



(12) 发明专利申请

(10) 申请公布号 CN 119955906 A

(43) 申请公布日 2025.05.09

(21) 申请号 202510148128.0

(51) Int.Cl.

(22) 申请日 2019.03.15

C12Q 1/6827 (2018.01)

(30) 优先权数据

C12Q 1/6886 (2018.01)

62/643,649 2018.03.15 US

C12N 15/11 (2006.01)

62/769,928 2018.11.20 US

(62) 分案原申请数据

201980032333.3 2019.03.15

(71) 申请人 格瑞尔公司

地址 美国加利福尼亚州

(72)发明人 卢煜明 赵慧君 陈君赐 盖万霞

姬露

(74) 专利代理机构 北京安信方达知识产权代理有限公司 11262

专利代理师 张晨 郑霞

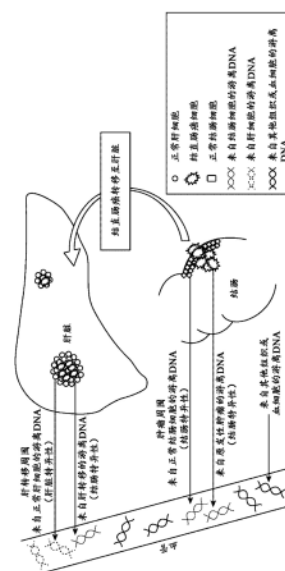
权利要求书2页 说明书26页
序列表(电子公布) 附图16页

(54) 发明名称

组织特异性的甲基化标志物

(57) 摘要

本申请涉及组织特异性的甲基化标志物。本文提供了包括用于识别游离核酸(例如,游离DNA分子)的起源组织的组织特异性标志物的组合物。本文还提供了通过确定包括组织特异性标志物的游离核酸的绝对量,识别游离核酸起源组织的方法、组合物和系统。本文还提供了通过分析组织特异性标志物来检测生物体组织中癌症的方法、组合物和系统。



1. 一种测量在一个或多个差异甲基化区域 (DMR) 中的胞嘧啶甲基化的方法, 所述方法包括:

(a) 从受试者的第一生物样品获得游离DNA分子; 和

(b) 对所述游离DNA分子进行测定, 以对所述一个或多个DMR中的每一个测量在所述DMR中包括靶序列的第一甲基化状态的所述游离DNA分子的量;

其中 (i) 所述一个或多个DMR包括Sestrin 3 (SESN3) 和蛋白酪氨酸激酶2 β (PTK2B) 中的一者或两者的靶序列; 和 (ii) 所述一个或多个DMR中的每一个的所述靶序列包括一个或多个CpG位点。

2. 根据权利要求1所述的方法, 其中所述测定包括将包括所述靶序列的所述游离DNA分子与探针杂交。

3. 根据权利要求1所述的方法, 其中所述测定包括用一对或多对引物来扩增所述游离DNA分子。

4. 根据权利要求1所述的方法, 其中所述测定包括将所述游离DNA分子中的非甲基化的胞嘧啶残基经亚硫酸氢盐转化为尿嘧啶。

5. 根据权利要求1所述的方法, 其中所述测定包括对来自所述第一生物样品的游离DNA分子进行甲基化感知测序。

6. 根据权利要求1所述的方法, 其中所述靶序列包括至少3个、至少4个、至少5个、至少6个、至少7个、至少8个、至少9个或至少10个CpG甲基化位点。

7. 根据权利要求1所述的方法, 其中所述第一甲基化状态包括所述靶序列内单个位点的甲基化密度、甲基化/非甲基化位点在所述靶序列内连续区域上的分布或所述靶序列内每个单个甲基化位点的甲基化模式或水平。

8. 根据权利要求1所述的方法, 其中所述靶序列在所述受试者的第一组织中比在所述受试者的第二组织中包括更高的甲基化密度。

9. 根据权利要求8所述的方法, 其中所述靶序列在所述第一组织中包括的甲基化密度大于50%。

10. 根据权利要求1所述的方法, 其中所述靶序列包括与SEQ ID NO: 1具有至少60%同一性的PTK2B的多核苷酸序列。

11. 根据权利要求1所述的方法, 其中所述测定包括 (i) 使用包括SEQ ID NO: 2的引物、包括SEQ ID NO: 3的引物或两者进行扩增, 或 (ii) 使用包括SEQ ID NO: 4的可检测地标记的探针用于检测所述靶序列。

12. 根据权利要求1所述的方法, 其中所述扩增包括 (i) 使用包括SEQ ID NO: 5的引物、包括SEQ ID NO: 6的引物或两者进行扩增, 或 (ii) 使用包括SEQ ID NO: 7的可检测地标记的探针用于检测所述靶序列。

13. 根据权利要求1所述的方法, 其中所述靶序列包括与SEQ ID NO: 8具有至少60%同一性的SESN3的多核苷酸序列。

14. 根据权利要求1所述的方法, 其中所述测定包括 (i) 使用包括SEQ ID NO: 9的引物、包括SEQ ID NO: 10的引物或两者进行扩增, 或 (ii) 使用包括SEQ ID NO: 11的可检测地标记的探针用于检测所述靶序列。

15. 根据权利要求1所述的方法, 其中所述测定包括 (i) 使用包括SEQ ID NO: 12的引

物、包括SEQ ID NO: 13的引物或两者进行扩增,或(ii)使用包括SEQ ID NO: 14的可检测地标记的探针用于检测所述靶序列。

16.根据权利要求1所述的方法,其中所述一个或多个DMR中的每一个的所述靶序列包括外显子序列。

17.根据权利要求1所述的方法,其中所述一个或多个DMR包括有包括SEQ ID NO:1的一个或多个CpG位点的靶序列。

18.根据权利要求1所述的方法,其中所述一个或多个DMR包括有包括SEQ ID NO:8的一个或多个CpG位点的靶序列。

19.根据权利要求1所述的方法,其中所测量的包括第一甲基化状态的所述游离DNA分子的量是绝对量。

20.根据权利要求1所述的方法,其中在所述DMR中包括所述靶序列的所述第一甲基化状态的所述游离DNA分子包括从癌症转移部位释放的DNA分子。

21.一种用于确定来自生物体肝脏的生物样品中游离DNA分子的量的组合物,包括一对引物,所述一对引物用于基于肝脏特异性标志物的甲基化水平扩增所述肝脏特异性标志物,其中所述肝脏特异性标志物包括与SEQ ID NO: 1具有至少60%、70%、80%、90%、95%、98%或99%同一性的多核苷酸序列。

22.根据权利要求21所述的组合物,其中所述一对引物包括有包括SEQ ID NO: 2的引物和包括SEQ ID NO: 3的引物。

23.根据权利要求21或22所述的组合物,还包括有包括SEQ ID NO: 4的可检测地标记的探针,用于检测所述肝脏特异性标志物。

24.根据权利要求21-23中任一项所述的组合物,还包括有包括SEQ ID NO: 5的引物和包括SEQ ID NO: 6的引物。

25.根据权利要求21-24中任一项所述的组合物,还包括有包括SEQ ID NO: 7的可检测地标记的探针,用于检测所述肝脏特异性标志物。

26.一种用于确定来自生物体结肠的生物样品中游离DNA分子的量的组合物,包括一对引物,所述一对引物用于基于结肠特异性标志物的甲基化水平扩增所述结肠特异性标志物,其中所述结肠特异性标志物包括与SEQ ID NO: 8具有至少60%、70%、80%、90%、95%、98%或99%同一性的多核苷酸序列。

27.根据权利要求26所述的组合物,其中所述一对引物包括有包括SEQ ID NO: 9的引物和包括SEQ ID NO: 10的引物。

28.根据权利要求26或27所述的组合物,还包括有包括SEQ ID NO: 11的可检测地标记的探针,用于检测所述结肠特异性标志物。

29.根据权利要求26-28中任一项所述的组合物,还包括有包括SEQ ID NO: 12的引物和包括SEQ ID NO: 13的引物。

30.根据权利要求26-29中任一项所述的组合物,还包括有包括SEQ ID NO: 14的可检测地标记的探针,用于检测所述结肠特异性标志物。

组织特异性的甲基化标志物

[0001] 本申请是申请日为2019年3月15日,申请号为201980032333.3,发明名称为“组织特异性的甲基化标志物”的申请的分案申请。

交叉引用

[0002] 本申请要求于2018年3月15日提交的美国临时申请号62/643,649,以及于2018年11月20日提交的美国临时申请号62/769,928的权益,其各自通过引用整体并入本文。

背景技术

[0003] 从不同组织到循环DNA的DNA定量测量可以潜在地提供有关许多不同病理状况的存在的重要信息。然而,涉及全基因组亚硫酸氢盐测序的现有的方法是相对昂贵的,并且可能对分析带来挑战。用于测量来自不同组织的DNA的更具成本效益的方法将是有益的。

[0004] 来自癌细胞的循环游离DNA的检测,通常称为液体活检,越来越多地用于癌症患者的管理。例如,血浆中表皮生长因子受体(EGFR)突变的检测可以与肿瘤组织中的突变状态很好地关联,并且可以预测对EGFR酪氨酸激酶抑制剂的反应性。除点突变外,其他与癌症相关的遗传和基因组改变,包括拷贝数改变和裂解方式改变,也可以在癌症患者的游离血浆中检测到。与未进行筛查的患者相比,通过筛查血浆DNA识别的患者可能具有明显的早期阶段分布和更优异的无进展生存期。

发明内容

[0005] 本公开内容的一个方面提供了一种确定患有第一组织癌症的生物体是否患有位于第二组织的癌症的方法,所述方法包括:(a)从患有第一组织癌症的所述生物体的第一生物样品获得游离DNA分子;(b)对所述游离DNA分子进行测定,以确定在所述游离DNA分子中靶序列的第一甲基化状态,其中所述靶序列的所述第一甲基化状态表明包括所述靶序列的游离DNA分子来自所述生物体的第二组织,其中所述第一组织和所述第二组织是不同的;(c)从包括具有所述第一甲基化状态的所述靶序列的所述第一生物样品确定游离DNA分子的绝对量;以及(d)根据所述绝对量确定所述生物体是否患有位于所述第二组织的癌症。

[0006] 在一些情况下,所述甲基化状态包括甲基化水平。在一些情况下,所述测定包括从所述第一生物样品中分离包括所述靶序列的所述游离DNA分子。在一些情况下,所述测定包括在油乳剂中分离包括所述靶序列的所述游离DNA分子。在一些情况下,所述测定包括将包括所述靶序列的所述游离DNA分子与探针杂交。在一些情况下,所述探针与所述靶序列杂交。在一些情况下,所述探针与所述靶序列的杂交亲和力取决于所述第一生物样品中所述靶序列的所述第一甲基化状态。在一些情况下,当所述靶序列的甲基化位点在所述第一生物样品中被甲基化时,所述探针与所述靶序列杂交。在一些情况下,当所述靶序列的甲基化位点在所述第一生物样品中未被甲基化时,所述探针与所述靶序列杂交。在一些情况下,所述测定包括检测所述探针与所述靶序列的所述杂交。

[0007] 在一些情况下,所述测定包括扩增所述游离DNA分子。在一些情况下,所述扩增包括使用一对引物。在一些情况下,所述一对引物中的至少一个引物与所述靶序列的杂交亲

和力取决于所述靶序列的所述第一甲基化状态。在一些情况下,当所述靶序列的甲基化位点在所述第一生物样品中被甲基化时,所述一对引物中的所述至少一个引物与所述靶序列杂交。在一些情况下,当所述靶序列的甲基化位点在所述第一生物样品中未被甲基化时,所述一对引物中的所述至少一个引物与所述靶序列杂交。在一些情况下,所述测定包括将所述游离DNA分子中的非甲基化的胞嘧啶残基经亚硫酸氢盐转化为尿嘧啶。在一些情况下,所述测定包括对来自所述第一生物样品的游离DNA分子进行甲基化感知测序。

[0008] 在一些情况下,所述靶序列包括至少3个、至少4个、至少5个、至少6个、至少7个、至少8个、至少9个或至少10个甲基化位点。在一些情况下,所述靶序列包括至少5个甲基化位点。在一些情况下,所述第一甲基化状态包括所述靶序列内单个位点的甲基化密度、甲基化/非甲基化位点在所述靶序列内连续区域上的分布、所述靶序列内每个单个甲基化位点的甲基化模式或水平、或非CpG甲基化。在一些情况下,所述靶序列在所述第一组织中比在所述第二组织中包括更高的甲基化密度。在一些情况下,所述第一甲基化状态包括所述靶序列内单个位点的甲基化密度,并且所述靶序列在所述第一组织中的甲基化密度为至少50%、至少60%、至少70%、至少80%、至少90%或至少95%。在一些情况下,所述靶序列在所述第一组织中包括的甲基化密度大于50%。在一些情况下,所述靶序列在所述第二组织中包括的甲基化密度为至多50%、至多40%、至多30%、至多20%、至多10%、至多5%或0%。在一些情况下,所述靶序列在所述第二组织中包括的甲基化密度小于20%。在一些情况下,所述靶序列在所述第一组织中比在所述第二组织中包括更低的甲基化密度。在一些情况下,所述靶序列在所述第一组织中包括的甲基化密度为至多50%、至多40%、至多30%、至多20%、至多10%、至多5%或0%。在一些情况下,所述靶序列在所述第一组织中包括的甲基化密度小于50%。在一些情况下,所述靶序列在所述第二组织中包括的甲基化密度为至少50%、至少60%、至少70%、至少80%、至少90%、至少95%或100%。在一些情况下,所述靶序列在所述第二组织中包括的甲基化密度大于80%。

[0009] 在一些情况下,所述第一组织包括肝脏组织,并且所述靶序列包括与SEQ ID NO: 1具有至少60%、70%、80%、90%、95%、98%或99%同一性的多核苷酸序列。在一些情况下,所述第一组织包括肝脏组织,并且其中所述测定包括使用包括SEQ ID NO: 2的引物、包括SEQ ID NO: 3的引物或两者进行扩增,或使用包括SEQ ID NO: 4的可检测地标记的探针用于检测所述靶序列。在一些情况下,所述扩增还包括使用包括SEQ ID NO: 5的引物、包括SEQ ID NO: 6的引物或两者,或使用包括SEQ ID NO: 7的可检测地标记的探针用于检测所述靶序列。

[0010] 在一些情况下,所述第一组织包括结肠组织,并且所述靶序列包括与SEQ ID NO: 8具有至少60%、70%、80%、90%、95%、98%或99%同一性的多核苷酸序列。在一些情况下,所述第一组织包括结肠组织,并且其中所述扩增包括使用包括SEQ ID NO: 9的引物、包括SEQ ID NO: 10的引物或两者,或使用包括SEQ ID NO: 11的可检测地标记的探针用于检测所述靶序列。在一些情况下,甲基化特异性扩增还包括使用包括SEQ ID NO: 12的引物、包括SEQ ID NO: 13的引物或两者,或使用包括SEQ ID NO: 14的可检测地标记的探针用于检测所述靶序列。

[0011] 在一些情况下,所述癌症选自:膀胱癌、骨癌、脑肿瘤、乳腺癌、宫颈癌、结直肠癌、食道癌、胃肠道癌、造血系统恶性肿瘤、头颈部鳞状细胞癌、白血病、肝癌、肺癌、淋巴瘤、骨

髓瘤、鼻咽癌、口腔癌、口咽癌、卵巢癌、前列腺癌、肉瘤、胃癌、黑素瘤和甲状腺癌。在一些情况下,所述癌症包括肝细胞癌或结直肠癌。

[0012] 在一些情况下,所述方法还包括确定所述第二组织中所述癌症的分类。在一些情况下,所述确定所述第二组织中癌症的所述分类包括评估来自所述生物体的第二生物样品的游离核酸分子。在一些情况下,所述评估包括确定来自所述第二生物样品的游离核酸分子的甲基化谱、拷贝数变异、单多态性 (SNP) 谱或裂解方式。在一些情况下,所述评估包括确定来自病原体的所述第二生物样品的游离核酸分子的量。在一些情况下,所述第二生物样品与所述第一生物样品相同。在一些情况下,其中所述第二生物样品与所述第一生物样品不同。

[0013] 本公开内容的一个方面提供了一种被配置为执行本文所提供的所述方法的系统。

[0014] 本公开内容的一个方面提供了一种非暂时性计算机可读介质,其包括用于控制计算机系统以执行本文所公开的所述方法的一系列指令。

[0015] 本公开内容的一个方面提供了一种分析生物体的生物样品的方法。所述方法可以包括:基于第一组织特异性标志物的甲基化状态,在来自所述生物样品的游离DNA分子中扩增第一组织特异性标志物,其中所述第一组织特异性标志物包括具有一个或多个差异甲基化位点的预定序列,所述一个或多个差异甲基化位点在所述生物体的第一组织中具有第一甲基化状态,并且在所述生物体的其他组织中具有第二甲基化状态,并且其中所述第一甲基化状态和第二甲基化状态不同;通过检测所述第一组织特异性标志物的所述扩增来识别所述游离DNA分子的起源组织;并且确定所述生物样品中的来自所述生物体的所述第一组织的游离DNA分子的绝对量。

[0016] 在一些情况下,所述方法还包括在所述扩增之前将非甲基化的胞嘧啶残基经亚硫酸氢盐转化为尿嘧啶。在一些情况下,所述方法还包括在所述扩增之前将所述游离DNA分子与所述生物样品中的其他DNA分子分离。在一些情况下,所述扩增包括使用与所述第一组织特异性标志物互补的甲基化特异性引物,并退火至所述一个或多个差异甲基化位点的至少一部分。在一些情况下,所述识别所述起源组织包括如果所述游离DNA分子中的所述第一组织特异性标志物被引物(所述引物被配置为扩增以所述第一甲基化状态被甲基化的第一组织特异性标志物)扩增,则确定所述游离DNA分子来自所述第一组织。

[0017] 在一些情况下,所述一个或多个差异甲基化位点在所述第一组织中比在所述第二组织中具有更高的甲基化密度。在一些情况下,所述一个或多个差异甲基化位点在所述第一组织中的甲基化密度为至少50%、至少60%、至少70%、至少80%、至少90%或至少95%。在一些情况下,所述一个或多个差异甲基化位点在所述第一组织中具有大于50%的甲基化密度。在一些情况下,所述一个或多个差异甲基化位点在所述第二组织中具有的甲基化密度为至多50%、至多40%、至多30%、至多20%、至多10%、至多5%或0%。在一些情况下,所述一个或多个差异甲基化位点在所述第二组织中的甲基化密度小于20%。

[0018] 在一些情况下,所述一个或多个差异甲基化位点在所述第一组织中比在所述第二组织中具有更低的甲基化密度。在一些情况下,所述一个或多个差异甲基化位点在所述第一组织中的甲基化密度为至多50%、至多40%、至多30%、至多20%、至多10%、至多5%或0%。在一些情况下,所述一个或多个差异甲基化位点在所述第一组织中具有小于50%的甲基化密度。在一些情况下,所述一个或多个差异甲基化位点在所述第二组织中的甲基化密度为至少

50%、至少60%、至少70%、至少80%、至少90%、至少95%或100%。在一些情况下,所述一个或多个差异甲基化位点在所述第二组织中具有大于80%的甲基化密度。

[0019] 在一些情况下,所述第一组织包括肝脏组织,并且所述第一组织特异性标志物包括与SEQ ID NO: 1具有至少60%、70%、80%、90%、95%、98%或99%同一性的多核苷酸序列。在一些情况下,所述第一组织包括肝脏组织,并且其中所述扩增包括使用包括SEQ ID NO: 2的引物、包括SEQ ID NO: 3的引物或两者,或使用包括SEQ ID NO: 4的可检测地标记的探针用于检测所述第一组织特异性标志物。在一些情况下,所述扩增还包括使用包括SEQ ID NO: 5的引物、包括SEQ ID NO: 6的引物或两者,或使用包括SEQ ID NO: 7的可检测地标记的探针用于检测所述第一组织特异性标志物。在一些情况下,所述第一组织包括结肠组织,并且所述第一组织特异性标志物包括与SEQ ID NO: 8具有至少60%、70%、80%、90%、95%、98%或99%同一性的多核苷酸序列。在一些情况下,所述第一组织包括结肠组织,并且其中所述扩增包括使用包括SEQ ID NO: 9的引物、包括SEQ ID NO: 10的引物或两者,或使用包括SEQ ID NO: 11的可检测地标记的探针用于检测所述第一组织特异性标志物。在一些情况下,甲基化特异性扩增还包括使用包括SEQ ID NO: 12的引物、包括SEQ ID NO: 13的引物或两者,或使用包括SEQ ID NO: 14的可检测地标记的探针用于检测所述第一组织特异性标志物。

