

(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
2016 年 6 月 23 日 (23.06.2016)



(10) 国际公布号
WO 2016/095736 A1

- (51) 国际专利分类号:
C12N 15/11 (2006.01)
- (21) 国际申请号: PCT/CN2015/096790
- (22) 国际申请日: 2015 年 12 月 9 日 (09.12.2015)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201410796113.7 2014 年 12 月 18 日 (18.12.2014) CN
- (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)。深圳华大基因科技有限公司 (BGI SHENZHEN CO., LIMITED) [CN/CN]; 中国广东省深圳市盐田区北山路 146 号北山工业区 11 栋 2 楼, Guangdong 518083 (CN)。
- (72) 发明人: 郭晶 (GUO, Jing); 中国广东省深圳盐田区北山工业区综合楼, Guangdong 518083 (CN)。郭荣荣 (GUO, Rongrong); 中国广东省深圳市盐田区北山工业区综合楼, Guangdong 518083 (CN)。李梅艳 (LI, Meiyang); 中国广东省深圳市盐田区北山工业区综合楼, Guangdong 518083 (CN)。耿春雨 (GENG,

Chunyu); 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)。
- (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, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, 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,

[见续页]

(54) Title: TARGET REGION ENRICHMENT METHOD BASED ON MULTIPLEX PCR, AND REAGENT

(54) 发明名称: 一种基于多重 PCR 的目标区域富集方法和试剂

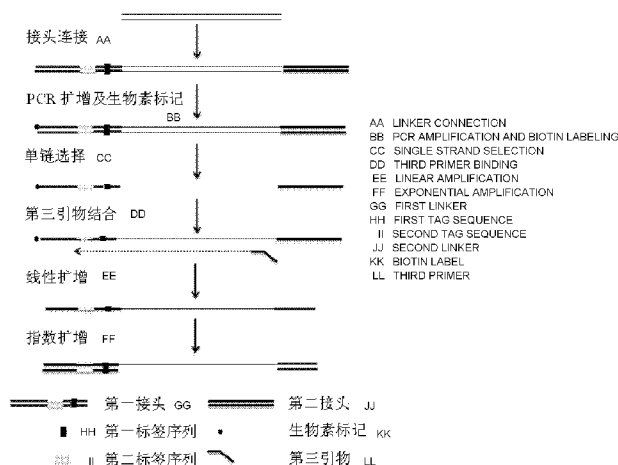


图 1 / FIG 1

(57) Abstract: Provided are a target region enrichment method based on multiplex PCR, and a reagent, the method comprising: connecting a first linker and a second linker respectively at two ends of a nucleic acid segment containing target regions to be enriched so as to obtain a linker-connected product; performing a PCR amplification on the linker-connected product using a first primer specifically bound to the first linker and a second primer specifically bound to the second linker to obtain an amplified product, the first primer or the second primer having a first affinity label; capturing a single strand having the first affinity label in the amplified product using a solid phase carrier; performing single primer linear amplification using a third primer with the captured single strand as a template; performing exponential amplification using the third primer and the first primer, with the linearly amplified product as the template, to obtain a product containing the target regions.

(57) 摘要:

[见续页]

WO 2016/095736 A1

CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, **本国际公布:**

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)。

- 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))。

提供一种基于多重 PCR 的目标区域富集方法和试剂，该方法包括：在含待富集目标区域的核酸片段两端分别连接第一接头和第二接头，得到接头连接产物；使用特异性结合第一接头的第一引物和特异性结合第二接头的第二引物，对接头连接产物进行 PCR 扩增，得到扩增产物，其中第一引物或第二引物具有第一亲和标记；通过固相载体捕获扩增产物中带有第一亲和标记的单链；以捕获的单链为模板，使用第三引物进行单引物线性扩增；以线性扩增产物为模板，使用第三引物和第一引物进行指数扩增，得到包含目标区域的产物。

一种基于多重 PCR 的目标区域富集方法和试剂

技术领域

本发明涉及分子生物学技术领域，尤其涉及一种基于多重 PCR 的目标区域富集方法和试剂。

背景技术

随着 DNA 测序技术的发展，高通量测序技术已被广泛地应用于生命科学研究的各个领域。尽管随着测序技术的不断更新和普及，测序技术的成本也越来越低，但全基因组测序技术本身的花费还是很昂贵。用于调和这一问题的较好的方案是将感兴趣的目标区域富集出来，再进行高通量测序。传统的序列捕获技术，通常是将高通量测序文库构建好后，利用探针进行目标区域的文库富集再测序。

除此之外 Life Tech 公司基于多重 PCR 技术开发了一种称为 Ampliseq 的试剂盒，可以通过多重 PCR 过程实现目标区域的富集，从而大大缩减了杂交捕获这个耗时的操作过程，可以称为区域捕获技术的一项革新。

