



(12) 发明专利

(10) 授权公告号 CN 108138235 B

(45) 授权公告日 2022. 04. 15

(21) 申请号 201680050336.6	(72) 发明人 哈提姆·T·阿拉维 大卫·A·阿尔奎斯特 威廉·R·泰勒 约翰·B·基谢尔 特雷西·C·亚伯 道格拉斯·W·马奥尼
(22) 申请日 2016.08.31	
(65) 同一申请的已公布的文献号 申请公布号 CN 108138235 A	
(43) 申请公布日 2018.06.08	(74) 专利代理机构 北京康信知识产权代理有限 责任公司 11240 代理人 张英 沈敬亭
(30) 优先权数据 62/212,221 2015.08.31 US	(51) Int.Cl. C12Q 1/6886 (2018.01)
(85) PCT国际申请进入国家阶段日 2018.02.28	
(86) PCT国际申请的申请数据 PCT/US2016/049653 2016.08.31	(56) 对比文件 US 2003224040 A1,2003.12.04 US 2014137274 A1,2014.05.15 US 2014274748 A1,2014.09.18
(87) PCT国际申请的公布数据 W02017/040627 EN 2017.03.09	审查员 谢庆宁
(73) 专利权人 梅约医药教育及研究基金会 地址 美国明尼苏达州 专利权人 精密科学发展公司	权利要求书1页 说明书46页 序列表28页 附图22页

(54) 发明名称
检测胃肿瘤

(57) 摘要
本发明提供了一种检测肿瘤的技术,特别是
用于检测类似胃癌等癌前和恶性肿瘤的方法、组
合物和相关用途,但不仅限于此。

1. 检测由ELM01、C13orf18和ZNF569的差异甲基化区域DMR组成的三种标记的組合的甲基化状态的试剂在用于制备用于表征从患有胃癌的人类对象获得的样本的系统的用途,所述系统包括:分析元件,被配置来确定所述样本的ELM01、C13orf18和ZNF569的差异甲基化区域DMR的甲基化状态;软件元件,被配置来将所述样本的所述甲基化状态与记录在数据库中的对照样本或参考样本的甲基化状态进行比较;和警报元件,被配置来基于所述样本的甲基化状态的組合确定单个值,并告知用户有关胃癌的甲基化状态;其中所述ELM01的差异甲基化区域DMR为染色体7胰岛范围37487867-37488498;所述C13orf18的DMR为染色体13胰岛范围46960767-46961669;和所述ZNF569的DMR为染色体19胰岛范围37957760-37958046。

2. 根据权利要求1所述的用途,其中,所述样本包括:含有DMR的核酸。
3. 根据权利要求1所述的用途,其中所述系统进一步包括用于分离核酸的元件。
4. 根据权利要求1所述的用途,其中所述系统进一步包括用于收集样本的元件。
5. 根据权利要求1所述的用途,其中,所述样本是粪便样本、组织样本、血浆样本或尿液样本。
6. 根据权利要求1所述的用途,其中,所述样本是胃组织样本。
7. 根据权利要求1所述的用途,其中,所述数据库包括:含有DMR的核酸序列。
8. 根据权利要求1所述的用途,其中,所述数据库包括:无胃癌的对象的核酸序列。

检测胃肿瘤

[0001] 相关申请的交互参照

[0002] 本申请是2016年8月31日提交的美国临时申请No. 62/212,221申请的优先权,其全部内容被纳入此处作为参考。

技术领域

[0003] 本发明提供了一种检测肿瘤的技术,特别是用于检测胃癌等癌前和恶性肿瘤的方法、组合物和相关用途,但并不局限于此。

[0004] 发明背景

[0005] 胃癌是世界上第三大最常见的癌症死亡原因(例如参照世界卫生组织的癌症信息发布文件No 297.WHO),而且在西方国家仍然很难治愈,主要是因为大多数患者都患具晚期疾病。在美国,胃恶性肿瘤是目前最常见的第十五种癌症(例如参照国家癌症研究所的监测、流行病学和最终结果SEER Stat信息发布文件:胃癌)。然而,在大多数发达国家,胃癌发病率在过去半个世纪中急剧下降。

[0006] 胃癌减少的原因是由于广泛使用制冷,其产生了若干有益的影响:增加了新鲜水果和蔬菜的消费量;减少了食盐的摄入量,盐被用作食品防腐剂;以及减少了因未冷藏肉类产品腐烂而产生的致癌化合物对食品的污染。食盐和盐类食物可能会破坏胃粘膜,导致炎症,以及相关的DNA合成和细胞增殖的增加。其他可能导致胃癌发病率下降的因素包括:由于卫生条件和抗生素使用的改善,慢性幽门螺杆菌感染率降低,以及一些国家加强了筛查(例如参照美国癌症协会的“全球癌症事实&数据显示”3rd ed.)。

[0007] 然而,在西方国家,胃癌仍然很难治愈,这主要是因为大多数患者都患有晚期疾病。即使是处于最有利状态的患者,也经常死于复发性疾病。

[0008] 因此,需要改进胃癌的早期检测方法和技术。

发明内容

[0009] 甲基化DNA作为一种潜在的生物标记,在大多数肿瘤类型的组织中得到了广泛的研究。在许多情况下,DNA甲基转移酶在胞嘧啶-磷酸盐-鸟嘌呤(CpG)胰岛位点向DNA中添加甲基,作为基因表达的外遗传控制。在具有生物吸引的机制中,肿瘤抑制基因启动子区获得的甲基化事件可被认为沉默基因表达,从而促使肿瘤发生。DNA甲基化可能是一种比RNA或蛋白质表达更具化学和生物稳定性的诊断工具(Laird (2010) “全基因组DNA甲基化分析的原则和挑战”Nat Rev Genet 11:191-203)。此外,在散发性结肠癌等其他癌症中,甲基化标记比单个DNA突变具有更好的特异性和更广泛的信息性和敏感性(Zou等人(2007) “结直肠癌肿瘤中的高度甲基化基因:用于筛查的暗示”癌症流行病学生物标记Prev 16:2686-96)。

[0010] 对CpG胰岛的分析在应用于动物模型和人类细胞系时已经取得了重要的结果。例如,Zhang和他的同事发现,来自同一CpG胰岛的不同部分的扩增产物可能存在不同程度的甲基化(Zhang等人(2009) “单碱基对和单等位基因解析的21号染色体基因启动子的DNA甲基化分析”PLoS Genet 5:e1000438)。此外,甲基化水平在高度甲基化序列和非甲基化序列

之间呈双模式分布,进一步支持DNA甲基转移酶活性的二进制转换图案(Zhang等人(2009)“单碱基对和单等位基因解析的21号染色体基因启动子的DNA甲基化分析”PLoS Genet 5:e1000438)。对小鼠体内和体外细胞系的分析表明,只有约0.3%的高CpG密度启动子(HCP,定义为在300个碱基对区域有>7%CpG序列)被甲基化,其中低CpG密度区域(定义为在300个碱基对区域内具有<5%CpG序列)往往以动态组织特异性图案被频繁地甲基化(Meissner等人(2008)“多能和分化细胞的基因组尺度DNA甲基化图谱”Nature 454:766-70)。HCPs包括普遍存在的管家基因的启动子和高度调控的发育基因。在>50%甲基化的HCP位点中,有几个已建立的标记,如Wnt 2,NDRG2,SFRP2和BMP3(Meissner等人(2008)“多能和分化细胞的基因组尺度DNA甲基化图谱”Nature 454:766-70)。

[0011] 总的来说,胃肠道癌症占癌症死亡率比任何其他器官系统都高。然而,目前只有结直肠癌在国内进行筛查。美国每年死于上消化道癌的死亡率超过90000,而结直肠癌的死亡率约为50000。胃腺癌是全世界第二常见的癌症死亡原因,每年有近一百万人死亡。在太平洋、亚洲、俄罗斯和拉丁美洲的一些地区最为普遍。由于大多数患者具晚期疾病,因此预后较差(5年生存率<5-15%)。

[0012] 胃癌的基因组机制是复杂和不完全了解的。幽门螺杆菌感染引起的慢性胃炎与环境暴露有关。染色体和小随体不稳定是与胃癌和表观遗传改变有关。在24%~47%的病例中也描述了CpG甲基化表型。通常被观察到的突变包括基因p53,APC, β -连环蛋白,和k-ras,但这些相对罕见,且具异质性。

[0013] 胃癌会使细胞和DNA散布至消化道,最终在大便中排出。以前曾用高敏感性化验来检测胃癌患者粪便中的突变DNA,这些患者的切除肿瘤已知含有相同的序列。然而,突变标记的一些限制与检测的潜在异质性和笨拙过程有关;通常,必须对多个基因的每个突变位点进行单独化验,才能获得高敏感性。

[0014] DNA甲基转移酶活性的胞嘧啶-磷酸盐-鸟嘌呤(CpG)胰岛位点的表观基因甲基化已被认为是大多数肿瘤组织中潜在的一类生物标记。在具有生物吸引的机制中,肿瘤抑制基因启动子区获得的甲基化事件可被认为沉默基因表达,从而促使肿瘤发生。DNA甲基化可能是一种比RNA或蛋白质表达更具化学和生物稳定性的诊断工具。此外,在散发性结肠癌等其他癌症中,异常甲基化标记比单个DNA突变具有更广泛的信息性和敏感性并具有更好的特异性。

[0015] 一些方法可用于寻找新的甲基化标记。虽然微阵列的CpG甲基化是一种合理的、高通量的方法,但这种策略偏向于已知的感兴趣区域,主要建立的肿瘤抑制基因启动子。在过去的十年里,DNA甲基化的全基因组分析的替代方法已经开发出来。有三种基本方法。第一种方法是利用DNA限制酶切来识别特定的甲基化位点,然后是几种可能的分析技术,这些技术提供的甲基化数据仅限于酶识别位点或以量化步骤扩增DNA的引物(如甲基化特异性PCR;MSP)。第二种方法是利用定向于甲基胞嘧啶或其他甲基化特异性结合域的抗体来丰富基因组DNA的甲基化部分,然后进行微阵列分析或测序,将片段映射到参考基因组。这种方法不提供片段中所有甲基化位点的单核苷酸解析。第三种方法从DNA的重亚硫酸盐处理开始,将所有未甲基化的胞嘧啶转化为尿嘧啶,然后在耦合到适配体后限制酶切以及对所有片段进行完全测序。选择限制酶可以丰富CpG稠密区的片段,减少分析过程中可能映射到多个基因体位的冗余序列数。后一种方法称为减少代表性重亚硫酸盐序列(RRBS),但据了解

尚未用于胃癌的研究。

[0016] RRBS在80~90%的CpG胰岛和大多数中高解读覆盖率的肿瘤抑制启动子的单核苷酸解析下产生CpG甲基化状态数据。在癌症病例对照研究中,对这些数据进行分析,可以识别差异甲基化区域(DMRs)。在先前对胰腺癌样本的RBS分析中,发现了数百例DMRs,其中一些从未与癌变相关,且其中一些未加标注。对于独立组织样本集的进一步有效性研究证实了标记CpGs在性能上具100%敏感性和特异性。

[0017] 高鉴别标记的临床应用可能会产生很大影响。例如,在诸如粪便或血液等远隔介质中的标记化验,可以为胃肿瘤的检测提供准确的筛查或诊断工具。

[0018] 在开发本技术实施例的过程中所进行的实验列出并描述从胃癌样本以类似方式所生成的123新DNA甲基化标记。连同癌、正常胃组织、正常结肠组织和正常白细胞DNA一起被测序。这些标记在较大的样本集上被验证,以确定及优化最强壮的候选对象。另外的实验确定了用于检测胃癌的10个最佳标记(ARHGEF4,ELM01,ABCB1,CLEC11A,ST8SIA1,SFMBT2,CD1D,CYP26C1,ZNF569和C13ORF18;见示例2和表4、5和6)。另外的实验确定了用于检测胃癌的30个最佳标记(ELM01,ARHGEF4,EMX1,SP9,CLEC11A,ST8SIA1,BMP3,KCNA3,DMRTA2,KCNK12,CD1D,PRKCB,CYP26C1,ZNF568,ABCB1,ELOVL2,PKIA,SFMBT2(893),PCBP3,MATK,GRN2D,NDRG4,DLX4,PPP2R5C,FGF14,ZNF132,CHST2(7890),FLI1,c13orf18,或ZNF569;见示例1和表1、2和3)。另外的实验确定了用于在血浆中检测胃部癌(例如胃癌)的12个最佳标记(ELM01,ZNF569,C13orf18,CD1D,ARHGEF4,SFMBT2,PPP25RC,CYP26C1,PKIA,CLEC11A,LRRCA和ST8SIA1;见示例3和表7和表8)。

[0019] 因此,本发明提供一种胃癌筛查标记的技术,当从对象的样本(例如粪便样本、胃组织样本、血浆样本)中被检测时提供高信号噪声比和低基质水平。标记是在通过将具胃癌的对象的胃组织或血浆样本的DNA标记的甲基化状态与对照对象的相同DNA标记的甲基化状态进行比较的病例对照研究中被确定(例如,正常胃组织、正常结肠上皮细胞,正常血浆和正常白细胞衍生的DNA;参照示例1、2和3、表1-8)(例如,血浆组织;参照示例3)。

[0020] 如本文所述,本技术提供了一些甲基化DNA标记及其子集(例如,2、3、4、5、6、7、8、9、10、11、12或更多标记),总体上对胃癌具有很高的鉴别能力。实验将筛查滤波器应用于候选标记,以识别标记提供高信噪比和低基质水平,为胃癌筛查或诊断提供高特异性。

[0021] 在一些实施例中,本技术涉及评估生物样本(例如胃组织、血浆样本)中所识别的一个或多个标记的存在和甲基化状态。这些标记包括在此所述的一个或多个差异甲基化区域(DMR),例如,如表1、2、4和/或7所提供的。甲基化状态在本技术的实施例中被评估。因此,在此提供的技术并不局限于基因甲基化状态被测量的方法。例如,在一些实施例中,甲基化状态是通过基因组扫描方法被测量。例如,一种方法涉及限制性标志基因组扫描(Kawi等人(1994)Mol.Cell.Biol.14:7421-7427),且另一个示例涉及甲基化敏感性随机引物PCR(Gonzalogo et al.(1997)Cancer Res.57:594-599)。在一些实施例中,特定CpG位点处甲基化图案中的变化通过基因组DNA的甲基化敏感性限制酶切被监测,然后是兴趣区域的Southern分析(酶切-Southern法)。在一些实施例中,分析甲基化图案中的变化涉及基于PCR的过程,其涉及在PCR扩增之前基因组DNA的甲基化敏感性限制酶切(Singer-Sam等人(1990)Nucl.Acids Res.18:687)。此外,其他技术也被说明,利用DNA的重亚硫酸盐处理作为甲基化分析的起点。其包括甲基化特异性PCR(MSP)(Herman等人(1992))

Proc.Natl.Acad.Sci.USA 93:9821-9826) 和从重亚硫酸盐转化的DNA中扩增的PCR产物的限制酶切(Sadri和Hornsby (1996) Nucl.Acids Res.24:5058-5059;and Xiong and Laird (1997) Nucl.Acids Res.25:2532-2534)。PCR技术被发展来用于检测基因突变(Kuppuswamy等人 (1991) Proc.Natl.Acad.Sci.USA 88:1143-1147) 和等位基因特异性表达的量化(Szabo和Mann (1995) Genes Dev.9:3097-3108;以及Singer-Sam等人 (1992) PCR Methods Appl.1:160-163)。这些技术使用内部引物,其对PCR生成的模板进行退火,并立即终止将化验的5'的单核苷酸。在一些实施例中使用的采用“定量的Ms-SNuPE化验”的方法在美国专利No.7,037,650中被说明。

[0022] 在评估甲基化状态时,甲基化状态通常表示为在特定定位点相对于含有该特定定位点的样本中的DNA总数被甲基化的DNA单个链的部分或百分比(例如,在单核苷酸、在特定区域或位点、在较长感兴趣序列,例如,长达a~100-bp, 200-bp, 500-bp, 1000-bpDNA序列或更长)。一般,非甲基化核酸的数量是通过校准器经PCR被确定的。然后,对已知量的DNA进行重亚硫酸盐处理,并使用实时PCR或其他指数式扩增来确定生成的甲基化特异性序列,例如QuARTS化验(例如由美国专利No.8,361,720和美国专利申请公开Nos.2012/0122088和2012/0122106提供,其被纳入此处作为参考)。

[0023] 例如,在一些实施例中,方法包括使用外部标准来生成用于非甲基化目标的标准曲线。标准曲线至少由两点构成,并将非甲基化DNA的实时Ct值与已知的定量标准联系起来。然后,用于甲基化目标的第二标准曲线由至少两个点和外部标准构成。该第二标准曲线将甲基化DNA的Ct与已知的定量标准联系起来。然后,测试样本Ct值被确定来用于甲基化和非甲基化的群体,并根据前两步产生的标准曲线来计算DNA的基因组等价物。感兴趣位点处的甲基化百分比通过将群体中DNA的总数量与甲基化DNA的数量相比较被计算,例如(甲基化DNA数量)/(甲基化DNA数量+非甲基化DNA数量)×100。

[0024] 在此,还提供了用于实践这些方法的组合物和试剂盒。例如,在一些实施例中,针对一个或多个标记的试剂(例如,引物、探针)被单独或成套提供(例如用来扩增多个标记的一组引物)。还可以提供用于执行检测化验的其他试剂(例如酶、缓冲剂、用于执行QuARTS、PCR、测序、重亚硫酸盐或其他化验的阳性和阴性对照)。在一些实施例中,提供一种试剂盒,其包含执行在此所述的方法所需的、足够的或有用的一个或多个试剂。此外还提供含有多个试剂的主混合试剂组件,所述多个试剂可以彼此添加和/或添加到测试样本中以完成反应混合物。

[0025] 在一些实施例中,在此所述的技术与可编程器件相关联,该可编程器件被设计用来执行在此所述的方法所提供的序列运算或逻辑操作。例如,本技术的一些实施例与计算机软件和/或计算机硬件相关联(例如由此被执行)。在一个方面,该技术涉及一种计算机,包括存储器形式、执行运算和逻辑操作的单元,以及用于执行一系列指令(例如,此处提供的方法)的处理单元(例如微处理器)来用于读取、操作和存储数据。在一些实施例中,微处理器是用于确定甲基化状态的系统的一部分(例如,一个或多个DMR,例如表1、2、4和7中提供的DMR1-274);比较甲基化状态(例如,一个或多个DMR的甲基化状态,例如,表1、2、4和7中提供的DMR1-274);生成标准曲线;确定Ct值;计算甲基化的部分、频率或百分比(例如一个或多个DMR,例如一个或多个DMR,表1、2、4和7中提供的DMR 1-274);鉴别CpG胰岛;确定化验或标记的特异性和/或敏感性;计算ROC曲线和相关的AUC;序列分析;在此所述的或本领域

已知的一切。

[0026] 在一些实施例中,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位置。

[0027] 在一些实施例中,软件或硬件组件接收多个化验的结果,并确定单一值的结果,以根据多个化验的结果(例如确定多个DMR的甲基化状态,如表1、2、4和7中提供的)向用户报告基于癌症风险的结果。相关实施例基于多个化验结果的数学组合(例如,加权组合、线性组合)计算风险因子,例如,确定多个标记的甲基化状态(例如,如表1、2、4和7中提供的多个DMR)。在一些实施例中,DMR的甲基化状态定义了维度,并可以在多维空间中具有值,且由多个DMR的甲基化状态定义的坐标作为与癌症风险相关的结果向用户报告。

[0028] 一些实施例包括存储介质和存储器部件。存储器部件(例如,易失性和/或非易失性存储器)可用于存储指令(例如,此处提供的处理的实施例)和/或数据(例如工作部件,诸如甲基化测量、序列以及与其相关的统计描述等)。一些实施例还涉及包括一个或多个CPU、图形卡和用户界面的系统(例如,包括诸如显示器的输出设备和诸如键盘的输入设备)。

[0029] 与本技术相关的可编程器件包括现有技术和正在开发或尚未开发的技术(例如量子计算机、化学计算机、DNA计算机、光学计算机、基于自旋电子的计算机等)。

[0030] 在一些实施例中,该技术包括用于发送数据的有线(例如金属电缆、光纤)或无线传输介质。例如,一些实施例涉及通过网络(例如,局域网(LAN)、广域网(WAN)、ad-hoc网络、因特网等)进行数据传输。可编程器件存在于该网络上,并且在一些实施例中,可编程器件与客户机/服务器相联。

[0031] 在一些实施例中,数据被存储在计算机可读存储介质上,如硬盘、闪存、光学介质、软盘等。

[0032] 在一些实施例中,在此提供的技术与多个可编程设备相关联,这些可编程设备协同工作以执行在此所述的方法。例如,在一些实施例中,多个计算机(例如,由网络连接)可以并行工作来收集和处理数据,例如,通过诸如以太网、光纤或无线网络技术的常规网络接口与网络(个人、公共或因特网)连接的完整计算机(带有机载CPU、存储、电源、网络接口等),以集群计算或网格计算或其他分布式计算机体系结构执行。

[0033] 例如,一些实施例提供包括计算机可读媒体的计算机。该实施例包括连接到处理器的随机存取存储器(RAM)。处理器执行存储在内存中的计算机可执行程序指令。这些处理器可以包括微处理器、ASIC、状态机或其他处理器,也可以是许多计算机处理器中的任何一个,例如加利福尼亚州圣克拉拉英特尔公司和伊利诺伊州Schaumburg的摩托罗拉公司的处理器。这些处理器包括或可以与媒体通信,例如计算机可读媒体,当由处理器执行时其可存储指令,使处理器执行在此所述的步骤。

[0034] 计算机可读媒体的实施例包括能够提供具计算机可读指令的处理器电子、光学、磁性或其他存储或传输设备,但并不局限于此。其他合适的媒体的示例包括软盘、CD-ROM、DVD、磁盘、存储器芯片、ROM、RAM、ASIC、配置处理器、所有光学媒体、所有磁带或其他磁媒体,或计算机处理器可从中读取指令的任何其他媒体,但并不局限于此。此外,各种其他形式的计算机可读媒体可以向计算机发送或传送指令,包括路由器、个人或公用网络或其他有线和无线传输设备或信道。这些指令可包括来自任何合适的计算机编程语言的代码,例如C、C++、C#、可视化Basic语言、Java、Python、Perl和JavaScript。

[0035] 在一些实施例中计算机被连接至网络。计算机还可以包括一些外部或内部设备，如鼠标、光盘、DVD、键盘、显示器或其他输入或输出设备。计算机的示例为个人电脑、数字助理、个人数字助理、手机、移动电话、智能电话、寻呼机、数字平板电脑、笔记本电脑、互联网设备和其他基于处理器的设备。一般，与本发明提供的技术的各个方面相关的计算机可以是任何类型的基于处理器的平台，这些平台可以在任何操作系统上运行，如Microsoft Windows、Linux、UNIX、MacOSX等，能够支持包含在此提供的技术的一个或多个程序。一些实施例包括执行其他应用程序(例如，应用程序)的个人计算机。这些应用程序可以包含在存储器中，并且可以包括：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、即时信使应用程序、演示应用程序、因特网浏览器应用程序、日历/管理者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。

[0036] 在此所述的与本技术相关的所有这些组件、计算机和系统都可以是逻辑或虚拟机。

[0037] 因此，提供一种技术，涉及在从对象获得的样本中筛查肿瘤的方法，该方法包括化验从对象(例如胃组织)获得的样本中的标记的甲基化状态(例如血浆样本)。当标记的甲基化状态不同于在没有肿瘤的对象中化验的标记的甲基化状态时，确定该对象为肿瘤。其中，所述标记包括从表1、2、4和/或7的DMR 1-274构成的组中选出的差异甲基化区域(DMR)中的碱基。本技术与识别和鉴别胃癌有关，例如，在一些实施例中，肿瘤是胃部癌(例如胃癌)。一些实施例提供包括化验多个标记的方法，例如，包括化验2至11至100或200或274标记。

[0038] 本技术并不局限于甲基化状态的评估。在一些实施例中，评估样本中标记的甲基化状态包括确定一个碱基的甲基化状态。在一些实施例中，化验样本中标记的甲基化状态包括确定多个碱基处的甲基化程度。此外，在一些实施例中，标记的甲基化状态包括相对于标记的正常甲基化状态，标记的甲基化增加。在一些实施例中，标记的甲基化状态包括相对于标记的正常甲基化状态，标记的甲基化减少。在一些实施例中，标记的甲基化状态包括相对于标记的正常甲基化状态，标记的甲基化的不同图案。

[0039] 此外，在一些实施例中，标记是100个或更少的碱基的区域，标记是500个或更少的碱基的区域，标记是1000个或更少的基的区域，标记是5000个或更少的基的区域，或者。在一些实施例中，标记是一个基。在一些实施例中，标记位于高cpg密度启动子中。

[0040] 本技术不受样本类型的限制。例如，在一些实施例中，样本是粪便样本、组织样本(例如胃组织样本)、血样本(例如血浆、血清、全血)、排泄物或尿液样本。

[0041] 此外，本技术在确定甲基化状态的方法上也不受限制。在一些实施例中，化验包括使用甲基化特异性聚合酶链式反应、核酸测序、质谱、甲基化特异性核酸酶、基于质量的分离或目标捕获。在一些实施例中，化验包括甲基化特异性寡核苷酸的使用。在一些实施例中，本技术使用大规模并行测序(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、微珠乳液测序、纳米粒测序等。

[0042] 该技术提供了用于化验DMR的试剂，例如，在一些实施例中，提供了一组包括由SEQ ID NO:1-109提供的序列的寡核苷酸。在一些实施例中，提供了一种寡核苷酸，该寡核苷酸包括与DMR中具有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。

[0043] 本技术提供各种标记组，例如，在一些实施例中，该标记包括具有标注的染色体区

域,标注为ARHGEF4,ELM01,ABCB1,CLEC11A,ST8SIA1,SFMBT2,CD1D,CYP26C1,ZNF569,或C13ORF18,且其包括标记(见表4)。此外,实施例提供一种分析表4中的DMR的方法,即DMR No.49,43,196,66,1,237,249,250,251或252。一些实施例提供确定标记的甲基化状态的方法,其中一个染色体区域具有一个标注,即ELM01,ARHGEF4,EMX1,SP9,CLEC11A,ST8SIA1,BMP3,KCNA3,DMRTA2,KCNK12,CD1D,PRKCB,CYP26C1,ZNF568,ABCB1,ELOVL2,PKIA,SFMBT2(893),PCBP3,MATK,GRN2D,NDRG4,DLX4,PPP2R5C,FGF14,ZNF132,CHST2(7890),FLI1,c13orf18,或ZNF569,且其包括标记(见表2)。此外,实施例提供一种分析表2中的DMR的方法,即DMR No.253,251,254,255,256,249,257,258,259,260,261,262,250,263,1,264,265,196,266,118,267,268,269,270,271,272,46,273,252,或237。在一些实施例中,所获得的样本是血浆样本,所述标记包括具有标注的染色体区域,标注为ELM01,ZNF569,C13orf18,CD1D,ARHGEF4,SFMBT2,PPP25RC,CYP26C1,PKIA,CLEC11A,LRRC4,及ST8SIA1,且其包括标记。此外,在样本是血浆样本的实施例中,提供从表1、2、4和7中分析DMR的方法,即DMR No.253,251,254,255,256,249,257,258,259,260,261,262,250,263,1,264,265,196,266,118,267,268,269,270,271,272,46,273,252,或237。在一些实施例中,所述方法包括确定两个标记的甲基化状态,例如表1、2、4或7行中提供的一对标记。

[0044] 提供试剂盒实施例,例如试剂盒包含:重亚硫酸盐试剂;以及控制核酸,其包含选自DMR1-274(来自表1、2、4或7)的DMR的序列,并且具有与无癌症(例如胃癌)的对象相关的甲基化状态。在一些实施例中,试剂盒包括本文所述的重亚硫酸盐试剂和寡核苷酸。在一些实施例中,试剂盒包括重亚硫酸盐试剂;和控制核酸,其包含选自DMR1-274(来自表1、2、4或7)的DMR的序列,并且具有与患有癌症(例如胃癌)的受试者相关的甲基化状态。一些试剂盒实施方式包括:用于从对象获得样本的样本收集器(例如粪便样本;胃组织样本;血浆样本);用于从样本中分离核酸的试剂;重亚硫酸盐试剂;和在此所述的寡核苷酸。

[0045] 本技术涉及组合物(例如,反应混合物)的实施例。在一些实施例中,提供一种组合物,该组合物包括:含有DMR的核酸和重亚硫酸盐试剂。一些实施例提供一种组合物,该组合物包括:含有DMR的核酸和在此所述的寡核苷酸。一些实施例提供一种组合物,包括:含有DMR的核酸和甲基化敏感性限制酶。一些实施例提供一种组合物,包括:含有DMR的核酸和聚合酶。

[0046] 提供了其他相关方法实施例,用于在从对象(例如胃组织样本;血浆样本;粪便样本)获得的样本中筛查肿瘤,例如。一种方法,包括确定样本中的标记的甲基化状态,该标记包括DMR中的一个碱基,即DMR1-274中的一个或多个(来自表1、2、4或7);将对象样本中的标记的甲基化状态与无癌症的对象的正常对照样本中的标记的甲基化状态进行比较;以及确定对象样本和正常对照样本的甲基化状态中的置信区间和/或p值差异。在一些实施例中,置信区间为90%,95%,97.5%,98%,99%,99.5%,99.9%或99.99%,且p值为0.1,0.05,0.025,0.02,0.01,0.005,0.001,或0.0001。方法的一些实施例提供以下步骤,包括:将含有DMR的核酸与重亚硫酸盐试剂反应,生成重亚硫酸盐反应核酸;将重亚硫酸盐反应核酸测序,提供重亚硫酸盐反应核酸的核苷酸序列;将重亚硫酸盐反应核酸的核苷酸序列与含有无癌症的对象的DMR的核酸的核苷酸序列进行比较,来鉴别两个序列中的差异;当存在差异时,鉴别出所述对象具有肿瘤。

[0047] 本发明提供一种在从对象获得的样本中筛查肿瘤的系统。系统的示例性实施例包

括,例如,在从对象获得的样本中筛查肿瘤的系统(例如,胃组织样本;血浆样本;粪便样本),该系统包括分析元件,被配置来确定样本的甲基化状态;软件元件,被配置来将样本的甲基化状态与记录在数据库中的对照样本或参考样本的甲基化状态进行比较;警报元件,被配置来告知用户有关胃癌的甲基化状态。在一些实施例中,警报通过软件元件被确定,其从多个化验中接收到结果(例如,确定多个标记的甲基化状态,如DMR,例如表1、2、或7中提供的)并计算值或结果,从而基于多个结果进行报告。一些实施例提供与在此提供的各DMR相关的加权参数的数据库,用于计算值或结果和/或警报向用户报告(例如,如医生,护士,临床医生等)。在一些实施例中,报告多个化验的所有结果,且在一些实施例中,使用一个或多个结果来提供评分、值或结果,其基于多个化验的一个或多个结果的组合,指示出对象中的癌症风险。

[0048] 在系统的一些实施例中,样本包括:含有DMR的核酸。在一些实施例中,系统还包括用于分离核酸的元件、用于收集样本的元件(例如用于收集粪便样本的元件)。在一些实施例中,系统包括:含有DMR的核酸序列。在一些实施例中,数据库包括无癌症的对象的核酸序列。此外还提供核酸,例如,一组核酸,各核酸具有包含DMR的序列。在一些实施例中,提供一组核酸,其中各核酸具有无癌症的对象的序列。相关的系统实施例包括在此所述的一组核酸和与所述一组核酸相关的核酸序列的数据库。一些实施例还包括重亚硫酸盐试剂。并且,一些实施例还包括核酸测序器。

[0049] 在一些实施例中,提供一种用于表征人类患者的样本(例如胃组织样本;血浆样本;粪便样本)的方法。例如,在一些实施例中,这些实施例包括:从人类患者的样本中获取DNA;化验DNA甲基化标记的甲基化状态,所述标记包括从表1、2或7的DMR1-274构成的组中选出的差异甲基化区域(DMR)中的碱基;将化验的一个或多个DNA甲基化标记的甲基化状态与无胃肿瘤的人类患者的一个或多个DNA甲基化标记的甲基化水平参考值进行比较。

[0050] 这种方法并不局限于特定类型的人类患者样本。在一些实施例中,样本是胃组织样本。在一些实施例中,样本是血浆样本。在一些实施例中,样本是粪便样本、组织样本、胃组织样本、血样或尿液样本。

[0051] 在一些实施例中,这些方法包括化验多个DNA甲基化标记。在一些实施例中,这些方法包括化验2至11个DNA甲基化标记。在一些实施例中,这些方法包括化验12至107DNA甲基化标记。在一些实施例中,这些方法包括化验样本中的一个或多个DNA甲基化标记的甲基化状态,其中包括确定一个碱基的甲基化状态。在一些实施例中,这些方法包括化验样本中的一个或多个DNA甲基化标记的甲基化状态,其中包括确定多个碱基的甲基化程度。在一些实施例中,这些方法包括化验正向链的甲基化状态或化验反向链的甲基化状态。

[0052] 在一些实施例中,DNA甲基化标记为100个或更少碱基的区域。在一些实施例中,DNA甲基化标记为500个或更少碱基的区域。在一些实施例中,DNA甲基化标记为1000个或更少碱基的区域。在一些实施例中,DNA甲基化标记为5000个或更少碱基的区域。在一些实施例中,DNA甲基化标记是一个碱基。在一些实施例中,DNA甲基化标记位于高CpG密度启动子中。

[0053] 在一些实施例中,所述化验包括使用甲基化特异性聚合酶链式反应、核酸测序、质谱、甲基化特异性核酸酶、基于质量的分离或目标捕获。

[0054] 在一些实施例中,所述测定包括甲基化特异性寡核苷酸的使用。在一些实施方式

中,甲基化特异性寡核苷酸从SEQ ID NO:1-109构成的组中被选出。

[0055] 在一些实施例中,具有从ARHGEF4,ELM01,ABCB1,CLEC11A,ST8SIA1,SFMBT2,CD1D,CYP26C1,ZNF569,或C13ORF18构成的组中选出的标注的染色体区域包括DNA甲基化标记。在一些实施方式中,DMR来自表4并从DMR No.49,43,196,66,1,237,249,250,251及252构成的组中被选出。

[0056] 在一些实施例中,具有从ELM01,ARHGEF4,EMX1,SP9,CLEC11A,ST8SIA1,BMP3,KCNA3,DMRTA2,KCNK12,CD1D,PRKCB,CYP26C1,ZNF568,ABCB1,ELOVL2,PKIA,SFMBT2(893),PCBP3,MATK,GRN2D,NDRG4,DLX4,PPP2R5C,FGF14,ZNF132,CHST2(7890),FLI1,c13orf18,及ZNF569构成的组中选出的标注的染色体区域包括DNA甲基化标记。

[0057] 在一些实施例中,DMR来自表2,并从DMR No.253,251,254,255,256,249,257,258,259,260,261,262,250,263,1,264,265,196,266,118,267,268,269,270,271,272,46,273,252,及237构成的组中被选出。

[0058] 在一些实施例中,所获得的样本是血浆样本,所述标记包括具有标注的染色体区域,所述标注为ELM01,ZNF569,C13orf18,CD1D,ARHGEF4,SFMBT2,PPP25RC,CYP26C1,PKIA,CLEC11A,LRRC4,及ST8SIA1,且其包括标记。