[0020] 在一些情况下,所述方法还包括基于第二组织特异性标志物的甲基化模式,确定所述生物样品中的源自所述第二组织的游离DNA分子的量,其中所述第一组织和第二组织是不同的。在一些情况下,所述第二组织属于所述生物体。

[0021] 在一些情况下,所述方法还包括基于所述生物样品中的来自所述生物体的所述第一组织的所述游离DNA分子的绝对量,诊断、监测或预测所述第一组织中的癌症。在一些情况下,所述癌症选自:膀胱癌、骨癌、脑肿瘤、乳腺癌、宫颈癌、结直肠癌、食道癌、胃肠道癌、造血系统恶性肿瘤、头颈部鳞状细胞癌、白血病、肝癌、肺癌、淋巴瘤、骨髓瘤、鼻咽癌、口腔癌、口咽癌、卵巢癌、前列腺癌、肉瘤、胃癌、黑素瘤和甲状腺癌。在一些情况下,所述癌症包括肝细胞癌或结直肠癌。在一些情况下,所述诊断或监测包括基于所述生物样品中的来自所述生物体的所述第一组织的所述游离DNA分子的绝对量,确定所述第一组织中的肿瘤的大小。在一些情况下,所述诊断或监测包括基于所述生物样品中的来自所述生物体的所述第一组织的所述游离DNA分子的绝对量,确定所述癌症是否转移至所述第一组织。

[0022] 在一些情况下,所述第一组织包括移植的器官。在一些情况下,本文提供的方法还包括基于所述生物样品中的来自所述第一组织的所述游离DNA分子的绝对量,评估器官移植。

[0023] 本公开内容的另一方面提供了一种用于确定生物样品中的来自生物体肝脏的游离DNA分子的量的组合物。所述组合物可以包括一对引物,所述一对引物用于基于所述肝脏特异性标志物的甲基化状态/水平扩增肝脏特异性标志物,其中所述肝脏特异性标志物包括与SEQ ID NO: 1具有至少60%、70%、80%、90%、95%、98%或99%同一性的多核苷酸序列。在一些情况下,所述一对引物包括有包括SEQ ID NO: 2的引物和包括SEQ ID NO: 3的引物。在一些情况下,所述组合物还包括有包括SEQ ID NO: 4的可检测地标记的探针,用于检测所述肝脏特异性标志物。在一些情况下,所述组合物还包括有包括SEQ ID NO: 5的引物和包括SEQ ID NO: 6的引物。在一些情况下,所述组合物还包括有包括SEQ ID NO: 7的可检测

地标记的探针,用于检测所述肝脏特异性标志物。

[0024] 本公开内容的另一方面提供了一种用于确定生物样品中的来自生物体结肠的游离DNA分子的量的组合物,其包括一对引物,该一对引物用于基于所述结肠特异性标志物的甲基化状态/水平扩增所述结肠特异性标志物,其中所述结肠特异性标志物包括与SEQ ID NO: 8具有至少60%、70%、80%、90%、95%、98%或99%同一性的多核苷酸序列。在一些情况下,所述一对引物包括有包括SEQ ID NO: 9的引物和包括SEQ ID NO: 10的引物。在一些情况下,所述组合物还包括有包括SEQ ID NO: 11的可检测地标记的探针,用于检测所述结肠特异性标志物。在一些情况下,所述组合物还包括有包括SEQ ID NO: 12的引物和包括SEQ ID NO: 13的引物。在一些情况下,所述组合物还包括有包括SEQ ID NO: 14的可检测地标记的探针,用于检测所述结肠特异性标志物。

援引并入

[0025] 本说明书中提到的所有出版物、专利和专利申请均通过引用并入本文,其程度如同特别地且单独地指出每一个单独的出版物、专利或专利申请均通过引用而并入。

附图说明

[0026] 在所附权利要求书中具体阐述了本公开的新颖性特征。通过参考对在其中利用到本公开的原理的说明性实施方案加以阐述的以下详细描述和附图,将会获得对本公开的特征和优点的更好的理解,在附图中:

图1是结直肠癌肝转移患者中血浆游离DNA起源的示意图。

[0027] 图2示出了基于肝脏特异性甲基化标志物分析(液滴数字PCR)和通过测序的供者特异性等位基因分析,肝脏移植受者血浆中肝脏源性DNA的分级浓度之间的相关性。

[0028] 图3A和图3B示出了在健康受试者、慢性HBV携带者、肝硬化患者和HCC患者的血浆中使用液滴数字PCR获得的肝脏源性DNA的绝对浓度(图3A)和分级浓度(图3B)。

[0029] 图4A和图4B示出了在HCC患者血浆中使用液滴数字PCR获得的、肿瘤的最大尺寸与肝脏源性DNA的绝对浓度(图4A)和分级浓度(图4B)之间的相关性。

[0030] 图5A-图5D示出了在健康受试者和有或没有肝转移的CRC患者中通过液滴数字PCR获得的结肠源性DNA和肝脏源性DNA的血浆浓度。(图5A)结肠源性DNA的绝对浓度,(图5B)结肠源性DNA的分级浓度,(图5C)肝脏源性DNA的绝对浓度和(图5D)肝脏源性DNA的分级浓度。

[0031] 图6示出了用于使用肝脏源性DNA和结肠源性DNA的绝对浓度和分级浓度来区分有无肝转移的CRC患者的ROC曲线。“AUC”表示曲线下面积。

[0032] 图7示出了蛋白酪氨酸激酶2 β (PTK2B) 基因区域内CpG位点的甲基化密度。

[0033] 图8示出了Sestrin 3 (SESN3) 基因区域内CpG位点的甲基化密度。

[0034] 图9示出了用于使用肝脏源性DNA的绝对浓度和分级浓度来区分HCC患者与非HCC患者的ROC曲线。

[0035] 图10示出了可以被编程或以其他方式配置以实现本文提供的方法的计算机控制系统。

[0036] 图11示出了本文公开的方法和系统的图。

具体实施方式

[0037] 本文提供了用于使用组织特异性标志物定量来自特定组织的游离核酸分子(例如,游离DNA分子,例如血浆DNA)的方法、组合物和系统。本文还提供了这些标志物的临床应用,例如,但不限于,癌症的诊断、监测和预测,转移癌的检测,以及在一些情况下的器官移植评估。

[0038] 除了癌症特异性的变化之外,从受癌症影响的器官释放到循环中的DNA可能普遍增加。例如,在肝癌患者中肝脏源性DNA的水平可以升高。不希望受到某种理论的束缚,受癌症影响的器官释放的DNA水平升高,可能是由于DNA从肿瘤细胞直接释放或被癌症侵袭的非肿瘤细胞更新增加。这种组织特异性DNA释放的增加也可以在由于细胞更新增加而导致经受急性排斥的器官移植接受者中观察到。在一些情况下,通过甲基化反褶积确定的移植器官内细胞对血浆DNA的比例贡献与基于供者特异性等位基因分析所确定的比例贡献有很好的相关。

[0039] 在一些情况下,由于相对较高的成本和较长的运转时间,使用全基因组亚硫酸氢盐测序的甲基化反褶积可能具有挑战性。此外,在一些情况下,甲基化反褶积可以确定不同器官的相对或比例贡献,而不是源于每个器官的DNA的绝对浓度。在将来自多于一个器官的DNA释放到循环中的情况下,测量源自特定器官的DNA的绝对浓度可以提供更多信息。例如,在结直肠癌(CRC)转移到肝脏的患者中,DNA增加的量将从肝脏释放到循环系统中。然而,由于由源自结肠的肿瘤细胞释放的DNA有甚至更大程度的增加,血浆中肝脏DNA的比例可能出现自相矛盾的减少。因此,开发能够准确确定具有组织特异性甲基化模式的DNA的绝对量的方法可以是有用的。

[0040] I. 组织特异性标志物

本文提供了可以识别游离DNA分子的起源组织的组织特异性标志物。在一些情况下,组织特异性标志物是生物体基因组的多核苷酸序列。在一些情况下,组织特异性标志物包括差异甲基化区域(DMR),其基于标志物多核苷酸序列内含有的一个或多个差异甲基化位点的甲基化状态来识别。在一些情况下,一个或多个差异甲基化位点包括一个或多个CpG位点。在一些情况下,一个或多个差异甲基化位点包括一个或多个非CpG位点。在一些情况下,本文讨论的组织特异性标志物可以称为靶序列。

[0041] 在一些情况下,组织特异性标志物的差异甲基化位点在生物体的第一组织中具有第一甲基化状态,而在生物体的不同的第二组织中具有第二甲基化状态。第一甲基化状态和第二甲基化状态可以不同,从而可以基于组织特异性标志物的甲基化状态来区分第一组织和第二组织。

[0042] 在一些情况下,组织特异性标志物的差异甲基化位点在生物体的第一组织中具有第一甲基化状态,而在生物的所有其他组织中具有第二甲基化状态。第一甲基化状态和第二甲基化状态可以不同,从而可以基于组织特异性标志物的甲基化状态将第一组织与生物体的所有其他组织进行区分。

[0043] 在一些情况下,组织特异性标志物包括至少1个、至少2个、至少3个、至少4个、至少5个、至少6个、至少7个、至少8个、至少9个、至少10个、至少12个、至少15个、至少20个、至少25个、至少30个、至少50个差异甲基化位点。在一些情况下,组织特异性标志物包括至少5个差异甲基化位点。甲基化核苷酸或甲基化核苷酸碱基可以指在核苷酸碱基上存在甲基部

分,其中在公认的典型核苷酸碱基中不存在甲基部分。例如,胞嘧啶在其嘧啶环上不含有甲基部分,但5-甲基胞嘧啶在其嘧啶环的5位含有甲基部分。在这种情况下,胞嘧啶不是甲基化的核苷酸,而5-甲基胞嘧啶是甲基化的核苷酸。在另一个示例中,胸腺嘧啶在其嘧啶环的5位含有甲基部分,然而,出于本文目的,当胸腺嘧啶存在于DNA中时,它不被认为是甲基化的核苷酸,因为胸腺嘧啶是DNA的典型核苷酸碱基。DNA的典型核苷碱基为胸腺嘧啶、腺嘌呤、胞嘧啶和鸟嘌呤。RNA的典型碱基是尿嘧啶、腺嘌呤、胞嘧啶和鸟嘌呤。相应地,“甲基化位点”可以是靶基因核酸区域中甲基化发生或可能发生的位置。例如,含有CpG的位置是甲基化位点,其中胞嘧啶可以被甲基化或可以不被甲基化。“位点”可以对应于单个位点,其可以是单个碱基位置或一组相关的碱基位置,例如,CpG位点。甲基化位点可以指有可能被甲基化的DNA分子的CpG位点或非CpG位点。CpG位点可以是DNA分子的如下一区域:在该区域中,胞嘧啶核苷酸后面是沿其5'至3'方向以碱基线性序列的鸟嘌呤核苷酸,并且该区域通过在体内自然发生的事件或通过体外进行的通过化学作用使核苷酸甲基化的事件都易于甲基化。非CpG位点可以是不具有CpG二核苷酸序列,但也易于通过在体内自然发生的事件或在体外进行的通过化学作用使核苷酸甲基化的事件甲基化的区域。基因座(locus)或区域可以对应于包括多个位点的区域。

[0044] 组织特异性标志物的甲基化状态可以包括:标志物区域内各个位点的甲基化密度、标志物内连续区域上甲基化/非甲基化位点的分布、含有多于一个位点的标志物内每个单个甲基化位点的甲基化模式或水平、以及非CpG甲基化。在一些情况下,组织特异性标志物的甲基化状态包括各个差异甲基化位点的甲基化水平(或甲基化密度)。对于给定的甲基化位点,甲基化密度可以指在给定的甲基化位点甲基化的核酸分子在含有该甲基化位点的感兴趣的核酸分子总数中的分数。例如,肝脏组织中第一甲基化位点的甲基化密度可以指在第一位点甲基化的肝脏DNA分子在总肝脏DNA分子中的分数。在一些情况下,甲基化状态包括各个差异甲基化位点之间甲基化/非甲基化状态的连贯性。

[0045] 在一些情况下,组织特异性标志物包括在第一组织中被高甲基化而在第二组织中被低甲基化的甲基化位点。例如,组织特异性标志物可以包括一个或多个在肝脏组织中被高甲基化的甲基化位点,由此其可以表示一个或多个甲基化位点在肝脏组织中具有至少50%、至少60%、至少70%、至少80%、至少90%或至少95%的甲基化密度;相反,一个或多个甲基化位点可以在其他组织(例如但不限于,血细胞、肺、食道、胃、小肠、结肠、胰腺、膀胱、心脏和大脑)中具有至多50%、至多40%、至多30%、至多20%、至多10%、至多5%或0%的甲基化密度。组织特异性标志物可以包括至少1个、至少2个、至少3个、至少4个、至少5个、至少6个、至少7个、至少8个、至少9个、至少10个、至少12个、至少15个、至少20个、至少25个、至少30个、至少30个、至少50个在第一组织中被高甲基化而在第二组织中被低甲基化的甲基化位点。在一些情况下,组织特异性标志物包括至少5个甲基化位点,其在第一组织中被高甲基化,而在第二组织中被低甲基化。

[0046] 组织特异性标志物可以包括至多300个碱基对(bp)、至多250 bp、至多225 bp、至多200 bp、至多190 bp、至多185 bp、至多180 bp、至多175 bp、至多170 bp、至多169 bp、至多168 bp、至多167 bp、至多166 bp、至多165 bp、至多164 bp、至多163 bp、至多162 bp、至多161 bp、至多160 bp、至多150 bp、至多140 bp、至多120 bp或至多100 bp。在一些情况下,组织特异性标志物包括至多166 bp。

[0047] 在一些情况下,本文提供的肝脏特异性标志物包括至少1个、至少2个、至少3个、至少4个、至少5个、至少6个、至少7个、至少8个、至少8个、至少9个、至少10个、至少12个、至少15个、至少20个、至少25个、至少30个、至少50个在肝脏组织中被高甲基化而在其他组织中被低甲基化的甲基化位点。在一些情况下,本文提供的肝脏特异性标志物包括至少5个在肝脏组织中被高甲基化而在其他组织中被低甲基化的甲基化位点。在肝脏组织中被高甲基化的每个甲基化位点在肝脏组织中可以具有至少50%、至少60%、至少70%、至少80%、至少90%或至少95%的甲基化密度,并且在其他组织(例如但不限于,血液、大脑、胸腺、胰腺、肾脏)中的甲基化密度为至多50%、至多40%、至多30%、至多20%、至多10%、至多5%或0%。在一些情况下,在肝脏组织中被高甲基化的每个甲基化位点在肝脏组织中的甲基化密度可以高于50%,而在其他组织(例如但不限于,血细胞、肺、食道、胃、小肠、结肠、胰腺、膀胱、心脏和大脑)中的甲基化密度低于20%。

[0048] 在一些情况下,组织特异性标志物包括在第一组织中被低甲基化而在第二组织中被高甲基化的甲基化位点。在一些情况下,组织特异性标志物包括在第一组织中被低甲基化而在其他组织中被高甲基化的甲基化位点。例如,组织特异性标志物可以包括一个或多个在肝脏组织中被低甲基化的甲基化位点,这可以意味着一个或多个甲基化位点在肝脏组织中具有至多50%、至多40%、至多30%、至多20%、至多10%、至多5%或0%的甲基化密度;相反,一个或多个甲基化位点在其他组织(例如但不限于,血细胞、肺、食道、胃、小肠、结肠、胰腺、膀胱、心脏和大脑)中可以具有至少50%、至少60%、至少70%、至少80%、至少90%或至少95%的甲基化密度。组织特异性标志物可以包括至少1个、至少2个、至少3个、至少4个、至少5个、至少6个、至少7个、至少8个、至少9个、至少10个、至少12个、至少15个、至少20个、至少25个、至少30个、至少30个、至少50个在第一组织中被低甲基化而在第二组织中被高甲基化的甲基化位点。在一些情况下,组织特异性标志物包括至少5个在第一组织中被低甲基化而在第二组织中被高甲基化的甲基化位点。

在一些情况下,本文提供的肝脏特异性标志物包括至少1个、至少2个、至少3个、至少4个、至少5个、至少6个、至少7个、至少8个、至少9个、至少10个、至少12个、至少15个、至少20个、至少25个、至少30个、至少50个在肝脏组织中被低甲基化而在其他组织中被高甲基化的甲基化位点。在一些情况下,本文提供的肝脏特异性标志物包括至少5个在肝脏组织中被低甲基化而在其他组织中被高甲基化的甲基化位点。在肝脏组织中被低甲基化的每个甲基化位点在肝脏组织中可以具有至多50%、至多40%、至多30%、至多20%、至多10%或至多5%的甲基化密度,并且在其他组织(例如但不限于,血液、大脑、胸腺、胰腺、肾脏)中的甲基化密度为至少50%、至少60%、至少70%、至少80%、至少90%、至少95%或100%。在一些情况下,在肝脏组织中被低甲基化的每个甲基化位点在肝脏组织中可以具有低于50%的甲基化密度,而在其他组织(例如但不限于,血细胞、肺、食道、胃、小肠、结肠、胰腺、膀胱、心脏和大脑)中的甲基化密度高于80%。

[0049] 在一些情况下,本文提供的结肠特异性标志物包括至少1个、至少2个、至少3个、至少4个、至少5个、至少6个、至少7个、至少8个、至少9个、至少10个、至少12个、至少15个、至少20个、至少25个、至少30个、至少50个在结肠组织中被高甲基化而在其他组织中被低甲基化的甲基化位点。在一些情况下,本文提供的结肠特异性标志物包括至少5个在结肠组织中被高甲基化而在其他组织中被低甲基化的甲基化位点。在结肠组织中被高甲基化的每个甲基

化位点在结肠组织中可以具有至少50%、至少60%、至少70%、至少80%、至少90%或至少95%的甲基化密度,并且在其他组织(例如但不限于,血液、大脑、胸腺、胰腺、肾脏)中的甲基化密度为至多50%、至多40%、至多30%、至多20%、至多10%、至多5%或0%。在一些情况下,在结肠组织中被高甲基化的每个甲基化位点在结肠组织中可以具有高于50%的甲基化密度,而在其他组织(例如但不限于,血细胞、肺、食道、胃、小肠、肝脏、胰腺、膀胱、心脏和大脑)中的甲基化密度低于20%。

[0050] 在一些情况下,本文提供的结肠特异性标志物包括至少1个、至少2个、至少3个、至少4个、至少5个、至少6个、至少7个、至少8个、至少9个、至少10个、至少12个、至少15个、至少20个、至少25个、至少30个、至少50个在结肠组织中被低甲基化而在其他组织中被高甲基化的甲基化位点。在一些情况下,本文提供的结肠特异性标志物包括至少5个在结肠组织中被低甲基化而在其他组织中被高甲基化的甲基化位点。在结肠组织中被低甲基化的每个甲基化位点在结肠组织中可以具有至多50%、至多40%、至多30%、至多20%、至多10%或至多5%的甲基化密度,并且在其他组织(例如但不限于,血液、大脑、胸腺、胰腺、肾脏)中的甲基化密度为至少50%、至少60%、至少70%、至少80%、至少90%、至少95%或100%。在一些情况下,在结肠组织中被低甲基化的每个甲基化位点在结肠组织中可以具有低于50%的甲基化密度,而在其他组织(例如但不限于,血细胞、肺、食道、胃、小肠、肝脏、胰腺、膀胱、心脏和大脑)中的甲基化密度高于80%。

[0051] 本文还提供了用于识别肝脏源性DNA分子的肝脏特异性标志物。肝脏特异性标志物可以位于8号染色体上的蛋白酪氨酸激酶2 β (PTK2B) 基因的外显子区域中。肝脏特异性DMR中的八个CpG位点可以在肝脏中被高甲基化,而在其他组织和血细胞中被低甲基化。本文提供的肝脏特异性标志物可以包括与SEQ ID NO: 1具有至少约50%、60%、70%、80%、90%、95%、98%或99%同一性的多核苷酸序列。本文提供的肝脏特异性标志物可以包括SEQ ID NO: 1。

