

(19)中华人民共和国国家知识产权局



(12)发明专利申请

(10)申请公布号 CN 108085395 A

(43)申请公布日 2018. 05. 29

(21)申请号 201810156947.X

(22)申请日 2018.02.24

(71)申请人 韩林志

地址 410154 湖南省长沙市开福区沙坪街
道中青路1048号山河医药健康产业园
标准厂房第17栋4楼西头

(72)发明人 韩林志 肖芳

(51)Int.Cl.

C12Q 1/6886(2018.01)

C12Q 1/6869(2018.01)

C12Q 1/6806(2018.01)

权利要求书2页 说明书12页
序列表6页 附图1页

(54)发明名称

基于高通量测序的宫颈癌多基因甲基化检测的引物组、试剂盒及方法

(57)摘要

本发明涉及一种基于高通量测序的宫颈癌多基因甲基化检测的引物组。所述引物组为用于特异性扩增十个目标扩增子的引物组合,包括分别针对FAM19A4、miR124-2、C13ORF18、EPB41L3、JAM3、SST、ZIC1、GHSR、GFRA1及ANKRD18CP扩增子的上游引物及下游引物。本发明涉及一种基于高通量测序的宫颈癌多基因甲基化检测的试剂盒。所述试剂盒包括含有上述引物的PCR多重扩增反应液。本发明还涉及一种基于高通量测序的宫颈癌多基因甲基化检测的方法。本发明提供的试剂盒及其检测方法具有多基因甲基化联合检测优势、灵敏性高、特异性好、通量高、检测时间快速的优点。

1. 一种基于高通量测序的宫颈癌多基因甲基化检测的引物组, 其特征在于, 包括分别针对如SEQ ID NO.1所示FAM19A4扩增子、如SEQ ID NO.2所示miR124-2扩增子、如SEQ ID NO.3所示C13ORF18扩增子、如SEQ ID NO.4所示EPB41L3扩增子、如SEQ ID NO.5所示JAM3扩增子、如SEQ ID NO.6所示SST扩增子、如SEQ ID NO.7所示ZIC1扩增子、如SEQ ID NO.8所示GHSR扩增子、如SEQ ID NO.9所示GFRA1扩增子及如SEQ ID NO.10所示ANKRD18CP扩增子的引物组合, 其中:

FAM19A4上游引物: 如SEQ ID NO.11所示,
FAM19A4下游引物: 如SEQ ID NO.12所示,
miR124-2上游引物: 如SEQ ID NO.13所示,
miR124-2下游引物: 如SEQ ID NO.14所示,
C13ORF18上游引物: 如SEQ ID NO.15所示,
C13ORF18下游引物: 如SEQ ID NO.16所示,
EPB41L3上游引物: 如SEQ ID NO.17所示,
EPB41L3下游引物: 如SEQ ID NO.18所示,
JAM3上游引物: 如SEQ ID NO.19所示,
JAM3下游引物: 如SEQ ID NO.20所示,
SST上游引物: 如SEQ ID NO.21所示,
SST下游引物: 如SEQ ID NO.22所示,
ZIC1上游引物: 如SEQ ID NO.23所示,
ZIC1下游引物: 如SEQ ID NO.24所示,
GHSR上游引物: 如SEQ ID NO.25所示,
GHSR下游引物: 如SEQ ID NO.26所示,
GFRA1上游引物: 如SEQ ID NO.27所示,
GFRA1下游引物: 如SEQ ID NO.28所示,
ANKRD18CP上游引物: 如SEQ ID NO.29所示,
ANKRD18CP下游引物: 如SEQ ID NO.30所示。

2. 一种基于高通量测序的宫颈癌多基因甲基化检测的试剂盒, 其特征在于, 包括含有如权利要求1所述的上游引物及下游引物的PCR多重扩增反应液, 所述PCR多重扩增反应液包括: AmpliSeq HiFi Mix、如权利要求1所述的SEQ ID NO.11-30引物依次等比混合组成的AmpliSeq Primer Pool及无核酸酶水。

3. 根据权利要求2所述的基于高通量测序的宫颈癌多基因甲基化检测的试剂盒, 其特征在于, 所述PCR多重扩增反应液的终成分为: $2 \times \text{AmpliSeq HiFi Mix}$ $10\mu\text{l}$ 、AmpliSeq Primer Pool $2\mu\text{l}$ 、目标基因组DNA 40ng - 70ng 、其余无核酸酶水补足至 $20\mu\text{l}$; 其中AmpliSeq Primer Pool为SEQ ID NO.11-30引物依次按摩尔比 $1:1:2:2:1:1:4:4:2:2:1:1:2:2:4:4:2:2:2:2$ 混合组成。

4. 根据权利要求2-3中任一所述的基于高通量测序的宫颈癌多基因甲基化检测的试剂盒, 其特征在于, 还包括阳性对照品及阴性对照品, 所述阳性对照品为宫颈癌细胞株CaSki, 所述阴性对照品为宫颈癌细胞株C33A。

5. 一种基于高通量测序的宫颈癌多基因甲基化检测的方法, 其特征在于, 包括如下步

骤:

步骤一:取待检测样本抽提的DNA,对其进行转化处理,转化后的DNA作为PCR的模板;

步骤二:提供如权利要求4所述的试剂盒,对所述模板进行PCR多重扩增,并在扩增产物两端接上合适的接头,完成目标扩增子测序文库构建;

步骤三:对所述扩增子测序文库进行质检,进而对样本的十个靶基因进行高通量测序分析,根据检测结果对样本进行判断。

6.根据权利要求5所述的基于高通量测序的宫颈癌多基因甲基化检测的方法,其特征在于,所述步骤一中转化处理所采用的试剂为亚硫酸盐;所述步骤二中PCR多重扩增反应程序为:99℃2min;25cycles,99℃15sec,60℃4min。

基于高通量测序的宫颈癌多基因甲基化检测的引物组、试剂盒及方法

技术领域

[0001] 本发明涉及体外诊断技术领域,特别的,涉及一种基于高通量测序的宫颈癌多基因甲基化检测的引物组、试剂盒及方法。

背景技术

[0002] 宫颈癌一直是困扰女性的重要恶性肿瘤。正常的宫颈细胞经过癌前病变发展成为宫颈癌,可能长达几十年。在这整个过程中部分HPV病毒的感染可被自身免疫系统修复而自动转归为正常,而另外一部分与宫颈癌联系更紧密的癌前病变则需及时进行干预来阻止癌前病变发展成为宫颈癌。目前我国临床上应用的主要筛查手段为中国女性的健康做出巨大贡献的同时也都存在其不可避免的缺陷。

