

## (12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织  
国际局

(43) 国际公布日  
2020 年 7 月 2 日 (02.07.2020)



(10) 国际公布号  
**WO 2020/135347 A1**

- (51) 国际专利分类号:  
*C12Q 1/68* (2018.01) *G01N 33/00* (2006.01)
- (21) 国际申请号: PCT/CN2019/127537
- (22) 国际申请日: 2019 年 12 月 23 日 (23.12.2019)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:  
201811642291.9 2018 年 12 月 29 日 (29.12.2018) CN
- (71) 申请人: 深圳华大生命科学研究院 (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: 王冀(WANG, Ji); 中国广东省深圳市盐田区北山工业区综合楼, Guangdong 518083 (CN)。王欧(WANG, Ou); 中国广东省深圳市盐田区北山工业区综合楼, Guangdong 518083 (CN)。章文蔚(ZHANG, Wenwei); 中国广东省深圳市盐田区北山工业区综合楼, Guangdong 518083 (CN)。

陈奥(CHEN, Ao); 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); 中国广东省深圳市福田区金田路与福华路交汇处现代商务大厦 1810, 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, 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,

(54) Title: METHOD FOR DETECTING DNA METHYLATION, TEST KIT, DEVICE AND APPLICATION

(54) 发明名称: 一种 DNA 甲基化检测的方法、试剂盒、装置和应用

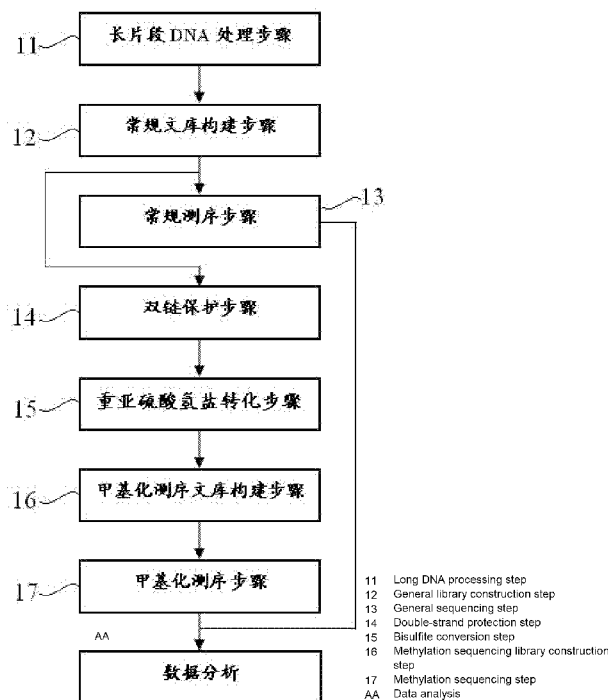


图 1

(57) Abstract: Provided by the present application is a method for detecting DNA methylation, comprising: a long DNA processing step of breaking long DNA; a general library construction step of using an stLFR technology to construct a library; a double-strand protection step of taking the recovered stLFR library, denaturing the DNA, and converting a molecular tag or linker sequence into double-strand by annealing or extension; a bisulfite conversion step of adding a double-stranded protective agent and bisulfite to perform deamination; a methylation sequencing library construction step of performing conventional PCR or rolling circle amplification on a deaminated product to obtain a methylation sequencing library; and a methylation sequencing step of sequencing the methylation sequencing library, and analyzing methylation information of genes according to a sequencing result. Also provided by the present application are a corresponding test kit, a device and an application. The method according to the present application uses the stLFR technology to obtain long-fragment methylation information. By means of a double-strand protection strategy, the molecular tag sequence may be added in advance to construct a library, which may effectively track a molecular source.

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))。

---

**(57) 摘要：**本申请提供了一种DNA甲基化检测的方法，包括长片段DNA处理步骤，将长片段DNA打断；常规文库构建步骤，用stLFR技术建库；双链保护步骤，取回收的stLFR文库，DNA变性，用退火或延伸方法将分子标签或接头序列转化为双链；重亚硫酸盐转化步骤，加入双链保护剂和重亚硫酸盐，进行脱氨基；甲基化测序文库构建步骤，对脱氨基产物进行常规PCR或环化后滚环扩增，得到甲基化测序文库；甲基化测序步骤，对甲基化测序文库测序，根据测序结果分析基因的甲基化信息；本申请还提供了相应试剂盒、装置和应用。本申请方法利用stLFR技术，获得长片段甲基化信息；通过双链保护策略，可事先加入分子标签序列建库，可有效追踪分子来源。

# 一种 DNA 甲基化检测的方法、试剂盒、装置和应用

## 技术领域

本申请涉及 DNA 甲基化检测领域，特别是涉及一种 DNA 甲基化检测的方法、试剂盒、装置和应用。

## 背景技术

分子生物学的快速发展，使基因测序技术成为现代生物科学研究中重要的手段之一，广泛应用于肿瘤的预防、筛查、诊断、治疗和预后中。基因测序技术能够真实地反映基因组上全部 DNA 遗传信息，进而比较全面地揭示肿瘤的发生机制及其发展过程，因而在肿瘤的科学研究中具有十分重要的地位。第一代测序技术是 1977 年 Sanger 等发明的双脱氧核苷酸末端终止法和 Gilbert 等发明的化学降解法；第二代测序技术包括 Roche 公司的 454 技术、Illumina 公司的 Solexa 技术、ABI 公司的 SOLiD 技术和华大基因的纳米球测序技术等；Helicos 公司和 Pacbio 的单分子测序技术被称为第三代测序技术。由于第三代测序技术对文库的要求较高，测序成本也较高，因此目前应用最广泛的是第二代测序技术。例如全基因组测序应用于无创产前基因检测、目标区域捕获测序应用于肿瘤靶向用药基因检测、单细胞基因组和转录组测序应用于肿瘤组织的异质性和发生发展机制的研究、长片段测序应用于无创地贫检测的研究等，多种临床检测及基础研究都是在第二代测序技术的平台上开展进行。第二代测序技术，也称高通量测序技术，为临床实验室的分子检测带来革命性的变化；因此，如何构建更高质量的测序文库，以得到更好的测序数据，成为高效检测和广泛临床应用的重要问题。

近年来，表观遗传学研究再次成为基因诊断和精准医疗研究的热点。表观遗传是指在不改变基因的核苷酸序列的情况下，发生的可遗传的基因表达的变化，例如甲基化修饰。针对甲基化修饰的测序检测称为甲基化测序。甲基化测序所提供的表观遗传学信息，在基因印记、胚胎发育、癌症预防等研究领域都有重要作用。在众多表观遗传因子中，碱基的甲基化，特别是胞嘧啶的甲基化信息，是目前实验手段最容易获得，且研究最为深入的表观遗传学分子级标志物。基于重亚硫酸盐的原理，对非甲基化的胞嘧啶进行化学转化，而甲基化的胞嘧啶不会被转化，从而得到基因组的甲基化信息。利用甲基化分析流程，与已知的参考基因组进行比对，即可确定甲基化位点。

目前，经典的全基因组甲基化测序文库的构建都是通过物理方式、酶切或者化学降解的方式将双链 DNA 长片段随机打断成几百 bp 的小片段，再将双链 DNA 变性成单链后，进行重亚硫酸氢钠处理，将没有甲基化的胞嘧啶通过脱氨基作用转化为尿嘧啶，通过引物退火延伸或加接头后延伸得到双链 DNA 分子，然后进行末端修复、

加“A”、加“接头”、PCR扩增等步骤，最终得到用于测序的文库。在获得测序数据后，因为大部分未甲基化的胞嘧啶转变为了尿嘧啶，因此需要将所得文库与已知的原始基因组进行比对，以确定甲基化位点或获得靶标区域的甲基化水平。

受限于脱氨基反应的反应原理，一般经过重亚硫酸盐处理后DNA降解严重，降解率甚至可以达到90%以上，因此会得到众多小片段DNA分子，为后期的基因组比对带来困难；并且，经过这个过程会损失一部分信息，容易丢失全长信息，虽然可以对基因的特定区域进行精确的甲基化测序，但是为全基因组的甲基化信息组装增加了难度。其次，因为重亚硫酸盐处理会改变非甲基化的DNA序列信息，因此无法事先加入分子标签（barcode）序列进行建库，以至于难以追踪分子来源，这对于多倍体生物的同源染色体甲基化信息提取带来困难，也会降低数据比对的可信度。虽然，部分测序试剂盒通过采用人工合成的已经甲基化的测序接头，将加接头步骤安排在片段化之前，可以一定程度上解决上述分子来源追踪的问题；但是，甲基化的测序接头合成成本高昂，不利于大规模的高通量测序操作。此外，现有的甲基化测序，其主要时间成本体现在测序后的数据分析上；对于没有参考基因组的物种，直接进行甲基化分析会非常困难。

## 发明内容

本申请的目的是提供一种新的DNA甲基化检测的方法、试剂盒、装置及其应用。

本申请具体采用了以下技术方案：

本申请的第一方面公开了一种DNA甲基化检测的方法，包括以下步骤，

长片段DNA处理步骤，包括将长片段DNA打断成小片段DNA；

常规文库构建步骤，包括采用stLFR技术进行文库构建，获得常规文库；

双链保护步骤，包括取用常规文库构建步骤制备的常规文库，磁珠回收，将DNA变性为单链，然后利用退火或延伸的方法将分子标签序列或测序接头序列转化为双链；其中，对常规文库的磁珠回收，本申请的一种实现方式中，在进行stLFR技术建库时，是采用带有分子标签序列的磁珠对DNA片段进行捕获，然后直接将磁珠和捕获DNA片段一起进行stLFR建库，获得常规文库；由于建库过程中涉及到PCR扩增等步骤，因此，常规文库中有两部分核酸，一部分是始终与磁珠结合的DNA，另一部分则是通过PCR扩增产生的游离DNA；用于甲基化测序的则是其中与磁珠结合的DNA，因此，需要取常规文库进行磁珠回收；而游离DNA则根据需求进行后续的常规测序，获得正常基因测序结果；

重亚硫酸盐转化步骤，包括向双链保护步骤的产物中加入双链保护剂，并加入重亚硫酸盐，进行脱氨基反应；

甲基化测序文库构建步骤，包括对脱氨基反应的产物进行常规PCR扩增或环化后滚环扩增，得到甲基化测序文库；

甲基化测序步骤，包括对甲基化测序文库进行测序，并根据测序结果分析获得基因的甲基化信息。其中，根据测序结果分析甲基化信息，具体的是指，将甲基化文库的测序结果与正常基因测序结果进行比对，获得基因的甲基化信息；对于有参考基因组的物种，用于比对的正常基因测序结果就是参考基因组；对于没有参考基因组的物种，用于比对的正常基因测序结果则是前述游离 DNA 进行常规测序获得的测序结果。

需要说明的是，本申请的 DNA 甲基化检测方法，利用 stLFR 技术进行常规文库构建，通过该技术采用的 co-barcoding，获得长片段的甲基化信息，解决了现有的甲基化测序读长偏短的问题；与此同时，采用双链保护的方法对分子标签序列或测序接头序列进行保护，而不影响待测片段的重亚硫酸盐处理，简单而有效的解决了现有的甲基化测序无法事先加入分子标签序列进行建库的问题，不仅可以有效追踪分子来源，也可以方便的对多倍体生物的同源染色体甲基化信息进行提取。

优选的，本申请的 DNA 甲基化检测方法还包括常规测序步骤，常规测序步骤包括，取用常规文库构建步骤采用 stLFR 技术构建的常规文库，将 DNA 变性为单链，然后再进行常规 PCR 扩增或环化后滚环扩增，获得测序文库，对测序文库进行全基因组测序，获得正常基因测序结果；因此，甲基化测序步骤还包括，将甲基化测序文库的测序结果与常规文库获得的正常基因测序结果比对，从而获得基因的甲基化信息。