但是 Ampliseq 对于细胞游离 DNA 的目标区域富集却无能为力，因为细胞游离 DNA 具有片段小的特点，其长度甚至已经小于 Ampliseq 试剂盒 PCR 产物的长度；除此之外，基于 Ampliseq 的区域捕获技术只有在 PCR 后才能够实现样本之间的混合建库，这将使一个比较耗时的过程。

发明内容

本发明提供一种基于多重 PCR 的目标区域富集方法和试剂，能够实现细胞游离 DNA 这种短片段的区域捕获，并且大大减小区域捕获的时间和成本。

根据本发明的第一方面，本发明提供一种基于多重 PCR 的目标区域富集方法，包括如下步骤：

在连接酶作用下，在含待富集目标区域的核酸片段两端分别连接第一接头和第二接头，得到接头连接产物，其中第一接头包括用于后续 PCR 扩增的第一接头固定序列、用于标记不同样本的第一标签序列和用于标记不同目标区域分子序列的第二标签序列；

使用特异性结合第一接头固定序列的第一引物和特异性结合第二接头的第二引物，对接头连接产物进行 PCR 扩增，得到扩增产物，其中第一引物或第二

引物具有第一亲和标记；

通过固相载体捕获扩增产物中带有第一亲和标记的单链，其中固相载体上带有能够与第一亲和标记亲和结合的第二亲和标记；

以捕获的单链为模板，使用第三引物进行单引物线性扩增，其中第三引物包括 5'端的第三引物固定序列和 3'端的目标区域特异性结合序列；

以线性扩增产物为模板，使用第三引物和第一引物进行指数扩增，得到包含目标区域的产物。

作为本发明的优选方案，第一标签序列为 5-10 个碱基长度的序列，第二标签序列为 10-12 个碱基长度的随机序列。

作为本发明的优选方案，含待富集目标区域的核酸片段为末端修复过的核酸片段。

作为本发明的优选方案，第一引物具有第一亲和标记。

作为本发明的优选方案，第一亲和标记为生物素标记，第二亲和标记为链霉亲和素标记，固相载体为磁珠。

根据本发明的第二方面，本发明提供一种基于多重 PCR 的目标区域富集试剂，包括如下组成部分：

第一接头和第二接头，用于在连接酶作用下分别连接到含待富集目标区域的核酸片段两端，以得到接头连接产物，其中第一接头包括用于后续 PCR 扩增的第一接头固定序列、用于标记不同样本的第一标签序列和用于标记不同目标区域分子序列的第二标签序列；

第一引物和第二引物，用于对接头连接产物进行 PCR 扩增，得到扩增产物，其中第一引物特异性结合第一接头固定序列，第二引物特异性结合第二接头，并且第一引物或第二引物具有第一亲和标记；

固相载体，用于捕获扩增产物中带有第一亲和标记的单链，其中固相载体上带有能够与第一亲和标记亲和结合的第二亲和标记；

第三引物，用于以捕获的单链为模板，进行单引物线性扩增，其中第三引物包括 5'端的第三引物固定序列和 3'端的目标区域特异性结合序列；

其中，第三引物和第一引物还用于以线性扩增产物为模板，进行指数扩增，得到包含目标区域的产物。

作为本发明的优选方案，第一标签序列为 5-10 个碱基长度的序列，第二标

签序列为 10-12 个碱基长度的随机序列。

作为本发明的优选方案，含待富集目标区域的核酸片段为末端修复过的核酸片段。

作为本发明的优选方案，第一引物具有第一亲和标记。

作为本发明的优选方案，第一亲和标记为生物素标记，第二亲和标记为链霉亲和素标记，固相载体为磁珠。

本发明基于多重 PCR 的目标区域富集方法，首先将连接有标签序列的产物进行目标区域的线性扩增富集，然后通过特定引物（第三引物）将目标区域进行 PCR 富集，从而使得多重 PCR 不再受限于短片段核酸序列的影响，能实现细胞游离 DNA 这种短片段分子的区域捕获；连接有特定标签序列（第一标签序列）的产物可以在目标区域富集之前，就实现多样本之间的混合操作，这大大降低了样本处理的成本，并提高样本处理的通量。

此外，本发明除对不同样本间进行不同标签（第一标签序列）的标记以外，还使用高重数目分子标签（第二标签序列）对同一样本内不同目标区域来源的分子进行标记。基于高重数目分子标签可以对是否 PCR 扩增产生的相同分子进行识别，从而在高深度测序下可以分辨低突变率（如千分之一突变率）的分子序列和正常序列，以便于对低突变的样本进行筛查和研究。

附图说明

图 1 为本发明基于多重 PCR 的目标区域富集方法的一个实施方案的原理示意图。

具体实施方式

下面通过具体实施方式对本发明作进一步详细说明。除非特别说明，下面实施例中所使用的技术均为本领域内的技术人员已知的常规技术；所使用的仪器设备和试剂等，均为本领域内的技术人员可以通过公共途径如商购等获得的。

