



(12)发明专利申请

(10)申请公布号 CN 108073790 A

(43)申请公布日 2018.05.25

(21)申请号 201610991671.8

(22)申请日 2016.11.10

(71)申请人 安诺优达基因科技(北京)有限公司

地址 100176 北京市北京经济技术开发区
科创六街88号院8号楼2单元701室

申请人 浙江安诺优达生物科技有限公司
安诺优达(义乌)医学检验有限公司

(72)发明人 刘卉 侯光远 郭璟 王秀莉
魏少华 玄兆伶 李大为 梁峻彬
陈重建

(74)专利代理机构 北京坤瑞律师事务所 11494
代理人 闫桑田

(51)Int.Cl.
G06F 19/22(2011.01)

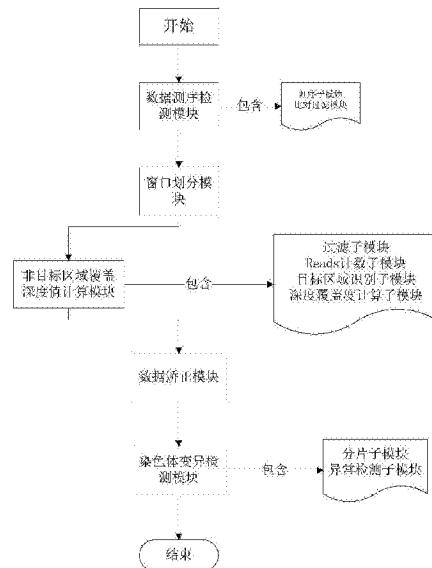
权利要求书2页 说明书7页 附图3页

(54)发明名称

一种染色体变异检测装置

(57)摘要

本发明涉及一种染色体变异检测装置，该染色体变异检测装置通过将目标区域捕获测序数据转化为低深度全基因组测序数据，从而基于捕获测序数据实现低深度全基因组检测的功能。本发明的染色体变异检测装置包括数据获取模块、窗口划分模块、数据转换模块、数据矫正模块、以及染色体变异判定模块。



1. 一种染色体变异检测装置,其包含下述模块:

数据获取模块,用于获取对测试样本进行目标区域捕获测序而得到目标区域捕获测序数据;

窗口划分模块,其与所述数据获取模块相连接,用于针对不同的参考基因组,按照给定的窗口长度对参考基因组序列划分窗口,生成各个窗口对应的GC含量以及匹配率;

数据转换模块,其与所述窗口划分模块相连接,用于将目标区域捕获测序数据转换成低深度全基因组测序数据;

数据矫正模块,其与所述数据转换模块相连接,用于利用划分窗口后计算得到的GC值以及匹配率值,分别进行LOESS矫正;以及

染色体变异判定模块:其与所述数据矫正模块相连接,用于判定是否存在染色体变异。

2. 根据权利要求1所述的染色体变异检测装置,其中,所述数据转换模块包括下述子模块:

目标区域识别子模块,其与所述窗口划分模块相连接,用于对目标区域的测序片段进行标记,并去除测试样本和参考基因组中被标记的目标区域;以及

数据转换核心子模块,其与所述目标区域识别子模块相连接,用于计算非补偿深度覆盖度值及补偿DOC值,

所述非补偿DOC值是指窗口内去除目标区域部分的DOC值,

所述补偿DOC值根据下述公式计算,

$$\text{compensated DOC} = \text{uncompensated DOC} * \text{binsize} / (\text{binsize} - x),$$

该公式中,compensated DOC表示补偿DOC值,uncompensated DOC表示非补偿DOC值,binsize表示窗口的长度,x表示窗口内属于目标区域的长度。

3. 根据权利要求1或2所述的染色体变异检测装置,其中,该染色体变异检测装置还包括过滤模块,所述数据获取模块和所述窗口划分模块通过所述过滤模块相连接,所述过滤模块用于将不合格及低质量的测序数据滤除。

4. 根据权利要求1~3中任一项所述的染色体变异检测装置,其中,所述窗口划分模块首先检测设定的窗口划分长度是否满足1Kb的整数倍,如果是,则根据测试样本的参考基因组,将参考基因组按染色体、窗口大小进行窗口划分,并给出每个窗口对应的GC值、匹配率。

5. 根据权利要求1~4中任一项所述的染色体变异检测装置,其中,所述数据矫正模块利用所述窗口划分模块生成的窗口GC含量和匹配率对所述数据转换模块计算得到的补偿DOC值进行LOESS矫正,得到DOC_{loess}值。

6. 根据权利要求5中任一项所述的染色体变异检测装置,其中,所述数据矫正模块对所述DOC_{loess}值再按如下公式进行一次中值矫正:

$$\log_2 \text{DOC}_{\text{loess}} = \log_2 (\text{DOC}_{\text{loess}} / \text{median}(\text{DOC}_{\text{loess}})),$$