需要说明的是，对于有参考基因组的物种，在甲基化测序步骤后，可以直接将甲基化测序文库的测序结果与参考基因进行比对分析，从而获得基因的甲基化信息。而对于没有参考基因组的物种，本申请的优选方案基于共建库的原理，将常规文库构建步骤获得的常规文库分为两部分，一部分进行全基因组测序，获得正常基因测序结果；另一部分经过处理后进行甲基化测序；将甲基化测序的测序结果与全基因组测序的正常基因测序结果进行比对分析，获得基因的甲基化信息。本申请的共建库的方案，可以在无参考文库或无参考基因组的情况下，在分子级别对处理前后的文库进行一对一对比，从而得到更加精确的甲基化信息。

优选的，长片段 DNA 处理步骤还包括，采用带有分子标签序列的载体对小片段 DNA 进行捕获，然后再将捕获的 DNA 用于常规文库构建步骤进行 stLFR 技术文库构建。

优选的，载体为磁珠。

可以理解，磁珠载体是本领域比较常规使用的载体，不排除还可以使用其它载体。

优选的，长片段 DNA 处理步骤中，将长片段 DNA 打断成小片段 DNA 具体采用 Tn5 转座酶。

可以理解，Tn5 转座酶打断获得小片段 DNA 是本领域常规使用的技术，不排除还可以采用其它物理或化学手段打断 DNA。

优选的，双链保护步骤中，延伸的方法具体包括，利用分子标签序列或测序接头序列自身的 3' 端序列作为引物或加入外源短片段作为引物，在 DNA 聚合酶作用下延伸，将分子标签序列或测序接头序列变为双链结构。

优选的，重亚硫酸盐转化步骤中，双链保护剂为盐离子、聚乙二醇、酶和有机溶剂中的至少一种，重亚硫酸盐为重亚硫酸氢钠。其中，盐离子例如 NaCl，聚乙二醇例如聚乙二醇 8000。

需要说明的是，在重亚硫酸盐进行脱氨基处理的过程，需要保持一定温度，防止单链 DNA 形成二级结构从而影响反应效率；但是，本申请采用了将分子标签序列或测序接头序列转化为双链的双链保护策略；因此，在重亚硫酸盐的反应温度下，需要保障转化的双链不被破坏，因此，本申请优选的方案中加入了双链保护剂，其作用就是使得已经双链化的分子标签序列或测序接头序列能够在高温下始终保持双链状态，以避免其参与重亚硫酸盐反应，从而起到有效保护的作用。可以理解，凡是能够起到以上作用的双链保护剂都可以用于本申请。

本申请的另一面公开了一种 DNA 甲基化检测的试剂盒，该试剂盒中包括第一组试剂、第二组试剂、第三组试剂、第四组试剂和第五组试剂；第一组试剂包括用于将长片段 DNA 打断成小片段 DNA 的试剂；第二组试剂包括用于 stLFR 技术建库的试剂，该试剂用于制备常规文库；第三组试剂包括用于通过退火或延伸的方法将分子标签序列或测序接头序列转化为双链的试剂，该试剂用于形成双链结构；第四组试剂包括双链保护剂和重亚硫酸盐，用于进行双链保护和脱氨基反应；第五组试剂包括用于对脱氨基反应的产物进行常规 PCR 扩增或环化后滚环扩增的试剂，该试剂用于制备甲基化测序文库。

需要说明的是，本申请的试剂盒实际上就是根据本申请的 DNA 甲基化检测方法，将其中核心的试剂有机的组合在一起，一方面，可以方便本申请的 DNA 甲基化检测方法的进行；另一方面，可以提高各试剂的使用效率，减少浪费，从而降低本申请的 DNA 甲基化检测的成本。

可以理解，本申请试剂盒中的各试剂都可以单独通过市售购买，但是，本申请的 DNA 甲基化检测方法涉及的试剂较多，单独购买一方面是耗费时间、繁琐、使用不便；另一方面，由于各试剂都不是按照本申请检测方法用量配比进行的剂量设计，所以，必然会存在大量的试剂浪费，增加检测成本。因此，本申请以本申请 DNA 甲基化检测方法为基础，将其使用的主要核心试剂组合为本申请的试剂盒。

优选的，本申请的试剂盒还包括第六组试剂，该第六组试剂包括用于对常规文库进行常规 PCR 扩增或环化后滚环扩增的试剂，该试剂用于制备常规测序文库，以

便于进行全基因测序，获得正常基因测序结果。可以理解，第六组试剂中，还可以根据需求增加常规测序所需的试剂。

需要说明的是，本申请的试剂盒是根据本申请的 DNA 甲基化检测方法而研发的，因此，第一组试剂、第二组试剂、第三组试剂、第四组试剂和第五组试剂，实际上就是依序用于本申请 DNA 甲基化检测方法中的长片段 DNA 处理步骤、常规文库构建步骤、双链保护步骤、重亚硫酸盐转化步骤和甲基化测序文库构建步骤；第六组试剂用于常规测序步骤的常规文库构建。可以理解，由于常规测序步骤并非必须的步骤，例如已经有参考基因组的情况下，则不需要进行常规测序；所以本申请的试剂盒中也可以选择性的增加或不增加第六组试剂，又或者将第六组试剂组合到本申请的试剂盒中，可以根据需求选择性的使用。

还需要说明的是，测序试剂盒是单独的试剂盒产品，因此没有包含在本申请的试剂盒中，例如常规测序有独立的常规测序试剂盒，甲基化测序步骤有甲基化测序试剂盒。

优选的，第一组试剂还包括带有分子标签序列的载体，用于对 DNA 片段进行捕获。

优选的，载体为磁珠，第一组试剂中还包括磁珠纯化的试剂。其中，磁珠纯化的试剂，例如磁珠缓冲液、清洗液、洗脱液等。

优选的，第一组试剂中，用于将长片段 DNA 打断成小片段 DNA 的试剂为 Tn5 转座酶及其缓冲液。

优选的，第三组试剂中，用于通过退火或延伸的方法将分子标签序列或测序接头序列转化为双链的试剂，具体包括外源短核苷酸片段、DNA 聚合酶及其缓冲液，以及用于 DNA 延伸的 dNTPs。

其中，外源短核苷酸片段是作为引物使用的，可以根据需求选择使用，也可以不添加到第三组试剂盒中；例如，本申请的一种实现方式中，可以直接利用分子标签序列或测序接头序列自身的 3' 端序列作为引物进行延伸，因此，外源短核苷酸片段可以只是作为备选方案放到第三组试剂中。

优选的，第四组试剂中，双链保护剂为盐离子、聚乙二醇、酶和有机溶剂中的至少一种，重亚硫酸盐为重亚硫酸氢钠。

可以理解，本申请的 DNA 甲基化检测方法，其全部或部分功能可以通过硬件的方式实现，也可以通过计算机程序的方式实现。当通过计算机程序的方式实现时，该程序可以存储于一计算机可读存储介质中，存储介质可以包括：只读存储器、随机存储器、磁盘、光盘、硬盘等，通过计算机执行该程序以实现本申请的方法。例如，将程序存储在设备的存储器中，当通过处理器执行存储器中程序，即可实现本申请的方法。当本申请的方法中全部或部分功能通过计算机程序的方式实现时，该程序也可以存储在服务器、另一计算机、磁盘、光盘、闪存盘或移动硬盘等存储介

质中，通过下载或复制保存到本地设备的存储器中，或对本地设备的系统进行版本更新，当通过处理器执行存储器中的程序时，即可实现本申请 DNA 甲基化检测方法的全部或部分功能。

因此，本申请的另一面公开了一种 DNA 甲基化检测的装置，包括长片段 DNA 处理模块、常规文库构建模块、常规测序模块、双链保护模块、重亚硫酸盐转化模块、甲基化测序文库构建模块和甲基化测序模块；

长片段 DNA 处理模块，包括用于将长片段 DNA 打断成小片段 DNA；该模块可以采用常规的物理、化学或酶的方式对长片段 DNA 进行打断；对于物理方式，例如可借鉴超声波破碎；对应化学或酶的方式，该模块一方面可以自动取样或配制反应液，另一方面可以为反应提供温度等反应条件；

常规文库构建模块，包括用于采用 stLFR 技术进行文库构建；stLFR 技术，即 single-tube Long Fragment Read，也就是长片段 DNA 信息获取技术，可参考专利文件 W02019217452A1；另外，stLFR 技术进行文库构建可以采用华大智造发布的 MGIEasy stLFR 文库制备试剂盒进行，具体参考试剂盒使用说明；

常规测序模块，包括用于取用常规文库构建模块中采用 stLFR 技术构建的常规文库，将 DNA 变性为单链，然后再进行常规 PCR 扩增或环化后滚环扩增，获得测序文库，对测序文库进行全基因组测序，获得正常基因测序结果；该模块仅仅在没有参考基因的情况下使用，当然，如果为了获得更精准的处理前后的结果，也可以采用该模块；该模块可以参考现有的自动建库和测序设备或装置；

双链保护模块，包括用于取用常规文库构建模块中采用 stLFR 技术构建的常规文库，磁珠回收，将 DNA 变性为单链，然后利用退火或延伸的方法将分子标签序列或测序接头序列转化为双链；该模块一方面可以参考自动取样或配制反应液，另一方面可以为退火或延伸提供温度等反应条件控制；

重亚硫酸盐转化模块，包括用于向经过双链保护模块处理的产物中加入双链保护剂，并加入重亚硫酸盐，进行脱氨基反应；该模块同样可以参考现有的自动取样和反应液配制装置或设备，并为脱氨基反应提供相应的反应条件；

甲基化测序文库构建模块，包括用于对脱氨基反应的产物进行常规 PCR 扩增或环化后滚环扩增，得到甲基化测序文库；该模块可以参考现有的测序模块；

甲基化测序模块，包括用于对甲基化测序文库进行测序，将甲基化测序结果与正常基因测序结果比对，获得基因的甲基化信息；该模块可以参考现有的甲基化测序和分析模块或装置。其中，正常基因测序结果可以是常规测序模块获得的测序结果，也可以是已经存在的参考基因组。

本申请的 DNA 甲基化检测装置，根据本申请 DNA 甲基化检测方法的需求，将检测方法中的各步骤通过自动化的模块实现 DNA 甲基化的自动检测，即本申请的 DNA 甲基化检测装置，为 DNA 甲基化检测提供了个性化的检测系统，能够方便有效的进

行 DNA 甲基化检测。并且，本申请的 DNA 甲基化检测装置同样具有本申请的 DNA 甲基化检测方法的各项优点，例如解决了现有的甲基化测序读长偏短的问题，解决了现有的甲基化测序无法事先加入分子标签序列进行建库、不能追踪分子来源的问题。

优选的，本申请的 DNA 甲基化检测装置中，其长片段 DNA 处理模块，还包括用于采用带有分子标签序列的载体对小片段 DNA 进行捕获，然后再将捕获的 DNA 用于常规文库构建模块采用 stLFR 技术进行文库构建；

优选的，载体为磁珠。

优选的，本申请的 DNA 甲基化检测装置，其长片段 DNA 处理模块中，将长片段 DNA 打断成小片段 DNA 具体采用 Tn5 转座酶。

优选的，本申请的 DNA 甲基化检测装置，其双链保护模块中，延伸的方法具体包括，利用分子标签序列或测序接头序列自身的 3' 端序列作为引物或加入外源短片段作为引物，在 DNA 聚合酶作用下延伸，将分子标签序列或测序接头序列变为双链结构。

优选的，本申请的 DNA 甲基化检测装置，其重亚硫酸盐转化模块中，双链保护剂为盐离子、聚乙二醇、酶和有机溶剂中的至少一种，重亚硫酸盐为重亚硫酸氢钠。

本申请的再一面还公开了一种 DNA 甲基化检测装置，该装置包括存储器和处理器；存储器用于存储程序；处理器用于通过执行存储器存储的程序以实现本申请的 DNA 甲基化检测方法。

本申请的再一面还公开了一种计算机可读存储介质，包括存储于其中的程序，该程序能够被处理器执行以实现本申请的 DNA 甲基化检测的方法。

本申请的再一面还公开了本申请的 DNA 甲基化检测方法、本申请的试剂盒或本申请的 DNA 甲基化检测装置在基因印记检测、胚胎发育检测或癌症预防检测中的应用。

可以理解，本申请的 DNA 甲基化检测方法、试剂盒和装置，能够获得长片段的甲基化信息，不仅可以有效追踪分子来源，也可以方便的对多倍体生物的同源染色体甲基化信息进行提取；因此，能够方便的应用于基因印记检测、胚胎发育检测、癌症预防检测等领域。

