



(12) 发明专利申请

(10) 申请公布号 CN 118186078 A

(43) 申请公布日 2024.06.14

(21) 申请号 202410180308.2

(22) 申请日 2024.02.18

(71) 申请人 杭州翱锐基因科技有限公司

地址 310000 浙江省杭州市杭州经济技术
开发区6号大街260号18幢301室

(72) 发明人 余韶华 张琼 朱友杰 徐博
郑文渊

(74) 专利代理机构 杭州天昊专利代理事务所
(特殊普通合伙) 33283

专利代理师 向庆宁

(51) Int. Cl.

C12Q 1/6886 (2018.01)

C12N 15/11 (2006.01)

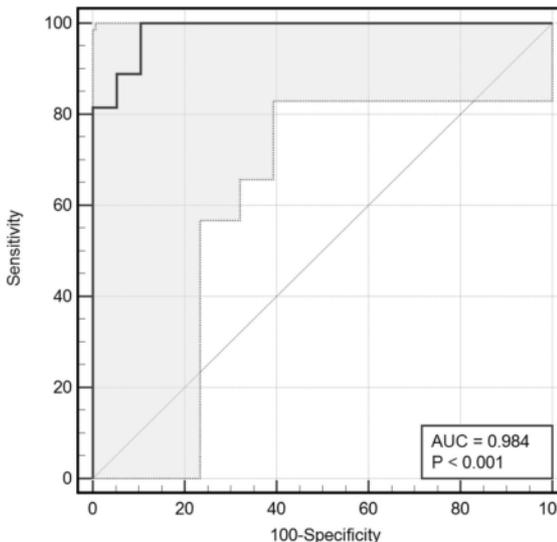
权利要求书2页 说明书31页 附图6页

(54) 发明名称

一种新型多靶点肺癌辅助诊断的标志物组合及其应用

(57) 摘要

本发明提供了一种新型多靶点肺癌诊断的组合标志物及其应用, 通过从肺癌、肺良性病变以及健康人样本的高深度全基因组亚硫酸盐测序(WGBS)数据中找到了系列新型的、能高效区分肺癌和肺良性病变患者的甲基化位点。通过对外周血中这些标志物的检测, 对于早期的肺癌能够有很好的诊断性能。同时, 在甲基化检测的基础上引入肺癌临床常用的3个蛋白标志物(CEA、CYFRA21-1和ProGRP), 多组学联合检测能进一步提高诊断的准确性。



1. 一种标志物组合用于制备肺癌辅助诊断试剂的用途,其特征在于,所述标志物组合包括Seq ID NO.1~Seq ID NO.18中的任意一种或多种所示的核苷酸序列或其完全互补序列的组合。

2. 如权利要求1所述的用途,其特征在于,所述标志物组合为选自Seq ID NO.1~Seq ID NO.18中的任意2种或任意6种所示的核苷酸序列或其完全互补序列的组合。

3. 如权利要求1所述的用途,其特征在于,所述标志物组合包括如Seq ID NO.7和/或Seq ID NO.18所示的核苷酸序列或其完全互补序列的组合。

4. 如权利要求1所述的用途,其特征在于,所述标志物组合包括如Seq ID NO.5、Seq ID NO.7、Seq ID NO.10、Seq ID NO.12、Seq ID NO.15和Seq ID NO.18中的任意一种或多种所示的核苷酸序列或其完全互补序列的组合。

5. 如权利要求1所述的用途,其特征在于,所述标志物组合为由Seq ID NO.1~Seq ID NO.18所示的核苷酸序列或其完全互补序列的组合。

6. 如权利要求1~5任一项所述的用途,其特征在于,所述标志物还包括蛋白标志物,所述蛋白标志物由CEA、CYFRA21-1和ProGRP组成。

7. 一种检测肺癌的引物组合,其特征在于,所述引物组合为选自如下表所示的18组引物、探针组合中的任意一组或多组:

组别	靶序列	正向引物	反向引物	探针
1	Seq ID NO.1	Seq ID NO.23	Seq ID NO.25	Seq ID NO.24
2	Seq ID NO.2	Seq ID NO.26	Seq ID NO.28	Seq ID NO.27
3	Seq ID NO.3	Seq ID NO.29	Seq ID NO.31	Seq ID NO.30
4	Seq ID NO.4	Seq ID NO.32	Seq ID NO.34	Seq ID NO.33
5	Seq ID NO.5	Seq ID NO.35	Seq ID NO.37	Seq ID NO.36
6	Seq ID NO.6	Seq ID NO.38	Seq ID NO.40	Seq ID NO.39
7	Seq ID NO.7	Seq ID NO.41	Seq ID NO.43	Seq ID NO.42
8	Seq ID NO.8	Seq ID NO.44	Seq ID NO.46	Seq ID NO.45
9	Seq ID NO.9	Seq ID NO.47	Seq ID NO.49	Seq ID NO.48
10	Seq ID NO.10	Seq ID NO.50	Seq ID NO.52	Seq ID NO.51
11	Seq ID NO.11	Seq ID NO.53	Seq ID NO.55	Seq ID NO.54
12	Seq ID NO.12	Seq ID NO.56	Seq ID NO.58	Seq ID NO.57
13	Seq ID NO.13	Seq ID NO.59	Seq ID NO.61	Seq ID NO.60
14	Seq ID NO.14	Seq ID NO.62	Seq ID NO.64	Seq ID NO.63
15	Seq ID NO.15	Seq ID NO.65	Seq ID NO.67	Seq ID NO.66
16	Seq ID NO.16	Seq ID NO.68	Seq ID NO.70	Seq ID NO.69
17	Seq ID NO.17	Seq ID NO.71	Seq ID NO.73	Seq ID NO.72
18	Seq ID NO.18	Seq ID NO.74	Seq ID NO.76	Seq ID NO.75

8. 一种用于检测肺癌的试剂盒,其特征在于,包括如权利要求7所述的引物和探针组合。

9. 一种用于检测肺癌的标记物组合,其特征在于,包括甲基化位点和蛋白标志物,所述甲基化位点为选自如序列列表Seq ID NO.1~Seq ID NO.18中的任意一种或多种;所述蛋白

标志物包括CEA、CYFRA21-1和ProGRP。

10.一种用于检测肺癌的标记物组合,其特征在于,包括Seq ID NO.5、Seq ID NO.7、Seq ID NO.10、Seq ID NO.12、Seq ID NO.15和Seq ID NO.18所示的核苷酸序列或其完全互补序列的组合。

一种新型多靶点肺癌辅助诊断的标志物组合及其应用

技术领域

[0001] 本发明涉及癌症检测领域,具体而言,涉及一种新型多靶点肺癌辅助诊断的标志物组合及其应用。

背景技术

[0002] 肺癌是全球最常见和最致命的癌症之一。Global cancer statistics 2020显示,全球约有220.7万肺癌新发病例,位居第二位;约有179.6万肺癌死亡病例,位居第一位(CA: a cancer journal for clinicians,2021,71(3):209-249)。中国肺癌发病数和死亡数分别占全球的37%和39.8%,肺癌防治是我国恶性肿瘤防控面临的重大挑战。近年来肺癌治疗取得了较大的进展,我国肺癌的5年生存率也从16.1%提高到19.7%,但仍有约75%的患者在诊断时处于肺癌晚期,错过了最佳根治性手术治疗时机(The Lancet Global Health, 2018,6(5):e555-e567)。早期诊断可显著提高肺癌患者预后生存,从肺癌患者术后5年生存率可以看出,I期患者术后5年生存率在77%~92%,而IIIA~IVA期患者仅为10%~36%,5年生存率存在显著差异(Chin J Tuberc Respir Dis,January 2023,Vol.46,No.1)。综上,提高肺癌生存率的关键在于早期发现、早期诊断和早期治疗,而早期诊断又是重中之重。

[0003] 肺部影像学检查方法主要包括胸部X线、胸部CT、磁共振显像(MRI)、正电子发射计算机断层显像(PET-CT)等,是临床上常用的肺癌早期诊断和诊断方法。胸部X线简便易行,但分辨率低,不易检出肺部小结节和纵隔、心包后等隐蔽部位的病灶,对早期肺癌的检出有一定的局限性(Jama,2021,325(10):962-970.)。胸部CT是当前肺癌诊断、分期、疗效评价及治疗后随访中最重要和最常用的影像检查方法,具有更高的分辨率,可检出肺部微小病灶和普通X线胸片难以显示部位的病变。其中,低剂量螺旋CT(low-dose computed tomography,LDCT)的诊断价值已被国家肺部诊断试验(NLST)、荷兰-比利时随机肺癌诊断试验(NELSON)和多中心意大利肺部检测试验(MILD)等研究证实。但是随着LDCT在国内外的广泛应用,很多问题逐渐突显出来,包括高危人群的定义标准不一、假阳性率过高、过度诊断及成本效益问题等。MRI和PET-CT虽然具有高灵敏度和特异度的特点,但是检测成本较高,临床推广十分困难。

[0004] 液体活检是一种微创方法,能够实时监测和早期识别从恶性病变转移到体液中的细胞或细胞产物的变化。同时,由于液体活检的微创性,可以预防组织活检带来的并发症。液体活检包括循环肿瘤细胞(CTCs)、循环肿瘤DNA(ctDNA)和其他肿瘤遗传物质,如细胞外囊泡(Evs)的富集和分离(Mol Cancer,2023.22(1):p.7)。ctDNA由于采集简单、信号特征丰富,已成为癌症早期检测的研究热点。ctDNA包含多种改变,包括甲基化、突变和拷贝数变化,可用于早期癌症检测。然而,与基因组改变(如突变和拷贝数变化)相比,ctDNA的甲基化改变发生得更早,呈现丰富的癌症特异性和组织特异性,并在体液中表现出显著的稳定性(Nat Rev Clin Oncol,2018.15(5):p.292-309)。Zhao等(Advanced Science,2023:2206518)开发了一套多重数字甲基化特异PCR(multiplex digital methylation-specific PCR,mdMSP)方法,基于研究团队前期发现的4个甲基化标志物(SOX17,CD01,

TAC1, 和HOXA7), 上述方法甲基化检测方法的最低检测限可以低至0.0005%, 对于肺癌的AUC可以达到0.86。这一研究结果证实了基于血液游离DNA甲基化进行肺癌早期诊断的可行性。

[0005] 虽然单一组学的检测技术已经得到广泛的研究和应用, 但是恶性肿瘤的发生涉及基因、表观遗传、转录组、微生物、蛋白和代谢等多个不同层次和不同维度的病理过程, 如果只对单一组学的特征进行分析, 会给靶点的筛选带来较大的局限性。多组学特征整合分析一方面能够相互印证, 增强检测结果的说服力; 另一方面也能够多维度、全方位地捕捉癌症早期信号, 提高检测灵敏度。在多组学单癌种检测方面, 2014年, 美国FDA批准了全球第一个通过粪便隐血和基因多靶点联合检测 (FIT-DNA) 用于结直肠癌诊断的产品Cologuard。该产品已被美国疾病预防工作组等多个权威组织推荐将其应用于适龄的无症状人群结直肠肿瘤早期诊断。作为Exact Science的核心产品, Cologuard利用了三种类型的生物标志物: NDRG4和BMP3基因的甲基化, KRAS基因的点突变和大便隐血中的血红蛋白。最终产品性能上, Cologuard表现出对结直肠癌92.3%的灵敏度, 对进展期腺瘤42.4%的灵敏度, 特异性为87% (N Engl J Med, 2014. 371 (2) : p. 187)。同时, 该司于2022年ESMO发布肝癌检测平台Oncoguard®Liver上线。该产品是通过检测外周血游离DNA甲基化及蛋白标志物进行肝癌早期诊断。2022年, Exact Sciences在Clinical Gastroenterology and Hepatology杂志公布了Oncoguard®Liver的最新研究成果, Chalasani等 (Clin Gastroenterol Hepatol, 2022. 20 (1) : p. 173-182. e7) 的这项研究共计纳入540例患者的血液样本 (136例病例+404例对照), 检测包含3种DNA甲基化标志物 (HOXA1、TSPYL5和B3GALT6) 和1种蛋白标志物 (AFP)。分析结果显示Oncoguard®Liver检测肝癌的灵敏度为88%, 特异性为87%, 目前正在开展前瞻性验证试验。以上研究表明, 以ctDNA甲基化为基础的多组学联合检测技术, 可以作为癌症检测的有利工具。

[0006] 目前国内获批的肺癌甲基化检测产品有透景生物的人SHOX2、RASSF1A基因甲基化DNA检测试剂盒 (PCR荧光法) 以及艾克伦生物的SHOX2/RASSF1A/PTGER4基因甲基化检测试剂盒 (PCR-荧光探针法), 肺癌产品有凯保罗的七种肺癌相关抗体检测试剂盒 (酶联免疫法), 但是上述产品都只是基于甲基化或者患者自身抗体等单一组学的检测, 临床上对于肺癌诊断的准确性还不够。

发明内容

[0007] 针对现有技术中存在的问题, 本发明提供了一种新型多靶点肺癌辅助诊断的组合标志物和试剂盒及其在肺癌诊断领域的应用。通过从肺癌及其癌旁组织的高深度全基因组甲基化 (WGBS) 数据中找到了系列新型的、能高效区分肺癌和良性病变患者的甲基化位点, 可用于肺癌的高效检测; 将该新型基因甲基化位点与临床上常用的蛋白标志物组合 (CEA、CYFRA21-1和ProGRP), 能进一步提高AUC值, 提高肺癌诊断的灵敏度和特异性。

[0008] 通过检测本发明提供的系列新型甲基化位点的甲基化水平, 并结合血清样本中的CEA、CYFRA21-1和ProGRP蛋白水平, 能够更加灵敏、特异的区分肺癌患者和非肺癌患者。通过检测临床肺癌及非肺癌患者样本的数据显示, 本发明提供的组合可以非常有效地区分肺癌患者和非肺癌患者, AUC值最高能达到0.984, 灵敏度达到89.5%, 特异性达到99.0%。

[0009] 甲基化, 是指从活性甲基化合物上将甲基催化转移到其他化合物的过程, 可形成

各种甲基化合物,或是对某些蛋白质或核酸等进行化学修饰形成甲基化产物。在生物系统内,甲基化是经酶催化的,这种甲基化涉及重金属修饰、基因表达的调控、蛋白质功能的调节以及核糖核酸加工。

[0010] CpG是胞嘧啶(C)一磷酸(p)一鸟嘌呤(G)的缩写,在哺乳动物中CpG以两种形式存在:一种是分散于DNA序列中;另一种呈现高度聚集状态,人们称之为CpG岛(CpG island),在正常组织里,70%~90%散在的CpG是被甲基修饰的,而与之相反,大小为100-1000bp左右且富含CpG二核苷酸的CpG岛,则往往非甲基化的。CpG岛经常出现在真核生物的编码基因的调控区,在其它地方出现时会由于CpG中的C易被甲基化而形成5'-甲基胞嘧啶,脱氨基后形成尿嘧啶,经DNA复制或扩增后转化为胸腺嘧啶(T),由于T本身就会存在于DNA中,因此不易被修复,所以被淘汰。故CpG在基因组中是以岛的形式分布的。

[0011] 发生甲基化修饰后的C不会被亚硫酸氢盐转化,未被甲基化修饰的胞嘧啶(C)会被亚硫酸氢盐转化为尿嘧啶(U),碱基互补配对的时候U会和A配对,C会和G配对。而且理论上,肿瘤患者的这些位点全都会被甲基化修饰,一般肿瘤患者的这些位点甲基化修饰是普遍现象,不会出现只有部分修饰的情况,基因组上距离较近的CpG位点会同时被甲基化或是去甲基化,这就意味着只能将一簇CpG位点作为整体进行分析。

