

(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
2019 年 1 月 24 日 (24.01.2019)



(10) 国际公布号
WO 2019/014936 A1

- (51) 国际专利分类号:
C12Q 1/68 (2018.01)
- (21) 国际申请号: PCT/CN2017/093914
- (22) 国际申请日: 2017 年 7 月 21 日 (21.07.2017)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (71) 申请人: 深圳华大基因研究院(BGI SHENZHEN) [CN/CN]; 中国广东省深圳市盐田区北山工业区综合楼, Guangdong 518083 (CN)。
- (72) 发明人: 杨林(YANG, Lin); 中国广东省深圳市盐田区北山工业区综合楼, Guangdong 518083 (CN)。高雅(GAO, Ya); 中国广东省深圳市盐田区北山工业区综合楼, Guangdong 518083 (CN)。

蒲丹丹(PU, Dandan); 中国广东省深圳市盐田区北山工业区综合楼, Guangdong 518083 (CN)。张海萍(ZHANG, Haiping); 中国广东省深圳市盐田区北山工业区综合楼, Guangdong 518083 (CN)。程云阳(CHENG, Yunyang); 中国广东省深圳市盐田区北山工业区综合楼, Guangdong 518083 (CN)。陈芳(CHEN, Fang); 中国广东省深圳市盐田区北山工业区综合楼, Guangdong 518083 (CN)。蒋慧(JIANG, Hui); 中国广东省深圳市盐田区北山工业区综合楼, Guangdong 518083 (CN)。

(74) 代理人: 深圳鼎合诚知识产权代理有限公司(DHC IP ATTORNEYS); 中国广东省深圳福田区金田路与福华路交汇处现代商务大厦2201, Guangdong 518048 (CN)。

(54) Title: TARGETED ENRICHMENT METHOD AND KIT FOR DETECTING LOW-FREQUENCY MUTATION

(54) 发明名称: 用于检测低频突变的靶向富集方法和试剂盒

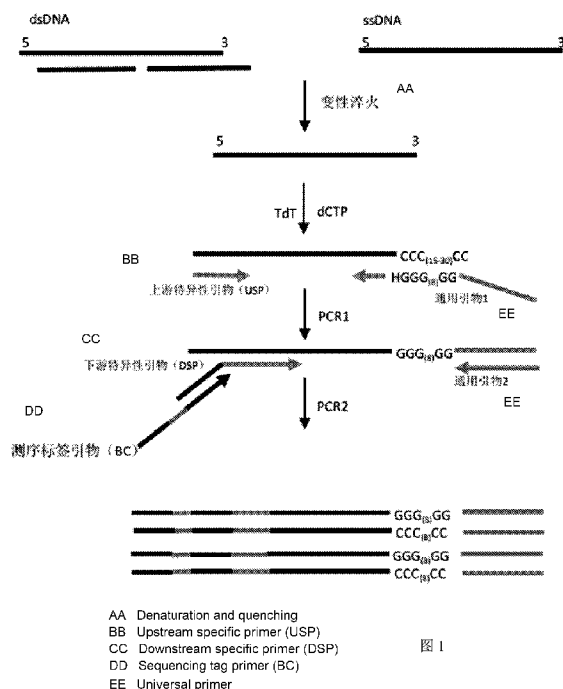


图 1

(57) Abstract: Provided are a targeted enrichment method and kit for detecting a low-frequency mutation. The method comprises: adding a mononucleotide or polynucleotide tail of a single base to the 3' end of single-stranded DNA and/or of each strand of double-stranded DNA under the action of a terminal transferase; performing a first instance of PCR amplification on a region where a mutation target site is located by using an upstream specific primer and a first universal primer, wherein the first universal primer comprises a sequencing primer sequence 1 at the 5' end thereof and continuous single bases, complementary to the mononucleotide or polynucleotide tail, at the 3' end thereof; and performing, by using a downstream specific primer and a second universal primer, a second instance of PCR amplification on the product, which has undergone the first instance of PCR amplification, wherein the second universal primer comprises the sequencing primer sequence 1 of the first universal primer.

(57) 摘要: 本发明提供了一种用于检测低频突变的靶向富集方法和试剂盒。该方法包括: 在末端转移酶的作用下, 在单链DNA和/或双链DNA每条链的3'端加上一段单一碱基的单聚核苷酸尾巴; 以上游特异性引物和第一通用引物对突变目标位点所在的区域进行第一次PCR扩增, 其中第一通用引物包括5'端的测序引物序列1和3'端的与单聚核苷酸尾巴互补的连续单一碱基; 以下游特异性引物和第二通用引物对第一次PCR扩增的产物进行第二次PCR扩增, 其中第二通用引物包括第一通用引物的测序引物序列1。

(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))。

用于检测低频突变的靶向富集方法和试剂盒

技术领域

本发明涉及分子生物学技术领域，具体涉及一种用于检测低频突变的靶向富集方法和试剂盒。

背景技术

二代测序（NGS）技术的发展，为现代基因组学的研究打开了新的局面，然而全基因组测序的成本和分析的复杂程度还是让科研人员倍感困难，尽管二代测序的通量越来越高，而费用越来越低，但它仍不是大多数遗传实验室的可行选择。对于复杂疾病的研究更是如此，这类研究至少需要数百个样本，以实现足够的统计能力。然而，这么多样本的全基因组测序，无论从成本考虑，还是从数据分析考虑，都是相对困难的。

目标区域靶向富集测序（Target region sequencing）技术的出现，缓解了上述问题。目标区域靶向富集测序技术，是对感兴趣的目标基因进行富集并结合二代测序技术进行测序的技术，得到目标区域的碱基信息，已达到检测疾病的目的，相对于全基因组测序，靶向富集测序技术可以降低测序成本，简化信息分析流程，提高目标区域测序深度，提高检测的灵敏性和准确性。

目前市场上针对目标区域进行富集的技术主要有两种，一种是基于探针的捕获富集技术，另一种是基于多重 PCR 的富集技术。基于探针的捕获富集技术，利用核酸分子碱基互补杂交的原理，根据目标区域设计反向互补的寡聚核苷酸探针，然后打断基因组 DNA，加上用于测序的接头后与探针杂交，洗脱未杂交上的 DNA，回收目标 DNA 片段，再进行文库制备进行 DNA 测序。这种技术需要样本起始 DNA 量高（一般需要达到微克），实验操作时间长，实验繁琐，不利于自动化建库，数据利用率低，成本高。基于多重 PCR 的富集技术，是根据目标区域设计引物，然后通过多重 PCR 富集目标区域，再对 PCR 产物进行文库制备进行 DNA 测序。这种技术虽然实验时间短，但前期需要进行复杂的引物设计工作，并且在后期需要进行大量繁琐的引物优化工作。此外，上述两种技术对模板的量和完整度有严格的要求，对细胞游离 DNA、高度降解 DNA、石蜡包埋甲醛固定的医学等样本无能为力。因此开发一种简单的针对短片段 DNA 能够进行有效富集的靶向富集方法显得尤为重要。

发明内容

本发明提供一种用于检测低频突变的靶向富集方法和试剂盒，能够对短片段、单链、双链及损失的 DNA 进行有效富集，结合二代测序技术可以检测发生在 DNA 中的低频突变。

根据第一方面，一种实施例中提供一种用于检测低频突变的靶向富集方法，包括：

在末端转移酶的作用下，在单链 DNA 和/或双链 DNA 每条链的 3'端加上一段单一碱基的单聚核苷酸尾巴；

以上游特异性引物和第一通用引物对突变目标位点所在的区域进行第一次 PCR 扩增，其中上述上游特异性引物以与其互补的目标序列为锚定位点，上述第一通用引物以上述单聚核苷酸尾巴为锚定位点，且上述第一通用引物包括 5'端的测序引物序列 1 和 3'端的与上述单聚核苷酸尾巴互补的连续单一碱基；

以下游特异性引物和第二通用引物对上述第一次 PCR 扩增的产物进行第二次 PCR 扩增，其中上述下游特异性引物位于上述上游特异性引物的下游且 5'端带有测序引物序列 2，上述第二通用引物包括上述第一通用引物 5'端的测序引物序列 1。

根据第二方面，一种实施例中提供一种用于检测低频突变的靶向富集试剂盒，包括：

末端转移酶和单一碱基核苷酸，用于在末端转移酶的作用下，在单链 DNA 和/或双链 DNA 每条链的 3'端加上一段单一碱基的单聚核苷酸尾巴；

上游特异性引物和第一通用引物，用于对突变目标位点所在的区域进行第一次 PCR 扩增，其中上述上游特异性引物以与其互补的目标序列为锚定位点，上述第一通用引物以上述单聚

核苷酸尾巴为锚定位点，且上述第一通用引物包括 5'端的测序引物序列 1 和 3'端的与上述单聚核苷酸尾巴互补的连续单一碱基；

下游特异性引物和第二通用引物，用于对上述第一次 PCR 扩增的产物进行第二次 PCR 扩增，其中上述下游特异性引物位于上述上游特异性引物的下游且 5'端带有测序引物序列 2，上述第二通用引物包括上述第一通用引物 5'端的测序引物序列 1。

本发明的方法，能够有效富集短片段单链 DNA、双链 DNA 以及带缺口损伤的 DNA 且具有极高的模板利用率；具有高度的检测灵敏性，可以检测低至 0.1%的低频突变；能够一次对多个靶向区域进行有效富集，并且能够保证很好的特异性、均一性和稳定性。