本申请的有益效果在于：

本申请的 DNA 甲基化检测方法，利用 stLFR 技术的优点，能够获得长片段的甲基化信息；采用双链保护的方法对分子标签序列或测序接头序列进行保护，可以事先加入分子标签序列进行建库，从而可以有效追踪分子来源，并方便的对多倍体生物的同源染色体甲基化信息进行提取。

附图说明

图 1 是本申请实施例中 DNA 甲基化检测方法的流程示意图；

图 2 是本申请实施例中 DNA 甲基化检测装置的构造示意图；

图 3 是本申请实施例中基于磁珠捕获的双链保护分子标签的示意图；

图 4 是本申请实施例中不同形式的双链保护的示意图；

图 5 是本申请实施例中不同形式的双链保护的示意图；

图 6 是本申请实施例中不同形式的双链保护的示意图；

图 7 是本申请实施例中构建的 stLFR 常规文库 (WGS) 和甲基化测序文库 (WGBS) 的琼脂糖凝胶电泳图。

## 具体实施方式

现有的甲基化测序读长偏短，并且，一般常规的甲基化测序不能事先加入分子标签序列进行建库，无法追踪分子来源，难以实现多倍体生物的同源染色体甲基化信息提取。虽然有研究提出可以采用人工合成的甲基化的测序接头，解决事先加入分子标签的问题，但是，人工合成甲基化测序接头成本高昂，难以用于大规模的高通量测序。

基于以上问题，本申请创造性的借鉴 stLFR 技术进行文库构建，解决读长偏短的问题；然后，创造性的提出双链保护策略，对事先添加的分子标签序列或测序接头序列进行保护，避免其被重亚硫酸盐处理破坏，以解决不能事先加入分子标签序列进行建库的问题；利用同一文库两次建库的策略，解决了无参考基因组条件下的甲基化文库比对问题。根据以上构思，本申请研发了一种 DNA 甲基化检测的方法，如图 1 所示，包括长片段 DNA 处理步骤 11、常规文库构建步骤 12、双链保护步骤 14、重亚硫酸盐转化步骤 15、甲基化测序文库构建步骤 16、甲基化测序步骤 17；在优选的方案中，基于共建库的原理，本申请的方法还包括常规测序步骤 13。

其中，长片段 DNA 处理步骤 11，包括将长片段 DNA 打断成小片段 DNA；包括采用物理、化学或酶的方式将长片段 DNA 打断，例如本申请的一种实现方式中，具体采用 Tn5 转座酶将 DNA 打断成小片段。

常规文库构建步骤 12，包括采用 stLFR 技术进行文库构建，获得常规文库。该步骤可以参考专利文献 W02019217452A1 或者采用华大智造发布的 MGIEasy stLFR 文库制备试剂盒进行常规建库。在本申请的一种实现方式中，在此之前还包括采用带有分子标签序列的载体对打断的 DNA 片段进行捕获，然后再进行 stLFR 技术建库。其中，带有分子标签序列的载体即带有多拷贝特定标签序列或分子标签的载体，是指在一个载体上，带有相同序列的多个拷贝的寡聚核苷酸序列。

本申请的一种实现方式中，具体采用磁珠作为载体，为了保障磁珠和 DNA 的稳固结合，本申请通过双链霉亲和素蛋白嵌合体连接，使用化学修饰的 DNA 与磁珠上的化学基团反应，这些化学修饰包括但不限于 I-linker、Amino-modified oligo、

Thiol-modified oligo、Acrydite-modified oligo 等。可以理解，磁珠在后续的反应中需要进行两轮 PCR 扩增，因此要求 DNA 与磁珠的偶联必须能够耐酸碱、耐高温且不易降解。

常规测序步骤 13, 包括取用常规文库构建步骤采用 stLFR 技术构建的常规文库，将 DNA 变性为单链，然后再进行常规 PCR 扩增或环化后滚环扩增，获得测序文库，对测序文库进行测序，获得正常基因测序结果。该步骤即正常的 DNA 建库和测序步骤，其目的是为了便于没有已知参考基因的物种，可以通过甲基化处理前后对比，准确的分析甲基化位点。可以理解，如果待测物种已经有比较准确的参考基因序列，则可以不用该步骤。

双链保护步骤 14, 包括取用常规文库构建模块中采用 stLFR 技术构建的常规文库，磁珠回收，将 DNA 变性为单链，然后利用退火或延伸的方法将分子标签序列或测序接头序列转化为双链。

双链保护可以采用退火或延伸的方式实现，其中，延伸的方式实现，例如，利用分子标签序列或测序接头序列自身的 3' 端序列作为引物，如图 3 的 A 图所示，或者，加入外源短片段作为引物，如图 3 的 B 图所示，在 DNA 聚合酶作用下延伸，将分子标签序列或测序接头序列变为双链结构。

另外，双链保护根据不同的载体或者 DNA 的不同形态，还可以有多种保护形式，如图 4 至图 6 所示。图 4 左边一列为与磁珠结合的 DNA 的双链保护，右边一列为游离 DNA 的双链保护；“●”表示磁珠，深色双实线表示互补 DNA 双链，浅色双实线表示非互补 DNA 双链；a 为单端保护、b 为双端保护、c 为桥式保护、d 为 3' 和/或 5' 端颈环结构保护、e 为 3' 和/或 5' 端颈环桥式结构保护、f 为分子内颈环结构保护。图 5 为环状 DNA 的双链保护机制，同样的，深色双实线表示互补 DNA 双链；a 为引入单链引物的双链保护、b 为通过 Y 接头引物进行的双链保护、c 为单链自形成双链保护。图 6 为以固体表面为载体的 DNA 的双链保护机制，深色双实线表示互补 DNA 双链，浅色双实线表示非互补 DNA 双链；其中，a 和 b 是不同位置的单端保护、c 为双端保护、d 为颈环式单端保护、e 为桥式结构保护、f 为 3' 或 5' 端颈环桥式结构保护、g 为分子间桥式结构保护。具体采用怎样的保护机制可以根据 DNA 序列的设计而定，或者根据不同的载体设计不同的序列，以采取不同的双链保护机制。

重亚硫酸盐转化步骤 15, 包括向双链保护步骤的产物中加入双链保护剂，并加入重亚硫酸盐，进行脱氨基反应。

在 DNA 进行重亚硫酸盐，例如重亚硫酸氢钠，处理过程中，需要保持一定温度，防止单链 DNA 形成二级结构从而影响反应效率，而在本申请中，文库同时存在双链和单链结构，因此需要加入一种 DNA 的双链保护试剂，使得已经双链化的序列能够在高温下始终保持双链状态，以起到有效的保护作用。本申请可以采用的双链保护

试剂包括盐离子、聚乙二醇、酶和有机溶剂等能够有效的防止高温下双链 DNA 变性，同时能够保持单链的转化效率。

甲基化测序文库构建步骤 16，包括对脱氨基反应的产物进行常规 PCR 扩增或环化后滚环扩增，得到甲基化测序文库。甲基化测序步骤 17，包括对甲基化测序文库进行测序，然后进行数据分析，并根据测序结果分析获得基因的甲基化信息。

本申请的 DNA 甲基化检测方法，相比现有技术，有以下优点：

- 1) 通过 stLFR 自身具有的 co-barcoding 技术，获得长片段的甲基化信息，解决了甲基化测序中读长偏短的问题；
- 2) 可以进行长片段 DNA 序列检测，利于新物种的从头测序，基因组组装；
- 3) 可以在不需要参考序列的前提下，对某一特定 DNA 片段进行快速准确的比对；
- 4) 可以对多倍型基因组不同染色体的甲基化水平进行测定；
- 5) 适用于对基于 oxBS-seq 原理的 5mC 的测序；
- 6) 适用于去甲基酶 Tet 和 AID/APOBEC 参与的 5hmC 的测序；
- 7) 适用于 TAPS-seq 原理的甲基化测序（5mC）、羟甲基化测序（5hmC）、羧基胞嘧啶（5caC）的测序。

本申请的 DNA 甲基化检测方法同样适用于甲基化测序（5mC）和羟甲基化测序（5hmC），详细说明如下：

1. 采用自 BGI 自产的甲基化测序试剂盒，对 DNA 样品进行常规的 Bisulfite 处理，按照专利 W02019217452A1 所描述建库方法建库上机，可以获得 5mC 测序结果。Bisulfite 试剂的主要成分是重重亚硫酸氢钠，其可以将非甲基化的胞嘧啶转化为 U，而 5mC 和 5hmC 则不会被转化，从而获得 DNA 的甲基化信息。Bisulfite 测序无法区分甲基化的类型是 5mC 还是 5hmC，得到的实际上是一个混合的信号。

2. 利用 oxBS-seq 的原理，先将文库进行氧化处理，将 5hmC 转化为羧基胞嘧啶（5caC），再进行 Bisulfite 处理，未甲基化的 C 和 5caC 都会转化为 U，而 5mC 则不会被转化，从而得到完整的 5mC 测序信息。

3. 利用 TAB-seq 或 ACE-seq 原理，先对 5hmC 进行糖基化保护，再用 Tet 酶或 A3A 酶对 DNA 进行去甲基化处理，将其他没有糖基化保护的 5mC 转化为非甲基化的 C，再进行 Bisulfite 处理，从而获得完整的 5hmC 的测序信息。

4. 利用 TAPS-seq 的原理，先将文库进行氧化处理，将 5mC 和 5hmC 转化为羧基胞嘧啶（5caC），再进行硼烷处理，将其转化为 U，从而获得 5mC 和 5hmC 的测序信息。

基于本申请的 DNA 甲基化检测方法，本申请研发了一种 DNA 甲基化检测装置，如图 2 所示，包括长片段 DNA 处理模块 21、常规文库构建模块 22、常规测序模块 23、双链保护模块 24、重亚硫酸盐转化模块 25、甲基化测序文库构建模块 26 和甲

基化测序模块 27；长片段 DNA 处理模块 21，包括用于将长片段 DNA 打断成小片段 DNA；常规文库构建模块 22，包括用于采用 stLFR 技术进行文库构建；常规测序模块 23，包括用于取用采用 stLFR 技术构建的常规文库，将 DNA 变性为单链，然后再进行常规 PCR 扩增或环化后滚环扩增，获得测序文库，对测序文库进行测序，获得正常基因测序结果；双链保护模块 24，包括用于取用部分采用 stLFR 技术构建的常规文库，磁珠回收，将 DNA 变性为单链，然后利用退火或延伸的方法将分子标签序列或测序接头序列转化为双链；重亚硫酸盐转化模块 25，包括用于向经过双链保护模块处理的产物中加入双链保护剂，并加入重亚硫酸盐，进行脱氨基反应；甲基化测序文库构建模块 26，包括用于对脱氨基反应的产物进行常规 PCR 扩增或环化后滚环扩增，得到甲基化测序文库；甲基化测序模块 27，包括用于对甲基化测序文库进行测序，并进行数据分析，将甲基化测序结果与正常基因测序结果比对，获得基因的甲基化信息。

本申请的 DNA 甲基化检测方法和装置，不但具有 stLFR 的优点，同时还能够给出精确的甲基化信息：

1) 在现有的甲基化测序中，因实验原理限制，通常只能得到较短的读长；而本申请通过 co-barcoding 技术，可以将短读长组装成长片段，给出更加完整的甲基化图谱；

