

(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  
2024年3月7日 (07.03.2024)



(10) 国际公布号  
WO 2024/045170 A1

(51) 国际专利分类号:  
C12Q 1/6886 (2018.01) G01N 33/574 (2006.01)

(21) 国际申请号: PCT/CN2022/116795

(22) 国际申请日: 2022年9月2日 (02.09.2022)

(25) 申请语言: 中文

(26) 公布语言: 中文

(71) 申请人: 深圳华大基因股份有限公司 (BGI GENOMICS CO., LTD.) [CN/CN]; 中国广东省深圳市盐田区洪安三街21号华大综合园7栋7层-14层, Guangdong 518083 (CN)。

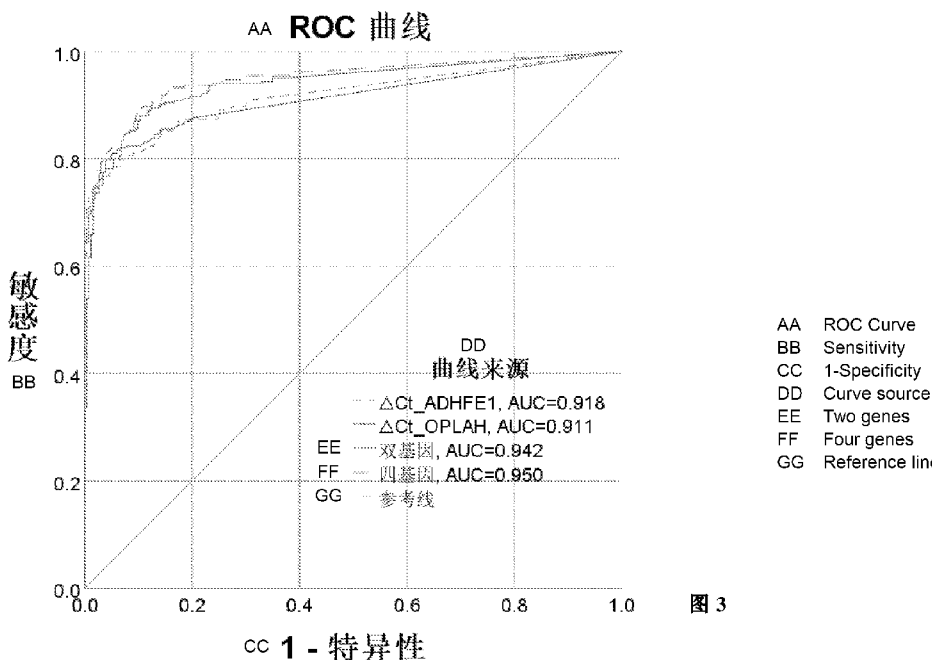
(72) 发明人: 李志隆 (LI, Zhilong); 中国广东省深圳市盐田区洪安三街21号华大综合园7栋7层

-14层, Guangdong 518083 (CN)。汪宇盈 (WANG, Yuying); 中国广东省深圳市盐田区洪安三街21号华大综合园7栋7层-14层, Guangdong 518083 (CN)。彭佳茜 (PENG, Jiaxi); 中国广东省深圳市盐田区洪安三街21号华大综合园7栋7层-14层, Guangdong 518083 (CN)。蒋睿婧芳 (JIANG, Ruijingfang); 中国广东省深圳市盐田区洪安三街21号华大综合园7栋7层-14层, Guangdong 518083 (CN)。孙健洸 (SUN, Jianlong); 中国广东省深圳市盐田区洪安三街21号华大综合园7栋7层-14层, Guangdong 518083 (CN)。

(74) 代理人: 中国贸促会专利商标事务所有限公司 (CCPIT PATENT AND TRADEMARK LAW)

(54) Title: COMBINATION OF DIFFERENTIALLY METHYLATED REGIONS, KIT, AND USE THEREOF

(54) 发明名称: 差异性甲基化区域组合、试剂盒和用途



(57) Abstract: Provided are a combination of differentially methylated regions, a kit, and use thereof. The combination of differentially methylated regions comprises the following differentially methylated regions: (1) chr8: 145106171-145107467 and (2) chr8:67344198-67345563. The combination of differentially methylated regions can be used for effective detection of colorectal cancer or precancerous lesions of colorectal cancer, or for evaluating prognostic risks of colorectal cancer patients, exhibiting high sensitivity and specificity.

(57) 摘要: 提供差异性甲基化区域组合、试剂盒和用途; 所述差异性甲基化区域组合, 其包含如下的差异性甲基化区域: (1) chr8: 145106171-145107467 和 (2) chr8:67344198-67345563; 所述的差异性甲基化区域组合能够用于有效地检测结直肠癌或结直肠癌前病变, 或者用于评估结直肠癌患者的预后风险, 具有良好的灵敏度和特异性。

**OFFICE**: 中国北京市复兴门内大街158号远洋大厦F10层, Beijing 100031 (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, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, 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, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, 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, SC, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

- 包括国际检索报告(条约第21条(3))。
- 包括说明书序列表部分(细则5.2(a))。

## 差异性甲基化区域组合、试剂盒和用途

### 技术领域

5 本发明属于生物医药领域，涉及 OPLAH 基因和 ADHFE1 基因的差异性甲基化区域的组合、试剂盒和用途。具体地，所述试剂盒为检测该差异性甲基化区域的试剂盒。具体地，所述用途为用于检测结直肠癌，或者用于结直肠癌患者的预后风险评估的用途。

### 背景技术

10 发生在下消化道结肠、直肠部位的结直肠癌是一种常见的恶性肿瘤。据最新统计数字显示，全世界范围内结直肠癌的新发病例数在男性中排第三位、在女性中排第二位，死亡病例数在男性中排第四位、女性中排第三位（Bray 等，2018）。在我国，结直肠癌的防控形势也十分严峻，为恶性肿瘤新发病例的第三位（38.8 万人）、死亡病例的第五位（18.7 万人）（郑荣寿等，2019）。结直肠癌发展缓慢，一般会经历息  
15 肉、腺瘤、肠癌等过程，从腺瘤发展成为肠癌时间可长达 5-10 年。而如在结直肠癌发展早期进行干预，可显著降低死亡率。I 期结直肠癌患者的五年生存率可达 90% 以上，而 IV 期结直肠癌患者的五年生存率低于 20%（Marzieh Araghi 等，2020）。

传统的结直肠癌早筛技术主要有肠镜、大便潜血试验、肿瘤标志物 CEA 和 CA19-  
20 9 等，这些技术手段目前均存在一定的局限性。肠镜检查是肠癌诊断的金标准，通过从肛门插入镜头及光源，循腔经直肠、乙状结肠等部位可达回肠末端，图像可实时传输到显示器上，供操作医生观察。肠镜的图像直观、清晰，可观察到癌、息肉、溃疡、出血等多种病变，还可对息肉等进行镜下摘除等操作。但其为一种侵入性的检测技术，准备过程复杂，需控制饮食及进行肠道清洁，待检者的依从性低。且对设备和人员有要求，需要在医院专业人员和麻醉师操作，部分受检者会有不适感并且有产生并发症  
25 的风险（3-5 例/1000 人）。大便潜血试验是一种无创的方法，通过检测大便中的血液成分（血红蛋白）来判别有无消化道出血，以便医生判断肠癌风险。该方法快捷简便，待检者依从性较高。但该检测方法易受到饮食中的肝脏、血液制品、绿叶蔬菜等影响出现假阳性，同时对出血量较少的早期结直肠癌患者的灵敏度低，易出现漏诊。而 CEA 等作为广谱的肿瘤标志物，特异度低，易产生较多的假阳性，灵敏度亦有限。因此，

需要建立一种准确、简便、经济的结直肠癌早期筛查方法。

DNA 甲基化是一种重要的基因表达调控机制，能够调节基因的表达和沉默，在肿瘤的发生发展中具有重大的影响。癌症相关基因的异常甲基化常出现于癌症发生的早期，因此 DNA 甲基化信号被认为是有潜力的肿瘤早期筛查标志物。人的粪便样本中混杂有肠道细胞 DNA，这些 DNA 携带有甲基化等信息，通过检测和分析这些信息，可推断受检者罹患肠癌的可能性。已有部分基因被证明了可以作为粪便中肠癌检测的甲基化标志物，例如 SFRP2，SEPT9，NDRG4 和 SDC2 等（Hannes M Müller 等，2004；Jie Chen 等，2019），但它们的性能尚不能完全满足临床需求，例如 SFRP2 对肠癌的灵敏度和特异性均仅为 77%（Hannes M Müller 等，2004）。

OPLAH 基因编码 5-羟脯氨酸酶，并影响谷氨酸补救路径和异常表达。有研究表明，OPLAH 异常高甲基化引起的表达下调，影响谷胱甘肽的代谢并伴随着肿瘤发生（Liu Y 等，2014）。

ADHFE1 编码一种转氢酶，由该基因的过甲基化引起的异常调控能缩短肠癌细胞周期并促进癌细胞增殖（Tae CH, Ryu KJ, Kim SH, Kim HC, Chun HK, Min BH, et al. Alcohol dehydrogenase, iron containing, 1 promoter hypermethylation associated with colorectal cancer differentiation. BMC Cancer. 2013;13:142.）。

目前，尚需要开发新的用于结直肠癌检测的标志物。

#### 发明内容

本发明人经过深入的研究和创造性的劳动，发现了 OPLAH 基因的差异性甲基化区域（Differentially methylated region, DMR）以及 ADHFE1 基因的差异性甲基化区域。本发明人惊奇地发现，所述差异性甲基化区域的组合能够用于有效地检测结直肠癌或结直肠癌前病变，或者用于评估结直肠癌患者的预后风险，具有良好的灵敏度和特异性。由此提供了下述发明：

本发明的一个方面涉及差异性甲基化区域，其包含如下的差异性甲基化区域：

- (1) chr8: 145106171-145107467 (DMR1)，和
- (2) chr8: 67344198-67345563 (DMR2)。

上面的 (1) 为 OPLAH 基因的差异性甲基化区域。

上面的 (2) 为 ADHFE1 基因的差异性甲基化区域。

上面的 (1) 和 (2) 也可以称为差异性甲基化区域组合。

在本发明的一些实施方式中, 所述的差异性甲基化区域, 其由差异性甲基化区域 (1) 和 (2) 组成。

5 在本发明的一些实施方式中, 所述的差异性甲基化区域, 其中:

第 (1) 项所示的差异性甲基化区域 (DMR1) 的碱基序列如 SEQ ID NO: 1 或其互补序列所示,

第 (2) 项所示的差异性甲基化区域 (DMR2) 的碱基序列如 SEQ ID NO: 2 或其互补序列所示。

10 DMR1 的碱基序列:

CGCCGTAGCTGCGTCCCCAGAACCGGGAGACTTAAGGCATCTTTATTGC  
GGGATCCTCACACGGCCTCCTGGGCCC GGCGATACTCATAGACGCTGCCGT  
GCTCGGGAAAGGCCAGTGCTTGC GGGGGCGACCCCGGCGGTGGGGCGGGG  
TCCTCCGGGTCCCCATAGCCACCGCCGCGGGCGTGTGGAGACAGAACACA  
15 TCCTGTTGGCGCGGGGGGGGGCGGGGAGGCGGGCTCAGTGCAGGCGTGGC  
CCCTCCCTGTACCTGCGCAGTCCCCTCCCCACCGGGCAGCGTCCATCACCC  
GCCCCATGGACCACGGAGGCCCGGTACGCGCCCCCTTTCCTTCCGGGGCC  
AAATTA AACTCGGCCTCTGGTCAGGTCTGCAACTGTTCCCCCTCTGTGGGAT  
GGGGCCGGGAGGGGAGGGGAGGGGAGGGGGACAGGACAGGCCGCGGCCC  
20 TTACCCCGGGGTACACGGTCACCGACGTCTTGCCGCCCAGATTACCGTCC  
GGCCGTTTTTGC GGATCAGCAGGTTTAGGCCGCGGGCGCCAGGCTCGCCCC  
CTGCGGAGGGAGGCGAGGAGTCCAGAGAGACCAGGGGCGGGGGTGGGGGC  
GGGGGCGGGGAGGGCCGCGAGGGCTGGGGAACCGCGCACCGTGGAGCCCG  
TATGGCCGGAAGGCGCGGCGCTCGGTCAGCACTGACAGCAGCGCCTCCTCA  
25 CGAAAGAGCAGCTCGCGGGTGACGCCGTCGCCGCTCGGAAGCGGCCTCTG  
CCCCCGAGCCCCCGCCG CAGCTCGAAGCGGCGCAGGATGACCGGGTACC  
TGCGAGGGCGAGGACGAGGGCAAAGACTCAGCGTGGCCCGAGCCCCGGCT  
CCCACCCGGCCGTGCCCCAGCCTCCGCGCACGCCGCTCACCGGCTCTCCA  
GGATCTCAGGGTCGGTGATGCGTGTGTTGGTCATGTGGCTGTGCACACCGC

TGCGCCCGTGCCAGCTGGGACCCGCGCCCGCGCCCGCCCGCCACCGTCTCGT  
 AGTAGCCCATGTGGGCGTTGCCAGGGTCACGTTGTTTCATGCAGCCCTGGT  
 GAGGGCACCTGGTCGCTGCGCTGGGCTGGGGCAGGGCGTCCCCACCTCGCC  
 CTCCCGCTCCTACTCCAGCCTGCCACCCAGTCCGCCCCCGAGCTGCGCCCA  
 5 CCCCGCCCCCGCACCTGGGAGGCGGGCGCAGGCCCCAAAGGCCCCCAGGATG  
 ACATCCACCACGCGCTGCGACGTGAGCACGTTGCCGCCACCACCGCCGCC  
 TCGGGCGACGGGTCCAGGATGGAGCCTC (SEQ ID NO: 1)

DMR2 的碱基序列:

10 AGTCTAATCATGGCAAGTTTTAACTATTGATGTAAGAGTTTTAACTATTG  
 ATTTTGCTTTCATCAAAAGCAAAGCCATTCACACAGTATTTAACAAAAACCC  
 TCGGGCTCTGCTAGGTTCTTGTTTACTGTAAAACAGTTTAAAGGCAGGGTC  
 TTGATCACTTAAAATACTCTTAGAAGGGAAAGTAACATTTTAAAGGATTAAC  
 AGTGGA AAAAGAATATCCCTGTTCTTCTTCCCCCAAGTCCTCAACATG  
 15 TGAAAAATAGAGTTGATAAGTAAGGAGATTTAAGGCAGAACCCGAGGCCTA  
 CGGAGCAGTTACCTTCTACGGCAATTTCAAGGGTGGATGGTGCAGCGCCG  
 CTGGGGCAGCTGGCGTTCTGGTTCTTACTCCGTGGGAAAATGGCCCTGAGC  
 CCGACTGGCTTGAGGCTTAGACAGGTGACCCCGCGAAGCGGGTGGGCAGGC  
 GCGGCCGAGGGGCGGGAGGCGGGCAGCCTCCGTGATTGGCCGGCGGAGGG  
 20 GGCGGTGTAGCGACCCGAGGAGGGAAGAGGACTCCAAGCGCCATGGCCGC  
 TGCCGCCCGAGCCCGGGTCGCGTACTTGCTGAGGCAACTGCAACGCGCAGC  
 GTGAGTGCGGGGCGGGCGGGCAGGGGACCGCAGGGTGCCACGCAG  
 CCCCCTGGGTGTGACCCGCTCAGATCCTGCGCAGTCTTCTCCGCTTACTTTC  
 AAGAATTCGGAGAATTTGGATAACCGCCCTGGGCTCCACCGCGCGCCAGGC  
 25 GGAACAGCGAGGTCACAGAGTGCGCGATGCAGCCTCTCAGGGCGCAGTGAG  
 GCAGCCTTTACTCTGCTTGTCTACCGTTAAAGAACTGAGGCCCAGAGAGGA  
 CCTGGTCCCACATCGTAGCAGGCAGTGGAGAGCTTGGCACGGTGGCCAAGA  
 CAAATTCATGTATGCAATTGACACATATTTATTGGGCATATATTATGTGCCA  
 AAAACTATTCTAGGTTCTGGGGATACAAAGATGAAAGGCTCTGGAGCAGTTT

GAGGTCTAGTGGGGAATGGGATGAGGGGACGAGGCTAAAACCGGCTATAGC  
 AGACTGGGTGTAGAAGGGAAGGATGGATGTGAGAAGTCGCAGAGTCACACT  
 TCCTGGCATGTACAGGGAGTCAGGGTATTCAATGTCATGACACGCCTGCCTC  
 CCAGCAGAAATGTGAGGTACATCCTGGTCACCTGATAAAATACTTGCTGACTG  
 5 TGCAAGTGCACAAGTCTGCCTGGGAAACATGGAAAAGTTTGAGCTCACAGG  
 TGCAGATCAGTAGGAGGAAAGGCGAGGGACTGGCGAGGAACGTCAGGCAG  
 AGAGAAGAGCAACACCAGCGATGGAGATGCAGTCC (SEQ ID NO: 2)

10 在本发明的一些实施方式中，所述的差异性甲基化区域，其用于检测结直肠癌或结直肠癌前病变，或者用于结直肠癌患者的预后风险评估。

在本发明的一些实施方式中，所述的差异性甲基化区域，其经过了如下处理：

亚硫酸氢盐处理、甲基化敏感性内切酶处理、甲基化限制性内切酶处理或甲基化修饰酶处理。

15 本发明的差异性甲基化区域是用于结直肠癌检测的 OPLAH 基因甲基化标志物或 ADHFE1 基因甲基化标志物。所述差异性甲基化区域组合为用于结直肠癌检测的 OPLAH 基因甲基化的标志物和 ADHFE1 基因甲基化的标志物的组合。

20 本发明的另一方面涉及一种试剂盒，包括检测本发明中任一项所述的差异性甲基化区域的试剂或者所述差异性甲基化区域内的甲基化位点的试剂。

在本发明的一些实施方式中，所述的试剂盒，其中，所述试剂包括：

1) 检测第 (1) 项所示的差异性甲基化区域或者所述差异性甲基化区域内的甲基化位点的引物和探针；

25 2) 检测第 (2) 项所示的差异性甲基化区域或者所述差异性甲基化区域内的甲基化位点的引物和探针；

优选地，所述试剂还包括：

3) 内参基因的引物和探针。

在本发明的一些实施方式中，所述的试剂盒，其中：

1) 中引物的序列选自 SEQ ID NOs: 3-79, 探针的序列选自 SEQ ID NOs: 160-188;

和 2) 中引物的序列选自 SEQ ID NOs: 80-159, 探针的序列选自 SEQ ID NOs: 189-209;

优选地:

1) 中引物的序列选自 SEQ ID NO: 30 和 SEQ ID NO: 32, 探针的序列选自 SEQ ID NO: 170; 和 2) 中引物的序列选自 SEQ ID NOs: 108-109, 探针的序列选自 SEQ ID NO: 197;

1) 中引物的序列选自 SEQ ID NOs: 25-26, 探针的序列选自 SEQ ID NO: 164; 和 2) 中引物的序列选自 SEQ ID NOs: 144-145, 探针的序列选自 SEQ ID NO: 204;

1) 中引物的序列选自 SEQ ID NOs: 26-27, 探针的序列选自 SEQ ID NO: 168; 和 2) 中引物的序列选自 SEQ ID NOs: 154-155, 探针的序列选自 SEQ ID NO: 206;

1) 中引物的序列选自 SEQ ID NOs: 33-34, 探针的序列选自 SEQ ID NO: 171; 和 2) 中引物的序列选自 SEQ ID NOs: 156-157, 探针的序列选自 SEQ ID NO: 209;

1) 中引物的序列选自 SEQ ID NOs: 56-57, 探针的序列选自 SEQ ID NO: 184; 和 2) 中引物的序列选自 SEQ ID NOs: 158-159, 探针的序列选自 SEQ ID NO: 203;

1) 中引物的序列选自 SEQ ID NOs: 61-62, 探针的序列选自 SEQ ID NO: 183; 和 2) 中引物的序列选自 SEQ ID NOs: 108-109, 探针的序列选自 SEQ ID NO: 191;

1) 中引物的序列选自 SEQ ID NOs: 33-34, 探针的序列选自 SEQ ID NO: 171; 和 2) 中引物的序列选自 SEQ ID NOs: 108-109, 探针的序列选自 SEQ ID NO: 191;

或者,

1) 中引物的序列选自 SEQ ID NOs: 25-26, 探针的序列选自 SEQ ID NO: 164; 和 2) 中引物的序列选自 SEQ ID NOs: 158-159, 探针的序列选自 SEQ ID NO: 203;

优选地, 3) 中内参基因的引物的序列如 SEQ ID NOs: 210-211 所示, 内参基因的探针的序列如 SEQ ID NO: 212 所示。

在本发明的一些实施方式中, 所述的试剂盒, 其还包含选自如下的一项或者多项: 粪便样本采样器、核酸提取试剂、甲基化检测样本前处理试剂、PCR 试剂和粪便潜血检测试剂;

优选地, 所述粪便潜血检测试剂为粪便免疫化学检测法检测试剂、愈创木脂检测法检测试剂、四甲基联苯胺法检测试剂或 ELISA 双抗夹心法检测试剂;

优选地, 所述粪便免疫化学检测法检测试剂为免疫胶体金法粪便潜血检测试剂。



在本发明的一些实施方式中，所述的试剂盒，其中，所述甲基化检测样本前处理试剂为亚硫酸氢盐、甲基化敏感性内切酶、甲基化限制性内切酶或甲基化修饰酶。

5 本发明的再一方面涉及本发明的差异性甲基化区域或者检测本发明中任一项所述的差异性甲基化区域或者所述差异性甲基化区域内的甲基化位点的试剂在制备用于检测结直肠癌或结直肠癌前病变，或者用于结直肠癌患者的预后风险评估的药物中的用途。

10 在本发明的一些实施方式中，所述的用途，其中，所述试剂为用于检测本发明中任一项所述的差异性甲基化区域或者所述差异性甲基化区域内的甲基化位点的引物和探针。

本发明的再一方面涉及一种检测结直肠癌或结直肠癌前病变或者一种评估结直肠癌患者的预后风险的方法，包括如下步骤：

15 检测本发明中任一项所述的差异性甲基化区域的含量或者其中的片段的含量；其中，所述片段包含所述差异性甲基化区域内的一个或多个甲基化位点。

本领域技术人员能够理解，检测差异性甲基化区域中片段的含量，例如通过 PCR 方法检测所述的片段的含量，能够代表差异性甲基化区域的含量。

在本发明的一些实施方式中，所述的方法，其中，通过荧光定量 PCR 方法检测本发明中任一项所述的各差异性甲基化区域或者其中的片段的含量。

20 在本发明的一些实施方式中，所述的方法，其中，

采用内参基因 Ct 值评估样本中人源 DNA 的含量，当内参基因 Ct 值 $>37$  时，判定样本人源 DNA 含量过低，本次检测失败，当内参基因 Ct 值 $\leq 37$  时，评估样本合格，可进行后续分析。

在本发明的一些实施方式中，所述的方法，其中，对于评估合格的样本：

25 当各差异性甲基化区域的 Ct 值大于 38 或无检出 Ct 值时，判定为检测阴性；

当各差异性甲基化区域的 Ct 值小于等于 38 时，进一步计算各差异性甲基化区域与内参基因 Ct 值的差值 $\Delta Ct$  值。

在本发明的一些实施方式中，所述的方法，其中：

如果第（1）项或第（2）项的 $\Delta Ct \leq$  截断值，判定为结直肠癌高风险，否则判定为

低风险；

其中 $\Delta Ct = (\text{所述差异性甲基化区域的 Ct 值或者扩增片段的 Ct 值}) - \text{内参基因 Ct 值}$ ；

其中，所述扩增片段包含所述差异性甲基化区域内的一个或多个甲基化位点。

5 在本发明的一些实施方式中，所述的方法，其中：

第（1）项的截断值为 7，和/或

第（2）项的截断值为 12。

在本发明的一些实施方式中，所述的方法，其中还包括联合粪便潜血检测结果进行评估的步骤。

10 在本发明的一些实施方式中，所述的方法，其中，所述样本为人体组织样本、血液样本、细胞样本、分泌物样本或排泄物样本例如粪便样本。

根据本发明的一些实施例，采用粪便样本作为检测对象。这样，粪便样本含有肠道脱落细胞，符合检测样本的需求，且具有无创性，对患者更友好。

15 根据本发明的一些实施例，可选择的检测手段包括但不限于：Sanger 测序、焦磷酸测序、高通量测序、核酸质谱、PCR、荧光定量 PCR、ddPCR 等。由于核酸序列未发生变化，甲基化修饰难以被直接检测，实验过程中往往需要对核酸进行预处理，此处预处理手段包括但不限于：亚硫酸氢盐处理、甲基化敏感性内切酶处理、甲基化限制性内切酶处理或甲基化修饰酶处理等。同时，第三代测序技术已被证实可用于直接测定 DNA 的甲基化修饰状态，随着技术发展，未来亦可作为检测手段。荧光定量 PCR  
20 技术是目前分子诊断常用且发展成熟的检测技术之一，具有快速、准确和成本低的特点。本技术方案采用了亚硫酸氢盐处理结合荧光定量 PCR 的方法，并对 OPLAH 基因和 ADHFE1 基因设计了若干甲基化特异性 PCR 引物和探针，能够特异性地识别甲基化的 DNA 分子，实现了对样本甲基化状态的检测，进而评估受检者的患癌风险。

25 通过对结直肠癌组织和对照组织样本进行全基因组甲基化测序和靶向甲基化测序，得到与结直肠癌高度相关的甲基化差异区域，其中 DMR1（chr8: 145106171-145107467）和 DMR2（chr8:67344198-67345563）的甲基化率差异尤为显著。针对所述区域进行引物、探针设计，可用于快捷、准确地检测 OPLAH 基因和 ADHFE1 基因的甲基化情况。本发明提供了若干引物和探针序列，如表 1 所示：

表 1：OPLAH 基因和 ADHFE1 基因的引物、探针序列

序列号	序列 (5'-3')	说明
3	GAATTAGGCGTCGTAGTTGC	OPLAH 基因引物
4	CGCAAACACTAACCTTTCCC	
5	GGGTTCGGCGATATTTATAGAC	OPLAH 基因引物
6	CCCCACCGCCGAAATC	
7	AGTACGGTAGCGTTTATGAGTATC	OPLAH 基因引物
8	AACTAAACGCCGTAACACTACG	
9	TTATCGTCGGGGTCGTTTTTC	OPLAH 基因引物
10	ACATCTTTATTACGAAATCCTCACAC	
11	TCGTCGGGTTTAGGAGGTC	OPLAH 基因引物
12	ACTCGAAACACGAACTAAACG	
13	TATTTTTATTGCGGGATTTTTATAC	OPLAH 基因引物
14	CAAATATATTCTATCTCCACACGC	
15	CGTAGTTGCGTTTTTAGAATCG	OPLAH 基因引物
16	TATAAAAACCCGAAAAACCCCG	
17	GGGTAGTGCGTTTGTAGTTTC	OPLAH 基因引物
18	CGACAACGTCTATAAATATCGC	
19	TTTTGGAGGGGTTTACGGATTATC	OPLAH 基因引物
20	CCAAAAACCGTATAAAAATCCCG	
21	AAGGTAAGTATAGATGAGGGGCG	OPLAH 基因引物
22	CGCAAACACTAACCTTTCCCG	
23	GAGGTCGTGTGAGGATTTTCG	OPLAH 基因引物
24	AACAATACGCCTACAACCTCCG	
25	GGGTAGTGCGTTTGTAGTTTCG	OPLAH 基因引物
26	ACACGACAACGTCTATAAATATCGC	
27	TCGGGTAGTGCGTTTGTAGTTTC	OPLAH 基因引物
26	ACACGACAACGTCTATAAATATCGC	
28	GGTTTACGGATTATCGGGTAGTGC	OPLAH 基因引物
29	CGACAACGTCTATAAATATCGCCG	