该公式中,log₂DOC_{loess}表示最终的DOC值,median(DOC_{loess})表示所有窗口的DOC_{loess}值的中值。

7. 根据权利要求5或6所述的染色体变异检测装置,其中,染色体变异判定模块包括:

分片子模块:其与所述数据矫正模块相连接,用于将所述log₂DOC_{loess}值按照一定的规则进行分片,从而确定CNV断点的位置并将其展现出来;以及

异常报告子模块:其与所述分片子模块相连接,用于利用分片后的数据进行差异性分

析,报告不同片段的异常状态信息。

8.根据权利要求1~7中任一项所述的染色体变异检测装置,其中,所述数据转换模块还包括下述子模块:

过滤子模块,其与所述窗口划分模块相连接,用于过滤除去不合格的测序片段;

测序片段计数子模块,其与所述过滤子模块相连接,用于统计过滤后剩下的测序片段,并将其存放到指定文件中;

所述目标区域识别子模块与所述测序片段计数子模块相连接。

一种染色体变异检测装置

技术领域

[0001] 本发明涉及全基因组检测领域,尤其涉及一种将目标区域捕获测序数据转换为低深度的全基因组测序数据,并能用于全基因组相关检测的数据转换处理装置及方法。

背景技术

[0002] 通过基于全基因组或者全外显子测序的染色体拷贝数变异(copy number variant,CNV)检测,能够获得染色体异常的信息。上述检测中,通过对整个基因组或者全部的外显子区域测序,利用callCNV的方法来获取整个染色体的CNV信息,以此推断是否存在染色体变异。因此,这种方法产出的测序数据量非常大(其中包含大量无法利用的数据),实验和分析过程的各项成本也比较高。

[0003] 另一方面,捕获测序是通过预先设计好的探针,对基因组上的目标区域DNA片段进行抓取,并对抓取到的DNA片段进行测序。捕获测序具有低成本、高效率的特点。但是,捕获测序中,只有40%-60%左右的数据(其反映了有外显子覆盖的目标区域的信息)被利用。

发明内容

[0004] 本发明人发现,捕获测序中未被利用的数据并非毫无意义,其反映了染色体上非目标区域的信息,如在测序数据处理过程中将其转化为低深度全基因组测序可利用的形式,则可以基于捕获测序数据实现低深度全基因组检测的功能。

[0005] 因此,本发明的目的在于提供一种能够基于捕获测序数据实现低深度全基因组检测的功能的染色体变异检测装置。

[0006] 为了实现上述目的,本发明提供:

[0007] 1. 一种染色体变异检测装置,其包含下述模块:

[0008] 数据获取模块,用于获取目标区域捕获测序数据;

[0009] 窗口划分模块:其与所述数据获取模块相连接,用于针对不同的参考基因组,按照用户给定的窗口长度对参考基因组序列划分窗口,生成各个窗口对应的GC含量以及匹配率(mappability);

[0010] 数据转换模块:其与所述窗口划分模块相连接,用于将目标区域捕获测序数据转换成近似于低深度全基因组测序数据;

[0011] 数据矫正模块:其与所述非目标区域覆盖深度值计算模块相连接,用于利用划分窗口后计算得到的GC值以及匹配率值,分别进行LOESS矫正;该矫正当能够减少因芯片不同区域捕获情况不同造成的差异,从而降低数据的波动性;

[0012] 染色体变异判定模块:其与所述数据矫正模块相连接,用于判定是否存在染色体变异。染色体变异判定模块例如利用现有的拷贝数变异检测R包(DNAcopy)将矫正后的数据按染色体以及DOC值进行划分,同一个染色体上位置相邻且DOC值相近的区域划分到一起,这个划分过程具体实现算法是循环二元分割算法(Circular binary segmentation,CBS),从而实现将平均DOC值不同的区域区分开,再利用现有的隐马尔科夫算法(HMM)给这些划分

后得到的DOC值进行定性判断,从而确定哪些染色体片段是变异的,哪些是正常的。

[0013] 2.根据项1所述的染色体变异检测装置,其中,所述数据转换模块包括下述子模块:

[0014] 目标区域识别子模块,其与所述窗口划分模块相连接,用于对目标区域的reads进行标记(例如,可以利用现有的工具MACS 1.4),并去除测试样本和参考基因组中被标记的目标区域(peaks)(可利用例如现有的工具bedtools);

[0015] 数据转换核心子模块,其与所述目标区域识别子模块相连接,用于计算非补偿深度覆盖度值(uncompensated DOC)及补偿DOC值(compensated DOC),

[0016] 所述非补偿DOC值是指窗口内去除peaks区域部分的DOC值(例如,可利用Rsamtools工具进行计算),

[0017] 所述补偿DOC值根据下述公式计算

[0018] $\text{compensated DOC} = \text{uncompensated DOC} * \text{binsize} / (\text{binsize} - x)$,

[0019] 该公式中,compensated DOC表示补偿DOC值,uncompensated DOC表示非补偿DOC值,binsize表示窗口的长度,x表示窗口内属于peaks区域的长度。

