

# (12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织  
国际局

(43) 国际公布日  
2020 年 1 月 9 日 (09.01.2020)



(10) 国际公布号  
**WO 2020/007089 A1**

(51) 国际专利分类号:

C12Q 1/6806 (2018.01) C12Q 1/6855 (2018.01)

C12Q 1/6827 (2018.01) C12Q 1/6876 (2018.01)

C12Q 1/6886 (2018.01)

(21) 国际申请号: PCT/CN2019/082233

(22) 国际申请日: 2019 年 4 月 11 日 (11.04.2019)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:

201810712104.3 2018 年 7 月 3 日 (03.07.2018) CN

(71) 申请人: 中国医学科学院肿瘤医院 (CANCER HOSPITAL, CHINESE ACADEMY OF MEDICAL SCIENCES) [CN/CN]; 中国北京市朝阳区潘家园南里 17 号中国医学科学院肿瘤医院, Beijing 100021 (CN)。北京泛生子基因科技有限公司 (GENETRON HEALTH (BEIJING) CO, LTD.) [CN/CN]; 中国北京市昌平区生命科学园生命园路 8 号院一区 11 号楼, Beijing 102206 (CN)。

(72) 发明人: 焦宇辰(JIAO, Yuchen); 中国北京市朝阳区潘家园南里 17 号中国医学科学院肿瘤医院, Beijing 100021 (CN)。曲春枫(QU, Chunfeng); 中国北京市朝阳区潘家园南里 17 号中国医学科学院肿瘤医院, Beijing 100021 (CN)。王沛(WANG, Pei); 中国北京市昌平区生命科学园生命园路 8 号院一区 11 号楼, Beijing 102206 (CN)。陈坤(CHEN, Kun); 中国北京市朝阳区潘家园南里 17 号中国医学科学院肿瘤医院, Beijing 100021 (CN)。王宇婷(WANG, Yuting); 中国北京市昌平区生命科学园生命园路 8 号院一区 11 号楼, Beijing 102206 (CN)。宋欠欠(SONG, Qianqian); 中国北京市昌平区生命科学园生命园路 8 号院一区 11 号楼, Beijing 102206 (CN)。王思振(WANG, Sizhen); 中国北京市昌平区生命科学园生命园路 8 号院一区 11 号楼, Beijing 102206 (CN)。阎海

(YAN, Hai); 中国北京市昌平区生命科学园生命园路 8 号院一区 11 号楼, Beijing 102206 (CN)。

(74) 代理人: 北京纪凯知识产权代理有限公司(JEEKAI & PARTNERS); 中国北京市丰台区广安路 9 号 5 号楼 15A 层, Beijing 100055 (CN)。

(81) 指定国(除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW。

(84) 指定国(除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

- 包括国际检索报告(条约第 21 条(3))。
- 包括说明书序列表部分(细则 5.2(a))。

(54) Title: CONSTRUCTION AND SEQUENCING DATA ANALYSIS METHOD FOR CTDNA LIBRARY FOR SIMULTANEOUSLY DETECTING VARIOUS COMMON MUTATIONS IN LIVER CANCER

(54) 发明名称: 一种同时检测多种肝癌常见突变的ctDNA文库构建和测序数据分析方法

(57) Abstract: Provided in the present invention is a construction and sequencing data analysis method for a ctDNA library for simultaneously detecting various common mutations in liver cancer, the method comprising sequentially subjecting a DNA sample to end repair and 3' end plus A treatment, ligating the treated DNA sample to a linker mixture, and obtaining the library by means of PCR amplification.

(57) 摘要: 本发明提供了一种同时检测多种肝癌常见突变的ctDNA文库构建和测序数据分析方法, 包括将DNA样本依次进行末端修复和3'端加A处理, 将处理后的DNA样本与接头混合物连接, 经过PCR扩增后得到文库。



WO 2020/007089 A1

## 一种同时检测多种肝癌常见突变的 ctDNA 文库构建和测序数据分析方法

### 技术领域

5 本发明涉及一种同时检测多种肝癌常见突变的 ctDNA 文库构建和测序数据分析方法。

### 背景技术

ctDNA (circulating tumor DNA), 即循环肿瘤 DNA, 指存在于血液、脑脊液等体液中, 游离于细胞外的肿瘤 DNA。ctDNA 通常和血液中来源于正常细胞的游离 DNA 混杂在一起, 被称为 cfDNA (cellfree DNA 增加空格)。通过检测 ctDNA 10 的突变, 可指导靶向用药, 治疗的监测以及癌症的早期筛查等等。基于 ctDNA 的检测方法包括 1) 基于 PCR 的热点突变检测法, 通常检测一个或几个热点突变或已知突变, 无法检测基因融合等复杂突变, 无法检测未知突变。2) 捕获/二代测序法: 可以检测较多基因的位置突变, 包括复杂突变, 但捕获试剂盒一般价格昂贵、操作复杂、耗时较长。在以上两种方法的背景下, 目前 ctDNA 检测 15 存在以下困难: 1) 一次抽血获得 ctDNA 标本量有限, 通常仅够支持一次检测, 这就造成 ctDNA 临床检测中通常是单平台、一次性的, 使用低成本的热点突变法检测到一个突变就无法再检测其他突变。而临床检测中经常需要根据首次检测的结果判断后续检测的目标和方案, 这就需要在后续检测中重新抽血。另外, ctDNA 相关临床检测或研究中经常需要比较多种技术的优劣, 这就需要数倍于正 20 常抽血量的标本, 患者通常无法接受。2) 无论是 PCR 法还是捕获法, 扩增过程中产生的噪音突变会严重干扰 ctDNA 低频突变的检测, 造成假阳性结果, 误导患者的诊断和治疗。3) ctDNA 突变含量低, 操作过程中易发生污染, 造成假阳性结果。

肝癌是世界第五高发和第二致死的常见肿瘤, 世界一半以上的肝癌发生在中国, 且以乙肝相关性肝癌为主。乙肝相关性肝癌几乎没有 KRAS、BRAF 等热点 25 突变, 其突变主要为 TP53、CTNNB1 等几个基因的编码区突变, TERT 高 GC 启动子区突变, 还包括 HBV 整合、TERT 拷贝数变异等复杂突变形式。使得目前没有简单、低成本、可靠的方案系统检测肝癌 ctDNA 突变。通过 ctDNA 检测对肝癌进行早期筛查、病情追踪、疗效评估、预后预测等具有重要临床意义。

### 30 发明公开

本发明的目的是提供一种同时检测多种肝癌常见突变的 ctDNA 文库构建和测序数据分析方法。

本发明提供一种测序文库构建方法, 依次可包括如下步骤:

- (1) 将 DNA 样本依次进行末端修复和 3' 端加 A 处理;
- 35 (2) 将步骤 (1) 处理后的 DNA 样本与接头混合物连接, 经过 PCR 扩增后得到文库;

所述接头混合物可由 n 个接头组成;

每个接头均由一条上游引物甲和一条下游引物甲形成部分双链结构得到；上游引物甲中具有测序接头甲、随机标签、锚定序列甲和位于 3' 末端的碱基 T；下游引物甲中具有锚定序列乙和测序接头乙；所述部分双链结构由上游引物中的锚定序列甲和下游引物中的锚定序列乙反向互补形成；

- 5        所述测序接头甲和测序接头乙为根据不同测序平台选择对应的测序接头；  
所述随机标签可为 8-14bp 的随机碱基；

所述锚定序列甲长度可为 14-20bp，连续重复碱基 $\leq 3$  个；

n 个接头采用 n 个不同的锚定序列甲，并且相同位置的碱基平衡，错配碱基数 $>3$ ；

- 10        n 可为 $\geq 8$  的任意自然数。

所述锚定序列不与引物其它部分相互作用（如形成发卡结构、二聚体等）。

所述上游引物甲自 5' 端依次可为测序接头甲、随机标签、锚定序列甲和碱基 T。

所述下游引物甲自 5' 端依次可为锚定序列乙和测序接头乙。

- 15        上文中，通常建库用的接头由两条序列退火而成，呈“Y”型结构，两条序列间互补配对的部分（即锚定序列甲和锚定序列乙）称之为锚定序列。所述锚定序列可作为序列固定的内置标签，用于标记原始模板分子。

所述“相同位置的碱基平衡”可为所述接头混合物中的 n 个锚定序列甲之间从起始碱基至末尾碱基的每个位置碱基平衡，即 A、T、C 和 G 均匀分布。

- 20        所述“错配碱基数 $>3$ ”可为所述接头混合物包含 n 个锚定序列甲，各个锚定序列甲之间的碱基至少有 3 个不同。该不同可为位置不同或顺序不同。

所述 DNA 样本可为基因组 DNA、cDNA、ct DNA 或 cf DNA 样本。

所述 n 具体可为 12。

所述随机标签具体可为 8bp 的随机碱基。

- 25        所述锚定序列甲长度具体可为 12bp。

当 n=12 时，所述锚定序列甲的核苷酸序列具体分别可为序列表的序列 1 自 5' 端第 30-41 位、序列表的序列 3 自 5' 端第 30-41 位、序列表的序列 5 自 5' 端第 30-41 位、序列表的序列 7 自 5' 端第 30-41 位、序列表的序列 9 自 5' 端第 30-41 位、序列表的序列 11 自 5' 端第 30-41 位、序列表的序列 13 自 5' 端第 30-41 位、序列表的序列 15 自 5' 端第 30-41 位、序列表的序列 17 自 5' 端第 30-41 位、序列表的序列 19 自 5' 端第 30-41 位、序列表的序列 21 自 5' 端第 30-41 位、序列表的序列 23 自 5' 端第 30-41 位所示。

- 30        所述测序接头甲具体可为来自 Illumina 公司的 Truseq 测序试剂盒的测序接头。所述测序接头甲具体可如序列表的序列表的序列 1 自 5' 端第 1-29 位所示。  
35        所述测序接头乙具体可为来自 Illumina 公司的 nextera 测序试剂盒的测序接头。所述测序接头乙具体可如序列表的序列表的序列 2 自 5' 端第 13-41 位所

示。

当  $n=12$  时，所述 12 个接头如下：

接头 1 可由序列表的序列 1 所示的单链 DNA 分子和序列 2 所示的单链 DNA 分子形成部分双链结构得到；接头 2 可由序列表的序列 3 所示的单链 DNA 分子和序列 4 所示的单链 DNA 分子形成部分双链结构得到；接头 3 可由序列表的序列 5 所示的单链 DNA 分子和序列 6 所示的单链 DNA 分子形成部分双链结构得到；接头 4 可由序列表的序列 7 所示的单链 DNA 分子和序列 8 所示的单链 DNA 分子形成部分双链结构得到；接头 5 可由序列表的序列 9 所示的单链 DNA 分子和序列 10 所示的单链 DNA 分子形成部分双链结构得到；接头 6 可由序列表的序列 11 所示的单链 DNA 分子和序列 12 所示的单链 DNA 分子形成部分双链结构得到；接头 7 可由序列表的序列 13 所示的单链 DNA 分子和序列 14 所示的单链 DNA 分子形成部分双链结构得到；接头 8 可由序列表的序列 15 所示的单链 DNA 分子和序列 16 所示的单链 DNA 分子形成部分双链结构得到；接头 9 可由序列表的序列 17 所示的单链 DNA 分子和序列 18 所示的单链 DNA 分子形成部分双链结构得到；接头 10 可由序列表的序列 19 所示的单链 DNA 分子和序列 20 所示的单链 DNA 分子形成部分双链结构得到；接头 11 可由序列表的序列 21 所示的单链 DNA 分子和序列 22 所示的单链 DNA 分子形成部分双链结构得到；接头 12 可由序列表的序列 23 所示的单链 DNA 分子和序列 24 所示的单链 DNA 分子形成部分双链结构得到。

所述接头可以通过将上游引物甲和下游引物甲进行退火得到。

所述接头混合物中，每个接头可等摩尔混合。

所述方法还可包括对步骤（2）得到的文库进行扩增的步骤。所述扩增的引物是根据接头序列设计的。具体的，所述扩增的引物至少要有一段序列与接头的某段序列匹配。所述扩增采用的引物对具体可由序列表的序列 25 和序列 26 所示的两条单链 DNA 分子组成。

本发明还保护以上所述的方法构建得到的 DNA 文库。

本发明还保护一种用于构建测序文库的试剂盒，包括以上任一所述的接头混合物。

所述试剂盒还可包括用于 DNA 提取的试剂、用于 DNA 建库的试剂、用于文库纯化的试剂、用于文库捕获的试剂等用于文库构建的材料。

本发明还保护一种用于检测 DNA 样本中肝癌突变的试剂盒，包括以上任一所述接头混合物和引物组合；所述引物组合包括引物组 I、引物组 II、引物组 III 和引物组 IV；

所述引物组 I 和所述引物组 II 中的各个引物是根据与肝癌相关的区域设计的特异性引物，作用是定位于基因组特定位置，实现目标区域的 PCR 富集；

所述引物组 III 和所述引物组 IV 中的各个引物的核苷酸序列均由“接头序列+特异序列”组成，特异序列用于目标区域进行进一步富集，接头序列用于通过

PCR 形成完整的可以测序的文库分子；

所述引物组III和所述引物组 I 可为“巢式”关系；所述引物组IV和所述引物组 II 可为“巢式”关系。

5 所述与肝癌相关的区域具体可为肝癌高频突变基因（TP53、CTNNB1、AXIN1、TERT）的相关区域和针对 HBV 整合热点区域。

所述引物组 I 具体可由序列表的序列 28 至序列 105 所示的单链 DNA 组成。

所述引物组 II 具体可由序列表的序列 106 至序列 187 所示的单链 DNA 组成。

所述引物组III具体可由序列表的序列 191 至序列 265 所示的单链 DNA 组成。

所述引物组IV具体可由序列表的序列 266 至序列 344 所示的单链 DNA 组成。

10 所述引物组合具体可由所述引物组 I、所述引物组 II、所述引物组III和所述引物组IV组成。

所述试剂盒还可包括用于 DNA 提取的试剂、用于 DNA 建库的试剂、用于文库纯化的试剂、用于文库捕获的试剂等用于文库构建的材料。

15 本发明还保护上述任一所述引物组合。所述引物组合的用途可为制备用于检测 DNA 样本中肝癌突变的试剂盒。

本发明还保护上述任一所述引物组合在制备用于检测 DNA 样本中肝癌突变的试剂盒中的应用。

本发明还保护一种检测 DNA 样本中目标突变的方法，包括如下步骤：

（1）按照上述任一所述的方法构建文库；

20 （2）对步骤（1）得到的文库进行两轮巢式 PCR 扩增，对产物进行测序，根据测序结果分析 DNA 样本中目标突变发生情况；

所述步骤（2）中，采用引物组合甲进行第一轮 PCR 扩增；

引物组合甲由上游引物甲和下游引物组合甲组成；

所述上游引物甲为文库扩增引物，用于步骤（1）的文库扩增；

25 所述下游引物组合甲为根据 n 个目标靶点设计的 n 条引物的组合；

以第一轮 PCR 的产物为模板，采用引物组合乙进行第二轮 PCR 扩增；

引物组合乙由上游引物乙、下游引物组合乙和 index 引物组成；

所述上游引物乙部分序列为文库扩增引物，用于第一轮 PCR 的产物的扩增；

30 所述下游引物组合乙中的引物与下游引物组合甲中检测相同目标靶点的引物形成巢式关系，各条引物中具有与 index 引物结合的区段；

所述 index 引物中包含与下游引物组合乙中各条引物结合的区段和 index 序列。

所述上游引物甲的部分序列与“每个接头的上游引物甲的测序接头甲”的序列完全相同。

35 所述上游引物乙用于补全文库分子的接头序列，使得扩增产物可以直接进行测序。所述上游引物乙和所述上游引物甲（第一轮 PCR 扩增使用的引物）的部分核苷酸序列完全相同。