30	GGGCGTTGTTTAGGGTTAC	OPLAH 基因引物
31	AAACGCAACTCGAAAACG	
30	GGGCGTTGTTTAGGGTTAC	OPLAH 基因引物
32	CCCTACCCCAACCCAAC	
33	CGTTCGTGTTAGTTGGGATTC	OPLAH 基因引物
34	CATAACAACGTAACCCTAAACAAC	
35	GGGACGTTTTGTTTTAGTTTAGC	OPLAH 基因引物
36	CCCGTACCAACTAAAACCC	
37	TATTTGCGAGGGCGAGGAC	OPLAH 基因引物
38	CCTAAAAAACCGATAAACGACG	
39	CGTTCGTGTTAGTTGGGATGC	OPLAH 基因引物
34	CATAACAACGTAACCCTAAACAAC	
33	CGTTCGTGTTAGTTGGGATTC	OPLAH 基因引物
40	TAAACAACGTAACCCTAAACACCG	
41	GGGCGTTGTTTAGGGTTGC	OPLAH 基因引物
32	CCCTACCCCAACCCAAC	
30	GGGCGTTGTTTAGGGTTAC	OPLAH 基因引物
42	CCTACCCCAACCCACCG	
43	CGTTCGTGTTAGTTGGGCTTC	OPLAH 基因引物
34	CATAACAACGTAACCCTAAACAAC	
33	CGTTCGTGTTAGTTGGGATTC	OPLAH 基因引物
44	TAAACAACGTAACCCTAAACCACG	
45	CGTTCGTGTTAGTTGGGATCC	OPLAH 基因引物
34	CATAACAACGTAACCCTAAACAAC	
46	CGTTCGTGTTAGTTGGGGTTC	OPLAH 基因引物
34	CATAACAACGTAACCCTAAACAAC	
33	CGTTCGTGTTAGTTGGGATTC	OPLAH 基因引物
47	TAAACAACGTAACCCTAAACGACG	
33	CGTTCGTGTTAGTTGGGATTC	

48	TAAACAACGTAACCCTAAACAACG	OPLAH 基因引物
30	GGGCGTTGTTTAGGGTTAC	OPLAH 基因引物
49	CCTACCCCAACCCAACG	
50	AAGCGGCGTAGGATGATC	OPLAH 基因引物
51	CCGAAACTCGAACCACG	
50	AAGCGGCGTAGGATGATC	OPLAH 基因引物
52	AAACCGAAACTCGAACCAC	
33	CGTTCGTGTTAGTTGGGATTC	OPLAH 基因引物
53	TTTAAAACCTACGCCGCCT	
33	CGTTCGTGTTAGTTGGGATTC	OPLAH 基因引物
54	ACGAAATAAAAACGCCCTACC	
33	CGTTCGTGTTAGTTGGGATTC	OPLAH 基因引物
31	AAACGCAACTCGAAAACG	
33	CGTTCGTGTTAGTTGGGATTC	OPLAH 基因引物
55	ACTCGAAAACGAACTAAATAAACAAACT	
56	TGAATAACGTGATTTTGGGTAACG	OPLAH 基因引物
57	CGACTCTCCAAAATCTCAAAATCG	
58	GGGTAAAGATTTAGCGTGGTTC	OPLAH 基因引物
48	TAAACAACGTAACCCTAAACAACG	
33	CGTTCGTGTTAGTTGGGATTC	OPLAH 基因引物
32	CCCTACCCCAACCCAAC	
33	CGTTCGTGTTAGTTGGGATTC	OPLAH 基因引物
59	ACGAAATAAACGCAACTCG	
33	CGTTCGTGTTAGTTGGGATTC	OPLAH 基因引物
60	CGCAACTCGAAAACGAACT	
61	AGGATGATCGGGTATTTGCG	OPLAH 基因引物
62	CAACCACATAACCAACACACG	
61	AGGATGATCGGGTATTTGCG	

63	GCGAATCCCAACTAACACG	OPLAH 基因引物
64	GACGAGGGTAAAGATTTAGCG	OPLAH 基因引物
63	GCGAATCCCAACTAACACG	
65	GGTAAAGATTTAGCGTGGTTCG	OPLAH 基因引物
48	TAAACAACGTAACCCTAAACAACG	
66	TCGGTTTTTTAGGATTTTAGGGTCG	OPLAH 基因引物
67	CCTCACCAAACTACATAACAACG	
68	TTGTATGAATAACGTGATTTTGGGTAACG	OPLAH 基因引物
69	CACCGACTCTCCAAAATCTCAAAATCG	
70	TTATCGATTTTGAGATTTTGGAGAGTCG	OPLAH 基因引物
71	CTCGATCAACACTAACAACAACG	
72	GCGTAGGATGATCGGGTATTTGC	OPLAH 基因引物
73	CGAATCCCAACTAACACGAACG	
74	TTTGCGGAGGGAGGCG	OPLAH 基因引物
75	GCGATTCCCAACCCTCG	
76	GGAGTTTAGAGAGATTAGGGGCG	OPLAH 基因引物
77	CGACCATACGAACTCCACG	
78	TCGTGGAGTTCGTATGGTGC	OPLAH 基因引物
79	CCCGCGAACTACTCTTTCG	
80	TCGTTGGGGTAGTTGGC	ADHFE1 基因引物
81	CCTATCTAAACCTCAAACCAATCG	
82	GTTTTGGGTTTTTATCGCGC	ADHFE1 基因引物
83	GCCCTAAAAAACTACATCGCG	
84	AGGTTTAGATAGGTGATTTTCGC	ADHFE1 基因引物
85	AAACTACCCGCCTCCC	
86	TTGGCGTTTTTGGTTTTTATTC	ADHFE1 基因引物
81	CCTATCTAAACCTCAAACCAATCG	
87	GATTTAAGGTAGAATTCGAGGTTTAC	

88	CACGAAATAAAAACCAAACG	ADHFE1 基因 引物
89	CGTAGTTTTTTGGGTGTGATTC	ADHFE1 基因 引物
90	AAAACGATATCCAAATTCTCCG	
91	CGTATTTTGTGATTTTCGTTGTTTC	ADHFE1 基因 引物
92	AAAAATTCGAAAAATTTAAATACCG	
93	CGTTTGTTTATTCGTTTCGC	ADHFE1 基因 引物
94	CGCCGCTAAAACAACCTAACG	
95	CGTTTGTGAGAGGTTGTATCGC	ADHFE1 基因 引物
92	AAAAATTCGAAAAATTTAAATACCG	
96	GTTTAGGGCGGTATTTAAATTTTTC	ADHFE1 基因 引物
97	ATATAACCCGCTCAAATCCTACG	
98	TGTATCGCGTATTTTGTGATTC	ADHFE1 基因 引物
99	CCGCTTACTTTCAAAAATTCG	
100	GGTAATTTTAAGGGTGGATGGTG	ADHFE1 基因 引物
101	TCACCTATCTAAACCTCAAACCA	
102	GGGGTTATTTGTTTAAGTTTAAAGTTAGT	ADHFE1 基因 引物
103	AAACAATTACCTTCTACRACAATTCA	
104	AAGTAAGGAGATTTAAGGTAGAATT	ADHFE1 基因 引物
105	TTAAAATCCTCTTCCCTCCTC	
106	ATATGTGAAAAATAGAGTTGATAAGTAAGG	ADHFE1 基因 引物
107	CTTAAAATCCTCTTCCCTCCTC	
108	GATGGTGCGAGCGTC	ADHFE1 基因 引物
109	CTATCTAAACCTCAAACCAATCG	
108	GATGGTGCGAGCGTC	ADHFE1 基因 引物
110	TACACCGCCCCCTCC	
108	GATGGTGCGAGCGTC	ADHFE1 基因 引物
111	TCCGCCGACCAATCAC	
108	GATGGTGCGAGCGTC	

112	CCGAACTCGAACGACAAC	ADHFE1 基因 引物
113	AYGGTAATTTTAAGGGTGGATGGTGYGAGYGT	ADHFE1 基因 引物
114	CRCCCCCTCCRCCRACCAATCACRAAAACT	
115	AAGTAAGGAGATTTAAGGTAGAATTYGAGGTTTAYG GAGTAGT	ADHFE1 基因 引物
116	ACCATAACRCTTAAAATCCTCTTCCCTCCTCRAATCR CT	
117	GGAGATTTAAGGTAGAATTCGAGGTTTAC	ADHFE1 基因 引物
118	CTCGAACGACAACGACCATAAC	
119	GAATTCGAGGTTTACGGAGTAGT	ADHFE1 基因 引物
120	CTCCTCGAATCGCTACACC	
121	TTTAAGGGTGGATGGTGCGAGC	ADHFE1 基因 引物
122	CTCAACAAATACGCGACCC	
123	CGGTTAATTACGGAGGTTGTTC	ADHFE1 基因 引物
124	CCTACGAAACAATTACCTTCTACGA	
123	CGGTTAATTACGGAGGTTGTTC	ADHFE1 基因 引物
125	AAAACAAAACCCGAAACCTAC	
126	TAGTAAGTACGCGATTCGGGTTC	ADHFE1 基因 引物
127	GAAACCTACGAAACAATTACCTTCTAC	
128	CGGAGTAGTTATTTTTTACGG	ADHFE1 基因 引物
129	AACGACAACGACCATAAC	
128	CGGAGTAGTTATTTTTTACGG	ADHFE1 基因 引物
130	GTTACAATTACCTCAACAAATACG	
131	GGTAGAATTCGAGGTTTACG	ADHFE1 基因 引物
132	TCTTCCCTCCTCGAATCG	
131	GGTAGAATTCGAGGTTTACG	ADHFE1 基因 引物
133	CAATCACGAAAACCTACCCG	
108	GATGGTGCGAGCGTC	



134	GAACTCGAACGACAACG	ADHFE1 基因 引物
108	GATGGTGCGAGCGTC	ADHFE1 基因 引物
135	AACTCGAACGACAACGAC	
108	GATGGTGCGAGCGTC	ADHFE1 基因 引物
136	CGTTACAATTACCTCAACAAATACG	
108	GATGGTGCGAGCGTC	ADHFE1 基因 引物
137	TTAAAATCCTCTTCCCTCCTCG	
108	GATGGTGCGAGCGTC	ADHFE1 基因 引物
138	ATCCTCTTCCCTCCTCG	
139	AGGTAGAATTCGAGGTTTACG	ADHFE1 基因 引物
140	TACCCACCCGCTTCG	
141	TGGGAAAATGGTTTTGAGTTCG	ADHFE1 基因 引物
137	TTAAAATCCTCTTCCCTCCTCG	
142	GAGGGAAGAGGATTTTAAGCG	ADHFE1 基因 引物
143	CCAAAACGATATCCAAATTCTCCG	
144	GGATCGTAGGGTGTTTACG	ADHFE1 基因 引物
145	CTACGCCCTAAAAAACTACATCG	
146	GAATTCGGAGAATTTGGATATCG	ADHFE1 基因 引物
147	AACTCTCCACTACCTACTACG	
148	GCGTATTTGTTGAGGTAATTGTAACG	ADHFE1 基因 引物
149	ACTCTATAACCTCGCTATTCCG	
150	AGGAGGGAAGAGGATTTTAAGCG	ADHFE1 基因 引物
151	CCAAAACGATATCCAAATTCTCCG	
152	GGGATCGTAGGGTGTTTACG	ADHFE1 基因 引物
153	CTCACTACGCCCTAAAAAACTACATCG	
154	TCGGTTAATTACGGAGGTTGTTTCG	ADHFE1 基因 引物
155	AACCTACGAAACAATTACCTTCTACG	
156	GAGTTTAGGGCGGTATTTAAATTTTTCG	

157	AAACTCCAAACGCCATAACCG	ADHFE1 基因 引物
158	TTAAGCGTTATGGTCGTTGTCG	ADHFE1 基因 引物
159	GCGCGATAAAAACCCAAAACG	
160	ACACGACAACGTCTATAAATATCGCCGAACC	OPLAH 基因探 针
161	CCCGACGATACTCATAAACGCTACCGTA	OPLAH 基因探 针
162	CCGTAACTACGTCCCCAAAACCGAAA	OPLAH 基因探 针
163	TCGGGTTTAGGAGGTCGTGTGAGG	OPLAH 基因探 针
164	AACCGTATAAAAATCCCGCAATAA	OPLAH 基因探 针
165	AAACGCAACTACGACGCCTAATTCGTAC	OPLAH 基因探 针
166	TTGGCGCGTATAGATACGATTCGGAGT	OPLAH 基因探 针
167	TAGTGCGTTTGTAGTTTCGAGT	OPLAH 基因探 针
168	CTCCCGATTCTAAAAACGCAA	OPLAH 基因探 针
169	ACGTAGTTACGGCGTTT	OPLAH 基因探 针
170	CAACGACCAAATACCCTCA	OPLAH 基因探 针
171	TTCGCGTCGTTTCGTTATCGTTT	OPLAH 基因探 针

172	AGCGTGGTTCGAGTTTCGGTTT	OPLAH 基因探 针
173	ATTTGGTCGTTGCGTTGGGTTGG	OPLAH 基因探 针
174	CCCGCCACCGTCTCGTAAT	OPLAH 基因探 针
175	TTGCGAGGGCGAGGACGAGG	OPLAH 基因探 针
176	CGAGGACGAGGGTAAAGATTTAGCGT	OPLAH 基因探 针
177	AACGCAACGATATACACAACCACATAACCAAC	OPLAH 基因探 针
178	TTCGAGTTTCGGTTTTTATTCGGTTCGTGT	OPLAH 基因探 针
179	TTCGGTTTTTATTCGGTTCGTG	OPLAH 基因探 针
180	CACCGACCCTAAAATCCTAAAAAACCGAT	OPLAH 基因探 针
181	ACCGATAAACGACGTACGCGAAAAC	OPLAH 基因探 针
182	CTACGAAACGATAACGAACGACGCGAA	OPLAH 基因探 针
183	ACGAGGGTAAAGATTTAGCGT	OPLAH 基因探 针
184	CACACCGCTACGCCGTACCAA	OPLAH 基因探 针
185	TGAGCGGCGTGCGCGGAG	OPLAH 基因探 针

