

人类HLA染色体区域杂合性缺失的分析方法和分析处理 装置

[0001] 本申请要求于2020年12月4日申请的、申请号为202011404354.4的中国专利申请的优先权，其全部内容通过引用结合在本申请中。

技术领域

[0002] 本申请涉及生物信息分析领域，具体涉及一种人类HLA染色体区域杂合性缺失的分析方法和分析处理装置。

背景技术

[0003] 染色体杂合性缺失(Loss Of Heterozygosity, LOH)在恶性肿瘤中是一种常见的染色体畸变。发生于HLA(人类白细胞抗原)染色体区域的LOH称为HLA Loss，发生HLA Loss后会影响到HLA的正常表达，使肿瘤细胞逃避机体免疫细胞的杀伤，导致肿瘤细胞的免疫逃逸。HLA Loss在实体瘤如非小细胞肺癌中常见，也可发生在血液肿瘤中，在造血干细胞移植手术后若肿瘤细胞发生HLA Loss则会导致该肿瘤细胞不被免疫细胞杀伤，导致该种血液病复发。至2017年为止，在半相合造血干细胞移植案例中，由于HLA Loss导致的移植后复发占有复发病例的33%。HLA Loss国际多中心协作组的结果显示，在396例有效入组病例中，HLA Loss导致的复发共51例，其中35例来自半相合移植，12例来自HLA错配的无关供者移植，4例来自全相合无关供者移植，脐血移植HLA Loss为0。多中心研究的结果显示，HLA

Loss是造血干细胞移植后免疫逃逸和复发的主要机制之一，且HLA Loss的比例与供患之间HLA的错配位点数量有关。

[0004] 移植后HLA Loss的检测对于明确复发原因，指导复发治疗有非常重要的意义。由于错配位点的丢失，常规的复发治疗手段，如减少免疫抑制剂使用量及供者淋巴细胞输注(DLI)，在HLA Loss复发中并不适用并需严格禁止，因为HLA Loss的复发中，DLI输注不会产生移植物抗白血病效应，但会产生严重的移植物抗宿主病反应。基于此，移植复发后HLA Loss是否存在，应该成为常规临床开

展项目，甚至，在复发前应该进行HLA Loss动态监测，以尽早的预测复发，对复发的治疗方案的制定也至关重要。

[0005] HLA Loss导致的移植后复发是近年新发现的现象，目前国际范围内可对其进行检测的移植中心屈指可数，有部分机构尝试采用荧光定量PCR方法对其检测，然而该方法仅能覆盖约70%人群的HLA型别。国内更是一片空白，尚无一种高效、精准、成熟可靠的检测方法和系统可对其进行检测和分析。基于以上原因，目前亟需一种高效、精准、成熟可靠且适合所有人群的HLA Loss检测方法和系统。

发明概述

技术问题

[0006] 本申请针对检测HLA Loss导致的移植后复发的需求，提供了一种高效、精准、成熟可靠且适合所有人群的HLA Loss分析方法和分析处理装置。

问题的解决方案

技术解决方案

[0007] 为此，本申请一方面提供了一种HLA染色体区域的杂合性缺失的分析方法，包括以下步骤：

[0008] 对样本数据进行拆分和过滤，得到过滤后的样本序列；

[0009] 获得各HLA基因测序数据文件，所述测序数据文件包含各HLA基因的样本数据；

[0010] 生成受者移植术前HLA基因型的第一个等位基因序列作为参考基因；

[0011] 获得各HLA基因中受者移植术后所占百分比；

[0012] 获得受者移植术后HLA基因染色体区域的百分比，以HLA%表示；

[0013] 获得受者移植术后细胞中各染色体总体百分比，以STR%表示；

[0014] 受者移植术后HLA Loss阴阳性判断。

[0015] 在本申请的实施方案中，各步骤具体为：

[0016] 对测序完成后的样本下机数据进行拆分和过滤，得到过滤后的样本序列；

[0017] 根据设定的比对参数，将过滤后的样本序列比对至各HLA基因参考基因序列，将比对上的测序序列拆分至各HLA测序数据文件，所述测序数据文件包含测序完成后各HLA基因的样本下机数据；