本发明中，任何情况下使用的“第一”、“第二”和“第三”等概念都不应当理解为具有顺序和技术的含义，其作用仅在于将其与其它对象区别开来。

请参考图 1，本发明一个实施方案中，基于多重 PCR 的目标区域富集方法的主要步骤包括：接头连接；PCR 扩增及生物素标记；单链选择；第三引物结合；线性扩增；和指数扩增。具体地：

接头连接，即在连接酶作用下，在含待富集目标区域的核酸片段两端分别

连接第一接头和第二接头，得到接头连接产物，其中第一接头包括用于后续 PCR 扩增的第一接头固定序列、用于标记不同样本的第一标签序列和用于标记不同目标区域分子序列的第二标签序列。

上述核酸片段可以是任何来源的含待富集目标区域的核酸片段，典型但非限定性的例子比如：基因组 DNA 经物理打断或转座酶包埋复合体打断得到的随机片段、PCR 扩增产物、全基因组扩增（whole genome amplification, WGA）产物等，特别是细胞游离 DNA 这种短片段分子也可作为本发明的核酸片段用于富集目标区域。

上述方法产生的核酸片段可能需要进行末端修复，尤其是物理打断产生的核酸片段，需要使用 DNA 聚合酶进行末端补平，使用 DNA 磷酸化激酶进行磷酸化处理，以便于后续接头连接反应顺利进行。

第一接头和第二接头均是双链短核苷酸序列，长度从十几个碱基对到几十个碱基对不等。

对于第一接头，其第一接头固定序列位于序列 5'端，用于后续 PCR 扩增，即作为后续 PCR 扩增的引物退火位点，第一标签序列和第二标签序列位于第一接头固定序列的下游（即 3'端），至于第一标签序列和第二标签序列的顺序（或者位置）可以不限定，即第一标签序列可以位于第二标签序列的上游（即 5'端）或下游（即 3'端），反之亦然。所谓“固定序列”，是指这一段序列是固定不变的，而其它部分序列可以变化，因此“固定序列”的概念也可以称作“不变序列”或“通用序列”等。此外，第一接头固定序列、第一标签序列和第二标签序列中两个相邻的序列可以直接相连，也可以有几个碱基的间隔，即通过“连接序列”连接起来形成第一接头。

第一标签序列用于标记不同样本，也就是说同一样本来源的核酸片段经接头连接以后会带上相同的第一标签序列，而不同样本来源的核酸片段经接头连接以后会带上不同的第一标签序列，从而使得本发明不像传统方法那样只有在通过 PCR 对目标区域富集后才能够进行不同样本之间的混合建库，而是可以在目标区域富集之前就进行多样本之间的混合操作，这大大降低了样本处理的成本，并可以提高样本处理的通量。

第二标签序列用于标记不同目标区域分子序列，即同一样本内的目标区域核酸片段经接头连接以后会带上不同的第二标签序列，由于第二标签序列是随

机序列，长度一般在 10-12 个碱基之间，理论上对于具有 N 个碱基的随机序列，第二标签序列的种类应该有 4^N 种，足以标记数以百万计以上的目标区域。因此，本发明这种基于高重数目分子标签（第二标签序列）可以对是否 PCR 扩增产生的相同分子进行识别，从而在高深度测序下可以分辨低突变率（如千分之一突变率）的分子序列和正常序列，以便于对低突变的样本进行筛查和研究。

基于上述说明，本发明中第一标签序列、第二标签序列这样的概念并不是指一个具有特定序列的序列段，而是基于样本来源和核酸片段分子来源而有变化的一组“标签集”，相应的第一接头也并非一个具有特定序列的序列段，而是指一组“接头集”。然而，第二接头却是一个具有特定序列的序列段，在本发明中是通用的，用于 PCR 扩增富集目标区域。

虽然接头连接步骤的意图在于在含待富集目标区域的核酸片段两端分别连接第一接头和第二接头，但是由于第一接头和第二接头可能由于不具有选择性连接的能力，而导致两个第一接头或两个第二接头同时连接在同一个核酸片段两端的情况发生。尽管如此，也不影响本发明的有效性，因为接头连接以后的 PCR 扩增步骤中使用的第一引物特异性结合第一接头固定序列，而第二引物特异性结合第二接头，从而将那些期望的两端分别连接第一接头和第二接头的核酸片段富集起来。当然，也可以通过对连接方式的设计实现第一接头和第二接头分别特异性连接到核酸片段的两端，比如可以采用的连接方式有：平端连接或粘端连接，如果第一接头和第二接头分别采用不同的连接方式，即可避免同一核酸片段两端连接同一接头的情况发生。

PCR 扩增及生物素标记，即使用特异性结合第一接头固定序列的第一引物和特异性结合第二接头的第二引物，对接头连接产物进行 PCR 扩增，得到扩增产物，其中第一引物或第二引物具有第一亲和标记。