[0052] 本文还提供了用于识别结肠源性DNA分子的结肠特异性标志物。结肠特异性标志物可以位于11号染色体上的Sestrin 3 (SESN3) 基因的外显子区域中。位于结肠特异性DMR内的所有六个CpG位点均可以在结肠中被高甲基化,而在其他组织和血细胞中被低甲基化。结肠特异性标志物可以包括与SEQ ID NO: 8具有至少约50%、60%、70%、80%、90%、95%、98%或99%同一性的多核苷酸序列。

[0053] 如本文所用,两个或多个核苷酸或氨基酸序列之间的术语“同一性”或“同一性百分比”可以通过比对序列来确定,以达到最佳比较目的(例如,可以在第一序列的序列中引入缺口)。然后可以比较相应位置处的核苷酸,两个序列之间的同一性百分比可以是所述序列共享的相同位置数的函数(即,同一性%=相同位置数/位置总数 $\times$ 100)。例如,第一序列中的位置可以被与第二序列中对应位置的相同核苷酸占据,那么分子在该位置是相同的。两个序列之间的同一性百分比可以是所述序列共享的相同位置数的函数,考虑到缺口数量和每个缺口的长度,这需要引入以实现两个序列的最佳比对。在一些情况下,为了比较目的而比对的序列的长度至少为约:参考序列长度的30%、40%、50%、60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或95%。BLAST[®] 搜索可以确定两个序列之间的同源性。同源性可以在两个序列的整个长度之间或两个序列的整个长度的部分之间。这两个序列可以是基因、核苷酸序列、蛋白质序列、肽序列、氨基酸序列或其片段。两个序列的实际

比较可以通过众所周知的方法,例如,使用数学算法来完成。这种数学算法的非限制性示例可以是在Karlin, S.和Altschul, S., Proc. Natl. Acad. Sci. USA, 90- 5873-5877 (1993)中所描述的那些数学算法。这种算法可以结合到NBLAST和XBLAST程序(2.0版)中,如在Altschul, S.等人, Nucleic Acids Res., 25:3389-3402 (1997)中所描述的。当使用BLAST和Gapped BLAST程序时,可以使用相应程序(例如,NBLAST)的任何相关参数。例如,序列比较的参数可以设置为得分= 100,字长= 12,或者可以更改(例如,W=5或W=20)。其他示例包括Myers和Miller的算法、CABIOS (1989)、ADVANCE、ADAM、BLAT和FASTA。

[0054] 本文还提供了识别组织特异性标志物的方法。该方法可以包括比较不同组织样品中基因组上的甲基化状态。公开可利用的数据库,如路线图表观基因组学项目 (RoadMap Epigenomics Project) (Roadmap Epigenomics Consortium等人Nature 2015;518:317-30)和蓝图 (BLUEPRINT) 项目 (Martens等人Haematologica 2013;98:1487-9)中的数据库,可以用于生物信息学分析,以筛选潜在的组织特异性标志物。在一些情况下,需要进行实验验证。例如,可以进行甲基化感知测序,如亚硫酸氢盐测序,以验证不同组织之中的甲基化状态。在一些情况下,甲基化特异性扩增也可以用于相对更靶向性的验证。

[0055] II. 分析组织中游离DNA分子的方法

本文提供了分析生物体的生物样品的方法。本文提供的方法可以确定来自生物体组织的游离核酸分子,例如游离DNA分子的绝对量。方法可以包括当游离DNA分子包括具有第一甲基化状态的第一组织特异性标志物时,将来自生物样品的游离DNA分子识别为来自生物体第一组织的游离DNA分子。在一些情况下,第一组织特异性标志物包括具有一个或多个差异甲基化位点的预定序列。在一些情况下,第一组织特异性标志物可以在生物体的第一组织中具有第一甲基化状态,并且在生物体的其他组织中具有第二甲基化状态,并且第一甲基化状态和第二甲基化状态可以不同。

[0056] 如本文所提供的,方法可以包括评估游离DNA分子的甲基化状态。在一些情况下,方法还可以包括在扩增前将非甲基化胞嘧啶残基经亚硫酸氢盐转化成为尿嘧啶。在一些情况下,方法可以包括通过任何其他方法转化甲基化或非甲基化的胞嘧啶残基,使得可以通过随后的检测方法,例如基于引物的扩增,来区分转化的残基。在一些情况下,可以用甲基化敏感酶消化游离的DNA分子,当甲基化位点被甲基化或非甲基化时,甲基化敏感酶在一个或多个特定的甲基化位点消化DNA分子。因此,甲基化敏感酶可以用于区分甲基化和非甲基化的DNA分子。可用于本文提供的方法中的甲基化敏感酶的非限制性示例可以包括Aat II、Acc II、Aor13H I、Aor51H I、BspT104 I、BssH II、Cfr10 I、Cla I、Cpo I、Eco52 I、Hae II、Hap II、Hha I、Mlu I、Nae I、Not I、Nru I、Nsb I、PmaC I、Psp1406 I、Pvu I、Sac II、Sal I、Sma I、SnaB I及其任何组合。

[0057] 本文提供的方法可以包括基于第一组织特异性标志物的甲基化状态,扩增来自生物样品的游离DNA分子中的第一组织特异性标志物。该方法还可以包括通过检测第一组织特异性标志物的扩增来识别游离DNA分子的起源组织。方法还可以包括确定来自生物体的第一组织的生物样品中的游离DNA分子的绝对量。扩增反应可以指一次或多次复制核酸的过程。在一些情况下,扩增方法包括但不限于聚合酶链反应 (PCR)、自我维持序列反应、连接酶链反应、cDNA末端快速扩增、聚合酶链反应和连接酶链反应、Q- β 噬菌体扩增、链置换扩增或剪接重叠延伸聚合酶链反应。在一些情况下,单个核酸分子,例如,通过数字PCR被扩增。

[0058] 在一些情况下,扩增包括使用与第一组织特异性标志物互补的甲基化特异性引物,并退火至一个或多个甲基化位点的至少一部分。甲基化特异性引物可以指可以区分甲基化和非甲基化靶序列的引物。对于含有甲基化位点的给定靶序列,可以将甲基化特异性引物设计成覆盖至少一部分甲基化位点。在一些情况下,当在扩增前进行亚硫酸氢盐转化时,可以将给定甲基化位点的甲基化特异性引物上的核苷酸残基设计为与该位点未转化的胞嘧啶残基互补,以检测甲基化的靶序列,而对于检测非甲基化的靶序列,可以将甲基化特异性引物上的核苷酸残基设计为与该位点的转化残基(例如,尿嘧啶残基)互补。

[0059] 在一些情况下,方法还可以包括在扩增之前将游离DNA分子与生物样品中的其他DNA分子分离。通过物理分离游离DNA分子,可以对生物样品中游离DNA分子进行“数字”测定,例如,但不限于,数字PCR,例如,液滴数字PCR。分离方法可以是技术人员已知的任何方法,例如,但不限于,通过微孔板、毛细管、油乳剂和小型化室的阵列进行分离。可以使用本领域中任何已知的技术进行数字PCR反应,如基于微流体的,或基于乳液的,例如,BEAMing (Dressman等人*Proc Natl Acad Sci USA* 2003; 100: 8817-8822)。

[0060] 在一些情况下,第一组织包括肝脏组织,并且该方法包括使用上述肝脏特异性标志物。在一些情况下,肝脏特异性标志物,例如SEQ ID NO: 1,包括在肝脏中被高甲基化但在其他组织中被低甲基化的位点。在这些情况下,该方法可以包括用于检测甲基化肝脏组织特异性标志物的引物(“用于甲基化测定的引物”)。例如,用于甲基化测定的引物可以包括有包括SEQ ID NO: 2的引物、包括SEQ ID NO: 3的引物或两者。在一些情况下,本文提供的引物用于亚硫酸氢盐转化后的扩增反应。可替代地或累积地,该方法可以包括使用包括SEQ ID NO: 4的可检测地标记的探针,以检测甲基化肝脏组织特异性标志物。任选地,该方法还可以包括使用用于检测非甲基化的肝脏组织特异性标志物的引物(“用于非甲基化测定的引物”)。用于非甲基化测定的引物可以包括有包括SEQ ID NO: 5的引物、包括SEQ ID NO: 6的引物或两者。可替代地或累积地,该方法可以包括使用包括SEQ ID NO: 7的可检测地标记的探针,以检测非甲基化肝脏组织特异性标志物。

[0061] 引物、探针或寡核苷酸在本文中可以互换使用,并且可以指多于一个,例如,通过磷酸二酯键连接在一起的2个、4个、6个、8个、10个、14个、18个、20个或40个核苷酸或化学修饰的核苷酸。引物可以包括约20至约30个核苷酸。引物可以包括至少6个、8个、10个、12个、14个、16个、18个、20个、22个、24个、26个、28个、30个、32个、34个、36个、38个或40个核苷酸。可以引入对引物组成核苷酸的化学修饰,以修饰引物的某些性质,例如,增加其稳定性、增加杂交特异性、以及用可检测的信号标记。可以在本文提供的组合物中使用的化学修饰可以包括共价修饰、附着化学和核苷酸碱基修饰。化学修饰可以包括磷酸化、添加生物素、胆固醇基-TEG、氨基修饰剂(例如,C6、C12或dT)、叠氮化物、炔烃、硫醇修饰剂、荧光团、暗猝灭剂和间隔子。引物可以包括一个或多个硫代磷酸酯键,或一个或多个修饰的碱基,如2-氨基嘌呤、2,6-二氨基嘌呤(2-氨基-dA)、5-溴dU、脱氧尿苷、反向dT、反向双脱氧-T、双脱氧-C、5-甲基dC、脱氧鸟苷、superT®、superG®、锁核酸、5-硝基嘌呤、2'-O-甲基RNA碱基、羟甲基dC、iso-dC、iso-dG、氟碱基(例如,氟C、U、A、G、T)和2'-O-甲氧基-乙基碱基(例如,2-甲氧基乙氧基A、G、U、C、T)。可检测地标记的探针可以包括本领域技术人员已知的任何可检测标记,例如,任何合适的荧光团。PCR反应(例如,数字PCR或实时PCR)可以通过任何合适的光学方法、磁方法、电子方法或本领域技术人员可用的任何其他技术监测。

[0062] 在一些情况下,第一组织包括结肠组织,并且该方法包括使用上述的结肠特异性标志物。在一些情况下,结肠特异性标志物,例如SEQ ID NO: 8,包括在结肠中被高甲基化但在其他组织中被低甲基化的位点。在这些情况下,用于甲基化测定的引物可以包括有包括SEQ ID NO: 9的引物、包括SEQ ID NO: 10的引物或两者。在一些情况下,本文提供的引物用于亚硫酸氢盐转化后的扩增反应。可替代地或累积地,该方法可以包括使用包括SEQ ID NO: 11的可检测地标记的探针,以检测甲基化结肠特异性标志物。任选地,该方法还可以包括使用用于检测非甲基化结肠特异性标志物的引物(“用于非甲基化测定的引物”)。用于非甲基化测定的引物可以包括有包括SEQ ID NO: 12的引物、包括SEQ ID NO: 13的引物或两者。可替代地或累积地,该方法可以包括使用包括SEQ ID NO: 14的可检测地标记的探针,以检测非甲基化结肠特异性标志物。

[0063] 本文提供的方法,例如,使用数字PCR技术,可以使得直接确定靶DNA分子的实际数量,而无需使用校准器。其他技术,如某些基于测序的方法,例如,但不限于,亚硫酸氢盐测序和使用PacBio测序平台的基于非亚硫酸氢盐的甲基化感知测序,可以确定靶组织的DNA相对于其他组织的相对或分级浓度。绝对量可以指DNA分子的绝对计数,或者在一些情况下,也可以指DNA分子的浓度,例如,每体积的数量、摩尔或重量,例如,拷贝数/mL、摩尔/L或mg/L。如本文提供的绝对量分析在将从多于一种类型的组织中释放增加量的DNA的情况下可以是有用的。另一方面,基于游离核酸分子测序的甲基化反褶积分析,如美国专利申请号14/803,692中公开的,可以分数贡献形式提供游离核酸的来源组织的读数,例如,第一组织贡献来自生物样品的A%的游离核酸,并且第二组织贡献来自同一生物样品的B%的游离核酸。

[0064] 在一些情况下,本文提供的方法、组合物和系统还可以利用诸如实时PCR、测序和微阵列等技术对游离核酸进行甲基化分析。在一些情况下,带有组织特异性标志物的游离核酸的绝对数量,如在数字PCR测定中对阳性反应的计数,可能无法通过一些技术直接从甲基化分析中得出。然而,这种绝对数量可以基于携带组织特异性标志物的游离核酸的浓度(相对或分级)间接计算,例如,通过考虑给定体积的生物样品中游离核酸的总数或浓度。在一些情况下,在本文提供的方法中可以使用的测序可以包括链终止测序、杂交测序、Illumina测序(例如,使用可逆终止子染料)、离子洪流半导体测序、质谱分析测序、大规模平行测序(MPSS)、Maxam-Gilbert测序、纳米孔测序、polony测序、焦磷酸测序、鸟枪法测序、单分子实时(SMRT)测序、SOLiD测序(使用四个荧光标记的双碱基探针杂交)、通用测序或其任何组合。具有靶向甲基化位点的探针的微阵列也可以在本文提供的方法中用于分析游离DNA分子的甲基化状态。

[0065] 在一些情况下,本文提供的方法还包括基于第二组织特异性标志物的甲基化模式,确定生物样品中的源自第二组织的游离DNA分子的量,其中第一组织和第二组织是不同的。第二组织可以属于同一生物体。第二组织也可以来自不同的生物体,例如,孕妇体内的胎儿。

[0066] III. 诊断、监测和预测癌症的方法

本文提供了诊断、监测、预测癌症的方法。方法可以包括:如上所述确定来自第一组织的游离DNA分子的绝对量;以及基于生物样品中的来自生物体的第一组织的游离DNA分子的绝对量,诊断、监测和预测第一组织中的癌症。

[0067] 来自第一组织的游离DNA分子的绝对量可以与第一组织的状况相关。例如,由于因肿瘤生长导致的肝脏组织中DNA分子释放的增加,肝脏源性血浆DNA分子的量可增加。在其他情况下,例如,由于器官移植而导致的细胞更新增加也可能导致伴随移植植物从组织释放的血浆DNA增多。

[0068] 本文提供的方法可以包括基于生物样品中的来自生物体的第一组织的游离DNA分子的绝对量,确定第一组织中的肿瘤的大小。可以将靶游离DNA分子的量与肿瘤大小相关联的预定比较图用于肿瘤大小的确定。肿瘤大小的检测可以帮助癌症的诊断、监测和预测。

[0069] 本文提供的方法可以包括基于生物样品中的来自生物体的第一组织的游离DNA分子的绝对量,确定癌症是否转移至第一组织。与靶DNA分子的分级量相比,通过本文提供的方法确定的靶DNA分子的绝对量可以在有转移和无转移的癌症患者之间提供期望的区分。

[0070] 本文提供的方法、组合物和系统可适用的癌症类型可以包括膀胱癌、骨癌、脑肿瘤、乳腺癌、宫颈癌、结直肠癌、食道癌、胃肠道癌、造血系统恶性肿瘤、头颈部鳞状细胞癌、白血病、肝癌、肺癌、淋巴瘤、骨髓瘤、鼻咽癌、口腔癌、口咽癌、卵巢癌、前列腺癌、肉瘤、胃癌或甲状腺癌。通过本文提供的方法评估的转移组织可以包括膀胱、骨骼、大脑、乳腺、子宫颈、结肠、食道、胃肠道、血液、头、颈、肝脏、肺、淋巴结、鼻、鼻咽、口、口咽、卵巢、前列腺、皮肤、胃或甲状腺。

[0071] 如本领域技术人员将容易理解的,癌细胞可以通过移入附近的正常组织而局部扩散,可以局部扩散到附近的淋巴结、组织或器官,并且可以扩散到身体的远处部位。癌症从最初的第一组织到第二组织的扩散可以称为转移,因此这种癌症可以称为转移癌。本文提供的方法、组合物和系统可应用的示例性癌症转移类型可以包括发生在表1所列部位的转移。

[0072] 表1. 示例性的癌症转移部位

癌症类型	转移部位
膀胱	骨骼、肝脏、肺
乳腺	骨骼、大脑、肝脏、肺
结肠	肝脏、肺、腹膜
肾	肾上腺、骨骼、大脑、肝脏、肺
肺	肾上腺、骨骼、大脑、肝脏、其他肺
黑色素瘤	骨骼、大脑、肝脏、肺、皮肤、肌肉
卵巢	肝脏、肺、腹膜
胰腺	肝脏、肺、腹膜
前列腺	肾上腺、骨骼、肝脏、肺
直肠	肝脏、肺、腹膜
胃	肝脏、肺、腹膜
甲状腺	骨骼、肝脏、肺
子宫	骨骼、肝脏、肺、腹膜、阴道

[0073] 在一些情况下,当与本领域技术人员可用的其他技术结合时,本文提供的方法、组

合物和系统可以用于诊断、监测和预测癌症。在一些情况下,其他分子标志物的检测(例如,在核酸,例如,DNA、RNA中),如拷贝数畸变(CNA)、单核苷酸多态性(SNP)、遗传突变、种系突变、体细胞突变、来自病原体(例如,病毒,例如,爱泼斯坦-巴尔病毒)的核酸、游离核酸的大小以及游离核酸的裂解方式,也可以与本文提供的方法、组合物和系统组合应用。技术的组合可以帮助促进癌症水平的检测,包括,但不限于,癌症是否存在、癌症的阶段、肿瘤的大小、涉及多少个染色体区域的缺失或扩增(例如,复制或三倍)和/或其他癌症严重程度的衡量。癌症水平可以是许多其他特征。水平可以为零。癌症水平还可以包括与缺失或扩增相关的恶化前或癌前状况。

[0074] 在一些情况下,拷贝数畸变(CNA)的检测,如美国专利号8,741,811中公开的方法,可以与本文提供的方法结合使用。如上所述,在一些情况下,基于来自第一组织的游离核酸的绝对量,本文提供的方法可以确定癌症是否已经转移到第一组织中。另一方面,CNA的检测可以帮助识别第一组织中转移癌细胞的起源。在一些情况下,游离核酸的裂解方式分析,如美国专利申请号15/218,497中公开的方法,可以与本文提供的方法、组合物和系统结合使用。在一些情况下,主题方法、组合物和系统可以与任何可用方法结合用于检测、监测或预测受试者的癌症。除了上述检测方法外,还可以进行任何适当的测试,例如肿瘤生物标志物测试(例如,针对肝癌的甲胎蛋白(AFP)、针对非小细胞肺癌的ALK基因、针对前列腺癌的前列腺特异性抗原(PSA)和针对甲状腺癌的甲状腺球蛋白)、体检、影像学成像(例如,计算机断层扫描、磁共振成像、正电子发射断层扫描(PET))、超声检查、内镜检查、活检或细胞学测试。

[0075] 在一些情况下,可以将主题方法、组合物或系统用于以定期、半定期或非定期的时间表监测受试者的癌症。例如,受试者可以基于每周、每月、每季度或每年进行使用主题方法、组合物或系统的癌症监测检查。在一些情况下,受试者可以每1个月、2个月、3个月、4个月、5个月、6个月、7个月、8个月、9个月、10个月、11个月或超过12个月进行此类检查。在一些情况下,可以基于最近检查的结果,例如,在一些情况下,根据医生的处方或医疗建议确定两次连续检查之间的间隔。

[0076] IV. 评估器官移植的方法

本文提供了基于确定来自移植组织的游离DNA分子的绝对量来评估器官移植的方法。如本文所述的移植组织被认为是关注的受试者的组织。

[0077] 如本文提供的方法可以利用来自移植组织的游离DNA分子的量与移植组织中的细胞更新率之间的相关性。因此,细胞更新率可以作为评估器官移植的标准。

[0078] V. 用于分析来自组织的游离DNA分子的组合物

本文还提供了用于分析来自特定组织(例如,骨骼、肝脏、肺、大脑、腹膜、肾上腺、皮肤、肌肉、阴道、结肠、膀胱、乳腺、肾脏、黑色素瘤、卵巢、胰腺、前列腺、直肠、胃、甲状腺或子宫)的游离DNA分子的组合物。

