(19)中华人民共和国国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 104254618 A
(43)申请公布日 2014.12.31

(21)申请号 201380013054.5
(22)申请日 2013.03.08
(30)优先权数据
61/608,623 2012.03.08 US
61/621,451 2012.04.06 US
(85)PCT国际申请进入国家阶段日
2014.09.05
(86)PCT国际申请的申请数据
PCT/IB2013/000312 2013.03.08
(87)PCT国际申请的公布数据
WO2013/132305 EN 2013.09.12
(71)申请人 香港中文大学
地址 中国香港新界
(72)发明人 卢煜明 陈君赐 郑文莉 江培勇 廖嘉炜 赵慧君
(74)专利代理机构 北京英赛嘉华知识产权代理有限责任公司 11204
代理人 王达佐 洪欣
(51)Int.Cl.
C12Q 1/68(2006.01)
G06F 19/18(2006.01)
G06F 19/20(2006.01)

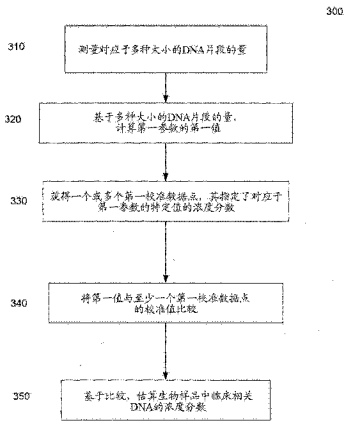
权利要求书4页 说明书22页 附图22页

(54)发明名称

母体血浆中胎儿DNA分数的基于大小的分析

(57)摘要

基于多种大小的DNA片段的量，确定来自生物样品的DNA混合物中临床相关DNA的浓度分数。例如，可以确定母体血浆中胎儿DNA或患者血浆中肿瘤DNA的浓度分数。已表明样品中DNA片段的大小分别与胎儿DNA的比例和肿瘤DNA的比例相关。校准数据点（例如，作为校准函数）指示了大小参数值与临床相关DNA的浓度分数值之间的对应性。对于给定的样品，大小参数的第一值可从样品中DNA片段的大小确定。第一值与校准数据点的比较可以提供对临床相关DNA的浓度分数的估算。



CN 104254618 A

1. 估算生物样品中临床相关 DNA 的浓度分数的方法，所述生物样品包含所述临床相关 DNA 和其他 DNA，所述方法包括：

对于多种大小中的每种大小：

测量来自所述生物样品的对应于所述大小的多个 DNA 片段的量；

使用计算机系统，基于多种大小的 DNA 片段的量，计算第一参数的第一值，所述第一参数提供了所述生物样品中 DNA 片段大小模式的统计学度量；

获得一个或多个第一校准数据点，其中每个第一校准数据点指定了对应于所述第一参数校准值的临床相关 DNA 的浓度分数，并且其中所述一个或多个校准数据点自多个校准样品确定而来；

将所述第一值与至少一个校准数据点的校准值比较；以及

基于所述比较，估算所述生物样品中临床相关 DNA 的浓度分数。

2. 如权利要求 1 所述的方法，其中所述多个 DNA 片段对应于基因组的一个或多个预定区域。

3. 如权利要求 1 所述的方法，其中所述第一参数代表相对于大 DNA 片段丰度的小 DNA 片段丰度，并且其中短 DNA 片段具有比所述大 DNA 片段更小的大小。

4. 如权利要求 1 所述的方法，还包括

基于多种大小的 DNA 片段的量，计算一个或多个第二参数的一个或多个第二值，所述一个或多个第二参数提供了所述生物样品中 DNA 片段大小模式的不同统计学度量；

获得对应于所述一个或多个第二参数的一个或多个第二校准数据点；

将所述一个或多个第二值与所述第二校准数据点的对应的第二校准值比较；并且

基于涉及所述第一值和所述一个或多个第二值的比较，估算所述生物样品中临床相关 DNA 的浓度分数。

5. 如权利要求 4 所述的方法，其中所述第一校准数据点和所述第二校准数据点是多维曲线上的点，并且所述比较包括确定具有对应于所述第一值和所述一个或多个第二值的坐标的多维点。

6. 如权利要求 1 所述的方法，其中所述第一校准数据点形成校准曲线。

7. 如权利要求 1 所述的方法，其中每个第一校准数据点自对应于不同校准样品的柱形图确定而来，其中柱形图提供了多种大小 DNA 片段的量，并且其中所述不同校准样品的至少一部分具有不同的浓度分数。

8. 如权利要求 1 所述的方法，其中测量对应于所述大小的 DNA 片段的量包括：

对于来自所述生物样品的多个 DNA 片段中的每个：

测量所述 DNA 片段的大小，

其中测量 DNA 片段的大小包括：

进行所述 DNA 片段的双端测序，以获得配对序列读数；

将所述配对序列读数与参照基因组比对；以及

使用比对的位置确定所述 DNA 片段的大小。

9. 如权利要求 1 所述的方法，其中测量对应于所述大小的 DNA 片段的量包括使用电泳。

10. 如权利要求 1 所述的方法，还包括：

通过如下计算所述一个或多个第一校准数据点：

对于所述多个校准样品中的每个：

测量所述校准样品中临床相关 DNA 的浓度分数；

测量对应于所述多种大小的 DNA 片段的量；以及

基于多种大小 DNA 片段的量，计算所述第一参数的校准值，所述校准样品的校准数据点包括所述校准值和测量的浓度分数。

11. 如权利要求 10 所述的方法，还包括：

确定函数，其近似于所述第一校准数据点在多个浓度分数间的校准值。

12. 如权利要求 11 所述的方法，其中所述函数为线性函数。

13. 如权利要求 10 所述的方法，其中所述生物样品来自怀有胎儿的孕妇，其中所述临床相关的 DNA 为胎儿 DNA，并且其中测量临床相关 DNA 的浓度分数包括以下的至少一种：

测量所述孕妇基因组中不存在的父本遗传的序列；和

测量胎儿特异性的表观遗传标志物。

14. 如权利要求 13 所述的方法，其中所述胎儿特异性表观遗传标志物包括母体血浆或血清中的展现胎儿或胎盘特异性 DNA 甲基化模式的 DNA 序列。

15. 如权利要求 10 所述的方法，其中所述临床相关的 DNA 为源自获取生物样品的患者肿瘤的 DNA。

16. 如权利要求 15 所述的方法，其中测量临床相关 DNA 的浓度分数包括：

鉴定一个或多个基因座，其中所述患者为杂合的，并且其中所述肿瘤展现杂合性丢失 (LOH) 使得等位基因缺失；

确定所述生物样品中所述一个或多个基因座处，具有未缺失的等位基因的序列读数的第一量 A；

确定所述生物样品中所述一个或多个基因座处，具有缺失的等位基因的序列读数的第二量 B；以及

使用比例 (A－B)/A 将临床相关 DNA 的浓度分数 F 计算为所述第一量与所述第二量之比。

17. 如权利要求 15 所述的方法，其中测量临床相关 DNA 的浓度分数包括：

鉴定一个或多个基因座，其中所述患者为杂合的，并且其中所述肿瘤展现一个等位基因的重复；

确定所述生物样品中所述一个或多个基因座处，具有非重复等位基因的序列读数的第一量 A；

确定所述生物样品中所述一个或多个基因座处，具有重复等位基因的序列读数的第二量 B；以及

使用比例 (B－A)/A 将临床相关 DNA 的浓度分数 F 计算为所述第一量与所述第二量之比。

18. 如权利要求 15 所述的方法，其中测量临床相关 DNA 的浓度分数包括：

鉴定一个或多个基因座，其中所述患者为纯合的，并且其中肿瘤组织中存在单核苷酸突变；

确定所述生物样品中所述一个或多个基因座处，具有野生型等位基因的序列读数的第一量 A；

确定所述生物样品中所述一个或多个基因座处，具有突变等位基因的序列读数的第二量 B ；

使用比例 2B/(A+B) 将临床相关 DNA 的浓度分数 F 计算为所述第一量与所述第二量之比。

19. 如权利要求 1 所述的方法，其中所述测量的大小为长度、分子量或与长度成比例的测量参数。

20. 如权利要求 1 所述的方法，其中所述多种大小中的至少一种对应于范围。

21. 计算机产品，包括存储多条指令的非临时性计算机可读介质，当执行时，所述指令控制计算机系统估算生物样品中临床相关 DNA 的浓度分数，所述生物样品包含所述临床相关的 DNA 和其他 DNA，所述指令包括 ：

对于多种大小中的每种大小 ：

计算来自所述生物样品的对应于所述大小的多个 DNA 片段的量 ；

基于多种大小的 DNA 片段的量，计算第一参数的第一值，所述第一参数提供了所述生物样品中 DNA 片段的大小模式的统计学度量 ；

获得一个或多个第一校准数据点，其中每个第一校准数据点指定了对应于所述第一参数的校准值的临床相关 DNA 的浓度分数，并且其中所述一个或多个校准数据点自多个校准样品确定而来 ；

将所述第一值与至少一个校准数据点的校准值比较 ；以及

基于所述比较，估算所述生物样品中临床相关 DNA 的浓度分数。

22. 分析生物体的生物样品的方法，所述生物样品包含源自正常细胞以及可能来自癌症相关细胞的 DNA，其中所述 DNA 中的至少一些为所述生物样品中无细胞的，所述方法包括 ：

对于多种大小中的每种大小 ：

测量来自生物样品的对应于所述大小的第一组 DNA 片段的量 ；

基于多种大小的 DNA 片段的量，计算第一参数的第一值，所述第一参数提供了所述生物样品中 DNA 片段的大小模式的统计学度量 ；

将所述第一值与参考值比较 ；以及

基于所述比较，确定所述生物体中癌症等级的分级。

23. 如权利要求 22 所述的方法，其中所述第一组 DNA 片段对应于所述生物体基因组的一个或多个预定区域。

24. 如权利要求 23 所述的方法，还包括 ：

鉴定来自所述生物样品的其他组的 DNA 片段，其中每组 DNA 片段对应于不同的预定区域 ；

测量对应于所述多种大小的 DNA 片段的量 ；

计算所述其他组 DNA 片段的第一参数的大小值 ；

将每个大小值与各自的参考值比较 ；以及

确定这样的预定区域 ：其中相应的大小值相比各自的参考值具有统计学差异。

25. 如权利要求 24 所述的方法，还包括 ：

使用鉴定的预定区域确定一种或多种可能的癌症类型，其中所述可能的癌症类型与所

述确定的预定区域相关。

26. 如权利要求 24 所述的方法，其中基于所述比较确定所述生物体中癌症等级的分级包括：

确定所鉴定的预定区域的数目，其中相比各自的参考值，所述相应的大小值具有统计学差异；以及

将所述数目与阈值区域数比较，以确定所述生物体中癌症等级的分级。

27. 如权利要求 24 所述的方法，其中所述各自的参考值中的至少两个是不同的。

28. 如权利要求 22 所述的方法，其中所述确定的分级对应于肿瘤大小或肿瘤数目。

29. 如权利要求 22 所述的方法，其中所述生物样品获自治疗后的生物体，并且其中所述参考值对应于治疗前采集的样品确定的第一参数的值。

30. 如权利要求 22 所述的方法，其中所述参考值对应于当推测所述生物体未患癌症时从样品确定的第一参数的值。

31. 如权利要求 22 所述的方法，其中所述参考值从获自一个或多个健康生物体的一个或多个生物样品确立。

32. 如权利要求 22 所述的方法，其中所述分级为所述生物体未患癌症或癌症等级已降低。

33. 如权利要求 22 所述的方法，其中所述分级为所述生物体确实患有癌症或者癌症等级已增加。

34. 如权利要求 22 所述的方法，其中所述生物体为人。

母体血浆中胎儿 DNA 分数的基于大小的分析

[0001] 相关申请的交叉引用

[0002] 本申请为 2012 年 3 月 8 日递交的标题为“SIZE-BASED ANALYSIS OF FETAL DNA FRACTION IN MATERNAL PLASMA(母体血浆中胎儿 DNA 分数的基于大小的分析)”的第 61/608,623 号美国临时专利申请,和 2012 年 4 月 6 日递交的标题为“SIZE-BASED ANALYSIS OF FETAL DNA FRACTION IN MATERNAL PLASMA(母体血浆中胎儿 DNA 分数的基于大小的分析)”的第 61/621,451 号美国临时专利申请的非临时性申请,并要求它们的权益,其通过引用整体并入本文,用于所有目的。

[0003] 发明背景

[0004] 母体血浆中无细胞的胎儿 DNA 的发现开创了非侵入性产前诊断的新的可能 (Lo YMD et al.Lancet 1997 ;350:485-487)。胎儿 DNA 均值 / 中值浓度分数被报导为约 3% -10% (Lo YMD et al.Am J Hum Genet 1998 ;62:768-775 ;Lun FMF et al.Clin Chem 2008 ;54:1664-1672)。胎儿 DNA 浓度分数是影响使用母体血浆 DNA 的非侵入性产前诊断测试性能的重要的参数。例如,对于胎儿染色体非整倍性 (例如 21 三体、18 三体或 13 三体) 的非侵入性的产前诊断,胎儿 DNA 浓度分数越高,母体血浆中来源于非整倍性染色体的 DNA 序列的过度表现越高。确实,已证明母体血浆中胎儿 DNA 浓度分数每减少 2 倍,将需要计数 4 倍的分子数来获得非整倍性检测 (Lo YMD et al.Proc Natl Acad Sci USA 2007 ;104:13116-13121)。

[0005] 对于通过随机大规模并行测序进行的胎儿三体非侵入性产前检测,样品的胎儿 DNA 浓度分数将影响获得较强检测所需要进行的测序量 (Fan HC and Quake SR.PLoS One 2010 ;5:e10439)。确实,一些研究组已列入了质量控制步骤,其中首先测量胎儿 DNA 浓度分数,并且只有含有大于最小胎儿 DNA 浓度分数的样品才有资格产生诊断结果 (Palomaki GE et al.Genet Med 2011 ;13:913-920)。其他研究组已在他们的诊断算法中列入了胎儿 DNA 浓度分数,用于估算特定母体血浆样品获自非整倍性妊娠的风险 (Sparks AB et al.Am J Obstet Gynecol 2012 ;206:319.e1-9)。

[0006] 除了非整倍性检测,胎儿 DNA 浓度分数还类似地影响使用母体血浆 DNA 进行的用于检测单基因疾病如血红蛋白病 (Lun FMF et al.Proc Natl Acad Sci USA 2008 ;105:19920-19925) 和血友病 (Tsui NBY et al.Blood 2011 ;117:3684-3691) 的非侵入性产前诊断测试。胎儿 DNA 浓度分数还影响构建胎儿全基因组基因图谱和突变图谱以及胎儿全基因组测序所需要进行的测序深度 (Lo YMD et al.Sci Transl Med 2010 ;2:61ra91 和美国专利申请 2011/0105353)。

[0007] 已描述了多种测量胎儿 DNA 浓度分数的方法。一种方法为测量母本基因组不存在的、胎儿特异性的、父本遗传的序列浓度。此类序列的实例包括男性胎儿中存在的 Y 染色体上的序列,和来自 Rhesus D 阴性孕妇怀有的 Rhesus D 阳性胎儿中的 RHD 基因的序列。还可使用母亲和胎儿中均存在的序列测量母体总血浆 DNA。为了得到胎儿 DNA 浓度分数,接着可以计算胎儿特异性的、父本遗传的序列浓度相比母体总血浆 DNA 浓度的比率。

[0008] 可使用的序列的另一实例包括利用单核苷酸多态性 (Lo YMD et al.Sci Transl

Med 2010 ;2:61ra91)。使用用于测量胎儿 DNA 浓度分数的遗传标志物的缺点是没有哪一组遗传标志物是所有胎儿 - 母亲对特征性的。然而可采用的另一方法是使用母体血浆中展现胎儿或胎盘特异性 DNA 甲基化模式的 DNA 序列 (Nygren AO et al.Clin Chem 2010 ; 56:1627-1635)。使用 DNA 甲基化标志物的可能缺点是可能存在 DNA 甲基化水平的个体间差异。此外,用于检测 DNA 甲基化标志物的方法通常较复杂,包括使用甲基化敏感性限制酶消化 (Chan KCA et al.Clin Chem 2008 ;52:2211-2218),或亚硫酸盐转化 (Chim SSC et al.Proc Natl Acad Sci USA2005 ;102:14753-14758),或甲基化 DNA 免疫沉淀 (MeDIP) (Papageorgiou EA et al.Nat Med 2011 ;17:510-513)。

[0009] 由于胎儿 DNA 浓度分数是重要的数值,用其他的方法和系统来确定该值是可取的。

[0010] 发明概述

[0011] 实施方案能够提供基于多种大小的 DNA 片段的量,估算来自生物样品的 DNA 混合物中临床相关 DNA 的浓度分数的方法和系统。例如,可以确定母体血浆中的胎儿 DNA 的浓度分数或患者血浆中的肿瘤 DNA 的浓度分数。已表明 DNA 片段的大小与胎儿 DNA 的比例和肿瘤 DNA 的比例相关。校准数据点 (例如,作为校准函数) 指示了大小参数值与临床相关 DNA 的浓度分数值之间的对应性。对于给定的样品,大小参数的第一值可自样品中 DNA 片段的大小确定而来。第一值与校准数据点的比较提供了关于临床相关 DNA 的浓度分数的估算。

[0012] 根据一个实施方案,方法估算生物样品中临床相关 DNA 的浓度分数,所述生物样品包含所述临床相关的 DNA 和其他 DNA。对于多种大小中的每一种大小,测量了来自生物样品的对应于所述大小的多个 DNA 片段的量。计算机系统基于多种大小的 DNA 片段的量,计算第一参数的第一值。第一参数提供了生物样品中 DNA 片段的大小模式的统计学度量。获得一个或多个第一校准数据点。每个第一校准数据点指定了对应于第一参数的校准值的临床相关 DNA 的浓度分数。所述一个或多个校准数据点自多个校准样品确定而来。将第一值与至少一个校准数据点的校准值比较。基于所述比较估算生物样品中临床相关 DNA 的浓度分数。

[0013] 根据另一个实施方案,方法分析了生物体的生物样品。所述生物样品包含源自正常细胞和可能来自癌症相关细胞的 DNA。所述 DNA 中的至少一些在所述生物样品中是无细胞的。对于多种大小中的每种大小,测量了来自生物样品的对应于所述大小的多个 DNA 片段的量。计算机系统基于多种大小 DNA 片段的量,计算了第一参数的第一值。第一参数提供了生物样品中 DNA 片段的大小模式的统计学度量。将第一值与参考值比较。基于所述比较确定生物体中癌症等级的分级。

[0014] 其他实施方案涉及系统、便携式用户装置和与本文所述方法相关的计算机可读介质。

[0015] 参考以下详细描述和附图可获得对本发明的性质和优势的更好的理解。

附图说明

[0016] 图 1 显示了根据本发明实施方案,母体血浆中循环无细胞 DNA 的大小分布的图 100。

[0017] 图 2A 显示了根据本发明实施方案，具有不同胎儿 DNA 浓度分数的两个母体血浆样品（妊娠的第一个三个月）中胎儿 DNA 的大小分布的图 200。

[0018] 图 2B 显示了根据本发明实施方案，具有不同胎儿 DNA 浓度分数的两个母体血浆样品（妊娠的第二个三个月）中 DNA 片段的大小分布的图 250。

[0019] 图 3 是方法 300 的流程图，其阐示了根据本发明实施方案估算生物样品中临床相关 DNA 的浓度分数的方法。

[0020] 图 4 是图 400，其显示了根据本发明实施方案，使用电泳获得的母体血浆 DNA 的大小分布（电泳图）。

[0021] 图 5A 是图 500，其显示了根据本发明实施方案，母体血浆中具有多种胎儿 DNA 百分比的样品的 150bp 或更小 DNA 片段的比例。

[0022] 图 5B 是图 550，其显示了≤ 150bp 的 DNA 片段与 163bp-169bp 的 DNA 的量的大小比，标示为 (CF(大小≤ 150)/ 大小 (163-169))。

[0023] 图 6A 是图 600，其显示了 140bp-146bp 的 DNA 片段与 163bp-169bp DNA 的量的大小比，标示为（大小 (140-146)/ 大小 (163-169))。

[0024] 图 6B 是图 650，其显示了 140bp-154bp 的 DNA 片段与 163bp-169bp DNA 的量的大小比，标示为（大小 (140-154)/ 大小 (163-169))。

[0025] 图 7 是图 700，其显示了 100bp-150bp 的 DNA 片段与 163bp-169bp DNA 的量的大小比，标示为（大小 (100-150)/ 大小 (163-169))。

[0026] 图 8 是图 800，其显示了根据本发明实施方案，母体血浆中具有多种胎儿 DNA 百分比的样品的 150bp 或更小 DNA 片段的比例。

[0027] 图 9A 是图 900，其显示了≤ 150bp 的 DNA 片段与 163bp-169bp DNA 的量的大小比，标示为 (CF(大小≤ 150)/ 大小 (163-169))。

[0028] 图 9B 是图 950，其显示了 140bp-146bp 的 DNA 片段与 163bp-169bp DNA 的量的大小比，标示为（大小 (140-146)/ 大小 (163-169))。

[0029] 图 10A 是图 1000，其显示了 140bp-154bp 的 DNA 片段与 163bp-169bp DNA 的量的大小比，标示为（大小 (140-154)/ 大小 (163-169))。

[0030] 图 10B 是图 1005，其显示了 100bp-150bp 的 DNA 片段与 163bp-169bp DNA 的量的大小比，标示为（大小 (100-150)/ 大小 (163-169))。

[0031] 图 11 的图显示了根据本发明实施方案，对于所示大小的重复元件，大小比相对于胎儿 DNA 百分比作图。

[0032] 图 12A 是电泳图 1200，其根据本发明实施方案可用于确定大小比。

[0033] 图 12B 是图 1250，其显示了根据本发明实施方案，母体血浆中具有多种胎儿 DNA 百分比的样品的 200bp-267bp 的 DNA 片段与 290bp-294bp DNA 的量的大小比。

[0034] 图 13 是根据本发明实施方案，由校准样品产生的测量结果确定校准数据点的方法 1300 的流程图。

[0035] 图 14A 是根据本发明实施方案，针对训练组，大小比相对于胎儿 DNA 浓度分数的图 1400。

[0036] 图 14B 是根据本发明实施方案，从图 14A 的线性函数 1410 推导（估算）的浓度分数相对于使用胎儿特异性序列测得的浓度分数的图 1450。

[0037] 图 15A 是图 1500，其显示了根据本发明实施方案，肿瘤切除之前和之后的两名肝细胞癌 (HCC) 患者血浆中具有多种肿瘤 DNA 百分比的样品的 150bp 或更小的 DNA 片段的比例。

[0038] 图 15B 是图 1550，其显示了肿瘤切除之前和之后的两名 HCC 患者的≤ 150bp 的 DNA 片段与 163bp-169bp DNA 的量的大小比，标示为 (CF(大小≤ 150)/ 大小 (163-169))。

[0039] 图 16A 是图 1600，其显示了肿瘤切除之前和之后的两名 HCC 患者的 140bp-146bp 的 DNA 片段与 163bp-169bp DNA 的量的大小比，标示为 (大小 (140-146)/ 大小 (163-169))。