2) 基于共建库的原理，对于 WGS 和 WGBS 文库进行比对，可以在无参考文库的情况下，在分子级别对处理前后的文库进行一对一比对，得到更加精确的甲基化信息；

3) 通过双链保护机制，有效的避免分子标签信息变异，节约时间及合成成本。

下面通过具体实施例对本申请作进一步详细说明。以下实施例仅对本申请进行进一步说明，不应理解为对本申请的限制。

### 实施例

本例以人基因组 DNA 测序为例，进行 DNA 甲基化检测，其过程包括：首先提取长片段基因组 DNA，再用 Tn5 转座酶将 DNA 打断成小片段，用带有分子标签序列的磁珠进行捕获。采用 stLFR 技术进行文库构建，获得常规文库。得到常规建库的文库后，进行常规的 PCR 扩增，吸取 PCR 产物上清得到文库 L，进行常规的上机测序操作得到 WGS 数据。同时回收常规文库的磁珠用于甲基化测序建库；将回收的磁珠上的 DNA 分子充分变性为单链，然后利用磁珠上分子中自身的 3' 端序列作为引物或加入一个外源的短片段作为引物，如图 3 所示，在 DNA 聚合酶作用下，将其中的 DNA 分子标签序列变为双链结构，加入双链保护剂并进行重亚硫酸氢钠转化反应，反应完成后进行常规的 PCR 扩增得到文库 M，对文库 M 进行上机测序操作得到 WGBS 数据；最后将 WGS 和 WGBS 数据进行一对一比对，确定甲基化位点。本例的 DNA 甲基化检测详细如下：

## 一、制备带有多拷贝的分子标签的载体

(1) 将接头序列 (linker) 通过生物素-链霉亲和素作用连接到修饰有链霉亲和素蛋白的磁珠上。Linker 为双链 DNA 分子, 由两条单链 DNA 链退火在一起, 两条单链 DNA 的退火条件为 70℃ 退火 1 分钟, 然后以 0.1℃/s 的速度缓慢降温至 20℃, 20℃ 反应 30 分钟。

Linker 的双链 DNA 分子中正义链为 Linker-F, 反义链为 Linker-R; 其中, Linker-F 为 SEQ ID NO.1 所示序列, Linker-F 的 3' 端具有 Bio 修饰, Linker-R 为 SEQ ID NO.2 所示序列。

SEQ ID NO.1:

5' -CTTCCGGCAGAACGACATGGCTACGAAAAAAAAA-3' Bio

SEQ ID NO.2: 5' -CGTAGCCATGTCGTTCTGCC-3'。

本例采用的带有链霉亲和素的磁珠为 Dynabeads M-280 streptation (112.06D, Invitrogen)。按照 Linker (50 μM) 2 μL 与 M280 磁珠 30 μL 的比例混合; 磁珠的保存液替换成 1 倍浓度的磁珠结合缓冲液, 25℃, 在垂直混匀仪上混合 1 个小时, 然后用低盐磁珠洗涤缓冲液洗涤 2 次, 最后用 12.5 μL 的 1 倍浓度的连接缓冲液重悬磁珠。

其中, 磁珠结合缓冲液由 50mM 的 Tris-HCl、150mM 的 NaCl 和 0.1mM EDTA 组成; 低盐磁珠洗涤缓冲液由 50mM 的 Tris-HCl、150mM 的 NaCl 和 0.02% 的 Tween-20 组成; 正常配制的连接缓冲液为 10 倍浓度, 10 倍浓度的连接缓冲液, 其由聚乙二醇 8000 30%、Tris-HCl 150mM、ATP 1mM、BSA 0.15mg/mL、MgCl<sub>2</sub> 30mM、DTT 1.5mM 组成。

(2) 将不同编号, 即 1 至 1536 号, 的序列 1 和辅助序列 1 分装在 384 孔板的不同孔中进行退火, 1:1 退火, 一共 4 μL/孔, 共计 4 个 384 孔板。

标签序列 1 与辅助序列 1 的退火条件是: 100 μM 的序列 1 与辅助序列 1 按体积比 1:1 混合后置于 PCR 仪器上, 70℃ 1 分钟, 然后以 0.1℃/s 的速度缓慢降温至 20℃, 20℃ 反应 30 分钟。

其中, 标签序列 1 为 SEQ ID NO.3 所示序列, 辅助序列 1 为 SEQ ID NO.4 所示序列。标签序列 1 的 5' 端含有分子标签序列 "Barcode", 辅助序列 1 的第 22 碱基和第 23 碱基之间插入分子标签序列 "Barcode"。

SEQ ID NO.3: 5' -[Barcode]ACCCTGACTAGGTCGC-3'

SEQ ID NO.4:

5' -GGAAGGGCGACCTAGTCAGGGT[Barcode]CGCAGA-3'。

(3) 使用 DNA 连接酶将载体上的接头序列与序列 1 进行连接。该序列包含一段特定的 DNA 序列, 标记为分子标签 1。具体步骤为将步骤 (1) 中带有 Linker 的 M280 磁珠平均分配到 4 块 384 孔板的 1536 个孔中, 每孔 2.5 μL。再在每孔补上 3.5

$\mu\text{L}$  连接混合液，连接混合液含有  $1\mu\text{L}$  连接酶 (T4 DNA ligase 600U/ $\mu\text{L}$ , Enzymatics)， $1.5\mu\text{L}$  超纯水和  $1\mu\text{L}$  连接缓冲液 ( $10\times\text{T4 DNA ligation buffer}$ , Enzymatics)，以每个 384 孔板  $10\mu\text{L}$  的总反应体积进行连接反应，反应温度  $25^\circ\text{C}$ ，反应时间 1 个小时。使用  $100\mu\text{L}$  高盐磁珠洗涤缓冲液洗涤一次，再使用  $100\mu\text{L}$  低盐磁珠洗涤缓冲液洗涤一次，用于去除反应中的连接酶与未能完全反应的寡聚核苷酸。

其中，高盐磁珠洗涤缓冲液由  $50\text{mM Tris-HCl}$ 、 $500\text{mM NaCl}$ 、 $0.02\%$  Tween-20 组成；低盐磁珠洗涤缓冲液由  $50\text{mM Tris-HCl}$ 、 $150\text{mM NaCl}$ 、 $0.02\%$  Tween-20 组成。

(4) 将步骤 (3) 清洗后的磁珠通过磁力架收集，然后用 1 倍浓度的连接缓冲液重悬磁珠，重悬后 Linker 的浓度为  $1.6\mu\text{M}$ ，并使用震荡混匀仪混匀。

(5) 将不同编号，即 1 至 1536 号，的标签序列 2 和辅助序列 2 分装在全新的 384 孔板的不同孔中退火，共计 4 个 384 孔板。

标签序列 2 与辅助序列 2 的退火条件是：标签序列  $2100\mu\text{M}$  与辅助序列  $2100\mu\text{M}$  各  $2\mu\text{L}$  混合于 384 孔板中，然后置于 PCR 仪器上， $70^\circ\text{C}$  1 分钟，然后以  $0.1^\circ\text{C/s}$  的速度缓慢降温至  $20^\circ\text{C}$ ， $20^\circ\text{C}$  反应 30 分钟。

其中，标签序列 2 为 SEQ ID NO.5 所示序列，辅助序列 2 为 SEQ ID NO.6 所示序列。标签序列 2 的第 47 碱基和第 48 碱基之间插入分子标签序列“Barcode”，辅助序列 2 的 3' 端含有分子标签序列“Barcode”。

SEQ ID NO.5:

5

-GTACTGAGGGCTGGCGACCTTGAGTGCTTCGACTGGCCGTCGTTTTTA[Barcode]TCTGCG-3'

SEQ ID NO.6: 5' -TAAAACGACG[Barcode]-3'。

(6) 将步骤 (4) 混匀后的磁珠按每孔  $2.5\mu\text{L}$  的量分配到步骤 (5) 的 384 孔板的各个孔中。然后加入  $3.5\mu\text{L}$  连接混合液，连接混合液含有  $1\mu\text{L}$  连接酶 (T4 DNA ligase 600U/ $\mu\text{L}$ , Enzymatics)， $1.5\mu\text{L}$  超纯水和  $1\mu\text{L}$  连接缓冲液 ( $10\times\text{T4 DNA ligation buffer}$ , Enzymatics)，在  $25^\circ\text{C}$  反应 1 个小时。

(7) 使用  $100\mu\text{L}$  高盐磁珠洗涤缓冲液洗涤一次，再使用  $100\mu\text{L}$  低盐磁珠洗涤缓冲液洗涤一次，用于去除反应中的连接酶与未能完全反应的寡聚核苷酸；最后，用磁力架收集磁珠，再将磁珠重悬于低盐磁珠洗涤缓冲液中， $4^\circ\text{C}$  可保存 1 年。

至此，完成 2359296 种分子标签磁珠载体的制备，磁珠的浓度为  $19000000$  个/ $\mu\text{L}$ 。

## 二、stLFR 技术建库

(1) 准备带有标签的磁珠，取 1 千万个磁珠，即  $5.3\mu\text{L}$  磁珠，用磁力架去除磁珠中的低盐磁珠洗涤液，用  $50\mu\text{L}$  低盐磁珠洗涤液洗涤一次，然后用强碱变性缓冲液  $50\mu\text{L}$  重悬磁珠，在常温下孵育 5 分钟，用磁力架吸附磁珠 2 分钟，去除强碱

变性缓冲液, 再加入强碱变性缓冲液 50  $\mu$ L 洗涤磁珠一次, 磁力架吸附磁珠 2 分钟, 去除强碱变性缓冲液, 用低盐磁珠洗涤缓冲液 50  $\mu$ L 洗涤一次, 用杂交缓冲液 50  $\mu$ L 洗涤一次, 最后用杂交缓冲液 50  $\mu$ L 重悬磁珠。

其中, 强碱变性缓冲液由 1.6M 的 KOH 和 1mM 的 EDTA 组成; 杂交缓冲液由 50mM Tris-HCl、1000mM NaCl、0.05% Tween-20 组成。

(2) 取 20  $\mu$ L 长片段 DNA (2ng/ $\mu$ L), 缓慢加入 27  $\mu$ L 超纯水、12  $\mu$ L 打断缓冲液 (5 $\times$ TI Buffer, MGIEasy stLFR) 和 1  $\mu$ L 转座复合体 (2 pmol/ $\mu$ L, MGIEasy stLFR), 共 60  $\mu$ L 反应液, 55 $^{\circ}$ C 孵育 10 分钟后置于冰上。取 15  $\mu$ L 反应液, 加入 35  $\mu$ L 超纯水, 稀释至 50  $\mu$ L, 标记为打断 DNA 溶液。

将 50  $\mu$ L 的打断 DNA 溶液与步骤 (1) 中带标签 (Barcode) 的磁珠 50  $\mu$ L 混合, 60 $^{\circ}$ C 反应 1 分钟, 然后让反应液置于室温, 让其自然降温并置于垂直混匀仪, 25 $^{\circ}$ C 混合反应 1 个小时, 标记为杂交捕获体系。

(3) 使用转座复合体 1 时, 在 1 个小时的杂交时间结束后, 加入 100  $\mu$ M 的封闭序列 1 (Blocker1) 2  $\mu$ L 到 100  $\mu$ L 杂交捕获体系中, 置于垂直混匀仪上 25 $^{\circ}$ C 反应 0.5 个小时。然后用低盐磁珠洗涤液洗涤磁珠两次, 加入聚合酶和连接酶的混合反应液重悬磁珠, 置于 20 $^{\circ}$ C 反应 0.5 小时。

其中, 混合反应液中含有: DNA 聚合酶 T4 DNA Polymerase (3U/ $\mu$ L, Enzymatics) 6  $\mu$ L、连接酶 T7 DNA ligase (30000U/ $\mu$ L, NEB) 1  $\mu$ L、2 $\times$ T7 DNA ligase buffer (NEB) 25  $\mu$ L、25mM dNTP 0.5  $\mu$ L、10mM ATP 5  $\mu$ L, 补足体积的水至 50  $\mu$ L。

2 $\times$ T7 DNA ligase buffer 中含有: 132mM Tris-HCl、20mM MgCl<sub>2</sub>、2mM ATP、2mM DTT、15% Polyethyleneglycol (聚乙二醇 6000)。

封闭序列 1 为 SEQ ID NO.7 所示序列。

SEQ ID NO.7:

5' -CCTAGCATGGACTATCGATCCTTGGTGATCATGTCGTCAGTGC-3' Bio

反应结束后加入 5  $\mu$ L 的 0.44% SDS 室温下孵育 10 分钟使转座酶变性从 DNA 上释放下来, 然后用高盐磁珠洗涤液和低盐磁珠洗涤液分别洗涤一次。加入聚合酶试剂混合液重悬磁珠, 置于 72 $^{\circ}$ C 反应 10 分钟。