- [0018] 获取移植术前受者和供者各HLA等位基因型别，比对后，获得受者与供者之间各HLA基因的SNP差异，生成受者移植术前HLA基因型的第一个等位基因序列作为参考基因；
- [0019] 将所述测序数据文件比对至所述参考基因，统计各HLA基因SNP位置的受者及供者分别的测序深度，将各HLA基因全部SNP位置受者的深度频率取平均值，得到各HLA基因中受者移植术后所占百分比；
- [0020] 取各HLA基因受者百分比的平均值，即为受者移植术后HLA染色体区域的百分比，以HLA%表示；
- [0021] 获得受者移植术后细胞中各染色体总体百分比，以STR%表示；
- [0022] 阴阳性判断：当 $HLA\% \leq 0.5\%$ ，且 $STR\% \geq 3\%$ 时，判断为HLA Loss阳性；当 $HLA\% \geq 3\%$ 时，判断为HLA Loss阴性；当 $0.5\% \leq HLA\% < 3\%$ 时提示无法判断，受者移植术后细胞HLA染色体可能处于正在缺失时期。
- [0023] 在本申请的一实施方案中，所述分析方法还包括以下步骤：
- [0024] 将HLA各基因百分比、平均百分比、阴阳性判断结果输出至报告文件。
- [0025] 在本申请的一实施方案中，各HLA基因分别为HLA-A、HLA-B、HLA-C、HLA-DRB1、HLA-DQB1和HLA-DPB1基因。
- [0026] 在本申请的一实施方案中，HLA-A基因比对至A*01:01:01:01、HLA-B基因比对至B*07:02:01:01、HLA-C基因比对至C*01:02:01:01、HLA-DRB1基因比对至DRB1*01:02:01:01、HLA-DQB1基因比对至DQB1*05:01:01:01、HLA-DPB1基因比对至DPB1*01:01:01:01。
- [0027] 本申请另一方面提供了一种HLA染色体区域的杂合性缺失的分析处理装置，其包括以下模块：
- [0028] 拆分和过滤模块，其对样本数据进行拆分和过滤，得到过滤后的样本序列；
- [0029] 基因测序数据文件获取模块，用于获得各HLA基因测序数据文件，所述测序数据文件包含各HLA基因的样本数据；
- [0030] 参考基因生成模块，用于生成受者移植术前HLA基因型的第一个等位基因序列作为参考基因；
- [0031] 各HLA基因所占百分比计算模块，用于获得各HLA基因中受者移植术后所占百分

比；

- [0032] HLA基因染色体区域百分比计算模块，用于获得受者移植术后HLA基因染色体区域的百分比，以HLA%表示；
- [0033] HLA基因各染色体总体百分比计算模块，用于获得受者移植术后细胞中各染色体总体百分比，以STR%表示；
- [0034] 阴阳性判断模块，用于进行受者移植术后HLA Loss阴阳性判断。
- [0035] 在本申请的一实施方案中，各模块具体为：
- [0036] 拆分和过滤模块，其对测序完成后的样本下机数据进行拆分和过滤，得到过滤后的样本序列；
- [0037] 基因测序数据文件获取模块，其根据设定的比对参数，将过滤后的样本序列比对至各HLA基因参考基因序列，将比对上的测序序列拆分至各HLA测序数据文件，所述测序数据文件包含测序完成后各HLA基因的样本下机数据；
- [0038] 参考基因生成模块，用于获取移植术前受者和供者各HLA等位基因型别，比对后，获得受者与供者之间各HLA基因的SNP差异，生成受者移植术前HLA基因型的第一个等位基因序列作为参考基因；
- [0039] 各HLA基因所占百分比计算模块，用于将所述测序数据文件比对至所述参考基因，统计各HLA基因SNP位置的受者及供者分别的测序深度，将各HLA基因全部SNP位置受者的深度频率取平均值，得到各HLA基因中受者移植术后所占百分比；
- [0040] HLA基因染色体区域百分比计算模块，用于取各HLA基因受者百分比的平均值，即为受者移植术后HLA染色体区域的百分比，以HLA%表示；
- [0041] HLA基因各染色体总体百分比计算模块，用于获得受者移植术后细胞中各染色体总体百分比，以STR%表示；
- [0042] 阴阳性判断模块，用于进行受者移植术后HLA Loss阴阳性判断，当 $HLA\% \leq 0.5\%$ ，且 $STR\% \geq 3\%$ 时，判断为HLA Loss阳性；当 $HLA\% \geq 3\%$ 时，判断为HLA Loss阴性；当 $0.5\% \leq HLA\% < 3\%$ 时提示无法判断，受者移植术后细胞HLA染色体可能处于正在缺失时期。
- [0043] 在本申请的一实施方案中，所述分析处理装置还包括以下模块：
- [0044] 报告文件生成模块，用于将HLA各基因百分比、平均百分比、阴阳性判断结果

