试剂iTRAQ(Applied Biosystems, Foster City, Calif.)的更近期的变体,继之以多维液相色谱(LC)和串联质谱(MS/MS)分析可以为实施本披露的方法提供进一步的方法论。

[0069] 使用化学发光抗体的化学发光测定法可以用于对蛋白质的灵敏的非放射性检测。用荧光物标记的抗体也可能是合适的。荧光物的实例包含但不限于DAPI、荧光素、Hoechst 33258、R-藻蓝蛋白、B-藻红蛋白、R-藻红蛋白、罗丹明、德克萨斯红和丽丝胺。间接标记包含本领域熟知的各种酶,例如辣根过氧化物酶(HRP)、碱性磷酸酶(AP)、β-半乳糖苷酶、尿素酶,以及诸如此类。使用适于辣根过氧化物酶、碱性磷酸酶、β-半乳糖苷酶的底物的检测系统为本领域所熟知。

[0070] 来自直接或间接标记的信号可以,例如使用显微镜(例如荧光显微镜或荧光扫描显微镜)进行分析。替代地,分光光度计可用来检测显色底物的颜色;辐射计数器可用来检测辐射,例如γ计数器用于检测<sup>125</sup>I;或荧光计可用来在特定波长的光的存在下检测荧光。如果需要,用于实施本披露的方法的测定法可以自动化或者由机器人执行,并且来自多个样本的信号可以同时检测。

[0071] 在一些实施方式中,该生物标志物是免疫荧光标志物。在一些实施方式中,该免疫荧光标志物包括对上皮细胞特异性的标志物。在一些实施方式中,该免疫荧光标志物包括对白细胞(WBC)特异性的标志物。在一些实施方式中,该免疫荧光标志物中的一种或多种包括CD45和CK。

[0072] 在一些实施方式中,有核细胞(如CTC或WBC)中免疫荧光标志物的存在或不存在导致区别性免疫荧光染色模式。CTC和WBC的免疫荧光染色模式可能因检测相应细胞中上皮细胞还是WBC标志物而不同。在一些实施方式中,确定一种或多种免疫荧光标志物的存在或不包括使用,例如明显地确认WBC的CD45的免疫荧光染色,比较CTC的区别性免疫荧光染

色与WBC的区别性免疫荧光染色。还有其他的与WBC的各种亚群结合的可检测标志物或可检测标志物的组合。这些可以以各种组合的形式使用,包含与CD45的免疫荧光染色组合或作为其替代。

[0073] 在一些实施方式中,CTC包括与周围的有核细胞不同的形态特征。在一些实施方式中,该形态特征包括细胞核大小、细胞核形状、细胞大小、细胞形状;和/或细胞核与细胞质的比率。在一些实施方式中,该方法进一步包括通过细胞核细节、细胞核轮廓、细胞核的存在或不存在、细胞质的质量、细胞质的数量、免疫荧光染色模式的强度来分析该有核细胞。本领域的普通技术人员理解,本披露的该形态特征可以包含可以确定并与CTC的检测相关联的细胞的任何特性、性质、特征或方面。

[0074] CTC数据可以用本领域已知的任何微观方法来生成。在一些实施方式中,该方法通过荧光扫描显微术进行。在某些实施方式中,显微方法提供CTC及其周围WBC的高分辨率图像(参见,例如Marrinucci D.等人,2012,Phys.Biol.9 016003)。在一些实施方式中,通过荧光扫描显微镜扫描用来自样本(如非富集血液样本)的单层有核细胞涂覆的载片,并且记录来自免疫荧光标志物和核染色物的荧光强度以允许测定每个免疫荧光标志物的存在或不存在以及评估该有核细胞的形态。在一些实施方式中,显微数据收集和分析以自动方式进行。

[0075] 在一些实施方式中,CTC数据包含对一种或多种生物标志物(例如CK和CD45)的检测。如果生物标志物在所使用的相应检测方法的背景噪声以上是可检测到的(例如是该背景的2倍、3倍、5倍或10倍;例如相比背景 $2\sigma$ 或 $3\sigma$ ),则其被认为“存在”于细胞中。在一些实施方式中,如果生物标志物在所用的检测方法的背景噪声以上是不可检测到的,则认为其是“不存在的”(例如,是背景信号的 $<1.5$ 倍或 $<2.0$ 倍;例如相比背景 $<1.5\sigma$ 或 $<2.0\sigma$ )。

[0076] 在一些实施方式中,有核细胞中免疫荧光标志物的存在或不存在通过选择荧光扫描过程期间的暴露时间来确定,使得所有免疫荧光标志物在视野中在WBC上达到预设的荧光水平。在这些条件下,CTC特异性免疫荧光标志物即使不存在于WBC上,在该WBC内作为具有固定高度的背景信号也是可见的。此外,CTC上不存在的WBC特异性免疫荧光标志物在该CTC内作为具有固定高度的背景信号是可见的。如果一种细胞针对相应的免疫荧光标志物的荧光信号显着高于固定的背景信号(例如,是该背景的2倍、3倍、5倍或10倍;例如相比背景 $2\sigma$ 或 $3\sigma$ ),则认为该细胞对于该标志物是阳性的(即该标志物被认为是存在的)。例如,有核细胞针对CD 45的荧光信号显着高于背景信号,则该细胞被认为是CD 45阳性(CD 45<sup>+</sup>)的。如果一种细胞针对相应免疫荧光标志物的荧光信号不显着高于背景信号(例如,是背景信号的 $<1.5$ 倍或 $<2.0$ 倍;例如,相比背景 $<1.5\sigma$ 或 $<2.0\sigma$ ),则该细胞被认为对于该标志物是阴性的(即该标志物被认为是不存在的)。