其中, 聚合酶试剂混合液中含有: 聚合酶 Standard Taq polymerase (5U/ $\mu$ L, NEB) 1  $\mu$ L、10 $\times$ thermopol buffer (NEB) 5  $\mu$ L、25mM dNTP (Enzymatics) 0.8  $\mu$ L, 总体积 50  $\mu$ L。反应结束后用磁力架吸附磁珠, 回收上清液。

10 $\times$ thermopol buffer 中含有: 200mM Tris-HCl、100mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>、100mM KCl、20mM MgSO<sub>4</sub>、1% Triton®X-100。

(4) 使用转座复合体 2 时, 在 1 个小时的杂交时间结束后, 加入 100  $\mu$ M 的封闭序列 2 (Blocker2) 2  $\mu$ L 到 100  $\mu$ L 的杂交捕获体系中, 置于垂直混匀仪上 25 $^{\circ}$ C 反应 0.5 个小时。加入 0.44% SDS 10  $\mu$ L 室温下孵育 10 分钟使转座酶变性从 DNA

分子上释放下来。用 100  $\mu$ L 高盐磁珠洗涤液和低盐磁珠洗涤液各洗涤一次，用 18  $\mu$ L 的 TE 缓冲液重悬磁珠，再加入 USER 酶 (1U/ $\mu$ L, NEB) 2  $\mu$ L，置于垂直混匀仪上 37 $^{\circ}$ C 反应 0.5 个小时，在 37 $^{\circ}$ C 环境下，用 37 $^{\circ}$ C 预热的高盐磁珠洗涤液和低盐磁珠洗涤缓冲液各洗涤一次。用连接酶试剂混合液重悬磁珠，置于垂直混匀仪上 25 $^{\circ}$ C 反应 1 个小时。连接酶试剂混合液中包含连接酶，其组成为：T4 DNA ligase (600U/ $\mu$ L, Enzymatics) 5  $\mu$ L、3 倍浓度的连接缓冲液 10  $\mu$ L 和 16.7  $\mu$ M 接头 21.5  $\mu$ L，总体积为 30  $\mu$ L。

其中，3 倍浓度的连接缓冲液中含有：30% 聚乙二醇 8000、150mM Tris-HCl、1mM ATP、0.15mg/mL BSA、30mM MgCl<sub>2</sub>、1.5mM DTT。16.7  $\mu$ M 接头，即接头 2，由正义链接头 2-F 和反义链接头 2-R 退火形成。接头 2-F 为 SEQ ID NO.9 所示序列，接头 2-F 的 5' 端具有磷酸化修饰，3' 端为双脱氧碱基；接头 2-R 为 SEQ ID NO.10 所示序列，接头 2-R 的 3' 端为双脱氧碱基。

SEQ ID NO.9: 5' -phos-TCTGCTGAGTCGAGAACGTCTddC-3'

SEQ ID NO.10: 5' -CTCGACTCAGCAGddA-3'

反应结束后高盐磁珠洗涤液和低盐磁珠洗涤缓冲液各洗涤一次。用 2 $\times$  T7 ligase buffer (NEB) 10  $\mu$ L、1  $\mu$ L 浓度 100  $\mu$ M 的引物 2、1  $\mu$ L 浓度 100  $\mu$ M 的引物 3 和 8  $\mu$ L 水配制成 20  $\mu$ L 退火混合液，用退火混合液重悬磁珠，70 $^{\circ}$ C 1 分钟，以 0.1 $^{\circ}$ C/秒的速度缓慢降温至 20 $^{\circ}$ C，20 $^{\circ}$ C 反应 30 分钟。然后，直接加入聚合酶和连接酶混合液试剂 20 $^{\circ}$ C 反应 0.5 个小时。

引物 2 为 SEQ ID NO.12 所示序列，引物 2 的 5' 端具有磷酸化修饰；引物 3 为 SEQ ID NO.8 所示序列，封闭序列 1 的 3' 端具有 Bio 修饰。

SEQ ID NO.8:

5' -CGTAGCCATGTCGTTCTGCCGAAGGGCGACCTAGTCAGGGT-3'

SEQ ID NO.12: 5' -phos-GAGACGTTCTCGACTCAGCAGA-3'。

其中，聚合酶和连接酶混合试剂包含：聚合酶 (T4 DNA Polymerase 3U/ $\mu$ L, Enzymatics) 6  $\mu$ L、连接酶 (T7 DNA ligase 3000U/ $\mu$ L, NEB) 1  $\mu$ L、2 $\times$ T7 DNA ligase buffer (NEB) 15  $\mu$ L、25mM dNTP 0.5  $\mu$ L、10mM ATP 5  $\mu$ L，补足体积的水至 30  $\mu$ L。

反应结束后用 100  $\mu$ L 高盐磁珠洗涤液和低盐磁珠洗涤缓冲液各洗涤一次。然后加入有链置换活性的聚合酶试剂混合液重悬磁珠，置于 65 $^{\circ}$ C 反应 5 分钟。聚合酶试剂混合液中含有：聚合酶 (Bst2.0 WarmStart<sup>®</sup> DNA Polymerase 8U/ $\mu$ L, NEB) 1  $\mu$ L、10 $\times$ Isothermal Amplification Buffer (NEB) 5  $\mu$ L、25mM dNTP (Enzymatics) 0.8  $\mu$ L、100mM MgSO<sub>4</sub> 1.5  $\mu$ L，总体积 50  $\mu$ L。反应结束后用磁力架吸附磁珠，回收上清液。

其中，10 $\times$ Isothermal Amplification Buffer 中含有：200mM Tris-HCl、100mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>、500mM KCl、20mM MgSO<sub>4</sub>、1% Tween<sup>®</sup>20。

(5) 使用引物 1 和引物 2 进行 DNA 分子聚合酶链式扩增 13 个循环。

反应体系如下：

50  $\mu$ L 步骤 (4) 中回收的上清液，75  $\mu$ L 2 $\times$  KAPA HiFi Master Mix (Kapa Biosystems)，0.75  $\mu$ L 引物 1 (100  $\mu$ mol/L)，0.75  $\mu$ L 引物 2 (100  $\mu$ mol/L)，23.5  $\mu$ L 超纯水。

反应条件如下：

I. 98 $^{\circ}$ C 3 分钟；II. 95 $^{\circ}$ C 30 秒，58 $^{\circ}$ C 30 秒，72 $^{\circ}$ C 2 分钟，重复步骤 (II) 13 次；III. 72 $^{\circ}$ C 5 分钟；IV. 4 $^{\circ}$ C 保存。

其中，引物 1 为 SEQ ID NO. 11 所示序列；引物 2 为 SEQ ID NO. 12 所示序列，引物 2 的 5' 端具有磷酸化修饰。

SEQ ID NO. 11: 5' -CGTAGCCATGTCGTTCTG-3'。

### 三、常规测序

将“二、stLFR 技术建库”得到的 PCR 产物置于磁力架上，小心吸取上清液，用 AMPure XP beads (Agencourt AMPure XP-Medium, A63882, AGENCOURT) 进行纯化，纯化方法参考官方提供的标准操作规程，纯化结束后回收产物为带有分子标签的适合测序的短片段分子。

由于本例采用 BGISEQ-500 进行测序，因此，需要对磁珠纯化的产物进行成环反应，操作详情参照 BGISEQ-500 标准 DNA 小片段建库流程的成环步骤。然后，对环化的产物进行 BGISEQ-500 测序。

### 四、双链 DNA 保护

将 stLFR 文库磁珠转移至 1.5mL 离心管中，用 500  $\mu$ L 的 1 $\times$  Buffer D (0.1M NaOH) 充分洗涤 2 次，将离心管置于磁力架上，弃掉上清液，回收磁珠。再用低盐磁珠洗涤缓冲液充分洗涤 2 次，弃掉上清液，回收磁珠。向回收的磁珠中加入 47  $\mu$ L 的超纯水、50  $\mu$ L 的 2 $\times$  PfuC $\alpha$  Mix buffer (Agilent)，2  $\mu$ L 的 PfuC $\alpha$  聚合酶 (Agilent)，以及 1  $\mu$ L 的 ME 引物。其中，ME 引物是对应于文库中靠近 barcode 的 3' 末端的一段互补序列，ME 引物为 SEQ ID NO. 13 所示序列。

SEQ ID NO. 13: 5' -CTGTCTCTTATACACATCT-3'。

将混合物及磁珠充分混匀，离心并转移至 250  $\mu$ L 的 PCR 管中，在 PCR 仪器上运行如下程序：98 $^{\circ}$ C 3 分钟，50 $^{\circ}$ C 1 分钟，72 $^{\circ}$ C 10 分钟，室温保持。

将反应物取出，至于冰上，并用高盐磁珠洗涤液洗涤 1 次，低盐磁珠洗涤缓冲液洗涤 2 次，移去上清，回收磁珠于 1.5mL 的离心管中。

### 五、重亚硫酸氢钠转化和甲基化测序文库

按照 BGI 甲基化测序试剂盒的实验步骤，完成重亚硫酸氢钠转化反应，详细如下：

1. 加入 20  $\mu$ L 双链 DNA 保护 buffer, 加入 130  $\mu$ L 的 Bisulfite Buffer (Zymo BioTech), 充分混匀, 用锡箔纸覆盖管体, 避光。

2. 将离心管置于旋转混合仪上, 在 60°C, 孵育 1 个小时。

3. 将离心管置于磁力架上 2 分钟, 弃去上清液, 加入 400  $\mu$ L 的 Wash buffer (Zymo BioTech), 前后置换离心管位置, 充分洗涤磁珠后, 弃去上清液。

4. 加入 200  $\mu$ L 的 Desulphonation buffer (Zymo BioTech), 在室温下孵育 25-30 分钟, 不时轻微震荡离心管, 防止磁珠沉降。

5. 将磁珠置于磁力架上 2 分钟, 弃去上清液, 加入 400  $\mu$ L 的低盐磁珠洗涤缓冲液, 前后置换离心管位置, 充分洗涤磁珠后, 弃去上清液, 将磁珠重悬于 TE buffer 中。

6. 取 10M 磁珠加入 PCR Mix 并补水至体积为 50  $\mu$ L。

其中, PCR Mix 包括: 25  $\mu$ L 的 2 $\times$ PfuCx Mix (Agilent)、0.25  $\mu$ L 的浓度为 100  $\mu$ M 的 PCR Forward primer、0.25  $\mu$ L 的浓度为 100  $\mu$ M 的 PCR Reverse primer、0.75  $\mu$ L 的 PfuCx polymerase (Agilent)。

PCR Forward primer 为 SEQ ID NO.14 所示序列, PCR Reverse primer 为 SEQ ID NO.15 所示序列。

SEQ ID NO.14: 5' -TGTGAGCCAAGGAGTTG-3'

SEQ ID NO.15: 5' -GAGACGTTCTCGACTCAGCAGA-3'。

将 PCR 管置于 PCR 仪器上运行如下程序: 98°C 3 分钟, 然后进入 10 个循环: 98°C 30 秒、58°C 30 秒、72°C 2 分钟, 循环结束后, 72°C 延伸 10 分钟, 然后降温至 4°C 保存。

7. 用 0.8 $\times$  的 AMPure XP beads (Agencourt AMPure XP-Medium, A63882, AGENCOURT) 对 PCR 产物进行纯化, 并将 PCR 产物重悬至 20  $\mu$ L 超纯水中, 即得到 LFR 甲基化测序文库。

## 六、甲基化测序

由于本例采用 BGISEQ-500 进行测序, 因此, 需要对磁珠纯化的产物进行成环反应, 操作详情参照 BGISEQ-500 标准 DNA 小片段建库流程的成环步骤。然后, 对环化的产物进行 BGISEQ-500 测序。