附图说明

图 1 为本发明一种实施例的用于检测低频突变的靶向富集方法的原理示意图；

图 2 为本发明另一种实施例的用于检测低频突变的靶向富集方法的原理示意图；

图 3 为本发明一种实施例中的分子标签校正示意图；

图 4 为本发明实施例 1 中靶向富集文库的安捷伦 2100 质检结果图；

图 5 为本发明实施例 1 中 HBB 各扩增子均一性结果图；

图 6 为本发明实施例 2 中靶向富集文库的安捷伦 2100 质检结果图；

图 7 为本发明实施例 2 中 10 个扩增子区域测序数据深度分布图；

图 8 为本发明实施例 2 中分子标签校正后的 10 个扩增子区域测序数据深度分布图；

图 9 为本发明实施例 2 中突变检测一致性结果图。

具体实施方式

下面通过具体实施方式结合附图对本发明作进一步详细说明。在以下的实施方式中，很多细节描述是为了使得本发明能被更好的理解。然而，本领域技术人员可以毫不费力的认识到，其中部分特征在不同情况下是可以省略的，或者可以由其他原材料、方法所替代。在某些情况下，本发明相关的一些操作并没有在说明书中显示或者描述，这是为了避免本发明的核心部分被过多的描述所淹没，而对于本领域技术人员而言，详细描述这些相关操作并不是必要的，他们根据说明书中的描述以及本领域的一般技术知识即可完整了解相关操作。

本文中为部件所编序号本身，例如“第一”、“第二”等，仅用于区分所描述的对象，不具有任何顺序或技术含义。

图 1 示出了本发明一种实施例的用于检测低频突变的靶向富集方法的原理，该方法包括：

步骤I：在末端转移酶的作用下，在单链 DNA 和/或双链 DNA 每条链的 3'端加上一段单一碱基的单聚核苷酸尾巴。

具体地，在图 1 所示的实施例中，双链 DNA 分子（dsDNA）通过变性解链后得到单链 DNA（ssDNA），通过末端转移酶在单链 DNA 的 3'端加入 15-30 个单一碱基 C（或者 A、T、G 中的任意一种），得到一个单聚核苷酸尾巴，此时所有单链、双链及损伤 DNA 模板的 3'端都有一段由单一碱基组成的长度为 15-30bp 的单聚核苷酸尾巴。需要说明的是，双链 DNA 分子不经变性也可以在末端转移酶的作用下直接在每条链的 3'端加入单聚核苷酸尾巴。

步骤II：以上游特异性引物 and 第一通用引物对突变目标位点所在的区域进行第一次 PCR 扩增，其中上游特异性引物以与其互补的目标序列为锚定位点，第一通用引物以单聚核苷酸尾巴为锚定位点，且第一通用引物包括 5'端的测序引物序列 1 和 3'端的与单聚核苷酸尾巴互补的连续单一碱基。

具体地，在图 1 所示的实施例中，在突变目标位点前 25-150bp 设计上游特异性引物（USP）。一端以单聚核苷酸尾巴为锚定位点，另一端以上游特异性引物（USP）互补的目标序列为锚定位点进行 10-20 个循环的扩增。以单聚核苷酸尾巴为锚定位点的引物序列称为“第一通用引物”（即图 1 中的通用引物 1）包括：5'端的测序引物序列 1（例如 BGISEQ-500、Illumina 或 Proton 等测序平台的测序引物序列），中间的连续碱基 G（或者 T、A、C 中的任意一种），

连续碱基 C 或 G 的个数为 11-15 个, T_m 值范围为 54-70°C, 个数太多不利于 PCR 扩增; 连续碱基 A 或 T 的个数为 25-35 个, T_m 值范围为 54-62°C, 个数太多不利于 PCR 扩增。在图 1 所示的优选的实施例中, 通用引物 1 中间是 13 个连续碱基 G, 并且, 优选地, 最后加上一个简并碱基 H (在其他情况下可以是 V、B 或 D), 这个末端的简并碱基用来固定 3'端产物的长度, 将末端为不同长度 C 碱基的模板经过带有简并碱基的引物进行 PCR 扩增以后, 得到 3'端为固定长度 C (在其他情况下可以是 A、T 或 G) 碱基的产物。得到的 PCR 产物经过磁珠纯化后, 产物从 5'端到 3'端依次是目标上游特异性引物序列、目标区域序列、连续的单聚核苷酸 C (在其他情况下可以是 A、T 或 G), 以及最后的测序引物序列。需要说明的是, 本发明实施例的方法, 适用于各种突变类型的靶向富集, 包括单核苷酸多态性 (SNP)、插入和删除 (INDEL) 以及拷贝数变异 (CNV) 等。

步骤III: 以下游特异性引物和第二通用引物对第一次 PCR 扩增的产物进行第二次 PCR 扩增, 其中下游特异性引物位于上游特异性引物的下游且 5'端带有测序引物序列 2, 第二通用引物包括第一通用引物 5'端的测序引物序列 1。

具体地, 在图 1 所示的实施例中, 在上游特异性引物 (USP) 的下游 (例如 USP 的下游 0-10bp 的区域) 设计下游特异性引物 (DSP), 并在下游特异性引物 (DSP) 的 5'端加上测序引物序列 2 (例如 BGISEQ-500、Illumina 或 Proton 等测序平台的测序引物序列), 用带有测序引物序列 2 的下游特异性引物 (DSP) 和通用引物 2 (也称为“第二通用引物”) 进行 15-30 个循环的扩增。所述第二通用引物包括所述第一通用引物 5'端的测序引物序列 1。

作为优选技术方案, 还可以引入一个测序标签引物 (BC), 所述测序标签引物 3'端包括所述下游特异性引物 5'端的测序引物序列 2。在 PCR 的第一个循环中, 下游特异性引物 (DSP) 和通用引物 2 进行扩增, 由于下游特异性引物 (DSP) 5'端也有测序引物序列 2, 因此得到产物的两端都带有测序引物序列; 从第二个循环开始, 以测序引物序列为锚定位点, 两个通用引物 (即通用引物 2 和测序标签引物 (BC)) 与下游特异性引物同时进行扩增, 得到靶向测序文库。引入测序标签引物的好处在于: (1) 减少特异性扩增步骤, 增加通用扩增步骤, 有利于不同扩增子区域的均一性; (2) 减少特异性引物引入的接头序列, 更多的接头序列由测序标签引物引入。

本实施例的方法具有以下有益效果:

(1) 对单链和双链 DNA 都能有效富集且具有极高的模板利用率。具体而言, 末端转移酶对游离单链、双链及损失 DNA 都能够进行碱基的添加, 且对模板的添加效率能够达到 99% 以上, 即 99% 以上的模板都能在 3'端加上单聚核苷酸尾巴。通过特异性引物和单聚核苷酸尾巴对靶向区域进行富集, 只需要模板上有特异性引物结合位点就能够产生有效富集。对于大小为 165bp 左右的游离 DNA, 引物所占范围在 25bp 左右, 对模板的利用率可以达到 $(165-25\text{bp})/165\text{bp}=85\%$, 因此对于一些高度片段化单链、双链以及损失 DNA 都能够达到很好的富集效果。

(2) 能够对多个靶向区域进行富集。具体而言, 采用特异性引物和锚定引物进行多重 PCR 扩增, 特异性引物可以是针对不同靶向区域的引物, 锚定引物是固定的引物序列, 通过混合的特异性引物和固定的靶向引物进行多个区域的富集, 并且采用两轮巢式 PCR 可以提高富集的特异性, 得到多个区域的靶向测序文库。

图 2 示出了本发明另一种实施例的用于检测低频突变的靶向富集方法的原理, 该方法包括:

双链 DNA 分子通过变性解链后得到单链 DNA, 通过末端转移酶对单链 DNA 模板的 3'端随机加入 5-10 个随机碱基, 作为分子标签, 用于标记原始模板。其中加上的分子标签的种类达到百万种 ($N=5^4+6^4+7^4+\dots+10^4=1397760$), 远远超过模板的个数, 因此每条模板大于 99.9% 概率都会加上一个唯一的由 A、T、C、G 四种碱基组成的 5-10bp 的分子标签, 用以对原始模板进行标记。通过磁珠纯化去除掉残留的 dNTP。

然后再通过末端转移酶在分子标签的基础上再加入 15-30 个单一碱基 C (也可以是 A、T、

G 中的任意一种), 得到一个单聚核苷酸尾巴, 然后再用磁珠纯化去除掉残留的 dCTP (或 dATP、dTTP、dGTP)。此时所有的游离单链、双链及损伤 DNA 模板的 3'端都带有一个由四种碱基组成的 5-10bp 的分子标签和由一种碱基组成的长度为 15-30bp 的单聚核苷酸尾巴。需要说明的是, 双链 DNA 分子不经变性也可以在末端转移酶的作用下直接在每条链的 3'端加入分子标签和单聚核苷酸尾巴。

在突变目标位点前 25-150bp 设计上游特异性引物 (USP)。一端以单聚核苷酸尾巴为锚定位点, 另一端以上游特异性引物 (USP) 互补的目标序列为锚定位点进行 10-20 个循环的扩增。以单聚核苷酸尾巴为锚定位点的引物序列称为“第一通用引物”(即图 2 中的通用引物 1) 包括: 5'端的测序引物序列 1 (例如 BGISEQ-500、Illumina 或 Proton 等测序平台的测序引物序列), 中间的连续碱基 G (或者 T、A、C 中的任意一种), 连续碱基 C 或 G 的个数为 11-15 个, T_m 值范围为 54-70℃, 个数太多不利于 PCR 扩增; 连续碱基 A 或 T 的个数为 25-35 个, T_m 值范围为 54-62℃, 个数太多不利于 PCR 扩增。在图 2 所示的优选的实施例中, 通用引物 1 中间是 13 个连续碱基 G, 并且, 优选地, 最后加上一个简并碱基 H (在其他情况下可以是 V、B 或 D), 这个末端的简并碱基用来固定 3'端产物的长度, 将末端为不同长度 C 碱基的模板经过带有简并碱基的引物进行 PCR 扩增以后, 得到 3'端为固定长度 C (在其他情况下可以是 A、T 或 G) 碱基的产物。得到的 PCR 产物经过磁珠纯化后, 产物从 5'端到 3'端依次是目标上游特异性引物序列、目标区域序列、5-10 个随机碱基组成的分子标签、连续的单聚核苷酸 C (在其他情况下可以是 A、T 或 G), 以及最后的测序引物序列。需要说明的是, 本发明实施例的方法, 适用于各种突变类型的靶向富集, 包括单核苷酸多态性 (SNP)、插入和删除 (INDEL) 以及拷贝数变异 (CNV) 等。