[0003] 首先,目前的方法是沿用了大半个世纪的细胞学检测方法,包括巴氏涂片法及近十多年发展起来的液基细胞学(TCT)。因为整个检测流程中从取样、抹片制备、观察、结果判读多个环节受到人为影响的可能性较大,造成漏检的几率较高。

[0004] 其次,自从宫颈癌归因于HPV病毒持续感染后发展起来的HPV-DNA检测。因为其检测的仅仅是HPV病毒的存在与否,而HPV病毒的存在,大部分是一过性的,也就是意味着这些被自体免疫清除掉的HPV感染也会被判定为阳性,导致过度治疗。

[0005] 所以,发明人期望从传统的形态病理向深层次的分子病理发展,从仅仅检测治病病毒的感染与否向检测是否真正发生了病变发展,从以上角度挖掘出能与癌前病变紧密相关的分子病理的变化指标,以指示患癌的风险。

[0006] 大量研究证明,随着宫颈癌病变进展,宿主相关细胞的甲基化水平升高。宫颈癌基因甲基化检测,就是通过检测患者自身宫颈细胞中DNA甲基化水平,与细胞学和HPV检测联合应用,可更加准确的预判患者患癌风险、进行早期筛查、做好分流管理和术后追踪,进一步为妇女健康保驾护航。

[0007] DNA甲基化是指在DNA甲基转移酶的作用下,S-腺苷甲硫氨酸的甲基(-CH₃)共价结合到DNA分子的胞嘧啶(C)碱基的5位碳原子上,形成5-甲基胞嘧啶(5mC),而并不改变DNA的序列。研究表明,FAM19A4、miR-124-2基因启动子甲基化与宫颈癌的发生和发展密切相关,是宫颈癌特异的分子标记物。因此,检测宫颈癌FAM19A4、miR-124-2基因启动子的甲基化状态,对宫颈癌的间接诊断、治疗、预后判断等方面具有重要意义。

[0008] 随着高通量测序技术的发展,测序成本的降低,使测序技术逐渐成为基础生物学研究和医学检测的常规实验方法。但全基因组测序结构复杂,数据量大,周期长,费用高,限制了它的发展。扩增子测序(Amplicon Sequencing)是只对目标区域进行测序研究的一项测序方法。通过设计感兴趣的基因组区域的引物,进行PCR扩增,对目标区域进行富集,然后针对特定长度的PCR产物或者捕获的片段进行建库,高通量测序,分析序列中的变异。扩增子测序能根据研究者目的,针对目标区域进行高覆盖度的测序,还可以检测到低频突变。采用扩增子测序,研究者可以对基因组中感兴趣的关键区域进行重点研究。这种高针对性的

方法可以高效的发现、验证及筛选关注区域的基因组变异。相比全基因组测序,扩增子测序对目标区域有更准确快速的研究,适用于大量样本的特定基因组区域研究。

[0009] 目前传统的甲基化检测方法包括:甲基化特异性PCR及亚硫酸氢盐测序法。然而这些方法存在通量过低、操作繁琐、准确性低及敏感性不高的缺点,限制了其在临床实验室的广泛应用;此外,市面上未见有针对FAM19A4、miR124-2、C13ORF18、EPB41L3、JAM3、SST、ZIC1、GHSR、GFRA1及ANKRD18CP十个目标扩增子靶基因甲基化联合检测的试剂盒产品。

[0010] 因此,有必要提供一种新的基于高通量测序的宫颈癌多基因甲基化检测的引物组、试剂盒及方法来克服上述缺陷。

发明内容

[0011] 为了解决上述传统甲基化检测方法存在通量过低、操作繁琐、准确性低及敏感性不高的技术问题,本发明提供一种操作简单、准确性高及敏感性高且基于高通量二代测序法的宫颈癌多基因甲基化检测的引物组、试剂盒及其方法。

[0012] 本发明提供了一种基于高通量测序的宫颈癌多基因甲基化检测的引物组,包括分别针对如SEQ ID NO.1所示FAM19A4扩增子、如SEQ ID NO.2所示miR124-2扩增子、如SEQ ID NO.3所示C13ORF18扩增子、如SEQ ID NO.4所示EPB41L3扩增子、如SEQ ID NO.5所示JAM3扩增子、如SEQ ID NO.6所示SST扩增子、如SEQ ID NO.7所示ZIC1扩增子、如SEQ ID NO.8所示GHSR扩增子、如SEQ ID NO.9所示GFRA1扩增子及如SEQ ID NO.10所示ANKRD18CP扩增子的引物组合,其中:

[0013] FAM19A4上游引物:如SEQ ID NO.11所示,

[0014] FAM19A4下游引物:如SEQ ID NO.12所示,

[0015] miR124-2上游引物:如SEQ ID NO.13所示,

[0016] miR124-2下游引物:如SEQ ID NO.14所示,

[0017] C13ORF18上游引物:如SEQ ID NO.15所示,

[0018] C13ORF18下游引物:如SEQ ID NO.16所示,

[0019] EPB41L3上游引物:如SEQ ID NO.17所示,

[0020] EPB41L3下游引物:如SEQ ID NO.18所示,

[0021] JAM3上游引物:如SEQ ID NO.19所示,

[0022] JAM3下游引物:如SEQ ID NO.20所示,

[0023] SST上游引物:如SEQ ID NO.21所示,

[0024] SST下游引物:如SEQ ID NO.22所示,

[0025] ZIC1上游引物:如SEQ ID NO.23所示,

[0026] ZIC1下游引物:如SEQ ID NO.24所示,

[0027] GHSR上游引物:如SEQ ID NO.25所示,

[0028] GHSR下游引物:如SEQ ID NO.26所示,

[0029] GFRA1上游引物:如SEQ ID NO.27所示,

[0030] GFRA1下游引物:如SEQ ID NO.28所示,

[0031] ANKRD18CP上游引物:如SEQ ID NO.29所示,

[0032] ANKRD18CP下游引物:如SEQ ID NO.30所示。

[0033] 本发明提供了一种基于高通量测序的宫颈癌多基因甲基化检测的试剂盒,包括含有上述的上游引物及下游引物的PCR多重扩增反应液,所述PCR多重扩增反应液包括:AmpliSeq HiFi Mix、上述的SEQ ID NO.11-30引物依次等比混合组成的AmpliSeq Primer Pool及无核酸酶水。

[0034] 在本发明提供的所述基于高通量测序的宫颈癌多基因甲基化检测的试剂盒一较佳实施例中,所述PCR多重扩增反应液的终成分为:2×AmpliSeq HiFi Mix 10μl、AmpliSeq Primer Pool 2μl、目标基因组DNA 40ng-70ng、其余无核酸酶水补足至20μl;其中AmpliSeq Primer Pool为SEQ ID NO.11-30引物依次按摩尔比1:1:2:2:1:1:4:4:2:2:1:1:2:2:4:4:2:2:2:2混合组成。

