



(12) 发明专利申请

(10) 申请公布号 CN 118147307 A

(43) 申请公布日 2024.06.07

(21) 申请号 202410376668.X

哈特姆·T·阿拉维

(22) 申请日 2018.02.27

(74) 专利代理机构 成都超凡明远知识产权代理

(30) 优先权数据

62/464,800 2017.02.28 US

有限公司 51258

专利代理人 张大皓

(62) 分案原申请数据

201880014641.9 2018.02.27

(51) Int.Cl.

C12Q 1/6886 (2018.01)

C12N 15/11 (2006.01)

(71) 申请人 梅约医学教育与研究基金会

地址 美国明尼苏达州

申请人 精密科学公司

(72) 发明人 大卫·A·阿尔奎斯特

威廉·R·泰勒 约翰·B·基谢尔

特蕾西·C·亚卜

道格拉斯·W·马奥尼

布赖恩·A·杜克

马修·T·格特曼

权利要求书1页 说明书77页

序列表 (电子公布)

(54) 发明名称

用于表征来自人患者的样品的方法

(57) 摘要

本申请涉及用于表征来自人患者的样品的方法。所述方法包括：A) 从人患者的样品获得DNA；B) 测定DNA甲基化标志物的甲基化状态，所述DNA甲基化标志物包含选自由来自表3或表22的DMR 1-140组成的组的差异甲基化区域(DMR)中的碱基；C) 将所述一种或多种DNA甲基化标志物的测定甲基化状态与未患有前列腺癌的人患者的一种或多种DNA甲基化标志物的甲基化水平参考值进行比较。

1. 一种用于表征来自人患者的样品的方法,所述方法包括:
 - a) 从人患者的样品获得DNA;
 - b) 测定DNA甲基化标志物的甲基化状态,所述DNA甲基化标志物包含选自由来自表3或表22的DMR 1-140组成的组的差异甲基化区域(DMR)中的碱基;
 - c) 将所述一种或多种DNA甲基化标志物的测定甲基化状态与未患有前列腺癌的人患者的一种或多种DNA甲基化标志物的甲基化水平参考值进行比较。
2. 如权利要求1所述的方法,其中所述样品是粪便样品、组织样品、前列腺组织样品、血液样品、血浆样品或尿液样品。
3. 如权利要求1所述的方法
其中所述样品是前列腺组织样品,并且所述DMR选自DMR号63、3、64、70、7、39、8、10、11、12、14、41、81、16、17、18、20、21、44、25和47;或者
其中所述样品是血浆样品,并且所述DMR选自DMR号17、12、45和47。
4. 如权利要求1所述的方法,其包括测定多个DNA甲基化标志物。
5. 如权利要求1所述的方法,其包括测定2至11种DNA甲基化标志物。
6. 如权利要求1所述的方法,其包括测定12至140种DNA甲基化标志物。
7. 如权利要求1所述的方法,其中测定所述样品中的所述一种或多种DNA甲基化标志物的甲基化状态包括确定一个碱基的甲基化状态。
8. 如权利要求1所述的方法,其中测定所述样品中的所述一种或多种DNA甲基化标志物的甲基化状态包括确定多个碱基处的甲基化程度。
9. 如权利要求1所述的方法,其包括测定正向链的甲基化状态或测定反向链的甲基化状态。
10. 如权利要求1所述的方法,其中所述DNA甲基化标志物是具有100个或更少碱基的区域。

用于表征来自人患者的样品的方法

[0001] 本申请是申请日为2018年02月27日发明名称为“检测前列腺癌”的申请号为201880014641.9的发明专利申请的分案申请。

[0002] 相关申请的交叉参考

[0003] 本申请要求2017年2月28日提交的美国临时专利申请号62/464,800的优先权,所述专利的内容以引用的方式整体并入。

技术领域

[0004] 本文提供了用于前列腺癌筛查的技术,并且具体地但非排他地涉及用于检测前列腺癌的存在的方法、组合物和相关用途。

背景技术

[0005] 前列腺癌 (PCa) 是男性中第二常见的诊断癌症,2008年全球新发病例数为903,000例并且死亡人数为258,000例。虽然PCA很常见,但所述疾病在临床行为方面也是多样化的。据估计,大约每6名美国男性中就有1人被诊断患有PCA,但美国男性的PCA死亡率仅为2.8% (36名男性中有1人) (参见例如,Strand SH等人,Int J Mol Sci 2014;15:16544-16576)。

[0006] PCA的存活取决于许多因素。早期诊断为不太晚期的疾病为大多数男性提供了治愈性治疗的最佳机会。事实上,通过使用前列腺特异性抗原 (PSA) 测试促进了早期PCA诊断。在PSA测试时代,PCA的阶段和等级迁移已发生,从而导致鉴定更适合于确定性治疗的疾病。尽管PSA在PCA的诊断和治疗中是有益的,但使用PSA的筛查也被认为是有争议的。使用PSA的筛查可能导致更多惰性、低风险肿瘤的诊断,并且随后的过度治疗可能使男性经受不必要的生活质量危害(勃起功能障碍、尿失禁)。因此,需要新的生物标志物来辅助PCA的诊断。还需要新的测试来为男性提供有关其癌症的改善的预后信息。

[0007] 本发明解决了这些需要。

发明内容

[0008] 甲基化DNA已被研究作为大多数肿瘤类型组织中潜在类别的生物标志物。在许多情况下,DNA甲基转移酶在胞嘧啶-磷酸-鸟嘌呤 (CpG) 岛位点处向DNA添加甲基作为基因表达的表观遗传控制。在生物学上有吸引力的机制中,肿瘤抑制基因的启动子区域中的获得性甲基化事件被认为是沉默表达,因此有助于肿瘤发生。DNA甲基化可能是比RNA或蛋白质表达更具化学和生物学稳定性的诊断工具 (Laird (2010) Nat Rev Genet 11:191-203)。此外,在其他癌症(像散发性结肠癌)中,甲基化标志物提供了极好的特异性,并且比单个DNA突变具有更广泛的信息和敏感性 (Zou等人 (2007) Cancer Epidemiol Biomarkers Prev 16:2686-96)。

[0009] 当应用于动物模型和人细胞系时,CpG岛的分析已产生了重要的发现。例如,Zhang及其同事发现,来自同一CpG岛不同部位的扩增子可具有不同的甲基化水平 (Zhang等人 (2009) PLoS Genet 5:e1000438)。此外,甲基化水平在高度甲基化与未甲基化序列之间双

峰状分布,进一步支持DNA甲基转移酶活性的二进制开关样模式 (Zhang等人 (2009) PLoS Genet 5:e1000438)。对体内小鼠组织和体外细胞系的分析表明,仅约0.3%的高CpG密度启动子 (HCP, 定义为在300个碱基对区域内具有>7% CpG序列) 被甲基化,而低CpG密度的区域 (LCP, 定义为在300个碱基对区域内具有<5% CpG序列) 倾向于经常以动态组织特异性模式被甲基化 (Meissner等人 (2008) Nature 454: 766-70)。HCP包括普遍存在的管家基因和高度调节的发育基因的启动子。在>50% 甲基化的HCP位点中有几种已建立的标志物,如Wnt 2、NDRG2、SFRP2和BMP3 (Meissner等人 (2008) Nature 454:766-70)。

[0010] 通过DNA甲基转移酶在胞嘧啶-磷酸-鸟嘌呤 (CpG) 岛位点处的DNA的表观遗传甲基化已被研究作为大多数肿瘤类型组织中潜在类别的生物标志物。在生物学上有吸引力的机制中,肿瘤抑制基因的启动子区域中的获得性甲基化事件被认为是沉默表达,从而有助于肿瘤发生。DNA甲基化可能是比RNA或蛋白质表达更具化学和生物学稳定性的诊断工具。此外,在其他癌症 (像散发性结肠癌) 中,异常甲基化标志物比单个DNA突变具有更广泛的信息和敏感性,并提供极好的特异性。

[0011] 有几种方法可用于寻找新的甲基化标志物。虽然基于微阵列的CpG甲基化询问是合理的高通量方法,但此策略偏向于已知的兴趣区域,主要是已建立的肿瘤抑制启动子。在过去十年中,已开发了用于DNA甲基化的全基因组分析的替代方法。存在三种基本方法。第一种采用通过识别特定甲基化位点的限制性内切酶消化DNA,随后提供几种可能的分析技术,其提供限于酶识别位点的甲基化数据或用于在定量步骤 (如甲基化特异性PCR;MSP) 中扩增DNA的引物。第二种方法使用针对甲基-胞嘧啶或其他甲基化特异性结合结构域的抗体富集基因组DNA的甲基化部分,随后进行微阵列分析或测序以将片段定位于参考基因组。此方法不提供片段内所有甲基化位点的单核苷酸分辨率。第三种方法开始于亚硫酸氢盐处理DNA以将所有未甲基化的胞嘧啶转化为尿嘧啶,随后进行限制性内切酶消化并在偶联至衔接子配体后对所有片段进行完全测序。限制性内切酶的选择可富集CpG密集区域的片段,从而减少可能在分析期间映射到多个基因位置的冗余序列的数量。

[0012] RRBS以所有CpG岛的80% -90%的单核苷酸分辨率和中等至高读数覆盖的大多数肿瘤抑制启动子产生CpG甲基化状况数据。在癌症病例-对照研究中,对这些读数的分析导致鉴定差异甲基化区域 (DMR)。在先前对胰腺癌样本的RRBS分析中,发现了数百种DMR,其中许多从未与致癌作用相关联,并且其中许多未被注释。对独立组织样品的进一步验证研究确认了标志物CpG,其在性能方面是100%敏感和特异的。

[0013] 本文提供了用于前列腺癌筛查的技术,并且具体地但非排他地涉及用于检测前列腺癌的存在的方法、组合物和相关用途。

[0014] 事实上,如实施例I-VIII中所述,在鉴定本发明实施方案的过程中进行的实验鉴定了一组新的73种差异甲基化区域 (DMR),用于区分前列腺来源的DNA与非肿瘤性对照DNA的癌症。另外,鉴定了10种新的DMR,其在前列腺上皮细胞 (癌症和正常) 中但不在正常白细胞DNA样品中甲基化。从CpG富集的亚硫酸氢盐转化的肿瘤和正常DNA的下一代测序研究鉴定了两组区域。肿瘤样品包括较少侵袭性的格里森6和更具侵袭性的格里森7+模式。使用专有过滤器和分析管道选择DMR,并使用新的甲基化特异性PCR (MSP) 测定法在独立的组织样品组中进行验证。这73种生物标志物测定证明在组织中具有优异的检测能力,并具有广泛的临床特异性-一些用于许多不同器官部位的癌症,另一些仅对前列腺癌特异。

[0015] 此类实验列出并描述了区分前列腺癌组织和良性前列腺组织的120种新的DNA甲基化标志物(表3)。从这120种新的DNA甲基化标志物,进一步的实验鉴定了73种能够区分侵袭性前列腺癌组织(例如,格里森评分7+)和良性前列腺组织的标志物。更具体地,鉴定了能够区分前列腺癌组织和良性前列腺组织的标志物和/或标志物组(例如,具有选自以下的注释的染色体区域:ACOXL、AKR1B1_3644、ANXA2、CHST11_2206、FLJ45983、GAS6、GRASP、HAPLN3、HCG4P6、HES5_0822、ITPRIPL1、KCNK4、MAX.chr1.61519554-61519667、MAX.chr2.97193166-97193253、MAX.chr3.193、MAX.chr3.72788028-72788112、RAI1_7469、RASSF2、SERPINB9_3389、SLC4A11和TPM4_8047)(参见实施例I-VI)。

[0016] 在开发本发明实施方案的过程中进行的另外的实验涉及鉴定能够区分前列腺癌组织和良性前列腺组织的标志物(例如,具有选自以下的注释的染色体区域:SERPINB9_3479、FL0T1_1665、HCG4P6_4618、CHST11_2206、MAX.chr12.485、GRASP_0932、GAS6_6425、MAX.chr3.193、MAX.chr2.971_3164、MAX.chr3.727_8028、HES5_0840、TPM4_8037、SLC03A1_6187、ITPRIPL1_1244、AKR1B1_3644、RASGRF2_6325、ZNF655_6075、PAMR1_7364、ST6GALNAC2_1113、CCN JL_9070、KCNB2_9128、IGFBP7_6412和WNT3A_5487)能够区分前列腺癌组织与良性前列腺组织(参见实施例VIII;表20)。

[0017] 在开发本发明实施方案的过程中进行的另外的实验涉及鉴定能够区分侵袭性前列腺癌组织(例如,格里森评分7+)和较少侵袭性前列腺癌组织(例如,格里森评分6)的标志物(例如,具有选自以下的注释的染色体区域:SERPINB9_3479、GRASP_0932、SLC03A1_6187、ITPRIPL1_1244、AKR1B1_3644、RASGRF2_6325、ZNF655_6075、PAMR1_7364、ST6GALNAC2_1113、CCN JL_9070、KCNB2_9128、IGFBP7_6412和WNT3A_5487)能够区分前列腺癌组织与良性前列腺组织(参见实施例VIII;表20)。

[0018] 在开发本发明实施方案的过程中进行的另外的实验涉及鉴定能够检测血液样品(例如,血浆样品)中前列腺癌的存在或不存在的标志物。事实上,鉴定了能够检测血浆样品中前列腺癌组织的存在或不存在的标志物和/或标志物组(例如,具有选自max.chr3.193、HES5、SLC03A1和TPM4_8047的注释的染色体区域)(参见实施例I-VI)。

[0019] 如本文所述,所述技术提供了许多甲基化DNA标志物及其子集(例如,2、3、4、5、6、7或8个标志物的集合),其对前列腺癌总体上具有高度区分。实验将选择过滤器应用于候选标志物以鉴定提供高信噪比和低背景水平的标志物,从而提供用于前列腺癌筛查或诊断目的的高特异性。

[0020] 在一些实施方案中,技术涉及评估生物样品(例如,前列腺组织、血浆样品)中本文鉴定标志物中的一种或多种的存在和甲基化状态。这些标志物包含如本文所讨论,例如,如表3和5中所提供的一个或多个差异甲基化区域(DMR)。在所述技术的实施方案中评估甲基化状态。因此,本文提供的技术不限于通过其测量基因甲基化状态的方法。例如,在一些实施方案中,通过基因组扫描方法测量甲基化状态。例如,一种方法涉及限制性标志性基因组扫描(Kawai等人(1994)Mol.Cell.Biol.14:7421-7427),并且另一个实例涉及甲基化敏感性任意引物PCR(Gonzalgo等人(1997)Cancer Res.57:594-599)。在一些实施方案中,通过用甲基化敏感性限制性内切酶消化基因组DNA,随后对感兴趣区域进行Southern分析(消化-Southern方法),监测特定CpG位点处甲基化模式的变化。在一些实施方案中,分析甲基化模式的变化涉及基于PCR的过程,其涉及在PCR扩增之前用甲基化敏感性限制性内切酶消

化基因组DNA (Singer-Sam等人 (1990) *Nucl. Acids Res.* 18:687)。另外,已报道了利用DNA的亚硫酸氢盐处理作为甲基化分析的起点的其他技术。这些包括甲基化特异性PCR (MSP) (Herman等人 (1992) *Proc. Natl. Acad. Sci. USA* 93:9821-9826) 和从亚硫酸氢盐转化的DNA扩增的PCR产物的限制性内切酶消化 (Sadri和Hornsby (1996) *Nucl. Acids Res.* 24:5058-5059; 以及Xiong和Laird (1997) *Nucl. Acids Res.* 25:2532-2534)。已开发了PCR技术用于检测基因突变 (Kuppuswamy等人 (1991) *Proc. Natl. Acad. Sci. USA* 88:1143-1147) 和等位基因特异性表达的定量 (Szabo和Mann (1995) *Genes Dev.* 9:3097-3108; 以及Singer-Sam等人 (1992) *PCR Methods Appl.* 1:160-163)。此类技术使用内部引物,其与PCR产生的模板退火并立即终止待测定的单个核苷酸的5'。在一些实施方案中使用了使用如美国专利号7,037,650中所述的“定量Ms-SNuPE测定”的方法。

[0021] 在评估甲基化状态时,甲基化状态经常表示为相对于包含特定位点的样品中的总DNA群体,在所述特定位点处(例如,在单个核苷酸处、在特定区域或基因座处、在较长的感兴趣的序列(例如,长达约100-bp、200-bp、500-bp、1000-bp的DNA子序列或更长的序列)处)被甲基化的单个DNA链的分数或百分比。传统上,使用校准物通过PCR确定未甲基化核酸的量。然后,已知量的DNA经亚硫酸氢盐处理,并使用实时PCR或其他指数扩增(例如QuARTS测定)确定所得甲基化特异性序列(例如,如由美国专利号8,361,720;和美国专利申请公布号2012/0122088和2012/0122106所提供,其以引用的方式并入本文)。

[0022] 例如,在一些实施方案中,方法包括通过使用外部标准产生未甲基化靶标的标准曲线。标准曲线由至少两个点构成,并将未甲基化DNA的实时Ct值与已知的定量标准相关。然后,从至少两个点和外部标准构建甲基化靶标的第二标准曲线。此第二标准曲线将甲基化DNA的Ct与已知的定量标准相关。接下来,测定甲基化和未甲基化群体的测试样品Ct值,并且由通过前两个步骤产生的标准曲线计算DNA的基因组当量。由甲基化DNA的量相对于群体中DNA的总量计算感兴趣位点处的甲基化百分比,例如,(甲基化DNA数量)/(甲基化DNA数量+未甲基化DNA数量)×100。

[0023] 本文还提供了用于实施所述方法的组合物和试剂盒。例如,在一些实施方案中,单独或成组(例如,用于扩增多个标志物的引物对的组)提供对一种或多种标志物特异的试剂(例如,引物、探针)。还可提供用于进行检测测定的另外的试剂(例如,酶、缓冲液、用于进行QuARTS、PCR、测序、亚硫酸氢盐或其他测定的阳性和阴性对照)。在一些实施方案中,提供了含有一种或多种必要、足够或可用于进行方法的试剂的试剂盒。还提供了含有试剂的反应混合物。进一步提供了含有多种试剂的主混合试剂组,所述试剂可彼此添加和/或与测试样品一起添加以完成反应混合物。

[0024] 在一些实施方案中,将本文所述的技术与可编程机器相关联,所述可编程机器被设计为执行由本文所述的方法提供的一系列算术或逻辑运算。例如,将所述技术的一些实施方案与计算机软件和/或计算机硬件相关联(例如,在其中实现)。在一个方面,所述技术涉及一种计算机,其包括存储器形式、用于执行算术和逻辑运算的元件,以及用于执行一系列指令(例如,本文提供的方法)的处理元件(例如,微处理器)以读取、操作和存储数据。在一些实施方案中,微处理器是系统的一部分,所述系统用于确定(例如,如表3和表22中提供的一个或多个DMR,例如DMR 1-140)的甲基化状态;比较(例如,如表3和表22中提供的一个或多个DMR,例如DMR 1-140)的甲基化状态;生成标准曲线;确定Ct值;计算(例如,一个或多

个DMR,例如,(例如,如表3和表22中提供的一个或多个DMR,例如DMR 1-140)的甲基化的分数、频率或百分比;鉴定CpG岛;确定测定或标志物的特异性和/或灵敏度;计算ROC曲线和相关的AUC;序列分析;所有如本文所述或本领域已知的。

[0025] 在一些实施方案中,微处理器或计算机使用算法中的甲基化状态数据来预测癌症的部位。

[0026] 在一些实施方案中,软件或硬件组件接收多个测定的结果并确定单个值结果以基于多个测定的结果向用户报告指示癌症风险(例如,确定(例如,如表3和表5中提供的)多个DMR的甲基化状态)。相关实施方案基于来自多个测定的结果的数学组合(例如,加权组合、线性组合)计算风险因子,所述多个测定例如,确定多个标志物(如(例如,如表3和表5中提供的)多个DMR)的甲基化状态。在一些实施方案中,DMR的甲基化状态定义维度并可具有多维空间中的值,并且由多个DMR的甲基化状态定义的坐标是例如向用户报告,例如与癌症风险相关的结果。

[0027] 一些实施方案包括存储介质和存储器组件。存储器组件(例如,易失性和/或非易失性存储器)可用于存储指令(例如,如本文提供的方法的实施方案)和/或数据(例如,工件,如甲基化测量、序列和与之相关的统计描述)。一些实施方案涉及还包括CPU、图形卡和用户界面中的一个或多个的系统(例如,包括输出设备(如显示器)和输入设备(如键盘))。

[0028] 与所述技术相关的可编程机器包括常规现有技术和正在开发或尚待开发的技术(例如,量子计算机、化学计算机、DNA计算机、光学计算机、基于自旋电子学的计算机等)。

[0029] 在一些实施方案中,技术包括用于传输数据的有线(例如,金属电缆、光纤)或无线传输介质。例如,一些实施方案涉及通过网络(例如,局域网(LAN)、广域网(WAN)、ad-hoc网络、因特网等)的数据传输。在一些实施方案中,可编程机器作为对等体存在于这样的网络上,并且在一些实施方案中,可编程机器具有客户机/服务器关系。

[0030] 在一些实施方案中,数据存储在计算机可读存储介质(如硬盘、闪存、光学介质、软盘等)上。

[0031] 在一些实施方案中,本文提供的技术与多个可编程设备相关联,所述多个可编程设备协同操作以执行如本文所述的方法。例如,在一些实施方案中,(例如,通过网络连接的)多个计算机可例如,在通过常规网络接口(如以太网、光纤)或通过无线网络技术连接到网络(专用、公共或互联网)的集群计算或网格计算的实现中或依赖于完整计算机(具有板载CPU、存储、电源、网络接口等)的一些其他分布式计算机体系结构中并行工作以收集和处理数据。

[0032] 例如,一些实施方案提供了一种包括计算机可读介质的计算机。所述实施方案包括耦合到处理器的随机存取存储器(RAM)。处理器执行存储在存储器中的计算机可执行程序指令。此类处理器可包括微处理器、ASIC、状态机或其他处理器,并且可以是许多计算机处理器中的任一种,如来自加利福尼亚州圣克拉拉的英特尔公司(Intel Corporation of Santa Clara, California)和伊利诺斯州绍姆堡的摩托罗拉公司(Motorola Corporation of Schaumburg, Illinois)的处理器。此类处理器包括或可与介质(例如计算机可读介质)通信,所述介质存储指令,当由处理器执行时,所述指令使处理器执行本文所述的步骤。

[0033] 计算机可读介质的实施方案包括但不限于能够向处理器提供计算机可读指令的电子、光学、磁性或其他存储或传输设备。合适介质的其他实例包括但不限于软盘、CD-ROM、

DVD、磁盘、存储器芯片、ROM、RAM、ASIC、配置的处理器、所有光学介质、所有磁带或其他磁性介质,或计算机处理器可从其读取指令的任何其他介质。另外,各种其他形式的计算机可读介质可向计算机传输或携带指令,包括路由器、专用或公共网络,或有线和无线的其他传输设备或信道。指令可包括来自任何合适的计算机编程语言的代码,包括例如C、C++、C#、Visual Basic、Java、Python、Perl和JavaScript。

[0034] 计算机在一些实施方案中连接到网络。计算机还可包括许多外部或内部设备,如鼠标、CD-ROM、DVD、键盘、显示器或其他输入或输出设备。计算机的实例是个人计算机、数字助理、个人数字助理、蜂窝电话、移动电话、智能电话、寻呼机、数字平板电脑、膝上型计算机、因特网应用工具和其他基于处理器的设备。通常,与本文提供的技术方面相关的计算机可以是在任何操作系统上操作的任何类型的基于处理器的平台,如Microsoft Windows、Linux、UNIX、Mac OS X等,其能够支持包括本文提供的技术的一个或多个程序。一些实施方案包括执行其他应用程序(例如,应用程序)的个人计算机。应用程序可包含在存储器中,并且可包括例如文字处理应用程序、电子表格应用程序、电子邮件应用程序、即时消息应用程序、演示应用程序、因特网浏览器应用程序、日历/组织器应用程序以及能够由客户端设备执行的任何其他应用程序。

[0035] 与所述技术相关联的本文所述所有此类组件、计算机和系统可以是逻辑的或虚拟的。

[0036] 因此,本文提供的技术涉及一种在从受试者获得的样品中筛查前列腺癌的方法,所述方法包括测定从受试者获得的样品(例如,前列腺组织)(例如,血浆样品)中的标志物的甲基化状态,并且当标志物的甲基化状态不同于在未患有前列腺癌的受试者中测定的标志物的甲基化状态时,将受试者鉴定为患有前列腺癌,其中标志物包含选自由如表3和表22中提供的DMR1-140组成的组的差异甲基化区域(DMR)中的碱基。

[0037] 在其中从受试者获得的样品是前列腺组织的一些实施方案中,标志物选自ACOXL、AKR1B1_3644、ANXA2、CHST11_2206、FLJ45983、GAS6、GRASP、HAPLN3、HCG4P6、HES5_0822、ITPRIPL1、KCNK4、MAX.chr1.61519554-61519667、MAX.chr2.97193166-97193253、MAX.chr3.193、MAX.chr3.72788028-72788112、RAI1_7469、RASSF2、SERPINB9_3389、SLC4A11和TPM4_8047。

[0038] 在其中从受试者获得的样品是前列腺组织的一些实施方案中,标志物选自SERPINB9_3479、FL0T1_1665、HCG4P6_4618、CHST11_2206、MAX.chr12.485、GRASP_0932、GAS6_6425、MAX.chr3.193、MAX.chr2.971_3164、MAX.chr3.727_8028、HES5_0840、TPM4_8037、SLC03A1_6187、ITPRIPL1_1244、AKR1B1_3644、RASGRF2_6325、ZNF655_6075、PAMR1_7364、ST6GALNAC2_1113、CCN JL_9070、KCNB2_9128、IGFBP7_6412和WNT3A_5487。

[0039] 在其中从受试者获得的样品是血浆的一些实施方案中,标志物选自max.chr3.193、HES5、SLC03A1和TPM4_8047。

[0040] 所述技术涉及鉴定和区分前列腺癌。一些实施方案提供了包括测定多种标志物(例如,包括测定2至11至100或140种标志物)的方法。

[0041] 所述技术不限于评估的甲基化状态。在一些实施方案中,评估样品中标志物的甲基化状态包括确定一个碱基的甲基化状态。在一些实施方案中,测定样品中标志物的甲基化状态包括确定多个碱基处的甲基化程度。此外,在一些实施方案中,标志物的甲基化状态

包括标志物相对于标志物的正常甲基化状态的甲基化增加。在一些实施方案中,标志物的甲基化状态包括标志物相对于标志物的正常甲基化状态的甲基化降低。在一些实施方案中,标志物的甲基化状态包括标志物相对于标志物的正常甲基化状态的不同甲基化模式。

[0042] 此外,在一些实施方案中,标志物是具有100个或更少碱基的区域,标志物是具有500个或更少碱基的区域,标志物是具有1000个或更少碱基的区域,标志物是具有5000个或更少碱基的区域,或者在一些实施方案中,标志物是一个碱基。在一些实施方案中,标志物在高CpG密度启动子中。

[0043] 所述技术不受样品类型的限制。例如,在一些实施方案中,样品是粪便样品、组织样品(例如,前列腺组织样品)、血液样品(例如,血浆、血清、全血)、排泄物或尿液样品。

[0044] 此外,所述技术不限于用于确定甲基化状态的方法。在一些实施方案中,测定包括使用甲基化特异性聚合酶链反应、核酸测序、质谱、甲基化特异性核酸酶、基于质量的分离,或靶捕获。在一些实施方案中,测定包括使用甲基化特异性寡核苷酸。在一些实施方案中,技术使用大规模平行测序(例如,下一代测序)来确定甲基化状态,例如,通过合成测序、实时(例如,单分子)测序、珠乳液测序、纳米孔测序等。

[0045] 所述技术提供了用于检测DMR的试剂,例如,在一些实施方案中,提供了一组寡核苷酸,其包含由SEQ ID NO:1-146和/或SEQ ID NO:147-234提供的序列。在一些实施方案中,提供了一种寡核苷酸,其包含与具有DMR中的碱基的染色体区域互补的序列,例如,对DMR的甲基化状态敏感的寡核苷酸。