[0059] 在一些实施例中,DMR来自表1,2,4或7,并从DMR No.253,237,252,261,251,196,250,265,256,249,及274构成的组中被选出。

[0060] 在一些实施例中,这些方法包括确定两个DNA甲基化标记的甲基化状态。在一些实施例中,这些方法包括确定表1、2、4或7行中提供的一对DNA甲基化标记的甲基化状态。

[0061] 在一些实施例中,本技术提供一种用于表征从人类患者获得的样本的方法。在一些实施例中,这些方法包括:确定样本中DNA甲基化标记的甲基化状态,所述样本包括从表1、2、4和7的DMR 1-274构成的组中选出的DMR中的碱基;将患者样本中的DNA甲基化标记的甲基化状态与无胃癌的人类对象的正常对照样本中的DNA甲基化标记的甲基化状态进行比较;并确定人类对象和正常对照样本的甲基化状态中的置信区间和/或p值差异。在一些实施例中,置信区间为90%,95%,97.5%,98%,99%,99.5%,99.9%或 $\geq 99.9\%$,p值为0.1,0.05,0.025,0.02,0.01,0.005,0.001,或0.0001。

[0062] 在一些实施例中,本技术提供一种用于表征从人类对象获得的样本(例如胃组织样本;血浆样本;粪便样本)的方法,该方法包括:将含有DMR的核酸与重亚硫酸盐试剂反应,生成重亚硫酸盐反应核酸;将重亚硫酸盐反应核酸测序,提供重亚硫酸盐反应核酸的核苷酸序列;将重亚硫酸盐反应核酸的核苷酸序列与含有无胃癌的对象的DMR的核酸的核苷酸序列进行比较,来鉴别两个序列中的差异。

[0063] 在一些实施例中,本技术提供一种用于表征从人体对象(例如胃组织样本;血浆样本;粪便样本)获得的样本的系统,该系统包括:分析元件,被配置来确定样本的甲基化状态;软件元件,被配置来将样本的甲基化状态与记录在数据库中的对照样本或参考样本的甲基化状态进行比较;警报元件,被配置来基于甲基化状态的组合确定单个值,并告知用户有关胃癌的甲基化状态。在一些实施例中,所述样本包括含有DMR的核酸。

[0064] 在一些实施例中,该系统还包括用于分离核酸的元件。在一些实施例中,该系统还包括用于收集样本的元件。

[0065] 在一些实施例中,所述样本是粪便样本、组织样本、胃组织样本、血样本或尿液样

本。

[0066] 在一些实施例中,数据库包括:含有DMR的核酸序列。在一些实施例中,数据库包括无胃癌的对象的核酸序列。

[0067] 基于在此所包含的内容,进一步的实施例对于本领域中的技术人员来说是显而易见的。

附图说明

[0068] 以下参照附图进行说明,从而可以更好地理解本技术的上述及其他特征、方面、及优点:

[0069] 图1示出ELMO 1对胃癌有很高的辨别能力,如示例2所述。

[0070] 图2示出FRET盒的寡核苷酸序列,用于经QuARTs (定量等位基因特异性实时目标和信号扩增)检测甲基化DNA特征。各FRET序列包括荧光团和猝灭剂,其可以在3个单独的化验中被一起多重复用。

[0071] 图3示出受试者工作特征曲线图,突出显示与单个标记曲线相比较的3个标记组(ELMO1,ZNF569和c13orf18)的血浆中的性能。在100%特异性中,该组检测到的胃癌的敏感性为86%。

[0072] 图4A示出血浆中甲基化ELMO 1的对数绝对链计数,阶段性示出从正常粘膜到阶段4的胃癌进展情况。定量标记水平随GC阶段被增加。

[0073] 图4B示出条形图表,按3个标记组(ELMO 1、ZNF 569和c13orf18)阶段(100%特异性)显示出血浆中胃癌的敏感性。

[0074] 图5A,图5B和图5C示出12个胃癌标记(ELMO1,ZNF569,C13orf18,CD1D,ARHGEF4,SFMBT2,PPP25RC,CYP26C1,PKIA,CLEC11A,LRR4,和ST8SIA1;ELMO1被运行两次,第二次为biplex格式)的性能。在74个血浆样本中以90%特异性(图5A)、95%特异性(图5B)和100%特异性(图5C)的矩阵格式。标记被垂直列出且样本水平排列。样本的排列为正常群位于左边且癌症群位于右边。肯定的命中为浅灰色,而错过的为深灰色。在此,表现最好的标记-ELMO 1-被列为第一为90%特异性,且其余组为100%特异性。该图允许标记以互补方式被评估。

[0075] 图6示出74个血浆样本中的3个胃癌标记(最终组)在100%特异性下的性能。标记被垂直列出且样本水平排列。样本的排列为正常群位于左边且癌症群位于右边。癌症按阶段排列。肯定的命中为浅灰色,而错过的为深灰色(ELMO1,ZNF569和c13orf18)。

[0076] 图7A-图7M示出血浆中12个胃癌标记(ELMO1,ZNF569,C13orf18,CD1D,ARHGEF4,SFMBT2,PPP25RC,CYP26C1,PKIA,CLEC11A,LRR4,和ST8SIA1;ELMO1被运行两次,第二次为biplex格式)的框图(线性比例)。对照样本(N=38)被列于左边,且胃癌病例(N=36)被列于右边。垂直轴是归一化为 β -肌动蛋白链的甲基化百分比。

[0077] 应理解,这些数字不一定是按比例绘制的,数字中的物体也不一定是按比例绘制的。这些图旨在使在此公开的设备、系统、组合物和方法的各种实施例更加清晰和容易理解。在可能的情况下,所有附图使用相同的参照符号来表示相同或相似的部件。此外,应理解,这些图画并不打算以任何方式限制本发明的范围。

具体实施方式

[0078] 本发明提供一种检测肿瘤的技术,特别是用于检测胃癌等癌前和恶性肿瘤的方法、组合物和相关用途,当并不局限于此。正如在此所述的技术,所使用的段标题仅用于组织结构目的,不应被解释为以任何方式限制本技术内容。

[0079] 在各种实施例的详细说明中,为了解释的目的,提出了许多具体细节,以提供对所公开的实施例的彻底理解。然而,本领域的技术人员应理解,这些不同的实施例可以在这些具体细节的情况下进行实践,但也可以不使用这些具体细节。在其他情况下,结构和设备以框图形式被示出。此外,本领域的技术人员可以很容易地认识到,呈现和执行方法的特异性序列是说明性的,且序列可以是多样化的,但仍然保留在此所公开的各种实施例的精神和范围内。

[0080] 本申请中引用的所有文献和类似材料,包括专利、专利申请、文章、书籍、论文和互联网网页,其全部内容都被纳入此处作为参照。除非另有定义,在此使用的所有技术和科学术语的含义与在此所述的各种实施例所属的技术领域中普通技术人员所理解的含义相同。当引用的术语定义与本技术中提供的定义不同时,应以本技术中提供的定义为准。

[0081] 定义

[0082] 为了促进对本技术的理解,以下对若干术语和短语进行了定义。详细说明中还列出了其他定义。

[0083] 在整个说明和权利要求中,除非上下文另有明确规定,以下术语具有此处相关的明确含义。在此使用的短语“在一个实施例中”不一定是指同一实施例,但也可以是同一实施例。此外,本文中使用的“在另一个实施例中”不一定指不同的实施例,但也可以是不同的实施例。因此,如下文所述,本发明的各种实施例可以很容易地结合在一起,且不偏离本发明的范围或精神。

[0084] 此外,如此处所使用的,“或”一词是包含性的“或“作用词,并等同于”和/或“,除非上下文另有明确规定。“基于”一词并不是排他性的,允许基于未描述的其他因素,除非上下文另有明确规定。此外,在整个说明书中,“a”、“an”和“the”的含义包括复数引用。“in”的意思包括“in”和“on”。

[0085] 如本文所述,“核酸”或“核酸分子”一般是指任何核糖核酸或脱氧核糖核酸,其可以是未经修改或修改的DNA或RNA。“核酸”不限于单链和双链核酸。如本文所用,“核酸”一词还包括如上所述的含有一个或多个修改碱基的DNA。因此,为了稳定性或其他原因而主链修改的DNA是一种“核酸”。此处使用的“核酸”一词包括化学、酶学或代谢修改的核酸形式,以及病毒和细胞的DNA特征的化学形式,例如简单和复杂的细胞。

[0086] 术语“寡核苷酸”或“多核苷酸”或“核苷酸”或“核酸”是指具有两个或两个以上脱氧核糖核酸或核糖核酸的分子,优选是超过三个,通常超过十个。确切的大小将取决于许多因素,反过来取决于最终的功能或使用的寡核苷酸。寡核苷酸能够以任何方式产生,包括化学合成、DNA复制、逆转录或上述组合。典型的脱氧核糖核酸是胸腺嘧啶、腺嘌呤、胞嘧啶和鸟嘌呤。典型的核糖核酸核苷酸有尿嘧啶、腺嘌呤、胞嘧啶和鸟嘌呤。

[0087] 如本文所述,核酸的“位点”或“区域”是指核酸的亚区,例如染色体上的基因、单核苷酸、CpG岛等。

[0088] 术语“互补”和“互补性”是指碱基对规则相关的核苷酸(如一个核苷酸)或多核苷

酸(例如一个核苷酸序列)。例如,序列5'-A-G-T-3'与序列3'-T-C-A-5'是互补的。互补可以是“部分的”,其中只有部分核酸的碱基根据碱基对规则进行匹配。或者,核酸之间可能存在“完全”或“完整”互补。核酸链之间的互补程度影响核酸链间杂交的效率和强度。这在扩增反应和依赖于核酸结合的化验方法中特别重要。

[0089] 术语“基因”一词是指核酸(例如,DNA或RNA)序列,其中包含产生RNA或多肽或其前体所必需的编码序列。功能多肽可以由全长编码序列或编码序列的任何部分编码,只要所述多肽的所需活性或功能特性(例如酶活性、配体结合、信号转导等)被保留。“部分”一词用于指某一基因时,指的是该基因的片段。这些片段的大小可以是减去一个核苷酸的几个核苷酸到整个基因序列。因此,“包括至少一部分基因的核苷酸”可以是包括该基因的片段或整个基因。

[0090] 术语“基因”一词还包括结构基因的编码区域,包括相邻于5'和3'端编码区的序列,例如两端距离约为1kb的序列。从而该基因对应于全长mRNA的长度(例如,包含编码、调控、结构和其他序列)。位于编码区5'且存在于mRNA上的序列称为5'非转译序列或未转译序列。位于编码区3'或下游并存在于mRNA上的序列称为3'非转译序列或3'未转译序列。“基因”一词包括基因的cDNA和基因组形式。在一些生物(如真核生物)中,基因的基因组形式或克隆包含被称为“内含子”或“中介区域”或“中介序列”的非编码序列不规则编码区。内含子是基因的片段,转录成核RNA(hnRNA);内含子可包含调控元素,如增强子。核子或最初转录中的内含子被移除或“剪接”;因此,信使RNA(mRNA)转录中没有内含子。mRNA在转译过程中的作用是指定新生多肽中氨基酸的序列或顺序。

[0091] 除了含有内含子外,基因的基因组形式还可包括位于RNA转录上序列的5'和3'端的序列。这些序列被称为“侧翼”序列或区域(这些侧翼序列位于mRNA转录上非转译序列的5'或3')。5'侧翼区可含有调控序列,如调控或影响基因转录的启动子和增强子。3'侧翼区可包含直接终止转录、转录后切割、以及聚腺苷酸化的序列。

[0092] 术语“野生型”一词指的是从自然发生源中分离出来的具有基因特征的基因。“野生型”一词涉及基因产物时是指从自然发生源中分离出来的具有基因特征的基因产物。“自然发生”一词适用于物体时是指物体可以在自然界中被发现的事实。例如,存在于生物体(包括病毒)中的多肽或多核苷酸序列可以从自然界的源中分离出来,并且没有被实验室中的人有意修改,且其是自然发生的。野生型基因通常是在群体中最常见的基因或等位基因,因此被任意指定为基因的“正常”或“野生型”形式。相反,“修改”或“突变”一词指的是基因或基因产物,分别指相对于野生型基因或基因产物,基因或基因产物显示出序列和/或功能特性(例如,改变的特性)被修改。应注意,自然发生的突变体是可以分离出来的;这些突变体是通过与野生型基因或基因产物相比较时具改变的特性被识别的。

[0093] 术语“等位基因”一词是指基因的变异;这些变异包括变异体、多态位点和单核苷酸多态位点、移码和剪接突变,但并不局限于此。等位基因可自然出现在群体中,也可出现在群体的任何特定个体的生命周期中。

[0094] 因此,“变异”和“突变”这两个术语在提及核苷酸序列时,是指经一个或多个核苷酸的不同于另一个通常相关的核苷酸序列的核酸序列。“变异”是两个不同的核苷酸序列之间的差异;通常,一个序列为参考序列。

[0095] “扩增”是核酸复制的一种特例,涉及模板特异性。其与非特异模板复制(例如,复

制依赖模板但不依赖于特异模板)形成对比。模板特异性与复制精确性(例如,合成适当的多核苷酸序列)和核苷酸(ribo-或deoxyribo-)特异性是不同的。模板特异性经常被描述为“目标”特异性。目标序列是“目标”,在某种意义上是从其他核酸中分离出来的。扩增技术主要是为这种分类而设计的。

[0096] 核酸的扩增一般是指生成核苷酸的多个拷贝,或多核苷酸的一部分,通常从少量多核苷酸(例如,单一的多核苷酸分子,10到100个拷贝的多核苷酸分子,这可以或也可以不完全一样),其中扩增产物扩增一般可以被检测。多核苷酸的扩增包括各种化学和酶过程。在聚合酶链反应(PCR)或连接酶链反应(LCR;参照例如美国专利No.5,494,810;其全部内容被纳入此处作为参考)中,从目标或模板DNA的一个或一些拷贝中生成的多个DNA拷贝都为扩增。扩增的其他类型包括:等位基因特异性PCR(参照例如美国专利No.5,639,611;其全部内容被纳入此处作为参考),集合PCR(参照例如美国PCR专利No.5,965,408;其全部内容被纳入此处作为参考),解旋酶依赖性扩增(参照例如美国专利No.7,662,594;其全部内容被纳入此处作为参考),热启动PCR(参照例如美国专利Nos.5,773,258和5,338,671;其全部内容被纳入此处作为参考),intersequence特异性PCR、反向PCR(参照例如Triglia等人(1988)的核酸研究16:8186;其全部内容被纳入此处作为参考),连接介导的PCR(见,例如,Guilfoyle等人的核酸研究25:1854-1858(1997);美国专利No.5,508,169;其全部内容被纳入此处作为参考),甲基化特异性PCR(参照例如Herman等人(1996)PNAS 93(13)9821-9826;其全部内容被纳入此处作为参考),miniprimer PCR、多重连接探针扩增技术(见,例如,Schouten等人(2002)的核酸研究30(12):e57;其全部内容被纳入此处作为参考),多重PCR(参照例如Chamberlain等人,(1988)的核酸研究16(23)11141-11156;Ballabio等人(1990)人类遗传学84(6)571-573;Hayden等人(2008)的BMC遗传学9:80;其全部内容被纳入此处作为参考),巢式PCR,重叠延伸PCR(参照例如,Higuchi等人(1988)的核酸研究16(15)7351-7367;其全部内容被纳入此处作为参考),实时荧光定量PCR(参照例如,Higuchi等人(1992)的生物技术10:413-417;Higuchi等人(1993)的生物技术11:1026-1030;其全部内容被纳入此处作为参考)、逆转录聚合酶链反应(参照例如Bustin,S.A.(2000)J.Molecular Endocrinology 25:169-193;其全部内容被纳入此处作为参考),固相PCR、热不对称交错PCR和Touchdown PCR(参照例如,Don等人的核酸研究(1991)19(14)4008;Roux,K.(1994)生物技术16(5)812-814;Hecker等人(1996)的生物技术20(3)478-485;其全部内容被纳入此处作为参考)。多核苷酸扩增也可以使用数字PCR被完成(参照例如Kalinina等人的核酸研究25:1999-2004,(1997);Vogelstein和Kinzle,Proc Natl Acad Sci USA.96:9236-41,(1999);国际专利公开No.W005023091A2;美国专利申请公开No.20070202525;其全部内容被纳入此处作为参考),但不局限于此。

[0097] 术语“聚合酶链式反应”(PCR)是指K.B.Mullis美国专利Nos.4,683,195,4,683,202和4,965,188的方法,其描述了一种不进行克隆或纯化而提高基因组DNA混合物中目标序列片段浓度的方法。目的序列的扩增过程包括在含有所需目标序列的DNA混合物中引入两个过量的寡核苷酸引物,然后在DNA聚合酶存在下进行精确的热循环序列。该两个引物与各自的双链靶序列互补。为了进行扩增,混合物被变性,然后将引物退火到目标分子中的互补序列。经过退火,引物被延伸到聚合酶,从而形成新的一对互补链。变性、引物退火和聚合酶延伸的步骤可以重复多次(即变性、退火和延伸构成一个“周期”;可以有許多“循环”),以

获得较高浓度的所需目标序列扩增片段。所需目标序列的扩增片段的长度由引物相对于彼此的相对体位决定,因此,该长度是可控参数。由于重复过程,该方法被称为“聚合酶链反应”(PCR)。由于所需的目标序列的扩增片段成为混合物中的主要序列(就浓度而言),因此被称为“PCR扩增”,以及“PCR产物”或“扩增子”。

[0098] 通过酶的选择模板特异性在大多数扩增技术中被实现。扩增酶是指在使用条件下,只处理核酸在不同核酸混合物中的特异性序列的酶。例如,在Q-beta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])。其他核酸不会被这种扩增酶复制。同样,在T7 RNA聚合酶的情况下,这种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,该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模板在连接接合处不匹配(Wu和Wallace(1989)的Genomics 4:560)。最后,依赖热稳定模板的DNA聚合酶(如Taq和Pfu DNA聚合酶)被发现由于其具有耐高温的能力,对于由引物限界及定义的序列显示出较高的特异性;高温可实现有利于引物与目标序列杂交而不与非目标序列杂交的热力条件(H.A.Erlich(ed.),PCR Technology,Stockton Press[1989])。

[0099] 如本文所用,术语“核酸检测化验”是指确定感兴趣核酸的核苷酸组成的任何方法。核酸化验方法包括DNA测序法、探针杂交法、结构特异性切割化验(如Hologic的INVADER化验),其在美国专利Nos.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等人美国PNAS,USA,97:8272(2000);以及US 2009/0253142中被说明,其全部内容被纳入此处作为参考);酶错配切割法(例如Variagenics美国专利Nos.6,110,684,5,958,692,5,851,770,其全部内容被纳入此处作为参考);聚合酶链反应;分支杂交方法(例如Chiron美国专利5,849,481,5,710,264,5,124,246,and 5,624,802,其全部内容被纳入此处作为参考);滚动循环复制(例如美国专利Nos.6,210,884,6,183,960和6,235,502,其全部内容被纳入此处作为参考);NASBA(例如美国专利No.5,409,818,其全部内容被纳入此处作为参考);分子信标技术(例如美国专利No.6,150,097,其全部内容被纳入此处作为参考);电子传感器技术(摩托罗拉美国专利Nos.6,248,229,6,221,583,6,013,170,and 6,063,573,其全部内容被纳入此处作为参考);循环探测技术(例如美国专利Nos.5,403,711,5,011,769,and5,660,988,其全部内容被纳入此处作为参考);Dade Behring信号扩增方法(例如美国专利Nos.6,121,001,6,110,677,5,914,230,5,882,867和5,792,614,其全部内容被纳入此处作为参考);连接酶链式反应(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No.5,288,609,其全部内容被纳入此处作为参考),但不局限于此。

[0100] 术语“可扩增核酸”是指可以通过任何扩增方法被扩增的核酸。在此“可扩增核酸”通常包含“样本模板”。

[0101] 术语“样本模板”是指源自于样本的核酸,该样本被分析用于存在的“目标”(定义如下)。相反,“基质模板”用于指样本模板以外的核酸,可以或也可以不存在于样本中。基质模板通常是不注意的。这可能是接转的结果,也可能是由于核酸污染物的存在,旨在从样本中被净化。例如,来自生物体的核酸,不同于被检测到的可作为测试样本中的基质。

[0102] 术语“引物”是指寡核苷酸,可以是自然产生的,如纯化的限制性内切酶或合成产物。当在引物延伸产物的合成与诱导的核酸链互补的条件下被配置时,能够作为合成的起

始点, (例如存在核苷酸和诱导剂, 如DNA聚合酶, 在合适的温度和pH下)。引物优选是单链, 以达到最大的扩增效率, 但也可以是双链。如果是双链, 则先对引物进行处理以分离其的链, 然后再用于准备延伸产物。优选是引物是寡脱氧核糖核酸。引物的长度必须足够长, 以在诱导剂存在下引发延伸产物的合成。引物的准确长度将取决于温度、引物源和方法的使用等多种因素。

[0103] 术语“探针”一词是指一种寡核苷酸(例如, 核苷酸序列), 无论是自然发生在纯化的限制性内切酶中, 还是通过合成、重组或PCR扩增产生的, 都能够与另一个感兴趣的寡核苷酸杂交。探针可以是单链的, 也可以是双链的。探针在化验、识别和分离特异基因序列(例如“捕获探针”)方面很有用。设想在本发明中使用的任何探针在某些实施例中可被任何“报告分子”标记, 以便在任何化验系统中都可以化验到, 包括但不限于酶(例如, ELISA以及酶组织化学分析)、荧光、放射性和发光系统。本发明无意仅限于任何特定的化验系统或标签。

[0104] 如本文所用, 术语“甲基化”是指胞嘧啶的C5或N4体位处的胞嘧啶甲基化物, 腺嘌呤的N6体位或其他类型的核酸甲基化物。在试管内扩增的DNA通常是非甲基化的, 因为典型的试管内DNA扩增方法不保留扩增模板的甲基化模式。然而, “非甲基化DNA”或“甲基化DNA”也可以分别指原模板未甲基化或甲基化的扩增DNA。

[0105] 因此, 在此使用的“甲基化核苷酸”或“甲基化核苷酸碱基”是指核苷酸碱基上存在在甲基部分, 其中甲基部分不存在于认可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, 而5-甲基胞嘧啶在其嘧啶环5体位处含有甲基部分。因此, 胞嘧啶不是甲基化核苷酸, 且5-甲基胞嘧啶是甲基化核苷酸。在另一个示例中, 胸腺嘧啶在其嘧啶环的5体位处含有甲基部分; 然而, 在此, 由于胸腺嘧啶是DNA的典型核苷酸碱基, 因此, 当胸腺嘧啶存在于DNA中时, 胸腺嘧啶不被视为甲基化核苷酸。

[0106] 在此使用的“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。

[0107] 在此, 核酸分子的“甲基化状态”、“甲基化谱”和“甲基化状况”指的是核酸分子中没有一个或多个甲基化核苷酸碱基的存在。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为甲基化)。不含任何甲基化核苷酸的核酸分子则被认为是非甲基化的。

[0108] 特定核酸序列(例如在此所述的基因标记或DNA区域)的甲基化状态可以指示该序列中每个碱基的甲基化状态, 也可以是指序列中的碱基子集的甲基化状态(例如, 一个或多个胞嘧啶), 或者可以是指有关该序列中区域甲基化密度的信息, 可以或也可以不提供该序列中甲基化发生的位置的精确信息。

[0109] 核酸分子中核苷酸位点的甲基化状态是指核酸分子中特定位点上存在或不存在甲基化核苷酸。例如, 当核酸分子中的第七核苷酸处存在的核苷酸为5-甲基胞嘧啶时, 核酸分子中的第七核苷酸处的胞嘧啶的甲基化状态为甲基化。同样, 当核酸分子中的第七核苷酸处存在的核苷酸为胞嘧啶(且不是5-甲基胞嘧啶)时, 核酸分子中的第七核苷酸处的胞嘧啶的甲基化状态为非甲基化。

[0110] 甲基化状态可以选择性地通过“甲基化值”(例如, 表示为甲基化频率、分数、比率、百分比等)被表示或示出。例如, 可以通过在取决于甲基化的限制酶切后定量完整核酸的数量, 或在重亚硫酸盐反应后比较扩增谱, 或比较重亚硫酸盐处理的和未处理的核酸的序列, 来生成甲基化值。因此, 一个值, 例如甲基化值, 代表甲基化状态, 并可以作为一个位点的多

个拷贝的甲基化状态的定量指示。当需要将样本中序列的甲基化状态与阈值或参考值进行比较时,其特别有用。

[0111] 如本文所使用的,“甲基化频率”或“甲基化百分比(%)”是指甲基化的分子或位点的实体数相对于未甲基化的分子或位点的实体数。

[0112] 因此,甲基化状态描述核酸(例如基因组序列)的甲基化状态。另外,甲基化状态指的是与甲基化有关的特异基因组位点处的核酸片段的特征。这些特征包括,DNA序列内的任何胞嘧啶(C)残基是否甲基化;甲基化C残基的位置;在核酸的任何特定区域中甲基化C的频率或百分比;以及由于等位基因的来源的差异而引起的甲基化的等位基因差异,但不局限于此。术语“甲基化状态”、“甲基化谱”和“甲基化状况”还涉及生物样本中核酸的任何特定区域中的甲基化C或未甲基化C的相对浓度、绝对浓度或图案。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“降低的甲基化”。同样,如果核酸序列内的胞嘧啶(C)残基与另一核酸序列(例如,来自不同区域或来自不同个体等)甲基化,则与另一核酸序列相比,该序列被认为是超甲基化或具有增加的甲基化。或者,如果DNA序列内的胞嘧啶(C)残基与另一核酸序列(例如,来自不同区域或来自不同个体等)未甲基化,则与另一核酸序列相比,该序列被认为是低甲基化或具有降低的甲基化。另外,本文使用的术语“甲基化图案”是指在核酸区域上甲基化和未甲基化核苷酸的集合部位。当甲基化和未甲基化核苷酸的数目在整个区域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图案。序列被说明为“差异甲基化”或具有“甲基化差异”或具有“不同的甲基化状态”表示其在程度上(例如,相对于另一个具有增加或降低的甲基化)、频率、或甲基化图案不同。术语“差异甲基化”是指与癌症阴性样本中的核酸甲基化水平或模式相比,癌症阳性样本中核酸甲基化水平或模式的差异。也可以指在手术后复发癌症的患者与没有复发的患者之间的水平或模式的差异。DNA甲基化的差异甲基化和特异性水平或图案是预后和预测生物标记,例如一旦正确的剪切或预测特征被确定。

[0113] 甲基化状态频率可以用来描述个体的群体或单一个体的样本。例如,甲基化状态频率为50%的核苷酸位点在50%的实体中被甲基化以及在50%的实体中未甲基化。例如,这种频率可以用来描述核苷酸位点或核酸区域在个体的群体或核酸集合中的甲基化程度。因此,当第一群体或核酸分子集合体中的甲基化不同于第二群体或核酸分子集合体中的甲基化时。第一群体或集合体中的甲基化状态频率将不同于第二群体或集合体中的甲基化状态频率。例如,该频率也可以用来描述核苷酸位点或核酸区域在单一个体中的甲基化程度。例如,该频率可以用来描述组织样本中的一组细胞在核苷酸位点或核酸区域被甲基化或非甲基化的程度。

[0114] 在此,“核苷酸位点”是指核苷酸在核酸分子中的位置。甲基化核苷酸的核苷酸位点是指甲基化核苷酸在核酸分子中的位置。

[0115] 通常,人类DNA的甲基化发生在二核苷酸序列上,其中包括相邻的鸟嘌呤和胞嘧啶,胞嘧啶位于鸟嘌呤的5'处(也称为CpG二核苷酸序列)。CpG二核苷酸中的大多数胞嘧啶在人类基因组中甲基化,但在特异性CpG二核苷酸丰富的基因组区域中,即CpG胰岛中未甲基化(参照例如Antequera等人(1990) (1990) Cell 62:503-514)。

[0116] 如本文所使用的,“CpG胰岛”是指基因组DNA的G:C-富集区,含有相对于总基因组DNA的数量增加的CpG二核苷酸。CpG胰岛在长度上可以至少有100,200个或更多的碱基对,其中该区域的G:C含量至少为50%,且观察到的CpG频率与预期频率的比率为0.6;在一些情况下。CpG胰岛在长度上可以至少有500个碱基对,其中该区域的G:C含量至少为55%),且观测到的CpG频率与预期频率的比率为0.65。观测到的CpG频率与预期频率的比率可根据Gardiner-Garden等人(1987) J.Mol.Biol.196:261-281中提供的方法计算。例如可以根据公式 $R = (A \times B) / (C \times D)$ 计算观测到的CpG频率与预期频率的比值,其中R是观测到的CpG频率与预期频率的比率。A是分析的序列中CpG二核苷酸的个数,B是分析的序列中核苷酸的总数,C是分析序列中C核苷酸的总数目。D是分析的序列中G核苷酸的总数。甲基化状态通常在CpG胰岛上被确定,例如在启动子区域。应理解,人类基因组中的其他序列容易发生DNA甲基化,如CpA和CpT(参照Ramsahoye (2000) Proc.Natl.Acad.Sci.USA 97:5237-5242);Salmon和Kaye (1970) Biochim.Biophys.Acta.204:340-351;Grafstrom (1985) 核酸研究13:2827-2842;Nyce (1986) 核酸研究14:4353-4367;Woodcock (1987) 生物化学与生物物理研究通讯.145:888-894)。

[0117] 如本文所使用的,试剂用于修改核酸分子的核苷酸来作为核酸分子的甲基化状态的一部分,或是甲基化特异性试剂是指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、组合物或其他试剂。使用试剂处理核酸分子的方法可以包括:将核酸分子与试剂接触,如果需要,再加上附加步骤以完成所需核苷酸序列的改变。核酸分子核苷酸序列的这种变化会导致核酸分子中的每个甲基化核苷酸被修改成不同的核苷酸。核酸核苷酸序列的这种变化可能导致核酸分子,其中每个非甲基化核苷酸被修改成不同的核苷酸。核酸核苷酸序列中的这种变化可导致核酸分子,其中每个未甲基化的选定核苷酸(例如,每一个未甲基化的胞嘧啶)都被修改成不同的核苷酸。使用这种试剂改变核酸核苷酸序列可导致核酸分子,其中每个核苷酸是甲基化核苷酸(例如,每一个甲基化胞嘧啶)被修改成不同的核苷酸。在此使用的试剂修改选定的核苷酸,指的是修改核酸分子(DNA的C、G、T和A,以及RNA的C、G、U和A)中的四个典型存在的核苷酸中的一个核苷酸的试剂,从而试剂修改一个核苷酸而不修改其他三个核苷酸。在一个示例性实施例中,该试剂修改非甲基化的选择的核苷酸以生成不同的核苷酸。在另一个示例性实施例中,该试剂可以对未甲基化的胞嘧啶核苷酸进行脱氨。一种典型的试剂是重亚硫酸盐。

[0118] 本文所使用的术语“重亚硫酸盐试剂”是指包括一些实施例中的重亚硫酸盐,二硫化物,硫化氢,或上述组合的试剂,用来区分甲基化和非甲基化胞嘧啶,例如在CpG二核苷酸序列中。

[0119] 术语“甲基化化验”是指在核酸序列中确定一个或多个CpG二核苷酸序列甲基化状态的任何方法。

[0120] 术语“MS AP-PCR”(甲基化敏感性任意启动的聚合酶链式反应)是指本领域认可的技术,允许使用富含CG的引物对基因组进行整体扫描,集中于最有可能包含cpG二核苷酸的区域,并在Gonzalgo等人(1997) Cancer Research 57:594-599中被说明。

[0121] 术语“MethyLightTM”是指基于荧光的实时PCR技术,在EADS等人的(1999) Cancer Res.59:2302-2306中被说明。

[0122] 术语“HeavyMethylTM”指的是一种化验,其中,用于覆盖扩增引物之间的或被其遮

盖的CpG体位的甲基化特异性阻断探针(也称为阻断剂)允许核酸样本的甲基化特异性选择性扩增。

[0123] 术语“HeavyMethyl™ MethyLight™”化验涉及HeavyMethyl™ MethyLight™化验, 其是MethyLight™化验的一种变化。其中MethyLight™化验与用于覆盖扩增引物之间的CpG体位的甲基化特异性阻断探针相结合。

[0124] 术语“Ms-SNuPE”(甲基化敏感性单核苷酸引物延伸)指Gonzalzo&Jones(1997) Nucleic Acids Res.25:2529-2531中所说明的本领域认可的化验。

[0125] 术语“MSP”(甲基化特异性PCR)是指Herman et al.(1996) Proc.Natl.Acad.Sci.USA 93:9821-9826以及美国专利No.5,786,146中所说明的本领域认可的甲基化化验。

[0126] 术语“COBRA”(联合亚硫酸亚铁限制分析)是指Xiong&Laird(1997) Nucleic Acids Res.25:2532-2534中所说明的本领域认可的甲基化化验。

[0127] 术语“MCA”(甲基化CpG胰岛扩增)是指Toyota等人(1999) Cancer Res.59:2307-12,以及WO 00/26401A1中所说明的甲基化化验。

[0128] 如本文所用的,“选择的核苷酸”是指核酸分子(DNA的C、G、T和A,以及RNA的C、G、U和A)中的四个典型核苷酸中的一个核苷酸。并且可以包括典型存在的核苷酸的甲基化衍生物(例如,当C是选择的核苷酸时,表示甲基化和非甲基化的C都可以是选择的核苷酸)。而甲基化的选择的核苷酸具体是指甲基化的典型存在的核苷酸,而非甲基化的选择的核苷酸则具体指未甲基化的典型存在的核苷酸。

[0129] “甲基化特异性限制酶”或“甲基化敏感性限制酶”是指一种酶,根据其识别部位的甲基化状态选择性地酶切核酸。当识别部位没有甲基化或半甲基化为特殊的限制性内切酶时,则该切割将不发生或是识别部位甲基化时以明显降低的效率发生。当识别部位甲基化为特殊的限制性内切酶时,则该切割将不发生或是识别部位没有甲基化时以明显降低的效率发生。优选是甲基化特异性限制酶,其识别序列包含CG二核苷酸(例如,类似CGCG或CCCGGG的识别序列)。另外,一些实施例中优选是限制酶,当该二核苷酸中的胞嘧啶在碳原子C5处甲基化时则限制酶不切割。

[0130] 在此使用的“不同核苷酸”是指与选择的核苷酸在化学上不同的核苷酸,通常使不同的核苷酸具有与所选择的核苷酸不同的Watson-Crick碱基对特性。因此,与选择的核苷酸互补的典型存在的核苷酸不同于与其他核苷酸互补的典型存在的核苷酸。例如,当C是选择的核苷酸时,U或T可以是其他核苷酸,例如C与G的互补和U或T与A的互补。与选择的核苷酸互补或与其他核苷酸互补的核苷酸,是指该核苷酸碱基对在高度严格的条件下与选择的核苷酸或其他核苷酸的亲和力高于与四个典型存在的核苷酸中的三个核苷酸互补的核苷酸碱基对。互补的示例是DNA中的Watson-Crick碱基对(例如A-T和C-G)和RNA(例如A-U和C-G)。因此,例如,G碱基对,在高度严格的条件下,与C的亲和力高于G碱基对与G,A或T,因此,当C是选择的核苷酸时,G是与选择的核苷酸互补的核苷酸。