[0040] 图 16B 是图 1650，其显示了肿瘤切除之前和之后的两名 HCC 患者的 140bp-154bp 的 DNA 片段与 163bp-169bp DNA 的量的大小比，标示为 (大小 (140-154)/ 大小 (163-169))。

[0041] 图 17 是图 1700，其显示了肿瘤切除之前和之后的两名 HCC 患者的 100bp-150bp 的 DNA 片段与 163bp-169bp DNA 的量的大小比，标示为 (大小 (100-150)/ 大小 (163-169))。

[0042] 图 18A 是图 1800，其显示了肿瘤切除之前和之后的 HCC 患者的 150bp 或更小的 DNA 片段的比例。

[0043] 图 18B 是图 1850，其显示了肿瘤切除之前和之后的 HCC 患者的≤ 150bp 的 DNA 片段与 163bp-169bp DNA 的量的大小比，标示为 (CF(大小≤ 150)/ 大小 (163-169))。

[0044] 图 19A 是图 1900，其显示了肿瘤切除之前和之后的 HCC 患者的 140bp-146bp 的 DNA 片段与 163bp-169bp DNA 的量的大小比，标示为 (大小 (140-146)/ 大小 (163-169))。

[0045] 图 19B 是图 1950，其显示了肿瘤切除之前和之后的 HCC 患者的 140bp-154bp 的 DNA 片段与 163bp-169bp DNA 的量的大小比，标示为 (大小 (140-154)/ 大小 (163-169))。

[0046] 图 20 是图 2000，其显示了肿瘤切除之前和之后的 HCC 患者的 100bp-150bp 的 DNA 片段与 163bp-169bp DNA 的量的大小比，标示为 (大小 (100-150)/ 大小 (163-169))。

[0047] 图 21 是流程图，其阐示了根据本发明实施方案分析生物体的生物样品以确定癌症等级的分级的方法 2100。

[0048] 图 22 是表 2200，其显示了可见于多种类型的癌症中的一些常见的染色体畸变。

[0049] 图 23 显示了可用于根据本发明实施方案的系统和方法的示例性计算机系统 2300 的方块图。

[0050] 定义

[0051] 如本文所用的术语“生物样品”是指取自对象 (例如，人，如孕妇) 且含有一种或多种目标核酸分子的任何样品。实例包括血浆、唾液、胸膜液、汗液、腹水、胆汁、尿、血清、胰液、粪便和宫颈刮片样品。生物样品可获自人、动物或其他合适的生物体。“校准样品”对应于这样的生物样品，其临床相关 DNA 分数是已知的，或可通过校准方法如使用临床相关 DNA 特异性的等位基因确定。临床相关 DNA 的实例为母体血浆中的胎儿 DNA 或患者血浆中的肿瘤 DNA。

[0052] 如本文所用，术语“基因座 (locus)”或其复数形式“基因座 (loci)”是在基因组间存在变异的任何长度的核苷酸 (或碱基对) 的位点或地址。术语“序列读数”是指获自核酸分子 (例如，DNA 片段) 的全部或一部分的序列。在一个实施方案中，仅对片段的一端测序。可选地，可对片段的两端 (例如，从每端起约 30bp) 测序以生成两个序列读数。然后可将成对的序列读数与参照基因组比对，这可提供片段长度。在又一实施方案中，例如，通过连接，可将线性 DNA 片段环化，并且可对跨越连接位点的部分测序。

[0053] 术语“通用测序”是指这样的测序，其中将适配子添加至片段的末端，并且将测序引物连接至适配子。因此，可用相同的引物对任何片段测序，因此测序可为随机的。

[0054] 术语胎儿 DNA 浓度分数与术语胎儿 DNA 比例及胎儿 DNA 分数可互换使用，并指存在于生物样品（例如，母体血浆或血清样品）中的源自胎儿的胎儿 DNA 分子的比例 (Lo YMD et al. Am J Hum Genet 1998 ;62:768-775 ;Lun FMF et al. Clin Chem 2008 ; 54:1664-1672)。类似地，术语肿瘤 DNA 浓度分数可与术语肿瘤 DNA 比例和肿瘤 DNA 分数互换使用，并指生物样品中存在的肿瘤 DNA 分子的比例。

[0055] 术语“大小模式 (size profile)”通常涉及生物样品中 DNA 片段的大小。大小模式可为提供多种大小的 DNA 片段的量的分布的柱形图。可将多种统计学参数（也称为大小参数或仅称为参数）用于区分一种大小模式与另一种。一个参数为特定大小或大小范围的 DNA 片段相对于所有 DNA 片段或相对于另一大小或范围的 DNA 片段的百分比。

[0056] “临床相关的”DNA 的实例包括母体血浆中的胎儿 DNA 和患者血浆中的肿瘤 DNA。另一实例包括移植患者血浆中的移植物相关的 DNA 量的测量结果。其他实例包括对象血浆中的造血与非造血 DNA 的相对量的测量结果。该后一实施方案可用于检测或监测或预测病理进程或造血和 / 或非造血组织相关的损伤。

[0057] “校准数据点”包括目标 DNA（即，临床相关的 DNA）的“校准值”和测量的或已知的浓度分数。校准值是测定的校准样品的大小参数的值，所述校准样品的临床相关 DNA 的浓度分数是已知的。校准数据点可以多种方式定义，例如，定义为离散点或校准函数（也称为校准曲线或校准面）。

[0058] 术语“癌症等级”可指癌症是否存在、癌症阶段、肿瘤大小、涉及多少缺失或扩增的染色体区域（例如，双倍性或三倍性），和 / 或癌症严重性的其他度量。癌症等级可为数字或其他特征。该水平可为 0。癌症等级还包括与缺失或扩增相关的恶化前的或癌症前期的状况。

[0059] 发明详述

[0060] 已经知道母体血浆中的无细胞胎儿 DNA 分子通常比母体来源的分子短 (Chan KCA et al. Clin Chem 2004 ;50:88-92 ;Lo YMD et al. Sci Transl Med 2010 ;2:61ra91)。胎儿 DNA 的存在导致母体血浆 DNA 的整体大小分布改变，并且改变的程度与胎儿 DNA 的浓度分数相关。通过测量母体血浆 DNA 的大小模式的特定值，实施方案可获得母体血浆中的胎儿 DNA 浓度分数。

[0061] 除应用于非侵入性的产前诊断外，实施方案还可用于测量生物体液中可用于临床的不同大小的核酸种类的浓度分数，其可用于癌症检测、移植和医疗监测。先前已证明癌症患者血浆中肿瘤来源的 DNA 比非癌症来源的 DNA 短 (Diehl F et al. Proc Natl Acad Sci USA 2005 ;102:16368-16373)。在移植环境下，已证明造血来源的 DNA 比非造血来源的 DNA 短 (Zheng YW et al. Clin Chem 2012 ;58:549-558)。例如，如果患者从供体接受了肝，则来源于肝（成体中的非造血器官）的 DNA 将比血浆中造血来源的 DNA 短 (Zheng YW et al. Clin Chem 2012 ;58:549-558)。类似地，在患有心肌梗死或中风的患者中，预期受损的非造血器官（即，分别为心脏和脑）释放的 DNA 将导致血浆 DNA 的大小模式向较短的范围转变。

[0062] I. 大小分布

[0063] 为了说明实施方案，我们在以下实例中表明可测量大小模式，例如，通过双端大规

模并行测序或通过电泳（例如，使用生物分析仪）。后一实例尤其有用，因为使用生物分析仪的电泳是较快而且相对便宜的方案。这将允许在对血浆 DNA 样品进行相对较贵的测序方法前，快速进行该分析来作为一种质量控制度量。

[0064] 图 1 显示了根据本发明实施方案，母体血浆中循环无细胞 DNA 的大小分布的图 100。大小分布可通过测量 DNA 片段大小，然后对多种大小 DNA 片段（例如，50 个碱基至约 220 个碱基范围内）的数目计数获得。图 100 显示了两种分布。分布 110 是针对母体血浆样品中所有的 DNA 片段，而分布 120 是仅针对来自胎儿的 DNA。水平轴是 DNA 片段碱基对 (bp) 的大小。垂直轴是测量的 DNA 片段的百分比。

[0065] 在图 1 中，已证明母体血浆中胎儿来源的 DNA 的大小分布比母体来源的分子短 (Chan KC et al.ClinChem 2004 ;50:88-92)。最近，我们使用双端大规模并行测序分析测定了孕妇中胎儿特异性 DNA 和总 DNA（主要来源于母亲）的高分辨率大小分布。我们证明两种 DNA 间的主要差异为 :对于胎儿来源的 DNA，166bp DNA 片段分数减小，且 150bp 以下的较短 DNA 的比例增加 (Lo YM et al.Sci Transl Med 20102:61ra91)。

[0066] 在本文中，我们概述了母体血浆样品（生物样品的一个实例）中总 DNA 片段的大小分布的分析如何有利于确定母体血浆中胎儿 DNA 的浓度分数。母体血浆中胎儿 DNA 浓度分数的增加将导致总 DNA 的整体大小分布缩短。在一个实施方案中，约 144bp DNA 片段和约 166bp DNA 片段的相对丰度（参数的一个实例）可用于反映胎儿 DNA 的浓度分数。在另一实施方案中，关于大小模式的其他参数或参数组合可用于反映血浆 DNA 的大小分布。

[0067] 图 2A 显示了根据本发明实施方案，具有不同胎儿 DNA 浓度分数的两个母体血浆样品（妊娠的第一个三个月）中胎儿 DNA 的大小分布的图 200。这两名孕妇均怀有男性胎儿。胎儿 DNA 浓度分数由来自 Y 染色体的序列在总测序 DNA 片段中的比例而确定。两个样品均采自妊娠第一个三个月的孕妇。个例 338（实线，胎儿 DNA 浓度分数 10％）具有比个例 263（虚线，胎儿 DNA 浓度分数 20％）低的胎儿 DNA 浓度分数。当与个例 263 相比时，个例 338 在 166bp 处具有较高的峰，而对于 150bp 以下的大小峰较低。换句话说，个例 263 中短于 150bp 的 DNA 片段更为丰富，而个例 338 中约 166bp 的片段更为丰富。这些观察与假设一致，即长 DNA 和短 DNA 的相对量可能与胎儿 DNA 浓度分数相关。

[0068] 图 2B 显示了根据本发明实施方案，具有不同胎儿 DNA 浓度分数的两个母体血浆样品（妊娠的第二个三个月）中 DNA 片段的大小分布的图 250。两个样品均采自第二个三个月的孕妇。这两名孕妇均怀有男性胎儿。胎儿 DNA 浓度分数由来自 Y 染色体的序列在总测序 DNA 片段中的比例而确定。类似于之前的实例，个例 5415（虚线，具有较高的胎儿 DNA 浓度分数 19％）的 150bp 以下的大小具有较高的峰，而个例 5166（实线，具有较低的胎儿 DNA 浓度分数 12％）在 166bp 处具有较高的峰。

[0069] 大小参数的不同值与胎儿 DNA 浓度分数值的相关性显示在下面的数据图中。另外，肿瘤 DNA 片段的大小与具有肿瘤 DNA 片段和来自正常细胞的 DNA 片段的样品中肿瘤 DNA 片段的百分比相关。因此，肿瘤片段大小还可用于确定样品中肿瘤片段的百分比。

[0070] II. 方法

[0071] 因为 DNA 片段大小与浓度分数（也称为百分比）相关，实施方案可使用该相关性来确定样品中具体类型的 DNA（例如，胎儿 DNA 或来自肿瘤的 DNA) 的浓度分数。具体类型的 DNA 是临床相关的，因为其为待估算的浓度分数。因此，方法可基于测得的 DNA 片段大小，

估算生物样品中临床相关 DNA 的浓度分数。

[0072] 图 3 是方法 300 的流程图，其阐示了根据本发明实施方案，估算生物样品中临床相关 DNA 的浓度分数的方法。生物样品包含临床相关的 DNA 和其他 DNA。生物样品可获自患者，例如，怀有胎儿的女性对象。在另一实施方案中，患者可患有或疑似患有肿瘤。在一个实施方案中，可将生物样品接收于仪器，例如，测序仪，其输出可用于确定 DNA 片段大小的测量数据（例如，序列读数）。方法 300 可全部或部分用计算机系统进行，如同本文所述其他方法所能进行的那样。

[0073] 在方框 310 中，测量了对应于多种大小的 DNA 片段的量。对于多种大小中的每种大小，可测量生物样品的对应于所述大小的多个 DNA 片段的量。例如，可测量具有 140 个碱基长度的 DNA 片段的数目。所述量可保存为柱形图。在一个实施方案中，测量了来自生物样品的多种核酸中的每种的大小，其可基于个体进行（例如，通过单分子测序）或基于群组进行（例如，通过电泳）。所述大小可对应于范围。因此，量可针对具有特定范围大小的 DNA 片段。

[0074] 可随机挑选多种 DNA 片段，或优选地，从基因组的一个或多个预定区域挑选多种 DNA 片段。例如，可进行靶向富集，如上文所述。在另一实施方案中，可对 DNA 片段随机测序（例如，使用通用测序），并且可将得到的序列读数与对应于对象（例如，参照的人基因组）的基因组比对。然后，可仅将序列读数与一个或多个预定区域对齐的 DNA 片段用于确定大小。

[0075] 在多个实施方案中，大小可为质量、长度或其他合适的大小度量。测量可以多种方式进行，如本文所述。例如，可进行双端测序和 DNA 片段比对，或可使用电泳。可测量统计学显著数目的 DNA 片段，以提供生物样品的精确大小模式。统计学显著数目的 DNA 片段的实例包括大于 100,000 ;1,000,000 ;2,000,000，或其他合适的值，这可取决于所需的精确度。

[0076] 在一个实施方案中，可将获自物理测量如双端测序或电泳的数据接收于计算机，并分析以实现 DNA 片段大小测量。例如，可分析（例如，通过比对）来自双端测序的序列读数来确定大小。再例如，可分析产生自电泳的电泳图以确定大小。在一个实施方案中，DNA 片段的分析确实包括实际的测序过程或对 DNA 片段进行电泳，但其他实施方案可仅进行所得数据的分析。

[0077] 在方框 320 中，基于多种大小 DNA 片段的量，计算第一参数的第一值。在一个方面，第一参数提供了生物样品中 DNA 片段的大小模式的统计学度量（例如，柱形图）。所述参数可称为大小参数，因为其自多种 DNA 片段的大小确定而来。

[0078] 第一参数可具有多种形式。此类参数为特定大小的 DNA 片段数除以片段总数，其可从柱形图（任何数据结构，提供了特定大小片段的绝对或相对计数）获得。再例如，参数可为特定大小或特定范围片段的数目除以另一大小或范围片段的数目。该除法可用作标准化，以解释针对不同样品分析的 DNA 片段的不同数目。标准化可通过针对每个样品分析相同数目的 DNA 片段实现，其有效地提供了与除以分析的片段的总数相同的结果。本文描述了参数的其他实例。

[0079] 在方框 330 中，获得了一个或多个第一校准数据点。每个第一校准数据点可指定对应于第一参数的特定值（校准值）的临床相关 DNA 的浓度分数。浓度分数可指定为特定浓度或浓度范围。校准值可对应于从多个校准样品确定的第一参数（即，特定大小参数）

的值。校准数据点可自具有已知浓度分数（其可通过本文描述的多种技术测量）的校准样品确定而来。校准样品中的至少一些具有不同的浓度分数，但一些校准样品可具有相同的浓度分数。

[0080] 在多个实施方案中，一个或多个校准点可定义为一个离散点、一组离散点、函数、一个离散点和函数，或离散或连续数值组的任何其他组合。例如，校准数据点可自具有特定浓度分数的样品的大小参数（例如，特定大小或大小范围片段的数目）的一个校准值确定而来。可使用多个柱形图，每个校准样品具有不同的柱形图，其中校准样品中的一些可具有相同的浓度分数。

[0081] 在一个实施方案中，可将从相同浓度分数的多个样品测得的相同大小参数的值组合，以确定特定浓度分数的校准数据点。例如，可从相同浓度分数的样品的大小数据获得大小参数数值的平均值，以确定特定校准数据点（或提供对应于校准数据点的范围）。在另一个实施方案中，具有相同校准值的多个数据点可用于确定平均浓度分数。

[0082] 在一个实施方案中，测量了多个校准样品的 DNA 片段的大小。确定了每个校准样品的相同大小参数的校准值，其中可将所述大小参数针对样品的已知的浓度分数作图。然后可将函数与图的数据点拟合，其中所述函数拟合确定了用于确定新样品的浓度分数的校准数据点。

[0083] 在方框 340 中，将第一值与至少一个校准数据点的校准值比较。比较可以多种方式进行。例如，比较可为第一值是否高于或低于校准值。比较可包括与校准曲线（由校准数据点组成）比较，因此比较可确定具有第一参数的第一值的曲线上的点。例如，计算的第一参数的数值 X（如从测得的新样品中 DNA 的大小确定的）可用作函数 F(X) 的输入，其中 F 为校准函数（曲线）。F(X) 的输出为浓度分数。可提供误差范围，其对于每个 X 值可能是不同的，从而提供了 F(X) 的输出值的范围。

[0084] 在步骤 350 中，生物样品中临床相关 DNA 的浓度分数基于比较来估算。在一个实施方案中，可以确定第一参数的第一值是大于还是小于阈值校准值，从而能确定估算的本样品的浓度分数是大于还是小于对应于阈值校准值的浓度分数。例如，如果计算的生物样品的第一值 X_1 大于校准值 X_C，则生物样品的浓度分数 FC_1 可确定为大于对应于 X_C 的浓度分数 FC_C。该比较可用于确定生物样品中是否存在进行其他检测（例如，检测胎儿非整倍性）的足够的浓度分数。该大于和小于的关联可取决于参数如何定义。在此类实施方案中，可能仅需要一个校准数据点。

[0085] 在另一个实施方案中，通过输入第一值至校准函数来实现比较。校准函数可通过确定对应于第一值的曲线上的点，有效地比较第一值与校准值。然后可将估算的浓度分数提供为校准函数的输出值。

[0086] 在一个实施方案中，可确定生物样品的多于一个参数的值。例如，可确定第二参数的第二值，其对应于生物样品中 DNA 片段大小模式的不同的统计学度量。第二值可使用 DNA 片段的相同的大小测量或不同的大小测量确定。每个参数可对应于不同的校准曲线。在一个实施方案中，可将不同的值独立地与不同的校准曲线比较，以获得多个估算的浓度分数，然后可将其平均或用于提供作为输出的范围。

[0087] 在另一实施方案中，可使用多维校准曲线，其中可将不同的参数值有效地输入至输出浓度分数的单个校准函数。单个校准函数可产生自获自校准样品的所有数据点的函数

拟合。因此，在一个实施方案中，第一校准数据点和第二校准数据点可为多维曲线上的点，其中比较包括确定具有对应于第一值和一个或多个第二值的坐标的多维点。

[0088] III. 测定大小

[0089] 可测定血浆DNA的大小分布，例如但不限于，使用实时PCR、电泳和质谱分析。在多个实施方案中，所测的大小为长度、分子量或测量的与长度或质量成比例的参数，如电泳图谱中的迁移性和在电泳或质谱仪中移动固定距离所需的时间。在另一个实例中，可用嵌入性荧光染料如溴化乙锭或SYBR Green对DNA染色，其中染料结合的量与DNA分子的长度成比例。可以通过UV光照射于样品上时发出的荧光的强度，确定结合的染料的量。测量大小的一些实例以及得到的数据描述如下。

[0090] A. 使用测序的第一胎儿样品集

[0091] 表1显示了以胎儿DNA分数为例的样品信息和测序分析。血浆样品取自80名孕妇，每名怀有一个男性胎儿。在这80名孕妇中，39名怀有整倍体胎儿，18名怀有21三体(T21)胎儿，10名怀有18三体(T18)胎儿，且13名怀有13三体(T13)胎儿。使用双端大规模并行测序确定血浆DNA的大小分布。母体血浆DNA的测序文库按先前所述构建(Lo YM et al. Sci Transl Med 2010;2:61ra91)，除了通过三引物PCR扩增将6个碱基的标识符引入至每个血浆样品的DNA分子。

[0092] 将两个样品引入一个测序道(即，2倍测序)。在其他实施方案中，可将多于两个样品引入一个测序道，例如，6或12或20个，或多于20个。所有文库均通过基因组分析仪IIx(Illumina)使用36-bp×2PE格式测序。进行了另外的7轮测序以解译每个测序的血浆DNA分子上的索引序列。使用短寡核苷酸比对程序2(Short Oligonucleotide Alignment Program 2，SOAP2)(soap.genomics.org.cn)，将36-bp的序列读数与非重复掩蔽的(non-repeat-masked)人参照基因组(Hg18)(genome.ucsc.edu)比对。确定了具有单独的成员的双端(PE)读数，所述成员在流动池(flow cell)的相同簇位置上测序，且以正确方向和无任何核苷酸错配地、唯一地与人基因组中的单个位置对齐。在其他实施方案中，比对可能不唯一的且可允许错配。

[0093] 仅回收展现插入物大小≤600bp的PE读数用于分析。利用这些标准，这些实验中分析的血浆DNA片段的大小范围为36bp-600bp。每个测序的DNA片段的大小从测序片段每端的最外面的核苷酸坐标推导而来。

[0094]

个例类型	个例数	孕龄(周)中值(范围)	PE读数数量(百万)中值(范围)	胎儿DNA分数(%)中值(范围)
整倍体	39	13.2(11.3-15.1)	4.7(1.8-12.0)	15.7(5.9-25.7)
T21	18	13.0(12.1-17.9)	5.2(2.5-8.9)	13.8(7.4-27.2)
T18	10	13.3(12.1-14.2)	4.9(3.6-6.2)	7.2(4.8-16.7)
T13	13	12.4(11.5-16.4)	5.3(2.7-7.7)	7.5(3.2-14.1)
总体	80	13.1(11.3-17.9)	4.9(1.8-12.0)	13.7(3.2-27.2)

[0095] 表1显示了多种非整倍性状态样品的数据。数据包括个例数、孕龄中值和范围，及

双端读数数量中值和范围，以及胎儿 DNA 分数。

[0096] 母体血浆样品中胎儿DNA的浓度分数按先前所述从与Y染色体对齐的序列的量推导而来 (Chiu RW et al.BMJ 2011 ;342:c7401)。该技术是校准方法的一个实例。因此，表1中测量的胎儿DNA分数可用于评估新样品中的胎儿DNA分数的校准数据点。用于收集表1中的数据的样品可被认为是校准样品。

[0097] B. 使用靶向测序的第二胎儿样品集

[0098] 表2显示了根据本发明实施方案，母体血浆DNA的样品信息和靶向富集。血浆样品采自48名孕妇，每名怀有一个胎儿。在这48名孕妇中，21名怀有整倍体胎儿，17名怀有21三体 (T21) 胎儿，9名怀有18三体 (T18) 胎儿，且1名怀有13三体 (T13) 胎儿。这些数据连同以下的实例证明实施方案可使用靶向技术。血浆DNA的大小分布可使用双端大规模并行测序确定。在其他实施方案中，血浆DNA的大小分布可例如但不限于使用实时PCR、电泳和质谱分析确定。

