(19)中华人民共和国国家知识产权局

(12)发明专利

(10)授权公告号 CN 105229168 B
(45)授权公告日 2020.07.17

(21)申请号 201480009730.6

(22)申请日 2014.02.19

(65)同一申请的已公布的文献号
申请公布号 CN 105229168 A

(43)申请公布日 2016.01.06

(30)优先权数据
61/767,219 2013.02.20 US

(85)PCT国际申请进入国家阶段日
2015.08.20

(86)PCT国际申请的申请数据
PCT/US2014/017226 2014.02.19

(87)PCT国际申请的公布数据
WO2014/130589 EN 2014.08.28

(73)专利权人 生物纳米基因有限公司
地址 美国宾夕法尼亚州

(72)发明人 曹涵 亚历克斯·R·黑斯蒂
欧内斯特·梓-森·莱姆

(74)专利代理机构 北京路浩知识产权代理有限公司 11002
代理人 瞿卫军 王朋飞

(51)Int.Cl.
C12Q 1/6883(2018.01)
C12Q 1/6809(2018.01)
G01N 33/50(2006.01)
C12M 1/34(2006.01)
C12M 3/00(2006.01)

(56)对比文件
CN 101765462 A,2010.06.30,
US 2004197843 A1,2004.10.07,
Ernest T Lam等.Genome mapping on nanochannel arrays for structural variation analysis and sequence assembly.《Nature Biotechnology》.2012,第30卷(第8期),第771-776页.

审查员 王瑶

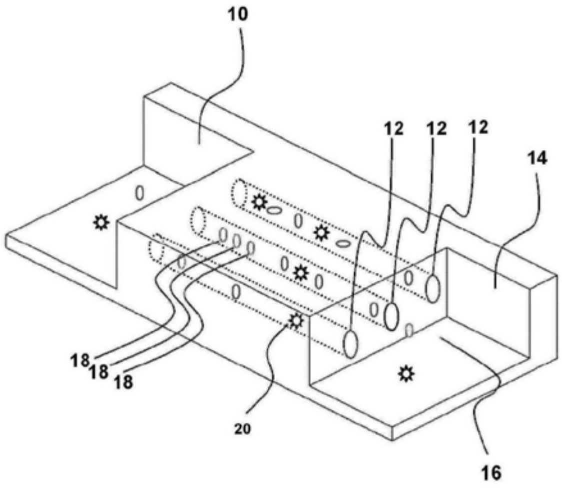
权利要求书6页 说明书16页 附图7页

(54)发明名称

纳米流体中分子的表征

(57)摘要

本发明提供了用于使用流体学检测和定量分子的方法。在优选实施方案中,该方法包括分析血液以检测来自胎儿或肿瘤的循环DNA或细胞的存在。



CN 105229168 B

1.用于非诊断目的的表征样品的方法,所述方法包括:

用至少第一标记标记多个样品分子,其中所述样品分子包含一个感兴趣的基因组片段或多个感兴趣的基因组片段的多核苷酸序列;

提供多个标记的参照分子,其中所述参照分子包含所述基因组片段的多核苷酸序列;

通过流体通道使多个标记的样品分子线性化;

通过所述流体通道使多个标记的参照分子线性化;

检测来自所述标记的样品分子和所述标记的参照分子的信号,所述信号包括感兴趣的基因组片段的模式特征;

将所述模式与感兴趣的参照基因组片段的模式进行比较;和

确定所述感兴趣的基因组片段上所述样品分子的覆盖深度;和所述基因组片段上所述参照分子的覆盖深度,

以确定所述样品中感兴趣的基因组片段的拷贝数。

2.如权利要求1所述的方法,其中进行标记包括用第一标记标记所述样品分子,并且其中所述参照分子包含第二标记,

其中所述第一标记经配置产生一种第一模式或多种第一模式,

其中所述第二标记经配置产生一种第二模式或多种第二模式,并且

其中所述第一标记和所述第二标记互不相同。

3.如权利要求1所述的方法,其中进行标记包括用第一标记进行标记,

其中一种第一模式或多种第一模式和一种第二模式或多种第二模式各包含所述第一标记,以及

其中一种第一模式或多种第一模式和一种第二模式或多种第二模式互不相同。

4.如权利要求1所述的方法,其还包括标记参照分子,以产生所述标记的参照分子,其中所述标记的参照分子包含一种第二模式或多种第二模式。

5.如权利要求1所述的方法,其中所述标记的参照分子和样品分子来自所述样品。

6.如权利要求1所述的方法,其中所述标记的参照分子来自与所述样品相同的生物体的不同组织。

7.如权利要求1所述的方法,其中来自所述标记的参照分子的信号包含电子或光学储存的值或值集。

8.如权利要求1所述的方法,其还包括:

用至少所述第一标记标记来自第二样品的多个第二样品分子,其中所述多个第二样品分子包含所述一个感兴趣的第一基因组片段或多个感兴趣的第一基因组片段的多核苷酸序列,其中已知所述多个第二样品分子与染色体异常不对应;

通过所述流体通道使所述多个第二样品分子易位。

9.如权利要求1所述的方法,

其中所述一个感兴趣的第一基因组片段或多个感兴趣的第一基因组片段包含性染色体或其至少一个片段,并且

其中所述一个参照基因组片段或多个参照基因组片段包含常染色体或其至少一个片段。

10.如权利要求1所述的方法,

其中所述一个感兴趣的第一基因组片段或多个感兴趣的第一基因组片段包含第一常染色体或其至少一个片段，其选自：人类21号染色体、人类13号染色体、人类14号染色体、人类15号染色体、人类16号染色体、人类18号染色体和人类22号染色体以及其片段，并且

其中所述一个参照基因组片段或多个参照基因组片段包含第二常染色体或其至少一个片段，其中所述第二常染色体或其片段与所述第一常染色体或其片段不同。

11.如权利要求1所述的方法，其中所述样品分子来自包含可能的基因组异常的样品，以及

其中所述一个参照基因组片段或多个参照基因组片段包含第一染色体或其片段，并且其中所述参照基因组片段来自已知不包含可能的基因组异常的第二样品。

12.如权利要求1-11中任一项所述的方法，其中所述第一标记包含荧光标记、放射性标记、磁性标记或非光学标记中的至少一种。

13.如权利要求2所述的方法，其中所述第二标记包含荧光标记、放射性标记、磁性标记或非光学标记中的至少一种。

14.如权利要求1-11中任一项所述的方法，其中进行标记包括：

用切口核酸内切酶在第一序列基序处使双链DNA的一条链产生切口；和

标记所述DNA。

15.如权利要求14所述的方法，还包括修复所述DNA上的至少一些切口。

16.如权利要求14所述的方法，其中不修复所述切口。

17.如权利要求14所述的方法，其中所述标记包含转录终止子。

18.如权利要求1-11中任一项所述的方法，其中用所述第一标记进行标记包括，用DNA结合实体和甲基转移酶对所述样品分子的至少一个序列基序加标签，所述DNA结合实体选自：非切割限制酶、锌指蛋白、抗体、转录因子、转录激活剂样结构域、DNA结合蛋白、聚酰胺、形成三股螺旋的寡核苷酸以及肽核酸。

19.如权利要求1-11中任一项所述的方法，其中用所述第一标记进行标记包括，用甲基转移酶对所述样品分子的至少一个序列基序加标签。

20.如权利要求1-11中任一项所述的方法，其还包括用非序列特异性标记标记所述样品分子。

21.如权利要求1-11中任一项所述的方法，其中流体纳米通道包含长度为至少10nm且横截面直径小于5000nm的通道。

22.如权利要求1-11中任一项所述的方法，其中所述样品选细菌、病毒体、DNA分子、RNA分子、核酸聚合物、蛋白质、肽和多糖。

23.如权利要求1-11中任一项所述的方法，其中所述样品来源于母体血液，并且其中所述参照分子来源于除血液之外的母体样品。

24.如权利要求1-11中任一项所述的方法，其中所述样品包含核苷酸，并且其中至少两种标记位于所述核苷酸中感兴趣区域的任一端。

25.如权利要求1-11中任一项所述的方法，其中所述检测包括测定实物计数、强度、波长或所述标记的大小。

26.如权利要求1-11中任一项所述的方法，其中所述检测包括测定所述样品中至少一个标记区域的长度。

27.如权利要求1-11中任一项所述的方法，其还包括测定由包含所述样品或所述样品部分的集合产生的信号。

28.如权利要求1-11中任一项所述的方法，其中进行比较包括使用由多个样品或样品部分产生的信号S1、S2…Sn与由参照产生的信号C之间的比例K：K1＝S1/C、K2＝S2/C…Kn＝Sn/C。

29.如权利要求28所述的方法，其中K1和Kn之间的差异用于鉴定胎儿样品的存在。

30.如权利要求28所述的方法，其中K1和Kn之间的差异用于鉴定来自肿瘤或其它癌源的DNA的存在。

31.如权利要求28所述的方法，其中K1和Kn之间的差异用于鉴定所述样品中遗传异常的存在。

32.如权利要求1-11中任一项所述的方法，其中所述感兴趣的参照基因组片段来源于已知的二倍体或单倍体染色体。

33.如权利要求1-11中任一项所述的方法，其中将来自所述样品的信号与来自元基因组的或微生物组研究的种群分布关联。

34.如权利要求1-11中任一项所述的方法，其中所述流体通道是纳米通道。

35.如权利要求1-11中任一项所述的方法，其中将所述流体通道与基底表面平行设置。

36.如权利要求1-11中任一项所述的方法，所述方法还包括产生直方图分布，以反映样品的覆盖深度。

37.如权利要求1-11中任一项所述的方法，其中所述样品包含循环胎儿细胞、循环肿瘤细胞或体液或组织。

38.如权利要求1-11中任一项所述的方法，其中易位包括对所述标记的样品施加驱动力，所述驱动力选自流体流、辐射场、电渗透力、电泳力、电动力、温度梯度、表面性质梯度、毛细管流、压力梯度、磁场、电场、后退弯月面、表面张力、热梯度、拉力、推力及其组合。

39.如权利要求1-11中任一项所述的方法，其中每个所述样品分子的长度小于500个核苷酸。

40.表征样品的系统，其包括：

用于用至少两种标记标记样品分子的一个或多个区域；

用于使所述标记的样品分子易位的流体通道，其中所述流体通道经配置延长所述样品分子的至少一部分，并且其中所述流体通道的长度为至少10nm且横截面直径小于5000nm；以及

用于检测所述流体通道中由所述标记的样品产生的信号的装置，

其中所述系统被配置为生成直方图以反映所述样品的覆盖深度，以确定所述样品中感兴趣的基因组片段的拷贝数。

41.表征样品的系统，其包括：

用于标记样品核酸分子的一个或多个区域；

用于使所述标记的样品核酸分子易位的流体纳米通道，其中所述流体纳米通道经配置延长所述样品核酸分子的至少一部分，并且其中所述流体纳米通道的长度为至少10nm且横截面直径小于1000nm；以及

用于检测所述流体通道中由所述样品核酸分子产生的信号的装置，

其中所述系统被配置为生成直方图以反映所述样品的覆盖深度，以确定所述样品中感兴趣的基因组片段的拷贝数。

42.表征样品的系统，其包括：

用于标记样品DNA上多个序列特异性位置的区域；

用于使所述样品DNA的至少一部分线性化的区域；以及

用于量化由所述样品DNA上的标记产生的信号的装置，

其中所述系统被配置为生成直方图以反映所述样品的覆盖深度，以确定所述样品中感兴趣的基因组片段的拷贝数。

43.如权利要求40-42中任一项所述的系统，其中所述样品选自细菌、病毒体、DNA分子、RNA分子、核酸聚合物、蛋白质、肽和多糖。

44.如权利要求40-42中任一项所述的系统，其中所述样品来源于母体血液。

45.如权利要求40-42中任一项所述的系统，其中所述样品包含核苷酸，并且所述至少两种标记位于所述核苷酸中感兴趣的区域的任一端。

46.如权利要求40-42中任一项所述的系统，其中所述标记选自荧光标记、放射性标记、磁性标记或其组合。

47.如权利要求40-41中任一项所述的系统，其中所述检测信号包括测定实物计数、强度、波长或所述标记的大小。

48.如权利要求40-41中任一项所述的系统，其中所述检测信号包括测定所述样品中至少一种标记的区域的长度。

49.如权利要求40-42中任一项所述的系统，其中将信号关联包括测定由样品集合或样品部分的集合产生的信号。

50.如权利要求40-42中任一项所述的系统，其中所述系统经配置用于将信号关联，其中使用由多种样品或样品部分产生的信号S1、S2…Sn与由参照产生的信号C之间的比例K将信号关联：K1＝S1/C、K2＝S2/C…Kn＝Sn/C。