输出至报告文件。

[0045] 在本申请的一实施方案中，各HLA基因分别为HLA-A、HLA-B、HLA-C、HLA-DRB1、HLA-DQB1和HLA-DPB1基因。

[0046] 在本申请的一实施方案中，HLA-A基因比对至A*01:01:01:01、HLA-B基因比对至B*07:02:01:01、HLA-C基因比对至C*01:02:01:01、HLA-DRB1基因比对至DRB1*01:02:01:01、HLA-DQB1基因比对至DQB1*05:01:01:01、HLA-DPB1基因比对至DPB1*01:01:01:01。

发明的有益效果

有益效果

[0047] 由上述描述可知，本申请基于二代测序数据，提供的分析方法和分析处理装置能够方便地进行流程化和批量化操作，人工解读工作量小，能够准确地检测样本中是否发生HLA染色体区域杂合性缺失，对检测HLA Loss导致的移植后复发具有重要的意义。

对附图的简要说明

附图说明

[0048] 图1为本申请分析方法流程图。

[0049] 图2为本申请分析处理装置中各模块相互关系显示图。

[0050] 图3为实施例1中原始数据整体数据质量显示图。

[0051] 图4为实施例1中过滤后数据整体数据质量显示图。

[0052] 图5为实施例2中原始数据整体数据质量显示图。

[0053] 图6为实施例2中过滤后数据整体数据质量显示图。

[0054] 图7为实施例3的线性拟合曲线图。

发明实施例

本发明的实施方式

[0055] 下面通过实施例对本申请作进一步的详细说明，旨在用于说明本申请而非限定本申请。应当指出，对于本领域技术人员而言，在不脱离本申请原理的前提下，还可以对本申请进行若干改进和修饰，这些改进和修饰也同样落入本申请的

保护范围之内。

- [0056] 如图1所示，图1为本申请提供的一种HLA染色体区域的杂合性缺失的分析方法，包括以下步骤：
- [0057] 对样本数据进行拆分和过滤，得到过滤后的样本序列；
- [0058] 获得各HLA基因测序数据文件，所述测序数据文件包含各HLA基因的样本数据；
- [0059] 生成受者移植术前第一个等位基因序列作为参考基因；
- [0060] 获得各HLA基因中受者移植术后所占百分比；
- [0061] 获得受者移植术后HLA基因染色体区域的百分比，以HLA%表示；
- [0062] 获得受者移植术后细胞中各染色体总体百分比，以STR%表示；
- [0063] 受者移植术后HLA Loss阴阳性判断。
- [0064] 如图2所示，图2为本申请提供的一种HLA染色体区域的杂合性缺失的分析处理装置中各模块相互关系显示图。该分析处理装置包括以下模块：
- [0065] 拆分和过滤模块，其对样本数据进行拆分和过滤，得到过滤后的样本序列；
- [0066] 基因测序数据文件获取模块，用于获得各HLA基因测序数据文件，所述测序数据文件包含各HLA基因的样本数据；
- [0067] 参考基因生成模块，用于生成受者移植术前HLA基因型的第一个等位基因序列作为参考基因；
- [0068] 各HLA基因所占百分比计算模块，用于获得各HLA基因中受者移植术后所占百分比；
- [0069] HLA基因染色体区域百分比计算模块，用于获得受者移植术后HLA基因染色体区域的百分比，以HLA%表示；
- [0070] HLA基因各染色体总体百分比计算模块，用于获得受者移植术后细胞中各染色体总体百分比，以STR%表示；
- [0071] 阴阳性判断模块，用于进行受者移植术后HLA Loss阴阳性判断。
- [0072] 通过以下实施例对上述HLA染色体区域的杂合性缺失的分析方法以及分析处理装置进行具体说明。
- [0073] 实施例1：造血干细胞移植术后HLA染色体区域的杂合性缺失阴性受者的检测
- [0074] 1、对造血干细胞移植受者术后进行靶向测序检测，获取配型和嵌合结果