[0035] 在本发明提供的所述基于高通量测序的宫颈癌多基因甲基化检测的试剂盒一较佳实施例中,还包括阳性对照品及阴性对照品,所述阳性对照品为宫颈癌细胞株CaSki,所述阴性对照品为宫颈癌细胞株C33A。

[0036] 本发明还提供了一种基于高通量测序的宫颈癌多基因甲基化检测的方法,包括如下步骤:

[0037] 步骤一:取待检测样本抽提的DNA,对其进行转化处理,转化后的DNA作为PCR的模板;

[0038] 步骤二:提供上述的试剂盒,对所述模板进行PCR多重扩增,并在扩增产物两端接上合适的接头,完成目标扩增子测序文库构建;

[0039] 步骤三:对所述扩增子测序文库进行质检,进而对样本的十个靶基因进行高通量测序分析,根据检测结果对样本进行判断。

[0040] 在本发明提供的所述基于高通量测序的宫颈癌多基因甲基化检测的方法一较佳实施例中,所述步骤一中转化处理所采用的试剂为亚硫酸盐;所述步骤二中PCR多重扩增反应程序为:99℃2min;25cycles,99℃15sec,60℃4min。

[0041] 相较于现有技术,本发明提供的基于高通量测序的宫颈癌多基因甲基化检测的引物组、试剂盒及方法的有益效果在于:

[0042] 一、通过设计特异性高的引物组,并配置成使用方便且检测结果可靠的试剂盒,再设计出科学合理的多重PCR反应体系,使得本发明具有快速、高通量、灵敏及特异性好的特点,实现对宫颈癌FAM19A4、miR124-2、C13ORF18、EPB41L3、JAM3、SST、ZIC1、GHSR、GFRA1及ANKRD18CP十个目标扩增子靶基因的甲基化程度快速及准确测量,以便对宫颈癌间接进行及时、有效的诊断和治疗,降低医疗成本,节约社会资源。

[0043] 二、通过设计用于特异性扩增目标扩增子的引物组,通过将接头序列、测序引物序列、Barcode、目标片段扩增引物序列这四种序列整合在上、下游扩增的一对引物上,使得利用该引物能够通过一步扩增完成文库构建,不仅提高了甲基化测序文库质量和建库的效率,而且构建所得的扩增子测序文库因具有常规测序平台上所使用的接头序列及测序引物序列,使得测序文库具有常规测序所用的测序试剂和引物进行测序。本发明相同的样本具有相同的barcode,不同的样本具有不同的barcode。一次可以对多个样本多个基因进行测序分析,相比一代测序提高了测序的速度。

[0044] 三、基于高通量二代测序的甲基化DNA检测技术,对病人宫颈细胞10个靶基因甲基化进行检测,所述的目标扩增子来源于多方面基于人群的验证和二代高通量筛选后的结

果,所选的目标扩增子均与宫颈癌的进程密切相关。宫颈癌的发生是个多诱因多步骤的过程,本发明的靶扩增子包括FAM19A4、miR124-2、C13ORF18、EPB41L3、JAM3、SST、ZIC1、GHSR、GFRA1、ANKRD18CP,多基因甲基化联合检测综合评价,灵敏度、特异性均要优于单基因检测。此外,甲基化发生并不局限于一两个碱基,高密度甲基化致使蛋白表达明显减低。本发明基于高通量二代测序的甲基化DNA检测技术,每个靶基因的甲基化区域检测均针对核心调控区域多个CpG,相对单个碱基的检测减少了碱基变化的随机性。

附图说明

[0045] 图1为基于多重PCR法设计的引物结构示意图;

[0046] 图2为扩增子测序文库构建完成后扩增产物的分布图。

具体实施方式

[0047] 为了使本发明的目的、技术方案及优点更加清楚明白,下面结合附图及实施例,对本发明进行进一步的详细说明。应当理解,此处所描述的具体实施例仅用以解释本发明,并不用于限定本发明。

[0048] 实施例1:试剂盒的制备

[0049] 一、目标扩增子及其序列

[0050] 目标扩增子包括十个扩增子。所述目标扩增子来源于多方面的文献验证和二代高通量筛选后的结果,所选的目标扩增子均与宫颈癌的进程密切相关。

[0051] 目标扩增子及其序列,如下表所示:

目标 扩增子	扩增子序列	序列号
FAM19A4 扩增子	YGGGYGGTTYGGTTAATTAGYGYGTTTYGYGGYG GGTTYGGTTTTTTTTGGTTAGYGYGTTAGTTGGGTT YGGTTTYGTATTGTAGTTGYGYGTYGTTTTGGAY GGGGYGTATYGATTGYGYGYGYGGTTGYGGGTAA ATATYG	SEQ ID NO.1
miR124-2 扩增子	AYGTYGTTATTTGAAAATTAGATTTATAGGTTTATGT ATGTTTTTAGGYGTGTGTTGTAAATGGTATGGAGAT ATATGTATATGTATAYGTAGGTATAYGTATYGTTTATA TTTTTAYGGAATAGATTAATTAATAGYGG	SEQ ID NO.2
C13ORF18 扩增子	TTGTTAGGTTTTTGATTTTTGAAAGYGTYGTTGYGT TTYGYGTYGYGGGTAGGTAGGGYGGGATTTTAGG AGGATYGGTAGAGGYGYGTATAGGTGYGTGGTGTT	SEQ ID NO.3

[0053]