[0012] 胞嘧啶5号位的甲基化修饰,是广泛存在于真核细胞生物中的一种DNA修饰方式,DNA上的甲基化修饰不仅在生物的生长发育中起到重要作用,在细胞的癌变过程中也起着非常重要的作用。由于和胞嘧啶具有一样的碱基互补配对性质,5-甲基胞嘧啶无法通过一代测序或高通量测序的方式来直接测定。目前最常用的检测5-甲基胞嘧啶(C)的方法是将待测DNA通过亚硫酸氢盐转化,在碱性水解后,未甲基化的胞嘧啶被转化为尿嘧啶(U),而5-甲基胞嘧啶(C)则不会发生转化。尿嘧啶在碱基互补配对的时候会与腺嘌呤互补配对,区别于胞嘧啶与鸟嘌呤的互补配对,因此通过对经过亚硫酸氢盐处理的DNA进行检测时,借助PCR技术,就可以确定未发生转化的胞嘧啶(C),从而确定在原始DNA分子中,哪些胞嘧啶发生了甲基化。因此甲基化位点标志物是整个序列基因组上距离较近的CpG位点的所有C都发生甲基化,所有的C都是甲基化位点。

[0013] 本发明还引入了蛋白标志物,以提高检测的准确性,蛋白标志物由癌胚抗原(CEA)、细胞角蛋白19的可溶性片段(CYFRA21-1)和胃泌素释放肽前体(ProGRP)组成。因为这些蛋白标志物都是肺癌临床上常用的检测指标,可以提供相应的检测信息。

[0014] 一方面,本发明提供了一种标志物用于制备肺癌检测试剂的用途,所述标志物为选自如Seq ID NO.1~Seq ID NO.18中的任意一种或多种核苷酸序列或其完全互补序列的组合。

[0015] 本发明通过对肺癌组织、对应的癌旁组织以及正常人cfDNA样本进行高深度全基因组亚硫酸盐测序,筛选出区分肺癌和肺良性病变患者的甲基化位点,然后进一步通过大量的肺癌和肺良性病变患者血浆样本验证,最终发现并确定了在肺癌患者中发生甲基化异常的18个靶序列。经临床验证,本发明提供的18个新型靶序列,在肺癌和肺良性病变患者中甲基化程度差异显著。

[0016] 18个靶序列为双链DNA,存在互补序列,可以理解的是,本发明提供的甲基化位点靶序列可以是其正义链,也可以是其反义链。

[0017] 在一些方式中,所述标志物为选自如序列表Seq ID NO.1~Seq ID NO.18中的1

种、2种、3种、4种、5种、6种、7种、8种、9种、10种、11种、12种、13种、14种、15种、16种、17种或18种核苷酸序列或其完全互补序列的组合。

[0018] 本发明找到了在肺癌和肺良性病变患者中甲基化程度具有显著差异的18个甲基化位点,并从中进一步精确筛选得到数量为2个、6个的甲基化位点的组合,实现了在较少的标志物检测情况下对于肺癌的高效检出。

[0019] 进一步的,所述标志物组合为选自Seq ID N0.1~Seq ID N0.18中的任意2种或任意6种所示的核苷酸序列或其完全互补序列的组合。

[0020] 对任意2种或任意6种的甲基化位点的组合进行了筛选,得到了11组高AUC值、高灵敏度和特异性的甲基化位点,其中包括2种甲基化位点的有3组,所述标志物组合可以选自下述3组中的任意一组所示的核苷酸序列或其完全互补序列的组合:

[0021] 1、Seq ID N0.3、Seq ID N0.7

[0022] 2、Seq ID N0.7、Seq ID N0.18

[0023] 3、Seq ID N0.13、Seq ID N0.18

[0024] 其中包括6种甲基化位点的有8组,所述标志物组合可以选自下述8组中的任意一组所示的核苷酸序列或其完全互补序列的组合:

[0025] 4、Seq ID N0.5、Seq ID N0.7、Seq ID N0.10、Seq ID N0.12、Seq ID N0.15和Seq ID N0.18

[0026] 5、Seq ID N0.2、Seq ID N0.7、Seq ID N0.8、Seq ID N0.11、Seq ID N0.14和Seq ID N0.15

[0027] 6、Seq ID N0.3、Seq ID N0.8、Seq ID N0.10、Seq ID N0.11、Seq ID N0.13和Seq ID N0.17

[0028] 7、Seq ID N0.1、Seq ID N0.4、Seq ID N0.12、Seq ID N0.14、Seq ID N0.15和Seq ID N0.17

[0029] 8、Seq ID N0.5、Seq ID N0.7、Seq ID N0.8、Seq ID N0.12、Seq ID N0.13和Seq ID N0.17

[0030] 9、Seq ID N0.1、Seq ID N0.3、Seq ID N0.6、Seq ID N0.8、Seq ID N0.10和Seq ID N0.11

[0031] 10、Seq ID N0.3、Seq ID N0.7、Seq ID N0.13、Seq ID N0.8、Seq ID N0.9和Seq ID N0.16

[0032] 11、Seq ID N0.3、Seq ID N0.4、Seq ID N0.5、Seq ID N0.10、Seq ID N0.13、Seq ID N0.15

[0033] 优选的,所述标志物为包含Seq ID N0.7和/或Seq ID N0.18的核苷酸序列或其完全互补序列的组合。Seq ID N0.7或Seq ID N0.18对应的甲基化位点的AUC值、检测灵敏度和特异性都明显高于其他16个甲基化位点。

[0034] 优选的,所述标志物为包含Seq ID N0.7和/或Seq ID N0.18的2种、6种或18种核苷酸序列或其完全互补序列的组合。

[0035] 进一步地,所述标志物为包括Seq ID N0.5、Seq ID N0.7、Seq ID N0.10、Seq ID N0.12、Seq ID N0.15和Seq ID N0.18所示的核苷酸序列或其完全互补序列的组合。

[0036] 经大量临床肺癌及肺良性病样本验证,本发明从上述18个靶序列中,找到6个能够

特别灵敏、特异地区分肺癌患者、肺良性病患者、良性对照个体的全新的高甲基化位点,分别具有如序列表Seq ID NO.5、Seq ID NO.7、Seq ID NO.10、Seq ID NO.12、Seq ID NO.15和Seq ID NO.18所示的序列。通过检测肺癌及肺良性病临床血浆样本的数据显示,仅仅采用这6个甲基化位点进行肺癌检测,其AUC值就能达到0.918,灵敏度达到85.2%,特异性达到99.0%;实验数据表明,该6个甲基化位点的肺癌诊断性能与18个甲基化位点组合的性能相近,而且检测过程更简便高效,最具应用前景。

[0037] 进一步地,所述标志物还包括蛋白标志物,所述蛋白标志物由癌胚抗原(CEA)、细胞角蛋白19的可溶性片段(CYFRA21-1)和胃泌素释放肽前体(ProGRP)组成

[0038] 由于技术的局限性,单独通过甲基化位点检测癌症的性能有一定的天花板,结合基因组学、表观遗传学、蛋白质组学等多个类别的多组学已成为发展趋势。已有多个临床研究表明,多组学标志物检测的敏感性和特异性优于单一组学标志物。因此在本发明提供的甲基化位点的基础上,结合蛋白水平检测,进一步提高肺癌检测的AUC值,提高检测灵敏度和特异性。

[0039] 本发明提供的蛋白标志物CEA、CYFRA21-1和ProGRP在肺癌病人血清中的水平,与肺良性病变病人血清中蛋白标志物的水平比较有着较显著的差异。研究证明,通过检测本发明提供的系列新型甲基化位点的甲基化水平,并结合血清样本中的CEA、CYFRA21-1和ProGRP,能够更加准确的区分肺癌患者和非癌个体。通过检测临床肺癌及肺良性病样本的数据显示,本发明提供的组合可以非常有效地区分肺癌和非癌的患者,AUC值最高能达到0.984,灵敏度达到89.5%,特异性达到99.0%。

[0040] 另一方面,本发明提供了一种检测肺癌的引物组合,所述引物组合为下表所示的18组引物、探针组合中的任意一组或多组:

[0041]

组别	靶序列	正向引物	反向引物	探针
1	Seq ID NO.1	Seq ID NO.23	Seq ID NO.25	Seq ID NO.24
2	Seq ID NO.2	Seq ID NO.26	Seq ID NO.28	Seq ID NO.27
3	Seq ID NO.3	Seq ID NO.29	Seq ID NO.31	Seq ID NO.30
4	Seq ID NO.4	Seq ID NO.32	Seq ID NO.34	Seq ID NO.33
5	Seq ID NO.5	Seq ID NO.35	Seq ID NO.37	Seq ID NO.36
6	Seq ID NO.6	Seq ID NO.38	Seq ID NO.40	Seq ID NO.39
7	Seq ID NO.7	Seq ID NO.41	Seq ID NO.43	Seq ID NO.42
8	Seq ID NO.8	Seq ID NO.44	Seq ID NO.46	Seq ID NO.45
9	Seq ID NO.9	Seq ID NO.47	Seq ID NO.49	Seq ID NO.48
10	Seq ID NO.10	Seq ID NO.50	Seq ID NO.52	Seq ID NO.51
11	Seq ID NO.11	Seq ID NO.53	Seq ID NO.55	Seq ID NO.54
12	Seq ID NO.12	Seq ID NO.56	Seq ID NO.58	Seq ID NO.57
13	Seq ID NO.13	Seq ID NO.59	Seq ID NO.61	Seq ID NO.60
14	Seq ID NO.14	Seq ID NO.62	Seq ID NO.64	Seq ID NO.63
15	Seq ID NO.15	Seq ID NO.65	Seq ID NO.67	Seq ID NO.66
16	Seq ID NO.16	Seq ID NO.68	Seq ID NO.70	Seq ID NO.69
17	Seq ID NO.17	Seq ID NO.71	Seq ID NO.73	Seq ID NO.72

18	Seq ID NO.18	Seq ID NO.74	Seq ID NO.76	Seq ID NO.75
----	--------------	--------------	--------------	--------------

[0042] 在一些方式中,所述引物组合为选自表2所示的引物、探针组合中的任意2种或6种。

[0043] 进一步地,所述引物组合为如下表所示的6组引物、探针组合:

组别	靶序列	正向引物	反向引物	探针
4	Seq ID NO.5	Seq ID NO.35	Seq ID NO.37	Seq ID NO.36
10	Seq ID NO.7	Seq ID NO.41	Seq ID NO.43	Seq ID NO.42
12	Seq ID NO.10	Seq ID NO.50	Seq ID NO.52	Seq ID NO.51
13	Seq ID NO.12	Seq ID NO.56	Seq ID NO.58	Seq ID NO.57
15	Seq ID NO.15	Seq ID NO.65	Seq ID NO.67	Seq ID NO.66
18	Seq ID NO.18	Seq ID NO.74	Seq ID NO.76	Seq ID NO.75

[0045] 进一步地,所述引物组合中还包含内参基因的引物和探针,所述内参基因COL2A1具有如序列表Seq ID NO.19所示的序列,其正向引物具有如序列表Seq ID NO.20所示的序列,反向引物具有如序列表Seq ID NO.22所示的序列,探针具有如序列表Seq ID NO.21所示的序列。

[0046] 再一方面,本发明提供了一种用于检测肺癌的试剂盒,所述试剂盒包括如上所述的任一种引物和探针组合。

[0047] 进一步地,所述试剂盒或芯片还包括用于检测蛋白的试剂或材料,所述蛋白包括CEA、CYFRA21-1和ProGRP。

[0048] 再一方面,本发明提供了一种用于检测肺癌的标记物组合,所述标记物组合包括甲基化位点和蛋白标志物,所述甲基化位点为选自如序列表Seq ID NO.1~Seq ID NO.18中的任意一种或多种;所述蛋白标志物包括CEA、CYFRA21-1和ProGRP。

[0049] 在一些方式中,所述标志物组合中的甲基化位点为包含Seq ID NO.7和/或Seq ID NO.18的2种、6种或18种核苷酸序列或其完全互补序列的组合。

[0050] 在一些方式中,所述标志物组合中的甲基化位点为包括Seq ID NO.3和Seq ID NO.7所示的核苷酸序列或其完全互补序列的组合,或包括Seq ID NO.7和Seq ID NO.18所示的核苷酸序列或其完全互补序列的组合,或包括Seq ID NO.13和Seq ID NO.18所示的核苷酸序列或其完全互补序列的组合,或包括Seq ID NO.5、Seq ID NO.7、Seq ID NO.10、Seq ID NO.12、Seq ID NO.15和Seq ID NO.18所示的核苷酸序列或其完全互补序列的组合,或包括Seq ID NO.2、Seq ID NO.7、Seq ID NO.8、Seq ID NO.11、Seq ID NO.14和Seq ID NO.15所示的核苷酸序列或其完全互补序列的组合,或包括Seq ID NO.3、Seq ID NO.8、Seq ID NO.10、Seq ID NO.11、Seq ID NO.13和Seq ID NO.17所示的核苷酸序列或其完全互补序列的组合,或包括Seq ID NO.1、Seq ID NO.4、Seq ID NO.12、Seq ID NO.14、Seq ID NO.15和Seq ID NO.17所示的核苷酸序列或其完全互补序列的组合,或包括Seq ID NO.5、Seq ID NO.7、Seq ID NO.8、Seq ID NO.12、Seq ID NO.13和Seq ID NO.17所示的核苷酸序列或其完全互补序列的组合,或包括Seq ID NO.1、Seq ID NO.3、Seq ID NO.6、Seq ID NO.8、Seq ID NO.10和Seq ID NO.11所示的核苷酸序列或其完全互补序列的组合,或包括Seq ID NO.3、Seq ID NO.7、Seq ID NO.8、Seq ID NO.9、Seq ID NO.13和Seq ID NO.16所示的核苷酸序列或其完全互补序列的组合,或包括Seq ID NO.3、Seq ID NO.4、Seq ID

N0.5、Seq ID N0.10、Seq ID N0.13和Seq ID N0.15所示的核苷酸序列或其完全互补序列的组合,或包括Seq ID N0.1~Seq ID N0.18所示的核苷酸序列或其完全互补序列的组合。

[0051] 在一些方式中,所述标志物组合中的甲基化位点为包括Seq ID N0.5、Seq ID N0.7、Seq ID N0.10、Seq ID N0.12、Seq ID N0.15和Seq ID N0.18所示的核苷酸序列或其完全互补序列的组合。

[0052] 再一方面,本发明提供了一种用于检测肺癌的标记物组合,所述标记物组合包括Seq ID N0.5、Seq ID N0.7、Seq ID N0.10、Seq ID N0.12、Seq ID N0.15和Seq ID N0.18所示的核苷酸序列或其完全互补序列的组合。

[0053] 通过本发明提供的甲基化位点进行体外检测肺癌的方法包括以下步骤:

[0054] 1) 分离待测生物样本中的基因组DNA或血浆游离DNA;

[0055] 2) 检测所述甲基化位点或者甲基化位点组合的甲基化状态;

[0056] 3) 通过所述目标基因甲基化位点状态,进行生物样品状态的判断,并实现对肺癌的体外检测。

[0057] 通过本发明提供的甲基化位点和蛋白组合进行体外检测肺癌的方法包括以下步骤:

[0058] 1) 分离待测生物样本中的基因组DNA或血浆游离DNA和血清;

[0059] 2) 检测所述甲基化位点或者甲基化位点组合的甲基化状态和血清CEA、CYFRA21-1和ProGRP蛋白水平;