51.如权利要求50所述的系统，其中K1和Kn之间的差异用于鉴定胎儿样品的存在。

52.如权利要求50所述的系统，其中K1和Kn之间的差异用于鉴定来自肿瘤或其它癌源的DNA的存在。

53.如权利要求50所述的系统，其中K1和Kn之间的差异用于确定所述样品中遗传异常的存在。

54.如权利要求53所述的系统，其中所述遗传异常是非整倍体。

55.如权利要求53所述的系统，其中所述遗传异常是易位、添加、扩增、颠换或倒置。

56.如权利要求40-42中任一项所述的系统，其还包括来源于已知的二倍体或单倍体染色体的参照。

57.如权利要求40-42中任一项所述的系统，其中将来自所述样品的信号与来自元基因组的或微生物组研究的种群分布关联。

58.如权利要求40-42中任一项所述的系统，其中所述流体通道是纳米通道。

59.如权利要求40-42中任一项所述的系统，其中将所述流体通与基底表面平行设置。

60.如权利要求40-42中任一项所述的系统，其中所述样品包含循环胎儿细胞、循环肿瘤细胞或体液或组织。

61.如权利要求40-42中任一项所述的系统,其中所述易位包括对所述标记的样品施加驱动力,所述驱动力选自流体流、辐射场、电渗透力、电泳力、电动力、温度梯度、表面性质梯度、毛细管流、压力梯度、磁场、电场、后退弯月面、表面张力、热梯度、拉力、推力及其组合。

62.如权利要求40-41中任一项所述的系统,其中所述样品分子为核酸,每个的长度小于500个核苷酸。

63.如权利要求42所述的系统,其中所述样品DNA的长度小于500个核苷酸。

64.用于非诊断目的的表征包含多核苷酸序列的样品的方法,所述方法包括:

用至少第一标记标记多个样品分子,其中所述样品分子包含一个感兴趣的基因组片段或多个感兴趣的基因组片段的多核苷酸序列,并且其中所述一个感兴趣的基因组片段或多个感兴趣的基因组片段与可能异常的基因组区域对应;

提供多个标记的参照分子,其中所述参照分子包含一个参照基因组片段或多个参照基因组片段的多核苷酸序列,并且其中已知所述一个参照基因组片段或多个参照基因组片段与所述可能异常的基因组区域不对应;

通过一个或多个流体通道使所述多个标记的样品分子和所述多个标记的参照分子易位;

检测来自所述标记的样品分子和所述标记的参照分子的信号,以至少确定:

所述一个感兴趣的基因组片段或多个感兴趣的基因组片段特有的一种第一模式或多种第一模式;和

所述一个参照基因组片段或多个照基因组片段特有的一种第二模式或多种第二模式;以及

将确定所述一种第一模式或多种第一模式的信号与确定所述一种第二模式或多种第二模式的信号关联,其中所述关联包括将一个感兴趣的基因组片段或多个感兴趣的基因组片段的覆盖深度与一个参照基因组片段或多个照基因组片段的覆盖深度进行比较,以确定所述一个感兴趣的基因组片段或多个感兴趣的基因组片段的拷贝数。

65.如权利要求64所述的方法,其中所述样品来源于母体血液,并且其中所述参照分子来源于除血液之外的母体样品。

66.如权利要求64所述的方法,其中所述样品包含循环胎儿细胞。

67.如权利要求64所述的方法,其还包括产生直方图分布,以反映样品的覆盖深度。

68.如权利要求64所述的方法,其还包括标记参照分子,以产生所述标记的参照分子,其中所述标记的参照分子包含所述一种第二模式或多种第二模式。

69.如权利要求64所述的方法,其中来自所述标记的参照分子的信号包含电子或光学储存的值或值集。

70.如权利要求64所述的方法,其中所述一个感兴趣的基因组片段或多个感兴趣的基因组片段包含性染色体或其至少一个片段,并且

其中所述一个参照基因组片段或多个参照基因组片段包含常染色体或其至少一个片段。

71.如权利要求64所述的方法,其中所述第一标记包含荧光标记、放射性标记、磁性标记或非光学标记中的至少一种。

72.如权利要求64所述的方法,其中进行标记包括:

用切口核酸内切酶在第一序列基序处使双链DNA的一条链产生切口，由此在所述双链DNA上产生切口；和

标记所述DNA。

73.如权利要求72所述的方法，还包括修复所述双链DNA上的至少一些切口。

74.如权利要求64所述的方法，其中所述多个标记的参照分子包含不同于所述第一标记的标记。

75.如权利要求64所述的方法，其中所述多个标记的参照分子包含与所述第一标记相同的标记。

76.如权利要求64所述的方法，其中用所述第一标记进行标记包括用甲基转移酶对所述样品分子的至少一个序列基序加标签。

77.如权利要求64所述的方法，其还包括用非序列特异性标记标记所述样品分子。

78.如权利要求64-77中任一项所述的方法，其中每个所述样品分子的长度小于500个核苷酸。

79.用于非诊断目的的表征包含多核苷酸序列的样品的方法，所述方法包括：

在样品分子的多核苷酸序列上标记多个序列特异性位置；

在流体通道中使所述样品分子的至少一部分线性化；

量化来自所述样品分子上的所述标记的信号；

将所述样品分子的多核苷酸序列的覆盖深度与参照分子的多核苷酸序列的覆盖深度进行比较，以确定所述样品分子的拷贝数；以及

当所述样品分子的覆盖深度与参照分子的覆盖深度不同时，确定所述样品分子中遗传异常是否存在。

80.如权利要求79所述的方法，其中所述样品包含循环胎儿细胞。

81.如权利要求79所述的方法，其还包括产生直方图分布，以反映样品的覆盖深度。

82.如权利要求79所述的方法，其中所述样品分子和所述参照分子来自相同生物体的不同组织。

83.如权利要求79所述的方法，其中来自所述参照分子的信号的数量包含电子或光学储存的值或值集。

84.如权利要求79所述的方法，其中所述遗传异常包括易位、添加、扩增、颠换、倒置、非整倍体、多倍体、单体性、三体性、21三体、13三体、14三体、15三体、16三体、18三体、22三体、三倍体、四倍体和性染色体非整倍体中的至少一种。

85.如权利要求79-84中任一项所述的方法，其中所述样品包含植物组织。

86.如权利要求79-84中任一项所述的方法，其中所述样品包含来源于环境样品的细菌或病毒体。

纳米流体中分子的表征

[0001] 相关申请的交叉引用

[0002] 本申请要求于2013年2月20日提交的美国临时申请第61767,219 号的权益，其全部内容通过引用并入本文。

[0003] 发明概述

[0004] 根据本文中的一些实施方案，提供了表征样品的方法。该方法可包括用至少第一标记标记多个样品分子，其中样品分子包含一个感兴趣的第一基因组片段或多个感兴趣的第一基因组片段的多核苷酸序列，并且其中一个感兴趣的第一基因组片段或多个感兴趣的第一基因组片段与样品的可能异常的基因组区域对应。该方法可以包括提供多个标记的参照分子，其中参照分子包含一个参照基因组片段或多个参照基因组片段的多核苷酸序列，并且其中已知一个参照基因组片段或多个参照基因组片段与可能异常的基因组区域不对应。本文所用的“与可能异常的基因组区域对应”和该字根(root term)的变型包括与异常染色体区域重叠或由其包括的基因组片段，包括但不限于复制、缺失、倒置、易位和/或非整倍体染色体或其片段。这样，基因组片段或片段可与存在的(例如，复制) 或不存在的(例如，缺失，例如如果该基因组片段包括在删除的区域中或与其重叠)的异常基因组区域对应。该方法可包括通过流体通道使多个标记的样品分子和多个标记的参照分子易位。该方法可以包括检测来自标记的样品分子的信号，从而确定一个感兴趣的第一基因组片段或多个感兴趣的第一基因组片段特有的至少一种第一模式或多种第一模模式；和一个参照基因组片段或多个参照基因组片段特有的一种第二模式或多种第二式。该方法可包括将确定一种第一模式或多种第一模式的信号与确定一种第二模式或多种第二模式的信号关联。在一些实施方案中，进行标记包括用第一标记标记样品分子，且其中参照分子包含第二标记，其中第一标记经配置产生一种第一模式或多种第一模式，并且其中第二标记经配置产生一种第二模式或多种第二模式，以及其中第一标记和第二标记互不相同。在一些实施方案中，进行标记包括用第一标记进行标记，其中一种第一模式或多种第一模式和一种第二模式或多种第二模式各包含第一标记，并且其中一种第一模式或多种第二模式和一种第二模式或多种第二模式互不相同。在一些实施方案中，该方法还包括标记参照分子，以产生标记的参照分子，其中标记的参照分子包含一种第二模式或多种第二模式。在一些实施方案中，标记的参照分子和样品分子来自样品。在一些实施方案中，标记的参照分子与样品分子来自相同生物体的不同组织。在一些实施方案中，标记的参照分子和样品分子来自不同的生物体。在一些实施方案中，来自标记的参照分子的信号包含电子或光学储存的值或值集。在一些实施方案中，该方法还包括用至少第一标记标记来自第二样品的多个第二样品分子，其中多个第二样品分子包含一个感兴趣的第一基因组片段或多个感兴趣的第一基因组片段的多核苷酸序列，其中已知多个第二样品分子与染色体异常不对应；通过流体通道使多个第二样品分子易位，以及检测来自标记的样品分子的信号，从而确定一个感兴趣的第一基因组片段或多个感兴趣的第一基因组片段特有的至少一种第一模式或多种第一模式，和一个参照基因组片段或多个参照基因组片段特有的一种第二模式或多种第二模式。在一些实施方案中，该方法还包括将所述模式的位置与参照基因组中模式的位置对齐。在一些实施方案

中,样品分子来自包含可能的基因组异常的样品,并且一个参照基因组片段或多个参照基因组片段包含第一染色体或其片段,其中参照基因组片段来自已知不包含基因组异常的第二样品。在一些实施方案中,一个感兴趣的第一基因组片段或多个感兴趣的第一基因组片段包含性染色体或其至少一个片段,并且一个参照基因组片段或多个参照基因组片段包含常染色体或其至少一个片段。在一些实施方案中,一个感兴趣的第一基因组片段或多个感兴趣的第一基因组片段包含第一常染色体或其至少一个片段,其选自:人类21号染色体、人类13号染色体、人类14号染色体、人类15号染色体、人类16号染色体、人类18号染色体和人类22号染色体以及其片段,并且一个参照基因组片段或多个参照基因组片段包含第二常染色体或其至少一个片段,其中第二常染色体或其片段与第一常染色体或其片段不同。在一些实施方案中,将信号关联包括使用由多个标记的样品分子或其部分产生的信号(S1,S2…Sn)和由参照产生的信号(C)之间的比例(K):K1＝S1/C,K2＝S2/C…Kn＝Sn/C。在一些实施方案中,第一标记包含荧光标记、放射性标记、磁性标记或非光学标记中的至少一种。在一些实施方案中,第二标记包含荧光标记、放射性标记、磁性标记或非光学标记中的至少一种。在一些实施方案中,进行标记包括用切口核酸内切酶在第一序列基序处使双链DNA的一条链产生切口;和标记该DNA。在一些实施方案中,该方法还包括修复 DNA上的至少一些切口。在一些实施方案中,不修复切口。在一些实施方案中,标记包含转录终止子。在一些实施方案中,用第一标记进行标记包括用DNA结合实体和甲基转移酶对样品分子的至少一种序列基序加标签,DNA结合实体选自:非切割限制酶、锌指蛋白、抗体、转录因子、转录激活剂样结构域、DNA结合蛋白、聚酰胺、形成三股螺旋的寡核苷酸以及肽核酸。在一些实施方案中,用第一标记进行标记包括用甲基转移酶对样品分子的至少一种序列基序加标签。在一些实施方案中,该方法还包括用非序列特异性标记标记样品分子。非序列特异性标记可以与第一标记和第二标记不同。在一些实施方案中,可能异常的基因组区域包括易位、添加、扩增、颠换、倒置、非整倍体、多倍体、单体性 (monosomy)、三体性、21三体、13三体、14三体、15三体、16三体、 18三体、22三体、三倍体、四倍体或性染色体非整倍体中的至少一种。在一些实施方案中,遗传异常包括三体性或单体性中的至少一种。

[0005] 根据本文中的一些实施方案,提供了表征样品的方法。该方法可以包括在样品分子的多核苷酸序列上标记多个序列特异性位置。该方法可以包括在流体通道中使样品分子的至少一部分线性化。该方法可包括量化来自样品分子上标记的信号。该方法可以包括将来自样品分子的信号的数量与来自参照分子的信号的数量进行比较。该方法可以包括当来自样品分子的信号的数量与由参照分子产生的信号的数量不同时,确定样品DNA中遗传异常是否存在。在一些实施方案中,样品分子和参照分子来自同一生物体。在一些实施方案中,样品分子和参照分子来自同一生物体的不同组织。在一些实施方案中,样品分子和参照分子来自不同生物体。在一些实施方案中,来自参照分子信号数量的信号包括电子或光学储存的值或值集。在一些实施方案中,样品分子包含DNA。在一些实施方案中,遗传异常包括易位、添加、扩增、颠换、倒置、非整倍体、多倍体、单体性、三体性、21三体、13三体、14三体、15三体、16三体、18三体、22三体、三倍体、四倍体或性染色体非整倍体中的至少一种。在一些实施方案中,遗传异常包括三体性或单体性中的至少一种。在一些实施方案中,进行标记包括用荧光标记、放射性标记、磁性标记或非光学标记中的至少一种标记多核苷酸。在一些实施方案中,进行标记包括:用切口核酸内切酶在第一序列基序处使双链DNA的一条链产生