本发明的一个实施方案中，以生物素作为第一亲和标记来标记第一引物或第二引物；相应地，以链霉亲和素作为第二亲和标记来标记固相载体。生物素-链霉亲和素系统是生物学中应用最广泛、效果最优异的生物亲和标记方法。当然，本发明并不局限于此。作为第一亲和标记，可以是生物学上常用的生物结合反应的一个组成部分，比如抗原或抗体，双链 DNA 短片段的一条链，生物素或链霉亲和素，等等。在第一亲和标记选用了抗原的情况下，第二亲和标记选用与该抗原结合的抗体，反之亦然；在第一亲和标记选用了双链 DNA 短片段的

一条链的情况下，第二亲和标记选用与该链互补配对的另一条链，反之亦然；在第一亲和标记选用了生物素的情况下，第二亲和标记选用与生物素结合的链霉亲和素，反之亦然。

本发明的一个实施方案中，固相载体是磁珠，作为固相载体还可以选用微孔板、芯片等，只要能将第二亲和标记（如链霉亲和素）固定到相应的固相载体即可，本发明最优选的是磁珠。

单链选择，即通过固相载体捕获扩增产物中带有第一亲和标记的单链，其中固相载体上带有能够与第一亲和标记亲和结合的第二亲和标记。

本发明的一个实施方案中，固相载体为磁珠，第一亲和标记为生物素标记，第二亲和标记为链霉亲和素标记。因此，使用链霉亲和素标记的磁珠通过链霉亲和素与生物素的结合，来选择 PCR 扩增产物中带有生物素标记的单链，再用于下游目标区域富集操作。

第三引物结合和线性扩增，即以捕获的单链为模板，使用第三引物进行单引物线性扩增，其中第三引物包括 5'端的第三引物固定序列和 3'端的目标区域特异性结合序列。

该步骤中，第三引物固定序列，是指这一段序列是固定不变的，而其它部分序列可以变化，因此“固定序列”的概念也可以称作“不变序列”或“通用序列”等。第三引物的 3'端的目标区域特异性结合序列，能够特异性结合待富集目标区域，再通过 PCR 扩增实现目标区域的富集。

指数扩增，即以线性扩增产物为模板，使用第三引物和第一引物进行指数扩增，得到包含目标区域的产物。至此，实现了目标区域的富集。

目标区域的富集后产物可以应用于后续高通量测序或其它分析平台。本发明的方法，也同样可以应用于其它各种组织样本，甚至石蜡包埋甲醛固定的医学样本。

下面通过实施例详细说明本发明。

本实施例中所用的接头和引物序列如表 1 所示。

表 1 接头和引物序列

名称	序列	SEQ ID NO:
第一接头	CCATCTCATCCCTGCGTGTCTCCGACTCAGNNNNNN	1

	NNNNCTAAGGTAACGAT (正义链)	
	ATCGTTACCTTAGNNNNNNNNNNCTGAGTCGGAGA CACGCAGGGATGAGATGGTT (反义链)	2
第二接头	AAGTCGGAGGCCAAGCGGTCTTAGGA (正义链)	3
	TCCTAAGACCGCTTGGCCTCCGACT (反义链)	4
第一引物	/5-bio/CCATCTCATCCCTGCGTGTC	5
第二引物	TCCTAAGACCGCTTGGCCTCCGACT	6
第三引物	CCACTACGCCTCCGCTTTCCTCTCTATGGGCAGTCGG TGATT <u>TCTATGGTGTGTCCCCCAACT</u>	7
	CCACTACGCCTCCGCTTTCCTCTCTATGGGCAGTCGG TGATT <u>GGGAGCCCCCGTTCTATATCA</u>	8
	CCACTACGCCTCCGCTTTCCTCTCTATGGGCAGTCGG TGATT <u>GCGCCACAGAGAAGTTGTTGA</u>	9
	CCACTACGCCTCCGCTTTCCTCTCTATGGGCAGTCGG TGATT <u>TCTCTCGGAGGAAGGACTTGAG</u>	10
	CCACTACGCCTCCGCTTTCCTCTCTATGGGCAGTCGG TGATT <u>GAGAGAACGCGGAATTGGTCTA</u>	11
	CCACTACGCCTCCGCTTTCCTCTCTATGGGCAGTCGG TGATT <u>TTTCAATTATCATCTTTGTTCATCAGCTGAA</u>	12
	CCACTACGCCTCCGCTTTCCTCTCTATGGGCAGTCGG TGATT <u>AGCCCCAGTGATCTTCCAGATA</u>	13
	CCACTACGCCTCCGCTTTCCTCTCTATGGGCAGTCGG TGATT <u>TTTAAATACAGAAGGCATAAATATTCCAGCAGA</u>	14
	CCACTACGCCTCCGCTTTCCTCTCTATGGGCAGTCGG TGATT <u>TGCTATTCTCAGATGACTCTGTGTTTTT</u>	15
	CCACTACGCCTCCGCTTTCCTCTCTATGGGCAGTCGG TGATT <u>GTGTTTGACTCTAGATGCTGTGAGAA</u>	16
	CCACTACGCCTCCGCTTTCCTCTCTATGGGCAGTCGG TGATT <u>GCTCATACAGCAGGCCATAGA</u>	17