[0046] 所述技术提供了各种标志物组,例如,在一些实施方案中,标志物包括具有注释以及包含标志物的染色体区域,所述注释是ACOXL、AKR1B1_3644、ANXA2、CHST11_2206、FLJ45983、GAS6、GRASP、HAPLN3、HCG4P6、HES5_0822、ITPRIPL1、KCNK4、MAX.chr1.61519554-61519667、MAX.chr2.97193166-97193253、MAX.chr3.193、MAX.chr3.72788028-72788112、RAI1_7469、RASSF2、SERPINB9_3389、SLC4A11和TPM4_8047(参见实施例I-VI)。

[0047] 在一些实施方案中,标志物包括具有注释以及包含标志物的染色体区域,所述注释是SERPINB9_3479、FL0T1_1665、HCG4P6_4618、CHST11_2206、MAX.chr12.485、GRASP_0932、GAS6_6425、MAX.chr3.193、MAX.chr2.971_3164、MAX.chr3.727_8028、HES5_0840、TPM4_8037、SLC03A1_6187、ITPRIPL1_1244、AKR1B1_3644、RASGRF2_6325、ZNF655_6075、PAMR1_7364、ST6GALNAC2_1113、CCNL_9070、KCNB2_9128、IGFBP7_6412和WNT3A_5487(参见实施例VIII)。

[0048] 在其中获得的样品是血浆样品的一些实施方案中,标志物包括具有注释以及包含标志物的染色体区域,所述注释是max.chr3.193、HES5、SLC03A1和TPM4_8047。

[0049] 提供了试剂盒实施方案,例如,包含亚硫酸氢盐试剂;和包含来自DMR的序列的对照核酸的试剂盒,所述DMR选自由DMR1-140(来自表3或22)组成的组,并具有与未患有前列腺癌的受试者相关的甲基化状态。在一些实施方案中,试剂盒包含亚硫酸氢盐试剂和如本文所述的寡核苷酸。在一些实施方案中,试剂盒包含亚硫酸氢盐试剂;和包含来自DMR的序列的对照核酸,所述DMR选自由DMR 1-140(来自表3或22)组成的组,并具有与患有前列腺癌的受试者相关的甲基化状态。一些试剂盒实施方案包括用于从受试者获得样品(例如,粪便样品;前列腺组织样品;血浆样品)的样品收集器;用于从样品分离核酸的试剂;亚硫酸氢盐试剂;和如本文所述的寡核苷酸。

[0050] 所述技术涉及组合物(例如,反应混合物)的实施方案。在一些实施方案中,提供了包含含有DMR的核酸和亚硫酸氢盐试剂的组合物。一些实施方案提供了包含含有DMR的核酸和如本文所述的寡核苷酸的组合物。一些实施方案提供了包含含有DMR的核酸和甲基化敏感性限制性内切酶的组合物。一些实施方案提供了包含含有DMR的核酸和聚合酶的组合物。

[0051] 提供了另外的相关方法实施方案以用于筛查从受试者获得的样品(例如,前列腺组织样品;血浆样品;粪便样品)中的前列腺癌,例如一种方法,其包括确定样品中包含DMR中碱基的标志物的甲基化状态,所述DMR是(来自表3或表22的)DMR 1-140中的一种或多种;将来自受试者样品的标志物的甲基化状态与来自未患有前列腺癌的受试者的正常对照样品的标志物的甲基化状态进行比较;并且确定受试者样品和正常对照样品的甲基化状态差异的置信区间和/或p值。在一些实施方案中,置信区间是90%、95%、97.5%、98%、99%、99.5%、99.9%或99.99%并且p值是0.1、0.05、0.025、0.02、0.01、0.005、0.001或0.0001。方法的一些实施方案提供以下步骤:使包含DMR的核酸与亚硫酸氢盐试剂反应以产生亚硫酸氢盐反应的核酸;对亚硫酸氢盐反应的核酸进行测序,以提供亚硫酸氢盐反应的核酸的核苷酸序列;将亚硫酸氢盐反应的核酸的核苷酸序列与包含来自未患有前列腺癌的受试者的DMR的核酸的核苷酸序列进行比较,以鉴定两个序列中的差异;并且当存在差异时将受试者鉴定为患有前列腺癌。

[0052] 用于在从受试者获得的样品中筛查前列腺癌的系统由所述技术提供。系统的示例性实施方案包括,例如,用于筛查从受试者获得的样品(例如,前列腺组织样品;血浆样品;粪便样品)中的前列腺癌的系统,所述系统包括:分析组件,该分析组件被配置为确定样品的甲基化状态;软件组件,该软件组件被配置为将样品的甲基化状态与对照样品或记录在数据库中的参考样品甲基化状态进行比较;以及警报组件,该警报组件被配置为警告前列腺癌相关甲基化状态的用户。在一些实施方案汇中,通过软件组件确定警报,所述软件组件接收来自多个测定的结果(例如,确定多个标志物(例如,(例如,表3或表5中提供的)DMR)的甲基化状态)并基于多个结果计算待报告的值或结果。一些实施方案提供与本文提供的每个DMR相关联的加权参数的数据库,用于计算待向用户(例如,如医生、护士、临床医生等)报告的值或结果和/或警报。在一些实施方案中,报告了来自多个测定的所有结果,并且在一些实施方案中,使用一个或多个结果以提供基于来自说明受试者中癌症风险的多个测定的一个或多个结果的复合的得分、值或结果。

[0053] 在系统的一些实施方案中,样品包含含有DMR的核酸。在一些实施方案中,系统进一步包括用于分离核酸的组件、用于收集样品的组件,如用于收集粪便样品的组件。在一些实施方案中,系统包含含有DMR的核酸序列。在一些实施方案中,数据库包含来自未患有前列腺癌的受试者的核酸序列。还提供了核酸,例如,一组核酸,每种核酸具有包含DMR的序列。在一些实施方案中,组核酸,其中每种核酸具有来自未患有前列腺癌的受试者的序列。相关的系统实施方案包括如所述的一组核酸和与所述组核酸相关的核酸序列的数据库。一些实施方案进一步包括亚硫酸氢盐试剂。并且,一些实施方案进一步包括核酸测序仪。

[0054] 在某些实施方案中,提供了用于表征来自人患者的样品(例如,前列腺组织样品;血浆样品;粪便样品)的方法。例如,在一些实施方案中,此类实施方案包括从人患者的样品获得DNA;测定DNA甲基化标志物的甲基化状态,所述DNA甲基化标志物包含选自由来自表3或表22的DMR 1-140组成的组的差异甲基化区域(DMR)中的碱基;并且将一种或多种DNA甲

基化标志物的测定甲基化状态与未患有前列腺癌的人患者的一种或多种DNA甲基化标志物的甲基化水平参考值进行比较。

[0055] 此类方法不限于来自人患者的特定类型的样品。在一些实施方案中,样品是前列腺组织样品。在一些实施方案中,样品是血浆样品。在一些实施方案中,样品是粪便样品、组织样品、前列腺组织样品、血液样品或尿液样品。

[0056] 在一些实施方案中,此类方法包括测定多个DNA甲基化标志物。在一些实施方案中,此类方法包括测定2至11种DNA甲基化标志物。在一些实施方案中,此类方法包括测定12至140种DNA甲基化标志物。在一些实施方案中,此类方法包括测定样品中的一种或多种DNA甲基化标志物的甲基化状态,包括确定一个碱基的甲基化状态。在一些实施方案中,此类方法包括测定样品中的一种或多种DNA甲基化标志物的甲基化状态包括确定多个碱基处的甲基化程度。在一些实施方案中,此类方法包括测定正向链的甲基化状态或测定反向链的甲基化状态。

[0057] 在一些实施方案中,DNA甲基化标志物是具有100个或更少碱基的区域。在一些实施方案中,DNA甲基化标志物是具有500个或更少碱基的区域。在一些实施方案中,DNA甲基化标志物是具有1000个或更少碱基的区域。在一些实施方案中,DNA甲基化标志物是具有5000个或更少碱基的区域。在一些实施方案中,DNA甲基化标志物是一个碱基。在一些实施方案中,DNA甲基化标志物在高CpG密度启动子中。

[0058] 在一些实施方案中,测定包括使用甲基化特异性聚合酶链反应、核酸测序、质谱、甲基化特异性核酸酶、基于质量的分离,或靶捕获。

[0059] 在一些实施方案中,测定包括使用甲基化特异性寡核苷酸。在一些实施方案中,甲基化特异性寡核苷酸选自由SEQ ID NO:1-146和/或SEQ ID NO:147-234组成的组。

[0060] 在一些实施方案中,具有选自下组的注释的染色体区域包含DNA甲基化标志物,该组由以下组成:ACOXL、AKR1B1_3644、ANXA2、CHST11_2206、FLJ45983、GAS6、GRASP、HAPLN3、HCG4P6、HES5_0822、ITPRIPL1、KCNK4、MAX.chr1.61519554-61519667、MAX.chr2.97193166-97193253、MAX.chr3.193、MAX.chr3.72788028-72788112、RAI1_7469、RASSF2、SERPINB9_3389、SLC4A11和TPM4_8047。在一些实施方案中,DMR来自表5。

[0061] 在一些实施方案中,具有选自下组的注释的染色体区域包含DNA甲基化标志物,该组由以下组成:SERPINB9_3479、FL0T1_1665、HCG4P6_4618、CHST11_2206、MAX.chr12.485、GRASP_0932、GAS6_6425、MAX.chr3.193、MAX.chr2.971_3164、MAX.chr3.727_8028、HES5_0840、TPM4_8037、SLC03A1_6187、ITPRIPL1_1244、AKR1B1_3644、RASGRF2_6325、ZNF655_6075、PAMR1_7364、ST6GALNAC2_1113、CCN JL_9070、KCNB2_9128、IGFBP7_6412和WNT3A_5487。

[0062] 在其中获得的样品是血浆样品的一些实施方案中,标志物包括具有注释以及包含标志物的染色体区域,所述注释是max.chr3.193、HES5、SLC03A1和TPM4_8047。

[0063] 在一些实施方案中,此类方法包括确定两个DNA甲基化标志物的甲基化状态。在一些实施方案中,此类方法包括确定在表3或表5的一行中提供的一对DNA甲基化标志物的甲基化状态。

[0064] 在某些实施方案中,所述技术提供了用于表征从人患者获得的样品的方法。在一些实施方案中,此类方法包括确定样品中DNA甲基化标志物的甲基化状态,所述DNA甲基化

标志物包含选自由来自表3和表22的DMR 1-140组成的组的DMR中的碱基；将来自患者样品的DNA甲基化标志物的甲基化状态与来自未患有前列腺癌的人受试者的正常对照样品的DNA甲基化标志物的甲基化状态进行比较；并且确定人患者和正常对照样品的甲基化状态差异的置信区间和/或p值。在一些实施方案中，置信区间是90%、95%、97.5%、98%、99%、99.5%、99.9%或99.99%并且p值是0.1、0.05、0.025、0.02、0.01、0.005、0.001或0.0001。

[0065] 在某些实施方案中，所述技术提供了用于表征从人受试者获得的样品（例如，前列腺组织样品；血浆样品；粪便样品）的方法，所述方法包括使包含DMR的核酸与亚硫酸氢盐试剂反应以产生亚硫酸氢盐反应的核酸；对亚硫酸氢盐反应的核酸进行测序，以提供亚硫酸氢盐反应的核酸的核苷酸序列；将亚硫酸氢盐反应的核酸的核苷酸序列与包含来自未患有前列腺癌的受试者的DMR的核酸的核苷酸序列进行比较，以鉴定两个序列中的差异。

[0066] 在某些实施方案中，所述技术提供了用于表征从人受试者获得的样品（例如，前列腺组织样品；血浆样品；粪便样品）的系统，所述系统包括：分析组件，该分析组件被配置为确定样品的甲基化状态；软件组件，该软件组件被配置为将样品的甲基化状态与对照样品或记录在数据库中的参考样品甲基化状态进行比较；以及警报组件，该警报组件被配置为基于甲基化状态的组合确定单个值并警告前列腺癌相关甲基化状态的用户。在一些实施方案中，样品包含含有DMR的核酸。

[0067] 在一些实施方案中，此类系统进一步包括用于分离核酸的组件。在一些实施方案中，此类系统进一步包括用于收集样品的组件。

[0068] 在一些实施方案中，样品是粪便样品、组织样品、前列腺组织样品、血液样品或尿液样品。

[0069] 在一些实施方案中，数据库包含含有DMR的核酸序列。在一些实施方案中，数据库包含来自未患有前列腺癌的受试者的核酸序列。

[0070] 相关领域技术人员将基于本文中所含有的教义而明白另外实施方案。

具体实施方式

[0071] 本文提供了用于前列腺癌筛查的技术，并且具体地但非排他地涉及用于检测前列腺癌的存在的方法、组合物和相关用途。由于本文描述了所述技术，使用的章节标题仅用于组织的目的并且不应被理解为以任何方式限制主题。

[0072] 在各种实施方案的此详细描述中，出于解释的目的，阐述了许多具体细节以提供对所公开实施方案的透彻理解。然而，本领域技术人员将理解，可在具有或没有这些具体细节的情况下实践这些各种实施方案。在其他情况下，结构和设备以框图形式示出。此外，本领域技术人员可容易地理解，其中呈现和执行方法的特定序列是说明性的，并且预期所述序列可变化但仍保持在本文公开的各种实施方案的精神和范围内。

[0073] 本申请中引用的所有文献和类似材料，包括但不限于专利、专利申请、文章、书籍、论文和因特网网页，均出于任何目的明确地以引用的方式整体并入。除非另外定义，否则本文使用的所有技术和科学术语具有与本文所述各种实施方案所属领域的普通技术人员通常所理解的相同含义。当并入的参考文献中的术语的定义似乎与本教导中提供的定义不同时，应以本教导中提供的定义为准。

[0074] 定义

[0075] 为了有助于理解本技术,以下定义一些术语和短语。另外的定义在整个详细描述中阐述。

[0076] 在整个说明书和权利要求书中,除非上下文另有明确规定,否则以下术语采用本文明确相关的含义。如本文所用的短语“在一个实施方案中”不一定是指相同的实施方案,尽管其可以是。此外,如本文所用的短语“在另一个实施方案中”不一定是指不同的实施方案,尽管其可以是。因此,如下所述,在不脱离本发明的范围或精神的情况下,可容易地组合本发明的各种实施方案。

[0077] 另外,如本文所用,术语“或”是包含性的“或”操作符并且等同于术语“和/或”,除非上下文另有明确规定。除非上下文另有明确规定,否则术语“基于”是非排他性的并且允许基于未描述的另外因素。另外,在整个说明书中,“一个”、“一种”和“所述”的含义包括复数引用。“在……中”的含义包括“在……中”和“在……上”。

[0078] 如本文所用,“核酸”或“核酸分子”通常是指任何核糖核酸或脱氧核糖核酸,其可以是未修饰的或修饰的DNA或RNA。“核酸”包括但不限于单链和双链核酸。如本文所用,术语“核酸”还包括如上所述的含有一个或多个修饰的碱基的DNA。因此,具有为稳定性或其他原因而修饰的骨架的DNA是“核酸”。如本文所用的术语“核酸”包括核酸的此类化学、酶促或代谢修饰的形式,以及病毒和细胞(包括例如,简单和复杂细胞)所特有的DNA的化学形式。

[0079] 术语“寡核苷酸”或“多核苷酸”或“核苷酸”或“核酸”是指具有两个或更多个脱氧核糖核苷酸或核糖核苷酸的分子,优选多于三个,并且通常多于十个。确切的大小将取决于许多因素,而所述因素又取决于寡核苷酸的最终功能或用途。寡核苷酸可以任何方式产生,包括化学合成、DNA复制、逆转录或其组合。DNA的典型脱氧核糖核苷酸是胸腺嘧啶、腺嘌呤、胞嘧啶和鸟嘌呤。RNA的典型核糖核苷酸是尿嘧啶、腺嘌呤、胞嘧啶和鸟嘌呤。

[0080] 如本文所用,术语核酸的“基因座”或“区域”是指核酸的子区域,例如,染色体上的基因、单核苷酸、CpG岛等。

[0081] 术语“互补”和“互补性”是指通过碱基配对规则相关的核苷酸(例如,1个核苷酸)或多核苷酸(例如,核苷酸序列)。例如,序列5'-A-G-T-3'与序列3'-T-C-A-5'互补。互补性可以是“部分的”,其中根据碱基配对规则仅匹配核酸碱基中的一些。或者,核酸之间可存在“完全”或“总”互补性。核酸链之间的互补程度影响核酸链之间的杂交的效率和强度。这在依赖于核酸之间结合的扩增反应和检测方法中特别重要。

[0082] 术语“基因”是指核酸(例如,DNA或RNA)序列,其包含产生RNA或多肽或其前体所必需的编码序列。功能性多肽可由全长编码序列或由编码序列的任何部分编码,只要多肽的所希望活性或功能特性(例如,酶活性、配体结合、信号转导等)得到保留即可。当用于提及基因时,术语“部分”是指所述基因的片段。片段的大小范围可从几个核苷酸到整个基因序列减去一个核苷酸。因此,“包含基因的至少一部分的核苷酸”可包含所述基因的片段或整个基因。

[0083] 术语“基因”还涵盖结构基因的编码区域并包括与编码区域相邻位于5'和3'两个末端上,例如在任一个末端上达约1kb距离的序列,使得所述基因对应于全长mRNA的长度(例如,包括编码、调节、结构和其他序列)。位于编码区域的5'并存在于mRNA上的序列称为5'非翻译或未翻译序列。位于编码区域的3'或下游并存在于mRNA上的序列称为3'非翻译或3'未翻译序列。术语“基因”涵盖基因的cDNA和基因组两种形式。在一些生物体(例如,真核

生物)中,基因的基因组形式或克隆含有用被称为“内含子”或“插入区域”或“插入序列”的非编码序列间断的编码区域。内含子是转录到核RNA(hnRNA)中的基因片段;内含子可含有调控元件(如增强子)。内含子从核或原始转录物中去除或“剪除”;因此,内含子不存在于信使RNA(mRNA)转录物中。mRNA在翻译期间起作用以指定新生多肽中的氨基酸的序列或顺序。

[0084] 除了含有内含子之外,基因的基因组形式还可包括位于RNA转录物上存在的序列的5'和3'末端上的序列。这些序列被称为“侧接”序列或区域(这些侧接序列位于存在于mRNA转录物上的非翻译序列5'或3')。5'侧接区域可含有控制或影响基因转录的调控序列(如启动子和增强子)。3'侧接区域可含有引导转录终止、转录后裂解和多聚腺苷酸化的序列。

[0085] 当提及基因时,术语“野生型”是指具有从天然存在来源分离的基因特征的基因。当提及基因产物时,术语“野生型”是指具有从天然存在来源分离的基因产物特征的基因产物。应用于物体的术语“天然存在的”是指物体可在自然界中找到的事实。例如,可从自然中的来源分离并且不借助于人在实验室中有意修饰的存在于生物体(包括病毒)中的多肽或多核苷酸序列是天然存在的。野生型基因经常是在群体中最经常观察到的基因或等位基因,并且因此被任意指定为所述基因的“正常”或“野生型”形式。相反,当提及基因或基因产物时,术语“修饰的”或“突变体”分别是指当与野生型基因或基因产物相比时,展示序列和/或功能特性的修饰(即,改变特征)的基因或基因产物。应注意天然发生的突变体可予以分离;这些突变体通过以下事实来鉴定:当与野生型基因或基因产物相比时,其具有改变的特征。

[0086] 术语“等位基因”是指基因的变异;变异包括但不限于变体和突变体、多态性基因座和单核苷酸多态性基因座、移码和剪接突变。等位基因可在群体中天然存在,或者其可在群体的任何特定个体的生命期间出现。

[0087] 因此,当用于提及核苷酸序列时,术语“变体”和“突变体”是指与另一个通常相关的核苷酸序列相差一个或多个核苷酸的核酸序列。“变异”是两个不同核苷酸序列之间的差异;通常,一个序列是参考序列。

[0088] “扩增”是涉及模板特异性的核酸复制的特殊情况。其与非特定模板复制(例如,为模板依赖但不依赖于特定模板的复制)形成对比。模板特异性在此区别于复制的保真度(例如,适当的多核苷酸序列的合成)和核苷酸(核糖核酸或脱氧核糖核酸)特异性。模板特异性经常根据“靶”特异性来描述。靶序列是“靶标”,意思是它们被寻求从其他核酸中分选出来。扩增技术主要被设计用于这种分选。

[0089] 核酸的扩增通常是指多核苷酸的多个拷贝或多核苷酸的一部分的产生,通常从少量多核苷酸(例如,单个多核苷酸分子、多核苷酸分子的10至100个拷贝,其可或可不完全相同)开始,其中扩增产物或扩增子通常是可检测的。多核苷酸的扩增涵盖多种化学和酶促过程。在聚合酶链反应(PCR)或连接酶链反应(LCR;参见例如,美国专利号5,494,810;其以引用的方式整体并入本文)期间从靶标或模板DNA分子的一个或几个拷贝产生多个DNA拷贝是扩增的形式。另外类型的扩增包括但不限于等位基因特异性PCR(参见例如,美国专利号5,639,611;其以引用的方式整体并入本文);装配PCR(参见例如,美国专利号5,965,408;其以引用的方式整体并入本文)、解旋酶依赖性扩增(参见例如,美国专利号7,662,594;其以引用的方式整体并入本文)、热启动PCR(参见例如,美国专利号5,773,258和5,338,671;其各

自以引用的方式整体并入本文)、序列间特异性PCR、反向PCR(参见例如,Triglia等人(1988)Nucleic Acids Res.,16:8186;其以引用的方式整体并入本文)、连接介导的PCR(参见例如,Guilfoyle,R.等人,Nucleic Acids Research,25:1854-1858(1997);美国专利号5,508,169;其各自以引用的方式整体并入本文)、甲基化特异性PCR(参见例如,Herman等人,(1996)PNAS 93(13)9821-9826;其以引用的方式整体并入本文)、迷你引物PCR、多重连接依赖性探针扩增(参见例如,Schouten等人,(2002)Nucleic Acids Research 30(12):e57;其以引用的方式整体并入本文)、多重PCR(参见例如,Chamberlain等人,(1988)Nucleic Acids Research 16(23)11141-11156;Ballabio等人,(1990)Human Genetics 84(6)571-573;Hayden等人,(2008)BMC Genetics 9:80;其各自以引用的方式整体并入本文)、嵌套PCR、重叠-延伸PCR(参见例如,Higuchi等人,(1988)Nucleic Acids Research 16(15)7351-7367;其以引用的方式整体并入本文)、实时PCR(参见例如,Higuchi等人,(1992)Biotechnology 10:413-417;Higuchi等人,(1993)Biotechnology 11:1026-1030;其各自以引用的方式整体并入本文)、逆转录PCR(参见例如,Bustin,S.A.(2000)J.Molecular Endocrinology 25:169-193;其以引用的方式整体并入本文)、固相PCR、热不对称交错PCR和降落PCR(参见例如,Don等人,Nucleic Acids Research(1991)19(14)4008;Roux,K.(1994)Biotechniques 16(5)812-814;Hecker等人,(1996)Biotechniques 20(3)478-485;其各自以引用的方式整体并入本文)。多核苷酸扩增还可使用数字PCR完成(参见例如,Kalinina等人,Nucleic Acids Research.25;1999-2004,(1997);Vogelstein和Kinzler,Proc Natl Acad Sci USA.96;9236-41,(1999);国际专利公布号W005023091A2;美国专利申请公布号20070202525;其各自以引用的方式整体并入本文)。

[0090] 术语“聚合酶链反应”(“PCR”)是指K.B.Mullis美国专利号4,683,195、4,683,202和4,965,188的方法,其描述了一种用于在不进行克隆或纯化的情况下增加基因组DNA混合物中靶序列片段浓度的方法。这种用于扩增靶序列的方法包括将大量过量的两个寡核苷酸引物引入含有所希望靶序列的DNA混合物中,随后在DNA聚合酶存在下进行精确的热循环序列。所述两个引物与它们各自的双链靶序列的链互补。为了实现扩增,使混合物变性,并且然后使引物与靶分子内的其互补序列退火。在退火后,用聚合酶延伸引物,以便形成一对新的互补链。变性、引物退火和聚合酶延伸的步骤可重复多次(即,变性、退火和延伸构成一个“循环”;可存在许多“循环”)以获得高浓度所希望靶序列的扩增片段。所希望靶序列的扩增片段的长度由引物相对于彼此的相对位置确定,并且因此,此长度是可控参数。由于所述过程的重复方面,所述方法被称为“聚合酶链反应”(“PCR”)。因为靶序列的所希望扩增片段成为混合物中的主要序列(就浓度而言),因此它们被称为“PCR扩增的”并且是“PCR产物”或“扩增子”。

[0091] 通过选择酶在大多数扩增技术中实现模板特异性。扩增酶是在其使用的条件下将仅处理核酸的异质混合物中的特定核酸序列的酶。例如,在Q-β复制酶的情况下,MDV-1RNA是复制酶的特异性模板(Kacian等人,Proc.Natl.Acad.Sci.USA,69:3038[1972])。此扩增酶将不复制其他核酸。类似地,在T7 RNA聚合酶的情况下,此扩增酶对其自身的启动子具有严格的特异性(Chamberlin等人,Nature,228:227[1970])。在T4 DNA连接酶的情况下,所述酶将不连接两个寡核苷酸或多核苷酸,其中寡核苷酸或多核苷酸底物与连接接头处的模板之间存在错配(Wu和Wallace(1989)Genomics 4:560)。最后,发现热稳定模板依赖性DNA聚

合酶(例如,Taq和Pfu DNA聚合酶)凭借其在高温下发挥作用的能力,对由引物界定并因此限定的序列展示出高特异性;高温导致有利于引物与靶序列杂交并且不与非靶序列杂交的热力学条件(H.A.Erlich(编辑),PCR Technology,Stockton Press[1989])。

[0092] 如本文所用,术语“核酸检測测定”是指确定感兴趣的核酸的核苷酸组成任何方法。核酸检測测定包括但不限于DNA测序方法、探针杂交方法、结构特异性裂解测定(例如,INVADER测定,Hologic公司)并且描述于例如,美国专利号5,846,717、5,985,557、5,994,069、6,001,567、6,090,543和6,872,816;Lyamichev等人,Nat.Biotech.,17:292(1999),Hall等人,PNAS,USA,97:8272(2000)和US 2009/0253142中,其各自出于所有目的以引用的方式整体并入本文);酶错配裂解方法(例如,Variagenics,美国专利号6,110,684、5,958,692、5,851,770,其以引用的方式整体并入本文);聚合酶链反应;分枝杂交方法(例如,Chiron,美国专利号5,849,481、5,710,264、5,124,246和5,624,802,其以引用的方式整体并入本文);滚环复制(例如,美国专利号6,210,884、6,183,960和6,235,502,其以引用的方式整体并入本文);NASBA(例如,美国专利号5,409,818,其以引用的方式整体并入本文);分子信标技术(例如,美国专利号6,150,097,其以引用的方式整体并入本文);E-传感器技术(Motorola,美国专利号6,248,229、6,221,583、6,013,170和6,063,573,其以引用的方式整体并入本文);循环探针技术(例如,美国专利号5,403,711、5,011,769和5,660,988,其以引用的方式整体并入本文);Dade Behring信号放大方法(例如,美国专利号6,121,001、6,110,677、5,914,230、5,882,867和5,792,614,其以引用的方式整体并入本文);连接酶链反应(例如,Barnay Proc.Natl.Acad.Sci USA 88,189-93(1991));和夹心杂交方法(例如,美国专利号5,288,609,其以引用的方式整体并入本文)。

[0093] 术语“可扩增的核酸”是指可通过任何扩增方法扩增的核酸。预期“可扩增的核酸”通常将包含“样品模板”。

[0094] 术语“样品模板”是指源自分析“靶”(下文定义)存在的样品的核酸。相反,“背景模板”用于指代除了样品模板之外的核酸,其可或可不存在于样品中。背景模板最经常是无意的。这可能是携带的结果,或者其可能是由于试图从样品纯化的核酸污染物的存在。例如,来自除了待检测的那些生物体之外的生物体的核酸可作为测试样品中的背景存在。

