

(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
2021 年 3 月 25 日 (25.03.2021)



(10) 国际公布号
WO 2021/051665 A1

- (51) 国际专利分类号:
C12Q 1/6806 (2018.01)
- (21) 国际申请号: PCT/CN2019/123892
- (22) 国际申请日: 2019 年 12 月 9 日 (09.12.2019)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201910896457.8 2019 年 9 月 20 日 (20.09.2019) CN
- (71) 申请人: 上海臻迪基因科技有限公司
(SHANGHAI ZENISIGHT LTD.) [CN/CN]; 中国上海市浦东新区高科东路 777 弄 10 号 2613 室, Shanghai 201210 (CN)。
- (72) 发明人: 郭志伟 (GUO, Zhiwei); 中国上海市浦东新区高科东路 777 弄 10 号 2613 室, Shanghai 201210 (CN)。 李英辉 (LI, Yinghui); 中国上海

市浦东新区高科东路 777 弄 10 号 2613 室, Shanghai 201210 (CN)。 陈倩 (CHEN, Qian); 中国上海市浦东新区高科东路 777 弄 10 号 2613 室, Shanghai 201210 (CN)。 胡荣君 (HU, Rongjun); 中国上海市浦东新区高科东路 777 弄 10 号 2613 室, Shanghai 201210 (CN)。

(74) 代理人: 上海德禾翰通律师事务所 (DEHEHANTONG LAW OFFICE); 中国上海市徐汇区宜山路 425 号光启城 24 层, Shanghai 200233 (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,

(54) Title: ENRICHMENT METHOD AND SYSTEM FOR GENE TARGET REGION

(54) 发明名称: 基因目标区域的富集方法及体系

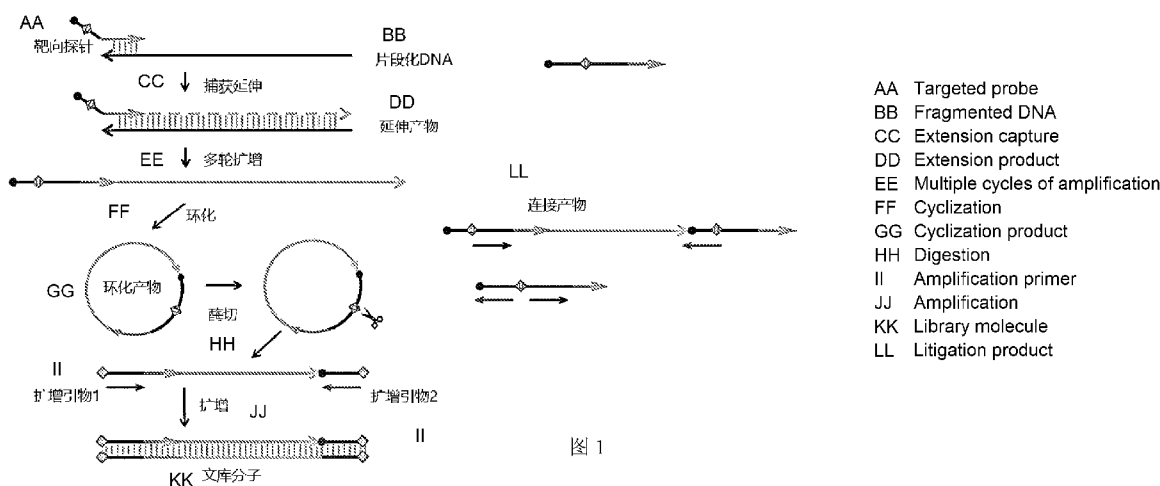


图 1

(57) Abstract: The present invention provides an enrichment method for a gene target region, the method comprising: (1) amplifying, by means of a specific probe, fragmented DNA including a target region, and providing a captured extension product, wherein the specific probe comprises a sequence that is complementary to the target region of the fragmented DNA, and a 3'-end nucleotide and a 5'-end nucleotide of the probe are both modified; and (2) adding a ligase to the captured extension product provided in step (1), and providing a ligation product. The invention further provides an enrichment system for a gene target, the system being applicable to the enrichment method for a gene target region provided in the present invention.

(57) 摘要: 本发明提供了一种基因目标区域的富集方法, 包括: (1) 通过特异性探针扩增包含目标区域的片段化DNA, 以提供捕获延伸产物, 所述的特异性探针包括与所述片段化DNA的目标区域互补的序列, 且其3'末端和5'末端核苷酸均被修饰; (2) 向步骤(1)所提供的捕获延伸产物中加入连接酶, 以提供连接产物。本发明还提供了一种基因目标区域的富集体系, 适用于本发明提供的基因目标区域的富集方法。

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))。
- 包括经修改的权利要求 (条约第19条(1))。
- 包括说明书序列表部分 (细则5.2(a))。

基因目标区域的富集方法及体系

技术领域

本发明涉及生物技术领域，特别是涉及一种基因目标区域的富集方法及体系。

背景技术

基因测序技术自上世纪七十年代问世以来已经历了近半个世纪，1985年PCR技术的横空出世推动了整个分子生物学领域的发展。下一代测序技术（NGS）具有准确、灵敏、通量高的优点，随着测序成本的不断降低，其应用范围不断扩大，但其应用也受到了需求多样化且费时费力的文库构建这一步骤的制约。在对临床样本比如血浆样本的建库过程中，DNA的存在形式通常是短片段的，受损的，单链或部分双链的，对于这些存在形式的尤其是片段小于200bp的DNA来说，现有的PCR技术并不能做到很好的捕获和富集。

对于微小DNA片段的富集，现有技术仍主要使用传统PCR建库方法，或先加接头连接再扩增比如杂交捕获法。但对于前者来说，由于需要双端引物，极大限制了适合扩增的片段的长度，且存在扩增中的偏好性导致产物的高不均一性，以及指数扩增累积的错误导致后续测序结果不准确的问题；而对于后者，虽然对待富集片段长度的要求没有PCR高，但需要先进行连接反应，而连接的效率通常只有20%~50%，导致捕获效率低下，还存在因连接困难容易丢失稀有分子的问题。为了解决NGS的建库问题，近期发展起来的技术还包括诸如分子倒置探针、多重PCR等。分子倒置探针与杂交捕获技术相比，特异性较好，但其口袋状探针设计复杂，也不适用于微小DNA片段的富集。多重PCR技术适合大规模样本，应用最为广泛，但要么引物设计要求极高且扩增子均一性差，要么扩增产物均一性较好但对起始样本浓度要求很高，同样不适用于起始浓度低的微小DNA片段的富集。这些现有技术构建文库时通常需要双端引物，为了去除接头二聚体污染必然引入纯化步骤，这导致小片段的双链DNA、受损伤的双链DNA和单链DNA分子的信息丢失。然而，在一些转录活跃的基因组区域恰恰是这些存在形式的DNA。综上所述，现有技术对片段化的，尤其是200bp以下的微小DNA进行富集的需求严重未能满足，而找到高效富集片段化DNA目标区域的方法，突破连接效率对富集效果的瓶颈，抑制非目的连接产物的产生，最大程度的捕获稀有分子并保持产物的均一性，是本发明要解决的主要技术问题。申请人在申请号为2019100024085的申请中，已先行确立了先用特异性探针线性扩增目标区域再连接接头实现富集的技术方案，主要的应用方向是基于二代测序的核酸检测。本发明将用另一思路解决片段化DNA靶向富集的问题。

发明内容

鉴于以上所述现有技术的缺点，本发明的一个目的在于提供一种基因目标区域的富集方

法，用于解决现有技术中的问题。

为实现上述目的及其他相关目的，本发明首先提供一种基因目标区域的富集方法，包括：

(1) 通过特异性探针扩增包含目标区域的片段化 DNA，以提供捕获延伸产物；

所述特异性探针包括与所述片段化 DNA 的目标区域互补和不互补的序列，且所述特异性探针的 3'末端和 5'末端核苷酸均被修饰；

(2) 向步骤(1)所提供的捕获延伸产物中加入连接酶，以提供连接产物，所述连接产物包括环化连接产物和线性连接产物。

本发明另一个目的在于提供一种用于富集片段化 DNA 目标区域的体系，包括适用于本发明提供的基因目标区域富集方法的特异性探针和连接酶。

附图说明

图 1 是本发明实施例中对于目标区域富集方法的流程示意图。

图 2 是本发明实施例中构建的分子的示意图。

具体实施方式

本发明发明人经过大量探索性研究，提供了一种基因目标区域的富集方法，所述基因目标区域的富集方法操作简单、结果可靠，尤其对于短片段核酸具有良好的富集效果，在此基础上完成了本发明。

本发明第一方面提供一种基因目标区域的富集方法，包括：

(1) 通过特异性探针扩增包含目标区域的片段化 DNA，以提供捕获延伸产物，所述特异性探针既包括与所述片段化 DNA 的目标区域互补也包括与所述目标区域不互补的序列，且所述特异性探针的 3'末端和 5'末端核苷酸均被修饰，用于阻止所述特异性探针结合模板前其 3'末端发生连接反应，主要是和自身 5'末端发生的自连反应；

(2) 将步骤(1)所提供的捕获延伸产物中加入连接酶，以提供连接产物。这一步的连接产物包含两部分，其一是捕获延伸产物分子内连接形成的环化连接产物，其二是捕获延伸产物与所述特异性探针分子间连接形成的线性连接产物。不论是环化连接产物或是线性连接产物都包含需要扩增的目标区域，均适用于目标分子的文库构建。

本发明所提供的基因目标区富集方法中，通过特异性探针扩增包含目标区域的片段化 DNA，包含目标区域的片段化 DNA 可以是一个，也可以是多个。通常来说，特异性探针与包含目标区域的片段化 DNA 是一一对应的，即反应体系中特异性探针的数量可以是一种，也可以多种。此步骤所述的延伸指的是对样本完整的线性预扩增步骤，包括单轮或多轮变性、退火、延伸步骤的循环。优选的实施例中，此步骤实施多轮循环，比如 2~100、2~10、10~20、

20~30、30~40、40~60、60~80、80~100 个循环，有效地增加包含目标区域的分子数。

所述基因目标区域的富集方法中，所述片段化 DNA 可以是双链 DNA、单链 DNA 和 cDNA 等，所述 cDNA 通常可以由 RNA 反转录获得。对于双链 DNA 来说，特异性探针可以包括与所述片段化 DNA 其中一链的目标区域互补的序列，也包括。所以本发明的富集方法也适用于片段化 RNA，本领域技术人员可以将 RNA 反转录成 cDNA 即可通过本发明所提供的富集方法进行后续操作。所述片段化 DNA 的长度可以是 25~200bps、25~40bps、40~60bps、60~80bps、80~100bps、100~120bps、120~140bps、140~160bps、160~180bps、或 180~200bps。