[0060] 3) 通过所述目标基因甲基化位点状态以及蛋白标志物水平,进行生物样品状态的判断,并实现对肺癌的体外检测。

[0061] 在一些方式中,所述方法还包括以下步骤:

[0062] 1) 分离待测生物样本的血清和血浆,提取待测生物样本的血浆游离DNA;

[0063] 2) 使用试剂处理步骤1)得到的DNA样品,使5位未甲基化的胞嘧啶碱基转化为尿嘧啶,转化为尿嘧啶以后的碱基在杂交能力上不同于5位未甲基化的胞嘧啶,并且是可以检测的;

[0064] 3) 将步骤2)处理过的DNA样品与聚合酶链式反应体系组合,聚合酶链式反应体系包含以下一种等几种组分:DNA聚合酶、所述目标靶序列的引物或者引物组合、相对应的探针或探针组合、聚合酶链式反应缓冲液,发生聚合酶链式反应后,产生扩增产物;

[0065] 4) 用被荧光标记的一种探针或几种探针的组合检测扩增产物,如果探针和扩增产物结合,则可以产生荧光信号;如果探针无法和扩增产物结合,则不能产生荧光信号;

[0066] 5) 基于是否产生荧光信号,确定所述目标基因靶序列的至少一个CpG的甲基化状态;

[0067] 6) 采用磁微粒化学发光免疫分析夹心法,测定人血清中CEA、CYFRA21-1和ProGRP的浓度;

[0068] 在一些方式中,所述的聚合酶链式反应体系中,DNA聚合酶包括耐热DNA聚合酶、热启动的DNA聚合酶、或缺乏5'-3'外切酶活性的聚合酶。

[0069] 所述目标基因靶序列中至少一个CpG的甲基化状态是由PCR反应的循环阈值Ct值或靶基因的Ct值之间的差值进行确定。通过利用PCR反应分析生物样本中DNA的甲基化状态,能够方便地实现一个或多个目标基因靶序列的甲基化状态的检测。

[0070] 在一些方式中,所述将DNA的5位未甲基化的胞嘧啶转化为尿嘧啶,优选采用的试剂为亚硫酸氢盐。

[0071] 胞嘧啶5号位的甲基化修饰,是广泛存在于真核细胞生物中的一种DNA修饰方式,DNA上的甲基化修饰不仅在生物的生长发育中起到重要作用,在细胞的原癌化过程中也起着非常重要的作用。由于和胞嘧啶具有一样的碱基互补配对性质,5-甲基胞嘧啶无法通过一代测序或高通量测序的方式来直接测定。目前最常用的检测5-甲基胞嘧啶的方法是将待测DNA通过亚硫酸氢盐转化,在碱性水解后,未甲基化的胞嘧啶被转化为尿嘧啶,而5-甲基胞嘧啶则不会发生转化。尿嘧啶在碱基互补配对的时候会与腺嘌呤互补配对,区别于胞嘧啶与鸟嘌呤的互补配对,因此通过对经过亚硫酸氢盐处理的DNA进行检测时,借助测序技术、聚合酶链式反应技术或者DNA分子杂交相关的技术,可以确定剩余的未发生转化的胞嘧啶,从而确定在原始DNA分子中,哪些胞嘧啶发生了甲基化。因此,本发明优选采用亚硫酸氢盐作为甲基化转化试剂,对待检的DNA样本进行处理后,通过测序、聚合酶链式反应或DNA分子杂交等相关技术,确定目标基因靶序列内的CpG二核苷酸序列的甲基化状态。

[0072] 在一些方式中,本发明的方法适用于分析混合状态的样本,例如血液、粪便或组织中存在的低浓度肿瘤细胞。因此,当分析这类样本中的CpG二核苷酸序列的甲基化状态时,本领域技术人员可以使用定量测定的方法来确定CpG二核苷酸序列的甲基化水平,如百分比、比率、分数或程度等,而不是单核苷酸分子的甲基化修饰状态。相应的,本发明中所述的甲基化状态应被认为是包括单核苷酸分子的甲基化修饰状态在内的,包含定量甲基化水平来反应的甲基化状态。

[0073] 在一些方式中,本发明采用实时荧光定量PCR方式来测定甲基化状态,诸如:使用Taqman探针的实时荧光定量PCR,使用荧光染料的实时荧光定量PCR,使用甲基化特异性PCR(MSP)等方法,用于测定目标基因靶序列的至少一个CpG二核苷酸的甲基化状态。由于不同甲基化状态的基因目标靶序列的碱基互补配对能力不同,因此可以通过实时荧光定量PCR进行基因组DNA样品中甲基化状态的定量测试,其中序列区分发生在探针杂交水平上。

[0074] 作为对照,本发明中使用COL2A1基因,通过设计引物探针均不覆盖任何CpG二核苷酸的位置,对于投入的经试剂处理后的基因组DNA进行检测。实时荧光定量PCR可以与任何适合的探针一起使用,如Taqman探针、MGB探针、蝎形探针等。所述的荧光探针常规包含一个发光基团、一段核酸序列、一个猝灭基团,及有必要时所进行的一些化学修饰或者特殊核苷酸,如硫代核苷酸、锁核酸等。

[0075] 一般实时荧光定量PCR检测过程中,探针会被设计成熔解温度超过正向和反向引物10°C,这使得探针在退火及延伸过程中会完全结合在PCR产物上。典型的,例如Taqman探针,会在具有5' -3' 外切酶活性的DNA聚合酶的作用下,在延伸过程中发生探针的水解,使得探针中的荧光基团和猝灭基团远离,从而破坏荧光基团和猝灭基团间的共振能量传递,使得荧光基团发出的荧光可以被仪器检测到,同时随着PCR产物的逐渐增多,荧光信号会在一定时间内呈现指数级别的上升,最后在荧光定量PCR仪上呈现“S”型的扩增曲线。

[0076] 用于实时荧光定量PCR的反应试剂包括但不限于:目标基因靶序列的正反向引物、Taqman荧光探针、优化的PCR缓冲液、三磷酸脱氧核苷酸以及具有5' -3' 外切酶活性的DNA聚合酶等。

[0077] 本发明检测目标蛋白水平采用但不限于磁微粒化学发光免疫分析夹心法,其他常

用的检测方法例如流式荧光发光法、酶联免疫吸附法等。

[0078] 在一些方式中,通过结合上述位点或者位点组合的甲基化以及蛋白标志物水平的状态来判断所检测生物样本是否呈阳性。

[0079] 本发明提供的用于肺癌辅助诊断的甲基化位点具有以下有益效果:

[0080] 1、提供了18个全新的差异甲基化位点,其在肺癌血浆游离DNA中的甲基化状态,与肺良性病变病人血浆游离DNA的甲基化状态有显著差异;

[0081] 2、从18个甲基化位点中,任选2、6种组合的甲基化位点,都能用于高灵敏度、高特异性地检测肺癌;

[0082] 3、从18个甲基化位点中,找到Seq ID N0.7、Seq ID N0.18两个甲基化位点的肺癌检测AUC值较高,灵敏度和特异性较好;

[0083] 4、找到6个甲基化位点(Seq ID N0.5、Seq ID N0.7、Seq ID N0.10、Seq ID N0.12、Seq ID N0.15和Seq ID N0.18),可选地与蛋白标志物CEA、CYFRA21-1和ProGRP组合,用于肺癌的无创、快速体外检测,能够更有效地区分肺癌和其余肺病患者,AUC值最高能达到0.984,灵敏度达到89.5%,特异性达到99.0%;

[0084] 5、方便、快捷、检测结果与临床金标准检测结果高度一致。

附图说明

[0085] 图1为实施例1中的肺癌、癌旁组织以及正常人血液WGBS测序数据热图,其中,tumor为肺癌组织样本;adjacent为癌旁组织样本,normal为正常人cfDNA样本;

[0086] 图2为实施例1中的所有cfDNA样本在筛选的靶标区域内的甲基化情况;

[0087] 图3为TCGA数据库中肺腺癌/癌旁样本在候选靶标区域内CpG位点的甲基化率分布情况;

[0088] 图4为TCGA数据库中肺鳞癌/癌旁样本在候选靶标区域内CpG位点的甲基化率分布情况;

[0089] 图5为实施例1中的18个甲基化位点在不同类型样本中甲基化水平对比的箱线图;

[0090] 图6为实施例2中的肺癌患者样本荧光定量PCR反应检测结果示意图;

[0091] 图7为实施例2中的肺良性病患者样本荧光定量PCR反应检测结果示意图;

[0092] 图8为实施例2中的2个甲基化位点结合CEA、CYFRA21-1和ProGRP蛋白检测归一化结果获得的ROC曲线;

[0093] 图9为实施例3中的6个甲基化位点结合CEA、CYFRA21-1和ProGRP蛋白检测归一化结果获得的ROC曲线。

具体实施方式

[0094] 下面结合附图和实施例对本发明作进一步详细描述,需要指出的是,以下所述实施例旨在便于对本发明的理解,而对其不起任何限定作用。本实施例中使用的试剂均为已知产品,通过购买市售产品获得。

[0095] 下述各实施例均进行了多组平行试验和多次重复试验,试验耗时长、需要的样本量大,因此各实施例中涉及的肺癌患者、肺良性疾病患者、正常人的肺癌组织及其癌旁组织样本均来自不同批次。

[0096] 实施例1基因甲基化位点的筛选

[0097] 本实施例从产品的实际临床应用出发,发现并确定了18个在肺癌患者中发生甲基化异常的靶序列。肺癌甲基化qPCR候选靶标的具体筛选过程如下:

[0098] 筛选过程包括两个阶段,第一阶段:对肺癌组织、对应的癌旁组织以及正常人cfDNA样本进行高深度全基因组亚硫酸盐测序(WGBS),通过分析这些数据,得到在肺癌组织中表现出异常高甲基化,同时在正常人cfDNA中表现出极低甲基化的基因组区域;第二阶段:合成针对以上区域的甲基化捕获探针,对肺癌和肺良性患者的cfDNA样本的靶区域进行高深度测序,通过分析这些数据,获得能有效区分肺癌患者cfDNA和肺良性疾病患者cfDNA的基因组区域并评估区分性能。

[0099] 在第一阶段,我们对18对肺癌组织及其癌旁组织样本,20例正常人cfDNA样本进行了高深度的WGBS。相比于使用450K芯片数据和基于甲基化敏感限制性内切酶处理后的测序数据,WGBS数据可以覆盖更多CpG位点的甲基化信息。同时,我们将肺癌组织数据与癌旁组织数据进行比较,可以有效规避直接将肺癌组织数据与正常人cfDNA数据比较时导致的正常组织特异甲基化区域的不足的情况。在筛选差异甲基化区域(DMR)时,我们将基因组切割成100bp(50bp overlap)的滑动窗口,对于每一区域,比较它们在每一对肺癌组织/癌旁组织中的甲基化差异情况,然后选择在至少一半的成对肺癌组织/癌旁组织中表现出差异甲基化(且甲基化的变化在两种样本中的方向是一致的,即在肺癌组织中表现出较高甲基化,在癌旁组织中平均甲基化率小于0.1)的区域。这是一种FDMR(Frequently Differential Methylation Region),与直接将不同类型的样本分成两组直接比较相比,我们的方法既可以充分利用配对样本的信息,同时也可避免个别样本的离群值对最终结果的影响。此外,对于通过以上方法获得的滑动窗口区域,我们又用20例正常人cfDNA的WGBS数据进行了第二步的筛选,即我们仅需要在正常人cfDNA中平均甲基化率小于0.01的滑动窗口,这可排除cfDNA中常见高甲基化区域可能对结果造成的混淆。通过以上方法,我们获得了一批在肺癌组织中表现出异常高甲基化、同时在正常人cfDNA中表现出极低甲基化的区域。图1展示了这些滑动窗口区域在肺癌组织、癌旁组织以及正常人cfDNA样本中的平均甲基化情况。

[0100] 在第二阶段,我们将通过文献调研得到的和肺癌相关的异常高甲基化区域与以上候选滑动窗口区域合并,针对合并后的区域,我们设计了5种甲基化捕获探针,即针对原始序列的探针、基因组正链CpG位点完全甲基化的探针和完全去甲基化的探针、基因组负链CpG位点完全甲基化和完全去甲基化的探针,仅针对CpG位点完全甲基化和完全去甲基化的情况设计探针,符合肿瘤DNA协同甲基化的生物学特征,也简化了我们分析数据的复杂程度。利用这些探针,我们对90例肺癌患者cfDNA样本和82例肺良性疾病患者cfDNA样本进行了捕获和高深度测序,靶区域的平均覆盖深度超过10000X。由于早期肿瘤患者外周血中的仅含有微量的ctDNA,对样本进行超高深度测序,有助于捕捉到微小的早期肿瘤甲基化变化信号,提高敏感性。

[0101] 分析这些捕获测序数据时,考虑到相互临近的CpG位点往往表现出相似的甲基化变化,我们在每一个候选滑动窗口区域内,把任意连续的3个CpG位点作为一个marker(3CpG),对于每一个marker,统计3个CpG位点均被甲基化或均未甲基化的数目。具体地,a.找到该区域内所有连续的3个CpG位点的坐标;b.对于任意一组连续的3个CpG位点,从bam文件夹中提取出相应的甲基化状态信息(bam文件XM tag中的甲基化状态字符串),当仅有

read1比对到这3个CpG位点时,使用read1的甲基化状态信息,当仅有read2比对到这3个CpG位点时,使用read2的甲基化状态信息,当read1和read2同时比对到这3个CpG位点时,按照bismark的建议,使用read1的甲基化状态信息,对于read中有indel的情况,需要校正XM tag中甲基化状态字符串的位置以匹配基因组坐标;c.同时使用比对到正链和负链的信息;d.由于我们的探针仅针对完全甲基化和完全去甲基化的状态,所以计算上述的methylation rate时并不考虑在这连续的3个CpG位点呈现出各种中间甲基化状态的read pair(例如:甲基化-非甲基化-甲基化、非甲基化-甲基化-甲基化等等)。这种处理方法有效整合了临近CpG位点的共甲基化信息,使获得的特征信号更robust。

[0102] 考虑到肺癌患者ctDNA的入血量各不相同,从cfDNA样本计算出的甲基化率并不能直接比较,所以,我们对3CpG数据进行0/1编码,即如果至少有一条read在这3个连续的CpG位点都是甲基化的,这个marker被编码为1;如果所有read在这3个连续的CpG位点都是去甲基化的,这个marker被编码为0;如果没有read覆盖这3个连续的CpG位点,则标记为NA(缺失值)。这样,我们就得到了172例cfDNA样本在靶区域内所有3CpG marker上的编码值。根据样本缺失值的数目,把所有样本分成两批,即缺失值比例小于等于(所有3CpG marker)1/10的和比例大于(所有3CpG marker)1/10的。为了挑选能够用来有效区分肺癌患者cfDNA和肺良性疾病患者cfDNA的基因组区域,我们进行了如下10折交叉验证。首先,把第一批样本随机分成10份,每次用其中9份构建随机森林模型,计算出不同候选靶标区域的重要性指标,并对其余份样本类型进行预测;这个过程一共进行10次。第二,累计在10次交叉验证中,不同候选靶标区域重要性指标大于1的次数,选择累积次数大于5的候选靶标区域。第三,统计在这些靶标区域内,表现出明显甲基化的肺良性疾病患者cfDNA样本的数目,挑选该数目不大于5的靶标区域作为最终的候选区域。图2展示了所有cfDNA样本在这些靶标区域内的甲基化情况。在图2中,每一行代表一个cfDNA样本,每一列代表一个3CpG marker,属于不同候选靶标区域的多个连续的3CpG合并在一起。