所述上游引物甲的核苷酸序列具体可如序列表的序列 27 所示。

所述上游引物乙的核苷酸序列具体可如序列表的序列 188 所示。

所述 index 引物自 5' 端可包括区段 A、index 序列和区段 B。所述 index 引物具体可由所述区段 A、所述 index 序列和所述区段 B 组成。所述区段 A 的核苷酸序列可如序列表的序列 189 所示。所述区段 B 的核苷酸序列可如序列表的序列 190 所示。

当所述目标突变为肝癌突变时，所述引物组合甲由以上任一所述引物组 I 和引物组 II 组成；所述引物组合乙由以上任一所述引物组 III 和引物组 IV 组成。使用引物组 I 和引物组 II 分别对模板进行第一轮 PCR 扩增，使用引物组 I 进行扩增的产物作为第二轮扩增的模板采用引物组 III 进行扩增，使用引物组 II 进行扩增的产物作为第二轮扩增的模板采用引物组 IV 进行扩增，然后将扩增产物等体积混合。

所述测序结果的分析方法可为：将随机标签序列相同、DNA 插入片段长度相同、DNA 插入片段两端断点（即 DNA 片段的序列一致（除所含的突变））均相同的 DNA 分子测序数据回溯到一个分子簇，如果簇内分子数 > 5 个并且簇内分子突变一致率 > 80% 并且簇个数 ≥ 5，则该突变为来自原始 DNA 样本的真实突变。

本发明还保护一种检测 DNA 样本中多种目标突变的方法，可包括如下步骤：

（1）按照上述任一所述的方法构建文库；

（2）对步骤（1）的文库进行靶区域富集并进行测序，根据测序结果分析 DNA 样本中目标突变发生情况。

所述靶区域富集可采用现有商购的靶向捕获试剂盒（例如 Agilent sureselect XT 靶向捕获试剂盒，Agilent 5190-8646）进行，将其中最后一步 PCR 扩增的引物对更换为由引物甲和引物乙组成的引物对。所述引物甲的核苷酸序列可如序列表的序列 345 所示。所述引物乙可包括区段甲、index 序列和区段乙。所述引物乙具体可由所述区段甲、所述 index 序列和所述区段乙组成。所述区段甲的核苷酸序列可如序列表的序列 346 所示。所述区段乙的核苷酸序列可如序列表的序列 347 所示。

所述测序结果的分析方法可为：将 DNA 插入片段长度相同、DNA 插入片段两端断点、两端的锚定序列均相同的初始 DNA 单链测序数据回溯到一个分子簇；将插入片段长度一致并且除突变点外的序列一致、分子簇两端的锚定序列相同但位置相反的同一起始 DNA 双链的分子簇标记为一对 duplex 分子簇；对某一突变来说，若至少有一对 duplex 分子簇支持，即可判断为真，若无 duplex 分子簇支持，那么至少有 4 个分子簇支持，可判断为真。

上文中，通常多个不同样本的文库会混合在一起测序，所述 index 序列用于标记不同的样本。测序完成后，根据不同的 index 序列对总的测序数据进行拆分。Index 的设计原则与之前描述的锚定序列的设计原则基本相似。

本发明由于采取以上技术方案，其具有以下优点：

1. 在不需要捕获的情况下同时检测肝癌 ctDNA 中的点突变、插入缺失突变、HBV 整合等多种突变形式。与捕获法相比，此技术只需数条 DNA 引物，不需昂贵的捕获探针和杂交试剂，成本极大降低；操作流程简单，可以从捕获法的 36 小时缩短到 8 小时。

5        2. 适合超小靶区的高效捕获，靶区可以小至捕获法最小靶区的 10%，极大的提高了测序效率。例如，肝癌常见突变 TP53、CTNNB1、AXIN1、TERT、HBV 整合的组合就是一个适合本技术的超小靶区，使用捕获法富集此靶区的上靶率 <10%，而本技术可以达到 80% 以上，极大的提高了测序效率，降低测序成本。

3. 在进行一次检测后，扩增后的文库可以支持 10-20 次的后续检测，每次检测的结果都可以代表全部原始 ctDNA 标本的突变状况，不会造成灵敏度和特异性的降低。

4. 在文库构建过程中将 DNA 条码 barcode 连接到起始 ctDNA 分子上，配合生物信息分析流程，实现 ctDNA 低频突变的高特异性检测。

5. 本技术所构建的文库可以同时用于 PCR 的热点检测和捕获法测序，且一份标本构建出来的文库可以同时支持多次检测，加入的 DNA barcode 可以有效滤除假阳性突变，实现基于 duplex 高特异性测序。

本发明对于肝癌的早期筛查、病情追踪、疗效评估、预后预测等具有重要临床意义。

### 附图说明

20        图 1 为 adapter 和引物架构示意图。

图 2 为 RaceSeq 靶区富集和文库构建示意图。

图 3 为 MC 文库进行捕获和进行 duplex 测序的示意图。

### 实施发明的最佳方式

25        以下的实施例便于更好地理解本发明，但并不限定本发明。下述实施例中的实验方法，如无特殊说明，均为常规方法。下述实施例中所用的试验材料，如无特殊说明，均为自常规生化试剂商店购买得到的。以下实施例中的定量试验，均设置三次重复实验，结果取平均值。

实施例 1、MC 文库的构建

一、cfDNA 分子的平末端修复和加 A 处理

30        取 10-45ng cfDNA，按照表 1 所示配置反应体系，然后按照表 2 的程序在 PCR 仪进行末端修复及 3' 末端加 A 处理，得到反应产物（4℃ 保存）。

表 1. 反应体系

| 成分                                              | 体积         |
|-------------------------------------------------|------------|
| cfDNA                                           | 50 $\mu$ l |
| End Repair & A-Tailing Buffer (KAPA KK8505)     | 7 $\mu$ l  |
| End Repair & A-Tailing Enzyme Mix (KAPA KK8505) | 3 $\mu$ l  |
| 总体积                                             | 60 $\mu$ l |

表 2. 反应程序

| 温度  | 时间    |
|-----|-------|
| 20℃ | 30min |

|     |       |
|-----|-------|
| 65℃ | 30min |
|-----|-------|

## 二、cfDNA 与 adapter 连接

按照表 3 配置反应体系，20℃ 反应 15min，得到连接产物（4℃ 保存）。

表 3. 反应体系

| 成分                           | 体积          |
|------------------------------|-------------|
| 步骤一得到的反应产物                   | 60 $\mu$ l  |
| Adapter Mix (50 $\mu$ M)     | 1.5 $\mu$ l |
| DNase/RNase-Free Water       | 8.5 $\mu$ l |
| Ligation Buffer(KAPA KK8505) | 30 $\mu$ l  |
| DNA Ligase(KAPA KK8505)      | 10 $\mu$ l  |
| 总体积                          | 110 $\mu$ l |

Adapter Mix 序列信息见表 4。

- 5 将表 4 中的单链 DNA 用 TE 溶解稀释至终浓度为 100  $\mu$ M。将同一组中的两条单链 DNA 等体积混合（各 50  $\mu$ l），进行退火（退火程序：95℃，15min；25℃，2h），得到 12 组 DNA 溶液，将 12 组 DNA 溶液等体积混合，得到 Adapter Mix。

表 4. adapterMix 序列信息

| 分组 | 编号 | 名称    | 序列（5' -3'）                                       |
|----|----|-------|--------------------------------------------------|
| 1  | 1  | R21_F | GACACGACGCTCTCCGATCTNNNNNNNNCCACTAGTAGCCT（序列 1）  |
|    | 2  | R21_R | GGCTACTAGTGGCTGTCTCTTATACACATCTCCGAGCCCAC（序列 2）  |
| 2  | 3  | R22_F | GACACGACGCTCTCCGATCTNNNNNNNNGGACTGTGTCGGT（序列 3）  |
|    | 4  | R22_R | CCGACACAGTCCCTGTCTCTTATACACATCTCCGAGCCCAC（序列 4）  |
| 3  | 5  | R23_F | GACACGACGCTCTCCGATCTNNNNNNNNGGTACTGACAGGT（序列 5）  |
|    | 6  | R23_R | CCTGTCAGTACCCTGTCTCTTATACACATCTCCGAGCCCAC（序列 6）  |
| 4  | 7  | R24_F | GACACGACGCTCTCCGATCTNNNNNNNNCCTAGTACAGCCT（序列 7）  |
|    | 8  | R24_R | GGCTGTACTAGGCTGTCTCTTATACACATCTCCGAGCCCAC（序列 8）  |
| 5  | 9  | R25_F | GACACGACGCTCTCCGATCTNNNNNNNNGGTAGTCAGAGGT（序列 9）  |
|    | 10 | R25_R | CCTCTGACTACCCTGTCTCTTATACACATCTCCGAGCCCAC（序列 10） |
| 6  | 11 | R26_F | GACACGACGCTCTCCGATCTNNNNNNNNTTCTCAGTGTTT（序列 11）  |
|    | 12 | R26_R | AACACGTGAGAACTGTCTCTTATACACATCTCCGAGCCCAC（序列 12） |
| 7  | 13 | R27_F | GACACGACGCTCTCCGATCTNNNNNNNNAACTCCACGTAAT（序列 13） |
|    | 14 | R27_R | TTACGTGGAGTTCTGTCTCTTATACACATCTCCGAGCCCAC（序列 14） |
| 8  | 15 | R28_F | GACACGACGCTCTCCGATCTNNNNNNNNTTCTCGAGAATTT 序列 15） |
|    | 16 | R28_R | AATTCTCGAGAACTGTCTCTTATACACATCTCCGAGCCCAC（序列 16） |
| 9  | 17 | R29_F | GACACGACGCTCTCCGATCTNNNNNNNNAACTCTTCCAAT（序列 17）  |

|    |    |       |                                                           |
|----|----|-------|-----------------------------------------------------------|
|    | 18 | R29_R | <u>TTGGAAGAGTTTCTGTCTCTTATACACATCTCCGAGCCCAC</u> (序列 18)  |
| 10 | 19 | R30_F | GACACGACGCTCTTCCGATCTNNNNNNNNNTTGAACGTCTTT (序列 19)        |
|    | 20 | R30_R | <u>AAGACGTTC</u> CAACTGTCTCTTATACACATCTCCGAGCCCAC (序列 20) |
| 11 | 21 | R31_F | GACACGACGCTCTTCCGATCTNNNNNNNNCCGGACTCCTCCT (序列 21)        |
|    | 22 | R31_R | <u>GGAGGAGTCCGGCTGTCTCTTATACACATCTCCGAGCCCAC</u> (序列 22)  |
| 12 | 23 | R32_F | GACACGACGCTCTTCCGATCTNNNNNNNNAAGGAGGAGTAAT (序列 23)        |
|    | 24 | R32_R | <u>TTACTCCTCCTTCTGTCTCTTATACACATCTCCGAGCCCAC</u> (序列 24)  |

表 4 中,8 个 N 表示 8bp 的随机标签。实际应用中,随机标签长度可为 8-14bp。

下划线表示 12bp 的锚定序列,每一组的上游序列和下游序列中,下划线部分反向互补,通过退火可使上游序列和下游序列结合在一起形成接头。同时,锚定序列可作为序列固定的内置标签,用于标记原始模板分子。实际应用中,锚定序列长度可为 12-20bp,连续重复碱基不超过 3 个,且不能与引物其它部分相互作用(如形成发卡结构、二聚体等),12 组每一个位置碱基平衡, mismatch 碱基数>3。

上游序列中末端加粗的 T 与原始分子末端加的“A”互补,进行 TA 连接。

上游序列中,自 5' 端第 1 至 21 位为(来自 Illumina 公司的 Truseq 测序试剂盒)为测序引物结合序列,其中,自 5' 端第 1 至 19 位为文库扩增引物部分。

下游序列中,非下划线部分(来自 Illumina 公司的 nextera 测序试剂盒)为测序引物结合序列,其中,自 3' 端第 1 至 22 位为设计文库扩增引物的部分。

表 4 中共包含 12 组接头,可以形成 12×12=144 种标记组合,结合分子本身的序列信息,足以区分原始样品中的所有分子,实际应用中也可适当增加(合成成本增高)或减少(区分效果略弱)组数。

连接产物结构如图 1 所示。其中,a 为接头部分,b 和 f 分别为文库扩增引物,c 为 8bp 随机标签(表 4 中的 8 个 N 表示),d 为 12bp 锚定序列(表 4 中的下划线表示),e 为插入片段(cfDNA)。

### 三、连接产物纯化

步骤二得到的连接产物中加入 110 μl AMPure XP 磁珠(贝克曼 A63880),涡旋混匀,室温放置 10min,磁力架吸附 5min;待溶液澄清后弃上清,然后加入 200 μl 80%(体积百分含量)乙醇水溶液清洗 2 次,弃上清;待乙醇晾干后,加入 30 μl DNase/RNase-Free Water,涡旋混匀,室温放置 10min,磁力架吸附 5min,吸取上清溶液至 PCR 管内,作为 PCR 模板。

### 四、文库扩增及纯化

1、取步骤三得到的 PCR 模板,按照表 5 配置反应体系,按照表 6 进行 PCR 扩增,得到 PCR 扩增产物(4℃保存)。

表 5. 反应体系

| 成分                 | 体积          |
|--------------------|-------------|
| HIFI (KAPA KK8505) | 35 $\mu$ l  |
| MC_F (33 $\mu$ M)  | 2.5 $\mu$ l |
| MC_R (33 $\mu$ M)  | 2.5 $\mu$ l |
| 模板                 | 30 $\mu$ l  |
| 总体积                | 70 $\mu$ l  |

表 5 中, 引物信息如下:

MC\_F (序列 25): GACACGACGCTCTTCCGAT (5' -3');

MC\_R (序列 26): GTGGGCTCGGAGATGTGTATAA (5' -3').