[0079] 用于确定生物样品中的来自生物体肝脏的游离DNA分子的量的组合物可以包括一对引物,该一对引物用于基于肝脏特异性标志物的甲基化状态扩增肝脏特异性标志物。肝脏特异性标志物可以包括与SEQ ID NO: 1具有至少60%、70%、80%、90%、95%、98%或99%同一性的多核苷酸序列。在一些情况下,该对引物包括有包括SEQ ID NO: 2的引物和包括SEQ ID NO: 3的引物。在一些情况下,该组合物还包括有包括SEQ ID NO: 4的可检测地标记的

探针,以检测肝脏特异性标志物。在一些情况下,该组合物还包括有包括SEQ ID NO: 5的引物和包括SEQ ID NO: 6的引物。在一些情况下,该组合物还包括有包括SEQ ID NO: 7的可检测地标记的探针,以检测肝脏特异性标志物。

[0080] 本文提供的组合物可以包括一对用于基于结肠特异性标志物的甲基化状态扩增肝脏特异性标志物的引物。结肠特异性标志物可以包括与SEQ ID NO: 8具有至少60%、70%、80%、90%、95%、98%或99%同一性的多核苷酸序列。在一些情况下,该对引物包括有包括SEQ ID NO: 9的引物和包括SEQ ID NO: 10的引物。在一些情况下,该组合物还包括有包括SEQ ID NO: 11的可检测地标记的探针,以检测结肠特异性标志物。在一些情况下,该组合物还包括有包括SEQ ID NO: 12的引物和包括SEQ ID NO: 13的引物。在一些情况下,该组合物还包括有包括SEQ ID NO: 14的可检测地标记的探针,以检测结肠特异性标志物。

[0081] VI. 生物样品

本文提供的方法中使用的生物样品可以包括来自活着的或死去的受试者的任何组织或材料。生物样品可以是游离的样品。生物样品可以包括核酸(例如,DNA,例如,基因组DNA或线粒体DNA或RNA)或其片段。样品中的核酸可以是游离核酸。样品可以是液体样品或固体样品(例如,细胞或组织样品)。生物样品可以是体液,如血液、血浆、血清、尿液、阴道液、鞘膜积液(例如,睾丸鞘膜积液)、阴道冲洗液、胸水、腹水、脑脊液、唾液、汗液、眼泪、痰、支气管肺泡灌洗液、乳头排出液、身体不同部位(例如,甲状腺、乳腺)的抽吸液等。也可以使用粪便样品。在各种实施方案中,已富集游离DNA的生物样品(例如,通过离心方案获得的血浆样品)中的大多数DNA可以是游离的(例如,大于50%、60%、70%、80%、90%、95%或99%的DNA可以是游离的)。生物样品可以被处理,以物理破坏组织或细胞结构(例如,离心和/或细胞裂解),从而将细胞内组分释放到溶液中,该溶液还可以含有酶、缓冲液、盐、洗涤剂等,用于准备样品进行分析。

[0082] 本文提供的方法、组合物和系统可以用于分析生物样品中的核酸分子。核酸分子可以是细胞核酸分子、游离核酸分子,或两者。本文提供的方法所用的游离核酸可以是生物样品中细胞外的核酸分子。游离核酸分子可以存在于各种体液,例如,血液、唾液、精液和尿液中。由于各种组织中的细胞死亡可以产生游离的DNA分子,这可能是由健康状况和/或疾病(例如,肿瘤入侵或生长、器官移植后的免疫排斥)引起的。

[0083] 本文提供的方法中使用的游离核酸分子,例如,游离DNA,可以存在于血浆、尿液、唾液或血清中。游离DNA可以以短片段的形式自然出现。游离DNA片段化可以指当产生或释放游离DNA分子时,高分子量DNA(如细胞核中的DNA)被切割、断裂或消化为短片段的过程。在一些情况下,本文提供的方法、组合物和系统可以用于分析细胞核酸分子,例如,当患者患有白血病、淋巴瘤或骨髓瘤时,来自肿瘤组织的细胞DNA或来自白细胞的细胞DNA。根据本公开内容的一些示例,可以对取自肿瘤组织的样品进行测定和分析。

[0084] VII. 受试者

本文提供的方法、组合物和系统可以用于分析来自受试者,例如,生物体,例如,宿主生物体的样品。受试者可以是任何人类患者,如癌症患者、有患癌症风险的患者或具有家族或个人癌症史的患者。在一些情况下,受试者处于癌症治疗的特定阶段。在一些情况下,受试者可以患有或疑似患有癌症。在一些情况下,受试者是否患有癌症是未知的。

[0085] 受试者可以患有任何类型的癌症或肿瘤。在一个示例中,受试者可以患有结肠癌

或大肠癌。在另一个示例中,受试者可以患有结直肠癌,或结肠癌和直肠癌。在另一个示例中,受试者可以患有肝癌,例如,肝细胞癌。癌症的非限制性示例可以包括,但不限于,肾上腺癌、肛门癌、基底细胞癌、胆管癌、膀胱癌、血液癌、骨癌、脑瘤、乳腺癌、支气管癌、心血管系统癌、宫颈癌、结肠癌、结直肠癌、消化系统癌、内分泌系统癌、子宫内膜癌、食道癌、眼癌、胆囊癌、胃肠道肿瘤、肝细胞癌、肾癌、造血系统恶性肿瘤、喉癌、白血病、肝癌、肺癌、淋巴瘤、黑色素瘤、间皮瘤、肌肉系统癌、骨髓增生异常综合征(MDS)、骨髓瘤、鼻腔癌、鼻咽癌、神经系统癌、淋巴系统癌、口腔癌、口咽癌、骨肉瘤、卵巢癌、胰腺癌、阴茎癌、垂体瘤、前列腺癌、直肠癌、肾盂癌、生殖系统癌、呼吸系统癌、肉瘤、唾液腺癌、骨骼系统癌、皮肤癌、小肠癌、胃癌、睾丸癌、喉癌、胸腺癌、甲状腺癌、肿瘤、泌尿系统癌、子宫癌、阴道癌或外阴癌。淋巴瘤可以是任何类型的淋巴瘤,包括B细胞淋巴瘤(例如,弥漫性大B细胞淋巴瘤、滤泡淋巴瘤、小淋巴细胞淋巴瘤、套细胞淋巴瘤、边缘区B细胞淋巴瘤、伯基特淋巴瘤、淋巴浆细胞性淋巴瘤、毛细胞白血病或原发性中枢神经系统淋巴瘤)或T细胞淋巴瘤(例如,前体T淋巴细胞淋巴瘤或外周T细胞淋巴瘤)。白血病可以是任何类型的白血病,包括急性白血病或慢性白血病。白血病的类型包括急性髓细胞性白血病、慢性髓细胞性白血病、急性淋巴细胞性白血病、急性未分化型白血病或慢性淋巴细胞性白血病。在一些情况下,癌症患者不患有特定类型的癌症。例如,在一些情况下,患者可以患有非乳腺癌的癌症。

[0086] 癌症的示例包括引起实体瘤的癌症以及不引起实体瘤的癌症。此外,本文提及的任何癌症可以是原发癌(例如,以其最初开始生长的身体部位命名的癌症)或继发性或转移癌(例如,源自身体另一部位的癌症)。

[0087] 通过本文所述的任何方法诊断的受试者可以是任何年龄,并且可以是成人、婴儿或儿童。在一些情况下,受试者是0岁、1岁、2岁、3岁、4岁、5岁、6岁、7岁、8岁、9岁、10岁、11岁、12岁、13岁、14岁、15岁、16岁、17岁、18岁、19岁、20岁、21岁、22岁、23岁、24岁、25岁、26岁、27岁、28岁、29岁、30岁、31岁、32岁、33岁、34岁、35岁、36岁、37岁、38岁、39岁、40岁、41岁、42岁、43岁、44岁、45岁、46岁、47岁、48岁、49岁、50岁、51岁、52岁、53岁、54岁、55岁、56岁、57岁、58岁、59岁、60岁、61岁、62岁、63岁、64岁、65岁、66岁、67岁、68岁、69岁、70岁、71岁、72岁、73岁、74岁、75岁、76岁、77岁、78岁、79岁、80岁、81岁、82岁、83岁、84岁、85岁、86岁、87岁、88岁、89岁、90岁、91岁、92岁、93岁、94岁、95岁、96岁、97岁、98岁或99岁,或在其范围内(例如,2岁至20岁,20岁至40岁或40岁至90岁)。可以受益的特定类别的患者可以是40岁以上的患者。可以受益的另一特定类型的患者可以是儿科患者。此外,通过本文所述的任何方法或组合物诊断的受试者可以是男性或女性。

[0088] 本文公开的任何方法也可以在非人类受试者,如实验室或农场动物,或来自本文公开的生物体的细胞样品上进行。非人类受试者的非限制性示例包括狗、山羊、豚鼠、仓鼠、小鼠、猪、非人类灵长类动物(例如,大猩猩、猿、猩猩、狐猴、或狒狒)、大鼠、绵羊、牛或斑马鱼。

[0089] 如上所述,主题方法、组合物和试剂盒可以用于癌症治疗各个阶段的受试者。使用主题方法、组合物和试剂盒对受试者的生物样品中的游离核酸进行分析的结果可以用于指导受试者的治疗计划。在一些情况下,可能需要用于治疗或治愈受试者癌症的药物或疗法。示例性的治疗选项可以包括化疗、放疗、手术切除肿瘤组织、免疫治疗、靶向治疗、激素治疗和干细胞治疗。在一些情况下,可以提供关于选择不同类型的治疗选项的指导。在一些非限

制性示例中,患者可以已经完成第一癌症的治疗,例如,手术切除受患病的肝叶中的肿瘤组织,并且可以使用主题方法、组合物或试剂盒对该患者进行常规监测测试,以检查是否有肝癌复发或转移。在这些情况下,测试结果可以用于指导患者是否将需要癌症的进一步治疗,并且如果发生肝癌复发或到其他组织的转移,可以采用哪些治疗选项。在一些情况下,可以提供有关治疗的具体剂量或给药方案的指导。例如,来自某个组织的游离核酸的量可以与要施用于患者的药物剂量或药物施用的频率/间隔(例如,每天、每周、每双周或每月)关联。在一些情况下,上次分析的结果可以用作评价和设计治疗方案以及后续监测分析的基础。

[0090] VIII. 计算机系统

本文公开的任何方法可以由一个或多个计算机系统执行和/或控制。在一些示例中,本文公开的方法的任何步骤可以由一个或多个计算机系统整体地、单独地或顺序地执行和/或控制。本文提到的任何计算机系统都可以利用任何合适数量的子系统。在一些实施方案中,计算机系统包括单个计算机设备,其中子系统可以是计算机设备的组件。在其他实施方案中,计算机系统可以包括多个计算机设备,每个计算机设备是具有内部组件的子系统。计算机系统可以包括台式计算机和膝上型计算机、平板电脑、移动电话和其他移动设备。

[0091] 子系统可以通过系统总线互连。附加子系统包括打印机、键盘、存储设备和耦合到显示器适配器的监视器。耦合到I/O控制器的外围设备和输入/输出(I/O)设备可以通过本领域已知的任何数量的连接部例如输入/输出(I/O)端口(例如,USB、FireWire®)连接到计算机系统。例如,可以使用I/O端口或外部接口(例如,以太网、Wi-Fi等)将计算机系统连接到广域网,如Internet、鼠标输入设备或扫描仪。通过系统总线的互连允许中央处理器与每个子系统进行通信,并且控制来自系统存储器或存储设备(例如,诸如硬盘或光盘等固定盘)的多个指令的执行,以及子系统之间的信息交换。系统存储器和/或存储设备可以体现计算机可读介质。另一个子系统是数据收集设备,如照相机、麦克风、加速计等。本文提到的任何数据都可以从一个组件输出到另一个组件,并可以输出给用户。

[0092] 计算机系统可以包括多个例如通过外部接口或内部接口连接在一起的相同的组件或子系统。在一些实施方案中,计算机系统、子系统或设备可以通过网络进行通信。在这种情况下,一台计算机可以被视为客户端,另一台计算机可以被视为服务器,其中每台计算机都可以被视为同一计算机系统的一部分。客户端和服务器可以各自包括多个系统、子系统或组件。

[0093] 本公开内容提供了被编程为实现本公开内容的方法的计算机控制系统。图10示出了计算机系统101,该计算机系统101被编程或以其他方式配置为如本文所述地确定来自生物体组织的游离核酸分子的绝对量。计算机系统101可以实现和/或调节本公开内容中提供的方法的各个方面,例如,控制来自生物样品的核酸分子的测序、进行测序数据的生物信息学分析的各个步骤(如本文所述)、集成数据收集、分析和结果报告、以及数据管理。计算机系统101可以是用户的电子设备或相对于电子设备远程放置的计算机系统。该电子设备可以是移动电子设备。

[0094] 计算机系统101包括中央处理单元(CPU,在本文中也称为“处理器”和“计算机处理器”)105,其可以是单核或多核处理器,或用于并行处理的多个处理器。计算机系统101还包括存储器或存储器单元110(例如,随机存取存储器、只读存储器、闪存)、电子存储单元115

(例如,硬盘)、用于与一个或多个其他系统通信的通信接口120(例如,网络适配器)、以及外围设备125(如缓存、其他存储器、数据存储器和/或电子显示器适配器)。存储器110、存储单元115、接口120和外围设备125通过诸如母板的通信总线(实线)与CPU 105通信。存储单元115可以是用于存储数据的数据存储单元(或数据存储库)。计算机系统101可以借助于通信接口120可操作地耦合到计算机网络(“网络”)130。网络130可以是Internet、互联网和/或外部网、或与Internet通信的内部网和/或外部网。在一些情况下,网络130是电信和/或数据网络。网络130可以包括一个或多个计算机服务器,其可以实现分布式计算,如云计算。在一些情况下,网络130可以在计算机系统101的帮助下实现对等网络,该对等网络可以使耦合到计算机系统101的设备充当客户端或服务器。

[0095] CPU 105可以执行一系列机器可读指令,其可以体现在程序或软件中。指令可以被存储在存储器单元中,如存储器110。指令可以被导送到CPU 105,CPU 105可以随后对CPU 105进行编程或以其他方式配置,以实现本公开内容的方法。由CPU 105执行的操作的示例可以包括获取、解码、执行和回写。

[0096] CPU 105可以是电路,如集成电路的一部分。系统101的一个或多个其他组件可以包括在电路中。在一些情况下,该电路是专用集成电路(ASIC)。

[0097] 存储单元115可以存储文件,如驱动程序、库和保存的程序。存储单元115可以存储用户数据,例如用户偏好和用户程序。在一些情况下,计算机系统101可以包括计算机系统101外部的一个或多个附加数据存储单元,如位于通过内部网或Internet与计算机系统101通信的远程服务器上。

[0098] 计算机系统101可以通过网络130与一个或多个远程计算机系统通信。例如,计算机系统101可以与用户的远程计算机系统(例如,安装有应用程序的智能电话,该应用程序接收并显示从计算机系统101发送的样品分析结果)通信。远程计算机系统的示例包括个人计算机(例如,便携式PC)、平板电脑或平板PC(例如,Apple® iPad、Samsung® Galaxy Tab)、电话、智能电话(例如,Apple® iPhone、Android使能性设备、Blackberry®)或个人数字助理。用户可以通过网络130访问计算机系统101。

[0099] 本文所述的方法可以通过存储在计算机系统101的电子存储单元上,例如,在存储器110或电子存储单元115上的机器(例如,计算机处理器)可执行代码的方式来实现。机器可执行或机器可读代码可以以软件的形式提供。在使用期间,代码可以由处理器105执行。在一些情况下,代码可以从存储单元115检索并被存储在存储器110上,以备处理器105访问。在一些情况下,可以排除电子存储单元115,并且将机器可执行指令存储在存储器110上。

[0100] 代码可以被预编译并配置为用于具有适于执行该代码的处理器器的机器,或者可以在运行时期间被编译。可以用可以选择的编程语言提供代码,以使代码能够以预编译或实时编译的方式执行。

[0101] 本文提供的系统和方法的方面,如计算机系统101,可以体现在编程中。技术的各个方面可以被认为是在机器可读介质上或以机器可读介质的类型体现的通常以机器(或处理器)可执行代码和/或关联数据形式的“产品”或“制品”。机器可执行代码可以存储在电子存储单元,如存储器(例如,只读存储器、随机存取存储器、闪存)或硬盘上。“存储”类型的介质可以包括计算机、处理器等的任何或所有有形存储器,或其相关模块,如各种半

导体存储器、磁带驱动器、磁盘驱动器等,它们可以随时为软件编程提供非暂时性存储。软件的全部或部分有时可以通过Internet或其他各种电信网络进行通信。例如,此类通信可以使得软件能够从一台计算机或处理器加载到另一台计算机或处理器,例如,从管理服务器或主机加载到应用服务器的计算机平台中。因此,可以承载软件元素的另一种介质包括光波、电波和电磁波,其例如通过有线和光纤陆线网络以及经各种空中链路在本地设备之间的物理接口上使用。载送此类波的物理元件,如有线或无线链路、光学链路等也可以被认为是承载软件的介质。如本文所用,除非限于非暂时性的有形“存储”介质,否则诸如计算机或机器“可读介质”等术语是指参与向处理器提供指令以供执行的任何介质。

[0102] 因此,机器可读介质,如计算机可执行代码,可以采取许多形式,包括但不限于,有形存储介质、载波介质或物理传输介质。非易失性存储介质包括,例如,光盘或磁盘,如任何计算机中的任何存储设备等,如可以用于实现附图中所示的数据库等。易失性存储介质包括动态存储器,如这种计算机平台的主存储器。有形传输介质包括同轴电缆;铜线和光纤,包括有包括计算机系统内总线的电线。载波传输介质可以采用电信号或电磁信号或声波或光波的形式,如在射频(RF)和红外(IR)数据通信期间所生成的那些。因此,计算机可读介质的常见形式包括例如:软磁盘、软盘、硬盘、磁带、任何其他磁介质、CD-ROM、DVD或DVD-ROM、任何其他光学介质、穿孔卡纸带、带孔模式的任何其他物理存储介质、RAM、ROM、PROM和EPROM、FLASH-EPROM、任何其他存储芯片或盒带、传输数据或指令的载波、传输此类载波的电缆或链路或计算机可以从中读取编程代码和/或数据的任何其他介质。许多这些形式的计算机可读介质可以涉及将一个或多个指令的一个或多个序列传送到处理器以供执行。

[0103] 计算机系统101可以包括电子显示器135或与电子显示器135通信,该电子显示器135包括用户界面(UI) 140,以提供,例如,样品分析的结果,例如但不限于图形显示来自不同组织的游离核酸的相对和/或绝对量、来自某些组织的游离核酸的对照或参考量、检测量与参考量之间的比较,以及读出是否存在癌症转移。UI的示例包括,但不限于,图形用户界面(GUI)和基于web的用户界面。

[0104] 本公开内容的方法和系统可以通过一种或多种算法来实现。在由中央处理单元105执行时,可以通过软件来实现算法。例如,算法可以控制来自样品的核酸分子的测序、直接收集测序数据、分析测序数据或基于测序数据的分析确定病理学分类。

[0105] 在一些情况下,如图11所示,可以从诸如人类受试者的受试者201获得样品202。可以对样品202进行如本文所述的一种或多种方法,如进行测定。在一些情况下,测定可以包括杂交、扩增、测序、标记、表观遗传修饰碱基或其任何组合。方法的一个或多个结果可以输入到处理器204中。一个或多个输入参数,如样品识别、受试者识别、样品类型、参考或其他信息可以输入到处理器204中。测定的一个或多个度量可以输入到处理器204中,使得处理器可以产生结果,如病理学分类(例如,诊断)或治疗建议。处理器可以将结果、输入参数、度量、参考或其任何组合发送到显示器205,如视觉显示器或图形用户界面。处理器204可以(i)将结果、输入参数、度量或其任何组合发送到服务器207,(ii)从服务器207接收结果、输入参数、度量或其任何组合,(iii)或其组合。

[0106] 本公开内容的各方面可以使用硬件(例如,专用集成电路或现场可编程门阵列)和/或使用具有以模块化或集成方式的一般可编程处理器的计算机软件以控制逻辑的形式实现。如本文所用,处理器包括单核处理器、在同一集成芯片上的多核处理器、或在单个电

路板上或联网的多个处理单元。基于本文提供的公开内容和教导,本领域普通技术人员将知道并理解使用硬件以及硬件和软件的组合来实现本文所述的实施方案的其他方式和/或方法。