[0099] 为获得靶标区域的高倍测序覆盖，在一个实施方案中使用了 Agilent SureSelect 靶标富集系统设计探针来捕获来自 chr7(0.9Mb 区域）、chr13(1.1Mb 区域）、chr18(1.2Mb 区域）和 chr21(1.3Mb 区域）的 DNA 分子。在探针设计中，首先挑选 chr7、chr13、chr18 上的外显子和 chr21 上的唐氏综合症决定区 (21q22.1-q22.3) 作为靶标区域。因为 chr13、chr18 和 chr21 比 chr7 具有更少的外显子区域，引入了 chr13、chr18 上的其他非外显子区域和 chr21 上的唐氏综合症决定区来平衡上述 4 条染色体间的靶向区域的总长度。所选的非外显子区域长度为 120bp，可唯一地映射，GC 含量接近 0.5，且在所靶向的染色体上均匀分布。

[0100] 将所有上述外显子和非外显子区域的坐标提交至 Agilent eArray 平台用于探针设计。将 500ng 的每种母体血浆 DNA 文库在 65℃下与捕获探针孵育 24h。杂交后，洗脱所靶向的DNA分子并根据厂商说明书通过12个循环的PCR进行扩增。使用50-bp×2PE格式在GA IIx(Illumina) 上对有靶标富集的文库编索引并测序。进行了另外的 7 轮测序以解译每个测序的血浆 DNA 分子上的索引序列。使用短寡核苷酸比对程序 2(SOAP2)(soap.genomics.org.cn)，将 50-bp 序列读数与非重复掩蔽的人参照基因组 (Hg18)(genome.ucsc.edu) 比对。将具有单独的成员的 PE 读数在流动池的相同簇位置上进行测序，并以正确方向与人基因组中的单个位置唯一地对齐。允许有两个错配 ;靶标富集后测序文库的复杂度明显降低。

[0101] 仅回收展现插入物大小≤ 600bp 的 PE 读数用于分析。利用这些标准，本研究中分析的血浆DNA片段的大小范围为36bp-600bp。每个测序的DNA片段的大小从测序片段每端最外面的核苷酸坐标推导而来。从携带胎儿特异性等位基因的片段和携带与各自母亲共有的等位基因的片段的比估算母体血浆样品中胎儿 DNA 的浓度分数。

[0102]

个例类型	个例数	孕龄(周)中值(范围)	PE 读数数量(百万)中值(范围)	胎儿 DNA 分数(%)中值(范围)
整倍体	21	13.0 (12.0 –13.3)	2.2 (1.7-3.0)	13.5 (8.4-22.0)
T21	17	13.6 (12.6 – 20.9)	2.1 (1.5-2.7)	15.4 (8.7-22.7)
T18	9	12.7 (11.9 –13.7)	1.9 (1.7-3.1)	10.5 (7.2-16.3)
T13	1	13	1.6	9.2
总体	48	13.1 (11.9 – 20.9)	2.1 (1.5-3.1)	13.4 (7.2-22.7)

[0103] 表 2 显示了来自多种非整倍性状态样品的靶向测序的数据。

[0104] C. 胎儿样品的电泳

[0105] 除了使用大规模并行测序，血浆 DNA 大小分布的分析还可通过电泳实现。电泳测量片段穿过介质的时间。不同大小的颗粒穿过介质所花的时间不同。因此，在一个实施方案中，可进行母体血浆 DNA 测序文库的微流体电泳来确定母体血浆 DNA 的大小分布。

[0106] 图 4 是图 400，其显示了根据本发明实施方案，使用电泳获得的母体血浆 DNA 的大小分布（电泳图）。微流体电泳使用 Agilent 2100 生物分析仪进行。两个样品的测序文库的电泳图显示在图 400 中。X 轴代表 DNA 到达传感器所花的时间长度，且对应于 DNA 片段的大小。Y 轴代表特定时间 DNA 片段通过传感器的荧光单位 (FU)。

[0107] DNA 片段到达传感器所花的时间长度与 DNA 片段的大小正相关。生物分析仪通过将检测样品的运行时间与具有已知长度的 DNA 片段混合物（即，DNA 梯）的运行时间比较，可自动地将时间长度转化为片段大小。然后使用大规模并行测序对 DNA 测序文库进行测序，并将 Y 染色体序列的分数用于确定这些样品的胎儿 DNA 浓度分数。

[0108] 在图 400 中，实线 410 代表样品 UK92797，其具有的胎儿 DNA 浓度分数为 8.3%，且虚线 420 代表样品 UK94884，其具有的胎儿 DNA 浓度分数为 20.3%。相比样品 UK92797，样品 UK94884（该样品具有较高的胎儿 DNA 分数）具有相对较高量的电泳时间间隔为 63 秒 -73 秒的 DNA（区域 A）（对应于 200bp-267bp 的 DNA 大小），和相对较低量的电泳时间为 76s 的 DNA（区域 B）（对应于～ 292bp 的 DNA 大小）。

[0109] 根据厂商的方案，将总大小为 122bp 的 DNA 适配子和引物组引入血浆 DNA 用于测序文库构建。因此，区域 A 对应于约 78bp-145bp 的血浆 DNA 片段，且区域 B 对应于约 170bp 的血浆 DNA 片段。此类减除可适合用于 DNA 文库构建的不同方案。例如，在 Illumina 单端读数测序文库制备过程中，将引入来自适配子 / 引物组的 92bp 总大小，而对于标准的双端测序文库制备该大小将为 119bp。

[0110] 在另一实施方案中，血浆 DNA 可通过本领域技术人员已知的全基因组扩增系统如 Rubicon Genomics PlasmaPlex WGA 试剂盒 (www.rubicongenomics.com/products) 进行扩增。然后可通过生物分析仪分析扩增的产物。在其他实施方案中，扩增的产物可通过来自例如 Caliper(www.caliperls.com/products/labchip-systems) 的电泳系统进行分析。在其他实施方案中，可使用例如，基于纳米孔的测序仪（例如，来自 Oxford Nanopore Technologies(www.nanoporetech.com)）或 Helico DNA 测序仪 (www.helicosbio.com)，不

经过扩增直接分析血浆 DNA 的大小分布。

[0111] IV. 大小参数

[0112] 如上文所述，多种参数可提供生物样品中 DNA 片段的大小模式的统计学度量。可使用分析的所有 DNA 片段或仅一部分的大小定义参数。在一个实施方案中，参数提供了短和长 DNA 片段的相对丰度，其中所述短和长 DNA 可对应于特定的大小或大小范围。

[0113] 为了研究母体血浆 DNA 的整体大小分布是否可用于反映胎儿 DNA 浓度分数，我们使用了不同的参数来定量短和长 DNA 的相对丰度，并确定了这些参数与胎儿 DNA 浓度分数间的相关性。这些研究的结果提供在以下部分中。为了说明的目的，我们使用的用于反映短 DNA 的相对丰度的参数包括：

[0114] i. 150bp 或更小 DNA 片段的比例，其标示为 CF(大小≤ 150))。CF 是指累积频率。因此，CF(大小≤ 150) 是指小于或等于 150bp 的片段的累积频率；

[0115] ii. ≤ 150bp 的 DNA 片段与 163bp-169bp DNA 的量之比，其标示为 (CF(大小≤ 150)/ 大小 (163-169))；

[0116] iii. 140bp-146bp 的 DNA 片段与 163bp-169bp DNA 的量之比，其标示为 (大小 (140-146)/ 大小 (163-169))；

[0117] iv. 140bp-154bp 的 DNA 片段与 163bp-169bp DNA 的量之比，其标示为 (大小 (140-154)/ 大小 (163-169))；和

[0118] v. 100bp-150bp 的 DNA 片段和 163bp-169bp DNA 的量之比，其标示为 (大小 (100-150)/ 大小 (163-169))。

[0119] 参数的其他实例为柱形图的频率计。在一个实施方案中，可使用多个参数。例如，每个参数的值可给出差异百分比，然后可以确定平均百分比。在另一实施方案中，每个参数对应于多维校准函数的不同维度，其中新样品的参数值对应于相应的多维面上的坐标。

[0120] V. 大小与浓度分数的相关性

[0121] 使用测序的两个样品集被用于阐明多种大小参数与浓度分数的相关性。还提供了重复元件大小的分析。电泳数据还显示了大小参数与浓度分数间的相关性。

[0122] A. 第一样品集

[0123] 图 5A 是图 500，其显示了根据本发明实施方案，母体血浆中具有多种胎儿 DNA 百分比的样品的 150bp 或更小 DNA 片段的比例。针对 80 份母体血浆样品，将 DNA ≤ 150bp 的比例相对于胎儿 DNA 浓度分数作图。整倍体样品表示为实心圆。13 三体 (T13) 样品表示为空心三角形。18 三体 (T18) 样品表示为空心菱形，且 21 三体 (T21) 样品表示为倒空心三角形。

[0124] 对于所有样品，胎儿 DNA 浓度分数与 DNA ≤ 150bp 的比例间存在正相关性 (皮尔森相关系数＝ 0.787)。该大小参数与胎儿 DNA 浓度分数间的正相关性在具有不同胎儿染色体状态的样品间呈现为一致的。这些结果表明大小参数的分析可用于估算母体血浆样品中的胎儿 DNA 浓度分数。因此，图 5 中的数据点可用作方法 300 的校准数据点。那么，如果新样品的参数 CF(大小≤ 150) 确定为 30，则胎儿 DNA 百分比可估算为约 7％ -16％。图 5 中的数据点还可用于确定拟合所示的原始数据点的校准函数。

[0125] 图 5B 是图 550，其显示了≤ 150bp 的 DNA 片段与 163bp-169bp DNA 的量的大小比，标示为 (CF(大小≤ 150)/ 大小 (163-169))。针对 80 份母体血浆样品，将 CF(大小≤ 150)/

大小 (163-169) 比相对于胎儿 DNA 浓度分数作图。对于所有的样品，胎儿 DNA 浓度分数与 CF(大小≤ 150)/ 大小 (163-169) 比间存在正相关性 (皮尔森相关系数＝ 0.815)。该大小参数与胎儿 DNA 浓度分数间的正相关性在具有不同胎儿染色体倍性状态的样品间是一致的。

[0126] 图 6A 是图 600，其显示了 140bp-146bp 的 DNA 片段与 163bp-169bp DNA 的量的大小比，标示为 (大小 (140-146)/ 大小 (163-169))。针对 80 份母体血浆样品，将大小 (140-146)/ 大小 (163-169) 比相对于胎儿 DNA 浓度分数作图。对于所有样品，胎儿 DNA 浓度分数与大小 (140-146)/ 大小 (163-169) 比间存在正相关性 (皮尔森相关系数＝ 0.808)。该大小参数与胎儿 DNA 浓度分数之间的正相关性在具有不同胎儿染色体倍性状态的样品间是一致的。

[0127] 图 6B 是图 650，其显示了 140bp-154bp 的 DNA 片段与 163bp-169bp DNA 的量的大小比，标示为 (大小 (140-154)/ 大小 (163-169))。针对 80 份母体血浆样品，将大小 (140-154)/ 大小 (163-169) 比相对于胎儿 DNA 浓度分数作图。对于所有样品，胎儿 DNA 浓度分数与大小 (140-154)/ 大小 (163-169) 比间存在正相关性 (皮尔森相关系数＝ 0.802)。该大小参数与胎儿 DNA 浓度分数之间的正相关性在具有不同胎儿染色体倍性状态的样品间呈现为一致的。

[0128] 图 7 是图 700，其显示了 100bp-150bp 的 DNA 片段与 163bp-169bp DNA 的量的大小比，标示为 (大小 (100-150)/ 大小 (163-169))。针对 80 份母体血浆样品，将大小 (100-150)/ 大小 (163-169) 比相对于胎儿 DNA 浓度分数作图。对于所有样品，胎儿 DNA 浓度分数与大小 (100-150)/ 大小 (163-169) 比间存在正相关性 (皮尔森相关系数＝ 0.831)。该大小参数与胎儿 DNA 浓度分数之间的正相关性在具有不同胎儿染色体倍性状态的样品间是一致的。

[0129] B. 第二样品集

[0130] 图 8 是图 800，其显示了根据本发明实施方案，母体血浆中具有多种胎儿 DNA 百分比的样品的 150bp 或更小的 DNA 片段的比例。针对 48 份母体血浆样品，将 DNA ≤ 150bp 的比例相对于胎儿 DNA 浓度分数作图，所述样品在靶标富集后进行大规模并行双端测序。整倍体样品表示为实心圆。13 三体 (T13) 样品表示为空心三角形。18 三体 (T18) 样品表示为空心菱形，且 21 三体 (T21) 样品表示为倒空心三角形。对于所有的样品，胎儿 DNA 浓度分数与 DNA ≤ 150bp 的比例间存在正相关性 (皮尔森相关系数＝ 0.816)。该大小参数与胎儿 DNA 浓度分数之间的正相关性在具有不同胎儿染色体状态的样品间是一致的。这些结果表明大小参数的分析可用于估算母体血浆样品中的胎儿 DNA 浓度分数。

[0131] 图 9A 是图 900，其显示了≤ 150bp 的 DNA 片段与 163bp-169bp DNA 的量的大小比，标示为 (CF(大小≤ 150)/ 大小 (163-169))。针对 48 份母体血浆样品，将 CF(大小≤ 150)/ 大小 (163-169) 比相对于胎儿 DNA 浓度分数作图。对于所有样品，胎儿 DNA 浓度分数与 CF(大小≤ 150)/ 大小 (163-169) 比间存在正相关性 (皮尔森相关系数＝ 0.776)。该大小参数与胎儿 DNA 浓度分数之间的正相关性在具有不同胎儿染色体倍性状态的样品间是一致的。

[0132] 图 9B 是图 950，其显示了 140bp-146bp 的 DNA 片段与 163bp-169bp DNA 的量的大小比，标示为 (大小 (140-146)/ 大小 (163-169))。针对 48 份母体血浆样品，将大小

(140-146)/大小(163-169)比相对于胎儿DNA浓度分数作图。对于所有样品,胎儿DNA浓度分数与大小(140-146)/大小(163-169)比间存在正相关性(皮尔森相关系数=0.790)。该大小参数与胎儿DNA浓度分数之间的正相关性在具有不同胎儿染色体倍性状态的样品间是一致的。

[0133] 图10A是图1000,其显示了140bp-154bp的DNA片段与163bp-169bp DNA的量的大小比,标示为(大小(140-154)/大小(163-169))。针对48份母体血浆样品,将大小(140-154)/大小(163-169)比相对于胎儿DNA浓度分数作图。对于所有样品,胎儿DNA浓度分数与大小(140-154)/大小(163-169)比间存在正相关性(皮尔森相关系数=0.793)。该大小参数与胎儿DNA浓度分数之间的正相关性在具有不同胎儿染色体倍性状态的样品间是一致的。

[0134] 图10B是图1005,其显示了100bp-150bp的DNA片段与163bp-169bp DNA的量的大小比,标示为(大小(100-150)/大小(163-169))。针对48份母体血浆样品,将大小(100-150)/大小(163-169)比相对于胎儿DNA浓度分数作图。对于所有样品,胎儿DNA浓度分数与大小(100-150)/大小(163-169)比间存在正相关性(皮尔森相关系数=0.798)。该大小参数与胎儿DNA浓度分数之间的正相关性在具有不同胎儿染色体倍性状态的样品间是一致的。

[0135] C.重复

[0136] 在上文,我们已证明母体血浆中的所有可映射的DNA片段的大小与胎儿DNA浓度分数相关。在本节中,我们研究了基因组中重复元件的大小的分析是否也能够用于估算血浆中的胎儿DNA浓度分数。在本实例中,我们分析了映射至基因组的Alu重复的DNA片段的大小分布。

[0137] 图11的图显示了根据本发明实施方案,对于所示大小的重复元件,大小比相对于胎儿DNA百分比作图。该实例使用100bp-150bp的DNA片段与163bp-169bp DNA的量的比(大小(100-150)/大小(163-169))来反映大小分布相对胎儿DNA百分比的改变。该大小比与胎儿DNA浓度分数间存在正相关性(皮尔森相关系数=0.829)。该结果表明重复元件的大小分析也可用于确定母体样品中的胎儿DNA浓度分数。

[0138] 除了使用大规模并行测序,其他方法如PCR、实时PCR和质谱分析也可用于确定母体血浆中重复元件(例如,Alu重复)的大小分布。在一个实施方案中,可将母体血浆样品中的DNA连接至接头。然后,可使用一种Alu序列特异性引物和另一种接头特异性引物进行PCR。在PCR后,可分析PCR产物的大小,例如通过电泳、质谱或大规模并行测序。这将允许读取母体血浆中来源于Alu重复的序列的大小。该策略可用于其他靶标序列或序列家族。此外,PCR之后可进行巢式PCR,其涉及另一Alu特异性引物,结合相同的接头特异性引物或接头内部的巢式引物。此类巢式PCR具有的优势为能增加针对目标序列(该情况下为Alu序列)的扩增特异性。

[0139] 使用重复元件的一个优势为它们具有相对高的拷贝数,因此它们可能更易于进行分析。例如,可能能够使用较少的扩增循环。另外,具有较高的拷贝数,其分析精确度可能更高。潜在的劣势为某些类型的重复元件可能在个体间具有不同的拷贝数。

[0140] D.电泳

[0141] 图12A是电泳图1200,其根据本发明实施方案可用于确定大小比。对于所有分析

的 DNA 文库，在约 292bp 处存在尖锋，其后的第二个峰范围为 300bp-400bp。由于大小范围曲线下的面积可以代表来自该区域的 DNA 片段的相对量，我们使用区域 A(200bp-267bp) 与 B(290bp-294bp) 的面积比定量短和长 DNA 片段的相对丰度。我们首先手动地将荧光单位 (FU) 的基线调整至 0，然后生成所选区域的面积。

[0142] 图 12B 是图 1250，其显示了根据本发明实施方案，母体血浆中具有多种胎儿 DNA 百分比的样品的 200bp-267bp 的 DNA 片段与 290bp-294bp DNA 的量的大小比（即，电泳图上显示的区域 A 和 B 的面积比）。存在一个 T13 个例，其显示低的 292-bp 峰，FU 值为 6.1，而所有其他个例均显示 FU 值≥ 20FU。由于低 FU 值会使得面积测量不精确，在分析中忽略了该个例。针对所有其他 79 份母体血浆样品，将区域 A 和 B 的面积比相对于胎儿 DNA 浓度分数作图。对于这些样品，胎儿 DNA 浓度分数与面积 A 和 B 的比间存在正相关性（皮尔森相关系数＝ 0.723)。

[0143] VI. 确定校准数据点

[0144] 如上文所述，校准数据点可以多种方式定义。另外，校准数据点可以多种方式获得。例如，可简单地从存储器将校准数据点读取为参数以及相应的浓度分数的一系列的校准值。同样，可从存储器（例如，具有预定函数形式的线性或非线性函数）读取校准函数，其中所述函数定义校准数据点。在一些实施方案中，可从校准样品测得的数据计算校准数据点。

[0145] A. 方法

[0146] 图 13 是根据本发明实施方案，由校准样品产生的测量结果确定校准数据点的方法 1300 的流程图。校准样品包含临床相关的 DNA 和其他 DNA。

[0147] 在方框 1310 中，接收了多个校准样品。校准样品可按本文所述获得。可通过单独的实验或通过一些确定分子来自哪个样品的鉴定方法（例如，为 DNA 片段标上条码），单独对每个样品进行分析。例如，可将校准样品接收于仪器，例如，测序仪，其输出可用于确定 DNA 片段大小的测量数据（例如，序列读数），或将其接收于电泳仪。

[0148] 在方框 1320 中，测量了多个校准样品中的每个的临床相关 DNA 的浓度分数。在测量胎儿 DNA 浓度的多个实施方案中，可使用父本遗传的序列或胎儿特异性表观遗传标志物。例如，父本遗传的等位基因不存在于孕妇的基因组中，且能够以与胎儿 DNA 浓度分数成比例的百分比在母体血浆中检测到。胎儿特异性表观遗传标志物可包括母体血浆中展现胎儿或胎盘特异性 DNA 甲基化模式的 DNA 序列。

[0149] 在方框 1330 中，测量来自每个校准样品的多种大小的 DNA 片段的量。大小可按本文所述的进行测量。可对大小计数、作图，用于制作柱形图或其他分选程序，以获得有关校准样品大小模式的数据。

[0150] 在方框 1340 中，基于多种大小的 DNA 片段的量，计算参数的校准值。可计算每个校准样品的校准值。在一个实施方案中，将相同的参数用于每个校准值。然而，实施方案可使用本文所述的多个参数。例如，小于 150 个碱基的 DNA 片段的累积分数可用作参数，且具有不同浓度分数的样品可能具有不同的校准值。可以确定每个样品的校准数据点，其中所述校准数据点包括校准值和测得的样品的浓度分数。这些校准数据点可用于方法 300，或可用于确定最终的校准数据点（例如，如通过函数拟合定义的）。

[0151] 在方框 1350 中，确定了近似于多个浓度分数间的校准值的函数。例如，可将线性

函数拟合至校准值，作为浓度分数的函数。线性函数可定义将在方法 300 中使用的校准数据点。

[0152] 在一些实施方案中，可以计算每个样品的多个参数的校准值。样品的校准值可定义多维坐标（其中每个维度用于每个参数），其联合浓度分数可提供数据点。因此，在一个实施方案中，可将多维函数拟合至所有的多维数据点。因此，可使用多维校准曲线，其中不同的参数值可被有效地输入至能够输出浓度分数的单个校准函数。并且，单个校准函数可产生自所有获自校准样品的数据点的函数拟合。

[0153] B. 测量肿瘤 DNA 浓度

[0154] 如同所述的，实施方案还可应用于生物样品中肿瘤 DNA 的浓度。接下来为涉及确定肿瘤 DNA 的浓度分数的实例。

[0155] 我们从肿瘤手术切除之前和之后的两名患有肝细胞癌 (HCC) 的患者采集了血浆样品。使用双端 (PE) 大规模并行测序进行大小分析。按先前所述构建母体血浆 DNA 的测序文库 (Lo YM et al. Sci Transl Med 2010 ;2:61ra91)。所有的文库均通过 HiSeq 2000(Illumina) 使用 50-bp×2PE 格式测序。使用短核苷酸比对程序 2(SOAP2)(soap.genomics.org.cn)，将 50-bp 序列读数与非重复掩蔽的人参照基因组 (Hg18)(http://genome.ucsc.edu) 进行比对。每个测序片段的大小自比对片段每端的最外面的核苷酸的坐标推导而来。

[0156] 我们使用 Affymetrix SNP6.0 微阵列系统，对提取自 HCC 患者的血细胞和肿瘤样品的 DNA 的进行基因型分型。对于每个个例，使用 Affymetrix Genotyping Console v4.0，基于 SNP 基因座的不同等位基因的强度，确定了肿瘤组织中展现杂合性丢失 (LOH) 的区域。使用下式从 LOH 区域中携带缺失的和未缺失的等位基因的序列的量的差异，估算了肿瘤来源 DNA 的浓度分数 (F) :F = (A - B)/A x 100%，其中 A 是 LOH 区域中携带杂合 SNP 的未缺失等位基因的序列读数的数量，并且 B 是 LOH 区域中携带杂合 SNP 的缺失等位基因的序列读数的数量。表 3 显示了其结果。