[0077] 通常,每个显微镜视野都含有CTC和WBC。在某些实施方式中,该显微镜视野示出了至少1、5、10、20、50或100个CTC。在某些实施方式中,该显微镜视野示出的WBC是CTC的至少10、25、50、100、250、500或1,000倍。在某些实施方式中,该显微镜视野包括由至少10、50、100、150、200、250、500、1,000或更多个WBC包围的一个或多个CTC或CTC簇。

[0078] 在本文所述的方法的一些实施方式中,该CTC数据的生成包括对存在于该血液样本中的CTC的计数。在一些实施方式中,本文所述的方法涵盖检测至少1.0个CTC/mL血液、1.5个CTC/mL血液、2.0个CTC/mL血液、2.5个CTC/mL血液、3.0个CTC/mL血液、3.5个CTC/mL血

液、4.0个CTC/mL血液、4.5个CTC/mL血液、5.0个CTC/mL血液、5.5个CTC/mL血液、6.0个CTC/mL血液、6.5个CTC/mL血液、7.0个CTC/mL血液、7.5个CTC/mL血液、8.0个CTC/mL血液、8.5个CTC/mL血液、9.0个CTC/mL血液、9.5个CTC/mL血液、10个CTC/mL血液或更多。

[0079] 在本文所述的方法的一些实施方式中,该CTC数据的生成包括检测CTC(包含非传统CTC)的不同亚型。在一些实施方式中,本文所述的方法涵盖检测至少0.1个CTC簇/mL血液、0.2个CTC簇/mL血液、0.3个CTC簇/mL血液、0.4个CTC簇/mL血液、0.5个CTC簇/mL血液、0.6个CTC簇/mL血液、0.7个CTC簇/mL血液、0.8个CTC簇/mL血液、0.9个CTC簇/mL血液、1个CTC簇/mL血液、2个CTC簇/mL血液、3个CTC簇/mL血液、4个CTC簇/mL血液、5个CTC簇/mL血液、6个CTC簇/mL血液、7个CTC簇/mL血液、8个CTC簇/mL血液、9个CTC簇/mL血液、10个簇/mL或更多。在具体的实施方式中,本文所述的方法包括检测至少1个CTC簇/mL血液。

[0080] 在一些实施方式中,所披露的方法涵盖预测模型的使用。在进一步的实施方式中,所披露的方法涵盖将可测量特征与参考特征进行比较。如本领域技术人员可以理解的,这种比较可以是与该参考特征的直接比较或者已经将该参考特征合并到该预测模型中的间接比较。在进一步的实施方式中,分析可测量值涵盖以下的一种或多种:线性判别分析模型、支持向量机分类算法、递归特征消除模型、微阵列模型的预测分析、逻辑回归模型、CART算法、柔性树算法(flex tree algorithm)、LART算法、随机森林算法、MART算法、机器学习算法、惩罚回归方法或其组合。在具体的实施方式中,分析包括逻辑回归。在另外的实施方式中,确定被表达为风险分数。

[0081] 分析分类过程可以使用各种统计分析方法中的任何一种来操作定量数据并提供对样本的分类。有用的方法的例子包括线性判别分析、递归特征消除、微阵列的预测分析、逻辑回归、CART算法、柔性树算法、LART算法、随机森林算法、MART算法、机器学习算法和本领域技术人员已知的其他方法。

[0082] 分类可以根据设定用于确定样本属于给定类别的概率的阈值的预测建模方法进行。该概率优选为至少50%;或者至少60%;或者至少70%;或者至少80%;或者至少90%;或者更高。分类还可以通过确定所获得的数据集与参考数据集之间的比较是否产生统计学显著性差异来进行。如果是,则从中获得该数据集的样本被分类为不属于该参考数据集类别。相反,如果这样的比较与该参考数据集在统计学上没有显著差异,则从中获得该数据集的样本被分类为属于该参考数据集类。

[0083] 模型的预测能力可以根据其提供质量度量,例如AUROC(ROC曲线下面积)或特定值或值范围的准确度的能力来评估。所述曲线下面积测量值可用于比较整个数据范围内分类器的准确度。具有较大AUC的分类器具有更大的在两组相关的组之间正确分类未知数的能力。ROC分析可以用于在各种临床情况下选择最佳阈值,从而平衡特异性和敏感性之间存在的内在折衷(inherent tradeoff)。在一些实施方式中,期望的质量阈值是预测模型,其将以至少约0.7、至少约0.75、至少约0.8、至少约0.85、至少约0.9、至少约0.95或更高的准确度进行样本分类。作为替代的措施,期望的质量阈值可以指代预测模型,其将以至少约0.7、至少约0.75、至少约0.8、至少约0.85、至少约0.9或更高的AUC进行样本分类。

[0084] 如本领域已知的,预测模型的相对灵敏度和特异性可以经调整以支持特异性度量或灵敏性度量,其中这两个度量具有反比关系。上述模型中的限值可以依据正在执行的测试的具体要求,经调整以提供选定的灵敏性或特异性水平。灵敏性和特异性之一或二者可

以是至少约0.7、至少约0.75、至少约0.8、至少约0.85、至少约0.9或更高的值。

[0085] 原始数据可以通过通常以一式三份或多种一式三份,测量每个可测量特征或生物标志物的值进行初始分析。可以对该数据进行操作,例如,可以使用标准曲线对原始数据进行转换,并将三次测量的平均值用于计算每位患者的平均值和标准偏差。这些值可以在用于模型之前被转换,例如,对数转换、Box-Cox转换(Box和Cox, *Royal Stat. Soc.*, B系列26: 211-246 (1964))。然后将该数据输入到预测模型中,其将根据状态对该样本进行分类。所得到的信息可以传达给患者或医疗保健提供者。在一些实施方式中,该方法具有>60%、>70%、>80%、>90%或更高的特异性。