## 七、实验结果

### 1. 文库电泳结果

甲基化测序文库及对应的 stLFR 文库制备的结果, 本例用 1.5% 琼脂糖凝胶 140V、30 分钟跑电泳, 结果如图 7 所示。图 7 中, 第一泳道为 DNA marker, 本例采用的是 Thermofisher 1kb plus DNA ladder; 第二和第三泳道为 WGS 的 S1 和 S2, 即正常基因测序的 stLFR 测序文库的两个重复; 第四和第五泳道为 WGBS 的 S1 和 S2, 即甲基化测序文库的两个重复。

## 2. 甲基化测序结果

本例采用相同的人基因组 DNA，按照现有的甲基化测序方法进行测序，作为对比。结果如表 1 所示。

表 1 测序结果比对

| 试验             | Clean_reads | Clean_bases(bp) | Mapped_reads | Mapping_rate(%) |
|----------------|-------------|-----------------|--------------|-----------------|
| 试验组一（常规文库）     | 156770568   | 15677056800     | 149334322    | 95.26           |
| 试验组二（甲基化文库）    | 156868710   | 15686871000     | 150202688    | 95.75           |
| 对照组 SRR6006942 | 157796950   | 23827339450     | 133582687    | 84.65           |

表 1 中，试验组一为本例 stLFR 技术建库后用于常规测序获得的数据，试验组二为本例 stLFR 技术建库后进行甲基化建库和测序获得的数据，对照组即按照常规的甲基化测序获得的数据。与对照组相比较，本例常规测序和甲基化测序结果的 mapping 效率都高于常规甲基化测序的对照组。并且，在无参考基因组的前提下，本例从试验组一和试验组二中分别抽取分子标签序列一致且长度一致的 reads 进行一对一比对，即可获得甲基化信息。例如，对于已知的低甲基化水平 LINE 基因片段：

试验组一的分子标签序列:1320\_499\_969 序列如 SEQ ID NO.16 所示，

SEQ ID NO.16:

5' -ATGGAACTGAACAACCTGCTCCTGAATGACTACTGGGTACATAACGAA-3'

试验组二的分子标签序列:1320\_499\_969 序列如 SEQ ID NO.17 所示，

SEQ ID NO.17:

5' -ATGGAAATGAATAATTTGTTTTGAATGATTATTGGGTATATAATGAA-3'

直接比较 SEQ ID NO.16 和 SEQ ID NO.17，即可确认该序列没有甲基化位点。而无需与人基因组参考序列进行比对，这对于本身没有参考基因组序列的物种而言，具有重要意义。

需要说明的是，以上试验只是本申请的一种具体实现方式，试验中采用的具体的反应体系、反应条件只是一个常规的参考，在此基础上可以根据具体试验进行改变。以上试验中试验的 DNA 片段化的酶，除 Tn5 转座酶以外，还可以是 Tn 转座酶家族的其他酶，如 Tn7；或者其他转座酶家族，如 Mu 家族；甚至不限于转座酶或者酶制剂，只要其能够使 DNA 被片段化同时将一段序列连接至 DNA 上即可。本实例涉及到的酶、缓冲液、使用的分子标签数量、载体、载体表面修饰、分子标签结构、加接头方案等可参考 W02019217452A1。

本申请的 DNA 甲基化检测可以用于全基因组甲基化测序、基因未知区域甲基化测序、已知特定区域的甲基化测序、同源染色体上等位基因的甲基化信息对比、单

倍体全基因组甲基化测序及组装。并且，可以是线粒体、叶绿体等的 DNA 甲基化检测。

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

## 权 利 要 求

1. 一种 DNA 甲基化检测的方法，其特征在于：包括以下步骤，  
长片段 DNA 处理步骤，包括将长片段 DNA 打断成小片段 DNA；  
常规文库构建步骤，包括采用 stLFR 技术进行文库构建，获得常规文库；  
双链保护步骤，包括取用所述常规文库，磁珠回收，将 DNA 变性为单链，然后利用退火或延伸的方法将分子标签序列或测序接头序列转化为双链；  
重亚硫酸盐转化步骤，包括向所述双链保护步骤的产物中加入双链保护剂，并加入重亚硫酸盐，进行脱氨基反应；  
甲基化测序文库构建步骤，包括对脱氨基反应的产物进行常规 PCR 扩增或环化后滚环扩增，得到甲基化测序文库；  
甲基化测序步骤，包括对甲基化测序文库进行测序，并根据测序结果分析获得基因的甲基化信息。
2. 根据权利要求 1 所述的方法，其特征在于：还包括常规测序步骤，  
所述常规测序步骤，包括取用所述常规文库构建步骤中采用 stLFR 技术构建的所述常规文库，将 DNA 变性为单链，然后再进行常规 PCR 扩增或环化后滚环扩增，获得测序文库，对测序文库进行全基因组测序，获得正常基因测序结果；  
所述甲基化测序步骤还包括，将甲基化测序文库的测序结果与所述常规文库获得的正常基因测序结果比对，从而获得基因的甲基化信息。
3. 根据权利要求 1 或 2 所述的方法，其特征在于：所述长片段 DNA 处理步骤还包括，采用带有分子标签序列的载体对 DNA 片段进行捕获，然后再将捕获的 DNA 用于所述常规文库构建步骤进行 stLFR 技术文库构建。
4. 根据权利要求 3 所述的方法，其特征在于：所述载体为磁珠。
5. 根据权利要求 1 或 2 所述的方法，其特征在于：所述长片段 DNA 处理步骤中，将长片段 DNA 打断成小片段 DNA 具体采用 Tn5 转座酶。
6. 根据权利要求 1 或 2 所述的方法，其特征在于：所述双链保护步骤中，延伸的方法具体包括，利用分子标签序列或测序接头序列自身的 3' 端序列作为引物或加入外源短片段作为引物，在 DNA 聚合酶作用下延伸，将分子标签序列或测序接头序列变为双链结构。
7. 根据权利要求 1 或 2 所述的方法，其特征在于：所述重亚硫酸盐转化步骤中，所述双链保护剂为盐离子、聚乙二醇、酶和有机溶剂中的至少一种，所述重亚硫酸盐为重亚硫酸氢钠。
8. 根据权利要求 1-7 任一项所述的方法在基因印记检测、胚胎发育检测或癌症预防检测中的应用。

9. 一种 DNA 甲基化检测的试剂盒，其特征在于：包括第一组试剂、第二组试剂、第三组试剂、第四组试剂和第五组试剂；

所述第一组试剂包括用于将长片段 DNA 打断成小片段 DNA 的试剂；

所述第二组试剂包括用于 stLFR 技术建库的试剂，该试剂用于制备常规文库；

所述第三组试剂包括用于通过退火或延伸的方法将分子标签序列或测序接头序列转化为双链的试剂；

所述第四组试剂包括双链保护剂和重亚硫酸盐，用于进行双链保护和脱氨基反应；

所述第五组试剂包括用于对脱氨基反应的产物进行常规 PCR 扩增或环化后滚环扩增的试剂，该试剂用于制备甲基化测序文库。

10. 根据权利要求 9 所述的试剂盒，其特征在于：还包括第六组试剂，所述第六组试剂包括用于对常规文库进行常规 PCR 扩增或环化后滚环扩增的试剂，该试剂用于制备常规测序文库，以便于进行全基因组测序，获得正常基因测序结果。

11. 根据权利要求 9 所述的试剂盒，其特征在于：所述第一组试剂还包括带有分子标签序列的载体，用于对 DNA 片段进行捕获。

12. 根据权利要求 11 所述的试剂盒，其特征在于：所述载体为磁珠，第一组试剂中还包括磁珠纯化的试剂。

13. 根据权利要求 9 所述的试剂盒，其特征在于：所述第一组试剂中，用于将长片段 DNA 打断成小片段 DNA 的试剂为 Tn5 转座酶及其缓冲液。

14. 根据权利要求 9 所述的试剂盒，其特征在于：所述第三组试剂中，用于通过退火或延伸的方法将分子标签序列或测序接头序列转化为双链的试剂，具体包括外源短核苷酸片段、DNA 聚合酶及其缓冲液，以及用于 DNA 延伸的 dNTPs。

15. 根据权利要求 9-14 任一项所述的试剂盒，其特征在于：所述第四组试剂中，双链保护剂为盐离子、聚乙二醇、酶和有机溶剂中的至少一种，重亚硫酸盐为重亚硫酸氢钠。

16. 根据权利要求 9-15 任一项所述的试剂盒在基因印记检测、胚胎发育检测或癌症预防检测中的应用。

17. 一种 DNA 甲基化检测的装置，其特征在于：包括长片段 DNA 处理模块、常规文库构建模块、常规测序模块、双链保护模块、重亚硫酸盐转化模块、甲基化测序文库构建模块和甲基化测序模块，

长片段 DNA 处理模块，包括用于将长片段 DNA 打断成小片段 DNA；

常规文库构建模块，包括用于采用 stLFR 技术进行文库构建，获得常规文库；

常规测序模块，包括取用所述常规文库构建模块中采用 stLFR 技术构建的所述常规文库，将 DNA 变性为单链，然后再进行常规 PCR 扩增或环化后滚环扩增，获得测序文库，对测序文库进行全基因组测序，获得正常基因测序结果；

双链保护模块，包括用于取用所述常规文库构建模块中采用 stLFR 技术构建的所述常规文库，磁珠回收，将 DNA 变性为单链，然后利用退火或延伸的方法将分子标签序列或测序接头序列转化为双链；

重亚硫酸盐转化模块，包括用于向经过所述双链保护模块处理的产物中加入双链保护剂，并加入重亚硫酸盐，进行脱氨基反应；

甲基化测序文库构建模块，包括用于对脱氨基反应的产物进行常规 PCR 扩增或环化后滚环扩增，得到甲基化测序文库；

甲基化测序模块，包括用于对甲基化测序文库进行测序，将甲基化测序结果与正常基因测序结果比对，获得基因的甲基化信息。

18. 根据权利要求 17 所述的装置，其特征在于：所述长片段 DNA 处理模块，还包括用于采用带有分子标签序列的载体对 DNA 片段进行捕获，然后再将捕获的 DNA 用于所述常规文库构建模块采用 stLFR 技术进行文库构建。

19. 根据权利要求 18 所述的装置，其特征在于：所述载体为磁珠。

20. 根据权利要求 17 所述的装置，其特征在于：所述长片段 DNA 处理模块中，将长片段 DNA 打断成小片段 DNA 具体采用 Tn5 转座酶。

21. 根据权利要求 17 所述的装置，其特征在于：所述双链保护模块中，延伸的方法具体包括，利用分子标签序列或测序接头序列自身的 3' 端序列作为引物或加入外源短片段作为引物，在 DNA 聚合酶作用下延伸，将分子标签序列或测序接头序列变为双链结构。

22. 根据权利要求 17-21 任一项所述的装置，其特征在于：所述重亚硫酸盐转化模块中，所述双链保护剂为盐离子、聚乙二醇、酶和有机溶剂中的至少一种，所述重亚硫酸盐为重亚硫酸氢钠。