CCACTACGCCTCCGCTTTCCTCTCTATGGGCAGTCGG TGATT <u>TCACCTTCAGAAGTCACAGAATGATTTT</u>	18
CCACTACGCCTCCGCTTTCCTCTCTATGGGCAGTCGG TGATT <u>AAAAAGCCATTTGACCGTGGAG</u>	19
CCACTACGCCTCCGCTTTCCTCTCTATGGGCAGTCGG TGATT <u>CTATATGTAGAGGCTGTTGGAAGCTG</u>	20
CCACTACGCCTCCGCTTTCCTCTCTATGGGCAGTCGG TGATT <u>CATACTTACCATGCCACTTTCCCTT</u>	21
CCACTACGCCTCCGCTTTCCTCTCTATGGGCAGTCGG TGATT <u>AGATTGCACCGGTCGACAAA</u>	22
CCACTACGCCTCCGCTTTCCTCTCTATGGGCAGTCGG TGATT <u>CACCAGCGTGTCCAGGAA</u>	23
CCACTACGCCTCCGCTTTCCTCTCTATGGGCAGTCGG TGATT <u>TCCACCATGACTTTGAGGTTGAG</u>	24
CCACTACGCCTCCGCTTTCCTCTCTATGGGCAGTCGG TGATT <u>ACTGTTTCGTATTTATAGCTGATTTGATGGA</u>	25
CCACTACGCCTCCGCTTTCCTCTCTATGGGCAGTCGG TGATT <u>ACGTCTTCCTTCTCTCTCTGTCA</u>	26
CCACTACGCCTCCGCTTTCCTCTCTATGGGCAGTCGG TGATT <u>GACTATGTCCGGGAACACAAAGA</u>	27

注：①第一接头序列中下划线部分为用于标记不同样本的第一标签序列，“N”部分为用于标记不同目标区域分子序列的第二标签序列，其它部分为用于后续 PCR 扩增的第一接头固定序列（除与第一标签序列相邻的三个碱基以外）；②第三引物序列中下划线部分为 3'端的目标区域特异性结合序列，其它部分为 5'端的第三引物固定序列；③所有序列均按 5'至 3'形式列出。

本实施例具体实验过程如下：

1、将外周血肿瘤细胞样本 DNA（100ng）进行末端修复后通过平末端或者粘末端方式连接特殊序列的接头（第一接头和第二接头），并进行 5-7 个循环的扩增。其中，PCR 扩增的一条引物（第一引物）上带有生物素修饰或其它可以进行特异链分离的标记。

1.1 末端修复反应体系如下表 2 所示，20℃反应 30min。

表 2

组分	用量 (μL)
片段化DNA	X (100ng)
T4DNA聚合酶(NEB公司)	5
T4PNK (NEB公司)	5
dNTP混合物	1.6
T4 PNK buffer (NEB公司) (10×)	10
总量	100

1.2连接反应体系如下表3所示，20℃反应30min。

表3

组分	用量 (μL)
末端修复后DNA	X
T4DNA连接酶 (NEB公司)	1
T4DNA连接酶缓冲液 (NEB公司)	5
第一接头	5
第二接头	5
ddH ₂ O	余量
总量	50

1.3 PCR扩增体系如下表4所示：

表4

组分	用量 (μL)
DNA连接产物	X
dNTP(25mM)	3
Platinum pfx DNA polymerase (life technology)	0.6
第一引物	3
第二引物	3
MgSO ₄	4
10×pfx buffer	10

ddH ₂ O	余量
总量	100

PCR扩增程序如下表5所示：

表5

温度	时间	循环
72℃	20min	1 个循环
95℃	5min	1 个循环
95℃	30s	7 个循环
58℃	30s	
70℃	1min	
72℃	5min	1 个循环
4℃	∞	1 个循环

2、取上述连接有特殊接头（第一接头和第二接头）的PCR扩增产物1pmol，用链霉亲和素磁珠进行单链DNA的捕获，单链的捕获过程采用0.1M NaOH洗脱的方式。

3、针对上述捕获的生物素标记的PCR扩增产物单链，采用靶向结合目标区域的单引物（第三引物）进行线性扩增。线性扩增体系和程序分别如下表6和表7所示：

表6

组分	用量（μL）
捕获的单链DNA	6
5× Ion AmpliSeq™ HiFi Mix（life technology）	4
第三引物	10
总量	20

表7

温度	时间	循环
95℃	3min	1 个循环
95℃	30s	40 个循环
65℃	1min	
70℃	5min	1 个循环

4、将线性扩增后的产物进行多重指数扩增，扩增引物的上游引物为1.3步骤的PCR扩增体系中生物素标记链特异匹配的引物（第一引物），下游引物为线性扩增的引物组（第三引物）。多重指数扩增的体系和程序如下表8和表9所示：