切口;和标记该DNA。在一些实施方案中,进行标记还包括修复第一 DNA上的至少一些切口。在一些实施方案中,不修复切口。在一些实施方案中,标记包含转录终止子。在一些实施方案中,进行标记包括用DNA 结合实体和甲基转移酶对样品分子的至少一种序列基序加标签,DNA结合实体选自:非切割限制酶、锌指蛋白、抗体、转录因子、转录激活剂样结构域、DNA结合蛋白、聚酰胺、形成三股螺旋的寡核苷酸以及肽核酸。在一些实施方案中,用第一标记进行标记包括用甲基转移酶对样品分子的至少一种序列基序加标签。

[0006] 根据本文中的一些实施方案,提供了表征样品的方法。该方法可以包括标记样品核酸分子。该方法可以包括通过流体纳米通道使标记的样品核酸分子易位,其中流体纳米通道经配置延长样品核酸分子的至少一部分,并且其中流体纳米通道的长度为至少10nm,横截面直径为小于 1000nm。该方法可包括检测流体通道中由样品核酸分子产生的信号。该方法可包括确定样品核酸分子上标记的位置。该方法可包括将样品核酸分子上标记的位置与参照基因组中标记的位置对齐,其中参照基因组由来自与样品分子相同的生物体的第二样品获得。

[0007] 在一些实施方案中,本文中的任一种方法的流体纳米通道包括长度为至少10nm且横截面直径为小于5000nm的通道。在一些实施方案中,流体通道包括纳米通道。在一些实施方案中,将流体通道与基底表面平行设置。在一些实施方案中。在一些实施方案中,易位包括对标记的样品施加驱动力,驱动力选自流体流、辐射场、电渗透力、电泳力、电动力、温度梯度、表面性质梯度、毛细管流、压力梯度、磁场、电场、后退弯月面(receding meniscus)、表面张力、热梯度、拉力、推力及其组合。

[0008] 在一些实施方案中,本文中的任一种方法的样品选自细菌、病毒体、 DNA分子、RNA分子、核酸聚合物、蛋白质、肽和多糖。在一些实施方案中,本文中任一种方法的样品来源于母体血液,并且其中参照分子来源于除血液之外的母体样品。在一些实施方案中,本文中的任一种方法的样品包含核苷酸,并且其中至少两种标记位于核苷酸中感兴趣区域的任一端。在一些实施方案中,本文中的任一种方法的样品包含循环胎儿细胞、循环肿瘤细胞或体液或组织。

[0009] 在一些实施方案中,本文中的任一种方法包括光学检测,其包括测定实物计数(physical count)、强度、波长或标记的大小。在一些实施方案中,本文中的任一种方法包括光学检测,其包括测定样品中至少一个标记区域的长度。在一些实施方案中,本文中的任一种方法还包括测定由包含样品或样品部分的集合(pool)产生的信号。

[0010] 在一些实施方案中,本文中的任一种方法包括使用由多个样品或样品部分产生的信号(S1,S2…Sn)和由参照产生的信号(C)之间的比例 (K):K1=S1/C,K2=S2/C…Kn=Sn/C。在一些实施方案中,K1和Kn之间的差异用于鉴定胎儿样品的存在。在一些实施方案中,K1和Kn之间的差异用于鉴定来自肿瘤或其它癌源的DNA的存在。在一些实施方案中, K1和Kn之间的差异用于确定样品中遗传异常的存在。在一些实施方案中,遗传异常是非整倍体。在一些实施方案中,遗传异常是易位、添加、扩增、颠换或倒置。

[0011] 在一些实施方案中,本文中的任一种方法包括来源于已知二倍体或单倍体染色体的参照。在一些实施方案中,本文中的任一种方法包括将来自样品的信号与来自元基因组的或微生物组研究的种群分布关联。在一些实施方案中,本文中的任一种方法包括产生直方图分布,以反映样品的覆盖深度。

[0012] 在一些实施方案中，提供了表征样品的系统。该系统可以包括用至少两种标记标记样品分子的一个或多个区域。该系统可以包括用于使标记的样品分子易位的流体通道，其中该流体通道经配置延长样品分子的至少一部分，并且其中该流体通道的长度为至少10nm且横截面直径小于5000nm。该系统可包括用于检测流体通道中由标记的样品产生的信号的装置。

[0013] 在一些实施方案中，提供了表征样品的系统。该系统可以包括用于标记样品核酸分子的一个或多个区域。该系统可以包括用于使标记的样品核酸分子易位的流体纳米通道，其中流体纳米通道被配置来延长样品核酸分子的至少一部分，并且其中流体纳米通道的长度为至少10nm和横截面直径为小于1000nm。该系统可包括用于检测流体通道中由样品核酸分子产生的信号的装置。

[0014] 在一些实施方案中，提供了表征样品的系统。该系统可以包括用于标记样品DNA上多个序列特异性位置的区域。该系统可以包括用于使样品DNA的至少一部分线性化的区域。该系统可以包括用于量化由样品 DNA上标记产生的信号的装置。

[0015] 在一些实施方案中，提供了表征样品的系统。该系统可以包括用于用至少两种标记标记样品分子的部件(means)。该系统可以包括用于使标记的样品分子线性化的部件。该系统可以包括用于检测流体通道中由标记的样品产生的信号的部件。

[0016] 在一些实施方案中，提供了表征样品的系统。该系统可以包括用于标记样品核酸分子的部件。该系统可以包括用于使标记的样品核酸分子线性化的部件。该系统可包括用于检测流体通道中由样品核酸分子产生的信号的部件。

[0017] 在一些实施方案中，提供了表征样品的系统。该系统可以包括用于标记样品DNA上多个序列特异性位置的部件。该系统可以包括用于使样品DNA的至少一部分线性化的部件。该系统可以包括用于量化样品DNA 上由标记产生的信号的部件。

[0018] 在一些实施方案中，本文中描述的任一种系统能够表征选自细菌、病毒体、DNA分子、RNA分子、核酸聚合物、蛋白质、肽和多糖的样品。在一些实施方案中，本文中描述的任一种系统可以表征来源于母体血液的样品，并且其中参照分子来源于除血液之外的母体样品。在一些实施方案中，本文中描述的任一种系统能够表征包含核苷酸的样品，并且其中至少两种标记位于核苷酸中感兴趣的区域的任一端。在一些实施方案中，本文中描述的任一种系统能够表征包含循环胎儿细胞、循环肿瘤细胞或体液或组织的样品。

[0019] 在一些实施方案中，本文中描述的任一种系统可以包括选自荧光标记、放射性标记、磁性标记或其组合的标记。在一些实施方案中，本文中描述的任一种系统可经配置用于光学检测，其中光学检测包括测定实物计数、强度、波长或标记的大小。在一些实施方案中，光学检测包括测定样品中至少一种标记的区域的长度。在一些实施方案中，本文中描述的任一种系统经配置可以用于将信号关联，其中将信号关联包括测定由样品池或样品部分的集合产生的信号。在一些实施方案中，本文中描述的任一种系统经配置可以用于将信号关联，其中将信号关联包括使用由多种样品或样品部分产生的信号(S1,S2…Sn)和由参照产生的信号(C) 之间的比例(K):K1＝S1/C,K2＝S2/C…Kn＝Sn/C。在一些实施方案中，K1 和Kn之间的差异用于鉴定胎儿样品的存在。在一些实施方案中，K1和 Kn之间的差异用于鉴定来自肿瘤或其它癌源的DNA的存在。在一些实施方案中，K1和Kn之间的差异用于确定样品中遗传异常的存在。在一些实施方案中，遗传异常是非整倍体。在一些实施方案中，遗传异常

是易位、添加、扩增、颠换或倒置。

[0020] 在一些实施方案中，本文中描述的任一种系统可包括来源于已知二倍体或单倍体染色体的参照。

[0021] 在一些实施方案中，本文中描述的任一种系统可将来自样品的信号与来自元基因组的或微生物组研究的种群分布关联。

[0022] 在一些实施方案中，本文中描述的任一种系统的流体通道包括纳米通道。在一些实施方案中，将本文中描述的任一种系统的流体通道与基底表面平行设置。在一些实施方案中，易位包括对标记的样品施加驱动力，驱动力选自流体流、辐射场、电渗透力、电泳力、电动力、温度梯度、表面性质梯度、毛细管流、压力梯度、磁场、电场、后退弯月面、表面张力、热梯度、拉力、推力及其组合。

[0023] 在一些实施方案中，本文中描述的任一种系统经配置产生直方图分布，以反映样品的覆盖深度。

[0024] 在一些实施方案中，提供了用于实施本文中描述的任一种方法的试剂盒。

[0025] 在一些实施方案中，提供了用于使用本文中描述的任一种系统的试剂盒。

[0026] 附图简要说明

[0027] 图1是根据本文中的一些实施方案，说明流经纳米流体通道的样品分子或颗粒(卵形体)和参照或比较分子或颗粒(球体)的示意图。

[0028] 图2是说明检测由标记的分子或颗粒发射的信号以对样品和参照分子或颗粒的量、强度和结构(configuration)制表的成像设置的实施方案的示意图。

[0029] 图3a说明由PCR产生、在单个纳米流动通道中荧光着色、流动和成像的具有已知大小(233bp、498bp和834bp)的小双链DNA片段的一系列图像。图3b示出在相同纳米流动通道中混合在一起、流动和成像的相同双链DNA片段。在直方图中标绘荧光信号。

[0030] 图4是说明描绘由具有已知大小(233bp、498bp和834bp)的单独标记的DNA分子发射的光子的高斯曲线的一系列图。总计数和强度与质量和/或分子大小成线性比例。在线性动态范围内，可通过该方法推测未知分子大小和数量。

[0031] 图5是使用来自图4的信息在线性动态范围内说明对未知分子大小和数量的推测的一系列散点图。

[0032] 图6是说明按参照基因组(人类基因组版本19)标绘的基因组DNA片段的直方图。y-轴示出特异性染色体区域的覆盖深度。除了没有序列信号 (如着丝粒和端粒)的区域之外，观察到整个基因组的均匀分布。

[0033] 图7a是说明与1号染色体对齐的来自人类雄性样品的二倍体基因组片段的图。y-轴提供了覆盖数量。x-轴提供了核苷酸位置。平均覆盖深度为5X。图7b是示出显示出平均覆盖深度为2X-2.5X(大致为二倍体常染色体深度的一半)的来自相同雄性样品的单倍体性染色体X的图，显示了使用本文中一些实施方案的方法和平台的定量测量。

[0034] 发明详述

[0035] 胎儿的小DNA片段脱落进入母体血流。也已发现肿瘤将DNA释放到血液中。本文一些实施方案的方法是用于分析血液中的多核苷酸片段如DNA片段以检测来自胎儿或肿瘤的多核苷酸或细胞的存在。本文一些实施方案的方法还用于分析母体血液中胎儿DNA以检测遗传异常。在一些优选实施方案中，本文中描述的方法需要使用基于纳米流体的单分子检

测平台以鉴定遗传异常。本文一些实施方案的方法和装置具有分析小或大分子如小或大DNA分子的优点。在一些实施方案中,用至少一种模式标记感兴趣的分子或区域,且用至少一种模式标记感兴趣的参照分子或区域。该分子可在微流体通道中被线性化,且可将感兴趣的分子或区域的覆盖深度与参照分子的覆盖深度比较,以确定感兴趣的分子的拷贝数。

[0036] 据估计母体血液中约3-15%的短DNA是胎儿来源的。本文描述了使用并入流体学的方法而容易地检测和定量小分子包括短DNA片段的方法。在一些优选实施方案中,该方法包括在没有测序或组装的情况下量化短DNA片段。

[0037] 现有产前测试包括针穿刺以吸取羊水,可导致流产和其它并发症。此外,许多目前的癌检测方法也包括侵入性程序如活组织检查。根据本文中一些实施方案,提供了非侵入性的产前测试方法。在一些实施方案中,该方法用于测试血液。在一些实施方案中,该方法仅测试血液样品,且并不测试来自其它组织的样品。

[0038] 本文还描述了使用并入流体学的方法而检测较大分子包含较长DNA 片段并追踪其来源的方法。例如,在一些实施方案中,DNA片段可被追溯回至肿瘤或癌的其它来源。在一些优选实施方案中,该方法用于追踪 DNA片段的来源以鉴定或表征遗传异常。

[0039] 在一个优选实施方案中,分析来自母体血液样品的循环DNA以鉴定或定量相对母体基因组的胎儿DNA。在一些实施方案中,该信息用于在没有侵入性测试的情况下确定胎儿基因组健康状况(如21三体)。其全部内容通过引用并入本文的Yahya-Graison等(Yahya-Graison et al., Classification of Human Chromosome 21Gene-Expression Variations in Down Syndrome:Impact on Disease Phenotypes(唐氏综合症中人类21号染色体基因-表达变异的分类:对疾病显形的影响),Am J Hum Genet 2007, 81(3):475-491)描述的HSA21寡核苷酸试验中提供了用于检测非整倍体的试验中使用的适合寡核苷酸的实例。