23. 根据权利要求 17-22 任一项所述的装置在基因印记检测、胚胎发育检测或癌症预防检测中的应用。

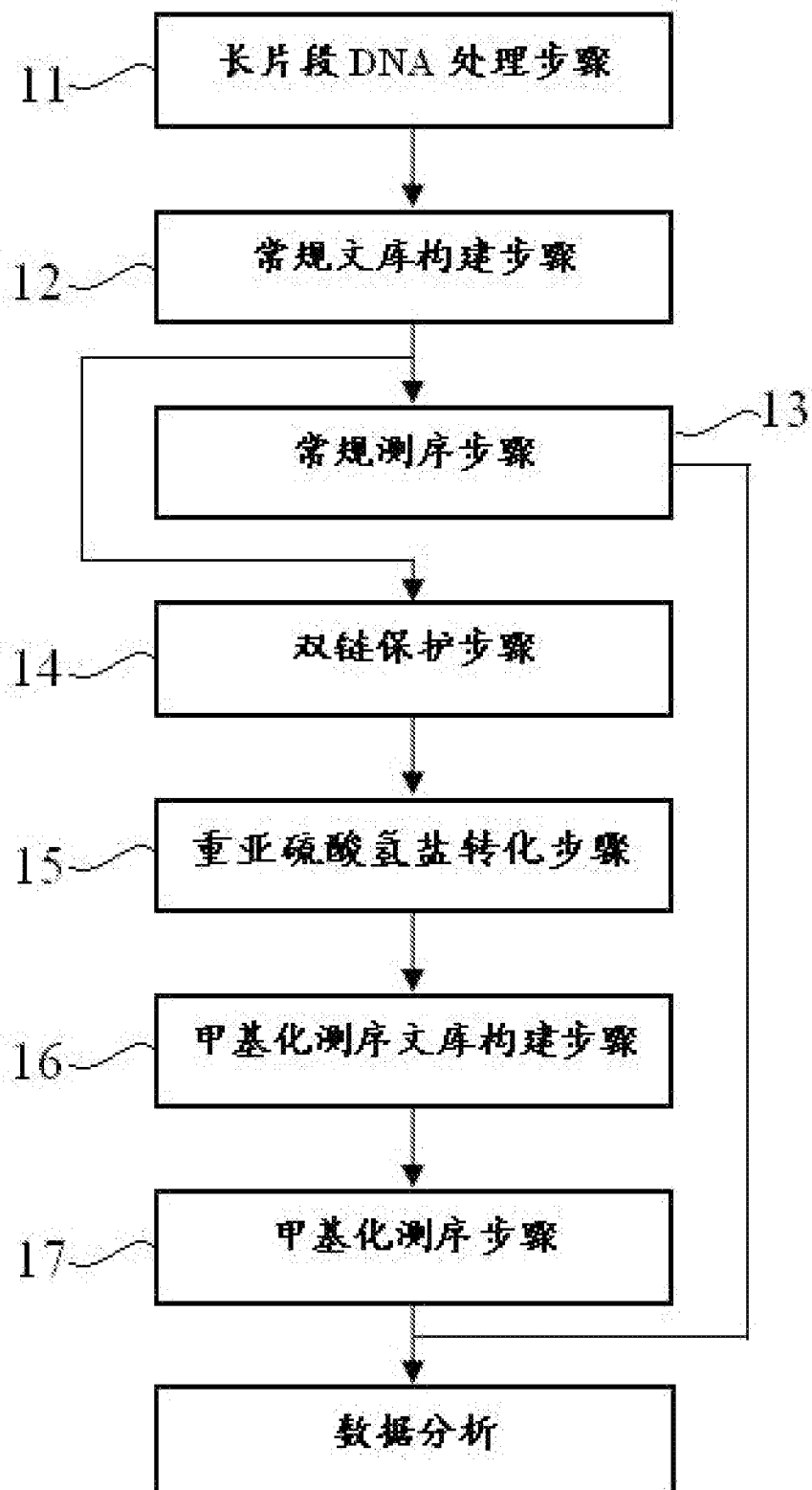


图 1

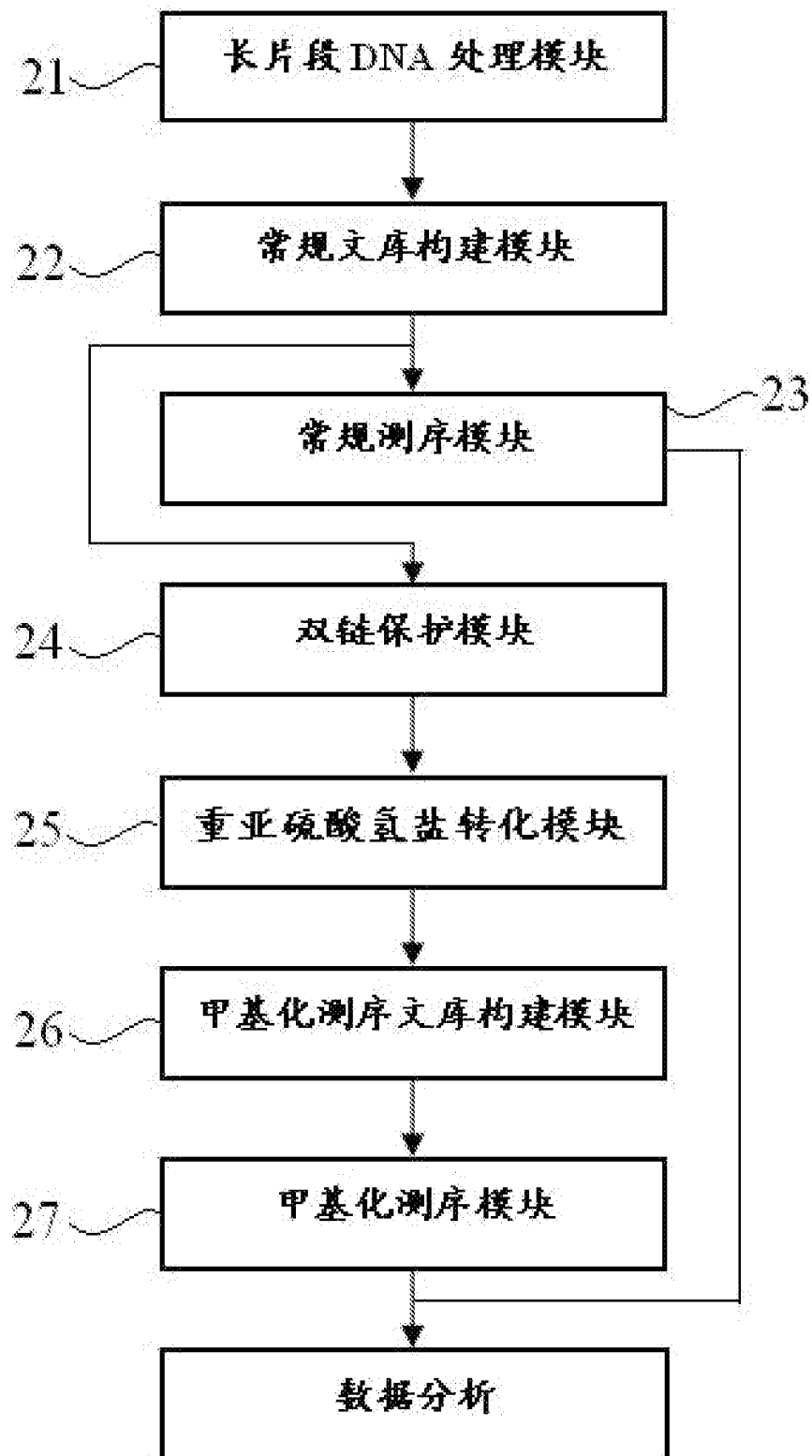


图 2

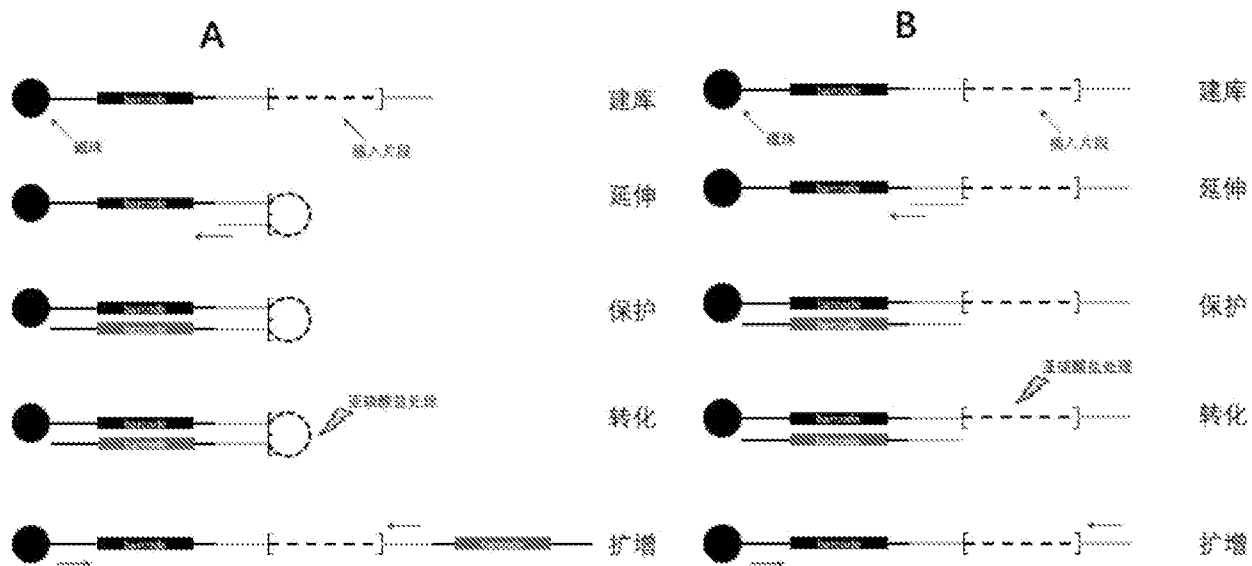


图 3

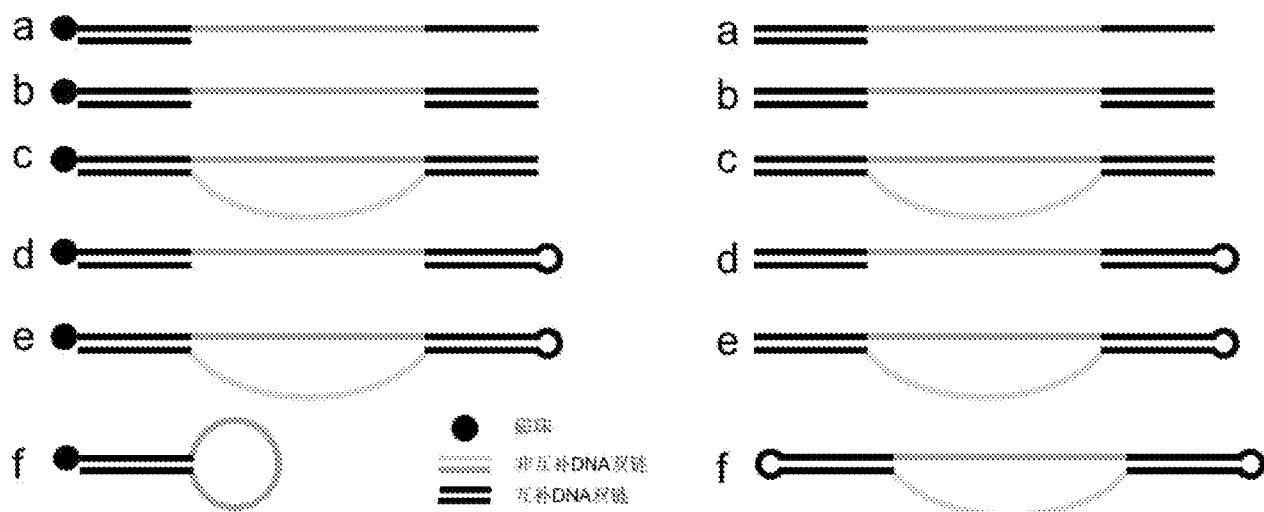


图 4

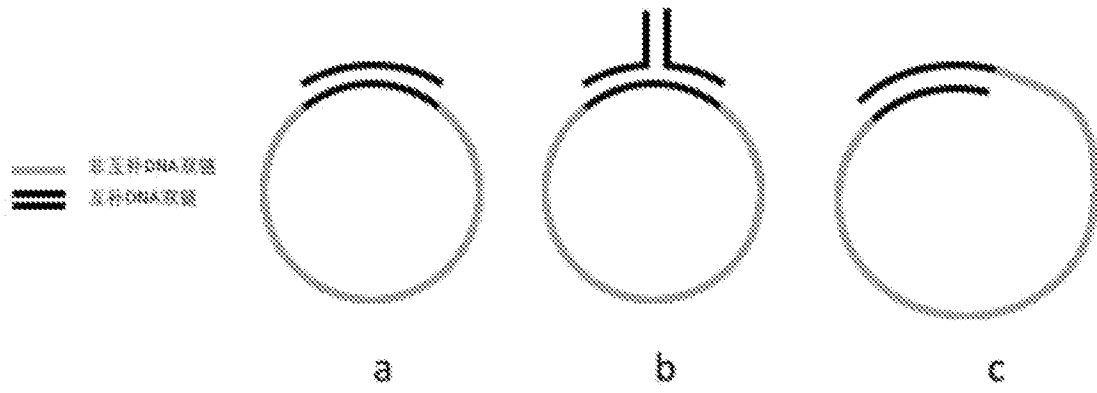


图 5

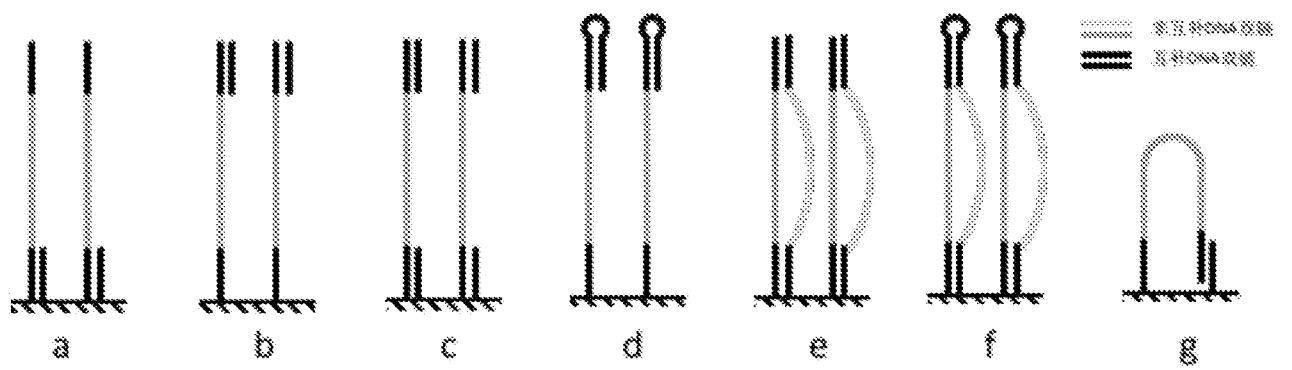


图 6

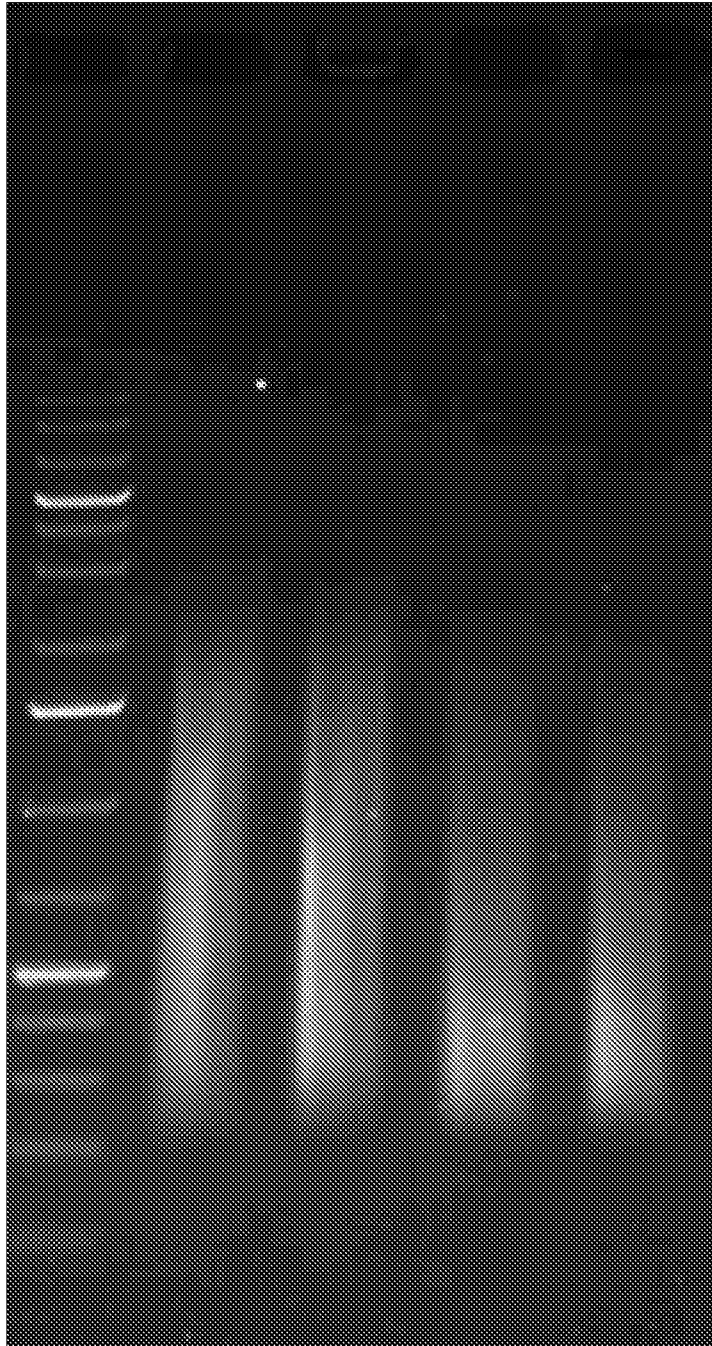


图 7

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2019/127537

**A. CLASSIFICATION OF SUBJECT MATTER**

C12Q 1/68(2018.01)i; G01N 33/00(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)

C12Q; G01N

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

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

DWPI, SIPOABS, CNABS, CNKI, ISI Web of knowledge, NCBI, Google Scholar: 甲基化, 文库, 标签, 条码, 重亚硫酸盐, 保护, 双链, 保护, stLFR, Methylation, library, tag, barcode, bisulfite, protection, double-strand, protection

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

| Category* | Citation of document, with indication, where appropriate, of the relevant passages                                     | Relevant to claim No. |
|-----------|------------------------------------------------------------------------------------------------------------------------|-----------------------|
| A         | CN 103088433 A (SHENZHEN BGI CO., LTD.; SHENZHEN BGI RESEARCH INSTITUTE) 08 May 2013 (2013-05-08)<br>entire document   | 1-23                  |
| A         | CN 103233072 A (OCEAN UNIVERSITY OF CHINA) 07 August 2013 (2013-08-07)<br>entire document                              | 1-23                  |
| A         | CN 102409408 A (SHENZHEN BGI CO., LTD.; SHENZHEN BGI RESEARCH INSTITUTE) 11 April 2012 (2012-04-11)<br>entire document | 1-23                  |
| A         | WO 2018175202 A1 (THE JOHNS HOPKINS UNIVERSITY) 27 September 2018 (2018-09-27)<br>entire document                      | 1-23                  |

☐ 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

13 February 2020

Date of mailing of the international search report

04 March 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

Authorized officer

Facsimile No. (86-10)62019451

Telephone No.

**INTERNATIONAL SEARCH REPORT**

International application No.

**PCT/CN2019/127537****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/CN2019/127537**

| Patent document<br>cited in search report |            |    | Publication date<br>(day/month/year) | Patent family member(s) |            |    | Publication date<br>(day/month/year) |
|-------------------------------------------|------------|----|--------------------------------------|-------------------------|------------|----|--------------------------------------|
| CN                                        | 103088433  | A  | 08 May 2013                          | WO                      | 2013064066 | A1 | 10 May 2013                          |
|                                           |            |    |                                      | CN                      | 103088433  | B  | 24 September 2014                    |
| CN                                        | 103233072  | A  | 07 August 2013                       | CN                      | 103233072  | B  | 02 July 2014                         |
| CN                                        | 102409408  | A  | 11 April 2012                        | WO                      | 2012037878 | A1 | 29 March 2012                        |
|                                           |            |    |                                      | HK                      | 1168393    | A1 | 22 November 2013                     |
|                                           |            |    |                                      | CN                      | 102409408  | B  | 07 August 2013                       |
| WO                                        | 2018175202 | A1 | 27 September 2018                    | CN                      | 110582580  | A  | 17 December 2019                     |

## 国际检索报告

国际申请号

PCT/CN2019/127537

|                                                                                                                                                                                                                                                                                                                                                                                                  |                                                                             |                             |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------|-----------------------------|
| <b>A. 主题的分类</b><br>C12Q 1/68(2018.01)i; G01N 33/00(2006.01)i<br><br>按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类                                                                                                                                                                                                                                                                                              |                                                                             |                             |
| <b>B. 检索领域</b><br>检索的最低限度文献(标明分类系统和分类号)<br>C12Q; G01N<br><br>包含在检索领域中的除最低限度文献以外的检索文献<br><br>在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用))<br>DWPI、SIP0ABS、CNABS、CNKI、ISI Web of knowledge、NCBI、Google Scholar: 甲基化、文库、标签、条码、重亚硫酸盐、保护、双链、保护、stLFR、Methylation、library、tag、barcode、bisulfite、protection、double-strand、protection                                                                            |                                                                             |                             |