[0107] 在本申请中所述的任何软件组件或功能都可以被实现为将由处理器使用任何合适的计算机语言(例如,Java、C、C++、C#、Objective-C、Swift)或脚本语言(例如,Perl或Python),使用,例如,传统或面向对象技术执行的软件代码。软件代码可以作为一系列指令或命令存储在计算机可读介质上,以进行存储和/或传输。合适的非暂时性计算机可读介质可以包括随机存取存储器(RAM)、只读存储器(ROM)、磁介质如硬盘驱动器或软盘,或光学介质如光盘(CD)或DVD(数字多功能磁盘)、闪存等。计算机可读介质可以是此类存储或传输设备的任何组合。

[0108] 此类程序还可以使用载波信号进行编码和传输,该载波信号适合于通过符合包括Internet的各种协议的有线网络、光学网络和/或无线网络进行传输。这样,可以使用由此类程序编码的数据信号创建计算机可读介质。用程序代码编码的计算机可读介质可以与兼容设备打包,或者与其他设备分开提供(例如,通过Internet下载)。任何此类计算机可读介质可以驻留在单个计算机产品(例如,硬盘驱动器、CD或整个计算机系统)上或内部,并且可以存在于系统或网络内的不同计算机产品上或内部。计算机系统可以包括监视器、打印机或用于向用户提供本文提及的任何结果其他合适的显示器。

[0109] 本文所述的任何方法可以用包括一个或多个处理器的计算机系统完全或部分执行,该计算机系统可以被配置为执行步骤。因此,实施方案可以针对被配置为执行本文所述的任何方法的步骤的计算机系统,其中不同的组件执行相应的步骤或相应的步骤组。尽管以编号步骤呈现,但本文的方法的步骤可以同时或以不同顺序执行。此外,这些步骤的一部分可以与其他方法的其他步骤的一部分一起使用。此外,步骤的全部或部分可以是可选的。此外,任何方法的任何步骤都可以使用用于执行这些步骤的模块、单元、电路或其他方法来执行。

[0110] 实施例

以下实例进一步说明了所述的实施方案,而不限制本公开内容的范围。

[0111] 实施例1.材料和方法

该实施例描述了用于实施例2-5的多种方法。

[0112] 受试者

表2列出了被招募的受试者的人口统计数据。所有被招募的受试者都签署了书面同意书。

[0113] 表2.研究中分析的受试者人口统计数据

健康和疾病状况	数量	性别 (男性:女性)	年龄 中值 (范围)
健康受试者	30	16 : 14	57 (26 - 73)
肝移植受者	13	12 : 1	67 (54 - 73)
HCC 患者	40	36 : 4	60 (39 - 82)
HBV 携带者	20	20 : 0	52 (43 - 77)
肝硬化患者	9	9 : 0	62 (53 - 75)
CRC 肝转移患者	27	20 : 7	62 (34 - 78)
无肝转移的 CRC 患者	35	18 : 17	67 (51 - 83)

[0114] 样品制备

对于每个受试者,将10 mL外周血收集到含有EDTA的管中。采血后6小时内处理血液样品,以分离血浆和血沉棕黄层。按照制造商的协议,使用QIAamp DSP DNA Mini Kit (Qiagen)从血浆中提取DNA。使用Epitect Plus Bisulfite Kit (Qiagen),将从2 mL至4 mL血浆中提取的DNA进行两轮亚硫酸氢盐处理。将亚硫酸氢盐转化的DNA在50 μ L水中洗脱,用于下游分析。

[0115] 肝脏特异性甲基化标志物和结肠特异性甲基化标志物的识别

将感兴趣的组织(即肝脏或结肠)的甲基化谱与其他血细胞和组织的甲基化谱进行比较,以挖掘组织特异性甲基化标志物。从肺、食道、小肠、结肠、胰腺、膀胱、心脏和肝脏的路线图表观基因组项目的数据库以及成红细胞、中性粒细胞、B淋巴细胞和T淋巴细胞的蓝图项目的数据库中检索了不同细胞类型的甲基化谱,。

[0116] 建立了以下甲基化标志物的标准。

[0117] 1. 如果CpG位点的甲基化密度在靶组织(即肝脏或结肠)中>50%且在其他血细胞和组织中<20%,则将CpG位点定义为在靶组织中高甲基化。

[0118] 2. 至少5个高甲基化的CpG位点的片段将在差异甲基化区域(DMR)内,以提高后者的信噪比和甲基化标志物的分析特异性。

[0119] 3. DMR应该短于166 bp,因为大多数循环DNA分子是短片段,并且峰大小为166 bp。

[0120] 肝脏和结肠的DMR

使用上述标准,识别了一种肝脏特异性标志物和一种结肠特异性标志物。肝脏特异性DMR位于8号染色体上的蛋白酪氨酸激酶2 β (PTK2B)基因的外显子区域。肝脏特异性DMR内的八个CpG位点在肝脏中被高甲基化,而在其他组织和血细胞中被低甲基化(图7)。PTK2B基因位于8号染色体上,并且CpG位点的基因组坐标显示在图7的X轴上。与其他组织相比,位于DMR(两条垂直虚线之间的区域)内的所有八个CpG位点对于肝脏都被高甲基化。黄色高亮区域在数字PCR测定的荧光探针内含有三个CpG位点(参见表3)。DMR内高亮区域每一侧的其他CpG位点均被数字PCR测定的引物覆盖。结肠特异性DMR在11号染色体上的Sestrin 3

(SESN3) 基因的外显子区域内。结肠特异性DMR内的所有六个CpG位点在结肠组织中均被高甲基化(图8),而在其他组织中均被低甲基化。SESN3位于11号染色体上,CpG位点的基因组坐标显示在图8的X轴上。与其他组织相比,位于DMR(两条竖直虚线之间的区域)内的所有六个CpG位点对于结肠都被高甲基化。黄色高亮区域内的三个CpG位点被结肠特异性甲基化测定的荧光探针覆盖(参见表3)。DMR内高亮区域每一侧的其他CpG位点均被数字PCR测定的引物覆盖。为了便于说明,在图7和图8未显示对肝脏、膀胱、食道、心脏、肺、胰腺和小肠的单独结果8,其平均值用“其他组织”表示。

[0121] 液滴数字PCR的设计和性能

开发了两种液滴数字PCR测定以通过肝脏特异性甲基化标志物和结肠特异性甲基化标志物中每一种定量甲基化DNA分子和非甲基化DNA分子。用于测定的引物和探针的序列在表3中列出(引物和探针中带下划线的核苷酸是在CpG位点的差异甲基化胞嘧啶)。两种液滴数字PCR测定可以分别使用标记为FAM和VIC的探针定量甲基化(来自靶组织)和非甲基化(来自非靶组织)。肝脏特异性标志物为PTK2B基因标志物位点(chr8: 27,183,116-27,183,176),并且结肠特异性标志物为SESN3基因标志物位点(chr11: 94,965,508-94,965,567)。

[0122] 对于每个样品,进行重复的数字PCR测定。制备总体积为20 μL 的反应混合物,含有8 μL 亚硫酸氢盐转化的DNA、各自的最终浓度为450 nM的正向引物和反向引物、最终浓度为250 nM的非甲基化特异性探针、和最终浓度为350 nM(肝脏测定)或250 nM(结肠测定)的用于结肠测定的甲基化特异性探针。在使用BioRad QX200 ddPCR液滴生成器进行PCR反应前,将反应混合物进行液滴生成。在每个板上运行通用甲基化DNA(来自EMD Millipore的CpGenome人甲基化DNA)和通用非甲基化DNA(来自Qiagen的EpiTect非甲基化人对照DNA)作为阳性对照和阴性对照。温度分布为:95°C持续10分钟;然后进行94°C的45个循环,持续15秒;和60°C(肝脏测定)或56°C(结肠测定)持续1分钟;以及最后在98°C温育10分钟。PCR完成后,通过QX200液滴读取器分析每个样品的液滴,并使用QuantaSoft (1.7版)软件解释结果。阳性荧光信号的截止值参照对照确定。使用来自重复孔的合并计数然后进行泊松校正,来计算每个样品中甲基化DNA序列和非甲基化DNA序列的数量。血浆中甲基化DNA序列或非甲基化DNA序列的浓度计算如下:

$$Cr = -\ln\left(1 - \frac{P}{R}\right) \times \left(\frac{Ve}{Vd}\right)$$

[0123] 其中Cr表示血浆中靶分子的浓度(即甲基化DNA序列或非甲基化的DNA序列),P表示含有靶分子(甲基化DNA序列或非甲基化DNA序列)放大信号的液滴数,R表示分析的总液滴数(带有和不带有放大信号), V_d 表示液滴的平均体积(即当前实施例中为 $0.9 \times 10^{-3} \mu\text{L}$), V_e 表示用于实验的血浆体积(即当前实施例中为320 μL)。

[0124] 表3.用于数字PCR测定的寡核苷酸序列

引物/探针/标志物	序列(5' -3')	SEQ ID NO
<i>PTK2B</i> 基因标志物位点	CGGCCGACTTACCTGTACTTGCCGCGTCCCGGCTCACCTGGCGGTGCCCCAG GAGTAGTC	1
甲基化测定（用于肝脏源性DNA）		
正向引物(F1)	TTTATTGTTCGGTTCGATTATTTGTA	2
反向引物(R1)	AACGACTACTCCTCGAACACCG	3
荧光探针(P1)	5' -FAM-TTGTCTCGTTTCGGTT-MGB-3'	15
荧光探针(P1)序列	TTGTCTCGTTTCGGTT	4
非甲基化测定（用于非肝脏源性DNA）		
正向引物(F2)	TGTATATTTATTTGTTTGGTTGATTATTTGTA	5
反向引物(R2)	CCAACAACACTACTCCTCAAACACCA	6
荧光探针(P2)	5' -VIC-TTGTGTGTGTTTGGTTTA-MGB-3'	16
荧光探针(P2)序列	TTGTGTGTGTTTGGTTTA	7
<i>SES3</i> 基因标志物位点	CGCAACCTGCTCCCGAGTGAGAACAAGGCCGGCGCAGCGGAGAACCTGATT GGTGCCC	8
甲基化测定（用于结肠源性DNA）		
正向引物(F1)	CGTAATTTGTTTTCGAGTGAGAATAA	9
反向引物(R1)	AAAAACCGAACACCAATCAAATTC	10
荧光探针(P1)	5' -FAM-TCGGCGTAGAGCGA-MGB-3'	17
荧光探针(P1)序列	TCGGCGTAGAGCGA	11
非甲基化测定（用于非结肠源性DNA）		
正向引物(F2)	GTTTGTAAATTTGTTTTCGAGTGAGAATAA	12
反向引物(R2)	AAAAACCAACACCAATCAAATTC	13
荧光探针(P2)	5' -VIC-AGGTGGTGTAGAGTGA-MGB-3'	18
荧光探针(P2)序列	AGGTGGTGTAGAGTGA	14

[0125] 分析来自不同类型样品的DNA

从香港威尔斯亲王医院解剖和细胞病理学科检索10种组织(例如,肝脏、肺、食道、胃、小肠、结肠、胰腺、膀胱、心脏和大脑)的经福尔马林固定的石蜡包埋(FFPE)样品。经组织学检查,这些组织被证实为正常。从健康受试者收集血沉棕黄层样品。使用QIAamp DNA Mini Kit(Qiagen)从FFPE组织中提取DNA。使用QIAamp DNA Blood Mini Kit(Qiagen)提取血沉棕黄层的DNA。使用1 μg细胞DNA进行亚硫酸氢盐转化。将经转化的DNA在20 μL水中洗脱,然后稀释50倍以进行下游分析。

[0126] 肝脏移植受者血浆中供者源性DNA的测量

使用Illumina iScan系统分析从供者的肝脏组织和受者的血沉棕黄层中提取的DNA,以确定供者和受者的基因型信息。从每个受者的4 mL血浆中提取的DNA用于测序文库制备。按照制造商的说明,用KAPA Library Preparation Kit(KAPA Biosystems)制备血浆

DNA测序文库。然后使用Illumina HiSeq 2500平台 (75×2个循环) 对索引的文库进行多路复用和测序。每个样品获得至少2000万对配对末端读数。使用短寡核苷酸比对程序2 (SOAP2) 将配对末端的读数与未屏蔽重复的人参考基因组 (GRCh 37/hg 19) 进行比对。仅包括配对末端的读数,其两端均以正确的方向与同一染色体比对并与人基因组中的单个位置比对。检索插入大小≤600 bp的配对末端读数,进行分析。如果多于一对的读数映射到相同的基因组位置 (即重复的读数), 则仅保留一对读数用于后续分析。配对末端读数的任何一个成员允许最多两个核苷酸错配。通过对具有单核苷酸多态性 (SNP) 等位基因的测序读数进行计数来确定循环中供者特异性DNA的分级浓度,这些单核苷酸多态性 (SNP) 等位基因在受者中是纯合的,并且在供者中是杂合的。

[0127] 实施例2. 肝脏标志物和结肠标志物的组织特异性

对于肝脏特异性标志物和结肠特异性标志物,源自靶组织的DNA分子将被高甲基化,而来自非靶组织的DNA分子将被低甲基化。因此,在肝脏测定中,总分子的百分比表示为甲基化的为L%,并且在结肠测定中,总分子的百分比表示为甲基化的为C%。为了确认肝脏标志物和结肠标志物的特异性,使用这两种数字PCR测定组分析从血沉棕黄层样品和10种正常组织中提取的DNA。对于每种类型的组织,包括来自不同个体的4个样品。

[0128] 肝脏组织的平均L%为67% (范围:57%-76%), 并且其他组织类型的平均L%为0.6% (范围:0.0%-2.2%)。表4总结了每种组织类型的结果。这些结果表明,肝脏测定能够特异性检测肝脏源性DNA。

[0129] 表4. 肝脏源性DNA (L%) 和结肠源性DNA (C%) 的平均分级浓度

组织	平均 L%	平均 C%
肝脏	67	0.8
结肠	0.4	22
大脑	1.9	0.8
食管	0.3	1.4
心脏	0.6	0.2
肺	0.3	1.4
胰腺	0.3	0.9
小肠	0.3	4.1
胃	0.1	1.3
膀胱	2.2	1.3
血沉棕黄层	0.0	0.1

[0130] 结肠组织的平均C%为22% (范围:17%-33%)。所有其他组织的平均C%为1.2% (范围:0.1%-4.1%), 表明甲基化序列的特异性是结肠源性的。结肠组织中相对较低的C%可能是由于结肠组织的细胞组分不均一所致。当将相同的测定用于比较不同疾病状态的受试者的水平时,结肠组织中相对较低的C%不会显著妨碍其临床应用。

[0131] 实施例3.肝脏移植受者中肝脏源性DNA的血浆浓度

通过肝脏移植受者血浆的分析验证了肝脏特异性测定的定量准确性。在这些受试者中,可以使用下一代测序根据携带供者特异性等位基因的血浆DNA分子的比例准确确定来自移植肝脏的DNA的分级浓度。通过肝脏特异性甲基化标志物和测序分析从13位接受肝脏移植的患者收集的14份血浆样品。在这两种方法测定的浓度之间观察到呈线性正相关($R = 0.99, P < 0.0001$, Pearson相关, 图2), 表明肝脏特异性甲基化标志物可以准确反映血浆中肝脏源性DNA的浓度。

[0132] 如本文证明的,通过肝脏特异性甲基化标志物测量的肝脏DNA浓度百分比贡献与基于供者特异性等位基因的测量的结果良好地相关。这些结果证实了肝脏特异性标志物在反映血浆中肝脏源性DNA浓度方面的准确性。

[0133] 实施例4.HCC患者血浆中肝脏源性DNA增加

通过靶向40位HCC患者、9位肝硬化患者、20位慢性HBV携带者和30位健康受试者的肝脏特异性甲基化模式的序列的数字PCR,确定了肝脏源性DNA的绝对浓度和分级浓度。

[0134] 健康受试者、慢性HBV携带者、肝硬化患者和HCC患者的肝脏源性甲基化序列的中值浓度分别为40拷贝数/mL(四分位间距(IQR): 18-86)、122拷贝数/mL(IQR: 47-185)、118拷贝数/mL(IQR: 86-159)和487(IQR: 138-1151)(图3A)。四组的浓度差异显著($P < 0.001$, Kruskal Wallis试验)。在事后分析中,HCC患者肝脏源性DNA的血浆浓度显著高于健康受试者($P < 0.001$, Dunn's试验)和慢性HBV携带者($P = 0.015$, Dunn's试验),但不高于肝硬化患者组($P = 0.248$, Dunn's试验)。健康受试者、慢性HBV携带者和肝硬化患者的浓度无统计学差异($P > 0.05$, Dunn's试验)。

[0135] 对于健康受试者、慢性HBV携带者、肝硬化患者和HCC患者,血浆中肝脏源性DNA的中值分级浓度分别为1.4%(IQR: 0.94%-3.2%)、4.6%(IQR: 1.7%-6.0%)、3.0%(四分位间距: 1.8%-7.3%)和9.4%(IQR: 4.1%-16.0%)(图3B)。四组中的分级浓度差异显著($P < 0.001$, Kruskal Wallis试验)。在事后分析中,HCC患者肝脏源性DNA的血浆浓度显著高于健康受试者($P < 0.001$, Dunn's试验),而不是慢性HBV携带者($P = 0.129$, Dunn's试验)和肝硬化患者($P = 0.592$, Dunn's试验)。健康受试者、慢性HBV携带者和肝硬化患者的浓度无统计学差异。

[0136] 这些结果表明,血浆中肝脏源性DNA的绝对浓度和分级浓度的分析均可以将HCC患者与非HCC受试者(包括健康受试者、慢性HBV携带者和肝硬化患者)区分。为了进一步确定绝对浓度或分级浓度对区分HCC受试者和非HCC受试者是否更好,进行了受者工作特征(ROC)曲线分析(图9)。绝对浓度和分级浓度的曲线下面积(AUC)分别为0.82和0.78。AUC的差异具有统计学意义($P = 0.022$, DeLong试验)。

[0137] 进一步分析了HCC患者血浆中肝脏源性DNA的浓度(绝对和分级)与肿瘤最大尺寸(通过计算机断层扫描或肿瘤切除后测定)之间的相关性。有趣的是,肿瘤的最大尺寸与绝对浓度($R = 0.74, P < 0.0001$, Spearman相关)显示出比与分级浓度($R = 0.56, P = 0.0002$, Spearman相关)更强的正相关(图4A和图4B)。肝脏源性DNA的浓度与HCC患者肿瘤的最大尺寸呈正相关,表明从肝脏释放的DNA量将反映肿瘤负荷。

[0138] 实施例5.在有肝转移和无肝转移的CRC患者中肝脏源性DNA和结肠源性DNA的分析在30名健康受试者、35名无肝转移的CRC患者和27名有肝转移的CRC患者中测量了

肝脏源性DNA和结肠源性DNA的血浆浓度。三组的结肠源性DNA的中值血浆浓度分别为0拷贝数/mL (IQR: 0-0)、4拷贝数/mL (IQR: 0-31) 和138拷贝数/mL (IQR: 0-6850) (图5A)。三组间的浓度显著不同 ($P < 0.001$, Kruskal Wallis 试验)。在事后分析中,有肝转移的CRC患者和无肝转移的CRC患者的浓度显著高于健康受试者 (分别为 $P < 0.001$ 和 $P = 0.042$, Dunn' s 试验)。有肝转移和无肝转移的CRC患者之间的差异无统计学意义 ($P = 0.079$, Dunn' s 试验)。

[0139] 对于健康对照受试者、无肝转移的CRC患者和有肝转移的CRC患者,血浆中结肠源性DNA的中值分级浓度分别为0% (IQR: 0%-0%)、0.09% (IQR: 0%-1.1%) 和0.84% (IQR: 0%-49.5%) (图5B)。三组间的分级浓度显著不同 ($P < 0.001$, Kruskal Wallis 试验)。在事后分析中,有肝转移的CRC患者和无肝转移的CRC患者的浓度显著高于健康受试者 (分别为 $P < 0.001$ 和 $P = 0.041$, Dunn' s 试验)。有肝转移和无肝转移的CRC患者之间的差异无统计学意义 ($P = 0.084$, Dunn' s 试验)。