[0086] 如本领域技术人员将理解的,分析分类过程可以使用各种统计分析方法中的任何一种来操作该定量数据并提供对该样本的分类。有用的方法的例子包括但不限于线性判别分析、递归特征消除、微阵列的预测分析、逻辑回归、CART算法、柔性树算法、LART算法、随机森林算法、MART算法和机器学习算法。

[0087] 在另一个实施方式中,本披露内容提供了用于测量生物标志物水平的试剂盒,其包括含有至少一种标记的探针、蛋白质或特异性结合样本中至少一种表达的生物标志物的抗体的容器。这些试剂盒还可以包含具有用于测定的其他相关试剂的容器。在一些实施方式中,试剂盒包括含有用于结合生物标志物的标记的单克隆抗体或核酸探针和至少一种校准物组合物的容器。该试剂盒可以进一步包括检测该可检测标记(例如酶或底物)所必需的组分。该试剂盒还可以含有可以进行测定并与测试样本进行比较的对照样本或一系列对照样本。该试剂盒的每个组件都可以封装在单独的容器内,并且各种容器全部可以含在单个包装件内,并附带用于解释使用该试剂盒进行的测定的结果的说明书。

[0088] 根据前面的描述,将显而易见的是,可以对本文所述的发明进行变更和修改,以将其用于各种用法和条件。这些实施方式也在以下权利要求书的范围内。

[0089] 本文变量的任何定义中的元素列表的叙述包含将所述变量定义为所列元素的任何单个元素或组合(或子组合)。本文实施方式的叙述包含实施方式作为任何单个实施方式或与任何其他实施方式或其部分进行组合。

[0090] 本说明书中提及的所有专利和出版物均通过引用并入本文,如同将每个独立的专利和出版物具体且单独地指示为通过引用并入。

[0091] 以下实施例是作为说明而非限制提供的。

## 实施例

[0092] 实施例1.

[0093] 如先前报道的,使用Epic Sciences Platform对CTC进行样本评估。Marrinucci等人 *Phys Biol* 9:016003, 2012; Marrinucci *Physiol Biol* 9:016003, 2012。Epic CTC收集和检测过程流程如下:(1)将来自血液样本的血液裂解有核细胞置于载片上;(2)将载片储存在-80°C生物储存库中;(3)用CK、CD45、DAPI和AR将载片染色;(4)扫描载片;(5)运行多参数数字病理学算法,以及(6)通过软件和人类读取来确认CTC对生物标志物的表达进行定量。在随后的CTC恢复和基因组图谱分析工作流程中,单个细胞被分离,经受全基因组扩增,以及NGS文库制备。在Illumina NextSeq 500上进行测序。

[0094] 血液样本经受溶血、离心、重悬并涂覆于载片上,接着是-80°C储存。在分析之前,



将载片解冻,通过免疫荧光(广谱细胞角蛋白、CD45、DAPI)标记,并通过自动荧光透视法成像,然后由病理学家培训的技术人员手动验证(MSL)。Marrinucci等人Phys Biol 9:016003,2012。如前所述,DAPI(+)、CK(+)和CD45(-)强度在CTC计数期间被归类为特性。

[0095] 更具体地,外周血液样本被收集在无细胞DNA BCT(美国内布拉斯加州奥马哈的Streck)中并在环境温度下立即运送至Epic Sciences(美国加州圣地亚哥)。在收到之后,如前所述(Marrinucci等人Hum Pathol 38(3):514-519(2007);Marrinucci等人Arch Pathol Lab Med 133(9):1468-1471(2009);Mikolajczyk等人J Oncol 2011:252361.(2011);Marrinucci等人Phys Biol 9(1):016003(2012);Werner等人J Circ Biomark 4:3(2015)),将红细胞裂解并将有核细胞分配到显微镜载玻片上,以及储存于-80℃直至染色。根据样本的白细胞计数和所使用的RBC裂解后细胞悬浮液的体积计算每个载片上涂覆的血液的毫升当量。如所述(Marrinucci等人,2007,同上;Marrinucci等人,2009,同上;Mikolajczyk等人,2011,同上;Marrinucci等人,2012,同上;Werner等人,2015,同上)的,通过免疫荧光确认循环肿瘤细胞。在随后的CTC恢复和基因组图谱分析工作流程期间,单个细胞被分离,经受全基因组扩增,以及NGS文库制备。在Illumina NextSeq 500上进行测序。

[0096] 图1至4以及相应的附图说明描述了更多的实验细节。

[0097] 本文变量的任何定义中的元素列表的叙述包含将所述变量定义为所列元素的任何单个元素或组合(或子组合)。本文实施方式的叙述包含实施方式作为任何单个实施方式或与任何其他实施方式或其部分进行组合。