[0095] 术语“引物”是指寡核苷酸,无论是在纯化的限制性消化物中天然存在还是合成产生的,当置于其中诱导了合成与核酸链互补的引物延伸产物的条件下(例如,在核苷酸和诱导剂(如DNA聚合酶)存在下并在合适的温度和pH下)时,其能够作为合成起始点起作用。引物优选地是单链的,以获得最大的扩增效率,但也可以是双链的。如果是双链的,则在用于制备延伸产物之前,首先处理引物以分离其链。优选地,引物是寡脱氧核糖核苷酸。引物必须足够长以在诱导剂存在下引发延伸产物的合成。引物的确切长度将取决于许多因素,包括温度、引物来源和方法的使用。

[0096] 术语“探针”是指寡核苷酸(例如,核苷酸序列),无论是在纯化的限制性消化物中天然存在还是合成、重组或通过PCR扩增产生的,其能够与另一种感兴趣的寡核苷酸杂交。探针可以是单链或双链的。探针可用于特定基因序列的检测、鉴定和分离(例如,“捕获探针”)。在一些实施方案中,预期本发明中使用的任何探针可用任何“报告分子”标记,使得在任何检测系统(包括但不限于酶(例如,ELISA,以及基于酶的组织化学测定)、荧光、放射性和发光系统)中均可检测到。并不旨在使本发明限于任何特定的检测系统或标记。

[0097] 如本文所用,“甲基化”是指胞嘧啶位置C5或N4处的胞嘧啶甲基化、腺嘌呤的N6位置或其他类型的核酸甲基化。体外扩增的DNA通常是未甲基化的,因为典型的体外DNA扩增方法不保留扩增模板的甲基化模式。然而,“未甲基化DNA”或“甲基化DNA”还可指扩增的DNA,其原始模板分别是未甲基化或甲基化的。

[0098] 因此,如本文所用,“甲基化核苷酸”或“甲基化核苷酸碱基”是指在核苷酸碱基上存在甲基部分,其中所述甲基部分不存在于公认的典型核苷酸碱基中。例如,胞嘧啶在其嘧啶环上不含甲基部分,但5-甲基胞嘧啶在其嘧啶环的位置5处含有甲基部分。因此,胞嘧啶不是甲基化核苷酸,并且5-甲基胞嘧啶是甲基化核苷酸。在另一个实例中,胸腺嘧啶在其嘧啶环的位置5处含有甲基部分;然而,出于本文的目的,当存在于DNA中时胸腺嘧啶不被认为是甲基化核苷酸,因为胸腺嘧啶是DNA的典型核苷酸碱基。

[0099] 如本文所用,“甲基化核酸分子”是指含有一个或多个甲基化核苷酸的核酸分子。

[0100] 如本文所用,核酸分子的“甲基化状态”、“甲基化分布”和“甲基化状况”是指核酸分子中存在不存在一个或多个甲基化核苷酸碱基。例如,含有甲基化胞嘧啶的核酸分子被认为是甲基化的(例如,核酸分子的甲基化状态被甲基化)。不含任何甲基化核苷酸的核酸分子被认为是未甲基化的。

[0101] 在提供或不提供甲基化发生的序列内位置的精确信息的情况下,特定核酸序列(例如,如本文所述的基因标志物或DNA区域)的甲基化状态可指示序列中每个碱基的甲基化状态或者可指示序列内(例如,一个或多个胞嘧啶的)碱基子集的甲基化状态,或者可指示关于序列内区域甲基化密度的信息。

[0102] 核酸分子中核苷酸基因座的甲基化状态是指核酸分子中特定基因座处甲基化核苷酸的存在或不存在。例如,当核酸分子中第7个核苷酸处存在的核苷酸是5-甲基胞嘧啶时,核酸分子中第7个核苷酸处的胞嘧啶的甲基化状态是甲基化的。类似地,当核酸分子中第7个核苷酸处存在的核苷酸是胞嘧啶(而不是5-甲基胞嘧啶)时,核酸分子中第7个核苷酸处的胞嘧啶的甲基化状态是未甲基化的。

[0103] 甲基化状况可任选地由(例如,表示甲基化频率、分数、比率、百分比等的)“甲基化值”表示或指示。例如,通过在用甲基化依赖性限制性内切酶限制消化后定量存在的完整核酸的量,或通过比较亚硫酸氢盐反应后的扩增曲线,或通过比较亚硫酸氢盐反应后的扩增曲线,或通过比较亚硫酸氢盐处理和未处理的核酸序列,可产生甲基化指。因此,值(例如甲基化值)表示甲基化状况,并且因此可用作跨基因座的多个拷贝的甲基化状况的定量指示。当希望将样品中序列的甲基化状况与阈值或参考值进行比较时,这是特别有用的。

[0104] 如本文所用,“甲基化频率”或“甲基化百分比(%)”是指相对于分子或基因座未甲基化的情况的数量,其中分子或基因座被甲基化的情况的数量。

[0105] 因此,甲基化状态描述了核酸(例如,基因组序列)的甲基化状态。另外,甲基化状态是指与甲基化相关的特定基因组基因座处的核酸片段的特征。此类特征包括但不限于此DNA序列中的任何胞嘧啶(C)残基是否被甲基化、甲基化C残基的位置、甲基化C在核酸的任何特定区域中的频率或百分比,以及由于例如等位基因起源的差异,甲基化中的等位基因差异。术语“甲基化状态”、“甲基化分布”和“甲基化状况”还指甲基化C或未甲基化C在整个生物样品中核酸的任何特定区域中的相对浓度、绝对浓度或模式。例如,如果核酸序列内的胞嘧啶(C)残基被甲基化,则其可被称为“高甲基化”或具有“增加的甲基化”,而如果DNA序

列内的胞嘧啶 (C) 残基未甲基化, 则其可被称为“低甲基化”或具有“减少的甲基化”。同样, 如果与另一个核酸序列(例如, 来自不同区域或来自不同个体等)相比, 核酸序列内的胞嘧啶 (C) 残基被甲基化, 则所述序列与所述另一个核酸序列相比被认为是高甲基化的或具有增加的甲基化。可替代地, 如果与另一个核酸序列(例如, 来自不同区域或来自不同个体等)相比, DNA序列内的胞嘧啶 (C) 残基未甲基化, 则所述序列与所述另一个核酸序列相比被认为是低甲基化的或具有减少的甲基化。另外, 如本文所用的术语“甲基化模式”是指核酸区域上甲基化和未甲基化核苷酸的集体位点。当甲基化和未甲基化核苷酸的数量在整个区域中相同或相似但甲基化和未甲基化核苷酸的位置不同时, 两种核酸可具有相同或相似的甲基化频率或甲基化百分比但具有不同的甲基化模式。当序列在甲基化程度(例如, 相对于另一个具有增加或减少的甲基化)、频率或模式方面不同时, 它们被称为“差异甲基化”或具有“甲基化差异”或具有“不同甲基化状态”。术语“差异甲基化”是指与癌症阴性样品中核酸甲基化的水平或模式相比, 癌症阳性样品中核酸甲基化的水平或模式的差异。其还可指在手术后癌症复发的患者相比于无复发的患者之间的水平或模式的差异。例如, 一旦定义了正确的截止或预测特征, 则差异甲基化和DNA甲基化的特定水平或模式是预后和预测性生物标志物。

[0106] 甲基化状态频率可用于描述个体群体或来自单个个体的样品。例如, 甲基化状态频率为50%的核苷酸基因座以50%的情况被甲基化, 并且以50%的情况未甲基化。例如, 这种频率可用于描述个体群体或核酸集合中核苷酸基因座或核酸区域被甲基化的程度。因此, 当第一群体或核酸分子池中的甲基化不同于第二群体或核酸分子池中的甲基化时, 第一群体或池的甲基化状态频率将不同于第二种群体或池的甲基化状态频率。例如, 这种频率还可用于描述单个个体中核苷酸基因座或核酸区域被甲基化的程度。例如, 这种频率可用于描述来自组织样品的一组细胞在核苷酸基因座或核酸区域处被甲基化或未甲基化的程度。

[0107] 如本文所用, “核苷酸基因座”是指核苷酸在核酸分子中的位置。甲基化核苷酸的核苷酸基因座是指甲基化核苷酸在核酸分子中的位置。

[0108] 通常, 人DNA的甲基化发生在包括相邻鸟嘌呤和胞嘧啶的二核苷酸序列上, 其中胞嘧啶位于鸟嘌呤的5'(也称为CpG二核苷酸序列)。CpG二核苷酸内的大多数胞嘧啶在人基因组中被甲基化, 然而一些在特定的富含CpG二核苷酸的基因组区域(称为CpG岛)中保持未甲基化(参见例如, Antequera等人(1990)Ce11 62:503-514)。

[0109] 如本文所用, “CpG岛”是指基因组DNA中相对于总基因组DNA含有增加数量的CpG二核苷酸的富含G:C的区域。CpG岛的长度可以是至少100、200或更多个碱基对, 其中所述区域的G:C含量是至少50%, 并且观察到的CpG频率与期望频率的比率是0.6; 在一些情况下, CpG岛的长度可以是至少500个碱基对, 其中所述区域的G:C含量是至少55%, 并且观察到的CpG频率与期望频率的比率是0.65。观察到的CpG频率与预期频率可根据Gardiner-Garden等人(1987)J.Mol.Biol.196:261-281中提供的方法计算。例如, 观察到的CpG频率与预期频率可根据公式 $R = (A \times B) / (C \times D)$ 计算, 其中R是观察到的CpG频率与预期频率的比率, A是分析序列中CpG二核苷酸的数量, B是分析序列中核苷酸的总数, C是分析序列中C核苷酸的总数, 并且D是分析序列中G核苷酸的总数。甲基化状态通常在例如启动子区域处的CpG岛中确定。应理解, 人基因组中的其他序列易于进行DNA甲基化, 如CpA和CpT(参见Ramsahoye(2000)

Proc. Natl. Acad. Sci. USA 97: 5237-5242; Salmon 和 Kaye (1970) Biochim. Biophys. Acta. 204: 340-351; Grafstrom (1985) Nucleic Acids Res. 13: 2827-2842; Nyce (1986) Nucleic Acids Res. 14: 4353-4367; Woodcock (1987) Biochem. Biophys. Res. Commun. 145: 888-894)。

[0110] 如本文所用,作为核酸分子的甲基化状态的函数修饰核酸分子的核苷酸的试剂或甲基化特异性试剂是指可以反映核酸分子的甲基化状态的方式改变核酸分子的核苷酸序列的化合物或组合物或其他试剂。用这种试剂处理核酸分子的方法可包括使核酸分子与试剂接触,如果希望,与另外的步骤结合,以完成核苷酸序列的所希望变化。核酸分子的核苷酸序列中的这种变化可导致其中每个甲基化核苷酸被修饰成不同核苷酸的核酸分子。核酸核苷酸序列中的这种变化可导致其中每个未甲基化核苷酸被修饰成不同核苷酸的核酸分子。核酸核苷酸序列中的这种变化可导致其中每个未甲基化的所选择核苷酸(例如,每个未甲基化胞嘧啶)被修饰成不同核苷酸的核酸分子。使用这种试剂来改变核酸核苷酸序列可导致其中为甲基化核苷酸的每个核苷酸(例如,每个甲基化胞嘧啶)被修饰成不同核苷酸的核酸分子。如本文所用,使用修饰所选择核苷酸的试剂是指修饰核酸分子中四个通常存在的核苷酸中的一个核苷酸(DNA的C、G、T和A,以及RNA的C、G、U和A),使得所述试剂修饰一个核苷酸而不修饰其他三个核苷酸的试剂。在一个示例性实施方案中,这种试剂修饰未甲基化的所选择核苷酸以产生不同核苷酸。在另一个示例性实施方案中,这种试剂可使未甲基化的胞嘧啶核苷酸脱氨基。示例性试剂是亚硫酸氢盐。

[0111] 在一些实施方案中,如本文所用,术语“亚硫酸氢盐试剂”是指包括亚硫酸氢盐、焦亚硫酸盐、重亚硫酸盐或其组合以例如在CpG二核苷酸序列中区分甲基化和未甲基化胞苷的试剂。

[0112] 术语“甲基化测定”是指用于确定核酸序列内的一种或多种CpG二核苷酸序列的甲基化状态的任何测定。

[0113] 术语“MS AP-PCR”(甲基化敏感性任意引物聚合酶链反应)是指本领域公认的技术,其允许使用富含CG的引物对基因组进行全局扫描以聚焦于最可能含有CpG二核苷酸的区域,并由Gonzalgo等人(1997)Cancer Research 57: 594-599描述。

[0114] 术语“MethyLightTM”是指由Eads等人(1999)Cancer Res. 59: 2302-2306描述的本领域公认的基于荧光的实时PCR技术。

[0115] 术语“HeavyMethylTM”是指其中覆盖扩增引物之间的或被扩增引物覆盖的CpG位置的甲基化特异性阻断探针(在本文中也称为阻断剂)能够进行核酸样品的甲基化特异性选择性扩增的测定。

[0116] 术语“HeavyMethylTMMethyLightTM”测定是指HeavyMethylTMMethyLightTM测定,其是MethyLightTM测定的变型,其中MethyLightTM测定与覆盖扩增引物之间的CpG位置的甲基化特异性阻断探针组合。

[0117] 术语“Ms-SNuPE”(甲基化敏感性单核苷酸引物延伸)是指由Gonzalgo和Jones(1997)Nucleic Acids Res. 25: 2529-2531描述的本领域公认的测定。

[0118] 术语“MSP”(甲基化特异性PCR)是指由Hermann等人(1996)Proc. Natl. Acad. Sci. USA 93: 9821-9826和由美国专利号5,786,146描述的本领域公认的甲基化测定。

[0119] 术语“COBRA”(组合的亚硫酸氢盐限制性分析)是指由Xiong和Laird(1997) Nucleic Acids Res. 25:2532-2534描述的本领域公认的甲基化测定。

[0120] 术语“MCA”(甲基化CpG岛扩增)是指由Toyota等人(1999) Cancer Res. 59:2307-12 并且在WO 00/26401A1中描述的甲基化测定。

[0121] 如本文所用,“所选择核苷酸”是指核酸分子中四个通常存在的核苷酸中的一个核苷酸(DNA的C、G、T和A,以及RNA的C、G、U和A),并且可包括通常存在的核苷酸的甲基化衍生物(例如,当C是所选择核苷酸时,甲基化和未甲基化的C两者均包括在所选择核苷酸的含义内),而甲基化的所选择核苷酸特异性地是指甲基化的通常存在的核苷酸,并且未甲基化的所选择核苷酸特异性地是指未甲基化的通常存在的核苷酸。

[0122] 术语“甲基化特异性限制性内切酶”或“甲基化敏感性限制性内切酶”是指依赖于其识别位点的甲基化状态选择性消化核酸的酶。在如果识别位点未甲基化或半甲基化则特异性地切割的限制性内切酶的情况下,如果识别位点被甲基化,则切割将不发生或将以显著降低的效率发生。在如果识别位点被甲基化则特异性地切割的限制性内切酶的情况下,如果识别位点未甲基化,则切割将不发生或将以显著降低的效率发生。优选的是甲基化特异性限制性内切酶,其识别序列含有CG二核苷酸(例如识别序列,如CGCG或CCCGGG)。对于一些实施方案,进一步优选的是如果此二核苷酸中的胞嘧啶在碳原子C5处甲基化则不切割的限制性内切酶。

[0123] 如本文所用,“不同核苷酸”是指与所选择核苷酸在化学上不同的核苷酸,通常使得不同核苷酸具有与所选择核苷酸不同的Watson-Crick碱基配对特性,由此通常存在的与所选择核苷酸互补的核苷酸与通常存在的与不同核苷酸互补的核苷酸是不同的。例如,当C是所选择核苷酸时,U或T可以是不同核苷酸,其例如是C与G的互补性和U或T与A的互补性。如本文所用,与所选择核苷酸互补或与不同核苷酸互补的核苷酸是指在高严格条件下,与四种通常存在的核苷酸中的三种互补核苷酸的碱基配对相比,以更高亲和力与所选择核苷酸或不同核苷酸碱基配对的核苷酸。互补性的实例是DNA(例如,A-T和C-G)和RNA(例如,A-U和C-G)中的Watson-Crick碱基配对。因此,例如,在高严格条件下,G碱基对C比G碱基对G、A或T具有更高的亲和力,并且因此,当C是所选择核苷酸时,G是与所选择核苷酸互补的核苷酸。

[0124] 如本文所用,给定标志物的“灵敏度”是指报告DNA甲基化值高于区分肿瘤和非肿瘤样品的阈值的样品的百分比。在一些实施方案中,阳性被定义为组织学证实的瘤形成,其报告DNA甲基化值高于阈值(例如,与疾病相关的范围),并且假阴性被定义为组织学证实的瘤形成,其报告DNA甲基化值低于阈值(例如,与无疾病相关的范围)。因此,灵敏度值反映了从已知的患病样品获得的给定标志物的DNA甲基化测量将在疾病相关测量范围内的概率。如这里所定义,计算的灵敏度值的临床相关性表示当应用于患有所述病状的受试者时给定标志物将检测到临床病状存在的概率的估计。

[0125] 如本文所用,给定标志物的“特异性”是指报告DNA甲基化值低于区分肿瘤和非肿瘤样品的阈值的非肿瘤样品的百分比。在一些实施方案中,阴性被定义为组织学证实的非肿瘤样品,其报告DNA甲基化值低于阈值(例如,与无疾病相关的范围),并且假阳性被定义为组织学证实的非肿瘤样品,其报告DNA甲基化值高于阈值(例如,与疾病相关的范围)。因此,特异性值反映了从已知的非肿瘤样品获得的给定标志物的DNA甲基化测量将在无疾病

相关测量范围内的概率。如这里所定义,计算的特异性值的临床相关性表示当应用于不具有所述病状的患者时给定标志物将检测到临床病状不存在的概率的估计。

[0126] 如本文所用的术语“AUC”是“曲线下面积”的缩写。具体地,其是指接受者操作特征(ROC)曲线下面积。ROC曲线是诊断测试的不同可能切割点的真阳性率相比于假阳性率的图。其示出了取决于所选择切割点的灵敏度与特异性之间的权衡(灵敏度的任何增加将伴随着特异性的降低)。ROC曲线下面积(AUC)是诊断测试准确度的度量(面积越大越好;最佳值是1;随机测试将具有位于对角线上的ROC曲线,面积为0.5;供参考:J.P. Egan. (1975) *Signal Detection Theory and ROC Analysis*, Academic Press, New York)。

[0127] 如本文所用,术语“肿瘤”是指“组织的异常质量,其生长超过并且与正常组织的生长不协调”参见例如,Willis RA, “The Spread of Tumors in the Human Body”, London, Butterworth&Co, 1952。

[0128] 如本文所用,术语“腺瘤”是指腺源性的良性肿瘤。虽然这些生长是良性的,但随着时间的推移它们可能会变得恶化。

[0129] 术语“癌前期”或“肿瘤前期”及其等同物是指正在经历恶性转化的任何细胞增殖性病症。

[0130] 肿瘤、腺瘤、癌症等的“部位”是其中肿瘤、腺瘤、癌症等所在的受试者体内的组织、器官、细胞类型、解剖区域、身体部位等。

[0131] 如本文所用,“诊断”测试应用包括检测或鉴定受试者的疾病状态或状况,确定受试者将收缩给定疾病或病状的可能性,确定患有疾病或病状的受试者将响应于治疗的可能性,确定患有疾病或病状的受试者的预后(或其可能的进展或退化),并确定治疗对患有疾病或病状的受试者的影响。例如,诊断可用于检测收缩肿瘤的受试者的存在或可能性,或者这样的受试者将有利地响应于化合物(例如医药,例如药物)或其他治疗的可能性。

[0132] 如本文所用,术语“标志物”是指例如基于其甲基化状态,能够通过区分癌细胞与正常细胞来诊断癌症的物质(例如,核酸或核酸区域)。

[0133] 当与核酸相关使用时,术语“分离的”,如“分离的寡核苷酸”,是指从与其天然来源通常相关的至少一种污染核酸鉴定和分离的核酸序列。分离的核酸以与其天然存在的形式或设置不同的形式或设置存在。相反,非分离的核酸(如DNA和RNA)以它们存在于自然界中的状态被发现。非分离的核酸的实例包括:在邻近基因附近的宿主细胞染色体上发现的给定DNA序列(例如,基因);RNA序列,如编码特定蛋白质的特定mRNA序列,其在细胞中发现为与编码多种蛋白质的许多其他mRNA的混合物。然而,编码特定蛋白质的分离的核酸包括,例如,通常表达所述蛋白质的细胞中的这种核酸,其中所述核酸处于与天然细胞不同的染色体位置中,或者另外侧接与自然界中发现的核酸序列不同的核酸序列。分离的核酸或寡核苷酸可以单链或双链形式存在。当利用分离的核酸或寡核苷酸表达蛋白质时,寡核苷酸将以最低限度含有有义链或编码链(即,寡核苷酸可以是单链的),但可含有有义链和反义链两者(即,寡核苷酸可以是双链的)。分离的核酸在从其天然或典型环境中分离后,可与其他核酸或分子组合。例如,分离的核酸可存在于宿主细胞中,其中已将其置于其中,例如用于异源表达。

[0134] 术语“纯化的”是指从其天然环境中去除、分离或分开的核酸或氨基酸序列的分子。因此,“分离的核酸序列”可以是纯化的核酸序列。“基本上纯化的”分子至少60%不含、

优选至少75%不含并且更优选至少90%不含与其天然相关联的其他组分。如本文所用,术语“纯化的”或“进行纯化”还指从样品中去除污染物。去除污染蛋白质导致样品中感兴趣的多肽或核酸的百分比增加。在另一个实例中,重组多肽在植物、细菌、酵母或哺乳动物宿主细胞中表达,并且通过去除宿主细胞蛋白纯化多肽;由此,样品中重组多肽的百分比增加。

[0135] 术语“包含给定多核苷酸序列或多肽的组合物”泛指含有给定多核苷酸序列或多肽的任何组合物。所述组合物可包含含有盐(例如NaCl)、洗涤剂(例如SDS)和其他组分(例如,Denhardt溶液、干乳、鲑鱼精DNA等)的水性溶液。

[0136] 术语“样品”使用的是其最广泛的意义。在某种意义上,其可指动物细胞或组织。在另一种意义上,其意指包括从任何来源获得的样本或培养物,以及生物和环境样品。生物样品可从植物或动物(包括人)获得,并且涵盖流体、固体、组织和气体。在一些实施方案中,样品是血浆样品。在一些实施方案中,样品是前列腺组织样品。在一些实施方案中,样品是粪便样品。环境样品包括环境材料,如表面物质、土壤、水和工业样品。这些实施例不能解释为限制适用于本发明的所述样品类型。

[0137] 如本文所用,在一些情况下使用的“远程样品”涉及从不是样品的细胞、组织或器官来源的部位间接收集的样品。例如,当在粪便样品(例如,不是直接取自前列腺的样品)中评估源自胰腺的样品材料时,所述样品是远程样品。

[0138] 如本文所用,术语“患者”或“受试者”是指有待经受由所述技术提供的各种测试的生物体。术语“受试者”包括动物,优选哺乳动物,包括人。在一个优选实施方案中,受试者是灵长类动物。在甚至更优选的实施方案中,受试者是人。

[0139] 如本文所用,术语“试剂盒”是指输送材料的任何输送系统。在反应测定的情况下,此类输送系统包括允许存储、运输或将反应试剂(例如,适当容器中的寡核苷酸、酶等)和/或支持材料(例如,缓冲剂、执行测定的书面指令等)从一个位置输送至另一个位置的系统。例如,试剂盒包括一个或多个套罩(例如,盒),其含有相关反应试剂和/或支持材料。如本文所用,术语“分段试剂盒”是指包括两个或更多个单独容器的输送系统,每一容器含有总试剂盒组分的子部分。这些容器可共同或单独输送至预定接受者。例如,第一容器可含有用于测定的酶,而第二容器含有寡核苷酸。术语“零散试剂盒”旨在涵盖含有根据联邦食品、药物和化妆品法(Federal Food, Drug, and Cosmetic Act)的部分520(e)来管理的分析物特异性试剂(ASR)的试剂盒,但不限于此。事实上,包含各自含有全部试剂盒组分的子部分的两个或更多个单独容器的任何输送系统包括在术语“零散试剂盒”中。相比之下,“组合试剂盒”是指将反应测定的所有组分包含于单一容器(例如,容纳每个所需组分的单一盒)的输送系统。术语“试剂盒”包括分段试剂盒和组合试剂盒两者。

[0140] 所述技术的实施方案

[0141] 本文提供了用于前列腺癌筛查的技术,并且具体地但非排他地涉及用于检测前列腺癌的存在的方法、组合物和相关用途。

[0142] 事实上,如实施例I-VI中所述,在鉴定本发明实施方案的过程中进行的实验鉴定了一组新的73种差异甲基化区域(DMR),用于区分前列腺来源的DNA与非肿瘤性对照DNA的癌症。另外,鉴定了10种新的DMR,其在前列腺上皮细胞(癌症和正常)中但不在正常白细胞DNA样品中甲基化。从CpG富集的亚硫酸氢盐转化的肿瘤和正常DNA的下一代测序研究鉴定了两组区域。肿瘤样品包括较少侵袭性的格里森6和更具侵袭性的格里森7+模式。使用专有

过滤器和分析管道选择DMR，并使用新的甲基化特异性PCR (MSP) 测定法在独立的组织样品组中进行验证。这73种生物标志物测定证明在组织中具有优异的检测能力，并具有广泛的临床特异性-一些用于许多不同器官部位的癌症，另一些仅对前列腺癌特异。

[0143] 此类实验列出并描述了区分前列腺癌组织和良性前列腺组织的120种新的DNA甲基化标志物(表3)。从这120种新的DNA甲基化标志物，进一步的实验鉴定了73种能够区分侵袭性前列腺癌组织(例如，格里森评分7+)和良性前列腺组织的标志物。更具体地，鉴定了能够区分前列腺癌组织和良性前列腺组织的标志物和/或标志物组(例如，具有选自以下的注释的染色体区域:ACOXL、AKR1B1_3644、ANXA2、CHST11_2206、FLJ45983、GAS6、GRASP、HAPLN3、HCG4P6、HES5_0822、ITPRIPL1、KCNK4、MAX.chr1.61519554-61519667、MAX.chr2.97193166-97193253、MAX.chr3.193、MAX.chr3.72788028-72788112、RAI1_7469、RASSF2、SERPINB9_3389、SLC4A11和TPM4_8047) (参见实施例I-VI)。

[0144] 在开发本发明实施方案的过程中进行的另外的实验涉及鉴定能够区分前列腺癌组织和良性前列腺组织的标志物(例如，具有选自以下的注释的染色体区域:SERPINB9_3479、FL0T1_1665、HCG4P6_4618、CHST11_2206、MAX.chr12.485、GRASP_0932、GAS6_6425、MAX.chr3.193、MAX.chr2.971_3164、MAX.chr3.727_8028、HES5_0840、TPM4_8037、SLC03A1_6187、ITPRIPL1_1244、AKR1B1_3644、RASGRF2_6325、ZNF655_6075、PAMR1_7364、ST6GALNAC2_1113、CCN JL_9070、KCNB2_9128、IGFBP7_6412和WNT3A_5487)能够区分前列腺癌组织与良性前列腺组织(参见实施例VIII;表20)。

[0145] 在开发本发明实施方案的过程中进行的另外的实验涉及鉴定能够区分侵袭性前列腺癌组织(例如，格里森评分7+)和较少侵袭性前列腺癌组织(例如，格里森评分6)的标志物(例如，具有选自以下的注释的染色体区域:SERPINB9_3479、GRASP_0932、SLC03A1_6187、ITPRIPL1_1244、AKR1B1_3644、RASGRF2_6325、ZNF655_6075、PAMR1_7364、ST6GALNAC2_1113、CCN JL_9070、KCNB2_9128、IGFBP7_6412和WNT3A_5487)能够区分前列腺癌组织与良性前列腺组织(参见实施例VIII;表20)。

[0146] 在开发本发明实施方案的过程中进行的另外的实验涉及鉴定能够检测血液样品(例如，血浆样品)中前列腺癌的存在或不存在的标志物。事实上，鉴定了能够检测血浆样品中前列腺癌组织的存在或不存在的标志物和/或标志物组(例如，具有选自max.chr3.193、HES5、SLC03A1和TPM4_8047的注释的染色体区域) (参见实施例I-VI)。