所述基因目标区域的富集方法中，所述步骤(1)的扩增体系中可以包括特异性探针、DNA 聚合酶和 dNTP。通过特异性探针扩增包含目标区域的片段化 DNA 的反应通常可以在 DNA 聚合酶存在的条件下进行。3'端封闭修饰的探针在高保真聚合酶的作用下结合模板后，封闭基团被切除，探针被活化，从而可以对目标序列进行有效延伸。所述 DNA 聚合酶可以具有 3'-5'外切酶活性，从而可以切除结合模板后的探针 3'端的取代基团，使探针能够沿模板延伸，优选可以为高保真 DNA 聚合酶，用于进一步提高产物的扩增效率和纯度。所述 DNA 聚合酶也可以是普通的 DNA 聚合酶，即只有聚合酶活性，没有 3'-5'外切酶活性。这种情况下，所述步骤(1)的扩增体系中还包括活性物质，所述活性物质可以用于切除结合目标区域后所述特异性探针的 3'末端修饰基团，故可以在捕获延伸体系中与 DNA 聚合酶（例如，普通的 DNA 聚合酶）组合，以提升线性扩增体系的效率，所述活性物质优选为核酸酶。通过特异性探针扩增包含目标区域的片段化 DNA 的反应通常可以在 dNTP 存在的条件下进行，所述 dNTP 通常可以是耦联标记分子的 dNTP，所述耦联标记分子可以是包括但不限于生物素等，所述 dNTP 可以是包括但不限于 dCTP、dATP 等。所述 dNTP 还可以耦联有标记分子，所述标记分子可以是生物素等，所述标记分子通常可以用于捕获延伸产物的纯化。

所述基因目标区域的富集方法中，所述的特异性探针包括与所述片段化 DNA 的目标区域互补的序列，从而可以实现对片段化 DNA 的目标区域的特异性扩增，本领域技术人员可选择合适的片段化 DNA 的目标区域，并根据片段化 DNA 的目标区域，设计合适的互补的序列，互补的序列通常靠近探针 3'末端。所述的特异性探针通常还包括与所述片段化 DNA 不互补的序列，从而可以实现酶切位点和通用序列的引入，不互补的序列通常靠近探针 5'末端。所述特异性探针可以是 3'末端核苷酸被修饰的特异性探针，用于阻止所述特异性探针的 3'末端与其他基团发生连接反应，从而可以避免游离探针的自连或连接其它非目的分子。本领域技术人员可选择合适的修饰基团以实现特异性探针 3'末端的修饰，例如，所修饰的基团可以取代特异性探针 3'末端核苷酸上的天然基团（例如，羟基、甲氧基等），以阻止所述特异性探

针 3'端发生连接反应,所使用的修饰基团通常可以是封闭基团。探针通过互补序列与模板上的目标区域结合后,3'端的修饰基团可以被酶切除,使得探针被活化,从而可以对目标序列进行有效延伸。特异性探针的 3'末端修饰基团可以是包括但不限于氢原子、C3 Spacer 基团、C6 Spacer 基团、磷酸基团(PO_4)、氨基(NH_2)等。不同取代基团的选择对于探针的捕获效果有明显差异,在本发明一优选实施例中,将探针 3'端羟基取代为 C3 Spacer 得到的效果最佳,与其它取代基团相比具有明显优势。所述特异性探针还包括通用序列,所述通用序列通常能被测序系统识别,从而可以将后续提供的连接产物通过测序系统进行测序,例如,对于 Ion Torrent 测序系统,所述通用序列可以是对应的 P1 序列和 A 序列的反向互补序列。这种能被测序系统识别的序列的应用,使本发明完成建库后的文库能够通过高通量测序平台进行测序,以提供各种后续研究和临床应用所需要的信息。在本发明一优选实施方式中,所述特异性探针 3'端尾部区域的碱基可以包含错配,错配的可以是探针 3'端最后一个碱基,也可以是接近 3'端的碱基;错配的可以是一个碱基,也可以是复数个。该错配不仅不影响探针的特异性和结合效率,还更有利于提高高保真 DNA 聚合酶的切割效率和保真度。

所述基因目标区域的富集方法中,所述步骤(1)中,还可以包括纯化捕获延伸产物。本领域技术人员可选择合适的方法对捕获延伸产物进行纯化,例如,捕获延伸产物的纯化方法可以是包括硅胶柱纯化、加热处理、磁珠纯化等。在本发明一具体实施例中,可以针对 dNTP 耦联的标记分子对捕获延伸产物进行纯化,纯化过程可以使用亲和素或者链霉亲和素包被的磁珠纯化等。

本发明所提供的基因目标区富集方法中,还可以包括:向步骤(1)所提供的捕获延伸产物中加入连接酶,以提供连接产物。所述连接酶是单链连接酶,优选为 T4 RNA 连接酶或热稳定性 RNA 连接酶等。所述连接产物可以是包括捕获延伸产物分子内连接形成的环化连接产物,也可以是捕获延伸产物与特异性探针分子间连接形成的线性连接产物。

所述基因目标区域的富集方法中,所述特异性探针可以是 5'末端核苷酸被修饰的、且在步骤(2)的反应温度下为单链结构,从而可以在单链连接酶的催化下,通过捕获延伸产物的 3'端羟基与其自身或特异性探针的 5'末端的修饰基团生成共价键,得到连接产物。在本发明一具体实施方式中,特异性探针的 5'末端核苷酸(例如,具体可以为的 5'末端核苷酸的 5 位羟基)被磷酸基团取代,从而形成磷酸化修饰,在单链连接酶的催化下,捕获延伸产物的 3'端羟基与与磷酸化的 5'末端生成共价键,得到连接产物。此时特异性探针有两种状态,一种是结合上模板形成了捕获延伸产物的具有磷酸化的 5'末端的特异性探针,另一种游离的特异性探针。所以对应的连接反应也可以是两种情况,既可以是捕获延伸产物的 3'端羟基连接自

身的 5'端磷酸基团，分子内首尾连接成环；也可以是捕获延伸产物的 3'端羟基连接游离探针的 5'端磷酸基团，连接产物仍然是线性产物。在本发明另一具体实施方式中，特异性探针的 5'末端核苷酸（例如，具体可以为的 5'末端核苷酸的 5 位羟基）被腺苷基团取代，从而形成腺苷酸化修饰，在 5'App DNA/RNA 热稳定连接酶的催化下，捕获延伸产物也可以分子内环化连接，或者游离探针的 5'末端腺苷基团可以与捕获延伸产物的 3'端相连接。在本发明另一具体实施方式中，特异性探针的 5'末端核苷酸（例如，具体可以为 5'末端核苷酸的 5 位羟基）被磷酸基团取代，从而形成磷酸化修饰，在热稳定 RNA 连接酶的催化下，特异性探针的 5'末端可以与捕获延伸产物的 3'端发生连接反应，也产生分子内环化的和分子间线性的两种连接产物。所述特异性探针也可以是 5'端区域具有黏性末端的部分双链结构，其 5'端区域的黏性末端具有单链性质，从而可以通过如上所述的方法对 5'端进行修饰，在合适的单链连接酶催化下，与捕获延伸产物连接。所述基因目标区域的富集方法中，所述步骤（2）还可以包括纯化连接产物。本领域技术人员可选择合适的方法对连接产物进行纯化，例如，连接产物的纯化方法可以是包括但不限于硅胶柱纯化、加热处理、磁珠纯化等。

本发明所提供的基因目标区域的富集方法，还可以包括：（3）在所述特异性探针与所述目标区域不互补的序列中，通常可以是靠近 5'端的序列中设置酶切位点，并在所述步骤（2）之后加入内切酶，用于在所述酶切位点切割所述步骤（2）形成的连接产物中的环化连接产物，使环化连接产物转变为线性连接产物，有利于后续的 PCR 扩增和 PCR 检测反应的进行。在本发明一具体实施方式中，所述酶切位点是替换了原胸腺嘧啶 T 的尿嘧啶 U，所述内切酶是 USER 酶或 SSDNA 内切酶。在本发明另一具体实施例中，特异性探针 5'端附近有酶切位点 U，当捕获延伸产物首尾相连形成环化产物后，USER 酶或 SSDNA 内切酶在酶切位点 U 切开环化产物使之成为线性连接产物，后续可以使用 PCR 扩增引物通过互补序列结合上酶切后的线性连接产物实现指数扩增。

本发明所提供的基因目标区域的富集方法，还可以包括：（4）PCR 扩增步骤（2）提供的连接产物。所述 PCR 扩增引物具有与所述特异性探针互补的序列，所述互补的序列与所述目标区域的序列不互补。本发明所提供的基因目标区域的富集方法，还可以包括：（5）在所述步骤（3）后，PCR 扩增所述步骤（2）和（3）提供的连接产物，所述 PCR 扩增引物具有与所述酶切位点两侧的序列互补的序列；优选的，所述酶切位点两侧的序列为测序通用序列通常可以通过 DNA 聚合酶和 PCR 扩增引物对连接产物进行 PCR 扩增，通过扩增连接产物，可以进一步富集含有目标区域的 DNA 的产物。本领域技术人员可选择合适的方法和体系对步骤（2）或步骤（3）所提供的连接产物进行扩增，例如，在本发明一实施例中，PCR 扩增

引物具有与所述特异性探针 5'端附近未结合模板的序列中互补和反向互补的序列，可以将或步骤（2）产生的环化和线性连接产物一起扩增。在本发明另一实施例中，特异性探针 5'端附近有酶切位点 U，PCR 扩增的前后引物分别与特异性探针上的酶切位点 U 两侧的序列互补和反向互补。当 USER 酶或 ssDNA 内切酶在酶切位点 U 切环化连接产物使之成为线性连接产物后，PCR 扩增引物可通过互补序列结合上酶切后的线性连接产物的两端实现指数扩增，如图 1 所示。所述基因目标区域的富集方法中，所述步骤（4）或（5）还可以包括纯化扩增产物。本领域技术人员可选择合适的方法对扩增产物进行纯化，例如可以是包括但不限于硅胶柱纯化、加热处理、磁珠纯化等。

本发明所提供的基因目标区域富集方法，还可以包括：（6）用检测引物 1、检测引物 2 和探针 3 检测步骤（4）或（5）提供的连接产物，以非二代测序而是 PCR 检测的手段快速提供目标区域的检测结果。所述的检测引物 1、检测引物 2 和探针 3 中，至少其一包含基因特异性序列，也即是说，对于不同的基因目标区域，三种引物/探针有关特异性序列的组合可以是单特异性、双特异性或三特异性。在本发明一个实施例中，所述检测引物 1 包含基因特异性序列，检测引物 2 和探针 3 包含通用序列，或检测引物 1 和检测引物 2 包含基因特异性序列；或检测引物 1 和探针 3 包含基因特异性序列，或检测引物 1、检测引物 2 和探针 3 均包含基因特异性序列。探针 3 还可以包含标记分子，比如可以是荧光分子，且探针 3 的序列与检测引物 1 或 2 不互补。