[0140] 对于健康对照受试者、无肝转移的CRC患者和有肝转移的CRC患者,血浆中肝脏源性DNA的中值浓度分别为40拷贝数/mL (IQR: 18-86)、23拷贝数/mL (IQR: 13-108) 和233拷贝数/mL (IQR: 56-2290) (图5C)。三组间的分级浓度显著不同 ($P < 0.001$; Kruskal-Wallis 试验)。在事后分析中,有肝转移的CRC患者的浓度显著高于无肝转移的CRC患者和健康对照组 (分别为 $P < 0.001$ 和 $P < 0.001$; Dunn 试验)。有趣的是,无肝转移的患者与健康对照组之间无显著差异 ($P = 1.0$; Dunn 试验)。

[0141] 对于健康对照受试者、无肝转移的CRC患者和有肝转移的CRC患者,血浆中肝脏源性DNA的中值分级浓度分别为0.8% (IQR: 0.3%-2.8%)、1.4% (IQR: 0.9%-3.3%) 和3.1% (IQR: 1.5%-5.3%) (图5D)。三组间的分级浓度显著不同 ($P < 0.001$, Kruskal-Wallis 试验)。在事后分析中,有肝转移的CRC患者的浓度显著高于无肝转移的CRC患者 ($P < 0.003$, Dunn' s 试验)。健康受试者的分级浓度与有肝转移的CRC患者和无肝转移的CRC患者在统计学上无显著差异 (分别为 $P = 0.114$ 和 $P = 0.717$)。

[0142] 由于在有肝转移和无肝转移的CRC患者之间观察到血浆中肝脏源性DNA和结肠源性DNA的绝对浓度和分级浓度存在显著差异,因此使用ROC曲线分析来确定哪个参数对区分两组最有用。肝脏源性DNA的绝对浓度和分级浓度的AUC分别为0.85和0.75 ($P = 0.01$, Delong 试验),并且结肠源性DNA的绝对浓度和分级浓度的AUC分别为0.69和0.69 ($P = 0.75$, Delong 试验) (图6)。

[0143] 在ROC分析中区分有肝转移和无肝转移的CRC患者,肝脏源性DNA的绝对浓度的分析优于分级浓度 (AUC: 0.85对0.75, $P = 0.01$, 图6)。不受理论的约束,可能的解释是,在一些转移至肝脏的CRC患者中,肝脏源性DNA和结肠源性DNA的绝对浓度都会增加。在一些患者中,尽管肝脏源性DNA的绝对浓度增加了,但由于结肠源性DNA的增加更大,肝脏的分级浓度保持不变或降低。类似地,还表明,与分级浓度相比,肝脏源性DNA的绝对浓度与肿瘤大小的相关性更好 ($R = 0.74$ 对0.56, Spearman 相关性, 图4A-图4B)。

[0144] 尽管本文中已经示出并描述了本公开内容的优选实施方案,但对于本领域技术人员显而易见的是,这些实施方案仅以示例的方式提供。本领域技术人员在不脱离本公开内容的前提下现将想到多种变化、改变和替代。应当理解,本文中所述的本公开内容的实施方案的各种可替代方案可用于实施本公开。所附权利要求旨在限定本公开的范围,并由此涵盖这些权利要求范围内的方法和结构及其等同方案。