表8

组分	用量（ μL ）
线性扩增后的DNA	6
5×Ion AmpliSeq™ HiFi Mix（life technology）	4
第一引物和第三引物	10
总量	20

表9

温度	时间	循环
95℃	3min	1 个循环
95℃	30s	20 个循环
65℃	1min	
70℃	5min	1 个循环

5、最终的多重指数扩增后的产物纯化后进行上机测序分析，分析内容包括目标区域富集比例，覆盖均一性以及变异等内容。结果如下表10所示：

表10

参数	数值
样本	外周血肿瘤细胞样本
总的读取数目	1253091
靶数目	1009600
靶数目比例	80.56%
靶数目覆盖度	99.97%
目标区域捕获效率	66%
单碱基突变数目	43893

注：总的读取数目是指测序中读的到总的读段数目总数；靶数目指目标区域读取的读段数目总数；靶数目比例是指目标区域总数占总读取数目的比值；靶数目覆盖度指目标区域覆盖区域比例的情况；目标区域捕获效率指目标区域占整个参考基因组目标区域的比例情况；单碱基突变数目指目标区域检测到的

单个碱基突变数目的情况。

本申请还提供如下所示的另一实施例，以表明本方法的可行性。该实施例除了以肺癌样本代替外周血肿瘤细胞样本以外，其它方面包括实验过程和材料与上面的实施例相同。结果如下表 11 所示：

表 11

参数	数值
样本	肺癌样本
总的读取数目	3163794
平均测序深度	32204.13
靶数目覆盖度	100.00%
目标区域捕获效率	89.10%

表 12 示出了在一个基因位置（chr7:55242469）上，采用常规杂交捕获建库和本发明的基于多重 PCR 的目标区域富集方法得到的等位基因频率的结果。

表 12

肺癌样本					杂交捕获方法	多重富集方法
基因位置	基因序列	基因型	变异	基因	等位基因频率	等位基因频率
chr7:55242469	TTAAGAGAAG CAACATCTC	TTAAGAGAAG CAACATCTC/T	INDEL	EGFR	0.185397 28	0.27706

表 11 和表 12 的结果表明，本发明的基于多重 PCR 的目标区域富集方法在目标区域捕获效率上高达 89.10%，能检出与常规杂交捕获建库同样的结果，且频率大致相近并且结果一致，由此可以确定本发明的方法可以应用于目标区域富集。

以上内容是结合具体的实施方式对本发明所作的进一步详细说明，不能认定本发明的具体实施只局限于这些说明。对于本发明所属技术领域的普通技术人员来说，在不脱离本发明构思的前提下，还可以做出若干简单推演或替换。

权 利 要 求 书

1. 一种基于多重 PCR 的目标区域富集方法，包括如下步骤：

在连接酶作用下，在含待富集目标区域的核酸片段两端分别连接第一接头和第二接头，得到接头连接产物，其中所述第一接头包括用于后续 PCR 扩增的第一接头固定序列、用于标记不同样本的第一标签序列和用于标记不同目标区域分子序列的第二标签序列；

使用特异性结合所述第一接头固定序列的第一引物和特异性结合所述第二接头的第二引物，对所述接头连接产物进行 PCR 扩增，得到扩增产物，其中所述第一引物或第二引物具有第一亲和标记；

通过固相载体捕获所述扩增产物中带有所述第一亲和标记的单链，其中所述固相载体上带有能够与所述第一亲和标记亲和结合的第二亲和标记；

以捕获的所述单链为模板，使用第三引物进行单引物线性扩增，其中所述第三引物包括 5'端的第三引物固定序列和 3'端的目标区域特异性结合序列；

以所述线性扩增产物为模板，使用所述第三引物和所述第一引物进行指数扩增，得到包含目标区域的产物。

2. 根据权利要求 1 所述的基于多重 PCR 的目标区域富集方法，其特征在于，所述第一标签序列为 5-10 个碱基长度的序列，所述第二标签序列为 10-12 个碱基长度的随机序列。

3. 根据权利要求 1 所述的基于多重 PCR 的目标区域富集方法，其特征在于，所述含待富集目标区域的核酸片段为末端修复过的核酸片段。

4. 根据权利要求 1 所述的基于多重 PCR 的目标区域富集方法，其特征在于，所述第一引物具有第一亲和标记。

5. 根据权利要求 1 所述的基于多重 PCR 的目标区域富集方法，其特征在于，所述第一亲和标记为生物素标记，所述第二亲和标记为链霉亲和素标记，所述固相载体为磁珠。

6. 一种基于多重 PCR 的目标区域富集试剂，包括如下组成部分：

第一接头和第二接头，用于在连接酶作用下分别连接到含待富集目标区域的核酸片段两端，以得到接头连接产物，其中所述第一接头包括用于后续 PCR 扩增的第一接头固定序列、用于标记不同样本的第一标签序列和用于标记不同目标区域分子序列的第二标签序列；