	GGGTTYGGGYGTYGYGGTATYGGTGTAGGAGYGY GTATTTTATAGAGTTTTTTTATTTGGGYGAYGTTTY GGTGTGTTGYGGYGGGAAYGGYG	
EPB41L3 扩增子	YGGAGGTYGGGGTTTAGGTTTTGYGYGTYGGTTTA GTTATTATTGYGTYGYGGYGGGYGGAGYGGGYGG GGGGYGYGGYGYGTAGGTTYGGTTYGGTGGGGGT TTYGGYGAGYGGGAGGGYGGTTGGGGATTTYGGT YGYGTYGGGYGYGGGGTTYGGGATTY	SEQ ID NO.4
JAM3 扩增子	YGGTTTAGAGTATYGTTGTATTYGTAAGTAGTTAGA TTTAGTTTTTTTTTGTATTATGGTGTGTYGGTTYGGTT GGGTTYGGYGGTYGTTATGGTAATTGGGGYGGGTY GTAGGGTTTTGGTAGGTTGGGYGTATGYGYGYGGG GATTATAAGTYGYGTYGYGTTGTGTYGTTGGTTTTTA GTAATTTTYGATAIGGYGTTGAGGYGGTTATYGYGA TTYGGTTTTGYGTTYGGTTGTTTGATTTTTTTTTG T	SEQ ID NO.5
SST 扩增子	YGGTTTTYGYGGYGTYGAGATGTTGTTTTGTYGTT TTTAGTGYGYGTTGGTTGYGTTGTTTATYGTTTTGG TTTTGGGTTGTGTTATYGGYGTTTTTTYGGATTTTA GATTTYGTTAGTTTTTGTAGAAGTTTTTGGTTGTTG TYGYG	SEQ ID NO.6
ZIC1 扩增子	GTTTTTTTTYGGGTAGTTTTGAYGYGYGGTTTTTY GGTAGAGATTGAGYGGYGAGAAAGTGYGAGTYGG GTYGGTAGAATTTGTTTGGYGGGYGTTGGAGTTTG YGTIATTYGYGGTTYGTAGTYGTTYGGTTATTTIGY GTTTGGTTYGGTTAGYGTTYGGGYGYGTYGY	SEQ ID NO.7
GHSR 扩增子	YGTTYGTTTTTGGTAGTATYGGTTTTGGAATTTYGG YGATTTTTTTTGTAATTTTTTAAATTYGTAGTGAG AGTTGTATTTAYGTTAYGGTGTATTATTATAGYGT TGAGYGTYGAGYGTTATTTYGTTATTTGTTTTTAT TTYG	SEQ ID NO.8
GFRA1 扩增子	YGYGGTYGGTTGTAGAATATTGYGTGTTTTGTGYG YGAGTAYGGGTTYGGTYGYGGGTTTTATYGTAAAGY	SEQ ID NO.9

[0054]		GGGYGTGTGTAGTTGYGGTTGTTTGGYGGGYGAG TTTAGGAGGGGAGTTGGAGTTGYGGAGYGGYGGGA ATAGGAGTAGGTYGAG	
	ANKRD18 CP 扩增子	YGGGAGAYGTTTGGGTAGGYGTTTTTGAGTTTAA TGGATTAAAGAGTAYGYGGGTGTTTATATATY GGGGATTGGGAATTGYGGAAGATTATAGGGTGGT TATTAAGGGYGAYGTYGYGGAGGTGGAGYGTGT TTGATTYGTAGGTTTYGGGATTGGATGTTYGYGAT AGAAAGGATAGGTAGYGGGGGTTTAGTTYGYGGT GGGAGGGGGTTTTAGGTTTTGTTTGTTYGTAGYG TTTGAGGYGGGGGTTTTGGAGGTYGTYGGGTTTTG GAGTYGY	SEQ ID NO.10

[0055] 备注:扩增子序列为亚硫酸处理后的序列;Y为简并碱基,表示C/T。

[0056] 二、用于特异性扩增目标扩增子的引物组设计与合成

[0057] 通过将接头序列、测序引物序列、Barcode、目标片段扩增引物序列整合在上、下游扩增的一对引物上,使得利用该引物能够通过一步扩增完成文库构建,不仅提高了甲基化测序文库质量和建库的效率,而且构建所得的扩增子测序文库因具有常规测序平台上所使用的接头序列及测序引物序列,使得测序文库具有常规测序所用的测序试剂和引物进行测序。参见图1-图2,其中图1为基于多重PCR法设计的引物结构示意图,图2为扩增子测序文库构建完成后扩增产物的分布图。

[0058] 本发明的人宫颈癌特异性甲基化的PCR引物组,是根据NCBI (美国国立生物技术信息中心) 公开的人类全基因组序列,使用Primer Premier3.0及Methyl Primer press v1.0设计。

[0059] 特异性引物组序列,如下表所示:

[0060]

序列名称	寡核苷酸序列	序列号
FAM19A4 上游引物	AATGATACGGCGACCAACCGAGATCTACACCTCTCTATTC GTCGGCAGCGTCAGATGTGTATAAGAGACAGGGGGTTA	SEQ ID NO.11

[0061]

	GGTAGGGATAGGAGTAGT	
FAM19A4 下游引物	CAAGCAGAAGACGGCATAACGAGATAGGCTCCGGTCTCG TGGGCTCGGAGATGTGTATAAGAGACAGGACTACGGGT ATCTAATCCAAACGACAACATAAAACAAAACCTCC	SEQ ID NO.12
miR124-2 上游引物	AATGATACGGCGACCACCGAGATCTACACCTCTCTATTC GTCGGCAGCGTCAGATGTGTATAAGAGACAGATTAGAG GGGTAATTAATTTGGATT	SEQ ID NO.13
miR124-2 下游引物	CAAGCAGAAGACGGCATAACGAGATAGGCTCCGGTCTCG TGGGCTCGGAGATGTGTATAAGAGACAGGACTACGGGT ATCTAATCC TCATCTCTAACACATCTACCAAAA	SEQ ID NO.14
C13ORF18 上游引物	AATGATACGGCGACCACCGAGATCTACACCTCTCTATTC GTCGGCAGCGTCAGATGTGTATAAGAGACAGATTTGGG GTTGGGAAGTTATT	SEQ ID NO.15
C13ORF18 下游引物	CAAGCAGAAGACGGCATAACGAGATAGGCTCCGGTCTCG TGGGCTCGGAGATGTGTATAAGAGACAGGACTACGGGT ATCTAATCC AAAAATCCTCAAAAAAACA AAAA	SEQ ID NO.16
EPB4IL3 上游引物	AATGATACGGCGACCACCGAGATCTACACCTCTCTATTC GTCGGCAGCGTCAGATGTGTATAAGAGACAGGATTTTAG GGAGGTTTTTAGGT	SEQ ID NO.17
EPB4IL3 下游引物	CAAGCAGAAGACGGCATAACGAGATAGGCTCCGGTCTCG TGGGCTCGGAGATGTGTATAAGAGACAGGACTACGGGT ATCTAATCC ATCTCCAATAAACTAATCCC	SEQ ID NO.18
JAM3 上游 引物	AATGATACGGCGACCACCGAGATCTACACCTCTCTATTC GTCGGCAGCGTCAGATGTGTATAAGAGACAGAGTTTATT GAAAGAGAATTTATGTGT	SEQ ID NO.19
JAM3 下游 引物	CAAGCAGAAGACGGCATAACGAGATAGGCTCCGGTCTCG TGGGCTCGGAGATGTGTATAAGAGACAGGACTACGGGT ATCTAATCC AACTCACCCCTAAAAACAACA	SEQ ID NO.20
SST 上游 引物	AATGATACGGCGACCACCGAGATCTACACCTCTCTATTC GTCGGCAGCGTCAGATGTGTATAAGAGACAGGAGTTTG ATTAGTTATTTTTTAGTT	SEQ ID NO.21
SST 下游	CAAGCAGAAGACGGCATAACGAGATAGGCTCCGGTCTCG	SEQ ID