表 6. 反应程序

| 温度   | 时间   | 循环数      |
|------|------|----------|
| 98°C | 45s  |          |
| 98°C | 15s  | 10cycles |
| 60°C | 30s  |          |
| 72°C | 30s  |          |
| 72°C | 5min |          |

5        2、步骤 1 得到的 PCR 扩增产物中加入 90  $\mu$  l AMPure XP 磁珠, 涡旋混匀, 室温放置 10min, 磁力架吸附 5min; 待溶液澄清后弃上清, 然后加入 200  $\mu$  l 80% (体积百分含量) 乙醇水溶液清洗 2 次, 弃上清; 待乙醇晾干后, 加入 100  $\mu$  l DNase/RNase-Free Water, 涡旋混匀, 室温放置 10min, 磁力架吸附 5min, 吸取上清溶液, 得到产物 (-20°C 储存)。产物即为可长期保存、反复使用的 MC 文库。

经检测, MC 文库可以支持 10-20 次的后续检测, 每次检测的结果都可以代表全部原始样本的突变状况, 不会造成灵敏度和特异性的降低。同时, 文库的构建方法不仅适用于 cfDNA 样本, 也适用于基因组 DNA 或 cDNA 样本。

#### 实施例 2、RaceSeq 富集靶区并构建测序文库

15        如图 2 所示, 使用针对我国肝癌高频突变基因 (TP53、CTNNB1、AXIN1、TERT) 的相关区域和针对 HBV 整合热点区域设计的引物, 和固定引物配合, 对 MC 文库进行两轮 PCR 扩增, 扩增产物即为测序文库。

图 2 中, a 为第一轮文库扩增上游引物, b 为第二轮文库扩增上游引物, c 为第一轮文库扩增下游引物库, 用于特异目标序列富集, d 为第二轮文库扩增下游引物库, 用于特异目标序列富集, e 为 index primer, 用于添加 index 序列。

20        1、取 300ng 实施例 1 制备的 MC 文库, 分为两份, 配置表 7 的反应体系 (一份加入 GSP1A mix, 另一份加入 GSP1B mix), 按照表 9 的反应程序进行第一轮 PCR 扩增, 得到第一轮扩增产物 (共得到两份第一轮扩增产物, 一份为 GSP1A mix 的扩增产物, 一份为 GSP1B mix 的扩增产物)。

表 7. 反应体系

| 成分                  | 体积         |
|---------------------|------------|
| Hifi (KAPA KK8505)  | 15 $\mu$ l |
| 上游引物 1355           | 3 $\mu$ l  |
| GSP1A mix/GSP1B mix | 2 $\mu$ l  |
| MC 文库               | 10 $\mu$ l |
| 总体积                 | 30 $\mu$ l |

表 7 中，引物信息如下：

上游引物 1355（序列 27）：TCTTCCCTACACGACGCTCTTCCGAT（5' -3'）。

5 GSP1A mix：将表 8 中属于引物池 GSP1A 中的各条引物用 TE 溶解，浓度为 100  $\mu$ M，然后等体积混合，用 TE 稀释至 0.3  $\mu$ M。引物池 GSP1A 中的引物用于扩增模板的正链。

GSP1B mix：将表 8 中属于引物池 GSP1B 中的各条引物用 TE 溶解，浓度为 100  $\mu$ M，然后等体积混合，用 TE 稀释至 0.3  $\mu$ M。引物池 GSP1B 中的引物用于扩增模板的负链。

10 引物池 GSP1A 和引物池 GSP1B 中，相同编号的引物从正负两个方向检测同一突变位点，同时使用可最大限度的富集原始分子信息。

表 8. 引物信息

| 基因名称   | 引物池   | 引物编号   | 引物序列（5' -3'）                         |
|--------|-------|--------|--------------------------------------|
| AXIN1  | GSP1A | HA1009 | TGTATTAGGGTGCAGCGCTC（序列 28）          |
| AXIN1  | GSP1A | HA1010 | CGCTCGGATCTGGACCTG（序列 29）            |
| AXIN1  | GSP1A | HA1011 | TGGAGCCCTGTGACTCGAA（序列 30）           |
| AXIN1  | GSP1A | HA1012 | GTGACCAGGACATGGATGAGG（序列 31）         |
| AXIN1  | GSP1A | HA1013 | TCCTCCAGTAGACGGTACAGC（序列 32）         |
| AXIN1  | GSP1A | HA1014 | TGCTGCTTGTCCCCACAC（序列 33）            |
| AXIN1  | GSP1A | HA1015 | CCGCTTGGCACCCTTCC（序列 34）             |
| AXIN1  | GSP1A | HA1016 | GGCACGGGAAGCACGTAC（序列 35）            |
| AXIN1  | GSP1A | HA1017 | CCTTGCAGTGGGAAGGTG（序列 36）            |
| CTNNB1 | GSP1A | HA1018 | GACAGAAAAGCGGCTGTTAGTCA（序列 37）       |
| TERT   | GSP1A | HA1019 | CCGACCTCAGCTACAGCAT（序列 38）           |
| TERT   | GSP1A | HA1020 | ACTTGAGCAACCCGAGTCTG（序列 39）          |
| TERT   | GSP1A | HA1021 | CTCCTAGCTCTGCAGTCCGA（序列 40）          |
| TERT   | GSP1A | HA1022 | GCGCCTGGCTCCATTTCC（序列 41）            |
| TERT   | GSP1A | HA1023 | CGCCTGAGAACCTGCAAAGAG（序列 42）         |
| TERT   | GSP1A | HA1024 | GTCCAGGGAGCAATGCGT（序列 43）            |
| TERT   | GSP1A | HA1025 | CGGGTTACCCACAGCCTA（序列 44）            |
| TERT   | GSP1A | HA1026 | GGCTCCCAGTGGATTTCG（序列 45）            |
| TERT   | GSP1A | HA1027 | GTCCTGCCCCTTACCTT（序列 46）             |
| HBV-C  | GSP1A | HA1028 | CCGACTACTGCCTCACCCTATAT（序列 47）       |
| HBV-C  | GSP1A | HA1029 | GGGTTTTTCTTGTGACAAGAATCCT（序列 48）     |
| HBV-C  | GSP1A | HA1030 | CCAACCTCCAATCACTACCAA（序列 49）         |
| HBV-C  | GSP1A | HA1031 | GGCGTTTTATCATATTCCTCTTCATCCT（序列 50）  |
| HBV-C  | GSP1A | HA1032 | CTACTTCCAGGAACATCAACTACCAG（序列 51）    |
| HBV-C  | GSP1A | HA1033 | CTGCACTTGTATTCCCATCCCAT（序列 52）       |
| HBV-C  | GSP1A | HA1034 | TCAGTTTACTAGTGCCATTTGTTTCAGT（序列 53）  |
| HBV-C  | GSP1A | HA1035 | TACAACATCTTGAGTCCCTTTTACCTC（序列 54）   |
| HBV-C  | GSP1A | HA1036 | AGAATTGTGGGTCTTTTGGGCTT（序列 55）       |
| HBV-C  | GSP1A | HA1037 | TGTAAACAATATCTGAACCTTTACCTGTT（序列 56） |
| HBV-C  | GSP1A | HA1038 | GCATGCGTGAACCTTTGTG（序列 57）           |
| HBV-C  | GSP1A | HA1039 | AACTCTGTTGCTCTCTCGGAA（序列 58）         |
| HBV-C  | GSP1A | HA1040 | CTGAATCCCGCGGACGAC（序列 59）            |

|       |       |        |                                         |
|-------|-------|--------|-----------------------------------------|
| HBV-C | GSP1A | HA1041 | CCGTCTGTGCCTTCTCATCTG (序列 60)           |
| HBV-C | GSP1A | HA1042 | GAACGCCACCAGGCTTG (序列 61)               |
| HBV-C | GSP1A | HA1043 | CCTTGAGGCGTACTTCAAAGACTG (序列 62)        |
| HBV-C | GSP1A | HA1044 | GGAGGCTGTAGGCATAAAATTGGT (序列 63)        |
| HBV-C | GSP1A | HA1045 | GTCCTACTGTTCAAGCCTCCAA (序列 64)          |
| HBV-C | GSP1A | HA1046 | GGGCTTCTGTGGAGTTACTCTC (序列 65)          |
| HBV-C | GSP1A | HA1047 | TTGTATCGGGAGGCCTTAGAGT (序列 66)          |
| HBV-C | GSP1A | HA1048 | TTCTGTGTTGGGGTGAGTTGA (序列 67)           |
| HBV-C | GSP1A | HA1049 | CCAGCATCCAGGGAATTAGTAGTCA (序列 68)       |
| HBV-C | GSP1A | HA1050 | TTCCTGTCTTACCTTTGGAAGAGAAAC (序列 69)     |
| HBV-C | GSP1A | HA1051 | CCGGAAGTACTGTTGTTAGACGTA (序列 70)        |
| HBV-C | GSP1A | HA1052 | CGTCGCAGAAGATCTCAATCTCG (序列 71)         |
| HBV-C | GSP1A | HA1053 | AAACTCCCTCCTTCTAACATTCAATTT (序列 72)     |
| HBV-C | GSP1A | HA1054 | TATGCCTGCTAGGTTCTATCCTAACC (序列 73)      |
| HBV-C | GSP1A | HA1055 | GGCATTATTTACATACTCTGTGGAAGG (序列 74)     |
| HBV-C | GSP1A | HA1056 | GTTGGTCTTCCAAACCTCGACA (序列 75)          |
| HBV-C | GSP1A | HA1057 | TTCAACCCCAACAAGGATCACT (序列 76)          |
| HBV-C | GSP1A | HA1058 | TTCCACCAATCGGCAGTCAG (序列 77)            |
| HBV-B | GSP1A | HA1059 | GCCCTGCTCAGAATACTGTCT (序列 78)           |
| HBV-B | GSP1A | HA1060 | ATTCGCAGTCCCAAATCTCC (序列 79)            |
| HBV-B | GSP1A | HA1061 | CATCTTCTCTGCATCCTGCT (序列 80)            |
| HBV-B | GSP1A | HA1062 | TTCCAGGATCATCAACCACCAG (序列 81)          |
| HBV-B | GSP1A | HA1063 | GTCCCTTTATGCCGCTGT (序列 82)              |
| HBV-B | GSP1A | HA1064 | ACCCTTATAAAGAATTTGGAGCTACTGTG (序列 83)   |
| HBV-B | GSP1A | HA1065 | CTCCTGAACATTGCTCACCTCA (序列 84)          |
| TP53  | GSP1A | HA1071 | AGACTGCCTTCCGGGTCA (序列 85)              |
| TP53  | GSP1A | HA1072 | CCTGTGGGAAGCGAAAATTCCA (序列 86)          |
| TP53  | GSP1A | HA1073 | ACCTGGTCTCTGACTGCT (序列 87)              |
| TP53  | GSP1A | HA1074 | AAGCAATGGATGATTTGATGCTGT (序列 88)        |
| TP53  | GSP1A | HA1075 | GACCCAGGTCCAGATGAAGC (序列 89)            |
| TP53  | GSP1A | HA1076 | TCCTGGCCCCTGTCATCT (序列 90)              |
| TP53  | GSP1A | HA1077 | GTGCCCTGACTTTCAACTCTGT (序列 91)          |
| TP53  | GSP1A | HA1078 | CAACTGGCCAAGACCTGC (序列 92)              |
| TP53  | GSP1A | HA1079 | CGCCATGGCCATCTACAAGC (序列 93)            |
| TP53  | GSP1A | HA1080 | GGTCCCCAGGCCTCTGAT (序列 94)              |
| TP53  | GSP1A | HA1081 | GAGTGGAAGGAAATTTGCGTGT (序列 95)          |
| TP53  | GSP1A | HA1082 | GCACTGGCCTCATCTTGGG (序列 96)             |
| TP53  | GSP1A | HA1083 | CCATCCACTACAACATCATGTGTAAC (序列 97)      |
| TP53  | GSP1A | HA1084 | TTTCCTTACTGCCTCTTGCTTCTC (序列 98)        |
| TP53  | GSP1A | HA1085 | GGGACGGAACAGCTTTGAGG (序列 99)            |
| TP53  | GSP1A | HA1086 | CACAGAGGAAGAGAATCTCCGCA (序列 100)        |
| TP53  | GSP1A | HA1087 | TGCCTCAGATTCACTTTTATCACCTT (序列 101)     |
| TP53  | GSP1A | HA1088 | CTCAGGTACTGTGTATATACTTACTTCTCC (序列 102) |
| TP53  | GSP1A | HA1089 | CGTGAGCGCTTCGAGATGT (序列 103)            |
| TP53  | GSP1A | HA1090 | GTGATGTCATCTCTCCTCCCTG (序列 104)         |
| TP53  | GSP1A | HA1091 | TGAAGTCCAAAAAGGGTCAGTCTAC (序列 105)      |