[0157]

个例编号	取样时间	测序读数的数量	血浆中肿瘤 DNA 的浓度分数(%)
1	肿瘤切除前	448M	51.60
	肿瘤切除后	486M	0.90
2	肿瘤切除前	479M	5.60
	肿瘤切除后	542M	0.90

[0158] 表 3 显示了血浆样品中肿瘤 DNA 的测序信息和测得的浓度分数。

[0159] 在另一实施方案中，可使用展现重复的基因座。例如，肿瘤可展现两条同源染色体中的一条增加一个拷贝，使得等位基因重复。然后，可以确定一个或多个杂合基因座（例如，SNP）处具有非重复等位基因的序列读数的第一量 A，以及杂合基因座处具有重复等位基因的序列读数的第二量 B。可使用比值 (B - A)/A，将临床相关 DNA 的浓度分数 F 计算为第一量与第二量之比。

[0160] 在另一实施方案中，可使用一个或多个纯合基因座。例如，可以确定一个或多个这样的基因座 :其中患者是纯合的，并且其中肿瘤组织中存在单核苷酸突变。然后，可以确定

一个或多个纯合基因座处具有野生型等位基因的序列读数的第一量A。并且，可以确定一个或多个纯合基因座处具有突变等位基因的序列读数的第二量B。可以使用比例2B/(A+B)，将临床相关DNA的浓度分数F计算为第一量与第二量之比。C. 数据点的函数拟合的实例

[0161] 现在描述进行从校准样品确定的参数值的函数拟合的实例。分析了来自80名孕妇（每名怀有一个男性胎儿）的血浆样品。在这80名孕妇中，39名怀有整倍体胎儿，13名怀有13三体(T13)胎儿，10名怀有18三体(T18)胎儿，且18名怀有21三体(T21)胎儿。这些孕妇的中值孕龄为13周加1天。从血浆样品提取DNA，并使用Illumina HiSeq2000平台按照所述的测序(Zheng YW et al.Clin Chem.2012;58:549-58)，除了测序以8倍格式进行。对于每个DNA分子，从两端中的每端测序50个核苷酸，并与参照基因组(hg18)比对。

[0162] 然后从两端最外面核苷酸的坐标推导每个测序分子的大小。对于每个样品，测序了中值为1.11千万的片段，并唯一地与参照基因组对齐。通过将100bp-150bp大小的DNA分子的比例除以163bp-169bp大小的DNA分子的比例，计算比值，并且该比值称为大小比。由于所有80名孕妇均怀有男性胎儿，将与Y染色体唯一对齐的序列读数的比例用于确定每个血浆DNA样品中胎儿DNA的浓度分数。

[0163] 将样品随机分成2组，即训练组和验证组。使用线性回归，基于训练组中的样品，确立胎儿DNA浓度分数与大小比间的关系。然后，使用线性回归公式，将大小比用于推导验证组的胎儿DNA浓度分数。验证在下一节中论述。

[0164] 图14A是根据本发明实施方案，针对训练组，大小比相对于胎儿DNA浓度分数的图1400。如上文所述，通过将100bp-150bp大小的DNA分子的比例除以163bp-169bp大小的DNA分子的比例，计算大小比。将大小比针对胎儿DNA的浓度分数作图，如数据点1405所示。空心圆代表整倍体个例。实心符号代表非整倍性个例（正方形为T13，圆为T18，且三角形为T21)。线性回归线1410由对数据点的函数拟合来产生。可通过任何合适的技术，例如，最小二乘法，进行函数拟合。线1410可用于估算测量的其他样品而非训练组样品的参数的值。线1410的每一部分可被认为是一个校准数据点。

[0165] VII. 与校准数据点的比较

[0166] 如上文所述，校准数据点可用于确定临床相关DNA的浓度分数。例如，图14A中的原始数据点1405可用于提供特定校准值的DNA浓度分数（图14A中标记为大小比）的范围，其中所述范围可用于确定浓度分数是否大于阈值量。代替范围，可使用特定大小比的浓度分数平均值。例如，可将对应于新样品的1.3的大小比的测量结果的浓度分数，确定为从1.3的两个数据点计算的平均浓度。在一个实施方案中，可使用函数拟合（例如，线1410)。

[0167] 图14B是根据本发明实施方案，从图14A的线性函数1410推导（估算）的浓度分数相对于使用胎儿特异性序列测得的浓度分数的图1450。使用基于训练组的数据确定的回归方程（即，线1410)，将确定的验证样品的大小比用于推导验证组样品的胎儿DNA浓度分数。测量的浓度分数对应于血浆DNA样品中Y染色体序列的比例（即，对齐Y染色体的序列读数的比例）。

[0168] 线1460代表了两组值间的完全相关。数据点1455的偏差表示估算的精确度，线1460上的点是完全精确的。如本文所述，估算不必为完全精确的，因为所需的检测可能仅是为了确定生物样品中是否存在足够百分比的临床相关DNA。空心圆代表整倍体个例。实心符号代表非整倍性个例（正方形为T13，圆为T18，且三角形为T21)。从大小比推导的胎儿

DNA 浓度分数与从 Y 染色体序列的比例测得的浓度分数间的中值差异为 2.1%。在 90%的样品中，差异小于 4.9%。

[0169] 将具有不同倍性状态的样品用于校准组和验证组。如图 14A 中所示，大小比与胎儿 DNA 浓度分数间的关系在具有不同倍性状态的样品间是一致的。因此，可在没有关于样品倍性状态的先前了解的情况下，从样品的大小比推导出胎儿 DNA 浓度分数，如图 14B 所示。一条校准曲线可用于具有不同倍性状态的样品，因此，我们在使用实施方案确定胎儿 DNA 浓度分数前不需要知道样品的倍性状态。

[0170] VIII. 癌症

[0171] 如本文所述，实施方案可用于评估生物样品中肿瘤 DNA 的浓度分数。如同胎儿实例，校准样品可用于确定相关性数据点，例如，通过将函数（例如，线性函数）拟合至显示大小参数值与测得的浓度分数间的相关性的数据点。

[0172] A. 大小与肿瘤 DNA 浓度的相关性

[0173] 图 15A 是图 1500，其显示了根据本发明实施方案，肿瘤切除之前和之后的两名 HCC 患者的血浆中具有多种肿瘤 DNA 百分比的样品的 150bp 或更小的 DNA 片段的比例。针对肿瘤切除之前（实心圆）和之后（空心圆）的两名 HCC 患者，将 DNA ≤ 150bp 的比例相对于肿瘤 DNA 浓度分数作图。两个空心圆在位置上彼此非常接近（实际上在彼此上面）。这些结果表明大小参数的分析可用于估算 HCC 患者血浆样品中的肿瘤 DNA 浓度分数。肿瘤切除后肿瘤 DNA 浓度分数和≤ 150bp DNA 片段的比例均下降。实心圆 1505 对应于具有低得多的肿瘤 DNA 百分比（其与较小的肿瘤大小相关）的样品。换句话说，具有较大肿瘤的患者具有较高比例的短 DNA，其体现在相比具有较小肿瘤的患者的较高比值的 CF（≤ 150bp）。

[0174] 图 15B 是图 1550，其显示了肿瘤切除之前和之后的两名 HCC 患者的≤ 150bp 的 DNA 片段与 163bp-169bp DNA 的量的大小比，标示为 (CF（大小≤ 150）/ 大小 (163-169))。针对肿瘤切除之前（实心圆）和之后（空心圆）的两名 HCC 患者，将 CF（大小≤ 150）/ 大小 (163-169) 比相对于肿瘤 DNA 浓度分数作图。两个空心圆在位置上彼此非常接近。肿瘤切除后肿瘤 DNA 浓度分数和大小比均下降。

[0175] 图 16A 是图 1600，其显示了肿瘤切除之前和之后的两名 HCC 患者的 140bp-146bp 的 DNA 片段与 163bp-169bp DNA 的量的大小比，标示为（大小 (140-146)/ 大小 (163-169))。针对肿瘤切除之前（实心圆）和之后（空心圆）的两名 HCC 患者，将大小 (140-146)/ 大小 (163-169) 比相对于肿瘤 DNA 浓度分数作图。肿瘤切除后肿瘤 DNA 浓度分数和大小比均下降。

[0176] 图 16B 是图 1650，其显示了肿瘤切除之前和之后的两名 HCC 患者的 140bp-154bp 的 DNA 片段与 163bp-169bp 的 DNA 的量的大小比，标示为（大小 (140-154)/ 大小 (163-169))。针对肿瘤切除之前（实心圆）和之后（空心圆）的两名 HCC 患者，将大小 (140-154)/ 大小 (163-169) 比相对于肿瘤 DNA 浓度分数作图。肿瘤切除后肿瘤 DNA 浓度分数与大小比均下降。

[0177] 图 17 是图 1700，其显示了肿瘤切除之前和之后的两名 HCC 患者的 100bp-150bp 的 DNA 片段与 163bp-169bp DNA 的量的大小比，标示为（大小 (100-150)/ 大小 (163-169))。针对肿瘤切除之前（实心圆）和之后（空心圆）的两名 HCC 患者，将大小 (100-150)/ 大小 (163-169) 比相对于肿瘤 DNA 浓度分数作图。肿瘤切除后肿瘤 DNA 浓度分数与大小比均下

降。

[0178] B. 治疗引起的大小下降

[0179] 图 18A 是图 1800，其显示了肿瘤切除之前和之后的 HCC 患者的 150bp 或更小的 DNA 片段的比例。来自同一癌症患者的一对样品表示为通过虚线连接的相同的符号。肿瘤切除后的癌症患者中血浆 DNA 的 DNA ≤ 150bp 的比例整体下降。

[0180] 治疗前和治疗后比例值的间隔证明了肿瘤存在与大小参数值间的相关性。治疗前和治疗后的值的间隔可用于确定治疗的成功度，例如，通过将该比例与阈值比较，其中低于阈值的比例可指示成功。在另一个实例中，可将治疗前和治疗后间的差异与阈值比较。

[0181] 还可将比例（或大小参数的任何其他值）用于检测肿瘤的出现。例如，可以确定大小参数的基线值。然后，在晚些时候，可再次测量大小参数的值。如果大小参数的值显示显著改变，则患者可能存在患有肿瘤的较高风险。如果大小参数的值在个体间变化不大（图 18A 表明比例变化不大（即，因为治疗后的值是相同的）），则可将相同的基线值用于其他患者。因此，不需要采集每名患者的基线值。

[0182] 图 18B 是图 1850，其显示了肿瘤切除之前和之后的 HCC 患者的≤ 150bp 的 DNA 片段与 163bp-169bp DNA 的量的大小比，标示为 (CF(大小≤ 150)/ 大小 (163-169))。来自同一癌症患者的一对样品表示为通过虚线连接的相同的符号。肿瘤切除后两例的该大小比均下降。

[0183] 图 19A 是图 1900，其显示了肿瘤切除之前和之后的 HCC 患者的 140bp-146bp 的 DNA 片段与 163bp-169bp DNA 的量的大小比，标示为（大小 (140-146)/ 大小 (163-169))。来自同一癌症患者的一对样品表示为通过虚线连接的相同的符号。肿瘤切除后两例的该大小比均下降。

[0184] 图 19B 是图 1950，其显示了肿瘤切除之前和之后的 HCC 患者的 140bp-154bp 的 DNA 片段与 163bp-169bp DNA 的量的大小比，标示为（大小 (140-154)/ 大小 (163-169))。来自同一癌症患者的一对样品表示为通过虚线连接的相同的符号。肿瘤切除后两例的该大小比均下降。

[0185] 图 20 是图 2000，其显示了肿瘤切除之前和之后的 HCC 患者的 100bp-150bp 的 DNA 片段与 163bp-169bp DNA 的量的大小比，标示为（大小 (100-150)/ 大小 (163-169))。来自同一癌症患者一对样品的表示为通过虚线连接的相同的符号。肿瘤切除后两例的该大小比均下降。

[0186] C. 方法

[0187] 图 21 是流程图，其阐示了根据本发明实施方案，分析生物体的生物样品以确定癌症等级分级的方法 2100。方法 2100 可分析生物体（例如，人）的生物样品。生物样品包含源自正常细胞和可能来自癌症相关细胞的 DNA。所述 DNA 中的至少一些为生物样品中无细胞的。方法 300 和 1300 的方面可与方法 2100 的实施方案一起使用。

[0188] 在方框 2110 中，测量对应于多种大小的 DNA 片段的量。对于多种大小中的每种大小，可以测量来自生物样品的对应于所述大小的多个 DNA 片段的量，如方法 300 所述。所述多个 DNA 片段可随机挑选，或优选地从基因组的一个或多个预定区域挑选。例如，可进行靶向富集，或者可使用来自基因组的特定区域的序列读数的选择，例如，如上文所述的。

[0189] 在方框 2120 中，基于多种大小的 DNA 片段的量，计算第一参数的第一值。在一个

方面，第一参数提供了生物样品中 DNA 片段的大小模式的统计学度量（例如，柱形图）。参数可称为大小参数，因为其从多个 DNA 片段的大小确定而来。本文提供了参数的实例。可使用多个参数，其在本文中也有描述。

[0190] 在方框 2130 中，将第一值与参考值比较。参考值的实例包括正常值和与正常值具有指定距离（例如，采用标准偏差的单位）的截止值。可从来自相同生物体（例如，当已知该生物体健康时）的不同样品确定参考值。因此，参考值可对应于当推测生物体未患癌症时从样品确定的第一参数的值。在一个实施方案中，生物样品获自治疗后的生物体，且参考值对应于从治疗前采集的样品确定的第一参数的值（例如，如上所述的）。还可从其他健康生物体的样品确定参考值。

[0191] 在方框 2140 中，基于所述比较确定生物体中的癌症等级的分级。在多个实施方案中，分级可为数值的、文本的或任何其他指标。分级可提供关于癌症的是或否的二元结果、可能性或其他得分，其可为绝对值或相对值，例如，相对于早些时候生物体的先前的分级。在一个实施方案中，分级为生物体未患癌症，或者癌症等级降低。在另一实施方案中，分级为生物体确实患有癌症，或癌症等级增加。

[0192] 如本文所述，癌症等级可包括癌症存在、癌症阶段或肿瘤大小。例如，第一值超标（例如，大于或小于，取决于第一参数如何定义）是否可用于确定是否存在癌症，或至少一种可能性（例如，百分比可能性）。超过阈值的程度可提供增加的可能性，其可导致使用多个阈值。另外，超过的程度可对应于不同的癌症等级，例如，更多的肿瘤或更大的肿瘤。因此，实施方案可诊断、定级、预测或监测生物体中癌症等级的进展。

[0193] D. 确定特定区域的大小分布

[0194] 如同其他实施方案，第一组 DNA 片段可对应于生物体基因组的一个或多个预定区域。因此，还可对选定区域进行大小分析，例如，特定染色体、染色体臂或相同长度例如 1Mb 的多个区域（单元格）。例如，可以聚焦于目标癌症类型中通常发生改变的区域。图 22 的表 2200 显示了可见于多种类型的癌症中的一些常见的染色体畸变。获得是指特定节段中具有一个或多个另外拷贝的染色体的扩增，丢失是指特定节段中一个或两个同源染色体的缺失。

[0195] 在一个实施方案中，可从生物样品确定其他组的 DNA 片段。每组 DNA 片段可对应于不同的预定区域，如表 2200 中所示的区域。还可使用与癌症不相关的区域，例如，用以确定参考值。可以确定对应于多种大小的 DNA 片段的量，并且可以确定每个其他组 DNA 片段的参数的大小值，如本文所述。因此，可以确定每个基因组区域的不同的大小值，其中一组 DNA 片段与基因组区域间存在一一对应的关系。

[0196] 可将每个大小值与各自的参考值比较。可以确定这样的预定区域：其中对应的大小值相比各自的参考值存在统计学差异。当参考值为正常值时，可通过比较大小值与截止值（例如，其中截止值为基于假定的或测量的统计学分布，与正常值的标准偏差的具体数字），进行统计学差异的确定。不同区域各自的参考值可以是相同的或不同的。例如，不同的区域可具有不同的大小正常值。

[0197] 在一个实施方案中，相比参考值具有统计学差异的区域数目可用于确定分级。因此，可以确定鉴定其中相应的大小值相比各自的参考值具有统计学差异的预定区域的数目。可将该数目与区域的阈值数比较来确定生物体中癌症等级的分级。可以基于正常样品

中和癌症样品中的方差确定阈值数。

[0198] 如表 2200 中所突出显示的，不同的癌症与基因组的不同的部分相关。因此，当可能的癌症类型与鉴定的区域相关时，可将具有统计学差异的区域用于确定一种或多种可能的癌症类型。例如，如果发现来自染色体节段 7p 的 DNA 片段的大小值显著低于正常值（例如，如通过截止值所确定的），当分级表明存在癌症时，则将结肠直肠癌确定为可能的癌症。应当注意染色体节段 7p 的大小值可用作确定分级的唯一指标，或可使用多个区域。在一个实施方案中，仅在整体分级指示癌症时，才会将染色体节段 7p 的大小值用于确定结肠直肠癌为可能的癌症。

[0199] IX. 计算机系统

[0200] 本文提及的任何计算机系统可以使用任何合适数量的子系统。此类子系统的实例示于图 23 的计算机设备 2300 中。在一些实施方案中，计算机系统包括单一的计算机设备，其中子系统可以为计算机设备的部件。在其他的实施方案中，计算机系统可以包括多个具有内部部件的计算机设备（每个为子系统）。

[0201] 图 23 所示的子系统通过系统总线 2375 相互连接。示出与显示适配器 2382 耦合的其他子系统，例如打印机 2374、键盘 2378、硬盘 2379、监控器 2376 等。外部和输入 / 输出 (I/O) 装置（其与 I/O 控制器 2371 耦合）可以通过本领域已知的任意数量的手段（例如串行端口 2377）与计算机系统连接。例如，串行端口 2377 或外部界面 2381（例如 Ethernet、Wi-Fi 等）可以用于将计算机系统 2300 与诸如 Internet 的广域网路、滑鼠输入装置或扫描器连接。通过系统总线 2375 的相互连接允许中央处理器 2373 与各个子系统连通，并控制来自系统存储器 2372 或硬盘 2379 的指令的执行，以及子系统之间的信息交换。系统存储器 2372 和 / 或硬盘 2379 可以表现为计算机可读介质。本文提及的任何值都可以由一个部件输出至另一个部件，并可以输出给用户。

[0202] 计算机系统可以包括例如通过外部界面 2381 或通过内部界面连接在一起的多个相同的部件或子系统。在一些实施方案中，计算机系统、子系统或设备可以在网络上连通。在这种情况下，一个计算机可以被认为是客户端，而另一个计算机为服务器，其中它们均可以为同一计算机系统的一部分。客户端和服务器均可以包括多个系统、子系统或部件。

[0203] 应该理解的是，本发明的任意实施方案均可以使用硬件（例如，专用集成电路或现场可编程门阵列）以控制逻辑的形式实施和 / 或使用具有通用的可编程处理器的计算机软件以模块或集成的方式实施。基于本文所提供的公开和教导，本领域的普通技术人员将了解并领会使用硬件以及硬件和软件的组合来实施本发明的实施方案的其他方式和 / 或方法。

[0204] 本申请中所描述的任意软件部件或功能都可以作为使用例如传统或面向对象的技术、使用任何合适的计算机语言（例如 Java, C++ 或 Perl）通过处理器执行的软件代码来实施。软件编码可以在用于储存和 / 或传输的计算机可读介质上以一系列指令或命令的形式储存，合适的介质包括随机存取存储器 (RAM)、只读存储器 (ROM)、磁介质（例如硬盘驱动器或软盘）、或光学介质（例如光盘 (CD) 或 DVD（数字通用光盘））、闪存等。计算机可读介质可以为此类存储或传输装置的任意组合。

[0205] 此类程序还可以使用适用于通过有线、光学和 / 或无线网络（遵守多种协议，包括 Internet）传输的载波信号来编码和传输。由此，可以使用由此类程序编码的数据信号来创

建根据本发明的实施方案的计算机可读介质。使用程序代码所编码的计算机可读介质可以使用相容的装置包装，或者与其他装置分开提供（例如通过 Internet 下载）。任意此类的计算机可读介质可以属于或者在单一的计算机程序产品（例如硬盘驱动器、CD 或整个计算机系统）内，并且可以在系统或网络内的不同计算机程序产品上或该产品内存在。计算机系统可以包括监控器、打印机或用于向用户提供本文所提及的任意结果的其他合适的显示器。

[0206] 本文所述的任何方法可以使用计算机系统（包括一个或多个处理器）全部或部分实施，所述系统可以被配置成实施多个步骤。因此，多个实施方案可以涉及计算机系统，该系统被配置成潜在地使用实施各个步骤或各组步骤的不同部件来实施本文所述的任意方法的步骤。尽管本文所述的方法步骤以编号的步骤呈现，但是这些方法的步骤可以同时或以不同的次序实施。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，步骤的全部或一部分可以是任选的。此外，任意方法的任意步骤可以使用模块、电路或用于实施这些步骤的其他手段来实施。

[0207] 在不脱离本发明实施方案的精神和范围的条件下，特定实施方案的具体详情可以以任何合适的方式组合。但是，本发明的其他实施方案可以涉及与各单一的方面或这些单一方面的特定组合相关的特定实施方案。

[0208] 为了说明和描述的目的，已经呈现本发明的示例性实施方案的上述描述。无意于穷举或将本发明限定于所述的精确形式，并且根据上文的教导，许多修改和变化也是可能的。选择并描述多个实施方案，以便最好地说明本发明的原理及其实践应用，由此使本领域的技术人员能够在多个实施方案以及各种修改（其适用于所考虑的特定用途）中最好地利用本发明。