- [0075] 本实施例采用造血干细胞移植受者术后血液标本，提取其基因组DNA为模板，进行二代靶向测序检测，检测结果采用本申请分析方法进行分析。
- [0076] 由移植前配型报告获取移植前受者HLA配型结果为：A*02:10, A*24:02, B*15:01, B*40:01, C*03:03, C*07:02, DRB1*09:01, DRB1*14:05, DQB1*05:03, DQB1*03:03；供者HLA配型结果为：A*24:02, A*26:01, B*15:01, B*40:01, C*03:03, C*07:02, DRB1*14:05, DRB1*15:02, DQB1*05:03, DQB1*06:01。供者和受者的B基因和C基因相合，本申请方法只分析计算A基因、DRB1基因和DQB1基因三个基因。由移植后嵌合率报告获取血液中受者细胞占比STR%为41.27%。
- [0077] 2、对测序获得的序列进行拆分和过滤，获得高质量的样本序列
- [0078] 参照附图1所示的流程图，使用bcl2fastq软件（网址为https://support.illumina.com/sequencing/sequencing_software/bcl2fastq-conversion-software.html）对二代测序下机数据进行拆分，并用Trimmomatic软件（网址为<http://www.usadellab.org/cms/?page=trimmomatic>）对数据进行过滤。过滤参数如下：过滤平均质量阈值：30；滑窗剪切（Slidingwindow）：5；得到过滤后的样本序列；开头及结尾切除质量值低于阈值的碱基数（Leading, Trailing）：5；长度阈值（Minlen）：50，过滤后得到高质量的样本序列。数据过滤前后对比情况参见附图3和附图4，可以明显看出原始数据在参考基因第1位以及第51至第75位碱基有部分reads，数据质量下降，而经过数据过滤之后这些位置质量差的部分均被删去。
- [0079] 3、将过滤后的样本序列比对至HLA基因参考基因，获得测序数据文件
- [0080] 由IMGT-HLA公共数据库（网址为<https://www.ebi.ac.uk/ipd/imgt/hla/>）查询获得A*01:01:01:01, DRB1*01:02:01:01, DQB1*05:01:01:01序列，使用bwa软件（网址为<https://sourceforge.net/projects/bio-bwa/>）将上一步过滤后的样本序列比对至各HLA基因参考基因（使用默认参数进行比对），将比对上的测序序列拆分至HLA-A、HLA-DRB1和HLA-DQB1共3个测序数据文件。
- [0081] 4、对比受者与供者各基因中不同型别之间的SNP差异，获得snp文件
- [0082] 使用MUSCLE软件（网址为<http://www.drive5.com/muscle/>）对比供者与受者A

基因中3个型别A*02:10, A*24:02, A*26:01的SNP差异(使用默认参数进行比对), 共有4个SNP, 保存为A.snp文件; 对比供者与受者DRB1基因中3个型别DRB1*09:01, DRB1*14:05, DRB1*15:02的SNP差异, 共有37个SNP, 保存为DRB1.snp文件; 对比供者与受者DQB1基因中3个型别DQB1*05:03, DQB1*03:03, DQB1*06:01的SNP差异, 共有19个SNP, 保存为DQB1.snp文件。