[0098] 本说明书中提及的所有专利和出版物均通过引用并入本文,如同将每个独立的专利和出版物具体且单独地指示为通过引用并入。

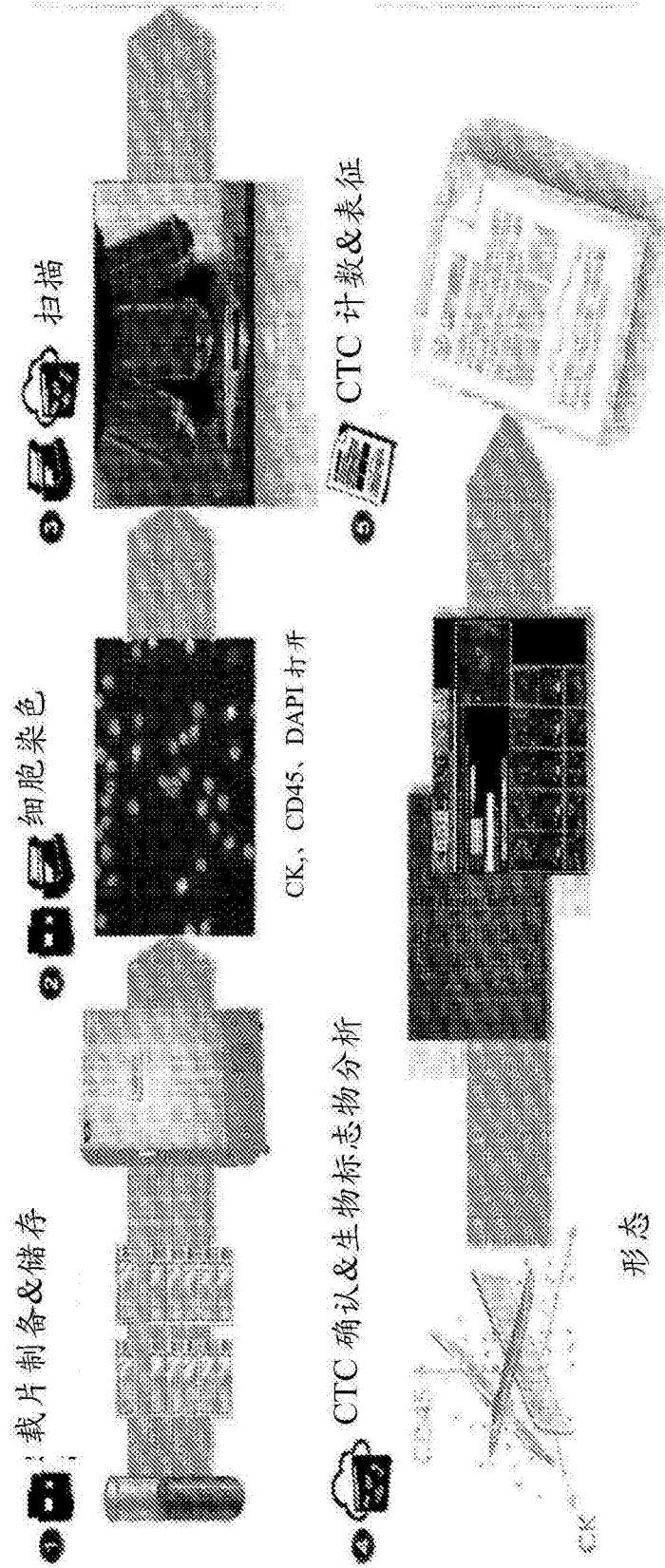


图1

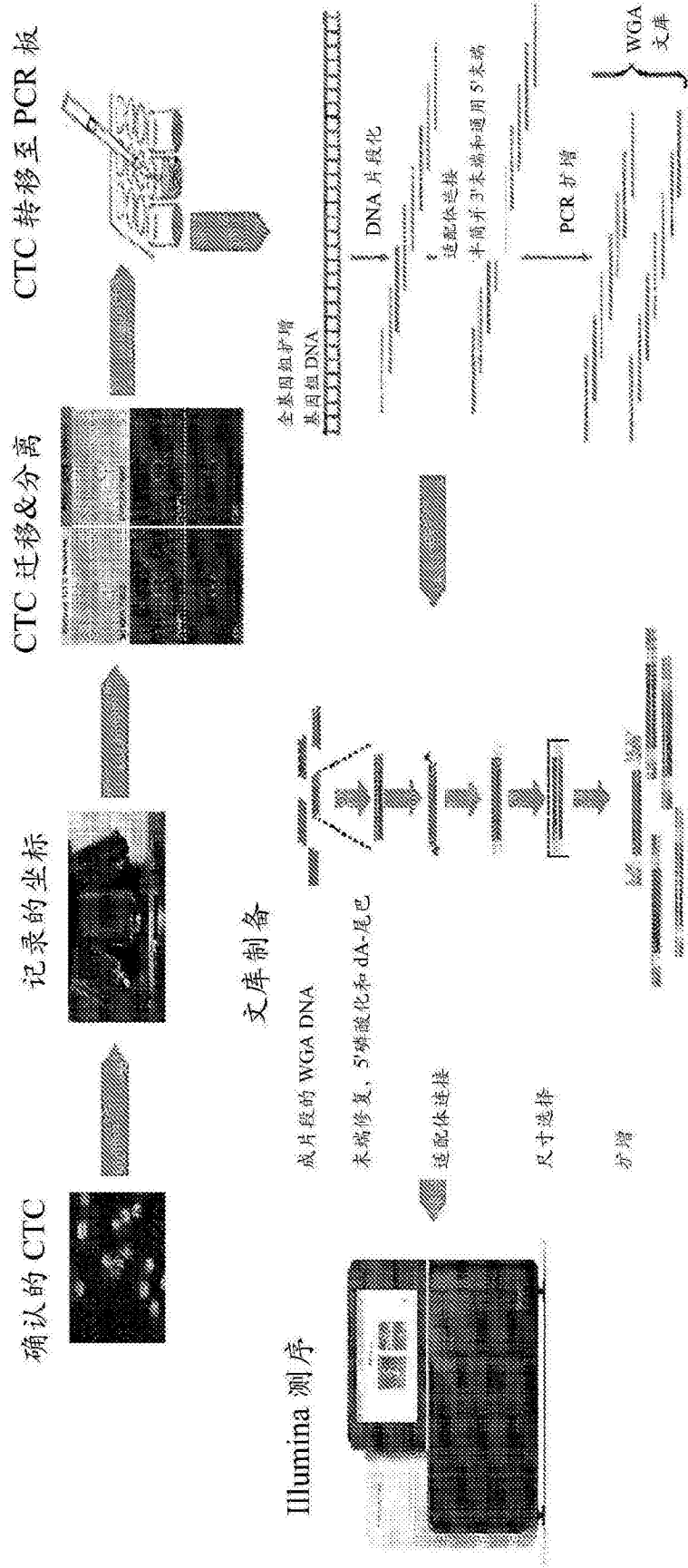


图1B

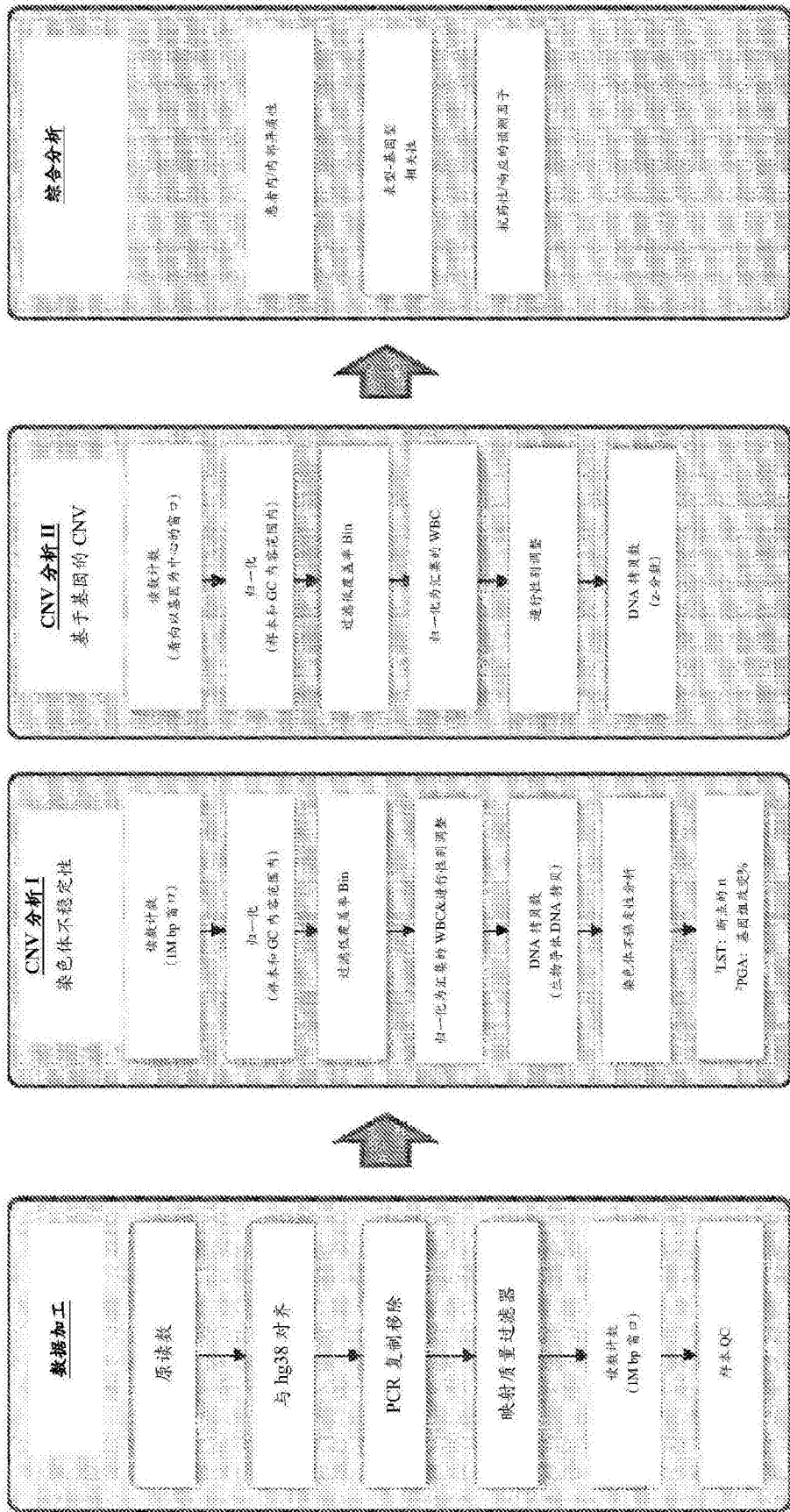


图2

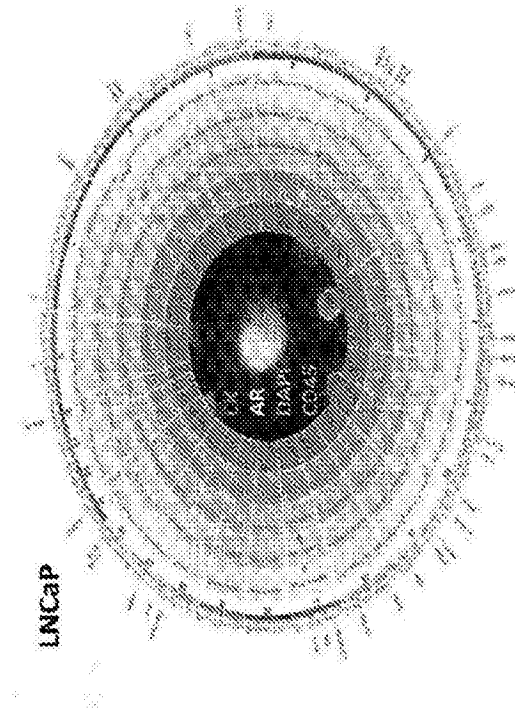