[0020] 该数据转换核心模块在去除了peaks区域的reads后,计算非目标区域的深度覆盖值(Depth of Coverage,DOC),首先利用例如Rsamtools计算窗口内去除peaks区域部分的DOC值,记为非补偿深度覆盖度值(uncompensated DOC),再计算窗口内属于peaks区域的长度x,所以整个窗口的DOC值经过公式转换记为“ ~~læ ø 1 1 ° - # 2 &~~(补偿DOC值)”,具体转换公式如下:

[0021] $\text{compensated DOC} = \text{uncompensated DOC} * \text{binsize} / (\text{binsize} - x)$ 。

[0022] 所述数据转换模块实现了下述功能:将窗口内属于peaks区域的reads去除后计算得到的非peaks区域的DOC值等比例扩展到整个窗口,即实现了用非目标区域的测序数据表示整个基因组,且转换后得到的数据能够排除基因突变的影响,真实的反映出整个染色体的情况。

[0023] 3.根据项1或2所述的染色体变异检测装置,其中,该染色体变异检测装置还包括过滤模块,所述数据获取模块和所述窗口划分模块通过所述过滤模块相连接,所述过滤模块用于将不合格及低质量的测序数据滤除。

[0024] 4.根据项1~3中任一项所述的染色体变异检测装置,其中,所述窗口划分模块首先检测用户设定的窗口划分长度是否满足1Kb的整数倍,如果是,则根据测试样本要用到的参考基因组比如人群样本用到的hg19参考基因组,将参考基因组按染色体、窗口大小(例如20Kb)进行窗口划分,并给出每个窗口对应的GC值、匹配率。

[0025] 5.根据项1~4中任一项所述的染色体变异检测装置,其中,所述数据矫正模块利用所述窗口划分模块生成的窗口GC含量和匹配率对所述数据转换模块计算得到的补偿DOC值进行LOESS矫正,得到DOC_{loess}值。这样能够减小数据波动。

[0026] 6.根据项5中任一项所述的染色体变异检测装置,其中,所述数据矫正模块对所述DOC_{loess}值再按如下公式进行一次中值矫正:

[0027] $\log_2 \text{DOC}_{\text{loess}} = \log_2 (\text{DOC}_{\text{loess}} / \text{median}(\text{DOC}_{\text{loess}}))$,

[0028] 该公式中, $\log_2 \text{DOC}_{\text{loess}}$ 表示最终的DOC值,median($\text{DOC}_{\text{loess}}$)表示所有窗口的DOC_{loess}值的中值。

- [0029] 这样能够更明显地区分正常区域和非正常区域。
- [0030] 7. 根据项5或6所述的染色体变异检测装置,其中,染色体变异判定模块包括:
- [0031] 分片子模块:其与所述数据矫正模块相连接,用于将所述 $\log_2\text{DOC}_{\text{loess}}$ 值按照一定的规则进行分片,从而确定CNV断点的位置并将其展现出来;
- [0032] 异常报告子模块:其与所述分片子模块相连接,用于利用分片后的数据进行差异性分析,报告不同片段的异常状态信息。所述异常状态信息的内容包括但不限于加倍(duplication)和缺失(deletion)。
- [0033] 8. 根据项1~7中任一项所述的染色体变异检测装置,其中,所述数据转换模块还包括下述子模块:
- [0034] 过滤子模块,其与所述窗口划分模块相连接,用于过滤除去不合格的测序片段;例如可以利用samtools工具计算测试样本的bam文件中各个reads的phred值,并将phred值小于37的reads过滤掉;
- [0035] 测序片段计数子模块,其与所述过滤子模块相连接,用于统计过滤后剩下的测序片段,并将其存放到指定文件(例如新的bam文件)中;
- [0036] 所述目标区域识别子模块与所述测序片段计数子模块相连接。
- [0037] 9. 一种染色体变异检测方法,其包括:
- [0038] 数据获取步骤,获取目标区域捕获测序数据;
- [0039] 窗口划分步骤,针对不同的参考基因组,按照用户给定的窗口长度对参考基因组序列划分窗口,生成各个窗口对应的GC含量以及匹配率;
- [0040] 数据转换步骤,将目标区域捕获测序数据转换成近似于低深度全基因组测序数据;
- [0041] 数据矫正步骤,利用划分窗口后计算得到的GC值以及匹配率值,分别进行LOESS矫正;该矫正能够减少因芯片不同区域捕获情况不同造成的差异,从而降低数据的波动性;
- [0042] 染色体变异判定步骤,判定是否存在染色体变异。例如,可以利用现有的拷贝数变异检测R包(DNAcopy)将矫正后的数据按染色体以及DOC值进行划分,同一个染色体上位置相邻且DOC值相近的区域划分到一起,这个划分过程具体实现算法是循环二元分割算法(Circular binary segmentation,CBS),从而实现将平均DOC值不同的区域区分开,再利用现有的隐马尔科夫算法(HMM)给这些划分后得到的DOC值进行定性判断,从而确定哪些染色体片段是变异的,哪些是正常的。
- [0043] 10. 根据项9所述的染色体变异检测方法,其中,所述数据转换步骤包括:
- [0044] 过滤子步骤,过滤除去不合格的reads;例如可以利用samtools工具计算测试样本的bam文件中各个reads的phred值,并将phred值小于37的reads过滤掉;
- [0045] reads计数子步骤,统计过滤后剩下的reads,并将其存放到指定文件(例如新的bam文件)中;
- [0046] 目标区域识别子步骤,对目标区域的reads进行标记(例如,可以利用现有的工具MACS 1.4),并去除测试样本和参考基因组中被标记的目标区域peaks(可利用例如现有的工具bedtools);
- [0047] 数据转换核心子步骤,计算非补偿深度覆盖度值(uncompensated DOC)及补偿DOC值(compensated DOC),