|        |       |        |                                      |
|--------|-------|--------|--------------------------------------|
| AXIN1  | GSP1B | HB1009 | GGGAGCATCTTCGGTGAAAC (序列 106)        |
| AXIN1  | GSP1B | HB1010 | CAGGCTTATCCCATCTTGGTCA (序列 107)      |
| AXIN1  | GSP1B | HB1011 | TTGGTGGCTGGCTTGGTC (序列 108)          |
| AXIN1  | GSP1B | HB1012 | GCTGTACCGTCTACTGGAGGA (序列 109)       |
| AXIN1  | GSP1B | HB1013 | GCTTGTTCTCCAGCTCTCGGA (序列 110)       |
| AXIN1  | GSP1B | HB1014 | GGGAAGTGGTGCCAAGCG (序列 111)          |
| AXIN1  | GSP1B | HB1015 | GCACACGCTGTACGTGCT (序列 112)          |
| AXIN1  | GSP1B | HB1016 | GCCTCCACCTGCTCCTTG (序列 113)          |
| AXIN1  | GSP1B | HB1017 | CCCTCAATGATCCACTGCATGA (序列 114)      |
| CTNNB1 | GSP1B | HB1018 | CTCATACAGGACTTGGGAGGTATC (序列 115)    |
| TERT   | GSP1B | HB1019 | CACAACCGCAGGACAGCT (序列 116)          |
| TERT   | GSP1B | HB1020 | CTCCAAGCCTCGGACTGC (序列 117)          |
| TERT   | GSP1B | HB1021 | GCCTCACACCAGCCACAAC (序列 118)         |
| TERT   | GSP1B | HB1022 | TCCCCACCATGAGCAAACCA (序列 119)        |
| TERT   | GSP1B | HB1023 | GTGCCTCCCTGCAACACT (序列 120)          |
| TERT   | GSP1B | HB1024 | GCACCACGAATGCCGGAC (序列 121)          |
| TERT   | GSP1B | HB1025 | GTGGGGTAACCCGAGGGA (序列 122)          |
| TERT   | GSP1B | HB1026 | GAGGAGGCGGAGCTGGAA (序列 123)          |
| TERT   | GSP1B | HB1027 | AGCGCTGCCTGAAACTCG (序列 124)          |
| TERT   | GSP1B | HB1028 | CGCACGAACGTGGCCAG (序列 125)           |
| HBV-C  | GSP1B | HB1029 | GAGCCACCAGCAGGAAAGT (序列 126)         |
| HBV-C  | GSP1B | HB1030 | CTAGGAATCCTGATGTTGTGCTCT (序列 127)    |
| HBV-C  | GSP1B | HB1031 | CGCGAGTCTAGACTCTGTGGTA (序列 128)      |
| HBV-C  | GSP1B | HB1032 | ATAGCCAGGACAAATTGGAGGACA (序列 129)    |
| HBV-C  | GSP1B | HB1033 | GACAAACGGGCAACATACCTT (序列 130)       |
| HBV-C  | GSP1B | HB1034 | CCGAAGGTTTGTACAGCAACAA (序列 131)      |
| HBV-C  | GSP1B | HB1035 | CTGAGCCAGGAGAAACGGACTGA (序列 132)     |
| HBV-C  | GSP1B | HB1036 | GGGACTCAAGATGTTGTACAGACTTG (序列 133)  |
| HBV-C  | GSP1B | HB1037 | GTTAAGGGAGTAGCCCCAACG (序列 134)       |
| HBV-C  | GSP1B | HB1038 | CAGGCAGTTTTTCGAAAACATTGCTT (序列 135)  |
| HBV-C  | GSP1B | HB1039 | TAAAGCAGGATAGCCACATTGTGTA (序列 136)   |
| HBV-C  | GSP1B | HB1040 | GGCAACAGGGTAAAGGTTTCAGATAT (序列 137)  |
| HBV-C  | GSP1B | HB1041 | CCACAAAGGTTCCACGCAT (序列 138)         |
| HBV-C  | GSP1B | HB1042 | TGGAAAGGAAGTGTACTTCCGAGA (序列 139)    |
| HBV-C  | GSP1B | HB1043 | GTCGTCCGCGGATTGAG (序列 140)           |
| HBV-C  | GSP1B | HB1044 | AAGGCACAGACGGGGAGA (序列 141)          |
| HBV-C  | GSP1B | HB1045 | TCACGGTGGTCTCCATGC (序列 142)          |
| HBV-C  | GSP1B | HB1046 | GGTCGTTGACATTGCTGAGAGT (序列 143)      |
| HBV-C  | GSP1B | HB1047 | AACCTAATCTCCTCCCCAACT (序列 144)       |
| HBV-C  | GSP1B | HB1048 | GCAGAGGTGAAAAAGTTGCATGG (序列 145)     |
| HBV-C  | GSP1B | HB1049 | CCACCCAAGGCACAGCTT (序列 146)          |
| HBV-C  | GSP1B | HB1050 | ACTCCACAGAAGCCCCAA (序列 147)          |
| HBV-C  | GSP1B | HB1051 | GCCTCCCGATACAAAGCAGA (序列 148)        |
| HBV-C  | GSP1B | HB1052 | GATTTCATCAACTACCCCAACACA (序列 149)    |
| HBV-C  | GSP1B | HB1053 | ACATAGCTGACTACTAATTCCCTGGAT (序列 150) |
| HBV-C  | GSP1B | HB1054 | ATCCCACTCCAAAAGACACCAAAT (序列 151)    |

|       |       |        |                                      |
|-------|-------|--------|--------------------------------------|
| HBV-C | GSP1B | HB1055 | GCGAGGGAGTTCTTCTTAGG (序列 152)        |
| HBV-C | GSP1B | HB1056 | CAGTAAAGTTTCCCACCTTGAGT (序列 153)     |
| HBV-C | GSP1B | HB1057 | CCTCTGTAAATGAATGTTAGGAAAGG (序列 154)  |
| HBV-C | GSP1B | HB1058 | GTTTAATGCCTTTATCCAAGGGCAAA (序列 155)  |
| HBV-C | GSP1B | HB1059 | CTCTTATATAGAATCCCAGCCTTCCAC (序列 156) |
| HBV-C | GSP1B | HB1060 | CTTGTCGAGGTTTGGGAAGACCA (序列 157)     |
| HBV-C | GSP1B | HB1061 | GTTTGAGTTGGCTCCGAACG (序列 158)        |
| HBV-C | GSP1B | HB1062 | CTGAGGGCTCCACCCCAA (序列 159)          |
| HBV-C | GSP1B | HB1063 | GTGAAGAGATGGGAGTAGGCTGT (序列 160)     |
| HBV-B | GSP1B | HB1064 | CCCATCTTTTGTGTTTGAGGGTTT (序列 161)    |
| HBV-B | GSP1B | HB1065 | TTAAAGCAGGATATCCACATTGCGTA (序列 162)  |
| HBV-B | GSP1B | HB1066 | TTGCTGAAAGTCCAAGAGTCCT (序列 163)      |
| HBV-B | GSP1B | HB1067 | GGTGAGCAATGTTTCAGGAGATTC (序列 164)    |
| HBV-B | GSP1B | HB1068 | ACTACTAGATCCCTGGACGCTG (序列 165)      |
| HBV-B | GSP1B | HB1069 | GGTGGAGATAAGGGAGTAGGCTG (序列 166)     |
| TP53  | GSP1B | HB1071 | TGCCCTTCCAATGGATCCAC (序列 167)        |
| TP53  | GSP1B | HB1072 | GTCCCCAGCCCAACCTT (序列 168)           |
| TP53  | GSP1B | HB1073 | CTCTGGCATTCTGGGAGCTT (序列 169)        |
| TP53  | GSP1B | HB1074 | TGGTAGGTTTTCTGGGAAGGGA (序列 170)      |
| TP53  | GSP1B | HB1075 | TGTCCCAGAATGCAAGAAGCC (序列 171)       |
| TP53  | GSP1B | HB1076 | GGCATTGAAGTCTCATGGAAGCCA (序列 172)    |
| TP53  | GSP1B | HB1077 | ACCTCCGTCATGTGCTGTGA (序列 173)        |
| TP53  | GSP1B | HB1078 | CTCACCATCGCTATCTGAGCA (序列 174)       |
| TP53  | GSP1B | HB1079 | GCAACCAGCCCTGTCGTC (序列 175)          |
| TP53  | GSP1B | HB1080 | GCACCACCACACTATGTCGAA (序列 176)       |
| TP53  | GSP1B | HB1081 | TTAACCCTCTCCCAGAGAC (序列 177)         |
| TP53  | GSP1B | HB1082 | TTCCAGTGTGATGATGGTGAGGAT (序列 178)    |
| TP53  | GSP1B | HB1083 | CAGCAGGCCAGTGTGCAG (序列 179)          |
| TP53  | GSP1B | HB1084 | CCGGTCTCTCCCAGGACA (序列 180)          |
| TP53  | GSP1B | HB1085 | GTGAGGCTCCCCTTTCTTGC (序列 181)        |
| TP53  | GSP1B | HB1086 | TGGTCTCCTCCACCGCTTC (序列 182)         |
| TP53  | GSP1B | HB1087 | GAACTTTCCACTTGATAAGAGGTCC (序列 183)   |
| TP53  | GSP1B | HB1088 | CTCCCCCTGGCTCCTTC (序列 184)           |
| TP53  | GSP1B | HB1089 | GGGGAGTAGGGCCAGGAAG (序列 185)         |
| TP53  | GSP1B | HB1090 | GCCCTTCTGTCTTGAACATGAGT (序列 186)     |
| TP53  | GSP1B | HB1091 | GTGGGAGGCTGTCACTGG (序列 187)          |

表 9. 反应程序

| 温度   | 时间    | 循环数     |
|------|-------|---------|
| 98°C | 3min  | 9cycles |
| 98°C | 15s   |         |
| 60°C | 90s   |         |
| 72°C | 120s  |         |
| 72°C | 10min |         |

2、步骤 1 得到的两份第一轮扩增产物分别利用 AMPure XP 磁珠按 1: 1.3 的比例进行纯化, 25  $\mu$  l DNase/RNase-Free Water 洗脱, 得到两份第一轮纯化产物。

3、分别以步骤 2 得到的两份第一轮纯化产物为模板，配置表 10 的反应体系（采用 GSP1A mix 扩增产物作为模板时，采用 GSP2A mix 扩增；采用 GSP1Bmix 扩增产物作为模板时，采用 GSP2B mix 扩增），按照表 12 的反应程序进行第二轮 PCR 扩增，得到第二轮扩增产物（4℃ 保存）。

5

表 10. 反应体系

| 成分                       | 体积    |
|--------------------------|-------|
| KapaHifi                 | 15 μl |
| 上游引物 3355                | 2 μl  |
| GSP2Amix/ GSP2Bmix       | 1 μl  |
| Index 引物 (10 μM)         | 2 μl  |
| 模板 (GSP1A mix/ GSP1Bmix) | 10 μl |
| 总体积                      | 30 μl |

表 10 中，引物信息如下：

上游引物 3355 （序列 188）：  
AATGATACGGCGACCACCGAGATCTACACTCTTCCCTACACGACGCTCT（5′-3′）；下划线部分为第一轮的上游引物 1355 相同的部分，3355 及 1355 均为 Illumina 测序平台测序的固定序列（也可更换为可在其他测序平台测序的序列）。

10

GSP2A mix：将表 11 中属于引物池 GSP2A 中的各条引物用 TE 溶解，浓度为 100 μM，然后等体积混合，用 TE 稀释至 0.3 μM。引物池 GSP2A 中的引物用于扩增模板的正链。

GSP2B mix：将表 11 中属于引物池 GSP2B 中的各条引物用 TE 溶解，浓度为 100 μM，然后等体积混合，用 TE 稀释至 0.3 μM。引物池 GSP2B 中的引物用于扩增模板的负链。

15

表 11 中，自 5′ 端第 1 至 20 位为与 Index 引物结合的部分。

GSP2A mix 与 GSP1A mix 中引物编号相同的引物为针对同一突变位点设计，两条引物形成巢式关系。

GSP2B mix 与 GSP2A mix 中引物编号相同的引物为针对同一突变位点设计，两条引物形成巢式关系。

20

Index 引物：CAAGCAGAAGACGGCATACGAGAT（序列 189）  
\*\*\*\*\*GTGACTGGAGTTCCTTGGCACCCGAGAATTCCA（序列 190）；下划线部分为与 GSP2 mix 结合的部分。\*\*\*\*\*为 index 序列位置，index 的长度为 6-8bp，作用是区分样本间的序列，方便多个样本混合测序。除 index 序列外，其余都是来自 Illumina 的 small RNA 测序试剂盒的固定序列。

25

表 11. 引物信息

| 基因名称  | 引物池   | 引物编号   | 引物序列（5′-3′）                                      |
|-------|-------|--------|--------------------------------------------------|
| AXIN1 | GSP2A | HA2009 | CTTGGCACCCGAGAATTCCATTGTTCTTGACGCAGAG（序列 191）    |
| AXIN1 | GSP2A | HA2010 | CTTGGCACCCGAGAATTCCAGACCTGGGGTATGAGCCTGA（序列 192） |
| AXIN1 | GSP2A | HA2011 | CTTGGCACCCGAGAATTCCAAGGCTGAAGCTGGCGAGA（序列 193）   |
| AXIN1 | GSP2A | HA2012 | CTTGGCACCCGAGAATTCCATGAGGACGATGGCAGAGACG（序列 194） |