[0131] 如本文所用的,给定标记的“敏感性”涉及样本的百分比,示出DNA甲基化值高于阈值,该阈值区分肿瘤样本和非肿瘤样本。在一些实施例中,阳性被定义为组织学证实的肿瘤,示出DNA甲基化值高于阈值(例如与疾病相关的范围)。伪阴性被定义为组织学证实的肿瘤,示出DNA甲基化值低于阈值(例如与无疾病相关的范围)。因此,敏感性值反映了从已知

的疾病样本中获得的给定标记的DNA甲基化测量的概率将在疾病相关测量的范围内。如在此所定义的,计算的敏感性值的临床相关性表示给定标记可检测出具临床条件时应用于具有该条件的患者的概率估计。

[0132] 如本文所述,给定标记的“特异性”指的是非肿瘤样本的百分比,示出DNA甲基化值低于阈值,区分肿瘤和非肿瘤样本。在一些实施例中,阴性被定义为组织学证实的非肿瘤样,示出低于阈值的DNA甲基化值。(例如与无疾病相关的范围),且伪阳性定义为组织学证实的非肿瘤样,示出高于阈值的DNA甲基化值(例如与疾病相关的范围)。因此,特异性的值反映了从已知的非肿瘤样本获得的给定标记的DNA甲基化测量的概率将在非疾病相关测量的范围内。如在此所定义的,计算的特异性值的临床相关性表示给定标记可检测出存在临床条件时应用于不具该条件的患者的概率估计。

[0133] 此处使用的术语“AUC”是“曲线下面积”的缩写。特别是指受试者工作特征(ROC)曲线下的面积。ROC曲线是诊断试验的不同可能切点的真阳性率与伪阳性率的曲线。显示出敏感性和特异性之间的权衡取决于所选择的切点(敏感性的任何增加都会伴随着特异性的下降)。ROC曲线下面积(AUC)是诊断试验准确性的衡量标准(面积越大越好;最优值为1;随机检验的ROC曲线位于对角线上,面积为0.5;参照J.P.Egan.(1975)信号检测理论和ROC分析,Academic Press,New York)。

[0134] 如此处所用,“肿瘤”一词是指“组织的异常肿块,其生长超过正常组织,且与正常组织的生长不协调”,参照Willis RA“肿瘤在人体中的扩散”London,Butterworth&Co,1952。

[0135] 在此,“腺瘤”一词是指起源于腺的良性肿瘤。虽然这些生长是良性的,但随着时间的推移,其可能会变成恶性。

[0136] 术语“癌前”或“肿瘤前”及其对应物是指正在发生恶性转化的任何细胞增殖性疾病。

[0137] 肿瘤、腺瘤、癌症等的“部位”是指肿瘤、腺瘤、癌症等所在的对象身体中的组织、器官、细胞类型、解剖区域、身体部位等。

[0138] 如本文所使用的,“诊断”测试应用包括检测或识别被试的疾病状态或状态,确定被试感染某一特定疾病或状况的可能性。确定有疾病或病情的患者对治疗有反应的可能性,确定有疾病或状态的患者的预后(或其可能的进展或倒退),并确定治疗对有疾病或病情的患者的影响。例如,诊断可以用于检测患肿瘤的对象的存在或可能性,或对象对化合物(例如药物,例如药物)或其他治疗作出有利反应的可能性。

[0139] 本文所用的术语“标记”是指能够通过区分癌细胞与正常细胞(例如,基于其甲基化状态)来诊断癌症的物质(例如,核酸或核酸的区域)。

[0140] 术语“分离”与核酸有关时,如在“分离的寡核苷酸”中,是指至少从一种通常与其天然源相关联的污染物核酸中被识别和分开的核酸序列。分离的核酸以不同于自然界中的核酸的形式或环境存在。相反,非分离的核酸,如DNA和RNA,则存在于自然界中。非分离核酸的示例包括:在邻近基因附近的宿主细胞染色体上发现的特定DNA序列(例如基因);RNA序列,例如编码特定蛋白质的特定mRNA序列。在细胞中被发现,作为具许多其他mRNAs的混合物,编码大量的蛋白质。然而,编码特定蛋白质的分离核酸包括,例如在通常表达该蛋白的细胞中含有这种核酸,其中核酸位于与自然细胞不同的染色体位置。或者两侧是不同于自

自然界中发现的其他核酸序列。分离的核酸或寡核苷酸以单链或双链形式存在。当分离的核酸或寡核苷酸被用来表达蛋白质时,该寡核苷酸将至少包含官能或编码链(即,该寡核苷酸可以是单链),但可同时包含官能和抗官能链(即寡核苷酸可以是双链)。从自然或典型环境中分离出来的核酸,可与其他核酸或分子结合。例如,分离的核酸可存在于宿主细胞中,例如用于异源表达。

[0141] 术语“纯化的”是指核酸或氨基酸序列分子从自然环境中被移除、分离或分开。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%游离,优选是至少75%游离,更优选时至少90%游离,来自相关天然的其他成分。如在此使用的,术语“纯化的”或“纯化”还指从样本中除去污染物。污染蛋白质的去除导致样本中所关注的多肽或核酸的百分比增加。在另一个示例中,重组多肽在植物、细菌、酵母或哺乳动物宿主细胞中被表达,并且通过去除宿主细胞蛋白来纯化多肽;由此样本中重组多肽的百分比增加。

[0142] 术语“组合物包括”给定的多核苷酸序列或多肽,广义上指含有给定的多核苷酸序列或多肽的任何组合物。该组合物可包括含有盐(如NaCl)、洗涤剂(例如SDS)和其他成分(例如Denhardt溶液、干牛奶、鲑鱼精子DNA等)的水溶液。

[0143] 术语“样本”以其最广泛的意义被使用。在一种意义上,其可以指动物细胞或组织。在另一种意义上,其意味着包括从任何源获得的样本或培养物,以及生物和环境样本。生物样本可从植物或动物(包括人类)获得,并包括流体、固体、组织和气体。在一些实施例中,样本是血浆样本。在一些实施例中,样本是胃组织样本。在一些实施例中,样本是粪便样本。环境样本包括环境材料,例如表面物质、土壤、水和工业样本。这些示例不应解释为限制应用于本发明的样本类型。

[0144] 在此使用的“远程样本”涉及从部位不是细胞、组织或器官的样本中间接收集的样本。例如,当来自胰腺的样本材料在粪便样本(例如,不是直接从胰腺提取的样本)中被评估时,样本为远程样本。

[0145] 如此处所使用的,术语“患者”或“对象”是指接受本技术提供的各种测试的生物体。术语“对象”包括动物,优选是哺乳动物,包括人类。在优选的实施例中,对象是灵长类动物。在更优选的实施方式中,对象是人类。

[0146] 如本文所用,术语“试剂盒”是指用于交付材料的任何交付系统。在反应化验方面,这种交付系统包括允许从一个地点至另一地点储存、运输或交付反应试剂(例如寡核苷酸、酶等,在适当的容器中)和/或辅助材料(例如缓冲剂、进行化验的书面指令等)的系统。例如,试剂盒包括一个或多个外壳(例如盒子),其中包含相关的反应试剂和/或辅助材料。如本文所用,“零碎试剂盒”是指由两个或多个单独的容器组成的交付系统,每个容器包含整个试剂盒成分的子部分。容器可以一起或单独交付给预定的接收方。例如,第一容器可包含用于化验的酶,而第二容器包含寡核苷酸。“零碎试剂盒”旨在包括由联邦食品药品和化妆品法第520(e)节规定的含有分析专用试剂(ASR's)的试剂盒,但不局限于此。事实上,任何具有两个或两个以上的独立容器且各容器包括整个试剂盒成分的子部分的交付系统,都表示“零碎试剂盒”一词。与此相反,“组合试剂盒”是指在单一容器中包括反应化验的所有成分交付系统(例如,各所需成分被包括在单一盒子中)。术语“试剂盒”既包括零碎试剂盒,也包括组合试剂盒。

[0147] 本技术的实施例

[0148] 本发明提供了一种检测肿瘤的技术,特别是用于检测胃癌等癌前和恶性肿瘤的方法、组合物和相关用途。在病例对照研究中,通过将胃部癌(如胃癌)患者的肿瘤DNA标记的甲基化状态与来自对照对象的相同DNA标记的甲基化状态(参照示例1和示例2)进行比较,来鉴定标记。另一些实验鉴定10个用于检测胃癌的最佳标记(ARHGEF4,ELM01,ABCB1,CLEC11A,ST8SIA1,SFMBT2,CD1D,CYP26C1,ZNF569,和C13orf18;参照示例2和表4)。另一些实验鉴定30个用于检测胃癌的最佳标记(参照示例1和表2)。另一些实验确定了在血浆中化验胃癌(例如胃癌)的12个最佳标记(ELM01,ZNF569,C13orf18,CD1D,ARHGEF4,SFMBT2,PPP25RC,CYP26C1,PKIA,CLEC11A,LRR4,和ST8SIA1;参照示例3和表1、2、4和7)。

[0149] 标记和/或标记组(例如,具有从ARHGEF4,ELM01,ABCB1,CLEC11A,ST8SIA1,SFMBT2,CD1D,CYP26C1,ZNF569,或C13orf18中选出的标注的染色体区域并包括标记(参照表4))在病例对照研究中,通过将具胃部癌(如胃癌)的对象的胃组织的DNA标记的甲基化状态与来自对照对象的相同DNA标记的甲基化状态(参照例2)进行比较被鉴定。

[0150] 标记和/或标记组(例如,具有从ELM01,ARHGEF4,EMX1,SP9,CLEC11A,ST8SIA1,BMP3,KCNA3,DMRTA2,KCNK12,CD1D,PRKCB,CYP26C1,ZNF568,ABCB1,ELOVL2,PKIA,SFMBT2(893),PCBP3,MATK,GRN2D,NDRG4,DLX4,PPP2R5C,FGF14,ZNF132,CHST2(7890),FLI1,c13orf18,或ZNF569中选出的标注的染色体区域并包括标记(参照表2))在病例对照研究中,通过将DNA标记的甲基化状态(例如具胃部癌(如胃癌)的对象的胃组织)与来自对照对象的相同DNA标记的甲基化状态(参照示例1)进行比较被鉴定。

[0151] 标记和/或标记组(例如,具有从ELM01,ZNF569,C13orf18,CD1D,ARHGEF4,SFMBT2,PPP25RC,CYP26C1,PKIA,CLEC11A,LRR4,和ST8SIA1中选出的标注的染色体区域并包括标记(参照表1、2、4和7))在病例对照研究中,通过将DNA标记的甲基化状态(例如具胃部癌(如胃癌)的对象的血浆)与来自对照对象的相同DNA标记的甲基化状态(参照示例3)进行比较被鉴定。

[0152] 此外,本技术提供各种标记组,例如,在一些实施例中,该标记包括具有标注的染色体区域,其标注为ARHGEF4,ELM01,ABCB1,CLEC11A,ST8SIA1,SFMBT2,CD1D,CYP26C1,ZNF569,或C13orf18,并包括标记(参照表4)。此外,实施例提供用于分析表4中的DMR的方法,即DMR No.49,43,196,66,1,237,249,250,251或252。一些实施例提供用于确定标记的甲基化状态的方法,其中染色体区域具有标注,即ELM01,ARHGEF4,EMX1,SP9,CLEC11A,ST8SIA1,BMP3,KCNA3,DMRTA2,KCNK12,CD1D,PRKCB,CYP26C1,ZNF568,ABCB1,ELOVL2,PKIA,SFMBT2(893),PCBP3,MATK,GRN2D,NDRG4,DLX4,PPP2R5C,FGF14,ZNF132,CHST2(7890),FLI1,c13orf18,或ZNF569,并包括标记(参照表2)。此外,实施例提供用于分析表2中的DMR的方法,即DMR No.253,251,254,255,256,249,257,258,259,260,261,262,250,263,1,264,265,196,266,118,267,268,269,270,271,272,46,273,252,或237。在一些实施例中,所获得的样本是血浆样本,所述标记包括具有标注的染色体区域,其标注为ELM01、ZNF569、C13orf18、CD1D、ARHGEF4、SFMBT2、PPP25RC、CYP26C1、PKIA、CLEC11A、LRR4和ST8SIA1,并包括标记。此外,在该实施例中,样本是血浆样本,方法是从表1、2、4和7中分析DMR,即DMR No.253,251,254,255,256,249,257,258,259,260,261,262,250,263,1,264,265,196,266,118,267,268,269,270,271,272,46,273,252,或237。在一些实施例中,所述方法包括确定两个标记的甲基化状态,例如表1、2、4或7行中提供的一对标记。

[0153] 虽然这里公开的是一些示出的实施例,但应该理解,这些实施例是通过示例而不是通过限制来呈现的。

[0154] 在具体方面,本技术提供用于鉴定、确定和/或分类诸如胃部癌(例如胃癌)的癌症的组合物和方法。该方法包括确定从对象(例如,粪便样本、胃组织样本、血浆样本)中分离的生物样本中至少一个甲基化标记的甲基化状态,其中标记的甲基化状态的变化指示出胃部癌(如胃癌)的存在、类别、或部位。特定实施例涉及含有差异甲基化区域(DMR,例如DMR 1-274,参照表1、2、4和7)的标记,用于诊断(例如,筛查)肿瘤细胞增生性疾病(例如癌症),包括疾病的癌前阶段的早期检测。

[0155] 此外,这些标记被用来区别肿瘤和良性细胞增殖性疾病。特别是,本技术公开一种将肿瘤细胞增生性疾病与良性细胞增殖性疾病区分开来的方法。

[0156] 除了实施例中至少一个标记,标记的区域,或含有DMR的标记的碱基(如DMR,例如DMR 1-274)的甲基化分析被提供并在表1,2,4或7中被列出,本技术还提供含有至少一个标记,标记的区域,或含有DMR的标记的碱基的标记组来用于检测癌症,特别是胃癌。

[0157] 本技术的一些实施例基于至少一个标记、标记的区域、或含有DMR的标记的碱基的CpG甲基化状态分析。

[0158] 在一些实施例中,本技术提供一种使用重亚硫酸盐技术结合一个或多个甲基化化验方法来确定至少一个含有DMR的标记(例如DMR 1-274,参照表1、2、4和7)中的CpG二核苷酸序列的甲基化状态。基因组CpG二核苷酸可以是甲基化的,也可以是非甲基化的(分别称为上甲基和下甲基)。然而,本发明的方法适用于在远程样本(例如血液、器官排出液或粪便)的基质内分析非均匀性质的生物样本,例如低浓度的肿瘤细胞或来自其的生物材料。因此,当分析该样本中CpG位点的甲基化状态时,可以使用定量化验来确定特定CpG体位的甲基化水平(例如百分比、分数、比率、比例或程度)。

[0159] 根据本技术,检测含有DMR的标记中的CpG二核苷酸序列的甲基化状态,对胃癌等癌症的诊断和定性都有一定的实用价值。

[0160] 标记的组合

[0161] 在一些实施例中,该技术涉及评估含有来自表1(例如DMR Nos.1-248)或表4(例如DMR Nos.,.49,43,196,66,1,237,249,250,251或252)或表2(DMR Nos.253,251,254,255,256,249,257,258,259,260,261,262,250,263,1,264,265,196,266,118,267,268,269,270,271,272,46,273,252,或237)或表7(DMR No.274)的标记组合或含有DMR的更多标记的甲基化状态。在一些实施例中,评估一个以上标记的甲基化状态提高了用于鉴定对象中的肿瘤(例如胃癌)的筛查或诊断的特异性和/或敏感性。在一些实施例中,标记或标记组合区分肿瘤的类型和/或位置。

[0162] 各种癌症通过各种标记组合被预测,例如,通过与预测的特异性和敏感性有关的统计技术来识别。该技术提供了用于鉴定一些癌症的预测组合和有效预测组合的方法。

[0163] 用于化验甲基化状态的方法

[0164] 在5-甲基胞嘧啶存在的情况下,最常用的分析核酸的方法是基于重亚硫酸盐的方法,其经Frommer等人被说明用于检测DNA中5-甲基胞嘧啶(Frommer等人(1992) Proc.Natl.Acad.Sci.USA 89:1827-31,其全部内容被明确纳入此处作为参照)或其变异。映射5-甲基胞嘧啶的重亚硫酸盐方法是基于观察胞嘧啶,而不是5-甲基胞嘧啶与亚硫酸氢

离子(又称重亚硫酸盐)反应。该反应通常按照以下步骤进行:首先,胞嘧啶与亚硫酸氢根离子反应生成磺化胞嘧啶。然后,磺化反应中间体的自发脱氨反应生成磺化尿嘧啶。最后,在碱性条件下对磺化尿嘧啶进行脱硫,生成尿嘧啶。由于尿嘧啶与腺嘌呤形成碱基对(从而表现为胸腺嘧啶)因此可进行检测,而5-甲基胞嘧啶碱基对与鸟嘌呤形成碱基对(从而表现为胞嘧啶)。从而能够通过重亚硫酸盐基因组测序(Grigg G,&Clark S,Bioessays (1994) 16:431-36;Grigg G,DNA Seq. (1996) 6:189-98)或是例如美国专利No.5,786,146中公开的甲基化特异性PCR(MSP),来辨别甲基化胞嘧啶与非甲基化胞嘧啶。

[0165] 一些常规技术涉及到将要分析的DNA封在琼脂糖基质中,从而防止DNA(重亚硫酸盐只与单链DNA反应)的扩散和复性,以及通过快速透析取代沉淀和纯化步骤(Olek A等人(1996)“一种修改和改进的重亚硫酸盐基胞嘧啶甲基化的分析方法”核酸研究24:5064-6)。因此,可以分析单个细胞的甲基化状态,示出该方法的实用性和敏感性。用于检测5-甲基胞嘧啶的常规方法在Rein等人(1998)核酸研究26:2255中被提出。

[0166] 重亚硫酸盐技术通常涉及在重亚硫酸盐处理之后扩增已知核酸的短、特异片段,然后通过测序对产物进行化验(Olek&Walter (1997) Nat. Genet. 17:275-6)。或者引物延伸反应(Gonzalzo&Jones (1997) Nucleic Acids Res. 25:2529-31;WO 95/00669; U.S. Pat. No. 6,251,594)来分析单个胞嘧啶体位。一些方法使用酶切(Xiong&Laird (1997) 核酸研究25:2532-4)。在本技术领域还描述了杂交检测技术(Olek et al., WO 99/28498)。此外,使用重亚硫酸盐技术检测单个基因的甲基化被说明(Grigg&Clark (1994) Bioessays 16:431-6,;Zeschnick等人(1997) Hum Mol Genet. 6:387-95;Feil等人(1994) 核酸研究22:695;Martin等人(1995) Gene 157:261-4;WO 9746705;WO 9515373)。

[0167] 各种甲基化化验程序在本领域为已知技术,并可根据本技术与重亚硫酸盐处理相结合。这些化验可检测出核酸序列中一个或多个CpG二核苷酸(例如CpG胰岛)的甲基化状态。除其他技术外,这些化验包括重亚硫酸盐处理核酸的测序、PCR(用于序列特异性扩增)、Southern blot杂交分析和甲基化敏感性限制酶的使用。

[0168] 例如,基因组测序被简化来分析甲基化图案和5-甲基胞嘧啶分布,使用重亚硫酸盐处理(Frommer等人(1992) Proc. Natl. Acad. Sci. USA 89:1827-1831)。此外,从重亚硫酸盐转化的DNA中扩增出的PCR产物的限制酶切法可用于评估甲基化状态,如Sadri&Hornsby (1997) Nucl. Acids Res. 24:5058-5059中被说明或COBRA(结合的重亚硫酸盐限制酶切分析)(Xiong&Laird (1997) Nucleic Acids Res. 25:2532-2534)已知的方法中被示出。

[0169] COBRATM分析是有利于检测少量的基因组DNA中特定位点处的DNA甲基化水平的定量甲基化化验(Xiong&Laird, 核酸研究25:2532-2534, 1997)。简单地说,限制酶切可以用来揭示重亚硫酸盐处理的DNA的PCR产物中甲基化引起的序列差异。根据Frommer等人(Proc. Natl. Acad. Sci. USA 89:1827-1831, 1992)所说明的程序,甲基化引起的序列差异经标准重亚硫酸盐处理首先引入基因组DNA。PCR扩增重亚硫酸盐转化的DNA,然后使用特别针对感兴趣的CpG胰岛的引物,随后是限制内切酶酶切,凝胶电泳,并使用特定的,标签的杂交探针进行检测。原始DNA样本中的甲基化水平通过宽频谱DNA甲基化水平的线性量化方式中的酶切和未酶切PCR产物的相对数量被示出。此外,该技术还可以应用于从显微切割石蜡包埋的组织样本中获得的DNA。

[0170] 典型试剂(例如,可在典型的基于COBRATM的试剂盒中可以找到)。COBRATM分析可包

括:针对特定位点的PCR引物(如特异基因、标记、DMR、基因区域、标记区域、重亚硫酸盐处理的DNA序列、CpG胰岛等);限制酶和适当的缓冲剂;基因杂交寡核苷酸;对照杂交寡核苷酸;寡核苷酸探针的激酶标签试剂盒;和标签的核苷酸,但不局限于此。此外重亚硫酸盐转化试剂可包括:DNA变性缓冲剂;磺化缓冲剂;DNA复原试剂或试剂盒(例如沉淀、超滤、亲和柱);脱硫缓冲剂;以及DNA复原成分。

[0171] 优选是类似“MethyLight™”的化验(基于荧光的实时PCR技术)(EADS等人的癌症研究59:2302-2306,1999),Ms-SNuPE™(甲基化敏感性单核苷酸引物延伸)反应(Gonzalogo& Jones,核酸研究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No.5,786,146),以及甲基化CpG胰岛扩增(“MCA”;Toyota等人的癌症研究59:2307-12,1999)被单独使用,或与其中一个或多个方法结合使用。

[0172] “HeavyMethyl™”化验技术是一种定量方法,基于重亚硫酸盐处理的DNA的甲基化特异性扩增来评估甲基化差异。甲基化特异性阻断探针(“阻断剂”)覆盖位于扩增引物之间或由其覆盖的CpG体位,实现核酸样本的甲基化特异性选择性扩增。

[0173] 术语“HeavyMethyl™ MethyLight™”化验是指HeavyMethyl™化验,其是MethyLight™化验的一种变化。其中MethyLight™化验与覆盖扩增引物之间的CpG体位甲基化特异性阻断探针相结合。HeavyMethyl™化验也可用来与甲基化特异性扩增引物相结合。

[0174] 典型试剂(例如,在典型的基于MethyLight™的试剂盒中可以找到)。HeavyMethyl™分析可以包括:针对特定位点的PCR引物(例如,特异基因、标记、DMR、基因区域、标记区域、重亚硫酸盐处理的DNA序列、CpG胰岛或重亚硫酸盐处理的DNA序列或CpG胰岛等);阻断寡核苷酸;优化的PCR缓冲剂和脱氧核苷酸;以及Taq聚合酶,但不局限于此。

[0175] MSP(甲基化特异性PCR)允许评估CpG胰岛内CpG位点的任何基的甲基化状态,而不依赖甲基化敏感性限制酶的使用(Herman等人Proc.Natl.Acad.Sci.USA 93:9821-9826,1996;美国专利No.5,786,146)。简单地说,DNA经重亚硫酸盐被修改,其将未甲基化的胞嘧啶转化为尿嘧啶,然后用甲基化和非甲基化DNA特异性引物扩增产物。MSP只需少量DNA,对给定的cpG胰岛位点的0.1%甲基化等位基因具敏感性,并可在从石蜡包埋的样本中提取的DNA上进行。用于MSP分析的典型试剂(例如可在典型的MSP试剂盒中被找到)但局不限于此:针对特定位点的甲基化和非甲基化PCR引物(例如,特异基因、标记、DMR、基因区域、标记区域)。重亚硫酸盐处理的DNA序列,CpG胰岛等;优化的PCR缓冲剂和脱氧核苷酸,以及特异性探针。

[0176] MethyLight™化验是高通量的定量甲基化化验,其利用基于荧光的实时PCR(例如TaqMan®),其在PCR步骤之后不需要进一步的操作(EADS等人的癌症研究59:2302-2306,1999)。简单地说,MethyLight™过程开始于基因组DNA的混合样本,其根据标准过程在重亚硫酸盐反应中被转化为甲基化引起的序列差异混合群体(重亚硫酸盐过程将未甲基化的胞嘧啶残基转化为尿嘧啶)。然后在“偏置”反应中执行基于荧光的PCR,例如通过PCR引物重叠已知的CpG二核苷酸。序列辨别均以扩增过程水平和荧光检测过程水平进行。

[0177] MethyLight™化验用作核酸中甲基化图案的定量测试,例如基因组DNA样本,以探针杂交水平进行序列辨别。在定量版本中,PCR反应提供甲基化特异性扩增,当存在荧光探针时,该探针重叠于特定的甲基化部位。一种反应提供对DNA输入量的无偏置对照,在这种

反应中,引物和探针都不覆盖任何CpG二核苷酸。此外通过使用不覆盖已知甲基化部位的对照寡核苷酸(例如HeavyMethyl™和MSP技术的基于荧光版本)或通过覆盖潜在甲基化部位的寡核苷酸来探测偏置PCR群体,从而实现基因组甲基化的定性测试。

[0178] MethyLight™过程与任何合适的探针一起使用(例如“TaqMan”)。®“探测仪,一个骑着灯的人。®例如,在某些应用中,双链基因组DNA用重亚硫酸盐处理,并用Taqman进行两组PCR反应之一。®探针,例如MSP引物和/或重甲基阻滞剂寡核苷酸和TaqMan。®探测器。塔克曼。®探针是用荧光“报告”和“猝灭”分子双重标记的,被设计成专为相对高GC含量的区域而设计,使其在PCR循环中在10℃左右的温度下熔化,比正引物或反向引物高。这允许TaqMan。®探针在PCR退火/延伸过程中保持完全杂交。当Taq聚合酶在PCR过程中合成一条新的链时,它最终会到达退火后的TaqMan。®探测器。Taq聚合酶5‘至3’内切酶活性将取代TaqMan。®探针通过消化,释放荧光报告分子,用实时荧光化验系统定量化验其非猝灭信号。

[0179] 用于MethyLight™分析的典型试剂(例如,可在典型的MethyLight™试剂盒中被找到)可以包括:针对特定位点的PCR引物(例如,特异基因、标记、DMR、基因区域、标记区域、重亚硫酸盐处理的DNA序列、CpG岛等); TaqMan®或Lightcycler®探针;优化的PCR缓冲剂和脱氧核苷酸;Taq聚合酶,但不局限于此。

[0180] QM™(定量甲基化)化验是用于基因组DNA样本中甲基化图案的替代性定量测试,其中以探针杂交水平进行序列辨别。在该定量版本中,PCR反应提供无偏置扩增,当存在荧光探针时,该探针重叠于特定的甲基化部位。一种反应提供对DNA输入量的无偏置对照,在这种反应中,引物和探针都不覆盖任何CpG二核苷酸。此外通过使用不覆盖已知甲基化部位的对照寡核苷酸(例如HeavyMethyl™和MSP技术的基于荧光版本)或通过覆盖潜在甲基化部位的寡核苷酸来探测偏置PCR群体,从而实现基因组甲基化的定性测试。

[0181] QM™过程可以与任何合适的探针一起使用,例如,“TaqMan®”探针, Lightcycler®探针,在扩增过程中。例如,双链基因组DNA经重亚硫酸盐被处理,并接受无偏置引物和TaqMan®探针。TaqMan®探针是用荧光“报告基因”和“猝灭剂”分子被双重标签,被设计成专为相对高GC含量的区域而设计,使其在PCR循环中以相比正向或反向引物高10℃的温度下熔化。这允许TaqMan®探针在PCR退火/延伸过程中保持完全杂交。当Taq聚合酶在PCR期间合成新的链时,将最终实现退火的TaqMan®探针。Taq聚合酶5‘至3’内切酶活性将通过酶切取代TaqMan®探针,释放荧光报告基因分子,通过使用实时荧光检测系统定量检测其非猝灭信号。用于QM™分析的典型试剂(例如,可在典型的基于QM™的试剂盒中被找到)可以包括:针对特定位点的PCR引物(例如,特异基因、标记、DMR、基因区域、标记区域、重亚硫酸盐处理的DNA序列、CpG岛等); TaqMan®或Lightcycler®探针;优化的PCR缓冲剂和脱氧核苷酸;Taq聚合酶,但不局限于此。

[0182] Ms-SNuPE™技术是一种定量的方法,用于评估特定CpG部位的甲基化差异,这种方法基于重亚硫酸盐处理的DNA,然后是单核苷酸引物延伸(Gonzalgo&Jones,核酸研究25: 2529-2531,1997)。简单地说,基因组DNA与重亚硫酸盐反应,将未甲基化的胞嘧啶转化为尿

嘧啶,而5-甲基胞嘧啶则保持不变。随后使用特别针对重亚硫酸盐转化的DNA的PCR引物执行目标序列的扩增,且产生物被分离并作为模板在感兴趣的cpG部位进行甲基化分析。少量的DNA可以被分析(例如,显微解剖的病理切片),避免利用限制酶来确定CpG部位的甲基化状态。

[0183] 用于SNUPE™分析的典型试剂(例如,可在典型的基于Ms-SNUPE™的试剂盒中被找到)可包括:针对特定位点的PCR引物(如特异基因、标记、DMR、基因区域、标记区域、亚硫酸亚铁处理DNA序列、cpG胰岛等);优化的PCR缓冲剂和脱氧核苷酸;凝胶提取试剂盒;阳性对照引物;针对特定位点的Ms-SNUPE™引物;反应缓冲剂(用于Ms-SNUPE反应);以及标签的核苷酸,但不局限于此。此外,重亚硫酸盐转化的试剂可包括:DNA变性缓冲剂;磺化缓冲剂;DNA复原试剂或试剂盒(例如沉淀、超滤、亲和柱);脱硫缓冲剂;以及DNA复原成分。

[0184] 简化基因组重亚硫酸盐测序(RRBS)从重亚硫酸盐处理开始,将所有未甲基化的胞嘧啶转化为尿嘧啶,然后酶切(例如,识别包含类似MspI的CG序列的部位的酶)和连接到适配剂配体后的片段完全测序。限制性内切酶的选择丰富了CpG稠密区的片段,减少了分析过程中可能映射到多个基因位置的冗余序列数。因此,RRBS通过选择限制性片段的子集(例如,通过制备凝胶电泳进行尺寸选择)进行测序,从而降低了核酸样本的复杂性。与全基因组重亚硫酸盐测序相比,酶切产生的每个片段都含有至少一个CpG二核苷酸的DNA甲基化信息。因此,RRBS丰富了启动子、CpG胰岛和其他基因组特征的样本,在这些区域具有高频率的限制酶切部位,从而为评估一个或多个基因组位点的甲基化状态提供了一种方法。

[0185] 典型的RRBS方案包括以下步骤:通过MspI的限制酶来酶切核酸样本、填充至悬垂和A尾、连接适配剂、重亚硫酸盐转化和PCR等。例如参照(2005)“以单核苷酸解析映射临床样本的基因组规模DNA甲基化”Nat Methods 7:133-6;Meissner等人(2005)“用于相对高分辨率DNA甲基化分析的简化基因组重亚硫酸盐测序”核酸研究33:5868-77。

[0186] 在一些实施例中,使用定量等位基因特异性实时目标和信号扩增(QuARTS)方法来评估甲基化状态。在各QuARTS化验中依次发生三种反应,包括:最初反应中的扩增(反应1)和目标探针裂解(反应2);第二反应中的FRET裂解和荧光信号生成(反应3)。当用特异引物扩增目标核酸时,带有皮瓣序列的特异性化验探针松散地与该扩增子结合。目标结合部位处特异侵袭性寡核苷酸的存在导致分裂,经探针与皮瓣序列之间的切割使皮瓣序列释放。所述皮瓣序列与相应的FRET盒的非簪部分互补。因此,该皮瓣序列在FRET盒上起到侵入性寡核苷酸的作用,并影响FRET盒荧光基团和猝灭剂之间的分裂,从而产生荧光信号。该分裂反应可以切割各目标的多个探针,从而各皮瓣释放出多个荧光基团,提供指数信号扩增。通过使用不同染料的FRET盒,QuARTS可以很好地检测出单个反应中的多个目标。例如参照Zou等人(2010)“具有新甲基化特异性技术的甲基化标记的敏感性量化”Clin Chem 56:A199; U.S.Pat.Appl.Ser.Nos.12/946,737,12/946,745,12/946,752和61/548,639。

[0187] 术语“重亚硫酸盐试剂”是指由重亚硫酸盐、二硫、亚硫酸氢或其组合组成的试剂,如在此所述的,用于区分甲基化和非甲基化CpG二核苷酸序列。所述处理的方法在本技术中是已知的(例如PCT/EP2004/011715,其被纳入此处作为参考)。重亚硫酸盐处理优选是在变性溶剂的存在下被执行,例如n-烷基乙二醇或二乙二醇二甲醚(DME)的存在下,或在二恶烷或二恶烷衍生物存在下进行,但并局限于此。在一些实施例中,在1%-35%(v/v)之间的浓度中使用变性溶剂。在一些实施例中,重亚硫酸盐反应是在清除剂(如但不限于铬衍生物)

的存在下进行的,例如6-羟基--2,5,7,8,-四甲基色烷2-羧酸或三羟苯酮酸及其衍生物,例如槲酸(见:pct/EP 2004/011715,其全部内容与参考文献相结合),但并局限于此。重亚硫酸盐转化优选是在30℃-70℃的反应温度下被进行,在该反应过程中温度在短时间内提高到85℃以上(参照PCT/EP2004/011715,其全部内容被纳入此处作为参考)。重亚硫酸盐处理后的DNA优选是在定量之前被纯化。其可通过本领域已知的任何手段被执行,类似超滤法,例如通过Microcon™(由Millipore™制造)进行,但并不局限于此。净化是根据经修改的制造商协议被进行(例如参照PCT/EP2004/011715,其全部内容被纳入此处作为参考)。

[0188] 在一些实施例中,使用根据本发明的一组引物寡核苷酸(例如参照表3和/或表5)和扩增酶来扩增经处理的DNA片段。一些DNA片段的扩增可以同时在一个和同一反应容器中进行。通常,扩增是使用聚合酶链反应(PCR)进行的。扩增的长度一般为100-2000个碱基对。

[0189] 在该方法的另一个实施例中,可以使用甲基化特异性引物寡核苷酸来检测含有DMR(例如表1、2、4和7所规定的DMR 1-274)的标记内或附近CpG体位的甲基化状态。这种技术(MSP)已经在赫尔曼的美国专利No.6,265,171中被描述。使用甲基化状态特异性引物来扩增重亚硫酸盐处理的DNA,可以区分甲基化和非甲基化核酸。MSP引物包含至少一个引物,与重亚硫酸盐处理的CpG二核苷酸杂交。因此,所述引物的序列包括至少一个CpG二核苷酸。针对非甲基化DNA的MSP引物在CpG中C体位的位置处的“T”。

[0190] 通过扩增获得的片段可以携带直接或间接检测的标签。在一些实施例中,标签是荧光标签、放射性核素或可分离的分子碎片,具有在质谱仪中可检测到的典型质量。在所述标签为质量标签的情况下,一些实施例提供了标签扩增子具有单一的正或负净电荷,允许在质谱仪中具有更好的选择性。可通过基体辅助激光解吸/电离质谱(MALDI)或电子喷雾质谱(ESI)等手段进行检测和可视化。

[0191] 适合于这些化验技术的DNA分离方法在本领域中为已知技术。具体而言,一些实施例包括核酸分离,其在美国专利申请No.13/470,251(“核酸的分离”)中被说明,其全部内容被纳入此处作为参考。