[0209] 除非明确作出相反的说明，描述“一个 (a)”、“一 (an)”或“所述 (the)”意指“一个或多个”。

[0210] 上文提及的所有专利、专利申请、出版物和描述在此以引用方式全文并入本文，用于所有的目的。这些文件均未确认为现有技术。

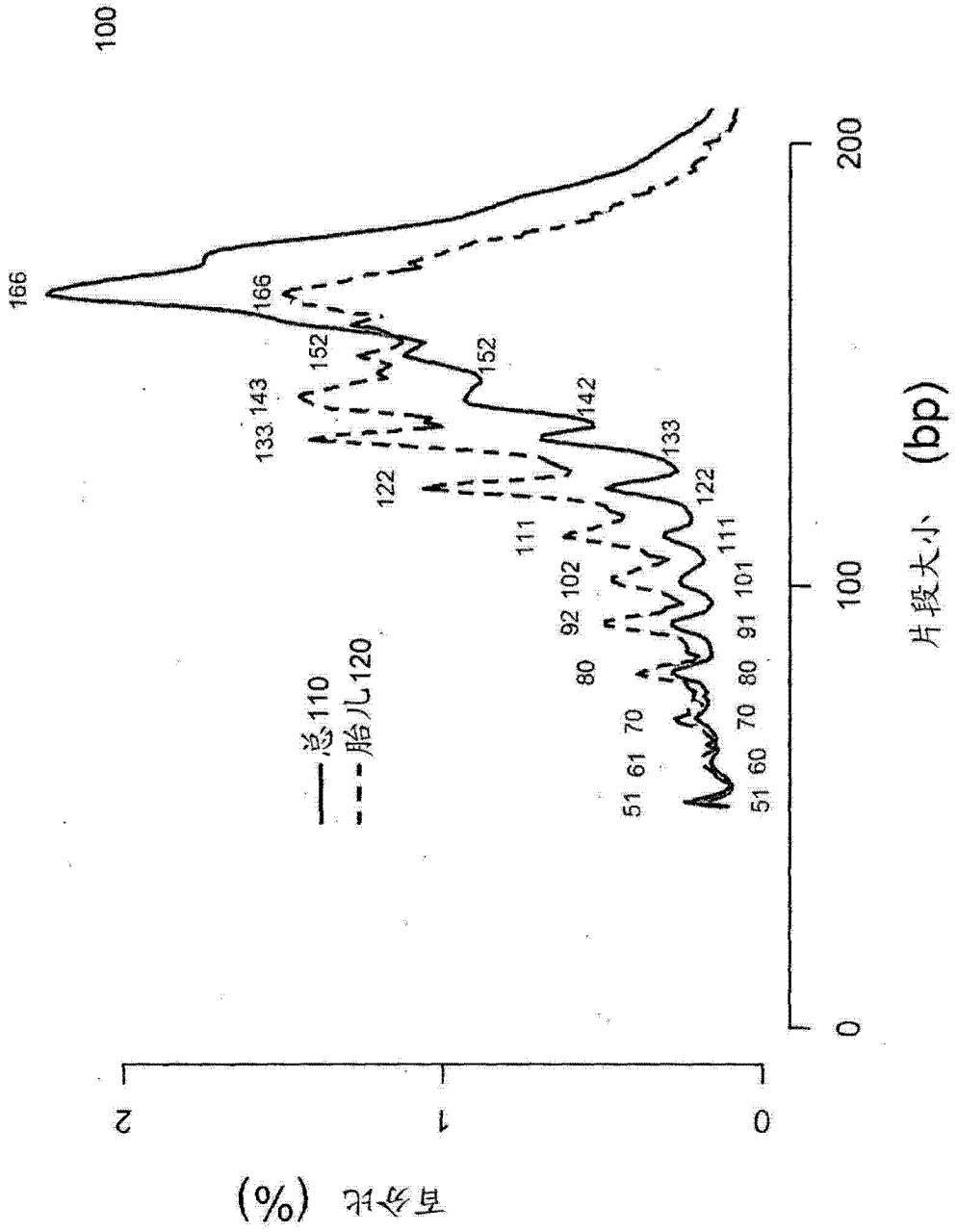

100
166
166
152
152
133 143
142
122
133
111
122
92 102
111
80
101
91
80
70
61 70
60 70
51
51
总110
胎儿120
0
100
200
片段大小 (bp)
0
1
2
百分比 (%)

图 1

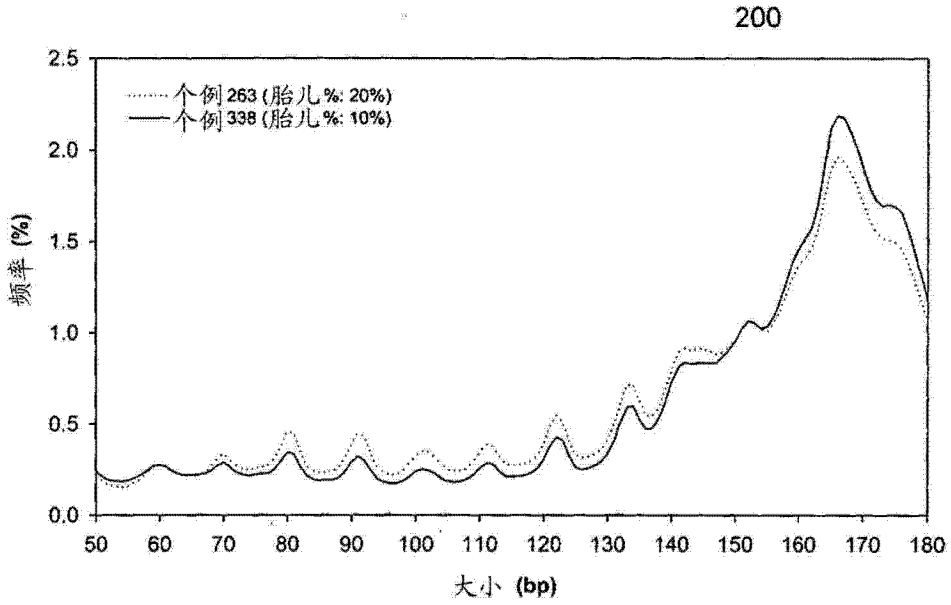

200
个例 263 (胎儿%: 20%)
个例 338 (胎儿%: 10%)
频率 (%)
大小 (bp)

图 2A

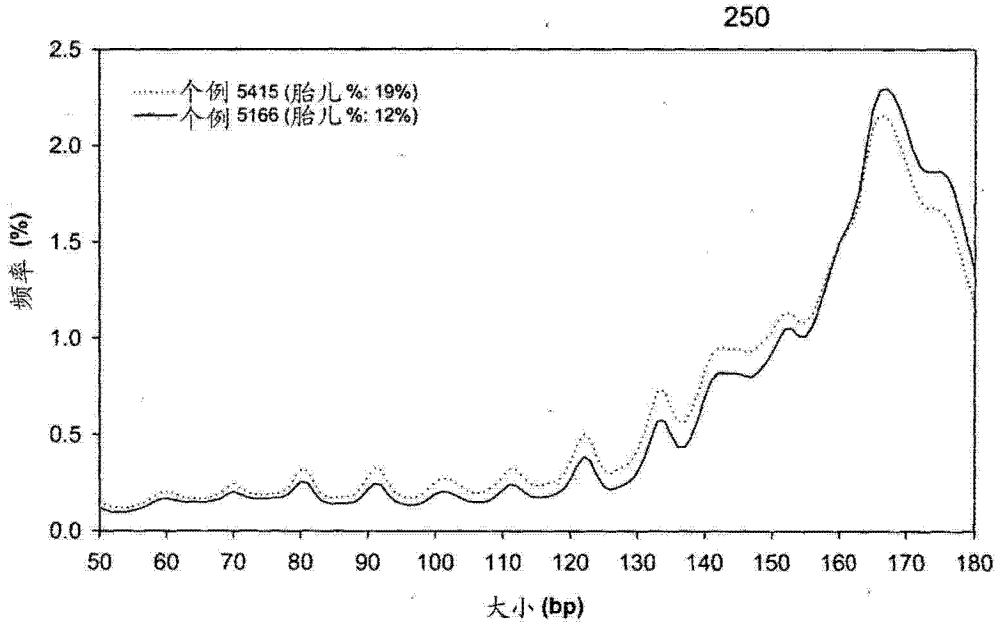

250
个例 5415 (胎儿%: 19%)
个例 5166 (胎儿%: 12%)
频率 (%)
大小 (bp)

图 2B

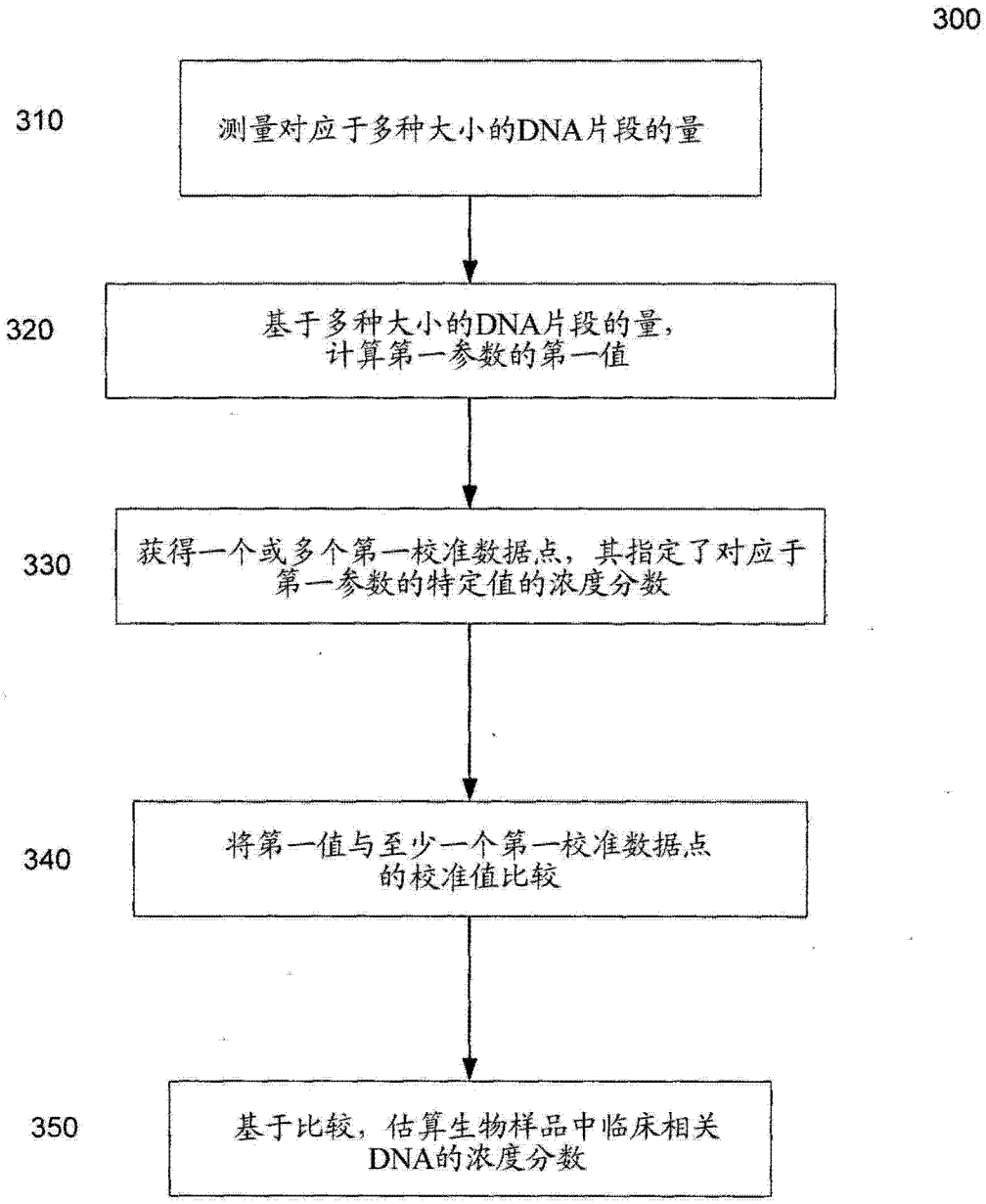

300
310
测量对应于多种大小的DNA片段的量
320
基于多种大小的DNA片段的量，
计算第一参数的第一值
330
获得一个或多个第一校准数据点，其指定了对应于
第一参数的特定值的浓度分数
340
将第一值与至少一个第一校准数据点
的校准值比较
350
基于比较，估算生物样品中临床相关
DNA的浓度分数

图 3

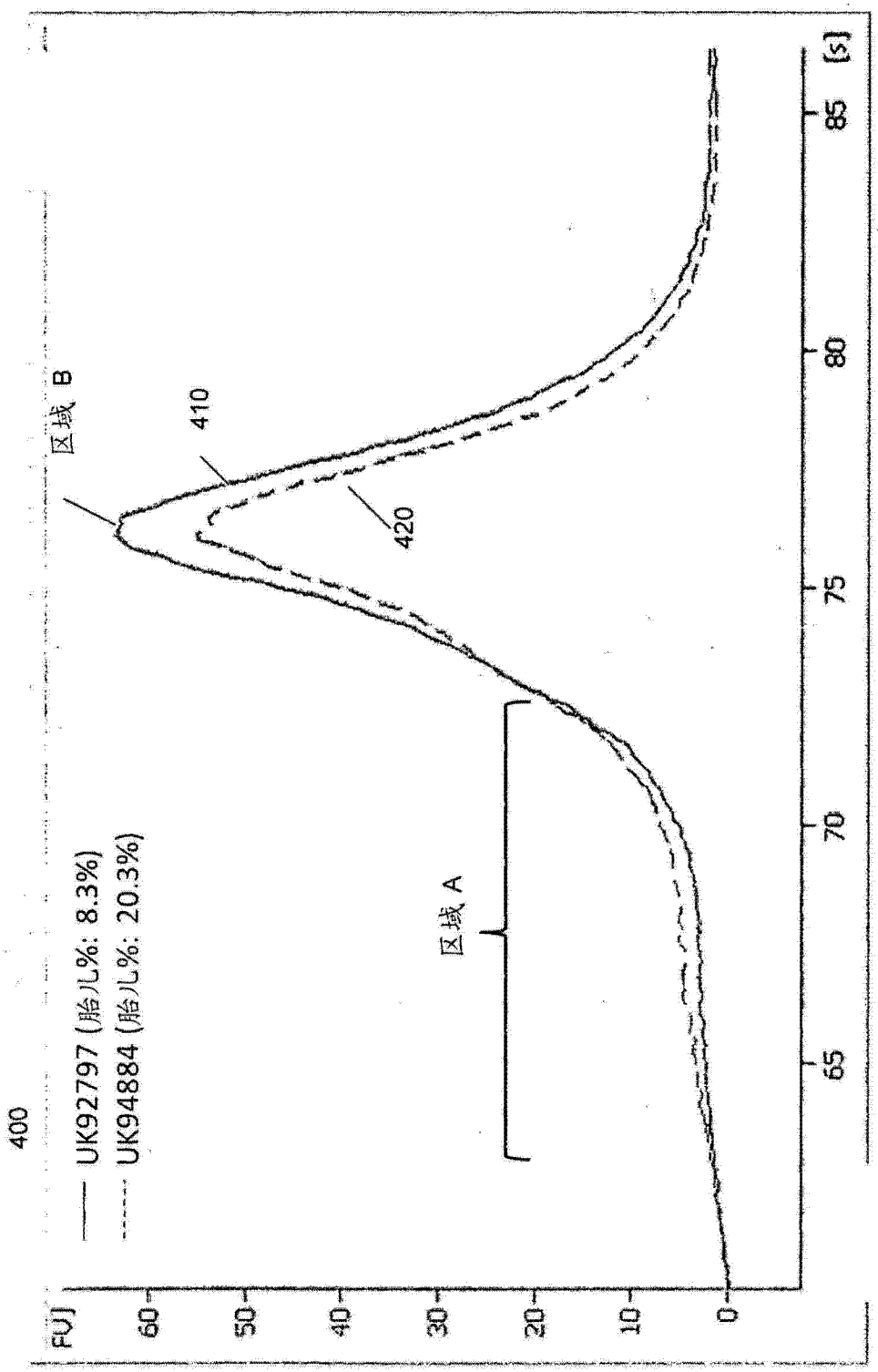

400
UK92797 (胎儿%: 8.3%)
UK94884 (胎儿%: 20.3%)
FVJ
60
50
40
30
20
10
0
区域 A
区域 B
410
420
65
70
75
80
85
[s]

图 4

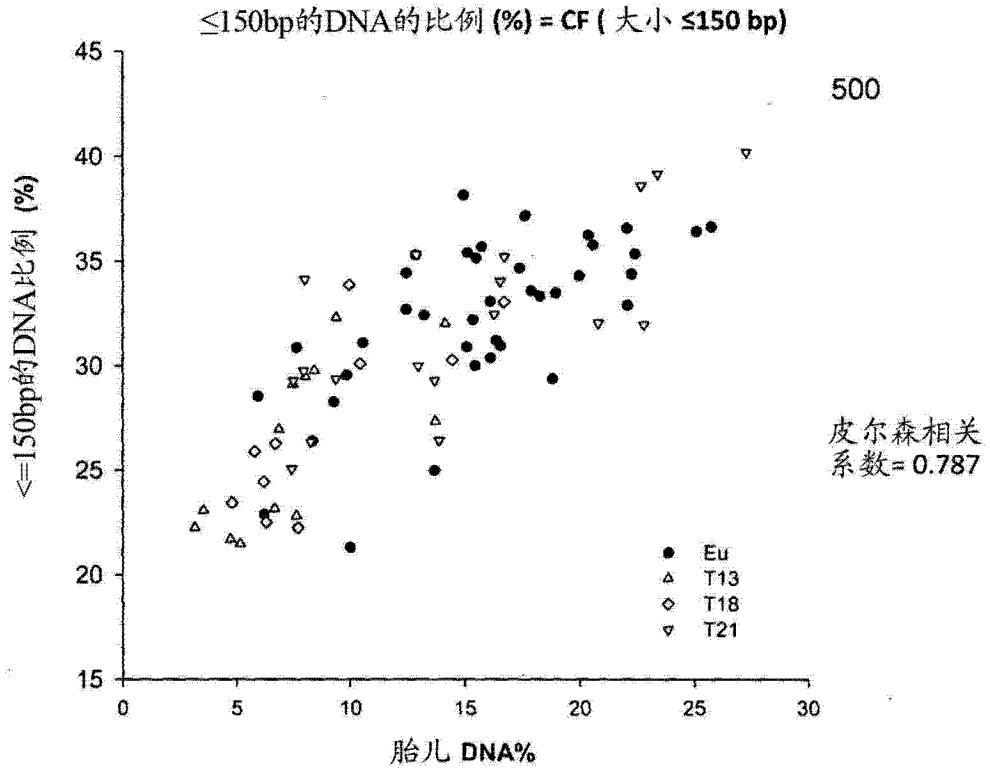

≤150bp的DNA的比例 (%) = CF (大小 ≤150 bp)
500
<=150bp的DNA比例 (%)
皮尔森相关
系数= 0.787
Eu
T13
T18
T21
胎儿 DNA%

图 5A

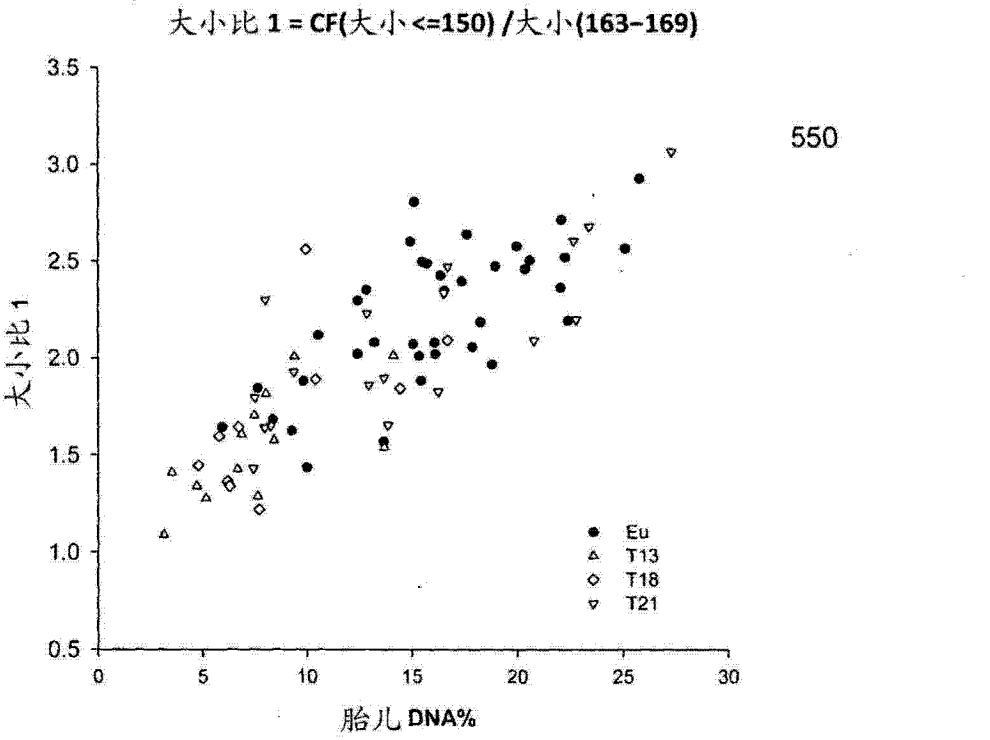

大小比 1 = CF(大小 <=150) /大小(163-169)
550
大小比 1
Eu
T13
T18
T21
胎儿 DNA%

图 5B

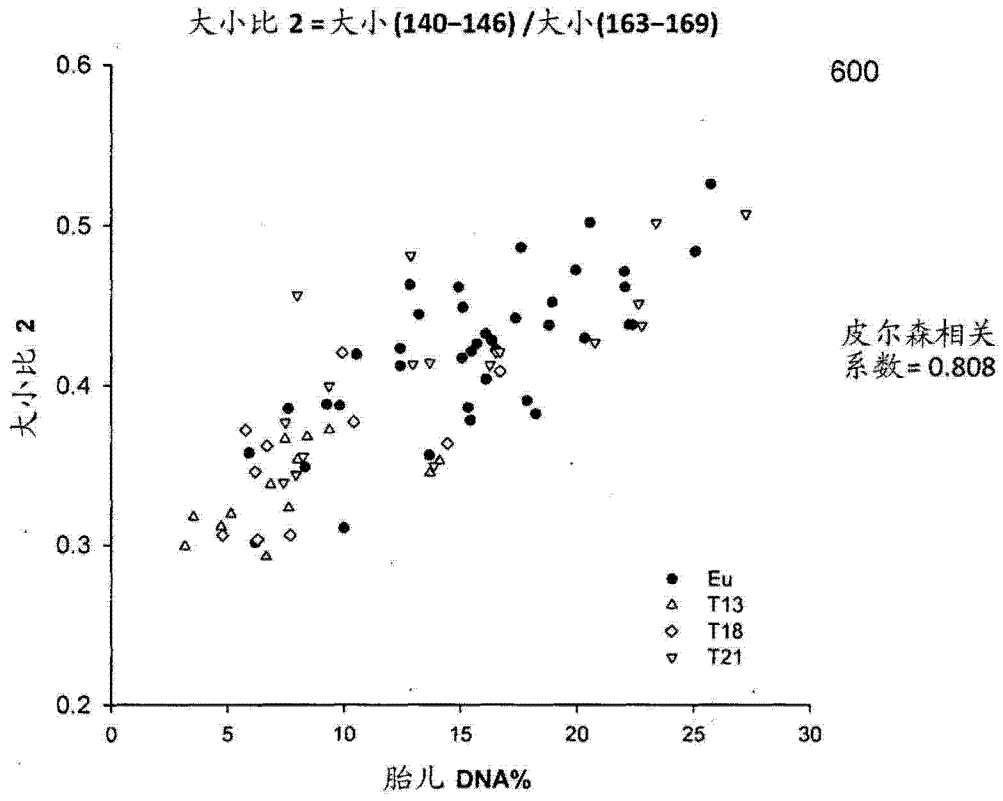

大小比 2 = 大小 (140–146) / 大小 (163–169)
600
皮尔森相关
系数 = 0.808
大小比 2
胎儿 DNA%
Eu
T13
T18
T21
0.6
0.5
0.4
0.3
0.2
0
5
10
15
20
25
30