|        |       |        |                                                             |
|--------|-------|--------|-------------------------------------------------------------|
| AXIN1  | GSP2A | HA2013 | CTTGGCACCCGAGAATTCCAGTACAGCGAAGGCAGAGAGT (序列 195)           |
| AXIN1  | GSP2A | HA2014 | CTTGGCACCCGAGAATTCCACACACAGGAGGAGGAAGGTGA (序列 196)          |
| AXIN1  | GSP2A | HA2015 | CTTGGCACCCGAGAATTCCATGTGTGGACATGGGCTGTG (序列 197)            |
| AXIN1  | GSP2A | HA2016 | CTTGGCACCCGAGAATTCCAACCCAAGTCAGGGGCGAA (序列 198)             |
| AXIN1  | GSP2A | HA2017 | CTTGGCACCCGAGAATTCCAGCGTGCAAAAGAAATGCCAAGAAG (序列 199)       |
| CTNNB1 | GSP2A | HA2018 | CTTGGCACCCGAGAATTCCATAGTCACTGGCAGCAACAGTC (序列 200)          |
| TERT   | GSP2A | HA2019 | CTTGGCACCCGAGAATTCCACTGCAAGGCCTCGGGAGA (序列 201)             |
| TERT   | GSP2A | HA2020 | CTTGGCACCCGAGAATTCCAATTCTCTGGGAAGTCCTCAGCT (序列 202)         |
| TERT   | GSP2A | HA2021 | CTTGGCACCCGAGAATTCCAGCTTGGAGCCAGGTGCCT (序列 203)             |
| TERT   | GSP2A | HA2022 | CTTGGCACCCGAGAATTCCACATTTCCACCCTTTCTCGACGG (序列 204)         |
| TERT   | GSP2A | HA2023 | CTTGGCACCCGAGAATTCCAACGGGCTGTGTCAAGGA (序列 205)              |
| TERT   | GSP2A | HA2024 | CTTGGCACCCGAGAATTCCAATGCGTCTCGGGTTCGT (序列 206)              |
| TERT   | GSP2A | HA2025 | CTTGGCACCCGAGAATTCCAAGCCTAGGCCGATTTCGAC (序列 207)            |
| TERT   | GSP2A | HA2026 | CTTGGCACCCGAGAATTCCAGATTTCGCGGCACAGACG (序列 208)             |
| TERT   | GSP2A | HA2027 | CTTGGCACCCGAGAATTCCATTCCAGTCCGCCTCCTC (序列 209)              |
| HBV-C  | GSP2A | HA2028 | CTTGGCACCCGAGAATTCCACCCATATCGTCAATCTTCTCGAGG (序列 210)       |
| HBV-C  | GSP2A | HA2029 | CTTGGCACCCGAGAATTCCATCACAGTACCACAGAGTCTAGACTC (序列 211)      |
| HBV-C  | GSP2A | HA2030 | CTTGGCACCCGAGAATTCCAAACCTCTTGTCTCCAATTTGTCC (序列 212)        |
| HBV-C  | GSP2A | HA2031 | CTTGGCACCCGAGAATTCCACCTGCTGTATGCCTCATCTTC (序列 213)          |
| HBV-C  | GSP2A | HA2032 | CTTGGCACCCGAGAATTCCACACGGGACCATGCAAGACC (序列 214)            |
| HBV-C  | GSP2A | HA2033 | CTTGGCACCCGAGAATTCCATTGGGCTTTTCGCAAGATTCCTAT (序列 215)       |
| HBV-C  | GSP2A | HA2034 | CTTGGCACCCGAGAATTCCACGTAGGGCTTTCCCCCACT (序列 216)            |
| HBV-C  | GSP2A | HA2035 | CTTGGCACCCGAGAATTCCACCTCTATTACCAATTTTCTTTGTCTTTGGG (序列 217) |
| HBV-C  | GSP2A | HA2036 | CTTGGCACCCGAGAATTCCAACACAATGTGGCTATCCTGCTT (序列 218)         |
| HBV-C  | GSP2A | HA2037 | CTTGGCACCCGAGAATTCCAGGCAACGGTCAGGTCTCT (序列 219)             |
| HBV-C  | GSP2A | HA2038 | CTTGGCACCCGAGAATTCCACTCTGCCGATCCATACTGCGGAA (序列 220)        |
| HBV-C  | GSP2A | HA2039 | CTTGGCACCCGAGAATTCCACACTTCCTTTCCATGGTGCTA (序列 221)          |
| HBV-C  | GSP2A | HA2040 | CTTGGCACCCGAGAATTCCACCGTTTGGGACTCTACCGT (序列 222)            |
| HBV-C  | GSP2A | HA2041 | CTTGGCACCCGAGAATTCCACGTGTGCACTTCGCTTCA (序列 223)             |
| HBV-C  | GSP2A | HA2042 | CTTGGCACCCGAGAATTCCATTGCCCAAGGTCTTACATAAGAGG (序列 224)       |
| HBV-C  | GSP2A | HA2043 | CTTGGCACCCGAGAATTCCAGTTTGTTTAAGGACTGGGAGGAGTT (序列 225)      |
| HBV-C  | GSP2A | HA2044 | CTTGGCACCCGAGAATTCCAGGTCTGTTCCACAGCACCATG (序列 226)          |
| HBV-C  | GSP2A | HA2045 | CTTGGCACCCGAGAATTCCACTGTGCCTTGGGTGGCTT (序列 227)             |
| HBV-C  | GSP2A | HA2046 | CTTGGCACCCGAGAATTCCATTGCCTTCTGATTCTTTCTTCTATT (序列 228)      |
| HBV-C  | GSP2A | HA2047 | CTTGGCACCCGAGAATTCCAGAGTCTCCGGAACATTGTTTACC (序列 229)        |
| HBV-C  | GSP2A | HA2048 | CTTGGCACCCGAGAATTCCAAGTTGATGAATCTGGCCACCT (序列 230)          |
| HBV-C  | GSP2A | HA2049 | CTTGGCACCCGAGAATTCCACAGCTATGTTAATGTTAATATGGGCCTA (序列 231)   |
| HBV-C  | GSP2A | HA2050 | CTTGGCACCCGAGAATTCCATATTTGGTGTCTTTGGAGTGTGGAT (序列 232)      |
| HBV-C  | GSP2A | HA2051 | CTTGGCACCCGAGAATTCCATAGAGGCAGGTCCCCTAGAAG (序列 233)          |
| HBV-C  | GSP2A | HA2052 | CTTGGCACCCGAGAATTCCACAATGTTAGTATCCCTTGACTCACA (序列 234)      |

|       |       |        |                                                          |
|-------|-------|--------|----------------------------------------------------------|
| HBV-C | GSP2A | HA2053 | CTTGGCACCCGAGAATTCCAACAGGAGGACATTATTGATAGATGTCA (序列 235) |
| HBV-C | GSP2A | HA2054 | CTTGGCACCCGAGAATTCCAAACCTTACCAAGTATTTGCCCTT (序列 236)     |
| HBV-C | GSP2A | HA2055 | CTTGGCACCCGAGAATTCATCTGTGGAAGGCTGGGATTCTATAT (序列 237)    |
| HBV-C | GSP2A | HA2056 | CTTGGCACCCGAGAATTCAGGGACAAATCTTTCTGTTCCCA (序列 238)       |
| HBV-C | GSP2A | HA2057 | CTTGGCACCCGAGAATTCAGGCCAGAGGCAAATCAGGT (序列 239)          |
| HBV-C | GSP2A | HA2058 | CTTGGCACCCGAGAATTCACAGTCAGGAAGACAGCCTACTC (序列 240)       |
| HBV-B | GSP2A | HA2059 | CTTGGCACCCGAGAATTCAAATACTGTCTCTGCCATATCGTCA (序列 241)     |
| HBV-B | GSP2A | HA2060 | CTTGGCACCCGAGAATTCAGTGTGTTTCATGAGTGGGAGGA (序列 242)       |
| HBV-B | GSP2A | HA2061 | NA                                                       |
| HBV-B | GSP2A | HA2062 | NA                                                       |
| HBV-B | GSP2A | HA2063 | NA                                                       |
| HBV-B | GSP2A | HA2064 | CTTGGCACCCGAGAATTCATTTGCCCTTCTGACTTCTTTCCGTC (序列 243)    |
| HBV-B | GSP2A | HA2065 | CTTGGCACCCGAGAATTCACACAGCACTCAGGCAAGCTA (序列 244)         |
| TP53  | GSP2A | HA2071 | CTTGGCACCCGAGAATTCAGTCACTGCCATGGAGGAGC (序列 245)          |
| TP53  | GSP2A | HA2072 | CTTGGCACCCGAGAATTCACCATGGGACTGACTTTCTGC (序列 246)         |
| TP53  | GSP2A | HA2073 | CTTGGCACCCGAGAATTCAACTGCTCTTTTACCCATCTACA (序列 247)       |
| TP53  | GSP2A | HA2074 | CTTGGCACCCGAGAATTCATGTCCCCGGACGATATTGAAC (序列 248)        |
| TP53  | GSP2A | HA2075 | CTTGGCACCCGAGAATTCACAGATGAAGCTCCAGAATGCC (序列 249)        |
| TP53  | GSP2A | HA2076 | CTTGGCACCCGAGAATTCATGTCATCTTCTGTCCCTTCCCA (序列 250)       |
| TP53  | GSP2A | HA2077 | CTTGGCACCCGAGAATTCACAACCTGTCTCCTTCTCTCCT (序列 251)        |
| TP53  | GSP2A | HA2078 | CTTGGCACCCGAGAATTCATGTGCAGCTGTGGGTTGAT (序列 252)          |
| TP53  | GSP2A | HA2079 | CTTGGCACCCGAGAATTCACAAGCAGTCACAGCACATGACG (序列 253)       |
| TP53  | GSP2A | HA2080 | CTTGGCACCCGAGAATTCACCTCTGATTCCTCACTGATTGCT (序列 254)      |
| TP53  | GSP2A | HA2081 | CTTGGCACCCGAGAATTCATTGCGTGTGGAGTATTTGGATG (序列 255)       |
| TP53  | GSP2A | HA2082 | CTTGGCACCCGAGAATTCATCTTGGGCCTGTGTATCTCCT (序列 256)        |
| TP53  | GSP2A | HA2083 | CTTGGCACCCGAGAATTCACATGTGTAACAGTTCCCTGCATGG (序列 257)     |
| TP53  | GSP2A | HA2084 | CTTGGCACCCGAGAATTCACCTTGCTTCTCTTTCTATCCTGAGT (序列 258)    |
| TP53  | GSP2A | HA2085 | CTTGGCACCCGAGAATTCACCTTTGAGGTGCGTGTGTTGTGC (序列 259)      |
| TP53  | GSP2A | HA2086 | CTTGGCACCCGAGAATTCAGCAAGAAAGGGAGCCTCA (序列 260)           |
| TP53  | GSP2A | HA2087 | CTTGGCACCCGAGAATTCACCTTTCCTTGCCTCTTTCC (序列 261)          |
| TP53  | GSP2A | HA2088 | CTTGGCACCCGAGAATTCATTCTCCCCCTCCTCTGTTGC (序列 262)         |
| TP53  | GSP2A | HA2089 | CTTGGCACCCGAGAATTCACCTTCGAGATGTTCCGAGAGCT (序列 263)       |
| TP53  | GSP2A | HA2090 | CTTGGCACCCGAGAATTCACCTCCCTGCTTCTGTCTCCTA (序列 264)        |
| TP53  | GSP2A | HA2091 | CTTGGCACCCGAGAATTCATCAGTCTACCTCCCGCCATA (序列 265)         |
| AXIN1 | GSP2B | HB2009 | CTTGGCACCCGAGAATTCAGAACTTGCTCCGAGGTCCA (序列 266)          |
| AXIN1 | GSP2B | HB2010 | CTTGGCACCCGAGAATTCACATCCAGCAGGGAATGCAGT (序列 267)         |
| AXIN1 | GSP2B | HB2011 | CTTGGCACCCGAGAATTCAGACACGATGCCATTGTTATCAAGA (序列 268)     |
| AXIN1 | GSP2B | HB2012 | CTTGGCACCCGAGAATTCACCTGTCTCCAGGAGCAGCTTC (序列 269)        |
| AXIN1 | GSP2B | HB2013 | CTTGGCACCCGAGAATTCACGGAGGTGAGTACAGAAAGTGG (序列 270)       |
| AXIN1 | GSP2B | HB2014 | CTTGGCACCCGAGAATTCAGGAGGCAGCTTGTGACACG (序列 271)          |

|        |       |        |                                                           |
|--------|-------|--------|-----------------------------------------------------------|
| AXIN1  | GSP2B | HB2015 | CTTGGCACCCGAGAATTCCACTCGTCCAGGATGCTCTCAG (序列 272)         |
| AXIN1  | GSP2B | HB2016 | CTTGGCACCCGAGAATTCCAGTGGTGGACGTGGTGGTG (序列 273)           |
| AXIN1  | GSP2B | HB2017 | CTTGGCACCCGAGAATTCCATGATTTCTGGTTCTTCTCCGCAT (序列 274)      |
| CTNNB1 | GSP2B | HB2018 | CTTGGCACCCGAGAATTCCAGAGGTATCCACATCTCTTCTCTCA (序列 275)     |
| TERT   | GSP2B | HB2019 | CTTGGCACCCGAGAATTCCAAGGACTTCCAGGAATCCAG (序列 276)          |
| TERT   | GSP2B | HB2020 | CTTGGCACCCGAGAATTCCAAGCTAGGAGGCCGACTT (序列 277)            |
| TERT   | GSP2B | HB2021 | CTTGGCACCCGAGAATTCCAACAACGGCCTTGACCCTG (序列 278)           |
| TERT   | GSP2B | HB2022 | CTTGGCACCCGAGAATTCCACCACCCCAATCTGTTAATCACC (序列 279)       |
| TERT   | GSP2B | HB2023 | CTTGGCACCCGAGAATTCCAAACACTTCCCCGCGACTTGG (序列 280)         |
| TERT   | GSP2B | HB2024 | CTTGGCACCCGAGAATTCCACGTGAAGGGGAGGACGGA (序列 281)           |
| TERT   | GSP2B | HB2025 | CTTGGCACCCGAGAATTCCAGGGGCCATGATGTGGAGG (序列 282)           |
| TERT   | GSP2B | HB2026 | CTTGGCACCCGAGAATTCCAAAGGTGAAGGGGACAGGAC (序列 283)          |
| TERT   | GSP2B | HB2027 | CTTGGCACCCGAGAATTCCAGCGAAAGGAAGGGGAGG (序列 284)            |
| TERT   | GSP2B | HB2028 | CTTGGCACCCGAGAATTCCAGCAGCACCTCGCGGTAG (序列 285)            |
| HBV-C  | GSP2B | HB2029 | CTTGGCACCCGAGAATTCCAGGAAAGTATAGGCCCTCACTC (序列 286)        |
| HBV-C  | GSP2B | HB2030 | CTTGGCACCCGAGAATTCCACTCTCCATGTTGCGGGCA (序列 287)           |
| HBV-C  | GSP2B | HB2031 | CTTGGCACCCGAGAATTCCAGAGGATTCTGTCAACAAGAAAAACCC (序列 288)   |
| HBV-C  | GSP2B | HB2032 | CTTGGCACCCGAGAATTCCAACAAGAGGTTGGTGAGTGATTGG (序列 289)      |
| HBV-C  | GSP2B | HB2033 | CTTGGCACCCGAGAATTCCAGTCCAGAAGAACCAACAAGAAGATGA (序列 290)   |
| HBV-C  | GSP2B | HB2034 | CTTGGCACCCGAGAATTCCACATAGAGGTTCTTGAGCAGGAATC (序列 291)     |
| HBV-C  | GSP2B | HB2035 | CTTGGCACCCGAGAATTCCCACTCCCATAGGAATCTTGCGAA (序列 292)       |
| HBV-C  | GSP2B | HB2036 | CTTGGCACCCGAGAATTCCACCCCAATACCACATCATCCATA (序列 293)       |
| HBV-C  | GSP2B | HB2037 | CTTGGCACCCGAGAATTCCAAGGTTCAAATGTATACCCAAAGACAA (序列 294)   |
| HBV-C  | GSP2B | HB2038 | CTTGGCACCCGAGAATTCCAAGTTTTAGTACAATATGTTCTTGCGGTA (序列 295) |
| HBV-C  | GSP2B | HB2039 | CTTGGCACCCGAGAATTCCACATTGTGTAAGGGGACAGCA (序列 296)         |
| HBV-C  | GSP2B | HB2040 | CTTGGCACCCGAGAATTCCATGTTTACACAGAAAGGCCTTGTAAGT (序列 297)   |
| HBV-C  | GSP2B | HB2041 | CTTGGCACCCGAGAATTCCACATGCGGCGATGGCCAATA (序列 298)          |
| HBV-C  | GSP2B | HB2042 | CTTGGCACCCGAGAATTCCATTCCGAGAGAGGACAACAGAGTTGT (序列 299)    |
| HBV-C  | GSP2B | HB2043 | CTTGGCACCCGAGAATTCCAGACGGGACGTAAACAAAGGAC (序列 300)        |
| HBV-C  | GSP2B | HB2044 | CTTGGCACCCGAGAATTCCAGGAGACCGCGTAAAGAGAGG (序列 301)         |
| HBV-C  | GSP2B | HB2045 | CTTGGCACCCGAGAATTCCAGTGCAGAGGTGAAGCGAAGT (序列 302)         |
| HBV-C  | GSP2B | HB2046 | CTTGGCACCCGAGAATTCCATCCAAGAGTCCTTATGTAAGACC (序列 303)      |
| HBV-C  | GSP2B | HB2047 | CTTGGCACCCGAGAATTCCACAACCTCTCCAGTCCTTAAACA (序列 304)       |
| HBV-C  | GSP2B | HB2048 | CTTGGCACCCGAGAATTCCAGGTGCTGGTGAACAGACCAA (序列 305)         |
| HBV-C  | GSP2B | HB2049 | CTTGGCACCCGAGAATTCCACTTGAGGCTTGAACAGTAGGA (序列 306)        |
| HBV-C  | GSP2B | HB2050 | CTTGGCACCCGAGAATTCCAAATCTTTATACGGGTCAATGTCCA (序列 307)     |
| HBV-C  | GSP2B | HB2051 | CTTGGCACCCGAGAATTCCACAGAGGCGGTGTCGAGGA (序列 308)           |
| HBV-C  | GSP2B | HB2052 | CTTGGCACCCGAGAATTCCAACACAGAACAGCTTGCCTGA (序列 309)         |
| HBV-C  | GSP2B | HB2053 | CTTGGCACCCGAGAATTCCACTGGGTCTTCCAAATTAATCCCA (序列 310)      |
| HBV-C  | GSP2B | HB2054 | CTTGGCACCCGAGAATTCCAGTTTCTCTTCCAAAGGTAAGACAGGA (序列 311)   |