[0147] 尽管本文的公开内容涉及某些示出的实施方案，但应理解，这些实施方案是以示例的方式而非限制的方式呈现的。

[0148] 在具体方面，本技术提供了用于鉴定、确定和/或分类癌症(如前列腺癌)的组合物和方法。所述方法包括确定从受试者分离的生物样品(例如，粪便样品、前列腺组织样品、血浆样品)中的至少一种甲基化标志物的甲基化状况，其中标志物的甲基化状态的变化指示前列腺癌的存在、类别或部位。具体实施方案涉及包含用于诊断(例如筛查)前列腺癌的差异甲基化区域(DMR，例如DMR 1-140，参见表3和表22)的标志物。

[0149] 除了其中分析了包含本文提供并在表3或表5中列出的DMR(例如，DMR，例如，DMR 1-140)的至少一种标志物、标志物的区域或标志物的碱基的甲基化分析的实施方案之外，所述技术还提供了标志物组，其包含含有具有用于检测癌症(特别是前列腺癌)的效用的DMR的至少一种标志物、标志物的区域或标志物的碱基。

[0150] 所述技术的一些实施方案基于对包含DMR的至少一种标志物、标志物的区域或标志物的碱基的CpG甲基化状况的分析。

[0151] 在一些实施方案中,本技术提供了亚硫酸氢盐技术与一种或多种甲基化测定组合用于确定包含DMR(例如,DMR 1-140,参见表3和表22)的至少一种标志物内CpG二核苷酸序列的甲基化状况。基因组CpG二核苷酸可以是甲基化或未甲基化的(可替代地分别称为上甲基化和下甲基化的)。然而,本发明的方法适用于在远程样品(例如,血液、器官流出物或粪便)的背景下分析具有异质性质的生物样品,例如,低浓度的肿瘤细胞或其生物材料。因此,当分析这种样品内CpG位置的甲基化状况时,可使用定量测定来确定特定CpG位置处甲基化的水平(例如,百分比、分数、比率、比例或程度)。

[0152] 根据本技术,在包含DMR的标志物中确定CpG二核苷酸序列的甲基化状况在诊断和表征癌症(如前列腺癌)两个方面均具有实用性。

[0153] 标志物的组合

[0154] 在一些实施方案中,技术涉及评估包含来自表3、表5或表22的DMR(例如,DMR号1-140)的标志物组合的甲基化状态。在一些实施方案中,评估多于一个标志物的甲基化状态增加了筛查或诊断以用于鉴定受试者中的肿瘤(例如,前列腺癌)的特异性和/或灵敏度。

[0155] 例如,如通过与预测的特异性和灵敏度相关的统计技术所鉴定,通过标志物的各种组合预测各种癌症。所述技术提供了用于鉴定一些癌症的预测组合和验证的预测组合的方法。

[0156] 用于测定甲基化状态的方法

[0157] 用于分析核酸中5-甲基胞嘧啶存在的最常用方法是基于由Frommer等人描述的用于检测DNA中5-甲基胞嘧啶的亚硫酸氢盐方法(Frommer等人(1992)Proc.Natl.Acad.Sci.USA 89:1827-31,其出于所有目的明确地以引用的方式整体并入本文)或其变型。映射5-甲基胞嘧啶的亚硫酸氢盐方法基于观察到胞嘧啶而非5-甲基胞嘧啶与亚硫酸氢根离子(也称为亚硫酸氢盐)反应。所述反应通常根据以下步骤进行:首先,胞嘧啶与亚硫酸氢盐反应以形成碘化胞嘧啶。接下来,碘化反应中间体的自发脱氨基作用产生碘化尿嘧啶。最后,碘化尿酸在碱性条件下脱碘化以形成尿嘧啶。检测是可能的,因为尿嘧啶与腺嘌呤形成碱基对(因此表现得像胸腺嘧啶),而5-甲基胞嘧啶与鸟嘌呤形成碱基对(因此表现得像胞嘧啶)。这使得通过例如,亚硫酸氢盐基因组测序(Grigg G和Clark S, Bioessays (1994) 16:431-36;Grigg G,DNA Seq. (1996) 6:189-98)或如例如美国专利号5,786,146中公开的甲基化特异性PCR(MSP),甲基化胞嘧啶与非甲基化胞嘧啶的区分成为可能。

[0158] 一些常规技术涉及包括将待分析的DNA封闭在琼脂糖基质中,从而防止DNA的扩散和复性(亚硫酸氢盐仅与单链DNA反应),并用快速透析代替沉淀和纯化步骤的方法(Olek A等人(1996)“A modified and improved method for bisulfite based cytosine methylation analysis”Nucleic Acids Res.24:5064-6)。因此可能的是分析单个细胞的甲基化状况,从而说明所述方法的实用性和灵敏度。用于检测5-甲基胞嘧啶的常规方法的概述由Rein,T.等人(1998)Nucleic Acids Res.26:2255提供。

[0159] 亚硫酸氢盐技术通常涉及在亚硫酸氢盐处理后扩增已知核酸的短的特异性片段,然后通过测序(Olek和Walter(1997)Nat.Genet.17:275-6)或引物延伸反应(Gonzalgo和

Jones (1997) Nucleic Acids Res. 25:2529-31; WO 95/00669; 美国专利号6,251,594) 测定产物以分析单个胞嘧啶位置。一些方法使用酶消化 (Xiong 和 Laird (1997) Nucleic Acids Res. 25:2532-4)。本领域中还描述了通过杂交进行的检测 (Olek 等人, WO 99/28498)。另外, 已描述了使用亚硫酸氢盐技术对个体基因进行甲基化检测 (Grigg 和 Clark (1994) Bioessays 16:431-6; Zeschinski 等人 (1997) Hum Mol Genet. 6:387-95; Feil 等人 (1994) Nucleic Acids Res. 22:695; Martin 等人 (1995) Gene 157:261-4; WO 9746705; WO 9515373)。

[0160] 各种甲基化测定程序是本领域已知的, 并且可与根据本技术的亚硫酸氢盐处理结合使用。这些测定允许确定核酸序列内一个或多个CpG二核苷酸 (例如, CpG岛) 的甲基化状态。除了其他技术之外, 此类测定涉及亚硫酸氢盐处理的核酸的测序, PCR (用于序列特异性扩增) 、Southern印迹分析和甲基化敏感性限制性内切酶的使用。

[0161] 例如, 通过使用亚硫酸氢盐处理, 简化了基因组测序以分析甲基化模式和5-甲基胞嘧啶分布 (Frommer 等人 (1992) Proc. Natl. Acad. Sci. USA 89:1827-1831)。另外, 从亚硫酸氢盐转化的DNA扩增的PCR产物的限制性内切酶消化可用于评估甲基化状态, 例如, 如由 Sadri 和 Hornsby (1997) Nucl. Acids Res. 24:5058-5059所述, 或如体现在称为COBRA (组合亚硫酸氢盐限制分析) 的方法中 (Xiong 和 Laird (1997) Nucleic Acids Res. 25:2532-2534)。

[0162] COBRATM分析是定量甲基化测定法, 其可用于确定少量基因组DNA中特定基因座处的DNA甲基化水平 (Xiong 和 Laird, Nucleic Acids Res. 25:2532-2534, 1997)。简而言之, 限制性内切酶消化用于揭示亚硫酸氢盐处理的DNA的PCR产物中甲基化依赖性序列差异。首先根据由Frommer 等人 (Proc. Natl. Acad. Sci. USA 89:1827-1831, 1992) 描述的程序通过标准亚硫酸氢盐处理将甲基化依赖性序列差异引入基因组DNA中。然后使用对感兴趣的CpG岛特异的引物进行亚硫酸氢盐转化的DNA的PCR扩增, 随后进行限制性内切核酸酶消化、凝胶电泳和使用特异性标记的杂交探针的检测。原始DNA样品中的甲基化水平由消化的和未消化的PCR产物的相对量以跨越广泛范围的DNA甲基化水平的线性定量方式表示。此外, 此技术可靠地应用于从显微切割的石蜡包埋的组织样品获得的DNA。

[0163] 用于COBRATM分析的典型试剂 (例如, 可在典型的基于COBRATM的试剂盒中找到) 可包括但不限于: 用于特定基因座的PCR引物 (例如, 特定基因、标志物、DMR、基因区域、标志物区域、亚硫酸氢盐处理的DNA序列、CpG岛等); 限制性内切酶和适当的缓冲液; 基因杂交寡核苷酸; 对照杂交寡核苷酸; 用于寡核苷酸探针的激酶标记试剂盒; 和标记的核苷酸。另外, 亚硫酸氢盐转化试剂可包括: DNA变性缓冲液; 磷酸缓冲液; DNA回收试剂或试剂盒 (例如, 沉淀、超滤、亲和柱); 脱磷酸缓冲液; 和DNA回收组分。

[0164] 优选地, 测定如“MethyLightTM” (基于荧光的实时PCR技术) (Eads 等人, Cancer Res. 59:2302-2306, 1999)、Ms-SNuPETM (甲基化敏感性单核苷酸引物延伸) 反应 (Gonzalgo 和 Jones, Nucleic Acids Res. 25:2529-2531, 1997)、甲基化特异性PCR (“MSP”; Herman 等人, Proc. Natl. Acad. Sci. USA 93:9821-9826, 1996; 美国专利号5,786,146) 和甲基化CpG岛扩增 (“MCA”; Toyota 等人, Cancer Res. 59:2307-12, 1999) 单独使用或与这些方法中的一种或多种组合使用。

[0165] “HeavyMethylTM”测定技术是基于亚硫酸氢盐处理的DNA的甲基化特异性扩增来评

估甲基化差异的定量方法。覆盖扩增引物之间的或被扩增引物覆盖的CpG位置的甲基化特异性阻断探针(阻断剂)能够进行核酸样品的甲基化特异性选择性扩增。

[0166] 术语“HeavyMethylTMMethyLightTM”测定是指HeavyMethylTMMethyLightTM测定,其是MethyLightTM测定的变型,其中MethyLightTM测定与覆盖扩增引物之间的CpG位置的甲基化特异性阻断探针组合。HeavyMethylTM测定还可与甲基化特异性扩增引物组合使用。

[0167] 用于HeavyMethylTM分析的典型试剂(例如,可在典型的基于MethyLightTM的试剂盒中找到)可包括但不限于:用于特定基因座的PCR引物(例如,特定基因、标志物、DMR、基因区域、标志物区域、亚硫酸氢盐处理的DNA序列、CpG岛或亚硫酸氢盐处理的DNA序列或CpG岛等);阻断寡核苷酸;优化的PCR缓冲液和脱氧核苷酸;和Taq聚合酶。

[0168] MSP(甲基化特异性PCR)允许评估CpG岛内几乎任何CpG位点组的甲基化状况,与甲基化敏感性限制性内切酶的使用无关(Herman等人Proc.Natl.Acad.Sci.USA 93:9821-9826,1996;美国专利号5,786,146)。简而言之,DNA被亚硫酸氢钠修饰,其将未甲基化而非甲基化的胞嘧啶转化为尿嘧啶,并且随后用对甲基化相对于未甲基化DNA特异的引物扩增产物。MSP仅需要少量DNA,对给定CpG岛基因座的0.1%甲基化等位基因敏感,并且可在从石蜡包埋的样品提取的DNA上进行。用于MSP分析的典型试剂(例如,可在典型的基于MSP的试剂盒中找到)可包括但不限于:用于特定基因座的甲基化和未甲基化PCR引物(例如,特定基因、标志物、DMR,基因区域、标志物区域、亚硫酸氢盐处理的DNA序列、CpG岛等);优化的PCR缓冲液和脱氧核苷酸,以及特异性探针。

[0169] MethyLightTM测定是一种高通量定量甲基化测定,其利用基于荧光的实时PCR(例如, **TaqMan**[®]),在PCR步骤后无需进一步操作(Eads等人,Cancer Res.59:2302-2306,1999)。简而言之,MethyLightTM过程开始于基因组DNA的混合样品,其在亚硫酸氢钠反应中根据标准程序转化为甲基化依赖性序列差异的混合池(亚硫酸氢盐过程将未甲基化胞嘧啶残基转化为尿嘧啶)。然后例如,使用重叠已知CpG二核苷酸的PCR引物在“偏向”反应中进行基于荧光的PCR。序列区分在扩增过程的水平和荧光检测过程的水平上发生。

[0170] MethyLightTM测定用作核酸(例如基因组DNA样品)中甲基化模式的定量测试,其中序列区分在探针杂交水平上发生。在定量版本中,PCR反应在荧光探针存在下提供甲基化特异性扩增,所述荧光探针重叠特定的推定甲基化位点。通过其中引物和探针均不覆盖任何CpG二核苷酸的反应提供对输入DNA量的无偏控制。可替代地,通过用不覆盖已知甲基化位点的对照寡核苷酸(例如,基于荧光的HeavyMethylTM和MSP技术版本)或用覆盖潜在甲基化位点的寡核苷酸探测有偏向的PCR池来实现基因组甲基化的定性测试。

[0171] MethyLightTM过程与任何合适的探针(例如“**TaqMan**[®]”探针、**Lightcycler**[®]探针等)一起使用。例如,在一些应用中,双链基因组DNA用亚硫酸氢钠处理,并经受使用**TaqMan**[®]探针(例如用MSP引物和/或重甲基阻断寡核苷酸和**TaqMan**[®]探针)的两组PCR反应之一。**TaqMan**[®]探针用荧光“报告基因”和“猝灭剂”分子双重标记,并且被设计为对相对较高的GC含量区域特异,使得其在PCR循环中在比正向或反向引物高约10°C的温度下熔化。这使得**TaqMan**[®]探针在PCR退火/延伸步骤期间保持完全杂交。由于Taq聚合酶在PCR过程中酶促合成新链,因此其最终将到达退火的**TaqMan**[®]探针。然后,Taq聚合酶5'至

3'核酸内切酶活性将通过消化 **TaqMan®** 探针取代其以释放荧光报道分子,以用于使用实时荧光检测系统定量检测其现在未淬灭的信号。

[0172] 用于MethyLight™分析的典型试剂(例如,可在典型的基于MethyLight™的试剂盒中找到)可包括但不限于:用于特定基因座的PCR引物(例如,特定基因、标志物、DMR、基因区域、标志物区域、亚硫酸氢盐处理的DNA序列、CpG岛等); **TaqMan®** 或 **Lightcycler®** 探针; 优化的PCR缓冲液和脱氧核苷酸; 和Taq聚合酶。

[0173] **QM™**(定量甲基化)测定是基因组DNA样品中甲基化模式的替代定量测试,其中序列区分在探针杂交水平上发生。在此定量版本中,PCR反应在荧光探针存在下提供无偏扩增,所述荧光探针重叠特定的推定甲基化位点。通过其中引物和探针均不覆盖任何CpG二核苷酸的反应提供对输入DNA量的无偏控制。可替代地,通过用不覆盖已知甲基化位点的对照寡核苷酸(基于荧光的HeavyMethyl™和MSP技术版本)或用覆盖潜在甲基化位点的寡核苷酸探测有偏向的PCR池来实现基因组甲基化的定性测试。

[0174] **QM™**过程可与扩增过程中的任何合适的探针(例如“**TaqMan®**”探针、**Lightcycler®**探针)一起使用。例如,双链基因组DNA用亚硫酸氢钠处理,并经受无偏引物和**TaqMan®**探针。**TaqMan®**探针用荧光“报告基因”和“猝灭剂”分子双重标记,并且被设计为对相对较高的GC含量区域特异,使得其在PCR循环中在比正向或反向引物高约10°C的温度下熔化。这使得**TaqMan®**探针在PCR退火/延伸步骤期间保持完全杂交。由于Taq聚合酶在PCR过程中酶促合成新链,因此其最终将到达退火的**TaqMan®**探针。然后,Taq聚合酶5'至3'核酸内切酶活性将通过消化**TaqMan®**探针取代其以释放荧光报道分子,以用于使用实时荧光检测系统定量检测其现在未淬灭的信号。用于**QM™**分析的典型试剂(例如,可在典型的基于**QM™**的试剂盒中找到)可包括但不限于:用于特定基因座的PCR引物(例如,特定基因、标志物、DMR、基因区域、标志物区域、亚硫酸氢盐处理的DNA序列、CpG岛等); **TaqMan®**或 **Lightcycler®** 探针; 优化的PCR缓冲液和脱氧核苷酸; 和Taq聚合酶。

[0175] **Ms-SNuPE™**技术是一种定量方法,用于评估基于亚硫酸氢盐处理DNA,随后进行单核苷酸引物延伸的特定CpG位点处的甲基化差异(Gonzalgo和Jones, *Nucleic Acids Res.* 25:2529-2531, 1997)。简而言之,使基因组DNA与亚硫酸氢钠反应以将未甲基化胞嘧啶转化为尿嘧啶,同时保持5-甲基胞嘧啶不变。然后使用对亚硫酸氢盐转化的DNA特异的PCR引物进行所希望靶序列的扩增,并且分离所得产物并用作在感兴趣的CpG位点处进行甲基化分析的模板。可分析少量的DNA(例如,显微切割的病理学部分),并且它避免使用限制性内切酶来确定CpG位点处的甲基化状况。

[0176] 用于**Ms-SNuPE™**分析的典型试剂(例如,可在典型的基于**Ms-SNuPE™**的试剂盒中找到)可包括但不限于:用于特定基因座的PCR引物(例如,特定基因、标志物、DMR、基因区域、标志物区域、亚硫酸氢盐处理的DNA序列、CpG岛等); 优化的PCR缓冲液和脱氧核苷酸; 凝胶提取试剂盒; 阳性对照引物; 用于特定基因座的**Ms-SNuPE™**引物; 反应缓冲液(用于**Ms-SNuPE**反应); 和标记的核苷酸。另外,亚硫酸氢盐转化试剂可包括:DNA变性缓冲液; 碘化缓冲液; DNA回收试剂或试剂盒(例如,沉淀、超滤、亲和柱); 脱磷酸缓冲液; 和DNA回收组分。

[0177] 简化代表性亚硫酸氢盐测序(RRBS)开始于亚硫酸氢盐处理核酸以将所有未甲基

化的胞嘧啶转化为尿嘧啶,然后在偶联到衔接子配体后进行限制酶消化(例如,通过识别包括诸如MspI的CG序列的位点的酶)和片段的完全测序。限制性内切酶的选择富集CpG密集区域的片段,从而减少可能在分析期间映射到多个基因位置的冗余序列的数量。因此,RRBS通过选择用于测序的限制性片段的子集(例如,通过使用制备性凝胶电泳的大小选择)来降低核酸样品的复杂性。与全基因组亚硫酸氢盐测序相反,通过限制性内切酶消化产生的每个片段均含有至少一个CpG二核苷酸的DNA甲基化信息。因此,RRBS富集样品中在这些区域中具有高频率的限制性内切酶切割位点的启动子、CpG岛和其他基因组特征,并且因此提供了一种评估一个或多个基因组基因座的甲基化状态的测定。

[0178] RRBS的典型方案包括用限制性内切酶(如MspI)消化核酸样品、填充突出端和A-尾部、连接衔接子、亚硫酸氢盐转化和PCR的步骤。参见例如,等人(2005)“Genome-scale DNA methylation mapping of clinical samples at single-nucleotide resolution”Nat Methods 7:133-6;Meissner等人(2005)“Reduced representation bisulfite sequencing for comparative high-resolution DNA methylation analysis”Nucleic Acids Res.33:5868-77。

[0179] 在一些实施方案中,将定量等位基因特异性实时靶标和信号放大(QuARTS)测定用于评估甲基化状态。在每个QuARTS测定中依次发生三个反应,包括初级反应中的扩增(反应1)和靶探针裂解(反应2);和二次反应中的FRET裂解和荧光信号产生(反应3)。当用特异性引物扩增靶核酸时,具有皮瓣序列的特异性检测探针松散地结合扩增子。特异性侵入性寡核苷酸在靶结合位点处的存在导致裂解酶通过在检测探针与皮瓣序列之间切割而释放皮瓣序列。皮瓣序列与对应FRET盒的非发夹部分互补。因此,皮瓣序列在FRET盒上起侵入性寡核苷酸的作用,并在FRET盒荧光团与猝灭剂之间实现裂解,这产生荧光信号。裂解反应可每个靶标切割多个探针,并且因此每个皮瓣释放多个荧光团,从而提供指数信号放大。通过使用具有不同染料的FRET盒,QuARTS可在单个反应孔中检测多个靶标。参见例如,Zou等人(2010)“Sensitive quantification of methylated markers with a novel methylation specific technology”Clin Chem 56:A199;美国专利申请序列号12/946,737、12/946,745、12/946,752和61/548,639。

[0180] 术语“亚硫酸氢盐试剂”是指包括亚硫酸氢盐、焦亚硫酸盐、重亚硫酸盐或其组合的试剂,其如本文所公开可用于区分甲基化和未甲基化CpG二核苷酸序列。所述处理的方法是本领域已知的(例如,PCT/EP2004/011715,其以引用的方式整体并入)。优选的是亚硫酸氢盐处理在变性溶剂(如但不限于正-烷二醇或二乙二醇二甲醚(DME))存在下或在二噁烷或二噁烷衍生物的存在下进行。在一些实施方案中,变性溶剂的使用浓度在1%与35%(v/v)之间。在一些实施方案中,亚硫酸氢盐反应在清除剂(如但不限于色满衍生物,例如6-羟基-2,5,7,8-四甲基色满2-羧酸或三羟基苯甲酸及其衍生物,例如没食子酸)存在下进行(参见:PCT/EP2004/011715,其以引用的方式整体并入)。亚硫酸氢盐转化优选地在30°C与70°C之间的反应温度下进行,由此在所述反应过程中温度升高至超过85°C持续短时间(参见:PCT/EP2004/011715,其以引用的方式整体并入)。优选地在定量之前纯化亚硫酸氢盐处理的DNA。这可通过本领域已知的任何方法进行,如但不限于超滤,例如,通过MicroconTM柱(由MilliporeTM制造)。根据改进的制造商的方案进行纯化(参见例如,PCT/EP2004/011715,其以引用的方式整体并入)。

[0181] 在一些实施方案中,使用根据本发明的引物寡核苷酸组(例如,参见表5和/或表7)和扩增酶扩增处理的DNA的片段。几个DNA片段的扩增可在同一个反应容器中同时进行。通常,使用聚合酶链反应(PCR)进行扩增。扩增子的长度通常是100至2000个碱基对。

[0182] 在所述方法的另一个实施方案中,可通过使用甲基化特异性引物寡核苷酸检测包含DMR(例如,如表3和表22中提供的DMR 1-140)的标志物内或附近的CpG位置的甲基化状况。此技术(MSP)已描述于Herman的美国专利号6,265,171中。使用甲基化状况特异性引物扩增亚硫酸氢盐处理的DNA允许区分甲基化与未甲基化核酸。MSP引物对含有至少一个与亚硫酸氢盐处理的CpG二核苷酸杂交的引物。因此,所述引物的序列包含至少一个CpG二核苷酸。对非甲基化DNA特异的MSP引物在CpG中的C位置处含有“T”。

[0183] 通过扩增获得的片段可携带直接或间接可检测的标记。在一些实施方案中,标记是荧光标记、放射性核素或可分离的分子片段,其具有可在质谱仪中检测的典型质量。在所述标记是质量标记的情况下,一些实施方案提供标记的扩增子具有单个正或负净电荷,从而允许质谱仪中更好的可检测性。可通过例如基质辅助激光解吸/电离质谱(MALDI)或使用电子喷雾质谱(ESI)进行所述检测并使其可视化。

[0184] 用于分离适用于这些测定技术的DNA的方法是本领域已知的。具体地,一些实施方案包括如美国专利申请序列号13/470,251(“Isolation of Nucleic Acids”)中所述地分离核酸,其以引用的方式整体并入本文。

[0185] 方法

[0186] 在一些实施方案中,提供了包括以下步骤的技术、方法:

[0187] 1)使从受试者获得的核酸(例如,例如从体液(如粪便样品或前列腺组织或血浆样品)分离的基因组DNA)与区分包含DMR(例如,例如表3和表22中提供的DMR 1-140)的至少一种标志物内的甲基化和非甲基化CpG二核苷酸的至少一种试剂或一系列试剂接触,并且

[0188] 2)检测前列腺癌(例如,提供有灵敏度大于或等于80%,并且特异性大于或等于80%)。

[0189] 在一些实施方案中,提供了包括以下步骤的技术、方法:

[0190] 1)使从受试者获得的核酸(例如,例如从体液(如粪便样品或前列腺组织)分离的基因组DNA)与区分选自具有注释的染色体区域的至少一种标志物内的甲基化和非甲基化CpG二核苷酸的至少一种试剂或一系列试剂接触,所述注释选自下组,该组由以下组成:AC0XL、AKR1B1_3644、ANXA2、CHST11_2206、FLJ45983、GAS6、GRASP、HAPLN3、HCG4P6、HES5_0822、ITPRIPL1、KCNK4、MAX.chr1.61519554-61519667、MAX.chr2.97193166-97193253、MAX.chr3.193、MAX.chr3.72788028-72788112、RAI1_7469、RASSF2、SERPINB9_3389、SLC4A11和TPM4_8047,并且

[0191] 2)检测前列腺癌(例如,提供有灵敏度大于或等于80%,并且特异性大于或等于80%)。

[0192] 在一些实施方案中,提供了包括以下步骤的技术、方法:

[0193] 1)使从受试者获得的核酸(例如,例如从体液(如粪便样品或前列腺组织)分离的基因组DNA)与区分选自具有注释的染色体区域的至少一种标志物内的甲基化和非甲基化CpG二核苷酸的至少一种试剂或一系列试剂接触,所述注释选自下组,该组由以下组成:SERPINB9_3479、FL0T1_1665、HCG4P6_4618、CHST11_2206、MAX.chr12.485、GRASP_0932、

GAS6_6425、MAX.chr3.193、MAX.chr2.971_3164、MAX.chr3.727_8028、HES5_0840、TPM4_8037、SLC03A1_6187、ITPR1PL1_1244、AKR1B1_3644、RASGRF2_6325、ZNF655_6075、PAMR1_7364、ST6GALNAC2_1113、CCNJL_9070、KCNB2_9128、IGFBP7_6412和WNT3A_5487，并且

[0194] 2) 检测前列腺癌(例如,提供有灵敏度大于或等于80%,并且特异性大于或等于80%)。

[0195] 在一些实施方案中,提供了包括以下步骤的技术、方法:

[0196] 1) 使从受试者获得的核酸(例如,例如从血浆样品分离的基因组DNA)与区分选自具有注释的染色体区域的至少一种标志物内的甲基化和非甲基化CpG二核苷酸的至少一种试剂或一系列试剂接触,所述注释选自下组,该组由以下组成: max.chr3.193、HES5、SLC03A1和TPM4_8047,并且

[0197] 2) 检测前列腺癌(例如,提供有灵敏度大于或等于80%,并且特异性大于或等于80%)。

[0198] 优选地,灵敏度是约70%至约100%、或约80%至约90%、或约80%至约85%。优选地,特异性是约70%至约100%、或约80%至约90%、或约80%至约85%。

[0199] 可通过任何方法分离基因组DNA,包括使用可商购获得的试剂盒。简而言之,其中感兴趣的DNA被细胞膜包封,生物样品必须被酶促、化学或机械方法破坏和裂解。然后例如可通过用蛋白酶K消化来清除DNA溶液中的蛋白质和其他污染物。然后从溶液回收基因组DNA。这可通过多种方法进行,包括盐析、有机提取或DNA与固相支持物的结合。方法的选择将受到若干因素的影响,包括时间、费用和DNA的所需数量。包含肿瘤物质或肿瘤前物质的所有临床样品类型均适用于本方法,例如细胞系、组织学载玻片、活组织检查、石蜡包埋的组织、体液、粪便、前列腺组织、结肠排出物、尿液、血浆、血清、全血、分离的血细胞、从血液分离的细胞及其组合。

[0200] 所述技术不限于用于制备样品并提供用于测试的核酸的方法。例如,在一些实施方案中,例如,如美国专利申请序列号61/485386中详述或通过相关方法,使用直接基因捕获从粪便样品或血液或血浆样品分离DNA。

[0201] 然后用区分包含DMR(例如,例如由表3和表22提供的DMR 1-140)的至少一种标志物内的甲基化和非甲基化CpG二核苷酸的至少一种试剂或一系列试剂处理基因组DNA样品。

[0202] 在一些实施方案中,试剂将在5'-位置处未甲基化的胞嘧啶碱基转化为尿嘧啶、胸腺嘧啶或在杂交行为方面与胞嘧啶不同的另一种碱基。然而在一些实施方案中,试剂可以是甲基化敏感性限制性内切酶。