图 6A

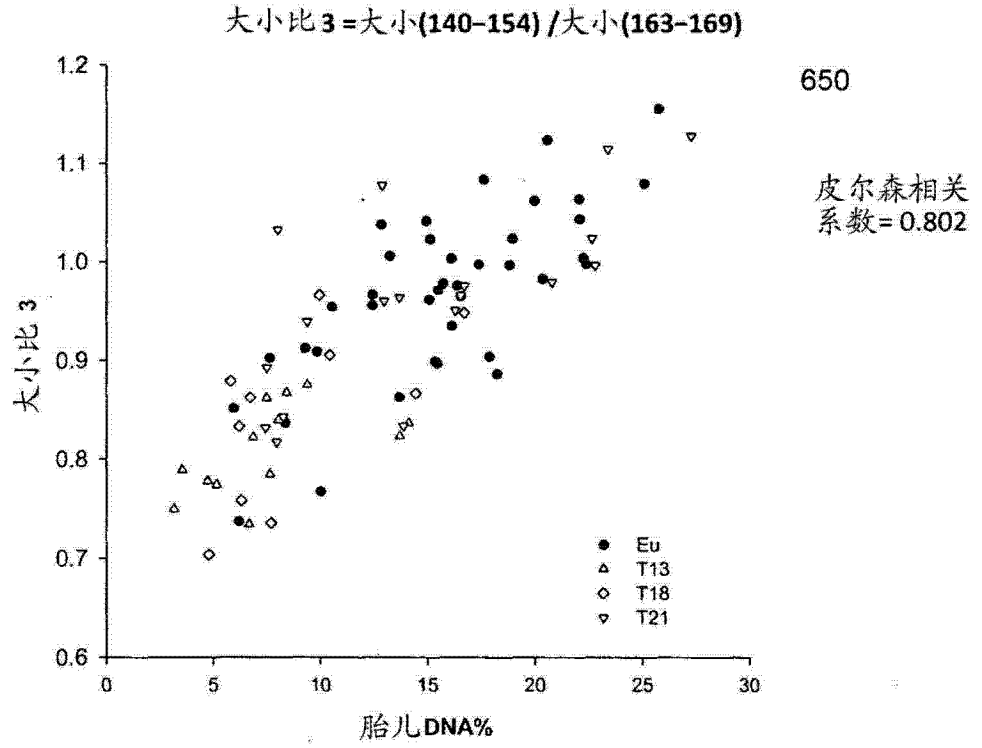

大小比 3 = 大小 (140–154) / 大小 (163–169)
650
皮尔森相关
系数 = 0.802
大小比 3
胎儿 DNA%
Eu
T13
T18
T21
1.2
1.1
1.0
0.9
0.8
0.7
0.6
0
5
10
15
20
25
30

图 6B

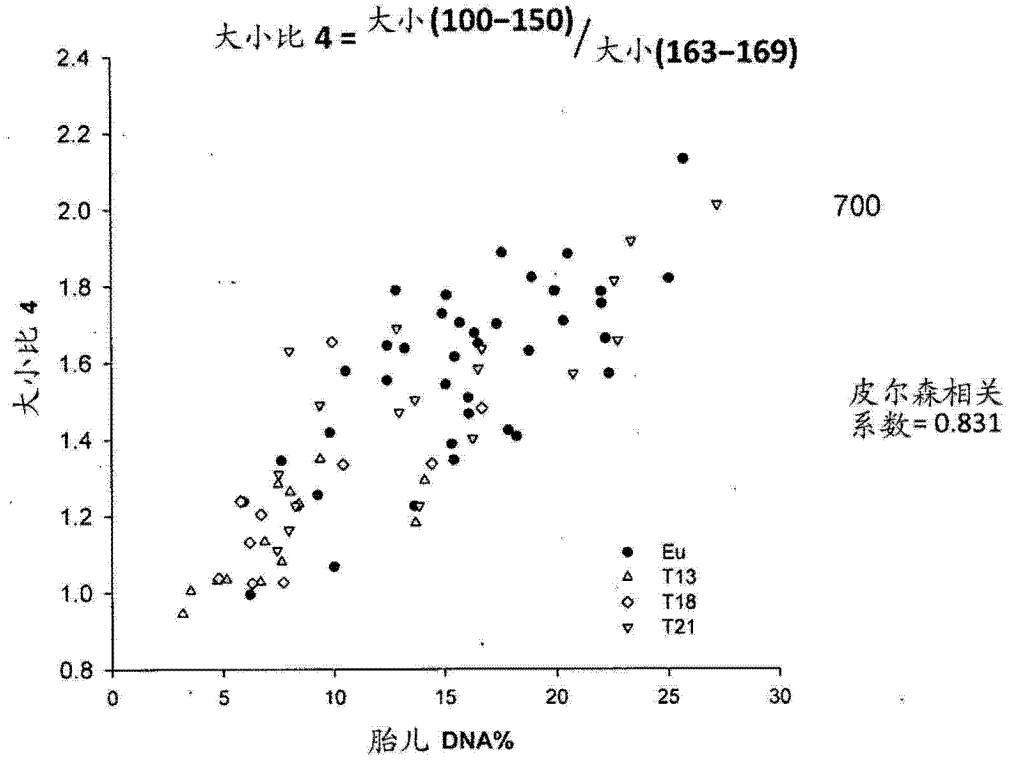

大小比 4 = 大小(100-150) / 大小(163-169)
700
皮尔森相关
系数= 0.831
大小比 4
胎儿 DNA%
Eu
T13
T18
T21

图 7

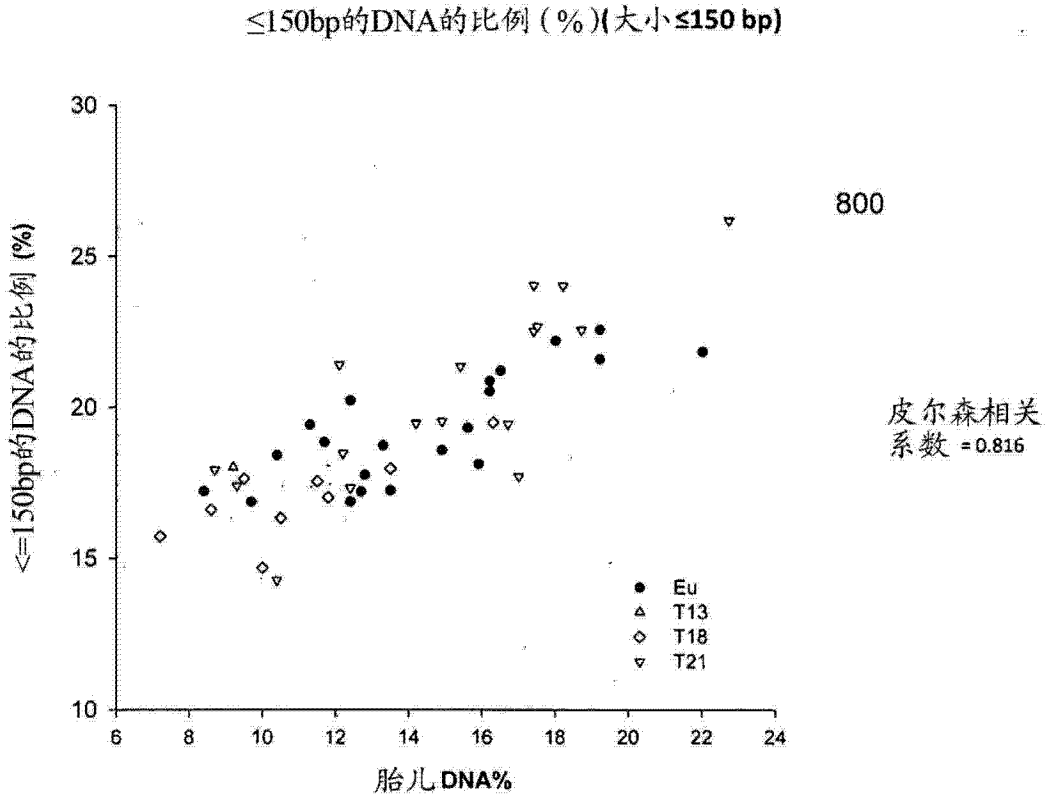

≤150bp的DNA的比例（%）(大小≤150 bp)
800
皮尔森相关
系数 = 0.816
<=150bp的DNA的比例 (%)
胎儿DNA%
Eu
T13
T18
T21

图 8

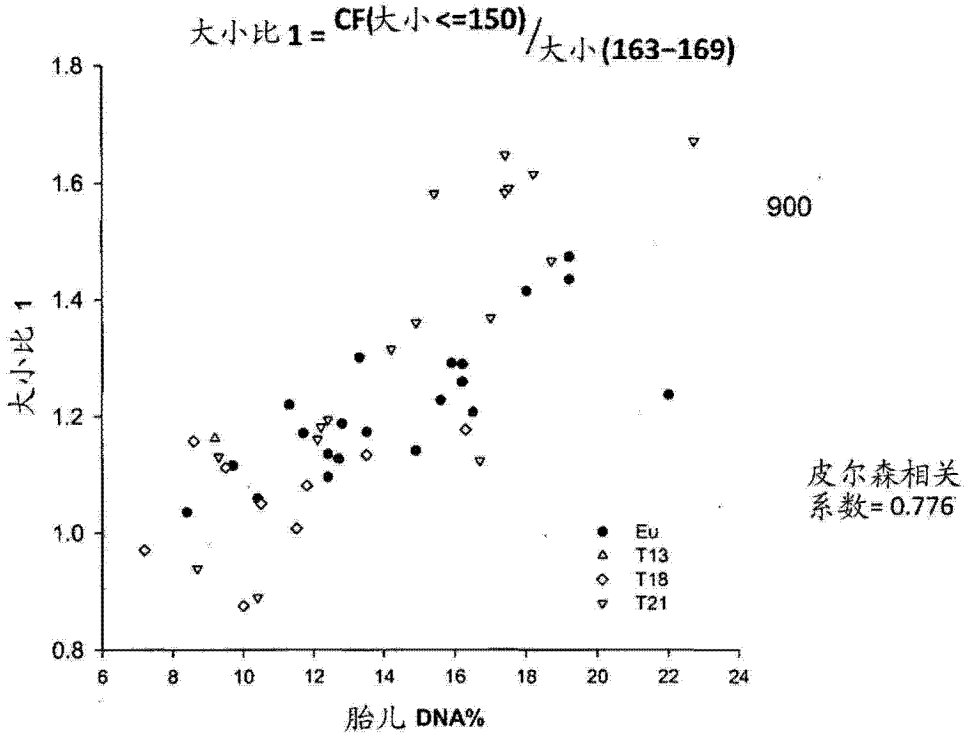

大小比 1 = CF(大小 <=150)/大小(163-169)
900
皮尔森相关
系数= 0.776
大小比 1
胎儿 DNA%
Eu
T13
T18
T21

图 9A

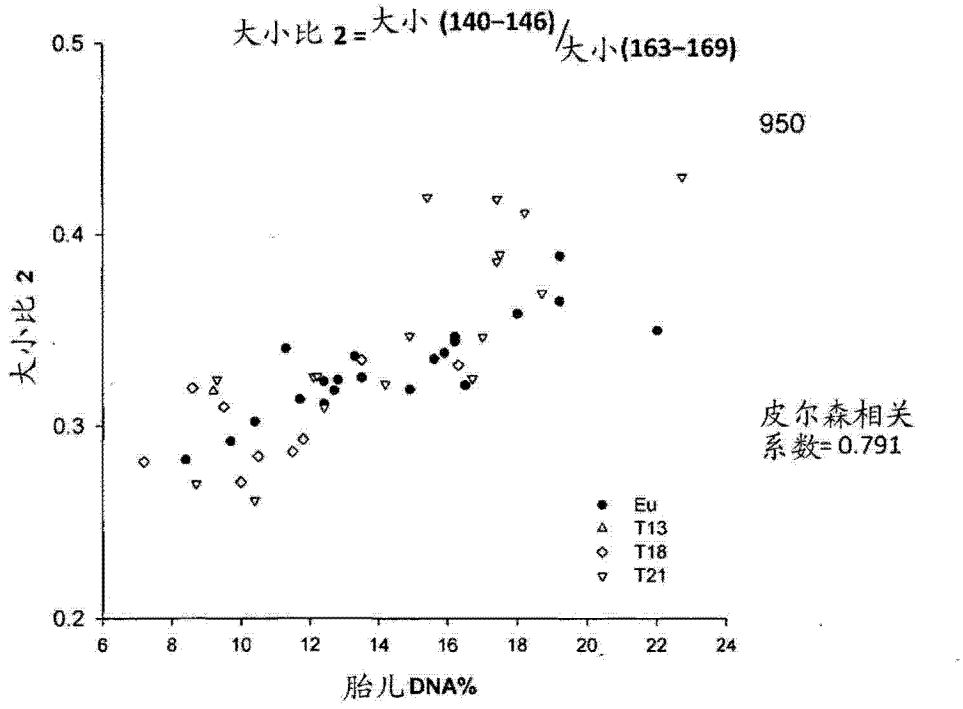

大小比 2 = 大小 (140-146)/大小(163-169)
950
皮尔森相关
系数= 0.791
大小比 2
胎儿DNA%
Eu
T13
T18
T21

图 9B

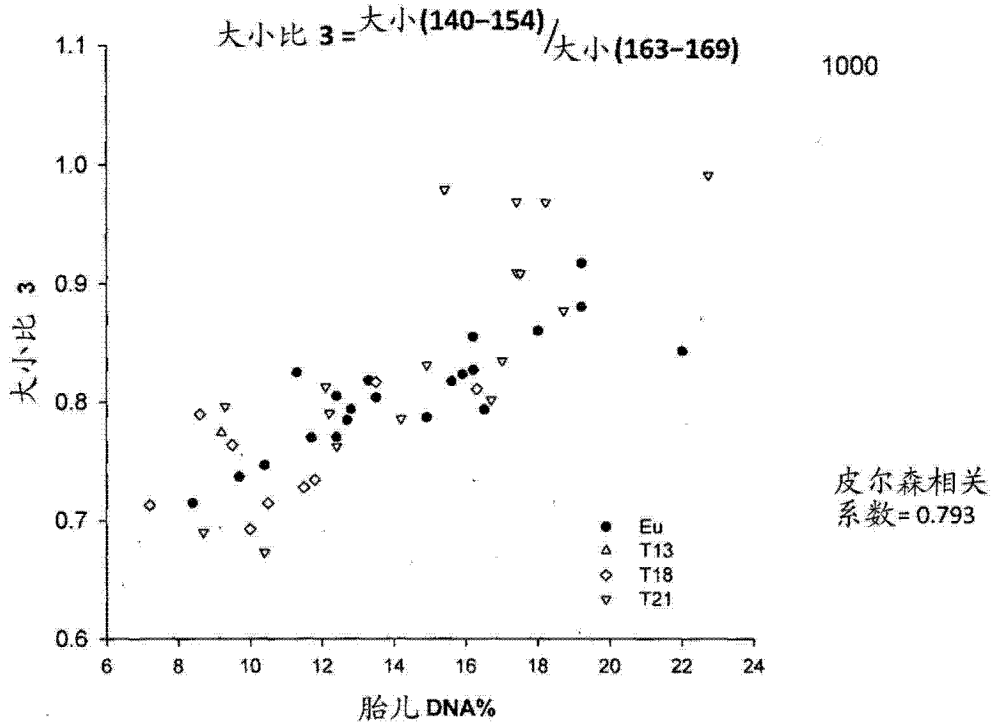

大小比 3 = 大小(140-154)/大小(163-169)
1000
大小比 3
胎儿 DNA%
皮尔森相关
系数 = 0.793
Eu
T13
T18
T21

图 10A

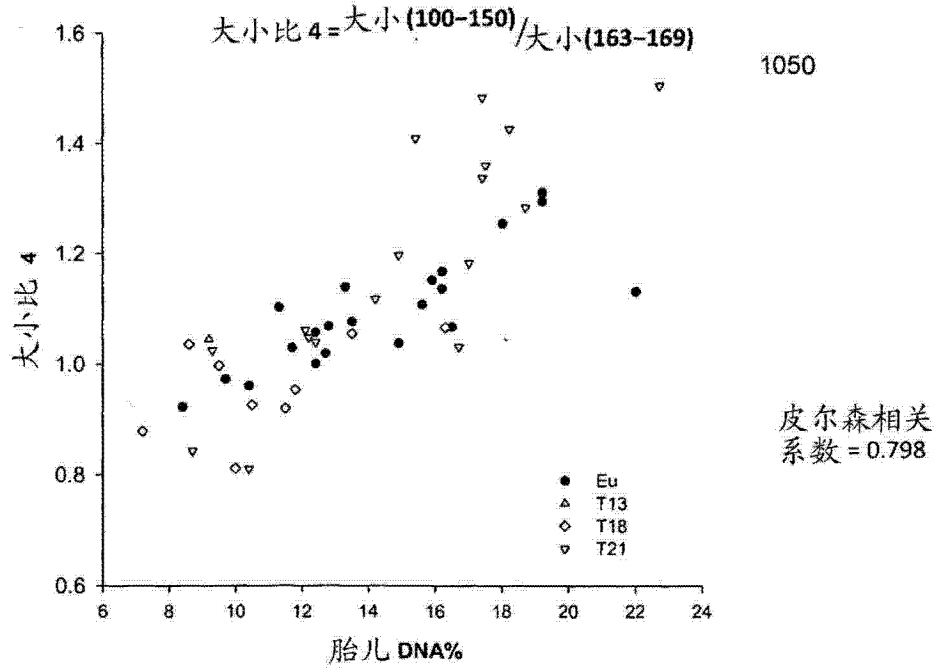

大小比 4 = 大小(100-150)/大小(163-169)
1050
大小比 4
胎儿 DNA%
皮尔森相关
系数 = 0.798
Eu
T13
T18
T21

图 10B

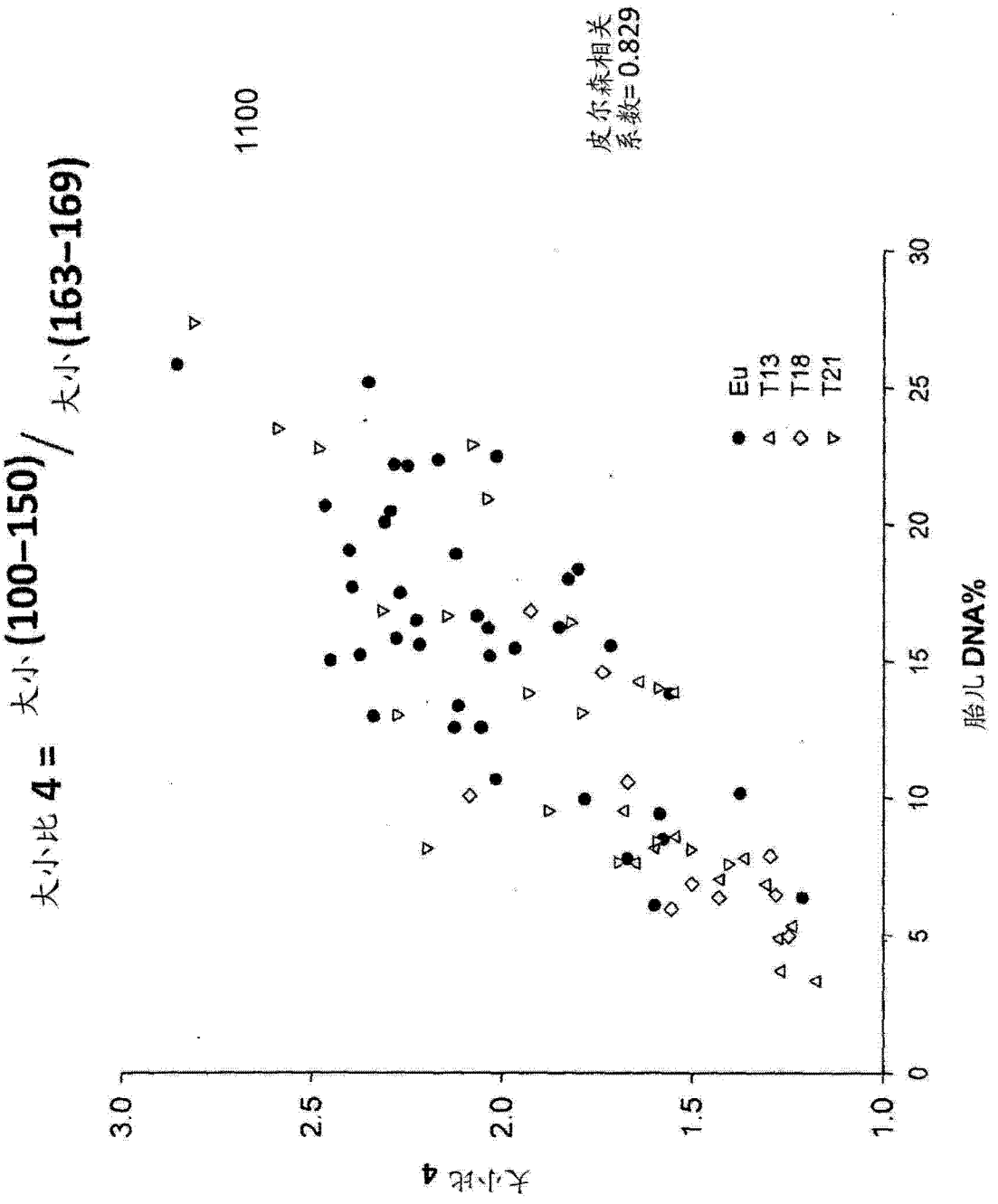

大小比 4 = 大小 (100–150) / 大小 (163–169)
1100
皮尔森相关
系数= 0.829
Eu
T13
T18
T21
胎儿 DNA%
大小比 4
3.0
2.5
2.0
1.5
1.0
0
5
10
15
20
25
30

图 11

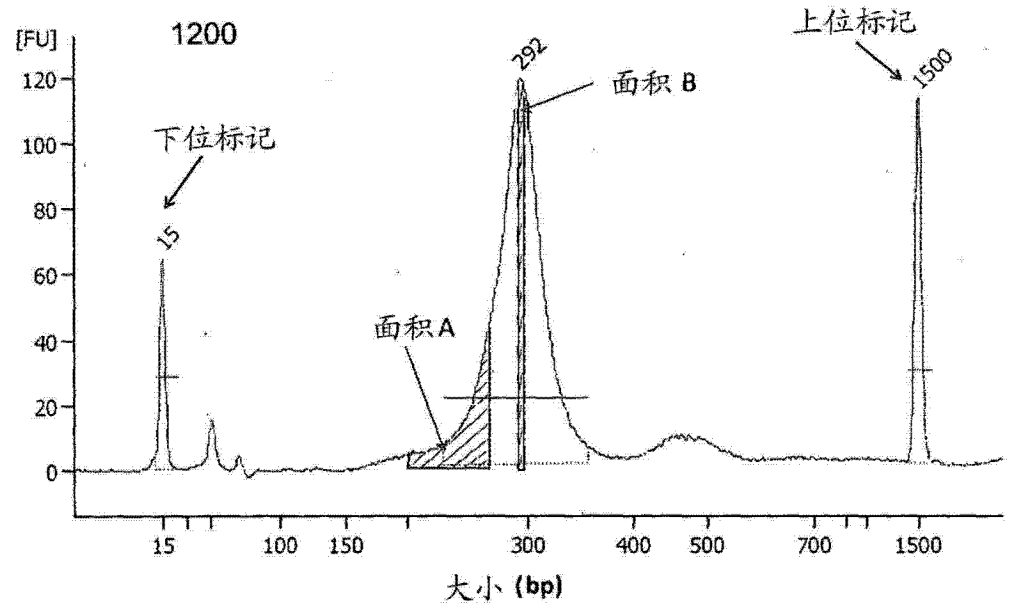

1200
[FU]
120
100
80
60
40
20
0
下位标记
15
面积A
292
面积 B
上位标记
1500
15
100
150
300
400
500
700
1500
大小 (bp)