186	AGTTTAGAGAGATTAGGGGCG	OPLAH 基因探 针
187	TACGCGATTCCCCAACCTCGCGAC	OPLAH 基因探 针
188	AAGGCGCGGCGTTCGGTTAG	OPLAH 基因探 针
189	TCAAAACCATTTTCCCACGAAATAAA	ADHFE1 基因 探针
190	CTCTATAACCTCGCTATTCCGCCT	ADHFE1 基因 探针
191	CCCACGAAATAAAAACCAAACGCCAAC	ADHFE1 基因 探针
192	ACTCAAACCATTTTCCCACGAA	ADHFE1 基因 探针
193	ACTCAAACCATTTTCCCACGAAAT	ADHFE1 基因 探针
194	AACGAAAACGCCACTACCCGATACGA	ADHFE1 基因 探针
195	AACGCCACTACCCGATACGAACTACACC	ADHFE1 基因 探针
196	AAAAACCAAACGCCAACTACCCCA	ADHFE1 基因 探针
197	CAAACGCCAACTACCCCAACGAC	ADHFE1 基因 探针
198	TGCGAGCGTCGTTGGGGTA	ADHFE1 基因 探针
199	TGGTCGTTGTCGTTTCGAGTTCGG	ADHFE1 基因 探针

200	CCTCGCTATTCCGCCTAACGCG	ADHFE1 基因 探针
201	ACTACGCAAATCTAAACGAATCA	ADHFE1 基因 探针
202	TATGGTCGTTGTCGTTTCGAGTT	ADHFE1 基因 探针
203	TACGTAAACACCCTACGATCCC	ADHFE1 基因 探针
204	TTCGTTTAGATTTTGCGTAG	ADHFE1 基因 探针
205	AACGATATCCAAATTCTCCGAAT	ADHFE1 基因 探针
206	TTCGTTTTTCGGTCGCGTTTGTT	ADHFE1 基因 探针
207	ATACGAACGCCGCTAAAACAACGTTCT	ADHFE1 基因 探针
208	TACCGCCCGAACCCGAATCGCGTA	ADHFE1 基因 探针
209	AAGCGGAGAAGATTGCGTA	ADHFE1 基因 探针
3	GAATTAGGCGTCGTAGTTGC	OPLAH 基因引 物
4	CGCAAACACTAACCTTTCCC	
5	GGGTTCGGCGATATTTATAGAC	OPLAH 基因引 物

在本文中，本发明人使用采样装置收集粪便样本，样本经过核酸提取和纯化步骤，获得包含人源 DNA 的核酸样本。进一步地，本发明人采用亚硫酸氢盐对所述核酸样本进行处理，使未发生甲基化的胞嘧啶转化为尿嘧啶，使甲基化位点能够被检测到。

5 针对所述基因，本发明人设计了如表 1 所述的若干引物和探针，能够特异性地扩增、

识别相关位点。所述引物、探针序列为连续的至少 15 个碱基长度的片段，而所述引物、探针的核酸序列等同、互补或杂交于经甲基化修饰预处理的所述甲基化区域。使用荧光定量 PCR 方法，本文提供的引物和探针能够有效地通过 Ct 值和/或 $\Delta$ Ct 值（目标基因 Ct 值与内参基因 Ct 值的差值）和/或曲线情况反映相关位点的甲基化情况，进而判断受检者相关基因的甲基化情况。进一步地，通过对若干癌症患者和非癌症患者样本的检测，可得到判定患癌风险判定的截断值，通过对比检测结果与截断值，可对受检者患癌风险进行评估。

所述优化后的引物、探针和 PCR 反应体系，可形成一种 OPLAH 基因和 ADHFE1 基因甲基化检测试剂盒，进一步的该试剂盒可用于癌症的风险评估。

10

关于粪便潜血检测：

对粪便样本中的血红蛋白进行检测，检测手段可包括粪便免疫化学检测法（FIT）、愈创木脂检测法（gFOBT）、四甲基联苯胺法和 ELISA 双抗夹心法等。

所述粪便潜血的检测方法中，FIT 利用单克隆或多克隆抗体直接检测人粪便中的血红蛋白，不受进食食物的影响。定性 FIT 是在粪便中血红蛋白含量超过一定阈值后会产生可视性的颜色变化，定量 FIT 则可测量数值，当超过一定的正常值范围后被定义为阳性。定性 FIT 检测可采用免疫胶体金法，具有检测方便、快捷的优点；定量 FIT 检测可采用乳胶免疫比浊法，具有较高灵敏度和特异性。

20

关于联合基因检测和粪便潜血检测的方法：

基因-粪便潜血联合检测的结果判定方法可以是基因检测结果和粪便潜血检测可分别定制判定阈值，综合以上两部分检测结果，受检者被判定为阳性或候选为阳性。

判定方法也可以是采用逻辑回归的方法，对 OPLAH 基因、ADHFE1 基因和粪便潜血的检测结果赋予不同的权重，通过逻辑函数计算。所述逻辑函数为  $g(y)=1/(1+e^{-y})$ 。其中  $g(y)$  是风险指数，当风险指数高于预设值时，判定为阳性或候选为阳性； $e$  为自然常数； $y=(\theta_0+\theta_1*X_1+\theta_2*X_2+\dots+\theta_n*X_n)$ ， $\theta_0$  为常数项， $\theta_i(i=1, 2, \dots, n)$  是待求系数， $X_i(i=1, 2, \dots, n)$  是基因和粪便潜血检测的结果。不同的权重  $\theta_i$  反映了自变量对因变量不同的贡献程度，不同的检测结果  $X_i$  可以包括但不限于基因检测的 Ct 值、 $\Delta$ Ct 值、甲基化率、粪便潜血定性检测结果和粪便潜血定量检测结果中的一个或一个以上的任意组合。

通过优化判定方法中的参数，可得到更高的检测灵敏度。

试剂盒包含基于荧光定量 PCR 法的 OPLAH 和 ADHFE1 基因甲基化联合检测试剂，可提供 OPLAH 和 ADHFE1 基因的甲基化检测结果。进一步地，试剂盒还可包含粪便潜血检测试剂，可提供粪便潜血的检测结果。通过计算 $\Delta Ct$  值或风险指数，与试剂盒预设的阳性判断值对比，可用于判断患结直肠癌的风险。

所述基于荧光定量 PCR 法的 OPLAH 和 ADHFE1 基因甲基化联合检测试剂中包含能够特异性扩增 OPLAH 和 ADHFE1 基因的引物和探针。所述试剂能够检测经甲基化修饰处理后的核酸中 OPLAH 和 ADHFE1 基因的甲基化情况。

所述基因甲基化联合检测试剂可搭配不同的粪便样本采样装置、核酸提取试剂和甲基化修饰预处理试剂使用。所述粪便采样装置可包含采样器和样本存储器，用于采集和存储样本。所述核酸提取试剂可以是能够提取人体生物样本中核酸的试剂盒，也可以是含有核酸捕获磁珠的试剂盒。所述甲基化修饰预处理试剂通过对核酸序列进行修饰，使发生甲基化修饰的胞嘧啶与未发生甲基化修饰的胞嘧啶产生便于检测的差异。通常的，所述甲基化修饰预处理试剂能够使未发生甲基化修饰的胞嘧啶转化为尿嘧啶，而发生甲基化修饰的胞嘧啶保持不变，序列的变化可被用于后续的检测。

所述粪便潜血检测试剂优选地，使用免疫胶体金法，可简便、快捷地获得粪便潜血定性检测结果。免疫胶体金法粪便潜血检测试剂可包含采样器、样本存储器和检测试纸。所述检测试纸包被有抗体和胶体金，当被胶体金标记的目标蛋白与试纸上预包被的抗体结合时会滞留在试纸上形成可视变化，从而判断样本中是否含有目标蛋白。

所述阳性判断值使用已知临床信息的样本划定，通过对若干阳性和阴性样本进行检测，构建分类模型，确定判定为阳性的条件。

本发明的再一方面涉及一种结直肠癌的诊断或评估模型，采用逻辑函数 $g(y)=1/(1+e^{-y})$ 进行计算，其中：

$g(y)$  是风险指数，当风险指数高于预设值时，判定为阳性或候选为阳性；

$e$  为自然常数；

$y=(\theta_0+\theta_1*X_1+\theta_2*X_2+\dots+\theta_n*X_n)$ ， $\theta_0$  为常数项， $\theta_i(i=1, 2, \dots, n)$  是待求系数， $X_i(i=1, 2, \dots, n)$  是本发明中任一项所述的差异性甲基化区域的检测结果和粪便潜血的检测结果。

本发明的再一方面涉及一种结直肠癌的诊断或评估系统，包括：

1) DMR检测单元：获取本发明中任一项所述的差异性甲基化区域或者其中的片段的含量数据；其中，所述片段包含所述差异性甲基化区域内的一个或多个甲基化位点；

5 2) 粪便潜血检测单元：获取粪便潜血检测结果数据；

3) 数据分析单元：将DMR检测单元和粪便潜血检测单元获得的数据输入本发明的结直肠癌的诊断或评估模型进行分析；

4) 风险评估单元：输出个体样本的风险指数或者判断结果。

10 本发明的再一方面涉及本发明的差异性甲基化区域(1)和差异性甲基化区域(2)在制备检测结直肠癌或者结直肠癌癌前病变的产品中的用途。

本发明中，如果没有特别说明，术语“结直肠癌”是指结肠癌和/或直肠癌，也简称为“肠癌”。

15 本发明中，如果没有特别说明，术语“早期结直肠癌”是指 I 期或 II 期结直肠癌。

本发明中，如果没有特别说明，检测差异性甲基化区域或者检测其中的片段，是指检测差异性甲基化区域的含量或者检测其中的片段的含量；其中，所述片段包含所述差异性甲基化区域内的一个或多个甲基化位点。

20 本发明中，如果没有特别说明，检测差异性甲基化区域内的甲基化位点，是指检测差异性甲基化区域内的甲基化位点的含量或者检测差异性甲基化区域内的甲基化位点的甲基化率。

### 发明的有益效果

本发明取得了如下的技术效果中的一项或者多项：

25 (1) 能够实现 OPLAH 基因和 ADHFE1 基因甲基化状态的快速检测；

(2) 能够检测、诊断或筛查结直肠癌，或者用于结直肠癌的预后风险评估，具有较高的灵敏度和/或特异性；优选地，同时兼具较高的灵敏度和特异性。

(3) 能够应用于早期结直肠癌或结直肠癌癌前病变的检测、诊断或筛查，具有较高的灵敏度和/或特异性；优选地，同时兼具较高的灵敏度和特异性。



(4) 可联合检测粪便中的血红蛋白，进一步提高结直肠癌的检测灵敏度。

(5) 与传统检测手段相比，本发明具有更高的敏感性，完全无创，检测过程非侵入性，无需肠道准备、无症状人群对该检测接受度高等优点。

5 附图说明

图 1: OPLAH 和 ADHFE1 基因在肠癌和癌旁组织样本中的差异甲基化区域。

图 2: OPLAH 和 ADHFE1 基因检测参考品结果。

图 3: OPLAH 和 ADHFE1 基因检测临床样本的 ROC 曲线。

图 4: 基因检测与基因+FIT 联检 ROC 曲线对比。

10 图 5: 检测灵敏度对比。

具体实施方式

下面将结合实施例对本发明的实施方案进行详细描述，但是本领域技术人员将会理解，下列实施例仅用于说明本发明，而不应视为限定本发明的范围。实施例中未注明具体条件者，按照常规条件或制造商建议的条件进行。所用试剂或仪器未注明生产厂商者，均为可以通过市购获得的常规产品。

部分相关定义参考中国临床肿瘤学会（CSCO）结直肠癌诊疗指南 2021，部分内容引用如下面的表 2：

20

表 2

结直肠癌	活检病理诊断为结肠癌或直肠癌
进展期腺瘤	直径 $\geq$ 1cm，或伴绒毛状结构，或伴高级别瘤变，伴重度异型增生、原位癌和黏膜内癌。
腺瘤患者	病理诊断为腺瘤的患者。
息肉患者	病理诊断为炎性息肉或增生性息肉，或未行病理的患者。
其他肠道疾病患者	包括肠炎、憩室、结肠黑变病等非息肉类病变的患者。

肠癌分期的标准采用 AJCC。

**实施例 1: 实时荧光定量 PCR 联合检测 OPLAH 和 ADHFE1 基因的甲基化水平**

发明人通过靶向甲基化捕获测序发现了 OPLAH 和 ADHFE1 基因的差异甲基化区域。所述差异甲基化区域在癌组织中具有高甲基化水平，而在癌旁组织中的甲基化水平较低。本实施例以荧光定量 PCR 法为例，针对所述区域，发明人设计了若干甲基化特异性引物和探针，能够特异性地扩增发生甲基化的核酸序列。经过体系引物筛选和体系优化，开发得到一种基于荧光定量 PCR 的 OPLAH 和 ADHFE1 基因联合检测方法。

具体步骤如下：

**1、靶向甲基化捕获测序发现差异甲基化区域：**

对来自于 38 人的 38 对结直肠癌样本和癌旁组织样本（其中就一对结直肠癌样本和癌旁组织样本而言，来自同一人），使用 DNeasy Blood & Tissue Kit(Qiagen, #69506) 进行核酸提取，用 Qubit 3.0 系统 (Invitrogen, 美国) 进行定量。

取 200ng 提取好的 DNA, 同时加入 1ng Unmethylated lambda DNA(PROMEGA, #D1521) 用于后续 C-U 转化质控，使用超声打断仪将 DNA 打断，再使用 AMPure XP (AGENCOURT, #A63882) 对 DNA 进行片段选择，使 DNA 片段大小集中在 160bp 左右。