[0083] 5、将测序数据文件分别与对应基因的参考序列进行比对, 生成各基因的比对文件

[0084] 由IMGT-HLA公共数据库(网址为<https://www.ebi.ac.uk/ipd/imgt/hla/>)查询获得A*02:10, DRB1*09:01:02:01, DQB1*05:03:01:01参考序列, 使用bwa软件(网址为<https://sourceforge.net/projects/bio-bwa/>)的mem算法将3个测序数据文件中序列按基因分别比对至A*02:10, DRB1*09:01:02:01, DQB1*05:03:01:01参考序列, 生成各基因比对文件A.bam、DRB1.bam、DQB1.bam。

[0085] 6、获得基因比对文件中各SNP位置测序深度, 计算各基因深度占比的平均值

[0086] 进一步, 调取snp文件, 使用bamreadcount软件(网址为<https://github.com/genome/bam-readcount>)调取bam文件中各SNP位置测序深度, 计算各位置属于受者的型别的深度占比。A基因4个SNP占比分别为47.88%, 38.11%, 51.75%, 46.32%, 平均数46.02%; DRB1基因37个SNP中共有18个SNP被检测, 占比分别为23.64%, 31.60%, 52.81%, 41.32%, 48.04%, 64.69%, 46.54%, 43.59%, 58.48%, 36.23%, 70.02%, 75.06%, 30.40%, 48.29%, 46.99%, 25.18%, 56.42%, 40.23%, 平均数46.64%; DQB1基因19个SNP中共有7个SNP被检测, 占比分别为46.23%, 44.99%, 38.79%, 36.90%, 37.59%, 32.15%, 36.68%, 平均数39.05%。

[0087] 7、获得全部基因的平均值, 判断移植受者术后HLA染色体区域是否发生杂合性缺失

[0088] 进一步, 取3个基因平均值43.90%。根据本申请阴阳性判断方法移植受者术后HLA染色体区域占比43.90%大于3%, 判断该检测者为HLA Loss阴性, 结果表明该检测者术后HLA染色体区域未发生杂合性缺失。

- [0089] 实施例2：造血干细胞移植术后HLA染色体区域的杂合性缺失阳性受者的检测
- [0090] 1、对造血干细胞移植术后受者进行靶向测序检测、获取配型和嵌合结果
- [0091] 本实施例采用造血干细胞移植术后受者血液标本，提取其基因组DNA为模板，进行二代靶向测序检测，检测结果采用本申请分析方法进行分析。
- [0092] 由移植前配型报告获取移植前受者HLA配型结果为：A*02:03, A*30:01, B*13:02, B*18:01, C*06:02, C*07:04, DRB1*07:01, DRB1*14:04, DQB1*05:03, DQB1*02:02；供者HLA配型结果为：A*11:01, A*30:01, B*13:02, B*15:02, C*06:02, C*08:01, DRB1*07:01, DRB1*13:02, DQB1*06:04, DQB1*02:02。由移植后嵌合率报告获取血液中受者细胞占比STR%为6.85%。
- [0093] 2、对测序获得的序列进行拆分和过滤，获得高质量的样本序列
- [0094] 参照附图1流程图，使用bcl2fastq软件（网址为https://support.illumina.com/sequencing/sequencing_software/bcl2fastq-conversion-software.html）对二代测序下机数据进行拆分，并用Trimmomatic软件（网址为<http://www.usadellab.org/cms/?page=trimmomatic>）对数据进行过滤，参数如下：过滤平均质量阈值：30；滑窗剪切（Slidingwindow）：5；得到过滤后的样本序列；开头及结尾切除质量值低于阈值的碱基数（Leading, Trailing）：5；长度阈值（Minlen）：50，过滤后得到高质量的样本序列。数据过滤前后对比情况参见附图5和附图6，可以明显看出原始数据在参考基因第61至第75位碱基有部分reads，数据质量下降，而经过数据过滤之后这些位置质量差的部分均被删去。
- [0095] 3、将过滤后的样本序列比对至HLA基因参考基因，获得测序数据文件
- [0096] 由IMGT-HLA公共数据库（网址为<https://www.ebi.ac.uk/ipd/imgt/hla/>）查询获得A*01:01:01:01, B*07:02:01:01, C*01:02:01:01, DRB1*01:02:01:01, DQB1*05:01:01:01序列，使用bwa软件（网址为<https://sourceforge.net/projects/bio-bwa/>）将上一步过滤后的样本序列比对至各HLA基因参考基因（使用默认参数进行比对），将比对上的测序序列拆分至HLA-A、HLA-B、HLA-C、HLA-DRB1、HLA-DQB1共5个测序数据文件。
- [0097] 4、对比供者与受者各基因中不同型别之间的SNP差异，获得snp文件
- [0098] 使用MUSCLE软件（网址为<http://www.drive5.com/muscle/>）对比供者与受者A