[0192] 方法

[0193] 在一些实施例中,本技术提供的方法包括以下步骤:

[0194] 1) 将从对象获得的核酸(例如基因组DNA,从类似粪便样本或胃组织或血浆样本的体液中分离)与至少一种试剂或一系列试剂接触,来区分含有DMR(例如DMR 1-274,如表1、2、4和7所示)的至少一个标记内甲基化和非甲基化的CpG二核苷酸,以及

[0195] 2) 检测肿瘤或增殖性疾病(例如,敏感性大于或等于80%,特异性大于或等于80%)。

[0196] 在一些实施例中,本技术提供了包括以下步骤的方法:

[0197] 1) 将从对象获得的核酸(例如基因组DNA,从类似粪便样本或胃组织的体液中分离)与至少一种试剂或一系列试剂接触,来区分至少一个标记内甲基化和非甲基化的CpG二核苷酸,其中标记从具有标注的染色体区域中被选出,所述标注从ARHGEF4,ELM01,ABCB1,CLEC11A,ST8SIA1,SFMBT2,CD1D,CYP26C1,ZNF569和C13ORF18构成的组中被选出,以及

[0198] 2) 检测胃癌(例如敏感性大于或等于80%,特异性大于或等于80%)。

[0199] 在一些实施例中,本技术提供的方法包括以下步骤:

[0200] 1) 将从对象获得的核酸(例如基因组DNA,从类似粪便样本或胃组织的体液中分离)

离)与至少一种试剂或一系列试剂接触,来区分至少一个标记内甲基化和非甲基化的CpG二核苷酸,其中标记从具有标注的染色体区域中被选出,所述标注从ELM01, ARHGEF4, EMX1, SP9, CLEC11A, ST8SIA1, BMP3, KCNA3, DMRTA2, KCNK12, CD1D, PRKCB, CYP26C1, ZNF568, ABCB1, ELOVL2, PKIA, SFMBT2 (893), PCBP3, MATK, GRN2D, NDRG4, DLX4, PPP2R5C, FGF14, ZNF132, CHST2 (7890), FLI1, c13orf18, 或ZNF569构成的组中被选出,以及

[0201] 2) 检测胃癌(例如,敏感性大于或等于80%,特异性大于或等于80%)。

[0202] 在一些实施例中,本技术提供的方法包括以下步骤:

[0203] 1) 将从对象获得的核酸(例如基因组DNA,从血浆样本中被分离)与至少一种试剂或一系列试剂接触,来区分至少一个标记内甲基化和非甲基化的CpG二核苷酸,其中标记从具有标注的染色体区域中被选出,所述标注从ELM01, ZNF569, C13orf18, CD1D, ARHGEF4, SFMBT2, PPP2R5C, CYP26C1, PKIA, CLEC11A, LRRC4和ST8SIA1构成的组中被选出,以及

[0204] 2) 检测胃癌(例如,敏感性大于或等于80%,特异性大于或等于80%)。

[0205] 优选是,敏感性约为70-100%,或者约为80-90%左右,或者约为80-85%。优选是,特异性约为70%-100%,或是约为80-90%,或是约为80-85%。

[0206] 基因组DNA可以通过任何方法被分离,包括使用商业销售的试剂盒。简单地说,当感兴趣的DNA经细胞膜被包围时,生物样本必须经酶、化学或机械手段被破坏和溶解。然后DNA溶液可以清除蛋白质和其他污染物,例如,通过蛋白酶K的酶切,然后基因组DNA从溶液中被复原。这可以通过多种方法进行,包括盐析、有机提取或DNA与固相载体的结合。方法的选择受时间、费用和所需DNA的量等因素的影响。含有肿瘤物质或癌前物质的所有临床样本类型适用于本方法,如细胞系,组织切片,活检,石蜡包埋组织,体液,粪便,胃组织,结肠排出液,尿液,血浆,血清。全血、分离的血细胞、从血液中分离出来的细胞及上述组合。

[0207] 本技术并不局限于制备样本和提供核酸测试的方法。例如,在一些实施例中,使用直接的基因捕获,例如美国专利申请No.61/485386中所说明的,或是相关方法将DNA从粪便样本或从血液或从血浆样本中被分离出来。

[0208] 然后通过至少一种试剂或一系列试剂处理基因组DNA样本,区分含有DMR的至少一个标记内的甲基化和非甲基化的CpG二核苷酸(例如DMR1-274,如表1、2、4和7所提供的)。

[0209] 在一些实施例中,试剂将5'体位处未甲基化的胞嘧啶碱基转化为尿嘧啶、胸腺嘧啶,或在杂交行为上与胞嘧啶不同的另一个碱基。然而,在一些实施例中,该试剂可以是甲基化敏感性限制酶。

[0210] 在一些实施例中,基因组DNA样本被以这样的方式处理,即5'体位处未甲基化的胞嘧啶碱基被转化为尿嘧啶、胸腺嘧啶,或在杂交行为上与胞嘧啶不同的另一个碱基。在一些实施例中,该处理是以重亚硫酸盐(亚硫酸氢,二硫酸钠)进行的,然后进行碱水解。

[0211] 然后对经过处理的核酸进行分析,以确定目标基因序列的甲基化状态(至少一个基因、基因组序列或来自含有DMR的标记的核苷酸,例如,从DMR1-274中选出的至少一个DMR,例如,如表1、2、4和7所提供的)。分析方法可以从本领域已知的分析方法中被选出,包括在此所列出的,例如在此所述的QuARTS和MSP。

[0212] 异常甲基化,更具体地说,含有DMR(例如DMR 1-274,如表1、2、4和7所提供的)的超甲基化与胃癌相关,并且在一些实施例中预测肿瘤部位。

[0213] 本技术涉及与胃癌相关的任何样本的分析。例如,在一些实施例中,样本包括从患

者获得的组织和/或生物液体。在一些实施例中,所述样本包括分泌物。在一些实施例中,该样本包括血液、血清、血浆、胃分泌物、胰液、胃肠活检样本、胃肠道活检的微解剖细胞、胃肠腔内的胃肠道细胞脱落和/或从粪便中复原的胃肠道细胞。在一些实施例中,对象是人类。这些样本可能来自上消化道、下消化道,或包括来自上消化道和下消化道的细胞、组织和/或分泌物。样本可包括肝、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。在一些实施例中,所述样本包括细胞液、腹水、尿液、粪便、胰液、内镜检查中获得的液体、血液、粘液或唾液。在一些实施例中,该样本是粪便样本。

[0214] 样本可以通过本领域中已知的任何手段获得,例如对技术人员来说是显而易见的。例如,尿液和粪便样本很容易获得,而血液、腹水、血清或胰液则可以通过针头和注射器获得。可通过将样本作为对象利用本领域技术人员已知的各种技术来获得无细胞或基本上没有细胞的样本,这些技术包括离心和过滤,但并不局限于此。虽然一般认为不使用侵入性技术来获取样本,但仍适合获取组织匀浆、组织切片和活检样本等样本。

[0215] 在一些实施例中,该技术涉及一种用于治疗患者(例如,患有胃部癌(例如胃癌)、早期胃癌或可能是胃癌的患者)的方法,该方法包括确定在此提供的一个或多个DMR的甲基化状态,并基于确定甲基化状态的结果对患者进行治疗。治疗可以是实施药化合物、疫苗、进行手术、对患者成像、进行其他测试。优选是,用于临床筛查的方法、预后评估方法、监测治疗结果的方法、识别最有可能对特定治疗做出反应的患者的方法、对患者或对象进行成像的方法、以及药物筛查和开发的方法。

[0216] 在该技术的一些实施例中,提供了一种用于诊断对象中的胃部癌(例如胃癌)的方法。在此使用的术语“作出诊断”和“诊断”是指本领域技术人员可以估计甚至确定对象是否患有某一特定疾病或状况或将来是否可能发展成某一特定疾病或状况的方法。本领域技术人员通常根据一个或多个诊断指标,例如生物标记(例如,在此所述的DMR)进行诊断,甲基化状态指示出病情的存在、严重性或不存在。

[0217] 随着诊断,临床癌症的预后关系到确定癌症的侵袭性和肿瘤复发的可能性,以规划最有效的治疗。如果能够做出更准确的预后,甚至可以评估患癌症的潜在风险,则可以选择适当的治疗方法,在一些情况下,可以为患者选择不太严重的治疗方法。癌症生物标记的甲基化状态评估(例如,确定甲基化状态)有助于将预后良好和/或不需要治疗或小部分治疗的低风险癌症患者与那些更有可能发展为癌症或需要更强化治疗的人中区分开来。

[0218] 因此,此处使用的“作出诊断”或“诊断”进一步包括确定患癌症的风险或确定预后,其可根据在此所述的诊断生物标记(例如DMR),提供临床结果预测(不论是否医疗),选择适当的治疗(或治疗是否有效),或监测当前的治疗并潜在地改变治疗。此外,在此公开的内容的一些实施例中,可以对生物标记进行多个确定,以便于诊断和/或预后。生物标记的时变可以用来预测临床结果,监测胃癌的进展和/或监测针对癌症的适当治疗的效果。例如,在一个实施例中,可以在有效治疗过程中观察到随着时间的推移在此所述的生物样本中一个或多个生物标记(例如,DMR)的甲基化状态发生变化(当被监测,可看到一个或多个附加生物标记)。

[0219] 在此公开的内容在一些实施例中还提供了用于确定是否在对象中发起或继续预防或治疗癌症的方法。在一些实施例中,所述方法包括:提供来自对象的一段时间内的一系

列生物样本;分析所述一系列生物样本,以确定各生物样本中的至少一个生物标记的甲基化状态;并比较各生物样本中的一个或多个生物标记的甲基化状态中的任何可测量的变化。生物标记甲基化状态的任何变化都可以用来预测癌症的发生风险,预测临床结果,决定是否开始或继续预防或治疗癌症,以及目前的治疗是否有效。例如,可以在治疗的开始之前选择第一时间点,并且可以在治疗开始后的某一时点选择第二时间点。甲基化状态可以从不同时间点采集的每一个样本中被测量,并观察质量和/或数量上的差异。不同样本中的生物标记水平的甲基化状态可以与胃癌的风险、预后、治疗效果和/或患者的癌症进展有关。

[0220] 在优选的实施例中,本发明的方法和组合物用于早期阶段的疾病治疗或诊断,例如,在疾病症状出现之前。在一些实施例中,本发明的方法和组合物用于在临床阶段的疾病治疗或诊断。

[0221] 如上所述,在一些实施例中,可以对一个或多个诊断或预后生物标记进行多次测定,并且可以利用标记中的时变来确定诊断或预后。例如,可以在初始时间确定并在第二时间再次确定诊断标记。在该实施例中,从初始时间到第二时间的标记的增加可以是诊断出特定类型或癌症的严重程度,或者给定的预后。同样,从初始时间到第二时间的标记的减少可以指示特定类型或癌症的严重程度,或给定的预后。此外,一个或多个标记的变化程度可能与癌症的严重程度和未来的不良事件有关。本领域的技术人员应理解,在一些实施例中,可以在多个时间点对同一生物标记进行比较测量,还可以在一个时间点测量给定的生物标记,在第二时间测量第二生物标记。对这些标记进行比较从而可以提供诊断信息。

[0222] 在此所使用的术语“确定预后”是指本领域的技术人员可以预测对象中某一状态的过程或结果的方法。“预后”并不是指以100%的准确度预测某一状态的进程或结果的能力,甚至给定进程或结果或多或少是基于生物标记(例如DMR)的甲基化状态被预测。相反,“预测”是指某一过程或结果发生的增加可能性;也就是说,与没有表现出这种情况的个体相比,某一过程或结果更有可能发生在这一特定情况下。例如,在没有表现出这种情况的个体中(例如,具有一个或多个DMR的正常甲基化状态),特定结果(例如,患胃癌)的机会可能非常低。

[0223] 在一些实施例中,统计分析将预测指标与不利结果的倾向相关联。例如,在一些实施例中,甲基化状态与从没有癌症的患者中获得的正常对照样本中的不同时,表示与对象相比,水平更接近于甲基化状态的受试者更有可能患上癌症。对照样本。由统计意义的水平决定的。此外,甲基化状态从基线(例如“正常”)水平的变化可以反映受试者的预后,甲基化状态的变化程度可能与不良事件的严重程度有关。统计显著性通常是通过比较两个或多个群体被确定并确定置信区间和/或p值。例如参照Dowdy and Wearden,“研究统计”,John Wiley&Sons,New York,1983,其全部内容被纳入此处作为参考。

[0224] 在其他实施例中,可以确定本文公开的预后或诊断生物标记的甲基化状态的阈值变化程度(例如,DMR),并且简单地将生物样本中的双体船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。本文提供的用于生物标记的甲基化状态的优选阈值变化为约5%、约10%、约15%、约20%、约25%、约30%、约50%、约75%、约100%和约150%。在另一些实施例中,可以建立“列线图”,通过该“列线图”,预后或诊断指标(生物标记或生物标记的组合)的甲基化状态直接涉及正向于给定结果的相关配置。本领域技术人员能够利用该列线图来关联两个数值,但应理解,由于参考个别样本测量而不是群体平均值,该测量

中的不确定性与标记浓度的不确定性相同。

[0225] 在一些实施例中,与生物样本同时分析对照样本,从而可以将从生物样本获得的结果与从对照样本获得的结果进行比较。另外,可以设想提供标准曲线,比较生物样本的化验结果。该标准曲线显示出生物标记的甲基化状态来作为化验个体的功能,例如当使用荧光标记时可以是荧光信号强度。使用来自多个供体的样本,可以提供标准曲线以控制正常组织中的一个或多个生物标记的甲基化状态,以及来自具有化生的供体或来自患有胃癌的供体的组织中的一个或多个生物标记的“风险”水平。在该方法的一些实施例中,当从对象获得的生物样本中提供的一个或多个DMR被鉴定为异常甲基化状态时,则该对象被鉴定为具有组织变形。在该方法的其他实施例中,当从对象获得的生物样本中的一个或多个生物标记被检测为异常甲基化状态时,则该对象被鉴定为具有癌症。

[0226] 标记的分析可以单独进行,也可以与一个测试样本中的附加标记同时进行。例如,几个标记可以组合成一个测试,以便有效地处理多个样本,并可提供更精确的诊断和/或预测准确性。此外,本领域的技术人员应认识到从同一对象测试多个样本(例如,在连续的时间点)的价值。这种序列样本测试可以识别出随时间标记甲基化状态的变化。在此甲基化状态的变化以及不存在甲基化状态的变化,可以提供关于疾病状态的有用信息,包括从事件开始的大约时间,可挽救组织的存在和数量。药物疗法的适当性,各种疗法的有效性,以及确定研究对象的结果,包括未来事件的风险,但并不局限于此。

[0227] 生物标记的分析可以以多种物理形式进行。例如,使用微滴度板或自动化可以方便大量测试样本的处理。或者,可以制定单一的样本格式,以便在流动运输或急诊室环境中及时进行治疗和诊断。

[0228] 在一些实施例中,当对象与对照甲基化状态相比,样本中至少一个生物标记的甲基化状态存在可测量的差异时,则被诊断为患有胃癌。相反,当在生物样本中没有发现甲基化状态的变化,则可以确定对象没有患胃癌、没有患癌的风险或患癌症的风险较低。在这方面,具有癌症或其风险的对象可以与低到基本没有癌症或其风险的对象相区分。那些具有患胃癌风险的对象可以被安排在更密集和/或定期的筛查计划,包括内镜监视。另一方面,那些风险较低或基本没有风险的对象可以避免接受内镜检查,直到将来进行筛查,例如根据本技术执行的筛查,指示出这些对象具有患胃癌的危险。

[0229] 如上所述,根据本技术的实施例,检测一个或多个生物标记甲基化状态的变化可以是定性测定,也可以是定量测定。因此,诊断对象具有胃部癌(例如胃癌)或有可能发展胃癌的步骤示出阈值测量被进行,例如,生物样本中一个或多个生物标记的甲基化状态与预定的对照甲基化状态不同。在该方法的一些实施例中,对照甲基化状态是任何可检测的生物标记的甲基化状态。在该方法的其他实施例中,对照样本与生物样本同时测试,预定的甲基化状态是对照样本中的甲基化状态。在该方法的其他实施例中,预定的甲基化状态基于和/或经标准曲线被鉴定。在该方法的其他实施例中,预定的甲基化状态为特定状态或状态范围。因此,可以在本领域技术人员明显可接受的限度内,部分基于被执行的方法的实施例和所需的特异性等,来选择预定的甲基化状态。

[0230] 进一步关于诊断方法,优选的对象是脊椎动物。优选的脊椎动物是温血动物;优选的温血动物是哺乳动物。优选的哺乳动物最优选是人。如在此使用的,术语“对象”包括人类和动物对象。因此,本文提供兽医治疗用途。因此,本技术提供了诸如人的哺乳动物的诊断,

以及那些诸如西伯利亚虎等濒危物种且具有重要意义的哺乳动物;具有经济重要性的哺乳动物,例如饲养在农场上以供人类食用的动物;和/或对人具有社会重要性的动物,例如作为宠物或动物园饲养的动物。这些动物的示例包括:食肉动物,如猫和狗;猪,如猪、肉猪和野猪;反刍动物和/或蹄类动物,如公牛、牛、绵羊、长颈鹿、鹿、山羊、野牛和骆驼;以及马,但不局限于此。因此,还提供了家畜的诊断和治疗,包括家养的猪、反刍动物、蹄类动物、马(包括竞赛马)等,但并不局限于此。本公开的技术还包括用于诊断对象中的胃部癌(例如胃癌)的系统。例如该系统可以作为商业试剂盒被提供,其可用来筛查胃癌的风险或诊断所收集的生物样本中的对象是否具有胃癌。根据本技术提供的示例性系统包括评估表1、2、4和7中提供的DMR的甲基化状态。

[0231] 示例

[0232] 示例1-用于检测胃癌的识别标记

[0233] 在用于提高本技术的实施例过程期间所执行的实验中针对胃癌鉴定了与248个DMR对应的123个DNA甲基化标记(例如DMR Nos.1-248;表1)。该DNA甲基化标记是由CpG胰岛富集与病例对照组织样本集的大规模并行测序相结合所产生的数据中被鉴定的。对照组包括正常胃组织、正常结肠上皮和源于DNA的正常白细胞。该技术采用了简化基因组重亚硫酸盐测序(RRBS)。第三步骤涉及依据覆盖切断来解析数据,不包括所有非信息部位,使用逻辑回归对比诊断亚组之间的甲基化百分比,基于定义的连续甲基化部位的组,建立生物信息CpG“胰岛”,以及受试者操作特征ROC分析。第四步骤过滤甲基化数据%(包括个别CPGs和生物信息簇)来选择标记,最大限度地提高信噪比,使基质甲基化最小化,对肿瘤异质性作出解释,并强调ROC的性能。此外,这一分析方法还确保了热点CPGs的鉴定,以方便和优化下游标记化验的设计-甲基化特异性PCR、小片段深度测序等。

[0234] 各样本产生约2-3百万个高质量的CpGs,经过分析和过滤,得到少于1000个的高度区分的个别部位。其被聚集成123个差异甲基化的局部区域,一些扩展到30-40个碱基,且一些大于一千碱基(参照表1)。12个DMRs没有标注,在文献中也没有任何参照。这些DMRs被命名为MAX,随后是染色体位置和坐标(参照表1)。所有DMRs具有0.8或以上的AUC值(病例与对照组的一些证实的精确分辨具有1.0的AUC)。除达到0.8的AUC阈值外,已鉴定的胃癌标记出肿瘤中的甲基化密度比正常胃或结肠粘膜高20倍以上,且非肿瘤性胃或结肠粘膜甲基化程度则<1.0%。

[0235] 表1.

[0236]

DMR No.	基因	染色体	胰岛范围	链
1	ABCB1	7	87229775-87229856	-
2	ADAM19	5	157002684-157002766	-
3	ADAM23	2	207307964-207308190	+
4	ADAM23	2	207308771-207308890	+
5	ADCY1	7	45613877-45613977	+
6	ADRA1B	5	159343458-159343934	+
7	AFF3	2	100721738-100721967	-
8	AGBL4	1	49242089-49242514	-
9	ALOX5	10	45914474-45914484	+
10	ANK1	8	41754327-41754528	-
11	ANKRD34B	5	79866019-79866185	-
12	ANO8	19	17439445-17439845	-
13	APOBEC3D	22	39417451-39417576	+
14	ARHGEF7	13	111767862-111767954	+
15	BACH2	6	91004976-91005235	-

[0237]

16	BARX1	9	96715112-96715262	-
17	BARX1	9	96715374-96715603	-
18	BARX1	9	96721501-96721818	-
19	BCL2	18	60985561-60985600	-
20	BCL2	18	60985788-60985828	-
21	BHLHE22	8	65493937-65494105	+
22	BMP3	4	81952135-81952303	+
23	BMP3	4	81952407-81952528	+
24	BMP6	6	7727566-7727907	+
25	BMP8A	1	39980533-39980619	+
26	C12orf39	12	21680721-21680828	+
27	C12orf42	12	103889269-103889591	-
28	C17orf46	17	43339242-43339498	-
29	C1orf70	1	1475560-1475650	-
30	C1orf70	1	1476065-1476127	-
31	C1orf95	1	226737123-226737246	+
32	C1QL3	10	16563608-16563892	-
33	C9orf50	9	132382706-132382909	-
34	CABIN1	22	24552655-24552724	+
35	CACNA1C	12	2162358-2162804	+
36	CACNA1H	16	1202428-1202560	+
37	CBFA2T3	16	89007265-89007441	-
38	CBLN1	16	49315385-49315725	-
39	CBLN1	16	49316162-49316268	-
40	CCDC106	19	56159280-56159589	+
41	CCDC42B	12	113592232-113592553	+
42	CCND2	12	4378158-4378413	+
43	CD1D	1	158150624-158151205	+
44	CDHR1	10	85954203-85954476	+
45	CELF2	10	11059947-11060085	+
46	CHST2	3	142838645-142839023	+
47	CHST2	3	142839223-142839495	+
48	CHST2	3	142839761-142840276	+
49	CLEC11A	19	51228264-51228394	+
50	CLEC11A	19	51228512-51228732	+
51	CLSTN2	3	139654045-139654299	+
52	CNR1	6	88876414-88876452	-
53	COMP	19	18901578-18901783	-
54	CPE	4	166300108-166300239	+
55	CPNE8	12	39299231-39299461	-
56	DAB2IP	9	124461296-124461420	+
57	DBX1	11	20177747-20178529	-
58	DOCK10	2	225906664-225906922	-
59	DOCK10	2	225907226-225907322	-
60	DOCK10	2	225907515-225907632	-
61	DOK1	2	74782380-74782580	+
62	DYNLRB2	16	80574542-80574687	+
63	DYNLRB2	16	80574797-80574927	+
64	EFNB3	17	7608691-7608815	+
65	EIF4E3	3	71803220-71804279	-
66	ELMO1	7	37487867-37488498	-

[0238]

67	ELOVL2	6	11044395-11044704	-
68	ELOVL5	6	53213597-53214212	-
69	EPB41L3	18	5542876-5542973	-
70	EYA4	6	133562791-133562878	+
71	EYA4	6	133563343-133563397	+
72	FAIM2	12	50297863-50297988	-
73	FAM78A	9	134152466-134152692	-
74	FBN2	5	127874257-127874290	-
75	FER1L4	20	34189084-34189185	-
76	FER1L4	20	34189488-34189566	-
77	FGF2	4	123748398-123748881	+
78	FLT3	13	28674537-28674770	-
79	FMN2	1	240255171-240255264	+
80	FMNL1	17	43298726-43298819	+
81	FOXD4L1	2	114256519-114257141	+
82	FRMD4A	10	13933638-13934256	-
83	GJC1	17	42907705-42907827	-
84	GLB1L3	11	134146132-134146380	+
85	GNG7	19	2702502-2702568	-
86	GPR135	14	59930725-59931142	-
87	GPX7	1	53067933-53068182	+
88	GRASP	12	52400919-52401166	+
89	GRM8	7	126891909-126892517	-
90	GSG1L	16	28074472-28074761	-
91	GSG1L	16	28075145-28075290	-
92	GUCY1A2	11	106888403-106888538	-
93	HECW2	2	197457425-197458011	-
94	HOPX	4	57521377-57521658	-
95	HOPX	4	57521822-57522182	-
96	HOXD8	2	176994098-176994972	+
97	IKZF1	7	50343838-50344453	+
98	INSM1	20	20348122-20348415	+
99	IRF4	6	391980-392182	+
100	ITGA4	2	182321830-182321983	+
101	KCNK12	2	47797187-47797452	-
102	KCNMA1	10	79397844-79397945	-
103	KCNQ3	8	133492336-133493033	-
104	KDM2B	12	122016688-122017150	-
105	KIF5C	2	149633046-149633131	+
106	KLHL29	2	23609989-23610119	+
107	LBH	2	30455594-30455655	+
108	LINGO3	19	2290273-2290393	-
109	LIPJ	10	90342642-90342873	+
110	LOC285548	4	13549245-13549556	-
111	LOC348840	3	197807208-197807313	-
112	LOC440461	17	66195303-66195773	+
113	LOC645323	5	87970260-87970894	-
114	LRRC10B	11	61276493-61276910	+
115	LRRFIP1	2	238600049-238600078	+
116	LY86-AS1	6	6546450-6546598	-
117	MARK1	1	220702225-220702345	+

[0239]

118	MATK	19	3785828-3786371	-
119	MAX.chr1.110626388-110626787	1	110626388-110626787	-
120	MAX.chr1.244012787-244012875	1	244012787-244012875	-
121	MAX.chr10.22624234-22624571	10	22624234-22624571	-
122	MAX.chr11.120435350-120435981	11	120435350-120435981	-
123	MAX.chr11.123301058-123301255	11	123301058-123301255	-
124	MAX.chr11.123301366-123301506	11	123301366-123301506	-
125	MAX.chr12.133484966-133485857	12	133484966-133485857	-
126	MAX.chr12.4273880-4274012	12	4273880-4274012	-
127	MAX.chr14.37116653-37117410	14	37116653-37117410	-
128	MAX.chr15.40616377-40616419	15	40616377-40616419	-
129	MAX.chr15.89942727-89942918	15	89942727-89942918	-
130	MAX.chr18.5630121-5630725	18	5630121-5630725	-
131	MAX.chr19.22034799-22034887	19	22034799-22034887	-
132	MAX.chr2.118981853-118981946	2	118981853-118981946	-
133	MAX.chr2.119067940-119068174	2	119067940-119068174	-
134	MAX.chr2.177503038-177503122	2	177503038-177503122	-
135	MAX.chr20.1784209-1784461	20	1784209-1784461	-
136	MAX.chr3.128336893-128337084	3	128336893-128337084	-
137	MAX.chr4.113626838-113626942	4	113626838-113626942	-
138	MAX.chr4.113627402-113627507	4	113627402-113627507	-
139	MAX.chr5.42951300-42951769	5	42951300-42951769	-
140	MAX.chr5.42992128-42992212	5	42992128-42992212	-
141	MAX.chr5.42992322-42992768	5	42992322-42992768	-
142	MAX.chr5.42995102-42995181	5	42995102-42995181	-
143	MAX.chr7.121950350-121951056	7	121950350-121951056	-
144	MAX.chr7.25896389-25896501	7	25896389-25896501	-
145	MAX.chr9.113945-114435	9	113945-114435	-
146	MAX.chr9.79627082-79627175	9	79627082-79627175	-
147	MAX.chr9.99983730-99984118	9	99983730-99984118	-
148	MEF2C	5	88185490-88185576	-
149	MIR129-2	11	43601090-43601396	+
150	MMP23B	1	1567269-1567478	+
151	MPZ	1	161275561-161275653	-
152	MPZ	1	161275766-161275996	-
153	MZB1	5	138728509-138729109	-
154	NDRG4	16	58497251-58497370	+
155	NDRG4	16	58497979-58498250	+
156	NEU1	6	31830234-31830331	-
157	NEU1	6	31830456-31830522	-
158	NEU1	6	31830682-31830917	-
159	NEUROG2	4	113437639-113437953	-

[0240]

160	NFATC4	14	24837540-24838009	+
161	NID2	14	52535758-52536437	-
162	NRXN2	11	64480878-64481067	-
163	OPLAH	8	145106742-145106921	-
164	OSR2	8	99952116-99952427	+
165	OSR2	8	99952591-99953023	+
166	PAX5	9	37027013-37027670	-
167	PDE8B	5	76506245-76506460	+
168	PEBP4	8	22722910-22722994	-
169	PHF21B	22	45405722-45405819	-
170	PLEKHO1	1	150122907-150123157	+
171	PLXNC1	12	94543138-94543232	+
172	PNMAL2	19	46996516-46996841	-
173	POMC	2	25391168-25391616	-
174	POU3F1	1	38510915-38511213	-
175	POU4F1	13	79176470-79176567	-
176	POU4F1	13	79176954-79177042	-
177	POU4F1	13	79177364-79177872	-
178	PPP2R5C	14	102247537-102247929	+
179	PRDM13	6	100051231-100051668	+
180	PRKCB	16	23847288-23847339	+
181	PRKCB	16	23847575-23847699	+
182	PRKCB	16	23847825-23848020	+
183	PRR16	5	119799921-119800390	+
184	PRR5L	11	36399221-36399473	+
185	PYGL	14	51410872-51411219	-
186	RASA3	13	114897059-114897135	-
187	RASSF2	20	4803267-4803703	-
188	RBPM52	15	65067593-65067619	-
189	RNF165	18	43913775-43914353	+
190	RPRM	2	154334604-154335314	-
191	RPSAP58	19	23941384-23941670	+
192	RSPO3	6	127440413-127441057	+
193	RUNX3	1	25256650-25257322	-
194	SEPTIN9	17	75368800-75369056	+
195	SEPTIN9	17	75369224-75369327	+
196	SFMBT2	10	7450406-7450831	-
197	SFMBT2	10	7451097-7451185	-
198	SFMBT2	10	7452358-7452452	-
199	SFMBT2	10	7452746-7452779	-
200	SFMBT2	10	7453455-7453511	-
201	SLC35F1	6	118228394-118228979	+
202	SLCO4C1	5	101631546-101631741	-
203	SLCO4C1	5	101632152-101632237	-
204	SND1	7	127671993-127672310	+
205	SNX32	11	65601248-65601514	+
206	SP9	2	175201823-175202389	+
207	ST8SIA4	5	100236601-100236784	-
208	ST8SIA6	10	17496306-17496417	-
209	ST8SIA6	10	17496572-17496711	-
210	TBC1D30	12	65218335-65218778	+

[0241]

211	TBC1D30	12	65220152-65220299	+
212	TCF24	8	67874670-67875083	-
213	TIAM1	21	32930248-32930322	-
214	TLX2	2	74742033-74742264	+
215	TM6SF1	15	83776207-83776706	+
216	TRIM9	14	51560661-51561240	-
217	TRPC3	4	122872067-122872455	-
218	TRPC3	4	122872626-122872724	-
219	TSC22D4	7	100075304-100075445	-
220	TTC28	22	29075259-29075493	-
221	UNC5CL	6	40995803-40996093	-
222	VCAN	5	82768837-82769031	+
223	VEGFC	4	177713711-177713910	-
224	WNT5A	3	55522631-55522773	-
225	WT1	11	32457189-32457408	-
226	ZBTB16	11	113930478-113930929	+
227	ZDBF2	2	207139497-207139592	+
228	ZNF132	19	58951402-58951530	-
229	ZNF134	19	58125597-58125764	+
230	ZNF256	19	58459002-58459219	-
231	ZNF331	19	54024023-54024436	+
232	ZNF347	19	53662191-53662618	-
233	ZNF354C	5	178487123-178487607	+
234	ZNF43	19	22018452-22018639	-
235	ZNF43	19	22018852-22019004	-
236	ZNF461	19	37157822-37158261	-
237	ZNF569	19	37957760-37958046	-
238	ZNF570	19	37960066-37960542	+
239	ZNF610	19	52839503-52840013	+
240	ZNF625	19	12267378-12267677	-
241	ZNF665	19	53696101-53696195	-
242	ZNF665	19	53696497-53696668	-
243	ZNF682	19	20149796-20149923	-
244	ZNF71	19	57106756-57106967	+
245	ZNF788	19	12203466-12203641	+
246	ZNF813	19	53970869-53971093	+
247	ZNF813	19	53971239-53971374	+
248	ZNF83	19	53193211-53193303	-

[0242] 在123个说明的标记中,有22个是根据AUC和倍数变化率(癌症vs.正常胃+正常结肠)被选择的,并与其他GI癌部位(结肠、胰腺、胆管、食道)的73个发现标记一起进行更大的组织验证研究。在这些范围为15-50样本的不同组中,包括癌症,正常群,以及适用的腺瘤。验证平台为定量甲基化特异性PCR,以及在甲基化对照集上被优化的引物。所有化验均在Roche 480LightCyclers上进行。结果分析了ROC特征、倍数变化和互补性。以敏感性(包括癌症和腺瘤)排列的前30项在表2中被示出。表3提供了表2中所示出的用于标记的引物信息。其中10项是在共同的Mayo韩国研究(参照示例2)中被选择用于进一步的组织验证。

[0243] 表2

[0244]

DMR No.	基因	敏感性 (100%特异性)	染色体	胰岛范围
253	ELMO1	95%	7	37487755-37488477
251	ARHGEF4	91%	2	131797843- 131797938
254	EMX1	88%	2	73147710-73147772
255	SP9	86%	2	175202025- 175202158
256	CLEC11A	86%	19	51228217-51228732
249	ST8SIA1	86%	12	22487528-22487620
257	BMP3	86%	4	81952348-81952402
258	KCNA3	84%	1	111217654- 111217816
259	DMRTA2	81%	1	50884349-50884499
260	KCNK12	81%	2	47797260-47797371
261	CD1D	79%	1	158150797- 158151205
262	PRKCB	79%	16	23847825-23848025
250	CYP26C1	79%	10	94822416-94822607
263	ZNF568	79%	19	37407197-37407365
1	ABCB1	79%	7	87229775-87229856
264	ELOVL2	77%	6	11044395-11044834
265	PKIA	77%	8	79428485-79428684
196	SFMBT2 (893)	74%	10	7450406-7450831
266	PCBP3	72%	21	47063793-47064177
118	MATK	72%	19	3785828-3786371
267	GRIN2D	70%	19	48918144-48918350
268	NDRG4	70%	16	58497395-58497451
269	DLX4	70%	17	48042426-48042820
270	PPP2R5C	67%	14	102247525- 102247929
271	FGF14	67%	13	103046885- 103047269
272	ZNF132	67%	19	58951402-58951775
46	CHST2 (7890)	65%	3	142838645- 142839023
273	FLI1	63%	11	128563956- 128564209

[0245]

252	c13orf18	63%	13	46960767-46961669
237	ZNF569	70% (92% spec)	19	37957760- 37958046

[0246] 表3.