在上游特异性引物 (USP) 的下游设计下游特异性引物 (DSP), 并在下游特异性引物 (DSP) 的 5'端加上测序引物序列 2 (例如 BGISEQ-500、Illumina 或 Proton 等测序平台的测序引物序列), 用带有测序引物序列 2 的下游特异性引物 (DSP) 和通用引物 2 (也称为“第二通用引物”) 进行 15-30 个循环的扩增。所述第二通用引物包括所述第一通用引物 5'端的测序引物序列 1。

作为优选技术方案, 还可以引入一个测序标签引物 (BC), 所述测序标签引物 3'端包括所述下游特异性引物 5'端的测序引物序列 2。在 PCR 的第一个循环中, 下游特异性引物 (DSP) 和通用引物 2 进行扩增, 由于下游特异性引物 (DSP) 5'端也有测序引物序列 2, 因此得到产物的两端都带有测序引物序列; 从第二个循环开始, 以测序引物序列为锚定位点, 两个通用引物 (即通用引物 2 和测序标签引物 (BC)) 与下游特异性引物同时进行扩增, 得到靶向测序文库。引入测序标签引物的好处在于: (1) 减少特异性扩增步骤, 增加通用扩增步骤, 有利于不同扩增子区域的均一性; (2) 减少特异性引物引入的接头序列, 更多的接头序列由测序标签引物引入。

图 2 所示的方法除具有图 1 所示的方法的有益效果以外, 还具有以下有益效果: 高度的检测灵敏性, 可以检测进一步降低至 0.1% 的低频突变。具体而言, 通过末端转移酶在游离 DNA 模板的 3'端随机加上的由四种碱基组成的长度为 5-10bp 的一段随机序列, 该序列的种类可达百万种, 可以对起始模板进行唯一标记, 通过分子标记结合信息分析方法可以检测进一步降低至 0.1% 的低频突变。

如图 3 所示, 对本实施例得到的靶向测序文库进行双端测序, 特异性引物那一端用来进行靶向富集区域的检测, 另一端用来读取分子标签信息用以标记模板, 通过结合特定的数据分析算法通过分子标签进行 PCR 错误和测序错误的去除, 实现检测非常低频的突变。

以下通过实施例详细说明本发明的技术方案和效果, 应当理解的是, 实施例仅是示例性的, 不能理解为对本发明保护范围的限制。

实施例 1: 地贫父源突变检测

针对与 beta 地贫相关的 HBB 基因设计 19 对引物, 对常见的 beta 地贫突变位点进行检测, 针对父母携带不同突变类型, 检测孕妇血浆游离 DNA 中胎儿是否携带父源突变来达到排除性诊断。

血浆来源是 12 周孕妇血浆游离 DNA，母亲携带 CD41/42 β E (del CTTT) 突变，父亲携带 CD71/72(Ins A) 突变，通过对血浆游离 DNA 捕获建库后测序，检测胎儿是否携带父亲来源的 (CD71/72) 突变。

实验步骤：

1. 寡聚核苷酸尾巴的添加

先将 cfDNA 在 95℃ 热变性 5 分钟，然后迅速插冰上，然后再进行酶反应。通过末端转移酶（美国 NEB 公司的 Terminal Transferase，货号 M0315S）在 DNA 的 3' 端加上寡聚核苷酸尾巴。

反应体系如表 1 所示：

表 1

试剂	体积
末端转移酶缓冲液 (10X)	5 μ l
二氯化钴 (10mM)	5 μ l
dCTP 混合物 (50 μ M)	5 μ l
末端转移酶	0.5 μ l
DNA	34.5 μ l
总共	50 μ l

37℃ 孵育 30 分钟，然后加入 10 μ l 浓度为 0.5M 的 EDTA 终止反应。加入 1.8 倍体积的 Agencourt AMPure XP 磁珠（美国贝克曼库尔特有限公司）108 μ l，按照说明书进行纯化，纯化后用 20 μ l 蒸馏水溶解 DNA。

2. 第一轮 PCR 扩增

PCR 酶采用 NEB 公司的 Q5® Hot Start High-Fidelity 2X Master Mix，货号：M0494L。

PCR 反应体系如表 2 所示：

表 2

试剂	体积
2XQ5 聚合酶混合液	25 μ l
上游特异性引物池 (10 μ M)	2.5 μ l
通用引物 1 (10 μ M)	2.5 μ l
上一步 DNA	20 μ l
总共	50 μ l

上游特异性引物池如表 3 所示，通用引物 1 如表 4 所示。

表 3

引物名称	引物序列(5-3')	引物编号
HBB-USP1	TGAGAGATGCAGGATAAGCAA	SEQ ID NO: 1
HBB-USP2	GTTGCCAATGTGCATTAGCT	SEQ ID NO: 2
HBB-USP3	TCCCAAGGTTTGAAGTAGCTC	SEQ ID NO: 3
HBB-USP4	TTAGGGAACAAAGGAACCTTTAAT	SEQ ID NO: 4
HBB-USP5	GTGGGAGGAAGATAAGAGGTATGA	SEQ ID NO: 5
HBB-USP6	GCTGCTATTAGCAATATGAAACCTC	SEQ ID NO: 6
HBB-USP7	TGATACATTGTATCATTATTGCCCTG	SEQ ID NO: 7
HBB-USP8	TAGTAATGTACTAGGCAGACTGTGT	SEQ ID NO: 8
HBB-USP9	TCATTCGTCTGTTTCCCATTTC	SEQ ID NO: 9
HBB-USP10	CCTTCCTATGACATGAACTTAACC	SEQ ID NO: 10
HBB-USP11	GCGTCCCATAGACTCACCC	SEQ ID NO: 11
HBB-USP12	CACCGAGCACTTTCTTGCC	SEQ ID NO: 12
HBB-USP13	GAAAATAGACCAATAGGCAGAGAGA	SEQ ID NO: 13

HBB-USP14	CCTTAAACCTGTCTTGTAACCTTGAT	SEQ ID NO: 14
HBB-USP15	CAGTAACGGCAGACTTCTCCTC	SEQ ID NO: 15
HBB-USP16	GTTGTGTCAGAAGCAAATGTAAGC	SEQ ID NO: 16
HBB-USP17	CTGACTTTTATGCCAGCC	SEQ ID NO: 17
HBB-USP18	CTAGGGTGTGGCTCCACAG	SEQ ID NO: 18
HBB-USP19	CAGCCGTACCTGTCCTTG	SEQ ID NO: 19

上游特异性引物池由表3所示的引物等摩尔数混合组成。

表4

引物名称	引物序列(5-3')	引物编号
通用引物1	<u>CGACATGGCTACGATCCGACTTGGGGGGGGGGGGGGH</u>	SEQ ID NO: 20
通用引物2	<u>Phos/GAACGACATGGCTACGATCCGACTT</u>	SEQ ID NO: 21
测序标签引物	<u>TGTGAGCCAAGGAGTTGCTGCGTACATTTGTCTTCCT</u> <u>AAGACCGCTTGGCCTCCGACTT</u>	SEQ ID NO: 22

扩增体系如下表5所示：

表5

步骤1	98℃ 5min
步骤2	98℃ 10s
步骤3	62℃ 5min
步骤4	重复步骤2-3，10个循环
步骤5	72℃ 5min

加入 1.8 倍体积的 Agencourt AMPure XP 磁珠（美国贝克曼库尔特有限公司）90μl，按照说明书进行纯化，纯化后用 20μl 蒸馏水溶解 DNA。

3. 第二轮 PCR 扩增

PCR 反应体系如下表 6 所示：

表 6

试剂	体积
2XQ5 聚合酶混合液	25μl
下游特异性引物池（10μM）	0.25μl
通用引物2（10μM）	2.5μl
测序标签引物（10μM）	2.5μl
上一步DNA	19.75μl
总共	50μl

下游特异性引物池如表7所示，通用引物2和测序标签引物如表4所示。

表7

引物名称	引物序列(5-3')	引物编号
HBB-DSP1	AAGACCGCTTGGCCTCCGACTT TGAGAGATGCAGGATAAGCAA	SEQ ID NO: 23
HBB-DSP2	AAGACCGCTTGGCCTCCGACTT GTTGCCAATGTGCATTAGCT	SEQ ID NO: 24
HBB-DSP3	AAGACCGCTTGGCCTCCGACTT TCCCAAGGTTTGAAGTAGCTC	SEQ ID NO: 25
HBB-DSP4	AAGACCGCTTGGCCTCCGACTT TTAGGGAACAAAGGAACCTTTAAT	SEQ ID NO: 26
HBB-DSP5	AAGACCGCTTGGCCTCCGACTT	SEQ ID NO: 27