[0203] 在一些实施方案中,以这样的方式处理基因组DNA样品:使得在5'位置处未甲基化的胞嘧啶碱基转化为尿嘧啶、胸腺嘧啶或在杂交行为方面与胞嘧啶不同的另一种碱基。在一些实施方案中,此处理用亚硫酸氢盐(重亚硫酸盐、焦亚硫酸盐)进行,随后进行碱水解。

[0204] 然后分析处理的核酸以确定靶基因序列(来自包含DMR(例如,选自例如表3和表22中提供的DMR 1-140的至少一种DMR)的标志物的至少一种基因、基因组序列或核苷酸)的甲基化状态。分析方法可选自本领域已知的那些,包括本文所列的那些,例如,如本文所述的QuARTS和MSP。

[0205] 异常甲基化,更具体地是包含DMR(例如,例如由表3和表22提供的DMR 1-140)的标志物的超甲基化与前列腺癌相关联。

[0206] 所述技术涉及与前列腺癌相关的任何样品的分析。例如,在一些实施方案中,样品包括从患者获得的组织和/或生物流体。在一些实施方案中,样品包括分泌物。在一些实施方案中,样品包括血液、血清、血浆、胃分泌物、胰液、胃肠活组织检查样品、来自前列腺活组织检查的显微切割细胞,和/或从粪便回收的前列腺细胞。在一些实施方案中,受试者是人。所述样品可包括来自前列腺、肝、胆管、胰腺、胃、结肠、直肠、食道、小肠、阑尾、十二指肠、息肉、胆囊、肛门和/或腹膜的细胞、分泌物或组织。在一些实施方案中,样品包括细胞液、腹水、尿液、粪便、胰液、内窥镜检查期间获得的流体、血液、粘液或唾液。在一些实施方案中,样品是粪便样品。

[0207] 此类样品可通过本领域已知(如对于技术人员是显而易见)的任何数量的方法获得。例如,尿液和粪便样品很容易获得,而血液、腹水、血清或胰液样品可通过例如针和注射器肠胃外获得。通过使样品经受本领域技术人员已知的包括但不限于离心和过滤的各种技术,可获得无细胞或基本上无细胞的样品。尽管通常优选不使用侵入性技术来获得样品,但仍可优选获得样品,如组织匀浆、组织切片和活组织检查样本。

[0208] 在一些实施方案中,技术涉及一种用于治疗患者(例如,患有前列腺癌、患有早期前列腺癌或可能发展前列腺癌的患者)的方法,所述方法包括确定如本文提供的一种或多种DMR的甲基化状态,并且基于确定甲基化状态的结果向患者施用治疗。治疗可以是施用药物化合物、疫苗、进行手术、对患者成像、进行另一次测试。优选地,所述用途在临床筛查方法、预后评估方法、监测治疗结果的方法、鉴定最可能响应于特定治疗处理的患者的方法、对患者或受试者进行成像的方法,以及药物筛选和开发的方法中。

[0209] 在所述技术的一些实施方案中,提供了一种用于诊断受试者中的前列腺癌的方法。如本文所用的术语“进行诊断”和“诊断”是指技术人员通过其可估计并且甚至确定受试者是否患有给定疾病或病状或是否可能在未来发展给定疾病或病状的方法。技术人员经常基于一个或多个诊断指标进行诊断,例如像生物标志物(例如,如本文所公开的DMR),其甲基化状态指示病状的存在、严重性或不存在。

[0210] 随着诊断,临床癌症预后涉及确定癌症的侵袭性和肿瘤复发的可能性以计划最有效的治疗。如果可进行更准确的预后或甚至可评估发展癌症的潜在风险,则可选择适当的治疗,并且在某些情况下可选择对患者不太严格的治疗。癌症生物标志物的评估(例如,确定甲基化状态)可用于分离具有良好预后和/或发展癌症的低风险的受试者,所述受试者将不需要来自那些更可能发展癌症或患有可能受益于更强化治疗的癌症复发的受试者的治疗或有限治疗。

[0211] 因此,如本文所用,“做出诊断”或“进行诊断”进一步包括确定发展癌症的风险或确定预后,这可基于本文公开的诊断生物标志物(例如,DMR)的量度,提供预测临床结果(具有或没有医学治疗)、选择适当的治疗(或治疗是否有效),或监测当前治疗并可能改变治疗。此外,在本公开主题的一些实施方案中,可进行生物标志物随时间的多次确定以促进诊断和/或预后。生物标志物的时间变化可用于预测临床结果、监测前列腺癌的进展和/或监测针对癌症的适当疗法的功效。在这样的实施方案中,例如,可预期在有效治疗过程中随时间观察生物样品中本文公开的一种或多种生物标志物(例如,DMR)(并且如果受到监测,可能还有一种或多种另外的生物标志物)的甲基化状态的变化。

[0212] 在一些实施方案中,本发明公开的主题进一步提供了一种用于确定是否开始或继

续预防或治疗受试者中的癌症的方法。在一些实施方案中,方法包括在一段时间内提供来自受试者的一系列生物样品;分析一系列生物样品以确定每个生物样品中本文公开的至少一种生物标志物的甲基化状态;并比较每个生物样品中一种或多种生物标志物的甲基化状态的任何可测量变化。在所述时间段内生物标志物的甲基化状态的任何变化可用于预测发展癌症的风险、预测临床结果、确定是否开始或继续预防或治疗癌症,以及当前疗法是否有效地治疗癌症。例如,可在开始治疗之前选择第一时间点,并且可在治疗开始后的某个时间选择第二时间点。可在从不同时间点取得的每个样品中测量甲基化状态,并记录定性和/或定量差异。来自不同样品的生物标志物水平的甲基化状态的变化可与受试者中的前列腺癌风险、预后、确定治疗功效和/或癌症进展相关联。

[0213] 在优选的实施方案中,本发明的方法和组合物用于在早期阶段(例如,在疾病症状出现之前)治疗或诊断疾病。在一些实施方案中,本发明的方法和组合物用于在临床阶段治疗或诊断疾病。

[0214] 如所指出的,在一些实施方案中,可进行一种或多种诊断或预后生物标志物的多次测定,并且标志物的时间变化可用于确定诊断或预后。例如,诊断标志物可在初始时间确定,并在第二时间再次确定。在此类实施方案中,标志物从初始时间到第二时间的增加可诊断癌症的特定类型或严重性,或给定的预后。同样地,标志物从初始时间到第二时间的减少可指示癌症的特定类型或严重性,或给定的预后。此外,一种或多种标志物的变化程度可与癌症的严重性和未来的不良事件相关。技术人员将理解,虽然在某些实施方案中,可在多个时间点对同一生物标志物进行比较的测量,也可在一个时间点测量给定的生物标志物,并在第二时间测量第二生物标志物,并且这些标志物的比较可提供诊断信息。

[0215] 如本文所用,短语“确定预后”是指技术人员可通过其预测受试者中病状的过程或结果的方法。术语“预后”并不是指以100%准确度预测病状的过程或结果的能力,或者甚至是基于生物标志物(例如,DMR)的甲基化状态可预测的或多或少可能发生的给定过程或结果的能力。相反,技术人员将理解术语“预后”是指某一过程或结果将发生的概率增加;也就是说,当与未表现出所述病状的那些个体相比时,在表现出给定病状的受试者中更可能发生过程或结果。例如,在未表现出病状(例如,具有一种或多种DMR的正常甲基化状态)的个体中,给定结果(例如,患有前列腺癌)的机会可能非常低。

[0216] 在一些实施方案中,统计分析将预后指标与不良结果的倾向相关联。例如,在一些实施方案中,如通过统计学显著性水平确定的,不同于从未患有癌症的患者获得的正常对照样品中的甲基化状态的甲基化状态可表示受试者比具有更类似于对照样品中的甲基化状态的水平的受试者更可能患有癌症。另外,甲基化状态从基线(例如,“正常”)水平的变化可反映受试者预后,并且甲基化状态的变化程度可与不良事件的严重性相关。经常通过比较两个或更多个群体并确定置信区间和/或p值来确定统计学显著性。参见例如,Dowdy and Wearden, *Statistics for Research*, John Wiley&Sons, New York, 1983, 其以引用的方式整体并入本文。本主题的示例性置信区间是90%、95%、97.5%、98%、99%、99.5%、99.9%和99.99%,而示例性p值是0.1、0.05、0.025、0.02、0.01、0.005、0.001和0.0001。

[0217] 在其他实施方案中,可建立本文公开的预后或诊断生物标志物(例如,DMR)的甲基化状态的阈值变化程度,并简单地将生物样品中双标志物的甲基化状态的变化程度与甲基化状态的阈值变化程度进行比较。本文提供的生物标志物的甲基化状态的优选阈值变化是

约5%、约10%、约15%、约20%、约25%、约30%、约50%、约75%、约100%和约150%。在又其他实施方案中,可建立“诺模图”,通过所述诺模图预后或诊断指标(生物标志物或生物标志物的组合)的甲基化状态与对给定结果的相关处置直接相关。技术人员熟悉使用此类诺模图来关联两个数值,同时理解此测量中的不确定性与标志物浓度的不确定性相同,因为参考了单个样品测量值,而不是群体平均值。

[0218] 在一些实施方案中,对照样品与生物样品同时进行分析,使得可将从生物样品获得的结果与从对照样品获得的结果进行比较。另外,预期可提供标准曲线,利用所述标准曲线可比较生物样品的测定结果。如果使用荧光标记,则此类标准曲线表示生物标志物的甲基化状态作为测定单位(例如荧光信号强度)的函数。使用取自多个供体的样品,可提供标准曲线用于正常组织中一种或多种生物标志物的对照甲基化状态,以及取自具有化生的供体或取自患有前列腺癌的供体的组织中的一种或多种生物标志物的“有风险”水平。在所述方法的某些实施方案中,在从受试者获得的生物样品中鉴定本文提供的一种或多种DMR的异常甲基化状态后,将受试者鉴定为具有化生。在所述方法的其他实施方案中,检测从受试者获得的生物样品中的一种或多种此类生物标志物的异常甲基化状态导致受试者被鉴定为患有癌症。

[0219] 标志物的分析可与一个测试样品内的另外标志物分开或同时进行。例如,可将若干标志物组合成一个测试,以有效地处理多个样品并且可能提供更好的诊断和/或预后准确度。另外,本领域技术人员将认识到测试来自相同受试者的多个样品(例如,在连续时间点)的价值。连续样品的这种测试可允许鉴定标志物甲基化状态随时间的变化。甲基化状态的变化,以及甲基化状态的变化的缺乏,可提供关于疾病状况的有用信息,其包括但不限于鉴定事件发生的大致时间、可挽救组织的存在和量、药物治疗的适当性、各种治疗的有效性,以及对受试者结果的鉴定,包括未来事件的风险。

[0220] 生物标志物的分析可以各种物理形式进行。例如,微量滴定板或自动化的使用可用于促进大量测试样品的处理。可替代地,可开发单一样品形式以便于及时地立即治疗和诊断,例如,在门诊运输或急诊室环境中。

[0221] 在一些实施方案中,如果与对照甲基化状态相比时,样品中至少一种生物标志物的甲基化状态存在可测量的差异,则将受试者诊断为患有前列腺癌。相反,当在生物样品中未鉴定出甲基化状态的变化时,可将受试者鉴定为未患有前列腺癌、没有患癌症的风险或者具有低的癌症风险。在这方面,患有癌症或其风险的受试者可与具有低至基本上没有癌症或其风险的受试者区分。具有发展前列腺癌风险的那些受试者可被置于更密集和/或定期的筛查计划中,包括内窥镜监视。另一方面,具有低至基本上没有风险的那些受试者可避免进行内窥镜检查,直到未来筛查(例如根据本技术进行的筛查)表明在这些受试者中出现了前列腺癌的风险。

[0222] 如上所述,取决于本技术方法的实施方案,检测一种或多种生物标志物的甲基化状态的变化可以是定性测定,或者其可以是定量测定。因此,诊断受试者患有前列腺癌或有发展前列腺癌风险的步骤表明某些阈值测量(例如,生物样品中一种或多种生物标志物的甲基化状态)与预定的对照甲基化状态不同。在所述方法的一些实施方案中,对照甲基化状态是生物标志物的任何可检测的甲基化状态。在其中对照样品与生物样品同时测试的方法的其他实施方案中,预定的甲基化状态是对对照品中的甲基化状态。在所述方法的其他实

施方案中,预定的甲基化状态基于标准曲线和/或由标准曲线鉴定。在所述方法的其他实施方案中,预定的甲基化状态是特定状态或状态范围。因此,部分地基于所实施的方法的实施方案和所希望的特异性等,可在对于本领域技术人员将显而易见的可接受限度内选择预定的甲基化状态。

[0223] 进一步关于诊断方法,优选的受试者是脊椎动物受试者。优选的脊椎动物是温血动物;优选的温血脊椎动物是哺乳动物。优选的哺乳动物最优先是人。如本文所用,术语“受试者”包括人和动物受试者两者。因此,本文提供了兽医治疗用途。因此,本技术提供了对哺乳动物,如人,以及由于濒临灭绝而具有重要性的那些哺乳动物,如西伯利亚虎;具有重要经济意义,如在农场饲养以供人消费的动物;和/或对人具有社会重要性的动物,如作为宠物或在动物园中饲养的动物的诊断。此类动物的实例包括但不限于:食肉动物,如猫和狗;猪,包括家猪、肉猪和野猪;反刍动物和/或有蹄类动物,如家牛、牛、羊、长颈鹿、鹿、山羊、野牛和骆驼;和马。因此,还提供了家畜的诊断和治疗,包括但不限于驯养的猪、反刍动物、有蹄类动物、马(包括赛马)等。本发明公开的主题进一步包括一种用于诊断受试者中的前列腺癌的系统。所述系统可例如作为商业试剂盒提供,其可用于筛查前列腺癌的风险或诊断已从其收集生物样品的受试者中的前列腺癌。根据本技术提供的示例性系统包括评估如表3和表5中提供的DMR的甲基化状态。

[0224] 实施例

[0225] 实施例I.

[0226] 此实施例提供了用于实施例II、III、IV、V和VI的材料和方法。

[0227] 实施例II、III、IV、V和VI证明甲基化DNA标志物可区分前列腺组织(例如,癌前列腺组织和/或非癌前列腺组织)与非前列腺组织(例如,白细胞),甲基化DNA标志物可区分癌前列腺组织与非癌前列腺组织,甲基化DNA标志物可区分高度侵袭性癌前列腺组织(例如,格里森评分等于或高于7.0(例如,7、8、9、10)与较低侵袭性癌前列腺组织(例如,格里森评分低于7(例如,6),以及甲基化DNA标志物可检测血液样品中的PCa。

[0228] 这些实验包括五个阶段。首先,使用简化代表性亚硫酸氢盐测序(RRBS)(参见例如,Gu H等人,Nat Methods 2010;7:133-6)对从前列腺癌(PCa)组织(格里森评分6和7+)、正常前列腺和来自健康志愿者的血沉棕黄层样品提取的DNA进行DNA甲基化标志物发现。其次,通过用于开发实时甲基化特异性PCR测定(qMSP)的严格的过滤标准和序列鉴定可区分的差异甲基化区域(DMR)。然后将这些测定应用于原始样品组以确保结果的再现性(技术验证)。第三,选择表现最佳的候选标志物进行对从独立档案案例和对照组织提取的DNA的qMSP生物验证。第四,将候选标志物序列在计算机上跨泛癌RRBS测序数据集进行比较,以测量每个标志物的位点特异性甲基化程度。第五,选择一组高效PCa标志物用于在盲法独立血浆样品中进行测试,以评估临床培养基中的PCa检测。

[0229] 研究受试者和样品

[0230] 所述研究得到了梅奥诊所机构审查委员会(Rochester, MN)的批准。新鲜冷冻(FF)组织、血浆和血沉棕黄层样品由IRB批准的患者生物库提供。由专家GI病理学家重新审查肿瘤组织切片以确认诊断并估计肿瘤细胞性。然后对切片进行宏观解剖。使用QiaAmp Mini试剂盒(Qiagen, Valencia CA)纯化基因组DNA,并且随后用AMPure XP试剂盒(Beckman Coulter, Brea CA)再纯化。

[0231] 简化代表性亚硫酸氢盐测序文库的制备

[0232] 将150ng每种样品的DNA稀释于26ul Te缓冲液 (5.77ng/uL) 中。将其在1X最终浓度 CutSmart 缓冲液 (New England Biolabs) 中通过1ul (20个单位) MspI 消化过夜。将3'突出端末端修复并用0.6ul 100mM dATP、0.06ul 100mM dCTP和0.06ul dGTP与2ul (10个单位) Klenow DNA聚合酶 (New England Biolabs) 的混合物加尾。将产物在30度下孵育20分钟、在37度下孵育20分钟，并保持在4度下。在末端修复后，通过2X Agencourt Ampure XP 珠 (Beckman Coulter) 纯化产物，用70% EtOH洗涤两次，并在20ul水中洗脱。使用在16度下孵育过夜的1X T4连接酶缓冲液中的1ul T4连接酶 (400个单位) 将 Illumina 衔接子连接到产物。将产物在65度下处理20分钟以加热使酶失活。在连接后，通过2X Agencourt Ampure XP 珠 (Beckman Coulter) 纯化产物，用70% EtOH洗涤两次，并在47ul水中洗脱。如其方案中所述，通过EZ-96DNA甲基化试剂盒 (Zymo Research) 亚硫酸氢盐转化45ul产物。通过2X Agencourt Ampure XP 珠 (Beckman Coulter) 纯化转化的产物，用70% EtOH洗涤两次，并在22ul水中洗脱。使用16ul亚硫酸氢盐转化的产物、1ul (2.5个单位) PfuTurbo Cx hotstart DNA聚合酶、0.5ul dNTP (每种25mM)、6ul Illumina 指数 (每种2.5uM) 和1X PfuTurbo Cx hotstart DNA聚合酶缓冲液以50ul总体积经由PCR加入 Illumina 指标。通过使用0.7X下的 Ampure XP 选择产物的双倍尺寸并收集上清液，而随后使用1.2X下的Ampure XP 并保留与珠结合的物质。最终产物在40ul中洗脱。通过Pico Green (Molecular Probes) 测定DNA质量产率，并在Tecan荧光计上测量。通过Agilent 2100 (Agilent) 上的高灵敏度DNA芯片测定DNA尺寸。计算产物的摩尔浓度，并将样品等摩尔地在4个样品/池中合并至10nM。

[0233] 大规模并行测序和生物信息学

[0234] 根据随机通道分配将样品装载到流动池上。测序由 Illumina HiSeq 2000 上的 Mayo Clinic Medical Genome Facility 处的 Next Generation Sequencing Core 进行。读取是单向的101个循环。每个流动细胞通道产生100-120百万个读数，足以使对齐序列的中值覆盖度为30-50倍测序深度。标准的 Illumina 管道软件称为基础，并以 fastq 格式生成读取。将 SAAP-RRBS (用于简化代表性亚硫酸氢盐测序的简化分析和注释管道) 用于序列读取评估和清除、与参考基因组的比对、甲基化状况提取和 CpG 报告和注释。排除具有低覆盖率 (≤ 10) 的 CpG。第三级分析包括去除无信息或低样品覆盖的 CpG，并在滑动 100bp 窗口内鉴定具有低背景和密集簇的甲基化 CpG 区域。读取深度标准基于所希望的统计功效以检测病例与对照之间甲基化 % 的 10% 差异。基于读数计数，通过每个 DMR 的甲基化百分比的逻辑回归来确定统计学显著性。为了考虑个体受试者的不同读取深度，使用过度分散的逻辑回归模型，其中使用来自拟合模型的残差的 Pearson 卡方统计来估计离散参数。如果对照组中甲基化 % 在癌症中 $\leq 1\%$ 且 $\geq 10\%$ ，则进一步考虑根据其显著性水平排列的 DMR。在大多数器官部位中，这导致了数百名潜在候选者。利用的另外过滤器是受试者操作特征曲线下面积 (AUC)、% 甲基化情况/对照倍数变化 (FC)，以及整个 DMR (并且在对照中缺乏) 中 CpG 的样品共甲基化的阳性样品。

[0235] 技术和生物组织验证

[0236] 甲基化特异性 PCR (MSP) 标志物测定是针对来自 PCa 发现数据集 (如由以上列出的标准确定) 的 120 种最有希望的 DMR 而开发的。引物由软件 (Methprimer-University of California, San Francisco CA) 或手工设计。通过 SYBR Green qPCR 对亚硫酸氢盐转化 (甲

基化和未甲基化的基因组DNA)、未转化的和非模板对照进行严格测试和优化测定。重新设计或丢弃与阴性对照交叉反应的测定。另外,进行溶解曲线分析以确保发生特异性扩增。对于技术验证阶段,通过qMSP重新测试用于RRBS发现的同一样品。将设计为甲基化盲的β-肌动蛋白测定用作代表总DNA拷贝的分母。通过逻辑回归分析数据,并将AUC和信号与背景结果与发现值进行比较。大约27%的标志物表现不佳并被淘汰。通过qMSP在扩展的独立组织样品组上测试剩余部分(N=72)。对结果进行逻辑分析,并且结果指标是AUC、FC和稳健案例样品%甲基化。

[0237] 跨器官验证

[0238] 为了评估最佳甲基化标志物如何在前列腺外进行,使用与之前测序的其他主要癌症(结肠癌、胰腺癌、食管癌、肝癌和胃癌)相比较的前列腺样品中验证DMR的测序读数构建了比较性CpG %甲基化基质。基于1)生物组织验证阶段的总体表现和2)跨其他癌症的标志物的位点特异性特征,选择最终的标志物组以在血浆中进行测试。为了最好地检测血液中的PCa,给定过量的非PCa DNA,选择了一种将表现出通用和前列腺特异性癌症信号的稳健标志物组。

[0239] QuART测定设计和血浆验证

[0240] 通过以下自动化二氧化硅珠方法从3-4mL堆积的冷冻血浆提取DNA(参见例如,美国专利申请序列号15/335,111) :

[0241] 表1

1	2ml Te 缓冲液 (1mM Tris 0.1mM EDTA)
2	100 μ l of 120 cp/ μ l 斑马鱼, 在 0.4 ng/ μ l 鱼 DNA 稀释液中
3	7ml 血浆裂解缓冲液 (4.3M GTC 10% IGEPAL)
4	2ml 血浆
5	在 55°C 下孵育 1 小时
6	加入 200 μ l 结合珠
7	加入 2.8ml 100% 异丙醇
8	在 30°C 下孵育 30 分钟。
9	磁化珠并除去上清液
10	加入 750 μ l 3M GuHCl 56.8% EtOH 以重悬结合珠
11	以 400 RPM 摆动 2 分钟。
12	结合珠并吸出上清液至废液
13	1000 μ l 洗涤 1 次 (80% ETOH), 在 30°C 下孵育 3 分钟, 结合珠并吸出上清液至废液
14	500 μ l 洗涤 1 次 (80% ETOH), 在 30°C 下孵育 3 分钟, 结合珠并吸出上清液至废液
15	250 μ l 洗涤 1 次 (80% ETOH), 在 30°C 下孵育 3 分钟, 结合珠并吸出上清液至废液
16	250 μ l 洗涤 1 次 (80% ETOH), 在 30°C 下孵育 3 分钟, 结合珠并吸出上清液至废液
17	在摇动下于 70°C 下干燥 15 分钟。
18	加入 125 μ l Te 缓冲液 (1mM Tris 0.1mM EDTA), 在摇动下于 65°C 下孵育 25 分钟
19	结合珠并将含有 DNA 的上清液转移到清洁管中
20	在 -20°C 下存储直到使用。
21	然后使用以下方法将DNA亚硫酸氢盐转化并纯化:
22	表2
23	1 5 μ l 0.36% BSA
24	2 70 μ l 样品
25	3 5 μ l 1.6N NaOH
26	4 在 42°C 下孵育 (变性) 20'
27	5 冷却 8'
28	6 加入 120 μ l 亚硫酸氢铵
29	7 在 65°C 下孵育 (转化) 75' (摇动 3')
30	8 加入 750 μ l 7M GuHCl
31	9 加入 50 μ l 结合珠

10	在摇动下于30°C下孵育30'
11	结合珠
12	吸出上清液至废液
13	加入1000ul 80% ETOH
14	在摇动下于30°C下孵育3'
15	结合珠
16	吸出上清液至废液
17	分配200ul脱碘化溶液
18	在摇动下于30°C下孵育7'
19	结合珠
20	吸出上清液至废液
21	加入250ul 80% ETOH
22	在摇动下于30°C下孵育3'
23	结合珠
24	吸出上清液至废液
25	在摇动下于70°C下将珠干燥15'
26	加入80ul Te缓冲液 (1mM Tris 0.1mM EDTA)
27	在摇动下于65°C下孵育25分钟
28	结合珠并将含有DNA的上清液转移到清洁管中
29	在-20°C下存储直到使用。

[0247] 然后使用从DMR序列开发的引物和探针,在QuARTs-X(参见例如,美国专利申请序列号15/335,096)格式的ABI实时PCR仪器上运行样品(10uL)。从Genscript获得含有感兴趣的标志物序列的质粒,并在1X QuART试剂中稀释至标称浓度为每15uL反应1个拷贝。将反应混合物分配至96个孔中的每个,在LightCycler上循环45个循环,并收集数据。孔被要求含有或不含有样品。将泊松随机变量设置为1,并通过反复试验输入平均成功率值,并用于计算所述值的累积概率。当累积概率等于具有信号的孔的百分比时,已找到了正确的平均成功率,在这种情况下是拷贝数。稀释这些质粒并用作测定标准。

[0248] 通过首先创建用多达12个靶标的引物进行的样品的预扩增板来进行QuART-X,所述靶标经历11个扩增循环。然后将此产物以1:9稀释并用作后续QuART反应的模板,其在三链体反应中仅含有三个靶标。用于计算链计数的标准品未经过预扩增。通过预扩增样品而不是标准品,增加了测定的灵敏度。

[0249] 通过回归分区(rPart)分析结果。使用逻辑回归将多个甲基化标志物组合成单个风险评分是标准技术。然而,在逻辑模型中很难发现和/或模拟标志物之间的高阶相互作用。当此类影响存在时,这限制了标志物组的预测能力。回归分区树(rPart)是一种决策树方法,其能够发现标志物之间的高阶相互作用,以这种方式最大化标志物组的预测准确性。

[0250] 实施例II.

[0251] 此实施例描述了RRBS结果和技术验证结果。

[0252] PCa产生了大量的可区分DMR,其中许多以前未被鉴定。比较PCa样品与正常前列腺的甲基化,鉴定出256个区域,其满足AUC>0.85,FC>20,并且p值<0.05截止值。这些区域中的

22个具有为1的AUC。当将PCa和正常前列腺甲基化与血沉棕黄层样品进行比较时,1895个区域高于截止值,其中827个具有完美的AUC评分。两次比较中的FC分别扩展到数百和数千。研究了区分格里森7+PCa(侵袭性,治疗显示癌症)相比于格里森6PCa(大多数情况下是惰性,通常不需要治疗)的潜在DMR。观察到129个DMR,具有 $FC > 2 (7+/6)$,其中最高 $FC = 72$ 。

[0253] 生物标志物开发过程的第二阶段是解决在初始发现阶段中相对较小的样品量引起的不确定性。在较少数量的DMR或候选标志物上使用不同的技术平台重新测试同一样品是迈向这一目的的第一步。使用SYBR Green的实时或定量甲基化特异性PCR(qMSP)是一种易于使用的方法,其具有高分析灵敏度和特异性。

[0254] 通过增加截止值来选择来自所有三个比较的最佳候选者来选择DMR,直到获得可管理数量的区域($N=120$) (参见表3)。

[0255] 表3.