基因中3个型别A*02:03, A*30:01, A*11:01的SNP差异(使用默认参数进行比对), 共有6个SNP, 保存为A.snp文件; 对比供者与受者B基因中3个型别B*13:02, B*18:01, B*15:02的SNP差异, 共有6个SNP, 保存为B.snp文件; 对比供者与受者C基因中3个型别C*06:02, C*07:04, C*08:01的SNP差异, 共有2个SNP, 保存为C.snp文件; 对比供者与受者DRB1基因中3个型别DRB1*07:01, DRB1*14:04, DRB1*13:02的SNP差异, 共有9个SNP, 保存为DRB1.snp文件; 对比供者与受者DQB1基因中3个型别DQB1*05:03, DQB1*02:02, DQB1*06:04的SNP差异, 共有14个SNP, 保存为DQB1.snp文件。

- [0099] 5、将测序数据文件分别与对应基因的参考序列进行比对, 生成各基因比对文件
- [0100] 由IMGT-HLA公共数据库(网址为<https://www.ebi.ac.uk/ipd/imgt/hla/>)查询获得A*02:03:01, B*13:02:01:01, C*06:02:01:01, DRB1*07:01:01:01, DQB1*05:03:01:01参考序列, 使用bwa软件(网址为<https://sourceforge.net/projects/bio-bwa/>)的mem算法默认参数将5个测序数据文件中序列按基因分别比对至A*02:03:01, B*13:02:01:01, C*06:02:01:01, DRB1*07:01:01:01, DQB1*05:03:01:01参考序列, 生成各基因比对文件A.bam、B.bam、C.bam、DRB1.bam、DQB1.bam。
- [0101] 6、获得基因比对文件中各SNP位置测序深度, 计算各基因深度占比的平均值
- [0102] 进一步, 调取snp文件, 使用bamreadcount软件(网址为<https://github.com/genome/bam-readcount>)查看bam文件中各SNP位置测序深度, 计算各位置属于受者的型别的深度占比, A基因6个SNP占比分别为0.53%, 0.52%, 0.35%, 0.35%, 0.52%, 0.38%, 平均数0.44%; B基因6个SNP中共有5个SNP被检测, 占比分别为0.11%, 0.16%, 0.41%, 0.44%, 0.29%, 平均数0.28%; C基因2个SNP占比分别为0.21%, 0.45%, 平均数0.33%; DRB1基因9个SNP中共有2个SNP被检测, 占比分别为0.00%, 0.43%, 平均数0.22%; DQB1基因14个SNP中共有7个SNP被检测, 占比分别为0.37%, 0.13%, 0.14%, 0.24%, 0.14%, 0.08%, 0.24%, 平均数0.19%。
- [0103] 7、获得全部基因的平均值, 判断移植受者术后HLA染色体区域是否发生杂合性

缺失

[0104] 进一步,取5个基因平均值0.29%。根据本申请阴阳性判断方法移植受者术后HLA染色体区域占比0.29%小于0.5%,且嵌合率百分比6.85%大于3%,判断该移植受者为HLA Loss阳性,结果表明该检测者移植术后HLA染色体区域已发生杂合性缺失。

[0105] 实施例3:验证本申请算法灵敏度与特异性

[0106] 本实施例通过人为地等比例混合两个健康人样本,模拟不同的供者与受者嵌合率,采用本申请算法检测在各浓度梯度时两个样本测序结果占比,通过线性拟合验证本申请算法灵敏度与特异性。