[0062]

引物	TGGGCTCGGAGATGTGTATAAGAGACAGGACTACGGGT ATCTAATCC AAAAAATCTCCTTACCTACTTCCC	NO.22
ZIC1 上游 引物	AATGATACGGCGACCACCGAGATCTACACCTCTCTATTC GTCGGCAGCGTCAGATGTGTATAAGAGACAGTTTTAGG GGGTTGAGATGTTTTAT	SEQ ID NO.23
ZIC1 下游 引物	CAAGCAGAAGACGGCATACGAGATAGGCTCCGGTCTCG TGGGCTCGGAGATGTGTATAAGAGACAGGACTACGGGT ATCTAATCC AAATCCTAACCTACAAACAATAAC	SEQ ID NO.24
GHSR 上 游引物	AATGATACGGCGACCACCGAGATCTACACCTCTCTATTC GTCGGCAGCGTCAGATGTGTATAAGAGACAGTTTTTTTT GTATGTTTTTTGGATT	SEQ ID NO.25
GHSR 下 游引物	CAAGCAGAAGACGGCATACGAGATAGGCTCCGGTCTCG TGGGCTCGGAGATGTGTATAAGAGACAGGACTACGGGT ATCTAATCC CTTAATAACCACCACCTTAACC	SEQ ID NO.26
GFRA1 上 游引物	AATGATACGGCGACCACCGAGATCTACACCTCTCTATTC GTCGGCAGCGTCAGATGTGTATAAGAGACAGAGAGAGT GTGTGTGTTTGGTAGTATT	SEQ ID NO.27
GFRA1 下 游引物	CAAGCAGAAGACGGCATACGAGATAGGCTCCGGTCTCG TGGGCTCGGAGATGTGTATAAGAGACAGGACTACGGGT ATCTAATCC CAAATTTCTTCTAACCACAAAAAACC	SEQ ID NO.28
ANKRD18 CP 上游引 物	AATGATACGGCGACCACCGAGATCTACACCTCTCTATTC GTCGGCAGCGTCAGATGTGTATAAGAGACAGTTATTATG AGGAAATTTTTTAGTTT	SEQ ID NO.29
ANKRD18 CP 下游引 物	CAAGCAGAAGACGGCATACGAGATAGGCTCCGGTCTCG TGGGCTCGGAGATGTGTATAAGAGACAGGACTACGGGT ATCTAATCC AACAACTAAACTCCATTTAACCTC	SEQ ID NO.30

[0063] 备注：其中Barcode序列为下划线部分且非唯一，不同的样本具有不同的Barcode。

[0064] 三、对照品选择

[0065] 阳性对照品为宫颈癌细胞株CaSki；阴性对照品为宫颈癌细胞株C33A。

[0066] 四、PCR反应液组成

[0067] 包括含有上述特异性引物组的PCR多重扩增反应液，所述PCR多重扩增反应液包括：AmpliSeq HiFi Mix、上述的SEQ ID NO.11-30引物依次等比混合组成的AmpliSeq Primer Pool及无核酸酶水；

[0068] 所述PCR多重扩增反应液的终成分为：2×AmpliSeq HiFi Mix 10μl、AmpliSeq

Primer Pool 2 μ l、目标基因组DNA 40ng-70ng、其余无核酸酶水补足至20 μ l；其中AmpliSeq Primer Pool为将引物干粉溶解成100uM，将SEQ ID NO.11-30所示二十条引物依次按摩尔比1:1:2:2:1:1:4:4:2:2:1:1:2:2:4:4:2:2:2:2混合组成。

[0069] 其中，所述AmpliSeq HiFi Mix、AmpliSeq Primer购自Thermo Fisher公司；高通量测序相关的接头、Barcode、测序试剂均购自Illumina公司；所述无核酸酶水为(Nuclease-Free Water)用DEPC(Diethyl pyrocarbonate, 焦碳酸二乙酯)处理过并经高温高压灭菌的超纯水。

[0070] 实施例2:利用上述试剂盒进行宫颈癌甲基化检测的方法

[0071] 一、检测方法

[0072] 步骤一:取待检测样本抽提的DNA,对其进行转化处理,转化后的DNA作为PCR的模板;

[0073] 其中,DNA提取试剂盒为QIAmp DNA Mini Kit(购自德国QIAGEN公司);转化处理所采用的试剂为亚硫酸氢盐或重亚硫酸盐,及其他辅助(相应试剂盒为购自德国QIAGEN公司的EpiTect Fast DNA Bisulfite Kit)。

[0074] 步骤二:提供上述的试剂盒,对所述模板进行PCR多重扩增,并在扩增产物两端接上合适的接头,完成目标扩增子测序文库构建;

[0075] 其中,PCR多重扩增反应程序为:99 $^{\circ}$ C 2min; 25cycles, 99 $^{\circ}$ C 15sec, 60 $^{\circ}$ C 4min。

[0076] 步骤三:对所述扩增子测序文库进行质检,进而对样本的十个靶基因进行高通量测序分析,根据检测结果对样本进行判断。

[0077] 具体检测方法,如下:

[0078] 一)收集生物样本:

[0079] 选自2017年1月-2017年6月期间在湖南省湘雅一医院就诊的53例C1N1宫颈脱落细胞、58例C1N2患者宫颈脱落细胞、27例C1N3+宫颈癌患者宫颈脱落细胞、90例正常人群的宫颈脱落细胞。

[0080] 二)提取组织DNA:

[0081] 提取样本组织DNA所用的试剂成分均来自QIAmp DNA Mini Kit(购自QIAGEN公司),具体操作参照说明书。

[0082] 三)DNA转化流程:

[0083] 所用的试剂成分均来自EpiTect Fast DNA Bisulfite Kit(购自QIAGEN公司)。

[0084] 1) 将提取的DNA配置亚硫酸盐转化体系配置:

[0085]

成分	反应体积(μ L)
提取的待测DNA	10 μ L
超纯水(RNase-free Water)	10 μ L
亚硫酸氢钠溶液(Bisulfite solution)	85 μ L
DNAprotect Buffer	35 μ L
总计	140 μ L

[0086] 2) 亚硫酸盐转化的运行体系:

[0087]