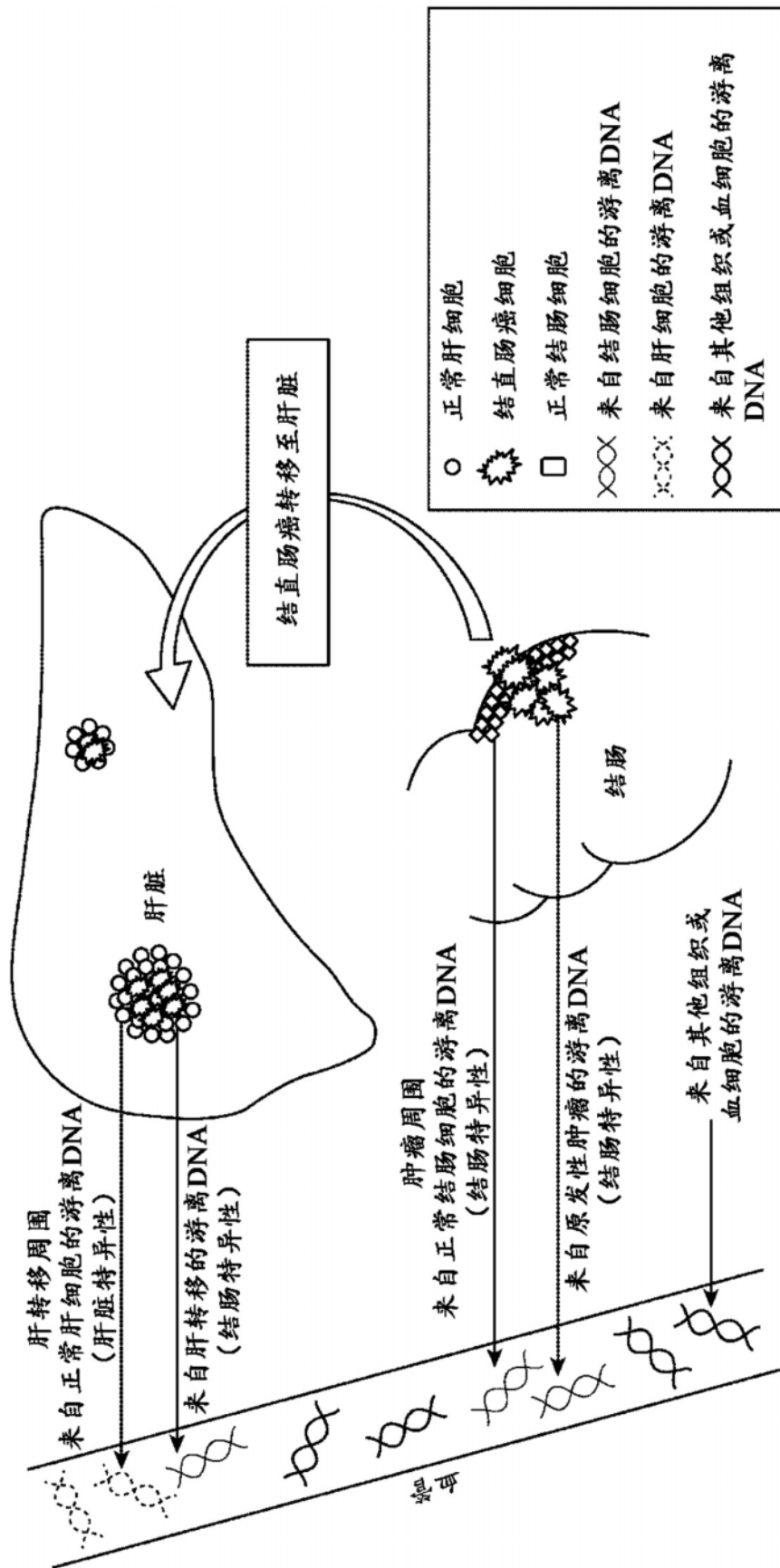


图1

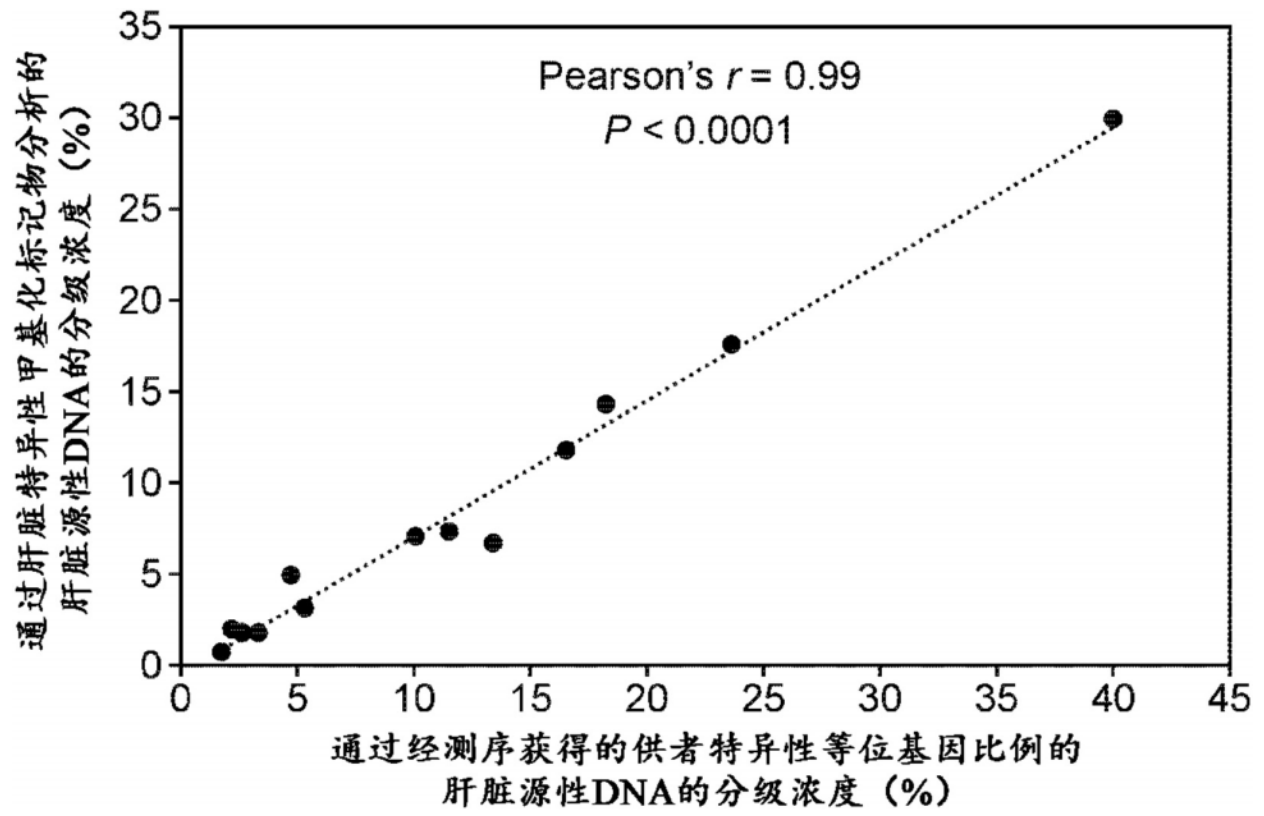


图2

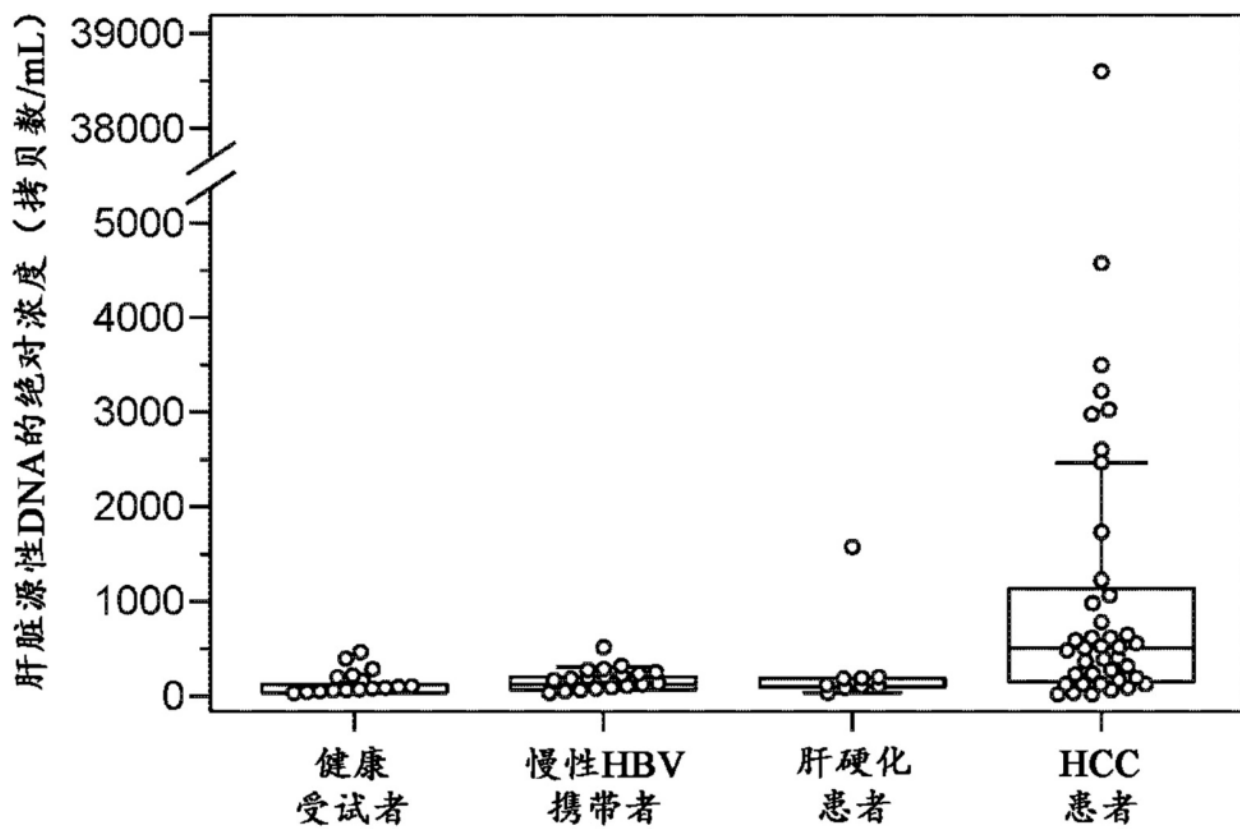


图3A

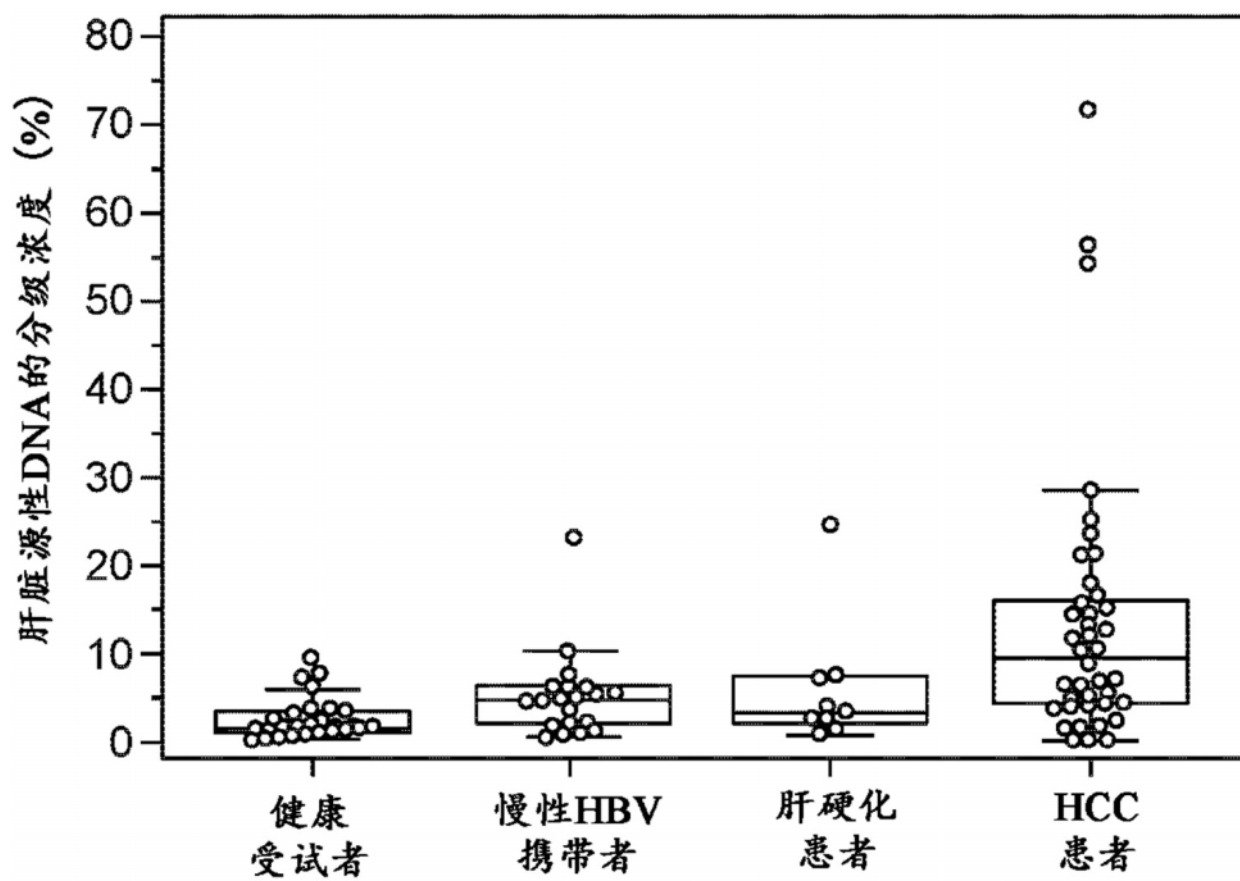


图3B

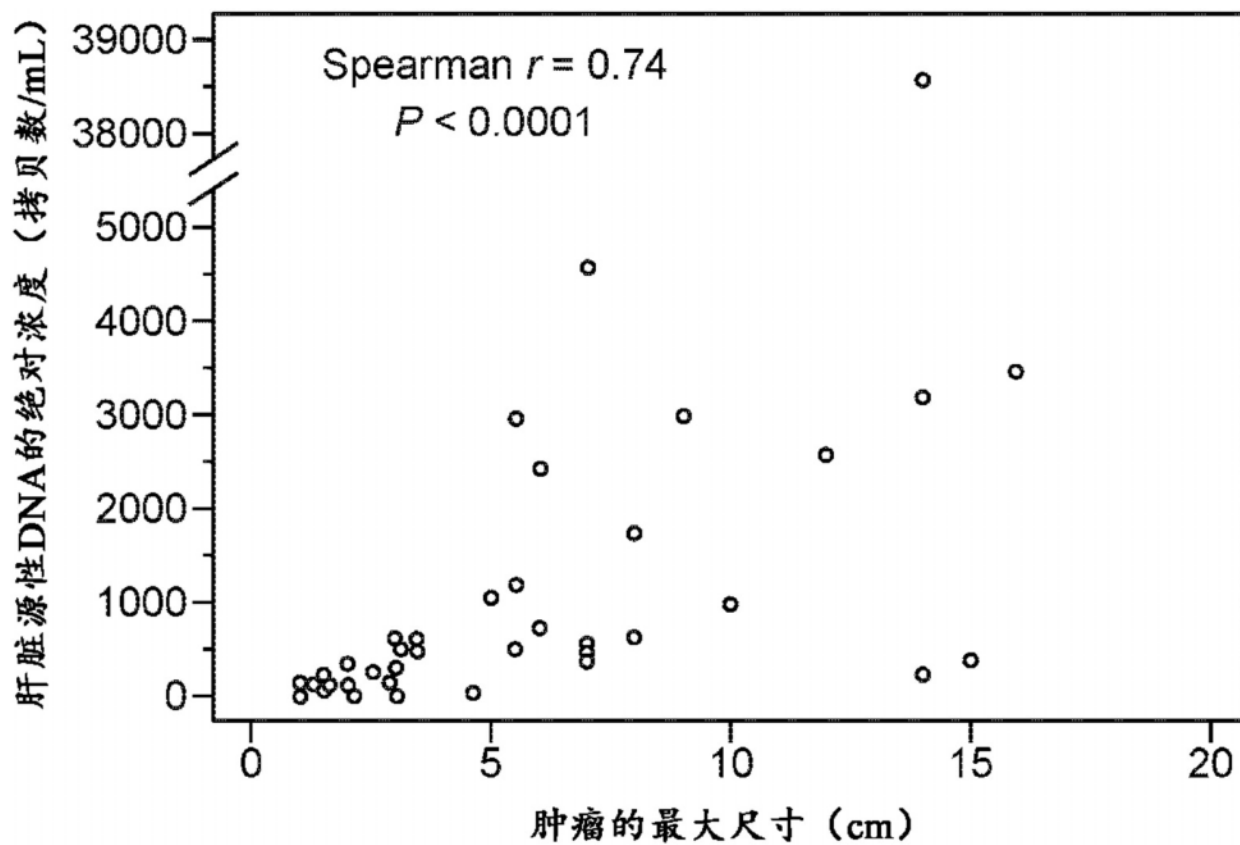


图4A

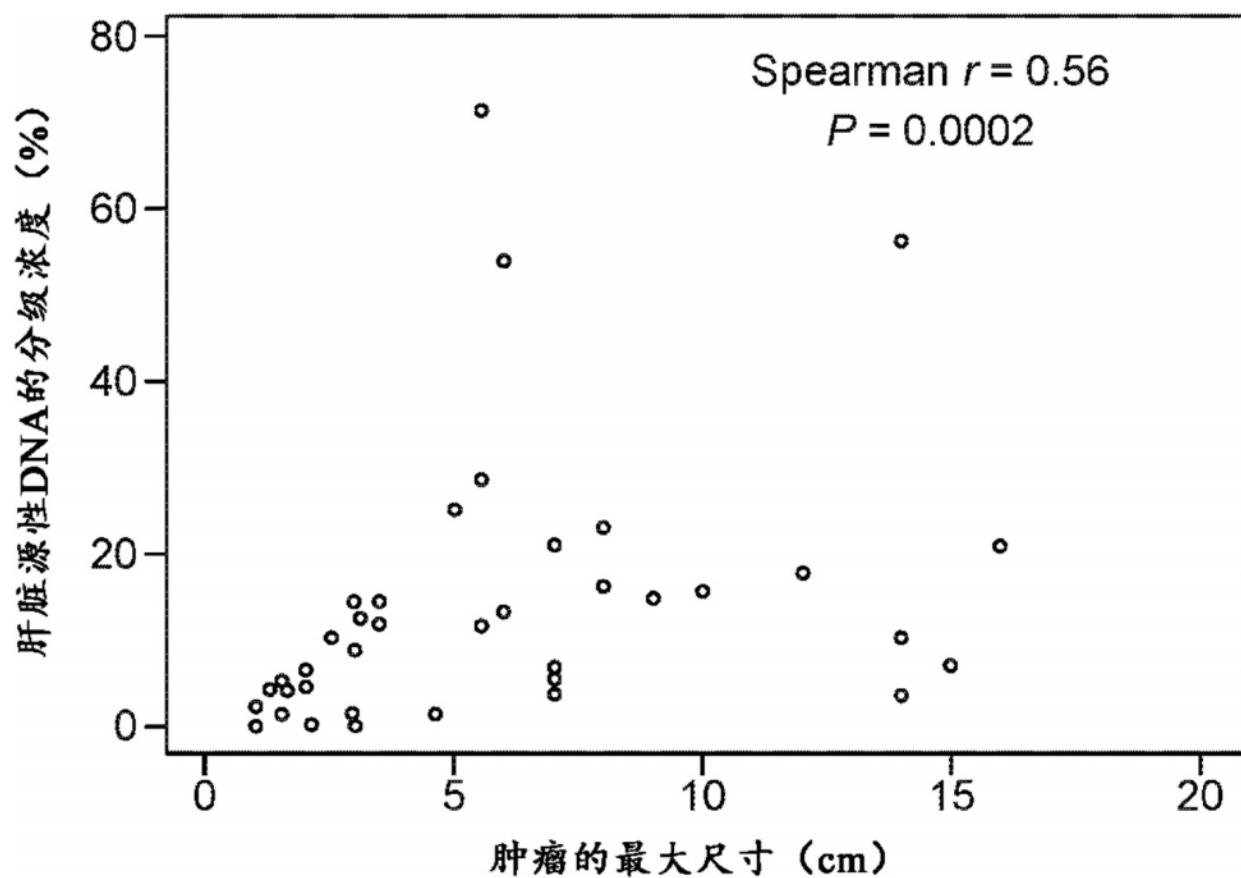


图4B

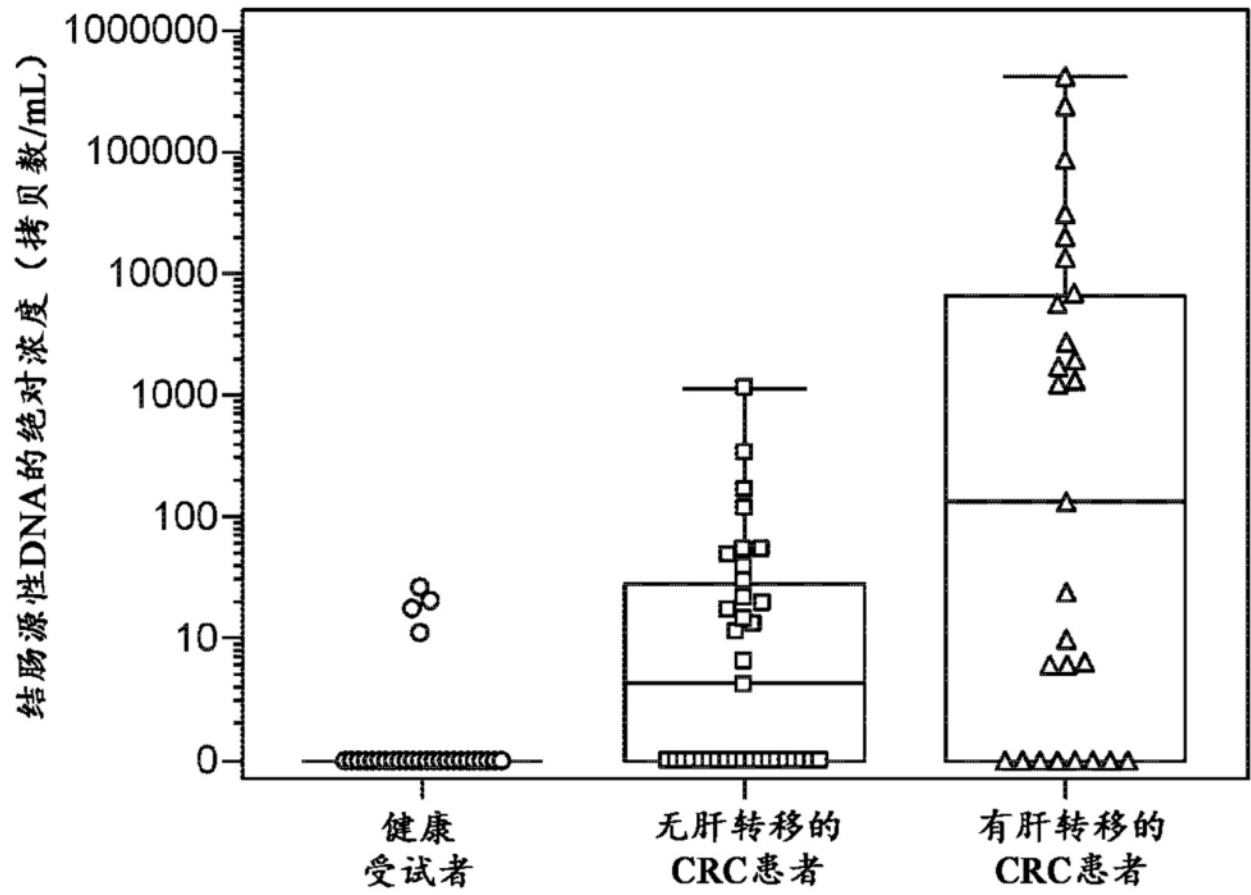


图5A

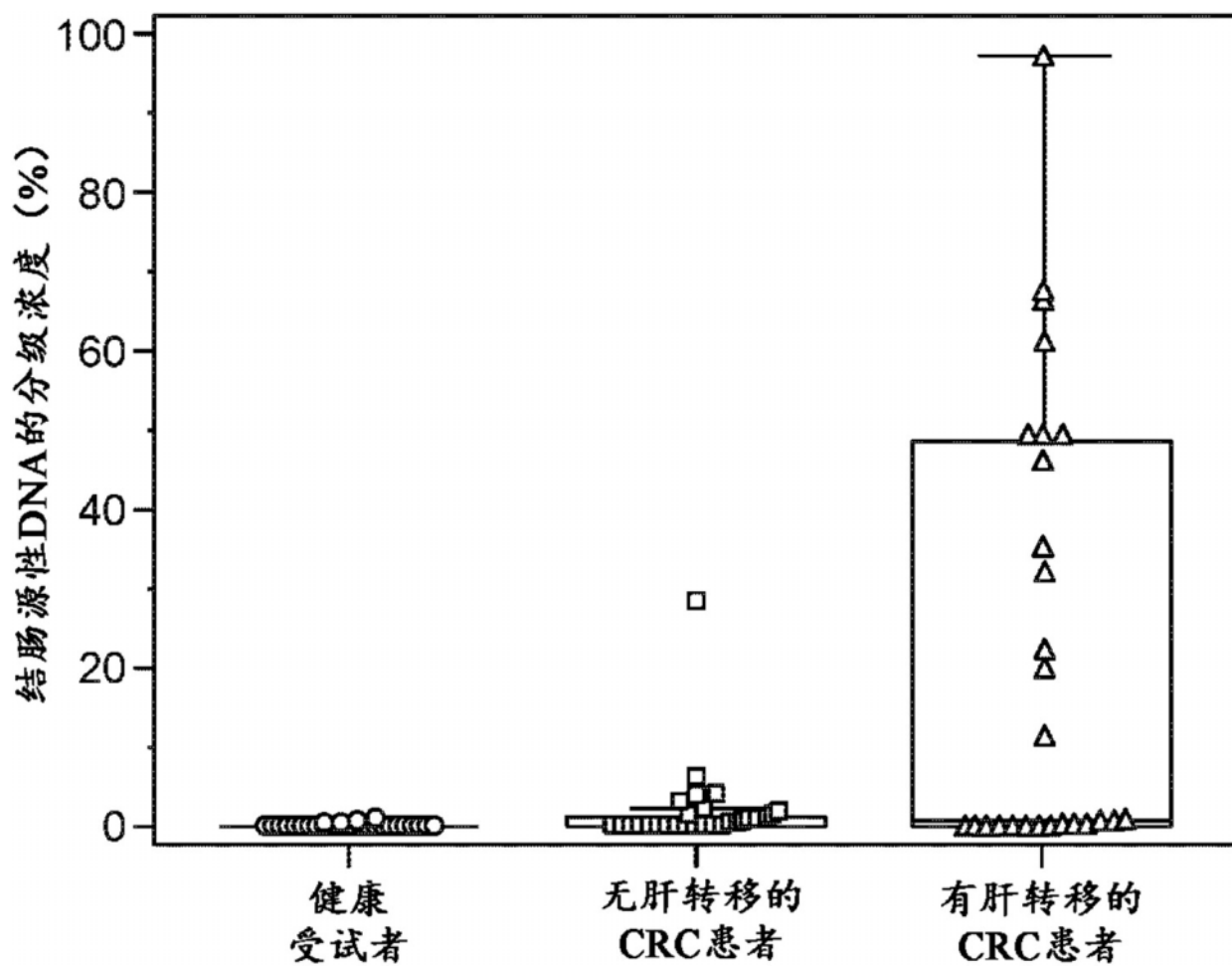


图5B