|       |       |        |                                                            |
|-------|-------|--------|------------------------------------------------------------|
| HBV-C | GSP2B | HB2055 | CTTGGCACCCGAGAATTCCAACCTGCCTCTACGTCTAACAACA (序列 312)       |
| HBV-C | GSP2B | HB2056 | CTTGGCACCCGAGAATTCCATTGTGAGTCCAAGGATACTAACATTG (序列 313)    |
| HBV-C | GSP2B | HB2057 | CTTGGCACCCGAGAATTCCAGGGAGTTTGCCACTCAGGATTAAA (序列 314)      |
| HBV-C | GSP2B | HB2058 | CTTGGCACCCGAGAATTCCAGGGCAAATACTTGGTAAGGTTAGGATA (序列 315)   |
| HBV-C | GSP2B | HB2059 | CTTGGCACCCGAGAATTCCACCTTCCACAGAGTATGTAAATAATGCCTA (序列 316) |
| HBV-C | GSP2B | HB2060 | CTTGGCACCCGAGAATTCCACTCCCATGCTGTAGCTCTTGTT (序列 317)        |
| HBV-C | GSP2B | HB2061 | CTTGGCACCCGAGAATTCCAGCTGGGTCCAAGTGGTGATC (序列 318)          |
| HBV-C | GSP2B | HB2062 | CTTGGCACCCGAGAATTCCACCCAAAAGACCACCGTGTG (序列 319)           |
| HBV-C | GSP2B | HB2063 | CTTGGCACCCGAGAATTCCATCTTCTGACTGCCGATTGGT (序列 320)          |
| HBV-B | GSP2B | HB2064 | NA                                                         |
| HBV-B | GSP2B | HB2065 | NA                                                         |
| HBV-B | GSP2B | HB2066 | CTTGGCACCCGAGAATTCCACAAGACCTTGGGCAGGTCC (序列 321)           |
| HBV-B | GSP2B | HB2067 | CTTGGCACCCGAGAATTCCAATTCTAAGGCTTCCCAGATACAGA (序列 322)      |
| HBV-B | GSP2B | HB2068 | CTTGGCACCCGAGAATTCCAACGCTGGATCTTCTAAATTATTACCC (序列 323)    |
| HBV-B | GSP2B | HB2069 | NA                                                         |
| TP53  | GSP2B | HB2071 | CTTGGCACCCGAGAATTCCAGATCCACTCACAGTTTCCATAGG (序列 324)       |
| TP53  | GSP2B | HB2072 | CTTGGCACCCGAGAATTCCACAGCCCAACCCTGTCTTA (序列 325)            |
| TP53  | GSP2B | HB2073 | CTTGGCACCCGAGAATTCCATGGGAGCTTCATCTGGACCTG (序列 326)         |
| TP53  | GSP2B | HB2074 | CTTGGCACCCGAGAATTCCAGAAGGGACAGAAGATGACAGG (序列 327)         |
| TP53  | GSP2B | HB2075 | CTTGGCACCCGAGAATTCCACAAGAAGCCAGACGGAAACC (序列 328)          |
| TP53  | GSP2B | HB2076 | CTTGGCACCCGAGAATTCCACCCCTCAGGGCAACTGAC (序列 329)            |
| TP53  | GSP2B | HB2077 | CTTGGCACCCGAGAATTCCAGTGCTGTGACTGCTGTAGATGGC (序列 330)       |
| TP53  | GSP2B | HB2078 | CTTGGCACCCGAGAATTCCAATCTGAGCAGCGCTCATGGTG (序列 331)         |
| TP53  | GSP2B | HB2079 | CTTGGCACCCGAGAATTCCACCCTGCTCTCTCCAGC (序列 332)              |
| TP53  | GSP2B | HB2080 | CTTGGCACCCGAGAATTCCACTATGTGAAAAGTGTCTGTCTATCC (序列 333)     |
| TP53  | GSP2B | HB2081 | CTTGGCACCCGAGAATTCCAGAGACCCAGTTGCAAAACCAG (序列 334)         |
| TP53  | GSP2B | HB2082 | CTTGGCACCCGAGAATTCCATGGGCCTCCGGTTCATGC (序列 335)            |
| TP53  | GSP2B | HB2083 | CTTGGCACCCGAGAATTCCAGTGCAGGGTGGCAAGTGG (序列 336)            |
| TP53  | GSP2B | HB2084 | CTTGGCACCCGAGAATTCCAGACAGGCACAAACACGCAC (序列 337)           |
| TP53  | GSP2B | HB2085 | CTTGGCACCCGAGAATTCCATTCTTGCGGAGATTCTCTTCTCT (序列 338)       |
| TP53  | GSP2B | HB2086 | CTTGGCACCCGAGAATTCCACGCTTCTGTCTGCTGCT (序列 339)             |
| TP53  | GSP2B | HB2087 | CTTGGCACCCGAGAATTCCAACCTTGATAAGAGGTCCCAAGACTTAG (序列 340)   |
| TP53  | GSP2B | HB2088 | CTTGGCACCCGAGAATTCCAAGCCTGGGCATCCTTGAG (序列 341)            |
| TP53  | GSP2B | HB2089 | CTTGGCACCCGAGAATTCCACAGGAAGGGGTGAGGTC (序列 342)             |
| TP53  | GSP2B | HB2090 | CTTGGCACCCGAGAATTCCACATGAGTTTTTATGGCGGGAGGT (序列 343)       |
| TP53  | GSP2B | HB2091 | CTTGGCACCCGAGAATTCCACAGTGGGGAACAAGAAGTGGA (序列 344)         |

其中, NA 表示没有引物。

表 12. 反应程序

| 温度   | 时间   | 循环数     |
|------|------|---------|
| 98°C | 3min |         |
| 98°C | 15s  | 9cycles |

|     |        |         |
|-----|--------|---------|
| 60℃ | 90s    | 8cycles |
| 72℃ | 90s    |         |
| 98℃ | 15s    |         |
| 60℃ | 30s    |         |
| 72℃ | 30s    |         |
| 72℃ | 10 min |         |

4、步骤 3 得到的采用 GSP2A mix 进行第二轮扩增的产物和采用 GSP1Bmix 进行第二轮扩增的产物等体积混合，利用 AMPure XP 磁珠按 1: 1.3 的比例进行纯化，50 μl DNase/RNase-Free Water 洗脱，得到第二轮纯化产物，即为可以在 Illumina HiSeq X 平台测序的测序文库，每个文库测序数据量为 2G，平均测序深度大于 60000 ×。

MC 文库上的 DNA 随机标签与 cfDNA 序列一起被添加到测序文库的 Read1 序列下游。在测序中，依次获得 DNA 随机标签序列、锚定序列、cfDNA 序列（图 1 中的 c、d、e 序列）。在数据分析时，将随机标签序列相同、DNA 插入片段长度相同、DNA 插入片段两端断点均相同的 DNA 分子测序数据回溯到一个分子簇，如果簇内分子数>5 个并且簇内分子突变一致率>80%并且簇个数≥5，则该突变为来自原始 DNA 样本的真实突变。

经 30ng 肝癌患者 cfDNA 的测序实验表明，此方法总计耗时仅需约 6 小时（手动操作约 1.5h），产生的 RaceSeq 文库的上靶率达到 80%，在 2Gb 数据量的情况下，测序深度达到 60000×，分子簇 5000 个，平均每个簇的测序分子数达到 12 个，具体可见表 13。

表 13. 测序数据展示

|               |            |            |            |            |
|---------------|------------|------------|------------|------------|
| 目标区域：20031 bp |            |            |            |            |
| 样品编号          | RG773      | RG774      | RG776      | RG777      |
| 样品起始量（ng）     | 30         | 30         | 30         | 30         |
| MC 文库产量（ng）   | 6800       | 7300       | 7200       | 6500       |
| 测序量（bp）       | 2002633889 | 1733018642 | 1725428157 | 1790632086 |
| 基因组比对率        | 0.9888     | 0.9896     | 0.9907     | 0.9896     |
| 上靶率           | 0.8015     | 0.8057     | 0.8295     | 0.7877     |
| 测序深度          | 80130.55   | 69705.42   | 71452.26   | 70411.13   |

#### 实施例 3、MC 文库的捕获及测序

如图 3 所示，可使用 Agilent sureselect XT 靶向捕获试剂盒（Agilent5190-8646）对实施例 1 的 MC 文库进行捕获（参照试剂盒说明书进行，也兼容其它品牌捕获试剂）将最后一步 PCR 扩增的引物更换为如下引物：

上游引物（5' -3'）：  
AATGATACGGCGACCAACGAGATCTACACTCTTCCCTACACGACGCTCTTCCGATCT（序列 345）  
（图 3 中的“a”），下划线部分为与引物 MC\_F 相同，作用为扩增文库，其余部分为 Illumina 测序平台测序所需的固定序列。

下游引物（5' -3'）：CAAGCAGAAGACGGCATACGAGAT（序列 346）  
\*\*\*\*\*GTCTCGTGGGCTCGGAGATGTGTATAA（序列 347）（图 3 中的“b”），下划线部分为与引物 MC\_R 相同，作用为扩增文库。\*\*\*\*\*为 index 序列位置，index

长度为 6-8bp, 作用是区分样本间的序列, 方便多个样本混合测序。其余部分为 Illumina 测序平台测序所需的固定序列。

捕获后的文库与 MC 文库具有相同的 DNA 随机标签序列、锚定序列和 cfDNA 序列, 依次位于 Read1 下游。将 DNA 插入片段长度相同、DNA 插入片段两端断点、两端的锚定序列均相同的初始 DNA 单链测序数据回溯到一个分子簇。同时, 将插入片段长度一致并且除突变点外的序列一致、分子簇两端的锚定序列相同但位置相反的同一起始 DNA 双链的分子簇标记为一对 duplex 分子簇。对某一突变来说, 若至少有一对 duplex 分子簇支持, 即可判断为真, 若无 duplex 分子簇支持, 那么至少有 4 个分子簇支持, 可判断为真。在一对 duplex 分子簇共同支持的突变可靠性更高, 可减少 90% 的假阳性突变。

#### 实施例 4、方法对比

收集了 5 例肝癌者 cf DNA 标本, 先按照实施例 1 中的方法进行 MC 文库构建, 然后同时按照实施例 2 中的方法进行 RaceSeq 靶区富集和实施例 3 中的常规的 Agilent sureselect XT 靶区富集、测序, 变异检测结果展示如表 13 和表 14 所示。

表 13. 两种方法 snv 检测对比

| 检测结果 | 样本编号  | 突变频率   | 基因    | 碱基位置    | 参考型 | 突变型 |
|------|-------|--------|-------|---------|-----|-----|
| 一致   | RG773 | 0.0013 | TERT  | 1295073 | C   | T   |
| 一致   | RG773 | 0.0009 | AXIN1 | 396583  | T   | A   |
| 一致   | RG773 | 0.0025 | TP53  | 7577539 | G   | A   |
| 一致   | RG775 | 0.0800 | TP53  | 7577534 | C   | A   |
| 一致   | RG776 | 0.0010 | AXIN1 | 347927  | C   | T   |
| 一致   | RG776 | 0.0024 | TP53  | 7577534 | C   | A   |