本发明所提供的基因目标区域的富集方法，可以用于核酸检测。通过扩增后的连接产物进一步进行检测的方法对于本领域技术人员来说是已知的。本发明目标区域富集的方法可以应用于基于 PCR 的基因序列检测中，例如，所述的目标区域包括：序列发生变异的位点，更具体可以是单碱基突变位点区域，碱基缺失位点区域，碱基插入位点区域，融合突变位点区域，表观遗传变异或基因特异性序列等。在本发明一实施例中，本发明所提供的基因目标区域的富集方法应用于 EGFR SNP-Q787 位点突变和男性特有 SRY 基因的检测，取得了理想的效果。

本发明第二方面提供一种用于富集片段化 DNA 目标区域的体系，包括适用于本发明第一方面所提供的基因目标区域的富集方法的特异性探针和连接酶。其中，连接酶可以是热稳定性 RNA 连接酶、T4 RNA 连接酶、或 5'App DNA/RNA 热稳定连接酶等。所述特异性探针的结构在本发明第一方面已经详细描述，在此不作赘述。

本发明所提供的体系中，还可以包括以下组分的一种或多种：耦联标记分子的 dNTP、DNA 聚合酶、核酸酶、内切酶等。其中，dTNP 可以是 dCTP，也可以是 dATP。耦联的标记分子可以是生物素。所述 DNA 聚合酶可以是具有 3'-5'外切酶活性的 DNA 聚合酶，优选高保

真 DNA 聚合酶。所述核酸酶具有 3'-5' 外切酶活性。所述内切酶可以是 USER 酶或者 ssDNA 内切酶。

本发明所提供的体系中，还可以包括 PCR 扩增引物，其序列通常与特异性探针的通用序列和环化酶切位点两侧的序列相配合，具体的可以是与位于环化酶切位点两侧的序列互补，并与所述特异性探针的通用序列至少部分互补的序列。

本发明所提供的体系中，还可以包括用于 PCR 检测的检测引物 1、检测引物 2 和探针 3，三者中至少其一含有基因特异性序列，具体的可以是单特异性、双特异性或三特异性的引物/探针组合。优选的实施例中，仅检测引物 1 包含基因特异性序列。其应用于非二代测序而是 PCR 检测的场景。

在本发明一优选实施例中，通过本发明的方法或体系进行文库构建后，文库分子在结构上依次包含以下序列：5'端测序通用序列、基因特异性探针序列、富集到的目标区域序列、3'端测序通用序列，如图 2 所示。其中富集到的目标区域包含富集前样本 DNA 的序列信息，其特点是：5'端在基因组上的位置是固定的，由特异性探针决定；而 3'端位置不固定，由建库初始的 DNA 片段化状态决定。因此，在富集后的数据分析中，该序列的 3'端在基因组上的位置也可以起到分子标签的作用。

本发明的有益效果在于，首先，在连接前先对所有片段化 DNA 样本的目标区域通过捕获延伸的手段进行预扩增，避免在连接阶段因连接酶连接效率不足所致的原始目标分子的损失或漏检，特别是小片段和稀有分子；预扩增阶段的延伸反应属于线性扩增，没有 PCR 扩增的偏好性，也不会累积 PCR 扩增所引入的错误，与常规 PCR 建库技术相比，产物均一性好；其次，预扩增阶段，只需为每个目标基因设计一条长约 30bp 的单链探针，避开了为短片段如 cfDNA 设计双端引物的困难，既提高了建库成功率还提升了建库便利性；而封闭探针 3'端，可以阻断探针与模板之外的非目的连接，有效降低游离探针引起的背景噪声，辅以耦联有标记分子的 dNTP 及其纯化系统，进一步提高目的产物纯度，使样本 DNA 分子达到最高转化率；再次，将特异性探针的 5'端修饰后，在单链连接酶催化下能很好地连接到延伸产物的 3'端但不会连接到游离探针经过封闭修饰的 3'端，避免游离探针的自连和错连，提高目的连接产物在终产物中的比例。最后，本发明的富集方法与申请人在申请号为 2019100024085 中公开的富集方法相比，省去了接头 DNA，原料准备上更具便利性，不同组分之间相互干扰生成非目的产物的几率更低，且进一步降低了成本。综上所述，本发明的富集方法及体系，操作简单，结果可靠，针对片段长度小于 200bp 的 DNA 使用，可以最大程度的减少原始分子尤其是稀有分子的损失，最高效地富集目标分子，且其保真性、特异性和灵敏度高，可检出突

变率低至 0.01% 的稀有突变分子。

以下通过特定的具体实例说明本发明的实施方式，本领域技术人员可由本说明书所揭露的内容轻易地了解本发明的其他优点与功效。本发明还可以通过另外不同的具体实施方式加以实施或应用，本说明书中的各项细节也可以基于不同观点与应用，在没有背离本发明的精神下进行各种修饰或改变。

实施例 1

实施例中所用寡聚核苷酸如表 1 所示：

表 1

名称	SEQ ID NO.	5'端修饰	3'端修饰	序列 (5' → 3')
Probe1	1	5'OH	3'OH	<u>GCTGAGTCGGAGACACGCAG</u> <u>GGATGAGACGCCACAACAGG</u> <u>CCTTCAGGCTCTCGCTGTGC</u> CCAATATTGTCTTTGTGTTC CCGGACATAGTCCAG
Probe2	2	5'PO4	3'OH	<u>GCTGAGTCGGAGACACGCAG</u> <u>GGATGAGACGCCACAACAGG</u> <u>CCTTCAGGCTCTCGCTGTGC</u> CCAATATTGTCTTTGTGTTC CCGGACATAGTCCAG
Probe3	3	5'PO4	3'C6 Spacer	<u>GCTGAGTCGGAGACACGCAG</u> <u>GGATGAGACGCCACAACAGG</u> <u>CCTTCAGGCTCTCGCTGTGC</u> CCAATATTGTCTTTGTGTTC CCGGACATAGTCCAG
Probe4	4	5'PO4	3'C3 Spacer	<u>GCTGAGTCGGAGACACGCAG</u> <u>GGATGAGACGCCACAACAGG</u> <u>CCTTCAGGCTCTCGCTGTGC</u> CCAATATTGTCTTTGTGTTC CCGGACATAGTCCAG
Probe5	5	5'PO4	3'NH2	<u>GCTGAGTCGGAGACACGCAG</u> <u>GGATGAGACGCCACAACAGG</u> <u>CCTTCAGGCTCTCGCTGTGC</u> CCAATATTGTCTTTGTGTTC CCGGACATAGTCCAG
Probe6	6	5'PO4	3'H	<u>GCTGAGTCGGAGACACGCAG</u> <u>GGATGAGACGCCACAACAGG</u> <u>CCTTCAGGCTCTCGCTGTGC</u> CCAATATTGTCTTTGTGTTC CCGGACATAGTCCAG
Probe7	7	5'PO4	3' PO4	<u>GCTGAGTCGGAGACACGCAG</u> <u>GGATGAGACGCCACAACAGG</u> <u>CCTTCAGGCTCTCGCTGTGC</u> CCAATATTGTCTTTGTGTTC CCGGACATAGTCCAG
Probe8	8	5'PO4	2'O-Methyl	<u>GCTGAGTCGGAGACACGCAG</u> <u>GGATGAGACGCCACAACAGG</u> <u>CCTTCAGGCTCTCGCTGTGC</u> CCAATATTGTCTTTGTGTTC CCGGACATAGTCCAG

Probe9	9	5'App	3'C3 Spacer	<u>GCTGAGTCGGAGACACGCAG</u> <u>GGATGAGACGCCACAACAGG</u> <u>CCTTCAGGCTCTCGCTGTGC</u> CCAATATTGTCTTTGTGTTC CCGGACATAGTCCAG
Probe10	10	5'PO4	3'C3 Spacer	<u>GCTGAGTCGGAGACACGCAG</u> <u>GGATGAGACGCCACAACAGG</u> <u>CCTTCAGGCTCTCGCTGTGC</u> CCAATATTGTCTTTGTGTTC CCGGACATArGTCCT
Probe11	11	5'PO4	3'C3 Spacer	<u>GCTGAGTCGGAGACACGCAG</u> <u>GGATGAGACGCCACAACAGG</u> <u>CCTTCAGGCTCTCGCTGTGC</u> CCAATATTGTCTTTGTGTTC CCGGACATAGTCCAGT
UF	12	5'OH	3'OH	CCTTCAGGCTCTCGCTGTGC
UR	13	5'OH	3'OH	CCTGTTGTGGCGTCTCATCC
F1	14	5'OH	3'OH	TCCCGGACAT AGTCCAGGAGGC
RX1	15	5'OH	3'OH	CGTGTGCCGCCTGCTGGG
MX1	16	5'FAM	3'MGB	CACGGTGGAGGTGAGGC
REF	17	5'OH	3'OH	GATCCTTCAGGCTCTCGCTG TGCCCAATATTGTCTTTGTG TTCCCGGACATAGTCCAGGA GGCAGCCGAAGGGCATGAGC TGC GTGATGAGCTGCACGGT GGAGGTGAGGCAGATGCCCA GCAGGCGGCACACGTGGGGG TTGCTGAGTCGGAGACACGC AGGGATGAGACGCCACAACA GGATC

此实施例中的探针都包含通用序列（下划线部分）。所有寡聚核苷酸由上海百力格生物有限公司合成。探针经 HPLC 两次纯化后，达到 99%以上的纯度。表中 SEQ ID NO.1-11 的探针的特异性序列都针对 EGFR 基因的第 20 外显子区域。

各探针 5'端核苷酸与 3'端核苷酸的修饰如表 1 中所示。其中，probe 6（SEQ ID NO.6）3'端核苷酸的 3 位羟基被替换为氢原子，变成了双脱氧核苷酸；probe 8（SEQ ID NO.8）3'端核苷酸的 2 位氢原子被甲氧基取代；probe 9（SEQ ID NO.9）5'端核苷酸的 5 位羟基被腺苷替代；probe 10（SEQ ID NO.10）序列中靠近 3'端的脱氧核糖核苷酸 G 被替换成核糖核苷酸 rG；probe 10 和 probe 11 和（SEQ ID NO.10，SEQ ID NO.11）的最后一个碱基均为错配碱基。