[0247]

DMR No.	基因	正向 MSP 引物(5'-3')	反向 MSP 引物(5'-3')
253	ELMO1	TTA TAT TTT TCG TTT TTA GTA ATT TCG CGT TAG C (SEQ ID NO: 21)	GAA AAC CCG CCG AAA CAT TTC GA (SEQ ID NO: 22)
251	ARHGEF4	TGT TTT CGC GGT CGT TAT ATA TTA CGT CGT (SEQ ID NO: 23)	GAA CTA TCC CCG AAC TCC GAC TCG A (SEQ ID NO: 24)
254	EMX1	CGG GTT TTA GCG ATG TTT ATT TTA GTT TCG T (SEQ ID NO: 25)	CCT TTT CGT TCG TAT AAA ATT TCG TT (SEQ ID NO: 26)
255	SP9	CGG GGG TAA AAA GGG TAG CGA TAG TGA TAC (SEQ ID NO: 27)	CGA AAA ATT CGA AAC GAA ACG TC (SEQ ID NO: 28)
256	CLEC11A	AGT TTG GCG TAG TCG GTA GAT C (SEQ ID NO: 29)	GCG CGC AAA TAC CGA ATA AAC G (SEQ ID NO: 30)
249	ST8SIA1	GAC GTT TGT CGT CGG GTT CGT TC (SEQ ID NO: 13)	AAA AAC CCT CCG CTA CCA CTT CGC (SEQ ID NO: 14)
257	BMP3	GTTTAATTTTCGGTTTCGTCGTC (SEQ ID NO: 31)	CGCTACGAAACACTCCGA (SEQ ID NO: 32)
258	KCNA3	TTT TGT TGG GTG GGT ATT TTC GCG (SEQ ID NO: 33)	ACA ACC CTC AAA ACC CCT CGA TAT C (SEQ ID NO: 34)
259	DMRTA2	TTT ATT TAT GGA TTA CGT TTT TAG CGA (SEQ ID NO: 35)	GAC GAC CGA ACG ATC ACG CA (SEQ ID NO: 36)
260	KCNK12	GCG TCG TTA GTA GTA CGA AGC (SEQ ID NO: 37)	GCA CCT CAA CGA AAA CAC CGA (SEQ ID NO: 38)
261	CD1D	GCG CGT AGC GGC GTT TC (SEQ ID NO: 39)	CCC ATA TCG CCC GAC GTA A (SEQ ID NO: 40)
262	PRKCB	GTC GTT TTT GGG CGT TTT AGA GGC (SEQ ID NO: 41)	TAA TCC CAA ACG AAC CGC CG (SEQ ID NO: 42)
250	CYP26C1	GGT TTT TTG GTT ATT TCG GAA TCG T (SEQ ID NO: 43)	TAT AAA AAC GCG CGT AAT CAA CGC T (SEQ ID NO: 44)
263	ZNF568	TTG AGA TGT TGG GTG AAG GCG ATT C (SEQ ID NO: 45)	CGC TAA CGC GAA AAA ATA ATT CGA CG (SEQ ID NO: 46)
1	ABCB1	GAT TTT GTT CGT CGT TAG TGC (SEQ ID NO: 9)	TCT CTA AAC CCG CGA ACG A (SEQ ID NO: 10)
264	ELOVL2	CGG TTT TAT TTA TTA TGA TTC GTA GCG G (SEQ ID NO:	CGA CTA CCC TAA ACA ACG CAT CGC (SEQ ID

[0248]

		47)	NO: 48)
265	PKIA	CGG GGA TGA TTT TAT GTA GTC GGA GTT TCG C (SEQ ID NO: 49)	CCC GCC GAA TAC TCG ATC AAC TCG (SEQ ID NO: 50)
196	SFMBT2 (893)	AAA ACG TTT TTA GGT ATT TGG TCG T (SEQ ID NO: 5)	ATA AAA AAA CCT CCT CCG AAC CCG C (SEQ ID NO: 6)
266	PCBP3	GGT CGC GTC GTT TTC GAT C (SEQ ID NO: 51)	GCC GCA AAC GCC GAC GA (SEQ ID NO: 52)
118	MATK	TGT ATA TTT CGA GGC GGT TTC GG (SEQ ID NO: 53)	CGC CCC CAA AAT AAA AAA ACG AA (SEQ ID NO: 54)
267	GRIN2D	TTT TTC GTT CGT TTT TTT ATT TTT TCG A (SEQ ID NO: 55)	AAC GCA TAC CAT CGA CTT CAA CTA CGA C (SEQ ID NO: 56)
268	NDRG4	CGGTTTTTCGTTCTGTTTTTCG (SEQ ID NO: 57)	CCGCCTTCTACGCGACTA (SEQ ID NO: 58)
269	DLX4	GTA TTT TTA ATA TTT GGT GAG TGC G (SEQ ID NO: 59)	TAC TAA AAC GTA CGA TAA ACA TCG T (SEQ ID NO: 60)
270	PPP2R5C	TCG ATT TTA TTT TTG TTG TCG TTG TAG ATT CGC (SEQ ID NO: 61)	GAA AAA ACT AAA AAA CGA CAA AAA AAC CCG ACG (SEQ ID NO: 62)
271	FGF14	GGA GGG AGT TTA ACG GGT TAG GTA C (SEQ ID NO: 63)	CGT AAA CAA AAA ACA ACG ACG ACG (SEQ ID NO: 64)
272	ZNF132	ATT GCG TCG TTG TTT AGG TAA CGT A (SEQ ID NO: 65)	CGA AAA TAC CTA TCT CCT CGA CT (SEQ ID NO: 66)
46	CHST2 (7890)	GGA ACG AGT GAT AGT CGG ATA GTT CGT C (SEQ ID NO: 67)	CGC CCG AAA ACG ACC CCG (SEQ ID NO: 68)
273	FLI1	GGG AGT GAG GGT AGG GCG TTC (SEQ ID NO: 69)	CTC GCA ACC CCT TCG AAT TAA CCC G (SEQ ID NO: 70)
252	c13orf18	TTT AGG GAA GTA AAG CGT CGT TTT C (SEQ ID NO: 19)	AAC GAC GTC TCG ATA CCT ACG A (SEQ ID NO: 20)
237	ZNF569	TGT GGA ATC GGG GTT TGT GTT CGC (SEQ ID NO: 11)	CCC ACC CAA CAC AAA AAA TCC GAC G (SEQ ID NO: 12)

[0249] 示例2-经新甲基化DNA标记的胃癌检测:来自美国和韩国的患者群体的组织验证

[0250] 在用于提高本技术的实施例过程期间所执行的实验中经整个甲基化测序鉴定了新甲基化DNA标记,然后从美国和韩国患者群体的组织中验证了最佳候选(参照表4、5和6)。

[0251] 在采用简化基因组重亚硫酸盐测序的整个甲基化发现工作中,从美国和韩国的患者中选出了17个甲基化DNA标记候选,用于胃组织组的盲验证。微切割组织中的DNA通过甲基化特异性聚合酶链式反应被化验,标记水平经 β -肌动蛋白(总的人类DNA)被标准化。GC病例包括来自美国的35名以及韩国的50名患者,具有病理证实的完整腺癌。对照组包括与病例匹配的人口统计的来自65名健康的美国患者的组织结构上正常的胃上皮细胞,以及来自相同的50名韩国病例患者的胃上皮细胞,但在远离最初GC的部位被获得。对于各标记,使用命名逻辑回归法计算ROC曲线下面积(AUC)。

[0252] 总体而言,前10位标记包括ARHGEF4,ELM01,ABCB1,CLEC11A,ST8SIA1,SFMBT2,CD1D,CYP26C1,ZNF569,和C13ORF18(参照表4)。在95%特异性下,这些标记组在100%的美国群体和94%的韩国群体中检测到GC。美国群体中的个别AUC各自为0.92,0.91,0.89,0.88,0.85,0.82,0.82,0.81,0.75,和0.66,且韩国群体中的个别AUC各自为0.73,0.90,0.78,0.75,0.80,0.94,0.87,0.79,0.75和0.79(参照表6)。有些标记,如ELM01,在各患者群中示出相似的高GC辨别力;其他标记,如ARHGEF4,由于患者群体中(参照图1)正常对照(NL)的高基质而示出更多变的辨别力。表5提供了用于表4的10个标记的正向和反向引物信息。

[0253] 表4

[0254]

DMR No.	基因	染色体	胰岛范围
49	CLEC11A	19	51228264-51228394
43	CD1D	1	158150624-158151205
196	SFMBT2 (893)	10	7450406-7450831
66	ELM01	7	37487867-37488498
1	ABCB1	7	87229775-87229856
237	ZNF569	19	37957760-37958046
249	ST8SIA1	12	22487528-22487620
250	CYP26C1	10	94822416-94822607
251	ARHGEF4	2	131797843-131797938
252	c13orf18	13	46960767-46961669

[0255] 表5

[0256]

DMR No.	基因	正向引物 (5'-3')	反向引物 (3'-5')
49	CLEC11A	AGT TTG GCG TAG TCG GTA GAT C (SEQ ID NO: 1)	GCG CGC AAA TAC CGA ATA AAC G (SEQ ID NO: 2)

[0257]

43	CD1D	GCG CGT AGC GGC GTT TC (SEQ ID NO: 3)	CCC ATA TCG CCC GAC GTA A (SEQ ID NO: 4)
196	SFMBT2 (893)	AAA ACG TTT TTA GGT ATT TGG TCG T (SEQ ID NO: 5)	ATA AAA AAA CCT CCT CCG AAC CCG C (SEQ ID NO: 6)
66	ELMO1	GGA GGG AGT TTA ACG GGT TAG GTA C (SEQ ID NO: 7)	CGT AAA CAA AAA ACA ACG ACG ACG (SEQ ID NO: 8)
1	ABCB1	GAT TTT GTT CGT CGT TAG TGC (SEQ ID NO: 9)	TCT CTA AAC CCG CGA ACG A (SEQ ID NO: 10)
237	ZNF569	TGT GGA ATC GGG GTT TGT GTT CGC (SEQ ID NO: 11)	CCC ACC CAA CAC AAA AAA TCC GAC G (SEQ ID NO: 12)
249	ST8SIA1	GAC GTT TGT CGT CGG GTT CGT TC (SEQ ID NO: 13)	AAA AAC CCT CCG CTA CCA CTT CGC (SEQ ID NO: 14)
250	CYP26C1	GGT TTT TTG GTT ATT TCG GAA TCG T (SEQ ID NO: 15)	TAT AAA AAC GCG CGT AAT CAA CGC T (SEQ ID NO: 16)
251	ARHGEF4	TGT TTT CGC GGT CGT TAT ATA TTA CGT CGT (SEQ ID NO: 17)	GAA CTA TCC CCG AAC TCC GAC TCG A (SEQ ID NO: 18)
252	c13orf18	TTT AGG GAA GTA AAG CGT CGT TTT C (SEQ ID NO: 19)	AAC GAC GTC TCG ATA CCT ACG A (SEQ ID NO: 20)

[0258] 表6

[0259]

DMR No.	基因	美国同生群曲线下的面积	韩国同生群曲线下的面积
49	CLEC11A	0.88	0.75
43	CD1D	0.82	0.87
196	SFMBT2 (893)	0.82	0.94
66	ELMO1	0.91	0.9
1	ABCB1	0.89	0.78
237	ZNF569	0.75	0.75
249	ST8SIA1	0.85	0.8
250	CYP26C1	0.81	0.79
251	ARHGEF4	0.92	0.73
252	c13orf18	0.66	0.79

[0260] 示例3-经对血浆中新甲基化DNA标记化验的胃癌检测

[0261] 本示例探讨了使用从示例I和II中选出的甲基化DNA标记 (MDM) 候选来化验血浆作为胃腺癌 (GC) 检测的方法。事实上, 该示例表明, 血浆中化验的一组新MDM能够准确地区分GC病例和正常对照。其示出在对照组中, 标记水平极小, 但在病例中被升高, 并且随着GC的阶段而逐渐升高。

[0262] 单中心的符合病例输入标准的归档血浆样本被包括在该研究中。病例包括37名患

者,在所有阶段、组织类型和胃部位具有病理证实的完整腺癌。病例血浆样本在切除、化疗或放射治疗前被采集。作为对照,来自38名年龄和性别匹配的健康志愿者的归档血浆样本被选择。DNA从2ml的血浆和重亚硫酸盐处理中被提取。然后,选择的MDMs (ELM01, ZNF569, C13orf18, CD1D, ARHGEF4, SFMBT2, PPP25RC, CYP26C1, PKIA, CLEC11A, LRRC4, 和ST8SIA1) (表7示出用于LRCC 4的DMR信息)通过使用非优化的定量等位基因实时目标和信号扩增方法被化验。这些标记在经临床敏感性和特异性排名的组织中表现出较高的性能、在将胃癌与正常胃组织比较时甲基化信号中的倍数变化,以及源于DNA的血细胞中的基质甲基化程度非常低。

[0263] 为了促进血浆中的最高分析性能,以前在组织中用于甲基化特异性PCR化验的标记被重新设计,使用原始DMR序列作为多重复用QuARTs (定量等位基因特异性实时目标和信号扩增) 化验。QuARTs化验包括PCR引物(参照表8)、检测探针(参照表8)和侵入性寡核苷酸(集成DNA技术)、GoTaq DNA聚合酶(Promega)、裂酶II (Hologic)和含有FAM、HEX和Quasar 670染料的荧光共振能量转移报告试剂盒(FRETs)。

[0264] 图2示出FRET盒的寡核苷酸序列,用于经QuARTs (定量等位基因特异性实时目标和信号扩增) 检测甲基化DNA特征。各FRET序列包括荧光团和猝灭剂,其可以在3个单独的化验中被一起多重复用。

[0265] 标准源于从质粒对照中培养及切割的目标序列。连续稀释标准的绝对拷贝从Poisson模型中获得。

[0266] DNA从2mL归档血浆样本中被纯化,随后重亚硫酸盐通过内部专有方法和自动化仪器被转化。在研究期间,过程对照基因被预先添加到所有样本中,以校正目标复原变异性。

[0267] 样本通过12个引物集被预扩增,稀释后进行QuARTs扩增。下一步的平台是LightCycler 480 (Roche)。用于各标记的链数被归一化为 β -肌动蛋白对照化验的链数,其被包括在多重反应中。

[0268] 针对结果进行逻辑分析,并以矩阵格式绘制,以评估互补性。

[0269] 图3示出受试者工作特征曲线图,突出与单个标记曲线相比较的3个标记组(ELM01, ZNF569和c13orf18)的血浆中的性能。在100%特异性下,该组检测到的胃癌的敏感性为86%。

[0270] 图4A示出血浆中甲基化ELM0 1的对数绝对链计数,阶段性示出从正常粘膜到阶段4的胃癌进展情况。定量标记水平随GC阶段被增加。

[0271] 图4B示出条形图表,按3个标记组(ELM0 1、ZNF569和c13orf18)阶段(100%特异性)显示出血浆中胃癌的敏感性。

[0272] 事实上,ELM0 1是最具鉴别性的标记,其AUC为94% (95%CI 89-99%)。在100%特异性下,由3个MDMs组(ELM01, ZNF569, 和C13orf18)检测到的GC的敏感性为86% (71-95%)。经GC阶段(图4A),在100%特异性下组的敏感性分别为50%、92%、100%,以及100%,分别用于阶段1 (n=8)、2 (n=13)、3 (n=4) 和4 (n=11) ($p=0.01$, 针对趋势)。定量MDM水平随GC阶段直接增加。甲基化ELM0 1(图4B)的分布表明,对照中正中血浆链计数/ml为0.4,且阶段1、2、3和4的GC病例中的链计数分别为16、111、101和213 ($p=0.01$, 针对趋势)。年龄和性别都不影响标记水平。

[0273] 图5A,图5B和图5C示出12个胃癌标记(ELM01, ZNF569, C13orf18, CD1D, ARHGEF4,

SFMBT2, PPP25RC, CYP26C1, PKIA, CLEC11A, LRRC4, 和 ST8SIA1; ELMO1 被运行两次, 第二次为 biplex 格式) 的性能。在 74 个血浆样本中以 90% 特异性 (图 5A)、95% 特异性 (图 5B) 和 100% 特异性 (图 5C) 的矩阵格式。标记被垂直列出且样本水平排列。样本的排列为正常群位于左边且癌症群位于右边。肯定的命中为浅灰色, 而错过的为深灰色。在此, 表现最好的标记 - ELMO 1 - 被列为第一为 90% 特异性, 且其余组为 100% 特异性。该图允许标记以互补方式被评估。

[0274] 图 6 示出 74 个血浆样本中的 3 个胃癌标记 (最终组) 在 100% 特异性下的性能。标记被垂直列出且样本水平排列。样本的排列为正常群位于左边且癌症群位于右边。癌症按阶段排列。肯定的命中为浅灰色, 而错过的为深灰色 (ELMO1, ZNF569 和 c13orf18)。

[0275] 图 7A-图 7M 示出血浆中 12 个胃癌标记 (ELMO1, ZNF569, C13orf18, CD1D, ARHGEF4, SFMBT2, PPP25RC, CYP26C1, PKIA, CLEC11A, LRRC4, 和 ST8SIA1; ELMO1 被运行两次, 第二次为 biplex 格式) 的框图 (线性比例)。对照样本 (N=38) 被列于左边, 且胃癌病例 (N=36) 被列于右边。垂直轴是归一化为 β -肌动蛋白链的甲基化百分比。

[0276] 表 7

DMR No.	基因	染色体	胰岛范围
274	LRCC4	7	127671993-127672310

[0278] 表 8.

标记	OLIGO 类型	OLIGO 名称	序列
PKIA	引物	F Primer PKIA	GCGAGTTGATCGAGTATTCGG (SEQ ID NO: 71)
	引物	R Primer PKIA	CTAATCAACGCAAATCCCCGC (SEQ ID NO: 72)
	探针	PKIA_Pb_A5	CCACGGACGCGCACGCCGCCA/3C6/ (SEQ ID NO: 73)
C13ORF18	引物	C13ORF18_FP_1	CGTAGTTTTTTAGAAAGTTTTTTAGGGAA GT (SEQ ID NO: 74)
	引物	C13ORF18_RP_1	CGACGTCTCGATACCTACGA (SEQ ID NO: 75)
	探针	C13ORF18_PROBE_1	CCACGGACGACGAAAACGACG /3C6/ (SEQ ID NO: 76)
CLEC11A	引物	CLEC_FP_1	GCGGGAGTTTGGCGTAG (SEQ ID NO: 77)
	引物	CLEC_RP_1	CGCGCAAATACCGAATAAACG (SEQ ID NO: 78)
	探针	CLEC_PROBE_1	CCACGGACGGTCGGTAGATCG/3C6/ (SEQ ID NO: 79)
ELMO	引物	F Primer ELMO1	GTTTTTCGGGTTTCGGGTTTTG (SEQ ID NO: 80)
	引物	R Primer ELMO1	CGCTAAAAATACGACGCCCG (SEQ ID NO: 81)
	探针	Probe A5 ELMO1 v1a	CCACGGACGGCGTTTTATTTCG /3C6/ (SEQ ID NO: 82)
SFMBT2 (896)	引物	F Primer SFMBT2_896 ver2	GCGTTTAGGTTGGTCGGAG (SEQ ID NO: 83)
	引物	R Primer SFMBT2_896 ver2	ACGCACTCAACCTACGAAC (SEQ ID NO: 84)
	探针	Probe A5 SFMBT2_896 ver2	CCACGGACGCCGAAAACTAC/3C6/ (SEQ ID NO: 85)
ZNF569	引物	ZNF569_FP_V2	AGAGTTCGGCGTTTAGAGTTAG (SEQ ID

[0280]

			NO: 86)
	引物	ZNF569 RP_V2	TCGAAACCTATATCCGCGAAC (SEQ ID NO: 87)
	探针	ZNF569 Pb A5_V2	CCACGGACGGCGGTTTTTCGA/3C6/ (SEQ ID NO: 88)
BTACT	引物	ACTB BT FP1	TTTGTTTTTTTGATTAGGTGTTTAAGA (SEQ ID NO: 89)
	引物	ACTB BT RP1	CACCAACCTCATAACCTTATC (SEQ ID NO: 90)
	探针	ACTB BT PB A3	GACGCGGAGATAGTGTTGTGG /3C6/ (SEQ ID NO: 91)
PPP2R5C	引物	F Primer PPP2R5C	TTCGATTTTATTTTTGTTGTCGTTGTAGA (SEQ ID NO: 92)
	引物	PPP2R5C RP_V2	ACGACAAAAAAACCCGACGA (SEQ ID NO: 93)
	探针	PPP2RC_Pb_A1	CGCCGAGGATTCGCGTGGGT/3C6/ (SEQ ID NO: 94)
ARHGEF4	引物	ARHGEF4 FP	CGTTCGCGTTATTTATTTTCGGCG (SEQ ID NO: 95)
	引物	ARHGEF4 RP	GCTCCTAATTCTCATCAACGTCGT (SEQ ID NO: 96)
	探针	ARHGEF4 Pb A1	CGCCGAGGGCGGCGTTTTGC/3C6/ (SEQ ID NO: 97)
CD1D	引物	CD1D_FP_2a	GAAGAGTGCGTAGGTTAGAGG (SEQ ID NO: 98)
	引物	CD1D_RP_2a	CATATCGCCCGACGTAAAAACC (SEQ ID NO: 99)
	探针	CD1D_PROBE_2a_A1	CGCCGAGGCTCGCGAAACGC/3C6/ (SEQ ID NO: 100)
ST8SIA1	引物	ST8SIA1_FP_ESO	AAATATGTGGTTCGTGGCGTT (SEQ ID NO: 101)
	引物	ST8SIA1_RP_ESO	ACGCAACAACGCGAAAAAC (SEQ ID NO: 102)
	探针	ST8SIA1_Pb A1_ESO	CGCCGAGGCGACGAAAAACG/3C6/ (SEQ ID NO: 103)
CYP26C1	引物	CYP26C1 FP	CGTTCGCGTTATTTATTTTCGGCG (SEQ ID NO: 104)
	引物	CYP26C1 RP	GCTCCTAATTCTCATCAACGTCGT (SEQ ID NO: 105)
	探针	CYP26C1 Pb A1	CGCCGAGGGCGGCGTTTTGC/3C6/ (SEQ ID NO: 106)
LRRC4	引物	LRRC4 FP	GCGTTAATTTTCGCGAGGTA (SEQ ID NO: 107)
	引物	LRRC4 RP	ACAATACTCTTATATATTAACGCCGCTC (SEQ ID NO: 108)
	探针	LRRC4 Pb A1	CGCCGAGGAGGCGACGGAGG/3C6/ (SEQ ID NO: 109)

[0281] 上述说明中提到的所有公开和专利,其全部内容被纳入此处作为参考。应理解,在不脱离所述技术的范围和精神下,本领域的技术人员可针对说明的组合物、方法和技术的使用进行各种修改和改变。虽然结合具体的示例性实施例对本技术进行了说明,但应理解,在此请求的本发明不应被解释为受限于这些具体实施例。实际上,对于那些在药理学、生物化学、医学或相关领域的技术人员来说可显而易见对实施本发明的所述模式进行各种修改,并属于后述权利要求的范围内。

- [0001] 序列表
- [0002] <110> 梅约医药教育及研究基金会 (MAYO FOUNDATION FOR MEDICAL EDUCATION AND RESEARCH)
- [0003] 精密科学公司 (EXACT SCIENCES CORPORATION)
- [0004] <120> 检测胃肿瘤
- [0005] <130> EXCTM-34505/WO-1/ORD
- [0006] <150> US 62/212,221
- [0007] <151> 2015-08-31
- [0008] <160> 119
- [0009] <170> PatentIn version 3.5
- [0010] <210> 1
- [0011] <211> 22
- [0012] <212> DNA
- [0013] <213> 人工序列 (Artificial Sequence)
- [0014] <220>
- [0015] <221> misc_feature
- [0016] <223> 合成的 (synthetic)
- [0017] <400> 1
- [0018] agtttggcgt agtcggtaga tc 22
- [0019] <210> 2
- [0020] <211> 22
- [0021] <212> DNA
- [0022] <213> 人工序列 (Artificial Sequence)
- [0023] <220>
- [0024] <221> misc_feature
- [0025] <223> 合成的 (synthetic)
- [0026] <400> 2
- [0027] gcgcgcaaata accgaataaa cg 22
- [0028] <210> 3
- [0029] <211> 17
- [0030] <212> DNA
- [0031] <213> 人工序列 (Artificial Sequence)
- [0032] <220>
- [0033] <221> misc_feature
- [0034] <223> 合成的 (synthetic)
- [0035] <400> 3
- [0036] gcgcgtagcg gcgttttc 17
- [0037] <210> 4

[0038]	<211> 19
[0039]	<212> DNA
[0040]	<213> 人工序列(Artificial Sequence)
[0041]	<220>
[0042]	<221> misc_feature
[0043]	<223> 合成的(synthetic)
[0044]	<400> 4
[0045]	cccatatcgc cgcacgtaa 19
[0046]	<210> 5
[0047]	<211> 25
[0048]	<212> DNA
[0049]	<213> 人工序列(Artificial Sequence)
[0050]	<220>
[0051]	<221> misc_feature
[0052]	<223> 合成的(synthetic)
[0053]	<400> 5
[0054]	aaaacgtttt taggtatttg gtcgt 25
[0055]	<210> 6
[0056]	<211> 25
[0057]	<212> DNA
[0058]	<213> 人工序列(Artificial Sequence)
[0059]	<220>
[0060]	<221> misc_feature
[0061]	<223> 合成的(synthetic)
[0062]	<400> 6
[0063]	ataaaaaaac ctcctccgaa cccgc 25
[0064]	<210> 7
[0065]	<211> 25
[0066]	<212> DNA
[0067]	<213> 人工序列(Artificial Sequence)
[0068]	<220>
[0069]	<221> misc_feature
[0070]	<223> 合成的(synthetic)
[0071]	<400> 7
[0072]	ggagggagtt taacgggtta ggtac 25
[0073]	<210> 8
[0074]	<211> 24
[0075]	<212> DNA
[0076]	<213> 人工序列(Artificial Sequence)

[0077] <220>
[0078] <221> misc_feature
[0079] <223> 合成的(synthetic)
[0080] <400> 8
[0081] cgtaaacaaa aaacaacgac gacg 24
[0082] <210> 9
[0083] <211> 21
[0084] <212> DNA
[0085] <213> 人工序列(Artificial Sequence)
[0086] <220>
[0087] <221> misc_feature
[0088] <223> 合成的(synthetic)
[0089] <400> 9
[0090] gattttgttc gtcgttagtg c 21
[0091] <210> 10
[0092] <211> 19
[0093] <212> DNA
[0094] <213> 人工序列(Artificial Sequence)
[0095] <220>
[0096] <221> misc_feature
[0097] <223> 合成的(synthetic)
[0098] <400> 10
[0099] tctctaaacc cgcgaaacga 19
[0100] <210> 11
[0101] <211> 24
[0102] <212> DNA
[0103] <213> 人工序列(Artificial Sequence)
[0104] <220>
[0105] <221> misc_feature
[0106] <223> 合成的(synthetic)
[0107] <400> 11
[0108] tgtggaatcg gggtttgtgt tcgc 24
[0109] <210> 12
[0110] <211> 25
[0111] <212> DNA
[0112] <213> 人工序列(Artificial Sequence)
[0113] <220>
[0114] <221> misc_feature
[0115] <223> 合成的(synthetic)

[0116] <400> 12
[0117] cccacccaac acaaaaaatc cgacg 25
[0118] <210> 13
[0119] <211> 23
[0120] <212> DNA
[0121] <213> 人工序列(Artificial Sequence)
[0122] <220>
[0123] <221> misc_feature
[0124] <223> 合成的(synthetic)
[0125] <400> 13
[0126] gacgtttgtc gtcgggttcg ttc 23
[0127] <210> 14
[0128] <211> 24
[0129] <212> DNA
[0130] <213> 人工序列(Artificial Sequence)
[0131] <220>
[0132] <221> misc_feature
[0133] <223> 合成的(synthetic)
[0134] <400> 14
[0135] aaaaaccctc cgctaccact tcgc 24
[0136] <210> 15
[0137] <211> 25
[0138] <212> DNA
[0139] <213> 人工序列(Artificial Sequence)
[0140] <220>
[0141] <221> misc_feature
[0142] <223> 合成的(synthetic)
[0143] <400> 15
[0144] ggttttttgg ttatttcgga atcgt 25
[0145] <210> 16
[0146] <211> 25
[0147] <212> DNA
[0148] <213> 人工序列(Artificial Sequence)
[0149] <220>
[0150] <221> misc_feature
[0151] <223> 合成的(synthetic)
[0152] <400> 16
[0153] tataaaaacg cgcgtaatca acgct 25
[0154] <210> 17

[0155] <211> 30
[0156] <212> DNA
[0157] <213> 人工序列(Artificial Sequence)
[0158] <220>
[0159] <221> misc_feature
[0160] <223> 合成的(synthetic)
[0161] <400> 17
[0162] tgtttttcgcg gtcgttatat attacgtcgt 30
[0163] <210> 18
[0164] <211> 25
[0165] <212> DNA
[0166] <213> 人工序列(Artificial Sequence)
[0167] <220>
[0168] <221> misc_feature
[0169] <223> 合成的(synthetic)
[0170] <400> 18
[0171] gaactatccc cgaactccga ctcga 25
[0172] <210> 19
[0173] <211> 25
[0174] <212> DNA
[0175] <213> 人工序列(Artificial Sequence)
[0176] <220>
[0177] <221> misc_feature
[0178] <223> 合成的(synthetic)
[0179] <400> 19
[0180] tttagggaag taaagcgtcg ttttc 25
[0181] <210> 20
[0182] <211> 22
[0183] <212> DNA
[0184] <213> 人工序列(Artificial Sequence)
[0185] <220>
[0186] <221> misc_feature
[0187] <223> 合成的(synthetic)
[0188] <400> 20
[0189] aacgacgtct cgatacctac ga 22
[0190] <210> 21
[0191] <211> 34
[0192] <212> DNA
[0193] <213> 人工序列(Artificial Sequence)

[0194] <220>
[0195] <221> misc_feature
[0196] <223> 合成的(synthetic)
[0197] <400> 21
[0198] ttatatattttt cgtttttagt aatttcgcgt tagc 34
[0199] <210> 22
[0200] <211> 23
[0201] <212> DNA
[0202] <213> 人工序列(Artificial Sequence)
[0203] <220>
[0204] <221> misc_feature
[0205] <223> 合成的(synthetic)
[0206] <400> 22
[0207] gaaaacccgc cgaaacattt cga 23
[0208] <210> 23
[0209] <211> 30
[0210] <212> DNA
[0211] <213> 人工序列(Artificial Sequence)
[0212] <220>
[0213] <221> misc_feature
[0214] <223> 合成的(synthetic)
[0215] <400> 23
[0216] tgttttcgcg gtcgttatat attacgtcgt 30
[0217] <210> 24
[0218] <211> 25
[0219] <212> DNA
[0220] <213> 人工序列(Artificial Sequence)
[0221] <220>
[0222] <221> misc_feature
[0223] <223> 合成的(synthetic)
[0224] <400> 24
[0225] gaactatccc cgaactccga ctcca 25
[0226] <210> 25
[0227] <211> 31
[0228] <212> DNA
[0229] <213> 人工序列(Artificial Sequence)
[0230] <220>
[0231] <221> misc_feature
[0232] <223> 合成的(synthetic)

[0233]	<400> 25
[0234]	cgggttttag cgatgtttat tttagtttcg t 31
[0235]	<210> 26
[0236]	<211> 26
[0237]	<212> DNA
[0238]	<213> 人工序列(Artificial Sequence)
[0239]	<220>
[0240]	<221> misc_feature
[0241]	<223> 合成的(synthetic)
[0242]	<400> 26
[0243]	ccttttcggtt cgtataaaaat ttcggtt 26
[0244]	<210> 27
[0245]	<211> 30
[0246]	<212> DNA
[0247]	<213> 人工序列(Artificial Sequence)
[0248]	<220>
[0249]	<221> misc_feature
[0250]	<223> 合成的(synthetic)
[0251]	<400> 27
[0252]	cgggggtaaa aagggtagcg atagtgatac 30
[0253]	<210> 28
[0254]	<211> 23
[0255]	<212> DNA
[0256]	<213> 人工序列(Artificial Sequence)
[0257]	<220>
[0258]	<221> misc_feature
[0259]	<223> 合成的(synthetic)
[0260]	<400> 28
[0261]	cgaaaaattc gaaacgaaac gtc 23
[0262]	<210> 29
[0263]	<211> 22
[0264]	<212> DNA
[0265]	<213> 人工序列(Artificial Sequence)
[0266]	<220>
[0267]	<221> misc_feature
[0268]	<223> 合成的(synthetic)
[0269]	<400> 29
[0270]	agtttggcgt agtcggtaga tc 22
[0271]	<210> 30

[0272]	<211> 22
[0273]	<212> DNA
[0274]	<213> 人工序列(Artificial Sequence)
[0275]	<220>
[0276]	<221> misc_feature
[0277]	<223> 合成的(synthetic)
[0278]	<400> 30
[0279]	gcgcgcaa at accgaataaa cg 22
[0280]	<210> 31
[0281]	<211> 23
[0282]	<212> DNA
[0283]	<213> 人工序列(Artificial Sequence)
[0284]	<220>
[0285]	<221> misc_feature
[0286]	<223> 合成的(synthetic)
[0287]	<400> 31
[0288]	gtttaatttt cggtttcgtc gtc 23
[0289]	<210> 32
[0290]	<211> 18
[0291]	<212> DNA
[0292]	<213> 人工序列(Artificial Sequence)
[0293]	<220>
[0294]	<221> misc_feature
[0295]	<223> 合成的(synthetic)
[0296]	<400> 32
[0297]	cgctacgaaa cactccga 18
[0298]	<210> 33
[0299]	<211> 24
[0300]	<212> DNA
[0301]	<213> 人工序列(Artificial Sequence)
[0302]	<220>
[0303]	<221> misc_feature
[0304]	<223> 合成的(synthetic)
[0305]	<400> 33
[0306]	ttttgttggg tgggtatttt cgcg 24
[0307]	<210> 34
[0308]	<211> 25
[0309]	<212> DNA
[0310]	<213> 人工序列(Artificial Sequence)

[0311] <220>
[0312] <221> misc_feature
[0313] <223> 合成的(synthetic)
[0314] <400> 34
[0315] acaaccctca aaaccctcg atatc 25
[0316] <210> 35
[0317] <211> 27
[0318] <212> DNA
[0319] <213> 人工序列(Artificial Sequence)
[0320] <220>
[0321] <221> misc_feature
[0322] <223> 合成的(synthetic)
[0323] <400> 35
[0324] tttatttatg gattacgttt ttagcga 27
[0325] <210> 36
[0326] <211> 20
[0327] <212> DNA
[0328] <213> 人工序列(Artificial Sequence)
[0329] <220>
[0330] <221> misc_feature
[0331] <223> 合成的(synthetic)
[0332] <400> 36
[0333] gacgaccgaa cgatcacgca 20
[0334] <210> 37
[0335] <211> 21
[0336] <212> DNA
[0337] <213> 人工序列(Artificial Sequence)
[0338] <220>
[0339] <221> misc_feature
[0340] <223> 合成的(synthetic)
[0341] <400> 37
[0342] gcgtcgtag tagtacgaag c 21
[0343] <210> 38
[0344] <211> 21
[0345] <212> DNA
[0346] <213> 人工序列(Artificial Sequence)
[0347] <220>
[0348] <221> misc_feature
[0349] <223> 合成的(synthetic)