反应温度	反应时间
95℃	5min
60℃	20min
95℃	5min
60℃	20min
20℃	保存最多20hours

[0088] 3) 亚硫酸盐转化后DNA纯化:

[0089] 对亚硫酸盐转化后DNA纯化所用的试剂成分均选自德国QIAGEN公司的EpiTect Fast DNA亚硫酸氢钠试剂盒,具体操作参照说明书。

[0090] 四) 多重PCR,目标区域富集流程

[0091] 1、使用的PCR仪器为美国ABI 7500型实时荧光定量PCR系统(购自Life Tech (applied biosystems公司),反应体系为20 μ l;

[0092] 2、PCR反应体系配制及条件,如下表所示:

[0093] PCR反应体系:

[0094]

成分	反应体积
2 \times AmpliSeq HiFi Mix	10 μ l
AmpliSeq Primer Pool	2 μ l
目标基因组DNA (20ng/ μ l)	总含量40ng-70ng
无核酸酶水	补足余量
总量	20 μ l

[0095] 3、上机检测,按照各荧光PCR仪说明书进行操作,设置PCR程序,PCR程序如下:

[0096] PCR反应程序:

[0097]

阶段	步骤	温度	时间
持续	酶活化	99℃	2min
循环数 (25cycles)	变性	99℃	15sec
	退火/延伸	60℃	4min
持续	--	10℃	∞

[0098] 五) Illumina配套试剂进行高通量测序和生物信息分析

[0099] 在得到扩增子测序文库后,对文库进行质检。检测所构建文库的片段大小是否合适,并根据浓度决定上机使用的量。本发明构建的文库可以利用Illumina Miseq、Miniseq配套试剂进行高通量测序。

[0100] 试剂盒检测结果满足质控要求,根据检测结果对样本进行判断。 β -actin测序质量和转化率达标, β -actin测序结果不合格Q30>80%或测序深度小于500 \times ,则样本不足或有

抑制物存在PCR反应无效或失败,需重复实验或采样。

[0101] 六) 诊断算法

[0102] FAM19A4、miR124-2、C13ORF18、EPB41L3、JAM3、SST、ZIC1、GHSR、GFRA1、ANKRD18CP 甲基化模式将用采用下列方式进行计算。

[0103] 单个CpG岛甲基化阳性状态判断:甲基化CpG岛阳性率 $\geq 2\%$;

[0104] 单基因的甲基化状态评判:目标序列的总甲基化CpG岛/总CpG岛 $\geq 30\%$ 。

[0105] 七) 宫颈癌脱落细胞的基因甲基化分布状况

[0106] 宫颈癌脱落细胞基因甲基化检测统计结果:

	样本	甲基化检测		P-value
		阴性	阳性	
	Normal (90)	88	2	
[0107]	CIN1(53)	45	8	
	CIN2(58)	42	16	
	CIN3(27)	2	25	
	鳞癌+腺癌 (30)	0	30	P<0.001

[0108] 本发明提供的基于高通量测序的宫颈癌多基因甲基化检测的引物组、试剂盒及方法的有益效果在于:

[0109] 一、通过设计特异性高的引物组,并配置成使用方便且检测结果可靠的试剂盒,再设计出科学合理的多重PCR反应体系,使得本发明具有快速、高通量、灵敏及特异性好的特点,实现对宫颈癌FAM19A4、miR124-2、C13ORF18、EPB41L3、JAM3、SST、ZIC1、GHSR、GFRA1及ANKRD18CP十个目标扩增子靶基因的甲基化程度快速及准确测量,以便对宫颈癌间接进行及时、有效的诊断和治疗,降低医疗成本,节约社会资源。

[0110] 二、通过设计用于特异性扩增目标扩增子的引物组,通过将接头序列、测序引物序列、Barcode、目标片段扩增引物序列这四种序列整合在上、下游扩增的一对引物上,使得利用该引物能够通过一步扩增完成文库构建,不仅提高了甲基化测序文库质量和建库的效率,而且构建所得的扩增子测序文库因具有常规测序平台上所使用的接头序列及测序引物序列,使得测序文库具有常规测序所用的测序试剂和引物进行测序。本发明相同的样本具有相同的barcode,不同的样本具有不同的barcode。一次可以对多个样本多个基因进行测序分析,相比一代测序提高了测序的速度。

[0111] 三、基于高通量二代测序的甲基化DNA检测技术,对病人宫颈细胞10个靶基因甲基化进行检测,所述的目标扩增子来源于多方面基于人群的验证和二代高通量筛选后的结果,所选的目标扩增子均与宫颈癌的进程密切相关。宫颈癌的发生是个多诱因多步骤的过程,本发明的靶扩增子包括FAM19A4、miR124-2、C13ORF18、EPB41L3、JAM3、SST、ZIC1、GHSR、GFRA1、ANKRD18CP,多基因甲基化联合检测综合评价,灵敏度、特异性均要优于单基因检测。此外,甲基化发生并不局限于一两个碱基,高密度甲基化致使蛋白表达明显减低。本发明基于高通量二代测序的甲基化DNA检测技术,每个靶基因的甲基化区域检测均针对核心调控区域多个CpG,相对单个碱基的检测减少了碱基变化的随机性。

[0112] 以上所述仅为本发明的实施例,并非因此限制本发明的专利范围,凡是利用本发明说明书内容所作的等效流程变换,或直接或间接运用在其它相关的技术领域,均同理包括在本发明的专利保护范围内。

序列表

<110> 韩林志

<120> 基于高通量测序的宫颈癌多基因甲基化检测的引物组、试剂盒及方法

<130> 2018

<160> 30

<170> SIP0SequenceListing 1.0

<210> 1

<211> 145

<212> DNA

<213> 人工序列(Artificial Sequence)

<400> 1

ygggyggtty ggtaattag ygygtttygy ggygggatty gttttttttg gttagygygt 60
tagttgggtt yggtttygta ttgttagttg ygygtygttt tggayggggy gtatygattg 120
ygygygyggt tgygggtaaa tatyg 145

<210> 2

<211> 140

<212> DNA

<213> 人工序列(Artificial Sequence)

<400> 2

aygtygttat ttgaaaatta gatttatagg tttatgtatg tttttagggyg tgtgttgtaa 60
atggatatgga gatatatgta tatgtatayg taggtatayg tatygtttat atttttaygg 120
aatagattaa ttaatagygg 140

<210> 3

<211> 198

<212> DNA

<213> 人工序列(Artificial Sequence)