UF 和 UR（SEQ ID NO.12，SEQ ID NO.13）是用于 PCR 扩增环化产物的前后引物，F1（SEQ ID NO.17）是用于检测 EGFR 基因第 20 外显子区域的前引物，RX1 和 MX1（SEQ ID NO.19 和 SEQ ID NO.20）分别是用于检测 EGFR 基因第 20 外显子区域的后引物和 MGB 探针。

主要试剂与材料

细胞 DNA 提取试剂盒和用于纯化单链扩增产物的 RNA Clean Kit 试剂盒购自天根生化科技（北京）有限公司。rTaq DNA 聚合酶、KOD plus 高保真 DNA 聚合酶和用于 PCR 检测的反应预混液（QPK-101）购自日本 TOYOBO 公司；RNase H2 热稳定性核酸酶购自美国 IDT 公司。APG ssDNA 连接酶、APG Enchant 高保真 DNA 聚合酶购自上海羿鸣生物科技有限公司。USER 酶、Q5 高保真 DNA 聚合酶、T4 RNA 连接酶与 5' App DNA/RNA 热稳定连接酶购自美国 NEB 公司；Agencourt AMPure 磁珠购自美国 Beckman 公司；相关过程中定量检测使用的校准品按照 REF 序列由上海百力格生物技术有限公司合成，按照一定浓度梯度配置于 20% 的甘油 TE 缓冲液中。

实验方法

步骤 1、样本准备

使用细胞 DNA 提取试剂盒提取健康人志愿者血细胞样本中 gDNA，超声打断成平均长度为 150bp 左右的片段，使用 qubit 定量后备用。使用去离子水作为对照样本 NC。

步骤 2、样本的定量检测

按照表 2 配制样本的 PCR 定量检测体系，并用表 3 的扩增条件，使用针对 EGFR20 外显子区域的检测体系定量检测样本的拷贝数浓度。

表 2. 样本 PCR 定量检测体系		
组分	用量 (μL)	终浓度
2×Taqman Master Mix (Toyobo)	25	1×
F1 (10 μM)	1.5	300 nM
RX1 (10 μM)	1.5	300 nM
MX1 (10 μM)	0.5	100 nM
去离子水	17.5	/
所有样本 / 校准品	4	/
总体积	50	/

表 3 样本 PCR 定量检测扩增条件		
循环数	温度	时间
1 cycle	95 °C	3 min
45 cycles	95 °C	10 sec
	60 °C	30 sec

步骤 3、线性扩增

3.1 按表 4、表 5、表 6 分别配制线性扩增反应体系。线性扩增的 PCR 程序如表 7 所示。

表 4 高保真线性扩增反应体系

组分	体积 (μL)	终浓度/总拷贝数
去离子水	3.1	/
5X Buffer	5	1×
probe (1 μM)	0.5	50 nM
dNTP (2 mM)	1.25	100 μM
APG Enchant 高保真聚合酶 (2U/μL)	0.15	
样本	15	20000 拷贝
Total	25	/

表 5 一般线性扩增反应体系

组分	体积 (μL)	终浓度/总拷贝数
去离子水	4	/
10 X Taq Buffer	2.5	1×
MgCl ₂	1.5	1.5 mM
probe (1 μM)	0.5	50 nM
dNTP (2 mM)	1.25	100 μM
rTaq 聚合酶 (2.5U/μL)	0.25	
样本	15	20000 拷贝
Total	25	/

表 6. 含核酸酶的线性扩增反应体系

组分	体积 (μL)	终浓度/总拷贝数
去离子水	3.8	/
10 X Taq Buffer	2.5	1×
MgCl ₂	1.5	1.5 mM
probe (1 μM)	0.5	50 nM
dNTP (2 mM)	1.25	100 μM
rTaq 聚合酶 (2U/uL)	0.25	
RNAse H2 核酸酶	0.2	
样本	15	20000 拷贝
Total	25	/

表 7. 线性扩增程序		
循环数	温度	时间
1 cycle	98 °C	2 min
1 或 20 或 40 或 80 cycles	98 °C	15 sec
	70 °C	45 sec
1 cycle	72 °C	4 min

3.2 线性扩增程序完成后，用 RNA Clean Kit 试剂盒对单链扩增产物进行纯化，去除未延伸的探针，用 30 μ L 洗脱缓冲液洗脱。

3.3 线性扩增效率检测

线性扩增产物的PCR检测体系如表2，检测线性扩增的PCR程序如表3所示。

步骤 4、环化连接

根据步骤 3 的结果，选择步骤 2 得到的部分线性扩增产物与对照样本 QC 分别按表 8、表 9 和表 10 进行环化反应。

表 8. APG ssDNA 连接酶环化体系			
组分	体积 (μ L)	终浓度	连接条件
10X ssDNA ligation Buffer	4	1X	60°C 1hr, 95°C 10 min 使连接酶 失活
50 mM MnCl ₂	2	2.5 mM	
APG ssDNA 连接酶 (200 U/ μ L)	1	5 U/ μ L	
线性扩增产物	25		
去离子水	8		
Total	40	/	

表 9. T4RNA 连接酶环化体系			
组分	体积 (μ L)	终浓度	连接条件
10X 连接 Buffer	4	1X	16°C 4hr, 60°C 15min 使连接酶 失活
T4 RNA 连接酶	2	5 U / μ L	
去离子水	1		
线性扩增产物	25		
50% PEG	8	10%	
Total	40	/	

表 10. 5'App 热稳定连接酶环化体系			
组分	体积 (μ L)	终浓度	连接条件

10X NE Buffer 1	4	1X	65°C 1hr, 95°C 15min 使连接酶 失活
50 mM MnCl ₂	2	2.5 mM	
5'App DNA/RNA 热稳定连接酶 (100 U/μL)	2	5 U/μL	
线性扩增产物	25		
50% PEG	8	10%	
Total	41	/	

步骤 5、环化产物得率检测

使用 TE 缓冲液将环化产物稀释 10 倍，按照表 11 配制 PCR 检测体系对步骤 4 的中的目标环化产物进行得率检测，同时检测不加模板的对照样本 NC。PCR 程序如表 12 所示。使用校准品对目标环化产物分子进行定量，评价目标环化产物得率。

表 11. 目标环化产物检测体系		
试剂	用量 (μL)	终浓度
2 X Taqman master Mix	12.5	1×
F1 (10 μM)	0.75	300 nM
UR (10 μM)	0.75	300 nM
MX1 (10 μM)	0.25	100 nM
去离子水	8.75	/
校准品/环化产物	2	/
Total	25	/

表 12. 目标环化产物检测 PCR 条件		
循环数	温度	时间
1 cycle	95 °C	3 min
45 cycles	95 °C	10 sec
	60 °C	30 sec

步骤 6、扩增环化产物

按表 13 的反应体系和表 14 的程序对步骤 4 得到的环化产物进行 PCR 扩增。

表13. 环化扩增体系		
组分	体积 (μL)	终浓度
环化产物	30	/
5× Q5 Buffer	10	1X

dNTP (10 mM each)	1	0.2 mM
Q5 DNA聚合酶 (2 U/ μ L) (NEB)	1	/
UF (10 μ M)	1	200 nM
UR (10 μ M)	1	200 nM
去离子水	6	/
Total	50	/

表 14. 文库扩增 PCR 程序		
循环数	温度	时间
1 cycle	98 °C	3 min
15 cycles	98 °C	10 sec
	72 °C	30 sec
1 cycle	72 °C	5 min

反应结束后，取 80 μ L AMPure XP 磁珠对扩增后文库进行纯化，参照操作说明，最终将纯化产物于 20 μ L 洗脱缓冲液（10 mM Tris-HCl，0.1mMEDTA，pH8.0）中洗脱，在磁力架上取上清转移到新的微量离心管。至此，包含通用序列的靶向文库分子扩库完成。

将扩增后的文库产物用 TE 缓冲液稀释 100 倍后，使用表 11 和表 12 的检测体系和条件检测文库产物。

实验结果及分析

本实施例用线性扩增倍数、环化连接效率、分子转化率与扩库倍数这四个参数评价本发明的方法和体系对目标区域扩增和连接的技术效果。具体计算方式如下：

线性扩增倍率 = 目的基因线性扩增后输出拷贝数 / 目的基因线性扩增前输入拷贝数

环化分子得率 = 目的基因环化产物输出拷贝数 / 目的基因线性扩增产物输入拷贝数 $\times 100\%$

分子转化率 = 目的基因环化产物输出拷贝数 / 扩增前输入拷贝数 $\times 100\%$

扩库倍数 = 扩库纯化后拷贝数 / 环化产物输出拷贝数 $\times 100\%$

结果 1、不同类型的聚合酶对不同修饰探针的线性扩增效率影响如表 15 所示

表 15. 不同修饰探针的线性扩增效率						
探针	DNA 聚合酶类型	核酸酶	循环数	线性扩增后输出拷贝数	扩增前输入拷贝数	线性扩增倍率
Probe1	高保真聚合酶	无	40	543000	20000	27.2
Probe2	高保真聚合酶	无	40	521000	20000	26.1

Probe3	高保真聚合酶	无	40	384000	20000	19.2
Probe4	高保真聚合酶	无	40	416000	20000	20.8
Probe5	高保真聚合酶	无	40	69300	20000	3.5
Probe6	高保真聚合酶	无	40	84200	20000	4.2
Probe7	高保真聚合酶	无	40	64000	20000	3.2
Probe8	高保真聚合酶	无	40	92000	20000	4.6
Probe9	高保真聚合酶	无	40	392000	20000	19.6
Probe10	高保真聚合酶	无	40	408000	20000	20.4
Probe11	高保真聚合酶	无	40	286000	20000	14.3
Probe1	普通聚合酶	无	40	623000	20000	31.2
Probe2	普通聚合酶	无	40	598000	20000	29.9
Probe3	普通聚合酶	无	40	23100	20000	1.2
Probe4	普通聚合酶	无	40	29200	20000	1.5
Probe5	普通聚合酶	无	40	31300	20000	1.6
Probe6	普通聚合酶	无	40	32800	20000	1.6
Probe7	普通聚合酶	无	40	27000	20000	1.4
Probe8	普通聚合酶	无	40	56000	20000	2.8
Probe9	普通聚合酶	无	40	21200	20000	1.1
Probe10	普通聚合酶	无	40	23800	20000	1.2
Probe11	普通聚合酶	无	40	22400	20000	1.1
Probe10	普通聚合酶	有	40	432000	20000	21.6
Probe11	普通聚合酶	有	40	18600	20000	0.9