使用 KAPA HyperPlus Library Preparation Kit (KAPA, #KK8510) 进行文库构建，使用 EZ DNA Metlylation-Gold Kit (Zymo Research, #D5006) 对接头连接产物进行亚硫酸氢盐转化，具体操作参考试剂盒说明书，其中接头连接步骤所使用的接头以及 PCR 步骤所使用的 PCR 引物更换为适用于 MGISEQ 平台的接头和引物，接头需要进行甲基化修饰（例如使用甲基化的 dCTP 合成接头），以避免亚硫酸氢盐处理改变序列。

使用 Seq Cap EZ Hybridization and Wash Kit (ROCHE, 5634253001) 和 SeqCap Epi CpGiant Enrichment Kit (ROCHE, 7138911001) 进行杂交、捕获和洗脱。由于使用了 MGI 平台的测序仪器，杂交过程中对接头序列的封闭须使用 MGI 测序平台对应的封闭序列。

使用 MGISEQ-2000 (MGI) 进行 PE100 测序。

对测序数据进行质控、过滤不合格测序序列、将测序序列与人类基因组序列的经计算机模拟的亚硫酸氢盐转化型式比对之后，统计所有序列并计算出每个 CpG 位点

的甲基化率 (Non-invasive lung cancer diagnosis and prognosis based on multi-analyte liquid biopsy, Chen Kezhong 等, 2020)。

特定位点甲基化率=甲基化 reads 测序深度/总测序深度。

5 本发明人发现 OPLAH 基因和 ADHFE1 基因存在显著的甲基化差异区域, 具有区分癌与非癌的潜力。图 1 展示了 OPLAH 和 ADHFE1 基因中各个 CpG 位点在癌组织和癌旁组织中的甲基化水平。

## 2、甲基化引物、探针设计

10 根据发现的差异甲基化区域, 本发明人针对在癌组织中高甲基化率且在癌旁组织中具有较低甲基化率的 CpG 位点, 设计了能够特异性扩增经甲基化修饰预处理后核酸序列的引物和探针, 其中一对引物中至少包含一条引物的 3'端具有至少一个 CpG 位点。所设计的引物、探针序列如表 1 所示。

## 3、荧光定量 PCR 能够联合检测 OPLAH 和 ADHFE1 基因的 DMR 的甲基化水平

15 本发明人选择经优化后的引物组和探针进行 OPLAH 和 ADHFE1 基因甲基化检测。本实施例中所述引物组包括: OPLAH 基因引物组 (SEQ ID NO: 30 和 SEQ ID NO: 32)、OPLAH 基因探针 (SEQ ID NO: 170)、ADHFE1 基因引物组 (SEQ ID NO: 108 和 SEQ ID NO: 109) 和 ADHFE1 基因探针 (SEQ ID NO: 197)。荧光定量 PCR 反应体系包括 1× PCR Buffer、2-5mM 镁离子、0.2-0.8mM dNTP、每个基因 0.2μM 引物、每个基因 50nM 探针和 1 单位 Taq 聚合酶。反应在 95℃ 预变性 10min, 然后进行 95℃ 变性 30 秒、55℃ 退火 30 秒和 72℃ 延伸 30 秒的循环 40 个。采用宏石 96S  
20 qPCR 仪进行检测, 基线、阈值按照默认设置, 在每个循环结束时收集荧光信号。

本发明人使用全甲基化标准品和全非甲基化标准品 (Merck, #S8001) 验证所述方法对 OPLAH 和 ADHFE1 基因甲基化水平的检测能力。全甲基化标准品是经过 M. SssI 甲基转移酶处理的细胞系 DNA, 所有 CpG 位点均被甲基化修饰; 全非甲基化标准品是甲基转移酶基因被敲除的细胞系 DNA, 其 CpG 位点的甲基化水平低于 5%。  
25 本发明人使用这两种标准品, 混合成不同甲基化水平的参考品, 包括 100%、50%、25%、12%、6%、3%、1% 和 0%。

对每个参考品取 20ng 进行 3 次检测, 首先使用甲基化检测样本前处理试剂盒 (BGI) 对 DNA 进行甲基化修饰前处理, 使未发生甲基化修饰的胞嘧啶转化为尿嘧啶。然后使用上述荧光定量 PCR 反应体系对参考品进行检测。

检测结果如表 3 所示。

表 3: OPLAH 和 ADHFE1 基因检测参考品数据

甲基化率	OPLAH 的 DMR 区域				ADHFE1 的 DMR 区域			
	Ct			平均 Ct	Ct			平均 Ct
100%	28.43	28.31	28.3	28.35	28.44	28.47	28.61	28.51
50%	29.12	29.14	29.3	29.19	29.81	29.58	29.77	29.72
25%	30.56	30.4	30.55	30.50	30.83	30.9	30.91	30.88
13%	31.3	31.33	31.2	31.28	32.6	32.67	32.55	32.61
6%	32.42	32.35	32.57	32.45	33.05	33.21	34.34	33.53
3%	33.17	32.8	33.45	33.14	36.12	35.62	34.91	35.55
1%	34.1	34.3	34.5	34.30	35.87	35.94	36	35.94
0%	NoCt	NoCt	NoCt	/	NoCt	NoCt	NoCt	/

结果显示，所有混入甲基化模板的参考品均被检出，检测限可低至 1%。而未加入甲基化模板的参考品则无发生扩增。

- 5 根据投入的甲基化模板量和平均 Ct 值，绘制标准曲线，如图 2 所示，标准曲线的 R<sup>2</sup>良好。说明本方法能够有效检出样本中的 OPLAH 和 ADHFE1 基因甲基化信号，且无非特异扩增。

#### 实施例 2: OPLAH 和 ADHFE1 基因甲基化联合检测试剂盒用于肠癌检测

- 10 本实施例以荧光定量 PCR 方法为例，公开一种通过联合检测 OPLAH 和 ADHFE1 基因甲基化水平来实现肠癌检测的试剂盒。

- 15 本试剂盒包含 PCR 扩增所需要的 Buffer、dNTP 和酶等试剂，还包括 OPLAH 和 ADHFE1 基因引物组和探针。本实施例中所述引物组包括：OPLAH 基因引物组 (SEQ ID NO: 30 和 SEQ ID NO: 32)、OPLAH 基因探针 (SEQ ID NO: 170)、ADHFE1 基因引物组 (SEQ ID NO: 108 和 SEQ ID NO: 109) 和 ADHFE1 基因探针 (SEQ ID NO: 197)。为了监测样本中是否含有足够的人源 DNA，本试剂盒中还加入了用于检测内参基因的引物和探针。所述内参基因一般为管家基因，其对应的引物和探针检测结果仅与人源 DNA 总量有关，而不受甲基化水平影响。本实施例中选用 GAPDH 基因作为内参基因，其引物和探针序列如表 4 所示。PCR 反应在 95℃ 预变性 10min，然后进

行 95℃变性 30 秒、55℃退火 30 秒和 72℃延伸 30 秒的循环 40 个。采用宏石 96S qPCR 仪进行检测，基线、阈值按照默认设置，在每个循环结束时收集荧光信号。

表 4：内参基因引物、探针序列表

SEQ ID NO:	序列 (5'-3')	说明
210	TTTAGGAGTGAGTGGAAGATAGA	内参基因引物
211	AAACCACACCATCCTAATTACCT	
212	CCCAAAACACATTTCTTCCATTC	内参基因探针

5 本发明共纳入 604 例粪便样本对本试剂盒进行性能评估，临床样本信息如表 5 所示。

表 5：样本信息表

样本类型	肠癌患者	健康人	息肉患者	腺瘤患者	其他肠道疾病患者	进展期腺瘤患者	总计
样本数	284	146	49	55	13	57	604
年龄中位数 (岁)	59 (32-84)	63 (29-78)	64 (37-74)	65 (30-76)	56 (49-86)	63 (38-75)	62 (29-86)
性别 (n)							
男性	170	61	29	29	10	25	324
女性	114	85	20	26	3	32	280
肠癌分期							
I 期	44 (15.49%)						
II 期	82 (28.87%)						
III 期	78 (27.46%)						
IV 期	22						

	(7.75%)
未知	58 (20.42%)

样本检测首先使用一次性粪便样本采样器 (BGI) 采集受检者的粪便样本。再使用粪便样本核酸提取试剂盒 (BGI) 提取粪便样本 DNA。下一步使用甲基化检测样本前处理试剂盒 (BGI) 对 DNA 进行甲基化修饰预处理, 得到转化后的 DNA。最后, 使用本试剂盒对转化后的 DNA 进行检测, 获得每个样本的内参、OPLAH 和 ADHFE1 基因的 Ct 值。

结果分析时, 内参基因阈值为 37, 当样本内参基因 Ct 值大于 37 时, 认为样本中 人源 DNA 含量不足, 判定为检测失败。当内参基因 Ct 值小于等于 37 时, 质检合格。对质检合格的样本, 当目标基因 Ct 值大于 38 或无检出 Ct 值时, 判定该基因检测阴性。当目标基因 Ct 值小于等于 38 时, 进一步计算该目标基因的  $\Delta$ Ct 值 (目标基因 Ct 值与内参基因 Ct 值的差值)。

阳性截断值的确定使用 SPSS 进行分析, 首先本发明人对单个基因使用 SPSS 软件绘制了 ROC 曲线, 如图 3 所示, OPLAH 和 ADHFE1 基因的 AUC 分别为 0.911 和 0.918。通过整合两个基因的检测结果进行分析, 得到了更高的 AUC, 达到了 0.942。进一步地, 本发明人使用相同方法对上述样本的 PPP2R5C 和 SDC2 基因进行检测, 得到了 PPP2R5C 和 SDC2 基因的 Ct 值和  $\Delta$ Ct 值。所述 PPP2R5C 和 SDC2 基因在癌组织和非癌组织中亦存在较显著的甲基化差异。进一步地本发明人使用 SPSS 软件分析了 OPLAH、ADHFE1、PPP2R5C 和 SDC2 四基因组合的检测效果, 检测结果如图 3 所示, 可见四基因组合检测与 OPLAH 和 ADHFE1 双基因检测的检测效果相近, AUC 为 0.950。由于双基因检测即可达到四基因组合检测的效果, 从成本、检测稳定性等方面考虑, 确定使用双基因联合检测的方法。最终确定当 OPLAH 基因的 DMR 的  $\Delta$ Ct 值小于等于 7, 或 ADHFE1 基因的 DMR 的  $\Delta$ Ct 值小于等于 12 时, 临床样本检测性能最佳, 具体性能如表 6 所示, 本试剂盒灵敏度达到 86.62%, 总体特异性 (健康人样本和息肉样本合计的特异性) 达到 92.78%, 进展期腺瘤的灵敏度达到 52.63%。本试剂盒对早期肠癌亦有较高灵敏度, 如表 7 所示, I 期、II 期肠癌灵敏度分别达到了 88.64% 和 89.02%。

表 6：临床样本检测性能

样本类型	检测阳性数	检测阴性数	样本总量	灵敏度	特异性
肠癌患者	246	38	284	86.62%	
健康人	11	135	146	/	92.47%
息肉患者	1	48	49	/	97.96%
腺瘤患者	5	50	55	/	90.91%
其他肠道疾病患者	2	11	13	/	84.62%
进展期腺瘤患者	30	27	57	52.63%	/
总数	295	309	604	86.09% (总符合率)	

其中总符合率=（金标准阴性检测正确的样本数+金标准阳性检测正确的样本数）/样本总数。金标准是指肠镜检测。

5

表 7：各分期肠癌样本的检测灵敏度

样本分期	检测阳性数	样本总数	灵敏度
I 期	39	44	88.64%
II 期	73	82	89.02%
III 期	71	78	91.03%
IV 期	18	22	81.82%
未知	45	58	77.59%
总体	246	284	86.62%

按照上述方法，本发明人还使用表 1 中的其它一些引物、探针组合对上述样本进行了测试，均获得良好的检测效果，具体结果见下面的表 8。

表 8

引物序列号	探针序列号	肠癌灵敏度	进展期腺瘤灵敏度	健康人样本和息肉样本合计的特异性

25、26、144 和 145	164、204	83.45%	50.88%	88.21%
26、27、154 和 155	168、206	85.21%	38.60%	89.73%
33、34、156 和 157	171、209	71.13%	36.84%	91.63%
56、57、158 和 159	184、203	88.03%	56.14%	80.99%
61、62、108 和 109	183、191	84.51%	43.86%	88.97%
33、34、108 和 109	171、191	79.23%	42.11%	90.11%
25、26、158 和 159	164、203	75.35%	36.84%	90.49%

### **实施例 3: OPLAH 和 ADHFE1 基因甲基化联合粪便潜血检测试剂盒用于肠癌检测**

5 本发明公开一种通过联合检测 OPLAH 和 ADHFE1 基因甲基化水平以及粪便潜血来实现肠癌检测的试剂盒。所述试剂盒包括 OPLAH 和 ADHFE1 基因甲基化水平检测模块，以及粪便潜血检测模块。

#### **1. OPLAH 和 ADHFE1 基因甲基化水平检测模块**

10 所述 OPLAH 和 ADHFE1 基因甲基化水平检测模块基于荧光定量 PCR 方法，包含 PCR 扩增所需要的 Buffer、dNTP 和酶等试剂，还包括 OPLAH 和 ADHFE1 基因引物组和探针。本实施例中所述引物组可从表 1 中挑选，本例中选择的引物组包括：OPLAH 基因引物组 (SEQ ID NO: 30 和 SEQ ID NO: 32)、OPLAH 基因探针 (SEQ ID NO: 170)、ADHFE1 基因引物组 (SEQ ID NO: 108 和 SEQ ID NO: 109) 和 ADHFE1 基因探针 (SEQ ID NO: 197)。为了监测样本中是否含有足够的人源 DNA，本试剂盒中还加入了用于检测内参基因的引物和探针。所述内参基因一般为管家基因，其对应的引物和探针检测结果仅与人源 DNA 总量有关，而不受甲基化水平影响。本实施例