DMR号	基因注释	染色体号	染色体上的区域 (开始碱基-结束碱基)
1	ADCY9	16	4165628-4165833
2	AKR1B1_3298	7	134143071-134143542
3	AKR1B1_3644	7	134143644-134143716
4	ANKRD35	1	145562809-14562898
5	CLIP4	2	29338053-29338117
6	CYBA_7733	16	88717482-88717805
7	FLJ45983	10	8097100-8097859
8	GRASP	12	52400932-52401144
9	GSTP1	11	67351212-67351638
10	HAPLN3	15	89438198-89438734
11	HCG4P6	6	29894504-29894683
12	HES5_0822	1	2460822-2460998
13	HES5_1047	1	2461047-2461876
14	ITPRIPL1	2	96990982-96991303
15	LRRC4	7	127671395-127672250
16	MAX.chr2.97193166-97193253	2	97193065-97193253
17	MAX.chr3.193	3	193776187-193776257
18	MAX.chr3.72788028-72788112	3	72788028-72788206
19	PTPRN2	7	157361654-157361753

[0256]

[0257]

20	RAI1_7469	17	17627094-17628093
21	RASSF2	20	4803273-4803687
22	Septin9_0074	17	75370074-75370187
23	Septin9_0492	17	75370492-75370581
24	SLC43A3	11	57194414-57194645
25	SLC4A11	20	3218937-3219001
26	SMAD3	15	67413704-67413772
27	SSBP4	19	18539756-18540408
28	TJP2	9	71788646-71789457
29	ABLIM1	10	116391692-116391769
30	AOX1	2	201450664-201450868
31	C3orf18	3	50604997-50605357
32	EIF5A2	3	170625931-170626391
33	EPSTI1	13	43566083-43566214
34	FBXO30	6	146136383-146136441
35	FLOT1_1586	6	30711586-30711681
36	FLOT1_1767	6	30711767-30711864
37	FLOT1_1904	6	30711904-30711966
38	FOSL1	11	65666866-65667056
39	GAS6	13	114566059-114566687
40	GSDMD	8	144640720-144640772
41	KCNK4	11	64059874-64059994
42	LOC100129726_1716	2	43451464-43452664
43	MAX.chr10.74079656-74079694	10	74079656-74079861
44	SERPINB9_3389	6	2903389-2903705
45	SLCO3A1	15	92396091-92396343
46	TPM4_7473	19	16187401-16187514
47	TPM4_8047	19	16187580-16188154
48	AGPS_7349	2	178257349-178257413
49	AGPS_7497	2	178257497-178257568
50	AGPS_7696	2	178257696-178257732
51	AMPD3	11	10472267-10472338
52	CH25H	10	90967594-90967655
53	GALR3	22	38214733-38214808
54	HDAC7_6722	12	48206722-48206761
55	MAX.chr1.227976339-227976430	1	227976339-227976430
56	MAX.chr1.8014263-8014319	1	8014263-8014319
57	MAX.chr19.34311051-34311120	19	34311051-34311120
58	NCRNA00092_4149	9	98784149-98784195
59	ZNF655_6084	7	99156084-99156145

[0258]

60	ZNF655_6545	7	99156545-99156606
61	ZNF655_6762	7	99156762-99156852
62	ABCB1	7	87229774-87229851
63	ACOXL	2	111875299-111875585
64	ANXA2	15	60690904-60690949
65	APBB1IP	10	26727818-26728120
66	ARPC1B_1906	7	98971906-98971950
67	ARPC1B_1967	7	98971967-98971997
68	ATP2B4	1	203598589-203598782
69	CHST11_2032	12	104852032-104852137
70	CHST11_2206	12	104852206-104852307
71	DOK1_2325	2	74782325-74782452
72	FLJ34208	3	194208259-194208471
73	HCG4P6	6	29894504-29894683
74	HEBP2	6	138724600-138724667
75	HEYL	1	40105264-40105646
76	KLF16	19	1855656-1855656
77	LAMA3	18	21269864-21270082
78	LIME1	20	62369129-62369383
79	LOC339674	22	42353799-42353881
80	LOC440925	2	171570371-171570463
81	MAX.chr1.61519554-61519667	1	61519406-61519667
82	MAX.chr1.61519679-61519759	1	61519679-61519759
83	MAX.chr12.48592041-48592162	12	48592041-48592162
84	MAX.chr17.77786640-77786733	17	77786640-77786733
85	PDE4D	5	58334709-58335864
86	PLCL2	3	16925808-16925889
87	SIM2	21	38119920-38120410
88	STX16	20	57224816-57225220
89	WNT3A	1	228225487-228225688
90	ABHD15	17	27893168-27893592
91	ADAP1	7	963082-963154
92	ADD3	10	111767345-111767579
93	AXIN1	16	374825-375308
94	BCL2L11_6887	2	111876440-111876914
95	BCL2L11_6935	2	51717908-51718147
96	BIN2_7908	12	51717908-51717961
97	CCDC88C	14	91790497-91790556
98	CTBP1	4	1210660-1210752
99	DOK1_2096	2	74782096-74782223

100	DOK1_2234	2	74782234-74782302
101	DOK1_2475	2	74782475-74782572
102	FAM129C	19	17634139-17634203
103	FAM78A_1379	9	134151379-134151451
104	FAM78A_8684	9	134148528-134148765
105	FNBP1	9	132650777-132650834
106	GNG7_1972	19	2561972-2562075
107	GNG7_2119	19	2562119-2562198
108	HMHA1_9304	19	1069304-1069391
109	INPP5D	2	233925169-233925288
110	LYL1	19	13210124-13210498
111	MAX.chr15.95128172-95128228	15	95128172-95128228
112	MAX.chr16.11327022-11327151	16	11327022-11327292
113	MAX.chr16.50308415-50308535	16	50308415-50308535
114	MAX.chr4.1049805-1049912	4	1049805-1049912
115	MAX.chr9.134128109-134128241	9	134128109-134128241
116	NCOR2	12	124950727-124950905
117	OSM	22	30662704-30662800
118	S1PR4_0092	19	3180092-3180379
119	S1PR4_8637	19	3178410-3178763
120	S1PR4_9843	19	3179843-3180058

[0260] DMR序列还必须在每个链基础上在整个寻址的CpG中表现出显著的共甲基化或连续甲基化。当所有被寻址的CpG被甲基化(在病例中)和未甲基化(在对照中)时,qMSP和其他基于扩增的方法效果最好。在对标准品(亚硫酸氢盐处理的通用甲基化基因组DNA)和对照样品(亚硫酸氢盐处理的未甲基化基因组DNA、未转化的基因组DNA等)进行QC测试后,99个区域以足够的线性、特异性和稳健性进行-并用于重新测试第1阶段样品。大多数测定的逻辑分析结果与来自测序阶段的甲基化%衍生数相当。所有样品中的总DNA链超过100,其中平均在1000。与正常血沉棕黄层样品相比,Z标志物(参见例如,美国专利申请序列号14/966,617)继续表现出为1的AUC和极高的FC。当研究Z-标志物候选物的癌症与良性比率时,约一半是1比1,其余的比率在2与10之间(中-Z)。在癌症与良性标志物中,34个具有0.95至1范围内的AUC。对于血沉棕黄层样品,所有均是阴性的。在第1阶段结果中将格里森7+癌症与格里森6癌症区分开来的标志物通常在验证测试中继续其表现。26种标志物具有超过2的FC,最高为292。

[0261] 表4示出了表3中鉴定的DMR 1) 格里森评分等于或高于6的前列腺细胞相比于良性前列腺细胞的曲线下面积(AUC),2) 格里森评分等于或高于6的前列腺细胞相比于良性前列腺细胞的倍数变化(FC),和3) 格里森评分等于或高于6的前列腺细胞相比于血沉棕黄层(正常)的倍数变化(FC)。

[0262] 表4.

[0263]

DMR 号	基因注释	格里森等于 或高于 6 的 前列腺细胞 相比于良性 前列腺细胞 的 AUC	格里森等于 或高于 6 的 前列腺细胞 相比于良性 前列腺细胞 的 FC	格里森等于 或高于 6 的 前列腺细胞 相比于血沉 棕黄层 (正 常) 的 FC
1	ADCY9	0.951	15.81	64.80
2	AKR1B1_3298	0.9857	24.80	108.96
3	AKR1B1_3644	0.9948	20.00	149.37
4	ANKRD35	0.9889	10.29	74.38
5	CLIP4	0.9905	14.31	68.96
6	CYBA_7733	0.9873	10.13	100.86
7	FLJ45983	0.9575	10.02	55.98
8	GRASP	1	65.34	54.18
9	GSTP1	1	29.28	54.11
10	HAPLN3	0.9984	27.21	83.77
11	HCG4P6	0.9531	18.31	69.37
12	HES5_0822	0.9698	13.91	74.47
13	HES5_1047	0.9714	18.20	63.37
14	ITPRIPL1	1	33.50	146.96
15	LRRC4	0.9802	244.30	23.20
16	MAX.chr2.97193166-97193253	0.9968	14.23	94.29
17	MAX.chr3.193	0.9762	18.09	80.44
18	MAX.chr3.72788028-72788112	0.981	19.72	64.91
19	PTPRN2	0.9841	28.83	57.62
20	RAI1_7469	0.9984	28.53	50.04
21	RASSF2	0.9921	12.47	163.35

[0264]

22	Septin9_0074	0.9524	11.13	52.90
23	Septin9_0492	0.9627	29.96	58.55
24	SLC43A3	0.9546	10.74	62.40
25	SLC4A11	0.9921	12.12	57.37
26	SMAD3	1	17.07	87.91
27	SSBP4	0.9841	31.78	164.83
28	TJP2	0.9849	26.83	79.08
29	ABLIM1	0.9873	13.10	11.21
30	AOX1	0.9889	56.18	27.46
31	C3orf18	0.9429	11.06	33.72
32	EIF5A2	0.9635	12.42	26.45
33	EPSTI1	0.9429	13.68	40.17
34	FBXO30	0.946	31.76	87.79
35	FLOT1_1586	0.9968	8.95	57.06
36	FLOT1_1767	0.981	12.92	31.67
37	FLOT1_1904	0.9984	15.87	36.70
38	FOSL1	0.9762	13.20	27.90
39	GAS6	1	31.31	41.78
40	GSDMD	0.9635	15.23	53.83
41	KCNK4	0.9905	9.67	17.52
42	LOC100129726_1716	0.9714	31.51	114.94
43	MAX.chr10.74079656-74079694	0.9603	16.06	52.28
44	SERPINB9_3389	0.9841	13.91	54.70
45	SLCO3A1	0.9619	10.74	26.11
46	TPM4_7473	0.9832	36.60	269.51
47	TPM4_8047	0.9911	29.55	70.39
48	AGPS_7349		9.80	279.97
49	AGPS_7497		13.66	81.68
50	AGPS_7696		3.23	60.01
51	AMPD3	0.7757	29.14	115.79
52	CH25H	0.6771	6.67	311.57
53	GALR3	0.6818	5.85	50.02
54	HDAC7_6722	0.7411	10.06	50.66
55	MAX.chr1.227976339-227976430	0.7386	10.84	61.80
56	MAX.chr1.8014263-8014319	0.7556	17.68	449.60
57	MAX.chr19.34311051-34311120	0.6258	8.23	79.12
58	NCRNA00092_4149	0.6765	5.17	79.44
59	ZNF655_6084	0.8301	4.72	166.97
60	ZNF655_6545	0.8833	20.32	145.27
61	ZNF655_6762	0.7778	6.34	68.68

[0265]

62	ABCB1	1	6.76	81.06
63	ACOXL	1	3.52	117.30
64	ANXA2	1	12.61	152.60
65	APBB1IP	1	1.41	112.40
66	ARPC1B_1906	1	2.12	295.90
67	ARPC1B_1967	1	1.45	228.10
68	ATP2B4	1	4.67	117.80
69	CHST11_2032	1	18.64	228.90
70	CHST11_2206	1	29.27	201.70
71	DOK1_2325	1	2.03	99.53
72	FLJ34208	1	2.81	137.00
73	HCG4P6	1	4.27	160.20
74	HEBP2	1	3.99	72.73
75	HEYL	1	5.64	71.91
76	KLF16	1	0.97	92.50
77	LAMA3	1	8.87	41.75
78	LIME1	1	2.98	112.30
79	LOC339674	1	2.48	52.99
80	LOC440925	1	1.69	187.50
81	MAX.chr1.61519554-61519667	1	2.71	131.00
82	MAX.chr1.61519679-61519759	1	2.32	157.20
83	MAX.chr12.48592041-48592162	1	8.24	73.55
84	MAX.chr17.77786640-77786733	1	8.18	42.27
85	PDE4D	1	3.85	91.96
86	PLCL2	1	3.25	101.40
87	SIM2	1	2.45	72.63
88	STX16	1	1.76	63.07
89	WNT3A	1	2.26	55.46
90	ABHD15	1	1.18	262.70
91	ADAP1	1	1.23	251.70
92	ADD3	1	38.62	342.40
93	AXIN1	1	1.08	499.60
94	BCL2L11_6887	1	1.39	325.50
95	BCL2L11_6935	1	1.42	274.50
96	BIN2_7908	1	1.07	413.00
97	CCDC88C	1	1.42	404.30
98	CTBP1	1	1.16	606.10
99	DOK1_2096	1	2.84	272.00
100	DOK1_2234	1	2.30	280.80
101	DOK1_2475	1	1.39	280.00

[0266]	102	FAM129C	1	0.97	255.10
	103	FAM78A_1379	1	1.03	899.10
	104	FAM78A_8684	1	1.00	524.10
	105	FNBP1	1	1.01	537.80
	106	GNG7_1972	1	1.09	689.20
	107	GNG7_2119	1	1.11	1503.00
	108	HMHA1_9304	1	0.91	261.30
	109	INPP5D	1	1.01	526.70
	110	LYL1	1	1.07	790.10
	111	MAX.chr15.95128172-95128228	1	1.13	330.20
	112	MAX.chr16.11327022-11327151	1	1.48	656.90
	113	MAX.chr16.50308415-50308535	1	0.98	407.50
	114	MAX.chr4.1049805-1049912	1	1.26	362.60
	115	MAX.chr9.134128109-134128241	1	1.01	356.90
	116	NCOR2	1	0.94	438.30
	117	OSM	1	1.07	504.30
	118	S1PR4_0092	1	1.00	453.00
	119	S1PR4_8637	1	1.02	461.40
	120	S1PR4_9843	1	1.03	575.00

[0267] 实施例III.

[0268] 选择通过实施例II中描述的实验鉴定的表现最佳的候选标志物进行对从独立档案案例和对照组织提取的DNA的qMSP生物验证。

[0269] 从第2阶段研究选择了73种标志物(参见表5),以在一组独立的前列腺组织(N=35种正常前列腺,19种格里森评分6,31种格里森评分7+)和正常血沉棕黄层(N=36)样品上运行。消除的27种标志物具有低于0.85的AUC(主要在癌症与良性组中),或在Z-标志物组中不完全阳性甲基化。大多数格里森7+与6标志物均被推进了。如前所述,通过qMSP测定所有样品。对于73种测定,格里森评分7+前列腺组织与正常良性前列腺组织的DMR基因组坐标和AUC列于表5中,并且各自的引物序列在表6中提供。

[0270] 表5.

标志物	Chr	链	AUC (7+/正常)	DMR 号
-----	-----	---	-------------	-------

[0272]	ABCB1	7	RS	0.91493	62
	ABLIM1	10	FS	0.88976	29
	ACOXL	2	FS	0.9592	63
	ADCY9	16	RS	0.86372	1
	ADD3	10	FS	0.92361	92
	AGPS_7349	2	FS	0.69705	48
	AGPS_7497	2	FS	0.6849	49
	AGPS_7696	2	FS	0.75174	50
	AKR1B1_3298	7	RS	0.92622	2
	AKR1B1_3644	7	RS	0.96732	3
	ANKRD35	1	FS	0.94618	4
	ANXA2	15	RS	0.96962	64
	ARPC1B_1906	7	FS	0.84028	66
	ARPC1B_1967	7	FS	0.86024	67
	AXIN1	16	RS	0.63411	93
	BIN2_7908	12	RS	0.58854	96
	CHST11_2206	12	RS	0.97861	70
	CTBP1	4	FS	0.46007	98
	EIF5A2	3	FS	0.92014	32
	FAM78A_8684	9	RS	0.57813	104
	FBXO30	6	FS	0.82899	34
	FLJ45983	10	FS	0.99049	7
	FLOT1_1586	6	FS	0.94355	35
	FLOT1_1904	6	FS	0.92929	37
	FNBP1	9	FS	0.61198	105
	FOSL1	11	FS	0.89583	38
	GALR3	22	RS	0.78559	53
	GAS6	13	FS	0.98099	39
	GNG7_2119	19	RS	0.75955	107
	GRASP	12	RS	0.96732	8
	GSDMD	8	FS	0.93576	40
	GSTP1	11	FS	0.94792	9
	HAPLN3	15	FS	0.95781	10
	HCG4P6	6	FS	0.9836	11
	HDAC7_6722	12	RS	0.81858	54
	HEBP2	6	RS	0.93403	74
	HES5_0822	1	RS	0.95247	12
	HES5_1047	1	RS	0.94097	13
	ITPRIPL1	2	RS	0.96078	14
	KCNK4	11	FS	0.9798	41
	KLF16	19	RS	0.53993	76
	LAMA3	18	RS	0.87413	77
	LOC100129726_1716	2	RS	0.87153	42

LOC339674	22	FS	0.90885	79
LOC440925	2	FS	0.82726	80
LRRC4	7	RS	0.94593	15
LYL1	19	RS	0.56597	110
MAX.chr1.61519554-61519667	1	RS	0.95486	81
MAX.chr1.8014263-8014319	1	FS	0.80729	56
MAX.chr10.74079656-74079694	10	RS	0.86372	43
MAX.chr12.48592041-48592162	12	RS	0.9321	83
MAX.chr15.95128172-95128228	15	FS	0.58681	111
MAX.chr16.11327022-11327151	16	FS	0.81684	112
MAX.chr17.77786640-77786733	17	RS	0.78125	84
MAX.chr2.97193166-97193253	2	FS	0.98099	16
MAX.chr3.193	3	FS	0.97683	17
MAX.chr3.72788028-72788112	3	FS	0.9798	18
NCRNA00092_4149	9	FS	0.84722	58
PDE4D	5.	RS	0.8776	85
RAI1_7469	17	RS	0.97656	20
RASSF2	20		0.97861	21
S1PR4_9843	19	FS	0.52778	120
Septin9_0492	17	FS	0.93142	23
SERPINB9_3389	6	FS	0.96019	44
SIM2	21	RS	0.81343	87
SLC4A11	20	FS	0.95399	25
SLCO3A1	15	FS	0.62448	45
SSBP4	19	FS	0.9401	27
STX16	20	FS	0.87153	88
TPM4_8047	19	RS	0.95722	47
WNT3A	1	FS	0.84462	89
ZNF655_6084	7	FS	0.94271	59
ZNF655_6545	7	FS	0.80339	60

[0273]

表6.

标志物	引物序列	DMR号
ABCB1	左 M 引物 TTGTTTTTGAGTCGCGGGC (SEQ ID NO: 99) 右 M 引物 ACCAATACGATTCTCCCTCCGAT (SEQ ID NO: 100)	62
ABLIM1	左 M 引物 TTTCGACGAGTAGGATTGAAGAAGGAACG (SEQ ID NO: 127)	29

[0275]

	右 M 引物 GCGAATCTATCTACCGAAACCGCGCT (SEQ ID NO: 128)	
ACOXL	左 M 引物 AGTTAAGTTAACGGGTGTGGCGG (SEQ ID NO: 93)	63
	右 M 引物 AAACGTCGATAAAACGAACGTCGTA (SEQ ID NO: 94)	
ADCY9	左 M 引物 TTTCGGCGTTTAGGTTCGTTTC (SEQ ID NO: 25)	1
	右 M 引物 GACTCAACGATACTCCCACCGCC (SEQ ID NO: 26)	
ADD3	左 M 引物 CGAGTTGTATAGTTAGAAGAGGACGT (SEQ ID NO: 1)	92
	右 M 引物 AACCGAAAAAACCTAATTGAAACG (SEQ ID NO: 2)	
[0276]	左 M 引物 GGGGTAGAGAATGTGAAGTTTAGACGT (SEQ ID NO: 63)	48
	右 M 引物 ACCCGCGACGACTTAACGACG (SEQ ID NO: 64)	
	左 M 引物 TTTTTATTGCGCGTTAGCGGTTCG (SEQ ID NO: 65)	49
	右 M 引物 CCGCCATAACTACCGCCTATACTACCG (SEQ ID NO: 66)	
	左 M 引物 TTAATGAGTGAAAGCGCGGAGAGTCG (SEQ ID NO: 67)	50
	右 M 引物 CGAAAATAACCGTAAACGCTACCGT (SEQ ID NO: 68)	
	左 M 引物 GATTGGCGTAGATAGGGACGT (SEQ ID NO: 143)	2
	右 M 引物 ACAACCGAACTACAAATACCTCGAA (SEQ ID NO: 144)	
	左 M 引物 GGGGTTCGTTTATATTTTCGCGC (SEQ ID NO: 27)	3
	右 M 引物 CAAAATCACTCGAAATCCCTCGCC (SEQ ID NO: 28)	
ANKRD35	左 M 引物 GGGAGGGTAGTTAGTTACGGTAATACGA (SEQ ID NO: 29)	4

[0277]

	右 M 引物 CTAACCACCAACGAACCCCGAA (SEQ ID NO: 30)	
ANXA2	左 M 引物 GGGTTAAAGTTATGGGTTTATTCGT (SEQ ID NO: 89)	64
	右 M 引物 TAAAACATCTCTACGCGACCTCGTA (SEQ ID NO: 90)	
ARPC1B_1906	左 M 引物 AGGGAGTTTCGTTGGTTGTCGAC (SEQ ID NO: 81)	66
	右 M 引物 ATACTACGAACTCCGCGCTCACACG (SEQ ID NO: 82)	
ARPC1B_1967	左 M 引物 AGGGAGTTTCGTTGGTTGTCGAC (SEQ ID NO: 83)	67
	右 M 引物 ATACTACGAACTCCGCGCTCACACG (SEQ ID NO: 84)	
AXIN1	左 M 引物 GGGTATTAGTTTTATGAGATTGCGT (SEQ ID NO: 3)	93
	右 M 引物 TAAAAAAATCCACCTATCGCTCGAA (SEQ ID NO: 4)	
BIN2_7908	左 M 引物 TTGTTAATTTCGGGGTCGTCGT (SEQ ID NO: 5)	96
	右 M 引物 GAAAACCCGCACTCCTCCTCGA (SEQ ID NO: 6)	
CHST11_2206	左 M 引物 TTTTTTAGTTAGATTCTGGACGT (SEQ ID NO: 85)	70
	右 M 引物 TAAAAATAAACCCATTCTCCGAT (SEQ ID NO: 86)	
CTBP1	左 M 引物 TATAGGATTAAAGTCGCGAACGT (SEQ ID NO: 7)	98
	右 M 引物 TCAACCTACTCCTCCTATATAACGAT (SEQ ID NO: 8)	
EIF5A2	左 M 引物 ATCGTTTATCGTAGAAGTCGGCGA (SEQ ID NO: 135)	32
	右 M 引物 TACGACCTAAACTAAATCCCCCGCA (SEQ ID NO: 136)	
FAM78A_8684	左 M 引物 GGAGTTAGAAGTTTTGGGAGGGGC (SEQ ID NO: 9)	104
	右 M 引物 AAACACGTAAACCCTCTACCCGAC	

	(SEQ ID NO: 10)	
FBXO30	左 M 引物 GTTTTCGTCGGTTAATTAGCGT (SEQ ID NO: 141)	34
	右 M 引物 AAAAATAACGAATCACAAACACCGTC (SEQ ID NO: 142)	
FLJ45983	左 M 引物 TAGTCGAGGTTATGGAGGTGACGGC (SEQ ID NO: 31)	7
	右 M 引物 ACTACCCGTTAACACGACGAA (SEQ ID NO: 32)	
FLOT1_1586	左 M 引物 TGTTTCGGAAGTTTAGTTGGGGATACGT (SEQ ID NO: 121)	35
	右 M 引物 AACACCAAACGTACCTAACACGCGAA (SEQ ID NO: 122)	
FLOT1_1904	左 M 引物 GTTATTAGGATTGGTAGAAGACGAI (SEQ ID NO: 119)	37
	右 M 引物 AACTACCAATCGAAAAACCGAA (SEQ ID NO: 120)	
FNBP1	左 M 引物 GCGTGATTGATGGGTGTATTACGT (SEQ ID NO: 11)	105
	右 M 引物 ATAAACTTCCGATCCCTACAACGAA (SEQ ID NO: 12)	
FOSL1	左 M 引物 GTTAGGAAGGGAGGGACGTTCCG (SEQ ID NO: 131)	38
	右 M 引物 CGAAAAACTACGAACACGTATCGAC (SEQ ID NO: 132)	
GALR3	左 M 引物 TGTAGTAGGATACGTTGAGTCGCGG (SEQ ID NO: 69)	53
	右 M 引物 GCGATAAAACTCCACGCCGTT (SEQ ID NO: 70)	
GAS6	左 M 引物 TAGTTAGTAGAGGGAGGGTCGCGG (SEQ ID NO: 117)	39
	右 M 引物 CGCGAAAAAACGAAAATCCGTT (SEQ ID NO: 118)	
GNG7_2119	左 M 引物 TAGTTCGGGGTTGCGGTGATTAC (SEQ ID NO: 13)	107
	右 M 引物 AAAATCCGAAACAAACATTACGCC (SEQ ID NO: 14)	
GRASP	左 M 引物 TGTTTCGGATACGGCGAGC (SEQ ID NO: 33)	8

[0279]

	右 M 引物 ACGAACGAAC TACGCGACGCT (SEQ ID NO: 34)	
GSDMD	左 M 引物 GTTCGTTAGAAGGTTCGCGTCGTATAC (SEQ ID NO: 145)	40
	右 M 引物 ACCTTCCCCGAAACCTAAAAC TCCG (SEQ ID NO: 146)	
GSTP1	左 M 引物 TCGTGATTAGTATTGGGGCGGAGC (SEQ ID NO: 35)	9
	右 M 引物 GAAACTCCAACGAAAACCTCGCGAC (SEQ ID NO: 36)	
HAPLN3	左 M 引物 AAGCGGTAAAGGGAGGAATT CGGTTC (SEQ ID NO: 37)	10
	右 M 引物 GACCCCCGAAAAC TCTAACCGTCG (SEQ ID NO: 38)	
HCG4P6	左 M 引物 GGATCGGAGTATTGGGATCGGAGTATAC (SEQ ID NO: 39)	11
	右 M 引物 AACTCTAATAATAAACGAC GCGAC (SEQ ID NO: 40)	
HDAC7_6722	左 M 引物 GTAGTATT TATT TATCGGGCGA (SEQ ID NO: 71)	54
	右 M 引物 AAAAAATCACACCTCCTTCAACGCT (SEQ ID NO: 72)	
HEBP2	左 M 引物 TTCGAGGTTTTCGGGCGAC (SEQ ID NO: 103)	74
	右 M 引物 CCCTCCTATCCGTTAACTT CGCGTA (SEQ ID NO: 104)	
HES5_0822	左 M 引物 GTGATT CGGCGGGATTGCG (SEQ ID NO: 41)	12
	右 M 引物 GACGAAAAAACGCTTCCCTACAAACGA (SEQ ID NO: 42)	
HES5_1047	左 M 引物 AAGAGTTGTATTAGGATTATAGCGA (SEQ ID NO: 43)	13
	右 M 引物 CGTAAAACGTCAAAAAC TACAC GAC (SEQ ID NO: 44)	
ITPRIPL1	左 M 引物 GGGATTAGGGTTAGGTTATTATCGT (SEQ ID NO: 45)	14
	右 M 引物 ACCCGC GTATCTCTTAAAATCGTA	

	(SEQ ID NO: 46)	
KCNK4	左 M 引物 TTAGTTAGGAAGTAGGGCGAGGCGA (SEQ ID NO: 125)	41
	右 M 引物 AAAACCGAACACGCAAAACGAA (SEQ ID NO: 126)	
KLF16	左 M 引物 TGAGATTCCGGTAAAGGAAGGGTC (SEQ ID NO: 95)	76
	右 M 引物 ACCTACCTATACGCCTCCAAACGAT (SEQ ID NO: 96)	
LAMA3	左 M 引物 GTTTTATTGGATGGTGTGGTCGT (SEQ ID NO: 115)	77
	右 M 引物 CCCGAACCTCTACTACAAATTAAATCGAA (SEQ ID NO: 116)	
LOC100129726_1716	左 M 引物 TACGTCGTCGTTATTAGATTATAATTTCGTC (SEQ ID NO: 133)	42
	右 M 引物 CCAAACCCCTACTACTACTACTACTAACTAC G (SEQ ID NO: 134)	
[0280] LOC339674	左 M 引物 TGGTGGATCGCGATTTCGTAAGAC (SEQ ID NO: 111)	79
	右 M 引物 CGCCGAAAACCAAATTATCGCG (SEQ ID NO: 112)	
LOC440925	左 M 引物 CGGTGAGTATT CGCGGTTTCGT (SEQ ID NO: 87)	80
	右 M 引物 AAATCGCCTCTCCCGAACGC (SEQ ID NO: 88)	
LRRC4	左 M 引物 TAATT CGCGAGGTAGGCGACGG (SEQ ID NO: 47)	15
	右 M 引物 CAATACTCTTATATATTACGCCGCT (SEQ ID NO: 48)	
LYL1	左 M 引物 TGTTCGTTAGTTATGAAGTATATCGG (SEQ ID NO: 15)	110
	右 M 引物 ACTTAATCGCGCAACAAACCGCA (SEQ ID NO: 16)	
MAX.chr1.61519 554-61519667	左 M 引物 GTCGC GTTTTTATTGTCGTTCGT (SEQ ID NO: 91)	81
	右 M 引物 ATAAAAATCGATTCTACCGCGTCGCC (SEQ ID NO: 92)	