对于所有对照样本，均没有检测到信号。

表 15 中所涉及的线性扩增反应体系和条件参见表 4 至表 7，其中，高保真聚合酶反应体系中使用的是 APG Enchant 高保真 DNA 聚合酶，具有 3'-5'外切酶活性，而普通聚合酶反应体系中使用的是 rTaq 酶，具有 5'-3'外切酶活性，但不具有 3'-5'外切酶活性。为了实现外切酶的功能，我们在普通聚合酶反应体系中添加了热稳定性核酸酶后，也能切除修饰后封闭的 3'末端，达到与使用高保真 DNA 聚合酶类似的技术效果。

从测试结果可知：1) 3'端核苷酸封闭修饰的探针在高保真聚合酶的作用下结合模板后，封闭基团被切除，探针被活化，从而可以对目标序列进行有效扩增；2) 热稳定性核酸酶与普通聚合酶 rTaq 合用，对核糖核苷酸基团修饰的探针也能起到活化效果，增强了本发明富集方法的适用性；3) 不同取代基团的选择对于线性扩增效果有明显差异，将探针 3'端羟基取代为 C3 Spacer 得到的效果理想，与其它取代基团相比具有明显优势。本实施例从结果 2 开始的数据主要是针对将 3'端羟基取代为 C3 Spacer 的探针得到的结果数据。4) 3'端核苷酸封闭修饰的探针的 3'端附近出现错配时 (Probe 11)，对于具有 proofreading 功能的高保真聚合酶驱动的线性扩增反应没有显著影响，或还有帮助。

结果 2、不同扩增循环数对线性扩增效率的影响如表 16 所示。

表 16 中所涉及的线性扩增反应体系和条件参见步骤 2，线性扩增效率检测方法参见步骤 3。

表 16. 不同循环数的扩增结果 (QC 样本)						
探针	DNA 聚合酶类型	核酸酶	循环数	线性扩增后输出拷贝数	扩增前输入拷贝数	线性扩增倍率
Probe1	高保真聚合酶	无	80	922000	20000	46.1
Probe2	高保真聚合酶	无	80	893400	20000	44.7
Probe4	高保真聚合酶	无	80	635000	20000	31.8
Probe11	高保真聚合酶	无	80	742000	20000	37.1
Probe10	普通聚合酶	有	80	821000	20000	41.1
Probe1	高保真聚合酶	无	40	543000	20000	27.2
Probe2	高保真聚合酶	无	40	521000	20000	26.1
Probe4	高保真聚合酶	无	40	416000	20000	20.8
Probe11	高保真聚合酶	无	40	286000	20000	14.3
Probe10	普通聚合酶	有	40	432000	20000	21.6
Probe1	高保真聚合酶	无	20	286000	20000	14.3
Probe2	高保真聚合酶	无	20	263000	20000	13.2
Probe4	高保真聚合酶	无	20	221000	20000	11.1
Probe11	高保真聚合酶	无	20	216000	20000	10.8
Probe10	普通聚合酶	有	20	258000	20000	12.9
Probe1	高保真聚合酶	无	1	33600	20000	1.7
Probe2	高保真聚合酶	无	1	32800	20000	1.6
Probe4	高保真聚合酶	无	1	30300	20000	1.5
Probe11	高保真聚合酶	无	1	29800	20000	1.5
Probe10	普通聚合酶	有	1	31600	20000	1.6

测试结果表明，无论是高保真聚合酶体系还是普通聚合酶加核酸酶的体系，都可通过循环数的增加可以有效提高 3'端封闭探针的线性扩增倍率。

结果 3、封闭探针环化结果的影响如表 17 所示

表 17 中所涉及环化反应体系见步骤 4，连接效率检测的方法见步骤 5。

表 17 不同探针环化后的测试结果							
探针	线性扩增用酶	循环数	环化连接酶	线性扩增后输入拷贝数	目标环化分子拷贝数	目标环化分子得率	分子转化率
Probe1	高保真聚合酶	80	APG ssDNA 连接酶	768333	0	0.00%	0.00%
Probe2	高保真聚合酶	80	APG ssDNA 连接酶	744500	521	0.07%	2.61%
Probe4	高保真聚合酶	80	APG ssDNA 连接酶	529167	28734	5.43%	143.67%
Probe1 1	高保真聚合酶	80	APG ssDNA 连接酶	618333	29124	4.71%	145.62%

Probe1 0	普通聚合酶+核酸酶	80	ssDNA 连接酶	684167	21414	3.13%	107.07%
-------------	-----------	----	-----------	--------	-------	-------	---------

结果表明: 1) 使用将 3'端核苷酸羟基替换为 C3 Spacer 的探针, 能够获得理想的目标环化分子得率和分子转化率, 后者可超过 100%, 表明用本发明的富集方法可以做到从原始分子到两端配置有测序通用序列的文库分子的完全无损; 2) 未经修饰的探针体系反应后, 只能检出极少量目标环化分子; 3) 探针的 5'端核苷酸只有被磷酸化修饰后, 在 APG ssDNA 连接酶作用下, 才能实现自身环化; 4) 对于 3'端附近被核糖核苷酸封闭修饰的探针, 使用普通聚合酶加核酸酶的体系, 也能取得较高的目标环化分子得率。

结果 4、不同 5'端修饰的探针和不同连接酶对环化效率的影响如表 18 所示。

表 18 中所涉及的连接反应体系参见表 8、表 9 和表 10, 环化效率检测方法参见步骤 5。

表 18. 不同 5'端修饰的探针和不同连接酶的测试结果							
探针	线性扩增用酶	循环数	环化连接酶	线性扩增后输入分子拷贝数	目标环化分子拷贝数	目标环化分子得率	分子转化率
Probe4	高保真聚合酶	80	ssDNA 连接酶	529167	28734	5.43%	143.67%
Probe4	高保真聚合酶	80	T4 RNA 连接酶	529167	2963	0.56%	14.82%
Probe9	高保真聚合酶	80	5'AppDNA/RNA 连接酶	504680	7318	1.45%	36.59%

结果表明, 使用 5'端磷酸化修饰的封闭探针配合 APG ssDNA 连接酶环化效果理想, 优于腺苷修饰的封闭探针配合其他连接酶的体系。

结果 5、连接产物扩库效率检测如表 19 所示。

表 19 不同探针环化后的测试结果							
探针	线性扩增用酶	循环数	环化连接酶	线性扩增后输入拷贝数	环化分子拷贝数	扩库后目标分子拷贝数	扩库倍数
Probe2	高保真聚合酶	80	ssDNA 连接酶	744500	521	1459	2.8
Probe4	高保真聚合酶	80	ssDNA 连接酶	529167	28734	3600339	125.3
Probe11	高保真聚合酶	80	ssDNA 连接酶	618333	29124	3145338	108.0
Probe10	普通聚合酶+核酸酶	80	ssDNA 连接酶	684167	21414	1893034	88.4

表 19 中所涉及的扩库反应体系与程序参见步骤 6。

结果表明, 封闭探针的目标环化分子可以得到较好的扩增。扩库倍数达到 100 倍左右,

显著优于未封闭探针。

结论：

探针 3'端核苷酸羟基的封闭修饰至关重要，可以防止探针自连，大幅提高目标分子的连接效率，而结合模板后的探针被 DNA 聚合酶切除修饰基团后还能进行多个循环的线性扩增。另外，热稳定性 APG ssDNA 连接酶的选用，环化后的纯化处理，以及环化后使用通用引物扩增也都能起到提高建库产物纯度和检测灵敏度的作用。

实施例 2

实施例中所用寡聚核苷酸如表 20 所示：

表 20.

名称	SEQ ID NO.	5'端修饰	3'端修饰	序列 (5' →3')
Probe4	4	5'PO4	3'C3 Spacer	<u>GCTGAGTCGGAGACACGCAG</u> <u>GGATGAGACGCCACAACAGG</u> <u>CCTTCAGGCTCTCGCTGTGC</u> CCAATATTGTCTTTGTGTTC CCGGACATAGTCCAG
Probe12	18	5'PO4	3'C3 Spacer	<u>GCTGAGTCGGAGACACGCAG</u> <u>GGATGAGACGCCACAACAGG</u> <u>CCUTCAGGCTCTCGCTGTGC</u> CCAATATTGTCTTTGTGTTC CCGGACATAGTCCAG
UF	12	5'OH	3'OH	CCTTCAGGCTCTCGCTGTGC
UR	13	5'OH	3'OH	CCTGTTGTGGCGTCTCATCC
F1	14	5'OH	3'OH	TCCCGGACAT AGTCCAGGAGGC
RX1	15	5'OH	3'OH	CGTGTGCCGCCTGCTGGG
MX1	16	5'FAM	3'MGB	CACGGTGGAGGTGAGGC
REF	17	5'OH	3'OH	GATCCTTCAGGCTCTCGCTG TGCCCAATATTGTCTTTGTG TTCCCGGACATAGTCCAGGA GGCAGCCGAAGGGCATGAGC TGCGTGATGAGCTGCACGGT GGAGGTGAGGCAGATGCCCA GCAGGCGGCACACGTGGGGG TTGCTGAGTCGGAGACACGC AGGGATGAGACGCCACAACA GGATC

同实施例 1，本实施例中的探针也包含通用序列。探针经 HPLC 两次纯化后，保证 99% 以上的纯度。

其中，probe 12（SEQ ID NO.18）两端的修饰与 probe4 相同，通用序列部分有一个碱基 T 被替代为 U。其余检测引物与探针与实施例 1 相同。

主要试剂与材料

APG ssDNA 内切酶购自上海羿鸣生物科技有限公司，USER 酶购自美国 NEB 公司，其余与实施例 1 相同。

实验方法

本实施例的样本准备，样本的定量检测，线性扩增与环化连接均与实施例 1 相同，其中初始模板为 20000 拷贝，线性扩增选用 APG Enchant 高保真聚合酶，80 个循环，环化连接选用 APG ssDNA 连接酶。

取环化产物加入 USER 酶或 APG ssDNA 内切酶，37 度酶切 1 小时，并以未酶切的环化产物作为对照。APG ssDNA 内切酶可以特异性酶切序列中包含 5' GGCC 3' 的单链 DNA。酶切后用 RNA Clean Kit 纯化酶切产物，用 35 μ L 洗脱缓冲液洗脱。

取 5 μ L 洗脱产物使用 TE 缓冲液将洗脱产物稀释 10 倍，采用实施例 1 中表 11 PCR 检测体系检测目标环化产物得率，PCR 程序如表 12 所示。

按照实施例 1 步骤 6 的方案，按表 13 的反应体系和表 14 的程序对步骤 4 得到的环化后洗脱产物进行 PCR 扩增。

反应结束后，取 80 μ L AMPure XP 磁珠对扩增后文库进行纯化，参照操作说明，最终将纯化产物于 20 μ L 洗脱缓冲液（10 mM TrisCl，0.1 mM EDTA，pH8.0）中洗脱，在磁力架上取上清转移到新的微量离心管。至此，包含通用序列的靶向文库分子扩库完成。