[0350] <400> 38
[0351] gcacctcaac gaaaacaccg a 21
[0352] <210> 39
[0353] <211> 17
[0354] <212> DNA
[0355] <213> 人工序列(Artificial Sequence)
[0356] <220>
[0357] <221> misc_feature
[0358] <223> 合成的(synthetic)
[0359] <400> 39
[0360] gcgcgtagcg gcgtttc 17
[0361] <210> 40
[0362] <211> 19
[0363] <212> DNA
[0364] <213> 人工序列(Artificial Sequence)
[0365] <220>
[0366] <221> misc_feature
[0367] <223> 合成的(synthetic)
[0368] <400> 40
[0369] cccatatcgc cgcacgtaa 19
[0370] <210> 41
[0371] <211> 24
[0372] <212> DNA
[0373] <213> 人工序列(Artificial Sequence)
[0374] <220>
[0375] <221> misc_feature
[0376] <223> 合成的(synthetic)
[0377] <400> 41
[0378] gtcgtttttg ggcgttttag aggc 24
[0379] <210> 42
[0380] <211> 20
[0381] <212> DNA
[0382] <213> 人工序列(Artificial Sequence)
[0383] <220>
[0384] <221> misc_feature
[0385] <223> 合成的(synthetic)
[0386] <400> 42
[0387] taatcccaaa cgaaccgccg 20
[0388] <210> 43

[0389] <211> 25
[0390] <212> DNA
[0391] <213> 人工序列(Artificial Sequence)
[0392] <220>
[0393] <221> misc_feature
[0394] <223> 合成的(synthetic)
[0395] <400> 43
[0396] ggtttttttg ttatttcgga atcgt 25
[0397] <210> 44
[0398] <211> 25
[0399] <212> DNA
[0400] <213> 人工序列(Artificial Sequence)
[0401] <220>
[0402] <221> misc_feature
[0403] <223> 合成的(synthetic)
[0404] <400> 44
[0405] tataaaaacg cgcgtaatca acgct 25
[0406] <210> 45
[0407] <211> 25
[0408] <212> DNA
[0409] <213> 人工序列(Artificial Sequence)
[0410] <220>
[0411] <221> misc_feature
[0412] <223> 合成的(synthetic)
[0413] <400> 45
[0414] ttgagatggt gggtgaaggc gattc 25
[0415] <210> 46
[0416] <211> 26
[0417] <212> DNA
[0418] <213> 人工序列(Artificial Sequence)
[0419] <220>
[0420] <221> misc_feature
[0421] <223> 合成的(synthetic)
[0422] <400> 46
[0423] cgctaacgcg aaaaaataat tcgacg 26
[0424] <210> 47
[0425] <211> 28
[0426] <212> DNA
[0427] <213> 人工序列(Artificial Sequence)

[0428] <220>
[0429] <221> misc_feature
[0430] <223> 合成的 (synthetic)
[0431] <400> 47
[0432] cggtttttatt tattatgatt cgtagcgg 28
[0433] <210> 48
[0434] <211> 24
[0435] <212> DNA
[0436] <213> 人工序列 (Artificial Sequence)
[0437] <220>
[0438] <221> misc_feature
[0439] <223> 合成的 (synthetic)
[0440] <400> 48
[0441] cgactaccct aaacaacgca tcgc 24
[0442] <210> 49
[0443] <211> 24
[0444] <212> DNA
[0445] <213> 人工序列 (Artificial Sequence)
[0446] <220>
[0447] <221> misc_feature
[0448] <223> 合成的 (synthetic)
[0449] <400> 49
[0450] cgactaccct aaacaacgca tcgc 24
[0451] <210> 50
[0452] <211> 24
[0453] <212> DNA
[0454] <213> 人工序列 (Artificial Sequence)
[0455] <220>
[0456] <221> misc_feature
[0457] <223> 合成的 (synthetic)
[0458] <400> 50
[0459] cccgccgaat actcgatcaa ctcg 24
[0460] <210> 51
[0461] <211> 19
[0462] <212> DNA
[0463] <213> 人工序列 (Artificial Sequence)
[0464] <220>
[0465] <221> misc_feature
[0466] <223> 合成的 (synthetic)

[0467] <400> 51
[0468] ggtcgcgtcg ttttcgatc 19
[0469] <210> 52
[0470] <211> 17
[0471] <212> DNA
[0472] <213> 人工序列(Artificial Sequence)
[0473] <220>
[0474] <221> misc_feature
[0475] <223> 合成的(synthetic)
[0476] <400> 52
[0477] gccgcaaacg ccgacga 17
[0478] <210> 53
[0479] <211> 23
[0480] <212> DNA
[0481] <213> 人工序列(Artificial Sequence)
[0482] <220>
[0483] <221> misc_feature
[0484] <223> 合成的(synthetic)
[0485] <400> 53
[0486] tgtatatattc gaggcggttt cgg 23
[0487] <210> 54
[0488] <211> 23
[0489] <212> DNA
[0490] <213> 人工序列(Artificial Sequence)
[0491] <220>
[0492] <221> misc_feature
[0493] <223> 合成的(synthetic)
[0494] <400> 54
[0495] cgcccccaaa ataaaaaac gaa 23
[0496] <210> 55
[0497] <211> 28
[0498] <212> DNA
[0499] <213> 人工序列(Artificial Sequence)
[0500] <220>
[0501] <221> misc_feature
[0502] <223> 合成的(synthetic)
[0503] <400> 55
[0504] tttttcgttc gtttttttat tttttcga 28
[0505] <210> 56

[0506] <211> 28
[0507] <212> DNA
[0508] <213> 人工序列(Artificial Sequence)
[0509] <220>
[0510] <221> misc_feature
[0511] <223> 合成的(synthetic)
[0512] <400> 56
[0513] aacgcatacc atcgacttca actacgac 28
[0514] <210> 57
[0515] <211> 21
[0516] <212> DNA
[0517] <213> 人工序列(Artificial Sequence)
[0518] <220>
[0519] <221> misc_feature
[0520] <223> 合成的(synthetic)
[0521] <400> 57
[0522] cggtttttcgt tcgttttttc g 21
[0523] <210> 58
[0524] <211> 18
[0525] <212> DNA
[0526] <213> 人工序列(Artificial Sequence)
[0527] <220>
[0528] <221> misc_feature
[0529] <223> 合成的(synthetic)
[0530] <400> 58
[0531] ccgccttcta cgcgacta 18
[0532] <210> 59
[0533] <211> 25
[0534] <212> DNA
[0535] <213> 人工序列(Artificial Sequence)
[0536] <220>
[0537] <221> misc_feature
[0538] <223> 合成的(synthetic)
[0539] <400> 59
[0540] gtattttttaa tatttggtga gtgcg 25
[0541] <210> 60
[0542] <211> 25
[0543] <212> DNA
[0544] <213> 人工序列(Artificial Sequence)

[0545] <220>
[0546] <221> misc_feature
[0547] <223> 合成的(synthetic)
[0548] <400> 60
[0549] tactaaaacg tacgataaac atcgt 25
[0550] <210> 61
[0551] <211> 33
[0552] <212> DNA
[0553] <213> 人工序列(Artificial Sequence)
[0554] <220>
[0555] <221> misc_feature
[0556] <223> 合成的(synthetic)
[0557] <400> 61
[0558] tcgattttat ttttgttgtc gttgtagatt cgc 33
[0559] <210> 62
[0560] <211> 33
[0561] <212> DNA
[0562] <213> 人工序列(Artificial Sequence)
[0563] <220>
[0564] <221> misc_feature
[0565] <223> 合成的(synthetic)
[0566] <400> 62
[0567] gaaaaaacta aaaaacgaca aaaaaacccg acg 33
[0568] <210> 63
[0569] <211> 25
[0570] <212> DNA
[0571] <213> 人工序列(Artificial Sequence)
[0572] <220>
[0573] <221> misc_feature
[0574] <223> 合成的(synthetic)
[0575] <400> 63
[0576] ggaggagatt taacgggtta ggtac 25
[0577] <210> 64
[0578] <211> 24
[0579] <212> DNA
[0580] <213> 人工序列(Artificial Sequence)
[0581] <220>
[0582] <221> misc_feature
[0583] <223> 合成的(synthetic)

[0584] <400> 64
[0585] cgtaaacaaa aaacaacgac gacg 24
[0586] <210> 65
[0587] <211> 25
[0588] <212> DNA
[0589] <213> 人工序列(Artificial Sequence)
[0590] <220>
[0591] <221> misc_feature
[0592] <223> 合成的(synthetic)
[0593] <400> 65
[0594] attgcgtcgt tgtttaggta acgta 25
[0595] <210> 66
[0596] <211> 23
[0597] <212> DNA
[0598] <213> 人工序列(Artificial Sequence)
[0599] <220>
[0600] <221> misc_feature
[0601] <223> 合成的(synthetic)
[0602] <400> 66
[0603] cgaaaatacc tatctcctcg act 23
[0604] <210> 67
[0605] <211> 28
[0606] <212> DNA
[0607] <213> 人工序列(Artificial Sequence)
[0608] <220>
[0609] <221> misc_feature
[0610] <223> 合成的(synthetic)
[0611] <400> 67
[0612] ggaacgagtg atagtcggat agttcgtc 28
[0613] <210> 68
[0614] <211> 18
[0615] <212> DNA
[0616] <213> 人工序列(Artificial Sequence)
[0617] <220>
[0618] <221> misc_feature
[0619] <223> 合成的(synthetic)
[0620] <400> 68
[0621] cgcccgaaaa cgaccccg 18
[0622] <210> 69

[0623] <211> 21
[0624] <212> DNA
[0625] <213> 人工序列(Artificial Sequence)
[0626] <220>
[0627] <221> misc_feature
[0628] <223> 合成的(synthetic)
[0629] <400> 69
[0630] gggagtgagg gtagggcggtt c 21
[0631] <210> 70
[0632] <211> 25
[0633] <212> DNA
[0634] <213> 人工序列(Artificial Sequence)
[0635] <220>
[0636] <221> misc_feature
[0637] <223> 合成的(synthetic)
[0638] <400> 70
[0639] ctcgcaaccc cttcgaatta acccg 25
[0640] <210> 71
[0641] <211> 21
[0642] <212> DNA
[0643] <213> 人工序列(Artificial Sequence)
[0644] <220>
[0645] <221> misc_feature
[0646] <223> 合成的(synthetic)
[0647] <400> 71
[0648] gcgagttgat cgagtattcg g 21
[0649] <210> 72
[0650] <211> 21
[0651] <212> DNA
[0652] <213> 人工序列(Artificial Sequence)
[0653] <220>
[0654] <221> misc_feature
[0655] <223> 合成的(synthetic)
[0656] <400> 72
[0657] ctaatcaacg caaatccccg c 21
[0658] <210> 73
[0659] <211> 22
[0660] <212> DNA
[0661] <213> 人工序列(Artificial Sequence)

[0662] <220>
[0663] <221> misc_feature
[0664] <223> 合成的 (synthetic)
[0665] <400> 73
[0666] ccacggacgc gcacgccgcc ac 22
[0667] <210> 74
[0668] <211> 30
[0669] <212> DNA
[0670] <213> 人工序列 (Artificial Sequence)
[0671] <220>
[0672] <221> misc_feature
[0673] <223> 合成的 (synthetic)
[0674] <400> 74
[0675] cgtagttttt tagaagtttt ttagggaagt 30
[0676] <210> 75
[0677] <211> 20
[0678] <212> DNA
[0679] <213> 人工序列 (Artificial Sequence)
[0680] <220>
[0681] <221> misc_feature
[0682] <223> 合成的 (synthetic)
[0683] <400> 75
[0684] cgacgtctcg atacctacga 20
[0685] <210> 76
[0686] <211> 22
[0687] <212> DNA
[0688] <213> 人工序列 (Artificial Sequence)
[0689] <220>
[0690] <221> misc_feature
[0691] <223> 合成的 (synthetic)
[0692] <400> 76
[0693] ccacggacga cgaaaacgac gc 22
[0694] <210> 77
[0695] <211> 17
[0696] <212> DNA
[0697] <213> 人工序列 (Artificial Sequence)
[0698] <220>
[0699] <221> misc_feature
[0700] <223> 合成的 (synthetic)

[0701] <400> 77
[0702] gcgggagttt ggcgtag 17
[0703] <210> 78
[0704] <211> 21
[0705] <212> DNA
[0706] <213> 人工序列(Artificial Sequence)
[0707] <220>
[0708] <221> misc_feature
[0709] <223> 合成的(synthetic)
[0710] <400> 78
[0711] cgcgcaaata ccgaataaac g 21
[0712] <210> 79
[0713] <211> 22
[0714] <212> DNA
[0715] <213> 人工序列(Artificial Sequence)
[0716] <220>
[0717] <221> misc_feature
[0718] <223> 合成的(synthetic)
[0719] <400> 79
[0720] ccacggacgg tcggtagatc gc 22
[0721] <210> 80
[0722] <211> 21
[0723] <212> DNA
[0724] <213> 人工序列(Artificial Sequence)
[0725] <220>
[0726] <221> misc_feature
[0727] <223> 合成的(synthetic)
[0728] <400> 80
[0729] gttttcgggt ttcgggtttt g 21
[0730] <210> 81
[0731] <211> 20
[0732] <212> DNA
[0733] <213> 人工序列(Artificial Sequence)
[0734] <220>
[0735] <221> misc_feature
[0736] <223> 合成的(synthetic)
[0737] <400> 81
[0738] cgctaaaaat acgacgcccg 20
[0739] <210> 82

[0740] <211> 22
[0741] <212> DNA
[0742] <213> 人工序列(Artificial Sequence)
[0743] <220>
[0744] <221> misc_feature
[0745] <223> 合成的(synthetic)
[0746] <400> 82
[0747] ccacggacgg cgttttattc gc 22
[0748] <210> 83
[0749] <211> 19
[0750] <212> DNA
[0751] <213> 人工序列(Artificial Sequence)
[0752] <220>
[0753] <221> misc_feature
[0754] <223> 合成的(synthetic)
[0755] <400> 83
[0756] gcgttttaggt tggtcggag 19
[0757] <210> 84
[0758] <211> 19
[0759] <212> DNA
[0760] <213> 人工序列(Artificial Sequence)
[0761] <220>
[0762] <221> misc_feature
[0763] <223> 合成的(synthetic)
[0764] <400> 84
[0765] acgcactcaa cctacgaac 19
[0766] <210> 85
[0767] <211> 22
[0768] <212> DNA
[0769] <213> 人工序列(Artificial Sequence)
[0770] <220>
[0771] <221> misc_feature
[0772] <223> 合成的(synthetic)
[0773] <400> 85
[0774] ccacggacgc cgaaaaacta cc 22
[0775] <210> 86
[0776] <211> 22
[0777] <212> DNA
[0778] <213> 人工序列(Artificial Sequence)

[0779] <220>
[0780] <221> misc_feature
[0781] <223> 合成的 (synthetic)
[0782] <400> 86
[0783] agagttcggc gtttagagtt ag 22
[0784] <210> 87
[0785] <211> 21
[0786] <212> DNA
[0787] <213> 人工序列 (Artificial Sequence)
[0788] <220>
[0789] <221> misc_feature
[0790] <223> 合成的 (synthetic)
[0791] <400> 87
[0792] tcgaaaccta tatccgcgaa c 21
[0793] <210> 88
[0794] <211> 22
[0795] <212> DNA
[0796] <213> 人工序列 (Artificial Sequence)
[0797] <220>
[0798] <221> misc_feature
[0799] <223> 合成的 (synthetic)
[0800] <400> 88
[0801] ccacggacgg cggtttttcg ac 22
[0802] <210> 89
[0803] <211> 27
[0804] <212> DNA
[0805] <213> 人工序列 (Artificial Sequence)
[0806] <220>
[0807] <221> misc_feature
[0808] <223> 合成的 (synthetic)
[0809] <400> 89
[0810] tttgtttttt tgattaggtg tttaaga 27
[0811] <210> 90
[0812] <211> 21
[0813] <212> DNA
[0814] <213> 人工序列 (Artificial Sequence)
[0815] <220>
[0816] <221> misc_feature
[0817] <223> 合成的 (synthetic)

[0818] <400> 90
[0819] caccaacctc ataaccttat c 21
[0820] <210> 91
[0821] <211> 22
[0822] <212> DNA
[0823] <213> 人工序列(Artificial Sequence)
[0824] <220>
[0825] <221> misc_feature
[0826] <223> 合成的(synthetic)
[0827] <400> 91
[0828] gacgcggaga tagtgttgtg gc 22
[0829] <210> 92
[0830] <211> 29
[0831] <212> DNA
[0832] <213> 人工序列(Artificial Sequence)
[0833] <220>
[0834] <221> misc_feature
[0835] <223> 合成的(synthetic)
[0836] <400> 92
[0837] ttcgatttta tttttgttgt cgttgtaga 29
[0838] <210> 93
[0839] <211> 20
[0840] <212> DNA
[0841] <213> 人工序列(Artificial Sequence)
[0842] <220>
[0843] <221> misc_feature
[0844] <223> 合成的(synthetic)
[0845] <400> 93
[0846] acgacaaaaa aacccgacga 20
[0847] <210> 94
[0848] <211> 21
[0849] <212> DNA
[0850] <213> 人工序列(Artificial Sequence)
[0851] <220>
[0852] <221> misc_feature
[0853] <223> 合成的(synthetic)
[0854] <400> 94
[0855] cgccgaggat tcgcgtgggt c 21
[0856] <210> 95

[0857] <211> 23
[0858] <212> DNA
[0859] <213> 人工序列(Artificial Sequence)
[0860] <220>
[0861] <221> misc_feature
[0862] <223> 合成的(synthetic)
[0863] <400> 95
[0864] cgttcgcgtt atttatttcg gcg 23
[0865] <210> 96
[0866] <211> 24
[0867] <212> DNA
[0868] <213> 人工序列(Artificial Sequence)
[0869] <220>
[0870] <221> misc_feature
[0871] <223> 合成的(synthetic)
[0872] <400> 96
[0873] gtcctaatt ctcatcaacg tcgt 24
[0874] <210> 97
[0875] <211> 21
[0876] <212> DNA
[0877] <213> 人工序列(Artificial Sequence)
[0878] <220>
[0879] <221> misc_feature
[0880] <223> 合成的(synthetic)
[0881] <400> 97
[0882] cgccgagggc ggcgttttgc c 21
[0883] <210> 98
[0884] <211> 21
[0885] <212> DNA
[0886] <213> 人工序列(Artificial Sequence)
[0887] <220>
[0888] <221> misc_feature
[0889] <223> 合成的(synthetic)
[0890] <400> 98
[0891] gaagagtgcg taggttagag g 21
[0892] <210> 99
[0893] <211> 22
[0894] <212> DNA
[0895] <213> 人工序列(Artificial Sequence)

[0896] <220>
[0897] <221> misc_feature
[0898] <223> 合成的 (synthetic)
[0899] <400> 99
[0900] catatcgccc gacgtaaaaa cc 22
[0901] <210> 100
[0902] <211> 21
[0903] <212> DNA
[0904] <213> 人工序列 (Artificial Sequence)
[0905] <220>
[0906] <221> misc_feature
[0907] <223> 合成的 (synthetic)
[0908] <400> 100
[0909] cgccgaggct cgcgaaacgc c 21
[0910] <210> 101
[0911] <211> 21
[0912] <212> DNA
[0913] <213> 人工序列 (Artificial Sequence)
[0914] <220>
[0915] <221> misc_feature
[0916] <223> 合成的 (synthetic)
[0917] <400> 101
[0918] aaatatgtgg ttcgtggcgt t 21
[0919] <210> 102
[0920] <211> 19
[0921] <212> DNA
[0922] <213> 人工序列 (Artificial Sequence)
[0923] <220>
[0924] <221> misc_feature
[0925] <223> 合成的 (synthetic)
[0926] <400> 102
[0927] acgcaacaac gcgaaaaac 19
[0928] <210> 103
[0929] <211> 21
[0930] <212> DNA
[0931] <213> 人工序列 (Artificial Sequence)
[0932] <220>
[0933] <221> misc_feature
[0934] <223> 合成的 (synthetic)

[0935] <400> 103
[0936] cgccgaggcg acgaaaaacg c 21
[0937] <210> 104
[0938] <211> 23
[0939] <212> DNA
[0940] <213> 人工序列(Artificial Sequence)
[0941] <220>
[0942] <221> misc_feature
[0943] <223> 合成的(synthetic)
[0944] <400> 104
[0945] cgttcgcggtt atttatttcg gcg 23
[0946] <210> 105
[0947] <211> 24
[0948] <212> DNA
[0949] <213> 人工序列(Artificial Sequence)
[0950] <220>
[0951] <221> misc_feature
[0952] <223> 合成的(synthetic)
[0953] <400> 105
[0954] gctcctaatt ctcatcaacg tcgt 24
[0955] <210> 106
[0956] <211> 21
[0957] <212> DNA
[0958] <213> 人工序列(Artificial Sequence)
[0959] <220>
[0960] <221> misc_feature
[0961] <223> 合成的(synthetic)
[0962] <400> 106
[0963] cgccgagggc ggcgttttgc c 21
[0964] <210> 107
[0965] <211> 19
[0966] <212> DNA
[0967] <213> 人工序列(Artificial Sequence)
[0968] <220>
[0969] <221> misc_feature
[0970] <223> 合成的(synthetic)
[0971] <400> 107
[0972] gcgttaattt cgcgaggta 19
[0973] <210> 108

[0974] <211> 28
[0975] <212> DNA
[0976] <213> 人工序列(Artificial Sequence)
[0977] <220>
[0978] <221> misc_feature
[0979] <223> 合成的(synthetic)
[0980] <400> 108
[0981] acaataactct tatatattaa cgccgctc 28
[0982] <210> 109
[0983] <211> 21
[0984] <212> DNA
[0985] <213> 人工序列(Artificial Sequence)
[0986] <220>
[0987] <221> misc_feature
[0988] <223> 合成的(synthetic)
[0989] <400> 109
[0990] cgccgaggag gcgacggagg c 21
[0991] <210> 110
[0992] <211> 29
[0993] <212> DNA
[0994] <213> 人工序列(Artificial Sequence)
[0995] <220>
[0996] <221> misc_feature
[0997] <223> 合成的(synthetic)
[0998] <400> 110
[0999] agccggtttt ccggctgaga ctccgcgctc 29
[1000] <210> 111
[1001] <211> 32
[1002] <212> DNA
[1003] <213> 人工序列(Artificial Sequence)
[1004] <220>
[1005] <221> misc_feature
[1006] <223> 合成的(synthetic)
[1007] <400> 111
[1008] tctagccggt ttccggctg agactccgcg tc 32
[1009] <210> 112
[1010] <211> 29
[1011] <212> DNA
[1012] <213> 人工序列(Artificial Sequence)

[1013]	<220>
[1014]	<221> misc_feature
[1015]	<223> 合成的(synthetic)
[1016]	<400> 112
[1017]	agccgggtttt ccggctgaga cgtccgtgg 29
[1018]	<210> 113
[1019]	<211> 32
[1020]	<212> DNA
[1021]	<213> 人工序列(Artificial Sequence)
[1022]	<220>
[1023]	<221> misc_feature
[1024]	<223> 合成的(synthetic)
[1025]	<400> 113
[1026]	tctagccggt tttccggctg agacgtccgt gg 32
[1027]	<210> 114
[1028]	<211> 28
[1029]	<212> DNA
[1030]	<213> 人工序列(Artificial Sequence)
[1031]	<220>
[1032]	<221> misc_feature
[1033]	<223> 合成的(synthetic)
[1034]	<400> 114
[1035]	agccgggtttt ccggctgaga cctcggcg 28
[1036]	<210> 115
[1037]	<211> 31
[1038]	<212> DNA
[1039]	<213> 人工序列(Artificial Sequence)
[1040]	<220>
[1041]	<221> misc_feature
[1042]	<223> 合成的(synthetic)
[1043]	<400> 115
[1044]	tctagccggt tttccggctg agacctcggc g 31
[1045]	<210> 116
[1046]	<211> 8
[1047]	<212> DNA
[1048]	<213> 人工序列(Artificial Sequence)
[1049]	<220>
[1050]	<221> misc_feature
[1051]	<223> 合成的(synthetic)

[1052]	<400> 116
[1053]	cgccgagg 8
[1054]	<210> 117
[1055]	<211> 9
[1056]	<212> DNA
[1057]	<213> 人工序列(Artificial Sequence)
[1058]	<220>
[1059]	<221> misc_feature
[1060]	<223> 合成的(synthetic)
[1061]	<400> 117
[1062]	gacgcggag 9
[1063]	<210> 118
[1064]	<211> 9
[1065]	<212> DNA
[1066]	<213> 人工序列(Artificial Sequence)
[1067]	<220>
[1068]	<221> misc_feature
[1069]	<223> 合成的(synthetic)
[1070]	<400> 118
[1071]	ccacggacg 9
[1072]	<210> 119
[1073]	<211> 8
[1074]	<212> DNA
[1075]	<213> 人工序列(Artificial Sequence)
[1076]	<220>
[1077]	<221> misc_feature
[1078]	<223> 合成的(synthetic)
[1079]	<400> 119
[1080]	gcgcgtcc 8