图 12A

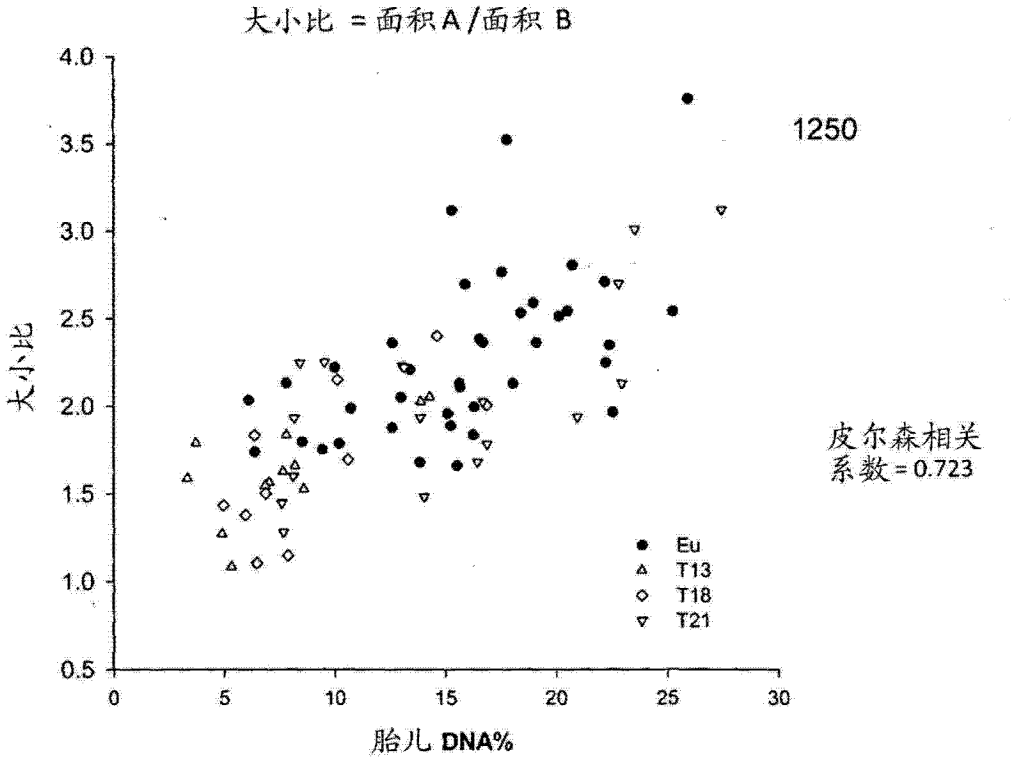

大小比 = 面积A/面积 B
1250
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
大小比
0
5
10
15
20
25
30
胎儿 DNA%
皮尔森相关
系数 = 0.723
Eu
T13
T18
T21

图 12B

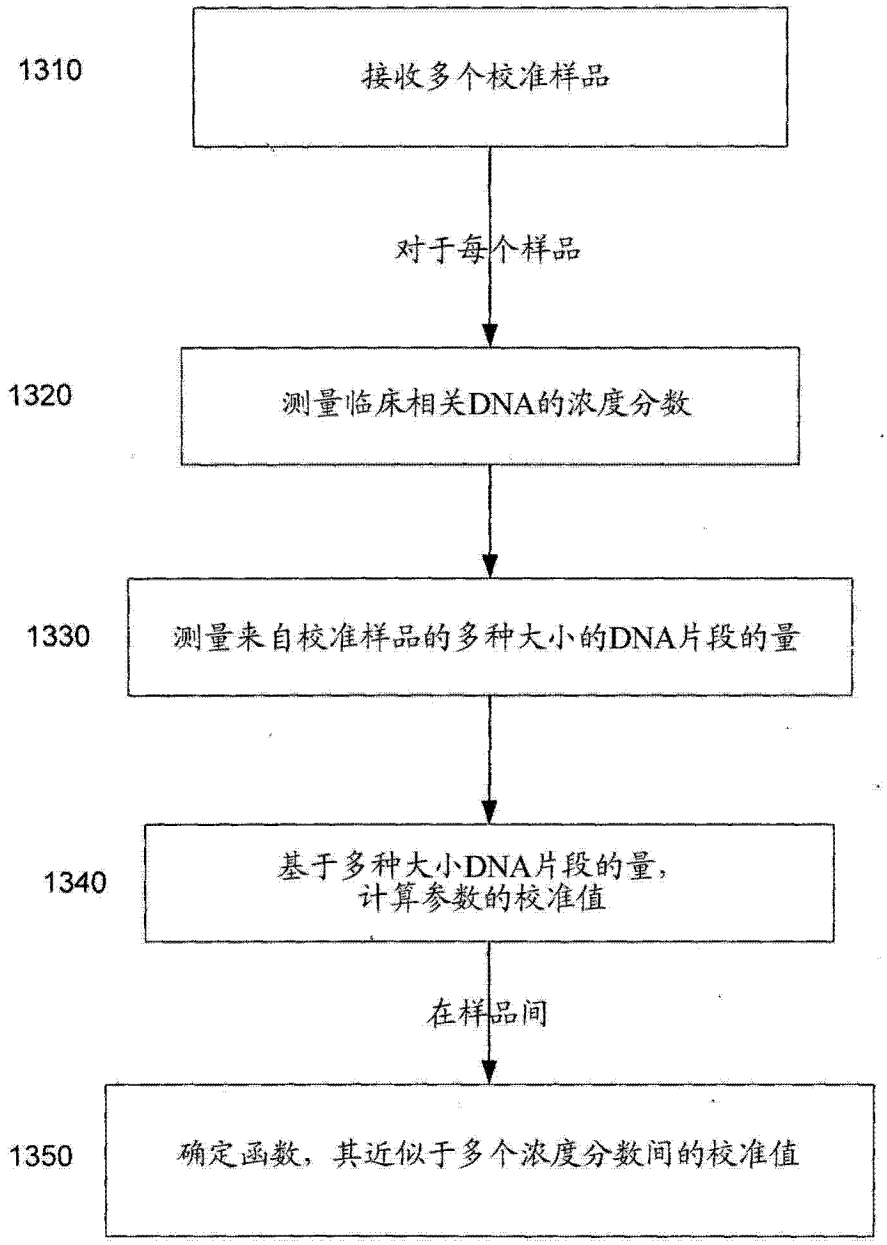

1310
接收多个校准样品
对于每个样品
1320
测量临床相关DNA的浓度分数
1330
测量来自校准样品的多种大小的DNA片段的量
1340
基于多种大小DNA片段的量，
计算参数的校准值
在样品间
1350
确定函数，其近似于多个浓度分数间的校准值

图 13

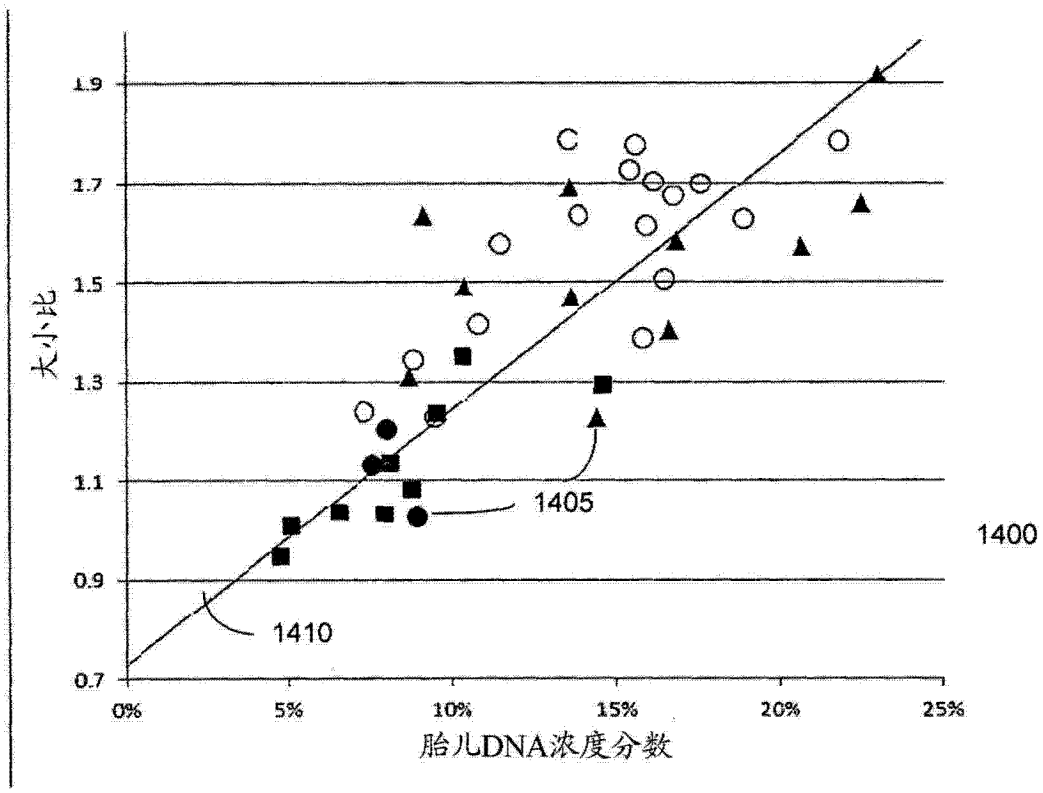

1.9
1.7
1.5
1.3
1.1
0.9
0.7
大小比
0%
5%
10%
15%
20%
25%
胎儿DNA浓度分数
1405
1410
1400

图 14A

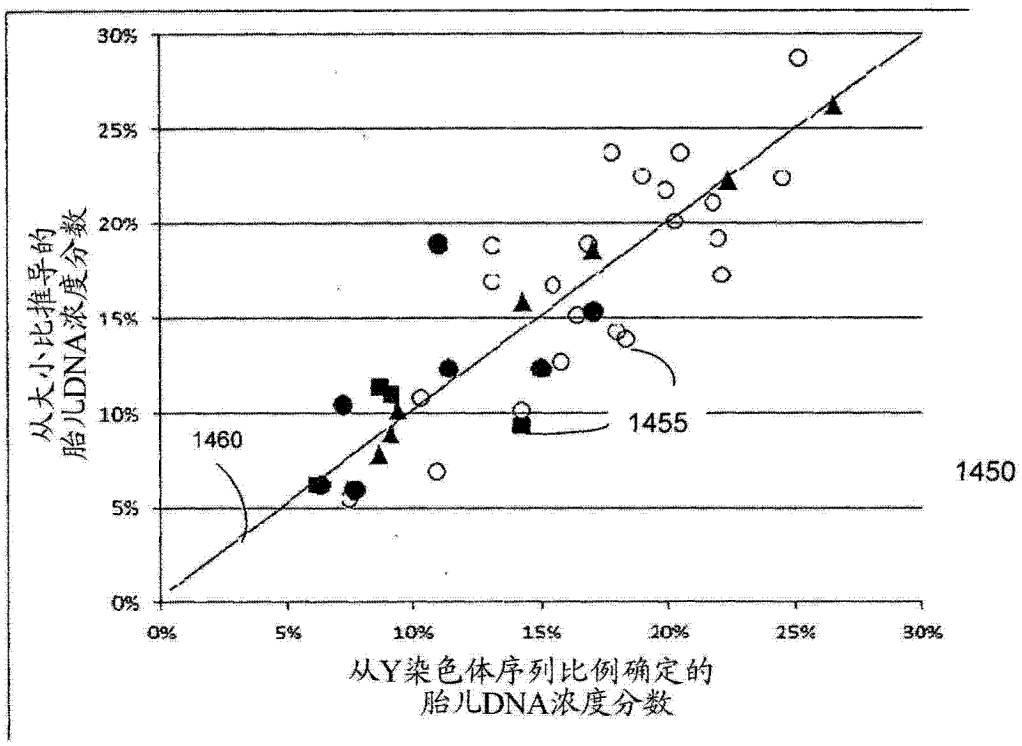

30%
25%
20%
15%
10%
5%
0%
从大小比推导的
胎儿DNA浓度分数
0%
5%
10%
15%
20%
25%
30%
从Y染色体序列比例确定的
胎儿DNA浓度分数
1455
1460
1450

图 14B

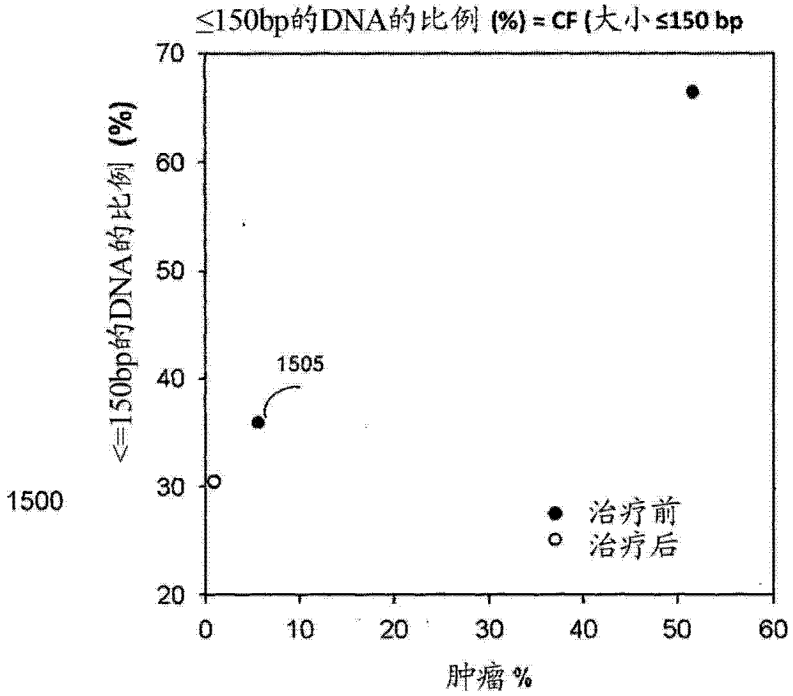

≤150bp的DNA的比例 (%) = CF (大小 ≤150 bp
<=150bp的DNA的比例 (%)
70
60
50
40
30
20
1505
1500
治疗前
治疗后
0
10
20
30
40
50
60
肿瘤 %

图 15A

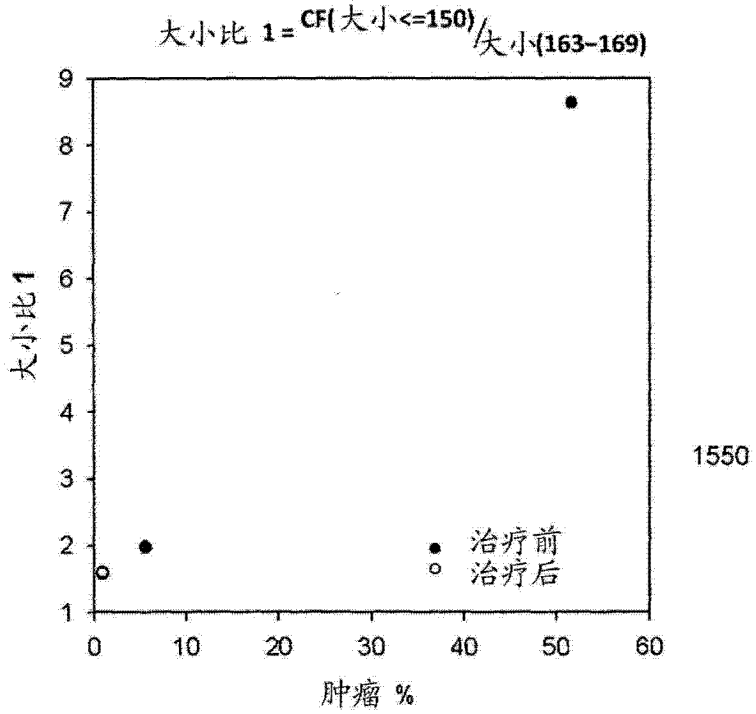

大小比 1 = CF(大小<=150)/大小(163-169)
大小比 1
9
8
7
6
5
4
3
2
1
1550
治疗前
治疗后
0
10
20
30
40
50
60
肿瘤 %

图 15B

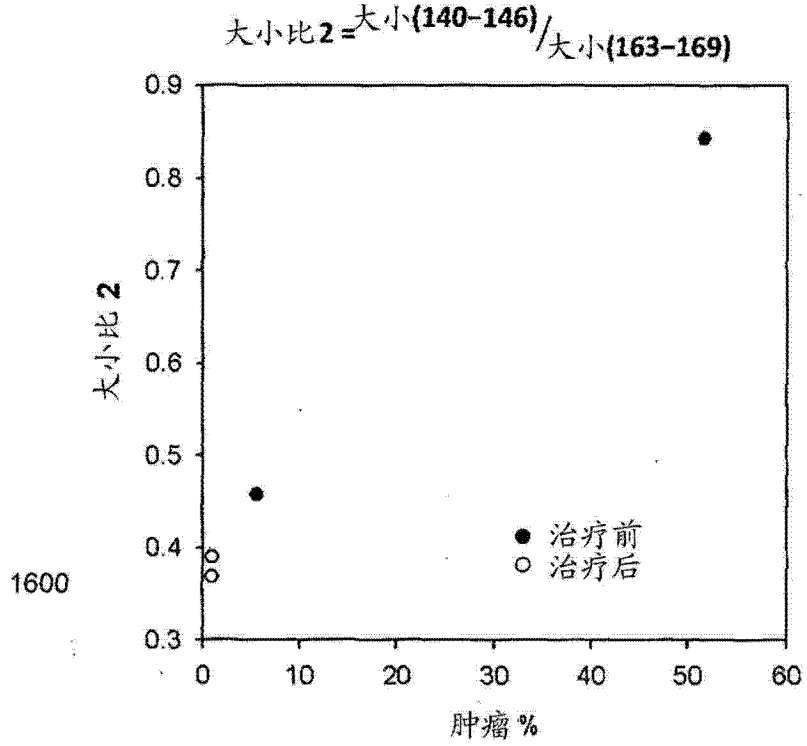

大小比2 = 大小(140-146) / 大小(163-169)
大小比 2
0.9
0.8
0.7
0.6
0.5
0.4
0.3
1600
治疗前
治疗后
0
10
20
30
40
50
60
肿瘤 %

图 16A

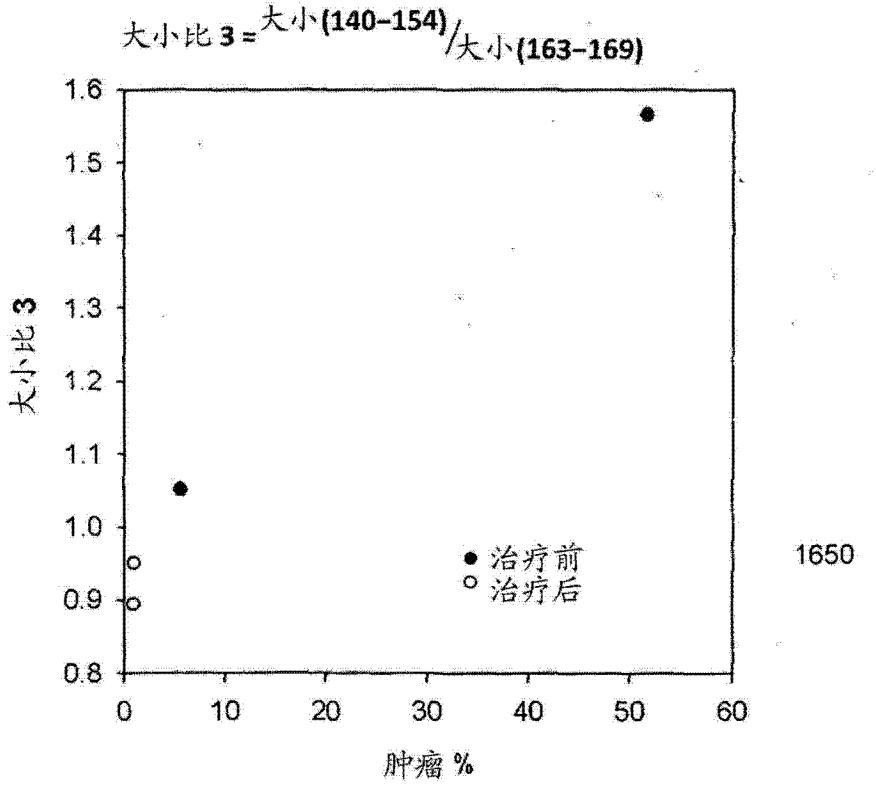

大小比 3 = 大小(140-154) / 大小(163-169)
大小比 3
1.6
1.5
1.4
1.3
1.2
1.1
1.0
0.9
0.8
治疗前
治疗后
1650
0
10
20
30
40
50
60
肿瘤 %

图 16B

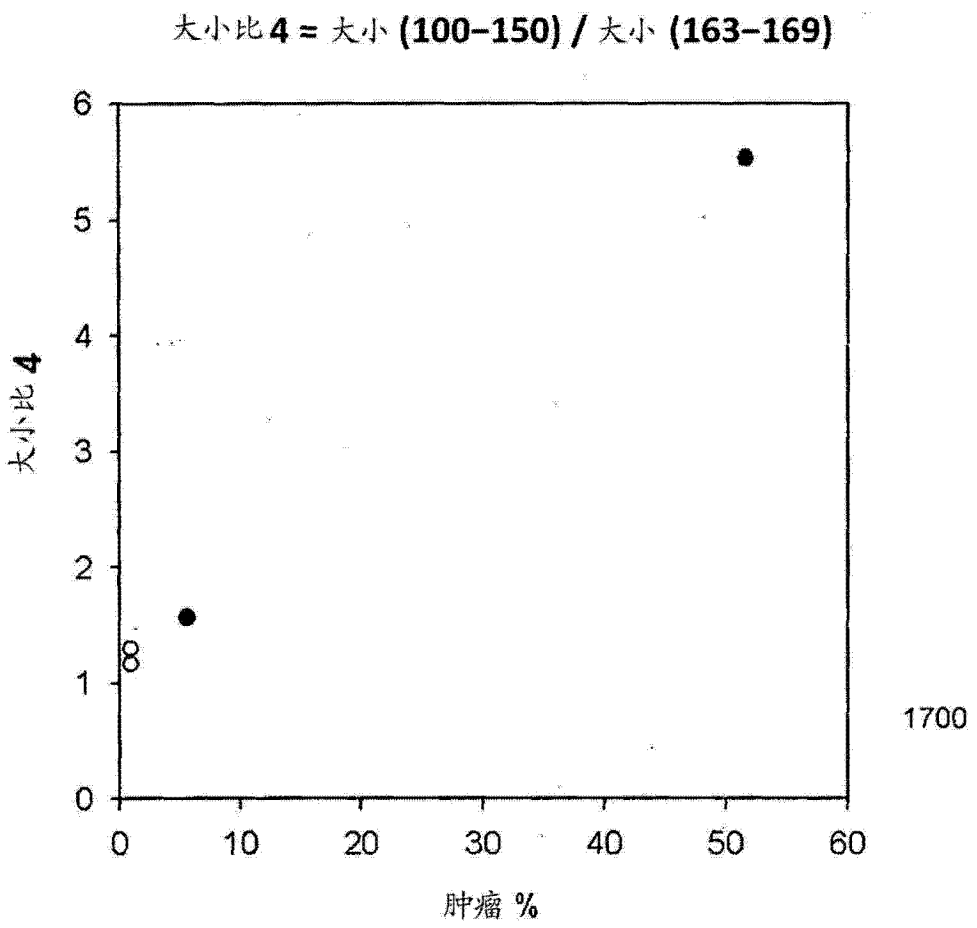

大小比4 = 大小 (100–150) / 大小 (163–169)
大小比4
6
5
4
3
2
1
0
0
10
20
30
40
50
60
肿瘤 %
8
1700