	GTGGGAGGAAGATAAGAGGTATGA	
HBB-DSP6	AAGACCGCTTGGCCTCCGACTT GCTGCTATTAGCAATATGAAACCTC	SEQ ID NO: 28
HBB-DSP7	AAGACCGCTTGGCCTCCGACTT TGATACATTGTATCATTATTGCCCTG	SEQ ID NO: 29
HBB-DSP8	AAGACCGCTTGGCCTCCGACTT TAGTAATGTACTAGGCAGACTGTGT	SEQ ID NO: 30
HBB-DSP9	AAGACCGCTTGGCCTCCGACTT TCATTCTGTCTGTTTCCCATTTC	SEQ ID NO: 31
HBB-DSP10	AAGACCGCTTGGCCTCCGACTT CCTTCCTATGACATGAACTTAACC	SEQ ID NO: 32
HBB-DSP11	AAGACCGCTTGGCCTCCGACTT GCGTCCCATAGACTCACCC	SEQ ID NO: 33
HBB-DSP12	AAGACCGCTTGGCCTCCGACTT CACCGAGCACTTTCTTGCC	SEQ ID NO: 34
HBB-DSP13	AAGACCGCTTGGCCTCCGACTT GAAAATAGACCAATAGGCAGAGAGA	SEQ ID NO: 35
HBB-DSP14	AAGACCGCTTGGCCTCCGACTT CCTTAAACCTGTCTTGTAACCTTGAT	SEQ ID NO: 36
HBB-DSP15	AAGACCGCTTGGCCTCCGACTT CAGTAACGGCAGACTTCTCCTC	SEQ ID NO: 37
HBB-DSP16	AAGACCGCTTGGCCTCCGACTT GTTGTGTCAGAAGCAAATGTAAGC	SEQ ID NO: 38
HBB-DSP17	AAGACCGCTTGGCCTCCGACTT CTGACTTTTATGCCAGCC	SEQ ID NO: 39
HBB-DSP18	AAGACCGCTTGGCCTCCGACTT CTAGGGTGTGGCTCCACAG	SEQ ID NO: 40
HBB-DSP19	AAGACCGCTTGGCCTCCGACTT CAGCCGTACCTGTCCTTGG	SEQ ID NO: 41

下游特异性引物池由表7所示的引物等摩尔数混合组成。

扩增体系如下表8所示：

表8

步骤1	98℃ 5min
步骤2	98℃ 10s
步骤3	62℃ 5min
步骤4	98℃ 10s
步骤5	62℃ 2min
步骤6	重复步骤4-5，25个循环
步骤7	72℃ 5min

加入 1 倍体积的 Agencourt AMPure XP 磁珠（美国贝克曼库尔特有限公司）50μl，按照说明书进行纯化，纯化后用 30μl 蒸馏水溶解 DNA。

4. 文库质检

用安捷伦 2100 检测靶向富集文库，其质检结果如图 4 所示。

5. 上机测序

质检合格后，采用 BGISEQ-500 测序平台进行测序，单端 100bp 测序，得到的下机数据

经过数据转换和质量过滤后采用如下信息分析。

6. 信息分析

首先将得到的数据去接头，得到单端测序结果，比对基因组（参考基因组 hg19），通过数据分析对目标位点的突变进行统计，得到目标位点的信息。结果如表 9-10 所示。根据目标区域位置统计各目标区域的深度，得出目标区域均一性信息，结果如图 5 所示，从图可以看到均一性好。

表 9：下机数据统计

样本	下机数据	比对率	捕获率	0.1X平均深度
1	6789141	93.2%	93.5%	100%

表10：Beta地贫检测结果

样本编号	母亲 (M) /父亲 (F)	父母的突变类型	孕妇血浆检测到的 突变类型和频率	胎儿是否正常
B333	M	CD41/42 βE	del CTTT 54.7%	正常
	F	CD71/72	Ins A 0.0%	

结论：检测得到胎儿不携带父亲的致病突变，通过sanger验证得到胎儿不携带父亲的致病突变，与本实施例的检测结果相同。可以通过这种方法来达到无创检测胎儿是否beta地贫父源突变。

实施例2：血浆游离DNA低频突变检测

针对肺癌相关的10个热点突变设计引物，对血浆游离DNA构建靶向测序文库，结合高通量测序和特定的信息分析对与肺癌相关的热点区域进行检测。

血浆游离DNA采用的是Horizon公司cfDNA标准品：0.1% Multiplex I cfDNA Reference Standard (货号: HD779)，突变信息如表11，起始用量10ng，按照下面进行实验。

表11：Horizon公司cfDNA标准品突变信息

基因	突变名称	突变类型	突变频率
BRAF	V600E	c.1799T>A(exon15)	0.00%
cKIT	D816V	c.2447A>T	0.00%
EGFR	G719S	c.2155G>A	0.00%
EGFR	T790M	c.2369C>T	0.10%
EGFR	L858R	c.2573T>G	0.10%
EGFR	ΔE746-A750	c.2235_2249del15 (Deletion)	0.10%
KRAS	G12D	c.35G>A	0.13%
KRAS	G13D	c.38G>A	0.00%
NRAS	Q61K	c.35G>A	0.13%
PIK3CA	E545K	c.35G>A	0.13%
PIK3CA	H1047R	c.35G>A	0.00%

实验步骤：

1. 分子标签的添加

先将 cfDNA 在 95℃ 热变性 5 分钟，然后迅速插冰上，然后再进行酶反应。通过末端转移酶（美国 NEB 公司的 Terminal Transferase，货号 M0315S）在 DNA 的 3'端加上 5-10 个随机碱基。

反应体系如下表 12 所示：

表 12

试剂	体积
末端转移酶缓冲液（10X）	5μl
二氯化钴（10mM）	5μl

dNTP混合物(50μM)	2μl
末端转移酶	0.5μl
DNA (10ng)	34.5μl
总共	50μl

37℃孵育30分钟,然后加入10μl浓度为0.5M的EDTA终止反应。加入1.8倍体积的Agencourt AMPure XP磁珠(美国贝克曼库尔特有限公司) 108μl,按照说明书进行纯化,纯化后用34μl蒸馏水溶解DNA。

2. 寡聚核苷酸尾巴的添加

反应体系如下表 13 所示:

表 13

试剂	体积
末端转移酶缓冲液 (10X)	5μl
二氯化钴 (10mM)	5μl
dCTP混合物(50μM)	5μl
末端转移酶	0.5μl
DNA	34.5μl
总共	50μl

37℃孵育30分钟,然后加入10μl浓度为0.5M的EDTA终止反应。加入1.8倍体积的Agencourt AMPure XP磁珠(美国贝克曼库尔特有限公司) 108μl,按照说明书进行纯化,纯化后用20μl蒸馏水溶解DNA。

3. 第一轮PCR扩增

PCR酶采用NEB公司的Q5® Hot Start High-Fidelity 2X Master Mix, 货号: M0494L。

反应体系如下表 14 所示:

表 14

试剂	体积
2XQ5聚合酶混合液	25μl
上游特异性引物池 (10μM)	2.5μl
通用引物1 (10μM)	2.5μl
上一步DNA	20μl
总共	50μl

上游特异性引物池如表15所示,通用引物1如表4所示。

表15

引物名称	引物序列(5-3')	引物编号
KRAS-G12D&13D-USP	CACCTGACCACTGTGTTCACCTT	SEQ ID NO: 42
EGFR-G719S-USP	TTGAGGATCTTGAAGGAACTGAAT	SEQ ID NO: 43
EGFR-E746-A750 &A751del-USP	GATCCCAGAAGGTGAGAAAGTTAA	SEQ ID NO: 44
EGFR-T790M-USP	TGCTGGGCATCTGCCTCA	SEQ ID NO: 45
EGFR-L858R-USP	GTGAAAACACCGCAGCATG	SEQ ID NO: 46
cKIT-D816V-USP	TCATGGTCGGATCACAAAGATT	SEQ ID NO: 47
NRAS-Q61K-USP	GCCTGTCCTCATGTATTGGTCT	SEQ ID NO: 48
BRAF-V600E-USP	GGATCCAGACAACCTGTTCAAACCTG	SEQ ID NO: 49
PIK3CA-H1047R-USP	GAGCAAGAGGCTTTGGAGTATTTC	SEQ ID NO: 50
PIK3CA-E545K-USP	CTCCATTTTAGCACTTACCTGTGAC	SEQ ID NO: 51

上游特异性引物池由表15所述的引物等摩尔数混合组成。

扩增体系如下表16所示:

表16

步骤1	98℃ 5min
步骤2	98℃ 10s
步骤3	62℃ 5min
步骤4	重复步骤2-3, 10个循环
步骤5	72℃ 5min

加入 1.8 倍体积的 Agencourt AMPure XP 磁珠（美国贝克曼库尔特有限公司）90 μ l，按照说明书进行纯化，纯化后用 20 μ l 蒸馏水溶解 DNA。

4. 第二轮 PCR 扩增：

PCR 反应体系如下表 17 所示：

表 17

试剂	体积
2XQ5聚合酶混合液	25 μ l
下游特异性引物池（10 μ M）	0.25 μ l
通用引物2（10 μ M）	2.5 μ l
测序标签引物(10 μ M)	2.5 μ l
上一步DNA	19.75 μ l
总共	50 μ l

下游特异性引物池如表18所示，通用引物2和测序标签引物如表4所示。

表18

引物编号	引物序列(5-3')	引物编号
KRAS-G12D&13D-DSP	AAGACCGCTTGGCCTCCGACTT CCTATTCCTTGGGAGGAAGCT	SEQ ID NO: 52
EGFR-G719S-DSP	AAGACCGCTTGGCCTCCGACTT TGAATTCAAAAAGATCAAAGTGCTG	SEQ ID NO: 53
EGFR-E746-A750 &A751del-DSP	AAGACCGCTTGGCCTCCGACTT GTTAAAATTCCCGTCGCTATCAA	SEQ ID NO: 54
EGFR-T790M-DSP	AAGACCGCTTGGCCTCCGACTT TCCACCGTGCAGCTCATCA	SEQ ID NO: 55
EGFR-L858R-DSP	AAGACCGCTTGGCCTCCGACTT AGCATGTCAAGATCACAGATTTTG	SEQ ID NO: 56
cKIT-D816V-DSP	AAGACCGCTTGGCCTCCGACTT TTTGTGATTTTGGTCTAGCCAGAG	SEQ ID NO: 57
NRAS-Q61K-DSP	AAGACCGCTTGGCCTCCGACTT TCTCTCATGGCACTGTACTCTTCTT	SEQ ID NO: 58
BRAF-V600E-DSP	AAGACCGCTTGGCCTCCGACTT GGACCCACTCCATCGAGATTT	SEQ ID NO: 59
PIK3CA-H1047R-DSP	AAGACCGCTTGGCCTCCGACTT ATTTTCATGAAACAAATGAATGATGC	SEQ ID NO: 60
PIK3CA-E545K-DSP	AAGACCGCTTGGCCTCCGACTT CTCCATAGAAAATCTTTCTCCTGC	SEQ ID NO: 61