第一引物和第二引物，用于对所述接头连接产物进行 PCR 扩增，得到扩增

产物，其中所述第一引物特异性结合所述第一接头固定序列，所述第二引物特异性结合所述第二接头，并且所述第一引物或第二引物具有第一亲和标记；

固相载体，用于捕获所述扩增产物中带有所述第一亲和标记的单链，其中所述固相载体上带有能够与所述第一亲和标记亲和结合的第二亲和标记；

第三引物，用于以捕获的所述单链为模板，进行单引物线性扩增，其中所述第三引物包括 5'端的第三引物固定序列和 3'端的目标区域特异性结合序列；

其中，所述第三引物和所述第一引物还用于以所述线性扩增产物为模板，进行指数扩增，得到包含目标区域的产物。

7. 根据权利要求 6 所述的基于多重 PCR 的目标区域富集试剂，其特征在于，所述第一标签序列为 5-10 个碱基长度的序列，所述第二标签序列为 10-12 个碱基长度的随机序列。

8. 根据权利要求 6 所述的基于多重 PCR 的目标区域富集试剂，其特征在于，所述含待富集目标区域的核酸片段为末端修复过的核酸片段。

9. 根据权利要求 6 所述的基于多重 PCR 的目标区域富集试剂，其特征在于，所述第一引物具有第一亲和标记。

10. 根据权利要求 6 所述的基于多重 PCR 的目标区域富集试剂，其特征在于，所述第一亲和标记为生物素标记，所述第二亲和标记为链霉亲和素标记，所述固相载体为磁珠。

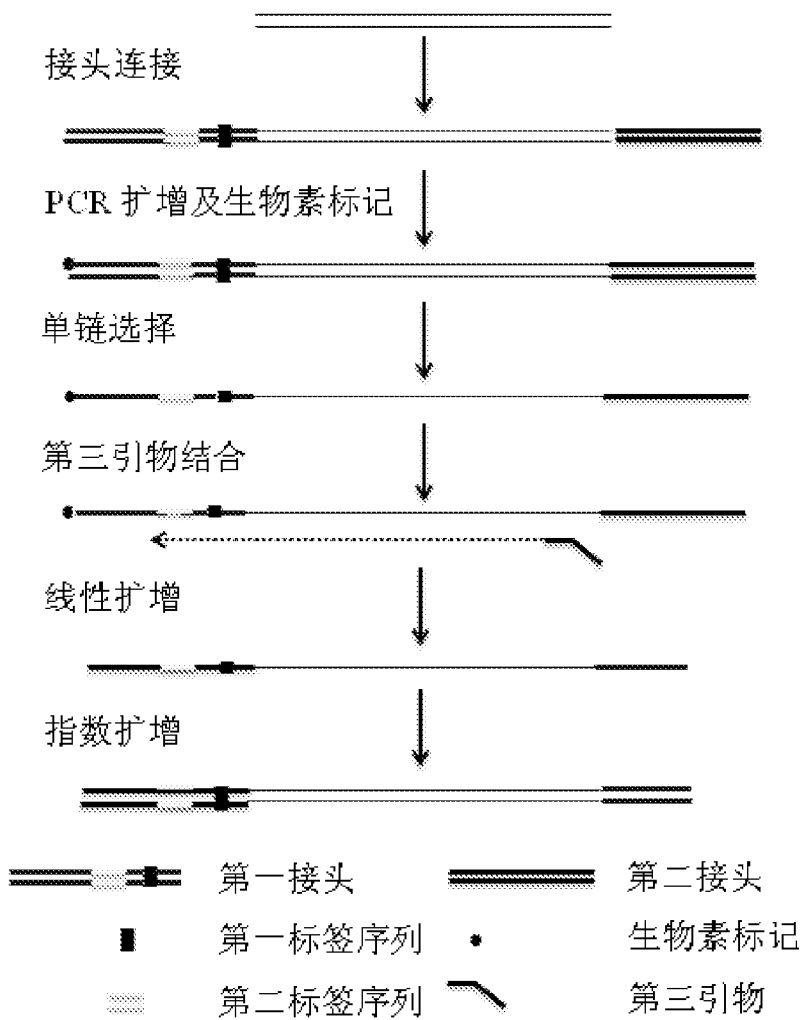


图 1

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2015/096790

A. CLASSIFICATION OF SUBJECT MATTER

C12N 15/11 (2006.01) 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)

C12N

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)

CPRSABS, CNABS, SIPOABS, WPI, EPODOC, CNKI, CA, PUBMED: PCR, adapter, primer

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN 102329876 A (SHENZHEN BGI TECHNOLOGY CO., LTD., BGI SHENZHEN) 25 January 2012 (25.01.2012) , see the abstract, claims 1-14	1-10
A	CN 102628082 A (KAIJING BIOTECHNOLOGY SUZHOU CO., LTD.) 08 August 2012 (08.08.2012) , see the whole document	1-10
A	CN 102533985 A (SHENZHEN BGI TECHNOLOGY CO., LTD., BGI SHENZHEN) 04 July 2012 (04.07.2012) , see the whole document	1-10
A	CN 101815792 A (COLUMBIA NEW YORK UNIVERSITY) 25 August 2010 (25.08.2010) , see the whole document	1-10
A	CN 101421410 A (COLUMBIA NEW YORK UNIVERSITY) 29 April 2009 (29.04.2009) , see the whole document	1-10