[0048] 所述非补偿DOC值是指窗口内去除peaks区域部分的DOC值(例如,可利用Rsamtools工具进行计算),

[0049] 所述补偿DOC值根据下述公式计算,

[0050] $\text{compensated DOC} = \text{uncompensated DOC} * \text{binsize} / (\text{binsize} - x)$,

[0051] 该公式中,compensated DOC表示补偿DOC值,uncompensated DOC表示非补偿DOC值,binsize表示窗口的长度,x表示窗口内属于peaks区域的长度。

[0052] 该子步骤在去除了peaks区域的reads后,计算非目标区域的深度覆盖值(Depth of Coverage,DOC),首先利用例如Rsamtools计算窗口内去除peaks区域部分的DOC值,记为非补偿深度覆盖度值(uncompensated DOC),再计算窗口内属于peaks区域的长度x,所以整个窗口的DOC值经过公式转换记为 $\text{compensated DOC} = \text{uncompensated DOC} * \text{binsize} / (\text{binsize} - x)$,具体转换公式如下:

[0053] $\text{compensated DOC} = \text{uncompensated DOC} * \text{binsize} / (\text{binsize} - x)$ 。

[0054] 数据转换步骤实现了下述功能:将窗口内属于peaks区域的reads去除后计算得到的非peaks区域的DOC值等比例扩展到整个窗口,即实现了用非目标区域的测序数据表示整个基因组,且转换后得到的数据能够排除基因突变的影响,真实的反映出整个染色体的情况。

[0055] 11.根据项9或10所述的染色体变异检测方法,其中,在所述窗口划分步骤之前还包括过滤步骤,将不合格及低质量的测序数据滤除。

[0056] 12.根据项9~11中任一项所述的染色体变异检测方法,其中,所述窗口划分步骤中,首先检测用户设定的窗口划分长度是否满足1Kb的整数倍,如果是,则根据测试样本要用到的参考基因组比如人群样本用到的hg19参考基因组,将参考基因组按染色体、窗口大小(例如20Kb)进行窗口划分,并给出每个窗口对应的GC值、匹配率。

[0057] 13.根据项9~12中任一项所述的染色体变异检测方法,其中,所述数据矫正步骤中利用所述窗口划分步骤生成的窗口GC含量和匹配率对所述数据转换步骤中计算得到的补偿DOC值进行LOESS矫正,得到 $\text{DOC}_{\text{loess}}$ 值。这样能够减小数据波动。

[0058] 14.根据项13所述的染色体变异检测方法,其中,所述数据矫正步骤中对所述 $\text{DOC}_{\text{loess}}$ 值再按如下公式进行一次中值矫正:

[0059] $\log_2 \text{DOC}_{\text{loess}} = \log_2 (\text{DOC}_{\text{loess}} / \text{median}(\text{DOC}_{\text{loess}}))$,

[0060] 该公式中, $\log_2 \text{DOC}_{\text{loess}}$ 表示最终的DOC值, $\text{median}(\text{DOC}_{\text{loess}})$ 表示所有窗口的 $\text{DOC}_{\text{loess}}$ 值的中值。

[0061] 这样能够更明显地区分正常区域和非正常区域。

[0062] 15.根据项14所述的染色体变异检测方法,其中,染色体变异判定步骤包括:

[0063] 分片子步骤:将所述 $\log_2 \text{DOC}_{\text{loess}}$ 值按照一定的规则进行分片,从而确定CNV断点的位置并将其展现出来;

[0064] 异常报告子步骤:利用分片后的数据进行差异性分析,报告不同片段的异常状态信息。所述异常状态信息的内容包括但不限于加倍和缺失。

[0065] 根据项9~15中任一项所述的染色体变异检测方法,其中,所述数据转换步骤在进行所述目标区域识别子步骤前还包括:

[0066] 过滤子步骤,过滤除去不合格的reads;例如可以利用samtools工具计算测试样本的bam文件中各个reads的phred值,并将phred值小于37的reads过滤掉;和