下游特异性引物池由表18所示的引物等摩尔数混合组成。

扩增体系如下表19所示：

表19

步骤1	98℃ 5min
-----	----------

步骤2	98℃ 10s
步骤3	62℃ 5min
步骤4	98℃ 10s
步骤5	62℃ 2min
步骤6	重复步骤4-5, 25个循环
步骤7	72℃ 5min

加入 1 倍体积的 Agencourt AMPure XP 磁珠（美国贝克曼库尔特有限公司）50 μ l，按照说明书进行纯化，纯化后用 30 μ l 蒸馏水溶解 DNA。

5. 文库质检

采用安捷伦 2100 检测靶向富集文库，其结果如图 6 所示。

6. 上机测序

质检合格后，采用 BGISEQ-500 测序平台进行测序，双端 50bp 测序，得到的下机数据经过数据转换和质量过滤后采用如下信息分析。

7. 信息分析

首先将得到的数据去接头，得到双端测序结果，一端测序结果用来比对基因组（参考基因组 hg19），另一端结果去除连续的 G 碱基后，从去掉一端开始截取 10 个碱基序列作为分子标签，用来标记前面一端的序列信息；进行基本参数统计（表 20）比对到基因组上的数据比例，根据目标区域位置统计各目标区域的深度，得出目标区域均一性信息（图 7-8），其中，图 7 示出了原始数据测序深度，图 8 示出了读段（reads）去重后的测序深度。结果显示均一性良好。统计目标位点的深度及四碱基比例，通过分子标签去除重复和测序错误、PCR 错误，通过特定的信息分析算法得到目标位点的突变信息，结果如表 21 所示。表 22 示出了检测结果一致性的结果。图 9 示出了突变检测一致性信息，表明目标位点的检出的突变信息和预期一致。

表20：下机数据统计

原始数据 (reads)	比对到基因组数据 (reads)	目标区域数据 (reads)	比对率	捕获率
1231815	1177615	1161129	95.6%	98.6%

表21：检测结果

突变类型（目标位点）	碱基类型	原始深度	原始突变比例	UID 校正后的深度	校正后的突变比例
V600E (GRCh38: 7:140753336)	四种碱基总深度	71017	0.60%	2343	0.00%
	A 碱基深度	120		0	
	T 碱基深度	70591		2343	
	C 碱基深度	111		0	
	G 碱基深度	195		0	
D816V (GRCh38: 4:54733155)	四种碱基总深度	219690	0.14%	1731	0.00%
	A 碱基深度	265		1731	
	T 碱基深度	25		0	

	C 碱基深度	17		0	
	G 碱基深度	219383		0	
G719S(GRCh38: 7:55174014)	四种碱基总深度	169580	0.20%	1616	0.00%
	A 碱基深度	256		0	
	T 碱基深度	39		0	
	C 碱基深度	45		0	
	G 碱基深度	169240		1616	
T790M(GRCh38: 7:55181378)	四种碱基总深度	169480	0.11%	2383	0.08%
	A 碱基深度	111		0	
	T 碱基深度	36		2	
	C 碱基深度	47		2383	
	G 碱基深度	169286		0	
L858R(GRCh38: 7:55191822)	四种碱基总深度	53421	0.47%	1910	0.10%
	A 碱基深度	181		0	
	T 碱基深度	23		1908	
	C 碱基深度	53172		0	
	G 碱基深度	45		2	
Δ E746-A750 (GRCh38: 7:55174772)	四种碱基总深度	177785	0.30%	2674	0.15%
	A 碱基深度	352		4	
	T 碱基深度	128		0	
	C 碱基深度	54		0	
	G 碱基深度	177251		2670	
G12D(GRCh38: 12:25245350)	四种碱基总深度	37715	0.07%	2431	0.12%
	A 碱基深度	37690		3	
	T 碱基深度	5		0	
	C 碱基深度	7		0	
	G 碱基深度	13		2428	
G13D(GRCh38: 12:25245351)	四种碱基总深度	71017	0.30%	2239	0.00%
	A 碱基深度	7		0	
	T 碱基深度	70805		0	
	C 碱基深度	10		0	
	G 碱基深度	195		2239	
Q61K(GRCh38: 1:114713909)	四种碱基总深度	219690	0.14%	1795	0.11%
	A 碱基深度	265		2	
	T 碱基深度	25		0	
	C 碱基深度	17		1793	
	G 碱基深度	219383		0	
E545K(GRCh38:	四种碱基总深	169580	0.20%	2592	0.15%

3:179218303)	度				
	A 碱基深度	256		4	
	T 碱基深度	39		0	
	C 碱基深度	45		0	
	G 碱基深度	169240		2588	
H1047R(GRCh38: 3:179234297)	四种碱基总深度	169480	0.11%	1714	0.00%
	A 碱基深度	37690		1714	
	T 碱基深度	5		0	
	C 碱基深度	7		0	
	G 碱基深度	13		0	

表 22：检测结果一致性

突变位点	原始突变	校正后的突变	标准品突变
V600E	0.60%	0.00%	0.00%
D816V	0.14%	0.00%	0.00%
G719S	0.20%	0.00%	0.00%
T790M	0.11%	0.08%	0.10%
L858R	0.47%	0.10%	0.10%
Δ E746-A750	0.30%	0.15%	0.10%
G12D	0.07%	0.12%	0.13%
G13D	0.30%	0.00%	0.00%
Q61K	0.14%	0.11%	0.13%
E545K	0.20%	0.15%	0.13%
H1047R	0.11%	0.00%	0.00%

结论：在标准品中，V600E、D816V、G719S、G13D、H1047R 均为阴性，突变比例为 0.00%，本实施例得到的结果也是阴性，突变比例也为 0.00%；在标准品中，T790M、L858R、 Δ E746-A750、G12D、Q61K、E545K 为阳性，突变比例分别为 0.10%、0.10%、0.10%、0.13%、

0.13%、0.13%，我们方法检测到的突变比例分别为 0.08%、0.10%、0.10%、0.15%、0.12%、0.11%、0.15%，检测值和真实值在 $\pm 0.02\%$ 误差范围内，表明方法可以准确检测低至 0.10%的突变。可以看到通过分子标记（UID）校正可以去除重复和测序错误，将测序背景从 0.60%降到 0.00%（此处需要说明的是 V600E 校正前的错误值最大，意即该方法错误率可以从 0.60%降到 0.00%）。

由图 9 可以看出，目标位点检出突变信息与预期一致，阴性突变点都未检出，阳性突变点都检出，并且检出频率和预期值相差不大，可以认为该方法在血浆游离 DNA 中可以检测低至 0.10%的突变。

以上应用了具体个例对本发明进行阐述，只是用于帮助理解本发明，并不用以限制本发明。对于本发明所属技术领域的技术人员，依据本发明的思想，还可以做出若干简单推演、变形或替换。

权 利 要 求 书

1. 一种用于检测低频突变的靶向富集方法，其特征在于，包括：

在末端转移酶的作用下，在单链 DNA 和/或双链 DNA 每条链的 3'端加上一段单一碱基的单聚核苷酸尾巴；

以上游特异性引物和第一通用引物对突变目标位点所在的区域进行第一次 PCR 扩增，其中所述上游特异性引物以与其互补的目标序列为锚定位点，所述第一通用引物以所述单聚核苷酸尾巴为锚定位点，且所述第一通用引物包括 5'端的测序引物序列 1 和 3'端的与所述单聚核苷酸尾巴互补的连续单一碱基；

以下游特异性引物和第二通用引物对所述第一次 PCR 扩增的产物进行第二次 PCR 扩增，其中所述下游特异性引物位于所述上游特异性引物的下游且 5'端带有测序引物序列 2，所述第二通用引物包括所述第一通用引物 5'端的测序引物序列 1。

2. 根据权利要求 1 所述的方法，其特征在于，所述第二次 PCR 扩增中还加入测序标签引物，所述测序标签引物 3'端包括所述下游特异性引物 5'端的测序引物序列 2。

3. 根据权利要求 1 所述的方法，其特征在于，所述方法还包括：在加所述单聚核苷酸尾巴之前，在末端转移酶的作用下，在单链 DNA 和/或双链 DNA 每条链的 3'端加上一段随机碱基。

4. 根据权利要求 3 所述的方法，其特征在于，所述随机碱基的长度是 5-10 个碱基。

5. 根据权利要求 1 所述的方法，其特征在于，所述单聚核苷酸尾巴的长度是 15-30 个单一碱基。

6. 根据权利要求 1 所述的方法，其特征在于，所述上游特异性引物距离所述突变目标位点 25-150bp。

7. 根据权利要求 1 所述的方法，其特征在于，所述第一次 PCR 扩增进行 10-20 个循环。

8. 根据权利要求 1 所述的方法，其特征在于，所述连续单一碱基是 11-15 个连续的 C 或 G 碱基，或 25-35 个连续的 A 或 T 碱基。

9. 根据权利要求 1 所述的方法，其特征在于，所述第一通用引物在所述连续单一碱基之后还包括一个简并碱基。

10. 根据权利要求 9 所述的方法，其特征在于，所述简并碱基是 H、V、B 或 D。

11. 根据权利要求 1 所述的方法，其特征在于，所述第二次 PCR 扩增进行 15-30 个循环。

12. 根据权利要求 1 所述的方法，其特征在于，所述突变目标位点包括单核苷酸多态性、插入和删除以及拷贝数变异。

13. 一种用于检测低频突变的靶向富集试剂盒，其特征在于，包括：

末端转移酶和单一碱基核苷酸，用于在末端转移酶的作用下，在单链 DNA 和/或双链 DNA 每条链的 3'端加上一段单一碱基的单聚核苷酸尾巴；

上游特异性引物和第一通用引物，用于对突变目标位点所在的区域进行第一次 PCR 扩增，其中所述上游特异性引物以与其互补的目标序列为锚定位点，所述第一通用引物以所述单聚核苷酸尾巴为锚定位点，且所述第一通用引物包括 5'端的测序引物序列 1 和 3'端的与所述单聚核苷酸尾巴互补的连续单一碱基；