[0103] 同时,为了验证这些区域的准确性,我们分别收集了TCGA数据库中肺腺癌/癌旁和肺鳞癌/癌旁样本的450K甲基化芯片数据,查看与最终候选靶标区域有交集的CpG位点的甲基化率的分布情况。如图3和图4所示,对于候选靶标区域内的CpG位点,肺癌样本具有更高的甲基化比率。使用肺癌组织、癌旁组织、肺良性结节血浆、肺癌血浆、正常血浆的数据结合文献调研,我们从NGS检测数据中筛选确定了36个启动子区肺癌高甲基化区域。

[0104] 根据区域设定,我们从NCBI数据库网站下载序列,并使用Methyl Primer Express v1.0软件将序列转化为甲基化转化后的序列,即保留“CG”序列中的“C”,并将所有其他的“C”转变为“T”。如下所示。为保证设计区域序列有较平衡的GC含量分布,在序列选择时同时考虑将反义链作为模板设计。根据通用引物设计原则,对36个高甲基化区域设计相应的检测引物和探针。使用以下浓度的模板对设计的引物进行筛选验证:

[0105] 1) 选择纯阳性模板MJ,该模板为经SSSI酶处理后纯化得到,所有CG位点均有甲基化修饰,是最佳的甲基化试验验证的阳性物质模板,选择100pg/u1及10pg/u1表征中浓度及低浓度水平。

[0106] 2) 选择纯阴性模板WBC,该模板取自健康人群的外周血白细胞gDNA,含有整套gDNA背景,经超声打断后,较好的模拟人cfDNA的状态。选择1ng/u1作为模拟真实高浓度样本。

[0107] 3) 低丰度混合模板MJ/WBC,该模板用于模拟真实样本中的混合状态,通常将5%作

为模拟低丰度样本的状态。

[0108] 根据引物筛选的设计组合,每种组合分别配制甲基化检测体系进行比较,以100pg/u1及10pg/u1甲基化阳性模板、1ng/u1甲基化阴性模板及5%丰度甲基化混合模板做综合筛选,筛选标准如下表1所示。最终以特异性、检出率以及扩增效果Ct值为判定指标,对引物组合进行选择。

[0109] 表1序列组合筛选标准

编号	模板类型	浓度	性能要求
1	WBC	1ng/u1	不检出
2	MJ	100pg/u1	CT值在29-31之间
3	MJ	10pg/u1	可检出
4	5%MJ/WBC	1ng/u1	可检出

[0111] 筛选结果如表2所示:

[0112] 表2:区域筛选结果

编号	区域	WBC-1ng/u1	MJ-100pg/u1	MJ-10pg/u1	MJ-5% (1ng/u1)
1	区域 1	37.1	30.51	34.61	36.1
2	区域 2	UD	31.78	UD	UD
3	区域 3	UD	30.82	34.61	35.6
4	区域 4	UD	30.66	34.11	35.77
5	区域 5	36.28	30.3	UD	36.81
6	区域 6	UD	30.06	UD	35.73
7	区域 7	UD	30.01	33.32	34.21
8	区域 8	UD	29.06	32.54	33.27
9	区域 9	UD	30.78	34.52	35.92
10	区域 10	UD	30.57	34.03	35.47

[0113]

	11	区域 11	UD	31.41	35.27	36.36
	12	区域 12	UD	31.98	35.54	37.34
	13	区域 13	UD	29	32.37	33.87
	14	区域 14	UD	30.3	33.38	34.21
	15	区域 15	UD	30.43	36.41	36.15
	16	区域 16	UD	30.54	35.24	37.91
	17	区域 17	UD	29.06	32.33	33.96
	18	区域 18	UD	29.41	33.08	34.02
	19	区域 19	UD	30.55	33.98	35.08
	20	区域 20	UD	30.35	34.27	35.17
	21	区域 21	33.59	31.58	35.41	34.38
	22	区域 22	34.07	31.33	35.03	34.78
[0114]	23	区域 23	UD	29.26	32.82	34.37
	24	区域 24	UD	29.25	32.97	34.86
	25	区域 25	UD	31.55	UD	36.21
	26	区域 26	UD	31.49	34.81	36.36
	27	区域 27	UD	30.98	34.84	36.12
	28	区域 28	UD	30.26	34.28	34.98
	29	区域 29	UD	32.17	UD	UD
	30	区域 30	UD	32.38	35.5	UD
	31	区域 31	UD	30.09	33.19	34.35
	32	区域 32	UD	30.65	33.55	34.31
	33	区域 33	UD	29.78	33.37	35.31
	34	区域 34	UD	30.56	33.82	35
	35	区域 35	UD	30.44	34.43	35.88
	36	区域 36	UD	30.69	33.6	35.25

[0115] UD:未检测到

[0116] 经过筛选,其中区域1、5、21、22存在非特异性扩增问题,将以上4个区域排除,区域2、6、11、12、25、26、29、30存在扩增灵敏度不足问题,同样将以上8个区域排除。剩余24个区域在检测性能上均达到较优水平。随后在临床样本上进行靶点的临床性能验证,使用样本包括肺癌样本50例,肺良性结节样本50例。对检测结果进行单区域的ROC分析,将单区域AUC<0.7的区域进行排除,以排除临床性能较差的区域,缩小区域数量,以满足qPCR检测同时,保证后续区域组合验证时,保持较佳的临床性能。

[0117] 表3:单区域AUC

编号	区域组合个数	AUC
1	区域 3	0.791
2	区域 4	0.751
3	区域 7	0.786
[0118] 4	区域 8	0.801
5	区域 9	0.759
6	区域 10	0.785
7	区域 13	0.788
8	区域 14	0.801
9	区域 15	0.659
10	区域 16	0.624
11	区域 17	0.526
12	区域 18	0.645
13	区域 19	0.800
14	区域 20	0.786
15	区域 23	0.733
[0119] 16	区域 24	0.710
17	区域 27	0.764
18	区域 28	0.712
19	区域 31	0.636
20	区域 32	0.610
21	区域 33	0.804
22	区域 34	0.791
23	区域 35	0.779
24	区域 36	0.820

[0120] 经过上述筛选,最终发现并确定了18个在肺癌患者中高甲基化的靶序列(具有如Seq ID NO.1~SeqID NO.18所示的核苷酸序列,详见说明书中表1)。这18个异常甲基化位点在肺癌患者和正常人血浆样本中甲基化水平对比的箱线图如图5所示,可见本发明筛选获得的18个甲基化位点能够有效区分肺癌和正常人。此外,单区域AUC<0.7的区域15、16、17、18、31、32的靶序列具有如SEQ ID NO.77~82所示的核苷酸序列。

[0121] 实施例2采用2个甲基化位点(或结合蛋白标志物)诊断肺癌

[0122] 本实施例从实施例1筛选的18个甲基化位点中选择其中2个用于肺癌的检测。并分别采用两种方式进行检测:一、采用2个甲基化位点检测肺癌;二、采用2个甲基化位点和蛋白标志物组合的方法检测肺癌。

[0123] 一、采用2个甲基化位点检测肺癌

[0124] 采用2个甲基化位点检测肺癌的具体方法包括如下步骤:

[0125] 第1步,分离血液样本的血清和血浆,使用磁珠法提取试剂,提取待测生物样本的血浆游离DNA,其中20例为肺癌患者,20例为肺良性病患者。

[0126] 第2步,使用亚硫酸氢盐为主要成分的甲基化转化试剂,对步骤1中提取的血浆游离DNA样本进行甲基化转化处理,投入5ng血浆游离DNA,将未发生甲基化的胞嘧啶转变为尿

嘧啶。

[0127] 第3步,将转化后的血浆游离DNA投入到含有检测基因靶序列的实时荧光定量PCR的反应体系中。其中用于检测2个靶序列的荧光探针分别使用FAM和ROX荧光染料标记、用于检测内参基因COL2A1的荧光探针使用VIC荧光染料标记。其中,上游引物、下游引物和探针是指如表2所示的,分别对应这2个靶序列的上游引物、下游引物和探针。

[0128] 荧光定量PCR检测时的体系是多个目的基因和内参COL2A1的引物探针混合,形成一个多重PCR体系,多个目的基因都与内参COL2A1同时检测。单管最多4个靶基因与内参同时检测,检测靶基因数量更多时则分多管进行检测。反应体系中,靶基因序列的正反向引物投入浓度为0.167 μ M,探针投入浓度为0.167 μ M,实时荧光定量PCR反应体系为30 μ L,内参基因序列的正反向引物投入浓度为0.083 μ M。

[0129] 第4步,设置荧光定量PCR反应检测程序如表4:

[0130] 表4、多重PCR反应程序

温度	时间	循环数	备注
95 $^{\circ}$ C	10 min	1	/
95 $^{\circ}$ C	15 s	45	/
57 $^{\circ}$ C	35 s		采集荧光信号

[0132] 第5步,获得荧光定量PCR反应检测结果。

[0133] 本实施例分别选用如下几组包含2个甲基化位点的组合进行检测:

[0134] 1、Seq ID N0.3、Seq ID N0.7

[0135] 2、Seq ID N0.7、Seq ID N0.18

[0136] 3、Seq ID N0.13、Seq ID N0.18

[0137] 当选用第2组Seq ID N0.7、Seq ID N0.18所示的核苷酸序列对应的两个甲基化位点时,检测结果如图6、7所示,其中图6为肺癌患者样本结果,所有待检测甲基化位点的荧光信号都有检出,为阳性;图7为肺良性病患者样本,仅有对照基因COL2A1的荧光信号被检出,其余待检测位点荧光信号未检出,为阴性。

[0138] 采用两个甲基化位点检测肺癌的结果分析判定方法:1)记录软件自动输出的各甲基化位点Ct值;2)分别计算样本中每个位点与内参COL2A1的Ct值,而后对Ct进行归一化处理: $\Delta Ct(\text{靶序列}) = |Ct(\text{COL2A1}) - Ct(\text{靶序列})|$;3)M个甲基化位点,第i个甲基化位点的score为 M_i 。 M_i 的取值分别为0或1,根据 $\Delta Ct(\text{靶序列})$ 值以及对应的Youden's index来判别 M_i 。如果 $\Delta Ct(\text{靶序列}) > \text{Youden's index}$ 设 $M_i = 1$,如果 $\Delta Ct(\text{靶序列}) < \text{Youden's index}$ 设 $M_i = 0$ 。甲基化的M-score = $\sum_i M_i$ 。检测结果如表2所示。

[0139] 根据logistic regression分析对使用所述试剂盒检测获得的结果进行校正和分析计算。M-score的阈值设定根据ROC曲线来设定。并通过综合两个目标甲基化位点的检测结果,对这40个样本进行10乘交叉验证,并取平均值,获得分类ROC曲线,计算AUC值、检测灵敏度和特异性。

[0140] 通过综合两个靶序列SEQ ID N0.7、SEQ ID N0.18甲基化的检测结果见表5,对这40个样本进行10乘交叉验证,并取平均值,获得分类ROC曲线。

[0141] 表5、2个甲基化位点甲基化水平检测结果

样本编号	临床诊断	Seq ID NO.7	Seq ID NO.18	COL2A1
S1	肺癌	45.00	45.00	合格的
S2	肺癌	34.51	38.84	合格的
S3	肺癌	31.58	34.57	合格的
S4	肺癌	31.99	32.44	合格的
S5	肺癌	32.23	40.57	合格的
S6	肺癌	38.99	45.00	合格的
S7	肺癌	35.44	33.32	合格的
S8	肺癌	32.91	32.52	合格的
S9	肺癌	33.92	33.05	合格的
S10	肺癌	45.00	45.00	合格的
S11	肺癌	36.88	39.63	合格的
S12	肺癌	31.51	45.00	合格的
S13	肺癌	45.00	39.45	合格的
S14	肺癌	45.00	45.00	合格的
S15	肺癌	29.79	29.02	合格的
[0142] S16	肺癌	35.34	45.00	合格的
S17	肺癌	29.94	34.99	合格的
S18	肺癌	28.73	45.00	合格的
S19	肺癌	40.17	39.17	合格的
S20	肺癌	45.00	45.00	合格的
S21	肺癌	37.22	36.09	合格的
S22	肺癌	30.57	41.80	合格的
S23	肺癌	34.34	45.00	合格的
S24	肺癌	37.47	36.58	合格的
S25	肺癌	37.64	38.37	合格的
S26	肺癌	45.00	38.86	合格的
S27	肺癌	36.83	37.62	合格的
S28	肺结节	40.05	45.00	合格的
S29	肺结节	45.00	45.00	合格的
S30	肺结节	45.00	45.00	合格的
S31	肺结节	45.00	45.00	合格的
S32	肺结节	45.00	45.00	合格的
S33	肺结节	45.00	45.00	合格的
S34	肺结节	45.00	45.00	合格的
S35	肺炎	45.00	45.00	合格的
[0143] S36	肺炎	45.00	38.64	合格的
S37	肺炎	45.00	45.00	合格的
S38	肺炎	45.00	45.00	合格的
S39	肺炎	45.00	40.05	合格的
S40	肺炎	45.00	45.00	合格的

[0144] 针对本次的样品检测结果,采用SEQ ID NO.7和SEQ ID NO.18的两个甲基化位点的组合的肺癌检测水平,其AUC=0.879,效果明显好于其他两个甲基化位点组合。第1、2、3组的检测分析结果详见后续实施例5及表8。

[0145] 二、采用两个甲基化位点和蛋白标志物组合的方法检测肺癌

[0146] 本实施例还分别采用1、2、3组的两个甲基化位点,再结合蛋白标志物的方法检测肺癌,其中,蛋白标志物包括CEA、CYFRA21-1和ProGRP。

[0147] 其中2个甲基化位点的检测方法参照上述的第1步~第5步,蛋白标志物CEA、

CYFRA21-1和ProGRP的检测还需增加第6步:取第1步分离的血清样本,采用磁微粒化学发光免疫分析夹心法,测定人血清中蛋白标志物的浓度。

[0148] 磁微粒化学发光免疫分析夹心法过程:将R1(分别采用CEA、CYFRA21-1和ProGRP相对应的抗体)、待测样本、M磁微粒混合孵育。样本中蛋白标志物的不同位点与磁珠上偶联的抗体结合,形成固相抗体-抗原复合物;通过清洗,后加入R2试剂(分别采用CEA、CYFRA21-1和ProGRP相对应的二抗)混合孵育,上述复合物和经标记的肿瘤标志物抗体结合,形成固相抗体-抗原-抗体夹心复合物;通过洗涤,未被结合的抗体以及其它物质被去除。在反应复合物中加入化学发光底物1和化学发光底物2,通过相对发光强度测定化学发光反应,所产生的发光强度与样本中肿瘤标志物的浓度成正比。对所述样品中的蛋白标志物进行磁微粒化学发光免疫分析夹心法测试并将得分确定为P值。检测结果如表6所示。

[0149] 分析判定方法:在结合上述的单独采用2个甲基化位点检测肺癌的结果分析判断方法,还需:

[0150] 1) 分别对每个样本的CEA、CYFRA21-1和ProGRP检测值进行归一化处理: $P1 = \log_{10} P_{CEA}$, $P2 = \log_{10} P_{CYFRA21-1}$, $P3 = \log_{10} P_{ProGRP}$;蛋白的P-score = $\sum_j \hat{p}(P_j)$;

[0151] 2) 根据logistic regression分析对使用所述试剂盒检测获得的结果进行校正和分析计算。M-score和P-score的阈值设定根据ROC曲线来设定。