<400> 3

ttgttaggtt tttgattttt gaaagygytyg ttgygtttyg ygytygygggt aggtagggyg 60
ggatttttag gaggatygggt agaggygygt ataggtgygt ggtgttgggt tygggygytyg 120
yggatyggt gtaggagygy gtatttttag agtttttttt atttgggyga ygtttyggtg 180
tttgygyggg gaayggyg 198

<210> 4

<211> 163

<212> DNA

<213> 人工序列(Artificial Sequence)

<400> 4

yggaggytyg ggttttaggtt ttgygygytyg gtttagttat tattgygytyg ygygggygg 60
agygggyggg gggygygygy ygtaggttyg gtttyggtggg ggtttyggyg agygggaggg 120

yggttgggga ttttggttyg gtygggygyg gggtttyggga tty 163

<210> 5

<211> 252

<212> DNA

<213> 人工序列(Artificial Sequence)

<400> 5

yggtttagag tatygttgta ttygtaagta gtagatTTTT agTTTTTTTT gttattatgg 60
tgtyggTTYG gttgggTTYG gyggygtta tggtaattgg ggygggtygt agggTTTTgg 120
taggttgggy gtatgygygy ggggattata agtygygtyg ygttgtygtt ggtTTTTtag 180
taatttTYga tatggygttg aggyggttat ygygatttyg gttttgygtt yggttgtttg 240
atTTTTTTTT gt 252

<210> 6

<211> 148

<212> DNA

<213> 人工序列(Artificial Sequence)

<400> 6

yggtttTYgy ggygtygaga tgttgTTTT tygtTTTTtag tgygygttg ttgygttgTT 60
tatygttttg gttttgggtt gtgttatygg ygttttTYg gatTTtagat ttggttagtt 120
ttttagaag tttttggttg ttgtygyg 148

<210> 7

<211> 172

<212> DNA

<213> 人工序列(Artificial Sequence)

<400> 7

gttttttTY gggtagtttt gaygygyggt ttttTYgta gagattgagy ggygagaaag 60
tgygagtygg gtyggtagaa tttgtttggy ggyggttgga gtttgygtta ttygyggtty 120
gtagtygtty gttattttg ygtttggtty ggtagygtt ygggygygty gy 172

<210> 8

<211> 150

<212> DNA

<213> 人工序列(Artificial Sequence)

<400> 8

ygttygtttt tgtagtaty ggttttgga tttYggygat ttttttgta aattttttta 60
attygttagt gagagttgta tttaygttay ggtgtttatt attatagygt tgagygyga 120
gygttattty gttatttggt ttttatttyg 150

<210> 9

<211> 154

<212> DNA

<213> 人工序列(Artificial Sequence)

<400> 9

ygyggttyggt ttagaatat tgygtgtttt gtgygygagt aygggttygg tygygggttt 60
tatygttaagy ggggtgtgt agttgyggtt gtttggyggg ygagtttagg agggagttgg 120
agttgyggag ygggggaat aggagtaggt ygag 154

<210> 10

<211> 286

<212> DNA

<213> 人工序列(Artificial Sequence)

<400> 10

ygggagaygt ttgggttagg ygtttttgag ttttatggat taagagtayg ygggtygggg 60
gtattataty ggggattggg aattgyggaa gatttatagg gtggttatta aggggyaygt 120
ygyggaggtg gagygttggt tgattygtag gtttygggat ttggatgtty gygatagaaa 180
ggataggtag ygggggttta gtttygygtg ggaggggggtt ttttaggtttt gtttgtttygt 240
agygtttgag gygggggttt tggaggttygt ygggttttgg agtygy 286

<210> 11

<211> 95

<212> DNA

<213> 人工序列(Artificial Sequence)

<400> 11

aatgatacgg cgaccaccga gatctacacc tctctattcg tcggcagcgt cagatgtgta 60
taagagacag ggggttaggt agggatagga gtagt 95

<210> 12

<211> 109

<212> DNA

<213> 人工序列(Artificial Sequence)

<400> 12

caagcagaag acggcatacg agataggetc cggtctcgtg ggctcggaga tgtgtataag 60
agacaggact acgggtatct aatccaaacg acaactaaaa caaaactcc 109

<210> 13

<211> 96

<212> DNA

<213> 人工序列(Artificial Sequence)

<400> 13

aatgatacgg cgaccaccga gatctacacc tctctattcg tcggcagcgt cagatgtgta 60
taagagacag attagagggg taattaattt ggattt 96

<210> 14

<211> 109

<212> DNA

<213> 人工序列(Artificial Sequence)

<400> 14

caagcagaag acggcatacg agataggctc cggctctcgtg ggctcggaga tgtgtataag 60
agacaggact acgggtatct aatcctcatc tctaacacat ctaccaaaa 109

<210> 15

<211> 92

<212> DNA

<213> 人工序列(Artificial Sequence)

<400> 15

aatgatacgg cgaccaccga gatctacacc tctctattcg tcggcagcgt cagatgtgta 60
taagagacag atttgggggtt gggaagttat tt 92

<210> 16

<211> 108

<212> DNA

<213> 人工序列(Artificial Sequence)

<400> 16

caagcagaag acggcatacg agataggctc cggctctcgtg ggctcggaga tgtgtataag 60
agacaggact acgggtatct aatccaaaaa ctctcaaaa aaacaaa 108

<210> 17

<211> 92

<212> DNA

<213> 人工序列(Artificial Sequence)

<400> 17

aatgatacgg cgaccaccga gatctacacc tctctattcg tcggcagcgt cagatgtgta 60
taagagacag gattttaggg aggttttttag gt 92

<210> 18

<211> 105

<212> DNA

<213> 人工序列(Artificial Sequence)

<400> 18

caagcagaag acggcatacg agataggctc cggctctcgtg ggctcggaga tgtgtataag 60
agacaggact acgggtatct aatccatctc caataaaact aacc 105

<210> 19

<211> 96

<212> DNA

<213> 人工序列(Artificial Sequence)

<400> 19

aatgatacgg cgaccaccga gatctacacc tctctattcg tcggcagcgt cagatgtgta 60
taagagacag agttttattga aagagaattt atgtgt 96

<210> 20

<211> 107

<212> DNA

<213> 人工序列(Artificial Sequence)

<400> 20

caagcagaag acggcatacg agataggctc cggctctcgtg ggctcggaga tgtgtataag 60
agacaggact acgggtatct aatccaactc acccctaaaa aacaaca 107

<210> 21

<211> 95

<212> DNA

<213> 人工序列(Artificial Sequence)

<400> 21

aatgatacgg cgaccaccga gatctacacc tctctattcg tcggcagcgt cagatgtgta 60
taagagacag gaggttgatt agttatTTTT tagtt 95