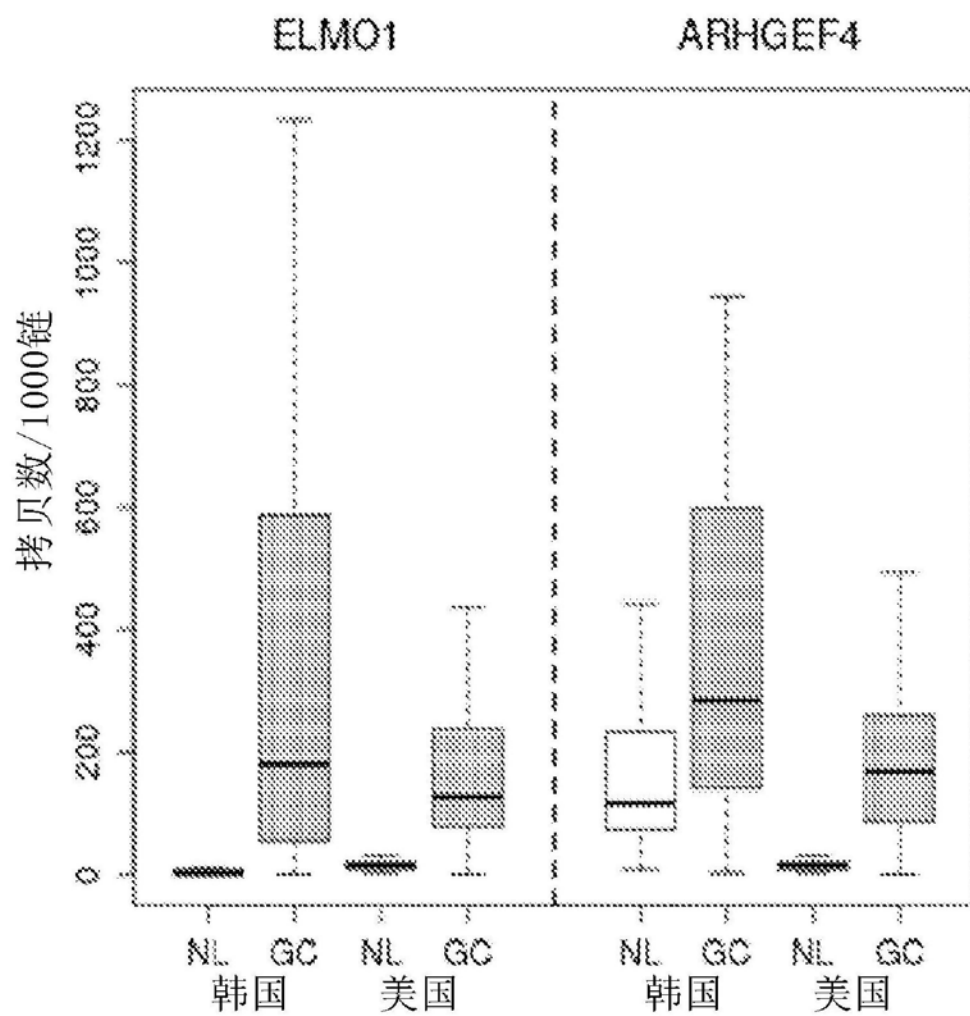


图1

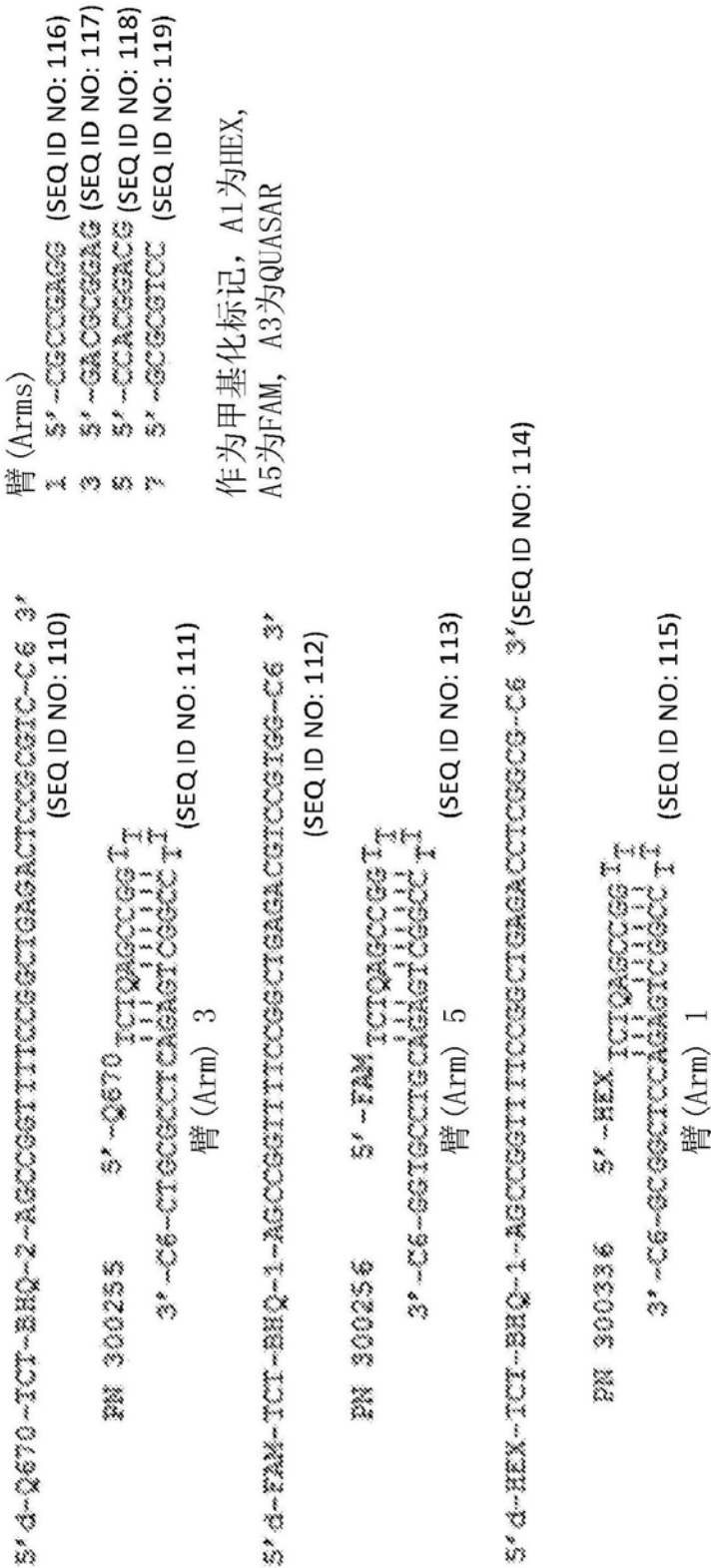


图2

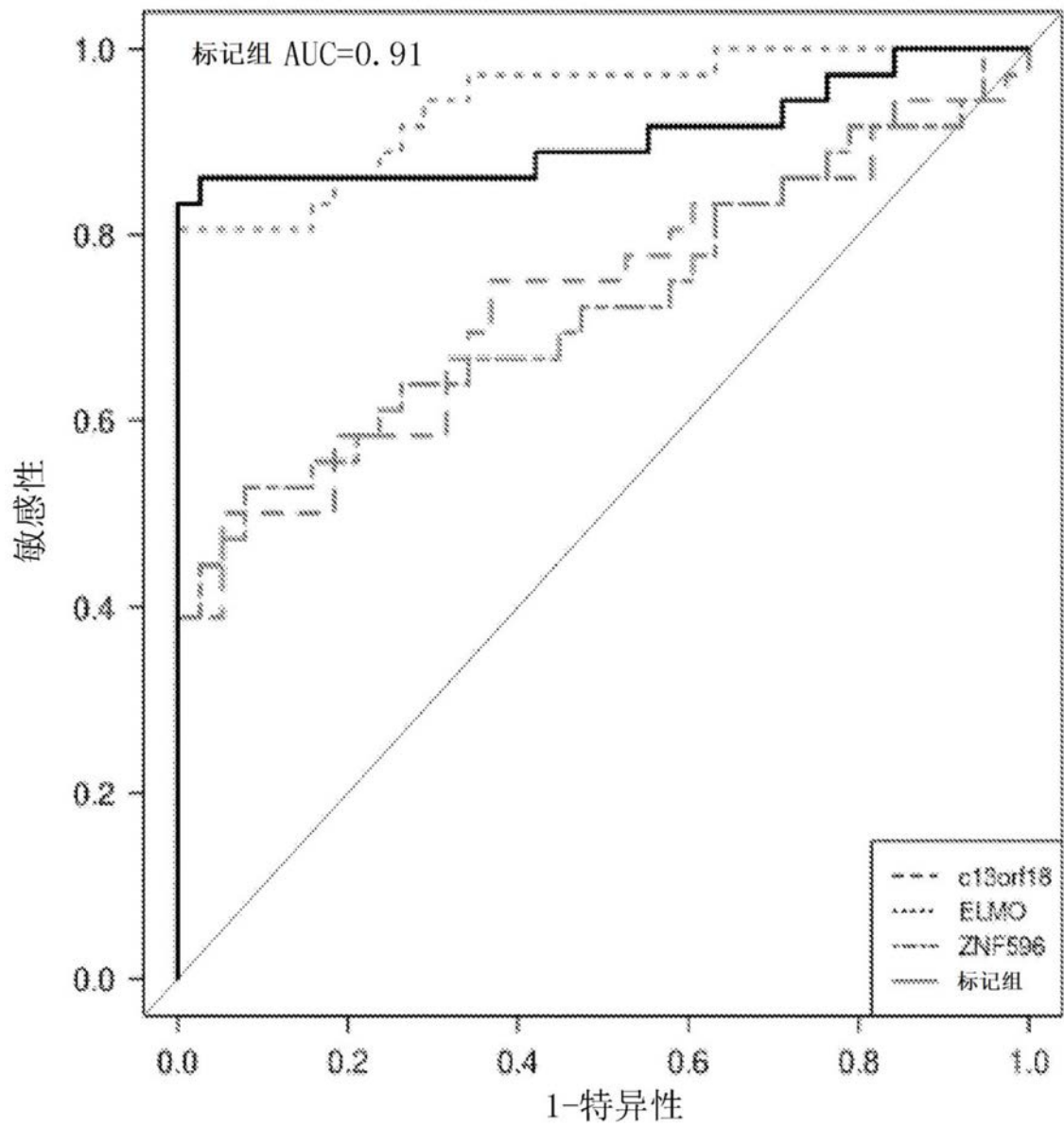


图3

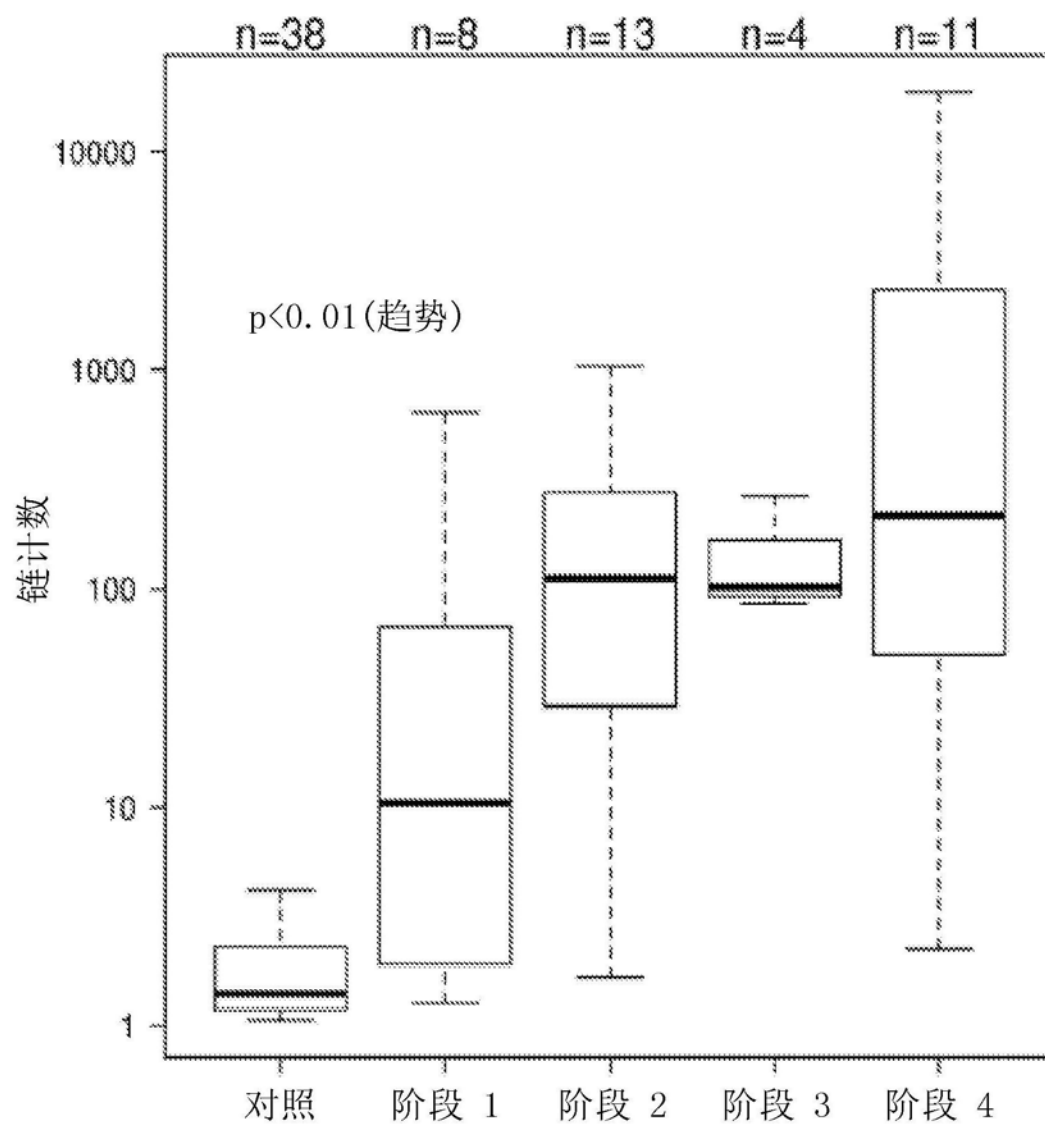


图4A

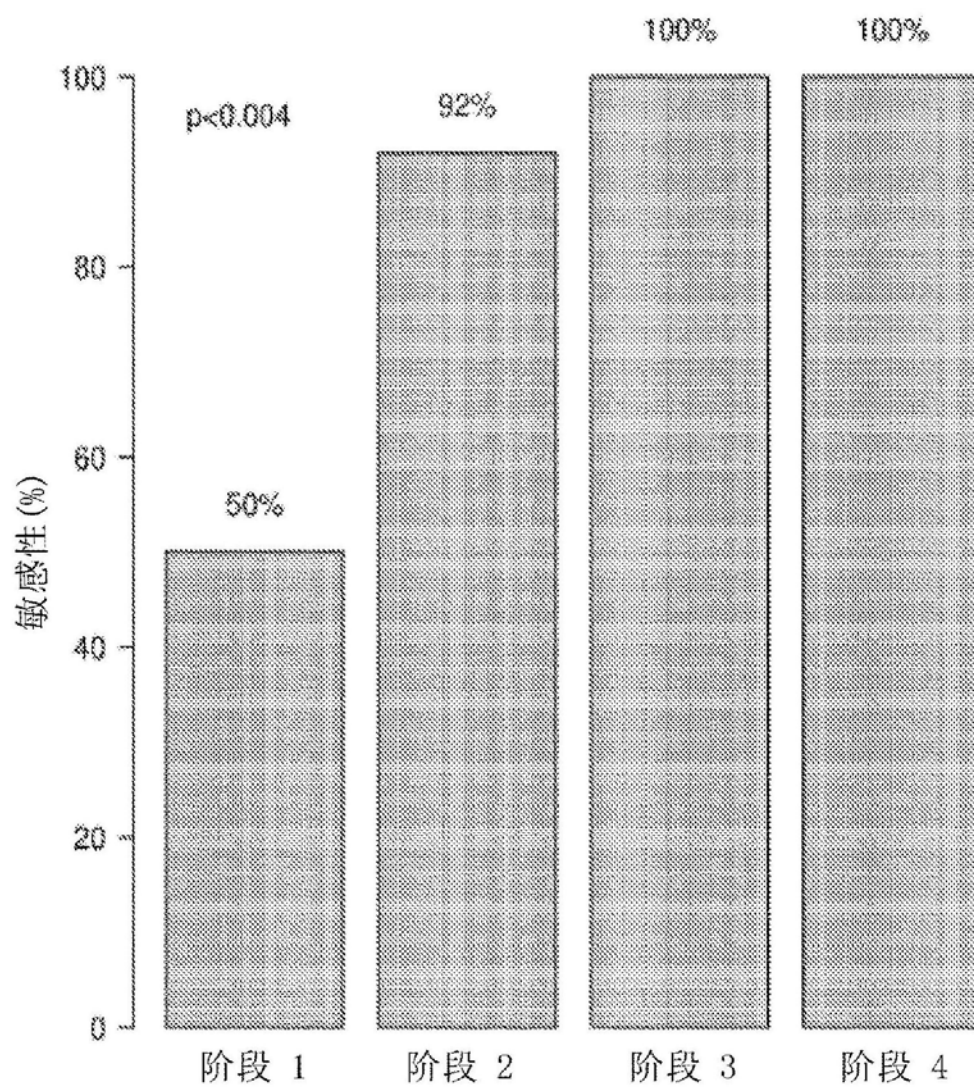


图4B

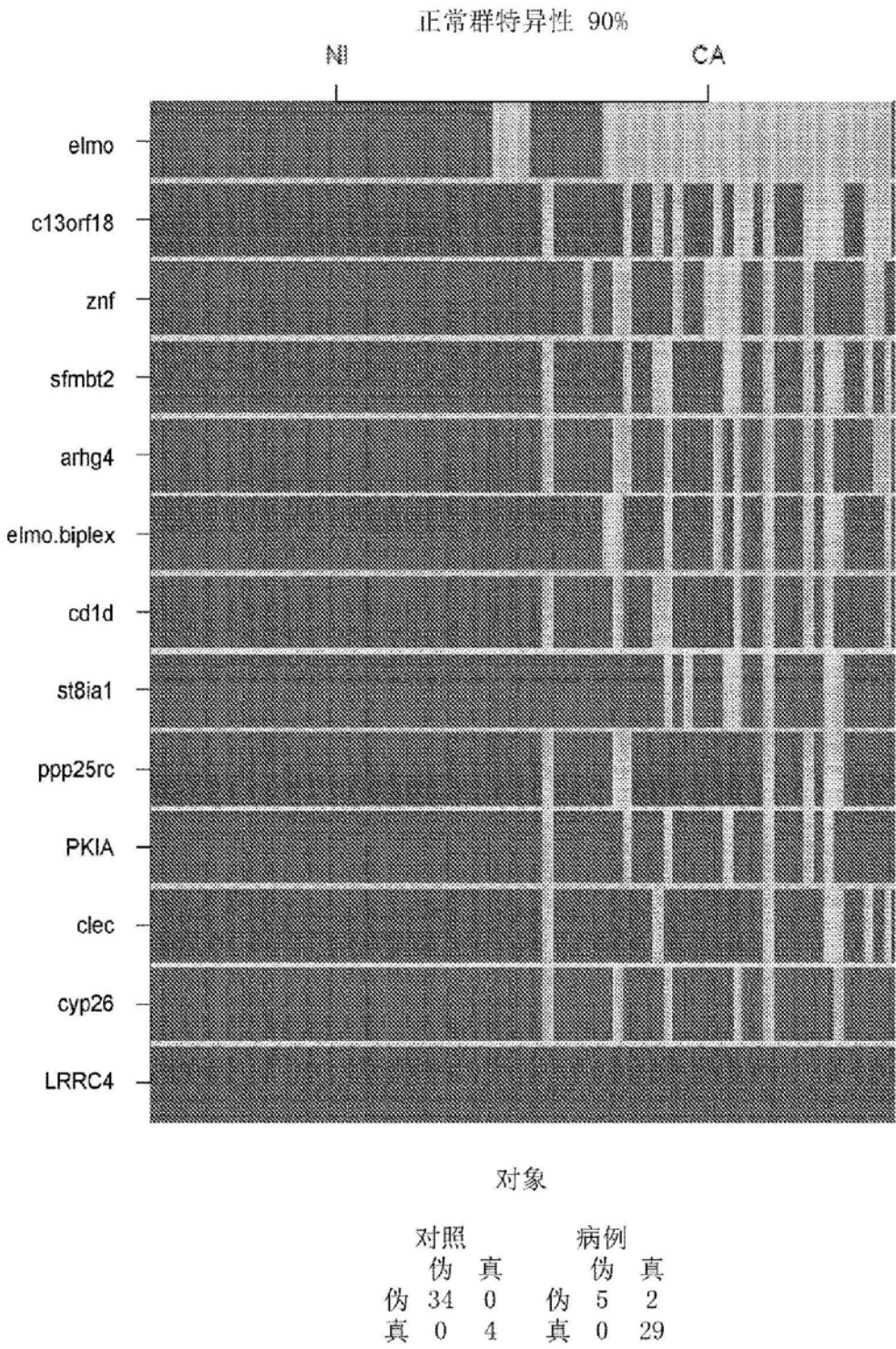
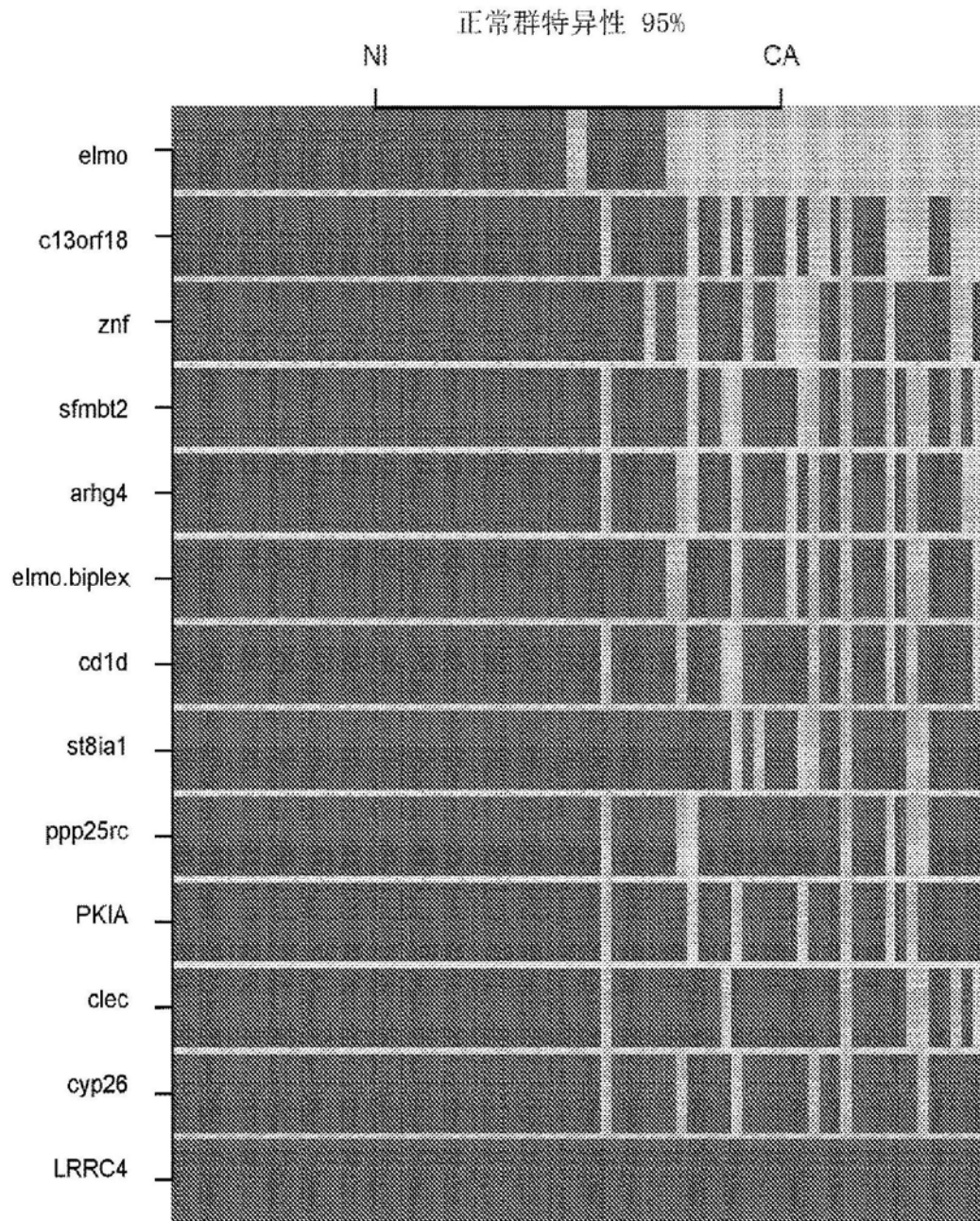