[0281]	MAX.chr1.80142 63-8014319	左 M 引物 AATACGCGACGGTTCGTTATTGC (SEQ ID NO: 73)	56
		右 M 引物 CATAACGTAAATCCACTTCCGACGAC (SEQ ID NO: 74)	
	MAX.chr10.7407 9656-74079694	左 M 引物 GGTTGTAAGGGGGTTGGGTACGC (SEQ ID NO: 139)	43
		右 M 引物 ATTCGAAAAAAACGCCGATACGA (SEQ ID NO: 140)	
	MAX.chr12.4859 2041-48592162	左 M 引物 TCGCGTTGACGGTTGTGACG (SEQ ID NO: 101)	83
		右 M 引物 AACCCCGATCCGAAAAACCGAA (SEQ ID NO: 102)	
	MAX.chr15.9512 8172-95128228	左 M 引物 CGGTTATATTATAAGAAAAGGAAGTTTCGT (SEQ ID NO: 17)	111
		右 M 引物 GAAAACCCAAACTACACACCCGCT (SEQ ID NO: 18)	
	MAX.chr16.1132 7022-11327151	左 M 引物 TTGGTTTTATTAAGTTATGTGACGA (SEQ ID NO: 19)	112
		右 M 引物 TAAAATTCCAAAAACGATAACGCT (SEQ ID NO: 20)	
	MAX.chr17.7778 6640-77786733	左 M 引物 GGGTGGATTTCGGGCGTTATAAATC (SEQ ID NO: 113)	84
		右 M 引物 CAAAACGACTCCCCGCCGAA (SEQ ID NO: 114)	
	MAX.chr2.97193 166-97193253	左 M 引物 GGTTTAGGGAAATATCGGCGT (SEQ ID NO: 49)	16
		右 M 引物 AACTCAAACCGAAAAATAATTGAT (SEQ ID NO: 50)	
	MAX.chr3.193	左 M 引物 AAAGGTTAGTTAAAGATGGAATCGT (SEQ ID NO: 51)	17
		右 M 引物 CTCGCGACGAAAAACCCGAA (SEQ ID NO: 52)	
	MAX.chr3.72788 028-72788112	左 M 引物 AGGATTGACGGAGTTATTCGT (SEQ ID NO: 53)	18
		右 M 引物 TAACATAACCACCCAACTCTCCCCGAA (SEQ ID NO: 54)	
	NCRNA00092_4 149	左 M 引物 CGTTAGGGGGTTCGAGCGTAGC (SEQ ID NO: 75)	58
		右 M 引物 CCCTAATTCCATCCTAAACGAATCGAC	

	(SEQ ID NO: 76)	
PDE4D	左 M 引物 AATTTCGTAGGAAGTAGTCGGTCGT (SEQ ID NO: 97)	85
	右 M 引物 TACCCTCCAAATTACCCAAACCGCT (SEQ ID NO: 98)	
RAI1_7469	左 M 引物 TTATAGTAGTCGTCGAAATATAAAGTTTCGT C (SEQ ID NO: 55)	20
	右 M 引物 CCGAAAAACCCAAAAAAACCCG (SEQ ID NO: 56)	
RASSF2	左 M 引物 GTCGTTACGGTATTGTTCGTTC (SEQ ID NO: 57)	21
	右 M 引物 ATCGCTAAAACCTAACCTAACGTC (SEQ ID NO: 58)	
S1PR4_9843	左 M 引物 TGAGGTTAAGGGATAGTTTCGCGG (SEQ ID NO: 21)	120
	右 M 引物 AACTACAACCTCAAATACTCCGCACGCT (SEQ ID NO: 22)	
[0282]	左 M 引物 GGTTTGCGTTTGCCTCGC (SEQ ID NO: 23)	23
	右 M 引物 CCATTATATAACTTCCCCTTCGCC (SEQ ID NO: 24)	
	左 M 引物 GGTTTATTATTCGGTGGTAGTCGG (SEQ ID NO: 129)	44
	右 M 引物 ACACGAAAACGACGACAACGCT (SEQ ID NO: 130)	
	左 M 引物 GGTCGTAGTCGGGAAGTTCGG (SEQ ID NO: 105)	87
	右 M 引物 ATTGACGAAAACACCGCGC (SEQ ID NO: 106)	
	左 M 引物 TTCGGTTGTTGTGTTCGTTGTCG (SEQ ID NO: 59)	25
	右 M 引物 CCGAATCTAAAAACGCTTCCTCTCGTA (SEQ ID NO: 60)	
	左 M 引物 GCGATTGAGCGTTGGATATCGA (SEQ ID NO: 137)	45
	右 M 引物 GTCGAAAACCCACGAACCGTT (SEQ ID NO: 138)	
SSBP4	左 M 引物 GTAGCGTCGGCGTAGAGCGTAGATC	27

	(SEQ ID NO: 61)	
	右 M 引物 ATAACGAATCCCCGCGAAAATTG (SEQ ID NO: 62)	
STX16	左 M 引物 CGGGAAATTTCGGAAAATATATACGT (SEQ ID NO: 107)	88
	右 M 引物 TACGAAATTCCAACAAAAAACGAA (SEQ ID NO: 108)	
TPM4_8047	左 M 引物 TTTTTTATTTTTATTTTTCGTCGT (SEQ ID NO: 123)	47
	右 M 引物 GACTTCTACTTACTTCCCTAACCGTT (SEQ ID NO: 124)	
[0283]	左 M 引物 CGGGCGGTTATACGTTTCGC (SEQ ID NO: 109)	89
WNT3A	右 M 引物 AAAAAAAATCCCCATTCAAACGCT (SEQ ID NO: 110)	
ZNF655_6084	左 M 引物 TTGCGAAAACGAGTTTCGAATTATGGAC (SEQ ID NO: 77)	59
	右 M 引物 CCCCCGAATATAATAACGACCCCCGAA (SEQ ID NO: 78)	
ZNF655_6545	左 M 引物 TTGGTTATTTATTCGCGGATCGA (SEQ ID NO: 79)	60
	右 M 引物 AAAACACGATCGCCGACTCCTAACG (SEQ ID NO: 80)	

[0284] AUC总体上是优秀的,虽然比早期的验证稍微少一些-这是在这个扩展的独立集合中预期的。如表5所示,20种标志物具有在0.95-0.99范围内的AUC值(格里森7+对比正常前列腺)和FC 17-164。除了存在单个异常值的GRASP之外,血沉棕黄层样品中的甲基化%可忽略不计。

[0285] 由于临床随访数据可用于前列腺病例,因此决定探索表观遗传标志物候选者的预后方面。使用回归分区(rPart),用于发现和/或建模逻辑模型内标志物之间的高阶相互作用的数学方法,选择五种预后标志物(FAM78A、WNT3A、GAS6、LOC100129726和MAX.chr3.727)。由甲基化DNA标志物定义的风险分组在预测相对于格里森评分的无进展存活方面增加了显著的预后内容($p<0.0001$),而格里森评分相对于甲基化DNA标志物风险分组没有增加值($p=0.2174$)。

[0286] 实施例IV.

[0287] 进行另外的实验以鉴定能够区分前列腺组织中PCa格里森评分高于7与格里森评分为6的标志物。此类实验利用QuARTs-X(定量等位基因特异性实时靶标和信号测定)(参见例如,美国专利申请序列号15/335,096)。表7示出了100%的标志物灵敏度和前列腺组织中PCa格里森7+相比于格里森6的倍数变化(寡核苷酸序列在表8中提供)。

[0288] 表7

DMR 号	标志物	前列腺组织样 品中格里森高 于 7 相比于格里 森 6 的灵敏度	前列腺组织样品中 格里森高于 7 相比 于格里森 6 的倍数 变化
[0289]	17 MAX.chr3.193	90%	1.26
	12 HES5_0822	84%	1.5
	45 SLCO3A1	70%	2
	47 TPM4_8047	68%	1.42
	83 MAX.chr12.48592041-48592162	86%	1.26
	39 GAS6	86%	1.02
	18 MAX.chr3.72788028-72788112	78%	1.37
	3 AKR1B1_3644	66%	1.18
	87 SIM2	34%	1.7
	11 HCG4P6	86%	0.99
	70 CHST11_2206	88%	1.7
	8 GRASP	82%	1.59
	44 SERPINB9_3389	74%	1.49
	85 PDE4D	62%	1.08
	14 ITPRIPL1	76%	2.16
	16 MAX.chr2.97193166-97193253	82%	1.34
	37 FLOT1_1904	86%	1.44

[0290] 表8

标志物	寡核苷 酸类型	寡核苷酸名称	序列
SERPINB9	引物	SERPINB9_FP	TTTATTTAGTCGTGCGCGG

[0292]	SERPINB9	引物	SERPINB9_RP	(SEQ ID NO: 147) ACGCGACACGAAAACGAC (SEQ ID NO: 148)
		探针	SERPINB9_Pb_A5	CCACGGACG CGACAACGCTAA/3C6/ (SEQ ID NO: 149)
		引物	FLOT1_FP	GTAGTGTGAGTTAAGTT GTTTCG (SEQ ID NO: 150)
	FLOT1	引物	FLOT1_RP	AACACCAAACGTACCTAATAC G (SEQ ID NO: 151)
		探针	FLOT1_Pb_A1	CGCCGAGG GCGAAAACGTAT/3C6/ (SEQ ID NO: 152)
		引物	HCG4P6_FP	CGGAGTATACGGAATATTAGG TTCG (SEQ ID NO: 153)
	HCG4P6	引物	HCG4P6_RP	AATAATAACGACGCGACATAA ACA (SEQ ID NO: 154)
		探针	HCG4P6_Pb_A5	CCACGGACG GCGTATAGATTG/3C6/ (SEQ ID NO: 155)
		引物	CHST11_2206_FP	GCGTCGAGGGCGTTT (SEQ ID NO: 156)
CHST11_2206	引物	CHST11_2206_RP		AATCTAAACTAAAAAA CGAAACTCGT (SEQ ID NO: 157)
		探针	CHST11_2206_Pb_A1	CGCCGAGG TCGCTTCCTAAA/3C6/ (SEQ ID NO: 158)
	MAX.chr12.485	引物	MAX.chr12.485_F P	GAGGAAATAGCGTGGTTCG (SEQ ID NO: 159)
		引物	MAX.chr12.485_R P	AAAACCGAAAAAAACTAATT GTCA (SEQ ID NO: 160)
		探针	MAX.chr12.485_P b_A5	CCACGGACG GCGTTGACGGTT/3C6/ (SEQ ID NO: 161)
GRASP	引物	GRASP_FP		CGCGCGGTATAGTCGG (SEQ ID NO: 162)
	引物	GRASP_RP		ACCCCCAACGAACGAAC (SEQ ID NO: 163)
	探针	GRASP_Pb_A1		CGCCGAGG CTATACCGCGACG/3C6/ (SEQ ID NO: 164)
GAS6	引物	GAS6_FP		GATTGAGGTTAGGTTTCGT

[0293]	GAS6	引物	GAS6_RP	CG (SEQ ID NO: 165) GAAAAACAAACGCTACTACCGC (SEQ ID NO: 166)
		探针	GAS6_Pb_A5	CCACGGACG GCGCGGAGTTGG/3C6/ (SEQ ID NO: 167)
		引物	HAPLN3_FP	GTTCGTATATAGTTTAGAAG TTTAGCGA (SEQ ID NO: 168)
	HAPLN3	引物	HAPLN3_RP	CGTCGAAAAACTACCTAAAAA CGAT (SEQ ID NO: 169)
		探针	HAPLN3_Pb_A1	CGCCGAGG TACGTAACCTAA/3C6/ (SEQ ID NO: 170)
		引物	SLCO3A1_FP	CGATTTGTAGCGTTGGATA TCG (SEQ ID NO: 171)
	SLCO3A1	引物	SLCO3A1_RP	TCGAAAACCCCACGAACC (SEQ ID NO: 172)
		探针	SLCO3A1_Pb_A5	CCACGGACG CGTTAATACCCC/3C6/ SEQ ID NO: 173)
		引物	MAX.chr3.193_FP	TTAAAGGTTAGTTAAAGATG GAATCGT (SEQ ID NO: 174)
93	MAX.chr3.1	引物	MAX.chr3.193_RP	CTCGCGACGAAAAAAACCC (SEQ ID NO: 175)
		探针	MAX.chr3.193_Pb_A1	CGCCGAGG CGAACTCCCAAC/3C6/ (SEQ ID NO: 176)
		引物	MAX.chr3.727_FP	GTGGTTTATTCGTTCGTT CG (SEQ ID NO: 177)
27	MAX.chr3.7	引物	MAX.chr3.727_RP	AAACTAACGAATATAACTCCG TCGA (SEQ ID NO: 178)
		探针	MAX.chr3.727_Pb_A1	CGCCGAGG GCGACGTTGAG/3C6/ (SEQ ID NO: 179)
		引物	HES5_FP	GCGAGAGGAGTAGGTTCG (SEQ ID NO: 180)
HES5	引物	HES5_RP	AACCTACGAACGCGCGA (SEQ ID NO: 181)	
	探针	HES5_Pb_A5	CCACGGACG ACCCGACGACCA/3C6/ (SEQ ID NO: 182)	
	PDE4D	引物	PDE4D_FP	GCGTACGGTCGCGTATT (SEQ ID NO: 183)

[0294]	PDE4D	引物	PDE4D_RP	CAATAACTCGACGAAACGCG (SEQ ID NO: 184)
		探针	PDE4D_Pb_A1	CGCCGAGG GCGAAATTCTAA/3C6/ (SEQ ID NO: 185)
	SIM2	引物	SIM2_FP	GTTTAGCGCGGGTTTCG (SEQ ID NO: 186)
		引物	SIM2_RP	CCCGAACTTCCCAGACT (SEQ ID NO: 187)
		探针	SIM2_Pb_A5	CCACGGACG GCGGTAGTGGTC/3C6/ (SEQ ID NO: 188)
	TPM4	引物	TPM4_FP	GGAGAAAGGCAGGGCG GA (SEQ ID NO: 189)
		引物	TPM4_RP	CGTTCCCAAAACGCGA (SEQ ID NO: 190)
		探针	TPM4_Pb_A1	CGCCGAGG ACGACGACGTAT/3C6/ (SEQ ID NO: 191)
	LRRC4_HC C	引物	LRRC4_HCC_FP	CGTCGTTCGTTCGTTTG (SEQ ID NO: 192)
		引物	LRRC4_HCC_RP	CCCCGCCCCCTCTAAAC (SEQ ID NO: 193)
		探针	LRRC4_HCC_Pb_A5	CCACGGACG CGAACGAAACGA/3C6/ (SEQ ID NO: 194)
	ITPRIPL1	引物	ITPRIPL1_FP	GTTAGTGGCGGTTAGGTAG (SEQ ID NO: 195)
		引物	ITPRIPL1_RP	TCACCGTCAATATTAATAAAAA ACACGA (SEQ ID NO: 196)
		探针	ITPRIPL1_Pb_A1	CGCCGAGG GCGTTGTTTTT/3C6/ (SEQ ID NO: 197)
	MAX.chrchr 2.97193166- 97193253	引物	MAX.chr2.chr2.97 193166-97193253_- FP	GCGAGGTTGCGGTTTG (SEQ ID NO: 198)
		引物	MAX.chr2.chr2.97 193166-97193253_- RP	CGAAAAAAATAATTGATTAAC CAAAACGC (SEQ ID NO: 199)
		探针	MAX.chr2.chr2.97 193166-97193253_- Pb_A5	CCACGGACG CCGATATTCCC/3C6/ (SEQ ID NO: 200)
	AKR1B1_3 644	引物	AKR1B1_FP	GTTCGTTTATATTTTCGCGC G

[0295]		引物	AKR1B1_RP	(SEQ ID NO: 201) CCGAACGTCCCGCAAAC (SEQ ID NO: 202)
		探针	AKR1B1_Pb_A1	CGCCGAGG CGAACTACTCAA/3C6/ (SEQ ID NO: 203)
		引物	FLJ45983_FP	GGGCGCGAGTATAGTCG (SEQ ID NO: 204)
	FLJ45983	引物	FLJ45983_RP	CAACGCGACTAATCCGC (SEQ ID NO: 205)
		探针	FLJ45983_Pb_A1	CGCCGAGG CCGTCACCTCCA/3C6/ (SEQ ID NO: 206)

[0296] 实施例V.

[0297] 进行实验,其中将候选标志物序列在计算机上跨泛癌RRBS测序数据集进行比较,以测量每个标志物的位点特异性甲基化程度。

[0298] DNA甲基化特征已被证明可准确预测人体解剖结构中肿瘤的部位。为了更好地确定器官部位相关的特异性,构建了针对多个癌症和器官组织中的73种DMR/标志物中的每个的RRBS衍生的甲基化值的计算机上CpG x样品基质。这些包括前列腺、肝脏、结肠直肠、胰腺、肺、食道、胃和胆管组织。可通过整个DMR中器官部位之间的定量甲基化差异以及连续的甲基化程度(由模式识别定义)来模拟位置特异性。如表9所示,8种标志物仅表明前列腺癌的特异性-而在另一个极端,11种标志物对所有癌症和组织均是通用的。其间是具有相同或不同程度特异性的簇,如前列腺/肝脏、前列腺/结肠/肝脏等。由于在其他癌症的比对读数中缺失DMR序列,前列腺标志物的子集仍然未定义。

[0299] 表9.

[0300]

DMR号	标志物	格里森高于 7 的前列腺组织 相比于良	格里森高于 7 的前列腺组织 相比于良	格里森等于或高于 6 的前列腺组织相 比于良	甲基化% (7+)	组织特异性
------	-----	------------------------	------------------------	---------------------------	--------------	-------

[0301]

		性前列腺组织的AUC	性前列腺组织的FC	比于血沉棕黄层的FC		
62	ABCB1	0.91493	16	3496	15	前列腺/肺
29	ABLIM1	0.88976	39	51	5	无数据
63	ACOXL	0.9592	18	59094554	37	前列腺/肺/胰腺
1	ADCY9	0.86372	19	38	4	无数据
92	ADD3	0.92361	2290	3075	20	前列腺特异性
48	AGPS_7349	0.69705	12	554	3	无数据
49	AGPS_7497	0.6849	7	7534	2	无数据
50	AGPS_7696	0.75174	58	52722	13	无数据
2	AKR1B1_3298	0.92622	32	604	37	前列腺/结肠/HCC
3	AKR1B1_3644	0.96732	128	6850	37	前列腺/结肠/HCC
4	ANKRD35	0.94618	5	4184	48	通用
64	ANXA2	0.96962	34	36464	66	前列腺/HCC
66	ARPC1B_1906	0.84028	5	45311	16	前列腺特异性
67	ARPC1B_1967	0.86024	8	5800	23	前列腺特异性
93	AXIN1	0.63411	1	16677	143	通用
96	BIN2_7908	0.58854	1	10778	90	通用
70	CHST11_2206	0.97861	145	41729	27	前列腺特异性
98	CTBP1	0.46007	1	1188	114	通用
32	EIF5A2	0.92014	32	7138	16	无数据
104	FAM78A_8684	0.57813	1	8026	55	通用
34	FBXO30	0.82899	13	1726	19	无数据
7	FLJ45983	0.99049	35	1480	54	前列腺/轻度通用
35	FLOT1_1586	0.94355	11	15687	64	前列腺特异性
37	FLOT1_1904	0.92929	65	322	38	前列腺特异性
105	FNBP1	0.61198	1	78760	97	通用
38	FOSL1	0.89583	69	481	5	无数据
53	GALR3	0.78559	31	34071	15	无数据
39	GAS6	0.98099	75	62587	47	前列腺特异性
107	GNG7_2119	0.75955	1	19012	100	通用

[0302]

8	GRASP	0.96732	78	37	30	前列腺/结肠 /HCC/肺
40	GSDMD	0.93576	28	14825	20	无数据
9	GSTP1	0.94792	21	361117	77	前列腺/HCC
10	HAPLN3	0.95781	60	969498	38	前列腺特异 性
11	HCG4P6	0.9836	30	73301	58	前列腺/轻度 通用
54	HDAC7_6722	0.81858	149	11130	12	无数据
74	HEBP2	0.93403	14	10185	21	前列腺特异 性
12	HES5_0822	0.95247	26	8884	29	前列腺特异 性
13	HES5_1047	0.94097	13	12636	28	前列腺特异 性
14	ITPRIPL1	0.96078	45	12163	19	前列腺/结肠 /HCC/肺
41	KCNK4	0.9798	34	682	29	无数据
76	KLF16	0.53993	1	1155	21	前列腺/HCC
77	LAMA3	0.87413	79	106	18	前列腺特异 性
42	LOC10012972 6_1716	0.87153	11	2594	16	无数据
79	LOC339674	0.90885	14	447	29	前列腺/HCC/ 肺 (仅小细 胞)
80	LOC440925	0.82726	3	2215	48	前列腺/肺
15	LRRC4	0.94593	61	19639	48	前列腺/结肠/ 肺/胰腺
110	LYL1	0.56597	1	23253	59	无数据
81	MAX.chr1.615 19554-615196 67	0.95486	5	5254	58	前列腺/肺
56	MAX.chr1.801 4263-8014319	0.80729	25	1770	10	前列腺/肺
43	MAX.chr10.74 079656-74079 694	0.86372	50	278757	14	无数据
83	MAX.chr12.48 592041-48592 162	0.9321	26	364	63	前列腺特异 性
111	MAX.chr15.95 128172-95128 228	0.58681	1	1363	104	通用
112	MAX.chr16.11 327022-11327	0.81684	3	31278	324	通用

[0303]	151					
	MAX.chr17.77 786640-77786 733	0.78125	16	67	7	前列腺/肺(仅 小细胞)
	MAX.chr2.971 93166-971932 53	0.98099	17	967	47	前列腺/结肠 /HCC/肺
	MAX.chr3.193	0.97683	43	1605	66	前列腺特异 性
	MAX.chr3.727 88028-727881 12	0.9798	87	633	52	前列腺/HCC
	NCRNA00092 4149	0.84722	11	1101	21	无数据
	PDE4D	0.8776	4	835	59	前列腺特异 性
	RAI1_7469	0.97656	28	320	56	通用
	RASSF2	0.97861	24	1528628	44	前列腺/结肠 /HCC
	S1PR4_9843	0.52778	1	722	86	通用
	Septin9_0492	0.93142	21	3348	33	无数据
	SERPINB9_33 89	0.96019	164	3993	50	前列腺特异 性
	SIM2	0.81343	3	323	51	前列腺特异 性
	SLC4A11	0.95399	20	768	16	前列腺/结肠 /HCC
	SLCO3A1	0.62448	4	6712	71	无数据
	SSBP4	0.9401	34	1043	25	前列腺/结肠 /HCC
	STX16	0.87153	4	279	34	前列腺/结肠
	TPM4_8047	0.95722	23	2578	31	前列腺特异 性
	WNT3A	0.84462	5	2144	28	前列腺/HCC
	ZNF655_6084	0.94271	34	(分母为 零)	11	无数据
	ZNF655_6545	0.80339	5		14	无数据

[0304] 实施例VI.

[0305] 进行实验,其中选择一组高效PCa标志物用于在盲法独立血浆样品中进行测试,以评估临床培养基中的PCa检测。

[0306] 选择二十五种DMR/标志物(参见表10)(其表现出适合于复杂生物介质(例如血浆)中的分析物检测的性能指标的最佳组合)用于开发多重QuARTs-X(定量等位基因特异性实时靶标和信号分析)-一种经常与其他癌症血浆研究一起使用的测试平台。在对合并血浆对照的DMR序列、设计过滤器和QC测试的独特性进行初步计算机测试后,17个设计被用于对来自Mayo前列腺癌生物库的回顾性收集的冷冻血浆样品进行测试。开发了9个三链体和1个双链体反应,每个反应均包括不受甲基化影响的对照β-肌动蛋白测定。还测试了另外两个过

程控制。QuARTs-X测定列于表10中(引物和探针信息见表8)。将最终标志物链(拷贝)归一化为 β -肌动蛋白对照,并表示为甲基化%。最佳的4个标志物是max.chr3.193、HES5、SLC03A1和TPM4_8047。结合使用rPart建模,血液样品中检测PCa的灵敏度和特异性分别是78%和91%。

[0307] 表10

[0308]

DMR号	标志物	AUC	p值
17	MAX.chr3.193	0.84539	0.1658
12	HES5_0822	0.76068	>0.0001
45	SLC03A1	0.67209	0.7921
47	TPM4_8047	0.63612	0.227
83	MAX.chr12.48592041-48592162	0.56359	0.037
39	GAS6	0.56073	0.0451
18	MAX.chr3.72788028-72788112	0.53893	0.2411
3	AKR1B1_3644	0.52	0.0165
87	SIM2	0.5101	0.3577
11	HCG4P6	0.5067	0.2852
70	CHST11_2206	0.50524	0.1351
8	GRASP	0.50282	0.2477
44	SERPINB9_3389	0.49417	0.1046
85	PDE4D	0.48252	0.9368
14	ITPRIPL1	0.47	0.7214
16	MAX.chr2.97193166-97193253	0.45301	0.7152
37	FLOT1_1904	0.42456	0.7204

[0309] 实施例VII.

[0310] 前列腺癌的治疗决策经常由格里森等级引导,这是主观的并且缺乏精确性。在发现和早期验证中,甲基化DNA标志物(MDM)被鉴定为具有预后关联(参见实施例I)。进行进一步的实验来评估新型MDM在使用来自独立组的档案组织预测生化复发方面的价值,其中根治性前列腺切除术(RP)后随访>12年。

[0311] 从2004年经历根治性前列腺切除术(RP)的737名男性中,随机选择446名,并且155名符合质量标准。利用福尔马林固定石蜡包埋(FFPE)组织块。一位专家病理学家使用更新的格里森标准以及用于宏观解剖的标记肿瘤以盲法重新审查所有样本。使用QiaAmp FFPE组织试剂盒(Qiagen)纯化基因组DNA,并通过Picogreen荧光定量。由于FFPE DNA可高度降解,因此还用100bp β -肌动蛋白扩增测定测试样品的可扩增基因组等同物。然后用亚硫酸氢钠处理DNA并纯化(Zymo研究)。

[0312] 选择二十三种MDM来测试样品。通过对使用73种MDM的独立组织验证结果运行递归分区分析(rPART)来得到二十三种MDM。所述研究(实施例I)中使用的所有患者在其临床记录中具有结果数据。具体地,实验运行了1000个引导带样品rPART(在计算机上)并寻找在建模树中最常出现的MDM。然后按频率(从高到低)对这些进行排序,并选择最佳MDM(参见表11)。如前所述,对普遍甲基化标准品和适当的阴性和阳性对照的稀释液盲法运行MSP测定。

将原始计数归一化为每种样品的总 β -肌动蛋白计数。将复发定义为PSA>0.4ng/mL。通过回归分区树模型选择最佳MDM以指定复发风险并通过四分位数分组(M1(最低)到M4(最高))。评估MDM和格里森等级组(GGG)的预后价值,并基于其与RP后结果的一致性进行比较。将以下标志物鉴定为预测复发率的最佳标志物:WNT3A、LOC100129726、FNBP1、GSDMD、ITPRIPL1、Chr1.61519554和Chr17.77786040。

[0313] 表11.