下游特异性引物和第二通用引物，用于对所述第一次 PCR 扩增的产物进行第二次 PCR 扩增，其中所述下游特异性引物位于所述上游特异性引物的下游且 5'端带有测序引物序列 2，所述第二通用引物包括所述第一通用引物 5'端的测序引物序列 1。

14. 根据权利要求 13 所述的试剂盒，其特征在于，所述试剂盒还包括测序标签引物，用于对所述第一次 PCR 扩增的产物进行第二次 PCR 扩增，所述测序标签引物 3'端包括所述下游特异性引物 5'端的测序引物序列 2。

15. 根据权利要求 13 所述的试剂盒，其特征在于，所述试剂盒还包括：混合核苷酸，用于在加所述单聚核苷酸尾巴之前，在末端转移酶的作用下，在单链 DNA 和/或双链 DNA 每条链的 3'端加上一段随机碱基。

16. 根据权利要求 15 所述的试剂盒，其特征在于，所述随机碱基的长度是 5-10 个碱基。

17. 根据权利要求 13 所述的试剂盒，其特征在于，所述单聚核苷酸尾巴的长度是 15-30 个单一碱基。

18. 根据权利要求 13 所述的试剂盒，其特征在于，所述上游特异性引物距离所述突变目标位点 25-150bp。

19. 根据权利要求 13 所述的试剂盒，其特征在于，所述第一次 PCR 扩增进行 10-20 个循环。

20. 根据权利要求 13 所述的试剂盒，其特征在于，所述连续单一碱基是 11-15 个连续的 C 或 G 碱基，或 25-35 个连续的 A 或 T 碱基。

21. 根据权利要求 13 所述的试剂盒，其特征在于，所述第一通用引物在所述连续单一碱基之后还包括一个简并碱基。

22. 根据权利要求 21 所述的试剂盒，其特征在于，所述简并碱基是 H、V、B 或 D。

23. 根据权利要求 13 所述的试剂盒，其特征在于，所述第二次 PCR 扩增进行 15-30 个循环。

24. 根据权利要求 13 所述的试剂盒，其特征在于，所述突变目标位点包括单核苷酸多态性、插入和删除以及拷贝数变异。

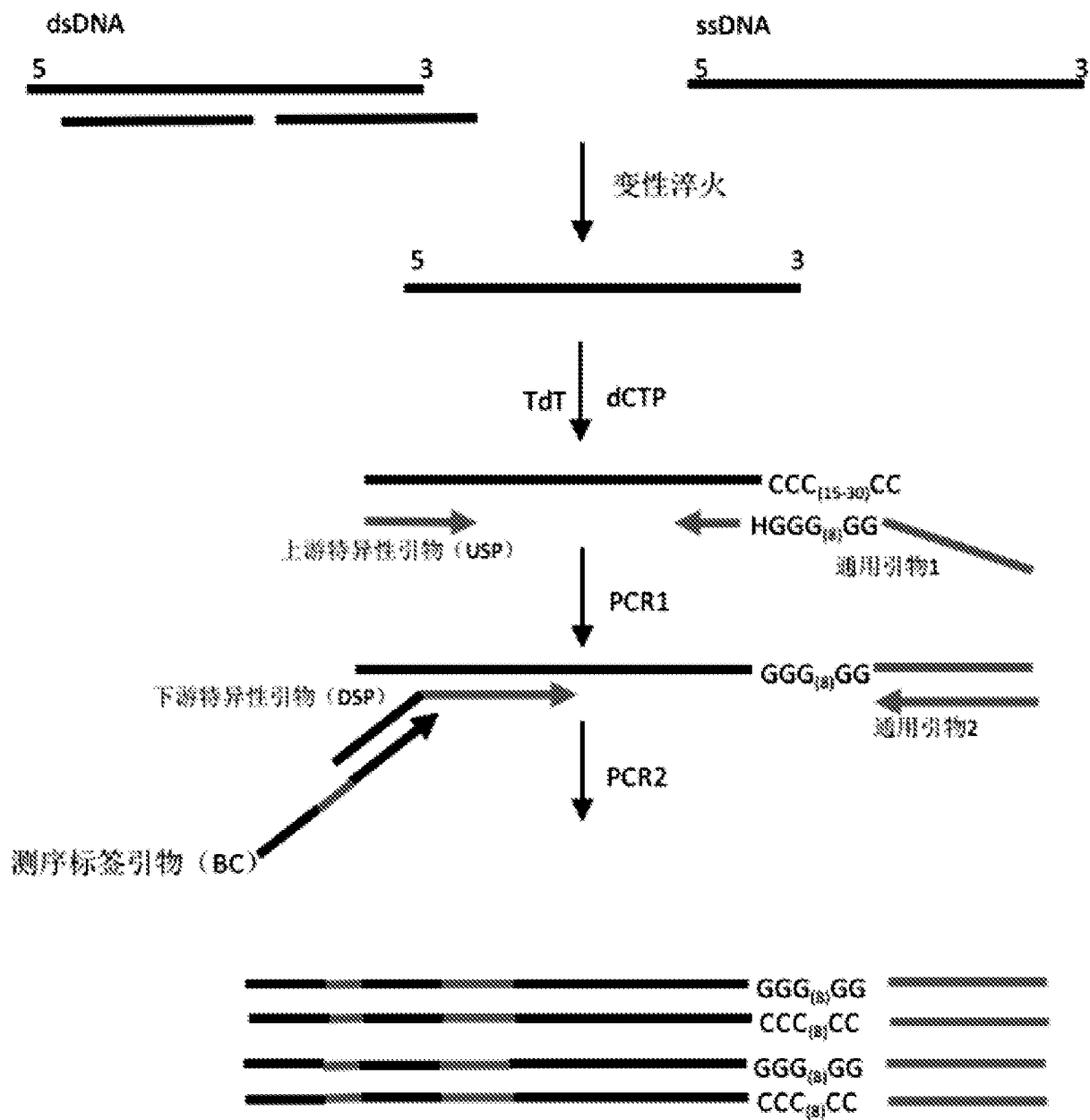


图 1

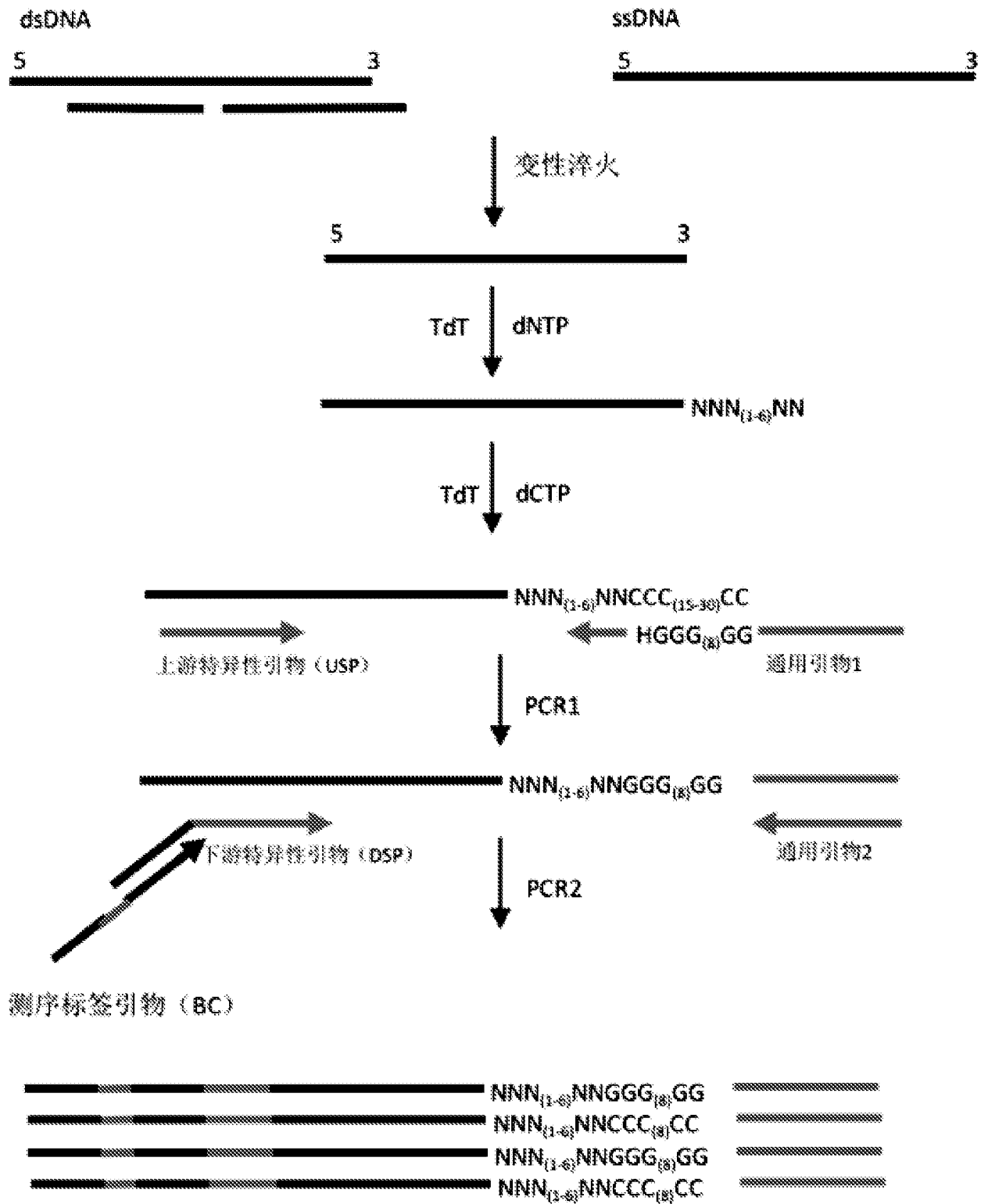


图 2

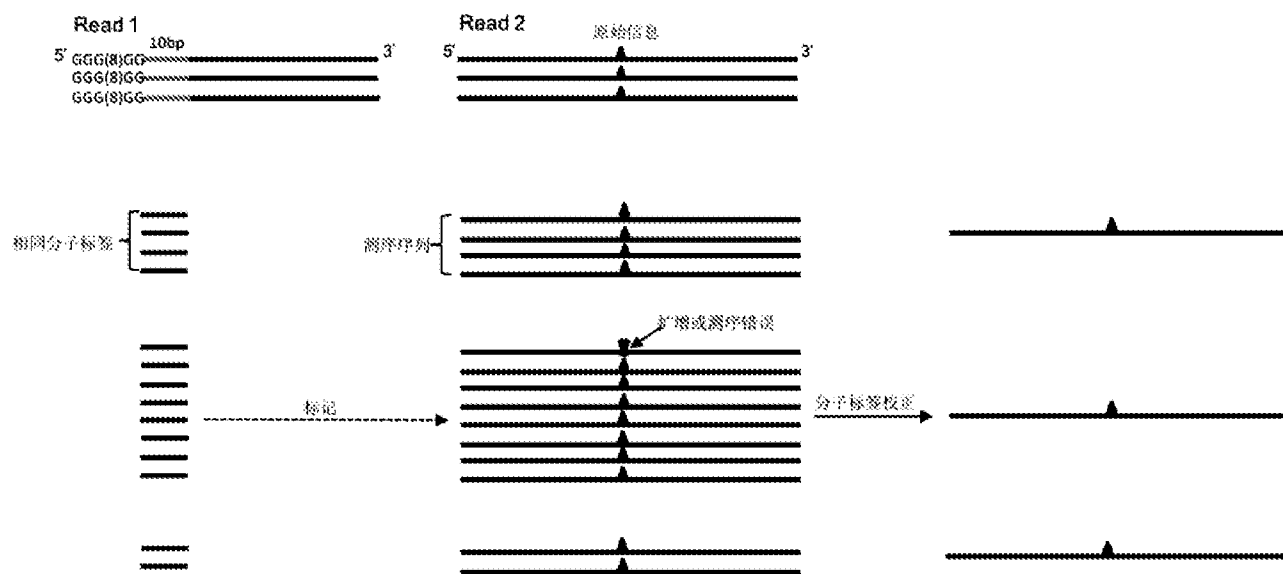


图 3

cfDNA文库

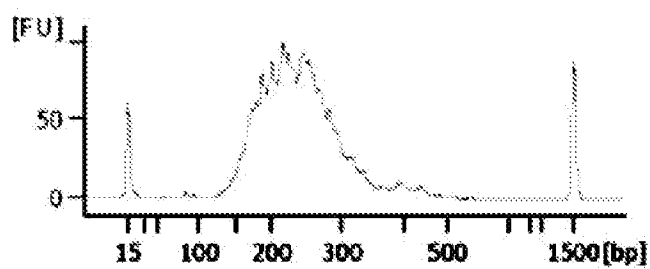


图 4

HBB各扩增子均一性

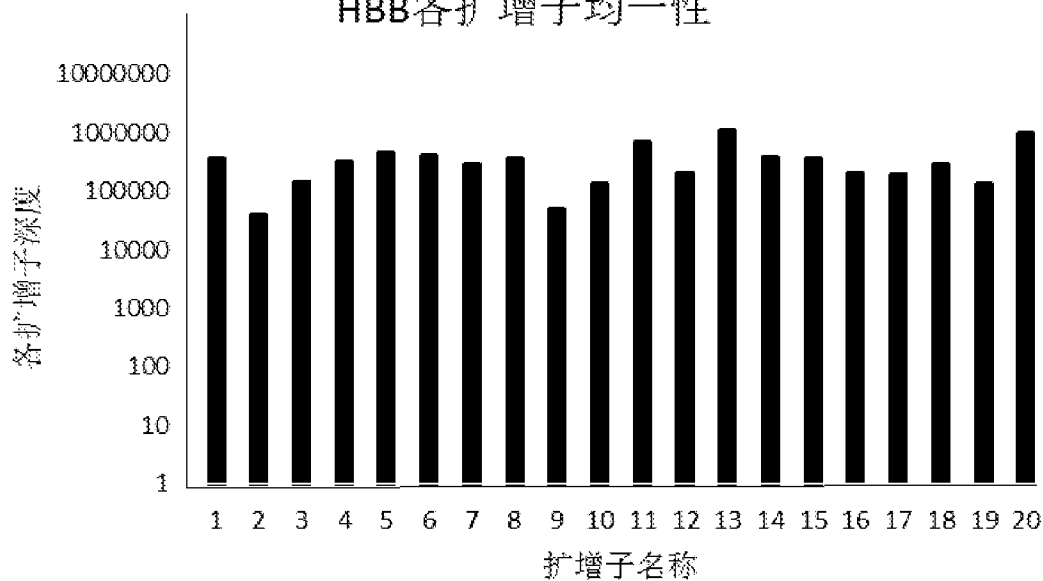


图 5

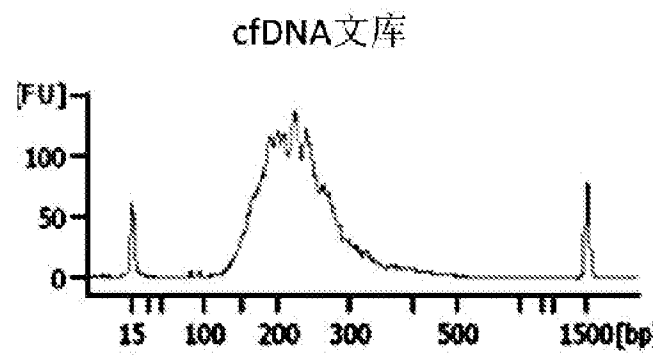


图 6

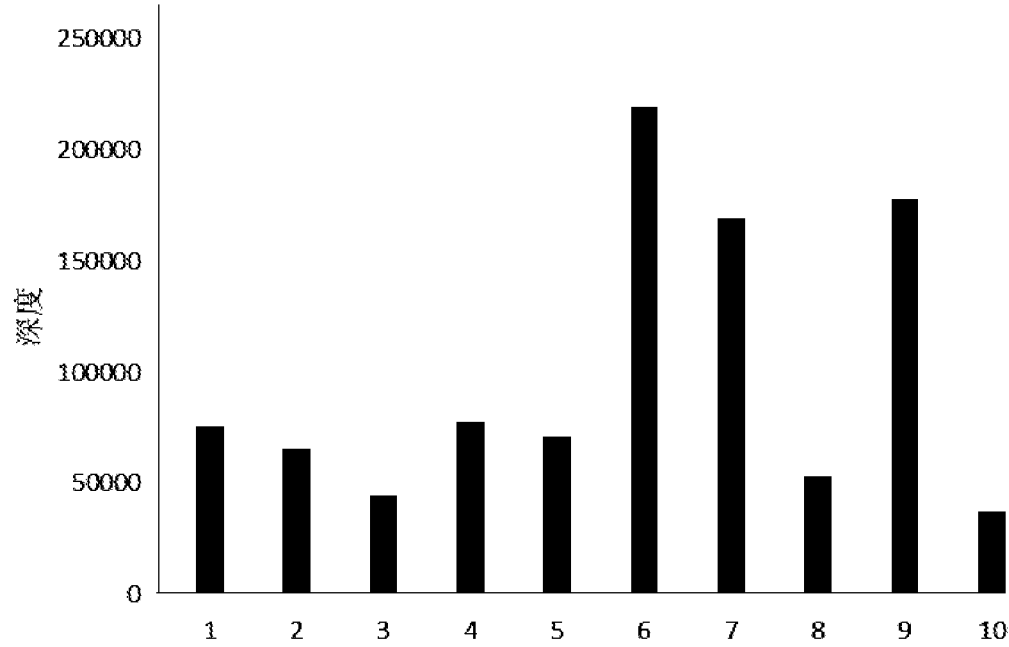


图 7

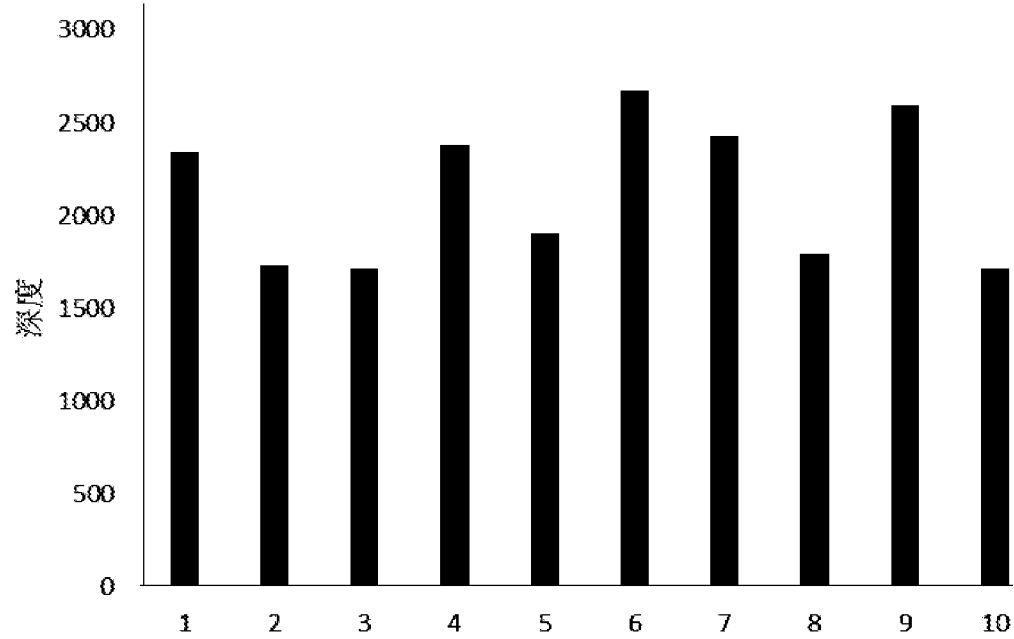


图 8

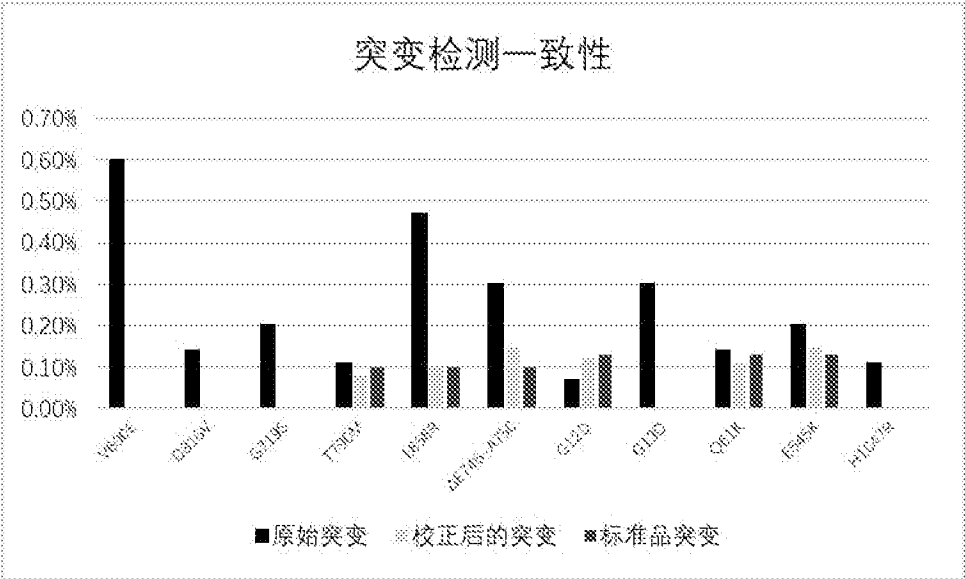


图9

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2017/093914

A. CLASSIFICATION OF SUBJECT MATTER

C12Q 1/68 (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)

CNMED; CPRSABS; CNABS; DWPI; SIPOABS; CNTXT; WOTXT; EPTXT; JPTXT; CNKI; Pubmed; Google scholar: 测序, 文库, 引物, 探针, 靶向富集, 多聚, 多重, 扩增, 尾巴, 尾, 加尾, 低频突变, 巢式, 锚定, 接头, 特异, PCR, enrich+, amplif+, multiplexed, multiplex, poly dT, tail, tailing, primer, poly dA, anchored multiplex PCR, nested, probe, search for sequence SEQ ID NOs: 1-61

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	CN 106192018 A (BGI-SHENZHEN), 07 December 2016 (07.12.2016), see abstract	1-24
Y	ZHENG, Z.L. et al., "Anchored Multiplex PCR for Targeted Next-Generation Sequencing", Nature Medicine, 20(12), 31 December 2014 (31.12.2014), pages 1479-1486, see figure 1	1-24
Y	US 2016115532 A1 (SEQUENTA INC.), 28 April 2016 (28.04.2016), see description, paragraphs [0025]-[0026], and figure 2	1-24