[0040] 在一些实施方案中,将感兴趣的样品与参照样品进行比较。在一些实施方案中,感兴趣的样品来源于母体血液样品。在一些实施方案中,参照样品是来自除血液之外的来源的母体样品。在一些实施方案中,母体参照样品包含多核苷酸如从除血液之外的二倍体组织分离的DNA。在一些实施方案中,母体参照样品包含口腔样品、唾液(saliva)样品、尿样品、痰(sputum)样品或眼泪样品。例如,在一些实施方案中,与母体口腔样品比较,检测母体血液样品中的21三体。

[0041] 在一些实施方案中,在进行本文中描述的方法之前,为了胎儿核酸富集感兴趣的样品。例如,在一些实施方案中,使用可被抗体沉淀(pull down)的胎儿细胞特异性标记物富集胎儿细胞。在一些实施方案中,对感兴趣的样品进行大小分级。然而,可使用领域技术人员已知的任何富集方法。

[0042] 在一些实施方案中,感兴趣的样品来源于肿瘤细胞或疑似肿瘤细胞或与肿瘤细胞流体连通的组织(例如血液)。在一些实施方案中,参照样品是来自健康细胞的样品。在一些实施方案中,参照样品来自与肿瘤细胞或疑似肿瘤细胞相同的生物体的健康细胞。在一些实施方案中,参照样品选自几乎不包含或不可能包含肿瘤细胞的组织或来自肿瘤细胞的核酸。

[0043] 本领域技术人员应认识到,感兴趣的样品可包括来自各种来源的核酸。在一些实施方案中,感兴趣的样品包含来源于环境样品的细菌或病毒体、动物或植物组织、血液或其

它体液。在一些实施方案中，DNA片段用于检测染色体异常或癌基因组。

[0044]　如本领域技术人员应认识到，本文中描述的方法可用于制备和分析来自循环胎儿或肿瘤细胞的DNA。例如，在一些实施方案中，在分析前对细胞进行细胞裂解以释放感兴趣的DNA。

[0045]　在一些实施方案中，试验或分析了全基因组。在一些实施方案中，仅试验或分析了一部分基因组。在一些实施方案中，试验或分析了全部染色体。在一些实施方案中，仅试验或分析了一部分染色体。在一些实施方案中，分析了全基因。在一些实施方案中，仅试验或分析了一部分基因。

[0046]　本文中描述的信号可包括任何适合的信号，包括光学信号，荧光信号、非光学信号、辐射信号、电信号、磁性信号、化学信号或其任何组合。在一些实施方案中，信号由以下产生：电子自旋共振分子、荧光分子、化学发光分子、放射性同位素、酶底物、生物素分子、抗生物素蛋白分子、电荷转移分子、半导体纳米晶体、半导体纳米颗粒、胶质金纳米晶体、配体、微珠、磁珠、顺磁性颗粒、量子点、显色底物、亲和分子、蛋白质、肽、核酸、碳水化合物、抗原、纳米线、半抗原、抗体、抗体片段、脂质或其组合。

[0047]　在一些实施方案中，通过使用一种或多种激发源来诱导荧光性、化学发光、磷光、生物发光或其任何组合来产生信号。适合的激发源包括激光、可见光源、红外光源、紫外光源或其任何组合。

[0048]　在一些实施方案中，对核苷酸或有关的信号(例如荧光团)的检测是定量的。在一些实施方案中，量化核苷酸的长度。在一些实施方案中，量化分子大小。在一些实施方案中，信号的强度与分子的长度相关。例如，如图3a中所示，较长DNA分子可产生比短DNA分子强的信号。在一些实施方案中，信号的强度与样品或流体通道中DNA的量相关。

[0049]　在一些实施方案中，分析了样品的复制数变化，例如，如美国专利公开号20130034546所述，该专利公开的全部内容通过引用并入本文。

[0050]　特定分子如来源于不同染色体的DNA片段的数量可在本文提供的方法中定量测量。在一些实施方案中，据观察，来源于二倍体常染色体的基因组DNA的量是来源于二倍体性染色体的两倍。在一些实施方案中，这种片段的数量反映了来源染色体的拷贝数。在一些实施方案中，使用两种或三种颜色标记。

[0051]　在一些实施方案中，检测染色体来源的片段，且用相对比例鉴定非整倍体。在一些实施方案中，使用比例K1＝S1/C和K2＝S2/C计算核苷酸的拷贝数，其中K1是第一样品与对照样品的信号比例，而K2是第二样品与对照样品的信号比例。预期来自参照样品的拷贝数是整数，且K1和 K2之间的差可显示出一种感兴趣的样品中的异常。在一些实施方案，通过将特定样品的比例与来自多种样品的平均比例进行比较来检测异常。该方法还预期，对照基因组序列包括每个基因组总长已知的不同部分，其中感兴趣的序列包括每个正常基因的长度已知的不同部分，且其中K1 和K2之间的显著差表明基因组中的遗传异常。在一些实施方案中，感兴趣的核苷酸可与三体连锁的染色体相关，其中对照基因组序列来自除三体连锁的染色体以外的染色体，和其中大约2:3或3:2的K1/K2比例指示三体基因型。在一些实施方案中，感兴趣的核苷酸序列包含一部分基因组的缺失。在一些实施方案中，感兴趣的核苷酸序列包含重复序列。这样，根据本文中的一些实施方案可确定重复序列的拷贝数。在一些实施方案中，第一样品包含母体血液(在没有被任一种理论限制的情况下，其可包括胎儿核

酸)，而第二样品包含除血液以外的母体组织(优选为几乎不包含或不可能包含胎儿核酸的组织)。

[0052] 在一些实施方案中，进行数字计数检测。在一些实施方案中，对颗粒(如珠)、细菌或病毒体颗粒进行数字计数检测。如本领域技术人员应认识到的，本文中描述的方法可以应用于可被唯一标记的各种靶标。在一些实施方案中，进行数字核型分析。例如，在一些实施方案中，对具有感兴趣的潜在非整倍体的染色体进行数字核型分析。本文中描述的方法可用于检测任何感兴趣的染色体变异，包括易位、添加、扩增、颠换、倒置、非整倍体、多倍体、单体性、三体性、21三体、13三体、14三体、 15三体、16三体、18三体、22三体、三倍体、四倍体或性染色体的异常，包括但不限于XO、XXY、XYY和XXX。

[0053] 在一些实施方案中，本文提供了方法，其中该方法足够敏感以致可检测在长度上约为几十至几百个核苷酸的“短”片段。在一些实施方案中，本文中描述的样品分子包含多核苷酸“短”片段。例如，在一些实施方案中，多核苷酸片段的长度为约10、15、20、25、30、35、40、45、 50、55、60、65、70、75、80、85、90、95或100个核苷酸。在一些实施方案中，多核苷酸片段的长度为约100、125、150、175、200、225、 250、275、300、325、350、375、400、425、450、475、500、525、550、 575、600、625、650、675、700、725、750、775、800、825、850、875、 900、925、950、975或1000个核苷酸。在一些实施方案中，感兴趣的分子的长度小于约1000、950、900、850、800、750、700、650、600、550、 500、450、400、350、300、250、200、150、100或50个核苷酸。在一些实施方案中，片段是双链的。在一些实施方案中，片段包含DNA。在一些实施方案中，片段包含与RNA杂交的DNA。在一些实施方案中，敏感性约高达检测与靶片段有关的单个荧光团。

[0054] 在一些实施方案中，感兴趣的核苷酸的长度为至少约500个核苷酸的片段，例如约500、600、700、800、900、1000、1100、1200、1300、 1400、1500、1600、1700、1800、1900或2000个核苷酸，包括所列值中任意两个之间的范围，例如长度为约500至约2000、约500至约1500、约500至约1000、约500至约900、约500至约700、约700至约2000、约700至约1500、约700至约1000、约700至约900、约1000至约2000、约1000至约1500或约1500至约2000个核苷酸。

[0055] 适用于本文中描述的方法和系统的分子包含聚合物、双链DNA、单链DNA、RNA、DNA-RNA杂交体、多肽、生物分子、蛋白质等。适合的聚合物包括均聚物、共聚物、嵌段共聚物、无规共聚物、支链共聚物、树状聚体或其任何组合。

[0056] 在一些实施方案中，本文中描述的方法足够敏感以致可检测组成小于母体血液样品中分子总数的约025％、0.5％、0.75％、1％、1.25％、1.5％、 1.75％、2％、2.25％、2.5％、2.75％、3％、3.25％、3.5％、3.75％、4％、 4.25％、4.5％、4.75％、5％、6％、7％、8％、9％、10％、11％、12％、13％、 14％、15％、20％或25％的胎儿分子。

[0057] 在一些实施方案中，标记针对序列基序或化学部分。标记可使用本领域技术人员已知的任何技术进行，包括化学或生化缀合。在一些实施方案中，本文中描述的标记与独特的序列基序结合。在一些实施方案中，本文中描述的标记与化学部分结合。在一些实施方案中，化学部分与特异性染色体有关。

[0058] 在本文的一些实施方案中，各标记独立地选自荧光团、量子点、树状聚体、纳米线、珠、半抗原、链霉亲和素、抗生物素蛋白、中性亲和素、生物素和反应性基团。在本文的一些

实施方案中，第一和第二标记独立地选自荧光团或量子点。在本文的一些实施方案中，标记中的至少一种包含非光学标记。在本文的一些实施方案中，用聚合酶进行标记。在本文的一些实施方案中，在包含标记的dNTP的存在下用聚合酶进行标记。在本文的一些实施方案中，聚合酶具有5'至3'核酸外切酶活性。在本文的一些实施方案中，聚合酶离开侧翼区域(flap region)，并且其中在用连接酶修复之前，去除侧翼区域以恢复可连接的切口。在本文的一些实施方案中，使用聚合酶的5'至3'核酸外切酶活性在至少一个核苷酸以有限浓度存在的条件下去除侧翼区域。在本文的一些实施方案中，使用聚合酶的5'至3'核酸外切酶活性在至少一个核苷酸从反应中省略的条件下去除侧翼区域。在本文的一些实施方案中，用侧翼核酸内切酶(flap endonuclease)去除侧翼区域。在本文的一些实施方案中，在至少一种dNTP 存在的条件下用聚合酶进行标记。在本文的一些实施方案中，至少一种 dNTP是单一种类的dNTP。在本文的一些实施方案中，本文描述的方法还包括通过调节温度、dNTP浓度、辅助因子浓度、缓冲液浓度或其任何组合来调节标记期间聚合酶的活性。在本文的一些实施方案中，使第一基序或第二基序产生切口包括用Nt.BspQI产生切口。在本文的一些实施方案中，除本文中描述的序列特异性标记或多种标记之外，还应用了非序列特异性标记例如多核苷酸主链标记。

[0059] 在一些实施方案中，本文描述的至少一种标记包括非光学标记。各种非光学标记可与本文的实施方案结合使用。在一些实施方案中，非光学标记包含电子标记。示例性电子标记包括但不限于具有强电荷的分子，例如离子诸如金属离子、带电氨基酸侧链或其它阳离子或阴离子。当将标记置于检测器中时，可例如通过导电性(或电阻率)来检测电子标记。在一些实施方案中，纳米通道包含电极，该电极被配置来通过测定通道中设置的物质的导电性或电阻率来确定电子标记是否存在。在一些实施方案中，非光学标记包含金属、金属氧化物(例如金属氧化物)或氧化硅部分。在一些实施方案中，非光学标记包含含金属、金属氧化物或其它氧化物的部分(例如纳米颗粒)。可以例如通过核磁共振检测特定金属或氧化物部分的存在。在一些实施方案中，标记被配置为根据特定条件(例如，pH变化)释放部分，例如质子或阴离子，并检测释放的部分是存在还是不存在。

[0060] 在一些实施方案中，两种或多种标记互不相同。例如，可以用第一标记标记第一基序以产生第一唯一模式，并且可以用与第一标记不同的第二标记标记与第一基序不同的第二基序，以产生第二唯一模式。在一些实施方案中，两种或多种标记是相同的。例如，可以用标记标记第一基序，并且也可以用相同标记标记与第一基序不同的第二基序以产生唯一模式。在一些实施方案中，用第一标记标记与感兴趣的第一染色体或区域对应的多个探针，并且用与第一标记不同的第二标记标记与感兴趣的染色体或区域(例如参照染色体或区域)对应的多个第二探针。这样，可以基于是否用第一标记或第二标记标记样品分子将包含来自感兴趣的第一染色体或区域的序列的标记的样品分子与包含来自感兴趣的第二染色体或区域的序列的样品分子区分开。