基因	模型数量

WNT3A	582
LOC100129726	496
FAM78A	476
KCNK4	438
KLF16.RS_FP	415
CHST11_2206	413
ARPC1B1906	387
MAX.chr1.61519554	383
ABCB1	377
AGPS_7696	375
LOC339674	368
SLCO3A1	368
LAMA3	367
MAX.chr17.77786640	366
AGPS_7497	352
MAX.chr3.727	351
GSDMD	339
GRASP	332 最佳 25%
LOC440925	326
SERPINB9_3389	319
EIF5A2	318
MAX.chr10.74079656	315
MAX.chr15.95128172	311
SLC4A11	311
ADD3	309
ABLIM1	307
ACOXL	306
ARPC1B_1967	294
ANXA2	293
FLOT1_1586	292
AKR1B1	287
HEBP2	287
S1PR4_9843	287
FOSL1	285
MAX.chr3.193776187	282
HCG4P6	277
ANKRD35	275 最佳 50%
GAS6	270
LYL1	267
AGPS_7349	265
PDE4D	264
MAX.chr2.97193166	249
RAI1_7469	237
FBXO30	232

[0315]

[0316]

STX16	232
ZNF655_6084	231
HAPLN3	228
MAX.chr1.8014263	228
FLOT1_1904	225
TPM4_8047	219
HDAC7_6722	218
SIM2	216
SSBP4	215
RASSF2	214
MAX.chr16.11327022	210
GALR3	209
AKR1B1_3644	208
ITPRIPL1	204
Septin9	200
ADCY9	196
AXIN1	188
LRRC4	185
GNG7_2119	182
CTBP1	175
ZNF655_6545	170
FLJ45983	166
HES5_1047	162
NCRNA00092	156
GSTP1	150
HES5_0822	148
FNBP1	132
BIN2_7908	128

[0317] 实施例VIII.

[0318] 进行另外的实验以鉴定能够区分1) 前列腺组织中PCa格里森评分高于7与格里森评分为6, 和2) PCa格里森评分高于6与非癌前列腺组织的标志物。此类实验利用QuARTs-X (定量等位基因特异性实时靶标和信号测定) (参见例如, 美国专利申请序列号15/335, 096)。

[0319] DNA提取

[0320] 具有已知临床信息的冷冻DNA组织样品获自Mayo Clinic储存库。使用来自Qiagen的DNeasy Blood&Tissue试剂盒, 按照制造商的方案从组织提取DNA。将大约100ng提取的DNA转移到亚硫酸氢盐转化反应中。

[0321] 亚硫酸氢盐转化和DNA纯化

[0322] 试剂制备:

[0323] 表12

组分缩写	名称	制剂
BIS SLN	亚硫酸氢盐转化溶液	56.6%亚硫酸氢铵
DES SLN	脱磺化溶液	70%异丙醇, 0.1 N NaOH
BND BDS	结合珠	Maxwell RNA 珠 (16 mg/mL), Promega
BND SLN	结合溶液	7 M 盐酸胍
CNV WSH	转化洗涤	10 mM Tris-HCl, 80%乙醇, 0.01%叠氮化钠
ELU BUF	洗脱缓冲液	10mM Tris, 0.1 mM EDTA, pH 8.0

[0325] 阻断溶液:

[0326] 表13

样品数量 (N)	NaOH (1.6 N) (uL)	BSA (350 ng/uL) (uL)
1	5	5
N		

[0328] 亚硫酸氢盐转化和DNA纯化遵循以下程序:

[0329] 1. 向深孔板 (DWP) 的每个孔中加入10μL阻断溶液。

[0330] 2. 将80μL每种样品加入DWP中。

[0331] 3. 通过用移液管组移液至30-40μL小心混合以避免气泡。

[0332] 4. 将DWP密封并以3000xg离心1分钟。

[0333] 5. 在42°C下孵育20分钟。

[0334] 6. 向每个孔中加入120μL BIS SLN。

[0335] 7. 8min冷却

[0336] 8. 在65°C下孵育75分钟, 同时在前3分钟内混合。

[0337] 9. 加入750μL BND SLN

[0338] 10. 预混合二氧化硅珠 (BND BDS) 并向DWP的孔中加入50μL二氧化硅珠 (BND BDS)。

[0339] 11. 在加热器振荡器上在30°C下以1,200rpm混合30分钟。

[0340] 12. 将珠在板磁铁上结合5分钟, 随后将溶液抽吸至废液。

[0341] 13. 加入1mL洗涤缓冲液 (CNV WSH), 然后将板移至加热器振荡器并以1,200rpm混合3分钟。

[0342] 14. 将珠在板磁铁上结合5分钟, 随后将溶液抽吸至废液。

[0343] 15. 加入0.25mL洗涤缓冲液 (CNV WSH), 然后将板移至加热器振荡器并以1,200rpm混合3分钟。

[0344] 16. 将珠在磁铁上结合2分钟, 随后将溶液抽吸至废液。

[0345] 17. 加入0.2mL脱磺化缓冲液 (DES SLN) 并在30°C下以1,200rpm混合7分钟。

[0346] 18. 将珠在磁铁上结合2分钟, 随后将溶液抽吸至废液。

[0347] 19. 加入0.25mL洗涤缓冲液 (CNV WSH), 然后将板移至加热器振荡器并以1,200rpm混合3分钟。

[0348] 20. 将珠在磁铁上结合2分钟, 随后将溶液抽吸至废液。

[0349] 21. 加入0.25mL洗涤缓冲液 (CNV WSH), 然后将板移至加热器振荡器并以1,200rpm

混合3分钟。

[0350] 22. 将珠在磁铁上结合2分钟,随后将溶液抽吸至废液。

[0351] 23. 通过移至加热器振荡器使板干燥并在70°C下孵育15分钟,同时以1,200rpm混合。

[0352] 24. 在DWP中的所有样品中加入80μL洗脱缓冲液 (ELU BFR)。

[0353] 25. 在65°C下孵育25分钟,同时以1,200rpm混合。

[0354] 26. 手动将洗脱液转移至96孔板,用箔密封件密封板,然后储存在-80°C下

[0355] 27. 可恢复/可转移体积是约65μL。

[0356] QuARTS-X用于甲基化DNA检测和定量

[0357] 多重PCR (mPCR) 设置:

[0358] 1- 为每种感兴趣的甲基化标志物制备含有正向和反向引物的10X引物混合物,每种最终浓度各自为750nM。使用10mM Tris-HCl, pH 8, 0.1mM EDTA作为稀释剂。

[0359] 2- 制备10X mPCR缓冲液,其含有100mM MOPS, pH 7.5, 75mM MgCl₂, 0.08% Tween 20, 0.08% IGEPAL CA-630, 2.5mM dNTP。

[0360] 3-如下制备mPCR主混合物:

[0361] 表14

[0362]	组分	每反应体积 (μL)
	水	9.62
	10X引物混合物 (各0.75uM)	7.5
	mPCR缓冲液	7.5
	GoTaq (5个单位/uL)	0.38
	总体积	25.0

[0363] 4-解冻DNA并旋转板。

[0364] 5-向96孔ABI Veriti板中加入25μL主混合物。

[0365] 6-将50μL每种样品转移到每个孔中,并通过上下移液几次混合每种样品。

[0366] 7-用铝箔密封件密封板

[0367] 8-放入加热盖热循环仪中,并使用以下配置文件“QX 12cycle”继续循环:

[0368] 表15

[0369]	阶段	温度/时间	循环次数
	预孵育	95°C/ 5 min	1
	扩增 1	95°C/ 30 sec	12
		64°C/ 60 sec	
	冷却	4°C/保持	1

[0370] 9-在孵育完成后,如下进行1至10倍稀释扩增子:

[0371] a. 获得深孔板并向每个孔中转移180μL 10mM Tris-HCl, pH 8, 0.1mM EDTA。

[0372] b. 用96孔印章小心地将小孔刺入扩增板的箔密封件。

[0373] c. 通过使用新鲜吸头和设置为50μL (不产生气溶胶) 的200μL移液器重复移液来混合75μL扩增样品 (不使用振荡器进行混合)

[0374] d. 使用新鲜吸头和设置为20μL (不产生气溶胶) 的20μL移液器将20μL扩增样品加

入到每个预填充孔中。

[0375] e. 通过使用新鲜吸头和设置为100 μ L(不产生气溶胶)的200 μ L移液器重复移液来混合稀释样品(不使用振荡器进行混合)

[0376] f. 用塑料密封件密封稀释的板。

[0377] g. 将稀释的板以1,000rpm离心1min。

[0378] h. 密封未用新铝箔密封件稀释的任何剩余mPCR产物。置于-80°C下。

[0379] 手动QuARTS分析设置:

[0380] 1-解冻鱼DNA稀释液(20ng/ μ L)并用于稀释测定中所需的质粒校准物。使用下表作为稀释指南:

[0381] 表16

	初始质粒浓度, 每 μ L拷贝	最终质粒浓度, 每 μ L拷贝	待加入的 质粒 μ L	待加入的 稀释剂 μ L	总体积, μ L
[0382]	1.00E+05	1.00E+04	5	45	50
	1.00E+04	1.00E+03	5	45	50
	1.00E+03	1.00E+02	5	45	50
	1.00E+02	1.00E+01	5	45	50

[0383] 2-使用下表为标志物A、B和C制备10X三链体QuARTS寡核苷酸混合物:

[0384] 表17

寡核苷酸	序列(5'-3)	浓度(uM)
标志物A正向引物	NA	2
标志物A反向引物	NA	2
标志物A探针-臂1	NA	5
标志物B正向引物	NA	2
标志物B反向引物	NA	2
标志物B探针-臂5	NA	5
标志物C正向引物	NA	2
标志物C反向引物	NA	2
标志物C探针-臂3	NA	5
A1 HEX FRET	/HEX/ TCT/BHQ-1/AGCCGGTTTCCGGCTGAG ACCTCGCGC/3C6/ (SEQ ID NO: 235)	5
A5 FAM FRET	/FAM/ TCT/BHQ-1/AGCCGGTTTCCGGCTGAG ACGTCCGTGG/3C6/ (SEQ ID NO: 236)	5
A3 QUASAR-670 FRET	/Q670/TCT/BHQ-2/AGCCGGTTTCCGGC TGAGACTCCGCGTC/3C6/ (SEQ ID NO: 237)	5

[0386]	dNTP混合物	2500
--------	---------	------

[0387] 3-使用下表制备QuARTS主混合物:

[0388] 表18

[0389]	组分	每反应体积 (μL)
	水	15.5
	10X三链体寡核苷酸混合物	3.0
	20X QuARTS酶混合物	1.5
	总体积	20.0

[0390] *20X酶混合物含有1个单位/μL GoTaq热启动聚合酶 (Promega)、146ng/μL Cleavase 2.0(Hologic)。

[0391] 4-使用96孔ABI板,向每个孔中移取20μL QuARTS主混合物。

[0392] 5-加入10μL适当的校准物或稀释的mPCR样品。

[0393] 6-用ABI透明塑料密封件密封板。

[0394] 7-使用3000rpm将板离心1分钟。

[0395] 8-将板置于ABI热循环仪中,编程以运行以下热协议“Quarts 5+40”,然后启动仪器

[0396] 表19

QuARTS 反应循环:				采集
阶段	温度/时间	升温速率 (每秒°C)	循环次数	
预孵育	95°C/3 min	4.4	1	无
扩增 1	95°C/ 20 sec	4.4	5	无
	63°C/ 30 sec	2.2		无
	70°C/ 30 sec	4.4		无
扩增 2	95°C/ 20 sec	4.4	40	无
	53°C/ 1 min	2.2		是
	70°C/ 30 sec	4.4		无
冷却	40°C/ 30 sec	2.2	1	无

[0398] A. 自动QuARTS设置:

[0399] 1-解冻需要放在Hamilton STARlet甲板上的1管鱼DNA稀释液 (20ng/μL)、2管1.62X寡核苷酸混合物和制备的校准物系列。

[0400] 2-在装载到Hamilton STARlet甲板之前,将所有试剂涡旋并离心。

[0401] 3-将含有样品的深孔板装载到磁体上。

[0402] 4-在平台上放置50μL CORE尖端的完整托盘。

[0403] 5-在平台上放置至少一整排1000μL CORE尖端。

[0404] 6-将空ABI 96孔板装载到STARlet平台上,其中条形码朝向机器前部(在左后角中具有A1孔)。

[0405] 7-按照屏幕上的平台布局和软件说明将试剂装载到指示的载体位置中。

[0406] 8-将2个无盖的条形码空管装载到平台上。

[0407] 9-在Hamilton上运行“QuARTSONLYV4.0_BA_20160127”方法。

[0408] 10-一旦方法完成,则移除96孔QuARTS板并用透明塑料盖密封。

[0409] 11-使用3000rpm将板离心1分钟。

[0410] 12-将板置于ABI热循环仪中,编程以运行以下热协议,然后启动仪器:“Quarts 5+40”。

[0411] 结果

[0412] 这些实验确定

[0413] 1)特异性甲基化DNA标志物 (SERPINB9_3479、GRASP_0932、SLC03A1_6187、ITPRIPL1_1244、AKR1B1_3644、RASGRF2_6325、ZNF655_6075、PAMR1_7364、ST6GALNAC2_1113、CCN JL_9070、KCNB2_9128、IGFBP7_6412和WNT3A_5487) 区分高度侵袭性癌前列腺组织(例如,格里森评分等于或高于7.0(例如,7、8、9、10)与较低侵袭性癌前列腺组织(例如,格里森评分低于7(例如,6),以及

[0414] 2)特异性甲基化DNA标志物 (SERPINB9_3479、FLOT1_1665、HCG4P6_4618、CHST11_2206、MAX.chr12.485、GRASP_0932、GAS6_6425、MAX.chr3.193、MAX.chr2.971_3164、MAX.chr3.727_8028、HES5_0840、TPM4_8037、SLC03A1_6187、ITPRIPL1_1244、AKR1B1_3644、RASGRF2_6325、ZNF655_6075、PAMR1_7364、ST6GALNAC2_1113、CCN JL_9070、KCNB2_9128、IGFBP7_6412和WNT3A_5487) 区分癌前列腺组织(例如,格里森评分等于或高于6.0(例如,6、7、8、9、10)与非癌前列腺组织。

[0415] 表20示出了正常组织的甲基化%、格里森评分为6的前列腺组织的甲基化%,以及格里森评分在7-10之间的前列腺组织的甲基化% (寡核苷酸序列在表21中提供;DMR信息在表22中提供)。

[0416] 表20.

标志物	甲基化% 正常 组织 (28个样品)	甲基化% 格里森 评分6 (24个样品)	甲基化格里森% 评分7-10 (42个样品)
SERPINB9_3479	0.42	12.77	19.45
FLOT1_1665	5.82	42.61	41.91
HCG4P6_4618	3.64	31.43	38.48
CHST11_2206	0.38	16.17	21.95
MAX.chr12.485	3.46	28.85	32.24
GRASP_0932	0.59	28.31	45.31
GAS6_6425	0.34	21.72	20.92
MAX.chr3.193	1.61	27.69	33.89

[0417]

[0418]	MAX.chr2.971_3164	2.08	19.89	24.66
	MAX.chr3.727_8028	0.72	19.76	27.73
	HES5_0840	0.76	13.81	15.08
	TPM4_8037	0.16	3.19	2.90
	SLCO3A1_6187	0.10	3.77	7.18
	ITPRIPL1_1244	0.18	8.77	17.76
	AKR1B1_3644	0.10	5.82	8.32
	JSRP1	0.09	2.16	7.61
	RASGRF2_6325	0.02	0.61	3.07
	ZNF655_6075	0.35	20.51	49.64
	PAMR1_7364	0.03	0.40	1.79
	ST6GALNAC2_1113	0.36	13.76	19.78
	CCNL_9070	0.15	3.33	7.85
	KCNB2_9128	0.19	8.48	17.22
	IGFBP7_6412	0.05	1.06	3.16
	WNT3A_5487	11.52	19.60	37.38

[0419] 注释:

[0420] 1.格里森评分7-10包括

[0421] a.格里森评分7的21个样品

[0422] b.格里森评分8的6个样品

[0423] c.格里森评分9的11个样品

[0424] d.格里森评分10的4个样品

[0425] 2.甲基化%是[(链标志物数量) ÷ (链肌动蛋白数量)] × 100

[0426] 使用相对于关于此数据的ACTB的甲基化%的逻辑回归分析拟合,使用100%截止值,使用两种标志物FL0T1和MAX.Chr3.193,其中AUC=0.99,以98.5%灵敏度从正常预测癌症是允许的。对于6对6+预测,逻辑回归分析相对于所述数据的ACTB拟合9种标志物的甲基化%,以92.8%灵敏度,以91.7%特异性预测格里森6+(AUC=0.96)。所述标志物是:GRASP、GAS6、MAX.chr3.193、MAX.chr2.971、TPM4、ITPRIPL2、AKR1B1、ZNF655、WNT3A。

[0427] 表21.

标志物	寡核苷酸类型	寡核苷酸名称	序列
JSRP1	引物	JSRP1_FP	GGGTCGTAGGAGTGTTCG (SEQ ID NO: 207)
	引物	JSRP1_RP	CCTCTCTAAAAACCGCTAAC (SEQ ID NO: 208)
	探针	JSRP1_Pb_A5	CCACGGACG CTCGTAAACGCC/3C6/ (SEQ ID NO: 209)
ZNF655_60 75	引物	ZNF655_6075_FP	AAGACGTGGAAAAGTTGCG (SEQ ID NO: 210)
	引物	ZNF655_6075_RP	CCGCGCGTCCATAATTG (SEQ ID NO: 211)
	探针	ZNF655_6075_Pb_A1	CCACGGACG CGAAAACTCGTT/3C6/ (SEQ ID NO: 212)
[0428] KCNB2_91 28	引物	KCNB2_9128_FP	GTAGGAGTGGTTGGCGC (SEQ ID NO: 213)
	引物	KCNB2_9128_RP	CCCACACCTCGACGAAAT (SEQ ID NO: 214)
	探针	KCNB2_9128_Pb_A5	CCACGGACG CGCGGAAGTTGA/3C6/ (SEQ ID NO: 215)
ST6GALN AC2_1113	引物	ST6GALNAC2_1113_FP	GGAGGAGAACGCGGATG (SEQ ID NO: 216)
	引物	ST6GALNAC2_1113_RP	GCGATCCCGCGAAAAAACG (SEQ ID NO: 217)
	探针	ST6GALNAC2_1113_Pb_A1	CCACGGACG GAACGCCCCAAA/3C6/ (SEQ ID NO: 218)
MAX.chr3.7 27_8028	引物	MAX.chr3.727_8028_FP	GTGGTTTATTCGTTCGTTCG (SEQ ID NO: 177)
	引物	MAX.chr3.727_8028_RP	AAACTAACGAATATAACTCCGTC GA (SEQ ID NO: 178) 或 CTAACTAAACTAACGAATATAAC TCCGTC (SEQ ID NO: 219)
	探针	MAX.chr3.727_8028_Pb_A1	CGCCGAGG GCGACGTTCGAG/3C6/ (SEQ ID NO: 218)

			NO: 179)
RASGRF2_6325	引物	RASGRF2_6325_FP	GTTAGGGCGGAGAGCGT (SEQ ID NO: 220)
	引物	RASGRF2_6325_RP	CGCGCGATAACAAAAACG (SEQ ID NO: 221)
	探针	RASGRF2_6325_Pb_A5	CGCCGAGG GCGAACTAAAAC/3C6/ (SEQ ID NO: 222)
PAMR1_7364	引物	PAMR1_7364_FP	ACGTTGGAGATTCGCGG (SEQ ID NO: 223)
	引物	PAMR1_7364_RP	CCCCCGCAACTCCTT (SEQ ID NO: 224)
	探针	PAMR1_7364_Pb_A5	CGCCGAGG GACGGCGGTTGT/3C6/ (SEQ ID NO: 225)
[0429]	引物	IGFBP7_6412_FP	GGGTCGTAGGTGTTCGAA (SEQ ID NO: 226)
	引物	IGFBP7_6412_RP	GCGCCCTACTCCTCGAC (SEQ ID NO: 227)
	探针	IGFBP7_6412_Pb_A5	CGCCGAGG CGCCGCTAAACT/3C6/ (SEQ ID NO: 228)
CCN JL_9070	引物	CCN JL_9070_F_P	GGTATCGTAGTTTCGCGGA (SEQ ID NO: 229)
	引物	CCN JL_9070_RP	CTCCTACGCCGCTCAA (SEQ ID NO: 230)
	探针	CCN JL_9070_Pb_A5	CGCCGAGG ATTAGAGGCGAT/3C6/ (SEQ ID NO: 231)
WNT3A_548787	引物	WNT3A_5487_FP	GTGTAAATGCGCGGGC (SEQ ID NO: 232)
	引物	WNT3A_5487_RP	CGCTTTAATTCAACACCGCG (SEQ ID NO: 233)
	探针	WNT3A_5487_Pb_A5	CGCCGAGG CGGTTATACGT/3C6/ (SEQ ID NO: 234)
FLOT1_1665	引物	FLOT1_1665_F_P	GTAGTGTGAGTTAAGTTGTT TCG (SEQ ID NO: 150)
	引物	FLOT1_1665_R_P	AACACCAAAACGTACCTAATACG (SEQ ID NO: 151)
	探针	FLOT1_1665_Pb_A1	CGCCGAGG GCGAAAACGTAT/3C6/ (SEQ ID NO: 152)

HCG4P6_4 618	引物	HCG4P6_4618_FP	CGGAGTATACGGAATATTAGGTT CG (SEQ ID NO: 153)
	引物	HCG4P6_4618_RP	AATAATAACGACGCGACATAAAC A (SEQ ID NO: 154)
	探针	HCG4P6_4618_Pb_A5	CCACGGACG GCGTATAGATTG/3C6/ (SEQ ID NO: 155)
SERPINB9_3479	引物	SERPINB9_3479_FP	TTTATTAGTCGTGCGCGG (SEQ ID NO: 147)
	引物	SERPINB9_3479_RP	ACGCGACACGAAAACGAC (SEQ ID NO: 148)
	探针	SERPINB9_3479_Pb_A5	CCACGGACG CGACAACGCTAA/3C6/ (SEQ ID NO: 149)
CHST11_22 06	引物	CHST11_2206_FP	GCGTCGAGGGCGTTT (SEQ ID NO: 156)
	引物	CHST11_2206_RP	AATCTAAACTAAAAAAACGA AACTCGT (SEQ ID NO: 157)
	探针	CHST11_2206_Pb_A1	CGCCGAGG TCGCTTCCTAAA/3C6/ (SEQ ID NO: 158)
[0430] MAX.chr12. 485	引物	MAX.chr12.485_FP	GAGGAATAGGCGTGGTTCG (SEQ ID NO: 159)
	引物	MAX.chr12.485_RP	AAAACCGAAAAAAACTAATTCTG CA (SEQ ID NO: 160)
	探针	MAX.chr12.485_Pb_A5	CCACGGACG GCGTGACGGTT/3C6/ (SEQ ID NO: 161)
GRASP_09 32	引物	GRASP_0932_FP	CGCGCGGTATAGTCGG (SEQ ID NO: 162)
	引物	GRASP_0932_RP	ACCCCCAACGAACGAAC (SEQ ID NO: 163)
	探针	GRASP_0932_Pb_A1	CGCCGAGG CTATACGCGACG/3C6/ (SEQ ID NO: 164)
GAS6_6425	引物	GAS6_6425_FP	GATTGAGGTTAGGTTCTCG (SEQ ID NO: 165)
	引物	GAS6_6425_RP	GAAAAAACGCTACTACCGC (SEQ ID NO: 166)
	探针	GAS6_6425_Pb_A5	CCACGGACG GCGCGGAGTTGG/3C6/ (SEQ ID NO: 167)
MAX.chr3.1	引物	MAX.chr3.193_	TTAAAGGTTAGTTAAAGATGGA

93		FP	ATCGT (SEQ ID NO: 174)	
	引物	MAX.chr3.193_RP	CTCGCGACGAAAAAAACCC (SEQ ID NO: 175)	
	探针	MAX.chr3.193_Pb_A1	CGCCGAGG CGAACTCCCAAC/3C6/(SEQ ID NO: 176)	
MAX.chr2.9 71	引物	MAX.chr2.971_FP	GCGAGGTTGCGGTTTG (SEQ ID NO: 198)	
	引物	MAX.chr2.971_RP	CGAAAAAATAATTGATTAAACCA AAACGC (SEQ ID NO: 199)	
	探针	MAX.chr2.971_Pb_A5	CCACGGACG CCGATATTCccc/3C6/ (SEQ ID NO: 200)	
HES5_0840	引物	HES5_0840_FP	GCGAGAGGAGTAGGTTCG (SEQ ID NO: 180)	
	引物	HES5_0840_RP	AACCTACGAACGCGCGA (SEQ ID NO: 181)	
	探针	HES5_0840_Pb_A5	CCACGGACG ACCGACGACCA/3C6/ (SEQ ID NO: 182)	
[0431]	TPM4_8037	引物	TPM4_8037_FP	GGAGAAAGGCGGGCG GA (SEQ ID NO: 189)
		引物	TPM4_8037_RP	CGTTCCCAAAAACGCGA (SEQ ID NO: 190)
		探针	TPM4_8037_Pb_A1	CGCCGAGG ACGACGACGTAT/3C6/ (SEQ ID NO: 191)
SLCO3A1_6187	引物	SLCO3A1_6187_FP	CGATTGTAGCGTTGGATATCG (SEQ ID NO: 171)	
	引物	SLCO3A1_6187_RP	TCGAAAACCCCACGAACC (SEQ ID NO: 172)	
	探针	SLCO3A1_6187_Pb_A5	CCACGGACG CGTTAATACCCC/3C6/ SEQ ID NO: 173)	
ITPRIPL1_1244	引物	ITPRIPL1_1244_FP	GTAGTGGCGGTTAGGTCG (SEQ ID NO: 195)	
	引物	ITPRIPL1_1244_RP	TCACCGTCAATATTAATAAAAAACACGA (SEQ ID NO: 196)	
	探针	ITPRIPL1_1244_Pb_A1	CGCCGAGG GC GTTGT TTTT/3C6/ (SEQ ID NO: 197)	
AKR1B1_3 644	引物	AKR1B1_3644_FP	GTTCGTTTATATTTCGCGCG (SEQ ID NO: 201)	

[0432]	引物	AKR1B1_3644 _RP	CCGAACGTCCGCGAAC (SEQ ID NO: 202)
	探针	AKR1B1_3644 _Pb_A1	CGCCGAGG CGAACTACTCAA/3C6/(SEQ ID NO: 203)

[0433] 表22.

DMR 号	基因注释	染色体号	染色体上的区域 (开始碱基-结束碱基)
121	WNT3A_5487	1	228225487-228225590
122	SERPINB9_3479	6	2903389-2903479
123	FLOT1_1665	6	30711586-30711665
124	HCG4P6_4618	6	29894618-29894693
125	GRASP_0932	12	52400932-52401020
126	GAS6_6425	13	114566425-114566518
127	SLCO3A1_6187	15	92396091-92396187
128	MAX.chr3.727_8028	3	72788028-72788112
129	HES5_0840	1	2460840-2460903
130	TPM4_8037	19	16188037-16188154
131	ITPRIPL1_1244	2	96991244-96991312
132	MAX.chr2.971_3164	2	97193164-97193252
133	AKR1B1_3644	7	134143644-134143721
134	RASGRF2_6325	5.	80256325-80256390
135	PAMR1_7364	11	35547364-35547423
136	ZNF655_6075	7	99156075-99156154
137	CCN JL_9070	5.	159739070-159739148
138	ST6GALNAC2_1113	17	74581113-74581238
139	IGFBP7_6412	4	57976412-57976506
140	KCNB2_9128	8	73449128-73449208
17	MAX.chr3.193	3	193776187-193776257
70	CHST11_2206	12	104852206-104852307
83	MAX.chr12.485	12	48592041-48592162

[0435] 出于所有目的,以上说明书提及的所有出版物和专利均以引用的方式整体并入本文。所描述技术的组合物、方法以及用途的各种改进和变化对于本领域技术人员将是显而易见的,而不背离所描述技术的范围和精神。虽然已结合特定示例性实施方案描述了所述技术,但应理解的是,要求保护的本发明不应不适当当地局限于此类特定实施方案。事实上,对药理学、生物化学、医学科学或相关领域技术人员来说显而易见的用于进行本发明的所描述模式的各种修改旨在于以下权利要求的范围内。