<210> 22

<211> 109

<212> DNA

<213> 人工序列(Artificial Sequence)

<400> 22

caagcagaag acggcatacg agataggctc cggctctcgtg ggctcggaga tgtgtataag 60
agacaggact acgggtatct aatccaaaaa atctccttac ctacttccc 109

<210> 23

<211> 94

<212> DNA

<213> 人工序列(Artificial Sequence)

<400> 23

aatgatacgg cgaccaccga gatctacacc tctctattcg tcggcagcgt cagatgtgta 60
taagagacag ttttaggggg ttgagatgtt ttat 94

<210> 24

<211> 109

<212> DNA

<213> 人工序列(Artificial Sequence)

<400> 24

caagcagaag acggcatacg agataggctc cggctctcgtg ggctcggaga tgtgtataag 60
agacaggact acgggtatct aatccaaatc ctaacctaca aacaataac 109

<210> 25

<211> 94

<212> DNA

<213> 人工序列(Artificial Sequence)

<400> 25

aatgatacgg cgaccaccga gatctacacc tctctattcg tcggcagcgt cagatgtgta 60
taagagacag ttttttttgt atgtttttgg attt 94

<210> 26

<211> 107

<212> DNA

<213> 人工序列(Artificial Sequence)

<400> 26

caagcagaag acggcatacg agataggctc cggtctcgtg ggctcggaga tgtgtataag 60
agacaggact acgggtatct aatcccttaa taaccaccac ctttaacc 107

<210> 27

<211> 96

<212> DNA

<213> 人工序列(Artificial Sequence)

<400> 27

aatgatacgg cgaccaccga gatctacacc tctctattcg tcggcagcgt cagatgtgta 60
taagagacag agagagtgtg tgtgttttgt agtatt 96

<210> 28

<211> 109

<212> DNA

<213> 人工序列(Artificial Sequence)

<400> 28

caagcagaag acggcatacg agataggctc cggtctcgtg ggctcggaga tgtgtataag 60
agacaggact acgggtatct aatcccaaatt ttcttctaac caaaaaacc 109

<210> 29

<211> 95

<212> DNA

<213> 人工序列(Artificial Sequence)

<400> 29

aatgatacgg cgaccaccga gatctacacc tctctattcg tcggcagcgt cagatgtgta 60
taagagacag ttattatgag gaaatttttt agttt 95

<210> 30

<211> 110

<212> DNA

<213> 人工序列(Artificial Sequence)

<400> 30

caagcagaag acggcatacg agataggctc cggtctcgtg ggctcggaga tgtgtataag 60
agacaggact acgggtatct aatccaacaa ctaaaactcc atttaacttc 110

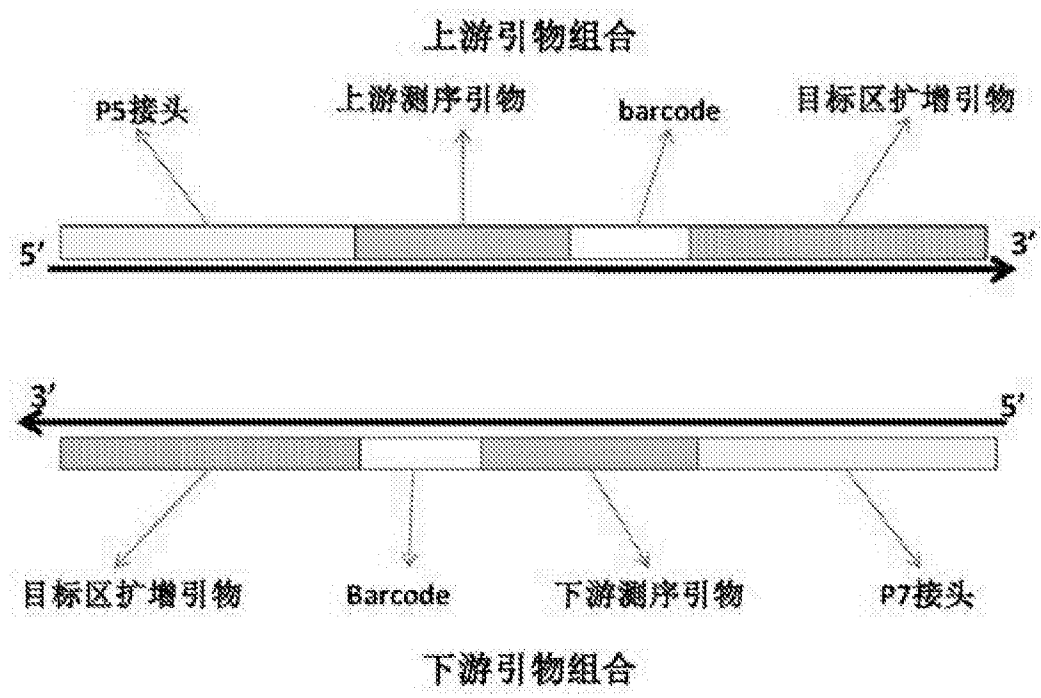


图1

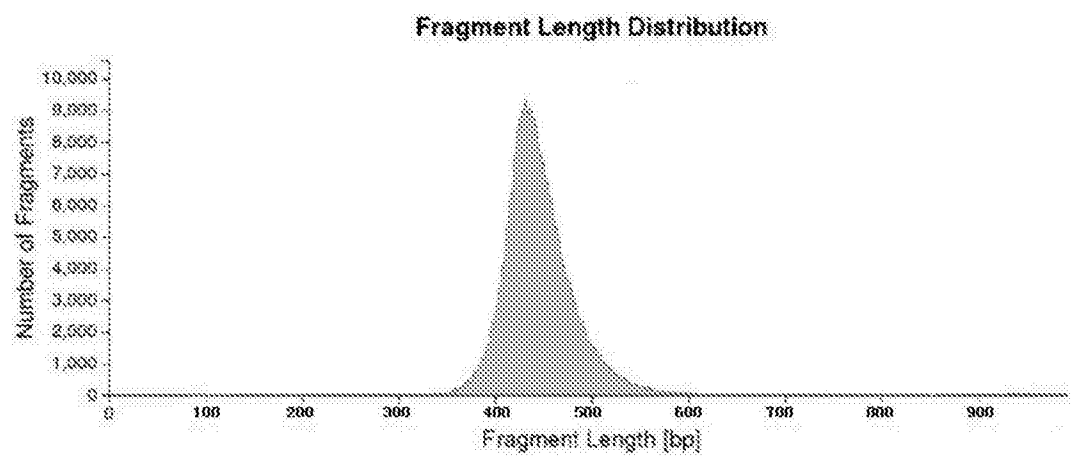


图2