[0067] reads计数子步骤,统计过滤后剩下的reads,并将其存放到指定文件(例如新的bam文件)中。

附图说明

[0068] 图1是本发明的染色体变异检测装置的优选实施方式的一例的示意图。

[0069] 图2是确定目标捕获区域后去除该区域reads后的示意图。

[0070] 图3是对样本做GC矫正后的示意图。

[0071] 图4是对样本做匹配率矫正后的示意图。

[0072] 图5是得到最终染色体变异检测结果的示意图。

发明的具体实施方式

[0074] 本说明书中提及的科技术语具有与本领域技术人员通常理解的含义相同的含义,如有冲突以本说明书中的定义为准。

[0075] 本发明的目的在于提供一种能够基于捕获测序数据实现低深度全基因组检测的功能的染色体变异检测装置。通过本发明的染色体变异检测装置,能够将目标区域捕获测序的数据转换成等价于低深度全基因组测序的数据,并利用这些转换后的数据检测染色体变异信息。下面将参考附图结合实施例来详细说明本发明。

[0076] 图1是本发明的染色体变异检测装置的优选实施方式的一例的示意图。该优选实施方式的染色体变异检测装置包括:

[0077] 数据获取模块,用于获取目标区域捕获测序数据。

[0078] 过滤模块,所述数据获取模块和所述窗口划分模块通过所述过滤模块相连接,所述过滤模块用于将不合格及低质量的测序数据滤除。在一个具体实施方式中,利用比对软件BWA比对到人类参考基因组,并利用samtools工具将比对后的reads存放在bam格式的文件中。比对完成后还要对原始的比对结果进行一个筛选,去除低质量以及重复的reads,得到用于输入给数据转换模块的unique bam文件。

[0079] 窗口划分模块:其与所述数据获取模块相连接,用于针对不同的参考基因组,按照用户给定的窗口长度对参考基因组序列划分窗口,生成各个窗口对应的GC含量以及匹配率;更具体地,对hg19人类参考基因组按照窗口长度为20Kb的长度划分窗口,计算窗口内的GC含量以及匹配率;

[0080] 数据转换模块:其与所述窗口划分模块相连接,用于将目标区域捕获测序数据转换成近似于低深度全基因组测序数据。在该实施方式中,所述数据转换模块包括下述子模块:

[0081] 过滤子模块,其与所述窗口划分模块相连接,用于过滤除去不合格的reads。具体地,用samtools读取unique bam文件中的reads,并计算reads的phred值,保留phred>37的reads。

[0082] reads计数子模块,其与所述过滤子模块相连接,用于统计过滤后剩下的reads,并将其存放到指定文件中。具体地,将这些reads按划分好的窗口计数后存入新的bam文件中。

[0083] 目标区域识别子模块,其与reads计数子模块相连接,用于对目标区域的reads进行标记,并去除测试样本和参考基因组中前一步中标记的目标区域peaks;具体地,利用MACS软件标记目标区域(peaks)中的reads,然后,对于带有peaks标记的reads,用bedtools

工具去除这些reads。图2为MACS识别标记目标区域以及剔除这些区域reads后的示意图。

[0084] 数据转换核心子模块,其与所述目标区域识别子模块相连接,用于计算非补偿深度覆盖度值(uncompensated DOC)及补偿DOC值(compensated DOC),

[0085] 所述非补偿DOC值是指窗口内去除peaks区域部分的DOC值(例如,可利用Rsamtools工具进行计算),

[0086] 所述补偿DOC值根据下述公式计算

[0087] $\text{compensated DOC} = \text{uncompensated DOC} * \text{binsize} / (\text{binsize} - x)$,

[0088] 该公式中,compensated DOC表示补偿DOC值,uncompensated DOC表示非补偿DOC值,binsize表示窗口的长度,这里为20Kb,x表示窗口内属于peaks区域的长度。

[0089] 根据之前划分好的窗口,通过Rsamtools工具计算各个窗口非peaks区域的DOC值,用窗口内非peaks区域计算得到的DOC值只能称为未补偿DOC值,不能完全表现整个窗口的DOC信息,所以利用数据转换装置对这一数据进行转换,得到新的称为补偿DOC值的数据。

[0090] 数据矫正模块:其与所述非目标区域覆盖深度值计算模块相连接,用于利用划分窗口后计算得到的GC值以及匹配率值,分别进行LOESS矫正。所述数据矫正模块利用所述窗口划分模块生成的窗口GC含量和匹配率对所述数据转换模块计算得到的补偿DOC值进行LOESS矫正,得到 $\text{DOC}_{\text{loess}}$ 值。然后,所述数据矫正模块对所述 $\text{DOC}_{\text{loess}}$ 值再按如下公式进行一次中值矫正:其中,LOESS表示局部(加权)线性回归;

[0091] $\log_2 \text{DOC}_{\text{loess}} = \log_2 (\text{DOC}_{\text{loess}} / \text{median}(\text{DOC}_{\text{loess}}))$,

[0092] 该公式中, $\log_2 \text{DOC}_{\text{loess}}$ 表示最终的DOC值,median($\text{DOC}_{\text{loess}}$)表示所有窗口的 $\text{DOC}_{\text{loess}}$ 值的中值。