[0061] 具有可逆终止子的核苷酸能形成第一磷酸二酯键，但在终止子逆转之前，不能形成(或具有有限能力以形成)第二磷酸二酯键。因此，具有可逆终止子的核苷酸可并入多核苷酸(例如在切口位点)，但直到终止子逆转，该核苷酸才可形成下游磷酸二酯键。可以使用本领域技术人员已知的技术进行逆转。例如，可通过能被切割的可切割连接子，例如通过电磁辐射，将终止子连接至核苷酸。如果使用包含3'可逆终止子的标记核苷酸进行切口修复，

可将单个标记的核苷酸并入切口，但终止子可防止其它标记的核苷酸并入切口。因此，可将切口标记限于每一切口一个标记的核苷酸。将切口标记限制为每一切口一个标记部分，能使被并入相同切口的多种标记的潜在偏差最小化。例如，如果采取方法将标记限制为每一切口一个标记部分，基于来自标记的相对强信号可解析非常接近的两个切口(即可以消除两种标记简单并入相同切口的可能性)。例如，如果需要定量评价切口数量，每一切口一个标记的方法可促进标记信号强度和切口数量之间的直接相关性。在含有可逆终止子的核苷酸上的标记可如本文中所述。在一些实施方案中，包含可逆终止子的核苷酸包含量子点。在一些实施方案中，包含可逆终止子的核苷酸包含荧光团。在一些实施方案中，包含可逆终止子的核苷酸包含非光学标记。

[0062] 在一些实施方案中，多种标记标记单一样品分子。在一些实施方案中，至少一种标记包含序列特异性标记。在一些实施方案中，至少一种标记包含非序列特异性标记。在一些实施方案中，至少一种标记包含序列特异性标记，并且至少一种标记包含非序列特异性标记。在一些实施方案中，至少一种标记不切割DNA的单链或双链。例如，在一些实施方案中，至少一种标记选自非切割限制酶、甲基转移酶、锌指蛋白、抗体、转录因子、DNA结合蛋白、发夹聚酰胺(hairpin polyamide)、形成三股螺旋的寡脱氧核苷酸、肽核酸或其组合。在一些实施方案中，序列特异性或非序列特异性标记都不切割DNA。

[0063] 在一些实施方案中，例如如果提供了荧光标记，使用敏感性照相机检测标记的进行。在一些实施方案中，例如如果提供了非光学标记，电子检测标记的进行。然而，可使用适用于对应标记的任一种检测方法。本文中描述的方法可包括在本文描述的分子的一个或多个区域中与荧光标记、放射性标记、磁性标记或其任何组合结合。可在标记与分子或分子的至少一部分或其它感兴趣的区域特异性互补的情况下完成结合。

[0064] 在一些实施方案中，切口酶产生随后使用例如标记的核苷酸或核核苷酸类似物标记的序列特异性切口。在一些实施方案中，荧光标记核苷酸或类似物。在一些实施方案中，通过在纳米通道中的限制使DNA线性化，从而导致均匀的线性化并允许严格且精确地测量DNA分子上包含特征模式(signature pattern)的切口标记之间的距离。在一些实施方案中，使用第二切口酶。在一些实施方案中，与第二标记色一起使用该第二切口酶。可根据本文中的实施方案使用的示例性切口酶包括但不限于 Nb.BbvCI、Nb.BsmI、Nb.BsrDI、Nb.BtsI、Nt.AlwI、Nt.BbvCI、Nt.BspQI、 Nt.BstNBI、Nt.CviPII及其组合。美国专利申请公开号2011/0171634和美国专利申请公开号2012/0237936中还提供了切口剂和实验规程的实例，上述专利申请的全部内容通过引用并入本文。

[0065] 在一些实施方案中，通过将探针与单链多核苷酸杂交，标记多核苷酸例如RNA或DNA。该探针可以与RNA或DNA的单链或其部分互补。在一些实施方案中，探针与特定序列基序互补。在一些实施方案中，提供多种探针以与多种特异性序列基序互补，例如至少2、3、4、5、6、7、 8、9、10、20、30、40、50、60、70、80、90、100、200、300、400、 500、600、700、800、900、1000、5,000或10,000个探针，包括所列值中任两种之间的范围。在一些实施方案中，探针具有随机序列。在一些实施方案中，提供了具有多种随机序列的探针。在一些实施方案中，探针包括有机荧光团、量子点、树状聚体、纳米线、珠、Au珠、顺磁性珠、磁性珠、放射性标记、聚苯乙烯珠、聚乙烯珠、肽、蛋白质、半抗原、抗体、抗原、链霉亲和素抗生物素蛋白、中性亲和素、生物素、核苷酸、寡核苷酸、序列特异性结合因子如工程限制酶、甲基转移酶、锌指结合

蛋白等中的一种或多种。在一些实施方案中，探针包括荧光团猝灭剂对。探针的一种配置可包括与探针第一端连接的荧光团和系至探针第二端的适合淬灭剂。这样，当探针未被杂交时，淬灭剂可防止荧光团发出荧光，而当探针与靶序列杂交时，探针被线性化，因此使淬灭剂远离荧光团并允许当由适当波长的电磁辐射激发的荧光团发荧光。在一些实施方案中，第一探针包括FRET对的第一荧光团，并且第二探针包括FRET对的第二荧光团。这样，第一探针和第二探针在彼此的FRET半径内与单一侧翼或与一对侧翼的杂交可允许通过FRET转移能量。在一些实施方案中，第一探针包括FRET对的第一荧光团，并且核苷酸上被并入以填充对应间隙的标记可包括FRET对的第二荧光团。这样，第一探针与侧翼的杂交和标记的核苷酸进入对应间隙允许通过FRET转移能量。

[0066] 在一些实施方案中，双链DNA可通过以下方式标记：首先通过增加温度或通过用有机溶剂操作使某些基因组区域双链之间的氢键熔融，以便打开所谓的D环，然后在退火回相对稳定形式之前，以等于或高于单链区域的亲和力与至少一种特异性探针杂交。这样，在一些实施方案中，在不使任一链产生切口或切割任一链的情况下，可通过本文描述的探针标记双链DNA。在一些实施方案中，多种D环可在单链上打开。这样，多种探针可退火至特定双链DNA。

[0067] 在一些实施方案中，进行标记包括通过甲基转移酶将标记转移至多核苷酸。在一些实施方案中，甲基转移酶特异性甲基化序列基序。这样，进行标记可包括通过甲基转移酶将标记转移至序列基序。示例性的适合的DNA甲基转移酶(MTase)包括但不限于M.BseCI(在5'-ATCGAT-3'序列内的N6处甲基化腺嘌呤)、M.Taql(在5'-TCGA-3'序列内的N6处甲基化腺嘌呤)和M.Hhal(在5'-GCGC-3'序列内的C5处甲基化第一胞嘧啶)。在一些实施方案中，两种或多种甲基转移酶提供了可以相同或不同的两种或多种标记。

[0068] 在一些实施方案中，进行标记包括通过甲基转移酶将标记转移至多核苷酸。在一些实施方案中，甲基转移酶特异性甲基化序列基序。这样，进行标记可包括通过甲基转移酶将标记转移至序列基序。示例性适合的 DNA甲基转移酶(MTase)包括但不限于M.BseCI(在5'-ATCGAT-3'序列内的N6处甲基化腺嘌呤)、M.Taql(在5'-TCGA-3'序列内的N6处甲基化腺嘌呤)和M.Hhal(在5'-GCGC-3'序列内的C5处甲基化第一胞嘧啶)。在一些实施方案中，两种或多种甲基转移酶提供了可以相同或不同的两种或多种标记。

[0069] 在一些实施方案中，通道包含微通道。在一些实施方案中，通道包含纳米通道。适合的流体纳米通道段的特征性横截面尺寸小于约1000 nm、小于约500nm或小于约200nm，或小于约100nm，或甚至小于约 50nm、约10nm、约5nm、约2nm，或甚至小于约0.5nm。适合的是，流体纳米通道段的特征性横截面尺寸不到分子回转半径的约两倍。在一些实施方案中，纳米通道的特征性横截面尺寸为至少约分子的持续长度。适用于本发明的流体纳米通道段的长度为至少约100nm、至少约500nm、至少约1000nm、至少约2微米、至少约5微米、至少约10微米、至少约1微米或甚至至少约10mm。在一些实施方案中，流体纳米通道段以以每立方分米至少1流体纳米通道段的密度存在。

[0070] 流体通道的实例可见于美国专利公开号2008/0242556中，该专利公开的全部内容通过引用并入本文。在一些实施方案中，分析了病毒体颗粒或细菌颗粒。例如，在一些实施方案中，使用微通道分析细菌细胞。在一些实施方案中，通道允许直径范围为几微米至几十微米的细胞流过。

[0071] 图1是根据本文中的一些实施方案，说明流体通道布置的示意图。该布置可包括样品输入室10。该布置可包括流体通道阵列12，例如流体纳米通道。该布置可包括样品输出室14。输出室可包含缓冲溶液16。纳米流体通道阵列12可以与输入室10流体连通。纳米流体通道阵列12可以与输出室14流体连通。感兴趣的样品分子或颗粒18可以设置在纳米流体通道阵列10中。感兴趣的对照或比较分子或颗粒18可以设置在纳米流体通道阵列10中。在一些实施方案中，流体通道阵列12将输入室 10与输出室14连接。在一些实施方案中，将感兴趣的样品分子或颗粒 18和感兴趣的对照或比较分子或颗粒20装载至样品输入室，并通过纳米流体通道阵列在缓冲溶液16中行进。在一些实施方案中，感兴趣的样品分子或颗粒18和感兴趣的对照或比较分子或颗粒20从流体通道阵列12 存放至样品输出室14。

[0072] 图2是根据本文中的一些实施方案说明用于检测感兴趣的样品分子或颗粒的布置的示意图。在一些实施方案中，所述布置包括第一样品入口或出口11、第二样品入口或出口11以及位于其之间并与第一和第二入口或出口11中每个流体连通的至少一种流体通道13。据本文预期，如果样品被装载至第一入口或出口11，该第一入口或出口11发挥入口的作用，并且第二入口或出口11可发挥出口的作用。据本文预期，如果样品被装载至第二入口或出口11，第二入口或出口11发挥入口作用，并且第一入口或出口11可发挥出口作用。在一些实施方案中，样品包含感兴趣的分子或颗粒18、感兴趣的对照或比较颗粒20或两种的组合。在一些实施方案中，感兴趣的分子或颗粒18、感兴趣的对照或比较颗粒20经过流体通道13。在一些实施方案中，流体通道13包含纳米通道。在一些实施方案中，流体通道13包含微通道。在一些实施方案中，流体通道13包含检测区22。在一些实施方案中，该系统包括设置在检测区域24上的覆盖物 (cover)24。在一些实施方案中，覆盖物24包含透明盖。在一些实施方案中，检测器26位于检测区域22和覆盖物24(如果存在的话)上。在一些实施方案中，例如，如果使用光学检测，检测器26包括光子检测/成像器。在一些实施方案中，透镜28以与检测区域22和检测器26光学连通定位。在一些实施方案中，透镜28位于检测区域22和检测器26之间。在一些实施方案中，双色镜30以与检测区域22、透镜28、检测器26和激发源 32光学连通定位，从而可以激发荧光标记(如果存在的话)，并且可以检测来自荧光标记(如果存在的话)的荧光。

[0073] 在一些实施方案中，以直方图的形式提供样品与参照样品的比较。在一些实施方案中，在直方图分布中对具有匹配参照或从头电子测序基因组组装的特定标记模式的分子的实物计数制表，以反映覆盖深度。如在遗传病的非整倍体或癌症中的结构变异的情况下，特定区域或全部染色体中比平均覆盖深度高或低反映出与正常倍性的偏差。

[0074] 其它可选实施方案

[0075] 本文中描述的一些实施方案可包括以下内容：表征样品的方法，包括：用至少两种标记标记样品分子的区域；通过流体通道使标记的样品分子易位，其中流体通道经配置延长样品分子的至少一部分，并且其中流体通道的长度为至少10nm且横截面直径为至少5000nm；检测流体通道中由标记的样品产生的信号；并且将由标记的样品产生的信号与由参照分子的对应区域产生的信号关联。该方法还可以包括：标记与样品分子的区域对应的参照分子的区域；通过流体通道使标记的参照样品分子易位，其中流体通道经配置延长样品分子的至少一部分，并且其中流体通道的长度为至少10nm且横截面直径为小于5000nm；以及检测流体通道中由标记的参照样品产生的信号，其中由参照分子的已知对应区域产生

的信号是由标记的参照样品产生的信号。

[0076] 在一些实施方案中，提供了表征样品的方法。该方法可以包括：标记样品核酸分子；通过流体纳米通道使标记的样品核酸分子易位，其中流体纳米通道经配置延长样品核酸分子的至少一部分，并且其中流体纳米通道的长度为至少10nm，且横截面直径为小于1000nm；检测流体通道中由样品核酸分子产生的信号；确定样品核酸分子上标记的位置；以及将样品核酸分子上标记的位置与参照基因组中标记的位置对齐。