Y	VARDI, O. et al., "Biases in the SMART-DNA Library Preparation Method Associated with Genomic Poly dA/dT Sequences", PLoS ONE, 12(2), 24 February 2017 (24.02.2017), article e0172769, see figure 1, and page 2, paragraph 2	1-24

☐ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents:	"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
"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	"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
"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)	"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
"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	"&" document member of the same patent family

Date of the actual completion of the international search 19 April 2018	Date of mailing of the international search report 04 May 2018
Name and mailing address of the ISA State Intellectual Property Office of the P. R. China No. 6, Xitucheng Road, Jimengqiao Haidian District, Beijing 100088, China Facsimile No. (86-10) 62019451	Authorized officer LI, Kangqi Telephone No. (86-10) 62411034

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2017/093914

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 106192018 A	07 December 2016	None	
US 2016115532 A1	28 April 2016	JP 2015524282 A	24 August 2015
		WO 2014026031 A1	13 February 2014
		EP 2882870 A1	17 June 2015
		EP 2882870 A4	16 March 2016

国际检索报告

国际申请号

PCT/CN2017/093914

A. 主题的分类 C12Q 1/68(2018.01) i 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. 检索领域 检索的最低限度文献(标明分类系统和分类号) C12Q 包含在检索领域中的除最低限度文献以外的检索文献 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)) CNMED;CPRSABS;CNABS;DWPI;SIPOABS;CNTXT;WOTXT;EPTXT;JPTXT;CNKI;Pubmed;Google scholar: 测序, 文库, 引物, 探针, 靶向富集, 多聚, 多重, 扩增, 尾巴, 尾, 加尾, 低频突变, 巢式, 锚定, 接头, 特异, PCR, enrich+, amplif+, multiplexed, multiplex, poly dT, tail, tailing, primer, poly dA, anchored multiplex PCR, nested, probe, 基于序列SEQ ID NO:1-61的检索																	
C. 相关文件 <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>Y</td> <td>CN 106192018 A (深圳华大基因研究院) 2016年 12月 7日 (2016 - 12 - 07) 参见摘要</td> <td>1-24</td> </tr> <tr> <td>Y</td> <td>ZHENG ZL. 等. "Anchored multiplex PCR for targeted next-generation sequencing" Nature Medicine, 第20卷, 第12期, 2014年 12月 31日 (2014 - 12 - 31), 第1479-1486页 参见图1</td> <td>1-24</td> </tr> <tr> <td>Y</td> <td>US 2016115532 A1 (SEQUENTA INC) 2016年 4月 28日 (2016 - 04 - 28) 参见说明书第[0025]-[0026]段, 图2</td> <td>1-24</td> </tr> <tr> <td>Y</td> <td>VARDI O. 