表 14. 两种方法 HBV 融合检测对比

| 检测方法     | 样本编号  | 基因 (人)                                     | 染色体号_右侧断点 | 碱基位置_右侧断点 | 病毒亚型  | 碱基位置_左侧断点 | 融合频率 | 结果 |
|----------|-------|--------------------------------------------|-----------|-----------|-------|-----------|------|----|
| 捕获法      | RG774 | PKDCC(dist=5233), LOC102723824(dist=78674) | 2         | 42290901  | HBV_C | 851       | 0.08 | 一致 |
| Race-seq | RG774 | PKDCC(dist=5233), LOC102723824(dist=78674) | 2         | 42290906  | HBV_C | 851       | 0.11 |    |
| 捕获法      | RG774 | SNTG1(dist=351419), PXDNL(dist=174040)     | 8         | 52058097  | HBV_C | 1811      | 0.15 | 一致 |
| Race-seq | RG774 | SNTG1(dist=351419), PXDNL(dist=174040)     | 8         | 52058096  | HBV_C | 1811      | 0.06 |    |

可以看出, Agilent sureselect XT 靶区富集方法和 RaceSeq 法对于单碱基突变和 HBV 插入的检测结果基本一致。

## 20 工业应用

本发明的发明人经过大量实验, 获得了用于检测 DNA 样本中肝癌突变的引物组合, 该引物组合由引物组 I、引物组 II、引物组 III 和引物组 IV 组成; 引物组 I 由序列

- 表的序列28至序列105所示的单链DNA组成；引物组 II 由序列表的序列106至序列187所示的单链DNA组成；引物组III由序列表的序列191至序列265所示的单链DNA组成；引物组IV由序列表的序列266至序列344所示的单链DNA组成。实验证明，在不需要捕获的情况下，上述引物组合可同时检测肝癌ctDNA中的点突变、插入缺失突变、
- 5 HBV整合等多种突变形式。由于不需昂贵的捕获探针和杂交试剂，成本极大降低，而且操作流程简单，检测时间仅为8h。本发明对于肝癌的早期筛查、病情追踪、疗效评估、预后预测等具有重要临床意义。

### 权利要求

1、一种测序文库构建方法，依次包括如下步骤：

(1) 将 DNA 样本依次进行末端修复和 3' 端加 A 处理；

5 得到文库；

所述接头混合物由 n 个接头组成；

每个接头均由一条上游引物甲和一条下游引物甲形成部分双链结构得到；  
上游引物甲中具有测序接头甲、随机标签、锚定序列甲和位于 3' 末端的碱基 T；  
下游引物甲中具有锚定序列乙和测序接头乙；所述部分双链结构由上游引物中的  
10 的锚定序列甲和下游引物中的锚定序列乙反向互补形成；

所述测序接头甲和测序接头乙为根据根据不同测序平台选择对应的测序接头；

所述随机标签为 8-14bp 的随机碱基；

所述锚定序列甲长度为 14-20bp，连续重复碱基 ≤ 3 个；

15 n 个接头采用 n 个不同的锚定序列甲，并且相同位置的碱基平衡，错配碱基数 > 3；

n 为 ≥ 8 的任意自然数。

2、如权利要求 1 所述的方法，其特征在于：n 为 12 时，

所述锚定序列甲的核苷酸序列分别为序列表的序列 1 自 5' 端第 30-41 位、  
20 序列表的序列 3 自 5' 端第 30-41 位、序列表的序列 5 自 5' 端第 30-41 位、序列表的序列 7 自 5' 端第 30-41 位、序列表的序列 9 自 5' 端第 30-41 位、序列表的序列 11 自 5' 端第 30-41 位、序列表的序列 13 自 5' 端第 30-41 位、序列表的序列 15 自 5' 端第 30-41 位、序列表的序列 17 自 5' 端第 30-41 位、序列表的序列 19 自 5' 端第 30-41 位、序列表的序列 21 自 5' 端第 30-41 位、序列  
25 表的序列 23 自 5' 端第 30-41 位所示；

接头 1 由序列表的序列 1 所示的单链 DNA 分子和序列 2 所示的单链 DNA 分子形成部分双链结构得到；接头 2 由序列表的序列 3 所示的单链 DNA 分子和序列 4 所示的单链 DNA 分子形成部分双链结构得到；接头 3 由序列表的序列 5 所示的单链 DNA 分子和序列 6 所示的单链 DNA 分子形成部分双链结构得到；接头 4  
30 由序列表的序列 7 所示的单链 DNA 分子和序列 8 所示的单链 DNA 分子形成部分双链结构得到；接头 5 由序列表的序列 9 所示的单链 DNA 分子和序列 10 所示的单链 DNA 分子形成部分双链结构得到；接头 6 由序列表的序列 11 所示的单链 DNA 分子和序列 12 所示的单链 DNA 分子形成部分双链结构得到；接头 7 由序列表的序列 13 所示的单链 DNA 分子和序列 14 所示的单链 DNA 分子形成部分双链结构  
35 得到；接头 8 由序列表的序列 15 所示的单链 DNA 分子和序列 16 所示的单链 DNA 分子形成部分双链结构得到；接头 9 由序列表的序列 17 所示的单链 DNA 分子和序列 18 所示的单链 DNA 分子形成部分双链结构得到；接头 10 由序列表的序列

19 所示的单链 DNA 分子和序列 20 所示的单链 DNA 分子形成部分双链结构得到；  
接头 11 由序列表的序列 21 所示的单链 DNA 分子和序列 22 所示的单链 DNA 分子  
形成部分双链结构得到；接头 12 由序列表的序列 23 所示的单链 DNA 分子和序  
列 24 所示的单链 DNA 分子形成部分双链结构得到。

5        3、如权利要求 1 或 2 所述的方法，其特征在于：所述方法还包括对步骤（2）  
得到的文库进行扩增的步骤。

4、如权利要求 3 所述的方法，其特征在于：所述扩增采用的引物对由序列  
表的序列 25 和序列 26 所示的两条单链 DNA 分子组成。

5、权利要求 1 至 4 任一所述的方法构建得到的 DNA 文库。

10       6、一种用于构建测序文库的试剂盒，包括权利要求 1 或 2 中所述的接头混  
合物。

7、一种用于检测 DNA 样本中肝癌突变的试剂盒，包括权利要求 1 或 2 中所  
述的接头混合物和引物组合；所述引物组合包括引物组 I、引物组 II、引物组 III  
和引物组 IV；

15       所述引物组 I 和所述引物组 II 中的各个引物是根据与肝癌相关的区域设计  
的特异性引物，作用是定位于基因组特定位置，实现目标区域的 PCR 富集；

所述引物组 III 和所述引物组 IV 中的各个引物的核苷酸序列均由“接头序列+  
特异序列”组成，特异序列用于目标区域进行进一步富集，接头序列用于通过  
PCR 形成完整的可以测序的文库分子；

20       所述引物组 III 和所述引物组 I 可为“巢式”关系；所述引物组 IV 和所述引  
物组 II 可为“巢式”关系。

8、如权利要求 7 所述的试剂盒，其特征在于：

所述引物组 I 由序列表的序列 28 至序列 105 所示的单链 DNA 组成；

所述引物组 II 由序列表的序列 106 至序列 187 所示的单链 DNA 组成；

25       所述引物组 III 由序列表的序列 191 至序列 265 所示的单链 DNA 组成；

所述引物组 IV 由序列表的序列 266 至序列 344 所示的单链 DNA 组成。

9、权利要求 7 或 8 中所述的引物组合。

10、权利要求 7 或 8 中所述的引物组合在制备用于检测 DNA 样本中肝癌突  
变的试剂盒中的应用。

30       11、一种检测 DNA 样本中目标突变的方法，包括如下步骤：

（1）按照权利要求 1 至 4 任一所述的方法构建文库；

（2）对步骤（1）得到的文库进行两轮巢式 PCR 扩增，对产物进行测序，  
根据测序结果分析 DNA 样本中目标突变发生情况；

所述步骤（2）中，采用引物组合甲进行第一轮 PCR 扩增；

35       引物组合甲由上游引物甲和下游引物组合甲组成；

所述上游引物甲为文库扩增引物，用于步骤（1）的文库扩增；

所述下游引物组合甲为根据 n 个目标靶点设计的 n 条引物的组合；

以第一轮 PCR 的产物为模板，采用引物组合乙进行第二轮 PCR 扩增；

引物组合乙由上游引物乙、下游引物组合乙和 index 引物组成；

所述上游引物乙为文库扩增引物，用于第一轮 PCR 的产物的扩增；

5 所述下游引物组合乙中的引物与下游引物组合甲中检测相同目标靶点的引物形成巢式关系，各条引物中具有与 index 引物结合的区段；

所述 index 引物中包含与下游引物组合乙中各条引物结合的区段和 index 序列。

12、如权利要求 11 所述的方法，其特征在于：

所述上游引物甲的核苷酸序列如序列表的序列 27 所示；

10 所述上游引物乙的核苷酸序列如序列表的序列 188 所示。

所述 index 引物自 5' 端包括区段 A、index 序列和区段 B；所述区段 A 的核苷酸序列如序列表的序列 189 所示，所述区段 B 的核苷酸序列如序列表的序列 190 所示。

13、如权利要求 11 或 12 所述的方法，其特征在于：

15 当所述目标突变为肝癌突变时，所述引物组合甲由引物组 I 和引物组 II 组成；所述引物组合乙由引物组 III 和引物组 IV 组成；

使用引物组 I 和引物组 II 分别对模板进行第一轮 PCR 扩增，使用引物组 I 进行扩增的产物作为第二轮扩增的模板采用引物组 III 进行扩增，使用引物组 II 进行扩增的产物作为第二轮扩增的模板采用引物组 IV 进行扩增，然后将扩增产物  
20 等体积混合。

14、如权利要求 11 至 13 任一所述的方法，其特征在于：所述测序结果的分析方法为：将随机标签序列相同、DNA 插入片段长度相同、DNA 插入片段两端断点均相同的 DNA 分子测序数据回溯到一个分子簇，如果簇内分子数>5 个并且簇内分子突变一致率>80%并且簇个数≥5，则该突变为来自原始 DNA 样本的真实  
25 突变。

15、一种检测 DNA 样本中多种目标突变的方法，包括如下步骤：

(1) 按照权利要求 1 至 4 任一所述的方法构建文库；

(2) 对步骤 (1) 的文库进行靶区域富集并进行测序，根据测序结果分析 DNA 样本中目标突变发生情况。

30 16、如权利要求 15 所述的方法，其特征在于：所述测序结果的分析方法为：将 DNA 插入片段长度相同、DNA 插入片段两端断点、两端的锚定序列均相同的初始 DNA 单链测序数据回溯到一个分子簇；将插入片段长度一致并且除突变点外的序列一致、分子簇两端的锚定序列相同但位置相反的同一起始 DNA 双链的分子簇标记为一对 duplex 分子簇；对某一突变来说，若至少有一对 duplex 分子簇支持，即可判断为真，若无 duplex 分子簇支持，那么至少有 4 个分子簇支持，可判断为真。  
35

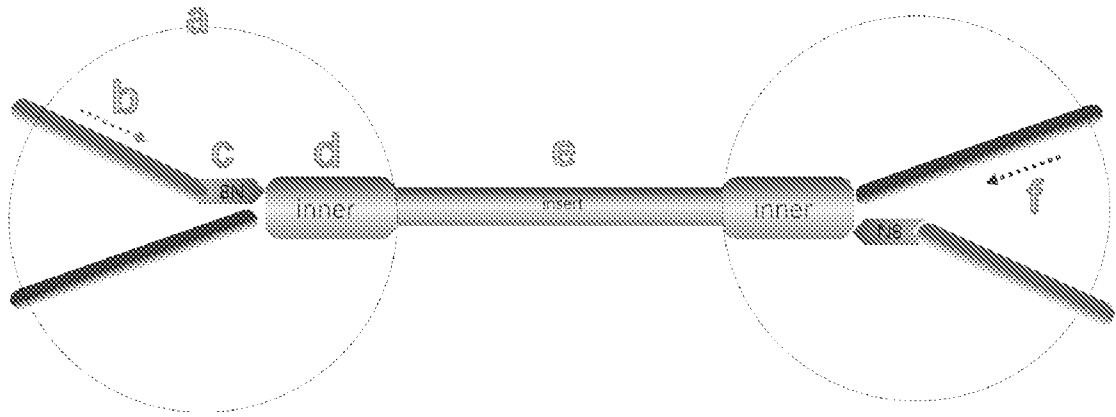


图 1

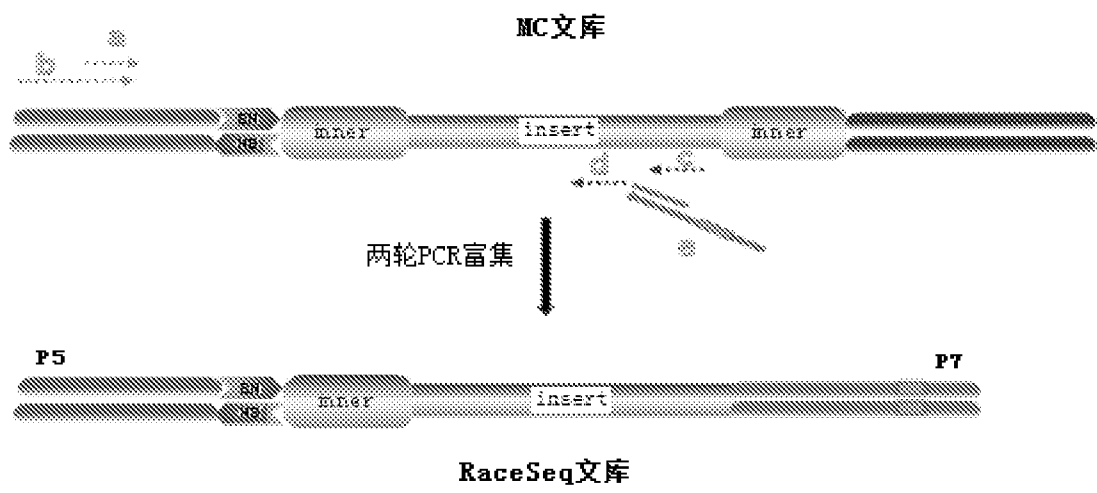


图 2

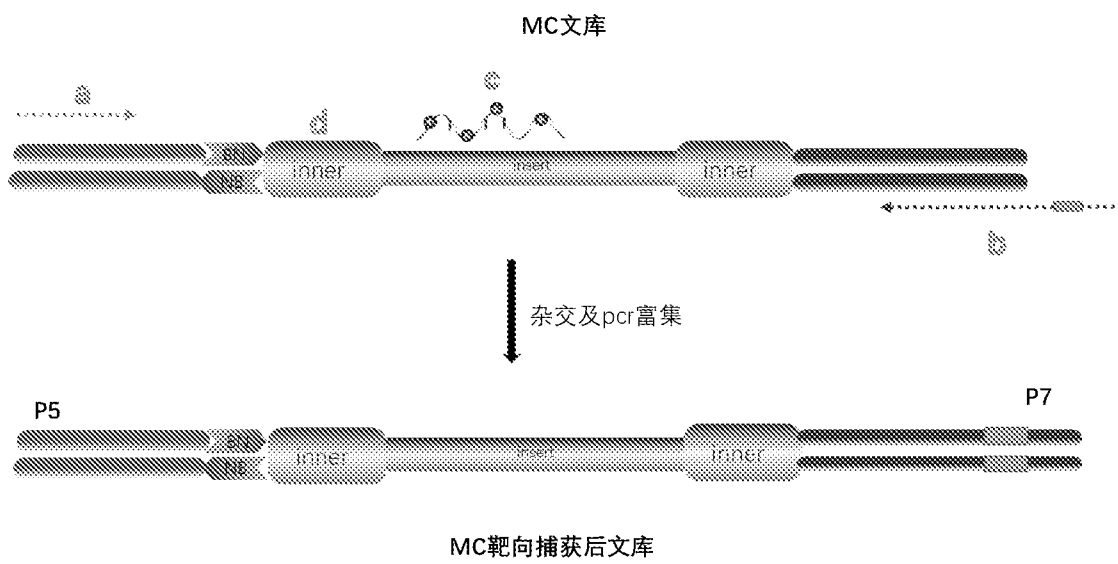


图 3

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2019/082233

**A. CLASSIFICATION OF SUBJECT MATTER**

C12Q 1/6806(2018.01)i; C12Q 1/6827(2018.01)i; C12Q 1/6886(2018.01)i; C12Q 1/6855(2018.01)i; C12Q 1/6876(2018.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