扩增后的文库产物，TE 缓冲液稀释 100 倍后，使用上文表 11 和表 12 检测体系和条件检测文库产物。

实验结果及分析

同实施例 1，本实施例也用线性扩增倍数、环化分子得率、分子转化率与扩库倍数这四个参数以评价本方法对目标区域扩增和连接的技术效果，具体计算方法见实施例 1。

结果 2.1 环化后酶切对检测的影响如表 21 所示。

表 21. 环化后酶切对检测的影响							
探针	线性扩增用酶	循环数	环化后酶切	线性扩增后输入分子拷贝数	目标环化分子拷贝数	环化分子得率	分子转化率
Probe4	高保真聚合酶	80	无	568000	38170	6.72%	190.85%
Probe4	高保真聚合酶	80	ssDNA 内切酶	568000	105818	18.63%	529.09%
Probe12	高保真聚合酶	80	无	375000	20775	5.54%	103.88%
Probe12	高保真聚合酶	80	USER 酶	375000	48638	12.97%	243.19%

结果表明，环化产物在酶切之后，检测出的环化分子得率与分子转化率均有显著提升。

结果 2.2 环化后酶切对扩库效率的影响如表 22 所示。

表 22. 环化后酶切对扩库效率的影响						
探针	线性扩增用酶	循环数	环化后酶切	目标环化分子 拷贝数	扩库后目标分 子拷贝数	扩库倍 数
Probe4	高保真聚合酶	80	无	38170	12748646	334
Probe4	高保真聚合酶	80	ssDNA 内切 酶	105818	101162390	956
Probe12	高保真聚合酶	80	无	20775	5152200	248
Probe12	高保真聚合酶	80	USER 酶	48638	38229075	786

结果表明，环化产物在酶切之后，扩库效率得到了显著的提升。

实施例 3

本实施例评估用本发明的富集方法和体系对真实样本进行线性扩增-环化建库后，针对微小片段的稀有分子和稀有变异分子的检出能力。

本实施例所用寡聚核苷酸序列如表 23 所示：

表 23

名称	SEQ ID NO.	5'端修饰	3'端修饰	序列 (5' → 3')
Probe4	4	5'PO4	3'C3 Spacer	<u>GCTGAGTCGGAGA</u> <u>CACGCAGGGATGA</u> <u>GACGCCACAACAG</u> <u>GCCTTCAGGCTCTC</u> <u>GCTGTGCCAATAT</u> TGTCTTTGTGTTCC CGGACATAGTCCA G
Probe13	19	5'PO4	3'C3 Spacer	<u>GCTGAGTCGGAGA</u> <u>CACGCAGGGATGA</u> <u>GACGCCACAACAG</u> <u>GCCTTCAGGCTCTC</u> <u>GCTGTGCGCTGCC</u> GAAGAATTGCAGT TTGCTTCCCGC
FX1	20	5'OH	3'OH	GCGTGATGAGT
FX2	21	5'OH	3'OH	CGCTTCGGTACTCT GC
RX2	22	5'OH	3'OH	GTCATCCCTGTACA ACCTG
MX2	23	5'FAM	3'MGB	GCGAAGTGCAACT GGACAA
UM	24	5'FAM	3'MGB	TGCGTGTCTCCGAC TCAGC
UF	12	5'OH	3'OH	CCTTCAGGCTCTCG CTGTGC
UR	13	5'OH	3'OH	CCTGTTGTGGCGTC TCATCC

RX1	15	5'OH	3'OH	CGTGTGCCGCCTGC TGGG
MX1	16	5'FAM	3'MGB	CACGGTGGAGGTG AGGC

其中探针 13 (SEQ ID NO.19) 特异性序列针对性染色体 Y 上的 SRY 基因。FX1 (SEQ ID NO.20) 为用来特异检测 EGFR 基因第 20 外显子区域 SNP-Q787 (2361 G>A) 的前引物, 前实施例 RX1 和 MX1 配对组成特异性检测 SNP-Q787 的 PCR 检测体系。FX2、RX2 和 MX2 (SEQ ID NO.21-23) 分别是用于检测 SRY 探针扩增产物的前引物, 后引物和 MGB 探针。UM (SEQ ID NO.24) 是用来检测通用序列的 MGB 探针。

主要试剂与材料

游离 DNA 提取试剂盒购自上海臻迪基因科技有限公司, 其他同实施例 1 和 2。

实验方法

步骤 1、靶向文库制备

使用游离 DNA 提取试剂盒提取健康人血浆样本中的 cfDNA, 使用 Qubit 定量后备用。通过 F1 和 RX1 扩增 EGFR-20 外显子区域的片段。使用一代测序方法检测其中 SNP-Q787 的状态。将 SNP 纯合子的男性 cfDNA 样本, 按照 1%, 0.1%, 0.03%, 0.01% 的质量比掺入到女性野生型 cfDNA 样本中, 同时以未掺入的女性野生型 cfDNA 样本作为空白对照 (QC)。以 50ng (约 15000 拷贝, 共 15ul) 作为单个样本总起始模板量进行后续靶向建库。每种梯度设置两个复孔, 最后以复孔结果的平均值作为该种样本的检测结果。

使用 APG ssDNA 内切酶对样本进行靶向建库。探针选用 probe4 和 probe13, 浓度为 50nM。线性扩增循环数为 80, 环化使用 APG ssDNA 连接酶, 酶切纯化后扩增 12 个循环, 再用 AMPure XP 磁珠纯化后获得 20 uL 文库产物。

步骤 2、不同特异性引物/探针组合的 qPCR 检测

对原始样本和靶向建库样本分别进行三特异性 PCR 检测体系检测, 再对靶向建库样本进行双特异性、单特异性 PCR 检测。取检测样本的平均 CT 值, 以 QC 样本的平均 CT 值为检测背景, 以两者之间的差值 ΔCT , 作为评价标准。当 $\Delta CT \geq 2.5$ 时认为具备显著性差异。

EGFR 基因 20 外显子 Q787Q 位点三特异性、双特异性、单特异性检测体系分别如表 24-26 所示。SRY 基因三特异性、双特异性、单特异性检测体系分别如表 27-29 所示。qPCR 检测程序如表 30 所示。

表 24. Q787Q 位点三特异性检测体系:

试剂	用量 (μ L)	终浓度
2×Taqman master Mix	10	1×
FX1 (10 μ M)	0.6	300 nM
RX1 (10 μ M)	0.6	300 nM

MX1 (10 μ M)	0.2	100 nM
H ₂ O	6.6	/
待测样本	2	/
Total	20	/

表 25. Q787Q 位点双特异性检测体系:

试剂	用量 (μ L)	终浓度
2 $\times$ Taqman master Mix	10	1 $\times$
FX1 (10 μ M)	0.6	300 nM
UR (10 μ M)	0.6	300 nM
MX1 (10 μ M)	0.2	100 nM
H ₂ O	6.6	/
待测样本	2	/
Total	20	/

表 26. Q787Q 位点单特异性检测体系:

试剂	用量 (μ L)	终浓度
2 $\times$ Taqman master Mix	10	1 $\times$
FX1 (10 μ M)	0.6	300 nM
UR (10 μ M)	0.6	300 nM
UM (10 μ M)	0.2	100 nM
H ₂ O	6.6	/
待测样本	2	/
Total	20	/

表 27. SRY 基因三特异性检测体系:

试剂	用量 (μ L)	终浓度
2 $\times$ Taqman master Mix	10	1 $\times$
FX2 (10 μ M)	0.6	300 nM
RX2 (10 μ M)	0.6	300 nM
MX2 (10 μ M)	0.2	100 nM
H ₂ O	6.6	/
待测样本	2	/
Total	20	/

表 28. SRY 基因双特异性检测体系:

试剂	用量 (μ L)	终浓度
2 $\times$ Taqman master Mix	10	1 $\times$
FX2 (10 μ M)	0.6	300 nM
UR (10 μ M)	0.6	300 nM
MX2 (10 μ M)	0.2	100 nM
H ₂ O	6.6	/
待测样本	2	/
Total	20	/

表 29. SRY 基因单特异性检测体系:

试剂	用量 (μL)	终浓度
2×Taqman master Mix	10	1×
FX2 (10 μM)	0.6	300 nM
UR (10 μM)	0.6	300 nM
UM (10μM)	0.2	100 nM
H ₂ O	6.6	/
待测样本	2	/
Total	20	/

表 30.

循环数	温度	时间
1 cycle	95°C	4 min
40 cycles	95°C	10 s
	61°C	30 s

实验结果及分析

针对原始样本和靶向建库后的样本使用不同特异性体系检测 EGFR SNP-Q787 位点和 SRY 基因的检测结果分别如表 31, 32 所示。

表 31、不同掺入比下的 EGFR SNP-Q787 位点检测结果

样本类型	掺入比	PCR 体系	CT1	CT2	CT 均值	ΔCT
原始样本	1%	三特异	30.23	31.06	30.65	7.29
	0.10%	三特异	34.35	35.56	34.96	2.98
	0.03%	三特异	36.43	36.85	36.64	1.29
	0.01%	三特异	37.94	37.61	37.78	0.16
	QC	三特异	38.34	37.52	37.93	0.00
文库样本	1%	三特异	25.12	24.98	25.05	9.67
	0.10%	三特异	28.35	29.65	29.00	5.72
	0.03%	三特异	31.52	32.06	31.79	2.93
	0.01%	三特异	32.84	33.52	33.18	1.54
	QC	三特异	34.42	35.02	34.72	0.00
文库样本	1%	双特异	24.02	24.26	24.14	10.16
	0.10%	双特异	27.42	27.58	27.50	6.80
	0.03%	双特异	30.16	29.54	29.85	4.45
	0.01%	双特异	31.25	31.69	31.47	2.83
	QC	双特异	34.56	34.04	34.30	0.00
文库样本	1%	单特异	23.56	23.43	23.50	10.04
	0.10%	单特异	27.54	26.92	27.23	6.30
	0.03%	单特异	29.22	28.95	29.09	4.45