图3A

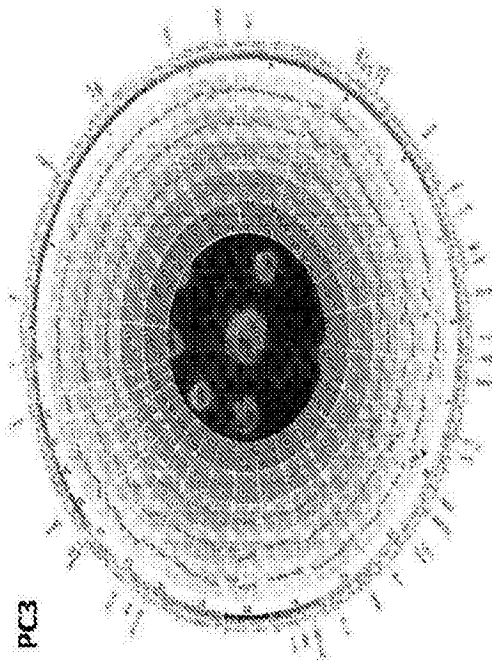


图3B

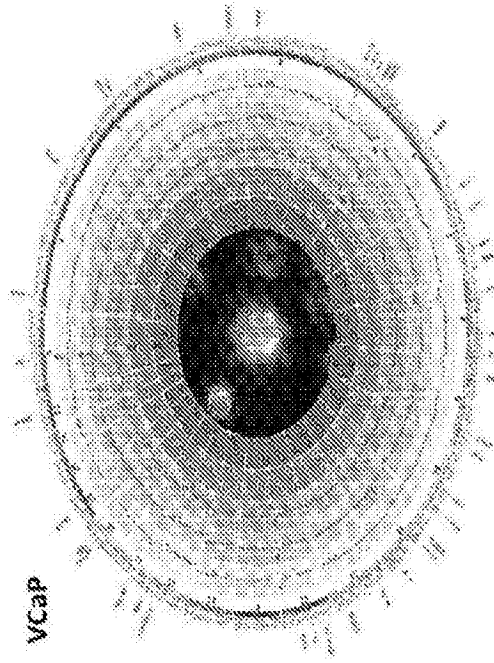


图3C

代表性细胞系	关键的基因组特征
MCF7	正常 (突变的) AR CNV; PTEN 杂合删除; wt p53
Vcap	AR 扩增; 完整的 PTEN; 正常 (突变的) p53 CNV
PC3	正常的 AR CNV; PTEN 纯合删除; p53 删除