**B. FIELDS SEARCHED**

Minimum documentation searched (classification system followed by classification symbols)

C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

DWPI, SIPOABS, CPRSABS, CPEA, CNABS, JPABS, CNMED, KRABS, CNTXT, USTXT, WOTXT, EPTXT, JPTXT, KRTXT, CNKI, ISI WEB OF SCIENCE, 中国专利生物序列检索系统, NATIONAL BIO-SEQUENCE DATABASE OF CHINESE PATENT, REGISTRY: 文库, 测序, 接头, 引物, 标签, 锚定, 肝癌, 突变, 循环肿瘤DNA, 无细胞DNA, library, sequencing, adaptor, primer, barcode, index, label, tag, tab, anchor, liver cancer, hepatoma, hepatic carcinoma, mutate, mutation, mutant, variation, ctDNA, circulating tumor DNA, cfDNA, cell-free DNA, SEQ ID NOs: 1-24

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

| Category* | Citation of document, with indication, where appropriate, of the relevant passages                                                                                                       | Relevant to claim No.  |
|-----------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------|
| X         | CN 106192018 A (BGI-SHENZHEN) 07 December 2016 (2016-12-07)<br>see abstract, and claims 1-10                                                                                             | 1, 3, 5-7, 9-11, 13-16 |
| A         | CN 106192018 A (BGI-SHENZHEN) 07 December 2016 (2016-12-07)<br>see entire document                                                                                                       | 2, 4, 8, 12            |
| X         | CN 107385042 A (GUANGZHOU FOREVERGEN HEALTH TECHNOLOGY CO., LTD.;<br>GUANGZHOU FOREVERGEN MEDICAL LABORATORY CO., LTD.) 24 November 2017<br>(2017-11-24)<br>see abstract, and claims 1-9 | 1, 3, 5-7, 9-11, 13-16 |
| A         | CN 107385042 A (GUANGZHOU FOREVERGEN HEALTH TECHNOLOGY CO., LTD.;<br>GUANGZHOU FOREVERGEN MEDICAL LABORATORY CO., LTD.) 24 November 2017<br>(2017-11-24)<br>see entire document          | 2, 4, 8, 12            |



Further documents are listed in the continuation of Box C.



See patent family annex.

\* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“E” earlier application or patent but published on or after the international filing date

“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

“O” document referring to an oral disclosure, use, exhibition or other means

“P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

“&amp;” document member of the same patent family

Date of the actual completion of the international search

10 July 2019

Date of mailing of the international search report

24 July 2019

Name and mailing address of the ISA/CN

China National Intellectual Property Administration  
No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing  
100088  
China

Facsimile No. (86-10)62019451

Authorized officer

Telephone No.

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2019/082233

**Box No. I** Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing:
  - a. ☒ forming part of the international application as filed:
    - ☒ in the form of an Annex C/ST.25 text file.
    - ☐ on paper or in the form of an image file.
  - b. ☐ furnished together with the international application under PCT Rule 13*ter*.1(a) for the purposes of international search only in the form of an Annex C/ST.25 text file.
  - c. ☐ furnished subsequent to the international filing date for the purposes of international search only:
    - ☐ in the form of an Annex C/ST.25 text file (Rule 13*ter*.1(a)).
    - ☐ on paper or in the form of an image file (Rule 13*ter*.1(b) and Administrative Instructions, Section 713).
2. ☐ In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that forming part of the application as filed or does not go beyond the application as filed, as appropriate, were furnished.
3. Additional comments:

**INTERNATIONAL SEARCH REPORT**  
**Information on patent family members**

International application No.

**PCT/CN2019/082233**

| Patent document<br>cited in search report |           |   | Publication date<br>(day/month/year) | Patent family member(s) | Publication date<br>(day/month/year) |
|-------------------------------------------|-----------|---|--------------------------------------|-------------------------|--------------------------------------|
| CN                                        | 106192018 | A | 07 December 2016                     | None                    |                                      |
| CN                                        | 107385042 | A | 24 November 2017                     | None                    |                                      |

## 国际检索报告

国际申请号

PCT/CN2019/082233

| <b>A. 主题的分类</b><br>C12Q 1/6806(2018.01)i; C12Q 1/6827(2018.01)i; C12Q 1/6886(2018.01)i; C12Q 1/6855(2018.01)i; C12Q 1/6876(2018.01)i<br>按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                                                                                             |                             |      |                   |         |   |                                                                            |                        |   |                                                                  |             |   |                                                                                             |                        |   |                                                                                    |             |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------|-----------------------------|------|-------------------|---------|---|----------------------------------------------------------------------------|------------------------|---|------------------------------------------------------------------|-------------|---|---------------------------------------------------------------------------------------------|------------------------|---|------------------------------------------------------------------------------------|-------------|
| <b>B. 检索领域</b><br>检索的最低限度文献(标明分类系统和分类号)<br>C12Q<br>包含在检索领域中的除最低限度文献以外的检索文献<br>在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用))<br>DWPI, SIPOABS, CPRSABS, CPEA, CNABS, JPABS, CNMED, KRABS, CNTXT, USTXT, WOTXT, EPTXT, JPTXT, KRTXT, CNKI, ISI Web of Science, 中国专利生物序列检索系统, REGISTRY: 文库, 测序, 接头, 引物, 标签, 锚定, 肝癌, 突变, 循环肿瘤DNA, 无细胞DNA, library, sequencing, adaptor, primer, barcode, index, label, tag, tab, anchor, liver cancer, hepatoma, hepatic carcinoma, mutate, mutation, mutant, variation, ctDNA, circulating tumor DNA, cfDNA, cell-free DNA, SEQ ID NOs: 1-24                                                                                                                                                       |                                                                                             |                             |      |                   |         |   |                                                                            |                        |   |                                                                  |             |   |                                                                                             |                        |   |                                                                                    |             |
| <b>C. 相关文件</b> <table border="1"> <thead> <tr> <th>类 型*</th> <th>引用文件, 必要时, 指明相关段落</th> <th>相关的权利要求</th> </tr> </thead> <tbody> <tr> <td>X</td> <td>CN 106192018 A (深圳华大基因研究院) 2016年 12月 7日 (2016 - 12 - 07)<br/>参见摘要, 权利要求1-10</td> <td>1, 3, 5-7, 9-11, 13-16</td> </tr> <tr> <td>A</td> <td>CN 106192018 A (深圳华大基因研究院) 2016年 12月 7日 (2016 - 12 - 07)<br/>参见全文</td> <td>2, 4, 8, 12</td> </tr> <tr> <td>X</td> <td>CN 107385042 A (广州永诺健康科技有限公司 广州永诺医学检验所有限公司) 2017年 11月 24日 (2017 - 11 - 24)<br/>参见摘要, 权利要求1-9</td> <td>1, 3, 5-7, 9-11, 13-16</td> </tr> <tr> <td>A</td> <td>CN 107385042 A (广州永诺健康科技有限公司 广州永诺医学检验所有限公司) 2017年 11月 24日 (2017 - 11 - 24)<br/>参见全文</td> <td>2, 4, 8, 12</td> </tr> </tbody> </table> |                                                                                             |                             | 类 型* | 引用文件, 必要时, 指明相关段落 | 相关的权利要求 | X | CN 106192018 A (深圳华大基因研究院) 2016年 12月 7日 (2016 - 12 - 07)<br>参见摘要, 权利要求1-10 | 1, 3, 5-7, 9-11, 13-16 | A | CN 106192018 A (深圳华大基因研究院) 2016年 12月 7日 (2016 - 12 - 07)<br>参见全文 | 2, 4, 8, 12 | X | CN 107385042 A (广州永诺健康科技有限公司 广州永诺医学检验所有限公司) 2017年 11月 24日 (2017 - 11 - 24)<br>参见摘要, 权利要求1-9 | 1, 3, 5-7, 9-11, 13-16 | A | CN 107385042 A (广州永诺健康科技有限公司 广州永诺医学检验所有限公司) 2017年 11月 24日 (2017 - 11 - 24)<br>参见全文 | 2, 4, 8, 12 |
| 类 型*                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | 引用文件, 必要时, 指明相关段落                                                                           | 相关的权利要求                     |      |                   |         |   |                                                                            |                        |   |                                                                  |             |   |                                                                                             |                        |   |                                                                                    |             |
| X                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      | CN 106192018 A (深圳华大基因研究院) 2016年 12月 7日 (2016 - 12 - 07)<br>参见摘要, 权利要求1-10                  | 1, 3, 5-7, 9-11, 13-16      |      |                   |         |   |                                                                            |                        |   |                                                                  |             |   |                                                                                             |                        |   |                                                                                    |             |
| A                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      | CN 106192018 A (深圳华大基因研究院) 2016年 12月 7日 (2016 - 12 - 07)<br>参见全文                            | 2, 4, 8, 12                 |      |                   |         |   |                                                                            |                        |   |                                                                  |             |   |                                                                                             |                        |   |                                                                                    |             |
| X                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      | CN 107385042 A (广州永诺健康科技有限公司 广州永诺医学检验所有限公司) 2017年 11月 24日 (2017 - 11 - 24)<br>参见摘要, 权利要求1-9 | 1, 3, 5-7, 9-11, 13-16      |      |                   |         |   |                                                                            |                        |   |                                                                  |             |   |                                                                                             |                        |   |                                                                                    |             |
| A                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      | CN 107385042 A (广州永诺健康科技有限公司 广州永诺医学检验所有限公司) 2017年 11月 24日 (2017 - 11 - 24)<br>参见全文          | 2, 4, 8, 12                 |      |                   |         |   |                                                                            |                        |   |                                                                  |             |   |                                                                                             |                        |   |                                                                                    |             |
| <input type="checkbox"/> 其余文件在C栏的续页中列出。 <input checked="" type="checkbox"/> 见同族专利附件。                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |                                                                                             |                             |      |                   |         |   |                                                                            |                        |   |                                                                  |             |   |                                                                                             |                        |   |                                                                                    |             |
| * 引用文件的具体类型:<br>“A” 认为不特别相关的表示了现有技术一般状态的文件<br>“E” 在国际申请日的当天或之后公布的在先申请或专利<br>“L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件(如具体说明的)<br>“O” 涉及口头公开、使用、展览或其他方式公开的文件<br>“P” 公布日先于国际申请日但迟于所要求的优先权日的文件<br>“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件<br>“X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性<br>“Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性<br>“&” 同族专利的文件                                                                                                                                                                                                                                                                                                                       |                                                                                             |                             |      |                   |         |   |                                                                            |                        |   |                                                                  |             |   |                                                                                             |                        |   |                                                                                    |             |
| 国际检索实际完成的日期<br>2019年 7月 10日                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                                                                                             | 国际检索报告邮寄日期<br>2019年 7月 24日  |      |                   |         |   |                                                                            |                        |   |                                                                  |             |   |                                                                                             |                        |   |                                                                                    |             |
| ISA/CN的名称和邮寄地址<br>中国国家知识产权局(ISA/CN)<br>中国北京市海淀区蓟门桥西土城路6号 100088<br>传真号 (86-10)62019451                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |                                                                                             | 授权官员<br>幸颖<br>电话号码 62411044 |      |                   |         |   |                                                                            |                        |   |                                                                  |             |   |                                                                                             |                        |   |                                                                                    |             |

第I栏

核苷酸和/或氨基酸序列 (续第1页第1. c项)

1. 关于国际申请中所公开的任何核苷酸和/或氨基酸序列, 国际检索是基于下列序列表进行的:

a. ☒ 作为国际申请的一部分提交的:

☒ 附件C/ST. 25文本文件形式

☐ 纸件或图形文件形式

b. ☐ 根据细则13之三. 1(a) 仅为国际检索目的以附件C/ST. 25文本文件形式与国际申请同时提交的:

c. ☐ 仅为国际检索目的在国际申请日之后提交的:

☐ 附件C/ST. 25文本文件形式 (细则13之三. 1(a))

☐ 纸件或图形文件形式 (细则13之三. 1(b) 和行政规程第713段)

2. ☐ 另外, 在提交/提供了多个版本或副本的序列表的情况下, 提供了关于随后提交的或附加的副本中的信息与申请时提交的作为申请一部分的序列表的信息相同或未超出申请时提交的申请中的信息范围 (如适用) 的所需声明。

3. 补充意见:

国际检索报告  
关于同族专利的信息

国际申请号

PCT/CN2019/082233

| 检索报告引用的专利文件 |           |   | 公布日<br>(年/月/日) | 同族专利 | 公布日<br>(年/月/日) |
|-------------|-----------|---|----------------|------|----------------|
| CN          | 106192018 | A | 2016年 12月 7日   | 无    |                |
| CN          | 107385042 | A | 2017年 11月 24日  | 无    |                |