[0107] 本实施例随机挑选了两例健康人血液标本,样本一的HLA配型结果为A*26:01, A*31:01, B*40:06, B*40:06, C*08:01, C*15:02, DRB1*09:01, DRB1*15:02, DQB1*06:02, DQB1*03:03, 样本二的HLA配型结果为A*24:02, A*24:02, B*40:01, B*40:03, C*03:04, C*07:02, DRB1*12:01, DRB1*15:01, DQB1*06:02, DQB1*03:01, 采用分光光度计检测样本DNA浓度后,根据DNA浓度采用浓度梯度稀释法将两个样本混合成1:1、1:4、1:8、1:16、1:32、1:64六个不同比例的样本,再进行二代建库及测序。

[0108] 混合样本中样本一理论比例分别为50.00%, 25.00%, 12.50%, 6.25%, 3.12%, 1.57%。应当指出该比例仅为参考值,由分光光度计检测吸光度后换算,非实际比例,仅作为混合样品参考,不作为本申请算法有效性的参考值。

[0109] 采用本申请分析方法对各比例样本下机测序数据进行分析,各比例各基因样本一占比参见表1,进一步对结果平均值进行线性拟合,结果参见附图7,拟合结果显示利用本申请算法进行HLA染色体区域的杂合性缺失检测具有非常高的灵敏度与特异性。

[0110] 表1:实施例3梯度混合结果表

[]

[表1]

理论比例	A结果	B结果	C结果	DRB1结果	DQB1结果	结果平均值
50.00%	49.49%	49.01%	45.95%	46.34%	45.78%	47.31%
25.00%	29.67%	28.77%	25.77%	27.59%	30.27%	28.41%
12.50%	18.14%	19.41%	13.88%	14.92%	17.54%	16.78%
6.25%	11.42%	10.04%	7.62%	7.03%	10.14%	9.25%
3.12%	8.74%	5.43%	4.65%	3.20%	7.15%	5.83%
1.57%	5.07%	2.93%	2.30%	1.48%	4.23%	3.20%

权利要求书

- [权利要求 1] 一种HLA染色体区域的杂合性缺失的分析方法，其中，包括以下步骤：
- 对样本数据进行拆分和过滤，得到过滤后的样本序列；
- 获得各HLA基因测序数据文件，所述测序数据文件包含各HLA基因的样本数据；
- 生成受者移植术前第一个等位基因序列作为参考基因；
- 获得各HLA基因中受者移植术后所占百分比；
- 获得受者移植术后HLA基因染色体区域的百分比，以HLA%表示；
- 获得受者移植术后细胞中各染色体总体百分比，以STR%表示；
- 受者移植术后HLA Loss阴阳性判断。
- [权利要求 2] 根据权利要求1所述的分析方法，其中，各步骤具体为：
- 对测序完成后的样本下机数据进行拆分和过滤，得到过滤后的样本序列；
- 根据设定的比对参数，将过滤后的样本序列比对至各HLA基因参考基因序列，将比对上的测序序列拆分至各HLA测序数据文件，所述测序数据文件包含测序完成后各HLA基因的样本下机数据；
- 获取移植术前受者和供者各HLA等位基因型别，比对后，获得受者与供者之间各HLA基因的SNP差异，生成受者移植术前HLA基因型的第一个等位基因序列作为参考基因；
- 将所述测序数据文件比对至所述参考基因，统计各HLA基因SNP位置的受者及供者分别的测序深度，将各HLA基因全部SNP位置受者的深度频率取平均值，得到各HLA基因中受者移植术后所占百分比；
- 取各HLA基因受者百分比的平均值，即为受者移植术后HLA染色体区域的百分比，以HLA%表示；
- 获得受者移植术后细胞中各染色体总体百分比，以STR%表示；
- 阴阳性判断：当 $HLA\% \leq 0.5\%$ ，且 $STR\% \geq 3\%$ 时，判断为HLA Loss阳性；当 $HLA\% \geq 3\%$ 时，判断为HLA Loss阴性；当 $0.5\% \leq HLA\% < 3\%$ 时提示