图3D

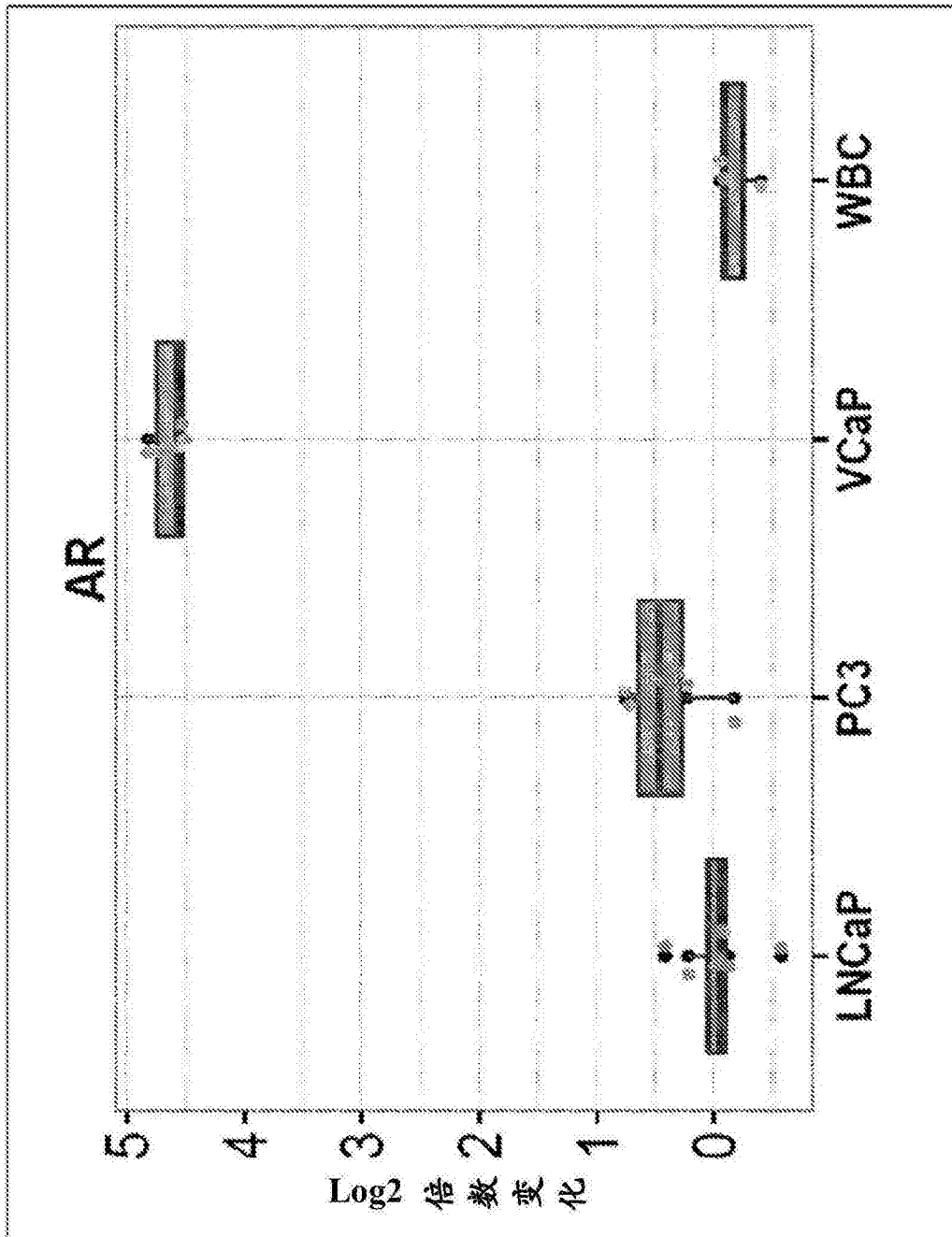


图4A