[0152] 通过整合甲基化和蛋白标志物两个互补的维度来提高检测性能。整合后的模型为LC-score = M-score + P-score。在一些实施方案中,在LC-score值等于或大于设定阈值时,结果表明患者中肺癌和/或早期肺癌的阳性检测。在一些实施方案中,在LC-score值小于阈值时,结果表明患者中肺癌和/或早期肺癌的阴性检测。

[0153] 3) 通过综合两个目标甲基化位点,以及CEA、CYFRA21-1和ProGRP蛋白检测结果的归一化结果,对这40个样本进行10乘交叉验证,并取平均值,获得分类ROC曲线。

[0154] 当选用第2组的Seq ID NO.7、Seq ID NO.18所示的核苷酸序列对应的两个甲基化位点时,结合CEA、CYFRA21-1和ProGRP蛋白检测后的归一化结果,对这40个样本进行10乘交叉验证,并取平均值,获得分类ROC曲线如图8所示,其中CEA、CYFRA21-1和ProGRP蛋白检测结果如表6所示。

[0155] 表6、不同样本蛋白水平检测结果

样本编号	临床诊断	proGRP(pg/mL)	CEA(ng/ml)	CYFRA21-1(ng/ml)
S1	肺癌	10.10	12.96	20.66
S2	肺癌	33.87	2.90	4.45
S3	肺癌	21.06	2.03	1.76
S4	肺癌	40.86	6.57	6.23
[0156] S5	肺癌	35.19	1.75	1.74
S6	肺癌	27.81	8.99	29.06
S7	肺癌	40.72	13.70	204.40
S8	肺癌	35.04	2.67	3.50
S9	肺癌	17.52	3.07	3.92
S10	肺癌	1.40	2.81	3.96

[0157]

S11	肺癌	38.83	1.69	2.82
S12	肺癌	22.90	2.90	3.41
S13	肺癌	5.00	13.82	10.09
S14	肺癌	42.73	6.89	2.07
S15	肺癌	22.76	187.40	148.50
S16	肺癌	44.34	4.30	82.06
S17	肺癌	63.20	4.61	4.33
S18	肺癌	30.32	77.74	4.90
S19	肺癌	53.40	39.02	17.84
S20	肺癌	38.91	5.47	3.57
S21	肺癌	30.82	6.16	5.58
S22	肺癌	60.08	1.78	5.02
S23	肺癌	58.98	19.15	14.36
S24	肺癌	5.84	10.68	5.00
S25	肺癌	38.57	1.90	2.38
S26	肺癌	46.28	3.81	3.39
S27	肺癌	45.26	5.37	10.80
S28	肺结节	3.89	3.14	6.41
S29	肺结节	42.76	53.77	12.59
S30	肺结节	21.82	2.28	2.10
S31	肺结节	23.86	1.86	3.38
S32	肺结节	32.53	3.69	3.90
S33	肺结节	23.60	1.99	2.42
S34	肺结节	22.49	8.70	2.50
S35	肺炎	43.94	2.61	2.12
S36	肺炎	36.65	1.50	3.33
S37	肺炎	29.29	2.27	4.87
S38	肺炎	30.13	10.26	2.21
S39	肺炎	345.04	10.22	145.17
S40	肺炎	23.96	5.30	2.30

[0158] 当采用第2组的两个甲基化位点,并结合CEA、CYFRA21-1和ProGRP蛋白检测后归一化结果,可以进一步提升肺癌检测效果,其AUC=0.940,灵敏度为81.5%,特异性为94.5%(图8)。

[0159] 第1、2、3组的2个甲基化位点,结合CEA、CYFRA21-1和ProGRP蛋白检测后的归一化结果,详见后续实施例5。

[0160] 实施例3采用6个甲基化位点(或结合蛋白标志物)诊断肺癌

[0161] 本实施例从实施例1筛选的18个甲基化位点中选择6个用于肺癌的检测。并分别采用两种方式进行检测:一、采用6个甲基化位点检测肺癌;二、采用6个甲基化位点和蛋白标志物组合的方法检测肺癌。

[0162] 一、采用6个甲基化位点检测肺癌

[0163] 本实施例分别选用如下几组包含6个甲基化位点的组合进行检测:

[0164] 4、Seq ID N0.5、Seq ID N0.7、Seq ID N0.10、Seq ID N0.12、Seq ID N0.15和Seq ID N0.18

[0165] 5、Seq ID N0.2、Seq ID N0.7、Seq ID N0.8、Seq ID N0.11、Seq ID N0.14和Seq ID N0.15

[0166] 6、Seq ID N0.3、Seq ID N0.8、Seq ID N0.10、Seq ID N0.11、Seq ID N0.13和Seq

ID NO.17

[0167] 7、Seq ID NO.1、Seq ID NO.4、Seq ID NO.12、Seq ID NO.14、Seq ID NO.15和Seq ID NO.17

[0168] 8、Seq ID NO.5、Seq ID NO.7、Seq ID NO.8、Seq ID NO.12、Seq ID NO.13和Seq ID NO.17

[0169] 9、Seq ID NO.1、Seq ID NO.3、Seq ID NO.6、Seq ID NO.8、Seq ID NO.10和Seq ID NO.11

[0170] 10、Seq ID NO.3、Seq ID NO.7、Seq ID NO.13、Seq ID NO.8、Seq ID NO.9和Seq ID NO.16

[0171] 11、Seq ID NO.3、Seq ID NO.4、Seq ID NO.5、Seq ID NO.10、Seq ID NO.13、Seq ID NO.15

[0172] 采用6个甲基化位点检测肺癌的具体方法包括如下步骤：

[0173] 第一步，得到40名肺良性病及肺癌患者，其中肺良性病样本为20例，肺癌样本为20例。分离提取样本的血清和游离血浆DNA。

[0174] 第二步，使用亚硫酸氢盐为主要成分的甲基化转化试剂，对血浆游离DNA样本进行甲基化转化处理，投入血浆游离DNA，将未发生甲基化的胞嘧啶转变为尿嘧啶。

[0175] 第三步，指如表2所示的分别对应这6个靶序列的上游引物、下游引物和探针，按照实施例2的方法进行实时荧光定量PCR检测。

[0176] 针对本次的样品检测结果，采用第7组的6个甲基化位点的效果明显好于其他组的肺癌检测水平。根据ROC曲线可得，其AUC=0.918。

[0177] 第4-11组的6个甲基化位点的检测分析结果详见后续实施例5。

[0178] 二、采用6个甲基化位点和蛋白标志物组合的方法检测肺癌

[0179] 在上述检测采用6个甲基化位点检测肺癌的基础上，还需增加：

[0180] 第四步，取第一步分离的血清样本，采用磁微粒化学发光免疫分析夹心法，测定人血清中CEA、CYFRA21-1和ProGRP蛋白标志物水平的检测，并对结果进行分析。

[0181] 第五步，将所检测到的结果中各基因甲基化和蛋白的值进行归一化，通过综合6个甲基化位点的 ΔCt (靶基因) 结果以及三个蛋白标志物水平的检测结果，对这40个样本进行10重交叉验证，并取平均值，获得分类ROC曲线 (图9)。

[0182] 第4-11组的6个甲基化位点，结合CEA、CYFRA21-1和ProGRP蛋白检测后的归一化结果详见后续实施例5。

[0183] 当采用第7组的6个甲基化位点，并结合CEA、CYFRA21-1和ProGRP蛋白检测后的归一化结果，可以进一步提升肺癌检测效果，其AUC=0.984，灵敏度为89.5%，特异性为99.0%。

[0184] 实施例4采用18个甲基化位点 (或结合蛋白标志物) 诊断肺癌

[0185] 本实施例采用实施例1筛选的18个甲基化位点进行肺癌的检测。并分别采用两种方式进行检测：一、采用18个甲基化位点检测肺癌；二、采用18个甲基化位点和蛋白标志物组合的方法检测肺癌。

[0186] 一、采用18个甲基化位点检测肺癌

[0187] 采用18个甲基化位点检测肺癌的具体方法包括如下步骤：

[0188] 第一步,得到40名肺良性病及肺癌患者,其中肺良性病样本为20例,肺癌样本为20例。分离提取样本的血清和游离血浆DNA。

[0189] 第二步,使用亚硫酸氢盐为主要成分的甲基化转化试剂,对血浆游离DNA样本进行甲基化转化处理,投入血浆游离DNA,将未发生甲基化的胞嘧啶转变为尿嘧啶。

[0190] 第三步,指如表2所示的18个靶序列的上游引物、下游引物和探针,按照实施例2的方法进行实时荧光定量PCR检测。

[0191] 采用18个甲基化位点进行检测的AUC达到0.919,详见实施例5。

[0192] 二、采用18个甲基化位点和蛋白标志物组合的方法检测肺癌

[0193] 采用18个甲基化位点和蛋白标志物组合的方法检测肺癌的具体方法包括如下步骤:在上述检测采用18个甲基化位点检测肺癌的基础上,还需增加:

[0194] 第四步,取第一步分离的血清样本,采用磁微粒化学发光免疫分析夹心法,测定人血清中CEA、CYFRA21-1和ProGRP蛋白标志物水平的检测,并对结果进行分析。

[0195] 第五步,将所检测到的结果中各基因甲基化和蛋白的值进行归一化,通过综合18个靶基因的 ΔCt (靶基因)结果以及三个蛋白标志物水平的检测结果,对这40个样本进行10乘交叉验证,并取平均值,获得分类ROC曲线,其AUC-甲基化+蛋白=0.982,灵敏度达到89.0%,特异性达到99.0%,详见实施例5。

[0196] 实施例5采用不同甲基化位点组合的性能比较分析

[0197] 对实施例1中得到的40例肺癌和肺良性病患者样本的18个甲基化位点(SEQ ID NO.1~18)的相对循环数 ΔCT 值进行不同位点组合的数学建模分析,以探讨18个甲基化位点和蛋白作为生物标记物组合对于检测肺癌的应用。

[0198] 首先,我们分别评估了上述18个甲基化位点中单个位点的模型对于诊断肺癌发生的性能,分别计算其AUC值,结果如表7所示。

[0199] 表7、单个甲基化位点诊断肺癌的性能比较

	SEQ ID NO	AUC
	1	0.776
[0200]	2	0.760
	3	0.760
	4	0.813

	5	0.798
	6	0.821
	7	0.862
	8	0.752
	9	0.732
	10	0.825
[0201]	11	0.773
	12	0.739
	13	0.837
	14	0.796
	15	0.811
	16	0.762
	17	0.774
	18	0.832

[0202] 由表7可以看出,实施例1提供的18个甲基化位点用于诊断肺癌的AUC值都较高,都具有较好的诊断性能,尤其是其中的Seq ID NO.7和Seq ID NO.18所示的核苷酸序列对应

的甲基化位点,是全新发现的可用于高效区分肺癌与其他肺病的高甲基化位点。

[0203] 其次,对比实施例2~4列举的不同组合甲基化位点,或结合蛋白标志物组合的诊断效能,此外加上了一组囊括实施例1中24个区域的甲基化位点,其中包括SEQ ID NO.1~18的18个甲基化位点和在实施例1中由于AUC值较低而被舍去的6个甲基化位点(区域15、16、17、18、31、32),分别对应具体序列如SEQ ID NO.77~82所示;最终结果如表8所示。

[0204] 表8、不同组合甲基化位点的模型对于诊断肺癌发生的比较

[0205]

组别	组合 (SEQ ID NO)	AUC-甲基化	AUC-甲基化+蛋白
1(实施例2)	3+7	0.825	0.877
2(实施例2)	7+18	0.879	0.940
3(实施例2)	13+18	0.858	0.910
4(实施例3)	5+7+10+12+15+18	0.918	0.984
5(实施例3)	2+7+8+11+14+15	0.902	0.952
6(实施例3)	3+8+10+11+13+17	0.911	0.961
7(实施例3)	1+4+12+14+15+17	0.907	0.957
8(实施例3)	5+7+8+12+13+17	0.917	0.967
9(实施例3)	1+3+6+8+10+11	0.899	0.949
10(实施例3)	3+7+13+8+9+16	0.907	0.954
11(实施例3)	3+4+5+10+13+15	0.909	0.949
12(实施例4)	1~18	0.919	0.982
13	1~18+77~82	0.919	0.982

[0206] 由表7和11可以看出,使用单个甲基化位点作为诊断模型的对比多个甲基化位点组合模型诊断性能较低,使用2个甲基化位点组合作为诊断模型比6个甲基化位点组合模型诊断性能要低,甲基化结合蛋白标志物的多组学模型的诊断性能要显著优于单一组学标志物;但是采用SEQ ID NO.1~18的18个甲基化位点,与采用18个甲基化位点和SEQ ID NO.77~82所示6个甲基化位点协同诊断,诊断性能没有明显差别,但选择18个甲基化位点,减少6个甲基化位点,能降低诊断过程的测试指标量和分析数据量。

[0207] 当选用甲基化位点协同蛋白共同进行诊断时,部分甲基化位点的组合与蛋白协同诊断的诊断性能显著增加,例如第4组采用Seq ID NO.5、Seq ID NO.7、Seq ID NO.10、Seq ID NO.12、Seq ID NO.15和Seq ID NO.18所示的核苷酸序列,结合蛋白进行多组学模型诊断时,AUC值达到了0.984,为所有模型中最高的,且相较于18个甲基化位点和蛋白,减少了12个甲基化位点,能降低诊断过程的测试指标量和分析数据量,为优选。

[0208] 当选用2个甲基化位点组合作为诊断模型时,第2组的AUC-甲基化和AUC-甲基化+蛋白都是最高的,可见第2组采用Seq ID NO.7、Seq ID NO.18所示的核苷酸序列对应的2个甲基化位点时,诊断性能明显高于其他2个甲基化位点组合。

[0209] 当选用6个甲基化位点组合作为诊断模型时,第4组的AUC-甲基化和AUC-甲基化+蛋白都是最高的,可见第4组采用Seq ID NO.5、Seq ID NO.7、Seq ID NO.10、Seq ID NO.12、Seq ID NO.15和Seq ID NO.18所示的核苷酸序列对应的6个甲基化位点时,诊断性能明显高于其他6个甲基化位点组合。

[0210] 当选用18个甲基化位点组合作为诊断模型时,其诊断性能与第4组的6个甲基化位

点组合的诊断性能非常接近,因此最优选采用第4组的6个甲基化位点组合。

[0211] 本实施例进一步对从中优选的不同的甲基化位点组合的诊断性能分别进行分析,灵敏度=患病被诊断患病的数量/患病总数量;特异度是指当真实情况未患病时诊断试验排除患病的能力,特异度=未患病未被诊断患病的数量/未患病总数量;Youden指数=灵敏度+特异度-1,结果如表9所示。

[0212] 表9、优选的多个甲基化位点的模型对于诊断肺癌发生的比较

组合 (SEQ ID NO)	AUC	灵敏度	特异度	Youden指数
7+蛋白	0.881	74.1%	100%	0.741
18+蛋白	0.861	74.1%	94.7%	0.688
7+18+蛋白	0.940	81.5%	94.5%	0.760
5+7+10+12+15+18+蛋白	0.984	89.5%	99.0%	0.885
2+7+8+11+14+15+蛋白	0.952	92.5%	92.8%	0.853
3+8+10+11+13+17+蛋白	0.961	93.5%	93.0%	0.865
1+4+12+14+15+17+蛋白	0.957	92.6%	92.9%	0.855
5+7+8+12+13+17+蛋白	0.967	93.9%	93.3%	0.872
1+3+6+8+10+11+蛋白	0.949	91.3%	91.6%	0.829
3+7+8+9+13+16+蛋白	0.954	91.7%	92.4%	0.841
3+4+5+10+13+15+蛋白	0.949	91.1%	92.2%	0.833
1~18+蛋白	0.982	89.0%	99.0%	0.880
1~18+77~82+蛋白	0.983	89.2%	99.0%	0.882