- 无法判断，受者移植术后细胞HLA染色体可能处于正在缺失时期。
- [权利要求 3] 根据权利要求1或2所述的分析方法，其中，还包括以下步骤：
将HLA各基因百分比、平均百分比、阴阳性判断结果输出至报告文件。
。
- [权利要求 4] 根据权利要求1-3中任一项所述的分析方法，其中，各HLA基因分别为HLA-A、HLA-B、HLA-C、HLA-DRB1、HLA-DQB1和HLA-DPB1基因。
- [权利要求 5] 根据权利要求1-4中任一项所述的分析方法，其中，HLA-A基因比对至A*01:01:01:01、HLA-B基因比对至B*07:02:01:01、HLA-C基因比对至C*01:02:01:01、HLA-DRB1基因比对至DRB1*01:02:01:01、HLA-DQB1基因比对至DQB1*05:01:01:01、HLA-DPB1基因比对至DPB1*01:01:01:01。
- [权利要求 6] 一种HLA染色体区域的杂合性缺失的分析处理装置，包括以下模块：
拆分和过滤模块，其对样本数据进行拆分和过滤，得到过滤后的样本序列；
基因测序数据文件获取模块，用于获得各HLA基因测序数据文件，所述测序数据文件包含各HLA基因的样本数据；
参考基因生成模块，用于生成受者移植术前第一个等位基因序列作为参考基因；
各HLA基因所占百分比计算模块，用于获得各HLA基因中受者移植术后所占百分比；
HLA基因染色体区域百分比计算模块，用于获得受者移植术后HLA基因染色体区域的百分比，以HLA%表示；
HLA基因各染色体总体百分比计算模块，用于获得受者移植术后细胞中各染色体总体百分比，以STR%表示；
阴阳性判断模块，用于进行受者移植术后HLA Loss阴阳性判断。
- [权利要求 7] 根据权利要求6所述的分析处理装置，其中，各模块具体为：
拆分和过滤模块，其对测序完成后的样本下机数据进行拆分和过滤，得到过滤后的样本序列；

基因测序数据文件获取模块，其根据设定的比对参数，将过滤后的样本序列比对至各HLA基因参考基因序列，将比对上的测序序列拆分至各HLA测序数据文件，所述测序数据文件包含测序完成后各HLA基因的样本下机数据；

参考基因生成模块，用于获取移植术前受者和供者各HLA等位基因型别，比对后，获得受者与供者之间各HLA基因的SNP差异，生成受者移植术前HLA基因型第一个等位基因序列作为参考基因；

各HLA基因所占百分比计算模块，用于将所述测序数据文件比对至所述参考基因，统计各HLA基因SNP位置的受者及供者分别的测序深度，将各HLA基因全部SNP位置受者的深度频率取平均值，得到各HLA基因中受者移植术后所占百分比；

HLA基因染色体区域百分比计算模块，用于取各HLA基因受者百分比的平均值，即为受者移植术后HLA染色体区域的百分比，以HLA%表示；

HLA基因各染色体总体百分比计算模块，用于获得受者移植术后细胞中各染色体总体百分比，以STR%表示；

阴阳性判断模块，用于进行受者移植术后HLA

Loss阴阳性判断，当 $HLA\% \leq 0.5\%$ ，且 $STR\% \geq 3\%$ 时，判断为HLA Loss阳性；当 $HLA\% \geq 3\%$ 时，判断为HLA Loss阴性；当 $0.5\% \leq HLA\% < 3\%$ 时提示无法判断，受者移植术后细胞HLA染色体可能处于正在缺失时期。

- [权利要求 8] 根据权利要求6或7所述的分析处理装置，其中，还包括以下模块：
报告文件生成模块，用于将HLA各基因百分比、平均百分比、阴阳性判断结果输出至报告文件。
- [权利要求 9] 根据权利要求6-8中任一项所述的分析处理装置，其中，各HLA基因分别为HLA-A、HLA-B、HLA-C、HLA-DRB1、HLA-DQB1和HLA-DPB1基因。
- [权利要求 10] 根据权利要求6-9中任一项所述的分析处理装置，其中，HLA-A基因比对至A*01:01:01:01、HLA-B基因比对至B*07:02:01:01、HLA-C基因比对至C*01:02:01:01、HLA-DRB1基因比对至DRB1*01:02:01:01、HLA-D

QB1基因比对至DQB1*05:01:01:01、HLA-DPB1基因比对至DPB1*01:01:01:01。