15



中选用 GAPDH 基因作为内参基因，其引物和探针序列如表 4 所示。PCR 反应在 95℃ 预变性 10min，然后进行 95℃ 变性 30 秒、55℃ 退火 30 秒和 72℃ 延伸 30 秒的循环 40 个。采用宏石 96S qPCR 仪进行检测，基线、阈值按照默认设置，在每个循环结束时收集荧光信号。

5 样本检测首先使用一次性粪便样本采样器 (BGI) 采集受检者的粪便样本。再使用粪便样本核酸提取试剂盒 (BGI) 提取粪便样本 DNA。下一步使用甲基化检测样本前处理试剂盒 (BGI) 对 DNA 进行甲基化修饰预处理，得到转化后的 DNA。最后，使用本试剂盒对转化后的 DNA 进行检测，获得每个样本的内参、OPLAH 和 ADHFE1 基因的 Ct 值。

10 结果分析时，内参基因阈值为 37，当样本内参基因 Ct 值大于 37 时，认为样本中人源 DNA 含量不足，判定为检测失败。当内参基因 Ct 值小于等于 37 时，质检合格。对质检合格的样本，当目标基因 Ct 值大于 38 或无检出 Ct 值时，判定该基因检测阴性。当目标基因 Ct 值小于等于 38 时，进一步计算该目标基因的  $\Delta Ct$  值 (目标基因 Ct 值与内参基因 Ct 值的差值)。

15 2. 粪便潜血模块

所述粪便潜血模块包含可定性检测粪便血红蛋白含量的试纸，其检测原理为胶体金法。具体地，使用 FIT 采样器采集粪便样本，与样本保存液混匀后取约 100 $\mu$ l 样本于试纸上，在一定时间内读取检测结果。

分析结果时，质控线应出现条带，否则检测失败；若质控线合格，则观察检测线，20 若检测线出现条带，则粪便潜血试验为阳性，无条带则为阴性。结果读取应在规定时间内完成，超时作废。

3. OPLAH 和 ADHFE1 基因甲基化检测联合粪便潜血检测的试剂盒的实现与性能评估

25 本实施例使用实施例 2 所述样本中具有粪便潜血检测结果的 318 例样本进行分析，样本的临床信息如表 9 所示。

表 9: 样本信息表

样本类型	肠癌患者	健康人	息肉患者	腺瘤患者	其他肠道疾病患者	进展期腺瘤患者	总计

样本数	149	96	21	21	2	29	318
年龄中位数 (岁)	59 (38-82)	63 (29-78)	64 (42-72)	63 (30-74)	55 (54-56)	62 (40-75)	61 (29-82)
性别 (n)							
男性	84	41	9	15	1	14	164
女性	65	55	12	6	1	15	154
肠癌分期							
I 期	19 (12.75%)						
II 期	32 (21.48%)						
III 期	45 (30.2%)						
IV 期	13 (8.72%)						
未知	40 (26.85%)						

阳性截断值的确定使用 SPSS 进行分析, 首先本发明人对基因检测结果使用 SPSS 软件绘制了 ROC 曲线, 如图 4 所示, OPLAH 和 ADHFE1 基因联检的 AUC 为 0.931, OPLAH 和 ADHFE1 基因在肠癌检测中具有较好的检测性能。进一步地, 本发明人将粪便潜血检测 (FIT) 结果纳入逻辑回归分析, 按照公式  $g(y)=1/(1+e^{-y})$  计算得到检测风险指数, 其中  $g(y)$  是风险指数, 当风险指数高于预设值时, 判定为阳性或候选为阳性;  $e$  为自然常数;  $y=(\theta_0+\theta_1*X_1+\theta_2*X_2+\dots+\theta_n*X_n)$ ,  $\theta_0$  为常数项,  $\theta_i(i=1, 2, \dots, n)$  是待求系数,  $X_i(i=1, 2, \dots, n)$  是基因和粪便潜血检测的结果。计算得到每个样本的检测风险指数后, 绘制 ROC 曲线确定最佳阳性截断值。如图 4 所示, 基因和 FIT 联检的 AUC 达到了 0.975, 基因与 FIT 联检能够有效提高检测性能。如表 10 所示, 当阳性截断值取检测风险指数 0.55 时, 达到最佳性能, 此时灵敏度和总体特异性分别为 95.97% 和 91.43%, 本试剂盒对进展期腺瘤具有较好的检出率, 达到了 58.62% 的灵敏度。本试

试剂盒对早期肠癌具有较高灵敏度，如表 11 所示，I 期、II 期肠癌灵敏度分别达到了 89.47% 和 100%。如图 5 所示，与双基因检测或 FIT 检测对比，本试剂盒对结直肠癌和进展期腺瘤都具有更高的检测灵敏度。

表 10：临床样本检测性能

样本类型	检测阳性数	检测阴性数	样本总量	灵敏度	特异性
肠癌患者	143	6	149	95.97%	
健康人	4	92	96	/	95.83%
息肉患者	2	19	21	/	90.48%
腺瘤患者	5	16	21	/	76.19%
其他肠道疾病患者	1	1	2	/	50.00%
进展期腺瘤患者	17	12	29	58.62%	/
总数	172	146	318	90.57% (总符合率)	

- 5 其中总符合率= (金标准阴性检测正确的样本数+金标准阳性检测正确的样本数) / 样本总数。金标准是指肠镜检测。

表 11：各分期肠癌样本的检测灵敏度

样本分期	检测阳性数	样本总数	灵敏度
I 期	17	19	89.47%
II 期	32	32	100.00%
III 期	43	45	95.56%
IV 期	12	13	92.31%
未知	39	40	97.50%
总体	143	149	95.97%

10

尽管本发明的具体实施方式已经得到详细的描述，本领域技术人员将会理解。根据已经公开的所有教导，可以对那些细节进行各种修改和替换，这些改变均在本发明的保护范围之内。本发明的全部范围由所附权利要求及其任何等同物给出。

## 权 利 要 求

1. 差异性甲基化区域，包括：

(1) chr8: 145106171-145107467，和

(2) chr8:67344198-67345563。

2. 根据权利要求 1 所述的差异性甲基化区域，其中：

第 (1) 项所示的差异性甲基化区域的碱基序列如 SEQ ID NO: 1 或其互补序列所示，和/或

第 (2) 项所示的差异性甲基化区域的碱基序列如 SEQ ID NO: 2 或其互补序列所示。

3. 根据权利要求 1 至 2 中任一权利要求所述的差异性甲基化区域，其用于检测结直肠癌或结直肠癌癌前病变，或者用于结直肠癌患者的预后风险评估。

4. 根据权利要求 1 至 3 中任一权利要求所述的差异性甲基化区域，其经过了如下处理：

亚硫酸氢盐处理、甲基化敏感性内切酶处理、甲基化限制性内切酶处理或甲基化修饰酶处理。

5. 一种试剂盒，包括检测权利要求 1 至 4 中任一权利要求所述的差异性甲基化区域的试剂或者所述差异性甲基化区域内的甲基化位点的试剂。

6. 根据权利要求 5 所述的试剂盒，其中，所述试剂包括：

1) 检测第 (1) 项所示的差异性甲基化区域或者所述差异性甲基化区域内的甲基化位点的引物和探针；

2) 检测第 (2) 项所示的差异性甲基化区域或者所述差异性甲基化区域内的甲基化位点的引物和探针；

优选地，所述试剂还包括：

3) 内参基因的引物和探针。

7. 根据权利要求 6 所述的试剂盒, 其中:

1) 中引物的序列选自 SEQ ID NOs: 3-79, 探针的序列选自 SEQ ID NOs: 160-188; 和 2) 中引物的序列选自 SEQ ID NOs: 80-159, 探针的序列选自 SEQ ID NOs: 189-209;

优选地:

1) 中引物的序列选自 SEQ ID NO: 30 和 SEQ ID NO: 32, 探针的序列选自 SEQ ID NO: 170; 和 2) 中引物的序列选自 SEQ ID NOs: 108-109, 探针的序列选自 SEQ ID NO: 197;

1) 中引物的序列选自 SEQ ID NOs: 25-26, 探针的序列选自 SEQ ID NO: 164; 和 2) 中引物的序列选自 SEQ ID NOs: 144-145, 探针的序列选自 SEQ ID NO: 204;

1) 中引物的序列选自 SEQ ID NOs: 26-27, 探针的序列选自 SEQ ID NO: 168; 和 2) 中引物的序列选自 SEQ ID NOs: 154-155, 探针的序列选自 SEQ ID NO: 206;

1) 中引物的序列选自 SEQ ID NOs: 33-34, 探针的序列选自 SEQ ID NO: 171; 和 2) 中引物的序列选自 SEQ ID NOs: 156-157, 探针的序列选自 SEQ ID NO: 209;

1) 中引物的序列选自 SEQ ID NOs: 56-57, 探针的序列选自 SEQ ID NO: 184; 和 2) 中引物的序列选自 SEQ ID NOs: 158-159, 探针的序列选自 SEQ ID NO: 203;

1) 中引物的序列选自 SEQ ID NOs: 61-62, 探针的序列选自 SEQ ID NO: 183; 和 2) 中引物的序列选自 SEQ ID NOs: 108-109, 探针的序列选自 SEQ ID NO: 191;

1) 中引物的序列选自 SEQ ID NOs: 33-34, 探针的序列选自 SEQ ID NO: 171; 和 2) 中引物的序列选自 SEQ ID NOs: 108-109, 探针的序列选自 SEQ ID NO: 191;

或者,

1) 中引物的序列选自 SEQ ID NOs: 25-26, 探针的序列选自 SEQ ID NO: 164; 和 2) 中引物的序列选自 SEQ ID NOs: 158-159, 探针的序列选自 SEQ ID NO: 203;

优选地, 3) 中内参基因的引物的序列如 SEQ ID NOs: 210-211 所示, 内参基因的探针的序列如 SEQ ID NO: 212 所示。

8. 根据权利要求 5 至 7 中任一权利要求所述的试剂盒, 其还包含选自如下的一项

或者多项：

粪便样本采样器、核酸提取试剂、甲基化检测样本前处理试剂、PCR 试剂和粪便潜血检测试剂；

优选地，所述粪便潜血检测试剂为粪便免疫化学检测法检测试剂、愈创木脂检测法检测试剂、四甲基联苯胺法检测试剂或 ELISA 双抗夹心法检测试剂；

优选地，所述粪便免疫化学检测法检测试剂为免疫胶体金法粪便潜血检测试剂。

9. 根据权利要求8所述的试剂盒，其中，所述甲基化检测样本前处理试剂为亚硫酸氢盐、甲基化敏感性内切酶、甲基化限制性内切酶或甲基化修饰酶。

10. 权利要求 1 至 4 中任一权利要求所述的差异性甲基化区域或者检测权利要求 1 至 4 中任一权利要求所述的差异性甲基化区域或者所述差异性甲基化区域内的甲基化位点的试剂在制备用于检测结直肠癌或结直肠癌前病变，或者用于结直肠癌患者的预后风险评估的药物中的用途。

11. 根据权利要求 10 所述的用途，其中，所述试剂为用于检测权利要求 1 至 4 中任一权利要求所述的差异性甲基化区域或者所述差异性甲基化区域内的甲基化位点的引物和探针。

12. 一种检测结直肠癌或结直肠癌前病变或者一种评估结直肠癌患者的预后风险的方法，包括如下步骤：

检测权利要求 1 至 4 中任一权利要求所述的差异性甲基化区域的含量或者其中的片段的含量；其中，所述片段包含所述差异性甲基化区域内的一个或多个甲基化位点。

13. 根据权利要求 12 所述的方法，其中，通过荧光定量 PCR 方法检测权利要求 1 至 4 中任一权利要求所述的各差异性甲基化区域或者其中的片段的含量。

14. 根据权利要求 13 所述的方法，其中，

采用内参基因 Ct 值评估样本中人源 DNA 的含量，当内参基因 Ct 值 $>37$  时，判

定样本人源 DNA 含量过低,本次检测失败,当内参基因 Ct 值 $\leq 37$ 时,评估样本合格,可进行后续分析。

15. 根据权利要求 14 所述的方法,其中,对于评估合格的样本:

当各差异性甲基化区域的 Ct 值大于 38 或无检出 Ct 值时,判定为检测阴性;

当各差异性甲基化区域的 Ct 值小于等于 38 时,进一步计算各差异性甲基化区域与内参基因 Ct 值的差值 $\Delta Ct$ 值。

16. 根据权利要求 13 至 15 中任一权利要求所述的方法,其中:

如果第(1)项或第(2)项的 $\Delta Ct \leq$ 截断值,判定为结直肠癌高风险,否则判定为低风险;

其中 $\Delta Ct =$ (所述差异性甲基化区域的 Ct 值或者扩增片段的 Ct 值)-内参基因 Ct 值;

其中,所述扩增片段包含所述差异性甲基化区域内的一个或多个甲基化位点。

17. 根据权利要求 12 至 16 中任一权利要求所述的方法,其中还包括联合粪便潜血检测结果进行评估的步骤。

18. 根据权利要求 12 至 17 中任一权利要求所述的方法,其中,所述样本为人体组织样本、血液样本、细胞样本、分泌物样本或排泄物样本例如粪便样本。

19. 一种结直肠癌的诊断或评估模型,采用逻辑函数 $g(y) = 1/(1+e^{-y})$ 进行计算,其中:

$g(y)$  是风险指数,当风险指数高于预设值时,判定为阳性或候选为阳性;

$e$ 为自然常数;

$y = (\theta_0 + \theta_1 * X_1 + \theta_2 * X_2 + \dots + \theta_n * X_n)$ ,  $\theta_0$ 为常数项,  $\theta_i (i=1, 2, \dots, n)$ 是待求系数,  $X_i (i=1, 2, \dots, n)$ 是权利要求 1 至 4 中任一权利要求所述的差异性甲基化区域的检测结果和粪便潜血的检测结果。

20. 一种诊断或评估结直肠癌的系统,包括:

1) DMR检测单元：获取权利要求1至4中任一权利要求所述的差异性甲基化区域或者其中的片段的含量数据；其中，所述片段包含所述差异性甲基化区域内的一个或多个甲基化位点；