[0214] 由表9可见,本发明提供的的几组甲基化位点+蛋白的组合都能够用于肺癌的高效检测,其中优选Seq ID NO.5、Seq ID NO.7、Seq ID NO.10、Seq ID NO.12、Seq ID NO.15和Seq ID NO.18和蛋白的多组学诊断模型,AUC最高能达到0.984,灵敏度达到89.5%,特异性达到99.0%,真正实现无创、全局、具有更高灵敏度和特异性的肺癌诊断,完全满足临床需求。

[0215] 实施例7训练集临床样本中采用6个甲基化位点(或结合蛋白标志物)诊断肺癌

[0216] 本实施例利用实施例1筛选的6个甲基化位点,并采用甲基化位点和蛋白标志物联合检测的方法用于训练集临床样本的肺癌检测。

[0217] 具体方法包括如下步骤:

[0218] 第1步,分离血液样本的血清和血浆,使用磁珠法提取试剂,提取待测生物样本的血浆游离DNA,其中185例为肺癌患者,70例为正常人对照。

[0219] 第2步,使用亚硫酸氢盐为主要成分的甲基化转化试剂,对步骤1中提取的血浆游离DNA样本进行甲基化转化处理,投入5ng血浆游离DNA,将未发生甲基化的胞嘧啶转变为尿嘧啶。

[0220] 第3步,将转化后的血浆游离DNA投入到含有检测基因靶序列的实时荧光定量PCR的反应体系中。其中用于检测6个靶序列的荧光探针分别使用TXD、FAM、和CY5荧光染料标记、用于检测内参基因COL2A1的荧光探针使用VIC荧光染料标记。其中,上游引物、下游引物和探针是指如表2所示的,分别对应这6个靶序列的上游引物、下游引物和探针。

[0221] 荧光定量PCR检测时的体系是多个目的基因和内参COL2A1的引物探针混合,形成

一个多重PCR体系,多个目的基因都与内参COL2A1同时检测。单管最多4个靶基因与内参同时检测,检测靶基因数量更多时则分多管进行检测。反应体系中,靶基因序列的正反向引物投入浓度为0.167 μ M,探针投入浓度为0.167 μ M,实时荧光定量PCR反应体系为30 μ L,内参基因序列的正反向引物投入浓度为0.083 μ M。

[0222] 第4步,设置荧光定量PCR反应检测程序如下:

	温度	时间	循环数	备注
[0223]	95 $^{\circ}$ C	10 min	1	/
	95 $^{\circ}$ C	15 s	45	/
	57 $^{\circ}$ C	35 s		采集荧光信号

[0224] 第5步,获得荧光定量PCR反应检测结果。

[0225] 第6步:取第1步分离的血清样本,采用磁微粒化学发光免疫分析夹心法,测定人血清中蛋白标志物的浓度。

[0226] 参考北京热景生物相应的蛋白检测试剂盒说明书,对所述样品中的蛋白标志物进行磁微粒化学发光免疫分析夹心法测试并将得分确定为P值。

[0227] 甲基化联合蛋白标志物检测肺癌的结果分析判定方法:

[0228] 1) 记录软件自动输出的各甲基化位点Ct值;

[0229] 2) 分别计算样本中每个位点与内参COL2A1的Ct值,而后对Ct进行归一化处理: $\Delta Ct(\text{靶序列}) = |Ct(\text{COL2A1}) - Ct(\text{靶序列})|$;

[0230] 3) 15个甲基化位点,第i个甲基化位点的score为 M_i 。 M_i 的取值分别为0或1,根据 $\Delta Ct(\text{靶序列})$ 值以及对应的Youden's index来判别 M_i 。如果 $\Delta Ct(\text{靶序列}) > \text{Youden's index}$ 设 $M_i = 1$,如果 $\Delta Ct(\text{靶序列}) < \text{Youden's index}$ 设 $M_i = 0$ 。甲基化的M-score = $\sum_{i=1}^{15} M_i$ (i=1-15)。

[0231] 4) 分别对每个样本的CEA、CYFRA21-1和ProGRP检测值进行归一化处理: $P_1 = \log_{10} P_{\text{CEA}}$, $P_2 = \log_{10} P_{\text{CYFRA21-1}}$, $P_3 = \log_{10} P_{\text{ProGRP}}$;蛋白的P-score = $\sum_{j=1}^3 P_j$ 。

[0232] 5) 根据logistic regression分析对使用所述试剂盒检测获得的结果进行校正和分析计算。M-score和P-score的阈值设定根据ROC曲线来设定。

[0233] 通过整合甲基化和蛋白标志物两个互补的维度来提高检测性能。整合后的模型为LC-score = M-score + P-score。在一些实施方案中,在LC-score值等于或大于设定阈值时,结果表明患者中肺癌和/或早期肺癌的阳性检测。在一些实施方案中,在LC-score值小于阈值时,结果表明患者中肺癌和/或早期肺癌的阴性检测。

[0234] 针对本次的训练集临床样本检测结果,采用6个甲基化位点并结合CEA、CYFRA21-1和ProGRP蛋白的标志物组合对于肺癌检测水平,其AUC=0.982,灵敏度为90.5%,特异性为98.4%。

[0235] 实施例8测试集临床样本中采用6个甲基化位点(或结合蛋白标志物)诊断肺癌

[0236] 本实施例利用实施例1筛选的6个甲基化位点,并采用甲基化位点和蛋白标志物联合检测的方法用于测试集临床样本的肺癌检测,其中158例为肺癌患者,70例为正常人对照,其他实验和检测方法同实施例7所示。

[0237] 针对本次的测试集临床样本检测结果,采用6个甲基化位点并结合CEA、CYFRA21-1和ProGRP蛋白的标志物组合对于肺癌检测水平,其AUC=0.983,灵敏度为91.1%,特异性为98.2%。

[0238] 本发明实施例优选的6个甲基化位点联合3个蛋白的组合能够用于肺癌的高效检测,在测试集临床样本中,其中AUC达到0.983,灵敏度达到91.1%,特异性达到98.2%,真正实现无创、全局、具有更高灵敏度、特异性的肺癌诊断,能够满足临床需求。

[0239] 虽然本发明披露如上,但本发明并非限于此。任何本领域技术人员,在不脱离本发明的精神和范围内,均可作各种更动与修改,因此本发明的保护范围应当以权利要求所限定的范围为准。

[0240] 序列表

[0241] Seq ID NO.1

[0242] 甲基化位点1:

[0243] AGAGTTAGGAGATCGGAAAGTTGAAGTAGTTGTTTTATTAAGTTATAGAGGAAGTGTTGAGTTAAGGGTTCGGGAGGTAGTTAGTTTACGGTAATACGAGAAAATTCGGGGTTCGTTGGTGGTTTAGGTTTAGGTTTGGGGTTAGGAGTTAAAGGTTTTGT

[0244] Seq ID NO.2

[0245] 甲基化位点2:

[0246] GAGGCGGTGGTCGAGGAGCGTCGTCGGAGCGGGATTGGTTCGTTTAGGTAGTCGAATAAGAGCGTTAGGTTAGCGAGATGCGGGGGCGTTTCGAGTAGTTTGAGAAAACGGTAGAGTTGTTGAAAGAGAAGATGGAGTATTTTATTGGGTTTTGTCGAGA

[0247] Seq ID NO.3

[0248] 甲基化位点3:

[0249] GTTAAGTTTTTCGCGGGAGGAGAGGTTTTTCGTAGGAGGAGACGTTAGAGGGTTTTATCGACGTTTTTCGGTTGACGAGTCGTTTTTCGGAGTTTCGTTTTTCGGGTTTTTCGTGCGTGCGAACGTTTCGGCGAACGTTACGGTCGATTTCGAGTAGTTGTTTCGC

[0250] Seq ID NO.4

[0251] 甲基化位点4:

[0252] GTAAGAGCGTAGTTACGGCGGGCGGTTAGTGCGTCGTTTTTTTTTTATTGGGTATTTTCGGGTTTCGAGGACGAGGCGCGCGGAGCGGTTGGCGGAGTTGGTGGCGTTGGAGGCGCGGAGTACGGCGACGTGTTGTAGTGGGTTTTTCGCGGATATTTTTT

[0253] Seq ID NO.5

[0254] 甲基化位点5:

[0255] CGCGATGTAAGTAATTGAGATCGGGAGTTGTTTTTCGGTAGAGCGTATTTATTTTCGGTTTTAGGTGGATTGAAGTTTAGAGCGGCGTTGTGTAGTTGGAAGGGCGCGGATAGTTTAAGTTAGAGGCGTTTTTCGGGGCGGCGGTAGGATATAAGATTTTTAAA

[0256] Seq ID NO.6

[0257] 甲基化位点6:

[0258] TTAGAATAGTAGGTTTTTGGAGTTTTTTTCGGGTTTTGGGTTTGTTCGTAGGAGCGCGAGTATAGCGCGTAGTTTCGGCGTTTCGTACGTACGGTGTGAGCGTTTTATTTCGTAGGATGTAGTTTTTTTCGGTTCGTTTGTTCGTTTTATGTTAGCGTGGTGTAG

[0259] Seq ID NO.7

[0260] 甲基化位点7:

[0261] TTTTCGGCGTTTAGTTCGTTTTATGCGGTTTCGTAGTTTTAAGTATAGTTGTTTTAGGGTTGGTGGCGT
AGGTTTTGTTATACGTCGAAGATTTTTTAGTCGGTCGGCGGCGTTTTTTCGGGTTTAGGGTTCGCGCGTTTAGTAG
TAGGGGCGAAAGGTA

[0262] Seq ID NO.8

[0263] 甲基化位点8:

[0264] TTGGTGTATGGCGCGGGGAAGGGGCGCGTTAGGACGGGTTTTTTTTTTTCGCGTTAGTTTCGGATTTTT
AGGGCGGAAGTCGGTGGTTGTTGGCGAGGCGGCGGCGGTTTATCGGGTTTTTTTTTTAAAAGGTTTCGTAATTA
TTAATGTGTTTTTTGT

[0265] Seq ID NO.9

[0266] 甲基化位点9:

[0267] GGTTTCGTAGTAGTTAGGGGTCGAGTTACGGTTACGGACGTTTTGGTGTTCGGTTCGTGTCGGGTTTTT
AGGCGGAGGAGGCGTTCGTTGGGTTTAGATTTTCGATTTTAGTTTCGGTTTTTCGGCGTTTGGGTTGCGCGGAGTTT
TTGTTTCGCGTTTCGG

[0268] Seq ID NO.10

[0269] 甲基化位点10:

[0270] TTGTTTTGTGTTGTTTTTTAGTTAGGGTTCGATCGGTTAGGTTTCGGTGAATTTTAGTAATTTTAGT
TTTGTTTAGGGATTTGTTTCGTTTCGTTGGGCGTCGAAAAGGTATATAGATGCGAGGGAATTTGCGGTTTGATTTT
TTGGGGATTTGTAATT

[0271] Seq ID NO.11

[0272] 甲基化位点11:

[0273] AGTATGTTTTTTTTAAAGTTTTTTTTGGGTGTTTTTTTTTATCGTTATCGTCGTCGGTTGTCGTTTTATT
ATTTTTGTTTTATTTTTTTTTTATATGTTTTATTTTAGACGGTGGTTTTTAGAAGTTTTGTTTTTTGATAGTT
GTCGTTTGGGTAGTTC

[0274] Seq ID NO.12

[0275] 甲基化位点12:

[0276] CGTCGGGAACGGAGCGCGTTTAATTTTTAGCGGGAGTCGTTAGGTTTGGTTTGGTCGGGGTTTTTTTC
GTTTCGTTATTTTCGATAAAGTATAGTCGAGTTCGGTTGGAAGGTAGAGTTTCGAAGTAGGTAGGACGGAGCGGAGT
AAAAGAATGCGGTTT

[0277] Seq ID NO.13

[0278] 甲基化位点13:

[0279] CGGTTTCGCGTTTCGCGTTTCGCGTTTTTAGCGTTGCGTTTTTCGTTTTCGTTTTCGTTTCGTTAGTT
TGGTTCGTTTAGCGATTGCGTTTATTTGAAGATCGTATTTAGGGTAGATGCGGAAATTGGTTTTAGTCGCGTTATG
TAGCGCGTTTTTCGTT

[0280] Seq ID NO.14

[0281] 甲基化位点14:

[0282] TAGGCGTCGGCGTTTAGGTCGTAGTCGGACGTAAAGGGTTTTGATAAAGGGGTTGTTGATATTGTAT
AAGTTCGGAACGGTCGAGGCGAAGGCGTCGGCGTTCGTTTCGTAGTCGTTTTTTGTGAGTTG

[0283] Seq ID NO.15

[0284] 甲基化位点15:

[0285] TAGGTGGTGGAGGTGGAGGTGGAGGACGACGAGAGGTTGGTCGCGGAGGGCGCGAGGGGGTTCGGTCCGGG
GAGCGCGCGGGCGAGTTTAGTACGCGCGTCAAGGGTTGTAGATAGCGGTCGGTCCGGCGTTTAGTTGTAGTTTTAG
GGGTCCGATGGGGTTGACGAAGTCGGTTTGTATTGTTTGGTCGCGACGATTTAGGTTTCGTCCGGTTTTGTTTATTTT
TTTTTAGCGTTTGTGTAGT

[0286] Seq ID NO.16

[0287] 甲基化位点16:

[0288] AGTACGTTGATGTTGGAGTTTAGGTTATTGCGGAGGGTTATGGCGGTAGGCGGGTTCGGTTCGGTTCG
GGCGGCGAGTAGGTAGTAGGGTCGGTCGGCGAGAGAGGCGTCGGGGTAAGTTTTTTTTTTCGGCGAAGTGGTTCGTT
TTTTAGTGAGTTTTTT

[0289] Seq ID NO.17

[0290] 甲基化位点17:

[0291] GGGCGTGTAGGTTGAGTGTCCGGGATAGGCGCGATATTGGTGTGGCGTTGGCGTTATAGATTTAGGG
TTGCGCGTGTTTTTTGGTTTTTAGTTGAGATCGGCGATGTTGGAGTTTTTGTGGTGGTTTTGGCGTTGTTGCG
GTCGTTATTATCGAGG

[0292] Seq ID NO.18

[0293] 甲基化位点18:

[0294] CGGCGGAAGTCGAATTCGCGGTTAGCGTGCGAGCGGTAGTTCGAAGGGCGGTGTTGGGAATATTATGT
AGGGCGCGTGC CGGTTAGGTGCGGATGTAGGTGGTGGTGC CGTGC GTTATAGCGTTGTTTAGTTGTAGTTGCGTT
TGAATTTGAAAGGATA

[0295] Seq ID NO.19

[0296] COL2A1:

[0297] TTTTTGTAAAGGAGGGATGTGGAGGGATAGAGGAGTAGTAGGTAAGGTTAGTAGGAGGTGATATAGGTA
GGGAGGATTAGGTTAAGGTTGGGAGGAGTTTATATTTGGTGT

[0298] Seq ID NO.20

[0299] COL2A1的正向引物:

[0300] ATGTGGAGGGATAGAGGAGTA

[0301] Seq ID NO.21

[0302] COL2A1的探针:

[0303] CACCTCCTACTAACCTTACC

[0304] Seq ID NO.22

[0305] COL2A1的反向引物:

[0306] CCTCCCAACCTTAACCTAAT

[0307] Seq ID NO.23

[0308] 甲基化位点1的正向引物:

[0309] GGTTGAGTTAAGGGTTCG

[0310] Seq ID NO.24

[0311] 甲基化位点1的探针:

[0312] AGTTTACGGTAATACGAGAA

[0313] Seq ID NO.25

- [0314] 甲基化位点1的反向引物：
[0315] CACCAACGAACCCCGAA
[0316] Seq ID NO.26
[0317] 甲基化位点2的正向引物：
[0318] GTCGTCGGAGCGGGGATT
[0319] Seq ID NO.27
[0320] 甲基化位点2的探针：
[0321] TCGTTTAGGTAGTCGAATAA
[0322] Seq ID NO.28
[0323] 甲基化位点2的反向引物：
[0324] CGCCCCGCATCTCGCT
[0325] Seq ID NO.29
[0326] 甲基化位点3的正向引物：
[0327] ATCGACGTTTTTCGCGTT
[0328] Seq ID NO.30
[0329] 甲基化位点3的探针：
[0330] CGAGTCGTTTTTCGGAGTTC
[0331] Seq ID NO.31
[0332] 甲基化位点3的反向引物：
[0333] GTTCGCCACGCACGAAAACCCGAA
[0334] Seq ID NO.32
[0335] 甲基化位点4的正向引物：
[0336] GGCGCGCGCGGAGCG
[0337] Seq ID NO.33
[0338] 甲基化位点4的探针：
[0339] GGCGGAGTTGGTGGCGTT
[0340] Seq ID NO.34
[0341] 甲基化位点4的反向引物：
[0342] CACTACAACACGTCGCCGTAC
[0343] Seq ID NO.35
[0344] 甲基化位点5的正向引物：
[0345] GAGCGTATTTATTTTCGGTTTTAG
[0346] Seq ID NO.36
[0347] 甲基化位点5的探针：
[0348] TAAACTATCGCGGCCCTTCCA
[0349] Seq ID NO.37
[0350] 甲基化位点5的反向引物：
[0351] CGCCGCGCCCCGAAACCG
[0352] Seq ID NO.38

- [0353] 甲基化位点6的正向引物:
[0354] GGAGCGCGAGTATAGCGC
[0355] Seq ID NO.39
[0356] 甲基化位点6的探针:
[0357] AGTTCGGCGTTCGTACG
[0358] Seq ID NO.40
[0359] 甲基化位点6的反向引物:
[0360] CTACGAAATAAAACGCTCAAC
[0361] Seq ID NO.41
[0362] 甲基化位点7的正向引物:
[0363] CGTCGAAGATTTTTTAGTCG
[0364] Seq ID NO.42
[0365] 甲基化位点7的探针:
[0366] TCGGCGCGTTTTTTGCG
[0367] Seq ID NO.43
[0368] 甲基化位点7的反向引物:
[0369] CCTACTACTAAACGCGCAACC
[0370] Seq ID NO.44
[0371] 甲基化位点8的正向引物:
[0372] GCGCGTTAGGACGGGTTTTT
[0373] Seq ID NO.45
[0374] 甲基化位点8的探针:
[0375] TTCGCGTTAGTTTCGGATT
[0376] Seq ID NO.46
[0377] 甲基化位点8的反向引物:
[0378] CGCCGCCGCTCGCCAACAA
[0379] Seq ID NO.47
[0380] 甲基化位点9的正向引物:
[0381] GGTCGAGTTACGGTTACGGAC
[0382] Seq ID NO.48
[0383] 甲基化位点9的探针:
[0384] GTTTCGGTTCGTGTCGG
[0385] Seq ID NO.49
[0386] 甲基化位点9的反向引物:
[0387] CGCCGAAAAACCGAAACTAAAATCGA
[0388] Seq ID NO.50
[0389] 甲基化位点10的正向引物:
[0390] GGGTTCGATCGGTTAGGTT
[0391] Seq ID NO.51

- [0392] 甲基化位点10的探针：
[0393] ACGCCCAACGAAACGAA
[0394] Seq ID NO.52
[0395] 甲基化位点10的反向引物：
[0396] TTCCCTCGCATCTATATAACC
[0397] Seq ID NO.53
[0398] 甲基化位点11的正向引物：
[0399] GGGTGTTTTTTTTTATCG
[0400] Seq ID NO.54
[0401] 甲基化位点11的探针：
[0402] TCGTCGTCGGTTGTCGTTT
[0403] Seq ID NO.55
[0404] 甲基化位点11的反向引物：
[0405] CACCGTCTAAAAATAAAACATATAA
[0406] Seq ID NO.56
[0407] 甲基化位点12的正向引物：
[0408] GGAGCGCGTTTAATTTTTAGC
[0409] Seq ID NO.57
[0410] 甲基化位点12的探针：
[0411] GTCGTTAGGTTTGGTTTGG
[0412] Seq ID NO.58
[0413] 甲基化位点12的反向引物：
[0414] TTATCCGAAAATAACGAACGA
[0415] Seq ID NO.59
[0416] 甲基化位点13的正向引物：
[0417] CGCGTTTCGCGTTTCGCG
[0418] Seq ID NO.60
[0419] 甲基化位点13的探针：
[0420] TTAGCGTTGCGTTTTTCG
[0421] Seq ID NO.61
[0422] 甲基化位点13的反向引物：
[0423] CGCAATCGCTAAACGAAC
[0424] Seq ID NO.62
[0425] 甲基化位点14的正向引物：
[0426] GCGTCGGCGTTTAGGTCGTAG
[0427] Seq ID NO.63
[0428] 甲基化位点14的探针：
[0429] CTTCGCCTCGACCGTTCCGAAC
[0430] Seq ID NO.64

- [0431] 甲基化位点14的反向引物:
[0432] CGACTACGAAACGAACGCCGACG
[0433] Seq ID NO.65
[0434] 甲基化位点15的正向引物:
[0435] GCGCGCGGGCGAGTTTAGT
[0436] Seq ID NO.66
[0437] 甲基化位点15的探针:
[0438] CGCGCGTCGAAGGGTTG
[0439] Seq ID NO.67
[0440] 甲基化位点15的反向引物:
[0441] CGTCGCGACCAAACAATACAAAC
[0442] Seq ID NO.68
[0443] 甲基化位点16的正向引物:
[0444] CGGTTCGGGCGGCGAGTAG
[0445] Seq ID NO.69
[0446] 甲基化位点16的探针:
[0447] CTCTCGCCGACCGACCC
[0448] Seq ID NO.70
[0449] 甲基化位点16的反向引物:
[0450] ACGACCACTTCGCCGAAAAAA
[0451] Seq ID NO.71
[0452] 甲基化位点17的正向引物:
[0453] GGCGTTGGCGTTATAGATTTAGG
[0454] Seq ID NO.72
[0455] 甲基化位点17的探针:
[0456] TGCGGCGTGTTTTTTGGTT
[0457] Seq ID NO.73
[0458] 甲基化位点17的反向引物:
[0459] CTCCAACATCGCCGATCTCAAAC
[0460] Seq ID NO.74
[0461] 甲基化位点18的正向引物:
[0462] AGCGTGGCGAGCGGTAGTTCG
[0463] Seq ID NO.75
[0464] 甲基化位点18的探针:
[0465] GCACCTAACCGCGCACGCGC
[0466] Seq ID NO.76
[0467] 甲基化位点18的反向引物:
[0468] GCTATAACGCACGCGCACCAC
[0469] Seq ID NO.77

[0470] 区域15甲基化位点

[0471] AGTTGCGATAGAGTGTGGGGTTGTTGATAAATGAATTGGTTATGGAGAAGGAGGTTATAGAGAAGTTGC
GGAAGTTTTTGGTTTTTTAGAGTAGCGTTTTTCGAGGGTTGTGGGATTGTTTGTTCGTAGATTTAGTGGGCGAGAGG
AGTGTATAAAGTAAAG

[0472] Seq ID NO.78

[0473] 区域16甲基化位点

[0474] TAAAAATTTGGTATTTAGATTTAGAAGTAGTGTGTTTAGGGTTGTTATTAGTGTAGTATTATCGTTTAT
TAGTATTTGGATTTGAATTGTATATTAGTTTGGGAAGAGGTTAATAGTATGATGGGTTGTTGTGGTTTTATTGAAG
TTTTGAGGAGGTTATT

[0475] Seq ID NO.79

[0476] 区域17甲基化位点

[0477] GTTTTTTTTAGTTCGAAGTTCGTAGTATATTCGTAGGTAAAGTTTTTTAAGTCGTTTAGGTAGTTAGG
GAGTTTTGCGTATTTGTTAGTACGGAGGTATTTTCGGGGTAGGGATATAATATATCGTTCGAGAGTTTGTTTTAGC
GAGCGTCGATTTTCGTT

[0478] Seq ID NO.80

[0479] 区域18甲基化位点

[0480] TTGGAGTATCGCGTATTTTCGCGCGGTGTAGGAGCGTTGGGGTTTTTTTTATTTATTGTAGCGTGTCGTT
TTCGAGATCGTCGGGGTCGGAGGATAGTTAGGAGAAATTCGTAAAGGTTTCGTTTTTTAGGGTGTAGTGGGAACGTT
TGCCTTTTAGTGCGAT

[0481] Seq ID NO.81

[0482] 区域31甲基化位点

[0483] AGGGAAGAGTTGTACGTGGAGCGGTTTCGGTTGGTGGTTTTTAGGGCGTTGAGGGCGTTGGCGAAAGAG
TCGTAGTTTCGCGTTTATTAGTTTTTTTTTTGTTTGGAGAGTATGGATATATTAATAAATATTTTTGTTTTTCGGAGAG
GAGTGTGCGAGAGATC

[0484] Seq ID NO.82

[0485] 区域32甲基化位点

[0486] TTTGCGAGATAAAAAATAATTATAGTTAGTTTTATTTAAGGGGGAGATTAGTTCGGTGTTTTTTCGGTGC
TTTTCGGGAGGAAAAGGGCGGGGAGTGGGGTAGGTCGGTCGGGTAGTTTAGTTTGTTCGGTTTAGGGTTTGATTATT
TCGGTTTTTTATTTGG