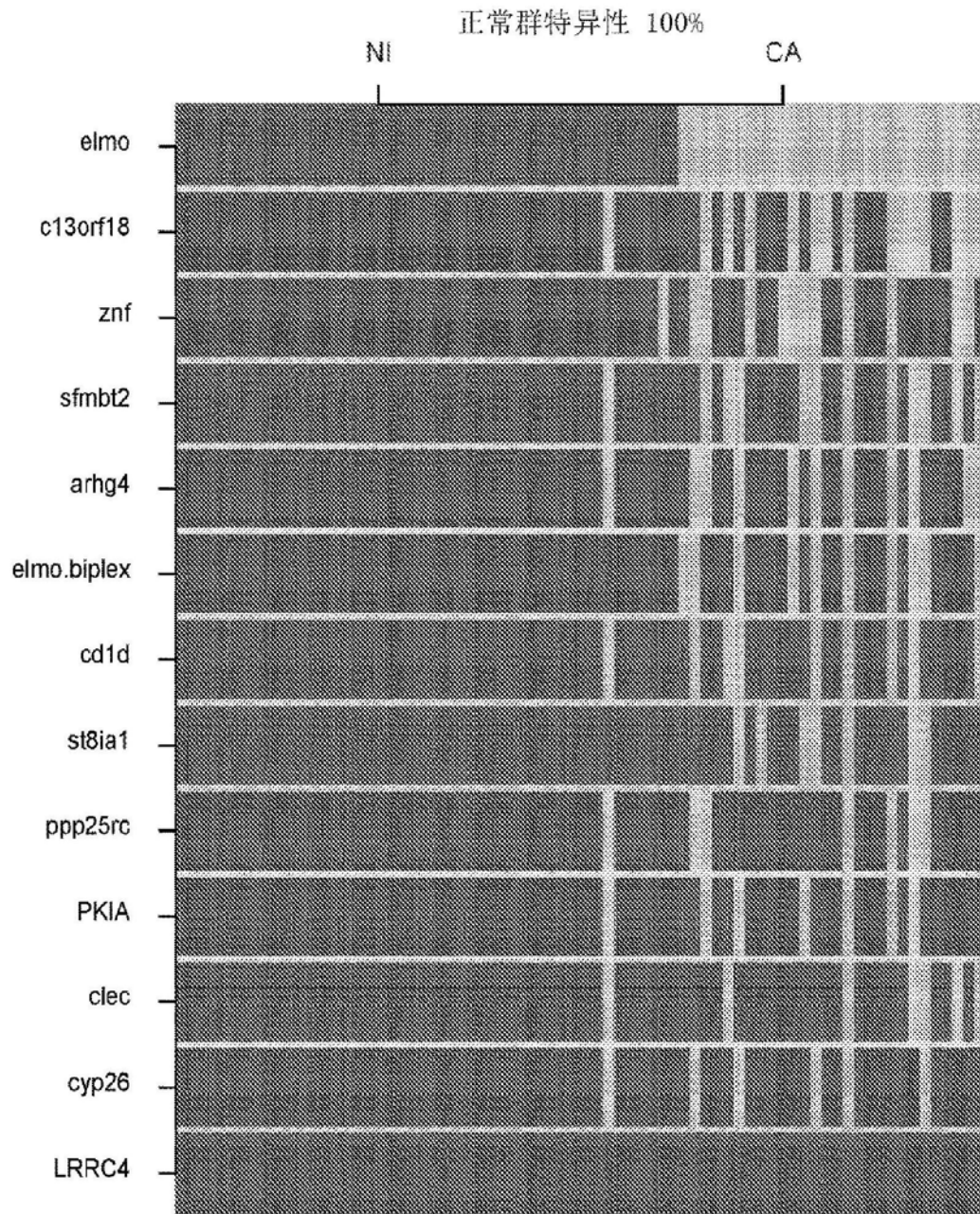
图5A



对象

对照			病例		
	伪	真		伪	真
伪	36	0	伪	5	2
真	0	2	真	0	29

图5B



对象

对照			病例		
伪	真		伪	真	
伪	38	0	伪	6	2
真	0	0	真	0	28

图5C

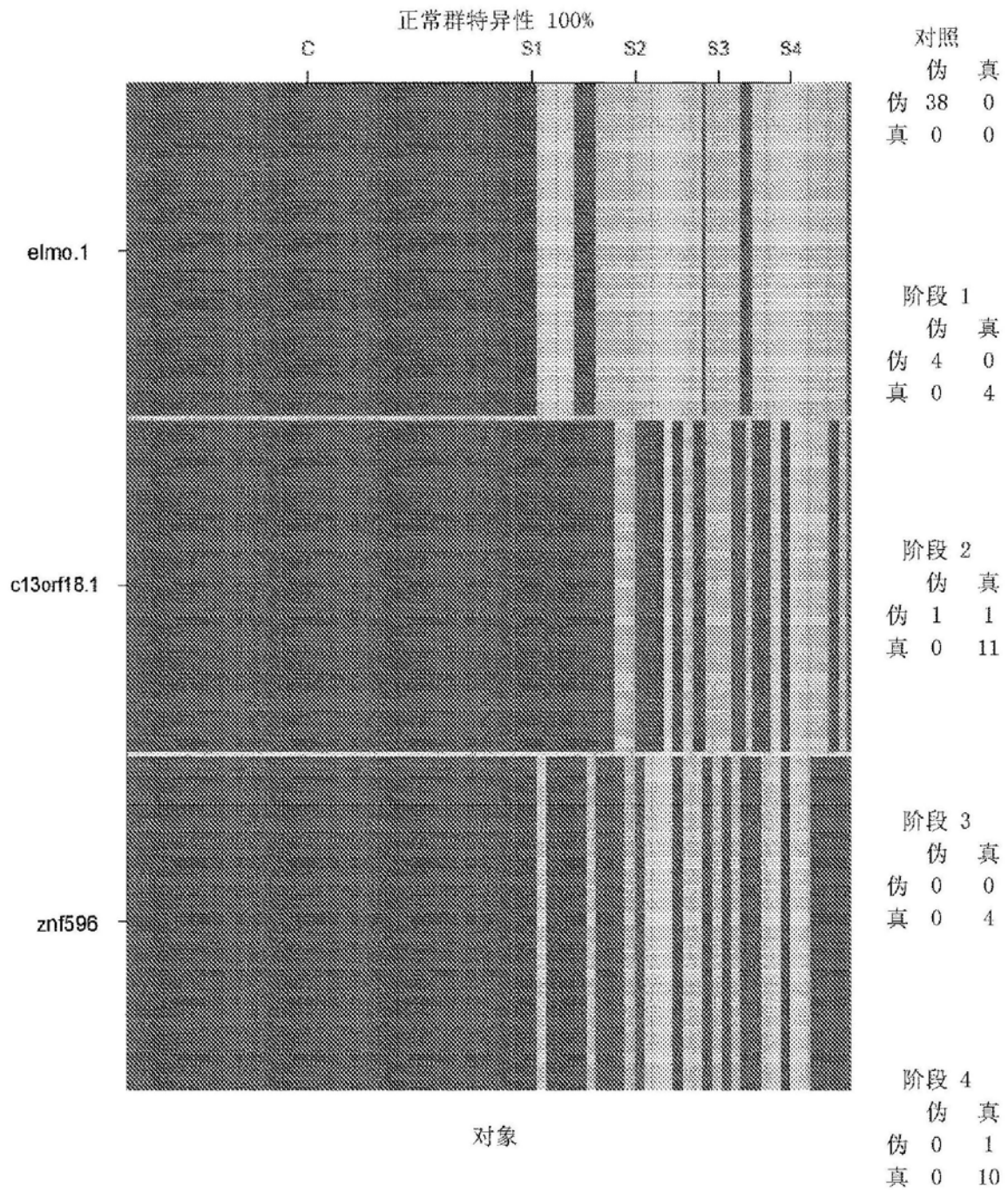


图6

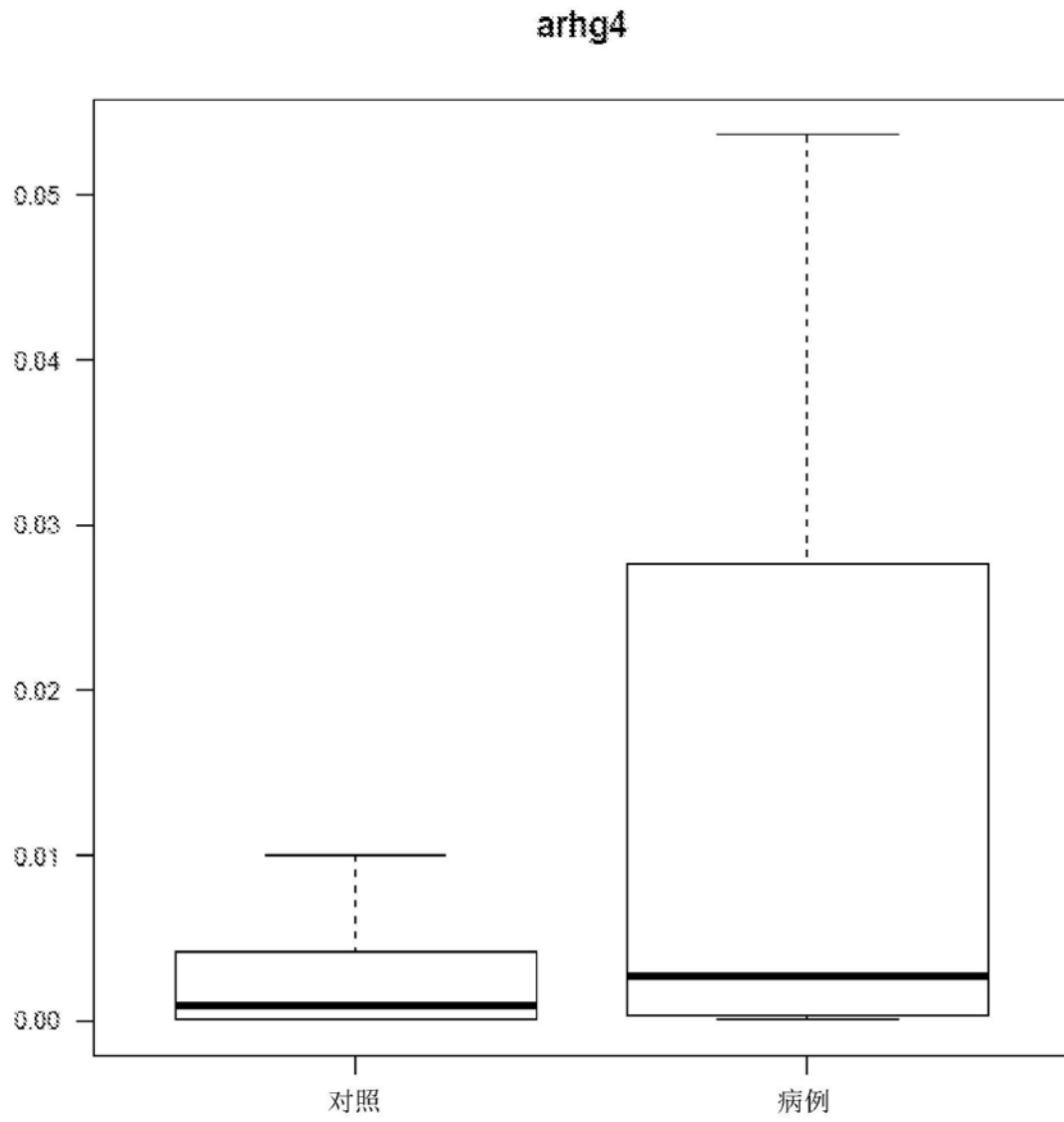


图7A

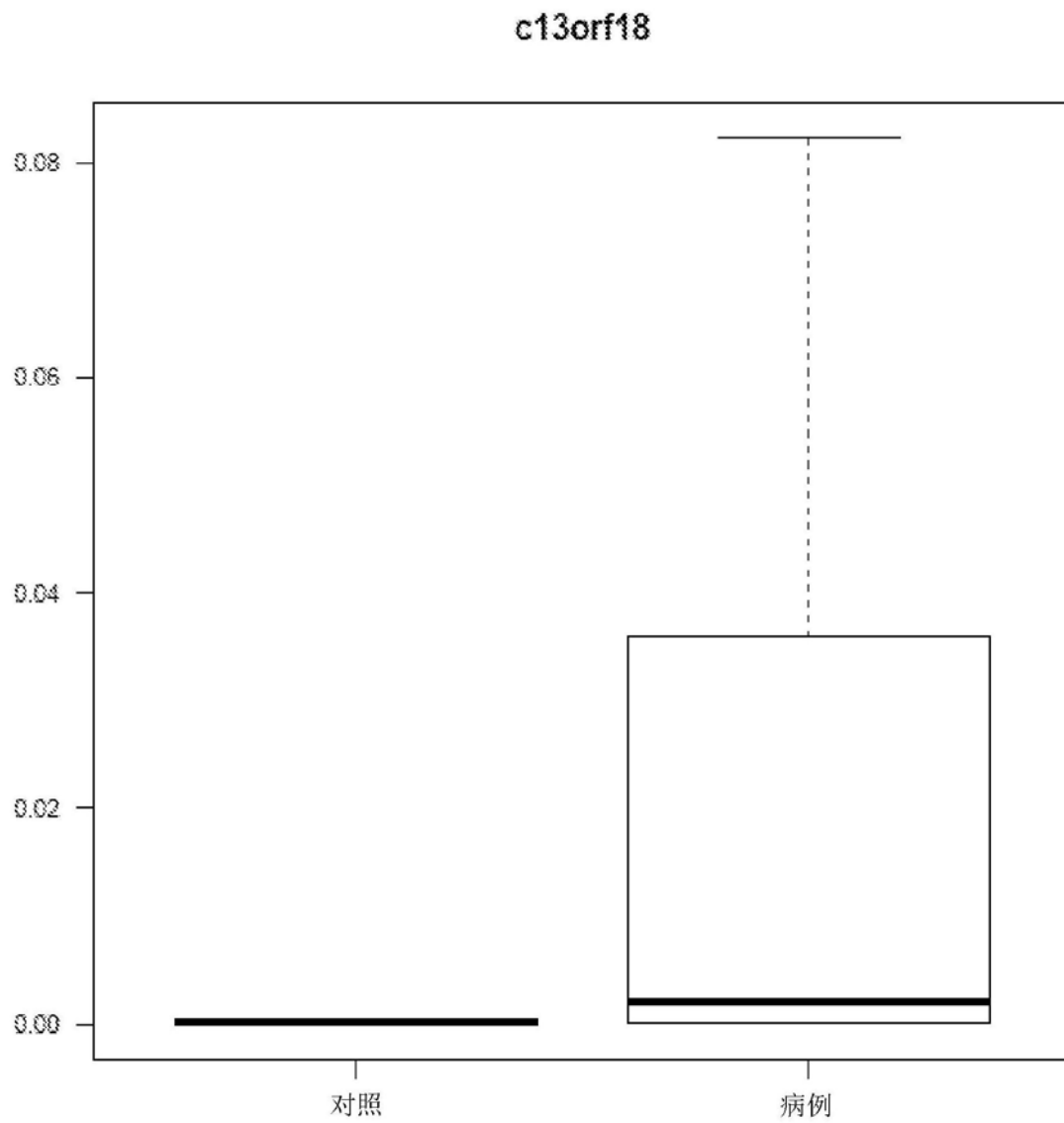


图7B

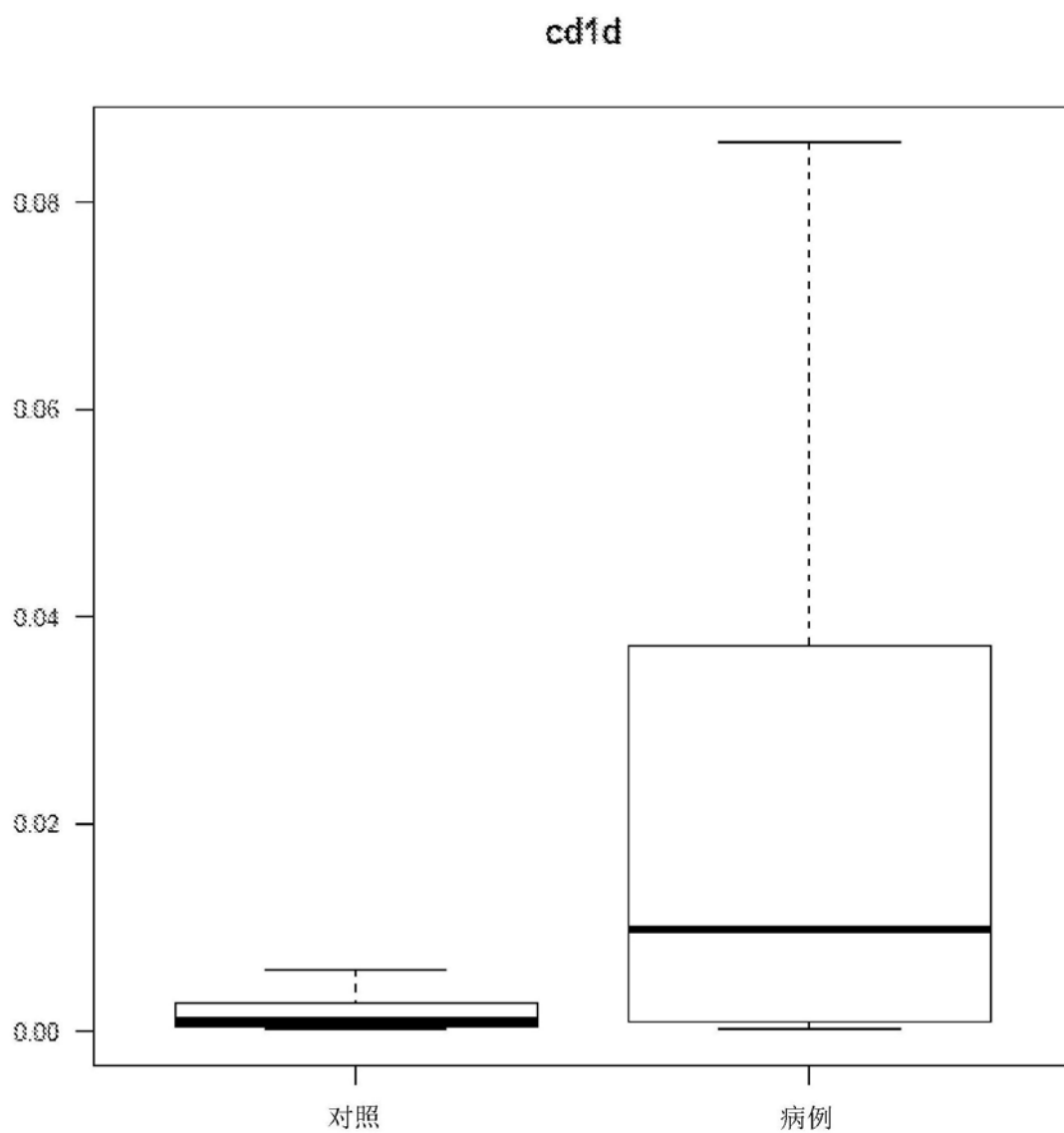


图7C

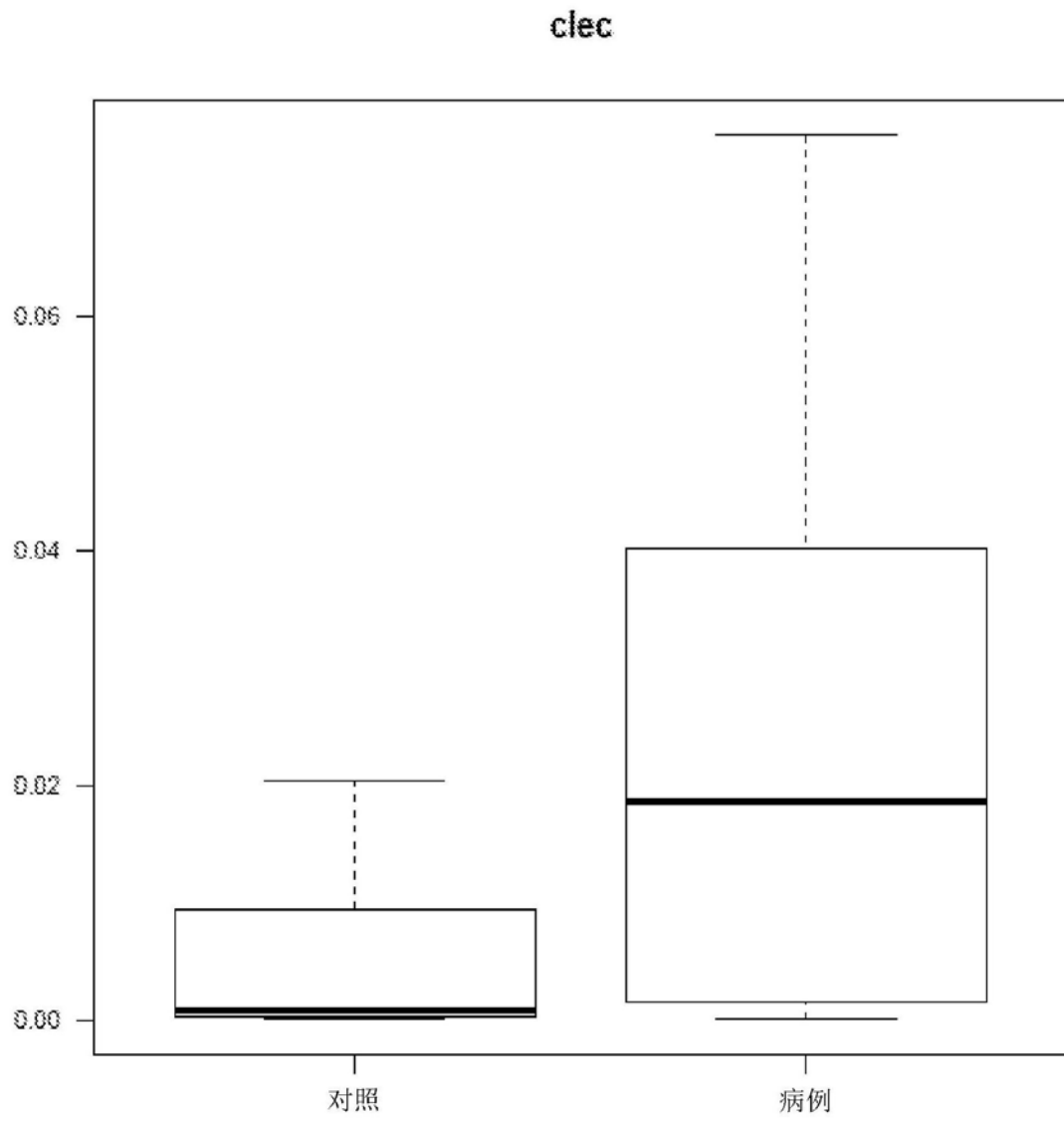


图7D

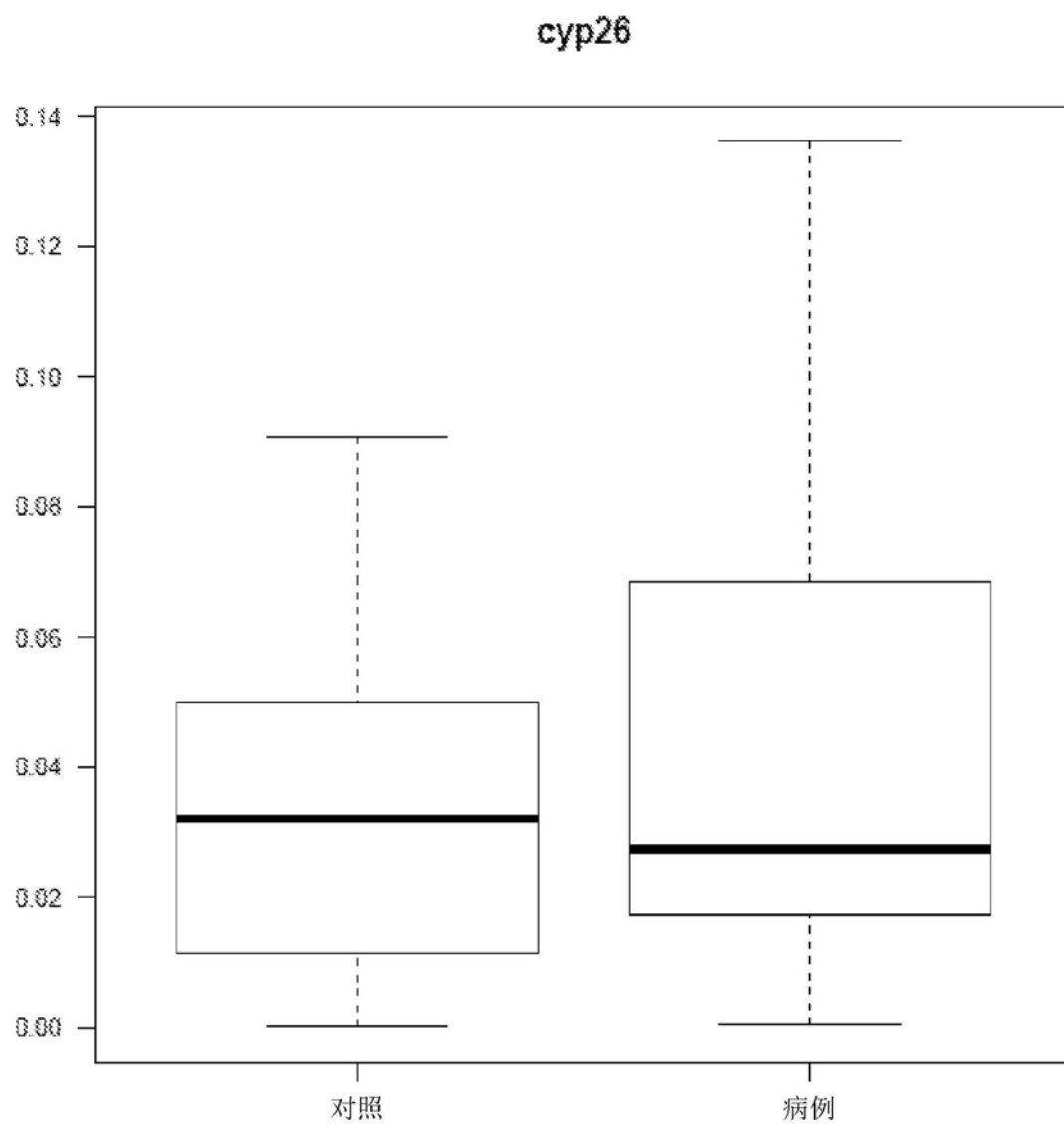


图7E

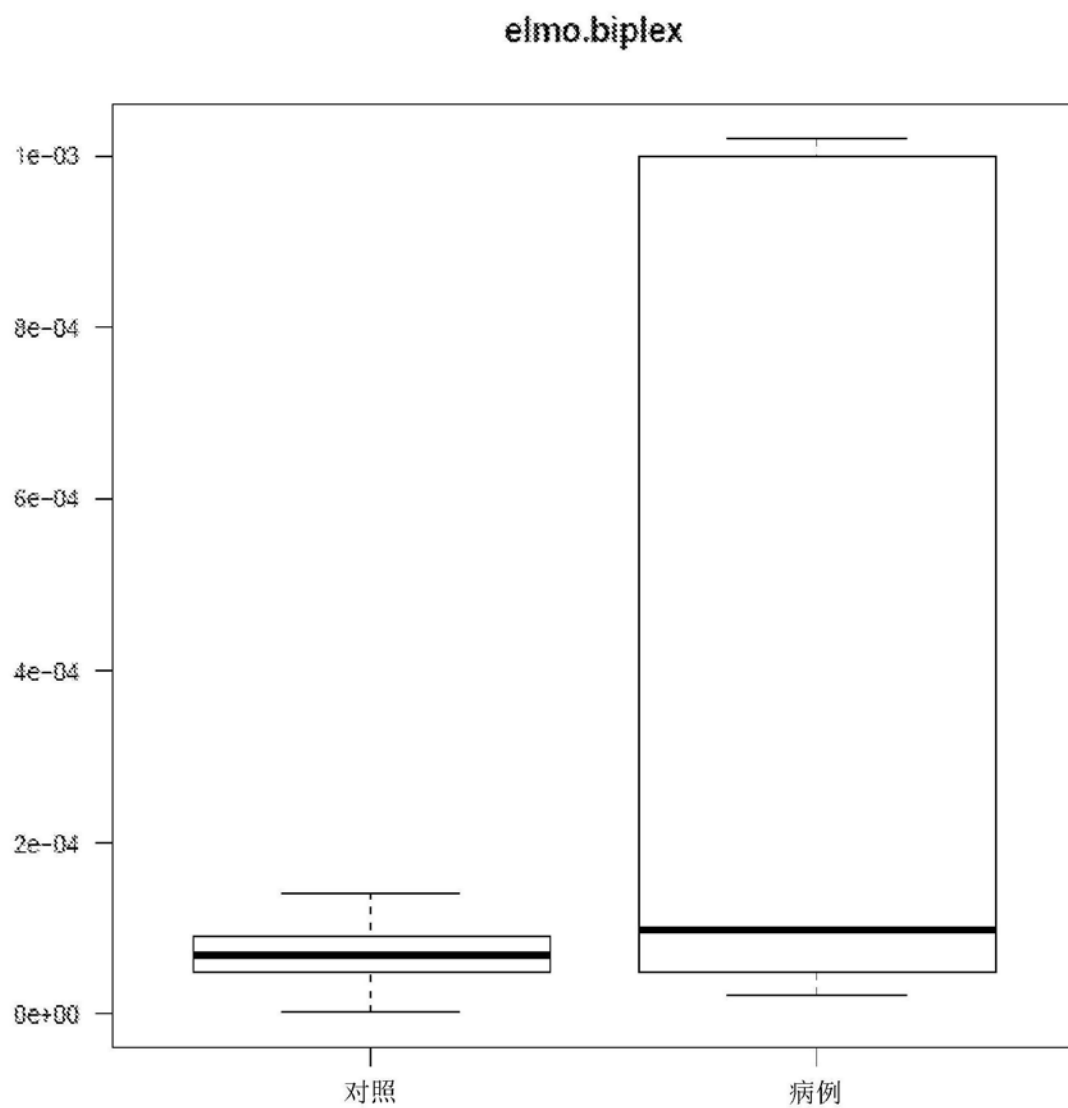


图7F

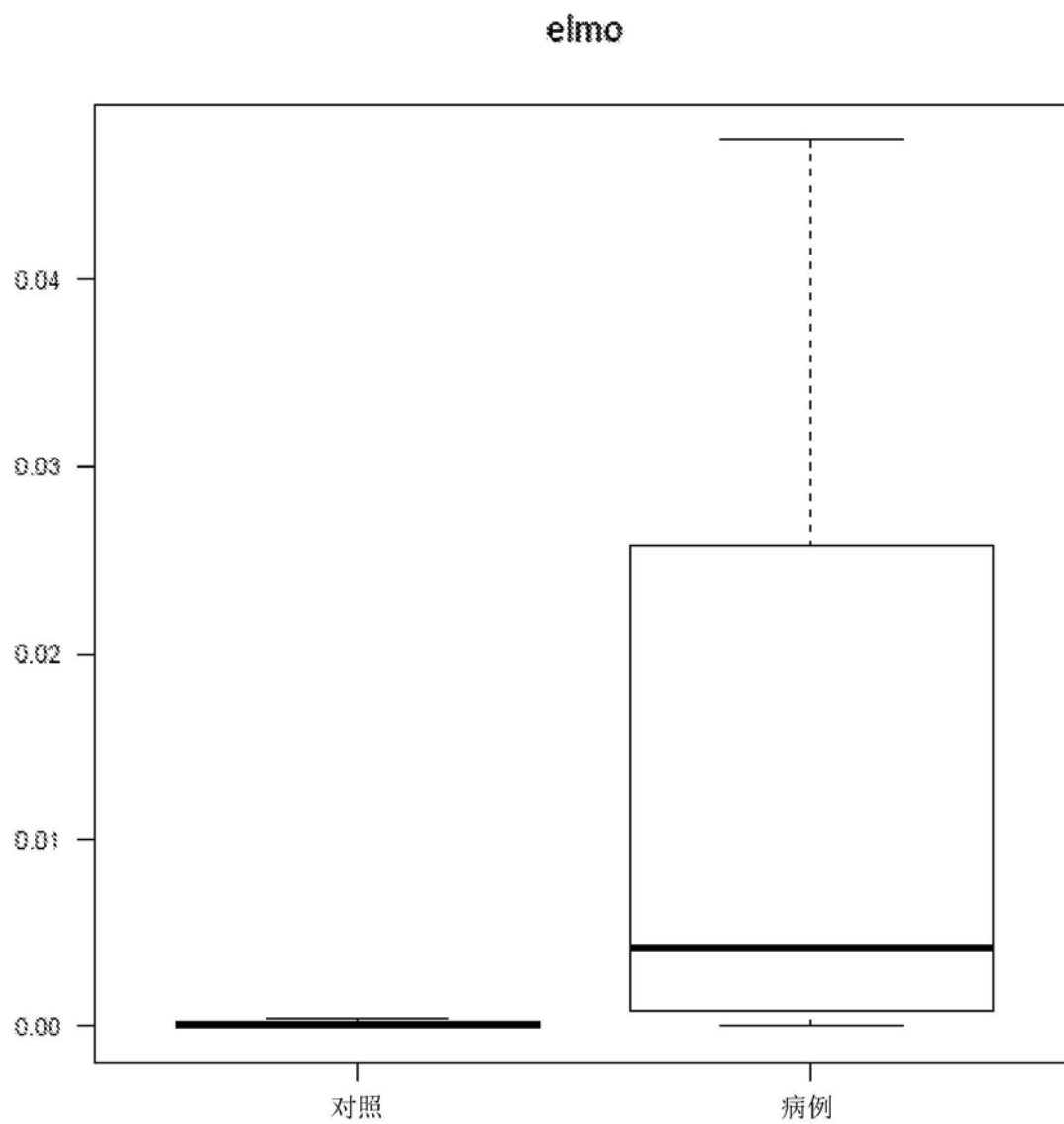


图7G

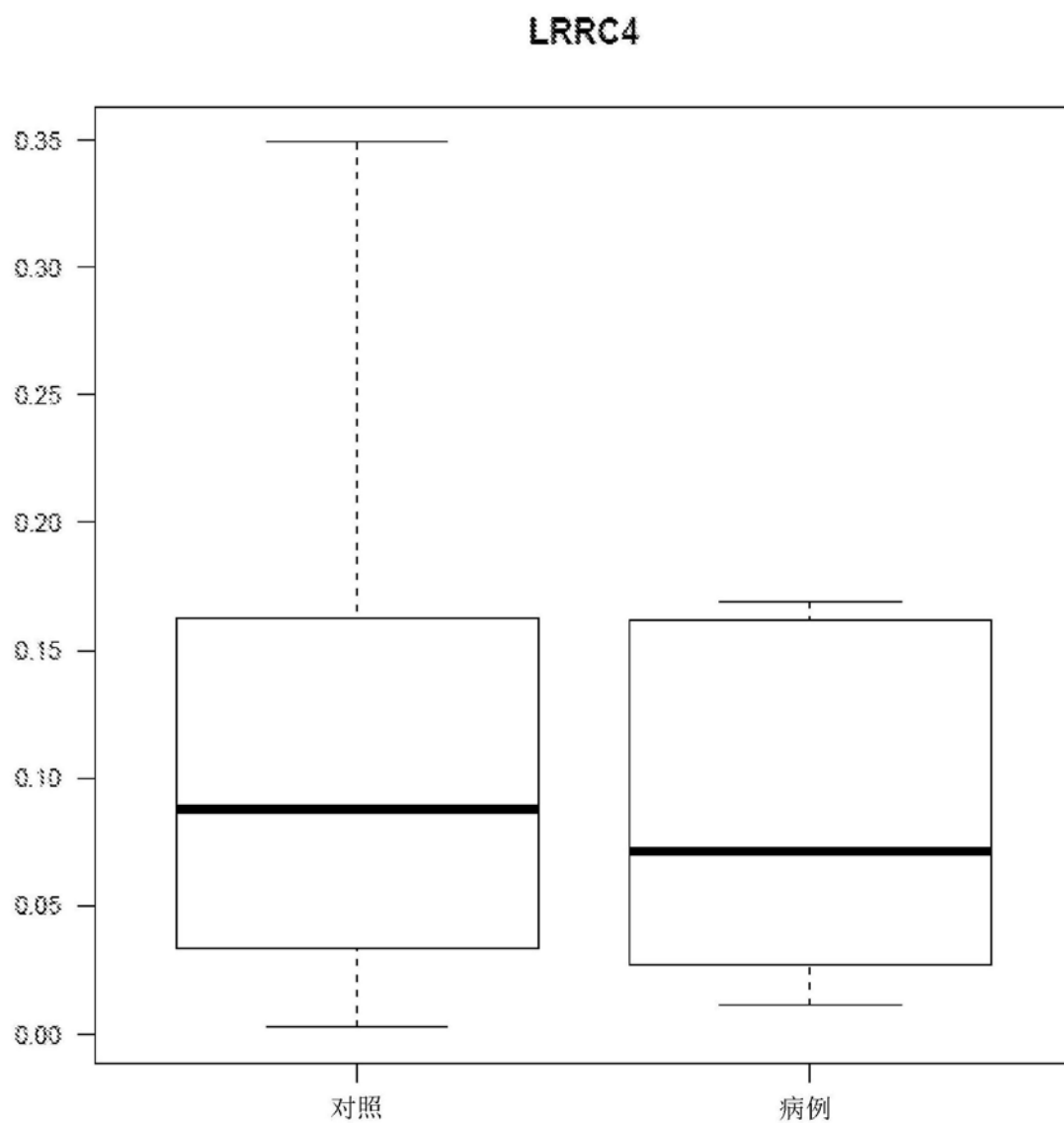


图7H

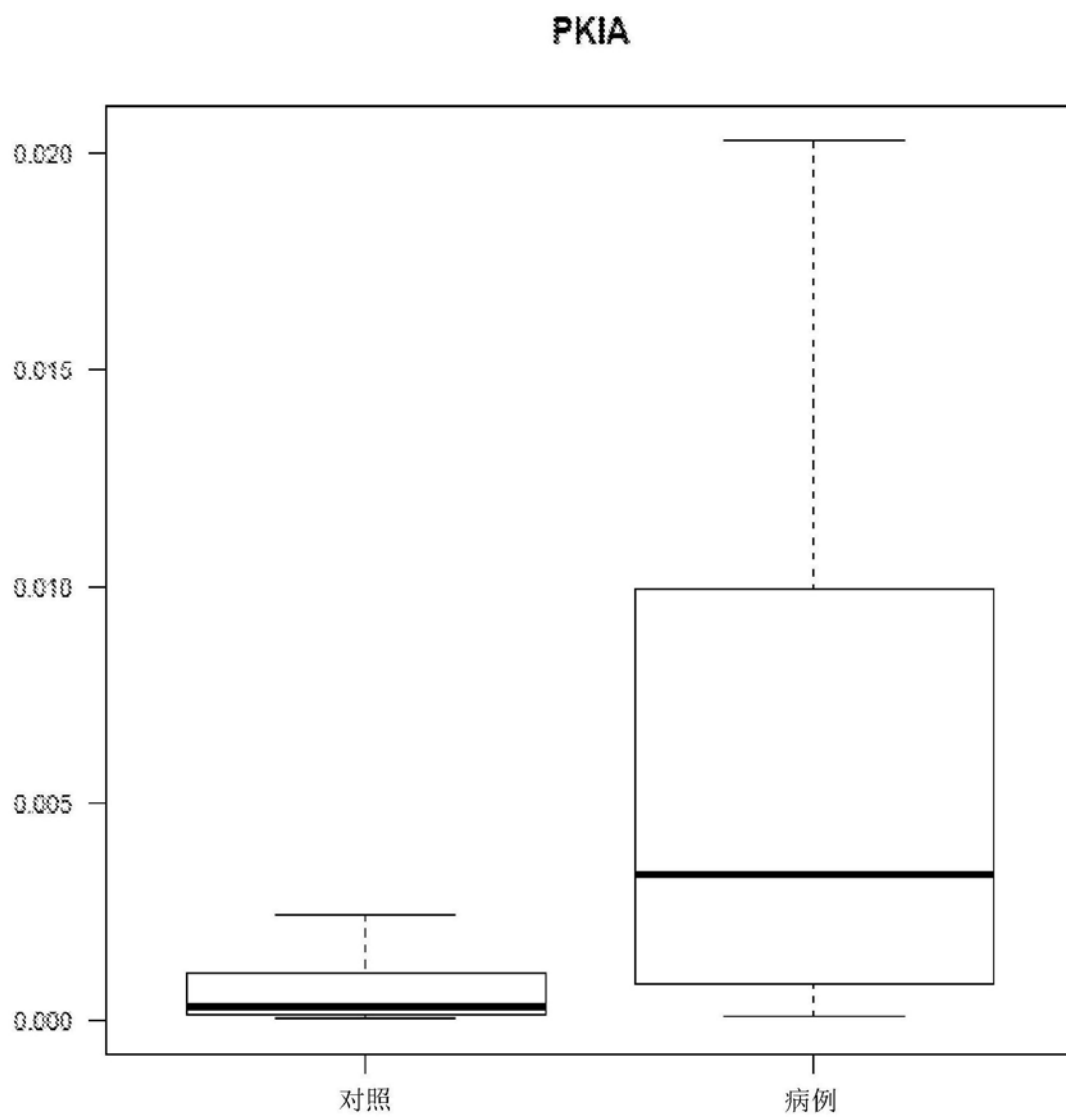


图7I

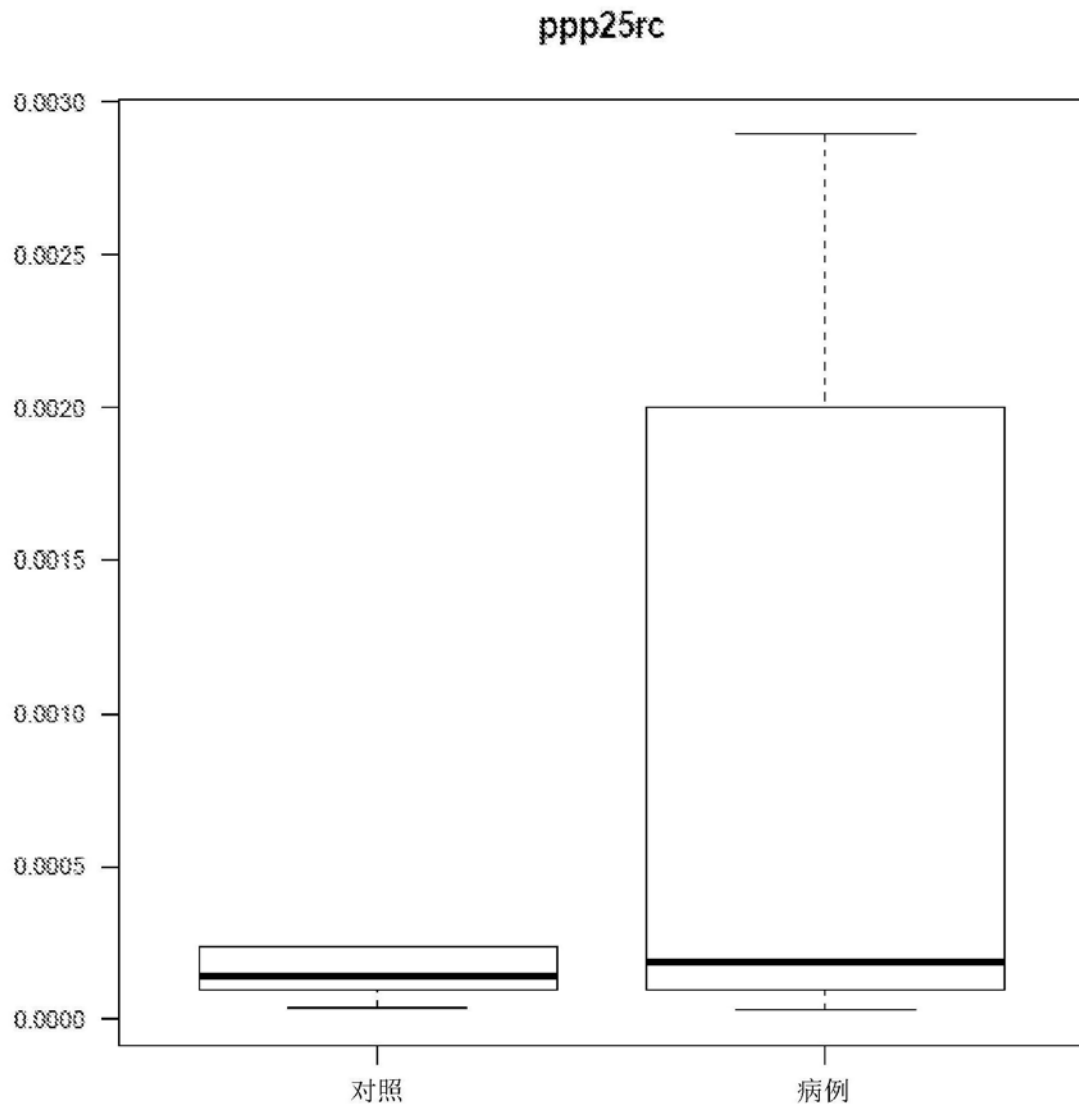


图7J

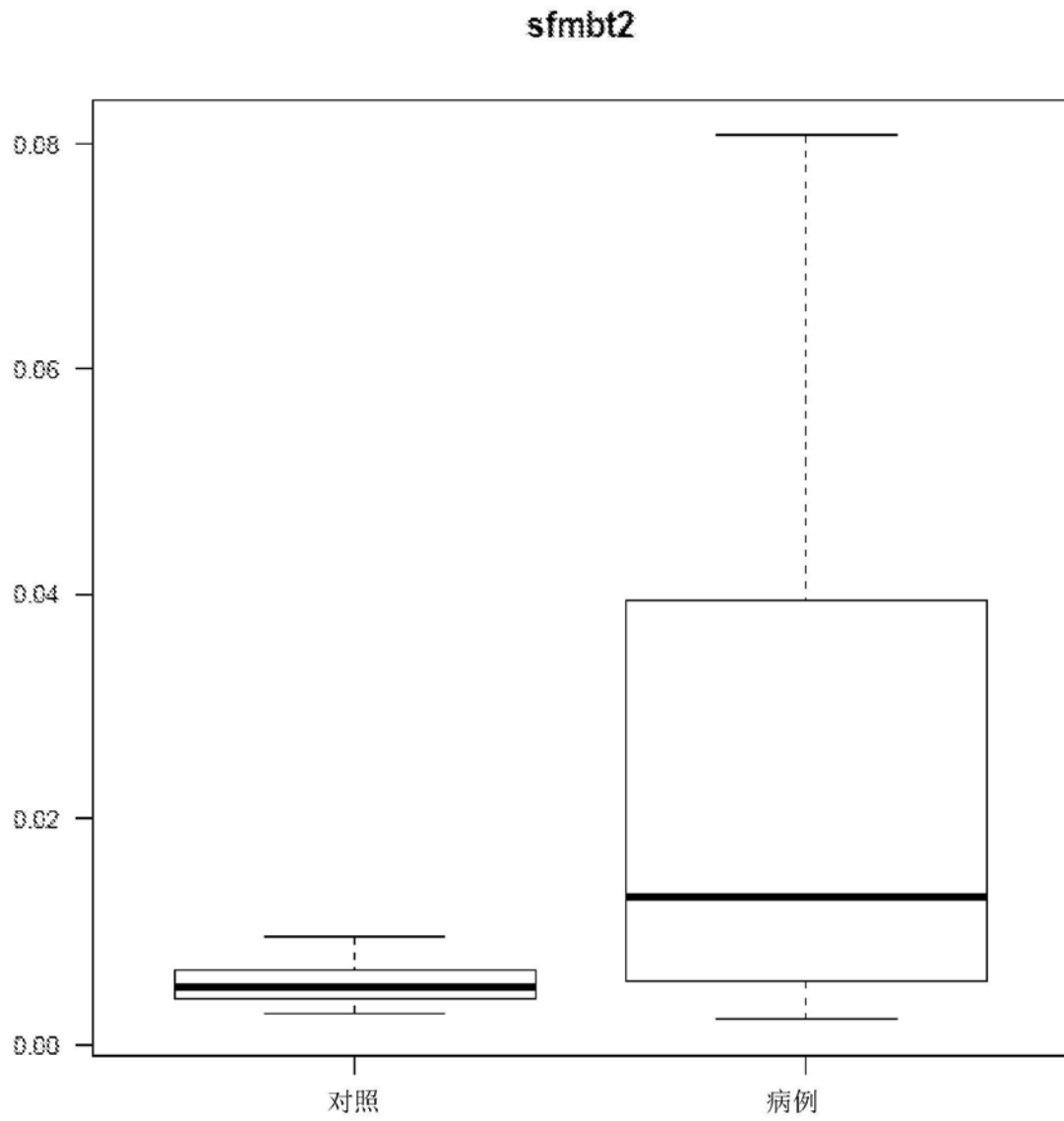


图7K

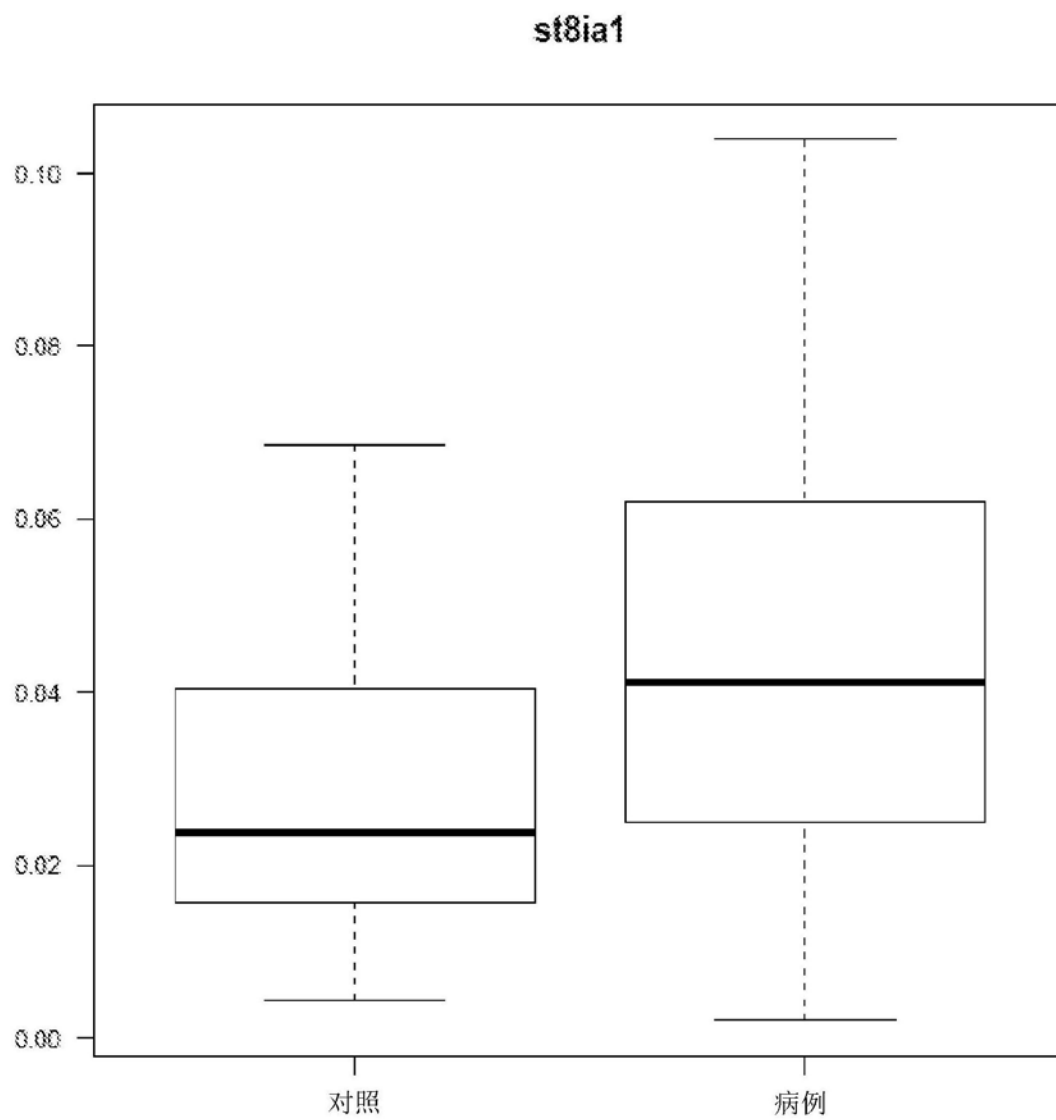


图7L

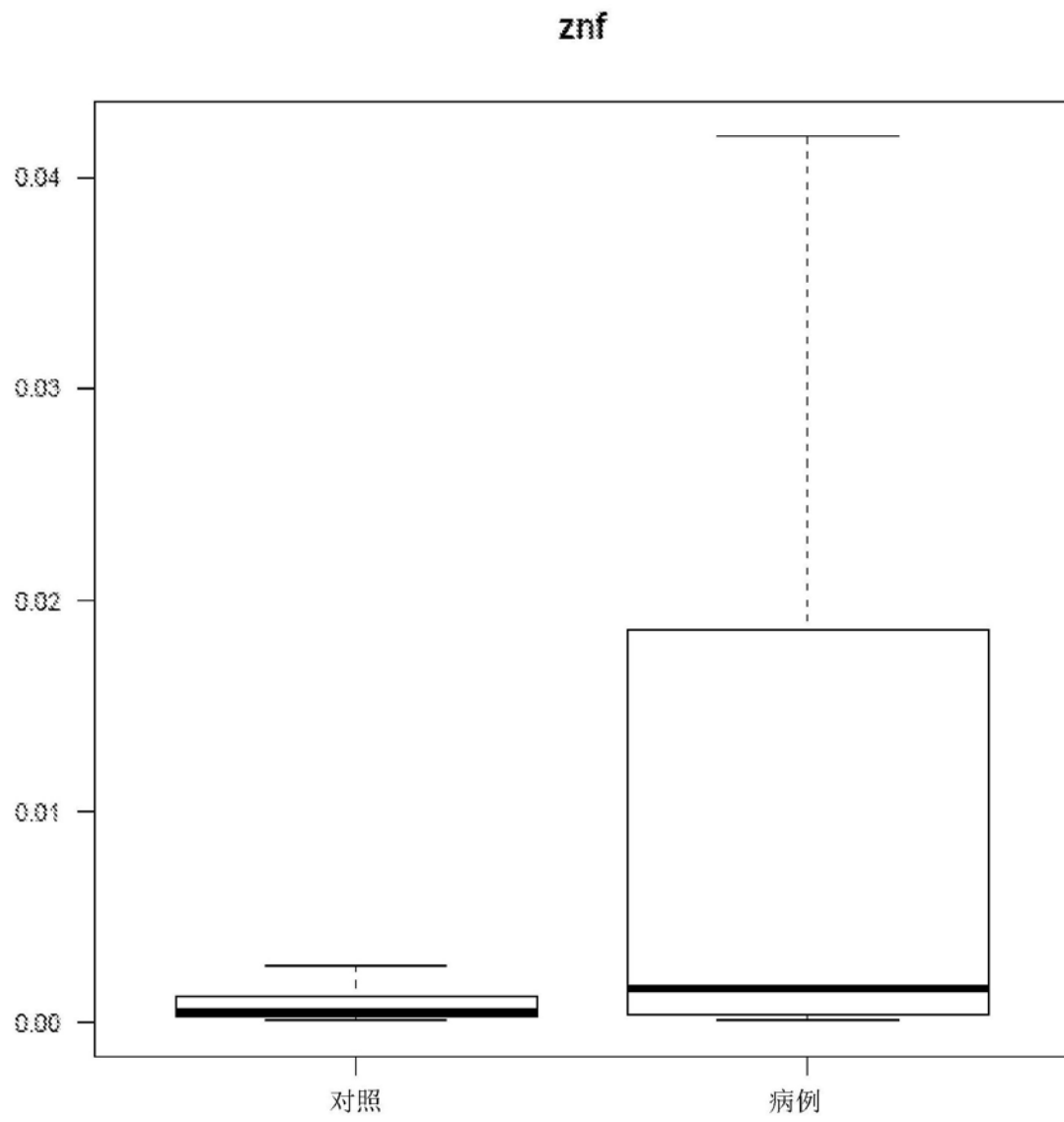


图7M