等. "Biases in the SMART-DNA library preparation method associated with genomic poly dA/dT sequences" PLoS ONE, 第12卷, 第2期, 2017年 2月 24日 (2017 - 02 - 24), 第e0172769篇 参见图1, 第2页第2段</td> <td>1-24</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	Y	CN 106192018 A (深圳华大基因研究院) 2016年 12月 7日 (2016 - 12 - 07) 参见摘要	1-24	Y	ZHENG ZL. 等. "Anchored multiplex PCR for targeted next-generation sequencing" Nature Medicine, 第20卷, 第12期, 2014年 12月 31日 (2014 - 12 - 31), 第1479-1486页 参见图1	1-24	Y	US 2016115532 A1 (SEQUENTA INC) 2016年 4月 28日 (2016 - 04 - 28) 参见说明书第[0025]-[0026]段, 图2	1-24	Y	VARDI O. 等. "Biases in the SMART-DNA library preparation method associated with genomic poly dA/dT sequences" PLoS ONE, 第12卷, 第2期, 2017年 2月 24日 (2017 - 02 - 24), 第e0172769篇 参见图1, 第2页第2段	1-24
类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求															
Y	CN 106192018 A (深圳华大基因研究院) 2016年 12月 7日 (2016 - 12 - 07) 参见摘要	1-24															
Y	ZHENG ZL. 等. "Anchored multiplex PCR for targeted next-generation sequencing" Nature Medicine, 第20卷, 第12期, 2014年 12月 31日 (2014 - 12 - 31), 第1479-1486页 参见图1	1-24															
Y	US 2016115532 A1 (SEQUENTA INC) 2016年 4月 28日 (2016 - 04 - 28) 参见说明书第[0025]-[0026]段, 图2	1-24															
Y	VARDI O. 等. "Biases in the SMART-DNA library preparation method associated with genomic poly dA/dT sequences" PLoS ONE, 第12卷, 第2期, 2017年 2月 24日 (2017 - 02 - 24), 第e0172769篇 参见图1, 第2页第2段	1-24															
☐ 其余文件在C栏的续页中列出。 ☒ 见同族专利附件。																	
* 引用文件的具体类型: "A" 认为不特别相关的表示了现有技术一般状态的文件 "E" 在国际申请日的当天或之后公布的在先申请或专利 "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) "O" 涉及口头公开、使用、展览或其他方式公开的文件 "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 "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 "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 "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 "&" 同族专利的文件																	
国际检索实际完成的日期 2018年 4月 19日		国际检索报告邮寄日期 2018年 5月 4日															
ISA/CN的名称和邮寄地址 中华人民共和国国家知识产权局 (ISA/CN) 中国北京市海淀区蓟门桥西土城路6号 100088 传真号 (86-10)62019451		受权官员 李康琦 电话号码 62411034															

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/CN2017/093914

检索报告引用的专利文件			公布日 (年/月/日)	同族专利			公布日 (年/月/日)
CN	106192018	A	2016年 12月 7日	无			
US	2016115532	A1	2016年 4月 28日	JP	2015524282	A	2015年 8月 24日
				WO	2014026031	A1	2014年 2月 13日
				EP	2882870	A1	2015年 6月 17日
				EP	2882870	A4	2016年 3月 16日