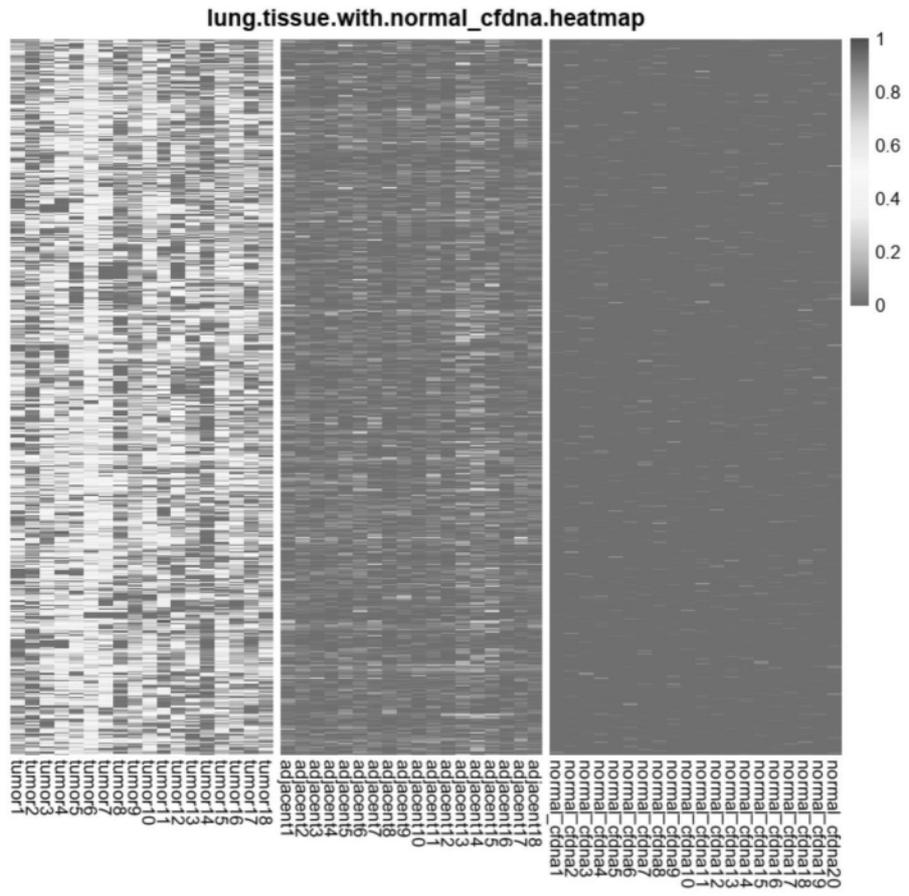


图1

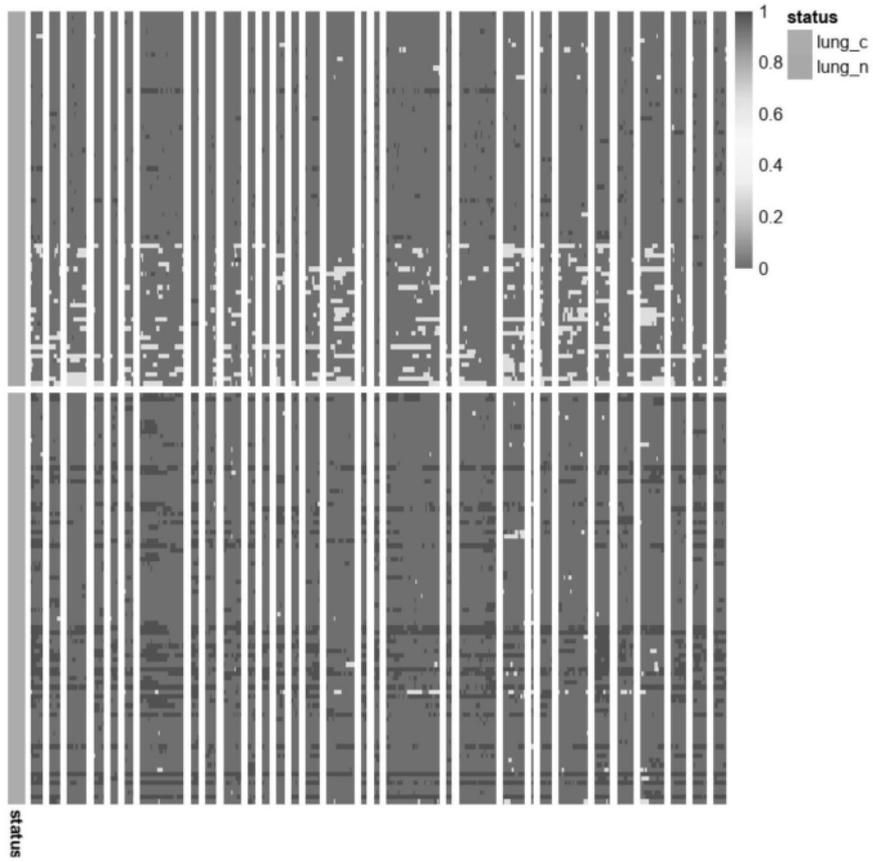


图2

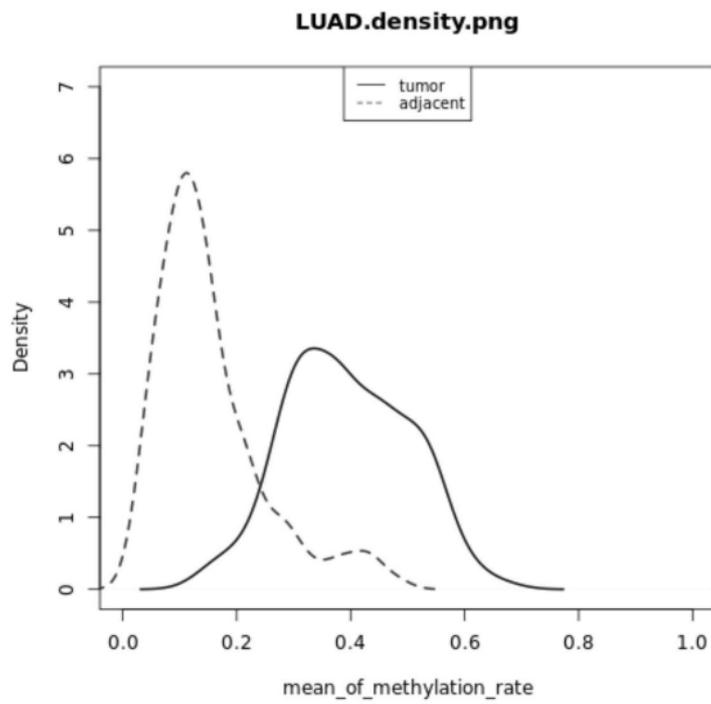


图3

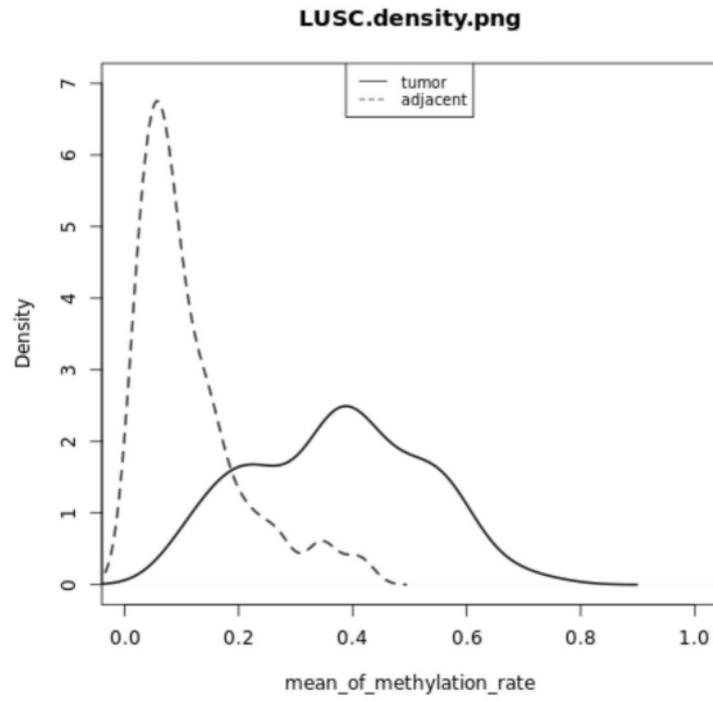
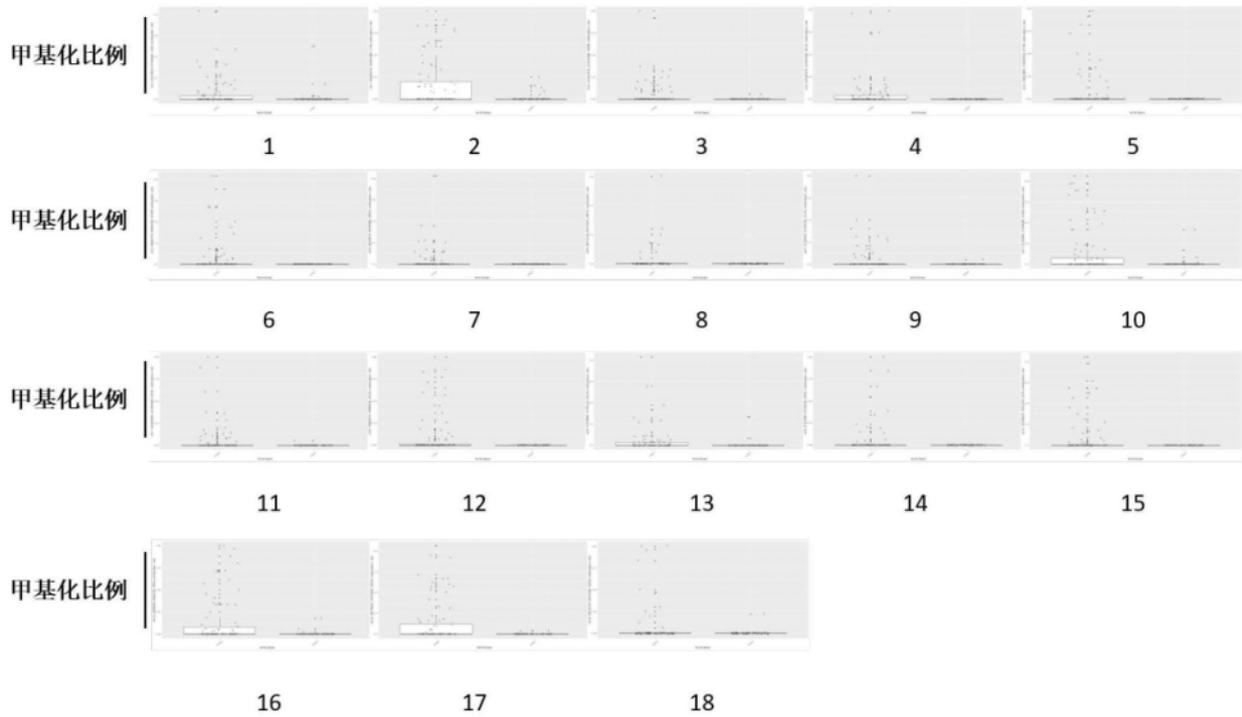


图4



(备注由左至右依次为：肺癌患者和正常人血浆样本)

图5

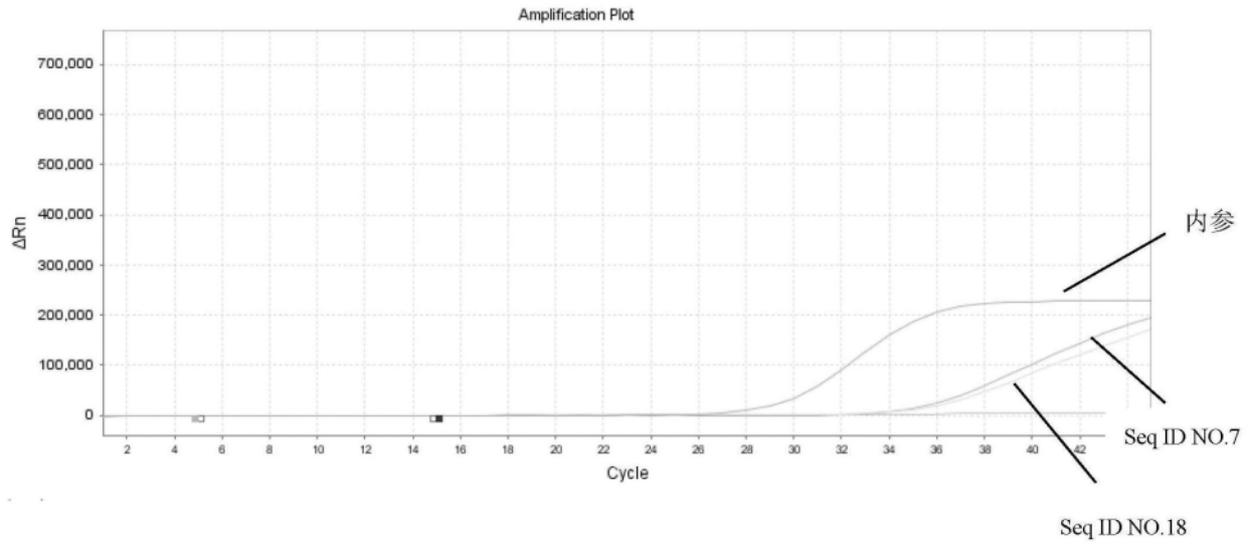


图6

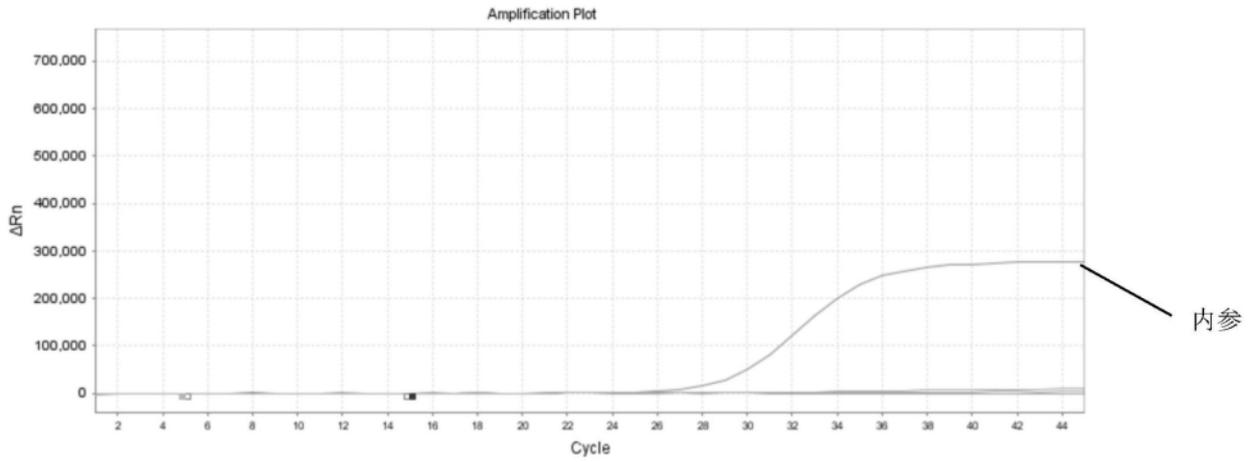


图7

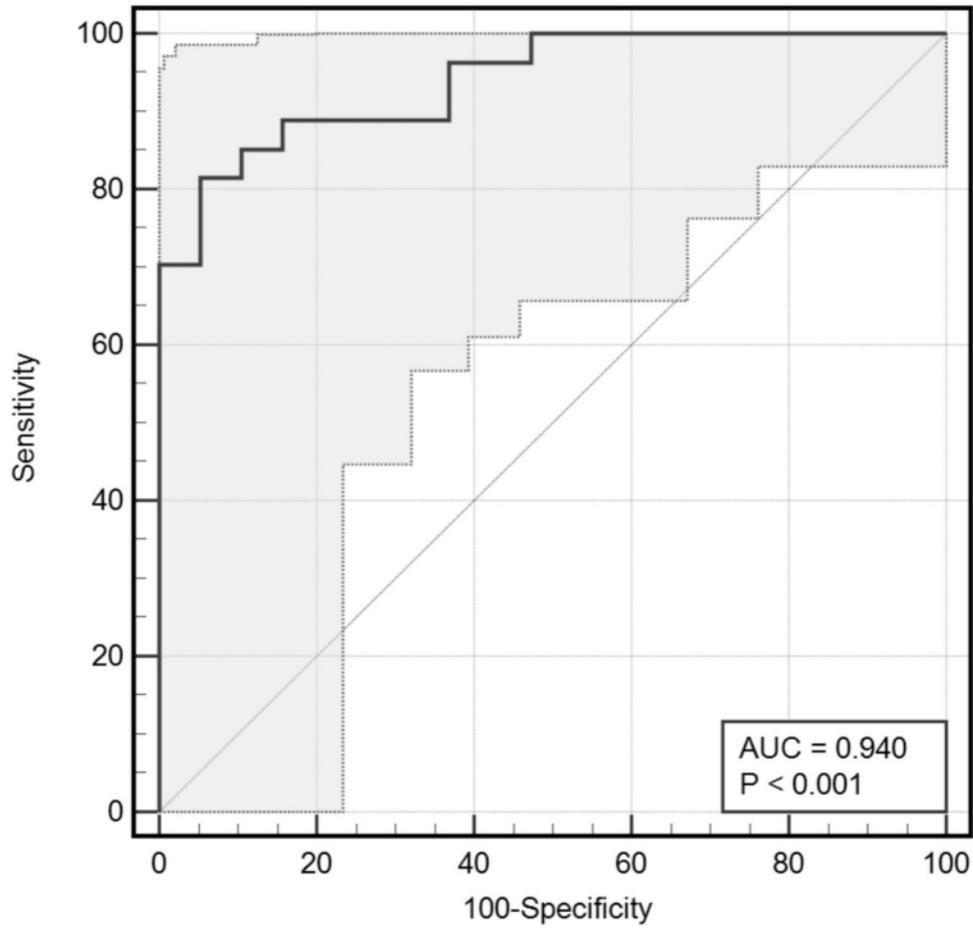


图8

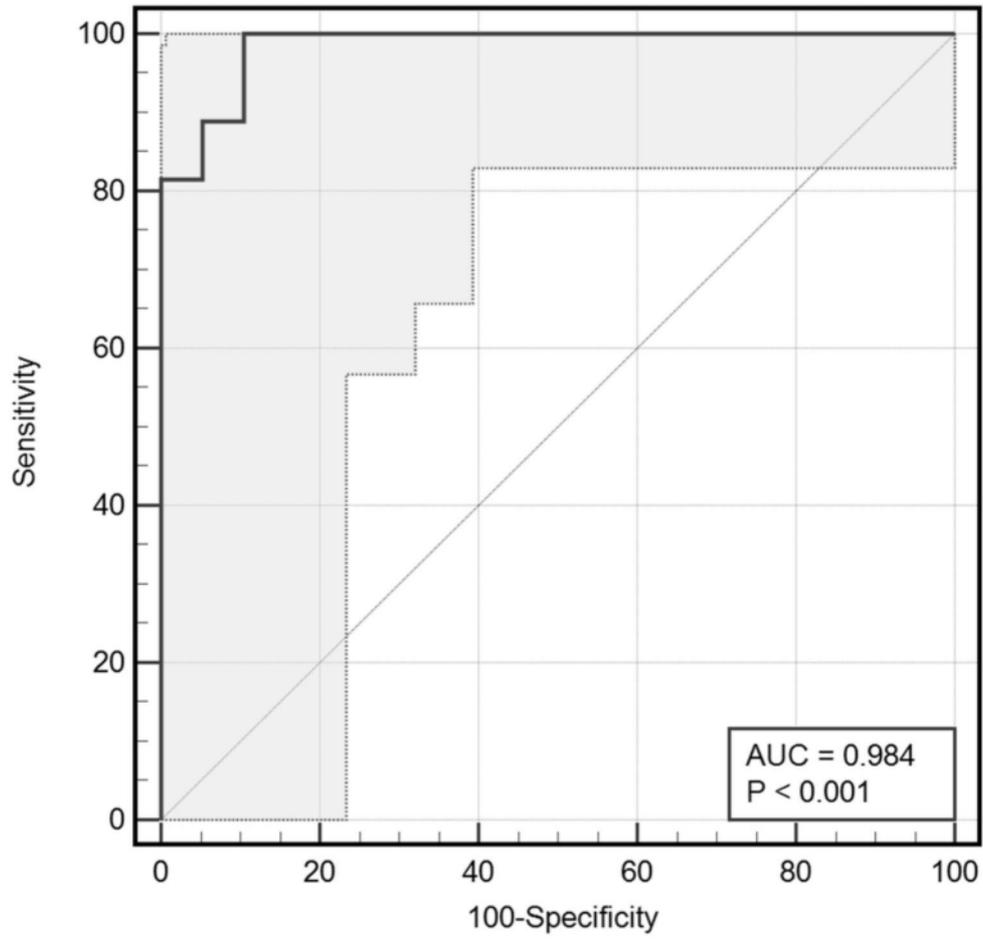


图9