| <b>C. 相关文件</b>                                                                                                                                                                                                                                                                                                                                                                                   |                                                                             |                             |
| 类 型*                                                                                                                                                                                                                                                                                                                                                                                             | 引用文件, 必要时, 指明相关段落                                                           | 相关的权利要求                     |
| A                                                                                                                                                                                                                                                                                                                                                                                                | CN 103088433 A (深圳华大基因科技有限公司 深圳华大基因研究院) 2013年 5月 8日 (2013 - 05 - 08)<br>全文  | 1-23                        |
| A                                                                                                                                                                                                                                                                                                                                                                                                | CN 103233072 A (中国海洋大学) 2013年 8月 7日 (2013 - 08 - 07)<br>全文                  | 1-23                        |
| A                                                                                                                                                                                                                                                                                                                                                                                                | CN 102409408 A (深圳华大基因科技有限公司 深圳华大基因研究院) 2012年 4月 11日 (2012 - 04 - 11)<br>全文 | 1-23                        |
| A                                                                                                                                                                                                                                                                                                                                                                                                | WO 2018175202 A1 (UNIV JOHNS HOPKINS) 2018年 9月 27日 (2018 - 09 - 27)<br>全文   | 1-23                        |
| <input type="checkbox"/> 其余文件在C栏的续页中列出。 <input checked="" type="checkbox"/> 见同族专利附件。                                                                                                                                                                                                                                                                                                             |                                                                             |                             |
| * 引用文件的具体类型:<br>“A” 认为不特别相关的表示了现有技术一般状态的文件<br>“E” 在国际申请日的当天或之后公布的在先申请或专利<br>“L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件(如具体说明的)<br>“O” 涉及口头公开、使用、展览或其他方式公开的文件<br>“P” 公布日先于国际申请日但迟于所要求的优先权日的文件<br>“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件<br>“X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性<br>“Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性<br>“&” 同族专利的文件 |                                                                             |                             |
| 国际检索实际完成的日期<br>2020年 2月 13日                                                                                                                                                                                                                                                                                                                                                                      |                                                                             | 国际检索报告邮寄日期<br>2020年 3月 4日   |
| ISA/CN的名称和邮寄地址<br>中国国家知识产权局(ISA/CN)<br>中国北京市海淀区蓟门桥西土城路6号 100088<br>传真号 (86-10)62019451                                                                                                                                                                                                                                                                                                           |                                                                             | 受权官员<br>张起<br>电话号码 62411036 |

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/CN2019/127537

| 检索报告引用的专利文件 |            |    | 公布日<br>(年/月/日) | 同族专利 |            |    | 公布日<br>(年/月/日) |
|-------------|------------|----|----------------|------|------------|----|----------------|
| CN          | 103088433  | A  | 2013年 5月 8日    | WO   | 2013064066 | A1 | 2013年 5月 10日   |
|             |            |    |                | CN   | 103088433  | B  | 2014年 9月 24日   |
| CN          | 103233072  | A  | 2013年 8月 7日    | CN   | 103233072  | B  | 2014年 7月 2日    |
| CN          | 102409408  | A  | 2012年 4月 11日   | WO   | 2012037878 | A1 | 2012年 3月 29日   |
|             |            |    |                | HK   | 1168393    | A1 | 2013年 11月 22日  |
|             |            |    |                | CN   | 102409408  | B  | 2013年 8月 7日    |
| WO          | 2018175202 | A1 | 2018年 9月 27日   | CN   | 110582580  | A  | 2019年 12月 17日  |