	0.01%	单特异	31.04	30.46	30.75	2.78
	QC	单特异	33.24	33.82	33.53	0.00

表 32. 不同掺入比下的 SRY 基因检测结果

样本类型	掺入比	PCR 体系	CT1	CT2	CT 均值	ΔCT
原始样本	1%	三特异	31.41	31.20	31.31	8.69
	0.10%	三特异	36.84	35.30	36.07	3.93
	0.03%	三特异	37.38	39.15	38.26	1.74
	0.01%	三特异	40.00	40.00	40.00	0.00
	QC	三特异	40.00	40.00	40.00	0.00
文库样本	1%	三特异	24.59	25.00	24.79	15.21
	0.10%	三特异	28.34	28.76	28.55	11.45
	0.03%	三特异	30.94	30.28	30.61	9.39
	0.01%	三特异	33.44	32.59	33.02	6.98
	QC	三特异	40.00	40.00	40.00	0.00
文库样本	1%	双特异	24.46	24.60	24.53	15.47
	0.10%	双特异	27.77	27.74	27.75	12.25
	0.03%	双特异	30.11	30.95	30.53	9.47
	0.01%	双特异	32.50	32.88	32.69	7.31
	QC	双特异	40.00	40.00	40.00	0.00
文库样本	1%	单特异	24.97	23.98	24.48	14.27
	0.10%	单特异	27.32	27.40	27.36	11.39
	0.03%	单特异	29.60	29.84	29.72	9.03
	0.01%	单特异	32.08	32.22	32.15	6.60
	QC	单特异	38.22	39.28	38.75	0.00

结论:

实验结果表明, 封闭探针对于本发明的环化富集方法至关重要, 能够显著提高 cfDNA 中稀有分子的检出率, 可将文库分子的检出率提升 10 倍以上, 原始样本的检出率一般在 0.1%, 而文库样本的检出率可达到 0.01%。

本实施例比较了传统基因三特异性体系检测方案, 和双特异性体系及单特异性两种引物/探针的体系。对单一碱基变异的 Q787 位点, 单特异性体系和双特异性体系相比三特异性体系而言, ΔCT 值更大, 具有更高的有效分辨率。对于特异性基因 SRY, 不同特异性引物/探针体系的检测灵敏度都较高, 三种体系都能达到 0.01%。

综上所述, 探针 3'端核苷酸中 3 位羟基的封闭修饰至关重要, 可以防止探针自连, 大幅提高环化效率。结合上模板的探针被切除修饰基团后, 还能进行多个循环的线性扩增, 最大程度的降低了原始分子的损失。另外, APG ssDNA 连接酶的选用, 环化后的酶切处理, 以及环化后使用通用引物扩增也都能起到提高建库产物纯度和检测灵敏度的作用, 分别/共同组成了本发明的技术方案。

1、一种基因目标区域的富集方法，包括：

(1) 通过特异性探针扩增包含目标区域的片段化 DNA，以提供捕获延伸产物；

其特征在于，所述特异性探针包括与所述片段化 DNA 的目标区域互补和不互补的序列，且所述特异性探针的 3'末端和 5'末端核苷酸均被修饰；

(2) 向步骤(1)所提供的捕获延伸产物中加入连接酶，以提供连接产物，所述连接产物包括环化连接产物和线性连接产物。

2、如权利要求 1 所述的基因目标区域的富集方法，其特征在于，

所述步骤(1)中，所述片段化 DNA 包含双链 DNA、单链 DNA 和 cDNA，所述片段化 DNA 的长度为 25~200bp；

和/或，所述步骤(1)的扩增体系中还包括 DNA 聚合酶和 dNTP。

3、如权利要求 2 所述的基因目标区域的富集方法，其特征在于，

所述 DNA 聚合酶具有 3'-5'外切酶活性；

和/或，所述 dNTP 还偶联有标记分子，所述标记分子优选为生物素。

4、如权利要求 1 所述的基因目标区域的富集方法，其特征在于，所述步骤(1)的扩增体系中还包括：

活性物质，所述活性物质在所述特异性探针结合所述片段化 DNA 的目标区域后，用于切除所述特异性探针的 3'末端修饰基团；

优选的，所述活性物质是核酸酶。

5、如权利要求 1 所述的基因目标区域的富集方法，其特征在于，

所述特异性探针还包括能被测序系统识别的通用序列；

和/或，所述特异性探针的 3'末端核苷酸的 3 位羟基被取代；

和/或，所述特异性探针的 3'末端核苷酸的 2 位甲氧基被取代；

优选的，所述特异性探针的 3'末端的取代基团选自氢原子、C3 Spacer 基团、C6 Spacer 基团、磷酸基团或氨基基团；

和/或，所述特异性探针 3'端尾部区域包含错配碱基；

和/或，所述特异性探针的 5'末端核苷酸的 5 位羟基被取代；

优选的，所述特异性探针的 5'末端的取代基团选自磷酸基团或腺苷基团。

6、如权利要求 1 所述的基因目标区域的富集方法，其特征在于，

所述步骤(2)的连接体系中包括单链连接酶,所述单链连接酶优选为 T4 RNA 连接酶或热稳定性 RNA 连接酶。

7、如权利要求 1 所述的基因目标区域的富集方法,其特征在于,

(3) 在所述特异性探针上与所述目标区域不互补的序列中设置酶切位点,并在步骤(2)后加入内切酶,用于在所述酶切位点切割步骤(2)提供的连接产物中的环化连接产物以提供线性连接产物;

优选的,所述酶切位点是尿嘧啶,所述内切酶是 ssDNA 内切酶或 USER 酶。

8、如权利要求 7 所述的基因目标区域的富集方法,其特征在于,还包括:

(4) 在所述步骤(2)后,PCR 扩增所述步骤(2)提供的连接产物;

优选的,所述步骤(4)中,PCR 扩增引物具有与所述特异性探针互补的序列,且与所述目标区域的序列不互补;

更优选的,与所述特异性探针互补的序列为测序通用序列;和/或,

(5) 在所述步骤(3)后,PCR 扩增所述步骤(2)和步骤(3)提供的连接产物;

优选的,所述步骤(5)中,PCR 扩增引物具有与所述酶切位点两侧的序列互补的序列;

更优选的,与所述酶切位点两侧的序列互补的序列为测序通用序列;

和/或,在在所述步骤(1)至(5)的任意步骤后进行产物纯化。

9、如权利要求 8 所述的基因目标区域的富集方法,其特征在于,还包括:

(6) 用检测引物 1、检测引物 2 和探针 3 检测步骤(4)或(5)提供的扩增产物,其特征在于,所述检测引物 1、检测引物 2 和探针 3 中,至少其一包含基因特异性序列;

优选的,所述检测引物 1 包含基因特异性序列,所述检测引物 2 和探针 3 包含通用序列;

和/或,所述检测引物 2 和/或探针 3 也包含基因特异性序列;

优选的,所述探针 3 包含标记分子,且所述探针 3 的序列与检测引物 1 或 2 不互补。

10、如权利要求 1 所述的基因目标区域的富集方法,其特征在于,所述基因目标区域的富集方法应用于核酸检测。

11、一种用于富集片段化 DNA 目标区域的体系,包括适用于如权利要求 1~10

任一权利要求所述的基因目标区域的富集方法的特异性探针和连接酶。

- 12、如权利要求 11 所述用于富集片段化 DNA 目标区域的体系,其特征在于,还包括以下组分的一种或多种: 偶联标记分子的 dNTP, DNA 聚合酶, 核酸酶, 内切酶;

和/或, 还包括 PCR 扩增引物, 所述扩增引物具有与所述特异性探针的通用序列中至少部分互补的序列;

和/或, 还包括检测引物 1、检测引物 2 和探针 3, 三者中至少其一含有基因特异性序列。

经修改的权利要求

国际局收到日：2020年09月29日（29.09.2020）

1、一种基因目标区域的富集方法，包括：

（1）通过特异性探针扩增包含目标区域的片段化 DNA，以提供捕获延伸产物；
其特征在于，所述特异性探针包括与所述片段化 DNA 的目标区域互补和不互补的序列，且所述特异性探针的 3'末端和 5'末端核苷酸均被修饰；

（2）向步骤（1）所提供的捕获延伸产物中加入连接酶，以提供连接产物，所述连接产物包括环化连接产物和线性连接产物；

所述步骤（1）的扩增体系中还包括：

活性物质，所述活性物质在所述特异性探针结合所述片段化 DNA 的目标区域后，用于切除所述特异性探针的 3'末端修饰基团；

优选的，所述活性物质是核酸酶；

所述特异性探针的 3'末端核苷酸的 3 位羟基被取代；

和/或，所述特异性探针的 3'末端核苷酸的 2 位甲氧基被取代；

优选的，所述特异性探针的 3'末端的取代基团选自氢原子、C3 Spacer 基团、C6 Spacer 基团、磷酸基团或氨基基团。

2、如权利要求 1 所述的基因目标区域的富集方法，其特征在于，

所述步骤（1）中，所述片段化 DNA 包含双链 DNA、单链 DNA 和 cDNA，

所述片段化 DNA 的长度为 25~200bp；

和/或，所述步骤（1）的扩增体系中还包括 DNA 聚合酶和 dNTP。

3、如权利要求 2 所述的基因目标区域的富集方法，其特征在于，

所述 DNA 聚合酶具有 3'-5'外切酶活性；

和/或，所述 dNTP 还偶联有标记分子，所述标记分子优选为生物素。

4、如权利要求 1 所述的基因目标区域的富集方法，其特征在于，

所述特异性探针还包括能被测序系统识别的通用序列；

所述特异性探针 3'端尾部区域包含错配碱基；

和/或，所述特异性探针的 5'末端核苷酸的 5 位羟基被取代；

优选的，所述特异性探针的 5'末端的取代基团选自磷酸基团或腺苷基团。

5、如权利要求 1 所述的基因目标区域的富集方法，其特征在于，

所述步骤（2）的连接体系中包括单链连接酶，所述单链连接酶优选为 T4 RNA

连接酶或热稳定性 RNA 连接酶。

6、如权利要求 1 所述的基因目标区域的富集方法，其特征在于，

(3) 在所述特异性探针上与所述目标区域不互补的序列中设置酶切位点，并在步骤 (2) 后加入内切酶，用于在所述酶切位点切割步骤 (2) 提供的连接产物中的环化连接产物以提供线性连接产物；

优选的，所述酶切位点是尿嘧啶，所述内切酶是 ssDNA 内切酶或 USER 酶。

7、如权利要求 6 所述的基因目标区域的富集方法，其特征在于，还包括：

(4) 在所述步骤 (2) 后，PCR 扩增所述步骤 (2) 提供的连接产物；

优选的，所述步骤 (4) 中，PCR 扩增引物具有与所述特异性探针互补的序列，且与所述目标区域的序列不互补；

更优选的，与所述特异性探针互补的序列为测序通用序列；和/或，

(5) 在所述步骤 (3) 后，PCR 扩增所述步骤 (2) 和步骤 (3) 提供的连接产物；

优选的，所述步骤 (5) 中，PCR 扩增引物具有与所述酶切位点两侧的序列互补的序列；

更优选的，与所述酶切位点两侧的序列互补的序列为测序通用序列；

和/或，在在所述步骤 (1) 至 (5) 的任意步骤后进行产物纯化。

8、如权利要求 7 所述的基因目标区域的富集方法，其特征在于，还包括：

(6) 用检测引物 1、检测引物 2 和探针 3 检测步骤 (4) 或 (5) 提供的扩增产物，其特征在于，所述检测引物 1、检测引物 2 和探针 3 中，至少其一包含基因特异性序列；