[0077] 在一些实施方案中，提供了表征样品的方法。该方法可以包括：处理双链DNA样品，以便生成从双链DNA样品移开的双链DNA样品的第一链的侧翼，其中侧翼的长度范围为约1个碱基至约1000个碱基，并且其中侧翼在与侧翼对应的双链DNA样品的第一链中引起间隙；将一个或多个碱基并入双链DNA，以清除间隙的至少一部分；用一种或多种标签标记处理的双链DNA的至少一部分；以及量化由双链DNA上的标记产生的信号；将由双链DNA产生的信号的数量与由参照DNA产生的信号的数量进行比较；以及当由双链DNA产生的信号的数量与由参照DNA 产生的信号的数量不同时，确定双链DNA中遗传异常的存在。

[0078] 在一些实施方案中，提供了表征样品的方法。该方法可以包括标记样品DNA上的多个序列特异性位置；使样品DNA的至少一部分线性化；量化由样品DNA上的标记产生的信号；将由样品DNA产生的信号的数量与由参照DNA产生的信号的数量进行比较；以及当由样品DNA产生的信号的数量与由参照DNA产生的信号的数量不同时，确定样品DNA 中遗传异常的存在。

[0079] 在一些实施方案中，提供了表征样品的系统。该系统可以包括：用于用至少两种标记标记样品分子的一个或多个区域；用于使标记的样品分子易位的流体通道，其中流体通道经配置延长样品分子的至少一部分，并且其中流体通道的长度为至少10nm且横截面直径为小于5000nm；以及用于检测流体通道中由标记的样品产生的信号的装置。

[0080] 在一些实施方案中，提供了表征样品的系统。该系统中以包括：用于标记样品核酸分子的一个或多个区域；用于使标记的样品核酸分子易位的流体纳米通道，其中流体纳米通道经配置延长样品核酸分子的至少一部分，并且其中流体纳米通道的长度为至少10nm且横截面直径为小于1000nm；以及用于检测流体通道中由样品核酸分子产生的信号的装置。

[0081] 在一些实施方案中，提供了表征样品的系统。该系统可以包括：用于处理双链DNA样品以便生成从双链DNA样品移开的双链DNA样品的第一链的侧翼的一个或多个区域，其中侧翼的长度的范围为约1个碱基至约1000个碱基，并且其中侧翼在与侧翼对应的双链DNA样品的第一链中引起间隙；用于将一个或多个碱基并入双链DNA以消除间隙的至少一部分的一个或多个区域；用于用一种或多种标签标记处理的双链 DNA的至少一部分的一个或多个区域；以及用于量化由双链DNA上标记产生的信号的装置。

[0082] 在一些实施方案中，提供了用于表征样品的系统。该系统可以包括：用于标记样品DNA上多个序列特异性位置的区域；用于使样品DNA的至少一部分线性化的区域；以及用于量化由样品DNA上标记产生的信号的装置。

[0083] 在一些实施方案中，提供了用于表征样品的系统。该系统可以包括：用于用至少两种标记标记样品分子的部件；用于使标记的样品分子线性化的部件；以及用于检测流体通道中由标记的样品产生的信号的部件。

[0084] 在一些实施方案中，提供了表征样品的系统。该系统可以包括：标记样品核酸分子

的部件;用于使标记的样品核酸分子线性化的部件;以及用于检测流体通道中由样品核酸分子产生的信号的部件。

[0085] 在一些实施方案中,提供了表征样品的系统。该系统可包括:用于处理双链DNA样品以生成从双链DNA样品移开的双链DNA样品的第一链的侧翼的部件,其中侧翼的长度的范围为约1至约1000个碱基,并且其中侧翼引起与侧翼对应的双链DNA样品的第一链中的间隙;用于将一个或多个碱基并入双链DNA以消除间隙的至少一部分的部件;用于用一种或多种标签标记处理的双链DNA的至少一部分的部件;以及用于量化由双链DNA上标记产生的信号定量的部件。

[0086] 在一些实施方案中,提供了表征样品的系统。该系统可以包括:用于表征样品的系统,其包括用于标记样品DNA上多个序列特异性位置的部件;使样品DNA的至少一部分线性化的部件;以及量化由样品DNA 上标记产生的信号的部件。

[0087] 根据一些实施方案,提供了本文描述的方法或系统,其中样品选自细菌、病毒体、DNA分子、RNA分子、核酸聚合物、蛋白质、肽以及多糖。

[0088] 根据一些实施方案,提供了本文描述的方法或系统,其中样品来源于母体血液,并且其中参照分子来源于除血液之外的母体样品。

[0089] 根据一些实施方案,提供了本文描述的方法或系统,其中样品包含核苷酸,并且其中至少两种标记位于核苷酸中感兴趣区域的任一端。

[0090] 根据一些实施方案,提供了本文描述的方法或系统,其中标记选自荧光标记、放射性标记、磁性标记或其组合。

[0091] 根据一些实施方案,提供了本文描述的方法或系统,其中光学检测包括测定实物计数、强度、波长或标记的大小。

[0092] 根据一些实施方案,提供了本文描述的方法或系统,其中光学检测包括测定样品中至少一种标记的区域的长度。

[0093] 根据一些实施方案,提供了本文描述的方法或系统,其中将信号关联包括测定由样品集合或部分样品的集合产生的信号。

[0094] 根据一些实施方案,提供了本文描述的方法或系统,其中将信号关联包括使用由多种样品或样品部分产生的信号(S1,S2…Sn)与由参照产生的信号(C)之间的比例(K):K1＝S1/C,K2＝S2/C…Kn＝Sn/C。在一些实施方案中,K1和Kn之间的差异用于鉴定胎儿样品的存在。在一些实施方案中,K1和Kn之间的差异用于鉴定来自肿瘤或其它癌源的DNA的存在。在一些实施方案中,K1和Kn之间的差异用于确定样品中遗传异常的存在。在一些实施方案中,遗传异常是非整倍体。在一些实施方案中,遗传异常是易位、添加、扩增、颠换或倒置。在一些实施方案中,参照来源于已知二倍体或单倍体染色体。在一些实施方案中,将来自样品的信号与来自元基因组的或微生物组研究的种群分布关联。

[0095] 根据一些实施方案,提供了本文描述的方法或系统,其中流体通道是纳米通道。在一些实施方案中,将流体通道与基底表面平行设置。在一些实施方案中,

[0096] 根据一些实施方案,提供了本文描述的方法或系统,其还包括生成直方图分布,以反映样品的覆盖深度。

[0097] 根据一些实施方案,提供了本文描述的方法或系统,其中样品包含循环胎儿细胞、循环肿瘤细胞或体液或组织。

[0098] 根据一些实施方案，提供了本文描述的方法或系统，其中易位包括对标记的样品施加驱动力，驱动力选自流体流、辐射场、电渗透力、电泳力、电动力、温度梯度、表面性质梯度、毛细管流、压力梯度、磁场、电场、后退弯月面、表面张力、热梯度、拉力、推力及其组合。

[0099] 根据一些实施方案中，提供了用于实施本文描述的方法的试剂盒。

[0100] 根据一些实施方案中，提供了用于使用前述权利要求中任一项所述系统的试剂盒。

[0101] 在本文提供的描述中，参考了构成说明书的一部分的附图。详细的描述、附图和权利要求中描述的说明性的实施方案不应是限制性的。在不偏离本文所列主题的精神或范围的情况下，可以利用其它实施方案，并且可以进行其它变化。应当容易地理解的是，本发明的方面，如本文中一般描述和附图中说明的，可以以多种不同配置被设置、替换、结合或设计，其所有都是可清楚地预期的并且都是本发明的一部分。

[0102] 除非另有定义，本文使用的技术和科学术语具有本发明所属领域中普通技术人员通常所理解的相同含义。

[0103] 本文中使用的术语"通道"指的是由边界限定的区域。这种边界可以是物理的、电学的、化学的、磁性的等。术语"纳米通道"用于说明某些通道被认为是某些尺寸的纳米级。

[0104] 本文中使用的术语"DNA"指的是任何长度(例如0.1Kb至1巨碱基) 的DNA。DNA可以是高纯度制剂、原材料或半原材料。DNA可来自任何生物源或可以是合成的。

[0105] 本文中使用的术语"核苷酸"指的是含有脱氧核糖核酸(例如DNA、 mtDNA、gDNA或cDNA)、核糖核酸(例如RNA或mRNA)或本领域中已知的核酸的任何其它变型的分子。术语"标记的核苷酸"指的是包含可检测的任何修饰的核苷酸。这包括但不限于具有与碱基相连的报告基因的核苷酸。报告基因包括但不限于荧光染料、半抗原、生物素分子或金纳米颗粒。术语"天然核苷酸"指的是未被修饰的核苷酸，或具有不干扰其并入DNA的轻微修饰。术语"t"、"c"、"a"、"g"和"u"指DNA中的核苷酸。

[0106] 术语"切口"指的是具有3'羟基端的发生于一条DNA链或另一条链上的磷酸二酯键断裂。

[0107] 本文中使用的术语"切口核酸内切酶"指的是天然存在的或工程改造的能破坏离开DNA单链上的磷酸二酯键从而在限定的序列留有3'-羟基的任何酶。切口核酸内切酶可以是天然存在的，通过修饰限制酶以消除一种DNA链的切割活性来工程改造，或通过将切口亚单位融合至DNA 结合域例如锌指和转录激活剂样效应子DNA识别域来产生。

[0108] 本文中使用的术语"标记位点"指的是具有暴露的3'羟基基团的任何DNA位点，在所述基团上聚合酶可以模板依赖性方式增加核苷酸。标记位点可以通过切口核酸内切酶、杂交探针或破坏任一DNA链上的磷酸二酯键的任何化学或物理手段产生。可以对在其生物源外的DNA或在 DNA提取之前发生破坏磷酸二酯键的手段，例如由于生物样品暴露于化学物质或外力如辐射。如果3'端不可延长，可以例如通过使用新英格兰生物实验室PreCR试剂盒来进行修复以恢复羟基。

[0109] 本文中使用的"样品"可包括例如血液、血清、血浆、唾液、灌洗流体、脑脊液流体、尿、精液、汗液、眼泪、痰等。本文中使用的"血液"、"血浆"和"血清"清楚地包含其成分(fraction)或处理部分，其中样品取自活组织检查、棉签(swab)、涂片等，"样品"清楚地包含来源于活组织检查、拭子、涂片等处理成分或部分。

[0110] 本文中使用的术语“染色体”指的是来源于染色质且包含DNA和蛋白成分(特别是组蛋白)的活细胞的承载遗传(heredity-bearing)的基因携带者。

[0111] 本领域技术人员应当认识到,当在使DNA分子经过纳米通道的情况下,“易位”可以与线性化交换使用。

[0112] 本文描述的方法、装置、系统和试剂盒等可并入以下参考文献的任一个中描述的方法、装置、系统和试剂盒:美国专利申请公开号2009/0305273、PCT公开号WO/2008/079169、美国专利申请公开号 2008/0242556、PCT公开号WO/2008/121828、美国专利申请公开号 2011/0171634、PCT公开号WO/2010/002883、美国专利申请公开号 2011/0296903、PCT公开号WO/2009/149362、美国专利申请公开号 2011/0306504、PCT公开号WO/2010/059731、美国专利申请公开号 2012/0097835、PCT公开号WO/2010/135323、PCT申请号 PCT/US11/57115、美国专利申请序列号13/606819、PCT申请号 PCT/US2012/054299、美国专利申请公开号2012/0244635、PCT公开号 WO/2011/038327、美国专利申请公开号2012/0237936、美国专利申请序列号13/503307、PCT公开号WO/2011/050147、美国专利申请序列号 61/734327、美国专利申请序列号61/761189以及美国专利申请序列号 61/713862,其全部内容通过引用并入本文的。

[0113] 实施例1

[0114] 通过PCR,由人类雄性样品产生基因组片段,进行标记并经过纳米通道。然后将检测的片段与用于各染色体的单一基因参照光学图对齐。基于对齐开始位点将分子分类。

[0115] 如图7a 中所示的,观察的二倍体常染色体(染色体1)的平均覆盖深度是5X,并且平均地分布于染色体。如果对分子的采样均匀,对齐开始位点会在整个染色体内任意分布,产生线性图。

[0116] 如图7b 中所示的,观察的来自相同雄性样品的单倍体性染色体(染色体X)的平均覆盖深度是2X-2.5X(大致为二倍体常染色体深度的一半),并且也平均地分布于整个染色体。该实施例说明使用本文中描述的方法和平台可实现定量测量。