[0093] 具体地,通过上述模块得到了样本按窗口划分计算的原始的DOC值,但是这些数据还存在着一些因染色体捕获差异以及特殊片段的结构差异造成的数据波动,为了消除这些数据波动,用窗口划分模块中计算得到的窗口GC含量和匹配率值分别进行LOESS校正,两次LOESS校正后得到图3,图4所示效果。做完LOESS校正后,再做一次中值校正,这样所有的矫正DOC值都更加接近0附近,便于后续的定性分析。

[0094] 染色体变异判定模块:其与所述数据矫正模块相连接,用于判定是否存在染色体变异。

[0095] 所述染色体变异判定模块包括:

[0096] 分片子模块:其与所述数据矫正模块相连接,用于将所述 $\log_2 \text{DOC}_{\text{loess}}$ 值按照一定的规则进行分片,从而确定CNV断点的位置并将其展现出来。具体而言,对于矫正后的DOC值做定性分析,即分成正常,缺失,加倍等类型,是一个明显的分类问题,普通的方法比如t检验,在处理这种位点数较大的数据上准确性不足,因此,先采用现有的R程序DNAcopy对矫正后的DOC数据进行分片,将位置相近且使得这些区域内的方差最小的所有位点归为一类,这样就相当于减少了要计算的位点数量,只需对分成一类的点做整体的差异性比较,就能得到该区域的定性结果,如图5所示。

[0097] 异常报告子模块:其与所述分片子模块相连接,用于利用分片后的数据进行差异性分析,报告不同片段的异常状态信息。

[0098] 图5中的结果显示1号染色体和8号染色体存在加倍的异常情况出现,这个结果和该样本通过fisher实验得到的结果(表1)是一致的。

[0099] 表1

[0100]