优选的，所述检测引物 1 包含基因特异性序列，所述检测引物 2 和探针 3 包含通用序列；

和/或，所述检测引物 2 和/或探针 3 也包含基因特异性序列；

优选的，所述探针 3 包含标记分子，且所述探针 3 的序列与检测引物 1 或 2 不互补。

9、如权利要求 1 所述的基因目标区域的富集方法，其特征在于，所述基因目标区域的富集方法应用于核酸检测。

10、一种用于富集片段化 DNA 目标区域的体系，包括适用于如权利要求 1~9 任一权利要求所述的基因目标区域的富集方法的特异性探针和连接酶。

11、如权利要求 10 所述用于富集片段化 DNA 目标区域的体系，其特征在于，还包括以下组分的一种或多种：偶联标记分子的 dNTP，DNA 聚合酶，核酸酶，内切酶；

和/或，还包括 PCR 扩增引物，所述扩增引物具有与所述特异性探针的通用序列中至少部分互补的序列；

和/或，还包括检测引物 1、检测引物 2 和探针 3，三者中至少其一含有基因特异性序列。

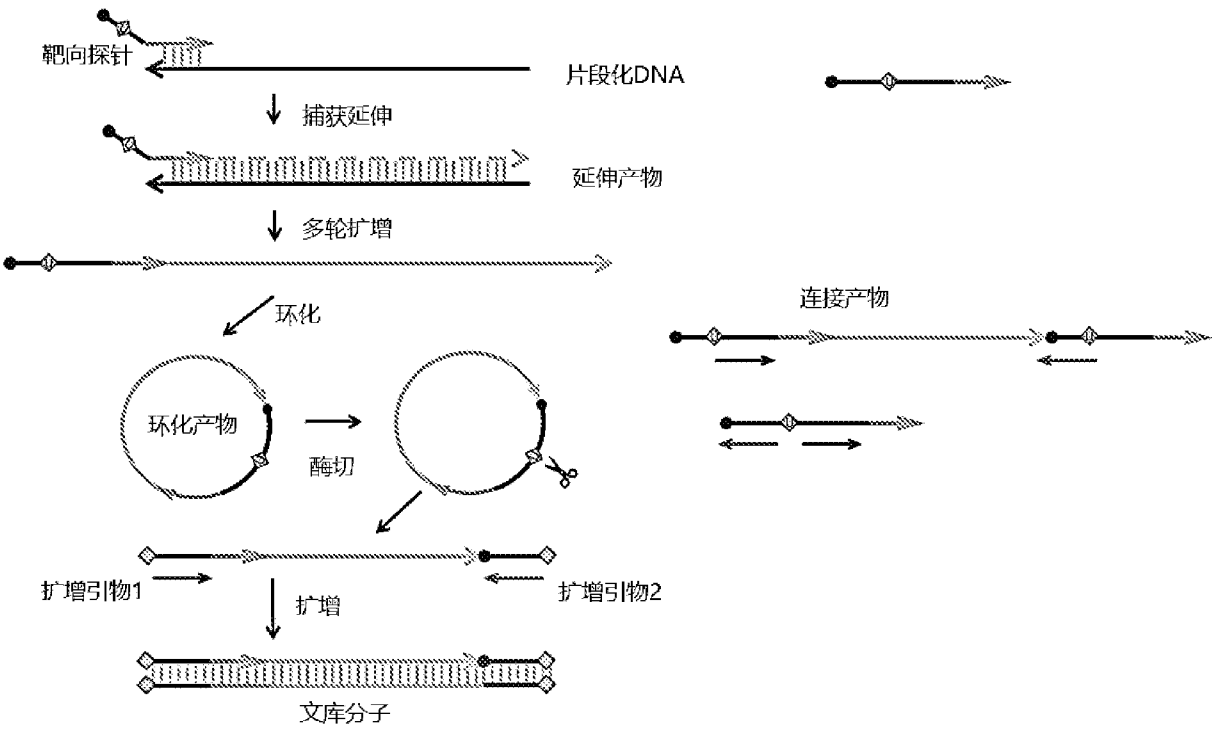


图 1

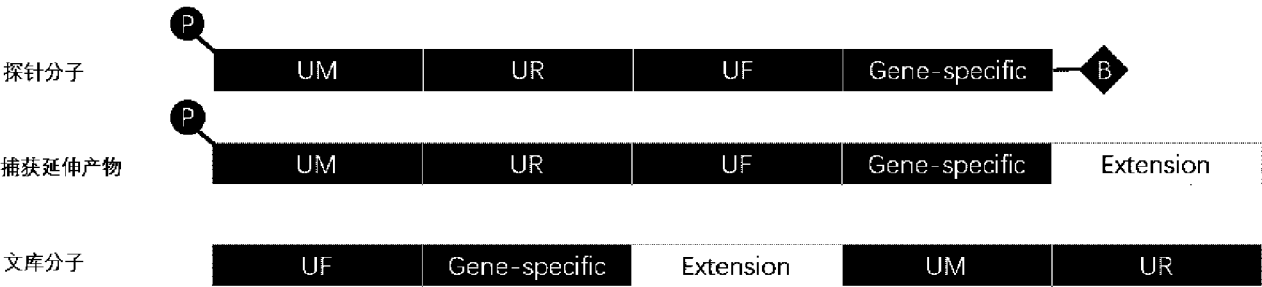


图 2

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2019/123892

A. CLASSIFICATION OF SUBJECT MATTER

C12Q 1/6806(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; TWABS; MOABS; TWMED; VEN; HKABS; CNABS; CPRSABS; SIPOABS; DWPI; VEN; CNTXT; EPTXT; USTXT; WOTXT; ISI-WEB OF SCIENCE; PUBMED; ELSEVIER; CNKI, WANFANG: 连接酶, 富集, 探针, 捕获, 延伸, 建库, 测序, 线性, 环状, 片段, cfDNA, extend+, ligase, NGS, sequencing, polymerase, probe, bp, library, capture, enrich

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 109593757 A (AMOY DIAGNOSTICS CO., LTD.) 09 April 2019 (2019-04-09) see abstract, claims 1-10, description paragraphs 58-76	1-12
A	CN 108396057 A (CHONGQING CANCER INSTITUTE) 14 August 2018 (2018-08-14) entire document	1-12
A	CN 102373288 B (SHENG, Sitong et al.) 11 December 2013 (2013-12-11) entire document	1-12
A	CN 105780129 A (BGI TIANJIN CO., LTD. et al.) 20 July 2016 (2016-07-20) entire document	1-12

☐ 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

“&” document member of the same patent family

Date of the actual completion of the international search

15 June 2020

Date of mailing of the international search report

23 June 2020

Name and mailing address of the ISA/CN

China National Intellectual Property Administration (ISA/CN)
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/CN2019/123892

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/CN2019/123892

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	109593757	A	09 April 2019	WO	2019062289	A1	04 April 2019
CN	108396057	A	14 August 2018	None			
CN	102373288	B	11 December 2013	CN	102373288	A	14 March 2012
CN	105780129	A	20 July 2016	CN	105780129	B	11 June 2019

国际检索报告

国际申请号

PCT/CN2019/123892

A. 主题的分类 C12Q 1/6806(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;TWABS;MOABS;TWMED;VEN;HKABS;CNABS;CPRSABS;SIPOABS;DWPI;VEN;CNTXT;EPTXT;USTXT;WOTXT;ISI-WEB OF SCIENCE;PUBMED;ELSEVIER;CNKI,WANFANG:连接酶, 富集, 探针, 捕获, 延伸, 建库, 测序, 线性, 环状, 片段, cfdna, extend+, ligase, NGS, sequencing, polymerase, probe, bp, library, capture, enrich																	
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>X</td> <td>CN 109593757 A (厦门艾德生物医药科技股份有限公司) 2019年 4月 9日 (2019 - 04 - 09) 参见摘要, 权利要求1-10、说明书第58-76段</td> <td>1-12</td> </tr> <tr> <td>A</td> <td>CN 108396057 A (重庆市肿瘤研究所) 2018年 8月 14日 (2018 - 08 - 14) 全文</td> <td>1-12</td> </tr> <tr> <td>A</td> <td>CN 102373288 B (盛司潼等) 2013年 12月 11日 (2013 - 12 - 11) 全文</td> <td>1-12</td> </tr> <tr> <td>A</td> <td>CN 105780129 A (天津华大基因科技有限公司等) 2016年 7月 20日 (2016 - 07 - 20) 全文</td> <td>1-12</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 109593757 A (厦门艾德生物医药科技股份有限公司) 2019年 4月 9日 (2019 - 04 - 09) 参见摘要, 权利要求1-10、说明书第58-76段	1-12	A	CN 108396057 A (重庆市肿瘤研究所) 2018年 8月 14日 (2018 - 08 - 14) 全文	1-12	A	CN 102373288 B (盛司潼等) 2013年 12月 11日 (2013 - 12 - 11) 全文	1-12	A	CN 105780129 A (天津华大基因科技有限公司等) 2016年 7月 20日 (2016 - 07 - 20) 全文	1-12
类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求															
X	CN 109593757 A (厦门艾德生物医药科技股份有限公司) 2019年 4月 9日 (2019 - 04 - 09) 参见摘要, 权利要求1-10、说明书第58-76段	1-12															
A	CN 108396057 A (重庆市肿瘤研究所) 2018年 8月 14日 (2018 - 08 - 14) 全文	1-12															
A	CN 102373288 B (盛司潼等) 2013年 12月 11日 (2013 - 12 - 11) 全文	1-12															
A	CN 105780129 A (天津华大基因科技有限公司等) 2016年 7月 20日 (2016 - 07 - 20) 全文	1-12															
☐ 其余文件在C栏的续页中列出。 ☒ 见同族专利附件。																	
<table border="0"> <tr> <td> * 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 </td> <td> “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 </td> </tr> </table>			* 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													
* 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 2020年 6月 15日		国际检索报告邮寄日期 2020年 6月 23日															
ISA/CN的名称和邮寄地址 中国国家知识产权局(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 田园 电话号码 62411047															

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/CN2019/123892

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	109593757	A	2019年 4月 9日	WO	2019062289	A1	2019年 4月 4日
CN	108396057	A	2018年 8月 14日	无			
CN	102373288	B	2013年 12月 11日	CN	102373288	A	2012年 3月 14日
CN	105780129	A	2016年 7月 20日	CN	105780129	B	2019年 6月 11日