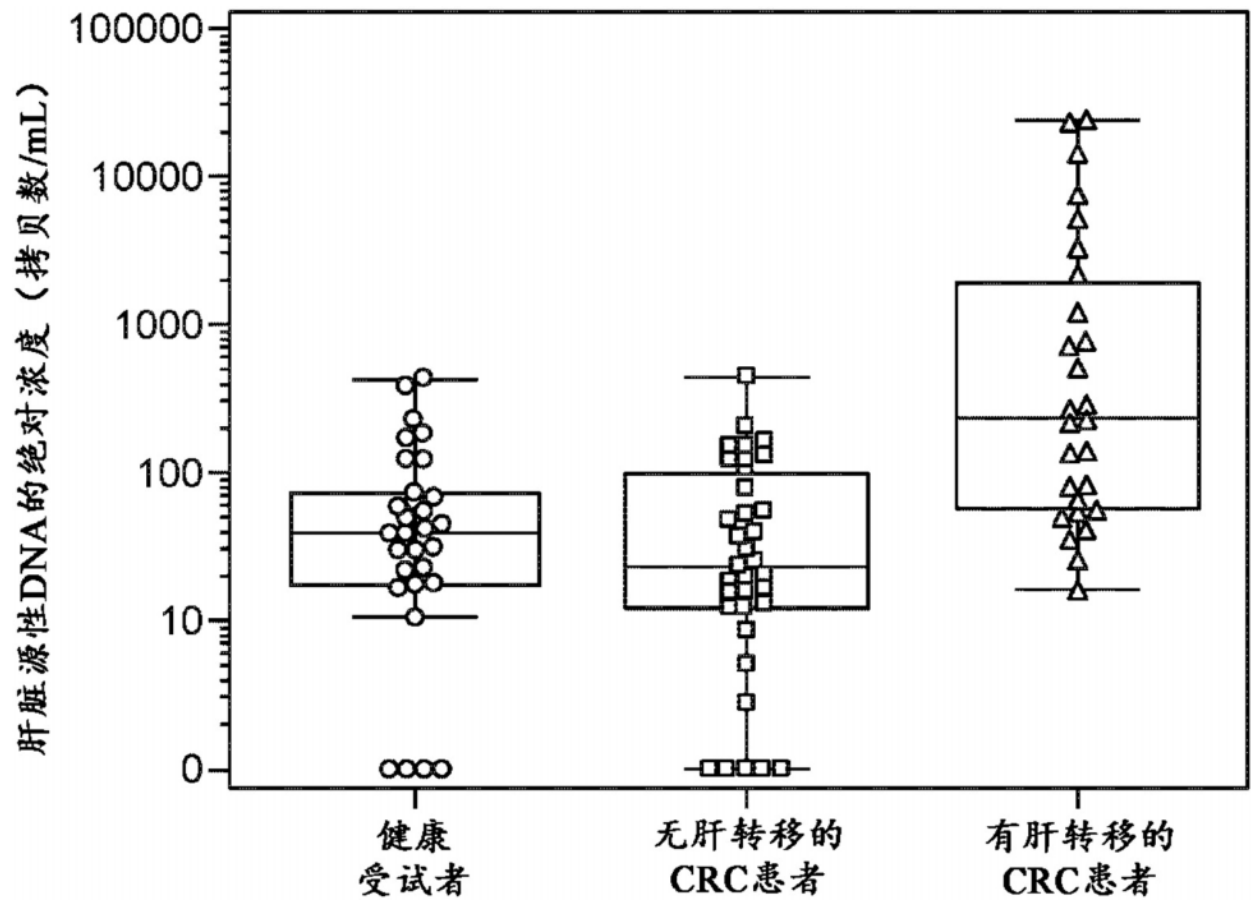


图5C

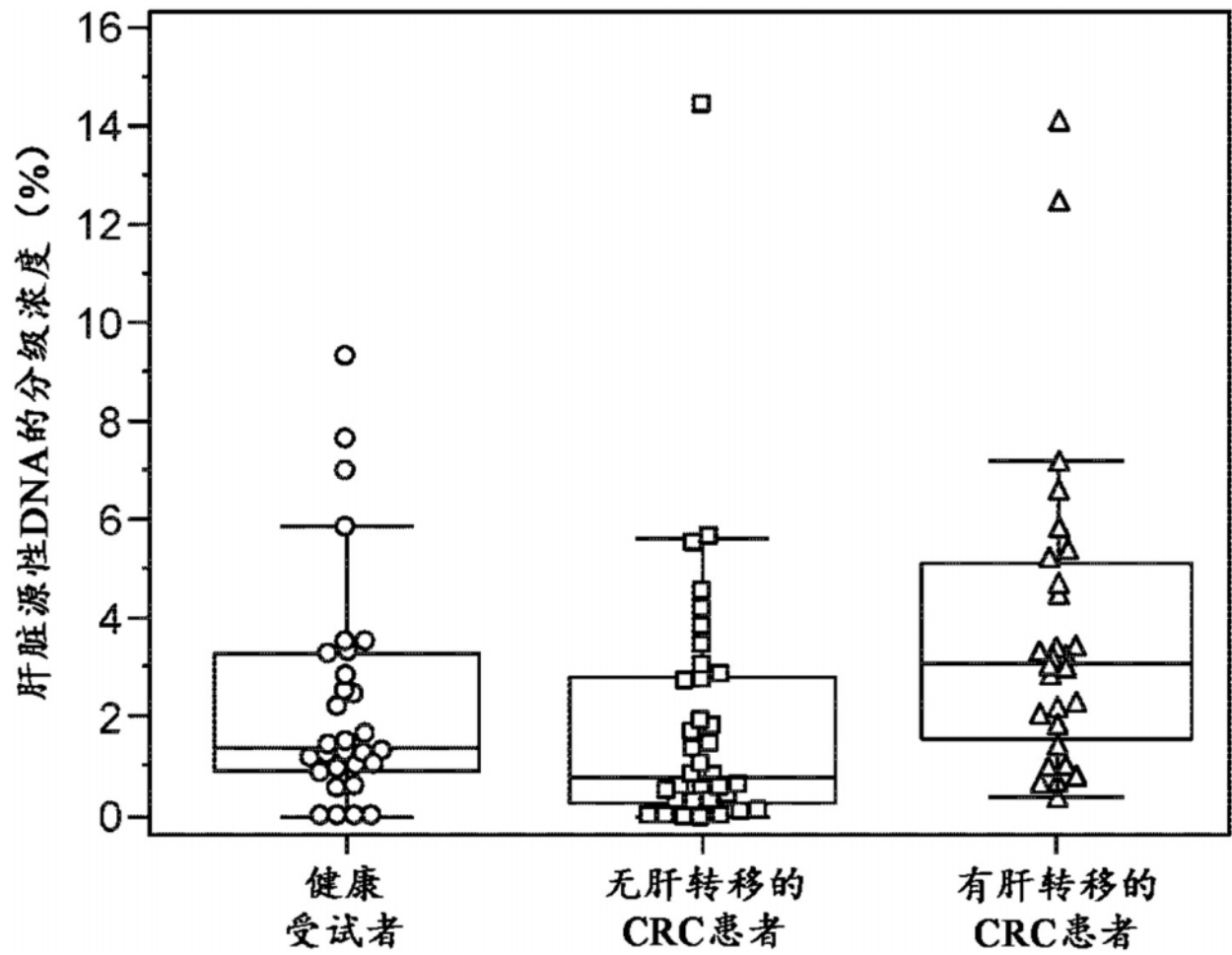


图5D

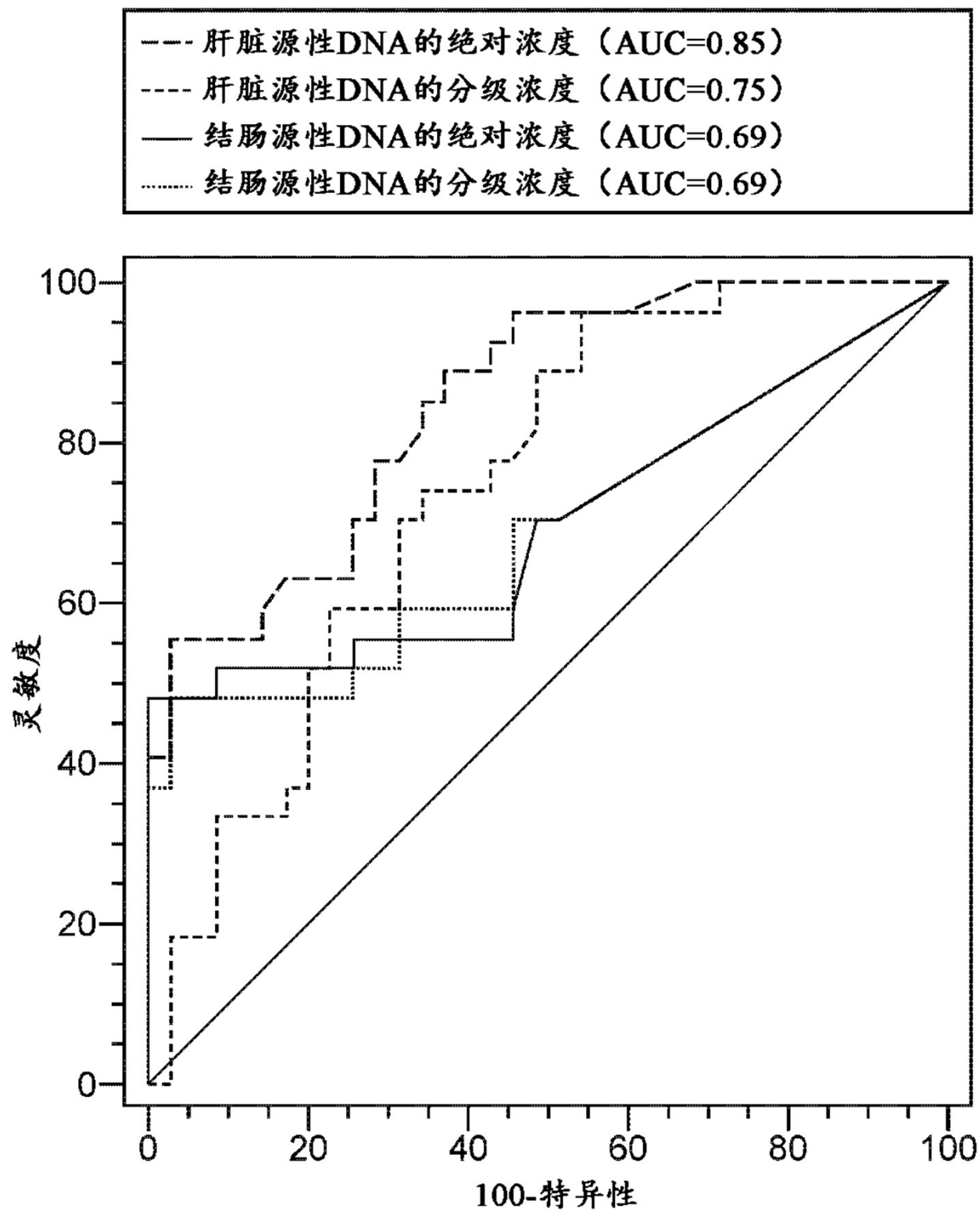


图6

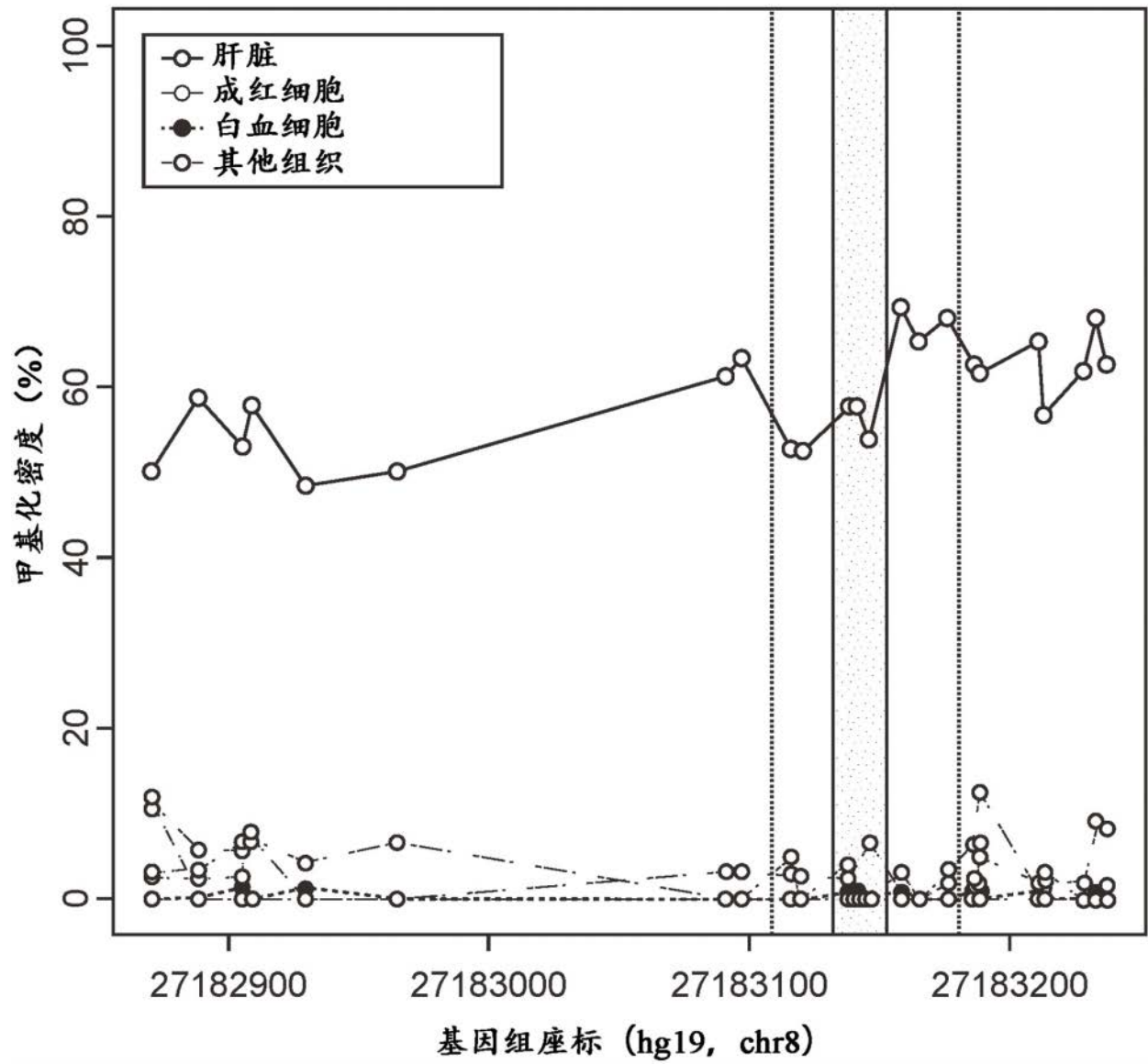


图7

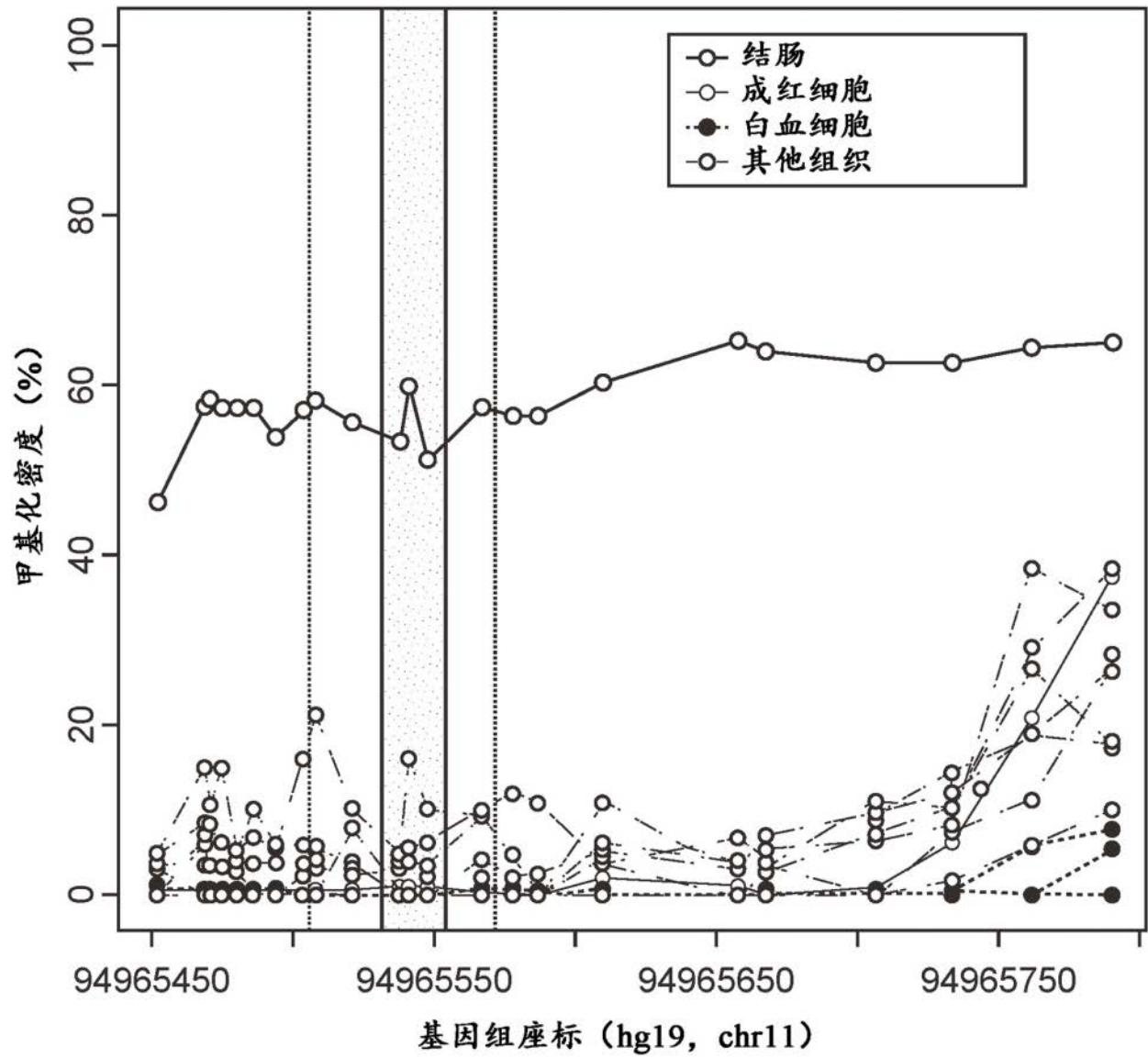


图8

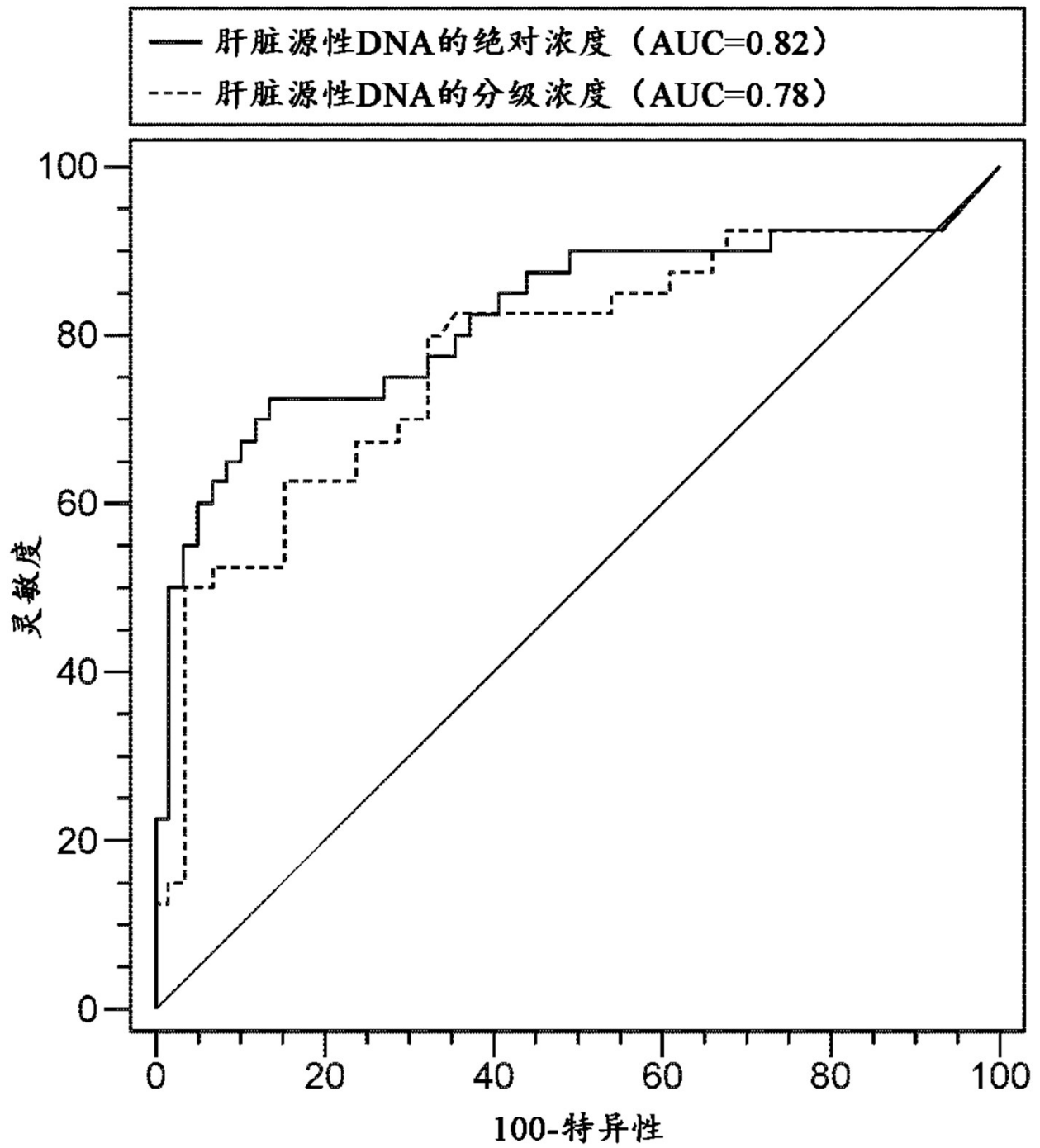


图9

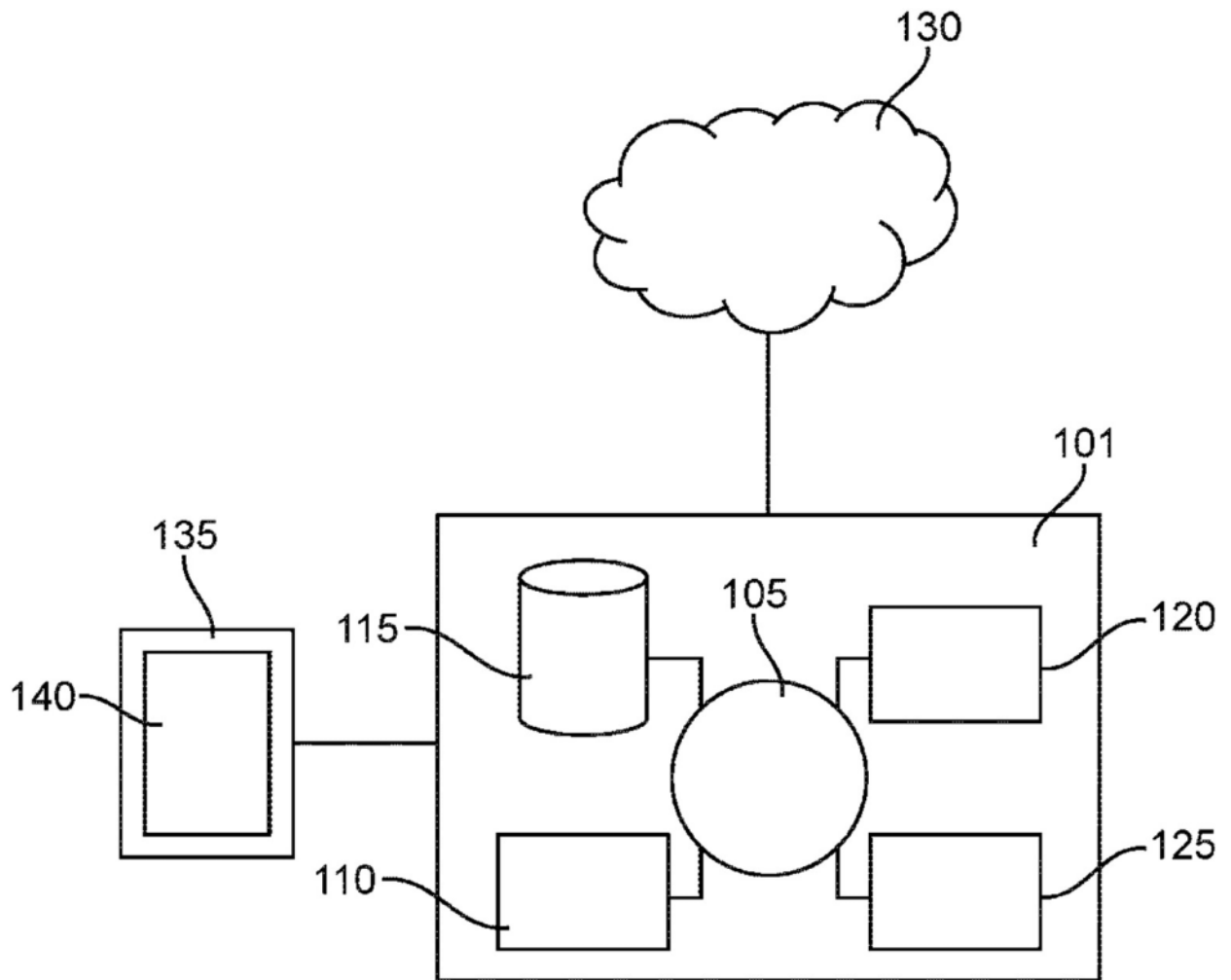


图10

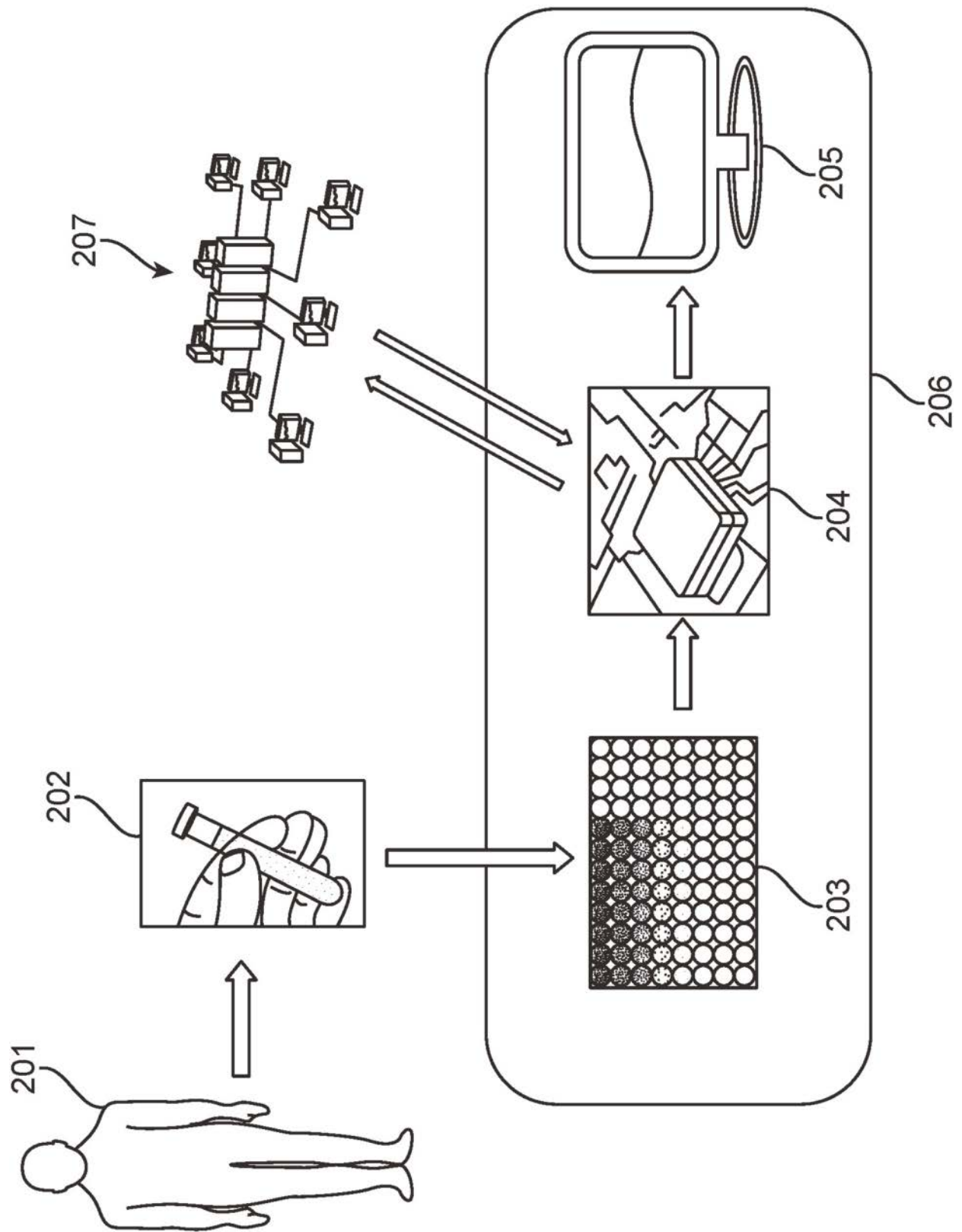


图11