☐ 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	"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
"E" earlier application or patent but published on or after the international filing date	"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
"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)	"&" document member of the same patent family
"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	

Date of the actual completion of the international search 24 February 2016	Date of mailing of the international search report 04 March 2016
Name and mailing address of the ISA State Intellectual Property Office of the P. R. China No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing 100088, China Facsimile No. (86-10) 62019451	Authorized officer SONG, Zhigang Telephone No. (86-10) 62411078

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2015/096790

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 102329876 A	25 January 2012	WO 2013053180 A1	18 April 2013
		WO 2013053182 A1	18 April 2013
		WO 2013053183 A1	18 April 2013
		WO 2013053207 A1	18 April 2013
		CN 102329876 B	02 April 2014
		CN 103874767 A	18 June 2014
		CN 103890189 A	25 June 2014
		US 2014249038 A1	04 September 2014
CN 102628082 A	08 August 2012	CN 102628082 B	17 September 2014
CN 102533985 A	04 July 2012	WO 2013091276 A1	27 June 2013
		CN 102533985 B	06 August 2014
CN 101815792 A	25 August 2010	WO 2009002891 A1	31 December 2008
		EP 2162555 A1	17 March 2010
		KR 20100058449 A	03 June 2010
		JP 2010530760 A	16 September 2010
		US 2010240549 A1	23 September 2010
		US 2010240549 A1	23 September 2010
CN 101421410 A	29 April 2009	WO 2007103910 A2	13 September 2007
		WO 2007103910 A3	29 November 2007
		EP 1994164 A2	26 November 2008
		KR 20080107464 A	10 December 2008
		US 2009203002 A1	13 August 2009
		JP 2009529330 A	20 August 2009
		JP 2009529330 A	20 August 2009

国际检索报告

国际申请号

PCT/CN2015/096790

A. 主题的分类

C12N 15/11 (2006.01) 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)

C12N

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))

CPRSABS、CNABS、SIPOABS、WPI、EPDOC、CNKI、CA、PUBMED: PCR, 接头, 引物, adapter, primer

C. 相关文件

类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求
A	CN 102329876 A (深圳华大基因科技有限公司, 深圳华大基因研究院) 2012年 1月 25日 (2012 - 01 - 25) 参见摘要、权利要求1-14	1-10
A	CN 102628082 A (晶生物科技苏州有限公司) 2012年 8月 8日 (2012 - 08 - 08) 参见全文	1-10
A	CN 102533985 A (深圳华大基因科技有限公司, 深圳华大基因研究院) 2012年 7月 4日 (2012 - 07 - 04) 参见全文	1-10
A	CN 101815792 A (纽约市哥伦比亚大学托管会) 2010年 8月 25日 (2010 - 08 - 25) 参见全文	1-10
A	CN 101421410 A (纽约市哥伦比亚大学托管会) 2009年 4月 29日 (2009 - 04 - 29) 参见全文	1-10

☐ 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

2016年 2月 24日

国际检索报告邮寄日期

2016年 3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

授权官员

宋智刚

传真号 (86-10) 62019451

电话号码 (86-10) 62411078

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2015/096790

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	102329876	A	2012年 1月 25日	WO	2013053180	A1	2013年 4月 18日
				WO	2013053182	A1	2013年 4月 18日
				WO	2013053183	A1	2013年 4月 18日
				WO	2013053207	A1	2013年 4月 18日
				CN	102329876	B	2014年 4月 2日
				CN	103874767	A	2014年 6月 18日
				CN	103890189	A	2014年 6月 25日
				US	2014249038	A1	2014年 9月 4日
CN	102628082	A	2012年 8月 8日	CN	102628082	B	2014年 9月 17日
CN	102533985	A	2012年 7月 4日	WO	2013091276	A1	2013年 6月 27日
				CN	102533985	B	2014年 8月 6日
CN	101815792	A	2010年 8月 25日	WO	2009002891	A1	2008年 12月 31日
				EP	2162555	A1	2010年 3月 17日
				KR	20100058449	A	2010年 6月 3日
				JP	2010530760	A	2010年 9月 16日
				US	2010240549	A1	2010年 9月 23日
CN	101421410	A	2009年 4月 29日	WO	2007103910	A2	2007年 9月 13日
				WO	2007103910	A3	2007年 11月 29日
				EP	1994164	A2	2008年 11月 26日
				KR	20080107464	A	2008年 12月 10日
				US	2009203002	A1	2009年 8月 13日
				JP	2009529330	A	2009年 8月 20日

表 PCT/ISA/210 (同族专利附件) (2009年7月)