2) 粪便潜血检测单元：获取粪便潜血检测结果数据；

3) 数据分析单元：将DMR检测单元和粪便潜血检测单元获得的数据输入权利要求19所述的结直肠癌的诊断或评估模型进行分析；

4) 风险评估单元：输出个体样本的风险指数或者判断结果。

21. 权利要求 1 至 4 中任一权利要求所述的差异性甲基化区域在制备检测结直肠癌或者结直肠癌前病变的产品中的用途。



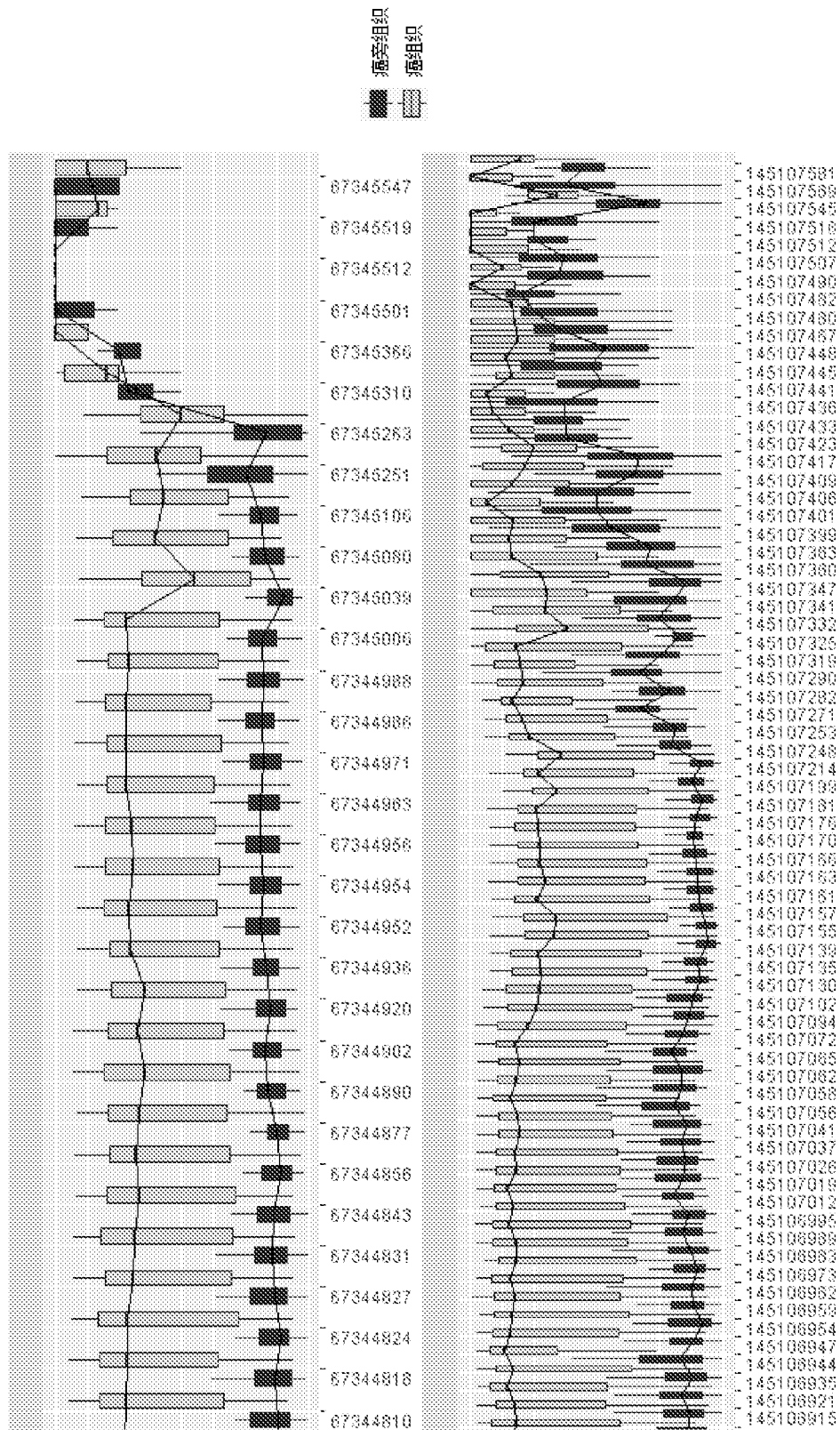


图 1

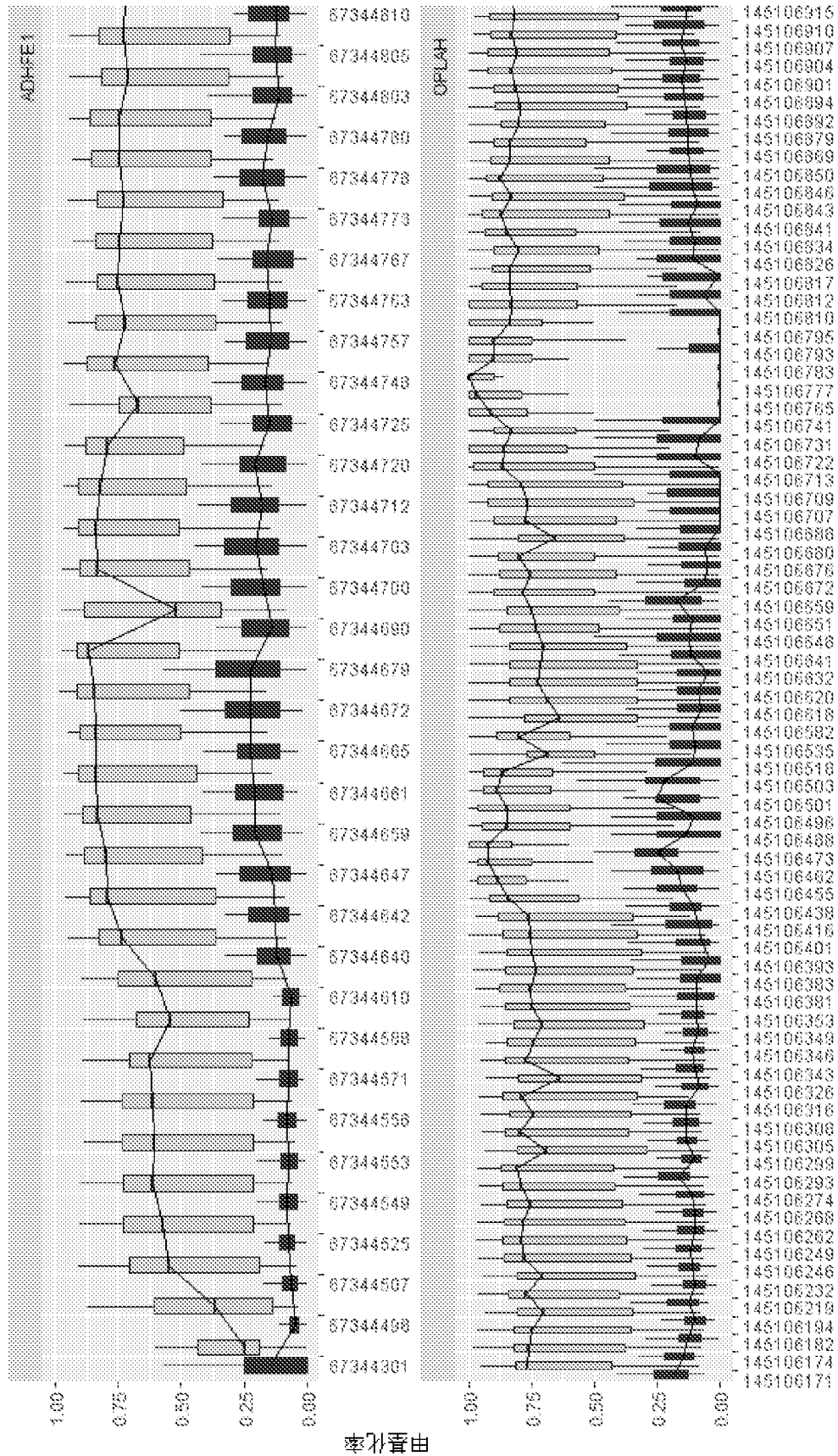


图 1 (续)