图 17

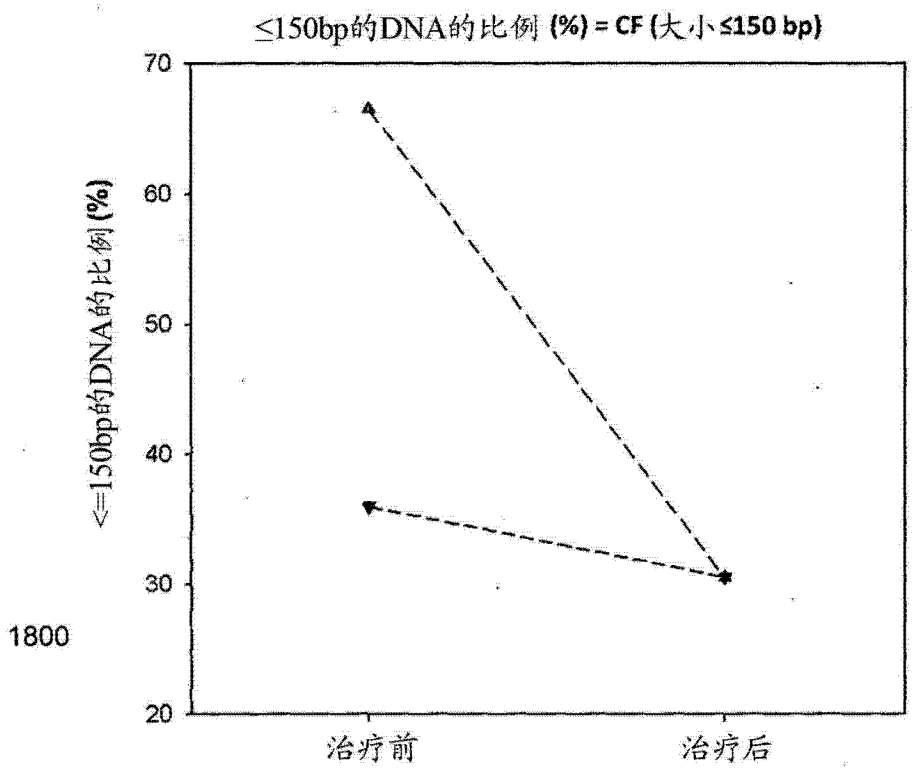

≤150bp的DNA的比例 (%) = CF (大小≤150 bp)
<=150bp的DNA的比例 (%)
70
60
50
40
30
20
1800
治疗前
治疗后

图 18A

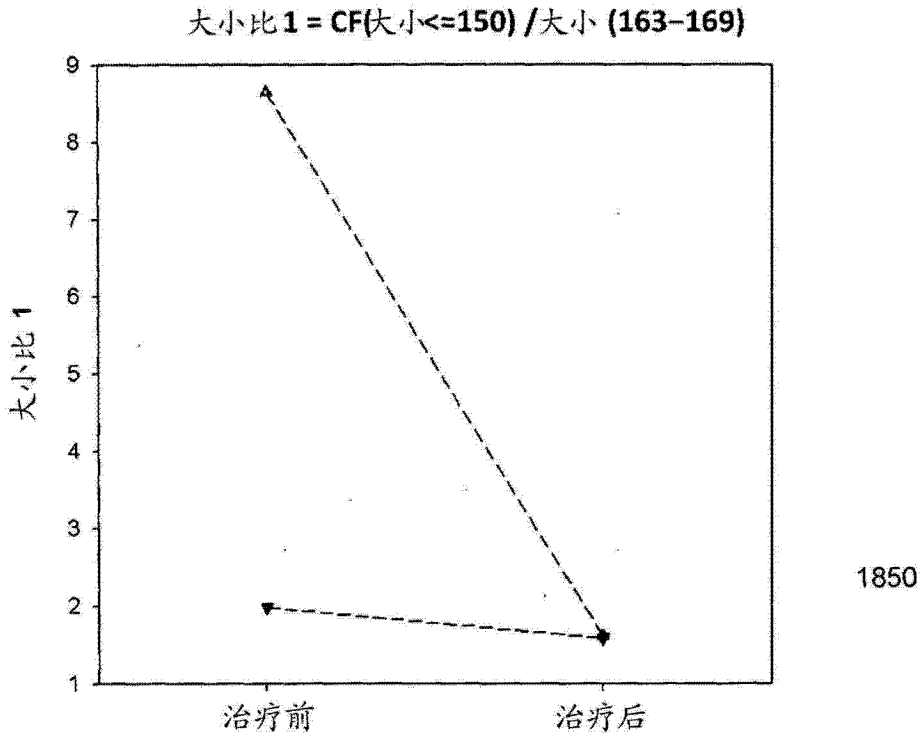

大小比 1 = CF(大小<=150) /大小 (163–169)
大小比 1
9
8
7
6
5
4
3
2
1
1850
治疗前
治疗后

图 18B

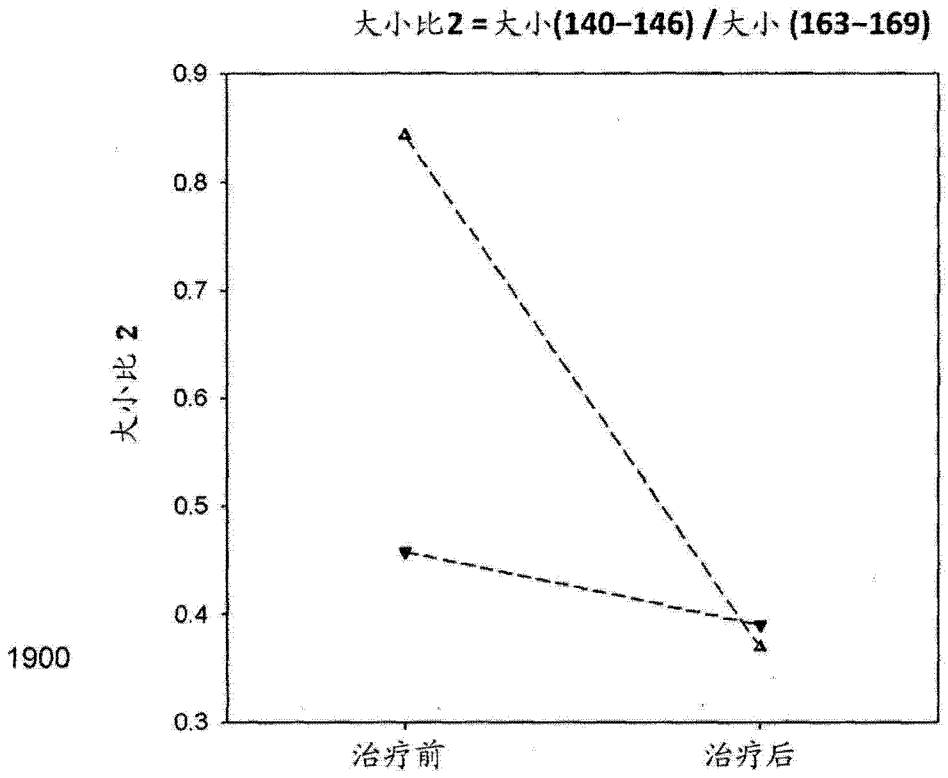

大小比2 = 大小(140-146) / 大小 (163-169)
大小比 2
0.9
0.8
0.7
0.6
0.5
0.4
0.3
治疗前
治疗后
1900

图 19A

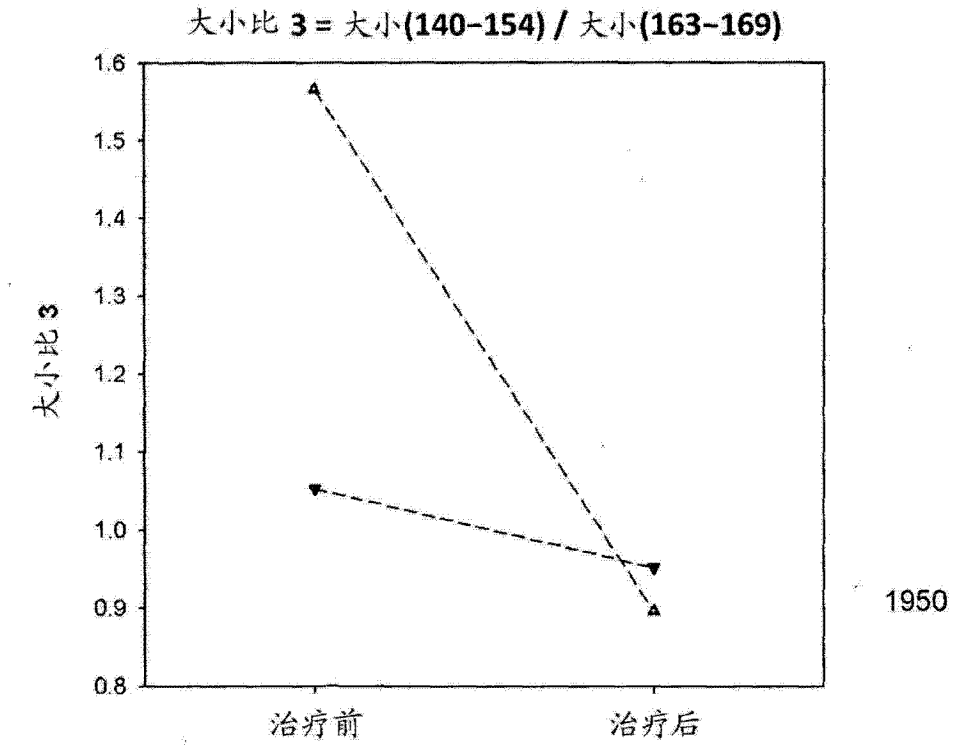

大小比 3 = 大小(140-154) / 大小(163-169)
大小比 3
1.6
1.5
1.4
1.3
1.2
1.1
1.0
0.9
0.8
治疗前
治疗后
1950

图 19B

大小比4 = 大小(100-150) / 大小(163-169)

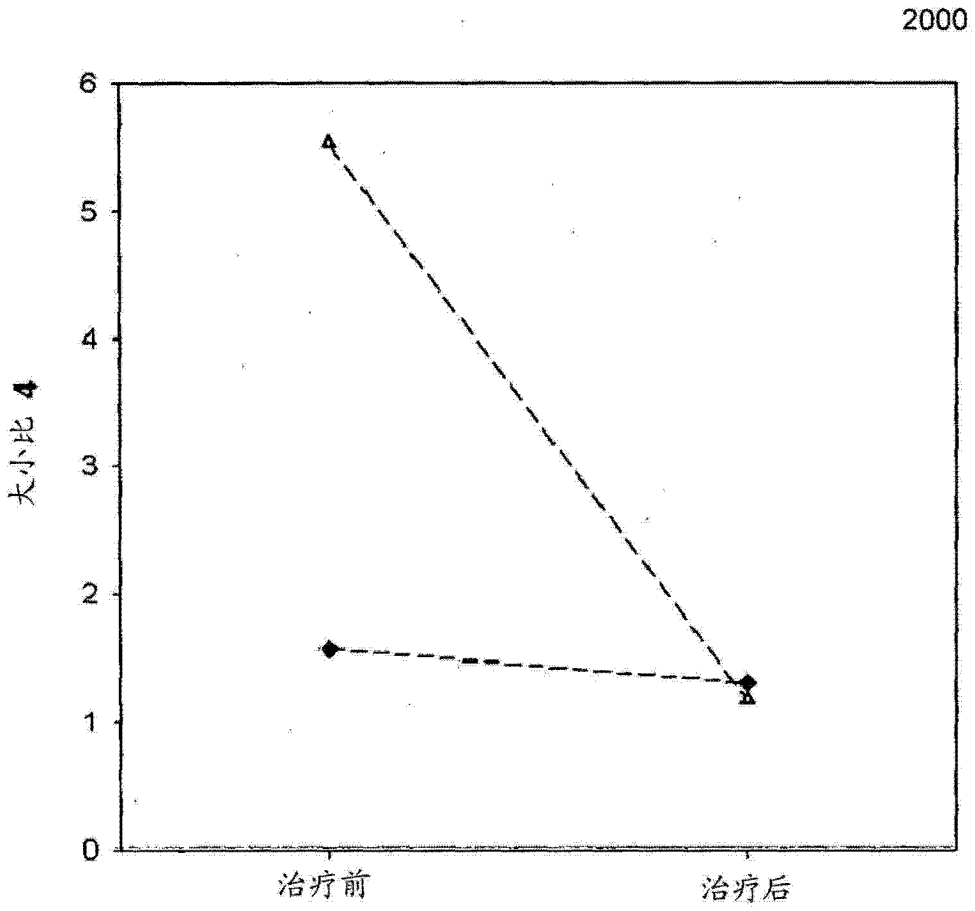


图 20

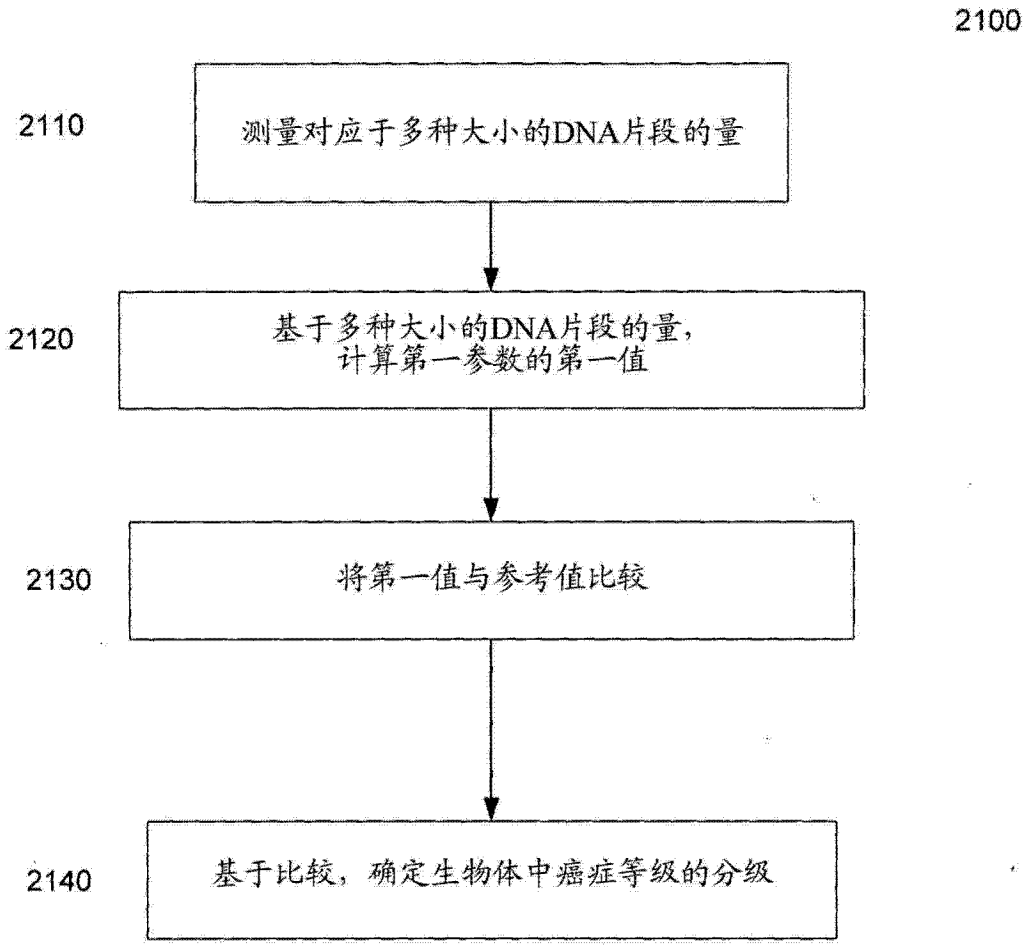

2100
2110
测量对应于多种大小的DNA片段的量
2120
基于多种大小的DNA片段的量，
计算第一参数的第一值
2130
将第一值与参考值比较
2140
基于比较，确定生物体中癌症等级的分级

图 21

	获得	丢失	参考文献
结肠直肠	7p, 7q, 8q, 11q, 13q, 和 20q,	5q, 8p, 17p, 18p, 18q 和 20p,	(Nakao et al. Carcinogenesis 2004;25: 1345-1357.) (Tsafrir et al. Cancer Res 2006; 66: 2129-2137)
乳腺	1q, 6p, 8q, 11q, 16p, 17q, 19, 和 20q	6q, 13q, 16q, 17p, 和 22q	(Tirkkonen et al. Gene Chromosome Canc 1998; 21: 177-184) (Richard et al. Int J Cancer 2000; 89: 305-310) (Pinkel et al. Nat Genet 1998; 20: 207-211) (Persson et al. Gene Chromosome Canc 1999; 25: 115-122) (Nishizaki et al. Int J Cancer 1997; 74: 513 - 517)
肺	1q, 3q, 5p, 和 8q	3p, 6q, 8p, 9p, 13q, 和17p	(Berrieman et al. Brit J Cancer 2004; 90: 900-905) (Luk et al. Cancer Genet Cytogen 2001; 125: 87 - 99) (Petersen et al. Cancer Res 1997; 57: 2331-2335) (Pei et al. Gene Chromosome Canc 2001; 31: 282-287)
HCC	1q, 8q, 17q 和 20q	4q, 6q, 8p, 13q, 16q和 17p	(Kusano et al. Cancer 2002; 94: 746-751) (Laurent-Puig et al. Gastroenterology 2001; 120: 1763-1773) (Moinzadeh et al. Brit J Cancer 2005; 92: 935-941)
卵巢	20q, 3q, 1q, 8q, 12p, 11q, 和17q	Xp, 18q, 4q, 9p, 和 13q	(Taetle et al. Gene Chromosome Canc 1999; 25: 290-300) (Schraml et al. Am J Pathol 2003; 163: 985 - 992) (Sonoda et al. Gene Chromosome Canc 1997; 20: 320-328)

2200

图 22

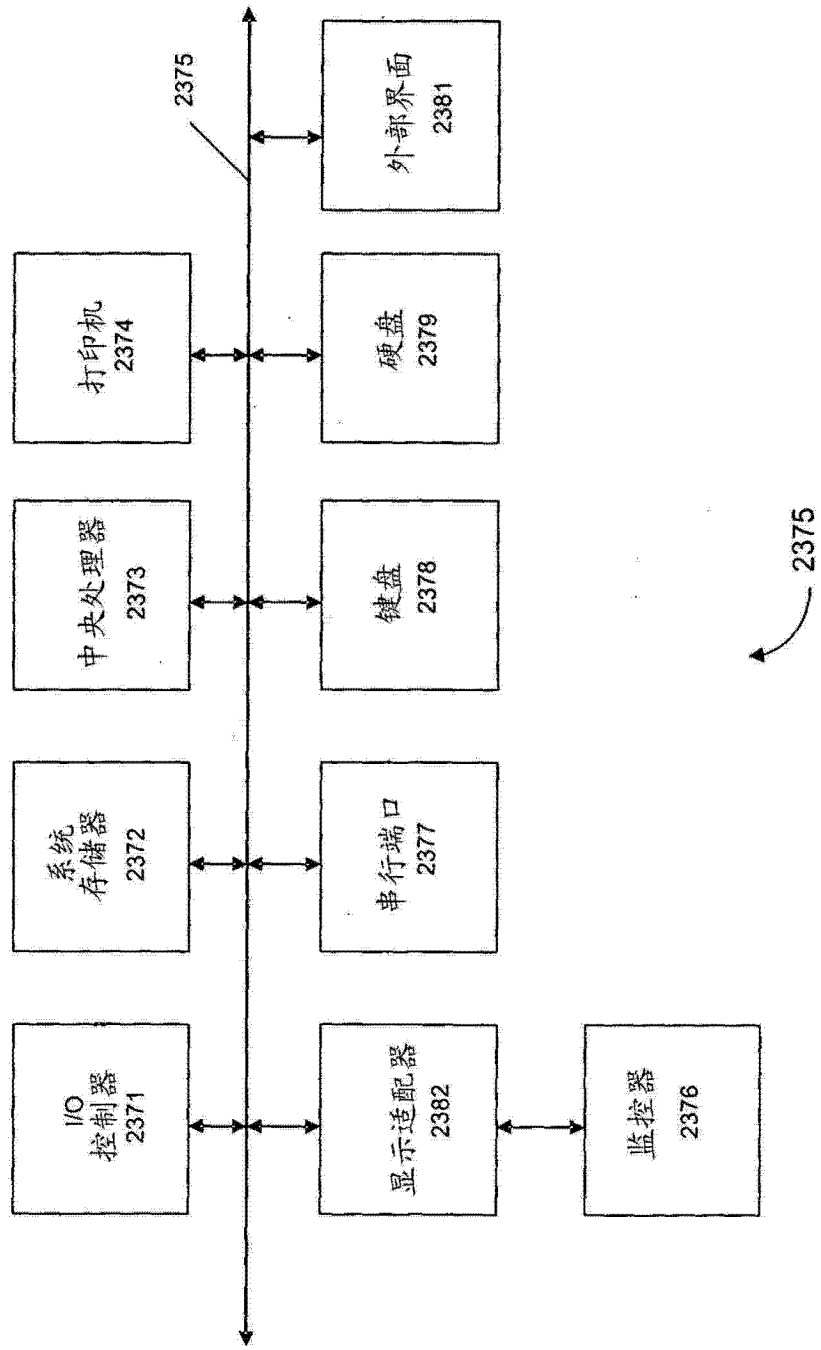

2375
外部界面
2381
打印机
2374
硬盘
2379
中央处理器
2373
键盘
2378
系统
存储器
2372
串行端口
2377
I/O
控制器
2371
显示适配器
2382
监控器
2376
2375

图 23