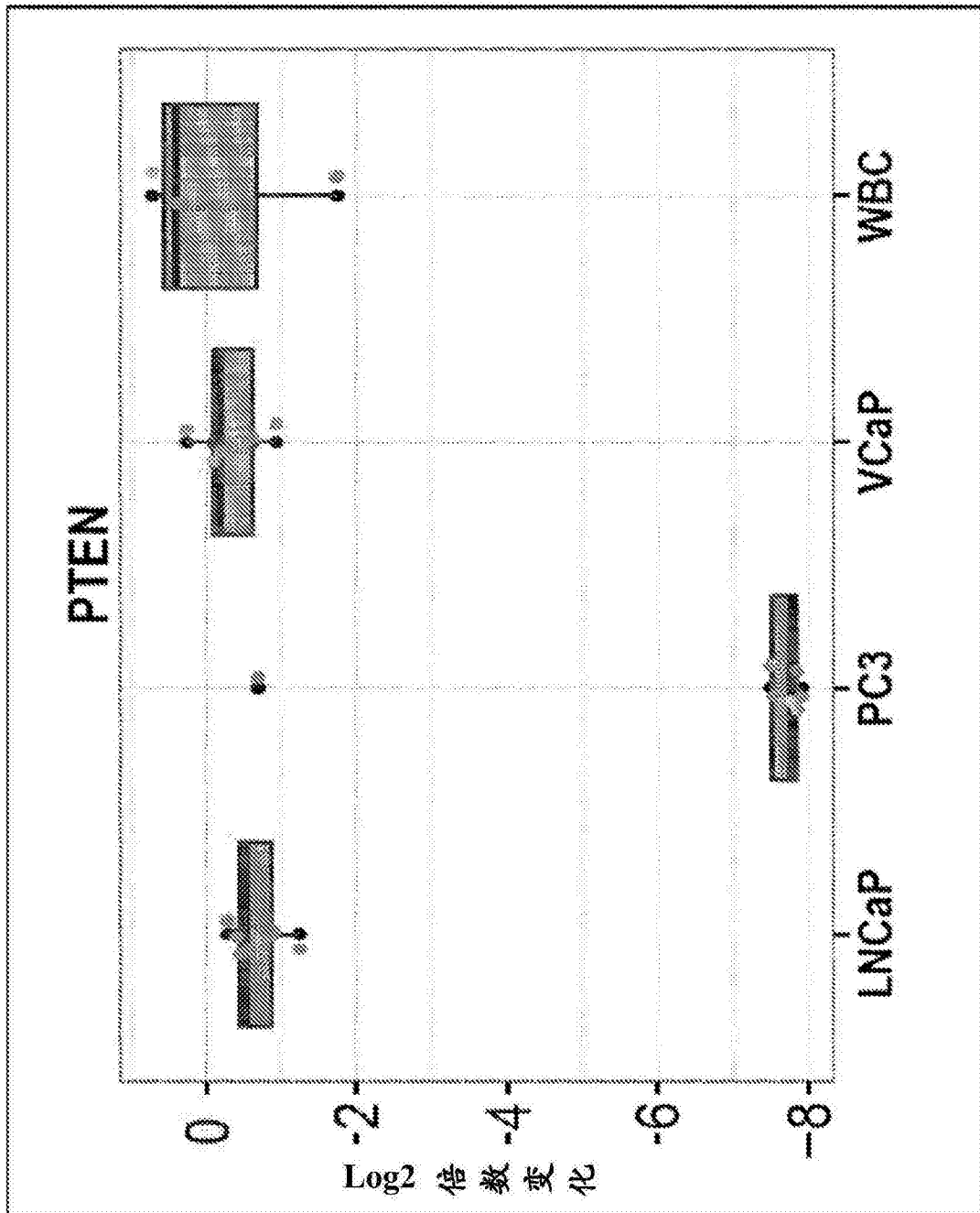


图4B



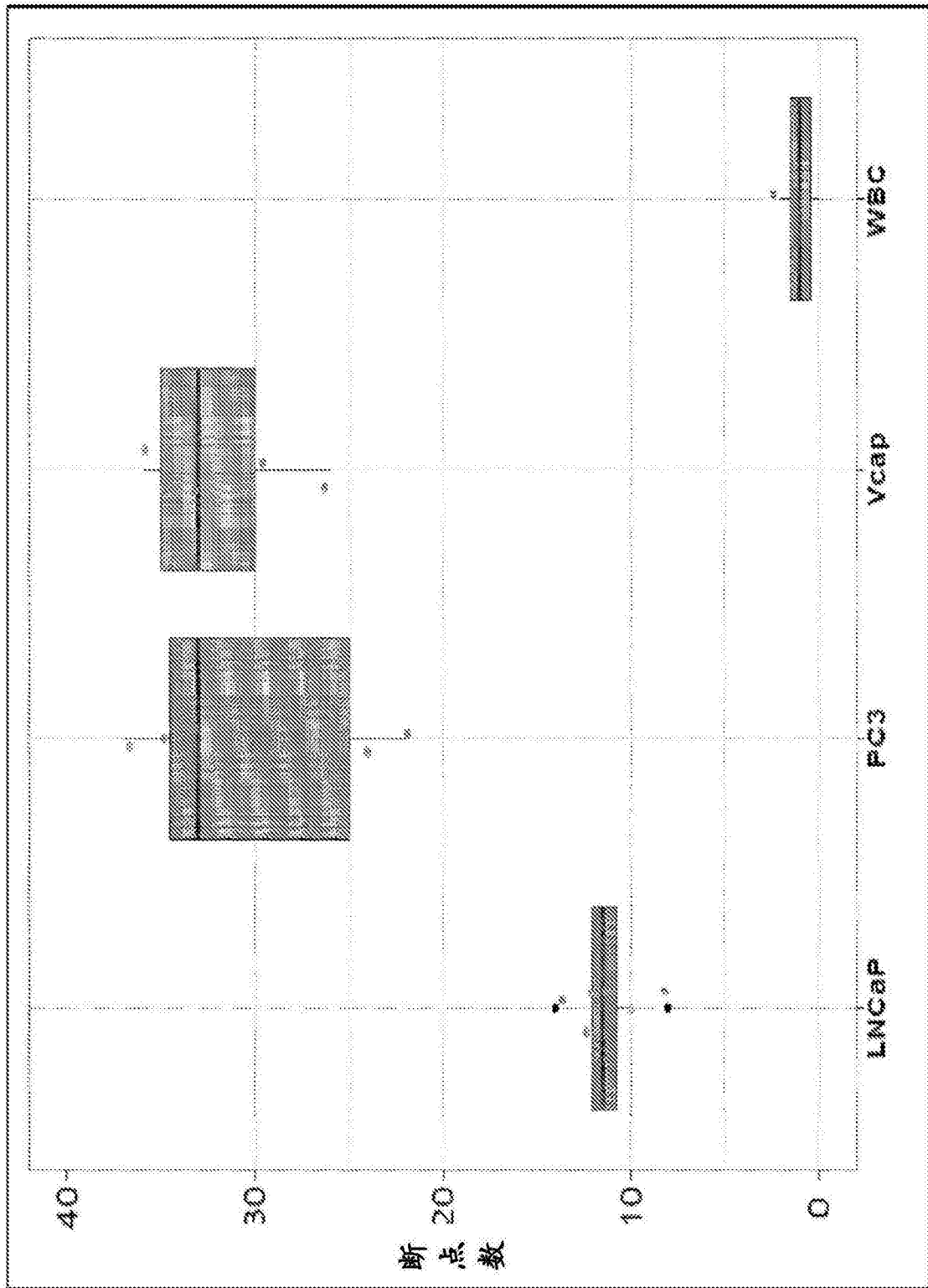