### OPLAH和ADHFE1基因检测参考品结果

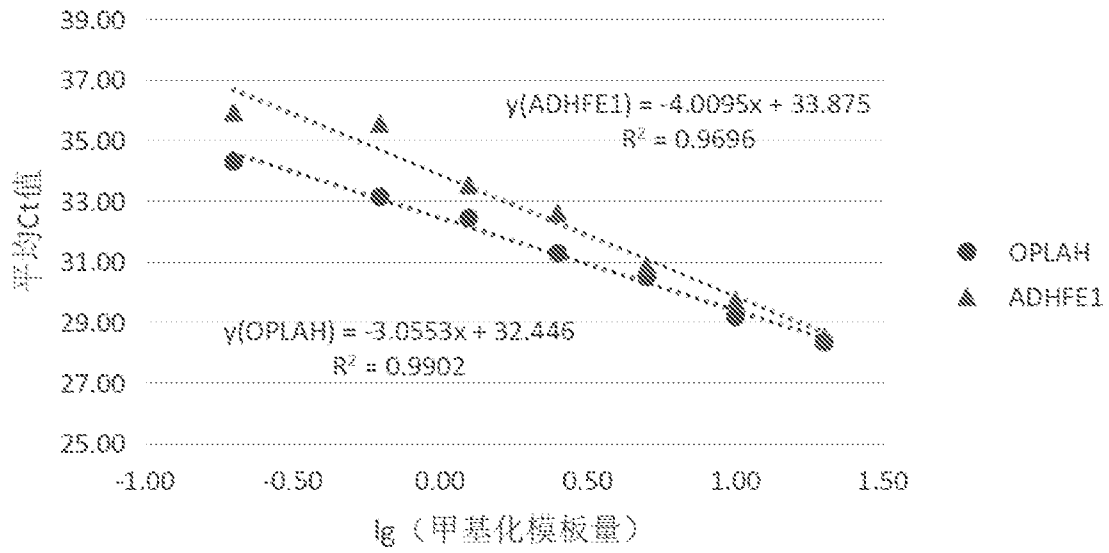


图 2

### ROC 曲线

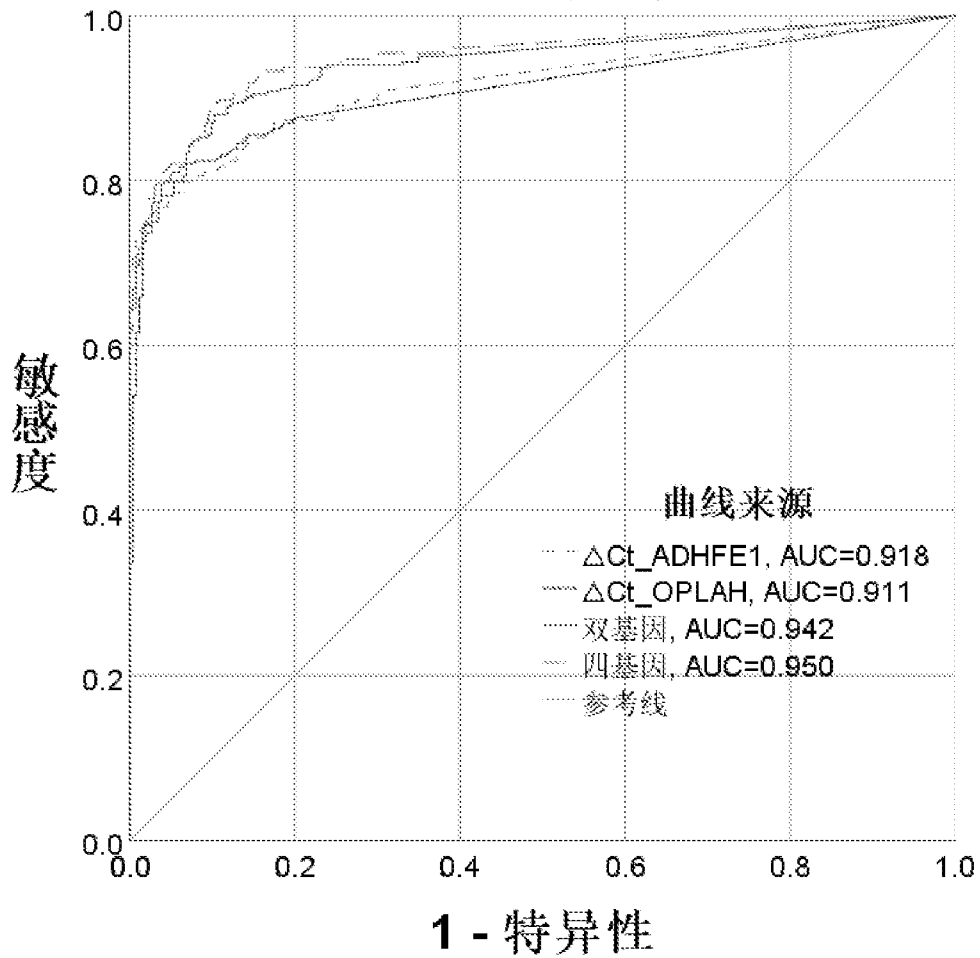


图 3

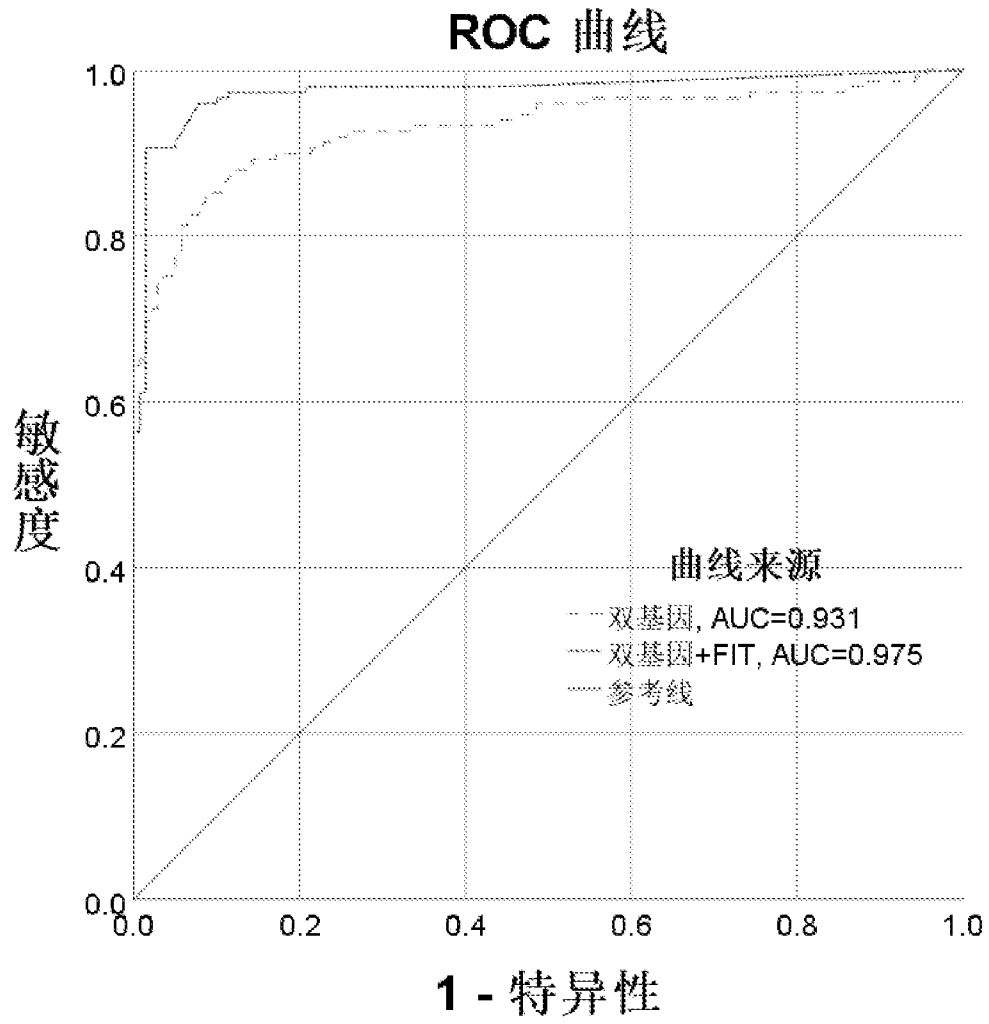


图 4

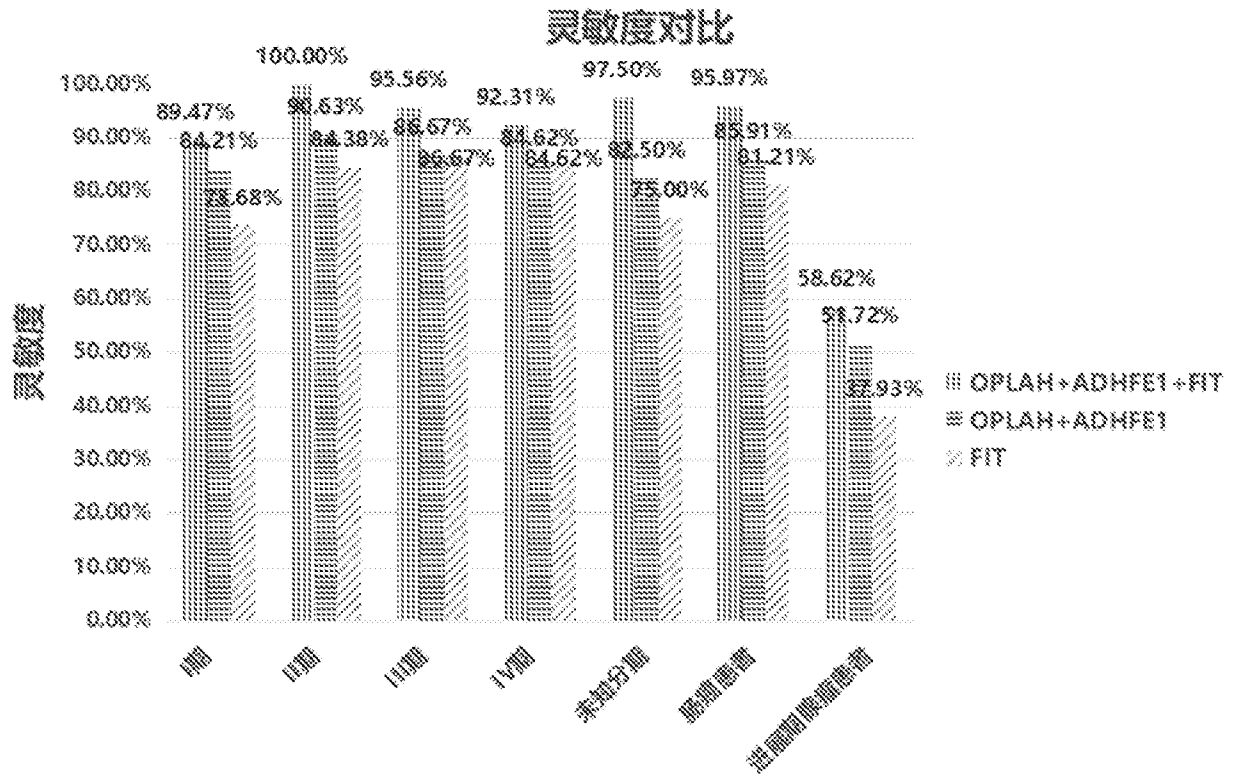


图 5

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2022/116795

<b>A. CLASSIFICATION OF SUBJECT MATTER</b> C12Q1/6886(2018.01)i;G01N 33/574(2006.01)i  According to International Patent Classification (IPC) or to both national classification and IPC		
<b>B. FIELDS SEARCHED</b> 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) VEN, CNABS, PubMed, ISI Web of Science, VCN, CNKI, 万方, WANFANG, 百度学术, BAIDU SCHOLAR, 读秀, DUXIU, STN, GenBank, 中国专利生物序列检索系统, China Patent Biological Sequence Search System: OPLAH, ADHFE1, methylation, biomarker, colorectal cancer, primer, probe, sequencing, 生物标志物, 甲基化, 引物, 探针, 癌, SEQ ID NOs: 1-2.		
<b>C. DOCUMENTS CONSIDERED TO BE RELEVANT</b>		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	BAHARUDIN, R. et al. "Epigenome-Wide DNA Methylation Profiling in Colorectal Cancer and Normal Adjacent Colon Using Infinium Human Methylation 450K" <i>Diagnostics</i> , Vol. 12, No. 98, 14 January 2022 (2022-01-14), abstract	1-21
X	NAUMOV, V. A. et al. "Genome-scale analysis of DNA methylation in colorectal cancer using Infinium HumanMethylation450 BeadChips" <i>Epigenetics</i> , Vol. 8, No. 9, 30 September 2013 (2013-09-30), abstract	1-21
A	CN 111500730 A (SUZHOU WEISHAN BIOTECHNOLOGY CO., LTD.) 07 August 2020 (2020-08-07) entire document	1-21
A	CN 110475875 A (EXACT SCIENCES DEVELOPMENT COMPANY LLC. et al.) 19 November 2019 (2019-11-19) entire document	1-21
<input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C. <input checked="" type="checkbox"/> See patent family annex.		
<p>* Special categories of cited documents:</p> <p>“A” document defining the general state of the art which is not considered to be of particular relevance</p> <p>“D” document cited by the applicant in the international application</p> <p>“E” earlier application or patent but published on or after the international filing date</p> <p>“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)</p> <p>“O” document referring to an oral disclosure, use, exhibition or other means</p> <p>“P” document published prior to the international filing date but later than the priority date claimed</p> <p>“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</p> <p>“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</p> <p>“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</p> <p>“&amp;” document member of the same patent family</p>		
Date of the actual completion of the international search <b>05 May 2023</b>		Date of mailing of the international search report <b>10 May 2023</b>
Name and mailing address of the ISA/CN <b>China National Intellectual Property Administration (ISA/ CN) China No. 6, Xitucheng Road, Jimenqiao, Haidian District, Beijing 100088</b>		Authorized officer   Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

**PCT/CN2022/116795**

<b>C. DOCUMENTS CONSIDERED TO BE RELEVANT</b>		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2021047694 A1 (THE BROAD INSTITUTE, INC. et al.) 18 February 2021 (2021-02-18) entire document	1-21
A	WO 2022075788 A1 (GENCURIX INC.) 14 April 2022 (2022-04-14) entire document	1-21
A	WO 2021091239 A1 (GENOMIC TREE, INC.) 14 May 2021 (2021-05-14) entire document	1-21

**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.
  - b.  furnished subsequent to the international filing date for the purposes of international search (Rule 13ter.1(a)),  
 accompanied by a statement to the effect that the sequence listing does not go beyond the disclosure in the international application as filed.
2.  With regard to any nucleotide and/or amino acid sequence disclosed in the international application, this report has been established to the extent that a meaningful search could be carried out without a WIPO Standard ST.26 compliant sequence listing.
3. Additional comments:



**INTERNATIONAL SEARCH REPORT**  
**Information on patent family members**

International application No.

**PCT/CN2022/116795**

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	111500730	A	07 August 2020	None			
CN	110475875	A	19 November 2019	WO	2018140781	A1	02 August 2018
				AU	2018211956	A1	25 July 2019
				CA	3049459	A1	02 August 2018
				US	2022056526	A1	24 February 2022
				JP	2020506680	A	05 March 2020
				EP	3574110	A1	04 December 2019
				EP	3574110	A4	13 January 2021
				US	2018245157	A1	30 August 2018
				US	11118228	B2	14 September 2021
				JP	2023025184	A	21 February 2023
				KR	20190111945	A	02 October 2019
US	2021047694	A1	18 February 2021	None			
WO	2022075788	A1	14 April 2022	KR	20220047065	A	15 April 2022
				KR	102472257	B1	30 November 2022
WO	2021091239	A1	14 May 2021	JP	2023500923	A	11 January 2023
				KR	20210055281	A	17 May 2021
				KR	102261606	B1	07 June 2021
				EP	4056715	A1	14 September 2022
				US	2023002830	A1	05 January 2023

<b>A. 主题的分类</b> C12Q1/6886 (2018.01) i; G01N 33/574 (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>B. 检索领域</b> 检索的最低限度文献 (标明分类系统和分类号) C12Q, G01N  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)) VEN, CNABS, PubMed, ISI Web of Science, VCN, CNKI, 万方, 百度学术, 读秀, STN, GenBank, 中国专利生物序列检索系统: OPLAH, ADHFE1, methylation, biomarker, colorectal cancer, primer, probe, sequencing, 生物标志物, 甲基化, 引物, 探针, 癌, SEQ ID NOs:1-2.		
<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
X	BAHARUDIN, R. 等. "Epigenome-Wide DNA Methylation Profiling in Colorectal Cancer and Normal Adjacent Colon Using Infinium Human Methylation 450K" Diagnostics, 第12卷, 第98期, 2022年1月14日 (2022 - 01 - 14), 摘要	1-21
X	NAUMOV, V. A. 等. "Genome-scale analysis of DNA methylation in colorectal cancer using Infinium HumanMethylation450 BeadChips" Epigenetics, 第8卷, 第9期, 2013年9月30日 (2013 - 09 - 30), 摘要	1-21
A	CN 111500730 A (苏州唯善生物科技有限公司) 2020年8月7日 (2020 - 08 - 07) 全文	1-21
A	CN 110475875 A (精密科学发展有限责任公司 等) 2019年11月19日 (2019 - 11 - 19) 全文	1-21
A	US 2021047694 A1 (The Broad Institute, Inc. 等) 2021年2月18日 (2021 - 02 - 18) 全文	1-21
<input checked="" 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: "A" 认为不特别相关的表示了现有技术一般状态的文件 "D" 申请人在国际申请中引证的文件 "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 2023年5月5日		国际检索报告邮寄日期 2023年5月10日
ISA/CN的名称和邮寄地址 中国国家知识产权局 中国北京市海淀区蓟门桥西土城路6号 100088		授权官员 李有朝  电话号码 (+86) 010-53961930

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	WO 2022075788 A1 (GENCURIX INC.) 2022年4月14日 (2022 - 04 - 14) 全文	1-21
A	WO 2021091239 A1 (GENOMICTREE, INC.) 2021年5月14日 (2021 - 05 - 14) 全文	1-21

第1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列表进行的:
- a.  作为国际申请的一部分提交的;
- b.  为国际检索的目的在国际申请日之后提交(细则13之三.1(a)),  
 附有说明序列列表不超出所提交国际申请公开范围的声明。
2.  本报告是在没有收到符合WIPO ST. 26标准的序列列表的情况下, 考虑了国际申请中披露的任何核苷酸和/或氨基酸序列, 在可进行有意义检索的范围内做出的。
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/CN2022/116795

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	111500730	A	2020年8月7日	无			
CN	110475875	A	2019年11月19日	WO	2018140781	A1	2018年8月2日
				AU	2018211956	A1	2019年7月25日
				CA	3049459	A1	2018年8月2日
				US	2022056526	A1	2022年2月24日
				JP	2020506680	A	2020年3月5日
				EP	3574110	A1	2019年12月4日
				EP	3574110	A4	2021年1月13日
				US	2018245157	A1	2018年8月30日
				US	11118228	B2	2021年9月14日
				JP	2023025184	A	2023年2月21日
				KR	20190111945	A	2019年10月2日
US	2021047694	A1	2021年2月18日	无			
WO	2022075788	A1	2022年4月14日	KR	20220047065	A	2022年4月15日
				KR	102472257	B1	2022年11月30日
WO	2021091239	A1	2021年5月14日	JP	2023500923	A	2023年1月11日
				KR	20210055281	A	2021年5月17日
				KR	102261606	B1	2021年6月7日
				EP	4056715	A1	2022年9月14日
				US	2023002830	A1	2023年1月5日