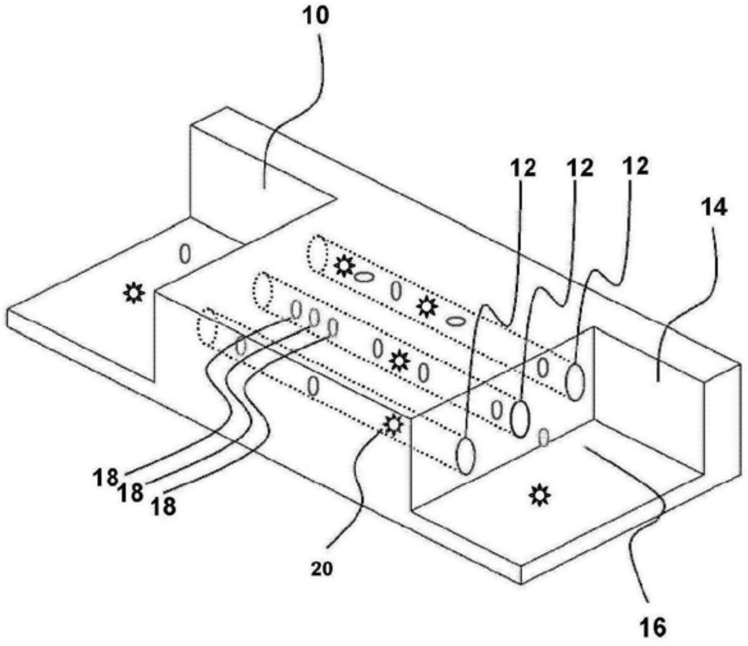


图1

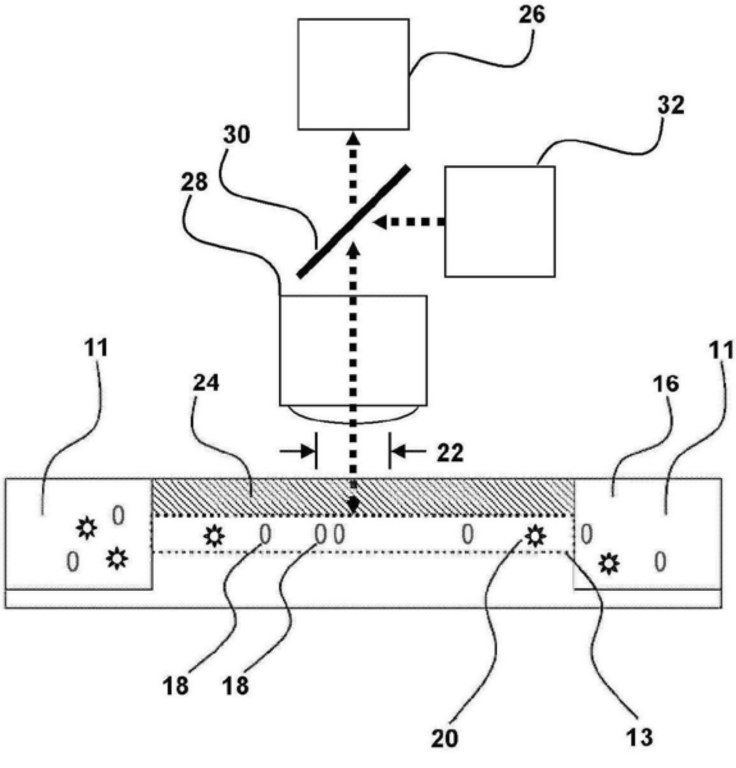


图2

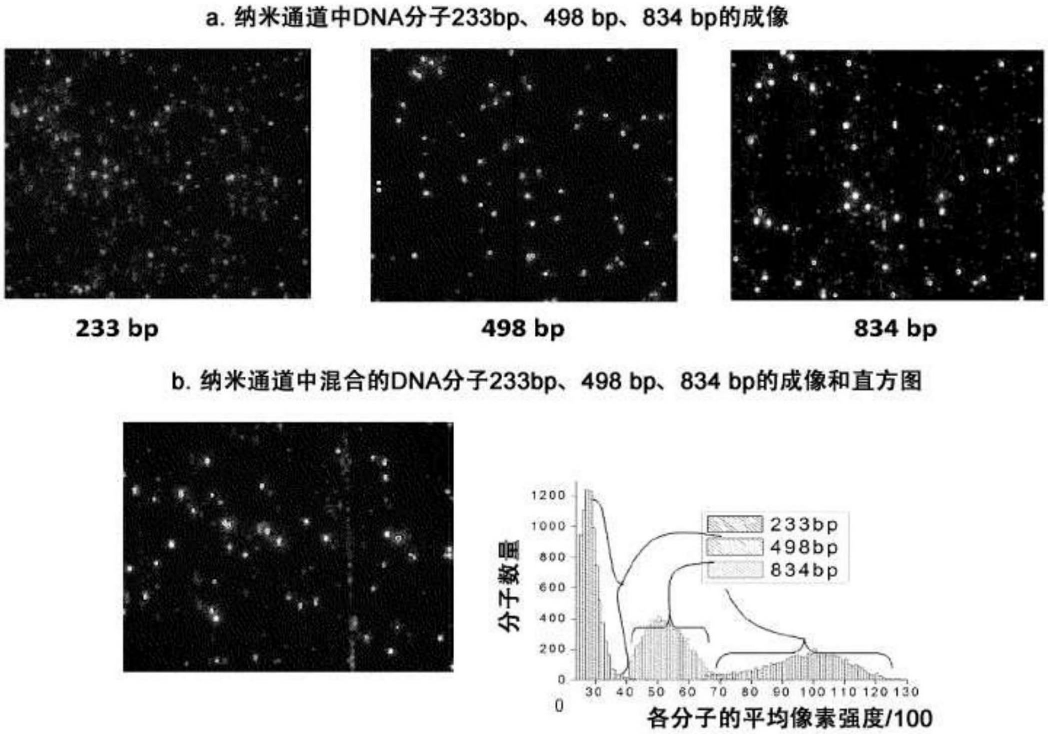

a. 纳米通道中DNA分子233bp、498 bp、834 bp的成像
233 bp
498 bp
834 bp
b. 纳米通道中混合的DNA分子233bp、498 bp、834 bp的成像和直方图
1200
1000
800
600
400
200
0
分子数量
233bp
498bp
834bp
30 40 50 60 70 80 90 100 110 120 130
0
各分子的平均像素强度/100

图3

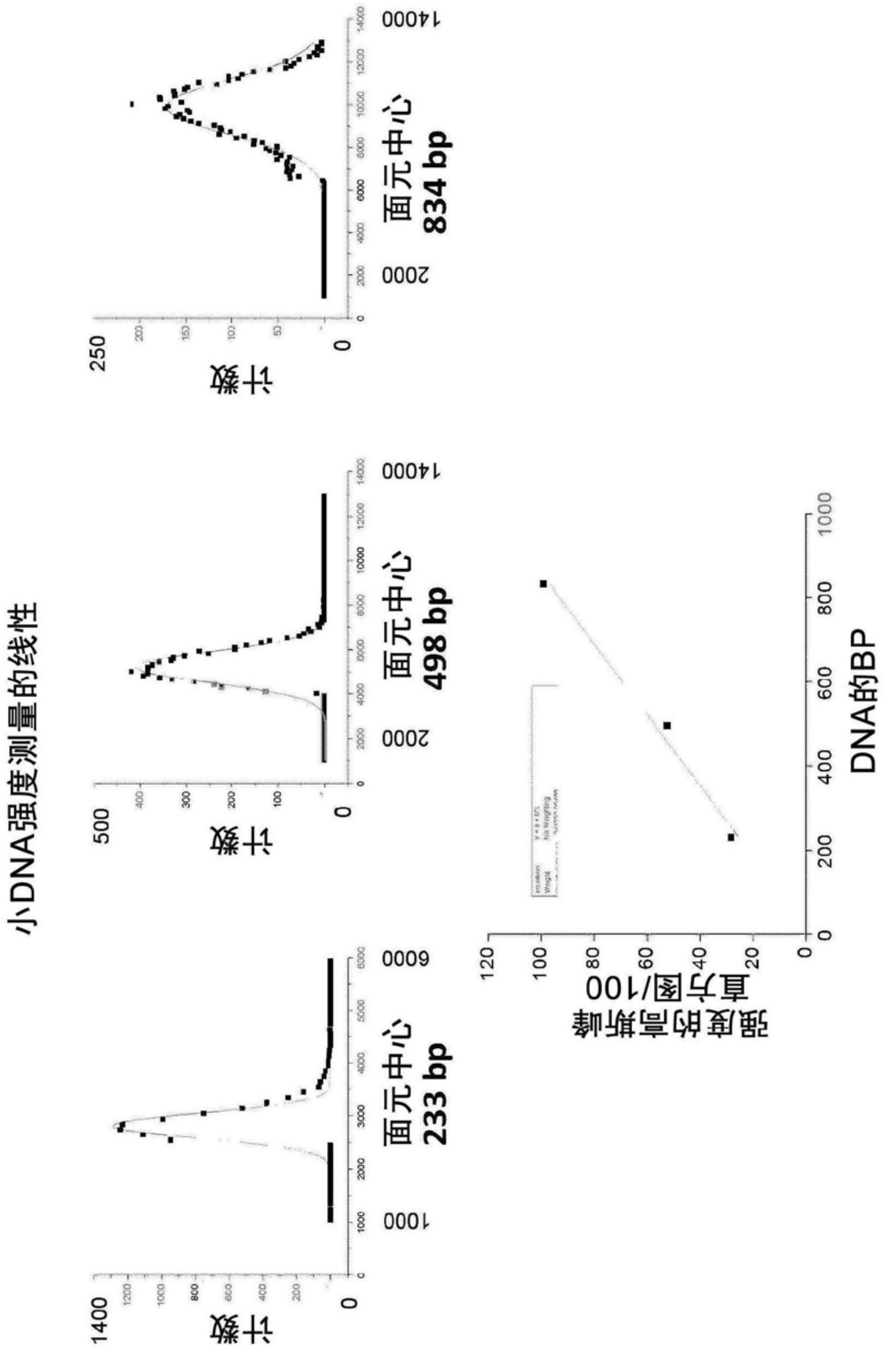

小DNA强度测量的线性
计数
1400
0
1000
6000
面元中心
233 bp
计数
500
0
2000
14000
面元中心
498 bp
计数
250
0
2000
14000
面元中心
834 bp
强度的高斯峰
直方图/100
120
100
80
60
40
20
0
0
200
400
600
800
1000
DNA的BP

图4

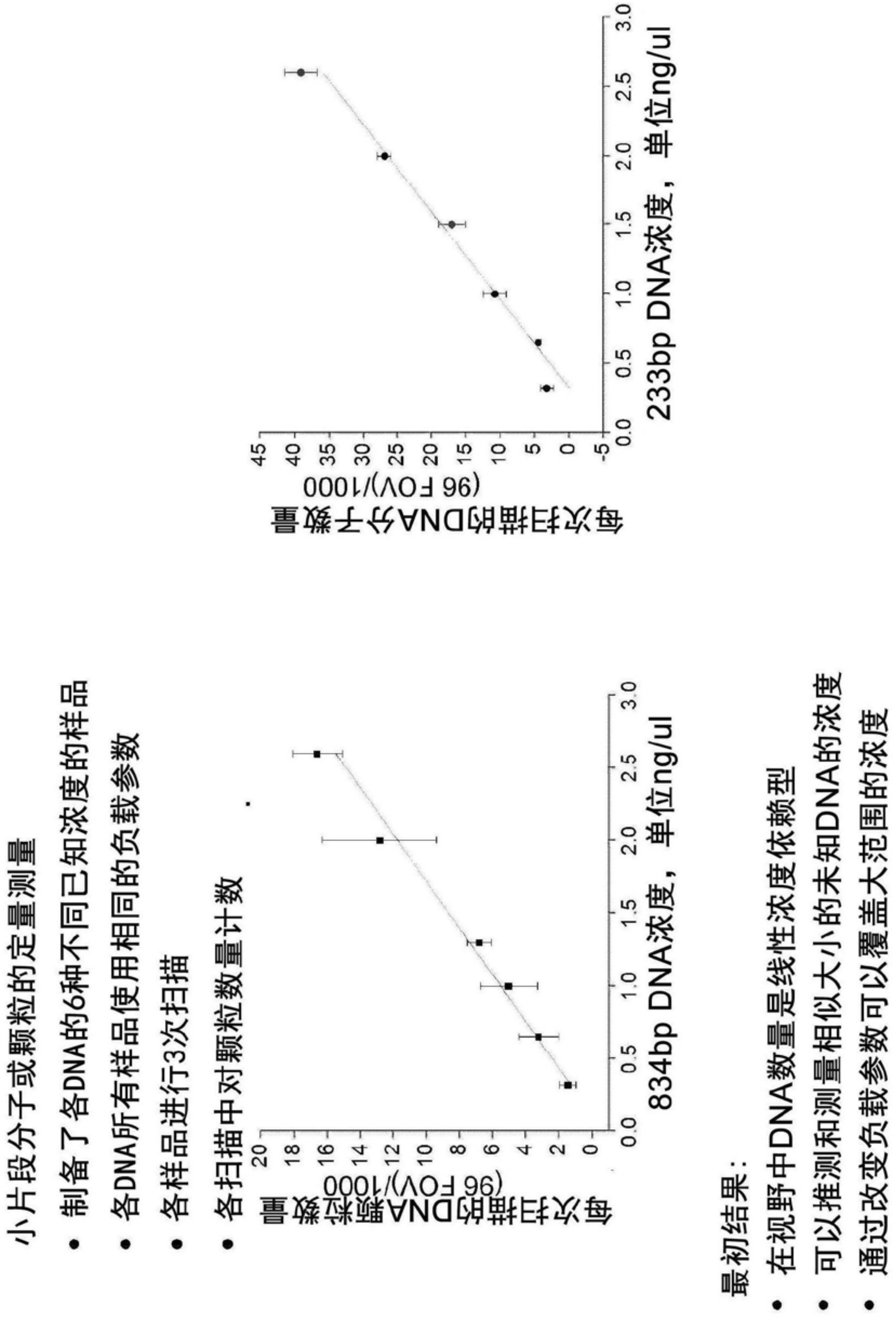

小片段分子或颗粒的定量测量
制备了各DNA的6种不同已知浓度的样品
各DNA所有样品使用相同的负载参数
各样品进行3次扫描
各扫描中对颗粒数量计数
每次扫描的DNA颗粒数量
(96 FOV)/1000
834bp DNA浓度，单位ng/ul
每次扫描的DNA分子数量
(96 FOV)/1000
233bp DNA浓度，单位ng/ul
最初结果：
在视野中DNA数量是线性浓度依赖型
可以推测和测量相似大小的未知DNA的浓度
通过改变负载参数可以覆盖大范围的浓度