样本编号	染色体异常情况	染色体编号	染色体编号
D136877	异常	+1	+8

[0101] 注:+表示加倍,-表示缺失

[0102] 工业实用性

[0103] 根据本发明，提供了一种能够基于捕获测序数据实现低深度全基因组检测的功能的染色体变异检测装置。

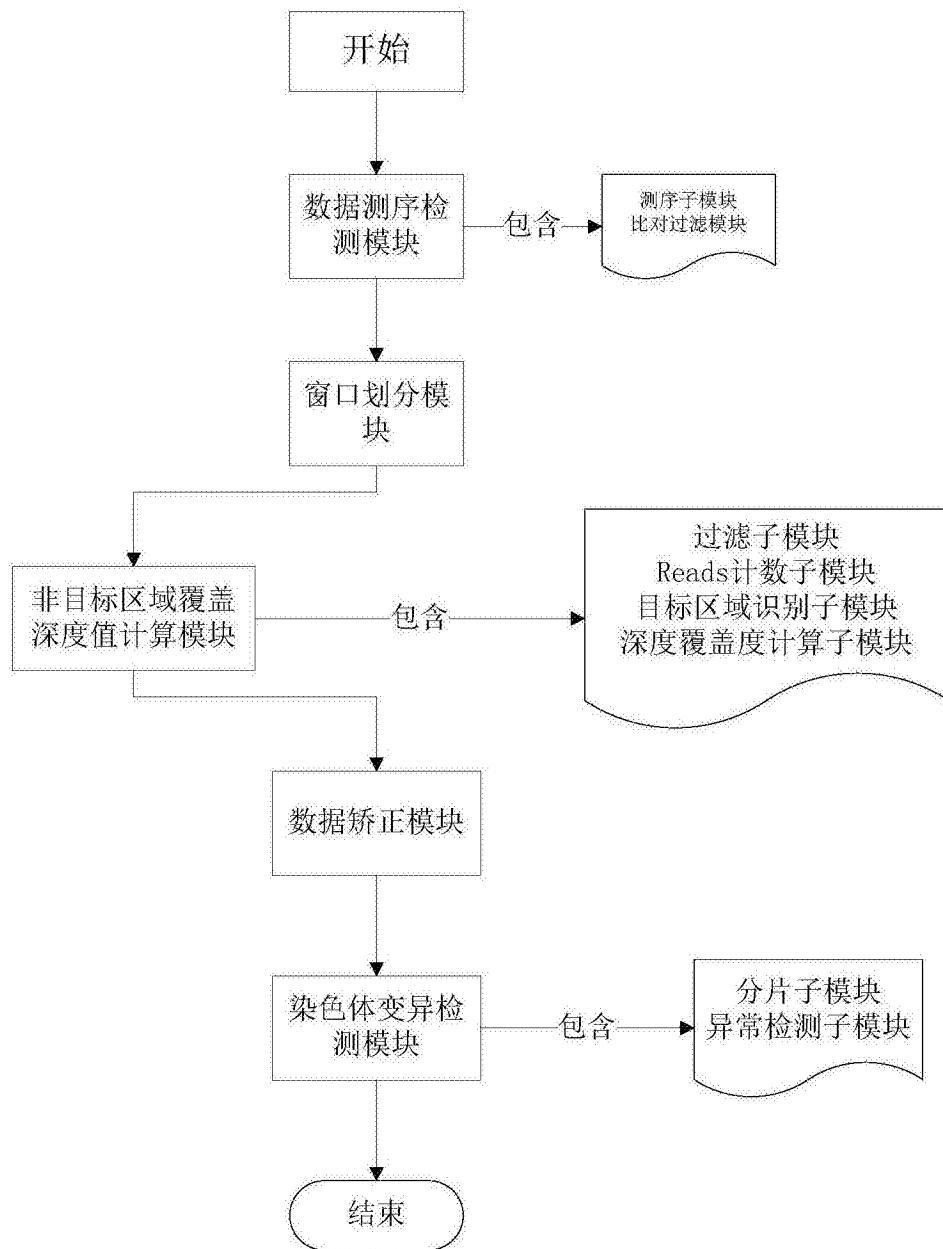


图1