图4C

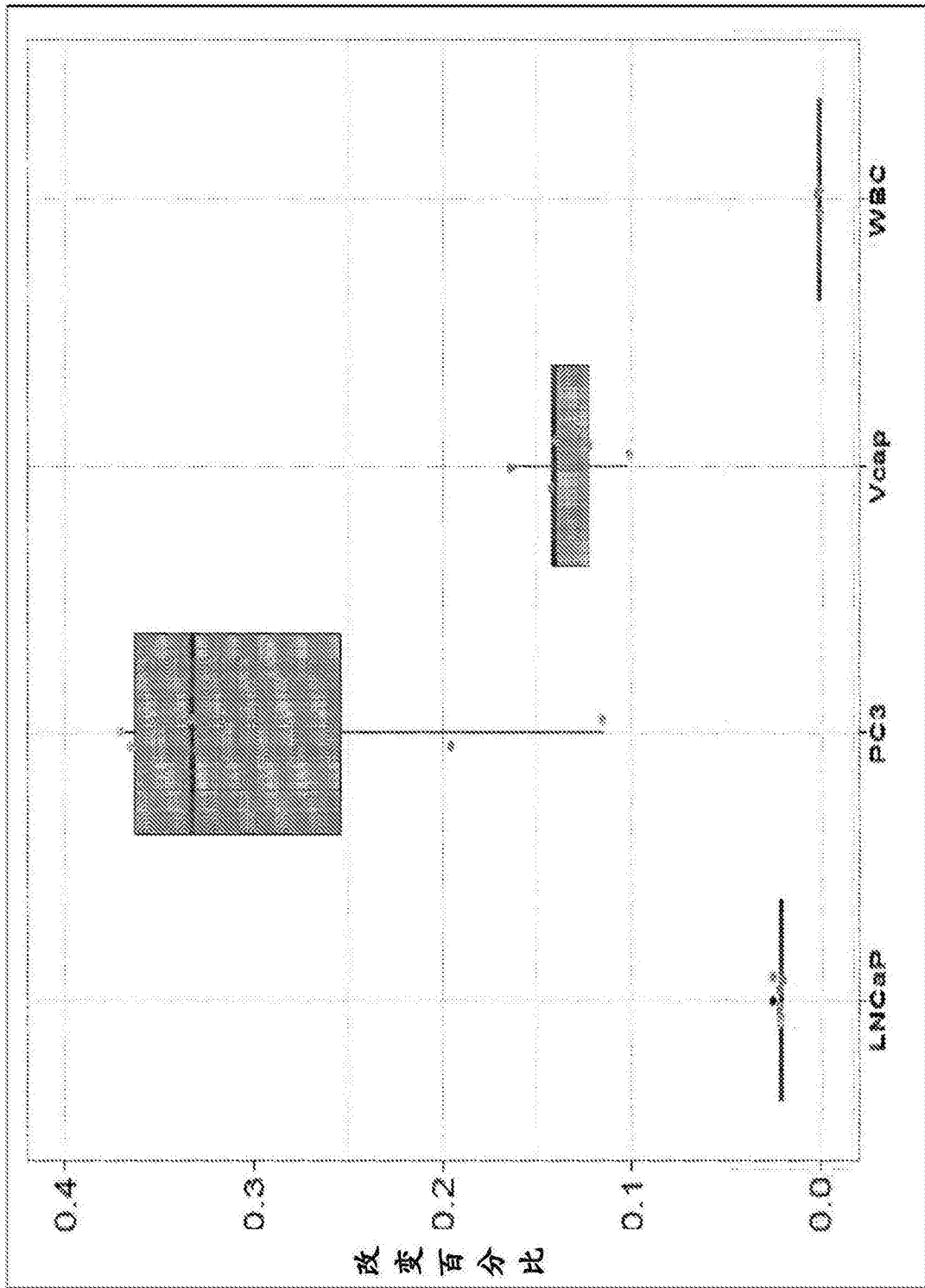


图4D