图5

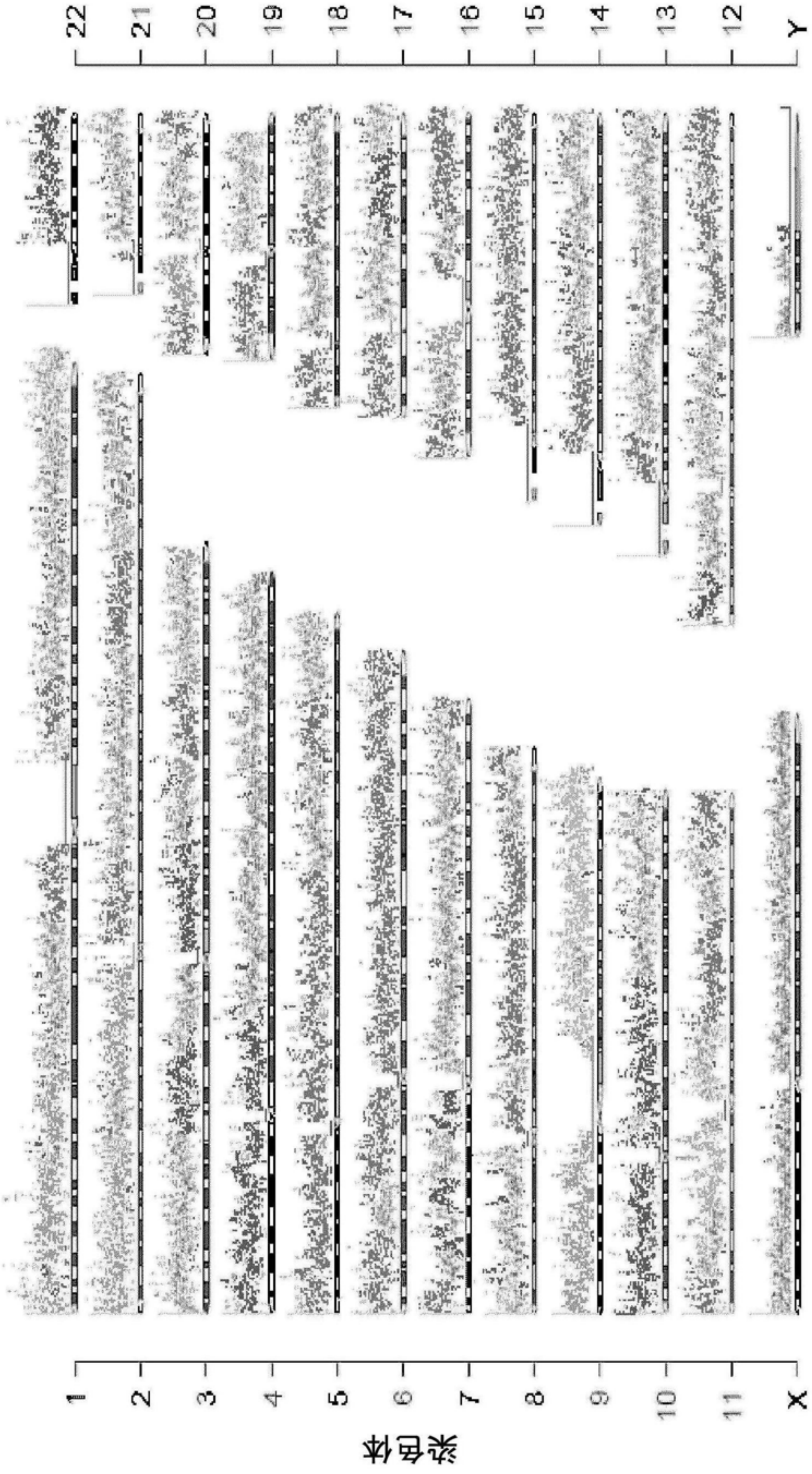


图6

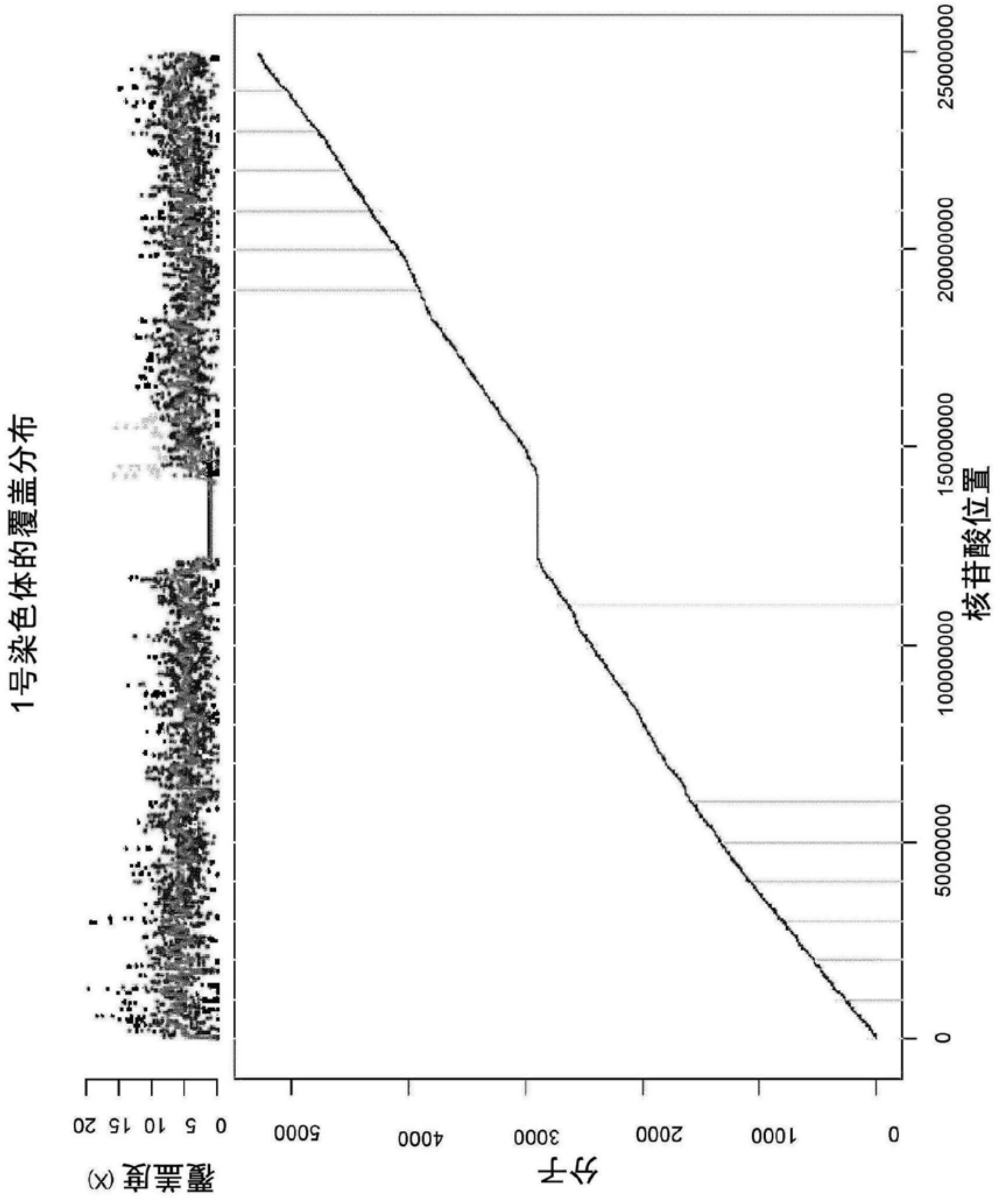

1号染色体的覆盖分布
核苷酸位置
0
50000000
100000000
150000000
200000000
250000000
分子
0
1000
2000
3000
4000
5000
覆盖度 (X)
0
5
10
15
20

图7a

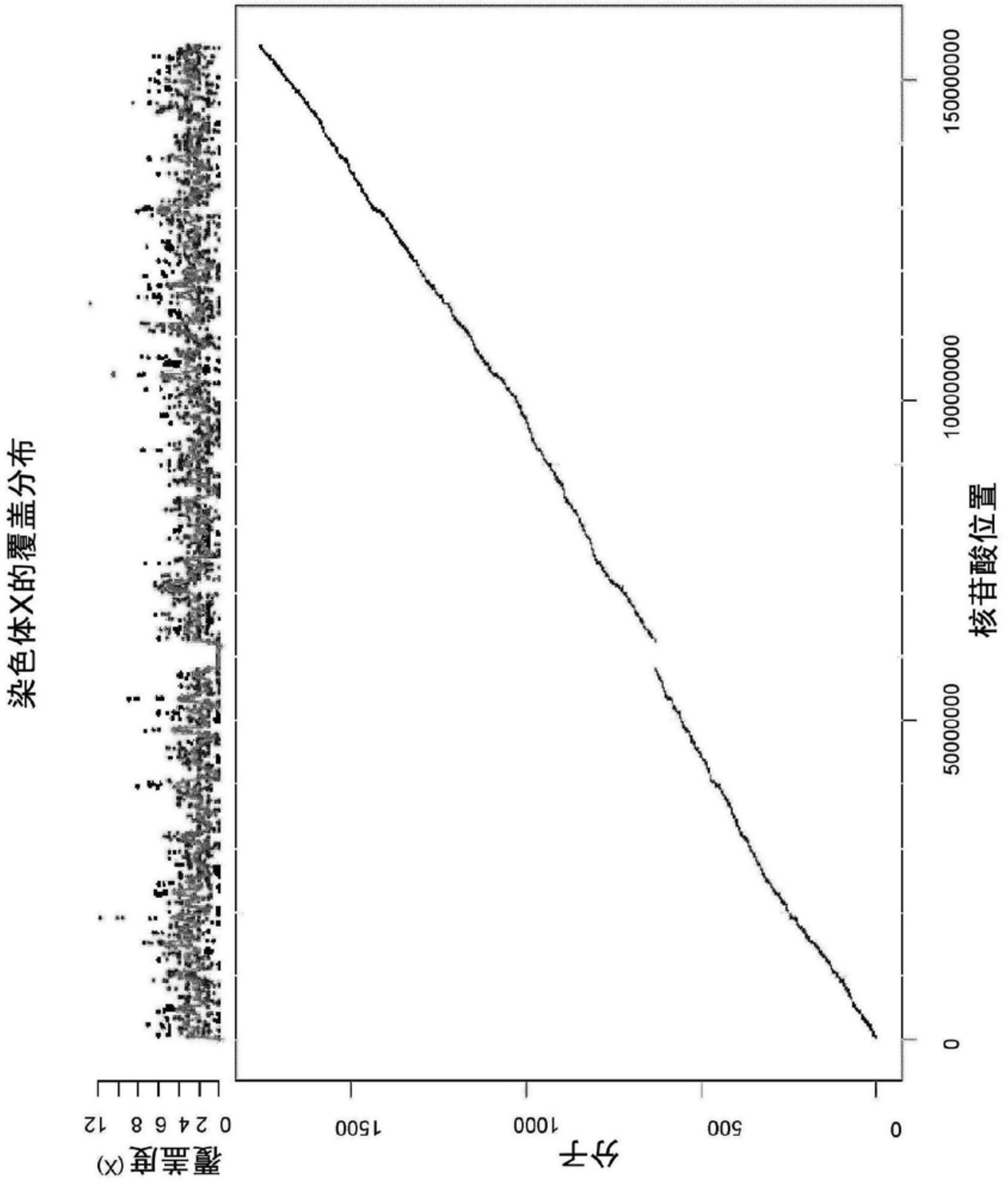

染色体X的覆盖分布
覆盖度(X)
0 2 4 6 8 12
分子
0
500
1000
1500
核苷酸位置
0
50000000
100000000
150000000

图7b