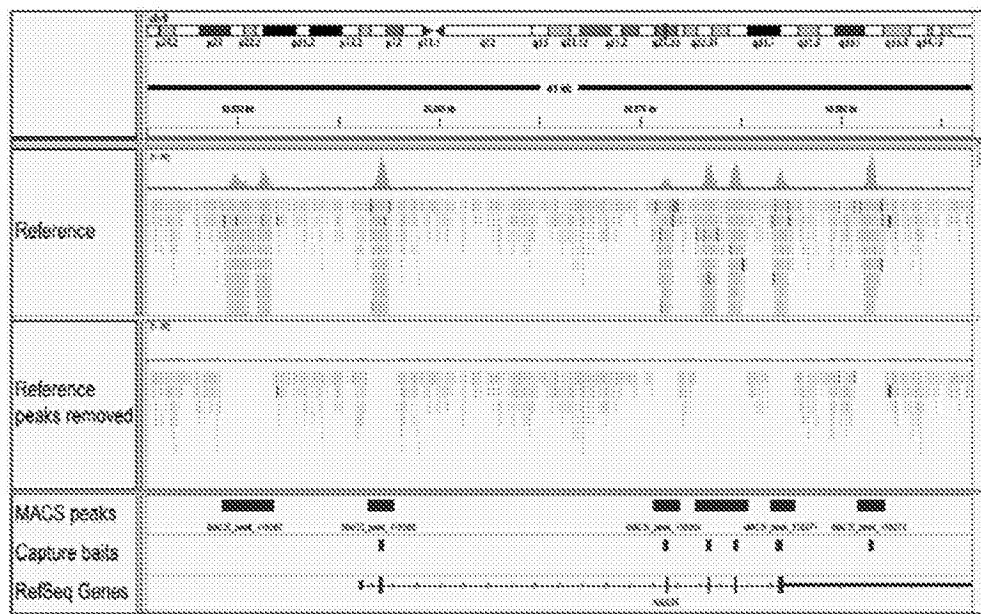


图2

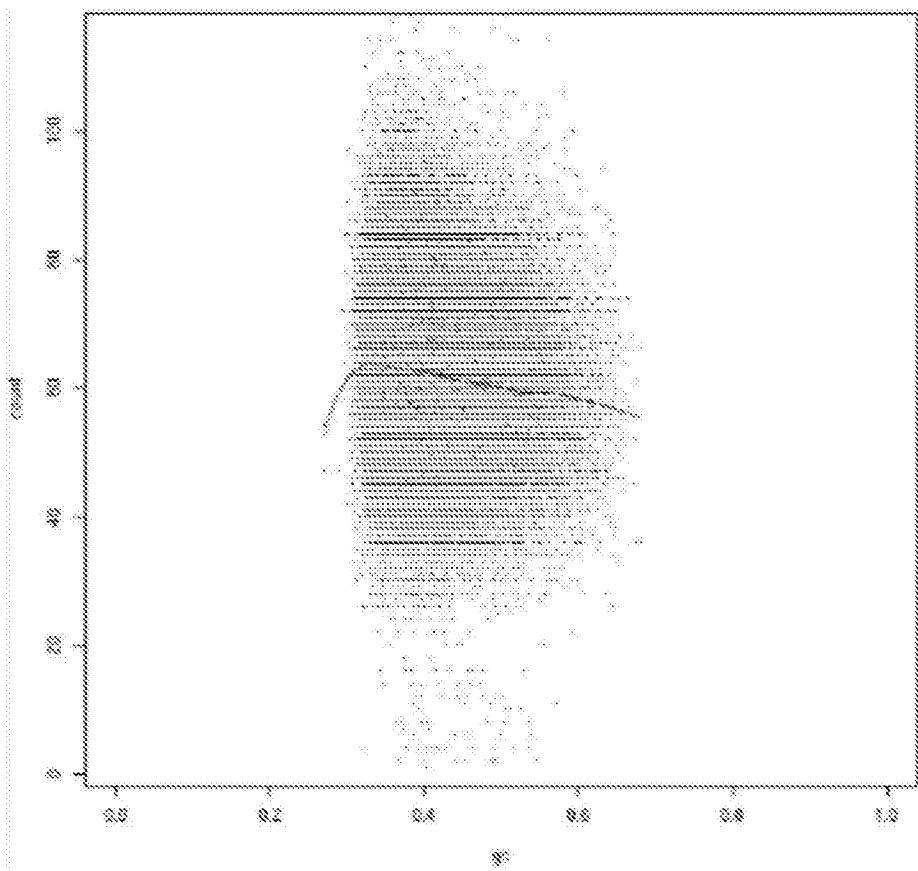


图3

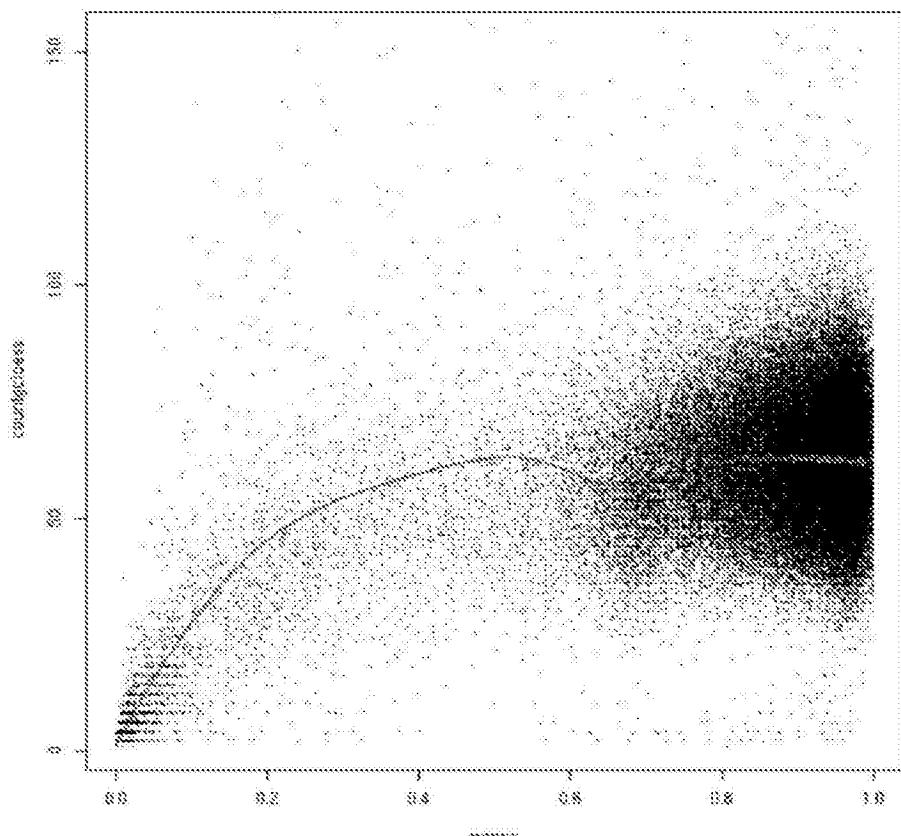


图4

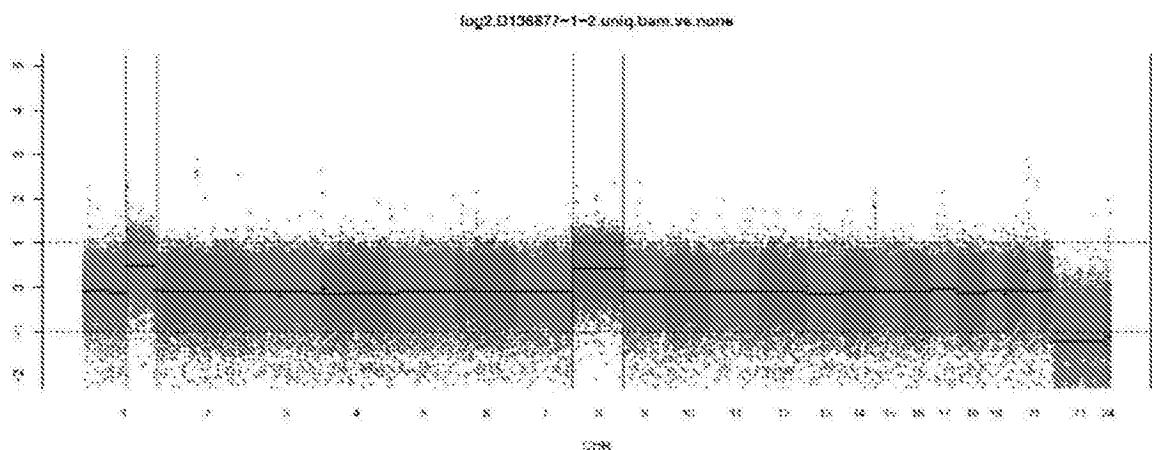


图5